International Journal of Clinical and Health Psychology 26 (2026) 100687

Contents lists available at ScienceDirect

International Journal of Clinical and Health Psychology

journal homepage: www.elsevier.es/ijchp

Timing of MVPA and psychosocial outcomes in adolescents with overweight/obesity

Siyu Pan[a,1], Peng Wang[a,b,c,1], Yiming Tao[a], Zhihao Zhang[a], Fabian Herold[d], Matthew Heath[e,f,g], Alyx Taylor[h], Vahid Farrahi[i], Cassandra J. Lowe[j], André O. Werneck[k], Brendon Stubbs[l], David Moreau[m], Myrto F. Mavilidi[n,o], Fred Paas[c,p], Tianze wuhao[a], Qian Yu[a], Jinming Li[a], Linjing Zhou[a], Ziquan Cai[a], Kaiqi Guan[a], Zijun Liu[a], Weijia Zhu[a], Xun Luo[a], Xia Xu[a,*], Liye Zou[a,*]

[a] Body-Brain-Mind Laboratory, School of Psychology, Wuhan Sports University, China
[b] Department of Language, Literature, and Communication, Faculty of Humanities, Vrije Universiteit Amsterdam, 1081 HV, Amsterdam, The Netherlands
[c] Department of Psychology, Education, and Child Studies, Erasmus University Rotterdam, Rotterdam, P.O. Box 1738, The Netherlands
[d] Department of Physiology, Faculty of Medicine, HMU Health and Medical University Erfurt, Erfurt, Thuringia, 99089, Germany
[e] School of Kinesiology, Faculty of Health Sciences, University of Western Ontario, 1151 Richmond St, London, ON N6A 3K7, Canada
[f] Graduate Program in Neuroscience, University of Western Ontario, 1151 Richmond St, London, ON N6A 3K7, Canada
[g] Canadian Centre for Activity and Aging, University of Western Ontario, 1201 Western Rd, London, ON, N6G 1H1 Canada
[h] School of Rehabilitation, Sport and Psychology, AECC University College, Bournemouth, BH5 2DF, UK
[i] Research Group of Data Analytics in Sport Science, Institute for Sports and Sport Science, TU Dortmund University
[j] Department of Psychology, University of Exeter, UK
[k] Center for Epidemiological Research in Nutrition and Health, Department of Nutrition, School of Public Health, Universidade de São Paulo, São Paulo, Brazil
[l] Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK
[m] School of Psychology and Centre for Brain Research, University of Auckland, Auckland, 1030 New Zealand
[n] Faculty of Education, Southern Cross University, Lismore, Australia
[o] Physical Activity, Sport and Exercise (PASER) Research Theme, Southern Cross University, Lismore, Australia
[p] School of Education, University of New South Wales, Sydney, NSW 1466, Australia

ARTICLE INFO

Keywords:
Adolescence
Overweight/obesity
MVPA timing
Accelerometry
SDQ
Causal forest
Heterogeneous exposure effects

ABSTRACT

Background: Adolescents with overweight/obesity have an elevated risk of mental health and behavioral difficulties. Exercise has been shown to impart psychological benefits to these individuals; however, whether the effects of moderate-to-vigorous physical activity (MVPA) delivery differ between weekdays and weekends is limited. Therefore, this study aimed to examine the effects of weekday and weekend MVPA at age 14 on internalizing and externalizing problems at age 17 among adolescents with overweight/obesity.

Methods: We analyzed data from two assessment waves of the UK Millennium Cohort Study (MSC): MCS6 (2015–2016; age ∼14) and MCS7 (2018–2019; age ∼17). Data were restricted to adolescents classified as overweight/obesity at MCS6 using the UK90 thresholds. Weekday and weekend MVPA were measured at age 14 using the wrist-worn GENEActiv accelerometer, including one pre-specified weekday and one weekend day. Outcomes at age 17 were parent-reported Strengths and Difficulties Questionnaire (SDQ) internalizing (emotional + peer problems) and externalizing (conduct + hyperactivity/inattention) composites. We estimated average treatment effects (ATEs) and conditional average treatment effect (CATEs, heterogeneous effects) using a causal forest framework (EconML) and adjusted for pre-exposure covariates (age, sex, body mass index, ethnicity, cognitive decision-making, household income, parental education, and parental mental health). Missing data were treated via K-Nearest Neighbors imputation.

Results: The analytic sample included 1,238 adolescents (mean age 14.25 years). Mean MVPA was higher on weekdays than weekends (135.74 ± 62.08 vs 113.80 ± 64.37 min/day). Covariate-adjusted average treatment effects (ATEs; per 1 min/day MVPA) were small and not statistically significant for internalizing or externalizing problems. Weekday MVPA ATEs were −0.0025 (95% CI −0.0062 to 0.0012) for internalizing and 0.0003

* Coresponding authors.
E-mail addresses: xiaxu0704@hotmail.com (X. Xu), liyezou123@gmail.com (L. Zou).
[1] Authors contributed to this manuscript equally.

https://doi.org/10.1016/j.ijchp.2026.100687
Received 24 February 2026; Accepted 15 April 2026
Available online 27 April 2026

(−0.0027 to 0.0033) for externalizing; weekend MVPA ATEs were −0.0008 (−0.0051 to 0.0035) and 0.0005 (−0.0026 to 0.0037), respectively. Heterogeneity was evident only for weekday MVPA effects on internalizing (22.98% with significant individual effects), with height as the strongest moderator (β = 0.0002; $p < 0.001$; $R^2 = 0.519$) and more negative CATEs among shorter adolescents (Q1 −0.004994 vs Q4 −0.001903; ANOVA F = 106.741, p < 0.001).

Conclusions: In adolescents with overweight/obesity, estimated average effects of weekday and weekend MVPA at the age of 14 on parent-reported internalizing and externalizing problems at age 17 were close to zero under a selection-on-observables framework. Any potential benefits may be subgroup-specific and context-dependent; the observed weekday-specific heterogeneity warrants replication with more reliable exposure measurement and a richer set of contextual covariates.

Introduction

Adolescence is a sensitive developmental period for both physical and mental health (Gordon et al., 2024). Adolescents living with overweight or obesity may face elevated risk of psychosocial difficulties (Mellin et al., 2002). This may be partly due to weight-related stigma (Puhl et al., 2020), body dissatisfaction (Gouveia et al., 2014), and peer-related challenges (Ames & Leadbeater, 2017) which contribute to internalizing symptoms (e.g., anxiety, low mood, social withdrawal) and externalizing problems (e.g., conduct problems, hyperactivity/inattention). This is concerning given the high and continually rising prevalence of adolescent overweight and obesity in the UK (over 30%) (NHS England, 2023) and worldwide (20%)(World Health Organization, 2024). For example, class III (severe) obesity has been especially prevalent in recent decades (BMI > 40) (Phelps et al., 2024). Obesity in adolescence frequently persists into adulthood and may therefore place young people on a trajectory of enduring physical and mental health disadvantage (Mumford et al., 2013). One of the most considered therapy for obesity/overweight is physical activity given its well-established efficacy in supporting psychological well-being. Public health guidelines recommend that children and adolescents accumulate an average of at least 60 minutes of moderate-to-vigorous physical activity (MVPA) per day(Bull et al., 2020).

Higher uptake of physical activity levels, especially MVPA, is often associated with better mental health in adolescents (Biddle & Asare, 2011), but findings are equivocal: Meta-analytic and longitudinal studies tend to find small associations with substantial between-study and between-sample variability (Biddle et al., 2019; Rodriguez-Ayllon et al., 2019). Two implications follow. First, much of this literature is observational, so mixed findings may reflect differences in confounding and reverse causality (for example, better mental health may enable greater activity), rather than the mental-health impact of MVPA itself. Second, even if MVPA is beneficial on average, the variability suggests that its effects are unlikely to be uniform and may depend on who is active and under what conditions. This heterogeneity may be particularly salient for adolescents with overweight/obesity, for whom psychosocial and physiological factors can shape whether MVPA is experienced as rewarding and protective (e.g., enjoyment, competence, social connection) or as stressful and avoidant (e.g., stigma, low perceived competence, negative peer experiences) (Puhl & Heuer, 2009). This provides the impetus for the field to move beyond investigating "average" effects and examine for whom and under what circumstances MVPA matters most. Recent work in causal machine learning adapts flexible prediction tools to estimate exposure-effect quantities (e.g., average and subgroup/individual differences) rather than only predicting outcomes (Wager & Athey, 2018). These methods combine ML flexibility with causal inference theories such as sample splitting ("honesty") and estimators designed to target heterogeneity. This helps avoid reliance on a small set of pre-specified interactions. One plausible and practically important source of variation worth exploring with this method is the timing of physical activity: whether physical activity is practiced during weekdays or weekends.

Weekday MVPA is often embedded in structured routines and settings (e.g., school physical education, commuting, organized sport), whereas weekend MVPA may be more discretionary and dependent on family resources, social opportunities, and access to safe environments (Fairclough et al., 2015). Given that typically weekdays and weekend days differ in structure, proportion of weekly physical activity volume, supervision, social contact, and predictability, the time spent on these days for MVPA may have different implications for health-related outcomes, as conceptualized in the structured day hypothesis (Brazendale et al., 2017). However, evidence regarding whether weekday versus weekend MVPA has differential downstream effects on internalizing and externalizing problems remains limited. This is particularly true among adolescents with overweight/obesity. Much of the prior literature is observational and thus vulnerable to confounding by a range of factors, such as socioeconomic status, family factors, baseline mental health, and correlated health behaviors (Bauman et al., 2012). In addition, standard analyses that focus only on population-average associations can mask meaningful individual differences (Kravitz et al., 2004).

