Study Title	The Walking In ScHools (WISH) Trial: a cross-border trial to evaluate a walking intervention in adolescent girls.
Funding organisation	Cross-border Healthcare Intervention Trials in Ireland Network (CHITIN)
Principal Investigator	Professor Marie Murphy University of Edinburgh & Ulster University mh.murphy@ulster.ac.uk
Co-Principal Investigator	Professor Alison Gallagher Ulster University am.gallagher@ulster.ac.uk
Co-Investigator	Dr Angela Carlin Ulster University a.carlin1@ulster.ac.uk
Co-Investigator	Dr Ian M. Lahart (trial statistician) University of Wolverhampton I.Lahart@wlv.ac.uk
Co-Investigator	Professor Russ Jago University of Bristol Russ.Jago@bristol.ac.uk
Co-Investigator	Dr Maria Faulkner Atlantic Technological University maria.faulkner@lyit.ie
WISH Trial Manager	Dr Maria O'Kane Ulster University m.okane@ulster.ac.uk
WISH Trial Manager	Dr Leanne Doherty Ulster University lc.doherty@ulster.ac.uk
Research Associate	Mr Gary McDermott Ulster University G.McDermott@ulster.ac.uk

[c h i t i n]
kite-tin
Cross-border
Healthcare Intervention Trials
in Ireland Network

Interreg
Northern Ireland - Ireland - Scotland

HSC Public Health Agency
Research and Development

HRB Health Research Board

A clustered randomised controlled trial to evaluate the effectiveness of a peer-led school-based walking intervention on adolescent girls' physical activity: the Walking In ScHools (WISH) study statistical analysis plan

Version: 1.0

Date: 09/03/2023

METHODS

Trial design

The WISH study is a school-based parallel-group clustered randomised controlled trial (c-RCT). Eighteen post-primary schools were recruited across the border region of Ireland and Northern Ireland. Following completion of baseline data collection, schools were randomised using a 1:1 allocation. Study outcomes were assessed at four timepoints: baseline (T0), mid-intervention (T1: 9-12 weeks post-baseline), end of intervention (T2: 18-21 weeks post-baseline), and follow-up (T3: 51-54 weeks post-baseline; see Figure 1). This current study received ethics approval from Ulster University Research Ethics Committee (Ref: REC/19/0020) on 20th June 2019. In addition to prospective registration at the International Standard Randomised Controlled Trial Number registry (1), the detailed protocol describing the methods of the study was published and is available online open access at BMC Public Health (2).

Sample size calculations

Sample size calculations were based on the WISH feasibility trial (3), which detected a mean difference of 11.41 min in total physical activity (light, moderate and vigorous intensity physical activity) measured using accelerometery between the intervention (2 schools consisting of a total of 79 children) and control (4 schools consisting of 84 children in total) groups. To detect a difference of 11.41 min total physical activity per day between groups, assuming a standard deviation (SD; pooled intervention and control group SD) of 25 min in total physical activity, a power of 80%, a significance of 0.05, an average cluster size of 20 children, an intra-class correlation of 0.03 [calculated from pilot data (3)], and a coefficient of variation of 0.46, we needed a sample size of 15 schools increasing to 16 schools (8 per study arm) to allow for 10% cluster attrition. To allow for 20% loss to follow-up we recruited at least 24 children per cluster (at least 384 children in total).

Based on recruitment rates of 34% in our feasibility study (3), all schools with at least 80 girls across years 9–10 (Northern Ireland) or total enrolment of >240 girls (Ireland) were eligible to participate in the study. From the Northern Ireland School Census (4), we estimated that 77 schools met these criteria and based on 2016 school recruitment data for Ireland (5) 54 post-primary schools met these criteria.

Randomisation and blinding

After data collection, schools were randomly allocated to control (n = 9) or intervention (n = 9). Randomisation was stratified by country, and to ensure allocation concealment, random allocation was performed by faculty staff who were blind to school identity and independent of the study team. Due to the nature of the intervention, it was not possible to blind participants.

