BMJ Open

Long-term effects of the Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010957
Article Type:	Research
Date Submitted by the Author:	08-Jan-2016
Complete List of Authors:	Anderson, Emma Louise; University of Bristol, MRC Integrative Epidemiology Unit Howe, Laura; University of Bristol, Social Medicine Kipping, Ruth; University of Bristol, Faculty of Medicine and Dentistry Campbell, Rona; University of Bristol, School of Social and Community Medicine Jago, Russ; University of Bristol, Centre for Exercise, Nutrition and Health Sciences Noble, Sian; University of Bristol, School of Social and Community Medicine Wells, Sian; University of Bristol, School of Social & Community Medicine Chittleborough, Catherine; University of Adelaide, Discipline of Public Health, School of Population Health and Clinical Practice Peters, Tim; University of Bristol, Department of Community Based Medicine Lawlor, Debbie; Department of Social Medicine, University of Bristol, MRC Integrative Epidemiology Unit
Primary Subject Heading:	Public health
Secondary Subject Heading:	Paediatrics, Nutrition and metabolism
Keywords:	children, randomised controlled trial, schools, physical activity, diet

Long-term effects of the Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial

Emma L Anderson,[1,2] Laura D Howe,[1,2] Ruth R Kipping,[1] Rona Campbell,[1] Russell Jago,[3] Sian M Noble,[1] Sian Wells,[1] Catherine Chittleborough,[4] Tim J Peters,[1,5] Debbie A Lawlor.[1,2]

[1] School of Social & Community Medicine, University of Bristol, 39 Whatley Road, Bristol, BS8 2PS

[2] MRC Integrative Epidemiology Unit at the University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN

[3] Centre for Exercise, Nutrition & Health, School for Policy Studies, University of Bristol, 8 Priory Road, Bristol BS8 1TZ

[4] School of Public Health, University of Adelaide, 178 North Terrace, Adelaide, South Australia 5005

[5] School of Clinical Sciences, University of Bristol, 69 St Michael's Hill, Bristol, BS2 8DZ

Corresponding author:

Debbie A Lawlor

MRC Integrative Epidemiology Unit at the University of Bristol

Oakfield House, Oakfield Grove

Bristol BS8 2BN, UK

Tel: +44 (0)117 33 10096

E-mail: d.a.lawlor@bristol.ac.uk

Acknowledgements: We thank all the students and teaching staff who took part in AFLY5. We thank all of the AFLY5 staff who include fieldworkers, administrative staff, computing and data management staff and the trainers who provided teacher training. We thank Dr Hugh Annett (retired Director of Public Health, NHS Bristol and Bristol City Council), Annie Hudson (former Strategic Director for Children, Young People and Skills, Bristol City Council) and Sheila Smith (Strategic Director for Children, Young People and Skills, North Somerset City Council) for their support of the Active for Life Year 5 study. We also thank the Chair and members of the trial steering committee for their advice and support.

The views expressed in this paper are those of the authors and not necessarily anyone in this acknowledgement list.

Funding: The AFLY5 RCT is funded by the UK National Institute for Health Research (NIHR) Public Health Research Programme (09/3005/04), which also paid the salary of ELA and SW. DAL and LDH work in a Unit that receives funds from UK Medical Research Council (MC_UU_12013/5). RRK and RC work in the Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research Centre of Excellence: joint funding (MR/KO232331/1) from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the Welsh Government and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. LDH is supported by a UK Medical Research Council Population Health Scientist fellowship (G1002375). DAL (NF-SI-0611-10196) and TJP (NF-SI-0512-10026) are NIHR Senior Investigators. This study was undertaken in collaboration with the Bristol Randomised Trials Collaboration (BRTC), a UKCRC Registered Clinical Trials Unit in receipt of National Institute for Health Research CTU support funding.

None of the funders had involvement in the Trial Steering Committee, the data analysis, data interpretation, data collection, or writing of the paper. DAL, LDH and ELA had access to all of the data in the study and DAL had the final responsibility for the decision to submit for publication.

The views expressed in this publication are those of the authors and not necessarily any of the funding bodies listed here.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: support from research funders in accordance with the funding statement included in this manuscript; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work, other than that RC is director of DECIPHer Impact, a not-for-profit company that is wholly owned by the Universities of Bristol and Cardiff and whose purpose is to licence and support the implementation of evidenced based health promotion interventions.

Ethical approval: Ethical approval was obtained from the University of Bristol Faculty of Medicine and Dentistry Committee for Ethics (reference number 101115).

Transparency declaration: DAL affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Data sharing: Full details of the study are available on the website (www.bristol.ac.uk/social-community-medicine/projects/afl/). We encourage anyone who would like to access these data for other projects to contact the corresponding author. We would be happy for external collaborators to access these data according to our data transfer agreement.

Contributors: DAL, together with RRK, RC, TJP, SN and RJ designed the AFLY5 study and obtained funds to complete it. DAL wrote the analysis plan used for this paper and LDH and ELA completed all analyses. SW managed the AFLY5 study, including managing data collection. ELA, LDH and DAL wrote the first draft of the paper, and DAL coordinated contributions from other co-authors. All authors contributed to the overall study aim and development of the design. All authors made critical comments on drafts of the paper.

ABSTRACT

Objective To investigate the long-term effectiveness of a school-based intervention to improve physical activity and diet in children.

Design Cluster randomised controlled trial.

Setting 60 primary schools in the south west of England.

Participants Primary school children who were aged 8-9 years at recruitment, 9-10 years during the intervention, and 10-11 years at the long-term follow-up assessment.

Intervention Teacher training, provision of lesson and child-parent interactive homework plans and teaching materials.

Main outcome measures Primary outcomes were accelerometer assessed minutes of moderate to vigorous physical activity (MVPA) per day, accelerometer assessed minutes of sedentary behaviour per day, and reported daily consumption of servings of fruit and vegetables.

Results 60 schools with 2221 eligible children were recruited. As in the previously published assessment immediately after the end of the intervention, none of the three primary outcomes differed between children in schools allocated to the intervention, compared to those in control schools at the end of the long-term follow-up (1-year after the end of the intervention). Differences in secondary outcomes were consistent with those at the immediate follow-up, with no evidence that these had diminished over time. Comparing intervention to control schools, the difference in mean child-reported screen viewing at the weekend was -16.03 minutes (95%CI: -32.82, 0.73), for servings of snacks per day the difference was -0.11 (95%CI: -0.39, 0.06), in servings of high energy drinks per day -0.20 (95%CI: -0.39, -0.01) and in servings of high fat foods per day -0.12 (95%CI: -0.39, 0.00). None of these reached our predefined level of statistical significance, especially after accounting for multiple testing.

Conclusion This theory driven school-based intervention may have some beneficial effects on reducing screen viewing time and consumption of snacks, high energy drinks and fatty foods that persist for up to 12 months after the end of the intervention.

Trial registration Current Controlled Trials ISRCTN50133740.

INTRODUCTION

Low levels of physical activity and fruit and vegetable consumption in childhood track into adulthood (1-3) and are associated with greater adiposity, adverse cardiometabolic risk factors, behavioural problems, low mood, and poorer academic attainment.(1-7) School-based interventions have the potential to efficiently change behaviours to healthier levels, or delay age-related changes in behaviour,(8) since most children attend school.

The Active for Life Year 5 (AFLY5) study(9) was a large school-based cluster randomised controlled trial (RCT) that was designed to addressed many of the limitations that had been identified in previous RCTs of interventions to improve physical activity and diet in children.(10-15) At the end of the intervention period (immediate follow-up), the intervention was ineffective at improving any of the three primary outcomes (time spent in moderate to vigorous physical activity, time spent in sedentary activity and fruit and vegetable consumption); however, it did result in improvements in three of the nine secondary outcomes (child-reported time spent screen-viewing, consumption of snacks and consumption of high-energy drinks).(16) A cluster randomised control trial design was necessary given the intervention is at the level of schools (rather than individual children).

In this paper, we report the long-term effects of the intervention on the primary and secondary outcomes that were assessed approximately 12 months post-intervention. Our aim was to determine whether any effects on primary outcomes emerged and whether effects on secondary outcomes that were observed immediately after the intervention were maintained, decreased or increased. In this and the previous paper the intervention is delivered at the cluster (school) level and outcomes measured and analysed on individual children, with the

clustering appropriately taken account of in the statistical analyses.

METHODS

Study design and participants

AFLY5 was a school-based, cluster RCT. A total of 60 state primary and junior schools were recruited between March and July 2011: 46 in Bristol and 14 in North Somerset, South West England. At the time of recruitment participants were aged 8-9. Full details of the trial have been published previously and will only be given in brief here.(9, 16, 17) The trial was registered prior to recruitment of schools or data collection (http://www.controlled-trials.com/ISRCTN50133740). Analyses have been undertaken in accordance with a published analytical plan that was approved by the Trial Steering Committee.(9, 17)

Ethical approval and consent

Ethical approval was obtained from the University of Bristol Faculty of Medicine and Dentistry Committee for Ethics (reference number 101115). Parents/guardians of children in Year 4 were sent a letter and information sheet about the study, with an opt-out consent form for each of the measurements and the opportunity to contact the research team to discuss the study as well as information about being able to withdraw at any stage. An information sheet for the child was sent at the same time that the letter was sent to the parents. Children were given a second copy of this information sheet at the time that measurements were undertaken and they were asked to give signed assent to each of the measurements.

Randomisation

Schools were defined as having high or low involvement in any initiatives aimed at increasing physical activity, reducing sedentary behaviour or increasing fruit and vegetable

consumption, based on their report of involvement in local or national initiatives, and also by thirds of their score on the English Index of Multiple Deprivation 2010 (IMD 2010).(18) Schools were grouped into six mutually exclusive strata by these two characteristics and randomly allocated to control or intervention within these strata.(9, 17) Randomisation was undertaken by DAL who was unaware of any other characteristics of the schools. School was concealed using the Bristol Randomised Trials Collaboration's automated (remote) system. After randomisation, one school refused to undertake the intervention; the head reported that they had hoped they would be randomised to control and did not have the time or capacity to accommodate the intervention. This school was retained in the relevant analyses on an intention-to-treat basis.

Intervention

Full details of the trial intervention have been published in the trial protocol and the paper reporting the immediate effect of the intervention.(9, 16) It comprised:

1. Training for classroom teachers and learning support assistants, provided by the trial manager, a nutritionist and physical education specialist. The training took place over a whole day (8-9 hours) in a location away from any of the schools and where the teachers / learning support assistants and those delivering the training would not be interrupted. Teachers / learning support assistants were given a choice of days to attend the training and schools were financially compensated for the cost of replacement teachers whilst their staff attended training. At the training days the rationale for the intervention was explained and each lesson and homework were discussed and then taught in interactive ways. Time was provided for questions and discussion. Teachers were instructed to deliver 16 lessons, 10 of which had associated homeworks. They were told that they could adapt the teaching plans and materials, as

they would with other lessons, for example, to suit their own style and the range of abilities in their class, but the aims and knowledge / skills to be imparted should not be changed.

2. Provision of 16 lesson-plans and teaching materials, including pictures, CDs and journals for teachers or learning support assistants to deliver over two out of the three school-terms (6-7 months). The 16 lessons included 9 lessons that were primarily related to how to be more active and less sedentary and why this was important, 6 to healthy nutrition and how to achieve this and 1 about reducing screen viewing. Each lesson did, however, combine different aspects of healthy behaviour. For example, in the physical activity lessons the children played games based on the food groups using photographs of food which reinforced the content of the nutrition lessons. Similarly, in the lesson (and associated homework) for reducing screen-viewing (called 'Freeze my TV') children were taught how to replace regular television watching with active play on some days.
3. Provision of 10 parental-child interaction homework activities. The homeworks were designed to involve parents and other family members in the behaviour change process by reinforcing the messages delivered during lessons. The homeworks included activities such as: 'Freeze my TV', in which a time / programme normally spent watching television would be replaced with physically active play involving the parents and other family members that the child would write a log about; cooking simple healthy food at home; playing 'Top Grubs' a card game based on trumps with pictures of food, such that higher scoring (trumping) foods are the healthier ones; and measuring the sugar content of drinks that the family have at home or include in school/work lunch packs.

4. Information was provided for schools to insert (as they wished) in the school newsletters about the importance of increasing physical activity, reducing sedentary behaviour and improving diet. The inserts were sent to all intervention schools on three occasions over the period of the intervention. Schools were free to edit these and insert none, all or some of them.
5. Written information for parents on how to encourage their children to eat healthily and be active was delivered via the school children at the start of the intervention.

The intervention took place when the children were aged 9-10 years (in UK school Year 5) after baseline assessment. Schools randomised to the control group continued standard education provision for the school year, and any involvement in additional health promoting activities, but had no access to the intervention teacher training or the teaching materials.

Outcomes

Box 1 lists the three primary and nine secondary outcomes.

Participant assessments

Baseline assessment (prior to intervention) was undertaken either between April and June 2011 or between September and November 2011, when the children were aged 8 to 9 years (i.e. before and after the school summer break). Immediate follow-up assessment was completed immediately post intervention approximately 12-months after the baseline assessment and the long-term assessment (with which this paper is concerned) took place 12-months after the immediate assessment, during which time the children were not exposed to the intervention. Every attempt was made to undertake the assessments in the same order so that the seasons would be similar at each assessment time.

Assessments measured primary and secondary outcomes, together with demographic characteristics and were conducted identically at each time point following published protocols.(9, 17) They were completed by trained fieldworkers who were blinded as to which arm of the trial schools had been allocated. Full details of these assessments have been published previously (9, 17) and are summarised here. Questionnaires asked for information on dietary intake and screen-time viewing and other characteristics and were administered in the classroom with at least one fieldworker present. Weight, height and waist circumference were measured in a private room by one of the trained fieldworkers, with a second fieldworker present in the room. All fieldworkers had passed Criminal Records Bureau checks, as required for working with children at the time that these data were collected. Physical activity was assessed using ActiGraph GT3X+ accelerometers (Actigraph LLC, Pensacola, Florida, USA) and time spent per day being sedentary and in moderate to vigorous activity were calculated using standard protocols as described previously.(9, 17)

Sample size calculation and account of multiple testing

Sample size calculations indicated that for the three primary outcome and nine secondary outcome measurements (including taking account of multiple testing with the secondary outcomes) a total of 60 schools with 1500 pupils (750 in each arm) needed to be recruited, so that 1275 (allowing for loss to follow-up) pupils could be included in the analyses.(9) This number - provided adequate power to detect what we considered to be minimally important effects.(9, 17) We recruited 60 schools and a total of 2,221 pupils, and included between 1066 and 2051 pupils in our analyses for different outcomes. Analyses for accelerometer based outcomes were on fewer participants than our sample size calculation suggested (N = 1066) because of a large proportion of participants not returning or not wearing the

accelerometer for at least eight hours for three days, the minimum required to be included in the study.(9, 17)

Statistical Analyses

Full details of the analysis plan have been published previously.(17) Briefly, main analyses assessing the effect of the intervention on the primary and secondary 12 months post-intervention were conducted as intention-to-treat, with missing data at baseline dealt with by including an indicator variable, as recommended by White et al.(19-21) A series of sensitivity analyses were conducted to test assumptions regarding the nature of missing data at baseline and at each of the follow-up assessments (see detailed analysis plan (17) for discussion of these assumptions and the sensitivity analyses). Multilevel regression models were used to account for clustering (non-independence) of children within schools.(17) All analyses included adjustment for the following baseline variables: age, sex, baseline measure of the outcome being analysed, involvement in other healthy behaviour promoting activities and school level deprivation. A secondary per-protocol analysis was undertaken, in which classes in the intervention arm were only included in analyses if teachers had taught at least 70% (11 of 16) of the AFLY5 lessons. There was one school for which we were unable to confirm how many lessons had been taught. For that school, we first did analyses assuming that they had been taught at least 11 lessons and then repeated them assuming that they had been taught fewer than 11; the results were identical whichever of these alternatives were used. We additionally assessed whether the effect of the intervention on accelerometer-assessed outcomes differed by week or weekend day and whether the results were affected by implausible values as defined previously. The researchers undertaking the analyses were blinded (unaware of) to whether schools had been allocated to intervention or control arms.

As detailed in the published statistical protocol (17) we initially planned to assess change in outcomes between baseline and the long-term follow-up using multilevel models to estimate a trajectory of the repeat measurements (baseline, immediate follow-up, long-term follow-up) within each individual, with random effects to quantify the estimated person-specific deviation from the study mean in terms of the intercept (baseline measurement) and rate of change (slope). However, when we attempted to run these models, they did not converge. This is likely because there were only three measurement occasions, meaning that the model did not have sufficient degrees of freedom. Therefore, we conducted analyses at a single time point as described above (that is, assessed the effect of the intervention on outcomes at the long-term follow-up) and plotted differences between the randomised groups at each time point in order to illustrate any notable changes in estimates of the primary and secondary outcomes between baseline and immediate and long-term follow-up.

RESULTS

Figure 1 shows the trial profile. Of the 2,242 potentially eligible children in the 60 participating schools, 10 left the school prior to randomisation and baseline data collection and for 11 their parents or carers did not provide consent to participate in any aspect of the study. All other children (N = 2,221; 1064 in the schools that were randomised to intervention and 1157 in those randomised to control schools), irrespective of whether or not we have all the data for them, are included in the analyses presented here. Proportions with data for each outcome were similar in intervention and control schools at both baseline and at the second follow-up assessment at 12 months post-intervention (**Figure 1**). Baseline characteristics were similar between children in intervention schools and those in control

schools (**Table 1**).

Figure 2 shows differences in means between the control and intervention group for the three primary and nine secondary outcomes at baseline, immediate follow-up and long-term (12-months) follow-up. These show that differences in means (and odds ratios for general and central overweight/obesity) between children in intervention and control schools were essentially the same at this long-term follow-up as they were immediately after the intervention. **Table 2** shows differences in means for all outcomes at the long-term follow-up from the main intention-to-treat analyses. None of the three primary outcomes differed between children in schools allocated to the AFLY5 intervention and those allocated to the control group at the end of the long-term follow-up. Differences in secondary outcomes were consistent with those seen at the end of the immediate follow-up, with no evidence that the previously reported beneficial effects for child-reported screen viewing at weekends, consumption of snacks and consumption of high energy drinks had notably diminished (or increased) over time. Consumption of high fat foods also appeared lower in children from intervention schools. However, none of these reached our predefined level of statistical significance after accounting for multiple testing. There was no evidence of an effect of the intervention on other secondary outcomes (screen-viewing during week days, mean BMI, mean waist circumference, general overweight/obesity or central overweight or obesity).

Results from the per-protocol analyses were consistent with the intention-to-treat analyses results (**Table 3**). Results were similar in all sensitivity analyses applying different assumptions about missing data (**Supplementary Tables S1-S4**). Results were also similar when we looked separately at time spent in MVPA and time spent in sedentary behaviour by weekday and weekend (**Supplementary Table S5**).

DISCUSSION

In this school-based cluster RCT, aimed at increasing physical activity, reducing sedentary behaviours and improving diet in school aged children, we found results at 12 months after the intervention had ended (that is, with no further lessons or teaching aimed at promoting healthy activity and dietary levels during that 12 months) were essentially the same those seem immediately after the end of the intervention. The lack of any effect on the three primary outcomes – time spent in MVPA, time spent in sedentary behaviour and fruit and vegetable consumption – was still observed 12 months later and the beneficial effects on three secondary outcomes (reported screen-viewing at weekends, consumption of snacks and of high energy drinks) were still present at 12-months post intervention. Some evidence of benefit in terms of consumption of high fat food was also observed in this long-term follow-up.

Meaning of study findings

Whilst the effects for these secondary outcomes were consistent in magnitude with those seen at the immediate follow-up, they did not reach our predefined level of statistical significance. Thus, our results suggest that the AFLY5 intervention may have some beneficial effect on childhood diet that is sustained for at least 12 months, though we cannot rule out that the long-term effect is due to chance.

As discussed in our previous publication of effects immediately at the end of the intervention, the lack of effect on primary outcomes, in particular on the objectively assessed accelerometer outcomes might highlight the importance of societal, structural changes to support greater levels of activity, over and above any intervention at a school level.(16) A

detailed process evaluation showed that fidelity of intervention implementation was good, but that teachers' enthusiasm for the AFLY5 programme was mixed despite them believing that the messages behind the lessons were important.(22) That evaluation highlighted that in general teachers did not like teaching physical activity, and had a tendency to delegate the activity lessons to teaching assistants. This might in part have contributed to the null effects, particularly for the activity outcomes. Despite developing an intervention that we had shown in pilot work fitted well with the primary school national curriculum in the UK,(23) and our process evaluation showing that on average 77% of the intervention lessons and homeworks were delivered and reached 95% of the children in intervention schools, teachers felt lack of time and the need to prioritise numeracy and literacy skills over the health promoting lessons of our intervention were important barriers to them and the children being more fully engaged with AFLY5.(22) Lastly, our process evaluation suggests that in the context of rapidly developing technologies the time taken to develop, test the feasibility of, and pilot, school-based interventions before completing large scale RCTs, as we have done in AFLY5, may mean that by the time school-based interventions get to the full scale RCT, the intervention is being implemented with out-of-date methods of delivery.(22, 24) Thus, whilst using schools for universal promotion of healthy behaviours is appealing, it may be that greater resources and support within schools, and wider engagement of the whole community, is necessary to achieve major shifts towards healthier behaviours.

Strengths and limitations

The study was designed to take account of known sources of bias in other RCTs in this area. A protocol was published before recruitment started, and a detailed analysis plan was written before any access to the study data. We developed an intervention according to guidelines for complex interventions, with the theoretical rationale for the intervention described in detail

elsewhere.(16) Our sample size calculation, which took account of the likely degree of clustering within schools, indicated that we needed a total of 1275 children to be included in the analyses. For all outcomes, except those related to accelerometer data, we achieved considerably higher numbers than this target. The number included in the main analyses for accelerometer based data was somewhat smaller than this at 1066. Sample size calculations are an approximation of the numbers needed, and we doubt that such a small difference will have had a major effect on our conclusions. Furthermore, wear time was similar in children in intervention and control schools; moreover, in sensitivity analyses using different approaches to dealing with missing data and which included 2051 children even for the accelerometer outcomes, the results were essentially the same as in the main analysis. One school refused to deliver any of the intervention, and others did not deliver all of the lessons. However, the per-protocol analysis, which did not differ from the main intention-to-treat analysis, shows that this does not explain the null results.

Conclusion

This long-term follow-up of a large well-conducted school based RCT has found very similar results to those found immediately after the intervention period, with no evidence of effect on the primary outcomes, but some suggestion that the intervention might be effective in reducing screen-viewing at weekends and reducing consumption of snacks, high-fat foods and high-energy drinks, though these effects on secondary outcomes might be due to chance. Overall, together with our process evaluation these findings suggest that curriculum-based interventions alone are unlikely to make a major impact on promoting healthy levels of physical activity and healthy diets in primary school children.

References

1. Lock K, Pomerleau J, Causer L, Altmann DR, McKee M. The global burden of disease attributable to low consumption of fruit and vegetables: implications for the global strategy on diet. Bull World Health Organ. 2005;83(2):100-8. Epub 2005/03/04.

2. Maynard M, Gunnell D, Emmett P, Frankel S, Davey Smith G. Fruit, vegetables, and antioxidants in childhood and risk of adult cancer: the Boyd Orr cohort. J Epidemiol Community Health. 2003;57(3):218-25. Epub 2003/02/21.

3. Ness AR, Maynard M, Frankel S, Smith GD, Frobisher C, Leary SD, et al. Diet in childhood and adult cardiovascular and all cause mortality: the Boyd Orr cohort. Heart. 2005;91(7):894-8. Epub 2005/06/17.

4. Boreham C, Riddoch C. The physical activity, fitness and health of children. J Sports Sci. 2001;19(12):915-29.

5. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act. 2010;7:40. Epub 2010/05/13.

6. Ness AR, Leary SD, Mattocks C, Blair SN, Reilly JJ, Wells J, et al. Objectively measured physical activity and fat mass in a large cohort of children. PLoS Med. 2007;4(3):e97. Epub 2007/03/29.

7. Ekelund U, Luan J, Sherar LB, Esliger DW, Griew P, Cooper A, et al. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. JAMA. 2012;307(7):704-12. Epub 2012/02/18.

8. Cooper AR, Goodman A, Page AS, Sherar LB, Esliger DW, van Sluijs EM, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). Int J Behav Nutr Phys Act. 2015;12:113. Epub 2015/09/18.

9. Lawlor DA, Jago R, Noble SM, Chittleborough CR, Campbell R, Mytton J, et al. The Active for Life Year 5 (AFLY5) school based cluster randomised controlled trial: study protocol for a randomized controlled trial. Trials. 2011;12:181. Epub 2011/07/26.

10. Dobbins M, Husson H, DeCorby K, LaRocca RL. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. Cochrane Database Syst Rev. 2013;2:CD007651. Epub 2013/03/02.

11. Metcalf B, Henley W, Wilkin T. Effectiveness of intervention on physical activity of children: systematic review and meta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). BMJ. 2012;345:e5888. Epub 2012/10/10.

12. Delgado-Noguera M, Tort S, Martinez-Zapata MJ, Bonfill X. Primary school interventions to promote fruit and vegetable consumption: a systematic review and meta-analysis. Prev Med. 2011;53(1-2):3-9. Epub 2011/05/24.

13. Evans CE, Christian MS, Cleghorn CL, Greenwood DC, Cade JE. Systematic review and meta-analysis of school-based interventions to improve daily fruit and vegetable intake in children aged 5 to 12 y. AM J CLIN NUTR. 2012;96(4):889-901. Epub 2012/09/07.

14. DeMattia L, Lemont L, Meurer L. Do interventions to limit sedentary behaviours change behaviour and reduce childhood obesity? A critical review of the literature. Obes Rev. 2007;8(1):69-81.

15. van Grieken A, Ezendam NP, Paulis WD, van der Wouden JC, Raat H. Primary prevention of overweight in children and adolescents: a meta-analysis of the effectiveness of interventions aiming to decrease sedentary behaviour. Int J Behav Nutr Phys Act. 2012;9:61. Epub 2012/05/30.

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.

17. Lawlor DA, Peters TJ, Howe LD, Noble SM, Kipping RR, Jago R. The Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial protocol detailed statistical analysis plan. Trials. 2013;14(1):234. Epub 2013/07/26.

18. Government DfCaL. The English Indices of Deprivation 2010. London: 2011.

19. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clinical trials. 2012;9(4):396-407. Epub 2012/07/04.

20. 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. Epub 2011/02/09.

21. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med. 2005;24(7):993-1007. Epub 2004/12/01.

22. Campbell R RE, Wells S, Kipping RR, Chittleborough CR, Peters TJ, Lawlor DA, Jago R. Intervention fidelity in a school-based diet and physical activity intervention in the UK: Active for Life Year 5. International Journal of Behavioral Nutrition and Physical Activity. 2015;12: 141.

23. Kipping RR, Payne C, Lawlor DA. Randomised controlled trial adapting US school obesity prevention to England. Archives of Disease in Childhood. 2008;93(6):469-73.

24. Jago R RE, Kipping RR, Wells S, Chittleborough CR, Peters TJ, Mytton J, Lawlor DA, Campbell R. Lessons learned from the AFLY5 RCT process evaluation: Implications for the design of physical activity and nutrition interventions in schools. BMC Public Health. 2015;15:946.

Box 1: AFLY5 primary and secondary outcomes

Primary outcomes

Accelerometer assessed mean time per day spent doing moderate/vigorous physical activity MVPA)

Accelerometer assessed mean time per day spent in sedentary activity

Self-reported (validated questionnaire) servings of fruit and vegetables consumed per day

Secondary outcomes

Self-reported (validated questionnaire) mean time spent screen viewing on a typical weekday

Self-reported (validated questionnaire) mean time spent screen viewing on a typical weekend day

Self-reported (validated questionnaire) servings of snacks consumed per day

Self-reported (validated questionnaire) servings of high fat foods consumed per day

Self-reported (validated questionnaire) servings of high energy drinks consumed per day

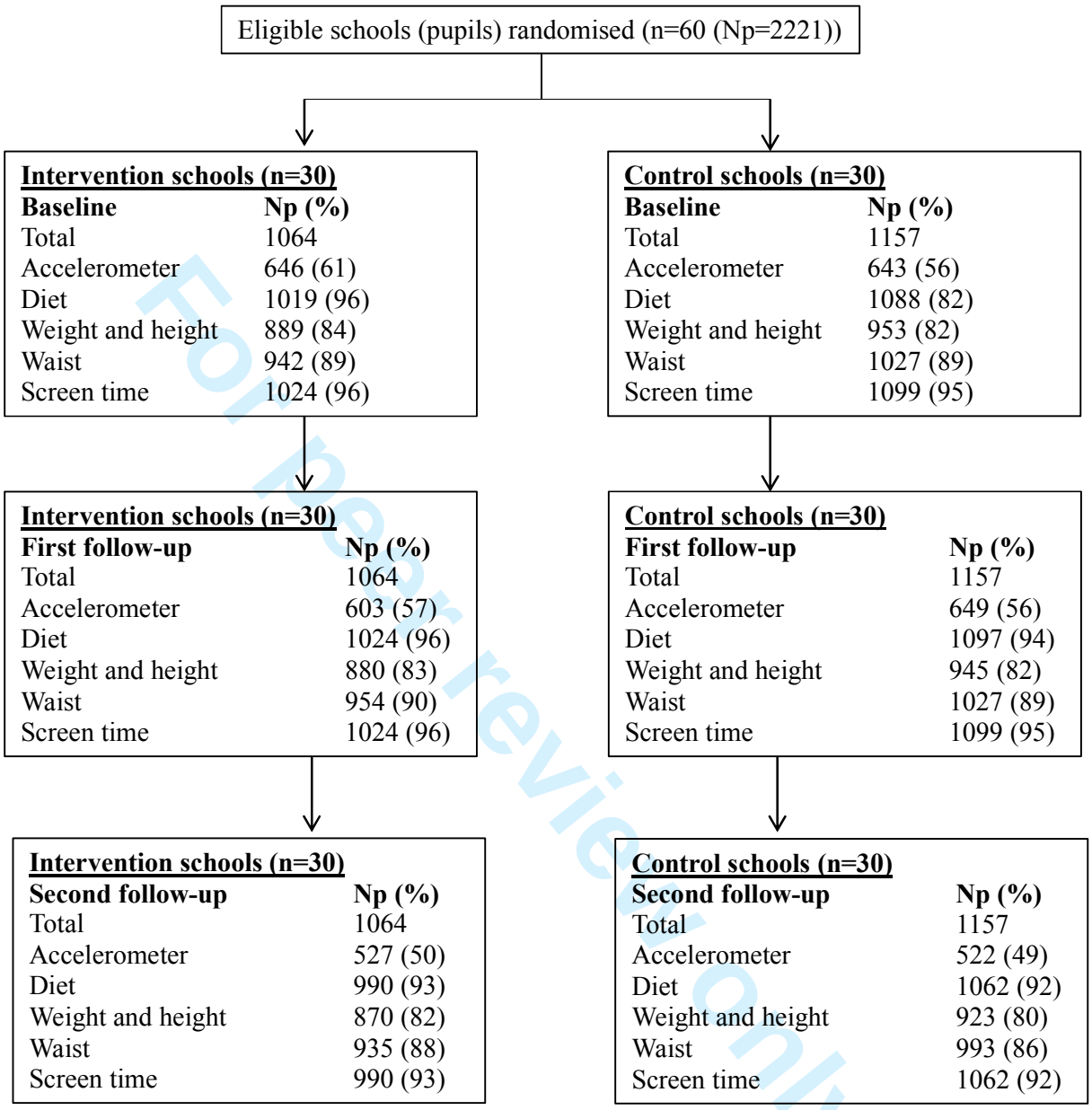
Body mass index determined from weight and height measured in classrooms by two study fieldworkers

Waist circumference measured in classrooms by two study fieldworkers

General overweight/obesity, determined by the International Obesity Task Force thresholds of body mass index for children (taking account of their age and sex)

Central overweight/obesity determined by thresholds of UK age and sex specific reference charts for waist circumference and defined by the International Diabetes Federation

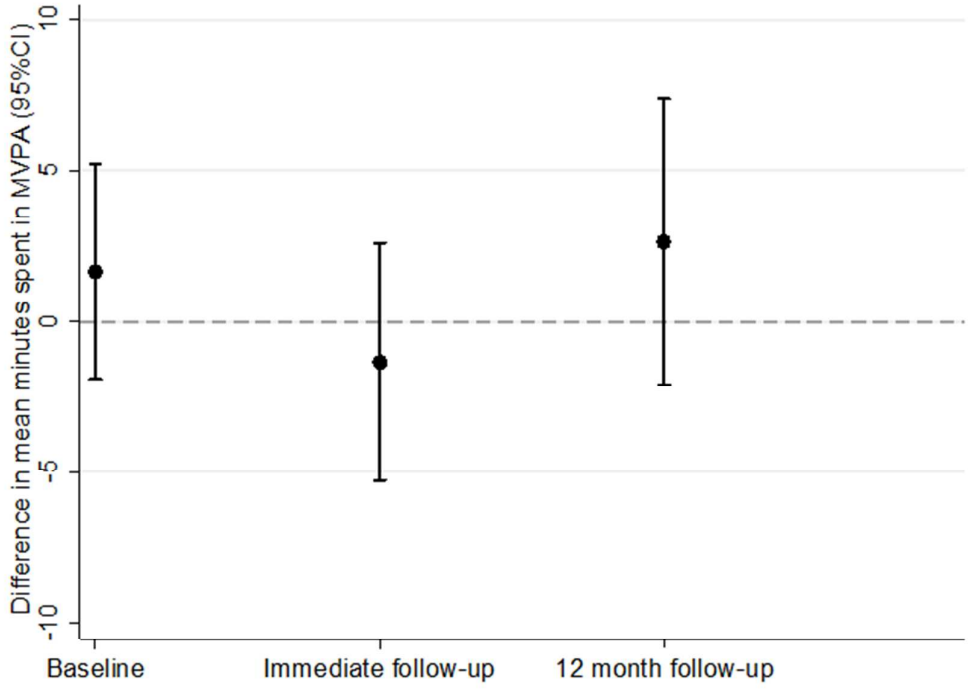
Figure 1 – Trial profile.