To address this complexity, we used a causal forest approach, which is well suited to settings where the psychosocial effects of MVPA may vary across adolescents with overweight/obesity and where population-average models may mask clinically meaningful heterogeneity. Conventional regression typically estimates a single mean effect under strong functional-form assumptions and a limited set of pre-specified interactions(Athey & Imbens, 2019). By contrast, causal forests flexibly estimate covariate-adjusted average and individual-level differences in effects by capturing higher-order interactions among potential moderators (Athey et al., 2019). Because the analysis remains observational, these estimates should be interpreted under the measured-covariate assumptions, and unmeasured confounding cannot be excluded (Imbens & Rubin, 2015).

Using data from the UK Millennium Cohort Study, we examined adolescents with overweight/obesity at approximately 14 years who were followed up at approximately 17 years. Our primary objective was to estimate the effects of weekday and weekend MVPA at age 14 on internalizing and externalizing problems at age 17. Our secondary objective was to examine whether these effects varied across individuals and to identify characteristics that may explain such heterogeneity. To do so, we applied causal forests to objectively measured weekday and weekend MVPA to estimate both average effects and effect heterogeneity under a selection-on-observables framework.

Methods

Data source and sample

Data for this study were sourced from the UK MCS database (University College London, 2024). This study is a nationally representative birth cohort that followed 19,519 children born between 2000 and 2002 across the four devolved nations of England, Wales, Scotland, and Northern Ireland. The MCS assessments cover multiple dimensions of the cohort members' lives including, but not limited to, family-related (e.g., parental socioeconomic status, occupation, and educational attainment) and child-related characteristics (e.g., habitual physical activity levels, cognitive and mental health-related outcomes). The MCS was conducted in full compliance with the Declaration of Helsinki and informed consent was obtained from the parents or primary caregivers of the cohort members. Further information about the MCS is available at: https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/.

This study focused on data obtained from two different assessment waves (MCS6: from January 2015 - April 2016; MCS7: January 2018 - May 2019). At MCS6, participants were approximately 14 years old and at MCS7 they were approximately 17 years old. The target sample consisted of adolescents with overweight/obesity, classified according to the UK 1990 growth reference thresholds: overweight BMI above the 85th percentile for children of the same age and sex; obesity: BMI above the 95th percentile for children of the same age and sex. Adolescents with overweight/obesity were excluded if: (1) data of the Strengths and Difficulties Questionnaire (SDQ) were missing or invalid; (2) valid accelerometer monitoring day(s) were missing at MCS7 and MCS6, respectively; (3) adolescents with MVPA values exceeding the mean of the group by more than 4 standard deviations.

Measures

Exposure: physical activity time (weekday and weekend MVPA)

Physical activity in MCS6 was assessed using a wrist-worn triaxial accelerometer (GENEActiv; Activinsights Ltd., Kimbolton, UK), a validated device widely used in cohort studies (Antczak et al., 2021; Chen et al., 2023). Raw accelerometry was recorded at 30 Hz and processed in 5-second epochs, with ENMO (Euclidean Norm Minus One) calculated for each epoch(Bakrania et al., 2016). Participants were randomly assigned to wear the accelerometer on one pre-specified weekday and one weekend day, selected by the computer-assisted personal interviewing (CAPI) program during the interviewer's visit (excluding the visit day and the two following days). Data were processed using the GGIR (Generalized Gait and Inactivity Recognition) package (Migueles et al., 2019), which applies auto-calibration procedures and non-wear detection. A day was considered valid if wear time was $\geq$10 hours (University College London, 2024). MVPA was defined as the total minutes in which mean acceleration exceeded 100 mg, consistent with large-scale cohort protocols (Brady et al., 2023).

Due to the MCS6 design, only one valid weekday and one valid weekend day were available; therefore, weekday and weekend MVPA were treated as context-specific snapshots rather than stable estimates of habitual activity. Single-day measures may differ in reliability by day type (e.g., greater within-person variability and/or different wear-time patterns on weekends), such that weekday–weekend contrasts may partly reflect measurement error. Differential reliability would be expected to bias effect estimates toward the null and could contribute to weaker or less stable detection of heterogeneity for weekend MVPA.

Outcomes: internalizing and externalizing problems

Internalizing and externalizing problems were assessed using the Strengths and Difficulties Questionnaire (SDQ), a validated behavioral screening tool that measures emotional and behavioral attributes (Goodman, 1997). The questionnaire consists of 25 items, each rated on a Likert scale, and depending on the context of the study and age of participants, can be completed by parents, teachers, or the participants themselves. The SDQ measures five key dimensions: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. Each dimension is scored on a scale from 0 to 10, with higher scores indicating more severe problems. Internalizing problems were calculated by summing scores from emotional symptoms and peer relationship problems, while externalizing problems were derived from conduct problems and hyperactivity/inattention. At age 14 (MCS6), data on these mental health problems were obtained from parent-completed SDQ questionnaires. To maintain consistency, parent-reported SDQ scores were selected as the outcome variables at age 17 (MCS7).

Covariates

Covariates were selected based on findings from previous studies (Hope et al., 2019; Terhaag et al., 2021; Yang et al., 2025), and included the following variables: age at interview (measured to the nearest 0.1 year), gender (male/female), body mass index (BMI) calculated from measured height and weight, ethnic group classification (i.e., White, Mixed, Indian, Pakistani and Bangladeshi, Black or Black British, and Other Ethnic Group, including Chinese and other backgrounds). Organisation for Economic Co-operation and Development (OECD)-equivalised weekly household income, parental education level (i.e., the highest National Vocational Qualification [NVQ]-equivalent level attained by either parent; in the MCS derived variable used here, Level 5 was the highest category), and parental psychological health assessed via the Kessler 6 (K6) scale (i.e., evaluating symptoms of depression and anxiety experienced over the past 30 days, with higher scores indicating greater psychological distress (Kessler et al., 2002). Cognitive competence (i.e., the child's/adolescent's quality of decision-making in the Cambridge Gambling Task) was also included it as a covariate given that executive functions, such as decision-making, are associated with emotional problems and health-related behavior engagement in adolescents (Terhaag et al., 2021). To reduce confounding by prior psychosocial functioning and regression-to-the-mean, we additionally adjusted for baseline SDQ internalizing and externalizing composites measured at MCS6 (parent-reported), matching each MCS7 outcome.

Data analysis

Descriptive statistics were performed for all variables of interest. Normally distributed continuous variables are presented as means (M) and standard deviations (SD), whereas non-normally distributed continuous variables are expressed as medians and interquartile range. Categorical variables are summarized using frequencies (N) and percentages (%). Missing data were handled using the k-nearest neighbors (KNN) imputation method, in which missing values are estimated based on the values of the most similar cases in the dataset (Beretta & Santaniello, 2016; Troyanskaya et al., 2001). This method is based on distance metrics between sample features. For each sample with missing values the method identifies the K most similar "neighbor" samples and imputes the missing values based on the observed values of these neighbors. All relevant variables were included in the imputation process and an appropriate K-value was determined through methods such as cross-validation to ensure the robustness of the imputation.

Causal machine learning framework: causal random forest

We used a causal machine learning approach to estimate the causal effect of MVPA on subsequent mental health outcomes and to assess effect heterogeneity. Specifically, we applied a causal random forest (also referred to as a causal forest within the generalized random forest family) to estimate both the average treatment effect (ATE) and conditional average treatment effects (CATEs)(Athey et al., 2019; Wager & Athey, 2018).

General analytic steps

1. Define the estimand and variables: Specify the exposure (MVPA), outcomes (internalizing and externalizing composites), and a pre-exposure covariate set to control confounding.
2. Estimate heterogeneous effects using causal forest: Fit a causal forest that uses sample splitting/bootstrapping to learn treatment effect variation across the covariate space while adjusting for measured confounders.
3. Summarize population and individual-level effects: Derive the ATE as an aggregate of individual-level effect estimates and obtain CATEs for each participant.

4. Interrogate effect heterogeneity: Examine whether specific characteristics systematically explain variation in CATEs using downstream models and descriptive comparisons.
5. Conduct inference and robustness checks: Use appropriate inference procedures for forest-based estimators and evaluate sensitivity to modeling choices where relevant.

Estimating the ATE and CATEs (causal forest)

We used a generalized random forest (causal forest) model to simultaneously estimate the ATE and individual CATEs. The model controlled for pre-specified confounders and used bootstrapped ensembles to produce participant-specific treatment effect estimates. The sample ATE was obtained by aggregating individual CATE estimates, and statistical inference was performed using the forest-based asymptotic inference procedures available for this estimator.

Heterogeneous treatment effect analysis

To explore potential sources of treatment effect variation, the estimated individual CATEs were subsequently used as the dependent variable in linear regression models. Candidate moderators were included as key predictors to test whether they systematically explained variation in CATEs, thereby evaluating whether individual characteristics moderated the psychological benefits of physical activity.

Statistical tools and implementation

All analyses were conducted in Python (version 3.11.14; Conda environment ai-py311)(Python Software Foundation, 2025). Key packages included pandas, numpy, scikit-learn, econml, shap, and statsmodels. Causal forest estimation was implemented using EconML (CausalForest module)(Harris et al., 2020; Lundberg & Lee, 2017; McKinney, 2010; Pedregosa et al., 2011; Seabold & Perktold, 2010).

