Integrating Multi-Ancestry Polygenic Risk Scores and Wearable Data to Detect Depression and Gene-Environment Interactions in Youth

Nathan Chen[1], Mahnoor Hyat[1], Jinhan Zhu[1], Anind Dey[2], & Jennifer K. Forsyth[1]

[1]Department of Psychology, University of Washington, Seattle, WA,

[2]Information School, University of Washington, Seattle, WA

Abstract

Background: Depression (DEP) has an estimated heritability of ~37%. Polygenic risk scores (PRS), which aggregate common genetic variant effects, account for up to 8.4% variance in case versus control status. Understanding the relationships between sleep, activity, and DEP symptoms, and whether these relationships differ depending on genetic predisposition is crucial for early identification, particularly during adolescence when DEP rates tend to increase.

Method: 2,768 adolescents (African [AFR], American Admixed [AMR], and European [EUR] ancestry) from the Adolescent Brain Cognitive Development study were examined. DEP-PRS was calculated using PRS-CSx. Average and variance sleep and activity measures were derived from Fitbit wearable data. Mixed effects models examined main and interaction effects of DEP-PRS and each Fitbit measure with DEP severity.

Results: DEP-PRS, mean resting heart rate, mean sedentary minutes, and sleep minutes variance were positively associated with DEP severity; mean intense active minutes, total steps, and sleep minutes were negatively associated. DEP-PRS showed interactions with sleep minutes variance, mean total steps, and mean intense active minutes, with stronger associations between these variables in individuals at higher genetic risk. In individuals at higher genetic risk with different ancestries, the association between mean sleep minutes and DEP severity was positive for AFR, negative for AMR, and non-significant for EUR.

Conclusion: The relationship between sleep, activity, and DEP severity appears to depend on genetic predisposition Further research with optimized PRS, research-grade wearables, and longitudinal design is needed.

Introduction

Depression (DEP) is a leading cause of illness and disability among adolescents but remains a frequently underrecognized mental health concern (Thapar 2012). Detecting DEP in youth is often complicated by the predominantly internalizing nature of symptoms, such as anhedonia, fatigue, and difficulty concentrating, which may not be readily apparent to parents, educators, or clinicians (Black 2022). Yet, as untreated depression can result in significant morbidity, heightened risk of self-harm, and impaired school and social functioning during a critical period of development, early detection is critical for facilitating prevention strategies and halting potential negative sequelae.

DEP risk is significantly influenced by both environmental and genetic factors, with twin studies measuring heritability at around 40 percent (Sullivan et al., 2000). Genome-wide association studies (GWAS) have now identified hundreds of common genetic variants that are associated with DEP. However, polygenic risk scores (PRS) for DEP, which estimate an individual's genetic risk for the disorder as a weighted sum of common risk variants, indicate that PRS can only explain up to 8.7% of the variance in DEP case-control status (Wray et al., 2018). Modest predictive accuracy of DEP PRS likely reflects the partial and strong contributions of environmental factors to DEP. Moreover, given that European ancestry individuals are strongly overrepresented in GWAS of DEP and allele frequencies and linkage disequilibrium patterns vary across global populations, DEP PRS show their highest predictive accuracy in individuals of European ancestry, with lower accuracy in underrepresented populations. This raises critical healthcare equity concerns (Martin et al., 2019). Thus, while PRS holds potential for improving early DEP risk prediction, the current clinical utility of PRS

remains limited, emphasizing the need for more inclusive and comprehensive approaches to DEP risk prediction.

One potential complementary path to elucidate the risk factors of DEP alongside PRS is to incorporate measurements of physical activity and sleep (Firth et al., 2020). Decreases in physical activity and increases in sedentary behavior are associated with increased risk of DEP in adolescents (Wang & Pieper, 2022; Hou et al., 2024). Sleep disturbances often precede DEP onset, which can in turn worsen sleep problems in a bidirectional manner (Goldstone et al., 2020, Lovato & Gradisar, 2014; Bacaro et al., 2024). However, a challenge in validating associations between activity and sleep relative to DEP is the discrepancy between self-report and wearable measures of activity and sleep (Kiss et al., 2025; Cerin et al., 2016; Prince et al., 2008). In healthy populations, systematic reviews indicate wearable measures of total sleep time, wake after sleep onset, resting heart rate, and total steps have no significant differences compared to polysomnography (PSG) and research-grade actigraphy devices (Haghayegh et al., 2019; Fuller et al., 2020). Meanwhile, several reviews have examined the lack of agreement between self-reports and PSG measures of total sleep and sleep onset latency, which has also been debated as a construct named sleep misperception (Harvey & Tang, 2012; Castelnovo et al., 2019). Other studies suggest the confounding effects of mood may explain the stronger associations between self-report sleep and DEP symptoms compared to wearable measures (George et al., 2019; Akre et al., 2025). Importantly, multiple studies have found longitudinal associations between wearable measures and DEP and its potential in remote monitoring and patient feedback applications (Nelson et al., 2022; Wainberg et al., 2021; Zhang et al., 2021; Borghare et al., 2024). Wearable measures may also capture a variety of dimensional phenotypes including

resting heart rate, step count, sleep variability, and macro sleep architecture, which may further elucidate associations between PRS and granular behavior measures.

In addition to wearable measures of activity and sleep potentially having utility for early detection of DEP, they may also be influenced by genetic risk for DEP, and their utility for early detection of DEP may depend on genetic risk. In one study that examined an adult population without psychiatric conditions, DEP PRS was associated with wearable measures including reduced walking and moderate activities and increased sedentary and sleep time (Dennison et al., 2021). However, the association between PRS and wearable measures for adolescents is unclear. One study found significant positive correlations between DEP PRS and parent-reported measures of disorders of initiating and maintaining sleep and disorders of excessive somnolence for adolescents (Ohi et al., 2021). Another study found that elevated risk for DEP based on having a parent with DEP was associated with average of 26 minutes more total sleep on weeknights than controls; however, familial-based risk of DEP may capture environmental influences to a greater extent than DEP PRS (Wescott et al., 2019). Two recent studies using self-report measures of sleep and activity in adults found that higher levels of recreational activity and sleep durations of 7 to 9 hours grouped with other health lifestyle factors, respectively, were protective of DEP severity regardless of genetic predisposition (Choi et al., 2020; Zhao et al., 2023). However, these studies utilized self-report lifestyle data from one baseline questionnaire and did not examine non-European populations due to the limitations of the selected PRS methods. Furthermore, it is unclear if similar relationships exist in adolescents. Therefore, investigating the associations between DEP PRS, wearable-derived activity and sleep measures, and DEP severity among adolescents with diverse ancestry, is a crucial next step for the field.

Using data from the ABCD Study, a U.S.-based large-scale study on adolescent development with genomic, clinical, demographic and Fitbit wearable data, this study therefore used a multi-ancestry optimized DEP PRS to examine whether: 1) DEP PRS is associated with wearable measures of activity and sleep in individuals without DEP; 2) whether activity and sleep patterns are differentially associated with current DEP severity depending on DEP PRS; 3) and whether these associations are concordant or discordant across different ancestries.

Methods

Participants and Ethical Considerations

Data from the ABCD study were used for the present investigation, which is a general population cohort of 11,878 children ages 8–11 years recruited from 21 sites throughout the U.S. (Barch et al., 2018). Each recruitment site obtained assent and consent from the children and their parent(s)/legal guardian(s), respectively, in accordance with local Institutional Reviews Boards. Data analyzed for the present study involved individuals of European (EUR), African (AFR), or American Admixed ancestry (AMR), and with valid Fitbit physical activity and sleep data, resulting in 2,768 unique participants at year 2. Year 2 data was used to obtain Fitbit wearable data from the full ABCD Fitbit protocol. Data was obtained from Release 5.1 from the ABCD consortium.

Depression History and Symptom Severity

Lifetime history of DEP (Lifetime-DEP) was assessed using the computerized Kiddie-Schedule for Affective Disorders–5 (K-SADS) based on endorsement from either child- or parent-report. A binary diagnosis of whether participants have ever had a K-SADS DEP

diagnosis up to year 2 of the ABCD study was used to validate and select the best performing PRS for each ancestry group. Subsequent analysis used DEP severity in the past six months, which was measured using scores from the "Depressive Problems" subscale of the Child Behavior Checklist (CBCL), a widely used measure of emotional problems and provides standardized scores based on national norms in children ages 6–18. The Depressive Problems subscale comprises items consistent with the DSM-5 criteria for DEP and was available from the parent report CBCL and not from the youth. The CBCL has been found to be valid and reliable for children with diverse backgrounds (Albores-Gallo et al., 2007).

Fitbit Data Quality Control and Summary Statistics

Participants in the ABCD study were invited to participate in an additional 3-week Fitbit study protocol. Daily and weekly summaries of physical activity and sleep were available, which includes total and average metrics by day such as the number of sedentary minutes or average resting heart rate during the daytime. This study utilized the daily Fitbit summary data, which the ABCD team had conducted preliminary quality control (QC) by pre-filtering daily sleep data to days with at least 300 minutes in nighttime wear time. Additional QC measures were performed by this study on the activity data following recommendations of ABCD Fitbit pilot study results (Wing et al., 2022). Days of activity data with less than 600 minutes in daytime wear time were excluded. Participants with less than 7 days of data or with data during Covid-19 starting on March 15th, 2020, were removed. Data was also restricted to the 3-week protocol period by removing additional days of data over 21 days, such that participants included in the analyses had a range of days with data between 7 days and 21 days (Table 1).

Summary statistics variables were constructed from daily-level time-series Fitbit data for the main analyses. Participant data were summarized using the average and variance statistics.

Each variable was z-score standardized to facilitate comparison of effect sizes and coefficients across variables, then utilized for mixed effects model analyses. To account for day-to-day variations in total wear time or total sleep time, variables' values proportional to total wear time (proportion) were created for main analyses, such as the percentage of total sleep time spent in deep sleep (mean proportion deep sleep minutes) or the percentage of total daytime wear minutes spent in sedentary time (mean proportion sedentary minutes).

GWAS Summary Statistics & Polygenic Risk Scores

GWAS summary statistics for DEP were obtained from meta-analyses and separate outputs for EUR, AFR, and AMR individuals (Wray 2018, Meng 2024). QC was conducted using the GenoPred pipeline on the GWAS summary statistics by only using high-quality variants with a minor allele frequency > 0.01, without ambiguity, no out-of-bound p- values, and no duplicates (Pain et al., 2024).

