THE COST-EFFECTIVENESS OF CHILD OBESITY INTERVENTIONS

WIN TECHAKEHAKIJ, M.D.

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF YORK

DEPARTMENT OF HEALTH SCIENCES

September 2011

Abstract

Childhood obesity has increasingly come to be recognized as a health problem globally. A number of interventions have been implemented in the attempt to reduce the prevalence of obesity in children, involving both curative and preventive measures. Efforts to tackle childhood obesity take a variety of forms, including medicinal, behavioural, school-based, and other types of interventions. Despite considerable efforts to reduce childhood obesity, however, there is as of yet no evidence-based standard protocol for either treatment or prevention.

One potential deterrent to the adoption of an intervention is the uncertainty of return on investment, due largely to unclear future benefits gained from the various interventions, as well as poor cost measurement. This is because most of the health benefits of child obesity interventions are not revealed until adulthood, making health gains from the interventions difficult to observe. Moreover, there is no method to assess the future medical costs saved due to the reduced prevalence of obesity in childhood. Modelling the costs and benefits of the interventions is therefore crucial, especially in making policy decisions as to which interventions are worth implementing using public funding. Cost-effectiveness analysis is an economic evaluation method used in resource allocation that compares the costs and benefits of competing alternatives.

The author explores evidence regarding the cost-effectiveness of child obesity interventions, with the aim of making contributions to the existing body of knowledge in this field. The first chapter provides an overview of obesity in childhood, addressing such issues as epidemiology, definition, aetiology, and physical and social consequences. The author then conducts a literature review of the cost-effectiveness of child obesity interventions, with results showing that only a handful of high-quality published studies are available. In order to perform the cost-effectiveness analysis, MEND 7-13, a child obesity intervention that aims to treat obese children aged 7 to 13, is used as a case study. The author explores the background of the MEND 7-13 programme and assesses the extent to which various characteristics of participants affect the benefits uptake of the programme. The data used in the analysis come from the MEND rollout phase, which lasted from January 2007 to December 2009 and included 6,828 participants, with an average follow-up duration of 10 weeks. The results showed that significant predictors of

BMI change in White children are gender, programme attendance rate, and baseline BMI; for Asian children, the only significant predictor is having parents who own their own houses; no evidence of association is shown in Black children; and total attendance and gender are significant predictors in the 'Other' group. The results also showed that the programme centre influences the change in BMI to some extent in all children.

The effectiveness of the MEND rollout programme, derived from the MEND rollout data, is used in the cost-effectiveness analysis. An economic modelling technique is developed and employed based on the assumption that MEND 7-13 is available to all children who meet the eligibility criteria in England in 2010. The intervention costs and medical costs saved due to the programme are estimated to be £551.2 million and £216 million, respectively. 200,910 incremental QALYs are gained, leading to an estimated ICER of £1,668.2 per QALY. The author performs a sensitivity analysis using Monte Carlo methods to assess the effect of parameter uncertainty on the cost-effectiveness outcomes.

The strengths and weaknesses of the economic modelling techniques are then discussed, with emphasis on the knowledge contributed by this thesis to the existing literature. The author also points out the research and policy implications of the results, alongside recommendations for future research.

List of contents

List of tables

List of figures

Acknowledgements

I am grateful for the support from a number of people during the time I spent working on this thesis. I thank my supervisor, Professor John Hutton, along with the members of my Research Advisory Group, Dr. Karen Bloor, Professor Alan Maynard, and Dr. Tom Johnston, for their advice and support. I am also thankful to Dr. Simon Crouch for his suggestions regarding the data and statistical analyses in Chapter 3 of this thesis. I am grateful to the MEND staff for providing me with the data used in this thesis, along with useful discussions concerning the methods and results of economic evaluation in Chapter 4 of this thesis. I wish to acknowledge the Collaborative Project to Increase Production of Rural Doctor (CPIRD), the Ministry of Public Health, Thailand, for providing me with a scholarship and funding to do my Ph.D.

I would like to acknowledge Judy Jou for her fantastic proofreading. I also appreciate the support and help of many people, including family, the department of Health Sciences staff, Ph.D. colleagues, friends, and many more people, who have encouraged me throughout the Ph.D. period.

Author's declaration

I declare that all the work presented in this thesis is my own, unless otherwise acknowledged. The contents and views expressed therein reflect the best of my own knowledge, investigation, and belief. The thesis contains no material previously published or written by other people, nor submitted for a degree or award elsewhere.

Win Techakehakij

September 2011

Chapter 1: Childhood obesity: a global rising problem

1.1 Introduction

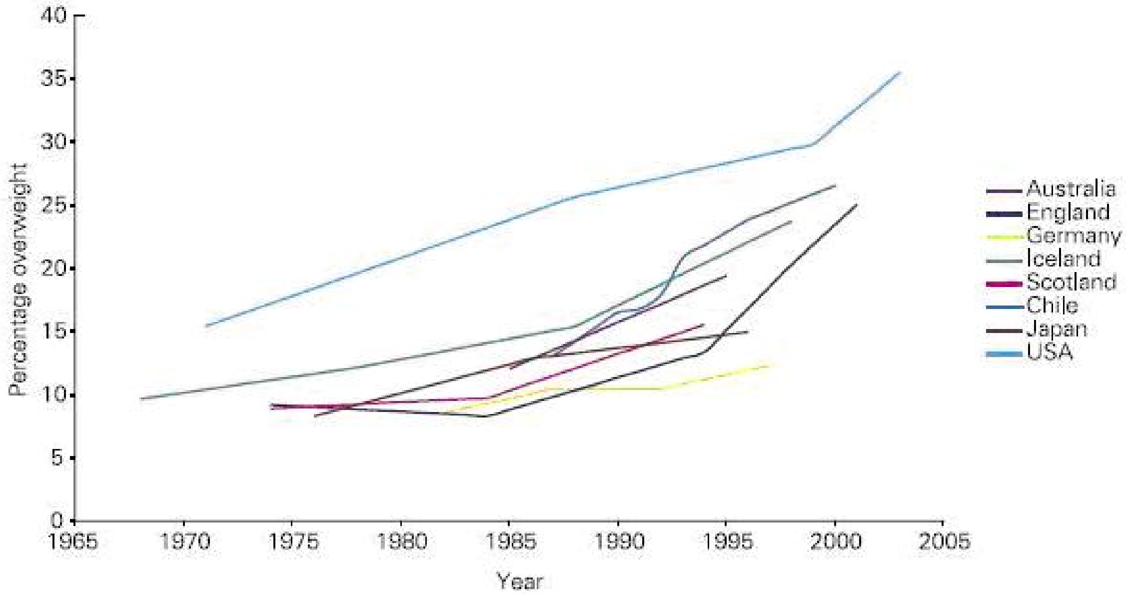
Childhood obesity has increasingly become a public health concern affecting a wide range of populations globally. The World Health Organization (WHO) (2011b) estimated in 2010 that more than 43 million children below the age of 5 were overweight, of which approximately 35 million were in developing countries. Approximately 10% of school-aged children worldwide were categorized as overweight or obese in 1990 by the International Obesity Task Force (IOTF) definition of paediatric overweight and obesity (Lobstein et al., 2004). While only less than 3% of children in Africa are overweight, prevalence is much higher in the US and Europe, where the prevalence of overweight and obesity in children is more than 30% and around 20% respectively (Lobstein et al., 2004).

In the US, the prevalence of overweight in children rose from 5.1% in 1974 to 13.9% in 2000 and 17.1% in 2004 (Ogden et al., 2007). The global prevalence and trend of overweight and obesity in children across countries are shown in Table 1.1 and Figure 1.1. As demonstrated, obesity has become more prevalent not only in the western world, but also in Asia and the Western Pacific. In Asia, the prevalence of childhood overweight in China was approximately 11 percent in 2000 (Wang et al., 2002, Lobstein and Leach, 2007), while in India, it was around 8-13 percent in children aged 5-17 in 2002 (Lobstein and Leach, 2007).

Table 1.1: Prevalence of childhood overweight and obesity worldwide, by WHO region, using the IOTF definition

	Most recent survey		**Projected 2006**		**Projected 2010**	
WHO regions	**overweight %**	**obesity %**	**overweight %**	**obesity %**	**overweight %**	**obesity %**
Africas (1987-2003)	1.6	0.2	-	-	-	-
Americas (1988-2002)	27.7	9.6	40.0	13.2	46.4	15.2
Eastern Med (1992-2001)	23.5	5.9	35.3	9.4	41.7	11.5
Europe (1992-2003)	25.5	5.4	31.8	7.9	38.2	10.0
South East Asia (1997-2002)	10.6	1.5	16.6	3.3	22.9	5.3
West Pacific (1993-2000)	12.0	2.3	20.8	5.0	27.2	7.0

Source: modified from Wang and Lobstein (2006)



Source: Butland et al. (2007)

Figure 1.1: The prevalence of overweight in pre-school and primary-school-age children in some countries

The dramatic increase in rates of overweight and obesity has occurred in the UK as well. In England, the percentage of overweight children increased from 10.5% in 1974 to 13.5% in 1994 and 23% in 2003 (Stamatakis et al., 2005). The Foresight programme reported the prevalence of childhood obesity in England to be 8-10% in 2004, which was projected to rise to 15% by 2025 and 25% by 2050 (Butland et al., 2007). The assumption that rates have increased steadily over time, however, may have led to an overestimation of overweight and obesity prevalence. Stamatakis et al. (2009), for instance, found that the prevalence of childhood overweight and obesity in England remained steady from 2002/3 to 2006/7. As the rise in prevalence tends to flatten, at least during the past decade, the projected prevalence of obesity in childhood should be lower than that reported in Butland et al. (2007).

The prevalence of obesity is influenced by social class. Though data demonstrating this link are largely unavailable for childhood obesity, the impact of social class on obesity prevalence is illustrated in adulthood obesity. According to the Health Survey for England data in 2004, the prevalence of obesity in the UK was approximately 18% and 10% in men and women in social class 1 (the highest), whereas 28% and 25% of men and women in social class 5 (the lowest) were obese (Butland et al., 2007).

Unlike social class, the data on the prevalence of obesity with respect to different ethnicities are very limited; this is due mainly to a small sample size, which does not facilitate reliable outcomes in the comparison (Butland et al., 2007).

It is noted that the estimated numbers and trends on childhood overweight and obesity from the WHO (Table 1.1) are based on national survey data that was collected sometime ago (Wang and Lobstein, 2006, World Health Organization, 2011b, Lobstein and Leach, 2007). The most recent national survey data used were collected around 2000-2003 (Wang and Lobstein, 2006, World Health Organization, 2011a). Moreover, the national survey data used in estimation vary in several domains, such as sampling methods, age of population, and the year of data collection. Each of these factors creates uncertainties in the outcomes in terms of the accuracy and reliability of the estimates. This point should be borne in mind when utilizing this information.

1.2 Definition and measurement of childhood obesity

Body mass index (BMI) is a standard tool that is widely used to measure overweight and obesity in youths and adults (Wyatt et al., 2006). BMI was first invented by Lambert Adolphe Quetelet, a Belgian mathematician, over 170 years ago (Hu, 2007). Also called the Quetelet index, BMI is calculated as weight in kilograms divided by the square of height in meters (World Health Organization, 2011b). The WHO recommends classification of adult obesity and overweight using BMI; people with a BMI not less than 25 are classified as overweight and those with a BMI not less than 30 are defined as obese (World Health Organization, 2011b).

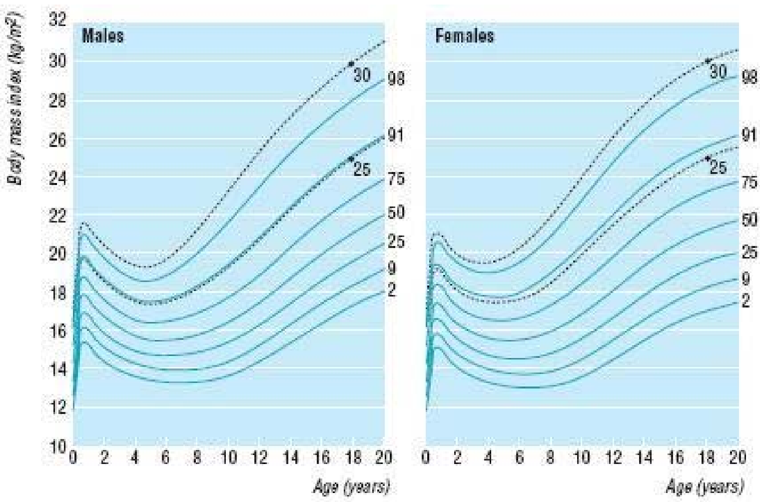
Using BMI to measure body fat is easy and has high accuracy because BMI is proven to be highly correlated with body fat percentage (Gallagher et al., 1996). Despite its benefits, however, some limitations still remain when using BMI to measure adiposity. Hu (2007) describes 4 limitations of using BMI as follows: 1) BMI cannot distinguish fat mass and lean body mass; 2) the validity of BMI is low in older people because lean body mass decreases immensely when aging (Roubenoff and Kehayias, 1991), making the validity of predicting obesity-related mortality using BMI lower in older adults; 3) the relationship between BMI and obesity-related mortality may be diverted by reverse causation, e.g. when a low BMI results from chronic illnesses; and 4) smoking is another confounder that influences the relationship (Hu, 2007).

Apart from BMI, there are other tools used to measure the level of obesity. Some are easier to employ, such as waist circumference and waist-hip ratio (World Health Organization, 2000). Others require specific equipment or high technology, together with technical skills, to measure body fat, e.g. skinfold thickness (Erselcan et al., 2000), bioimpedance analysis (Heymsfield et al., 1997), and dual energy x-ray absorbtiometry (Wyatt et al., 2006).

The accuracy and reliability of body fat measurements tend to increase alongside the difficulty of the measure, causing limitations in general application. On the other hand, the easily applicable tools, such as waist circumference and waist-hip ratio, contain low accuracy in determining body fat. Concerning this trade-off, BMI seems to be the measure that balances accessibility and accuracy, as it has high correlation with body fat percentage and requires only height and weight in the calculation.

Unlike in adults, there is no standard international agreement for the definition of childhood obesity. This is due to the fact that obesity is defined in relation to 2 components: 1) body fat content; and 2) increased risk of adverse health consequences (Reilly, 2005). These factors vary across ages and ethnicity, leading to several definitions being provided to account for these differences.

Although strong evidence has demonstrated the relationship between childhood obesity and morbidity/mortality in both the short and long terms (Reilly et al., 2003), the occurrence of morbidity and amount of body fat content do not consistently correspond to fixed BMI cut-offs in children as in adults. This is due to differences in physiological growth rate and body fat patterns (Power et al., 1997). The median and other averaged centiles for BMI in British children are as shown in Figure 1.2 As illustrated, the BMI curves fluctuate in the early years of life before rising consistently after the age of 6.



Source: Cole et al. (2000)

Figure 1.2: Centiles for body mass index for British males and females

To account for differences in physiological growth and adiposity patterns, it is widely accepted that overweight and obesity cut-offs in children need to be adjusted specifically for gender, age, and sometimes ethnicity (Power et al., 1997). Even though several methods have been proposed to define overweight and obesity in children, there is as yet no standard classification at present.

The recommendation for using BMI cut-offs in diagnosing childhood obesity can be classified by population reference into 2 methods: the population-specific BMI centile cut-off and the internationally standardized BMI cut-off. While considerable scientific evidence supports the population-specific definition of BMI, results are difficult to compare across populations; the main problem with using the international standard BMI definition, meanwhile, is the lack of supportive evidence in real practice.

The first method suggests applying population-specific BMI centile cut-offs to define overweight and obesity in childhood, wherein the national BMI reference data of particular populations are used in determining the cut-off points. The cut-off points of overweight and obesity for American children, for instance, are at the 85th and 95th percentiles (Barlow and Dietz, 1998, Centers for Disease Control and Prevention, 2009), while those for British children are at the 91st and 98th percentiles of the UK 1990 reference chart for age and sex (Scottish Intercollegiate Guidelines Network, 2010). To diagnose obesity and overweight, the Scottish Intercollegiate Guidelines Network also suggests the use of the 95th and 85th BMI percentile cut-offs for research purposes, whereas cut-offs of 98th and 91st BMI percentiles are for clinical use (Scottish Intercollegiate Guidelines Network, 2010).

Studies have examined the accuracy of using 85th and 95th centile cut-offs as a diagnostic tool. A literature review by Reilly et al., for instance, points out that they contain high specificity (Reilly et al., 2002). This means that using these cut-offs is unlikely to result in misdiagnosing non-obese children as obese (low false-positive). However, there is evidence indicating the opposite; that is, the numbers of overweight and obese children have been overestimated at the cut-off points of 85th and 95th percentiles (Jebb and Prentice, 2001). In the 1997 national diet and nutrition survey of young people, Jebb and Prentice (2001) found that the prevalence of obesity in 16-year-old adolescents was 14% using the 95th percentile of United Kingdom standards, whereas only 5.8% were found to be obese when the adult BMI cut-off of 30 was applied.

Given the above-mentioned evidence, it remains difficult to determine the accuracy of the 85th and 95th centile cut-offs. Despite high specificity being indicated by the review, applying the centile cut-offs in adolescents may lead to overestimation of obesity prevalence relative to the adult definition of BMI≥30. Further research, examining the accuracy of the population specific definition, is required to ascertain the appropriateness of the measure.

Using the population-specific BMI centile cut-off is controversial among researchers due to the trade-off between its benefits and drawbacks. Clinical application favours this method because it minimizes the risks of unnecessary treatment and stigmatization resulting from diagnosing non-obese children as obese (Reilly, 2005). On the other hand, the population-specific cut-off does not allow comparison across populations because of differences in the BMI centile cut-offs for overweight and obesity

(Cole et al., 2000). The population-specific cut-offs allow researchers to decide where to set the cut-off for each population, leading to inconsistency in defining obesity and difficulty comparing the results across studies. Moreover, it is worth noting that a centile cut-off assumes fixed prevalence of obesity in the population, as 5% of children will invariably be obese if the 95th percentile is set as a diagnostic cut-off of obesity (Power et al., 1997).

The population-specific definition is supported by a significant amount of evidence, particularly in terms of relevancy to clinical significance, sensitivity, specificity, and advantages and drawbacks, possibly due to the definition having been utilized for a long time. Not only the quantity, however, but also the quality of this supportive evidence needs to be taken into consideration. Most publications studying the consequences of childhood obesity are based on cross-sectional or short-term follow-up studies (Reilly et al., 2003, Puhl and Latner, 2007). This may reflect a practical limitation, as a considerable amount of time and financial resources are required in order to obtain long-term follow-up data. The information from the short-term studies may not provide strong evidence of association due to a failure to demonstrate a long-term temporal effect. Moreover, body growth is another important element affecting obesity status in childhood (Scottish Intercollegiate Guidelines Network, 2010). Obese children may only need to maintain their weights in order to become leaner, as they are growing taller. Using more long-term follow-up data, particularly those that cover the transition from childhood to adulthood, allows adjustment for age and body growth when drawing associations.

In response to concerns raised regarding the use of the population-specific centile cut-off, an alternative definition of obesity in children has been recommended: the internationally standardized BMI cut-off. Cole et al. (2000) analyzed the trends of child growth from 6 national surveys, including Brazil, Great Britain, Hong Kong, the Netherlands, Singapore, and the US, and calculated the international BMI cut-off points for overweight and obesity by age and gender. The international cut-offs, also known as International Obesity Task Force (IOTF) cut-points, were constructed with the purpose of creating a sole classification for childhood obesity without regard to population and ethnicity (Reilly, 2005). Table 1.2 shows the internationally standardized cut-offs for males and females between ages 2 and 18.

Table 1.2: International cut off points for BMI for overweight and obesity by gender between ages 2 and 18

Age (years)	Body mass index 25 kg/m^2		Body mass index 30 kg/m^2	
	Males	Females	Males	Females
2	18.41	18.02	20.09	19.81
2.5	18.13	17.76	19.8	19.55
3	17.89	17.56	19.57	19.36
3.5	17.69	17.4	19.39	19.23
4	17.55	17.28	19.29	19.15
4.5	17.47	17.19	19.26	19.12
5	17.42	17.15	19.3	19.17
5.5	17.45	17.2	19.47	19.34
6	17.55	17.34	19.78	19.65
6.5	17.71	17.53	20.23	20.08
7	17.92	17.75	20.63	20.51
7.5	18.16	18.03	21.09	21.01
8	18.44	18.35	21.6	21.57
8.5	18.76	18.69	22.17	22.18
9	19.1	19.07	22.77	22.81
9.5	19.46	19.45	23.39	23.46
10	19.84	19.86	24	24.11
10.5	20.2	20.29	24.57	24.77
11	20.55	20.74	25.1	25.42
11.5	20.89	21.2	25.58	26.05
12	21.22	21.68	26.02	26.67
12.5	21.56	22.14	26.43	27.24
13	21.91	22.58	26.84	27.76
13.5	22.27	22.98	27.25	28.2
14	22.62	23.34	27.63	28.57
14.5	22.96	23.66	27.98	28.87
15	23.29	23.94	28.3	29.11
15.5	23.6	24.17	28.6	29.29
16	23.9	24.37	28.88	29.43
16.5	24.19	24.54	29.14	29.56
17	24.46	24.7	29.41	29.69
17.5	24.73	24.85	29.7	29.84
18	25	25	30	30

Source: Cole et al. (2000)

Classifying overweight and obesity by the IOTF cut-offs has some advantages in terms of feasibility when comparing obesity prevalence across countries, or even within the same population over time (Jebb and Prentice, 2001, Reilly et al., 2002). Despite these benefits, however, using IOTF cut-offs has been criticized for its disadvantages. One concern addresses the lack of supportive evidence at the international level regarding health consequences (Reilly et al., 2002). Researchers have also argued that obesity cut-offs should be derived from a population-specific reference so that differences in ethnic characteristics can be accounted for (Wang and Lobstein, 2006). However, an increasing number of studies has been published in support of the IOTF definition. Neovius et al. (2004), for example, tests the diagnostic accuracy of the IOTF classification. The results showed that the IOTF definition of overweight has high specificity for both sexes (0.95-1), while sensitivity is low for females (0.22-0.25, compared with 0.72-0.84 for males). This supportive clinical evidence encourages the use of IOTF definition.

In practice, the adoption of overweight and obesity definitions in children varies globally across health organizations. Table 1.3 lists the definitions of childhood obesity used by selected health organizations.

Table 1.3: Comparing the definition of overweight and obesity in children and adolescents

Sources*	**Definition**		**Population reference**	**Others**
	Overweight	**Obesity**		
RCPCH/NOF 2002	UK 91^{st}, 98^{th} and 99.6^{th} centile lines / IOTF definition		-	Both the UK and IOTF definitions are suggested
NHMRC 2003	>85^{th} (CDC 2002)	>95^{th} (CDC 2002)	USA	-
SIGN 2003	>91^{st} (UK 1990)	>98^{st} (UK 1990)	UK	-
AAP 2003	>95^{th}		Population of interest	BMIs 85^{th}-95^{th} are classified as at risk of overweight
AHA 2005	≥95^{th} (CDC)	-	USA	-
RNAO 2005	85^{th}-94^{th}	≥95^{th} (CDC)	USA	-
USPSTF 2005	≥95^{th}	-	USA	BMIs 85^{th}-94^{th} are classified as at risk of overweight

* AAP, American Academy of Pediatrics; AHA, American Heart Association; CDC, Centers for Disease Control and Prevention; NOF, National Obesity Forum; NHMRC, National Health and Medical Research Council (Australia); RCPCH, Royal College of Paediatrics and Child Health; RNAO, Registered Nurses Association of Ontario; SIGN; Scottish Intercollegiate Guidelines Network; USPSTF, United States Preventive Services Task Force

Source: modified from NICE: Obesity (2006)

As shown by Table 1.3, the population-specific definitions have been adopted by health organizations more than the international definition. This might be due to the relatively recent introduction of the IOTF definition in 2000, which may need more time before becoming more widely used. This should not detract from the use of the IOTF definition, as the body of supportive evidence tends to increase over time.

One suggestion pertaining to the use of the IOTF definition is to consider the diversity of a given population. This argument is based on the IOTF definition being constructed from multi-national data. Due to differences in ethnic distribution, the IOTF definition should be better fitted to populations with higher ethnic diversity, such as the US and UK; in 2009, the percentages of White, Black, and Asian populations in the US were 79.6%, 12.9%, and 4.6%, respectively (U.S. Census Bureau, 2010); ethnic minorities are also projected to comprise an increasing proportion of the UK population, from approximately 8% in 2001 to 20% in 2051 (Office for National Statistics, 2011a, Wohland et al., 2010). In contrast, in populations with little ethnic diversity, using a population-specific definition seems more appropriate.

The author therefore concludes that using the IOTF definition in countries with great population diversity, such as the US and UK, tends to be more beneficial in terms of encouraging a single standard of the childhood obesity definition. The population-specific definition, meanwhile, should be applied in countries with little population diversity in order to increase the accuracy of obesity diagnosis in children. In Chapters 3 and 4 of this thesis, the analysis of the UK-based data is carried out using the IOTF definition.

1.3 The development of obesity: nature or nurture?

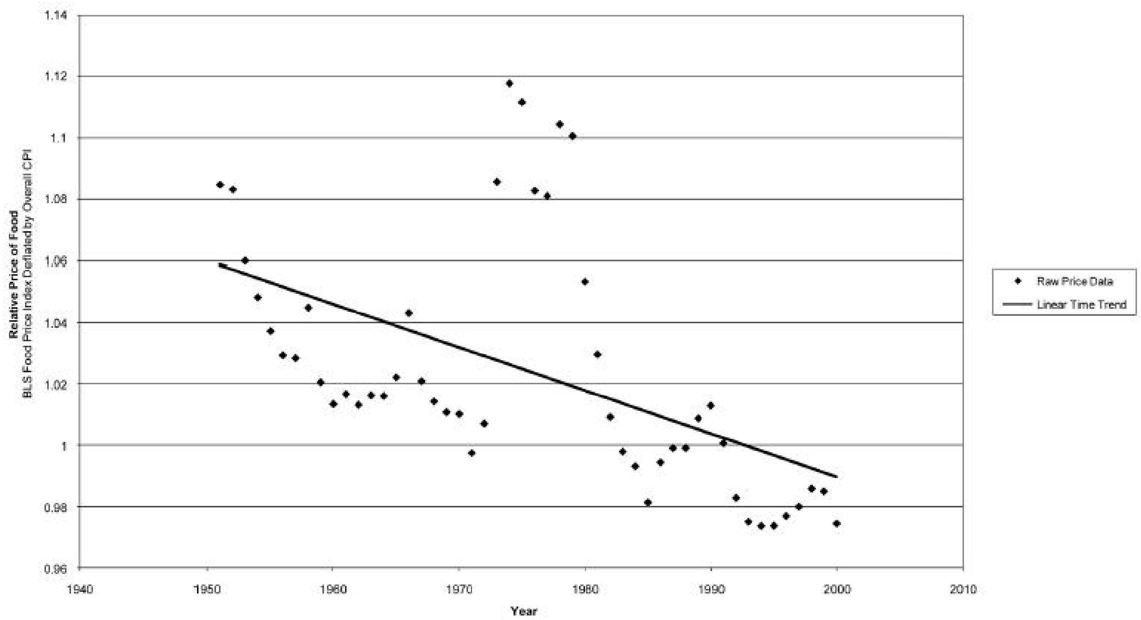
1.3.1 Economics of obesity: the influence of technology on energy counterbalance

Finkelstein et al. (2005) reviewed the causes of obesity and discussed the concept of energy counterbalance as a major mechanism. The development of obesity is regarded as an excess of energy input over outflow. The authors explain that the rise in the prevalence of obesity during the past decades is due to increased energy intake and decreased output in the general population (Finkelstein et al., 2005), with technological advancement considered to be the main contributing factor (Philipson, 2001, Lakdawalla and Philipson, 2002, Lakdawalla and Philipson, 2009). Energy expenditure in particular is likely to have been extensively affected by advancements in technology. A clear example is modern transportation, including both cars and public transit, which saves time and energy of travelling and shipping, compared to the past.

Lakdawalla and Philipson (2009) claim subsequently that technological change affects energy balance through changes in lifestyle and reductions in food prices. Modern lifestyles have seen a decrease in passive exercise due to a shift away from strenuous physical labour. In the past, workers completed passive exercise by doing routine strenuous work, which mostly required physical strength (Lakdawalla and Philipson, 2009). During the post-agricultural period, advanced technology has allowed the working process to become more efficient and less strenuous for humans, for both agricultural and industrial procedures (Lakdawalla and Philipson, 2009). To observe this trend, goods-producing industries, in contrast to service-providing industries, may be observed as a proxy of the strenuous working market (Finkelstein et al., 2005). The reduction in the proportion of employed wage and salary workers in goods-producing industries in the US, from 35% in 1960 to 27% and 19% in 1980 and 2000 respectively, reflects the shift away from manpower use in strenuous industries (Finkelstein et al., 2005). In addition to the decrease in passive physical activity through strenuous work, modern daily lifestyles have become much more convenient due to advanced technologies such as washing machines and dish washers (Finkelstein et al., 2005).

Not only have technological changes reduced time and energy spent in daily activities, but the hidden costs of doing physical activity have also increased (Lakdawalla and Philipson, 2009). This is due to the fact that modern workers would need to take time—and likely pay—away from their routine jobs in order to do enough exercise to maintain the same level of energy expenditure that was passively gained from strenuous work in the past. This would lead to loss of income or loss of leisure time for other activities.

The use of advanced technology in the agricultural process has also led to an increase in the supply of agricultural products, resulting in lowered food prices (Lakdawalla and Philipson, 2009). Lakdawalla and Philipson (2009) claim that the price of food has been decreasing since 1951, as seen in Figure 1.3, except during the 1970s when food prices rose in response to the oil shock. They argue that price reduction is related to an increase in food consumption and excessive energy intake (Lakdawalla and Philipson, 2009).



Source: Lakdawalla and Philipson (2009)

Figure 1.3: Changes in the relative price of food in the US, 1951-2000

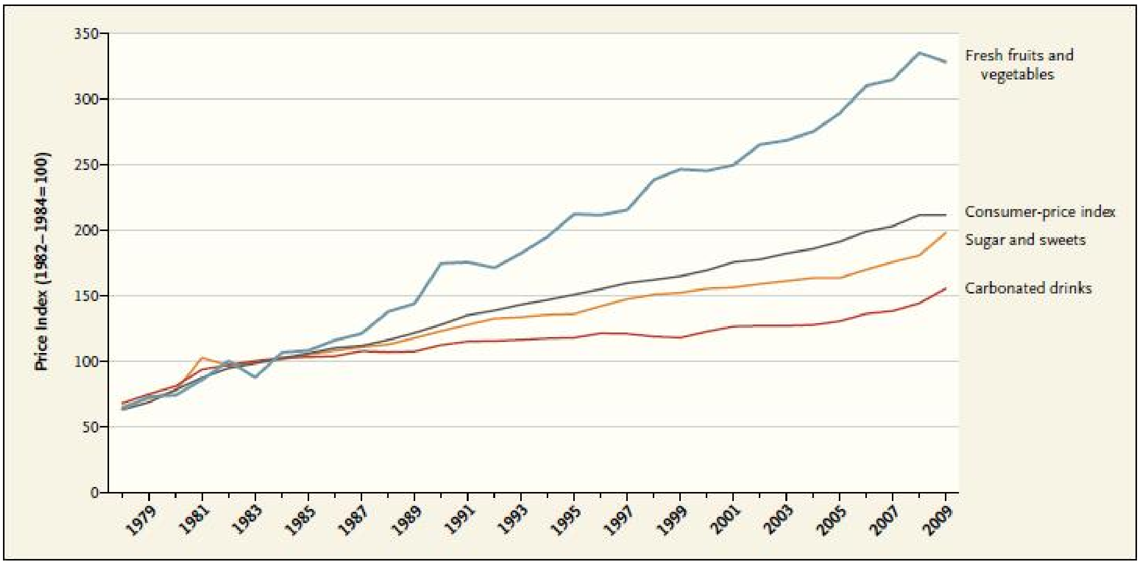
However, the claim that food prices have decreased relatively over time is arguable. The trend shown in Figure 1.3 raises 3 points of concern: 1) the starting point of the trend is too far in the past to make reliable comparisons when considering the consumer price index (CPI), or the relative price trend. A number of factors used in CPI estimation have changed over time, such as changes in spending patterns, price fluctuation, and new items added into the market (around 10,000 items each year) (USDA Economic Research Service, 2007). Adjusting for all these factors in the CPI calculation inevitably creates inconsistency in the factors and methods applied over a long period of time (USDA Economic Research Service, 2007). This is considered a limitation when comparing CPIs across extended periods. As seen in Figure 1.3, Lakdawalla and Philipson (2009) extended the plot over 50 years, making the CPI comparison vulnerable to these threats and the conclusion less reliable; 2) Lakdawalla and Philipson (2009) fail to provide the details of how the graph, Figure 1.3, is constructed, as well as the methods used to weigh the food and beverages items. In order to demonstrate this, Table 1.4 provides information about overall CPIs and CPIs of food and beverages, which were retrieved from the US Bureau of

Labor Statistics (Bureau of Labor Statistics, 2011). As shown, the CPIs of food and beverages have risen at the same rate as the overall rate of increase for all CPIs, with a relative ratio of approximately 1 during 1980-2010. This information is inconsistent with what is shown in Lakdawalla and Philipson (2009), which raises questions as to the reliability of the data and methods used to generate the graph (Figure 1.3); and 3) when separating the prices of food and beverages into categories, divergent price trends are seen. The relative price for fresh fruits and vegetables has increased constantly since 1985, while the prices of sugar and sweets and carbonated drinks have become relatively cheaper over time (Brownell and Frieden, 2009). Figure 1.4 displays the CPI trends for food and beverage products compared with overall CPIs.

Table 1.4: The US CPIs of all items and food and beverages from 1980 to 2010 (selected years)

Year	**Annual CPI**		**Relative ratio:**
	All items	**Food and beverages**	**Food&Beverages/All items**
1980	86.3	86.7	1.00
1985	109.3	105.6	0.97
1990	133.8	132.1	0.99
1995	153.5	148.9	0.97
2000	174.0	168.4	0.97
2005	196.8	191.2	0.97
2006	201.8	195.7	0.97
2007	210.0	203.3	0.97
2008	210.2	214.2	1.02
2009	215.9	218.2	1.01
2010	219.2	220.0	1.00

Source: Bureau of Labor Statistics (2011)



Source: Brownell and Frieden (2009)

Figure 1.4: Relative price changes for fresh fruits and vegetables, sugar and sweets, and carbonated drinks in the US, 1978-2009

Given the supportive evidence above, it is convincing that technological change has reduced energy expenditure by making work less strenuous and decreasing opportunities for passive exercise. However, the argument pertaining to the decrease in food prices due to advanced technologies is subject to a number of questions and inconsistencies. Contrary to what Lakdawalla and Philipson suggest, figures seem to indicate that, during the past 25 years, the relative prices of sugar and sweets and carbonated drinks have decreased, whereas the relative price of fresh fruits and vegetables has seen a noticeable rise.

Although changes in relative prices have been found, the effect of price change on energy counterbalance is still unclear. The next part explores the impact of price change on the demand for food, as well as health in terms of weight and BMI changes.

1.3.2 Price elasticity of demand for food and beverages

Price elasticity of demand (PED) is a concept that shows the extent to which consumption rates change in response to a change in the price of goods or services. Krugman and Wells (2009) define the price elasticity of demand as "the ratio of the percent change in the quantity demanded to the percent change in the price as we move along the demand curve (dropping the minus sign)" (Krugman and Wells, 2009). PED is calculated by:

$$PED = (\%\text{change in quantity demanded}) / (\%\text{ change in price})$$
$$= (\Delta Q/Q)/ (\Delta P/P)$$

Where, ΔQ is the change in quantity demanded

Q is the initial quantity demanded

ΔP is the change in price

P is the initial price

The PED is elastic when the absolute value of price elasticity is greater than 1, and inelastic when it is less than 1 (Krugman and Wells, 2009).

While the previous section shows the change in the prices of food and beverages, the impact of price on consumption is revealed by the PED for food and beverages. Andreyeva et al. (2010) conducted a systematic review of 160 studies on the price elasticity of demand for food. Of these, 99 studies used time series data, 34 were based on household survey data, and the other 27 were from scanner data (e.g. barcode scanner at the supermarket) (Andreyeva et al., 2010). Their results, displayed in Table 1.5, indicate that the demand for food is rather inelastic for all categories, as the absolute values of mean price elasticity estimates are below 1. In this regard, the demand for soft drinks and fruit is less inelastic than that for vegetables and sweets/sugar (Andreyeva et al., 2010).

Table 1.5: US price elasticity estimates, by food and beverage category, from 1938–2007

Food and beverage category*	**Absolute value of mean price elasticity estimate (95% CI)**	**Range**
Food away from home	0.81 (0.56,1.07)	0.23–1.76
Soft drinks	0.79 (0.33,1.24)	0.13–3.18
Juice	0.76 (0.55,0.98)	0.33–1.77
Beef	0.75 (0.67,0.83)	0.29–1.42
Pork	0.72 (0.66,0.78)	0.17–1.23
Fruit	0.7 (0.41,0.98)	0.16–3.02
Poultry	0.68 (0.44,0.92)	0.16–2.72
Dairy	0.65 (0.46,0.84)	0.19–1.16
Cereals	0.6 (0.43,0.77)	0.07–1.67
Milk	0.59 (0.40,0.79)	0.02–1.68
Vegetables	0.58 (0.44,0.71)	0.21–1.11
Fish	0.5 (0.30,0.69)	0.05–1.41
Fats/oils	0.48 (0.29,0.66)	0.14–1.00
Cheese	0.44 (0.25,0.63)	0.01–1.95
Sweets/sugar	0.34 (0.14,0.53)	0.05–1.00
Eggs	0.27 (0.08,0.45)	0.06–1.28

*Including restaurant meals and fast food

Note: The price elasticity of demand measures the percentage change in purchased quantity or demand with a 1% change in price.

Source: modified from Andreyeva et al. (2010)

As the evidence demonstrated earlier, the relative prices of fresh fruits and vegetables have increased during the past few decades, while an opposite trend is observed for soft drinks and sweets/sugar. Taking price elasticity into account, these changes in price subsequently lead to a decrease in the consumption of fruits and vegetables and increasing demand for soft drink and sweets/sugar. The change in consumption patterns as a result of price change is a mechanism used to explain the impact of food prices on energy counterbalance, weight, and BMI change.

It is noted that, apart from the effect of food prices on consumption demand, there are other factors, such as socio-economic status (SES) and individual consumer behaviours, which impact the magnitude of the price effect. Financial status may in part determine the size of the price effect; that is, increased prices tend to reduce consumption to a greater extent among the poor compared with the rich. In addition, consumption is considerably affected not only by SES, but also by consumer behaviour. Raising the prices of high-calorie foods may not reduce calorie consumption if consumers choose to buy in bulk at reduced prices, wait for the items to be on sale, or switch to other substitute foods with equally high calories.

Another issue worth noting with regard to this change in consumption patterns is that children do not really make all decisions themselves. Most behaviours, such as food consumption and exercise, are largely influenced by their parents, who act as both mediators and confounders (Golan and Crow, 2004). Even though the price effect leads to changes in consumption, parental influence is regarded as very crucial. This is to clarify the implication of evidence showing the effect of food price on child consumption that already accounts for the hidden impact of parents' decisions.

Sturm and Datar (2011) assessed the importance of food prices, including fruits and vegetables, fast food, and sugar-sweetened beverages, for child consumption using the US national cohort data from 1998-9 to 2004. They report that a significant decrease in the consumption of fruits and vegetables is found when their price increases by 1 standard deviation; however, no significant difference is seen for sugar-sweetened beverages (Sturm and Datar, 2011). Results from the study imply that an increase in fruits and vegetables prices reduce consumption, while evidence does not show any change in sugar-sweetened beverages consumption in response to a change in price.

The effect of the confounder, SES, might play a crucial role in explaining the price effect on the consumption demand of sugar-sweetened beverages. A US study examining the impact of sugar sweetened beverage taxes on weight reveals the stratification of the price effect by income level (Finkelstein et al., 2010). The study shows that taxes on sugar sweetened beverages affect the consumption and weight of the middle-income households the most, whereas only a small impact can be observed in the poorest and richest 25% of US households. The authors attribute this failure of the price mechanism to reduce consumption and weight in the poorest 25% to purchasing behaviour: facing higher prices, consumers might instead find cheaper alternatives, such as buying generic brands, purchasing in bulk, or waiting until the items are discounted (Finkelstein et al., 2010).

Concerning the impact of food price on body weight and BMI, a systematic review, including 9 studies, was conducted to explore this association (Powell and Chaloupka, 2009). Of these, only 4 studies, 2 cross-sectional and 2 longitudinal, examined the association in children and adolescents. Two longitudinal studies revealed a significant association between increased price of fresh fruits and vegetables and raised BMI in kindergarten children over a 6-year follow-up (Sturm and Datar, 2005, Sturm and Datar, 2008). However, the effect of price on weight and BMI change is very modest even if a significant association exists (Powell and Chaloupka, 2009).

A significant effect of fruit and vegetable price on BMI was also echoed in other studies (Powell and Bao, 2009, Sturm and Datar, 2011). Powell and Bao (2009) reported that a 10 percent increase in fruit and vegetable price is associated with 0.7 percent increase in BMI in children, or an approximate weight change in 12 year-old boys and girls by 0.4 and 0.5 pounds, respectively.

In this regard, the impact of price on the issue of equity is also worth noting; that is, studies show that the influence of fruit and vegetable price changes on consumption and BMI increases in children from low socio-economic background in comparison with their high SES peers (Powell and Bao, 2009, Sturm and Datar, 2011). This finding echoes the importance of SES in estimating the magnitude of price effect.

To examine the impact of food price on BMI over time, Goldman et al. (2009) use the US panel data to assess the effect of changes in the price of calories on BMI over short-term and long-term periods. They reported that, within 2 years, a 10% reduction in price per calorie results in an increase in BMI of 0.22 units. BMI would increase by 1.5 units after 30 years if price per calorie dropped by 10%. The long-term effect of a 10% price-per-calorie change is 1.9 units of BMI change (Goldman et al., 2009).

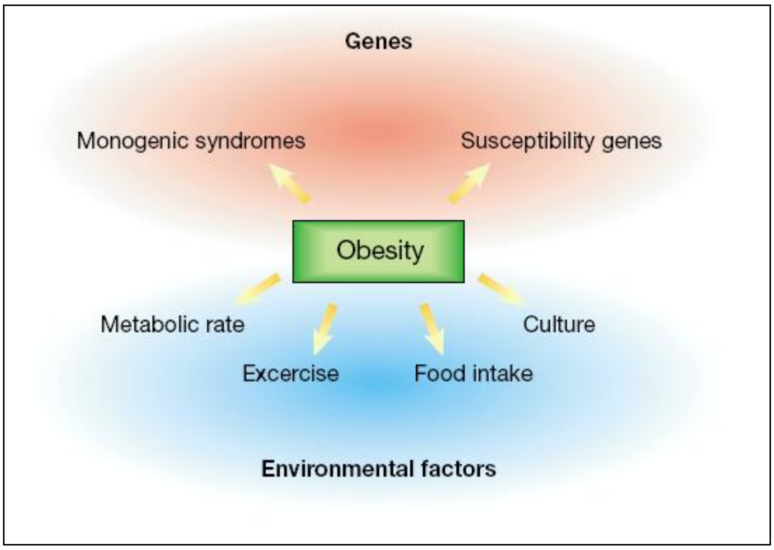
Goldman et al. demonstrate that only a modest effect on BMI will be seen in the short term when raising the price per calorie. The question then arises as to whether or not this tiny change in BMI has any clinically significant impact on health. Despite the magnitude of BMI change being greater in the long term, a BMI change of 1.5-1.9 units for life is nevertheless insufficient to make any assumptions regarding clinical significance. Furthermore, a number of other factors may also contribute to the change in BMI during this long period, e.g. changes in lifestyle, environment, social norms, and especially body growth in children. Food price alone, therefore, maybe not be an accurate predictor for BMI change in the long run, making the long-term estimates by Goldman et al. somewhat unconvincing. Therefore, from the policy perspective, raising price per calorie with the objective of reducing obesity may not be an effective policy.

In sum, existing literature indicates a modest effect of food price on BMI change. Some evidence indicates that an association exists between the price of fruits and vegetables and changes in consumption rate and BMI, though the degree of association is small. This implies that alterations of food and beverage prices during the past decades may be able to explain obesity in childhood to only a limited extent. Hence, there might be other factors involved in the increase in the prevalence of childhood obesity.

1.3.3 Other factors associated with the development of obesity

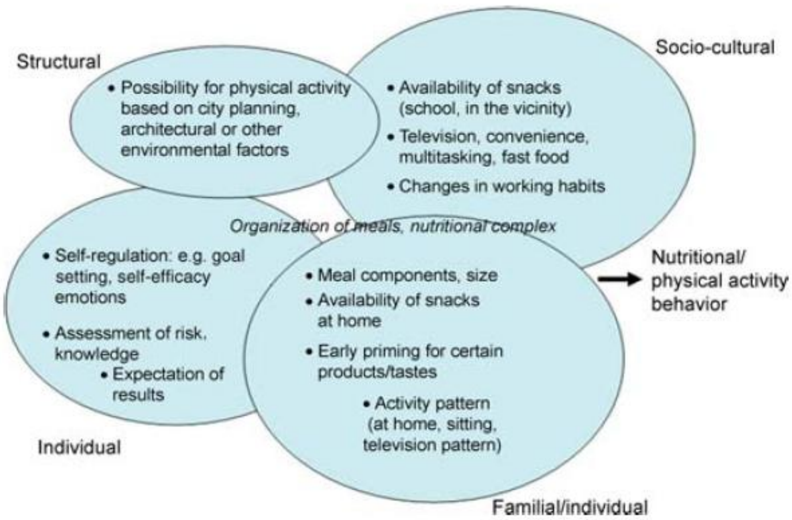
The modern conception of obesity development suggests that complex interactions of genetic, environmental, psychological and behavioural factors are responsible for the development of obesity (Kopelman, 2000, Yang et al., 2007). All these related attributes influence body weight through the physiological pattern of energy intake and outflow (Kopelman, 2000). This hypothesis has been supported by several studies (e.g. Kopelman, 2000, Parlesak and Kromker, 2008). Although there is evidence showing that the development of obesity is attributable to a number of causes, the relationships between these related factors have not yet been clearly demonstrated (Parlesak and Kromker, 2008).

Figure 1.5 shows the various factors influencing the development of obesity. It can be seen that related factors are mainly categorized into 2 groups: genes and environment. While energy balance is commonly used in describing the mechanism of obesity development, other associated genetic and environmental factors are silently involved in the process. Figure 1.6 shows the psychological attributes associated with obesity. As illustrated in the diagram, such environmental factors as family, workplace, and community play an important role in leading individuals to become obese.



Source: Kopelman (2000)

Figure 1.5: Factors influencing the development of obesity



Source: Parlesak and Kromker (2008)

Figure 1.6: Multifactorial psychosocial medley of causes in the development of overweight and obesity

1.3.4 Factors associated with obesity

1.3.4.1 Genetic factors

Genetic factors play a significant role in the development of obesity. The mechanism of susceptible genes for obesity was first mentioned almost 50 years ago by Neel (1962) with the concept of the thrifty genotype hypothesis, according to which the thrifty gene that increases susceptibility to obesity was selectively passed from primitive groups that had greater levels of activity or experienced starvation. Researchers have in fact demonstrated that obesity is somewhat as heritable as height and more heritable than many diseases that are commonly acknowledged as genetic disorders (Friedman, 2003). Stunkard et al. (1986) showed that the heights, weights and BMIs of male twins are highly correlated across time. Moreover, the impact of heritability was estimated to be 16-85% for BMI, 37-81% for waist circumference, 6-30% for waist-hip ratio and 35-63% for the percent of body fat (Yang et al., 2007).

In terms of genetic aetiology, obesity is classified into 3 groups as follows: 1) monogenic obesity; 2) syndromic obesity; and 3) polygenic obesity, or common obesity (Ichihara and Yamada, 2008, Bell et al., 2005b).

Monogenic obesity results from a single gene mutation whose related diseases usually occur during childhood (Mutch and Clement, 2006, Yang et al., 2007). This type of obesity accounts for a tiny fraction of obesity cases (around 200 reported cases until 2006); however, it leads to considerable adverse health conditions as well as health costs (Mutch and Clement, 2006, Farooqi and O'Rahilly, 2007). In their review, Yang et al. (2007) revealed that 176 single-gene-mutation-related obesity cases involving 11 different genes were reported up until 2005. The benefits of discovering this type of obesity are significant, as was shown in the treatment of a 9-year-old English girl with congenital leptin deficiency, a monogenic disease characterized by hyperphagia, excessive weight gain early in life, and severe obesity (Farooqi et al., 1999). After a year of treatment with daily subcutaneous leptin injections to the 94-kg girl, her appetite was dramatically reduced and her obesity successfully reversed, with sustained fat loss and a total weight loss of 16.4 kg (Farooqi et al., 1999).

Unlike monogenic obesity, the characteristics of syndromic obesity include distinct abnormalities other than obesity, such as mental retardation, dysmorphic features and organ-specific developmental abnormalities (Bell et al., 2005b). Abnormalities in genes or chromosomes lead to autosomal or X-linked disorders, which incur at least 20 genetic obesity syndromes (Bell et al., 2005b, Mutch and Clement, 2006). Among these syndromes, Prader-Willi syndrome (PWS) is the most frequent, with an incidence of 1 in 25,000 births (Bell et al., 2005b). PWS is characterized by obesity, hyperphagia, diminished fetal activity, muscular hypotonia, mental retardation, short stature and hypogonadotropic hypogonadism (Bell et al., 2005b, Ichihara and Yamada, 2008).

The last and most common type of genetic obesity is polygenic obesity. Polygenic obesity refers to obesity that is caused by the effects of various genes in a susceptible environment (Mutch and Clement, 2006, Ichihara and Yamada, 2008). Despite it being the major contributor to genetically related obesity, research regarding polygenic obesity has not advanced as much as for other types of genetic obesity (Farooqi and O'Rahilly, 2007, Ichihara and Yamada, 2008, Mutch and Clement, 2006). The issue of replication appears to be the main problem, as numerous genes and environmental factors are involved in the manifestation of polygenic obesity. Insufficient sample size and differences in study design have thus led to inconsistent results (Mutch and Clement, 2006, Ichihara and Yamada, 2008).

Despite their serious health consequences, the prevalence of monogenic and syndromic obesity is deemed very small in the population. In contrast, polygenic obesity is likely to account for a huge proportion of obesity. However, as stated earlier, the evidence supporting this genetic type as the cause of obesity is weak. Hence, it is still unclear as to what extent genetic factors affect the development of obesity.

1.3.4.2 Mental health factors

The connection between obesity and mental illnesses such as depression is explained by the anecdotal association with poor eating habits, physical inactivity, and low compliance with medication and recommended lifestyle modifications (Wyatt et al., 2006). A review exploring the association and relationship of obesity with psychiatric disorders found that major depressive disorders, as well as bipolar disorder, are associated with increased risk of being overweight in children and adolescents (McElroy et al., 2004). This association was demonstrated by a cohort study following 90 children aged 6-17 with major depression, compared with their 87 controlled peers, for 10-15 years (Pine et al., 2001). McElroy et al. (2004) also pointed out in their review that, though relatively weak, research has shown a causal relationship between obesity and certain types of mood disorders, with major depression in particular as a likely contributor to obesity.

Concerning psychiatric treatment, evidence shows that medications used in the treatment of mental disorders can also induce obesity. Schwartz et al. (2004) reviewed the effects of psychiatric medications on obesity and reported that using mood stabilizers, antipsychotics and antidepressants is associated with weight gain. The magnitude of the medications' effect on weight gain, however, varies depending upon the types of drugs used (Schwartz et al., 2004). The biochemical explanations for the aetiology of various psychiatric-medications-induced obesity are described in Virk et al. (2004).

The above-mentioned evidence showing particular psychiatric disorders, especially major depression, and psychiatric drugs to cause childhood obesity is derived from systematic reviews that were rated as good quality. It is thus reasonable to conclude that psychiatric problems and medical treatment lead to the development of obesity in children and adolescents. These findings alert those involved, such as parents and medical personnel, to monitor the effects of psychiatric conditions, with the aim of providing proper management for weight-related side effects alongside mental health treatments.

1.3.4.3 Socio-economic factors

The relationship between socio-economic status (SES) during early childhood and BMI in adulthood was examined by a longitudinal study (Laitinen et al., 2001). The results showed that significant predictors for being obese at age 31 include low SES of the child's family, a high maternal BMI before pregnancy, a high BMI during adolescence, and early menarche (Laitinen et al., 2001). Additional evidence indicates that the incidence of obesity is negatively correlated with SES (Cahnman, 1968), and that the prevalence of obesity and overweight is relatively high in people with low income and low education

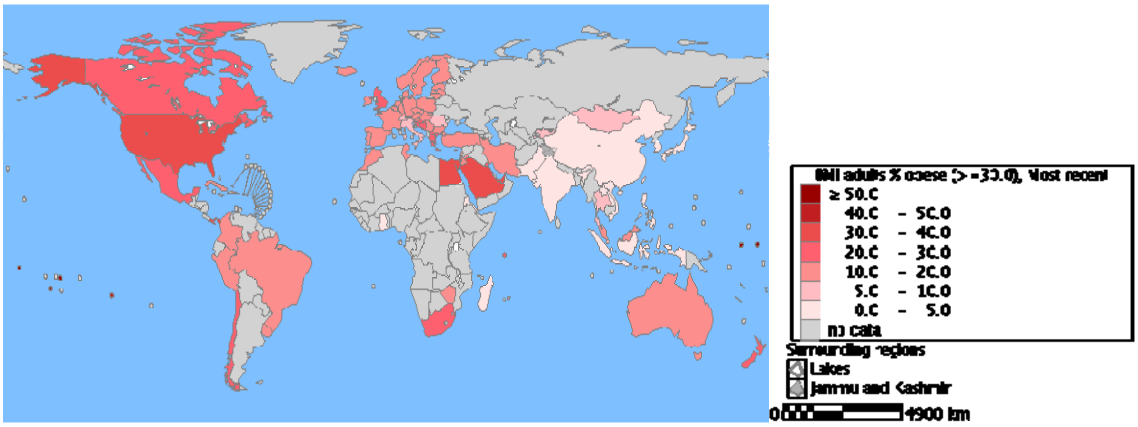
(Seidell and Flegal, 1997). This leads to the conclusion that low SES in childhood may be related to the occurrence of adult obesity.

The influence of SES on obesity is particularly remarkable in females. A US-based study by Goldblatt et al. (1965) reported that women and men of low SES were 6 and 2 times more likely to be obese than those of high SES, respectively. A study also showed the prevalence of obesity in females in England and Wales who came from low SES to be 25%, and those from high SES to be 10.7% (Kopelman, 2000).

It is evident that SES is inversely associated with the prevalence of obesity. One explanation could be the increased costs of physical activity as a result of technological changes. As discussed earlier, advanced technology has caused work to become less strenuous, leading to diminishing passive exercise from work. People need to find extra time, and perhaps even pay extra, to engage in levels of physical activity that would allow them to maintain energy balance. This then links to SES because lower-income individuals are more likely to be employed in low-paid positions with longer hours, limiting their opportunity for physical activity; extra costs present a further barrier. Reduced access to physical activity may explain the association between SES and obesity development.

1.3.4.4 Demographic factors: Age, Gender and Ethnicity

Demographic distribution appears to be another important factor associated with the prevalence of obesity. Figure 1.7 shows the geographic distribution of obesity in adults globally. As can be seen, obesity is prevalent in the western world, including the UK, Europe, and Australia.



Source: World Health Organization (2011a)

Figure 1.7: Geographic distribution of adulthood obesity worldwide

Moon et al. (2007) analyzed the 1998 and 1999 Health Survey for England data and reported the association between obesity and certain demographics, including age, gender, and ethnicity. They found an increasing trend of obesity and overweight in relation to age, which drops slightly in older age groups. Weight gain in different age groups also incurs different levels of risk. The results from the Framingham Heart Study showed that gaining one pound of weight between ages 30 and 42 increases the risk of death within 26 years by 1%; between ages 50 and 62, the risk increases by 2% (Kopelman, 2000).

Concerning gender, it has been clearly shown that females are more likely to be obese than males, whereas males have a higher prevalence of overweight than females (Moon et al., 2007, Seidell and Flegal, 1997, Wang and Beydoun, 2007). A systematic review and meta-analysis based on the US population showed that the prevalence of obesity in females is higher than in males (42% versus 34%), whereas males have a higher prevalence of overweight than females (74% versus 69%) (Wang and Beydoun, 2007). These findings correspond to studies conducted in the UK. Moon et al. (2007) revealed that obesity is found in 20% of females compared with approximately 17% in males, while nearly 50% of men in the UK are overweight compared with only 36% of women (Moon et al., 2007).

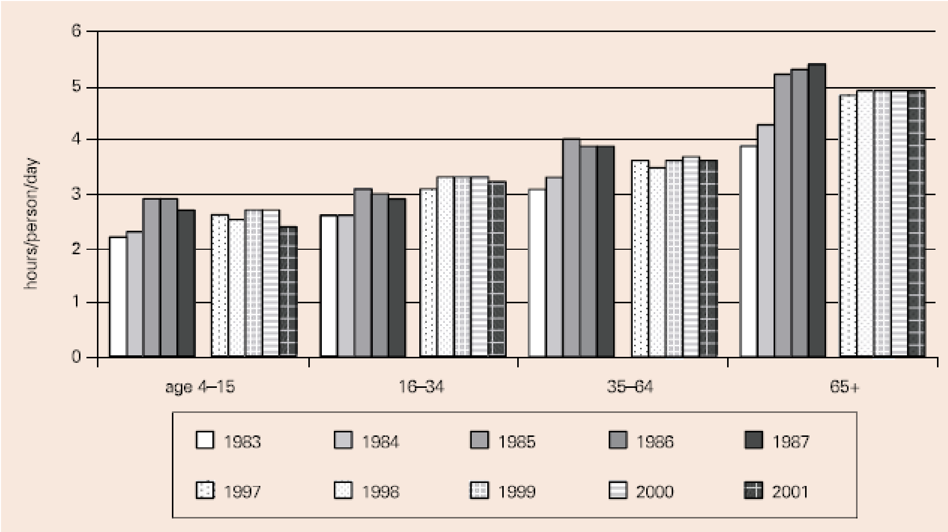
Regarding ethnicity, the prevalence of obesity and overweight varies across different ethnic groups, though the association between ethnicity and obesity is still unclear (Seidell and Flegal, 1997). Several US studies report that the prevalence of overweight and obesity in African-American populations is higher than that in Caucasian, Hispanic and Asian populations (Matthews, 2002, Gordon-Larsen et al., 2002). On the other hand, a UK-based study found no difference in the prevalence of obesity and overweight among White, Black, and Asian populations (Moon et al., 2007).

1.3.4.5 Sedentary behaviour – TV watching

TV watching is one of the most common sedentary behaviours to which obesity in adolescents and children might be attributed. Statistics show that children between ages 2 and 7 spend approximately 2.5 hours per day in front of the television; the number increases to 4.5 hours for children between the ages of 8-18 (Bryant et al., 2007, Robinson, 2001).

Epstein et al. (1995) conducted an RCT exploring the relationship of sedentary behaviours with physical activity in children and adolescents. They reported a significant reduction in time spent on sedentary behaviour, including television watching, in the intervention group (Epstein et al., 1995). The participants in the intervention also lost weight at a significant level (Epstein et al., 1995). Similar results were subsequently found by studies from Robinson (1999b) and Gortmaker et al. (1999b).

Nevertheless, evidence showing time spent on TV watching is somewhat inconsistent across studies. The Foresight report reveals the trend of TV watching time in the UK by age group, as demonstrated in Figure 1.8. From 1983 to 2001, the amount of time spent watching TV by children between ages 4 and 15 remained constant at 2-3 hours per day. In people between ages 16 and 34, less than an hour increase in TV watching time can be observed from 1983 to 2001. Despite the apparent association between increased TV watching time and obesity development, therefore, TV watching itself may not be considered the main contributing factor to childhood obesity in the UK. In addition, TV watching time in children and adolescents during the past decade may increasingly have been replaced by computer games and internet usage (Willoughby, 2008). Further research to explore the associations between time spent on computer games/internet and childhood obesity would shed more light on this.



Source: Foresight report (Lobstein and Leach, 2007)

Figure 1.8: Leisure-time broadcast television watching per person per day in the UK, 1983-1987 and 1997-2001, by age group

1.3.4.6 Physical activity and Exercise

Physical activity and exercise are major factors in the balance of energy expenditure. It has been estimated that 20-50% of energy outflow is attributable to physical activity (Kopelman, 2000). The relationship between exercise and weight loss has been suggested by several studies, although no causal inference has yet been found. Votruba et al. (2000) reviewed 11 articles pertaining to the effect of exercise on weight loss. They reported that exercise significantly contributed to some weight loss in 6 studies, whereas the other 5 studies showed no significant effect from exercise (Votruba et al., 2000). Another meta-analysis that explored the influence of exercise on weight reduction suggested that aerobic exercise alone can reduce weight (Garrow and Summerbell, 1995). Sherwood and colleagues (2000) conducted a cross-sectional study to examine the effects of calorie intake and exercise on body weight. The results showed that high-intensity physical activity was the only significant predictor in men, whereas daily energy intake and moderate- to high-intensity physical activity were both significant predictors in women (Sherwood et al., 2000).

To summarize, a number of studies have pointed out that factors other than technological changes also influence the development of childhood obesity. Evidence of association is found in mental health factors, particularly those related to major depression and the use of psychiatric medication, as well as in physical activity. Some evidence also indicates an association with genetic factors, SES, demographic factors, and TV watching. This information is helpful in designing interventions to reduce obesity in childhood more efficiently; with more knowledge of how childhood obesity develops, health professionals can better adopt different strategies to address its underlying causes, such as incorporating physical activity education and providing mental health support at schools in an attempt to reduce stress. Mental health support, for instance, can be in the form of building children's self-esteem and managing bullying. One example of a child obesity intervention that translates these attributes into practice is MEND 7-13, which is used as a case study for cost-effectiveness analysis in this thesis. Details of the intervention are described in Chapter 3.

The development of obesity in childhood involves multi-faceted interactions among individual and environmental factors. This aetiology implies that it may be unrealistic to expect a "silver bullet" solution that addresses the complexity of childhood obesity development. Collaboration among stakeholders, using multidisciplinary approaches to tackle the multi-faceted roots of the problem, is suggested to achieve the goal of reducing childhood obesity prevalence.

1.4 Undesirable health consequences of obesity in childhood

1.4.1 Medical problems

Comprehensive literature reviews of the health consequences of childhood obesity have been conducted by Reilly et al. (2003) and Han et al. (2010), which revealed that childhood obesity is associated with a number of adverse health outcomes in both the short and long terms.

The short term effects of childhood obesity include increased risks of psychological and psychiatric consequences, cardiovascular problems, and other undesirable health effects (Reilly et al., 2003, Han et al., 2010). In terms of long-term health impact, the reviews showed that childhood obesity is not only associated with increased risk of adverse health effects in adulthood, including persistent obesity, asthma, type 2 diabetes, cardiovascular risks, and mortality/ morbidity, but also related to low socio-economic status compared with the normal-weight population (Reilly et al., 2003, Han et al., 2010).

Focusing on the risk of obesity persistence, Dietz (1994) suggests a hypothetical mechanism of this association: because childhood is a critical period of adiposity rebound, childhood obesity leads to increased adiposity development in the future. Studies have investigated the relationship between childhood and adulthood obesity. Serdula and colleagues (1993) reviewed 17 prospective cohort studies that explore this association. They reported that all reviewed studies consistently showed a positive relationship between childhood and adulthood obesity. The results indicated that approximately one-third of obese preschool children and half of obese school-age children became obese when they reached adulthood. Moreover, the risk of becoming obese adults was 2 to 6 times higher for obese children than for non-obese peers (Serdula et al., 1993). The estimated effect of childhood obesity on the development of obesity in adults was found to be even stronger in a more recent study. Whitaker et al. (1997) reported that obesity in children between ages 10 and 17 increases the risk of young adulthood (ages 20-29) obesity by 17 to 22 times, regardless of parental obesity. More recently published studies also provide supportive evidence of the increased risk of obesity persistence in adulthood (Freedman et al., 2001, Guo et al., 2002).

