# Cluster randomised controlled trial and economic and process evaluation to determine the effectiveness and cost-effectiveness of a novel intervention [Healthy Lifestyles Programme (HeLP)] to prevent obesity in school children

Katrina Wyatt, Jenny Lloyd, Siobhan Creanor, Colin Green,
Sarah G Dean, Melvyn Hillsdon, Charles Abraham,
Richard Tomlinson, Virginia Pearson, Rod S Taylor, Emma Ryan,
Adam Streeter, Camilla McHugh, Alison Hurst, Lisa Price,
Louise Crathorne, Chris Krägeloh, Richard Siegert and Stuart Logan

**NHS**
*National Institute for Health Research*

# Cluster randomised controlled trial and economic and process evaluation to determine the effectiveness and cost-effectiveness of a novel intervention [Healthy Lifestyles Programme (HeLP)] to prevent obesity in school children

Katrina Wyatt,[1]* Jenny Lloyd,[1] Siobhan Creanor,[2]
Colin Green,[1] Sarah G Dean,[1] Melvyn Hillsdon,[3]
Charles Abraham,[1] Richard Tomlinson,[4]
Virginia Pearson,[5] Rod S Taylor,[1] Emma Ryan,[6]
Adam Streeter,[2] Camilla McHugh,[1] Alison Hurst,[1]
Lisa Price,[3] Louise Crathorne,[1] Chris Krägeloh,[7]
Richard Siegert[7] and Stuart Logan[1]

[1]Institute for Health Research, University of Exeter Medical School,
 South Cloisters, St Luke's Campus, Exeter, UK
[2]Peninsula Clinical Trials Unit and Medical Statistics, Plymouth University
 Peninsula Schools of Medicine & Dentistry, Plymouth, UK
[3]Sport and Health Sciences, College of Life and Environmental Sciences,
 University of Exeter, Exeter, UK
[4]Royal Devon & Exeter NHS Foundation Trust, Exeter, UK
[5]Devon County Council, Public Health Directorate, Exeter, UK
[6]Isca Academy, Exeter, UK
[7]Health Sciences, Auckland University of Technology, Auckland, New Zealand

*Corresponding author

# Public Health Research

### Criteria for inclusion in the *Public Health Research* journal
Reports are published in *Public Health Research* (PHR) if (1) they have resulted from work for the PHR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Public Health Research* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## PHR programme
The Public Health Research (PHR) programme, part of the National Institute for Health Research (NIHR), evaluates public health interventions, providing new knowledge on the benefits, costs, acceptability and wider impacts of non-NHS interventions intended to improve the health of the public and reduce inequalities in health. The scope of the programme is multi-disciplinary and broad, covering a range of interventions that improve public health. The Public Health Research programme also complements the NIHR Health Technology Assessment programme which has a growing portfolio evaluating NHS public health interventions.

For more information about the PHR programme please visit the website: http://www.nets.nihr.ac.uk/programmes/phr

## This report

# Abstract

## Cluster randomised controlled trial and economic and process evaluation to determine the effectiveness and cost-effectiveness of a novel intervention [Healthy Lifestyles Programme (HeLP)] to prevent obesity in school children

Katrina Wyatt,[1]* Jenny Lloyd,[1] Siobhan Creanor,[2] Colin Green,[1] Sarah G Dean,[1] Melvyn Hillsdon,[3] Charles Abraham,[1] Richard Tomlinson,[4] Virginia Pearson,[5] Rod S Taylor,[1] Emma Ryan,[6] Adam Streeter,[2] Camilla McHugh,[1] Alison Hurst,[1] Lisa Price,[3] Louise Crathorne,[1] Chris Krägeloh,[7] Richard Siegert[7] and Stuart Logan[1]

[1]Institute for Health Research, University of Exeter Medical School, South Cloisters, St Luke's Campus, Exeter, UK
[2]Peninsula Clinical Trials Unit and Medical Statistics, Plymouth University Peninsula Schools of Medicine & Dentistry, Plymouth, UK
[3]Sport and Health Sciences, College of Life and Environmental Sciences, University of Exeter, Exeter, UK
[4]Royal Devon & Exeter NHS Foundation Trust, Exeter, UK
[5]Devon County Council, Public Health Directorate, Exeter, UK
[6]Isca Academy, Exeter, UK
[7]Health Sciences, Auckland University of Technology, Auckland, New Zealand

*Corresponding author  k.m.wyatt@exeter.ac.uk

**Background:** Approximately one-third of children in England leave primary school overweight or obese. There is little evidence of effective obesity prevention programmes for children in this age group.

**Objective:** To determine the effectiveness and cost-effectiveness of a school-based healthy lifestyles programme in preventing obesity in children aged 9–10 years.

**Design:** A cluster randomised controlled trial with an economic and process evaluation.

**Setting:** Thirty-two primary schools in south-west England.

**Participants:** Children in Year 5 (aged 9–10 years) at recruitment and in Year 7 (aged 11–12 years) at 24 months' post-baseline follow-up.

**Intervention:** The Healthy Lifestyles Programme (HeLP) ran during the spring and summer terms of Year 5 into the autumn term of Year 6 and included four phases: (1) building a receptive environment, (2) a drama-based healthy lifestyles week, (3) one-to-one goal setting and (4) reinforcement activities.

**Main outcome measures:** The primary outcome measure was body mass index (BMI) standard deviation score (SDS) at 24 months post baseline measures (12 months post intervention). The secondary outcomes comprised waist circumference SDS, percentage body fat SDS, proportion of children overweight and obese at 18 and 24 months, accelerometer-assessed physical activity and food intake at 18 months, and cost-effectiveness.

**Results:** We recruited 32 schools and 1324 children. We had a rate of 94% follow-up for the primary outcome. No difference in BMI SDS was found at 24 months [mean difference –0.02, 95% confidence interval (CI) –0.09 to 0.05] or at 18 months (mean difference –0.02, 95% CI –0.08 to 0.05) between children in the intervention schools and children in the control schools. No difference was found between the intervention and control groups in waist circumference SDS, percentage body fat SDS or physical activity levels. Self-reported dietary behaviours showed that, at 18 months, children in the intervention schools consumed fewer energy-dense snacks and had fewer negative food markers than children in the control schools. The intervention effect on negative food markers was fully mediated by 'knowledge' and three composite variables: 'confidence and motivation', 'family approval/behaviours and child attitudes' and 'behaviours and strategies'. The intervention effect on energy-dense snacks was partially mediated by 'knowledge' and the same composite variables apart from 'behaviours and strategies'. The cost of implementing the intervention was approximately £210 per child. The intervention was not cost-effective compared with control. The programme was delivered with high fidelity, and it engaged children, schools and families across the socioeconomic spectrum.

**Limitations:** The rate of response to the parent questionnaire in the process evaluation was low. Although the schools in the HeLP study included a range of levels of socioeconomic deprivation, class sizes and rural and urban settings, the number of children for whom English was an additional language was considerably lower than the national average.

**Conclusions:** HeLP is not effective or cost-effective in preventing overweight or obesity in children aged 9–10 years.

**Future work:** Our very high levels of follow-up and fidelity of intervention delivery lead us to conclude that it is unlikely that school-based programmes targeting a single age group can ever be sufficiently intense to affect weight status. New approaches are needed that affect the school, the family and the wider environment to prevent childhood obesity.

**Trial registration:** Current Controlled Trials ISRCTN15811706.

# Contents

## CONTENTS

# List of tables

# List of figures

# List of boxes

# List of abbreviations

| | | | |
|---|---|---|---|
| BMI | body mass index | NCMP | National Child Measurement Programme |
| CACE | complier average causal effect | NICE | National Institute for Health and Care Excellence |
| CFA | confirmatory factor analysis | | |
| CFI | comparative fit index | NIHR | National Institute for Health Research |
| CHD | coronary heart disease | | |
| CI | confidence interval | OR | odds ratio |
| CONSORT | Consolidated Standards of Reporting Trials | PAG | Project Advisory Group |
| | | PE | physical education |
| CRC | colorectal cancer | QALY | quality-adjusted life-year |
| df | degrees of freedom | RCT | randomised controlled trial |
| EFA | exploratory factor analysis | RMSEA | root-mean-square error of approximation |
| EQ-5D | EuroQol-5 Dimensions | | |
| FIQ | Food Intake Questionnaire | SAT | Standard Assessment Test |
| HeLP | Healthy Lifestyles Programme | SD | standard deviation |
| ICC | intraclass correlation coefficient | SDS | standard deviation score |
| IMB | information, motivation and behavioural skills | SEM | structural equation modelling |
| | | SRMR | standardised root-mean-square residual |
| IMD | Index of Multiple Deprivation | | |
| m$g$ | milligravity | T2DM | type 2 diabetes mellitus |
| MLQ | My Lifestyle Questionnaire | TSC | Trial Steering Committee |
| MRC | Medical Research Council | | |

# Plain English summary

Childhood obesity is an important health issue. Working with teachers, families and children, we developed the Healthy Lifestyles Programme (HeLP), which aims to engage and support children and families to make healthy food and activity choices. We designed a study to understand whether or not HeLP can prevent children aged 9–10 years from becoming overweight or obese. The study involved 32 primary schools from Devon, half of which were randomly selected to receive the programme while the other half continued as usual. We measured children's weight and height, waist circumference and percentage body fat; assessed how active they were; and used questionnaires to look at the types of food and drink they consumed. We also asked what they understood about a healthy lifestyle and how they felt about it.

The study began when the children were 9–10 years old, in Year 5, and HeLP was delivered in the spring and summer terms of Year 5 and in the autumn term of Year 6. Children had their final set of measurements taken when they were at secondary school (aged 11–12 years). We were able to follow up 94% of children for their final set of measurements, an exceptionally high follow-up rate; we think that this is because schools, children and families helped us design the trial.

Despite HeLP's success in engaging children, families and teachers, there was no difference in weight status between children who had received the programme and those who had not. There was no difference in the amount of physical activity children did or in the amount of time they spent not being active. We saw a positive difference in some snacking behaviours, with children who had taken part in HeLP eating fewer unhealthy snacks and having less unhealthy foods generally.

Given that the programme failed to achieve sufficient change in behaviour to prevent overweight or obesity, we think that new approaches are needed to support families and children in making healthy lifestyle choices.

# Scientific summary

## Background

The Healthy Lifestyles Programme (HeLP) was a school-based intervention designed to prevent overweight and obesity in children. The intervention was developed using intervention mapping (involving extensive stakeholder involvement) and was guided by the information, motivation and behavioural skills (IMB) model. The intervention has four phases and runs over three school terms with 9- to 10-year-old children. HeLP aims to engage children, parents and schools through a mixture of educational activities, drama, goal-setting and reinforcement activities to help children to increase healthy behaviours and reduce the risk of overweight and obesity.

## Objectives

To estimate the effectiveness and cost-effectiveness of HeLP in preventing overweight and obesity in children.

## Specific objectives

1. To assess the effectiveness of HeLP in children aged 9–10 years by comparing between intervention and control schools:

   - body mass index (BMI) standard deviation score (SDS) at 24 months (primary outcome)
   - BMI SDS at 18 months
   - waist circumference SDS at 18 and 24 months
   - percentage body fat SDS at 18 and 24 months
   - proportion of children classified as underweight, overweight and obese at 18 and 24 months
   - physical activity [time, in average minutes per day, spent in sedentary, light, moderate, moderate to vigorous or vigorous physical activity, and the average volume of physical activity in milligravity (m*g*) units] at 18 months
   - food intake (number of energy-dense snacks, healthy snacks, negative and positive food markers) at 18 months.

2. To estimate the costs associated with the delivery of the HeLP intervention and its cost-effectiveness versus usual practice.
3. To conduct a mixed-methods process evaluation and a mediational analysis to explore the way in which the programme worked (i.e. how it was delivered, taken up and experienced, and the possible mediators of change).

## Methods

We undertook a cluster randomised controlled trial with follow-up at 18 and 24 months post baseline (6 and 12 months post intervention) in 32 primary schools in Devon that had at least 20 pupils in Year 5 (aged 9–10 years). We aimed to recruit half of the schools with ≥ 19% pupils eligible for free school meals (the national average in 2012). Schools were randomised post collection of baseline measures. Schools allocated to receive HeLP started the programme in the spring term of Year 5 with phase 1 activities to create a receptive context in the school and to engage the children and their families. Phases 2 and 3 (healthy lifestyles week and goal-setting) took place in the summer term of Year 5 and reinforcement

activities (phase 4) took place in the autumn term of Year 5. Schools allocated to the control group continued with their standard syllabus. For practical reasons, the trial was run over two cohorts (each with 16 schools). The primary outcome was BMI SDS at 24 months post baseline. The secondary outcomes were BMI SDS at 18 months, waist circumference SDS and percentage body fat SDS, weight status category at 18 and 24 months, accelerometer-assessed physical activity and child-reported consumption of food and drink (averaged across the week, as well as for weekday and weekend separately). Children had 18-month measures taken at the end of Year 6 (aged 10–11 years) and in the autumn term of Year 7 (aged 11–12 years) when the children had moved to secondary school.

Potential mediators were assessed using a bespoke questionnaire that was underpinned by the IMB model and was developed to capture knowledge, cognitions and behaviours that we anticipated would be related to changing diet and physical activity and, hence, weight status. An exploratory factor analysis generated four composite variables: 'confidence and motivation', 'peer norms', 'family approval/behaviours and child attitudes' and 'behaviours and strategies'. These, as well as 'knowledge', were used in the subsequent mediational analysis. We collected details on the cost of the intervention and developed a framework for estimating the cost-effectiveness of the intervention. Questionnaires, focus groups and interviews were conducted with parents, teachers, head teachers and children, and observations of intervention delivery were conducted to assess how HeLP was delivered and received, as well as the experience of being part of the programme.

## Results

We recruited 32 schools and 1324 children. Only 34 eligible children opted out of the study and we had a rate of 94% follow-up for our primary outcome at 24 months. No difference was found between children in the intervention schools and children in the control schools at 24 months for BMI SDS [mean difference (intervention minus control) –0.02, 95% confidence interval (CI) –0.09 to 0.05] for any anthropometric measure (all adjusted for baseline values) at 18 or 24 months or for physical activity at 18 months.

Children in the intervention schools reported eating significantly fewer energy-dense snacks per day (–0.37, 95% CI –0.66 to –0.07) and fewer negative food markers per day (–0.47, 95% CI –0.91 to –0.02) than children in the control schools, when these were averaged across the week. These differences were also evident for weekday, but not weekend, reporting of energy-dense snacks and negative food markers. The intervention increased children's knowledge, cognitions and behavioural skills. The effect of the intervention on reported weekday negative food markers was fully mediated ($p = 0.059$) by three composite variables, confidence and motivation, family approval/behaviours and child attitudes and behaviours and strategies, as well as knowledge. The intervention effect on weekday energy-dense snacks was partially mediated ($p = 0.041$) by the same variables, apart from behaviours and strategies.

The cost of implementing the intervention was estimated at approximately £210 per child. Costs were predominantly those associated with the time inputs for the HeLP co-ordinator and drama components of the programme. Given the lack of effectiveness, the programme is not cost-effective compared with the control group; results are unambiguous, with no expected health gains and schools' usual practice being less costly.

The process evaluation showed that the programme was delivered with a high degree of fidelity in respect of both the components of the intervention and the manner in which the components were delivered. Over 95% of children took part in at least 8 of the 10 sessions of healthy lifestyles week and set goals. Over 90% of children were considered to be actively engaged with the programme according to prespecified criteria, with no differences across socioeconomic groups. Similarly, 75% of parents were considered engaged, and all but three of the schools were deemed to be very engaged and enthusiastic. Parents and children reported changes that they had made to diet and activity, and the majority of barriers to and/or facilitators of achieving goals were related to parental support or lack of it.

## Limitations

We had a low response rate (25%) to the parental questionnaire that was part of the process evaluation. The schools involved in the trial were broadly similar to those in the south-west and England in terms of socioeconomic deprivation, class sizes and both rural and urban settings, but not in terms of ethnicity.

## Conclusions

HeLP is neither effective nor cost-effective in affecting BMI SDS or preventing overweight and obesity in 9- to 10-year-olds, nor is it effective in increasing physical activity. There is weak evidence of effectiveness in reducing the consumption of unhealthy foods (negative food markers and energy-dense snacks).

## Future research

Our findings show that, although it is an intensive intervention that was able to engage schools, children and their families, HeLP did not affect objectively measured physical activity or weight status. Schools are an obvious location for health promotion programmes but we consider it unlikely that any school-based programme that is not part of a wider co-ordinated whole-systems approach will be able to prevent overweight and obesity in a single age group. Research should address the effectiveness of programmes that incorporate whole-systems approaches.

The failure of this intensive intervention, with high levels of engagement, to have a sufficient impact on objectively measured behaviours in children aged 9–10 years still leaves open questions about whether or not it is possible to design effective interventions aimed at very young children whose behaviours may be more malleable or at older children who may have greater autonomy than younger children in making healthy choices.

## Trial registration

The trial is registered as ISRCTN15811706.

## Funding

# Chapter 1 Introduction

## Scientific background

Obesity is considered to be one of the greatest challenges facing public health in the 21st century. Currently, one-fifth of boys and girls in England start school overweight or obese, and one-third of children leave primary school (aged 11 years) overweight or obese.[1] Childhood obesity is strongly associated with socioeconomic status, with children from the least affluent decile being twice as likely to be obese as children from the most affluent decile.[2] Childhood obesity tracks into adulthood; a recent systematic review[3] reported that obese children are five times more likely than non-obese children to be obese as adults, and that 80% of obese adolescents were still obese in adulthood.

School-based obesity preventative interventions have the potential to reach a large number of children and families across the socioeconomic spectrum, and schools provide the organisational, social and communication structures to educate children and parents about healthy lifestyles. Systematic reviews of school-based interventions to prevent obesity and/or increase physical activity and reduce sedentary behaviours show, at best, moderate evidence of effectiveness, with the majority of studies conducted in the USA.[4] Several methodological shortcomings have also been identified, such as differential loss to follow-up, not having sufficient statistical power to detect clinically meaningful differences between groups, particularly on adiposity outcomes, and short-term follow-up.[5,6]

## Evidence for the effectiveness of school-based obesity prevention programmes

In 2012, Khambalia et al.[7] examined the quality of evidence and findings from existing systematic reviews and meta-analyses of school-based programmes in the control and prevention of childhood obesity published between 1990 and 2010. Eight reviews and 112 studies were examined in total, with the most recent literature search conducted in 2008. All eight reviews acknowledged that the studies they included had heterogeneous designs, outcomes and ages of participants. The reviews were judged for their quality, and five were considered to be high quality according to their prespecified criteria.

These five reviews included three meta-analyses: Cook-Cottone et al.,[8] which examined study factors and effect sizes in obesity prevention programmes in 6- to 12-year-olds; Gonzalez-Suarez et al.,[9] which focused on the management of obesity in 11- to 18-year-olds; and Katz,[10] which determined the effectiveness of school-based strategies for both obesity prevention and obesity management in 3- to 18-year-olds. The other two reviews were qualitative systematic reviews: Brown and Summerbell,[11] which assessed the effectiveness of prevention programmes focusing on changing dietary intake and/or physical activity levels in 5- to 18-year-olds; and Kropski et al.,[12] which provided a focused evaluation of the quality and results of long-term obesity prevention programmes in 2- to 19-year-olds.

Despite the heterogeneity of studies included in these reviews, the authors concluded that there was evidence to recommend that school-based obesity prevention programmes include both diet and physical activity components, involve the child's family and be of longer duration. However, they did suggest that, given no one single intervention will fit all school populations, further high-quality research needs to focus on identifying specific programme characteristics and approaches predictive of success.

Wang et al.[4] is the most recent systematic review and meta-analysis of childhood obesity prevention programmes. Of the 139 intervention studies identified, 115 (83%) were located in the primary school, of which 37 were school-based only. These studies assessed the effect of a combined diet and physical

activity intervention on weight-related outcomes. The approaches included intensive classroom physical activity lessons led by trained teachers, moderate to vigorous physical activity sessions, nutrition and education materials, and promoting and providing a healthy diet. There was weak evidence that these approaches were effective at reducing body mass index (BMI), BMI standard deviation score (SDS), prevalence of obesity and overweight, percentage body fat, waist circumference and skinfold thickness. Intervention studies that reported a significant effect tended to be of long duration (between 52 and 156 weeks), with the longer-term programmes having the greatest effect. A further 28 studies were school-based with a home component (e.g. family homework lessons, training and information sheets for parents, nutrition programme for children and family). These studies provided moderate evidence of effectiveness, with half reporting statistically significant beneficial intervention effects. However, in all studies, a range of adiposity measures was used and there was high study heterogeneity in terms of setting, design, sample size, characteristics, intervention approach and length of follow-up, which makes cross-comparisons challenging. The review was unable to identify specific programme characteristics and approaches predictive of success or to explore the comparative effectiveness of specific intervention approaches (e.g. educational interventions vs. environmental intervention).

Since the publication of the above reviews, a number of additional evaluations of school-based interventions[13–16] have been published involving children of a similar age to those in the Health Lifestyles Programme (HeLP) study. However, evidence of effectiveness in changing behaviours and/or weight status of children continues to be inconsistent and the content of the intervention varies greatly between studies.

In 2010, The Healthy Study Group[13] published its findings from a 3-year cluster randomised controlled trial (RCT) of a multicomponent programme addressing risk factors for diabetes among American children whose race or ethnic group and socioeconomic status placed them at high risk of obesity and type 2 diabetes mellitus (T2DM). The intervention consisted of four integrated components: (1) nutrition, (2) physical activity, (3) behavioural knowledge and skills, and (4) communications and social marketing. No significant group differences were observed in the prevalence of overweight and obesity (primary outcome), but children in the intervention schools had a greater reduction in the secondary outcome of BMI SDS (–0.01 vs. –0.05; $p = 0.04$) at 3-year follow-up.

In 2011, Jansen et al.[15] published their findings from an 8-month cluster RCT of a school-based intervention to reduce overweight and improve fitness in 6- to 12-year-old primary school children in a multiethnic, low-income, inner-city neighbourhood in Rotterdam, Netherlands. This was a multicomponent programme based on behavioural and ecological models, involving physical education sessions delivered by a specialist physical education teacher, additional sport and play activities outside school hours, and an educational programme. Significant positive intervention effects were found in terms of the percentage of overweight children, waist circumference and 20-minute shuttle run among those aged 6–9 years. The prevalence of overweight in 6 to 8-year-olds increased by 4.3% in the control group and by 1.3% in the intervention group. No significant effects were found for BMI for 9- to 12-year-olds.

In 2014, the results of the Active for Life-year 5 (AFLY5)[16] study, a UK primary school-based cluster RCT aimed at increasing physical activity, reducing sedentary behaviours and increasing fruit and vegetable consumption in 9- to 10-year-old children, were published. No differences were observed between the intervention and control groups in the three primary outcomes outlined above. The intervention was effective for three out of nine secondary outcomes after multiple testing was taken into account: self-reported time spent in screen viewing at the weekend, self-reported servings of snacks per day and servings of high-energy drinks per day were all reduced. The intervention was an adaptation of the American school-based intervention Planet Health,[17] consisting of 16 lessons (delivered by the class teacher) and 10 pieces of homework in which the children were encouraged to work with their parents.

Grydeland et al.[14] assessed the effect of a 20-month cluster RCT of a school-based intervention trial on the BMI of 11-year-old children in Norway. The intervention took a multilevel approach including collaboration with school principals and teachers, school health services and parent committees. School teachers were

the key in delivering the intervention components, which included lessons, fact sheets for parents, meetings and awareness raising of leisure time activities. There was no significant effect of the intervention overall, although significant effects were found for both BMI ($p = 0.024$) and BMI SDS ($p = 0.003$) in girls. Furthermore, children of higher-educated parents seemed to benefit more from the intervention, highlighting the need to develop interventions that do not widen the health inequalities that already exist between children from different socioeconomic groups.

## Rationale for HeLP

We began work to design, pilot and then fully evaluate a school-based obesity prevention programme in primary schools in 2006. Our aim was to create a novel programme of activities using an intervention mapping approach[18] and to follow the UK Medical Research Council's (MRC) framework for the development and evaluation of complex interventions.[19] The intervention mapping approach combines theoretical and stakeholder perspectives, empirical evidence and new research to assess and develop a potential set of possible solutions for a particular problem, rather than defining a practice or research agenda around a specific theory. We therefore searched the literature for appropriate systematic reviews of school-based obesity prevention programmes and carried out extensive stakeholder consultation with practitioners (teachers, head teachers and drama specialists), local policy-makers (director of public health and the public health policy lead), public health commissioners and the local healthy schools team, in order to understand the school, education and public health context at that time. The overarching aim of the research was to develop a school-based intervention, assess its feasibility and complete pilot work before we sought funding to run the full-scale effectiveness trial, were that deemed appropriate.[20]

Around the time of our initial research, the Brown and Summerbell review[11] concluded that interventions that increase activity and reduce sedentary behaviours may help children to maintain a healthy weight, although the results were short term and inconsistent. Their conclusions were that a combined approach was likely to be more effective in preventing children becoming overweight in the long term.

We wanted to incorporate these recommendations into our programme and address the methodological limitations reported in other trials, such as insufficient statistical power; biased recruitment from, and retention of, more affluent schools, children and families; short follow-up periods; and high levels of attrition. Thus, a crucial focus in developing HeLP was to promote the engagement of schools, children and their families throughout the intervention, as we believed this to be essential for behaviour change to occur. In line with the World Health Organization's Health Promoting Schools framework,[21] we aimed to develop activities that were compatible with the existing school curriculum and promote messages in a manner that could have an impact on the wider school culture, as well as specific behaviours of children and their families. First and foremost, we sought to build supportive and trusting relationships with teachers, children and their families by employing people with specific skills and competencies to deliver the intervention.

We developed an initial phase of the intervention to create a receptive context for the programme and utilised engaging delivery methods to try to increase school, child and family engagement with the programme. Previous research into preventing childhood obesity has found that it can be difficult to engage parents in order to affect change within the family;[22,23] thus, we believed that the delivery methods needed to be sufficiently dynamic, creative and empowering to motivate the children to talk about the activities at home with their parents and encourage parents to come into the school to attend key events. This thinking led us to explore the use of interactive drama, as it had shown promise in promoting positive attitudes towards a number of health behaviours[24] and was a means of delivering a range of behaviour change techniques in manner that could engage the children.

The development of the intervention was guided by the information, motivation and behavioural skills (IMB) model[25] and intervention activities were ordered to enable support and sustain behaviour change in accordance with the health action process model.[26]

Between 2006 and 2008, we worked with 8- to 11-year-old children in one local primary school to assess a range of delivery methods for key objectives identified in the intervention mapping process. We then went on to refine the newly developed intervention based on feedback from teachers, children and parents, and carried out a feasibility study (examining before-and-after intervention changes in the same children) in a second primary school from a more economically deprived ward of Exeter, Devon. In the feasibility study we focused solely on children aged 9–10 years, as our early work demonstrated that this age group was more receptive to the messages and, crucially, engaged their families to the greatest extent.[27] Teachers also informed us that the Year 5 curriculum had the flexibility to deliver the healthy lifestyles week, whereas in Year 6 there was pressure to focus on Standard Assessment Tests (SATs). Based on qualitative data from teachers and head teachers in the piloting stages, the intervention was extended to add activities at the beginning of Year 6. Teachers felt that this would not have an impact SATs teaching and was important in further motivating the children to change behaviours. Minor refinements were also made to the education lessons and the drama scripts to enhance delivery and continuity. In 2008 we carried out an exploratory RCT [funded by the National Institute for Health Research (NIHR) Research for Patient Benefit programme] with 202 children aged 9–10 years in four schools to assess 'proof of concept' and the feasibility of running a cluster RCT of HeLP. The trial design and the intervention were feasible and acceptable to schools, children and their families. At 18 months, compared with children in the control schools, children in the intervention schools consumed fewer energy-dense snacks and more healthy snacks, had less negative food markers and more positive food markers, had lower mean television/screen time and spent more time doing moderate to vigorous physical activity every day.[28] Children in the intervention group had lower average anthropometric measures at 18 and 24 months than children in the control group, with larger differences at 24 months than at 18 months for all measures.[28]

Funding for the definitive trial was awarded in March 2012 and the trial began in September 2012. None of the schools that were involved in the feasibility and pilot work were included in the definitive trial presented in this monograph.

## Rationale for our study design

The HeLP study was designed to address some of the methodological weaknesses in previous school-based RCTs to prevent obesity. Schools have been highlighted as providing a captive audience with the potential to access children across the socioeconomic spectrum; however, few studies took a relational approach to both the intervention and the trial design, whereby the building of relationships with all participants to enhance engagement with the programme and the trial was a key focus in recruitment, data collection, and intervention design and delivery.

Specifically, the design of the HeLP study ensured that (1) baseline measurements were collected before schools (and children) were allocated to the intervention or the control group; (2) a standardised objective measure with a significant follow-up time was utilised to assess obesity prevention (BMI SDS at 24 months); (3) an objective measure of physical activity, with evidence of high compliance, was used[29] (GeneActiv activity monitor; www.activinsights.com); (4) half of the schools had ≥ 19% of children eligible for free school meals (the national average at the time of the trial) to ensure that the results were generalisable across the socioeconomic spectrum; and (5) loss to follow-up was minimised by creating detailed standard operating procedures for the collection of all measures at each time point, with a focus on enhancing positive relationships with teachers, children and parents.

During intervention delivery we ensured that (1) delivery personnel had the necessary skills and competencies to build relationships; (2) there was one key contact person per school who co-ordinated the research hand delivery (the HeLP co-ordinator); (3) the delivery methods used were suitably engaging and dynamic for the target group, so that the children would be motivated to take the messages home and initiate discussion with their family; (4) the intervention used strategies to enhance identification with,

and ownership of, the healthy lifestyle messages and enabled children to explore possible solutions and propose changes in preparation for real-life situations; (5) information and guidance on promoting healthy eating and physical activity went home to parents throughout the intervention period; and (6) parents were involved in setting goals with their child.

The intervention aimed to change behaviours both inside and outside the school environment; these measurements were concerned with behaviours across the week, including weekends. We aimed to capture physical activity data across the whole week for children at baseline and at 18-month follow-up. To support this aim, we used wrist-worn activity monitors (the GeneActiv accelerometer), as feasibility studies had demonstrated high rates of compliance with this monitor in this age group.[30] We took account of the clustered nature of the trial design in the sample size calculation and analysis, and examined the effect of the intervention 6 months (18 months post baseline) and 12 months (24 months post baseline) after its completion in order to examine whether or not any observed effects of the intervention were sustained (see the published trial protocol[31]).

## Aim and objectives

The aim of this cluster RCT was to determine the effectiveness and cost-effectiveness of HeLP in preventing overweight and obesity in children.

Our objectives were:

1. to assess the effectiveness of HeLP in children aged 9–10 years by comparing between intervention and control schools:

   i. BMI SDS at 24 months (primary outcome)
   ii. BMI SDS at 18 months
   iii. waist circumference SDS at 18 and 24 months
   iv. percentage body fat SDS at 18 and 24 months
   v. proportion of children classified as underweight, overweight and obese at 18 and 24 months
   vi. physical activity (average time spent per day in sedentary, light, moderate, moderate to vigorous and vigorous activity) and average weekly volume of physical activity in milligravity (m*g*) units at 18 months
   vii. food intake (number of energy-dense snacks, healthy snacks, negative and positive food markers) at 18 months.

2. to estimate the costs associated with the delivery of the HeLP intervention and its cost-effectiveness versus usual practice
3. to conduct a mixed-methods process evaluation and mediational analysis to explore the way the programme worked (i.e. how it was delivered, taken up and experienced, and the possible mediators of change).

# Chapter 2 Trial design and methods

## Study design

We carried out a pragmatic, superiority, cluster RCT with blinded outcome assessment, allocating schools (1 : 1) to either the HeLP intervention or usual school practice. As the intervention was designed to be delivered in schools, a cluster design was used. Individual child measurements were collected at baseline, 12 months [My Lifestyle Questionnaire (MLQ) only, which assessed potential mediating variables], 18 months and 24 months (anthropometric only) post baseline. Alongside the evaluation of effectiveness, we carried out a parallel economic and process evaluation.

Following the exploratory trial[28] and after assessing delivery requirements, we decided to run the trial in two cohorts, each with the same number of intervention and control schools. All schools were recruited in spring 2012 and then allocated to cohort 1 (commencing the trial in September 2012) or cohort 2 (commencing the trial in September 2013).

### Ethics approval and research governance

We obtained ethics approval for the trial from the Peninsula College of Medicine and Dentistry Research Ethics Committee (reference number 12/03/140) in March 2012. Research governance approval was obtained in June 2012 from our sponsor, the Royal Devon & Exeter NHS Foundation Trust (study number 1304762). A Trial Steering Committee (TSC) was composed and chaired by Professor Martin White (editor-in-chief of the Public Health Research programme), with five other independent members and the research and development director for the Royal Devon & Exeter NHS Trust as the research sponsor (see *Appendix 1*). After discussion at the first TSC meeting (5 November 2012), it was agreed that a Data Monitoring Committee was not necessary because the risk of harm was considered low and no interim analyses were planned.

### Stakeholder involvement

The development of HeLP and the conduct of the trial benefited greatly from the extensive stakeholder involvement in both intervention and trial design. We worked with a group of teachers, head teachers, parents and children from the early piloting of HeLP[27] who became our Project Advisory Group (PAG). We invited teachers and parents from our PAG to be partners on our research bids for both the exploratory[28] and the definitive trial.[31]

Membership of the PAG increased as we progressed though the piloting phases. Meetings were held when required and at times convenient to the group (usually 4–6 p.m.), and all expenses were paid, including cover for teaching staff when necessary. Our PAG members advised us on what was feasible and acceptable when taking behavioural and anthropometric measures from 9- to 10-year-old children, and how to communicate with parents about the research process so that they would (1) receive the information, (2) understand it and (3) feel that they were able to engage with the researchers if they had any concerns or queries. In addition, it was important for us to understand how best to recruit schools and engage teachers with both the programme and the trial. The head teacher in our PAG suggested that we recruit schools via a regional network of primary school heads (the Devon Association of Primary Headteachers) during one of their quarterly briefing sessions, and a teacher involved in the exploratory trial[28] offered to talk to head teachers about her experiences of being involved in the programme during the session.

Our PAG also supported the intervention development and delivery, providing invaluable feedback on possible intervention activities and delivery methods and ensuring that they were acceptable to and feasible for schools, children and their families. It was important that any intervention we developed did not widen existing health inequalities and had the potential to engage children and their families from across the socioeconomic spectrum. Our PAG highlighted the importance of quality delivery by personnel

who were able to engage school staff, children and their families. Two parent members helped us to recruit the HeLP co-ordinators for the definitive trial, in addition to providing critical feedback during practise delivery of the parent engagement events.

### Eligibility and recruitment of schools and children

Schools from across Devon were eligible for inclusion, and they were recruited via the Devon Association of Primary Headteachers and local primary school learning community meetings between March and July 2012. The inclusion criteria were state primary and junior schools with children in at least one single Year 5 group of ≥ 20 children. At the start of the trial we estimated that approximately 125 schools were eligible. All children in all Year 5 classes in the school were invited to participate. We aimed to have half of the schools in the trial with at least the national average percentage of pupils eligible for free schools meals (19% at the time of recruitment of schools). Special schools (for children whose additional needs cannot be met in a mainstream setting) were excluded because they were unlikely to be teaching the standard national curriculum around which the intervention had been designed.

Schools that were willing to participate and fulfilled all of the inclusion criteria were then purposely sampled to represent a range of number of Year 5 classes (1–3 classes), locations (urban and rural) and deprivation (< 19% and ≥ 19% of children eligible for free school meals). Schools that were eligible but not sampled for the study were asked if they were prepared to go on a 'waiting list' in case any of the schools allocated to participate in cohort 2 dropped out during the interim 1-year period before commencing participation in the trial.

Children were recruited using an 'opt-out' system, in which detailed written information about the trial was sent directly home to parents/carers via the school, with parents returning an 'opt-out' form if they did not want their child to participate in the measures only or in either the intervention or the measures. Parents were given 3 weeks to return the opt-out form and class teachers regularly reminded the children during this period to encourage their parent(s) to read the pack. Parents were able to speak to the class teacher or the school's allocated HeLP co-ordinator at any time if they required further information, which was also made clear in the written information provided.

All children who were on the registration list at one of the recruited schools at the start of the autumn term of 2012/13, and whose parent/carer did not complete an opt-out form, were classed as participants.

### Randomisation, allocation concealment and blinding

Randomisation was by school. All schools were initially randomly allocated to intervention or control by a computer-generated sequence that was stratified by (1) the proportion of children eligible for free school meals (< 19% or ≥ 19%) and (2) school size (one Year 5 class or ≥ 2 Year 5 classes). For practical reasons, half of the schools commenced the study in 2012 (cohort 1) and the other half commenced it in 2013 (cohort 2). Randomisation was performed by a statistician in the UK Clinical Research Collaboration-registered Peninsula Clinical Trials Unit immediately after all schools had been recruited (i.e. in 2012) but each school's allocated group (intervention or control) was not communicated to the schools, parents or researchers until *after* the baseline measures had been taken in each cohort (2012 for cohort 1 and 2013 for cohort 2). The Peninsula Clinical Trials Unit ensured that there were equal numbers of control and intervention schools in both cohorts in order to facilitate trial delivery.

*Figure 1* shows the timeline cluster[33] for the HeLP study and *Table 1* provides the key to the figure.

### Intervention

Full details of the intervention have been published in the trial protocol and a paper describing the intervention mapping procedures.[34] In summary, following the literature review and stakeholder consultation (step 1 of the intervention mapping process), we mapped out programme objectives and their associated behavioural and environmental determinants (step 2), selected behaviour change techniques and delivery methods (step 3) and then produced intervention components and their associated materials for delivery (step 4). Steps 5 and 6 of the intervention mapping process involved feasibility and piloting. A summary of the intervention is given below.
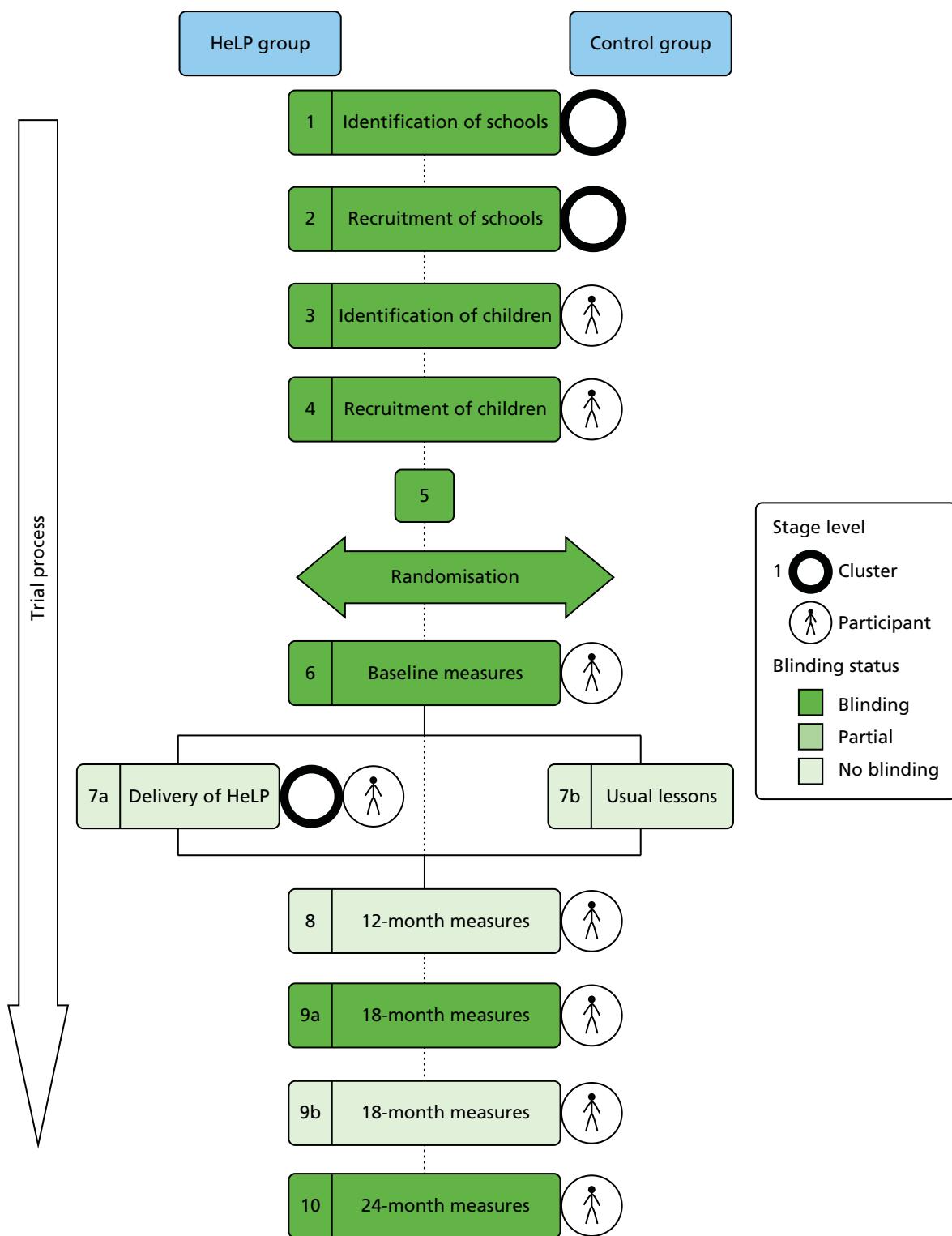
**FIGURE 1** Timeline cluster diagram for HeLP. Reproduced from Lloyd *et al.*[32] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

HeLP was a theory-based, multicomponent, school-based obesity prevention intervention delivered to all Year 5 children (aged 9–10 years) in a school. It consists of four phases delivered during three school terms, which have been ordered to enable and support behaviour change. HeLP delivers a general healthy lifestyle message, encouraging a healthy energy balance with a focus on behaviours relating to the consumption of sweetened fizzy drinks, healthy and unhealthy snacks, the amount of physical activity

**TABLE 1** Key to *Figure 1*

| Item | Description |
|------|-------------|
| 1 | *Identification of schools* |
| | Trial team attended Devon primary head teachers' meetings and network events |
| 2 | *Recruitment of schools* |
| | All Devon state primary and junior schools with children in at least one single Year 5 group of ≥ 20 children eligible to participate. Schools that expressed interest in participating were purposely sampled to represent a range of number of Year 5 classes (1–3 Year 5 classes), locations (urban and rural) and deprivation (< 19% and ≥ 19% of children eligible for free school meals). The remaining schools were placed on a waiting list for cohort 2 (two schools allocated to cohort 2 dropped out before commencing the trial and were replaced by a school from the waiting list) |
| 3 | *Identification of children* |
| | All children in Year 5 in each recruited school were eligible to participate |
| 4 | *Recruitment of children* |
| | Information sent to Year 5 child's parent/carer, who had the opportunity to opt out child from trial |
| 5 | *Randomisation* |
| | Allocation of schools to intervention or control group was done using a computer-generated sequence, stratified on school size (one Year 5 class vs. ≥ 2 Year 5 classes), and percentage of pupils eligible for free school meals (< 19% vs. ≥ 19%) as an indicator of school-level socioeconomic status. After randomisation to intervention or control, schools were allocated to cohort 1 or cohort 2 by a statistician from the Clinical Trials Unit, with equal numbers of intervention and control schools in both cohorts. Allocation remained concealed from all delivery personnel, schools and children until after baseline measures were captured in each cohort |
| 6 | *Baseline measures* |
| | Performed by trained assessors, who were blinded to schools' allocated groups |
| 7a | *Intervention delivery* |
| | HeLP intervention was delivered to schools and children who were allocated to the intervention group, with no blinding for schools, HeLP co-ordinators, children or parents/carers |
| 7b | *Usual care* |
| | Children in schools allocated to control group received usual care, with no blinding for schools, HeLP co-ordinators, children or parents/carers |
| 8 | *12-month measure* |
| | Self-completed MLQ |
| 9a | *18-month measures* |
| | Anthropometric measures collected by trained independent assessors blinded to group allocation |
| 9b | *18-month measures* |
| | Self-completed FIQ while children were still at primary school, 18 months post baseline. Physical activity data from a subset of children while they were still at primary school, 18 months post baseline, were objectively assessed |
| 10 | *24-month measures* |
| | Anthropometric measures were collected by trained independent assessors, blinded to group allocation, after children had moved to secondary school (secondary schools had a mix of children from intervention and control schools), 24 months post baseline (12 months post intervention) |

FIQ, Food Intake Questionnaire. Reproduced from Lloyd *et al.*[32] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

and the amount of screen time. An overarching message promoted throughout the programme is the '80/20' rule, which recommends that we should eat healthily and be active at least 80% of the time. Brief details of each phase are as follows.

- Four HeLP co-ordinators each organised the collection of measurements and delivery of HeLP in eight schools (four intervention and four control). The co-ordinators were also responsible for delivering components of the programme (*Tables 2* and *3*) and for building relationships with schools, children and families. It was believed that having one key contact person for each school would be crucial in building and strengthening relationships with teachers, children and parents during the 1-year intervention period and would increase engagement with the programme, which was believed to be necessary for behaviour change to occur.
- The HeLP co-ordinators were provided with a written delivery manual, and practised delivery of school assemblies, parent assemblies and goal setting, with critical feedback from the trial manager and the parent representatives on the PAG, before delivering the components to the children. The HeLP co-ordinator liaised with the school administrators, head teachers and teachers to discuss the programme and arrange timings for delivery at each intervention school. The teachers were provided with easy-to-read information leaflets at appropriate points throughout the intervention delivery period to ensure that they were aware of upcoming activities. A picture of the HeLP co-ordinator and their contact details were put up at the reception desk.

**TABLE 2** Intervention phases 1 and 2, function, BCTs, delivery methods and personnel

| Phase | Function | BCTs | Component (frequency and duration) | Delivery personnel |
|---|---|---|---|---|
| Phase 1<br><br>Creating a supportive context<br><br>Spring term (Year 5) January–March | • Establish relationships with schools, children and families<br>• Raise awareness and increase knowledge<br>• Promote positive attitudes and norms towards healthy eating and physical activity<br>• Increase self-efficacy for behaviour change | • Provide information on behaviour–health link<br>• Provide information on health–behaviour link<br>• Modelling/demonstrating behaviour<br>• Prompt identification as a role model<br>• Provide information on behaviour–health link<br>• Skill building | Whole school assembly (1 × 20 minutes)<br><br>Newsletter article<br><br>Literacy lesson (to create HeLP rap) (1 × 1 hour)<br><br>Activity workshops (2 × 1.5 hours)<br><br>Parent assembly (1 × 1 hour) involving child performances | HeLP co-ordinators<br><br>HeLP co-ordinators<br><br>Class teacher<br><br>Professional sportsmen/dancers<br><br>Class teachers/ HeLP co-ordinator/ drama group |
| Phase 2<br><br>Intensive Healthy Lifestyles Week – 1 week<br><br>Summer term (Year 5) April–June | • Strengthen relationships with schools, children and families<br>• Increase knowledge<br>• Increase self-awareness<br>• Increase self-efficacy<br>• Develop communication and problem-solving skills<br>• Increase social support (school, peer and family) | • Provide information on health–behaviour link<br>• Problem-solving/barrier identification<br>• Modelling/demonstrating behaviour<br>• Prompt identification as a role model<br>• Communication skills training<br>• Teach to use prompts/cues | Education lessons (5 × 1 hour) (morning)<br><br>Drama (5 × 2 hours) (afternoon) (forum theatre; role play; food tasting, discussions, games, etc.) | Class teacher<br><br>Drama group |

BCT, behaviour change technique.
Reproduced from Lloyd *et al.*[32] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

**TABLE 3** Intervention phases 3 and 4, function, BCTs, delivery methods and personnel

| Phase | Function | BCTs | Component (frequency and duration) | Delivery personnel |
|---|---|---|---|---|
| Phase 3<br><br>Personal goal setting with parental support<br><br>Summer term (Year 5), June–July | • Increase awareness of own behaviour<br>• Increase self-efficacy for change<br>• Develop planning skills<br>• Increase parental support | • Self-monitoring<br>• Goal setting (behaviour)<br>• Problem-solving/ barrier identification<br>• Plan social support<br>• Provide information on where and when to perform a behaviour<br>• Agree behavioural contract<br>• Prompt identification as a role model | Self-reflection questionnaire (1 × 40 minutes)<br><br>Goal-setting sheet to go home to parents to complete with child (1 × 10 minutes)<br><br>One-to-one goal-setting interview (1 × 10 minutes) (goals sent home to parents)<br><br>Forum theatre assembly (1 × 1 hour) | HeLP co-ordinator<br><br>HeLP co-ordinator/ parents<br><br>HeLP co-ordinator<br><br>HeLP co-ordinator/ drama group |
| Phase 4<br><br>Reinforcement activities<br><br>Autumn term (Year 6), September–December | • Increase self-awareness and prioritise healthy goals<br>• Consolidate social support<br>• Develop self-monitoring and coping skills<br>• Increase parental support | • Provide information on health–behaviour link<br>• Modelling/demonstrating behaviour<br>• Prompt identification as a role model<br>• Provide social approval<br>• Prompt self-monitoring<br>• Prompt intention formation<br>• Follow-up prompts<br>• Prompt review of behavioural goals<br>• Prompt barrier identification and resolution<br>• Coping plans | Education lesson (1 × 1 hour)<br><br>Drama workshop (1 × 1 hour), followed by a class-delivered assembly about the project to rest of school (1 × 20 minutes)<br><br>One-to-one goal-supporting interview to discuss facilitators/barriers and to plan new coping strategies (1 × 10 minutes) (renewed goals sent home to parents) | Class teacher<br><br>Drama group<br><br>HeLP co-ordinator<br><br>HeLP co-ordinator |

BCT, behaviour change technique. Reproduced from Lloyd *et al.*[32] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

- Phase 1 (spring term of Year 5) of the intervention aimed to *create a supportive context* by establishing relationships, raising awareness of the programme and setting the foundation for the successful delivery of subsequent components. Professional sports people and dancers came to the schools to run practical workshops and introduce the importance of healthy lifestyles. The aim was to create a 'buzz' in the school and help to set a positive atmosphere for future activities. At the end of this phase, children showcased the skills they had learnt in a parent assembly, during which the HeLP co-ordinator gave parents further information about the programme.
- Phase 2 (summer term of Year 5) was the intensive healthy lifestyles week, which involved education lessons (delivered by the class teacher) every morning and interactive drama activities (delivered by a local drama group) every afternoon during the week. *Table 4* lists the learning objectives for each education and drama session. Teachers were provided with a teaching manual containing all five lesson plans and associated resources, including a compact disc. No training was required for the teachers to deliver the lessons and each lesson plan was set out clearly with objectives linking to the national curriculum at Key Stage 2. Short and simple homework tasks were given at the end of each session for the children to complete in time for the next session. The drama framework was built around four characters (Disorganised Duncan, Fooball Freddie, Snacky Sam and Active Amy), each represented by

**TABLE 4** Learning objectives for each of the five education and drama sessions

| Session | Education lesson learning objectives | Drama session learning objectives |
|---|---|---|
| 1: The energy balance (energy in) | • Understand the energy balance and the importance of it for their health<br>• Understand what energy-dense food means<br>• Understand that different types of food provide energy for the body at different speeds and the impact of this<br>• Understand the balance of good health model | • Getting to know and building relations with characters<br>• Awareness of lifestyles in general and child's own<br>• Increase knowledge of healthy/ unhealthy snack foods and drinks<br>• Increase knowledge of food processing |
| 2: The energy balance (energy out) | • Understand the energy balance and the importance of it for their health<br>• Understand how their diet relates to the balance of good health model<br>• Understand how we use up energy<br>• Understand the activity pyramid and how it relates to the energy balance<br>• Understand that different amounts of energy are needed by the body for different amounts of activity | • Increase understanding of different activity levels<br>• Increase awareness of own activity levels and sedentary behaviour<br>• Raise awareness of importance of sleep<br>• Increase being open to trying of new foods<br>• Increase peer support |
| 3: Temptation and strategies | • Understand the challenges they may face in trying to eat more healthily and reduce screen time<br>• Understand some of the strategies they and their families can use to help them change their behaviours | • Increase awareness of food production<br>• How to read food labels<br>• Understanding and developing persuasive arguments<br>• Problem-solving |
| 4: Decision-making and responsibility | • Be able to reflect on their lifestyle and discuss changes they may need to make to be more healthy<br>• Understand how the media manipulates their decision making in terms of food choices<br>• Understand who is responsible for and who influences their food choices<br>• Understand strategies they can use to be more healthy | • Awareness/reflection of changes in characters<br>• Self-awareness of own healthy lifestyles choices<br>• Develop communication and social support |
| 5: Food marketing and goal-setting | • Understand the three key messages of HeLP<br>• Understand the 80 : 20 rule in terms of healthy lifestyles<br>• Understand the meaning of goals and the use of goal setting as a strategy to change behaviours<br>• Be aware of healthy alternatives to energy-dense snacks and drinks | • Reflection and self-awareness<br>• Goal-setting |

an actor whose attributes related to the three key programme behaviours. Children were asked to choose the character they felt they most resembled and, during the week, they worked closely with that actor to help the character change their behaviour. These sessions were dynamic and interactive and involved role play, games, dance, problem-solving, food tasting and forum theatre. Forum theatre[35] is a technique in which actors act out a family scene and children must focus on the behaviours of one character. If they notice that the character is not adhering to the healthy lifestyle messages, they can shout 'stop' and suggest a change the character could make to improve the outcomes. The child then enters the scene, taking on the role of the character, and the scene is rerun with the suggested change. This method brings the children into the performance, enabling them to have a direct input into the dramatic action they are watching. A drama facilitator led the drama sessions and co-ordinated the delivery of the activities within this component. All actors and the drama facilitator completed a 4-day training programme and were given a detailed manual of the drama scripts and an overview (verbal presentation and written document) of the HeLP intervention. Members of the drama team were

expected to liaise with each other to practise the scripts before delivering the sessions in each school. The themes for each lesson and drama session were:

- the energy balance (energy in)
- the energy balance (energy out)
- overcoming temptation
- decision-making and responsibility
- food marketing and goal-setting.

An information sheet for each topic above was sent home to parents in their child's book bag.

- Phase 3 (summer term of Year 5) was personal goal setting with parental support. Children were encouraged to reflect on their own behaviours and set goals (based on the HeLP messages) with their parents and this was followed up with each child having a 10-minute one-to-one goal-setting discussion with their school's HeLP co-ordinator. A sheet with each child's goals and the character they worked with was sent directly home to parents and a copy was also kept at school in the children's healthy lifestyles folder.
- Phase 4 (autumn term of Year 6) focused on reinforcing all of the messages covered during Year 5 and involved a range of components to refocus the children and their parents on the HeLP messages and behaviour change strategies. This phase included a further lesson (led by the class teacher), a drama workshop (delivered by the actors), a class-delivered assembly to the whole school about the programme and a second one-to-one goal discussion with the HeLP co-ordinator.
- Each phase of HeLP was designed to involve parents as much as possible. In phase 1, a newsletter was sent to parents and there was a parent assembly. In phase 2, an information leaflet was sent home to parents each day based on the theme covered in the drama session and parents were invited in to the school to watch work in progress during the last two drama sessions of the week. In phase 3, children set goals at home with their parent/carer on a 'goal-setting' sheet and returned to school with this sheet to discuss with the HeLP co-ordinator. Finalised goals were then typed up and sent directly home to parents along with the HeLP '80/20' fridge magnet as a further reminder of the programme. Another parent assembly took place after the completed goals sheet had been sent home. In phase 4, following the one-to-one goal-supporting session, a further sheet with the child's goals was, once again, sent home in the post.
- Some components of each phase were directed to the whole school rather than to only the Year 5 children. For example, there was a whole school assembly in phase 1 and a class-delivered assembly to the whole school in phase 4.
- HeLP was designed to allow for some flexibility, so that each activity could fit the context of the school. For example, schools were able to select the timings of parent assemblies and, throughout the year-long intervention delivery period, the HeLP co-ordinator worked closely with the teachers to understand how best to engage and involve the parents, which varied from school to school.

Phases 1–3 of the intervention took place when the children were in school Year 5 (aged 9–10 years) and phase 4 took place at the beginning of school Year 6 (aged 10–11 years). The intervention was designed to fit in with the national curriculum at Key Stage 2 and all lessons and drama sessions included learning objectives relating to personal, social and health education, science, numeracy and literacy. All learning objectives were clearly stated in the HeLP manuals and referenced against the objectives specified in the national curriculum at Key Stage 2.

### *Usual practice*
Schools randomised to the control group continued standard education provision throughout the intervention delivery period, including any involvement in additional health promoting activities, but had no access to the HeLP resources and scripts, which have not been published and were not made available by the research team anywhere except to the intervention schools.

### Sample size calculation

The sample size calculation was informed by the intraclass correlation coefficients (ICCs) for BMI SDS collected during the pilot/feasibility work[27] and by data from another large study involving the National Child Measurement Programme (NCMP).[36] For our primary outcome measure of BMI SDS, we determined the number of schools required (assuming an average of 35 Year 5 children per school, with coefficient of variation of 0.5, and an ICC of 0.02) to detect a between-group difference in BMI SDS of 0.25 units at 24 months with 90% power, a two-sided 5% significance level, a standard deviation (SD) of 1.3 and adjusting for baseline BMI SDS (assuming within-child correlation of 0.8). We allowed for a conservative 20% loss to follow-up or missing data at 24 months.[31] These calculations showed that we needed to recruit 28 schools with a total of at least 952 children to ensure that we had 24 outcome data from 762 children for the analysis. We decided to recruit 32 schools to ensure that we had a minimum of 28 schools completing the trial, each with an estimated average of 35 children.

### Outcome measures

Measurements with children were taken at baseline (before schools' allocated groups were revealed) and at 12 (MLQ only), 18 and 24 months post baseline. The primary outcome measure of effectiveness was change in BMI SDS at 24 months. The secondary outcome measures included BMI SDS at 18 months, waist circumference SDS, percentage body fat SDS and the percentage of children classified as underweight/healthy weight, overweight and obese at 18 and 24 months with raw anthropometric measures being presented for completeness and comparison with other studies. Physical activity was measured using accelerometry and self-reported food intake scores using the validated Food Intake Questionnaire (FIQ) at 18 months, weekday and weekend. (Although this was not explicitly stated in the trial registration or published protocol, the FIQ assesses both weekday and weekend dietary behaviours and it was always our intention to analyse these data separately as well as scores averaged across the week.) Independent blinded assessors were used for the collection of the 18- and 24-month anthropometric measures.

## Outcome assessments

Detailed standard operating procedures were created for the collection of each measure at each time point. Letters were sent home to parents at each assessment time point to remind them that measurements were going to take place. Within each cohort, all child-level measures were collected over an 8-week time period.

Baseline assessments (before revealing schools' allocated groups) were undertaken in the autumn term of school Year 5, between October and November in 2012 (cohort 1) and in 2013 (cohort 2). Outcome assessments were completed: immediately post intervention in October/November 2013/14 (12 months post baseline) for the MLQ only, at 6 months post intervention in June/July 2014/15 (18 months post baseline), which included all behavioural and anthropometric measures; and at 12 months post intervention in October/November (24 months post baseline), which included only anthropometric measures. Anthropometric measurements (height, weight, percentage body fat and waist circumference) were undertaken by trained assessors, co-ordinated by the HeLP co-ordinator (all had completed enhanced Criminal Records Bureau/Disclosure and Barring Service checks). Outcome assessors completed refresher training before each data collection time point. The coefficients of variation to assess inter-rater reliability for height and waist circumference were 0.2% and 1.3% at baseline, 0.1% and 1.2% at 18 months and 0.1% and 0.4% at 24 months, indicating high inter-rater reliability. The independent outcome assessors were blinded to the allocation of schools. At 24 months post baseline, children were measured in their secondary school, and thus secondary school classes contained a mix of intervention and control children. The outcome assessors did not know the child's group allocation. If a child said something to reveal their group allocation, the assessor was asked to record, on the pro forma, that the measurement had been taken unblinded.

