Physical Activity School-Specific PORTfolio intervention - PASSPORT

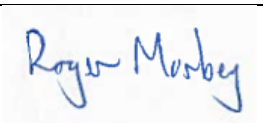
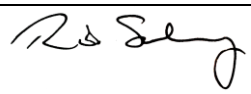
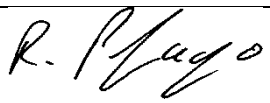
Statistical Analysis Plan

Version *1.1 (02/10/2026)*

Based on Protocol UKRI REF EP/X023508/1 *(24/10/2025)*

Version	Date	Author	Edits
1	22/05/26	Roger Morbey	Initial SAP, approved by reviewers listed below
1.1	02/10/26	Ruth Salway	Sections 3.5.7 and 4.8 added in response to changes to PE & Sports Premium for 2026/27

The following people have reviewed the Statistical Analysis Plan and are in agreement with the contents

	NAME	JOB TITLE	SIGNATURE	DATE
Author	Roger Morbey	Senior Research Associate		24/04/2026
Senior Statistician	Ruth Salway	Lecturer in Quantitative Public Health Improvement		28/04/2026
Study Chief Investigator	Russ Jago	Professor of Physical Activity & Public Health		28/04/2026
Statistical reviewer	Frank de Vocht	Professor in Epidemiology and Public Health		29/04/2026
Independent Statistical reviewer	Patricia Albers	Senior Research Associate		22/05/2026

Effective Date:	

Contents

List of abbreviations 3
1 Introduction 4
1.1 Purpose 4
1.2 Background 4
1.3 Objectives 5
1.4 Quantitative research questions 5
2 Synopsis of Study Design and procedures 6
2.1 Synopsis and rationale 6
2.2 Study plan 6
2.3 Description of intervention 8
2.4 Sample size justification 8
2.5 Randomisation procedures 10
2.6 Data collection 11
3 General Analysis considerations 13
3.1 Analysis population 13
3.2 Data cleaning and missing data 13
3.3 Outcome measures 13
3.4 Derived variables 15
3.5 Bias and confounding 16
3.6 Statistical software 18
4. Analysis 18
4.1 Descriptive analysis 18
4.2 Impact on primary outcome: weekday MVPA 18
4.3 Impact on other accelerometer outcomes 19
4.4 Impact on other secondary outcomes 19
4.5 Least active pupils 19
4.6 Intervention impact change over time 19
4.7 Mediation by active travel, clubs and motivation 20

4.8 Other analyses 20
5. Sensitivity analyses 21
6. Indicative Tables (subject to change) 21
6.1 Baseline data 22
6.2 Numbers analysed 22
6.3 Outcomes and model estimates 23
7. References 26

List of abbreviations

Acronym	Details
MVPA	*Moderate-vigorous physical activity*
SWCRT	*Stepped wedge cluster randomised trial*
ICC	*Intra-Cluster Correlation Coefficient*
SEN	Special Educational Needs
EAL	English as additional language
FSM	Free school meals
PP	Pupil premium (includes pupils eligible for free school meals)

1 Introduction

1.1 Purpose

This document outlines the statistical methods to be implemented during the analyses of quantitative data collected within the scope of the Physical Activity School-Specific PORTfolio (PASSPORT) project . This Statistical Analysis Plan (SAP) was prepared in accordance with the Protocol version *UKRI REF EP/X023508/1*, dated *24/10/2025*, trial registration DOI, https://www.isrctn.com/ISRCTN14226285.

The purpose of this Statistical Analysis Plan (SAP) is to provide a framework in which answers to the protocol objectives may be achieved in a statistically rigorous fashion, aimed at eliminating bias or analytical deficiencies. Specifically, this plan has the following purposes:

- To explain in detail how the data will be handled and analysed, adhering to commonly accepted standards and practices of statistical analysis.
- To prospectively outline the specific types of analyses and presentations of data that will form the basis for conclusions.

1.2 Background

Physical activity is associated with improved physical and mental health across the life course. Physical activity tracks from childhood into adulthood with more active children likely to engage in both a higher amount and wider range of physical activities in adulthood [1, 2]. The UK Chief Medical Officer's advice is that children and young children should engage in moderate or vigorous physical activity (MVPA) for an average of at least 60 minutes per day across the week [3]. There is a need to put in place programs that focus on increasing MVPA and building children's confidence and interest in physical activity. The end of primary school years (9-11 years of age) is a critical age for establishing these behaviours and is the focus of this project.

The PASSPORT project is a school-based intervention for primary schools to increase children's physical activity, which considers the school context. The intervention will consist of a core package plus six pathways. Primary schools will be provided with a tailored report, based on their context, recommending intervention components organised within specific pathways, from which they will choose those they consider most effective for their situation. Schools will be given a total budget of £7000 to implement their chosen interventions. The study will evaluate the impact of the PASSPORT intervention as a whole across a school year (September to July) (context assessment, recommendations, choices and implementation) on the accelerometer-measured physical activity of pupils in Year 5 (aged 9-10).

1.3 Objectives

The PASSPORT protocol describes inter-linked objectives to evaluate the impact of the PASSPORT intervention on children's physical activity and related measures. This SAP addresses the quantitative objectives.

1.4 Quantitative research questions

The primary outcome will be average weekday MVPA. The primary research questions are:

1. What is the overall impact of the PASSPORT intervention on Year 5 pupil's average weekday MVPA?
    a. Does this differ by pupil gender or by pupil premium (a measure of deprivation [4]?
    b. What is the impact of the PASSPORT intervention on the physical activity of the children who were least active at baseline?
    c. How does the impact of the PASSPORT intervention change over time?
    d. Is Y5 pupil MVPA mediated by pupil active travel, attendance at after school clubs or pupil motivation for physical activity?
2. What is the overall impact of the PASSPORT intervention on Year 5 pupil's total volume of physical activity, average weekday sedentary time, average weekday light physical activity and average weekend MVPA?
    a. Does this differ by pupil gender or by pupil premium?
    b. What is the impact of the PASSPORT intervention on the physical activity of the children who were least active at baseline?
    c. How does the impact of the PASSPORT intervention change over time?
3. What is the impact of the PASSPORT intervention on other outcomes:
    a. Pupil wellbeing?
        i. Does this differ by pupil gender or by pupil premium?
    b. Aggregated pupil attendance by school for the cohort year group?

