

https://dspace.org/

Using explainable machine learning and fitbit data to investigate predictors of adolescent obesity

O Kiss; Fiona Baker; Robert Palovics; Erin E Dooley; K Pettee Gabriel ...

https://hdl.handle.net/10539/41292
WOS

Downloaded from a DSpace repository. https://dspace.org/

scientific reports

Check for updates

OPEN

Using explainable machine learning and fitbit data to investigate predictors of adolescent obesity

Orsolya Kiss[1]✉, Fiona C. Baker[1,5], Robert Palovics[3], Erin E. Dooley[4], Kelley Pettee Gabriel[4] & Jason M. Nagata[2]

Sociodemographic and lifestyle factors (sleep, physical activity, and sedentary behavior) may predict obesity risk in early adolescence; a critical period during the life course. Analyzing data from 2971 participants (M = 11.94, SD = 0.64 years) wearing Fitbit Charge HR 2 devices in the Adolescent Brain Cognitive Development (ABCD) Study, glass box machine learning models identified obesity predictors from Fitbit-derived measures of sleep, cardiovascular fitness, and sociodemographic status. Key predictors of obesity include identifying as Non-White race, low household income, later bedtime, short sleep duration, variable sleep timing, low daily step counts, and high heart rates (AUC_{Mean} = 0.726). Findings highlight the importance of inadequate sleep, physical inactivity, and socioeconomic disparities, for obesity risk. Results also show the clinical applicability of wearables for continuous monitoring of sleep and cardiovascular fitness in adolescents. Identifying the tipping points in the predictors of obesity risk can inform interventions and treatment strategies to reduce obesity rates in adolescents.

Obesity is a significant public health problem for adolescents in the United States[1] with differences noted by race/ethnicity and socioeconomic status[1]. The prevalence of obesity in adolescence, which frequently persists into adulthood[2,3] has significantly increased in past decades. One in six adolescents is classified as having obesity and one-third are at risk for obesity[4] Early adolescence (10–14 years) is a particularly high-risk period for the development of obesity[5] and carries a high risk for adult obesity[2,3] and comorbid conditions, including earlier age onset[6]. Adolescent obesity also has immediate consequences, including advanced pubertal onset[7], menstrual irregularities in girls[8], and cardiovascular risk factors, including prediabetes, type 2 diabetes, high cholesterol levels, hypertension, and metabolic syndrome[9]. Additionally, adolescents with obesity may experience weight stigma and are at risk for psychological problems such as depression[10], anxiety[11], poor self-esteem and body image, poor peer relationships, and eating disorders[12]. The rarity of transitions from obesity to normal weight during adolescence[13] shows that without targeted efforts, obesity becomes a persistent condition with long-term detrimental consequences.

Wearable devices, such as fitness trackers, smartwatches, and other health monitoring devices, have gained popularity among various age groups, including adolescents[14]. Wearable devices can record sleep, physical activity, and physiological measures (e.g., heart rate) that may be associated with obesity. Recently, several studies documented the use and efficacy of wearable devices to prevent and treat obesity in adults[15–17].

Sleep problems are now recognized as being associated with obesity risk in adolescence[18]. Sleep is crucial for energy restorative processes and metabolism, mediating glucose metabolism[19], influencing food intake[20,21], and appetite regulation. Importantly, sleep changes considerably in adolescence, driven by biological and environmental/social factors leading to later bedtimes while still needing to get up early for school[22]. Consequently, youth are at risk for poor/insufficient sleep, especially those with high socioeconomic disparities[23,24] and there is a growing number of adolescents in the US who are not getting sufficient, good quality sleep. With the rise of wearables, there's an added advantage in continuously detecting and tracking measurable behaviors, such as sleep and movement patterns and indices of cardiovascular (CV) fitness, which could be pivotal risk factors for obesity in adolescents. These devices provide an objective assessment of physical activity and sleep that addresses the limitations associated with self-report methods, including records, recalls, or questionnaires.

[1]Center for Health Sciences, SRI International, 333 Ravenswood Ave, Menlo Park, CA 94025, USA. [2]Department of Pediatrics, University of California, San Francisco, San Francisco, CA, USA. [3]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA. [4]Department of Epidemiology, University of Alabama at Birmingham, 1665 University Boulevard, Birmingham, AL 35233, USA. [5]School of Physiology, University of the Witwatersrand, Parktown, Johannesburg, South Africa. ✉email: orsolya.kiss@sri.com

Physical inactivity in adolescents is another risk factor for obesity[25], and overweight and obese youth are more sedentary and less physically active than their peers with healthy weight[26]. Increasing physical activity is considered an important strategy to increase physical activity related energy expenditure[27], making exercise one of the most common strategies in the treatment and prevention of obesity[28]. In association with physical inactivity and obesity, heart rate variability tends to be lower and cardiovascular function is altered[29] since the heart has to work harder to pump blood to the excess fat tissue, which requires more oxygen and nutrients[30]. As a result, the heart rate increases to compensate for this extra workload[31].

Here we leverage the Adolescent Brain Cognitive Development (ABCD) study®, which collects data from a diverse adolescent population across the United States, to apply a data-driven approach relying on data collected with commercially available wearable Fitbit devices to identify the most important measurable predictors of obesity and investigate the non-linear associations of these predictors with obesity. This analysis allows us to pinpoint the tipping points in the predictors where the risk of obesity begins to escalate, shedding light on the complex interplay between lifestyle factors and health outcomes. In addition to sleep, the number of steps, cardiovascular fitness measures, and their day-to-day variability, we consider sociodemographic measures, shown to be relevant in the prediction of obesity[32]. We hypothesize that models trained primarily on wearable-derived measures will accurately predict adolescent obesity.

Methods

Participants

Data were obtained from the US-based, multi-site ABCD Study® (data version 5.0 released in June 2023), (21 research sites from 17 states—https://abcdstudy.org/study-sites/) of ~ 11,800 children aged 9–10 years at baseline[33]. Full details on the ABCD Study protocols, design, and recruitment procedures may be found elsewhere[33,34]. We included data from *2*,971 adolescents from the Year 2 assessment who had anthropometric data, and who provided data from the second and third weeks of the Fitbit protocol. As part of the preprocessing, we excluded participants who did not have at least two weekend and two weekday sleep data and participants with incorrect values (e.g., negative time in bed values). Written informed consent and assent were obtained from a parent/guardian and the adolescent, respectively. Procedures were approved by a centralized institutional review board (University of California, San Diego). All procedures complied with the Declaration of Helsinki. Table 1 provides a detailed breakdown of the demographics for our analytical sample. When compared to the ABCD sample of participants who were excluded due to insufficient data, the analytical sample consisted of slightly younger participants, with an average age of 11.94 years (SD = 0.64) as opposed to 12.02 years (SD = 0.66). The sex distribution was also marginally different, with females making up 49% of the analytical sample versus 47% of the excluded group. The analytical sample had a higher representation of White individuals (62.5% vs. 48.4% in the excluded group) and Asians (7.1% vs. 5.5% in the excluded group), while it had fewer Latinos (16.4% vs. 17.3%) and significantly fewer Black individuals (9.5% vs. 23.6%). Socioeconomically, both the analytical sample and the excluded group reported similar household incomes below $75,000, at 31.3% and 31.0% respectively.

Variable	ABCD Fitbit sub-sample N = 2971
Age (years) mean (range)	11.94 (10–14)
Sex	
Female	1470 (49.4%)
Male	1501 (50.5%)
Race/ethnicity	
Asian	212 (7.1%)
Black	287 (9.6%)
Latino	488 (16.4%)
Native American	91 (3.0%)
Other	23 (0.7%)
White	1859 (62.5%)
Unknown/not reported	11 (0.03%)
Household income	
Below $75,000 household income	930 (31.3%)
$75,000 or above household income	1857 (62.5%)
Unknown/not reported	184 (6.1%)
Parental education	
< High school diploma	240 (8.0%)
College or above	2726 (91.4%)
Unknown/not reported	5 (< 0.1%)

Table 1. Demographics collected at baseline for the subset of ABCD Study participants included in the current analysis (Fitbit sub-sample).