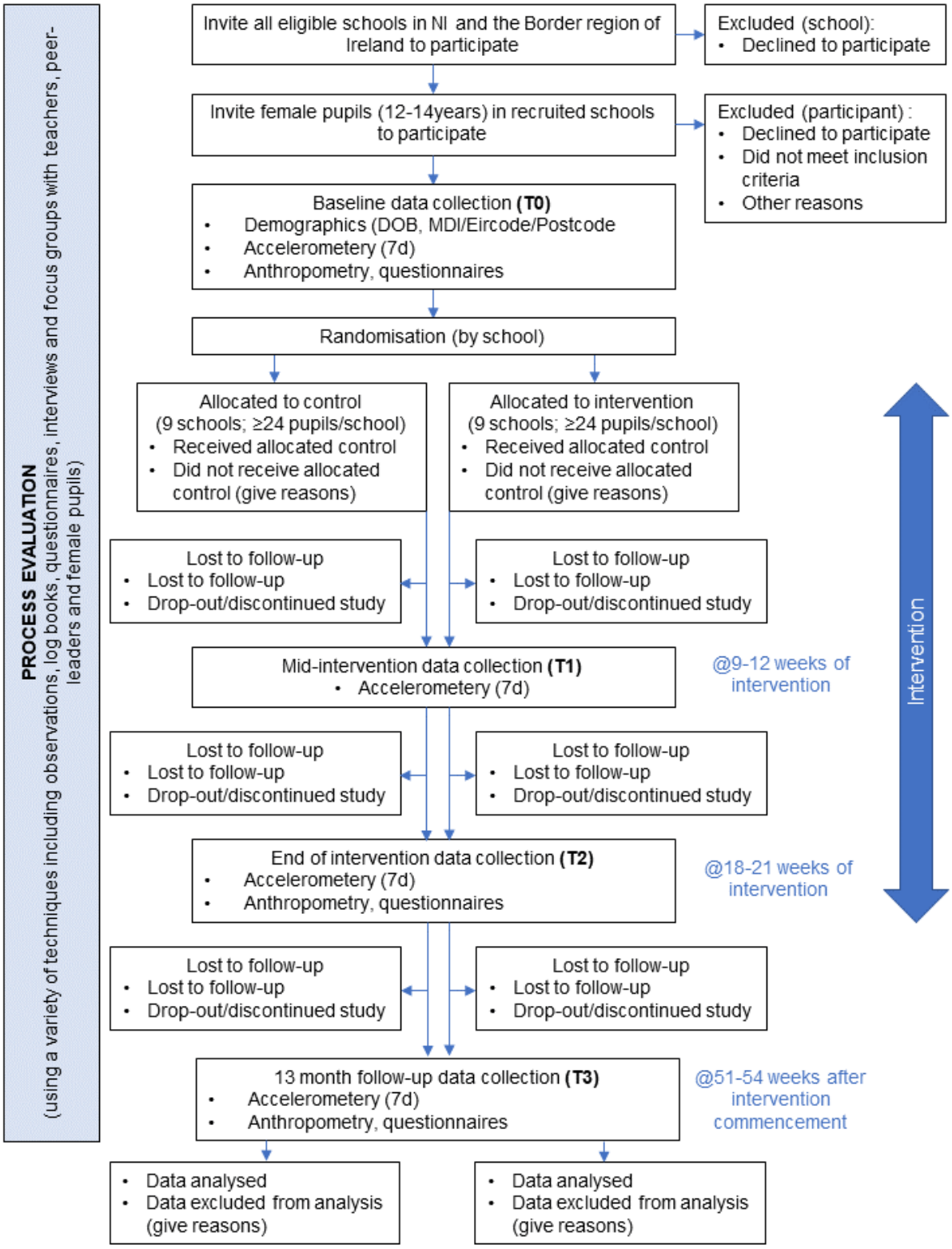

Figure 1. Flow of participants through study

Outcomes

Primary outcome

The primary outcome was total physical activity (counts per minute) of pupils at the end of the intervention (T2) measured using the Actigraph GT3X accelerometer (Actigraph LLC, Florida) worn for seven days (Table 1).

The device was placed on an elastic waist band and pupils were asked to always wear the accelerometer, remove it only for bathing, water-based activities, such as swimming and when asleep. Pupils were asked to wear the accelerometer on their right hip and to ensure that it was worn in the same position each day. Pupils were asked to wear the accelerometer for seven consecutive days and will be included in the analysis if they have ≥2 valid weekdays of data (500 mins/day) (6).

During measurement periods, pupils were asked to keep a log of when they wore and took off the accelerometer to encourage compliance with the wear-time protocol as recommended by Trost and colleagues (7). Minutes of total physical activity (light, moderate and vigorous) per day were estimated using the Evenson et al. (8) cut-off points. A sampling epoch of 15 s was employed during data collection. Periods of ≥60 min of zero counts was categorised as 'non-wear' and removed.

Secondary outcomes

It is also important to note, that there are no statistical power calculations for our secondary outcomes (Table 1) and as such all analyses will be presented with point estimates and 95% confidence intervals but without p-values.