2 Synopsis of Study Design and procedures

2.1 Synopsis and rationale

We use an open-cohort incomplete stepped wedge randomised control (SWRCT) design [5] as a pragmatic alternative to the traditional cluster randomised controlled trial . In this design, all schools begin in the control arm and transition in random order from control to intervention until all schools have received the intervention. This is an open cohort with mainly repeated measures from pupils, but allowing for the possibility that pupils may drop out of the study and new pupils may be recruited if necessary.

There are several advantages to a SWRCT in this setting [6]. Firstly, the increased statistical power means fewer schools are needed, making it feasible to recruit the necessary numbers. Secondly, the design's phased implementation of interventions makes intervention support and data collection logistically easier. Thirdly, as all schools receive the intervention and associated funds, they have more incentive not to drop-out of the study. Finally, the design enables analysis of secondary questions such as how the impact of the intervention varies over time. The SWRCT also allows us to account for the seasonal nature of physical activity.

2.2 Study plan

Eighteen schools will be recruited and randomly allocated to one of five sequences, two to the first sequence and four schools each to the other four sequences. Baseline recruitment and measurement will take place between 24th February 2026 and 22nd May 2026. Main trial measurements will take place between 3rd September 2026 and 23rd July 2027.

The following figure shows the planned sequence for allocation of interventions to schools.

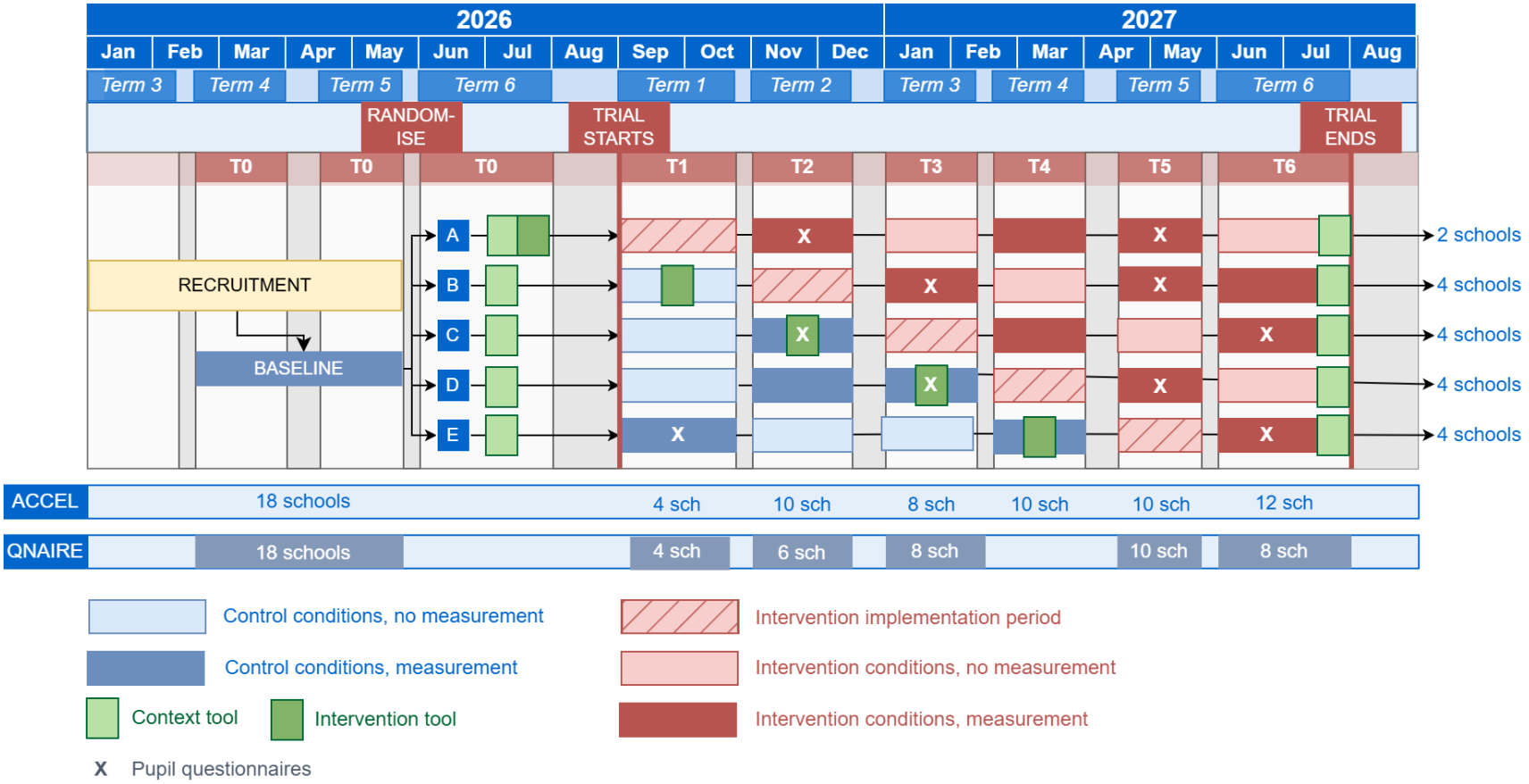


Figure 1 Sequence plan: 5 sequences are labelled A to E. Terms are labelled T1 to T6 within academic year. ACCEL – accelerometer measurements, QNAIRE – pupil questionnaire measurements.