In addition, there is some evidence that girls are at a higher risk of developing obesity in adulthood if obesity presents during adolescence (Dietz, 1994). A retrospective cohort study showed that 21-25% of obese adults were obese in early adolescence (Braddon et al., 1986, Freedman et al., 2001); among the obese adults, only 10% of males compared with 30% of females were obese at a young age (Braddon et al., 1986).

Obesity persistence and adverse health consequences of obesity in adulthood

Increase in risk of obesity persistence in adulthood is of particular interest because it is proven that adult obesity leads to a number of adverse health effects. Obesity and overweight increase risks of serious health problems in adulthood such as hypertension, dyslipidemia, diabetes, cardiovascular diseases (CVD), cerebrovascular diseases, osteoarthritis and some cancers, e.g., breast, prostate and colon cancer (National Heart Lung and Blood Institute, 1998). These risks are especially pronounced in moderate and severe obesity groups. Table 1.6 summarizes the increased risk of developing obesity-related diseases.

Table 1.6: The risk of developing diseases in obese adults relative to the non-obese

Disease	Relative risk	
	Women	Men
DM type 2	12.7	5.2
Hypertension	4.2	2.6
Myocardial infarction	3.2	1.5
Colon cancer	2.7	3
Angina	1.8	1.8
Gall Bladder diseases	1.8	1.8
Ovarian cancer	1.7	-
Osteoarthritis	1.4	1.9
Stroke	1.3	1.3

Source: the Comptroller and Auditor General (2001).

Calle and associates (2003) found that people with a BMI of at least 40 had an approximately 1.6 times higher risk of cancer death than those with a BMI of 18.5 to 24.9.

They also asserted that overweight and obesity in the US contributed to 14 and 20 percent of cancer deaths in men and women respectively in 1999-2000 (Calle et al., 2003).

While no association has been found between overweight and reduced life expectancy (Flegal et al., 2005), researchers have pointed out the relationship between obesity and a reduction in life expectancy. Both life insurance data and epidemiological studies describe obesity and overweight as strong predictors of life expectancy (Kopelman, 2000).

A study using data from the Framingham Heart Study showed the estimated impact of obesity on life expectancy for people between the ages of 30 and 49 (Peeters et al., 2003). They found that the life expectancy of obese participants decreases by 6 to 7 years relative to non-obese participants, resulting in 11 to 18 additional deaths per 100 people (Peeters et al., 2003).

As obesity reduces life expectancy, the number of deaths increases as a consequence when the obese population rises. Mokdad et al. (2004) reviewed the leading causes of death that can be prevented by behavioural change in the US and found obesity and overweight as the second leading cause after tobacco. They also reported that the number of obesity-related deaths rose from approximately 300,000 (14% of total deaths) in 1990 to 400,000 (16.6% of total deaths) in 2000 (Mokdad et al., 2004). However, the number of deaths attributable to obesity in the US varies among studies, possibly due to differences in the estimation methods used (Wyatt et al., 2006). For example, the annual number of US deaths linked to obesity by Allison el al. (1999) was 300,000, whereas Flegal and colleagues (2005) estimated the number of deaths per annum to be 110,000.

In the UK, it was estimated that obesity leads to an approximate 9-year reduction of life expectancy, with around 30,000 deaths per annum resulting from obesity (National Audit Office, 2001, Moon et al., 2007).

1.4.2 Sociological problems

The causal relationship between childhood obesity and social stigma was reviewed by Puhl and Latner (2007). This systematic review was chosen because of its recent update, as well as the comprehensiveness of information in the review. The authors split the social consequences of weight bias into categories and rate the level of evidence by the strength of association/relationship provided in the studies (Puhl and Latner, 2007).

Despite a dearth of rigorous evidence, the review identifies causal relationships between childhood obesity and bullying; children's negative attitudes toward themselves, as well as obese and overweight peers; and lower socio-economic status (Puhl and Latner, 2007).

Many cross-sectional studies have found evidence supporting the impact of weight bias on the bullying of obese children (Pearce et al., 2002, Hayden-Wade et al., 2005, Janssen et al., 2004). In this regard, the temporal sequence of the relationship has been demonstrated by a British longitudinal study that followed 8,210 children from age 7.5 to 8.5 to observe bullying patterns (Griffiths et al., 2006). They found that obese boys and girls are 1.5 times more likely to be overtly bullied than their average-weight peers. Moreover, obese boys are 1.66 times more likely to bully others overtly compared with their average-weight counterparts (Griffiths et al., 2006).

In addition to bullying and victimization, obese children are also subjected to negative attitudes from close contacts, including their friends, teachers, and parents. Biases and negative attitudes toward obesity have been found in children as early as age 3, as summarized in the review by Puhl and Latner (2007). The authors revealed that reviewed studies had applied different methods, such as telling stories, describing pictures, assessing playmate preferences, and direct interviews, to obtain the information used in the cross-sectional and experimental studies. The study showed that children and adolescents have negative impressions toward their obese peers, describing them as being mean, lazy, ugly, stupid, lying, less attractive, and having lower self-esteem (Puhl and Latner, 2007). In addition, cross-sectional studies have found that negative attitudes toward the obese are even more extensive in young adulthood. College-age students were interviewed about their attitudes toward obese colleagues, with results showing that, in addition to the negative impressions expressed in childhood, obese adolescents are claimed to be less attractive, less likely to be dating, sexually unskilled, and deserving of heavier and less attractive partners (Neumark-Sztainer et al., 1998, Harris, 1990, Regan, 1996, Puhl and Latner, 2007).

Weight bias from parents is perhaps the most unexpected finding. Two survey studies report weight-bias offence from family members (Neumark-Sztainer et al., 2002). The study, containing 4,746 adolescent samples, revealed that 47% and 34% of very overweight girls and boys reported being teased about weight by family members (Neumark-Sztainer et al., 2002). Evidence also shows that financial support given to overweight girls was less than those to their normal-weight peers (Puhl and Latner, 2007). One study interviewed the parents of overweight children to explore this matter and reported that parents felt guilty about their child's obesity (Pierce and Wardle, 1997). Anger and frustration also appear in some parents, especially those who did not know how to deal with the problem (Pierce and Wardle, 1997).

Additionally, childhood obesity leads to low SES in adulthood. Gortmaker et al. (1993) conducted a 7-year follow-up in overweight adolescents to explore social and economic consequences. They showed that overweight women received fewer years of schooling, were less likely to get married, had lower household incomes, and had higher rates of household poverty (Gortmaker et al., 1993).

The evidence provided above shows a strong link between childhood obesity and the problems of bullying and being bullied. Another concern worth noting is the increased risk of low SES attributable to obesity; together with low SES having been identified as a risk factor for childhood obesity, this evidence points to a vicious cycle between obesity and low SES, which urges closer attention to be paid to higher-risk groups. Efforts to stop this cycle may help reduce the impact on both the obese themselves and those they affect.

1.4.3 Psychological problems

Unlike physical effects, the impact of childhood obesity on psychological well-being is much less evident. This finding has been corroborated by a number of reviews (Friedman and Brownell, 1995, French et al., 1995, Ricciardelli and McCabe, 2001, Wardle and Cooke, 2005). A recent literature review, examining the psychological consequences of obesity in children and adolescents, summarizes research findings, which are mainly in terms of body satisfaction, self-esteem, and depression (Wardle and Cooke, 2005).

Studies during the past decades have revealed a small association between childhood obesity and body dissatisfaction (Friedman and Brownell, 1995, Ricciardelli and McCabe, 2001). The reviews reported a lack of supportive evidence for a temporal sequence of association, with the scarcity of longitudinal studies being noted (Ricciardelli and McCabe, 2001). These results are echoed in a more recent literature review by Wardle and Cooke (2005), consisting of 18 studies, 5 of which were prospective studies. Results from the review mirror the findings from the previous reviews, indicating a tiny degree of association (Wardle and Cooke, 2005).

Concerning self-esteem, a review of 35 studies reports a small effect of childhood obesity on self-esteem (French et al., 1995). The authors note that the esteem scores of obese children are within the normal range, despite being lower than those of the normal-weight (French et al., 1995). Similar conclusions were reached by a more recent literature review, which included 28 additional studies, 7 of which were prospective studies (Wardle and Cooke, 2005). In this regard, Wardle and Cooke point out the influence of gender and age on the self-esteem level: girls and older children may be at higher risk of having low self-esteem in relation to their counterparts (Wardle and Cooke, 2005).

There is also a lack of supportive evidence for the association between depression and childhood obesity (Wardle and Cooke, 2005). In the review, 28 articles – including 6 prospective studies – were summarized. With regard to the results, none of the prospective studies show a significant effect (Wardle and Cooke, 2005).

The review from Wardle and Cooke (2005) points out that body dissatisfaction may appear in obese children and adolescents, whereas problems regarding self-esteem and depression seem to be rarely observed. This could be due to confounding factors influencing the associations. For example, with regard to the impact of depression, existing studies explore the association between obesity and depressive symptoms or depression scores in general. However, obesity was found to contribute only to mood disorders or minor depression, which may affect some functional areas such as eating habits or learning ability. The idea of obesity affecting only specific conditions was proposed by Friedman and Brownell (1995) nearly 2 decades ago; as of yet, the gap remains unexplored, representing an area of potential further research.

In sum, childhood obesity leads to undesirable physical and social consequences. In terms of physical health, obesity in children increases the risks of many health conditions such as cardiovascular diseases, asthma, and obesity persistence in adulthood. The risk of obesity persistence is particularly important, since adult obesity leads to a number of chronic diseases, such as diabetes, hypertension, and certain types of cancer. Increased mortality risk is another consequence of obesity-related conditions. Concerning the social impact of obesity, weight bias is a potential cause of bullying, with obese children acting as both perpetrator and victim; parents may also be involved in bullying in either role. Unlike with physical and social consequences, however, little substantial evidence has been found regarding the impact of obesity on psychological problems in childhood, even now.

Overall, this section highlights the major consequences of child obesity, the severity of which will hopefully encourage public and private stakeholders alike to find solutions to the issue. The next part of the chapter explores various approaches to tackling childhood obesity.

1.5 Obesity: what can we do?

> *"The only way to keep your health is to eat what you don't want, drink what you don't like, and do what you'd rather not."*
>
> Mark Twain (1835 – 1910)

A number of measures have been devised to tackle the childhood obesity problem. Some of these measures aim at changing obesity-inducing environments, e.g. public policy interventions; others focus on changes at the individual level, e.g. curative and preventive interventions of childhood obesity.

1.5.1 Public policy to tackle childhood obesity

Public policy is one of the key measures that can help reduce childhood obesity by shaping the environment on a macro level. The goal of these public policies is to make the societal environment less conducive to the development of childhood obesity. Public policies are implemented with the expectation of encouraging or discouraging certain types of behaviour and include, for example, legislation and financial tools, taxation, and subsidies. Some of the public policies aimed at tackling childhood obesity include sugar-sweetened beverage (SSB) taxation and food price subsidies.

1.5.1.1 SSB taxation

> *"Sugar, rum, and tobacco, are commodities which are nowhere necessaries of life, which are become objects of almost universal consumption, and which are, therefore, extremely proper subjects of taxation."*
>
> Adam Smith, The Wealth of Nations, 1776

The principle of taxing SSBs to tackle obesity stems from a link between SSB consumption and weight gain, as shown in 2 systematic reviews. Vartanian et al. (2007) conducted a systematic review and meta-analysis exploring the association between soft drink consumption and health and nutrition. They showed that the effect size, or correlation, of SSB consumption on the average body weight of children, derived from 13 cross-sectional and 9 longitudinal studies, is 0.03, which is considered very small (Vartanian et al., 2007). They also found a significantly lower effect size reported in studies funded by the food industry compared with non-industry-funded ones (Vartanian et al., 2007). Another systematic review, funded by the food industry, explored the effect of SSB consumption on BMI change in children and adolescents below age 19 (Forshee et al., 2008). After reviewing 12 longitudinal and randomized controlled trial studies, the authors reported a very modest effect of SSB consumption on children's BMIs, with an estimated correlation of 0.004-0.017 (Forshee et al., 2008). This small effect of SSB consumption on children's weights was echoed by another recently published systematic review (Powell and Chaloupka, 2009).

SSB taxation, especially on soft drinks, was first proposed as an instrument of obesity reduction by Brownell et al. (2009). The authors point out that, in the US, soft drink taxation is based on sales tax, which is levied on goods as a proportion of price. Brownell et al. argue that this sales tax encourages consumers to buy products in large quantities at reduced prices, which results in excessive consumption of SSBs (Brownell et al., 2009). Instead of sales tax, they suggest a 1-penny-per-ounce excise tax, which levies tax on SSBs based on the weight or volume of the products (Brownell and Frieden, 2009). The excise tax can be levied directly on the producers and wholesalers, leading to increased SSB prices when the tax burden is transferred from producers to consumers. In contrast, sales tax is levied at retailers, which make it difficult to detect if tax is levied on the products properly (Brownell et al., 2009). Once the price rises, the consumption of SSB would be replaced by healthier beverages, which brings about a reduction of calorie intake and weight loss (Brownell and Frieden, 2009). The price policy would also be particularly effective in the poor, who are affected to a greater extent by price change. In addition, a penny-per-ounce tax is a revenue-generating policy that will increase tax income to the government (Brownell and Frieden, 2009).

The arguments above cast into doubt the ability of SSB taxation to effectively tackle childhood obesity, due to a few reasons. First and foremost, 3 systematic reviews, regardless of the funding body issue, clearly indicate that any effect of SSB consumption on weight/BMI change in childhood is minimal. This conclusion is echoed by another recent study (Fletcher et al., 2010). Thus, the assumption that a price increase will reduce the prevalence of obesity in childhood is not scientifically evident. Secondly, the claim that

SSB consumption would be replaced by other healthier drinks lacks supportive evidence. The amount of calorie intake may not differ much if children and adolescents drink other cheaper SSB brands instead of soft drinks after tax increases. This concern was also raised by Brownell and Frieden (2009). Third, for the purpose of tackling childhood obesity, this policy fails to discriminate its impact between children who will and will not benefit from it. In the UK, the prevalence of childhood obesity, as mentioned in the introduction section, is such that a majority of children are still not obese even in the no-SSB-taxation scenario. Therefore, while SSB taxation would only reduce or prevent childhood obesity in the obese and those at high risk of becoming obese, rather than all children, the policy affects the entire population. The price effect from the policy is theoretically effective among children and adolescents with low socio-economic status in particular, whereas the childhood obesity problem actually exists in all social classes. However, as mentioned in section 1.3.2, a recent study shows that the soft drink taxes significantly reduce soft drink consumption and weight gain only in middle-income households, while the effect of taxation on consumption and weight is very small in the highest 25% and lowest-income households (Finkelstein et al., 2010). Figure 1.9 illustrates this problem. Considering only the effect on childhood obesity, this policy not only unnecessarily affects no-risk children from middle socio-economic backgrounds, but also treats the obesity problem in children from lower- and higher-income families too lightly.

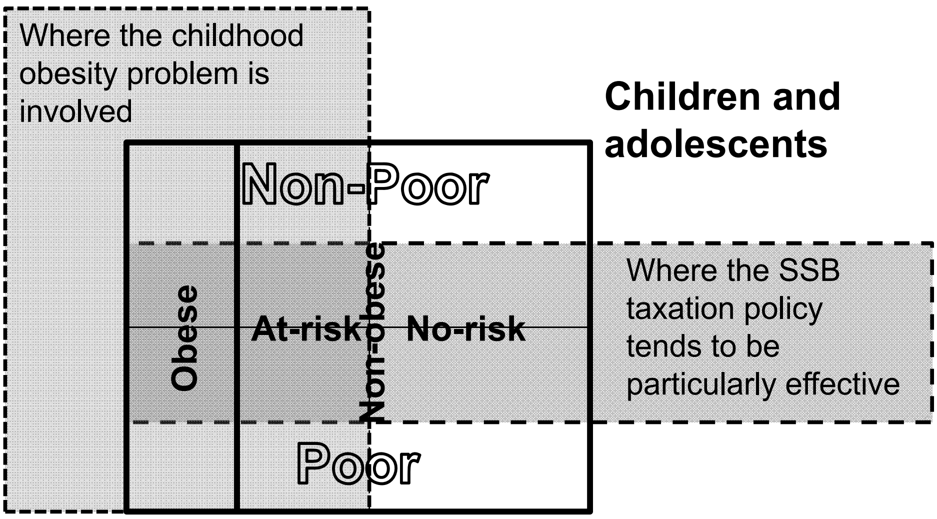


Figure 1.9: The childhood obesity problem and the effect of SSB taxation policy

1.5.1.2 Fiscal food pricing policy: Fruit and vegetable subsidy

As discussed in the 'Economics of obesity' section, some evidence supports the association between a rise in the price of fruits and vegetables and a decrease in consumption, weight gain, and increased BMI (Sturm and Datar, 2005, Sturm and Datar, 2008, Powell and Chaloupka, 2009, Powell and Bao, 2009, Sturm and Datar, 2011). Powell and Bao (2009) suggest that a fiscal subsidy to fruit and vegetable prices would reduce BMIs of children by 2.5-2.8%, leading to a decrease in the prevalence of childhood obesity as a consequence.

It is partly convincing that a fruit and vegetable price subsidy policy might contribute to weight loss and BMI reduction. Evidence clearly demonstrates a rise in the relative prices of fruits and vegetables over time, while the price elasticity of demand explains the reduced consumption. Studies also indicate that a decrease in fruits and vegetables consumption leads to an increase in weight and BMI in children.

However, considering the effect of the policy on childhood obesity only, subsidizing fruit and vegetable prices in general is deemed inefficient. This policy fails to discriminate between healthy weight children and those in the targeted group, which are at-risk and overweight/obese children. A considerable amount of financial resources would be required in order to subsidize fruit and vegetable prices in the market to an extent that results in clinical effectiveness.

One way of improving the idea to subsidize fruit and vegetable prices is to transfer financial incentives directly to the targeted populations. This means that, instead of spending the fiscal budget to lower prices in the market, a financial incentive should be directed to obese children or children with high risks of developing obesity in childhood. The financial aid can be in the form of either discounted coupon for fruits and vegetables, or cash. Studies have revealed the effectiveness of the direct payment method, i.e. conditional cash transfer, in improving health in the targeted populations, especially in low- and middle-income countries (Lagarde et al., 2007, Lagarde et al., 2009). This might be one way that helps reduce the prevalence of childhood obesity.

The assessment of the SSB taxation and fruits and vegetables subsidy proposals aptly reflects an old saying – "There's no such thing as a free lunch" – in the context of public health policy, as the benefits gained by one group of individuals may be at the cost of another. While SSB taxation might be effective in obese and at risk children of particular SES, the price effect inevitably affects the entire population at the same time. Similarly, a policy to subsidize fruits and vegetables would use a considerable amount of public revenues to decrease prices, with only small weight reductions expected. These collateral effects are sometimes neglected or dismissed, to the detriment of society as a whole.

Other examples of public policies related directly or indirectly to obesity in childhood and adulthood are those targeting smoking cessation and junk-food or fat taxation.

In investing in youth tobacco control, a number of strategies aimed at smoking cessation have been proposed (Lantz et al., 2000). However, one side effect of smoking cessation is weight gain, as shown in a systematic review (Klesges et al., 1989). Klesges et al. (1989) reviewed 70 cross-sectional and longitudinal studies, with a total sample size of over 350,000 subjects, evaluating the effects of smoking on weight. They reported that 55 studies (79%) found that smokers who stop smoking gain more weight than non-smokers (Klesges et al., 1989). Health benefits from smoking cessation are thus reduced by the risk of obesity.

Another proposed policy to tackle obesity in both children and adults is junk-food or fat taxation, which is aimed at reducing junk food consumption through the price mechanism (Leicester and Windmeijer, 2004). However, results from studies examining the effect of the fat tax do not support this argument (Chouinard et al., 2007, Yaniv et al., 2009). Chouinard et al. (2007) report that a fat tax leads to only a tiny change in fat consumption, which is indicative of its price inelasticity. In addition to the minimal effect on fat consumption, Yaniv et al. (2009) also show that a fat tax may increase obesity. This is probably because the increase in the junk foods prices may impact not only the consumption of fat and/or junk foods, but also the time spent on exercise, as time for physical activity may be reduced for the sake of shopping and preparing foods instead (Yaniv et al., 2009).

In light of these findings, the author emphasizes the importance of considering the collateral effects of policy proposals alongside their benefits, as it is likely that few benefits come without a cost.

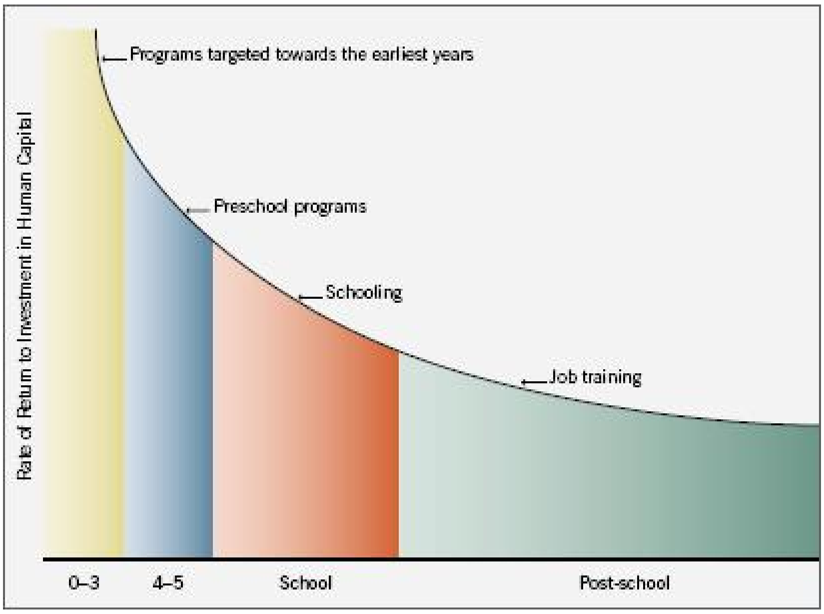
1.5.2 Human investment: treatment and prevention of obesity in childhood

The health investment for obesity is based on the concept of health capital and the demand for health from Grossman's model (Grossman, 1972). The model suggests that health can be accounted in terms of a stock of health, or health capital, which produces healthy time. Investment in health, by avoiding unhealthy behaviours or promoting healthy behaviours, results in health improvement and increased healthy time (Grossman, 1972). Inputs for gross investment include such factors as medical care, diet, exercise, and recreation (Grossman, 1972).

Even though investment in health affects health capital to some extent, the change in health stock does not directly reflect the level of investment. This is as Grossman states "… even though health capital falls over the life cycle, gross investment might increase, remain constant or decrease" (Grossman, 1972). Moreover, the model also shows that health depreciates with age, despite health expenditures increasing as individuals age (Grossman, 1972).

With regard to Grossman's model, the implications of obesity-related health investment can be summarized into 2 points. Firstly, because health depreciates with age, it can be assumed that adverse health effects from obesity, which is one of the factors leading to health deterioration in the elderly, will worsen when age increases. This idea also relates to obesity-related health expenditure, which includes the costs of treating obesity-related complications, productivity loss, and time lost by informal care-takers. As age increases health expenditure, obesity-related health expenditure will increase if obesity is left untreated. In other words, without regard to time preference and discounting, as discussed in Chapter 2, the model suggests that it is more efficient to treat and prevent obesity sooner rather than later because the health costs increase as time passes. Therefore, when considering the model, it is suggested that obesity should be detected and treated promptly or as soon as possible after confirming diagnosis. Early treatment and prevention of obesity will maximize health condition and reduce obesity-related health expenditure.

The concept of early investment on health is also supported by another modern economic theory, the Heckman equation (Heckman, 2010). Heckman (2010) proposes that investing in early childhood education would give the highest rate of return on investment. Early childhood education, supported by investment in parents and family environment, would lead to sustainable development to adulthood. This can be extrapolated to investment in the promotion of healthy lifestyles in high-risk children, which would result in social cost reduction in the long run (Heckman, 2010). Heckman also asserts that investing in early childhood education is a cost-effective strategy, especially during the critical period of 0-5 years of age (Heckman, 2010). Figure 1.10 illustrates the rate of return to investment by age.



Source: Heckman (2010)

Figure 1.10: Rate of return to human capital investment by age

Secondly, from the model, treatment and prevention of childhood obesity are the forms of health investment best able to improve health. Investing in treatment means to lessen the degree of obesity in obese children using appropriate tools, such as education, exercise, medication, surgery, or even direct financial incentives. Prevention of childhood obesity, meanwhile, applies to both obese and non-obese children, with the aim of preventing them from moving on to a higher degree of obesity. Prevention can be accomplished through lifestyle modification, avoiding unhealthy behaviours, and promoting healthy behaviours. Obesity-promoting factors include psychiatric problems and TV watching, whereas obesity-demoting factors consist of physical activity and exercise.

From Grossman's model and the Heckman equation, it can be concluded that investment in treating and preventing childhood obesity should be done as early as possible in order to achieve better health. However, health investment in obesity (in the form of behavioural modification) does not always reduce social costs and total health expenditure, as shown in the Heckman equation, even though it promotes health. This is due to the fact that the costs incurred from health investment may outweigh the medical costs of obesity. Similarly, the costs of interventions to prevent and treat childhood obesity could also be higher than total obesity-related health expenditures without the interventions. Figure 1.11 illustrates the association between health investment and obesity-related medical and non-medical costs with total health expenditure. As seen, the total obesity-related health expenditure stems from the sum of the costs of health investment and the incurred medical costs of obesity. The investment would be considered cost-saving if the costs of childhood obesity treatment and prevention are lower than the obesity-related medical and non-medical costs saved due to investment. Alternatively, investment in obesity interventions can still be deemed efficient, despite not being cost-saving, if the financial benefits from the investment are less than its costs when converting health benefits into monetary units.

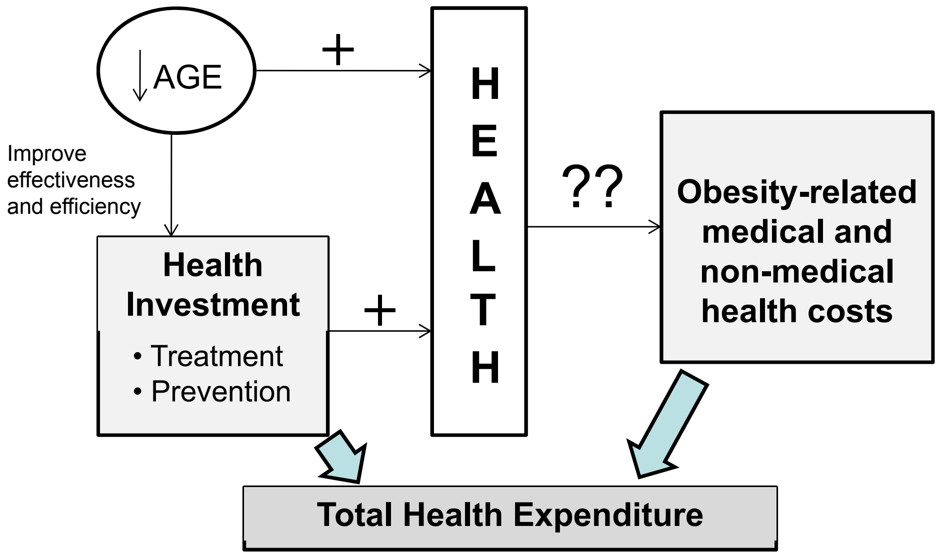


Figure 1.11: Factors influencing total health expenditure of childhood obesity from the Grossman's model

Despite health improvement being achieved by investment, it is uncertain whether the total health expenditure would be reduced with respect to the investment. In other words, the efficiency of the investment is the main factor of concern. The next issue of consideration is whether or not health investment is worthwhile in particular cases, especially when comparing with a do-nothing scenario or standard approaches. In order to answer this question, a tool that can consider both costs and benefits of health investment for all alternatives at the same time is required.

1.6 Economic evaluation

Economic evaluation, or cost-effectiveness analysis, is a method suited to this purpose. Gold et al. (1996) defines cost-effectiveness analysis as "… a method designed to assess the comparative impacts of expenditures on different health interventions."

Drummond et al. (2005) describe 4 types of economic evaluation: 1) cost-minimization analysis – the costs of the interventions are compared by assuming there is no difference between health benefits gained between alternatives; 2) cost-benefit analysis – the outcomes of the alternatives are estimated in terms of monetary units and considered together with the costs, while the net benefits of the alternatives are compared; 3) cost-effectiveness analysis – the costs of the interventions are calculated per natural unit of effectiveness. For example, program A reduces 3 units of BMI for children, with the total cost of £900 per child, whereas program B can reduce BMI by 5 units at a total cost of £2,000 per child. The result of the cost-effectiveness analysis is demonstrated in terms of cost per additional unit of benefit generated, or the incremental cost-effectiveness ratio (ICER). Thus, the ICER of program A is £900/3 = £300/one unit reduction of BMI, while the ICER of program B is £2,000/5 = £400/one unit reduction of BMI; and 4) cost-utility analysis – the ICER is used as a unit of comparison, while the benefits are measured in terms of additional quality-adjusted life years (QALYs) gained as a result of the intervention. Another alternative outcome unit that has been proposed is disability-adjusted life years (DALYs) (Drummond et al., 2005).

Health outcomes: QALYs versus DALYs

QALYs and DALYs have been widely used as health outcome units for cost-utility analysis during the past decades. Nonetheless, a number of differences with regard to their underlying concepts and assumptions results in slight differences in their utilization and interpretation. Understanding the concepts behind QALYs and DALYs helps researchers select the best unit when conducting cost-effectiveness analysis.

QALYs, developed in the 1960s by economists and psychologists (as cited in Gold et al., 2002), describes an individual's health gained from an intervention by accounting for both additional life years gained and health-related quality of life (HRQL) during that period. HRQL is multiplied by additional life years to arrive at the number of QALYs gained. While the conventional approach to evaluating health status is to measure it in separate dimensions or attributes of health, such as levels of activity, pain, or self-care, HRQL is estimated by considering all such attributes together. The method of estimation used, therefore, plays a crucial role in determining HRQL.

To measure HRQL, a number of instruments have been utilized: the Health Utility Index (HUI), the Quality of Well-Being Scale (QWB), and the EQ-5D. QALYs define health states ranging from 0 to 1, where 0 represents death and 1 represents full health. It is noted that health states lower than 0 are also possible if death is not defined as the worst health state (Drummond et al., 2005). The HRQL of the worse-than-death condition then can be assigned a negative value, for example, a vegetative stage in the brain-dead patients. The techniques commonly used to generate health score are standard gamble, time trade-off, and visual analogue scales. The HUI employs the time trade-off and multi-attribute utility scale method to assess health, whereas QWB and EQ-5D use multi-attribute utility scale and visual analogue scale (Gold et al., 2002). The use of different instruments in estimating HRQL unavoidably affects the resulting QALYs values (Gold et al., 2002); despite these variations, however (Kopec and Willison, 2003), the correlations among outcomes from the measures are still within a reasonable range (Gold et al., 2002).

DALYs, conversely, were invented by the collective efforts of the World Bank and World Health Organization in 1993 for the purpose of measuring burden of diseases and disability in populations, especially in developing countries (Murray et al., 1996, Gold et al., 2002). Unlike QALYs, which use HRQL to refer to the individual's overall health state, DALYs attach HRQL to particular diseases in order to improve utilization in real practice (Murray et al., 1996, Gold et al., 2002). Moreover, the values of health were generated by secondary data and expert opinions due to concerns regarding self-assessments of health. Health values are weighted by age and gender in the calculation. The DALYs scale ranges from 1 to 0, where 1 represents death and 0 represents full health. Health is valued by a person trade-off method, or the trade-off between people with particular diseases and healthy people. For example, experts may agree to trade 100 patients with end-stage renal disease with 20 full-health people. The discount rate for future adjustment of DALYs is generally 3% (Murray et al., 1996, Gold et al., 2002).

Though DALYs appear to be a pragmatic tool with which to compare health states among individuals, some fundamental concerns have been raised. Firstly, the estimation of DALYs is based on expert opinions and secondary data, not on the actual perceived health status of individuals like in QALYs. Despite concerns regarding biases in self-assessment of health, QALYs still reflect the actual health state of particular populations of interest, whereas secondary information is used to form DALYs. This may lead to some degree of uncertainty. Furthermore, in contrast with QALYs, DALYs cannot assess co-morbid conditions within the same individual, as health is already defined in relation to specific diseases and disability, not individual health state (Gold et al., 2002). This is particularly important in the elderly, who are often simultaneously affected by multiple underlying diseases, such as hypertension and diabetes. DALYs calculation also fails to account for side effects caused by treatments, such as drug allergy caused by the use of antibiotics to treat infectious diseases.

1.7 Conclusions

Increasing evidence has singled out childhood obesity as a health problem globally. BMI is currently a standard measure for obesity, even though there are some limitations, particularly on assessing body fat distribution. Regarding the definition of childhood obesity, both population-based and internationally standardized definitions have been used, even though there is as yet no unanimous agreement.

The concept of energy balance, or the balance between calorie intake and energy expenditure, has long been used to explain the development of obesity. However, during the past half century, a number of additional factors have been found to be related to obesity, contributing directly or indirectly toward the energy balance process. As discussed above, obesity-related factors can be classified into the following categories: 1) genetic factors; 2) psychiatric factors – mental health conditions and medications; 3) socio-economic-demographic factors – age, gender, ethnicity, socio-economic status; and 4) behavioral factors – sedentary behaviour, e.g., TV watching, and physical activity.

While obesity has been found to be somewhat heritable, the heritability of obesity requires not only specific genes, but also an obesity-prone environment, both of which combine to encourage the development of excessive body fat. Mental health factors, including such mood disorders as depression and bipolar disorder, also promote weight gain. Moreover, obesity is often an unfortunate side effect of psychiatric medication used to treat psychiatric disorders. Concerning socio-economic-demographic factors, characteristics that increase the risk of obesity include aging, being female, and low socio-economic status. Finally, evidence has shown that behaviours leading to obesity include sedentary behaviour such as TV watching, whereas physical activity and exercise induce weight loss.

Research has demonstrated obesity's multidimensional adverse consequences in terms of not only physical health, but also undesirable social impact. In terms of physical health, obese children and adolescents are affected by increased risks of cardiovascular diseases, asthma, and obesity persistence in adulthood. With regard to social impact, research has revealed that weight bias is associated with bullying and being bullied in obese children.

To respond to the problem of childhood obesity, interventions to treat and prevent obesity have been used. The human capital model from Grossman (1972) is applied in order to explain the benefits of obesity reduction. From the model, health is taken as a stock that can be invested in. Investment in health should be performed as early as possible to maximize benefits gained. This concept is supported by Heckman (2010). The chapter points out the measures that have been used to tackle childhood obesity, on both macro and individual levels. The examples of SSB taxation and food price subsidy policies are raised for discussion in the chapter.

Although the human capital model suggests that the prevention and treatment of childhood obesity improves health, it is uncertain whether the future benefits gained from interventions are worth the effort. This concern is especially important if the maintenance of intervention effectiveness is poor, as most adverse consequences of obesity occur in adulthood. While the costs of reducing obesity in childhood produce a certain amount of future health benefits in adulthood, it may be more efficient to invest comparatively fewer resources in treatment of adulthood obesity that produces better outcomes. Assessment of the costs and outcomes of the interventions is deemed necessary in this process. The chapter thus introduces the concept of economic evaluation, particularly cost-effectiveness analysis, which is a tool that can compare costs and outcomes among several interventions at the same time.

Despite its importance, only a handful of cost-effectiveness studies of child obesity interventions are available. Furthermore, no literature review of this matter exists. The lack of synthesis of evidence creates difficulty in improving the body of knowledge concerning the cost-effectiveness of child obesity interventions, particularly with regard to the health outcome estimation.

The purposes of this thesis are therefore as follows: 1) to explore the literature concerning the economic evaluation of child obesity interventions; 2) to demonstrate and improve an economic modelling technique used in the long-term health estimation of child obesity interventions in cost-effectiveness analysis; and 3) to point out the gaps, weaknesses and future research that are helpful for the improvement of economic evaluation. Relevant research questions, as well as exploration of the answers, are presented in sequence as follows:

Question 1 – What is the current evidence baseline of the long-term cost-effectiveness of child obesity interventions?

In response to this question, ***"Chapter 2: Cost-effectiveness of interventions for treating or preventing obesity in children and adolescents: a literature review"*** conducts a literature review on the cost-effectiveness of interventions to treat or prevent obesity in children and adolescents. Regarding this, the chapter shows the costs of illness resulting from obesity and highlights the importance of economic evaluation in the assessment of obesity interventions. The literature review focuses on cost-effectiveness analyses of child obesity interventions that represent the long-term health outcome in terms of life years, either QALYs or DALYs. A systematic reviewing approach, including keywords and inclusion criteria, is used. The author identifies selected articles and perform quality assessments using Drummond's critical assessment guidelines of economic evaluation (Drummond et al., 2005). All selected articles are critically investigated and compared in detail. In-depth discussion on the economic evaluation techniques used in the selected papers then takes place, including economic modelling techniques, underlying assumptions made, using QALYs versus DALYs, discount rate, medical cost offset, and other subjects.

Question 2 – The introduction of MEND 7-13 as one example of an intervention to tackle obesity in children: what factors affect programme effectiveness among participants?

To explore the economic evaluation of child obesity interventions more critically, an intervention aimed at reducing obesity in children, the MEND 7-13 programme, is introduced as a case study. ***"Chapter 3: The MEND 7-13 programme"*** begins with an introduction of MEND 7-13, including the programme's background and development, delivery, justification, and effectiveness. Afterward, questions are raised concerning health equity issues in the MEND rollout phase, as well as intrinsic/extrinsic factors influencing the benefit uptake of the programme. These factors include gender, age, programme attendance, and other socio-demographic attributes. To explore the associations among these factors, the chapter begins by exploring the data used in the analysis in terms of descriptive statistics, the data collection process, and the quality of the data. Comparison of the baseline characteristics of the MEND data with those in the 2008 Health Survey for England data follows, with respect to gender, age, ethnicity, and the level of deprivation. The chapter then discusses the issue concerning missing data, as well as the variable recoding process. This is followed by a demonstration of the statistics used in the analysis, simple and logistic regression, including the lists of variables employed in the regression model. The results of the analysis are shown and interpreted, followed by implications and recommendations drawn.

Question 3 – Is MEND 7-13 cost-effective in the long term?

The entirety of ***"Chapter 4: Economic Evaluation of the National Child Weight Management Programme: MEND 7-13"*** is dedicated to answering this question. The chapter traces the incurred costs and health outcomes of MEND 7-13 from the health care payer's perspective. The results of the analysis are displayed in terms of costs per QALY gained and net benefits. With regard to this, an economic model used to estimate QALY gained is developed and used in the economic evaluation. Afterward, the chapter performs sensitivity analyses to assess the robustness of the economic evaluation. The Monte-Carlo simulation is employed to assess the impact of parameter uncertainty in the model. The chapter also performs a subgroup analysis, based on the inputs from Chapter 3, to elucidate the economic effects of health disparity. An assessment of the economic evaluation approach applied in this chapter is performed using Drummond's critical assessment guidelines of economic evaluation. Critical discussion follows in order to point out certain issues, such as the limitations of the study, gaps and weaknesses, applicability of results, and future research to improve economic evaluation.

After the examination of economic evaluation, key messages learned from the previous chapters are summarized in ***"Chapter 5: Conclusions."*** This chapter answers the research questions raised at the beginning using the evidence learned from the previous chapters, followed by a summary of Chapters 1-4. The chapter then illustrates what could have been done differently and justifies the choices made. Afterward, the chapter recaps the contributions to knowledge made by the thesis, including research and policy implications, pointing out and filling gaps in the existing studies, and recommendations for future research. The final conclusions, with take-away messages from the thesis, are then made.

Chapter 2: Cost-effectiveness of interventions for treating or preventing obesity in children and adolescents: a literature review

2.1 Introduction

As discussed in Chapter 1, the global prevalence of childhood obesity, which has increased significantly during the past decades (World Health Organization, 2000, Lobstein et al., 2004), leads to many undesirable health consequences. Strong evidence has illustrated the repercussions of overweight and obesity for physical and mental health in children (Reilly et al., 2003, Weiss et al., 2004, Han et al., 2010). Overweight and obesity in childhood are known to be risk factors for certain health conditions, including psychological problems, asthma, chronic inflammation, and cardiovascular risk factors (Reilly et al., 2003, Han et al., 2010). Moreover, childhood obesity extends its adverse impacts to adulthood, as it increases the chance of becoming obese as adults. A systematic review by Reilly et al. (2003) showed that obesity in children increases the risks of adult obesity, adult morbidity and mortality, and cardiovascular risk factors. In addition, obesity in childhood also leads to poor socio-economic status, in terms of both income and educational attainment, later on in life (Reilly et al., 2003).

2.1.1 Costs of illness

Obesity not only incurs health problems, but also places a burden on health expenditures in both childhood and adulthood. The rising prevalence of obesity has increased obesity-related co-morbidities in adulthood, which often require costly long-term treatments (Wyatt et al., 2006). Wolf and Colditz (1998) showed that direct medical treatments account for more than half of obesity-related health expenditures, with diabetes consuming the largest share of spending (Wolf and Colditz, 1998, Visscher and Seidell, 2001). Thorpe et al (2005), meanwhile, commented that obesity-related diseases have raised health expenditures not only because of the surging prevalence of obesity, but also due to the advancement of medical technologies used to treat related conditions. Moreover, productivity loss due to obesity-related morbidity and mortality creates additional economic costs (Wyatt et al., 2006).

Concerning the cost of illness in childhood, John et al. (2010) reviewed 7 studies, published in 2007-8, demonstrating the economic burden of childhood obesity. Six of seven studies are based on the US population. Despite some inconsistency of results across studies, the review pointed out that childhood obesity incurs a considerable amount of health expenditures annually (John et al., 2010).

The economic burden of obesity appears to be more significant in adults than in youths. Withrow and Alter (2010) conducted a systematic review to show the economic burden of obesity worldwide. The review included 32 studies, 53% of which were US-based. They showed that approximately 0.7 to 2.8% of total healthcare expenditures in 12 countries is attributable to obesity (Withrow and Alter, 2010). These figures vary across countries. For example, Thompson and Wolf (2001) point out that obesity contributes to 5.5 to 7% of medical expenditures in the US, whereas in other countries, including Australia, Canada, France, New Zealand, and Portugal, obesity accounts for only 2 to 3.5 % of medical expenditures. Several other US studies have reported the total annual medical expenditure of obesity treatment to be around 5 to 7% (Finkelstein et al., 2003, Finkelstein et al., 2004, Wolf and Colditz, 1994, Wolf and Colditz, 1998).

Concerning the financial impact of obesity at the individual level, Withrow and Alter (2010) reviewed 32 studies estimating the medical costs of obesity in adults, 20 of which were modelling-based, while the estimations from the other 12 were database-based. Only 9 out of the 32 studies were based on longitudinal data. The results showed that obese people had approximately 30% (ranging from 6 to 45%) higher medical costs than do their normal weight counterparts.

Thompson et al. (1999) constructed a model to calculate the lifetime medical care costs of the obese compared with those of normal weight people. They reported that the total discounted lifetime medical care costs for 5 major obesity-related diseases – hypertension, hypercholesterolemia, type 2 diabetes, coronary heart disease, and stroke – were approximately 10,000 USD higher for obese people than for people of normal weight (Thompson et al., 1999).

It is noted that most studies estimated the medical costs of obesity by the increased risk of obesity-related diseases, as the direct medical costs of obesity are relatively small compared with the incurred costs of its consequences. The obesity-related medical costs shown in the studies thus largely depend on the diseases accounted for in the estimates. This may explain the wide range of the medical expenditure estimates across studies. More details of this are described in the discussion section of this chapter.

In the UK, the Comptroller and Auditor General (2001) estimated the total direct and indirect costs of obesity in England to be £2.63 billion, or 0.3% of the UK Gross Domestic Product, in 1998. Of these expenses, approximately £0.5 billion was attributed to direct medical costs, whereas indirect medical costs and productivity loss accounted for over £2.1 billion (The Comptroller and Auditor General, 2001). The total economic loss due to obesity in England rose to £3.34-3.72 billion in 2002, £0.99-1.12 million of which stemmed from the direct medical costs of obesity and related complications (House of Commons Health Committee, 2004). The Foresight programme projected the future NHS costs of obesity to be £3.9 billion, £5.3 billion and £7.1 billion in 2015, 2025, and 2050 respectively (McPherson et al., 2007).

2.1.2 Effectiveness of child obesity interventions

To reduce the negative consequences of childhood overweight and obesity, a number of interventions have been implemented to treat or prevent childhood obesity. Systematic reviews were then performed to determine the effectiveness of such interventions (Summerbell et al., 2005, Oude Luttikhuis et al., 2009).

Oude Luttikhuis et al. (2009) conducted a systematic review and meta-analysis of the effectiveness of interventions for treating obesity in children, which included 64 randomised controlled trials (RCTs) using medication, activity-based, behavioural, and dietary interventions. Of these, 37 studies were conducted in children below the age of 12, whereas the other studies focused on children ages 12 and older. The results showed that dietary interventions produced small but significant effectiveness at all ages, whereas activity-based interventions are ineffective in reducing weight at all groups. Regarding behavioural interventions, the meta-analyses showed that interventions are effective at 6 months follow-up for all ages, but only children ages 12 and above benefit from the weight reduction at 12 months follow-up. Drug interventions showed modest effects in reducing weight in children ages 12 and over. The authors concluded that lifestyle interventions in children and adolescents are associated with the reduction of weight with or without the use of medication (Oude Luttikhuis et al., 2009).

Another systematic review evaluating the effectiveness of interventions to prevent childhood obesity was conducted by Summerbell et al. (2005). Twenty-two studies were included with various durations of follow-up: 10 long-term (≥1 year) and 12 short-term (<1 year). Only 2 of 10 long-term studies showed a significant—though small—effect on obesity prevention. One study, using dietary education and physical activity intervention, showed success in preventing obesity only in girls, while the other, using a multimedia-based physical activity intervention, succeeded in preventing obesity in both. For the short-term studies, the effectiveness of the interventions was consistently low. Only 2 of 12, both using only physical activity interventions, were found to be slightly effective.

2.1.3 Cost-effectiveness of child obesity interventions

Although some evidence supports the effectiveness of interventions for the treatment and prevention of obesity in children and adolescents, it is still uncertain from a public health perspective whether or not these interventions should be promoted and publicly funded on a large scale. To answer this question, policymakers need to consider not only the effectiveness of the proposed interventions, but also their costs.

As stated in Chapter 1, cost-effectiveness analysis (CEA) is an economic evaluation method used in resource allocation that compares the costs and benefits of competing alternatives. Gold et al. (1996) define cost-effectiveness analysis as "... the method designed to assess the comparative impacts of expenditures on different health interventions" (Gold et al., 1996). A number of studies showing the cost-effectiveness of child obesity interventions have been published (Such as in John et al., 2010, Cawley, 2007). However, only a handful of those consider a more generic measure of long-term outcomes, for example using quality- or disability-adjusted life years (QALY or DALY), which would facilitate cost-utility analysis (Gold et al., 1996).

Although a few reviews have illustrated the cost-effectiveness of child obesity interventions, they provide only a limited amount of information (John et al., 2010, Cawley, 2010). Moreover, the reviews failed to consider issues related to the characteristics of the economic models and interventions, such as modelling assumptions, use of QALYs versus DALYs, and the appropriate discount rate, as well as discussion of cost-effectiveness analysis. There has yet to be a comprehensive literature review examining the long term cost-utility of childhood obesity interventions and appraising critically the estimation of costs and outcomes.

The purpose of this study is to assess the long term cost-effectiveness and cost-utility of the interventions that aim to treat or prevent obesity in children or adolescents, and to critically appraise the methods used in the cost-effectiveness analyses.

2.2 Methods

2.2.1 Search strategy

The search strategy was modified from keywords found in the existing reviews (Summerbell et al., 2005, Oude Luttikhuis et al., 2009). Four groups of keywords were identified: children, obesity, type of the intervention, and economic evaluation. The keywords used were gathered from the previous reviews to maximize the search coverage. MEDLINE (1950 To March 2010), EMBASE (1980 To March 2010), PsycINFO (1806 To March 2010) and Econlit (1969 To March 2010) were searched.

2.2.2 Inclusion criteria

The study included cost-effectiveness and cost-utility analyses and did not place restrictions based on perspective (e.g. societal, public health, customer, etc.). Selected interventions must have the aim of treating or preventing obesity in children. Interventions focusing solely on treating or preventing particular obesity-related diseases, such as diabetes, were excluded. Inclusion criteria placed no limitations on the type of intervention or study design determining the effectiveness data (i.e. the effectiveness of the interventions can be from RCTs, non-randomised trials, cohort studies, epidemiological studies, or other study designs).

No restrictions were applied to the comparator, but the literature must clearly identify in the analysis what the interventions were compared with. Only interventions that targeted populations under the age of 20 (at the start of the intervention) were included. No constraints were applied for the follow-up duration.

Studies must show the outcomes in terms of incremental costs of interventions divided by the incremental outcome units, which can be represented in terms of the incremental cost-effectiveness ratio (ICER), where the comparator can be either another intervention or no intervention; or dominance, if the intervention is cost-saving. Outcome units must be expressed in terms of QALYs gained, DALYs saved, or life years gained. Searches were restricted to articles written in English.

2.2.3 Data extraction and management

The author is the only person involved in the data extraction process. Data was extracted for the review using set standards, as seen in Tables 2.2-2.4. The QUOROM flowchart of this review is provided in Figure 2.1.

A total number of 1,706 articles were identified following a keyword search. Of these, 252 duplicates were removed, leaving 1,454 articles. After screening titles and abstracts, 22 articles were found to be relevant, and full texts were retrieved for a critical review. Seventeen articles were excluded after reading full texts due to criteria mismatch. Five articles matched the inclusion criteria and were included in the review.

The search was repeated to expand the coverage by removing the keyword related to ‘intervention’. The search was performed in the same 4 databases from the inception to June 2010. After paper selection, the same 5 articles were identified. The finding was also confirmed by 2 review articles published in 2010 with similar results (Cawley, 2010, John et al., 2010).

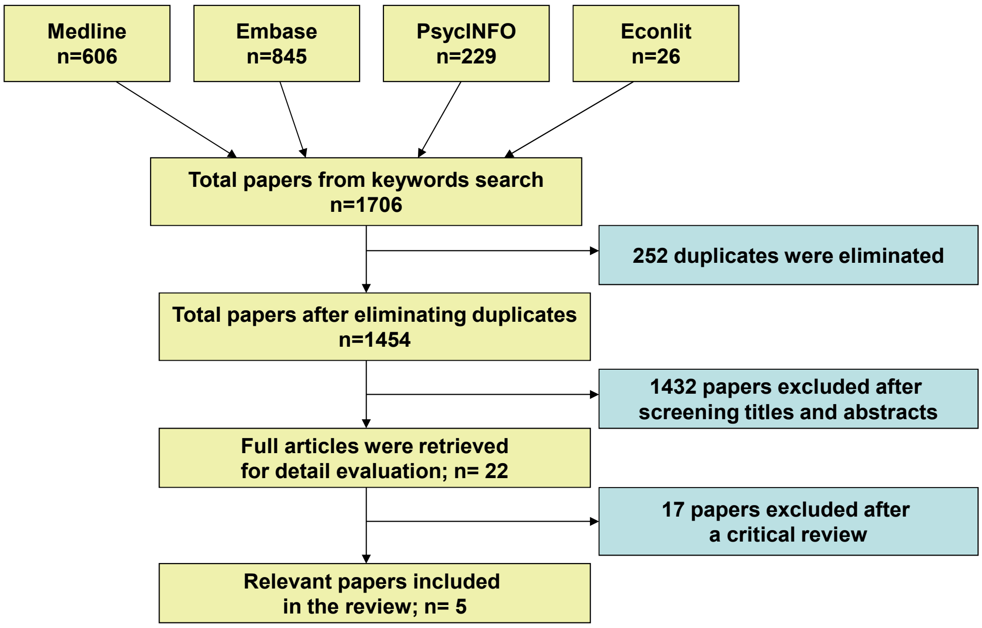


Figure 2.1: QUOROM flowchart

2.2.4 Updating relevant published literature

In order to update this chapter with recently published studies on the cost-effectiveness of child obesity intervention, another search was performed after the systematic search was done. Two relevant cost-effectiveness studies were found in this regard, both of which were available in Medline in July and August 2010, respectively.

2.2.5 Quality assessment of included studies

The assessment criteria for economic evaluations from Drummond et al. (2005) were used in the quality assessment of reviewed articles. The results are shown in Table 2.1, with more details in Appendix 3.

Although the Drummond checklist is one of the standard tools used in quality assessment for cost-effectiveness studies, the results from the checklist do not reflect the quality of the study very well. This is mainly due to the high complexity and diversity of the economic modelling applied to the studies. It is thus difficult to use a simple checklist to evaluate the study quality. Information gleaned from the Drummond checklist indicates how comprehensively a study is conducted, rather than its actual quality. An example can be seen in Table 2.1 on Questions 3 and 6 of the checklist. As demonstrated, all studies established the programme effectiveness well (from Question 3), but only one study valued the outcomes credibly (from Question 6). It is interesting to explore why the outcomes were not credible, given that the effectiveness was well-established. The main reasons are due to the assumptions made in the economic models being unreliable (Details are provided in Appendix 3). The Drummond checklist, however, fails to demonstrate this through the 10 questions.

Due to substantial variation across studies, e.g. types of intervention, outcome measurement, cost estimation, and targeted populations, it is not feasible to conduct a meta-analysis using these data.

Table 2.1: Assessing the reviewed articles for economic evaluation

Check-list for assessing economic evaluation	**Wang et al. (2003)**			**Brown et al. (2007)**			**Moodie et al. (2008)**			**Magnus et al. (2009)**			**Moodie et al. (2009)**			**Ananthapavan et al. (2010)**			**Moodie et al. (2010)**		
	Yes	No	Can't tell	Yes	No	Can't tell	Yes	No	Can't tell	Yes	No	Can't tell	Yes	No	Can't tell	Yes	No	Can't tell	Yes	No	Can't tell
1. was a well defined question posed in an answerable form?	A			A			A			A			A			A			A		
2. was a comprehensive description of the competing alternatives given (that is, can you tell who did what, to whom, where, and how often)?			A	A			A			A			A			A			A		
3. was there evidence that the programme's effectiveness had been established?	A			A			A			A			A			A			A		
4. were all the important and relevant costs and consequences for each alternative identified?			A			A	A			O	C		O		C	A				O	C
5. were costs and consequences measured accurately in appropriate physical units?	A			A			O		C	O		C	O	C		A			O	C	
6. were costs and consequences valued credibly?	C	O		C	O		O	C			A			A			A			A	
7. were costs and consequences adjusted for differential timing?	A			A			A			A			A			A			A		
8. was an incremental analysis of costs and consequences of alternatives performed?	A			A			A			A			A			A			A		
9. was allowance made for uncertainty in the estimates of costs and consequences?	A			A			A			A			A			A			A		
10. did the presentation and discussion of study results include all issues of concern to users?			A			A	A			A			A			A			A		

C = for the cost side; O = for the outcome side; A = all (both cost and outcome sides)

2.3 Results

2.3.1 Characteristics of included studies

Of the reviewed studies, two were based on the US population (Wang et al., 2003, Brown et al., 2007), whereas the other 5 were based on the Australian population (Magnus et al., 2009, Moodie et al., 2009, Moodie et al., 2008a, Ananthapavan et al., 2010, Moodie et al., 2010). Two studies (Wang et al., 2003, Brown et al., 2007) used decision tree modelling to estimate future health outcomes, while the others applied a Markov model (Magnus et al., 2009, Moodie et al., 2009, Moodie et al., 2008a, Ananthapavan et al., 2010, Moodie et al., 2010). These five studies were undertaken as part of the Assessing Cost-Effectiveness in obesity (ACE-obesity) project, which evaluated interventions aimed at reducing child obesity in Australia. The same outcome estimation methods, to estimate DALYs, were applied to the project-related studies.

Six studies accounted for the costs of interventions from the societal perspective, which considers both the personal costs to patients—including the costs of patient time, informal care giving, and transportation—and direct medical costs. The Magnus et al. study, meanwhile, employed the public health perspective, focusing on the direct medical costs without considering costs to the patients (Magnus et al., 2009).

The interventions appearing in the reviewed papers varied considerably in terms of their characteristics, target population and coverage, duration of follow-up, and program effectiveness.

School-based interventions were found in 2 studies (Wang et al., 2003, Brown et al., 2007) examining the treatment of overweight and obesity through health education. Another study focused on a family-based, GP-mediated intervention designed to treat obesity through behavioural modification (Moodie et al., 2008a). A fourth study evaluated a public program that encouraged children to walk to schools (Moodie et al., 2009), while another assessed the effects of a hypothetical intervention banning TV advertisement of high-fat/high-sugar foods and beverages (Magnus et al., 2009). The cost-effectiveness of laparoscopic adjustable gastric banding (LAGB) in severely obese adolescents was evaluated in a study (Ananthapavan et al., 2010), whereas the last paper examined the Active After-school Communities program (AASC), which encourages children to increase their levels of physical activity through after-school sessions (Moodie et al., 2010).

The effectiveness of the interventions was derived from randomised controlled trials in only 3 of the 7 studies (Wang et al., 2003, Brown et al., 2007, Moodie et al., 2008a); the other four estimated the effects of the interventions using the modelling method (Magnus et al., 2009, Moodie et al., 2009, Ananthapavan et al., 2010, Moodie et al., 2010). One modelling method assumed that removing TV advertising of energy-dense, nutrition-poor food and beverages reduces product consumption, resulting in lowering energy intake, body weight, and BMI (Magnus et al., 2009). The second estimated the impact of time spent walking to school on energy consumption, which led to weight loss and reduction in BMI as a consequence (Moodie et al., 2009). Another study estimated the effectiveness of LAGB from the available medical records of 28 patients in a hospital in Australia from 1996 to May 2005 (Ananthapavan et al., 2010), while the last study modelled the effect of time spent in physical activity after school on energy consumption, which led to weight loss and reduction in BMI as a consequence (Moodie et al., 2010).

The target populations of the interventions varied in age and gender, ranging from 5 to 19 years old. One study included both boys and girls in the analysis but focused solely on obese girls when considering program effectiveness, as the intervention showed no significant effect in boys (Wang et al., 2003). The duration of outcome follow-up in the reviewed articles ranged from 9 months to 3 years.

The definition of obesity and overweight in children and adults was inconsistent across studies. For the childhood definition, one study defined obesity as both BMI and triceps skinfolds (TSF) equal to or greater than the 85th percentile (Wang et al., 2003); another used the 85th BMI centile as a cut-off for diagnosing overweight (Brown et al., 2007); and a third defined childhood obesity in accordance with the International Obesity Task Force (IOTF) definition (Moodie et al., 2008a). For the adulthood definition, 2 studies (Wang et al., 2003, Brown et al., 2007) used an earlier definition of adulthood obesity, (BMI ≥27.3 in females and BMI ≥27.8 in males) instead of the current definition of BMI ≥30 kg/m^2 (World Health Organization, 2000). One study focused on severely obese adolescents and applied a BMI cut-off of 35 or greater for all participants (Ananthapavan et al., 2010). Three studies did not mention the definition of obesity used in the articles (Moodie et al., 2009, Magnus et al., 2009, Moodie et al., 2010).

All studies compared the interventions with the status quo; the costs and outcomes of alternative programs were not estimated. The ages of target populations and sample sizes of the studies varied considerably, particularly in the hypothesis-based studies by Magnus et al., Moodie et al., Ananthapavan et al, and Moodie et al. (Magnus et al., 2009, Moodie et al., 2009, Ananthapavan et al., 2010, Moodie et al., 2010). Two studies reported the costs of the intervention in US dollars (Wang et al., 2003, Brown et al., 2007), while the others reported costs in Australian dollars (Magnus et al., 2009, Moodie et al., 2009, Moodie et al., 2008a, Ananthapavan et al., 2010, Moodie et al., 2010). The currency year bases used in the reviewed articles ranged from 1996 to 2004. Of these, 5 studies from the ACE-obesity project were based on figures from 2001.

2.3.2 Characteristics of the economic analyses

The main differences in the estimation of health outcomes are due largely to the assumptions made in each model. These assumptions can be categorized into 2 types.

The first type of assumption appeared in the 2 US-based studies (Wang et al., 2003, Brown et al., 2007). In these studies, it was assumed that overweight children have a higher risk of developing obesity at ages 21-29 and 40-64 than do normal-weight children. Obesity at ages 40-64 increases the probability of death and reduces quality of life, leading to a reduction in life expectancy compared to those in the normal weight range. The long-term health outcomes were measured only between ages 40 and 64.

The second type of assumption, used in the 5 Australian-based studies (Magnus et al., 2009, Moodie et al., 2009, Moodie et al., 2008a, Ananthapavan et al., 2010, Moodie et al., 2010), was based on the link between the change in BMI and its impact on mortality and morbidity rates. The cohort in the model was divided by age (5-9, 10-14, 15-19) and gender (males, females), resulting in 6 different groups. Long-term health consequences were estimated for either the expected lifespan or until age 100.

Concerning the intervention costs, 2 studies approximated the incurred costs of intervention from the actual spending in the trials (Wang et al., 2003, Brown et al., 2007). One study estimated the costs in the cost-effectiveness analysis based on the trials but projected the intervention coverage to a national scale (Moodie et al., 2008a). The rest calculated the costs of the intervention from a hypothetical model (Magnus et al., 2009, Moodie et al., 2009, Ananthapavan et al., 2010, Moodie et al., 2010).

The techniques used in the reviewed studies to assess the cost offsets, or medical costs saved by the intervention, significantly affect the quality of the cost estimates. The cost offsets in the Wang et al. (2003) study were inferred from Gorsky et al. (1996), which estimated the medical costs of obesity-related diseases, including heart disease, hypertension, diabetes mellitus, gallstones, and osteoarthritis, of women between ages 40 and 65. In contrast, Brown et al. (2007) adopted the cost offsets from Oster et al. (1999), which calculated the medical costs of obesity in men and women between ages 35 and 64 from related diseases, including hypertension, hypercholesterolemia, type 2 diabetes, coronary heart disease, and stroke; in addition, they also conducted a sensitivity analysis using the medical costs from Gorsky et al. (1996). Contrary to Wang et al. and Brown et al., the other four studies stated the cost offsets in the analysis without mentioning the source of information or the method of calculation (Moodie et al., 2008a, Magnus et al., 2009, Moodie et al., 2009, Moodie et al., 2010). The study by Ananthapavan et al., meanwhile, referred the source of cost offset estimates to a paper by Cater et al. (2008). However, it is still unclear how the cost offsets were estimated.

Concerning the long-term health consequences, two studies estimated the outcomes in terms of QALYs (Wang et al., 2003, Brown et al., 2007). QALYs from both were approximated over a period of 25 years, between ages 40 and 64. The remainder of the studies illustrated the health benefits using DALYs, which accounted for life expectancy from the childhood period until death (Magnus et al., 2009, Moodie et al., 2009, Moodie et al., 2008a, Ananthapavan et al., 2010, Moodie et al., 2010). All selected papers applied a discount rate of 3% for both costs and outcomes in the base case analysis.

In reporting outcomes with regard to the cost-effectiveness of the interventions, 2 studies reported extreme values. Moodie et al. (2009) found an extremely high ICER of AU$760,000 per DALY. In contrast, Magnus et al. (2009) showed that the intervention was dominant, i.e. the intervention simultaneously reduced total costs and improved health outcomes. The extreme ICER estimates might be attributed to the fact that the effectiveness of the interventions was derived from hypothetical models, which may impact ICER estimates.