As for parental education, only 8% of the analytical sample had parents with a high school education or less, compared to 12.9% in the excluded group.

Input measures included in the models

Demographics

We included the following caregiver reported sociodemographic information from the baseline year, collected between 2016 and 2018: Biological sex (female/male), sex and gender minority status (yes/no), race/ethnicity (White, Latino, Black, Asian, Native American, other race), number of siblings, household income (high income/low income as defined based on the household income, which approximated the median US household income), highest educational attainment of parents/caregiver (college or above/high school or less), immigration status ("anyone in your child's family born outside of the United States"—yes/no), marital status of the caregiver (married/ not married, where married and living with a partner were combined; widowed, divorced, separated, and never married were categorized as not married/not living with a partner), access to food, telephone service, gas—electric service (e.g., "needed food but couldn't afford to buy it or couldn't afford to go out to get it"—yes/no), housing, eviction (e.g., "didn't pay the full amount of the rent or mortgage because you could not afford it"—yes/no), heath care—doctor ("needed to see a doctor or go to the hospital but didn't go because you could not afford it"—yes/no) and number of people living at the same address. Age collected in months at the Year 2 assessment was converted to years.

Medical history. As part of the Year 2 assessment, caregivers reported their child's recent medical visits and conditions. We included items about hospitalization, emergency room visits, medical visits, and medical conditions (sickle cell anemia, problems with the heart, kidney disease, diabetes, cerebral palsy, asthma).

Fitbit measures. At the Year 2 assessment (November 2018 and November 2020), ABCD Study participants were invited to wear a Fitbit Charge HR 2 (Fitbit Inc., San Francisco, CA) on their wrist for up to 21 days[35] (for more detailed description see Supplementary Section 1). Fitbit has been validated against research-grade, gold-standard devices in children and adolescents[36]. Fitbit Charge HR 2 devices use proprietary algorithms to continuously measure biobehavioral features at up to a one-second sampling rate[35] with a photoplethysmography and accelerometer. We utilized the daily aggregated physical activity and sleep measures, as it was released by the ABCD Data Management team (see Supplementary Section 1 for more information). Within each adolescent's 21-day study protocol, we excluded the first week (7 days) of data from our analysis of weekday variability in sleep and physical activity measures. This exclusion considered the influence of self-monitoring[37] or the awareness of being observed[38], which are known to alter behavior in response to being monitored. Our decision was based on our observation that during the first week, participants tended to have higher levels of physical activity compared to the subsequent weeks (Fig. S3, and Supplementary Section 2). We hypothesized that this may have been due to factors such as excitement or novelty associated with wearing the Fitbit device, as well as an increased awareness of being observed, which often leads individuals to modify their behavior in the short term. By focusing on data from the second and third weeks, we aimed to capture more habitual and stable patterns of behavior. This approach improves the accuracy of our estimates of weekday variability in sleep and cardiovascular fitness and physical activity measures, offering insights into adolescents' habitual behaviors rather than their temporary adjustments to new circumstances. In doing so, we mitigate the potential distortions in data caused by the initial reaction to the novelty of the device and the consciousness of being part of a study, ensuring that our analysis reflects participants' regular patterns of physical activity[39,40] and sleep[41,42], that is relevant in the prediction of obesity, and has less temporal variations. From the remaining Fitbit data, we included summary scores of sleep, activity, and cardiovascular fitness measures (mean) and calculated variability in all daily measures using data from the second and third weeks (standard deviation, difference between the minimum and maximum value for weekdays and weekends, and mean square successive difference[43] on weekdays).

Sleep Variables and variability measures included sleep timing (bedtime, wake-up time), sleep duration, sleep onset latency, sleep inertia (time taken to get out of bed) on weekdays and weekend days separately, and mid-sleep timing. In our study, mid-sleep is defined as the midpoint of the sleep cycle, calculated using the formula: Mid-Sleep = Sleep Onset Time + (Sleep Duration / 2), where Sleep Onset Time is the time at which the subject falls asleep, and Sleep Duration is the total time spent asleep, thereby providing a quantitative measure to assess individual chronotypes and natural sleep patterns. In addition, we calculated the sleep efficiency (the ratio of the total time spent asleep to the total amount of time spent in bed intended for sleep) and the relative and absolute social jetlag.

Physical activity—steps per day Prior research has demonstrated that Fitbit devices offer an accurate and reliable estimation of daily step count as a measure of accumulated physical activity in adolescents over time[35,36]. All days (weekdays and weekends separately) with a minimum of 1000 steps per day (and > 600 min of wear time) were included, in line with prior studies[44,45].

Cardiovascular fitness measures (Fitbit) Data included sleeping heart rate (HR) and daytime resting state HR, averaged across the second and third week, as well as day-to-day variability in these measures.

COVID-19. Given that the COVID-19 pandemic occurred during some of the Fitbit data collection period, we created a COVID data collection variable (yes/no) based on dates of wear, defining March 13, 2020, as the onset of the COVID-19 pandemic in the U.S.

Outcome measure—obesity

At the Year 2 Assessment, weight (Health-o-meter 844KL High-Capacity Digital Bathroom Scale; Jarden Corporation; Rye, NY) and height (Carpenter's square, steel tape measure) were assessed by the interviewer at each site. Body mass index (BMI) was computed using the standard formula, weight (kilograms) divided by height (meters) squared ($BMI = weight/height^2$). BMI was converted into sex- and age-specific percentiles in accordance with the Centers for Disease Control and Prevention (CDC) growth curves and definitions. Individuals were classified as having obesity if their BMI was ≥ 95th percentile[46].

Training and evaluation of classification models

During the feature selection process, we removed features having missing records above 98% and dropped quasi-constant (> 99%) features from the dataset. Even though Explainable Boosting Machine (EBM) algorithms, which were employed in our study, inherently possess a robustness to multicollinearity, we took steps to reduce collinearity among variables with the aim of simplifying the model and enhancing its interpretability. To identify redundant features and reduce collinearity among variables, we calculated the Spearman's rho (ρ) correlation matrix and identified pairs of variables with large correlation values ($|\rho| > 0.75$). We compared these features with the outcome variable and kept the feature that showed a higher correlation with the outcome in the model. Missing items for predictor variables were replaced using k-Nearest Neighbors[47] (implemented in the Python package sklearn.impute.KNNImputer with uniform weights). Figures were generated using the seaborn (Version 0.11) Python (Version 3.8.5) package[48].

Model performance depends on the selection of their hyper-parameters (e.g., learning rate). To find the best performing hyper-parameter configuration, we performed a grid search by evaluating all possible configurations in a predefined range.

As our primary method, we employed the Explainable Boosting Machine (EBM)[49] (implemented in the Python package InterpretML, 0.5.1), a non-linear machine learning model (Fig. 1) that can incorporate interactions among features. This ability is particularly useful when existing literature indicates potential interactions among the predictors, allowing us to construct a more informed and context-aware models. To identify the most important two-way interactions in the data, we included the number of interactions into the parameter search. For model evaluation, we used area under the receiver operating characteristic curve (AUC) metrics and accuracy (ACC). The AUC is a widely recommended measure for assessing models on unbalanced datasets due to its insensitivity to class distribution[50–52].

To assess the performance of the EBM model in comparison to both linear and nonlinear models, we separately trained Regularized Logistic Regression with L1 (Lasso) regularization (referred to as LassoReg) and Gradient Boosted Trees (GBT). A comprehensive search across several potential parameters was conducted to identify the best-fitting model for targeting obesity.

