Active-6

Assessing the impact of COVID-19 on the physical activity of Year 6 children and their parents: Identifying scalable actions to mitigate adverse impacts and provide rapid evidence to policy makers

This project is funded by the National Institute for Health Research (Project number 131847)

Statistical Analysis Plan

Version 0.1	16th April 2021	Ruth Salway	Initial draft
Version 0.2	25th May 2021	Ruth Salway	V0.2
Version 0.3	10th June 2021	Ruth Salway	Addressing comments from RJ + BT
Version 0.4	17th June 2021	Ruth Salway	Addressing comments from FDV
Version 0.5	13th August 2021	Ruth Salway	Incorporating comments from SMG
Version 1.0	28th September 2021	Ruth Salway	Approved by SSC 21st September 2021
Version 1.1	28th January 2022	Ruth Salway	Change to proposed main analysis to include adjustment for first data collection period[1]: 4.2.2 additional paragraph added 4.3.1 first data collection period added to list of adjustment variables 4.7.2 added sensitivity analysis

[1] This change was first proposed in November 2021 after analysing the interim data (see report September 2021) and seeing improved response rates in the second data collection period, but before the final Wave 1 dataset was constructed or analysed.

This project is funded by the National Institute for Health Research [Public Health Research Programme] (project number 131847). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR PHR Programme or the Department of Health.

Table of Contents

1. Introduction 5
2. Synopsis of study design and procedures 5
   2.1 Background and rationale 5
   2.2 Objectives and aims 5
      2.2.1 Primary Objective 5
      2.2.2 Secondary Objectives 5
   2.3 Study design 6
      2.3.1 Comparator group 6
      2.3.2 Schools 6
      2.3.3 Children and parents 6
   2.4 Data collection 7
      2.4.1 Accelerometer data 7
      2.4.2 Questionnaires 7
      2.4.3 Child height/weight measurements 7
      2.4.4 COVID-related variables 7
   2.5 Sample size 7
   2.6 0Outcome measures 8
      2.6.1 Primary outcome 8
      2.6.2 Secondary outcomes 8
   2.7 Interim analyses 8
3. general analysis considerations 9
   3.1 Analysis population 9
   3.2 Data cleaning 9
   3.3 Derived variables 9
   3.4 Bias and confounding 10
      3.4.1 Differences in pre and post COVID populations 10
      3.4.2 Secular Trend 10
      3.4.3 Confounding 11
      3.4.4 Adjustment for other covariates 11
   3.5 Missing data 11
   3.6 Reporting estimates and statistical significance 11
   3.7 Statistical software 12
4. Analysis 12
   4.1 Interim Analyses 12
   4.2 General Analysis Methods 12
      4.2.1 Model Checking/diagnostics 12

4.2.2 Adjustment for first data collection period ... 12
4.3 Acute effects of COVID-19: Comparison between T0 and T1 ... 12
4.3.1 Primary analysis: comparison of children's weekday MVPA between T0 and T1 ... 12
4.3.2 Secondary analyses ... 13
4.4 Chronic effects of COVID-19: Comparison between T0 and T2 ... 13
4.5 Effects of COVID-19 change over time: Comparisons across T0, T1 and T2 ... 13
4.6 Mediation Analysis ... 14
4.7 Sensitivity Analyses ... 14
4.7.1 Sensitivity to imputed data ... 14
4.7.2 Sensitivity to Wave 1 first data collection period adjustment ... 14
4.7.3 Differences in accelerometer wear duration protocol between B-Proact1v and Active 6 ... 15
4.7.4 Incomplete data on BMI ... 15
5. Sample Tables and Figures (subject to change, and excludes exploratory analyses) ... 16
5.1 Interim Reports ... 16
5.2 Populations ... 16
5.3 Acute effects of COVID: comparisons between T0 and T1 ... 16
5.4 Mid-term effects of COVID: comparisons between T0 and T2 ... 16
5.5 Effects of COVID-19 change over time: Comparisons across T0, T1 and T2 ... 16
5.6 Mediation Analysis ... 16
5.7 Sensitivity Analyses ... 16
6. References ... 25

List of Abbreviations

Acronym	Details
BIC	Bayesian Information Criterion (indicator of model fit)
BMI	Body mass index
CI	Confidence interval
CMO	Chief Medical Officer
ICC	Intra-cluster correlation
IMD	Index of multiple deprivation
IQR	Inter-quartile range
MAR	Missing at random
MVPA	Moderate to vigorous physical activity
PA	Physical activity
SAP	Statistical analysis plan
SD	Standard deviation
SEP	Socio-economic position
T0	Time 0: Pre-COVID-19 comparator group (B-Proact1v Y6 data)
T1	Time 1: Wave 1
T2	Time 2: Wave 2
zBMI	Standardised body mass index z-score

1. INTRODUCTION

This Statistical Analysis Plan (SAP) details the rules proposed and the presentation that will be followed, as closely as possible, when analysing and reporting the main results from Active-6. This document is written by Dr Ruth Salway who will conduct the statistical analysis for this study, and acts as a stand-alone document so that any statistician has the required information needed to perform the statistical analysis. The information in this document is adapted from the funder-approved study protocol.

The purpose of the plan is to:

- ensure that the analysis is appropriate for the aims of the study, reflects good statistical practice, and that interpretation of a priori and post hoc analyses respectively is appropriate.
- explain in detail how the data will be handled and analysed to enable others to perform the actual analysis in the event of sickness or other absence

Additional exploratory or auxiliary analyses of data not specified in the protocol are permitted but fall outside the scope of this analysis plan (although such analyses would be expected to follow Good Statistical Practice).

The analysis strategy will be made available if required by journal editors or referees when the main papers are submitted for publication. Additional analyses suggested by reviewers or editors will, if considered appropriate, be performed in accordance with the Analysis Plan, but if reported the source of such a post-hoc analysis will be declared.

Amendments to the statistical analysis plan will be described and justified in the final report of the study.

2. SYNOPSIS OF STUDY DESIGN AND PROCEDURES

2.1 Background and rationale

Physical activity is important for health among children and their parents. The COVID-19 pandemic and the associated social changes, such as the lockdown and subsequent phased release, have had a marked impact on physical activity patterns but without data it is impossible to know the magnitude of these effects, how they may differ by demographic variables or the possible solutions for those challenges. The aim of this project is to examine the impact of changes that have resulted from COVID-19 on the physical activity of Year 6 children and their parents. Specifically, we want to know if there is an acute change in physical activity and if that change is maintained, thereby indicating a chronic impact. This study uses a repeated cross-sectional design, with data on Year 6 children and their parents collected in two waves (T1: May 2021 – December 2021 and T2: January 2022 – July 2022) and compared to those who took part in the third phase of the B-Proact1v study between March 2017 and June 2018 (T0). Throughout this analysis plan we interpret the difference before and after the COVID-10 pandemic as 'the effect of the pandemic'.