Assumptions

Our causal interpretation relies on four assumptions: (1) consistency, (2) no interference, (3) conditional exchangeability given measured covariates, and (4) positivity/overlap across covariate strata. Consistency assumes that the accelerometer-derived MVPA measure corresponds to a well-defined exposure (i.e., differences in measured MVPA reflect meaningful differences in activity rather than measurement artifacts). No interference assumes one adolescent's MVPA does not directly affect another adolescent's SDQ outcomes; while peers may co-participate in activity, the cohort design does not explicitly model social network spillovers, so any remaining interference would primarily bias inference. Conditional exchangeability assumes that, after adjusting for the pre-exposure covariate set (including baseline mental health), there are no unmeasured common causes of MVPA and later SDQ outcomes. In this context, potentially important unmeasured confounders may include school-level PE policies and extracurricular provision, peer network influences, neighborhood safety and access to activity spaces, and family-level support or constraints (e.g., transport, time, encouragement), which may shape both MVPA and mental health. Positivity/overlap assumes adequate representation of different MVPA levels across covariate profiles; we examined overlap diagnostically and avoided reliance on extrapolation to covariate regions with limited support.

Results

Participant characteristics and descriptive statistics

In the MCS6 survey, a total of 11,884 cohort members participated. Among them, 3,830 members were categorized as being overweight or obese according to the UK90 criteria. This accounted for 32.23% of the total respondents. After data extraction, 1,238 adolescents had valid data for internalizing problems, externalizing problems, and MVPA time (see Table 1)

Average treatment effects of weekday and weekend MVPA on SDQ outcomes

The estimation results based on the causal forest model (Table 2) show that across different exposure-outcome combinations, the ATEs were all relatively small and did not reach statistical significance. These effect estimates should be interpreted as covariate-adjusted (conditional on the observed covariate set) and therefore rely on the assumption that important confounders were measured. We note that unmeasured confounding cannot be ruled out. Specifically, weekday MVPA showed a slight negative effect on internal psychological state (ATE = -0.0025, 95% CI [-0.0062, 0.0012]) but its confidence interval included zero. For other combinations, including the effect of weekday MVPA on external behavioral performance (ATE = 0.0003, 95% CI [-0.0027, 0.0033]), the effects of weekend MVPA on both internal psychological state (ATE = -0.0008, 95% CI [-0.0051, 0.0035]), and external behavioral performance (ATE = 0.0005, 95% CI [-0.0026, 0.0037]), their confidence intervals all spanned zero. This indicates that no statistically significant effect was observed at the overall population level. Because ATEs are scaled per 1-minute MVPA, the corresponding effect of a 30-minute difference is obtained by multiplying the ATE by 30. For example, the estimated effect of weekday MVPA on internalizing problems (−0.0025 SDQ points per minute) corresponds to approximately −0.075 SDQ

Table 1
Overview of the sample characteristics of adolescent with overweight/obesity. BMI: Body-Mass-Index; CGT: Cambridge Gambling Task.

Variables	Sample size (n)	Total
Age	1238	14.25 ± 0.34
Body height (cm)	1238	164.76 ± 8.01
Body weight (kg)	1238	70.04 [62.4, 75.8]
BMI (kg·m^{-2})	1238	25.75 [23.25, 27.39]
CGT Quality of Decision Making in MCS6	1174	0.88 [0.81, 1]
MVPA Time on Weekday (min per day)	1152	135.74 ± 62.08
MVPA Time on Weekend (min per day)	1144	113.80 ± 64.37
Internalizing Problems Score in MCS6	1214	4.00 [1, 6]
Externalizing Conduct Score in MCS6	1214	4.15 [2, 6]
Internalizing Problems Score in MCS7	1238	4.16 [1, 6]
Externalizing Conduct Score in MCS7	1237	3.54 [1, 5]
Obesity Flag (UK90 thresholds)	1238	
Overweight	558 (45.07%)	244 (43.49%)
Obese	680 (54.93%)	317 (56.51%)
Ethnic	1230	
Other Ethnic group	21 (1.71%)	12 (2.15%)
Black or Black British	52 (4.23%)	22 (3.94%)
Pakistani and Bangladeshi	75 (6.10%)	25 (4.48%)
Indian	27 (2.20%)	15 (2.69%)
Mixed	40 (3.25%)	24 (4.30%)
White	1015 (82.52%)	460 (82.44%)
Parents' mental health (K6)	1179	4.30 [1, 6]
Highest Parental Educational Attainment	1092	
OECD-equivalized weekly household income (£)	1237	415.00 ± 173.63

Note. Continuous variables are presented as mean ± standard deviation (SD) when approximately normally distributed, median [interquartile range] when non-normally distributed, and categorical variables as n (%). BMI = body mass index; CGT = Cambridge Gambling Task; MVPA = moderate-to-vigorous physical activity; MCS = Millennium Cohort Study.

Table 2
Estimated average treatment effects for different exposure-outcome combinations.

Exposure Variable	Outcome Variable	ATE	95% CI	Proportion of Sample with Significant Effect
MVPA_weekday	Internalizing problems at age 17 (parent-reported SDQ composite)	-0.0025	[-0.0062, 0.0012]	22.98%
MVPA_weekday	Externalizing problems at age 17 (parent-reported SDQ composite)	0.0003	[-0.0027, 0.0033]	0.00%
MVPA_weekend	Internalizing problems at age 17 (parent-reported SDQ composite)	-0.0008	[-0.0051, 0.0035]	5.65%
MVPA_weekend	Externalizing problems at age 17 (parent-reported SDQ composite)	0.0005	[-0.0026, 0.0037]	1.61%

internalizing points per 30 minutes, with uncertainty spanning zero. It is noteworthy that only for the effect of weekday MVPA on internalizing problems did 22.98% of individuals show a statistically significant effect estimate, suggesting meaningful heterogeneity (Fig. 1). Uncertainty in individual-level estimates is illustrated for a subset of participants in Fig. 2.

Heterogeneous effects and moderators

We analyzed the estimated individual conditional average treatment effects (CATE) to investigate the sources of heterogeneity in treatment effects. Among all the covariates examined, body height was identified as the most important factor explaining the variation in effects for the "weekday MVPA to internal psychological state" pathway (Fig. 3). Scatter plots visually indicated a clear linear trend between height and CATE (Fig. 4). We therefore constructed a linear regression model with CATE as the dependent variable and height as the independent variable. The results (Table 3) showed that height had a significant positive predictive effect on CATE ($\beta = 0.0002$, SE = 0.0000, $t = 16.307$, $p < 0.001$). The model demonstrated a good overall fit ($R^2 = 0.5194$, Adjusted $R^2 = 0.5175$, $F(1, 246) = 265.90$, $p < 0.001$), indicating that height could explain 51.94% of the variation in individual treatment effects and is a key explanatory factor for effect heterogeneity. This gradient is also evident when stratifying height into quartiles (Fig. 5)

Discussion

In this longitudinal cohort of UK adolescents with overweight/obesity, we found no clear evidence that either weekday or weekend