For model comparison, we used a 2×5 cross-validation technique. We fitted the final model on the entire dataset using the implementation in scikit-learn[53] (version 0.18.2) and Python 3.6.8.

Feature importance in final models

We analyzed the global feature importance values of the fine-tuned EBM model that we trained on the entire dataset (lr = 0.005, max_round = 10, max_bin = 16, number of interactions = 5). We ranked the features according to their importance (absolute value of the coefficients) to interpret the best-performing models. For the sake of interpretability, we categorized the top features into eight domains: demographics (e.g., age, sex); cardiovascular health (e.g., HR); physical activity (e.g., step counts); medical history (e.g., during the past year seeing the doctor other than for regular checkups); and sleep (e.g., sleep duration).

Implications and contribution

Using Explainable Boosting Machine algorithms, this study underscores the efficacy of wearable devices in predicting early adolescent obesity through sleep, physical activity, and socioeconomic factors. The wearables' convenience for health monitoring offers potential for early obesity intervention and promoting healthier behaviors, especially among socioeconomically disadvantaged adolescents.

Results

In our analysis, the Explainable Boosting Machine algorithm illuminated the intricate relationships between the predictors and adolescent obesity, including key interactions among sociodemographic factors, sleep, physical activity, and cardiovascular fitness highlighted in Fig. 3. Figures 5 and Fig. S2 underscore the non-linear nature of these associations, revealing specific values and patterns that contribute to varying degrees of risk. This granularity in interpretation, made possible by the EBM's adeptness at handling non-linear relationships, offers a nuanced understanding of the intertwined roles of sociodemographic factors, sleep, physical activity and cardiovascular fitness in the context of adolescent obesity.

The proportion of participants classified as having obesity in the Fitbit sub-sample analyzed here was lower (13.5%) than in the whole ABCD cohort at the Year 2 assessment (17.1%, Fig. 2). Results of the cross-validation indicated acceptable prediction performance for all our models; however, the EBM performed well both on the cross validation (mean AUC ≥ 0.726) and in the fine-tuned model (AUC ≥ 0.737) (Table 2).

Feature importance in the classification model

The final EBM model included 69 features. Figure 3 shows the top 20 features by mean absolute contribution score (C). The model identified five important interactions influencing obesity risk. The interaction between White race and overnight heart rate emerged as the top-ranked contribution score ($C = 0.019$) within the model.

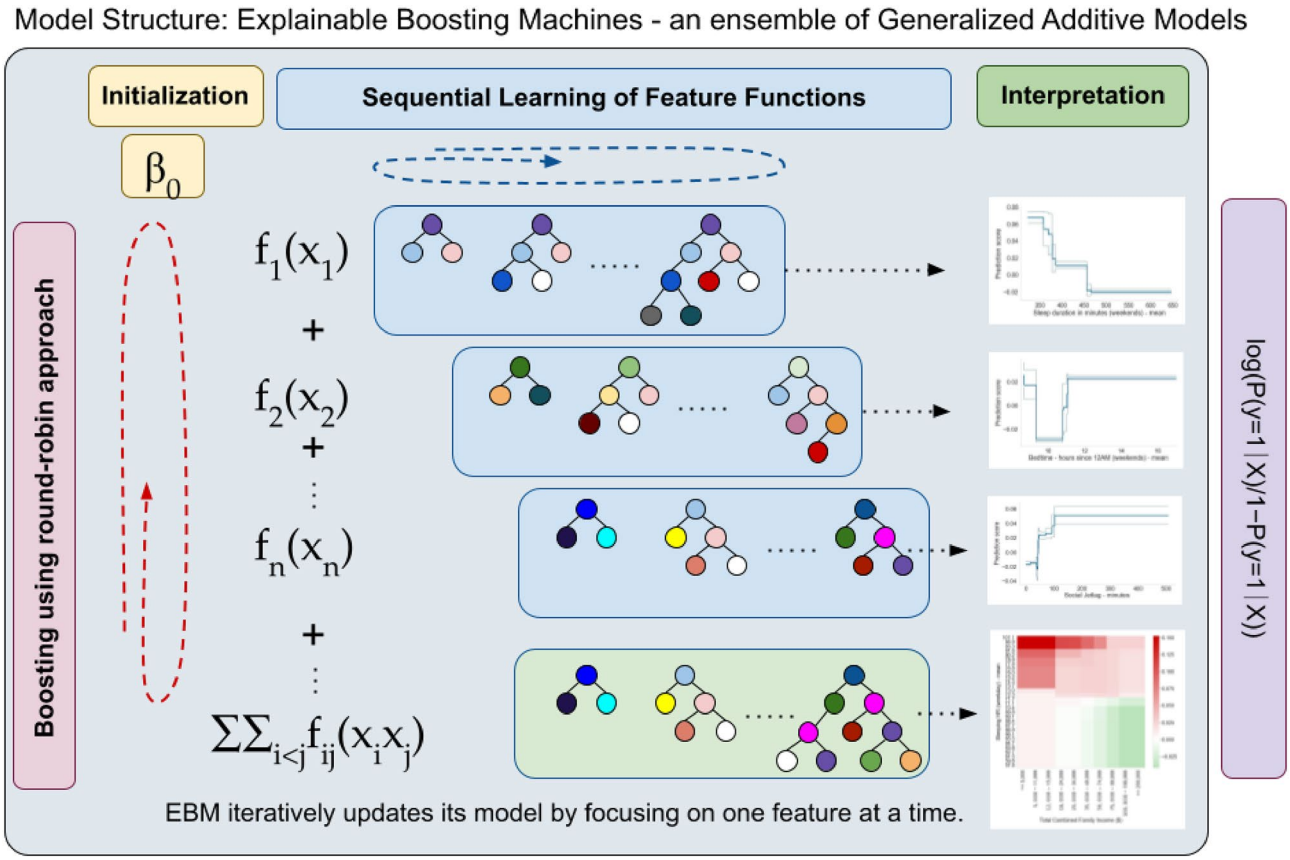


Figure 1. This figure illustrates the iterative process of building an Explainable Boosting Machine (EBM) for binary classification. Starting with a basic model, the EBM sequentially updates its predictions by cycling through each feature to learn its unique contribution via gradient boosting. This round-robin approach ensures gradual refinement, where each feature's effect is independently modeled and then aggregated to form the final predictive model. Interactions between pairs of features are also explored and integrated, enhancing the model's accuracy. The process emphasizes continuous learning from residuals, leading to a highly interpretable and precise predictive model that offers clear insights into the influence of individual features and their interactions on the outcome. P is the probability of the target variable belonging to the positive class (obesity), β0 is the intercept; fi (xi) represents the smooth function for the i-th feature; xi is the i-th feature; n is the total number of features[82]. Figures are just for illustration, generated using the seaborn (Version 0.11, URL: https://seaborn.pydata.org/generated/seaborn.heatmap.html) Python (Version 3.8.5) package[48].

This interaction reveals an elevated risk of obesity among individuals who identified as any race other than White who also have a sleeping HR greater than 72 BPM, and a corresponding increase in obesity risk among White individuals associated with a heart rate of 80 BPM or higher (Fig. 4A). Further interactions included White race with Resting HR ($C = 0.018$) and Black race with Sleeping HR ($C = 0.017$), both highlighting race-specific variations in heart rate associated with obesity risk (Fig. 4C-D). Another key interaction was between sleeping heart rate and economic hardship, measured by difficulties in affording food ($C = 0.019$). Participants facing food access issues with higher sleeping heart rates were at greater risk for obesity (Fig. 4B). This finding aligns with another interaction between Sleeping HR and Total Combined Family Income ($C = 0.016$), revealing nuanced risk profiles. Specifically, participants from families with incomes ranging between $5,000 or lower and $16,000 who also had elevated sleeping HR were found to be at the highest risk for obesity. Conversely, having a family income above $100,000 and a lower sleeping HR was associated with a decreased risk of obesity (Fig. 4E).