2.2 Objectives and aims

The Active-6 Protocol describes seven inter-linked objectives. This SAP addresses the five quantitative objectives; separate documents will be produced for the qualitative and health economics objectives.

2.2.1 *Primary Objective*

(Aim A) To assess the **acute** effect of the COVID-19 pandemic on the Weekday MVPA of Year 6 children by comparing new data (T1) to T0 data sampled from the same schools 3 years earlier. We will also assess whether effects differ by socioeconomic position and/or gender.

2.2.2 *Secondary Objectives*

B. To determine if there are differences between the physical activity and sedentary behaviour of Year 6 children and their parents when compared to T0 data sampled from the same schools 3 years earlier for the following secondary outcomes at the Time 1 assessment:
    i. Parent accelerometer measured weekday minutes of weekday MVPA
    ii. Child accelerometer measured weekend minutes of MVPA

iii. Parent accelerometer measured weekend minutes of MVPA
iv. Child accelerometer measured weekday sedentary minutes
v. Child accelerometer measured weekend sedentary minutes
vi. Parent accelerometer measured weekday sedentary minutes
vii. Parent accelerometer measured weekend sedentary minutes

C. To assess the **chronic** effects of the COVID-19 pandemic on primary and secondary outcomes when data are sampled from the same schools for the T2 assessment.

D. To examine the extent to which differences in total volume of physical activity and sedentary time at both T1 and T2 are explained by the variation in the frequency that the child is active, child physical activity enjoyment and motivation, mode of travel to school, child screen-time, after-school club attendance, parent physical activity motivation and self-efficacy., and to examine the specific impact of school walking, cycling and play provision, curriculum physical activity, school grounds and school physical activity policies on differences in physical activity.

E. To produce rapid interim reports from the project to UK policy makers to inform the development of effective strategies to increase physical activity in groups who may have been disproportionately affected by changes due to COVID-19.

2.3 Study design

Active-6 is a natural experiment, comparing data in a pre-COVID comparator group to new data collected at two time periods after the COVID-19 pandemic was declared. This study uses a repeated cross-sectional design, with data on Year 6 children and their parents collected in two waves (T1: May 2021 – December 2021 and T2: January 2022 – July 2022) and compared to those who took part in the third phase of the B-Proact1v study between March 2017 and June 2018 (T0). Although collected in the same schools and year groups, the children at each time point differ. For the interpretation of the results, we assume that any observed differences between time points result from the covid-19 pandemic (2020-21) and not from a secular trend over the 2018-2022 time period. Given the ubiquitous incidence of the pandemic no control group is available to test that assumption.

2.3.1 *Comparator group*

The pre-COVID comparator group, T0, will be the Year 6 data previously collected from the B-Proact1v study. These participants were recruited from 50 schools with a consent rate of 63% of children within schools and showed considerable variation in both parent and child physical activity and in SEP. Data is available on a total of 1296 child and parent pairs.

2.3.2 *Schools*

Data will be collected in the same primary schools as T0. All schools are located within the greater Bristol area. During the consent process for the B-Proact1v project, all schools agreed to be re-contacted in relation to future waves of the project. Schools that take part will be asked to sign a school study agreement in which they provide approval to join the study.

2.3.3 *Children and parents*

Participants will be Year 6 children (aged 10-11) and one parent per family. All Year 6 children and their parents will be invited to join the study. In families where there are two or more Year 6 children, all willing children will be invited and eligible to take part. Only one parent per family will be asked to provide data. Written informed parental consent to take part will be obtained prior to data collection. A person may withdraw from the study at any point and not provide any further data from that point.

2.4 Data collection

Data will be collected at two time points T1 and T2. At each time point, pupils and one parent will wear a waist-worn ActiGraph wGT3X-BT accelerometer for 7 days to estimate levels of PA and sedentary time, and complete a questionnaire. These devices will be sent home from the school in family-specific packs that also include details on how to complete the parent and child questionnaires. Due to logistical issues with data collection during the pandemic, these may not be the same seven days for all pupils within the same school, but the data are intended to proxy for habitual physical activity.

2.4.1 *Accelerometer data*

Recording of accelerometer data begins the day after the child/parent receives it to remove any initial atypical activity. Accelerometer data will be processed using the same processing protocols as used in the B-Proact1v study. Specifically, a valid day of data will be defined as at least 500 minutes of data, after excluding intervals of ≥60 minutes of zero counts allowing up to two minutes of interruptions (1). Note that any days where accelerometer data is missing (for example, if the accelerometer is not worn at all) is not a valid day. Child data will be characterised as sedentary, light or MVPA using Evenson population-specific cut points for children (2). Mean minutes of weekday and weekend MVPA will then be derived as well as mean minutes of sedentary time per day. A comparable process will be undertaken for the parents, but the data will be analysed using the Troiano adult accelerometer cut points (3). We are not considering step counts as an outcome measure.

2.4.2 *Questionnaires*

Child and parent questionnaires assessing a range of variables that may help to explain changes in physical activity and sedentary time will be completed online after consent has been obtained. These include demographics, the frequency with which that the child is active in different settings, child physical activity enjoyment and motivation, mode of travel to school, child screen-time, after-school club attendance, parent physical activity motivation and self-efficacy. The questionnaires will be administered via REDCap online and data will be automatically uploaded to a central University of Bristol server. Parents will be sent up to three reminders via email to complete the survey. If parents express a preference for a paper option, we will send them a form along with the accelerometers.

2.4.3 *Child height/weight measurements*

When it is possible to collect data in schools, child height and weight will be recorded by fieldworkers to the nearest 0.1 cm and 0.1 kg, using a Seca digital scale and a Seca stadiometer. To ensure data quality, measurements will be taken twice, and if they differ a third measurement will be taken.

2.4.4 *COVID-related variables*

Schools will be asked about COVID-19-related changes such as restructured physical education and after-school provision. In addition, we will collect data on any COVID restrictions in place at the time of data collection (national and/or local), whether children had been defined as clinically vulnerable or extremely vulnerable, and the numbers of staff and children in the year group who absent at the time of data collection due to COVID-19 or self-isolating.

2.5 Sample size

This is an unbalanced cluster design, where the baseline numbers are known and fixed, and the clusters are matched across time-points. All 50 schools from T0 will be included in the analysis. A simulation-based approach (4) was used to calculate the sample size required to detect a 5-minute difference in child Weekday MVPA, comparing new data collected to the known T0 data. A 5-minute difference has been suggested as the smallest difference that is likely to have a meaningful impact on health and is slightly higher than the average 4-minute impact of accelerometer based physical activity interventions (5).