QC-ed GWAS summary statistics from EUR, AFR, and AMR were used as inputs into the PRS-CS and PRS-CSx PRS methods. PRS-CS is a method that refines single nucleotide variant (SNP) effect sizes using continuous shrinkage on a single set of GWAS summary statistics while accounting for Linkage Disequilibrium (LD) using LD reference panels from the 1000 Genomes Project (Ge et al., 2019). The results of the refined posterior SNP effect sizes were then inputted into PLINK's linear scoring command to obtain PRS as Ge et al. (2019) suggested and z-scored standardized within each ancestry. The PRS outputs of PRS-CS will hereby be referred to as PRS-CS EUR, PRS-CS AFR, and PRS-CS AMR. PRS-CSx is the successor of PRS-CS by utilizing multiple ancestry-specific GWAS summary statistics to improve cross-population polygenic prediction (Ruan et al., 2022). PRS-CSx provides both ancestry-specific SNP effect sizes outputs (PRS-CSx EUR, PRS-CSx AFR, and PRS-CSx AMR)

and a meta option (PRS-CSx META) where an inverse weighted variance analysis is conducted on PRS-CSx EUR, PRS-CSx AFR, and PRS-CSx AMR outputs without fine-tuning on a target population. In total, 7 outputs from PRS-CS and PRS-CSx were compared in a validation dataset that includes ABCD participants without valid Fitbit data to select a single best-performing PRS for each ancestry before subsequent analysis with Fitbit variables and CBCL DEP severity.

Statistical Analyses

Mixed effects models, as implemented in the GENESIS package in R, were used to investigate associations among DEP-PRS, Fitbit variables, and CBCL DEP severity (Gogarten et al., 2019). Covariates included age, gender, income-to-needs ratio, ancestry principal components (PC), and a genetic relatedness matrix. Genetically derived ancestry PCs and genetic relatedness matrix were derived from a previous study that used the GENESIS package's pipeline for Principal Components Analysis (PCA) and kinship coefficients accounting for population structures in ABCD subjects (Hyat et al., 2025). Income-to-needs ratio was derived using the income bracket and family size information on each ABCD participant by dividing the mean of the selected income bracket by the federal poverty threshold depending on family size. Our primary analysis using mixed effects models involved three specific setups. First, DEP-PRS derived by PRS-CS and PRS-CSx were tested against lifetime-DEP in ABCD participants (n=7,681) to validate and pick the best-performing PRS for each ancestry. Then, the optimized DEP-PRS were tested for main effects with each Fitbit variable as the dependent variable, including or excluding participants above the subclinical threshold for CBCL DEP severity. Third, the main and interaction effects of each Fitbit variable and DEP-PRS were tested in the same model with CBCL DEP severity as the dependent variable. To account for multiple testing effects, false discovery rate (FDR) was applied to models separately by the type of main analyses

and by either activity or sleep. For variables with significant interaction effects, follow-up analyses were conducted within each stratified genetic risk group, defined by low risk = z-scored DEP-PRS < -1, medium risk = z-scored DEP-PRS <= 1, >= -1, and high risk = z-scored DEP-PRS > 1, to explain interaction results rather than using formal pairwise comparisons since there is no established risk categories based on PRS. Ancestry-specific analyses were conducted for the second and third part of the primary analyses. Sensitivity analyses were performed by excluding socioeconomic status as a covariate for ancestry-specific models due to significant differences in income-to-needs ratio between ancestries.

Results

Demographics

2,768 participants with valid Fitbit activity and sleep data, CBCL DEP severity assessment at year 2, genotype data, and EUR (n=2,013), AFR (n=350), and AMR (n=405) ancestry were included (Table 1). Participants who passed Fitbit quality control protocol had an average 21.04 days of activity data (SD=2.21) and 14.49 days of sleep data (SD=4.24).

DEP-PRS, lifetime history of DEP, and current DEP severity

DEP-PRS outputs from PRS-CS and PRS-CSx were compared by testing for association with KSADS lifetime DEP (Table 2). For EUR, PRS-CS EUR (OR = 1.32, $p < 0.001$, 95% CI [1.19, 1.46]) outperformed PRS-CSx outputs. For AFR, PRS-CSx EUR (OR = 1.14, $p = 0.13$, 95% CI [0.96, 1.35]) outperformed other PRS-CSx and PRS-CS outputs. For AMR, PRS-CSx META (OR = 1.23, $p = 0.02$, 95% CI [1.01, 1.50]) outperformed other PRS-CSx and PRS-CS outputs.

Associations between DEP-PRS and activity and sleep metrics

With each activity measure as the dependent variable, significant associations were observed between DEP-PRS and 9 out of 14 activity measures, 8 of which survived FDR correction. Higher DEP-PRS was associated with higher mean resting heart rate and mean proportion of sedentary minutes, and lower mean total steps and mean and variance of fairly and very active minutes (Figure 1.1A).

With each sleep measure as the dependent variable, significant associations were observed between DEP-PRS and 15 out of 20 sleep measures, 6 of which survived FDR correction (Figure 1.2A). Higher DEP-PRS was associated with lower mean proportion of wake counts and multiple sleep minutes mean measures including total, wake, light, and deep sleep. Higher DEP-PRS was also associated with greater variance in multiple sleep minutes measures including variance in total, light, and REM sleep. DEP-PRS was not associated with any sleep mean or variance measures of proportion in specific sleep stages.

Associations between DEP severity, DEP-PRS, activity, and sleep metrics

Activity

In models for each activity measure and DEP-PRS in the same model where CBCL DEP severity was the dependent variable, DEP-PRS was associated with DEP severity in all models (mean beta (SD) = 0.49 (0.02), FDR $p < 0.001$) and significant associations were observed between ten out of 14 Fitbit activity variables and CBCL DEP severity, nine of which survived FDR correction (Figure 2.1A). Among the seven significant mean variables, proportion of total steps, sedentary minutes, resting heart rate, proportion of fairly active minutes showed the highest magnitude of effects. For the mean measures of different categories of activity intensity,

fairly active minutes had the highest negative association, followed by lightly active minutes and very active minutes. The mean of sedentary minutes, very active minutes, and lightly active minutes each had significant associations while the variance of the respective variables had small effects and non-significance. For variance variables, proportion of fairly active minutes, total steps, and resting heart rate had significant associations.

Significant interactions were also observed between multiple activity measures and DEP-PRS, including fairly and very active minutes variance, with the former surviving FDR correction (Figure 2.2A). Follow-up analyses within PRS risk groups for very active minutes variance showed a positive association with DEP severity for subjects in the low risk group and negative associations for subjects in the medium or high PRS group (Figure 2.3A).

Sleep

In models with each sleep measure and DEP-PRS in the same model where CBCL DEP severity was the dependent variable, DEP-PRS and DEP severity were also significant associated, (mean beta (SD) = 0.49 (0.01), FDR $p < 0.001$), and significant associations were observed for 15 out of 20 Fitbit sleep variables, 13 of which survived FDR correction (Figure 3.1A). The highest magnitudes of effects were observed in mean heart rate during various sleep stages, variance in light, wake, or total sleep minutes, and variance and mean in wake counts. For the sleep variables in different sleep stages, associations were observed in all of the respective variance variables, while the mean measures for time spent in wake or total sleep were significant. The variance in time spent in light, REM, and deep sleep stages all had higher main effect sizes compared to their respective proportional variables.

In addition, four interactions between DEP-PRS and sleep metrics for current DEP severity were significant (Figure 3.2A). In particular, there were interactions between DEP-PRS

and variances in time spent in each of the four sleep stages, including proportional measures. Follow-up analyses within PRS risk groups showed the positive association between each sleep measure and CBCL DEP severity were stronger for subjects in the high PRS group, followed by the medium and low PRS group (Figure 3.3A). In the high PRS group and among the different sleep stages, variances in proportion of deep sleep minutes had the highest association with DEP severity, followed by wake, light sleep, and rem sleep minutes.

Ancestry-stratified associations between DEP severity and DEP-PRS, activity, and sleep metrics

Activity

Significant associations detected between DEP-PRS and activity measures in the multi-ancestry model became stronger and had higher effect sizes in the EUR only sample (Figure 1.1B). For AFR, the only significant association was between PRS and increased fairly active minutes variance, which was in the positive direction compared to the significant negative association for EUR (Figure 1.1C). For AMR, no significant associations were found (Figure 1.1D).

Significant main associations between each Fitbit activity variable and CBCL DEP severity remained significant with similar effect sizes in the EUR only sample (Figure 2.1B). In both AFR and AMR, there were less significant associations, but effect directions remained similar (Figure 2.1C & Figure 2.1D).

Significant interactions from the multi-ancestry model that remained significant in the EUR only sample included light, fairly, and very active minute variance, while the interactions for the mean of the respective variables became significant (Figure 2.2B). For AFR, nominally significant interactions include PRS and very active minutes mean, which was in the opposite, positive direction compared to EUR, as well as a negative interaction with light active minutes

mean, which were not found in EUR or AMR (Figure 2.2C & Figure 2.2D). Follow-up analyses suggested the association between very active minutes mean and DEP severity for AFR is positive for the high risk group and negative for the low risk group, and vice versa for the EUR follow-up analyses (Figure 2.3B & Figure 2.3C). Furthermore, follow-up analyses suggested the association between light active minutes mean and DEP severity for AFR was negative only for the high risk group. For AMR, no significant interactions were observed, while effect sizes were generally more similar to EUR rather than AFR (Figure 2.3D).

Sleep

Significant associations between DEP-PRS and sleep measures in the multi-ancestry sample became less significant but had similar directions and effect sizes in the EUR only sample (Figure 1.2B). For AFR, 2 significant associations were observed between DEP-PRS and decreased mean proportion of REM sleep minutes, which survived FDR correction, and increased mean proportion of light sleep minutes (Figure 1.2C). Both significant associations in AFR were not observed in full sample, EUR, or AMR, and had higher effect sizes compared to significant variables in other ancestries. For AMR, DEP-PRS was nominally positively associated with deep and REM sleep variance, where the latter association was observed in the full and EUR only samples (Figure 1.2D).

Significant main associations between each sleep variable and CBCL DEP severity remained significant with similar effect sizes in the EUR only sample (Figure 3.1B). The significant negative associations of total and light sleep minutes mean with DEP severity were only observed in EUR. The effect sizes for those variables were in the opposite direction but non-significant for other ancestries, specifically with light minutes mean for AFR and total sleep minutes mean for AMR (Figure 3.1C & Figure 3.1D). Significant associations between sleep

variance measures and DEP severity were in similar, positive directions for EUR and AMR, but stronger effect sizes were observed for AMR with total and deep sleep variance.

Significant interactions from the multi-ancestry model, specifically sleep variance measures, remained significant in the EUR sample (Figure 3.2B). Follow-up analyses for EUR suggest similar associations to the multi-ancestry sample as well (Figure 3.3B). For AFR, a nominally significant negative interaction between PRS and REM sleep minutes mean was observed, which was not observed in EUR and observed as an opposite, positive interaction for AMR (Figure 3.2C & Figure 3.2D). Follow-up analyses suggested that the association between REM minutes mean and DEP severity is strong and negative for the AFR high risk group, and negative for the AMR high risk group (Figure 3.3C & Figure 3.3D). For AMR, 4 nominally significant associations, including positive interactions with total and deep sleep minutes mean and negative interactions with light sleep minutes mean and wake counts variance, were not observed in other ancestries nor in the full sample. Follow-up analyses suggested positive associations between total and deep sleep minutes mean and DEP severity for the AMR high risk group, and negative associations for both variables in the AMR low risk group.

