Digital phenotyping from wearables using AI characterizes psychiatric disorders and identifies genetic associations

Jason Liu[1,2,*], Beatrice Borsari[1,2,*], Yunyang Li[3], Susanna Liu[1,2], Yuan Gao[4], Xin Xin[1,2], Shaoke Lou[1,2], Matthew Jensen[1,2], Diego Garrido-Martín[5], Terril Verplaetse[6], Garrett Ash[7,8,9], Jing Zhang[10], Matthew J. Girgenti[6], Walter Roberts[6,9,#], Mark Gerstein[1,2,3,9,11,#]

[1] Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, 06511, USA
[2] Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, 06511, USA
[3] Department of Computer Science, Yale University, New Haven, CT, 06511, USA
[4] Yale School of Public Health, Division of Health Informatics, Yale University, New Haven, CT, 06511, USA
[5] Department of Genetics, Microbiology and Statistics, Universitat de Barcelona (UB), Barcelona, 08028, Spain
[6] Department of Psychiatry, Yale University School of Medicine, New Haven, CT, 06511, USA
[7] Section of General Internal Medicine, Yale University School of Medicine, New Haven, CT, 06511, USA
[8] Center for Pain, Research, Informatics, Medical Comorbidities and Education Center (PRIME), VA Connecticut Healthcare System, West Haven, CT, 06516, USA
[9] Department of Biomedical Informatics and Data Science, Yale University, New Haven, CT, 06511, USA
[10] Department of Computer Science, University of California, Irvine, CA, 92697, USA
[11] Department of Statistics and Data Science, Yale University, New Haven, CT, 06511, USA

* Co-first authors
# Co-corresponding authors (pi@gersteinlab.org)

Abstract

Psychiatric disorders are complex and influenced by both genetic and environmental factors. However, studying the full spectrum of these disorders is hindered by practical limitations on measuring human behavior. This highlights the need for novel technologies that can measure behavioral changes at an intermediate level between diagnosis and genotype. Wearable devices are a promising tool in precision medicine, since they can record physiological measurements over time in response to environmental stimuli and do so at low cost and minimal invasiveness. Here we analyzed wearable and genetic data from a cohort of the Adolescent Brain Cognitive Development study. We generated >250 wearable-derived features and used them as intermediate phenotypes in an interpretable AI modeling framework to assign risk scores and classify adolescents with psychiatric disorders. Our model identifies key physiological processes and leverages their temporal patterns to achieve a higher performance than has been previously possible. To investigate how these physiological processes relate to the underlying genetic architecture of psychiatric disorders, we also utilized these intermediate phenotypes in univariate and multivariate GWAS. We identified a total of 29 significant genetic loci and 52 psychiatric-associated genes, including *ELFN1* and *ADORA3*. These results show that wearable-derived continuous features enable a more precise representation of psychiatric disorders and exhibit greater detection power compared to categorical diagnostic labels. In summary, we demonstrate how consumer wearable technology can facilitate dimensional approaches in precision psychiatry and uncover etiological linkages between behavior and genetics.

Introduction

Psychiatric disorders of childhood and adolescence currently affect 1 in 7 youths in the United States and globally[1,2]. Externalizing disorders such as attention-deficit/hyperactivity disorder, and internalizing disorders such as anxiety, are among the most prevalent and represent a wide spectrum of dysfunctional behavior patterns[3]. Treatment barriers are complex and multifaceted but major contributors include our limited understanding of psychiatric phenotypes and difficulty identifying youth individuals that experience these disorders.

Traditionally, psychiatric disorders have been conceptualized as categorical macrophenotypes, based on clinical manifestations of a disease which are defined according to the number and type of symptoms, and the presence of distress or impairment[4-6]. While this has practical benefits in terms of reliability and ease of diagnosis, it poses several challenges to the research of these disorders, and consequently to the development of treatments. In fact, psychiatric disorders are complex and often comorbid, and this high degree of heterogeneity is not always accurately translated into categorical diagnosis labels, which may be defined by arbitrary cut-offs. Instead, intermediate phenotypes (i.e., quantitative traits that are positioned between genotype and macrophenotype) may better capture the heterogeneity potentially missed by existing diagnostic categories[7-9]. Additionally, genetic penetrance is expected to be higher for these intermediate phenotypes compared to macrophenotypes, enabling improved dissection of the genetic architecture underlying psychiatric disorders[10]. Nevertheless, many genome-wide association studies (GWAS) aimed at identifying genetic variants or biomarkers for psychiatric disorders do not consider these intermediate phenotypes and instead rely on dichotomised (i.e., binary) traits. In fact, identifying intermediate phenotypes with clinical and biological relevance remains a challenge[11].

Therefore, to improve our understanding of psychiatric disorders it is important that we identify intermediate phenotypes that not only offer a more comprehensive representation of an individual's behavior with respect to their environment, but also relate well with existing clinical definitions and aid in diagnosis. Once identified, these intermediate phenotypes can then be also used to guide more comprehensive studies to identify genetic associations and biomarkers that may ultimately improve precision treatments.

To achieve this goal, it is important to leverage new emerging technologies that can quantitatively assess an individual's behavioral patterns[12]. Wearable sensors such as smartwatches collect data that reflect physical and physiological processes (e.g., movement, pulse, metabolic intake), and can be used to infer higher-order behavioral events (e.g., sleep, exercise) and their temporal dynamics. Because of the documented relationship between such higher-order behavioral events and mental health, and given their low cost and minimal invasiveness, wearable devices have emerged as promising tools for mental health monitoring and psychiatric evaluation[13-15].

Therefore, wearable sensors have the potential for capturing intermediate phenotypes relevant to behavior and psychiatric disorders, ultimately enabling improved GWAS. However, significant computational challenges remain in generating intermediate phenotypes from wearable-derived data that describe the full spectrum of a given psychiatric disorder. Moreover, further curation of these intermediate phenotypes is necessary to identify genetic associations that have clinical and biological relevance.

To address these limitations, we developed an AI modeling framework that flexibly leverages data from wearable devices to generate intermediate phenotypes in the form of static and dynamic digital features. We establish the validity of these digital features as intermediate phenotypes by classifying externalizing and internalizing disorders with an accuracy beyond baseline expectation, and even surpassing the performance of some other gold-standard intermediate phenotypes such as fMRI measurements[16-19]. Interpretability modules in our AI framework enable us to identify key temporal and physiological insights between clinical diagnosis and digital features, further supporting the validity of using these wearable-derived features as intermediate phenotypes. We further curate these intermediate phenotypes and employ them in GWAS models to identify genetic associations and biomarkers that capture the continuous spectrum of psychiatric disorders and behavioral patterns. Finally, we identify 29 significant loci, several of which overlap previously reported genetic variants associated with behavioral traits and mental illnesses and are proximal to genes with a documented role in neurodevelopmental and psychiatric disorders.

In sum, this work shows how wearable devices can advance our understanding of psychiatric disorders by establishing a more objective and dimensional approach that can ultimately lead to improved treatments in precision psychiatry.

Results

Leveraging the Adolescent Brain Cognitive Development cohort

To improve our understanding of psychiatric disorders, we leveraged and analyzed a dataset from a cohort of US adolescents recruited by the NIH Adolescent Brain Cognitive Development Consortium (ABCD) project, consisting of clinical, wearable, and genetic data (**Fig. 1**). The ABCD cohort consists of a total of 11,878 adolescents (5682 males and 6196 females), of age between nine and fourteen years and belonging to four different ethnicities. We identified nine categories of psychiatric phenotypes, which were established using a gold standard parent diagnostic semi-structured interview (Kiddie Schedule for Affective Disorders and Schizophrenia-5)[20]. The healthy controls represented adolescents who did not meet the criteria for any of those nine psychiatric disorders. We defined these clinical labels as the categorical macrophenotypes in the study (**Fig. 1A-B**). Our modeling framework also utilized data from cognitive tests (e.g., NIH Toolbox) and behavioral checklists (**Fig. 2A**).

Generating intermediate phenotypes from wearable-derived data

We processed data obtained from FitBit smartwatches, which comprise measurements of heart rate, calories, activity intensity, steps, metabolic equivalents, sleep level and sleep intensity (**Fig. 1C**)[21]. These measurements quantify an individual's physiological processes and their real-time changes in response to environmental stimuli, and can thus provide key information about an individual's behavior.

To reconstruct the full spectrum of an individual's behavioral functioning from these data, we applied two different feature engineering techniques, allowing us to generate wearable-derived dynamic and static features, which we consider as intermediate phenotypes. The dynamic features preserve the time-varying nature of the original data as a time series, enabling sequential and temporal patterns of the data to be retained. In contrast, the static features summarize patterns of

the digital data to produce time-invariant, quantitative features that are commonly used in downstream modeling[14,22].

To generate dynamic features, we performed signal imputation and processing after filtering the individuals with sparse data, and obtained 48 channels of time series (**Figure 2B**). Compared to the static features, this further processing allowed us to preserve both local and global temporal patterns potentially relevant to characterizing behavior and neurological response to stimuli[8,23].

To generate static features, we first collected a total of 49 FitBit summary-based features. We next applied descriptive statistics (e.g. mean, median, etc.) to each of these features and generated a total of 258 static features for each individual (**Fig. 2C**)[14,24]. We then grouped these static features into seven main clusters, each of which summarizes different aspects of physiological and behavioral processes, such as heart rate, sleep duration and quality, metabolic intake or physical activity (**Fig. 2D**).

Altogether, static and dynamic features represent the physiological and behavioral profiles of the adolescents, and can be leveraged as intermediate phenotypes in a wide range of analyses to better characterize psychiatric phenotypes, such as generating disorder-specific probability risk scores, macrophenotype classification, model interpretability, and biomarker identification via wearable GWAS.