No. of schools	School response rate	Power	Mean no. pupils in schools	Pupil response rate	Minimum sample size
50	100%	98%	13.7	33%	686

45	90%	97%	13.7	33%	617
40	80%	95%	13.7	33%	549
30	60%	90%	16.6	40%	498
30	60%	84%	13.7	33%	412
25	50%	89%	20.2	48%	505
25	50%	81%	15.8	38%	394
20	40%	82%	21.2	50%	422

Simulations were performed using 10,000 iterations based on the known T0 accelerometer data (complete data on 1125 pupils across 50 schools, mean weekday MVPA = 61mins, SD=23.1 mins) and simulated follow-up data for a mean weekday MVPA of 56 mins, based on a between school intra-cluster correlation of 0.15, total SD of 23.1, and within-school variability (school-level random effect) all estimated from the T0 data. The table shows estimates based on between 20 and 50 schools taking part in the study and gives the response rates, minimum number of pupils in each school and corresponding overall sample size to provide at least the specified power to detect a difference at a 5% two-sided significance level. For all scenarios, the overall number of children in schools is assumed to be lower than the average of 23 recruited in the T0 data, with the reduction in sample size proportionate across schools. The data in the table show that even if only 50% of schools take part with 16 pupils per school recruited (a response rate of 38% compared to 53% achieved at T0) for a sample size of 394 we would still have 80% power to detect a meaningful difference. These estimates are intentionally very conservative to allow for the possibility of lower numbers of schools and pupils taking part because of COVID-19. The power depends primarily on the school response rate, with pupil response rate to a lesser extent; if more schools participate this will increase the power (for example, an extra 5 schools with similar pupil response rate will achieve approximately 90% power).

2.6 Outcome measures

2.6.1 *Primary outcome*

The primary outcome will be child waist-worn-accelerometer-measured weekday minutes of moderate to vigorous physical activity (MVPA) at T1.

2.6.2 *Secondary outcomes*

There are 7 secondary outcomes for the T1 assessment:

1. Parent weekday minutes of weekday MVPA (Accelerometer)
2. Child weekend minutes of MVPA (Accelerometer)
3. Parent weekend minutes of MVPA (Accelerometer)
4. Child weekday sedentary minutes (Accelerometer)
5. Child weekend sedentary minutes (Accelerometer)
6. Parent weekday sedentary minutes (Accelerometer)
7. Parent weekend sedentary minutes (Accelerometer)

All these variables (plus the primary weekday child MVPA) will also be assessed at T2 and these will also be treated as secondary outcomes.

2.7 Interim analyses

In line with secondary aim E to provide timely evidence to policy makers, we will conduct rapid interim analyses. These will occur after data collection in every 12 schools is completed (approximately every 2-3 months) and at the end of each Wave (December 2021 and July 2022), and be shared with schools and policy makers. The proposed analyses are described in more detail below. Depending on the school participation rate, this will result in a maximum of four interim reports for each wave.

3. GENERAL ANALYSIS CONSIDERATIONS

3.1 Analysis population

The population for child-based analysis is all children with at least two valid weekdays and one valid weekend day of accelerometer data (missing data in other days will be included in accelerometer-derived measures as these are based on averages over valid days). For adults-based analyses that do not include any child measures the population is all adults with at least two valid weekdays and one valid weekend day of accelerometer data.

3.2 Data cleaning

Data will be cleaned using the same procedures as in B-Proact1v. We will perform internal consistency checks and plausibility checks to flag any inconsistencies, implausible or extreme values.

3.3 Derived variables

Derived variables will be calculated in the same way as in B-Proact1v where possible to allow direct comparison. The exceptions are child and parent screen-viewing and mode of travel to school, where Active-6 questions were simplified as a result of issues identified in the B-Proact1v questions.

Child age	Age will be calculated from parent-reported child date of birth and date of accelerometer measurement
Parent ethnicity	Parent ethnicity will be self-reported in one of 16 categories and regrouped into *White*, *Black/African/Caribbean/Black British*, *Asian/Asian British*, *Mixed*, *Other*
Highest household education	The highest educational qualification within the household will be reported by the parent and coded as: *Up to GCSE/O level or equiv*, *A level/NVQ or equiv*, *Degree/HND or equiv*, *Higher degree (MSc/PhD) or equiv*. For some reporting (eg interim reports) this will be regrouped as '*up to A level equiv*' and '*degree equiv or higher*'.
Index of Multiple Deprivation (IMD)	IMD score will be calculated from parent-reported home postcode. School IMD will be calculated from school postcode.
Parent BMI	Parent BMI will be calculated from self-reported height and weight
Parent physical activity motivation	Parents will complete the behavioural regulations in exercise questionnaire (BREQ-2) (6) assessing 5 subscales: external motivation, introjected motivation, identified motivation, intrinsic motivation and amotivation. We will calculate each subscale as the average of the items and the composite autonomous (mean of intrinsic & identified) and controlled (mean of introjected and external) motivation.
Child SV	In Active-6 parents will report total child screen-viewing for schoolwork and leisure and tv-viewing (on any device) for both weekdays and weekends. In B-Proact1v, TV, computer, phone and games console screen-viewing were reported separately in categories. These will be assigned to the midpoint of the category, summed and categorised into the same groups as Active-6 (less than 1h, 1-2h, 2-3h, 3-4h, 4-5h, 5+h) to produce total screen-viewing and TV viewing for weekdays and weekends separately.
Parent SV	Parent screen-viewing data will be treated in the same as above.
Number of siblings	Number of siblings for each child will be calculated from parent-reported number of children.
Child frequency activity score	Children will be asked how often they attend a sport club at school, outside of school, play outside, play at home or go for a walk etc with family. The first four questions were asked in B-Proact1v, and will be summed in the same way to create a child frequency of activity score.
Active after-school clubs	Children will report whether they attend an active after-school club on each day of the week. These will be summed to calculate the number of active after-school clubs attended.