The HeLP co-ordinators led the collection of the behavioural measures (FIQ and physical activity using accelerometry), as well of as the MLQ.

## Anthropometric measurements

All anthropometric measures were collected over the course of 1 day in each school. If a child was absent on the day of measurement, up to three further attempts to collect their data were made for up to a further 2 weeks from the day of absence. Height was measured using a SECA stadiometer (Hamburg, Germany), recorded to an accuracy of 1 mm. Weight was measured using the Tanita Body Composition Analyser SC-330 (Tanita, Amsterdam, Netherlands). Weight was recorded to within 0.1 kg and children were asked to take off their shoes and socks or tights. Percentage body fat was estimated from leg-to-leg bioelectric impedance analysis (using the Tanita Body Composition Analyser SC-330). Waist circumference was measured using a non-elastic flexible tape measure placed 4 cm above the umbilicus.

We were mindful that, at baseline, the HeLP co-ordinator had yet to develop a working relationship with the children, so, to put the children at ease and minimise any possible stigmatisation of overweight or sensitive children, the collection of these measurements formed part of a specially designed lesson that was based around measuring in general and ways in which information can be presented. The HeLP co-ordinator led the lesson, which provided them with a good opportunity to learn the children's names and allowed the children to become familiar with the HeLP co-ordinator. Each child, one at a time, left the classroom during the lesson to go to a private room and have their height, weight, waist circumference and percentage body fat by bioelectrical impedance measured by two other trained researchers. For the 18- and 24-month measurements, no special lesson took place, as the children were familiar with the measurement process and felt at ease with the HeLP co-ordinator, who co-ordinated the smooth running of the measurements in each school.

At each data collection time point, the children had the option to decline one or more measurements. For the anthropometric measures using the Tanita scales, a printout was produced that provided information on the child's weight, BMI and percentage body fat. While the measurements were taken, the electronic reading was covered so that the child was unable to read their results. This process had been developed during the pilot phases to reduce any stigmatisation of overweight/obese and underweight children and to minimise any discussion about weight.

## Accelerometer measurements

We used a wrist-worn triaxial accelerometer, called the GeneActiv,[37] to objectively measure physical activity and sedentary time for a subset of the recruited children (one randomly selected class per participating school). These accelerometers are waterproof so they do not need to be removed for swimming or showering. We asked children to wear the accelerometer continuously (including at night) for 8 consecutive days on the wrist of their non-dominant arm. To assist with adherence, information packs were sent to parents 1 week before their child was fitted with the GeneActivs, providing information on wearing the accelerometer as well as guidance to be distributed to sports coaches to prevent the accelerometer being removed during sport. On the day the accelerometers were issued, the HeLP co-ordinator spoke to groups of 10 children about how to comply with the procedures, and answered any questions. Following recommended guidance, children were included in the analysis if they had at least 4 days (including a weekend day) of 10 hours of wear time on each of those days.[38] In these analyses, non-wear was determined if at least two accelerometer axes had a SD of < 13 m$g$ and a range of < 50 m$g$ over 60-minute windows, using moving increments of 15 minutes.[39]

## Food Intake Questionnaire

Food intake was assessed using the adapted version of the validated FIQ.[40] The FIQ asks children about the food and beverages they consumed the previous day and allows an estimation of the *number* of different types of healthy and unhealthy food and drink items consumed per day. This questionnaire was developed for children of the same age as those in this study, and it can be adapted to capture information on the consumption of a number of food groups, depending on the focus of an intervention (Dr Allan Hackett, Liverpool John Moores University, 21 September 2006, personal communication). We adapted the questionnaire so that we could focus specifically on healthy (10 items) and unhealthy (13 items) snacks and drinks, and negative (25 items) and positive (22 items) food markers. Children had to answer yes or no as to whether or not they had consumed each food item the previous day. A number of items make up each category to provide a total score from

which an average number consumed per day is calculated. Questionnaire items, the scoring system and the tolerances for missing data can be seen in *Appendix 2*.

Children completed the FIQ twice, at baseline and at 18 months, to allow the assessment of weekday (completed Tuesday to Friday) and weekend (completed on a Monday) food intake. The HeLP co-ordinator led the two lessons required for the children to complete the questionnaires at each time point. Children were arranged in literacy tables to ensure that assistance could be given as efficiently as possible. Support was also provided by an additional researcher, the class teacher and the teaching assistant.

## My Lifestyle Questionnaire

The MLQ was designed to assess knowledge and a series of potential cognitive and behavioural mediators of any observed differences in outcome between the control and the intervention groups. The questionnaire included items designed to assess knowledge, self-efficacy, intentions, peer norms, family approval, attitudes towards restrictions on behaviour, parental provision and rules, goal-setting, self-monitoring and a range of relevant behavioural skills, including suggestions to and discussion with parents, shopping, cooking and trying new snacks. These items were derived from previous research applying the IMB and health action process models but tailored to the logic model underpinning the development of HeLP (see *Appendix 3*).

### *School assessment*

Information on the school-level characteristics and policies on physical activity and nutrition adopted in each school was collected at baseline (October/November 2012/13) and at 18 months (June/July 2014/15) (see *Appendix 4*) using a questionnaire that was completed by a member of staff and/or an administrator.

## Index of Multiple Deprivation

The Index of Multiple Deprivation (IMD) score was assigned to the lower super output area of each school and pupil as determined by their postcode.[41] The school's IMD score was included as a continuous measure within the mediational analysis; the child-level IMD score was also used as a continuous measure in the secondary analysis of the primary outcome to assess whether or not there was a differential effect on BMI SDS, and the child IMD score was used as quartiles when assessing whether or not there was any differential effect of engagement by socioeconomic status in the process evaluation.

### *Changes to trial protocol*

There were two substantial amendments to the protocol during the course of the trial. The first amendment was to clarify the inclusion criteria to include schools who had one single Year 5 class but who may have had a second class that mixed Year 5 and Year 6 children. The second amendment related to schools allocated to cohort 2 of the study, which were due to start the study in September 2013 (cohort 1 schools started in September 2012). These schools were re-contacted in July 2013, at which time two indicated that their circumstances had changed and that they were no longer able (or eligible) to participate in the trial. Schools that had been placed on the waiting list were then contacted to establish if they were still willing and eligible to participate, of which two were. Given the possibility of selection bias in the two withdrawn schools and in terms of potential imbalance between intervention and control groups in school-level confounders (known and unknown), the 16 schools in cohort 2 (i.e. after replacing the two withdrawn schools with the two schools from the waiting list) were all reallocated to the intervention or the control group. This was completed using a minimisation approach to ensure reasonable balance in the stratification factors between the allocated groups across the combined two cohorts.

## Statistical analysis

A detailed analysis plan was developed by the Trial Management Group and approved by the TSC in November 2015.[42] Three minor amendments were made following the TSC meeting in July 2016 (see *Appendix 5*). These amendments were approved by the TSC chairperson in September 2016. The analysis

plan and a summary of the amendments are published on the NIHR Public Health Research programme website, along with the study protocol.[43]

### General methods

The reporting and presentation of data from this trial are in accordance with the Consolidated Standards of Reporting Trials (CONSORT) guidelines for cluster randomised trials.[44] The primary comparative analyses were conducted on a complete-case intention-to-treat basis, with children analysed according to the group (intervention or control) to which their school was randomised. All comparative analyses allowed for the clustered nature of the data (i.e. children within schools). Unless otherwise specified, all adjusted comparative analyses were adjusted for the two stratification variables (the proportion of children eligible for free school meals and the number of Year 5 classes) and baseline values for the outcome under consideration, when available. The analyses were also adjusted for gender and cohort. Between-group differences with only adjustment for clustering are presented for completeness.[45] The 95% confidence intervals (CIs) for between-group comparisons were calculated and these are presented in *Chapter 3*. When given, *p*-values for statistical significance are two-sided and the significance level was set at ≤ 0.05.

Adjustments were not made for multiple testing as the primary outcome of interest was clearly defined a priori. As this is a trial of a complex intervention, the secondary outcomes are all potentially of interest and relevance to participants, parents and other stakeholders. Interpretation of the clinical significance of any differences between the two groups acknowledges the range of variables being measured.

Summaries of continuous/measurement variables comprise the number of schools or participants and either

 i. the mean and SD *or*
ii. the median and interquartile range

as appropriate for the distributional form of the data under consideration.

Summaries of categorical variables comprise the number of schools or participants and the number and percentage of observations in each category.

Participating schools were compared with state primary schools in Devon and England at the time of school recruitment into the trial (2012) in terms of the following characteristics:

- percentage of children eligible for free school meals
- average number of pupils per school
- percentage of children achieving level 4 at Key Stage 2
- proportion of pupils with English as an additional language.

The recruitment, flow and follow-up of schools and children in the trial are summarised using the CONSORT-style flow diagram appropriate for cluster trials. The extent and distribution of missing data for each variable were assessed and dealt with as detailed in *General methods*, *Missing outcome data*.

### Data sources and data entry

The data analysed came from a number of sources. Data collection for all sources followed standard operating procedures, as outlined in *Study design*, *Outcome measures*. Anthropometric measures were captured on a specifically designed data collection form.

All data were entered twice, first by the data manager and then by another member of the research team, and stored on a secure purposively designed database. Data queries were raised and resolved at data entry. Data discrepancies following second data entry were discussed and resolved with the trial manager.

Electronic data were extracted from the database during the course of the study for the purpose of checking (validating) and for study progress reports, as well as for the end-of-study statistical analyses.

### Comparison of baseline characteristics

Baseline characteristics were collected at the beginning of each cohort of the trial and appropriate summary statistics were computed to compare allocated groups for appropriate balance and to provide an overview of the study sample, at both school and child levels. At the school level, the characteristics included the percentage of children eligible for free school meals, the IMD score derived from the school's postcode, the number of Year 5 classes and the average educational attainment at Key Stage 2 SATs. At the child level, the characteristics included gender, age at baseline data collection, ethnicity, individual IMD value, all anthropometric measurements, physical activity and FIQ.

The formal statistical comparison at baseline of randomised groups is not good practice[46] and, thus, was not undertaken: only summary statistics are presented in *Chapter 3*. We prespecified that should there be any substantial imbalance between randomised groups at baseline, in terms of any relevant variables not already being adjusted for in the primary analysis, further adjusted sensitivity analyses may be performed, to allow for such variable(s), in addition to the prespecified variables for adjustment, to assess the robustness of the primary analysis.[46]

### Primary analysis of primary outcome measure (body mass index standard deviation scores at 24 months)

The primary analysis of the primary outcome, BMI SDS at 24 months, was based on the observed data only and utilised random-effects linear regression modelling, allowing for the clustered nature of the data. Adjusted analyses included the two school-level stratification variables as covariates, as well as baseline BMI SDS, gender and cohort. The means and SDs are presented for each group, together with the mean difference (intervention minus control) between groups, the 95% CI for the mean difference and the corresponding *p*-value. The ICC (with 95% CI) from the random-effects regression model for BMI SDS is also reported.

### Secondary analyses of the primary outcome

A small number of sensitivity analyses of the primary outcome were prespecified in the analysis plan to assess how robust the results of the primary analyses were to any biases from missing data or to children in the intervention group who were categorised as non-compliers. These sensitivity analyses were revised following the TSC meeting in July 2016. The proposed amendments were approved by the TSC (chairperson) prior to undertaking the sensitivity analyses outlined below.

### Amendment 1

Given the low number of missing BMI scores and the low number of data deemed missing at random, a sensitivity analysis was undertaken to look at the effect of missing data using a best-case/worst-case scenario analysis.

The first set of these analyses was based on hypothetically driven assumptions. Given the hypothetical preventative nature of the HeLP intervention, the best-case scenario:

- assumed no change between baseline and 24 months in BMI SDS for children allocated to the intervention group (i.e. the baseline BMI SDS value will be carried forward to replace the missing 24-month BMI SDS value)
- imputed missing 24-month BMI SDS values for children allocated to the control group with their corresponding baseline BMI SDS value plus the (marginal) mean change between baseline and 24 months for the children allocated to the control group with complete baseline and 24-month BMI SDS data.

The worst-case scenario:

- assumed that children allocated to the intervention group who were not obese at baseline were obese at the 24 month follow-up: the 24 month BMI SDS value will be set at the Public Health England threshold for obesity (i.e. the 95th percentile; this is 1.645). For children allocated to the intervention group who were obese at baseline, the baseline BMI SDS value will be carried forward to replace the missing BMI SDS value
- imputed missing 24-month BMI SDS values for children allocated to the control group with their corresponding baseline BMI SDS value plus the (marginal) mean change between baseline and 24 months for the children allocated to the control group with complete baseline and 24-month BMI SDS data.

After imputing the missing 24-month BMI SDS scores for both scenarios, the primary analyses model was fitted to the full intention-to-treat data set to allow us to ascertain if the missing primary outcome data significantly influenced the results of the primary effectiveness analysis.

In addition to the primary analyses, exploratory analyses of the following possible interactions were undertaken to assess whether any effect of the HeLP intervention was modified by (1) gender, (2) baseline BMI SDS, (3) number of Year 5 classes within school and (4) child-level socioeconomic status. These subgroup analyses were performed by adding the interaction term between allocated group and the subgroup variable into the random-effects regression model. A test of the interaction was also performed to assess whether or not there was evidence that the effect of the intervention differed across the two cohorts. As the study was not powered for these exploratory interaction analyses, the results have been interpreted with caution, based on the corresponding CIs for the subgroups.

Finally, a repeated measures model was fitted to all the observed BMI SDS data at baseline, 18 months and 24 months, including effects of time and the interaction term between allocated group and time, to assess whether or not there was evidence that any effect of the intervention differed across time (see *Appendix 6*). In this model, adjustment was made for gender, a child-level fixed effect, and the school-level factors comprising the two stratification variables. The best fitting model was determined through a chi-squared test of the likelihood ratio of the log-likelihoods for the above model for different nested covariance patterns of each child's block of visits in the residual matrix. Stepwise comparisons were made between covariance patterns of increasing parsimony up to an exchangeable pattern. The stopping rule for further pattern simplification was a change in the log-likelihood determined to be significant at the 5% level to the next more parsimonious covariance pattern.

### *Analysis of secondary outcomes*
Secondary outcomes were compared between groups based on the observed data only. Most of the secondary outcomes were of a continuous nature and so comparative analyses followed the approach detailed above for the primary outcome, using random-effects linear regression modelling, allowing for the clustered nature of the data and including the stratification factors, baseline value of the variable under consideration and gender and cohort. Binary outcomes (such as the proportion of children classified as obese at 24 months) were analysed using binary logistic regression, allowing for the clustered nature of the data and including the stratification factors, baseline BMI SDS and cohort. Ordinal outcomes (e.g. categorisation of weight status) were similarly analysed using ordinal logistic regression. For all models, corresponding distributional assumptions were investigated, as outlined below.

### Checking distributional and modelling assumptions
Initial frequency and normal probability plots helped to inform the selection of models fitted to each outcome and whether transformation or model-based transformation under a generalised linear model might be necessary. When outcomes were discrete, sparse categories were amalgamated to create new levels with a sufficient number of participants in each for subsequent modelling as ordinal outcomes.

The tenability of assumptions for modelling outcomes as normally distributed variables was inspected through frequency and normal probability plots of the residuals from the fitted models, as well as box

plots of residuals by allocated group and plots of residuals against fitted values. Model fit was judged through inspection of these plots along with plots of the observations against the fitted values. The distribution of the random school effect was checked by way of a normal probability plot of the best linear unbiased predictors from each model. Either the applied model was revised or a different model was sought if there were any marked deviations from the assumptions. Model stability and influence of the most extreme values were diagnosed through plotting the *dfbetas* calculated for each modelled factor and covariate, with careful attention paid to those for the allocated group. Any observations identified with sufficient influence to change the significance of the intervention effect would prompt further investigation, with results presented both with and without any such data points.

When the outcomes were deemed ordinal and fitted with an ordinal mixed-effects model, the assumption of proportional odds was tested through application of the generalised ordinal model (clustered on schools) to the data, using the *gologit2* package in Stata (version 14, StataCorp LP, College Station, TX, USA). Any significant changes in the coefficients across the levels of the outcome would be identified from significance testing of the Peterson–Harrell parameters.[47]

### Missing outcome data

The recruitment target for this trial allowed for a conservative 20% attrition rate by 24 months, with 24-month data required from 762 children to achieve 90% power in the primary analysis. Substantially more children were recruited than the target, given the higher than expected number of recruited schools with more than one Year 5 class: 1371 children were eligible for recruitment, compared with a pre-estimated target of around 950.

Various trial processes were put in place to minimise missing data. For example, missing data items, such as age and sex, were queried at the time of data entry and up to three visits were made to the school to take measurements for children who were absent on the first measurement day. For the FIQ, when participants were missing a subset of the items, the total score was extrapolated based on the average scores across the four categories (energy-dense snacks, healthy snack foods, negative food markers and positive food markers). To be included in the physical activity analysis, children needed to comply with the required minimum wear time of ≥ 10 hours per day for at least 3 weekdays and 1 weekend day. Non-wear was determined as outlined previously in *Study design*, *Outcome measures*, *Accelerometer measurements*.[39] To minimise missing data due to non-wear, each 15-minute period of device non-wear time was replaced by the children's own data from the same time of day, averaged across all other recorded days.[48] This approach provided a person-specific method for imputing missing data. Any time window with > 50% non-wear was treated as missing.[39]

With the measures to minimise missing data put in place, the likely reasons for a child having missing outcome data at the 18- or 24-month follow-up were as follows.

- The parent/carer opted the child out of the trial before follow-up data collection at either 18 or 24 months.
- The child refused to participate in the collection of anthropometric measures at 18 months but remained in the trial.
- The child refused to participate in the collection of anthropometric measures at 24 months.
- The child had moved out of the county before follow-up data collection at either 18 or 24 months.
- The anthropometric data were missing because the child was withdrawn (but had not left the school) before that time point.
- The child was absent on the day of measurement and on subsequent follow-up visits.

Missing some but not all of the anthropometric measures could occur if a child did not give their assent/ consent for a particular measure (e.g. the child did not want to remove their tights, meaning that it was not possible to collect the percentage body fat measurement). Therefore, the numbers of children with valid data for each of the baseline measurements varied. We explored whether or not the missing data for a particular measure were similar in the two allocated groups.

## Analysis populations and missing data

Full intention-to-treat analysis (i.e. including all participants in the main analyses) requires that all participants have complete follow-up data/no missing data. It was expected from the outset that a small proportion of children would be lost to follow-up by 24 months. As has been previously noted,[49] all statistical methods for handling missing data, including complete-case analysis (i.e. analysis of observed data), have assumptions on the nature or reasons for missingness, which are generally not testable, and so it is necessary to consider the assumptions within the particular context. During the development of the analysis plan, the Trial Management Group considered that the missing at random assumption, necessary for many of the common statistical methods for handling missing data, would be plausible: randomisation was at the school level, opt-out consent was used before baseline measures were collected and it was felt highly unlikely that the delivery of the intervention, or lack of the intervention programme in the control schools, would affect the likelihood of children being absent on days when study data were being collected. Within Devon, it is also known that movement between schools is relatively low.

There was, therefore, no strong a priori reason to assume that children who were lost to follow-up would be missing not at random. For the primary analysis, no imputation of missing anthropometric data was undertaken and the primary outcome analysis was based on the complete-case/observed outcomes data set[49] (i.e. a modified intention-to-treat approach). Given the assumption that any missing primary outcome measures at 24 months would be missing at random, a sensitivity analysis of the primary outcome measure was planned, with missing BMI SDS at 24 months to be imputed using multiple imputation and the analysis re-run on the imputed data set. However, as outlined above in *Statistical analysis*, *Secondary analyses of the primary outcome*, this sensitivity analysis was subsequently replaced by best-case/worst-case sensitivity analyses.

### *Derived outcome variables*

- Body mass index for each child was calculated from height and weight (i.e. weight/height$^2$). The BMI was then standardised by age and gender to obtain the BMI SDS (sometimes known as BMI z-scores); we used the British 1990 (UK90) growth reference charts, and we implemented the calculations using the package LMSgrowth, developed by Cole.[50]
- Categorisations of weight status into underweight, normal, obese or overweight was made based on the definitions from Cole *et al.*;[51] these are the categories used by Public Health England when measuring population of children (e.g. when reporting NCMP results). The weight status categories are defined using the following UK 1990 population cut-off points: underweight, $\leq$ 2nd centile; healthy weight, > 2nd to < 85th centile; overweight, $\geq$ 85th centile; and obese, $\geq$ 95th centile.
- Waist circumference was similarly standardised by age and gender, based on the British 1990 growth reference charts, and implemented using the LMSgrowth package.[52]
- Percentage body fat was estimated from leg-to-leg bioelectric impedance analysis (Tanita Body Composition Analyser SC-330) and standardised by age and gender, based on the British 1990 growth reference charts, and implemented using the LMSgrowth software.[53]
- Average time per day in sedentary, light, moderate, vigorous and moderate to vigorous physical activity intensities was calculated using published cut-off points for children at baseline and at 18 months.[30] In addition, the average volume of physical activity (m*g*) for all valid days was derived. This outcome was added following the first HeLP TSC meeting (November 2012).
- Details of the scoring of the FIQ are given in *Appendix 2*. From the FIQ data, the mean number of energy-dense snacks, healthy snacks, negative food markers and positive food markers for weekday and weekend was calculated, for baseline and 18 months.

# Chapter 3 Results (primary and secondary outcomes)

## Participant numbers through the HeLP study

*Figure 2* shows the flow diagram of the progress of schools and children through the phases of the trial. The target of recruiting 32 schools, to ensure that a minimum of 28 schools completed the trial, was met, with all 32 completing the trial. The number of children in each school was greater than we had anticipated before the recruitment of schools started (the prior assumption was a mean of 35 Year 5 children per school). Therefore, the total number of eligible children was greater than expected. All 1371 children in the 32 participating schools were eligible for the trial: 34 were opted out by their parent/carer and 13 left the school before baseline data collection (and randomisation). All of the remaining 1324 children, irrespective of whether or not they had (all) measurements collected, were included in the trial, and so the minimum requirement of 952 children recruited was far exceeded.

Up to four visits were made to each school in an attempt to obtain measures for children who had been absent on the original day of data collection, with only a small number of children absent at all subsequent visits. As described in *Chapter 2*, one class per school was randomly sampled to complete the physical activity data collection. At 24 months, there was a slightly greater loss to follow-up of children in the intervention group than in the control group (6.1% vs. 3.7%). The number of children included in the primary analysis of BMI SDS at 24 months was 1244, 63% greater than the required 762 children from the a priori sample size calculation.

*Figure 2* shows the trial profile. '*N*' refers to the number of schools (clusters) and '*n*' refers to the number of children (individual participants). Two schools that had been allocated to cohort 2 withdrew while they were waiting to commence the trial and so these were subsequently replaced with two of the four schools on the waiting list before cohort 2 commenced. All schools that started the trial remained in it, and so all of the randomised clusters are present at baseline and at each follow-up point. The percentage given in brackets for the proportion of children with data at both baseline and follow-up is calculated from the total number of recruited children in the schools at baseline. Not all children with a follow-up measure necessarily had a corresponding baseline measure (or vice versa) owing to different children being absent on the day of the main and additional assessments for each of the time points and/or owing to children leaving or moving schools. In all of the analyses, children were analysed in the group (intervention or control) to which the school they were enrolled in at baseline was randomised.

## Completeness of outcome measures through the HeLP study

*Figure 3* shows the completeness of the outcome measures at each of the time points. A small number of children at each time point did not assent to one or more of the anthropometric measures. In summary, between 93% and 99% of children provided data on BMI, waist circumference, percentage body fat and FIQ. The proportion of children who had usable physical activity data files was 96% at baseline and > 88% at 18 months; similarly, the proportion of children with valid physical activity data was 94% and 84% at baseline and at 18 months, respectively, following the application of the minimum wear requirements (3 weekdays and 1 weekend day, each with a minimum of 10 hours of wear time per day). There was no evidence of differences between allocated groups in terms of the completeness of outcome measures throughout the trial, with the exception of the small difference at 24 months for BMI SDS.

**FIGURE 2** Trial profile.

**Baseline**

Total (*n* = 1324)

BMI (height and weight) (*n* = 1312, 99.1%)
Waist circumference (*n* = 1309, 98.9%)
Body fat (*n* = 1309, 98.9%)
FIQ weekday (*n* = 1317, 99.5%)
FIQ weekend (*n* = 1322, 99.9%)
MLQ (*n* = 1310, 98.9%)
PA (analysed files)
(one class per school) (*n* = 851/886, 96.1%)
PA (valid data)
(one class per school) (*n* = 830/886, 93.7%)

**Allocation of schools**

Allocated to intervention (*N* = 16)
Children (*n* = 676)
Mean (coefficient of variation) 42.3 (0.49)
children per school

Allocated to control (*N* = 16)
Children (*n* = 648)
Mean (coefficient of variation) 40.5 (0.44)
children per school

**12 months**

MLQ (*n* = 652, 96.5%)

MLQ (*n* = 626, 96.6%)

**18 months**

BMI (height and weight) (*n* = 644, 95.3%)
Waist circumference (*n* = 645, 95.4%)
Body fat (*n* = 644, 95.3%)
FIQ weekday (*n* = 647, 95.7%)
FIQ weekend (*n* = 648, 95.9%)
PA (analysed files) (*n* = 384/428, 89.7%)
PA (valid data) (*n* = 359/428, 83.9%)

BMI (height and weight) (*n* = 621, 95.8%)
Waist circumference (*n* = 620, 95.7%)
Body fat (*n* = 619, 95.5%)
FIQ weekday (*n* = 625, 96.5%)
FIQ weekend (*n* = 626, 96.6%)
PA (analysed files) (*n* = 404/458, 88.2%)
PA (valid data) (*n* = 386/458, 84.3%)

**24 months**

BMI (height and weight) (*n* = 630, 93.2%)
Waist circumference (*n* = 629, 93.1%)
Body fat (*n* = 629, 93.1%)

BMI (height and weight) (*n* = 620, 95.7%)
Waist circumference (*n* = 618, 95.4%)
Body fat (*n* = 620, 95.7%)

**Analysis**

Primary analysis
Primary outcome BMI SDS (*n* = 628, 92.9%)
Children excluded as no baseline BMI (*n* = 2)

Primary analysis
Primary outcome BMI SDS (*n* = 616, 95.1%)
Children excluded as no baseline BMI (*n* = 4)

**FIGURE 3** Completeness of measures across time points. PA, physical activity.

All anthropometric measures were taken blind to group allocation. A further assessment of the success of allocation concealment was made at the 24-month data collection point, when there was a mixt of children from intervention and control primary schools in each secondary school, with the independent (blind) assessors requested to indicate whether or not the child had revealed their allocated group; no child had reported their allocated group.

*Figure 3* shows the completeness of the measures across the time points. '*N*' refers to the number of schools (clusters) and '*n*' refers to the number of children (individual participants). The percentage given in brackets for the proportion of children with data at baseline and follow-up is of the total number of recruited children in the schools at baseline. Not all children with a follow-up measure necessarily had a corresponding baseline measure (or vice versa) owing to different children being absent on the day of the main and additional assessments for each of the time points and/or owing to children leaving or moving schools. In all of the analyses, children were analysed in the group (intervention or control) to which the school they were enrolled in at baseline was randomised. There were no differences in missing anthropometric data by allocated group at each time point.

## Baseline characteristics

### School-level characteristics

*Table 5* shows how the HeLP study schools compare with schools in Devon and England in terms of the percentage of children eligible for free school meals, the school size, the percentage of children achieving level 4 at Key Stage 2 and the proportion of pupils with English as an additional language or who were non-white British.

Owing to the inclusion criteria for schools for this study (at least one Year 5 class of > 20 pupils and half of schools having ≥ 19% pupils eligible for free school meals), HeLP schools are larger and more deprived than average primary schools in Devon. Although broadly similar to other primary schools in Devon, HeLP schools have a considerably lower proportion of pupils with English as an additional language than do primary schools in England in general.

*Table 6* shows the school level-baseline characteristics. In the intervention group; there were an equal number of schools with single and multiple Year 5 classes, while in the control group, 9 out of the 16 schools had single Year 5 classes. In both groups, ≥ 19% of the children were eligible for free school meals in 7 out of 16 schools. The median school-level IMD scores were comparable between the allocated groups.

**TABLE 5** Comparison of HeLP schools with Devon schools and all primary schools in England

| | Schools | | |
| --- | --- | --- | --- |
| Characteristic | HeLP | Devon | In England[54] |
| Percentage of children eligible for free school meals | 20 | 12.7 | 19 |
| Average number of pupils/school | 306 | 168 | 251 |
| Percentage of children achieving level 4 at Key Stage | 79 | 81 | 81 |
| Proportion of pupils with English as an additional language/non-white British | 4.1 | 2.6 | 16.8 |

**TABLE 6** Comparison of school level baseline characteristics by randomised group

| | Group | | |
| --- | --- | --- | --- |
| Characteristic | Intervention (*N* = 16) | Control (*N* = 16) | All (*N* = 32) |
| Cohort | | | |
|   1 | 8 | 8 | 16 |
|   2 | 8 | 8 | 16 |
| Number of Year 5 classes | | | |
|   Single class | 8 | 9 | 17 |
|   More than one | 8 | 7 | 15 |
| Free school meals | | | |
|   < 19% of pupils | 9 | 9 | 18 |
|   ≥ 19% of pupils | 7 | 7 | 14 |
| Median school IMD score (IQR) | 14,380 (8640) | 13,341 (12,577.5) | 13,933 (11,592) |

### Individual-level characteristics

*Tables 7* and *8* show the comparison of individual baseline characteristics by allocated group. The two allocated groups were similar in terms of the child-level baseline characteristics, including physical activity and food intake scores. The baseline anthropometric measurements, although similar and with considerable overlap, were greater on average in the intervention group, with 12.1% and 14.6% of the intervention group classified as overweight and obese, respectively, compared with 10.7% and 12.6% of the control group.

**TABLE 7** Comparison of individual anthropometric baseline characteristics by randomised group

| Characteristic | Group, mean (SD) | | |
| --- | --- | --- | --- |
| | Intervention (*N* = 676) | Control (*N* = 648) | All (*N* = 1324), mean (SD) |
| Age (years) | 9.8 (0.3) | 9.7 (0.3) | 9.7 (0.3) |
| Gender, *n* (%) | | | |
|     Female | 336 (49.7) | 343 (52.9) | 679 (51.3) |
|     Male | 340 (50.3) | 305 (47.1) | 645 (48.7) |
| Child IMD, median (IQR) | 16,060 (9610) | 13,171 (14,141) | 14,935 (11,450) |
| Ethnicity: white, *n* (%) | 368 (95.3) | 390 (96.3) | 758 (95.8) |
| Height (cm) | 138.9 (6.7) | 137.8 (6.5) | 138.3 (6.6) |
| Weight (kg) | 34.4 (7.7) | 33.3 (7.2) | 33.8 (7.5) |
| BMI (kg/m²) | 17.7 (2.9) | 17.4 (2.8) | 17.5 (2.9) |
| BMI SDS | 0.31 (1.16) | 0.18 (1.14) | 0.25 (1.16) |
| Waist circumference | 61.8 (7.4) | 60.8 (7.5) | 61.3 (7.5) |
| Waist circumference SDS | 0.72 (1.11) | 0.55 (1.15) | 0.64 (1.13) |
| % body fat | 20.1 (7.3) | 20.0 (6.8) | 20.1 (7.1) |
| % body fat SDS | −0.61 (2.18) | −0.63 (2.38) | −0.62 (2.28) |
| % body fat SDS[a] | −0.39 (1.62) | −0.46 (1.52) | −0.42 (1.57) |
| Weight status,[b] *n* (%) | | | |
|     Underweight | 11 (1.6) | 10 (1.6) | 21 (1.6) |
|     Healthy | 479 (71.6) | 483 (75.1) | 962 (73.3) |
|     Overweight | 81 (12.1) | 69 (10.7) | 150 (11.4) |
|     Obese | 98 (14.6) | 81 (12.6) | 179 (13.6) |

IQR, interquartile range.
a   After excluding extreme body fat absolute SDS values of ≥ 5.
b   Weight status categories are defined using the Public Health England definitions.[55]
Reproduced from Lloyd *et al.*[32] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

**TABLE 8** Comparison of individual behavioural baseline characteristics by randomised group

| Characteristic | Group, mean (SD) | | All (*n* = 886), mean (SD) |
| --- | --- | --- | --- |
| | Intervention (*n* = 428) | Control (*n* = 458) | |
| Physical activity (one class per school) | | | |
| Average weekly volume (m*g*) | 49.0 (11.3) | 49.6 (10.9) | 49.3 (11.1) |
| Average daily total activity | 182.7 (36.7) | 185.0 (34.7) | 183.9 (35.7) |
| Average daily light | 129.4 (24.7) | 131.1 (24.2) | 130.3 (24.4) |
| Average daily moderate | 40.0 (12.1) | 40.4 (11.4) | 40.2 (11.7) |
| Average daily moderate to vigorous | 53.3 (16.8) | 53.9 (16.2) | 53.6 (16.5) |
| Average daily vigorous | 13.3 (6.2) | 13.5 (6.2) | 13.4 (6.2) |
| Average daily sedentary | 780.4 (36.1) | 778.2 (34.0) | 779.3 (35.1) |
| FIQ scores | *n* = 676 | *n* = 648 | *n* = 1324 |
| Energy-dense snacks: weekday | 4.0 (2.4) | 4.0 (2.4) | 4.0 (2.4) |
| Energy-dense snacks: weekend | 4.6 (2.5) | 4.4 (2.4) | 4.5 (2.5) |
| Energy-dense snacks: average | 4.2 (2.2) | 4.1 (2.2) | 4.1 (2.2) |
| Healthy snacks: weekday | 3.4 (1.8) | 3.2 (1.7) | 3.3 (1.8) |
| Healthy snacks: weekend | 3.2 (1.9) | 2.9 (1.8) | 3.1 (1.9) |
| Healthy snacks: average | 3.3 (1.6) | 3.1 (1.6) | 3.2 (1.6) |
| Positive food markers: weekday | 6.1 (2.9) | 5.7 (2.8) | 5.9 (2.8) |
| Positive food markers: weekend | 6.0 (3.0) | 5.5 (2.9) | 5.8 (3.0) |
| Positive food markers: average | 6.0 (2.7) | 5.7 (2.5) | 5.8 (2.6) |
| Negative food markers: weekday | 6.5 (3.7) | 6.7 (3.8) | 6.6 (3.7) |
| Negative food markers: weekend | 7.7 (4.0) | 7.1 (3.6) | 7.4 (3.8) |
| Negative food markers: average | 6.8 (3.4) | 6.8 (3.3) | 6.8 (3.4) |

## Primary intention-to-treat analysis of body mass index standard deviation score at 24 months

The unadjusted mean BMI SDS at 24 months was slightly higher in the intervention group, at 0.35, than in the control group, at 0.22, with SDs of 1.25 and 1.22, respectively (*Table 9*). The effect of the intervention was estimated allowing for clustering within schools, modelled as a school-centred random effect, and other prespecified variables. The results showed no evidence of an intervention effect, with an estimated between-group mean difference (intervention minus control) of –0.02 (95% CI –0.09 to 0.05) (see *Table 9*), and an estimated ICC of 0.014 (95% CI 0.003 to 0.069). Without adjustment for the baseline BMI scores and gender of each child, cohort and the stratification factors, but after allowing for the school-level clustering, the mean BMI SDS score was 0.11 (95% CI –0.11 to 0.33) greater in the intervention group than in the control group, but this difference was not statistically significant (see *Table 9*).

**TABLE 9** Primary intention-to-treat and sensitivity analyses of the effect of the HeLP intervention on the primary outcome, BMI SDS, assessed at 24 months post baseline

| Estimation method | Group | | | | Analysis | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Intervention | | Control | | Adjusted | | | Unadjusted | |
| | n | Mean (SD) | n | Mean (SD) | Total *n* in analysis | Mean difference (95% CI) | *p*-value | Mean difference (95% CI) | *p*-value |
| Random-effects estimation | | | | | | | | | |
| Complete data | 630 | 0.35 (1.25) | 620 | 0.22 (1.22) | 1244 | –0.02 (–0.09 to 0.05) | 0.567 | 0.11 (–0.11 to 0.33) | 0.332 |
| Worst-case imputation[a] | 671 | 0.43 (1.25) | 647 | 0.23 (1.23) | 1312 | 0.07 (–0.02 to 0.16) | 0.129 | – | – |
| Best-case imputation[a] | 671 | 0.34 (1.24) | 647 | 0.23 (1.23) | 1312 | –0.02 (–0.09 to 0.05) | 0.485 | – | – |
| GEE estimation | | | | | | | | | |
| Complete data | 630 | 0.35 (1.25) | 620 | 0.22 (1.22) | 1244 | –0.02 (–0.08 to 0.05) | 0.568 | – | – |

GEE, generalised estimating equations.
a  Missing BMI SDS at 24 months was imputed for all children with a baseline BMI SDS ($n = 1312$); there were an additional six children who had BMI SDS at 24 months ($n = 671 + 647 = 1318$) but did not have a baseline BMI SDS and so could not be included in the adjusted analyses ($n = 1312$).
**Notes**
Number of children and mean (SD) observed BMI SDS at 24 months for each group and the estimate of the effect of the intervention (95% CIs and *p*-values) for analyses with and without adjustment for stratification variables, cohort, baseline BMI SDS and gender.
Reproduced from Lloyd *et al.*[32] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

### Model diagnostics for the primary analysis of body mass index standard deviation score at 24 months

Following the fitting of the full random-effects model, plots of the residuals and best linear unbiased predictors for the random school intercept revealed no systematic deviation from the assumed distribution of the observations, required for linear mixed modelling. Inspection of the observed versus fitted values indicated a good fit within the random variation of the data. Inspection of the *dfbeta* statistics of individuals and schools did not reveal any extreme of influential values or instability in the estimation of the coefficients, including for the intervention effect.

### Secondary analysis of body mass index standard deviation score at 24 months

All of the secondary analyses that were undertaken produced results that were consistent with the primary analysis. An alternative computational approach using generalised estimating equations yielded an adjusted between-group mean difference of –0.02 (95% CI –0.08 to 0.05), allowing for clustering around schools and adjusting for the same variables as in the random-effects model in the primary analysis (see *Table 9*). The results from the two approaches are, thus, consistent with each other, with no evidence of an intervention effect.

The sensitivity analysis of the fully adjusted random-effects model based on the worst-case scenario for missing observations at 24 months estimated an increase in mean BMI SDS of 0.07 (95% CI –0.02 to 0.16) for the intervention group relative to the control group (see *Table 9*). Based on the best-case scenario for the same missing observations, the direction and size of the effect of the intervention changed to –0.02 (95% CI –0.09 to 0.05). Therefore, neither scenario provided any tentative evidence of a significant or meaningful intervention effect (see *Table 9*).

### Subgroup analyses of body mass index standard deviation score at 24 months

The results for the prespecified subgroups, defined by gender, school size, free school meal category and cohort, are presented in *Table 10*. The estimated intervention effects were small and in the same direction within subgroups for gender, cohort and school size, with little or no effect on the BMI SDS at 24 months and no evidence of a significant difference from the null effect. Interaction terms between each factor and the allocated group were added to the random-effects model, and the absence of modifying effects by each factor was confirmed by results from the test of the interaction term. All *p*-values were > 0.05. Similarly, there was no evidence of a modifying effect on the intervention by the BMI SDS at baseline or from the individual child-level IMD scores. The point estimate for the effect of the intervention in either level of the free school meal factor was not significantly different from the null, but, although this effect was negative in those schools with < 19% of their children eligible for free school meals, there was a small, non-significant, increase in the BMI SDS in those children from schools in which ≥ 19% of pupils were eligible for free school meals.

**TABLE 10** Primary intention-to-treat analyses of the effect of the HeLP intervention on the primary outcome, BMI SDS, assessed at 24 months post baseline, in prespecified subgroups

| | Group | | | | Adjusted analysis | | |
| | Intervention | | Control | | | | |
| Subgroup | *n* | Mean (SD) | *n* | Mean (SD) | Total *n* in analysis | Mean difference (95% CI) | *p*-value for interaction |
|---|---|---|---|---|---|---|---|
| School size | | | | | | | |
|   Single Year 5 class | 197 | 0.22 (1.28) | 249 | 0.11 (1.23) | 444 | –0.06 (–0.20 to 0.09) | 0.559 |
|   Multiple Year 5 classes | 433 | 0.41 (1.23) | 371 | 0.29 (1.20) | 800 | –0.01 (–0.13 to 0.11) | |
| Free school meal category | | | | | | | |
|   < 19% of children | 417 | 0.28 (1.21) | 359 | 0.05 (1.18) | 774 | –0.03 (–0.16 to 0.09) | 0.613 |
|   ≥ 19% of children | 213 | 0.47 (1.30) | 261 | 0.44 (1.23) | 470 | 0.04 (–0.15 to 0.22) | |
| Gender | | | | | | | |
|   Girls | 316 | 0.33 (1.28) | 328 | 0.27 (1.20) | 642 | –0.01 (–0.09 to 0.08) | 0.849 |
|   Boys | 314 | 0.37 (1.21) | 292 | 0.15 (1.23) | 602 | –0.03 (–0.13 to 0.07) | |
| Cohort | | | | | | | |
|   1 | 237 | 0.33 (1.32) | 381 | 0.22 (1.21) | 615 | –0.01 (–0.18 to 0.17) | 0.568 |
|   2 | 393 | 0.36 (1.20) | 239 | 0.21 (1.24) | 629 | –0.06 (–0.20 to 0.08) | |
| Individual IMD score | 630 | 0.35 (1.25) | 620 | 0.22 (1.22) | 1182 | 0.00 (–0.17 to 0.14) | 0.954 |
| Baseline BMI SDS | 630 | 0.35 (1.25) | 620 | 0.22 (1.22) | 1244 | 0.00 (–0.09 to 0.05) | 0.952 |

**Notes**
Number of children and mean (SD) BMI SDS for each randomised group; estimate of the effect of the intervention (95% CI) within each subgroup; *p*-values for the global test modelled as an interaction with the allocated group in the random-effects model.

In summary, there was no evidence of a significant effect of the intervention in any of the prespecified subgroups. As noted in *Chapter 2*, these subgroup analyses, although prespecified, are considered to be exploratory and the trial was not designed with the aim of being able to detect differences between subgroups.

## Longitudinal analysis of body mass index standard deviation score

The analysis was based on fitting a longitudinal model with observations at fixed time points (baseline, and 18 months and 24 months post baseline). The complexity of the between-time-point covariance for each child was reduced to an autoregressive pattern of order one. As the outcome of this longitudinal model comprised all BMI SDS measures, including those at baseline, the effect of the intervention was determined from the interaction between the allocated group and the time point.

According to this model, there was no significant difference between the two groups at baseline (intervention BMI SDS greater than controls by 0.117; $p = 0.170$). The BMI SDS increased significantly in both groups at each subsequent time point relative to baseline, by an average of 0.04 (95% CI 0.01 to 0.06) at 18 months and by 0.05 (95% CI 0.01 to 0.09) at 24 months, although, relative to the control group, the increase in the intervention group was attenuated, not significantly, by –0.03 (95% CI –0.08 to 0.01) at 18 months and by –0.02 (95% CI –0.07 to 0.04) at 24 months.

*Figure 4* shows the predicted marginal BMI SDS by allocated group across the time points of the trial.
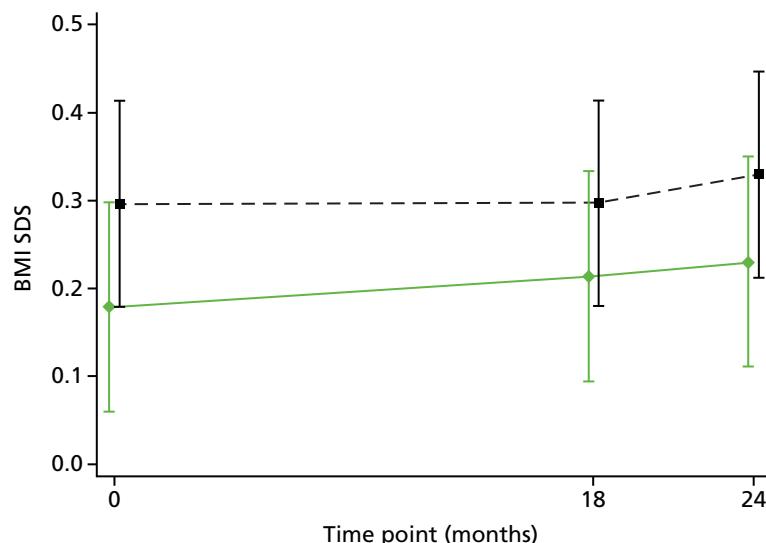


**FIGURE 4** Predicted marginal BMI SDS with 95% CIs in the intervention group (black with dashed connecting lines) and control group (green with solid connecting lines) across time points. According to the longitudinal model, allowing for hierarchical clustering by child within each school, modelling the within-child covariance between fixed time points as an autoregressive pattern of order one. Reproduced from Lloyd *et al.*[32] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

## Primary analyses of secondary outcomes at 24 months (anthropometric measures only)

*Table 11* shows the summary data of the remaining anthropometric measures at 24 months post baseline. Among the children in the intervention group, slightly more were classified as being overweight or obese, according to the Public Health England definition of categorisation of BMI scores at 24 months, relative to the control group (30.8% vs. 26.6%). However, analysing the weight status categories under the assumptions of an ordinal logistic model with school-level random effects, there was no evidence of an intervention effect in the fully adjusted model ($p = 0.472$). The odds of being in the higher weight categories were greater among the children in the intervention group than among those in the control group, although this change in odds was not statistically significant [odds ratio (OR) 1.15, 95% CI 0.79 to 1.67]. The intervention effect size was similar in both the unadjusted and adjusted models. Relaxing the assumption of proportional odds, the estimated intervention effect size from the generalised ordinal model, with a robust error variance clustered on schools, was comparable with that from the random-effects ordinal logistic model (results not shown). According to the generalised ordinal model (i.e. the model without assuming proportional odds across the weight status categories), the odds of being classified as overweight or obese relative to healthy or underweight were 1.14 (95% CI 0.77 to 1.68) in the intervention group relative to the control group. The odds were 1.17 (95% CI 0.74 to 1.84) for the risk of being obese relative to the lower weight status categories in the intervention group relative to the control group. The test of proportional odds under the Peterson–Harrell parameterisation indicated that there was no evidence of a change in the odds between binary comparisons of levels of weight status category ($p = 0.503$), indicating that the assumption of proportional odds for intervention effect had not been violated.

Consistent with the primary outcome measure, the average anthropometric measurements were generally slightly larger for the children in the intervention group at 24 months, as they were at baseline, relative to those in the control group (see *Table 11*). An exception was found in the standardised body fat measurements before the exclusion of the extreme values (implausible absolute SD scores of $\geq 5$). The unadjusted mean differences (intervention minus control) were positive for all anthropometric outcomes except the standardised body fat score that included extreme values. However, there was no evidence of an intervention effect from the unadjusted random-effects model, allowing for clustering on schools. After adjusting for the stratification variables, cohort, gender and the corresponding baseline measurements, all mean differences were negative, indicating lower, albeit non-significant, measurements in the intervention group at 24 months than in the control group.

## Primary analyses of outcomes at 18 months

*Tables 12–14* show the summarised data of the primary and secondary outcomes at 18 months post baseline. Overall, fewer children were classified as overweight or obese than at 24 months, but more children in the intervention group than in the control group were already overweight or obese at 18 months (28.9% vs. 25.4%). Although not statistically significant, this difference was reflected in both the adjusted and unadjusted analyses when the weight categories were modelled as an ordinal outcome. No violation of the proportional odds assumption was detected, according to the test of the Peterson–Harrell parameters from the generalised ordinal model (results not shown).

The anthropometric outcomes at 18 months displayed a similar pattern to those at 24 months, with larger mean values for children in the intervention group than for children in the control group, and positive mean between-group differences (intervention minus control) from the unadjusted analyses, except for the standardised measurement of percentage body fat with extreme values included (extreme values defined as absolute standard scores of $\geq 5$) (see *Table 12*). After adjustment for all of the prespecified variables, the mean between-group differences in anthropometric measurements were all negative, indicating lower values in the intervention group at 18 months than in the control group. However, there was no evidence of a significant intervention effect for any of the anthropometric measures at 18 months.

**TABLE 11** Primary analyses, both adjusted and unadjusted, for secondary outcomes at 24 months post baseline

| | Group | | | | Analysis | | | | | |
| | Intervention | | Control | | Adjusted | | | Unadjusted | | |
| **Weight status category**[a] | **n** | **% of group** | **n** | **% of group** | **Total n in analysis** | **OR (95% CI)** | **p-value** | **Total n in analysis** | **OR (95% CI)** | **p-value** |
|---|---|---|---|---|---|---|---|---|---|---|
| Underweight and healthy | 436 | 69.2 | 455 | 73.4 | | | | | | |
| Overweight | 89 | 14.1 | 84 | 13.5 | | | | | | |
| Obese | 105 | 16.7 | 81 | 13.1 | | | | | | |
| Overweight and obese | 194 | 30.8 | 165 | 26.6 | | | | | | |
| Ordinal logistic model | | | | | 1244 | 1.15 (0.79 to 1.67) | 0.472 | 1250 | 1.20 (0.83 to 1.74) | 0.329 |
| **Anthropometric outcomes** | **n** | **Mean (SD)** | **n** | **Mean (SD)** | **Total n in analysis** | **Mean difference (95% CI)** | **p-value** | **Total n in analysis** | **Mean difference (95% CI)** | **p-value** |
| BMI | 630 | 19.03 (3.55) | 620 | 18.68 (3.34) | 1244 | –0.04 (–0.23 to 0.15) | 0.653 | 1250 | 0.29 (–0.32 to 0.91) | 0.339 |
| Waist circumference (cm) | 629 | 65.47 (8.97) | 618 | 64.73 (8.55) | 1238 | –0.24 (–1.30 to 0.83) | 0.655 | 1247 | 0.76 (–0.96 to 2.49) | 0.374 |
| Waist circumference SDS | 629 | 0.63 (1.24) | 618 | 0.54 (1.21) | 1238 | –0.05 (–0.23 to 0.13) | 0.563 | 1247 | 0.09 (–0.15 to 0.33) | 0.441 |
| % body fat | 629 | 20.01 (8.21) | 620 | 19.76 (7.67) | 1240 | –0.03 (–0.61 to 0.55) | 0.917 | 1249 | 0.18 (–1.26 to 1.62) | 0.801 |
| Body fat SDS | 629 | –0.78 (2.16) | 620 | –0.78 (1.89) | 1240 | –0.04 (–0.29 to 0.22) | 0.758 | 1249 | –0.02 (–0.37 to 0.33) | 0.900 |
| Body fat SDS[b] | 612 | –0.59 (1.84) | 607 | –0.65 (1.69) | 1198 | –0.02 (–0.17 to 0.13) | 0.791 | 1219 | 0.04 (–0.23 to 0.32) | 0.752 |

a  Weight status categories are defined using the Public Health England definitions.[55]
b  After excluding extreme body fat absolute SDS values of ≥ 5.

**Notes**
Number and percentage of children within weight status categories and mean (SD) of anthropometric measures for each randomised group and estimate of the effect of the intervention (95% CIs and p-values) for analyses with and without adjustment for stratification variables, cohort, baseline measure and gender.

**TABLE 12** Primary analyses, both adjusted and unadjusted, for weight status and other anthropometric outcomes 18 months post baseline

| | Group | | | | Analysis | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Intervention | | Control | | Adjusted | | | Unadjusted | | |
| Weight status category[a] | n | % of group | n | % of group | Total n in analysis | OR (95% CI) | p-value | Total n in analysis | OR (95% CI) | p-value |
| Underweight and healthy | 458 | 71.1 | 463 | 74.6 | | | | | | |
| Overweight | 87 | 13.5 | 78 | 12.6 | | | | | | |
| Obese | 99 | 15.4 | 80 | 12.9 | | | | | | |
| Overweight and obese | 186 | 28.9 | 158 | 25.4 | | | | | | |
| Ordinal logistic model | | | | | 1259 | 0.99 (0.63 to 1.55) | 0.952 | 1265 | 1.18 (0.81 to 1.72) | 0.379 |
| Anthropometric outcome | n | Mean (SD) | n | Mean (SD) | Total n in analysis | Mean difference (95% CI) | p-value | Total n in analysis | Mean difference (95% CI) | p-value |
| BMI SDS | 644 | 0.32 (1.23) | 621 | 0.20 (1.21) | 1259 | –0.02 (–0.08 to 0.05) | 0.612 | 1265 | 0.11 (–0.12 to 0.33) | 0.335 |
| BMI | 644 | 18.68 (3.40) | 621 | 18.38 (3.24) | 1259 | –0.07 (–0.22 to 0.09) | 0.386 | 1265 | 0.26 (–0.35 to 0.86) | 0.391 |
| Waist circumference | 645 | 64.96 (8.58) | 620 | 63.93 (7.97) | 1257 | –0.31 (–1.40 to 0.78) | 0.563 | 1265 | 0.80 (–0.79 to 2.39) | 0.310 |
| Waist circumference SDS | 645 | 0.69 (1.18) | 620 | 0.57 (1.15) | 1257 | –0.07 (–0.27 to 0.12) | 0.438 | 1265 | 0.08 (–0.15 to 0.32) | 0.472 |
| % body fat | 644 | 19.33 (8.05) | 619 | 19.28 (7.70) | 1254 | –0.13 (–0.66 to 0.41) | 0.636 | 1263 | 0.03 (–1.42 to 1.48) | 0.963 |
| Body fat SDS | 644 | –0.99 (2.23) | 619 | –0.98 (2.03) | 1254 | –0.02 (–0.25 to 0.22) | 0.897 | 1263 | –0.02 (–0.38 to 0.35) | 0.917 |
| Body fat SDS[b] | 618 | –0.74 (1.84) | 593 | –0.75 (1.73) | 1189 | –0.02 (–0.16 to 0.12) | 0.765 | 1211 | 0.01 (–0.29 to 0.31) | 0.938 |

a Weight status categories are defined using the Public Health England definitions.[55]
b After excluding extreme body fat absolute SDs values ≥ 5.

**Notes**
Number and percentage of children within weight status categories and mean (SD) of anthropometric measures for each randomised group and estimate of the effect of the intervention (95% CIs and p-values) for analyses with and without adjustment for stratification variables, cohort, baseline BMI SDS and gender.

**TABLE 13** Primary analyses, both adjusted and unadjusted, for physical activity outcomes at 18 months post baseline

| | Group | | | | Analysis | | | | | |
| | Intervention | | Control | | Adjusted | | | Unadjusted | | |
| Physical activity outcome | n | Mean (SD) | n | Mean (SD) | Total *n* in analysis | Mean difference (95% CI) | *p*-value | Total *n* in analysis | Mean difference (95% CI) | *p*-value |
|---|---|---|---|---|---|---|---|---|---|---|
| Average weekly volume (m*g*) | 359 | 52.14 (13.95) | 386 | 51.47 (12.95) | 705 | 0.57 (–1.58 to 2.72) | 0.590 | 745 | 0.53 (–2.18 to 3.24) | 0.692 |
| Average daily total | 359 | 199.71 (43.94) | 386 | 198.05 (40.20) | 705 | 1.26 (–6.84 to 9.36) | 0.752 | 745 | 1.23 (–8.24 to 10.70) | 0.792 |
| Average daily light | 359 | 141.72 (27.80) | 386 | 141.07 (27.09) | 705 | 0.70 (–4.73 to 6.13) | 0.793 | 745 | 0.43 (–5.87 to 6.73) | 0.891 |
| Average daily moderate | 359 | 44.26 (16.24) | 386 | 43.46 (13.43) | 705 | 0.41 (–2.28 to 3.09) | 0.759 | 745 | 0.68 (–2.41 to 3.78) | 0.656 |
| Average daily moderate–vigorous | 359 | 57.99 (22.34) | 386 | 56.98 (19.39) | 705 | 0.56 (–2.76 to 3.89) | 0.731 | 745 | 0.85 (–3.24 to 4.94) | 0.674 |
| Average daily vigorous | 359 | 13.734 (7.66) | 386 | 13.52 (7.38) | 705 | 0.15 (–1.01 to 1.3) | 0.798 | 745 | 0.15 (–1.33 to 1.63) | 0.834 |
| Average daily sedentary | 359 | 764.50 (43.29) | 386 | 766.36 (39.88) | 705 | –1.39 (–9.45 to 6.68) | 0.727 | 745 | –1.46 (–10.91 to 8.00) | 0.755 |

**TABLE 14** Primary analyses, both adjusted and unadjusted, for FIQ outcomes at 18 months post baseline

| | Group | | | | Analysis | | | | | |
| | Intervention | | Control | | Adjusted | | | Unadjusted | | |
| Outcome | n | Mean (SD) | n | Mean (SD) | Total n in analysis | Mean difference (95% CI) | p-value | Total n in analysis | Mean difference (95% CI) | p-value |
|---|---|---|---|---|---|---|---|---|---|---|
| Average weekly FIQ | | | | | | | | | | |
| Energy-dense snacks | 646 | 3.72 (1.86) | 624 | 4.06 (2.07) | 1260 | −0.37 (−0.66 to −0.07) | 0.017 | 1270 | −0.29 (−0.63 to 0.06) | 0.098 |
| Healthy snacks | 637 | 3.61 (1.63) | 617 | 3.30 (1.50) | 1227 | 0.22 (−0.04 to 0.47) | 0.092 | 1254 | 0.31 (0.02 to 0.60) | 0.039 |
| Negative food markers | 647 | 5.90 (2.73) | 624 | 6.38 (3.00) | 1262 | −0.47 (−0.91 to −0.02) | 0.041 | 1271 | −0.40 (−0.94 to 0.14) | 0.143 |
| Positive food markers | 647 | 6.20 (2.36) | 624 | 5.77 (2.31) | 1256 | 0.26 (−0.12 to 0.64) | 0.173 | 1271 | 0.42 (0.01 to 0.84) | 0.045 |
| Weekday FIQ | | | | | | | | | | |
| Energy-dense snacks | 647 | 3.54 (2.03) | 625 | 3.99 (2.27) | 1265 | −0.47 (−0.84 to −0.11) | 0.013 | 1272 | −0.41 (−0.83 to 0.01) | 0.055 |
| Healthy snacks | 645 | 3.69 (1.77) | 624 | 3.38 (1.64) | 1257 | 0.23 (−0.08 to 0.54) | 0.143 | 1269 | 0.30 (−0.04 to 0.64) | 0.079 |
| Negative food markers | 647 | 5.54 (2.94) | 625 | 6.21 (3.28) | 1266 | −0.64 (−1.17 to −0.11) | 0.020 | 1272 | −0.61 (−1.25 to 0.03) | 0.061 |
| Positive food markers | 647 | 6.28 (2.55) | 625 | 5.87 (2.52) | 1266 | 0.27 (−0.18 to 0.73) | 0.231 | 1272 | 0.39 (−0.09 to 0.86) | 0.104 |
| Weekend FIQ | | | | | | | | | | |
| Energy-dense snacks | 647 | 4.17 (2.21) | 626 | 4.26 (2.35) | 1269 | −0.10 (−0.43 to 0.24) | 0.556 | 1273 | 0.01 (−0.37 to 0.38) | 0.977 |
| Healthy snacks | 639 | 3.42 (1.83) | 620 | 3.12 (1.73) | 1242 | 0.23 (−0.04 to 0.50) | 0.086 | 1259 | 0.31 (0.02 to 0.59) | 0.035 |
| Negative food markers | 648 | 6.79 (3.24) | 626 | 6.82 (3.36) | 1270 | −0.07 (−0.56 to 0.42) | 0.770 | 1274 | 0.12 (−0.44 to 0.68) | 0.671 |
| Positive food markers | 648 | 6.00 (2.66) | 626 | 5.52 (2.64) | 1264 | 0.36 (−0.13 to 0.84) | 0.145 | 1274 | 0.50 (0.05 to 0.95) | 0.030 |

There was no significant effect of the intervention on any of the physical activity outcomes (see *Table 13*). Measures of physical activity, average weekly volume and average time per day in each intensity category were all fractionally larger in the intervention group from a direct comparison of the mean values of the two groups and from both the unadjusted and adjusted analyses. Conversely, the accelerometry data classified as sedentary behaviour (between 6 a.m. and 10 p.m.) were, on average, lower in the intervention group. However, none of the observed differences was statistically significant.