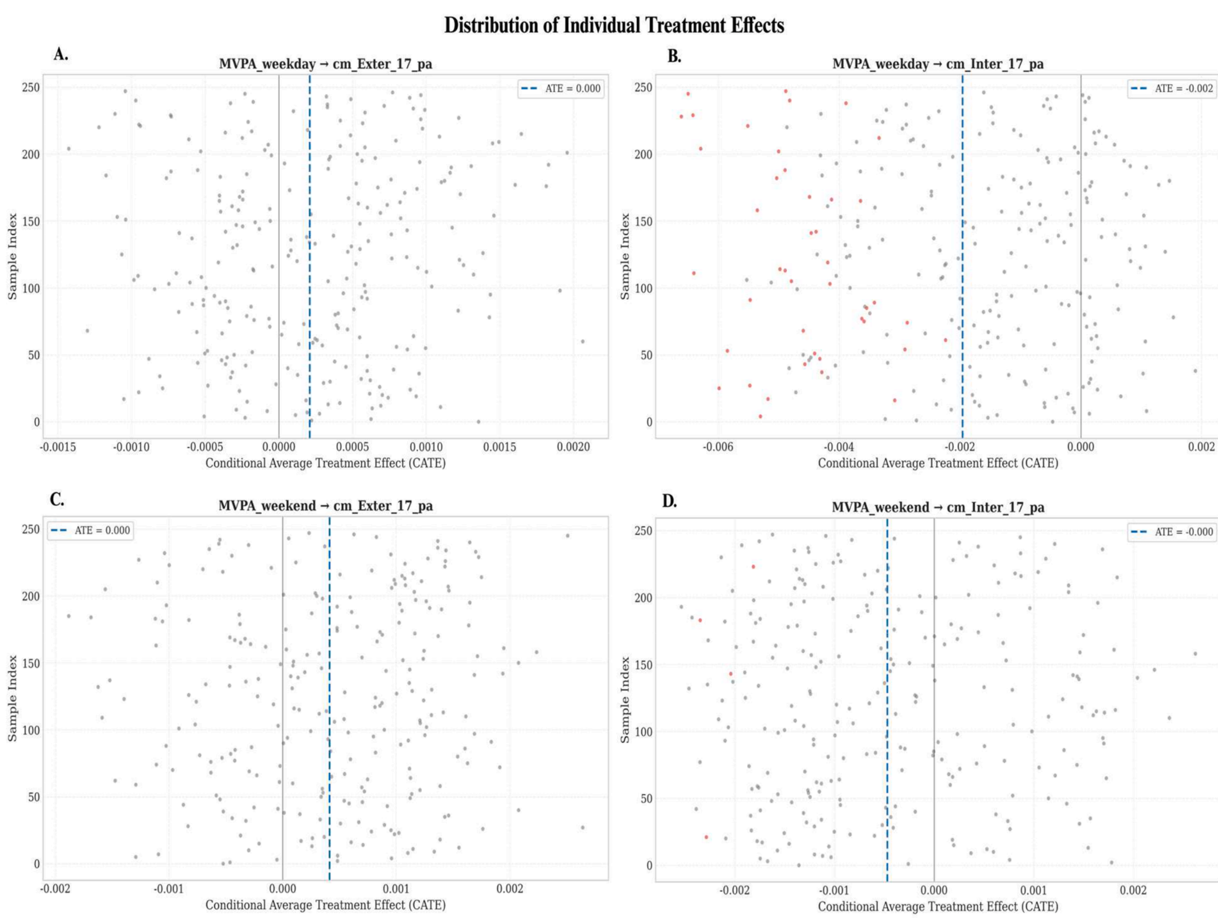

Fig. 1. Distribution of individual conditional average treatment effects (CATEs) of MVPA on adolescent mental health outcomes. Each dot represents one participant's estimated CATE from the causal forest, plotted against an arbitrary participant index (y-axis; shown only to separate points visually). The x-axis shows the CATE, interpreted as the estimated change in SDQ score at age 17 associated with a 1-minute increase in MVPA at age 14 (negative values indicate lower symptom scores). The blue dashed vertical line marks the sample average treatment effect (ATE). Red dots indicate individuals whose estimated CATE differs from zero at $p < 0.05$; gray dots indicate non-significant individual effects. Panels show effects for (A) weekday MVPA → externalizing, (B) weekday MVPA → internalizing, (C) weekend MVPA → externalizing, and (D) weekend MVPA → internalizing.

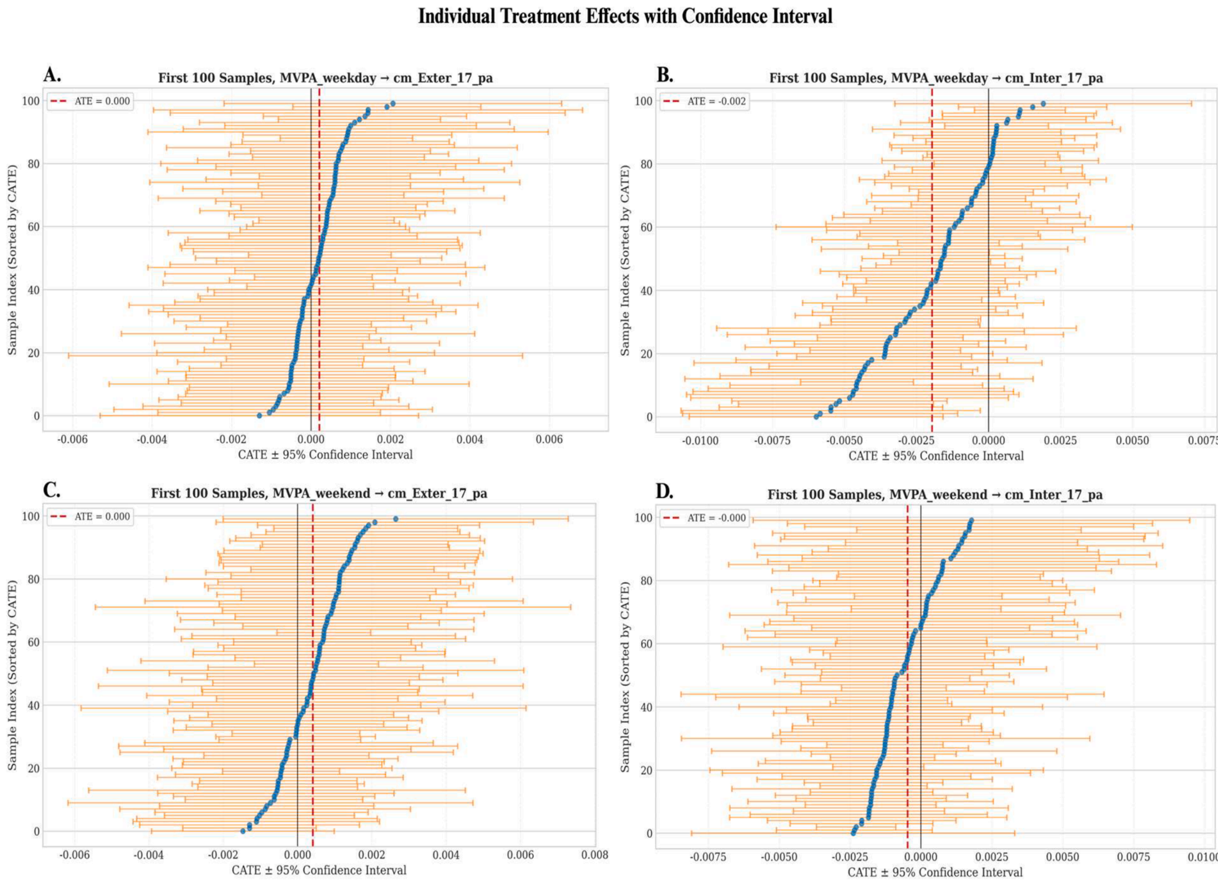

Fig. 2. Individual conditional average treatment effects (CATEs) with 95% confidence intervals for the first 100 participants. CATEs were estimated using the causal forest and represent the estimated change in SDQ score at age 17 associated with a 1-minute increase in MVPA at age 14 (negative values indicate lower symptom scores). Each horizontal bar shows the 95% confidence interval for an individual participant's CATE, and the blue curve/points indicate the corresponding point estimates. The vertical red line denotes zero effect, and the shaded band indicates the 95% confidence interval around the sample average treatment effect (ATE). Panels show effects for (A) weekday MVPA → externalizing, (B) weekday MVPA → internalizing, (C) weekend MVPA → externalizing, and (D) weekend MVPA → internalizing. Participant order is for display only.

MVPA at age ∼14 had a detectable *average* causal effect on internalizing or externalizing problems at age ∼17. Across all four exposure–outcome combinations, estimated ATEs were small and statistically non-significant with confidence intervals spanning zero. However, meaningful heterogeneity emerged for the weekday MVPA to internalizing pathway, suggesting that the potential psychosocial benefits of MVPA may be concentrated in specific subgroups rather than distributed uniformly across the population. Thus, the absence of an average effect did not imply that MVPA was irrelevant for all individuals.

Interpretation of average effects and timing differences

A key finding was the contrast between the null average effects and the presence of heterogeneity specifically for weekday MVPA and internalizing problems. This pattern is consistent with the possibility that context, rather than MVPA duration alone, shapes whether MVPA is psychologically beneficial for adolescents living with overweight/obesity, although residual confounding cannot be ruled out (Lubans et al., 2016). According to the structured day hypothesis, weekday MVPA is more likely to occur within structured environments such as school physical education, active commuting, or organized extracurricular activities (Brazendale et al., 2021). These settings may provide predictable routines, social contact, and adult supervision, which can support stress regulation and social belonging; however, they may also reflect broader contextual factors that influence both opportunities for MVPA and mental health, such as school PE policies and extracurricular provision, school climate, peer network characteristics, and neighborhood safety/access to facilities. In contrast, weekend MVPA may be more dependent on family resources (McMinn et al., 2013), neighborhood conditions (Stone et al., 2014), and peer inclusion (Sebire et al., 2018), leading to greater variability in the type, quality, and psychosocial "meaning" of the activity (Ortega et al., 2022). Greater heterogeneity in weekend contexts may contribute to less stable downstream associations and may partially explain why weekend MVPA effects were near zero on average.

Measurement considerations likely also contributed to the weak average estimates, particularly for weekend MVPA. MVPA was assessed using one valid weekday and one valid weekend day which we intentionally treated as context-specific snapshots rather than an indicator of regular behavior (Ma et al., 2024). Single-day measures can introduce substantial random error, and this error is plausibly larger on weekends when schedules and opportunities vary more (Sampson et al., 2016). Such measurement noise would be expected to attenuate effects toward the null (Bolger & Laurenceau, 2013). This limitation becomes even more consequential given the multi-year interval between exposure (∼14 years) and outcomes (∼17 years), during which physical activity patterns, school environments, peer networks, and developmental processes can change markedly.

Direction and magnitude of effects

Interpreting the direction of effects requires attention to outcome scaling. Internalizing and externalizing composites were derived from SDQ domains where higher scores reflect more severe problems.

