nature medicine

Article

https://doi.org/10.1038/s41591-023-02440-2

Wearable movement-tracking data identify Parkinson's disease years before clinical diagnosis

In the format provided by the authors and unedited

Supplement

Supplemental Table 1: Overview of study population

For each diagnosis we show summary statistics for the diseased and prodromal groups and their respective unaffected control matches as well as the whole cohort of unaffected controls. We report the mean and standard deviation of the average acceleration (no wear-time bias corrected, field 90087), age at accelerometer data collection, age at baseline data collection, and time to PD diagnosis. We also show the percent of male, depressed and PD cases in each group and the total size of the group.

Supplemental Table 2: Codes for the diseases

For each disease and prodromal symptom, we show the associated ICD10, ICD9, read codes and the code used in UKBB for self-report. The ICD9 and ICD10 codes were extracted with phecodes (Wu et al., 2018) and if available the UKBB curated list of codes. The read codes were mapped to the ICD10 codes using the TRUD NHS Read browser.

Supplemental Table 3: Differences in acceleration for every hour of the day

We report the statistics of the two-sided T-tests (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) for the comparison of every hour of the day between prodromal Parkinson's disease (PD) cases and their age- and sex-matched unaffected controls as well as diagnosed PD cases and their matched controls. Significance at 0.05 with Bonferroni correction is reached for p-values smaller than 5.2×10^{-3}.

Supplemental Table 4: Significant association of covariates with average acceleration

For the covariates we report the association with average acceleration ('No_wear_time_bias_adjusted_average_acceleration', field 90087). For the real valued covariates, age at accelerometer data collection and body mass index (BMI), we report the statistics for the pearson correlation and for the binary covariate, male sex, we report the statistics for the Mann-Whitney U together with their p-values (0.05 Bonferroni-corrected threshold: 0.0166).

Supplemental Table 5: Performance metrics for each of the models

For each model we show the mean and standard deviation of the area under precision recall curve (AUPRC) and area under receiver operator curve (AUROC) on the test sets of the five outer cross-validation folds.

Supplemental Table 6: Significant differences in performance

For each trained logistic regression model, we show the test statistics of the two-sided T-tests (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) of the area under precision recall curve (AUPRC) across the five outer folds of the nested cross-validation (N=5).

Supplemental Table 7: Performance metrics for each survival model

For each modality-specific survival model and the combined model, we show the mean, standard deviation, and standard error of the time-dependent area under receiver operator curve (AUROC) over the test sets of the five outer cross-validation folds (N=5).

Supplemental Table 8: Significant differences in performance for the survival models

For each modality-specific model and the combined models, we show the statistics of the two-sided T-tests (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) of the mean area under receiver operator curve (AUROC) on the tests sets across the five outer folds of the nested cross-validation (N=5).

Supplemental Table 9: Features for the prediction models

For each group (prodromal Parkinson's disease (PD), diagnosed PD, unaffected control matched for prodromal PD, unaffected control matched for diagnosed PD, all unaffected controls, population) we report the mean, standard deviation, and sample size for each predictor included in the logistic regression models (prodromal symptoms beforePD) and the survival models (prodromal symptoms beforeacc).

Supplemental Table 10: Group comparison for average acceleration

We report the two-sided T-test statistics (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) of the group comparisons for residual (age-, BMI-, and sex-corrected through unaffected control cohort) no-wear time bias corrected average acceleration after removal of cases diagnosed with comorbid depression or Parkinson's disease.

Supplemental Table 11: Group comparison for consecutive sleep

We report the two-sided T-test statistics (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) of the group comparisons for residual (age-, BMI-, and sex-corrected through unaffected control cohort) mean maximum consecutive sleep [hours] after removal of cases diagnosed with comorbid depression or Parkinson's disease.

Supplemental Table 12: Group comparison for daytime sleeping

We report the two-sided T-test statistics (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) of the group comparisons for residual (age-, BMI-, and sex-corrected through unaffected control cohort) mean number of naps during daytime (7am-11pm) after removal of cases diagnosed with comorbid depression or Parkinson's disease.

Supplemental Table 13: Group comparison for sleep duration

We report the two-sided T-test statistics (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) of the group comparisons for residual (age-, BMI-, and sex-corrected through unaffected control cohort) mean sleep duration [hours] after removal of cases diagnosed with comorbid depression or Parkinson's disease.

Supplemental Table 14: Group comparison for night-time waking

We report the two-sided T-test statistics (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) of the group comparisons for residual (age-, BMI-, and sex-corrected through unaffected control cohort) mean number of wakeups during night-time (11pm-7am) after removal of cases diagnosed with comorbid depression or Parkinson's disease.

Supplemental Table 15: Group comparison for variability of acceleration during sleep

We report the two-sided T-test statistics (T, degrees of freedom (dof), alternative, p-value, 95% CI, cohen's d, Bayes factor, power) of the group comparisons for residual (age-, BMI-, and sex-corrected through unaffected control cohort) standard deviation of acceleration during epochs classified as sleep after removal of cases diagnosed with comorbid depression or Parkinson's disease.

Supplemental Table 16: Description of derived physical activity features

A list of the calculated physical activity features included in the accelerometry modality additional to the features obtained from UKBB. A description for each is provided alongside an indicator to which activity class the feature belongs.

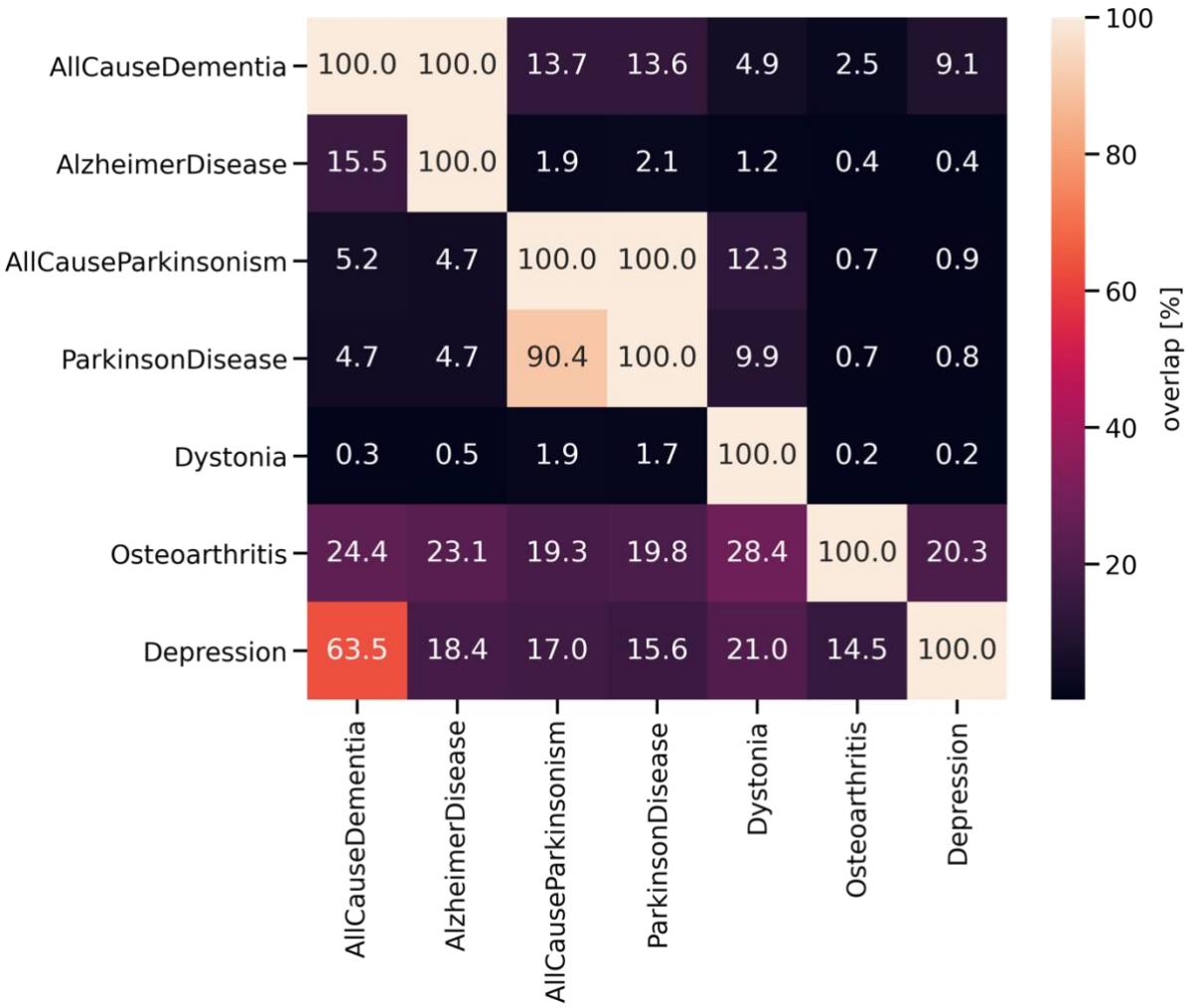
Supplemental Table 17: List of read codes for Anti-Parkinsonism drugs

A list of the read-codes used to identify individuals who were prescribed antiparkinsonism drugs at any timepoint. This list was extracted from the panel 2 UKBB provided algorithmically defined outcomes.