The mean scores of the FIQ outcomes averaged over the week were higher in the intervention group for healthy snacks and positive food markers and lower for the energy-dense snacks and negative food markers (see *Table 14*). These between-group differences were statistically significant in the unadjusted analysis for the healthy snacks and positive food markers. However, once the analysis was adjusted for baseline score, stratification variables and other potential confounders, these differences were no longer statistically significant, while the between-group differences for energy-dense snacks and negative food marker scores were statistically significant in the adjusted analyses, with the strongest evidence of an intervention effect observed for energy-dense snacks ($p = 0.017$). According to the adjusted analysis, the mean average weekly scores for the intervention group were 0.37 (95% CI –0.66 to –0.07) lower than for the control group for the energy-dense snacks score and 0.47 (95% CI –0.91 to –0.02) lower for the negative food markers score.

The pattern of differences observed for the weekly averaged FIQ outcomes was repeated for the weekday scores. There was weak evidence of between-group differences from the unadjusted analyses for the energy-dense snacks and negative food marker scores, as well as for the healthy snacks score. After adjustment for the prespecified variables there was evidence of statistically significant differences for the energy-dense snacks ($p = 0.013$) and negative food markers ($p = 0.020$). The intervention effect sizes were of a similar magnitude and direction for the two unhealthy food scores, with the mean average weekday energy-dense snacks score being 0.47 (95% CI –0.84 to –0.11) lower in the intervention group than in the control group, and 0.64 (95% CI –1.17 to –0.11) lower for the negative food markers score.

At the weekend, the pattern of the differences between the two groups was the same, with higher mean scores for healthy snacks and positive food markers and lower mean scores for energy-dense snacks and negative food markers for the intervention group than for the control group (see *Table 14*). The between-group differences for the healthy snacks and positive food markers scores were statistically significant in the unadjusted analyses ($p = 0.035$ and 0.030, respectively); however, after adjustment for the prespecified variables, neither between-difference was statistically significant ($p = 0.086$ and 0.145, respectively).

### Model diagnostics for the primary analysis of the secondary outcomes

Following the fitting of the full random-effects models, plots of the residuals and best linear unbiased predictors for the random school intercept revealed no systematic deviation from the assumed distribution of the observations, required for linear mixed modelling, for all the secondary outcomes, with the exception of the FIQ scores when the weekday and weekend scores were modelled separately. Given these model violations, the weekday and weekend FIQ scores were reanalysed as ordinal outcomes, after aggregating the sparser larger scores into a single category to leave five levels for each score. The results from the ordinal models are shown in *Table 15*. In summary, from the adjusted analyses, there remained evidence of statistically significant intervention effects for the weekday energy-dense snacks (OR 0.68, 95% CI 0.51 to 0.90) and negative food markers scores (OR 0.69, 95% CI 0.52 to 0.92), indicating that children in the intervention group were more likely to have lower scores than children in the control group for either of the weekday negative food behaviour categories. The OR for the intervention effect for the weekend healthy snacks score was 1.27 (95% CI 1.00 to 1.61), indicating that children in the intervention group may be more likely than children in the control group to have higher healthy snacks scores at the weekend, although the 95% CI does include the null value of one.

**TABLE 15** Primary analysis, both adjusted and unadjusted, of weekday and weekend FIQ scores, modelled as ordinal outcomes, at 18 months post baseline

| | Group | | | | Analysis | | | | | |
| | Intervention | | Control | | Adjusted | | | Unadjusted | | |
| Outcome | *n* | Mean (SD) | *n* | Mean (SD) | Total *n* in analysis | OR (95% CI) | *p*-value | Total *n* in analysis | OR (95% CI) | *p*-value |
|---|---|---|---|---|---|---|---|---|---|---|
| Weekday FIQ (5-level ordinal outcomes) | | | | | | | | | | |
| Energy-dense snacks | 647 | 2.24 (1.40) | 625 | 2.50 (1.43) | 1265 | 0.68 (0.51 to 0.90) | 0.007 | 1272 | 0.72 (0.54 to 0.98) | 0.034 |
| Healthy snacks | 645 | 2.46 (1.35) | 624 | 2.25 (1.35) | 1257 | 1.29 (0.97 to 1.70) | 0.077 | 1269 | 1.36 (1.00 to 1.86) | 0.052 |
| Negative food markers | 647 | 2.11 (1.35) | 625 | 2.39 (1.37) | 1266 | 0.69 (0.52 to 0.92) | 0.012 | 1272 | 0.70 (0.50 to 0.97) | 0.033 |
| Positive food markers | 647 | 2.35 (1.37) | 625 | 2.15 (1.35) | 1266 | 1.29 (0.98 to 1.69) | 0.072 | 1272 | 1.33 (0.99 to 1.79) | 0.057 |
| Weekend FIQ (5-level ordinal outcomes) | | | | | | | | | | |
| Energy-dense snacks | 647 | 2.29 (1.27) | 626 | 2.36 (1.28) | 1269 | 0.89 (0.69 to 1.16) | 0.395 | 1273 | 0.97 (0.73 to 1.28) | 0.816 |
| Healthy snacks | 639 | 2.22 (1.39) | 620 | 2.01 (1.38) | 1242 | 1.27 (1.00 to 1.61) | 0.047 | 1259 | 1.32 (1.01 to 1.72) | 0.043 |
| Negative food markers | 648 | 2.18 (1.41) | 626 | 2.20 (1.43) | 1270 | 0.97 (0.74 to 1.26) | 0.810 | 1274 | 1.05 (0.78 to 1.41) | 0.739 |
| Positive food markers | 648 | 2.29 (1.45) | 626 | 2.07 (1.45) | 1264 | 1.27 (0.94 to 1.72) | 0.124 | 1274 | 1.36 (1.02 to 1.80) | 0.034 |

**Notes**
Number and percentage of children within FIQ categories for each randomised group and estimate of the effect of the intervention (ORs, 95% CIs and *p*-values) for analyses with and without adjustment for stratification variables, cohort, baseline measure and gender.

As with the analyses of the weight status categories, the assumption of proportional odds was tested through fitting of the generalised ordinal model, with robust error variances, clustered on schools. While the assumption of proportional odds was violated for some of the adjusted factors, the intervention effects from these models were similar to the presented results from the random-effects models, differing in their coefficients by $< \pm 0.05$. The test of the Peterson–Harrell parameters indicated marginally weak evidence of a violation of the proportional odds assumption for the intervention effect in the model for weekday energy-dense snacks ($p = 0.049$). This was observed to be due to a different, larger effect in the sparser high energy-dense snack score categories.

## Intraclass correlation coefficients

*Table 16* shows the ICC and 95% CIs for each of the outcomes from the random-effects models. The estimated ICC for the primary outcome was 0.014 (95% CI 0.003 to 0.069), slightly lower than the estimate of 0.02 used in the sample size calculation. The estimated ICC for weight status when dichotomised (overweight and obese vs. healthy and underweight) was extremely small (estimated as 0.000). This was due to the estimated between-cluster variance for this outcome being particularly small relative to the variance of the standard logistic distribution, assumed to be that of the latent continuous distribution that determines the weight of status of individuals. The wide CI for this ICC is an artefact of the estimated parameter being so close to the boundary of the parameter space, and the logit function,[56] which yielded a lower bound very close to zero and an upper bound at one. The estimated ICCs for waist circumference and waist circumference SDS were higher than for the other anthropometric measures.

## Complier average causal effect analysis

The original analysis plan included a complier average causal effect (CACE) analysis to estimate the CACE of treatment, as a potentially unbiased estimate of receiving HeLP, with non-compliers defined as children who did not receive at least four sessions of drama and the one-to-one goal-setting session. The TSC subsequently deemed the planned analysis unnecessary, given the small number of children who 'switched' between allocated treatment groups: there were only 24 out of 675 children allocated to the intervention group who were categorised as non-compliers (3.6%) and 4 of these 24 did not provide primary outcome data.

**TABLE 16** Intraclass correlation coefficients from the random-effects models for the outcomes at 18 and 24 months post baseline

| Outcome | ICC (95% CI) |
|---|---|
| ***Primary outcome at 24 months*** | |
| BMI SDS | 0.014 (0.003 to 0.069) |
| ***Secondary outcomes at 24 months*** | |
| BMI | 0.011 (0.001 to 0.074) |
| Waist circumference (cm) | 0.093 (0.049 to 0.169) |
| Weight status[a] (overweight/obese vs. healthy/underweight) | 0.000 (0.000 to 1.000) |
| Waist circumference SDS | 0.116 (0.064 to 0.201) |
| % body fat | 0.021 (0.005 to 0.078) |
| Body fat SDS | 0.023 (0.007 to 0.075) |
| Body fat SDS[b] | 0.017 (0.003 to 0.078) |

**TABLE 16** Intraclass correlation coefficients from the random-effects models for the outcomes at 18 and 24 months post baseline (*continued*)

| Outcome | ICC (95% CI) |
|---|---|
| **Secondary outcomes at 18 months** <br> *Anthropometric* | |
| BMI | 0.008 (0.001 to 0.081) |
| BMI SDS | 0.013 (0.003 to 0.064) |
| Weight status[a] (overweight/obese vs. healthy/underweight) | 0.060 (0.011 to 0.262) |
| Waist circumference (cm) | 0.125 (0.069 to 0.215) |
| Waist circumference SDS | 0.178 (0.105 to 0.285) |
| % body fat | 0.019 (0.004 to 0.076) |
| Body fat SDS | 0.012 (0.002 to 0.076) |
| Body fat SDS[b] | 0.012 (0.002 to 0.080) |
| *Physical activity* | |
| Mean weekly volume | 0.053 (0.020 to 0.130) |
| Mean daily total activity | 0.084 (0.039 to 0.174) |
| Mean daily light | 0.095 (0.045 to 0.189) |
| Mean daily moderate | 0.076 (0.034 to 0.163) |
| Mean daily moderate–vigorous | 0.052 (0.020 to 0.130) |
| Mean daily vigorous | 0.047 (0.017 to 0.123) |
| Mean daily sedentary (6 a.m. to 10 p.m.) | 0.085 (0.039 to 0.175) |
| *Food intake* | |
| Mean daily energy-dense snacks | 0.023 (0.008 to 0.069) |
| Mean daily healthy snacks | 0.032 (0.012 to 0.083) |
| Mean daily negative food markers | 0.026 (0.009 to 0.072) |
| Mean daily positive food markers | 0.031 (0.012 to 0.080) |

a Weight status categories are defined using the Public Health England definitions.[55]
b After excluding extreme body fat absolute SD values of $\geq 5$.

# Chapter 4 Economic evaluation

## Introduction

The aims of the economic analyses were to estimate the expected resource use and related costs associated with delivery of HeLP (i.e. the HeLP intervention) in a policy-relevant setting, and to estimate the cost-effectiveness of the HeLP intervention where it is shown to be an effective intervention compared with usual practice alone.

Given the effectiveness profile for the HeLP intervention, presented in *Chapter 3*, and the absence of statistically significant differences on the primary outcome measure of BMI SDS, we begin here with a summary statement on the cost-effectiveness of HeLP versus usual practice. We find that there is no expectation, using the framework for economic evaluation developed/described below, of improvements in the likely incidence of weight-related health events [e.g. coronary heart disease (CHD)], or related cost savings to the health and social care system, from the HeLP intervention. As described in the following sections, HeLP incurs additional costs for delivery of the intervention with no identified or expected health gains, using the analytical framework developed and presented here, and therefore usual practice 'dominates' HeLP in a resource allocation context (i.e. HeLP is not cost-effective compared with usual practice).

Although we highlight the results of the cost-effectiveness analysis at the outset in this chapter, for completeness, and to inform the reader and to inform future research, we provide detail in this chapter on the methods used, and the results of the cost analysis estimating the cost for delivery of HeLP. We also provide a narrative and detail on the development of a decision-analytic modelling framework developed to assess the cost-effectiveness of HeLP, using this to provide illustrative and exploratory cost-effectiveness analyses.

Below we summarise the methods used to estimate resource use and costs for delivery of HeLP, and the methods used to develop a decision-analytic modelling framework. We then present results for these two areas of research. When describing development of a modelling framework, we present summary details of a literature review of model-based economic evaluations that model childhood obesity interventions through adult years, and we describe the development of a modelling framework in this context, and the areas of evidence synthesis required to populate the model.

## Methods

### Estimating resource use and costs for delivery of the HeLP intervention

In summary, the HeLP intervention has 24 components, with activities delivered as either school level (i.e. delivered once regardless of school size) ($n = 5$) or class level ($n = 19$) (*Table 17*). Key to the intervention are the drama activities, made up of eight components. Further details on the HeLP intervention and usual care comparator are given in *Chapter 2*.

The research questions here are:

- What is the estimated resource use and cost associated with delivery of the HeLP intervention?
- What is the estimated mean cost of HeLP per participant? (Per school? Per class? Per cohort of schools?)

**TABLE 17** Summary of the components of the HeLP intervention

| Phase | Activity type | School- or class-level activity | Drama (Y/N) |
|---|---|---|---|
| 1 | Whole school meeting | School | N |
| 1 | Whole class assembly | School | N |
| 1 | 2 × activity workshops | Class | N |
| 1 | Rap poetry | Class | N |
| 1 | Newsletters | School | N |
| 1 | Parents' assembly | School | N |
| 2 | 5 × PSHE lessons | Class | N |
| 2 | 5 × drama workshops | Class | Y |
| 3 | Goal setting #1 | Class | N |
| 3 | Self-monitoring | Class | N |
| 3 | Parents' forum | School | Y |
| 4 | 1 × drama workshop | Class | Y |
| 4 | 1 × class assembly | Class | Y |
| 4 | 1 × PSHE lesson | Class | N |
| 4 | Goal setting #2 | Class | N |

N, no; PSHE, personal, social, and health education; Y, yes.

Informed through prior development research, and through conduct of the earlier exploratory RCT of HeLP,[27,28] the delivery of the HeLP intervention involves resource use and cost primarily for the staff inputs required during delivery (contact and non-contact activities), for HeLP co-ordinators, the drama co-ordinator, actors and activity/workshop co-ordinators. Teaching staff are present for some activities and take an active role in others, but based on findings from prior pilot research their time input is not considered additional/incremental input over and above their expected teaching role. Therefore, costs associated with teaching input are not included in the base-case cost analysis; time inputs are described, and these are included in a sensitivity analysis.

Additional resource use and costs are incurred in the training of the delivery staff and for materials/consumables used to support the delivery of the intervention.

Data on resource use for delivery of the HeLP intervention were collected within the trial through completion of contact sheets by HeLP co-ordinators at every contact, by component of HeLP, describing who was involved (staff type) in the contact (component of HeLP) and the time inputs for each person involved, including time associated with planning/preparation and related travel time. Given the importance of the drama components in the HeLP intervention, alongside the data collected by HeLP co-ordinator contact sheets, the drama co-ordinator provided details of the estimated resource use associated with the delivery of drama across the HeLP intervention.

Estimates for resource use associated with the training required for those staff engaged in delivery of the HeLP intervention were provided by the HeLP study co-ordinator based on their experiences in the running of the trial. Estimates of materials/consumables needed to support delivery of the HeLP intervention are based on within-trial requirements and projected estimates for future delivery.

Estimates of resource use by unit (e.g. time/hours) are combined with published national unit costs, or cost estimates, to estimate the mean costs for delivery of HeLP by school and by class, and to estimate the mean participant-level costs for delivery of the HeLP intervention. Costs are presented in 2015 UK pounds sterling. Unit costs for staff inputs are predominantly taken from those published *Unit Costs for Health and Social Care* staff grades reported by Curtis and Burns.[57] *Table 18* reports the units costs used in the analyses, the source of the unit cost and/or the assumptions used in allocating unit costs to estimates of resource use. Cost estimates were determined in a 'bottom-up' approach, based on detailed micro-level resource use data collected within the trial.

**TABLE 18** Unit costs (GBP) used to estimate cost of delivery of HeLP intervention

| Resource | Unit cost (£) | Source |
|---|---|---|
| HeLP co-ordinator | 41.29 | Salary based on median FTE band 6 (£31,914) Agenda for Change health staff 2015, working 37.5 hours per week/41.7 weeks per annum[a] |
| Drama facilitator/HeLP co-ordinator | 41.29 | Salary based on median FTE band 6 (£31,943) Agenda for Change health staff 2014, working 37.5 hours per week/42 weeks per annum[a] |
| Actors[b] | 11.20 | Salary based on median FTE band 4 (£13,798) Agenda for Change health staff 2014, working 37 hours per week/41.9 weeks per annum. Calculation of unit cost per hour based on cost structure reported in Curtis and Burns[57] (excludes overheads) [e.g. Section 11.6 (p. 192)] |
| Actors *(sensitivity analysis)* | 25.34 | Salary based on median FTE band 5 (£25,764) Agenda for Change health staff 2014, working 37.5 hours pe week/42 weeks per annum. This band was chosen given rough equivalence to the Equity rate for actors of £440 per week. Calculation of unit cost per hour based on cost structure reported in Curtis and Burns[57] (excludes qualifications, non-staff costs and capital costs) [e.g. Section 10.2 (p. 171)] |
| Daily travel allowance (actors) | 19.00 | Per-day allowance based on the Equity weekly rate for travel outside London of £95.00 (5 days) and number of actors per session |
| Teaching staff *(sensitivity analysis)* | 21.00 | Based on supply teacher rate of £160 per day;[58] 37.5 hours (this assumption follows the length of day for other job types, i.e. 37.5 hours, although what this rate consists of or is based on is unclear) |
| Activity co-ordinator | 20.00 | Allowance of £20 per hour, including travel costs, has been made for activity co-ordinators (assumption based on within-trial experience) |
| Lead trainer | 41.29 | Salary based on median FTE band 6 Agenda for Change health staff 2014, working 37.5 hours per week/41.7 weeks per annum. Calculation of unit cost per hour based on cost structure reported in Curtis and Burns[57] (excludes qualifications and capital costs) |
| Training | – | See *Appendix 7* |
| Other costs | – | Consumables (see below) |

FTE, full-time equivalent; GBP, Great British pounds.
a  Calculation of unit cost per hour based on cost structure reported in Curtis and Burns[57] (excludes qualifications and capital costs) (e.g. Sections 10.1, 10.3, 10.4; pp. 169, 171–72). This unit cost includes costs associated with management and travel/transport.
b  Based on the delivery of HeLP in the RCT, when actors were relatively inexperienced or in training, we have applied an hourly rate at the level of the minimum wage, approximately £7 per hour, to the actors. A higher rate, as above, is applied in sensitivity analyses.

'Other' resources/costs involved with the delivery of the drama intervention included:

- costumes (drama) – it was assumed that three sets of drama costumes at £80 per set would be required (based on £20 per character)
- materials (equipment including fizzy drinks, water and laminated cards) – it was assumed that three sets of materials would be required at a cost of £150 per set
- public liability – annual cost of £250 (based on guidance from drama provider).

These other additional costs (totalling £940 across the cohort of 27 classes) were assumed to be for a 12-month period and were distributed across the cohort in base-case analyses.

### Development of modelling framework (Exeter Obesity Model) to estimate the cost-effectiveness of the HeLP intervention versus usual practice

As set out in the prespecified economic analysis plan,[43] the framework for estimating the cost-effectiveness of the HeLP intervention is based on development of and subsequent use of a decision-analytic model to predict the future costs and benefits associated with an expected between-group difference in the HeLP RCT primary outcome measure of BMI SDS. A two-stage economic model has been developed, described in more detail below, to predict future adult weight status, from weight status profiles at 24-month follow-up in the HeLP RCT (BMI SDS) at age 11–12 years, and thereafter, in stage 2, to predict a profile of future weight-related health events (e.g. T2DM or CHD) as a function of the predicted adult weight status profiles (i.e. for treatment cohort participants vs. controls). The aim of the modelling framework – the Exeter Obesity Model – was to capture the difference in costs and outcomes, over time in adult years, associated with weight status profiles at 24-month follow-up for treatment and control participants in the HeLP RCT at age 11–12 years.

A decision-analytic model was developed based on a review of published models in this area (i.e. modelling impacts of childhood obesity interventions through adult years) and discussions across stakeholder groups (see detail of the PAG in *Chapter 2*). The model was informed and populated based on literature review for parameter inputs, and against good practice modelling guidance.[59,60] The aim of the model-based framework was to estimate the cost-effectiveness of the HeLP intervention, assumed at this stage to be a relatively low-cost public health intervention, versus usual practice. The setting is a UK public health setting, involving a school-based intervention as a means of having an impact on future adult weight status and health status through reduced incidence of adverse health events. The decision-analytic context is the question on the cost-effectiveness of the HeLP intervention versus usual practice, and the modelling framework was not intended to be an accurate prediction of the life experiences of children by weight status through adult years. The development of a model-based framework was set out conceptually in a prespecified economic analysis plan.[43] A cost-effectiveness analysis was expected to synthesis data on predicted mean incremental costs and benefits, taking a primary perspective as that of the NHS and Personal Social Services (i.e. third-party payer), consistent with the methods recommended by the National Institute for Health and Care Excellence (NICE).[61] The primary economic end point was cost per quality-adjusted life-year (QALY), with a cost-effectiveness analysis also to be presented against cost per life-year gained and cost per weight-related event avoided.

## Results

### Estimating the resource use and cost of the HeLP intervention

The HeLP intervention was delivered to children in 16 schools, reflecting the delivery of the intervention to 27 classes (*Table 19*). Within-trial data on resource use for delivery were from 719 HeLP co-ordinator contact sheets, representing data on 94% of expected contacts with schools and accounting for school- and class-level activities.

**TABLE 19** Summary of school level characteristics for delivery of HeLP

| School code | Cohort | Pupils (baseline), (*n*) | Classes, (*n*) | Classes in analysis, (*n*) |
|---|---|---|---|---|
| 23 | 1 | 26 | 1 | 1 |
| 2 | 1 | 33 | 1 | 1 |
| 5 | 1 | 27 | 3 | 1[a] |
| 9 | 1 | 25 | 1 | 1 |
| 15 | 1 | 52 | 1.5 | 2[b] |
| 17 | 1 | 29 | 1 | 1 |
| 10 | 1 | 20 | 1 | 1 |
| 21 | 1 | 42 | 2 | 2 |
| 32 | 2 | 23 | 1 | 1 |
| 19 | 2 | 48 | 2 | 2 |
| 6 | 2 | 84 | 3 | 3 |
| 28 | 2 | 66 | 3 | 3 |
| 30 | 2 | 37 | 1.5 | 2[c] |
| 29 | 2 | 31 | 1 | 1 |
| 12 | 2 | 86 | 3 | 3 |
| 33 | 2 | 47 | 2 | 2 |
| Total: 16 | – | 676 | 28 | 27 |
| Mean | – | 24.14 | 1 | – |
| Mean | – | 25.01 | – | NA |

NA, not applicable.
a　School 5: three classes, combined for the delivery of the HeLP intervention (each activity delivered as though one class).
b　School 15: treated as two classes (class-level activities delivered × 2).
c　School 30: differences across Years 5 and 6, but assumed to be delivered × 2 for class-level activities.

Summary data on aggregate resource use across all 27 classes, by staff type and with time (hours) by type of time input (preparation, task and travel), are presented in *Table 20*. Data by staff type and by school-class configuration are presented in *Table 21*, showing economies of scale as school size increases. *Table 22* reports the estimated total costs for delivery of HeLP across all 16 schools (and 27 classes). Overall, it is estimated to cost approximately £144,749 for delivery of HeLP to the 16 schools, with 27 classes, giving a mean estimated cost per child at £214 (*n* = 676). Drama staff costs make up 41% of the estimated delivery costs, with HeLP co-ordinator costs accounting for 37% of the overall estimated costs for delivery (when combined, these primary cost components are 78% of the costs for delivery of HeLP).

Estimates of the cost for training required for staff involved in delivery of HeLP, and other costs, are based on expected requirements for delivery across a cohort of 16 schools (27 classes), with training for four HeLP co-ordinators (one trainer), two drama facilitators and eight (two teams) actors. The details of the cost estimates of training are provided in *Appendix 7*.

Estimated delivery costs for HeLP by school-class configuration are presented in *Table 23*, providing an estimate by school size (i.e. schools with one, two or three classes by year group). The estimated mean costs by class, and by child per class, show a mean cost per class (per child) that reduces as the school size increases, from £5724 (£229) for a school with one class to £15,390 (£205) for a school with three classes.

**TABLE 20** Resource use by staff type by type of contact (preparation, task and travel): total across 16 schools, 27 classes

| Delivery requirement | Type of staff hours (mean) | | | | |
| --- | --- | --- | --- | --- | --- |
| | HeLP co-ordinator | Teaching staff | Activity co-ordinator | Drama facilitator | Actor |
| Preparation time | | | | | |
| Class-level activities | 228.58 | 38.51 | 34.70 | 170.66 | 357.50 |
| School-level activities | 95.38 | 57.90 | 0.00 | 13.60 | 40.27 |
| Total | 323.96 | 96.41 | 34.70 | 184.26 | 397.77 |
| Task time | | | | | |
| Class-level activities | 536.60 | 278.41 | 153.00 | 379.16 | 1412.32 |
| School-level activities | 53.90 | 33.42 | 0.00 | 23.56 | 67.91 |
| Total | 590.50 | 311.83 | 153.00 | 402.71 | 1480.23 |
| Travel time | | | | | |
| Class-level activities | 304.81 | NA | NA | 309.37 | NA |
| School-level activities | 79.56 | NA | NA | 27.73 | NA |
| Total | 384.38 | NA | NA | 337.10 | NA |
| Total time (including travel) | 1298.83 | 408.24 | NA | 924.07 | NA |
| Total time (excluding travel) | 914.46 | 408.24 | 187.70 | 586.97 | 1877.99 |

NA, not applicable.
**Notes**
(1) No distinction has been made between contact and non-contact time in costing the intervention. (2) Although travel time was recorded in the trial, no travel time has been included in the costing for actors. Instead, daily travel expenses have been applied. (3) Activity co-ordinator time assumes two members of staff. (4) Actor time assumes (based on data collected) four actors in each of the five workshops in phase 2 and three actors for the remaining three drama activities.

However, it is expected that the provision and delivery of HeLP would be at the level of geographical region, involving a mix of school sizes.

## Sensitivity analysis

The following sensitivity analyses were conducted:

- sensitivity analysis 1 – input; actors' cost per hour increased (from £11.20 to £25.34)
- sensitivity analysis 2 – input; inclusion of teacher time/costs.

### Sensitivity analysis 1: actors unit costs uprated to £25.34 per hour (from £11.20 per hour)

The sensitivity analysis was conducted using a unit cost of £25.34 per hour for actor time inputs. This hourly rate is based on guidance from the drama provider for the potential salary level for professional actors, but is aligned here with salary structure for a similarly paid health-care worker using the Agenda for Change band 5 for health staff[57] (see *Table 20*) and unit costs from Curtis and Burns.[57]

In this scenario, the total costs for the delivery of HeLP increase markedly, from £214.13 per participant to £258.43 per participant, based on the aggregate costing framework (*Table 24*), and the total (mean) costs increase by school-class scenario from £5724 (£229) and £15,390 (£205) for schools with one class and three classes, respectively, to £6782 (£274) and £18,643 (£248).

**TABLE 21** Resource use by school–class configuration: resource use by staff type and by type of contact (preparation, task and travel)

| Time allocation | Type of staff | | | | | | | | | | | | | | |
| | HeLP co-ordinator | | | Teaching staff | | | Activity co-ordinator | | | Drama facilitator/HeLP co-ordinator | | | Actors | | |
| Number of schools/number of classes | 1/1 | 1/2 | 1/3 | 1/1 | 1/2 | 1/3 | 1/1 | 1/2 | 1/3 | 1/1 | 1/2 | 1/3 | 1/1 | 1/2 | 1/3 |
| Preparation time (hours) | | | | | | | | | | | | | | | |
|     Class | 8.47 | 16.93 | 25.40 | 1.43 | 2.85 | 4.28 | 1.29 | 2.57 | 3.86 | 6.32 | 12.64 | 18.96 | 13.24 | 26.48 | 39.72 |
|     School | 5.96 | 5.96 | 5.96 | 3.62 | 3.62 | 3.62 | 0.00 | 0.00 | 0.00 | 0.85 | 0.85 | 0.85 | 2.52 | 2.52 | 2.52 |
| Total | 14.43 | 22.89 | 31.36 | 5.05 | 6.47 | 7.90 | 1.29 | 2.57 | 3.86 | 7.17 | 13.49 | 19.81 | 15.76 | 29.00 | 42.24 |
| Task time (hours) | | | | | | | | | | | | | | | |
|     Class | 19.87 | 39.75 | 59.62 | 10.31 | 20.62 | 30.93 | 5.67 | 11.33 | 17.00 | 14.04 | 28.09 | 42.13 | 52.31 | 104.6 | 156.9 |
|     School | 3.37 | 3.37 | 3.37 | 2.09 | 2.09 | 2.09 | 0.00 | 0.00 | 0.00 | 1.47 | 1.47 | 1.47 | 4.24 | 4.24 | 4.24 |
| Total | 23.24 | 43.12 | 62.99 | 12.40 | 22.71 | 33.02 | 5.67 | 11.33 | 17.00 | 15.52 | 29.56 | 43.60 | 56.55 | 108.8 | 161.1 |
| Travel time (hours) | | | | | | | | | | | | | | | |
|     Class | 11.29 | 22.58 | 33.87 | NA | NA | NA | NA | NA | NA | 11.46 | 22.92 | 34.37 | NA | NA | NA |
|     School | 4.97 | 4.97 | 4.97 | NA | NA | NA | NA | NA | NA | 1.73 | 1.73 | 1.73 | NA | NA | NA |
| Total (hours) | 16.26 | 27.55 | 38.84 | NA | NA | NA | NA | NA | NA | 13.19 | 24.65 | 36.11 | NA | NA | NA |
| Total time (including travel) (hours) | 53.93 | 93.56 | 133.1 | NA | NA | NA | NA | NA | NA | 35.88 | 67.70 | 71.43 | NA | NA | NA |
| Total time (excluding travel) (hours) | 37.67 | 66.01 | 94.35 | 17.45 | 29.18 | 40.92 | 6.95 | 13.90 | 20.86 | 22.69 | 43.05 | 35.33 | 72.31 | 137.8 | 203.4 |

NA, not applicable.
**Notes**
(1) No distinction has been made between contact and non-contact time in costing the intervention. (2) Although travel time was recorded in the trial, no travel time has been included in the costing for actors. Instead, daily travel expenses have been applied. (3) Activity co-ordinator time assumes two members of staff. (4) Actor time assumes (based on data collected) four actors in each of the five workshops in phase 2 and three actors for the remaining three drama activities.

**TABLE 22** Estimated cost (GBP) for delivery of HeLP (total cost across 16 schools, 27 classes)

| Delivery requirement | Hours (mean) | Unit cost (£) | Total cost (£) |
|---|---|---|---|
| HeLP co-ordinator[a] | 1298.83 | 41.29[b] | 53,633.58 |
| Activity co-ordinator[a] | 187.70 | 20[b] | 3754.04 |
| HeLP drama | | | |
| Drama facilitator/HeLP co-ordinator[a] | 924.07 | 41.29[b] | 38,158.38 |
| Actors[c] | 1878 | 11.20[b] | 21,034.32 |
| Travel allowance (actors) | – | 19.00[d] | 14,250.00 |
| Other costs | – | – | 940.00 |
| Training | | | |
| Training HeLP co-ordinators[e] | – | – | 5145.54 |
| Training drama[f] | – | – | 7833.64 |
| Total cost | – | – | 144,749.50 |
| Cost per participant | – | – | 214.13 |

GBP, Great British pounds.
a Includes contact and non-contact (preparation and travel time).
b Per hour.
c Includes contact and non-contact (preparation time).
d Per day.
e Based on two HeLP co-ordinators.
f Based on two drama teams.

**TABLE 23** Estimated total cost (GBP) for delivery of HeLP by school-class configuration: one school and one, two and three classes

| Staff type | Unit cost (£) | One school, one class Hours (mean) | Cost (£) | One school, two classes Hours (mean) | Cost (£) | One school, three classes Hours (mean) | Cost (£) |
|---|---|---|---|---|---|---|---|
| HeLP co-ordinator[a] | 41.29 | 53.93 | 2227.05 | 93.56 | 3863.49 | 133.19 | 5499.93 |
| Activity co-ordinator[a] | 20.00 | 6.95 | 139.04 | 13.90 | 278.08 | 20.86 | 417.12 |
| HeLP drama | | | | | | | |
| Drama facilitator/HeLP co-ordinator[a] | 41.29 | 35.88 | 1481.50 | 67.70 | 2795.53 | 99.52 | 4109.57 |
| Actors[b] | 11.20 | 72.31 | 809.90 | 137.86 | 1544.07 | 203.41 | 2278.25 |
| Travel allowance (actors) | 19.00 | – | 551.00 | – | 1045.00 | – | 1539.00 |
| Other costs | 940 | – | 34.81 | – | 69.63 | – | 104.44 |
| Training | | | | | | | |
| Training HeLP co-ordinators[c] | 5145.54 | – | 190.58 | – | 381.15 | – | 571.73 |
| Training drama[d] | 7833.64 | – | 290.13 | – | 580.27 | – | 870.40 |
| Total cost | – | – | 5724.01 | – | 10,557.22 | – | 15,390.43 |
| Cost per participant[e] | – | – | 228.62 | – | 210.83 | – | 204.90 |

GBP, Great British pounds.
a Includes contact and non-contact (preparation and travel time).
b Includes contact and non-contact (preparation time).
c Based on two HeLP co-ordinators.
d Based on two drama teams.
e Based on mean class size from the HeLP study (25.04).

**TABLE 24** Sensitivity analysis 1: estimated cost (GBP) for delivery of HeLP (total cost across 16 schools, 27 classes) using a higher unit cost (£25.34) for actor time

| Delivery requirement | Hours (mean) | Unit cost (£) | Total cost (£) |
|---|---|---|---|
| HeLP co-ordinator[a] | 1298.83 | 41.29[b] | 53,633.58 |
| Activity co-ordinator[a] | 187.70 | 20[b] | 3754.04 |
| HeLP drama | | | |
| Drama facilitator/HeLP co-ordinator[a] | 924.07 | 41.29[b] | 38,158.38 |
| Actors[c] | 1878 | 25.34[b] | 47,588.39 |
| Travel allowance (actors) | – | 19.00[d] | 14,250.00 |
| Other costs | – | – | 940.00 |
| Training | | | |
| Training HeLP co-ordinators[e] | – | – | 5145.54 |
| Training drama[f] | – | – | 11,227.07 |
| Total cost | – | – | 174,697.07 |
| Cost per participant | – | – | 258.43 |

GBP, Great British pounds.
a  Includes contact and non-contact (preparation and travel time).
b  Per hour.
c  Includes contact and non-contact (preparation time).
d  Per day.
e  Based on two HeLP co-ordinators.
f  Based on two drama teams.

### Sensitivity analysis 2: inclusion of teacher time/costs, added to base-case estimates

Based on prior consideration and pilot research, it is not expected that teacher time inputs around HeLP activities are additional/incremental to usual activities (i.e. usual teaching practice). However, we have explored this in sensitivity analysis, and we have estimated costs here including teacher time with a unit cost of £21 per hour (£160 per day) for teacher time (assuming any additional time inputs will be covered using a supply teacher). This hourly rate is based on published guidance from the National Association of Schoolmasters Union of Women Teachers[58] on supply teacher rates of payment (see *Table 18*).

In this scenario, the total costs for the delivery of HeLP increase modestly, by < £13 per child, from £214.13 per participant to £227 per participant, based on the aggregate costing framework (*Table 25*), and the total (mean) costs increase by school-class scenario from £5724 (£229) and £15,390 (£205) for schools with one class and three classes, respectively, to £6090 (£243) and £16,249 (£216).

### A framework for the economic evaluation of HeLP

Here we describe the development of the Exeter Obesity Model, a decision-analytic model-based framework for assessment of the cost-effectiveness of the HeLP school-based intervention versus usual practice in children aged 9–10 years. The development of the model was based on a review of model-based economic evaluations of interventions related to childhood obesity, with a focus on models tracking from childhood through to adult years, and a summary of this systematic literature search is presented below. Thereafter we document the conceptual structure and the operational development of the model, describing data sources and the populating of the model to address the primary research question of whether or not the HeLP intervention could be considered cost-effective compared with usual practice.

TABLE 25 Sensitivity analysis 2: estimated cost (GBP) for delivery of HeLP (total cost across 16 schools, 27 classes) including costs for teacher time

| Delivery requirement | Hours (mean) | Unit cost (£) | Total cost (£) |
|---|---|---|---|
| HeLP co-ordinator[a] | 1298.83 | 41.29[b] | 53,633.58 |
| Teacher | 408.24 | 21.00[b] | 8572.94 |
| Activity co-ordinator[a] | 187.70 | 20.00[b] | 3754.04 |
| HeLP drama | | | |
| Drama facilitator/HeLP co-ordinator[a] | 924.07 | 41.29[b] | 38,158.38 |
| Actors[c] | 1878 | 11.20[b] | 21,034.32 |
| Travel allowance (actors) | – | 19.00[d] | 14,250.00 |
| Other costs | – | – | 940.00 |
| Training | | | |
| Training HeLP co-ordinators[e] | – | – | 5145.54 |
| Training drama[f] | – | – | 7833.64 |
| Total cost | – | – | 153,322.44 |
| Cost per participant | – | – | 226.81 |

GBP, Great British pounds.
a Includes contact and non-contact (preparation and travel time).
b Per hour.
c Includes contact and non-contact (preparation time).
d Per day.
e Based on two HeLP co-ordinators.
f Based on two drama teams.

## Review of decision-analytic models for use in cost-effectiveness analyses on obesity

A systematic review of model-based economic evaluations of interventions related to childhood obesity, as a means of preventing obesity in adults, was conducted in order to inform the structure and development of a modelling framework suitable for assessing the cost-effectiveness of the HeLP intervention.

## *Methods*

**Search strategy** Systematic searches were conducted in several databases (including MEDLINE, EMBASE, PsycINFO, NHS Economic Evaluation Database and Database of Abstracts of Reviews of Effects). The keywords and medical subject headings (MeSH) for obesity were combined with keywords for cost and economic analysis to capture relevant economic evaluations in all databases. The search strategy is detailed in *Appendix 8*. In addition, references within the included studies were reviewed to identify potentially overlooked studies. The searches were date limited 1990–2012 (original search). Language restrictions were not applied. Update searches were conducted in 2015.

**Eligibility criteria** Titles and abstracts for all articles identified by the literature search were screened. Studies were eligible for inclusion if they were economic evaluations (cost–benefit, cost–utility, cost-effectiveness and cost–consequence analyses) of interventions related to child/adolescent overweight/obesity as a means of preventing obesity in adults. Next, full texts of all articles deemed eligible from the previous stage were reviewed to further confirm their eligibility. At this stage, the exclusion criteria were burden-of-illness studies, cost analyses, study types other than economic evaluations, review articles, editorials or reports, and articles not written in English. Studies published as abstracts or conference presentations were included only if sufficient details were presented to allow both an appraisal of the methodology and an assessment of the results to be undertaken.

**Data extraction and synthesis**　For each study included in the final review, data were extracted into a table. The information obtained included authors, country, study year, comparators, population characteristics, study design, time horizon, perspective, data sources, effectiveness, cost, sensitivity analysis and study conclusions. The identified modelling studies are summarised, and the approaches available to model disease progression over time are discussed in a summary descriptive review.

## Results

**Studies identified**　A total of seven models (reported in 11 publications) were included in the final review.[62–72] The study identification process is summarised in *Figure 5*.

The study characteristics for the seven included models are summarised in *Table 26*.

**Summary of the literature identified on the modelling of the cost-effectiveness of interventions to address obesity in children**　The results from the literature review indicated that seven studies (reported in 11 publications) used model-based economic evaluations to assess the cost-effectiveness of interventions aimed at reducing obesity in children. Most of the models identified in the review were from North America,[62,63,67–72] with two European studies[64,65] (one being a multisite study across six countries in Europe[65]) and one from New Zealand.[66] The studies varied in terms of the modelling approach used, the range of health states (including weight-related events) included in the models, the techniques used to predict weight status in adulthood as a function of childhood weight status and the outcomes included in the models.
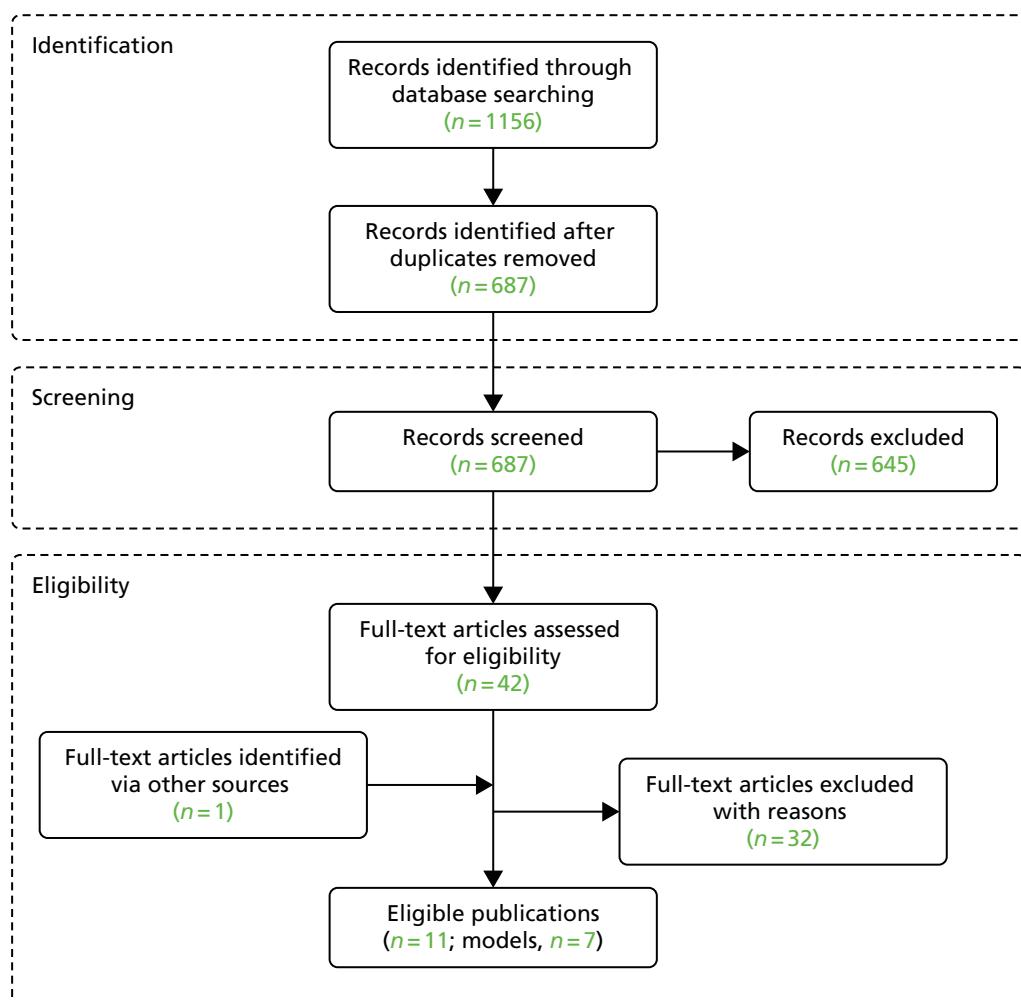


**FIGURE 5** The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

**TABLE 26** Summary characteristics of the identified models

| First author, year | Country setting | Intervention | Outcomes | Analytical approach, WS | Health events | Intervention effect | Time horizon/cycle length | Perspective |
|---|---|---|---|---|---|---|---|---|
| Wang et al., 2003[62] | USA | Planet Health | Cases of adulthood OW prevented; costs; QALYs; ICER | Extrapolation from childhood WS to adult (21–29 years) WS using Whitaker et al.[73] and from 21 to 29 years using NHANES EFS | Estimated 25-year direct health-care and medication costs[b] in women aged 40 years who maintain an OW status to age 65 years | Intervention effect expressed as probability of OW and at-risk OW at age 11 years | 25 years/NA | Societal perspective (health-care sector costs and costs related to productivity loss) |
| Brown et al., 2007[63] | USA | CATCH | Cases of adulthood OW prevented; costs; QALYs; ICER | Based on – same as – Wang et al.[62] | Estimated 25-year direct health-care and medication costs[b] in women aged 40 years who maintain an OW status to age 65 years | Intervention effect expressed as probability of OB and non-OB at age 14 years | 25 years/NA | Societal perspective (health-care sector costs and costs related to productivity loss) |
| Hollingworth et al., 2012[64] | UK | Lifestyle interventions | Costs; incidence of weight-related events; cost per life-year gained | Extrapolation over time estimated using a BMI growth equation simulated using Monte Carlo techniques | Events include CHD, stroke, diabetes, breast cancer, CRC | Intervention effect modelled as reduction in BMI SDS | Lifetime/1 year | Third-party payer (UK NHS) |
| Pil et al., 2014[65] | Six EU countries (Belgium, Bulgaria, Germany, Greece, Poland and Spain) | ToyBox | Costs; incidence of weight-related events; QALYs; ICER | Extrapolation from childhood WS to adult (30–34 years) WS using Venn et al.;[75] projection of the incidence of events (IASO) and mortality from age 30–34 years onwards (Markov model) | Event-free; CHD;[a] T2DM; stroke;[a] CRC;[a] breast cancer;[a] death<br><br>Suffering from > 1 condition not possible in the model | Intervention effect expressed as a relative risk reduction applied to country-specific probabilities of the weight categories in the adult general population | 70 years/1 year | Societal perspective (health-care sector costs and costs related to productivity loss) |
| Rush et al., 2014[66] | New Zealand | Project Energise | Costs; QALYs; ICER | Population (aged from 2 to 75 years categorised (HE, OW, OB) by fitting a log-normal distribution to mean BMI and SE reported in population survey data | Risk of 14 obesity-related conditions[c] using life tables, and data on RRs of disease incidence and mortality conditional on BMI | Intervention effect expressed as median difference in BMI between children receiving PE and historical control (not receiving PE) | Lifetime/NA | Third-party payer (health-care budget) |

| First author, year | Country setting | Intervention | Outcomes | Analytical approach, WS | Health events | Intervention effect | Time horizon/cycle length | Perspective |
|---|---|---|---|---|---|---|---|---|
| Tran et al., 2014[67] | Canada | APPLE | Costs; OW and OB cases prevented | Applied parameters from growth curve models to OW and OB prevalence rates of Alberta (APPLE vs. no APPLE) | Health events not measured. Estimated by multiplying the total direct health-care cost for OB by the proportion of OW/OB cases prevented by the intervention | Intervention effect expressed as the prevalence of obesity among children attending APPLE schools vs. those not attending APPLE schools | Lifetime/NA | Health-care perspective |
| Gortmaker et al., 2015[69] (multiple publications: Wright et al., 2015,[72] Sonneville et al., 2015,[71] Long et al., 2015,[70] Barrett et al., 2015[68]) | USA | SSB; TV AD; ECE; active PE | Cost per BMI change; costs; DALYs;[d] QALYs;[e] bet savings per dollar spent on an intervention[f] | Logic models link the intervention to behaviour change or change in energy levels and thus change in BMI | Excess health-care costs associated with obesity among children and adults | Evidence review; studies linking the intervention to behavioural change or shift in energy balance | 10 years or lifetime | Societal perspective |

Active PE, active physical education; APPLE, Alberta Project Promoting active Living and healthy Eating (in Schools); CATCH, Coordinated Approach to Child Health; CG, clinical guidelines; CRC, colorectal cancer; DALY, disability adjusted life-year; ECE, early care and education policy change; EFS, Epidemiologic Follow up Study; EU, European Union; HW, healthy weight; IASO, International Association of Obesity; ICER, incremental cost-effectiveness ratio; NHANES, National Health and Nutrition Examination Survey; OB, obese; OW, overweight; PE, Project Energise; RoW, rest of world; RR, relative risk; SE, standard error; SSB, sugar-sweetened beverage; TV AD, television advertising; WS, weight status.

a  The first year is a transitional state in which participants remain in for one cycle in this state, after which they can move to a follow-up state or die.
b  Medical costs estimated were those associated with events of fatal and non-fatal CHD, hypertension, diabetes, symptomatic gallstones, and osteoarthritis expected to occur during a 25-year period in each weight cohort.
c  Obesity-related conditions included chronic heart failure and T2DM.
d  DALY outcomes only reported for the SSB outcome as for the other three interventions participants will be aged < 30 years at follow-up and RRs of obesity-related diseases are 1.0 at the age of < 35 years.
e  QALY outcomes not reported for the ECE or Active PE interventions as few participants would be adults after 10 years.
f  Calculated for SSB, TV-AD and ECE interventions only.

Nine of the studies[64–72] were published in 2012 or later, and of note in the current report is that the initial searching of the literature in 2012 was used as the basis for the conceptual modelling framework for HeLP, as set out below, and it was, therefore, based on a limited literature (as described below and in *Table 26*). However, a number of the more recent studies adopt the methodological framework used in the earlier literature, which has influenced the conceptual framework used for the HeLP-specific research.

Before 2012 the evidence base for the assessment of the cost-effectiveness of interventions aimed at reducing obesity in children was small. Two models were identified.[62,63] Five models (reported in nine publications) were subsequently identified that were published between 2012 and 2015,[64–67,69] with the majority of these published in the last 2 years.

The earlier models[62,63] adopted a two-stage approach. The first stage used a decision tree structure and affixed a probability estimate that an individual (child) may change weight status between childhood and adult years. These predictions for adult weight status were based on longitudinal data from either Whitaker *et al.*[73] or NHANES I (National Health and Nutrition Examination Survey) (epidemiological study data).[74] The second stage of the model used a Markov-type approach, allocating the cohort in adult years to states by weight status (overweight, not overweight) and followed people up over time (through annual cycles) to compare a treatment versus control group in order to estimate the number of cases that were overweight in each group, and the medical care costs averted and number of QALYs saved per case of adulthood overweight prevented. Two models published in recent years[65,66] also adopted a similar two-stage approach. In one model,[65] the first stage takes a similar approach to that in Wang *et al.*[62] and Brown *et al.*[63] (i.e. the long-term extrapolation of the intervention effect on childhood weight status to adult age based on longitudinal data from Venn *et al.*[75]). The later models[65,66] adopted a different approach in the second stage. Rather than estimating the medical care costs averted per case of adulthood overweight prevented, both studies projected the incidence of chronic diseases associated with obesity and mortality in adulthood. A core set of weight-related events were included in the models: CHD, stroke, T2DM, breast cancer and colorectal cancer (CRC). The remaining three models[64,67,69] used epidemiological models, statistical[64,69] and prevalence profile[67] approaches. These models present as potentially useful modelling frameworks for the assessment of interventions for childhood obesity, but at the present time the detail reported for these approaches was not sufficient to fully understand the specific detail and methods employed, and, importantly, the detail available did not allow us to replicate the methods used.

## Model development (cost-effectiveness framework)

### Introduction/context

The effectiveness outcomes in the definitive trial of HeLP versus usual practice (e.g. BMI SDS) do not provide a strong basis for the direct estimation and presentation of trial-based cost-effectiveness analysis, with decision-makers requiring estimates of cost per QALY and or cost per life-year saved (i.e. longer-term outcomes) to consider and to compare results of cost-effectiveness analyses across a wide range of health-care interventions. Therefore, a modelling framework was developed to assess the cost-effectiveness of the HeLP intervention versus usual practice, in a UK school/education setting, in preventing childhood overweight and obesity. The model development process was informed and guided by the need to extrapolate from effectiveness data from the HeLP RCT (at 24-month follow-up) over a longer-term and policy-relevant time horizon. That is, the modelling aimed to predict the expected impact of a change in childhood BMI SDS (primary outcome), and any resulting change in the distribution of children by weight status (intervention vs. control) over time.

At the outset the HeLP intervention was expected to be a relatively low-cost (per child) intervention, and the objective was to develop a simple model to estimate the cost-effectiveness of HeLP. At the time of developing the model, the literature was sparse and there was no readily available modelling framework publicly available. We sought to develop a simple decision-analytic model, building on the published literature in this area and using weight-related health events of primary importance (when there was robust evidence of the relationship between incidence of events and weight status, applicable to a general population analysis), to assess the cost-effectiveness of the HeLP intervention.

## Model structure (summary)

The model development process is described in further detail below. In summary, the final model is a two-stage model (*Figure 6*). In stage 1 it predicts adult weight status (healthy weight, overweight and obese) from childhood weight status (BMI SD centile), and in stage 2 it predicts against a set of weight-related health events in adult years. A simplifying assumption is the use of only three weight status categories in adults: (1) healthy weight, (2) overweight and (3) obese. The use of three states, and the terminology used here (i.e. healthy weight), is consistent with the identified published model-based economic evaluations and also the available longitudinal data on the tracking of obesity from childhood to adult years. We acknowledge, and note here, that a small percentage of people will be underweight; however, for simplicity they are part of the group referred to as healthy weight and are treated in the same way as others in the healthy weight status category.

In stage 1 of the model, a decision tree structure is used to predict adult weight status as a function of childhood weight status. The model starts with a cohort of children distributed by weight status per the UK90 growth reference standards[51] using clinical assessment classifications ($\leq$ 91st centile; > 91st to $\leq$ 95th centile; > 95th to $\leq$ 98th centile; and > 98th centile), and predicts the expected distribution of the cohort by weight status in adulthood. Data from the UK longitudinal study reported by Power *et al.*,[76] which followed up participants from age 7 to age 33 years [see detail in *Model development (stage 1): predicting adult weight status from child weight status*], are used to predict adult weight status. Based on data reported by Power *et al.*,[76] the probability of children from a particular weight category remaining or changing weight status at the age of 33 years was estimated. These probabilities were applied to the starting cohort of children distributed by weight status in order to predict the likely distribution of people in adulthood, by weight status category. The impact of interventions, in this instance HeLP, can be considered/integrated in this first stage of the model via the specification of the starting distribution of children by weight status (e.g. an effective intervention will have fewer children in BMI SD centile categories associated with overweight and/or obese weight status). The model structure uses (starts with) inputs for a base cohort distribution of children by weight status categories and applies effectiveness data (e.g. hazard rates/relative risks) to proportions of children in categories of weight status associated with overweight and obesity in order to simulate the reported treatment effectiveness in a comparison of the same cohort of people both with and without the intervention (effectiveness).

In stage 2, a Markov model is used to model weight-related health events over time by adult weight status (*Figure 7*). Weight-related health events included in the model are T2DM, CHD, stroke and CRC [see detail in *Model development (stage 2): predicting the impact of overweight and obesity on incidence of weight-related health events*]. The model cycle length is 1 year and has a 30-year time horizon. The time horizon is chosen to reflect a longer-term adult time horizon, over which adult weight-related health events may be predicted (i.e. over age 33–62 years). Adults enter the model in one of three event-free health states depending on their weight status (healthy weight, overweight or obese). In the base-case model, adult overweight and obese weight status categories are BMI of 25.0–30.0 kg/m$^2$ and BMI of
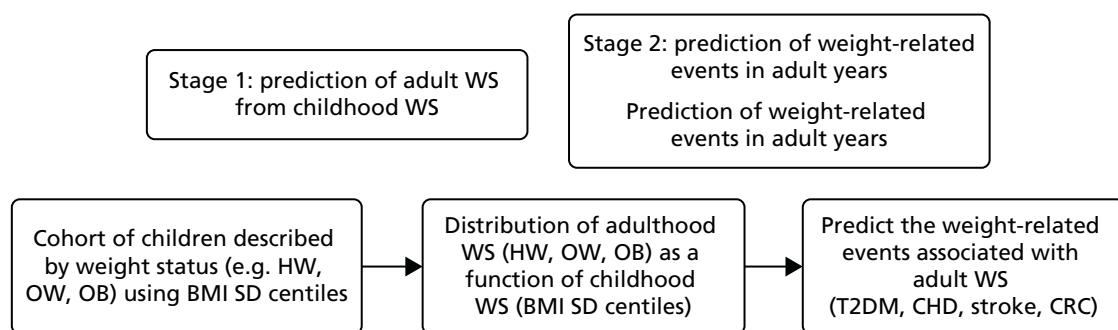


**FIGURE 6** Model overview. HW, healthy weight; OB, obese; OW, overweight; WS, weight status.
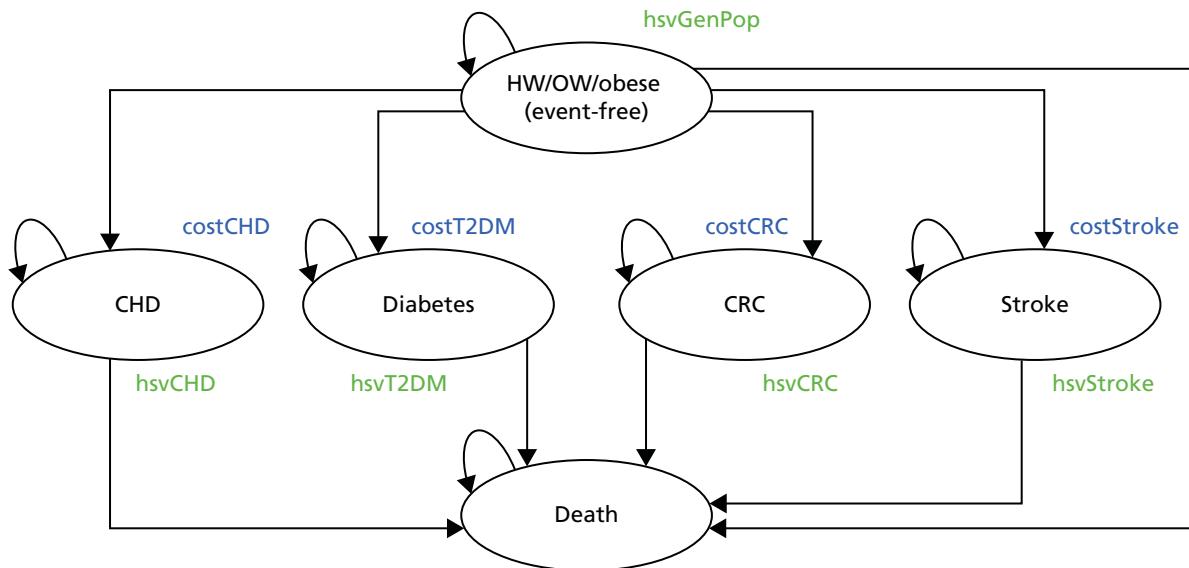
**FIGURE 7** Markov model structure. Costs by event/state: costCHD, costT2DM, costCRC and costStroke. Health state values (QALY weights) by event/state: hsvGenPop, hsvCHD, hsvT2DM, hsvCRC and hsvStroke. HW, healthy weight; OW, overweight.