Discussion

This study examined the interplay among DEP-PRS, wearable-based physical activity and sleep data, and DEP severity outcome measures. DEP-PRS was associated with lifetime DEP, activity, and sleep measures. DEP-PRS also showed interaction effects with wearable measures of activity and sleep in association with DEP severity, where stronger associations between wearable measures and DEP severity were observed in individuals at higher genetic risk. There are ancestry-specific differences in wearable associations in individuals at higher genetic risk.

Recent literature has suggested multi-ancestry methods, like PRS-CSx, could perform better than single-ancestry methods on diverse populations depending on method optimization, GWAS data selection, and ancestral population, but have not utilized psychiatric phenotypes (Wang et al., 2023; Gunn et al., 2025). Several studies on the ABCD dataset have associated multi-ancestry PRS and DEP measures using PRS-CS and the PRS-CSx META option (Ahern et al., 2023, Lee et al., 2025). This study corroborates the stronger association between PRS-CS EUR and KSADS lifetime DEP for EUR in the ABCD dataset but also indicates the ancestry-specific outputs of PRS-CSx, which were not used by previous studies, could be useful. PRS-CSx EUR outperformed both PRS-CSx META and PRS-CS AFR for AFR, while PRS-CSx META performed best for AMR. Alternatively, another study on the ABCD dataset used the same KSADS outcome measure and derived multi-ancestry PRS with SBayesRC, a PRS method that integrates functional annotations to pinpoint causal variants (Xu et al., 2024, Zheng et al., 2024). Xu et al. derived odds ratios for EUR and multi-ancestry populations at 1.33 ([1.15 – 1.52], $p<0.001$) and 1.31 ([1.15 – 1.50], $p<0.001$), respectively, while this study found odds ratios for the same outcome at 1.31 ([1.18 - 1.45], $p<0.001$) and 1.20 ([1.11 - 1.29], $p<0.001$). However, this study utilized ancestry-specific GWAS summary statistics to take advantage of the fine-mapping approach of PRS-CSx, while Xu et al. (2024) utilized a different set of meta-analyzed GWAS summary statistics (Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, 2025). Therefore, future studies are needed to systematically evaluate the use of existing PRS methods and GWAS summary statistics for diverse populations with psychiatric phenotypes, as more polygenicity and heterogeneity could implicate PRS accuracy.

In a previous systematic review of wearable features in association with DEP, 22 of 48 studies reported statistically significant associations for physical activity variables, with vigorous activity having one of the higher associations among all reported variables (Rohani et al., 2018). While studies from Rohani et al. (2018) were adult samples, recent studies using the Fitbit measures from the ABCD dataset have indicated higher mean resting heart rate, lower mean and variability in total steps, and lower average time spent in fairly and very active minutes are associated with higher internalizing symptoms (Nelson et al., 2022; Damme et al., 2024). Results from this study corroborate the main effects in association with CBCL DEP severity among all previously mentioned variables and increased mean sedentary minutes. Extending on previous studies, this study identified significant interaction effects in the multi-ancestry sample between PRS and average total steps and variance in fairly and very active minutes, with stronger negative associations as PRS increases. Existing studies on adult populations have found no interactions between a EUR-only PRS and self-report activity and indicate physical activity as a protective factor for DEP risk across different levels of genetic predisposition (Choi et al., 2020, Zhao et al., 2023). However, our study used multi-ancestry optimized PRS and objective measures of behavior and identified cross-sectional significant interactions. A previous study also suggests higher rates of vegetative symptoms in adolescent-onset DEP compared to adult recurrent DEP (Rice et al., 2019). Results of this study further indicate a need for longitudinal studies to examine the bidirectional relationship between lifestyle factors and DEP in adolescents depending on genetic risk. Alternatively, in ancestry-specific analysis, the association between mean very active minutes and DEP severity was more positive as PRS increases for AFR, but the opposite interaction effects were identified for EUR. This discordant association suggests caution in interpreting wearable-derived minutes of intense activity, as such measures can capture both

potentially protective physical activity and activity changes due to differences in socioeconomic factors (Sallis et al., 1996).

Existing systematic reviews suggest associations between wearable measures of sleep duration variability and sleep efficiency or sleep quality, rather than average sleep duration, with DEP symptoms (Zierer et al.,2024; Messman et al., 2024; De Angel et al., 2022). While the lack of associations for average sleep duration could be due to non-linear associations, a prior study in the ABCD dataset found no significant non-linear association for wearable measures of average sleep duration, and implicates higher variability with higher internalizing symptoms as well (Nelson et al., 2022). Our study is consistent with these prior findings and builds on them, in terms of observing consistent associations of variability in sleep duration, variability in wake time during sleep, as well as mean wake counts with DEP severity, as well as interaction effects w/ PRS, which suggests these variables are generally associated but even more relevant for individuals with higher genetic predisposition. In ancestry-specific analyses, while we found average sleep duration had non-significant main effects, average duration of sleep in in specific sleep stages show significant interactions with PRS, highlighting the importance of average sleep duration in specific sleep stages when considered within-ancestry and alongside PRS due to potential genetic and race/ethnicity factors. Furthermore, our analyses indicate that for the African ancestry population, PRS is negatively associated with the proportion of REM sleep as well as a significant interaction effect between higher REM sleep and higher PRS with higher DEP symptoms, while PRS was not associated with the proportion of REM sleep in the trans-ancestry model or for other ancestries. Meta-analysis of objective sleep measurements in association with DEP, including both sleep efficiency measures and sleep architecture measures, have further emphasized the importance of the sleep variability and efficiency measures, rather

than the non-significant effects for total deep sleep, REM sleep, REM density, and REM latency and significant heterogeneity in true effect sizes of these variables across the pooled studies (Lovato & Gradisar, 2014; Augustinavicius et al., 2014). However, these studies were not ancestry-specific, and there are studies on racial and ethnic differences in sleep architecture that warrants further investigation on whether these differences also present as risk factors for DEP (Johnson et al., 2019; Rao et al., 2019, Poland et al., 1999). Future studies should explore whether these sleep differences are risk factors for DEP, utilize ancestry-specific longitudinal analysis, or incorporate more accurate genetic influence measures.

The strengths of this study include the use of a large (n=2,768) and diverse (AFR, AMR, and EUR) population, multi-ancestry PRS methods, and wearable data to derive associations between activity, sleep, and DEP symptoms depending on genetic predisposition and ancestry. However, there are several limitations . First, this is a cross-sectional analysis on the ABCD dataset, where CBCL symptoms were measured right before the Fitbit protocol at year 2, which means causality cannot be established. Second, although multi-ancestry PRS methods performed better than single-ancestry PRS methods, the association between PRS and both KSADS lifetime history of DEP and CBCL DEP severity for the AFR population is still not significant and could hinder the validity of associations with behavior variables, which emphasizes the need for optimized PRS methods. Third, further investigation is needed to compare the accuracy of self-report and wearables with gold-standard measures concurrently, especially when clinical symptoms are present (Kawai et al., 2023). Several studies do indicate significant differences between wearables and gold-standard PSG in measuring sleep when individuals present clinical symptoms of DEP or insomnia (Cook et al., 2017; Kang et al., 2017). The accuracy of wearable measures of macro sleep architecture (light, deep, and REM sleep) is lower than total sleep time

and sleep efficiency, which could affect the validity of associations this study found for specific sleep stages (Haghayegh et al., 2019). Also, the accuracy for measuring energy expenditure, which is used to categorize the number of minutes in light, fairly, and very active minutes, is less than total steps (Fuller et al., 2020). Wearables capture all forms of movement and do not automatically infer the specific types of activities. Even though there is a significant negative association between specific intensity of active minutes and DEP severity, active minutes can capture activities that may have different associations with mental health if able to be classified separately (Butte et al., 2018). Furthermore, previous research highlighted that ABCD participants' socioeconomic demographics were significantly associated with their willingness to participate in the additional Fitbit study and adherence to device wear time, which could induce bias into ABCD wearable prediction studies (Kim et al., 2023).

In conclusion, this study was able to utilize a diverse U.S. adolescent population dataset with genotype data, Fitbit wearable data, and psychiatric assessments to elucidate the gene-environment interactions that warrant further longitudinal studies with research-grade activity and sleep measurements to test the effectiveness of lifestyle interventions for DEP or as predictors of DEP onset.

Tables & Figures

Table 1. Demographic information for ABCD participants in the whole dataset and in each ancestry that passed quality control for Fitbit data and have valid PRS.

Demographics	Whole Dataset (n = 2,768)	AFR (n = 350)	AMR (n = 405)	EUR (n = 2,013)
CBCL DEP Severity Score, Mean (SD)	53.67 (5.87)	53.3 (5.89)	53.8 (5.96)	53.7 (5.85)
Gender, n (%)				
Cis female	1,366 (49%)	189 (54%)	190 (47%)	980 (49%)
Cis male	1,402 (51%)	160 (45%)	215 (53%)	1027 (50%)
Gender Diverse	7 (1%)	1 (1%)	0 (0%)	6 (1%)
Age (in Months), Mean (SD)	143.42 (7.85)	143.0 (7.51)	142.0 (7.97)	144.0 (7.86)
Income-to-Needs Ratio, Mean (SD)	3.21 (1.85)	1.9 (1.63)	1.91 (1.57)	3.71 (1.69)
# of Days with Fitbit Activity Data, Mean (SD)	21.04 (2.21)	20.7 (2.39)	20.9 (2.27)	21.1 (2.16)
# of Days with Fitbit Sleep Data, Mean (SD)	14.49 (4.24)	12.8 (4.07)	14.2 (4.3)	14.8 (4.19)
Mean Daily Daytime Minutes with Fitbit Data, Mean (SD)	864.3 (48.64)	884.0 (52.0)	866.0 (46.8)	861.0 (47.6)
Mean Total Steps, Mean (SD)	9578.78 (3055.73)	9315.0 (2973.0)	8899.0 (3023.0)	9762.0 (3055.0)

Mean Total Sleep, Mean (SD)	449.81 (36.28)	423.0 (41.2)	447.0 (35.9)	455.0 (33.2)

Table 2. Single-ancestry PRS method (PRS-CS) utilized with each ancestry's specific GWAS summary statistics was compared against multi-ancestry PRS method (PRS-CSx) that used all 3 GWAS summary statistics in a large ABCD validation sample that includes participants without valid Fitbit data. OR represents the change in odds of KSADS lifetime history of DEP diagnosis per one standard deviation increase in the PRS. KSADS Lifetime History of DEP diagnoses is defined by having at least 1 diagnosis of DEP from baseline up to year 2 of ABCD study. P refers to nominal p value. The best-performing PRS for each ancestry (in bold) were selected to use in association with Fitbit wearable measures and CBCL DEP severity in the main statistical analyses.