Accelerometer data was used to calculate:

- Total physical activity (counts per minute) at mid-intervention (T1), and follow-up (T3)
- Time spent in sedentary behaviour and light, moderate and vigorous intensity physical activity at mid-intervention (T1), end of intervention (T2), and follow-up (T3)
- Proportion of pupils meeting current physical activity recommendations at mid-intervention (T1), end of intervention (T2), and follow-up (T3) (9)

The following secondary outcomes were measured at baseline (T0), end of intervention (T2), and follow-up (T3) to assess between group differences and changes over time:

- Height (cm) and weight (kg) was measured to the nearest 0.1 cm and 0.1 kg, respectively, using a freestanding stadiometer (Leicester Height Measure), and digital scales (Seca 877) to calculate body Mass Index (BMI). BMI was converted to an age-specific and gender-specific z-score (10,11).
- Waist and hip circumference was measured to the nearest 0.1 cm using an anatomical measuring tape and waist-to-hip ratio was calculated.
- Coping, resilience, and cognitive reappraisal (12).
- Sleep quality, duration, and efficiency (The Pittsburgh Sleep Quality Index) (13).
- Social media use, social integration, and emotional connection to social media (14).
- Body weight and appearance satisfaction (15).

At baseline (T0) the following descriptive data was collected:

- Date of birth
- School postcode/Eircode to derive Index of Multiple Deprivation (IMD)

Data storage and integrity

Data were entered electronically on a secure file storage system and password protected. Data were anonymised by assigning a unique identification number to each pupil. All paper data entry forms were entered and checked by another researcher. We inspected data files for omissions and errors. Double data checking was completed for the primary (total physical activity) and secondary outcomes. All inconsistencies or implausible values were investigated and rectified. Missing and out-of-range data were checked against paper data collection forms to confirm values and correct them if necessary.

Table 1. Quantitative outcome measures and assessment timepoints

Outcome measure	**Baseline (T0)**	**Mid-intervention (T1)**	**End of intervention (T2)**	**Follow-up (T3)**
Demographic information (age, postcode/Eircode)	X			
Total physical activity (accelerometery)	X	X	X	X
Time spent in sedentary behaviour, and low, moderate, and vigorous intensity physical activity (accelerometery)	X	X	X	X
Proportion of pupils meeting current physical activity recommendations (accelerometery)	X	X	X	X
Body Mass Index z-scores (height and weight)	X		X	X
Waist: hip ratio (waist and hip circumference)	X		X	X
Coping, resilience, and cognitive reappraisal (Emotion Regulation Questionnaire for Children)	X		X	X
Sleep quality, duration, and efficiency (Pittsburgh Sleep Quality Index)	X		X	X
Social media use, social integration, and emotional connection to social media	X		X	X
Body weight and appearance satisfaction	X		X	X
Self-efficacy for physical activity and walking	X		X	X
Health-related quality of life (Kidscreen-10)	X		X	X
Reasons for engaging in physical activity (BREQ-3)	X		X	X

STATISTICAL ANALYSIS

General analyses principles

The statistician performing the data analyses (IML) was blinded to allocation throughout the study and statistical analysis will only be undertaken when all data has been collected (T3). Outcome data will be reported in accordance with Consolidated Standards of Reporting Trials (CONSORT) guidelines (16).

Participant population

The analysis population will consist of all eligible, randomised children excluding pupils who left the trial geographical area prior to the collection of baseline measurement and pupils who withdrew from the study and were unwilling for their data to be used.

The main analysis for each outcome will be undertaken on a modified 'intention-to-treat' (ITT) basis, i.e. all pupils with a recorded outcome at T0 and T2 will be included in the analysis and will be analysed according to the group to which they were allocated. We will adopt a similar approach to the analysis of outcomes at T0 and T1 and T0 and T3. No pupils crossed over from their allocated group to another (i.e. change from intervention to control, or vice-versa). Pupils will be included in the analysis regardless of adherence to the physical activity intervention.

The population for the primary outcome analysis will consist of all randomised pupils for whom sufficient baseline accelerometer PA data were collected (i.e. pupils meeting the wear time criterion described above) and for whom data were available at T2 (i.e. complete case/observed outcomes data analysis). As the full analysis population will exclude the number of pupils lost to follow-up (due to withdrawal, unavailability, or insufficient accelerometer wear time), a sensitivity analysis will be performed to account for all randomised pupils if we can assume the data are likely missing at random. We have no a priori reason to assume that data will be missing at random, therefore, we do not anticipate that this sensitivity analysis will be appropriate. The number of and reasons for loss to follow-up will be reported, and we will also compare the baseline characteristics of pupils included in the analysis to those excluded from our analysis.