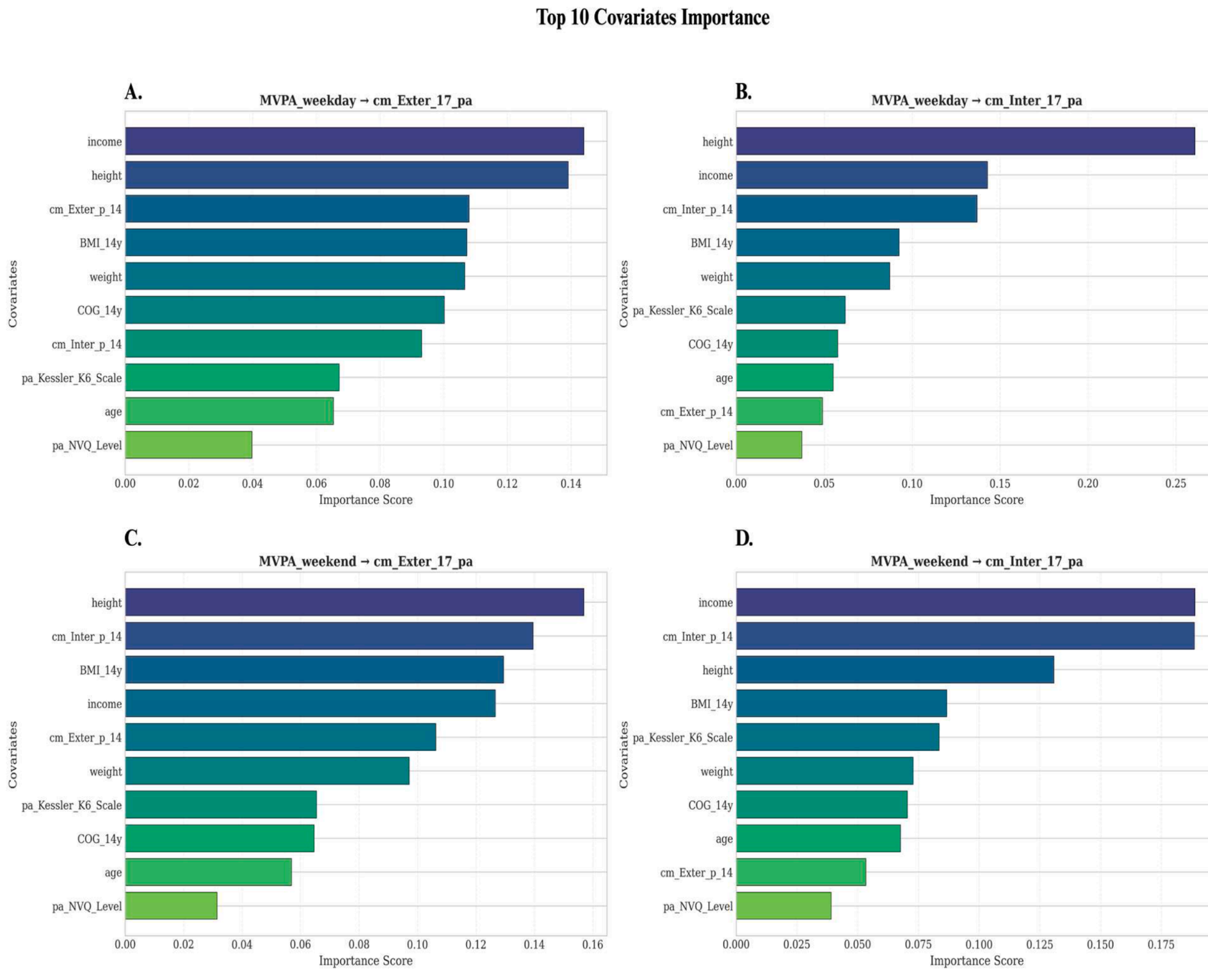

Fig. 3. Covariate importance for predicting heterogeneity in MVPA effects (CATEs). Bars show the top 10 covariates ranked by their contribution to predicting individual conditional average treatment effects (CATEs) from the causal forest, displayed separately for each exposure–outcome model: (A) weekday MVPA → externalizing, (B) weekday MVPA → internalizing, (C) weekend MVPA → externalizing, and (D) weekend MVPA → internalizing. Higher importance scores indicate covariates that were more influential in explaining variation in estimated treatment effects across individuals; importance reflects predictive contribution within the fitted forest and does not imply causal moderation on its own.

According to this scoring method, a negative MVPA effect indicates fewer internalizing symptoms (i.e., a beneficial association). Even where the point estimate suggested a benefit, the average magnitude was small and uncertain at the population level. This reinforces a practical message: increasing MVPA may not reliably translate into broad, average improvements in parent-reported psychosocial functioning among adolescents with overweight/obesity, because the psychological impact of activity likely depends on how, where, and with whom it occurs. For some adolescents, MVPA may take place in supportive settings that promote social connection and stress regulation; for others, activity contexts may involve social comparison, weight-related stigma, performance expectations, or peer pressure, which could blunt benefits or introduce new stressors. Still, regular MVPA remains important for physical health and may confer mental health benefits in specific subgroups or contexts (Jacob et al., 2021).

Why height may be associated with heterogeneity in weekday MVPA effects

Height was the strongest variable associated with heterogeneity in estimated weekday MVPA effects on internalizing problems, with estimated effects becoming less negative as height increased (i.e., smaller apparent benefit among taller adolescents). However, height should not be interpreted as a direct causal determinant of this heterogeneity. Rather, it may serve as a marker of correlated developmental and social processes (Schick & Steckel, 2010). At age ~14, height may partly reflect differences in pubertal timing and biological maturation, which in turn may relate to physical competence (McKinnon & Kolen, 2018), fatigue (Liu et al., 2025), sleep regulation (El Halal & Nunes, 2019), and perceived athletic ability (Beunen & Malina, 2008). The implications of pubertal timing may also differ by sex: earlier maturation may confer physical advantages and social status that support sport engagement among boys, whereas in girls it may be linked to reduced sport participation because of body image concerns, increased self-consciousness, and greater social evaluation in mixed-gender or performance-oriented settings. In this sense, height may be capturing broader developmental and social-contextual differences that shape how weekday MVPA is experienced, rather than exerting a direct effect on internalizing symptoms itself.

The pattern observed among shorter adolescents—namely, more negative estimated MVPA effects—may indicate that some subgroups have greater scope to benefit from supportive weekday activity contexts (Bedard et al., 2022; Sallis et al., 2000). One plausible explanation is that shorter or later-maturing adolescents enter school-based activity settings with lower perceived physical competence or greater psychosocial vulnerability, such that increases in MVPA are more strongly associated with gains in mastery, self-efficacy, or social belonging relevant to internalizing symptoms (Babic et al., 2014; Rebelo et al., 2025). Equally,

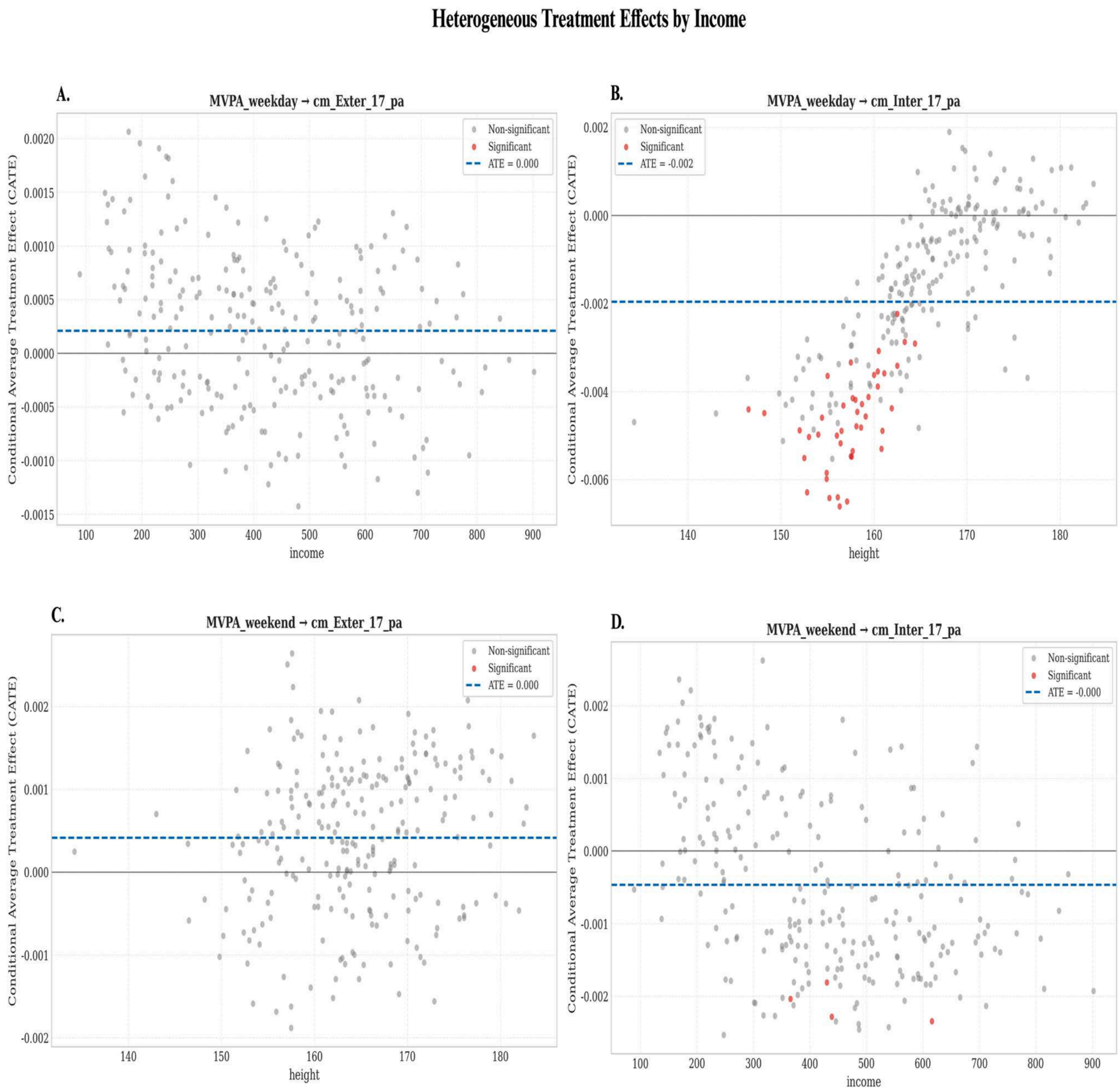

Fig. 4. Heterogeneous treatment effects of MVPA on mental health outcomes by body height. Scatterplots display individual conditional average treatment effects (CATEs) from the causal forest against height (cm). The y-axis shows the CATE, interpreted as the estimated change in SDQ score at age 17 associated with a 1-minute increase in MVPA at age 14 (negative values indicate lower symptom scores). Red points indicate individuals whose estimated CATE differs from zero at $p < 0.05$; gray points indicate non-significant individual effects. The blue dashed horizontal line marks the sample average treatment effect (ATE). Panels show effects for (A) weekday MVPA → externalizing, (B) weekday MVPA → internalizing, (C) weekend MVPA → externalizing, and (D) weekend MVPA → internalizing.