Supplemental Table 18: Calibration metrics for all logistic regression models

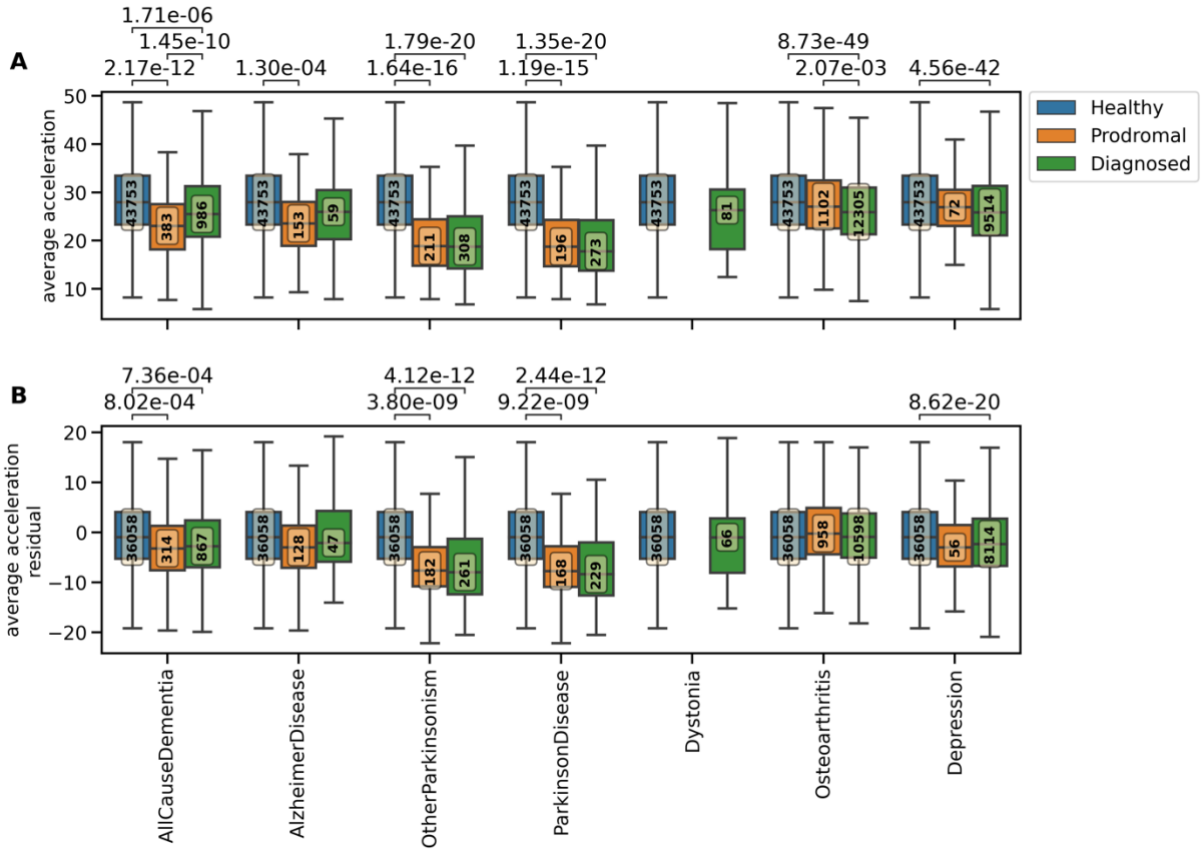
We report the calibration at large with the mean predicted probability and the prevalence, the calibration intercept and slope, and the binned mean predicted probability and fraction of positives

used to plot the calibration curve. This is provided for every outer cross-validation fold for each logistic regression model.



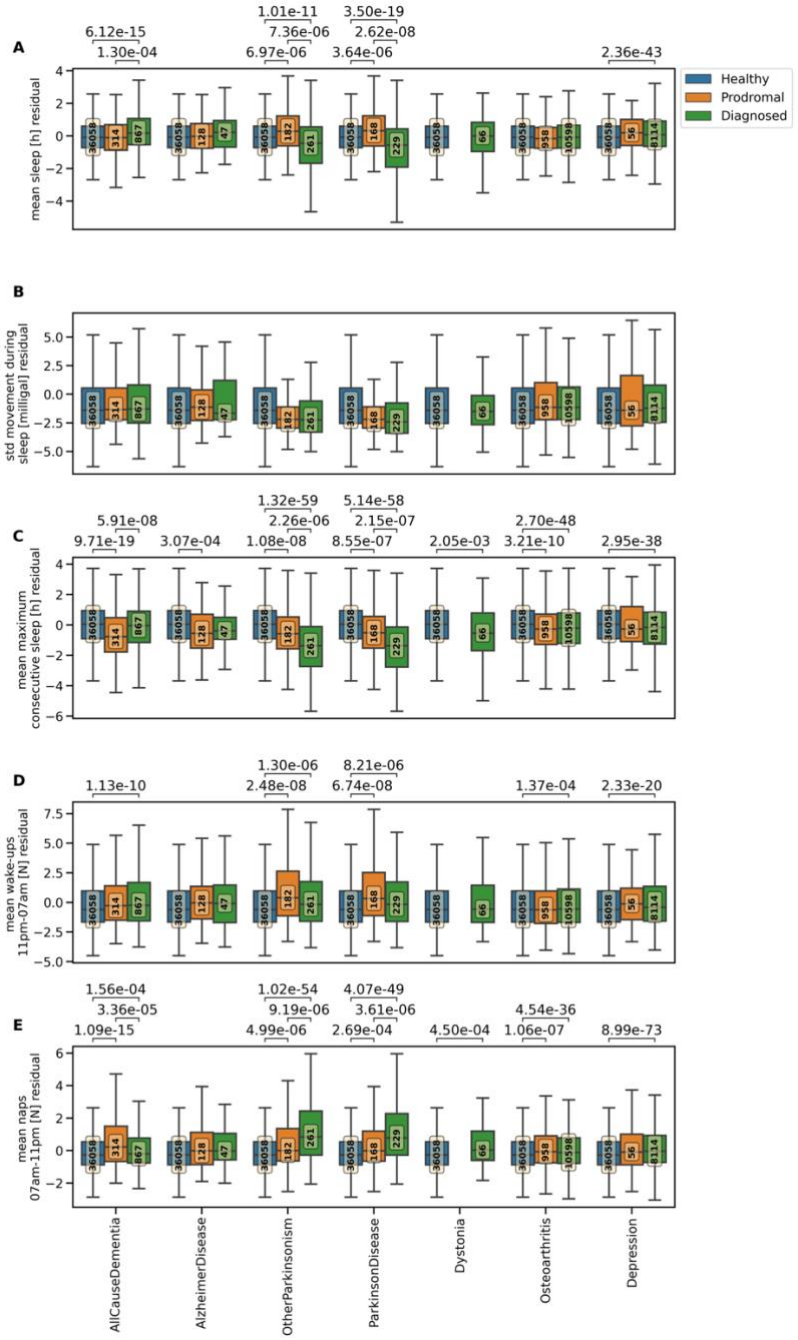
Supplemental Figure 1: Comorbidities in the UK Biobank.

Each column indicates the percentage of overlap of a disorder with the other disorders. For example, of all ‘AllCauseDementia’ subjects 63.5% also have a diagnosis of depression and 4.7% also a diagnosis of PD.



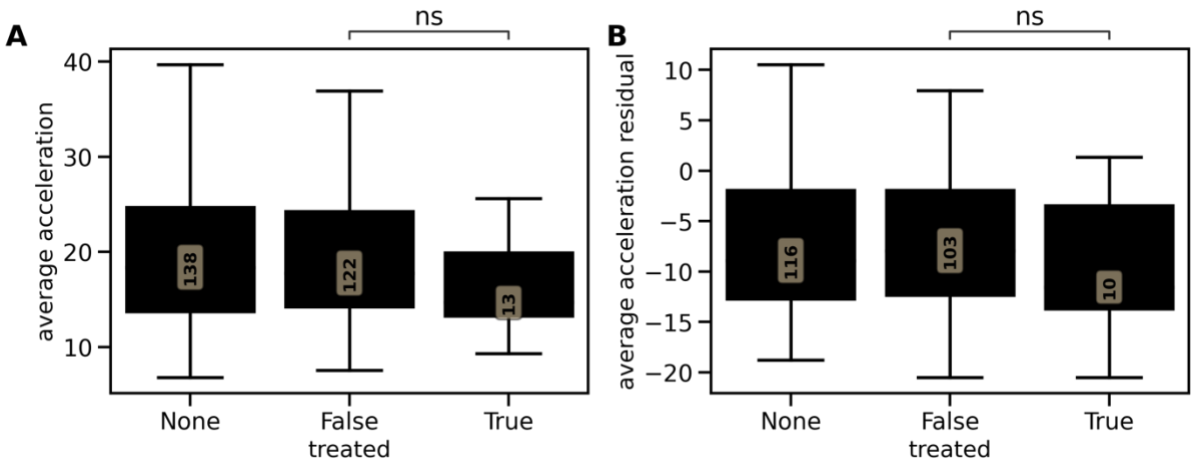
Supplemental Figure 2: Reduced acceleration across diseases

Boxplots for no wear-time bias adjusted average acceleration [A] before and [B] after correction for covariates are shown for seven disease groups and unaffected controls displaying the mean as the centre, the 25% and 75% quartiles as the bounds of the box, and the Q3 + 1.5*IQR/ Q1 - 1.5*IQR as the whiskers. For each disease group we differentiate between diagnosed (green), prodromal (orange), and unaffected (blue). The correction for covariates affects the number of subjects in each group (number in yellow box) as subjects with incomplete information about covariates were excluded. The correction includes subtracting recovered effects for each covariate from the variable of interest, thus changing the scale of the variable (y-axis). P-values are shown for group differences (two sided T-test, two sided Welch T-test when comparison with Healthy group) where 0.05 Bonferroni adjusted significance is reached at 2.38×10^{-3}.