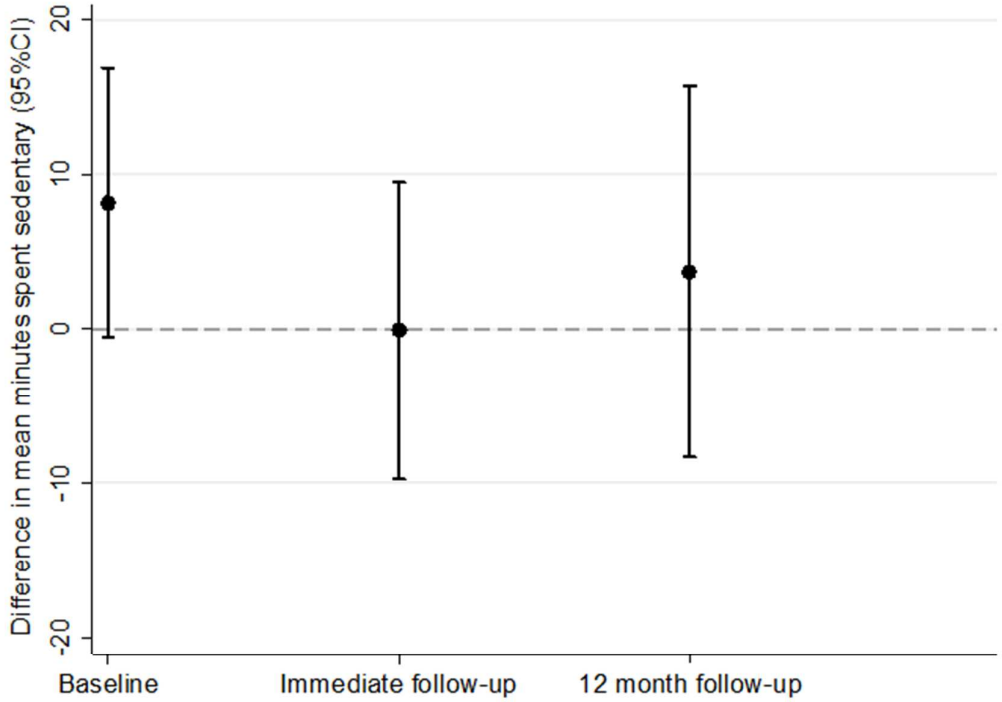
Np=number of participants (school pupils). No schools withdrew from study, so all randomised units are present at baseline and at both follow-up assessments. Percentages for proportions of children with each measurement at baseline and at follow-ups are of total number of children who were pupils in randomised schools at baseline. Not all pupils with follow-up measures necessarily had data on the same measure at baseline (or vice versa), because of different pupils being absent at baseline and follow-up assessments at each time point, and because of pupils leaving or moving between schools. In all analyses, those who were randomised were analysed in the group (intervention or control) to which they were randomised.

Figure 2: Difference in means between the control and intervention group for the three primary outcomes and nine secondary outcomes, assessed at baseline, first follow-up (conducted immediately after the end of the intervention) and second follow-up (12-months post-intervention).

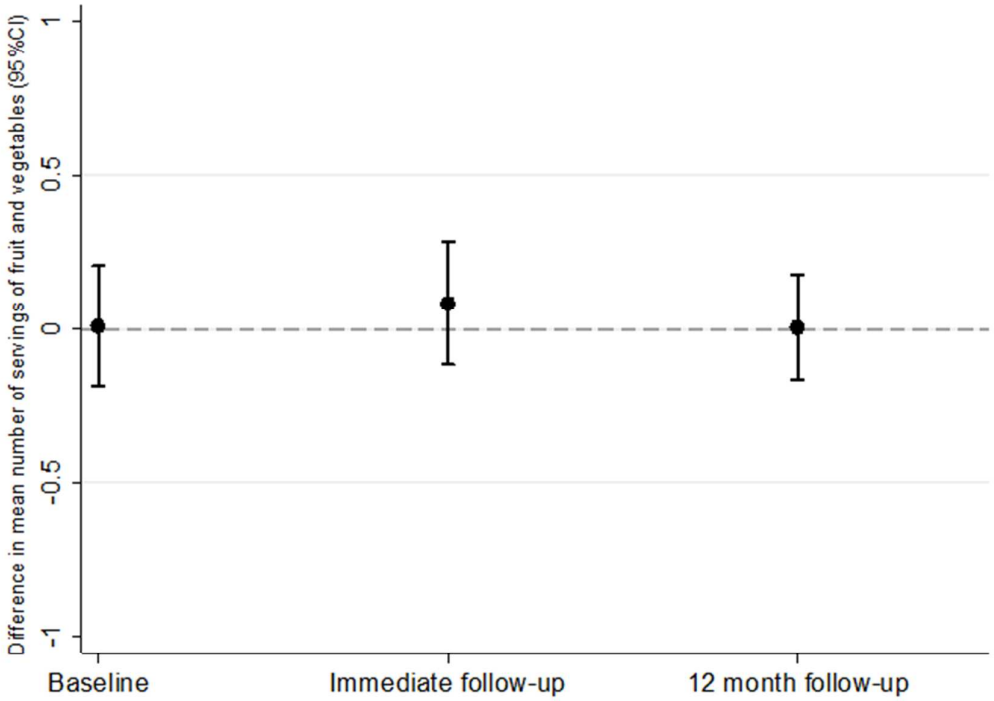
a. Accelerometer assessed time spent in moderate to vigorous physical activity



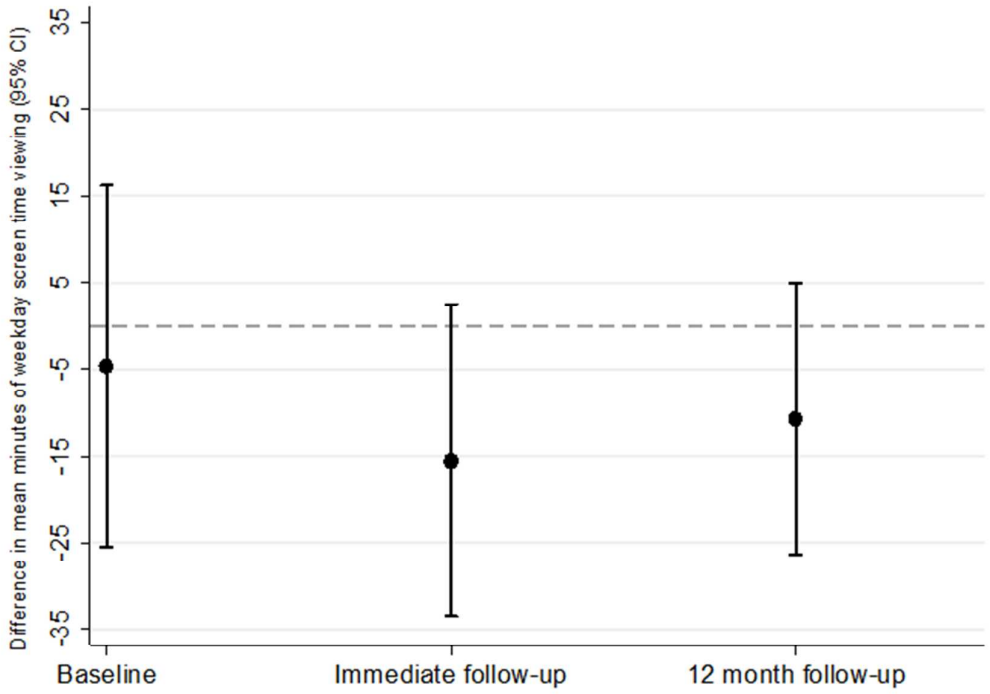
b. Time spent in sedentary behaviour



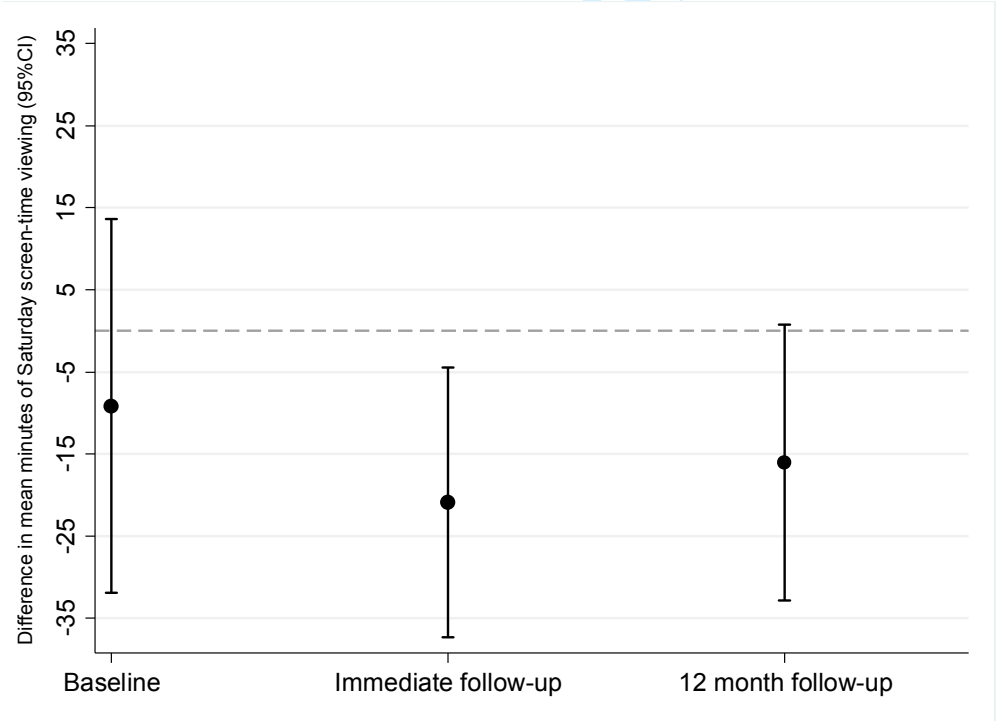
c. Servings of Fruit and Vegetables per day



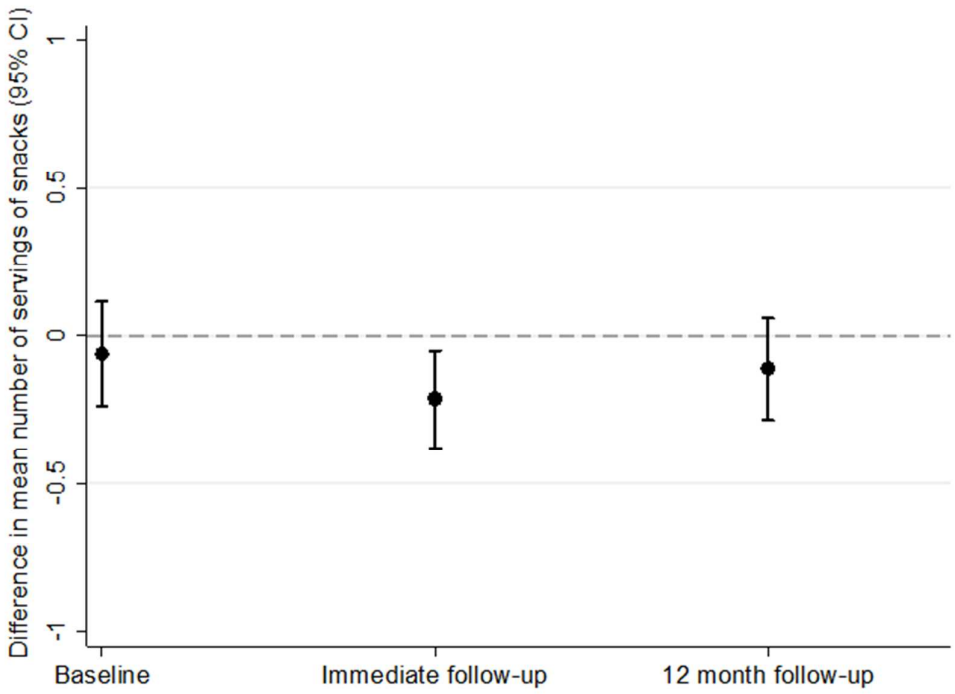
d. Time spent screen viewing on weekdays



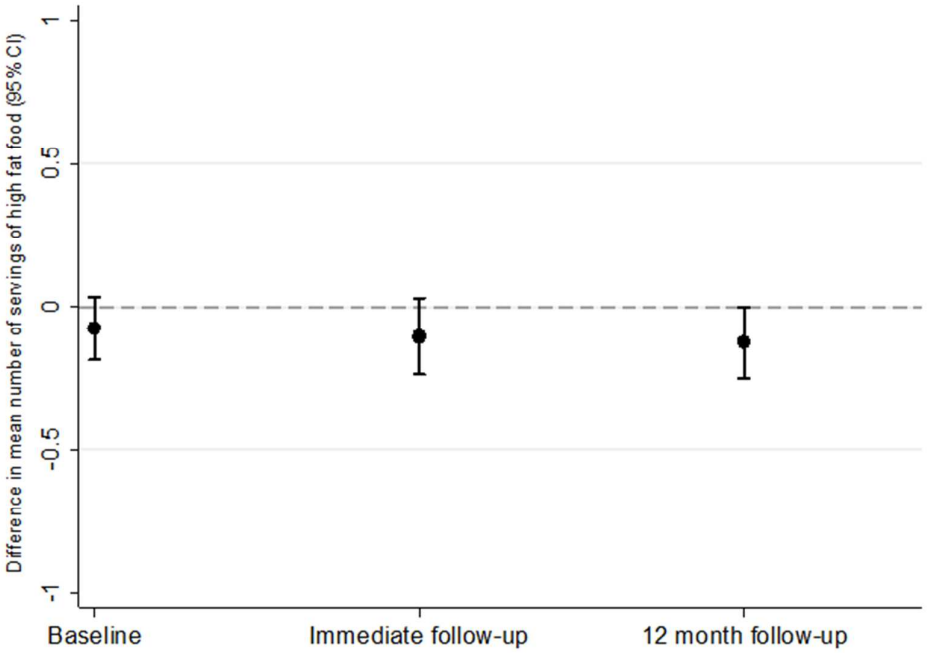
e. Time spent screen viewing on Saturdays



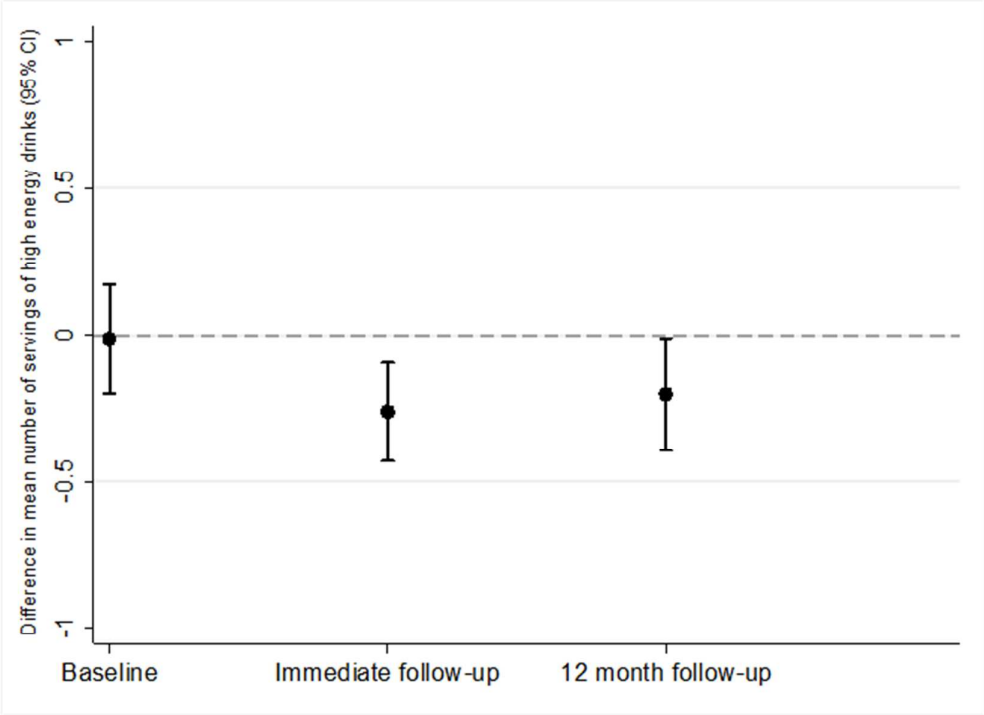
f. Servings of snacks per day



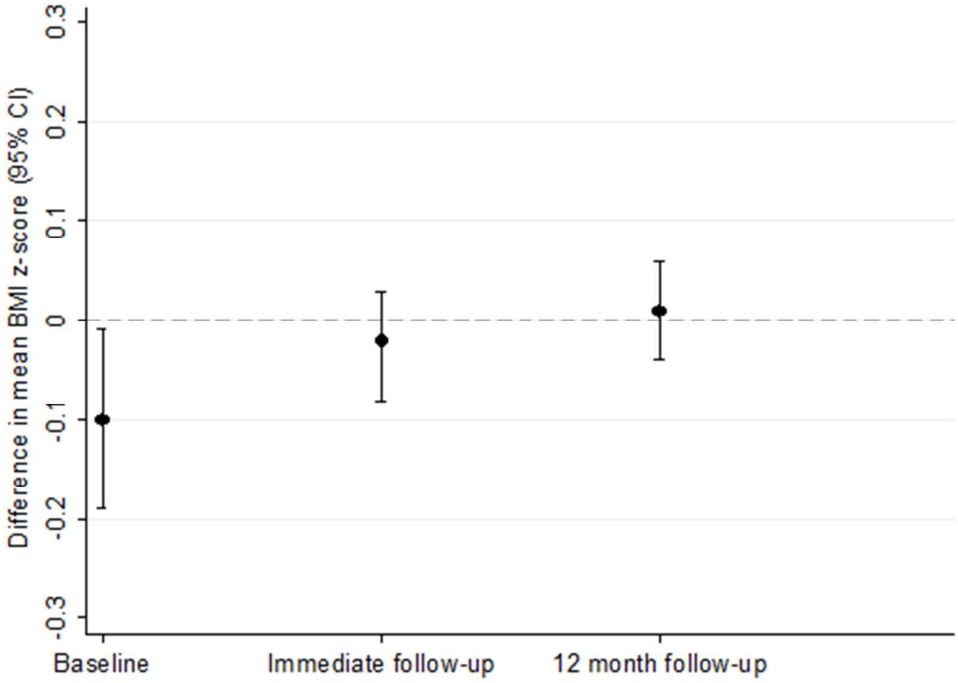
g. Servings of high fat foods per day



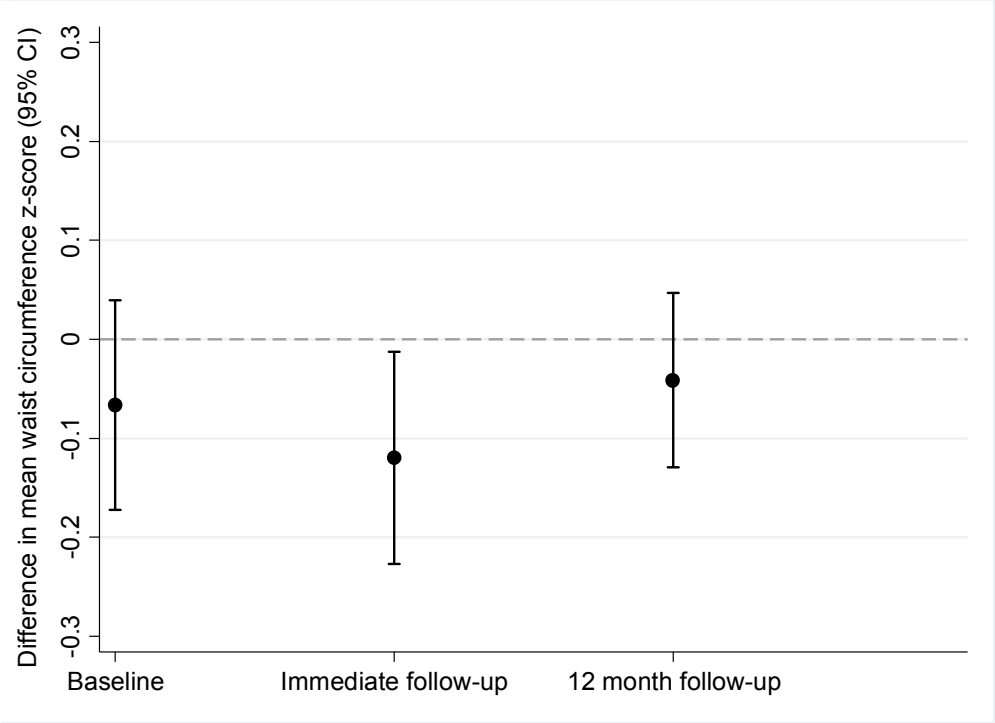
h. Servings of high energy drinks per day



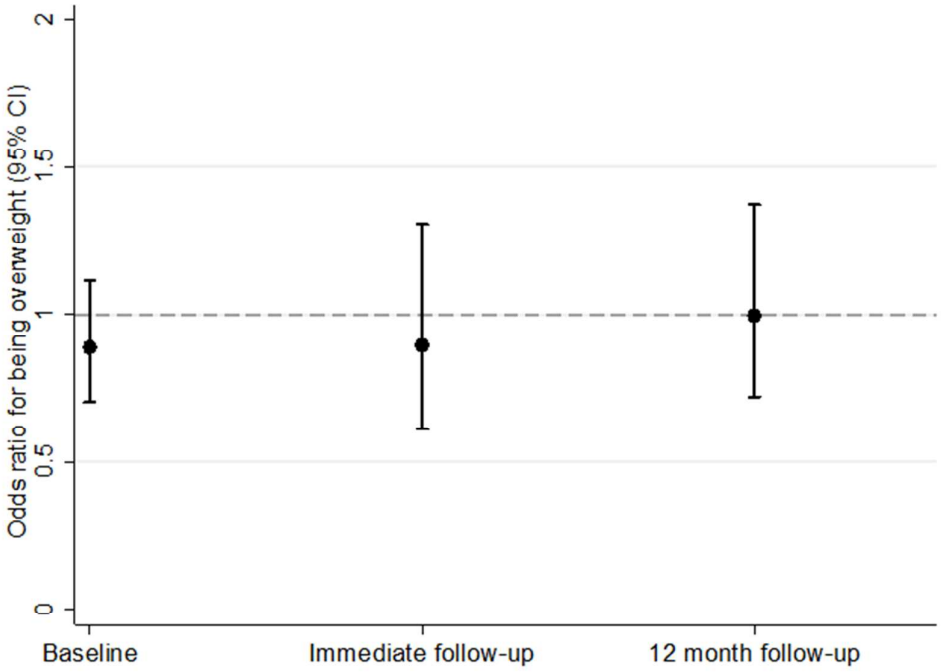
i. Body mass index (as a continuous variable)



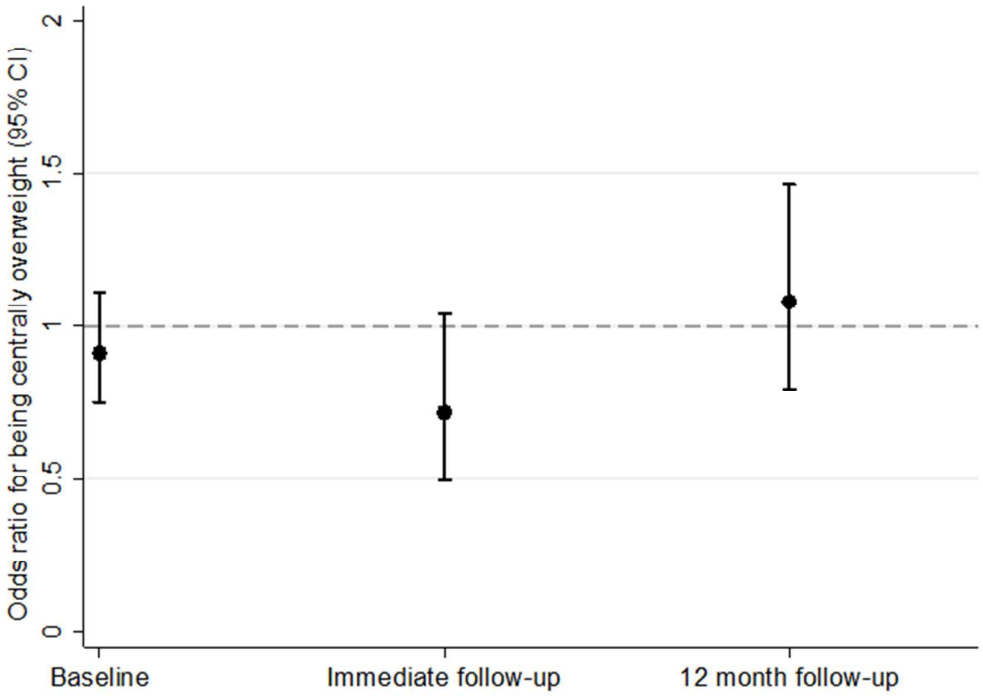
j. Waist circumference (as a continuous variable)



k. General overweight or obesity (based on BMI measurements)



l. Central overweight / obesity based on waist circumference measurements



The figures all show differences in means for continuous variables (graphs a to j) and odds ratios for binary outcomes (graphs k and l), comparing those in the intervention arm of the trial to those in the control arm (dots), together with 95% confidence intervals (vertical lines with horizontal caps representing the limits). The dashed horizontal lines represent the null values (zero for all differences in means of continuous variables and one for odds ratios of binary outcomes).

Table 1: Comparison of baseline characteristics by randomised group

Characteristic	Unit and type of summary measure	Intervention schools Np=1064		Control schools Np=1157	
		Number	Distribution	Number	Distribution
Age	Mean (SD) years	1024	9.5 (0.3)	1099	9.5 (0.3)
MVPA[a]	Mean (SD) minutes	912	59 (23)	928	56 (21)
Sedentary behaviour[a]	Mean (SD) minutes	912	422 (72)	928	416 (68)
Servings of fruit and vegetables	Median (IQR) number / day	1019	1 (0 to 2)	1088	1 (0 to 2)
Servings of snacks	Median (IQR) number / day	1019	2 (1 to 3)	1088	2 (1 to 3)
Servings of high fat foods	Median (IQR) number / day	1019	0 (0 to 1)	1088	1 (0 to 1)
Servings of high energy drinks	Median (IQR) number / day	1019	2 (1 to 3)	1088	2 (1 to 3)
BMI	Mean (SD) z-score	889	-0.06 (0.94)	953	0.05 (1.04)
WC	Mean (SD) z-score	942	-0.03 (0.97)	1027	0.03 (1.02)
Screen-viewing weekday	Median (IQR) minutes	1024	105 (45 to 240)	1099	105 (45 to 225)
Screen-viewing Saturday	Median (IQR) minutes	1024	90 (30 to 240)	1099	105 (30 to 240)
Total number of valid days of wearing accelerometer[b]	Median (IQR) days	912	3 (2 to 5)	928	3 (2 to 4)
Total number of valid weekdays of wearing accelerometer[b]	Median (IQR) days	979	2 (2 to 3)	1025	2 (1 to 3)
Total hours of wearing accelerometer on valid days[a]	Mean (SD) hours / day	912	11.6 (1.5)	928	11.5 (1.4)
Hours of wearing accelerometer on	Mean (SD) hours / day	896	11.8 (1.6)	919	11.7 (1.5)

valid weekdays[b]					
Categorical variables					
Gender	N (%) female	520	49%	608	52%
	N (%) male	544	51%	549	48%
General overweight / obesity	N (%) No	717	81%	743	78%
	N (%) Yes	172	19%	210	22%
Central overweight/obesity	N (%) No	601	64%	631	61%
	N (%) Yes	341	36%	396	39%
Returned accelerometer	N (%) No	85	8%	132	11%
	N (%) Yes	979	92%	1025	89%
Wore accelerometer for requested amount of time	N (%) No	820	77 %	953	82%
	N (%) Yes	244	23%	204	18%
Wore accelerometer for required amount of time	N (%) No	418	39%	514	44%
	N (%) Yes	646	61%	643	56%
School involved in other health promoting activities	N (%) No	264	25%	446	39%
	N (%) Yes	800	75%	711	61%
School deprivation score	N (%) low	315	30%	460	40%
	N (%) medium	368	35%	345	30%
	N (%) high	381	36%	352	30%

Np: number of participants; SD: standard deviation; MVPA: moderate or vigorous physical activity; IQR: interquartile range; BMI: body mass index; WC: waist circumference

[a]Including only participants with at least 3 days of valid data

[b]Including all valid days, regardless of the number of valid days

Note some % within categories do not sum to exactly 100 because of rounding

Table 2: Main intention-to-treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. Numbers of participants vary by outcome as indicated in the table.