2.3 Description of intervention

PASSPORT is a school-based intervention designed to increase children's physical activity.

The intervention will consist of a context assessment followed by a report with tailored recommendations, from which a school chooses components to build a bespoke programme of interventions that are specific to their context.

The context assessment is completed by schools at baseline to provide an understanding of the school's physical activity context. It includes questions on school environment, facilities, and resources for physical activity, current physical activity opportunities and pupil engagement, school policies and organisation, and school culture.

The context assessment is used to provide a tailored set of recommendations, provided by the research team, which will identify areas to target that would fit the context. This consists of a core package that is offered to all schools to build a school culture of physical activity, and a set of recommended physical activity intervention components based on the context assessment that are organised into six potential pathways that cover different physical activity opportunities:

1. Physically active and outdoor learning
2. Active breaks
3. Breaktime and lunchtime
4. Active travel
5. Before and after school activities
6. School and wider community

Schools choose from the recommendations, with support from the research team, to create their own bespoke action plan and milestone tracker to be delivered over the school year. Schools will be provided with resources and intervention support for their selected intervention components.

2.4 Sample size justification

2.4.1 Primary outcome: average weekday MVPA

We used a simulation-based approach to estimate power [6] for a two-sided test at the 5% significance level for the proposed incomplete stepped wedge design with five sequences (two in the first sequence and four in the remaining four sequences), shown in Figure 1, under a range of different scenarios for the number of participating pupils and schools. We estimated the power to detect a 5 min and 6 min difference in weekday MVPA, which represent meaningful effect sizes for an impact on health [7] and are achievable in a multi-component intervention of this type, with a variance of 540

(standard deviation of 23.2 min [8]). We allowed for clustering of pupils within schools (intra cluster correlation coefficient ICC of 0.10 based on the upper 95% confidence interval bound), repeated pupil measurements (individual autocorrelation IAC = 0.45) and correlation between schools over time (cluster autocorrelation CAC = 0.75 between measurements to give a conservative estimate of 0.4 for the overall CAC over a full year), with all correlation estimates based on estimates derived for accelerometer-based weekday MVPA in UK primary schools [8]. Simulations allowed for an open cohort of pupils based on a total eligible closed pupil population of 60 per school (equivalent to two classes of 30 pupils in a year group, which is a typical size of school in the trial area) with differing assumptions about the overall participation rates and proportion of repeat participants, which correspond to an average valid sample of 20, 30 and 40 pupils at each measurement period. We assumed a trial target of 18 schools, alongside power estimates for 16 schools, allowing two to drop out (estimates shown are for drop-outs from sequences B and D, which had lowest power). As MVPA is seasonal, we included a seasonal trend in MVPA based on a previous study [9] which varied by approximately 4 minutes between winter and summer. Power was estimated based on 50,000 simulations, using a generalised additive mixed effects model including clustering, repeated measures, school autocorrelation and a continuous thin-plate spline to capture seasonal and secular time trends. This reflects the primary analysis model described in the section on Impact on weekday MVPA, section 4.2. Further details of the simulation approach are given in Salway et al. 2025 [6].

Table 1 shows the estimated power for each combination of school and pupil sample sizes, for effect sizes of 5 and 6 min difference in MVPA.

Table 1: Power for 5 and 6 min differences in weekday MVPA by pupil sample sizes per school

	18 schools		**16 schools**	
pupils per school at each measurement	**5 min**	**6 min**	**5 min**	**6 min**
20	70%	82%	67%	79%
30	81%	91%	78%	88%
40	85%	93%	83%	92%

The simulations suggests that with the full target of 18 schools, we would be able to detect a minimum difference of 5 mins of weekday MVPA, if we are able to achieve valid accelerometer data in at least half the eligible pupils in each school at each measurement period. For lower pupil recruitment/retention rates, or if up to 2 schools drop out, we will still be able to detect a difference of 6 mins. Our pilot study achieved a 50-70% rate of valid data per measurement point, which suggests that these sample

sizes are achievable. Note that given the complexity of the simulations involved, all parameter estimates are on the conservative side.

2.4.2 Secondary outcomes: pupil questionnaire

Secondary outcomes are measured in the pupil questionnaire administered at three time points and so form a two-step complete stepped wedge design. A sample of 16-18 schools with 30 pupils at each measurement would therefore provide over 80% power to detect a minimum effect size of 0.25 for a 5% two-sided significance level (assuming a lower ICC of 0.02 typical of behavioural risk factors in schools [11], IAC=0.7 and a conservative choice of CAC=1). This corresponds to a detectable difference of ~1.5-2 in the Stirling well-being scale, with a difference of ~1 in the positive emotional state and positive outcome subscales.

2.4.3 School-level outcomes: attendance

School attendance will be reported at school-level with 9 measurements (from term 4 in 2025/26 to term 6 in 2026/27). These form a complete five-step stepped wedge design, and 16-18 schools would have approximately 80% power to detect a change in attendance rate of at least 0.6 percentage points (based on mean attendance rates of 92-95%) at a 5% two-sided significance level.

2.5 Randomisation procedures

To minimise bias, schools will be randomly allocated to clusters after baseline measurements. Furthermore, each measurement stage will be blinded for the analysis team (blinding for the schools and fieldwork team will not be possible due to the nature of the intervention).

To ensure that random allocation to the five sequences results in a balanced study, where schools that receive the intervention early are comparable to those that receive it late, we will use a covariate-constrained method [12]. We will check for balance across the following school covariates: number of pupils in year 4, % with pupil premium, % of female pupils, % with English as an additional language, % non-white ethnicity, and % with special educational needs or disabilities.