> 30.0 kg/m², respectively, with healthy weight being BMI of < 25 kg/m². Each year (1-year cycle) adults are at risk of remaining in the event-free health state, developing a weight-related event, or death. The model assumes that adults who experience an event will remain in that health state until death and are not exposed to any other events; this is a simplifying assumption, consistent with previously published model-based evaluations. People in the event-free health state are exposed to general population all-cause mortality risks. People in states defined by weight-related events are exposed to condition-specific mortality risks [see *Model development (stage 2): predicting the impact of overweight and obesity on incidence of weight-related health events*]. The costs and outcomes associated with weight-related events are estimated based on a review of the published literature; costs are presented as 2015 UK pounds sterling (£). Data on mortality are from national data sources (all-cause mortality) and from published literature on event-related mortality. Future costs and benefits are discounted at 3.5% per annum in the base-case analysis, with the option to vary this in sensitivity analyses, and the perspective of the analysis, consistent with the context of the decision problem, is that of the UK NHS and Personal Social Services (third-party payer perspective). The decision-analytic context is guided by the UK NHS decision-making context, using methods guidance for conduct of economic evaluations issued by NICE.[77]

### Model development (stage 1): predicting adult weight status from child weight status
Longitudinal studies were used to inform the prediction of the proportion of children that will be healthy weight, overweight or obese in adulthood based on their childhood weight status.

A literature search was undertaken to identify systematic reviews that reviewed the literature concerning the tracking of childhood weight status into adulthood. A total of three systematic reviews were identified.[3,78,79] Owing to methodological differences between the reviews, the bibliographies of included studies for all three reviews were scrutinised for cohort studies reporting cohort data tracking childhood weight status into adulthood. A total of eight studies were identified.[75,76,80–85] Of these, four[75,76,80,85] presented data projecting weight status from childhood into adulthood and included classifications by weight status (healthy weight, overweight), and were therefore selected for consideration to populate the model. The remaining four studies were excluded, either because they provided no usable data[81,82,84] or because the data were for a child age range that was inappropriate for the current research.[83] The study characteristics for the included studies are presented in *Table 27*, and a summary description is provided below.

**TABLE 27** Study characteristics of the included weight status tracking studies

| Author, year | Country | Reference standard | Age in years (assessment time points) | BMI centile group[a] | Weight status (BMI in kg/m²) | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Underweight | Healthy weight | Overweight | Obese[a] |
| Power et al., 1997[76] | UK | UK90<br><br>Cole et al.[51] | 11–33 | ≤ 91st<br><br>> 91st to ≤ 95th<br><br>> 95th to ≤ 98th<br><br>> 98th | < 20 | 20–24.9 | 25.0–30.0 | > 30.0 |
| Wright et al., 2001[80] | UK | UK90<br><br>Cole et al.[51] | 11–50 | < 25th<br><br>25–74th<br><br>75–90th<br><br>> 90th | – | – | > 24 | > 30 |
| Laitinen et al., 2001[85] | Finland | NHANES I and II<br><br>Whitaker et al.[73]<br><br>Frisancho[86] (cited in Whitaker) | 14–31 | < 85th<br><br>85th to < 95th (obese)<br><br>≥ 95th (very obese) | < 18.5 | 18.5–24.9 | 25.0–29.9 | ≥ 30 |
| Venn et al., 2007[75] | Australia | IOTF<br><br>Cole et al.[87] | 15–34 | ≥ 5th to 90th (normal)<br><br>≥ 90th to 97th (at risk OW)<br><br>> 97th (OW) | Not defined as a separate category | < 25.0 | 25.0–29.9 | ≥ 30[b] |

CDC, Centre for Disease Control; IOTF, International Obesity Task Force; NHANES, National Health and Nutrition Examination Survey; OW, overweight.
a   The centile indicates the relative position of the child's BMI among children of the same sex and age. The growth charts show the weight status categories used with children and teens: underweight, healthy weight, overweight and obese.
b   Also included a very obese group defined as having a BMI of ≥ 35 kg/m².

**Power et al. 1997 (1958 UK birth cohort)** Power et al.[76] followed up a cohort of children who were born in Britain in March 1958 (the 1958 UK birth cohort). In this longitudinal study, height and weight measurements were taken at various points between the ages of 7 and 33 years. The weight status of children was defined by the UK90 growth reference standards.[51] At the end of the data collection, measurements for analysis were available for 11,212 adults. Adult BMI was categorised as healthy weight (20.0–24.9 kg/m²), overweight (25.0–30.0 kg/m²) and obese (> 30.0 kg/m²). From the study, the authors reported the weight status of an adult aged 33 years as a function of child weight status at the age of 11 years.

**Wright et al. 2001** Wright et al.[80] reported data on a UK cohort of children who were born in Newcastle upon Tyne between May and June 1947. Height and weight measurements were taken at 9, 13 and 50 years. Weight status of children was defined by UK90 growth reference standards. The analyses were based on the proportion of adults who were overweight or obese at the age of 50 years as a result of weight status at 9 and 13 years, respectively.

**Laitinen *et al.* 2001** Laitinen *et al.*[85] followed a cohort of children from birth year 1966 in northern Finland. Height and weight measurements were taken at birth, and at 1, 14, and 31 years of age. Data at 14 years were obtained from a questionnaire that was sent to participants, and follow-up data at 31 years (1997–8) were obtained from a questionnaire, a clinical examination or both. Adolescence (age 14 years) overweight was defined as a BMI $\geq$ 85th to < 95th percentile and obesity as a BMI $\geq$ 95th percentile. Adult BMI at 31 years was classified as follows: underweight (< 18.5 kg/m²), normal weight (18.5–24.9 kg/m²), overweight (25.0–29.9 kg/m²) and obese ($\geq$ 30.0 kg/m²). The analysis was restricted to individuals for whom BMI data were available for all measurement points (males, $n = 2876$; females, $n = 3404$).

**Venn *et al.* 2007** Venn *et al.*[75] examined overweight and obesity in Australian children followed through to adulthood. A cohort study of 8498 children aged 7–15 years participated in the 1985 Australian Schools Health and Fitness Survey. Of these, 4571 completed a follow-up questionnaire at age 24–34 years in 2001–5. Height and weight were measured in 1985, and self-reported at follow-up. The accuracy of self-reported data was checked for 1185 participants. Overweight and obesity in childhood were defined according to international standard definitions for BMI, and, in adulthood, as a BMI of 25–29.9 kg/m² and $\geq$ 30 kg/m², respectively, after correcting for self-report error.

*Summary*
The longitudinal studies identified here presented data on tracking of weight status from childhood into adulthood. The results presented for each of the identified studies indicate that childhood overweight/obesity persists into adulthood, endorsing the broader evidence base which supports the use of childhood weight status as a strong predictor of adult weight status in a policy context (e.g. Whitaker *et al.*[73]). There were methodological differences between the included studies, for example the number of participants included in the studies, the ages at which authors predicted weight status to and from, prevalence of overweight and obese across the studies, the thresholds used to define weight status as a child, the centile categories used to describe the distribution of weight status, and the growth reference population. All studies have limitations; however, these are considered to be high-quality longitudinal studies.

Data from each of these four longitudinal studies have been used to predict adult weight status by childhood weight status, represented in the form of transition probabilities across the defined child to adult weight status categories (*Table 28*). For stage 1 of the Exeter Obesity Model transition probabilities derived from the longitudinal data reported by Power *et al.*,[76] for the 1958 British birth cohort, are used in the base-case analysis (model structure), predicting the proportions of children (aged 11–12 years) expected to be healthy weight, overweight or obese at 33 years. These probabilities were applied to the distribution of children at the start of the model, expected to be informed by the 24-month follow-up data in the HeLP study, in order to predict the expected distribution of the cohort by adult weight status at the age of 33 years.

Data from Wright *et al.*[80] (British cohort data to age 50 years) and Laitinen *et al.*[85] (Finnish cohort data to age 31 years) were set out for potential use in sensitivity analysis to assess structural uncertainty (see *Table 28*). Although not specifically appropriate for sensitivity analyses, the estimate transit probabilities using data reported by Venn *et al.*[75] (Australian cohort data to age 29 years) are presented to provide a comparison between a relatively recent set of data and the data from other cohort studies[76,80] that were collected in earlier years.

*Model development (stage 2): predicting the impact of overweight and obesity on incidence of weight-related health events*

**Weight-related events** Stage 2 of the model starts with a cohort of people with a given distribution by weight status (healthy weight, overweight or obese), as described in stage 1. All persons start stage 2 as disease free (i.e. free of the major weight-related health events modelled in stage 2). Stage 2 uses CHD, stroke, T2DM and CRC as the weight-related health events of interest. The selection and inclusion of these four events is considered a conservative assumption (i.e. it is expected that the incidence of a greater number of health events will be influenced by weight status). The selection and inclusion of these four events is based

**TABLE 28** Predicted probabilities of adult weight status by child BMI centile category

| | Weight status | | |
| BMI centile category | HW | OW | OB |
|---|---|---|---|
| **Age 11 years[a]** | **Weight status at age 33 years** | | |
| ≤ 91st | 0.59 | 0.32 | 0.08 |
| > 91st to ≤ 95th | 0.25 | 0.39 | 0.36 |
| > 95th to ≤ 98th | 0.24 | 0.37 | 0.39 |
| > 98th | 0.10 | 0.34 | 0.56 |
| **Age 13 years[b]** | **Weight status at age 50 years** | | |
| < 25th | 0.57 | 0.38 | 0.05 |
| ≥ 25th to ≤ 74th | 0.32 | 0.54 | 0.14 |
| > 74th to ≤ 90th | 0.23 | 0.55 | 0.23 |
| > 90th | 0.13 | 0.42 | 0.45 |
| **Age 14 years[c]** | **Weight status at age 31 years** | | |
| < 85th | 0.69 | 0.27 | 0.04 |
| ≥ 85th to < 95th | 0.28 | 0.48 | 0.23 |
| ≥ 95th | 0.15 | 0.33 | 0.51 |
| **Age 11 years[d]** | **Weight status at age 29 years** | | |
| ≥ 5th to < 90th | 0.62 | 0.29 | 0.09 |
| ≥ 90th to < 97th | 0.21 | 0.37 | 0.42 |
| ≥ 97th | 0.14 | 0.29 | 0.57 |

HW, healthy weight; OB, obese; OW, overweight.
a Data from Power et al.[76]
b Data from Wright et al.[80]
c Data from Laitinen et al.[85]
d Data from Venn et al.[75]

on the review of the previously published model-based economic evaluations, discussed above, which typically included a small core set of weight-related health events similar to those selected here, and consideration of the data on incidence of events by weight status (see below), plus discussion of the conceptual modelling framework across the HeLP stakeholder group(s). Of importance when identifying health events for inclusion in the model was the simplifying assumption in the model that once a person experienced a condition they were then not exposed further to additional events. This led to a focus on the main primary health conditions, rather than events such as hypertension, which, although relatively frequent (and serious), was considered to be analogous to an intermediate outcome increasing the future risk of events such as CHD, stroke and T2DM.

Incidence (probability) of the weight-related health events was informed using a targeted keyword literature search (in MEDLINE) to identify studies that estimated the incidence of comorbidities related to obesity and overweight. The search strategy was structured using the following terms: (obesity OR adipose tissue OR body mass index OR body composition) AND (comorbidit$OR multimorbidit$OR chronic disease). No date limits or language restrictions were applied to the search. Eligible studies were systematic reviews of cohort studies, reporting risk estimates based on the incidence of weight-related events including T2DM, cardiovascular disease (including hypertension, coronary artery disease, congestive heart failure, pulmonary embolism, stroke and dislipidaemia) and CRC. The searches retrieved 126 titles and abstracts. Ten reviews were considered at full-text stage. Of these, only one was considered eligible for inclusion.[88]

Guh *et al.*[88] conducted a literature search to provide an estimate of the incidence of each comorbidity related to obesity and overweight using a meta-analysis. Twenty comorbidities were initially included in this analysis: cancer (kidney, colorectal, prostate, ovarian, uterine/endometrial, esophageal, pancreatic and postmenopausal breast), T2DM, cardiovascular disease (hypertension, coronary artery disease, congestive heart failure, pulmonary embolism, stroke and dislipidaemia), gallbladder disease, chronic back pain, osteoarthritis, asthma and sleep apnoea.

For each of the weight-related events evaluated in Guh *et al.*,[88] we sourced the studies included in the review and extracted the data on number of incident cases and person-years from the primary papers contributing to the reported pooled relative risk estimate. We used these data to estimate the incidence of the events in the reference group (i.e. healthy weight), assuming a male-to-female gender ratio of 1 : 1. Using these raw data on incidence in the reference group, we then applied the risk ratios reported by Guh *et al.*[88] to estimate the incidence of the weight-related events in people with weight status at overweight and obese (*Table 29*).

**Mortality data** People are exposed to an annual mortality risk in each cycle of the model in stage 2. People in the disease-free state (no weight-related event) are subject to all-cause mortality risks based on age-specific data from the UK Office for National Statistics,[89] using data reported for 2012–14, assuming a cohort of people with a gender ratio of 1 : 1.

Data on mortality risks by health condition, whereby people are in health states defined by the weight-related health events, were informed by a targeted literature search of the published literature. We searched the literature using terms for the population of interest (i.e. T2DM, CHD, stroke and CRC) with terms for mortality rate using the MEDLINE database in December 2015. *Table 30* presents the data that we have used to populate the model on mortality risks for the health events. For stroke and CRC, we have used a mortality input for year 1 of the event, together with a separate mortality parameter for subsequent years in that condition/health state. This structure on mortality data is consistent with the findings from the literature search.

**TABLE 29** Estimated incidence rate (probability) of adult weight-related events by weight status: cases per 1000 person-years

| Weight-related event | Weight status | | |
| --- | --- | --- | --- |
| | HE | OW | OB |
| CRC | 0.00087 | 0.00129 | 0.00159 |
| T2DM | 0.00165 | 0.00456 | 0.01334 |
| Stroke | 0.00141 | 0.00188 | 0.00234 |
| CHD | 0.00237 | 0.00329 | 0.00482 |

HE, healthy; OB, obese; OW, overweight.
Data source: Guh *et al.*[88]

**TABLE 30** Mortality rates by condition of interest

| Health state | Annual mortality rate | Source |
| --- | --- | --- |
| CHD | 0.170 | Sigurdsson *et al.*, 1995[90] |
| Diabetes (T2DM) | 0.056 | Roper *et al.*, 2002[91] |
| Stroke: year 1 | 0.370 | Hankey *et al.*, 2000[92] |
| Stroke: subsequent years | 0.100 | Hankey *et al.*, 2000[92] |
| CRC: year 1 | 0.243 | Cancer Research UK, 2012[93] |
| CRC: subsequent years | 0.080 | Cancer Research UK, 2012[93] |

**Health state values (QALY weights)** Data used in the model on health state values to derive QALYs are reported in *Table 31*. The health state value used for the 'event free' state (0.86) is based on UK EuroQol-5 Dimensions (EQ-5D) data, consistent with methods guidance from NICE,[61] and is the general population mean health state value reported by Kind *et al.*[94] Data inputs for other model states were based on a literature review to identify published systematic reviews reporting health state values (utilities) associated with the health states (events, conditions) used in the model. The search method involved targeted keyword searching in MEDLINE (December 2015) for CHD, T2DM, stroke and CRC. Terms for health state values ('health state value$' and 'EQ-5D' and 'EQ5D') were combined with terms for the specified conditions of interest. A filter was applied to the searches to identify systematic reviews (review or 'systematic review' or 'literature review' or meta$). In addition, we used the data reported in the model-based economic evaluations for obesity identified earlier (see *Table 26*). In considering the appropriateness of the identified health state values, the following attributes were taken into account: the population describing the health state (e.g. age, sex, disease severity), the utility value elicitation technique (e.g. time trade-off, standard gamble, visual analogue score), the sample size and the country. Further detail on the literature identified is presented below. The data used in the model for health state values are broadly in line with published estimates used in other public health modelling contexts (e.g. see studies reported in *Table 27*).

## *Summary of literature identified, and selection of health state values as model inputs*

**Type 2 diabetes mellitus** One systematic review was identified[98] that identified a set of health state/ utility values for one or more complications in people with diabetes. The methodology of identified studies was considered against the requirements of the UK NICE reference case (2013) (i.e. criteria included use of EQ-5D questionnaire value set or other established measure of preference, health states should be determined by patients, and valuation performed using a societal valuation algorithm [preferences of general public]). The authors created a preferred input set of utility values. Of 16,574 records identified, they report on 61 studies, of which we considered 19 to be consistent with the NICE criteria.

The studies included in the Beaudet *et al.*[98] review were scrutinised and baseline T2DM utility values were extracted. For the base case for the current Exeter Obesity Model, the utility value of 0.78 reported in the study by Clarke *et al.*[95] was considered to be the most methodologically robust, given its large sample size, T2DM-specific population and use of the EQ-5D in a UK population. This estimate is also broadly consistent with other published estimates for T2DM without complications (e.g. NICE[99]).

**Coronary heart disease** One systematic review was identified[100] that reviewed utility data based on EQ-5D in cardiovascular disease. Sixty articles were identified that reported EQ-5D scores (health state values). There was wide variation in terms of cardiovascular disease subgroups, disease stage, age distribution and other methodological aspects. The majority of studies reported EQ-5D index scores using the UK-based algorithm. The majority of identified studies reported data for ischaemic heart disease and cerebrovascular disease. Stratification by disease severity was possible for ischaemic heart disease and heart failure participants.

**TABLE 31** Health state values

| Health state | Parameter value, mean (SD) | Source |
|---|---|---|
| Event-free | 0.86 (0.23) | Kind *et al.*, 1999[94] |
| T2DM | 0.78 (0.77) | Clarke *et al.*, 2002[95] |
| CHD | 0.67 (0.34) | Matza *et al.*, 2015[96] |
| Stroke | 0.52 (0.38) | Matza *et al.*, 2015[96] |
| CRC | 0.76 (NR) | Djalalov *et al.*, 2014[97] |

NR, not reported.

Supplementary searches in PubMed identified one additional study[96] reporting health state values for cardiovascular conditions, also distinguishing between acute impact including the cardiovascular event and the chronic post-event impact. Three cardiovascular conditions were described: stroke, acute coronary syndrome and heart failure. One-year acute health states represented the event and its immediate impact, and post-event health states represented chronic impact. Values were from a sample ($n = 200$) of the UK general population. The mean time trade-off health state value for acute coronary syndrome was 0.67 (SD 0.34) and for chronic post event health states was 0.82 (0.17).

For the base case in the Exeter Obesity Model, the time trade-off value of 0.67 (SD 0.34) reported by Matza *et al.*[96] was considered the most reflective of the health state in the model. This is in line with estimates used in other models of this type.[101–103]

**Stroke** One systematic review was identified[104] that reviewed evidence on health state values for stroke. Mean utilities of major stroke (Rankin scale 4–5) and minor stroke (Rankin scale 2–3) were presented, stratified by study design and elicitation method. A total of 23 studies were included in the review. Study populations in the included studies were healthy people, people at risk of stroke or stroke survivors. The studies included in the review reporting data from general public (non-patient, healthy participants) were considered most relevant to the health state used in the Exeter Obesity Model.[105–108] However, two of these studies[105,108] used only the visual analogue scale to elicit utilities, and the remaining studies used mixed samples and had relatively small sample sizes. The supplementary searches carried out to identify utility values for CHD identified one paper[96] that also reported health state values for stroke. In that study, a number of health state values are reported, including a mean (SD) value of 0.52 (0.38) for chronic stroke. We also reviewed previous health technology appraisals and clinical guidelines reported by NICE for stroke, identifying data presented for stroke in the appraisal of apixaban for the prevention of stroke and systemic embolism in people with non-valvular atrial fibrillation.[103,109]

Considering the varied evidence available, we use the data reported by Matza *et al.*[96] in the base case for the model presented here. This estimate is consistent with health state values used for stroke in prior public health modelling contexts, such as exercise referral schemes.[102,110]

**Colorectal cancer** One systematic review was identified[97] that reviewed evidence on utility weights for CRC health states. Twenty-six articles met the inclusion criteria for the review. The authors report a mean EQ-5D health state value of 0.76 (based on 33 reported utilities), and this is used in the current model framework. This estimate is broadly consistent with data used to inform the NICE technology appraisal of drugs CRC,[111] aligned to first-line and second-line treatment health states.[111]

## Cost data

As stated previously, the perspective applied for the economic analyses is that of the UK NHS and Personal Social Services, and cost estimates included for model inputs are consistent with that perspective (i.e. third-party payer).

**Intervention costs (HeLP)** The additional cost included for the delivery of the HeLP intervention, in the base-case analysis, is the estimated mean per participant cost of £214.14 (see *Table 22*).

**Costs for disease states (weight-related events)** An annual estimate of costs associated with model states by condition (weight-related event) is applied in the model when people are in these states. The estimates of annual costs used are presented in *Table 32*. The estimates used are informed by a literature search to identify published systematic reviews of the literature in each of the disease areas.

Targeted literature searches in MEDLINE and web searches combining terms for the population of interest (i.e. T2DM, CHD, stroke and CRC) with terms for costs (e.g. cost, annual cost, per patient cost, UK cost) were conducted in December 2015. Cost estimates used in the model were inflated/uprated to 2014/15 prices, when appropriate, using inflation indices from Curtis and Burns.[57]

**TABLE 32** Treatment costs per health state

| Health state/condition | Parameter value/costs (£) | Source |
|---|---|---|
| T2DM | 3717[a] per year | Kanavos et al., 2012[112] |
| CHD | 2852 per year | NHS Scotland, 2007[113] |
| Stroke | 2800[b] | Saka et al., 2009[114] |
| CRC | 8441 per year | Cancer Research UK, 2014[115] |

a 2010/11 costs.
b Reflects the cost of a rapid recovery case, applied as a one-off treatment cost (in year 1 of stroke).

**Type 2 diabetes mellitus** Kanavos et al.[112] report costs for diabetes (direct and indirect) across France, Germany, Italy, Spain and the UK. Estimates include inpatient, outpatient, primary and community costs, and medication-related costs. Annual UK costs for T2DM are estimated at £3717 mean per-person costs. These costs are based on high-quality data from large retrospective cohort studies for primary care and outpatient activity[116] and inpatient costs.[117]

**Coronary heart disease** A report from NHS Scotland[113] presents estimates of the resource use and costs associated with cardiovascular disease-related diagnoses. The cost per patient for CHD was estimated as £2852. This estimate was based on a comprehensive costing methodology, assessing the average cost per patient for the following principal diagnoses: unstable angina, £1760; acute myocardial infarction, £3990; subsequent myocardial infarction, £4240; chronic ischaemic heart disease unspecified, £2765; and stable angina, £1500.

**Stroke** Saka et al.[114] reported detailed data on the resource use and cost associated with stroke, estimating societal and direct costs of care. Direct care costs include diagnosis, inpatient care and outpatient care and are based on data from a population-based register. The data are predominantly presented as aggregate burden/cost of illness results; however, when considering an estimate of cost per case/per stroke patient (applying assumptions/methods from Bosanquet and Franks[118]), the authors present estimates of £2800 and £17,500 for costs associated with a rapid recovery case and a case with disability but discharged into the community. As a conservative assumption here we use the cost estimate of £2800 for the treatment of stroke events arising in the model presented here. This cost is applied as a one-off treatment cost (model payoff), rather than an annual treatment cost.

**Colorectal cancer** A report from Cancer Research UK[115] presents mean estimated costs for CRC of £8441, covering diagnosis and treatment costs, across disease stages 1–4. Cost estimates are based on a mapping of national treatment guidelines, national data sets and clinical audit data. This cost estimate is similar to the data presented by Trueman et al.[119] in a report commissioned by the UK Department of Health (mean cost approximately £8800).

### Modelling summary

The model structure has been described above, and the parameter inputs to populate the model have been presented together with a description of methods and rationale for data inputs and assumptions used. *Appendix 9* presents a summary of the model parameters and inputs used.

The Exeter Obesity Model has been developed as part of the research funded alongside the NIHR-funded clinical trial (RCT) on HeLP versus usual practice. The model is set out here as a parsimonious modelling framework appropriate for estimating cost-effectiveness for HeLP versus usual practice, given the expectation of a relatively low-cost intervention being used in a public health context, with benefits accruing over the longer term (i.e. in adult years, 20–50 years after the intervention is delivered). The specific focus of the model development has been the use of key weight-related events as outcomes of interest in the model,

the use of published data to populate the model, and the presentation of cost-effectiveness analyses using cost per QALY, and cost per life-year saved, in order to inform health policy and decision-making in a UK context. The conceptual model and the simple model design are based on the need to answer the a priori research question on whether or not HeLP is cost-effective versus usual practice, that is, whether it represents value for money to a UK third-party payer (such as the NHS) when assessed against commonly used estimates of willingness to pay for health benefits (i.e. cost-per-QALY thresholds).[77]

However, as referred to above, the HeLP intervention is not effective compared with usual practice, and therefore the modelling framework is presented here as a potentially useful generic contribution to the methods available to model the effectiveness and cost-effectiveness of interventions for childhood obesity, where it is important to extrapolate from relatively short-term effectiveness outcomes to longer-term impacts of interventions, into and over adult years.

### Cost-effectiveness analyses (HeLP vs. usual practice)

The above sections have presented results of the estimates of the costs associated with delivery of the HeLP intervention, and have described the modelling framework set out to estimate the cost-effectiveness of the HeLP intervention versus usual practice.

#### *Effectiveness of HeLP versus usual practice*

Data presented in *Chapter 3* have reported that there is no evidence that HeLP is more effective than usual practice. *Table 9* reports no statistically significant differences between HeLP participants and controls at 24-month follow up on the primary outcome measure of BMI SDS. The Exeter Obesity Model estimates cost-effectiveness with a starting premise that an intervention has shown a difference in effectiveness that can be translated to a reduction in the relative risk associated with being in BMI SDS weight status categories associated with overweight and/or obesity, using the BMI centile categories applied in the base data drawn from Power *et al.*[120] (see *Table 28*). This model structure allows a comparison of a cohort of children over time, with and without the intervention. The results from the HeLP RCT do not show reductions in the relative risk for being in states aligned to overweight and/or obesity, finding no statistically significant difference between the intervention group and the control group (see *Table 11*).

### Cost-effectiveness of HeLP versus usual practice

Given the effectiveness profile for the HeLP intervention, there is no expectation, using the Exeter Obesity Model framework, that there will be any improvements in the likely incidence of the weight-related health events (CHD, stroke, T2DM and CRC), or cost savings to the health and social care system associated with weight-related events. This profile, together with a certain expectation that the introduction of the HeLP intervention is associated with additional resource use and costs, in the short term, leads to the clear conclusion that HeLP is not considered to be a cost-effective alternative to usual practice.

HeLP incurs additional costs with no identified or expected health gains, using the analytical framework developed and presented here, and therefore usual practice 'dominates' HeLP in a resource allocation context (*Table 33*). HeLP does not offer value for money to the UK health-care system, compared with usual practice.

As this is a clear and unambiguous outcome, consistent with agreement with the TSC, no further detailed cost-effectiveness analysis is presented here. We have presented detailed cost analysis and results on the estimated expected mean incremental cost per child for delivery of the HeLP intervention, together with estimates of expected additional costs per school (per class).

#### *Exploratory cost-effectiveness analyses using the Exeter Obesity Model*

**Predicted results for control participants** In the HeLP RCT the distribution of children starting the trial and at 24-month follow-up by weight status categories, as centile categories used by Power *et al.*,[120] are reported in *Table 34*, together with their predicted weight status profile as adults aged 33 years (using standard adult BMI categories).

**TABLE 33** Summary of differences in outcomes and costs (HeLP treatment vs. control)

| Outcome | Group | | Adjusted difference (intervention vs. control) |
|---|---|---|---|
| | Intervention (HeLP) | Control (usual practice) | |
| Mean BMI SDS (SD)[a] (*n* = 650) | 0.35 (1.25) (*n* = 630) | 0.22 (1.22) (*n* = 620) | –0.02, 95% CI –0.082 to 0.045; *p* = 0.568 (control = lower); no statistically significant difference |
| Proportion (%) overweight at 24-month follow-up[b] (baseline)[c] | 14.1 (12.1) | 13.5 (10.7) | NA |
| Proportion obese at 24-month follow-up[b] (baseline)[c] | 16.7 (14.6) | 13.1 (12.6) | NA |
| Proportion (%) overweight or obese at 24-month follow-up[b] (baseline)[c] | 30.8 (26.8) | 26.6 (23.3) | OR (adjusted analyses) 1.15, 95% CI 0.79 to 1.67[b]  (Control lower odds than intervention)  No statistically significant difference |
| Physical activity outcomes, 18-month follow-up,[d] average daily total mean (SD) | 199.71 (43.9) (*n* = 359) | 198.05 (40.2) (*n* = 386) | 1.26, 95% CI –6.84 to 9.36; no statistically significant difference |
| Mean weekly FIQ (18-month follow-up)[d] | | | |
| Energy-dense snacks, mean number (SD) | 3.72 (1.86) | 4.06 (2.07) | –0.37, 95% CI –0.066 to –0.07; statistically significant improvement with intervention |
| Healthy snacks, mean number (SD) | 3.61 (1.63) | 3.30 (1.5) | 0.22, 95% CI –0.04 to 0.47; no statistically significant difference |
| Negative food markers, mean number (SD) | 5.90 (2.73) | 6.38 (3.0) | –0.47, 95% CI –0.91 to –0.02; statistically significant improvement with intervention |
| Positive food markers, mean number (SD) | 6.20 (2.36) | 5.77 (2.31) | 0.26, –0.12 to 0.64; no statistically significant difference |
| Additional treatment costs (mean, per participant)[e] | £214.13 | None | £214.13; intervention more costly than control |

NA, not applicable.
a  See *Table 9*.
b  See *Table 7*.
c  See *Table 11*.
d  See *Table 14*.
e  See *Table 22*.

**TABLE 34** Summary of weight status profile by group at childhood and by adult predicted profile

| Weight status category | Group | | | |
|---|---|---|---|---|
| | HeLP | | Control (usual practice) | |
| | Baseline | 24 months | Baseline | 24 months |
| BMI centile category (Power *et al.*[120])[a] | | | | |
| ≤ 91st | 80.42% | 76.51% | 82.27% | 80.32% |
| > 91st to ≤ 95th | 4.93% | 6.83% | 5.13% | 6.61% |
| > 95th to ≤ 98th | 6.28% | 7.46% | 6.38% | 4.84% |
| > 98th | 8.37% | 9.21% | 6.22% | 8.23% |
| Adult weight status[b] | | | | |
| Healthy weight | NA | 49.87% | NA | 51.37% |
| Overweight | NA | 33.31% | NA | 33.14% |
| Obese | NA | 16.82% | NA | 15.48% |

NA, not applicable.
a  Source: HeLP RCT, intention-to-treat participants.
b  Predictions using data from Power *et al.*[120]

When modelling the experiences of a cohort ($n = 1000$) of children defined using the weight status data for the control participants in the HeLP study, we predict that at age 33 years the cohort will result in an adult cohort with 49% overweight or obese adults and 51% healthy weight adults. We predict that 33% will be overweight and 15.5% will be obese, and this is consistent with the data reported by age in the Health Survey for England (2009 data).[121] Over the 30-year follow-up period, adult-years, there are 252 weight-related events (CHD, stroke, T2DM and CRC). These 252 events comprise 76 cases of CHD, 105 cases T2DM, 28 cases of CRC and 42 cases of stroke; 675 participants remain event free at age 62–63 years. The mean per participant cost associated with those events is estimated at £1140 (£6821 undiscounted; £3145 discounted at 1.5% per annum). Over the 30-year time horizon (30 adult-years) where events, costs and event related outcomes are estimated, in addition to all-cause mortality, the mean number of life-years per participant is 26.75, prior to discounting (4.47 when discounted at 3.5% per annum; 12.34 at 1.5% per annum), and the mean number of QALYs is 22.78 prior to discounting (3.8 when discounted at 3.5% per annum; 10.5 at 1.5% per annum). Of note, overall we are modelling life expectancy over a 50-year time horizon, with specific payoffs (future costs and outcomes for health events) captured in the latter 30 years.

The modelling framework developed to assess the cost-effectiveness of interventions such as HeLP is dependent on identifying a beneficial (effective) intervention, with effectiveness reflected through an improved weight status profile in childhood. That mechanism of effect provides an opportunity to consider the future potential impact of a change in childhood weight status, into adult years, and the estimation of that effect over adult years through a reduction in future incidence of a core set of weight-related health events. When applying the data from the HeLP exploratory trial (crude unadjusted relative risks, by BMI SDS category), the modelling framework indicates that HeLP has the potential to be a cost-effective intervention (see scenario analyses E in *Table 35*). This is primarily due to the reduced number of participants in the HeLP exploratory trial intervention arm, compared with the number in the control arm, who ended up in the childhood weight status category associated with obesity. However, the sample size in the exploratory trial was very small, and the distributions by weights status in the intervention and control arms were not balanced at baseline, given that the study design was for an exploratory trial, and feasibility research.

Here, to illustrate the use of the modelling framework, and to illustrate the potential scenarios in which an intervention such as HeLP may be considered cost-effective, we present a series of scenario analyses in *Table 35* describing alternative weight status profiles for a potential hypothetical treatment group, compared with the data reported for the HeLP control participants (standardised to a cohort of $n = 1000$).

In these exploratory analyses we have used the inputs on expected additional costs for the HeLP intervention, and the data on the weight status of the control participants in HeLP at 24-month follow-up, to explore the cost-effectiveness of the HeLP intervention at hypothetical scenarios on intervention effectiveness [i.e. relative risk(s) of being overweight and/or obese]. We use hypothetical estimates of the relative risk, between controls and intervention, and we present data on predicted distribution by adult weight status, and summary data on weight related events, cost per life-year saved and cost per QALY gained.

The exploratory results demonstrate that when the relative risks are $\geq 0.75$ and above ($\leq 25\%$ reduction in proportion in BMI SDS centile category) the cost per QALY is above the commonly referred to UK NHS willingness to pay for a QALY, of between £20,000 and £30,000 per QALY.[61] However, of note is that the number of children in the BMI SD centiles associated with overweight and obesity is relatively small (i.e. circa 6% in each category), which therefore requires a relatively high risk reduction in order to make a meaningful impact on the actual numbers of children, and thereafter adults, in the overweight and/or obesity related weight status categories (e.g. a relative risk of 0.75 for childhood obesity represents a difference of approximately 20 cases per 1000 children). We therefore suggest that relatively modest effectiveness results may lead to a scenario in which an intervention is cost-effective, regardless of the apparent magnitude of the estimated relative risk (for small proportions of the cohort). We advise that the exploratory results presented here should not be interpreted in a way that indicates that interventions may be cost-effective only when dramatic effects are seen. The exploratory results also illustrate that when considering the impacts of such a public health intervention over time, the mean incremental costs and outcomes are very small, and the cost per life-year gained, and cost per QALY gained,

**TABLE 35** Scenario analyses: controls (HeLP RCT) vs. hypothetical comparison scenarios for effectiveness

| | BMI SD centile category (%) | | | | Adult WS distributions (%) | | | Events[a] (*n*) | Mean costs (£) | Mean life-years | Mean QALYs | Cost per LYG (£)[a] | Cost per QALY (£)[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ≤ 91st | > 91st to ≤ 95th | > 95th to ≤ 98th | > 98th | HW | OW | OB | | | | | | |
| Control, 24 months | 80.3 | 6.6 | 4.8 | 8.2 | 51.4 | 33.1 | 15.5 | 252 | 1140 | 4.47 | 3.81 | N/A | N/A |
| A: vs. control, 25% reduction in OB (RR 0.75) | 82.3 | 6.6 | 4.8 | 6.2 | 52.4 | 33.1 | 14.5 | 249 | 1335 | 4.476 | 3.812 | 44,723 (13,540) | 46,654 (14,124) |
| B: vs. control 50% reduction in OB (RR 0.50) | 84.4 | 6.6 | 4.8 | 4.1 | 53.4 | 33.1 | 13.5 | 246 | 1318 | 4.480 | 3.816 | 20,264 (4673) | 21,139 (4875) |
| C: 25% reduction in OB and OW (RRs 0.75) | 83.6 | 6.6 | 3.6 | 6.2 | 52.83 | 33.04 | 14.13 | 248 | 1328 | 4.478 | 3.813 | 30,850 (8513) | 32,183 (8881) |
| D: 50% reduction in OW and OB (RRs 0.50) | 86.9 | 6.6 | 2.4 | 4.1 | 54.29 | 32.93 | 12.78 | 243 | 1303 | 4.484 | 3.819 | 13,330 (2162) | 13,906 (2255) |
| E: Applying data from HeLP exploratory RCT[c] (crude RRs) | 87.6 | 5.2 | 4.7 | 2.6 | 54.7 | 32.9 | 12.4 | 242 | 1295 | 4.486 | 3.821 | 10,996 (1302) | 11,439 (1361) |

HW, healthy weight; LYG, life-year gained; OB, obese; OW, overweight; RR, relative risk; WS, weight status.
a  Events comprising CHD, stroke, T2DM and CRC.
b  Cost per life-year and QALY at base-case discount rate of 3.5% per annum, with figure in brackets the estimated cost per life-year and QALY when using a discount rate of 1.5% per annum.
c  Lloyd et al.[28]
**Note**
Baseline/starting distribution (standardised to cohort, *n* = 1000): 82.3% BMI SDS centile ≤ 91st, 5.1% BMI SDS centile > 91st to ≤ 95th, 6.3% BMI SDS centile > 95th to ≤ 98th, 6.2% BMI SDS centile > 98th.

estimates are very sensitive to small changes in the mean incremental inputs to the cost-effectiveness ratios. Furthermore, the exploratory results highlight that the results are sensitive to the parameter used to discount future costs and outcomes, with cost-effectiveness estimates markedly different (more attractive) when using a discount rate of 1.5% per annum (vs. the base case 3.5% per annum).

## Discussion

We have estimated the additional cost for delivery of the HeLP intervention to be £214 per participant, based on delivery to a cohort of schools (classes) similar to that in the HeLP study. The estimated cost is based on good-quality data collected within the trial, describing the staff inputs and resource use required to deliver the intervention, alongside realistic estimates of the costs associated with the resource inputs. The drama component, integral to the HeLP intervention, is the primary area of resource use and cost, and, together with the costs associated with the HeLP co-ordinator role, accounts for over three-quarters of the cost for delivery of HeLP. There are economies of scale when delivering HeLP in larger schools, given that some components are delivered once at the school level, with costs spread over a larger number of children receiving the intervention. We would anticipate that interventions such as HeLP would be commissioned (funded) across a cohort of schools, and in that setting we report an estimated mean cost per class of approximately £5350. This is a significant additional cost per class, at the level of the school budget, but we have shown in the illustrative cost-effectiveness analyses that a relatively small reduction in the predicted number of adverse weight-related health events (e.g. T2DM and CHD) in future adult years demonstrates that this investment represents value for money using a health and social care sector perspective.

However, in the trial we found no evidence that the HeLP intervention was effective in terms of a difference in mean BMI SDS and in preventing overweight or obesity at 24-month follow-up. Given the scenario of additional costs, together with an absence of demonstrated benefits in weight status, the result of the cost-effectiveness analyses is unambiguous, with usual care dominating the HeLP intervention.

As part of the research process, we have described the development of a framework for assessing the cost-effectiveness of the HeLP intervention (or similar interventions). We have described the Exeter Obesity Model, which builds on the published evidence base for methods surrounding the model-based economic evaluation of interventions for childhood obesity, in a public health context. When seeking to estimate the cost-effectiveness of interventions such as HeLP, there is no readily available decision-analytic modelling framework to utilise. We have therefore set down such a framework and described it in detail to support any future research in a similar context. The modelling framework is simple, using good-quality longitudinal data, and limiting its scope to a small number of health events where there is robust evidence that the incidence of disease has a strong association with weight status, and where there is a published meta-analysis reporting relative risks by weight status categories. However, the modelling framework makes a number of simplifying assumptions, which are deserving of further sensitivity analyses and research. We have not presented a full detailed set of cost-effectiveness results owing to the effectiveness scenario for the HeLP intervention. We had anticipated presenting cost-effectiveness analyses (results) when addressing structural uncertainty, and applying alternative estimates of the predictions of adult weight status as a function of childhood weight status (BMI SD centiles) (i.e. alternatives to the data use in the base-case analysis[120]). Such sensitivity analyses have not been undertaken, but we do report the estimated transition probabilities for alternative sources of data to inform future modelling developments. Based on our preliminary modelling, when using data from Wright *et al.*[80] or Laitinen *et al.*[85] (see *Table 27*) we would expect the outcome for estimates of cost-effectiveness (i.e. binary outcome where intervention is either considered cost-effective or not) to be similar to that when using the base case data from Power *et al.*[120]

Modelling the cost-effectiveness of public health interventions typically results in the use of a long-term time horizon, often over 30–40 years into the future, and given the expectation that interventions are population based with only a small number of identifiable beneficiaries (in terms of hard outcomes), the mean incremental differences in costs and outcomes are relatively small (averaged across a large target group for the intervention). This leads to difficulties in articulating the expected benefits at a population

level, and to difficulties associated with sensitivity in cost- effectiveness ratios (e.g. cost per QALY estimates) to very small changes to the estimated incremental costs and effects. We see these scenarios played out here in the exploratory cost-effectiveness analyses and cost-per-life-year and cost-per-QALY estimates. For example, in scenario B (see *Table 35*), we report a gain of 10 life-years (approximately 6.5 QALYs) per 1000 children exposed to the HeLP intervention, owing to the prevention of approximately six adverse health events (including preventing five cases of T2DM). At a geographical region, such as in Devon, we would see this multiplied using the target population of 9- to 10-year-olds, approximately 10,000 children, and benefits would be in the region of 100 life-years saved per year. This estimate is adjusted using a discount rate, but an unadjusted (non discounted) estimate would be in the region of 520 life-years saved per 10,000 children receiving an intervention (or 240 life-years saved, using a discount rate of 1.5% per year). These illustrations are speculative and hypothetical, and scenario B assumes an effectiveness scenario with a relative risk of 0.5 for obesity, so we must urge caution here when we seek to illustrate a potential analytical framework (although the HeLP exploratory RCT suggested an effect that was much higher than that in this scenario).

A related characteristic of the modelling framework, although more specific to estimates of effectiveness, is that the magnitude of the relative risk estimates is likely to be quite sensitive to changes in weight status for relatively small numbers of children. That is, given the small numbers of children considered obese in the starting control cohort (e.g. approximately 60–80 per 1000 children), and the large numbers in other weight status categories, the relative risk statistic can be in the region of 0.5 when there are differences of only 20 children in differing directions of weight status (when comparing two similarly distributed cohorts). The exploratory data presented in *Table 35* give some indication of this prevailing characteristic of the effectiveness inputs (albeit in a hypothetical setting), although further consideration is recommended.

Specific challenges in modelling the impacts of childhood obesity interventions, using the framework suggested here, include the profiles for the incidence of health events by weight status. The differences in the distributions by weight status are small in both child and adult populations, with only around one-fifth of the population likely to be overweight or obese. This, combined with the data available on the difference in rates of events by weight status, presents an analytical challenge. For example, the incidence of T2DM and CHD is approximately 13.3 cases and 4.8 cases, respectively, per 1000 person-years in obese states, compared with estimates of 1.7 cases and 2.4 cases per 1000 person-years in healthy weight status. Of note here is the 8.5-fold risk of T2DM for obese versus healthy weight status, compared with the twofold risk of CHD for obese versus healthy weight status. We see this reflected in the modelling results, with the vast majority of the differences in predicted events due to cases of T2DM, with very little difference in the incidence of the other events when comparing a potential treatment cohort with controls. Such results provide a rationale for modelling only those health events that exhibit a potentially meaningful difference in incidence by weight status (i.e. meaningful in a mathematical modelling context).

The model presented here has simplifying assumptions, which in many instances are due to the parsimonious structure of the modelling framework, although the absence of data in some areas is also a mitigating factor. We suggest that the model is likely to be sufficient to answer the research question on the cost-effectiveness of an intervention versus usual practice (or other), when the intervention is shown to be effective on weight status outcomes, but we do not propose this modelling framework for more detailed examination of the consequences of overweight and obesity.

The HeLP intervention has been shown to have positive outcomes in areas of secondary importance, and the overall feedback and satisfaction with HeLP from the school setting is very positive. That there is no specific policy-relevant evidence to indicate it is cost-effective, when considering weight status, points commissioners to the dominance of usual practice. However, given that there is a wide range of other interventions being used across the school setting, in the absence of evidence of their effectiveness and cost-effectiveness, it may be that commissioners consider the potential use of HeLP, or HeLP-like interventions, in a school-based setting, and if so here we have provided strong evidence on the resource use and cost associated with delivery of HeLP, and a framework that may be used in any further future assessment of HeLP or similar interventions.

# Chapter 5 Process evaluation

## Introduction

The overarching aim of the process evaluation for the cluster RCT of HeLP was to contextualise the trial effectiveness and cost-effectiveness results with respect to how the intervention may or may not have worked.

### Aims

There were two main aims of the process evaluation:

1. to assess the uptake and fidelity of the intervention
2. to assess whether or not the intervention worked in the way it was expected to in terms of the intervention logic model.

### Research questions

To address the aims of the process evaluation, we devised the following research questions.

Aim 1: to assess uptake and fidelity of the HeLP intervention:

1. How much of HeLP did children and families receive?
2. Was the programme delivered as designed?

Aim 2: to assess whether or not the intervention worked in the way it was expected to in terms of the intervention logic model (*Figure 8*):

3. Did schools, children and parents engage with HeLP?
4. How were the attempts to change behaviours experienced by the children?
5. Were there reported changes to children's eating and physical activity behaviours?
6. Do statistical models combining potential cognitive and behavioural changes mediate observed between group differences in outcomes?

### Logic model

The logic model for HeLP (see *Figure 8*) attempts to visually represent the theoretical underpinnings of the intervention, the content of the intervention, the process by which the intervention was assumed to work, the context in which the intervention was delivered and the outcomes it was hoped would be achieved. From the outset we wanted to develop a programme that had an impact on both the school and the family environment, as well as affecting individual processes related to health behaviours.[122] The underpinning assumptions were to create supportive environments for healthy lifestyles choices and positive relationships at the level of the school, child and family. In line with the World Health Organization's Health Promoting Schools framework, we aimed to develop activities that had an impact on the school environment and the curriculum, and that linked to families.[21] The delivery methods used were highly interactive and encouraged identification with and ownership of the healthy lifestyle messages so that the children were motivated to take them home to their parents and effect change within the family.

The HeLP intervention used the IMB model[123] as a guide to selecting cognitive and behavioural antecedents targeted by the intervention, antecedents that could lead to behaviour change. Intervention activities were then ordered to enable, support and sustain behaviour change in accordance with the health action process model.[26]

| Resources | Activities | Short-term outcomes | Medium-term outcomes | | Long-term outcomes |
|---|---|---|---|---|---|

Drama group (four actors and a drama facilitator per Healthy Lifestyle Week)

HeLP co-ordinator ×4

Multidisciplinary research team

Project Advisory Group (stakeholders)

Local network of primary school head teachers

Funding (intervention and research)

HeLP co-ordinator manual and resources

Actor manual and drama resource box

Teacher manual and resources

Training week for pool of actors and facilitators

Ongoing training for the HeLP co-ordinators

Regular HeLP team meetings

HeLP intervention (four phases over three school terms)

Establish relationships (schools, children and parents)

'Buy in' (schools, children and parents)

Awareness of HeLP messages (schools, children and parents)

Buzz and excitement in school about HeLP (school, children and parents)

Strengthen relationships (schools, children, families)

Skill building (children)

Confidence building (children)

Increasing knowledge (children, parents and teachers)

Increasing self-awareness (children, parents, teachers)

Increase social support (peers, school, and family)

Talking to parents about healthy lifestyles (children)

Reflect on snacking and physical activity behaviours (children and parents)

Setting goals and self-monitoring (children)

Resisting temptation children (children)

Buy healthy snacks and drinks and make them available (parents)

Shopping, cooking and trying new healthy food (children)

Reinforce rules around screen time and around snacking (parents)

Monitor child behaviours (parents)

Adapt goals (children)

Reinforce behaviour change (teachers, head teacher)

Increase physical activity (children and family)

Decrease sedentary behaviours (children and family)

Maintenance of increased consumption of healthy snacks/ positive food markers (children and family)

Maintenance of decreased consumption of energy-dense snacks/ negative food markers (children and family)

Family leads a more healthy lifestyle generally

Improved quality of life

Reduced incidence of obesity

**FIGURE 8** The logic model for HeLP.

The process evaluation explored whether or not the intervention worked in the way it was expected to by carrying out a thematic analysis,[124] using a Framework approach,[125] of the qualitative data collected from focus groups with children and interviews with teachers and parents, as well as a mediation analysis[126] using structural equation modelling (SEM)[127] with the MLQ data from both the intervention and the control group.

### General methods

The process evaluation of the HeLP study distinguishes between methods and analyses relating to (1) the intervention arm only and (2) the overall trial (including data from intervention and control arms). The analytical approach for (1) is hypothesis-raising only, synthesising data from field notes, observations, interviews, focus groups and questionnaires and for (2) it is hypothesis testing using a meditation analysis. For ease of reading, this chapter is presented in two sections. The first presents all process data collected from the intervention arm of the trial and the second presents the mediation analysis, which includes data from both the intervention and the control arm. *Table 36* provides a summary of the data sources for each process evaluation research question and the process evaluation dimension it assessed.

## Section 1: process data collected from the intervention arm of the trial

### Methods

#### Sampling and recruitment
This involved collecting data from Year 5 teachers, children and parents of participating children.

**TABLE 36** Summary of research questions for the process evaluation

| Research question | Process evaluation dimension assessed | Data source |
|---|---|---|
| How much of HeLP did the children and families receive? | Uptake | Child and/or family registers for each component |
| Was the programme delivered as designed? | Delivery: fidelity to form and function | Observation checklists |
| | | Field notes |
| Did schools, children and parents engage with HeLP? | Engagement | Observations |
| | | Field notes |
| | | Parental signature on goal-setting sheet |
| | | Qualitative evaluation (interviews and focus groups with teachers, children and parents) |
| | | Parent questionnaire |
| How were attempts to change behaviours experienced by the children? | Experience | Qualitative evaluation (interviews and focus groups with teachers, children and parents) |
| | | Parent questionnaire |
| Were there reported changes to children's eating and physical activity behaviours? | Impact | Qualitative evaluation (interviews and focus groups with teachers, children and parents) |
| | | Parent questionnaire |
| Do statistical models combining potential cognitive and behavioural changes mediate observed between group differences in outcomes? | Statistical mediators of group differences in outcomes | Variables derived from measures included in the MLQ for intervention and control children |

All Year 5 teachers from each of the 16 intervention schools participated in a semistructured interview. Children were selected to participate in a focus group by the HeLP co-ordinator based on the child's level of engagement (see *Observations*). The HeLP co-ordinator selected groups to ensure that there were equal numbers of boys and girls when possible. Each school had at least two focus groups of between six and eight children (one with children who were categorised as being engaged and one with children who were less engaged).

All parents of participating children were sent a questionnaire about the programme directly to their home address, along with a stamped addressed envelope. Parents were invited to participate in a semistructured interview, and they indicated their willingness to do so by completing a section at the end of the questionnaire.

### Data collection
All data were anonymised, and any comments or observations relating to specific individuals or schools in a way that could allow them to be recognised were removed.

### Parent questionnaire
A parent questionnaire (see *Appendix 10*) was distributed to all parents in December 2013/14 after the end of the intervention. The aim of the questions was to obtain data on any discussions that parents had had with their child about healthy lifestyles; their knowledge about the programme's aims; their awareness of their child's goals and how easy or difficult they felt it was for the child to achieve them; and details of any changes made to behaviours at a child level and/or a family level and whether or not these had been maintained.

### Interviews
All Year 5 teachers in each intervention school participated in a 45- to 60-minute semistructured interview, on school premises, about their attitudes towards, and experiences of, participating in the HeLP intervention and any impact that they believed it to have had on them, the children and the parents, as well as on the school as a whole (see *Appendix 11*). All teacher interviews were carried out by the trial manager in June/July 2013/14 following phase 3 of the intervention and before schools broke up for the summer holidays.

Those parents indicating that they were happy to be interviewed were contacted via e-mail and/or telephone by the HeLP co-ordinator to arrange the interview. The interviews were conducted either by telephone or face to face at either the school or the parent's home, whichever was more convenient for the parent. The HeLP co-ordinators carried out all parent interviews for their respective schools. The aim of the interviews was to obtain detailed data on parents' views of HeLP, discussions that took place at home about healthy lifestyles, the barriers to and facilitators of the achievement of personalised goals, and any child and/or family changes made as a result of the programme and whether or not these were being maintained (see *Appendix 12* for the parent interview schedule). These interviews were carried out between January and March 2014/15, during the spring term, after the intervention had finished.

### Focus groups
A selection of children from each school participated in a focus group that consisted of between six and eight children. The HeLP co-ordinator for the school carried out two or three focus groups per school, depending on school size. One focus group was carried out with children who were considered to be engaged and another was carried out with children considered to be less engaged children. Details of how these categories were given are presented in the following section (see *Observations*). The aim of the focus groups was to obtain detailed data on what children had learnt; whether or not they talked about healthy lifestyles at home with their family and peer group and what they discussed; what they thought about the programme and how they felt it differed from the usual curriculum; how they felt about choosing and working with a character like them; what it was like to try to set and achieve their goals; what strategies they used to help achieve their goals; and whether or not there had been any changes at a family level (see *Appendix 13* for the focus group schedule).

Focus groups took place during lesson time when the children were in Year 6 (between February and April in 2014 for cohort 1 and in 2015 for cohort 2). They were led by the HeLP co-ordinator for that school and facilitated by an additional HeLP co-ordinator, who took notes and supported the management of the group. To help children remember the details of the programme, visual cues were provided along with a short summary of the activities they participated in for each phase.

## Observations

Observations of intervention components were carried out to obtain data on fidelity to form (i.e. were all of the HeLP components delivered in the correct order?) and fidelity to function (i.e. were key components delivered in the 'spirit' of HeLP?). To assess fidelity to form, a yes/no checklist was completed (by the HeLP co-ordinator) for all HeLP components to indicate whether or not a component had been delivered (see *Appendix 14* for an example checklist).

'Spirit' was defined as enthusiastic delivery, open body language, responsiveness to child/school needs, and clear and friendly communication. The key components observed to assess fidelity to function were the parent assembly (phase 1), the healthy lifestyles week (phase 2), the parent assembly (phase 3) and the class-delivered assembly (phase 4).

To assess fidelity to function, a score between 1 and 10 was given for (1) delivery, (2) child responsiveness, (3) parent responsiveness and (4) teacher responsiveness for each of the four key components observed. At the beginning of data collection, the trial manager and the principal investigator independently scored fidelity to function for the parent assembly (phase 1) across three schools. No discrepancy in scoring was observed. Thereafter the majority of observations were carried out by the trial manager. The HeLP co-ordinator assessed the majority of the healthy lifestyle week components (phase 2) after they had carried out initial assessments alongside the trial manager. Once again, no discrepancies in scoring were observed (see *Appendix 15* for an example checklist).

Each HeLP co-ordinator also collected informal observational data, in the form of field notes (see below), for child and school engagement. Assessment of a child's engagement was based on the interaction of the HeLP co-ordinator with the child during the one-to-one goal-setting interview and from more general observations during the course of the intervention. The HeLP co-ordinator gave each child an engagement score between 0 and 3. The criteria for scoring were:

- 0 = uninterested/unaware goals needed to be set
- 1 = reluctant/needs a lot of prompting
- 2 = enthusiastic and happy to chat about goals and how they will achieve them
- 3 = very enthusiastic; has discussed goals at home and has clear strategies for achieving them.

School-level engagement was assessed using three scores based on the HeLP co-ordinator interaction with and observations of the head teacher, the Year 5 teacher(s) and the school support staff.

A score between 0 and 3 was given to each staff member:

- 0 = unengaged/unco-operative
- 1 = supportive
- 2 = enthusiastic and supportive
- 3 = very enthusiastic and used HeLP in other aspects of teaching/school activities.

## Field notes

Each HeLP co-ordinator kept recorded notes in a diary of their informal interactions with and observations of staff and children during the intervention, which fed into their assessment of staff and child engagement. In addition, the HeLP co-ordinators recorded any unintended consequences of the programme.

## Registers of attendance

The HeLP co-ordinator kept registers of attendance for all intervention components and for parental attendance at parental engagement events (i.e. parents invited to the school to observe activities). The Year 5 teachers were asked to keep a register for each personal, social and health education lesson that they delivered during the healthy lifestyles week, which was given to the HeLP co-ordinator.

## Parental signature

In phase 3 of the intervention, children were asked to set goals at home with their parents on a goal-setting pro forma before they had their one-to-one discussion with the HeLP co-ordinator. Parents were asked to sign the form once discussions had taken place. All goal-setting pro formas were collected by the HeLP co-ordinator following the one-to-one discussion, and a copy of the goals was sent back to each parent/carer.

## Analysis

All qualitative and quantitative process data were analysed blind to trial outcome and, initially, analysed separately. The different data sources were then combined to address each research question. Details of the analyses and subsequent synthesis are described in the sections below.

## Data from registers, parent questionnaire and goal-setting sheets

Child and parental attendance at events, parental signature on the goal-setting sheet and both quantitative and more open qualitative responses from the parent questionnaire were entered into a Microsoft Access® (2014 version, Microsoft Corporation, Redmond, WA, USA) database. The data were then exported into Microsoft Excel® (2014 version, Microsoft Corporation, Redmond, WA, USA) and imported into NVivo (version 11, QSR International, Warrington, UK) or Stata.

The parental engagement score was measured using two sources of data: attendance at one or more parent events and/or signature on the goal-setting sheet. A score between 0 and 2 was given to each parent:

- 0 = did not attend any activity/did not sign the goal-setting sheet
- 1 = attended one or more events or signed the goal-setting sheet (but not both)
- 2 = attended one or more events and signed the sheet.

Parents were then dichotomised into two groups ($\geq$ 1 = engaged parent and < 1 = less engaged parent).

## Qualitative data (interviews and focus groups)

Interviews and focus groups were digitally recorded and verbatim transcripts were prepared from the sound files. The transcripts were checked for accuracy against the sound files and corrections were made if required. Any comments that could identify people or schools were anonymised before the transcripts were imported into NVivo. A deductive approach was used to code the data. Four transcripts from the cohort 1 focus groups were selected (two engaged and two less engaged) from each HeLP co-ordinator's group of schools; these were read by the cohort 1 HeLP co-ordinators and the trial manager, and an initial code framework was agreed. Minor edits were made to the topic guide for the cohort 2 focus groups. A similar process of coding was used with the cohort 2 focus groups, and the initial codes were discussed, refined or amended and a new coding framework produced by HeLP co-ordinators and trial manager. The HeLP co-ordinators and one independent researcher then used this coding framework to code the remaining focus groups. The parent and teacher interviews were coded using a similar approach, with the coding framework being edited at each stage. The trial manager coded 20% of the transcripts, with the principal investigator providing verification for half of the 20% checked by the trial manager. The codes were then categorised (second cycle coding) to identify emerging themes and subthemes. Data from all sources (parent and teacher interviews and focus groups) were collated for each theme/subtheme and transferred into tables. The resulting tables were then analysed for agreement, partial agreement, silence or dissonance from the different data sources.