Travel to school	Children will report how they usually travel to and from school as either walk/cycle/scoot (active travel) or bus/car (non-active travel). This will be combined as most days non-active travel/most days active travel in one direction/ most days active travel in both directions. In B-Proact1v children were asked each mode of transport (walk/cycle/scoot/ bus/car) for each day of the week in each direction. This will be recoded as mostly active travel to/from school if the child uses walk/cycle/scoot on 3 or more days and then combined as above.
Child BMI	When COVID restrictions allow, height and weight will be measured by fieldworkers in schools, and child BMI calculated. These will converted to age- and sex-specific standard deviation scores based on UK reference curves (7)
Mean weekday MVPA, light activity and sedentary time	Physical activity will be measured using waist-worn ActiGraph accelerometers worn over 7 days. Periods of ≥ 60 minutes of zero counts will be recorded as “non-wear” and removed and a valid day is defined as 500 minutes of data between 6am and 12pm. Minutes of daily MVPA, light activity and sedentary time will be estimated using Evenson cut points (2) for children and Troiano cut points (3) for parents. Weekday means will be calculated for those participants providing at least two valid weekdays, by taking the average of all valid weekdays.
Mean weekend MVPA, light activity and sedentary time	Daily MVPA, light activity and sedentary time will be estimated as described above, and weekend means will be calculated for those participants providing at least one valid weekend day, by taking the average of all valid weekend days.
Mean weekly MVPA, light activity and sedentary time	Daily MVPA, light activity and sedentary time will be estimated as described above, and means will be calculated for those participants providing at least two valid weekdays and one valid weekend day, by taking the average of all valid days.
Child/parent meets CMO PA guidelines	A binary variable for whether the child meets the current 2019 CMO PA guidelines (8) of a daily average of at least 60 minutes of MVPA, average across the week. The child’s mean daily MVPA from above will be coded as 0 if it is less than 60 mins (does not meet CMO PA guidelines) or 1 if it is 60 mins or greater (does meet CMO PA guidelines). Parents meeting CMO PA guidelines for adults will be calculated to be consistent with B-Pract1v data as a daily average of at least 30 minutes of MVPA, averaged across the week.

3.4 Bias and confounding

3.4.1 *Differences in pre and post COVID populations*

The study design is matched on school and child age, and since Year 6 school-level distributions of key demographics are unlikely to change substantially over three years, this will reduce bias due to differences in pre- and post-COVID populations. However, there may be bias in pre- and post-COVID samples due to differential response rates both for pupils and for schools.

To determine whether there is imbalance, demographic data (child gender, child age, household education, IMD, parent age, parent gender and parent ethnicity) at pupil and school level will be compared between T0, T1 and T2 by reporting means (SD), medians (Inter-quartile-range; IQR) or number (%) depending on the nature of the data and the respective distribution. Note that these variables are included as covariates in all adjusted regression models, so any imbalance is accounted for in the model. To explore any bias due to different underlying populations, we will repeat the main analyses restricting data to the same school at each time point.

3.4.2 *Secular Trend*

We are unable to adjust for any background trend in PA over the 2018-2022 time period, due to lack of a control group for the pandemic. We will interpret any observed differences between waves as due to the COVID-19 pandemic (2020-21) rather than from a secular trend, but include this a possible alternative in discussion and a limitation of the analyses.

3.4.3 *Confounding*

All adjusted models will be adjusted for confounders known to be associated with children's physical activity: child gender, household education and child zBMI (where available). Parent models will be adjusted for parent age, gender, BMI and household education.

3.4.4 *Adjustment for other covariates*

All analyses will be adjusted for accelerometer wear time and hours of daylight at time of data collection. We will also adjust for seasonality using harmonic sine/cosine functions, a method used previously with the B-Proact1v data (9). We will not adjust for day order as children receive accelerometers on different days, and activity estimates are averaged over multiple days.

3.5 Missing data

We will report the extent of missing data in accelerometer data and other key variables. We will investigate patterns of missingness for any variable with more than 10% missing data to determine if a missing at random (MAR) assumption is valid.

One key source of missing data will be child zBMI in Wave 1 due to not being able to measure height/weight for some participants due to COVID restrictions. These data will be missing at school level rather than due to individual characteristics, and as it is likely that this will depend on the order in which schools are recruited rather than pupil demographics, a MAR assumption is reasonable; this will be checked by looking at missing data patterns in relation to school demographics. To maximise information and reduce school-level bias, we will use multilevel joint modelling multiple imputation via the package jomo in R (10) to impute missing zBMI, using child, parent and school data and retaining the multilevel structure. The model will include all variables associated with missingness and be stratified by child gender. Child height and weight will be imputed directly, with BMI and zBMI computed passively. If it is not possible to collect any zBMI data in Wave 1, then zBMI will excluded from analyses. We anticipate that fieldwork will be possible for Wave 2, but if there are further restrictions, we will apply the same approach.

As multiple imputation will be necessary for zBMI at T1, we will impute any other missing data on covariates (but not accelerometer-derived outcome measures) at the same time. At T2, we will use multiple imputation if any non-outcome variable has more than 10% missingness. Sensitivity analyses will be conducted to compare all imputed analyses to complete case analyses.

Data may be missing for whole schools if they participate at one timepoint but not another. We will report school-level characteristics (size, IMD) and pupil demographics (% female, % white, % household education categories, % overweight/obese, mean weekday/weekend MVPA/light/sedentary, % pupils meeting PA guidelines) for the 50 schools at T0 and participating schools at T1 and T2. Summaries and interim reports will be reported for participating schools only, but an advantage of the proposed multilevel modelling for the main analyses is that we can include data from schools at any time point. To explore any bias due to different underlying populations, we will repeat the main analyses restricting data to the same school at each time point.

3.6 Reporting estimates and statistical significance

For the main analyses comparing primary and secondary outcomes before and after COVID-19, we will conduct two-sided hypothesis tests for the difference, at a 5% significance level and report the p-value as well as the estimate and 95% confidence interval. For exploratory analyses we will present estimates and 95% confidence intervals and interpret with caution. While p-values will be reported, as they are often requested by reviewers, we will interpret findings based on the magnitude and precision of estimates, rather than statistical significance.

3.7 Statistical software

A mixture of statistical software will be used depending on the analysis. Data cleaning and descriptive tables will be produced in Stata (v15 or higher) and all multilevel models will be run in MLwiN (v3.05 or higher) via the Stata command runmlwin. Multiple imputation for multilevel models will be undertaken in R (v 4.0.5 or higher) via the package jomo. Generalised additive mixed models will be run in R via the package mgcv.

4. ANALYSIS

4.1 Interim Analyses

We will produce a 2-page comparison sheet describing the levels of physical activity in children and parents (mean and standard deviation of minutes of sedentary, light and MVPA and the percentage meeting CMO physical activity guidelines) in the schools and the comparable data for the same schools at T0. Data will be aggregated across schools and presented for all participants and by gender and socio-economic sub-groups. Based on the distribution of household education in B-Proact1v (GCSE: 21%, A level: 28%, degree: 35%; higher degree: 16%) we will report household education in two groups: *Up to A level* or equivalent and *Degree level equivalent or higher* to avoid small samples in subgroups.

These data will be shared with schools and policy makers after every 12 schools (approximately 3-4 months) with the amount of data shared becoming greater over time. These analyses are purely descriptive and may be limited in generalisability, especially initially. However, as these will still be useful to policy makers, they will be accompanied with appropriate warnings and caveats. To preserve data confidentiality, no data will be reported for gender or education subgroups with sample sizes less than 10.