Predicting psychiatric phenotypes from wearable-derived intermediate phenotypes

To demonstrate the validity of static and dynamic features as clinically relevant intermediate phenotypes and to evaluate their utility as a diagnostic tool, we employed these features in an array of classification tasks to identify individuals with either an externalizing (ADHD) or internalizing (anxiety) disorder from their typically developing peers. We selected ADHD and anxiety due to their high prevalence in adolescents, which is mirrored in the cohort (**Fig. 1B**)[25].

We applied a gradient boosting machine learning algorithm, XGBoost, for classification tasks using static features (**Fig. 2C-D**)[26]. On the other hand, to fully leverage the time series nature of the dynamic features, we used a convolutional neural network for time series, featuring depthwise separable convolution, called Xception **(Fig. 2E)**[27]. Variable convolutional filters and residual (skip) connections, coupled with efficient parametrization, allow information encoded in both small and large receptive fields to be more optimally leveraged. Practically, this framework takes into account local and global patterns of physiology and behavior when performing downstream classification of psychiatric disorders. In both modeling approaches we included covariates that accounted for demographic features, family history of disorders, and other clinical information (**Fig. 2A**). To assess the benefit, in terms of model performance, of including wearable-derived data, we also trained a baseline model using just the covariates, which served as a comparison to the models including static or dynamic features. In practice, this comparison allowed us to determine whether wearable-derived features can improve diagnostic accuracy relative to that achievable using only a widely used broadband behavior rating scale.

After data filtering, we first used static features to classify 216 individuals with ADHD (an externalizing disorder) versus 1,737 of their typically developing peers (healthy controls) (**Fig. 3A)**. Using static features with XGBoost, we achieved an average area under the receiver operating

characteristic curve (AUROC) of 0.87 and precision of 0.79. When using the dynamic features and Xception, we were able to achieve an average AUROC of 0.89 and precision of 0.83. The baseline model consisting of only the covariates achieved an average AUROC of 0.83, suggesting that the inclusion of wearable-derived features facilitates a clinically meaningful improvement in diagnostic accuracy. This improvement between the baseline and dynamic features model demonstrates statistical significance (one-sided t-test between baseline and dynamic features model, p value = 0.0022).

Second, we evaluated the performance of our model using static or dynamic features in the classification of 666 individuals diagnosed with anxiety disorder (internalizing disorder) versus 1,737 of their typically developing peers (healthy controls) (**Fig. 3B)**. Here, we again repeated the use of the same modeling framework, i.e., static features with XGBoost and dynamic features with Xception, and compared it to the baseline covariate model. We found that static and dynamic features achieve an average AUROC of 0.69 and 0.71 and precision of 0.64 and 0.68, respectively. In both models, the performance was greater than that of the baseline model (average AUROC of 0.67), with the dynamic features model showing the most significant performance improvement (one-sided t-test between baseline and dynamic feature model, p value = 0.00016). Overall, the fact that the models using dynamic features achieved the highest performance suggests the usefulness of the temporal patterns intrinsic to wearable-derived data towards understanding behavior.

Interpreting wearable features prioritized by the deep learning model

Deep learning methods are typically characterized by complex internal structures that cannot be easily interpreted by humans. While maximizing the classification accuracy is one crucial aspect for characterizing complex phenotypes, understanding which features are most important in terms of their individual contribution to performance is also critical. To this end, we utilized ablation techniques to determine the relative contribution of each individual feature to model performance. For the ADHD classification task, heart rate was the most important feature (largest change in AUROC), followed by other dynamic features (i.e., sleep, steps, METs) as well as covariates such as demographics, family history, and cognitive scores from picture memory and stop-signal reaction time tests (**Fig. 3C**).

On the other hand, the ablation study for the anxiety classification task revealed a different set of important features. In this case, sleep quality and stage, calories, and step count were the most important dynamic features, whereas heart rate features, which were extremely important for classifying ADHD, were not prioritized in the anxiety model (**Fig. 3D**). Additionally, while the anxiety model prioritized some covariates that were relevant also for the ADHD model (e.g., sex, family history, and family divorce), cognitive scores from tests such as picture memory did not appear to be important for the identification of individuals diagnosed with anxiety, consistent with theory-driven accounts of neurocognitive aspects of anxiety disorders[28].

Additionally, we assessed the importance for model performance of dynamic features at various times throughout the day by adapting gradient-weighted class activation mapping (Grad-CAM) strategies[29]. We calculated the relative importance of each time point in our dynamic features. For ADHD, we observed enriched significance of the heart rate dynamic feature around the early afternoon, potentially suggesting stronger behavioral differences between adolescents with ADHD

and their typically developing peers (healthy controls) during this time of day (**Fig. 3E**). This is consistent with clinical research demonstrating time-of-day effects on ADHD symptom expression[30]. In contrast, sleep-related dynamic features during the night are much more informative in classifying anxiety, consistent with clinical expectations (**Fig. 3F**)[31]. Together, the ablation studies suggest a role for wearable-derived features to not only serve as quantitative intermediate phenotypes, but also to more closely reveal insights into the behavioral and physiological temporal patterns related to categorical macrophenotypes.

Using wearable-derived features as intermediate phenotypes for wearable GWAS

Our accurate classification of both internalizing and externalizing psychiatric phenotypes based on wearable-derived features suggests that these features can serve as useful intermediate phenotypes and may be leveraged to identify genetic associations with psychiatric conditions (**Fig. 4**). To this end, we first focused specifically on ADHD, given the higher predictive power observed with our models (**Fig. 3A-B**) and its higher estimated heritability compared to anxiety (77-88% vs. 30-60%)[32,33]. We selected 1,191 individuals (137 individuals with ADHD and 1,054 healthy control individuals) with genetic and wearable data available, and performed a GWAS using the continuous prediction scores obtained from our wearable modeling framework. In practice, these scores represent risk probabilities for ADHD[34,35]. We identified 10 genome-wide (p value $< 5 \cdot 10^{-8}$) significant loci and 21 psychiatric or brain-related genes (**Fig. 4B**). Three of the identified genes (*ADORA3*, *PSMD11* and *DLG4*) have been previously associated with ADHD, bolstering the overall functional significance of the results[36-38]. Furthermore, several of these loci overlap with previously reported GWAS SNPs related to ADHD, neuroticism, sleep disruption and other clinically relevant traits (**Fig. 4B**). Note that here we used a continuous risk score as a univariate response variable for the GWAS. In comparison, when performing a traditional case–control GWAS for ADHD on the same set of individuals using the binary diagnostic label (presence/absence of disorder) as response variable, we did not identify any significant loci (**Fig. 1A, Fig. 4A**). This result is consistent with the higher statistical power of continuous measurements over dichotomized (i.e., binary) traits, and with the findings that intermediate phenotypes show higher genetic penetrance compared to macrophenotypes[7,39-41].

While the analysis above collapses wearable-derived features into a single continuous variable that summarizes the risk score for a particular disorder (i.e., GWAS with a univariate response), it is also possible to directly use the full set of wearable-derived features as a more comprehensive and richer phenotype to represent the continuum of psychiatric disorders and their latent manifestations (i.e., GWAS with a multivariate response). In fact, these features can collectively capture behavioral patterns by measuring physiological processes and their real-time changes in response to environmental stimuli, and unlike disease risk scores, are not restricted to a specific cohort of individuals[42,43]. In what follows, we employed a multivariate nonparametric test of association to regress the vector of wearable-derived features on the genotype of each genetic variant, employing a larger cohort that spans healthy controls and individuals with any psychiatric disorder (n = 2,410)[44]. From this novel type of GWAS we identified 19 significant loci and 31 genes with a documented role in neurodevelopmental and psychiatric disorders (**Fig. 4C**). Many of these loci overlap previously identified GWAS SNPs related to heart, sleep, metabolism, and brain traits (**Fig. 4C**). This aligns with the close association between physiological functions, the central nervous system, and individual behavior.

Functionally dissecting and interpreting novel wearable GWAS loci

To further investigate the loci identified by the behavioral GWAS, we dissected the variants using a battery of publicly available genomic resources[45,46]. Many of these loci overlap either GTEx expression quantitative trait loci (eQTLs) or ENCODE candidate cis-regulatory elements (cCREs), suggesting a link between the biochemical activity of these variants and their functional impact on macrophenotype. We also explored the impact of these loci beyond behavioral traits and their relationship with clinical psychopathology. For example, behavioral traits significantly associated with a specific genetic variant may correlate with clinical symptoms of a specific psychiatric cohort. Indeed, in some cases we show that the genetic variant in question is also differentially enriched between that specific psychiatric cohort and healthy individuals.

For instance, we found the minor allele (G) at rs365990 to be significantly associated with an increase in mean heart rate and a decrease in interday heart rate variation (**Fig. 5A, left**). The variant, missense for *MYH6*, had been previously linked to atrial fibrillation, ventricular tachycardia and resting heart rate, and the entire locus shows a significant enrichment of chromatin features in heart samples compared to other tissues and organs[47-50]. We also found the same allele to be enriched in the bipolar/psychotic disorder cohort compared to healthy controls (**Fig. 5A, right**). This cohort included youth meeting criteria for bipolar or unspecified psychotic spectrum disorder, and such severe pathology is known to be associated with characteristic irregularities in heart activity[51-53]. SNP rs365990 is also a GTEx eQTL for the *CMTM5* gene, which is highly expressed in brain subregions and has been implicated in stress response and childhood adversity, further supporting the relevance of this locus for psychiatric conditions in addition to heart pathophysiology[46,54].

In a similar fashion, we explored variants rs113525298 and rs147959551. The minor allele at rs113525298 is associated with prolonged periods in bed and shorter vigorously active time during the day, and appears at a lower frequency in the ADHD cohort compared to healthy individuals (**Fig. 5B**). This suggests a potential protective role of the allele against hyperactivity disorders, further supported by the proximity of the SNP to *ELFN1*, previously implicated in the pathophysiology of ADHD[55,56]. On the other hand, we found the minor allele at rs147959551 associated with shorter sedentary time at night and a shorter period of time identified as sleep based on heart rate, two features suggestive of sleep disruption (**Fig. 5C)**. The same allele is also enriched in individuals with depression disorder, consistent with growing evidence implicating sleep impairment as a transdiagnostic feature of many forms of adolescent psychopathology (**Fig. 5C, right)**[57,58].