Both White race ($C = 0.014$) and household income ($C = 0.012$) were identified as important predictors in our analysis. Having a high sleeping heart rate ($C = 0.016$) and a high daytime resting heart rate ($C = 0.013$) were also in the top 20 factors associated with obesity. Additionally, physical activity measures represented by lower weekend ($C = 0.011$) and weekday step count ($C = 0.009$) were also associated with obesity. Notably, adolescents with weekend step counts below 5,000 were found to have an elevated risk, and those who achieved at least 15,000 steps per day exhibited a relatively low risk for obesity (Fig. 5).

Sleep variables, comprising 8 out of the top-ranked variables, were strong predictors of obesity risk. Factors such as late bedtime (after 12 AM) ($C = 0.012$), higher social jetlag ($C_{Relative\ Social\ Jetlag} = 0.010$, $C_{Absolute\ Social\ Jetlag} = 0.010$), shorter sleep duration (< 6 h) on weekends ($C = 0.009$) and on weekdays (<7 hours, $C = 0.008$), higher variability in sleep onset on weekdays ($C = 0.009$), higher variability in mid-sleep timing ($C = 0.010$), and higher variability in sleep end on weekends ($C = 0.008$) were all associated with an elevated risk for obesity (Fig. S2). Conversely,

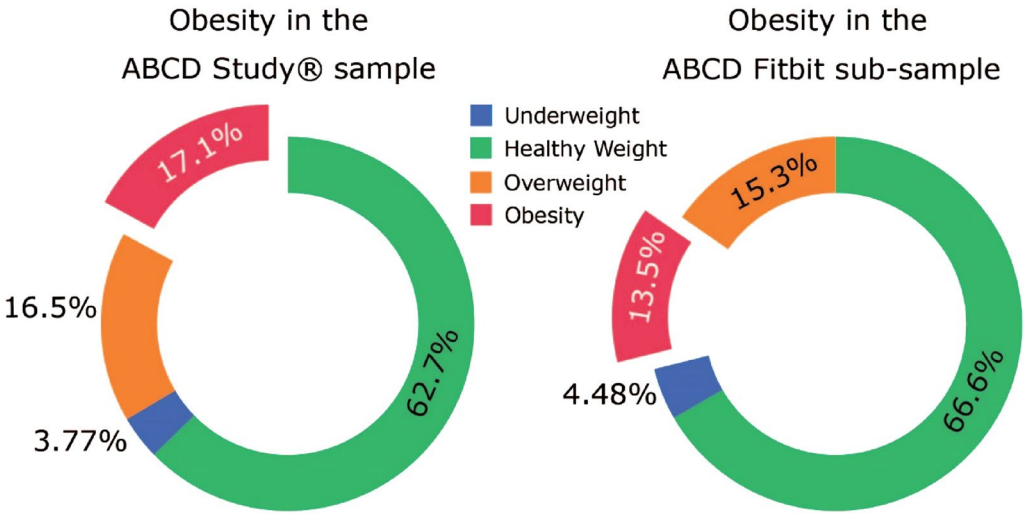


Figure 2. Body mass index distribution for the complete ABCD Study sample at the Year 2 assessment (N = 7552) and for the sub-sample of participants in the analytical sample with Fitbit and weight data (n = 2971) based on the CDC guidelines for BMI percentiles (Underweight: less than the 5th percentile; Healthy Weight: 5th percentile to less than the 85th percentile; Overweight: 85th percentile to less than the 95th percentile; Obesity: 95th percentile or greater).

	Evaluation	LassoReg (Lr = 0.08)	Gradient boosted Tree (n_tree = 60, lr = 0.1, depth = 2)	Explainable boosting machine (max_round = 10, Lr = 0.005, max_bin = 16, interaction = 5)
2 × 5 fold cross validation mean performance	AUC	0.735	0.708	0.726
Best performing model	AUC	0.684	0.664	0.737
	ACC	0.86	0.65	0.86

Table 2. Results of the cross-validation and the hyperparameters for the final fine-tuned models predicting obesity in the ABCD sub-sample of adolescents with Fitbit and BMI data (N = 2971).

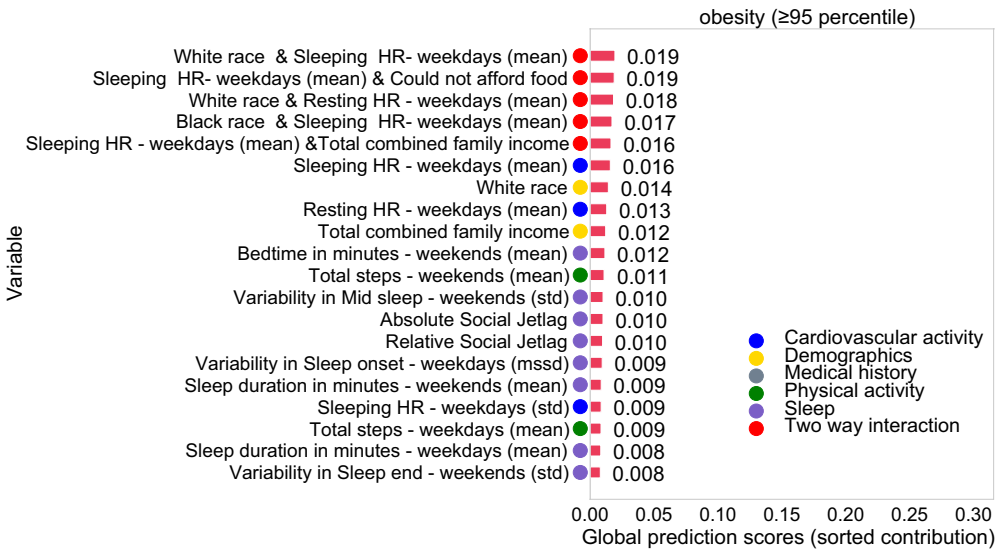


Figure 3. Mean absolute contribution score of each term (feature or interaction) in the best performing EBM model trained to predict obesity in young adolescents (n = 2971). The top 20 features are sorted by their contribution (C_i), where C_i represents the contribution score for the i-th feature to the prediction. These scores are expressed in terms of log odds, derived from a combination of univariate models for each feature, considering their individual contributions as well as interactions with other features.