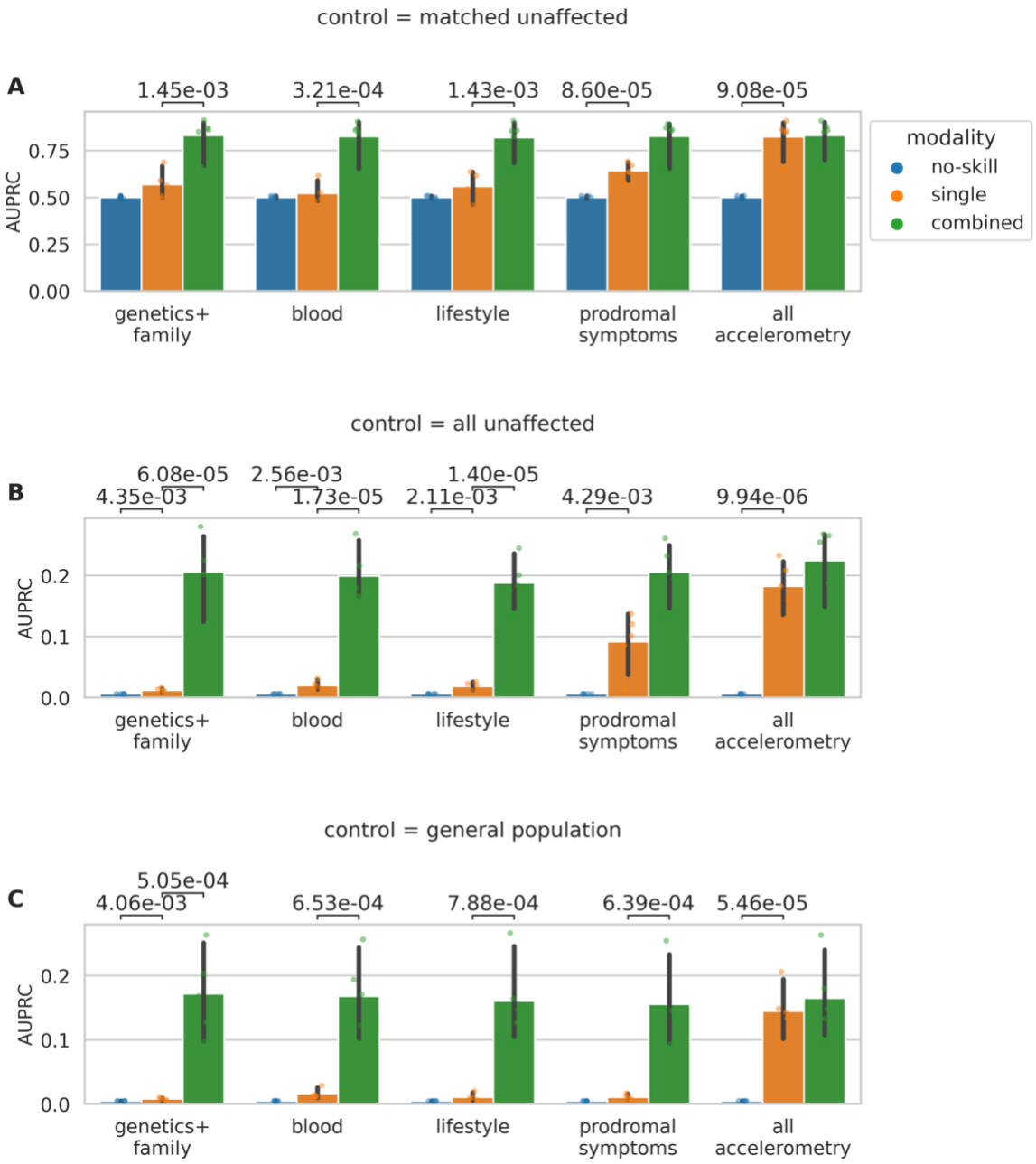
Supplemental Figure 3: Quality and length of sleep in several disorders

Group comparisons between prodromal, diagnosed, and unaffected controls performed with two sided T-tests (two sided Welch T-test when comparison with Healthy group) are shown before removal of comorbid depressed and PD subjects. The measures are age, BMI, and sex corrected through parameters learned from the unaffected population thus the residual is displayed and does not reflect the true value of the variable leading to potentially negative values. The boxplots show the mean as the centre, the 25% and 75% quartiles as the bounds of the box, and the Q3 + 1.5*IQR/ Q1 - 1.5*IQR as the whiskers. The yellow boxes show the number of subjects in each group. The p-values are shown when 0.05 Bonferroni-corrected threshold is reached at 2.38×10^{-3}.



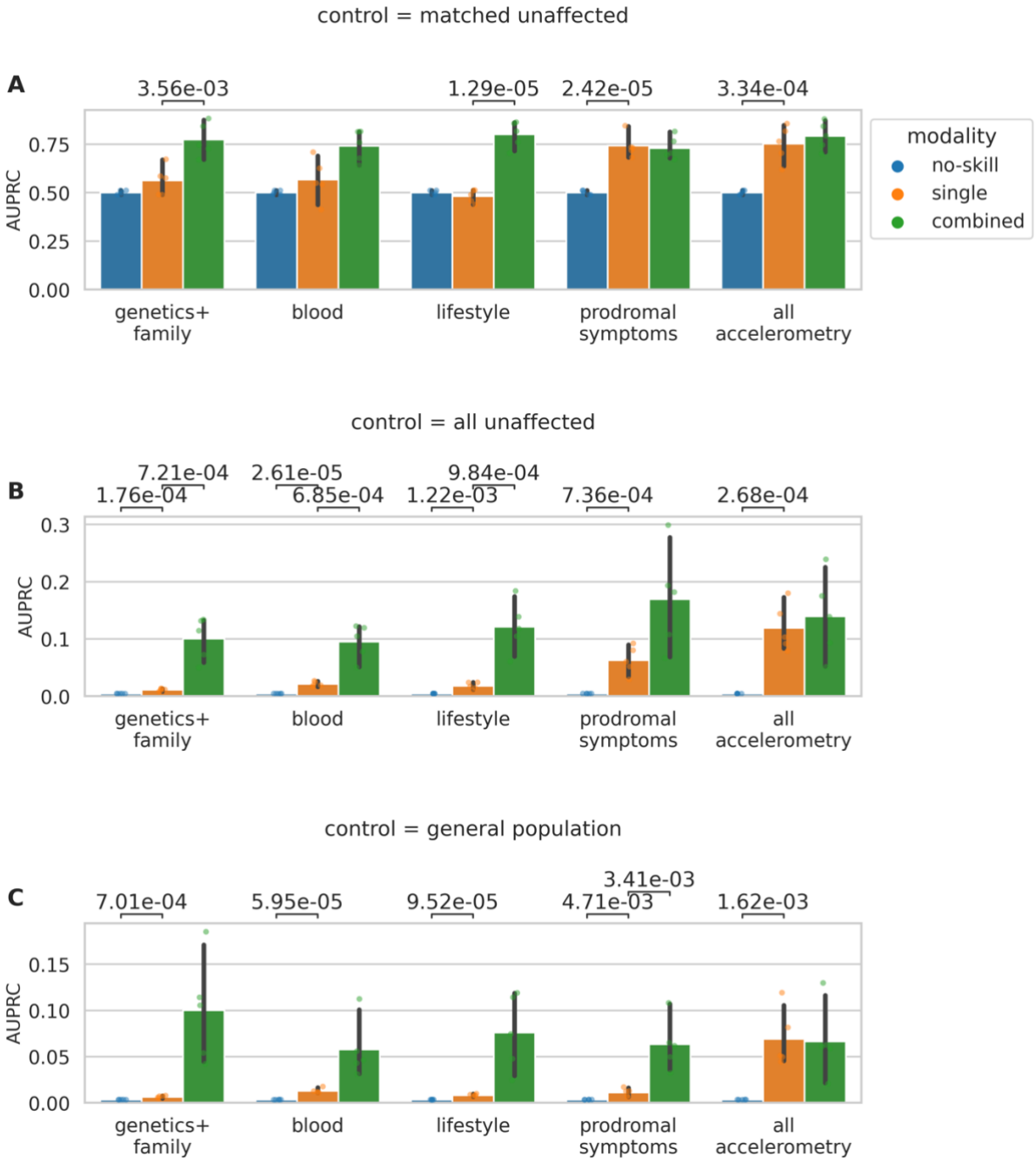
Supplemental Figure 4: Medication effect on acceleration

We plot the [A] average acceleration and the [B] residual average acceleration corrected for age, sex, and BMI as learned from unaffected controls for treated (N = 13, N = 10) versus non-treated diagnosed Parkinson's disease cases (N = 122, N = 103). Treated means here that no more than 10 weeks prior to accelerometry data collection a prescription for antiparkinsonism drugs was found in the associated primary care records. Subject with no GP prescription records available were excluded here (N = 138, N = 116). The boxplots show the mean as the centre, the 25% and 75% quartiles as the bounds of the box, and the Q3 + 1.5*IQR/ Q1 - 1.5*IQR as the whiskers. Group comparison was performed with two-sided T-tests with a significance threshold of 0.05.