4.2 General Analysis Methods

We will employ a repeated cross-sectional analysis, matched on schools, comparing the T1 data with the T0 data, using linear multilevel models with children nested within time periods (0, 1 or 2) nested within schools. The difference in average daily weekday MVPA following COVID-19 will be estimated as the regression coefficient for an indicator variable for time period.

4.2.1 *Model Checking/diagnostics*

Previous analyses have found the primary and secondary outcomes to be approximately normally distributed and so linear models are likely to suitable. However, model assumptions will be checked via visual inspection of the residuals from the fixed part of the multi-level model and the random effects at the cluster level and if necessary, we will consider appropriate transformations.

We will not remove any outliers from the main analyses. However, if the residual plots indicate the presence of outliers, we will investigate the leverage of these points, and undertake a sensitivity analysis omitting points of high leverage.

4.2.2 *Adjustment for first data collection period*

Data collection in Wave 1 is in two time periods: May-July 2021 and September-December 2021, with all data collection being undertaken remotely in the first period, with additional restriction in schools. As COVID restrictions and data collection protocols differ by period, we will include an additional time period indicator in the main analyses. Unadjusted models will be reported as sensitivity analyses.

4.3 Acute effects of COVID-19: Comparison between T0 and T1

We will report means and 95% CIs of all primary and secondary outcomes overall and by gender and household education for T0 and T1.

4.3.1 *Primary analysis: comparison of children's weekday MVPA between T0 and T1*

Comparisons between T0 and T1 will be made using a linear multilevel model to allow for the complex clustering design, and all models will be adjusted for accelerometer wear time, hours of daylight, seasonality and first data collection period. We will fit the following models:

- Model 1: unadjusted for confounders
- Model 2: adjusted for gender and SEP (household education and IMD)
- Model 3: as model 2 but additionally adjusted for additional COVID restrictions and staff and pupil absences at point of measurement, where applicable

We will also fit a model adjusted for gender, SEP and child zBMI. However, as child zBMI will only be available at T1 for a subset of children due to COVID data collection restrictions, this will rely heavily on imputed data, and so we will report this as a supplementary analysis rather than a main analysis. Sensitivity analyses will compare result with complete case data, and Models 1 and 2 for the subset of children who have zBMI data available.

Sample sizes will not be large enough to allow stratification by gender or SEP, and previous B-Proact1v analysis has not suggested any substantial gender or SEP-specific associations with other covariates. Therefore, to assess whether effects differ by gender, we will repeat the above models including a time-gender interaction term, and test whether this is non-zero (that is, there exist differences in COVID impact between boys and girls). We will also assess whether effects differ by household education similarly; depending on sample sizes it may be necessary to use a less detailed household education variable (eg up to A level versus degree or higher). Interaction terms rather than subgroup analyses will be used due to potentially small sample sizes.

4.3.2 *Secondary analyses*

The child secondary outcomes (child weekend MVPA, child weekday sedentary time and child weekend sedentary time) will be analysed using the same modelling approach as for the primary outcome. The parent secondary outcomes (parent weekday/weekend MVPA and sedentary time) will be analysed similarly, controlling for parent confounders: parent age, parent gender, parent BMI and household education.

The same parent may appear in T0 and either T1 or T2 (or both). For example, analysis of the B-Proact1v data shows that 9% of participating children have a sibling three years younger. For analysis of parent outcomes, we will investigate the extent of this, and if necessary, use robust standard errors to account for the dependence.

4.4 Chronic effects of COVID-19: Comparison between T0 and T2

The primary and secondary analyses described above will be repeated for a comparison between T0 and T2, with the addition of the model adjusted for gender, SEP and child zBMI as a main analysis, unless further COVID restrictions during Wave 2 continue to restrict data collection by fieldworkers.

4.5 Effects of COVID-19 change over time: Comparisons across T0, T1 and T2

We will combine data from all three waves to explore how MVPA changes across time. We will fit models 1-3 as in 4.3.1 (unadjusted, adjusted for PA confounders, adjusted for PA confounders and covid restrictions). We will include gender and/or household education interaction terms in models 2 & 3 if they were found to be statistically significant in 4.3 and 4.4.

Firstly, we will use a discrete wave variable to estimate MVPA at each time point. This will use a multi-level model similar to above, but with three time points instead of two. Due to the timings of data collection, some of the children in T1 may also be repeated in T2. The multilevel model will thus be cross-classified by child and time period.

Secondly, we will extend this to use the date of data collection as a continuous measure of time, which will allow a more nuanced investigation of how any COVID effects vary with time and restrictions. This will use a generalised additive mixed model to explore nonlinear changes over time using penalised splines (11), maintaining the multilevel structure. The model is flexible and model 3 will include national and local COVID restrictions at different times so as be able to include additional infection-control measures or future COVID-19 waves if necessary, to estimate their effect on physical activity. We will use a combination of the estimated degrees of freedom and associated p-value from the model, an indicator of model fit (BIC: Bayesian Information Criterion) and plots of the spline for the time variable to assess whether any change in MVPA over time is linear.

Due to the increased complexity of these models, we will not use imputed data, which will unfortunately limit the extent to which we are able to include child zBMI. We will adopt the same approach as in 4.4.1, focusing primarily on models without zBMI that use the full data, but exploring a subset of models that include zBMI, with sensitivity analyses.

This analysis will be treated as exploratory, and so we will present estimates and 95% confidence intervals and interpret with caution. While p-values will be reported, as they are often requested by reviewers, we will interpret findings based on the magnitude and precision of estimates, rather than statistical significance.

4.6 Mediation Analysis

To understand the extent to which differences in MVPA between T0 and T2 can be explained by other factors we will undertake a mediation analysis. This analysis will focus on the long-term impact (T0 to T2) due to the unavailability of child zBMI for a large proportion of children at T1. However, we will also investigate whether similar patterns are evident in T0-T1 differences using the smaller subset of children for whom zBMI measurements are available.

We will investigate individual factors (physical activity enjoyment and motivation, mode of travel to school, screen-time and after-school club attendance), parental factors (parent physical activity, motivation) and school factors (school walking, cycling and play provision, curriculum physical activity and extra-curricular programs) and include individual and school-level confounders such as school size, deprivation, and location.