Outcome (primary/secondary)	Control group (reference group)		Intervention group		Main comparison between the two groups (Intervention versus Control)		
	Np	Mean (SD) or number (%)	Np	Mean (SD) or number (%)	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes:							
Time spent in MVPA (minutes per day)	522	52.56 (20.67)	527	54.37 (22.23)	1049	2.48 (-1.80, 6.77)	0.26
Time spent in sedentary behaviour (minutes per day)	522	461.78 (66.33)	527	465.46 (70.61)	1049	2.79 (-7.78, 13.37)	0.60
Servings of fruit and vegetables (number per day)	1062	1.80 (1.55)	990	1.82 (1.59)	2051	0.01 (-06, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	1062	148.01 (126.39)	990	138.88 (125.00)	2051	-10.74 (-26.30, 4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	1062	180.52 (164.82)	990	167.71 (156.28)	2051	-16.03 (-32.82, 0.73)	0.06
Body mass index (z-score)	923	0.03 (1.02)	870	-0.03 (0.97)	1793	0.01 (-0.04, 0.06)	0.72
Waist circumference (z-score)	993	0.03 (1.04)	935	-0.03 (0.95)	1928	-0.04 (-0.13, 0.05)	0.36
Servings of snacks (number per day)	1062	2.11 (1.55)	990	1.99 (1.47)	2051	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	1062	0.86 (0.94)	990	0.74 (1.07)	2051	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	1062	2.38 (1.58)	990	2.19 (1.45)	2051	-0.20 (-0.39, -0.01)	0.04
Binary outcomes							
Generally overweight/obese	923	194 (21.02)	870	175 (20.11)	1793	1.00 (0.72, 1.37)	0.98
Centrally overweight/obese	993	421 (42.40)	935	394 (42.14)	1928	1.08 (0.80, 1.46)	0.62

Np: number of participants; SD: standard deviation; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance after taking account of multiple testing)

All differences in means / odds ratios with their 95% CIs have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline/school stratifying variables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses participants were included for each outcome if they had a follow-up measurement of that outcome; for missing baseline data we used an indicator variable as described by White & Thompson,(21) which means for each outcome participants are included even if they do not have a baseline measurement.

Table 3: Per-protocol analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. Numbers vary by outcome as indicated in the table.

Outcome (**primary**/secondary)	**Control group (reference group)**		**Intervention group**		**Main comparison between the two groups (Intervention versus Control)**		
	Np	**Mean (SD) or number (%)**	**Np**	**Mean (SD) or number (%)**	**Np**	**Difference in means or odds ratio (95%CI)**	**p-value**
Continuous outcomes:							
Time spent in MVPA (minutes per day)	522	52.56 (20.67)	356	54.15 (22.27)	878	2.63 (-2.10, 7.37)	0.28
Time spent in sedentary behaviour (minutes per day)	522	461.78 (66.33)	356	466.17 (70.58)	878	3.67 (-8.32, 15.66)	0.55
Servings of fruit and vegetables (number per day)	1062	1.80 (1.55)	701	1.91 (1.66)	1762	0.05 (-0.15, 0.25)	0.63
Time spent screen-viewing (minutes per day weekday)	1062	148.01 (126.39)	701	134.98 (120.94)	1762	-8.97 (-26.81, 8.87)	0.32
Time spent screen-viewing (minutes per day Saturday)	1062	180.52 (164.82)	701	159.35 (149.97)	1762	-21.73 (-41.19, -2.26)	0.03
Body mass index (z-score)	923	0.03 (1.02)	612	-0.03 (0.98)	1535	0.01 (-0.05, 0.07)	0.69
Waist circumference (z-score)	993	0.03 (1.04)	657	-0.04 (0.94)	1650	-0.03 (-0.13, 0.06)	0.52
Servings of snacks (number per day)	1062	2.11 (1.55)	701	2.07 (1.48)	1762	-0.03 (-0.23, 0.16)	0.72
Servings of high fat foods (number per day)	1062	0.86 (0.94)	701	0.75 (1.15)	1762	-0.11 (-0.26, 0.04)	0.14
Servings of high energy drinks (number per day)	1062	2.38 (1.58)	701	2.22 (1.43)	1762	-0.18 (-0.41, 0.5)	0.12
Generally overweight/obese	923	194 (21.02)	612	121 (19.77)	1535	0.98 (0.68, 1.41)	0.91
Centrally overweight/obese	993	421 (42.40)	657	272 (41.40)	1650	1.06 (0.76, 1.49)	0.72

Np: number of participants; SD: standard deviation; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Per-protocol analysis defined as teaching at least 70% (11 out of the 16) AFLY5 lessons. All participants from the intervention schools where the teacher taught fewer than 11 lessons are excluded from these analyses.

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance after taking account of multiple testing)

All differences in means/odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline/school stratifying variables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses, after removal of schools that did not teach at least 11 out of 16 of the lessons, participants were only included for each outcome if they had a follow-up measurement of that outcome. For partial missing baseline data we used an indicator variable as described by White & Thompson,(21) which means for each outcome participants are included even if they do not have a baseline measurement.

Supplementary Table S1: Sensitivity analysis: intention-to-treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes 12 months post-intervention. Numbers vary by outcome as indicated in the table. In these analyses participants were only included for each outcome if they had a baseline and a follow-up measurement of that outcome.

Outcome	Main comparison between the two groups (Intervention versus Control)		
Primary / secondary	**Np**	**Difference in means or odds ratio (95%CI)**	**p-value**
Continuous outcomes:			
Time spent in MVPA (minutes per day)	1000	3.05 (-1.33, 7.44)	0.17
Time spent in sedentary behaviour (minutes per day)	1000	2.21 (-8.28, 12.71)	0.68
Servings of fruit and vegetables (number per day)	1953	0.02 (-0.15, 0.19)	0.83
Time spent screen-viewing (minutes per day weekday)	1965	-10.53 (-26.1, 5.05)	0.19
Time spent screen-viewing (minutes per day Saturday)	1965	-17.3 (-33.71, -0.88)	0.04
Body mass index (z(sd)-score)	1563	0 (-0.05, 0.04)	0.95
Waist circumference (z(sd)-score)	1748	-0.03 (-0.12, 0.05)	0.47
Servings of snacks (number per day)	1953	-0.13 (-0.3, 0.04)	0.13
Servings of high fat foods (number per day)	1953	-0.13 (-0.25, 0)	0.04
Servings of high energy drinks (number per day)	1953	-0.21 (-0.4, -0.02)	0.03
Binary outcomes:			
Generally overweight/obese	1563	0.83 (0.56, 1.22)	0.35
Centrally overweight/obese	1748	1.01 (0.73, 1.4)	0.93

Np: number of participants; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

Supplementary Table S2: Sensitivity analysis: intention-to-treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. In these analyses participants were only included for each outcome if they had a baseline and a follow-up measurement for all three primary outcomes. Numbers included are identical for the three primary outcomes (N = 757) but can vary by outcome for secondary outcomes (though none of these can be higher than 757) as indicated in the table.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	757	1.28 (-3.22, 5.78)	0.58
Time spent in sedentary behaviour (minutes per day)	757	0.60 (-10.44, 11.63)	0.92
Servings of fruit and vegetables (number per day)	757	-0.13 (-0.34, 0.09)	0.26
Time spent screen-viewing (minutes per day weekday)	757	0.20 (-17.54, 17.94)	0.98
Time spent screen-viewing (minutes per day Saturday)	757	-8.46 (-28.49, 1.56)	0.41
Body mass index (z(sd)-score)	682	0.00 (-0.06, 0.07)	0.80
Waist circumference (z(sd)-score)	728	-0.01 (-0.12, 0.09)	0.90
Servings of snacks (number per day)	757	-0.13 (-0.38, 0.13)	0.33
Servings of high fat foods (number per day)	757	-0.13 (-0.33, 0.07)	0.19
Servings of high energy drinks (number per day)	757	-0.12 (-0.37, 0.12)	0.32
Generally overweight/obese	680	1.09 (0.64, 1.85)	0.76
Centrally overweight/obese	728	11.35 (0.81, 2.23)	0.25

Np: number of participants; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

Missing baseline data for secondary outcomes (once those with missing baseline primary outcomes are excluded) were managed as in the main analyses.

Supplementary Table S3: Sensitivity analysis: intention-to-treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention, with missing data for either baseline or follow-up measure of an outcome assumed to be 10% healthier than the average value in the study sample.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	2051	0.74 (-1.59, 3.07)	0.53
Time spent in sedentary behaviour (minutes per day)	2051	1.78 (-4.63, 8.20)	0.59
Servings of fruit and vegetables (number per day)	2051	0.01 (-0.16, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	2051	-10.74 (-26.30, 4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	2051	-16.03 (-32.82, 0.76)	0.06
Body mass index (z(sd)-score)	2051	0.01 (-0.04, 0.06)	0.70
Waist circumference (z(sd)-score)	2051	-0.02 (-0.11, 0.06)	0.56
Servings of snacks (number per day)	2051	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	2051	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	2051	-0.20 (-0.39, -0.01)	0.04
Generally overweight/obese	2051	0.98 (0.76, 1.26)	0.87
Centrally overweight/obese	2051	1.05 (0.77, 1.43)	0.78

Np: number of participants; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses participants all participants are included (N = 2,221 (the number of participants recruited to the study). Missing baseline data is managed as in the main analyses

and missing outcome data are imputed on the basis of those with missing data being 10% healthier than all participants in the study for a given outcome.

Supplementary Table S4: Sensitivity analysis: intention-to-treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention, with missing data for either baseline or follow-up measure of an outcome assumed to be 10% less healthy than the average value in the study sample.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	2051	1.04 (-1.18, 3.26)	0.36
Time spent in sedentary behaviour (minutes per day)	2051	-0.72 (-6.39, 4.95)	0.80
Servings of fruit and vegetables (number per day)	2051	0.01 (-0.16, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	2051	-10.74 (-26.30,4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	2051	-16.03 (-32.82, 0.76)	0.06
Body mass index (z(sd)-score)	2051	0.01 (-0.04, 0.06)	0.70
Waist circumference (z(sd)-score)	2051	-0.02 (-0.11, 0.06)	0.56
Servings of snacks (number per day)	2051	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	2051	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	2051	-0.20 (-0.39, -0.01)	0.04
Generally overweight/obese	2051	0.98 (0.76, 1.26)	0.87
Centrally overweight/obese	2051	1.05 (0.77, 1.43)	0.78

Np: number of participants; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses participants all participants are included (N = 2,221 (the number of participants recruited to the study). Missing baseline data is managed as in the main table and

missing outcome data are imputed on the basis of those with missing data being 10% less healthy than all participants in the study for a given outcome.

Supplementary Table S5: Main intention-to-treat analyses of the effect of AFLY5 intervention on accelerometer-assessed outcomes separately for week and weekend days. Numbers vary by outcome as indicated in the table.

Outcome	Main comparison between the two groups (Intervention versus Control) on week days			Main comparison between the two groups (Intervention versus Control) on weekend days		
	Np	Difference in means or odds ratio (95%CI)	p-value	Np	Difference in means or odds ratio (95%CI)	p-value
Time spent in MVPA (minutes per day)	1627	2.47 (-1.37, 6.32)	0.21	972	3.26 (-3.62, 10.14)	0.35
Time spent in sedentary behaviour (minutes per day)	1627	1.87 (-8.51, 12.24)	0.72	972	3.07 (-10.91, 17.06)	0.67

Np: number of participants; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

All differences in means with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses, participants were only included for each outcome if they had a follow-up measurement of that outcome. For partial missing baseline data we used an indicator variable as described by White & Thompson,(1) which means for each outcome participants are included even if they do not have a baseline measurement.

Only participants included in the main analyses (i.e. with at least 3 valid days of accelerometer data) are included in this sensitivity analysis.

References

1. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med. 2005;24(7):993-1007. Epub 2004/12/01.

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[1,2]	See table 2	5-6
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	7
	2b	Specific objectives or hypotheses	Whether objectives pertain to the the cluster level, the individual participant level or both	7-8
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	8-9
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		**None so no reporting (protocol is published)**
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	8-9
	4b	Settings and locations where the data were collected		**8-9**
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	9-10
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and	Whether outcome measures pertain to the cluster level, the individual participant level or both	11 & Box 1 page 22

		when they were assessed		
	6b	Any changes to trial outcomes after the trial commenced, with reasons		**None so no reporting (protocol is published)**
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or *k*), and an indication of its uncertainty	12-13
	7b	When applicable, explanation of any interim analyses and stopping guidelines		**N/A**
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		**8-9**
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	8-9
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	8-9
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	8-9
	10b		Mechanism by which individual participants were included in clusters for the purposes of the	8-9

			trial (such as complete enumeration, random sampling)	
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		**12 & 13**
	11b	If relevant, description of the similarity of interventions		**N/A**
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	13
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		**13-14**
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	14 & Figure 1 on page 23
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	14 & Figure 1 page 23
Recruitment	14a	Dates defining the periods of recruitment and follow-up		**11**
	14b	Why the trial ended or was stopped		**N/A**
Baseline data	15	A table showing baseline	Baseline characteristics for the	

		demographic and clinical characteristics for each group	individual and cluster levels as applicable for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	31
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	33-34
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		**33-34**
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		**35 and supplementary material**
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[3])		**N/A – intervention was integrated into school teaching curriculum**
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		**17-18**
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	17-18
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		**17-19**

Other information			
Registration	23	Registration number and name of trial registry	**6 & 8**
Protocol	24	Where the full trial protocol can be accessed, if available	**Referenced throughout the paper – reference numbers 9 and 17 in reference list which starts on page19**
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	**2-3**

** Note: page numbers optional depending on journal requirements*

Table 2: Extension of CONSORT for abstracts[1,2] to reports of cluster randomised trials

Item	Standard Checklist item	Extension for cluster trials
Title	Identification of study as randomised	**Identification of study as cluster randomised**
Trial design	Description of the trial design (e.g. parallel, cluster, non-inferiority)	
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	**Eligibility criteria for clusters**
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	**Whether objective or hypothesis pertains to the cluster level, the individual participant level or both**
Outcome	Clearly defined primary outcome for this report	**Whether the primary outcome pertains to the cluster level, the individual participant level or both**
Randomization	How participants were allocated to interventions	**How clusters were allocated to interventions**
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants randomized to each group	**Number of clusters randomized to each group**
Recruitment	Trial status[1]	
Numbers analysed	Number of participants analysed in each group	**Number of clusters analysed in each group**
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	**Results at the cluster or individual participant level as applicable for each primary outcome**
Harms	Important adverse events or side effects	
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register	
Funding	Source of funding	

[1] Relevant to Conference Abstracts

REFERENCES

1 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008, 371:281-283

2 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG at al (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 5(1): e20

3 Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.

BMJ Open

Long-term effects of the Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010957.R1
Article Type:	Research
Date Submitted by the Author:	19-Apr-2016
Complete List of Authors:	Anderson, Emma Louise; University of Bristol, MRC Integrative Epidemiology Unit Howe, Laura; University of Bristol, Social Medicine Kipping, Ruth; University of Bristol, Faculty of Medicine and Dentistry Campbell, Rona; University of Bristol, School of Social and Community Medicine Jago, Russ; University of Bristol, Centre for Exercise, Nutrition and Health Sciences Noble, Sian; University of Bristol, School of Social and Community Medicine Wells, Sian; University of Bristol, School of Social & Community Medicine Chittleborough, Catherine; University of Adelaide, Discipline of Public Health, School of Population Health and Clinical Practice Peters, Tim; University of Bristol, Department of Community Based Medicine Lawlor, Debbie; Department of Social Medicine, University of Bristol, MRC Integrative Epidemiology Unit
Primary Subject Heading:	Public health
Secondary Subject Heading:	Paediatrics, Nutrition and metabolism
Keywords:	children, randomised controlled trial, schools, physical activity, diet

Long-term effects of the Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial

Emma L Anderson,[1,2] Laura D Howe,[1,2] Ruth R Kipping,[1] Rona Campbell,[1] Russell Jago,[3] Sian M Noble,[1] Sian Wells,[1] Catherine Chittleborough,[4] Tim J Peters,[1,5] Debbie A Lawlor.[1,2]

[1] School of Social & Community Medicine, University of Bristol, 39 Whatley Road, Bristol, BS8 2PS

[2] MRC Integrative Epidemiology Unit at the University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN

[3] Centre for Exercise, Nutrition & Health, School for Policy Studies, University of Bristol, 8 Priory Road, Bristol BS8 1TZ

[4] School of Public Health, University of Adelaide, 178 North Terrace, Adelaide, South Australia 5005

[5] School of Clinical Sciences, University of Bristol, 69 St Michael's Hill, Bristol, BS2 8DZ

Corresponding author:

Debbie A Lawlor

MRC Integrative Epidemiology Unit at the University of Bristol

Oakfield House, Oakfield Grove

Bristol BS8 2BN, UK

Tel: +44 (0)117 33 10096

E-mail: d.a.lawlor@bristol.ac.uk

ABSTRACT

Objective To investigate the long-term effectiveness of a school-based intervention to improve physical activity and diet in children.
Design Cluster randomised controlled trial.
Setting 60 primary schools in the south west of England.
Participants Primary school children who were aged 8-9 years at recruitment, 9-10 years during the intervention, and 10-11 years at the long-term follow-up assessment.
Intervention Teacher training, provision of lesson and child-parent interactive homework plans and teaching materials.
Main outcome measures Primary outcomes were accelerometer assessed minutes of moderate to vigorous physical activity (MVPA) per day, accelerometer assessed minutes of sedentary behaviour per day, and reported daily consumption of servings of fruit and vegetables.
Results 60 schools with 2221 eligible children were recruited. As in the previously published assessment immediately after the end of the intervention, none of the three primary outcomes differed between children in schools allocated to the intervention, compared to those in control schools at the end of the long-term follow-up (1-year after the end of the intervention). Differences in secondary outcomes were consistent with those at the immediate follow-up, with no evidence that these had diminished over time. Comparing intervention to control schools, the difference in mean child-reported screen viewing at the weekend was -16.03 minutes (95%CI: -32.82, 0.73), for servings of snacks per day the difference was -0.11 (95%CI: -0.39, 0.06), in servings of high energy drinks per day -0.20 (95%CI: -0.39, -0.01) and in servings of high fat foods per day -0.12 (95%CI: -0.39, 0.00). None of these reached our predefined level of statistical significance, especially after accounting for multiple testing.
Conclusion This theory driven school-based intervention may have some beneficial effects on reducing screen viewing time and consumption of snacks, high energy drinks and fatty foods that persist for up to 12 months after the end of the intervention.

Trial registration Current Controlled Trials ISRCTN50133740.

Study strengths and limitations

Strengths

- The study was designed to take account of known sources of bias in other RCTs in this area.
- A protocol was published before recruitment started, and a detailed analysis plan was written before any access to the study data.

- Random allocation was concealed and outcome assessors were blinded to which group the schools and children and been randomised to.
- Accelerometers were used to objectively assess time spend in moderate to vigorous activity and sedentary behaviour.
- Our sample size calculation, which took account of the likely degree of clustering within schools.

Limitations

- The study was undertaken in state schools in the South West of England that covered a range of deprivation levels and both urban and rural communities, but results may not be generalizable to more ethnically diverse populations in the UK or beyond the UK
- There was missing data for the accelerometer assessed outcomes, but a range of sensitivity analyses did not alter our findings and levels of weartime and valid accelerometer data were similar in intervention and control arms

INTRODUCTION

Low levels of physical activity and fruit and vegetable consumption in childhood track into adulthood (1-3) and are associated with greater adiposity, adverse cardiometabolic risk factors, behavioural problems, low mood, and poorer academic attainment.(1-7) School-based interventions have the potential to efficiently change behaviours to healthier levels, or delay age-related changes in behaviour,(8) since most children attend school. However, previous randomised controlled trials (RCTs) of such interventions have potentially important sources of bias and few have explored long-term outcomes beyond the end of the intervention period.

A systematic review and meta-analysis of 44 school-based RCTs found beneficial effects on moderate or vigorous physical activity (MVPA) during school hours, but the authors noted that benefit might have been exaggerated due to the outcome assessment being self-/parental-reported and not blind to school allocation in most trials and because of the marked loss to follow-up in several trials.(9) In many of those RCTs the intervention included extra compulsory physical activity lessons or activities during school break-times. Those have the advantage that they do not interrupt the school curriculum, but in the absence of any long-term follow-up beyond the intervention period it is impossible to determine whether the greater time spent in MVPA is simply as a result of a level of compulsion to be more active. Evidence from observational epidemiological studies suggests that compulsory physical activity in lessons or break-time in school are associated with more school-based activity, but not with more activity otherwise.(10, 11) A systematic review restricted to studies that had used objectively accelerometers assessed activity and did not restrict the outcome to activity during school hours found some evidence of benefit of a similar magnitude in both family focused and school curriculum interventions, but noted that the magnitude of effect was

modest.(12) Reviews of interventions to reduce time spent in sedentary behaviour have similarly noted some evidence of effect, but cautioned about likely sources of bias, including lack of adequate concealment of random allocation, subjective outcome measurements with no blinding of participants and little evaluation that effects were sustained long-term post intervention.(13,14) Likewise, two systematic reviews of school-based interventions to increase fruit and vegetable consumption found some possible evidence of modest effect but were concerned about lack of adequate concealment of random allocation and failure to take account of clustering within analyses.(15,16)

The Active for Life Year 5 (AFLY5) study (17) was a large school-based cluster randomised controlled trial (RCT). It was designed to addressed many of the limitations that had been identified in previous RCTs of interventions to improve physical activity and diet in children(9-16) by objectively measureing physical activity and sedentary behavoiur and by determing effects on outocmes both immediately after the end of the intervention as well as 12 months later. At the end of the intervention period (immediate follow-up), the intervention was ineffective at improving any of the three primary outcomes (time spent in moderate to vigorous physical activity, time spent in sedentary activity and fruit and vegetable consumption); however, it did result in improvements in three of the nine secondary outcomes (child-reported time spent screen-viewing at weekends, consumption of snacks and consumption of high-energy drinks).(18) A cluster randomised control trial design was necessary given the intervention is at the level of schools (rather than individual children).

In this paper, we report the long-term effects of the intervention on the primary and secondary outcomes that were assessed approximately 12 months post-intervention. Our

initial aim when designing the study was to be able to determine whether any effects of the intervention would last beyond the period of the intervention. Given we now know the immediate post intervention results,(18) our aim in this paper was to determine whether any effects on primary outcomes emerged at the 12 month follow-up assessment (i.e. whether there was a delayed effect of the intervention on the primary outcomes) and whether effects on secondary outcomes that were observed immediately after the intervention were maintained, decreased or increased 12-months after the intervention. In this and the previous paper the intervention is delivered at the cluster (school) level and outcomes measured and analysed on individual children, with the clustering appropriately taken account of in the statistical analyses.

METHODS

Study design and participants

AFLY5 was a school-based, cluster RCT. Clustering was at the level of the schools, with eligibility for study entry being: (i) any state primary or junior schools that (ii) provided education to children aged 8 to 11 years and (iii) were within the Bristol City and North Somerset administrative areas (both areas in the South West of England). All children in UK school year 4 (age 8-9 years) at the time of recruitment were eligible for recruitment if their parents provided consent and they assented (see below).

A total of 60 state primary and junior schools were recruited between March and July 2011: 46 in Bristol and 14 in North Somerset, South West England. At the time of recruitment participants were aged 8-9. Full details of the trial have been published previously so only a brief summary will be given here.(17-19) The trial was registered prior to recruitment of

schools or data collection (http://www.controlled-trials.com/ISRCTN50133740). Analyses have been undertaken in accordance with a published analytical plan that was approved by the Trial Steering Committee.(17-19)

Ethical approval and consent

Ethical approval was obtained from the University of Bristol Faculty of Medicine and Dentistry Committee for Ethics (reference number 101115). Parents/guardians of children in Year 4 were sent a letter and information sheet about the study, with an opt-out consent form for each of the measurements and the opportunity to contact the research team to discuss the study as well as information about being able to withdraw at any stage. An information sheet for the child was sent at the same time that the letter was sent to the parents. Children were given a second copy of this information sheet at the time that measurements were undertaken and they were asked to give signed assent to each of the measurements.

Randomisation

Schools were defined as having high or low involvement in any initiatives aimed at increasing physical activity, reducing sedentary behaviour or increasing fruit and vegetable consumption, based on their report of involvement in local or national initiatives. Schools were also split into tertiles based on their score on the English Index of Multiple Deprivation 2010 (IMD 2010).(20) Schools were grouped into six mutually exclusive strata by these two characteristics and randomly allocated to control or intervention within these strata.(17-19) Randomisation was undertaken by DAL who was unaware of any other characteristics of the schools. School was concealed using the Bristol Randomised Trials Collaboration's automated (remote) system. After randomisation, one school refused to undertake the intervention; the head reported that they had hoped they would be randomised to control and

did not have the time or capacity to accommodate the intervention. This school was retained in the relevant analyses on an intention-to-treat basis.

Intervention

The intervention was adapted from a previously evaluated US intervention(21) and is based on Social Cognitive Theory,(22) with a particular emphasis on increasing the children's self-efficacy (perceived competence) to be physically active and eat a healthy diet.(23) Full details of the trial intervention have been published in the trial protocol and the paper reporting the immediate effect of the intervention.(17, 18) It comprised:

1. Training for classroom teachers and learning support assistants, provided by the trial manager, a nutritionist and physical education specialist. The training took place over a whole day (8-9 hours) in a non-school location and where the teachers / learning support assistants and those delivering the training would not be interrupted. Teachers / learning support assistants were given a choice of days to attend the training and schools were financially compensated for the cost of replacement teachers whilst their staff attended training. At the training days the rationale for the intervention was explained and each lesson and homework activity was discussed and then taught in interactive ways. Time was provided for questions and discussion. Teachers were instructed to deliver 16 lessons, 10 of which had associated homework. They were told that they could adapt the teaching plans and materials, as they would with other lessons, for example, to suit their own style and the range of abilities in their class, but the aims and knowledge / skills to be imparted should not be changed.
2. Provision of 16 lesson-plans and teaching materials, including pictures, CDs and journals for teachers or learning support assistants to deliver over two out of the three school-terms (6-7 months). The 16 lessons included 9 that were primarily related to

how to be more active and less sedentary and why this was important, 6 to healthy nutrition and how to achieve this and 1 about reducing screen viewing. Each lesson did, however, combine different aspects of healthy behaviour. For example, in the physical activity lessons the children played games based on the food groups using photographs of food which reinforced the content of the nutrition lessons. Similarly, in the lesson (and associated homework) for reducing screen-viewing (called 'Freeze my TV') children were taught how to replace regular television watching with active play on some days.

3. Provision of 10 parental-child interaction homework activities. The activities were designed to involve parents and other family members in the behaviour change process and reinforced the messages delivered during lessons. The homeworks included activities such as: 'Freeze my TV', in which a time normally spent watching television would be replaced with physically active play involving the parents and other family members that the child would write a log about; cooking simple healthy food at home; playing 'Top Grubs' a card game based on trumps with pictures of food, such that higher scoring (trumping) foods are the healthier ones; and measuring the sugar content of drinks that the family have at home or include in school/work lunch packs.
4. Information was provided for schools to insert (as they wished) in their school newsletters about the importance of increasing physical activity, reducing sedentary behaviour and improving diet. The inserts were sent to all intervention schools on three occasions over the period of the intervention. Schools were free to edit these and insert none, all or some of them.
5. Written information for parents on how to encourage their children to eat healthily and be active was delivered via the school children at the start of the intervention.

The intervention took place when the children were aged 9-10 years (in UK school Year 5) after baseline assessment. Schools randomised to the control group continued standard education provision for the school year, and any involvement in additional health promoting activities, but had no access to the intervention teacher training or the teaching materials.

Outcomes

Box 1 lists the three primary and nine secondary outcomes.

Participant assessments

Baseline assessment (prior to intervention) was undertaken either between April and June 2011 or between September and November 2011, when the children were aged 8 to 9 years (i.e. before and after the school summer break). Immediate follow-up assessment was completed immediately post intervention approximately 12-months after the baseline assessment and the long-term assessment (with which this paper is concerned) took place 12-months after the immediate assessment, during which time the children were not exposed to the intervention. Every attempt was made to undertake the assessments in the same order so that the seasons would be similar at each assessment time.

Assessments measured primary and secondary outcomes, together with demographic characteristics and were conducted identically at each time point following published protocols.(17,19) They were completed by trained fieldworkers who were blinded as to which arm of the trial schools had been allocated. Full details of these assessments have been published previously (17, 19) and are summarised here. Questionnaires asked for information on dietary intake and screen-time viewing and other characteristics and were administered in the classroom with at least one fieldworker present. Weight, height and waist circumference

were measured in a private room by one of the trained fieldworkers, with a second fieldworker present in the room. All fieldworkers had passed Criminal Records Bureau checks, as required for working with children at the time that these data were collected. Physical activity was assessed using ActiGraph GT3X+ accelerometers (Actigraph LLC, Pensacola, Florida, USA) and time spent per day being sedentary and in moderate to vigorous activity were calculated using standard protocols as described previously.(17, 19)

Sample size calculation and account of multiple testing

Sample size calculations indicated that for the three primary outcome and nine secondary outcome measurements (including taking account of multiple testing with the secondary outcomes) a total of 60 schools with 1500 pupils (750 in each arm) needed to be recruited, so that 1275 (allowing for loss to follow-up) pupils could be included in the analyses.(17) This number - provided adequate power to detect what we considered to be minimally important effects.(17, 19) We recruited 60 schools and a total of 2,221 pupils, and included between 1066 and 2052 pupils in our analyses for different outcomes. Analyses for accelerometer based outcomes were on fewer participants than our sample size calculation suggested (N = 1066) because of a large proportion of participants not returning or not wearing the accelerometer for at least eight hours for three days, the minimum required to be included in the study.(17, 19)

Statistical Analyses

Full details of the analysis plan have been published previously.(19) Briefly, main analyses assessing the effect of the intervention on the primary and secondary 12 months post-intervention were conducted as intention-to-treat, with missing data at baseline being replaced with a value of 999 and a variable to indicate missing data at baseline (0=not

missing, 1=missing) being included in regression models, as recommended by White et al.(24-26) For primary outcomes the level of statistical significance used was $p< 0.05$ and for secondary outcomes the level of statistical significance used was $p<0.01$, after correcting for multiple testing.(19) A series of sensitivity analyses were conducted to test assumptions regarding the nature of missing data at baseline and at each of the follow-up assessments (see detailed analysis plan (19) for discussion of these assumptions and the sensitivity analyses). Multilevel regression models were used to account for clustering (non-independence) of children within schools.(19) All analyses included adjustment for the following baseline variables: age, sex, baseline measure of the outcome being analysed, involvement in other healthy behaviour promoting activities and school level deprivation. A secondary per-protocol analysis was undertaken, in which classes in the intervention arm were only included in analyses if teachers had taught at least 70% (11 of 16) of the AFLY5 lessons. There was one school for which we were unable to confirm how many lessons had been taught. For that school, we first did analyses assuming that they had been taught at least 11 lessons and then repeated them assuming that they had been taught fewer than 11; the results were identical whichever of these alternatives were used. We additionally assessed whether the effect of the intervention on accelerometer-assessed outcomes differed by week or weekend day and whether the results were affected by implausible values as defined previously. The researchers undertaking the analyses were blinded to (unaware of) whether schools had been allocated to intervention or control arms.