### *Observational data (checklists and field notes)*

**Delivery to function**  The four scores (enthusiastic delivery, engaged/participating children, engaged parents and engaged school staff) per observation in each school were averaged (mean score) to create a single score out of 10 for each of the four components observed (the parent assembly in phase 1, the healthy lifestyles week in phase 2, the parent assembly in phase 3 and the class-delivered assembly in phase 4). These four scores were then averaged again to produce a single delivery mean score per school. A score of ≥ 8 was prespecified to indicate that the intervention had been delivered as designed (i.e. in the 'spirit' of HeLP).

**Delivery to form**  If all components for each phase were delivered (represented as a tick on the checklist), then it was recorded that 100% of HeLP components had been delivered in that school. If a minor activity (e.g. the practise of a chant or a missed scene from a drama workshop) was missing from a particular component, 2% was deducted for this from the overall total (see *Table 38*).

**Child-level engagement**  Individual child scores were dichotomised to create two groups (≤ 1 = less engaged children and > 1 = engaged children).

**School-level engagement**  Individual scores for the head teacher, the Year 5 teacher and the support staff were aggregated to give a score out of 9 for each school. Schools were then dichotomised into two groups (0–3 = less engaged school and 4–9 = engaged school).

The handwritten field notes were typed up by each HeLP co-ordinator and entered into Microsoft Excel and then imported into NVivo once any comments that could identify people or schools had been anonymised.

### *Results*

The results from this section of the process evaluation are presented under each aim and their associated research questions.

## Aim 1: to assess the uptake and fidelity of the HeLP intervention

- Research question 1: how much of HeLP did children and families receive?

Across cohorts 1 and 2, 676 children were randomised to receive the intervention. *Table 37* shows the percentage of children participating in each phase of HeLP and the percentage of children receiving the four drama sessions in phase 2 and the one-to-one goal-setting in phase 3 (considered to be the key components of the intervention essential for behaviour change to occur) delivered in the manner in which HeLP had been designed.

**TABLE 37** Uptake of HeLP

| Cohort | Phase (%) | | | | Percentage of children receiving four drama sessions (phase 2) and goal-setting (phase 3)[a] delivered in the spirit of HeLP[b] |
|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | |
| 1 | 91.2 | 94.1 | 91.1 | 92.1 | 93.7 |
| 2 | 94.7 | 93.7 | 92.5 | 91.4 | 92.7 |
| Total | 93.4 | 93.9 | 92.0 | 91.6 | 93.0 |

a  Dose of HeLP deemed to be essential for behaviour change.
b  Enthusiastic delivery, open body language, responsive to child/school needs, and clear and friendly communication.
Reproduced from Lloyd *et al.*[32] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

Across the programme there were six invitations for parents and carers to come into the school and take part in the programme. These included the parent assembly and observation of the two activity workshops in phase 1, observation of work in progress in the final two drama sessions of the healthy lifestyles week in phase 2 and the forum theatre assembly in phase 3 (see *Tables 2* and *3*).

Just over half of children (52%, 353/676) had family attending at least one parent event.

- Research question 2: was the programme delivered as designed?

*Table 38* shows that we can be confident that all intervention schools received a complete or near-complete programme that was delivered as designed (i.e. all schools scored > 8 for fidelity to function).

### Aim 2: to assess whether or not the intervention worked in the way it was expected to in terms of the intervention logic model (see *Figure 8*)

All Year 5 teachers from the 16 schools that received the intervention were interviewed ($n = 28$) and all parents of participating children in the intervention schools were sent a questionnaire in the post following the intervention. Just over one-quarter of parents (26%, 176/676) returned the questionnaire, of which 80 (45%) indicated that they were happy to be interviewed. As there is some evidence that health promotion programmes can (unintentionally) widen health inequalities, we wanted to see whether or not engagement with HeLP and the trial process was the same across all socioencomic groups. Analysis by individual level IMD rank (i.e. based on the child's postcode) showed that there was an equal spread across the socioeconomic spectrum for questionnaire returns. Fifty-two parents (of which two were fathers) were interviewed, with slightly more parents from the higher two than the lower two IMD quartiles participating (60% and 40%). Forty-five parents (87%) interviewed were categorised as engaged, 81% of whom also had an engaged child. Thirteen per cent of parents interviewed were less engaged, of whom 2% (one parent) had a less engaged child.

**TABLE 38** Fidelity of delivery of HeLP (form and function)

| School [cohort 1 (1–8); cohort 2 (9–16)] | Per cent of components delivered in complete form (fidelity to form) | HeLP delivery score (fidelity to function) (maximum score of 10) |
|---|---|---|
| 1 | 100 | 8.9 |
| 2 | 100 | 9.0 |
| 3 | 100 | 9.1 |
| 4 | 100 | 9.8 |
| 5 | 96 | 9.7 |
| 6 | 94 | 9.4 |
| 7 | 96 | 9.5 |
| 8 | 98 | 9.6 |
| 9 | 94 | 9.0 |
| 10 | 94 | 9.1 |
| 11 | 96 | 9.4 |
| 12 | 96 | 9.7 |
| 13 | 98 | 9.2 |
| 14 | 96 | 8.1 |
| 15 | 100 | 9.7 |
| 16 | 94 | 9.2 |

A total of 35 focus groups (18 with engaged children and 17 with less engaged children) were conducted across both cohorts.

- Research question 3: did schools, children and parents engage with HeLP?

Child, parent and school engagement scores, as well as the qualitative data from the focus groups and interviews relating to enjoyment and engagement of the programme, are presented here. Evidence of possible mechanisms leading to engagement/enjoyment (e.g. knowledge and understanding, discussions at home about healthy lifestyles, delivery methods and identification with the characters) is also presented here. Each quotation presented is referenced with the source (school number, P = parent, T = teacher, LEC = less engaged child, EC = engaged child). We also present relevant data from the parent questionnaire (see *Appendix 10*).

### Child engagement

Ninety-six per cent (652/676) of children set goals with the HeLP co-ordinator in phase 3 and, of these children, 63% (411/652) had parental support, as indicated by a parent's signature on the goal-setting sheet and/or written comments regarding how the parent would support the child in achieving their goals.

Twenty-four children had missing engagement scores (13 children had moved out of the area, eight children had changed schools before the one-to-one goal-setting discussion and three children were absent on multiple visits by the HeLP co-ordinator) and had not set goals.

Based on the child engagement scoring system, 92% (602/652) of children were deemed to be engaged with HeLP. Similar percentages of boys and girls were considered engaged (91% and 94%, respectively); however, those children in schools that had more than one Year 5 class had a greater percentage of engaged children than those with only one Year 5 class (97% and 82%; respectively). *Table 39* shows that HeLP was able to engage children across the socioeconomic spectrum, although there were slightly more children from the most deprived quartiles in the less engaged category.

There was very clear evidence from all sources (teachers, parents and children) in the interviews and focus groups that children really enjoyed and engaged with all aspects of the the programme across all schools. All children spoke postively about HeLP:

> *I thought it was much more different than I thought it would be, because I didn't really exactly imagine it as a boring old literacy lesson or anything but I didn't really think it was going to be that joyful and exciting. It was much more than I expected.*

> *Female LEC, school 14*

**TABLE 39** Child engagement by IMD rank[a]

| Deprivation quartile | Number (%) of less engaged children | Number (%) of engaged children | Total number of children |
|---|---|---|---|
| 1 (most deprived) | 16 (33) | 156 (26) | 172 |
| 2 | 15 (31) | 143 (24) | 158 |
| 3 | 8 (16) | 147 (24) | 155 |
| 4 (least deprived) | 10 (20) | 155 (26) | 165 |
| Total | 49 | 601 | 650 |

a Two children could not be included in the analysis of engagement by IMD rank as we did not have their postcodes.

*Amazing, fun, healthy, extraordinary and the best!*

*Female EC, school 7*

*It was brilliant it was such good fun; the children reacted to it really positively. In fact I have not heard them say anything negative about it at all. They've never said 'oh no we're doing this again' they've always been really open and engaged . . .*

*T, school 15*

*She was coming home and really enthusiastic and sort of you know explaining to us about what she'd made in school and wanted us . . . wanting us to try it, to do it and share it with her, so you know I think that certainly helped a lot you know and . . . she wanted us to be a part of it.*

*P, school 4*

The healthy lifestyles week seemed to be the most enjoyable aspect, with the majority of children reporting that they liked the interactive and dynamic nature of the sessions and working with the actors and characters, with whom they identified:

*We had lots and lots of fun and everyone was getting excited for the next day and the next and the next, and for that week of activities, I was really excited about actually going to school, usually I'm just 'oh I want to stay in bed' but that week I was really excited about going to school.*

*Male LEC, school 12*

*The reason I liked the Healthy Lifestyles Week was because you were actually seeing what, sort of like a made-up version of four different people who have trouble and the ways you can improve it by just following them.*

*female LEC, school 12*

*Um, my favourite part . . . was probably in the drama was like when we were, because it was so totally different to what we normally do. In school we obviously do literacy and maths when we have to write stuff down but you could really express your emotions through drama, I really liked that.*

*Female EC, school 6*

*Awesome, it was really fun though because we got to act and then they told us what was in stuff.*

*Male EC, school 11*

However, a minority of children were less keen on the acting aspect:

*I didn't like doing the acting but it was OK when we were doing the non-acting.*

*Male EC, school 3*

*Well they were really good but when the characters asked us to come up do a bit of acting and stuff it was a bit embarrassing, because I put my hand up and I didn't know what I was in for.*

*Female LEC, school 9*

The teachers reported that the delivery of the healthy lifestyles week and the activity workshops engaged the children and, in some cases, even the ones who were usually shy or disruptive:

*I thought that the cleverness of the drama was that they did it from a child's perspective. They acted out as children, but they were like the cool teenage children which children can relate to. Interestingly enough, there are several children in my class that could switch off so easily but they didn't, because they immediately were drawn in to this character, this role play character. I thought the idea of having*

*the Duncan and the Active Amy and all the rest of it, initially when you heard it you thought oh this isn't going to work, and then you saw them go straight in to role, the acting was fantastic . . .*

*T, school 13*

*Yes, yeah there was one boy who would, well no two of them in fact that would never ever ever get up and do any drama or anything and they were up taking part in everything. I was so touched with emotion I had to run out and tell mum at the end of the day. So I think that it bought him out 'cause he, he was one of the ones that did it in the hall as well.*

*T, school 10*

*And the acting just lifted it 100 times more and to them it was so important and to them . . . we did the lesson in the morning but they couldn't wait until the afternoon. They were so linked to those characters and it was such a clever thing.*

*T, school 15*

In the parent questionnaire, almost two-thirds of parents reported that their child had talked a lot about the programme at home (see *Appendix 10*). Parents who were interviewed reported that their child was enthused and motivated to make changes, and that discussions had taken place at home about what they had been doing in school:

*She was always coming back and telling me what had been discussed. One of the main things, she was talking about exercise and that was one of the topics at the time and she was coming back and she said that part of her plan was to spend less time looking at the TV or on the computer and spend more time playing, and she went straight out on the trampoline with her little brother and since then – it's like, she's out now.*

*P, school 3*

*Yeah, he came home with oh saying these are the things that we should be looking for on the shopping list [Interviewer: OK] he was saying about how, you know, what was good to eat and what's not good to eat and he was particularly talks about it when we all sit round for tea, 'cause when we – I make sure that at the end of the day, we are all sat together eating the evening meal together [Interviewer: Yeah] so, I think that's really important cos I think you can all talk and just . . .*

*P, school 2*

*Yeah they'll stay with him for life 'cause he talks about it a lot [Interviewer: Excellent] and you know I think that they have continued to talk about it at school and he's taken it on board.*

*P, school 7*

Children's knowledge and understanding of the key messages were considered to provide further evidence that children had engaged with the programme; there was also evidence from the focus groups of their understanding of the messages around marketing, moderation and food labelling:

*What helped me a massive amount was looking at the ingredients and looking at what's inside stuff, like if it says like fruit on it, it may not actually be made out of fruit. Like fruit winders and stuff and some stuff and my mum said like if they make the big front of the packet really appetising and want to make you feel like you want to buy them but then the back is like all small and you can hardly read it so they are trying to trick you to get the really unhealthy stuff but make it look really appetising.*

*Male EC, school 16*

*I think my favourite part about it was doing the food machines. I especially liked how they used acting to show how the foods were made and what process they go through.*

*Male LEC, school 12*

*Yeah I just think it's really raised their awareness, yes the 80 : 20 so yes and he did talk about that with his brother and sister and you know that was the bit that really raise his awareness. You know impacted us . . .*

*P, school 14*

### Parent engagement

Based on the parental engagement scoring system, just over three-quarters (77%, 520/676) of parents were deemed to have engaged with HeLP. *Table 40* shows a similar pattern of parental engagement and IMD rank to that observed for child engagement, with more parents from the lowest IMD ranking considered to be 'less engaged' than the highest IMD ranking, although similar proportions were engaged with the programme.

Of the 92% of 'engaged' children, the majority had parents who were also considered to be engaged. Of the 130 parents who were considered less engaged, 20 also had less engaged children, while the other 30 less engaged children had engaged parents.

The majority of parents interviewed spoke positively about the programme, reporting that it was a worthwhile project and supported the messages it was trying to promote at home. However, we do acknowledge that 87% of the parents interviewed were categorised as engaged, and thus were probably less likely to be critical in their interviews. The majority of responses from the parent questionnaire were very favourable towards HeLP, with 99% of parents indicating on the parent questionnaire that they were aware of their child's goals:

*. . . it gives me more reinforcement so when I feel like I'm nagging her about you know you really need to eat a wider range of food and la la la, then she'll kind of go well yeah 'cause you're supposed to have blah blah. And it's good 'cause she's learnt it from somewhere else. And if she's learnt it at school it must be true! [Laughing] Whereas if you learn it at home you can take it or leave it, so it is good reinforcement.*

*Engaged parent, school 10*

*One of the good things I liked about the programme was that you weren't giving out lists of things that were quite expensive or difficult for people to buy or achieve um and I've seen that before and I find that frustrating and people can't always afford to get you know certain fruits and vegetables and food products. It's not always easy for them to do that and a lot of the food suggestions that you made particularly in terms of 'replace this with this' were simple and achievable . . .*

*Less engaged parent, school 16*

**TABLE 40** Parent engagement by IMD rank

| Deprivation quartile | Number (%) of less engaged parents | Number (%) of engaged parents | Total number of parents |
|---|---|---|---|
| 1 (most deprived) | 45 (35) | 127 (24) | 172 |
| 2 | 31 (24) | 127 (24) | 158 |
| 3 | 25 (19) | 130 (25) | 155 |
| 4 (least deprived) | 29 (22) | 136 (26) | 165 |
| Total | 130 | 520 | 650 |

Similar to what was seen in the children, engagement was also evidenced by parents' knowledge and understanding of the key messages:

*Yes we'd look at the cereal boxes and say 'right better not have that cereal, that's got too much percentage fat' I can't remember what [Interviewer: Sugar I think is it, yeah] sugar yeah um 'we'll have that one instead'. Yeah so it did make everybody think 'ooh!'. Check on stuff that you think some things you think are healthy are actually, got a lot more sugar content than you expected.*

*Engaged parent, school 14*

*No I think it's been a really good programme actually and just to raise awareness of you know what you need to do to be healthy and you don't have to stop all of those things because ultimately people think 'oh you can't have this and you can't have that' and they need to know that there's a balance actually and that you can have it you just need to be a little bit aware of what you're eating and yeah so I think's it's been brilliant and I wouldn't change any of it . . .*

*Engaged parent, school 16*

However, 18 parents reported some negative feelings towards certain aspects of the programme in their questionnaire responses. One parent, who felt that their child was a fussy eater, had a smaller range of foods to choose from as he wanted to cut down on less healthy options. A couple of parents expressed concern about the cost of fruit and vegetables and a few parents reported that their 'already healthy' children took the programme messages to the extreme. One parent felt that it was the responsibility of the parents to educate their children about physical activity and health.

A small number of data from teachers also suggested that a few parents were less engaged with the programme. Two teachers from two different schools reported disappointing parental attendence, and in one school (school 7) a teacher informally reported to the HeLP co-ordinator that a parent had talked to the teaching assistant about not wanting to be lectured at about how to look after her child. Another teacher from a different school (school 9) reported that she had heard from a minority of children that their parents had thrown the project leaflets in the bin.

## School engagement

School engagement scores ranged from 9 (maximum score) to 2. Out of the 16 schools, only three were categorised as 'less engaged' according to the school engagement scoring system (schools 1, 3 and 14). Characteristics of the less engaged schools included lack of senior leadership within the school generally, absent teacher during of the drama workshops and absent Year 5 teacher owing to illness during phases of the programme.

There was nothing to suggest that variation in school engagement was linked to differences in school context, or school characteristics and policies as assessed by the School Characteristics and Policies Questionnaire (see *Appendix 16*). *Table 41* provides a summary overview of each school by cohort in relation to size, deprivation, location and engagement.

## Teacher engagement

Regardless of the HeLP co-ordinator's perception of teacher engagement (see *Table 41*), all Year 5 teachers spoke positively about the programme in the interviews and there was strong evidence that teachers were engaged with it. Many believed that the programme's compatibility with the national curriculum made it feasible to deliver and helped to reinforce healthy lifestyle messages that the curriculum already promoted. All teachers reported that the HeLP co-ordinator assigned to their school provided excellent ongoing support, which really helped in the delivery of the programme:

*She's [HeLP co-ordinator] ever so good, she's obviously worked with teachers quite a lot I think, she knows how it is. I don't have to say 'oh because we're doing this' she says 'oh yeah yeah I realise so you know she's really really good at fitting the scheme, we're trying to fit in with you' and she helped*

TABLE 41 School engagement information

| School [cohort 1 (1–8); cohort 2 (9–16)] | Number of Year 5 classes | Free school meals, % (national average 19%) | Urban/rural | School staff engagement score | Overall school engagement score |
|---|---|---|---|---|---|
| 1 | 1 | < 19 | Urban/rural | Head teacher = 2, teacher = 0, administrator = 1 | 3 |
| 2 | 1.5 | ≥ 19 | Urban | Head teacher = 1, teacher = 1, administrator = 2 | 4 |
| 3 | 1 | ≥ 19 | Urban | Head teacher = 0, teacher = 1, administrator = 1 | 2 |
| 4 | 2 | ≥ 19 | Urban | Head teacher = 1, teacher = 3, administrator = 2 | 6 |
| 5 | 1 | ≥ 19 | Urban/rural | Head teacher = 1, teacher = 2, administrator = 3 | 6 |
| 6 | 1 | ≥ 19 | Urban | Head teacher = 1, teacher = 2, administrator = 2 | 5 |
| 7 | 1 | ≥ 19 | Urban | Head teacher = 3, teacher = 3, administrator = 1 | 7 |
| 8 | 1 | < 19 | Rural | Head teacher = 2, teacher = 2, administrator = 1 | 5 |
| 9 | 2 | < 19 | Urban | Head teacher = 2, teacher = 3, administrator = 3 | 8 |
| 10 | 3 | < 19 | Urban | Head teacher = 1, teacher = 2, administrator = 2 | 5 |
| 11 | 3 | < 19 | Urban/rural | Head teacher = 1, teacher = 3, administrator = 2 | 6 |
| 12 | 2 | < 19 | Rural | Head teacher = 2, teacher = 2, administrator = 2 | 6 |
| 13 | 1 | < 19 | Urban | Head teacher = 3, teacher = 3, administrator = 2 | 8 |
| 14 | 1.5 | < 19 | Urban | Head teacher = 0, teacher = 1, administrator = 2 | 3 |
| 15 | 1 | ≥ 19 | Rural | Head teacher = 3, teacher = 3, administrator = 3 | 9 |
| 16 | 3 | < 19 | Urban | Head teacher = 2, teacher = 2, administrator = 2 | 6 |

*fit in with us so you know it worked really well. That flexibility is really important and you know she's brilliant and she knows her stuff and she knows . . . I've been really impressed she's been brilliant, really helpful and ever so amenable. We've had children with possible problems with their eating and you know she was brilliant talking to the parents . . . nothing was too much trouble it was like 'oh yeah I'll sit down and chat with them' and she was brilliant when she was sort of talking to the parents and I think it is really important that you have that point of contact. For me it was so easy that I could you know send an e-mail or [HeLP co-ordinator] would e-mail me and I could e-mail back and you know that's really really useful and very good and helps the thing run really well.*

*T, school 15*

Many teachers commented that the way in which the lessons were delivered and the personnel used for delivery really helped the children think more deeply about their health and how they could make small, achievable changes to their eating and actvity behaviours:

> *I think they really engaged with those lessons. You know I was thinking about some of the messages that were coming across from the actors, when we had the actors come in for that week you know we tell them those messages all the time and they don't sink in as much as they did then. And I think that coming from a younger perspective, the street dance obviously and [name of sports group] you know they nailed it really. Yeah they loved that and yeah I think you know its one thing having, I've got great relationships with the kids we all have but it's one thing having Mrs [XXXX] tell them this and it's quite another thing to experience it from younger people, or people they perceive to be younger. The actors were teenagers and they loved them. So yeah I think it had a deeper impact because it came from a different perspective. Than their parents or their teachers really.*
>
> *T, school 11*

● Research question 4: how were the attempts to change behaviours experienced by the children?

Whereas phase 1 and phase 2 of the HeLP intervention focused on introducing the messages and engaging children and parents with how changes could be made using the creative delivery methods discussed in *Chapter 2* (see *Study design*, *Intervention*), phase 3 focused on the children setting personalised goals around the key messages.

Ninety-six per cent of children set goals with the HeLP co-ordinator in phase 3 and, of these, 63% had parental support, which was indicated by a parent's signature on the goal-setting sheet and/or written comments regarding how the parent would support the child in achieving their goals. Children were encouraged to set three goals based on snacking, fizzy drink consumption and physical activity, depending on the area in which they needed to make changes. *Table 42* shows the number and percentage of goals set for each target behaviour by gender.

As *Table 42* shows, the majority of goals set by both genders were focused on increasing their physical activity and snacking. 'Other' goals included helping others to be healthy, going to bed earlier, studying, doing homework/household chores, trying new foods, eating more and drinking more water. No marked differences were observed in the type of goal set by gender or engagement.

Children reported mixed feelings about setting goals, with some finding it easy and others finding it more challenging for various reasons:

> *I felt setting the goals, um one of my goals is really easy to set but the other two were really hard. 'Cause one of them was do, ask my mum to get healthy stuff and we don't have that much money so we couldn't get any . . .*
>
> *Male EC, school 9*

**TABLE 42** Number and percentage of goals[a] set for each target behaviour by gender

| Goal type | Number (%) of goals set by | |
|---|---|---|
| | Boys | Girls |
| Snacking | 251 (26) | 276 (28) |
| Drinks | | |
|   Fizzy | 60 (6) | 74 (8) |
|   Other | 8 (1) | 16 (2) |
| Physical activity | 391 (40) | 336 (35) |
| Other | 265 (27) | 267 (28) |

a  Each child could set a maximum of three goals. Overall, 1944 goals were set (975 by boys and 969 by girls).

*I think because my mum understood it, like [Interviewer: Go on] I don't really know, umm, my mum kind of knew about it, so if I went to speak to her then it was quite easy because she knew what I was talking about . . .*

*Female LEC, school 1*

*I found it quite easy because my mum was really supportive.*

*Male E, school 13*

*I didn't have much, well I did talk to my mum but I didn't have much time to talk to my mum because my brother and sister needed mum too. My brother needed help with his homework, my sister needed help with her homework, my brother needed to read to my mum and my sister needed to read to my mum, and my mum needed to sleep with my sister before going to bed and reading her a story, so it was a bit busy and I did have a quick little talk to her, but I didn't have much of a talk so it was a bit difficult and we only did about half of it.*

*Female EC, school 7*

*I found it quite hard, I thought it was a bit too open, like the sheets we took home, it was really open so it wasn't really, 'cause people could easily just put tick, tick, cross, tick, even if they weren't actually trying. And because we especially did it over the weekend and the holidays, I found that extremely hard because in the holidays and that, I kind of let myself go and like, have a lot of chocolate, that sort of thing. So if we had done it during the week I think I would have found it a lot easier, so maybe next time.*

*Male LEC, school 14*

*I thought it was hard 'cause my mum worked late, so did my dad, so I didn't have anyone to talk to about it.*

*Female EC, school 16*

In the parent questionnaire, 99% (174/176) of parents reported being aware of their child's goals, 59% (103/176) reported that they thought that it was 'quite easy' for their child to make changes based on their goals, and 20% (35/176) reported that they thought that it was 'quite difficult'. During interviews, parents reported how their children engaged with and took responsibility for their goals:

*Especially the school dinners because she does have a school dinner every day and she would always choose the chocolate cake or the pudding or something and then she, I didn't say it to her, she made the choice to go for the fruit salad or yoghurt or you know the healthy thing and every day she'd come back, I mean we have always said what have you had for lunch and she'd say whatever, and then she would say but I had fruit salad and she was very proud of herself when she managed to choose the healthy option.*

*P, school 9*

*Yeah so she really owned her goals as well. She was really bought in from the start. She was very responsible about it which was good because it gave her something to be responsible for and it was only kind of small changes so um . . . she's not allowed coke at home but at her dad's she was allowed coke. So she made the decision to go for fruit juice with fizzy water instead and that's carried on as well so um and she's . . . she was the one who kind of raised it there and she was making those choices herself it wasn't kind of like 'oh no I've got to do this or you must be doing this because it's on your goal' it was about 'yeah I need to do this because that's what I've said I'm going to do'.*

*P, school 16*

In the children's focus groups, both engaged and less engaged children reported mixed levels of difficulty in achieving their goals; some children reported a number of barriers that made it difficult to make changes and others found it easier to achieve their goals:

> *I found it difficult because I found out that actually I was quite busy, so I didn't really get time to pick when I was going to do it and choose what I ate. I kind of picked anything 'cause I didn't have the time.*
>
> *Female LEC, school 1*

> *Well I didn't find it difficult but it was hard to change it, because it was like a routine and trying to get them in, get away from the other stuff and trying to get new stuff.*
>
> *Male LEC, school 7*

> *Sometimes I found it hard, but sometimes it was quite easy because when I got more in to the programme I found it easier because I started doing it more, in the beginning I found it really hard to make any changes.*
>
> *Female EC, school 6*

> *I found two of them I think easier, but there was one which was really difficult, it was one which I could only have I think it was three sweets a week, and it got really difficult because it was the summer holidays and that's the time when you normally get the most sweets so it was quite difficult to keep that one. I think I kept it half the time and half the time I didn't.*
>
> *Female EC, school 7*

As the HeLP logic model (see *Figure 8*) illustrates, it was assumed that changes in children's eating and activity behaviours would be facilitated by a number of child and parent behaviours. Specific to the goal-setting phase, it was hoped that the children would use a number of strategies they had learnt during phase 2 to help them achieve their goals (e.g. putting their goals up on the fridge, or putting a note on their video game console to encourage them to limit their time). It was also hoped that parents would support them by increasing access to and availability of healthy snacks and physical activity, and enforcing rules about screen time.

When the children discussed their experiences of making changes in the focus groups, some commented that their parents had supported them in this way:

> *My parents helped a lot with my goals because my mum made a rota which is so I only get an hour or half an hour on screen time every day.*
>
> *Male EC, school 6*

> *Yeah I told my dad to stop giving me loads of sweets on Fridays, and then he did do that for me and now I don't eat as much sweets.*
>
> *Male LEC, school 8*

A few teachers also reported that parents were supporting their children:

> *I spoke to parents as well, parents have even said that actually they've gone home and they're getting out the humus and the carrot sticks rather than the crisps and the chocolate bars. So you know everyone's kind of seen a change and improvement. Which is great.*
>
> *T, school 11*

Many children commented on using a variety of strategies to help them achieve their goals:

> *With the magnet that you sent out, because I wasn't eating healthy so my mum put it at my height on the fridge.*
>
> *Female LEC, school 15*

*Well with mine, we photocopied our sheets so when I go round to my friend [name]'s house he's got a copy of my sheet and when he comes round to my house we've got a copy of his sheet. We photocopied it as well so my granddad's got a photocopy of my sheet so every time I go down there for lifeguard training I have my tea there and I have a look on the sheet and see what I can have 'cause we made another one and what I can eat if you get what I mean.*

*Male LEC, school 9*

*I found it like, we wrote a little note or I think one of my friends did it or we like look at the fridge with our magnet, the thing what you gave us, and then we, sometimes when I open the fridge I'm like, then I remember all the stuff I've been taught, and that's how I remember and how I achieve my goals.*

*Female EC, school 12*

We know that changes at a family level can have a positive impact on children making changes to their eating and activity behaviours,[128] and there was strong evidence for this from parents and children alike. The results of the parent questionnaire indicated that 74% of families had made changes (see *Appendix 17*), with 78% of parents providing further details. These included providing less of the unhealthy foods in the household: 'we don't buy sweets/crisps to be kept in the house' (parent of an engaged girl, school 6); 'fizzy drinks replaced with squash and sparkling water' (parent of an engaged girl, school 14). Parents also reported spending more family time doing physical activity: 'we spend more time at weekends going out. Sometimes cycling or walking or going on scooters' (parent of an engaged girl, school 9); 'extra family walks, general family health and fitness' (parent of an engaged girl, school 11); 'try not to have unhealthy snacks in the house, make conscious effort to get more exercise' (parent of an engaged boy, school 1).

This was also supported by the data from the parent interviews:

*I'd always felt really guilty about them eating loads of biscuits, it's just they are a really easy go-to food in the tin and kind of this the healthy eating lifestyle if nothing else it just made me really sit up and think about making changes as a family and so as a family you've got to change that and if we change nothing else you do need to change that.*

*P, school 13*

*Yeah definitely yeah because like you know if she would get a snack and she has a carrot, and then thinks oh yeah you will all have carrots, so in that respect you know she has pulled the whole family to a healthier snack rather than just herself. [Interviewer: OK] Yes and you know obviously when she does go outside she'll take at least one of her sisters or brother with her, so yes I suppose in that respect I . . . no so much as discussing it with them or talk to them about it just by her actions really . . .*

*P, school 1*

*It's mainly the 80 : 20 [Interviewer: Yep] and we still talk about it now um we, we always have healthy food because that's our 80 [Interviewer: Yeah] um and er we don't have, we, I now don't have any snacks in the house at all um where I would have done before. I think it's helped also that the children aren't having packed lunches anymore as well, that helps not having snacks in the house [Interviewer: Yeah] and if they want to go and get a snack then that's their 20 and they can walk to the shop and get it.*

*P, school 7*

*Yes we definitely decreased the fruit juice that we have. I think all of us. I think I knew but it's just having it reinforced 'cause you hear all sorts of things in the news.*

*P, school 10*

In the focus groups, children (both the engaged and the less engaged group) also reported family changes:

*Well after I started being more healthy, the whole family kind of started, so now there is no chocolate in the way so you won't be tempted and we buy more healthy things.*

*Female EC, school 6*

*I've noticed that normally before the Healthy Lifestyles Programme we used to have our cupboards stocked with different unhealthy foods like chocolate bars and cakes but now 75% of them have been replaced now.*

*Male LEC, school 9*

*Um yeah, we started doing more as a whole family together and going out more, like going on family walks on the weekend, instead of just having one person go out by themselves, and literally just going out all together.*

*Male LEC, school 12*

- Research question 5: were there reported changes to children's eating and physical activity behaviours?

The focus group and interviews suggested that children had made changes to their diet and/or activity/ sedentary behaviours:

*The one area that I did really notice that he picked up on is looking at ingredients of food. Looking at the percentage of fat in something and he'll come home from school and rather than just asking for a biscuit he'll go and get a huge carrot to have with it.*

*P, school 14*

*So um she does tend to eat more fruit now and sort of she's interested in trying new things and I think that's probably one of the best things that's come out of it [Interviewer: Ahum] we've always had a battle with [name] eating and that but since she's done the programme she's been willing to try.*

*P, school 4*

Teachers also reported being aware of changes the children had made:

*It was perfect timing, to have that farm trip after the 80 : 20 programme to see if they had adjusted their lunchboxes to be a little bit more, there was more water and less fizzy drinks, less sweets, more fruit. Because I went around looking at the packed lunches as we were sitting on these bales of hay, and it was a kind of relaxed fashion, and the children were just openly talking about it straight away. And I thought wow the impact has been good.*

*T, school 13*

The evidence from the focus groups and interviews of reported changes was supported by data from the parent questionnaire. Seventy-seven per cent of parents noticed changes in their child's choice of snacks and 60% noticed changes in their choice of drinks. Fifty-five per cent of parents noticed changes to screen time, with 67% reporting changes to physical activity. In response to the question asking whether or not any observed changes had been maintained, 78% reported that they had (see *Appendix 17*). When given the opportunity to provide details of these changes, 89% of parents responded and provided a brief description of the change they had observed in their child. The majority of responses focused on changes in snack and drink choice; for instance, 'X will now opt for a piece of fruit instead of a chocolate bar/ biscuit etc.' (Parent of an engaged girl, school 5) or 'not drinking coke anymore' (parent of an engaged girl, school 6). Parents also reported a raised awareness in their child around healthy choices: 'X is more aware of healthy choices and is much easier to divert to healthy alternatives' (parent of an engaged boy, school 10).

Parents identified increases in physical activity behaviour and their children taking up more clubs: 'she increased time on the trampoline/outdoors regardless of the weather' (parent of an engaged girl, school 11); 'understands more about the need to be physically active, i.e. joined the swimming club' (parent of an engaged girl, school 8).

Parents also reported changes in screen time, although to a lesser degree: 'screen time has reduced and sports has increased' (parent of an engaged boy, school 11); 'there was always a limit on screen play but we did not have to monitor it as much as he was more aware' (parent of an engaged boy, school 15).

### Adverse events

There was one adverse event reported during the 5-year trial. A HeLP co-ordinator received an e-mail from a mother of a child from intervention school 16, following the healthy lifestyles week, regarding her daughter's unhealthy exercising and eating behaviours. The trial manager spoke to the mother and followed up with an e-mail detailing the actions taken as a result of her concerns. The mother was happy for her child to continue with the programme, and the HeLP co-ordinator subsequently set goals with the child that did not focus on changing physical activity or dietary behaviours (which were already meeting the 80 : 20 balance). The HeLP co-ordinator informed both the school head teacher and the class teacher of this event and all were very supportive of the programme and the way the messages were delivered.

The trial manager interviewed this mother after the intervention was completed, and the latter said the following about this event:

> Towards the end of the programme what was very good was that the team spoke to her after we chatted, and made that clear to her and suggested that she come up with goals that would help the family and that was a really good thing to do as it enabled her to use her imagination and to take what she'd learnt and put it into practice but without inflicting on herself which she didn't need to do . . . Overall I think the message that they were left with was pretty good. I was happy that it kind of reinforced what we've always tried to teach them at home as well.

### Summary

The HeLP intervention was delivered with a very high degree of fidelity to content and in the spirit of HeLP; moreover, 95% of children received what was considered to be the necessary 'dose' for behaviour change. Prespecified assessment criteria of engagement and interview and focus group data from teachers, parents and children showed that HeLP appeared to engage schools, children and their families across the socioeconomic spectrum.

There was no evidence to support the possibility of differential take up or engagement with the programme by school, gender or child-level socioeconomic status. This may be due to the engagement scores we developed not being sufficiently nuanced to capture actual engagement, and data from parent questionnaires, interviews and focus groups all suggested high levels of engagement as well as reported changes in the targeted behaviours. However, there was evidence, from the focus groups in particular, that children experienced most difficulty in setting and achieving their goals if they did not have sufficient support from home. All teachers were very positive about the programme and had no negative comments about the delivery or the impact on their workload.

Despite the reported changes, the results of the trial showed no differences in physical activity or sedentary behaviours between the children in the intervention schools and those in the control schools, although a difference was observed in favour of the intervention fewer types of energy-dense snacks and negative food markers. There are very few data to explain why the reported changes did not translate into measurable differences, especially with regard to physical activity. This was the most common goal set by the children, although fewer examples of how this was achieved were reported in the interviews and focus groups. The most likely explanation is that the changes that were discussed in the focus groups, interviews and questionnaires were not of sufficient intensity or duration for their effect to be captured at the 18-month follow-up.

A limitation of the process evaluation was the low response rate (25%) from parents to the questionnaire (which also formed the pool for parent interviews), which limits the strength of the data in relation to reported behaviour change, as well as the ability of the data to represent all participating parents. Eighty-seven per cent of parents interviewed were considered to be 'engaged' with HeLP, based on prespecified criteria, compared with 77% of the whole sample, although their socioeconomic status was similar to that of all parents.

## Section 2: mediation analyses

### Background

The mediation analysis tested whether or not changes in targeted cognitions and behaviour patterns could statistically explain any observed between group differences in trial outcomes. Our analysis plan stipulated that we would assess potential mediators of any observed group differences in the main primary and/or secondary outcomes and, therefore, the mediation analyses were carried out once the results of the trial had been revealed.

Significant intervention–control group differences were observed on two self-report dietary outcome measures, namely the average number of different types of energy-dense snacks consumed during the week (0.4 mean difference in favour of the intervention) and the average number of different types of unhealthy foods consumed during the week (negative food markers), which was a mean difference of 0.6 in favour of the intervention group. Mediation analyses were conducted to explore whether or not changes children's self-reported knowledge, cognitions and behaviours could statistically account for these self-reported dietary differences.

- Research question 6: do statistical models combining potential cognitive and behavioural changes mediate observed between-group differences in the number of weekday energy-dense snacks consumed at 18 months and/or the number of unhealthy foods (negative food markers) at 18 months?

### The My Lifestyle Questionnaire

We developed the MLQ to capture possible mediators of diet and physical activity behaviours that we hypothesised would need to change to affect weight status (see *Appendix 3*). The MLQ development was based on the HeLP intervention mapping process, and the 50-item questionnaire comprised a number of sections designed to assess knowledge, individual cognitions and motivations, perception of parental behaviours, and the use of behavioural strategies and change techniques, as well as specific behaviours targeted by the programme (see *Appendix 18* for full details). Some MLQ items were theoretically derived and based on previous measures of theoretically defined constructs (such as attitude, injunctive and descriptive norms, self-efficacy and intention). Other items (including questions assessing knowledge and behaviours related to discussing healthy eating and physical activity at home with the family, helping to choose healthy options when shopping with the family and participating in cooking) were designed to assess changes in the short- and medium-term outcomes that HeLP was expected to generate.

### Methods

#### Sampling and recruitment

Data were collected from all participating children in both intervention and control schools (see *Chapter 2*) completing the MLQ at baseline and at 12 months post baseline (i.e. immediately after the end of the intervention).

#### Data collection

The HeLP co-ordinators collected the MLQ data during a dedicated lesson and read the questions from the front of the class, with the children completing the questionnaire at the same time and in silence. Clarifications were given for specific questions, and a poster was produced to demonstrate percentages so that the children were able to understand a question relating to proportions (question 4 in the knowledge section; see *Appendix 3*). Children who required additional support completed the questionnaire in a smaller group outside the classroom with an additional researcher.

## Analysis

### *Factor analyses to create composite My Lifestyle Questionnaire variables*

An exploratory factor analysis (EFA) was conducted on half of the baseline MLQ data set, followed by a confirmatory factor analysis (CFA) (in the remaining data set) to generate composite variables that loaded together and could be taken into the subsequent mediation analyses. We adopted a parsimonious approach to the construction of these variables, summarising across theoretically derived and HeLP-specific MLQ items (see *Appendix 18*).

*Table 43* shows the five statistically generated composite variables from the factor analyses, the MLQ constructs that they came from, the questions that fed into the constructs and the corresponding score ranges. The variable referred to as 'knowledge' was originally designed as the sum of five questions, with each question having several examples for the children to answer. We have called the four composite variables derived from the factor analysis 'confidence and motivation', 'peer norms', 'family approval/behaviours and child attitudes' and 'behaviours and strategies'.

**TABLE 43** Derivation of MLQ composite variables

| Composite variable (score range) | MLQ construct | Number of items | Question number in the MLQ | Minimum–maximum score |
|---|---|---|---|---|
| Knowledge (0–20) | Healthy snack/drink alternatives | 1 | 1 | 0–6 |
| | Food group proportions | 1 | 2 | 0–5 |
| | Lifestyle physical activity | 1 | 3 | 0–4 |
| | Energy balance | 1 | 4 | 0–2 |
| | Strategies for change | 1 | 5 | 0–3 |
| Confidence and motivation (9–36) | Self-efficacy to make healthy eating and activity choices | 3 | 6–8 | 3–12 |
| | Intentions to make healthy eating and activity choices | 6 | 9–14 | 6–24 |
| Peer norms (8–32) | Peer norms | 3 | 15–17 | 3–12 |
| | Peer approval | 5 | 18–22 | 5–20 |
| Family approval/behaviours and child attitudes (9–36) | Family approval | 3 | 23–25 | 3–12 |
| | Attitudes towards restrictions on behaviour | 3 | 26–28 | 3–12 |
| | Parental provision and rules around food and physical activity | 3 | 29,30,32 | 3–12 |
| Behaviours and strategies (18–72) | Goal setting | 6 | 33–38 | 6–24 |
| | Self-monitoring | 4 | 39–42 | 4–16 |
| | Discussion about healthy lifestyles with parents/family | 3 | 43–45 | 3–12 |
| | Encouraging the family to be more healthy | 2 | 46–47 | 2–8 |
| | Helping parents to choose healthy alternatives when shopping | 1 | 48 | 1–4 |
| | Helping with cooking at home | 1 | 49 | 1–4 |
| | Trying new, healthy food and drinks | 1 | 50 | 1–4 |

### *Structural equation modelling*

Structural equation modelling was used for the mediation analyses. SEM allows the assessment of the relationships among a set of variables using multi-item scales, multiple variables and multiple outcomes.[129,130] The variables included in the mediational analyses are baseline and 12-month BMI SDS, gender, school IMD score, class size and baseline and 18-month energy-dense snacks and negative food markers, as well as the five composite variables from the MLQ (see *Table 43*). The independent variable is the allocation group: intervention versus control.

There are two dependent variables: the number of weekday energy-dense snacks consumed per day at 18 months and the number of weekday unhealthy foods consumed per day (negative food markers) at 18 months. These were the secondary outcome variables from the main trial that had statistically significant between-group differences. We took a parsimonious approach to the modelling exercise and hence did not select the two other statistically significant secondary outcomes (the average number of energy-dense snacks and negative food markers foods consumed across the whole week at 18 months). Furthermore, we did not select healthy snacks at the weekend as the adjusted between-group difference from the linear random-effects regression model was not statistically significant (mean difference 0.23, 95% CI –0.04 to 0.50) [see *Chapter 3*, *Primary analyses of secondary outcomes at 24 months (anthropometric measures only)*].

Eleven variables were controlled at baseline. These included baseline scores for the outcome variables and the potential composite mediating variables, as well as four demographic measures – gender, BMI SDS, school size and school-level deprivation – that could have influenced (enhanced or diminished) scores on the dependent outcomes.

Variables in the path model to be tested using SEM are shown in *Figure 9*.

Preparation of the data included examination of, and possible replacement of, any missing values. For the 'knowledge' items, if a child did not answer an item it was assumed that they did not know the answer and were scored zero for that item. For the composite variables that contained Likert scales (see *Table 43*),



**FIGURE 9** Pathway analysis conceptual model. B&S, behaviours and strategies; C&M, confidence and motivation; EDS, energy-dense snacks; FAB&CA, family approval/behaviours and child attitudes; NFM, negative food markers; PN, peer norms; SES, socioeconomic status.

the following rules were used to replace missing values: if up to 75–80% of items were available for a particular variable, this was averaged, and the rounded average was used to replace the 20–25% of missing values, thereby retaining the ordinal nature of the individual items.

The model shown in *Figure 9* was tested twice: once for the outcome variable weekday energy-dense snacks at 18 months and once for weekday negative food markers at 18 months. According to our proposed model, the intervention may have a direct effect on the outcome variable at 18 months but also indirect effects through changes in the five composite mediating variables at 12 months.

Structural equation modelling[127] was used to test this model, specifically whether the mediating variables function as full or partial mediators between the intervention and the outcome variables. Evidence for full mediation would be that a previously significant effect of the intervention on the outcome variable is no longer present when the composite mediating variables have been added to the model. On the other hand, if a significant association still remains between the intervention and the outcome variables after adding the composite mediating variables, this would indicate partial mediation, at best.

### Software and type of factor analysis and goodness-of-fit indices used

Structural equation modelling was conducted with the software package AMOS version 23 (IBM Corporation, Armonk, NY, USA) using maximum likelihood estimation. As chi-squared is known to be inflated with sample size,[131] it was not used as an indicator of goodness of fit (and misfit means that there is major deviation between what the model specifies and what exists in the actual data). Instead, a set of indicators was used to provide a fair assessment of goodness of fit. These indicators are known to be sensitive to different aspects of model misfits. The following criteria were used:

- root-mean-square error of approximation (RMSEA) of $< 0.06$[132]
- comparative fit index (CFI) of $> 0.95$[132]
- standardised root-mean-square residual (SRMR) of $< 0.08$[132]
- relative/normed chi-squared [$\chi^2$/degrees of freedom (df)] of $< 5$ (although this criterion is commonly used and well known, it is equally sensitive to sample size[133]).

### Results

Of the 1324 children, 105 had missing data that did not meet the rules for replacing missing values. A conservative approach was taken by discarding these participants, which reduced the sample size from 1324 to 1219.

A descriptive summary of data for the five composite variables at baseline and at 12 months is provided in *Tables 44* and *45*. None of the variables demonstrated elevated levels of kurtosis and skewness.

**TABLE 44** Descriptive summary of the five composite variables at baseline

| Variable (number of included items) | Score range | Mean | SD | Kurtosis | Skewness | Inter-item correlation range (mean) | Cronbach's alpha |
|---|---|---|---|---|---|---|---|
| Knowledge (5) | 1–19 | 8.47 | 2.69 | –0.11 | 0.09 | NA | NA |
| Confidence and motivation (9) | 9–36 | 23.95 | 6.38 | –0.56 | –0.28 | 0.18–0.59 (0.37) | 0.84 |
| Peer norms (9) | 5–32 | 19.63 | 4.49 | –0.22 | –0.12 | 0.07–0.48 (0.22) | 0.69 |
| Family approval/behaviours and child attitudes (8) | 9–36 | 26.72 | 5.92 | –0.19 | –0.56 | 0.17–0.55 (0.33) | 0.82 |
| Behaviours and strategies (18) | 3–72 | 41.19 | 12.79 | –0.54 | 0.22 | 0.18–0.61 (0.40) | 0.92 |

NA, not applicable.

**TABLE 45** Descriptive summary of the five composite variables at 12 months

| Variable (number of included items) | Score range | Mean | SD | Kurtosis | Skewness | Inter-item correlation range (mean) | Cronbach's alpha |
|---|---|---|---|---|---|---|---|
| Knowledge (5) | 2–20 | 10.49 | 3.18 | –0.26 | –0.05 | NA | NA |
| Confidence and motivation (9) | 9–36 | 24.21 | 5.79 | –0.49 | –0.29 | 0.21–0.65 (0.35) | 0.83 |
| Peer norms (9) | 7–30 | 19.55 | 4.25 | –0.32 | –0.23 | 0.00–0.55 (0.21) | 0.70 |
| Family approval/behaviours and child attitudes (8) | 9–36 | 27.22 | 5.80 | 0.04 | –0.70 | 0.17–0.66 (0.35) | 0.83 |
| Behaviours and strategies (18) | 18–72 | 42.33 | 11.78 | –0.55 | 0.16 | 0.18–0.68 (0.39) | 0.92 |

NA, not applicable.

Cronbach's alpha was good for all of the composite variables except peer norms, which approached the consensual cut-off value of 0.70 for acceptable internal consistency.

*Tables 46* and *47* show the correlation coefficients for the five mediating variables at baseline and at 12 months, respectively. Knowledge had very low correlations, and in one case no correlation, with the other variables. The correlations among the confidence and motivation, peer norms, family approval/ behaviours and child attitudes and behaviours and strategies variables were small to medium.

Preliminary analyses indicated that the conditions for mediation analysis[126] had been met: (1) the intervention had a significant effect on each of the five 12-month composite variables, namely knowledge [$t(1217) = 15.27$;

**TABLE 46** Pearson correlation coefficients of the composite variables at baseline

| Variable | Knowledge | Confidence and motivation | Peer norms | Family approval/behaviours and child attitudes |
|---|---|---|---|---|
| Confidence and motivation | 0.06* | – | | |
| Peer norms | 0.09** | 0.37** | – | |
| Family approval/behaviours and child attitudes | 0.22** | 0.49** | 0.41** | – |
| Behaviours and strategies | –0.02 | 0.61** | 0.34** | 0.40** |

*$p < 0.05$, **$p < 0.01$.

**TABLE 47** Pearson correlation coefficients of the composite variables at 12 months

| Variable | Knowledge | Confidence and motivation | Peer norms | Family approval/behaviours and child attitudes |
|---|---|---|---|---|
| Confidence and motivation | 0.16** | – | | |
| Peer norms | 0.15** | 0.36** | – | |
| Family approval/behaviours and child attitudes | 0.22** | 0.53** | 0.38** | – |
| Behaviours and strategies | 0.10** | 0.69** | 0.33** | 0.51** |

*$p < 0.05$, **$p < 0.01$.

$p < 0.01$], confidence and motivation [$t(1217) = 4.00$; $p < 0.01$], peer norms [$t(1217) = 4.01$; $p < 0.01$], family approval/behaviours and child attitudes [$t(1217) = 3.40$; $p < 0.01$] and behaviours and strategies [$t(1217) = 5.66$; $p < 0.01$]; and (2) the composite variable scores at 12 months were significantly ($p < 0.01$) correlated with the two outcome variables: weekday energy-dense snacks and negative food markers (*Table 48*). As the outcome variable negative food markers at 18 months as well as the negative food markers score at baseline revealed slightly elevated kurtosis (1.86 and 2.18, respectively), the scores were logarithmically transformed before the SEM was conducted.

Our longitudinal path analysis contained data from three time points: (1) at baseline, including the demographic variables gender, baseline BMI SDS, school size and school-level deprivation, with direct paths to the outcome variable as well as the baseline scores of the outcome and demographic variables; (2) the 12-month mediating variables; and (3) the 18-month outcome data. The baseline score of the outcome variable (energy-dense snack or negative food marker) as well as the child's allocated group (independent variable) also had direct paths to the outcome variable. The model shows that baseline composite variables predicted scores on these same variables at 12 months, which in turn predicted the outcome variable at 18 months. The intervention versus control variable also predicted the scores on the five composite variables at 12 months. Several of the composite variables were correlated at baseline (guided by modification indices of > 10). The error variances of the mediating variables at 12 months were correlated in an identical pattern to those pairs of mediating variables at baseline.

*Table 49* shows the goodness-of-fit indices for the two longitudinal path analysis models (with energy-dense snacks and negative food markers as outcome variables). Overall, the model fits were acceptable for both outcome variables, and although the CFI results were slightly below the 0.95 criterion for an excellent fit, the other indices met the criteria.

### Path analysis for weekday energy-dense snacks (18 months)
*Figure 10* presents a diagram of the path analysis for energy-dense snacks. Not shown are the correlations of the variables at baseline, error covariances and non-significant regression weights (dashed lines).

**TABLE 48** Summary of Pearson correlation coefficient between the outcome variables energy-dense snacks and negative food makers with the mediating variables at 12 months

| Variable | Energy-dense snacks | Negative food markers |
|---|---|---|
| Knowledge | −0.16** | −0.16** |
| Confidence and motivation | −0.19** | −0.20** |
| Peer norms | −0.12** | −0.10** |
| Family approval/behaviours and child attitudes | −0.28** | −0.27** |
| Behaviours and strategies | −0.13** | −0.12** |
| **$p < 0.01$. | | |

**TABLE 49** Goodness-of-fit indices for the longitudinal path analysis models with energy-dense snacks and negative food markers at 18 months as outcome variables

| Goodness-of-fit indices | Energy-dense snacks | Negative food markers |
|---|---|---|
| Satorra–Bentler scaled $\chi^2$/df | 4.303 | 4.336 |
| RMSEA | 0.052 | 0.052 |
| CFI | 0.943 | 0.940 |
| SRMR | 0.057 | 0.061 |

**FIGURE 10** Partial mediation model with significant standardised regression weights of pathways for weekday energy-dense snack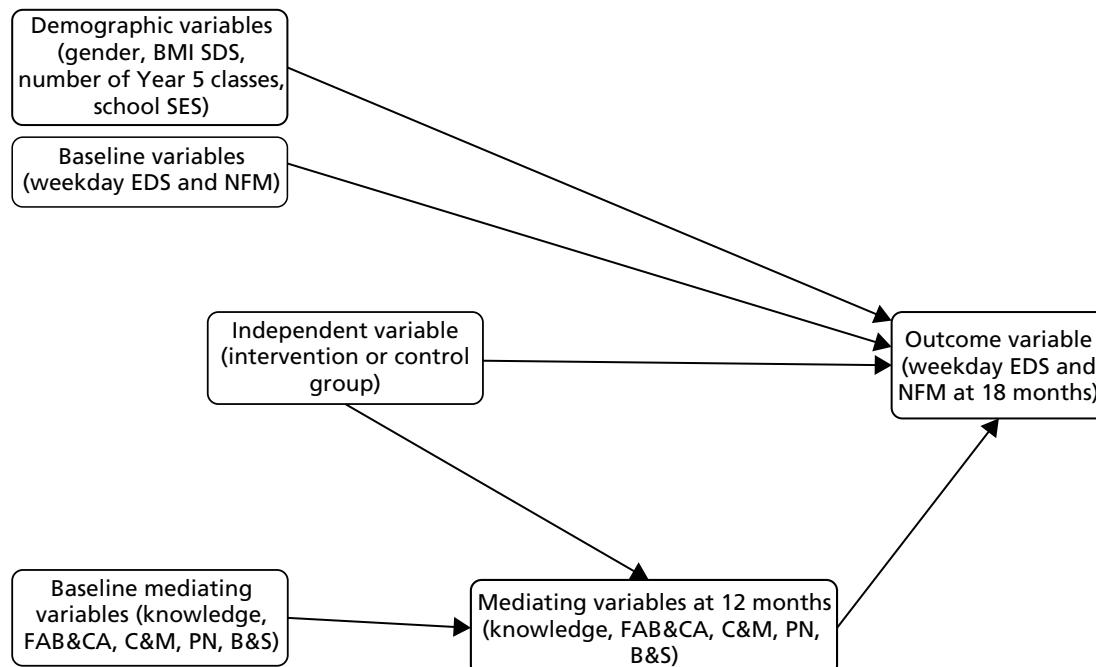s at 18 months. B&S, behaviours and strategies; C&M, confidence and motivation; EDS, energy-dense snacks; FAB&CA, family approval/behaviours and child attitudes; NFM, negative food markers; PN, peer norms; SES, socioeconomic status. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

The following composite variables were significantly ($p < 0.05$) correlated at baseline: knowledge with family approval/behaviours and child attitudes and confidence and motivation; family approval/behaviours and child attitudes with confidence and motivation, peer norms and behaviours and strategies; confidence and motivation with peer norms and behaviours and strategies; and peer norms with behaviours and strategies. All paths shown in the diagram were significant ($p < 0.05$), except for the associations between gender and energy-dense snacks (18 months), between school-level deprivation and energy-dense snacks (18 months), between peer norms and energy-dense snacks (18 months) and between behaviours and strategies and energy-dense snacks (18 months). The direct effect of the intervention on energy-dense snacks (18 months) was significant ($p = 0.041$), indicating partial mediation by knowledge, family approval/behaviours and child attitudes and confidence and motivation. Of these three mediating variables, the effect of family approval/behaviours and child attitudes was the strongest, with a standardised regression weight of –0.18, followed by confidence and motivation ($\beta = -0.09$) and knowledge ($\beta = -0.07$).

## Path analysis for weekday negative food markers (18 months)

*Figure 11* shows the full mediation model with significant standardised regression weights of pathways for weekday negative food markers at 18 months. Not shown are the correlations of the variables at baseline, error covariances and non-significant regression weights (dashed lines). The results for negative food markers were generally similar, with the same variables correlated at baseline and at 12 months. Similar to the results for energy-dense snacks, there were no significant associations between gender and negative food markers (18 months), between school-level deprivation and negative food markers (18 months) and between peer norms and negative food markers (18 months). However, the association between behaviours and strategies and negative food markers (18 months) was significant ($p < 0.05$), and there was no significant association between school size and negative food markers (18 months). Additionally, the direct effect of the intervention on negative food markers (18 months) was not significant ($p = 0.059$) indicating full mediation by knowledge, family approval/behaviours and child attitudes, confidence and motivation and behaviours and strategies.