Formal mediation tests can be problematic and introduce confounding in multilevel models as mediators may act at different levels (12), and so an overall estimate of a mediation effect will not be possible. We will therefore fit a number of models and use a combination of model fit (BIC), magnitude and direction of estimates, statistical significance, and interpret the combined evidence to assess the extent to which mediating factors are present, and if so whether they could be responsible for all, a large or a small part of any reduction in MVPA. Specifically, we use:

- how MVPA differs between T0 and T2, adjusting for all confounders (gender, household education and IMD and child zBMI (analysis from 4.5)
- how the potential mediating factor differs between T0 and T2, adjusting for relevant confounders
- how MVPA differs between T0 and T2, adjusting for confounders and potential mediator
- how MVPA differs between T0 and T2, adjusting for confounders and all potential mediators

Potential mediators may be continuous or categorical; we will investigate the distributions and use transformations/grouping of small cells where necessary. In addition, we will explore school and wave-level random effects to explore how much of the total between-wave variability in MVPA is explained by unobserved individual and school-level factors, and consider the extent to which sources of unmeasured confounding may be responsible for potential mediators. Although statistical significance is one of a number of criteria to be considered in interpreting the mediation analysis, no formal hypothesis testing will be presented.

4.7 Sensitivity Analyses

A number of analyses are proposed to assess the sensitivity of the above analyses to various assumptions. These are described below. Sensitivity analyses will be presented alongside the corresponding main analyses in order for them to be compared and contrasted. As these will be exploratory in nature, 95% confidence intervals and p-values will be presented but will be interpreted with due caution.

4.7.1 *Sensitivity to imputed data*

The procedure for multiple imputation is described in section 3.5. We will repeat all analyses that use imputed data using complete case data only.

4.7.2 *Sensitivity to Wave 1 first data collection period adjustment*

To explore the effect of any differences between the first and subsequent data collection periods in Wave 1, we will repeat the main analysis without the first data collection period adjustment, and also a sensitivity analysis excluding data from the first period.

4.7.3 *Differences in accelerometer wear duration protocol between B-Proact1v and Active 6*

B-Proact1v (T0) and Active-6 (T1 & T2) differ in the number of days the accelerometer is worn (5 days at T0 and 7 at T1 & T2). In both cases, this affects weekday wear rather than weekend wear, with both protocols specifying two days of weekend data.

For analysis, the primary/secondary outcomes are derived from accelerometer data averaged across valid days. However, we will explore the extent to which the change in protocol from five days of data collection to seven days affects results by randomly picking 3 consecutive weekdays per school and restricting accelerometer data to those days only (this is equivalent to the process followed in B-Proact1v). We will recalculate and report estimates of mean weekday and overall MVPA, light and sedentary times for boys and girls, and for parents, at T1 and T2 for five days and compare with the original seven days of accelerometer data. If these estimates are substantially different (differ by more than 10%) we will repeat all main analyses for five days of accelerometer data.

4.7.4 *Incomplete data on BMI*

The main analyses at T1 will use imputed zBMI data, and these will be compared to complete case data in 4.9.1. However, for models that include zBMI, the complete case dataset will be substantially smaller and so it will be difficult to directly compare to models without zBMI. To investigate whether the subset of data with child zBMI differs from the main dataset, we will repeat other models using the zBMI subset of data.

5. SAMPLE TABLES AND FIGURES (SUBJECT TO CHANGE, AND EXCLUDES EXPLORATORY ANALYSES)

5.1 Interim Reports

Interim reports will be produced every 12 schools and at end of each wave. Tables 1&2 will be produced in Wave 1, with Tables 3 &4 in Wave 2.

Table 1	Comparison of children's physical activity between T0 and T1 aggregated over [N] schools
Table 2	Comparison of parent's physical activity between T0 and T1 aggregated over [N] schools
Table 3	Comparison of children's physical activity between T0, T1 and T2 aggregated over [N] schools
Table 4	Comparison of parent's physical activity between T0, T1 and T2 aggregated over [N] schools

5.2 Populations

Figure 1	Flow diagram
Table 5	Comparison of key pupil demographics at T0, T1 and T2
Table 6	Comparison of school-level characteristics for participating schools in T0, T1 and T2
Table 7	Summary of missing data

5.3 Acute effects of COVID: comparisons between T0 and T1

Table 8	Estimates of primary and secondary outcomes for T0 and T1
Table 9	Comparison of child weekday MVPA between T0 and T1

Tables 8 & 9 will be repeated for secondary outcomes

5.4 Mid-term effects of COVID: comparisons between T0 and T2

Tables will follow same format as acute effects, with comparisons between T0 and T2

5.5 Effects of COVID-19 change over time: Comparisons across T0, T1 and T2

Table 10	Comparison of child weekday MVPA between T0, T1 and T2
Table 11	Comparison of child weekday MVPA with continuous change over time
Figure 2	Smooth estimate of change in MVPA over time

5.6 Mediation Analysis

Table 12	Comparison of potential mediators between T0 and T2
Table 13	Association of MVPA and potential mediators

5.7 Sensitivity Analyses

Table 11	Comparison of key characteristics between imputed and observed data
Table 12	Comparison of accelerometer estimates for 7 days versus 5 days of data
Table 13	Comparison of key characteristics for children with and without zBMI measurements at T0

Also, key tables of models from 5.3 and 5.4 to be repeated for complete case data (no imputation), children with zBMI measurements, and restricting to the same schools at each time point. We will also produce key tables of models from 5.3 and 5.4 for 5 days of accelerometer data (section 4.7.2) and with high leverage points omitted (section 4.2.1), if deemed necessary (see relevant sections).

Table 1: Comparison of children's physical activity between T0 and T1 aggregated over [N] schools

		Pre-COVID Comparator group[1] (Mar 2017 - May 2018)	Wave 1 (May 2021 –[DATE])
All children		N=	N=
	Mean weekday MVPA (SD)		
	Mean weekday light (SD)		
	Mean weekday sedentary (SD)		
	Mean weekend MVPA (SD)		
	Mean weekend light (SD)		
	Mean weekend sedentary (SD)		
	Meets PA guidelines[2]: N (%)		
Child Gender			
Boys		N=	N=
	Mean weekday MVPA (SD)		
	…		
Girls		N=	N=
	…		
Household education			
	Up to A level or equivalent	N=	N=
	…		
	Degree equivalent or higher	N=	N=
	…		

[1] Comparator group of the same schools

[2] CMO recommends at least 60 minutes MVPA per day for children, averaged across the week

Table 2: Comparison of parent's physical activity between T0 and T1 aggregated over [N] schools

		Pre-COVID Comparator group[1] (Mar 2017 - May 2018)	Wave 1 (May 2021 –[DATE])
All parents		N=	N=
	Mean weekday MVPA (SD)		
	Mean weekday light (SD)		
	Mean weekday sedentary (SD)		
	Mean weekend MVPA (SD)		
	Mean weekend light (SD)		
	Mean weekend sedentary (SD)		
	Meets PA guidelines[2]: N (%)		
Parent gender			
Male		N=	N=
	Mean weekday MVPA (SD)		
	…		
Female		N=	N=
	…		
Household education			
	Up to A level or equivalent	N=	N=
	…		
	Degree equivalent or higher	N=	N=
	…		