PRS Method	**EUR (n=5,015)**		**AFR (n=1,513)**		***AMR (n=1,101)***	
	OR	***P***	***OR***	***P***	***OR***	***P***
PRS-CS	**1.309**	**2.3e-07**	1.091	0.19	*0.923*	*0.41*
PRS-CSx META	*1.105*	*0.05*	*1.041*	*0.61*	***1.259***	***0.02***
PRS-CSx EUR	*1.094*	*0.08*	***1.139***	***0.14***	*1.174*	*0.14*
PRS-CSx AFR	0.99	0.86	1.062	0.41	*1.03*	*0.70*
PRS-CSx AMR	1.001	0.98	0.899	0.15	*1.116*	*0.25*

Figure 1.1. Results of association between DEP-PRS and each Fitbit activity variable as the dependent variable, with FDR correction applied across all models at once. Each association was conducted with all participants included as well as excluding participants who had CBCL DEP Severity score above the designated clinical cut-off score (>= 65), where the participant would be considered as a DEP case. W/ SES refers to the model including income-to-needs ratio as a covariate alongside with age, gender, ancestry PC, and GRM for random effects. The 4 plots include A (Whole dataset), B (EUR), C (AFR), and D (AMR).

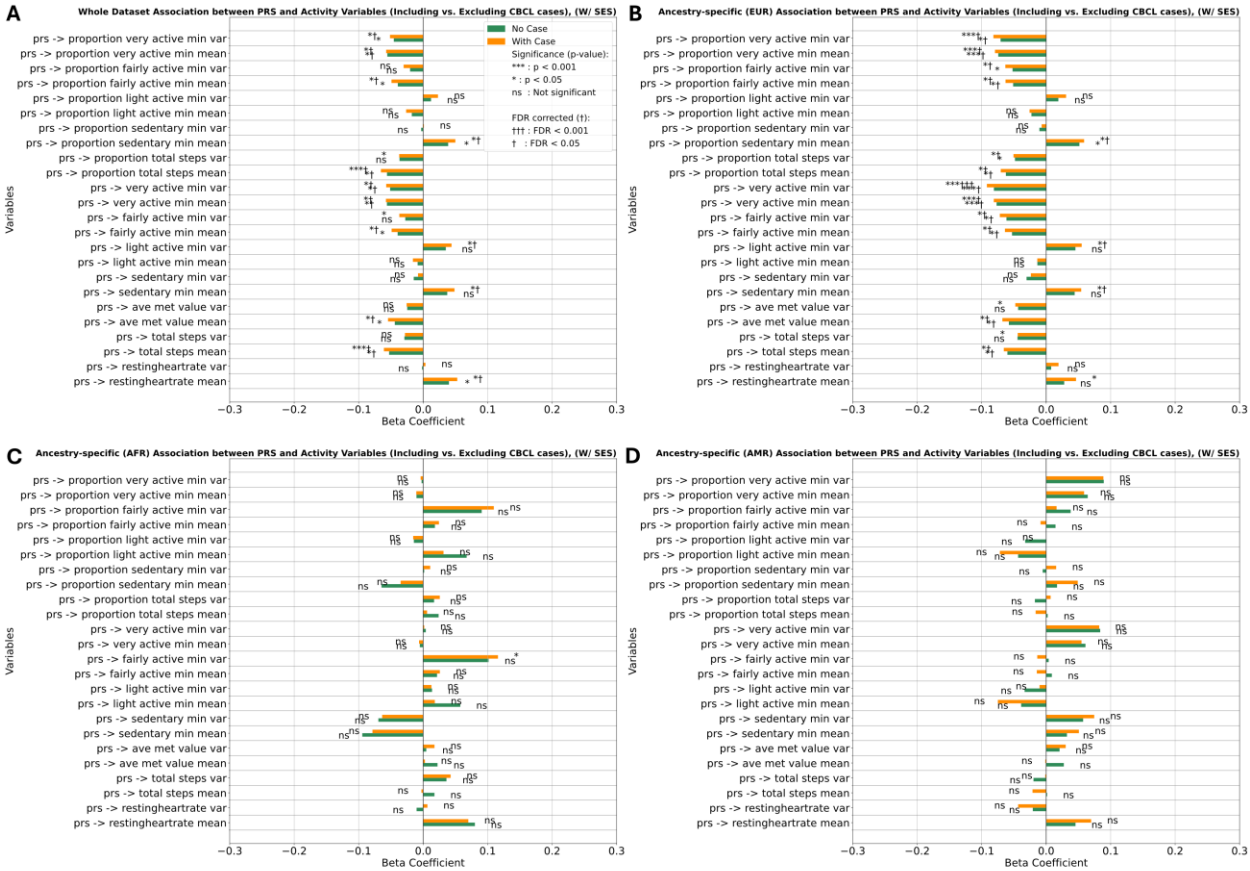


Figure 1.2. Results of association between DEP-PRS and each Fitbit sleep variable as the dependent variable, with FDR correction applied across all models at once. Each association was conducted with all participants included as well as excluding participants who had CBCL DEP Severity score above the designated clinical cut-off score (>= 65), where the participant would be considered as a DEP case. W/ SES refers to the model including income-to-needs ratio as a covariate alongside with age, gender, ancestry PC, and GRM for random effects. The 4 plots include A (Whole dataset), B (EUR), C (AFR), and D (AMR).

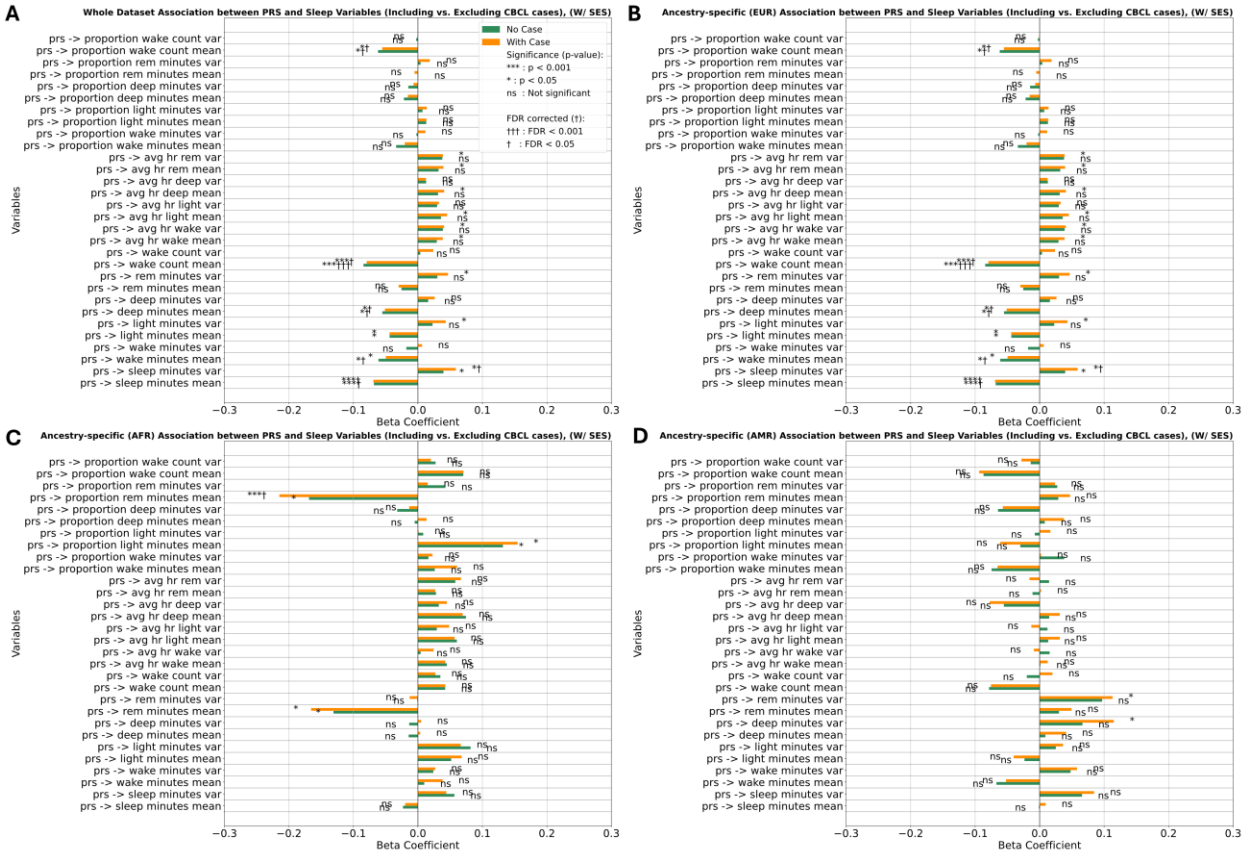


Figure 2.1. Results of main effects of DEP-PRS and each Fitbit activity variable in association with CBCL DEP Severity. The mean and SD beta for DEP-PRS is calculated with the betas of DEP-PRS across different models, where each model has one Fitbit activity measure and DEP-PRS. Interaction effects were also included in each model, and the results are shown in Figure 2.2. FDR correction was applied across all main and interaction effects of activity measures at once. W/ SES refers to the model including income-to-needs ratio as a covariate alongside with age, gender, ancestry PC, and GRM for random effects. The 4 plots include A (Whole dataset), B (EUR), C (AFR), and D (AMR).