2.5.1 Eligibility criteria

An allocation would be considered sufficiently balanced if each covariate is within the chosen covariate tolerance when comparing measurements taken before and after intervention. Covariate tolerances will be chosen after recruitment of schools based on the variation in the covariates, to ensure that there are sufficient allocations to choose from.

2.5.2 Sensitivity analysis

As a precaution the set of eligible allocations will be examined to determine if the criteria have been over-restrictive. Attention will be given to ensure no two schools are always allocated to the same sequence or never allocated to the same sequence. If this occurs then the allocation can be re-run with relaxed criteria, i.e. higher tolerances for one or more covariates.

2.5.3 Choosing an allocation

Allocations will be randomly generated and checked against the eligibility criteria for each covariate. A single allocation will then be selected at random from a set of 100,000 eligible allocations using random number generation in the statistical programming software R. To ensure objectivity, the process for selection of an allocation will be conducted by an independent statistician. The R script and random number seed will be published on the Open Science Framework for transparency.

2.6 Data collection

2.6.1 Accelerometers

Physical activity will be measured at baseline and on three subsequent points depending on the school's sequence allocation (see Study plan), section 2.2.

Physical activity will be collected using the wrist-worn AX3 (or AX6) tri-axial accelerometer, manufactured by Axivity Ltd (UK). The accelerometer is set for recording at the default values of 100Hz sampling frequency with ±8g sampling range. The accelerometer device is worn on the wrist of the participant's non-dominant hand and can be worn 24h a day including for sleeping, bathing or swimming. Participants are encouraged not to remove the accelerometer, but if removed it should be replaced on the same wrist and in the same orientation as demonstrated by the fieldworker and described in the instructions.

The accelerometer is worn from Day 1, the day of the fieldworker's visit. However, recording does not begin until 05:00 the following morning (Day 2). The device is worn for 7 days, for a total of 6 days of recorded accelerometer data. Data is extracted from the device using the Open Movement GUI software (https://openmovement.dev) and the raw data exported in cwa format. A valid day is defined as having at least 10 hours of data; for inclusion a pupil must have at least two valid weekday readings.

2.6.2 School questionnaires

School questionnaire data will be collected via the school study agreement during recruitment.

Data collected will include the following demographics data:

- Number of primary school aged pupils (Reception – Year 6)
- Number of pupils in Year 4 (stratified by Girls / Boys /Other)
- Number of classes in Year 4
- % primary school aged pupils with English as an additional language
- % primary school aged pupils with non-white ethnicity
- % primary school aged pupils with special educational needs or disabilities
- % primary school aged pupils eligible for individual free school meals
- % primary school aged pupils eligible for Pupil Premium

2.6.3 Pupil questionnaires

Pupil questionnaire data will be collected at baseline, mid-point and end of study as described in the Study plan, section 2.2. Data will be collected on pupil's wellbeing and physical activity motivation. The Stirling Children's Wellbeing Score [13] will be used to measure wellbeing. Physical activity motivation will be measured using the Relative Autonomy Index (RAI) based on the Behavioural Regulations in Exercise Questionnaire (BREQ-2) [14] adapted for children. The questionnaires also include information on whether the pupils attend extra-curricular clubs involving physical activity at school and in the community, and whether they typically use active travel to get to school.

2.6.4 Pupil consent forms

On recruitment, parents will be asked to provide consent for pupils to participate in the study. Information collected will include:

- Pupil gender
- Pupil ethnicity
- Special educational needs or disabilities
- Receipt of pupil premium
- English as an additional language

2.6.5 School attendance

Schools will be asked to provide information on the pupil attendance rates for the trial cohort year group (year 4 at baseline, and year 5 for the main trial) for each term of the trial from term 4 in 2025/26 to term 6 in 2026/27. Schools will be asked to provide this data retrospectively at a mid-point in the trial and at the end of the trial. In addition to the main outcome of percentage absence, data will also be collected on percentage of:

- Persistent absences (% pupils who are absent >10% of school days)
- Severe absences (% pupils who are absent >50% of school days)
- Authorised absences (% of sessions)
- Unauthorised absences (% of sessions)

3 General Analysis considerations

3.1 Analysis population

The analysis population includes all pupils recruited prior to May 2026 from the participating schools. Pupils for schools will not be continually recruited each term but with an additional follow-up wave of recruitment prior to the first wave of data collection in year 5 if necessary. Whilst the intention is for a closed cohort design with the same pupils in the same schools returning measurements each time, it is acknowledged that a few pupils may drop out of the study part way through. If during the study period a pupil transfers from one school to another within the study, only their measurements at the first school will be collected and analysed. Analysis will be on an intention-to-treat (ITT) basis, defined as analysing schools to the sequence as randomised.

3.2 Data cleaning and missing data

We will perform internal consistency checks and plausibility checks to flag any inconsistencies, implausible or extreme values.

We will report the extent of missing data in accelerometer data and other key variables. Pupil questionnaires are paper-based and completed and checked by fieldworkers in school when accelerometers are handed out, and so we expect very low levels of missing data. 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.

3.3 Outcome measures

3.3.1 Primary and secondary accelerometer measures

The primary measure of pupil activity is the pupil's mean weekday number of minutes spent in moderate or vigorous physical activity (MVPA).

Secondary outcomes measures will include total weekday volume of activity, mean weekday sedentary time, mean weekday time spent in light physical activity and mean weekend MVPA.

Accelerometer data processing

Raw accelerometer data will be processed using the open-source R package GGIR [15] version 3.3-0. The parameters used in GGIR are saved as a configuration file for future reference.

MVPA includes activity above the Hildebrand child threshold of 191.6. Hildebrand et al defined MVPA thresholds for children wearing wrist monitors [16, 17], using GENEActiv monitors which are comparable to the AX3/AX6 monitors [17].

Table 2: Key non-default parameter settings for GGIR package.