Table 2.2: Summary of the modelling characteristics

	Wang et al. (2003)	Brown et al. (2007)	Moodie et al. (2008)	Magnus et al. (2009)	Moodie et al. (2009)	Ananthapavan et al. (2010)	Moodie et al. (2010)
Modelling characteristics							
Population base	The US	The US	Australia	Australia	Australia	Australia	Australia
Model	Decision tree	Decision tree	Markov model	Markov model	Markov model	Markov model	Markov model
Perspective	Societal	Societal	Societal	Public Health	Societal	Societal	Societal
Main assumption of the model	Overweight in children at age 14 increases the risks of overweight between ages 21-29 and 40-64; overweight between ages 40-64 reduces life years and quality of life, increases probability of death	Overweight in children at age 11 increases the risks of overweight between ages 25-29 and 40-64; overweight between ages 40-64 reduces life years and quality of life, increases probability of death	DALYs calculation was based on the association between the decrease in BMI and the reduction in mortality and morbidity rates due to obesity	DALYs calculation was based on the association between the decrease in BMI and the reduction in mortality and morbidity rates due to obesity	DALYs calculation was based on the association between the decrease in BMI and the reduction in mortality and morbidity rates due to obesity	DALYs calculation was based on the association between the decrease in BMI and the reduction in mortality and morbidity rates due to obesity	DALYs calculation was based on the association between the decrease in BMI and the reduction in mortality and morbidity rates due to obesity
Outcome unit	QALY	QALY	DALY	DALY	DALY	DALY	DALY
Outcome calculation	QALYs were calculated over 25 years, from age 40 to 64	QALYs were calculated over 25 years, from age 40 to 64	DALYs were estimated for either the entire lifespan or until age 100 years	DALYs were estimated for either the entire lifespan or until age 100 years	DALYs were estimated for either the entire lifespan or until age 100 years	DALYs were estimated for either the entire lifespan or until age 100 years	DALYs were estimated for either the entire lifespan or until age 100 years
Discount rate	3% for costs and outcomes	3% for costs and outcomes	3% for costs and outcomes	3% for costs and outcomes	3% for costs and outcomes	3% for costs and outcomes	3% for costs and outcomes

Table 2.3: Summary of the intervention characteristics

	Wang et al. (2003)	Brown et al. (2007)	Moodie et al. (2008)	Magnus et al. (2009)	Moodie et al. (2009)	Ananthapavan et al. (2010)	Moodie et al. (2010)
Intervention characteristics							
Overview of the intervention	"Planet Health", a school-based intervention tailored to reduce childhood obesity through health education (diet, physical activity, reduction of TV viewing time)	"CATCH" (Coordinated Approach to Child Health), a school-based program including a classroom curriculum at each level, a physical education program, school food-service modification, and family- and home-based programs	"LEAP" (live, eat and play), a family-based GP-mediated intervention focused on lifestyle, diet, and exercise	"Removing TV advertising to reduce childhood obesity", a hypothetical intervention based on the current practice in Australia with a one-year follow-up projection	"Walking School Bus" program, an intervention encouraging children to walk to schools under the supervision of safety volunteers. The assumed follow-up duration is one year.	Laparoscopic adjustable gastric banding (LAGB), a bariatric surgery, performing in severel obese adolescents to reduce weight.	"AASC" (Active After-school Communities program, a school-based measure to increase physical activity in after school from 3pm to 5:30pm for 2-3 sessions a week for 8 weeks per school term, 4 terms a year.
Modelling assumption of the impact from the intervention on BMI change	Planet Health leads to a healthy lifestyle and promotes a healthy weight in children. The prevalence of childhood overweight is therefore decreased.	CATCH promotes a healthy lifestyle and weight in children, resulting in the reduction of childhood overweight prevalence.	LEAP encourages healthy lifestyle and eating habit in overweight and obese children. Behavioral modificantion leads to the decrease in weight and BMI consequently.	Removing TV advertising of energy-dense, nutrition-poor (EDNP) food and beverages reduces product consumption. Changes in energy intake consequently decrease body weight and BMI.	Walking time increases energy expenditure, resulting in weight loss and reduction in BMI as a result.	LAGB results in weight reduction and decrease in BMI as a consequence.	AASC increases energy expenditure, resulting in weight loss and reduction in BMI as a result.
Sources of the outcome estimates	RCT	RCT	RCT	Modelling	Modelling	Clinical data from the medical record	Modelling
Aim of the intervention	Prevention	Prevention	Treatment	Prevention+Treatment	Prevention+Treatment	Treatment	Prevention+Treatment

	Wang et al. (2003)	Brown et al. (2007)	Moodie et al. (2008)	Magnus et al. (2009)	Moodie et al. (2009)	Ananthapavan et al. (2010)	Moodie et al. (2010)
Intervention characteristics							
Characteristics of sample in the model (eligibility criteria)	overweight middle school girls (≈12 years old)	overweight third-grade children (age 8)	obese and overweight children aged 5-9 years except very obese children (BMI z-score ≥ 3.0)	all children aged 5-14 years in Australia	children between ages 5-7 in Australia	Australian adolescents aged 14-19 with BMI ≥35	Australian children aged 5-11
Follow-up duration from the clinical evidence	2 years	3 years	9 months	-	-	3 years	-
Obesity/ overweight definition	*Childhood overweight: BMI≥85th centile Adulthood: F ages 21-29: BMI≥27.3 F after age40: BMI≥25*	*Childhood overweight: BMI≥95th centile Adulthood: Ages 25-29: F: BMI≥27.3; M: BMI≥27.8 Age>40: BMI > 30*	Childhood overweight: IOTF definition	-	-	Definition of severe obesity between ages 14-19: BMI ≥35	-
Effectiveness of the intervention	*Intervention*: prevalence of overweight among girls decreased 3.2%; *Control*: prevalence of overweight increased 2.2%; No difference among boys	*Intervention*: prevalence of overweight increased by 1% for boys and 2% for girls; *Control*: prevalence of overweight increased by 9% for boys and 13% for girls.	*Intervention*: mean BMI decreased by 0.31 *Control*: no change	*Intervention*: BMI reduced by 0.26 *Control*: no change	*Intervention*: BMI reduced by 0.014 and 0.015 for boys and girls respectively *Control*: no change	*Intervention*: the mean BMI reduced by 13.93 *Control*: no change	*Intervention*: BMI reduced by 0.085-0.112 units *Control*: no change
Total intervention costs per participant	US$28	US$104	AU$650.5	AU$0.054	AUD$2,900	AUD$31,000	AUD$488.5
Comparator	No intervention	No intervention	No intervention	No intervention	No intervention	No intervention	No intervention

	Wang et al. (2003)	Brown et al. (2007)	Moodie et al. (2008)	Magnus et al. (2009)	Moodie et al. (2009)	Ananthapavan et al. (2010)	Moodie et al. (2010)
Intervention characteristics							
Sample size in the trial	*Intervention*: 310 girls *Control*: hypothetical 310 girls	*Intervention*: 423 children (F 199; M 224) *Control*: 473 children (F 224; M 249)	*Intervention*: 73 children *Control*: 80 children	-	-	Intervention: 2 boys; 26 girls	-
Targeted population in the CEA model (only children)	*Intervention*: 310 girls *Control*: hypothetical 310 girls	*Intervention*: 423 children (F 199; M 224) *Control*: 473 children (F 224; M 249)	Assuming the eligible children aged 5-9 of 9,685 were covered	Assuming children aged 5-14 were all covered = 2,400,000	Assuming eligible children in the recruited schools were included = 15,681	Assuming the eligible adolescents of 4,120 were included	Assuming the eligible children of 82,500 were included

Table 2.4: Summary of the reviewed literature: results

	ICER	Year based for currency value	Miscellaneous
Wang et al. (2003)	US$4,305	1996	Net benefit was estimated to be US$7,313
Brown et al. (2007)	US$900	2004	Net benefit was estimated to be US$68,125
Moodie et al. (2008)	AU$4,670	2001	The intervention was modeled based on the assumption that the coverage of LEAP extended to 9,685 children with a follow-up duration of 1 year.
Magnus et al. (2009)	dominant	2001	-
Moodie et al. (2009)	AU$0.76M	2001	-
Ananthapavan et al. (2010)	AU$4,400	2001	The effectiveness of LAGB reflects the practices of 1 group of surgeons in Australia.
Moodie et al. (2010)	AU$82,000	2001	The intervention costs per participant are calculated by total costs ($40.3M)/ total participants(82,500) = $488.5

2.4 Discussion

Results from the reviewed studies showed that most interventions are cost-effective with respect to the country-specific threshold, except those in Moodie et al. (2009) and Moodie et al. (2010). However, it might not be possible to determine which types of intervention tend to be more cost-effective by only looking at the intervention design or effectiveness, as number of components need to be taken into account in the economic evaluation; these include the economic modelling techniques applied, the costs of intervention, medical costs offset, and the duration of follow-up.

As evidenced by the results, the major factors affecting the quality of the studies reviewed are the methods and assumptions used to estimate future costs and outcomes of obesity. Appendix 3 provides the critical assessment of the reviewed articles. In this regard, certain factors influencing the results and quality of economic evaluation should be emphasized in particular for a better understanding of their impact. This section aims to discuss these points in order to highlight the strengths, weaknesses, and gaps of the economic models, which will benefit future research.

2.4.1 The modelling assumptions

Two types of economic model were employed in the reviewed articles to estimate long-term health outcomes: the decision tree and the Markov model.

Based on the decision tree technique, the first model estimates health outcomes in terms of QALYs. This method was used by Wang et al. (2003) and Brown et al. (2007). The flowchart of QALYs calculation is shown in Figure 2.2. As can be seen, QALYs gained from the intervention were evaluated only during the 25-year period between ages 40 and 64.

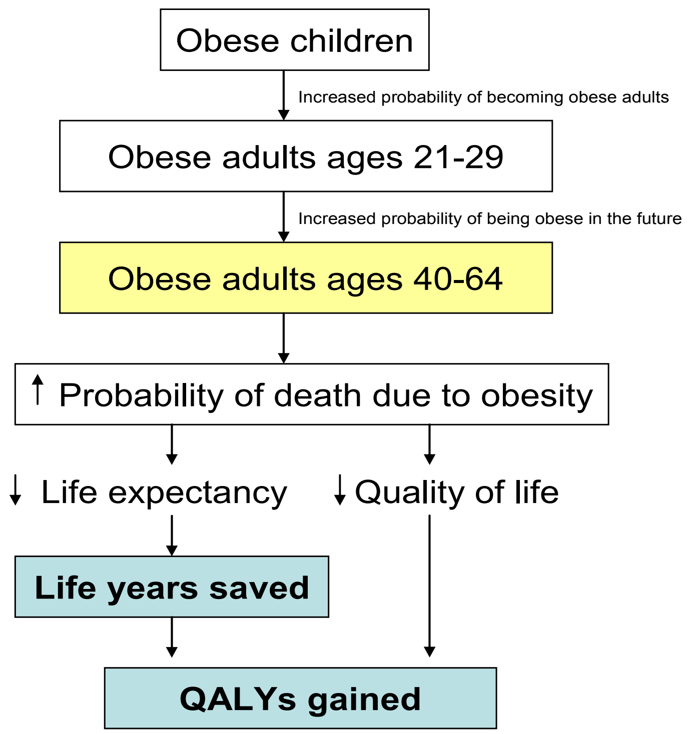


Figure 2.2: Flowchart of the methods used in estimating QALYs

Four points of concern arise pertaining to the application of this modelling method. Firstly, it assumed that all children survived in the model until they died of obesity-related conditions between ages 40 and 64. Failure to consider the possibility of children leaving the model before age 40 due to other causes of mortality led to an overestimation of the rate of survival. The age- and gender-specific mortality rates for all causes should be adjusted in the model to improve the accuracy of estimation.

Secondly, the model limited the duration of the health outcome measurement to ages 40 to 64, while in fact obesity increases the risks of morbidity and mortality in adulthood at all ages (Flegal et al., 2005). This raises an issue concerning the underestimation of mortality and morbidity attributable to obesity. Not only were obesity-related health consequences occurring before age 40 dismissed, but shortened life expectancy due to obesity after age 64 was also unaccounted for. The measure of obesity's effects on health could be improved if the model extends the duration of the outcome measurement.

Thirdly, the impact of obesity on health is over-generalized by the model. Evidence shows that adverse health consequences are positively correlated with the severity of obesity (World Health Organization, 2000), i.e. the more obese, the higher the risk of co-morbidities. This leads to differences in quality of life among those experiencing different severities (Sach et al., 2006). However, Wang et al. (2003) and Brown et al. (2007) failed to account for this in their models. The estimated impact of obesity can be improved if the severity of obesity is taken into consideration.

Finally, defining the effectiveness of interventions as the reversal of obesity leads to underestimation of the program benefits. Interventions can produce benefits in the form of weight reduction even if children are still categorized as obese. For example, an extremely obese child losing 15 kg due to an intervention has improved quality of life and lowered health risks, even though he is still obese. Additionally, overweight children who participated in the intervention cannot be counted as successes, as they are already below the obesity cut-off from the start.

Another economic model, used in estimating DALYs in the other 5 studies (Magnus et al., 2009, Moodie et al., 2009, Moodie et al., 2008a, Ananthapavan et al., 2010, Moodie et al., 2010), is the Markov model. The methods used to estimate DALYs in these 5 studies were as reported in Haby et al. (2006), which estimated DALYs saved based on one-unit changes in BMI.

Figure 2.3 shows the summary of the DALYs calculation. DALY is the sum of years of life lost (YLL) and years lived with disability (YLD), calculated by adjusting the expected life years with the disability weight. The DALY estimations were based on 6 different cohorts stratified by age (5-9; 10-14; and 15-19) and gender (male and female). The DALYs saved were derived mainly from the potential impact fraction (PIF), which is "the percentage reduction in disease or death that would take place if exposure to the risk factor were reduced to the counterfactual distribution" (Ezzati et al., 2002). In this case, PIF is calculated by comparing the normal population with a hypothetical non-obese population. As illustrated by Figure 2.4, an easy form of PIF can be calculated by comparing the incidence rate of the population with that of the unexposed, which is $(I_p - I_u)/ I_p$. In this case, I_p and I_u are the mortality rates of the total population and the non-obese population, respectively.

For instance, if the mortality and morbidity rate from all causes = 1% and PIF of obesity = 0.05, the mortality and morbidity rate would be decreased to (1-0.05)*1% = 0.95% if obesity and its effects were eliminated from the population.

Although this method compares health outcomes between populations with and without adverse effects from obesity, some interventions are aimed only at obese children, not the entire population. In Figure 2.4, the difference in the incidence rate would thus be $I_e - I_u$, instead of $I_p - I_u$, where I_e in this case is the mortality rate of the obese population. Haby et al. (2006) suggested approximating I_e by adjusting I_p with the correction factor. An example in the paper states "If in the example we found that 25% of the total amount of disease attributed to BMI occurred in the top 10% of the BMI distribution, the correction factor was 25/10 = 2.5" (Haby et al., 2006).

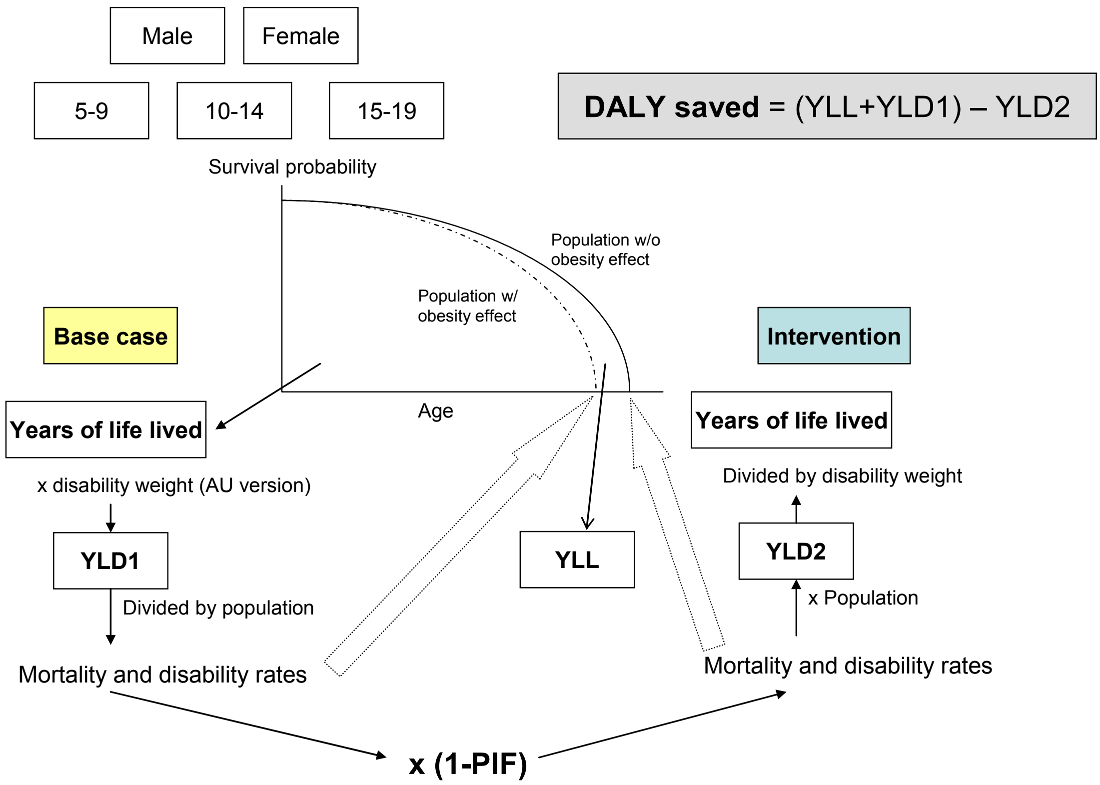


Figure 2.3: Summary of the methods used to estimate DALYs

The DALYs saved are represented by the area under the curve between the 2 survival curves with and without obesity risks included. In Figure 2.3, said area represents YLL, and the area adjusted by the disability weight is YLD (YLD2 – YLD1). Therefore, the DALYs saved = YLL – YLD = (YLL + YLD1) – YLD2.

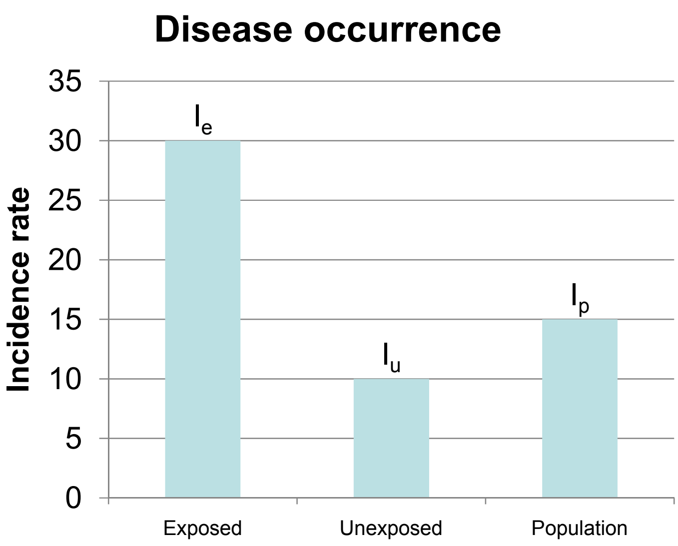


Figure 2.4: Comparing the incidence of disease occurrence in the exposed, unexposed, and population

Concerns with regard to the use of this model stem from the interpretation and implications of the results. The impact of obesity on disability and mortality was estimated in terms of DALYs regardless of the BMI baseline. On one hand, it is clear that both overweight and obesity increase the risk of comorbidity, as higher BMI increases risk regardless of categorization (World Health Organization, 2000). However, a clear relationship was shown between increased BMI and mortality risk only in people with BMI ≥ 30 (Calle et al., 2003, Flegal et al., 2005). It is still ambiguous whether or not overweight (BMI 25-30) increases the mortality risk (Garfinkel, 1986, Calle et al., 2003, Flegal et al., 2005). Therefore, it may not be accurate to use this method to assume the health benefits gained from reducing BMI when the target population is not obese.

An example demonstrating this fallacy can be seen in Moodie et al. (2009)'s study. The population targeted by the walking school bus program was children between ages 5 and 7 with average BMIs of 16.2-16.5, whereas the BMI cut-offs for obesity in children in this age range, according to the IOTF definition, are between 19 and 21 (Cole et al., 2000). Hence, the assumption that a decrease in BMI improves health by reducing obesity-related risks may not be relevant, as children were not obese to begin with. This pitfall also appears in Moodie et al. (2010), where none of the targeted population were considered obese. The BMIs of all boys were lower than the overweight cut-offs according to the IOTF definition, while only a small number of girls were considered overweight.

Therefore, this is another example emphasizing the need for the DALY estimation method to be used carefully, especially when evaluating non-obese children.

2.4.2 QALYs versus DALYs

Chapter 1 illustrates the theoretical concepts relevant to the use of QALYs and DALYs, alongside their advantages and drawbacks. In considering the assessment of long-term health effects of obesity in children, it is evident that evaluation involves a number of obesity-related co-morbidities and health risks, which highlight the theoretical limitations in the use of DALYs. QALYs are therefore considered a fitter measure in economic evaluation than DALYs in this context.

In reality, however, certain issues affect the theoretical advantages in the use of QALYs over DALYs. Firstly, although economic evaluation is in favour of the use of QALYs in relation to DALYs, no literature has explored to what extent the use of DALYs changes the results compared with their counterpart. Preference is still based on the assumptions regarding the construction of both measures. The lack of evidence showing the impact of obesity-related co-morbidities on changes in DALYs compared with QALYs raises the question as to whether the two measures are truly different, as well as the ability of the HRQL to account for co-morbidity in QALYs estimation. Future research may be required to explore this question. Secondly, the DALYs estimation technique was developed for use in one project (ACE-obesity), which evaluated the cost-effectiveness of many child obesity interventions at the same time. Despite their inability to account for co-morbidities and health risks of obesity, however, DALYs estimation proves to be the one measure that is consistently applicable across studies. In contrast, there is as of yet no single agreed-upon technique to assess health outcomes in terms of QALYs. In conducting economic evaluation, using the same methods to estimate health outcome units facilitates results comparison; for the use in outcome comparison, therefore, using DALYs may be advantageous in this regard. These observations also emphasize the need of a standardized economic method to assess QALYs in child obesity interventions in order to maximize the comparability of results across studies. Chapter 5 of this thesis elaborates upon this idea and suggests further research to develop a standardized economic model to estimate

QALYs of child obesity interventions based on the cost-effectiveness model developed in this thesis.

2.4.3 Discount rate

Experts almost uniformly apply discounting to the cost when conducting cost-effectiveness analysis (such as in Gold et al., 1996, Drummond et al., 2005); however, the recommended discount rates vary. Drummond et al. (2005) conclude that discounting should be applied in the base case analysis, with alternative scenarios applying a rate of 0-5% to be included in the sensitivity analysis.

Unlike cost, there is no consensus on whether or not the future health outcome should be discounted to the present time. Drummond et al. (2005) present 3 main reasons for not discounting outcomes: 1) it is not pragmatic to think that people invest in health for the future or trade healthy years over time; 2) discounting implies less weight to life in the future; and 3) there is evidence showing that health and cost should be discounted at different rates. However, the authors also provided arguments to support discounting future health, mainly due to the inconsistency of the analysis (Gold et al., 1996, Drummond et al., 2005), which was demonstrated in the papers by Weinstein and Stason (1977) and Keeler and Cretin (1983). These fallacies led to some impossible conclusions, such as delaying the building of a hospital because the health benefits gained are still the same (no discount) while the capital investment is cheaper every year (discount). Another impossible scenario from Drummond et al. (2005) describe "… a health programme giving rise to $1 of health benefits each and every year stretching into the future would be worthwhile whatever the size of the initial capital sum …".

In addition to these individual findings, applying no discount rate on health outcomes causes inconsistency in policy decision making. This is particularly the case when comparing health outcomes with other commodities in resource allocation, as economic evaluation traditionally discounts the value of commodities (Drummond et al., 2005).

Debates about discounting rates are ongoing, as seen in the example of a recent publication (Claxton et al., 2011). Claxton et al. suggest a number of factors that should be taken into consideration in deciding whether or not the discount rate of the health outcome should be equal to that of the costs. These factors include the change in the ICER threshold, the technological growth in healthcare, and healthcare budgets. The authors conclude that the discount rate for costs and health outcomes should be the same only if the ICER threshold is expected to be constant over time; however, they also present supportive evidence indicating that the discount rate applied to health should be lower than that applied to costs (Claxton et al., 2011).

An issue worth noting is the application of differential discounting to the public health projects that produce health benefits in the long-term future, such as child obesity interventions or vaccination programmes. Inconsistency in applying discounting may result from individual perceptions of the value of the interventions. For example, in order to favour future generations, policymakers may deliberately bias the cost-effectiveness results by reducing the discount rate on benefits in order to decrease the ICER.

Despite ongoing disagreements regarding discounting effects, the current recommendation is to discount effects at the same rate as costs (Viscusi, 1995, Gold et al., 1996, National Institute for Health and Clinical Excellence, 2008). All selected studies applied the discount rate of 3% on both costs and health outcomes.

2.4.4 Medical cost offsets

Assessing the medical costs of obesity is another challenge in conducting cost-effectiveness analysis of child obesity intervention, which could easily result in uncertainty across studies. Though it is acknowledged that obesity increases the risk of many chronic diseases, resulting in the rise of mortality, obesity rarely causes death directly. Therefore, the costs of obesity largely depend on the number of adverse effects it induces. Accounting for the undesirable health consequences of obesity is dependent on how researchers determine the impact of obesity on the disease, and the costs of treatment. For example, Gorsky et al. (1996) derived the medical costs of obesity from 5 related diseases: coronary heart disease, hypertension, diabetes, gallstones, and osteoarthritis. Oster et al. (1999), meanwhile, included 5 obesity-related diseases in their cost estimate, 3 of which overlapped with Gorsky et al.'s: hypertension, hypercholesterolemia, type 2 diabetes, coronary heart disease, and stroke. Table 2.5 shows the relative risk of certain diseases from obesity, as also mentioned earlier in Chapter 1. It is worth noting that Gorsky et al. and Oster et al. excluded the treatment costs of colorectal cancer and ovarian cancer when evaluating the medical costs of obesity, despite higher relative risks, but included the costs of osteoarthritis and stroke. This is probably because the treatment costs of the 2 cancers may not be as high as for chronic conditions, as cancer patients have a short life expectancy compared with those with stroke or osteoarthritis. Another possibility is due to a lack of the information needed to estimate the costs of these 2 cancers.

Table 2.5: The relative risk of developing diseases in obese people, from a literature review

	Relative risk	
Disease	**Women**	**Men**
DM type 2	12.7	5.2
Hypertension	4.2	2.6
Myocardial infarction	3.2	1.5
Colon cancer	2.7	3
Angina	1.8	1.8
Gall Bladder diseases	1.8	1.8
Ovarian cancer	1.7	-
Osteoarthritis	1.4	1.9
Stroke	1.3	1.3

Source: the Comptroller and Auditor General (2001)

Inconsistency in estimating the medical cost offsets inevitably affects the cost estimation in cost-effectiveness analysis to some extent. This is particularly an issue when the cost-effectiveness papers do not refer to the data source or state how the cost offset was calculated, such as in the reviewed articles (Moodie et al., 2008a, Magnus et al., 2009, Moodie et al., 2009, Moodie et al., 2010). This affects the quality of the results as a consequence.

To deal with this problem, analyses should account for as many obesity-related conditions as possible to make the cost estimate comprehensive. A good example of the estimation of the costs of obesity appears in McPherson et al. (2007), where 10 obesity-related diseases were taken into account in the cost estimate.

2.4.5 Sources of outcome estimates

The reliability of intervention effectiveness is dependent on the strength of the evidence. Three of the reviewed studies measured the effectiveness of interventions using randomised controlled trials, while the rest approximated intervention effects using hypothetical models. Intervention effectiveness derived from a hypothetical model is considered weak evidence, which affects the reliability of the results and future utilization.

2.4.6 Duration of follow-up

The duration of follow-up is an important factor in determining the sustainability of intervention effectiveness. It is evident that all reviewed papers pertain to studies with little or no long-term follow-up. This is a major concern, as seen in two responses (Dalziel and Segal, 2006, Segal and Dalziel, 2007) to the studies by Wang et al. (2003) and Haby et al. (2006). A longer follow-up period indicates a higher degree of sustainability, especially if follow-up extends until adulthood. While the short follow-up duration is of concern, however, it is only fair to acknowledge that difficulty in attaining long-term funding leads to difficulties in extending the follow-up period over a decade or more. One proposed solution for this is to use routinely collected data in following up child obesity interventions. This would show the long-term health outcomes of participants with reduced time and costs spent on the follow-up.

2.5 Conclusions

Childhood obesity incurs high health expenditures. Some evidence shows that interventions designed to treat and prevent obesity in children are effective in reducing childhood obesity. However, the cost of intervention is still an issue. Cost-effectiveness analysis is a method that assesses the costs and outcomes of interventions of interest. This literature review summarizes the existing cost-effectiveness studies of child obesity interventions up to the present day. As can be seen, only a small fraction of existing child obesity interventions has been the subject of economic evaluation with the ICER provided. This could be due to difficulties in conducting economic evaluation, particularly with regard to estimating outcomes in terms of life years. In addition, the results of economic evaluation need to take into consideration the uncertainties in the calculation caused by the modelling methods, outcome units, discount rate, the medical cost offsets, sources of outcome, and duration of follow-up. The cost-effectiveness of child obesity interventions should be assessed more regularly to provide economic information that helps policymakers allocate resources more efficiently.

This literature review acknowledges its linguistic limitations, which potentially reduces the number of articles found. Moreover, the databases containing gray literature or papers in press were excluded from the search.

While there are limitations associated with modelling, modelled evaluations of an intervention's effectiveness and cost-effectiveness analysis do fill a gap in the literature. The results try to answer the question of how efficiently resources have been used to tackle child obesity via the interventions. As the available techniques used in evaluating future health outcomes still have a number of limits and weaknesses, future research should focus on developing new techniques to improve the future health outcome estimation for child obesity interventions.

Chapter 3: The MEND 7-13 programme

3.1 Introduction

Chapter 2 states that only a handful of cost-effectiveness studies of child obesity interventions are available, possibly due to difficulty in estimating the future health outcomes of childhood obesity. The chapter also describes the current existing economic evaluation techniques used to assess future health outcomes, as well as their limitations. Despite the techniques' weaknesses, it is essential for child obesity interventions to incorporate cost-effectiveness research to fill a gap in decision making processes. This thesis addresses this issue by not only producing a cost-effectiveness study of a child obesity programme, but also improving an existing economic modelling technique, allowing for a more comprehensive assessment of the future health outcomes of the interventions.

In order to demonstrate cost-effectiveness analysis, MEND 7-13, a child obesity intervention, is used as a case study. MEND 7-13 has been operating for about a decade with the aim of treating obesity in children aged 7-13. The programme has been widely implemented in England and Wales.

Understanding the background of MEND 7-13 before starting the cost-effectiveness analysis will help give insight into how the intervention has been implemented. Therefore, Chapter 3 presents an overview of MEND 7-13, including background, programme effectiveness, and internal factors associated with the benefit uptake of the participants. The first part of this chapter introduces MEND 7-13, with emphasis on the development of the programme, the delivery of treatment, and evidence of effectiveness. The second part of the chapter then uses statistical analyses to assess internal proxies influencing the effectiveness of MEND 7-13 in terms of the BMI change rate and obesity status.

Part 1: Background of MEND 7-13

3.2 What is MEND 7-13?

The MEND (Mind, Exercise, Nutrition, Do It!) 7-13 programme is an intervention to treat childhood overweight and obesity. The target of the MEND 7-13 programme is overweight and obese children, with BMIs above the 91st centile of the UK 1990 BMI chart, between ages 7 and 13 in England and Wales. Supported by families and communities, MEND 7-13 aims to implement healthy lifestyles in children and their families through food and physical education. The programme is comprised of 20 2-hour sessions over a 10-week period, followed by a 2-year support system to promote programme sustainability (MEND Central Ltd., 2010).

The MEND Programme was first developed in 2001 at the Institute of Child Health, University College London, and the Great Ormond Street Hospital for Children. The programme was devised and tested in the feasibility phase and in an RCT from 2005-7 (Sacher et al., 2005, Sacher et al., 2010). In 2007, MEND 7-13 was publicly implemented with over 5,000 obese young participants and their parents during the so-called roll-out phase.

While participants may have heard of MEND 7-13 through several channels, such as newspaper, TV, or radio, the only method available for all participants to enrol in MEND 7-13 is via online registration at the MEND website.

3.2.1 Programme Content

Implemented over a 2-year duration, the MEND programme is comprised of: 1) a 10-week intervention period (MEND Programme): 20 2-hour programme sessions by the local delivery team; 2) 42-week transition: follow-up at 6 and 12 months, with a recap of the lessons recommended by MEND Central; and 3) 12-month maintenance: media-based support materials such as a magazine and website, with voluntary access (MEND Central Ltd., 2010).

The key elements of the programme include:

1. **Mind**: The contents of these theoretically-oriented sessions are aimed at informing families about social learning theory and behaviour modification. Children, parents, and staff work together at the beginning of the 10-week period to set targets and incentives, with the goal of promoting long-term behavioural change. There are also parent-only sessions, which aim to develop skills in particular topics such as internal and external triggers (e.g., TV viewing and stimulus control), understanding the difference between hunger and craving, being a role model on eating habits and physical activity, helping their children to develop self-esteem, and addressing potential problems related to bullying (MEND Central Ltd., 2010).
2. **Nutrition**: Based on government-issued healthy eating guidelines, the programme introduces practical skills targeting nutrition. These include regular and balanced healthy eating, learning about portion sizes, and changing unhealthy behaviours through various methods. Children and parents learn practical skills through interactive games, visual demonstrations, and other activities, which include touring supermarkets, reading and understanding food labels, and trying healthy recipes (MEND Central Ltd., 2010).
3. **Exercise**: Physical activities consist of multi-skill team games that are non-competitive, structured, and progressive, which aim to improve balance, agility, and coordination. Sessions are designed to appeal and engage young people who are not motivated by traditional activities by delivering in a fun and game-based manner (MEND Central Ltd., 2010).
4. **Family/carer involvement**: It is mandatory for parents or carers to participate in the programme with children throughout the process, during both theoretical sessions and activities. The programme also engages with children to set individual rewards (MEND Central Ltd., 2010).

After the 10-week programme, MEND 7-13 continues to provide resources and support to families and children to maintain healthy lifestyle changes and achieve a healthier weight through MEND World. MEND World is a standard part of MEND's offerings and is managed by the Sustainable Health Outcomes team. Resources provided by MEND World include the centralized and local levels as follows:

Centralized MEND World resources – 1) MEND World passport – an A5 sheet of paper that allows children to track their changes and connect to other MEND graduates via the MEND World website; 2) My World magazines – a quarterly magazine containing inspiring stories of MEND graduates, practical information, and interactive tools to help families achieve and maintain healthy behaviours; 3) MEND World website – a self-directed learning and behavioural change strategy website to help guide families through health-based challenges; and 4) continued support – from the MEND Sustainable Health Outcomes team, which remains available to promote a long-term healthy lifestyle.

Local MEND World activities – activities provided by local delivery teams to support families after graduation from their local programme. Activities suggested by MEND at the local MEND World level include the following: 1) a minimum of 10-12 weekly graduate exercise sessions; 2) free or discounted access to local physical activity sessions or leisure services; 3) healthy growth check every 3 months; and 4) local strategies to promote ongoing contact and support networks (MEND Central Ltd., 2010).

3.2.2 Programme Delivery

Local people are trained, equipped, and employed to deliver MEND 7-13 in their own communities. The delivery team consists of a theory (Mind and Nutrition) leader, exercise leader, and programme assistant, who are supervised by a Programme Manager. Standardized training is delivered to all programme managers and delivery teams. In this regard, team members must attend a 4-day training course prior to the start of the programme (MEND Central Ltd., 2010).

In addition to the provision of training, MEND supplies media support for staff in order to achieve quality enhancement. Delivery teams receive a full set of program delivery manuals, complete with session plans, objectives, direct teaching notes, and desired outcomes. Moreover, teams are supported by an internet-based programme management portal, which permits monitoring of health outcomes, as well as facilitating the consistency of delivery and monitoring (MEND Central Ltd., 2010).

Even when delivered at volume from community settings by non-specialists, MEND 7-13 shows the same effectiveness observed in clinical trials at 3-month follow-up (Kolotourou et al., 2009). With regard to this, it is noted that the MEND eligibility criteria in the trial is BMI $\geq 98^{th}$ centile, whereas the rollout programme extends eligibility to children with BMI $\geq 91^{st}$ centile. Currently, the programme is being delivered in over 300 community venues in the UK, and some 30,000 people, including children and parents, have participated to date.

3.2.3 Justification for the MEND 7-13 Programme

The MEND 7-13 programme was tailored based on the published evidence from 3 sources as follows (MEND Central Ltd., 2010): 1) evidence from systematic reviews of the effectiveness of interventions to treat obesity in children (Summerbell et al., 2003, Oude Luttikhuis et al., 2009); 2) guidance and recommendations from expert consensus, especially the National Institute for Health and Clinical Excellence (NICE) guidance for the prevention and treatment of child obesity; and 3) delivering the programme in a community setting in order to maximize the number of participating families according to evidence and expert recommendations for safe implementation (MEND Central Ltd., 2010).

Family involvement is one of the requirements of the MEND 7-13 programme because parents are an important primary mediator of change in making behavioural treatment more effective (Golan and Crow, 2004). A review also found that parents act as role models and change agents for their children in developing healthy eating habits to treat and prevent obesity (Golan and Crow, 2004).

Unlike other obesity treatments, the MEND 7-13 programme does not take a diet-restriction approach to prevent subsequent problems (NICE, 2006), as eating restriction is associated with an increased risk of eating disorders and body dissatisfaction in obese children.

Additionally, MEND 7-13 is a group-based community intervention, which is the most commonly used delivery format. There is evidence showing that group-based treatments are more effective than individual treatments because they are more efficient, enhancing therapeutic interactions between participants, increasing attendance rates, and being more cost-effective (Robinson, 1999a). Community interventions provide greater access to minority groups, counter stigma, and provide a social support network

MEND 7-13 not only complies with existing guidelines for the treatment of childhood obesity, but has also been updated several times to maintain effectiveness and reflect participant feedback.

In addition to programme effectiveness, MEND 7-13 was also tailored with cost-effectiveness in mind. One approach to reduce programme costs is to lower the standard level of skills and expertise of the staff, which increases feasibility when scaling up the coverage. The qualifications and staff expertise at each stage of the programme are summarized in Table 3.1.

Table 3.1: Professionals and expertises involved in MEND 7-13 at different phases

Programme phase	Programme staff
Feasibility	Paediatric dietician, clinical psychologist and physiotherapist
RCT	Health, social education, and exercise professionals
Rollout	Many additional non-specialists e.g. community health workers, YMCA staff

3.2.4 Programme Effectiveness

The programme attendance and retention rates are as shown in Table 3.2.

Table 3.2: Attendance and retention rates of MEND 7-13 in the feasibility study and RCT

	Feasibility	**RCT**
N of children	11	116
Mean programme attendance (%)	78%	86%
Drop out after recruitment (%)	9%	3%

Source: Sacher et al. (2005) and Sacher et al. (2010)

A randomized controlled trial of MEND 7-13 was conducted between 2005 and 2007, involving 116 children with BMI ≥ 98th centile of the UK 1990 BMI chart (Sacher et al., 2010). Change in waist circumference was the primary outcome, while the secondary outcomes were changes in BMI and % body fat. The RCT calculated the sample size with respect to the difference in waist circumference of 3 cm from the feasibility test at a significance level of 0.05 and the power of 80%. Ten percent of the required sample was added to account for drop outs, resulting in a final required sample of 48 per treatment group. The baseline values of health indicators were applied for outcome adjustment using linear regression. A significance level of 0.05 was applied in the analysis.

The study had 2 phases: 1) randomization, in which participants were randomly assigned into 2 groups (60 – intervention, 56 – control) and followed up for 6 months; and 2) after randomization, wherein all children in the control group received the intervention, while those in the intervention group during the first phase were followed for another 6 months (total of 1 year). Therefore, the randomization happened only during the first 6 months.

Concerning the attrition rate, 6 cases, all from the intervention group, dropped out before the trial started: 3 for medical reasons, 2 for social reasons, and one for no reason provided. At 6 months, 17 children from the intervention and 11 from the control groups did not attend the follow-up, leaving 37 children (62%) in the intervention group and 45 (80%) in the control group to be assessed at 6 months. The analysis was carried out without accounting for uncertainty due to missing data, or imputation analysis, in the sensitivity analysis. Therefore, this is considered a methodological weakness.

Table 3.3 illustrates the results of the RCT at 6 months. As can be seen, 14 health indicators were tested in the trial. Significant improvement in health indicators was reported for children who attended the programme compared with their control counterparts at the 6-month follow-up. These health indicators include waist circumference (4.1 cm reduction), BMI (1.2 units reduction), recovery heart rate (20.3 beats per minute reduction), physical activity hours (3.9 hours increase), and sedentary hours (5.1 hours decrease) (Sacher et al., 2010).

As the trial aimed to assess several health indicators at the same time, the significance level for each individual test needs to be adjusted before interpreting the results. However, Sacher et al. (2010) failed to do so, which is considered a weakness of the methodology. A method commonly used to adjust the significance level in analyses containing multiple comparison is the Bonferroni correction (Dunn, 1961). The Bonferroni correction simply divides the significance level (0.05) by the number of tests, which in this case is 14. The significance level, adjusted by the Bonferroni correction, is 0.05/14 = 0.0036. Using the adjusted significance level, it is seen that supportive evidence for improvement in health indicators still appears only in the recovery heart rate and time spent in sedentary activity, without baseline adjustment. When compared with baseline adjustment, evidence of outcome improvement is seen only in waist circumference, waist circumference z-score, BMI, BMI z-score, and recovery heart rate.

Table 3.3: Results from the MEND 7-13 randomized controlled trial at 6 months

Parameters	Intervention		Control		Difference (unadjusted)		Total	Difference (adjusted for baseline)	
	n	**Mean (sd)**	**n**	**Mean (sd)**	**Mean (CI)**	**p-value**	**n**	**Mean (CI)**	**p-value**
Waist circumference (cm)	37	77.7(7.2)	45	82.0(8.6)	-4.3(-7.8,-0.8)	0.02	81	-4.1(-5.6,-2.7)	<0.0001
Waist circumference z-score	37	2.53(0.58)	45	2.76(0.61)	-0.23(-0.5,0.03)	0.09	81	-0.37(-0.49,-0.25)	<0.0001
BMI (kg/m2)	37	25.7(3.3)	45	27.7(5.2)	-1.9(-3.8,0)	0.05	82	-1.2(-1.8,-0.6)	<0.0001
BMI (z-score)	37	2.47(0.5)	45	2.75(0.66)	-0.28(-0.54,-0.02)	0.03	82	-0.24(-0.34,-0.13)	<0.0001
Lean body mass (kg)	23	35.7(5.9)	22	36.2(7.4)	-0.5(-4.5,3.5)	0.8	43	-0.8(-2.6,0.9)	0.3
Fat mass (kg)	23	21.8(4.5)	22	23.8(9.7)	-2.1(-6.7,2.6)	0.4	43	-2.4(-4.8,0)	0.05
Body fat (%)	23	37.9(4.8)	22	38.6(7.7)	-0.7(-4.6,3.1)	0.7	43	-1.6(-5,1.9)	0.7
Maternal BMI (kg/m2)	27	28.8(5.6)	33	29.9(6.8)	-1.1(-4.3,2.2)	0.5	60	0.4(-0.4,1.3)	0.3
Systolic blood pressure (mm Hg)	36	111.1(10.2)	45	112.5(9.0)	-1.5(-5.7,2.8)	0.5	81	-1(-6.4,4.4)	0.7
Diastolic blood pressure (mm Hg)	36	60.7(7.9)	45	64.5(7.8)	-3.9(-7.4,-0.4)	0.03	81	-3.9(-8.1,0.4)	0.07
Recovery heart rate (beats/min)	37	92(84,100)	45	108(88,136)	-	0.001	79	-20.3(-34.2,-6.3)	0.003
Physical activity (h/week)	37	14.2(8.2)	45	11.0(7.8)	3.2(-0.3,6.7)	0.07	82	3.9(0.1,7.8)	0.04
Sedentary activity (h/week)	37	15.9(7.2)	45	21.7(9.2)	-5.8(-9.5,-2.2)	0.002	82	-5.1(-9,-1.1)	0.01
Global self-esteem score (max 4)	37	3.2(0.7)	44	2.9(0.7)	0.3(0,0.6)	0.05	81	0.3(0,0.7)	0.04

Source: Sacher et al. (2010)

In 2007, MEND 7-13 extended access to the programme to many areas in England and broadened the eligibility criteria from BMI ≥ 98th centile to BMI ≥ 91st centile. The extended programme is the so-called MEND roll-out phase. The changes in health indicators before and after this phase of the intervention were compared using paired-samples t-tests. Table 3.4 illustrates the health outcomes comparisons. As shown, the health indicators of children at the end of the programme are significantly improved compared to their baselines. However, it is unclear as to whether the improvement in health during the roll-out phase actually resulted from the programme effect, due to the lack of a control group. This is especially because MEND 7-13 has expanded the coverage to overweight children, BMI between the 91st and 98th centile, without supportive evidence from the RCT.

Table 3.4: Comparing BMI, waist circumference, recovery heart rate, and time spent on physical activity and sedentary behaviours before and after the intervention (10-12 weeks) using paired-samples t-test

Variable	N	Mean*	[95% CI]		p-value
BMI	6819	-0.812	-0.789	-0.835	<0.0001
Waist circumference	6777	-2.934	-2.787	-3.080	<0.0001
Recovery heart rate	6014	-9.330	-8.840	-9.820	<0.0001
Physical activity hours	5648	3.802	4.075	3.529	<0.0001
Sedentary behaviour hours	5648	-6.159	-5.889	-6.429	<0.0001

* Mean difference = mean(after) – mean(before)

Part 2: Characteristics associated with benefit uptakes of the participants

Although some evidence supports the effectiveness of MEND 7-13 in treating childhood obesity, certain intrinsic factors may play a silent role in influencing benefit uptakes. Examples of these factors include age, gender, housing and family statuses, ethnicity, and even the programme delivery site itself. Learning the extent to which these characteristics influence programme effectiveness not only reveals their impact on health improvement, but also improves understanding of the implications of the economic evaluation in the next chapter. The purpose of this study is to explore the characteristics that influence the effectiveness of MEND 7-13, in terms of the BMI change and change in obesity status.

The objectives of the analysis are as follows: 1) to explore factors that influence the BMI change of children in MEND 7-13; and 2) to explore factors that are associated with the probability of children in MEND 7-13 transitioning from obese to non-obese.

3.3 Methodology

3.3.1 Characteristics of the MEND data

This study used the secondary data from the MEND 7-13 rollout programme. The programme participants are children between ages 7 and 13 whose BMIs matched the eligibility criteria, a BMI equal to or greater than the 91st centile of the UK 1990 reference chart for age and sex. The dataset used in the analysis was collected between January 2007 and December 2009 in England and Wales, containing 6,828 observations from 1,228 different programme sites at baseline.

The local programme delivery teams were responsible for collecting all participant data at baseline, after attending the programme (10-12 weeks), and at 6 and 12 months follow-up. The teams measured children's heights and weights and collected other health indicators, including recovery heart rate, time spent on physical activity and sedentary behaviour per week, nutrition score, and other related variables. Tables 3.5, 3.6, and 3.7 summarize the descriptive statistics of these variables. The duration of follow-up is calculated as the time from the first measurement date (1st session) to the last measurement date (last session) (MEND Central Ltd., 2010). Table 3.5 illustrates the demographic characteristics of children in the programme. As seen, the average age of participants is approximately 10-11 years old. Most participants attended almost all sessions, with a total attendance median of 16 times (max 20 times). White children were the majority among participants, while the number of females was slightly higher than males. It is also seen that more than a third of ethnicity, owning home status, and programme attendance data are missing.

Table 3.5: Descriptive statistics of data characteristics at baseline

Baseline characteristics	**MEND Rollout**	
	N	**Values***
Total observations	6,828	
Baseline age (years)	6,828	Mean: 10.4 (SD: 1.7 years) (Max 6.5; Min 14)
Duration of follow-up (weeks)	6,828	Mean: 9.9 (SD: 1.4 week) Median 10.13
Gender:	6,828	
Male		44.54%
Female		55.46%
Ethnicity (defined later in this chapter)	4,301	
White		76.66%
Asian		5.39%
Black		5.32%
Other		12.62%
Single parent status	6,828	
No		68.44%
Yes		31.56%
Owning home	4,292	
No		45.55%
Yes		54.45%
Total number of programme attendance	5,162	Median 16 (P10-90:12-19)
Programme site	6,828	Unique values = 1,228

*P10-90 = the 10th - 90th percentile.

Table 3.6 compares the obesity status of children when using the IOTF definition versus the UK 1990 definition. It is noted that the obesity cut-off of the IOTF definition is higher than that of those of the UK 1990 definition, leaving a smaller number of children defined as obese at baseline (5,073 for IOTF versus 6,062 for the UK one). Using different definitions of obesity also impacts the measurement of programme success. The effectiveness of MEND 7-13 to revert obesity in children is 15.3% using the IOTF definition, while the effectiveness drops to 9.07% when applying the UK 1990 definition. This difference emphasizes the influence of the definition applied in the analysis on programme effectiveness.

Table 3.6: Obesity status comparing between definitions - A complete case analysis* for BMI before and after the intervention

At baseline: obese or not?	After the intervention: obese or not?		Total
	Yes	**No**	
For the IOTF definition			
Yes	4,297 (84.70%)	**776 (15.30%)**	5,073
No	32 (1.83%)	1,714 (98.17%)	1,746
Total	4,329	2,490	6,819
For the UK 1990 definition			
Yes	5,512 (90.93%)	**550 (9.07%)**	6,062
No	30 (3.96%)	727 (96.04%)	757
Total	5,542	1,227	6,819

* A complete case analysis includes only children whose BMI data at pre- and post-measurements were collected. In this case, 9 observations did not have BMI data and thus were excluded.

Table 3.7 shows health indicators at baseline in comparison to those after the intervention. As demonstrated, the data indicate an improvement in all health parameters, even though the degrees of change vary across variables. In this regard, missing values are found in many variables. A number of variables contain more than 1,000 missing values – up to almost 4,000 – including time spent in physical activity and sedentary behaviours, nutrition score, psychological symptoms, body esteem scale, and self-esteem scale. A later section in this chapter explores the follow up process, which helps explain the source of the missing values.

Table 3.7: Descriptive statistics of the MEND data comparing pre-post measurement

Characteristics	Baseline		After the intervention	
	N	Values*	N	Values*
Total number of observations	6,828		6,828	
BMI (kg/m2)	6,828	Mean: 27.4 (SD: 4.5)	6,819	Mean: 26.6 (SD: 4.6)
Waist circumference (cm)	6,797	Mean: 86.6 (SD: 14.8)	6,790	Mean: 83.8 (SD: 12.2)
Obesity status by gender (IOTF definition)	6,828		6,819	
Male	3,041	74.19%	3,038	63.03%
Female	3,787	74.60%	3,781	63.85%
Fitness heart rate test (recovery heart rate)(beats/min)	6,234	Mean: 110.3 (SD:22.5) (P10-90:84-139)	6,198	Mean: 100.9 (SD:21.1) (P10-90:77-128)
Time spent in physical activity per week (hrs/wk)	6,657	Mean: 11.4 (SD:9.1) (P10-90:3.5-21)	5,687	Mean: 15.2 (SD:9.8) (P10-90:6-26.25)
Time spent in sedentary behaviours (hrs/wk)	6,657	Mean: 16.6 (SD:11) (P10-90:4-30)	5,687	Mean: 10.7 (SD:7.6) (P10-90:3-20)
Nutrition score	6,650	P50 = 17 (P10-90:11-22)	5,771	P50 = 24 (P10-90:18-28)
Psychological symptom (SDS score)	6,619	P50 = 13 (P10-90:5-23)	5,703	P50 = 9 (P10-90:3-19)
Body esteem scale	5,841	P50 = 8 (P10-90:3-18)	4,876	P50 = 14 (P10-90:4-22)
Rosenberg Self-Esteem Scale (adapted)	3,383	P50 = 16 (P10-90:7-26)	2,952	P50 = 20 (P10-90:10-29)
Using the IOTF standard BMI cut-offs				
Percent of obese children	5,081	74.41%	4,329	63.40%
Using the UK 98th BMI centile cut-offs				
Number of obese children (%)	6,070	88.90%	5,542	81.17%

***P10-90 = the 10th - 90th percentile; P50 = the 50th percentile, or median.**

Figures 3.1 to 3.7 show the distributions of BMI, waist circumference, and recovery heart rate at pre- and post-measurements. From the figures, decreases in BMI, waist circumference, and recovery heart rate can been seen after the intervention.

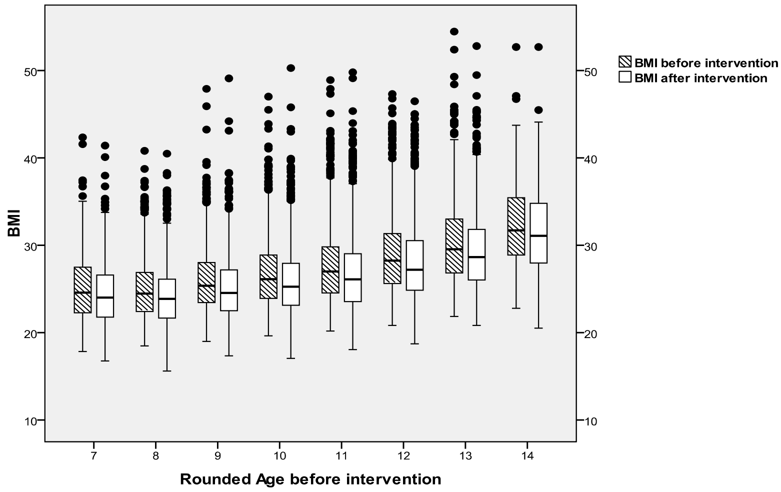


Figure 3.1: Box plots showing BMI distributions before and after the intervention by age

*The rounded age is to the nearest 0.5

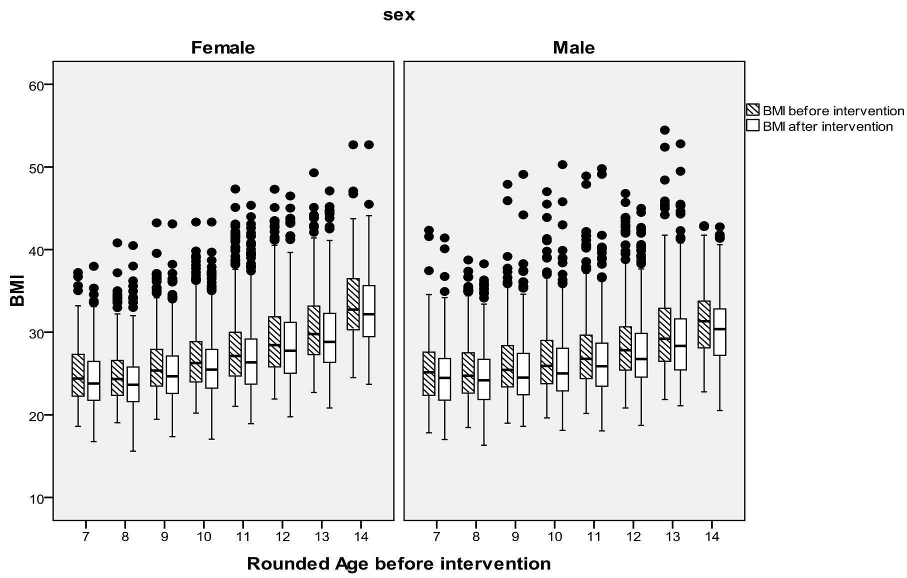


Figure 3.2: Box plots showing BMI distributions before and after the intervention by age and gender

*The rounded age is to the nearest 0.5

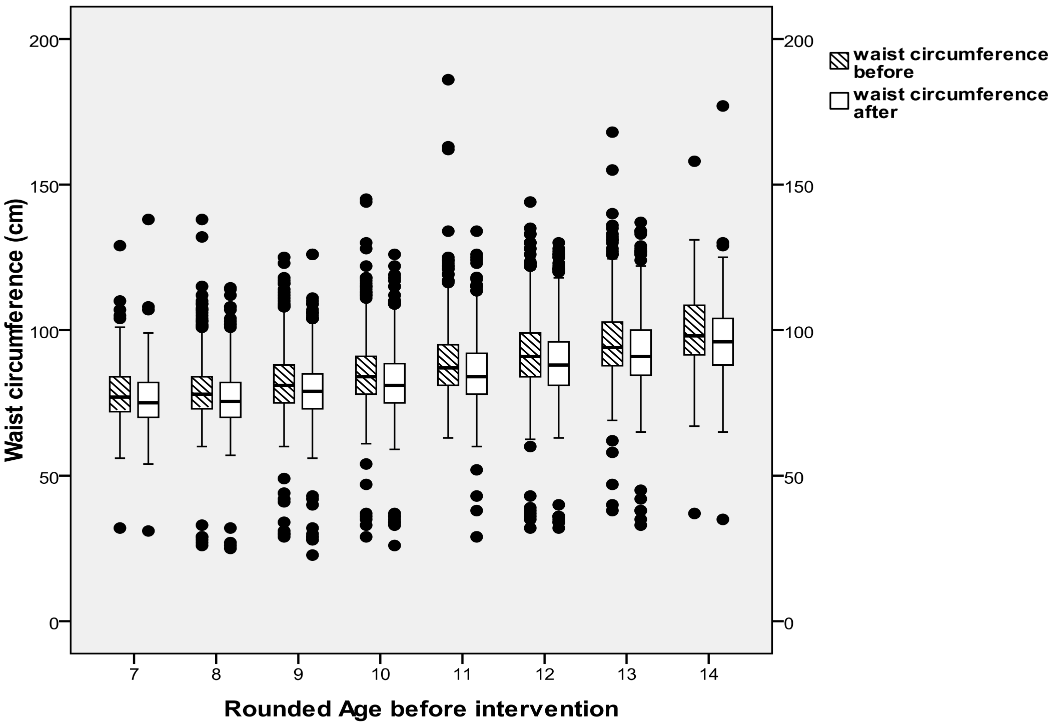


Figure 3.3: Box plots showing waist circumference distribution before and after the intervention by age

*The rounded age is to the nearest 0.5

**An observation with waist circumference of 728 cm was considered an outlier, possibly due to miscoding. As it tends to influence the analysis, this observation was excluded before plotting the graph.

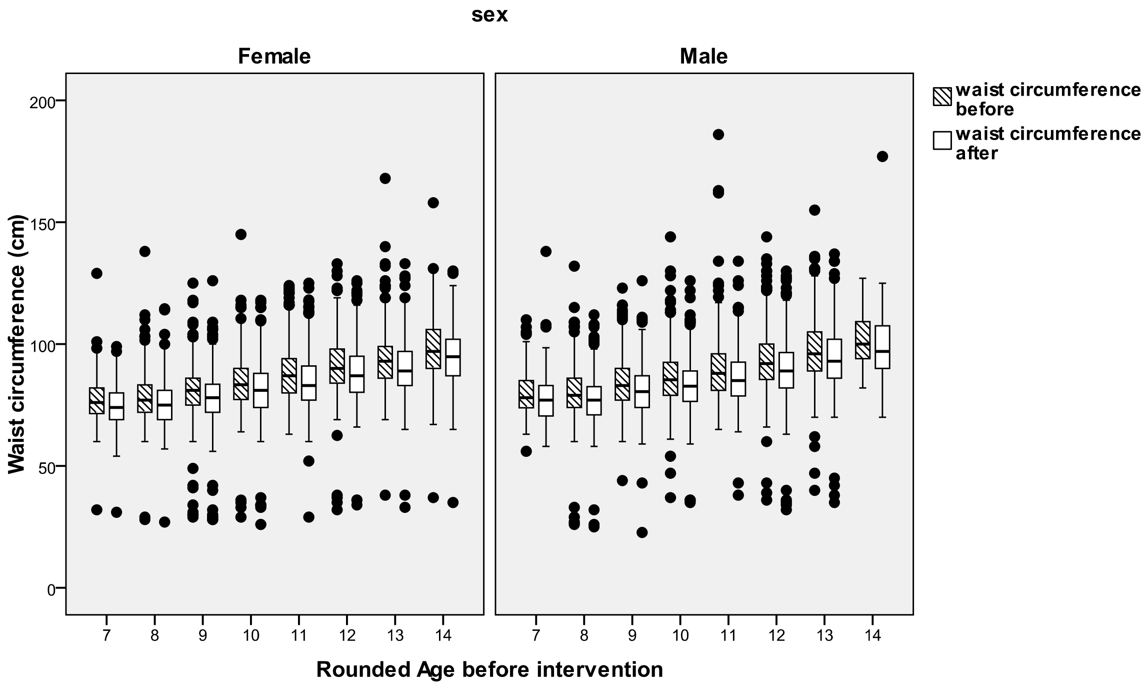


Figure 3.4: Box plots showing waist circumference distribution before and after the intervention by age and gender

*The rounded age is to the nearest 0.5

**An observation with waist circumference of 728 cm was considered an outlier, possibly due to miscoding. As it tends to influence the analysis, this observation was excluded before plotting the graph.