Overall, these results highlight how wearable-derived features can be leveraged as intermediate phenotypes in GWAS and enable the identification of genetic variants relevant to clinical psychiatry with significant effects on exhibited behavior in adolescents.

Discussion

Psychiatric disorders have been traditionally described with diagnostic categories based on retrospective self-report of symptom sets. However, current efforts in the field are increasingly leveraging novel technologies to transition from retrospective self-reporting and fixed symptom sets to more dimensional conceptualizations, aiming to capture the complex and heterogeneous nature of psychiatric disorders for more accurate research into their underlying structure[6]. One approach to enhancing dimensional models is the use of intermediate phenotypes—quantitative traits linked more closely to a disorder's underlying molecular pathways. Although intermediate phenotypes have been derived from cellular, tissue, and organ levels of information, computational strategies that generate useful intermediate phenotypes in the behavioral domain are currently limited. Wearable biosensors such as smartwatches offer a unique opportunity to objectively study psychiatric disorders in a non-invasive way by measuring their underlying physiological foundations of behavior over time.

Towards this end, we used wearable data to generate static and dynamic features that were employed by our AI modeling framework as intermediate phenotypes to distinguish between adolescents with and without psychiatric disorders. Models utilizing these wearable-derived intermediate phenotypes performed comparably to those based on more expensive data sources such as fMRI measurements[18,59]. To gain critical theoretical insights and inform treatment development efforts, we augmented the modeling framework with interpretability modules, allowing us to pinpoint temporal and functional regions of the time series that were highly correlated with overall disease state[60]. These interpretability modules have the potential to facilitate mechanistic studies that offer deeper insight into the underlying complexities of these disorders. For example, our interpretability modules revealed that heart rate time series held high importance in predicting ADHD. This finding aligns with the clinical manifestation of ADHD – affected children are characterized by episodes of heightened arousal that are often incongruent with environmental demands[61]. Conversely, the interpretability modules identified sleep intensity and quality as key predictors in our anxiety disorder models, in line with known disruptions in sleep patterns and circadian rhythms commonly seen in youth with anxiety disorders[62].

Wearable-derived intermediate phenotypes are not just effective for detecting the presence of psychiatric disorders in individuals; they also serve as a valuable research tool for understanding the correspondence between behavior patterns and molecular attributes. This comprehensive approach helps to uncover the foundational elements of pathological behavior patterns. In this context, we focused on establishing links with genetics. Specifically, we showed that these intermediate phenotypes can serve as response variables in GWAS models. Their continuous nature enhances statistical power compared to categorical diagnostic labels. Furthermore, we took advantage of the features' correlated structure to create multivariate response variables for GWAS. This strategy is statistically advantageous because it mitigates the multiple testing burden associated with evaluating the numerous (>250) independent features. Conversely, from a biological standpoint, these wearable GWAS allowed us to explore triaxial associations encompassing genetic, physiological, and psychiatric factors. Utilizing our framework, we successfully identified a significant association between a missense variant of the *MYH6* gene, which encodes the cardiac muscle myosin, and heart rate patterns. Heart activity receives complex inputs from the CNS, which implies behavioral influence and, in combination with our GWAS, supports the notion of a gene-behavior-disorder pathway[63]. Building on this finding, we discovered

enrichment of the same genetic variant among individuals with bipolar/psychotic disorders, psychiatric conditions known to be associated with characteristic irregularities in heart activity[51]. While additional research is needed to confirm such associations, our findings resonate with the objectives of the RDoC initiative[6]. Specifically, wearable-derived intermediate phenotypes serve as objective markers of behavior, bridging lower-level biological systems like genetics to broader psychiatric disorders.

While we have employed these wearable-derived intermediate phenotypes in a targeted research context (i.e., to enhance a psychiatric GWAS), their broad applicability make them promising for other domains of health research. For example, the risk scores generated by our AI-modeling framework could be used to assess disorder severity, and the genetic variants identified by our wearable GWAS could be employed to construct more comprehensive polygenic risk scores for behavioral and psychiatric disorders. Unlike other diseases (e.g., cancer) where objective biomarkers are more common, psychiatry faces a significant barrier in treatment due to the lack of objective and sensitive screening methods[64]. Therefore, these physiological and genetic features could be leveraged as objective biomarkers to more accurately subtype patients within diagnostic categories, which in turn could help move towards precision treatment delivery in psychiatry.

Although the results presented in this study require further experimental validation, they illuminate the transformative potential of wearable devices combined with AI modeling frameworks for deepening our understanding of complex behavioral and psychiatric traits. We anticipate that further development of our AI modeling framework, coupled with an expanded array of wearable devices, could fundamentally transform how psychiatric disorders are measured and understood in both research and clinical settings. This could lead to more nuanced digital intermediate phenotypes and open new avenues for the study of human behavior.

Acknowledgements

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive DevelopmentSM (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

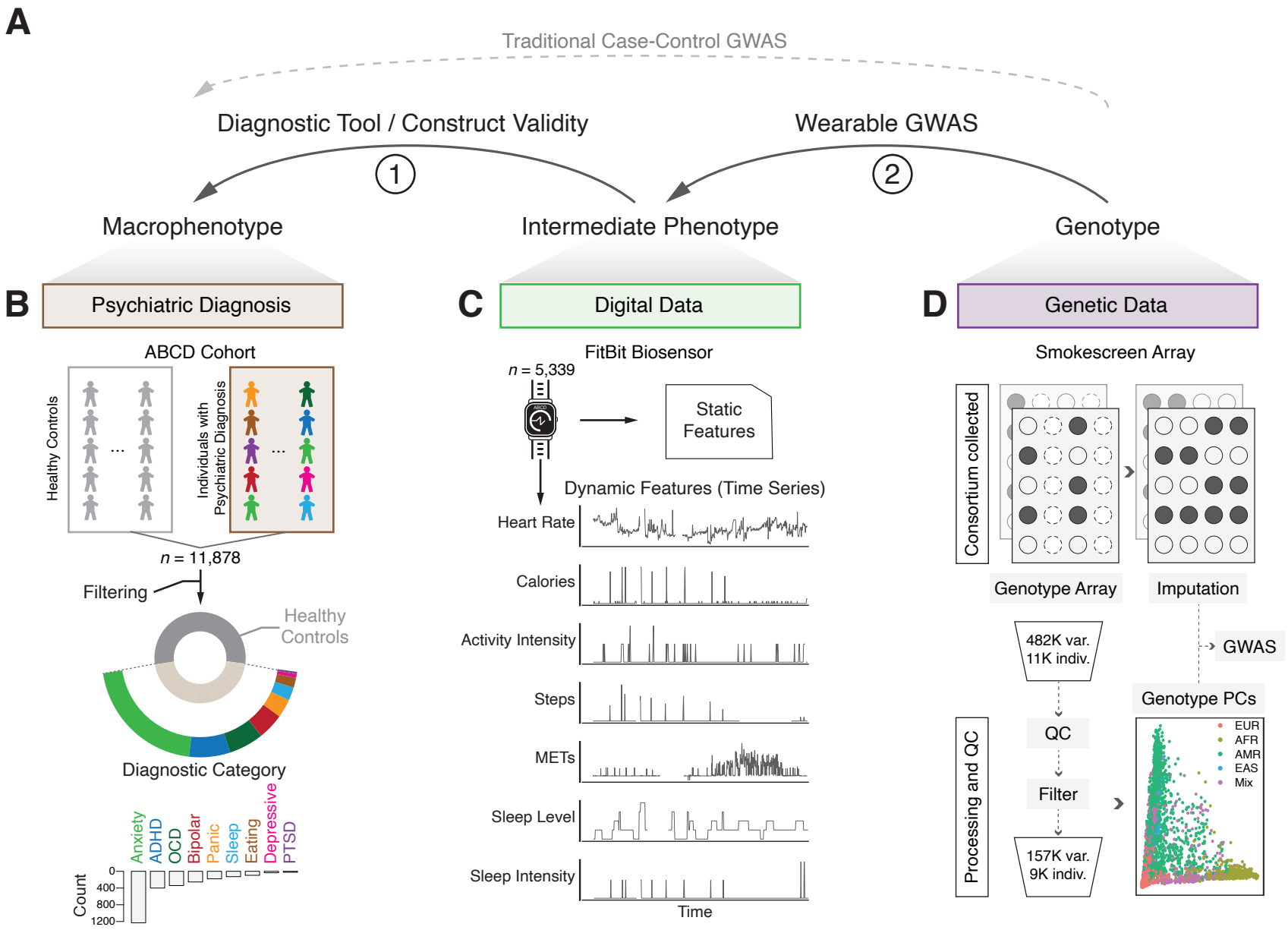
A
Traditional Case-Control GWAS
Diagnostic Tool / Construct Validity
Wearable GWAS
1
2
Macrophenotype
Intermediate Phenotype
Genotype
B
Psychiatric Diagnosis
ABCD Cohort
Healthy Controls
Individuals with Psychiatric Diagnosis
n = 11,878
Filtering
Healthy Controls
Diagnostic Category
Anxiety
ADHD
OCD
Bipolar
Panic
Sleep
Eating
Depressive
PTSD
Count
0
400
800
1200
C
Digital Data
FitBit Biosensor
n = 5,339
Static Features
Dynamic Features (Time Series)
Heart Rate
Calories
Activity Intensity
Steps
METs
Sleep Level
Sleep Intensity
Time
D
Genetic Data
Smokescreen Array
Consortium collected
Genotype Array
Imputation
GWAS
482K var.
11K indiv.
QC
Filter
157K var.
9K indiv.
Processing and QC
Genotype PCs
EUR
AFR
AMR
EAS
Mix