Parameter	Setting	Description
Do.enmo	True	uses Euclidean Norm minus one as acceleration metric (for compatibility with Hildebrand cut points)
Acc.metric	"ENMO"	
Mvpaththreshold	191.6	Hildebrand threshold for moderate or vigorous activity in children
Threshold.lig	56.3	Threshold for light activity
Threshold.mod	191.6	Threshold for moderate activity
Threshold.vig	695.8	Threshold for vigorous activity
Includedaycrit	c(10, 10)	Minimum number of hours of day needed to assess sleep period and provide summary reports.
Includedaycrit.part5	10	
Minimum_MM_length.part5	10	Minimum length of day (midnight to midnight) for summary reports.

The GGIR package outputs summary data which includes the amount of time each day a pupil spent in moderate or vigorous exercise in minutes. Other information includes total weekday volume, measured as average acceleration, accelerometer wear time and time spent in sedentary or in light activity. Further documentation on GGIR can be found at: https://wadpac.github.io/GGIR/

Post GGIR processing, data will be extracted from the "part 5-day summaries" csv files. Exclusion criteria will ensure we are using only data where at least 10 hours during the day is available and at least two weekday measurements are required.

3.3.2 Pupil wellbeing

Pupil wellbeing will be measured using data collected from the pupil questionnaires. The Stirling Children's Wellbeing Score has been validated for this age group and consists of 15 questions. Each question is scored on a Likert scale from 1 to 5.

Three of the wellbeing questions are included to determine if respondents are responding based on the socially desirability of their answers, forming a social desirability sub-scale. The Stirling Scale scores should be interpreted with caution when scores on the social desirability sub-scale total 3 or 14-15 as this indicates that the respondent may be selecting answers based on social desirability or undesirability [18]. Therefore, pupil questionnaires with a social desirability score of 3, 14 or 15 are excluded from further analysis of wellbeing or motivation.

The other 12 wellbeing questions split into positive emotional state and positive outlook sub-components. The Stirling score is the total of these 12 questions with an overall scale of 12-60. The two sub-components of the Stirling score will also be presented, to summarise pupil's positive emotional state and positive outlook.

3.4 Derived variables

The methods for calculation of derived variables are summarised in the following table:

Table 4 Derived variables

Type	Variable	Method
Physical activity	Weekday MVPA	MVPA includes activity above the Hildebrand child threshold of 191.6. Total minutes spent in MVPA during waking periods on weekdays are calculated for days with at least 10 hours of data. , the mean weekday daily minutes will be calculated for each pupil across the week, for those providing at least 2 valid weekdays .
	Total physical activity	Mean acceleration during weekdays.
	Weekday sedentary time	Mean daily time spent awake on weekdays but with physical activity below the threshold of 56.3.
	Weekday light physical activity	Mean daily time spent awake on weekday with activity levels between 56.3 and 191.6. Excluding days with less than 10 hours and weeks where less than two valid weekdays are available.
	Average weekend MVPA	As weekday MVPA but for weekends. Valid days must have 10 hours of recordings, and at least two weekdays within the week.
	Least active pupils	The 25% of pupils with the lowest MVPA recorded at baseline, stratified by gender.
Pupil characteristics and questionnaire	Gender	Self-reported on parental consent form.
	Pupil premium	Whether child receives free school meals or pupil premium, as reported by parental consent form.
	Pupil wellbeing	Stirling wellbeing score: Sum of 12 wellbeing questions in pupil questionnaire, each scored out on Likert scale of 1-5. Excluding any results where the sum of the 3 social desirability questions is 3, 14 or 15. The two subscales for emotional and outlook wellbeing will also be recorded.
	Pupil motivation	Relative autonomy index: Sum of 12 motivation questions. The Intrinsic and Identified scores

		are given positive weights of 2 and 1, whilst the introjected and external scores are given negative weights of -1 and -2 respectively. Excluding any results where sum of 3 social desirability questions is 3, 14 or 15.
	Active travel	Whether pupil uses active travel as the typical mode of transport to and from school, from pupil questionnaire.
	After school club	Whether pupil attends at least 1 after school club per week which involved physical activity, from pupil questionnaire.
	Outside school club	Whether pupil attends at least 1 outside school club per week which involved physical activity, from pupil questionnaire.
School-level data	Attendance	% attendance of pupils in school year group during term.

3.4.1 Sub-group analysis

A sub-group analysis for the least active pupils will be defined as those in the bottom 25% of pupils by MVPA at baseline, stratified by gender.

3.5 Bias and confounding

3.5.1 Alterations to schedule

If any changes to the scheduled data collection occur, then time will be recorded as when collection occurred. Time is treated as a continuous variable so this will not bias results.

Analysis will be intention-to-treat. If there are delays to intervention implementation, then a per protocol sensitivity analysis (see Sensitivity analyses, section 5) will be conducted if either five or more schools in total, or if more than 50% of schools in a sequence, do not implement as scheduled.

3.5.2 Retention rates

Potential bias could occur if too many pupils drop out during the study and only provide measurements for the early stages in the study. Particularly, if they only provide measurements prior to the intervention and are structurally different, e.g. have lower MVPA on average, than other pupils. Therefore, we will measure the number of pupils who do not complete the study, including whether they provided both pre-intervention and intervention measurements.

3.5.3 Unbalanced random allocation

Bias could occur if, for example, the random allocation puts all the schools with high special needs in the same cluster. Bias due to unbalanced allocations will be reduced due to the covariate constrained randomisation process (see section 2.5).

3.5.4 Secular and seasonal trends

SWRCTs are partially confounded by time, and so it is important to model time in the analysis. Additionally, physical activity is known to exhibit seasonality [21], and so we will model time continuously as a nonlinear smooth function via a thin plate spline [22], which will capture both seasonal and secular trends. Note that the specific incomplete design used in this trial has been chosen to minimise any bias arising due to seasonality [6].