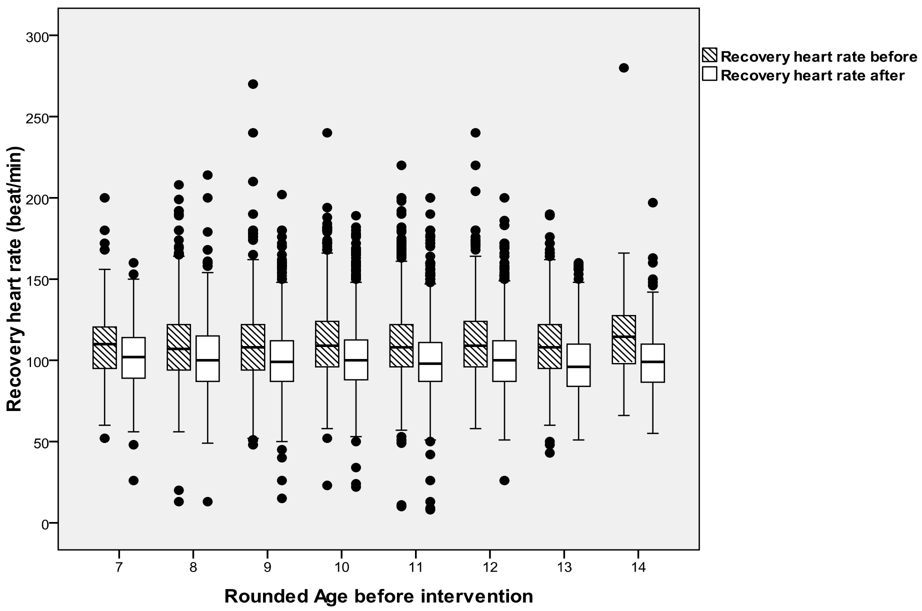


Figure 3.5: Box plots showing recovery heart rate distribution before and after the intervention by age

*The rounded age is to the nearest 0.5

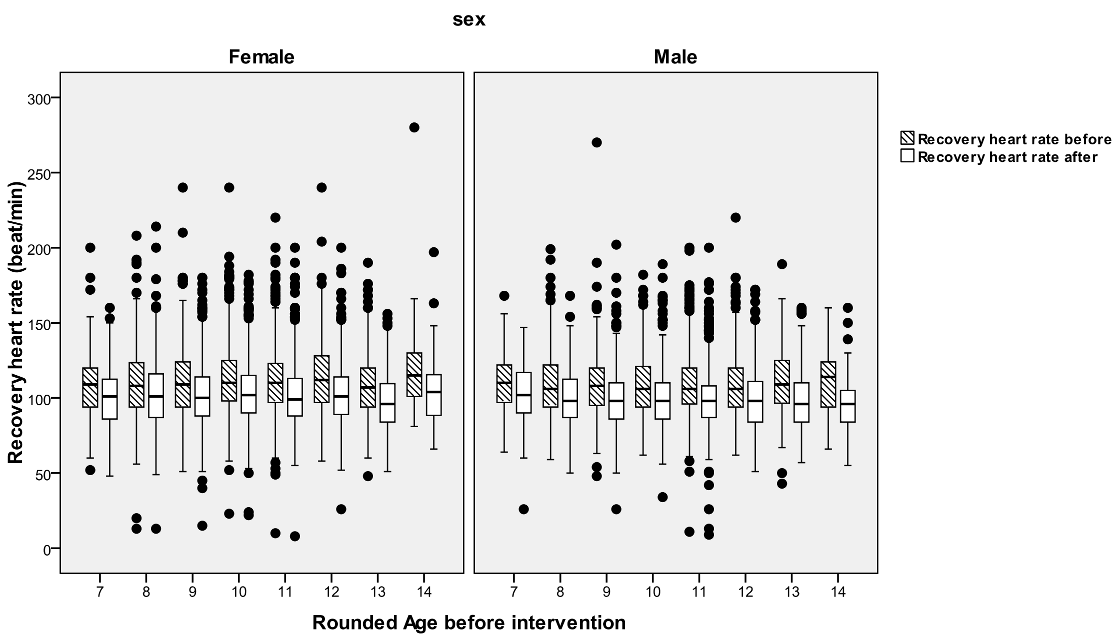


Figure 3.6: Box plots showing recovery heart rate distribution before and after the intervention by age and gender

*The rounded age is to the nearest 0.5

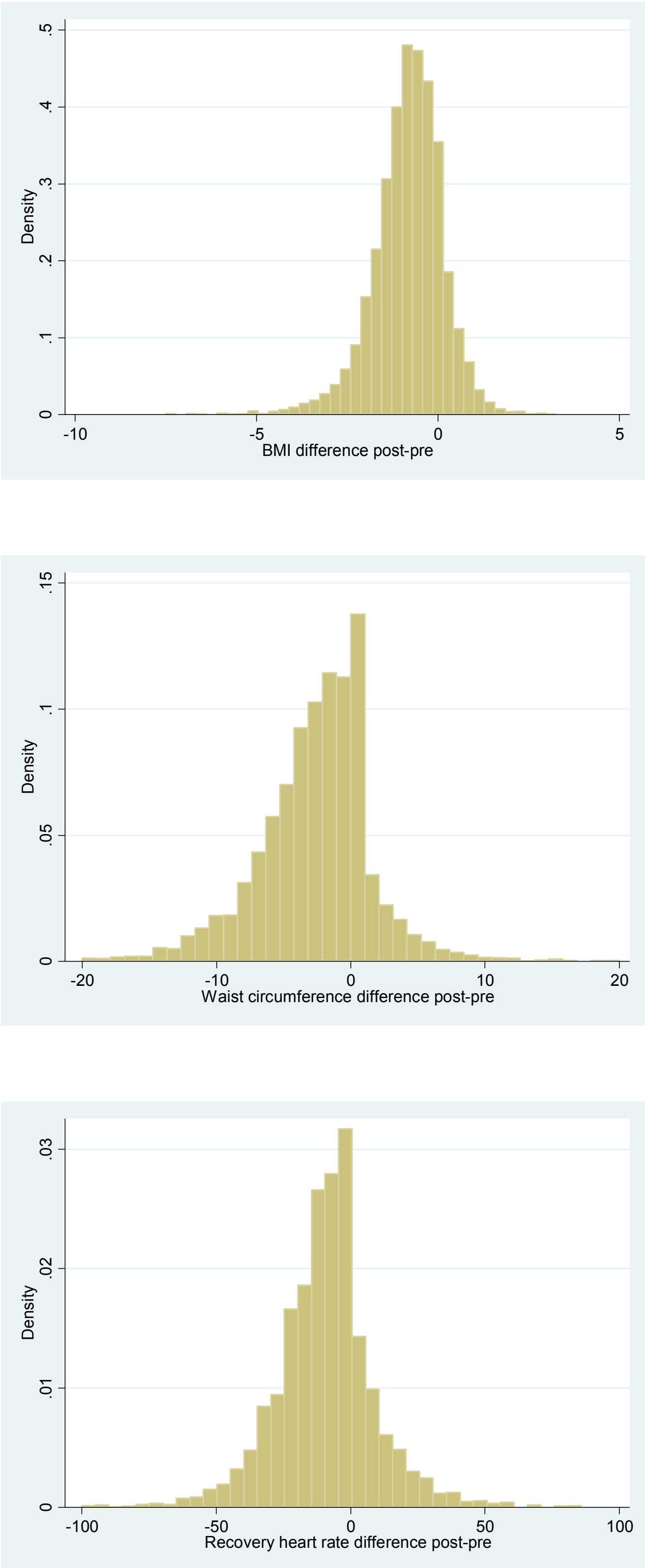


Figure 3.7: Histograms of differences in BMI, waist circumference, and heart rate pre-post

3.3.2 Health measurement

MEND Central provides local programme delivery teams with detailed recommendations for measuring children's health outcomes. All health indicators should be measured during in-person visits, not by interview. Details of the measurement techniques for each health parameter were given; for example, children should take off shoes before measuring weight and height, and waist circumference is measured at 2 finger widths above the umbilicus. Weight and height are measured in kilogram and centimetre units respectively. BMI is calculated by (weight in kg)/(height in meter)^2.

To evaluate cardiovascular fitness, the YMCA Step Test or Progressive Pulse Ratio Test is applied. The YMCA Step Test is a validated measure that MEND uses to assess the baseline and progress of children's cardiovascular fitness throughout the programme (MEND Central Ltd., 2010). Children are asked to step up and down on a 20-30cm-height step, depending on age, at a rate of 24 steps per minute for 3 minutes. After finishing the last step, children have to sit down immediately and remain still. The heart beat is then counted for one full minute. The number of post-exercise heartbeats in one minute, or the recovery heart rate, is the score of the test. The lower the score, the fitter the subject's cardiovascular status is.

3.3.3 Follow-up process

The local programme delivery teams have full authority to select the time and methods used to follow up children. Normally, the location of the programme site tends to be at nearby schools, and the sessions usually start immediately after school in order to accommodate parents picking up their children. Regarding the follow-up process, MEND Central advises that all measurements taken pre- and post-intervention be repeated at follow-up. However, height and weight seem to be the only consistent measurements taken when children come to follow up, especially at 6 and 12 months. Other health indicators are mostly missing at this stage.

Another problem concerning the follow-up is the recalling method. MEND Central has no data about the frequency of the follow-up, the method used to contact children and parents (phone, email, or post), the number of attempts made to contact participants, the follow-up location, the persons who conduct the follow-ups, or the health indicators collected during the follow-up sessions. All these activities are dependent on the local staff, i.e. whether or not they follow up children, and if so, how they do it. Moreover, activities held by local staff during the follow-up sessions vary considerably across the programme sites. Some centres run specific MEND follow-up sessions that incorporate refreshers of MEND topics, while others only offer activity sessions or review the lessons without physical activities (MEND Central Ltd., 2010). One possible reason for inconsistency in follow-up is that MEND Central does not require the local programme delivery teams to report a follow-up session or health parameters recorded during the follow-up, even though 6- and 12-month follow-ups are specified in the programme regulations. Another possibility is that children and parents may move out of the area, making it difficult to contact them for the follow-up sessions.

The local delivery teams are responsible for finding a location for the programme centre. The choice of locations varies across centres, depending upon availability, budget, special deals or discounts the teams might receive from the local authorities (e.g. local gyms or swimming pools), and the teams' decisions. As the local delivery teams are not obliged to report where the programme is held, information about the location of the programme site is missing from the data (MEND Central Ltd., 2010). The only area variable presented in the data is the children's home postcode, which is collected at the time of application.

3.3.4 Comparing socio-demographic characteristics of MEND 7-13 to HSE 2008

To determine whether the samples in the MEND data are representative of overweight children in the English population, data from Health Survey for England (HSE) 2008 were used. The comparison was performed in 2 groups: children with BMI $\geq 91^{st}$ centile from the UK 1990 BMI chart, and those categorized as obese according to the IOTF definition. The socio-demographic distribution of the subjects in the HSE data was compared to that of obese children in MEND, controlling for age and gender. This study uses a complete case analysis to compare subjects' characteristics between MEND and

HSE2008; the comparison includes only cases for which all measurements of interest were recorded.

Table 3.8 shows the number of obese children from each dataset included in the complete case comparison. As illustrated, only the MEND data contained missing values, especially in the cases of ethnicity and level of deprivation.

Table 3.8: Comparing the numbers of cases in MEND and HSE2008 at baseline in a complete case analysis

Characteristics	N of children with BMI ≥91st centile		N of obese children (IOTF definition)	
	MEND	**HSE2008**	**MEND**	**HSE2008**
Total obese children	6,828	936	5,081	259
Gender	6,828	936	5,081	259
Age	6,828	936	5,081	259
Ethnicity	4,301	936	3,233	259
Level of deprivation	6,239	936	4,620	259

Note: The number of cases in MEND data was counted at baseline, while the BMI data in 9 cases were missing at follow-up.

Figures 3.8-3.13 show the socio-demographic distributions of the samples in the MEND and HSE 2008 data. As the figures illustrate, small differences can be seen between samples when observing gender and ethnicity distributions. However, differences in the distributions between samples are noticeable. MEND 7-13 has a high proportion of children aged 9-11 in the programme compared with the HSE 2008 distribution. In addition, while the distributions of the level of deprivation appear similar for all samples, obese children in MEND 7-13 came from less deprived areas compared to those in HSE 2008.

To examine these between-sample differences in socio-demographic distributions, the Chi-square test was carried out. For children with BMI ≥91st centile, evidence shows that there are significant differences in some socio-demographic distributions, including age and ethnicity, between MEND and HSE 2008 (p-values are <0.001, and 0.004, respectively). However, there is no evidence indicating a difference in gender (p-value = 0.121) and level of deprivation (p-value = 0.266) between 2 samples. A conclusion that can be inferred here is that the programme recruitment process might not successfully target children from deprived areas, though they are more likely to have low socioeconomic status and be more prone to obesity. When comparing only the obese children with respect to the IOTF definition between samples, the evidence indicates a difference in gender, age, and level of deprivation (p-values are 0.028, <0.001, and 0.033, respectively), while little evidence is seen for differences in ethnicity (p-value = 0.09).

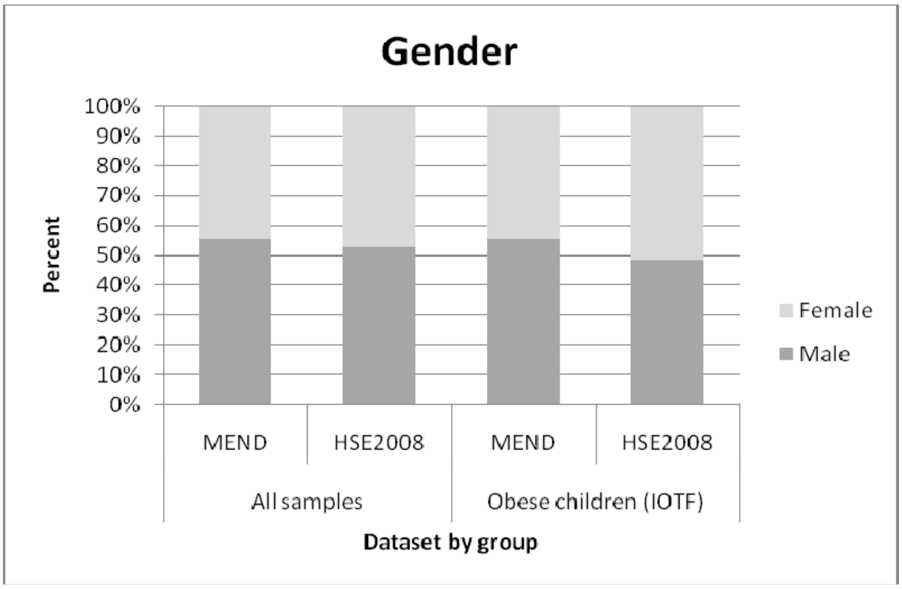
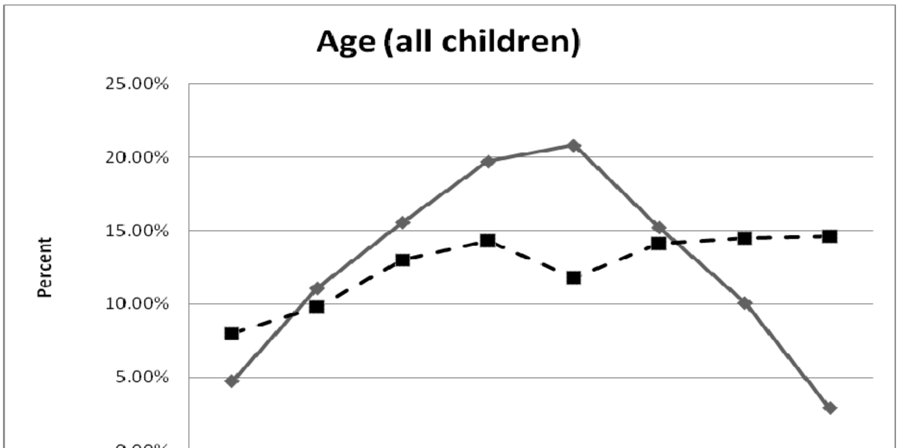


Figure 3.8: Comparing gender of samples between MEND and HSE2008



	7	8	9	10	11	12	13	14
MEND	4.73%	11.06%	15.55%	19.70%	20.80%	15.20%	10.06%	2.90%
HSE2008	7.98%	9.76%	12.98%	14.31%	11.78%	14.10%	14.49%	14.60%

Figure 3.9: Comparing ages of children with BMI ≥91st centile between MEND and HSE2008

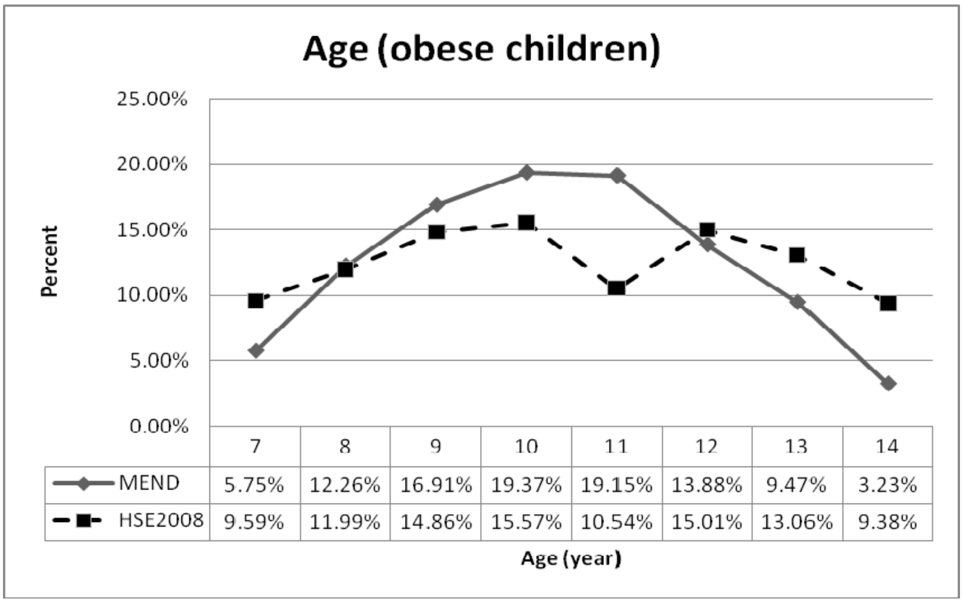


Figure 3.10: Comparing ages of obese children (IOTF definition) between MEND and HSE2008

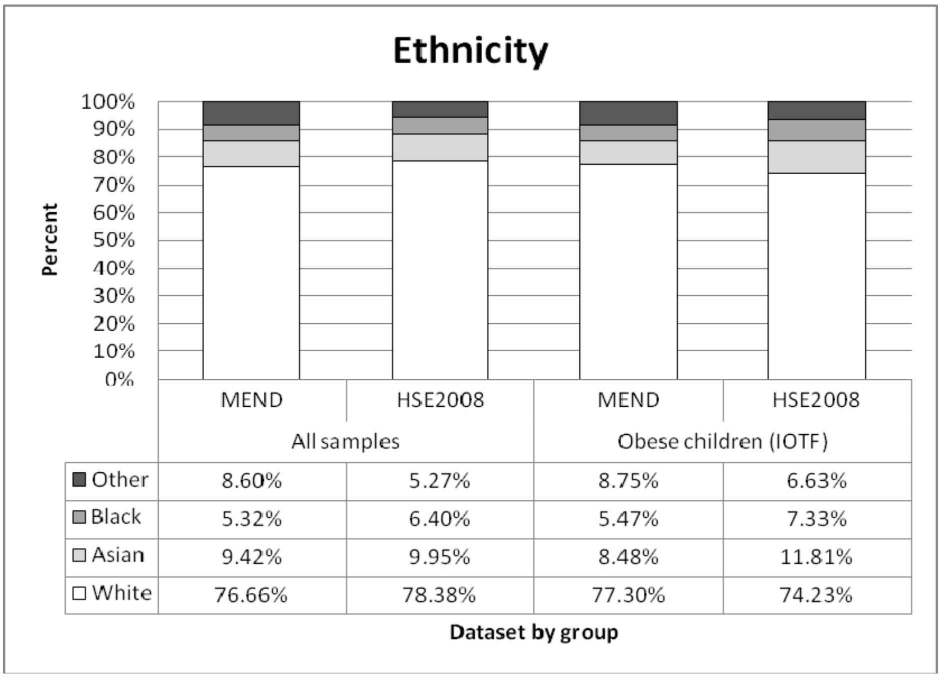


Figure 3.11: Comparing ethnicity of samples between MEND and HSE2008

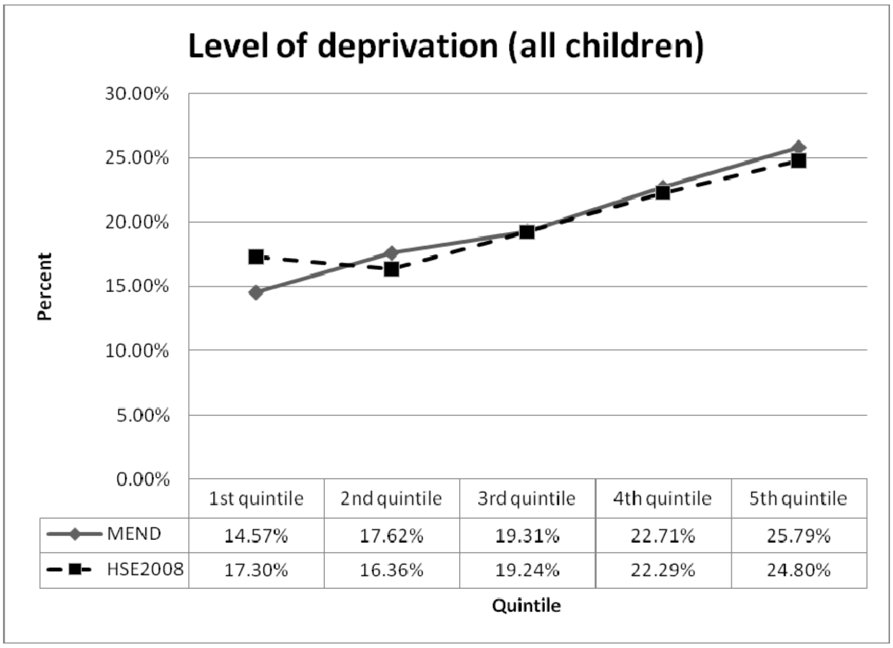


Figure 3.12: Comparing level of deprivation of children with BMI ≥91st centile between MEND and HSE2008 (1st quintile = least deprived area, 5th quintile = most deprived area)

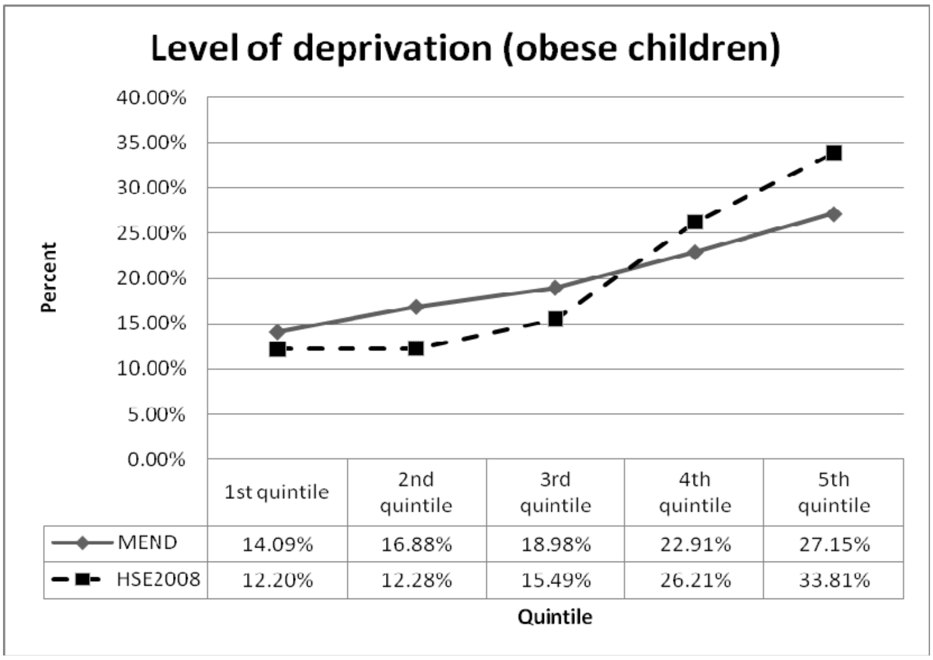


Figure 3.13: Comparing level of deprivation of obese children (IOTF definition) between MEND and HSE2008 (1st quintile = least deprived area, 5th quintile = most deprived area)

Differences in the distributions of some baseline characteristics may reflect selection biases during the recruitment process. One source of selection bias is the enrolment method, which is only available through the MEND website. Statistics show that approximately 10 million adults in the UK had never accessed the internet in 2009 (the data were collected in 2007-2009) (Office for National Statistics, 2011b). As online registration is the only programme registration method, these 10 million adults would never be able to register any obese children they know, including perhaps their own children, to the programme. Moreover, internet access is associated with education level, which subsequently links to socio-economic status. Supportive evidence reveals a higher proportion of internet access among educated adults in comparison with those without any formal qualification, 97% versus 45%, respectively (Office for National Statistics, 2011b). Another source of selection bias is the availability of the programme site. Although more than a thousand programme sites have been made available throughout England and Wales during the rollout phase, there are many more areas that have not had access to MEND 7-13. Inquiry about specific programme site locations is also limited owing to the unavailability of data about programme location, as stated earlier. These biases should be taken into consideration when interpreting the results.

3.3.5 Justification for the selection of factors in the analysis

Chapter 1 presents a number of factors associated with the development of childhood obesity, including age, gender and ethnicity. These factors were thus incorporated into the regression analysis to assess their impact on the effectiveness of MEND 7-13. In addition, other proxies may also play an important role in benefit uptake, such as having single parents or living in their own houses. The number of programme sessions attended is another factor that seems directly related to effectiveness, by allowing participants more time to learn from the programme. Additionally, differences in the programme centres may also affect the benefit uptakes of children, owing to the particular characteristics of each centre, such as environments that encourage learning ability and local staff's skills in delivering messages and creating an atmosphere conductive to learning. This chapter explores the association of programme effectiveness with all these features by including them into the regression model as explanatory variables.

3.3.6 Constructing new variables

Some variables were constructed or recoded based on the existing ones, for use in the regression analysis. This section details the construction of the new variables.

3.3.6.1 Ethnicity

The original dataset of MEND 7-13 contains a variable, 'Ethnicity', which displays the ethnic groups of children. The variable is split into 17 categories, with 2,527 missing values. These categories include: 1) African; 2) Bangladeshi; 3) British; 4) Caribbean; 5) Chinese; 6) Indian; 7) Pakistani; 8) Irish; 9) White; 10) Mixed White and Asian; 11) Mixed White and Black African; 12) Mixed White and Black Caribbean; 13) Any other Black Background; 14) Any other Asian background; 15) Any other White background; 16) Other than Mixed background; and 17) Other Ethnic Groups.

Due to small sample sizes in most ethnic groups (10 of 17 categories have fewer than 70 observations), a new variable, 'ethnic1', was constructed. The seventeen categories in the 'Ethnicity' variable were re-categorized into 4 broader ethnic origins in order to create fewer categories with more samples in each. The new variable 'ethnic1' consisted of the 'White', 'Asian', 'Black', and 'Other' categories. It is noted that there is no straightforward way to do this. Recoding ethnicity into groups was based on standard categorization and supportive evidence regarding differences in the prevalence of childhood obesity among different ethnic groups. The criteria used to combine the old categories into the new ones are as shown in Table 3.9.

Table 3.9: recoded ethnic variable into 4 categories

New variable	Categories in the old variable (Ethnicity) that incorporate into the recoded variable	Frequency
White	Any other White background, British, Irish, White	3,297
Asian	Any other Asian background, Bangladeshi, Chinese, Pakistani, Indian	405
Black	African, Any other Black Background	229
Other	Caribbean, Other Ethnic Groups, Other than Mixed background, Mixed White and Asian, Mixed White and Black African, Mixed White and Black Caribbean	370

3.3.6.2 Postcode and the level of deprivation

As the only available area variable is the postcode of children's houses, the effect of the area variable applied in the regression therefore comes from the effect of children's housing areas, not the area of the programme site.

Using postcodes directly in the regression analysis raises 2 concerns: 1) the majority of postcodes are unique for each participant (5,942 unique values out of 6,828 observations), which leads to the potential problem of over-fitting and thus decreases the generalizability of the regression model; and 2) using the postcode directly does not give much meaningful information when interpreting the results. Thus, this study employs the deprivation level of the household area instead, using the home postcode to make a credible estimate.

To convert the postcode into the level of deprivation, an online geography matching and conversion tool, GeoConvert, was employed (Census Dissemination Unit, 2011). Postcodes in the dataset were first matched to The National Statistics Postcode Directory (NSPD), which provides all the current and terminated postcodes that have ever existed in the UK (ESRC Census Programme, 2011). The latest edition of the NSPD when conducting the research, in February 2010, was applied in the conversion. With regard to this process, 246 observations did not match the postcode standard and were thus excluded. The author investigated these unmatched cases and found that miscoding is the main reason for this.

After retrieving the matched postcodes, the next step was to convert them into the level of deprivation using GeoConvert. In this regard, the study chose the Index of Multiple Deprivation 2007 (IMD 2007) to represent the level of deprivation in the analysis. The IMD 2007 measures the level of deprivation by ranking scores based on small-area geography, so-called the Lower Super Output Area (Communities and Local Government, 2011b). It considers 7 domains: 1) Income deprivation; 2) Employment deprivation; 3) Health deprivation and disability; 4) Education, skills, and training deprivation; 5) Barriers to housing and services; 6) Crime; and 7) Living environment deprivation (Communities and Local Government, 2011a). Postcodes corresponding to more deprived areas receive higher IMD 2007 scores and lower ranks.

The results from the IMD conversion showed that 343 observations were not matched with the IMD 2007, making the total number of participants with completed deprivation scores 6,239. The first possible reason for this mismatch is that the deprivation score is only available for postcodes in England, while MEND 7-13 was rolled out in both England and Wales. Hence, children whose addresses are in Wales were excluded for this reason. The second possible reason is that the deprivation scores are not available for all English postcodes, leaving certain postcodes unmatched even when the addresses are in England.

The IMD 2007 is a ranking rather than an actual measure, which is inconvenient when interpreting the results of the analysis; a conventional way to deal with this problem is to categorize the IMD ranking by quintile. A new variable, imd_q, was created based on the quintiles of the IMD 2007, where the first quintile represents the least deprived areas and the last, or fifth, quintile refers to the most deprived areas.

3.3.6.3 Differences in BMI

The variable showing the difference in BMI before and after the intervention was constructed. The difference is calculated by the value at post-intervention less that at pre-intervention.

$$BMI\ difference = BMI\ post - BMI\ pre$$

A negative measure means that children had reduced weight at follow-up, which determines the success of the programme. A positive measure implies the opposite.

3.3.6.4 Obesity status

The obesity status variable was constructed and used as a proxy of the health outcome. As stated in Chapter 1, a number of definitions using different BMI cut points have been proposed to define childhood obesity, e.g. using the BMI centile cut points (at the 91^{st}, 95^{th}, or 98^{th} centiles) for specific populations, as well as the international (IOTF) definition, which created international BMI cut-offs using cross-sectional data from 6 countries (Brazil, Great Britain, Hong Kong, the Netherlands, Singapore, and the United States).

Although MEND 7-13 recruited participants into the rollout programme based on the 91^{st} BMI centile from UK 1990 BMI chart, applying the IOTF definition better fits the context of this thesis. As one purpose of this chapter is to inform the economic evaluation in the next chapter, consistency in the definition between two chapters is deemed important. Concerning the economic evaluation, using the UK definition hinders comparison of the cost-effectiveness results, which are usually compared with those from other studies both inside and outside the UK; it is unlikely for studies outside the UK to use the UK definition. As a result, assigning the IOTF cut-offs provides more meaningful results for the whole thesis.

The IOTF definition was applied to the BMIs of children both before and after the intervention. Based on children's obesity status prior to and after the intervention, four categories were created. The details of the categories are shown in Table 3.10. Nine observations were excluded due to missing values, which were due to missing BMIs at follow-up.

Table 3.10: Obesity status

Obesity status	Obese or not (IOTF definition)		Frequency
	Pre-intervention	Post-intervention	
0	Yes	Yes	4,297
1	Yes	No	776
2	No	Yes	32
3	No	No	1,714

Since one purpose of the analysis is to explore the effects of factors associated with obesity reversion, only children who were obese at baseline were counted. A binary variable, 'obstat1', containing only the first 2 categories of the obesity status in Table 3.10, was created, with 2 categories: Obesity reverted after the programme or not (No – 4,297, Yes – 776).

It is noted that although the statistical results in this chapter, shown in section 3.4.2, were all based on the IOTF definition, the author realizes the impact of the definition applied and thus conducted the analysis using the UK definition as well. However, after testing the goodness of fit in the logistic regression, the models from both the IOTF and UK 1990 definition do not well represent the MEND dataset. Details of this are mentioned later, in section 3.4.2.

3.3.6.5 Duration of follow-up

The follow-up duration is calculated as the time from the first measurement date (first session) to the post-intervention measurement date (last session). The data provide the duration of follow-up in months. This study converted the time scale of the duration into weeks by multiplying time in months by 4.33. Table 3.11 shows the relationship between the duration of programme attendance and attendance frequency. As seen, the mean follow-up duration is approximately 10 weeks regardless of the total number of programme sessions attended. One possible explanation of this is that most children might attend the first and last few sessions of the programme while missing some sessions in the middle. The correlation between change in BMI and the duration of follow-up, presented as 'periodwk' in Table 3.15, is tested in the later part of this chapter, with results showing that there is no association between these factors. Therefore, the duration of follow-up is not included in the regression model.

Table 3.11: Means of duration in the programme (from the 1st to last session) with respect to the total number of programmes attended

Total attendance	Duration from 1st to last session (wk)		Percent (of 5,162 non-missing values)
	Means	SD	
5	10.05	1.47	0.3%
6	10.32	1.97	0.9%
7	10.33	1.67	0.6%
8	10.30	2.38	1.0%
9	10.12	1.45	1.6%
10	9.93	1.98	1.5%
11	9.98	1.34	2.2%
12	9.81	1.64	3.4%
13	9.92	1.36	4.9%
14	9.90	1.58	8.1%
15	10.01	1.45	10.7%
16	9.97	1.40	15.2%
17	9.92	1.25	16.3%
18	9.93	1.29	19.1%
19	9.90	0.99	6.6%
20	9.75	1.04	7.5%

3.3.7 Missing data and other problems

The MEND data suffer from numerous missing values. Table 3.12 shows the number of missing values for some variables in the dataset. As seen, a great amount of data was missing at 6 and 12 months follow-up. Compared with the baseline, less than 7% and 1% of overall data still remain at 6 and 12 months follow-up respectively.

Table 3.12: Missing values in the MEND 7-13 data

Variable names	Missing values at a different time (of 6,828)			
	baseline	3 month	6 month	12 month
Gender	0			
Postcode	1			
Area-specific deprivation score	589			
Total attendance		1,666		
Programme site	0			
Ethnicity	2,527			
Singleparent	2,557			
Ownhome	2,536			
Age	0	0	6,365	6,764
Height	0	6	6,365	6,764
Weight	0	6	6,365	6,764
BMI	0	9	6,365	6,764
Waist circumference	31	38	6,370	6,764
Heart rate	594	630	6,538	6,785
Physical activity hours	171	1,141	6,731	6,821
Nutrition score	178	1,057	6,736	6,819

One concern with regard to the missing data is that it does not occur at random, but rather varies according to the programme site. Figure 3.14 demonstrates the patterns of missing values in 3 variables—singleparent, ethnicity, and ownhome—with respect to the programme centre (1,228 sites). As can be seen, most centres tend to either collect or miss all the data, whereas only some record partial data. The figure at the right lower corner shows the number of centres that did not collect any of the 3 variables from participants (left column), compared to those that collected all the data from all participants (right column) and those that collected incomplete data (middle column). A consistent pattern of missing and completed data can be noticed here; that is, more than 350 centres did not collect any of these 3 variables, while almost 500 centres did complete the data collection in its entirety.

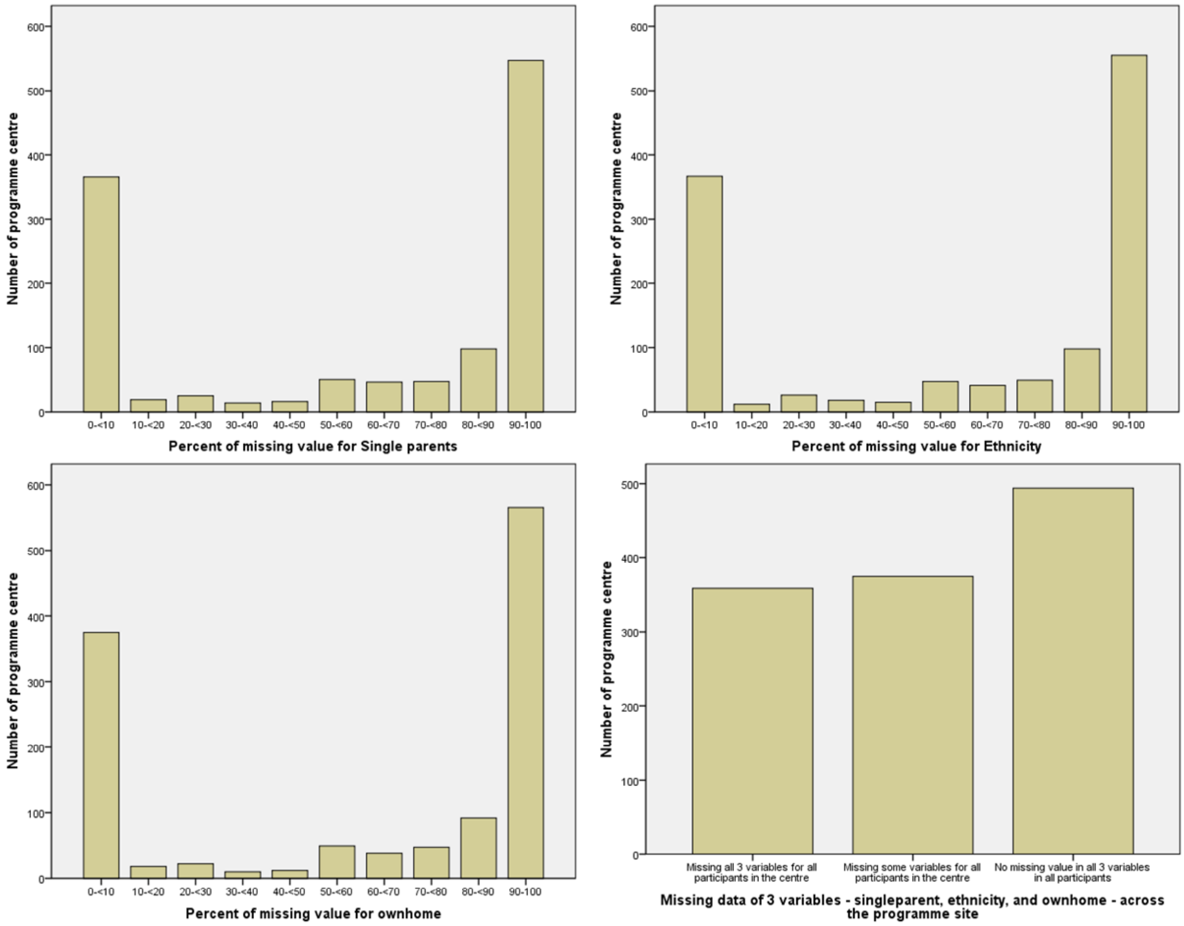


Figure 3.14 Tendency of the outcome data being missed at 10-week follow-up by the programme centre (1,228 sites) when considering 3 variables – single parent, ethnicity, and own home

It is also noted that some parameters, such as age and health indicators, were measured over time, whereas demographic variables were collected only once at baseline.

Due to a massive amount of missing data at follow-up, particularly at 6 and 12 months, it is impossible to obtain any meaningful results from analysing the data at these checkpoints. Therefore, the analysis focused on the changes in health outcomes at 2-3 months after attending the programme to avoid biases in the outcomes and increase the accuracy and reliability of the analysis.

Missing data often causes problems during analysis not only because it leads to possible bias in estimates, but also because the analysis loses power due to the omission of subjects with missing data. A number of methods have been employed to deal with missing data, such as complete-case analysis, single imputation, and multiple imputation. Introduced in the 1970's, multiple imputation (MI) is a technique that has been increasingly used in recent decades (Rubin, 1976, StataCorp, 2009).

The purpose of conducting the multiple imputation analysis in this chapter is to account for uncertainty from the missing data; that is, to observe whether the results from the MI analysis are different from those from the conventional analysis. In order to do so, there are 3 steps: imputation step, completed-data analysis, and pooling step (StataCorp, 2009). The first step aims at filling missing values in the dataset with the values generated from the imputation model. The imputation models are chosen with respect to the distribution of the imputed variables, alongside other variables required in the models as predictors. During this step, M complete datasets are created and ready to be used in the analysis in the next step (the analyses in this chapter create 20 imputations). After imputing all missing variables in M datasets, the next step is to perform the statistical analysis on each imputed dataset separately. Then, in the last step, results from all datasets are combined into a single multiple-imputation result (StataCorp, 2009).

Multiple imputation with chained equations (MICE), a specific type of imputation, is used in this chapter to assess the impact of the missing values in the MEND data. MICE imputes missing values using an iterative multivariable regression technique called switching regression. Some advantages of using MICE, in relation to multivariate normal approach, are: 1) it does not require the assumption of multivariate joint distribution; that is, the joint distribution of all variables used in the imputation step is not assumed to be normal, as is in MI; and 2) the sample size required is smaller (2011).

In addition to the missing data, there is also an issue regarding re-registration to the program that affects the quality of this dataset. It is evident that MEND 7-13 does not encourage children who graduated from the programme once to register again. Nevertheless, it is not easy for MEND to check if any children repeat the programme, due to technical issues with the database system. As far as the MEND staff could detect, a small number of the MEND graduates—approximately 80 cases—re-registered for the programme (MEND Central Ltd., 2010). Of these, some were removed from the dataset for both visits, some kept records for either the first or the second visit (selected at random), and some kept both the first and the second registers for the same children (MEND Central

Ltd., 2010). The author has decided to use the MEND data as sent from MEND Central and treats all children in the dataset as separate individuals. Errors in the dataset resulting from having children who attended the programme twice are admitted as a minor weakness of the analysis.

3.3.8 Labelling and coding variables

Variable names sometimes use abbreviations so as to make them fit the available space. The descriptions of the variable names used in the analysis are shown in Table 3.13 to clarify these abbreviations. Furthermore, as seen above, some categorical variables were re-categorized and re-labelled before being incorporated into the analysis. Table 3.14 lists the labels of the variables used in the analysis.

Table 3.13: Descriptions of the variable names used in the analysis

Variable names	Descriptions
bmidif	The BMI change of children during programme attendance
obstat1	Obesity status at post-measurement (last session) in children who were obese at baseline (IOTF definition)
sex	Gender
age1	Age at baseline
singleparent	Children having single parent or not
ownhome	Children living with parents who own their houses or not
ethnic1	Ethnic groups
bmipre	Baseline BMI
TotalAttendance	A total number of programme attendance
imd_q	Quintile of the deprivation level

Table 3.14: Labels of some categorical variables

Variable (var name)	Coding	
Obesity status	0*	Obese before and after intervention
(obstat1)	1	Obese before/ not obese after
Level of deprivation (quintile)	0*	1st quintile (least deprived)
(imd_q)	1	2nd quintile
	2	3rd quintile
	3	4th quintile
	4	5th quintile (most deprived)
Gender	1*	Female
(sex)	2	Male
Ethnicity	1*	White
(ethnic1)	2	Asian
	3	Black
	4	Others
Single parent or not	0*	No
(singleparent)	1	Yes
Own home or not	0*	No
(ownhome)	1	Yes

*Taken as baseline in the regression model

3.3.9 Objectives and methods of the analysis

Objective 1: To explore factors that influence the BMI change of children in MEND 7-13

Main outcome variable

The outcome variable is the BMI change

Explanatory variables

The independent variables employed in the regression analysis are the following:

1. Baseline BMI
2. Gender (1=Female; 2=Male)
3. Age (age at baseline was applied in the regression)
4. Ethnicity (1=White; 2=Asian; 3=Black; 4=Other)
5. Single parent status or not (0=No; 1=Yes)

6. Living in their own houses or not (0=No; 1=Yes)
7. The total number of programme attendance
8. Level of deprivation (quintile)
9. Programme site (applied only in the multilevel analysis)

Statistical analysis

First, a simple linear regression was used to assess the influence of associated factors on the BMI change of children. Owing to the difference in the BMI change across ethnicities shown in the simple linear regression, the author conducted the analysis by ethnic stratification. The diagnostic tests were carried out, including testing the multicollinearity, normality, heteroscedasticity, and independence of observation.

To address the concern regarding the extent to which the missing data affect the results of the analysis, imputation analyses – specifically, multiple imputation with chained equations (MICE) – were carried out, with 20 imputations produced.

With regard to programme delivery, it is possible that the degree of success for each programme site may vary due to its particular characteristics. To account for the effect of the different programme sites on success, a multilevel linear regression was applied. It is noted from the data that a huge number of 1,228 unique programme sites corresponding to participants in the dataset was used in the analysis. Therefore, a random effects analysis was employed in the multilevel analysis when accounting for the programme site. A significance level of α=0.05 was applied to determine the strength of evidence.

Objective 2: To explore factors that are associated with the probability of children in MEND 7-13 transitioning from obese to non-obese

Main outcome variable

The obesity status before and after intervention was used as an outcome variable in the regression model. With regard to the use of obesity status, only obese children (in accordance with the IOTF definition) were included in the analysis, making the outcome variable binary (obese at follow-up or not: Yes/No).

Explanatory variables

The independent variables employed in the regression analysis are the following:

1. BMI difference from the obesity threshold
2. Gender (1=Female; 2=Male)
3. Age (age at baseline was applied in the regression)
4. Ethnicity (1=White; 2=Asian; 3=Black; 4=Other)
5. Single parent status or not (0=No; 1=Yes)
6. Living in their own houses or not (0=No; 1=Yes)
7. The total number of programme attendance
8. Level of deprivation (quintile)

Statistical analysis

A logistic regression was applied to evaluate the effects of the explanatory variables on obesity status. Owing to the difference in the obesity status change rate across ethnicities shown in the simple logistic regression, the author conducted the analysis by ethnic stratification. The tests for goodness of fit and discriminating ability of the model were performed.

All statistical analyses were carried out using Stata version 11 (StataCorp, 2011).

3.4 Results

3.4.1 Factors influencing the BMI change rate

The correlations between the variables in the regression models were tested before starting the regression analysis. Correlations among the variables are shown in Table 3.15. As seen, moderate correlation was found between baseline BMI and baseline age, and single parent status and owning home status. Nevertheless, the degree of correlation is deemed insignificant.

Table 3.15: Correlations between variables with p-values

Variables	bmidif	Period wk	sex	age1	single parent	Own home	ethnic1	bmipre	Total Atten dance	imd_q
bmidif	**1**									
periodwk	**-0.0095**	**1**								
	0.4346									
sex	**-0.0425**	**-0.003**	**1**							
	0.0004	*0.8036*								
age1	**-0.0235**	**-0.0201**	**0.0471**	**1**						
	0.0519	*0.0968*	*0.0001*							
single parent	**0.0061**	**-0.0065**	**0.0024**	**0.0107**	**1**					
	0.6888	*0.6703*	*0.8747*	*0.4838*						
ownhome	**-0.0231**	**0.0138**	**-0.0077**	**0.0749**	**-0.3156**	**1**				
	0.1308	*0.3668*	*0.6143*	*0*	*0*					
ethnic1	**0.0936**	**0.0084**	**0.0096**	**-0.097**	**0.0633**	**-0.1322**	**1**			
	0	*0.5802*	*0.5273*	*0*	*0*	*0*				
bmipre	**-0.0748**	**-0.0027**	**-0.0021**	**0.392**	**0.0488**	**-0.058**	**-0.0037**	**1**		
	0	*0.8239*	*0.8613*	*0*	*0.0014*	*0.0001*	*0.8092*			
Total Attendance	**-0.1191**	**-0.0405**	**-0.0036**	**0.012**	**-0.0608**	**0.084**	**-0.1035**	**-0.0519**	**1**	
	0	*0.0036*	*0.7935*	*0.387*	*0.0003*	*0*	*0*	*0.0002*		
imd_q	**0.0706**	**-0.0188**	**0.0002**	**-0.0682**	**0.1566**	**-0.3109**	**0.261**	**0.0733**	**-0.0447**	**1**
	0	*0.1373*	*0.9853*	*0*	*0*	*0*	*0*	*0*	*0.0022*	

First, a simple linear regression was carried out to explore the association between the BMI change and covariates. The diagnostic tests for the model were also performed, including testing for multicollinearity, normality, heteroscedasticity, and the independence of observations (Farrar and Glauber, 1967). Concerning multicollinearity testing, the variance inflation factor was applied, with no evidence of multicollinearity when considering the results. Unlike multicollinearity, testing the distribution of residuals uncovered a problem regarding heteroscedasticity, which indicates a lack of consistency of variances across observations (Cook and Weisberg, 1983). Heteroscedasticity affects the preliminary assumption of the regression and thus impacts the accuracy of the analysis. Hence, the robust standard errors, Huber-White, were applied in the analysis to resolve this problem.

Table 3.16 shows the results from a simple linear regression exploring the association between the change in BMI and other variables. As shown in the table, there is weak evidence indicating that the BMI change of boys is slightly higher than that of girls. Strong evidence suggests a decrease in programme effectiveness in 'Asian', 'Black' and 'Other' groups, compared with 'White'. Moreover, there is strong evidence showing that the BMI change increases in children who frequently attended the programme and in those with higher baseline BMIs.

From these results, it can be concluded that children with particular characteristics experience a higher degree of the effectiveness of MEND 7-13 in reducing BMIs. These characteristics are: male, White ethnicity, high baseline BMI, and frequent attendance of the programme.

Table 3.16: Association between the BMI change and other factors (robust SE)

Variables	Coefficient	SE	p-value
Gender	-0.067	0.035	0.053
Age	0.016	0.012	0.173
Ethnicity (compared with "White")			
Asian	0.209	0.062	0.001
Black	0.275	0.074	<0.001
Others	0.141	0.061	0.020
Single parent	0.004	0.040	0.928
Own home	-0.021	0.040	0.592
Baseline BMI	-0.023	0.005	<0.001
Total attendance	-0.044	0.007	<0.001
Deprivation level (compared with the least deprived area "1st quintile")			
2nd quintile	0.140	0.061	0.022
3rd quintile	0.109	0.062	0.078
4th quintile	0.068	0.059	0.255
5th quintile	0.134	0.061	0.028
constant	0.268	0.186	0.150

Next, interaction terms with the 'ethnicity' variable were added into the regression to explore their impact on the model. Ethnicity was selected to create interaction terms due to its high magnitude (coefficient) of association, as shown in Table 3.16. After incorporating interaction terms into the model, the results suggest that the effectiveness of MEND 7-13 is different among children of different ethnic origins. Table 3.17 compares the means of BMI change by ethnicity. As seen, the BMI changes of Asian and Black children tend to be lower than those of White and Other. Compared to White, the percent change in BMI of the Asian, Black, and Other groups are 62.1%, 61.6%, and 78.5%, respectively.

Table 3.17: Comparing means of the BMI change by ethnicity

Ethnicity	Mean BMI reduction during the 10-week programme	SE	Comparing the percent of BMI reduction to 'White'
White	0.858	0.017	100%
Asian	0.533	0.045	62.1%
Black	0.529	0.064	61.6%
Other	0.673	0.050	78.5%

To account for differences in the BMI change across ethnicities, the analysis was performed by ethnic stratification: White, Asian, Black and Other. Table 3.18 shows the impact of various characteristics on the BMI change among different ethnicities. Owing to ethnic stratification, only 2,391 observations for 'White', 271 for 'Asian'; 156 for 'Black', and 264 for 'Other' were left in the analyses.

Table 3.18: Comparing the BMI change rate across ethnic groups (robust SE)

Variables	Coefficient	SE	p-value
White			
Gender	-0.066	0.040	0.096
Age	0.024	0.014	0.083
Single parent	-0.053	0.046	0.248
Own home	-0.079	0.046	0.088
Baseline BMI	-0.027	0.006	<0.001
Total attendance	-0.049	0.008	<0.001
Deprivation 2nd quintile	0.133	0.063	0.037
Deprivation 3rd quintile	0.091	0.065	0.164
Deprivation 4th quintile	0.059	0.063	0.351
Deprivation 5th quintile	0.144	0.066	0.029
constant	0.430	0.221	0.052
Asian			
Gender	0.052	0.114	0.647
Age	0.001	0.039	0.973
Single parent	0.172	0.165	0.298
Own home	0.236	0.124	0.059
Baseline BMI	-0.024	0.015	0.117
Total attendance	-0.017	0.021	0.416
Deprivation 2nd quintile	0.192	0.390	0.622
Deprivation 3rd quintile	0.288	0.394	0.466
Deprivation 4th quintile	0.323	0.382	0.399
Deprivation 5th quintile	0.231	0.377	0.540
constant	-0.163	0.580	0.779
Black			
Gender	-0.008	0.134	0.954
Age	-0.043	0.045	0.340
Single parent	0.198	0.141	0.162
Own home	0.329	0.172	0.057
Baseline BMI	-0.021	0.014	0.149
Total attendance	0.013	0.021	0.537
Deprivation 2nd quintile	0.722	0.517	0.165
Deprivation 3rd quintile	0.271	0.490	0.581
Deprivation 4th quintile	-0.023	0.478	0.962
Deprivation 5th quintile	0.247	0.469	0.599

constant	-0.120	0.667	0.858
Others			
Gender	-0.321	0.116	0.006
Age	-0.012	0.044	0.789
Single parent	0.089	0.124	0.470
Own home	-0.123	0.140	0.381
Baseline BMI	-0.005	0.012	0.670
Total attendance	-0.063	0.016	<0.001
Deprivation 2nd quintile	-0.218	0.325	0.502
Deprivation 3rd quintile	-0.046	0.311	0.882
Deprivation 4th quintile	-0.301	0.309	0.330
Deprivation 5th quintile	-0.219	0.311	0.481
constant	0.906	0.637	0.156

Regarding the impact of gender in the simple regression, the results fail to reveal any evidence of significant associations in all ethnic groups but ‘White’ and ‘Other’. There is also some evidence showing the association between the change in BMI and age, but only in White children. In contrast, some evidence indicates that having parents who own their own houses is a significant predictor in ‘White’, ‘Asian’ and ‘Black’ groups; however, these associations do not go in the same directions. Owning a home increases the BMI change rate in White children but decreases it in Black and Asian peers. Moreover, strong evidence supporting the association between the BMI change and baseline BMI remained only in White children after ethnic stratification. Lastly, the results show that strong evidence supports the association with programme attendance only in children categorized as ‘White’ and ‘Other’.

Due to a great amount of missing data, MICE was carried out to account for the missing values in the analysis. Stata/SE version 11 statistical software package was used in this regard. Three types of registration were applied: 1) Imputed – used when there are missing values that need to be imputed before running the analysis; 2) Regular – used when there is no missing value, so the variable is used in the model as it is; and 3) Passive – used when there are missing values, but no imputation is processed before the analysis. Instead, the software fills in missing values based on other regular and imputed variables.

In this analysis, variables containing no missing value were registered as regular and then used in the imputation process. For the variables with missing values, the imputed registration was applied, meaning that the missing values will be imputed before running the analysis. None of the variables were registered as passive. Table 3.19 lists the types of registration for each variable. As shown, gender, age, and baseline BMI were registered as regular, as no missing value is found in them. Data imputation is carried out for the rest of the variables, and registered as imputed.

Table 3.19: Registering variables for imputation

Registration types	Variables
Imputed	singleparent, ownhome, Totalattendance, bmidif, imd_q
Regular	sex, age1, bmipre

Table 3.20 illustrates the methods and predictors used in the imputation process. Selection of imputation method is based on the distribution of the variable; that is, multivariate normal regression is used for continuous variables, ordered logistic regression for ordinal variables, and logistic regression for binary variables. Concerning the predictors used in the imputation process, gender and baseline age were applied in all imputations.

Twenty imputations were produced for each variable.

Table 3.20: Methods and predictors used in the imputation process

Variables	Imputed methods	Predictors
Totalattendance	multivariate normal regression	sex, age1, singleparent, ownhome
bmidif	multivariate normal regression	sex, age1, bmipre
imd_q	ordered logistic regression	sex, age1, singleparent, ownhome
singleparent	logistic regression	sex, age1, ownhome
ownhome	logistic regression	sex, age1, singleparent

After registering the variables, a simple linear regression was carried out again with the ethnicity stratification. Table 3.21 demonstrates the results from the analysis of the imputed data.

Table 3.21: Imputation analysis comparing the BMI change rate across ethnic groups (robust SE)

Variables	Coefficient	SE	p-value
White			
Gender	-0.077	0.033	0.021
Age	0.028	0.011	0.013
Single parent	-0.027	0.039	0.491
Own home	-0.055	0.039	0.156
Baseline BMI	-0.025	0.005	<0.001
Total attendance	-0.037	0.007	<0.001
Deprivation 2nd quintile	0.109	0.056	0.051
Deprivation 3rd quintile	0.032	0.056	0.570
Deprivation 4th quintile	0.100	0.055	0.071
Deprivation 5th quintile	0.142	0.056	0.012
constant	0.141	0.190	0.458
Asian			
Gender	-0.004	0.092	0.967
Age	0.017	0.030	0.574
Single parent	0.124	0.132	0.346
Own home	0.258	0.099	0.010
Baseline BMI	-0.022	0.011	0.050
Total attendance	-0.010	0.018	0.566
Deprivation 2nd quintile	0.201	0.270	0.457
Deprivation 3rd quintile	0.119	0.277	0.668
Deprivation 4th quintile	0.264	0.261	0.312
Deprivation 5th quintile	0.241	0.261	0.356
constant	-0.355	0.464	0.445
Black			
Gender	-0.088	0.128	0.493
Age	-0.064	0.041	0.117
Single parent	0.086	0.131	0.512
Own home	0.124	0.159	0.439
Baseline BMI	0.013	0.019	0.501
Total attendance	0.016	0.025	0.530
Deprivation 2nd quintile	0.617	0.573	0.285
Deprivation 3rd quintile	0.227	0.543	0.677
Deprivation 4th quintile	-0.059	0.479	0.902
Deprivation 5th quintile	0.117	0.456	0.798
constant	-0.630	0.719	0.383
Others			
Gender	-0.291	0.103	0.005
Age	-0.029	0.037	0.434
Single parent	-0.031	0.110	0.777
Own home	-0.101	0.114	0.373
Baseline BMI	0.007	0.012	0.571

Total attendance	-0.048	0.017	0.006
Deprivation 2nd quintile	-0.198	0.284	0.486
Deprivation 3rd quintile	-0.003	0.269	0.991
Deprivation 4th quintile	-0.247	0.262	0.347
Deprivation 5th quintile	-0.194	0.263	0.461
constant	0.526	0.538	0.329

The results from the imputation analysis confirm that the missing values in the data barely affect the results. In the 'White' group, imputation analysis diminished the effect of having parents who own their own houses on the BMI change rate, while the effects of other variables remained the same. In the 'Asian' group, the only change that appeared after the imputation analysis was some evidence indicating a positive association between baseline BMI and the BMI change rate. In the 'Black' group, evidence of association between the BMI change and home ownership no longer existed. In the 'Other' group, no change can be observed.

Lastly, a multilevel analysis was performed to account for the impact of the programme site on the BMI change rate. In this regard, the programme site was added into the model as a class variable. Table 3.22 illustrates the results from the multilevel analysis.

Table 3.22: Multilevel analysis showing the association between the BMI change rate and other factors across ethnicity (robust SE)

Variables	Coefficient	SE	p-value
White			
Fixed-effects parameters			
Gender	-0.067	0.038	0.075
Age	0.023	0.013	0.094
Single parent	-0.055	0.046	0.227
Own home	-0.071	0.050	0.159
Baseline BMI	-0.027	0.006	<0.001
Total attendance	-0.055	0.009	<0.001
Deprivation 2nd quintile	0.142	0.067	0.033
Deprivation 3rd quintile	0.078	0.068	0.247
Deprivation 4th quintile	0.066	0.066	0.317
Deprivation 5th quintile	0.129	0.070	0.066
constant	0.548	0.233	0.019
Random-effects parameters			
Programme site			
Variance of the model	0.763		

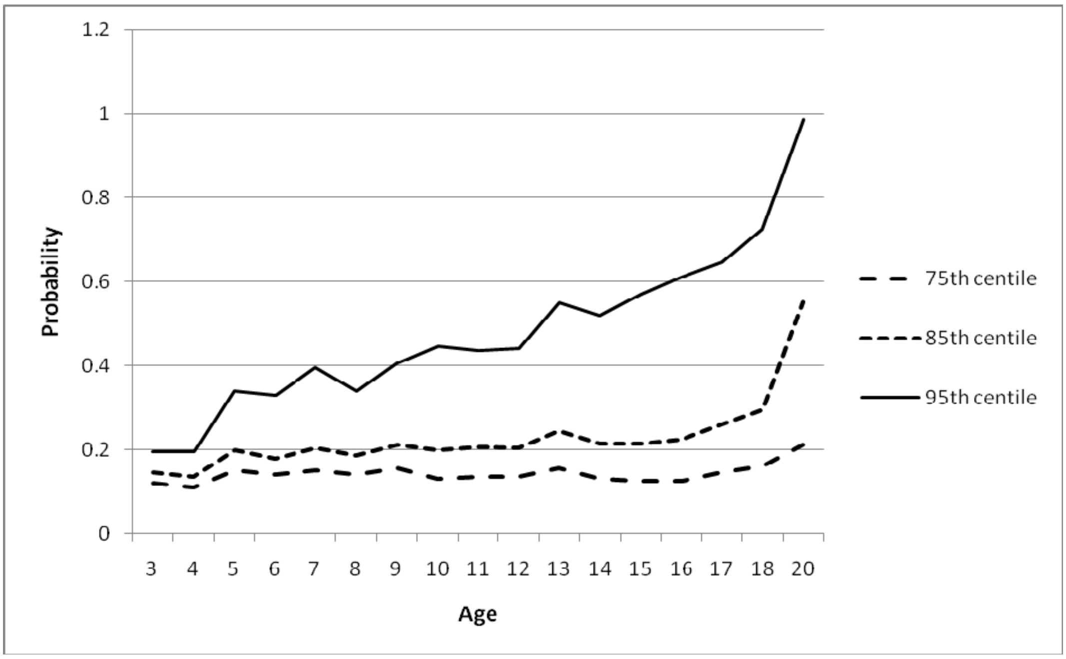
Variance of the residual	0.154		
Proportion of variance explained by the program site	0.168		
Asian			
Fixed-effects parameters			
Gender	0.024	0.116	0.839
Age	-0.002	0.032	0.959
Single parent	0.158	0.159	0.322
Own home	0.239	0.121	0.049
Baseline BMI	-0.027	0.014	0.065
Total attendance	-0.017	0.020	0.400
Deprivation 2nd quintile	0.180	0.388	0.644
Deprivation 3rd quintile	0.261	0.383	0.497
Deprivation 4th quintile	0.291	0.384	0.448
Deprivation 5th quintile	0.236	0.376	0.531
constant	-0.032	0.608	0.959
Random-effects parameters			
Programme site			
Variance of the model	0.777		
Variance of the residual	0.071		
Proportion of variance explained by the program site	0.084		
Black			
Fixed-effects parameters			
Gender	-0.016	0.134	0.903
Age	-0.054	0.046	0.235
Single parent	0.164	0.152	0.280
Own home	0.293	0.189	0.121
Baseline BMI	-0.019	0.015	0.202
Total attendance	0.014	0.021	0.488
Deprivation 2nd quintile	0.742	0.512	0.147
Deprivation 3rd quintile	0.256	0.469	0.586
Deprivation 4th quintile	0.019	0.451	0.967
Deprivation 5th quintile	0.258	0.447	0.564
constant	-0.055	0.642	0.932
Random-effects parameters			
Programme site			
Variance of the model	0.570		
Variance of the residual	0.080		
Proportion of variance explained by the program site	0.123		
Others			
Fixed-effects parameters			
Gender	-0.306	0.121	0.011
Age	-0.006	0.038	0.877
Single parent	0.062	0.126	0.623

Own home	-0.129	0.133	0.333
Baseline BMI	-0.005	0.012	0.699
Total attendance	-0.063	0.018	<0.001
Deprivation 2nd quintile	-0.242	0.338	0.473
Deprivation 3rd quintile	-0.061	0.335	0.855
Deprivation 4th quintile	-0.317	0.324	0.328
Deprivation 5th quintile	-0.227	0.329	0.491
constant	0.848	0.656	0.196
Random-effects parameters			
Programme site			
Variance of the model	0.686		
Variance of the residual	0.088		
Proportion of variance explained by the program site	0.114		

As seen in the table, the coefficients of the fixed-effect parameters changed to some extent after accounting for the difference in programme centre. In the 'White' group, some evidence indicates that boys tend to reduce BMI at a higher rate than do girls, whereas there is strong evidence indicating that frequent programme attendance and high BMI baseline are associated with a rise in the BMI change. While having parents who own their houses and frequent programme attendance are significant predictors for the BMI change in the 'Asian' group, no evidence of association is seen in the 'Black' group. In the 'Other' group, there is strong evidence for a positive association between age and attendance rate and BMI change after accounting for the programme site effect.

With regard to the random-effects parameters, the results indicate the variation between sites to be attributable to the programme by 16.8%, 8.4%, 12.3%, and 11.4% in 'White', 'Asian', 'Black', and 'Other', respectively. The results imply that the effect of the variation between programme sites on BMI change is high in White children compared with their peers.