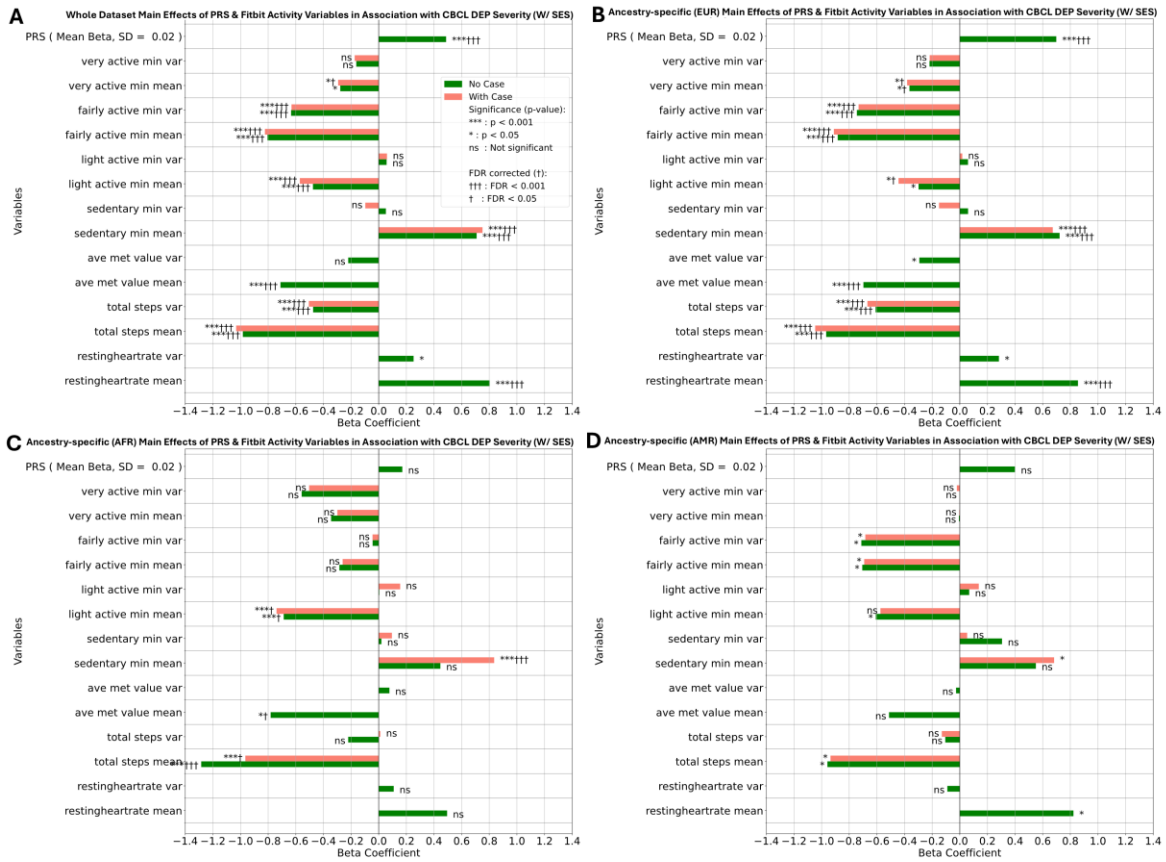


Figure 2.2. Results of interaction effects of DEP-PRS and each Fitbit activity variable in association with CBCL DEP Severity. Follow-up analyses for interaction effects by PRS risk group are in Figure 2.3. W/ SES refers to the model including income-to-needs ratio as a covariate alongside with age, gender, ancestry PC, and GRM for random effects. The 4 plots include A (Whole dataset), B (EUR), C (AFR), and D (AMR).

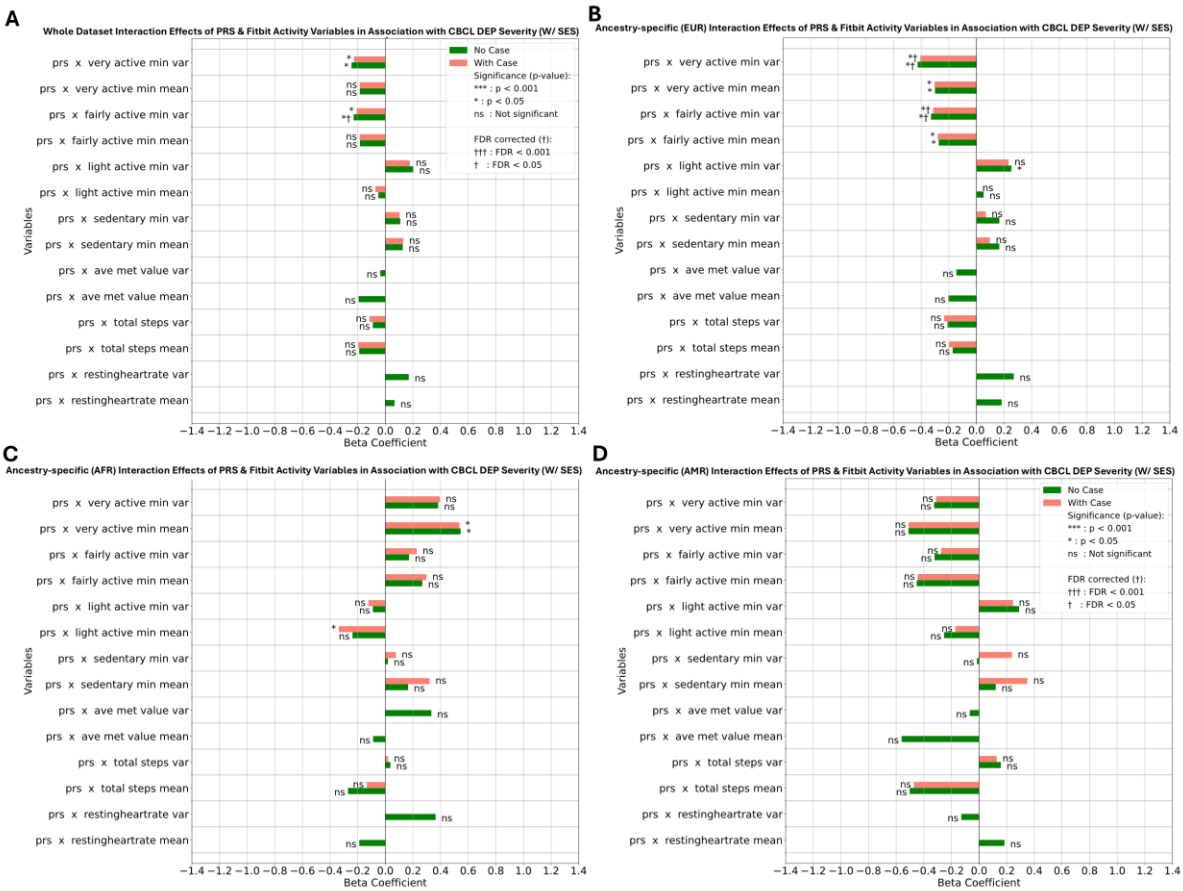


Figure 2.3. Follow-up analyses of interaction effects with results of main effects between each Fitbit activity measure and CBCL DEP severity within each PRS risk group (low, medium, high). NOTE: no formal pairwise comparisons were conducted, as these analyses are meant for checking interaction results. Activity variables that had significant interactions with DEP-PRS were indicated by shaded rows. W/ SES refers to the model including income-to-needs ratio as a covariate alongside with age, gender, ancestry PC, and GRM for random effects. The 4 plots include A (Whole dataset), B (EUR), C (AFR), and D (AMR).

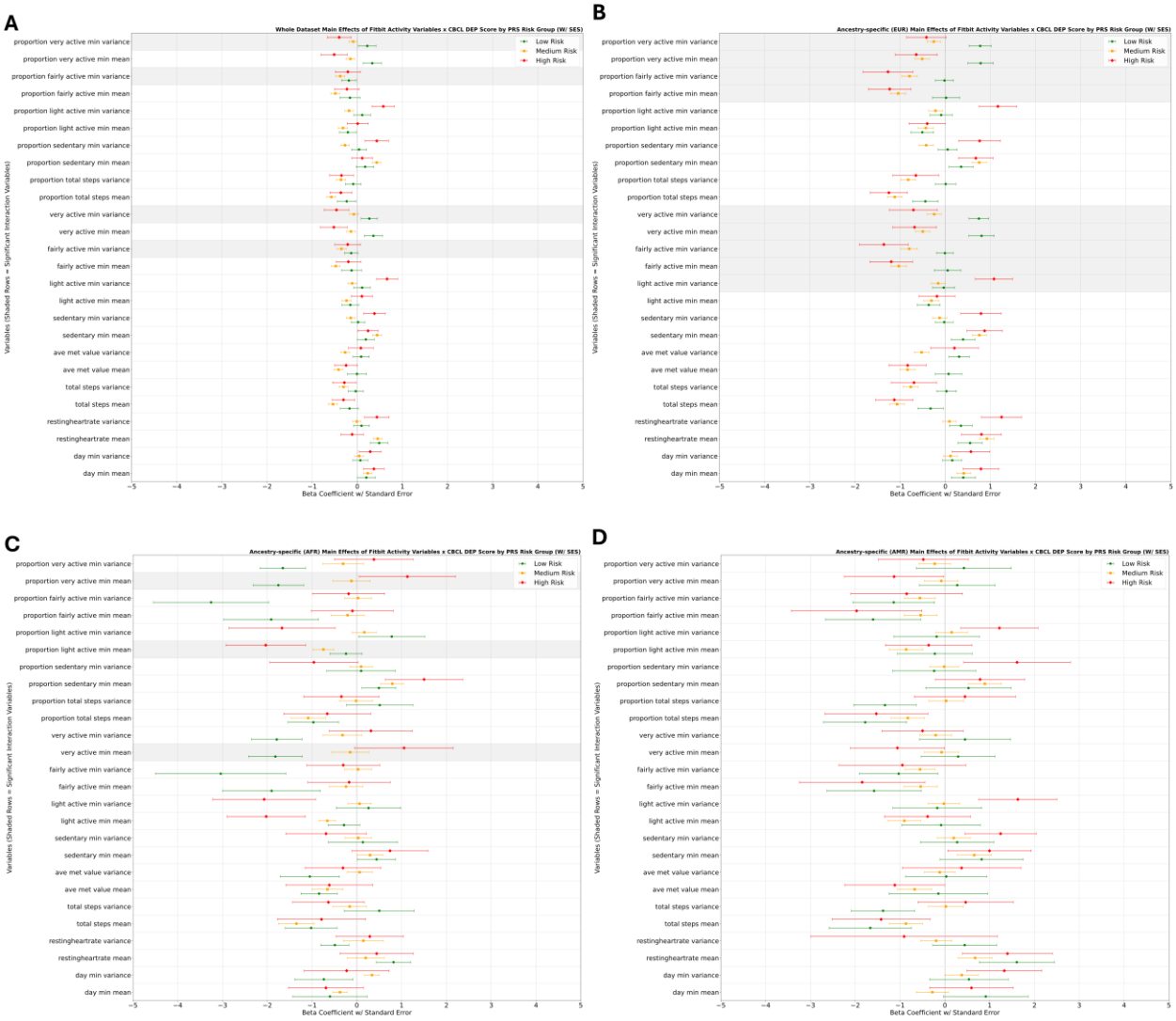


Figure 3.1. Results of main effects of DEP-PRS and each Fitbit sleep variable in association with CBCL DEP Severity. The mean and SD beta for DEP-PRS is calculated with the betas of DEP-PRS across different models, where each model has one Fitbit sleep measure and DEP-PRS. Interaction effects were also included in each model, and the results are shown in Figure 3.2. FDR correction was applied across all main and interaction effects of sleep measures at once. W/ SES refers to the model including income-to-needs ratio as a covariate alongside with age, gender, ancestry PC, and GRM for random effects. The 4 plots include A (Whole dataset), B (EUR), C (AFR), and D (AMR).