Figure 1

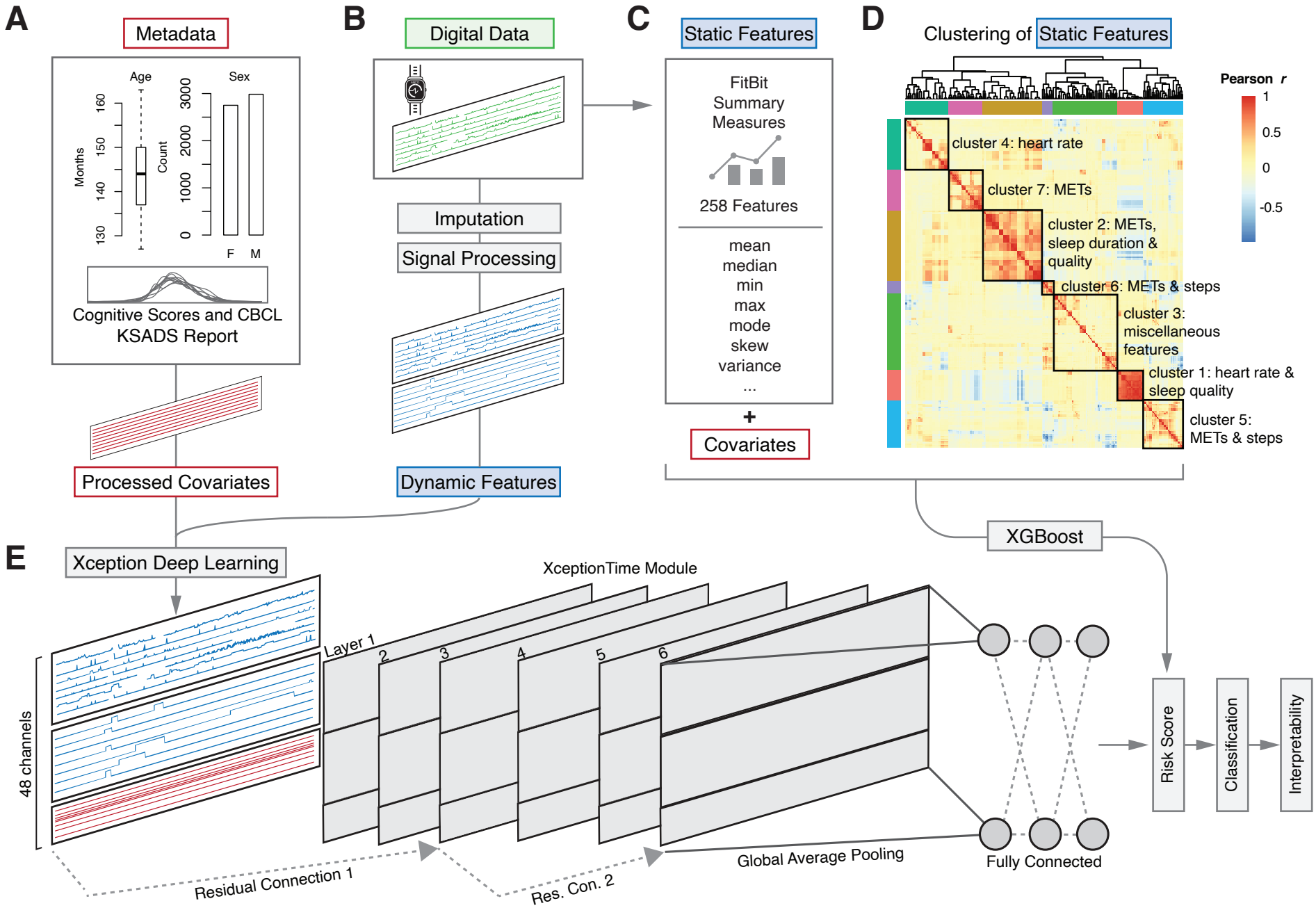
A
Metadata
Age
Sex
Months
Count
F
M
Cognitive Scores and CBCL
KSADS Report
Processed Covariates
B
Digital Data
Imputation
Signal Processing
Dynamic Features
C
Static Features
FitBit
Summary
Measures
258 Features
mean
median
min
max
mode
skew
variance
...
+
Covariates
D
Clustering of Static Features
Pearson r
cluster 4: heart rate
cluster 7: METs
cluster 2: METs,
sleep duration &
quality
cluster 6: METs & steps
cluster 3:
miscellaneous
features
cluster 1: heart rate &
sleep quality
cluster 5:
METs & steps
XGBoost
E
Xception Deep Learning
XceptionTime Module
48 channels
Layer 1
Residual Connection 1
Res. Con. 2
Global Average Pooling
Fully Connected
Risk Score
Classification
Interpretability

Figure 2

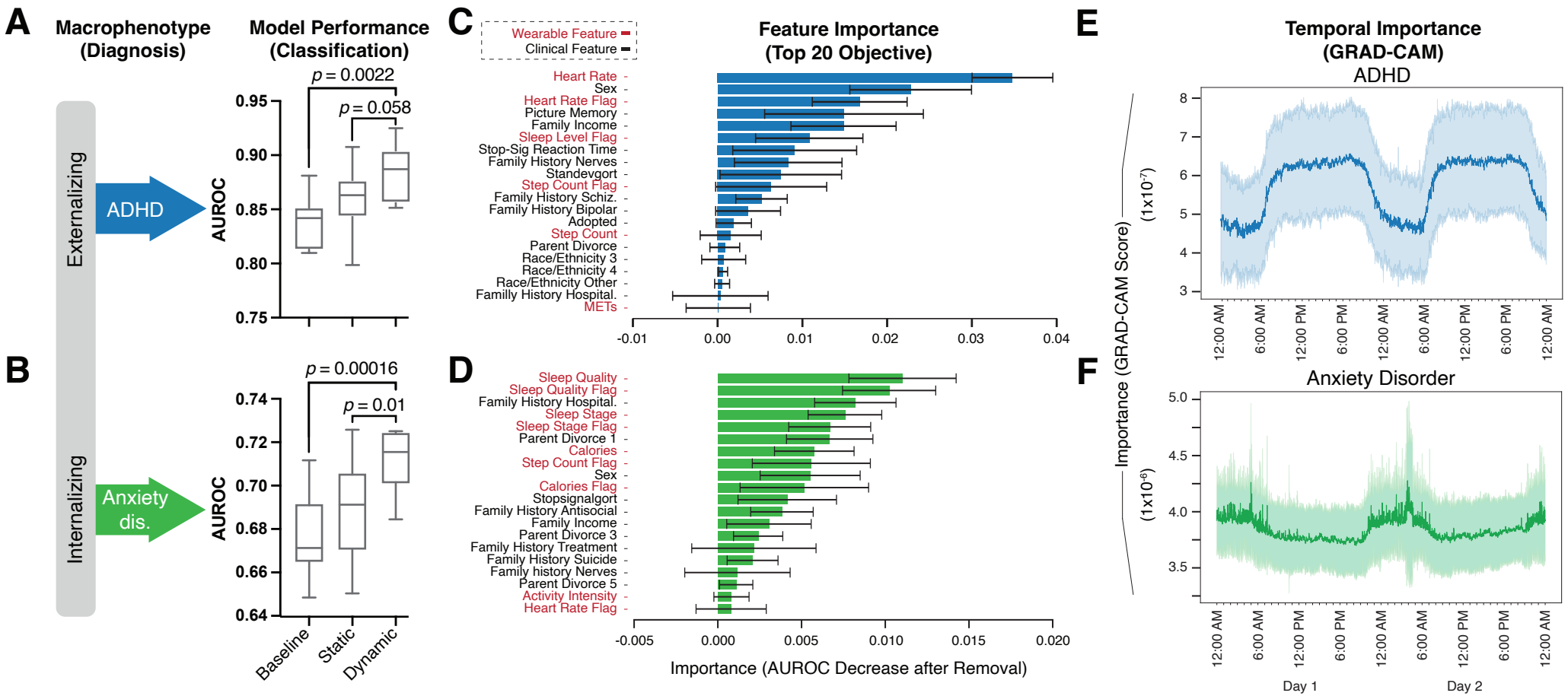
A
Macrophenotype (Diagnosis)
Model Performance (Classification)
Externalizing
ADHD
p = 0.0022
p = 0.058
AUROC
0.95
0.90
0.85
0.80
0.75
B
Internalizing
Anxiety dis.
p = 0.00016
p = 0.01
AUROC
0.74
0.72
0.70
0.68
0.66
0.64
Baseline
Static
Dynamic
C
Wearable Feature
Clinical Feature
Feature Importance (Top 20 Objective)
Heart Rate
Sex
Heart Rate Flag
Picture Memory
Family Income
Sleep Level Flag
Stop-Sig Reaction Time
Family History Nerves
Standevgort
Step Count Flag
Family History Schiz.
Family History Bipolar
Adopted
Step Count
Parent Divorce
Race/Ethnicity 3
Race/Ethnicity 4
Race/Ethnicity Other
Familly History Hospital.
METs
-0.01
0.00
0.01
0.02
0.03
0.04
D
Sleep Quality
Sleep Quality Flag
Family History Hospital.
Sleep Stage
Sleep Stage Flag
Parent Divorce 1
Calories
Step Count Flag
Sex
Calories Flag
Stopsignalgort
Family History Antisocial
Family Income
Parent Divorce 3
Family History Treatment
Family History Suicide
Family history Nerves
Parent Divorce 5
Activity Intensity
Heart Rate Flag
-0.005
0.000
0.005
0.010
0.015
0.020
Importance (AUROC Decrease after Removal)
E
Temporal Importance (GRAD-CAM)
ADHD
Importance (GRAD-CAM Score)
(1x10^-7)
8
7
6
5
4
3
12:00 AM
6:00 AM
12:00 PM
6:00 PM
12:00 AM
6:00 AM
12:00 PM
6:00 PM
12:00 AM
F
Anxiety Disorder
(1x10^-6)
5.0
4.5
4.0
3.5
Day 1
Day 2