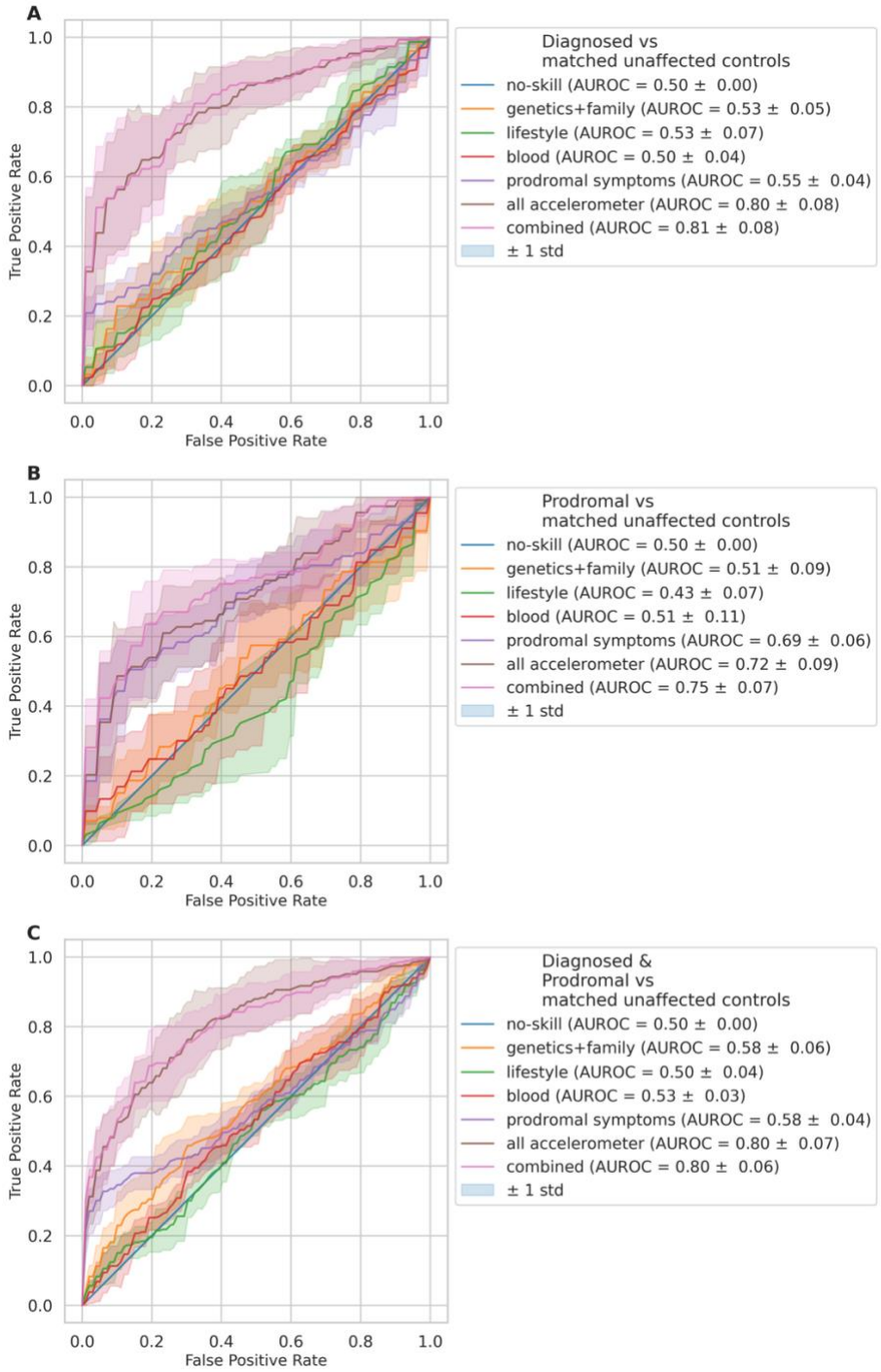
Supplemental Figure 5: Model comparison to identify diagnosed Parkinson's disease

The performance comparisons with two-sided T-tests (N=5) for each modality are shown. We compare no-skill (intercept) with single (modality-specific) and single with combined (modality + accelerometry modality) in terms of the area under precision recall curve (AUPRC). For the accelerometry modality, the combined model is the one merging all modalities. We show this for each control group setting [A-C]. The barplots show the mean performance across 5-folds and the error bars indicate the 95% Bonferroni-adjusted confidence interval. The dots display the individual performances per fold. P-values are shown when 0.05 Bonferroni-corrected significance is reached at $5x10^{-3}$.



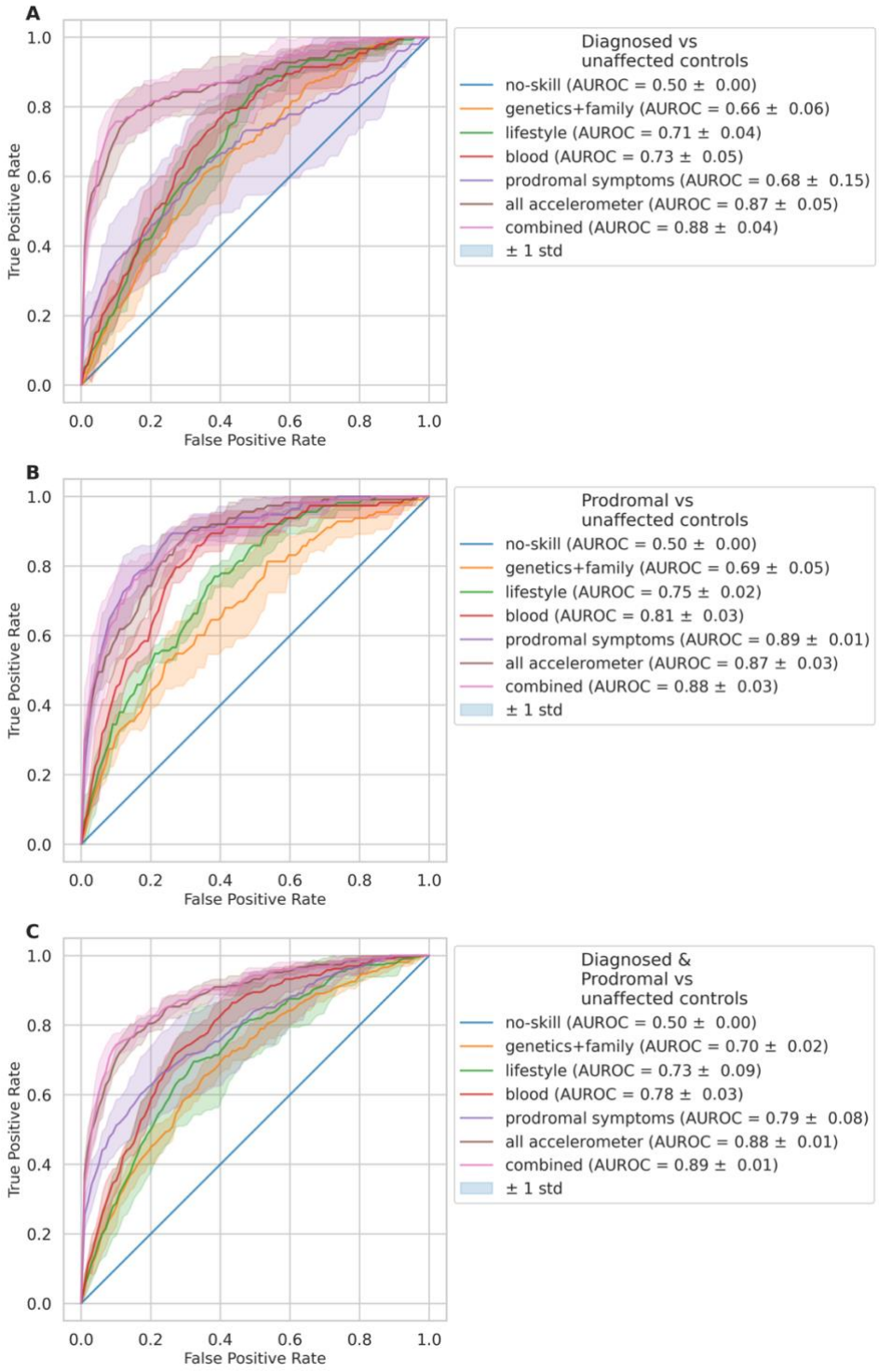
Supplemental Figure 6: Model comparison to identify prodromal Parkinson's disease

The performance comparisons with two-sided T-tests (N=5) for each modality are shown. We compare no-skill (intercept) with single (modality-specific) and single with combined (modality + accelerometry modality) in terms of the area under precision recall curve (AUPRC). For the accelerometry modality, the combined model is the one merging all modalities. We show this for each control group setting [A-C]. The barplots show the mean performance across 5-folds and the error bars indicate the 95% Bonferroni-adjusted confidence interval. The dots display the individual performances per fold. P-values are shown when 0.05 Bonferroni-corrected significance is reached at $5x10^{-3}$.