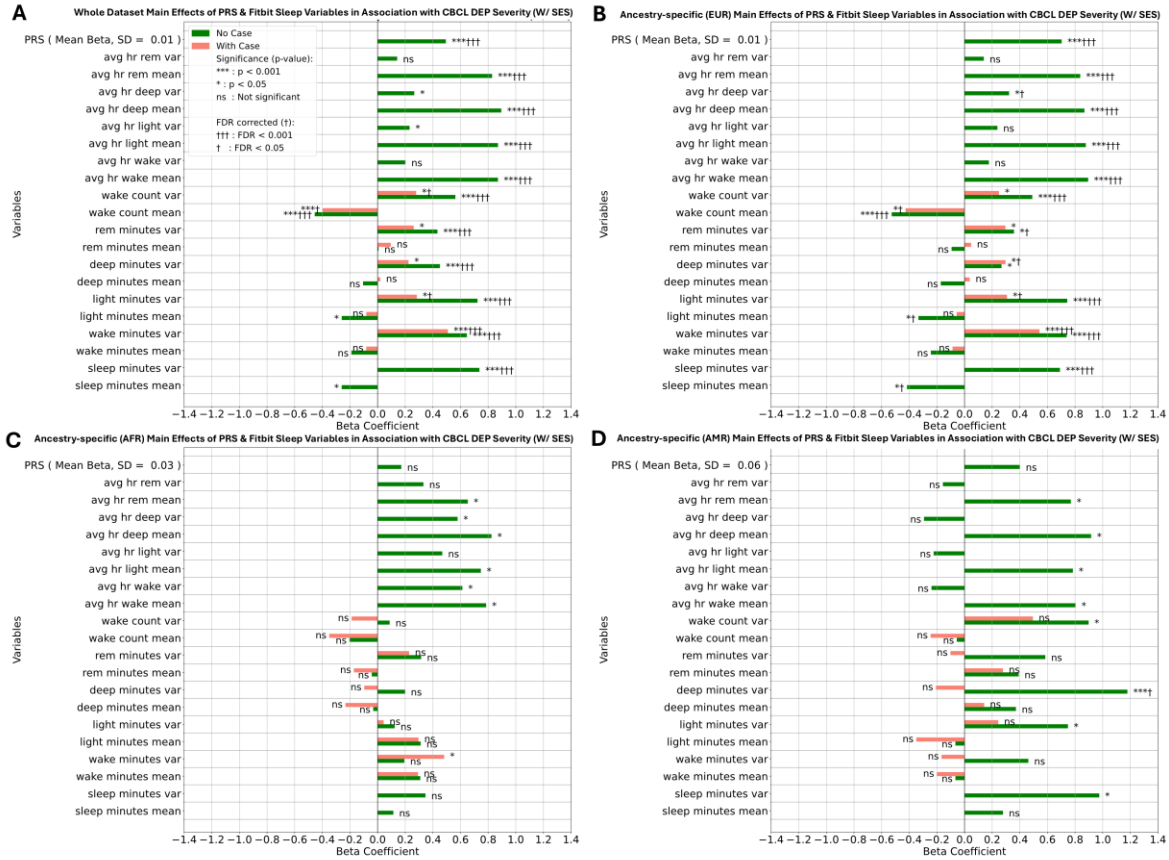


Figure 3.2. Results of interaction effects of DEP-PRS and each Fitbit sleep variable in association with CBCL DEP Severity. Follow-up analyses for interaction effects by PRS risk group are in Figure 3.3. W/ SES refers to the model including income-to-needs ratio as a covariate alongside with age, gender, ancestry PC, and GRM for random effects. The 4 plots include A (Whole dataset), B (EUR), C (AFR), and D (AMR).

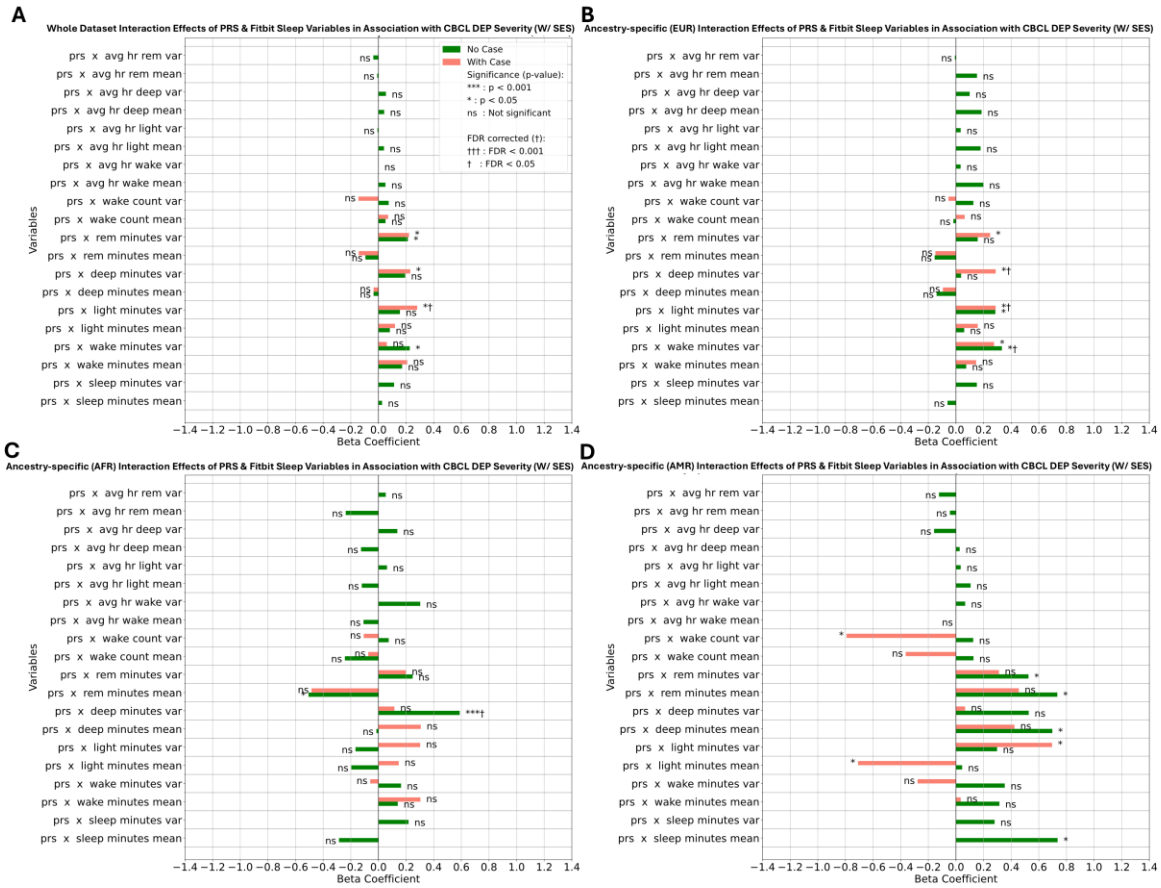
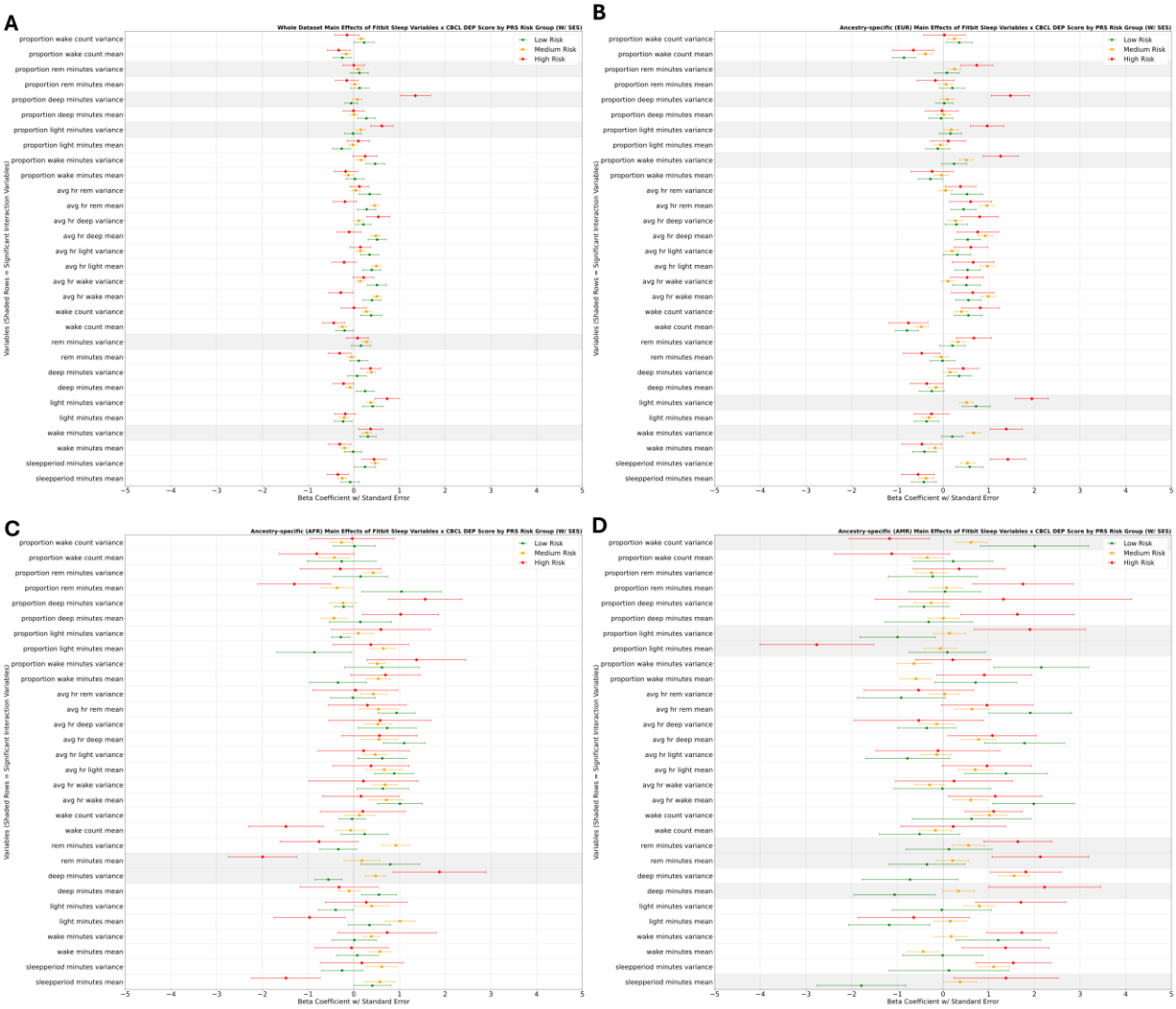


Figure 3.3. Follow-up analyses of interaction effects with results of main effects between each Fitbit sleep measure and CBCL DEP severity within each PRS risk group (low, medium, high). NOTE: no formal pairwise comparisons were conducted, as these analyses are meant for checking interaction results. Sleep variables that had significant interactions with DEP-PRS were indicated by shaded rows. W/ SES refers to the model including income-to-needs ratio as a covariate alongside with age, gender, ancestry PC, and GRM for random effects. The 4 plots include A (Whole dataset), B (EUR), C (AFR), and D (AMR).