Table 3
Linear regression results of conditional average treatment effect on height.

Variable	Coefficient	Standard Error	t-value	p-value
Intercept	-0.0280	0.0015	-18.364	<0.001
height	0.0002	0.0000	16.307	<0.001

height may simply be acting as a proxy for unmeasured characteristics such as pubertal stage, physical self-concept, sport participation opportunities, or peer dynamics (Sherar et al., 2010; Stodden et al., 2008). Accordingly, the height-related pattern should be interpreted as hypothesis-generating evidence of subgroup variation, not as evidence that height itself causally modifies the psychological effects of MVPA or as a basis for targeting interventions by stature alone.

Methodological considerations

A major strength of this study is the use of causal forests to estimate heterogeneous effects while adjusting for a broad set of pre-exposure covariates. This framework is well-suited to uncovering exposure effect heterogeneity that would be masked by models focused solely on population averages. Nevertheless, inference remains contingent on key identification assumptions, particularly exchangeability (no unmeasured confounding) and positivity (Heckman et al., 2025; Imbens & Rubin, 2015). Physical activity is embedded in complex social and environmental systems that may not be fully measured, including parenting practices, school policies, peer networks, and neighbourhood resources (Wang et al., 2023). Therefore, residual confounding therefore remains possible. Relatedly, differential adherence and compliance may also bias estimates. Adolescents who feel self-conscious about weight (Castonguay et al., 2015), experience weight-related stigma (Pearl et al., 2021), or face peer pressure (Salvy et al., 2008) may be less likely to participate fully in activity opportunities (or may avoid structured settings such as PE), which could affect both measured MVPA and mental health outcomes. In addition, non-random accelerometer wear compliance (e.g., lower wear time among adolescents who are less comfortable wearing the device (Wang et al., 2025) could contribute to measurement error and selection, further limiting causal interpretation.

In addition, our explanatory analysis linking estimated CATEs to

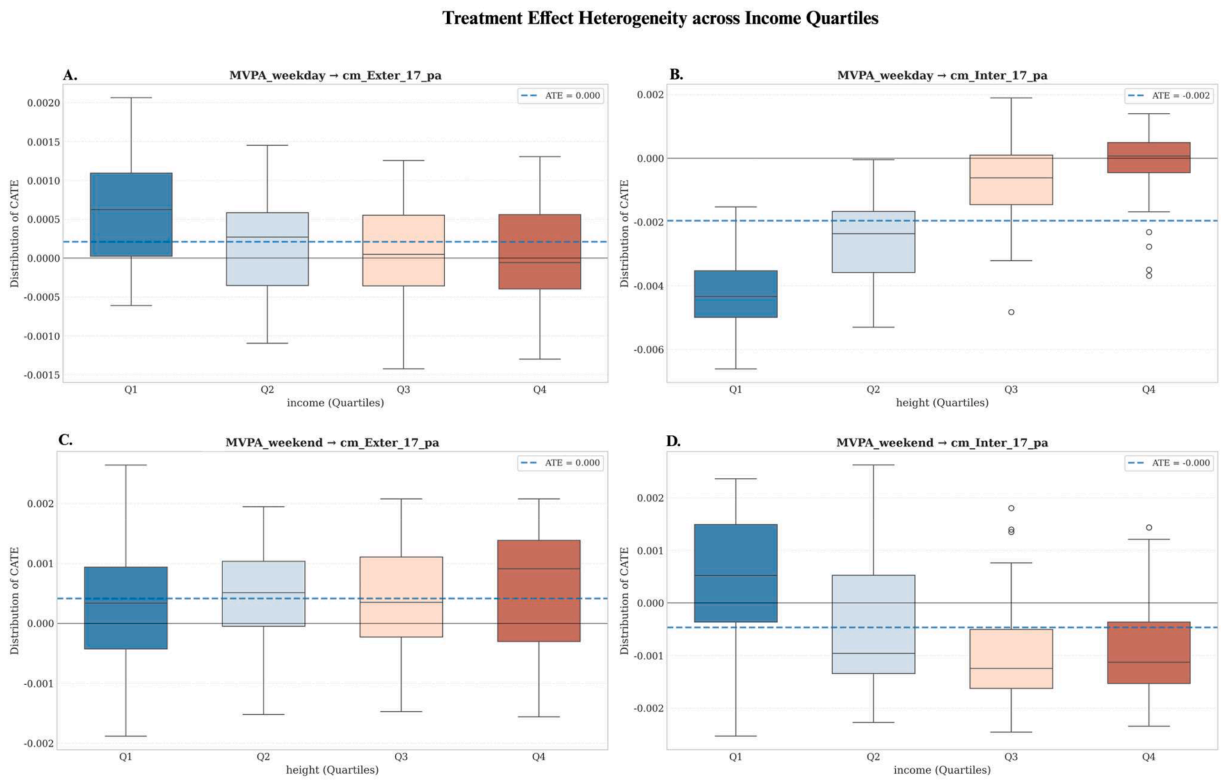

Fig. 5. Treatment effect heterogeneity across body height quartiles. Boxplots summarize the distribution of individual conditional average treatment effects (CATEs) within height quartiles (Q1 = shortest; Q4 = tallest). CATEs are interpreted as the estimated change in SDQ score at age 17 associated with a 1-minute increase in MVPA at age 14 (negative values indicate lower symptom scores). The central line in each box denotes the median, boxes indicate the interquartile range, whiskers show the spread beyond the quartiles, and points denote outliers. The blue dashed horizontal line marks the sample average treatment effect (ATE). Panels show effects for (A) weekday MVPA → externalizing, (B) weekday MVPA → internalizing, (C) weekend MVPA → externalizing, and (D) weekend MVPA → internalizing.

height improves interpretability but should be viewed cautiously because CATEs are themselves estimated quantities with uncertainty. Replication using approaches that more fully propagate first-stage uncertainty—and ideally using repeated measures of exposure and richer contextual information—would strengthen confidence in the specific moderator pattern observed.

Practical implications

Two practical implications follow. First, the absence of clear average effects suggests that “one-size-fits-all” MVPA recommendations may be less than optimal for adolescents with overweight/obesity, even though promoting MVPA remains a key element for fostering physical and mental health (Wang et al., 2024). Second, the weekday-specific heterogeneity suggests that any psychosocial benefits may depend not only on the amount of MVPA but also on the context in which it occurs. Interventions embedded in weekday routines—such as school physical education, active commuting initiatives, and after-school programs—may therefore be especially promising if they prioritize psychological safety, reduce weight-based stigma, and support perceived competence, enjoyment, and social inclusion.

Research implications and future directions

Several research priorities emerge. Future studies should assess physical activity across multiple weekdays and weekend days to obtain more reliable estimates of MVPA and should collect richer contextual information on the type of activity (e.g., organized sport, physical education, active transport, informal play), the social setting (e.g., alone, with peers, with family, adult-supervised), and the subjective experience of participation (e.g., perceived enjoyment, competence, inclusion, and weight-related stigma). Studies should also include direct measures of pubertal stage, physical self-concept, and school- or neighborhood-level opportunity structures to clarify why height-related heterogeneity emerged in the present analyses. In addition, future work should examine potential non-linear dose-response relationships and outcomes that are more proximal to activity, such as short-term mood, stress biomarkers (e.g., hair cortisol concentration), and sleep. Finally, quasi-experimental and intervention studies that deliberately vary weekday MVPA opportunities and contexts would provide a stronger basis for evaluating whether weekday MVPA has distinct effects on internalizing problems in adolescents with overweight/obesity.

Limitations

This study has several limitations. First, weekday and weekend MVPA were measured on a single day, respectively. This likely introduces substantial measurement error and attenuates average effects (Heckman et al., 2026), particularly for weekend activity. Second, although we adjusted for a broad set of pre-exposure covariates, residual confounding remains possible (e.g., pubertal timing, physical self-concept, peer experiences, school policies, neighborhood environment). Third, outcomes were parent-reported SDQ scores which may not adequately capture adolescents’ internal experiences across contexts and may vary in accuracy across developmental stages. Fourth, causal-forest estimates of heterogeneity are inherently model-dependent (Credit & Lehnert, 2024); exploratory moderator patterns should be interpreted as hypothesis-forming and require replication. Finally, missing-data handling and attrition may affect generalizability (Wang et al., 2025); sensitivity analyses using alternative missingness strategies and (where appropriate) survey weights are important to evaluate robustness. In addition, because the Millennium Cohort Study is

UK-based, these findings should be interpreted in the context of British adolescents and policy, school, and cultural environments; associations may differ in other countries where social norms, activity opportunities, and weight-related stigma and supports differ.

Conclusions

In this sample of UK adolescents with overweight/obesity, weekday and weekend MVPA at age 14 showed no clear average association with parent-reported internalizing or externalizing problems at age 17 after covariate adjustment. However, weekday MVPA showed meaningful heterogeneity for internalizing problems, suggesting that potential psychosocial benefits may be concentrated in specific subgroups rather than shared uniformly across adolescents. These findings indicate that increasing MVPA alone may be insufficient as a broad mental health strategy and that the psychosocial impact of activity likely depends on individual and contextual factors. Future studies should replicate these patterns using multi-day accelerometry and richer measures of activity context and development.