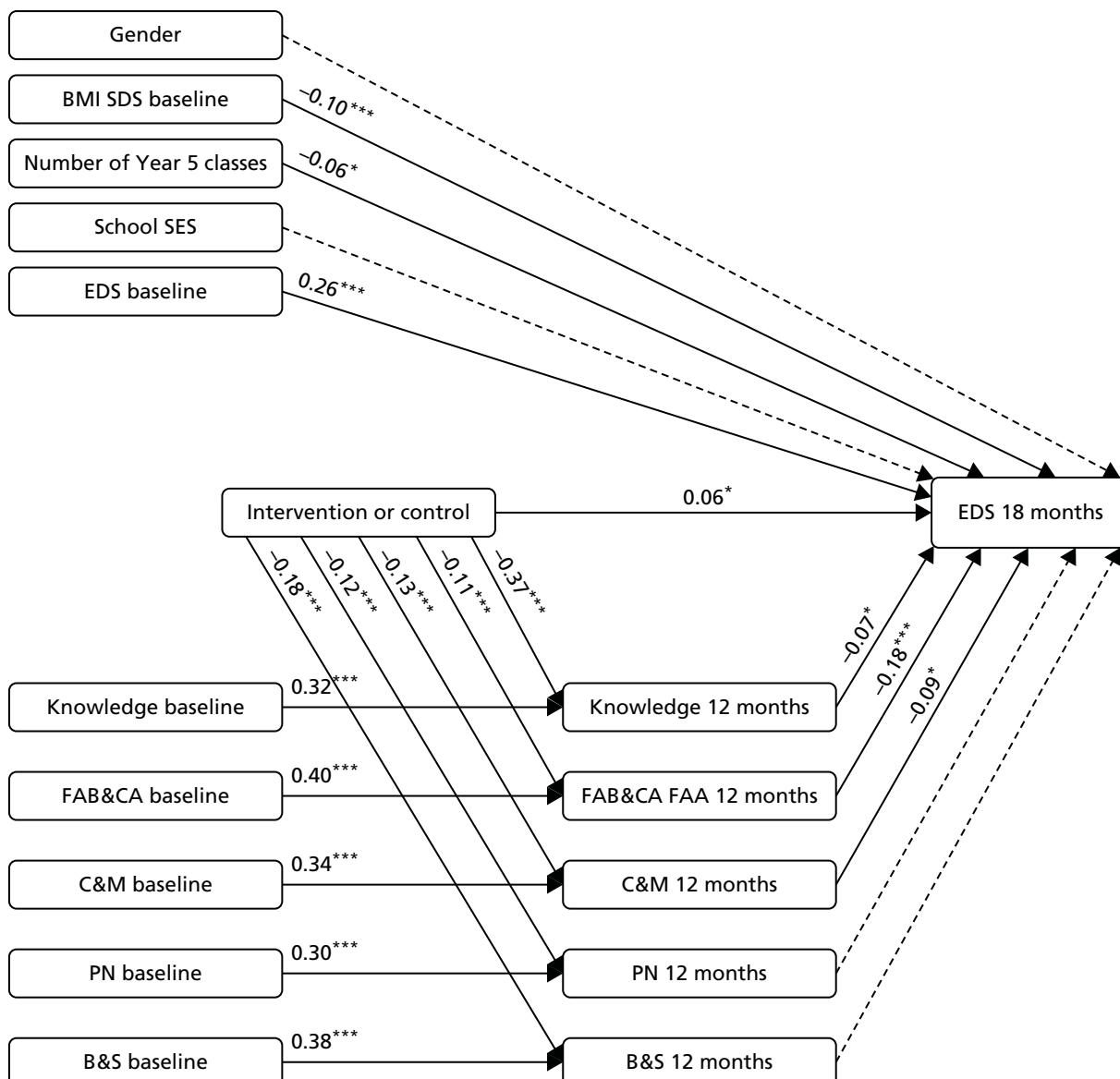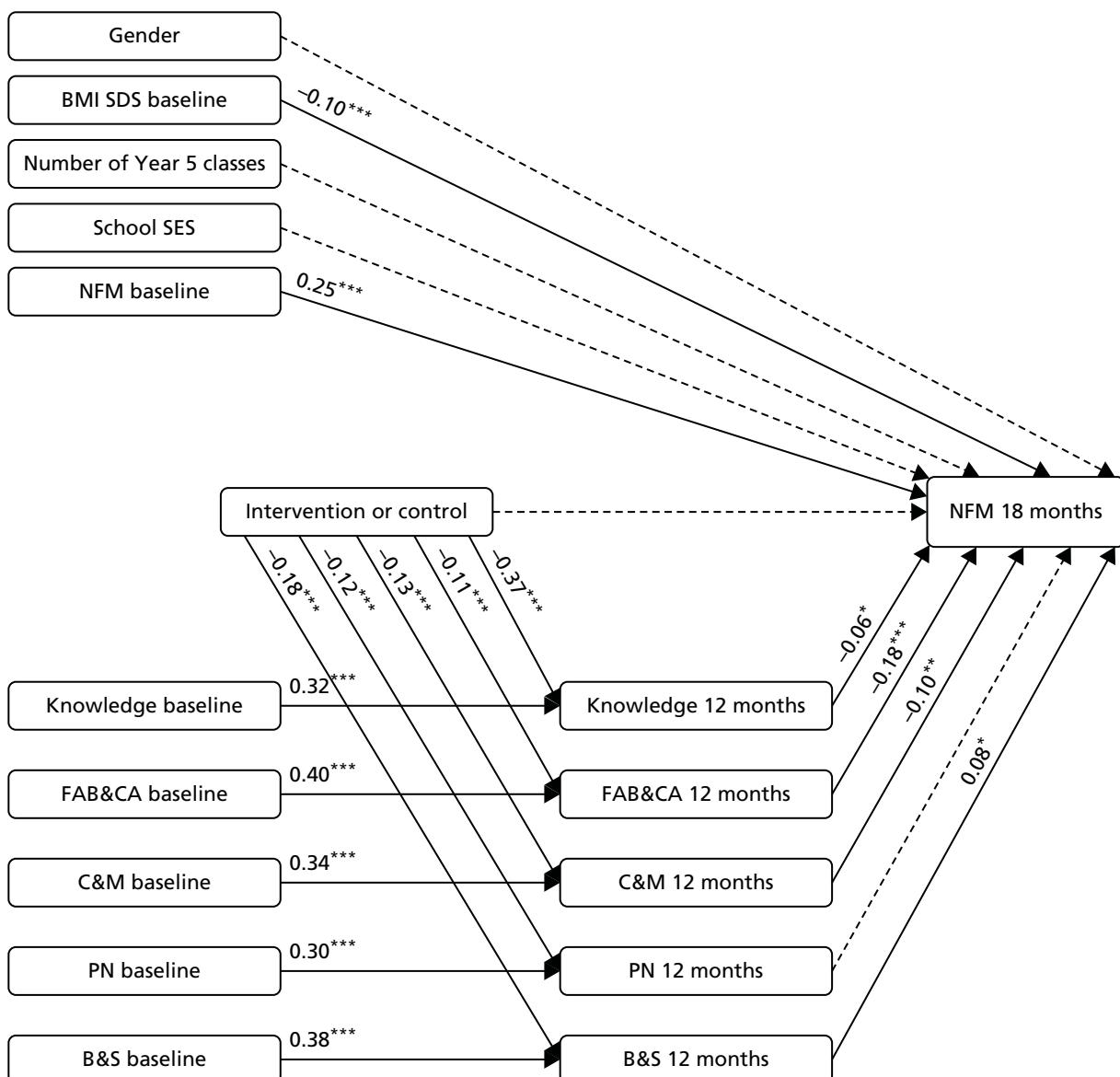**FIGURE 11** Full mediation model with significant standardised regression weights of pathways for weekday negative food markers at 18 months. B&S, behaviours and strategies; C&M, confidence and motivation; EDS, energy-dense snacks; FAB&CA, Family approval/behaviours and child attitudes; NFM, negative food markers; PN, peer norms; SES, socioeconomic status.*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Of these four variables, the effect of family approval/behaviours and child attitudes was the strongest, with a standardised regression weight of –0.18, followed by confidence and motivation ($\beta = -0.10$), behaviours and strategies ($\beta = 0.08$) and knowledge ($\beta = -0.06$).

Overall, the effects of the composite mediating variables on the outcome variables were fairly small. The largest association was between family approval/behaviours and child attitudes at 12 months and the outcome variables. For both paths (between family approval/behaviours and child attitudes and energy-dense snacks and between family approval/behaviours and child attitudes and negative food markers), $\beta$ was –0.18, which means that an increase of 1 SD of the family approval/behaviours and child attitudes variable was associated with a reduction of 0.18 in energy-dense snacks and negative food markers. Additionally, the results of the behaviours and strategies variable need to be interpreted with caution. As outlined in *Appendix 18*, this variable was the least psychometrically robust of MLQ variables. The path between behaviours and strategies at 12 months and energy-dense snacks was not significant, but the path between behaviours and strategies at 12 months and negative food markers resulted in a significant change in the opposite direction to those of the other mediating variables. This counterintuitive result is most likely a result of collinearity; as shown in *Table 48*, behaviours and strategies was negatively correlated with negative food markers. However, in the context of the other MLQ variables, this association became positive ($\beta$ was 0.08; see *Figure 11*). Such seemingly paradoxical cases have been described in the literature[134] and are most likely due to collinearity with other predictor variables or the operation of suppressor variables. *Table 47* shows that behaviours and strategies and confidence and motivation were correlated to an extent ($r = 0.69$) that could question the statistical independence of the variables and further multiple-linear regression analyses confirmed that the coefficient of behaviours and strategies does change direction when confidence and motivation is added as a second predictor. Further exploration of these composite variables could clarify which items are responsible for these effects.

### *Summary*
We developed and evaluated a self-report tool, the MLQ, which aimed to capture changes in knowledge, cognitions and behaviours that could explain changes in weight gain. The MLQ contains items that are relevant, acceptable and feasible for 9- to 10-year-olds to complete in a timely manner and it has undergone psychometric testing, although future additional evaluation of the MLQ could include further tests for reliability and construct validity. Five composite variables emerged from the evaluation of the MLQ and these were used as the mediating variables in the two longitudinal path analyses. The first analysis was for the number of weekday unhealthy foods consumed per day (weekday negative food markers at 18 months) and showed statistical evidence for full mediation as the previous significant effect of the intervention on this outcome variable was no longer present ($p = 0.059$) after all five mediators had been entered into the model. This result suggests that the composite variables arising from the MLQ could provide a possible explanation of how the intervention generated the observed differences in dietary behaviour. However, further exploration could clarify this explanation. The analysis for the number of weekday energy dense snacks consumed per day at 18 months revealed partial mediation ($p = 0.041$); one possible explanation for this is that other mediating variables exist that have not been measured in this study. Overall, it is noted that the full and partial mediation effects were both close to the cut-off point for statistical significance ($p = 0.05$).

The findings arising from our use of composite variables require cautious interpretation; we are able only to conclude that children have changed in a combination of ways for two of the secondary outcomes in the trial. Future work could include more theoretically based modelling work (which would be possible given the theoretical underpinning of many of the items in the MLQ) as well as cluster analyses[135] to investigate which specific mediator and which moderator variables might predict healthy outcomes in the whole HeLP cohort.

## Conclusions from the process evaluation

Data from the process evaluation show that HeLP was delivered as designed in all 16 intervention schools, with high uptake and engagement from schools, children and their families across the socioeconomic spectrum. Qualitative data suggest that HeLP worked in accordance with the underpinning IMB model and the logic model by changing children's self-reported levels of knowledge, cognitions and behaviours as intended; however, these changes were insufficient to affect physical activity behaviours at 18 months or BMI SDS at either 18 or 24 months post baseline. The mediation analyses show that the intervention effects on the consumption of weekday energy-dense snacks were mediated by knowledge and two composite variables, namely family approval/behaviours and child attitudes and confidence and motivation, whereas the intervention effect on weekday consumption of unhealthy foods (negative food markers) was mediated by the same variables as well as the composite variable behaviours and strategies.

# Chapter 6 Discussion and conclusions

## Summary of findings

We evaluated the effectiveness and cost-effectiveness of HeLP, a novel school-based obesity prevention programme, in a cluster RCT involving 32 schools and > 1300 children. We found no evidence of a difference in BMI SDS at 24 months or that participating in HeLP reduced the likelihood that children would be overweight or obese compared to children not receiving the intervention. Similarly, no differences between the intervention and control groups were observed in either anthropometric measures or physical activity objectively assessed using accelerometers at 18 months post baseline. Self-reported weekly average consumption of different types of energy-dense snacks was lower in those attending intervention schools (0.4 mean difference in score), as was the number of different types of negative food markers (0.5 mean difference), small differences that were statistically significant. These differences were largely accounted for by reported differences in weekday consumption.

The cost of implementing HeLP was estimated at approximately £210 per child. Assumptions are reported regarding the proportions of children needing to move weight category for cost-effectiveness to be achieved using NICE cost-per-QALY methodology.

## Comparison with other studies

Wang *et al.*[4] is the most recent systematic review and meta-analysis of childhood obesity prevention programmes targeting both diet and physical activity. The review identified 139 intervention studies that had weight-related outcomes, of which 115 were located in the primary school. The 37 studies that were purely school-based and did not have a family component showed a low strength of evidence for reducing BMI, BMI SDS, prevalence of obesity and overweight, percentage body fat, waist circumference and skinfold thickness. However, studies that also included a family component provided moderate evidence of effectiveness, with half reporting statistically significant beneficial intervention effects. Other systematic reviews and meta-analyses also suggest that school-based obesity prevention interventions can have a modest effect on BMI SDS and it is unclear whether such effect sizes (typically < 0.1) are clinically meaningful at either a population or an individual level.[7,136]

We are not aware of any recent, well-conducted, school-based obesity prevention RCTs, using objective outcome measures, for this age group, that have shown a clinically relevant effect on adiposity measures at 2-year follow-up. A very recent school-based trial (Active for Life-year 5)[16] involving 60 schools and > 2000 children (aged 9–10 years), which aimed to increase physical activity, reduce sedentary behaviour and increase fruit and vegetable consumption at 2-year follow-up, found no effect of the intervention on any of these primary outcomes or on weight status.

The exploratory trial of HeLP demonstrated the feasibility and acceptability of the programme, trial processes and provided evidence of 'proof of concept' for HeLP. Furthermore, the exploratory trial showed changes in diet and physical activity behaviours and weight status; however, these were not replicated in the main trial.[28] No changes were made to the programme or trial design between the exploratory trial and the definitive RCT, with the manuals and standard operating procedures, developed in the exploratory trial, used for delivery and assessment in the main trial. In addition, HeLP was delivered as designed in all intervention schools with very high levels of engagement, as was also seen in the exploratory trial. The findings from the process evaluation allow us to be confident that the difference in results between the exploratory and the definitive trial are not due to scale-up issues of delivery. The only difference in trial design was that we had to run the main trial in two cohorts; however, there was no evidence of a

differential effect of the intervention between the two cohorts on the primary outcome, indicating that this logistical requirement did not affect the overall findings, and our follow-up rates at 18 and 24 months were similar across both trials.

## Understanding the lack of effectiveness

Conducting health promotion interventions within schools has the obvious potential advantage of being able to reach virtually all children. The behaviours that underlie the development of obesity and overweight in children and adolescents result from a complex interaction of individual, family and social factors. This is particularly relevant for children of primary school age, as their ability to influence their diet and activity is directly limited by decisions made by their parents/carers, as well as being affected by wider social influences. We therefore aimed to develop an intervention that would influence not only the children themselves, but also their parents and the school environment, as we felt that this would have a higher likelihood of being effective.

The design of the intervention followed the MRC's guidance for the development and evaluation of complex interventions[19] and was based on a review of research evidence and on the use of intervention mapping[18] to draw on evidence regarding potentially effective methods of achieving behaviour change. Our review of existing evidence suggested that multifaceted interventions were more likely to be effective when they addressed both diet and exercise, were of significant duration and involved the family, although the strength of these conclusions was limited owing to the paucity of existing high-quality studies. We were also aware that a common reason for failure in health promotion programmes is a failure to persuade the target group to participate and to stay involved, so strategies to achieve engagement by children, parents and schools were fundamental in the design of both the intervention and the trial. In addition, for public health interventions to have an impact, they need to be deliverable without disrupting normal activities and at an affordable cost. In particular, in the context of crowded school curricula, interventions need to reduce rather than increase teachers' workloads and, ideally, help to achieve educational targets if schools are to embrace them. We therefore worked closely with children from the target age group, parents, teachers and education advisors at all stages of the design.

We assessed the extent to which children, parents and teachers actually engaged with the programme using prespecified criteria for engagement, as well as conducting focus groups with children and interviews with parents and teachers. The results suggested very high levels of engagement across all socioeconomic groups and considerable enthusiasm for the programme. We succeeded in achieving extremely high levels of participation: only 34 out of 1371 children opted out of the trial and only 80 out of 1324 who started the study were lost to follow-up, and hence > 94% of children provided anthropometric data at 2 years and we have accelerometry data at 18 months post baseline on 84% of children (3 weekdays and 1 weekend day).

The lack of any effect of the intervention on our primary outcome measure is particularly disappointing given the high levels of engagement and that the programme was developed with substantial stakeholder involvement and reflected the current best evidence regarding techniques to change behaviour. There is a number of potential explanations: we recognise the importance of wider family and social factors in driving health behaviours, which may limit the potential effects of interventions that are delivered primarily at the level of the individual. We are also aware that although the increase in overweight among children is perceived by policy-makers as constituting a major threat to health, it is less clear whether or not parents share this view. It has been repeatedly reported that a large proportion of parents of overweight children perceive their children as being of normal weight.[137] Stakeholder consultation in the development of HeLP clearly showed that focusing on weight as a health issue would not be acceptable to head teachers, teachers or parents, and thus the programme was careful to avoid discussing obesity or weight in any context. Furthermore, we believe that our very high levels of engagement support the programme's focus on healthy lifestyle behaviours for all. However, it may be the case that the HeLP messages regarding diet

and physical activity behaviours were not really seen by parents as referring to a problem that required significant and sustained behaviour change.

We chose to target children aged 9–10 years for the intervention on the basis of our early pilot work. Broadly speaking, our initial development work showed that younger children were keen to engage but did not appear to successfully absorb the messages such that they engaged their parents with them, while older children were less easy to engage. However, it is the case that, at this age, the ability of children to actually influence their own diet and activity may be limited, as, inevitably, most decisions will be made by their parents. Parental involvement in obesity prevention programmes has been frequently cited as one of the key characteristics associated with behaviour change, and throughout the programme there were activities specifically aimed at engaging parents across the socioeconomic spectrum.

Our data showed that > 50% of parents attended at least one parental engagement event, and three-quarters either attended an event or signed support for their child on the goal-setting sheet. Although it is possible to engage children (and assess their levels of engagement) in school-based programmes, the lack of direct contact with parents means that parental engagement is more difficult to achieve and assess. Many parents reported (in questionnaire responses and interview) that they had made changes at a family level and gave examples of how they were supporting their child, but this was only a subset of all of the parents involved. Given the children's comments about the need for parental support, it is likely that the programme did not enable parents or families to make the types of sustained behaviour changes necessary to affect the weight status of their child. We consider it unlikely that school-based obesity prevention programmes will ever be of sufficient intensity to affect the family environment, and more direct approaches would be required.

We were mindful from the outset that the programme should complement educational activities (by meeting national curriculum objectives) and not be a burden to teachers. It is possible that, although the external delivery of the majority of programme activities (by sport and dance groups, actors and the HeLP co-ordinators) was welcomed by teachers and children, this delivery mode meant that the programme was probably not embedded across the school and hence could not systematically affect the school environment.

Our development and early piloting work suggested that a more intensive intervention with greater parental involvement in the programme would be less acceptable or unfeasible for schools, suggesting that those planning future obesity prevention interventions may need to consider new approaches and settings if they are seeking to affect family behaviours and the home environment.

## Trial strengths and limitations

There are well-recognised potential sources of bias associated with cluster RCTs,[33] including recruitment bias, performance bias and detection bias. We tried to reduce the likelihood of these biases by recruiting schools and children and collecting baseline measures before the known allocation of schools to intervention or control (to reduce differential take-up); by capturing evidence of changes in school policies around food or physical activity through school completion of a checklist at the beginning of the trial and at 18 months (see *Appendix 16*); and by using assessors blinded to allocation to measure the anthropometric outcomes. The primary outcome for the trial was based on measures taken when the children had transferred to secondary school, and hence secondary schools included children from both the intervention and the control groups. No child revealed their group allocation to the blinded assessor at 24-month follow-up. However, it should be noted that although the MLQ and FIQ were completed before the group allocation at baseline and at 12 and 18 months, respectively, the children were aware of their group allocation, as shown in *Figure 1*.

Sample size calculations, which allowed for the anticipated level of clustering as estimated from the exploratory trial and NCMP data,[28,138] suggested that we needed measures from approximately 760 children at 24 months to detect a 0.25 difference in BMI SDS. We estimated that we needed 28 schools, each with

approximately 35–40 children, to complete the study and so, allowing for a 20% loss to follow-up at 24 months, we sought to recruit 32 schools to allow for both school-level and child-level dropouts. No schools dropped out of the study. Few eligible children (34/1371) opted out of the study and we achieved follow-up rates of between 84% and 96% for all outcome measures, and hence there is a very low risk of attrition bias.

Reviews and recent studies of school-based obesity prevention and management trials in children have highlighted low participation, differential dropout and loss to follow-up rates.[4,5,9,139] For example, completeness of anthropometric data in school-based obesity prevention programmes has ranged between 70% and 80% for follow-up of ≥ 24 months[13,140,141] and the percentage of children providing a representative pattern of their physical activity levels across the entire week (at least 3 weekdays and 1 weekend day of 10 hours' wear time) tends to be much lower (between 40% and 60%).[38] In the HeLP trial, 84% of children met the minimum wear time criteria of 3 weekdays and 1 weekend day, and 79% provided data on all 7 days at 18 months. The HeLP trial, therefore, has one of the most complete follow-up rates and physical activity compliance of recent obesity prevention trials for children in this age group. We attribute this to the relational approach that we took in the trial delivery, as well as to the stakeholder involvement we had in developing the intervention and the trial design.[27]

We weighted our recruitment of schools to achieve a similar proportion of pupils receiving free school meals as the national average in 2012 (19%), which is higher than the average for Devon (12.7%), and the schools in the HeLP trial were larger than the average Devon primary school. In other respects, the schools that participated in the trial are representative of Devon, and the anthropometric data from the children are broadly similar to the Devon NCMP Year 6 data.[142] This gives us confidence that the results are mostly applicable to a wider population. The one exception is that, although the included schools reflected the proportion of children with English as an additional language typical for Devon (2.6%), this is substantially lower than the national figure of 16.8% of pupils with English as an additional language in England.

A further strength of this trial was its robust process evaluation involving qualitative and quantitative data and its prespecified logic model and mediating variables questionnaire to understand whether or not HeLP was working in the way it was intended, with the qualitative data being analysed prior to the outcome result being known. We also attempted to assess engagement (school, child and parent); whereas school and child engagement was assessed using different methods (observation of intervention sessions and response to the programme), parent engagement was assessed as attending one of six possible parent events and/or providing a signature on the child's goal-setting sheet. We acknowledge that this is a weak assessment of engagement and does not capture engagement with the messages of the programme. The low response rate (25%) from parents to the questionnaire about HeLP, which formed part of the process evaluation, limits the strength of these data and also suggests that the sample interviewed (who were recruited from responders to the questionnaire) is unlikely to be representative of all parents. This suggests that the qualitative data from parents cannot be considered as representative of all parents receiving the programme. However, the children recruited for the process evaluation were representative of the overall sample, and all Year 5 teachers were interviewed.

## Research recommendations

The disappointing results of this study and other recent well-designed, UK-based trials suggest to us that it is unlikely that affordable, school-based interventions that are inherently time-limited can achieve clinically meaningful change in weight status/anthropometric outcomes in a single targeted age group. This view is reinforced by other research using data from the NCMP that suggested that there was no consistent school effect on childhood obesity.[143] Future studies need to develop interventions that aim to specifically affect the family environment, perhaps employing a more widely co-ordinated whole-systems approach, although it is unlikely that such an approach would be feasible to deliver within a school setting.

We recognise that one reason for the lack of effectiveness of this intervention may be that, despite engaging and affecting knowledge and motivation, children of this age have limited ability to affect their own diet and levels of physical activity. This suggests that researchers should consider whether or not other age groups offer greater possibilities for affecting behaviour. It is possible that older children, who have a greater degree of autonomy, if engaged, may be more able to make changes in their behaviour. It has also been suggested that pregnant women and parents of very young children may be more amenable to adopting health messages. Examples include successful campaigns to promote smoking cessation, the avoidance of soft cheeses during pregnancy and the 'Back to Sleep' campaign. Programmes addressing this group (e.g. the Henry Programme: www.henry.org.uk/homepage/) have claimed impact, but rigorous evaluations are lacking and are urgently needed.

## Conclusions

This cluster randomised trial provides rigorous evidence about the effects of HeLP, a novel school-based obesity prevention programme aimed at 9- to 10-year-old children, in a sample of children broadly representative of the UK population. The programme achieved high levels of engagement with children, schools and parents, but did not affect their BMI or their likelihood of being overweight or obese 2 years later compared with children attending schools in which the intervention was not delivered.

# Acknowledgements

We thank all of the pupils, teachers, head teachers, teaching assistants and parents/carers for participating in the study. We thank the HeLP co-ordinators, the drama teams, the director of Headbangers Theatre Company, Just4Funk, Attik Dance, Exeter Chiefs and Plymouth Raiders for their enthusiasm and support in delivering the programme. We thank all of the HeLP study staff, trained assessors and trainers who delivered assessment training, as well as data management and administrative staff who provided fabulous support throughout the trial. We are indebted to our Public Advisory Group whose advice and ongoing support were invaluable in ensuring that the trial design and intervention were feasible and acceptable to schools, children and their families. We also thank the chairperson and members of the TSC for their advice and support throughout the trial. We would also like to acknowledge the NIHR Collaboration for Leadership in Applied Health Research and Care for providing non-financial methodological support during the transition from the exploratory trial to the definitive evaluation.

## Trial registration and ethics approval

## Funding

## Contributions of authors

**Katrina Wyatt** (Associate Professor of Health Research, University of Exeter Medical School) led the study, co-designed the HeLP intervention and co-led the process evaluation. She was involved in all stages of the HeLP trial, including conception, design, interpretation of data, drafting and critical revision of the report for important intellectual content and approval of the final version, and co-ordinated responses from other authors.

**Jenny Lloyd** (Senior Research Fellow Child Health, University of Exeter Medical School) co-led the project, co-designed the HeLP intervention, led the process evaluation and managed the delivery of the trial. She was involved in all stages of the HeLP trial, including conception, design, interpretation of data, drafting and critical revision of the report for important intellectual content and approval of the final version.

**Siobhan Creanor** [Associate Professor (Reader) in Clinical Trials & Medical Statistics, Plymouth University Peninsula Schools of Medicine and Dentistry] co-led the project, led the statistical analyses and was involved in all stages of the HeLP trial, including conception, design, interpretation of data, drafting and critical revision of the report for important intellectual content and approving of the final version.

**Colin Green** (Professor of Health Economics, University of Exeter Medical School) co-led the project, led the economic evaluation and was involved in all stages of the HeLP trial, including conception, design, interpretation of data, drafting and critical revision of the report for important intellectual content and approval of the final version.

**Chris Krägeloh** (Associate Professor Health Sciences, Auckland University of Technology) co-led the psychometric evaluation of the MLQ and carried out the meditational analyses.

**Richard Siegert** (Professor of Psychology and Rehabilitation, Auckland University of Technology) co-led the psychometric evaluation of the MLQ and carried out the meditational analyses.

**Stuart Logan** (Director of Institute for Health Research, University of Exeter Medical School) co-led the project, co-designed the intervention and was involved in all stages of the HeLP trial, including conception, design, interpretation of data, drafting and critical revision of the report for important intellectual content and approval of the final version.

All authors made critical revisions to the monograph.

## Publications

Wyatt KM, Lloyd JJ, Abraham C, Creanor S, Dean S, Densham E, *et al.* The Healthy Lifestyles Programme (HeLP), a novel school-based intervention to prevent obesity in school children: study protocol for a randomised controlled trial. *BMC Trials* 2013;**14**:95.

Lloyd JJ, Wyatt KM. Qualitative findings from an exploratory trial of the Healthy Lifestyles Programme (HeLP) and their implications for the process evaluation in the definitive trial. *BMC Public Health* 2014;**14**:578.

Lloyd J, Wyatt K. The Healthy Lifestyles Programme (HeLP) – an overview of and recommendations arising from the conceptualisation and development of an innovative approach to promoting healthy lifestyles for children and their families. *IntJ Environ Res Public Health* 2015;**12**:1003–19.

Lloyd J, Wyatt K. Uptake, retention and engagement of children participating in the cluster randomised controlled trial of the Healthy Lifestyles Programmes (HeLP). *Educ Health* 2015;**33**:88–95.

Creanor S, Lloyd J, Hillsdon M, Dean S, Green C, Taylor RS, *et al.* Detailed statistical analysis plan for a cluster randomised controlled trial of the Healthy Lifestyles Programme (HeLP), a novel school-based intervention to prevent obesity in school children. *Trials* 2016;**17**:599.

Lloyd J, Creanor S, Logan S, Green C, Dean SG, Hillsdon M, *et al.* Effectiveness of the Healthy Lifestyles Programme (HeLP) to prevent obesity in UK primary-school children: a cluster randomised controlled trial [published online ahead of print November 28 2017]. *Lancet Child Adolesc Health* 2017.

## Data sharing statement

We are keen for these data to be used widely by the scientific community. Details of the study can be found on the website (http://medicine.exeter.ac.uk/help) and anyone interested in using these data should contact the corresponding author in the first instance. We do not have funds for data extraction and putting data files together for collaborators and so may have to charge for providing this service.

# References

1. NHS Digital. *National Child Measurement Programme England, 2015/16 School Year*. Leeds: NHS Digital; 2016. URL: http://content.digital.nhs.uk/catalogue/PUB22269/nati-chil-meas-prog-eng-2015–2016-rep.pdf (accessed 1 October 2016).

2. National Obesity Observatory. *Data Factsheet: Child Obesity and Socioeconomic Status 2012*. NHS; 2012. URL: http://docplayer.net/21226653-Child-obesity-and-socioeconomic-status.html (accessed 6 September 2017).

3. Simmonds M, Llewellyn A, Owen CG, Woolacott N. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. *Obes Rev* 2016;**17**:95–107. https://doi.org/10.1111/obr.12334

4. Wang Y, Cai L, Wu Y, Wilson RF, Weston C, Fawole O, *et al.* What childhood obesity prevention programmes work? A systematic review and meta-analysis. *Obes Rev* 2015;**16**:547–65. https://doi.org/10.1111/obr.12277

5. Wang Y, Wu Y, Wilson RF, Bleich S, Cheskin L, Weston C, *et al. Childhood Obesity Prevention Programs: Comparative Effectiveness Review and Meta-Analysis*. Rockville, MD: Agency for Healthcare Research and Quality; 2013.

6. Summerbell CD, Waters E, Edmunds LD, Kelly S, Brown T, Campbell KJ. Interventions for preventing obesity in children. *Cochrane Database Syst Rev* 2005;**3**:CD001871. https://doi.org/10.1002/14651858.CD001871.pub2

7. Khambalia AZ, Dickinson S, Hardy LL, Gill T, Baur LA. A synthesis of existing systematic reviews and meta-analyses of school-based behavioural interventions for controlling and preventing obesity. *Obes Rev* 2012;**13**:214–33. https://doi.org/10.1111/j.1467-789X.2011.00947.x

8. Cook-Cottone C, Casey CM, Feeley TH. A meta-analytic review of obesity prevention in the schools 1997–2008. *Psychol Sch* 2009;**46**:695–719. https://doi.org/10.1002/pits.20409

9. Gonzalez-Suarez C, Worley A, Grimmer-Somers K, Dones V. School-based interventions on childhood obesity: a meta-analysis. *Am J Prev Med* 2009;**37**:418–27. https://doi.org/10.1016/j.amepre.2009.07.012

10. Katz DL. School-based interventions for health promotion and weight control: not just waiting on the world to change. *Annu Rev Public Health* 2009;**30**:253–72. https://doi.org/10.1146/annurev.publhealth.031308.100307

11. Brown T, Summerbell C. Systematic review of school-based interventions that focus on changing dietary intake and physical activity levels to prevent childhood obesity: an update to the obesity guidance produced by the National Institute for Health and Clinical Excellence. *Obes Rev* 2009;**10**:110–41. https://doi.org/10.1111/j.1467-789X.2008.00515.x

12. Kropski JA, Keckley PH, Jensen GL. School-based obesity prevention programs: an evidence-based review. *Obesity* 2008;**16**:1009–18. https://doi.org/10.1038/oby.2008.29

13. Healthy Study Group, Foster GD, Linder B, Baranowski T, Cooper DM, Goldberg L, *et al.* A school-based intervention for diabetes risk reduction. *N Engl J Med* 2010;**363**:443–53. https://doi.org/10.1056/NEJMoa1001933

14. Grydeland M, Bjelland M, Anderssen SA, Klepp KI, Bergh IH, Andersen LF, *et al.* Effects of a 20-month cluster randomised controlled school-based intervention trial on BMI of school-aged boys and girls: the HEIA study. *Br J Sports Med* 2014;**48**:768–73. https://doi.org/10.1136/bjsports-2013-092284

15. Jansen W, Borsboom G, Meima A, Zwanenburg EJ, Mackenbach JP, Raat H, Brug J. Effectiveness of a primary school-based intervention to reduce overweight. *Int J Pediatr Obes* 2011;**6**:e70–7. https://doi.org/10.3109/17477166.2011.575151

16. Kipping RR, Howe LD, Jago R, Campbell R, Wells S, Chittleborough CR, *et al.* Effect of intervention aimed at increasing physical activity, reducing sedentary behaviour, and increasing fruit and vegetable consumption in children: active for Life Year 5 (AFLY5) school based cluster randomised controlled trial. *BMJ* 2014;**348**:g3256. https://doi.org/10.1136/bmj.g3256

17. Gortmaker SL, Peterson K, Wiecha J, Sobol AM, Dixit S, Fox MK, *et al.* Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. *Arch Pediatr Adolesc Med* 1999;**153**:409–18. https://doi.org/10.1001/archpedi.153.4.409

18. Bartholomew LK, Markham CM, Ruiter R, Fernández ME, Kok G, Parcel GS. *Planning Health Promotion Programmes: An Intervention Mapping Approach*. 4th edn. San Francisco, CA: Jossey-Bass; 2016.

19. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655. https://doi.org/10.1136/bmj.a1655

20. Baranowski Y, Cerin E, Baranowski J. Steps in the design, development and formative evaluation of obesity prevention-related behavior change trials. *Int J Behav Nutr Phys Act* 2009;**6**. https://doi.org/10.1186/1479-5868-6-6

21. Langford R, Bonell CP, Jones HE, Pouliou T, Murphy SM, Waters E, *et al.* The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement. *Cochrane Database Syst Rev* 2014;**4**:CD008958. https://doi.org/10.1002/14651858.CD008958.pub2

22. Hesketh K, Waters E, Green J, Salmon L, Williams J. Healthy eating, activity and obesity prevention: a qualitative study of parent and child perceptions in Australia. *Health Promot Int* 2005;**20**:19–26. https://doi.org/10.1093/heapro/dah503

23. Sonneville K, La Pelle N, Taveras E, Gillman M, Prosser L. Economic and other barriers to adopting recommendations to prevent childhood obesity: results of a focus group study with parents. *BMC Pediatr* 2009;**9**:81. https://doi.org/10.1186/1471-2431-9-81

24. Joronen K, Rankin SH, Astedt-Kurkip P. School-based drama interventions in health promotion for children and adolescents; systematic review. *J Adv Nurs* 2008;**63**:116–31. https://doi.org/10.1111/j.1365-2648.2008.04634.x

25. Fisher J, Fisher W. The Information-Motivation-Behavioral Skills Model. In DiClemente R, Crosby R, Kegler M, editors. *Emerging Theories in Health Promotion Practice and Research*. San Francisco, CA: Jossey-Bass; 2002. pp. 40–70.

26. Schwarzer R. Self-Efficacy in the Adoption and Maintenance of Health Behaviors: Theoretical Approaches and a New Model. In Schwarzer R, editor. *Self-Efficacy: Thought Control of Action*. London: Hemisphere; 1992. pp. 217–43.

27. Wyatt KM, Lloyd JJ, Creanor S, Logan S. The development, feasibility and acceptability of a school-based obesity prevention programme: results from three phases of piloting. *BMJ Open* 2011;**1**:e000026. https://doi.org/10.1136/bmjopen-2010-000026

28. Lloyd JJ, Wyatt KM, Creanor S. Behavioural and weight status outcomes from an exploratory trial of the Healthy Lifestyles Programme (HeLP): a novel school-based obesity prevention programme. *BMJ Open* 2012;**2**:e000390. https://doi.org/10.1136/bmjopen-2011-000390

29. Fairclough SJ, Noonan R, Rowlands AV, Van Hees V, Knowles Z, Boddy LM. Wear compliance and activity in children wearing wrist- and hip-mounted accelerometers. *Med Sci Sports Exerc* 2016;**48**:245–53. https://doi.org/10.1249/MSS.0000000000000771

30. Phillips LR, Parfitt G, Rowlands AV. Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *J Sci Med Sport* 2013;**16**:124–8. https://doi.org/10.1016/j.jsams.2012.05.013

31. Wyatt KM, Lloyd JJ, Abraham C, Creanor S, Dean S, Densham E, *et al.* The Healthy Lifestyles Programme (HeLP), a novel school-based intervention to prevent obesity in school children: study protocol for a randomised controlled trial. *BMC Trials* 2013;**14**:95. https://doi.org/10.1186/1745-6215-14-95

32. Lloyd J, Creanor S, Logan S, Green C, Dean SG, Hillsdon M, *et al.* Effectiveness of the Healthy Lifestyles Programme (HeLP) to prevent obesity in UK primary-school children: a cluster randomised controlled trial [published online ahead of print November 28 2017]. *Lancet Child Adolesc Health* 2017. https://doi.org/10.1016/s2352-4642(17)30151-7

33. Caille A, Kerry S, Tavernier E, Leyrat C, Eldridge S, Giraudeau B. Timeline cluster: a graphical tool to identify risk of bias in cluster randomised trials. *BMJ* 2016;**354**:i4291. https://doi.org/10.1136/bmj.i4291

34. Lloyd JJ, Logan S, Greaves CJ, Wyatt KM. Evidence, theory and context – using intervention mapping to develop a school-based intervention to prevent obesity in children. *Int J Behav Nutr Phys Act* 2011;**8**:73. https://doi.org/10.1186/1479-5868-8-73

35. Boal A. *Theatre of the Oppressed*. London: Pluto; 1979.

36. Williams AJ, Wyatt KM, Williams CA, Logan S, Henley WE. A repeated cross-sectional study examining the school impact on child weight status. *Prev Med* 2014;**64**:103–7. https://doi.org/10.1016/j.ypmed.2014.04.003

37. ActivInsights Ltd. *Kimboton U*. URL: www.geneactiv.org/ (accessed 10 September 2016).

38. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc* 2005;**37**(Suppl. 11):531–43. https://doi.org/10.1249/01.mss.0000185657.86065.98

39. van Hees VT, Gorzelniak L, Dean León EC, Eder M, Pias M, Taherian S, *et al.* Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. *PLOS ONE* 2013;**8**:e61691. https://doi.org/10.1371/journal.pone.0061691

40. Johnson B, Hackett AF. Eating habits of 11–14-year-old schoolchildren living in less affluent areas of Liverpool, UK. *J Hum Nutr Diet* 1997;**10**:135–44. https://doi.org/10.1046/j.1365-277X.1997.00046.x

41. Department for Communities and Local Government. *The English Indices of Deprivation 2010*. London: Department for Communities and Local Government; 2011. URL: www.communities.gov.uk/documents/statistics/pdf/1871208.pdf (accessed 12 September 2016).

42. Creanor S, Lloyd J, Hillsdon M, Dean S, Green C, Taylor RS, *et al.* Detailed statistical analysis plan for a cluster randomised controlled trial of the Healthy Lifestyles Programme (HeLP), a novel school-based intervention to prevent obesity in school children. *Trials* 2016;**17**:599. https://doi.org/10.1186/s13063-016-1737-y

43. National Institute for Health Research. *National Institute for Health Research Project PHR – 10301001*. URL: www.nets.nihr.ac.uk/projects/phr/10301001 (accessed 1 March 2016).

44. Campbell MK, Elbourne DR, Altman DG, CONSORT group. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004;**328**:702–8. https://doi.org/10.1136/bmj.328.7441.702

45. Committee for Proprietary Medicinal Products. *Points to Consider on Adjustment for Baseline Covariates*. London: The European Agency for the Evaluation of Medicinal Products; 2003.

46. Senn S. Testing for baseline balance in clinical trials. *Stat Med* 1994;**13**:1715–26. https://doi.org/10.1002/sim.4780131703

47. Peterson B, Harrell FE Jr. Partial proportional odds models for ordinal response variables. *Applied Statistics* 1990;**39**:205–17. https://doi.org/10.2307/2347760

48. Catellier DJ, Hannan PJ, Murray DM, Addy CL, Conway TL, Yang S, Rice JC. Imputation of missing data when measuring physical activity by accelerometry. *Med Sci Sports Exerc* 2005;**37**(Suppl. 11):555–62. https://doi.org/10.1249/01.mss.0000185651.59486.4e

49. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011;**342**:d40. https://doi.org/10.1136/bmj.d40

50. Cole TJ. The LMS method for constructing normalized growth standards. *Eur J Clin Nutr* 1990;**44**:45–60.

51. Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. *Arch Dis Child* 1995;**73**:25–9. https://doi.org/10.1136/adc.73.1.25

52. McCarthy HD, Jarrett KV, Crawley HF. The development of waist circumference percentiles in British children aged 5.0-16.9 y. *Eur J Clin Nutr* 2001;**55**:902–7. https://doi.org/10.1038/sj.ejcn.1601240

53. McCarthy HD, Cole TJ, Fry T, Jebb SA, Prentice AM. Body fat reference curves for children. *Int J Obes* 2006;**30**:598–602. https://doi.org/10.1038/sj.ijo.0803232

54. UK Government. *Compare School and College Performance*. URL: www.compare-school-performance.service.gov.uk/download-data (accessed 24 October 2016).

55. Public Health England. *Measuring and Interpreting BMI in Children*. London: Public Health England; 2012. URL: http://webarchive.nationalarchives.gov.uk/20170110171049/https://www.noo.org.uk/NOO_about_obesity/measurement/children (accessed 6 September 2017).

56. Eldridge SM, Ukoumunne OC, Carlin JB. The intra-cluster correlation coefficient in cluster randomized trials: a review of definitions. *Int Stat Rev* 2009;**77**:378–94. https://doi.org/10.1111/j.1751-5823.2009.00092.x

57. Curtis L, Burns A. *Unit Costs of Health and Social Care 2015*. Canterbury: Personal Social Services Research Unit, University of Kent; 2015.

58. *National Association of Schoolmasters Union of Women Teachers: The Teachers' Union*. URL: www.nasuwt.org.uk/ (accessed 3 June 2016).

59. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *PharmacoEconomics* 2006;**24**:355–71. https://doi.org/10.2165/00019053-200624040-00006

60. Caro JJ, Briggs AH, Siebert U, Kuntz KM, Force I-SMGRPT. Modeling good research practices –overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Med Decis Making* 2012;**32**:667–77. https://doi.org/10.1177/0272989X12454577

61. National Institute for Health and Care Excellence. *Managing Overweight and Obesity Among Children and Young People: Lifestyle Weight Management Services*. NICE guidelines PH47. London: NICE; 2013. URL: http://guidance.nice.org.uk/PH47/Guidance/pdf/English (accessed 14 June 2016).

62. Wang LY, Yang Q, Lowry R, Wechsler H. Economic analysis of a school-based obesity prevention program. *Obes Res* 2003;**11**:1313–24. https://doi.org/10.1038/oby.2003.178

63. Brown HS, Pérez A, Li YP, Hoelscher DM, Kelder SH, Rivera R. The cost-effectiveness of a school-based overweight program. *Int J Behav Nutr Phys Act* 2007;**4**:47. https://doi.org/10.1186/1479-5868-4-47

64. Hollingworth W, Hawkins J, Lawlor DA, Brown M, Marsh T, Kipping RR. Economic evaluation of lifestyle interventions to treat overweight or obesity in children. *Int J Obes* 2012;**36**:559–66. https://doi.org/10.1038/ijo.2011.272

65. Pil L, Putman K, Cardon G, De Bourdeaudhuij I, Manios Y, Androutsos O, *et al.* Establishing a method to estimate the cost-effectiveness of a kindergarten-based, family-involved intervention to prevent obesity in early childhood. The ToyBox-study. *Obes Rev* 2014;**15**(Suppl. 3):81–9. https://doi.org/10.1111/obr.12179

66. Rush E, Obolonkin V, McLennan S, Graham D, Harris JD, Mernagh P, Weston AR. Lifetime cost effectiveness of a through-school nutrition and physical programme: Project Energize. *Obes Res Clin Pract* 2014;**8**:e115–22. https://doi.org/10.1016/j.orcp.2013.03.005

67. Tran BX, Ohinmaa A, Kuhle S, Johnson JA, Veugelers PJ. Life course impact of school-based promotion of healthy eating and active living to prevent childhood obesity. *PLOS ONE* 2014;**9**:e102242. https://doi.org/10.1371/journal.pone.0102242

68. Barrett JL, Gortmaker SL, Long MW, Ward ZJ, Resch SC, Moodie ML, *et al.* cost effectiveness of an elementary school active physical education policy. *Am J Prev Med* 2015;**49**:148–59. https://doi.org/10.1016/j.amepre.2015.02.005

69. Gortmaker SL, Long MW, Resch SC, Ward ZJ, Cradock AL, Barrett JL, *et al.* Cost effectiveness of childhood obesity interventions: evidence and methods for CHOICES. *Am J Prev Med* 2015;**49**:102–11. https://doi.org/10.1016/j.amepre.2015.03.032

70. Long MW, Gortmaker SL, Ward ZJ, Resch SC, Moodie ML, Sacks G, *et al.* Cost effectiveness of a sugar-sweetened beverage excise tax in the U.S. *Am J Prev Med* 2015;**49**:112–23. https://doi.org/10.1016/j.amepre.2015.03.004

71. Sonneville KR, Long MW, Ward ZJ, Resch SC, Wang YC, Pomeranz JL, *et al.* BMI and healthcare cost impact of eliminating tax subsidy for advertising unhealthy food to youth. *Am J Prev Med* 2015;**49**:124–34. https://doi.org/10.1016/j.amepre.2015.02.026

72. Wright DR, Kenney EL, Giles CM, Long MW, Ward ZJ, Resch SC, *et al.* Modeling the cost effectiveness of child care policy changes in the U.S. *Am J Prev Med* 2015;**49**:135–47. https://doi.org/10.1016/j.amepre.2015.03.016

73. Whitaker RC, Wright JA, Pepe MS, Seidel KD, Dietz WH. Predicting obesity in young adulthood from childhood and parental obesity. *N Engl J Med* 1997;**337**:869–73. https://doi.org/10.1056/NEJM199709253371301

74. Cox CS, Rothwell ST, Madans JH, Finucane FF, Freid VM, Kleinman JC, *et al.* Plan and operation of the NHANES I Epidemiologic Followup Study, 1987. *Vital Health Stat 1* 1992;**27**:1–190.

75. Venn AJ, Thomson RJ, Schmidt MD, Cleland VJ, Curry BA, Gennat HC, Dwyer T. Overweight and obesity from childhood to adulthood: a follow-up of participants in the 1985 Australian Schools Health and Fitness Survey. *Med J Aust* 2007;**186**:458–60.

76. Power C, Hertzman C, Matthews S, Manor O. Social differences in health: life-cycle effects between ages 23 and 33 in the 1958 British birth cohort. *Am J Public Health* 1997;**87**:1499–503. https://doi.org/10.2105/AJPH.87.9.1499

77. National Institute for Health and Care Excellence. *Guide to the Method of Technology Appraisal 2013*. London: NICE; 2013. URL: www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781 (accessed 1 November 2016).

78. Singh AS, Mulder C, Twisk JW, van Mechelen W, Chinapaw MJ. Tracking of childhood overweight into adulthood: a systematic review of the literature. *Obes Rev* 2008;**9**:474–88. https://doi.org/10.1111/j.1467-789X.2008.00475.x

79. Brisbois TD, Farmer AP, McCargar LJ. Early markers of adult obesity: a review. *Obes Rev* 2012;**13**:347–67. https://doi.org/10.1111/j.1467-789X.2011.00965.x

80. Wright CM, Parker L, Lamont D, Craft AW. Implications of childhood obesity for adult health: findings from thousand families cohort study. *BMJ* 2001;**323**:1280–4. https://doi.org/10.1136/bmj.323.7324.1280

81. Hawk LJ, Brook CG. Influence of body fatness in childhood on fatness in adult life. *Br Med J* 1979;**1**:151–2. https://doi.org/10.1136/bmj.1.6157.151

82. Cheung YB, Machin D, Karlberg J, Khoo KS. A longitudinal study of pediatric body mass index values predicted health in middle age. *J Clin Epidemiol* 2004;**57**:1316–22. https://doi.org/10.1016/j.jclinepi.2004.04.010

83. McCarthy A, Hughes R, Tilling K, Davies D, Smith GD, Ben-Shlomo Y. Birth weight; postnatal, infant, and childhood growth; and obesity in young adulthood: evidence from the Barry Caerphilly Growth Study. *Am J Clin Nutr* 2007;**86**:907–13.

84. Reilly JJ, Bonataki M, Leary SD, Wells JC, Davey-Smith G, Emmett P, *et al.* Progression from childhood overweight to adolescent obesity in a large contemporary cohort. *Int J Pediatr Obes* 2011;**6**:e138–43. https://doi.org/10.3109/17477166.2010.497538

85. Laitinen J, Power C, Järvelin MR. Family social class, maternal body mass index, childhood body mass index, and age at menarche as predictors of adult obesity. *Am J Clin Nutr* 2001;**74**:287–94.

86. Frisancho AR. *Anthropometric Standards for the Assessment of Growth and Nutritional Status*. Ann Arbor, MI: University of Michigan Press, 1990. https://doi.org/10.3998/mpub.12198

87. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;**320**:1240. https://doi.org/10.1136/bmj.320.7244.1240

88. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC Public Health* 2009;**9**:88. https://doi.org/10.1186/1471-2458-9-88

89. Office for National Statistics. *National Life Tables, United Kingdom: 2012–2014*. Newport: Office for National Statistics; 2015. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2015-09-23 (accessed 3 April 2016).

90. Sigurdsson E, Sigfusson N, Agnarsson U, Sigvaldason H, Thorgeirsson G. Long-term prognosis of different forms of coronary heart disease: the Reykjavik Study. *Int J Epidemiol* 1995;**24**:58–68. https://doi.org/10.1093/ije/24.1.58

91. Roper NA, Bilous RW, Kelly WF, Unwin NC, Connolly VM, South Tees Diabetes Mortality Study. Cause-specific mortality in a population with diabetes: South Tees Diabetes Mortality Study. *Diabetes Care* 2002;**25**:43–8. https://doi.org/10.2337/diacare.25.1.43

92. Hankey GJ, Jamrozik K, Broadhurst RJ, Forbes S, Burvill PW, Anderson CS, Stewart-Wynne EG. Five-year survival after first-ever stroke and related prognostic factors in the Perth Community Stroke Study. *Stroke* 2000;**31**:2080–6. https://doi.org/10.1161/01.STR.31.9.2080

93. Cancer Research UK. *Bowel Cancer Survival Statistics*. Oxford: Cancer Research UK; 2012. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival (accessed 20 June 2016).

94. Kind P, Hardman G, Macran S. *UK Population Norms for EQ-5D*. York: University of York; 1999.

95. Clarke P, Gray A, Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). *Med Decis Making* 2002;**22**:340–9. https://doi.org/10.1177/0272989X0202200412

96. Matza LS, Stewart KD, Gandra SR, Delio PR, Fenster BE, Davies EW, *et al.* Acute and chronic impact of cardiovascular events on health state utilities. *BMC Health Serv Res* 2015;**15**:173. https://doi.org/10.1186/s12913-015-0772-9

97. Djalalov S, Rabeneck L, Tomlinson G, Bremner KE, Hilsden R, Hoch JS. A review and meta-analysis of colorectal cancer utilities. *Med Decis Making* 2014;**34**:809–18. https://doi.org/10.1177/0272989X14536779

98. Beaudet A, Clegg J, Thuresson PO, Lloyd A, McEwan P. Review of utility values for economic modeling in type 2 diabetes. *Value Health* 2014;**17**:462–70. https://doi.org/10.1016/j.jval.2014.03.003

99. National Institute for Health and Care Excellence (NICE). *Section 6 Health Economics; Economic Modelling*. London: NICE; 2006.

100. Dyer MT, Goldsmith KA, Sharples LS, Buxton MJ. A review of health utilities using the EQ-5D in studies of cardiovascular disease. *Health Qual Life Outcomes* 2010;**8**:13. https://doi.org/10.1186/1477-7525-8-13

101. Anokye NK, Trueman P, Green C, Pavey TG, Hillsdon M, Taylor RS. The cost-effectiveness of exercise referral schemes. *BMC Public Health* 2011;**11**:954. https://doi.org/10.1186/1471-2458-11-954

102. Pavey TG, Anokye N, Taylor AH, Trueman P, Moxham T, Fox KR, *et al.* The clinical effectiveness and cost-effectiveness of exercise referral schemes: a systematic review and economic evaluation. *Health Technol Assess* 2011;**15**(44). https://doi.org/10.3310/hta15440

103. National Institute for Health and Care Excellence. *Apixaban for the Prevention of Stroke and Systemic Embolism in People with Non-valvular Atrial Fibrillation: STA Report*. London: NICE; 2012.

104. Post PN, Stiggelbout AM, Wakker PP. The utility of health states after stroke: a systematic review of the literature. *Stroke* 2001;**32**:1425–9. https://doi.org/10.1161/01.STR.32.6.1425

105. Adar R, Adar R, Cohen E, Kreitler S. Carotid endarterectomy for symptom-free stenosis: the patient's point of view. *Cardiovasc Surg* 1994;**2**:582–5.

106. Hallan S, Asberg A, Indredavik B, Widerøe TE. Quality of life after cerebrovascular stroke: a systematic study of patients' preferences for different functional outcomes. *J Intern Med* 1999;**246**:309–16. https://doi.org/10.1046/j.1365-2796.1999.00531.x

107. Lenert LA, Soetikno RM. Automated computer interviews to elicit utilities: potential applications in the treatment of deep venous thrombosis. *J Am Med Inform Assoc* 1997;**4**:49–56. https://doi.org/10.1136/jamia.1997.0040049

108. Man-Son-Hing M, Laupacis A, O'Connor AM, Coyle D, Berquist R, McAlister F. Patient preference-based treatment thresholds and recommendations: a comparison of decision-analytic modeling with the probability-tradeoff technique. *Med Decis Making* 2000;**20**:394–403. https://doi.org/10.1177/0272989X0002000403

109. Edwards SJ, Hamilton V, Trevor N, Nherera L, Karner C, Thurgar E. *Apixaban for the Prevention of Stroke and Systemic Embolism in People with Non-valvular Atrial Fibrillation: A Single Technology Appraisal*. London: BMJ-TAG; 2012. URL: www.nice.org.uk/guidance/ta275/documents/stroke-and-systemic-embolism-prevention-nonvalvular-atrial-fibrillation-apixaban-evidence-review-group-report2 (accessed 1 November 2016).

110. National Institute for Health and Care Excellence. *Modelling the Cost Effectiveness of Physical Activity Interventions*. London: NICE; 2006 (accessed 15 May 2016).

111. National Institute for Health and Care Excellence. *Cetuximab for the First-line Treatment of Metastatic Colorectal Cancer*. London: NICE; 2009 (accessed 15 May 2016).

112. Kanavos P, van den Aardweg S, Schurer W. *Diabetes Expenditure, Burden of Disease and Management in 5 EU Countries*. London: London School of Economics (Health); 2012.

113. NHS Scotland. *Management of Coronary Heart Disease: A National Clinical and Resource Impact Assessment*. Edinburgh: NHS Scotland; 2007.

114. Saka O, McGuire A, Wolfe C. Cost of stroke in the United Kingdom. *Age Ageing* 2009;**38**:27–32. https://doi.org/10.1093/ageing/afn281

115. Cancer Research UK. *Saving Lives through Research*. London: Cancer Research UK; 2014. URL: www.cancerresearchuk.org/sites/default/files/saving_lives_averting_costs.pdf (accessed 17 June 2016).

116. Currie CJ, Gale EA, Poole CD. Estimation of primary care treatment costs and treatment efficacy for people with Type 1 and Type 2 diabetes in the United Kingdom from 1997 to 2007*. *Diabet Med* 2010;**27**:938–48. https://doi.org/10.1111/j.1464-5491.2010.03040.x

117. Morgan CL, Peters JR, Dixon S, Currie CJ. Estimated costs of acute hospital care for people with diabetes in the United Kingdom: a routine record linkage study in a large region. *Diabet Med* 2010;**27**:1066–73. https://doi.org/10.1111/j.1464-5491.2010.03086.x

118. Bosanquet N, Franks P. *Stroke Care: Reducing the Burden of Disease*. London: The Stroke Association; 1998.

119. Trueman P, Lowson K, Bending M, Ganderton M, Chaplin S, Wright DH, *et al. Bowel Cancer Services: Costs and Benefits (Summary Report to the Department of Health)*. York: York Health Economics Consortium and School of Health and Related Research (University of Sheffield); 2007.

120. Power C, Lake JK, Cole TJ. Body mass index and height from childhood to adulthood in the 1958 British born cohort. *Am J Clin Nutr* 1997;**66**:1094–101.

121. Health and Social Care Information Centre. *Health Survey for England 2009: Health and Lifestyles, Summary of Key Findings*. URL: http://content.digital.nhs.uk/catalogue/PUB00414/heal-surv-heal-life-eng-2009-rep-v1.pdf (accessed 6 September 2017).

122. Bonell CP, Fletcher A, Jamal F, Wells H, Harden A, Murphy S, Thomas J. Theories of how the school environment impacts on student health: systematic review and synthesis. *Health Place* 2013;**24**:242–9. https://doi.org/10.1016/j.healthplace.2013.09.014

123. Fisher J, Fisher W. Modifiable determinants of behaviour: information, motivation, behaviour skills model. *Psychol Bull* 1992.

124. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. https://doi.org/10.1191/1478088706qp063oa

125. Richie J, Spencer L. Qualitative Data Analysis for Applied Policy Research. In Bryman A, Burgess R, editors. *Analysing Qualitative Data*. London: Routledge; 1994. pp. 173–94. https://doi.org/10.4324/9780203413081_chapter_9

126. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**:1173–82. https://doi.org/10.1037/0022-3514.51.6.1173

127. Cole DA, Maxwell SE. Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *J Abnorm Psychol* 2003;**112**:558–77. https://doi.org/10.1037/0021-843X.112.4.558

128. Pyper E, Harrington D, Manson H. The impact of different types of parental support behaviours on child physical activity, healthy eating, and screen time: a cross-sectional study. *BMC Public Health* 2016;**16**:568. https://doi.org/10.1186/s12889-016-3245-0

129. Aneshensel CS. *Theory-Based Data Analysis for the Social Sciences*. Los Angeles, CA: Sage; 2012.

130. Tennant AH, Horton M, Pallant J. *An Introduction to Structural Equation Modelling (SEM) using AMOS*. Leeds: Psychometric Laboratory for Health Sciences, University of Leeds; 2013.

131. Marsh HW, Balla JR, McDonald RP. Goodness-of-fit indices in confirmatory factor analysis: the effect of sample size. *Psychol Bull* 1988;**103**:391–410. https://doi.org/10.1037/0033-2909.103.3.391

132. Hu L, Bentler P. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol Methods* 1998;**3**:424–53. https://doi.org/10.1037/1082-989X.3.4.424

133. Wheaton B. Assessment of fit in overidentified models with latent variables. *Sociol Methods Res* 1987;**16**:118–54. https://doi.org/10.1177/0049124187016001005

134. Kock N, Gaskins L. Simpson's paradox, moderation, and the emergence of quadratic relationships in path models: an information systems illustration. *Int J Appl Nonlinear Sci* 2016;**2**:200–34. https://doi.org/10.1504/IJANS.2016.077025

135. Snell DL, Surgenor LJ, Hay-Smith EJ, Williman J, Siegert RJ. The contribution of psychological factors to recovery after mild traumatic brain injury: is cluster analysis a useful approach? *Brain Inj* 2015;**29**:291–9. https://doi.org/10.3109/02699052.2014.976594

136. Williamson DA, Champagne CM, Harsha D, Han H, Martin CK, Newton R, *et al.* Louisiana (LA) Health: design and methods for a childhood obesity prevention program in rural schools. *Contemp Clin Trials* 2008;**29**:783–95. https://doi.org/10.1016/j.cct.2008.03.004

137. Doolen J, Alpert PT, Miller SK. Parental disconnect between perceived and actual weight status of children: a metasynthesis of the current research. *J Am Acad Nurse Pract* 2009;**21**:160–6. https://doi.org/10.1111/j.1745-7599.2008.00382.x

138. Williams AJ. *Determination of School-based Contextual Factors and Their Associations with the Prevalence of Overweight and Obese Children*. PhD thesis. Exeter: Peninsula College of Medicine and Dentistry; 2014.