References

Ahern, J., Thompson, W., Fan, C. C., & Loughnan, R. (2023). Comparing Pruning and Thresholding with Continuous Shrinkage Polygenic Score Methods in a Large Sample of Ancestrally Diverse Adolescents from the ABCD Study®. *Behavior genetics*, *53*(3), 292–309. https://doi.org/10.1007/s10519-023-10139-w

Albores-Gallo, L., Lara-Muñoz, C., Esperón-Vargas, C., Cárdenas Zetina, J. A., Pérez Soriano, A. M., & Villanueva Colin, G. (2007). Validity and reliability of the CBCL/6-18. Includes DSM scales. *Actas espanolas de psiquiatria*, *35*(6), 393–399.

Akre, S., Cohen, Z. D., Welborn, A., Zbozinek, T. D., Balliu, B., Craske, M. G., & Bui, A. A. T. (2025). Comparing self reported and physiological sleep quality from consumer devices to depression and neurocognitive performance. *NPJ digital medicine*, *8*(1), 92. https://doi.org/10.1038/s41746-025-01493-6

Augustinavicius, J. L., Zanjani, A., Zakzanis, K. K., & Shapiro, C. M. (2014). Polysomnographic features of early-onset depression: a meta-analysis. *Journal of affective disorders*, *158*, 11–18. https://doi.org/10.1016/j.jad.2013.12.009

Bacaro, V., Miletic, K., & Crocetti, E. (2024). A meta-analysis of longitudinal studies on the interplay between sleep, mental health, and positive well-being in adolescents. International journal of clinical and health psychology : IJCHP, 24(1), 100424. https://doi.org/10.1016/j.ijchp.2023.100424

Barch, D. M., Albaugh, M. D., Avenevoli, S., Chang, L., Clark, D. B., Glantz, M. D., Hudziak, J. J., Jernigan, T. L., Tapert, S. F., Yurgelun-Todd, D., Alia-Klein, N., Potter, A. S., Paulus, M. P., Prouty, D., Zucker, R. A., & Sher, K. J. (2018). Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. Developmental cognitive neuroscience, 32, 55–66. https://doi.org/10.1016/j.dcn.2017.10.010

Borghare, P. T., Methwani, D. A., & Pathade, A. G. (2024). A Comprehensive Review on Harnessing Wearable Technology for Enhanced Depression Treatment. *Cureus*, *16*(8), e66173. https://doi.org/10.7759/cureus.66173

Butte, N. F., Watson, K. B., Ridley, K., Zakeri, I. F., McMurray, R. G., Pfeiffer, K. A., Crouter, S. E., Herrmann, S. D., Bassett, D. R., Long, A., Berhane, Z., Trost, S. G., Ainsworth, B. E., Berrigan, D., & Fulton, J. E. (2018). A Youth Compendium of Physical Activities: Activity Codes and Metabolic Intensities. *Medicine and science in sports and exercise*, *50*(2), 246–256. https://doi.org/10.1249/MSS.0000000000001430

Canali, S., Schiaffonati, V., & Aliverti, A. (2022). Challenges and recommendations for wearable devices in digital health: Data quality, interoperability, health equity, fairness. *PLOS digital health*, *1*(10), e0000104. https://doi.org/10.1371/journal.pdig.0000104

Castelnovo, A., Ferri, R., Punjabi, N. M., Castronovo, V., Garbazza, C., Zucconi, M., Ferini-Strambi, L., & Manconi, M. (2019). The paradox of paradoxical insomnia: A theoretical review towards a unifying evidence-based definition. *Sleep medicine reviews*, *44*, 70–82. https://doi.org/10.1016/j.smrv.2018.12.007

Cerin, E., Cain, K. L., Oyeyemi, A. L., Owen, N., Conway, T. L., Cochrane, T., VAN Dyck, D., Schipperijn, J., Mitáš, J., Toftager, M., Aguinaga-Ontoso, I., & Sallis, J. F. (2016). Correlates of Agreement between Accelerometry and Self-reported Physical Activity.

Medicine and science in sports and exercise, *48*(6), 1075–1084. https://doi.org/10.1249/MSS.0000000000000870

Choi, K. W., Zheutlin, A. B., Karlson, R. A., Wang, M. J., Dunn, E. C., Stein, M. B., Karlson, E. W., & Smoller, J. W. (2020). Physical activity offsets genetic risk for incident depression assessed via electronic health records in a biobank cohort study. *Depression and anxiety*, *37*(2), 106–114. https://doi.org/10.1002/da.22967

Cook, J. D., Prairie, M. L., & Plante, D. T. (2017). Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: A comparison against polysomnography and wrist-worn actigraphy. *Journal of affective disorders*, *217*, 299–305. https://doi.org/10.1016/j.jad.2017.04.030

Damme, K. S. F., Vargas, T. G., Walther, S., Shankman, S. A., & Mittal, V. A. (2024). Physical and mental health in adolescence: novel insights from a transdiagnostic examination of FitBit data in the ABCD study. *Translational psychiatry*, *14*(1), 75. https://doi.org/10.1038/s41398-024-02794-2

De Angel, V., Lewis, S., White, K., Oetzmann, C., Leightley, D., Oprea, E., Lavelle, G., Matcham, F., Pace, A., Mohr, D. C., Dobson, R., & Hotopf, M. (2022). Digital health tools for the passive monitoring of depression: a systematic review of methods. *NPJ digital medicine*, *5*(1), 3. https://doi.org/10.1038/s41746-021-00548-8

Dennison, C. A., Legge, S. E., Bracher-Smith, M., Menzies, G., Escott-Price, V., Smith, D. J., Doherty, A. R., Owen, M. J., O'Donovan, M. C., & Walters, J. T. R. (2021). Association of genetic liability for psychiatric disorders with accelerometer-assessed physical activity in the UK Biobank. *PloS one*, *16*(3), e0249189. https://doi.org/10.1371/journal.pone.0249189

Firth, J., Solmi, M., Wootton, R.E., Vancampfort, D., Schuch, F.B., Hoare, E., Gilbody, S., Torous, J., Teasdale, S.B., Jackson, S.E., Smith, L., Eaton, M., Jacka, F.N., Veronese, N., Marx, W., Ashdown-Franks, G., Siskind, D., Sarris, J., Rosenbaum, S., Carvalho, A.F. and Stubbs, B. (2020), A meta-review of "lifestyle psychiatry": the role of exercise, smoking, diet and sleep in the prevention and treatment of mental disorders. World Psychiatry, 19: 360-380. https://doi.org/10.1002/wps.20773

Fuller, D., Colwell, E., Low, J., Orychock, K., Tobin, M. A., Simango, B., Buote, R., Van Heerden, D., Luan, H., Cullen, K., Slade, L., & Taylor, N. G. A. (2020). Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review. *JMIR mHealth and uHealth*, *8*(9), e18694. https://doi.org/10.2196/18694

Ge, T., Chen, C. Y., Ni, Y., Feng, Y. A., & Smoller, J. W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature communications*, *10*(1), 1776. https://doi.org/10.1038/s41467-019-09718-5

George, M. J., Rivenbark, J. G., Russell, M. A., Ng'eno, L., Hoyle, R. H., & Odgers, C. L. (2019). Evaluating the Use of Commercially Available Wearable Wristbands to Capture Adolescents' Daily Sleep Duration. *Journal of research on adolescence : the official journal of the Society for Research on Adolescence*, *29*(3), 613–626. https://doi.org/10.1111/jora.12467

Gogarten, S. M., Sofer, T., Chen, H., Yu, C., Brody, J. A., Thornton, T. A., Rice, K. M., & Conomos, M. P. (2019). Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics (Oxford, England)*, *35*(24), 5346–5348. https://doi.org/10.1093/bioinformatics/btz567

Goldstone, A., Javitz, H. S., Claudatos, S. A., Buysse, D. J., Hasler, B. P., de Zambotti, M., Clark, D. B., Franzen, P. L., Prouty, D. E., Colrain, I. M., & Baker, F. C. (2020). Sleep Disturbance Predicts Depression Symptoms in Early Adolescence: Initial Findings From the Adolescent Brain Cognitive Development Study. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine*, *66*(5), 567–574. https://doi.org/10.1016/j.jadohealth.2019.12.005

Gunn, S., Wang, X., Posner, D. C., Cho, K., Huffman, J. E., Gaziano, M., Wilson, P. W., Sun, Y. V., Peloso, G., & Lunetta, K. L. (2025). Comparison of methods for building polygenic scores for diverse populations. *HGG advances*, *6*(1), 100355. https://doi.org/10.1016/j.xhgg.2024.100355

Haghayegh, S., Khoshnevis, S., Smolensky, M. H., Diller, K. R., & Castriotta, R. J. (2019). Accuracy of Wristband Fitbit Models in Assessing Sleep: Systematic Review and Meta-Analysis. *Journal of medical Internet research*, *21*(11), e16273. https://doi.org/10.2196/16273

Harvey, A. G., & Tang, N. K. (2012). (Mis)perception of sleep in insomnia: a puzzle and a resolution. *Psychological bulletin*, *138*(1), 77–101. https://doi.org/10.1037/a0025730

Hou, J., Deng, Q., Sha, L., Zhu, J., Xiang, R., Zhao, X., Xiao, C., Qin, C., Qu, Y., Han, T., Song, X., Yang, B., Yu, T., Zhou, J., Zheng, S., Fan, M., Yan, P., & Jiang, X. (2025). Physical activity and risk of depression in adolescents: A systematic review and meta-analysis of prospective observational studies. *Journal of affective disorders*, *371*, 279–288. https://doi.org/10.1016/j.jad.2024.11.065

Hyat, M., Zhu, J., Boltz, T. A., Conomos, M. P., Hughes, D. E., Fohner, A. E., Foster, K. T., Bigdeli, T. B., & Forsyth, J. K. (2025). Independent versus joint effects of polygenic or family-based schizophrenia risk in diverse ancestry youth in the ABCD study. *Psychological medicine*, *55*, e327. https://doi.org/10.1017/S0033291725102304