As detailed in the published statistical protocol (19) we initially planned to assess change in outcomes between baseline and the long-term follow-up using multilevel models to estimate a trajectory of the repeat measurements (baseline, immediate follow-up, long-term follow-up) within each individual, with random effects to quantify the estimated person-specific

deviation from the study mean in terms of the intercept (baseline measurement) and rate of change (slope). However, when we attempted to run these models, they did not converge. This is likely because there were only three measurement occasions, meaning that the model did not have sufficient degrees of freedom. Therefore, we conducted analyses at a single time point as described above (that is, assessed the effect of the intervention on outcomes at the long-term follow-up) and plotted differences between the randomised groups at each time point in order to illustrate any notable changes in estimates of the primary and secondary outcomes between baseline and immediate and long-term follow-up.

RESULTS

Figure 1 shows the trial profile. Of the 2,242 potentially eligible children in the 60 participating schools, 10 left the school prior to randomisation and baseline data collection and for 11 their parents or carers did not provide consent to participate in any aspect of the study. All other children (N = 2,221; 1064 in the schools that were randomised to intervention and 1157 in those randomised to control schools), irrespective of whether or not we have all the data for them, are included in the analyses presented here (with numbers differing for each outcome in the main analyses as a result of some missing data). Proportions with data for each outcome were similar in intervention and control schools at both baseline and at the second follow-up assessment at 12 months post-intervention (**Figure 1**). Baseline characteristics were similar between children in intervention schools and those in control schools (**Table 1**).

Figures 2a to 2l shows differences in means or odds ratios between the control and intervention group for the three primary and nine secondary outcomes at baseline, immediate follow-up and long-term (12-months) follow-up. These show that differences in means (and

odds ratios for general and central overweight/obesity) between children in intervention and control schools were essentially the same at this long-term follow-up as they were immediately after the intervention, when examining point estimates. Differences in the primary outcomes were consistent with the null hypothesis (**Figures 2a to 2c**). Differences in secondary outcomes were consistent with those seen at the end of the immediate follow-up,(**Figures 2d to 2l**) with no evidence that the previously reported beneficial effects for child-reported screen viewing at weekends,(**Figure 2e**) consumption of snacks (**Figure 2f**) and consumption of high energy drinks (**Figure 2h**) had notably diminished (or increased) in magnitude over time (Figures 2. However, there was no strong statistical support for any effect of the intervention on primary and secondary outcomes at 12 months after the intervention. **Table 2** shows differences in means or odds ratios for all outcomes at the long-term follow-up from the main intention-to-treat analyses. None of the three primary outcomes differed, nor the nine secondary outcomes, reached our predefined level of statistical significance for an effect after accounting for multiple testing.

Results from the per-protocol analyses were consistent with the intention-to-treat analyses results (**Table 3**). Results were similar in all sensitivity analyses applying different assumptions about missing data (**Supplementary Tables S1-S4**). Results were also similar when we looked separately at time spent in MVPA and time spent in sedentary behaviour by weekday and weekend (**Supplementary Table S5**).

DISCUSSION

In this school-based cluster RCT, aimed at increasing physical activity, reducing sedentary behaviours and improving diet in school aged children, we found results at 12 months after the intervention had ended (that is, with no further lessons or teaching aimed at promoting

healthy activity and dietary levels during that 12 months) were essentially the same as those seen immediately after the end of the intervention in terms of size of effect. The lack of any effect on the three primary outcomes – time spent in MVPA, time spent in sedentary behaviour and fruit and vegetable consumption – was still observed 12 months later and the beneficial effects on three secondary outcomes (reported screen-viewing at weekends, consumption of snacks and of high energy drinks) were still somewhat present at 12-months post intervention. However, slight attenuation of the effect on these secondary outcomes meant that at this long-term follow-up none of our outcomes (primary or secondary) reached our pre-specified level of statistical significance.

Meaning of study findings

Whilst the effects for these secondary outcomes were consistent in magnitude with those seen at the immediate follow-up, they did not reach our predefined level of statistical significance. Thus, these results suggest that apparent benefits on these secondary outcomes are due to chance.

As discussed in our previous publication of effects immediately at the end of the intervention,(18) the lack of effect on primary outcomes, in particular on the objectively assessed accelerometer outcomes, might highlight the importance of societal and structural changes to support greater levels of activity, over and above any intervention at a school level.(18) Our intervention was based on theory,(22, 23) built on a similar intervention that had been previously shown to work in the US(21) and in pilot work, conducted by us, it was shown to fit well with the primary school national curriculum in the UK.(27) Furthermore, the detailed process evaluation conducted as part of the full AFLY5 RCT, in which we used quantitative measures of intervention delivery and qualitative focus groups with children and

in-depth interviews with teachers and parents,(28), showed that on average 77% of the intervention lessons and homeworks were delivered and reached 95% of the children in intervention schools. However, teachers felt lack of time and the need to prioritise numeracy and literacy skills over the health promoting lessons of our intervention were important barriers to them and the children being more fully engaged with AFLY5.(28) The process evaluation also highlighted that in general teachers did not like teaching physical activity, and had a tendency to delegate such lessons to teaching assistants. This might also have contributed to the null effects, particularly for the activity outcomes. Lastly, our process evaluation suggests that in the context of rapidly developing technologies the time taken to develop, test the feasibility of, and pilot, school-based interventions before completing large scale RCTs, as we have done in AFLY5, may mean that by the time school-based interventions get to the full scale RCT, the intervention is being implemented with out-of-date methods of delivery.(28, 29) Thus, whilst using schools for universal promotion of healthy behaviours is appealing, it may be that greater resources and support within schools, and wider engagement of the whole community, is necessary to achieve major shifts towards healthier behaviours.

Strengths and limitations

The study was designed to take account of known sources of bias in other RCTs in this area. A protocol was published before recruitment started, and a detailed analysis plan was written before any access to the study data. We developed an intervention according to guidelines for complex interventions, with the theoretical rationale for the intervention described in detail elsewhere.(18) Our sample size calculation, which took account of the likely degree of clustering within schools, indicated that we needed a total of 1275 children to be included in the analyses. For all outcomes, except those related to accelerometer data, we achieved

considerably higher numbers than this target. The number included in the main analyses for accelerometer based data was somewhat smaller than this at 1066. Sample size calculations are an approximation of the numbers needed, and we doubt that such a small difference will have had a major effect on our conclusions. Furthermore, wear time was similar in children in intervention and control schools; moreover, in sensitivity analyses using different approaches to dealing with missing data and which included 2052 children even for the accelerometer outcomes, the results were essentially the same as in the main analysis. One school refused to deliver any of the intervention, and others did not deliver all of the lessons. However, the per-protocol analysis, which did not differ from the main intention-to-treat analysis, shows that this does not explain the null results.

Conclusion

This long-term follow-up of a large well-conducted school based RCT has found similar results to those found immediately after the intervention period. None of the primary or secondary outcomes reached our predefined levels of statistical significance, suggesting that apparent benefits on some secondary outcomes are due to chance. Overall, together with our process evaluation these findings suggest that curriculum-based interventions alone are unlikely to make a major impact on promoting healthy levels of physical activity and healthy diets in primary school children.

Acknowledgements: We thank all the students and teaching staff who took part in AFLY5. We thank all of the AFLY5 staff who include fieldworkers, administrative staff, computing and data management staff and the trainers who provided teacher training. We thank Dr Hugh Annett (retired Director of Public Health, NHS Bristol and Bristol City Council), Annie Hudson (former Strategic Director for Children, Young People and Skills, Bristol City

Council) and Sheila Smith (Strategic Director for Children, Young People and Skills, North Somerset City Council) for their support of the Active for Life Year 5 study. We also thank the Chair and members of the trial steering committee for their advice and support.

The views expressed in this paper are those of the authors and not necessarily anyone in this acknowledgement list.

Funding: The AFLY5 RCT is funded by the UK National Institute for Health Research (NIHR) Public Health Research Programme (09/3005/04), which also paid the salary of ELA and SW. DAL and LDH work in a Unit that receives funds from UK Medical Research Council (MC_UU_12013/5). RRK and RC work in the Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research Centre of Excellence: joint funding (MR/KO232331/1) from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the Welsh Government and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. LDH is supported by a UK Medical Research Council Population Health Scientist fellowship (G1002375). DAL (NF-SI-0611-10196) and TJP (NF-SI-0512-10026) are NIHR Senior Investigators. This study was undertaken in collaboration with the Bristol Randomised Trials Collaboration (BRTC), a UKCRC Registered Clinical Trials Unit in receipt of National Institute for Health Research CTU support funding.

None of the funders had involvement in the Trial Steering Committee, the data analysis, data interpretation, data collection, or writing of the paper. DAL, LDH and ELA had access to all of the data in the study and DAL had the final responsibility for the decision to submit for publication.

The views expressed in this publication are those of the authors and not necessarily any of the funding bodies listed here.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: support from research funders in accordance with the funding statement included in this manuscript; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work, other than that RC is director of DECIPHer Impact, a not-for-profit company that is wholly owned by the Universities of Bristol and Cardiff and whose purpose is to licence and support the implementation of evidenced based health promotion interventions.

Ethical approval: Ethical approval was obtained from the University of Bristol Faculty of Medicine and Dentistry Committee for Ethics (reference number 101115).

Transparency declaration: DAL affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Data sharing: Full details of the study are available on the website (www.bristol.ac.uk/social-community-medicine/projects/afl/). We encourage anyone who would like to access these data for other projects to contact the corresponding author. We would be happy for external collaborators to access these data according to our data transfer agreement.

Contributors: DAL, together with RRK, RC, TJP, SN and RJ designed the AFLY5 study and obtained funds to complete it. DAL wrote the analysis plan used for this paper and LDH and ELA completed all analyses. SW managed the AFLY5 study, including managing data collection. ELA, LDH and DAL wrote the first draft of the paper, and DAL coordinated contributions from other co-authors. All authors contributed to the overall study aim and development of the design. All authors made critical comments on drafts of the paper.

Funding: This trial was funded by UK National Institute for Health Research (NIHR) Public Health Research Programme (09/3005/04). The funders had no role in the study design, data collection, analysis or interpretation of results.

References

1. Lock K, Pomerleau J, Causer L, Altmann DR, McKee M. The global burden of disease attributable to low consumption of fruit and vegetables: implications for the global strategy on diet. Bull World Health Organ. 2005;83(2):100-8. Epub 2005/03/04.

2. Maynard M, Gunnell D, Emmett P, Frankel S, Davey Smith G. Fruit, vegetables, and antioxidants in childhood and risk of adult cancer: the Boyd Orr cohort. J Epidemiol Community Health. 2003;57(3):218-25. Epub 2003/02/21.

3. Ness AR, Maynard M, Frankel S, Smith GD, Frobisher C, Leary SD, et al. Diet in childhood and adult cardiovascular and all cause mortality: the Boyd Orr cohort. Heart. 2005;91(7):894-8. Epub 2005/06/17.

4. Boreham C, Riddoch C. The physical activity, fitness and health of children. J Sports Sci. 2001;19(12):915-29.

5. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act. 2010;7:40. Epub 2010/05/13.

6. Ness AR, Leary SD, Mattocks C, Blair SN, Reilly JJ, Wells J, et al. Objectively measured physical activity and fat mass in a large cohort of children. PLoS Med. 2007;4(3):e97. Epub 2007/03/29.

7. Ekelund U, Luan J, Sherar LB, Esliger DW, Griew P, Cooper A, et al. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. JAMA. 2012;307(7):704-12. Epub 2012/02/18.

8. Cooper AR, Goodman A, Page AS, Sherar LB, Esliger DW, van Sluijs EM, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). Int J Behav Nutr Phys Act. 2015;12:113. Epub 2015/09/18.

9. Dobbins M, Husson H, DeCorby K, LaRocca RL. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. Cochrane Database Syst Rev. 2013;2:CD007651. Epub 2013/03/02.

10. Mallam KM, Metcalf BS, Kirkby J, Voss LD, Wilkin TJ. Contribution of timetabled physical education to total physical activity in primary school children: cross sectional study. BMJ. 2003;327:592-3.

11. Cleland V, Dwyer T, Blizzard L, Venn A. The provision of compulsory school physical activity: associations with physical activity, fitness and overweight in childhood and twenty years later. Int J Behav Nutr Phys Act. 2008;5:14

12. Metcalf B, Henley W, Wilkin T. Effectiveness of intervention on physical activity of children: systematic review and meta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). BMJ. 2012;345:e5888. Epub 2012/10/10.

13. DeMattia L, Lemont L, Meurer L. Do interventions to limit sedentary behaviours change behaviour and reduce childhood obesity? A critical review of the literature. Obes Rev. 2007;8(1):69-81.

14. van Grieken A, Ezendam NP, Paulis WD, van der Wouden JC, Raat H. Primary prevention of overweight in children and adolescents: a meta-analysis of the effectiveness of interventions aiming to decrease sedentary behaviour. Int J Behav Nutr Phys Act. 2012;9:61. Epub 2012/05/30.

15. Delgado-Noguera M, Tort S, Martinez-Zapata MJ, Bonfill X. Primary school interventions to promote fruit and vegetable consumption: a systematic review and meta-analysis. Prev Med. 2011;53(1-2):3-9. Epub 2011/05/24.

16. Evans CE, Christian MS, Cleghorn CL, Greenwood DC, Cade JE. Systematic review and meta-analysis of school-based interventions to improve daily fruit and vegetable intake in children aged 5 to 12 y. AM J CLIN NUTR. 2012;96(4):889-901. Epub 2012/09/07.

17. Lawlor DA, Jago R, Noble SM, Chittleborough CR, Campbell R, Mytton J, et al. The Active for Life Year 5 (AFLY5) school based cluster randomised controlled trial: study protocol for a randomized controlled trial. Trials. 2011;12:181. Epub 2011/07/26.

18. 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.

19. Lawlor DA, Peters TJ, Howe LD, Noble SM, Kipping RR, Jago R. The Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial protocol detailed statistical analysis plan. Trials. 2013;14(1):234. Epub 2013/07/26.

20. Department for Communities and Local Government. The English Indices of Deprivation 2010. London: 2011.

21. Gortmaker SL, Cheung LW, Peterson KE, Chomitz G, Cradle JH, Dart H, et al. Impact of a school-based interdisciplinary intervention on diet and physical activity among urban primary school children: eat well and keep moving. Arch Pediatr Adolesc Med. 1999;153:975-83

22. Bandura A: Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall; 1986

23. Bandura A: Self-efficacy: The exercise of control. New York: Freeman; 1997.

24. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clinical trials. 2012;9(4):396-407. Epub 2012/07/04.

25. 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. Epub 2011/02/09.

26. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med. 2005;24(7):993-1007. Epub 2004/12/01.

27. Kipping RR, Payne C, Lawlor DA. Randomised controlled trial adapting US school obesity prevention to England. Archives of Disease in Childhood. 2008;93(6):469-73.

28. Campbell R RE, Wells S, Kipping RR, Chittleborough CR, Peters TJ, Lawlor DA, Jago R. Intervention fidelity in a school-based diet and physical activity intervention in the UK: Active for Life Year 5. International Journal of Behavioral Nutrition and Physical Activity. 2015;12: 141.

29. Jago R RE, Kipping RR, Wells S, Chittleborough CR, Peters TJ, Mytton J, Lawlor DA, Campbell R. Lessons learned from the AFLY5 RCT process evaluation: Implications for the design of physical activity and nutrition interventions in schools. BMC Public Health. 2015;15:946.

Box 1: AFLY5 primary and secondary outcomes

Primary outcomes

Accelerometer assessed mean time per day spent doing moderate/vigorous physical activity MVPA (minutes per day)

Accelerometer assessed mean time per day spent in sedentary activity (minutes per day)

Self-reported (validated questionnaire) servings of fruit and vegetables consumed per day (servings per day; treated in all analyses as a continuous variable)

Secondary outcomes

Self-reported (validated questionnaire) mean time spent screen viewing on a typical weekday (minutes)

Self-reported (validated questionnaire) mean time spent screen viewing on a typical weekend day (minutes)

Self-reported (validated questionnaire) servings of snacks consumed per day (servings per day; treated in all analyses as a continuous variable)

Self-reported (validated questionnaire) servings of high fat foods consumed per day (servings per day; treated in all analyses as a continuous variable)

Self-reported (validated questionnaire) servings of high energy drinks consumed per day (servings per day; treated in all analyses as a continuous variable)

Body mass index determined from weight and height measured in classrooms by two study fieldworkers (kg/m^2; treated in all analyses as a standard deviation z-score)

Waist circumference measured in classrooms by two study fieldworkers (mm; treated in all analyses as a standard deviation z-score)

General overweight/obesity, determined by the International Obesity Task Force thresholds of body mass index for children (taking account of their age and sex) (binary outcome)

Central overweight/obesity determined by thresholds of UK age and sex specific reference charts for waist circumference and defined by the International Diabetes Federation. (binary outcome)

Figure legends

Figure 1 – Trial profile.

Footnote to Figure 1

Np=number of participants (school pupils). No schools withdrew from study, so all randomised units are present at baseline and at both follow-up assessments. Percentages for proportions of children with each measurement at baseline and at follow-ups are of total number of children who were pupils in randomised schools at baseline. Not all pupils with follow-up measures necessarily had data on the same measure at baseline (or vice versa), because of different pupils being absent at baseline and follow-up assessments at each time point, and because of pupils leaving or moving between schools. In all analyses, study participants were analysed in the group (intervention or control) to which they were randomised.

Figure 2: Difference in means and odds ratios for the intervention compared to the control group for the three primary outcomes and nine secondary outcomes, assessed at baseline, first follow-up (conducted immediately after the end of the intervention) and second follow-up (12-months post-intervention).

a. Accelerometer assessed time spent in moderate to vigorous physical activity

b. Time spent in sedentary behaviour

c. Servings of Fruit and Vegetables per day

d. Time spent screen viewing on weekdays

e. Time spent screen viewing on Saturdays

f. Servings of snacks per day

g. Servings of high fat foods per day

h. Servings of high energy drinks per day

i. Body mass index z-score (as a continuous variable)

j. Waist circumference z-score (as a continuous variable)

k. General overweight or obesity (based on BMI measurements)

l. Central overweight / obesity based on waist circumference measurements

Footnote to Figure 2

The figures all show differences in means for continuous variables (graphs a to j) and odds ratios for binary outcomes (graphs k and l), comparing those in the intervention arm of the trial to those in the control arm (dots), together with 95% confidence intervals (vertical lines with horizontal caps representing the limits). The dashed horizontal lines represent the null values (zero for all differences in means of continuous variables and one for odds ratios of binary outcomes).

Table 1: Comparison of baseline characteristics by randomised group

Characteristic	Unit and type of summary measure	Intervention schools N participants=1064 N schools = 30		Control schools N participants=1157 N schools = 30	
		Number	Distribution	Number	Distribution
Age	Mean (SD) years	1024	9.5 (0.3)	1099	9.5 (0.3)
MVPA[a]	Mean (SD) minutes	912	59 (23)	928	56 (21)
Sedentary behaviour[a]	Mean (SD) minutes	912	422 (72)	928	416 (68)
Servings of fruit and vegetables	Median (IQR) number / day	1019	1 (0 to 2)	1088	1 (0 to 2)
Servings of snacks	Median (IQR) number / day	1019	2 (1 to 3)	1088	2 (1 to 3)
Servings of high fat foods	Median (IQR) number / day	1019	0 (0 to 1)	1088	1 (0 to 1)
Servings of high energy drinks	Median (IQR) number / day	1019	2 (1 to 3)	1088	2 (1 to 3)
BMI	Mean (SD) z-score	889	-0.06 (0.94)	953	0.05 (1.04)
WC	Mean (SD) z-score	942	-0.03 (0.97)	1027	0.03 (1.02)
Screen-viewing weekday	Median (IQR) minutes	1024	105 (45 to 240)	1099	105 (45 to 225)
Screen-viewing Saturday	Median (IQR) minutes	1024	90 (30 to 240)	1099	105 (30 to 240)
Total number of valid days of wearing accelerometer[b]	Median (IQR) days	912	3 (2 to 5)	928	3 (2 to 4)
Total number of valid weekdays of wearing accelerometer[b]	Median (IQR) days	979	2 (2 to 3)	1025	2 (1 to 3)
Total hours of wearing accelerometer on valid days[a]	Mean (SD) hours / day	912	11.6 (1.5)	928	11.5 (1.4)
Hours of wearing accelerometer on	Mean (SD) hours / day	896	11.8 (1.6)	919	11.7 (1.5)

valid weekdays[b]					
Categorical variables					
Gender	N (%) female	520	49%	608	52%
	N (%) male	544	51%	549	48%
General overweight / obesity	N (%) No	717	81%	743	78%
	N (%) Yes	172	19%	210	22%
Central overweight/obesity	N (%) No	601	64%	631	61%
	N (%) Yes	341	36%	396	39%
Returned accelerometer	N (%) No	85	8%	132	11%
	N (%) Yes	979	92%	1025	89%
Wore accelerometer for requested amount of time	N (%) No	820	77 %	953	82%
	N (%) Yes	244	23%	204	18%
Wore accelerometer for required amount of time	N (%) No	418	39%	514	44%
	N (%) Yes	646	61%	643	56%
School involved in other health promoting activities	N (%) No	264	25%	446	39%
	N (%) Yes	800	75%	711	61%
School deprivation score	N (%) low	315	30%	460	40%
	N (%) medium	368	35%	345	30%
	N (%) high	381	36%	352	30%

SD: standard deviation; MVPA: moderate or vigorous physical activity; IQR: interquartile range; BMI: body mass index; WC: waist circumference

[a]Including only participants with at least 3 days of valid data

[b]Including all valid days, regardless of the number of valid days

Note some % within categories do not sum to exactly 100 because of rounding

Table 2: Main intention-to-treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. Numbers of participants vary by outcome as indicated in the table.

Outcome (primary/secondary)	Control group (reference group)		Intervention group		Main comparison between the two groups (Intervention versus Control)		
	Np	Mean (SD) or number (%)	Np	Mean (SD) or number (%)	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes:							
Time spent in MVPA (minutes per day)	522	52.56 (20.67)	527	54.37 (22.23)	1049	2.48 (-1.80, 6.77)	0.26
Time spent in sedentary behaviour (minutes per day)	522	461.78 (66.33)	527	465.46 (70.61)	1049	2.79 (-7.78, 13.37)	0.60
Servings of fruit and vegetables (number per day)	1062	1.80 (1.55)	990	1.82 (1.59)	2052	0.01 (-06, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	1062	148.01 (126.39)	990	138.88 (125.00)	2052	-10.74 (-26.30, 4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	1062	180.52 (164.82)	990	167.71 (156.28)	2052	-16.03 (-32.82, 0.73)	0.06
Body mass index (z-score)	923	0.03 (1.02)	870	-0.03 (0.97)	1793	0.01 (-0.04, 0.06)	0.72
Waist circumference (z-score)	993	0.03 (1.04)	935	-0.03 (0.95)	1928	-0.04 (-0.13, 0.05)	0.36
Servings of snacks (number per day)	1062	2.11 (1.55)	990	1.99 (1.47)	2052	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	1062	0.86 (0.94)	990	0.74 (1.07)	2052	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	1062	2.38 (1.58)	990	2.19 (1.45)	2052	-0.20 (-0.39, -0.01)	0.04
Binary outcomes							
Generally overweight/obese	923	194 (21.02)	870	175 (20.11)	1793	1.00 (0.72, 1.37)	0.98
Centrally overweight/obese	993	421 (42.40)	935	394 (42.14)	1928	1.08 (0.80, 1.46)	0.62

Np: number of participants; SD: standard deviation; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance after taking account of multiple testing)

All differences in means / odds ratios with their 95% CIs have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline/school stratifying variables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses participants were included for each outcome if they had a follow-up measurement of that outcome; for missing baseline data we used an indicator variable as described by White & Thompson,(21) which means for each outcome participants are included even if they do not have a baseline measurement.

Table 3: Per-protocol analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. Numbers vary by outcome as indicated in the table.

Outcome (**primary**/secondary)	**Control group (reference group)**		**Intervention group**		**Main comparison between the two groups (Intervention versus Control)**		
	Np	**Mean (SD) or number (%)**	**Np**	**Mean (SD) or number (%)**	**Np**	**Difference in means or odds ratio (95%CI)**	**p-value**
Continuous outcomes							
Time spent in MVPA (minutes per day)	522	52.56 (20.67)	356	54.15 (22.27)	878	2.63 (-2.10, 7.37)	0.28
Time spent in sedentary behaviour (minutes per day)	522	461.78 (66.33)	356	466.17 (70.58)	878	3.67 (-8.32, 15.66)	0.55
Servings of fruit and vegetables (number per day)	1062	1.80 (1.55)	701	1.91 (1.66)	1762	0.05 (-0.15, 0.25)	0.63
Time spent screen-viewing (minutes per day weekday)	1062	148.01 (126.39)	701	134.98 (120.94)	1762	-8.97 (-26.81, 8.87)	0.32
Time spent screen-viewing (minutes per day Saturday)	1062	180.52 (164.82)	701	159.35 (149.97)	1762	-21.73 (-41.19, -2.26)	0.03
Body mass index (z-score)	923	0.03 (1.02)	612	-0.03 (0.98)	1535	0.01 (-0.05, 0.07)	0.69
Waist circumference (z-score)	993	0.03 (1.04)	657	-0.04 (0.94)	1650	-0.03 (-0.13, 0.06)	0.52
Servings of snacks (number per day)	1062	2.11 (1.55)	701	2.07 (1.48)	1762	-0.03 (-0.23, 0.16)	0.72
Servings of high fat foods (number per day)	1062	0.86 (0.94)	701	0.75 (1.15)	1762	-0.11 (-0.26, 0.04)	0.14
Servings of high energy drinks (number per day)	1062	2.38 (1.58)	701	2.22 (1.43)	1762	-0.18 (-0.41, 0.5)	0.12
Binary outcomes							
Generally overweight/obese	923	194 (21.02)	612	121 (19.77)	1535	0.98 (0.68, 1.41)	0.91
Centrally overweight/obese	993	421 (42.40)	657	272 (41.40)	1650	1.06 (0.76, 1.49)	0.72

Np: number of participants; SD: standard deviation; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Per-protocol analysis defined as teaching at least 70% (11 out of the 16) AFLY5 lessons. All participants from the intervention schools where the teacher taught fewer than 11 lessons are excluded from these analyses.

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance after taking account of multiple testing)

All differences in means/odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline/school stratifying variables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses, after removal of schools that did not teach at least 11 out of 16 of the lessons, participants were only included for each outcome if they had a follow-up measurement of that outcome. For partial missing baseline data we used an indicator variable as described by White & Thompson,(21) which means for each outcome participants are included even if they do not have a baseline measurement.

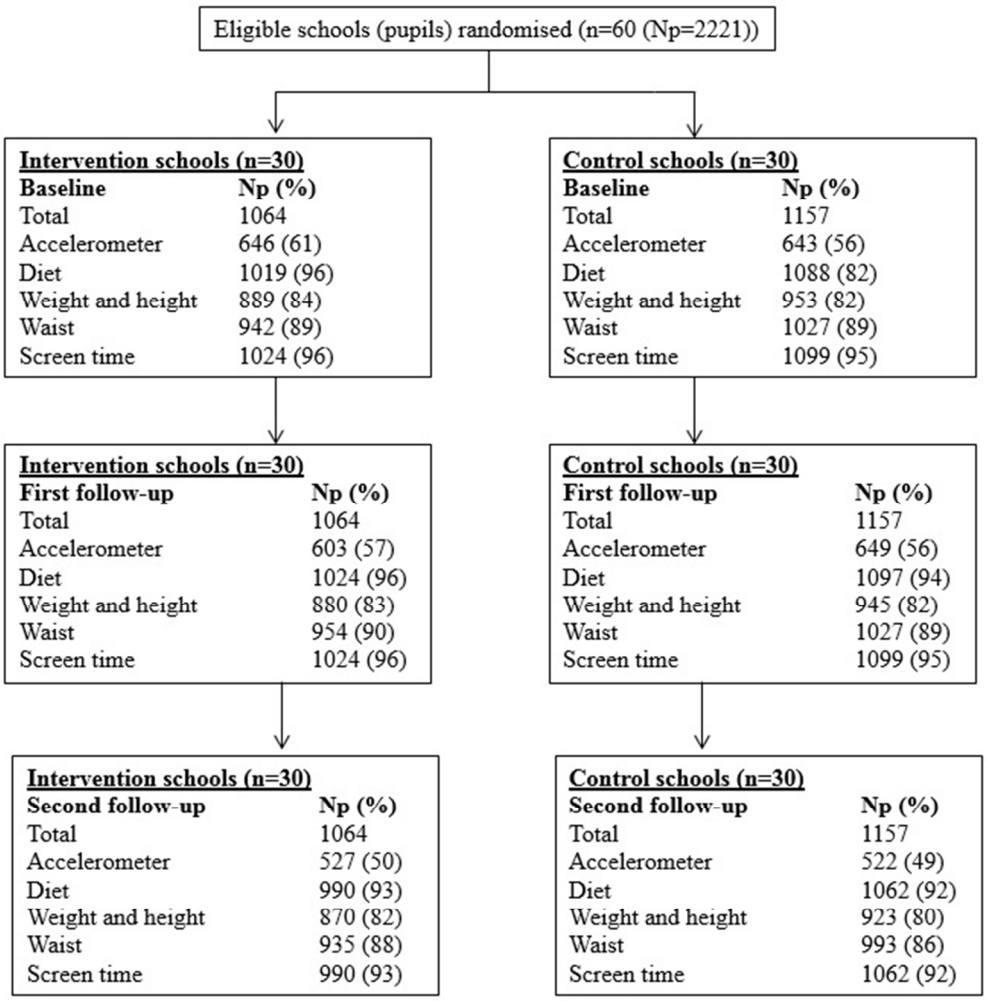


Figure 1 - Trial Profile
Footnote to the figure provide
55x56mm (300 x 300 DPI)

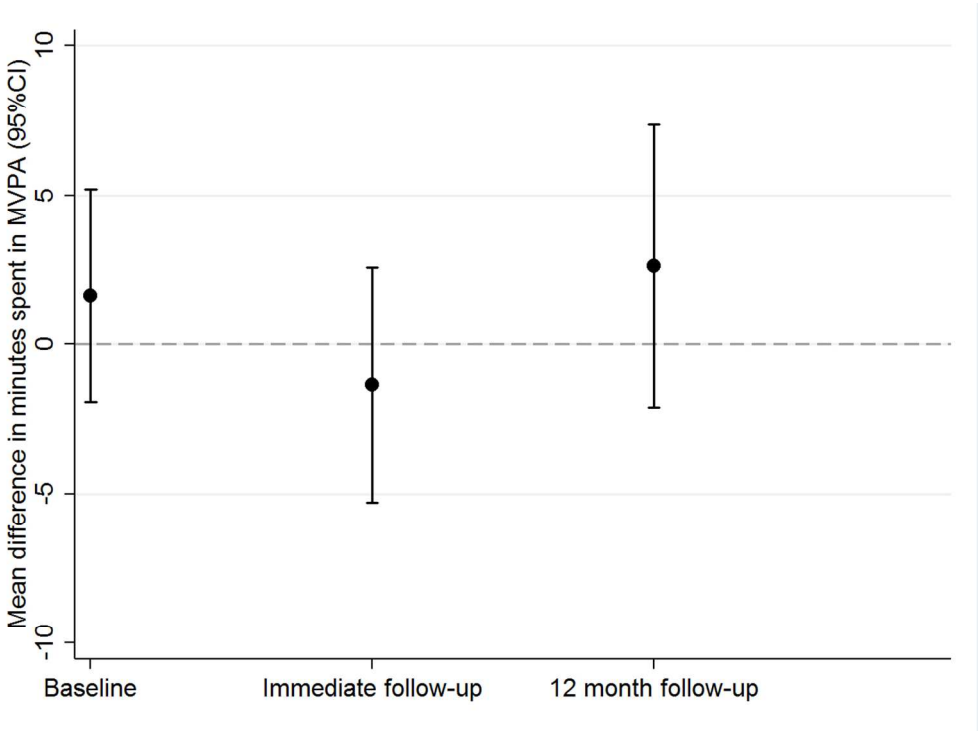
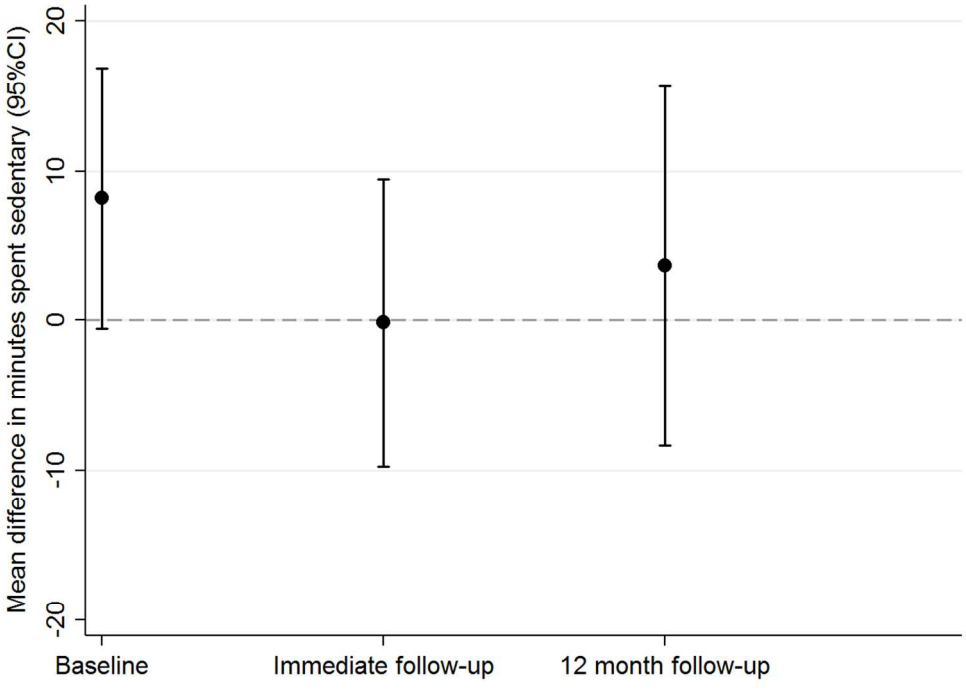


Figure 2: Difference in means and odds ratios for the intervention compared to the control group for the three primary outcomes and nine secondary outcomes, assessed at baseline, first follow-up (conducted immediately after the end of the intervention) and second follow-up (12-months post-intervention).