Figure 3

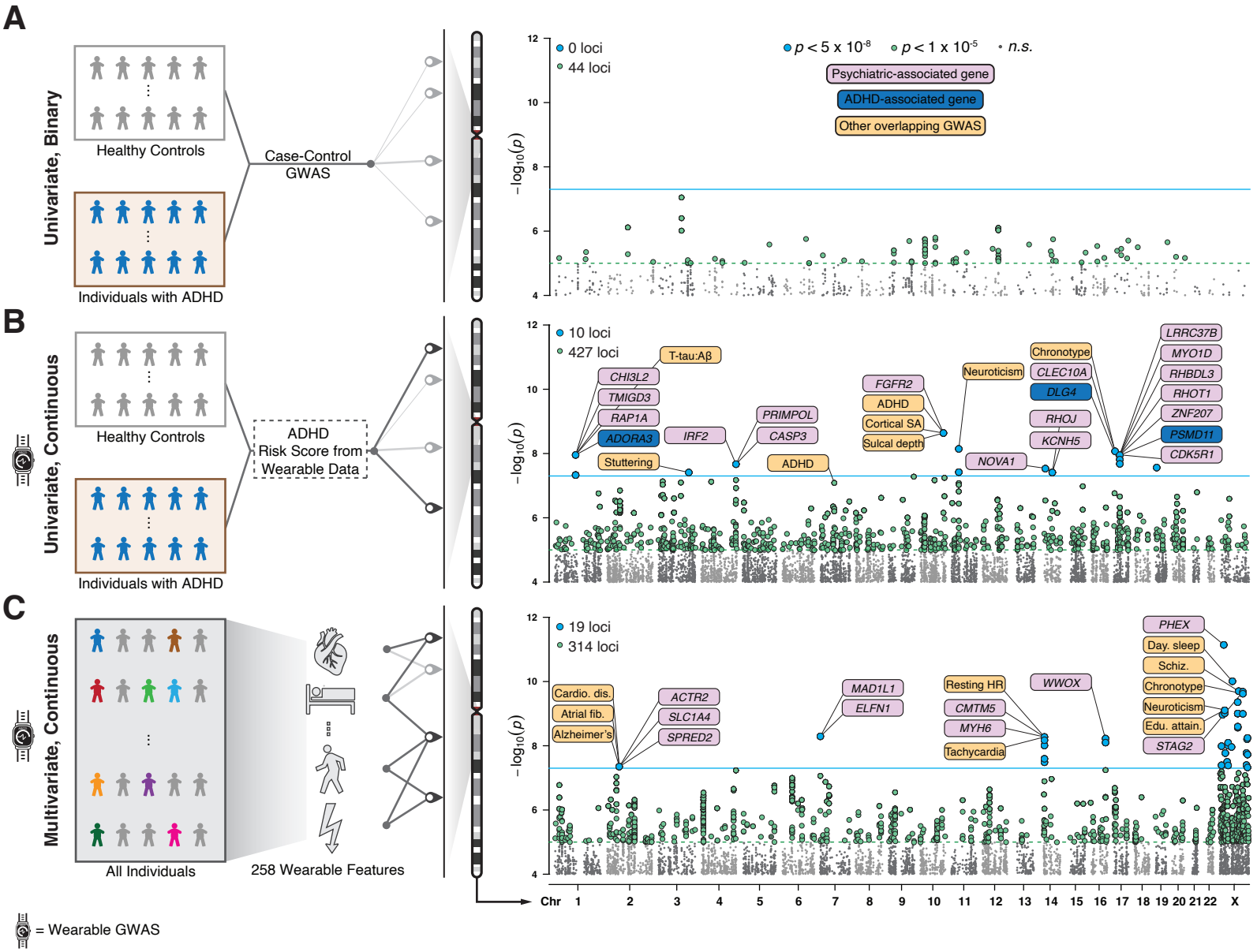
A
Univariate, Binary
Healthy Controls
Individuals with ADHD
Case-Control GWAS
0 loci
44 loci
p < 5 x 10-8
p < 1 x 10-5
n.s.
Psychiatric-associated gene
ADHD-associated gene
Other overlapping GWAS
−log10(p)
B
Univariate, Continuous
Healthy Controls
Individuals with ADHD
ADHD Risk Score from Wearable Data
10 loci
427 loci
T-tau:Aβ
CHI3L2
TMIGD3
RAP1A
ADORA3
Stuttering
IRF2
PRIMPOL
CASP3
ADHD
FGFR2
ADHD
Cortical SA
Sulcal depth
Neuroticism
NOVA1
Chronotype
CLEC10A
DLG4
RHOJ
KCNH5
LRRC37B
MYO1D
RHBDL3
RHOT1
ZNF207
PSMD11
CDK5R1
C
Multivariate, Continuous
All Individuals
258 Wearable Features
19 loci
314 loci
Cardio. dis.
Atrial fib.
Alzheimer's
ACTR2
SLC1A4
SPRED2
MAD1L1
ELFN1
Resting HR
CMTM5
MYH6
Tachycardia
WWOX
PHEX
Day. sleep
Schiz.
Chronotype
Neuroticism
Edu. attain.
STAG2
Chr 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
= Wearable GWAS

Figure 4

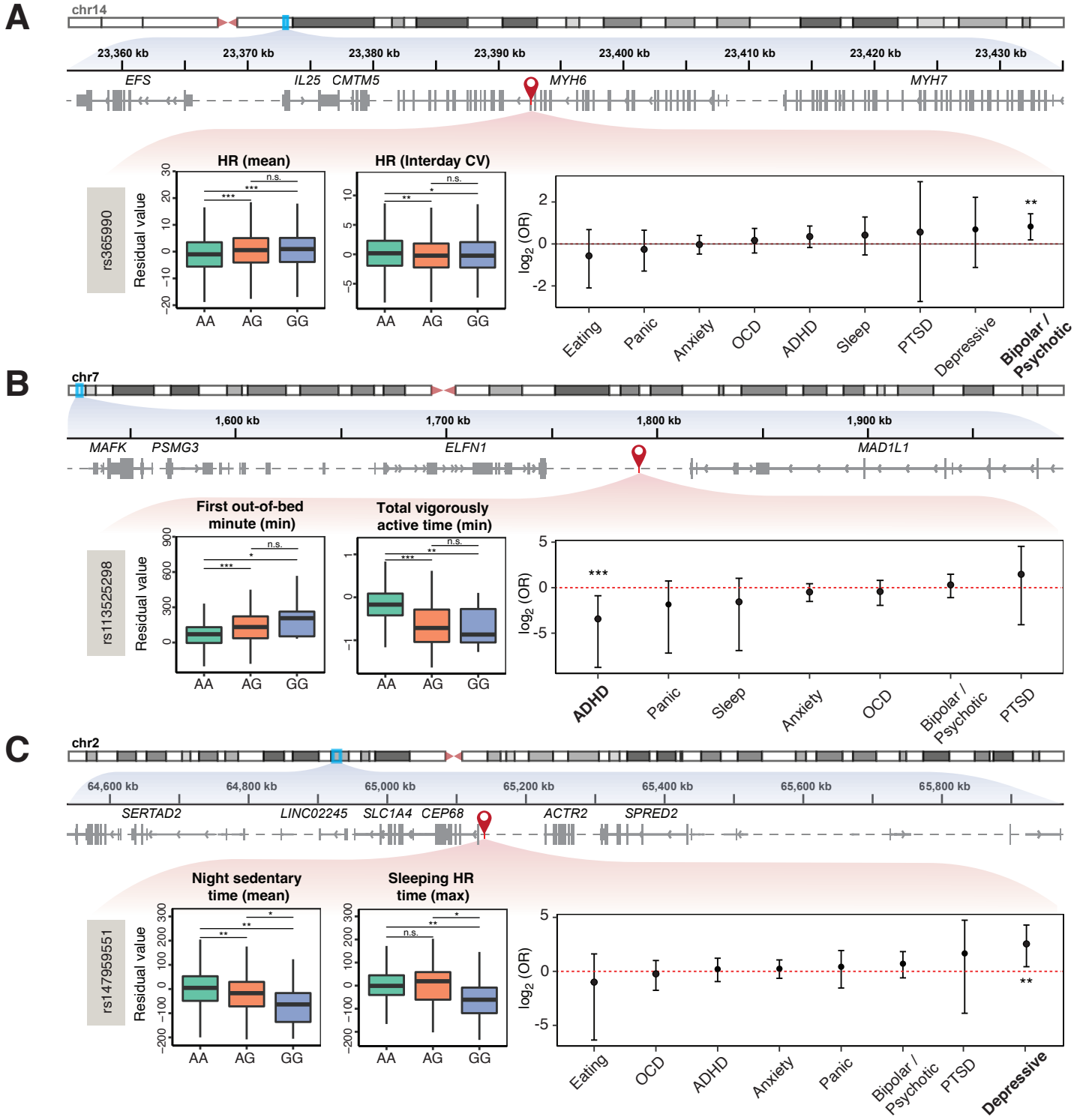
A
chr14
23,360 kb
23,370 kb
23,380 kb
23,390 kb
23,400 kb
23,410 kb
23,420 kb
23,430 kb
EFS
IL25
CMTM5
MYH6
MYH7
rs365990
HR (mean)
HR (Interday CV)
Residual value
AA
AG
GG
log2 (OR)
Eating
Panic
Anxiety
OCD
ADHD
Sleep
PTSD
Depressive
Bipolar / Psychotic
B
chr7
1,600 kb
1,700 kb
1,800 kb
1,900 kb
MAFK
PSMG3
ELFN1
MAD1L1
rs113525298
First out-of-bed minute (min)
Total vigorously active time (min)
C
chr2
64,600 kb
64,800 kb
65,000 kb
65,200 kb
65,400 kb
65,600 kb
65,800 kb
SERTAD2
LINC02245
SLC1A4
CEP68
ACTR2
SPRED2
rs147959551
Night sedentary time (mean)
Sleeping HR time (max)

Figure 5

Figure Captions

Figure 1

Leveraging clinical, digital, and genetic data of the ABCD cohort to improve characterization of psychiatric disorders.