Declaration of competing interest

There are no conflicts of interest.

References

Ames, M., & Leadbeater, B. (2017). Overweight and isolated: The interpersonal problems of youth who are overweight from adolescence into young adulthood. *International Journal of Behavioral Development, 41*(3), 390–404.

Antczak, D., Lonsdale, C., del Pozo Cruz, B., Parker, P., & Sanders, T. (2021). Reliability of GENEActiv accelerometers to estimate sleep, physical activity, and sedentary time in children. *International Journal of Behavioral Nutrition and Physical Activity, 18*(1), 73. https://doi.org/10.1186/s12966-021-01143-6

Athey, S., & Imbens, G. W. (2019). Machine Learning Methods That Economists Should Know About. *Annual Review of Economics, 11*, 685–725. https://doi.org/10.1146/annurev-economics-080217-053433, 11, 2019.

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *The Annals of Statistics, 47*(2), 1148–1178. https://doi.org/10.1214/18-AOS1709

Babic, M. J., Morgan, P. J., Plotnikoff, R. C., Lonsdale, C., White, R. L., & Lubans, D. R. (2014). Physical activity and physical self-concept in youth: Systematic review and meta-analysis. *Sports Medicine, 44*(11), 1589–1601. https://doi.org/10.1007/s40279-014-0229-z

Bakrania, K., Yates, T., Rowlands, A. V., Esliger, D. W., Bunnewell, S., Sanders, J., Davies, M., Khunti, K., & Edwardson, C. L. (2016). Intensity Thresholds on Raw Acceleration Data: Euclidean Norm Minus One (ENMO) and Mean Amplitude Deviation (MAD) Approaches. *PLOS ONE, 11*(10), Article e0164045. https://doi.org/10.1371/journal.pone.0164045

Bauman, A. E., Reis, R. S., Sallis, J. F., Wells, J. C., Loos, R. J., & Martin, B. W. (2012). Correlates of physical activity: Why are some people physically active and others not? *The Lancet, 380*(9838), 258–271. https://doi.org/10.1016/S0140-6736(12)60735-1

Bedard, C., King-Dowling, S., Obeid, J., Timmons, B. W., & Ferro, M. A. (2022). Correlates of Moderate-to-Vigorous Physical Activity in Children With Physical Illness and Physical–Mental Multimorbidity. *Health Education & Behavior, 49*(5), 780–788. https://doi.org/10.1177/10901981221100697

Beretta, L., & Santaniello, A. (2016). Nearest neighbor imputation algorithms: A critical evaluation. *BMC Medical Informatics and Decision Making, 16*(3), 74. https://doi.org/10.1186/s12911-016-0318-z

Beunen, G., & Malina, R. M. (2008). Growth and biologic maturation: Relevance to athletic performance. *The Young Athlete, 1*, 3–17.

Biddle, S. J., & Asare, M. (2011). Physical activity and mental health in children and adolescents: A review of reviews. *British Journal of Sports Medicine, 45*(11), 886–895.

Biddle, S. J. H., Ciaccioni, S., Thomas, G., & Vergeer, I. (2019). Physical activity and mental health in children and adolescents: An updated review of reviews and an analysis of causality. *Psychology of Sport and Exercise, 50 Years of FEPSAC: Current and Future Directions to Sport and Exercise Psychology Research, 42*, 146–155. https://doi.org/10.1016/j.psychsport.2018.08.011

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research.* Guilford press.

Brady, R., Brown, W. J., & Mielke, G. I. (2023). Day-to-day variability in accelerometer-measured physical activity in mid-aged Australian adults. *BMC Public Health, 23*(1), 1880. https://doi.org/10.1186/s12889-023-16734-0

Brazendale, K., Beets, M. W., Armstrong, B., Weaver, R. G., Hunt, E. T., Pate, R. R., Brusseau, T. A., Bohnert, A. M., Olds, T., Tassitano, R. M., Tenorio, M. C. M., Garcia, J., Andersen, L. B., Davey, R., Hallal, P. C., Jago, R., Kolle, E., Kriemler, S., Kristensen, P. L., … van Sluijs, E. M. F. (2021). Children's moderate-to-vigorous physical activity on weekdays versus weekend days: A multi-country analysis. *The International Journal of Behavioral Nutrition and Physical Activity, 18*, 28. https://doi.org/10.1186/s12966-021-01095-x

Brazendale, K., Beets, M. W., Weaver, R. G., Pate, R. R., Turner-McGrievy, G. M., Kaczynski, A. T., Chandler, J. L., Bohnert, A., & von Hippel, P. T. (2017). Understanding differences between summer vs. school obesogenic behaviors of children: The structured days hypothesis. *International Journal of Behavioral Nutrition and Physical Activity, 14*(1), 100. https://doi.org/10.1186/s12966-017-0555-2

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., … Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine, 54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

Castonguay, A. L., Pila, E., Wrosch, C., & Sabiston, C. M. (2015). Body-related self-conscious emotions relate to physical activity motivation and behavior in men. *American Journal of Men's Health, 9*(3), 209–221.

Chen, M., Landré, B., Marques-Vidal, P., Hees, V. T.van, Gennip, A. C. E.van, Bloomberg, M., Yerramalla, M. S., Benadjaoud, M. A., & Sabia, S. (2023). Identification of physical activity and sedentary behaviour dimensions that predict mortality risk in older adults: Development of a machine learning model in the Whitehall II accelerometer sub-study and external validation in the CoLaus study. *eClinicalMedicine*, 55. https://doi.org/10.1016/j.eclinm.2022.101773

Credit, K., & Lehnert, M. (2024). A structured comparison of causal machine learning methods to assess heterogeneous treatment effects in spatial data. *Journal of Geographical Systems, 26*(4), 483–510. https://doi.org/10.1007/s10109-023-00413-0

El Halal, C., dos, S., & Nunes, M. L. (2019). Sleep and weight-height development. *Jornal de Pediatria, 95*(Suppl 1), S2–S9.

Fairclough, S. J., Boddy, L. M., Mackintosh, K. A., Valencia-Peris, A., & Ramirez-Rico, E. (2015). Weekday and weekend sedentary time and physical activity in differentially active children. *Journal of Science and Medicine in Sport, 18*(4), 444–449. https://doi.org/10.1016/j.jsams.2014.06.005

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Gordon, J. A., Dzirasa, K., & Petzschner, F. H. (2024). The neuroscience of mental illness: Building toward the future. *Cell, 187*(21), 5858–5870. https://doi.org/10.1016/j.cell.2024.09.028

Gouveia, M. J., Frontini, R., Canavarro, M. C., & Moreira, H. (2014). Quality of life and psychological functioning in pediatric obesity: The role of body image dissatisfaction between girls and boys of different ages. *Quality of Life Research, 23*(9), 2629–2638.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature, 585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Heckman, J. J., Tian, H., Zhang, Z., & Zhou, J. (2026). *Measuring the Growth of Skills (Working Paper 34737)*. National Bureau of Economic Research. https://doi.org/10.3386/w34737

Heckman, J., Tian, H., Zhang, Z., & Zhou, J. (2025). *Dynamic Complementarity.* ROCKWOOL Foundation Berlin. RFBerlin.

Hope, S., Rougeaux, E., Deighton, J., Law, C., & Pearce, A. (2019). Associations between mental health competence and indicators of physical health and cognitive development in eleven year olds: Findings from the UK Millennium Cohort Study. *BMC Public Health, 19*(1), 1461. https://doi.org/10.1186/s12889-019-7789-7

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference for statistics, social, and biomedical sciences: An introduction (pp. xix, 625)*. Cambridge University Press. https://doi.org/10.1017/CBO9781139025751

Jacob, C. M., Hardy-Johnson, P. L., Inskip, H. M., Morris, T., Parsons, C. M., Barrett, M., Hanson, M., Woods-Townsend, K., & Baird, J. (2021). A systematic review and meta-analysis of school-based interventions with health education to reduce body mass index in adolescents aged 10 to 19 years. *The International Journal of Behavioral Nutrition and Physical Activity, 18*, 1. https://doi.org/10.1186/s12966-020-01065-9

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L. T., Walters, E. E., & Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine, 32*(6), 959–976. https://doi.org/10.1017/S0033291702006074

Kravitz, R. L., Duan, N., & Braslow, J. (2004). Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *The Milbank Quarterly, 82*(4), 661–687. https://doi.org/10.1111/j.0887-378X.2004.00327.x

Liu, C., Liu, Y., Tian, Y., Pang, L., & Liu, J. (2025). How drop height- and fatigue-dominated overtraining affect performance and its underlying mechanisms. *Intelligent Sports and Health, 1*(4), 194–200. https://doi.org/10.1016/j.ish.2025.08.001

Lubans, D., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., Kelly, P., Smith, J., Raine, L., & Biddle, S. (2016). Physical Activity for Cognitive and Mental Health in Youth: A Systematic Review of Mechanisms. *Pediatrics, 138*(3), Article e20161642. https://doi.org/10.1542/peds.2016-1642

Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions* (arXiv:1705.07874). arXiv. https://doi.org/10.48550/arXiv.1705.07874.