3.5.5 Within cluster contamination

To reduce the risk of within cluster contamination, new pupils joining a school or class during the study period will not be recruited into the study if they were already part of the study in another school, only their measurements in the original school will be retained.

3.5.6 Measurement error

Measurement error could occur due to monitor bands being worn in the incorrect orientation. During the pilot study approximately half of the pupils removed the bands at some point and then put them back on. Assuming they are just as likely to put them back on in the correct orientation as upside down, a worst-case scenario would be 25% of measurements coming from incorrectly orientated bands. An in-house test comparing correctly and incorrectly oriented bands, found only small differences. Thus, it is estimated that measurement error due to incorrectly oriented bands is likely to be less than 1 minute for the primary outcome of mean weekday MVPA.

3.5.7 Changes to PE and Sports Premium

On 21 May 2026 the Department for Education announced changes to the PE and Sports Premium funding provided to primary schools to support PE, sport and physical activity in school. These changes are expected to be operational from Spring 2027, with transition arrangements from September 2026, and will affect the amount received, the funding model and how it can be spent, for the 2026/27 academic year (and beyond). These changes will occur after baseline data collection of the PASSPORT study and before the interventions begin and may affect physical activity provision within schools, and hence the primary MVPA outcome.

We have explored the potential impact of the changes on the statistical power of the study and potential bias in the primary effect parameter of interest via simulation. The

staggered intervention implementation of the stepped wedge design provides both control and intervention measurements throughout 2026/27 and the analysis includes estimation of a secular trend in MVPA, and so the planned design and analysis are robust to these policy changes. We estimate a small drop in power of 0-2 percentage points and negligible bias in the effect parameter of up to 1-2%, depending on the magnitude of the policy impact on MVPA (2-5 mins). As a result, there are no changes needed to the main analysis.

3.6 Statistical software

Statistical analysis will be undertaken in R [23] using the package GGIR [15] for processing accelerometer data and mgcv for modelling generalised additive mixed models [24].

4. Analysis

4.1 Descriptive analysis

Data collected from accelerometers and pupil and school questionnaires will be summarised, recording the number of missing observations, and for numeric variables, minimum, maximum, mean, median, interquartile range and standard deviations (see indicative tables, section 6). For categorical variables the number and percentage in each category will be included in the summary. Pupil and school characteristics will be summarised for each sequence at baseline.

4.2 Impact on primary outcome: weekday MVPA

Average weekday MVPA will be modelled using a generalised additive mixed effects model, (GAMM) with a linear link function, with a binary variable indicating whether schools were in the intervention period. School, pupil and school-time random effects will be included in a mixed effects model, following Hooper et al [25]. Trends will be modelled using a thin-plate spline to capture a combined secular trend and seasonality [5, 26]. Wear time will be included as a potential confounding variable.

The Stepped-wedge design provides high power for treatment heterogeneity with respect to individual characteristics. Analyses of primary and secondary physical activity outcomes and pupil wellbeing will be repeated using covariate and interaction terms added to the basic model to enable comparison of groups.

Model assumptions will be checked via visual inspection of residuals from the fixed part of the multi-level model and the random effects. If residual plots indicate the presence of outliers, we will investigate the leverage of these points and if necessary, undertake a sensitivity analysis with the points omitted.

4.3 Impact on other accelerometer outcomes

We will repeat the GAMM model with the secondary accelerometer outcome variables:

- Mean weekday daily volume (weighted average of mean acceleration during waking period.)
- Mean weekday time spent sedentary
- Mean weekday time spent in light physical activity
- Mean weekend time spent in MVPA

4.4 Impact on other secondary outcomes

4.4.1 Pupil wellbeing

Pupil wellbeing will be measured at three points using the pupil questionnaires: at baseline, midpoint, and at the end of the study. This forms a two-step SWRCT, and so we will use a linear mixed effects model for Stirling Children's Wellbeing Score with fixed effects for intervention and step (baseline, mid, end), and school, pupil and school-time random effects as above. As time trends are expected to be much weaker, and the number of steps is smaller, we replace the spline for time with fixed discrete time effects for each step.

4.4.2 School attendance

Analysis of school attendance will be at school level, with measurements available at each term. This forms a stepped wedge design analysis with five steps and multiple pre-intervention measurements. As percentage absence is small (<10%), analysis will use a Poisson mixed effects model for absence counts, with an offset for total number of sessions, fixed effects for intervention and discrete time, and a random effect for repeated school measurements.

4.5 Least active pupils

Pupils identified as being least active at baseline (in the bottom 25% of MVPA, stratified by gender) will be analysed as a separate subgroup. Note that this analysis will be underpowered compared to the main analysis.

4.6 Intervention impact change over time

The main analysis for the primary outcome will be extended to explore whether the intervention impact changes over time. We will use models with a variable for time since start of intervention effects. We will model using discrete time effects based on the post-intervention terms. However, we will also use a natural cubic spline model to model any changes to the intervention effect over time, if there is sufficient data from the start of the intervention to the end of the study [27]. The spline will be constrained to

pass through the origin (i.e. zero at time intervention first implemented). The modelled splines, with confidence intervals, will be used to describe how intervention effects vary over time after intervention.

4.7 Mediation by active travel, clubs and motivation

If the primary analysis finds an intervention effect on weekday MVPA, we will explore whether this is mediated by pupil measures of physical activity motivation, active travel and/or attendance of clubs. These mediators are taken from the pupil questionnaire which is collected at three points: baseline, mid-point and end of trial.