A) Framework schematic describing how intermediate phenotypes from wearable-derived data are leveraged to better understand the association between macrophenotype and genotype. The link between intermediate phenotype and macrophenotype serves as construct validity and aid in diagnostics. Wearable GWAS is performed through genotype-to-intermediate-phenotype association studies. **B)** The Adolescent Brain Cognitive Development (ABCD) cohort contains 11,878 individuals spanning nine different categorical macrophenotypes based on clinical diagnosis from the Kiddie Schedule for Affective Disorders and Schizophrenia-5. A breakdown of the counts of each disorder is shown in the bottom bar graph, with anxiety disorder and ADHD being the most prevalent. "Bipolar" refers to bipolar or psychotic disorders. **C)** Digital data from FitBit biosensors are collected for 5,339 individuals. The collected time series data are then processed into dynamic and static features, with information spanning various physiological and higher order processes. **D)** Genetic data are collected by the ABCD consortium through Smokescreen genotyping array. Imputed genotypes are used for downstream GWAS analyses. The genotype arrays are subjected to best-practice processing and QC to ensure included individuals and SNPs are of high quality. PCA performed on 8,791 individuals and 157,556 genotyped SNPs reveals distinct ancestral clusters across the cohort and the inferred genotype principal components (PCs) are used as covariates in downstream analyses.

Figure 2

Workflow for data processing, feature engineering, and model architecture.

A) ABCD cohort metadata including various demographic features, cognitive test scores, and clinical characteristics are used as covariates and represent the input features used in our baseline comparison model. Features shown in this plot correspond to the filtered set of individuals with wearable data. **B)** Digital data collected by wearable biosensors are used to generate dynamic features after signal processing and imputation steps. Together with the processed covariates, these time series features represent the input features for the dynamic model. **C)** Summary statistics applied to digital data collected by wearables are used to generate a total of 258 static features. In addition to the covariates, these are the input features used in the static model. The static model leverages the machine learning framework, XGBoost, for downstream tasks such as risk score generation, classification, model interpretability, and wearable GWAS. **D)** Hierarchical clustering of the static features yields seven distinct physiological clusters of wearable data. **E)** The dynamic model is based on the Xception deep learning framework, and uses the generated 48 channels from the dynamic features and covariates as input into a convolution-like model. The architecture consists of six inception layers and residual connections. Global average pooling and a fully connected layer allow for similar downstream tasks as mentioned in C).

Figure 3

Performance and interpretability of psychiatric phenotype classification models.

A-B) Model performance for baseline, static, and dynamic models employed for classifying individuals with ADHD (blue, top) or individuals with anxiety disorder (green, bottom) versus healthy controls. *P* values were calculated using one-sided t-test. **C-D)** Feature importance based

on ablation studies for the dynamic model for ADHD (blue, top) and anxiety disorder (green, bottom) classification. Wearable-derived dynamic features are shown in red font and clinical features (covariates) are shown in black font. Feature importance is equivalent to the decrease in model performance (AUROC) after removal of the given feature. **E-F)** Temporal importance during a 48-hour period for dynamic features in ADHD (blue, top) or anxiety disorder (green, bottom) classification based on the GRAD-CAM interpretability module. Importance is represented as the GRAD-CAM score, based on each time points contribution towards model performance.

Figure 4

Manhattan plots summarizing the results of the univariate and multivariate wearable GWAS.

A) Left panel: Case-control GWAS on 1,191 individuals from the ABCD cohort. We employed the clinical diagnosis label as the univariate binary response variable for the GWAS (1 = "individual with ADHD", n = 137; 0 = "healthy control individual", n = 1,054). Right panel: Manhattan plot showing the $-\log_{10} p$ value of association between the genetic variants and the univariate binary response variable. No genetic variants passed the genome-wide significance threshold (p value < $5 \cdot 10^{-8}$; blue line). Genetic variants with a suggestive p value (< 10^{-5}) are represented as green dots. **B)** Analogous representation to panel A) using the wearable-derived risk scores for ADHD as univariate continuous response variable. The GWAS was performed on the same set of 1,191 individuals and using the same set of covariates as in panel A). 10 and 427 loci passed the p value thresholds of $5 \cdot 10^{-8}$ and $1 \cdot 10^{-5}$, respectively. Loci chr1:111,372,165-111,482,359, chr17:7,101,607-7,101,608, and chr17:32,256,997-32,283,356 are proximal to genes *ADORA3* (72 Kb), *DLG4* (86 Kb) and *PSMD11* (174 Kb)(highlighted in dark blue) respectively, which have been previously associated with ADHD. Other proximal genes related to other psychiatric disorders are highlighted in pink (evidence obtained from OpenTargets). Brain-related traits associated with genetic variants overlapping the ten genome-wide significant loci are highlighted in orange. GWAS associations were obtained from the EBI-NHGRI GWAS catalog. **C)** Analogous representation to panel A) using the 258 wearable-derived static features as multivariate continuous response variable. The GWAS was performed on a set of 2,410 individuals (both healthy controls and individuals with any disorder). 19 and 314 loci passed the p value thresholds of $5 \cdot 10^{-8}$ and $1 \cdot 10^{-5}$, respectively. Neuropsychiatric-related genes proximal to the identified loci are shown in pink. Brain-, heart-, and sleep-related traits with associated variants overlapping the 19 loci are highlighted in orange.

Figure 5

Exploring the genetic-physiological-psychiatric axis with wearable GWAS.

A) Left panel: rs365990 (chr14:23,392,602, A/G) is located in exon 25 of *MYH6*, and is associated with changes in wearable-derived heart rate features (multivariate GWAS p value = 5.33E-09). The boxplots show distributions of covariate-adjusted mean and interday coefficient of variation (CV) for heart rate across genotype groups at rs365990 (AA n individuals = 1,228; AG n individuals = 1,509; GG n individuals = 519). p values (two-sided Wilcoxon Rank-Sum test) for each pairwise comparison are also displayed, encoded as follows: *** ($p \leq 0.001$), ** ($0.001 < p \leq 0.01$), * ($0.01 < p \leq 0.05$), n.s. ($p > 0.05$). For visualization purposes, outliers are not shown. Right panel: enrichment, displayed as odds-ratio (log2 OR; y axis) of the minor allele (G) in individuals with different psychiatric disorders (x axis) compared to healthy controls. OR estimates

and 95% confidence interval (error bar) are displayed. The red horizontal dashed line indicates no enrichment. The G allele is significantly more enriched in individuals with bipolar/psychotic disorder compared to healthy controls (two-sided Fisher test *p* value: 8.00E-03; FDR-adjusted *p* value: 7.00E-02). **B)** Similar representation for rs113525298 (chr7:1,791,353; AA *n* individuals = 2,294; AG *n* individuals = 101; GG *n* individuals = 15). rs113525298 is located 125 Kb from *ELFN1*, a gene that encodes a postsynaptic protein involved in the temporal dynamics of interneuron recruitment[65,66]. *Elfn1* mutant mice exhibit hyperactivity that is treatable by psychostimulant medication[55,56]. The G allele at rs113525298 is associated with increased minimum number of first-out-of-bed minutes and decreased minimum number of total-vigorously-active minutes (multivariate GWAS *p* value = 5.10E-09), and is significantly more enriched in healthy controls compared to individuals with ADHD (two-sided Fisher test *p* value: 9.00E-04; FDR-adjusted *p* value: 6.00E-03). **C)** Similar representation for rs147959551 (chr2:65,140,366; AA n individuals = 2,279; AG n individuals = 117; GG n individuals = 14), located near a cluster of genes relevant for several psychiatric disorders, such as *ACTR2* (schizophrenia; 87 Kb), *SLC1A4* (schizophrenia, bipolar disorder, major depressive disorder; 117 Kb) and *SPRED2* (schizophrenia, OCD; 170 Kb)[67-77]. The G allele of rs147959551 is associated with decreased mean number of sedentary-time-at-night minutes and decreased maximum number of sleep-based-on-heart-rate minutes (multivariate GWAS *p* value = 4.47E-08), and is significantly more enriched in individuals with depressive disorder compared to healthy controls (two-sided Fisher test *p* value: 9.74E-03; FDR-adjusted *p* value: 7.80E-02).

References

1 Zablotsky, B., Terlizzi, E. P. & National Center for Health Statistics (U.S.). in *NCHS data brief,* 1 online resource (7 pages, 1 unnumbered page).

2 UNICEF. *Impact of COVID-19 on poor mental health in children and young people 'tip of the iceberg' – UNICEF*, <https://www.unicef.org/philippines/press-releases/impact-covid-19-poor-mental-health-children-and-young-people-tip-iceberg-unicef#:~:text=According%20to%20the%20latest%20available,needs%20and%20mental%20health%20funding> (2021).