Level of confidence and p values

Statistical tests and confidence intervals will be two-sided. Between-group comparisons will be calculated and presented with 95% confidence intervals wherever possible. The statistical significance level will be set at the 5% level for our primary analysis. For our secondary outcome analyses, it is also important to note, however, that there are no statistical power calculations for these assessments and as such all analyses will be presented with point estimates and 95% confidence intervals but without p-values. To control for multiple testing in our secondary analysis, we will adjust the false discovery rate using the Benjamini-Hochberg Procedure (17).

Unadjusted vs. adjusted analyses

All comparative analyses will allow for the clustered nature of the of the data to ensure correct confidence intervals and type I error rates are calculated. For each outcome, unless otherwise specified, the main analysis will be the covariate-adjusted analysis, which will adjust for the child-level covariates (level 1): baseline value for the outcomes of interest [e.g., baseline total physical activity

(counts per minute) in the primary analysis], age, and BMI z-score. In the interests of parsimony, covariates will only be retained if their inclusion results in a significant improvement in fit statistics. Unadjusted between-group differences will be presented for completeness.

Missing data

If a pupil was not available for the collection of outcome measures, a second visit was organised to try to capture the missing measures. However, if the participant was absent at the second visit, they were recorded as absent for that timepoint and further visits for that specific timepoint were not arranged. The proportion of children missing each outcome will be summarised for each allocated group and at each time point, with reasons for missing outcomes documented wherever possible.

The main analysis of the primary outcome uses the accelerometer PA data at post-intervention (T2) months which could be missing for a number of reasons:

1. Parent/carer opts pupil out of trial before follow-up data collection.
2. Pupil refused to wear accelerometer or return an accelerometer.
3. Pupil wore accelerometer for less than the required amount of time [i.e. ≥2 valid weekdays of data (500 mins/day].
4. Accelerometer malfunction.
5. Pupil moved out of the trial geographical area before follow-up data collection.
6. Pupil is absent on day of measurement and subsequent follow-up visits.
7. Pupil withdrew from the study.
8. Some other reason.

Due to the nature of collection of secondary outcomes (via Qualtrics on iPads), we do not anticipate missing data for secondary outcomes, but potential reasons for missing secondary data may include:

- Pupil declined anthropometric measurement.
- Pupil declined to answer a question on a paper based self-report measure.
- Pupil entered data incorrectly.
- Pupil is absent on day of measurement and subsequent follow-up visits.
- Pupil withdrew from the study.
- Equipment malfunction, paper questionnaire used.

The patterns of missing data for both primary and secondary outcomes and reasons for missingness (if known), will be explored and summarised for both study arms. There is no a priori reason to assume that pupils who are lost to follow-up are missing at random. However, if we find no reason to suggest that data are not missing at random, then in a sensitivity analysis, data will be imputed using a

multivariate normal distribution imputation model to allow a full ITT analysis. Missing values will be imputed separately by allocated randomisation group and will comply with the multi-level character of the data. However, if there is evidence that the data are 'missing not at random', then no sensitivity analysis will be performed. We will compare descriptively differences in baseline physical activity, BMI, and deprivation of those missing and those not missing data.

Presentation of comparative analyses

For the primary analysis and each of the other continuous outcomes. The mean and standard deviation for each allocated group will be presented, together with the mean between-group difference, 95% confidence interval for the differences, and the p-value (primary analysis only). For binary outcomes, the percentage and frequency of pupils in the outcome category of interest (e.g. percentage meeting physical activity guidelines) will be presented for each allocated group, along with the odds ratio for the intervention effect, and 95% confidence interval for the odds ratio. Similarly, for ordinal outcomes, the percentage and frequency of pupils in each outcome category will be reported for each allocated group, along with the odds ratios, and 95% confidence intervals for the odds ratios. In addition, the intracluster correlation coefficient (ICC) will be reported for each outcome, based on the adjusted analyses, together with 95% confidence interval.

Proposed analyses

Baseline

Baseline characteristics, collected at the time of baseline assessments (T0), will be cross tabulated according to the randomised group to check for appropriate balance and to provide an overview of the study population, both at the school and pupil levels.

At the pupil level, variables will include age at baseline data collection, baseline measures of all physical activity, anthropometric measurements, and all self-report outcomes. IMD is the only school level variable. The baseline characteristics of each group will be summarised as the mean, standard deviation, and range for continuous, approximately normally distributed variables (investigated via Shapiro-Wilkes normality tests and visual inspection of visual inspection of histograms and Q-Q plots); medians, interquartile range, and minimum-maximum for continuous, non-normal, skewed variables; frequencies and percentages of pupils/schools in each category for categorical variables.