Ma, T., Sirard, J., Yang, L., Li, Y., Tsang, S., & Fu, A. (2024). Revisiting the concept of bout: Associations of moderate-to-vigorous physical activity sessions and non-sessions with mortality. *The International Journal of Behavioral Nutrition and Physical Activity, 21*, 81. https://doi.org/10.1186/s12966-024-01631-5

McKinney, W. (2010). Data Structures for Statistical Computing in Python. *SciPy, 2010*. https://doi.org/10.25080/Majora-92bf1922-00a
McKinnon, M. F., & Kolen, A. M. (2018). The Impact of Growth in Height on the Physical Competencies of Children. *Journal of Physical Activity Research, 3*(1), 6–10. https://doi.org/10.12691/jpar-3-1-2
McMinn, A. M., Griffin, S. J., Jones, A. P., & van Sluijs, E. M. F. (2013). Family and home influences on children's after-school and weekend physical activity. *The European Journal of Public Health, 23*(5), 805–810. https://doi.org/10.1093/eurpub/cks160
Mellin, A. E., Neumark-Sztainer, D., Story, M., Ireland, M., & Resnick, M. D. (2002). Unhealthy behaviors and psychosocial difficulties among overweight adolescents: The potential impact of familial factors. *Journal of Adolescent Health, 31*(2), 145–153.
Migueles, J. H., Rowlands, A. V., Huber, F., Sabia, S., & Hees, V. T.van (2019). GGIR: A Research Community–Driven Open Source R Package for Generating Physical Activity and Sleep Outcomes From Multi-Day Raw Accelerometer Data. *Journal for the Measurement of Physical Behaviour*. https://journals.humankinetics.com/view/journals/jmpb/2/3/article-p188.xml.
Mumford, E. A., Liu, W., Hair, E. C., & Yu, T.-C. (2013). Concurrent trajectories of BMI and mental health patterns in emerging adulthood. *Social Science & Medicine, 98*, 1–7.
NHS England. (2023). *National Child Measurement Programme, England, 2022/23 School Year*. October 9. NHS England Digital https://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme/2022-23-school-year.
Ortega, A., Bejarano, C. M., Cushing, C. C., Staggs, V. S., Papa, A. E., Steel, C., Shook, R. P., Conway, T. L., Saelens, B. E., Glanz, K., Cain, K. L., Frank, L. D., Kerr, J., Schipperijn, J., Sallis, J. F., & Carlson, J. A. (2022). Location-specific psychosocial and environmental correlates of physical activity and sedentary time in young adolescents: Preliminary evidence for location-specific approaches from a cross-sectional observational study. *The International Journal of Behavioral Nutrition and Physical Activity, 19*, 108. https://doi.org/10.1186/s12966-022-01336-7
Pearl, R. L., Wadden, T. A., & Jakicic, J. M. (2021). Is weight stigma associated with physical activity? A systematic review. *Obesity, 29*(12), 1994–2012.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*(85), 2825–2830.
Phelps, N. H., Singleton, R. K., Zhou, B., Heap, R. A., Mishra, A., Bennett, J. E., Paciorek, C. J., Lhoste, V. P., Carrillo-Larco, R. M., Stevens, G. A., Rodriguez-Martinez, A., Bixby, H., Bentham, J., Cesare, M. D., Danaei, G., Rayner, A. W., Barradas-Pires, A., Cowan, M. J., Savin, S., … Ezzati, M. (2024). Worldwide trends in underweight and obesity from 1990 to 2022: A pooled analysis of 3663 population-representative studies with 222 million children, adolescents, and adults. *The Lancet, 403*(10431), 1027–1050. https://doi.org/10.1016/S0140-6736(23)02750-2
Puhl, R. M., & Heuer, C. A. (2009). The stigma of obesity: A review and update. *Obesity, 17*(5), 941–964. https://doi.org/10.1038/oby.2008.636
Puhl, R. M., Himmelstein, M. S., & Pearl, R. L. (2020). Weight stigma as a psychosocial contributor to obesity. *American Psychologist, 75*(2), 274.
Python Software Foundation. (2025). *Python 3.11.14 [Computer software]*. https://www.python.org/downloads/release/python-31114/.
Rebelo, M., Serrano, J., Honório, S., Santos, J., Marques, C., & Batista, M. (2025). Associations Between Body Mass Index, Physical Activity, Perceived School Competence, and Academic Performance in Portuguese Elementary Students. *Children, 12*(12), 1601. https://doi.org/10.3390/children12121601
Rodriguez-Ayllon, M., Cadenas-Sánchez, C., Estévez-López, F., Muñoz, N. E., Mora-Gonzalez, J., Migueles, J. H., Molina-García, P., Henriksson, H., Mena-Molina, A., Martínez-Vizcaíno, V., Catena, A., Löf, M., Erickson, K. I., Lubans, D. R., Ortega, F. B., & Esteban-Cornejo, I. (2019). Role of Physical Activity and Sedentary Behavior in the Mental Health of Preschoolers, Children and Adolescents: A Systematic Review and Meta-Analysis. *Sports Medicine, 49*(9), 1383–1410. https://doi.org/10.1007/s40279-019-01099-5
Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents. *Medicine and Science in Sports and Exercise, 32*(5), 963–975. https://doi.org/10.1097/00005768-200005000-00014
Salvy, S.-J., Bowker, J. W., Roemmich, J. N., Romero, N., Kieffer, E., Paluch, R., & Epstein, L. H. (2008). Peer influence on children's physical activity: An experience sampling study. *Journal of Pediatric Psychology, 33*(1), 39–49.
Sampson, J. N., Matthews, C. E., Freedman, L., Carroll, R. J., & Kipnis, V. (2016). Methods to Assess Measurement Error in Questionnaires of Sedentary Behavior. *Journal of Applied Statistics, 43*(9), 1706–1721. https://doi.org/10.1080/02664763.2015.1117593
Schick, A., & Steckel, R. (2010). *Height as a Proxy for Cognitive and Non-Cognitive Ability (W16570*. National Bureau of Economic Research, Article w16570. https://doi.org/10.3386/w16570
Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *SciPy, 2010*. https://doi.org/10.25080/Majora-92bf1922-011
Sebire, S. J., Jago, R., Banfield, K., Edwards, M. J., Campbell, R., Kipping, R., Blair, P. S., Kadir, B., Garfield, K., Matthews, J., Lyons, R. A., & Hollingworth, W. (2018). Results of a feasibility cluster randomised controlled trial of a peer-led school-based intervention to increase the physical activity of adolescent girls (PLAN-A). *The International Journal of Behavioral Nutrition and Physical Activity, 15*, 50. https://doi.org/10.1186/s12966-018-0682-4
Sherar, L. B., Cumming, S. P., Eisenmann, J. C., Baxter-Jones, A. D. G., & Malina, R. M. (2010). Adolescent biological maturity and physical activity: Biology meets behavior. *Pediatric Exercise Science, 22*(3), 332–349. https://doi.org/10.1123/pes.22.3.332
Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A Developmental Perspective on the Role of Motor Skill Competence in Physical Activity: An Emergent Relationship. *Quest, 60*(2), 290–306. https://doi.org/10.1080/00336297.2008.10483582
Stone, M. R., Faulkner, G. E., Mitra, R., & Buliung, R. N. (2014). The freedom to explore: Examining the influence of independent mobility on weekday, weekend and after-school physical activity behaviour in children living in urban and inner-suburban neighbourhoods of varying socioeconomic status. *The International Journal of Behavioral Nutrition and Physical Activity, 11*, 5. https://doi.org/10.1186/1479-5868-11-5
Terhaag, S., Fitzsimons, E., Daraganova, G., & Patalay, P. (2021). Sex, ethnic and socioeconomic inequalities and trajectories in child and adolescent mental health in Australia and the UK: Findings from national prospective longitudinal studies. *Journal of Child Psychology and Psychiatry, 62*(10), 1255–1267. https://doi.org/10.1111/jcpp.13410
Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., & Altman, R. B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics, 17*(6), 520–525. https://doi.org/10.1093/bioinformatics/17.6.520
University College London, U. I. O. E.. (2024). *Millennium Cohort Study (Version 16th Release)*. UK Data Service. https://doi.org/10.5255/UKDA-SERIES-2000031
Wager, S., & Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association, 113*(523), 1228–1242. https://doi.org/10.1080/01621459.2017.1319839
Wang, P., Liu, A., & Sun, X. (2025). Integrating emotion dynamics in mental health: A trimodal framework combining ecological momentary assessment, physiological measurements, and speech emotion recognition. *Interdisciplinary Medicine, 3*(3), Article e20240095. https://doi.org/10.1002/INMD.20240095
Wang, P., Liu, H., Shi, Y., Liu, A., Zhu, Q., Albu, I., Pacholec, M., Cheng, L., Sun, X., & Chi, X. (2025). Harnessing Small-Data Machine Learning for Transformative Mental Health Forecasting: Towards Precision Psychiatry With Personalised Digital Phenotyping. *Med Research*. https://doi.org/10.1002/mdr2.70017
Wang, P., Wang, Z., & Qiu, S. (2024). Universal, school-based transdiagnostic interventions to promote mental health and emotional wellbeing: A systematic review. *Child and Adolescent Psychiatry and Mental Health, 18*(1), 47. https://doi.org/10.1186/s13034-024-00735-x
Wang, Y., Steenbergen, B., van der Krabben, E., Kooij, H.-J., Raaphorst, K., & Hoekman, R. (2023). The Impact of the Built Environment and Social Environment on Physical Activity: A Scoping Review. *International Journal of Environmental Research and Public Health, 20*(12), 6189. https://doi.org/10.3390/ijerph20126189
World Health Organization. (2024). March 1 *Obesity and overweight* https://www.who.int/zh/news-room/fact-sheets/detail/obesity-and-overweight.
Yang, M., Violato, M., & Carson, C. (2025). The changing role of family income in mental health from childhood to adolescence: Findings from a UK longitudinal study. *Archives of Public Health, 83*(1), 224. https://doi.org/10.1186/s13690-025-01702-4