3 CDC. *Data and Statistics on Children's Mental Health*, <https://www.cdc.gov/childrensmentalhealth/data.html> (2023).

4 McGorry, P. D. & Nelson, B. Transdiagnostic psychiatry: premature closure on a crucial pathway to clinical utility for psychiatric diagnosis. *World Psychiatry* **18**, 359-360 (2019). https://doi.org/10.1002/wps.20679

5 Hartmann, J. A. *et al.* Pluripotential Risk and Clinical Staging: Theoretical Considerations and Preliminary Data From a Transdiagnostic Risk Identification Approach. *Front Psychiatry* **11**, 553578 (2020). https://doi.org/10.3389/fpsyt.2020.553578

6 Insel, T. *et al.* Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *Am J Psychiat* **167**, 748-751 (2010). https://doi.org/10.1176/appi.ajp.2010.09091379

7 Flint, J., Timpson, N. & Munafo, M. Assessing the utility of intermediate phenotypes for genetic mapping of psychiatric disease. *Trends Neurosci* **37**, 733-741 (2014). https://doi.org/10.1016/j.tins.2014.08.007

8 Leuchter, A. F., Hunter, A. M., Krantz, D. E. & Cook, I. A. Intermediate phenotypes and biomarkers of treatment outcome in major depressive disorder. *Dialogues Clin Neurosci* **16**, 525-537 (2014). https://doi.org/10.31887/DCNS.2014.16.4/aleuchter

9 Sanchez-Roige, S. & Palmer, A. A. Emerging phenotyping strategies will advance our understanding of psychiatric genetics. *Nat Neurosci* **23**, 475-480 (2020). https://doi.org/10.1038/s41593-020-0609-7

10 Meyer-Lindenberg, A. & Weinberger, D. R. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nat Rev Neurosci* **7**, 818-827 (2006). https://doi.org/10.1038/nrn1993

11 Meyer-Lindenberg, A. Intermediate or brainless phenotypes for psychiatric research? *Psychol Med* **40**, 1057-1062 (2010). https://doi.org/10.1017/s0033291709991929

12 Koppe, G., Guloksuz, S., Reininghaus, U. & Durstewitz, D. Recurrent Neural Networks in Mobile Sampling and Intervention. *Schizophr Bull* **45**, 272-276 (2019). https://doi.org/10.1093/schbul/sby171

13 Gomes, N., Pato, M., Lourenco, A. R. & Datia, N. A Survey on Wearable Sensors for Mental Health Monitoring. *Sensors (Basel)* **23** (2023). https://doi.org/10.3390/s23031330

14 Dunn, J. *et al.* Wearable sensors enable personalized predictions of clinical laboratory measurements. *Nat Med* **27**, 1105-1112 (2021). https://doi.org/10.1038/s41591-021-01339-0

15 Insel, T. R. Digital Phenotyping: Technology for a New Science of Behavior. *JAMA* **318**, 1215-1216 (2017). https://doi.org/10.1001/jama.2017.11295

16 Cronbach, L. J. & Meehl, P. E. Construct validity in psychological tests. *Psychol Bull* **52**, 281-302 (1955). https://doi.org/10.1037/h0040957

17 Jablensky, A. Psychiatric classifications: validity and utility. *World Psychiatry* **15**, 26-31 (2016). https://doi.org/10.1002/wps.20284

18 Sen, B., Borle, N. C., Greiner, R. & Brown, M. R. G. A general prediction model for the detection of ADHD and Autism using structural and functional MRI. *PLoS One* **13**, e0194856 (2018). https://doi.org/10.1371/journal.pone.0194856

19 Kim, W. P. *et al.* Machine Learning-Based Prediction of Attention-Deficit/Hyperactivity Disorder and Sleep Problems With Wearable Data in Children. *JAMA Netw Open* **6**, e233502 (2023). https://doi.org/10.1001/jamanetworkopen.2023.3502

20 van Dijk, M. T., Murphy, E., Posner, J. E., Talati, A. & Weissman, M. M. Association of Multigenerational Family History of Depression With Lifetime Depressive and Other Psychiatric Disorders in Children: Results from the Adolescent Brain Cognitive Development (ABCD) Study. *JAMA Psychiatry* **78**, 778-787 (2021). https://doi.org/10.1001/jamapsychiatry.2021.0350

21 Anttila, V. *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018). https://doi.org/10.1126/science.aap8757

22 Mishra, T. *et al.* Pre-symptomatic detection of COVID-19 from smartwatch data. *Nat Biomed Eng* **4**, 1208-1220 (2020). https://doi.org/10.1038/s41551-020-00640-6

23 Liu, J. *et al.* Bayesian structural time series for biomedical sensor data: A flexible modeling framework for evaluating interventions. *PLoS Comput Biol* **17**, e1009303 (2021). https://doi.org/10.1371/journal.pcbi.1009303

24 Bent, B. *et al.* The digital biomarker discovery pipeline: An open-source software platform for the development of digital biomarkers using mHealth and wearables data. *J Clin Transl Sci* **5**, e19 (2020). https://doi.org/10.1017/cts.2020.511

25 Salari, N. *et al.* The global prevalence of ADHD in children and adolescents: a systematic review and meta-analysis. *Ital J Pediatr* **49**, 48 (2023). https://doi.org/10.1186/s13052-023-01456-1

26 Chen, T. Q. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. *Kdd'16: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785-794 (2016). https://doi.org/10.1145/2939672.2939785

27 Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. *Proc Cvpr Ieee*, 1800-1807 (2017). https://doi.org/10.1109/Cvpr.2017.195

28 Kofler, M. J. *et al.* Working memory and information processing in ADHD: Evidence for directionality of effects. *Neuropsychology* **34**, 127-143 (2020). https://doi.org/10.1037/neu0000598

29 Selvaraju, R. R. *et al.* Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *Ieee I Conf Comp Vis*, 618-626 (2017). https://doi.org/10.1109/Iccv.2017.74

30 Imeraj, L. *et al.* Time-of-day effects in arousal: disrupted diurnal cortisol profiles in children with ADHD. *J Child Psychol Psychiatry* **53**, 782-789 (2012). https://doi.org/10.1111/j.1469-7610.2012.02526.x

31 Alvaro, P. K., Roberts, R. M. & Harris, J. K. A Systematic Review Assessing Bidirectionality between Sleep Disturbances, Anxiety, and Depression. *Sleep* **36**, 1059-1068 (2013). https://doi.org/10.5665/sleep.2810

32 Grimm, O., Kranz, T. M. & Reif, A. Genetics of ADHD: What Should the Clinician Know? *Curr Psychiatry Rep* **22**, 18 (2020). https://doi.org/10.1007/s11920-020-1141-x

33 Purves, K. L. *et al.* A major role for common genetic variation in anxiety disorders. *Mol Psychiatry* **25**, 3292-3303 (2020). https://doi.org/10.1038/s41380-019-0559-1

34 Yang, L., Wang, S. & Altman, R. B. POPDx: an automated framework for patient phenotyping across 392 246 individuals in the UK Biobank study. *J Am Med Inform Assoc* **30**, 245-255 (2023). https://doi.org/10.1093/jamia/ocac226

35 Yang, L., Sadler, M. C. & Altman, R. B. Genetic association studies using disease liabilities from deep neural networks. *medRxiv* (2023). https://doi.org/10.1101/2023.01.18.23284383

36 Chen, Y. C. *et al.* Neuroanatomic, epigenetic and genetic differences in monozygotic twins discordant for attention deficit hyperactivity disorder. *Mol Psychiatry* **23**, 683-690 (2018). https://doi.org/10.1038/mp.2017.45

37 Karlsson Linner, R. *et al.* Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nat Neurosci* **24**, 1367-1376 (2021). https://doi.org/10.1038/s41593-021-00908-3

38 Fan, Z. *et al.* DLGAP1 and NMDA receptor-associated postsynaptic density protein genes influence executive function in attention deficit hyperactivity disorder. *Brain Behav* **8**, e00914 (2018). https://doi.org/10.1002/brb3.914

39 Royston, P., Altman, D. G. & Sauerbrei, W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* **25**, 127-141 (2006). https://doi.org/10.1002/sim.2331

40 Breitling, L. P. & Brenner, H. Odd odds interactions introduced through dichotomisation of continuous outcomes. *J Epidemiol Community Health* **64**, 300-303 (2010). https://doi.org/10.1136/jech.2009.089458

41 Schmitz, S., Adams, R. & Walsh, C. The use of continuous data versus binary data in MTC models: a case study in rheumatoid arthritis. *BMC Med Res Methodol* **12**, 167 (2012). https://doi.org/10.1186/1471-2288-12-167

42 Schandry, R. Heart beat perception and emotional experience. *Psychophysiology* **18**, 483-488 (1981). https://doi.org/10.1111/j.1469-8986.1981.tb02486.x

43 Hsueh, B. *et al.* Cardiogenic control of affective behavioural state. *Nature* **615**, 292-299 (2023). https://doi.org/10.1038/s41586-023-05748-8

44 Garrido-Martin, D., Calvo, M., Reverter, F. & Guigo, R. A fast non-parametric test of association for multiple traits. *Genome Biol* **24**, 230 (2023). https://doi.org/10.1186/s13059-023-03076-8

45 Consortium, E. P. *et al.* Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699-710 (2020). https://doi.org/10.1038/s41586-020-2493-4

46 Consortium, G. T. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318-1330 (2020). https://doi.org/10.1126/science.aaz1776

47 Landrum, M. J. *et al.* ClinVar: improvements to accessing data. *Nucleic Acids Res* **48**, D835-D844 (2020). https://doi.org/10.1093/nar/gkz972

48 Fu, T. *et al.* Association of the MYH6 Gene Polymorphism with the Risk of Atrial Fibrillation and Warfarin Anticoagulation Therapy. *Genet Test Mol Biomarkers* **25**, 590-599 (2021). https://doi.org/10.1089/gtmb.2021.0025

49 Rangaraju, A. *et al.* Genetic variants in post myocardial infarction patients presenting with electrical storm of unstable ventricular tachycardia. *Indian Pacing Electrophysiol J* **18**, 91-94 (2018). https://doi.org/10.1016/j.ipej.2018.01.003

50 Eijgelsheim, M. *et al.* Genome-wide association analysis identifies multiple loci related to resting heart rate. *Hum Mol Genet* **19**, 3885-3894 (2010). https://doi.org/10.1093/hmg/ddq303

51 Faurholt-Jepsen, M., Kessing, L. V. & Munkholm, K. Heart rate variability in bipolar disorder: A systematic review and meta-analysis. *Neurosci Biobehav Rev* **73**, 68-80 (2017). https://doi.org/10.1016/j.neubiorev.2016.12.007

52 Carr, O., de Vos, M. & Saunders, K. E. A. Heart rate variability in bipolar disorder and borderline personality disorder: a clinical review. *Evid Based Ment Health* **21**, 23-30 (2018). https://doi.org/10.1136/eb-2017-102760

53 Stautland, A. *et al.* Heart rate variability as biomarker for bipolar disorder. *medRxiv*, 2022.2002.2014.22269413 (2022). https://doi.org/10.1101/2022.02.14.22269413