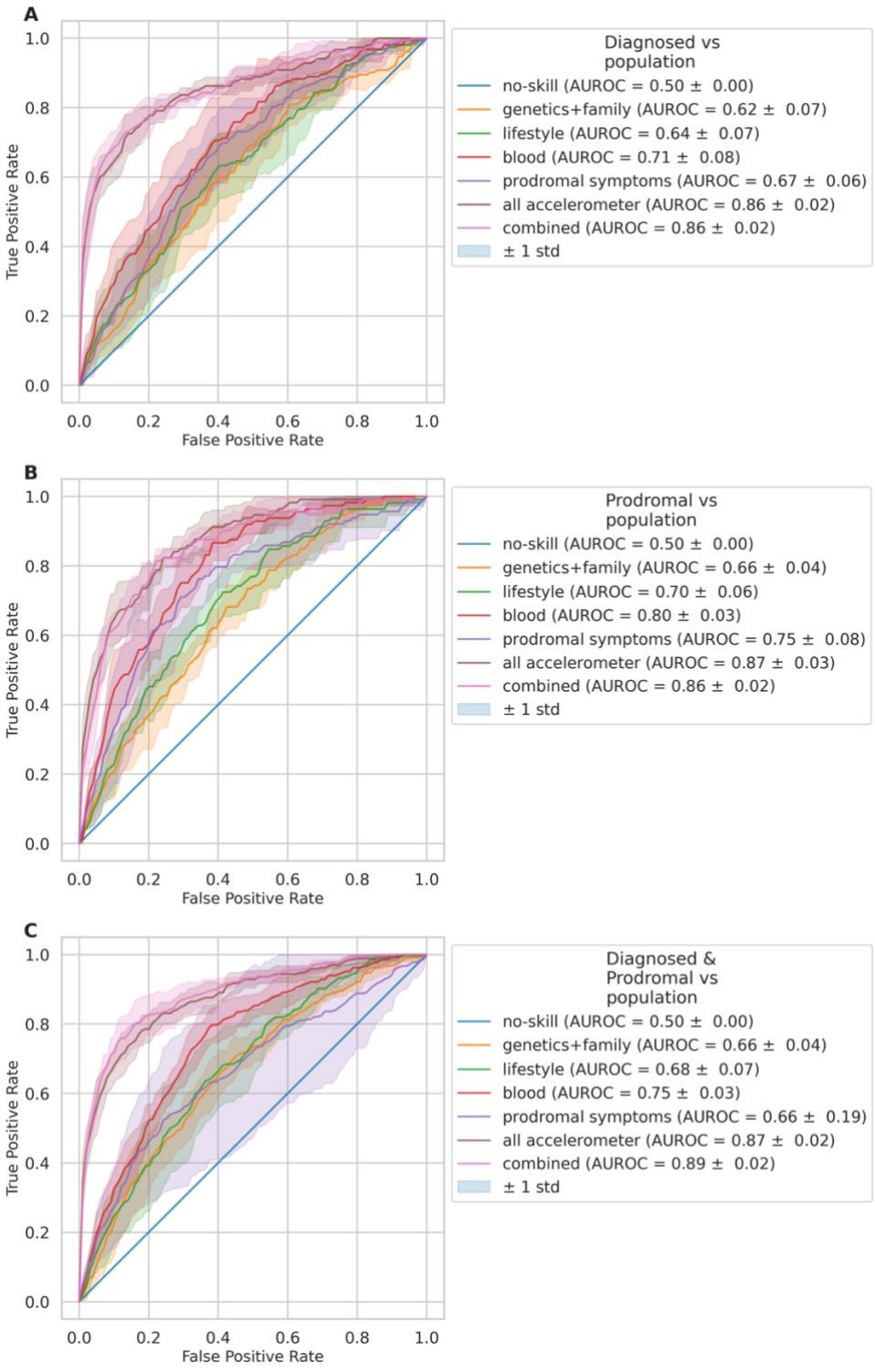
Supplemental Figure 7: Receiver operator curves for matched healthy control models

The mean receiver operator curves across the outer five-folds are shown together with their standard deviation. We show this for each model type: [A] identifying diseased cases, [B] identifying prodromal cases, [C] identifying diseased and prodromal cases from matched healthy controls for each tested modality.



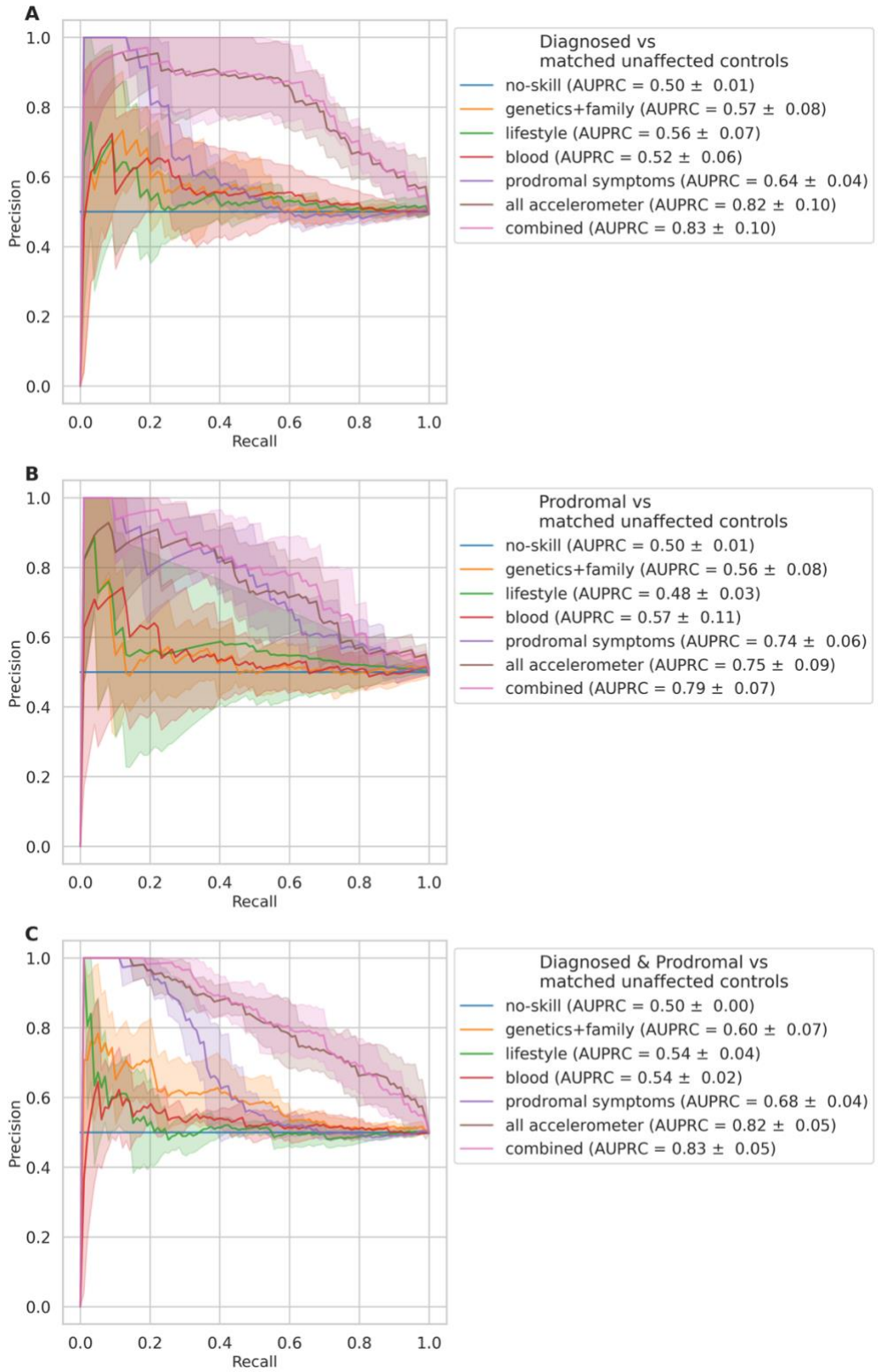
Supplemental Figure 8: Receiver operator curves for unmatched healthy control models

The mean receiver operator curves across the outer five-folds are shown together with their standard deviation. We show this for each model type: [A] identifying diseased cases, [B] identifying prodromal cases, [C] identifying diseased and prodromal cases from unmatched healthy controls for each tested modality.



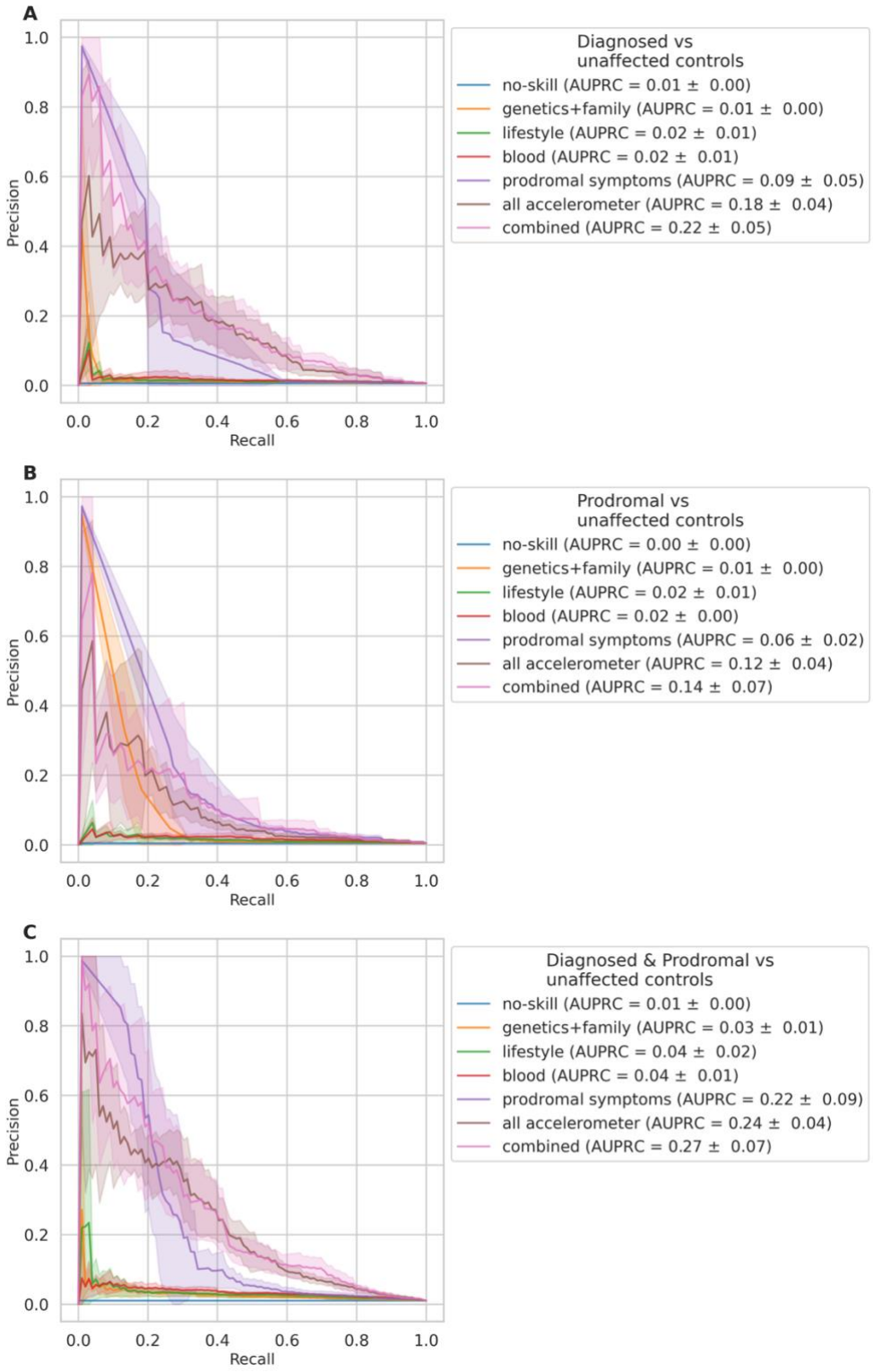
Supplemental Figure 9: Receiver operator curves for population-based models

The mean receiver operator curves across the outer five-folds are shown together with their standard deviation. We show this for each model type: [A] identifying diseased cases, [B] identifying prodromal cases, [C] identifying diseased and prodromal cases from the population for each tested modality.