In the next part, this study assesses the intrinsic factors affecting the change in obesity status. This is because the influence of BMI changes on health also depends on baseline BMIs, particularly with regard to the probability of adult obesity, which is the core assumption of this thesis. The impact of change in obesity status in childhood and reduced probabilities of adult obesity can be seen in Figure 3.15. It is noted that probabilities of adult obesity decrease drastically when BMIs shift from the 95th percentile to 85th percentile, compared with shifting from 85th percentile to 75th percentile.



Source: modified from Guo et al. (2002)

Figure 3.15: Probability of becoming obese at age 35 for children with BMI values at 75th, 85th, and 95th percentile

As obesity status in childhood is a significant factor in determining the probability of becoming obese in adulthood, the economic modelling in the next chapter is built around this association. Prior to evaluating future health using obesity status in childhood, this chapter explores the characteristics that influence the change in obesity status.

3.4.2 Factors influencing the change in obesity status

Firstly, this study expresses particular concern when applying obesity cut-offs to assess the change in obesity status. The chances of changing into the non-obese category for obese children are not equal; rather they depend on baseline BMIs, with children whose BMIs are slightly above the obesity cut-offs having a higher chance of obesity reversion compared with extremely obese peers. Figure 3.16 demonstrates the association between changes in obesity status and baseline BMIs from the MEND data. It is evident that changes in obesity status occurred only in obese children whose BMI differences from the cut-offs are less than 6 kg/m^2.

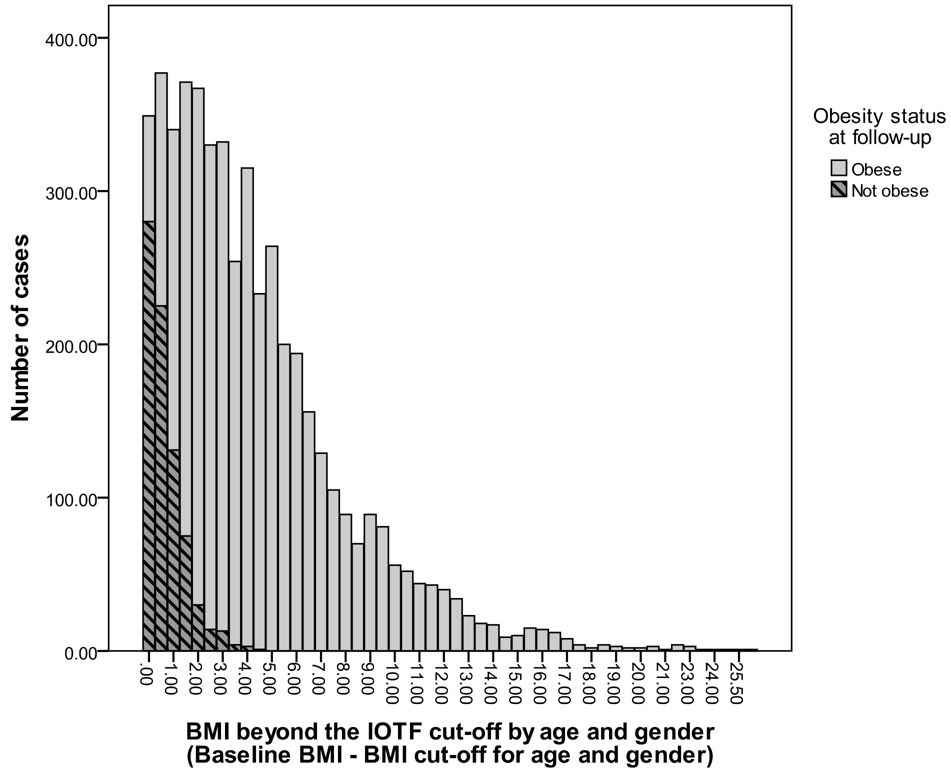


Figure 3.16: The influence of differences in baseline BMI and obesity cut-offs on the change in obesity status of obese children at baseline

To begin with, correlations between variables in the subgroup data were elucidated, as shown in Table 3.23. As shown, there is a high correlation between baseline age and baseline BMI, while the BMI difference from the obesity threshold resolves this problem.

In order to account for the impact of this on the analysis, this study carried out a subgroup analysis by focusing on obese children (IOTF definition) whose baseline BMIs differ from the cut-offs by less than 7 units. Justifications for doing so are: 1) this subgroup still contains up to 80% of the total obese children (4,082 observations); and 2) the subgroup analysis reduces the effect of those with BMIs farther away from the cut-offs who remain obese at follow-up. This variable, BMI difference from the obesity threshold, was used in the analysis.

Table 3.23: Correlations between variables and p-values using the subgroup data (used in the logistic regression)

Variables	obstat1	sex	age1	single parent	own home	ethnic1	bmidif	bmipre	Total Attendance	imd_q
obstat1	**1**									
sex	**0.012**	**1**								
	0.430									
age1	**0.032**	**0.064**	**1**							
	0.039	*0.000*								
single parent	**-0.026**	**0.001**	**0.015**	**1**						
	0.198	*0.948*	*0.446*							
ownhome	**0.045**	**0.001**	**0.075**	**-0.303**	**1**					
	0.024	*0.968*	*0.000*	*0.000*						
ethnic1	**-0.052**	**0.013**	**-0.106**	**0.077**	**-0.133**	**1**				
	0.009	*0.519*	*0.000*	*0.000*	*0.000*					
bmidif	**-0.544**	**-0.008**	**-0.012**	**0.036**	**-0.040**	**0.025**	**1**			
	0.000	*0.617*	*0.435*	*0.069*	*0.042*	*0.206*				
bmipre	**-0.358**	**-0.010**	**0.714**	**0.035**	**0.027**	**-0.060**	**0.684**	**1**		
	0.000	*0.513*	*0.000*	*0.080*	*0.177*	*0.002*	*0*			
Total Attendance	**0.030**	**-0.008**	**0.018**	**-0.068**	**0.069**	**-0.076**	**-0.017**	**-0.001**	**1**	
	0.091	*0.643*	*0.330*	*0.002*	*0.001*	*0.001*	*0.340*	*0.975*		
imd_q	**0.073**	**-0.006**	**0.071**	**-0.153**	**0.319**	**-0.255**	**-0.063**	**0.007**	**0.043**	**1**
	0.000	*0.736*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.657*	*0.021*	

A simple logistic regression was applied to assess an association of change in obesity status with other factors, as shown in Table 3.24. As shown by the table, evidence supports an association of obesity reversion with BMI difference from the threshold and total attendance, whereas no evidence is provided for other factors. In this regard, strong evidence indicates that for every single unit increase in BMI difference from the threshold, the odds of becoming non-obese decrease by 0.84 with a 95% CI (0.80, 0.87). There is also weak evidence showing that the odds of becoming non-obese are 1.069 higher for every additional instance of programme attendance, controlling for other factors.

Table 3.24: Association between change in obesity status and other factors (robust SE)

Variables	Odds ratio	95% Confidence Interval	
		Lower	Upper
Gender	0.768	0.551	1.070
Age	1.010	0.919	1.110
Ethnicity (compared with "White")			
Asian	0.557	0.303	1.024
Black	0.426	0.179	1.011
Others	0.829	0.428	1.608
Single parent	1.121	0.781	1.609
Own home	1.051	0.735	1.505
BMI difference from the obesity threshold	0.160	0.127	0.202
Total attendance	1.069	1.006	1.137
Deprivation level (compared with the least deprived area "1st quintile")			
2nd quintile	1.225	0.748	2.007
3rd quintile	1.378	0.828	2.292
4th quintile	1.232	0.729	2.083
5th quintile	1.006	0.563	1.799

The test for goodness of fit of the simple logistic model indicated that the predicted values from the model are significantly different from the real data, representing the poor fit of the model. This affects the model's predictability and result interpretation. In other words, the results from the logistic regression do not reflect well the characteristics of the dataset. The study also tests the discriminating ability of the model using the Receiver Operating Characteristic (ROC) curve, resulting in the ROC the model of 0.93[1]. The author chose not to carry out the logistic regression analysis, as results would not represent the MEND data due to poor fit.

3.5 Discussion

This chapter reviews the background and development of MEND 7-13, as well as evidence of the programme's effectiveness. It also elucidates the impacts of intrinsic

[1] The Receiver Operating Characteristic (ROC) analysis informs how well the model can discriminate between 2 categories (binary outcome) by considering a trade-off between the true positive rate and false positive rate. The area under the ROC curve represents the discriminating ability of the model, compared with 0.5 (no discriminating ability or as random as flipping a coin).

factors on benefit uptakes of the programme in terms of change in BMI and obesity status change. The results show that programme effectiveness, in terms of the BMI change, is affected to a large extent by the ethnic origin of participants. Furthermore, the results from the analysis highlight issues regarding health disparities for programme effectiveness among participants. It is noted that, due to the poor fit of the logistic model, the results from the regression do not aptly represent the influences of participants' characteristics and thus are not interpreted here.

Concerning the change in BMI, the interpretation of the results is based on multilevel analysis due to its comprehensiveness. Significant predictors in White children are found to be gender, frequent attendance, and baseline BMI. The BMI reduction in boys is approximately 0.07 units higher than that in girls; every additional instance of programme attendance results in an additional 0.06 units of BMI reduction; and every unit increase in baseline BMI leads to an additional BMI reduction of 0.03 units. In the 'Asian' group, some evidence shows that having parents who own their own houses and baseline BMI are associated with BMI change. Parental home ownership reduces the BMI change in children by 0.24 units, compared with their counterparts; every additional unit increase in baseline BMI results in an additional 0.03-unit reduction of BMI. No evidence of association is found in Black children. Total attendance and gender are significant predictors in the 'Other' group, where the BMI reduction in boys is approximately 0.3 units higher than that in girls and every additional instance of programme attendance results in an additional 0.06 units of BMI reduction. The results also show that the programme centre has some influence on the change in BMI per week in all children, though this effect varies slightly across ethnicities; that is, the variation between sites is attributable to differences in the programme centres by 16.8%, 8.4%, 12.3%, and 11.4% in 'White', 'Asian', 'Black', and 'Other', respectively. The influence of the programme site on programme effectiveness may stem from the location and staff consistency, with programme location playing an important role in creating a successful learning atmosphere. For instance, the programmes that have access to swimming pools, playgrounds, and modern gyms might be more attractive than those located in poorly-maintained local gyms. Consistency of the local staff also affects the quality of learning. Even though MEND Central equips the local staff with a 4-day training course and manuals before allowing them to lead the programme sessions, there may be a considerable amount of difference in the teams' ability to convey the materials to programme attendees, depending on staff's characteristics and their pre-existing skills. Additional consideration of the consistency of programme delivery across sites may reduce inequity in programme quality and improve the effectiveness of MEND 7-13.

The major issue with regard to statistically significant predictors is whether or not their impact on the BMI change is also clinically significant. The magnitudes of all significant predictors are extremely low; as seen in Table 3.22, the coefficients of most significant predictors are below 0.1 units. In addition, body growth of the children also makes the interpretation more difficult, as BMI is dependent on height, which increases with age. Interpreting the clinical significance of BMI changes must therefore be done with respect to age, which did not come into play in this chapter due to a short follow-up period.

Hence, it can be concluded only that equity issues may exist among the participants of the programme with respect to different socio-demographic factors and programme site.

Certain points pertaining to the MEND 7-13 programme can be clarified and emphasized for a better understanding of the programme and outcome utilization. These include potential threats against the accuracy and reliability of the results in the analysis, such as supportive evidence of the programme effectiveness and the quality of the data used in the analysis.

The main evidence supporting the effectiveness of MEND 7-13 comes from a 6-month randomised-controlled trial in obese children, with BMI $\geq 98^{th}$ centile of the UK 1990 BMI chart (Sacher et al., 2010). It is noted, however, that the rollout phase expands the eligibility criteria to cover overweight children (BMI $\geq 91^{st}$ centile), while only obese children (BMI $\geq 98^{th}$ centile) were included in the trial period. Although this study performed a t-test to compare health parameters before and after the intervention, the regression to the mean bias cannot be ruled out, meaning that BMIs of children may tend to decrease to the normal limit regardless of the intervention. Moreover, the skills and expertises of the staff employed during the MEND rollout phase were diminished compared with those in the trial. Despite increased scalability and possible cost-effectiveness, stronger evidence should still be required to demonstrate the effectiveness of the rollout programme, e.g. conducting a trial comparing the rollout programme with a control group.

Moreover, uncertainty of the results is increased by the fact that the analysis is based on short follow-up data. The average duration of follow-up is around 10 weeks, which is deemed an insufficiently long period when considering how fast weight and BMI can change, especially in children and adolescents. Further evidence, particularly from a long follow-up study, is required to provide more information about the long-term effectiveness of the programme.

Another limitation of the results stems from the quality of the data used, which affects the accuracy and reliability of the analysis. As mentioned in the chapter, a number of inconsistencies appeared in the data collecting process, such as follow-up methods, re-registration, and massive amounts of missing data. Furthermore, the comparison of socio-demographic distributions between MEND 7-13 and HSE 2008 suggests the possibility of selection bias in the recruitment process. The results show that obese children of MEND 7-13 might come from less deprived areas than did those in HSE 2008. Therefore, it is acknowledged that the utility and generalisability of the results might be limited due to these factors.

In addition to concerns regarding the rationale, study design, and quality of the data, the interpretation of results is also an issue of consideration. Even though some evidence of association is present in the analysis, the clinical significance of the results remains in doubt. This is especially true when the BMI change rate is an outcome variable, where the associations found between variables were miniscule. Results interpretation and utilization should be done carefully in conjunction with clinical implication.

Given the weaknesses of the programme, several recommendations to improve the MEND 7-13 rollout can be made in an attempt to enhance programme effectiveness and accountability.

First, re-assessment of the effectiveness of the MEND 7-13 rollout programme, particularly in the long run, is required. The MEND 7-13 rollout diminishes the qualifications of the staff in order to make the programme cost-effective and expands the eligibility criteria from the BMI cut-off of the 98th centile to the 91st centile. However, the only evidence supporting the programme effectiveness are the changes in health indicators compared with their baselines over 10-week follow-up. The analysis to assess MEND 7-13's effectiveness was carried out using a paired-samples t-test comparing pre- and post-intervention, without a control group. The results from a paired-samples t-test, assuming a flat comparator for the control, are not enough to demonstrate programme effectiveness, due essentially to the regression to the mean bias. Moreover, the sample size of the MEND RCT is deemed small, especially when 14 health indicators were examined at the same time. In addition, the duration of follow-up is considered too short. Programme effectiveness derived from a 10-week follow-up cannot be used to predict long-term outcomes, as a considerable number of long-term uncertainties are involved. Owing to these weaknesses, inferences drawn from the short-term data are limited in their applicability, particularly for use in the policy decision making process.

To evaluate the robustness of programme effectiveness, a long-term randomized controlled study with an adequate sample size in the MEND 7-13 rollout setting is needed. Ideally, the duration of follow-up should last until adulthood. However, a realistic aim for follow-up duration may be 1-2 years, as the routine health assessment at 12 months is already present in the protocol. Reducing the loss at follow-up at 12 months, and thus the amount of missing data, would help immensely in determining programme effectiveness with greater accuracy by providing more useful information about MEND 7-13 in comparison with the currently used data.

When assessing multiple health indicators together, the significance level needs to be adjusted using statistical techniques; one technique that can be employed in this regard is the Bonferroni correction.

Concerning the registration process, other methods of programme enrolment, apart from the website, should be made available in order to reduce the selection bias. Possible alternatives include registration venues at schools and communities, or by post. In designing the registration process, MEND Central should also target children from low SES, who are especially vulnerable to obesity. Additionally, the registration and database system should be re-designed to prevent re-registration. These actions would improve the quality of the programme delivery and data, even though additional costs may come alongside.

Another recommendation pertains to programme evaluation, which should be done more regularly and comprehensively. A number of outcome parameters have been collected to show the performance of MEND 7-13, such as the frequency of attendance and health indicators, at the end of the initial 10-week sessions. However, almost all parameters were missing at 6- and 12-months follow-up. Many variables also contained a large number of missing values at the end of the programme (10-12 weeks). The considerable amount of missing data prompts the question of whether the follow-up sessions at 6 and 12 months were in fact carried out, and if so, how they were conducted and why most health indicators were not collected or recorded. Moreover, the follow-up method should be assessed to determine if one method may be more effective than the others, such as phone, email, text message, or post.

In addition to outcome evaluation, process evaluation is also deemed important to ensure the quality of the delivery process. Information used to evaluate the process of delivery, however, is missing from the data. For example, the number of sessions attended by the children is tabulated as an indicator of the duration of programme exposure. However, it is also important to assess the quality of the contents delivered during the sessions, such as the number of topics successfully conveyed and the learning atmosphere in classes. Measuring and providing indicators that reflect the process of instruction during the sessions would better ensure the quality of content delivery.

In addition to the above-mentioned recommendations, the characteristics of the programme centres should also be examined for the future development of MEND 7-13. Results from the statistical analysis shed light on the impact of the different programme centre characteristics on the BMI change rate. Characteristics of the programme site affecting the performance include the qualification, experience, and individual abilities of the local staff, as well as the programme location. Comparing the factors that influence programme effectiveness with respect to particular outcome aspects would reveal the extent to which specific components of the programme centre enhance the quality of outcomes. This information would be helpful in allowing MEND 7-13 to achieve better performance.

Despite certain weaknesses as stated above, this chapter provides detailed insights into MEND 7-13 and the implications of particular characteristics that influence programme effectiveness in children. Not only can the economic evaluation in the next chapter benefit from this understanding, but the information gained may also highlight issues regarding the disparities in programme effectiveness, allowing programme executives to take these considerations into account in the future.

Chapter 4: Economic Evaluation of the National Child Weight Management Programme: MEND 7-13

4.1 Introduction

Chapter 3 introduces the MEND 7-13 programme and explores the extent to which different factors influence benefit uptake. Despite its effectiveness, there is still a lack of evidence to indicate whether MEND should be widely implemented. In order to make this decision, the costs of the programme need to be considered alongside its benefits. Economic evaluation is a tool that can be used to help answer this policy question. The objective of this chapter is to evaluate the long-term health economic impact of MEND 7-13 in the rollout phase.

4.2 Methodology

4.2.1 Cost-Effectiveness Framework and Assumption

The long-term economic evaluation of the MEND 7-13 Programme was performed using the cost-effectiveness framework, which shows the result in terms of cost per additional unit of benefit generated, or the incremental cost-effectiveness ratio (ICER) (Gold et al., 1996, Drummond et al., 2005). The benefits were measured in terms of the additional quality-adjusted life years (QALYs) gained as a result of the intervention.

ICER = the additional costs of MEND / the additional QALY gained from MEND

The main advantage of this approach is that it allows comparison with the benefits from other types of intervention and other health programmes. This is accepted practice in the NHS, in line with NICE methodological guidance (National Institute for Health and Clinical Excellence, 2008).

In this study, the long-term costs and health consequences of MEND were projected under the assumption that MEND 7-13 is available for all children who are eligible for the programme according to its BMI cut-off criteria for recruitment: children in England between ages 7 and 13 with BMI ≥ 91st centile in 2010. This study selected 2010 as a year base so that the results of the study will be most relevant to the current situation. The costs and outcomes of MEND 7-13 were calculated separately.

MEND was not compared with alternative approaches to reducing obesity, as there is no consensus on standard treatment for obesity in children (National Institute for Health and Clinical Excellence, 2006, Scottish Intercollegiate Guidelines Network, 2010). Due to the lack of a standard approach to treating childhood obesity, a flat comparator, or current routine care, is assumed. Costs and benefits were estimated in terms of any changes from current demographic and epidemiological trends that might result from the introduction of MEND 7-13.

4.2.2 Cost Estimates

The cost analysis was based on the health care payer perspective (Gold et al., 1996). Even though the standard recommendation for cost-effectiveness analysis suggests that economic evaluation be conducted from the societal perspective, the health care payer perspective was chosen in this study due to practicality (Gold et al., 1996, Drummond et al., 2005, National Institute for Health and Clinical Excellence, 2008). The results of cost-effectiveness analysis are most useful when they contribute to the decision-making process. As MEND 7-13 is tailored as a public health programme to treat childhood obesity, the perspective applied in the analysis should appropriately cover the social costs and benefits that help policymakers in allocating public resources. Another perspective that may be applicable in this case is from the NICE Public Health guidance, which also includes the expenditures and savings of the non-health public sector (National Institute for Health and Clinical Excellence, 2009). However, the use of the NICE Public Health guidance is practically limited in this case, as the source of the estimates for obesity-related non-health costs in the public sector is very limited. This may be due to the fact that obesity leads to a number of adverse medical conditions, each of which may itself result in extensive incurred costs to the public sector. Currently, there is no existing literature estimating the non-health costs of obesity in the public sector. Therefore, it is not feasible to adopt the NICE Public Health perspective in this study.

The additional costs of the programme were calculated as the costs of delivery less the future health care costs avoided by the reduction in the future number of obese adults.

Costs = (**The costs of MEND in 2010 prices**) – (**The reduction of obesity-related medical costs in the future due to MEND**)*discount rate[2]

4.2.3 The costs of MEND

The author has limited access to the cost data of MEND 7-13. As a result, all information used in this thesis comes from communication with MEND rather than direct data. The inability to break down the programme costs is therefore acknowledged as a weakness of the analysis.

The costs of MEND are taken as the average costs of the programme per child, which are calculated by the total programme costs divided by the number of children in the programme. Because the analysis compares the intervention with the status quo, there is no estimation of the costs of alternative programmes.

Due to the varying scales at which the MEND programme can be delivered, cost calculation is based on two types of programme: small contract and large scale rollout. The total costs of the programme were then weighted by the number of children in each category. The average costs of MEND 7-13 per child, after adjusting for the size of the contract, were given by MEND.

[2] The NICE guidance suggests the discount rate of 3.5% for both cost and outcome.

Concerning the costs of data collection, the activity-based costs were accrued in the 4 main phases of the programme:

- Programme set-up (staff training, exercise and MEND kits, and system setup);
- Participant recruitment;
- Treatment; and
- Programme resource, delivery support, and evaluation and monitoring.

For each activity, fixed costs, such as office rental, and variable costs, such as staffing, equipment, and venues, were included.

4.2.4 The reduction of obesity-related medical costs in the future due to MEND

The cost of MEND was offset by the extent to which the programme reduced the obesity-related medical costs of MEND beneficiaries. The MEND programme not only reduces the number of obese children, but, through changing behaviour in the longer term, also reduces the number of obesity cases in adults and the consequent adverse health effects (Guo et al., 2002, Guo et al., 1994, Freedman et al., 2001, Braddon et al., 1986). The relationship between childhood and adult obesity was estimated from a cohort study by Freedman et al. (2001), with an average follow-up period of 17 years.

The direct medical costs of obesity in the UK were estimated from the Foresight report (McPherson et al., 2007) because of the following reasons: 1) the methods used in estimating direct medical costs of obesity are comprehensive – based on 10 high-cost obesity-related diseases such as diabetes, stroke, coronary heart disease, certain kinds of cancer, and other diseases; 2) this cost estimation is the most recent publication that estimates the medical costs of obesity in the UK; and 3) the results fit the economic model in this study.

In McPherson et al.'s study, it was estimated that the NHS will spend £5.3 billion and £7.1 billion on obesity-related medical costs in 2025 and 2050, respectively (McPherson et al., 2007). The proportion of these costs that might be avoided through

MEND was projected from the reduction in future adult obesity cases achieved by the programme.

To estimate the medical costs of obesity saved by MEND, a 17-year projection using 2010 as a year base was performed. The sum of the total medical costs of obesity between 2027 and 2079 was calculated to estimate the total medical costs of obesity through the lifespan. This is due to the fact that the medical costs of obesity can be incurred at any time in life, and the life expectancy of the UK population is approximately 79 years (World Health Organization, 2010). The analysis assumed that all those eligible for the MEND programme were between the ages of 7 and 13 in 2010 and thus will be between ages 24 and 30 in 2027. The scope was assumed for reduction in obesity-related costs for the rest of the cohort lifetime: 2027 to 2079.

Only part of the obese population – those with ages corresponding to MEND's eligibility criteria in 2010 – will benefit from MEND. The prevalence of obesity in each age group was assumed to be constant over time. The proportion of people aged 24-30 in the total obese adult population was calculated using Health Survey for England (HSE) data from 2008 (NHS, 2010), as HSE is the one of the most comprehensive sources of data about the UK population, and HSE 2008 is the most recent year in which the analysis was conducted. The future medical costs were discounted at 3.5% per year with base year 2010.

4.2.5 Measurement of Health Outcomes

On the outcome side, the short-term outcome of the programme is defined as the number of obesity cases avoided, i.e., the number of obese children diverted into the non-obese group after attending the programme compared with the status quo.

A number of methods have been proposed to define childhood obesity, as was seen in Chapter 1. In this study, the BMI cut-offs definition was used. Two BMI cut-offs were applied in this study: the 98th BMI centile cut-offs of the UK 1990 reference chart for age and sex (Scottish Intercollegiate Guidelines Network, 2010, NICE, 2006) and the international standard BMI cut-offs, or International Obesity Task Force (IOTF) definition

(Cole et al., 2000). In this study, both definitions were employed in the short-term analysis, and only the IOTF definition was applied in the long-term economic evaluation.

It is noted that the UK national survey data were used in constructing the IOTF definition for overweight and obesity. The IOTF sets obesity cut-offs at slightly higher levels than does the UK 1990 reference – for example, around 99th BMI centile at age 18 (Cole et al., 2000).

The observed number of obesity cases avoided from the short-term analysis was used to project the long-term health benefits of the programme. The main concepts applied to the future projection were as follows:

1. Childhood obesity increases the risk of becoming obese in adulthood
2. Obese adults have high risks of mortality in comparison to normal-weight adults.
3. The health-related quality of life (HRQL) of obese people is lower than that of normal-weight people.

1. Childhood obesity increases the risk of becoming obese in adulthood

Strong evidence indicates that obese children and adolescents are more likely to become obese adults than the non-obese (Singh et al., 2008). As found by the Freedman et al. (2001) cohort, obese children have a 4.3 times higher risk of becoming obese in adulthood than do non-obese children.

Even though many longitudinal studies have shown the strength of the relationship between childhood obesity and increased risk of becoming obese in adulthood, however, the results remain inconsistent across them.

Braddon et al. (1986) determined the probability of 11-year-old obese children becoming obese at age 36 to be 0.21, whereas Guo et al. (2002) estimated the probabilities of obese boys and girls between ages 7 and 13 of becoming obese at age 35 at 0.22-0.46 and 0.41-0.59 respectively. Another study by Freedman et al. (2001) revealed a probability of 0.77 for 10-year-old ($\pm$3) obese children becoming obese at age 27 ($\pm$5).

In this study, the probabilities used to project persistent obesity in adulthood were drawn from the Freedman et al. (2001) study for the following reasons:

- The probabilities from Braddon et al. (1986) were excluded because the age group used in the sample did not match the average age of MEND 7-13. The relevant starting ages provided in the Braddon et al. study are 7 and 11, whereas the other 2 studies showed the transitional probabilities at age 10 (MEND covers children between ages 7 and 13, or 10 on average). Moreover, the definitions of obesity and childhood obesity in the Braddon et al. study are inconsistent with those used in this study. Braddon et al. define childhood obesity as having weight >130% of the index of relative weight (this study uses BMI), while the cut-off for obesity in women is defined as BMI >29.1 (versus the standard BMI definition for adults of 30 being used in this thesis) (Braddon et al., 1986);

- Unlike in Braddon et al., both Freedman et al. (2001) and Guo et al. (2002) apply the BMI cut-offs of 30 for adulthood obesity and the 95th centile of the US population for childhood obesity. The definitions of adulthood obesity from these articles are thus consistent with that of this study. However, this study applies the IOTF definition for childhood obesity, with cut-offs set at approximately the 99th BMI centile of the UK 1990 reference chart for age and sex. The author thus acknowledges as a limitation a small inconsistency in the definition of childhood obesity.

- Freedman et al. (2001) followed up 2,617 samples, whereas the sample size in Guo et al. (2002) is only 347 subjects. Uncertainty from inadequate sample size is reduced by using the parameters from Freedman et al. (2001), rather than Guo et al. (2002); and

- The follow-up period in the Freedman et al. cohort is shorter than the others, resulting in less uncertainty from the shorter period of projection. Moreover, the risk of obesity increases with age (Seidell and Flegal, 1997, Kopelman, 2000). Freedman et al. (2001) followed up children from the average age of 10 to 27, while the follow-up durations in Guo et al. (2002) start from ages 3-20 to 35. Therefore, projecting the chance of becoming obese to a younger age group of adults also discounts the influence of age on the risk of obesity.

2. ***Obese adults have high risks of mortality in relation to normal-weight adults.***

The increased mortality risk of obesity from Flegal et al. (2005) was used in this study due to its comprehensiveness. Flegal et al. (2005) used the data from the National Health and Nutrition Examination Survey I-III in the US, which included over 35,000 samples. They calculated the relative risks of death by BMI level and showed that the relative risk of death increased up to 1.83 times when comparing people with normal BMI to those with BMI≥35 (Flegal et al., 2005). The details of the relative risks are shown in Table 4.5.

3. ***The health-related quality of life (HRQL) of obese people is lower than that of normal-weight people.***

Studies have demonstrated the association between higher BMI and lower HRQL (such as Sach et al., 2006, Wang et al., 2010). In this study, EQ-5D is the HRQL measure of choice, as it not only provides all quality of life information required, but also uses a convenient collection method, which improves the accuracy and reliability of the data.

In this study, the HRQL was drawn from Sach et al. for two reasons, as follows:

- The HRQL in Wang et al. (2010) was based on the US population, while the HRQL reported in Sach et al. (2006) was derived from 1,865 British adults aged 45 and over. The UK-based HRQL is better fit to this study, which focuses on child obesity interventions in the UK; and
- Wang et al. (2010) assumed all people with BMI greater than 30 to have the same quality of life without adjusting for the level of obesity; however,

assuming equal quality of life for people who are mildly obese and morbidly obese is considered inappropriate, as morbidly obese people are likely to have more difficulties than the mildly obese, leading to lower quality of life. In contrast, Sach et al. (2006) reported quality of life of the obese by BMI level, splitting subjects into categories with BMIs of 30-<35, 35-<40, and ≥40. Owing to its comprehensiveness, quality of life data from Sach et al. (2006) was used in this study.

In addition, some further assumptions were made when constructing the model as follows:

1. The persistence of obesity in adulthood begins at the age of 27 and lasts for the entire lifetime;
2. Non-obese adults at the age of 27 are assumed to remain non-obese for life. In other words, adults in both obese and non-obese groups will stay in their respective groups until they die and leave the model;
3. The short-term effects of MEND are sustained throughout the childhood period; and
4. The discount rates for the future costs and life years are 3.5% per year.

It is acknowledged that the assumptions made in order to construct the economic model may create some uncertainties; these issues are addressed in the discussion section of this chapter. Concerning the selection of the discount rates applied in the economic evaluation, the discount rate of 3.5% for both costs and life years has been chosen for this analysis among those recommended; this follows the NICE guidance (2008) as it is the national guidance for economic evaluation in the UK, which fits in the context of MEND 7-13.

The projection of the additional life years gained in the future was derived from the survival probability of the UK population. The survival probability used in this study was from the Office for National Statistics (2010) because of the reliability of the source and the applicability of the probability data to the model.

To find the additional life years gained, the survival probabilities from the life table (Office for National Statistics, 2010) were discounted at the rate of 3.5% and were then used to construct the survival curves. The areas under the curves were estimated to find the discounted life years. Afterward, QALYs were calculated by multiplying discounted life years by HRQL weights for obese and non-obese states.

4.2.6 Transforming QALY into monetary units

The contingent valuation approach was used to assign money values to health outcomes (QALYs) (Drummond et al., 2005). According to this approach, the healthy time gained is valued by maximum willingness to pay according to the health care payer perspective, which is £20,000-30,000 per QALY in the UK (National Institute for Health and Clinical Excellence, 2008). The net benefits of MEND were calculated as the monetary value of health gained less the net cost of the programme.

The ICERs of the previous cost-effectiveness studies in Chapter 2 were used to compare with the ICER of MEND, in order to show the cost-effectiveness of the MEND programme.

4.2.7 Sensitivity analysis

The multivariate sensitivity analysis was performed using the Monte Carlo simulation, 1,000 simulations for each scenario, to assess the uncertainty of the model assumptions. Both Microsoft Excel and Stata, a statistical package, were used in this regard. The sensitivity analysis accounted for the uncertainty of variables, including the probabilities of children turning into obese adults, relative risks of death at different ages and BMI levels, and EQ5D scores at different BMI levels.

The maintenance of programme effectiveness, in terms of BMI reduction, was assumed to be 100%, 50%, 25%, and 12.5%. The confidence interval level of 95% was applied to all uncertainties. Threshold analysis was conducted to ascertain the percentage of the maintenance of programme effectiveness needed in order to break even with the costs, given the ICER thresholds of £20,000 to £30,000 from NICE. Furthermore, the sensitivity analysis shows the importance of the discount rate applied in the calculation. Variations of the discount rate, ranging from 0-7%, were applied in order to demonstrate its influence on the ICER estimates.

The sensitivity analysis also reveals the implications of programme access among participants for health equity, as described in Chapter 3; that is, children from less deprived areas are more likely to enrol for the programme. In addition, health disparities may come into play not only during the recruitment process, but can also affect the effectiveness and cost-effectiveness of the programme as a consequence. This equity consideration is particularly important in economic evaluation (Williams and Cookson, 2006, Cookson et al., 2009). In order to address health disparities among participants, subgroup analysis was performed to determine ICERs with respect to the differences in programme effectiveness for particular characteristics. These features include ethnicity, home owning status, and single parent status. To demonstrate variations resulting from the definition of childhood obesity used, a sensitivity analysis was carried out, which applied the definition from the UK 1990 reference chart for age and sex.

The data used was from the above quoted sources and the rollout of the MEND programme. The analysis was performed using Stata ver.11 and Microsoft Excel.

4.3 Results

To calculate the costs and outcomes of MEND, the number of children eligible for the MEND programme (eligibles) and the effectiveness of the programme were first assessed. The cost and outcome sides of economic evaluation were then calculated. Finally, the results were estimated in terms of ICERs.

4.3.1 MEND Eligibles

Table 4.1 shows the estimated number of potential MEND beneficiaries, children between ages 7 and 13 with BMI≥ 91st centile, in 2010. As can be seen, approximately 1.3 million children in England were included under MEND's eligibility criteria.

Table 4.1: Estimated number of children who were eligible for MEND in England in 2010

Children with BMI≥91th UK centile in England in 2010	Boys	Girls
Ages 6-10		
BMI >95th centile	277,657	342,449
85th centile < BMI ≤ 95 centile	215,147	197,401
so, 91st centile ≤ BMI ≤ 95th centile	107,574	98,701
Total children with BMI ≥ 91st centile	385,231	441,150
Total children ages 6-10	826,380	
So, total children ages 7-10	**661,104**	
Ages 11-15		
BMI >95th centile	479,519	396,325
85th centile < BMI ≤ 95 centile	221,859	241,566
so, 91st centile ≤ BMI ≤ 95th centile	110,930	120,783
Total children with BMI ≥ 91st centile	590,449	517,108
Total children ages 11-15	1,107,557	
So, total children ages 11-13	**664,534**	
So, total children betw ages 7-13 with BMI ≥ 91st centile in 2010	**1,325,638**	

Source: modified from Zaninotto et al. (2006)

4.3.2 The short-term effectiveness of MEND

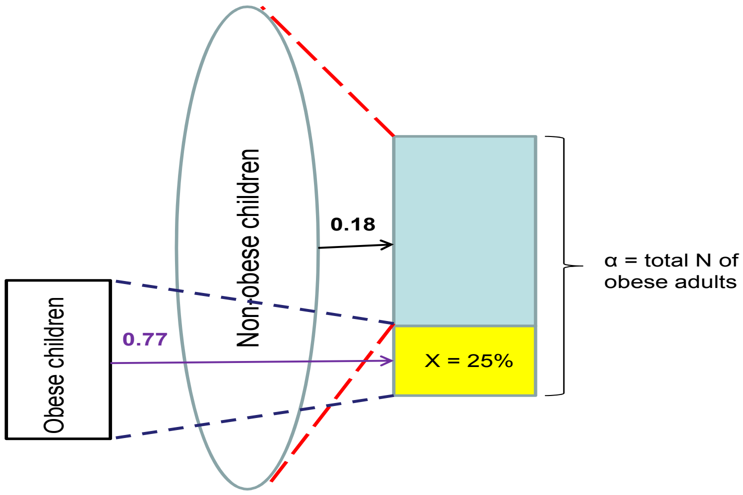
The short-term MEND rollout data were analysed to determine the effectiveness of MEND in reducing obesity cases. Almost 7,000 observations were included in the analysis, with a follow-up period of 3 months. The results of the analysis are as shown in Table 4.2. The observations were well distributed by gender, with an average age of 10.4. After the completion of the programme, the number of obese children decreased by 11% according to the international definition and 7.7% according to the UK definition.

Concerning the change in obesity status, 15.3% of obese children became non-obese according to the international definition. The percent reduction was a bit lower, 9.06%, when the UK definition was applied. In this study, the effectiveness of MEND is based on the international definition.

Table 4.2: Summary of the outcome characteristics

Characteristics	Rollout
Number of observations	6,828
Age at the start (years)	10.4 (ranging 6.5-14)
Mean time of the follow-up period (months)	2.3 (SD: 0.3)
Gender:	
Male	44.5%
Female	55.5%
Using the international BMI cut-offs	
Number of obese children (%)	
Before	5,081 (74.41%)
After	4,329 (63.4%)
Obesity reverted into non-obesity (%)	776 **(15.30%)**
Using the UK 98th BMI centile cut-offs	
Number of obese children (%)	
Before	6,070 (88.9%)
After	5,542 (81.17%)
Obesity reverted into non-obesity (%)	550 **(9.06%)**

The data from Freedman et al. (2001) indicate that 25% of obese adults age 27 were obese in the childhood period**.** Moreover, the probability of obese children becoming obese adults is 0.77, while that of non-obese children becoming obese adults is 0.18 (Freedman et al., 2001).



Source: modified from Freedman et al. (2001)

Figure 4.1: Probabilities of becoming obese in obese and non-obese children

It is seen that obese children have a higher chance of becoming obese in adulthood than do non-obese children of approximately 0.77/0.18 = 4.3 times (Calculated from the figures in Freedman et al., 2001). As shown in Figure 4.1, if obese children were treated and converted into the non-obese, the total number of obese adults attributable to childhood obesity will be reduced from 25% to 25%*0.18/0.77 = 25%/4.3 = 6%

It can therefore be estimated that if all obese children are treated and converted into the non-obese, the percentage of all obese adults would consequently be reduced by (25% - 25%/4.3) = 19%.

Given the 15.3% success rate of the programme in reverting obesity cases in children and the 19% reduction of obese adults if MEND is implemented, the percent of obesity cases reduced in this cohort by 2027 if the MEND programme were implemented in 2010 for all eligible children between ages 7 and 13, with BMI≥ 91^{st} is 15.3%*19% = 2.9%. Figure 4.2 illustrates the flow chart of the aforementioned percent reduction of obesity cases.

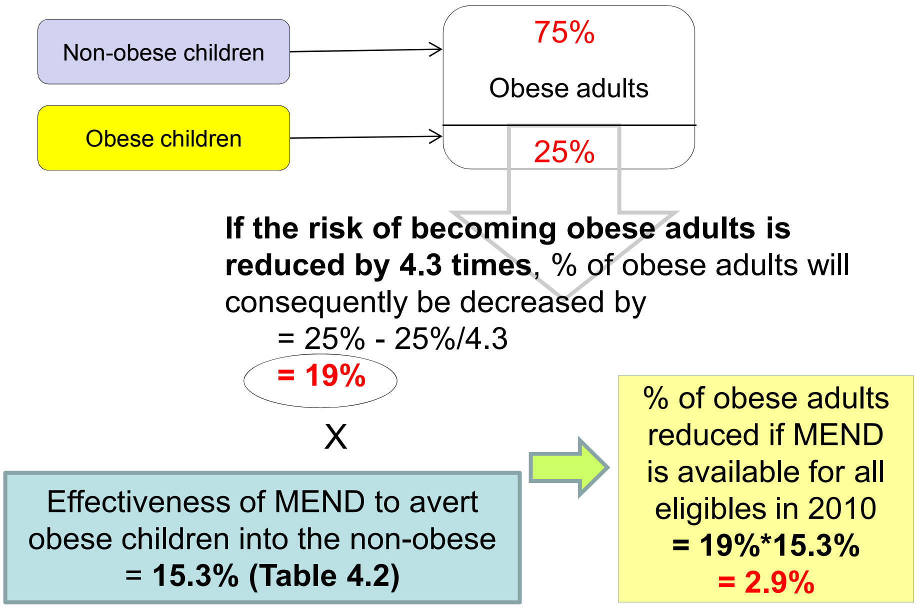


Figure 4.2: Estimation of the percent of obesity cases reduced

4.3.3 Costs

4.3.3.1 The costs of MEND

The average costs of the MEND programme are shown in Table 4.3. After adjusting by the size of the programmes and the number of participants in each, the average costs of the programme were approximately £415.77 in 2010.

Table 4.3: Summary of the programme costs

Items	Values
Children in the large rollout programme	19,600
Children in the small contract programme	108
Average costs per child	**£415.77**

Source: MEND programme

The total amount spent on MEND in 2010 is derived from the costs of MEND per child times the total number of obese children in England who are eligible for MEND. Therefore, the total costs of the MEND programme in 2010 are 1,325,638*£415.77 = £551.2 million.

4.3.3.2 The reduction of obesity-related medical costs in the future due to MEND

The estimated medical costs of obesity from 2027 to 2079 after discounting are shown in Appendix 4. As can be seen, the sum of the total medical costs of obesity between 2027 and 2079 is £93.48 billion after discounting (Butland et al., 2007).

To find the direct medical costs of obesity saved by MEND, the percentage of obese adults between ages 24 and 30 is required. This was calculated from the 2008 Health Survey for England data (NHS, 2010).

Some of the key statistics used in the calculation are displayed in Table 4.4. As shown, the total number of obese adults in 2008 was 10.14 million, whereas the estimated number of obese people between ages 16-24 and ages 25-34 were 677,000 and 1,222,000 respectively (NHS, 2010).

From Table 4.4, **the estimated number of obese people between ages 24 and 30** was (677,000/9) + (1,222,000*6/10) = **808,422**, or approximately **7.97% of the total obese adults**.

Table 4.4: Estimation of the percent of obese adults between ages 24 and 30 over the total obese adult population (≥16yrs) in England in 2008

Obese adults* (Males + Females)		**Numbers**
All (ages ≥16)		**10,146,000**
Ages 16-24		677,000
	So, obese adults age 24	75,222
Ages 25-34		1,222,000
	So, obese adults ages 25-30	733,200
	So, obese adults ages 24-30	**808,422**
So, the percent of the obese ages 24-30 over the total obese adults		**7.97%**

Source: modified from NHS (2010)

* Obesity is defined as having BMI≥30

From Figure 4.2, the reduction in obesity cases if MEND is available for all eligible children is 2.9%. Therefore, **the estimated direct medical costs of obesity saved if MEND is implemented in 2010** is 7.97%*2.9%*£93.48 billion = **£216 million** in 2010.

4.3.3.3 The net costs

The net costs of MEND are **£551.2 million – £216 million = £335.15 million.**

4.3.4 Health Outcomes

To find the long-term outcome of the programme, the total number of obese adults avoided by MEND was estimated. The additional QALYs gained were then calculated by multiplying the number of obese adults avoided with the additional QALYs gained per case avoided.

From the life table in Annex 1, the probability of survival for children aged 10 to 27 is approximately **0.994**, leading to the approximate number of survivors among eligible children in 2027 of 0.994*1,325,638 = **1,317,922**.

As shown in Table 4.3, the percent of obesity cases reverted after the programme is 15.3% using the international definition. Thus, the number of obese children reverted in 2010 = 1,317,922*15.3% = **201,642**.

From the Freedman et al. (2001) study, the probability of obese children becoming obese adults is 0.77, while that of non-obese children becoming obese adults is 0.18. Thus, if obese children are converted into non-obese, and their risk of becoming obese adults thereby lowered, the population probability of becoming obese adults will be reduced by 0.77-0.18 = **0.59**.

Therefore, it can be estimated that **implementing MEND in 2010 decreases the number of obese adults in 2027** by 201,642*0.59 = **119,862**.

In order to estimate the additional life years gained from preventing obesity in adulthood, the mean age of 27 was used to represent the age of the beneficiary in 2027.

The survival probabilities from the life table[3] in Annex 1 were applied to construct the survival curves. While the survival curves of normal-weight people were constructed using the survival probabilities of general population, the survival curves of the obese were based on the survival probabilities adjusted by the relative risks of death due to increased BMI (Flegal et al., 2005). The relative risks of death by BMI level are shown in Table 4.5.

Table 4.5: Relative risks of death at different BMI category by age

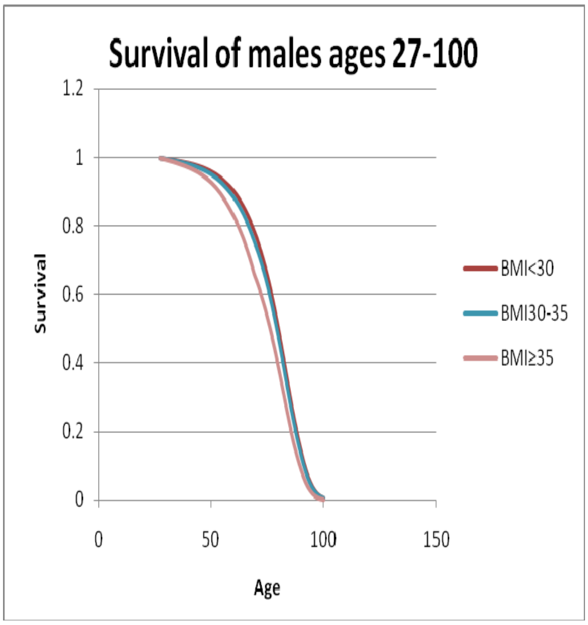
BMI level	Relative risk								
	25-59			60-69			≥70		
	RR	95%CI		RR	95%CI		RR	95%CI	
		L	U		L	U		L	U
18.5-25	1			1			1		
30- <35	**1.2**	0.84	1.72	**1.13**	0.89	1.42	**1.03**	0.91	1.17
≥35	**1.83**	1.27	2.62	**1.63**	1.16	2.3	**1.17**	0.94	1.47

Source: Modified from Flegal et al. (2005)

In this regard, the expected life years of people in all groups are estimated under the assumption that all people will die by the age of 100.

Figure 4.3 shows the survival curves of males and females between ages 27 and 100 by BMI category. As can be seen from both figures, the survival curves of people in the lower BMI group lie above those of people in the higher BMI group.

3 The model assumed the survival of the UK general population in 06-08 as the survival of the normal-weight people. Thus, the actual probability of survival in the normal-weight should be slightly higher than the estimate in the model if the probability of survival of the obese is excluded.



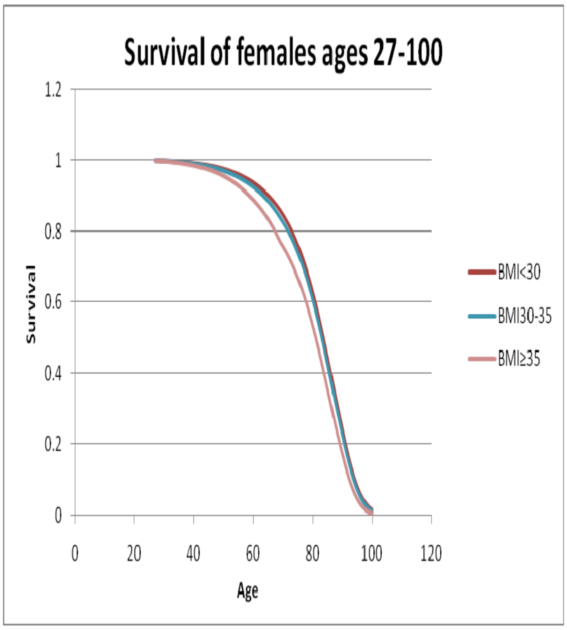


Figure 4.3: Survival curve of males and females ages 27-100 by BMI category

In the NICE guidance for technology appraisal, a discount rate of 3.5% is recommended to apply to future life years. Thus, the discounted survival is calculated by (the prob. of survival at year X)/(1.035)^(X-2010). For example, the discounted survival of a man age 27 (in 2027) = (0.999182)/(1.035)^(2027-2010) = 0.55703. Figure 4.4 shows the survival curve after discounting.

The area under the survival curve (AUC) represents the expected life years of each person in a particular BMI group. Concerning this, the differences in the AUC between the non-obese and the obese, BMI 30-<35 and BMI≥35, are the additional expected life years gained from preventing obesity. Stata is used to find the AUC. The AUC of the discounted curves in Figure 4.4 are as shown in Table 4.6.

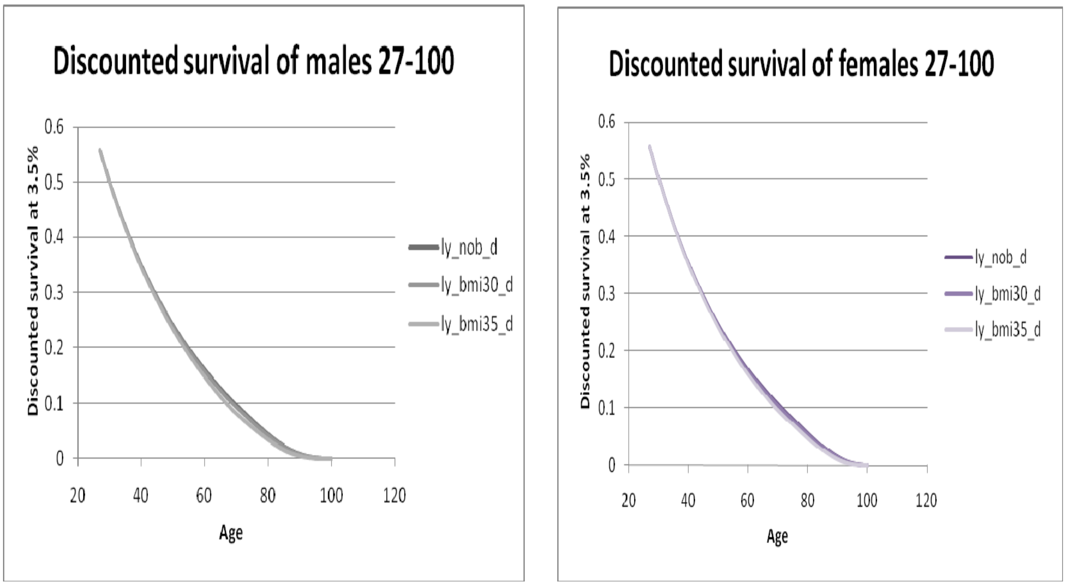


Figure 4.4: Discounted survival curve of males and females ages 27-100 by BMI category

The total lifetime QALY is calculated by multiplying the expected life years with the HRQL for each BMI group. The values of HRQL by BMI level are shown in Table 4.7. The estimated QALYs of each BMI category are also shown in Table 4.6. As seen in Table 4.6, applying the discount rate in the long-term evaluation affects the future life year estimation to a great extent. Discounting the future life years between ages 27 and 100 to base year 2010 results in an approximation of 12-13 years of life. This example demonstrates how the effect of the discount rate exponentially increases when considering the outcomes over a long period of time. It is also noted that the model does not account for the life years from the start of the programme until age 27, which is the duration in which the effect of discounting is not as high as in the later period.

Table 4.6: Expected life years and QALYs by BMI group

BMI status	Males		Females	
	LYs	QALYs	LYs	QALYs
Normal weight	12.984811	10.4268032	13.425904	10.7810009
BMI 30-<35	12.849829	9.04627962	13.331727	9.38553581
BMI ≥35	12.418577	8.09070292	13.019429	8.48215799

Table 4.7: HRQL (using EQ-5D) for different BMI group

BMI	HRQL
Normal (18.5-25)	**0.803**
30-<35	**0.704**
35-<40	0.682
≥40	0.621
≥35	**0.6515**

Source: Modified from Sach et al. (2006)

Table 4.8 lists the additional QALYs gained by the normal-weight in comparison to the obese. As shown in the table, a person with normal BMI has approximately 1.39 and 2.32 higher QALYs than does one with BMI 30-<35 and ≥35 respectively. These gains result from longer survival and higher perceived quality of life during survival.

Table 4.8: Additional QALYs gained by preventing obesity

QALYs lost when comparing with the normal-weight	QALYs
Males with BMI30-<35	1.3805
Females with BMI 30-<35	1.3955
Average BMI 30-<35	**1.3880**
Males with BMI≥35	2.3361
Females with BMI≥35	2.2988
Average BMI≥35	**2.3175**

Among obese people with BMI≥30, the proportion of people with BMI 30-<35 was approximately 69% (Flegal et al., 2005). The total QALYs gained from MEND were calculated by the sum of the additional QALYs gained from preventing obesity cases in both BMI 30-<35 and ≥35 groups.

Hence, the QALYs gained from MEND are (1.388*69%*119,862) + (2.3175*(1-69%)*119,862) = **200,905 QALYs** in 2010.

4.3.5 Economic Evaluation

From the cost and outcome analysis, the ICER of MEND is calculated as (£335.15 million)/ (200,905 QALYs) = **£1,668.2 per QALY gained.**

This ICER is considered cost-effective according to the NICE guidance (National Institute for Health and Clinical Excellence, 2008) and is also low compared with several other obesity treatments (see Table 4.9 and 4.10).

Table 4.9: Cost-effectiveness ratios for obesity interventions

Interventions	Cost/QALY Range
MEND 7-13	£1,668.2
Surgery	£6,289-8,527
Pharmacotherapy	£6,349-24,431
Non-pharmacological approaches (diet and physical activity-based)	£174-9,971
Public health interventions	£265-3,018

Source: NICE (2006); and Counterweight Project Team (2010)

Table 4.10: Cost-effectiveness ratios for child obesity intervention from Chapter 2

Articles	ICERs	Yearbase	CPI		Adjusted ICERs (to 2009)	Exchange rates to Pounds in 2009	Adjusted ICERs in Pounds
			Base	2009			
Wang et al. (2003)	US$4,305	1996	156.9*	214.5*	$5,885.4	0.64	£3,772.6
Brown et al. (2007)	US$900	2004	188.9*	214.5*	$1,022.0	0.64	£655.1
Moodie et al. (2008)	AU$4,670	2001	132.2**	166.4**	$5,878.1	0.50	£2,965.5
Magnus et al. (2009)	dominant	2001	132.2**	166.4**	dominant	0.50	dominant
Moodie et al. (2009)	AU$760,000	2001	132.2**	166.4**	$956,611.2	0.50	£482,610.3
Ananthapavan et al. (2010)	AU$4,400	2001	132.2**	166.4**	$5,538.3	0.50	£2,794.1
Moodie et al. (2010)	AU$82,000	2001	132.2**	166.4**	$103,213.3	0.50	£52,071.1
MEND	£1,668.20	2010					

Abbreviations: ICERs - incremental cost-effectiveness ratios, CPI - consumer price indices
*Based on the US consumer price indices
**Based on the Australian consumer price indices

To estimate net benefits, QALYs gained were converted into monetary units by multiplying by social willingness to pay, i.e. the threshold value of £20,000-30,000 for NICE. The benefit gained in terms of monetary units ranges from £20,000*200,905 to £30,000*200,905 = **£4,018.11 to 6,027.16** million in 2010.

Hence, when compared with the net costs of £335.15m, the MEND programme produces a net saving of **£3,682.96 to 5,692.01 million**.

4.3.6 Sensitivity analysis

The multivariate sensitivity analysis, Monte Carlo simulation, was performed to illustrate the uncertainty in the analysis. The parameters used are shown in Table 4.11. In addition to parameter uncertainty, the sensitivity analysis also tests the cost-effectiveness of MEND 7-13 when short-term programme effectiveness is diminished in the long run.

Table 4.11: Parameters used in the multivariate sensitivity analysis

Parameters	Means	95% CI		References
		Lower	Upper	
Transitional probabilities				
Prob of children with BMI ≥ 95th turning into obese adults	0.77	0.71	0.83	(Freedman et al., 2001)
Prob of children with BMI < 95th turning into obese adults	0.18	0.16	0.2	(Freedman et al., 2001)
Relative risk of death BMI18.5-25 VS BMI30-35 at ages 25-59	1.2	0.84	1.72	(Flegal et al., 2005)
Relative risk of death BMI18.5-25 VS BMI30-35 at ages 60-69	1.13	0.89	1.42	(Flegal et al., 2005)
Relative risk of death BMI18.5-25 VS BMI30-35 at ages ≥70	1.03	0.91	1.17	(Flegal et al., 2005)
Relative risk of death BMI18.5-25 VS BMI ≥35 at ages 25-60	1.83	1.27	2.62	(Flegal et al., 2005)
Relative risk of death BMI18.5-25 VS BMI ≥35 at ages 60-70	1.63	1.16	2.3	(Flegal et al., 2005)
Relative risk of death BMI18.5-25 VS BMI ≥35 at ages ≥70	1.17	0.94	1.47	(Flegal et al., 2005)
EQ5D quality of life for normal-weight people	0.803	0.819	0.787	(Sach et al., 2006)
EQ5D quality of life for people BMI30-35	0.704	0.743	0.665	(Sach et al., 2006)
EQ5D quality of life for people BMI≥35	0.667	0.728	0.605	(Sach et al., 2006)

Table 4.12 shows the sensitivity analysis of the programme. As demonstrated, the MEND programme is still cost-effective even when only 25% of programme effectiveness is maintained, with the willingness to pay of £20,000 in the lower 95% confidence interval.

The threshold analysis is illustrated in Table 4.13 and Figure 4.5. As seen, the maintenance of a short-term effectiveness of 9.14% and 13.47% would be required in order to make MEND 7-13 cost-effective with respect to the thresholds of £20,000 and £30,000 per QALY, respectively.

Table 4.12: Results of the sensitivity analysis

Parameter	Values		
	Base case	**95% CI**	
Health outcomes			
Males			
Estimated QALYs of a normal-weight person aged 27-100	10.4297	10.4185	10.4409
Estimated QALYs of an obese person with BMI 30-35 aged 27-100	9.0406	9.0134	9.0679
Estimated QALYs of an obese person with BMI ≥35 aged 27-100	8.2789	8.2321	8.3257
QALYs gained from reverting BMI30-35 into the normal range	1.3891	1.3596	1.4185
QALYs gained from reverting BMI ≥35 into the normal range	2.1508	2.1027	2.1990
Females			
Estimated QALYs of a normal-weight person aged 27-100	10.77439	10.76307	10.78572
Estimated QALYs of an obese person with BMI 30-35 aged 27-100	9.378684	9.351939	9.405428
Estimated QALYs of an obese person with BMI ≥35 aged 27-100	8.661134	8.613464	8.708805
QALYs gained from reverting BMI30-35 into the normal range	1.395709	1.366674	1.424743
QALYs gained from reverting BMI ≥35 into the normal range	2.113258	2.064267	2.16225
Males + Females			
QALYs gained from reverting BMI30-35 into the normal range	*1.3924*	*1.3631*	*1.4216*
QALYs gained from reverting BMI ≥35 into the normal range	*2.1321*	*2.0835*	*2.1806*
Number of surviving children in 2027	1,317,922		
Total costs of MEND (£million)	551.16		
Number of obese adults attributable to childhood obesity	783,415	672,140	883,008
Sensitivity analysis			
If the maintenance of effect = 100%	15.30%		
The medical costs saved due to the programme (£million)	216.01		
Net costs of MEND (£million)	335.15		
Number of reduced obese adults	119,863	102,837	135,100
QALYs gained	194,379	163,147	223,849
ICERs (£/QALY)	*1,724.2*	*2,054.3*	*1,497.2*
Net benefits at a threshold of £20,000 (£million)	*3,552.4*	*2,927.8*	*4,141.8*
Net benefits at a threshold of £30,000 (£million)	*5,496.2*	*4,559.3*	*6,380.3*
If the maintenance of effect = 50%	7.65%		
The medical costs saved due to the programme (£million)	108.01		
Net costs of MEND (£million)	443.15		
Number of reduced obese adults	59,931	51,419	67,550
QALYs gained	97,190	81,574	111,925

Parameter	Values		
	Base case	95% CI	
ICERs (£/QALY)	*4,559.7*	*5,432.6*	*3,959.4*
Net benefits at a threshold of £20,000 (£million)	*1,500.6*	*1,188.3*	*1,795.3*
Net benefits at a threshold of £30,000 (£million)	*2,472.5*	*2,004.1*	*2,914.6*
If the maintenance of effect = 25%	3.83%		
The medical costs saved due to the programme (£million)	54.00		
Net costs of MEND (£million)	497.16		
Number of reduced obese adults	29,966	25,709	33,775
QALYs gained	48,595	40,787	55,962
ICERs (£/QALY)	*10,230.7*	*12,189.2*	*8,883.8*
Net benefits at a threshold of £20,000 (£million)	*474.7*	*318.6*	*622.1*
Net benefits at a threshold of £30,000 (£million)	*960.7*	*726.4*	*1,181.7*
If the maintenance of effect = 12.5%	1.91%		
The medical costs saved due to the programme (£million)	27.00		
Net costs of MEND (£million)	524.16		
Number of reduced obese adults	14,983	12,855	16,888
QALYs gained	24,297	20,393	27,981
ICERs (£/QALY)	*21,572.6*	*25,702.4*	*18,732.6*
Net benefits at a threshold of £20,000 (£million)	*-38.2*	*-116.3*	*35.5*
Net benefits at a threshold of £30,000 (£million)	*204.8*	*87.6*	*315.3*

Table 4.13: Threshold analysis for the maintenance of the programme effectiveness in the long term in order to break even with the NICE threshold

	% of effect maintenance		
		95% CI	
NICE threshold	Base case	Upper	Lower
£20,000 per QALY	13.47%	15.84%	11.75%
£30,000 per QALY	9.14%	10.35%	7.95%

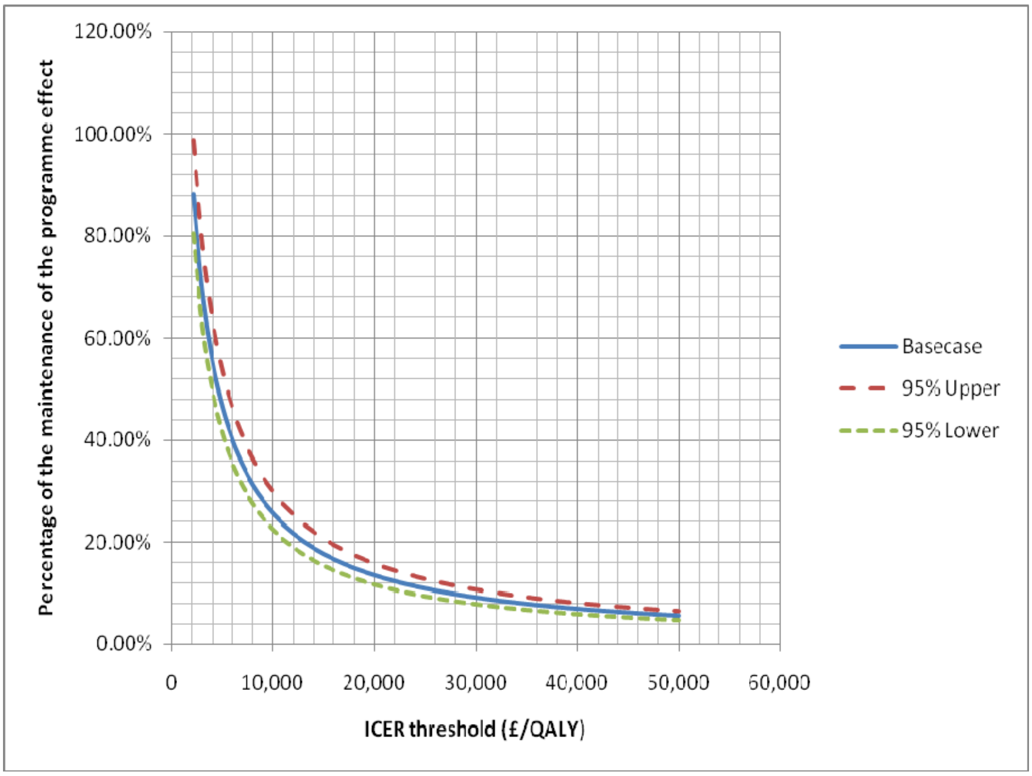


Figure 4.5: Threshold analysis showing the percentage of programme effect maintenance required in order to break even with the ICER threshold

A sensitivity analysis of time preference is also performed, using variations of the discount rate applied to the costs and health outcomes of MEND 7-13, in order to demonstrate impact of the discount rate on the ICER estimates. Table 4.14 shows the ICERs of MEND 7-13 when different discount rates, ranging from 0-7%, were employed. Not only were the equal rates of discounting applied, but selected rates of differential discounting were also demonstrated, including the discount rates of 6% for costs and 0% and 1.5% for health outcomes (National Institute for Clinical Excellence, 2001). As illustrated, the effect of the discount rate on the ICER estimation is very crucial. While the intervention can be cost-saving if no discounting is applied, comparing the estimates using the equal discount rates of 1.5% and 7% results in an almost 40-fold increase in the ICER.