54 Dieckmann, L., Cole, S. & Kumsta, R. Stress genomics revisited: gene co-expression analysis identifies molecular signatures associated with childhood adversity. *Transl Psychiatry* **10**, 34 (2020). https://doi.org/10.1038/s41398-020-0730-0

55 Hesdorffer, D. C. *et al.* ADHD as a risk factor for incident unprovoked seizures and epilepsy in children. *Arch Gen Psychiatry* **61**, 731-736 (2004). https://doi.org/10.1001/archpsyc.61.7.731

56 Dolan, J. & Mitchell, K. J. Mutation of Elfn1 in mice causes seizures and hyperactivity. *PLoS One* **8**, e80491 (2013). https://doi.org/10.1371/journal.pone.0080491

57 Chawner, S. *et al.* Sleep disturbance as a transdiagnostic marker of psychiatric risk in children with neurodevelopmental risk genetic conditions. *Transl Psychiatry* **13**, 7 (2023). https://doi.org/10.1038/s41398-022-02296-z

58 Arns, M., Kooij, J. J. S. & Coogan, A. N. Review: Identification and Management of Circadian Rhythm Sleep Disorders as a Transdiagnostic Feature in Child and Adolescent Psychiatry. *J Am Acad Child Adolesc Psychiatry* **60**, 1085-1095 (2021). https://doi.org/10.1016/j.jaac.2020.12.035

59 Wang, Z., Zhou, X., Gui, Y., Liu, M. & Lu, H. Multiple measurement analysis of resting-state fMRI for ADHD classification in adolescent brain from the ABCD study. *Transl Psychiatry* **13**, 45 (2023). https://doi.org/10.1038/s41398-023-02309-5

60 Obermeyer, Z., Powers, B., Vogeli, C. & Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**, 447-453 (2019). https://doi.org/10.1126/science.aax2342

61 Dane, A. V., Schachar, R. J. & Tannock, R. Does actigraphy differentiate ADHD subtypes in a clinical research setting? *J Am Acad Child Adolesc Psychiatry* **39**, 752-760 (2000). https://doi.org/10.1097/00004583-200006000-00014

62 Alfano, C. A., Ginsburg, G. S. & Kingery, J. N. Sleep-related problems among children and adolescents with anxiety disorders. *J Am Acad Child Adolesc Psychiatry* **46**, 224-232 (2007). https://doi.org/10.1097/01.chi.0000242233.06011.8e

63 Beauchaine, T. P. & Thayer, J. F. Heart rate variability as a transdiagnostic biomarker of psychopathology. *Int J Psychophysiol* **98**, 338-350 (2015). https://doi.org/10.1016/j.ijpsycho.2015.08.004

64 Singh, I. & Rose, N. Biomarkers in psychiatry. *Nature* **460**, 202-207 (2009). https://doi.org/10.1038/460202a

65 Dunn, H. A., Patil, D. N., Cao, Y., Orlandi, C. & Martemyanov, K. A. Synaptic adhesion protein ELFN1 is a selective allosteric modulator of group III metabotropic glutamate receptors in trans. *Proc Natl Acad Sci U S A* **115**, 5022-5027 (2018). https://doi.org/10.1073/pnas.1722498115

66 Girgenti, M. J. *et al.* Transcriptomic organization of the human brain in post-traumatic stress disorder. *Nat Neurosci* **24**, 24-33 (2021). https://doi.org/10.1038/s41593-020-00748-7

67 Datta, D., Arion, D., Roman, K. M., Volk, D. W. & Lewis, D. A. Altered Expression of ARP2/3 Complex Signaling Pathway Genes in Prefrontal Layer 3 Pyramidal Cells in Schizophrenia. *Am J Psychiatry* **174**, 163-171 (2017). https://doi.org/10.1176/appi.ajp.2016.16020204

68 Yao, W. *et al.* ERVWE1 Reduces Hippocampal Neuron Density and Impairs Dendritic Spine Morphology through Inhibiting Wnt/JNK Non-Canonical Pathway via miR-141-3p in Schizophrenia. *Viruses* **15** (2023). https://doi.org/10.3390/v15010168

69 Deng, X. *et al.* Association study of polymorphisms in the neutral amino acid transporter genes SLC1A4, SLC1A5 and the glycine transporter genes SLC6A5, SLC6A9 with schizophrenia. *BMC Psychiatry* **8**, 58 (2008). https://doi.org/10.1186/1471-244X-8-58

70 Li, Y. X. *et al.* Inhibitors of the Neutral Amino Acid Transporters ASCT1 and ASCT2 Are Effective in In Vivo Models of Schizophrenia and Visual Dysfunction. *J Pharmacol Exp Ther* **367**, 292-301 (2018). https://doi.org/10.1124/jpet.118.251116

71 Lin, E., Lin, C. H. & Lane, H. Y. Logistic ridge regression to predict bipolar disorder using mRNA expression levels in the N-methyl-D-aspartate receptor genes. *J Affect Disord* **297**, 309-313 (2022). https://doi.org/10.1016/j.jad.2021.10.081

72 Weis, S. *et al.* Changes in region- and cell type-specific expression patterns of neutral amino acid transporter 1 (ASCT-1) in the anterior cingulate cortex and hippocampus in schizophrenia, bipolar disorder and major depression. *J Neural Transm (Vienna)* **114**, 261-271 (2007). https://doi.org/10.1007/s00702-006-0544-0

73 Lee, P. H. *et al.* Multi-locus genome-wide association analysis supports the role of glutamatergic synaptic transmission in the etiology of major depressive disorder. *Transl Psychiatry* **2**, e184 (2012). https://doi.org/10.1038/tp.2012.95

74 Lin, C. H., Huang, M. W., Lin, C. H., Huang, C. H. & Lane, H. Y. Altered mRNA expressions for N-methyl-D-aspartate receptor-related genes in WBC of patients with major depressive disorder. *J Affect Disord* **245**, 1119-1125 (2019). https://doi.org/10.1016/j.jad.2018.12.016

75 Malki, K. *et al.* Highly polygenic architecture of antidepressant treatment response: Comparative analysis of SSRI and NRI treatment in an animal model of depression. *Am J Med Genet B Neuropsychiatr Genet* **174**, 235-250 (2017). https://doi.org/10.1002/ajmg.b.32494

76 Hepbasli, D. *et al.* Genotype- and Age-Dependent Differences in Ultrasound Vocalizations of SPRED2 Mutant Mice Revealed by Machine Deep Learning. *Brain Sci* **11** (2021). https://doi.org/10.3390/brainsci11101365

77 Jia, Y. F. *et al.* Astrocytic Glutamate Transporter 1 (GLT1) Deficiency Reduces Anxiety- and Depression-Like Behaviors in Mice. *Front Behav Neurosci* **14**, 57 (2020). https://doi.org/10.3389/fnbeh.2020.00057

78 Löning, M. *et al.* sktime: A Unified Interface for Machine Learning with Time Series. *33rd Conference on Neural Information Processing Systems (NeurIPS 2019)* (2019).

79 Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Adv Neur In* **32** (2019).
80 Oguiza, I. *tsai - A state-of-the-art deep learning library for time series and sequential data*, <https://github.com/timeseriesAI/tsai> (2022).
81 Marees, A. T. *et al.* A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int J Methods Psychiatr Res* **27**, e1608 (2018). https://doi.org/10.1002/mpr.1608
82 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575 (2007). https://doi.org/10.1086/519795
83 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015). https://doi.org/10.1186/s13742-015-0047-8
84 Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017). https://doi.org/10.1038/s41467-017-01261-5
85 Demontis, D. *et al.* Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat Genet* **55**, 198-208 (2023). https://doi.org/10.1038/s41588-022-01285-8
86 Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet* **51**, 63-75 (2019). https://doi.org/10.1038/s41588-018-0269-7
87 Di Tommaso, P. *et al.* Nextflow enables reproducible computational workflows. *Nat Biotechnol* **35**, 316-319 (2017). https://doi.org/10.1038/nbt.3820
88 Manrai, A. K., Ioannidis, J. P. A. & Patel, C. J. Signals Among Signals: Prioritizing Nongenetic Associations in Massive Data Sets. *Am J Epidemiol* **188**, 846-850 (2019). https://doi.org/10.1093/aje/kwz031
89 Kurilshikov, A. *et al.* Large-scale association analyses identify host factors influencing human gut microbiome composition. *Nat Genet* **53**, 156-165 (2021). https://doi.org/10.1038/s41588-020-00763-1
90 Frankish, A. *et al.* GENCODE: reference annotation for the human and mouse genomes in 2023. *Nucleic Acids Res* **51**, D942-D949 (2023). https://doi.org/10.1093/nar/gkac1071
91 Ochoa, D. *et al.* The next-generation Open Targets Platform: reimagined, redesigned, rebuilt. *Nucleic Acids Res* **51**, D1353-D1359 (2023). https://doi.org/10.1093/nar/gkac1046
92 Vosa, U. *et al.* Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet* **53**, 1300-1310 (2021). https://doi.org/10.1038/s41588-021-00913-z
93 Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362** (2018). https://doi.org/10.1126/science.aat8464
94 Bryois, J. *et al.* Cell-type-specific cis-eQTLs in eight human brain cell types identify novel risk genes for psychiatric and neurological disorders. *Nat Neurosci* **25**, 1104-1112 (2022). https://doi.org/10.1038/s41593-022-01128-z
95 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010). https://doi.org/10.1093/bioinformatics/btq033

96 Rozowsky, J. *et al.* The EN-TEx resource of multi-tissue personal epigenomes & variant-impact models. *Cell* **186**, 1493-1511 e1440 (2023). https://doi.org/10.1016/j.cell.2023.02.018

97 Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* **10**, e1004383 (2014). https://doi.org/10.1371/journal.pgen.1004383

98 Jansen, P. R. *et al.* Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet* **51**, 394-403 (2019). https://doi.org/10.1038/s41588-018-0333-3