[1] Comparator group of the same schools

[2] CMO recommends at least 30 minutes MVPA per day for adults, averaged across the week

Table 3: Comparison of children's physical activity between T0, T1 and T2 aggregated over [N] schools

		Pre-COVID Comparator group[1] (Mar 2017 - May 2018)	Wave 1[1] (May 2021-Dec 2021)	Wave 2 (Jan 2022 –[DATE])
All children		N=	N=	N=
	Mean weekday MVPA (SD)			
	Mean weekday light (SD)			
	Mean weekday sedentary (SD)			
	Mean weekend MVPA (SD)			
	Mean weekend light (SD)			
	Mean weekend sedentary (SD)			
	Meets PA guidelines[2]: N (%)			
Child Gender				
Boys		N=	N=	N=
	Mean weekday MVPA (SD)			
	…			
Girls		N=	N=	N=
	…			
Household education				
Up to A level or equivalent		N=	N=	N=
	…			
Degree equivalent or higher		N=	N=	N=
	…			

[1] Comparator group of the same schools

[2] CMO recommends at least 60 minutes MVPA per day for children, averaged across the week

Table 4: Comparison of parent's physical activity between T0, T1 and T2 aggregated over [N] schools

		Pre-COVID Comparator group[1] (Mar 2017 - May 2018)	Wave 1[1] (May 2021-Dec 2021)	Wave 2 (Jan 2022 –[DATE])
All parents		N=	N=	N=
	Mean weekday MVPA (SD)			
	Mean weekday light (SD)			
	Mean weekday sedentary (SD)			
	Mean weekend MVPA (SD)			
	Mean weekend light (SD)			
	Mean weekend sedentary (SD)			
	Meets PA guidelines[2]: N (%)			
Parent gender				
Male		N=	N=	N=
	Mean weekday MVPA (SD)			
	…			
Female		N=	N=	N=
	…			
Household education				
Up to A level or equivalent		N=	N=	N=
	…			
Degree equivalent or higher		N=	N=	N=
	…			

[1] Comparator group of the same schools

[2] CMO recommends at least 60 minutes MVPA per day for adults, averaged across the week

Figure 1: Flow Diagram

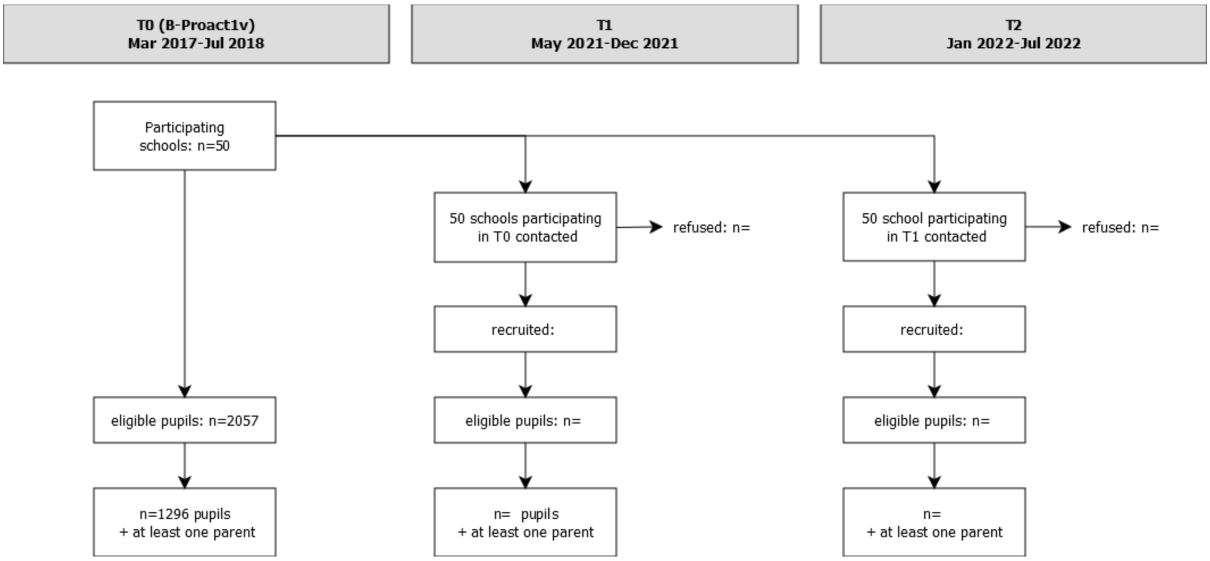

Table 5: Comparison of key pupil demographics between Waves

	T0: Pre-COVID	T1	T2
	N=	N=	N=
Child age: mean (SD)			
Child gender			
% male			
% female			
Parent age			
Parent gender			
% male			
% female			
Parent ethnicity			
% White			
% Black/African/Caribbean/ Black British			
% Asian/Asian British			
% Mixed			
% Other			
Household education			
Up to GCSE/equiv %			
A level/equiv %			
Degree/HND/ equiv %			
Higher degree %			
IMD score: mean (SD)			

Table 6: Comparison of school-level characteristics between participating schools in T0, T1 and T2

	T0: Pre-COVID	T1	T2
	N=	N=	N=
School characteristics			
School size: median (IQR)			
Average school IMD: mean (SD)			
…			
Aggregated Pupil characteristics			
Average child age: mean (SD)			
Average % girls: mean (SD)			
Average % white ethnicity: mean (SD)			
Average % up to A level household education: mean (SD)			
Average pupil IMD: mean (SD)			

Table 7: Summary of missing data

	T0: Pre-COVID	T1	T2
	N (%) missing	N (%) missing	N (%) missing
Child			
Age			
Gender			
Weekday accelerometer data			
Weekend accelerometer data			
Parent			
Age			
Gender			
Ethnicity			
Household education			
Weekday accelerometer data			
Weekend accelerometer data			
Household			
IMD score: mean (SD)			

Table 8: Estimates of primary and secondary outcomes for T0 and T1

	Pre-COVID		Wave 1	
	Mean/%	95% CI	Mean/%	95% CI
All children	N=		N=	
Mean weekday MVPA				
Mean weekday light				
Mean weekday sedentary				
Mean weekend MVPA				
Mean weekend light				
Mean weekend sedentary				
% Meets PA guidelines[2]:				
Child Gender				
Boys	N=			N=
Mean weekday MVPA				
…				
Girls	N=			N=
…				
Household education				
Up to A level or equivalent	N=			N=
…				
Degree equivalent or higher	N=			N=
…				

Table 9: Comparison of child weekday MVPA between T0 and T1

	Difference between pre-COVID and T1			
	N	Estimate	95% CI	p-value
Model 1: unadjusted				[1]
Model 2				[1]
Model 3				[1]
Interaction Effects[2]				
Gender				[3]
Household education				[3]

All models are adjusted for accelerometer wear time, hours of daylight and seasonality.
Model 2 also controls for gender and household education, and Model 3 additionally includes COVID-restrictions and staff/pupil absences.