Specifically for this figure
Figure 2 a. Accelerometer assessed time spent in moderate to vigorous physical activity
Footnote in main document for
108x79mm (300 x 300 DPI)



b. Time spent in sedentary behaviour

108x78mm (300 x 300 DPI)

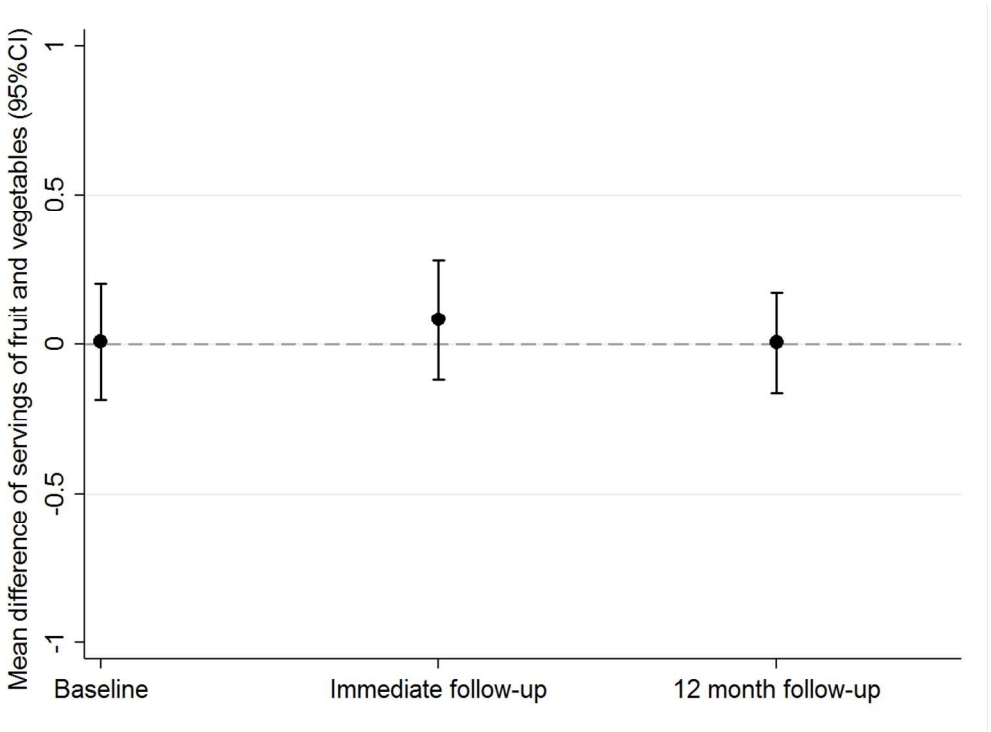


Figure 2 c. Servings of Fruit and Vegetables per day

113x82mm (300 x 300 DPI)

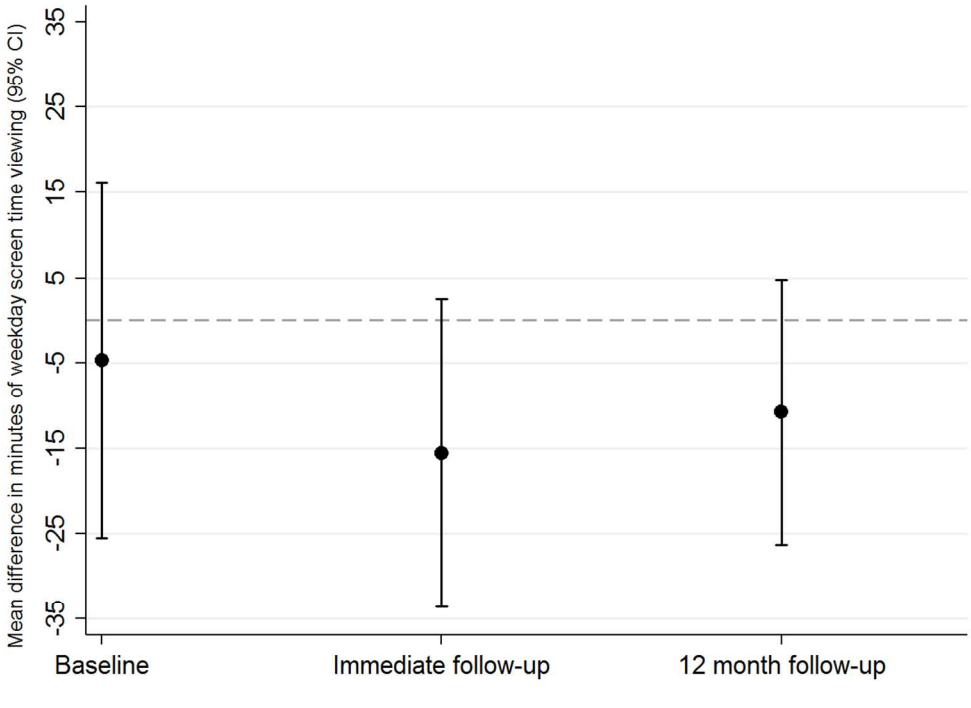


Figure 2 d. Time spent screen viewing on weekdays

113x82mm (300 x 300 DPI)

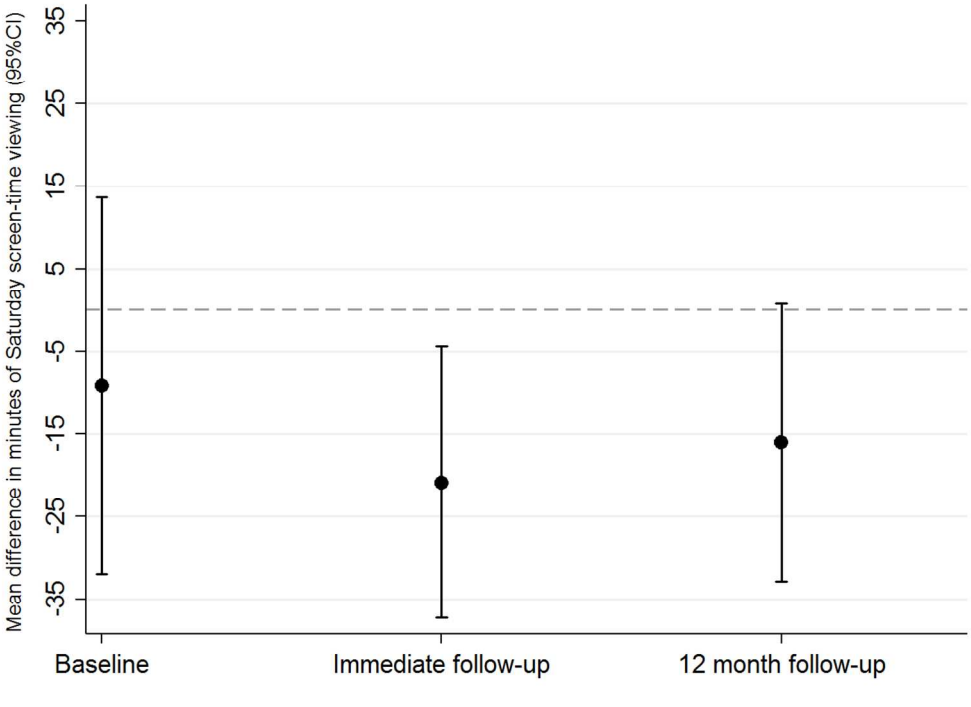


Figure 2 e. Time spent screen viewing on Saturdays

113x82mm (300 x 300 DPI)

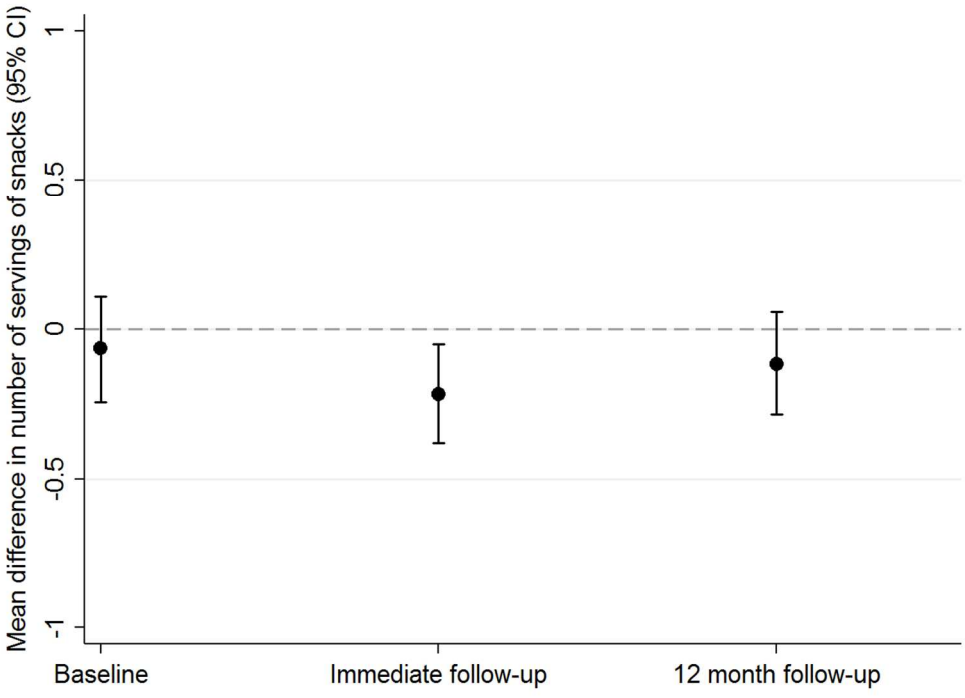


Figure 2 f. Servings of snacks per day

113x82mm (300 x 300 DPI)

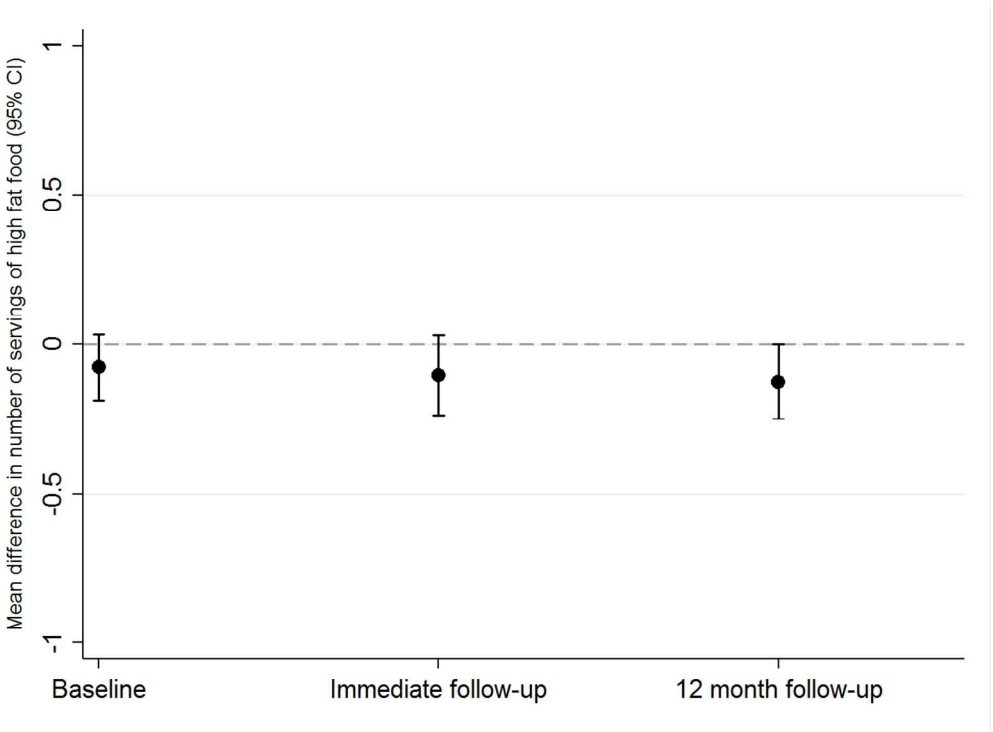


Figure 2 g. Servings of high fat foods per day

113x82mm (300 x 300 DPI)

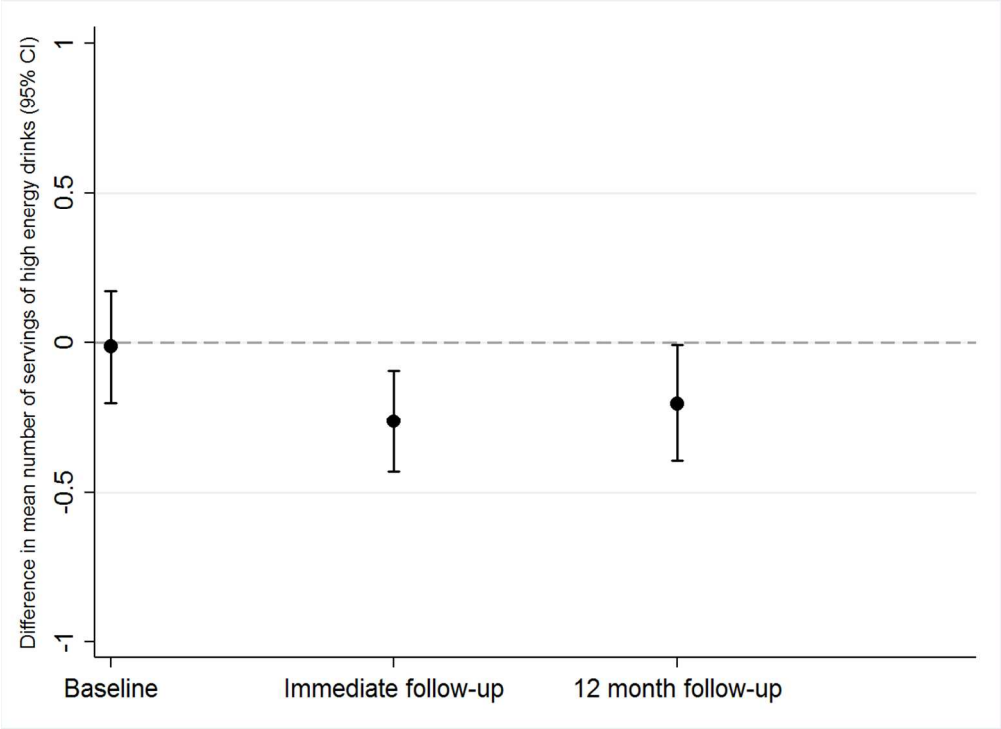


Figure 2 h. Servings of high energy drinks per day

109x79mm (300 x 300 DPI)

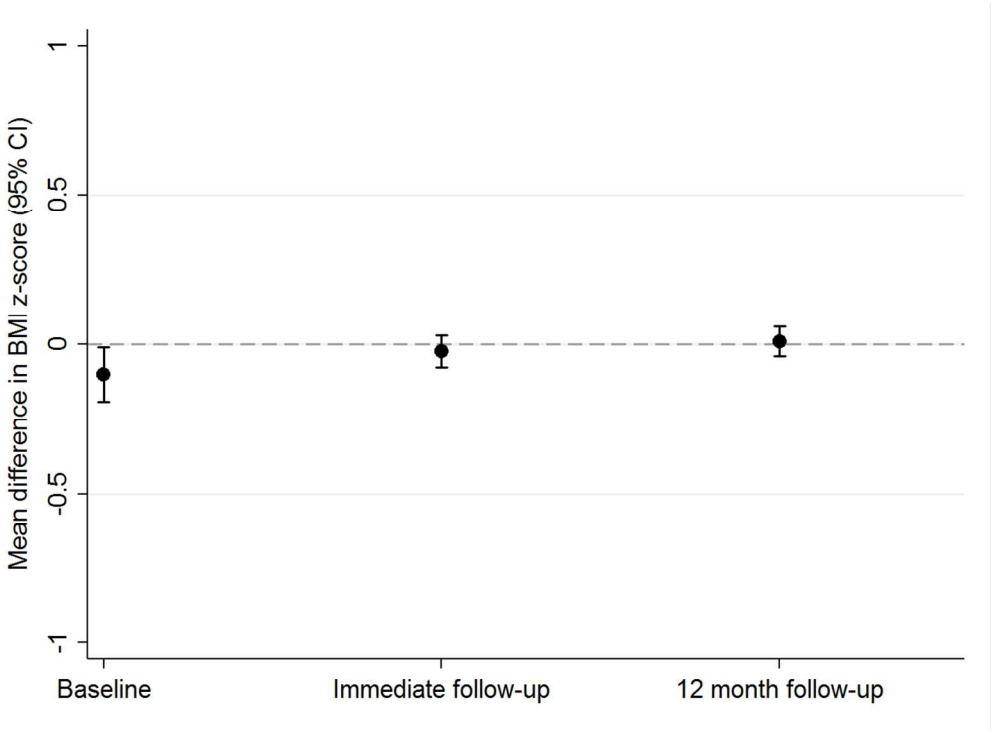


Figure 2 i. Body mass index z-score (as a continuous variable)

113x82mm (300 x 300 DPI)

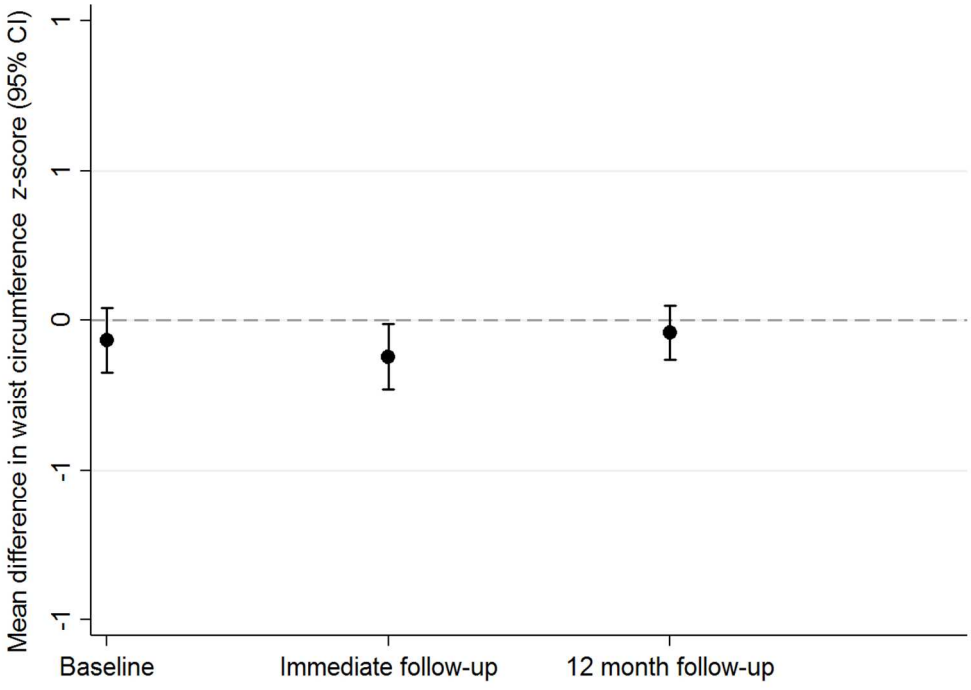


Figure 2 j. Waist circumference z-score (as a continuous variable)

113x82mm (300 x 300 DPI)

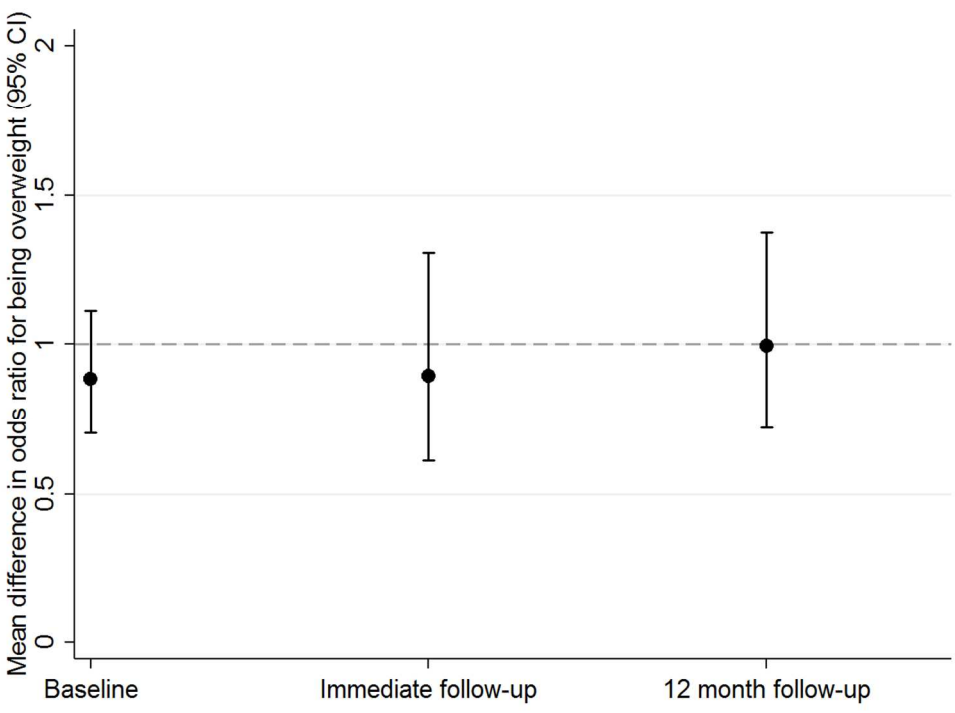


Figure 2 k. General overweight or obesity (based on BMI measurements)

113x82mm (300 x 300 DPI)

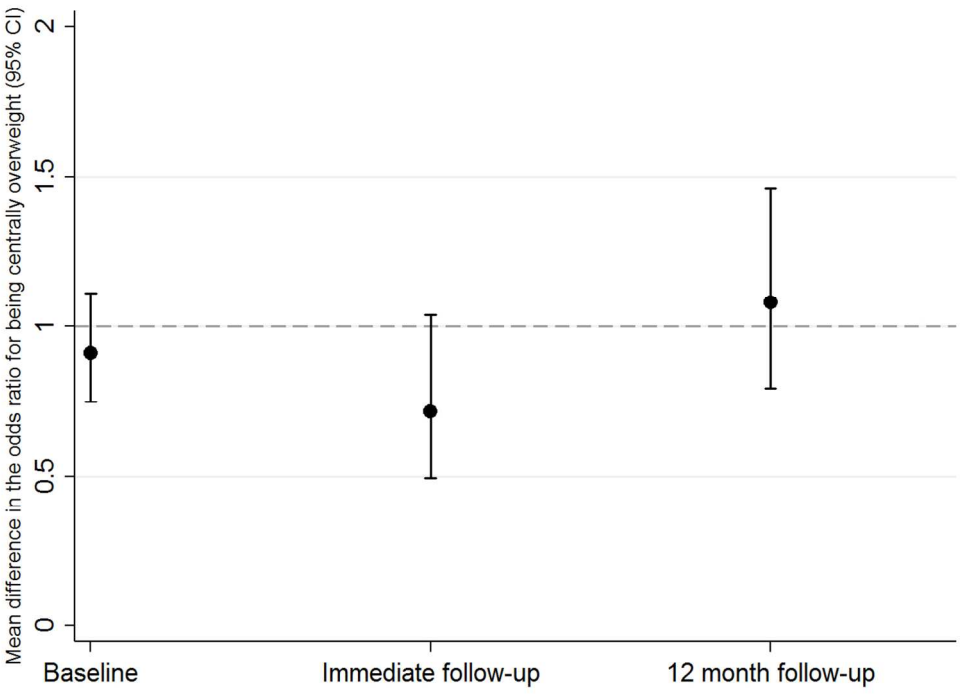


Figure 2 I. Central overweight / obesity based on waist circumference measurements
Footnote to all parts of Figur
113x82mm (300 x 300 DPI)

Supplementary Table S1: Sensitivity analysis: intention to treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes 12 months post-intervention. Numbers vary by outcome as indicated in the table. In these analyses participants were only included for each outcome if they had a baseline and a follow-up measurement of that outcome.

Outcome	Main comparison between the two groups (Intervention versus Control)		
Primary / secondary	**Np**	**Difference in means or odds ratio (95%CI)**	**p-value**
Continuous outcomes			
Time spent in MVPA (minutes per day)	1000	3.05 (-1.33, 7.44)	0.17
Time spent in sedentary behaviour (minutes per day)	1000	2.21 (-8.28, 12.71)	0.68
Servings of fruit and vegetables (number per day)	1953	0.02 (-0.15, 0.19)	0.83
Time spent screen-viewing (minutes per day weekday)	1965	-10.53 (-26.1, 5.05)	0.19
Time spent screen-viewing (minutes per day Saturday)	1965	-17.3 (-33.71, -0.88)	0.04
Body mass index (z(sd)-score)	1563	0 (-0.05, 0.04)	0.95
Waist circumference (z(sd)-score)	1748	-0.03 (-0.12, 0.05)	0.47
Servings of snacks (number per day)	1953	-0.13 (-0.3, 0.04)	0.13
Servings of high fat foods (number per day)	1953	-0.13 (-0.25, 0)	0.04
Servings of high energy drinks (number per day)	1953	-0.21 (-0.4, -0.02)	0.03
Binary outcomes			
Generally overweight/obese	1563	0.83 (0.56, 1.22)	0.35
Centrally overweight/obese	1748	1.01 (0.73, 1.4)	0.93

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed), BMI: body mass index, WC: waist circumference, F&V fruit and vegetables.

Supplementary Table S2: Sensitivity analysis: intention to treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. In these analyses participants were only included for each outcome if they had a baseline and a follow-up measurement for all three primary outcomes. Numbers included are identical for the three primary outcomes (N = 757) but can vary by outcome for secondary outcomes (though none of these can be higher than 757) as indicated in the table.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	757	1.28 (-3.22, 5.78)	0.58
Time spent in sedentary behaviour (minutes per day)	757	0.60 (-10.44, 11.63)	0.92
Servings of fruit and vegetables (number per day)	757	-0.13 (-0.34, 0.09)	0.26
Time spent screen-viewing (minutes per day weekday)	757	0.20 (-17.54, 17.94)	0.98
Time spent screen-viewing (minutes per day Saturday)	757	-8.46 (-28.49, 1.56)	0.41
Body mass index (z(sd)-score)	682	0.00 (-0.06, 0.07)	0.80
Waist circumference (z(sd)-score)	728	-0.01 (-0.12, 0.09)	0.90
Servings of snacks (number per day)	757	-0.13 (-0.38, 0.13)	0.33
Servings of high fat foods (number per day)	757	-0.13 (-0.33, 0.07)	0.19
Servings of high energy drinks (number per day)	757	-0.12 (-0.37, 0.12)	0.32
Binary outcomes			
Generally overweight/obese	680	1.09 (0.64, 1.85)	0.76
Centrally overweight/obese	728	11.35 (0.81, 2.23)	0.25

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed), BMI: body mass index, WC: waist circumference, F&V fruit and vegetables.
Missing baseline data for secondary outcomes (once those with missing baseline primary outcomes are excluded) were managed as in the main analyses.

Supplementary Table S3: Sensitivity analysis: intention to treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention, with missing data for either baseline or follow-up measure of an outcome assumed to be 10% healthier than the average value in the study sample.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	2052	0.74 (-1.59, 3.07)	0.53
Time spent in sedentary behaviour (minutes per day)	2052	1.78 (-4.63, 8.20)	0.59
Servings of fruit and vegetables (number per day)	2052	0.01 (-0.16, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	2052	-10.74 (-26.30, 4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	2052	-16.03 (-32.82, 0.76)	0.06
Body mass index (z(sd)-score)	2052	0.01 (-0.04, 0.06)	0.70
Waist circumference (z(sd)-score)	2052	-0.02 (-0.11, 0.06)	0.56
Servings of snacks (number per day)	2052	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	2052	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	2052	-0.20 (-0.39, -0.01)	0.04
Binary outcomes			
Generally overweight/obese	2052	0.98 (0.76, 1.26)	0.87
Centrally overweight/obese	2052	1.05 (0.77, 1.43)	0.78

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed), BMI: body mass index, WC: waist circumference, F&V fruit and vegetables.
In these analyses participants all participants are included (N = 2,221 (the number of participants recruited to the study). Missing baseline data is managed as in the main analyses and missing outcome data are imputed on the basis of those with missing data being 10% healthier than all participants in the study for a given outcome.