139. Waters E, de Silva-Sanigorski A, Hall BJ, Brown T, Campbell KJ, Gao Y, *et al.* Interventions for preventing obesity in children. *Cochrane Database Syst Rev* 2011;**12**:CD001871. https://doi.org/10.1002/14651858.CD001871.pub3

140. Te Velde SJ, Brug J, Wind M, Hildonen C, Bjelland M, Perez-Rodrigo C, *et al.* Effects of a comprehensive fruit- and vegetable-promoting school-based intervention in three European countries: the Pro Children Study. *Br J Nutr* 2008;**99**:893–903. https://doi.org/10.1017/S000711450782513X

141. Lytle LA, Murray DM, Perry CL, Story M, Birnbaum AS, Kubik MY, Varnell S. School-based approaches to affect adolescents' diets: results from the TEENS study. *Health Educ Behav* 2004;**31**:270–87. https://doi.org/10.1177/1090198103260635

142. Public Health England. *NCMP Local Authority Profile*. URL: http://fingertips.phe.org.uk/profile/national-child-measurement-programme/data#page/0/gid/8000011/pat/6/par/E12000009/ati/102/are/E10000008 (accessed 1 November 2016).

143. Williams AJ, Henley WE, Williams CA, Hurst AJ, Logan S, Wyatt KM. Systematic review and meta-analysis of the association between childhood overweight and obesity and primary school diet and physical activity policies. *Int J Behav Nutr Phys Act* 2013;**10**:101. https://doi.org/10.1186/1479-5868-10-101

144. Lloyd J. *Preventing Childhood Obesity: Developing Complex Interventions for Behaviour Change*. Exeter: University of Exeter; 2013.

145. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*: Cambridge: Cambridge University Press; 2011. https://doi.org/10.1017/CBO9780511996214

146. Parcel GS, Edmundson E, Perry CL, Feldman HA, O'Hara-Tompkins N, Nader PR, *et al.* Measurement of self-efficacy for diet-related behaviors among elementary school children. *J Sch Health* 1995;**65**:23–7. https://doi.org/10.1111/j.1746-1561.1995.tb03335.x

147. Sallis JF, Grossman RM, Pinski RB, Patterson TL, Nader PR. The development of scales to measure social support for diet and exercise behaviors. *Prev Med* 1987;**16**:825–36. https://doi.org/10.1016/0091-7435(87)90022-3

148. Schutz R, Smoll F, Carre F, Mosher R. Inventories and norms for children's attitudes toward physical activity. *Res Q Exerc Sport* 1985;**56**:256–65. https://doi.org/10.1080/02701367.1985.10605372

149. Biddle S, Akande D, Armstrong N, Ashcroft M, Brooke R, Goudas M. The self-motivation inventory modified for children: evidence on psychometric properties and its use in physical exercise. *Int J Sport Psychol* 1996;**27**:237–50.

150. Fisher JD, Fisher WA, Bryan A, Misovich SJ. Information-motivation-behavioral skills model-based HIV risk behavior change intervention for inner-city school youth. *Health Psychol* 2002;**21**:177–86. https://doi.org/10.1037/0278-6133.21.2.177

151. Jöreskog KG, Sörbom D. *LISREL 8: Structural Equation Modelling with the SIMPLIS Command Language*. Hillsdale, NJ: Erlbaum Associates; 1993.

152. Jöreskog KG. New developments in LISREL: analysis of ordinal variables using polychoric correlations and weighted least squares. *Qual Quant* 1990;**24**:387–404. https://doi.org/10.1007/BF00152012

153. Milfont TL, Fischer R. Testing measurement invariance across groups: applications in cross-cultural research. *Int J Psychol Res* 2010;**3**:111–21. https://doi.org/10.21500/20112084.857

154. Fletcher A, Wolfenden L, Wyse R, Bowman J, McElduff P, Duncan S. A randomised controlled trial and mediation analysis of the 'Healthy Habits', telephone-based dietary intervention for preschool children. *Int J Behav Nutr Phys Act* 2013;**10**:43. https://doi.org/10.1186/1479-5868-10-43

# Appendix 1 Trial Steering Committee

| Name (title) | Position | Department/faculty | Institution |
|---|---|---|---|
| Professor Martin White[a] (chairperson) | Professor of Public Health | MRC Epidemiology Unit | University of Cambridge |
| Dr Sara Brookes[a] | Senior Lecturer in Medical Statistics | Department of Social Medicine | University of Bristol |
| Dr Russ Jago[a] | Professor of Paediatric Physical Activity & Public Health | Exercise, Nutrition and Health Sciences | University of Bristol |
| Dr Chris Moudiotis[a] | Consultant Paediatrician | Acute Paediatrics | Royal Devon and Exeter NHS Foundation Trust, Exeter |
| Mr Christopher Gardner | Research & Development Directorate Manager | Research and Development | Royal Devon and Exeter NHS Foundation Trust, Exeter |
| Dr Katrina Wyatt (principal investigator) | Associate Professor in Child Health | Institute for Health Research (Child Health Group) | University of Exeter Medical School |
| Dr Jennifer Lloyd (trial manager) | Senior Research Fellow in Child Health | Institute for Health Research (Child Health Group) | University of Exeter Medical School |
| Anne Billington[a] | Head teacher | | Offwell Primary School |
| Claire O'Keeffe[a] | Parent | | |

a Independent member.

# Appendix 2 Food Intake Questionnaires

## Food Intake Questionnaire: weekday

Date:                                            Time point:

## What did YOU eat and drink yesterday?

Child ID:

School:

Class:

## WEEK

We are interested in finding out what you ate yesterday. Please read and listen to the questions very carefully and answer as honestly as you can. There are no right or wrong answers and all your responses will be kept secret.

Please answer all questions by putting a circle around either **YES** or **NO**.

## Yesterday did you have anything to;

1. Eat or drink before leaving home to come to school?    Y   N

2. Eat or drink on your way to school?    Y   N

## Yesterday, did you;

3. Eat a school lunch?    Y   N

4. Eat a packed lunch from home?    Y   N

5. Go home for your lunch?    Y   N

## Did you at any time yesterday eat any amount of the following?

### Breakfast Cereals

6. Any cereals with:
Chocolate flavouring, sugar frosted coatings; a crunchy, crisp texture; hoop shapes; added nuts/honey/maple; fruit fillings?    Y   N

7. Any plain cereals
e.g. Branflakes, Weetabix, All Bran, Nestle Force, Fruit 'n' Fibre, Sultana Bran (including supermarket own brands), Shredded Wheat, muesli, malted wheat cereals, porridge, instant oat cereal e.g. Oatso Simple, Ready Brek, Cornflakes, Rice Crispies, Rice Pops, Puffed Wheat, Special K?    Y   N

## Bread

8. White bread (slices or rolls)?

     Y    N

9. Brown or wholemeal bread any type (slices or rolls)?

     Y    N

10. Malt/fruit bread?

     Y    N

11. Bread sticks/crackers?

     Y    N

## Butter or Margarine

12. Butter/ Margarine?

     Y    N

## Biscuits

13. Biscuits e.g. malted milk, Digestives, Rich Tea etc.?

     Y    N

14. Chocolate covered biscuits, e.g. Kit Kat, Penguin, Club etc?

     Y    N

# Did you at any time yesterday eat any amount of the

## Cakes and Puddings

15. Any sort of cake: Swiss roll (plain or chocolate), doughnuts, scones, individual pies, jam/custard tarts etc?

     Y    N

16. Any sort of pudding: Fruit pie, sponge pudding, tinned fruit, jelly, trifle, lemon meringue, cheese cake, milk pudding (like rice, semolina, custard) etc?

     Y    N

## Sweets and chocolates

17. Sweets such as boiled sweets, fruit gums or pastilles, liquorice, jelly sweets, chews, toffees, etc?    Y    N

18. Chocolates or chocolate bars like Quality Street, Rolos, Mars Bars, Twix?    Y    N

19. Ice cream, choc ices, ice lollies, ice pops?    Y    N

## Added Sugar

20. Sugar (white or brown) in any drink such as tea, coffee, hot chocolate?   Y   N

21. Sugar (white or brown) on any food such as cornflakes or pancakes?   Y   N
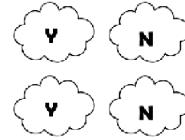
## Potatoes

22. Boiled potatoes?   Y   N

23. Mashed potatoes?   Y   N

24. Baked or jacket potatoes?   Y   N

25. Roast potatoes?   Y   N

26. Chips?   Y   N

27. Crisps (any type or flavour)?   Y   N

## Starchy Foods

28. Pasta?   Y   N

29. Rice?   Y   N

30. Noodles?   Y   N

31. Homemade pizza?   Y   N

32. Shop bought pizza?   Y   N

## Nuts

Nuts such as peanuts, cashews, almonds, brazils, walnuts, etc?

33. Salted?   Y   N

34. Unsalted?   Y   N

## Did you at any time yesterday eat any amount of the following?

## Fruit

35. Any fresh fruit such as apples, oranges (any type), pears, bananas, plums etc?   Y   N

36. Any dried fruit such as raisins, dried apricots, sultanas etc?   Y   N
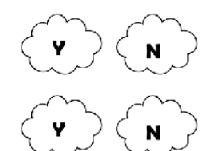
## Vegetables

37. Any type of salad such as celery, tomatoes, lettuce, cucumber etc?

Y    N

38. Any fried vegetables e.g. fried onions, fried mushrooms or fried tomatoes etc?

Y    N

39. Any other vegetables e.g. peas, cabbage, carrots, leeks, green beans, kidney beans, parsnips, tinned tomatoes, cauliflower, turnips, sprouts etc?

Y    N

## Fish

46. Fish fried in batter?

Y    N

47. Any other types of fish e.g. fish fingers or tinned (like sardines, tuna, pilchards) etc?

Y    N

## Cheese

50. Hard Cheese e.g. Cheddar, Red Leicester, Cheshire?

Y    N

51. Processed cheese e.g. Cheese Strings, cheese slices, Dairylea?

Y    N

## Did you at any time yesterday eat any amount of the following?

## Meat

40. Shop bought burger?

Y    N

41. Shop bought sausages?

Y    N

42. Homemade burger?

Y    N

43. Homemade sausages?

Y    N

44. Meat pie, Cornish pasty, pork pie or sausage roll etc?

Y    N

45. Any other type of meat e.g. minced meat, steak, ham, chicken etc?

Y    N

## Eggs

48. Boiled, poached or scrambled?

Y    N

49. Fried?

Y    N

### Yoghurt

52.    Yoghurt?    Y    N

## Did you at any time yesterday eat any amount of any of the following?

### Takeaway food

53.    Chip shop food or Chinese, Indian, Curries, Pizza, Kebabs etc?

Y    N

54.    Did you put any salt on your food?

Y    N

### Still squashes
(which you add water to like orange squash, Ribena, barley water etc)?

56.    No sugar squash?

Y    N

57.    Regular squash?

Y    N

### Sweetened Fizzy drinks
(like lemonade, coke, fanta, 7UP, soda stream etc)?

55.    Sweetened fizzy drink?

Y    N

### Milk
(including milk in tea, coffee, milkshakes, flavoured milk, hot chocolate or on cereals etc)?

58.    Ordinary full fat milk (Blue top)?    Y    N

59     Semi skimmed or skimmed milk (Green or red top)?    Y    N

60.    Water?    Y    N

## Food Intake Questionnaire: weekend

Date:                                          Time point:

# What did YOU eat and drink yesterday?

Child ID:

School:

Class:

# WEEKEND

We are interested in finding out what you ate yesterday. Please read and listen to the questions very carefully and answer as honestly as you can. There are no right or wrong answers and all your responses will be kept secret.

Please answer all
questions by
putting a circle around
either **YES** or **NO**.

## Yesterday, did you:

1.  Eat lunch?    Y    N

## Did you at any time yesterday eat any amount of the following?

### Breakfast Cereals

2.  Any cereals with:
    Chocolate flavouring, sugar frosted coatings; a crunchy,
    crisp texture; hoop shapes; added nuts/honey/maple;
    fruit fillings?    Y    N

3.  Any plain cereals
    e.g. Branflakes, Weetabix, All Bran, Nestle Force, Fruit 'n'
    Fibre, Sultana Bran (including supermarket own brands),
    Shredded Wheat, muesli, malted wheat cereals, porridge,
    instant oat cereal e.g. Oatso Simple, Ready Brek, Cornflakes,
    Rice Crispies, Rice Pops, Puffed Wheat, Special K?    Y    N
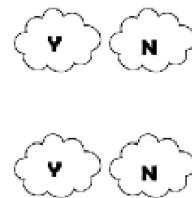
### Bread

4.  White bread (slices or rolls)?    Y    N

5.  Brown or wholemeal bread any type (slices or rolls)?    Y    N

6.  Malt/fruit bread?    Y    N

7.  Bread sticks/crackers?    Y    N

## Did you at any time yesterday eat any amount of the following?
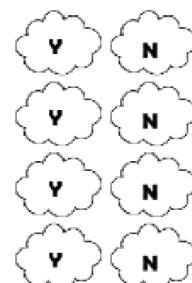
### Butter or Margarine

8.　　Butter / Margarine?　　Y　N

### Cakes and Puddings

11.　　Any sort of cake: Swiss roll (plain or chocolate), doughnuts, scones, individual pies, jam/custard tarts etc?

Y　N

12.　　Any sort of pudding: Fruit pie, sponge pudding, tinned fruit, jelly, trifle, lemon meringue, cheesecake, milk pudding (like rice, semolina, custard) etc?

Y　N

### Biscuits

9.　　Biscuits e.g. malted milk, Digestives, Rich Tea etc.?

Y　N

10.　　Chocolate covered biscuits e.g. Kit Kat, Penguin, Club etc?

Y　N

### Sweets and chocolates

13.　　Sweets such as boiled sweets, fruit gums or pastilles, liquorice, jelly sweets, chews, toffees, etc?　　Y　N

14.　　Chocolates or chocolate bars like Quality Street, Rolos, Mars Bars, Twix?　　Y　N

15.　　Ice cream, choc ices, ice lollies, ice pops?　　Y　N

## Added Sugar

16.     Sugar (white or brown) in any drink such as tea, coffee, hot chocolate?     Y    N

17.     Sugar (white or brown) on any food such as cornflakes or pancakes?     Y    N
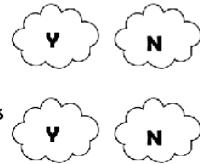
## Potatoes

18.     Boiled potatoes?    Y    N

19.     Mashed potatoes?    Y    N

20.     Baked or jacket potatoes?    Y    N

21.     Roast potatoes?    Y    N

22.     Chips?    Y    N

23.     Crisps (any type or flavour)?    Y    N

## Starchy Foods

24.     Pasta?    Y    N

25.     Rice?    Y    N

26.     Noodles?    Y    N

27.     Homemade pizza?    Y    N

28.     Shop bought pizza?    Y    N

## Nuts

Nuts such as peanuts, cashews, almonds, brazils, walnuts, etc?

29.     Salted?    Y    N

30.     Unsalted?    Y    N

## Did you at any time yesterday eat any amount of the following?

## Fruit

31.     Any fresh fruit such as apples, oranges (any type), pears, bananas, plums etc?    Y    N

32.     Any dried fruit such as raisins, dried apricots, sultanas etc?    Y    N

# Did you at any time yesterday eat any amount of the following?
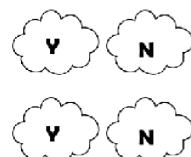
## Vegetables

33. Any type of salad such as celery, tomatoes, lettuce, cucumber etc?

Y    N

34. Any fried vegetables e.g. fried onions, fried mushrooms or fried tomatoes etc?

Y    N

35. Any other vegetables e.g. peas, cabbage, carrots, leeks, green beans, kidney beans, parsnips, tinned tomatoes, cauliflower, turnips, sprouts etc?

Y    N
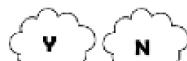
## Fish

42. Fish fried in batter?    Y    N

43. Any other types of fish e.g. fish fingers or tinned (like sardines, tuna, pilchards) etc?

Y    N

## Cheese

46. Hard Cheese e.g. Cheddar, Red Leicester, Cheshire?    Y    N

47. Processed cheese e.g. Cheese Strings, cheese slices, Dairylea?

Y    N

## Meat

36. Shop bought burger?

Y    N

37. Shop bought sausages?

Y    N
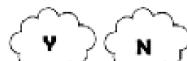
38. Homemade burger?

Y    N

39. Homemade sausages?

Y    N

40. Meat pie, Cornish pasty, pork pie or sausage roll etc?

Y    N

41. Any other type of meat e.g. minced meat, steak, ham, chicken etc?

Y    N

## Eggs

44. Boiled, poached or scrambled?

Y    N

45. Fried?    Y    N

**Yoghurt**

48.  Yoghurt?   Y   N

# Did you at any time yesterday eat any amount of any of the following?

**Takeaway food**

49.  Chip shop food or Chinese, Indian, Curries, Pizza, Kebabs etc?

Y   N

50.  Did you put any salt on your food?

Y   N

**Still squashes**
(which you add water to like orange squash, Ribena, barley water etc)?

52.  No sugar squash?

Y   N

53.  Regular squash?

Y   N

**Sweetened Fizzy drinks**
(like lemonade, coke, fanta, 7UP, soda stream etc)?

51.  Sweetened fizzy drink?

Y   N

**Milk**
(including milk in tea, coffee, milkshakes, flavoured milk, hot chocolate or on cereals etc)?

54.  Ordinary full fat milk (Blue top)?   Y   N

55.  Semi skimmed or skimmed milk (Green or red top)?   Y   N

56.  Water?   Y   N

## Food Intake Questionnaire scoring

| Type of food | FIQ weekday item | FIQ weekend item |
|---|---|---|
| Energy dense snacks (*n* = 13) | | |
| High sugar cereals | 6 | 2 |
| Plain biscuits | 13 | 9 |
| Chocolate biscuits | 14 | 10 |
| Cakes | 15 | 11 |
| Puddings | 16 | 12 |
| Boiled sweets | 17 | 13 |
| Chocolate | 18 | 14 |
| Ice cream | 19 | 15 |
| Crisps | 27 | 23 |
| Salty nuts | 33 | 29 |
| Pies/pasties | 44 | 40 |
| Processed cheese | 51 | 47 |
| Sweetened fizzy drinks | 55 | 51 |
| Healthy snack foods (*n* = 10) | | |
| Brown/wholemeal bread | 9 | 5 |
| Malt/fruit bread | 10 | 6 |
| Breadsticks/crackers | 11 | 7 |
| Unsalted nuts | 34 | 30 |
| Fresh fruit | 35 | 31 |
| Dried fruit | 36 | 32 |
| Salad | 37 | 33 |
| Hard cheese | 50 | 46 |
| Yoghurt | 52 | 48 |
| Semiskimmed milk | 59 | 55 |
| Negative marker foods (*n* = 25) | | |
| Sugar cereals | 6 | 2 |
| Butter/margarine | 12 | 8 |
| Biscuits | 13 | 9 |
| Chocolate biscuits | 14 | 10 |
| Cakes | 15 | 11 |
| Puddings | 16 | 12 |
| Boiled sweets | 17 | 13 |
| Chocolate | 18 | 14 |
| Ice cream | 19 | 15 |
| Sugar added to drinks | 20 | 16 |
| Sugar added to food | 21 | 17 |

| Type of food | FIQ weekday item | FIQ weekend item |
|---|---|---|
| Roast potatoes | 25 | 21 |
| Crisps | 27 | 23 |
| Chips | 26 | 22 |
| Salted nuts | 33 | 29 |
| Fried vegetables | 38 | 34 |
| Shop-bought burger | 40 | 36 |
| Shop-bought sausage | 41 | 37 |
| Pies and pasties | 44 | 40 |
| Fried fish | 46 | 42 |
| Fried egg | 49 | 45 |
| Processed cheese | 51 | 47 |
| Takeaways | 53 | 49 |
| Salt added to food | 54 | 50 |
| Sweet fizzy drink | 55 | 51 |
| Positive marker foods ($n = 22$) | | |
| Low-sugar cereals | 7 | 3 |
| Brown/wholemeal bread | 9 | 5 |
| Malt/fruit loaf | 10 | 6 |
| Breadsticks/crackers | 11 | 7 |
| Boiled potatoes | 22 | 18 |
| Mashed potatoes | 23 | 19 |
| Baked potatoes | 24 | 20 |
| Pasta | 28 | 24 |
| Rice | 29 | 25 |
| Noodles | 30 | 26 |
| Homemade pizza | 31 | 27 |
| Unsalted nuts | 34 | 30 |
| Fresh fruit | 35 | 31 |
| Dried fruit | 36 | 32 |
| Salad | 37 | 33 |
| Vegetables | 39 | 35 |
| Homemade burgers | 42 | 38 |
| Homemade sausages | 43 | 39 |
| Yogurt | 52 | 48 |
| No-sugar squash | 56 | 52 |
| Semiskimmed milk | 59 | 55 |
| Water | 60 | 56 |

# Appendix 3 My Lifestyle Questionnaire

## My Lifestyle Questionnaire

Date: _____　　　　Time point: _____

# My Lifestyle Questionnaire

Child ID: _____

Child DOB: _____

School: _____

Class: _____

We are interested in finding out about your lifestyle. Please listen to the questions very carefully and answer as best you can. Please note that none of your answers will be shared.

## <u>Section A</u>

For questions 1 and 2 you are given 2 lists (A and B). You have to **match up the items in list B to list A** by writing in the correct number as shown in the example below.

E.g. Sport and equipment

Example

| A—Sport | | B—Equipment |
|---|---|---|
| Tennis | __5__ | 1. Goggles |
| Football | __2__ | 2. Boots |
| Snooker | __4__ | 3. Beam |
| Swimming | __1__ | 4. Cue |
| Gymnastics | __3__ | 5. Racquet |

**1.    Unhealthy food and healthy replacement.**

From list B, please choose what you believe to be the healthy alternative that **best matches** the unhealthy food item.

| A – Unhealthy Food | | B—Healthy Alternatives |
|---|---|---|
| Orange tango | _____ | 1. Portion of hard cheese |
| Packet of crisps | _____ | 2. Dried fruit e.g. apricots/raisins |
| Biscuits | _____ | 3. Ham sandwich (wholemeal bread) |
| Jelly sweets | _____ | 4. Bread sticks |
| Jam sandwich (white bread) | _____ | 5. Sweet oatcakes |
| Cheese strings | _____ | 6. Sugar free orange squash and |
| | |     fizzy water |

**2.  Food group and the percentage of each that should make up your diet.**

From List B choose the correct percentage for each food group.
You may use a percentage more than once. Your answers should add up to 100%.

| A | | B | |
|---|---|---|---|
| Fruits and vegetables | _____% | 12% | 33% |
| Fatty and sugary foods | _____% | 7% | 11% |
| Meat, fish and alternatives | _____% | 40% | 15% |
| Milk and dairy products | _____% | 50% | 2% |
| Bread, other cereals and potatoes | _____% | | |

**3.** The activity triangle below has 4 levels to show very little time at the top with more time in each level as you go down. In the box to the right of the triangle there are 4 different types of activity. **Please write the letter linked to each activity type in one section of the triangle to show how much time you should spend on that type of activity.** You can only use each letter once and have one letter in each level of the pyramid.

**Very little time**

**Lots of time**
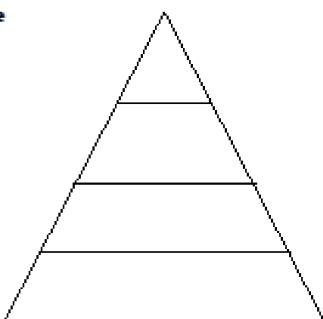
| A | Games and activities that make the heart beat faster e.g. playing football/netball |
| B | Daily activities e.g. Walking |
| C | Not moving around for more than 2 hours in the day |
| D | Activities which involve stretching and building strength e.g. gymnastics, judo, weight training |

**4.** What is the **smallest** amount of healthy foods we should be eating? What is the **largest** amount of unhealthy foods we should be eating?

Choose your answers from the list below and write your answer in the appropriate box. Your answer should add up to 100%.

15% 20% 25% 30% 35%

40% 45% 50% 55% 60%

65% 70% 75% 80% 85%

| Healthy food | to | Unhealthy food |
| % | | % |
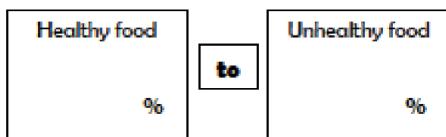
**5.** Imagine you have decided to reduce your screen time on a school day evening by playing outside instead. In the box below **write down 3 things that would help you to remember** to play outside.

1.

2.

3.

## Section B

Questions 6-8 are statements about **how easy or hard it would be for you to do certain things.** You must <u>TICK</u> the box that shows how much you agree with the statement.

- *Unhealthy snacks include biscuits, cakes, crisps, sweets and sweetened fizzy drinks (e.g. coke, fanta, lemonade)*
- *Healthy snacks include fruit, crackers, breadsticks, oatcakes, dried fruit, carrot sticks, wholemeal bread*
- *Screen based activities include TV, computer, mobile phone, Xbox, Nintendo DS, Ipad etc.*

| No. | Question | Strongly agree | Slightly agree | Slightly disagree | Strongly disagree |
|-----|----------|----------------|----------------|-------------------|-------------------|
| 6 | I would find it **very easy** to give up half of my screen based activities and swap them with being more active everyday | | | | |
| 7 | I would find it **very easy** to eat only one unhealthy snack every day | | | | |
| 8 | If offered my favourite unhealthy snack I would find it **very easy** to say 'no thanks' | | | | |

Questions 9-14 are statements **about whether you plan to do certain things in the next two weeks.**
You must TICK the box that shows how much you agree with the statement.

| No. | Question | Strongly agree | Slightly agree | Slightly disagree | Strongly disagree |
|-----|----------|----------------|----------------|-------------------|-------------------|
| 9 | I plan to give up at least half of **my screen based activities** and replace them with **physical activity** | | | | |
| 10 | I plan to only eat **one unhealthy snack** every day | | | | |
| 11 | If offered my favourite unhealthy snack, I will say **'no thanks'** | | | | |
| 12 | I plan to eat **new** healthy snacks | | | | |
| 13 | I plan to talk to my parents about how the whole family can **eat more healthily** | | | | |
| 14 | I plan to try to persuade my parents to be **more active** | | | | |

Questions 15-17 are statements **about how many of your friends do certain things.**
You must <u>TICK</u> the box which shows how many of your friends do these things.

| No. | Question | None | A few | Lots | All |
|---|---|---|---|---|---|
| 15 | How many of your friends eat more than two **healthy** snacks at school each day? | | | | |
| 16 | How many of your friends drink a fizzy drink (e.g. coke, lemonade , Fanta) **every day**? | | | | |
| 17 | How many of your friends **play on a computer, mobile phone or watch TV** when they get home from school? | | | | |

Questions 18-22 are statements about **whether your friends think it's ok to do certain things.** You must <u>TICK</u> the box that shows how much you agree with the statement.

| No. | Question | Strongly agree | Slightly agree | Slightly disagree | Strongly disagree |
|---|---|---|---|---|---|
| 18 | Most of my friends think it is ok to **play on the computer/mobile phone** for most of the time at home | | | | |
| 19 | Most of my friends think it is ok to **help their parents cook a meal** | | | | |
| 20 | Most of my friends think it is ok to **help their parents shop for healthy foods** | | | | |
| 21 | Most of my friends think its ok to **watch TV** for most of the time at home | | | | |
| 22 | Most of my friends think it is ok to **help their parents with the household chores** e.g. hoovering, laying the table, tidying their bedroom | | | | |

Questions 23-25 are statements **about whether your family think it is ok to do certain things.**
You must <u>TICK</u> the box which shows how much you agree with the statement.

| No. | Question | Strongly agree | Slightly agree | Slightly disagree | Strongly disagree |
|---|---|---|---|---|---|
| 23 | Most of my family think it is ok for me to eat more than three **unhealthy** snacks (e.g. crisps, chocolate, biscuits, sweets) every day | | | | |
| 24 | Most of my family think it is ok for me to drink more than two **sweetened fizzy drinks** each day | | | | |
| 25 | Most of my family think it is ok for me to <u>spend more than 2 hours</u> each evening on **screen-based activities** | | | | |

Questions 26-28 are statements **about eating and activity**.
You must <u>TICK</u> the box which shows how much you agree with the statement.

| No. | Question | Strongly agree | Slightly agree | Slightly disagree | Strongly disagree |
|---|---|---|---|---|---|
| 26 | If you want to eat lots of food high in salt, sugar and fat then you should be able to | | | | |
| 27 | If you want to drink sweetened fizzy drinks then you should be able to | | | | |
| 28 | If you want to spend more than 3 hours a day doing screen based activities then you should be able to | | | | |

## Section C

Questions 29-32 are statements **about your home.**
You must <u>TICK</u> the box which shows how much you agree with the statement.

| No. | Question | Strongly agree | Slightly agree | Slightly disagree | Strongly disagree |
|---|---|---|---|---|---|
| 29 | My parents/carers provide mostly **healthy snacks** for me at home | | | | |
| 30 | My parents/carers provide mostly **unhealthy drinks** for me at home | | | | |
| 31 | My parents/carers provide **many opportunities** for me to be physically active | | | | |
| 32 | My parents/carers have **rules at home** to help me eat more healthily and to reduce my screen time e.g. I can only spend half an hour on a computer game each day or I am not allowed to help myself to crisps, chocolate, biscuits and sweets whenever I want. | | | | |

## Section D

Questions 33–50 are statements **about what you do**.
You must tick the box which shows how often you do these things.

| No. | Question | All of the time | Most of the time | Some of the time | Never |
|---|---|---|---|---|---|
| 33 | I set targets to help me cut down on the number of unhealthy snacks I eat | | | | |
| 34 | I set targets to help me cut down on the number of sweetened fizzy drinks I have | | | | |
| 35 | I set targets to help me cut down on my screen time | | | | |
| 36 | I set targets to help me do more physical activity | | | | |
| 37 | If I set a target, I write it down | | | | |
| 38 | If I set a target, I tell/show my parents/carers and ask them to agree it | | | | |
| 39 | I try to remember each type of food I eat to make sure I do not eat too many unhealthy snacks | | | | |
| 40 | I try to remember each type of food I eat to make sure I eat enough healthy food | | | | |
| 41 | I try to remember how much screen-based activity I do to make sure I don't do too much | | | | |
| 42 | I try to remember how much physical activity I do to make sure I do enough | | | | |
| 43 | I talk to my parents/carers about eating more healthily | | | | |
| 44 | I talk to my parents/carers about increasing my physical activity | | | | |
| 45 | I talk to my parents/carers about spending less time on screen-based activities | | | | |
| 46 | I make suggestions to my parents/carers about how we can be more active as a family | | | | |
| 47 | I make suggestions to my parents/carers about how we can eat more healthily | | | | |
| 48 | I go food shopping with my parents/carers and help them choose healthy foods and drinks to buy | | | | |
| 49 | I help my parents/carers to cook at home | | | | |
| 50 | I try new healthy foods and drinks | | | | |

## My Lifestyle Questionnaire scoring

High scores represent health promoting cognition, motivation, attitudes and behaviours

**Section A – Knowledge (one point for each correct answer)**

Answers

**Q1**

Orange tango – 6

Packet of crisps – 4

Biscuits – 5

Jelly sweets – 2

Jam sandwich – 3

Cheese strings - 1

**Q2**

Fruits and veg – 33%

Fatty foods and sugary foods – 7%

Meat, fish and alternatives – 12%

Milk and dairy products – 15%

Bread, other cereals and potatoes – 33%

**Q3**

From top of triangle to bottom

C

D

A

B

**Q4**

Healthy food – 80%

Unhealthy food – 20%

**Q5**

Each strategy listed below will be counted as 1 point. In order to score 3, there needs to be 3 separate strategies. For example if a child writes removing the X Box, removing the TV and hiding the remote control for her 3 strategies this will only count as 1 point.

Any answer indicating leaving a **note/reminders** to themselves *e.g. putting a post it note on the computer telling me to play outside instead*

Any answer suggesting they **write their goal** (of replacing screen time on a school day evening with playing outside instead) on a piece of paper. This would include getting their parents to sign their goal

Any answer suggesting a **stimulus cue** e.g. placing a football/kit/trainers etc in a place that reminds them they should play outside rather than do screen-based activity

Any answer suggesting **removing or hiding TV/computer/phone/Xbox/DS** from a room

Any answer suggesting getting **parents/friends/sibling to remind them** of their goals

**Sections B-D**

| Points | 4 | 3 | 2 | 1 |
|---|---|---|---|---|
| | Strongly agree | Slightly agree | Slightly disagree | Strongly disagree |
| | All | Lots | A few | None |
| | All of the time | Most of the time | Some of the time | Never |

**Reverse score items:** 16, 17, 18, 21, 23, 24, 25, 26, 27, 28, 30

Range of possible scores and associated questions for each construct

| Construct | Minimum score | Maximum score | Questions |
|---|---|---|---|
| Knowledge | 0 | 20 | 1-5 |
| Self efficacy | 3 | 12 | 6-8 |
| Intentions | 6 | 24 | 9-14 |
| Peer norms | 3 | 12 | 15-17 |
| Peer approval | 5 | 20 | 18-22 |
| Family approval | 3 | 12 | 23-25 |
| Attitudes towards restrictions on behaviours | 3 | 12 | 26-28 |
| Parental provision and rules | 4 | 16 | 29-32 |
| Goal setting | 6 | 24 | 33-38 |
| Self monitoring | 4 | 16 | 39-42 |
| Discussion with parents | 3 | 12 | 43-45 |
| Suggestions to parents | 2 | 8 | 46-47 |
| Child shopping | 1 | 4 | 48 |
| Child cooking | 1 | 4 | 49 |
| Child trying new healthy snacks | 1 | 4 | 50 |

# Appendix 4 School characteristics and policy questionnaire

School Characteristics and Policy Questionnaire (SCPQ)

School: _____

Time point: _____

| | |
|---|---|
| **Playing Field** | |
| **Gym or Hall** | |
| **Healthy school standard/HCQM** | |
| **PEDPASS or Active Lifestyle Award** | |
| **Daily Take 10 or equivalent** | |
| **Breakfast club** | |
| **School travel plan** | |
| **Provision of space for wet play** | |
| **Play equipment freely available** | |
| **Number of after school Physical Activity clubs** | |

| Policies and Initiatives | |
| --- | --- |
| **School Physical Activity**<br><br>**policies/initiatives** | |
| **School Nutrition**<br><br>**policies/initiatives** | |

# Appendix 5 Analysis plan amendment document

**Amendments to the HeLP Analysis Plan V5 (3/11/15) September 14th 2016**

| Section 1: Summary of Amendments Original Planned Analysis | Section Number (page) in Analysis Plan V5 (date) | Summary of Amendment |
|---|---|---|
| **Secondary analysis of primary outcome - multiple imputation** | **7.1.3 (pages 17-18)** | **Multiple imputation of missing BMI SDS data at 24 months will not be undertaken. Alternative "Best Case / Worst Case" sensitivity analyses will be performed:** <br> Best Case: will assume no change in BMI SDS for children allocated to intervention group; change in BMI SDS for children allocated to the control group will be imputed from (marginal) mean change in BMI SDS for the control group <br> Worst Case: will assume that children who were not obese at baseline and who were allocated to the intervention group, were obese at 24 months, with BMI SDS imputed from Public Health England thresholds; change in BMI SDS for children allocated to the control group will be imputed from (marginal) mean change in BMI SDS for the control group |
| **Secondary analysis of primary outcome – CACE analysis** | **7.1.3 (pages 18-19)** | **CACE analysis will not be undertaken** |
| **Mediational analysis** | **9.1.3 (page 24)** | Both exploratory and confirmatory factor analyses will be carried out on the MLQ. In the Analysis Plan only 'confirmatory factor analyses' was specified. |

Section 2: Rationale and Further Details of Amendments

 *Secondary analyses of the primary outcome*

(a) Multiple imputation modelling

It was stated in analysis plan V5 that multiple imputation would be used to impute any missing BMI scores at 24 months (primary outcome) based on the assumption of data being missing at random, see section 6.2.1. Given the low proportion of missing BMI data at 24 months (5.6% (74/1324) of recruited children; 5.2% (68/1312) of children with baseline BMI SDS), compounded by the low amount of data deemed missing at random, the appropriateness of this planned sensitivity analysis was discussed at the HeLP Trial Steering Committee held on 20th July 2016. Following discussions, it was agreed that multiple imputation modelling was not required, however, an alternative sensitivity analyses approach to examine the effect of missing primary outcome data, was agreed, namely a "Best Case / Worst Case" scenario analyses. The first set of these analyses are based on hypothetically driven assumptions. Given the hypothetical preventative nature of the HeLP intervention, the *best case scenario* will:

assume no change between baseline and 24 months in BMI SDS for children allocated to the intervention group i.e. the baseline BMI SDS value will be carried forward to replace the missing 24 month BMI SDS value

impute missing 24 month BMI SDS values for children allocated to the control group with their corresponding baseline BMI SDS value plus the (marginal) mean change between baseline and 24 months for the children allocated to the control group with complete baseline and 24 month BMI SDS data.

The *worst case scenario* will:

assume that children allocated to the intervention group who were not obese at baseline were obese at the 24 month follow-up: the 24 month BMI SDS value will be set at the Public Health England threshold for obesity (i.e. the 95th percentile; this is 1.645). For children allocated to the intervention group who were obese at baseline, the baseline BMI SDS value will be carried forward to replace the missing BMI SDS value.

impute missing 24 month BMI SDS values for children allocated to the control group with their corresponding baseline BMI SDS value plus the (marginal) mean change between baseline and 24 months for the children allocated to the control group with complete baseline and 24 month BMI SDS data.

After imputing the missing 24 month BMI sds scores for both scenarios, the primary analyses model will then be fitted to the full intention-to-treat data set, to ascertain if the missing primary outcome data significantly influenced the results of the primary effectiveness analysis.

(b) Complier Average Causal Effect of Treatment (CACE) analysis

In analysis plan V5, it was stated that a CACE analysis would be undertaken to "estimate the complier average causal effect of treatment (CACE), as a potentially unbiased estimate of receiving HeLP." Non-compliers were defined as children who did not receive at least 4 sessions of drama and the 1-1 goal setting session. Given the very proportion of children categorised as non-compliers (there were only 24 out of 675 children allocated to the intervention group who were categorised as non-compliers (3.6%) and four of these children did not provide primary outcome data), the appropriateness of this planned sensitivity analysis was discussed at the HeLP Trial Steering Committee held on 20th July 2016. Following discussions, it was agreed that a CACE analysis will not be undertaken.

### 9.1.3 Mediational analysis

In analysis plan V5, it was stated that a "**confirmatory** factor analysis" would initially be undertaken on the MLQ. This was an error in the analysis plan, which should have read "factor analysis". We designed the MLQ at two theoretical levels; the first overarching theoretical framework was the IMB model of behaviour change (Information – Motivation – Behavioural Skills) and the second was a more detailed level comprising between six and 14 potential constructs, three of which were single items. We therefore did not have a firm idea of the exact number of constructs in the new measure and there are no published findings from other factor analyses to draw upon. Hence it is appropriate to use exploratory factor analysis in the first half of the data set and follow this with a confirmatory factor analysis in the second half.

# Appendix 6 Group by time repeated measures model

This is a repeated measures model used to assess whether or not there was evidence that any effect of the intervention differed over time.

Body mass index standard deviation scores for each child, $i$, nested within each school, $s$, at visit, $t$, were regressed on school-level factors, *Time point* and *Group*, and their interaction, adjusting for $k$ school-level factors $X$, modelling each school, and each child nested within school, as random intercepts, $\gamma_s$ and $\gamma_{i(s)}$, respectively:

$$BMIsd_{i(s),t} \sim \beta_{0,s} + \sum_{k,s}\beta_{k,s}X_{k,s} + \beta_{Group}Group_s + \beta_t Visit_s + \beta_{s,t}(Group_s.Visit_s) + \beta_{i,Gender}Gender_i + \gamma_s \tag{1}$$

$$+ \gamma_{i(s)} + \varepsilon(pattern, Visit_i),$$

where child-specific residual error, $\varepsilon \sim N(0, f(pattern, visit))$, was a matrix function of *Visit* pertaining to a particular covariance pattern, and $\beta$ and $\gamma$ denote fixed and random effects, respectively.

# Appendix 7 Training costs for staff involved in the delivery of HeLP

## Estimate training costs: HeLP co-ordinators

Based on within-trial experience on training requirements, it is assumed that training requirements for HeLP co-ordinators consist of 5 full days (10 sessions) of training, with two sessions focusing on the organisational set-up required for the intervention and the remaining sessions focused on specific skills related to the delivery of the intervention. It is expected/assumed that the training of four HeLP co-ordinators will be delivered by one experienced trainer.

Costs associated with training are estimated assuming that costs are distributed over a 12-month period (for base-case cost estimates; over 27 classes). No venue costs were included in the estimated costings, and an allowance of £500 for other costs (e.g. materials, subsistence) has been included. Using these assumptions, the estimated cost for the delivery of the training component for the HeLP co-ordinators is £191 per class or approximately £8 per participant (*Table 50*).

## Estimate training costs: drama team

Based on within-trial experience on training requirements, it is assumed that training requirements for drama delivery staff will consist of 5 full days (37.5 hours for drama facilitators, 30 hours for actors). The delivery of drama is carried out by a drama team comprising one drama facilitator and four actors. It is anticipated that the drama training, for two drama teams, will be delivered by one trainer.

Costs associated with training are estimated assuming that costs are distributed over a 12-month period (for base-case cost estimates; over 27 classes). No venue costs were included in the costing, and an allowance of £500 for other costs (e.g. materials, subsistence) has been included. Using these assumptions, the estimated cost for the delivery of the training component for the drama teams is £290 per class or £11.60 per participant (*Table 51*).

**TABLE 50** Training resources/costs: HeLP co-ordinators

| Staff type | Staff, (*n*) | Hours | Unit cost (£) | Total cost (£) | Notes |
|---|---|---|---|---|---|
| Lead trainer | 1 | 37.5 | 41.29 | 1548.51 | 1 × lead trainer × 37.5 hours |
| HeLP co-ordinator | 4 | 37.5 | 41.29 | 3097.03 | 4 × HeLP co-ordinators × 37.5 hours[a] |
| Venue | NA | NA | NA | NA | Costs not included |
| Other course costs[b] | NA | NA | NA | 500.00 | |
| Total cost | | | | 5145.54 | |
| *Cost per HeLP co-ordinator* | | | | *2572.77* | |
| *Cost per class (participant)* | | | | *190.58 (7.61)* | Based on base-case cohort of 27 classes, *n* = 25 per class |

NA, not applicable.
a We train four HeLP co-ordinators in order to provide two co-ordinators (core) for delivery of the intervention across the 27 classes, with two additional people trained and available to cover the co-ordinator role when needed (sickness/absence cover).
b Allowance for course materials and subsistence (e.g. refreshments).

**TABLE 51** Training resources/costs: drama teams

| Staff type | Staff, (*n*) | Hours | Unit cost (£) | Total cost (£) | Notes |
|---|---|---|---|---|---|
| Lead trainer | 1 | 37.5 | 41.29 | 1548.51 | 1 × lead trainer × 37.5 hours |
| Drama facilitator/HeLP co-ordinator | 2 | 37.5 | 41.29 | 3097.03 | 2 × drama facilitators × 37.5 hours |
| Actors | 8 | 30 | 11.20 | 2688.10 | 8 × actors × 30 hours |
| Venue | NA | NA | NA | NA | Costs not included |
| Other course costs[a] | NA | NA | NA | 500.00 | |
| Total cost | | | | 7833.64 | |
| *Cost per drama team*[b] | – | – | – | *3916.82* | |
| *Cost per class (participant)* | – | – | – | *290.13 (11.59)* | Based on base-case cohort of 27 classes, *n* = 25 per class |

NA, not applicable.
a  Allowance for course materials and subsistence (e.g. refreshments).
b  Drama team = 1 × drama facilitator + 4 × actors.

# Appendix 8  Review of cost-effectiveness analysis models: literature search strategy

## Sample MEDLINE search strategy (searches translated for other databases)

Date of search: August 2015.

Date range searched: 1990–current.

### Search strategy
exp obesity/ or exp obesity, abdominal/ or exp obesity, morbid/

exp overweight/

obes$.ti,ab.

body weight changes/ or weight gain/ or weight loss/

exp body fat distribution/ or exp body mass index/ or exp body size/ or exp skinfold thickness/ or exp waist-hip ratio/

((bmi or body mass index) adj2 (gain* or loss* or chang*)).ti,ab.

child/

child$.ti,ab.

teen$.ti,ab.

boys.ti,ab.

girls.ti,ab.

youth$.ti,ab.

adolescent/

child nutrition disorders/ or overnutrition/

cost$.ti.

((cost$ or economic) adj2 (effective$ or utility$ or benefit$ or minimis$ or analysis or evaluation)).ab.

or/1-6

or/7-14

15 or 16

17 and 18 and 19

Limit 20 to yr=’1990 -Current)

# Appendix 9  Summary of model input parameters for base-case Exeter Obesity Model

| Model parameter | Input/value | Source |
|---|---|---|
| Cohort size, $n$ | 1000 | Analysis input |
| Gender, % male | 50 | Assumption |
| Distribution by weight status: child (BMI SDS centile categories) | – | HeLP RCT, controls |
| Probability of adult weight status, by child weight status | – | Power et al., 1997[120] (see Table 28) |
| Follow-up duration, adult years | 30 | Analysis input |
| Adult start age (years) | 33 | Power et al., 1997[120] |
| Incidence of weight-related events per 1000 person-years: by weight status – healthy weight/overweight/obese | | |
|     CHD | 0.0024/0.0033/0.00048 | Guh et al., 2009[88] |
|     T2DM | 0.0017/0.0046/0.0133 | Guh et al., 2009[88] |
|     Stroke | 0.0014/0.0019/0.0023 | Guh et al., 2009[88] |
|     CRC | 0.0009/0.0013/0.0016 | Guh et al., 2009[88] |
| Mortality data, annual mortality rate/probability: | | |
|     All-cause general population mortality, by age (weighted for gender) | – | Office for National Statistics Life Tables 2012–14[89] |
|     CHD | 0.17 | Sigurdsson et al., 1995[90] |
|     T2DM | 0.06 | Roper et al., 2002[91] |
|     Stroke: year 1 | 0.37 | Hankey et al., 2000[92] |
|     Stroke: subsequent years | 0.10 | Hankey et al., 2000[92] |
|     CRC: year 1 | 0.24 | Bowel Cancer UK Mortality, 2012[93] |
|     CRC: subsequent years | 0.08 | Bowel Cancer UK Mortality, 2012[93] |
| Health state values/utilities | | |
|     General population | 0.86 | Kind et al., 1999[94] |
|     Coronary heart disease | 0.67 | Matza et al., 2015[96] |
|     T2DM | 0.78 | Clarke et al., 2002[95] |
|     Stroke | 0.52 | Matza et al., 2015[96] |
|     CRC | 0.76 | Djalalov et al., 2014[97] |
| Cost inputs (£, 2015) | | |
|     Intervention (HeLP) cost | 214.13 | HeLP RCT/analysis |
|     CHD, cost per year | 3403 | NHS Scotland, 2007[113] |
|     T2DM, cost per year | 4004 | Kanavos et al., 2012[112] |
|     Stroke, one-off treatment cost | 3126 | Saka et al., 2009[114] |
|     CRC, cost per year | 8861 | Cancer Research UK, 2014[115] |

| Model parameter | Input/value | Source |
|---|---|---|
| Discount rate: future costs, % | 3.5 | NICE, 2013[77] |
| Discount rate: future costs, % | 3.5 | NICE, 2013[77] |
| Half-cycle correction | 0.5 | Assumption |

**Assumptions**
Adult weights status: weight status category constant over time.
Events: first event is an absorbing state (for events), and participants remain in that state, with probability of death each cycle, but not subject to further risks for other events.

# Appendix 10 Parent questionnaire

### The Healthy Lifestyles Programme
### PARENTS/CARERS QUESTIONNAIRE

The Healthy Lifestyles Programme (HeLP) would like to hear about your personal experience of the Programme as a parent/carer and family. It is crucial in helping us to understand how HeLP affected you and your family. We would be grateful, therefore, if you can spare the time to complete the simple questionnaire below. Please return it in the envelope provided and it will be entered in the prize draw to win a £50 high street voucher!

**1. How did you become aware of your child's involvement in the Healthy Lifestyles Programme?**

**2. Did you and your child talk about the Healthy Lifestyles Programme at all?**

☐　A lot

☐　A little

☐　Not at all

**3. What do you think the Programme was trying to do?**

**4. Part of the Programme was to encourage the children to make changes to some of their lifestyle behaviours and state what they were going to do differently. We called these their HeLP goals.**

\* Were you aware of these goals?　　☐ Yes　　☐ No

\* How easy do you think it was for your child to make these changes?

　　☐ Easy　　☐ Quite Easy　　☐ Quite Difficult　　☐ Difficult

**5. Did you notice any change in your child's:**

*Choice of snacks                                      ☐ Yes  ☐ No

*Choice of drinks                                      ☐ Yes  ☐ No

*Time spent doing screen based activities (e.g. computer, iPod,
phone,   iPad, TV )                                    ☐ Yes  ☐ No

*Time spent doing physical activity?      ☐ Yes  ☐ No

*Other (please specify below)

**6. If you have said yes to any of the above, please could you write a little about the changes you noticed below:**

**7. If your child made some changes, have they managed to stick to any of them?**

☐        Yes

☐        No

☐        Unsure

**8. Have you or the rest of your family made any changes because of your child taking part in HeLP?**

☐        No changes

☐        A few

☐        Lots

**9. If any changes have been made (by you or the rest of the family) , please could you write what these are below:**

**10. If there is anything else you would like to add, please write it below:**

**11. If you are happy to talk and give more detail to you HeLP Coordinator we would love to hear your thoughts. Please provide your details below and we will be in touch shortly.**

*Name

*Telephone number

*Email

*Child's name

*Child's school and class

*Best time for us to call you

# Appendix 11 Year 5 teacher interview schedule

**The HeLP Programme**

**Year 5 Teacher Interview – Intervention Schools**

### INTRODUCTION

Thank you for agreeing to give a brief interview on your experience of your involvement in HeLP. Understanding your experience of HeLP is vital to the further development of the programme, so please feel free to be as frank as you wish! This interview will be relatively unstructured to allow you to just talk about your views and experiences of participating in the programme. So that I can concentrate on what you are saying is it ok for me to record this interview? All your comments will remain anonymous and will only be used for the purposes of this research project.

1. How did you feel when you heard that your Y5 class was going to receive the Healthy Lifestyles Programme?
   a. Was there anything that made you anxious?
   b. Was there anything that you were excited about?
   c. Is there anything we could have done to allay any fears in the first instance?

2. Do you think it is necessary to have programmes like HeLP given everything else going on in schools?
   a. Probe added value – what makes HeLP distinct

3. Having been through the whole year 5 programme could you give us some indication of the workload/hassle factor?
   a. Has it added to your workload and, if so, in what way?
   b. What could we do differently to alleviate your workload?
   c. Probe how the PSHE lessons were received
   d. How much do the HeLP activities overlap with the year 5 curriculum? (try and get teachers to give a percentage overlap)

4. One of the key roles of the HeLP Coordinator is to build relationships with schools, children and families and especially to support you.
   a. Did you feel supported?
   b. Was there more we could have done?

5. How do you think the children found the programme?
   a. How many drama sessions did you observe and what did you think?
   b. Did you notice any impact at a class/individual child level?

        i. Probe possible reasons for teacher observations re whole class or individual children (need to tease out what is it about HeLP that led to certain behaviours – programme differentiation)

6. Do you think HeLP is known/understood in the wider school context?
   a. Probe examples

7. Do you have any sense of how parents perceived the programme?

8. Have you used any aspect of HeLP in your day to day teaching and/or interactions with the children?

9. Had HeLP had any effect on you personally?

10. Would you recommend HeLP to your colleagues or other schools?

11. Is there anything else you would like to say/comment on?


Thank you so much for sharing your views about the Healthy Lifestyles Programme today.

# Appendix 12 Parent interview schedule

Parent Interview Schedule (intervention schools only)

**Introduction**

Thank you so much for agreeing to participate in this interview in order to expand on your questionnaire responses, we really appreciate your time. The aim of the interview is to understand yours and your family's experience of HeLP and any impact if may have had so that we can learn for future development. We are happy to hear all types of feedback, so please feel free to express anything that you think we could have done differently to improve the programme.

The interview will take around 15 minutes and will be audio taped to ensure that we do not miss anything. All information you provide will be anonymised. Before we begin please feel free to ask any further questions.

Once I switch the recorder on I will say your name and the date and then we can begin.

**Questions**

1. What were your overall views of the programme? (probe whether they attended any parental engagement activities)

2. You mentioned that XXX talked about the Healthy Lifestyles programme (a lot; a little); please could you expand on the sort of conversations you had.

3. We asked in the questionnaire about the HeLP goals, which were set around the three key messages of reducing fizzy drink consumption, replacing unhealthy snacks with healthy alternatives and reducing screen time. You thought that it was easy/quite easy/quite difficult/difficult for XX to make these changes. Would you be able to expand on the reasons for this?

4. You indicated that you noticed changes in XX choice of snacks/drinks/screen time/physical activity. Please could you tell me a little more about these changes?

5. Is XX managing to maintain these changes at all? (if so probe how; if not probe why this might be).

6. You mention that the family have made some changes as a result of HeLP, please could you expand on these.

7. Are the family/you managing to maintain these changes?

8. Is there anything else you would like to say about HeLP in particular or healthy lifestyles in general? Probe if there is anything we could have done differently to help us refine the programme for roll out.

# Appendix 13 Focus group schedule

## Focus groups with Year 6: programme schools

**FOCUS GROUPS WITH YEAR 6 – Programme Schools**

Last year you were involved in our Healthy Lifestyles Programme, we are here today to find out what you remember about the Programme and how you felt about it, in research we call this a 'focus group' which means to have a chat or discussion about your opinion and feelings about the programme. Each of you is here because we randomly chose names out of a hat. It is really important to get some honest feedback, good and bad, from you as this will help us make the Programme better in the future.

There are no right or wrong answers and we will make sure that each of you has the opportunity to speak- it's very important we hear everyone's views. So that we don't talk over each other we are going to use this carrot, you can talk only when you are holding the carrot when you want to speak put your hand up and we will pass the carrot to you. To help us we are also going to use the ground rules we used during Healthy Lifestyles week, do you remember what they were?

No personal comments

No put downs

Respect and listen to each other

Ok so now we are going to transport you back in time, like the actors did during Healthy Lifestyles week with their time machine, we have some photos here to help remind you of the things we did during the programme and we would like you to look at these and spend some time remembering the activities, thinking about the character you worked with and the goals you set. We will ask questions about your home and your parents and what they thought of the programme too. *Refer to poster-* these are the things we want to focus on today. The photos coming round are of the activity workshops (rugby/basketball/dance) and the HLW.

*Pass the photos around the table*

*Have a copy of goals sheet visible*

*Give the children a Post-it note each and a pen*

Before we get started we are going to ask you to imagine you are describing the Healthy Lifestyles Programme to some Year 5 children who are just about to start but don't know what it's all about.  Can you come up with 3 words that you would use to describe the whole experience to other children?

There is a voice recorder in the room so we can write up notes about the things you say, please don't worry about it being on. We will go round the room and each say our name, so let's get started…..

What did you enjoy about the HeLP Programme? (one thing from each child) Probe reasons for why they enjoyed these things

Did you talk about HeLP at home? *Probe what was said and with whom and why they wanted to talk about these things*

What do you think your parents/family think of HeLP? *Probe events they*

*came to and if anything was said/discussed as a result*

Did you talk about HeLP at school e.g. in the playground or in other lessons? *Probe what was said and with whom and why they discussed these things*

What did you think about the drama and the other activities during the Healthy Lifestyles Week?

How was it different to what you usually do in school? *Probe why and the impact this had on them*

How easy was it to pick a character? *Draw out why they identified with their character*

Did working with your chosen character help you to set goals? *Probe reasons why*

How did you feel when you had a different actor playing your character in the year 6 activities? *Probe reasons for these feelings*

How did you find setting goals at home with your parents? *Probe reasons for their experiences. Probe if anything was discussed/said at this time*

What was it like trying to achieve your goals? (at home and at school). *Probe reasons for why it was difficult/easy*

Could you tell us anything you or your friends and family did to help you achieve your goals? (at home and at school). *Probe what they were and the impact it had*

Have you noticed any changes in yourself from being involved in the Programme? *Probe what they were. Make it clear to the children this can include their thoughts and attitudes as well as their behaviours*

Have you noticed any changes at home since the Programme? *Probe what they were. This can also include discussions and how they relate to their family as well*

Was there anything about the Programme you didn't enjoy so much? *Probe reasons why*

Is there anything else you would like to say?

Thank you so much for sharing your views and feelings today. If there is anything else you would like to say to me or share about being involved in HeLP then please stay behind after this session.