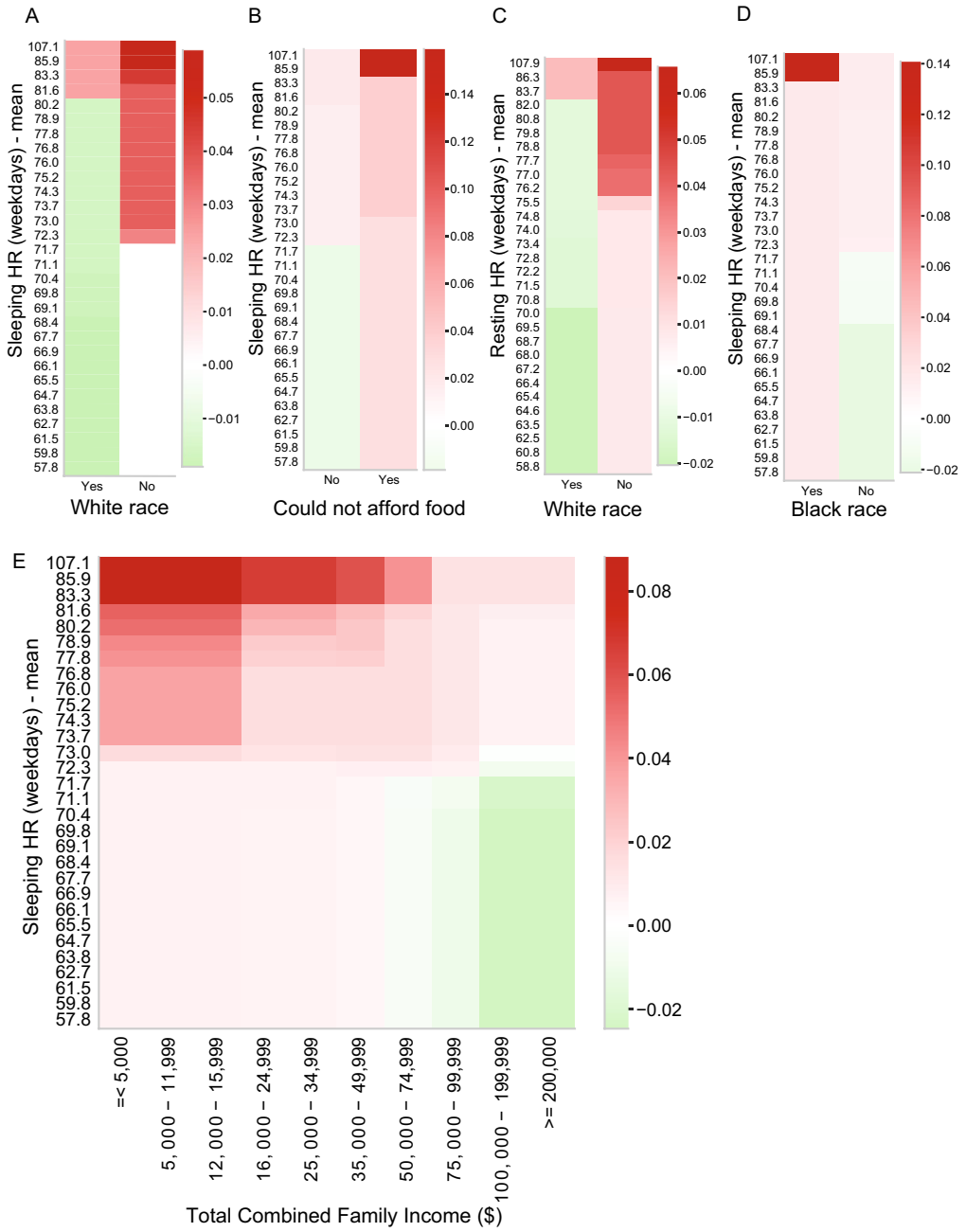


Figure 4. Interaction terms in the EBM model: two-way interaction between the Sleeping heart rate and White race (**A**), Sleeping heart rate and food access (**B**), White race and Resting heart rate (**C**), Black race and Sleeping heart rate (**D**) and the two-way interaction between the Sleeping heart rate and Household income (**E**) in the prediction of obesity. The color of the heatmap represents the direction of the effect (red: values associated with higher prediction score and increased risk, green: values associated with decreased risk). Figures were generated using the seaborn (Version 0.11, URL: https://seaborn.pydata.org/generated/seaborn.heatmap.html) Python (Version 3.8.5) package[48].

longer sleep duration (> 7.5 h) emerged as a protective factor, underscoring the multifaceted relationship between sleep patterns and obesity (Fig. 5).

Discussion

Utilizing a nonlinear approach, we identified crucial tipping points in the predictors where specific behaviors began to put adolescents at heightened risk for obesity in this demographically diverse sample of U.S. 10–14-year-olds. We found that obesity was relatively common (~ 14% in the Fitbit subgroup) and was accurately detected in a machine-learning model with wearable-based measures of short sleep duration, low step counts, and high

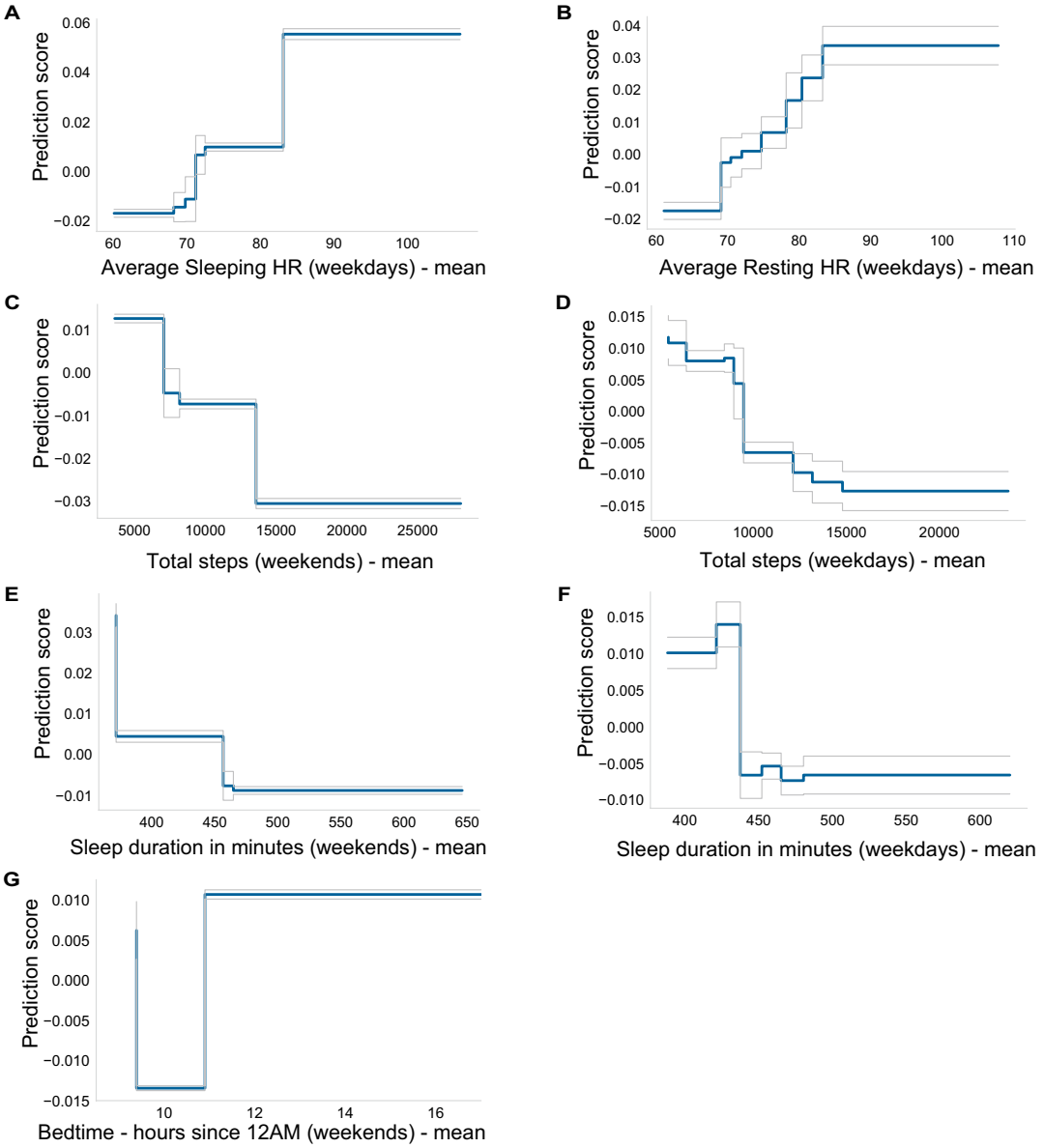


Figure 5. Contribution to the prediction (C) (blue line) for the predictors with the highest contribution in the best performing EBM: (**A**) Average Sleeping heart rate—weekdays (HR), (**B**) Average Resting heart rate—weekdays (HR). (**C**) Total step count—weekends, D. Total step count—weekdays, (**E**) Sleep duration in minutes—weekends, (**F**) Sleep duration in minutes—weekdays. (**G**) Bedtime in hours—weekends. The grey lines represent the standard error. These graphs showcase the nuanced effect of each variable on obesity prediction, highlighting zones where their influence shifts from being protective to risky, or vice versa, allowing us to identify tipping points.

heart rate together with sociodemographic indicators of disparities. Results indicate the value of using wearables over multiple days to determine associations of healthy behaviors like sleep and physical activity with obesity, and the relevance of sociodemographic factors in adolescents.