Table 4.14: Sensitivity analysis of the discount rate applied to the costs and health outcomes

Discount rate		Total medical costs saved by MEND in 2010 (million)	Total incremental costs (MEND costs - total medical costs saved) (million)	Total QALYs gained	ICERs (£/QALY)
Cost	**Outcome**				
0.0%	0.0%	895.96	-344.80	863,359	**cost-saving**
1.5%	1.5%	468.23	82.93	440,290	**188.35**
3.0%	3.0%	259.67	291.49	241,760	**1,205.71**
3.5%	3.5%	216.01	335.15	200,905	**1,668.20**
5.0%	5.0%	128.68	422.48	119,815	**3,526.14**
6.0%	6.0%	93.49	457.67	87,305	**5,242.15**
7.0%	7.0%	69.19	481.97	64,835	**7,433.91**
6.0%	0.0%	93.49	457.67	863,359	**530.10**
6.0%	1.5%	93.49	457.67	440,290	**1,039.46**

In addition to the multivariate sensitivity analysis, maintenance of the effect, and the discount rate, sensitivity analyses also estimate ICERs by subgroup populations based on the base case parameters. The programme effectiveness of particular groups was used to calculate ICERs, as can be seen in Table 4.15. The results indicated that certain characteristics are potentially associated with programme effectiveness, leading to changes in cost-effectiveness ratios. Characteristics associated with high benefit uptakes of the programme include being White, having parents who own a home, not being from a single-parent household, and not living in a deprived area.

Table 4.15: Sensitivity analysis of the programme effectiveness and ICERs by subgroups (IOTF definition)

Subgroup	N of obese children	Percent of obesity reverted*	ICERs
Total participants	5,073	15.30%	£1,668.19
Ethnicity			
White	2,495	15.75%	£1,589.81
Asian	273	10.26%	£3,015.82
Black	177	7.91%	£4,231.23
Others	283	12.01%	£2,419.71
Parents owning home or not			
No	1,526	12.45%	£2,296.20
Yes	1,692	16.37%	£1,488.88
Having single parent or not			
No	2,167	15.23%	£1,680.80
Yes	1,044	13.12%	£2,124.03
Level of deprivation (quintile)			
1st quintile (least deprived)	652	19.48%	£1,079.52
2nd quintile	779	17.20%	£1,365.15
3rd quintile	876	18.72%	£1,167.00
4th quintile	1,060	14.43%	£1,833.59
5th quintile (most deprived)	1,253	10.61%	£2,880.87

Note: Percent of obesity reverted is calculated from complete-case samples.
*Using the IOTF definition

Table 4.16 shows the subgroup analysis when the 98th BMI centile cut-offs of the UK 1990 reference chart for age and sex are applied instead of the IOTF definition of childhood obesity. Comparing with the results in Table 4.15, it is seen that the effectiveness of the programme is affected immensely by the definition of childhood obesity, falling from 15.3% to 9% when the UK definition is applied instead. When applying the UK definition, Asian children turn out to be the group that most highly responds to the intervention in relation to their peers. Nevertheless, the direction of association remains the same for homeowners, single parents, and deprivation score, regardless of the definition used – that is, MEND 7-13 tends to be less effective in children whose parents do not own their houses, have single parents, and live in a deprived area.

Table 4.16: Sensitivity analysis of the programme effectiveness and ICERs by subgroups (UK 1990 definition)

Subgroup	N of obese children	Percent of obesity reverted*	ICERs
Total participants	6,062	9.07%	£3,552.57
Ethnicity			
White	2,953	8.70%	£3,749.38
Asian	349	11.46%	£2,587.44
Black	211	4.74%	£7,780.03
Others	332	8.73%	£3,732.80
Parents owning home or not			
No	1,768	7.18%	£4,770.74
Yes	2,063	9.84%	£3,190.44
Having single parent or not			
No	2,588	9.23%	£3,472.35
Yes	1,225	7.27%	£4,698.37
Level of deprivation (quintile)			
1st quintile (least deprived)	791	12.01%	£2,419.71
2nd quintile	955	11.41%	£2,603.49
3rd quintile	1,070	8.88%	£3,651.58
4th quintile	1,268	8.36%	£3,945.59
5th quintile (most deprived)	1,440	7.22%	£4,738.35

Note: Percent of obesity reverted is calculated from complete-case samples.
*Using the UK 1990 definition

The sensitivity analysis also explores the impact of the eligibility criteria on the ICER of MEND 7-13 by assuming that only children aged 7-13 with BMI $\geq 98^{th}$ centile were eligible for the programme. As programme effectiveness is defined as the reversion of obesity status, narrowing down the eligibility criteria from BMI $\geq 91^{st}$ centile to BMI $\geq 98^{th}$ centile would only affect the intervention cost in the ICER calculation. The obesity-related medical costs saved due to MEND 7-13, £216 million; the programme effectiveness, 15.3%; and the QALYs gained, 200,905 QALYs; still remain the same in the estimation.

Table 4.17 shows the estimation of the number of children between the ages of 7 and 13 with BMI $\geq 98^{th}$ centile in England in 2010. As demonstrated, 408,636 children would be eligible for MEND 7-13 when new eligibility criteria are applied, resulting in total intervention costs of 408,636*£415.77 = £169.9 million. This would make the programme cost-saving when considering the total intervention costs less the medical costs saved due to the MEND 7-13, with an approximate saving of £169.9m -£216m = -£46.1 million.

Table 4.17: Estimated number of children between ages 7 and 13 with BMI ≥98th centile in England in 2010

Children with BMI≥98th UK centile in England in 2010		Boys	Girls
Ages 6-10			
	BMI >95th centile	277,657	342,449
	Total children with BMI ≥ 98th centile	111,063	136,980
	Total children ages 6-10	248,042	
	So, total children ages 7-10	**198,434**	
Ages 11-15			
	BMI >95th centile	479,519	396,325
	Total children with BMI ≥ 98th centile	191,808	158,530
	Total children ages 11-15	350,338	
	So, total children ages 11-13	**210,203**	
So, total children betw ages 7-13 c/ BMI ≥ 91st centile in 2010		**408,636**	

It can be seen from the sensitivity analysis that maintenance of programme effectiveness is an important factor in determining the long-term cost-effectiveness of MEND 7-13. Only an effect maintenance greater than 9.1% and 13.5% will make the programme cost-effective in the long run, with respect to the ICER thresholds of £20,000 and £30,000 per QALY, respectively. The discount rate applied in the cost and health outcome estimation, however, affects the ICER estimate to a considerable extent: ICERs increase when higher discount rates are applied, provided that the same rate is applied to both costs and outcomes. Moreover, the definition of obesity applied in the analysis plays a crucial role in calculating the effectiveness of MEND 7-13. Inconsistency of the definitions used in the analysis should be considered when interpreting results. In addition, the results showed that the programme is more cost-effective in particular subgroups—namely, White and/or Asian, children whose parents own their own houses, children who do not have single parents, and children whose houses are in less deprived areas. Programme improvement is urged, with the aim of ensuring equity in programme effectiveness among all participants.

4.4 Discussion

This chapter introduces an economic modelling approach that improves upon existing ones in the approximation of the long-term health outcome of child obesity interventions. The assessment criteria for economic evaluations from Drummond et al. (2005) were applied to the cost-effectiveness analysis of MEND 7-13, as demonstrated in Table 4.18.

The model used in this chapter improves the health outcome estimates derived from existing techniques, as used by Wang et al. (2003) and Haby et al. (2006). A shared weakness in all models is the failure to account for the dynamic change in BMI and obesity status before and after the starting point. Both models focus on the change in obesity status/BMI in childhood from the starting point to the end of the follow-up period without considering the obesity status/BMI before and during the intervention. As a result, the risks of ex-obese children who reduced weight and of obese children who gained weight during the follow-up period cannot be taken into account by the models.

Considering the first model, two main drawbacks of Wang et al.'s model are as follows: 1) limiting the range of future health considered to the 25 years between ages 40 and 64, rather than covering the entire adulthood period; and 2) the failure to account for different severities of obesity in children when determining quality of life. Note that Wang et al. have developed another model, which accounts for health from age 40 to the end of the life expectancy (Wang et al., 2010). This diminishes the first weakness mentioned, though accounting for future health prior to age 40 could further improve the model. Nevertheless, the latter concern still remains (Wang et al., 2010). The model proposed in this chapter accounts for these weaknesses, as: 1) the duration of future outcome is projected from age 27 to the entire lifespan or age 100; and 2) obese children are categorized into those with BMI between 30-<35 and ≥ 35, and are assigned suitable quality of life for different severities.

Similarly, Haby et al.'s model also contains two points of concern: 1) the model accounts for the benefits of BMI reduction without regard to obesity status, which has the potential to be a critical fallacy in health outcome estimation, as was pointed out in Chapter 2; and 2) the unit of health outcome from Haby et al.'s model is presented in terms of disability-adjusted life years (DALYs), not QALYs. Using DALYs to assess the health status of the obese may not be appropriate due to their inability to account for the complications of the disease, as discussed in Chapter 1, leading to problems in performing cost-utility analysis. The model presented in this chapter overcomes these issues because: 1) the future health projection is based on obesity status, as also seen in Wang et al. (Wang et al., 2003, Wang et al., 2010); and 2) the health outcome is calculated in terms of QALYs.

Despite several advantages over the previous models, certain issues remain of concern when using the new economic modelling approach.

First, the technique needs to be adjusted if the starting age of children changes. The model in this chapter estimates the future health of children aged 10 (7-13) using the probabilities in Freedman et al. (2001), which corresponds to the MEND sample. However, it may be problematic if the starting age group changes. One way to deal with this is to use probabilities from another paper that provide more comprehensive transitional probabilities for the model (Guo et al., 2002). However, some problems arise due to the small sample size contained in the study (347 samples in Guo et al. compared with 2,617 sample in Freedman et al.), which could produce uncertainty in the results.

Second, the health benefits gained from obesity reduction are based on the difference in life expectancy between the normal population and the normal population with increased risk of death from obesity. The life expectancy of the normal population is applied to the non-obese adults, as the mortality risk of obesity in the population is accounted for in the health outcome estimation. The model thus compares the relative risk of mortality in the obese population to that in the normal population, which includes both the obese and non-obese rather than the non-obese alone. However, the life expectancy of the obese group should ideally be derived from the obese population, not the normal population with increased risk of death from obesity, as illustrated by Figure 4.6. The author acknowledges the data limitation, which leads to a small difference in estimating future health benefits due to double counting the obesity-related mortality risk in the normal population.

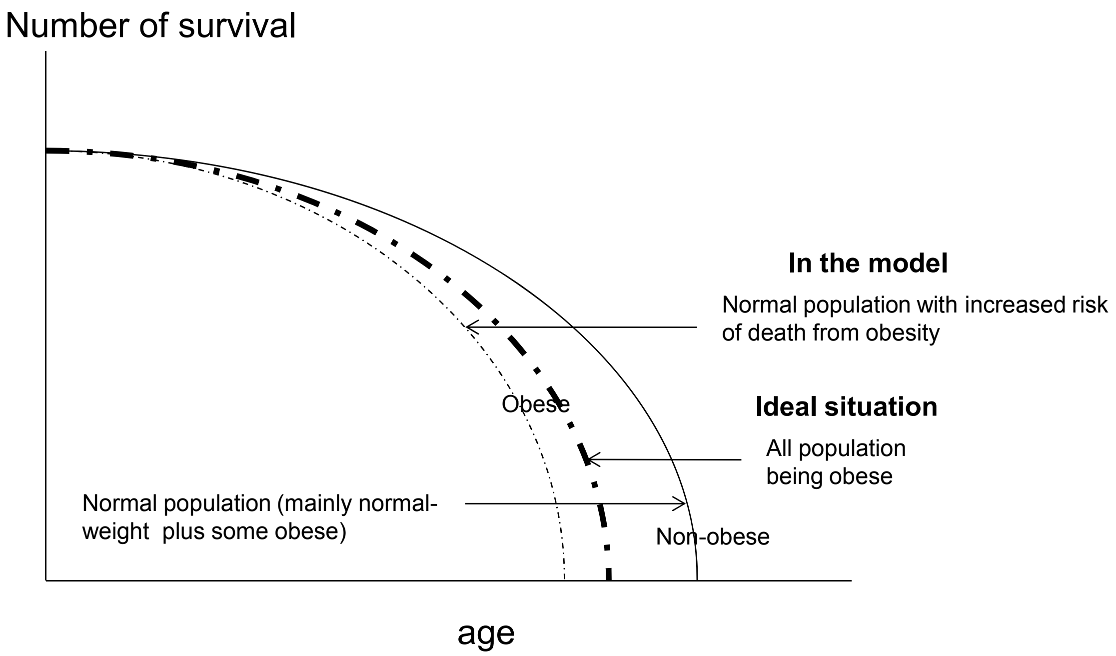


Figure 4.6: Comparing the assumption made in the model versus that in ideal situation

Third, concerning the effectiveness of MEND, the definition used of programme success does not account for the short-term health and other social benefits gained from attending the programme. These short-term benefits include reduced risk of illness in childhood and the consequent decrease in health care costs. The benefits of the programme to parents who participated with their children were also omitted. Moreover, this analysis fails to consider the health benefits delivered to obese children whose BMI is reduced as a result of attending the programme, yet not sufficiently to be re-classified as 'overweight'. Any reduction in obesity will have long-term cost benefits irrespective of whether the beneficiaries cross the obese/overweight cut-off. It is likely that the ICER of the intervention would be revised downward if consideration were given to these factors.

Lastly, the model assumes the obesity status of the samples to be persistent after the age of 27. However, obesity status can still change even in adulthood, resulting in an altered mortality risk. Failure to account for the possible fluctuation in mortality risks resulting from the change in obesity status inevitably affects the ICER estimate. Future research estimating the impact of obesity status change in the ex-obese and ex-non-obese is required to determine the degree of uncertainty attributable to this dynamic.

In addition to questions regarding health outcome estimation, concerns also arise regarding the cost estimate.

Firstly, the costs of MEND are from MEND Central Ltd. The author has limited access to data and all cost information was collected and summarized by MEND staff. Even though the methods used to calculate the costs of MEND have been satisfactorily assessed and verified by the author, a modicum of concern still remains, as the costs of the programme cannot be clearly derived in detail.

Secondly, the projected future costs of obesity in this study came from Butland et al. (2007) which assumes that the prevalence of childhood obesity will increase until 2050. However, future obesity trends could increase at a lower rate or even decrease, which would make the ICER higher than estimated.

With regard to the ICER, the results of this study show that the MEND programme is considered cost-effective when comparing its ICER with the standard threshold from the NICE guidance. In this regard, it is worth noting the assumptions that had to be made to reach this conclusion, which could be tested in future research.

The major concern pertaining to the assumptions made in the analysis is the sustainability of the programme effect. Data on the effectiveness of MEND came from a cohort study that followed children for 2-3 months on average. While estimating the ICER based on the short-term effect represents an optimistic outcome, it makes sense to perform a sensitivity analysis showing alternative scenarios where long-term maintenance of the programme effect is reduced, together with the threshold analysis. As the ICER in this case was derived from the 10-week effectiveness of the programme, projecting to the next 70 years, the estimated ICER will inevitably contain a number of biases, regardless of the robustness of the modelling techniques. Using this ICER must therefore be done with caution, especially in the policy decision making process.

Moreover, the possibility of diminishing marginal benefits with programme roll-out could potentially bias the results. As the economic model assumes that MEND 7-13 will be available to more than 1.3 million English children aged 7-13 in 2010, using the programme effectiveness derived from the rollout data containing 6,828 participants would result in very optimistic estimates. This is because it is unlikely that all obese children, along with their parents, are willing and able to participate in the programme. As a consequence, the programme attendance rate would be lower than that in the rollout phase, resulting in a decrease in programme effectiveness. Another practical constraint that would cause diminishing marginal benefits is the capacity of the staff, as the programme would require a considerable number of staff in order to deliver services to 1.3 million children and their parents. This would lead to practical limitations in terms of training sessions and quality control. Also of note is the inconsistency in programme effectiveness as a result of various definitions of childhood obesity being applied. As seen in the sensitivity analysis, the effectiveness of MEND 7-13 changes to some extent when using the 2 different definitions. Extrapolating the ICERs of child obesity interventions should be therefore done with careful consideration of any uncertainty resulting from different definitions of childhood obesity.

In addition to treating obesity, MEND 7-13 may also prevent overweight children from becoming obese in adulthood. However, apart from the lack of evidence, it is also beyond the scope of this analysis to consider the cost-effectiveness of obesity cases reverted in this manner, in the absence of data on the relationship between adult obesity and childhood overweight. It is likely that the ICER of the intervention could be revised downward if this factor were considered.

The sensitivity analysis demonstrates the effect of the discount rate applied in the economic evaluation on the ICER estimate. The selection of the discount rate should be done carefully when conducting the cost-effectiveness analysis in order to minimize the time preference bias.

With regard to results from the subgroup analysis, issues regarding health disparities arise. The results reveal not only the target populations that MEND 7-13 should focus on in order to maximize benefits, but also the particular characteristics of children who need more attention in order to achieve the goal of the intervention. Although the programme can increase its effectiveness by recruiting only children who are more responsive to the intervention, e.g., from the least deprived areas, some concern should be shown for health disparity issues. That is, obese children should have equal access to publicly-funded treatment for obesity regardless of their socio-economic-demographic characteristics. If MEND 7-13 were to target only obese children who will make the intervention most effective, it would be hard to ensure that the rest of the obese children are able to access obesity treatment from other sources. Therefore, it seems appropriate for MEND 7-13 to continue its current practice of using only obesity status and age for the eligibility criteria.

Another concern regards the baseline prevalence of childhood obesity. The best available data to estimate obesity in children in the model were based on figures from 1973-4, when the prevalence of childhood obesity was much lower than at present. If the proportion of obese adults attributable to childhood obesity is greater than estimated (25%), which would produce larger benefits from MEND, the estimated ICER in this study is likely to be too high.

As the economic evaluation employs the health care payer perspective in the analysis, the non-health-related components were excluded from the cost and health outcome estimations. Alternative perspectives, such as the societal or public health perspectives, would take these elements into account, leading then to revised ICERs. For example, there is an opportunity cost of time spent to attend the programme for both children and parents that would potentially increase the estimated ICER if it were to be incorporated into the analysis; the magnitude of this effect may vary depending upon the methods used to value the time cost. Another example is the cost-saving from non-NHS budgets that would revise the estimated ICER downward if taken into account. The magnitude of this effect, however, is unknown due mainly to the lack of supportive evidence.

The approach taken in this analysis makes a mixture of assumptions, some of which are favourable and some unfavourable to the MEND programme. As a result, the ICER estimated in this chapter should be used with caution in making policy decisions, due in large part to the effectiveness of the programme being derived from short-term data. Long-term follow-up data would be needed to determine both the sustainability and the cost-effectiveness of the programme.

Table 4.18: Assessing economic evaluation of MEND 7-13 from Drummond et al. (2005)

Check-list for assessing economic evaluation	CEA of MEND			Justification
	Yes	No	Can't tell	
1. was a well defined question posed in an answerable form?	A			The study explores the CEA of MEND 7-13 compared with status quo.
2. was a comprehensive description of the competing alternatives given (that is, can you tell who did what, to whom, where, and how often)?	A			See in Methodology.
3. was there evidence that the programme's effectiveness had been established?	A			The evidence of the programme's effectiveness came from the MEND 7-13 rollout phase.
4. were all the important and relevant costs and consequences for each alternative identified?	O	C		The author does not have access to the intervention costs of MEND 7-13. Information regarding cost collection and the total costs came from personal communication with MEND.
5. were costs and consequences measured accurately in appropriate physical units?	A			Costs were measured in Pound, while health outcomes were changes in weight, which were transformed into QALYs gained.
6. were costs and consequences valued credibly?		O	C	Limited access to the intervention costs is an issue affecting credibility of the cost data. Regarding the outcome, the major concern comes from the short duration of follow-up.
7. were costs and consequences adjusted for differential timing?	A			Both costs and outcomes were discounted at 3.5% per annum.
8. was an incremental analysis of costs and consequences of alternatives performed?	A			The programme is compared with status quo or flat comparator.
9. was allowance made for uncertainty in the estimates of costs and consequences?	A			See in Sensitivity analysis.
10. did the presentation and discussion of study results include all issues of concern to users?	A			See in Discussion.

C = for the cost side; O = for the outcome side; A = all (both cost and outcome sides)

Chapter 5: Conclusions

5.1 Overview

In response to the research questions initially presented in Chapter 1, this chapter summarizes the contents, main findings, and contributions to knowledge of this thesis, along with identifying future research possibilities. The novelty of this thesis covers 3 main areas:

1. The thesis identifies the current evidence base for the long-term cost-effectiveness of child obesity interventions. The author examines the quality of the evidence by conducting a literature review of the cost-effectiveness of child obesity interventions, as appears in Chapter 2. A number of aspects pertaining to economic evaluation are explored, including economic modelling techniques and other relevant issues. The results of the review not only reveal the complexity of the economic modelling required, but also point out the dearth of basic data essential for the determination of the cost-effectiveness of interventions to treat or prevent obesity in childhood and adolescents. This highlights the need to improve both the quality and the quantity of cost-effectiveness research by ensuring that well-designed effectiveness studies always include thorough costing of the interventions being evaluated.

2. The author critically assesses the design and conceptual basis of the MEND 7-13 programme, a child obesity intervention used as a case study in the cost-effectiveness section of this thesis. Developed in 2001, MEND 7-13 has examined its programme effectiveness in a feasibility test and a randomized controlled trial. However, in an attempt to scale up the programme coverage with cost-effectiveness in mind, MEND 7-13 lowered the staff qualifications requirements and expanded the eligibility criteria for programme participation. No evidence has as yet supported the effectiveness of the programme at this stage. Chapter 3 explores the supportive evidence for programme effectiveness, in terms of both process and outcome, with critical assessment provided. Moreover, characteristics influencing the benefit uptakes of the programme participants were examined using statistical

analyses. The results indicate a lack of evidence to support the effectiveness of MEND 7-13 during the rollout stage, reflecting the weakness of the programme's design and implementation processes. Finally, the author provides some recommendations for the future improvement of MEND 7-13.

3. The author improves an economic modelling technique, based on existing models, to assess the cost-effectiveness of child obesity interventions, especially for the long-term estimation of QALYs, as illustrated in Chapter 4. The author also calculates the incremental cost-effectiveness ratio (ICER) of MEND 7-13, as a case study, with sensitivity analysis being carried out to account for the uncertainty of the parameters used in the ICER calculation. Discussions on the developed model, showing the advantages and drawbacks of the technique, and issues concerning the ICER of MEND 7-13, follow. Due to a lack of data on the costs of the programme, as well as the poor quality and incompleteness of the effectiveness data, the author concludes that the use of the cost-effectiveness results requires careful and transparent management of basic data and its modelling if policy choices are not to be corrupted by the results of poor analysis.

The first part of this chapter presents a brief summary of the chapters in this thesis. Afterward, the author discusses potential variations in the analyses, with justifications for the selected choices. Research implications follow, with particular emphasis on the strengths and weaknesses of each type of economic model; suggestions to improve accuracy in the estimates are also provided. Regarding this, the author compares the model developed in Chapter 4 to other existing models in order to illustrate its contribution to existing knowledge. An economic model that combines advantages of the other models is proposed for future research. Gaps and additional future research opportunities are also pointed out, with emphasis on the MEND 7-13 programme and economic modelling. Lastly, policy implications are considered, followed by final conclusions.

5.2 Chapter Summary

Chapter 1 demonstrates the rising problem of childhood obesity worldwide. In the UK, the prevalence of childhood obesity was estimated to increase from 8-10% in 2004 to 15% in 2025 (Butland et al., 2007). However, it may be argued that the projected prevalence of childhood obesity has been overestimated, as some evidence reveals a flattening trend during the past decade (Stamatakis et al., 2009). Despite a lack data on childhood obesity in particular, the UK data point out the impact of social class on the prevalence of adulthood obesity; those in lower social classes are more prone to obesity compared to those in higher ones (Butland et al., 2007). Due to limited data, the association between ethnicity and the prevalence of obesity is unavailable.

The author introduces a variety of definitions of childhood obesity. Conceptually, all definitions compare children's BMIs to BMI centile cut-offs; what is different is the population reference referred to by the definition: population-specific, such as the UK or US population, or an international population, i.e., the International Obesity Task Force (IOTF) definition. The author suggests that using the population-specific definition has an advantage when there is not much ethnic variation in the population of interest, whereas in populations with considerable ethnic diversity, the IOTF definition is preferred due to convenience and better comparability across studies.

The author also examines the factors associated with the development of obesity. Advancements in technology have been raised as a main factor contributing to the increased prevalence of obesity (Lakdawalla and Philipson, 2002, Finkelstein et al., 2005). This is because advanced technology has made routine work become more efficient and less strenuous, causing passive exercise during work to be diminished. Attaining physical activity thus becomes more costly, as it requires effort to find extra time, and perhaps pay, for opportunities to exercise. In addition to the impact on physical activity, advanced technology also affects energy balance through the food production and price. The concept of price elasticity of demand was introduced to explain price changes and consumption (Krugman and Wells, 2009). The author concludes that overall food prices have remained the same during the past few decades, while a rise or drop in prices can be observed when considering subcategories separately; fruit and vegetables prices have increased, whereas the prices of sugar, sweets, carbonated drinks have decreased. The author also shows that the effect of food prices on consumption not only varies across social classes, but also depends on consumer behaviour. An example of this can be seen from a study revealing the effects on soft drink consumption of a tax to increase soft drink prices (Finkelstein et al., 2010). Finkelstein et al. (2010) show that an increase in soft drink prices affects middle-income households the most, whereas only little impact is seen in the households of the richest and poorest quartiles.

Afterward, the author explores other factors associated with the development of obesity, including genetics, mental health factors, socio-economic status, demographic factors – age, gender, and ethnicity – sedentary behaviour, and physical activity. Overall, the author points out that the development of childhood obesity is a complex interaction involving individual and environmental factors. Stakeholders' collaboration, using multidisciplinary approaches to tackle the multi-faceted roots of the problem, is suggested to achieve the goal of reducing childhood obesity prevalence.

The chapter continues on to illustrate the undesirable health effects of obesity in childhood and adulthood, highlighting the adverse consequences of childhood obesity for physical and psychological health, as well as its social impact. Obesity in youths leads to a number of adverse health consequences in both short and long terms, with examples including cardiovascular risks, diabetes, asthma, and obesity persistence in adulthood (Reilly et al., 2003, Han et al., 2010). The risk of obesity persistence is of particular concern, as adult obesity is associated with many chronic diseases and a rise in morbidity and mortality risks (Flegal et al., 2005). In addition to physical health, the author shows that childhood obesity also affects social well-being, in the form of bullying (Puhl and Latner, 2007). Obese children are more likely to bully and be bullied by those around them, including perhaps their parents. Unlike physical and social impact, however, no clear evidence has established the relationship between obesity and adverse psychological effects (Wardle and Cooke, 2005).

The author then explores the measures and interventions used to tackle childhood obesity, as well as the supportive theories behind them. With regard to public policies aimed at reducing obesity, sugar-sweetened beverage (SSB) taxation and fruit and vegetable subsidy policies are discussed. Despite revenue-generating benefits, some evidence shows that SSB taxation does not seem to be effective in reducing obesity in children and adolescents for 3 reasons. First, the effect size of the SSB tax on weight is tiny (Vartanian et al., 2007, Forshee et al., 2008). Second, consumers may switch to cheaper calorie-dense beverages if SSB prices are increased (Brownell and Frieden, 2009). Finally, the price effect of the policy targets not the obese, but those most sensitive to price effects. In other words, this policy has poor effect discrimination. In contrast, evidence shows that subsidizing fruits and vegetables may be an effective policy to tackle obesity (Powell and Bao, 2009). However, subsidizing all fruit and vegetable prices in the market may not be an efficient strategy. Instead, a delivery method to improve policy efficiency, the direct payment method, is suggested due to the method's effectiveness (Lagarde et al., 2009). The author emphasizes the necessity of evaluating the collateral effects of public policies alongside their benefits, as no/few benefits come without a cost.

After discussing public policies related to obesity, the author introduces the concept of human investment, as demonstrated by Grossman's model. It is implied from the model that the treatment and prevention of obesity are ways to invest in health, which can increase health stock and decrease future medical expenditures. Consequently, obesity should be treated as early as possible to maximize health and financial benefits. However, it is uncertain whether or not the costs of health investment actually outweigh obesity-related medical expenses. This raises questions as to whether these types of intervention are actually cost-effective and thus worth public investment. Economic evaluation is introduced as a tool to inform decision-making in this regard. A brief theoretical framework of economic evaluation is presented, following by a discussion on health outcome units: QALYs versus DALYs.

In addition to the adverse health and social consequences explored in Chapter 1, the economic burden of obesity is another point of concern from the public perspective, as a huge proportion of fiscal budgets in many countries has been spent on medical expenditures. Chapter 2 starts by providing brief evidence of the economic burden of obesity from the existing literature. Studies from several countries, such as the US, UK, and many European countries, reveal that 2-7% of national healthcare spending is attributable to obesity. In the UK, it was estimated in 1998 that 0.3% of its GDP was spent on the direct and indirect costs of obesity (The Comptroller and Auditor General, 2001).

The future costs of obesity for the NHS are approximated at £3.9 billion and £5.3 billion in 2015 and 2025, respectively (McPherson et al., 2007).

The author also discusses the effectiveness of interventions to treat and prevent childhood obesity from existing systematic reviews. Findings from a meta-analysis of these reviews indicate that, for behavioural programmes aimed at treating childhood obesity, effectiveness is maintained for children at all ages at the 6-month follow-up. However, only children aged 12 and above benefit from the weight reduction at the 12-month follow-up (Oude Luttikhuis et al., 2009). Despite some evidence showing the effectiveness of child obesity interventions, it is not yet clear whether or not investment in childhood obesity interventions is efficient.

To explore this matter, the author conducts a literature review of the cost-effectiveness of child obesity interventions. Cost-effectiveness studies published until March 2010 were included. With regard to the search, 5 articles were reviewed, of which 2 were based on US interventions and the other 3 on Australian ones from the ACE-obesity project. In order to update this chapter with recently published studies on the cost-effectiveness of child obesity interventions, another search was performed after the systematic search was completed. Two relevant cost-effectiveness studies were found in this regard, both of which are part of the same ACE-obesity project in Australia as featured in the 3 Australia-based articles in the literature review.

Only 3 studies estimated the effectiveness of interventions based on a trial, whereas the other 3 applied a hypothetical model to approximate the intervention effect; one study used clinical data from medical records. Health outcomes were expressed in terms of QALYs or DALYs in the reviewed articles. The results showed that incremental cost-effectiveness ratios (ICERs) vary considerably across studies. This discrepancy not only resulted from differences in the costs and outcomes of the interventions, but also largely depended on other factors, for example, the assumptions made in the economic models, sources of costs and outcomes estimates, and economic modelling techniques.

At the end of the chapter, the author makes three major recommendations for future research: 1) cost-effectiveness of child obesity interventions to be assessed more routinely to provide economic information for use in policy decision-making processes; 2) long-term follow-up data to be made available to validate the effectiveness of the interventions; and 3) economic modelling techniques to be improved for better accuracy and reliability.

In response to the suggestions for more and better cost-effectiveness research on child obesity interventions, as asserted in Chapter 2, a behavioural intervention to treat childhood obesity, MEND 7-13, was used as a case study. Chapter 3 introduces the development of the MEND 7-13 programme and identifies the components of the intervention that can potentially tackle and reverse childhood obesity.

MEND 7-13 is a behavioural intervention aimed at treating obesity in children aged 7-13 by changing their lifestyles through such methods as physical activity and dietary education. The effectiveness of MEND 7-13 was assessed in a 6-month trial conducted between 2005 and 2007, for which only children with BMI $\geq 98^{th}$ centile were eligible. After the trial, the programme was rolled out in many areas throughout England and Wales, covering both obese and overweight children (BMI $\geq 91^{st}$ centile). The author examines the effectiveness of the MEND rollout, using t-tests to compare health indicators at baseline with those at the 3-month follow-up. The results indicated significant short-term improvement of all health indicators, including BMI, waist circumference, recovery heart rate, physical activity hours and sedentary hours. Using the t-test to demonstrate the short-term effectiveness of MEND 7-13, however, contains a number of weaknesses, such as the lack of a control group and the regression to the mean bias. Therefore, the results from the t-tests cannot be used to infer the long-term effectiveness of the programme.

The next part of the chapter explores the extent to which intrinsic factors might be influencing the programme's success. The study assesses certain characteristics of participants that might affect the benefit uptakes of MEND 7-13, defined as the BMI change rate per week and obesity status change. The 10-week follow-up data from the rollout phase, collected from January 2007 to December 2009, were used in this regard. Prior to starting the analysis, an assessment of data quality was performed, revealing significant flaws such as selection bias, poor follow-up processes, re-registration, and missing values, which affect both quality of data and interpretation of results. This leads to very limited generalisability of the results. The author also compares the demographic characteristics of the MEND samples with those in the Health Survey for England 2008 data. After assessing the quality of the data, the author provides details regarding the recoding of variables, which were then used in the regression models.

Statistical techniques were used in elucidating the effects of the predictors. The analysis began with a simple regression, including simple linear regression for the BMI change rate and simple logistic regression for the obesity status change. In accordance with the preliminary analysis, the results suggested that ethnicity is an important factor affecting programme success, in terms of both change in BMI per week and obesity status. Hence, the analysis was stratified with respect to the ethnic origin of children: White, Asian, Black, and Other. The study also tested the impact of missing data values on outcomes. Imputation analysis was undertaken in this regard, with results showing that missing values had little effect on outcome estimates for both the BMI change rate and obesity status change. The last part of the analysis employed a multilevel analysis to account for the impact of the programme centre on programme effectiveness. The results showed that significant predictors of the BMI change rate in White children are gender, programme attendance rate, and baseline BMI; for Asian children, the only significant predictor is having parents who own their houses; no evidence of association is shown in Black children; and total attendance and gender are significant predictors in the 'Other' group. The results also showed that the programme centre influences the change in BMI per week to some extent in all children.

The author then explores the associations between intrinsic factors and obesity status change, based on the IOTF definition. Logistic analysis was applied in this regard. However, after testing the goodness of fit of the model, the results showed a poor fit problem. This implies that the predicted values from the model poorly represent the actual data. As the results from the logistic analysis do not give meaningful information about MEND 7-13, the author thus does not present the interpretation of these results. Lastly, the author summarizes the keys points from the analysis and discusses the weaknesses, gaps, and implications of the results. Recommendations to improve MEND 7-13, in terms of programme evaluation and delivery, are provided.

Chapter 4 is dedicated to the cost-effectiveness of MEND 7-13. This study assesses cost-effectiveness under the assumption that MEND 7-13 is accessible to all children in 2010 in England with a BMI ≥ 91^{st} centile of the UK 1990 reference chart. The costs and health outcomes were discounted at 3.5% per annum, according to the NICE guidance (National Institute for Health and Clinical Excellence, 2008), with base year 2010.

The author starts by calculating the ICER of the intervention, which is the ratio of the incremental costs to incremental QALYs gained as a result of the programme. The incremental costs of the programme were calculated by the total costs of MEND 7-13 less the future medical costs saved due to the intervention. Estimation of the intervention costs was calculated on the basis of activity, which include programme set-up, participant recruitment, treatment, and evaluation and monitoring. For each activity, the study derives the total costs into fixed and variable costs. It is noted that the author does not have access to the details of the intervention costs; only the total costs of the programme were provided by MEND Central. The total costs of MEND 7-13 per child equal £415.77, while the number of children who are eligible for the programme is 1,325,638, setting the total costs of MEND 7-13 in 2010 at £551.2 million. Pertaining to the future medical costs saved due to MEND 7-13, the number of obese adults prevented due to the programme was used to project the amount by which the medical costs of obesity are reduced. The study approximates the total future medical costs of obesity paid by the NHS from the Foresight report, which estimated the medical costs of obesity from the increased risks of 10 obesity-related conditions, such as heart disease, diabetes, and certain types of cancer (Butland et al., 2007). The total number of obese cases prevented is derived from the NHS (NHS, 2010). The results showed that the estimated direct medical costs saved by MEND 7-13 is £216 million, which results in a net incremental cost of £551.2 million – £216 million = £335.15 million. It is acknowledged that uncertainties from projecting the obesity prevalence and its costs into the future considerably affect the accuracy and reliability of the cost estimates, as the base case parameters assumed in the model – such as the obesity prevalence and treatment costs – may change over time.

Concerning the incremental health outcome estimation, the economic modelling is based on 3 main links. Firstly, obese children are more likely to be obese adults than their non-obese counterparts. Chapter 4 draws evidence of this association from Freedman et al. (2001), who analyzed longitudinal data of 2,617 children in the US with an average follow-up duration of 17 years. They report that the probabilities of becoming obese in obese and non-obese children are 0.77 and 0.18, respectively. Secondly, obesity increases the mortality risk in adulthood. Supportive evidence was drawn from a study assessing the association between BMI and mortality risk from survey data comprising more than 30,000 samples (Flegal et al., 2005). The results showed that mortality risk increases when BMI is greater than 30, compared with that in the normal-range BMI (18.5-25) (Flegal et al., 2005). Finally, through various mechanisms, obesity reduces quality of life for adults. Chapter 4 uses the results from a study consisting of 1,865 patients in the UK, which illustrated a negative association between BMI and quality of life (Sach et al., 2006).

Given the links above, the author estimates the number of obese adults prevented due to MEND 7-13 in 2027 at 119,862. After adjusting by life expectancy and quality of life, implementing MEND 7-13 throughout England leads to an incremental QALYs of 200,905 QALYs in 2010, resulting in an ICER of £1,668.2 per QALY gained. The ICER of MEND 7-13 is deemed cost-effective when compared with both the NICE threshold of £20,000-30,000 per QALY and other obesity/child obesity interventions.

The author then accounts for uncertainty of the estimates in the sensitivity analysis, which includes uncertainties of the variables, time preference, and the maintenance of programme effectiveness. The sensitivity analysis also performed a subgroup analysis with respect to the use of IOTF and UK 1990 definitions. The results showed that effectiveness maintenance, definition of childhood obesity, and particular characteristics of the subgroup influence the ICER estimates. A final discussion then follows, addressing the contribution of this study to the existing literature, weaknesses and gaps of the model, implications, and future research opportunities.

5.3 What could have been done differently?

Definition of childhood obesity used in both the epidemiologic and economic evaluation chapters

The first possible alternative is to change the definition of childhood obesity, which would affect the results of all epidemiologic analysis and economic evaluation. This thesis justifies the use of the IOTF definition and acknowledges the uncertainties caused by the choice of definition. Though the sensitivity analysis in economic evaluation assessed the impact of the UK 1990 definition on programme effectiveness and ICERs, the study ultimately applies only the IOTF definition to the epidemiologic study in Chapter 3 in order to make the results consistent with the base case economic evaluation. Using the UK 1990 definition would potentially affect the results of the epidemiologic study and economic evaluation through the change in programme effectiveness corresponding to the cut-off, compared with those appearing in this thesis, which are based on the IOTF definition.

Evidence of the transitional probability in childhood of becoming obese adults in the economic evaluation chapter

One of the backbone assumptions in the model used in the economic evaluation is the probability of being obese in adulthood with respect to obesity status in childhood, from Freedman et al. (2001). Chapter 4 acknowledges the limitations pertaining to this assumption, as the mean ages at baseline and follow-up are 10 and 27 years. The economic modelling is therefore inflexible when evaluating long-term health outcomes in interventions that do not focus on children at 10 years of age. One way to increase the flexibility of the model is by employing the findings from Guo et al. (2002), which provides transitional probabilities of being obese at age 35 from the starting ages of 3-20. Guo et al. (2002) reported the transitional probabilities of being obese that correspond to the 75^{th}, 85^{th}, and 95^{th} BMI centile of children at specific ages. This allows the model to change the starting age of children when it is applied to other interventions. However, the trade-offs for flexibility are as follows: 1) the sample size in Guo et al. (2002) is small, only 347 subjects compared with over 2,600 samples in Freedman et al., which may cause more uncertainty in the probability estimates; and 2) the transitional probability of normal-weight children is not provided in Guo et al. (2002), so the comparators of the transitional probability of children with BMI greater than the 95^{th} centile would be those of children with BMI between 75^{th}-85^{th} centile, not those with BMI below 85^{th} centile. Future research is needed to provide better data in order to obtain more precise estimates of the transitional probabilities.

Evidence regarding the probability of obesity in childhood leading to adult overweight in the economic evaluation chapter

Some evidence suggests that the probability of becoming overweight adults in obese children is higher than that in the non-obese (Guo et al., 2002). However, the economic modelling in Chapter 4 does not account for the effect of childhood obesity on becoming overweight adults due to its miniscule magnitude (Guo et al., 2002). Evidence has pointed out that adulthood overweight, BMI 25-30, is not significantly associated with an increased mortality risk (Garfinkel, 1986, Calle et al., 2003, Flegal et al., 2005). In other words, there is a lack of scientific proof that the life expectancy of overweight adults is different from that of the normal-weight. As QALYs are calculated by multiplying life years with quality of life, this implies that the difference in QALYs between overweight and the normal-weight adults depends on differences in the quality of life, given the same duration of life expectancy for both groups. Concerning weight-related quality of life, some evidence points out that the quality of life in overweight adults may be slightly lower than that in their normal-weight counterparts (Sach et al., 2006, Wang et al., 2010). Investigation of the significance of this difference is required before incorporating the results into the QALYs calculation. However, examining weight-related quality of life is out of the scope of this thesis. Therefore, the author considers as a future research opportunity the examination of the differences in weight-related quality of life between groups, with potential to incorporate it into the QALYs estimate.

Perspectives applied in the economic evaluation

The economic evaluation uses the health payer perspective, which best serves its purposes; despite this, the author realises that applying the societal and public health perspectives would each also lead to comprehensive results that slightly differ slightly from those presented in this thesis (Gold et al., 1996, National Institute for Health and Clinical Excellence, 2009). However, the non-health economic and social consequences of reducing childhood obesity may be difficult to account for, due mainly to the lack of data, and thus were not included in the analysis. Nonetheless, if alternative perspectives are taken, the costs and effectiveness of the intervention would be revised, resulting in new ICERs that can demonstrate programme efficiency in different ways.

5.4 Research implications and future research opportunities

For cost-effectiveness analysis

First and foremost, this study contributes to the body of knowledge on the economics of child obesity intervention in the form of cost-effectiveness information, with implications for the policy decision-making process. By considering cost-effectiveness, policymakers can take into account both intervention costs and outcomes when making policy decisions. Examining the cost-effectiveness of child obesity interventions therefore fills a gap in the literature used during the policymaking process.

In addition to filling this gap in the literature, the health outcomes from the economic study in this thesis also facilitate cost-utility analysis. Although there are some cost-effectiveness publications on child obesity interventions, as stated in Chapter 2, only two studies demonstrated the long-term health outcome in terms of QALYs, while DALYs were applied to the rest. Performing cost-utility analysis, which uses strictly QALYs as an outcome, accounts for the health outcome more comprehensively when comparing across cost-effectiveness studies, as discussed in Chapter 1.

Concerning the conceptual framework of health outcomes estimation, the economic modelling is based on one of two preliminary assumptions: 1) changes in BMI change health outcomes, and 2) changes in obesity status change health outcomes. Examples of the first approach can be seen in 5 of the reviewed articles (Moodie et al., 2008a, Magnus et al., 2009, Moodie et al., 2009, Ananthapavan et al., 2010, Moodie et al., 2010), whereas the latter model is seen in Wang et al. (2003), Brown et al. (2007), and Chapter 4 of this thesis. Currently, there is no existing method that can clearly account for the impact of a change in BMI and the original BMI baseline at the same time. Haby et al. (2006) tried to weight the impact of baseline BMI on DALYs calculation using a correction factor, i.e., high BMI contributes to DALYs lost more than low BMI does. However, they performed only a simple calculation in using the correction factor (shown in Chapter 2) and failed to demonstrate how the correction factor was derived for each case. This leaves the effect of using a correction factor unclear as a result.

The advantages and drawbacks of each assumption are flip sides of the same coin. The BMI-based model can account for the health benefits of all the subjects in the model equally as long as there is change in BMI. However, the model cannot clearly take into account the influence of baseline BMI on health risks. Furthermore, health outcomes are demonstrated in terms of DALYs, not QALYs, which do not facilitate cost-utility analysis. On the other hand, the model focusing on the change in obesity status clearly takes into account the importance of baseline BMI. However, it is also evident that changes in obesity status are more likely to occur in those with BMI slightly above the obesity cut-offs compared with those with BMIs farther away from the cut-offs.

In light of the discussion above, research opportunities to improve both types of model arise. For the BMI-based model, calibration and justification of the use of a correction factor is the key to accounting for baseline BMI. For the model based on obesity status, 2 aspects of the estimation method can be improved.

First, QALY estimation in childhood can be incorporated into the long-term QALYs calculation. Current existing models assume that the QALYs of obese children are the same as those of the non-obese, as no evidence shows that childhood obesity increases the mortality risk in childhood (Reilly et al., 2003, Han et al., 2010). However, research has shown that obesity in childhood causes some undesirable social consequences, with possible adverse psychological effects. This may decrease the health-related quality of life of obese children and adolescents, resulting in lower QALYs during youth. Nevertheless, little evidence has been published on the interpretation of the sociological and psychological impacts of childhood obesity in terms of health-related quality of life. Further research inquiring as to the health-related quality of life of obese children in comparison with their non-obese peers is recommended, with the expectation of incorporating it into the QALYs estimation in childhood.

Second, improving QALYs estimation in adulthood can be done by applying more BMI cut-offs with transitional probabilities of becoming obese in adulthood, in order to increase sensitivity to BMI changes. For example, if there are cut-offs for children who are severely obese, moderately obese, obese, and normal-weight with transitional probabilities, the model will be more sensitive in detecting any changes in obesity status in children, as seen in Figure 5.1. In scenario 1, only one cut-off is applied, which cannot account for benefits resulting from the shifting of moderately and severely obese children toward mild obesity. Scenario 2 accounts for the benefits gained from reducing BMI to mild obesity, but assumes equal benefits in reverting severe and moderate obesity; scenario 3 discriminates between these last two categories. Future research is suggested to supply the evidence needed to fill the gaps in both estimating a correction factor and showing transitional probabilities at different baseline BMIs.

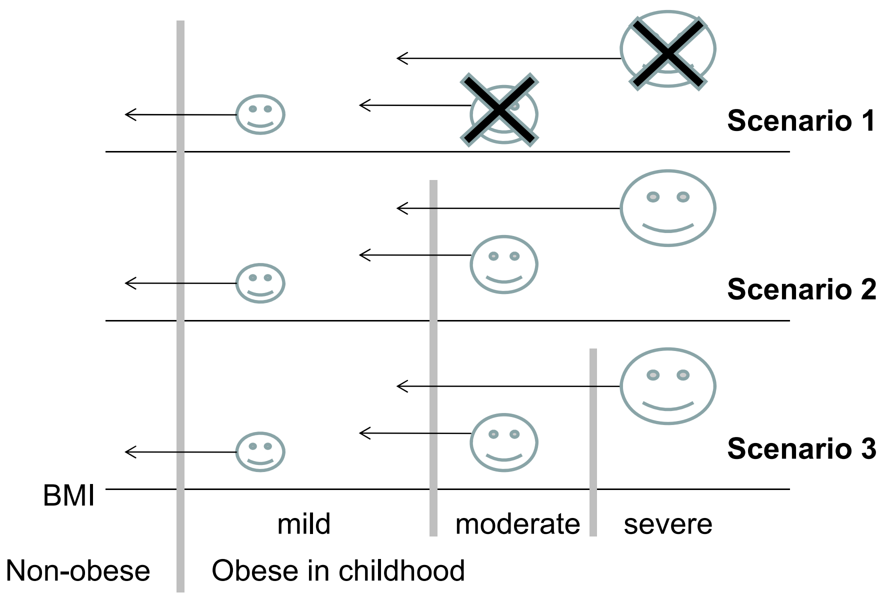


Figure 5.1: Applying more BMI cut-offs to account for degrees of change in BMI

Concerning the model based on obesity status change, this thesis improves the economic modelling from existing techniques. Table 5.1 compares the economic modelling of childhood obesity in this thesis with 3 other studies, with a proposed model included. It is noted that the table contains the model from Wang et al. (2010), which estimates QALYs based on obesity status change. However, their study was not included in the literature review in Chapter 2 because it does not conduct a cost-effectiveness analysis of the intervention. The author has decided to incorporate their model in the summary, as the model demonstrates the improvement of the QALYs calculation.

As seen in Table 5.1, the economic model proposed by this thesis improves the existing models in several aspects. First, compared with Wang et al. (2003) and Brown et al. (2007), the thesis updates the definition of obesity in both childhood and adulthood to those currently used. Second, the duration of QALYs calculation is longer than other existing models, from age 27 onward, even when comparing with the model from Wang et al. (2010). This results in the QALYs estimation being more comprehensive. Lastly, the HRQL applied in the model can distinguish the quality of life of people with BMI in the normal range from those with BMI 30-35 and ≥35. This makes the estimation more comprehensive than those in Wang et al. (2003), Brown et al. (2007), and Wang et al. (2010). It is noted that the model in Wang et al. (2010) compared the HRQL of those with BMI 18.5-25 with those of BMI 25-30 and BMI ≥30. Despite the impact of overweight on HRQL being unproven, as discussed in section 5.4, the author takes this as a model contribution and thus suggests incorporating it into the proposed model as a future research opportunity. Some details of the proposed model can be seen in Table 5.1.

Despite weaknesses attributed to preliminary assumptions, the strength of this economic model lies in its potential to simplify the QALYs estimation process for general application. This thesis improves the economic modelling technique used in the existing literature by allowing the technique to be adjusted for a specific starting age of children. (This is done by employing the transitional probability from Guo et al. (2002), as shown in Table 5.1) This improvement creates a state-of-the-art economic modelling technique, together with suggesting a future research opportunity in the form of a standardized league table illustrating QALYs gained from a child obesity intervention by age. The potential of this future research, based on the economic model developed in this thesis, would simplify QALYs calculation, as only 2 variables will be needed to determine QALYs gained: programme effectiveness and the starting age of children in the programme. It is expected that the standardized QALYs table may be used to encourage study of the cost-effectiveness of child obesity interventions as an ultimate goal.

Table 5.1: Comparison of the economic modelling to estimate the long-term health outcome of childhood obesity

Model characteristics	Wang et al. (2003)	Brown et al. (2007)	Wang et al. (2010)	Model in this thesis	The proposed model
Definition of childhood overweight/obesity	Overweight: BMI≥85th centile	Overweight: BMI≥95th centile	Overweight: 85th-<95th percentiles Obesity: ≥95th percentiles	Obesity: IOTF definition	Obesity: IOTF definition
Definition of adulthood obesity	Ages 21-29: F: BMI ≥27.3; M: BMI≥27.8, Age>40: BMI≥25	Ages 25-29: F: BMI ≥27.3; M: BMI≥27.8, Age>40: BMI > 30	BMI ≥30	BMI ≥30	BMI ≥30
Possible starting ages of the model	1-17 years old	1-17 years old	16-17 years old	7-13 years old	3-20 years old
Link from childhood to adulthood obesity	From ages 1-17 to 21-29 (Whitaker et al. (1997)), From ages 21-29 to 40 (the authors calculated from the NHANES I and NHEFS data)	From ages 1-17 to 25-29 (Whitaker et al. (1997)), From ages 25-29 to 40 (the authors calculated from the NHANES I and NHEFS data)	From ages 16-17 to 40 (Wang et al. (2008))	From ages 7-13 to 27 (Freedman et al. (2001))	From ages 3-20 to 35 (Guo et al. (2002))
Duration of QALYs calculation	From ages 40 to 64	From ages 40 to 64	From age 40 to either the entire lifespan or until age 100	From age 27 to either the entire lifespan or until age 100	From age 35 to either the entire lifespan or until age 100
Comparison of HRQL by BMI in adulthood	Comparing BMI<25 with BMI≥25 (From the authors' calculation)	Comparing BMI<30 with BMI≥30 (From the authors' calculation)	Comparing BMI 18.5-<25 with BMI 25-<30 and BMI≥30 (Derived from Shaw et al. (2005))	Comparing BMI 18.5-<25 with BMI 30-<35 and BMI≥35 (Sach et al. (2006))	Comparing BMI 18.5-<25 with BMI 25-<30,BMI 30-<35 and BMI≥35 (Sach et al. (2006))

Another potential area of improvement in estimating the quality of health outcomes comes from the lack of UK-based evidence. Due to unavailability of the UK evidence, the economic modelling shown in Chapter 4 employed several key parameters from US-based literature, such as the transitional probability of being obese in adulthood and the mortality risk of obesity in adulthood. Accuracy and reliability of the estimate would increase if UK-based studies can complement these gaps in the model.

For the MEND 7-13 programme

The major weaknesses of the MEND data result from selection bias, short follow-up duration, and the follow-up process, which immensely affect the reliability of the results. The MEND 7-13 programme might consider improving the quality of its data for research purposes; the detailed suggestions are provided in the next section.

In the epidemiological analysis, this thesis focuses on health outcomes that are relevant to the economic evaluation of MEND 7-13: BMI change rate and obesity status change. In addition to these outcomes, a number of health indicators have been left unexplored, such as waist circumference, physical activity hours, and sedentary hours. However, this is beyond the scope of this thesis. A research opportunity thus occurs for those who would like to assess the extent to which characteristics other than BMI and obesity status affect programme effectiveness.

In terms of research implication, the results from the study barely explore the impact of intrinsic factors due not only to data quality, but also to clinical insignificance. This is particularly when assessing the impact on the BMI change rate, where only small changes in BMI were shown in association with children's characteristics and other covariates. Careful utilization of results is recommended.

5.5 Policy implications

From the human capital perspective, MEND 7-13 is considered an investment on health by treating obesity in childhood. As mentioned in Chapter 1, Heckman (2010) proposes that the best time for investing in child education is before 5 years of age. However, the programme is open to children between ages 7 and 13, with a mean age of 10.4, as shown by the descriptive statistics in Chapter 3. In this regard, the starting age of MEND 7-13 might be slightly late for child education with respect to the Heckman equation. Nonetheless, considering the skills taught by MEND 7-13, e.g., supermarket tours, food-label reading, and healthy cooking with parents, the programme may not be well-suited for toddlers and children below 5 years old. Moreover, the rate of return to investment in human capital during the schooling age, shown in Chapter 1, is still high compared with investments made post-schooling.

In spite of the above-mentioned weaknesses, this thesis sheds light on health disparities as well as the cost-effectiveness of the MEND 7-13 programme, which is beneficial for programme improvement and policy advice.

Chapter 3 reveals an issue regarding the inconsistency of programme delivery across MEND 7-13 centres, which includes differences in locations and staff ability. MEND Central may consider this as an opportunity to learn from the centres that effectively reduce obesity in children relative to the others by focusing on differences in the characteristics of the locations and the local staff. This can be one way to improve the effectiveness of the programme.

Concerning health disparities, the results from Chapter 3 explore the extent to which certain characteristics influence changes in health indicators. In doing so, the author points out socio-demographic disparities as a potential predictor for gaining benefits from the programme. This not only alerts the MEND staff to be aware of health disparities, but can also be used to tailor the intervention to maximize health outcomes in two ways. On one hand, the programme can maximize its impact by selecting children who are most likely to benefit from the intervention. On the other hand, participants with particular characteristics that are prone to be less responsive to the intervention should be paid careful attention by staff, e.g., by spending more time to ensure message delivery, in order to increase programme effectiveness.

Recommendations to improve the quality of delivery and programme effectiveness are listed as follows. First, the effectiveness of the MEND 7-13 rollout should be re-evaluated based on the characteristics of the rollout programme, including the staff qualification and eligibility criteria. The study should have an adequate sample size with a control group in order to avoid biases. In the case that many health outcomes are tested using the same dataset, adjusting the level of significance using statistical techniques is needed when interpreting results from the statistical analysis.

Second, the programme should consider improving the quality of data in several aspects, especially the duration of follow-up. The long-term follow-up data should be available to incorporate into the assessment of programme effectiveness. Duration of follow-up should ideally last until adulthood, but a more realistic goal may be to reduce the missing values at 12 months. The data collecting process should also be examined not only to find out the cause of the missing data at 6 and 12 months, but also to reduce instances of re-registration. In addition, apart from health outcome assessments, the programme should provide information about the delivery process – such as the completion of topic delivery during the sessions and the learning atmosphere in classes – in order for the process evaluation to ensure the quality of delivery.

Third, MEND7 7-13 should compare different programme characteristics across centres – such as staff qualification and location – to learn the extent to which these factors influence programme effectiveness. This information would be helpful in setting up the local centres, as location and delivery team characteristics can be selected to maximize the health benefits in the context of the local settings.

Finally, to reduce health disparities, the programme should target obese children from low socio-economic status, who are less likely to register for the programme. This can be done through several ways. For instance, more registration methods, apart from website-based, should be made available; also, the programme sites should be located in deprived areas, making them more accessible to local children and their parents. However, the costs incurred from efforts to target disadvantaged children would make the programme more expensive, which affects its efficiency. This raises a question of whether or not MEND 7-13 would still be cost-effective after addressing health disparities.

Regarding the cost-effectiveness results, the ICER of MEND 7-13 in the base case analysis is highly cost-effective when compared with the NICE threshold. However, at £551.2m per annum, the total cost of the programme is deemed enormous for a child obesity intervention. This appears to corroborate some arguments discouraging the treatment of childhood obesity due to high costs and unsustainable effects (Ebbeling et al., 2002, Lobstein et al., 2004). Apart from financial constraints, questionable return on investment may cause policymakers to hesitate, as future uncertainties may affect the delayed benefits from child obesity interventions. Other perspectives, such as priority setting for health policy budgeting, societal opinions, and political considerations, may need to be incorporated when considering MEND 7-13 within the public funding scheme.

5.6 Final conclusions

One of the implications of this thesis concerns the fact that the effectiveness of child obesity interventions has been widely evaluated using different types of outcome measures, such as BMI, weight changes, and waist circumference. However, only a handful of studies focus on evaluating the cost-effectiveness of interventions to reduce childhood obesity, particularly using cost-utility analysis. One explanation for this could be due to the fact that reducing childhood obesity is a long-term investment that aims at preventing future obesity-related medical conditions. While assessing the short-term effectiveness of the child obesity programme is straightforward, three issues arise when considering the evaluation of long-term cost-effectiveness. First, evaluation of public health interventions in the long run requires a considerable amount of time and expense for follow-up, which may not be pragmatic. Assessment of the short-term effectiveness of interventions thus becomes a more practical alternative that can also provide policymakers with information on the programmes' performance. In addition to the time and costs incurred from long-term follow-ups, assessing the long-term health effects of obesity is very complicated, as obesity increases the risk of tens of medical conditions, e.g., diabetes, stroke, and various cancers. It becomes a challenge not only to assess the increased risks of disease occurrence with respect to the degree of obesity, but also to evaluate the subsequent costs and health outcome deterioration resulting from such disease. The difficulty of measuring the long-term costs and health consequences of obesity may therefore be another issue that discourages further study in this area. Lastly, apart from the limitations inherent in estimating long-term costs and outcome, researchers need to employ an appropriate economic model in order to derive cost-effectiveness results, including QALYs estimation. Moreover, there is as of yet no single agreed-upon technique to assess QALYs for child obesity interventions. Due to these concerns, most studies focus only on the effectiveness of child obesity interventions, whereas very few evaluate long-term cost-effectiveness results.

One solution to encourage programme efficiency is to improve the contract of the child obesity interventions in terms of definitions, enforcement, management, and evaluation. This is not only to ascertain that the programme has set delivery guidelines with expected results in terms of process and outcome quality, but also to ensure that programme evaluation can be done appropriately using complete data and analyses.

Economic evaluation is one of many tools that helps inform policymakers in the policymaking process. This thesis proposes an economic modelling technique and suggests opportunities for future research that would simplify the process of deriving the long-term health outcomes of childhood obesity, thus encouraging the production of cost-effectiveness research.

Cost-effectiveness analysis is not an indicator of whether or not an intervention should be funded and implemented. Other types of information, such as the societal perspective, ethical considerations, and political concerns should be considered alongside the results from the economic perspective in order to reflect concerns from various viewpoints. Doing so would achieve the purpose of providing multi-dimensional information to policymakers, which will facilitate their drafting and implementation of evidence-based policies.