# Appendix 14 Fidelity to form example checklist

**PHASE 1**

**Whole school Assembly Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to component content

| Activity – class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Introduction to HeLP | | |
| The energy balance – explanations and questions | | |
| The energy balance mime (boy and girl) | | |
| The HeLP rap | | |

General Comments:

**PHASE 1**

**Activity Workshop 1 Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:                              ACTIVITY:

Adherence to component content

| Activity – class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Intro and Q/A re the energy balance and HeLP messages | | |
| Sportsmen/dancers talking about the importance of eating healthily and being active | | |
| Practical session | | |

General Comments:

**PHASE 1**

**Activity Workshop 2 Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:                                    ACTIVITY:

Adherence to component content

| Activity – class teacher present | Yes/Partial/No | Notes-plus TA |
|---|---|---|
| Intro and Q/A re the energy balance and HeLP messages | | |
| Sportsmen/dancers talking about the importance of eating healthily and being active | | |
| Practical session | | |

General Comments:

**PHASE 1**

**Parents' assembly Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to component content

| Activity – class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Intro with rationale of the Programme | | |
| Summary of the HeLP Programme | | |
| Family invitations | | |
| Contacts details given | | |
| Dance performance | | |
| Sport performance (e.g.rugby/basketball) | | |
| Raps performance | | |

General comments:

**PHASE 2**

**Drama Workshop 1 Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to key content

| Session 1 – Class teacher present | Yes/partial/No | Notes |
|---|---|---|
| Introduce themes (80/20, energy balance) | | |
| Nail and tooth in coke | | |
| Introduce characters (alarm clock scene)/children choose character | | |
| Children working in small groups with their character | | |
| Introduce traffic light food system and quiz | | |
| Food machines | | |
| Key message chant | | |
| Distribution of Traffic Light Food System parent information sheet | | |

Quality of delivery and participant responsiveness (1=Low Quality        10=High Quality)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Enthusiastic delivery | | | | | | | | | | |
| Engaged/participating children | | | | | | | | | | |

General comments:

**PHASE 2**

**Drama Workshop 2 Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to key content

| Session 2 - Class teacher present | Yes/Partial/No | Notes Mark away on course so had a sub |
|---|---|---|
| 24 hour clock and discussion | | |
| 80/20 activities game | | |
| Lunchtime scene | | |
| Tasting healthy snacks and drinks | | |
| Rerun key messages and chant created in session | | |
| Distribution of breakfast parent information sheet | | |
| Encouragement from children to Duncan to pass this on to his parents | | |

Quality of delivery and participant responsiveness (1=Low Quality        10=High Quality)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Enthusiastic delivery | | | | | | | | | | |
| Engaged/participating children | | | | | | | | | | |

General comments:

**PHASE 2**

**Drama Workshop 3 Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to key content

| Session 3 - Class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Professor Grossen-Luncheon scene and discussion | | |
| Reading food labels and group discussion with character | | |
| Temptation T shirts | | |
| Temptation ladder | | |
| Aunt Betsy scene | | |
| Rerun key messages and chant created in session | | |
| Distribution of Snacks parent information sheet | | |
| Encouragement from children to Duncan to pass this on to his parents | | |

Quality of delivery and participant responsiveness (1=Low Quality       10=High Quality)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Enthusiastic delivery | | | | | | | | | | |
| Engaged/participating children | | | | | | | | | | |

**PHASE 2**

**Drama Workshop 4 Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to key content

| Session 4 - Class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Alarm clock scene | | |
| Small group practising 3 message chant | | |
| Snacky Sam's tea time scene and discussion | | |
| Duncan's tea time scene and Forum Theatre of the scene | | |
| Distribution of Lunch Box parent information sheet | | |
| Encouragement from children to Duncan to pass this on to his parents | | |

Quality of delivery and participant responsiveness (1=Low Quality       10=High Quality)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Enthusiastic delivery | | | | | | | | | | |
| Engaged/participating children | | | | | | | | | | |

General comments:

**PHASE 2**

**Drama Workshop 5 Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to key content

| Session 5 - Class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Adverts | | |
| Marketing strategy | | |
| Making adverts in small groups | | |
| Making smoothies | | |
| Role play how to overcome teasing | | |
| Tooth out of coke and discussion | | |
| Making a scene! (groups recap the week in a scene) | | |
| School disco scene | | |
| Distribution of Parent Power parent information sheet | | |
| Encouragement from children to Duncan to pass this on to his parents | | |

Quality of delivery and participant responsiveness (1=Low Quality          10=High Quality)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Enthusiastic delivery | | | | | | | | | | |
| Engaged/participating children | | | | | | | | | | |

General comments:

**PHASE 3**

**Forum Theatre Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to component content

| Activity – class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Intro – set the scene | | |
| The 80/20 rap | | |
| Food machines | | |
| Supermarket scenes | | |
| Forum theatre | | |

General comments

**PHASE 4**

**Drama Workshop Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to component content

| Activity – class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| 80:20 Rap with actions | | |
| How AA and DD have changed | | |
| Food machines | | |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Enthusiastic delivery | | | | | | | | | | |
| Engaged/participating children | | | | | | | | | | |

**PHASE 4**

**Yr 6 Class-Delivered Assembly Fidelity Check**

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to component content (to be completed with JL or KW if present)

| Activity – class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Whole *class* rap | | |
| Overview of HeLP | | |
| Before/after DD and AA showing change in behaviours. | | |
| Food Machines | | |
| Whole *school* rap | | |
| Presence of a senior member of staff (e.g. Head, Deputy Head) | | |
| Presence of year 6 teacher(s) | | |

Quality of delivery and participant responsiveness (1=Low Quality        10=High Quality)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Well organised assembly | | | | | | | | | | |
| Engaged children (Y6) | | | | | | | | | | |
| Engaged audience | | | | | | | | | | |

General comments:

# Appendix 15 Fidelity to function example checklist

**PHASE 1**

**Parent Assembly Fidelity Check**

NAME OF SCHOOL:

NAME OF OBSERVER:

DATE:

Quality of delivery and participant responsiveness (1=Low Quality            10=High Quality)

|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Enthusiastic delivery |  |  |  |  |  |  |  |  |  |  |
| Engaged/participating children |  |  |  |  |  |  |  |  |  |  |
| Engaged parents |  |  |  |  |  |  |  |  |  |  |
| Engaged school staff |  |  |  |  |  |  |  |  |  |  |

|  | Yes/No |
|---|---|
| Presence of a senior member of staff (e.g. Head, Deputy Head) |  |
| Year 5 teacher in charge |  |
| Presence of all year 5 teachers |  |
| Pupils raps |  |
| Pupil dance |  |
| Pupil sports skills |  |

General comments

Teacher's view of the evening:

Parent conversation notes

**PHASE 3**

**Forum Theatre Evening Fidelity Check**

NAME OF SCHOOL:

NAME OF OBSERVER:

DATE:

Quality of delivery and participant responsiveness (1=Low Quality          10=High Quality)

|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Enthusiastic delivery |  |  |  |  |  |  |  |  |  |  |
| Engaged/participating children |  |  |  |  |  |  |  |  |  |  |
| Engaged parents |  |  |  |  |  |  |  |  |  |  |
| Engaged staff |  |  |  |  |  |  |  |  |  |  |

|  | Yes/No |
|---|---|
| Presence of a senior member of staff (e.g. Head, Deputy Head) |  |
| Presence of all year 5 teachers |  |

General comments

Teacher's view of the evening:

Parent conversation notes

**PHASE 4**

### Yr 6 Class-Delivered Assembly Fidelity Check

NAME OF SCHOOL:

NAME OF HC:

DATE:

Adherence to component content ( to be completed with JL or KW if present)

| Activity – class teacher present | Yes/Partial/No | Notes |
|---|---|---|
| Whole *class* rap | | |
| Overview of HeLP | | |
| Before/after DD and AA showing change in behaviours. | | |
| Food Machines | | |
| Whole *school* rap | | |
| Presence of a senior member of staff (e.g. Head, Deputy Head) | | |
| Presence of year 6 teacher(s) | | |

Quality of delivery and participant responsiveness (1=Low Quality          10=High Quality)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Well organised assembly | | | | | | | | | | |
| Engaged children (Y6) | | | | | | | | | | |
| Engaged audience | | | | | | | | | | |

General comments:

# Appendix 16 Schools in the HeLP study

## Baseline characteristics

| School | Urban or rural? | Playing field? | Gym or hall? | Healthy school standard? | Space for wet play? | Play equipment available? | Number of after school clubs |
|---|---|---|---|---|---|---|---|
| 01 | Urban/rural | ✓ | – | – | – | – | 4 |
| 02 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 4 |
| 03 | Urban | ✗ | ✓ | ✓ | ✗ | ✓ | 4 |
| 04 | Rural | ✓ | ✓ | ✓ | ✓ | ✓ | 5 |
| 05 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 2 |
| 06 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 2 |
| 07 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 4 |
| 08 | Urban/rural | ✓ | ✓ | ✓ | ✓ | ✓ | 3 |
| 09 | Urban | ✓ | ✓ | ✗ | ✓ | ✓ | 3 |
| 10 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 19 |
| 11 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 15 |
| 12 | Urban/rural | ✓ | ✓ | – | ✓ | ✓ | 9 |
| 13 | Urban | ✓ | ✓ | ✗ | ✓ | ✓ | 4 |
| 14 | Urban | ✓ | ✓ | ✗ | ✓ | ✓ | 6 |
| 15 | Rural | ✓ | ✓ | ✓ | ✓ | ✓ | 11 |
| 16 | Rural | ✓ | ✓ | ✓ | ✓ | ✓ | 7 |
| 17 | Rural | ✓ | ✓ | ✓ | ✓ | ✓ | 6 |
| 18 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 5 |
| 19 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 1 |
| 20 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 6 |
| 21 | Urban | ✓ | ✓ | ✗ | ✓ | ✓ | 4 |
| 22 | Urban | ✗ | ✓ | ✓ | ✓ | ✓ | 9 |
| 23 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 6 |
| 24 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 10 |
| 25 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 3 |
| 26 | Urban | ✗ | ✓ | ✓ | ✓ | ✓ | 5 |
| 27 | Urban | ✗ | ✓ | – | ✓ | ✓ | 9 |
| 28 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 5 |
| 29 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 2 |
| 30 | Urban | ✓ | ✓ | ✓ | ✓ | ✓ | 10 |
| 31 | Urban/rural | ✓ | ✓ | ✓ | ✓ | ✓ | 6 |
| 32 | Rural | ✓ | ✓ | ✗ | ✓ | ✓ | 4 |

–, missing data.
Blue shading denotes control schools and green shading denotes cohort 2 schools.

# Physical activity and nutrition policies

| | PA policies | | | | | | | | | | | | Nutrition policies | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Fitness sessions/ aerobics | | Specialised coaches (including outside agency)/PE specialist (including sports/PE co-ordinator) | | Take '10'/ PEDPASS | | School travel/ bikeability | | Sports ambassador/ play teacher | | Gifted and talented sports programme | | Breakfast club/cookery club/classes | | Healthy snacks at break | | Food For Life | | School fruit and vegetable scheme | | Healthy lunchboxes | | Family eating set up at lunch | | |
| School | B | 18 | B | 18 | B | 18 | B | 18 | B | 18 | B | 18 | B | 18 | B | 18 | B | 18 | B | 18 | B | 18 | B | 18 | School awards |
| 01 | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | | | ✓ | ✓ | | | | | | | | | |
| 02 | | | | | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | |
| 03 | | | | | ✓ | ✓ | | | | | | | | | | | | | | | | | | | Health School plus |
| 04 | | | | | | | ✓ | ✓ | | | | | ✓ | ✓ | | | | | ✓ | ✓ | | | | | School Nutritional Standards award |
| 05 | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | ✓ | ✓ | | | | | | | Active Lifestyle Award. Healthy school award |
| 06 | | | | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | |
| 07 | | | ✓ | ✓✓ | | | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | Healthy School award. HCQM. Eco school award |
| 08 | | | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓✓ | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | ✓ | ✓ | Eco school Green Flag |
| 09 | | | | | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | |
| 10 | | | | | ✓ | ✓ | | | | | | | ✓ | ✓ | | | | | | | | | | | HCQM |
| 11 | | | | | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | ✓ | ✓ | | | | | | | | | HCQM Bronze |
| 12 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | | | ✓ | ✓ | | | | | | | Healthy school Status. Eco Flag Award |
| 13 | | | | | | | ✓ | ✓ | | | | | | | | | | | | | | | ✓ | ✓ | |
| 14 | | | | | | | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | ✓ | ✓ | | | |
| 15 | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | | | | | | | | | | | | | |
| 16 | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | ✓ | ✓ | | | | | | | | | |

| School | Fitness sessions/aerobics B | 18 | Specialised coaches (including outside agency)/PE specialist (including sports/PE co-ordinator) B | 18 | Take '10'/PEDPASS B | 18 | School travel/bikeability B | 18 | Sports ambassador/play teacher B | 18 | Gifted and talented sports programme B | 18 | Breakfast club/cookery club/classes B | 18 | Healthy snacks at break B | 18 | Food For Life B | 18 | School fruit and vegetable scheme B | 18 | Healthy lunchboxes B | 18 | Family eating set up at lunch B | 18 | School awards |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 17 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | ✓ | ✓ | | | | | | | | | Healthy Eating Awards |
| 18 | | | | ✓(red) | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | |
| 19 | | | | | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | | | | | | | | | | | Healthy Schools Plus |
| 20 | | | | | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | |
| 21 | | | | | | | | ✓(red) | | | | | ✓ | ✓ | | | | | | | | | | | Healthy School Award. Starting HCQM Bronze |
| 22 | ✓ | ✓ | | | ✓ | ✓ | | | | ✓(red) | | | ✓ | ✓ | ✓ | ✓ | | | ✓ | | | | | | Healthy Schools plus starting HCQM Bronze |
| 23 | | | | ✓(red) | ✓ | ✓ | ✓ | ✓ | | | | | | | ✓ | ✓ | ✓ | ✓ | | | | | | | |
| 24 | ✓ | ✓ | | ✓(red) | ✓ | ✓ | | | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | Healthy School Award |
| 25 | | | | ✓(red) | | | ✓ | ✓ | | | | | ✓ | ✓ | ✓ | ✓ | | | | | | | | | |
| 26 | | | | | ✓ | ✓ | | | | | | | ✓ | ✓ | | | | | | | | | | | |
| 27 | | | ✓ | ✓ | ✓ | ✓ | | | | | | | ✓ | ✓ | ✓ | ✓ | | | | | | | | | |
| 28 | | | | | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | Healthy school award |
| 29 | | | | | | | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | |
| 30 | | | ✓ | ✓ | | | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | |
| 31 | | | | | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ | | | | | | | | | | | |
| 32 | | | | | | | ✓ | ✓ | | | | | | | | | | | | | ✓ | ✓ | | | |

–, missing data; HCQM, Healthy Child Quality Mark; PA, physical activity; PE, physical education.
Blue shading denotes control schools and green shading denotes cohort 2 schools. Red ticks indicate policy difference from baseline.

# Appendix 17 Parent questionnaire responses to the quantitative questions

| Question number | Question | Response choice | n (%) |
|---|---|---|---|
| 2 | Did you and your child talk about the Healthy Lifestyles Programme at all? | A lot | 102 (58.0) |
| | | A little | 72 (40.9) |
| | | Not at all | 1 (0.6) |
| | | Spoilt/missing | 1 (0.6) |
| 4a | Were you aware of your child's goals? | No | 1 (0.6) |
| | | Yes | 175 (99.4) |
| 4b | How easy do you think it was for your child to make these changes? | Easy | 31 (17.6) |
| | | Quite easy | 103 (58.5) |
| | | Quite difficult | 35 (19.9) |
| | | Very difficult | 2 (1.1) |
| | | Spoilt/missing | 5 (2.8) |
| 5a | Did you notice any change in your child's choice of snacks? | No | 36 (20.5) |
| | | Yes | 136 (77.3) |
| | | Spoilt/missing | 4 (2.3) |
| 5b | Did you notice any change in your child's choice of drinks? | No | 66 (37.5) |
| | | Yes | 106 (60.2) |
| | | Spoilt/missing | 4 (2.3) |
| 5c | Did you notice any change in your child's choice of time spent doing screen-based activities? | No | 76 (43.2) |
| | | Yes | 97 (55.1) |
| | | Spoilt/missing | 3 (1.7) |
| 5d | Did you notice any change in your child's time spent doing physical activity? | No | 57 (32.4) |
| | | Yes | 117 (66.5) |
| | | Spoilt/missing | 2 (1.1) |
| 7 | Have they managed to stick to any of the changes they made? | No | 9 (5.1) |
| | | Yes | 138 (78.4) |
| | | Unsure | 19 (10.8) |
| | | Spoilt/missing | 10 (5.7) |
| 8 | Have the rest of family made changes? | No changes | 46 (26.1) |
| | | A few | 119 (67.6) |
| | | Lots | 11 (6.3) |

# Appendix 18 Development and evaluation of the My Lifestyle Questionnaire

Development and evaluation of the MLQ for 9- to 10-year-olds: an instrument designed to assess health behaviour change in school children taking part in the HeLP RCT.

## Background

At the start of the HeLP trial there was no instrument available to assess the range of health behaviours targeted by the HeLP intervention. As part of the process evaluation of this complex intervention trial, we developed and put in place a plan to evaluate such an instrument, called the MLQ.

## Objectives

To develop and evaluate (1) whether or not the MLQ contains items that are relevant, acceptable and feasible for 9- to 10-year-olds to complete and (2) that the MLQ is psychometrically robust and (3) fit for purpose to use in a mediational analysis as part of the process evaluation of the HeLP trial.

## Methods

The MLQ underwent two development phases and then a third phase for psychometric evaluation.

### Participants

Phase 1 was undertaken by the research team; phase 2 participants were 9- to 10-year-old children with a range of educational abilities. Twelve children took part initially and then a further 102 children from two primary schools who were a convenience sample based on contacts established during the pilot HeLP trial. In phase 3, the participants were the 1324 children who took part in the HeLP trial.

### Phase 1: development and establishment initial construct validity

The initial development of the HeLP MLQ was an iterative process[144] and involved four steps:[145] (1) definition and elaboration of the construct intended to be measured; (2) choice of measurement method; (3) selection and formulation of items; and (4) determination of the scoring system. Experts in health psychology and behaviour change within the research team discussed how to meet each step taking into consideration the following requirements.

- Year 5 children (9- to 10-year-olds) would have a 1-hour lesson made available so it should take no longer than 40 minutes to complete (estimated no more than 60 items).
- Readability (language and sentence construction) needed to be appropriate for target age group (9- to 10-year-olds) to avoid assessment of the child's cognitive ability rather than the question construct.
- Child-friendly format to keep the target group interested and focused on the task at hand.
- Items to allow for assessment of constructs over time, enabling the evaluation of any effects attributable to HeLP.
- Knowledge items needed to be sufficiently challenging to assess the key knowledge exchanged in HeLP.
- Assessment of remaining constructs (individual motivations and cognitions, perceived environment and health behaviours) to be measured at an ordinal level using Likert scales (as is typical in other questionnaires for children).

Two further sources also informed the selection and formulation of items: the intervention mapping process used to create the HeLP intervention[34] and previously published related questionnaires.[146–149]

The research team selected 52 items for first draft of the MLQ. The IMB model,[150] which underpinned the mapping of the HeLP intervention, provided the top-level categorisation of these items, which were then divided into further sections according to more specific behaviour change principles, theories or techniques.

## Phase 2: initial pilot and field testing for face validity and content validity

The first MLQ draft was piloted on 12 children who completed the form at home, and their parents were asked to provide feedback on the child's understanding of the questions and the time taken to complete it. This feedback was incorporated before a revised version of the MLQ was field tested in two schools with 102 children aged 9–10 years. In school A the 57 children were asked to complete the questionnaire in their own time with researchers and teachers available for support and guidance if needed. In school B the 45 children completed the questionnaire as a whole class activity and any child with additional learning needs was given extra support outside the main classroom. The purpose of this fieldwork was to further test the applicability, comprehensibility, ease of completion, appropriateness of instructions, items and response format, overall presentation of the questionnaire and time taken to complete. Any items or formatting issues that were found to be problematic by either the children or the teachers were removed or amended. Problematic items were defined as those that were ambiguously worded, or difficult to understand or score. This phase ended with 50 items and these were then used in the HeLP trial.

## Phase 3: psychometric evaluation

The 50-item MLQ was part of the baseline and 12-month data collection of all children involved in the HeLP trial (intervention and control schools from both cohorts). The MLQ was administered during a classroom session assigned to the HeLP intervention and children were supported to complete all sections by their teacher and the HeLP co-ordinator. The MLQ data were collected during a lesson when the HeLP co-ordinators read the questions from the front of the class and the children completed the questionnaire at the same time in silence. Clarifications were given for specific questions. A poster was produced to demonstrate percentages so that the children were able to understand question 4 in the 'knowledge' section. Children with additional needs completed the questionnaire in a smaller group outside the classroom with an additional researcher.

Data were double entered and cleaned with any reverse scoring items recoded. Spoilt and missing values were coded as missing data.

The first MLQ section contained five 'knowledge' items (with 'knowledge' relating to the information component of the IMB model). Responses from each child were summed to give a 'knowledge' total score (ranging from 0 to 20).

The remaining MLQ sections contained 45 Likert scale items. To allow checking of the psychometric properties of the MLQ, the baseline data set and the 12-month data set were then each split into two further data sets (A and B) by taking alternate entries from the master database; this ensured that each sample contained similar numbers from both cohorts of the HeLP trial. Each Likert scale item was then screened for reading ease score, mean, SD, ceiling and floor effects and interitem correlations. We prespecified the criteria[145] and agreed the following for identifying problems.

- Flesch–Kincaid reading ease score of < 60 was considered problematic (https://readability-score.com/text/; last accessed 21 October 2016).
- Items that lack a range of scores/lack of variation around the mean (mean score to be > 1.5 or < 3.5; SD to be > 0.8 or < 1.8).
- Ceiling and floor effects when > 30% of participants have the maximum or minimum scores.
- Individual items that correlate ≥ 0.6 with another item.

It was agreed that any item failing on more than two of these properties would be removed from further analysis.

The remaining items from baseline data set A were then analysed using EFA to propose a plausible factor structure as there were no published data for determining the factor constructs arising from our new items, and the developmental mapping exercise created a number of different levels of item categorisation, for example the IMB model level, a variety of theoretically derived constructs as well as three single items. Although using EFA as the first approach to the factor analysis is justified, it should be noted that our analysis plan had to be amended to reflect this first step rather than CFA; this amendment was agreed with the TSC.

The factors which emerged from the EFA were further assessed using a CFA with baseline data set B. EFA explored whether or not this factor structure was still plausible for the 12-month data set, again followed by a CFA to provide fit indices of the proposed model fits. Additionally, a multigroup CFA with invariance testing was to be used to explore the stability of the proposed factor solutions (comparing data set B at baseline with data set B at 12 months).

An iterative approach was taken that was guided by a set of goodness-of-fit indices but gave priority to the interpretability of any proposed factor solution. This also minimised the chances that conclusions were driven by spurious results, which may yield slightly superior fit indices but provide little practical or theoretical benefits. Additionally, while correlating error variances typically leads to substantially improved model fits, such correlations were applied cautiously and with the intention to identify method effects or item clusters with shared content.

As the data were ordinal in nature, the most appropriate CFA approach was using an asymptotically distribution-free method using polychoric correlations. This technique is also recommended for Likert-scale items with fewer than five options and the sample size requirements for this approach were met.

### Software and type of factor analysis and goodness-of-fit indices used

Exploratory factor analysis was conducted using principal components analysis with promax factor rotations using the software package SPSS (Statistical Package for the Social Sciences), version 23. The Kaiser criterion of eigenvalue of > 1 was applied to identify number of factors for an initial factor solution. The pattern of factor loadings and cross-loadings was inspected (ignoring factor loadings of < 0.20), together with conceptual criteria and face validity of the scale, to derive plausible factor solutions.

Confirmatory factor analyses were conducted with the software package Lisrel version 8.80 (Scientific Software Internationl Inc., Skokie, IL, USA)[151] using robust diagonally-weighted least-squares and polychoric correlations and asymptotic covariance matrices.[152] As chi-squared is known to be inflated with sample size,[131] it was not used as an indicator of goodness of fit. Instead, a set of indicators was used to provide a fair assessment of goodness of fit. These indicators are known to be sensitive to different aspects of model misfits. The following criteria were used:

- RMSEA of < 0.06[132]
- CFI of > 0.95[132]
- SRMR of < 0.08[132]
- relative/normed chi-squared ($\chi^2$/df) < 5 (although this criterion is commonly used and well known, it is equally sensitive to sample size[133]).

## Results

### Phases 1 and 2

An initial 52-item MLQ booklet was created in phase 1 from team discussion, mapping work and published sources.[144]

Qualitative feedback from the fieldwork (phase 2) and summary statistical data (e.g. means, SDs, floor and ceiling effects and item correlations) were discussed by the development team, resulting in a number of amendments and formatting changes. The resultant second draft of MLQ comprised 50 items in five sections (*Box 1*): the first section contained five knowledge items with a score range of 0–20. The remaining sections were related to individual motivations and cognitions (23 items), parental mediating behaviours (four items), child use of change techniques (10 items) and child mediating behaviours (eight items); all of these items had a four-point Likert rating scale (see *Appendix 3* for a copy of the MLQ and its scoring system). Eleven items require reverse scoring. Higher scores represent health promoting motivation and behavioural skills.

### Phase 3

There were 1324 children involved in the trial, randomised to intervention and control groups within two cohorts. The MLQ was administered as part of a school lesson, and the children were supported to complete the questionnaire by their class teacher and the HeLP co-ordinator; this process resulted in high level of completeness. However, some children had spoilt entries (partially marked answers, etc.) or some missing entries, particularly for the last items in the questionnaire. All spoilt entries were recoded as missing data. Overall, the proportion of MLQ missing data was small (< 1%), and imputation of missing values was not undertaken at this stage of the evaluation work.

**BOX 1** Sections of the MLQ

**Knowledge**

- Healthy snack/drink alternatives.
- Food group proportions.
- Lifestyle physical activity.
- Energy balance.
- Strategies for change.

**Individual motivations and cognitions**

- Self-efficacy to make healthy eating and activity choices.
- Intentions to make healthy eating and activity choices.
- Peer norms.
- Peer approval.
- Family approval.
- Attitudes towards restrictions on behaviour.

**Parental mediating behaviours**

- Parental provision of healthy/unhealthy snacks/drinks and physical activity opportunities.
- Parental rules for food consumption and screen time.

**Child mediating behaviours**

- Goal-setting.
- Self-monitoring.
- Discussion about healthy lifestyles with parents/family.
- Encouraging the family to be more healthy.
- Shopping with family.
- Cooking with family.
- Trying new healthy foods/drinks.

### Knowledge scores

The total 'knowledge' score (see *Appendix 3* for scoring system) at baseline (*n* = 1310) ranged from 0 to 19, with a mean of 8.40 (SD 2.72). At 12 months (*n* = 1278) the total knowledge score range was 0–20, with a mean of 10.48 (SD 3.19).

### Data screening and preliminary analyses of the My Lifestyle Questionnaire Likert scale items

The remaining analysis, focused on the 45 Likert scale items, was randomly split in half (*n* = 662 each) and designated data set A and data set B for both the baseline and the 12-month data.

*Table 52* summarises the screening exercise performed on each item for each of the four data sets. Two items (31 and 44) failed the Flesch–Kincaid reading ease test; item 31 also showed a high mean score in the 12-month data. For spread of scores (the SD should be > 0.8 or < 1.8), items 15 and 31 at baseline and items 15 and 17 at 12 months showed low SDs. The pattern of floor and ceiling effects were similar for the same time point data sets: in baseline data set A there were 12 items with floor effects and 19 items with ceiling effects; in baseline data set B the same pattern was repeated, albeit with 14 items showing floor effects and 18 showing ceiling effects. In the 12-month data sets fewer items showed floor effects (8 items in data set A and seven items in data set B) but a similar pattern of ceiling effects as before (16 items in data set A and 18 items in data set B). For individual item correlations, only one pair of items correlated > 0.06 in the baseline A data set (items 39 and 40). In both the 12-month data sets four item pairs correlated > 0.06 (items 7 and 10; items 8 and 11; items 26 and 27; and items 39 and 40). Overall, one item (31) failed on more than two of the criteria (reading ease, means, SDs and ceiling effects).

As data screening revealed issues with item 31, this was removed from all data sets for the subsequent factor analyses. Initially, the analysis team had agreed on data imputation rules in preparation for the confirmatory factor analyses. However, as polychoric correlations were used for these analyses, this was no longer necessary. Additionally, no data needed to be discarded, and the sample size could be maintained at 662.

### Results of exploratory factor analysis on baseline data set A

An initial EFA yielded 10 factors with eigenvalues of > 1.00, although the scree plot indicated that only three to six factors were likely to be plausible, and any factors beyond that would contribute very little to variance explained. A four-factor solution accounted for 42.68% cumulative variance. The loading pattern matched that of the questionnaires, although with some substantial cross-loadings; we named the four factors as:

- factor 1: behaviours and strategies (considerable cross-loading with factor 3)
- factor 2: family approval/behaviours and child attitudes (items 29 and 32 loaded most on factor 1)
- factor 3: confidence and motivations (with high cross-loadings to other factors)
- factor 4: peer norms (item 16 loading substantially more with factor 2).

## Results of exploratory factor analysis on baseline data set A

Replication with the 12-month data set A yielded a similar pattern. Cumulative variance was 42.60% with a four-factor solution. Again, substantial cross-loadings were noticeable:

- factor 1: additional behaviours and strategies (considerable cross-loading with factor 3)
- factor 2: family approval/behaviours and child attitudes (items 29, 30 and 32 loaded most on factor 1)
- factor 3: confidence and motivation (with high cross-loadings to other items especially item 6)
- factor 4: peer norms (item 16 loading substantially more with factor 2).

**TABLE 52** Screening summary of MLQ Likert scale items for baseline data sets A and B and 12-months data sets A and B

| MLQ item | Flesch-Kincaid reading ease (> 60) | Data set | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | A | | | | | B | | | | |
| | | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) |
| **Baseline MLQ items** | | | | | | | | | | | |
| 6 | 62.1 | 2.93 | 1.008 | 13.1 | 34.4 | 9 (0.55) | 2.95 | 0.936 | 10.7 | 30.7 | 9 (0.52) |
| 7 | 65.7 | 2.68 | 1.060 | 17.5 | 27.7 | 10 (0.57) | 2.70 | 1.055 | 16.4 | 28.7 | 10 (0.56) |
| 8 | 69 | 2.24 | 1.083 | 33.4 | 15.9 | 11 (0.57) | 2.26 | 1.106 | 33.3 | 18.1 | 11 (0.58) |
| 9 | 67.3 | 2.80 | 1.092 | 18.7 | 32.9 | 6 (0.55) | 2.86 | 1.014 | 13.7 | 31.5 | 6 (0.52) |
| 10 | 69.8 | 2.45 | 1.115 | 26.6 | 23.1 | 7 (0.57) | 2.58 | 1.124 | 23.4 | 27.5 | 7 (0.56) |
| 11 | 72.6 | 2.36 | 1.083 | 28.6 | 18.5 | 8 (0.57) | 2.35 | 1.124 | 31.0 | 20.7 | 8 (0.58) |
| 12 | 103 | 3.10 | 1.045 | 12.0 | 48.3 | 13 (0.45) | 3.08 | 1.017 | 10.5 | 45.4 | 13 (0.41) |
| 13 | 74.3 | 2.46 | 1.123 | 27.0 | 23.6 | 14 (0.46) | 2.44 | 1.135 | 28.7 | 23.3 | 43 (0.46) |
| 14 | 81.9 | 2.87 | 1.170 | 19.9 | 42.9 | 13 (0.46) | 2.84 | 1.164 | 21.0 | 39.9 | 13 (0.45) |
| 15 | 95.7 | 2.21 | 0.744 | 10.9 | 8.3 | 20 (0.27) | 2.16 | 0.672 | 10.3 | 5.4 | 18 (0.19) |
| 16 | 79.8 | 2.78 | 0.954 | 12.5 | 24.4 | 24 (0.35) | 2.89 | 0.934 | 10.0 | 28.8 | 24 (0.31) |
| 17 | 85 | 1.91 | 0.830 | 36.5 | 3.1 | 16 (0.26) | 1.94 | 0.814 | 33.6 | 2.9 | 16 (0.25) |
| 18 | 88.4 | 2.24 | 0.958 | 24.0 | 12.5 | 21 (0.46) | 2.23 | 0.961 | 26.0 | 11.4 | 21 (0.49) |
| 19 | 101.4 | 2.84 | 1.089 | 16.8 | 35.5 | 20 (0.44) | 2.78 | 1.102 | 19.1 | 32.9 | 20 (0.51) |
| 20 | 95.4 | 2.81 | 1.069 | 16.1 | 33.2 | 22 (0.45) | 2.84 | 1.106 | 17.2 | 37.1 | 19 (0.51) |
| 21 | 105 | 2.31 | 1.034 | 27.6 | 15.1 | 18 (0.46) | 2.34 | 1.040 | 26.4 | 16.4 | 18 (0.49) |
| 22 | 83.3 | 2.48 | 1.161 | 28.7 | 26.1 | 20 (0.45) | 2.51 | 1.194 | 29.4 | 29.2 | 20 (0.51) |
| 23 | 83.3 | 2.80 | 1.107 | 17.1 | 36.0 | 24 (0.52) | 2.88 | 1.068 | 14.2 | 37.4 | 25 (0.43) |
| 24 | 85 | 3.11 | 1.069 | 12.1 | 50.8 | 23 (0.52) | 3.12 | 1.056 | 11.8 | 50.3 | 27 (0.46) |

| MLQ item | Flesch-Kincaid reading ease (> 60) | Data set | | | | | | | | | |
| | | A | | | | | B | | | | |
| | | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 25 | 73 | 2.84 | 1.145 | 18.3 | 40.4 | 28 (0.50) | 2.92 | 1.116 | 16.3 | 42.2 | 28 (0.52) |
| 26 | 93.5 | 3.03 | 0.959 | 8.4 | 39.0 | 27 (0.60) | 3.08 | 0.923 | 7.3 | 39.3 | 28 (0.53) |
| 27 | 89.9 | 2.74 | 1.060 | 16.5 | 29.9 | 26 (0.60) | 2.73 | 1.026 | 14.2 | 28.6 | 26 (0.50) |
| 28 | 80.8 | 2.92 | 1.131 | 17.3 | 42.9 | 27 (0.51) | 2.91 | 1.107 | 16.2 | 41.0 | 26 (0.53) |
| 29 | 72.8 | 3.27 | 0.815 | 4.0 | 46.4 | 50 (0.34) | 3.28 | 0.840 | 4.8 | 48.1 | 50 (0.32) |
| 30 | 64.9 | 3.20 | 0.936 | 7.6 | 47.9 | 27 (0.48) | 3.20 | 0.922 | 6.5 | 47.6 | 27 (0.42) |
| 31 | 25.5 | 3.44 | 0.866 | 5.6 | 63.3 | 36 (0.32) | 3.52 | 0.783 | 3.7 | 66.9 | 36 (0.31) |
| 32 | 81.8 | 2.80 | 1.119 | 18.7 | 35.7 | 29 (0.28) | 2.71 | 1.162 | 23.0 | 34.0 | 25 (0.32) |
| 33 | 84.8 | 2.19 | 0.948 | 26.9 | 10.4 | 34 (0.57) | 2.14 | 0.943 | 27.6 | 10.8 | 34 (0.58) |
| 34 | 85.1 | 2.20 | 1.073 | 33.5 | 15.9 | 33 (0.57) | 2.24 | 1.076 | 31.6 | 17.1 | 33 (0.58) |
| 35 | 103 | 2.12 | 1.060 | 36.2 | 14.4 | 41 (0.51) | 2.16 | 1.050 | 33.9 | 14.6 | 41 (0.57) |
| 36 | 61.3 | 2.78 | 1.117 | 18.7 | 35.0 | 42 (0.54) | 2.74 | 1.104 | 17.9 | 33.3 | 42 (0.57) |
| 37 | 103.7 | 1.92 | 1.120 | 51.5 | 15.6 | 38 (0.53) | 1.98 | 1.155 | 50.5 | 17.4 | 38 (0.52) |
| 38 | 85.1 | 2.06 | 1.151 | 45.4 | 18.0 | 37 (0.53) | 2.04 | 1.142 | 44.9 | 18.2 | 37 (0.52) |
| 39 | 80.8 | 2.31 | 1.104 | 30.5 | 19.6 | 40 (0.62) | 2.31 | 1.098 | 30.5 | 19.1 | 40 (0.63) |
| 40 | 85.2 | 2.49 | 1.099 | 23.5 | 24.3 | 39 (0.62) | 2.47 | 1.096 | 24.2 | 23.3 | 39 (0.63) |
| 41 | 80.8 | 2.29 | 1.112 | 31.9 | 19.5 | 35 (0.51) | 2.24 | 1.129 | 35.5 | 19.0 | 35 (0.57) |
| 42 | 63.7 | 2.68 | 1.123 | 20.0 | 32.0 | 36 (0.54) | 2.68 | 1.155 | 20.9 | 34.8 | 40 (0.57) |
| 43 | 61.3 | 2.04 | 1.069 | 41.0 | 14.2 | 47 (0.59) | 2.08 | 1.109 | 40.8 | 16.8 | 47 (0.60) |
| 44 | 34.2 | 2.41 | 1.155 | 29.8 | 24.7 | 42 (0.53) | 2.35 | 1.159 | 31.8 | 23.9 | 42 (0.55) |

**TABLE 52** Screening summary of MLQ Likert scale items for baseline data sets A and B and 12-months data sets A and B (*continued*)

| MLQ item | Flesch-Kincaid reading ease (> 60) | Data set | | | | | | | | | |
| | | A | | | | | B | | | | |
| | | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 45 | 65.7 | 1.93 | 1.075 | 47.7 | 13.5 | 47 (0.51) | 1.98 | 1.092 | 46.4 | 14.4 | 47 (0.56) |
| 46 | 65.2 | 2.18 | 1.095 | 35.3 | 17.1 | 47 (0.58) | 2.16 | 1.091 | 35.5 | 17.1 | 47 (0.59) |
| 47 | 65.7 | 1.97 | 1.062 | 45.0 | 12.9 | 43 (0.59) | 1.99 | 1.059 | 42.8 | 13.8 | 43 (0.60) |
| 48 | 85.2 | 2.48 | 1.145 | 26.4 | 26.5 | 39 (0.47) | 2.45 | 1.145 | 27.2 | 25.9 | 47 (0.48) |
| 49 | 94.3 | 2.58 | 1.127 | 21.0 | 30.3 | 50 (0.37) | 2.50 | 1.131 | 23.6 | 27.7 | 48 (0.41) |
| 50 | (103) | 2.93 | 1.063 | 12.7 | 40.3 | 12 (0.43) | 2.83 | 1.065 | 13.1 | 36.6 | 48 (0.46) |
| *12-month MLQ items* | | | | | | | | | | | |
| 6 | 62.1 | 2.97 | 0.876 | 7.3 | 29.5 | 9 (0.59) | 2.99 | 0.900 | 8.8 | 31.0 | 9 (0.57) |
| 7 | 65.7 | 2.64 | 1.003 | 15.2 | 23.5 | 10 (0.63) | 2.66 | 1.029 | 16.0 | 25.6 | 10 (0.66) |
| 8 | 69 | 2.12 | 0.966 | 32.2 | 8.9 | 11 (0.64) | 2.24 | 1.005 | 28.6 | 12.7 | 11 (0.62) |
| 9 | 67.3 | 3.07 | 0.911 | 6.9 | 38.0 | 6 (0.59) | 3.02 | 0.930 | 8.2 | 35.8 | 6 (0.57) |
| 10 | 69.8 | 2.53 | 1.010 | 18.8 | 19.6 | 7 (0.63) | 2.58 | 1.039 | 18.9 | 22.7 | 7 (0.66) |
| 11 | 72.6 | 2.29 | 0.980 | 25.0 | 12.7 | 8 (0.64) | 2.38 | 1.009 | 23.4 | 15.6 | 8 (0.62) |
| 12 | 103 | 3.27 | 0.892 | 6.9 | 49.7 | 50 (0.44) | 3.27 | 0.875 | 6.0 | 49.0 | 13 (0.44) |
| 13 | 74.3 | 2.52 | 1.089 | 23.3 | 23.5 | 14 (0.56) | 2.58 | 1.033 | 18.8 | 22.2 | 14 (0.51) |
| 14 | 81.9 | 2.72 | 1.136 | 21.0 | 33.4 | 13 (0.56) | 2.80 | 1.141 | 19.4 | 37.5 | 13 (0.51) |
| 15 | 95.7 | 2.27 | 0.668 | 7.6 | 4.7 | 28 (0.18) | 2.32 | 0.673 | 5.2 | 6.6 | 12 (0.13) |
| 16 | 79.8 | 2.83 | 0.896 | 8.5 | 24.6 | 17 (0.31) | 2.88 | 0.909 | 7.9 | 28.2 | 24 (0.33) |
| 17 | 85 | 1.74 | 0.744 | 43.3 | 0.8 | 16 (0.31) | 1.75 | 0.751 | 42.7 | 0.8 | 16 (0.27) |
| 18 | 88.4 | 2.24 | 0.944 | 24.3 | 11.0 | 21 (0.54) | 2.30 | 0.957 | 22.5 | 12.7 | 21 (0.56) |

| MLQ item | Flesch-Kincaid reading ease (> 60) | Data set | | | | | | | | | |
| | | A | | | | | B | | | | |
| | | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) |
| 19 | 101.4 | 2.89 | 1.040 | 13.0 | 36.1 | 20 (0.43) | 2.88 | 1.042 | 13.4 | 35.7 | 20 (0.47) |
| 20 | 95.4 | 2.76 | 1.010 | 12.7 | 29.0 | 22 (0.45) | 2.73 | 1.027 | 14.2 | 28.6 | 19 (0.47) |
| 21 | 105 | 2.30 | 0.972 | 24.5 | 12.2 | 18 (0.54) | 2.38 | 1.021 | 24.5 | 15.8 | 18 (0.56) |
| 22 | 83.3 | 2.46 | 1.123 | 26.8 | 23.8 | 20 (0.45) | 2.40 | 1.098 | 27.4 | 20.8 | 20 (0.45) |
| 23 | 83.3 | 2.98 | 1.022 | 11.1 | 40.3 | 24 (0.48) | 3.03 | 1.020 | 10.9 | 42.5 | 24 (0.45) |
| 24 | 85 | 3.32 | 0.959 | 8.0 | 59.2 | 23 (0.48) | 3.35 | 0.919 | 6.3 | 59.5 | 27 (0.47) |
| 25 | 73 | 2.89 | 1.110 | 16.3 | 40.4 | 28 (0.57) | 2.89 | 1.066 | 14.5 | 37.3 | 28 (0.51) |
| 26 | 93.5 | 2.96 | 0.943 | 9.1 | 33.3 | 27 (0.64) | 3.00 | 0.936 | 8.3 | 35.6 | 27 (0.67) |
| 27 | 89.9 | 2.83 | 0.979 | 11.4 | 29.3 | 26 (0.64) | 2.83 | 1.020 | 12.9 | 31.9 | 26 (0.67) |
| 28 | 80.8 | 2.87 | 1.087 | 16.6 | 36.9 | 25 (0.57) | 2.86 | 1.074 | 15.7 | 36.0 | 27 (0.57) |
| 29 | 72.8 | 3.25 | 0.801 | 3.1 | 44.5 | 50 (0.36) | 3.21 | 0.855 | 4.9 | 44.5 | 40 (0.35) |
| 30 | 64.9 | 3.37 | 0.812 | 4.1 | 54.2 | 24 (0.46) | 3.30 | 0.868 | 5.0 | 52.2 | 24 (0.47) |
| 31 | 25.5 | 3.52 | 0.805 | 4.1 | 68.1 | 36 (0.33) | 3.61 | 0.687 | 2.1 | 71.1 | 36 (0.36) |
| 32 | 81.8 | 2.85 | 1.064 | 16.1 | 34.1 | 33 (0.39) | 2.79 | 1.067 | 16.4 | 32.2 | 35 (0.39) |
| 33 | 84.8 | 2.25 | 0.854 | 19.6 | 7.4 | 40 (0.52) | 2.26 | 0.844 | 17.9 | 8.0 | 34 (0.60) |
| 34 | 85.1 | 2.32 | 1.027 | 24.7 | 17.0 | 33 (0.50) | 2.34 | 1.010 | 24.3 | 15.4 | 33 (0.60) |
| 35 | 103 | 2.23 | 1.004 | 28.4 | 13.2 | 41 (0.54) | 2.22 | 0.971 | 26.1 | 12.3 | 41 (0.55) |
| 36 | 61.3 | 2.96 | 1.000 | 10.0 | 38.2 | 42 (0.52) | 2.92 | 0.990 | 10.3 | 35.0 | 42 (0.60) |
| 37 | 103.7 | 1.97 | 1.126 | 48.7 | 16.0 | 38 (0.48) | 2.02 | 1.113 | 44.6 | 16.4 | 38 (0.52) |
| 38 | 85.1 | 2.14 | 1.165 | 41.7 | 20.2 | 43 (0.50) | 2.11 | 1.119 | 40.8 | 16.9 | 43 (0.53) |

**TABLE 52** Screening summary of MLQ Likert scale items for baseline data sets A and B and 12-months data sets A and B  (*continued*)

| MLQ item | Flesch-Kincaid reading ease (> 60) | Data set | | | | | | | | | |
| | | A | | | | | B | | | | |
| | | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) | Mean (> 1.5 or < 3.5) | SD (> 0.8 or < 1.8) | Floor % (< 30%) | Ceiling % (< 30%) | Highest pair item (correlation coefficient) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 39 | 80.8 | 2.27 | 1.031 | 27.3 | 15.9 | 40 (0.63) | 2.33 | 1.037 | 26.2 | 16.4 | 40 (0.71) |
| 40 | 85.2 | 2.43 | 1.041 | 22.6 | 19.2 | 39 (0.63) | 2.51 | 1.006 | 18.7 | 19.3 | 39 (0.71) |
| 41 | 80.8 | 2.33 | 1.094 | 29.0 | 19.6 | 35 (0.54) | 2.30 | 1.045 | 27.1 | 17.0 | 35 (0.55) |
| 42 | 63.7 | 2.83 | 1.042 | 13.8 | 33.2 | 36 (0.52) | 2.81 | 1.039 | 14.7 | 31.3 | 6 (0.60) |
| 43 | 61.3 | 2.15 | 1.035 | 33.6 | 13.5 | 47 (0.65) | 2.11 | 0.992 | 32.6 | 11.8 | 47 (0.58) |
| 44 | 34.2 | 2.51 | 1.120 | 23.9 | 26.2 | 43 (0.53) | 2.45 | 1.097 | 24.6 | 23.1 | 42 (0.56) |
| 45 | 65.7 | 1.99 | 1.003 | 40.3 | 10.5 | 43 (0.57) | 1.96 | 0.992 | 41.2 | 10.2 | 41 (0.52) |
| 46 | 65.2 | 2.18 | 1.085 | 34.3 | 17.1 | 47 (0.57) | 2.16 | 1.063 | 34.3 | 15.3 | 47 (0.54) |
| 47 | 65.7 | 2.05 | 1.014 | 37.6 | 11.2 | 43 (0.64) | 2.03 | 0.980 | 36.5 | 10.0 | 43 (0.58) |
| 48 | 85.2 | 2.50 | 1.112 | 23.0 | 26.3 | 47 (0.52) | 2.50 | 1.073 | 21.5 | 23.5 | 40 (0.47) |
| 49 | 94.3 | 2.60 | 1.061 | 17.5 | 26.7 | 48 (0.39) | 2.60 | 1.055 | 17.0 | 26.6 | 50 (0.40) |
| 50 | 103) | 2.97 | 1.010 | 9.5 | 40.0 | 47 (0.48) | 2.94 | 0.988 | 8.7 | 37.5 | 48 (0.41) |

Shading denotes the items highlighted according to the prespecified criteria in the psychometric evaluation.

## Exploratory factor analysis conclusions

Based on the above-mentioned results, the following subscales were proposed for testing using CFA (subscales have been reordered to reflect item number sequence):

- confidence and motivation (items 6–14)
- peer norms (items 15–22)
- family approval/behaviours and child attitudes (items 23–32 without item 31)
- behaviours and strategies (items 33–50).

### *Results of confirmatory factor analyses*

### Confidence and motivation (items 6–14)

The fit indices of the first model (without any error variance correlations) showed that the solution was marginally adequate. However, the RMSEA was elevated, at 0.114 (*Table 53*). Using an iterative approach, error variances were correlated based on modification indices. However, such correlations were kept to a minimum and conducted based on shared meaning or structure of questions.

After a satisfactory model fit was obtained for the baseline data set B, the exact same model was fitted to the 12-month data set B. The results are also shown in *Table 53*.

The fit for the replication data set was slightly worse. Although the RMSEA was now marginally above the 0.060 cut-off point, the CI (90%) contained 0.060 (0.047 to 0.076). A multigroup invariance test (reported below) will specifically test generalisability of these fits to the 12-month data set.

### Peer norms (items 15–22)

The fit indices of the first model (without any error variance correlations) showed that the solution was inadequate. Additionally, the regression weights for items 15–17 were low, at 0.31, 0.32, and 0.38, respectively. After correlating the error terms for a cluster of three items with shared meaning (Items 19, 20, and 22 were all about helping parents), the fits improved substantially (*Table 54*). Additionally, the regression weights for items 15–17 increased to 0.34, 0.38, and 0.44, respectively.

After a satisfactory model fit has been obtained for the baseline data set B, the exact same model was fitted to the 12-month data set B (see *Table 54*), showing that this solution was also adequate for that data set.

**TABLE 53** Goodness-of-fit indices for the confidence and motivation subscale

| Fit index | Baseline data set B | | 12-month data set B |
| --- | --- | --- | --- |
| | No error variance correlation | Error variance correlation | Error variance correlations |
| Satorra–Bentler scaled $\chi^2$/df | 9.650 | 2.304 | 3.476 |
| RMSEA | 0.114 | 0.044 | 0.061 |
| CFI | 0.946 | 0.993 | 0.987 |
| SRMR | 0.078 | 0.041 | 0.043 |

Results are shown for fits without correlation of any error variances (first model) and with correlations (final model). Fit indices for the replication of the final model with the 12-month data set B are also shown. All results are shown to three decimal places.

TABLE 54 Goodness-of-fit indices for the peer norms subscale

| Fit index | Baseline data set B | | 12-month data set B |
| | No error variance correlation | Error variance correlation | Error variance correlation |
| --- | --- | --- | --- |
| Satorra–Bentler scaled $\chi^2$/df | 10.341 | 2.974 | 3.005 |
| RMSEA | 0.119 | 0.055 | 0.055 |
| CFI | 0.898 | 0.982 | 0.978 |
| SRMR | 0.085 | 0.046 | 0.045 |

Results are shown for fits without correlation of any error variances (first model) and with correlations (final model). Fit indices for the replication of the final model with the 12 months Dataset B are also shown. All results are shown to three decimal places.

## Family approval/behaviours and child attitudes (items 23–32 without item 31)

The fit indices of the first model (without any error variance correlations) showed that the solution was marginally adequate, although the RMSEA was slightly elevated, at 0.088. After correlating error terms for pairs of items with shared meaning, the fit became excellent (*Table 55*).

After a satisfactory model fit had been obtained for baseline data set B, the exact same model was fitted to the 12-month data set B (see *Table 55*), showing that this solution was also excellent for that data set.

## Behaviours and strategies (items 33–50)

The fit indices of the first model (without any error variance correlations) showed that the solution was marginally adequate, although the RMSEA was slightly elevated, at 0.079. After correlating error terms for pairs of items with shared meaning, the fit has become very good (*Table 56*).

After a satisfactory model fit had been obtained for baseline data set B, the exact same model was fitted to the 12-month data set B (see *Table 56*). The fit for the replication data set was worse, and the RMSEA was now > 0.060. The 90% CI (0.065 to 0.078) did not contain this 0.060 cut-off value. A multigroup invariance test (reported below) will provide more detailed information about the adequacy of this model for the 12-month data set.

TABLE 55 Goodness-of-fit indices for the family approval/behaviours and child attitudes subscale

| Fit index | Baseline data set B | | 12-month data set B |
| | No error variance correlation | Error variance correlation | Error variance correlation |
| --- | --- | --- | --- |
| Satorra–Bentler scaled $\chi^2$/df | 6.164 | 2.100 | 2.963 |
| RMSEA | 0.088 | 0.039 | 0.039 |
| CFI | 0.967 | 0.995 | 0.995 |
| SRMR | 0.061 | 0.034 | 0.034 |

Results are shown for fits without correlation of any error variances (first model) and with correlations (final model). Fit indices for the replication of the final model with the 12-month data set B are also shown. All results are shown to three decimal places.

**TABLE 56** Goodness-of-fit indices for the behaviours and strategies subscale

| Fit index | Baseline data set B | | 12-month data set B |
|---|---|---|---|
| | No error variance correlation | Error variance correlation | Error variance correlation |
| Satorra–Bentler scaled $\chi^2$/df | 5.094 | 3.055 | 4.376 |
| RMSEA | 0.079 | 0.056 | 0.072 |
| CFI | 0.977 | 0.990 | 0.981 |
| SRMR | 0.057 | 0.043 | 0.054 |

Results are shown for fits without correlation of any error variances (first model) and with correlations (final model). Fit indices for the replication of the final model with the 12-month data set B are also shown. All results are shown to three decimal places.

## Results of multigroup invariance test

The following analyses were conducted to test specifically the suitability of the factor solution obtained from baseline data set B with the 12-month data set B. These tests were conducted separately for each subscale. A baseline model (model 1) for each subscale tested for generalisability of the final factor solutions reported to the 12-month data set (configural invariance). In this model, the 12-month data set received separate factor loadings and error variances. In the next model (model 2), the factor loadings of the 12-month data set were fixed to the values estimated for the baseline data set. Comparing the fit between models 1 and 2 is thus a test of metric invariance. In other words, if model 2 does not provide a significantly improved fit (as assessed by a chi-squared difference test) compared with model 1, the factor loadings for the two data sets can be assumed to be equal. In the next step, invariance of error variation and the error term correlations was investigated. Model 3 fixed the general error variance and the values of error term correlations of the 12-month data set to the values for the baseline data set. Again, if there is no significant improvement compared with the preceding model (in this case model 2), residual invariance can be concluded. Refer to Milfont and Fischer[153] and Marsh *et al.*[131] for information about measurement invariance testing.

## Results of invariance tests

### Confidence and motivation (items 6–14)

The metric invariance test for the confidence and motivation subscale failed to demonstrate full metric invariance as the comparison of model 2 with model 1 was significant (*Table 57*). An inspection of the results of the CFA with the two data sets separately revealed a clear difference in the factor loading for item 10. For the baseline data set B, the factor loading of item 10 was 0.76, and for the 12-month data set B it was 0.63. As a result, partial invariance was testing whereby only the factor loading of item 10 was allowed to vary between samples (model 2a). The fit of model 2a was not a significant improvement on model 1, confirming partial metric invariance. Last, residual invariance was confirmed through model 3.

**TABLE 57** Summary of the results from the measurement invariance test of the confidence and motivation subscale

| Model | $\chi^2$ (df) | Model comparison | $\Delta\chi^2$ ($\Delta$df) | *p*-value |
|---|---|---|---|---|
| 1 (configural invariance) | 132.280 (46) | – | | |
| 2 (metric invariance) | 151.421 (55) | 1 | 19.141 (9) | 0.024* |
| 2a (partial metric invariance) | 147.492 (54) | 1 | 15.212 (8) | 0.055 |
| 3 (residual invariance) | 153.544 (63) | 2a | 6.052 (9) | 0.735 |

*$p < 0.05$, **$p < 0.01$.

## Peer norms (items 15–22)

The metric invariance test for the peer norms subscale also failed to demonstrate full metric invariance as model 2 provided a significant improvement on model 1 (*Table 58*). An inspection of the results of the CFA with the two data sets separately revealed a clear difference in the factor loading for item 15. For the baseline data set B, the factor loading of item 15 was 0.34, and for the 12-month data set B it was 0.10. As a result, partial invariance was testing whereby only the factor loading of item 15 was allowed to vary between samples (model 2a). The fit of model 2a did not provide a significant improvement on model 1, thereby confirming partial metric invariance. Last, residual invariance was confirmed through model 3.

As in the above CFA conducted separately for the two data sets, the regression weight for item 15 was also low here, suggesting that it may be advisable to consider removing this item from the scale.

## Family approval/behaviours and child attitudes (items 23–32 without item 31)

As with confidence and motivation and peer norms, the metric invariance test for the family approval/behaviours and child attitudes subscale failed to demonstrate full metric invariance. As shown in *Table 59*, model 2 provided a significant improvement on model 1. An inspection of the results of the CFA with the two data sets separately revealed a clear difference in the factor loading for item 24. For the baseline data set B, the factor loading of Item 24 was 0.61, and for the 12-month data set B it was 0.70. As a result, partial invariance was testing whereby only the factor loading of item 24 was allowed to vary between samples (model 2a). The fit of model 2a was not a significant improvement on model 1, confirming partial metric invariance. Last, residual invariance was confirmed through model 3.

## Behaviours and strategies (items 33–50)

As with the other subscales, full metric invariance could not be established for the behaviours and strategies subscale (*Table 60*). An inspection of the results of the CFA with the two data sets separately revealed no very large difference in factor loadings, and for that reason inspection of modification indices guided the selection of factor loadings that were going to be allowed to be vary in a subsequent partial invariance test (model 2a). However, even when eight factor loadings were allowed to vary freely in the second sample, the fit of model 2a was significantly different from that of model 1.

**TABLE 58** Summary of the results from the measurement invariance test of the peer norms subscale

| Model | $\chi^2$ (df) | Model comparison | $\Delta\chi^2$ ($\Delta$df) | *p*-value |
|---|---|---|---|---|
| 1 (configural invariance) | 101.612 (34) | – | | |
| 2 (metric invariance) | 119.083 (42) | 1 | 17.471 (8) | 0.026* |
| 2a (partial metric invariance) | 111.697 (41) | 1 | 10.085 (7) | 0.184 |
| 3 (residual invariance) | 116.319 (49) | 2a | 4.622 (8) | 0.797 |

*$p < 0.05$, **$p < 0.01$.

**TABLE 59** Summary of the results from the measurement invariance test of the family approval/behaviours and child attitudes subscale

| Model | $\chi^2$ (df) | Model comparison | $\Delta\chi^2$ ($\Delta$df) | *p*-value |
|---|---|---|---|---|
| 1 (configural invariance) | 100.026 (40) | – | | |
| 2 (metric invariance) | 117.789 (49) | 1 | 17.763 (9) | 0.038* |
| 2a (partial metric invariance) | 114.793 (48) | 1 | 14.767 (8) | 0.064 |
| 3 (residual invariance) | 119.690 (57) | 2a | 4.897 (9) | 0.843 |

*$p < 0.05$, **$p < 0.01$.

**TABLE 60** Summary of the results from the measurement invariance test of the behaviours and strategies subscale

| Model | $\chi^2$ (df) | Model comparison | $\Delta\chi^2$ ($\Delta$df) | *p*-value |
|---|---|---|---|---|
| 1 (configural invariance) | 931.872 (252) | - | | |
| 2 (metric invariance) | 988.929 (270) | 1 | 57.057 (18) | < 0.001** |
| 2a (partial metric invariance) | 958.43 (262) | 1 | 26.558 (10) | 0.003** |

*$p < 0.05$, **$p < 0.01$.

## Discussion

The HeLP trial provided the opportunity to develop a self-report tool to assess children's views on what factors influenced them in making healthy lifestyle choices. The development phases were an important starting point but had limited resources, and therefore much of psychometric evaluation took place with the main trial data set. Although it would have been better to have an established questionnaire before commencing the trial, the process that has been undertaken still exceeds the evaluation normally afforded to new trial measures in child health research.[154] The large sample size and the low proportion of missing data further enhance our confidence in the data obtained for our analysis, and demonstrate that we have met our first objective: that the MLQ contains items that are relevant, acceptable and feasible for 9- to 10-year-olds to complete.

For the second objective, we have gone some way towards establishing that the MLQ has psychometrically robust elements. The 45 Likert scale items have been comprehensively examined and the four-factor solution identified through the EFA was subsequently tested separately on the baseline and 12-month data sets B. With some minor modifications (correlation of error variance based on shared meaning and structure), very good to excellent fits could be obtained for all subscales. Testing of multigroup invariance revealed some very minor differences between the baseline and 12-months data sets. For the three of the subscales (confidence and motivation, peer norms and family approval/behaviours and child attitudes), the factor loading of one item each was required to vary in the replication sample, thus establishing at least partial metric invariance. The regression weight for item 15 was very low, however, indicating that it was a candidate for removal from the scale; however, we agreed to retain the item owing to its conceptual importance and function together with the other two peer norms items (16 and 17). Unfortunately, neither metric nor partial metric invariance could be established for the behaviours and strategies subscale. However, the fits for the individual CFAs were very good, so the subscale can still be used. The lack of metric invariance indicates that the behaviours and strategies subscale will need to be used with caution when making comparisons across time points. Furthermore, we have not, as yet, undertaken additional testing of the knowledge variable, which remains a simple summary score of children's responses to five questions. There are some additional limitations to the work undertaken, in that although a parallel test retest study was considered it proved not possible to undertake within the resources available, and as yet we have not tested the MLQ for associations with other similar published questionnaires in order to further establish the construct validity of the MLQ.

The overall strength of the MLQ evaluation is the emergence of five variables (knowledge, confidence and motivation, peer norms, family approval/behaviours and child attitudes, and behaviours and strategies) that link to the underpinning models for the intervention (the IMB model and the trial logic model). A further limitation of this work is that the predominantly statistically driven approach and the relatively small number of composite variables means that the theoretical detail underpinning some individual items has been lost. However, these five provide a useful parsimonious set of variables for the trial mediational analysis, suggesting that we have met our third objective of demonstrating that the MLQ is fit for purpose to use in a mediational analysis as part of the process evaluation of the HeLP trial.

## Conclusion

We have developed and evaluated a self-report tool, the MLQ, which assesses the components of the HeLP trial intervention. The MLQ contains items that are relevant, acceptable and feasible for 9- to 10-year-olds to complete in a timely manner, and much of the MLQ has been assessed for its psychometric properties. Five composite variables have emerged from the statistical analysis and have been deemed fit for use in a mediational analysis as part of the process evaluation of the HeLP trial. Future additional evaluation of the MLQ could include further tests for reliability and construct validity and further examination and use of the MLQ items that have a strong theoretical basis.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

*Published by the NIHR Journals Library*