Formal mediation analysis is not possible with such a complex model that involves random effects at multiple levels, changing control and intervention status over time and different measurement points for outcome and mediators. Instead, we will investigate mediation in an exploratory way, by fitting the following models:

(a) **impact of intervention on each mediator**: this will be based on a two-step SWRCT, with discrete time effects (following that used for the secondary outcome of pupil wellbeing, using linear or logistic models as appropriate)
(b) **impact of intervention on weekday MVPA, adjusted for each mediator**: this will also be a two-step SWRCT using only the measurement points where both MVPA and mediator are available, but will otherwise follow the same model as the primary analysis, with a spline for the time trend and random effects for school, pupil and school-time.

These models will be compared with the primary analysis of the effect of intervention on MVPA, using a combination of model fit (AIC), magnitude and direction of estimates, standard errors and p-values to assess the extent to which these mediating factors are present, and if so whether they could be responsible for all, a large or a small part of any increase in MVPA. If more than one mediator is identified, we will extend model (b) to include all identified mediators.

Due to the complexity, no formal hypothesis testing will be presented.

4.8 Other analyses

To explore the potential impact of the PE and Sports Premium changes, we will undertake an additional analysis to estimate the counterfactual MVPA under control conditions, by replacing the combined secular and seasonal spline in the main model with two separate splines. We will report the modelled secular trend spline with confidence intervals, and estimate an overall impact of the policy change over the academic year. This analysis will be exploratory in nature.

5. Sensitivity analyses

If schools fail to implement the intervention when scheduled, then a sensitivity analysis will check that changes to the planned schedule did not result in imbalance between control and intervention in terms of school covariates. The tests for school covariate balances used for randomisation will be reapplied using the actual timing of school interventions. Also, a sensitivity analysis will check for any significant changes to seasonality estimates if actual intervention times are used instead of intention to treat.

6. Indicative Tables (subject to change)

6.1 Baseline data

Table A	Pupil measures at baseline
Table B	Pupil characteristics at baseline
Figure C	Balance of school co-variates across sequence allocation

6.2 Numbers analysed

Table D	Missing data
Figure E	Numbers participating, flow diagram

6.3 Outcomes and model estimates

Table F	Outcome measures: raw data summaries
Table G	Outcome measures: modelled effect size
Table H	Sub-group analysis
Table I	Least active pupils: modelled effect size
Table J	Intervention effect over time
Table K	Mediating variables: modelled effect size
Table L	Weekday MVPA & mediator: modelled effect size

6.1 Baseline data

Table A Pupil measures at baseline

Measure	**N**	**Min**	**Mean**	**Max**	**SD**	**Sequence (mean)**				
						A	**B**	**C**	**D**	**E**
Weekday MVPA										
Total activity volume										
Weekday light activity										
Weekday sedentary activity										
Weekend MVPA										
Stirling wellbeing score										
Relative autonomy index										
% using Active Travel		**NA**		**NA**	**NA**					
% using after school club		**NA**		**NA**	**NA**					
% using outside school club		**NA**		**NA**	**NA**					

Table B Pupil characteristics at baseline

			Sequence				
	N	**All**	**A**	**B**	**C**	**D**	**E**
% girls							
% non-white							
% with SEN or disabilities							
% with pupil premium							
% with English as additional language							

Figure C: Balance of school co-variates across sequence allocation

6.2 Numbers analysed

Table D Missing Data

Measure	Number of records	% missing
School termly absentee data		
Pupil ethnicity		
Pupil SEN/disability		
Pupil English as additional language		
Pupil premium		
Pupil gender		
Well-being (Stirling score)		

Motivation (RAI)		
Weekday MVPA		
Weekday light activity		
Weekday sedentary activity		
Weekend MVPA		

Figure E Numbers participating, flow diagram

6.3 Outcomes and model estimates

Table F Outcome measures: raw data summaries

Outcome measure	Pre-intervention	Post-intervention
Weekday MVPA		
Total volume of activity		
Weekday light activity		
Weekday sedentary activity		
Weekend MVPA		
Stirling Wellbeing Score		
% Pupil absence		

Table G: Outcome measures: modelled effect size

Outcome measure	**Modelled effect size**	**p-value**	**95% confidence interval**
Weekday MVPA			
Total volume of activity			
Weekday light activity			
Weekday sedentary activity			
Weekend MVPA			
Stirling Wellbeing Score			
% Pupil absence			

Table H Sub-group analysis

Outcome measure	**Girls**			**Boys**		
	Modelled effect size	p-value	95% confidence interval	Modelled effect size	p-value	95% confidence interval
Weekday MVPA						
Total time active						

Weekday light activity						
Weekday sedentary activity						
Weekend MVPA						
Stirling Wellbeing Score						
Outcome measure	Pupil premium			No pupil premium		
	Modelled effect size	p-value	95% confidence interval	Modelled effect size	p-value	95% confidence interval
Weekday MVPA						
Total time active						
Weekday light activity						
Weekday sedentary activity						
Weekend MVPA						
Stirling Wellbeing Score						

Table I Least active pupils N = : modelled effect size

Outcome measure	**Modelled effect size**	**p-value**	**95% confidence interval**
Weekday MVPA			
Total time active			
Weekday light activity			
Weekday sedentary activity			
Weekend MVPA			

Table J Intervention effect over time

	After 1 term	**2 terms after intervention**	**3 terms after intervention**

Outcome measure	Modelled effect size	p-value	95% confidence interval	Modelled effect size	p-value	95% confidence interval	Modelled effect size	p-value	95% confidence interval
Weekday MVPA									
Total time active									
Weekday light activity									
Weekday sedentary activity									
Weekend MVPA									

Table K: Mediating variables: modelled effect size

Outcome measure	Modelled effect size	p-value	95% confidence interval
Active travel			
Attendance			
After school clubs			
Outside school clubs			
Relative autonomy index			

Table L: Weekday MVAP & mediator: modelled effect size

Mediating variable	Modelled effect size	p-value	95% confidence interval
Active travel			
Attendance			
After school clubs			
Outside school clubs			
Relative autonomy index			