Our results are consistent with prior research that has shown racial/ethnic disparities in adolescent obesity risk, with Black and Hispanic/Latinx adolescents at higher risk[54–56]. In a nationally representative U.S. cohort, Rendall et al.[56] investigated the distinct obesity trajectories between Hispanic and Black children and reported that one in four Hispanic and Black children was estimated to have obesity by 8th grade, and Hispanic and Black children with overweight and obesity are less likely to return to normal weight levels than are White children with overweight and obesity. In addition, consistent with prior research[57], we found that low household income and difficulties with food access were associated with an increased risk of obesity, highlighting the persistent socio-economic factors that may contribute to health disparities.

In our models, cardiovascular fitness measures were strongly associated with obesity, including higher resting and overnight heart rate and low day-to-day variability in heart rate. Daytime resting heart rate is commonly used as an indicator of long-term cardiovascular health and fitness[58], and a lower heart rate during sleep also reflects

robust cardiovascular health[59]. Elevated resting heart rate is associated with higher BMI[60], waist circumference[58], and body fat percentage in adolescents[61,62]. This relationship may be multifaceted and complex. On the one hand, a higher resting heart rate could be linked to lower physical activity levels[58,63], potentially contributing to weight gain and obesity[64]. On the other hand, higher BMI itself might be associated with an elevated heart rate due to the extra workload on the heart[30]. It is essential to note, however, that our cross-sectional data do not permit a clear determination of causal directionality.

Longitudinal observational studies reveal a complex relationship between physical activity, BMI, and resting heart rate in adolescents. More active adolescents tend to have a lower prospective risk for obesity (see reference[65] for a review). This is supported by our results on the physical activity measures. High physical activity levels measured on weekdays and weekends appeared to be protective (10,000 + daily average step count), and low activity was related to high obesity risk as measured by the step count, which has been previously reported as a reliable measure for energy expenditure in adults[66]. Furthermore, it was also reported that the resting heart rate and changes in the resting heart rate have prospective effects and may predispose individuals to the development of obesity both in adolescents[67] and in adults[68]. In our analysis, both physical activity and resting heart rate were strong and independent predictors of obesity.

Later bedtime, short sleep duration, variability of mid-sleep timing, and variability in sleep onset and the end-point of sleep (wake up) were identified as important measures associated with obesity. The association between short sleep duration and obesity in adolescence has been supported by cross-sectional[69–71] and longitudinal studies[72]. Most of the previously reported results were based on self-reported data[69,70] or utilized objective sleep measures with relatively low sample sizes[71]. This is the first study, to our knowledge, documenting the association between poor sleep health and obesity based on wearable sleep technology in a large sample of adolescents. Although there are multiple studies showing that sleep quality and timing are related to body fat percentage and obesity risk in youth[73], awareness about the importance of sleep remains low in the general population[74], leading to potential long-term consequences on their physical and mental health[75]. Our findings, which provide specific thresholds for bedtime and other health factors associated with obesity, can be particularly valuable for practitioners. These insights may inform both prevention and intervention strategies for obesity[76], enhancing the ability to manage this complex health issue through a better understanding of the role of sleep.

The strengths of the present study include (i) the large and diverse adolescent sample and the length of the behavioral time series collected and analyzed, which is extensively higher than most previous studies analyzing obesity in adolescents, thus increasing the ecological validity of the present results; (ii) our modeling efforts, including the cross-validation and testing for non-linear associations; (iii) the choice to focus on step counts rather than energy equivalents, a decision informed by potential reliability issues with photoplethysmography, especially in individuals with high BMI[77] or dark skin tone[78].

The study also has limitations. First, because the Fitbit accurately detects sleep–wake states[36,59], we focused on derived outcomes of sleep quantity and quality but did not investigate sleep architecture. These outcomes, however, cannot provide a full understanding of sleep behavior, and future studies are needed to get more insight into the sleep-related brain activities in obesity (i.e., polysomnography). Second, challenges in detecting sleep onset[79] and potentially reduced reliability in individuals with sleep problems[80] pose limitations on the Fitbit's ability to fully capture sleep patterns. Third, the lower representation of adolescents identifying as a race other than White who participated in the Fitbit sub-study[81] indicates that our sample is not as diverse as the whole ABCD cohort, and there may be barriers to collecting these data in minorities. Lastly, the data are observational and cannot determine causality. Additionally, the inclusion of food intake and food logs could potentially increase the model performance; however, the current analysis did not include dietary information. Future work can extend this analysis with longitudinal data to determine whether sleep and physical activity measures predict the development of obesity over time, and thus offer more definitive insights.

Conclusion

In conclusion, the present study demonstrated that objective sleep, physical activity, and cardiovascular measures collected by a commercial wearable device combined with sociodemographic factors are accurate predictors of early adolescent obesity. Our results showed that elevated resting heart rate, lower socioeconomic status (e.g., lower household income), minority race, shorter sleep duration and lower step count were associated with a higher risk for obesity. Measuring sleep, steps, and heart rate using wearable consumer devices is a viable alternative for large health monitoring systems because they are easier to employ than in-person clinic visits. Considering the prevalence of obesity and the associated sleep problems in adolescents in the general population, these devices could be incorporated into the clinical practice to increase awareness and support screening, diagnosis, and treatment in this age group to promote metabolic health, healthy sleep, and overall well-being, especially in the socially disadvantaged groups who may face barriers to in-person clinic visits. This could lead to earlier detection of obesity and associated health problems, as well as more effective treatment and management. Incorporating wearable devices into clinical practice may also promote healthy behaviors, such as increased physical activity and improved sleep, which were protective in our analysis and could have positive effects on overall health and well-being beyond obesity prevention.

Data availability

The datasets analyzed during the current study are from the ABCD data repository, available at the NIMH Data Archive Digital Object Identifier (DOI) https://doi.org/10.15154/8873-zj65. DOIs can be found at https://doi.org/10.15154/1518688.