We expected that children in both allocated groups will, on average, be similar, given the randomisation procedure. Formal statistical comparison at baseline of randomised groups is not recommended (18) and thus will not be undertaken—only descriptive data, as described above, will be presented. If substantial baseline imbalance between randomised groups is identified in terms of any relevant variables not already being adjusted for in the primary analysis, additional adjusted sensitivity analyses may be performed, to allow for such variable(s), in addition to the prespecified variables for adjustment, to assess the robustness of the primary analysis.

Primary analysis of primary outcome

To address our primary research question, a modified-ITT analysis will be conducted for the primary outcome measure (total physical activity, minutes per week, T0 to T2) using multi-level modelling. To

allow for the clustered nature of our data, hierarchical models with two levels (school, pupil) will be fitted, while controlling for baseline total physical activity at the pupil level. This analysis will include the school the pupil attended as a random effect.

At the school level we will only include group allocation (i.e., intervention, control) as an explanatory variable. At the pupil level, we will include baseline total physical activity (counts per minute), age, and BMI z-score as explanatory variables. We will also report the analysis unadjusted for baseline age and BMI z-scores (see 'Unadjusted and adjusted analyses' section). ML-win will be used to conduct the analysis (19). We will use R Statistical Software to obtain descriptive statistics (20).

REFERENCES

1. ISRCTN12847782: The walking in schools (WISH) trial [Internet]. 2019 [cited 2023 Mar 6]. Available from: https://www.isrctn.com/ISRCTN12847782

2. O'Kane SM, Carlin A, Gallagher AM, Lahart IM, Jago R, Faulkner M, et al. A study protocol for a clustered randomised controlled trial to evaluate the effectiveness of a peer-led school-based walking intervention on adolescent girls' physical activity: the Walking In ScHools (WISH) study. BMC Public Health. 2020 Apr 21;20(1):541.

3. Carlin A, Murphy MH, Nevill A, Gallagher AM. Effects of a peer-led Walking In ScHools intervention (the WISH study) on physical activity levels of adolescent girls: a cluster randomised pilot study. Trials. 2018 Jan 11;19:31.

4. School enrolments - 2018/19 statistical bulletins | Department of Education [Internet]. Education. 2018 [cited 2023 Mar 6]. Available from: https://www.education-ni.gov.uk/publications/school-enrolments-201819-statistical-bulletins

5. Department of Education [Internet]. [cited 2023 Mar 6]. Available from: https://www.gov.ie/en/organisation/department-of-education/?referrer=http://www.education.ie/en/Publications/Statistics/Data-on-Individual-Schools/#

6. Willis K, Tibbitts B, Sebire SJ, Reid T, MacNeill SJ, Sanderson E, et al. Protocol for a cluster randomised controlled trial of a Peer-Led physical Activity iNtervention for Adolescent girls (PLAN-A). BMC Public Health. 2019 May 28;19(1):644.

7. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc. 2005 Nov;37(11 Suppl):S531-543.

8. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci. 2008 Dec;26(14):1557–65.

9. Physical activity guidelines: UK Chief Medical Officers' report [Internet]. GOV.UK. 2020 [cited 2023 Mar 6]. Available from: https://www.gov.uk/government/publications/physical-activity-guidelines-uk-chief-medical-officers-report

10. Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. Arch Dis Child. 1995 Jul;73(1):25–9.

11. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ. 2000 May 6;320(7244):1240.

12. Gullone E, Taffe J. The Emotion Regulation Questionnaire for Children and Adolescents (ERQ–CA): A psychometric evaluation. Psychol Assess. 2012;24:409–17.

13. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. Psychiatry Res. 1989 May 1;28(2):193–213.

14. Woods HC, Scott H. #Sleepyteens: Social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. J Adolesc. 2016 Aug;51:41–9.

15. Kelly Y, Zilanawala A, Booker C, Sacker A. Social Media Use and Adolescent Mental Health: Findings From the UK Millennium Cohort Study. eClinicalMedicine. 2018 Dec 1;6:59–68.

16. Schulz KF, Altman DG, Moher D, the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 2010 Mar 24;8(1):18.

17. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995;57(1):289–300.

18. Bland JM, Altman DG. Comparisons against baseline within randomised groups are often used and can be highly misleading. Trials. 2011 Dec 22;12(1):264.

19. Charlton C, Rabash J, Browne WJ, Healy M, Cameron B. MLwiN Version 3.06. Centre for Multilevel Modelling, University of Bristol.; 2022.

20. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Internet]. 2022. Available from: https://www.R-project.org/