Annex 1: Life table based on year 2006-2008

Age	Males				Females			
x	q_x	S_x	$S_{10\text{-}27}$	$S_{27\text{-}100}$	q_x	S_x	$S_{10\text{-}27}$	$S_{27\text{-}100}$
10	0.000105	0.999895	0.999895		0.000093	0.999907	0.999907	
11	0.000119	0.999881	0.999776		0.000097	0.999903	0.999810	
12	0.000120	0.999880	0.999656		0.000100	0.999900	0.999710	
13	0.000161	0.999839	0.999495		0.000118	0.999882	0.999592	
14	0.000183	0.999817	0.999312		0.000119	0.999881	0.999473	
15	0.000248	0.999752	0.999064		0.000158	0.999842	0.999315	
16	0.000352	0.999648	0.998713		0.000173	0.999827	0.999142	
17	0.000526	0.999474	0.998187		0.000245	0.999755	0.998898	
18	0.000601	0.999399	0.997587		0.000271	0.999729	0.998627	
19	0.000644	0.999356	0.996945		0.000257	0.999743	0.998370	
20	0.000699	0.999301	0.996248		0.000241	0.999759	0.998130	
21	0.000659	0.999341	0.995592		0.000264	0.999736	0.997866	
22	0.000702	0.999298	0.994893		0.000263	0.999737	0.997604	
23	0.000724	0.999276	0.994172		0.000247	0.999753	0.997357	
24	0.000734	0.999266	0.993443		0.000294	0.999706	0.997064	
25	0.000750	0.999250	0.992698		0.000287	0.999713	0.996778	
26	0.000787	0.999213	**0.991916**		0.000337	0.999663	**0.996442**	
27	0.000818	0.999182		0.999182	0.000311	0.999689		0.999689
28	0.000833	0.999167		0.998350	0.000358	0.999642		0.999331
29	0.000883	0.999117		0.997468	0.000381	0.999619		0.998950
30	0.000943	0.999057		0.996528	0.000416	0.999584		0.998535
31	0.001012	0.998988		0.995519	0.000406	0.999594		0.998129
32	0.001063	0.998937		0.994461	0.000487	0.999513		0.997643
33	0.001128	0.998872		0.993339	0.000539	0.999461		0.997106
34	0.001176	0.998824		0.992171	0.000575	0.999425		0.996532
35	0.001289	0.998711		0.990892	0.000598	0.999402		0.995936
36	0.001290	0.998710		0.989614	0.000644	0.999356		0.995295
37	0.001339	0.998661		0.988289	0.000727	0.999273		0.994571
38	0.001454	0.998546		0.986852	0.000795	0.999205		0.993781
39	0.001515	0.998485		0.985357	0.000897	0.999103		0.992889
40	0.001603	0.998397		0.983777	0.000980	0.999020		0.991916
41	0.001751	0.998249		0.982054	0.001056	0.998944		0.990869
42	0.001892	0.998108		0.980196	0.001151	0.998849		0.989728
43	0.002017	0.997983		0.978219	0.001267	0.998733		0.988474
44	0.002165	0.997835		0.976102	0.001341	0.998659		0.987149
45	0.002417	0.997583		0.973742	0.001530	0.998470		0.985638
46	0.002570	0.997430		0.971240	0.001648	0.998352		0.984014
47	0.002861	0.997139		0.968461	0.001828	0.998172		0.982215
48	0.003043	0.996957		0.965514	0.002067	0.997933		0.980185
49	0.003336	0.996664		0.962293	0.002151	0.997849		0.978077
50	0.003706	0.996294		0.958727	0.002559	0.997441		0.975574
51	0.004114	0.995886		0.954783	0.002692	0.997308		0.972948
52	0.004370	0.995630		0.950610	0.002861	0.997139		0.970164
53	0.004969	0.995031		0.945887	0.003158	0.996842		0.967100
54	0.005410	0.994590		0.940769	0.003537	0.996463		0.963680
55	0.005956	0.994044		0.935166	0.003755	0.996245		0.960061
56	0.006470	0.993530		0.929116	0.004141	0.995859		0.956085
57	0.006876	0.993124		0.922727	0.004390	0.995610		0.951888
58	0.007510	0.992490		0.915797	0.004717	0.995283		0.947398

Age	Males				Females			
x	q_x	*Sx*	*S10-27*	*S27-100*	q_x	*Sx*	*S10-27*	*S27-100*
59	0.008062	0.991938		0.908414	0.005303	0.994697		0.942374
60	0.008872	0.991128		0.900355	0.005696	0.994304		0.937006
61	0.010017	0.989983		0.891336	0.006452	0.993548		0.930961
62	0.011204	0.988796		0.881349	0.006907	0.993093		0.924530
63	0.012388	0.987612		0.870431	0.007798	0.992202		0.917321
64	0.013577	0.986423		0.858613	0.008516	0.991484		0.909509
65	0.014944	0.985056		0.845782	0.009178	0.990822		0.901162
66	0.016315	0.983685		0.831983	0.010084	0.989916		0.892074
67	0.017649	0.982351		0.817300	0.011251	0.988749		0.882038
68	0.019784	0.980216		0.801130	0.012330	0.987670		0.871162
69	0.021402	0.978598		0.783984	0.013535	0.986465		0.859371
70	0.023352	0.976648		0.765677	0.015072	0.984928		0.846418
71	0.025827	0.974173		0.745902	0.016561	0.983439		0.832401
72	0.028749	0.971251		0.724458	0.018374	0.981626		0.817106
73	0.031659	0.968341		0.701522	0.020832	0.979168		0.800084
74	0.034966	0.965034		0.676993	0.023483	0.976517		0.781296
75	0.039290	0.960710		0.650394	0.025871	0.974129		0.761083
76	0.043633	0.956367		0.622015	0.029214	0.970786		0.738849
77	0.048538	0.951462		0.591824	0.032712	0.967288		0.714680
78	0.053864	0.946136		0.559946	0.036780	0.963220		0.688394
79	0.060471	0.939529		0.526085	0.041825	0.958175		0.659602
80	0.066775	0.933225		0.490956	0.047053	0.952947		0.628565
81	0.074375	0.925625		0.454441	0.052661	0.947339		0.595465
82	0.083149	0.916851		0.416655	0.058476	0.941524		0.560644
83	0.091383	0.908617		0.378580	0.066223	0.933777		0.523517
84	0.101365	0.898635		0.340205	0.074507	0.925493		0.484511
85	0.111253	0.888747		0.302356	0.083203	0.916797		0.444198
86	0.122792	0.877208		0.265229	0.092440	0.907560		0.403136
87	0.127281	0.872719		0.231470	0.101085	0.898915		0.362385
88	0.139239	0.860761		0.199241	0.114035	0.885965		0.321061
89	0.151278	0.848722		0.169100	0.124557	0.875443		0.281070
90	0.170113	0.829887		0.140334	0.140443	0.859557		0.241596
91	0.189938	0.810062		0.113679	0.160738	0.839262		0.202762
92	0.207416	0.792584		0.090100	0.179041	0.820959		0.166460
93	0.224041	0.775959		0.069914	0.197859	0.802141		0.133524
94	0.237285	0.762715		0.053325	0.215442	0.784558		0.104757
95	0.264643	0.735357		0.039213	0.234414	0.765586		0.080201
96	0.282137	0.717863		0.028149	0.255550	0.744450		0.059705
97	0.310074	0.689926		0.019421	0.271716	0.728284		0.043483
98	0.323622	0.676378		0.013136	0.300529	0.699471		0.030415
99	0.334547	0.665453		0.008741	0.314942	0.685058		0.020836
100	0.371257	0.628743		0.005496	0.341322	0.658678		0.013724

q_x is the mortality rate between age x and (x +1), that is the probability that a person aged x exact will die before reaching age (x +1).

S_x is the probability of survival between age x and (x+1), which equals (1- q_x)

$S_{10\text{-}27}$ is the probability of survival between ages 10 and 27.

$S_{27\text{-}100}$ is the probability of survival between ages 27 and 100.

Source: Modified from Office for National Statistics (2010)

Appendix 1: Keywords search in the literature review

Medline

1. exp Obesity/
2. exp Weight Gain/
3. exp Weight Loss/
4. obes$.tw.
5. (weight gain or weight loss).tw.
6. (overweight or over weight or overeat$ or over eat$).tw.
7. weight change$.tw.
8. ((bmi or body mass index) adj2 (gain or loss or change)).tw.
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. exp Child/
11. exp Adolescent/
12. exp Child, Preschool/
13. exp Infant/
14. (child$ or adolescen$ or infant$).tw.
15. (teenage$ or young people or young person or young adult$).tw.
16. (schoolchildren or school children).tw.
17. (pediatric$ or paediatric$).tw.
18. (boys or girls or youth or youths).tw.
19. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18
20. exp Behavior Therapy/
21. Family Therapy/
22. Psychotherapy, Group/
23. ((behavio?r or psychological or family) adj (therapy or modif$ or strateg$)).tw.
24. social support.tw.
25. Diet, Reducing/
26. Diet Therapy/
27. Caloric Restriction/
28. Fasting/
29. (diets or diet or dieting or fasting).tw.
30. (diet$ adj (modif$ or therapy or intervention$ or strateg$)).tw.
31. (high fat$ or low fat$ or fatty food$).tw.
32. (low calorie or calorie control or healthy eating).tw.
33. exp Exercise/
34. "Physical Education and Training"/
35. exercis$.tw.
36. (aerobics or physical therapy or physical activity or physical inactivity or fitness).tw.
37. (physical training or physical education).tw.
38. sedentary behavio?r.tw.
39. Health Promotion/
40. Health Education/
41. (health promotion or health education).tw.
42. ((obesity or overweight) adj2 (prevent$ or treat$)).tw.
43. Primary Prevention/
44. Secondary Prevention/
45. (primary prevention or secondary prevention).tw.
46. (preventive measure$ or preventative measure$).tw.
47. (preventive care or preventative care).tw.
48. (group therapy or family therapy or cognitive therapy).tw.
49. (program or programs or programme$ or intervention$).tw.
50. Counseling/
51. counsel?ing.tw.
52. exp Anti-Obesity Agents/
53. Metformin/
54. ((anti-obes$ or antiobes$) adj (agent$ or drug$ or medicine$)).tw.
55. sibutramine.tw.
56. rimonabant.tw.
57. metformin.tw.
58. reductil.tw.
59. lipstatin.tw.
60. orlistat.tw.
61. xenical.tw.
62. acomplia.tw.
63. glucophage.tw.
64. (weight loss adj2 agent$).tw.
65. anorexigenic agent$.tw.
66. exp Bariatric Surgery/
67. gastroplast$.tw.
68. bariatric surgery.tw.
69. gastric bypass$.tw.
70. (gastric adj2 band$).tw.
71. (surgery adj2 obes$).tw.
72. or/20-71
73. exp Cost-Benefit Analysis/
74. exp "Costs and Cost Analysis"/
75. (cost-effectiveness or cost-benefit or cost-utility or cost-analysis).tw.
76. (cost or costs or costing).tw.
77. (economic$ or net benefit$).tw.
78. Economics/
79. or/73-78
80. 9 and 19 and 72 and 79

Embase

1. exp obesity/
2. exp weight gain/
3. exp weight reduction/

4. obes$.tw.
5. (weight gain or weight loss).tw.
6. (overweight or over weight or overeat$ or over eat$).tw.
7. weight change$.tw.
8. ((bmi or body mass index) adj2 (gain or loss or change)).tw.
9. or/1-8
10. exp child/
11. exp adolescent/
12. exp preschool child/
13. exp infant/
14. (child$ or adolescen$ or infant$).tw.
15. (teenage$ or young people or young person or young adult$).tw.
16. (schoolchildren or school children).tw.
17. (pediatr$ or paediatr$).tw.
18. (boys or girls or youth or youths).tw.
19. or/10-18
20. exp behavior therapy/
21. family therapy/
22. social support/
23. behavior modification/
24. support group/
25. ((behavio?r or psychological or family) adj (therapy or modif$ or strateg$)).tw.
26. social support.tw.
27. diet restriction/
28. fat intake/
29. low fat diet/
30. low calory diet/
31. diet therapy/
32. (diets or diet or dieting or fasting).tw.
33. (diet$ adj2 (modif$ or therapy or intervention$ or strateg$)).tw.
34. (high fat$ or low fat$ or fatty food$).tw.
35. (low calorie or calorie control or healthy eating).tw.
36. exp exercise/
37. physical education/
38. physical activity/
39. exercis$.tw.
40. (aerobics or physical therapy or physical activity or physical inactivity or fitness).tw.
41. (physical training or physical education).tw.
42. sedentary behavio?r.tw.
43. health promotion/
44. health education/
45. nutrition education/
46. (health promotion or health education).tw.
47. ((obesity or overweight) adj2 (prevent$ or treat$)).tw.
48. primary prevention/
49. (primary prevention or secondary prevention).tw.
50. (preventive measure$ or preventative measure$).tw.
51. (preventive care or preventative care).tw.
52. (group therapy or family therapy or cognitive therapy).tw.
53. (program or programs or programme$ or intervention$).tw.
54. counseling/
55. counsel?ing.tw.
56. antiobesity agent/
57. sibutramine/
58. anorexigenic agent/
59. ((anti-obes$ or antiobes$) adj (agent$ or drug$ or medicine$)).tw.
60. sibutramine.tw.
61. rimonabant.tw.
62. metformin.tw.
63. reductil.tw.
64. lipstatin.tw.
65. orlistat.tw.
66. xenical.tw.
67. acomplia.tw.
68. glucophage.tw.
69. (weight loss adj2 agent$).tw.
70. anorexigenic agent$.tw.
71. exp bariatric surgery/
72. gastroplast$.tw.
73. bariatric surgery.tw.
74. gastric bypass$.tw.
75. (gastric adj2 band$).tw.
76. (surgery adj2 obes$).tw.
77. or/20-76
78. exp economic evaluation/
79. exp cost effectiveness analysis/
80. exp cost benefit analysis/
81. exp cost utility analysis/
82. economics/
83. (cost-effectiveness or cost-benefit or cost-utility or cost-analysis).tw.
84. (cost or costs or costing).tw.
85. (economic$ or net benefit$).tw.
86. or/78-85
87. 9 and 19 and 77 and 86

PsycINFO

1. exp Obesity/
2. exp Weight Gain/
3. exp Weight Loss/

4. obes$.tw.

5. (weight gain or weight loss).tw.

6. (overweight or over weight or overeat$ or over eat$).tw.

7. weight change$.tw.

8. ((bmi or body mass index) adj2 (gain or loss or change)).tw.

9. or/1-8

10. (child$ or adolescen$ or infant$).tw.

11. (teenage$ or young people or young person or young adult$).tw.

12. (schoolchildren or school children).tw.

13. (pediatr$ or paediatr$).tw.

14. (boys or girls or youth or youths).tw.

15. or/10-14

16. exp Behavior Therapy/

17. Social Support/

18. Family Therapy/

19. ((behavio?r or psychological or family) adj (therapy or modif$ or strateg$)).tw.

20. social support.tw.

21. (diets or diet or dieting or fasting).tw.

22. (diet$ adj2 (modif$ or therapy or intervention$ or strateg$)).tw.

23. exp Exercise/

24. exercis$.tw.

25. (aerobics or physical therapy or physical activity or physical inactivity or fitness).tw.

26. (physical training or physical education).tw.

27. sedentary behavio?r.tw.

28. Health Promotion/

29. Health Education/

30. (health promotion or health education).tw.

31. ((obesity or overweight) adj2 (prevent$ or treat$)).tw.

32. (primary prevention or secondary prevention).tw.

33. (preventive measure$ or preventative measure$).tw.

34. (preventive care or preventative care).tw.

35. (group therapy or family therapy or cognitive therapy).tw.

36. (program or programs or programme$ or intervention$).tw.

37. Counseling/

38. counsel?ing.tw.

39. ((anti-obes$ or antiobes$) adj (agent$ or drug$ or medicine$)).tw.

40. sibutramine.tw.

41. rimonabant.tw.

42. metformin.tw.

43. reductil.tw.

44. orlistat.tw.

45. xenical.tw.

46. acomplia.tw.

47. (weight loss adj2 agent$).tw.

48. anorexigenic agent$.tw.

49. exp Bariatric Surgery/

50. gastric bypass$.tw.

51. (gastric adj2 band$).tw.

52. (surgery adj2 obes$).tw.

53. gastroplast$.tw.

54. bariatric surgery.tw.

55. or/16-54

56. exp "Costs and Cost Analysis"/

57. (cost-effectiveness or cost-benefit or cost-utility or cost-analysis).tw.

58. (cost or costs or costing).tw.

59. (economic$ or net benefit$).tw.

60. or/56-59

61. 9 and 15 and 55 and 60

Econlit

1. obes$.tw.

2. (weight gain or weight loss).tw.

3. (overweight or over weight or overeat$ or over eat$).tw.

4. weight change$.tw.

5. ((bmi or body mass index) adj2 (gain or loss or change)).tw.

6. or/1-5

7. (child$ or adolescen$ or infant$).tw.

8. (teenage$ or young people or young person or young adult$).tw.

9. (schoolchildren or school children).tw.

10. (pediatr$ or paediatr$).tw.

11. (boys or girls or youth or youths).tw.

12. or/7-11

13. ((behavio?r or psychological or family) adj (therapy or modif$ or strateg$)).tw.

14. social support.tw.

15. (diets or diet or dieting or fasting).tw.

16. (diet$ adj2 (modif$ or therapy or intervention$ or strateg$)).tw.

17. (high fat$ or low fat$ or fatty food$).tw.

18. (low calorie or calorie control or healthy eating).tw.

19. exercis$.tw.

20. (aerobics or physical therapy or physical activity or physical inactivity or fitness).tw.

21. (physical training or physical education).tw.

22. sedentary behavio?r.tw.

23. (health promotion or health education).tw.

24. ((obesity or overweight) adj2 (prevent$ or treat$)).tw.

25. (primary prevention or secondary prevention).tw.

26. (preventive measure$ or preventative measure$).tw.

27. (preventive care or preventative care).tw.

28. (group therapy or family therapy or cognitive therapy).tw.

29. (program or programs or programme$ or intervention$).tw.

30. counsel?ing.tw.

31. ((anti-obes$ or antiobes$) adj (agent$ or drug$ or medicine$)).tw.

32. sibutramine.tw.

33. metformin.tw.

34. orlistat.tw.

35. or/13-34

36. (cost-effectiveness or cost-benefit or cost-utility or cost-analysis).tw.

37. (cost or costs or costing).tw.

38. (economic$ or net benefit$).tw.

39. 36 or 37 or 38

40. 6 and 12 and 35 and 39

Appendix 2: Excluded articles from the literature review

Articles	Reasons for exclusion
(Carter et al., 2009)	No intervention outcome was presented
(Cawley, 2007)	The research is not finished yet
(Charlwood and Gibbons, 1986)	No original work was carried out
(Dalziel and Segal, 2006)	Comment
(Finkelstein and Brown, 2005)	The intervention outcome was presented only in terms of cost-benefits, not in life years.
(Ganz, 2003)	No intervention
(Goldfield et al., 2001)	The outcome was shown in the change in BMI z-score, not in life years
(Haby et al., 2006)	DALYs saved of the interventions were shown without calculating cost-utility analysis
(Haby et al., 2007)	Response to the comment
(Kalavainen et al., 2009)	The intervention outcome was shown in BMI standard deviations score, not in life years.
(Moodie et al., 2008b)	No intervention
(Mooney, 2007)	No intervention
(Segal and Dalziel, 2007)	Comment
(Seidell et al., 2005)	No intervention
(Wake et al., 2008)	The outcome was not shown in the life year unit.
(Wang et al., 2008)	No intervention
(Wang et al., 2006)	Response to the comment

Appendix 3: critical assessment of economic evaluation

Assessing economic evaluation of Wang et al. (2003)

1. Was a well-defined question posed in answerable form?

YES

The authors describe the purpose of their study, which is to evaluate whether or not the reduced prevalence of obese children due to Planet Health, a school-based intervention to prevent obesity in children, is worth the additional costs (p.1314).

The authors state that they conducted CEA and CBA to compare the value of Planet Health with that of status quo in the control (p.1314).

They assert that a societal perspective was applied to identify the relevant costs of Planet Health. Although the direct medical costs were estimated from existing literature (Gorsky et al.), they failed to estimate some indirect costs, such as the costs of patients' travel time and informal care givers. No reason for exclusion is provided in the article.

Overall, the authors clearly define the question in answerable format, including the costs and outcome of the intervention as well as the comparator – the status quo in the controlled group. However, they did not comprehensively include all relevant costs in accordance with the societal perspective into the analysis, which might be due to the unavailability of data. This may be arbitrary, and it is therefore up to the reader to decide if the results are biased due to excluded costs.

2. Was a comprehensive description of the competing alternatives given? (that is, can you tell who did what to whom, where and how often?)

UNCLEAR

Details are described in Gortmaker et al. (1999a). The main objective of the randomized controlled trial is to assess the effectiveness of Planet Health, a school-based behavioral intervention to prevent obesity in middle school children. Planet Health was implemented in 5 schools in the US, and its effectiveness was compared with that of 5 other control schools over a period of 2 years.

The authors did not clearly state the sample size in the economic evaluation. Of a total starting number of 1,203 children in the trial, 310 girls and 331 boys completed the 2-year follow-up in the intervention schools, whereas 317 girls and 337 boys did so in the control schools (Gortmaker et al., 1999a). However, the change in the prevalence of obesity is significant only among girls, not boys (p.415) (Gortmaker et al., 1999a). Wang et al. (2003) assert that the intervention costs of 1,203 children were accounted for in the cost analysis (p.1322). Regarding the effectiveness, the 310 girls in the intervention were included in the outcome estimation and a hypothetical 310 girls were assumed as a comparator in the model, without mentioning the number of boys included (p.1315).

This information misled Dalziel and Segal (2006) when they commented on Wang et al.'s analysis. Dalzeil and Segal inferred that only female samples were included the rogram was proven to be effective, which biased the results (p.1481). This claim was countered by Wang et al. (2006), who claimed misunderstanding on the part of Dalziel and Segal's. They assert that they included the costs of all participants in the study, and that the model accounted for only the effectiveness of Planet Health for girls (p.1483). However, Wang el al. (2006) failed, for the second time, to clearly state how many boys they included when they referred to all participants. Logically, it can be assumed that the more boys included in the analysis, the less effective the Planet Health would be. The most likely number of boys should be the 331who completed the 2-year follow-up in the intervention arm. However, a more pessimistic viewpoint may see the hypothetical model include only 310 boys, the same number as girls; 337 boys, as in the control; or an average between the number of boys in the intervention and control groups. This affects the effectiveness of the program, as the reader would have to assess the extent to which results may be biased.

Gortmaker et al. (1999a) describe the intervention as comprising 16 lessons delivered by trained teachers to children per year, for a total of 32 lessons. One or two 45-minute classes were designed for each lesson, depending on the content (p.411) (Gortmaker et al., 1999a).

A do-nothing alternative seems to fit in this case given the type of intervention and other factors. It is understandable that the increasing trend of obesity in children is the major concern in this case. Even though many child obesity interventions have been employed to tackle this public health issue, none seem widely accepted as a standard practice.

3. Was the effectiveness of the programmes or services established?

YES

Details are described in Gortmaker et al. (1999a). The primary outcome of the randomized controlled trial is the obesity status, which is defined as a composite of BMI and triceps skinfolds (TSF). Student heights were measured without shoes to the nearest 0.1 cm, and weights were measured in light clothes to the nearest 0.1 kg. BMI was then calculated as weight (kg)/ height^2 (cm). TSF was measured twice by trained project staff using calibrated Holtain calipers to improve precision, according to accepted protocols. Obesity was defined as having both a BMI and a TSF that is equal to or greater than the 85th percentile of age- and sex-specific reference data. After the 2-year follow-up, the results showed that the prevalence of obesity rose from 21.5% to 23.7% in the control group, while that of girls in the intervention schools was reduced from 23.6% to 20.4% (p.411) (Gortmaker et al., 1999a).

The major concern of the outcome from a 2-year intervention is the sustainability of the effect. The duration of 2 years reflects the effectiveness of Planet Health in the short to intermediate terms; however, it is uncertain whether the reduced prevalence of obesity will be sustained throughout the childhood period.

The authors claim that a composite indicator of obesity could improve the accuracy of diagnosis, as studies have shown the limitations of using a single measure (either BMI or TSF) to diagnose obesity (Reilly et al., 2002, Reilly et al., 2003). They also make a good point about the trade-off between accuracy, with reference to the sensitivity and specificity of a diagnostic tool, and applicability. BMI is widely used as a standard measure of obesity because of its applicability, whereas the use of others measures, such as TSF or other radiographic scanning, is limited to studies that require high accuracy. This is mainly due to the high costs, complex techniques, and low availability of the measures. The subsequent drawback of using a composite indicator is comparability, as most obesity research tends to use BMI as an indicator. Inconsistency is likely if the definitions of obesity across studies are different. Therefore, despite some benefits from using a composite indicator as an obesity diagnostic tool, obesity should be measured using a sole indicator – BMI – in current practice.

The primary outcome of the study is the QALYs gained between ages 40 and 64. QALY calculation is based on the link between obesity in childhood and increased risks of being obese in adulthood, which leads to higher morbidity and mortality rates. Using QALY is considered a standard health outcome in the economic evaluation.

4. Were all the important and relevant costs and consequences for each alternative identified?

UNCLEAR

As mentioned earlier, the authors failed to mention the costs of patients' travel time and informal care givers, even though other relevant costs to the societal perspective, including training costs, overhead (except classroom), operation costs, future medical costs of obesity saved, and costs of productivity loss reverted, were mentioned and sensitivity analysis was performed. The missing information may not affect the result to a great extent, but it is difficult to tell.

5. Were costs and consequences measured accurately in appropriate physical units (for example, hours of nursing time, number of physician visits, lost work-days, gained life-years)?

YES

The costs of Planet Health were estimated retrospectively over the period of 2 years (p.1317). The obesity-related medical expenses and costs of productivity loss for females were measured for the lifelong duration (p.1320). All costs and outcomes were estimated for all participants of the cohort (p.1322). QALYs were estimated only during 25 years in adulthood, between the ages of 40 and 64. The range of estimate could be improved if the duration of interest is extended (p.1319).

The process of outcome data collection is described in Gortmaker et al. (1999a). The program activities were delivered at school. The assessment of the program adhesion is based on a survey. Students were asked to complete the survey independently in class under the supervision of trained teachers (p.411) (Gortmaker et al., 1999a). The validity of results from the survey is calibrated with teacher reports of implementation. The results revealed good validity compared with classroom observation (p.412) (Gortmaker et al., 1999a). The information gained from the trial is therefore considered high-quality.

The effectiveness of Planet Health is referred to Gortmaker et al.'s study (p.1314), which used an appropriate statistical test (t-test) to assess the change in obesity prevalence (p.415) (Gortmaker et al., 1999a). QALYs were calculated in a conventional way, i.e., the quality of life multiplied by life years (p.1316-7).

6. Were costs and consequences valued credibly?

YES – for cost; NO – for outcome

The intervention costs and medical costs are based on charges according to the societal perspective (p.1317, 1320). The medical costs of obesity were derived using an economic model by Gorsky et al. (1996) (p.1316), which estimates the costs of obesity based on 5 obesity-related diseases: heart disease, hypertension, diabetes mellitus, gallstones, and osteoarthritis (Gorsky et al., 1996). It is noted that Gorsky et al. estimated the medical costs of obesity only in women. However, the female-based results from Gorsky et al.'s do not affect the analysis because Health Planet showed the effectiveness only in girls.

The methods used to estimate the costs of obesity are controversial, as obesity not only invokes direct medical costs, but also increases risk of a number of chronic diseases, such as cardiovascular diseases, hypertension, and diabetes. Estimating the medical costs saved from converting obesity therefore depends on approximations of the number and extent of subsequent diseases to which obesity contributes. This is particularly complicated when estimating the future medical costs of obesity because most adverse consequences of obesity arise in adulthood. Hence, interpreting the estimated medical costs of childhood obesity must be done cautiously.

Inconsistency in the definition of obesity appears along the estimation of the long-term health outcome. The authors draw links among overweight in children, the increased risk of becoming overweight by ages 21-29, and said risk by age 40. The BMI cutoff for obesity at ages 21-29 is 27.8 for men and 27.3 for women (p.870) (Whitaker et al., 1997), whereas the rest of the analysis uses the BMI cutoff of 25 (p.1316) (Wang et al., 2003). Moreover, the definition of childhood obesity that appeared in Whitaker et al. (1997) is based on BMI, while the study used a composite definition of BMI and TSF to define overweight in children.

A note worth mentioning is that the interpretation of the CEA results should be done cautiously due to the difference in obesity definition. In the study, the probability of becoming obese in adulthood is drawn from Whitaker et al. (1997), which used the BMI cutoffs of 27.8 for men and 27.3 for women (p.870) (Whitaker et al., 1997); the BMI cutoff of obesity, however, is now 30 for all adults.

7. Were costs and consequences adjusted for differential timing?

YES

Both costs and consequences were discounted at an annual rate of 3% (p.1314).

8. Was an incremental analysis of costs and consequences of alternatives performed?

YES

An incremental analysis of costs and consequences of Planet Health is performed in the study (p.1320). The incremental cost and consequences are a result of the implementation of the program in relation to status quo.

9. Was allowance made for uncertainty in the estimates of costs and consequences?

YES

Both one-way and multivariate sensitivity analyses are demonstrated in the study (p.1321). Uncertainties included in the analysis are:
(1) The transitional probabilities of becoming obese from age 14 to 21-29.
(2) The transitional probabilities of becoming obese from age 21-29 to 40.
(3) Years of health life scores between ages 40 and 64.
(4) Annual workdays lost (for productivity loss estimation)
(5) Annual discount rate
(6) Medical costs of obesity saved

The multivariate analysis showed the uncertainty of the base case ICER of $4,305 per QALY to range from $1,612 to $9,010 per QALY. Net benefits of CATCH range from -$8,579 to $53,392 (p.1321).

The authors also assert in their conclusions that all ICERs shown in the sensitivity analyses are still far lower than $30,000 per QALY, which is generally considered a cost-effective threshold in the US (though it has moved up to $50,000-100,000 at present).

10. Did the presentation and discussion of study results include all issues of concern to users?

UNCLEAR

The discussion of the study is fairly comprehensive. The authors point out the applicability of the ICERs and compare them to the general ICER threshold of $30,000 (at that time). They also claim that the range of ICERs of Planet Health is lower than that of treatments of certain diseases such as asymptomatic hypertension (ICERs range from $14,185 to $39,867) and low-cholesterol diet therapy (ICERs range from $37,000 to $236,000) (p.1321-2).

They assert that the main problem of the analysis is due to the lack of information for projecting the long-term health outcomes. They also state other limitations of the study, including the intervention costs being collected retrospectively, the number of prevented overweight adults being estimated rather than measured, limited availability of data sources, the relapse of overweight not being accounted for in the model, inconsistency in the definition of childhood obesity, only effectiveness of the program from girls being included in the analysis, and neglect of medical costs of obesity during adolescence.

The authors point out that the study can be useful for policymakers in allocating public resources (p.1322).

The authors fail to address potential issues other than economic concerns, such as political, societal, and ethical considerations, that may influence the decision as to whether or not the program should be adopted.

Assessing economic evaluation of Brown et al. (2007)

1. Was a well-defined question posed in answerable form?

YES

The authors explain the context of their study, which aims to evaluate whether or not the reduced prevalence of obesity due to the CATCH program, a school-based intervention to prevent obesity in childhood, is worth the additional costs (p.2).

They state that CEA and CBA were used to compare the value of CATCH in the intervention group with that of the control group (p.1).

The authors clearly state that their method is similar to Wang et al. (2003), and that the societal perspective is used to identify the relevant costs of CATCH (p.2). They estimate the direct medical costs of obesity from existing literature (Gorsky et al.) as Wang et al. (2003) did and similarly failed to account for costs of patients' travel time and informal care givers. No reason for exclusion is provided in the article (p.6).

2. Was a comprehensive description of the competing alternatives given? (that is, can you tell who did what to whom, where and how often?)

YES

Details of the study are referred to Coleman et al. (2005)'s study, which randomized 8 selected schools into the intervention and control groups (4 each). Of these, 224 girls and 249 boys participated in the control schools, whereas 119 girls and 224 boys were in the intervention schools. The intervention was delivered to children over the three years of the CATCH program.

A subgroup analysis was conducted among Hispanic children, as they accounted for 93% of participants (p.2).

The results of the trial showed that the prevalence of overweight increased by 1% for boys and 2% for girls in the intervention group, while prevalence in the control group increased by 9% for boys and 13% for girls (p.2) (Brown et al., 2007).

A do-nothing alternative seems to fit in this case, given the type of intervention and other contexts. It is understandable that the increasing trend of obesity in children is the major concern in this case. Even though many child obesity interventions have been proposed to tackle this public health issue, none seems widely accepted as a standard practice.

3. Was the effectiveness of the programmes or services established?

YES

The primary outcome of the randomized controlled trial is the overweight status, which is defined as BMI at or above the 85^{th} percentile for sex and age (p.2). All anthropometric measurements were collected by trained research assistants, who attended and passed the reliability training to collect the data (p.2). The overweight status was converted to the prevalence of overweight in each group over the period of 3 years.

The major concern of the outcome from a 3-year intervention is the sustainability of the effect. The duration of 3 years reflects the effectiveness of CATCH in the short to intermediate terms; however, it is uncertain whether the reduced prevalence of obesity will be sustained throughout the childhood period.

The primary outcome of the study is the QALYs gained between ages 40 and 64. QALY calculation is based on the link between obesity in childhood and increased risks of being obese in adulthood, which leads to higher morbidity and mortality rates. Using QALY is considered a standard health outcome in the economic evaluation.

4. Were all the important and relevant costs and consequences for each alternative identified?

UNCLEAR

As mentioned earlier, the authors failed to mention the costs of patients' travel time and informal care givers, even though other relevant costs to the societal perspective, including training costs, overhead, operation costs, future medical costs of obesity saved, and costs of productivity loss reverted, were mentioned and sensitivity analysis was performed. The missing information may not affect the result to a great extent, but it is difficult to tell.

5. Were costs and consequences measured accurately in appropriate physical units (for example, hours of nursing time, number of physician visits, lost work-days, gained life-years)?

YES

The costs of CATCH were estimated for all participants over the period of 3 years (p.4). The obesity-related medical expenses and costs of productivity loss of both males and females were measured for life long (p.5).

The CATCH program comprises 4 components: CATCH PE (physical education); CATCH EAT SMART, implemented during the school meal; separate curricula for 3rd, 4th, and 5th grades implemented in the classroom; and HOME TEAM, implemented in the classroom and at home (Heath and Coleman, 2003).

The authors refer to the program evaluation process in another article (Heath and Coleman, 2002). Heath and Coleman (2002) describe the program monitoring as consisting of 4 aspects: 1) observation of physical activity – self-reported physical activity, heart rate monitoring, and motion sensors were used to assess the validity and reliability of observations from the trained observers, who were sent to evaluate children's physical activity; 2) menu and recipe collection – the head cook and cafeteria at each school were interviewed for details on how breakfast and lunch were provided, and the information was converted to nutrition density information using software; 3) structured interviews – specific details of program implementation were evaluated by a set of questionnaires. Information was gathered from interviewing school staff, students, and their families; and

4) process evaluation surveys – PE teachers, cafeteria staff, and classroom teachers completed well-designed questionnaires about the implementation of CATCH (p.446-8).

Overall, the trial contains high-quality measurements of the program delivery and data collection process, resulting in the reliability of the causal inference and outcomes reported.

6. Were costs and consequences valued credibly?

YES – for cost; NO – for outcome

The intervention costs and medical costs are based on charges according to the societal perspective (p.1, 4-5). The medical costs of obesity were derived by an economic model by Oster et al. (1999) that estimated the costs based on 5 obesity-related diseases: hypertension, hypercholesterolemia, type 2 diabetes, coronary heart disease, and stroke. Referring the medical costs of obesity from Oster et al.'s seems to affect accuracy of the cost estimate gradually. The medical costs of obesity in Oster et al. were estimated in the obese between the ages of 35 and 64, whereas QALYs in cost-effectiveness analysis were assessed between ages 40 and 65, leading to a small inconsistency of the age of interest. Moreover, the BMI cutoff of obesity applied in Oster et al. (1999) is 32.5 kg/m^2, while the authors use the BMI of 30 to define obesity in cost-effectiveness analysis (p.5).

The authors also conducted a sensitivity analysis showing the ICERs when applying the costs from Gorsky et al. (1996), instead of from Oster et al.. This improves the consistency of the estimate and allow the results to compare with Wang et al. (2003)'s (p.5). However, using Gorsky et al.'s in the medical cost saved necessitate the medical costs for females being substituted for males, as Gorsky et al. estimated the costs based in women only.

Inconsistency in the definition of obesity also affects the estimation of the long-term health outcome. The authors draw links among overweight in children, the increased risk of becoming overweight by ages 21-29, and said risk by age 40. The BMI cutoff for obesity at ages 21-29 is 27.8 for men and 27.3 for women (p.870) (Whitaker et al., 1997), whereas the rest of the analysis uses the BMI cutoff of 25 (p.1316) (Wang et al., 2003). Moreover, the definition of childhood obesity that appeared in Whitaker et al. (1997) is based on BMI, while the study used a composite definition of BMI and TSF to define overweight in children.

A note worth mentioning is that the interpretation of the CEA results should be done cautiously due differences in the definition of obesity. In the study, the probability of becoming obese in adulthood is drawn from Whitaker et al. (1997), which used the BMI cutoffs of 27.8 for men and 27.3 for women (p.870) (Whitaker et al., 1997); the BMI cutoff of obesity, however, is now 30 for all adults.

7. Were costs and consequences adjusted for differential timing?

YES

Both costs and consequences were discounted at an annual rate of 3% (p.2).

8. Was an incremental analysis of costs and consequences of alternatives performed?

YES

The incremental costs and consequences of CATCH are performed in the study (p.7). The incremental cost and consequences result from the implementation of the program in relation to status quo.

9. Was allowance made for uncertainty in the estimates of costs and consequences?

YES

A multivariate sensitivity analysis is performed in the study, which takes into account 48 variables (p.8). Among the included variables are the transitional probabilities of becoming obese from age 14 to 21-29, the transitional probabilities of becoming obese from ages 21-29 to 40, years of health life scores between ages 40 and 64, costs of productivity loss, and medical costs of obesity saved.

The multivariate analysis showed the uncertainty of the base case ICER of $900 per QALY to range from $900 to $1,143 per QALY. Net benefits of CATCH range from $23,707 to $27,453 (p.8).

The authors also asserted in the conclusion that all ICERs shown in the sensitivity analyses are still far much lower than $30,000 per QALY, which is generally considered a cost-effective threshold in the US (it moved up to $50,000-100,000 now).

10. Did the presentation and discussion of study results include all issues of concern to users?

UNCLEAR

The authors compare the results to those of the previous cost-effectiveness study by Wang et al. (2003) in several aspects (p.8-10). They also point out that the ICER of CATCH is lower than the threshold value of $30,000 and conclude that the CATCH program is very cost-effective. The authors assert that school-based health promotion interventions such as CATCH should be expanded, as its effectiveness was proven to be worth the costs (p.10).

The authors acknowledge two limitations of the study, which are the uncertainty of trends in medical costs and the limited availability of data regarding the medical costs of obesity (p.10).

The authors failed to describe other potential issues that may influence in decision making process whether or not the program should be adopted other than the economic perspective, such as political concern, societal issue, or ethical consideration.

Assessing economic evaluation of Moodie et al. (2008)

1. Was a well-defined question posed in answerable form?

YES

The authors describe the purpose of their study, which is to evaluate whether or not the reduced number of obese children due to the intervention of the LEAP trial, a family-based GP-mediated intervention to treat obesity in children, is worth the additional costs (p.364).

The authors state that they conducted CEA to compare the value of LEAP with that of the status quo in the control (p.363).

The authors assert that a societal perspective was applied to identify the relevant costs of LEAP for either the entire lifespan of 100 years, including intervention costs, time and travel costs of children and parents, and the future medical costs saved shown in terms of cost offsets (p.368-70, 372).

Overall, the authors clearly define the question in an answerable format, including the costs and outcome of the intervention as well as the comparator – the status quo in the controlled group.

2. Was a comprehensive description of the competing alternatives given? (that is, can you tell who did what to whom, where and how often?)

YES

The authors describe in the article that the LEAP trial is a small randomized controlled trial that aimed to treat obesity in children through lifestyle, diet, and exercise modification, as well as brief solution-focused techniques. They state that the trial was conducted in Melbourne in 2002-2005, involving 34 GPs and 73 children in the intervention arm and 80 children in the control group. The primary outcome of the trial is the BMI at 9 months. (p.364-5)

A do-nothing alternative seems to fit in this case, given the type of intervention and other contexts. It is understandable that the increasing trend of obesity in children is the major concern in this case. Even though many child obesity interventions have been proposed to tackle this public health issue, none seem widely accepted as a standard practice.

3. Was the effectiveness of the programmes or services established?

YES

An economic model was applied to estimate DALYs saved based on the change in BMI. The results of the trial showed that the mean BMI decreased by 0.25 kg/m^2 with 95%CI (-0.62, 0.12). One sample in the intervention group was considered an outlier and removed in the sensitivity analysis, leaving a change in the mean BMI of -0.31 kg/m^2 with 95%CI (-0.66,0.04) (p.365).

The major concern of the outcome from a 9-month intervention is the sustainability of the effect. The duration of 9 months reflects the effectiveness of LEAP in the short to intermediate terms; however, it is uncertain whether or not the reduced prevalence of obesity will be sustained throughout the childhood period.

The intervention in the LEAP trial also encounters selection bias. Though LEAP is a randomized controlled clinical trial, the recruitment criteria for children are not generalized to all obese children. The trial excluded extremely obese children, defined as a BMI z-score at or above 3.0 (p.364). This exclusion criterion does not seem to reflect what would happen in regular practice, thereby lowering the applicability of the results.

4. Were all the important and relevant costs and consequences for each alternative identified?

YES

The authors include all costs, including intervention costs, capital costs, operation costs, and medical costs illustrated in terms of cost offsets. Details of intervention related costs are as shown in Table 1 (p.369-70). However, the authors do not show how the medical costs saved were calculated. This could potentially create bias in the results.

Concerning the consequences, the authors specify DALY as a long-term health outcome of the study according to the societal perspective (p.365).

5. Were costs and consequences measured accurately in appropriate physical units (for example, hours of nursing time, number of physician visits, lost work-days, gained life-years)?

UNCLEAR – for cost; YES – for outcome

The costs of LEAP were estimated for all participants over the period of 9 months (p.363). However, the authors fail to specify the duration of the obesity-related medical expense estimate. As mentioned earlier, this could lead to potential bias on the cost side.

The authors describe the intervention delivery in detail, from GPs and children recruitment to program delivery. First, GPs were recruited and trained to do solution-focused family therapy. Obese and overweight children according to the IOTF definition between the ages of 5 and 9 whose BMI z-scores are less than 3.0 were recruited into the trial. It was then arranged for recruited children to have a 40-minute session (or longer) with the GP, followed by 3 20-to-40-minute visits over the next 3 months. Benefits of the program were assessed by health gain, measured primarily in terms of BMI at 9 months adjusted for baseline BMI. (p.364-5)

The changes in BMI were recorded from all participants and used to estimate DALYs saved. DALYs saved were approximated from the childhood period to the age of 100 or death.

6. Were costs and consequences valued credibly?

NO – for cost; YES – for outcome

The authors do not specify how they estimated the obesity-related costs in the study. The cost offsets of AU$3.3M were used to represent the future benefits of the intervention in monetary units. However, no source of cost offsets estimation is given (p.372). The obesity-related costs saved in the future are a potential factor influencing ICER estimation. The reader would need to assess the reliability of obesity-related costs estimation and potential biases.

The effectiveness of the LEAP study was drawn from a randomized controlled trial. Though outcome maintenance may remain an issue, it can be assumed that the authors already make use of the best available evidence. The authors refer the DALY estimating method to that in Haby et al. (2006) without deriving it in detail. The technique was specific to Australian population, which is the population of interest in the study. The DALYs lost were estimated from childhood period to either death or 100 years.

7. Were costs and consequences adjusted for differential timing?

YES

Both costs and consequences were discounted at an annual rate of 3% (p.364).

8. Was an incremental analysis of costs and consequences of alternatives performed?

YES

An incremental analysis of the costs and consequences of LEAP was performed in the study (p.372). The incremental cost and consequences are a result of the implementation of the program in relation to status quo.

9. Was allowance made for uncertainty in the estimates of costs and consequences?

YES

Monte Carlo simulations were conducted using a number of parameters, as shown in Table 2 (p.371). Parameters used in the analysis include duration of recruitment and training, costs of training and equipment, patient recruitment rate and costs, and compliance of intervention delivery (p.371). However, the authors state that uncertainty analysis does not meet its purpose, which is to address the uncertainties resulting from assumptions made due to a lack of information, and omit it (p.372). The ICER of LEAP is estimated to be $4,670 per DALY with 95% uncertainty interval ranging from being dominated (the intervention costs more for less effect) to $100,000 per DALY (p.372).

10. Did the presentation and discussion of study results include all issues of concern to users?

YES

The authors compare the ICER of $4,670 to the Australian threshold value of $50,000 and conclude that LEAP is cost-effective. They also identify the weaknesses of the study, which are due to the small sample size and short follow-up period, and which could increase the ICER substantially and result in the intervention being cost-ineffective (p.374-5).

The authors provide a comprehensive discussion about using the results in real practice. Several points of consideration with regard to the CEA, such as strength of evidence, equity, acceptability, feasibility, sustainability, and policy considerations, are included in Table 4 (p.373).

They suggest alternative scenarios that can reduce costs in real practice – for example, coordination of staff recruitment or training or sharing resources with other relevant programs (p.375).

Assessing economic evaluation of Magnus et al. (2009)

1. Was a well-defined question posed in answerable form?

YES

The authors state the purpose of the study, which is to assess the cost-effectiveness of banning television (TV) advertising in Australia for energy-dense, nutrient-poor food and beverages during children's peak viewing times. The study employs the public health perspective (p.1).

The evidence used in the analysis is derived mainly from the authors' hypothetical assumptions, since there is no existing trial assessing the benefits of an intervention banning TV advertising of high fat/ high sugar food and beverages. Therefore, the results of the study must be interpreted and extrapolated carefully.

2. Was a comprehensive description of the competing alternatives given? (that is, can you tell who did what to whom, where and how often?)

YES

The authors compare the hypothetical TV advertisement ban with the current practice in Australia – that is, the Children's TV standards 16-23, which allows advertisements up to 5 minutes to be shown during designated timeslots every 30 minutes for a total of 5 hours per week (p.2).

A do-nothing alternative seems to fit in this case, given the type of intervention and other contexts. It is understandable that the increasing trend of obesity in children is the major concern in this case. Even though many child obesity interventions have been proposed to tackle this public health issue, none seem widely accepted as a standard practice.

3. Was the effectiveness of the programmes or services established?

YES

The authors estimate the effectiveness of the intervention in terms of DALYs based on a hypothetical model using existing evidence. Since research illustrating the effectiveness of banning TV advertising is limited, assumptions were made in order to estimate the impact of the intervention.

First, as no randomized controlled trial assessing the effect of reducing TV advertising of high fat/high sugar food exists, the authors use the results from another trial (Gorn and Goldberg, 1982). Gorn and Goldberg (1982) conducted a randomized controlled trial in 288 children between ages 5 and 8 to explore the impact of food and beverage advertising during camp over a period of 2 weeks. They found that children who were exposed to commercials for sweets ate more sweets than those who did not. Children with commercials exposure also drank more sugar-sweetened KoolAid, rather than orange juice, compared to their peers (Gorn and Goldberg, 1982). In their cost-effectiveness study, Magnus et al. (2009) constructed a relative risk of sweet consumption in children exposed to advertisements compared with their counterparts. The authors also assume that the effect of food choice was maintained for all age groups from 5 to 14 years of age (p.3-4).

Second, to draw the link between high fat/high sugar food consumption and weight gain, the authors categorize foods and beverages as core (less sweet) and non-core (sweeter) based on an article by Bell et al. (2005a). They assume that the average energy gained from consuming non-core food and beverages is equal to energy gained from consuming high fat/high sugar food advertised on TV, while energy gained from core food and beverages is used as an energy intake baseline. The incremental amount of energy intake was converted into weight gain by referring to the association between percent change in energy consumption and weight. Weight gain resulting from additional energy consumption was used in predicting changes in BMI by comparing weight with the average heights of Australian boys and girls by age. Finally, the authors construct an economic model, based on estimated BMI changes, to estimate the long-term health outcomes in terms of DALY.

These assumptions used in estimating weight gain are considered weak and potentially biased for many reasons. Firstly, there is no trial to show the actual effect of banning TV advertising of high fat/ high sugar food and beverages. The authors base their hypotheses on a small randomized controlled trial involving only 288 children between ages 5 and 8. The follow-up duration of the trial is considered very short – only 2 weeks; however, they assume that its effect would be maintained for all children between ages 5 to 14 regardless of gender and age. It is not practical to assume that children everywhere, regardless of age and gender, would have the same response to the banning of advertisements, as children at different ages have different interests and would react to the advertisements in different ways. Sweets may be more popular in school-age children than in young teenagers, for example, or the impact of sweets advertisement may decrease in female teens due to weight concerns. To assume that the effect of the short-term study would be maintained is significantly biased.

Secondly, the authors estimate the amount of energy intake by creating categories of core and non-core food and beverages. However, there is some overlap between core food/beverages and high fat/ high sugar food and beverages, such as fruit or juice. Eggs, milk and cheese were counted as core food, while cheesecake, which is made from these ingredients, is not. Assuming that high fat/high sugar and non-core food and beverages fall into the same category, therefore, is a poor method of estimating the energy density of the subjects' diets.

Thirdly, the association between reduced energy intake and weight loss per participant does not make sense. The mechanism of weight control often refers to the concept of energy balance, which consists of energy intake and outflow. The amounts of both intake and outflow are important in determining excessive energy consumption, which then leads to weight gain. Using an existing article (Bell et al., 2005a), the authors estimate the average amount of energy density intake for each individual by type of food and beverages without regard as to the amount of intake. Moreover, the authors do not mention differences in the rate of energy outflow by age and gender. This assumption not only oversimplifies the scientific evidence, but also fails to adhere to the concept of energy balance.

Finally, calculating BMI from average weight gain and assuming average heights by age and gender is considerably biased. A very tall man could gain a lot of weight while keeping his BMI in a healthy range. Assuming mean weight and height for all children is not practical in reality.

To conclude, the method of outcome estimation is lacking in validity and practicality, resulting in the unreliability of the downstream outcomes.

4. Were all the important and relevant costs and consequences for each alternative identified?

NO – for cost; YES – for outcome

The authors identify relevant costs, including the intervention and medical costs, but failed to state the overhead and operating costs in their analysis (p.4). However, the authors do not show how the medical costs saved were calculated. This could potentially lead to bias in the results. Moreover, since it is a hypothetical scenario, it is unclear whether all relevant costs are accounted for.

Concerning the consequence, the authors specify DALY as a long-term health outcome of the study according to the public health perspective (p.4-5).

5. Were costs and consequences measured accurately in appropriate physical units (for example, hours of nursing time, number of physician visits, lost work-days, gained life-years)?

UNCLEAR – for cost; YES – for outcome

The authors include as intervention costs the costs of hiring 2 staff for overseeing regulation, and the cost offsets are represented by the future health sector costs of obesity saved. However, they failed to state the duration for which the 2 staff would be hired. In addition, they did not specify how they estimated the cost offsets, including the number of obesity-related conditions taken into consideration, duration of follow-up, and whether costs or charges are used in the calculation.

The changes in BMI were recorded from all participants and used to estimate DALYs saved. DALYs saved were approximated from the childhood period to the age of 100 or death.

6. Were costs and consequences valued credibly?

NO

The authors estimate the total intervention costs as the costs of hiring 2 staff to enforce and monitor the regulation, with costs totaling AU$130,000. These estimated costs are not pragmatic because the job description and the number of staff are unclear. Furthermore, the overhead and operating costs are excluded in the estimation.

The authors also estimate cost offsets, which represent the future health sector costs saved, as AU$300M without providing any reference or calculating methods. Additionally, they estimate the cost offsets based on the assumption that preventing an increased BMI in all children results in health gain in adulthood. However, the majority of children is not obese and will not gain any benefit from the prevention of increased BMI. Therefore, the actual cost offsets derived from obese children only should be far less than AU$300M.

Regarding outcome estimate, the intervention did not target only obese children, but rather all children aged 5-14 years. However, the authors account for health benefits gained from preventing increased BMI in terms of the reduced risk of adverse consequences of obesity. This is not logical, as not all children are obese to begin with. Normal-weight children will not gain any health benefits from avoiding risks of obesity when their BMIs decrease. Moreover, children with BMI in the healthy range are the majority, not those with high BMIs. This assumption does not make scientific sense.

7. Were costs and consequences adjusted for differential timing?

YES

Both costs and consequences were discounted at an annual rate of 3% (p.1).

8. Was an incremental analysis of costs and consequences of alternatives performed?

YES

An incremental analysis the costs and consequences of the intervention was performed in the study (p.4). The incremental cost and consequences result from the implementation of the program in relation to the status quo.

9. Was allowance made for uncertainty in the estimates of costs and consequences?

YES

The authors calculate ICERs in several alternative scenarios, including 1) changing the methods to categorize food and beverages; and 2) assuming 30 staff instead of 2 staff monitoring compliance. However, they did not provide information about the alternative methods used for food and beverages (p.5).

The ICERs of the base case scenario and sensitivity analysis indicated that the intervention is dominant, meaning that it costs less for more health gain (p.4-5).

10. Did the presentation and discussion of study results include all issues of concern to users?

YES

The authors compare the dominant ICER to the Australian threshold value of $50,000 and conclude that banning TV advertising is a very good buy. They also declare the weaknesses of the study to be due to the small sample size and short follow-up period, which could increase the ICER substantially and result in the intervention being cost-ineffective (p.5).

The authors provide a comprehensive discussion about the practical applicability of the results. Several aspects concerning the CEA, such as strength of evidence, equity, acceptability, feasibility, sustainability, and policy considerations, are included in Table 5 (p.6).

They discuss the possibility of a synergistic effect of the intervention with other obesity programs to reduce prevalence. They also state that it is uncertain what kind of advertising will replace the high fat/high sugar spots during children's timeslots if the ban is employed. The authors also suggest that the availability and pricing of alternative foods such as fruits and vegetables should render them competitive to sweets and snacks, and should therefore present no barrier to consumer choice. The limitations of the effectiveness estimate due to limited evidence are addressed as well. In addition, the authors consider whether the quality of children's programming will be lowered if revenue from food advertising is lost, claiming that the impact is found to be less than expected (though little information is provided) (p.5-8).

Assessing economic evaluation of Moodie et al. (2009)

1. Was a well-defined question posed in answerable form?

YES

The authors state the purpose of the study, which is to assess cost-effectiveness of the walking school bus (WSB) program for primary school children from 5 to 7 years of age in Australia. The societal perspective is employed in the study (p.1).

The evidence used in the analysis is derived mainly from the authors' hypothetical assumptions, as even though the walking school bus program does exist, no randomized controlled trial has ever explored its effectiveness. Therefore, the results of the study must be interpreted and extrapolated carefully.

2. Was a comprehensive description of the competing alternatives given? (that is, can you tell who did what to whom, where and how often?)

YES

The authors assert that WSB is compared with the do-nothing scenario (p.2). A do-nothing alternative seems to fit in this case given the type of intervention and other contexts. It is understandable that the increasing trend of obesity in children is the major concern in this case. Even though many child obesity interventions have been proposed to tackle this public health issue, none seem widely accepted as a standard practice.

3. Was the effectiveness of the programmes or services established?

YES

The authors estimate the effectiveness of the intervention based on a hypothetical model using existing evidence over a period of one year. Since research illustrating the effectiveness of WSB is limited, assumptions were made to estimate the impact of the intervention.

First, the extra time spent on walking to school (around 28 mins), additional energy spent (difference between energy spent on sitting and walking), and the number of participating days were used to calculate the additional energy expense per year. The proportion of increased energy expenditure was employed to estimate average weight loss from WSB. Then, the mean height and weight of children aged 5-7 years were used to calculate the baseline BMI and the estimated lowered BMI from WSB. Finally, the authors construct an economic model, based on estimated BMI changes, to estimate the long-term health outcomes in terms of DALY.

These assumptions used in estimating weight gain are considered weak and potentially biased for a number of reasons. Firstly, it is known that the energy expenditure of each individual varies based upon the metabolism rate by gender and age. Assuming the mean energy expense of walking for all leads to a biased estimation.

Secondly, the association between energy expenditure from WSB and weight loss per participant does not make sense. The mechanism of weight control is often represented by the concept of energy balance, which consists of energy intake and outflow. However, the authors use only the incremental energy outflow to predict the lowered weight, while ignoring the possibility that children might eat more after activities. This assumption not only oversimplifies the scientific evidence, but also fails to correspond to the concept of energy balance.

Thirdly, they assume a 100% sustainability of the program effectiveness, while the costs and outcomes were estimated over a period of one year. It is doubtful whether the effectiveness can be maintained until adulthood.

Fourthly, the authors premise health benefit evaluation on the wrong assumption. From the base case parameters, the authors show that the mean BMIs of boys and girls were 16.24 and 16.47 kg/m^2 (p.3). This implies that they are not obese according to either IOTF or common definitions of obesity. DALY calculation is based on the assumption that the reduction of BMI brings about health benefits from decreasing obesity-related health risks. Therefore, it does not make sense to assume that the risks of obesity-related conditions are reduced in non-obese children when their BMIs decrease.

Lastly, calculating BMI from average weight gain and assuming average heights by age and gender is considerably biased. A very tall man could gain a lot of weight while keeping his BMI in a healthy range. Assuming mean weight and height for all children is not practical in reality.

To conclude, the method of estimating outcomes is lacking in validity and practicality, resulting in the unreliability of the downstream outcomes.

4. Were all the important and relevant costs and consequences for each alternative identified?

UNCLEAR – for cost; YES – for outcome

The authors identify relevant costs, including the intervention costs, medical costs, overhead and operating costs in the analysis (p.4). However, the authors do not show how the medical costs saved were calculated. This could potentially lead to bias in the results. Moreover, since it is a hypothetical scenario, it is unclear whether all relevant costs are accounted for.

Concerning the consequence, the authors specify DALY as a long-term health outcome of the study according to the societal perspective (p.4).

5. Were costs and consequences measured accurately in appropriate physical units (for example, hours of nursing time, number of physician visits, lost work-days, gained life-years)?

NO – for cost; YES – for outcome

The authors show the operating pathway and its relevant costs in Figure 2 (p.6), including operating costs at all phases of the program. They assume that the program is in the steady state, so the set-up, research, and development costs prior to the recruitment were excluded (p.4). However, they do not clearly state whether or not capital costs were included.

The authors estimate the medical costs of obesity in terms of cost offsets (p.8). However, they do not specify how they estimate the cost offsets, including the number of obesity-related conditions taken into consideration, duration of follow-up, and whether costs or charges are used in the calculation.

The changes in BMI were recorded from all participants and used to estimate DALYs saved. DALYs saved were approximated from the childhood period to the age of 100 or death.

6. Were costs and consequences valued credibly?

NO

As discussed earlier, the authors do not clearly show if they included capital costs in the estimation. It is also unclear whether they use costs or charges in the analysis.

Regarding outcome estimate, the effectiveness of the intervention was derived from a hypothetical scenario, which contains uncertainties. Furthermore, the intervention did not target obese children, but rather those aged 5-7 years regardless of obesity status. However, the authors account for health benefits gained from preventing increased BMI in terms of the reduced risk of adverse consequences of obesity. This is not logical, as not all children are obese to begin with. Normal-weight children will not gain any health benefits from avoiding the risks of obesity when their BMIs decrease. Additionally, children with BMI in the healthy range are the majority, not the ones with high BMIs. This assumption does not make scientific sense.

7. Were costs and consequences adjusted for differential timing?

YES

Both costs and consequences were discounted at an annual rate of 3% (p.1).

8. Was an incremental analysis of costs and consequences of alternatives performed?

YES

An incremental analysis of costs and consequences of the intervention is performed in the study (p.8). The incremental cost and consequences result from the implementation of the program in relation to the status quo.

9. Was allowance made for uncertainty in the estimates of costs and consequences?

YES

The authors include alternatives in the sensitivity analysis, such as additional benefits from non-obesity related objectives, cost-cutting measures, and increasing the number of participants and schools. However, WSB is not cost-effective for most alternative scenarios (p.9).

10. Did the presentation and discussion of study results include all issues of concern to users?

YES

The authors discuss the possibility of WSB being cost-ineffective due to low recruitment and uptake of the program. They acknowledge that the purpose of WSB is not actually to reduce obesity, but to promote active transport behavior to students (p.8). They also address issues regarding the sustainability and feasibility of the program due to a lack of funding (p.10).

The authors provide a comprehensive discussion about the practical applicability of the results. Several aspects concerning the CEA, such as strength of evidence, equity, acceptability, feasibility, sustainability, and policy considerations, are included in Table 5 (p.8).

Assessing economic evaluation of Ananthapavan et al. (2009)

1. Was a well-defined question posed in answerable form?

YES

The authors state the purpose of the study, which is to assess whether or not the weight reduction from treating severely obese Australian adolescents with laparoscopic adjustable gastric banding (LAGB) is worth the additional costs. The societal perspective is employed in the study (p.377).

The authors state that in the CEA, laparoscopic adjustable gastric banding is compared with a conventional management of adolescent obesity, which is referred to as 'no treatment' in the paper (p.378-383).

The evidence used in the analysis is derived mainly from the authors' hypothetical assumptions, as even though laparoscopic adjustable gastric banding does exist, no randomized controlled trial has ever explored its effectiveness. Therefore, the results of the study must be interpreted and extrapolated carefully.

Overall, the authors clearly define the question in an answerable format, including the costs and outcome of the intervention as well as the comparator – conventional management or status quo in the controlled group.

2. Was a comprehensive description of the competing alternatives given? (that is, can you tell who did what to whom, where and how often?)

YES

The authors describe in the article that LAGB is a bariatric surgery used in severely obese people. LAGB is proven effective and causes fewer complications than other operations. The effectiveness of LAGB used in the CEA is based on the medical records of 28 adolescents (2 males, 26 females) aged 15-19 at the Centre for Bariatric Surgery, Avenue Hospital, Australia from 1996 to May 2005. The primary outcome of LAGB is changes in BMI (p.378)

A conventional management of adolescent obesity, which is do-nothing in real practice, alternative seems acceptable in this case. It is understandable that conventional management is most likely to be applied in managing adolescent obesity, while there is no standard practice for treating adolescent obesity. However, for those who suffer from morbid obesity, it is also likely that intensive treatment, such as medical treatment or other types of surgery, may eventually be considered an option, .

3. Was the effectiveness of the programmes or services established?

YES

An economic model was applied to estimate DALYs saved based on the change in BMI. The results from the 3-year follow-up data showed that the mean BMI decreased by 13.93 ± 2 kg/m^2.

The major concern of the outcome is the sustainability of the effect. The duration of 3 years, which was assumed to be sustained for 4 years in the model, reflects the effectiveness of LAGB in the intermediate term (p.378); however, it is uncertain whether or not the effectiveness will be sustained in the long term. Due to limited available information, the authors make a number of assumptions in the model to estimate the long-term outcome, such as technological improvement, surgical band failure, and other surgical complications. While it is unlikely that surgical technique will not improve during the next few decades, the lack of long-term data lead the authors to guess that the surgical bands will need to be replaced twice per life time, and that no unforeseen surgical complications will happen.

4. Were all the important and relevant costs and consequences for each alternative identified?

YES

The authors include all costs, including intervention costs, capital costs, operation costs, and medical costs illustrated in terms of cost offsets. The details of the intervention-related costs are as shown in Table 1 (p.379-81).

Concerning the consequences, the authors specify DALY as a long-term health outcome of the study according to the societal perspective (p.377).

5. Were costs and consequences measured accurately in appropriate physical units (for example, hours of nursing time, number of physician visits, lost work-days, gained life-years)?

YES

The authors demonstrate clearly the cost calculation of LAGB in the paper (p. 379-381). In this regard, the medical cost offsets were estimated from Carter et al. (2008) (p.381,383).

The changes in BMI were recorded from all participants and used to estimate DALYs saved. DALYs saved were approximated from the childhood period to the age of 100 or death.

6. Were costs and consequences valued credibly?

NO

In the economic model, the intervention was assumed to cover all severely obese adolescents between ages 14 and 19, with BMI ≥35, leading to a total number of 4,120 eligibles in 2001 in Australia. However, all costs and outcomes of LAGB in the model were derived from the medical records of 28 patients in a single hospital. Not only is the data used not randomized, but all 28 patients were also treated by the same hospital (p.378). This could potentially lead to a lack of generalizability and credibility in the estimate. The reader would need to assess the reliability of obesity-related costs estimation and potential biases.

The authors derive all cost components in the estimate, including fixed and variable direct medical costs from surgery, patients' and parents' time, and the cost offset (p.379-381).

The authors base their DALY estimating method on that in Haby et al. (2006) without deriving it in detail. The technique was specific to Australian population, which is the population of interest in the study. The DALYs lost were estimated from the childhood period to either death or 100 years.

7. Were costs and consequences adjusted for differential timing?

YES

Both costs and consequences were discounted at an annual rate of 3% (p.377).

8. Was an incremental analysis of costs and consequences of alternatives performed?

YES

An incremental analysis of the costs and consequences of LAGB was performed in the study (p.383). The incremental cost and consequences are a result of the implementation of the program in relation to the comparator.

9. Was allowance made for uncertainty in the estimates of costs and consequences?

YES

Monte Carlo simulations of 3,000 iterations were conducted using a number of parameters, as shown in Table 2 (p.383). Parameters used in the analysis include patient recruitment rate and costs, resource use, effectiveness of LAGB, and medical cost offsets (p.380-381). The ICER of LEAP is estimated to be $4,400 per DALY with 95% uncertainty interval $2,900-6,120 DALY (p.383).