Received: 10 October 2023; Accepted: 26 April 2024
Published online: 31 May 2024

References

1. Ogden, C. L. *et al.* Trends in obesity prevalence by race and Hispanic origin—1999-2000 to 2017–2018. *JAMA* **324**, 1208–1210 (2020).
2. Hedley, A. A. *et al.* Prevalence of overweight and obesity among US children, adolescents, and adults, 1999–2002. *Jama* **291**, 2847–2850 (2004).
3. Guo, S. S., Wu, W., Chumlea, W. C. & Roche, A. F. Predicting overweight and obesity in adulthood from body mass index values in childhood and adolescence. *Am. J. Clin. Nutr.* **76**, 653–658 (2002).
4. Wang, Y. & Beydoun, M. A. The obesity epidemic in the United States—gender, age, socioeconomic, racial/ethnic, and geographic characteristics: A systematic review and meta-regression analysis. *Epidemiol. Rev.* **29**, 6–28 (2007).
5. Jasik, C. B. & Lustig, R. H. Adolescent obesity and puberty: The "perfect storm". *Ann. N. Y. Acad. Sci.* **1135**, 265–279 (2008).
6. Thompson, D. R. *et al.* Childhood overweight and cardiovascular disease risk factors: The National Heart, Lung, and Blood Institute Growth and Health Study. *J. Pediatr.* **150**, 18–25 (2007).
7. De Leonibus, C., Marcovecchio, M. L. & Chiarelli, F. Update on statural growth and pubertal development in obese children. *Pediatr. Rep.* **4**, e35 (2012).
8. Witchel, S. F., Burghard, A. C., Tao, R. H. & Oberfield, S. E. The diagnosis and treatment of PCOS in adolescents: An update. *Curr. Opin. Pediatr.* **31**, 562 (2019).
9. Gurnani, M., Birken, C. & Hamilton, J. Childhood obesity: Causes, consequences, and management. *Pediatr. Clin.* **62**, 821–840 (2015).
10. Nemiary, D., Shim, R., Mattox, G. & Holden, K. The relationship between obesity and depression among adolescents. *Psychiatr. Ann.* **42**, 305–308 (2012).
11. Lindberg, L., Hagman, E., Danielsson, P., Marcus, C. & Persson, M. Anxiety and depression in children and adolescents with obesity: A nationwide study in Sweden. *BMC Med.* **18**, 30 (2020).
12. Rankin, J. *et al.* Psychological consequences of childhood obesity: Psychiatric comorbidity and prevention. *Adolesc. Health Med. Ther.* **7**, 125 (2016).
13. Wardle, J., Brodersen, N. H., Cole, T. J., Jarvis, M. J. & Boniface, D. R. Development of adiposity in adolescence: Five year longitudinal study of an ethnically and socioeconomically diverse sample of young people in Britain. *BMJ* **332**, 1130–1135 (2006).
14. Lupton, D. Young people's use of digital health technologies in the global north: Narrative review. *J. Med. Internet Res.* **23**, e18286 (2021).
15. Lugones-Sanchez, C. *et al.* Effectiveness of an mHealth intervention combining a smartphone app and smart band on body composition in an overweight and obese population: Randomized controlled trial (EVIDENT 3 study). *JMIR mHealth uHealth* **8**, e21771 (2020).
16. Kim, J.-W. *et al.* Impact of personal health records and wearables on health outcomes and patient response: Three-arm randomized controlled trial. *JMIR mHealth uHealth* **7**, e12070 (2019).
17. Jauho, A.-M. *et al.* Effect of wrist-worn activity monitor feedback on physical activity behavior: A randomized controlled trial in Finnish young men. *Prev. Med. Rep.* **2**, 628–634 (2015).
18. Fatima, Y., Doi, S. A. R. & Mamun, A. A. Longitudinal impact of sleep on overweight and obesity in children and adolescents: A systematic review and bias-adjusted meta-analysis. *Obes. Rev.* **16**, 137–149 (2015).
19. Koren, D. *et al.* Sleep architecture and glucose and insulin homeostasis in obese adolescents. *Diabetes Care* **34**, 2442–2447 (2011).
20. Spiegel, K., Tasali, E., Penev, P. & Cauter, E. V. Brief communication: Sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. *Ann. Intern. Med.* **141**, 846–850 (2004).
21. Burt, J., Dube, L., Thibault, L. & Gruber, R. Sleep and eating in childhood: A potential behavioral mechanism underlying the relationship between poor sleep and obesity. *Sleep Med.* **15**, 71–75 (2014).
22. Colrain, I. M. & Baker, F. C. Changes in sleep as a function of adolescent development. *Neuropsychol. Rev.* **21**, 5–21 (2011).
23. Frederick, C. B., Snellman, K. & Putnam, R. D. Increasing socioeconomic disparities in adolescent obesity. *Proc. Natl. Acad. Sci.* **111**, 1338–1342 (2014).
24. Krueger, P. M. & Reither, E. N. Mind the gap: Race/ethnic and socioeconomic disparities in obesity. *Curr. Diabetes Rep.* **15**, 1–9 (2015).
25. Hills, A. P., Andersen, L. B. & Byrne, N. M. Physical activity and obesity in children. *Br. J. Sports Med.* **45**, 866–870 (2011).
26. Daniels, S. R. Critical periods for abnormal weight gain in children and adolescents. *Handbook of Pediatric Obesity: Etiology, Pathophysiology and Prevention*, 67–78 (2005).
27. Morano, M., Colella, D. & Capranica, L. Body image, perceived and actual physical abilities in normal-weight and overweight boys involved in individual and team sports. *J. Sports Sci.* **29**, 355–362 (2011).
28. Vasconcellos, F. *et al.* Physical activity in overweight and obese adolescents: Systematic review of the effects on physical fitness components and cardiovascular risk factors. *Sports Med.* **44**, 1139–1152 (2014).
29. HenjeBlom, E., Olsson, E. M., Serlachius, E., Ericson, M. & Ingvar, M. Heart rate variability is related to self-reported physical activity in a healthy adolescent population. *Eur. J. Appl. Physiol.* **106**, 877–883 (2009).
30. Sun, K. & Scherer, P. E. Adipose tissue dysfunction: A multistep process. *Novel Insights into Adipose Cell Functions*, 67–75 (2010).
31. Piché, M.-E., Poirier, P., Lemieux, I. & Després, J.-P. Overview of epidemiology and contribution of obesity and body fat distribution to cardiovascular disease: An update. *Prog. Cardiovasc. Dis.* **61**, 103–113 (2018).
32. National Health and Nutrition Examination Survey 2017–March 2020 Prepandemic Data Files Development of Files and Prevalence Estimates for Selected Health Outcomes. in *National Health Statistics Reports* (ed. National Center for Health, S.). https://doi.org/10.15620/cdc:106273 (2021).
33. Garavan, H. *et al.* Recruiting the ABCD sample: Design considerations and procedures. *Dev. Cogn. Neurosci.* **32**, 16–22 (2018).
34. Casey, B. J. *et al.* The adolescent brain cognitive development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).
35. Bagot, K. *et al.* Current, future and potential use of mobile and wearable technologies and social media data in the ABCD study to increase understanding of contributors to child health. *Dev. Cogn. Neurosci.* **32**, 121–129 (2018).
36. Godino, J. G. *et al.* Performance of a commercial multi-sensor wearable (Fitbit Charge HR) in measuring physical activity and sleep in healthy children. *PLoS ONE* **15**, e0237719 (2020).
37. Page, E. J., Massey, A. S., Prado-Romero, P. N. & Albadawi, S. The use of self-monitoring and technology to increase physical activity: A review of the literature. *Perspect. Behav. Sci.* **43**, 501–514 (2020).
38. van Sluijs, E. M., van Poppel, M. N., Twisk, J. W. & van Mechelen, W. Physical activity measurements affected participants' behavior in a randomized controlled trial. *J. Clin. Epidemiol.* **59**, 404–411 (2006).
39. Matthews, C. E., Ainsworth, B. E., Thompson, R. W. & Bassett, D. R. Jr. Sources of variance in daily physical activity levels as measured by an accelerometer. *Med. Sci. Sports Exerc.* **34**, 1376–1381 (2002).
40. Hart, T. L., Swartz, A. M., Cashin, S. E. & Strath, S. J. How many days of monitoring predict physical activity and sedentary behaviour in older adults?. *Int. J. Behav. Nutr. Phys. Activity* **8**, 1–7 (2011).