Supplementary Table S4: Sensitivity analysis: intention to treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention, with missing data for either baseline or follow-up measure of an outcome assumed to be 10% less healthy than the average value in the study sample.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	2052	1.04 (-1.18, 3.26)	0.36
Time spent in sedentary behaviour (minutes per day)	2052	-0.72 (-6.39, 4.95)	0.80
Servings of fruit and vegetables (number per day)	2052	0.01 (-0.16, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	2052	-10.74 (-26.30,4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	2052	-16.03 (-32.82, 0.76)	0.06
Body mass index (z(sd)-score)	2052	0.01 (-0.04, 0.06)	0.70
Waist circumference (z(sd)-score)	2052	-0.02 (-0.11, 0.06)	0.56
Servings of snacks (number per day)	2052	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	2052	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	2052	-0.20 (-0.39, -0.01)	0.04
Binary outcomes			
Generally overweight/obese	2052	0.98 (0.76, 1.26)	0.87
Centrally overweight/obese	2052	1.05 (0.77, 1.43)	0.78

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed), BMI: body mass index, WC: waist circumference, F&V fruit and vegetables.
In these analyses participants all participants are included (N = 2,221 (the number of participants recruited to the study). Missing baseline data is managed as in the main table and missing outcome data are imputed on the basis of those with missing data being 10% less healthy than all participants in the study for a given outcome.

Supplementary Table S5: Main intention to treat analyses of the effect of AFLY5 intervention on accelerometer-assessed outcomes during 3 valid days, separately for week and weekend days. Numbers vary by outcome as indicated in the table.

Outcome	Main comparison between the two groups (Intervention versus Control) on week days			Main comparison between the two groups (Intervention versus Control) on weekend days		
	Np	Difference in means (95%CI)	p-value	Np	Difference in means (95%CI)	p-value
Time spent in MVPA (minutes per day)	1627	2.47 (-1.37, 6.32)	0.21	972	3.26 (-3.62, 10.14)	0.35
Time spent in sedentary behaviour (minutes per day)	1627	1.87 (-8.51, 12.24)	0.72	972	3.07 (-10.91, 17.06)	0.67

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

All differences in means with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed).

In these analyses, participants were only included for each outcome if they had a follow-up measurement of that outcome. For partial missing baseline data we used an indicator variable as describe by White & Thompson,(1) which means for each outcome participants are included even if they do not have a baseline measurement.

Only participants included in the main analyses (i.e. with at least 3 valid days of accelerometer data) are included in this sensitivity analysis.

References

1. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med. 2005;24(7):993-1007. Epub 2004/12/01.

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[1,2]	See table 2	5-6
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	8
	2b	Specific objectives or hypotheses	Whether objectives pertain to the the cluster level, the individual participant level or both	9
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	9
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		**None so no reporting (protocol is published)**
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	9
	4b	Settings and locations where the data were collected		9-10 & 13-14
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	11-13
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and	Whether outcome measures pertain to the cluster level, the individual participant level or both	13 & Box 1

		when they were assessed		
	6b	Any changes to trial outcomes after the trial commenced, with reasons		**None so no reporting (protocol is published)**
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or *k*), and an indication of its uncertainty	14
	7b	When applicable, explanation of any interim analyses and stopping guidelines		N/A
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		10-11
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	10-11
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	10-11
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	10-11
	10b		Mechanism by which individual participants were included in clusters for the purposes of the	10-11

			trial (such as complete enumeration, random sampling)	
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	10
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		13 & 14
	11b	If relevant, description of the similarity of interventions		N/A
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	14-16
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		14-16
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	16 & Figure 1
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	16 & Figure 1
Recruitment	14a	Dates defining the periods of recruitment and follow-up		13
	14b	Why the trial ended or was stopped		N/A
Baseline data	15	A table showing baseline	Baseline characteristics for the	Table 1; 35-36

		demographic and clinical characteristics for each group	individual and cluster levels as applicable for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	Table 1; 35-36 Table 2; 37-38 Table 3; 39-40
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	37-40
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		37-40
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		Supplementary material
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[3])		**N/A – intervention was integrated into school teaching curriculum**
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		19-20
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	19-20
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		20

Other information			
Registration	23	Registration number and name of trial registry	6
Protocol	24	Where the full trial protocol can be accessed, if available	Referenced throughout the paper – reference numbers 9 and 17 in reference list which starts on page19
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	2-3 & 6

** Note: page numbers optional depending on journal requirements*

Table 2: Extension of CONSORT for abstracts[1,2] to reports of cluster randomised trials

Item	Standard Checklist item	Extension for cluster trials
Title	Identification of study as randomised	**Identification of study as cluster randomised**
Trial design	Description of the trial design (e.g. parallel, cluster, non-inferiority)	
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	**Eligibility criteria for clusters**
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	**Whether objective or hypothesis pertains to the cluster level, the individual participant level or both**
Outcome	Clearly defined primary outcome for this report	**Whether the primary outcome pertains to the cluster level, the individual participant level or both**
Randomization	How participants were allocated to interventions	**How clusters were allocated to interventions**
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants randomized to each group	**Number of clusters randomized to each group**
Recruitment	Trial status[1]	
Numbers analysed	Number of participants analysed in each group	**Number of clusters analysed in each group**
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	**Results at the cluster or individual participant level as applicable for each primary outcome**
Harms	Important adverse events or side effects	
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register	
Funding	Source of funding	

[1] Relevant to Conference Abstracts

REFERENCES

1 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008, 371:281-283

2 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG at al (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 5(1): e20

3 Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.

BMJ Open

Long-term effects of the Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010957.R2
Article Type:	Research
Date Submitted by the Author:	05-Jun-2016
Complete List of Authors:	Anderson, Emma Louise; University of Bristol, MRC Integrative Epidemiology Unit Howe, Laura; University of Bristol, Social Medicine Kipping, Ruth; University of Bristol, Faculty of Medicine and Dentistry Campbell, Rona; University of Bristol, School of Social and Community Medicine Jago, Russ; University of Bristol, Centre for Exercise, Nutrition and Health Sciences Noble, Sian; University of Bristol, School of Social and Community Medicine Wells, Sian; University of Bristol, School of Social & Community Medicine Chittleborough, Catherine; University of Adelaide, Discipline of Public Health, School of Population Health and Clinical Practice Peters, Tim; University of Bristol, Department of Community Based Medicine Lawlor, Debbie; Department of Social Medicine, University of Bristol, MRC Integrative Epidemiology Unit
Primary Subject Heading:	Public health
Secondary Subject Heading:	Paediatrics, Nutrition and metabolism
Keywords:	children, randomised controlled trial, schools, physical activity, diet

Long-term effects of the Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial

Emma L Anderson,[1,2] Laura D Howe,[1,2] Ruth R Kipping,[1] Rona Campbell,[1] Russell Jago,[3] Sian M Noble,[1] Sian Wells,[1] Catherine Chittleborough,[4] Tim J Peters,[1,5] Debbie A Lawlor.[1,2]

[1] School of Social & Community Medicine, University of Bristol, 39 Whatley Road, Bristol, BS8 2PS

[2] MRC Integrative Epidemiology Unit at the University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN

[3] Centre for Exercise, Nutrition & Health, School for Policy Studies, University of Bristol, 8 Priory Road, Bristol BS8 1TZ

[4] School of Public Health, University of Adelaide, 178 North Terrace, Adelaide, South Australia 5005

[5] School of Clinical Sciences, University of Bristol, 69 St Michael's Hill, Bristol, BS2 8DZ

Corresponding author:

Debbie A Lawlor

MRC Integrative Epidemiology Unit at the University of Bristol

Oakfield House, Oakfield Grove

Bristol BS8 2BN, UK

Tel: +44 (0)117 33 10096

E-mail: d.a.lawlor@bristol.ac.uk

ABSTRACT

Objective To investigate the long-term effectiveness of a school-based intervention to improve physical activity and diet in children.

Design Cluster randomised controlled trial.

Setting 60 primary schools in the south west of England.

Participants Primary school children who were aged 8-9 years at recruitment, 9-10 years during the intervention, and 10-11 years at the long-term follow-up assessment.

Intervention Teacher training, provision of lesson and child-parent interactive homework plans and teaching materials.

Main outcome measures Primary outcomes were accelerometer assessed minutes of moderate to vigorous physical activity (MVPA) per day, accelerometer assessed minutes of sedentary behaviour per day, and reported daily consumption of servings of fruit and vegetables.

Results 60 schools with 2221 eligible children were recruited. As in the previously published assessment immediately after the end of the intervention, none of the three primary outcomes differed between children in schools allocated to the intervention, compared to those in control schools at the end of the long-term follow-up (1-year after the end of the intervention). Differences in secondary outcomes were consistent with those at the immediate follow-up, with no evidence that these had diminished over time. Comparing intervention to control schools, the difference in mean child-reported screen viewing at the weekend was -16.03 minutes (95%CI: -32.82, 0.73), for servings of snacks per day the difference was -0.11 (95%CI: -0.39, 0.06), in servings of high energy drinks per day -0.20 (95%CI: -0.39, -0.01) and in servings of high fat foods per day -0.12 (95%CI: -0.39, 0.00). None of these reached our predefined level of statistical significance, especially after accounting for multiple testing.

Conclusion School based curriculum interventions alone are unlikely to have a major public health impact on children's diet and physical activity.

Trial registration Current Controlled Trials ISRCTN50133740.

Funding: This trial was funded by UK National Institute for Health Research (NIHR) Public Health Research Programme (09/3005/04). The funders had no role in the study design, data collection, analysis or interpretation of results.

Study strengths and limitations

Strengths

- The study was designed to take account of known sources of bias in other RCTs in this area.
- A protocol was published before recruitment started, and a detailed analysis plan was written before any access to the study data.
- Random allocation was concealed and outcome assessors were blinded to which group the schools and children had been randomised to.
- Accelerometers were used to objectively assess time spent in moderate to vigorous activity and sedentary behaviour.
- Our sample size calculation, which took account of the likely degree of clustering within schools.

Limitations

- The study was undertaken in state schools in the South West of England that covered a range of deprivation levels and both urban and rural communities, but results may not be generalizable to more ethnically diverse populations in the UK or beyond the UK

- There was missing data for the accelerometer assessed outcomes, but a range of sensitivity analyses did not alter our findings and levels of weartime and valid accelerometer data were similar in intervention and control arms

INTRODUCTION

Low levels of physical activity and fruit and vegetable consumption in childhood track into adulthood (1-3) and are associated with greater adiposity, adverse cardiometabolic risk factors, behavioural problems, low mood, and poorer academic attainment.(1-7) School-based interventions have the potential to efficiently change behaviours to healthier levels, or delay age-related changes in behaviour,(8) since most children attend school. However, previous randomised controlled trials (RCTs) of such interventions have potentially important sources of bias and few have explored long-term outcomes beyond the end of the intervention period.

A systematic review and meta-analysis of 44 school-based RCTs found beneficial effects on moderate or vigorous physical activity (MVPA) during school hours, but the authors noted that benefit might have been exaggerated due to the outcome assessment being self-/parental-reported and not blind to school allocation in most trials and because of the marked loss to follow-up in several trials.(9) In many of those RCTs the intervention included extra compulsory physical activity lessons or activities during school break-times. Those have the advantage that they do not interrupt the school curriculum, but in the absence of any long-term follow-up beyond the intervention period it is impossible to determine whether the greater time spent in MVPA is simply as a result of a level of compulsion to be more active. Evidence from observational epidemiological studies suggests that compulsory physical activity in lessons or break-time in school are associated with more school-based activity, but not with more activity outside of school or if the activity stops being compulsory.(10, 11) A systematic review restricted to studies that had used objectively assessed activity using accelerometers and did not restrict the outcome to activity during school hours found some evidence of benefit of a similar magnitude in both family focused and school curriculum

interventions, but noted that the magnitude of effect was modest.(12) Reviews of interventions to reduce time spent in sedentary behaviour have similarly noted some evidence of effect, but cautioned about likely sources of bias, including lack of adequate concealment of random allocation, subjective outcome measurements with no blinding of participants and little evaluation that effects were sustained long-term post intervention.(13,14) Likewise, two systematic reviews of school-based interventions to increase fruit and vegetable consumption found some possible evidence of modest effect but were concerned about lack of adequate concealment of random allocation and failure to take account of clustering within analyses.(15,16)

The Active for Life Year 5 (AFLY5) study (17) was a large school-based cluster randomised controlled trial (RCT). It was designed to addressed many of the limitations that had been identified in previous RCTs of interventions to improve physical activity and diet in children(9-16) by objectively measuring physical activity and sedentary behavoiur and by determing effects on outocmes both immediately after the end of the intervention as well as 12 months later. At the end of the intervention period (immediate follow-up), the intervention was ineffective at improving any of the three primary outcomes (time spent in moderate to vigorous physical activity, time spent in sedentary activity and fruit and vegetable consumption); however, it did result in improvements in three of the nine secondary outcomes (child-reported time spent screen-viewing at weekends, consumption of snacks and consumption of high-energy drinks).(18) A cluster randomised control trial design was necessary given the intervention is at the level of schools (rather than individual children).

In this paper, we report the long-term effects of the intervention on the primary and secondary outcomes that were assessed approximately 12 months post-intervention. Our initial aim when designing the study was to be able to determine whether any effects of the intervention would last beyond the period of the intervention. Given we now know the immediate post intervention results,(18) our aim in this paper was to determine whether any effects on primary outcomes emerged at the 12 month follow-up assessment (i.e. whether there was a delayed effect of the intervention on the primary outcomes) and whether effects on secondary outcomes that were observed immediately after the intervention were maintained, decreased or increased 12-months after the intervention. In this and the previous paper the intervention is delivered at the cluster (school) level and outcomes measured and analysed on individual children, with the clustering appropriately taken account of in the statistical analyses.

METHODS

Study design and participants

AFLY5 was a school-based, cluster RCT. Clustering was at the level of the schools, with eligibility for study entry being: (i) any state primary or junior schools that (ii) provided education to children aged 8 to 11 years and (iii) were within the Bristol City and North Somerset administrative areas (both areas in the South West of England). All children in UK school year 4 (age 8-9 years) at the time of recruitment were eligible for recruitment if their parents provided consent and they assented (see below).

A total of 60 state primary and junior schools were recruited between March and July 2011: 46 in Bristol and 14 in North Somerset, South West England. At the time of recruitment

participants were aged 8-9. Full details of the trial have been published previously so only a brief summary will be given here.(17-19) The trial was registered prior to recruitment of schools or data collection (http://www.controlled-trials.com/ISRCTN50133740). Analyses have been undertaken in accordance with a published analytical plan that was approved by the Trial Steering Committee.(17-19)

Ethical approval and consent

Ethical approval was obtained from the University of Bristol Faculty of Medicine and Dentistry Committee for Ethics (reference number 101115). Parents/guardians of children in Year 4 were sent a letter and information sheet about the study, with an opt-out consent form for each of the measurements and the opportunity to contact the research team to discuss the study as well as information about being able to withdraw at any stage. An information sheet for the child was sent at the same time that the letter was sent to the parents. Children were given a second copy of this information sheet at the time that measurements were undertaken and they were asked to give signed assent to each of the measurements.

Randomisation

Schools were defined as having high or low involvement in any initiatives aimed at increasing physical activity, reducing sedentary behaviour or increasing fruit and vegetable consumption, based on their report of involvement in local or national initiatives. Schools were also split into tertiles based on their score on the English Index of Multiple Deprivation 2010 (IMD 2010).(20) Schools were grouped into six mutually exclusive strata by these two characteristics and randomly allocated to control or intervention within these strata.(17-19) Randomisation was undertaken by DAL who was unaware of any other characteristics of the schools. School was concealed using the Bristol Randomised Trials Collaboration's

automated (remote) system. After randomisation, one school refused to undertake the intervention; the head reported that they had hoped they would be randomised to control and did not have the time or capacity to accommodate the intervention. This school was retained in the relevant analyses on an intention-to-treat basis.

Intervention

The intervention was adapted from a previously evaluated US intervention(21) and is based on Social Cognitive Theory,(22) with a particular emphasis on increasing the children's self-efficacy (perceived competence) to be physically active and eat a healthy diet.(23) Full details of the trial intervention have been published in the trial protocol and the paper reporting the immediate effect of the intervention.(17, 18) It comprised:

1. Training for classroom teachers and learning support assistants, provided by the trial manager, a nutritionist and physical education specialist. The training took place over a whole day (8-9 hours) in a non-school location and where the teachers / learning support assistants and those delivering the training would not be interrupted. Teachers / learning support assistants were given a choice of days to attend the training and schools were financially compensated for the cost of replacement teachers whilst their staff attended training. At the training days the rationale for the intervention was explained and each lesson and homework activity was discussed and then taught in interactive ways. Time was provided for questions and discussion. Teachers were instructed to deliver 16 lessons, 10 of which had associated homework. They were told that they could adapt the teaching plans and materials, as they would with other lessons, for example, to suit their own style and the range of abilities in their class, but the aims and knowledge / skills to be imparted should not be changed.

2. Provision of 16 lesson-plans and teaching materials, including pictures, CDs and journals for teachers or learning support assistants to deliver over two out of the three school-terms (6-7 months). The 16 lessons included 9 that were primarily related to how to be more active and less sedentary and why this was important, 6 to healthy nutrition and how to achieve this and 1 about reducing screen viewing. Each lesson did, however, combine different aspects of healthy behaviour. For example, in the physical activity lessons the children played games based on the food groups using photographs of food which reinforced the content of the nutrition lessons. Similarly, in the lesson (and associated homework) for reducing screen-viewing (called 'Freeze my TV') children were taught how to replace regular television watching with active play on some days.
3. Provision of 10 parental-child interaction homework activities. The activities were designed to involve parents and other family members in the behaviour change process and reinforced the messages delivered during lessons. The homeworks included activities such as: 'Freeze my TV', in which a specific time that would normally be spent watching television would be replaced with physically active play involving the parents and other family members that the child would write a log about; cooking simple healthy food at home; playing 'Top Grubs' a card game based on trumps with pictures of food, such that higher scoring (trumping) foods are the healthier ones; and measuring the sugar content of drinks that the family have at home or include in school/work lunch packs.
4. Information was provided for schools to insert (as they wished) in their school newsletters about the importance of increasing physical activity, reducing sedentary behaviour and improving diet. The inserts were sent to all intervention schools on

three occasions over the period of the intervention. Schools were free to edit these and insert none, all or some of them.

5. Written information for parents on how to encourage their children to eat healthily and be active was delivered via the school children at the start of the intervention.

The intervention took place when the children were aged 9-10 years (in UK school Year 5) after baseline assessment. Schools randomised to the control group continued standard education provision for the school year, and any involvement in additional health promoting activities, but had no access to the intervention teacher training or the teaching materials.

Outcomes

Box 1 lists the three primary and nine secondary outcomes.

Participant assessments

Baseline assessment (prior to intervention) was undertaken either between April and June 2011 or between September and November 2011, when the children were aged 8 to 9 years (i.e. before and after the school summer break). Immediate follow-up assessment was completed immediately post intervention approximately 12-months after the baseline assessment and the long-term assessment (with which this paper is concerned) took place 12-months after the immediate assessment, during which time the children were not exposed to the intervention. Every attempt was made to undertake the assessments in the same order so that the seasons would be similar at each assessment time.

Assessments measured primary and secondary outcomes, together with demographic characteristics and were conducted identically at each time point following published protocols.(17,19) They were completed by trained fieldworkers who were blinded as to which

arm of the trial schools had been allocated. Full details of these assessments have been published previously (17, 19) and are summarised here. Questionnaires asked for information on dietary intake and screen-time viewing and other characteristics and were administered in the classroom with at least one fieldworker present. Weight, height and waist circumference were measured in a private room by one of the trained fieldworkers, with a second fieldworker present in the room. All fieldworkers had passed Criminal Records Bureau checks, as required for working with children at the time that these data were collected. Physical activity was assessed using ActiGraph GT3X+ accelerometers (Actigraph LLC, Pensacola, Florida, USA) and time spent per day being sedentary and in moderate to vigorous activity were calculated using standard protocols as described previously.(17, 19)

Sample size calculation and account of multiple testing

Sample size calculations indicated that for the three primary outcome and nine secondary outcome measurements (including taking account of multiple testing with the secondary outcomes) a total of 60 schools with 1500 pupils (750 in each arm) needed to be recruited, so that 1275 (allowing for loss to follow-up) pupils could be included in the analyses.(17) This number - provided adequate power to detect what we considered to be minimally important effects.(17, 19) We recruited 60 schools and a total of 2,221 pupils, and included between 1066 and 2052 pupils in our analyses for different outcomes. Analyses for accelerometer based outcomes were on fewer participants than our sample size calculation suggested (N = 1066) because of a large proportion of participants not returning or not wearing the accelerometer for at least eight hours for three days, the minimum required to be included in the study.(17, 19)

Statistical Analyses

Full details of the analysis plan have been published previously.(19) Briefly, main analyses assessing the effect of the intervention on the primary and secondary 12 months post-intervention were conducted as intention-to-treat, with missing data at baseline being replaced with a value of 999 and a variable to indicate missing data at baseline (0=not missing, 1=missing) being included in regression models, as recommended by White et al.(24-26) For primary outcomes the level of statistical significance used was $p< 0.05$ and for secondary outcomes the level of statistical significance used was $p<0.01$, after correcting for multiple testing.(19) A series of sensitivity analyses were conducted to test assumptions regarding the nature of missing data at baseline and at each of the follow-up assessments (see detailed analysis plan (19) for discussion of these assumptions and the sensitivity analyses). Multilevel regression models were used to account for clustering (non-independence) of children within schools.(19) All analyses included adjustment for the following baseline variables: age, sex, baseline measure of the outcome being analysed, involvement in other healthy behaviour promoting activities and school level deprivation. A secondary per-protocol analysis was undertaken, in which classes in the intervention arm were only included in analyses if teachers had taught at least 70% (11 of 16) of the AFLY5 lessons. There was one school for which we were unable to confirm how many lessons had been taught. For that school, we first did analyses assuming that they had been taught at least 11 lessons and then repeated them assuming that they had been taught fewer than 11; the results were identical whichever of these alternatives were used. We additionally assessed whether the effect of the intervention on accelerometer-assessed outcomes differed by week or weekend day and whether the results were affected by implausible values as defined previously. The researchers undertaking the analyses were blinded to (unaware of) whether schools had been allocated to intervention or control arms.

As detailed in the published statistical protocol (19) we initially planned to assess change in outcomes between baseline and the long-term follow-up using multilevel models to estimate a trajectory of the repeat measurements (baseline, immediate follow-up, long-term follow-up) within each individual, with random effects to quantify the estimated person-specific deviation from the study mean in terms of the intercept (baseline measurement) and rate of change (slope). However, when we attempted to run these models, they did not converge. This is likely because there were only three measurement occasions, meaning that the model did not have sufficient degrees of freedom. Therefore, we conducted analyses at a single time point as described above (that is, assessed the effect of the intervention on outcomes at the long-term follow-up) and plotted differences between the randomised groups at each time point in order to illustrate any notable changes in estimates of the primary and secondary outcomes between baseline and immediate and long-term follow-up.

RESULTS

Figure 1 shows the trial profile. Of the 2,242 potentially eligible children in the 60 participating schools, 10 left the school prior to randomisation and baseline data collection and for 11 their parents or carers did not provide consent to participate in any aspect of the study. All other children (N = 2,221; 1064 in the schools that were randomised to intervention and 1157 in those randomised to control schools), irrespective of whether or not we have all the data for them, are included in the analyses presented here (with numbers differing for each outcome in the main analyses as a result of some missing data). Proportions with data for each outcome were similar in intervention and control schools at both baseline and at the second follow-up assessment at 12 months post-intervention (**Figure 1**). Baseline characteristics were similar between children in intervention schools and those in control

schools (**Table 1**).

Figures 2a to 2l shows differences in means or odds ratios between the control and intervention group for the three primary and nine secondary outcomes at baseline, immediate follow-up and long-term (12-months) follow-up. These show that differences in means (and odds ratios for general and central overweight/obesity) between children in intervention and control schools were essentially the same at this long-term follow-up as they were immediately after the intervention, when examining point estimates. Differences in the primary outcomes were consistent with the null hypothesis (**Figures 2a to 2c**). Differences in secondary outcomes were consistent with those seen at the end of the immediate follow-up,(**Figures 2d to 2l**) with no evidence that the previously reported beneficial effects for child-reported screen viewing at weekends,(**Figure 2e**) consumption of snacks (**Figure 2f**) and consumption of high energy drinks (**Figure 2h**) had notably diminished (or increased) in magnitude over time (Figures 2. However, there was no strong statistical support for any effect of the intervention on primary and secondary outcomes at 12 months after the intervention. **Table 2** shows differences in means or odds ratios for all outcomes at the long-term follow-up from the main intention-to-treat analyses. None of the three primary outcomes differed, nor the nine secondary outcomes, reached our predefined level of statistical significance for an effect after accounting for multiple testing.

Results from the per-protocol analyses were consistent with the intention-to-treat analyses results (**Table 3**). Results were similar in all sensitivity analyses applying different assumptions about missing data (**Supplementary Tables S1-S4**). Results were also similar when we looked separately at time spent in MVPA and time spent in sedentary behaviour by weekday and weekend (**Supplementary Table S5**).

DISCUSSION

In this school-based cluster RCT, aimed at increasing physical activity, reducing sedentary behaviours and improving diet in school aged children, we found results at 12 months after the intervention had ended (that is, with no further lessons or teaching aimed at promoting healthy activity and dietary levels during that 12 months) were essentially the same as those seen immediately after the end of the intervention in terms of size of effect. The lack of any effect on the three primary outcomes – time spent in MVPA, time spent in sedentary behaviour and fruit and vegetable consumption – was still observed 12 months later and the beneficial effects on three secondary outcomes (reported screen-viewing at weekends, consumption of snacks and of high energy drinks) were still somewhat present at 12-months post intervention. However, slight attenuation of the effect on these secondary outcomes meant that at this long-term follow-up none of our outcomes (primary or secondary) reached our pre-specified level of statistical significance.

Meaning of study findings

Whilst the effects for these secondary outcomes were consistent in magnitude with those seen at the immediate follow-up, they did not reach our pre-specified level of statistical significance. Thus, these results suggest that apparent benefits on these secondary outcomes are due to chance.

As discussed in our previous publication of effects immediately at the end of the intervention,(18) the lack of effect on primary outcomes, in particular on the objectively assessed accelerometer outcomes, might highlight the importance of societal and structural changes to support greater levels of activity, over and above any intervention at a school

level.(18) Our intervention was based on theory,(22, 23) built on a similar intervention that had been previously shown to work in the US(21) and in pilot work, conducted by us, it was shown to fit well with the primary school national curriculum in the UK.(27) Furthermore, the detailed process evaluation conducted as part of the full AFLY5 RCT, in which we used quantitative measures of intervention delivery and qualitative focus groups with children and in-depth interviews with teachers and parents,(28), showed that on average 77% of the intervention lessons and homeworks were delivered and reached 95% of the children in intervention schools. However, teachers felt lack of time and the need to prioritise numeracy and literacy skills over the health promoting lessons of our intervention were important barriers to them and the children being more fully engaged with AFLY5.(28) The process evaluation also highlighted that in general teachers did not like teaching physical activity, and had a tendency to delegate such lessons to teaching assistants. This might also have contributed to the null effects, particularly for the activity outcomes. Lastly, our process evaluation suggests that in the context of rapidly developing technologies the time taken to develop, test the feasibility of, and pilot, school-based interventions before completing large scale RCTs, as we have done in AFLY5, may mean that by the time school-based interventions get to the full scale RCT, the intervention is being implemented with out-of-date methods of delivery.(28, 29)

Whilst using schools for universal promotion of healthy behaviours is appealing, a key implication of our findings is that this alone is unlikely to have benefit. Pressures on schools to deliver academic success and the fact that teachers do not necessarily feel equipped, responsible for, or in the case of physical activity, enjoy promoting health behaviours,(28) suggest that curriculum based health promotion alone is unlikely to benefit population health. Our RCT was large and well-conducted and the results suggest that further investment in

RCTs of curriculum based interventions (alone) to improve children's diet and activity are not wanted. Whether investing in extra-curricular activities, including in the necessary human resources (e.g. people who are appropriately trained and skilled), structural resources (appropriate space) and equipment, would be beneficial at a population level is unclear and may warrant further evaluation. Societal interventions such as those that were envisaged as a legacy of the 2012 Olympics, and the more recent 'sugar tax' may be beneficial but will require a natural experiment type approach,(30) rather than an RCT, for their evaluation. Evaluation of past major sporting events and early assessments of the 2012 Olympics, suggest that like our assessment of a school based curriculum, much more intense, comprehensive (across all levels of society – home, neighbourhoods, schools, work, government, transport systems) and long-term investments are required to support the next generation to be more active and eat healthier.(31-33)

Strengths and limitations

The study was designed to take account of known sources of bias in other RCTs in this area. A protocol was published before recruitment started, and a detailed analysis plan was written before any access to the study data. We developed an intervention according to guidelines for complex interventions, with the theoretical rationale for the intervention described in detail elsewhere.(18) Our sample size calculation, which took account of the likely degree of clustering within schools, indicated that we needed a total of 1275 children to be included in the analyses. For all outcomes, except those related to accelerometer data, we achieved considerably higher numbers than this target. The number included in the main analyses for accelerometer based data was somewhat smaller than this at 1066. Sample size calculations are an approximation of the numbers needed, and we doubt that such a small difference will have had a major effect on our conclusions. Furthermore, wear time was similar in children in

intervention and control schools; moreover, in sensitivity analyses using different approaches to dealing with missing data and which included 2052 children even for the accelerometer outcomes, the results were essentially the same as in the main analysis. One school refused to deliver any of the intervention, and others did not deliver all of the lessons. However, the per-protocol analysis, which did not differ from the main intention-to-treat analysis, shows that this does not explain the null results.

Conclusion

This long-term follow-up of a large well-conducted school based RCT has found similar results to those found immediately after the intervention period. None of the primary or secondary outcomes reached our predefined levels of statistical significance, suggesting that apparent benefits on some secondary outcomes are due to chance. Overall, together with our process evaluation these findings suggest that curriculum-based interventions alone are unlikely to make a major impact on promoting healthy levels of physical activity and healthy diets in primary school children.

Acknowledgements: We thank all the students and teaching staff who took part in AFLY5. We thank all of the AFLY5 staff who include fieldworkers, administrative staff, computing and data management staff and the trainers who provided teacher training. We thank Dr Hugh Annett (retired Director of Public Health, NHS Bristol and Bristol City Council), Annie Hudson (former Strategic Director for Children, Young People and Skills, Bristol City Council) and Sheila Smith (Strategic Director for Children, Young People and Skills, North Somerset City Council) for their support of the Active for Life Year 5 study. We also thank the Chair and members of the trial steering committee for their advice and support.