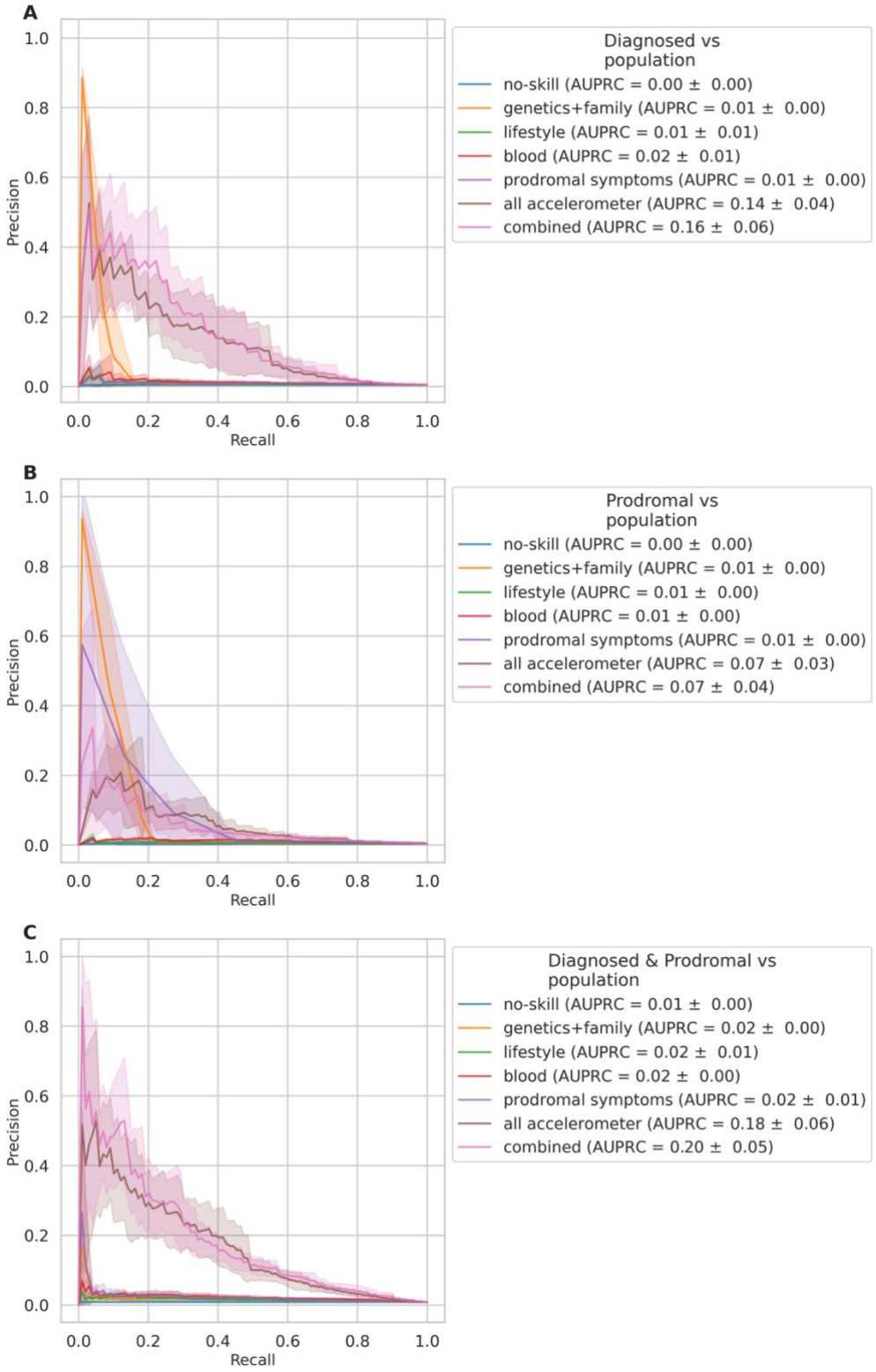
Supplemental Figure 10: Precision recall curves for matched healthy control models

The mean precision recall curves across the outer five-folds are shown with their standard deviation. We show this for each model type: [A] identifying diseased cases, [B] identifying prodromal cases, [C] identifying diseased and prodromal cases from matched healthy controls for each tested modality.



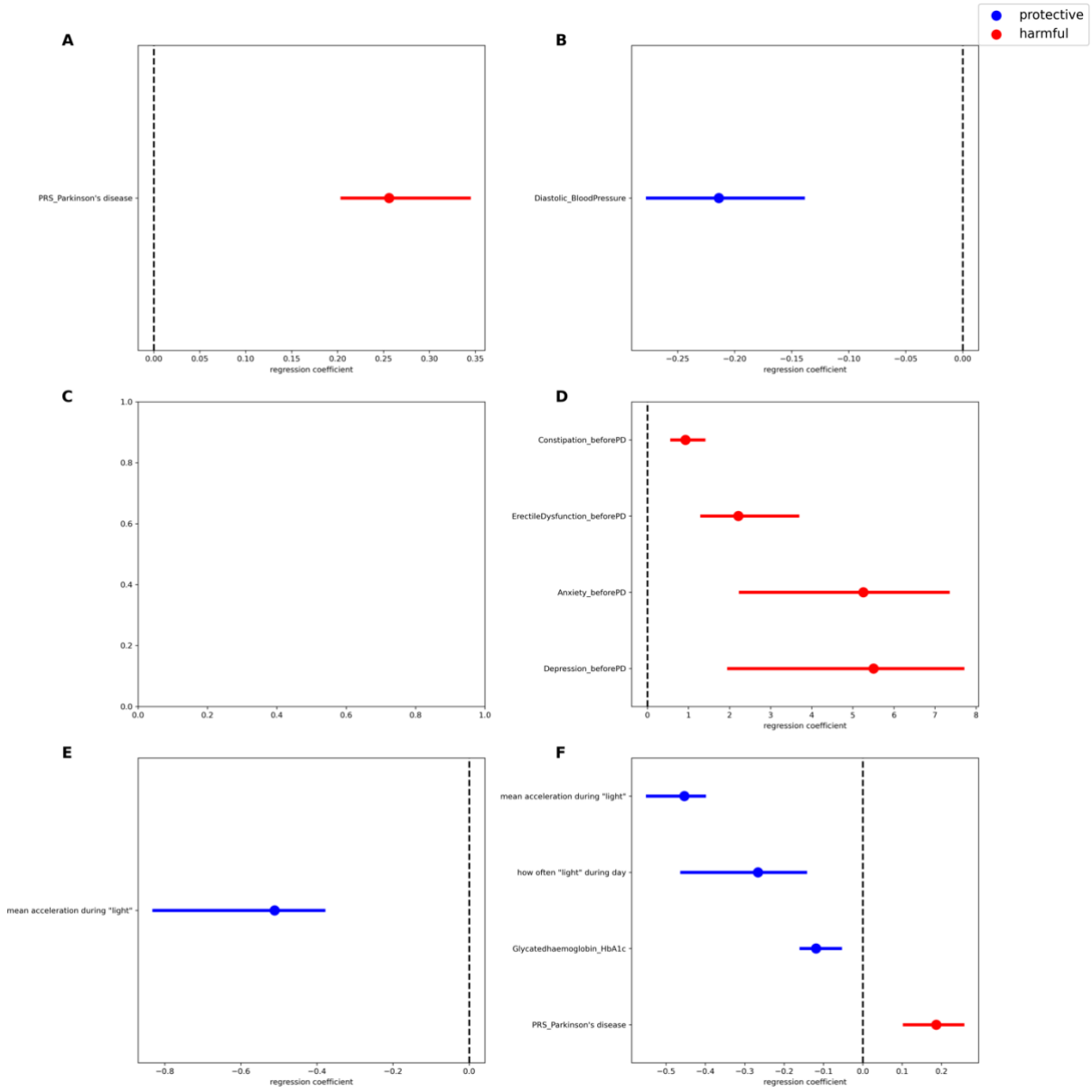
Supplemental Figure 11: Precision recall curves for unmatched healthy control models

The mean precision recall curves across the outer five-folds are shown with their standard deviation. We show this for each model type: [A] identifying diseased cases, [B] identifying prodromal cases, [C] identifying diseased and prodromal cases from unmatched healthy controls for each tested modality.