41. Aili, K., Åström-Paulsson, S., Stoetzer, U., Svartengren, M. & Hillert, L. Reliability of actigraphy and subjective sleep measurements in adults: The design of sleep assessments. *J. Clin. Sleep Med.* **13**, 39–47 (2017).
42. Van Someren, E. J. Improving actigraphic sleep estimates in insomnia and dementia: how many nights?. *J. Sleep Res.* **16**, 269–275 (2007).
43. Jahng, S., Wood, P. K. & Trull, T. J. Analysis of affective instability in ecological momentary assessment: Indices using successive difference and group comparison via multilevel modeling. *Psychol. Methods* **13**, 354 (2008).
44. St Fleur, R. G. *et al.* Use of Fitbit devices in physical activity intervention studies across the life course: Narrative review. *JMIR mHealth uHealth* **9**, e23411 (2021).
45. Hemphill, N. M., Kuan, M. T. & Harris, K. C. Reduced physical activity during COVID-19 pandemic in children with congenital heart disease. *Can. J. Cardiol.* **36**, 1130–1134 (2020).
46. Barlow, S. E. & Committee, E. Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: Summary report. *Pediatrics* **120**, S164–S192 (2007).
47. Troyanskaya, O. *et al.* Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**, 520–525 (2001).
48. Waskom, M. L. Seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
49. Nori, H., Jenkins, S., Koch, P. & Caruana, R. Interpretml: A unified framework for machine learning interpretability. arXiv:1909.09223 (2019).
50. Ling, C. X., Huang, J. & Zhang, H. AUC: A statistically consistent and more discriminating measure than accuracy. *IJCAI* **3**, 519–524 (2003).
51. Halimu, C., Kasem, A. & Newaz, S.S. Empirical comparison of area under ROC curve (AUC) and Mathew correlation coefficient (MCC) for evaluating machine learning algorithms on imbalanced datasets for binary classification. in *Proceedings of the 3rd International Conference on Machine Learning and Soft Computing*, 1–6 (2019).
52. He, H. & Ma, Y. *Imbalanced Learning: Foundations, Algorithms, and Applications* (Springer, 2013).
53. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
54. Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M. & Flegal, K. M. Prevalence of high body mass index in US children and adolescents, 2007–2008. *JAMA* **303**, 242–249 (2010).
55. Singh, G. K., Kogan, M. D., Van Dyck, P. C. & Siahpush, M. Racial/ethnic, socioeconomic, and behavioral determinants of childhood and adolescent obesity in the United States: Analyzing independent and joint associations. *Ann. Epidemiol.* **18**, 682–695 (2008).
56. Rendall, M. S., Weden, M., Fernandes, M. & Vaynman, I. H ispanic and black US children's paths to high adolescent obesity prevalence. *Pediatr. Obes.* **7**, 423–435 (2012).
57. de Oliveira Forkert, E. C. *et al.* Abdominal obesity and its association with socioeconomic factors among adolescents from different living environments. *Pediatr. Obes.* **12**, 110–119 (2017).
58. Kwok, S.-Y. *et al.* Resting heart rate in children and adolescents: Association with blood pressure, exercise and obesity. *Arch. Dis. Childhood* **98**, 287–291 (2013).
59. de Zambotti, M. *et al.* Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. *Physiol. Behav.* **158**, 143–149 (2016).
60. Sarganas, G., Rosario, A. S. & Neuhauser, H. K. Resting heart rate percentiles and associated factors in children and adolescents. *J. Pediatr.* **187**, 174–181 (2017).
61. Farah, B. Q. *et al.* Association between resting heart rate and cardiovascular risk factors in adolescents. *Eur. J. Pediatr.* **174**, 1621–1628 (2015).
62. Fernandes, R. A. *et al.* Resting heart rate: Its correlations and potential for screening metabolic dysfunctions in adolescents. *BMC Pediatr.* **13**, 1–7 (2013).
63. Silva, D. A. S., Lima, T. R. D. & Tremblay, M. S. Association between resting heart rate and health-related physical fitness in Brazilian adolescents. *BioMed Res. Int.* **2018**, 1–10 (2018).
64. Rauner, A., Mess, F. & Woll, A. The relationship between physical activity, physical fitness and overweight in adolescents: A systematic review of studies published in or after 2000. *BMC Pediatr.* **13**, 1–9 (2013).
65. Must, A. & Tybor, D. J. Physical activity and sedentary behavior: A review of longitudinal studies of weight and adiposity in youth. *Int. J. Obes.* **29**, S84–S96 (2005).
66. Adam Noah, J., Spierer, D. K., Gu, J. & Bronner, S. Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. *J. Med. Eng. Technol.* **37**, 456–462 (2013).
67. Palatini, P. *et al.* Resting heart rate as a predictor of body weight gain in the early stage of hypertension. *Obesity* **19**, 618–623 (2011).
68. Shigetoh, Y. *et al.* Higher heart rate may predispose to obesity and diabetes mellitus: 20-year prospective study in a general population. *Am. J. Hypertens.* **22**, 151–155 (2009).
69. Chen, X., Beydoun, M. A. & Wang, Y. Is sleep duration associated with childhood obesity? A systematic review and meta-analysis. *Obesity* **16**, 265 (2008).
70. Knutson, K. L. & Van Cauter, E. Associations between sleep loss and increased risk of obesity and diabetes. *Ann. N. Y. Acad. Sci.* **1129**, 287–304 (2008).
71. Pacheco, S. R. *et al.* Overweight in youth and sleep quality: Is there a link?. *Arch. Endocrinol. Metab.* **61**, 367–373 (2017).
72. Mitchell, J. A., Rodriguez, D., Schmitz, K. H. & Audrain-McGovern, J. Sleep duration and adolescent obesity. *Pediatrics* **131**, e1428–e1434 (2013).
73. Hart, C. N., Cairns, A. & Jelalian, E. Sleep and obesity in children and adolescents. *Pediatr. Clin.* **58**, 715–733 (2011).
74. Meltzer, L. J. *et al.* The children's report of sleep patterns (CRSP): A self-report measure of sleep for school-aged children. *J. Clin. Sleep Med.* **9**, 235–245 (2013).
75. Fallone, G., Owens, J. A. & Deane, J. Sleepiness in children and adolescents: Clinical implications. *Sleep Med. Rev.* **6**, 287–306 (2002).
76. Wang, W., Cheng, J., Song, W. & Shen, Y. The effectiveness of wearable devices as physical activity interventions for preventing and treating obesity in children and adolescents: Systematic review and meta-analysis. *JMIR mHealth uHealth* **10**, e32435 (2022).
77. McVeigh, J. A. *et al.* Convergent validity of the fitbit charge 2 to measure sedentary behavior and physical activity in overweight and obese adults. *J. Meas. Phys. Behav.* **4**, 39–46 (2021).
78. Nowara, E. M., McDuff, D. & Veeraraghavan, A. A meta-analysis of the impact of skin type and gender on non-contact photoplethysmography measurements. in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1148–1155 (2020).
79. de Zambotti, M., Cellini, N., Goldstone, A., Colrain, I. M. & Baker, F. C. Wearable sleep technology in clinical and research settings. *Med. Sci. Sports Exerc.* **51**, 1538–1557 (2019).
80. Kahawage, P., Jumabhoy, R., Hamill, K., de Zambotti, M. & Drummond, S. P. A. Validity, potential clinical utility, and comparison of consumer and research-grade activity trackers in Insomnia Disorder I: In-lab validation against polysomnography. *J. Sleep Res.* **29**, e12931 (2020).
81. Kim, E. H. *et al.* Association of demographic and socioeconomic indicators with the use of wearable devices among children. *JAMA Netw. Open* **6**, e235681–e235681 (2023).
82. Lou, Y., Caruana, R., Gehrke, J. & Hooker, G. Accurate intelligible models with pairwise interactions. in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 623–631 (2013).

Acknowledgements
J.M.N. was funded by the Doris Duke Charitable Foundation (2022056) and the National Heart, Lung, and Blood Institute (K08HL159350 and R01MH135492). O.K and F.B was supported by the National Institutes of Health (R01MH135492 and R01 MH128959) and additional federal partners under award number U01DA041022. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions
OK cleaned and analyzed the data, constructed the figures, and wrote the first draft of the manuscript. Conception and design of experiments and analysis: OK, JMN, RP, FCB. Writing—review & editing; FCB, JMN, EED, KPG. All authors critically reviewed the manuscript and approved the final manuscript for submission.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-60811-2.

Correspondence and requests for materials should be addressed to O.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024