The views expressed in this paper are those of the authors and not necessarily anyone in this acknowledgement list.

Funding: The AFLY5 RCT is funded by the UK National Institute for Health Research (NIHR) Public Health Research Programme (09/3005/04), which also paid the salary of ELA and SW. DAL and LDH work in a Unit that receives funds from UK Medical Research Council (MC_UU_12013/5). RRK and RC work in the Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research Centre of Excellence: joint funding (MR/KO232331/1) from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the Welsh Government and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. LDH is supported by a UK Medical Research Council Population Health Scientist fellowship (G1002375). DAL (NF-SI-0611-10196) and TJP (NF-SI-0512-10026) are NIHR Senior Investigators. This study was undertaken in collaboration with the Bristol Randomised Trials Collaboration (BRTC), a UKCRC Registered Clinical Trials Unit in receipt of National Institute for Health Research CTU support funding.

None of the funders had involvement in the Trial Steering Committee, the data analysis, data interpretation, data collection, or writing of the paper. DAL, LDH and ELA had access to all of the data in the study and DAL had the final responsibility for the decision to submit for publication.

The views expressed in this publication are those of the authors and not necessarily any of the funding bodies listed here.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: support from research funders in accordance with the funding statement included in

this manuscript; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work, other than that RC is director of DECIPHer Impact, a not-for-profit company that is wholly owned by the Universities of Bristol and Cardiff and whose purpose is to licence and support the implementation of evidenced based health promotion interventions.

Ethical approval: Ethical approval was obtained from the University of Bristol Faculty of Medicine and Dentistry Committee for Ethics (reference number 101115).

Transparency declaration: DAL affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Data sharing: Full details of the study are available on the website (www.bristol.ac.uk/social-community-medicine/projects/afl/). We encourage anyone who would like to access these data for other projects to contact the corresponding author. We would be happy for external collaborators to access these data according to our data transfer agreement.

Contributors: DAL, together with RRK, RC, TJP, SN and RJ designed the AFLY5 study and obtained funds to complete it. DAL wrote the analysis plan used for this paper and LDH and ELA completed all analyses. SW managed the AFLY5 study, including managing data collection. ELA, LDH and DAL wrote the first draft of the paper, and DAL coordinated contributions from other co-authors. All authors contributed to the overall study aim and development of the design. All authors made critical comments on drafts of the paper.

References

1. Lock K, Pomerleau J, Causer L, Altmann DR, McKee M. The global burden of disease attributable to low consumption of fruit and vegetables: implications for the global strategy on diet. Bull World Health Organ. 2005;83(2):100-8. Epub 2005/03/04.

2. Maynard M, Gunnell D, Emmett P, Frankel S, Davey Smith G. Fruit, vegetables, and antioxidants in childhood and risk of adult cancer: the Boyd Orr cohort. J Epidemiol Community Health. 2003;57(3):218-25. Epub 2003/02/21.

3. Ness AR, Maynard M, Frankel S, Smith GD, Frobisher C, Leary SD, et al. Diet in childhood and adult cardiovascular and all cause mortality: the Boyd Orr cohort. Heart. 2005;91(7):894-8. Epub 2005/06/17.

4. Boreham C, Riddoch C. The physical activity, fitness and health of children. J Sports Sci. 2001;19(12):915-29.

5. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act. 2010;7:40. Epub 2010/05/13.

6. Ness AR, Leary SD, Mattocks C, Blair SN, Reilly JJ, Wells J, et al. Objectively measured physical activity and fat mass in a large cohort of children. PLoS Med. 2007;4(3):e97. Epub 2007/03/29.

7. Ekelund U, Luan J, Sherar LB, Esliger DW, Griew P, Cooper A, et al. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. JAMA. 2012;307(7):704-12. Epub 2012/02/18.

8. Cooper AR, Goodman A, Page AS, Sherar LB, Esliger DW, van Sluijs EM, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). Int J Behav Nutr Phys Act. 2015;12:113. Epub 2015/09/18.

9. Dobbins M, Husson H, DeCorby K, LaRocca RL. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. Cochrane Database Syst Rev. 2013;2:CD007651. Epub 2013/03/02.

10. Mallam KM, Metcalf BS, Kirkby J, Voss LD, Wilkin TJ. Contribution of timetabled physical education to total physical activity in primary school children: cross sectional study. BMJ. 2003;327:592-3.

11. Cleland V, Dwyer T, Blizzard L, Venn A. The provision of compulsory school physical activity: associations with physical activity, fitness and overweight in childhood and twenty years later. Int J Behav Nutr Phys Act. 2008;5:14

12. Metcalf B, Henley W, Wilkin T. Effectiveness of intervention on physical activity of children: systematic review and meta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). BMJ. 2012;345:e5888. Epub 2012/10/10.

13. DeMattia L, Lemont L, Meurer L. Do interventions to limit sedentary behaviours change behaviour and reduce childhood obesity? A critical review of the literature. Obes Rev. 2007;8(1):69-81.

14. van Grieken A, Ezendam NP, Paulis WD, van der Wouden JC, Raat H. Primary prevention of overweight in children and adolescents: a meta-analysis of the effectiveness of interventions aiming to decrease sedentary behaviour. Int J Behav Nutr Phys Act. 2012;9:61. Epub 2012/05/30.

15. Delgado-Noguera M, Tort S, Martinez-Zapata MJ, Bonfill X. Primary school interventions to promote fruit and vegetable consumption: a systematic review and meta-analysis. Prev Med. 2011;53(1-2):3-9. Epub 2011/05/24.

16. Evans CE, Christian MS, Cleghorn CL, Greenwood DC, Cade JE. Systematic review and meta-analysis of school-based interventions to improve daily fruit and vegetable intake in children aged 5 to 12 y. AM J CLIN NUTR. 2012;96(4):889-901. Epub 2012/09/07.

17. Lawlor DA, Jago R, Noble SM, Chittleborough CR, Campbell R, Mytton J, et al. The Active for Life Year 5 (AFLY5) school based cluster randomised controlled trial: study protocol for a randomized controlled trial. Trials. 2011;12:181. Epub 2011/07/26.

18. 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.

19. Lawlor DA, Peters TJ, Howe LD, Noble SM, Kipping RR, Jago R. The Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial protocol detailed statistical analysis plan. Trials. 2013;14(1):234. Epub 2013/07/26.

20. Department for Communities and Local Government. The English Indices of Deprivation 2010. London: 2011.

21. Gortmaker SL, Cheung LW, Peterson KE, Chomitz G, Cradle JH, Dart H, et al. Impact of a school-based interdisciplinary intervention on diet and physical activity among urban primary school children: eat well and keep moving. Arch Pediatr Adolesc Med. 1999;153:975-83

22. Bandura A: Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall; 1986

23. Bandura A: Self-efficacy: The exercise of control. New York: Freeman; 1997.

24. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clinical trials. 2012;9(4):396-407. Epub 2012/07/04.

25. 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. Epub 2011/02/09.

26. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med. 2005;24(7):993-1007. Epub 2004/12/01.

27. Kipping RR, Payne C, Lawlor DA. Randomised controlled trial adapting US school obesity prevention to England. Archives of Disease in Childhood. 2008;93(6):469-73.

28. Campbell R RE, Wells S, Kipping RR, Chittleborough CR, Peters TJ, Lawlor DA, Jago R. Intervention fidelity in a school-based diet and physical activity intervention in the UK: Active for Life Year 5. International Journal of Behavioral Nutrition and Physical Activity. 2015;12: 141.

29. Jago R RE, Kipping RR, Wells S, Chittleborough CR, Peters TJ, Mytton J, Lawlor DA, Campbell R. Lessons learned from the AFLY5 RCT process evaluation: Implications for the design of physical activity and nutrition interventions in schools. BMC Public Health. 2015;15:946.

30. Glymour MM. Natural Experiments and Instrumental Variable Analyses in Social Epidemiology. In: Oakes JM, Kaufman JS. Methods in Social Epidemiology. John Whiley & Son, San Francisco; 2006

31. Weed M, Coren E, Fiore J, Mansfield L, Wellard I, CHatziefstathiou D, Dowse S. A systematic review of the Evidence base for Developing a Physical Activity and Health Legacy from the Londno 2012 Olympic and Paralympic Games. Department of Health 2009.

32. Griffiths M, Armour K. Physical education and youth sport in England: conceptual and practical foundations for an Olympic Legacy? Internatoinal Journal of Sport Policy and Politics 2013; 5: 213-227.

33. Mackintosh C, Darko N, Rutherford Z, Wilkins H-M. A qualitative study of the impact of the London 2012 Olympocs on families in the East Midlands of England: lessonds fo rsports development policy and practice. Sport, Education and Society 2015; 20: 1065-1087

Box 1: AFLY5 primary and secondary outcomes

Primary outcomes

Accelerometer assessed mean time per day spent doing moderate/vigorous physical activity MVPA (minutes per day)

Accelerometer assessed mean time per day spent in sedentary activity (minutes per day)

Self-reported (validated questionnaire) servings of fruit and vegetables consumed per day (servings per day; treated in all analyses as a continuous variable)

Secondary outcomes

Self-reported (validated questionnaire) mean time spent screen viewing on a typical weekday (minutes)

Self-reported (validated questionnaire) mean time spent screen viewing on a typical weekend day (minutes)

Self-reported (validated questionnaire) servings of snacks consumed per day (servings per day; treated in all analyses as a continuous variable)

Self-reported (validated questionnaire) servings of high fat foods consumed per day (servings per day; treated in all analyses as a continuous variable)

Self-reported (validated questionnaire) servings of high energy drinks consumed per day (servings per day; treated in all analyses as a continuous variable)

Body mass index determined from weight and height measured in classrooms by two study fieldworkers (kg/m^2; treated in all analyses as a standard deviation z-score)

Waist circumference measured in classrooms by two study fieldworkers (mm; treated in all analyses as a standard deviation z-score)

General overweight/obesity, determined by the International Obesity Task Force thresholds of body mass index for children (taking account of their age and sex) (binary outcome)

Central overweight/obesity determined by thresholds of UK age and sex specific reference charts for waist circumference and defined by the International Diabetes Federation. (binary outcome)

Figure legends

Figure 1 – Trial profile.

Footnote to Figure 1

Np=number of participants (school pupils). No schools withdrew from study, so all randomised units are present at baseline and at both follow-up assessments. Percentages for proportions of children with each measurement at baseline and at follow-ups are of total number of children who were pupils in randomised schools at baseline. Not all pupils with follow-up measures necessarily had data on the same measure at baseline (or vice versa), because of different pupils being absent at baseline and follow-up assessments at each time point, and because of pupils leaving or moving between schools. In all analyses, study participants were analysed in the group (intervention or control) to which they were randomised.

Figure 2: Difference in means and odds ratios for the intervention compared to the control group for the three primary outcomes and nine secondary outcomes, assessed at baseline, first follow-up (conducted immediately after the end of the intervention) and second follow-up (12-months post-intervention).

a. Accelerometer assessed time spent in moderate to vigorous physical activity

b. Time spent in sedentary behaviour

c. Servings of Fruit and Vegetables per day

d. Time spent screen viewing on weekdays

e. Time spent screen viewing on Saturdays

f. Servings of snacks per day

g. Servings of high fat foods per day

h. Servings of high energy drinks per day

i. Body mass index z-score (as a continuous variable)

j. Waist circumference z-score (as a continuous variable)

k. General overweight or obesity (based on BMI measurements)

l. Central overweight / obesity based on waist circumference measurements

Footnote to Figure 2

The figures all show differences in means for continuous variables (graphs a to j) and odds ratios for binary outcomes (graphs k and l), comparing those in the intervention arm of the

trial to those in the control arm (dots), together with 95% confidence intervals (vertical lines with horizontal caps representing the limits). The dashed horizontal lines represent the null values (zero for all differences in means of continuous variables and one for odds ratios of binary outcomes).

Table 1: Comparison of baseline characteristics by randomised group

Characteristic	Unit and type of summary measure	Intervention schools N participants=1064 N schools = 30		Control schools N participants=1157 N schools = 30	
		Number	Distribution	Number	Distribution
Age	Mean (SD) years	1024	9.5 (0.3)	1099	9.5 (0.3)
MVPA[a]	Mean (SD) minutes	912	59 (23)	928	56 (21)
Sedentary behaviour[a]	Mean (SD) minutes	912	422 (72)	928	416 (68)
Servings of fruit and vegetables	Median (IQR) number / day	1019	1 (0 to 2)	1088	1 (0 to 2)
Servings of snacks	Median (IQR) number / day	1019	2 (1 to 3)	1088	2 (1 to 3)
Servings of high fat foods	Median (IQR) number / day	1019	0 (0 to 1)	1088	1 (0 to 1)
Servings of high energy drinks	Median (IQR) number / day	1019	2 (1 to 3)	1088	2 (1 to 3)
BMI	Mean (SD) z-score	889	-0.06 (0.94)	953	0.05 (1.04)
WC	Mean (SD) z-score	942	-0.03 (0.97)	1027	0.03 (1.02)
Screen-viewing weekday	Median (IQR) minutes	1024	105 (45 to 240)	1099	105 (45 to 225)
Screen-viewing Saturday	Median (IQR) minutes	1024	90 (30 to 240)	1099	105 (30 to 240)
Total number of valid days of wearing accelerometer[b]	Median (IQR) days	912	3 (2 to 5)	928	3 (2 to 4)
Total number of valid weekdays of wearing accelerometer[b]	Median (IQR) days	979	2 (2 to 3)	1025	2 (1 to 3)
Total hours of wearing accelerometer on valid days[a]	Mean (SD) hours / day	912	11.6 (1.5)	928	11.5 (1.4)
Hours of wearing accelerometer on	Mean (SD) hours / day	896	11.8 (1.6)	919	11.7 (1.5)

valid weekdays[b]					
Categorical variables					
Gender	N (%) female	520	49%	608	52%
	N (%) male	544	51%	549	48%
General overweight / obesity	N (%) No	717	81%	743	78%
	N (%) Yes	172	19%	210	22%
Central overweight/obesity	N (%) No	601	64%	631	61%
	N (%) Yes	341	36%	396	39%
Returned accelerometer	N (%) No	85	8%	132	11%
	N (%) Yes	979	92%	1025	89%
Wore accelerometer for requested amount of time	N (%) No	820	77 %	953	82%
	N (%) Yes	244	23%	204	18%
Wore accelerometer for required amount of time	N (%) No	418	39%	514	44%
	N (%) Yes	646	61%	643	56%
School involved in other health promoting activities	N (%) No	264	25%	446	39%
	N (%) Yes	800	75%	711	61%
School deprivation score	N (%) low	315	30%	460	40%
	N (%) medium	368	35%	345	30%
	N (%) high	381	36%	352	30%

SD: standard deviation; MVPA: moderate or vigorous physical activity; IQR: interquartile range; BMI: body mass index; WC: waist circumference

[a]Including only participants with at least 3 days of valid data

[b]Including all valid days, regardless of the number of valid days

Note some % within categories do not sum to exactly 100 because of rounding

Table 2: Main intention-to-treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. Numbers of participants vary by outcome as indicated in the table.

Outcome (primary/secondary)	Control group (reference group)		Intervention group		Main comparison between the two groups (Intervention versus Control)		
	Np	Mean (SD) or number (%)	Np	Mean (SD) or number (%)	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes:							
Time spent in MVPA (minutes per day)	522	52.56 (20.67)	527	54.37 (22.23)	1049	2.48 (-1.80, 6.77)	0.26
Time spent in sedentary behaviour (minutes per day)	522	461.78 (66.33)	527	465.46 (70.61)	1049	2.79 (-7.78, 13.37)	0.60
Servings of fruit and vegetables (number per day)	1062	1.80 (1.55)	990	1.82 (1.59)	2052	0.01 (-06, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	1062	148.01 (126.39)	990	138.88 (125.00)	2052	-10.74 (-26.30, 4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	1062	180.52 (164.82)	990	167.71 (156.28)	2052	-16.03 (-32.82, 0.73)	0.06
Body mass index (z-score)	923	0.03 (1.02)	870	-0.03 (0.97)	1793	0.01 (-0.04, 0.06)	0.72
Waist circumference (z-score)	993	0.03 (1.04)	935	-0.03 (0.95)	1928	-0.04 (-0.13, 0.05)	0.36
Servings of snacks (number per day)	1062	2.11 (1.55)	990	1.99 (1.47)	2052	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	1062	0.86 (0.94)	990	0.74 (1.07)	2052	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	1062	2.38 (1.58)	990	2.19 (1.45)	2052	-0.20 (-0.39, -0.01)	0.04
Binary outcomes							
Generally overweight/obese	923	194 (21.02)	870	175 (20.11)	1793	1.00 (0.72, 1.37)	0.98
Centrally overweight/obese	993	421 (42.40)	935	394 (42.14)	1928	1.08 (0.80, 1.46)	0.62

Np: number of participants; SD: standard deviation; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance after taking account of multiple testing)

All differences in means / odds ratios with their 95% CIs have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline/school stratifying variables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses participants were included for each outcome if they had a follow-up measurement of that outcome; for missing baseline data we used an indicator variable as described by White & Thompson,(21) which means for each outcome participants are included even if they do not have a baseline measurement.

Table 3: Per-protocol analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. Numbers vary by outcome as indicated in the table.

Outcome (**primary**/secondary)	**Control group (reference group)**		**Intervention group**		**Main comparison between the two groups (Intervention versus Control)**		
	Np	**Mean (SD) or number (%)**	**Np**	**Mean (SD) or number (%)**	**Np**	**Difference in means or odds ratio (95%CI)**	**p-value**
Continuous outcomes							
Time spent in MVPA (minutes per day)	522	52.56 (20.67)	356	54.15 (22.27)	878	2.63 (-2.10, 7.37)	0.28
Time spent in sedentary behaviour (minutes per day)	522	461.78 (66.33)	356	466.17 (70.58)	878	3.67 (-8.32, 15.66)	0.55
Servings of fruit and vegetables (number per day)	1062	1.80 (1.55)	701	1.91 (1.66)	1762	0.05 (-0.15, 0.25)	0.63
Time spent screen-viewing (minutes per day weekday)	1062	148.01 (126.39)	701	134.98 (120.94)	1762	-8.97 (-26.81, 8.87)	0.32
Time spent screen-viewing (minutes per day Saturday)	1062	180.52 (164.82)	701	159.35 (149.97)	1762	-21.73 (-41.19, -2.26)	0.03
Body mass index (z-score)	923	0.03 (1.02)	612	-0.03 (0.98)	1535	0.01 (-0.05, 0.07)	0.69
Waist circumference (z-score)	993	0.03 (1.04)	657	-0.04 (0.94)	1650	-0.03 (-0.13, 0.06)	0.52
Servings of snacks (number per day)	1062	2.11 (1.55)	701	2.07 (1.48)	1762	-0.03 (-0.23, 0.16)	0.72
Servings of high fat foods (number per day)	1062	0.86 (0.94)	701	0.75 (1.15)	1762	-0.11 (-0.26, 0.04)	0.14
Servings of high energy drinks (number per day)	1062	2.38 (1.58)	701	2.22 (1.43)	1762	-0.18 (-0.41, 0.5)	0.12
Binary outcomes							
Generally overweight/obese	923	194 (21.02)	612	121 (19.77)	1535	0.98 (0.68, 1.41)	0.91
Centrally overweight/obese	993	421 (42.40)	657	272 (41.40)	1650	1.06 (0.76, 1.49)	0.72

Np: number of participants; SD: standard deviation; MVPA: moderate to vigorous physical activity (accelerometer assessed); CI: confidence interval

Per-protocol analysis defined as teaching at least 70% (11 out of the 16) AFLY5 lessons. All participants from the intervention schools where the teacher taught fewer than 11 lessons are excluded from these analyses.

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance after taking account of multiple testing)

All differences in means/odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline/school stratifying variables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

In these analyses, after removal of schools that did not teach at least 11 out of 16 of the lessons, participants were only included for each outcome if they had a follow-up measurement of that outcome. For partial missing baseline data we used an indicator variable as described by White & Thompson,(21) which means for each outcome participants are included even if they do not have a baseline measurement.

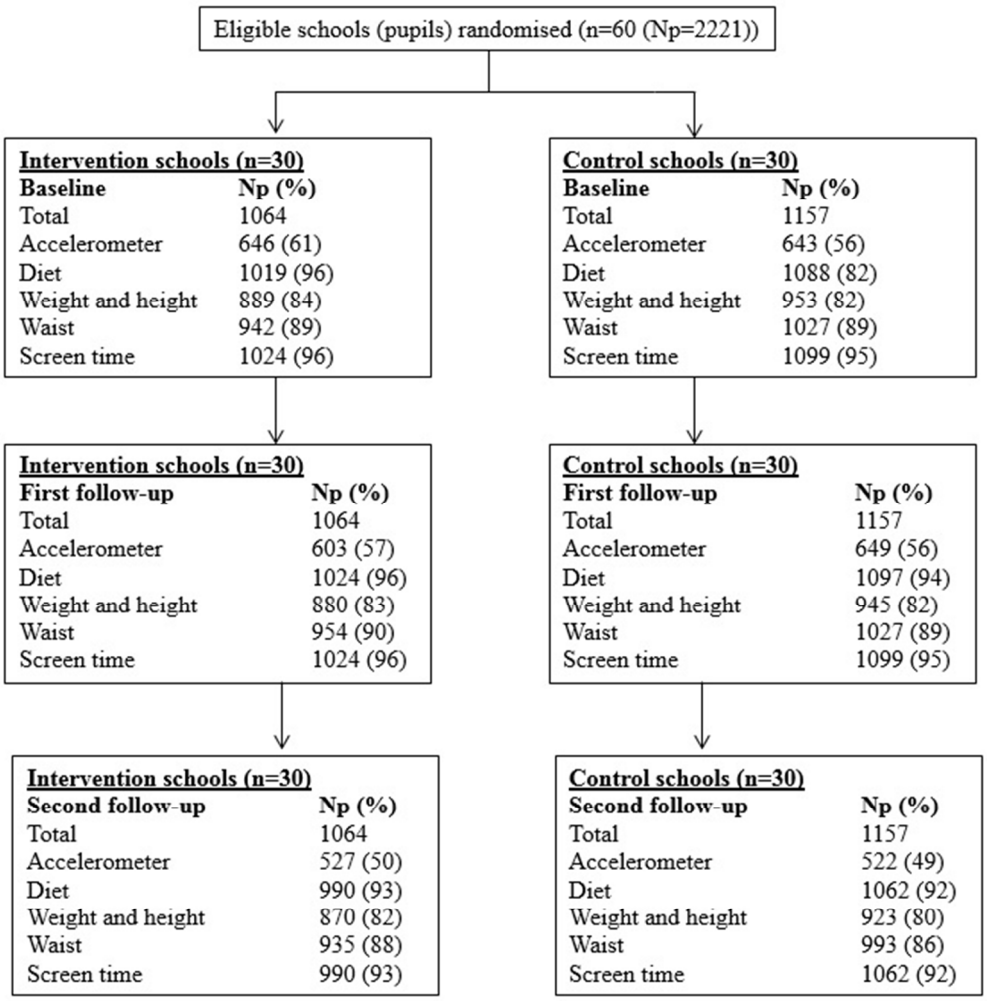


Figure 1 - Trial Profile
Footnote to the figure provide
55x56mm (300 x 300 DPI)

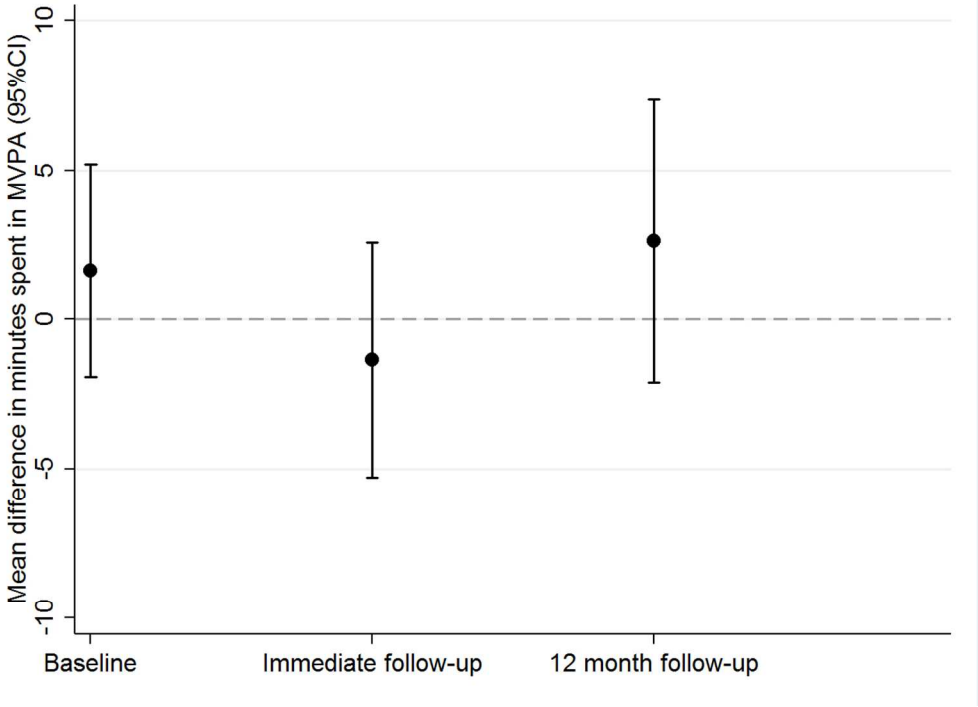
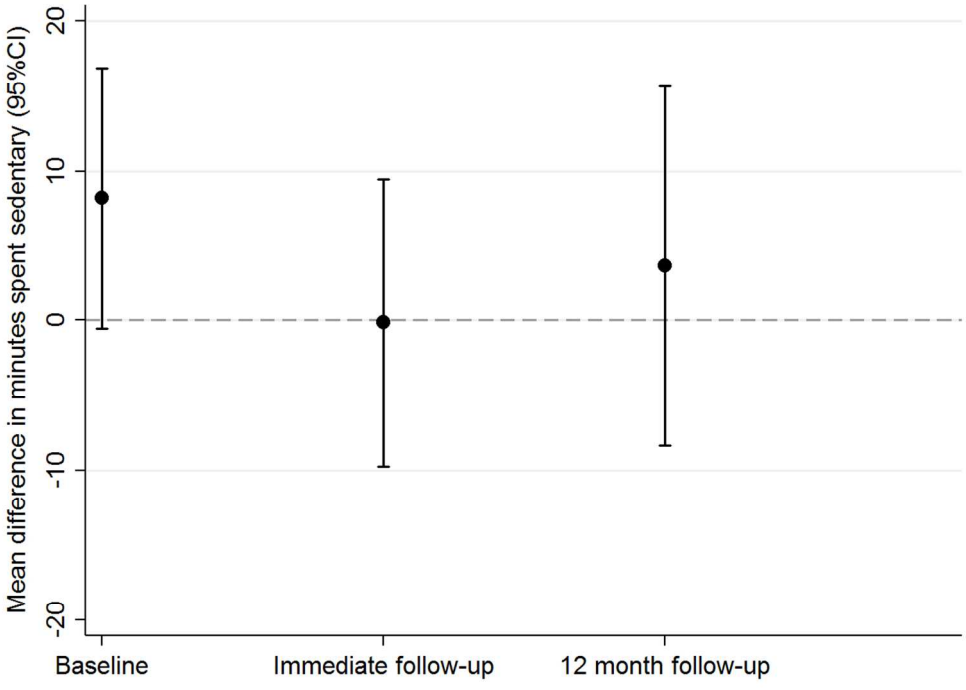


Figure 2: Difference in means and odds ratios for the intervention compared to the control group for the three primary outcomes and nine secondary outcomes, assessed at baseline, first follow-up (conducted immediately after the end of the intervention) and second follow-up (12-months post-intervention).