Johnson, D. A., Jackson, C. L., Williams, N. J., & Alcántara, C. (2019). Are sleep patterns influenced by race/ethnicity - a marker of relative advantage or disadvantage? Evidence to date. *Nature and science of sleep*, *11*, 79–95. https://doi.org/10.2147/NSS.S169312

Kang, S. G., Kang, J. M., Ko, K. P., Park, S. C., Mariani, S., & Weng, J. (2017). Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. *Journal of psychosomatic research*, *97*, 38–44. https://doi.org/10.1016/j.jpsychores.2017.03.009

Kawai, K., Iwamoto, K., Miyata, S., Okada, I., Fujishiro, H., Noda, A., Nakagome, K., Ozaki, N., & Ikeda, M. (2023). Comparison of Polysomnography, Single-Channel Electroencephalogram, Fitbit, and Sleep Logs in Patients With Psychiatric Disorders: Cross-Sectional Study. *Journal of medical Internet research*, *25*, e51336. https://doi.org/10.2196/51336

Kim, E. H., Jenness, J. L., Miller, A. B., Halabi, R., de Zambotti, M., Bagot, K. S., Baker, F. C., & Pratap, A. (2023). Association of Demographic and Socioeconomic Indicators With the Use of Wearable Devices Among Children. *JAMA network open*, *6*(3), e235681. https://doi.org/10.1001/jamanetworkopen.2023.5681

Kiss, O., Shaska, A., Müller-Oehring, E. M., Hasler, B. P., Franzen, P. L., FitzGerald, D. H., Clark, D. B., & Baker, F. C. (2025). Assessment of sleep measures and their agreement: youth-reported, caregiver-reported, and Fitbit-derived data in a large early adolescent cohort. *Sleep*, *48*(7), zsaf065. https://doi.org/10.1093/sleep/zsaf065

Lee, E., van Dijk, M. T., Kim, B. G., et al. (2025). Polygenic scores for psychiatric traits mediate the impact of multigenerational history for depression on offspring psychopathology. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-025-03221-8

Lovato, N., & Gradisar, M. (2014). A meta-analysis and model of the relationship between sleep and depression in adolescents: recommendations for future research and clinical practice. *Sleep medicine reviews*, *18*(6), 521–529. https://doi.org/10.1016/j.smrv.2014.03.006

Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Electronic address: andrew.mcintosh@ed.ac.uk, & Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium (2025). Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. *Cell*, *188*(3), 640–652.e9. https://doi.org/10.1016/j.cell.2024.12.002

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature genetics*, *51*(4), 584–591. https://doi.org/10.1038/s41588-019-0379-x

Messman, B. A., Wiley, J. F., Feldman, E., Dietch, J. R., Taylor, D. J., & Slavish, D. C. (2024). Irregular sleep is linked to poorer mental health: A pooled analysis of eight studies. *Sleep health*, *10*(4), 493–499. https://doi.org/10.1016/j.sleh.2024.03.004

Nelson, B. W., Flannery, J. E., Flournoy, J., Duell, N., Prinstein, M. J., & Telzer, E. (2022). Concurrent and prospective associations between fitbit wearable-derived RDoC arousal and regulatory constructs and adolescent internalizing symptoms. *Journal of child psychology and psychiatry, and allied disciplines*, *63*(3), 282–295. https://doi.org/10.1111/jcpp.13471

Ohi, K., Ochi, R., Noda, Y., Wada, M., Sugiyama, S., Nishi, A., Shioiri, T., Mimura, M., & Nakajima, S. (2021). Polygenic risk scores for major psychiatric and neurodevelopmental disorders contribute to sleep disturbance in childhood: Adolescent Brain Cognitive Development (ABCD) Study. *Translational psychiatry*, *11*(1), 187. https://doi.org/10.1038/s41398-021-01308-8

Pain, O., Al-Chalabi, A., & Lewis, C. M. (2024). The GenoPred pipeline: a comprehensive and scalable pipeline for polygenic scoring. *Bioinformatics (Oxford, England)*, *40*(10), btae551. https://doi.org/10.1093/bioinformatics/btae551

Poland, R. E., Rao, U., Lutchmansingh, P., McCracken, J. T., Lesser, I. M., Edwards, C., Ott, G. E., & Lin, K. M. (1999). REM sleep in depression is influenced by ethnicity. *Psychiatry research*, *88*(2), 95–105. https://doi.org/10.1016/s0165-1781(99)00080-3

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Connor Gorber, S., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *The international journal of behavioral nutrition and physical activity*, *5*, 56. https://doi.org/10.1186/1479-5868-5-56

Rao, U., Hammen, C. L., & Poland, R. E. (2009). Ethnic differences in electroencephalographic sleep patterns in adolescents. *Asian journal of psychiatry*, *2*(1), 17–24. https://doi.org/10.1016/j.ajp.2008.12.003

Rice, F., Riglin, L., Lomax, T., Souter, E., Potter, R., Smith, D. J., Thapar, A. K., & Thapar, A. (2019). Adolescent and adult differences in major depression symptom profiles. *Journal of affective disorders*, *243*, 175–181. https://doi.org/10.1016/j.jad.2018.09.015

Rohani, D. A., Faurholt-Jepsen, M., Kessing, L. V., & Bardram, J. E. (2018). Correlations Between Objective Behavioral Features Collected From Mobile and Wearable Devices

and Depressive Mood Symptoms in Patients With Affective Disorders: Systematic Review. *JMIR mHealth and uHealth*, *6*(8), e165. https://doi.org/10.2196/mhealth.9691

Ruan, Y., Lin, Y. F., Feng, Y. A., Chen, C. Y., Lam, M., Guo, Z., Stanley Global Asia Initiatives, He, L., Sawa, A., Martin, A. R., Qin, S., Huang, H., & Ge, T. (2022). Improving polygenic prediction in ancestrally diverse populations. *Nature genetics*, *54*(5), 573–580. https://doi.org/10.1038/s41588-022-01054-7

Sallis, J. F., Zakarian, J. M., Hovell, M. F., & Hofstetter, C. R. (1996). Ethnic, socioeconomic, and sex differences in physical activity among adolescents. *Journal of clinical epidemiology*, *49*(2), 125–134. https://doi.org/10.1016/0895-4356(95)00514-5

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: review and meta-analysis. *The American journal of psychiatry*, *157*(10), 1552–1562. https://doi.org/10.1176/appi.ajp.157.10.1552

Wainberg, M., Jones, S. E., Beaupre, L. M., Hill, S. L., Felsky, D., Rivas, M. A., Lim, A. S. P., Ollila, H. M., & Tripathy, S. J. (2021). Association of accelerometer-derived sleep measures with lifetime psychiatric diagnoses: A cross-sectional study of 89,205 participants from the UK Biobank. *PLoS medicine*, *18*(10), e1003782. https://doi.org/10.1371/journal.pmed.1003782

Wang, C. H., & Peiper, N. (2022). Association Between Physical Activity and Sedentary Behavior With Depressive Symptoms Among US High School Students, 2019. *Preventing chronic disease*, *19*, E76. https://doi.org/10.5888/pcd19.220003

Wang, Y., Kanai, M., Tan, T., Kamariza, M., Tsuo, K., Yuan, K., Zhou, W., Okada, Y., BioBank Japan Project, Huang, H., Turley, P., Atkinson, E. G., & Martin, A. R. (2023). Polygenic prediction across populations is influenced by ancestry, genetic architecture, and methodology. *Cell genomics*, *3*(10), 100408. https://doi.org/10.1016/j.xgen.2023.100408

Wescott, D. L., Morash-Conway, J., Zwicker, A., Cumby, J., Uher, R., & Rusak, B. (2019). Sleep in Offspring of Parents With Mood Disorders. *Frontiers in psychiatry*, *10*, 225. https://doi.org/10.3389/fpsyt.2019.00225

Wing, D., Godino, J. G., Baker, F. C., Yang, R., Chevance, G., Thompson, W. K., Reuter, C., Bartsch, H., Wilbur, A., Straub, L. K., Castro, N., Higgins, M., Colrain, I. M., de Zambotti, M., Wade, N. E., Lisdahl, K. M., Squeglia, L. M., Ortigara, J., Fuemmeler, B., Patrick, K., … Bagot, K. S. (2022). Recommendations for Identifying Valid Wear for Consumer-Level Wrist-Worn Activity Trackers and Acceptability of Extended Device Deployment in Children. *Sensors (Basel, Switzerland)*, *22*(23), 9189. https://doi.org/10.3390/s22239189

Wray, N. R., Ripke, S., Mattheisen, M., et al. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*, 668–681. https://doi.org/10.1038/s41588-018-0090-3

Xu, E. Y., Grimes, P. Z., Kwong, A. S. F., Lawrie, S. M., & Whalley, H. C. (2025). Longitudinal Prediction of Adolescent Depression from Environmental and Polygenic Risk Scores. *medRxiv : the preprint server for health sciences*, 2025.07.08.25331098. https://doi.org/10.1101/2025.07.08.25331098

Zhang, Y., Folarin, A. A., Sun, S., Cummins, N., Bendayan, R., Ranjan, Y., Rashid, Z., Conde, P., Stewart, C., Laiou, P., Matcham, F., White, K. M., Lamers, F., Siddi, S., Simblett, S., Myin-Germeys, I., Rintala, A., Wykes, T., Haro, J. M., Penninx, B. W., … RADAR-CNS Consortium (2021). Relationship Between Major Depression Symptom Severity and Sleep Collected Using a Wristband Wearable Device: Multicenter Longitudinal

Observational Study. *JMIR mHealth and uHealth*, *9*(4), e24604. https://doi.org/10.2196/24604

Zhao, Y., Yang, L., Sahakian, B. J., Langley, C., Zhang, W., Kuo, K., Li, Z., Gan, Y., Li, Y., Zhao, Y., Yu, J., Feng, J., & Cheng, W. (2023). The brain structure, immunometabolic and genetic mechanisms underlying the association between lifestyle and depression. Nature Mental Health, 1, Article 10. https://doi.org/10.1038/s44220-023-00120-1

Zheng, Z., Liu, S., Sidorenko, J., Wang, Y., Lin, T., Yengo, L., Turley, P., Ani, A., Wang, R., Nolte, I. M., Snieder, H., LifeLines Cohort Study, Yang, J., Wray, N. R., Goddard, M. E., Visscher, P. M., & Zeng, J. (2024). Leveraging functional genomic annotations and genome coverage to improve polygenic prediction of complex traits within and between ancestries. *Nature genetics*, *56*(5), 767–777. https://doi.org/10.1038/s41588-024-01704-y

Zierer, C., Behrendt, C., & Lepach-Engelhardt, A. C. (2024). Digital biomarkers in depression: A systematic review and call for standardization and harmonization of feature engineering. *Journal of affective disorders*, *356*, 438–449. https://doi.org/10.1016/j.jad.2024.03.163