10. Did the presentation and discussion of study results include all issues of concern to users?

YES

The authors compare the ICER of $4,470 to the Australian threshold value of $50,000 and conclude that LAGB is cost-effective. They also identify the weaknesses of the study, which are due to the short follow-up period, and assumptions made due to the lack of available data, which include the long-term effectiveness of LAGB, complications of surgery, and biased source of data from one medical setting (p.383-384).

The authors provide a comprehensive discussion about using the results in real practice. Several points of consideration with regard to the CEA, such as strength of evidence, equity, acceptability, feasibility, sustainability, and policy considerations, are included in Table 4 (p.382).

Assessing economic evaluation of Moodie et al. (2010)

1. Was a well-defined question posed in answerable form?

YES

The authors state the purpose of the study, which is to assess the cost-effectiveness of the Active After-school Communities (AASC) program for children from 5 to 14 years of age in Australia. The societal perspective is employed in the study (p.1585).

The evidence used in the analysis is derived mainly from the authors' hypothetical assumptions, as even though the AASC program does exist, no randomized controlled trial has ever explored its effectiveness. Therefore, the results of the study must be interpreted and extrapolated carefully.

Overall, the authors clearly define the question in an answerable format, including the costs and outcome of the intervention as well as the comparator – conventional management or status quo in the controlled group.

2. Was a comprehensive description of the competing alternatives given? (that is, can you tell who did what to whom, where and how often?)

YES

The authors state that AASC is an after-school physical activity program aimed at reducing obesity in children aged 5-14 (p.1585). The activity sites may include schools or organizations that participate in delivering services, which provide 2-3 sessions weekly, 8 weeks per semester, 4 semesters a year (p.1586). Economic evaluation narrows down the population of interest to children between the ages of 5 and 11 (p.1587).

The authors compare AASC with the do-nothing scenario (p.1586). A do-nothing alternative seems to fit in this case, given the type of intervention and other contexts. It is understandable that the increasing trend of obesity in children is the major concern in this case. Even though many child obesity interventions have been proposed to tackle this public health issue, none seem widely accepted as a standard practice.

3. Was the effectiveness of the programmes or services established?

YES

The authors estimate the effectiveness of the intervention based on a hypothetical model using existing evidence over a period of one year. Since research illustrating the effectiveness of AASC is limited, assumptions were made to estimate the impact of the intervention.

First, the extra time spent on the program (around 60 mins), additional energy spent (difference between energy spent on sitting and walking), and the number of participating days were used to calculate the additional energy expense per year. The proportion of increased energy expenditure was then used to estimate average weight loss from AASC. Then, the mean height and weight of children aged 5-9 and 10-11 years were used to calculate the baseline BMI and the estimated lowered BMI from WSB. Finally, the authors construct an economic model, based on estimated BMI changes, to estimate the long-term health outcomes in terms of DALY.

These assumptions used in estimating weight gain are considered weak and potentially biased for a number of reasons. Firstly, it is known that the energy expenditure of each individual varies according to the metabolism rate by gender and age. Assuming a mean energy expense from activities for all participants leads to a biased estimation at the individual level.

Secondly, the association between energy expenditure from AASC and weight loss per participant does not make sense. The mechanism of weight control is often represented by the concept of energy balance, which consists of energy intake and outflow. However, the authors use only the incremental energy outflow to predict the lowered weight, while ignoring the possibility that children might eat more after physical activity. This assumption not only oversimplifies the scientific evidence, but also fails to adhere to the concept of energy balance.

Thirdly, they assume a 100% sustainability of program effectiveness, while the costs and outcomes were estimated over a period of one year. It is doubtful as to whether the effectiveness can be maintained until adulthood.

Fourthly, the authors premise health benefit evaluation on the wrong assumption. From the base case parameters, the authors show that the mean BMIs of boys and girls were 16.83 and 17.30 kg/m^2 for children aged 5-9, and 18.06 and 18.87 kg/m^2 for children aged 10-11 (p.1587). This implies that none of them are obese according to either IOTF or common definitions of obesity, and only a small fraction is considered overweight. DALY calculation is based on the assumption that the reduction of BMI brings about health benefits from decreasing obesity-related health risks. Therefore, it does not make sense to assume that the risks of obesity-related conditions are reduced in non-obese children when their BMIs decrease.

Lastly, calculating BMI from average weight gain and assuming average heights by age and gender is considerably biased. A very tall man could gain a lot of weight while keeping his BMI in a healthy range. Assuming mean weight and height for all children is not practical in reality.

To conclude, the method of estimating outcomes is lacking in validity and practicality, resulting in the unreliability of the downstream outcomes.

4. Were all the important and relevant costs and consequences for each alternative identified?

UNCLEAR – for cost; YES – for outcome

The authors identify relevant costs, including the intervention costs, medical costs, overhead and operating costs in the analysis (p.1588). However, the authors do not show how the medical costs saved were calculated. This could potentially lead to bias in the results. Moreover, since it is a hypothetical scenario, it is unclear whether all relevant costs are accounted for.

Concerning the consequences, the authors specify DALY as a long-term health outcome of the study according to the societal perspective (p.1586-1587).

5. Were costs and consequences measured accurately in appropriate physical units (for example, hours of nursing time, number of physician visits, lost work-days, gained life-years)?

NO – for cost; YES – for outcome

The authors show the operating pathway and its relevant costs in Figure 2 (p.1588), including operating costs at all phases of the program. They assume that the program is in the steady state, so the set-up, research, and development costs prior to the recruitment were excluded (p.1588). However, they do not clearly state whether or not capital costs were included.

The authors estimate the medical costs of obesity in terms of cost offsets (p.1590). However, they do not specify how they estimate the cost offsets, including the number of obesity-related conditions taken into consideration, duration of follow-up, and whether costs or charges are used in the calculation.

The changes in BMI were recorded from all participants and used to estimate DALYs saved. DALYs saved were approximated from the childhood period to the age of 100 or death.

6. Were costs and consequences valued credibly?

NO

In the economic model, the intervention was assumed to extend eligibility to all Australian children aged 5-11, leading to a total number of 57,757 participants receiving benefits from it in 2001.

As discussed earlier, the authors do not clearly show if they included capital costs in the estimation. It is also unclear as to whether they use costs or charges in the analysis.

Regarding the outcome estimate, the effectiveness of the intervention was derived from a hypothetical scenario, which contains uncertainties. Furthermore, the intervention did not target obese children, but rather those aged 5-11 years regardless of obesity status. However, the authors account for health benefits gained from preventing increased BMI in terms of the reduced risk of adverse consequences of obesity. This is not logical, as not all children are obese to begin with. Normal-weight children will not gain any health benefits from avoiding the risks of obesity when their BMIs decrease. Additionally, children with a BMI in the healthy range are the majority, not the ones with high BMIs. This assumption does not make scientific sense.

7. Were costs and consequences adjusted for differential timing?

YES

Both costs and consequences were discounted at an annual rate of 3% (p.1585).

8. Was an incremental analysis of costs and consequences of alternatives performed?

YES

An incremental analysis of costs and consequences of the intervention is performed in the study (p.1589-1590). The incremental cost and consequences result from the implementation of the program in relation to the status quo.

9. Was allowance made for uncertainty in the estimates of costs and consequences?

YES

The authors include alternatives in the sensitivity analysis, such as the percent of schools participating the program, time spent in the activity, and energy spent during the program. Monte Carlo simulations of 3,000 iterations were conducted using a number of parameters, as shown in Table 2 (p.1589). The ICER of AASC is estimated to be $82,000 per DALY with a 95% uncertainty interval of $40,000-165,000 DALY (p.1590).

10. Did the presentation and discussion of study results include all issues of concern to users?

YES

The authors discuss the possibility of AASC being cost-ineffective due to the modest impact of physical activity on BMI in general, as supported by evidence they also provide (p.1589-1590). They discuss uncertainty resulting from assumptions made in the model that might lower the ICER, such as time spent and energy expenditure during the program (p.1590).

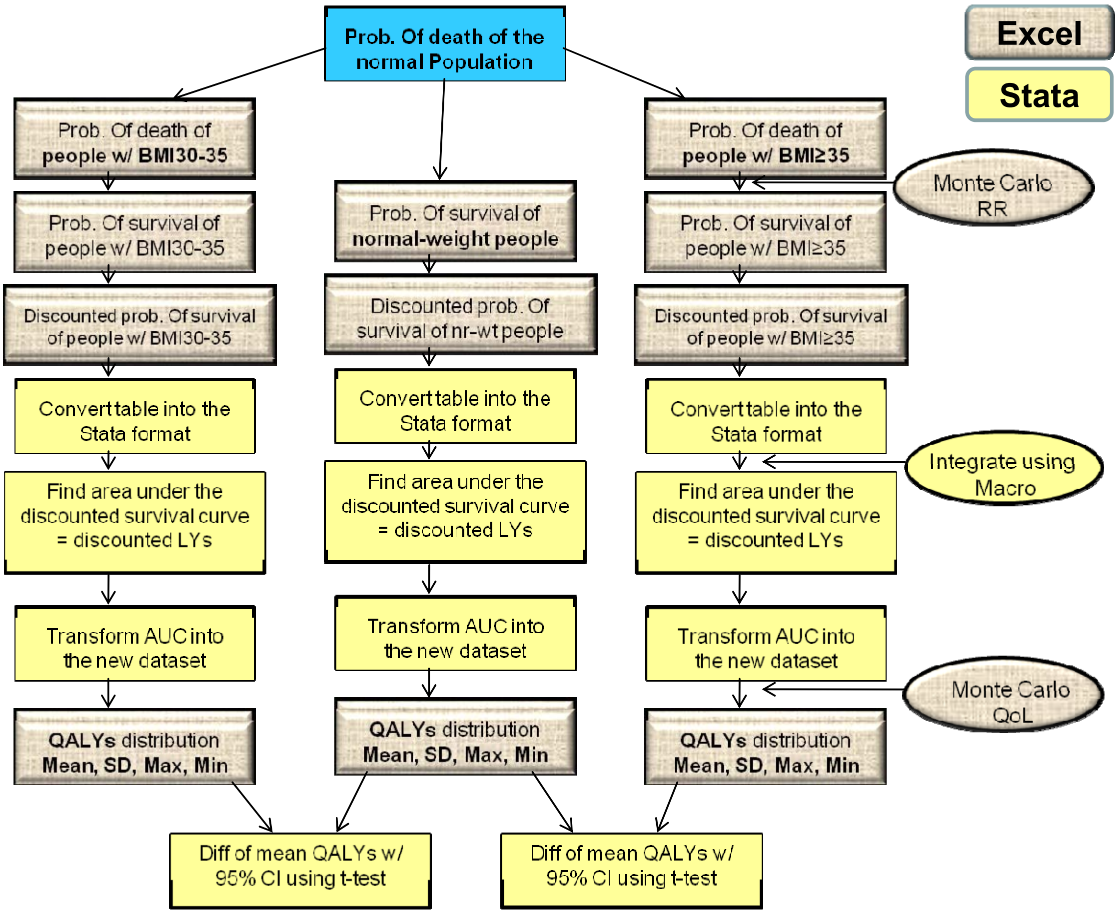
The authors provide a comprehensive discussion about the practical applicability of the results. Several aspects concerning the CEA, such as strength of evidence, equity, acceptability, feasibility, sustainability, and policy considerations, are included in Table 4 (p.1591).

Appendix 4: Estimation of the direct medical costs of obesity from 2027 to 2079

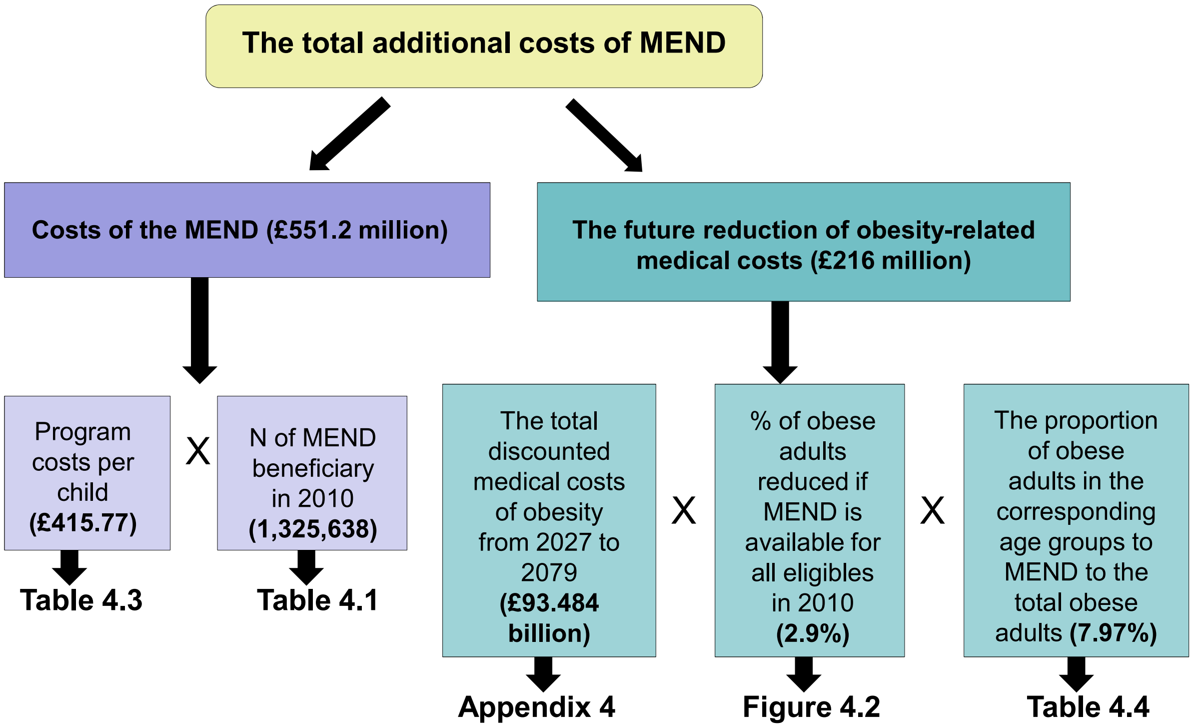
Year	Projected medical costs of obesity (billion)	Discounted (3.5%) medical cost into 2010 prices (billion)
2025	5.3	
2026	5.372	
2027	5.444	3.033
2028	5.516	2.970
2029	5.588	2.907
2030	5.66	2.845
2031	5.732	2.783
2032	5.804	2.723
2033	5.876	2.664
2034	5.948	2.605
2035	6.02	2.547
2036	6.092	2.491
2037	6.164	2.435
2038	6.236	2.380
2039	6.308	2.326
2040	6.38	2.273
2041	6.452	2.221
2042	6.524	2.170
2043	6.596	2.120
2044	6.668	2.070
2045	6.74	2.022
2046	6.812	1.974
2047	6.884	1.928
2048	6.956	1.882
2049	7.028	1.837
2050	7.1	1.793
2051	7.172	1.750
2052	7.244	1.708
2053	7.316	1.667
2054	7.388	1.626
2055	7.46	1.586
2056	7.532	1.548
2057	7.604	1.510
2058	7.676	1.472
2059	7.748	1.436
2060	7.82	1.400
2061	7.892	1.365
2062	7.964	1.331
2063	8.036	1.298
2064	8.108	1.265
2065	8.18	1.233
2066	8.252	1.202
2067	8.324	1.171

Year	Projected medical costs of obesity (billion)	Discounted (3.5%) medical cost into 2010 prices (billion)
2068	8.396	1.142
2069	8.468	1.113
2070	8.54	1.084
2071	8.612	1.056
2072	8.684	1.029
2073	8.756	1.002
2074	8.828	0.977
2075	8.9	0.951
2076	8.972	0.926
2077	9.044	0.902
2078	9.116	0.879
2079	9.188	0.856
Total discounted medical costs 2027-2079		**93.484**

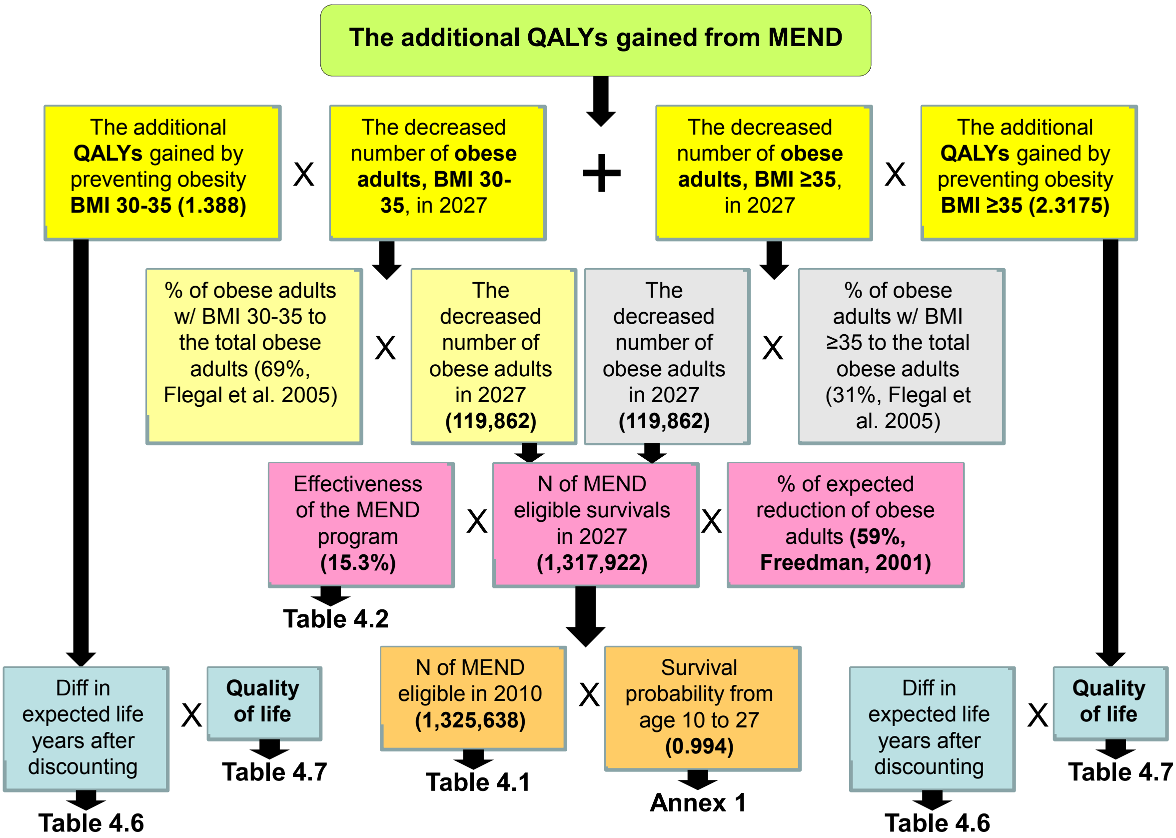
Appendix 5: Flow chart of the Monte Carlo simulation procedure in the sensitivity analysis in the economic evaluation


Prob. Of death of the normal Population
Excel
Stata
Prob. Of death of people w/ BMI30-35
Prob. Of survival of people w/ BMI30-35
Discounted prob. Of survival of people w/ BMI30-35
Convert table into the Stata format
Find area under the discounted survival curve = discounted LYs
Transform AUC into the new dataset
QALYs distribution Mean, SD, Max, Min
Prob. Of survival of normal-weight people
Discounted prob. Of survival of nr-wt people
Convert table into the Stata format
Find area under the discounted survival curve = discounted LYs
Transform AUC into the new dataset
QALYs distribution Mean, SD, Max, Min
Prob. Of death of people w/ BMI≥35
Monte Carlo RR
Prob. Of survival of people w/ BMI≥35
Discounted prob. Of survival of people w/ BMI≥35
Convert table into the Stata format
Integrate using Macro
Find area under the discounted survival curve = discounted LYs
Transform AUC into the new dataset
Monte Carlo QoL
QALYs distribution Mean, SD, Max, Min
Diff of mean QALYs w/ 95% CI using t-test
Diff of mean QALYs w/ 95% CI using t-test

Appendix 6: Flow chart for the long-term cost analysis in the economic evaluation


The total additional costs of MEND
Costs of the MEND (£551.2 million)
The future reduction of obesity-related medical costs (£216 million)
Program costs per child (£415.77)
X
N of MEND beneficiary in 2010 (1,325,638)
Table 4.3
Table 4.1
The total discounted medical costs of obesity from 2027 to 2079 (£93.484 billion)
X
% of obese adults reduced if MEND is available for all eligibles in 2010 (2.9%)
X
The proportion of obese adults in the corresponding age groups to MEND to the total obese adults (7.97%)
Appendix 4
Figure 4.2
Table 4.4

Appendix 7: Flow chart for the long-term effectiveness analysis in the economic evaluation



The additional QALYs gained from MEND
The additional QALYs gained by preventing obesity BMI 30-35 (1.388)
X
The decreased number of obese adults, BMI 30-35, in 2027
+
The decreased number of obese adults, BMI ≥35, in 2027
X
The additional QALYs gained by preventing obesity BMI ≥35 (2.3175)
% of obese adults w/ BMI 30-35 to the total obese adults (69%, Flegal et al. 2005)
X
The decreased number of obese adults in 2027 (119,862)
The decreased number of obese adults in 2027 (119,862)
X
% of obese adults w/ BMI ≥35 to the total obese adults (31%, Flegal et al. 2005)
Effectiveness of the MEND program (15.3%)
X
N of MEND eligible survivals in 2027 (1,317,922)
X
% of expected reduction of obese adults (59%, Freedman, 2001)
Table 4.2
Diff in expected life years after discounting
X
Quality of life
Table 4.6
Table 4.7
N of MEND eligible in 2010 (1,325,638)
X
Survival probability from age 10 to 27 (0.994)
Table 4.1
Annex 1
Diff in expected life years after discounting
X
Quality of life
Table 4.6
Table 4.7

List of abbreviations

Acronyms	Description
AUC	Area under the curve
BMI	Body mass index
CEA	Cost-effectiveness analysis
CI	Confidence interval
CPI	Consumer price index
CVD	Cardiovascular disease
DALYs	Disability-adjusted life years
EQ-5D	EuroQol 5 Dimensions
HRQL	Health-related quality of life
HSE0	Health Survey for England
HUI	Health Utility Index
ICER	Incremental cost-effectiveness ratio
IMD	Index of Multiple deprivation
IOTF	International Obesity Task Force
LAGB	Laparoscopic adjustable gastric banding
MI	Multiple imputation
MICE	Multiple imputation with chained equations
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NSPD	National Statistics Postcode Directory
PED	Price elasticity of demand
PIF	Potential impact fraction
QALYs	Quality-adjusted life years
QWB	Quality of Well-being
RCT	Randomized controlled trial
ROC	Receiver Operating Characteristics
RR	Relative risk
SE	Standard error
SES	Socio-economic status
SSB	Sugar-sweetened beverage
TSF	Tricep skinfolds
WHO	World Health Organization
YLD	Years of life lost
YLL	Years lived with disability

List of references

ALLISON, D. B., FONTAINE, K. R., MANSON, J. E., STEVENS, J. & VANITALLIE, T. B. (1999) Annual Deaths Attributable to Obesity in the United States. *JAMA,* 282**,** 1530-1538.

ANANTHAPAVAN, J., MOODIE, M., HABY, M. & CARTER, R. (2010) Assessing cost-effectiveness in obesity: laparoscopic adjustable gastric banding for severely obese adolescents. *Surgery for Obesity and Related Diseases,* 6**,** 377-385.

ANDREYEVA, T., LONG, M. W. & BROWNELL, K. D. (2010) The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. *Am J Public Health,* 100**,** 216-22.

BARLOW, S. & DIETZ, W. (1998) Obesity evaluation and treatment: expert committee recommendations. *Pediatrics,* 102**,** e29.

BELL, A., KREMER, P., MAGAREY, A. & SWINBURN, B. (2005a) Contribution of 'noncore'foods and beverages to the energy intake and weight status of Australian children. *European Journal of Clinical Nutrition,* 59**,** 639-645.

BELL, C. G., WALLEY, A. J. & FROGUEL, P. (2005b) The genetics of human obesity. *Nature Reviews Genetics,* 6**,** 221-234.

BRADDON, F. E., RODGERS, B., WADSWORTH, M. E. & DAVIES, J. M. (1986) Onset of obesity in a 36 year birth cohort study. *British medical journal,* 293**,** 299.

BROWN, H. S., PEREZ, A., LI, Y.-P., HOELSCHER, D., KELDER, S. & RIVERA, R. (2007) The cost-effectiveness of a school-based overweight program. *International Journal of Behavioral Nutrition and Physical Activity,* 4**,** 47.

BROWNELL, K. D., FARLEY, T., WILLETT, W. C., POPKIN, B. M., CHALOUPKA, F. J., THOMPSON, J. W. & LUDWIG, D. S. (2009) The Public Health and Economic Benefits of Taxing Sugar-Sweetened Beverages. *New England Journal of Medicine*.

BROWNELL, K. D. & FRIEDEN, T. R. (2009) Ounces of prevention--the public policy case for taxes on sugared beverages. *The New England Journal of Medicine,* 360**,** 1805.

BRYANT, M. J., LUCOVE, J. C., EVENSON, K. R. & MARSHALL, S. (2007) Measurement of television viewing in children and adolescents: a systematic review. *Obes Rev,* 8**,** 197-209.

BUREAU OF LABOR STATISTICS (2011) Consumer Price Index.

BUTLAND, B., JEBB, S., KOPELMAN, P., MCPHERSON, K., THOMAS, S., MARDELL, J. & PARRY, V. (2007) Tackling obesities: Future choices—Project report. *Government Office of the Chief Scientist,* http://www.foresight.gov.uk.

CAHNMAN, W. J. (1968) The stigma of obesity. *Sociological Quarterly***,** 283-299.

CALLE, E. E., RODRIGUEZ, C., WALKER-THURMOND, K. & THUN, M. J. (2003) Overweight, obesity, and mortality from cancer in a prospectively studied cohort of US adults.

CARTER, R., MOODIE, M., MARKWICK, A., MAGNUS, A., VOS, T., SWINBURN, B. & HABY, M. M. (2009) Assessing cost-effectiveness in obesity (ACE-obesity): an overview of the ACE approach, economic methods and cost results. *BMC Public Health,* 9**,** 419.

CARTER, R., VOS, T., MOODIE, M., HABY, M., MAGNUS, A. & MIHALOPOULOS, C. (2008) Priority setting in health: origins, description and application of the

Australian Assessing Cost-Effectiveness initiative. *Expert Rev Pharmacoecon Outcomes Res,* 8**,** 593-617.
CAWLEY, J. (2007) The cost-effectiveness of programs to prevent or reduce obesity: the state of the literature and a future research agenda. *Archives of Pediatrics & Adolescent Medicine,* 161**,** 611-4.
CAWLEY, J. (2010) The Economics Of Childhood Obesity. *Health Affairs,* 29**,** 364.
CENSUS DISSEMINATION UNIT (2011) GeoConvert.
CENTERS FOR DISEASE CONTROL AND PREVENTION (2009) About BMI for Children and Teens.
CHARLWOOD, R. & GIBBONS, K. (1986) The nutritional management of obese patients--a utilisation review. *Australian Clinical Review,* 6**,** 49-52.
CHOUINARD, H. H., DAVIS, D. E., LAFRANCE, J. T. & PERLOFF, J. M. (2007) Fat taxes: big money for small change. bepress.
CLAXTON, K., PAULDEN, M., GRAVELLE, H., BROUWER, W. & CULYER, A. J. (2011) Discounting and decision making in the economic evaluation of health-care technologies. *Health Economics,* 20**,** 2-15.
COLE, T. J., BELLIZZI, M. C., FLEGAL, K. M. & DIETZ, W. H. (2000) Establishing a standard definition for child overweight and obesity worldwide: international survey. *British medical journal,* 320**,** 1240.
COLEMAN, K., TILLER, C., SANCHEZ, J., HEATH, E., SY, O., MILLIKEN, G. & DZEWALTOWSKI, D. (2005) Prevention of the epidemic increase in child risk of overweight in low-income schools: the El Paso coordinated approach to child health. *Archives of pediatrics and adolescent medicine,* 159**,** 217.
COMMUNITIES AND LOCAL GOVERNMENT (2011a) The English Indices of Deprivation 2007.
COMMUNITIES AND LOCAL GOVERNMENT (2011b) Indices of Deprivation 2007.
COOK, R. D. & WEISBERG, S. (1983) Diagnostics for heteroscedasticity in regression. *Biometrika,* 70**,** 1.
COOKSON, R., DRUMMOND, M. & WEATHERLY, H. (2009) Explicit incorporation of equity considerations into economic evaluation of public health interventions. *Health Economics, Policy and Law,* 4**,** 231-245.
COUNTERWEIGHT PROJECT TEAM AND P. TRUEMAN, HAYNES, S. M., LYONS, G. F., MCCOMBIE, E. L., MCQUIGG, M. S. A., MONGIA, S., NOBLE, P. A., QUINN, M. F., ROSS, H. M., THOMPSON, F., BROOM, J. I., LAWS, R. A., RECKLESS, J. P. D., KUMAR, S., LEAN, M. E. J., FROST, G. S., FINER, N., HASLAM, D. W., MORRISON, D. & SLOAN, B. (2010) Long-term cost-effectiveness of weight management in primary care. *International Journal of Clinical Practice,* 64**,** 775-783.
DALZIEL, K. & SEGAL, L. (2006) Uncertainty in the Economic Analysis of School-Based Obesity Prevention Programs: Urgent Need for Quality Evaluation. *Obesity,* 14**,** 1481-1482.
DIETZ, W. H. (1994) Critical periods in childhood for the development of obesity. *Am J Clin Nutr,* 59**,** 955-9.
DRUMMOND, M. F., SCULPHER, M. J., TORRANCE, G. W., O'BRIEN, B. J. & STODDART, G. L. (2005) *Methods for the economic evaluation of health care programmes*, Oxford University Press, USA.
DUNN, O. J. (1961) Multiple comparisons among means. *Journal of the American Statistical Association,* 56**,** 52-64.
EBBELING, C. B., PAWLAK, D. B. & LUDWIG, D. S. (2002) Childhood obesity: public-health crisis, common sense cure. *The Lancet,* 360**,** 473-482.

EPSTEIN, L. H., VALOSKI, A. M., VARA, L. S., MCCURLEY, J., WISNIEWSKI, L., KALARCHIAN, M. A., KLEIN, K. R. & SHRAGER, L. R. (1995) Effects of decreasing sedentary behavior and increasing activity on weight change in obese children. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association,* 14**,** 109.

ERSELCAN, T., CANDAN, F., SARUHAN, S. & AYCA, T. (2000) Comparison of Body Composition Analysis Methods in Clinical Routine. *Annals of Nutrition & Metabolism,* 44**,** 243-248.

ESRC CENSUS PROGRAMME (2011) National Statistics Postcode Directory.

EZZATI, M., LOPEZ, A. D., RODGERS, A., VANDER HOORN, S. & MURRAY, C. J. L. (2002) Selected major risk factors and global and regional burden of disease. *The Lancet,* 360**,** 1347-1360.

FAROOQI, I. S., JEBB, S. A., LANGMACK, G., LAWRENCE, E., CHEETHAM, C. H., PRENTICE, A. M., HUGHES, I. A., MCCAMISH, M. A. & O'RAHILLY, S. (1999) Effects of recombinant leptin therapy in a child with congenital leptin deficiency.

FAROOQI, I. S. & O'RAHILLY, S. (2007) Genetic factors in human obesity. *Obes Rev,* 8 Suppl 1**,** 37-40.

FARRAR, D. & GLAUBER, R. (1967) Multicollinearity in Regression Analysis: The Problem Revisited. *The Review of Economics and Statistics,* 49**,** 92-107.

FINKELSTEIN, E., ZHEN, C., NONNEMAKER, J. & TODD, J. (2010) Impact of targeted beverage taxes on higher-and lower-income households. *Archives of Internal Medicine,* 170**,** 2028.

FINKELSTEIN, E. A. & BROWN, D. S. (2005) A cost-benefit simulation model of coverage for bariatric surgery among full-time employees. *American Journal of Managed Care,* 11**,** 641-6.

FINKELSTEIN, E. A., FIEBELKORN, I. C. & WANG, G. (2003) National medical spending attributable to overweight and obesity: how much, and who's paying. *Health Affairs,* 3**,** 219-226.

FINKELSTEIN, E. A., FIEBELKORN, I. C. & WANG, G. (2004) State-Level Estimates of Annual Medical Expenditures Attributable to Obesity*.

FINKELSTEIN, E. A., RUHM, C. J. & KOSA, K. M. (2005) Economic causes and consequences of obesity. *Annu Rev Public Health,* 26**,** 239-57.

FLEGAL, K. M., GRAUBARD, B. I., WILLIAMSON, D. F. & GAIL, M. H. (2005) Excess deaths associated with underweight, overweight, and obesity. *Jama,* 293**,** 1861-7.

FLETCHER, J. M., FRISVOLD, D. & TEFFT, N. (2010) CAN SOFT DRINK TAXES REDUCE POPULATION WEIGHT? *Contemporary Economic Policy,* 28**,** 23-35.

FORSHEE, R. A., ANDERSON, P. A. & STOREY, M. L. (2008) Sugar-sweetened beverages and body mass index in children and adolescents: a meta-analysis. *American Journal of Clinical Nutrition,* 87**,** 1662.

FREEDMAN, D. S., KHAN, L. K., DIETZ, W. H., SRINIVASAN, S. R. & BERENSON, G. S. (2001) Relationship of childhood obesity to coronary heart disease risk factors in adulthood: the Bogalusa Heart Study. *Pediatrics,* 108**,** 712.

FRENCH, S. A., STORY, M. & PERRY, C. L. (1995) Self-esteem and obesity in children and adolescents: a literature review. *Obesity research,* 3**,** 479-90.

FRIEDMAN, J. M. (2003) A war on obesity, not the obese.

FRIEDMAN, M. A. & BROWNELL, K. D. (1995) Psychological correlates of obesity: Moving to the next research generation. *Psychological Bulletin,* 117**,** 3.

GALLAGHER, D., VISSER, M., SEPULVEDA, D., PIERSON, R. N., HARRIS, T. & HEYMSFIELD, S. B. (1996) How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? *American Journal of Epidemiology,* 143**,** 228-239.
GANZ, M. L. (2003) The economic evaluation of obesity interventions: its time has come. *Obesity Research,* 11**,** 1275-7.
GARFINKEL, L. (1986) Overweight and mortality. *Cancer,* 58**,** 1826-1829.
GARROW, J. S. & SUMMERBELL, C. D. (1995) Meta-analysis: effect of exercise, with or without dieting, on the body composition of overweight subjects. *Eur J Clin Nutr,* 49**,** 1-10.
GOLAN, M. & CROW, S. (2004) Parents Are Key Players in the Prevention and Treatment of Weight-related Problems. *Nutrition reviews,* 62**,** 39-50.
GOLD, M. R., SIEGEL, J. E., RUSSELL, L. B. & WEINSTEIN, M. C. (1996) *Cost-effectiveness in health and medicine,* New York, Oxford University Press, USA.
GOLD, M. R., STEVENSON, D. & FRYBACK, D. G. (2002) HALYS AND QALYS AND DALYS, OH MY: Similarities and Differences in Summary Measures of Population Health. *Annual Review of Public Health,* 23**,** 115-134.
GOLDBLATT, P. B., MOORE, M. E. & STUNKARD, A. J. (1965) Social Factors in Obesity. *Jama,* 192**,** 1039-44.
GOLDFIELD, G., EPSTEIN, L., KILANOWSKI, C., PALUCH, R. & KOGUT-BOSSLER, B. (2001) Cost-effectiveness of group and mixed family-based treatment for childhood obesity. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity,* 25**,** 1843.
GOLDMAN, D., LAKDAWALLA, D. & ZHENG, Y. (2009) Food Prices and the Dynamics of Body Weight. *NBER Working Papers*.
GORDON-LARSEN, P., ADAIR, L. S. & POPKIN, B. M. (2002) Ethnic differences in physical activity and inactivity patterns and overweight status. *Obesity,* 10**,** 141-149.
GORN, G. J. & GOLDBERG, M. E. (1982) Behavioral evidence of the effects of televised food messages on children. *Journal of Consumer Research,* 9**,** 200-205.
GORSKY, R., PAMUK, E., WILLIAMSON, D., SHAFFER, P. & KOPLAN, J. (1996) The 25-year health care costs of women who remain overweight after 40 years of age. *American Journal of Preventive Medicine,* 12**,** 388.
GORTMAKER, S., PETERSON, K., WIECHA, J., SOBOL, A., DIXIT, S., FOX, M. & LAIRD, N. (1999a) Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. *Archives of pediatrics and adolescent medicine,* 153**,** 409.
GORTMAKER, S. L., MUST, A., PERRIN, J. M., SOBOL, A. M. & DIETZ, W. H. (1993) Social and economic consequences of overweight in adolescence and young adulthood. *N Engl J Med,* 329**,** 1008-12.
GORTMAKER, S. L., PETERSON, K., WIECHA, J., SOBOL, A. M., DIXIT, S., FOX, M. K. & LAIRD, N. (1999b) Reducing obesity via a school-based interdisciplinary intervention among youth Planet Health. Am Med Assoc.
GRIFFITHS, L. J., WOLKE, D., PAGE, A. S. & HORWOOD, J. P. (2006) Obesity and bullying: different effects for boys and girls. *Archives of Disease in Childhood,* 91**,** 121-125.
GROSSMAN, M. (1972) On the Concept of Health Capital and the Demand for Health. *The Journal of Political Economy,* 80**,** 223-255.

GUO, S., ROCHE, A., CHUMLEA, W., GARDNER, J. & SIERVOGEL, R. (1994) The predictive value of childhood body mass index values for overweight at age 35 y. *American Journal of Clinical Nutrition,* 59**,** 810.

GUO, S., WU, W., CHUMLEA, W. & ROCHE, A. (2002) Predicting overweight and obesity in adulthood from body mass index values in childhood and adolescence. *American Journal of Clinical Nutrition,* 76**,** 653.

HABY, M., VOS, T., CARTER, R., MOODIE, M., MARKWICK, A., MAGNUS, A., TAY-TEO, K. & SWINBURN, B. (2007) "A new approach to assessing the health benefit from obesity interventions in children and adolescents: The Assessing Cost-effectiveness in Obesity Project": Response to Segal and Dalziel. *International Journal of Obesity,* 31**,** pp.

HABY, M. M., VOS, T., CARTER, R., MOODIE, M., MARKWICK, A., MAGNUS, A., TAY-TEO, K. S. & SWINBURN, B. (2006) A new approach to assessing the health benefit from obesity interventions in children and adolescents: the assessing cost-effectiveness in obesity project. *Int J Obes,* 30**,** 1463-1475.

HAN, J. C., LAWLOR, D. A. & KIMM, S. Y. S. (2010) Childhood obesity. *The Lancet,* 375**,** 1737-1748.

HARRIS, M. B. (1990) Is Love Seen as Different for the Obese?1. *Journal of Applied Social Psychology,* 20**,** 1209-1224.

HAYDEN-WADE, H. A., STEIN, R. I., GHADERI, A., SAELENS, B. E., ZABINSKI, M. F. & WILFLEY, D. E. (2005) Prevalence, Characteristics, and Correlates of Teasing Experiences among Overweight Children vs. Non-overweight Peers[ast][ast]. *Obesity,* 13**,** 1381-1392.

HEATH, E. & COLEMAN, K. (2002) Evaluation of the institutionalization of the coordinated approach to child health (CATCH) in a US/Mexico border community. *Health Education & Behavior,* 29**,** 444.

HEATH, E. & COLEMAN, K. (2003) Adoption and institutionalization of the Child and Adolescent Trial for Cardiovascular Health (CATCH) in El Paso, Texas. *Health Promotion Practice,* 4**,** 157.

HECKMAN, J. (2010) The Heckman Equation: Invest in early childhood development:Reduce deficits, strengthen the economy.

HEYMSFIELD, S. B., NUNEZ, C. & PIETROBELLI, A. (1997) Bioimpedance Analysis: What Are the Next Steps? *Nutrition in Clinical Practice,* 12**,** 201.

HOUSE OF COMMONS HEALTH COMMITTEE (2004) Obesity: Third Report of Session 2003-4.

HU, F. B. (2007) Obesity and mortality: watch your waist, not just your weight. *Archives of Internal Medicine,* 167**,** 875.

ICHIHARA, S. & YAMADA, Y. (2008) Genetic factors for human obesity. *Cellular and Molecular Life Sciences (CMLS),* 65**,** 1086-1098.

JANSSEN, IAN, CRAIG, WENDY, M., BOYCE, WILLIAM, F., PICKETT & WILLIAM (2004) *Associations between overweight and obesity with bullying Behaviors in School-Aged Children,* Elk Grove Village, IL, ETATS-UNIS, American Academy of Pediatrics.

JEBB, S. & PRENTICE, A. (2001) Single definition of overweight and obesity should be used. *BMJ: British Medical Journal,* 323**,** 999.

JOHN, J., WENIG, C. M. & WOLFENSTETTER, S. B. (2010) Recent economic findings on childhood obesity: cost-of-illness and cost-effectiveness of interventions. *Current Opinion in Clinical Nutrition & Metabolic Care,* 13**,** 305-313 10.1097/MCO.0b013e328337fe18.

KALAVAINEN, M., KARJALAINEN, S., MARTIKAINEN, J., KORPPI, M., LINNOSMAA, I. & NUUTINEN, O. (2009) Cost-effectiveness of routine and group programs for treatment of obese children. *Pediatrics International,* 51**,** 606-11.
KEELER, E. B. & CRETIN, S. (1983) Discounting of Life-Saving and Other Nonmonetary Effects. *Management Science,* 29**,** 300-306.
KLESGES, R. C., MEYERS, A. W., KLESGES, L. M. & LA VASQUE, M. E. (1989) Smoking, body weight, and their effects on smoking behavior: a comprehensive review of the literature. *Psychol Bull,* 106**,** 204-30.
KOLOTOUROU, M., CHADWICK, P., COLE, T. J., LAWSON, M. S., SINGH, A. S. & SACHER, P. M. (2009) The MEND Programme: National effectiveness data. *Obesity Facts,* 2**,** 27-8.
KOPEC, J. A. & WILLISON, K. D. (2003) A comparative review of four preference-weighted measures of health-related quality of life. *Journal of Clinical Epidemiology,* 56**,** 317-325.
KOPELMAN, P. G. (2000) Obesity as a medical problem. *Nature,* 404**,** 635-43.
KRUGMAN, P. & WELLS, R. (2009) *Microeconomics,* New York, Worth Publishers.
LAGARDE, M., HAINES, A. & PALMER, N. (2007) Conditional cash transfers for improving uptake of health interventions in low- and middle-income countries: a systematic review. *Jama,* 298**,** 1900-10.
LAGARDE, M., HAINES, A. & PALMER, N. (2009) The impact of conditional cash transfers on health outcomes and use of health services in low and middle income countries. *Cochrane Database of Systematic Reviews.* Chichester, UK, John Wiley & Sons, Ltd.
LAITINEN, J., POWER, C. & JARVELIN, M. R. (2001) Family social class, maternal body mass index, childhood body mass index, and age at menarche as predictors of adult obesity. *American Journal of Clinical Nutrition,* 74**,** 287.
LAKDAWALLA, D. & PHILIPSON, T. (2002) The growth of obesity and technological change: a theoretical and empirical examination. *NBER Working Paper*.
LAKDAWALLA, D. & PHILIPSON, T. (2009) The growth of obesity and technological change. *Economics & Human Biology,* 7**,** 283-293.
LANTZ, P. M., JACOBSON, P. D., WARNER, K. E., WASSERMAN, J., POLLACK, H. A., BERSON, J. & AHLSTROM, A. (2000) Investing in youth tobacco control: a review of smoking prevention and control strategies. *Tobacco control,* 9**,** 47.
LEICESTER, A. & WINDMEIJER, F. (2004) The'fat tax': economic incentives to reduce obesity. *Open Access publications from University College London*.
LOBSTEIN, T., BAUR, L. & UAUY, R. (2004) Obesity in children and young people: a crisis in public health. *Obes Rev,* 5**,** 4-104.
LOBSTEIN, T. & LEACH, R. J. (2007) Tackling Obesities: Future Choices - International Comparisons of Obesity Trends, Determinants and Responses - Evidence Review. Department of Innovation Universities and Skills.
MAGNUS, A., HABY, M. M., CARTER, R. & SWINBURN, B. (2009) The cost-effectiveness of removing television advertising of high-fat and/or high-sugar food and beverages to Australian children. *International Journal of Obesity,* Nature Publishing Group**,** 10.
MATTHEWS, K. A. (2002) Body mass index in mid-life women: erlative influence of menopause, hormone use, and ethnicity. *International Journal of Obesity,* 26**,** 1150.
MCELROY, S. L., KOTWAL, R., MALHOTRA, S., NELSON, E. B., KECK, P. E. & NEMEROFF, C. B. (2004) Are mood disorders and obesity related? A review for the mental health professional. *J Clin Psychiatry,* 65**,** 634-51, quiz 730.

MCPHERSON, K., MARSH, T. & BROWN, M. (2007) Tackling Obesities: Future Choices–Modelling Future Trends in Obesity and the Impact on Health. *Foresight*.

MEND CENTRAL LTD. (2010) The MEND 7 - 13 Programme.

MOKDAD, A. H., MARKS, J. S., STROUP, D. F. & GERBERDING, J. L. (2004) Actual causes of death in the United States, 2000. Am Med Assoc.

MOODIE, M., HABY, M., GALVIN, L., SWINBURN, B. & CARTER, R. (2009) Cost-effectiveness of active transport for primary school children - Walking School Bus program. *Int J Behav Nutr Phys Act,* 6**,** 63.

MOODIE, M., HABY, M., WAKE, M., GOLD, L. & CARTER, R. (2008a) Cost-effectiveness of a family-based GP-mediated intervention targeting overweight and moderately obese children. *Economics & Human Biology,* 6**,** 363-76.

MOODIE, M., HABY, M., WAKE, M., GOLD, L. & CARTER, R. (2008b) Symposium on the Economics of Obesity: Cost-Effectiveness of a Family-Based GP-Mediated Intervention Targeting Overweight and Moderately Obese Children. *Economics and Human Biology,* 6**,** 363-76.

MOODIE, M. L., CARTER, R. C., SWINBURN, B. A. & HABY, M. M. (2010) The Cost-effectiveness of Australia's Active After-school Communities Program. *Obesity,* 18**,** 1585-1592.

MOON, G., QUARENDON, G., BARNARD, S., TWIGG, L. & BLYTH, B. (2007) Fat nation: deciphering the distinctive geographies of obesity in England. *Soc Sci Med,* 65**,** 20-31.

MOONEY, G. (2007) Economic evaluation of prevention: we need to do better but first we need to sort out what the good is. *International Journal of Public Health,* 52**,** 339-40.

MURRAY, C. J. L., LOPEZ, A. D. & MUNDIAL, B. (1996) *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*, Harvard University Press Cambridge, MA.

MUTCH, D. M. & CLEMENT, K. (2006) Unraveling the genetics of human obesity. *PLoS Genet,* 2**,** e188.

NATIONAL AUDIT OFFICE (2001) Tackling obesity in England.

NATIONAL HEART LUNG AND BLOOD INSTITUTE (1998) Clinical guidelines on the identification, evaluation and treatment of overweight and obesity in adults.

NATIONAL INSTITUTE FOR CLINICAL EXCELLENCE (2001) Guidance for Manufacturers and Sponsors.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE (2006) Obesity: guidance on the prevention, identification, assessment and management of overweight and obesity in adults and children.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE (2008) Guide to the methods of technology appraisal.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE (2009) Public health guidance process and method guides: Methods for development of NICE public health guidance 2009.

NEEL, J. V. (1962) Diabetes mellitus: a thrifty" genotype rendered detrimental by progress"? *Am J Hum Genet,* 14**,** 353-362.

NEOVIUS, M. G., LINNÉ, Y. M., BARKELING, B. S. & ROSSNER, S. O. (2004) Sensitivity and specificity of classification systems for fatness in adolescents. *The American journal of clinical nutrition,* 80**,** 597.

NEUMARK-SZTAINER, D., FALKNER, N., STORY, M., PERRY, C., HANNAN, P. & MULERT, S. (2002) Weight-teasing among adolescents: Correlations with weight status and disordered eating behaviors. *International Journal of Obesity*.
NEUMARK-SZTAINER, D., STORY, M. & FAIBISCH ED.D, L. (1998) Perceived stigmatization among overweight African-American and caucasian adolescent girls. *Journal of Adolescent Health,* 23**,** 264-270.
NHS (2010) Health Survey for England - 2008 trend tables.
NICE (2006) Obesity: the prevention, identification, assessment and management of overweight and obesity in adults and children.
OFFICE FOR NATIONAL STATISTICS (2010) Interim Life Tables.
OFFICE FOR NATIONAL STATISTICS (2011a) Ethnicity & Identity.
OFFICE FOR NATIONAL STATISTICS (2011b) Internet Access.
OGDEN, C. L., YANOVSKI, S. Z., CARROLL, M. D. & FLEGAL, K. M. (2007) The epidemiology of obesity. *Gastroenterology,* 132**,** 2087-102.
OSTER, G., THOMPSON, D., EDELSBERG, J., BIRD, A. & COLDITZ, G. (1999) Lifetime health and economic benefits of weight loss among obese persons. *American Journal of Public Health,* 89**,** 1536.
OUDE LUTTIKHUIS, H., BAUR, L., JANSEN, H., SHREWSBURY, V. A., O'MALLEY, C., STOLK, R. P. & SUMMERBELL, C. D. (2009) Interventions for treating obesity in children. *Cochrane Database Syst Rev***,** CD001872.
PARLESAK, A. & KROMKER, D. (2008) Obesity - a social and physical risk. *J Dtsch Dermatol Ges,* 6**,** 442-9.
PEARCE, M. J., BOERGERS, J. & PRINSTEIN, M. J. (2002) Adolescent Obesity, Overt and Relational Peer Victimization, and Romantic Relationships. *Obesity,* 10**,** 386-393.
PEETERS, A., BARENDREGT, J. J., WILLEKENS, F., MACKENBACH, J. P., AL MAMUN, A. & BONNEUX, L. (2003) Obesity in adulthood and its consequences for life expectancy: a life-table analysis. *Ann Intern Med,* 138**,** 24-32.
PHILIPSON, T. (2001) The world-wide growth in obesity: an economic research agenda. *Health Economics,* 10**,** 1-7.
PIERCE, J. W. & WARDLE, J. (1997) Cause and Effect Beliefs and Self-esteem of Overweight Children. *Journal of Child Psychology and Psychiatry,* 38**,** 645-650.
PINE, D. S., GOLDSTEIN, R. B., WOLK, S. & WEISSMAN, M. M. (2001) The Association Between Childhood Depression and Adulthood Body Mass Index. *Pediatrics,* 107**,** 1049-1056.
POWELL, L. M. & BAO, Y. (2009) Food prices, access to food outlets and child weight. *Economics & Human Biology,* 7**,** 64-72.
POWELL, L. M. & CHALOUPKA, F. J. (2009) Food Prices and Obesity: Evidence and Policy Implications for Taxes and Subsidies. *The Milbank Quarterly,* 87**,** 229-257.
POWER, C., LAKE, J. & COLE, T. (1997) Measurement and long-term health risks of child and adolescent fatness. *International Journal of Obesity,* 21**,** 507-526.
PUHL, R. & LATNER, J. (2007) Stigma, obesity, and the health of the nation's children. *Psychological bulletin,* 133**,** 557.
REGAN, P. C. (1996) Sexual Outcasts: The Perceived Impact of Body Weight and Gender on Sexuality1. *Journal of Applied Social Psychology,* 26**,** 1803-1815.
REILLY, J. (2005) Descriptive epidemiology and health consequences of childhood obesity. *Best Practice & Research Clinical Endocrinology & Metabolism,* 19**,** 327-341.

REILLY, J., WILSON, M., SUMMERBELL, C. & WILSON, D. (2002) Obesity: diagnosis, prevention, and treatment; evidence based answers to common questions. *Archives of Disease in Childhood,* 86**,** 392.
REILLY, J. J., METHVEN, E., MCDOWELL, Z. C., HACKING, B., ALEXANDER, D., STEWART, L. & KELNAR, C. J. H. (2003) Health consequences of obesity. *British Medical Journal,* 88**,** 748.
RICCIARDELLI, L. A. & MCCABE, M. P. (2001) Children's body image concerns and eating disturbance: A review of the literature. *Clinical Psychology Review,* 21**,** 325-344.
ROBINSON, T. N. (1999a) Behavioural treatment of childhood and adolescent obesity. *International Journal of Obesity,* 23**,** S52-S57.
ROBINSON, T. N. (1999b) Reducing children's television viewing to prevent obesity: a randomized controlled trial. *Jama,* 282**,** 1561-7.
ROBINSON, T. N. (2001) Television viewing and childhood obesity. *Pediatr Clin North Am,* 48**,** 1017-25.
ROUBENOFF, R. & KEHAYIAS, J. J. (1991) The meaning and measurement of lean body mass. *Nutrition reviews,* 49**,** 163-175.
RUBIN, D. B. (1976) Inference and Missing Data. *Biometrika,* 63**,** 581-592.
SACH, T. H., BARTON, G. R., DOHERTY, M., MUIR, K. R., JENKINSON, C. & AVERY, A. J. (2006) The relationship between body mass index and health-related quality of life: comparing the EQ-5D, EuroQol VAS and SF-6D. *International Journal of Obesity,* 31**,** 189-196.
SACHER, P. M., CHADWICK, P., WELLS, J. C. K., WILLIAMS, J. E., COLE, T. J. & LAWSON, M. S. (2005) Assessing the acceptability and feasibility of the MEND Programme in a small group of obese 7-11-year-old children. *Journal of Human Nutrition & Dietetics,* 18**,** 3-5.
SACHER, P. M., KOLOTOUROU, M., CHADWICK, P. M., COLE, T. J., LAWSON, M. S., LUCAS, A. & SINGHAL, A. (2010) Randomized Controlled Trial of the MEND Program: A Family-based Community Intervention for Childhood Obesity. *Obesity,* 18**,** S62-S68.
SCHWARTZ, T. L., NIHALANI, N., JINDAL, S., VIRK, S. & JONES, N. (2004) Psychiatric medication-induced obesity: a review. *Obes Rev,* 5**,** 115-21.
SCOTTISH INTERCOLLEGIATE GUIDELINES NETWORK (2010) Management of Obesity. A national clinical guideline. Scottish Intercollegiate Guidelines Network.
SEGAL, L. & DALZIEL, K. (2007) Economic evaluation of obesity interventions in children and adults. *Int J Obes,* 31**,** 1183-1184.
SEIDELL, J. C. & FLEGAL, K. M. (1997) Assessing obesity: classification and epidemiology. *Br Med Bull,* 53**,** 238-52.
SEIDELL, J. C., NOOYENS, A. J. & VISSCHER, T. L. S. (2005) Cost-effective measures to prevent obesity: epidemiological basis and appropriate target groups. *Proceedings of the Nutrition Society,* 64**,** 1-5.
SERDULA, M. K., IVERY, D., COATES, R. J., FREEDMAN, D. S., WILLIAMSON, D. F. & BYERS, T. (1993) Do Obese Children Become Obese Adults? A Review of the Literature. *Preventive Medicine,* 22**,** 167-177.
SHERWOOD, N. E., JEFFERY, R. W., FRENCH, S. A., HANNAN, P. J. & MURRAY, D. M. (2000) Predictors of weight gain in the Pound of Prevention study. *International journal of obesity,* 24**,** 395-403.
SINGH, A. S., MULDER, C., TWISK, J. W., VAN MECHELEN, W. & CHINAPAW, M. J. (2008) Tracking of childhood overweight into adulthood: a systematic review of the literature. *Obes Rev,* 9**,** 474-488.

STAMATAKIS, E., PRIMATESTA, P., CHINN, S., RONA, R. & FALASCHETI, E. (2005) Overweight and obesity trends from 1974 to 2003 in English children: What is the role of socio-economic factors? *British medical journal*.
STAMATAKIS, E., WARDLE, J. & COLE, T. J. (2009) Childhood obesity and overweight prevalence trends in England: evidence for growing socioeconomic disparities. *Int J Obes,* 34**,** 41-47.
STATA LIBRARY (2011) Multiple Imputation Using ICE. UCLA Academic Technology Services, Statistical Consulting Group.
STATACORP, L. E. (2009) *Stata Multiple-Imputation Reference Manual*, Stata Press.
STATACORP, L. E. (2011) Stata 11.
STUNKARD, A. J., FOCH, T. T. & HRUBEC, Z. (1986) A Twin Study of Human Obesity. *JAMA,* 256**,** 51-54.
STURM, R. & DATAR, A. (2005) Body mass index in elementary school children, metropolitan area food prices and food outlet density. *Public Health,* 119**,** 1059-1068.
STURM, R. & DATAR, A. (2008) Food prices and weight gain during elementary school: 5-year update. *Public Health,* 122**,** 1140-1143.
STURM, R. & DATAR, A. (2011) Regional price differences and food consumption frequency among elementary school children. *Public Health,* 125**,** 136-141.
SUMMERBELL, C. D., ASHTON, V., CAMPBELL, K. J., EDMUNDS, L., KELLY, S. & WATERS, E. (2003) Interventions for treating obesity in children. *Cochrane database of systematic reviews (Online)*.
SUMMERBELL, C. D., WATERS, E., EDMUNDS, L. D., KELLY, S., BROWN, T. & CAMPBELL, K. J. (2005) Interventions for preventing obesity in children. *Cochrane Database Syst Rev***,** CD001871.
THE COMPTROLLER AND AUDITOR GENERAL (2001) Tackling Obesity in England. *HC 220 Session 2000-2001.* National Audit Office.
THOMPSON, D., EDELSBERG, J., COLDITZ, G. A., BIRD, A. P. & OSTER, G. (1999) Lifetime Health and Economic Consequences of Obesity. *Arch Intern Med,* 159**,** 2177-2183.
THOMPSON, D. & WOLF, A. M. (2001) The medical-care cost burden of obesity. *Obesity Reviews,* 2**,** 189-197.
THORPE, K. E., FLORENCE, C. S., HOWARD, D. H. & JOSKI, P. (2005) The rising prevalence of treated disease: effects on private health insurance spending. *Health Affairs,* 24.
U.S. CENSUS BUREAU (2010) State & County QuickFacts.
USDA ECONOMIC RESEARCH SERVICE (2007) Food CPI and Expenditures: CPI Bias.
VARTANIAN, L. R., SCHWARTZ, M. B. & BROWNELL, K. D. (2007) Effects of soft drink consumption on nutrition and health: a systematic review and meta-analysis. *American Journal of Public Health,* 97**,** 667.
VIRK, S., SCHWARTZ, T. L., JINDAL, S., NIHALANI, N. & JONES, N. (2004) Psychiatric medication induced obesity: an aetiologic review. *Obesity reviews,* 5**,** 167-170.
VISCUSI, W. K. (1995) Discounting health effects for medical decisions. *Valuing health care***,** 125-147.
VISSCHER, T. L. & SEIDELL, J. C. (2001) The public health impact of obesity. *Annu Rev Public Health,* 22**,** 355-75.
VOTRUBA, S. B., HORVITZ, M. A. & SCHOELLER, D. A. (2000) The role of exercise in the treatment of obesity. *Nutrition,* 16**,** 179-188.

WAKE, M., GOLD, L., MCCALLUM, Z., GERNER, B. & WATERS, E. (2008) Economic evaluation of a primary care trial to reduce weight gain in overweight/obese children: the LEAP trial. *Ambulatory Pediatrics,* 8**,** 336-41.

WANG, L., CHYEN, D., LEE, S. & LOWRY, R. (2008) The association between body mass index in adolescence and obesity in adulthood. *Journal of Adolescent Health,* 42**,** 512-518.

WANG, L. Y., DENNISTON, M., LEE, S., GALUSKA, D. & LOWRY, R. (2010) Long-term Health and Economic Impact of Preventing and Reducing Overweight and Obesity in Adolescence. *Journal of Adolescent Health,* In Press, Corrected Proof.

WANG, L. Y., YANG, Q., LOWRY, R. & WECHSLER, H. (2003) Economic analysis of a school-based obesity prevention program. *Obes Res,* 11**,** 1313-24.

WANG, L. Y., YANG, Q., LOWRY, R. & WECHSLER, H. (2006) Counterpoint: uncertainty in the economic analysis of school-based obesity prevention programs: urgent need for quality evaluation. *Obesity,* 14**,** 1483-4.

WANG, Y. & BEYDOUN, M. A. (2007) The Obesity Epidemic in the United States--Gender, Age, Socioeconomic, Racial/Ethnic, and Geographic Characteristics: A Systematic Review and Meta-Regression Analysis. *Epidemiol Rev***,** mxm007.

WANG, Y. & LOBSTEIN, T. (2006) Worldwide trends in childhood overweight and obesity. *International Journal of Pediatric Obesity,* 1**,** 11-25.

WANG, Y., MONTEIRO, C. & POPKIN, B. M. (2002) Trends of obesity and underweight in older children and adolescents in the United States, Brazil, China, and Russia. *The American journal of clinical nutrition,* 75**,** 971.

WARDLE, J. & COOKE, L. (2005) The impact of obesity on psychological well-being. *Best Practice & Research Clinical Endocrinology & Metabolism,* 19**,** 421-440.

WEINSTEIN, M. C. & STASON, W. B. (1977) Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med,* 296**,** 716-721.

WEISS, R., DZIURA, J., BURGERT, T. S., TAMBORLANE, W. V., TAKSALI, S. E., YECKEL, C. W., ALLEN, K., LOPES, M., SAVOYE, M. & MORRISON, J. (2004) Obesity and the metabolic syndrome in children and adolescents. *The New England journal of medicine,* 350**,** 2362.

WHITAKER, R. C., WRIGHT, J. A., PEPE, M. S., SEIDEL, K. D. & DIETZ, W. H. (1997) Predicting Obesity in Young Adulthood from Childhood and Parental Obesity. *N Engl J Med,* 337**,** 869-873.

WILLIAMS, A. H. & COOKSON, R. A. (2006) *Equity-efficiency trade-offs in health technology assessment,* Cambridge, ROYAUME-UNI, Cambridge University Press.

WILLOUGHBY, T. (2008) A short-term longitudinal study of Internet and computer game use by adolescent boys and girls: Prevalence, frequency of use, and psychosocial predictors. *Developmental Psychology,* 44**,** 195.

WITHROW, D. & ALTER, D. A. (2010) The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. *Obesity Reviews*.

WOHLAND, P., REES, P., NORMAN, P., BODEN, P. & JASINSKA, M. (2010) *Ethnic Population Projections for the UK and Local Areas, 2001-2051*, School of Geography, University of Leeds.

WOLF, A. M. & COLDITZ, G. A. (1994) The cost of obesity: the US perspective. *Pharmacoeconomics,* 5**,** 34.

WOLF, A. M. & COLDITZ, G. A. (1998) Current estimates of the economic cost of obesity in the United States. *Obesity Research,* 6**,** 97.

WORLD HEALTH ORGANIZATION (2000) Obesity: Preventing and Managing the Global Epidemic. Geneva, Switzerland, World Health Organization.

WORLD HEALTH ORGANIZATION (2010) United Kingdom.
WORLD HEALTH ORGANIZATION (2011a) Global Database on Body Mass Index: an interactive surveillance tool for monitoring nutrition transition
WORLD HEALTH ORGANIZATION (2011b) Obesity.
WYATT, S. B., WINTERS, K. P. & DUBBERT, P. M. (2006) Overweight and obesity: prevalence, consequences, and causes of a growing public health problem. *Am J Med Sci,* 331**,** 166-74.
YANG, W., KELLY, T. & HE, J. (2007) Genetic epidemiology of obesity. *Epidemiologic reviews,* 29**,** 49.
YANIV, G., ROSIN, O. & TOBOL, Y. (2009) Junk-food, home cooking, physical activity and obesity: The effect of the fat tax and the thin subsidy. *Journal of Public Economics,* 93**,** 823-830.
ZANINOTTO, P., WARDLE, H., STAMATAKIS, E., MINDELL, J. & HEAD, J. (2006) *Forecasting obesity to 2010*, Department of Health.