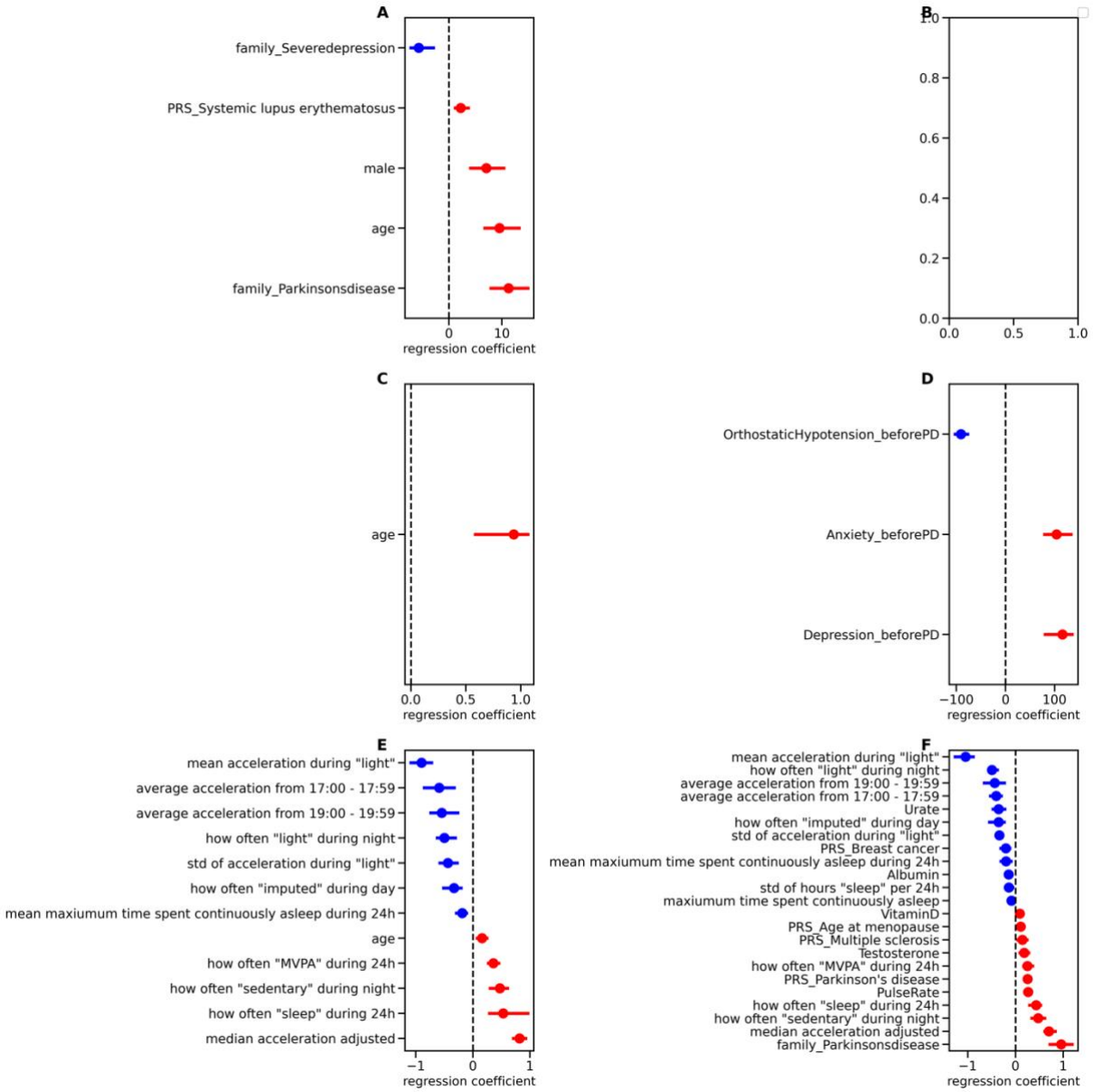
Supplemental Figure 12: Precision recall curves for population-based models

The mean precision recall curves across the outer five-folds are shown with their standard deviation. We show this for each model type: [A] identifying diseased cases, [B] identifying prodromal cases, [C] identifying diseased and prodromal cases from the population for each tested modality.



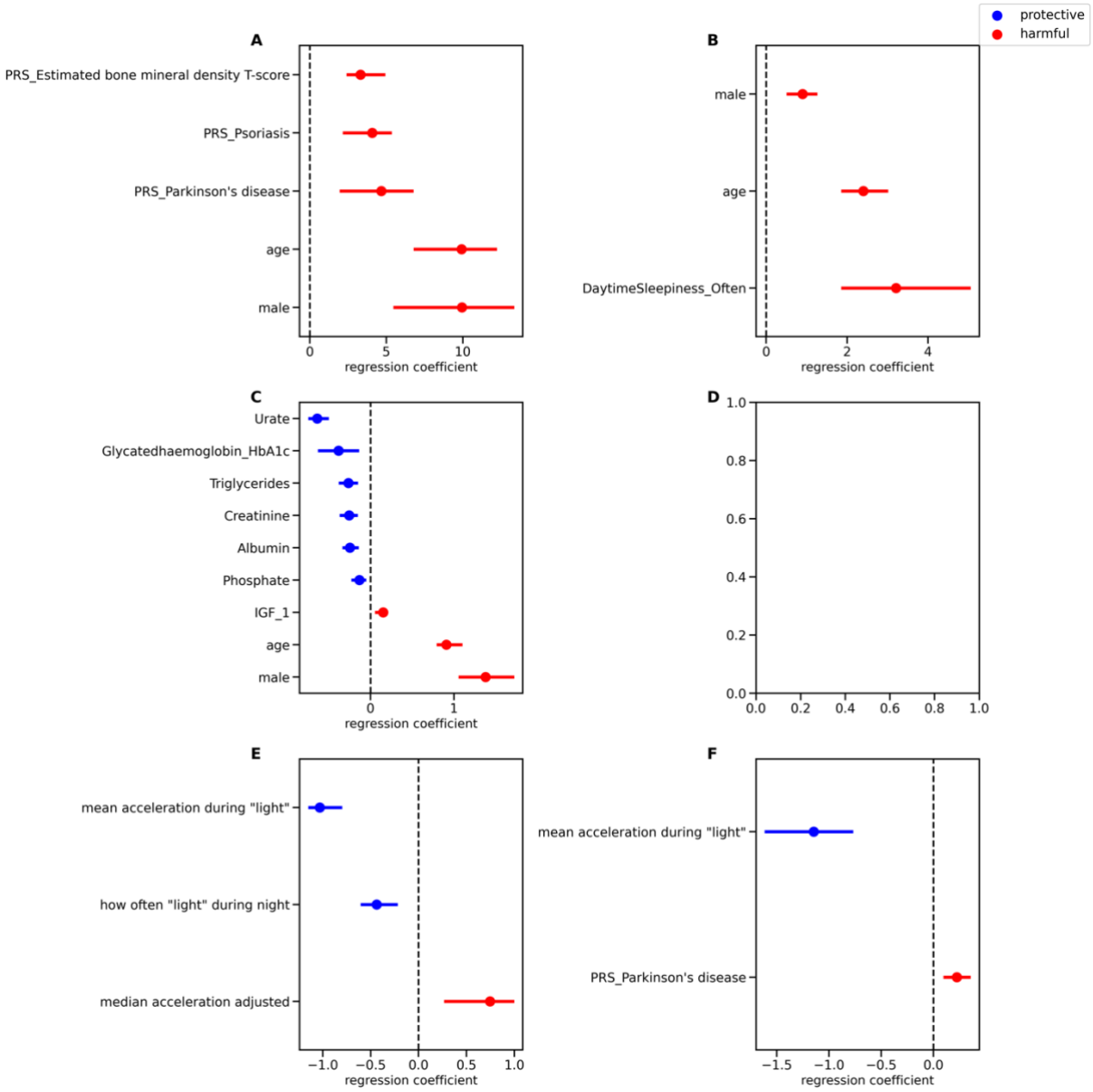
Supplemental Figure 13: Feature importance for each modality for the diagnosed vs matched unaffected controls models

The most important, significant features across the five outer cross-validation splits are shown for the diagnosed Parkinson's disease (PD) versus matched unaffected controls models. For each feature we show the mean regression coefficient across the outer five cross-validation folds and the 95% CI corrected for multiple comparisons with the Bonferroni-correction. Features that increase the likelihood of having PD are shown in red, whereas those that decrease the likelihood are shown in blue. Each plot shows the features of a model trained on a different feature set/modality: [A] genetics+family, [B] lifestyle, [C] blood, [D] prodromal symptoms, [E] all accelerometry, [F] combined.



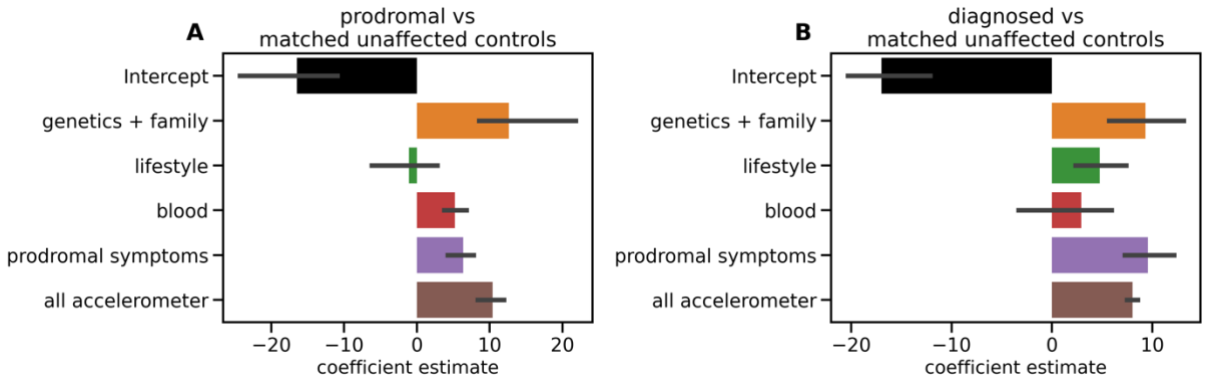
Supplemental Figure 14: Feature importance for each modality for the diagnosed vs all unaffected controls models

The most important, significant features across the five outer cross-validation splits are shown for the diagnosed Parkinson's disease (PD) versus matched unaffected controls models. For each feature we show the mean regression coefficient across the outer five cross-validation folds and the 95% CI corrected for multiple comparisons with the Bonferroni-correction. Features that increase the likelihood of having PD are shown in red, whereas those that decrease the likelihood are shown in blue. Each subplot shows the features of one model trained on a different feature set/modality: [A] genetics+family, [B] lifestyle, [C] blood, [D] prodromal symptoms, [E] all accelerometry, [F] combined.