Specifically for this figure
Figure 2 a. Accelerometer assessed time spent in moderate to vigorous physical activity
Footnote in main document for
108x79mm (300 x 300 DPI)



b. Time spent in sedentary behaviour

108x78mm (300 x 300 DPI)

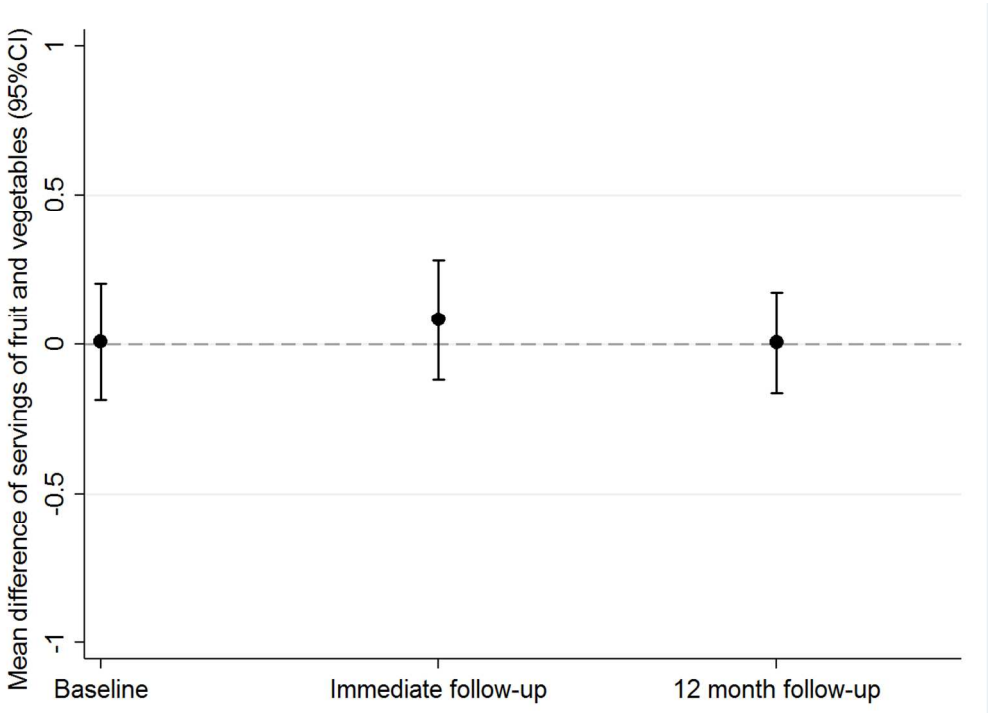


Figure 2 c. Servings of Fruit and Vegetables per day

113x82mm (300 x 300 DPI)

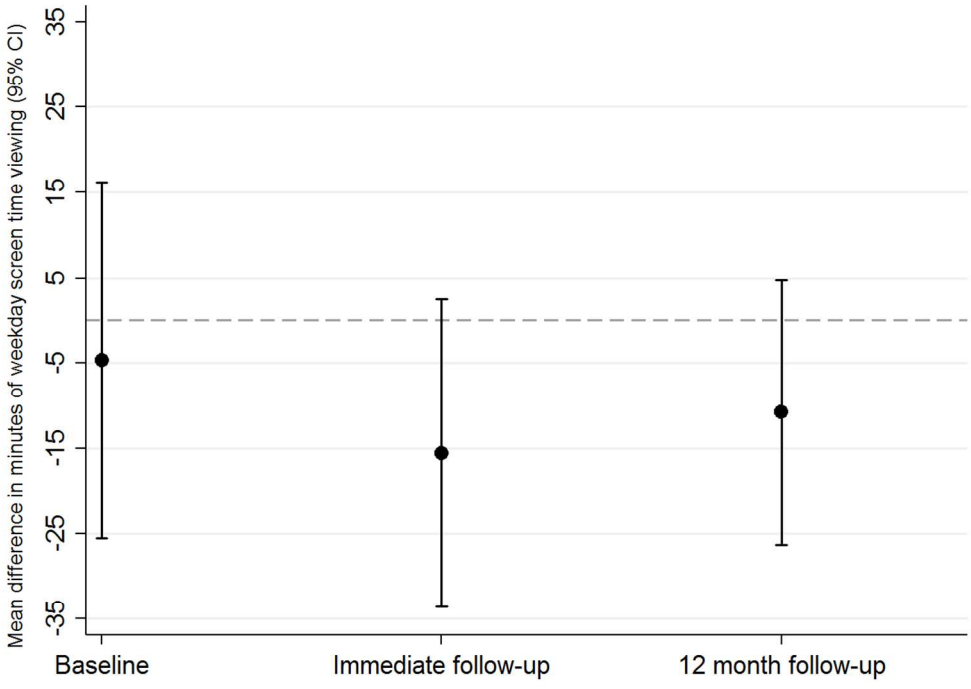


Figure 2 d. Time spent screen viewing on weekdays

113x82mm (300 x 300 DPI)

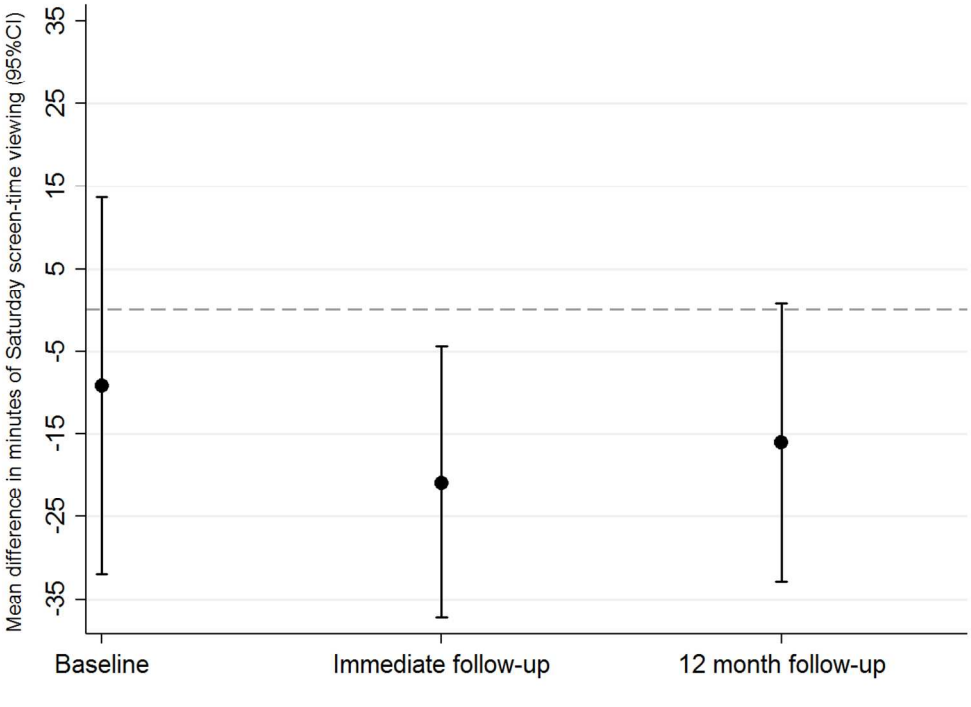


Figure 2 e. Time spent screen viewing on Saturdays

113x82mm (300 x 300 DPI)

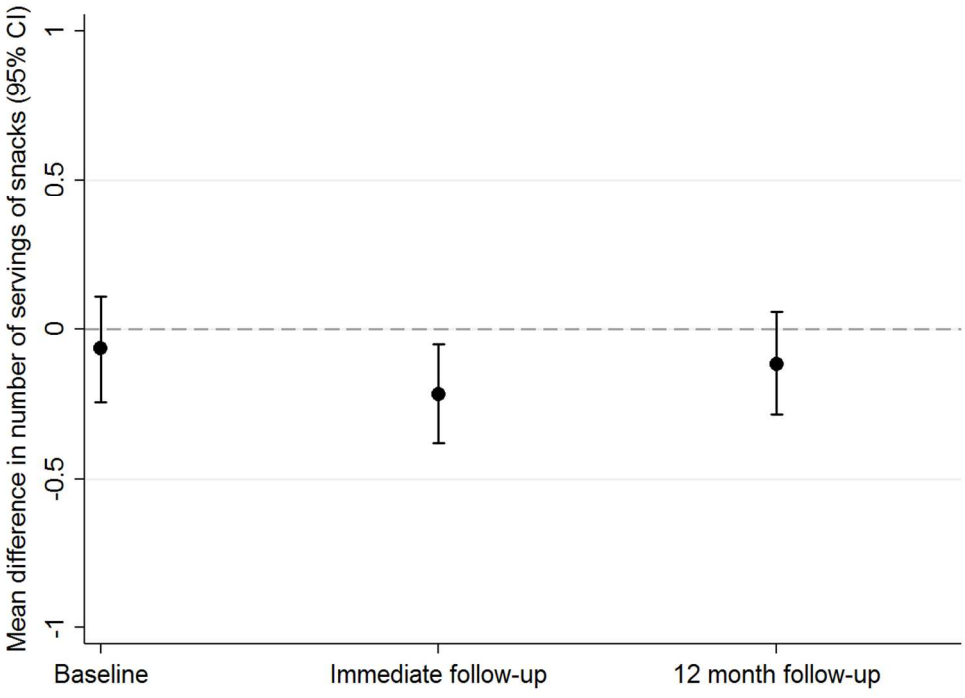


Figure 2 f. Servings of snacks per day

113x82mm (300 x 300 DPI)

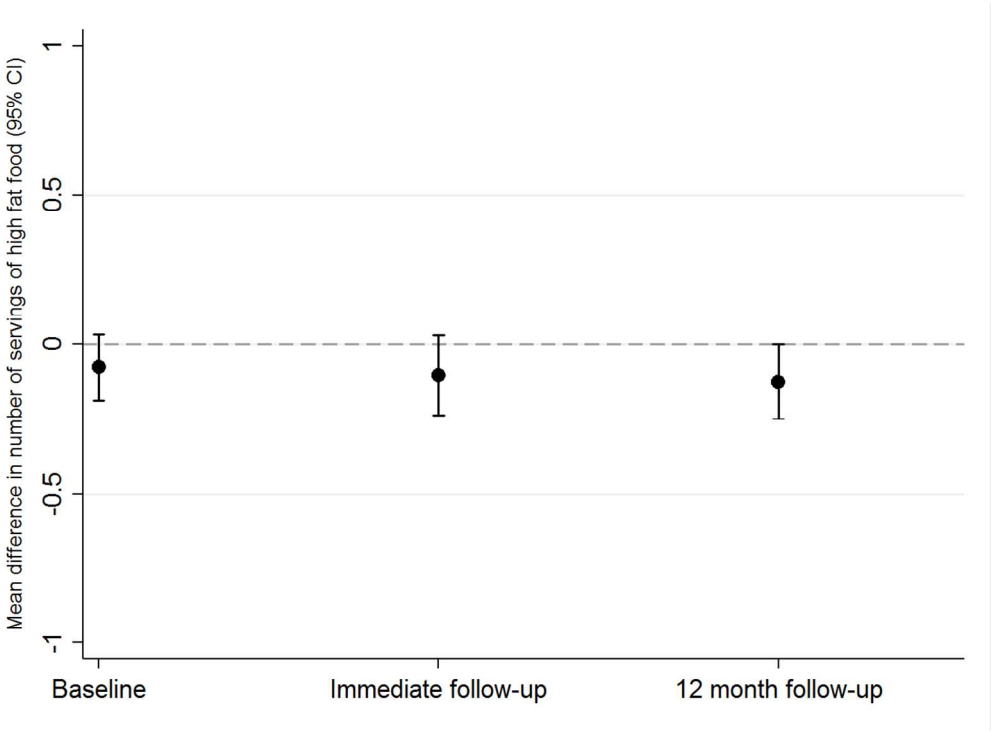


Figure 2 g. Servings of high fat foods per day

113x82mm (300 x 300 DPI)

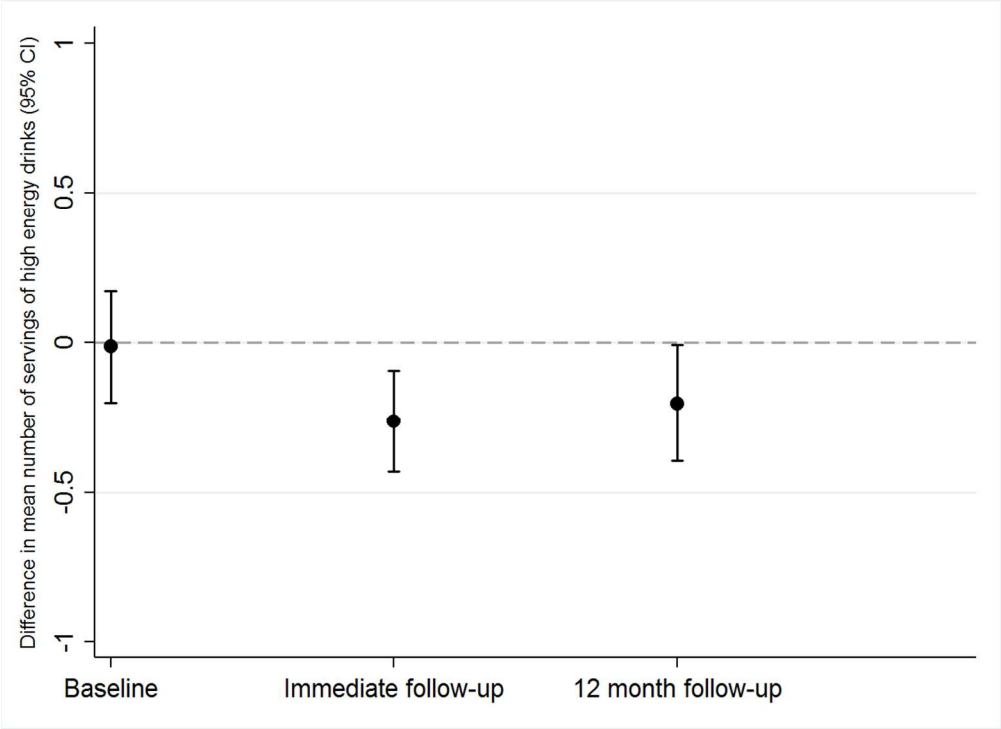


Figure 2 h. Servings of high energy drinks per day

109x79mm (300 x 300 DPI)

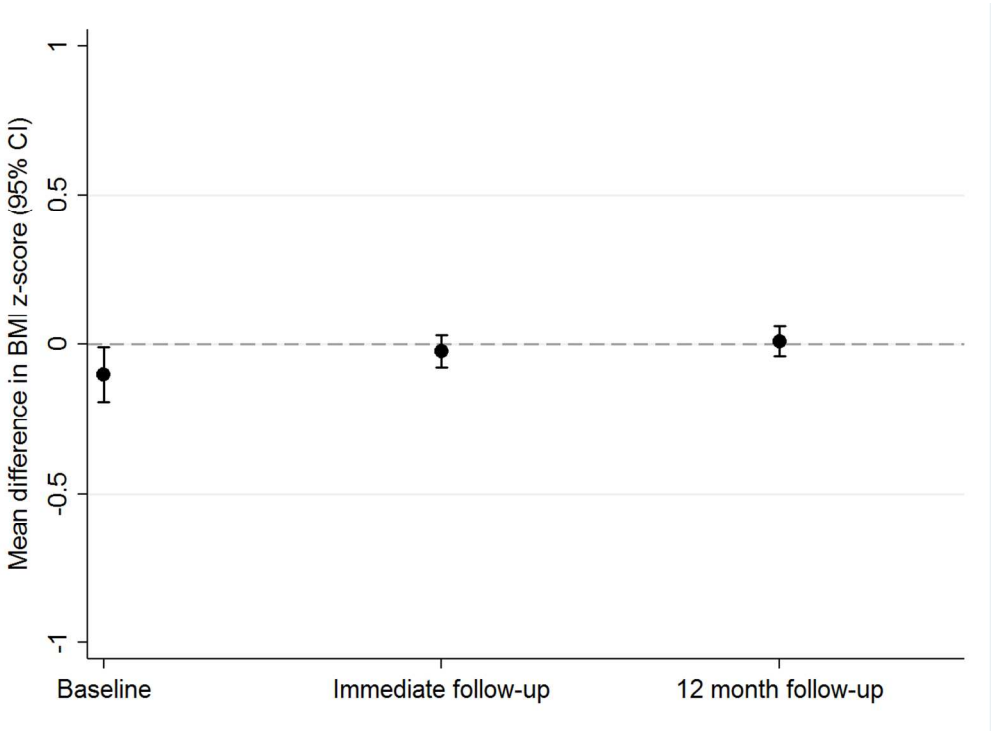


Figure 2 i. Body mass index z-score (as a continuous variable)

113x82mm (300 x 300 DPI)

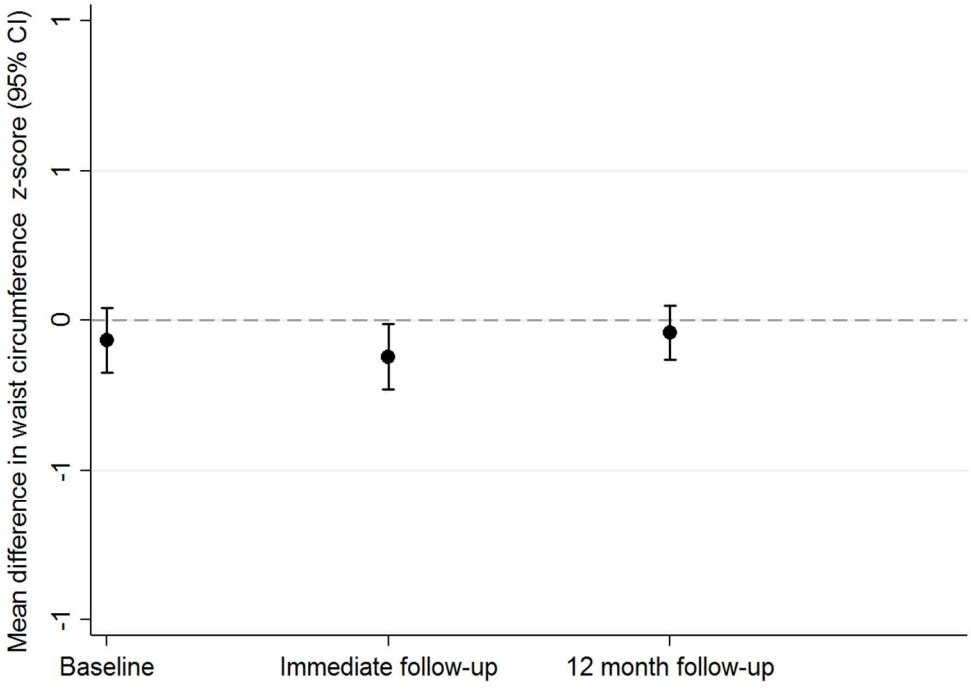


Figure 2 j. Waist circumference z-score (as a continuous variable)

113x82mm (300 x 300 DPI)

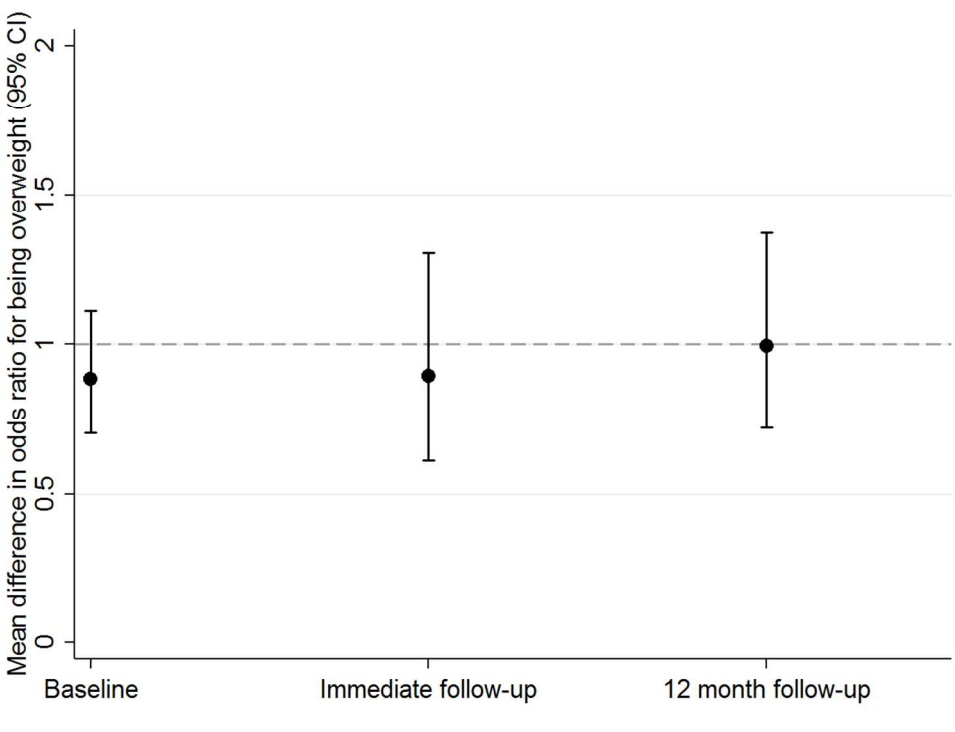


Figure 2 k. General overweight or obesity (based on BMI measurements)

113x82mm (300 x 300 DPI)

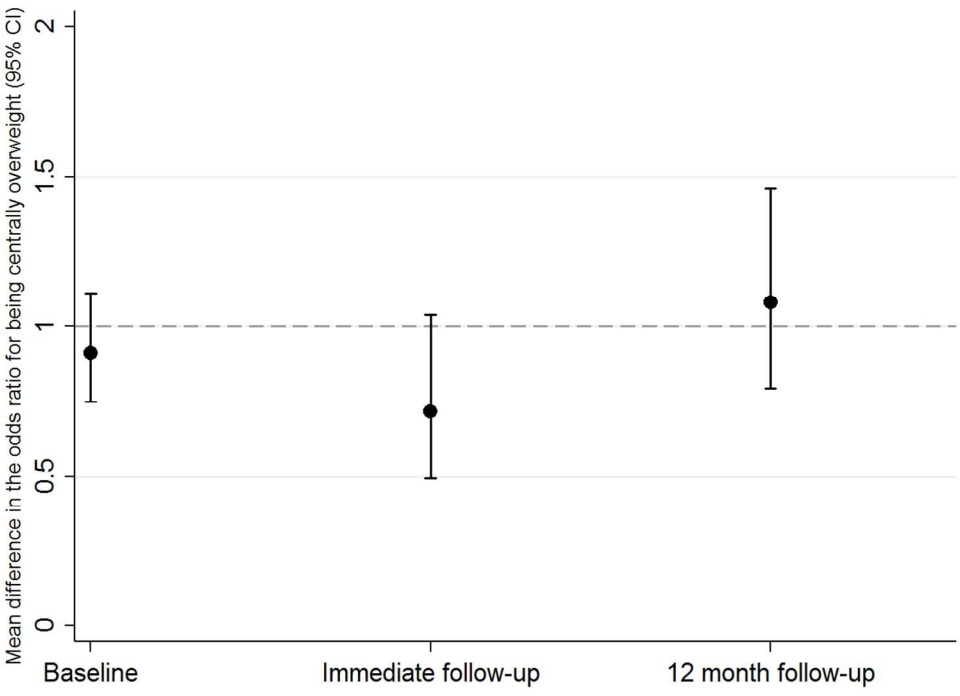


Figure 2 l. Central overweight / obesity based on waist circumference measurements
Footnote to all parts of Figur
113x82mm (300 x 300 DPI)

Supplementary Table S1: Sensitivity analysis: intention to treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes 12 months post-intervention. Numbers vary by outcome as indicated in the table. In these analyses participants were only included for each outcome if they had a baseline and a follow-up measurement of that outcome.

Outcome	Main comparison between the two groups (Intervention versus Control)		
Primary / secondary	**Np**	**Difference in means or odds ratio (95%CI)**	**p-value**
Continuous outcomes			
Time spent in MVPA (minutes per day)	1000	3.05 (-1.33, 7.44)	0.17
Time spent in sedentary behaviour (minutes per day)	1000	2.21 (-8.28, 12.71)	0.68
Servings of fruit and vegetables (number per day)	1953	0.02 (-0.15, 0.19)	0.83
Time spent screen-viewing (minutes per day weekday)	1965	-10.53 (-26.1, 5.05)	0.19
Time spent screen-viewing (minutes per day Saturday)	1965	-17.3 (-33.71, -0.88)	0.04
Body mass index (z(sd)-score)	1563	0 (-0.05, 0.04)	0.95
Waist circumference (z(sd)-score)	1748	-0.03 (-0.12, 0.05)	0.47
Servings of snacks (number per day)	1953	-0.13 (-0.3, 0.04)	0.13
Servings of high fat foods (number per day)	1953	-0.13 (-0.25, 0)	0.04
Servings of high energy drinks (number per day)	1953	-0.21 (-0.4, -0.02)	0.03
Binary outcomes			
Generally overweight/obese	1563	0.83 (0.56, 1.22)	0.35
Centrally overweight/obese	1748	1.01 (0.73, 1.4)	0.93

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed), BMI: body mass index, WC: waist circumference, F&V fruit and vegetables.

Supplementary Table S2: Sensitivity analysis: intention to treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention. In these analyses participants were only included for each outcome if they had a baseline and a follow-up measurement for all three primary outcomes. Numbers included are identical for the three primary outcomes (N = 757) but can vary by outcome for secondary outcomes (though none of these can be higher than 757) as indicated in the table.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	757	1.28 (-3.22, 5.78)	0.58
Time spent in sedentary behaviour (minutes per day)	757	0.60 (-10.44, 11.63)	0.92
Servings of fruit and vegetables (number per day)	757	-0.13 (-0.34, 0.09)	0.26
Time spent screen-viewing (minutes per day weekday)	757	0.20 (-17.54, 17.94)	0.98
Time spent screen-viewing (minutes per day Saturday)	757	-8.46 (-28.49, 1.56)	0.41
Body mass index (z(sd)-score)	682	0.00 (-0.06, 0.07)	0.80
Waist circumference (z(sd)-score)	728	-0.01 (-0.12, 0.09)	0.90
Servings of snacks (number per day)	757	-0.13 (-0.38, 0.13)	0.33
Servings of high fat foods (number per day)	757	-0.13 (-0.33, 0.07)	0.19
Servings of high energy drinks (number per day)	757	-0.12 (-0.37, 0.12)	0.32
Binary outcomes			
Generally overweight/obese	680	1.09 (0.64, 1.85)	0.76
Centrally overweight/obese	728	11.35 (0.81, 2.23)	0.25

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed), BMI: body mass index, WC: waist circumference, F&V fruit and vegetables.
Missing baseline data for secondary outcomes (once those with missing baseline primary outcomes are excluded) were managed as in the main analyses.

Supplementary Table S3: Sensitivity analysis: intention to treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention, with missing data for either baseline or follow-up measure of an outcome assumed to be 10% healthier than the average value in the study sample.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	2052	0.74 (-1.59, 3.07)	0.53
Time spent in sedentary behaviour (minutes per day)	2052	1.78 (-4.63, 8.20)	0.59
Servings of fruit and vegetables (number per day)	2052	0.01 (-0.16, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	2052	-10.74 (-26.30, 4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	2052	-16.03 (-32.82, 0.76)	0.06
Body mass index (z(sd)-score)	2052	0.01 (-0.04, 0.06)	0.70
Waist circumference (z(sd)-score)	2052	-0.02 (-0.11, 0.06)	0.56
Servings of snacks (number per day)	2052	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	2052	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	2052	-0.20 (-0.39, -0.01)	0.04
Binary outcomes			
Generally overweight/obese	2052	0.98 (0.76, 1.26)	0.87
Centrally overweight/obese	2052	1.05 (0.77, 1.43)	0.78

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed), BMI: body mass index, WC: waist circumference, F&V fruit and vegetables.
In these analyses participants all participants are included (N = 2,221 (the number of participants recruited to the study). Missing baseline data is managed as in the main analyses and missing outcome data are imputed on the basis of those with missing data being 10% healthier than all participants in the study for a given outcome.

Supplementary Table S4: Sensitivity analysis: intention to treat analyses of the effect of AFLY5 intervention on primary and secondary outcomes assessed 12 months post-intervention, with missing data for either baseline or follow-up measure of an outcome assumed to be 10% less healthy than the average value in the study sample.

Outcome	Main comparison between the two groups (Intervention versus Control)		
	Np	Difference in means or odds ratio (95%CI)	p-value
Continuous outcomes			
Time spent in MVPA (minutes per day)	2052	1.04 (-1.18, 3.26)	0.36
Time spent in sedentary behaviour (minutes per day)	2052	-0.72 (-6.39, 4.95)	0.80
Servings of fruit and vegetables (number per day)	2052	0.01 (-0.16, 0.17)	0.94
Time spent screen-viewing (minutes per day weekday)	2052	-10.74 (-26.30,4.81)	0.18
Time spent screen-viewing (minutes per day Saturday)	2052	-16.03 (-32.82, 0.76)	0.06
Body mass index (z(sd)-score)	2052	0.01 (-0.04, 0.06)	0.70
Waist circumference (z(sd)-score)	2052	-0.02 (-0.11, 0.06)	0.56
Servings of snacks (number per day)	2052	-0.11 (-0.29, 0.06)	0.19
Servings of high fat foods (number per day)	2052	-0.12 (-0.25, 0.00)	0.05
Servings of high energy drinks (number per day)	2052	-0.20 (-0.39, -0.01)	0.04
Binary outcomes			
Generally overweight/obese	2052	0.98 (0.76, 1.26)	0.87
Centrally overweight/obese	2052	1.05 (0.77, 1.43)	0.78

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

Outcomes in bold are primary outcomes ($p < 0.05$ indicates statistical significance); all others are secondary outcomes ($p < 0.01$ indicates statistical significance, after taking account of multiple testing).

All differences in means / odds ratios with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes and multi-level multivariable logistic regression was used for binary outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed), BMI: body mass index, WC: waist circumference, F&V fruit and vegetables.
In these analyses participants all participants are included (N = 2,221 (the number of participants recruited to the study). Missing baseline data is managed as in the main table and missing outcome data are imputed on the basis of those with missing data being 10% less healthy than all participants in the study for a given outcome.

Supplementary Table S5: Main intention to treat analyses of the effect of AFLY5 intervention on accelerometer-assessed outcomes during 3 valid days, separately for week and weekend days. Numbers vary by outcome as indicated in the table.

Outcome	Main comparison between the two groups (Intervention versus Control) on week days			Main comparison between the two groups (Intervention versus Control) on weekend days		
	Np	Difference in means (95%CI)	p-value	Np	Difference in means (95%CI)	p-value
Time spent in MVPA (minutes per day)	1627	2.47 (-1.37, 6.32)	0.21	972	3.26 (-3.62, 10.14)	0.35
Time spent in sedentary behaviour (minutes per day)	1627	1.87 (-8.51, 12.24)	0.72	972	3.07 (-10.91, 17.06)	0.67

Np: number of participants; MVPA: moderate or vigorous physical activity; CI: confidence interval

All differences in means with their 95%CI have been estimated using a multi-level model to account for clustering (non-independence) among children from the same school. Multi-level multivariable linear regression was used for effects of the intervention on continuously measured outcomes.

The following baseline / school stratifying covariables were included: age, gender, the baseline measure of the outcome under consideration, school involvement in other health promoting behaviours, school area level deprivation.

MVPA: moderate and vigorous physical activity (accelerometer assessed), SB: sedentary behaviour (accelerometer assessed).

In these analyses, participants were only included for each outcome if they had a follow-up measurement of that outcome. For partial missing baseline data we used an indicator variable as describe by White & Thompson,(1) which means for each outcome participants are included even if they do not have a baseline measurement.

Only participants included in the main analyses (i.e. with at least 3 valid days of accelerometer data) are included in this sensitivity analysis.

References

1. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med. 2005;24(7):993-1007. Epub 2004/12/01.

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[1,2]	See table 2	5-6
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	8
	2b	Specific objectives or hypotheses	Whether objectives pertain to the the cluster level, the individual participant level or both	9
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	9
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		**None so no reporting (protocol is published)**
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	9
	4b	Settings and locations where the data were collected		9-10 & 13-14
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	11-13
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and	Whether outcome measures pertain to the cluster level, the individual participant level or both	13 & Box 1

		when they were assessed		
	6b	Any changes to trial outcomes after the trial commenced, with reasons		**None so no reporting (protocol is published)**
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or *k*), and an indication of its uncertainty	14
	7b	When applicable, explanation of any interim analyses and stopping guidelines		N/A
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		10-11
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	10-11
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	10-11
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	10-11
	10b		Mechanism by which individual participants were included in clusters for the purposes of the	10-11

			trial (such as complete enumeration, random sampling)	
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	10
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		13 & 14
	11b	If relevant, description of the similarity of interventions		N/A
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	14-16
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		14-16
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	16 & Figure 1
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	16 & Figure 1
Recruitment	14a	Dates defining the periods of recruitment and follow-up		13
	14b	Why the trial ended or was stopped		N/A
Baseline data	15	A table showing baseline	Baseline characteristics for the	Table 1; 35-36

		demographic and clinical characteristics for each group	individual and cluster levels as applicable for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	Table 1; 35-36 Table 2; 37-38 Table 3; 39-40
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	37-40
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		37-40
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		Supplementary material
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[3])		**N/A – intervention was integrated into school teaching curriculum**
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		19-20
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	19-20
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		20

Other information			
Registration	23	Registration number and name of trial registry	6
Protocol	24	Where the full trial protocol can be accessed, if available	Referenced throughout the paper – reference numbers 9 and 17 in reference list which starts on page19
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	2-3 & 6

** Note: page numbers optional depending on journal requirements*

Table 2: Extension of CONSORT for abstracts[1,2] to reports of cluster randomised trials

Item	Standard Checklist item	Extension for cluster trials
Title	Identification of study as randomised	**Identification of study as cluster randomised**
Trial design	Description of the trial design (e.g. parallel, cluster, non-inferiority)	
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	**Eligibility criteria for clusters**
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	**Whether objective or hypothesis pertains to the cluster level, the individual participant level or both**
Outcome	Clearly defined primary outcome for this report	**Whether the primary outcome pertains to the cluster level, the individual participant level or both**
Randomization	How participants were allocated to interventions	**How clusters were allocated to interventions**
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants randomized to each group	**Number of clusters randomized to each group**
Recruitment	Trial status[1]	
Numbers analysed	Number of participants analysed in each group	**Number of clusters analysed in each group**
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	**Results at the cluster or individual participant level as applicable for each primary outcome**
Harms	Important adverse events or side effects	
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register	
Funding	Source of funding	

[1] Relevant to Conference Abstracts

REFERENCES

1 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008, 371:281-283

2 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG at al (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 5(1): e20

3 Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.