[1] p-value for a test for a difference between T0 and T1
[2] Interaction effects reported for model adjusted for gender or household education.
[3] p-value for a test for an interaction effect

Table 10: Estimates of child weekday MVPA between T0, T1 and T2

	Model 1: unadjusted (N=)	Model 2 (N=)	Model 3 (N=)
Difference between pre-COVID and T1: mean (95% CI)			
Difference between pre-COVID and T2: mean (95% CI)			
Confounders			
Child gender: female	-		
… [other confounders, including interaction effects if appropriate, plus any COVID restriction variables]	-		
Model fit			
BIC[1]			

All models are adjusted for accelerometer wear time, hours of daylight and seasonality.
Model 2 also controls for gender and household education, including interaction effects [if these were found to be statistically significant in Analyses 4.3.1 & 4.4.1], and Model 3 additionally includes COVID- restrictions and staff/pupil absences.
[1] model fit: lower values indicate better fit

Table 11: Estimates of child weekday MVPA with continuous change over time

	Model 1: unadjusted (N=)	Model 2 (N=)	Model 3 (N=)
Change over time compared to baseline			
September 2021: mean (95% CI)			
December 2021: mean (95% CI)			
… [time points chosen to describe spline]			
Confounders			
Child gender: female			
… [other confounders, including interaction effects if appropriate, plus any COVID restriction variables]			
Model fit			
estimated degrees of freedom for spline			
p-value[2] for spline			
BIC[1]			

All models are adjusted for accelerometer wear time, hours of daylight and seasonality.
Model 2 also controls for gender and household education, including interaction effects [if these were found to be statistically significant in Analyses 4.3.1 & 4.4.1], and Model 3 additionally includes COVID- restrictions and staff/pupil absences.
[1] model fit: lower values indicate better fit
[2] estimated degrees of freedom indicates 'wiggliness' of spline, with a value of 1 indicating linear; p-value for a test for a difference from a linear change over time

Figure 2: Smooth estimate of change in MVPA over time

[illustrative graph based on dummy data]

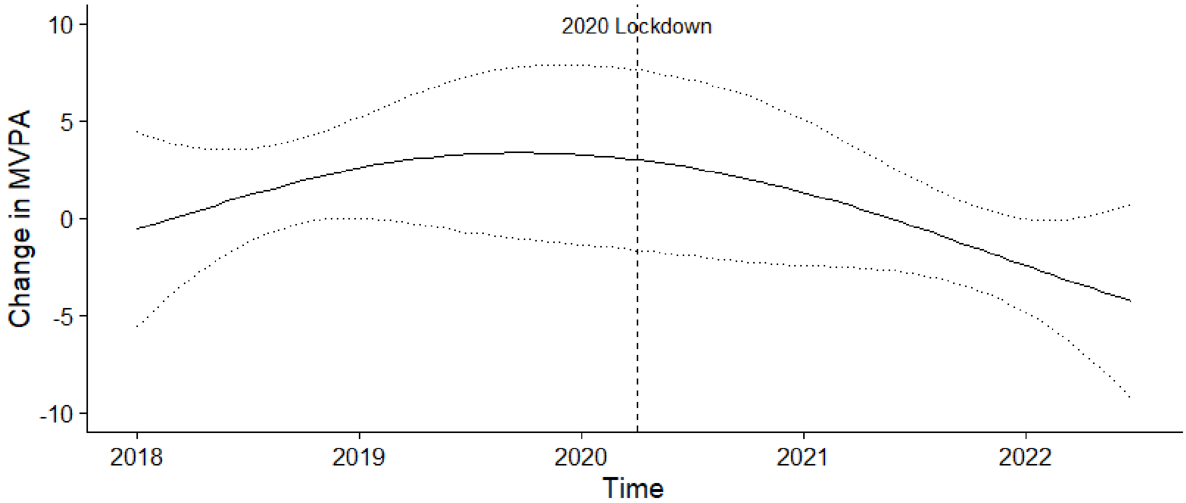

Table 12: Comparison of potential mediators between T0 and T2

		Difference between pre-COVID and T2		Model Fit (BIC)
	N	Estimate	95% CI	
Individual mediators				
Physical activity enjoyment				
Mode of travel to school				
…				
Parental mediators				
Parent physical activity				
Parent PA motivation				
…				
School mediators				
Curriculum physical activity				
Play provision				
…				

All models are adjusted for accelerometer wear time, hours of daylight and seasonality, individual confounders gender, household education and child zBMI, and school confounders size and deprivation.

Table 13: Association of MVPA and potential mediators

		Difference in weekday MVPA between pre-COVID and T2		Model Fit (BIC)
	N	Estimate	95% CI	
No mediators				
Individual mediators				
Physical activity enjoyment				
Mode of travel to school				
…				
Parental mediators				
Parent physical activity				
Parent PA motivation				
…				
School mediators				
Curriculum physical activity				
Play provision				
…				
ALL MEDIATORS				

All models are adjusted for accelerometer wear time, hours of daylight and seasonality, individual confounders gender, household education and child zBMI, and school confounders size and deprivation.

6. REFERENCES

1. 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.
2. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci. 2008;26(14):1557-65.
3. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Medicine and science in sports and exercise. 2008;40(1):181-8.
4. Arnold BF, Hogan DR, Colford JM, Hubbard AE. Simulation methods to estimate design power: an overview for applied research. BMC Med Res Methodol. 2011;11:94.
5. 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.
6. Markland D, Tobin V. A Modification to the Behavioural Regulation in Exercise Questionnaire to Include an Assessment of Amotivation. Journal of Sport and Exercise Psychology. 2004;26(19):1-196.
7. Cole T, Freeman J, Preece M. Body mass index reference curves for the UK, 1990. Arch Dis Child. 1995;73(1):25 - 9.
8. UK Chief Medical Officers. UK Chief Medical Officers' Physical Activity Guidelines. London, UK: Department of Health and Social Care; 2019.
9. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. Time series regression studies in environmental epidemiology. Int J Epidemiol. 2013;42(4):1187-95.
10. Quartagno M, Grund S, Carpenter J. jomo: A Flexible Package for Two-level Joint Modelling Multiple Imputation. The R Journal. 2019;11(2):205-28.
11. Wood S. Generalized additive models: an introduction with R: CRC Press; 2006.
12. Zhang Z, Zyphur MJ, Preacher KJ. Testing Multilevel Mediation Using Hierarchical Linear Models: Problems and Solutions. Organizational Research Methods. 2009;12(4):695-719.