7. References

1. Aira T, V.T., Heinonen OJ, et al., *Physical activity from adolescence to young adulthood: patterns of change, and their associations with activity domains and sedentary time.* Int J Behav Nutr Phys Act, 2021. **18(1)**: p. 85.
2. Twisk JW, K.H., van Mechelen W., *Tracking of activity and fitness and the relationship with cardiovascular disease risk factors.* Med Sci Sports Exerc., 2000. **32(8)**: p. 1455-1461.
3. UK Chief Medical Officers, *UK Chief Medical Officers' Physical Activity Guidelines*. 2019, Department of Health and Social Care,: London, UK.
4. Education, D.f., *Pupil premium: overview 2025*. 2025.
5. Hussey MA, H.J., *Design and analysis of stepped wedge cluster randomized trials.* Contemp Clin Trials., 2007. **Feb;28(2):**: p. 182-91.
6. Salway, R., et al., *Designing stepped wedge trials to evaluate physical activity interventions in schools: methodological considerations.* Int J Behav Nutr Phys Act, 2025. **22**(1): p. 22.
7. Metcalf, B., W. Henley, and T. Wilkin, *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**: p. e5888.
8. Salway, R., et al., *School-level intra-cluster correlation coefficients and autocorrelations for children's accelerometer-measured physical activity in England by age and gender.* BMC Med Res Methodol, 2024. **24**(1): p. 179.
9. Jago, R., et al., *Association of BMI category with change in children's physical activity between ages 6 and 11 years: a longitudinal study.* Int J Obes (Lond), 2020. **44**(1): p. 104-113.
10. Salway, R., et al., *School-level intra-cluster correlation coefficients and autocorrelations for children's accelerometer-measured physical activity in England by age and gender.* BMC Medical Research Methodology, 2024. **24**(1): p. 179.
11. Hale, D.R., et al., *School-level variation in health outcomes in adolescence: analysis of three longitudinal studies in England.* Prev Sci, 2014. **15**(4): p. 600-10.
12. Moulton, L.H., et al., *Statistical design of THRio: a phased implementation clinic-randomized study of a tuberculosis preventive therapy intervention.* Clinical Trials, 2007. **4**(2): p. 190-199.
13. Liddle, I. and G.F.A. Carter, *Emotional and psychological well-being in children: the development and validation of the Stirling Children's Well-being Scale.* Educational Psychology in Practice, 2015. **31**(2): p. 174-185.
14. Moustaka F.C., V.S.P., Vazou S., Kaperoni M., Markland D., *Initial validity evidence for the Behavioural Regulation in Exercise Questionnaire – 2 among Greek exercise participants.* European Journal of Psychological Assessment, 2010. **26**: p. 269-276.
15. van Hees V, M.J., Sabia S, Patterson M, Fang Z, Heywood J, Capdevila Pujol J, Kushleyeva L, Chen M, Yerramalla M, Bos P, Sanders T, Zhao C, Meneghel Danilevicz I, Barreto Mesquita V, Segantin G,, *GGIR: Raw Accelerometer Data Analysis.* 2025, Medical Research Council UK, Accelting, French National Research Agency (2025).

16. Hildebrand, M., et al., *Evaluation of raw acceleration sedentary thresholds in children and adults*. Scandinavian Journal of Medicine & Science in Sports, 2017. **27**(12): p. 1814-1823.
17. HILDEBRAND, M., et al., *Age Group Comparability of Raw Accelerometer Output from Wrist- and Hip-Worn Monitors*. Medicine & Science in Sports & Exercise, 2014. **46**(9): p. 1816-1824.
18. Child Outcomes Research Consortium. *The Stirling Children's Wellbeing Scale*. [cited 2025 03/09/2025]; Available from: https://www.corc.uk.net/outcome-measures-guidance/directory-of-outcome-measures/stirling-childrens-wellbeing-scale/.
19. Sebire SJ, J.R., Fox KR, Edwards MJ, Thompson JL, *Testing a self-determination theory model of children's physical activity motivation: a cross-sectional studay.* International Journal of Behavioral Nutrition and Physical Activity, 2013. **10**.
20. Bangor University. *Scoring the BREQ*. [cited 2025 03/09/2025]; Available from: http://exercise-motivation.bangor.ac.uk/breq/brqscore.php.
21. Atkin, A.J., et al., *Seasonal Variation in Children's Physical Activity and Sedentary Time*. Med Sci Sports Exerc, 2016. **48**(3): p. 449-56.
22. Wood, S.N., *Thin plate regression splines*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2003. **65**(1): p. 95-114.
23. R Core Team, *R: A Language and Environment for Statistical Computing*. 2025, R Foundation for Statistical Computing: Vienna, Austria.
24. Wood, S.N., *Generalized Additive Models: An Introduction with R (2nd edition)*. 2017: Chapman and Hall/CRC.
25. Hooper R, T.S., de Hoop E, Eldridge S., *Sample size calculation for stepped wedge and other longitudinal cluster randomised trials*. Stat Med., 2016. **Nov 20**(35(26)): p. 4718-4728.
26. Karla Hemming, M.T., Joanne E McKenzie, Richard Hooper, Andrew Copas, Jennifer A Thompson, Mary Dixon-Woods, Adrian Aldcroft, Adelaide Doussau, Michael Grayling, Caroline Kristunas, Cory E Goldstein, Marion K Campbell, Alan Girling, Sandra Eldridge, Mike J Campbell, Richard J Lilford, Charles Weijer, ,Andrew B Forbes, Jeremy M Grimshaw, *Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration*. BMJ, 2018. **363**.
27. Kenny A, V.E., Xia F, Heagerty PJ, Hughes JP., *Analysis of stepped wedge cluster randomized trials in the presence of a time-varying treatment effect*. Stat Med., 2022 **Sep 30**(41(22)): p. 4311-4339.