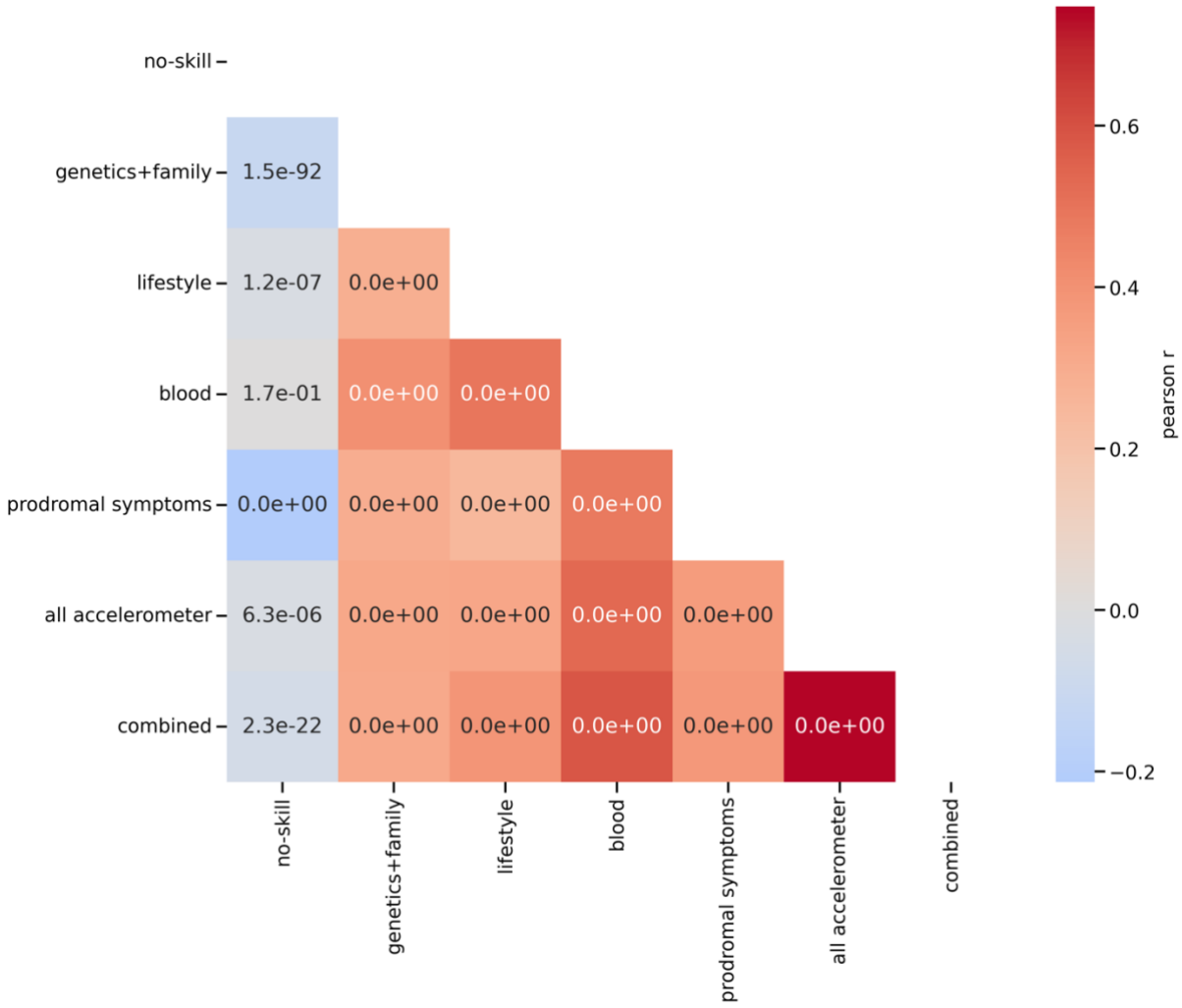
Supplemental Figure 15: Feature importance for each modality for the diagnosed vs population models

The most important, significant features across the five outer cross-validation splits are shown for the diagnosed Parkinson's disease (PD) versus matched unaffected controls models. For each feature we show the mean regression coefficient across the outer five cross-validation folds and the 95% CI corrected for multiple comparisons with the Bonferroni-correction. Features that increase the likelihood of having PD are shown in red, whereas those that decrease the likelihood are shown in blue. Each plot shows the features of a model trained on a different feature set/modality: [A] genetics+family, [B] lifestyle, [C] blood, [D] prodromal symptoms, [E] all accelerometry, [F] combined.



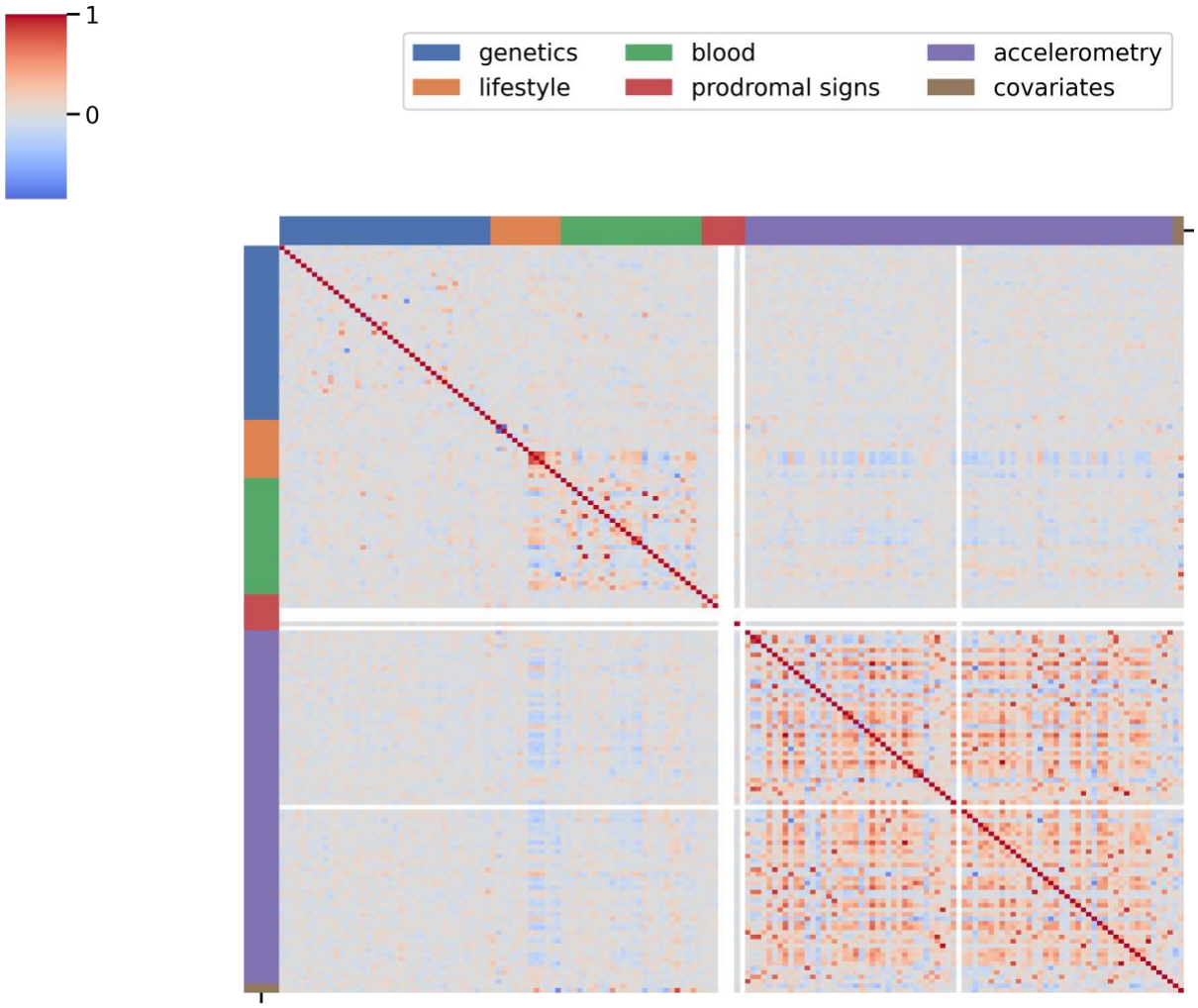
Supplemental Figure 16: Modality importance for the stacked model

The coefficients assigned to the modality specific predictions are shown. We plot the mean coefficient across the five outer folds and the Bonferroni-adjusted 95% Confidence Interval across those five folds (N=5). [A] shows the coefficients of the prodromal Parkinson's disease (PD) versus matched unaffected controls model and [B] the ones of the diagnosed PD versus matched unaffected controls model.



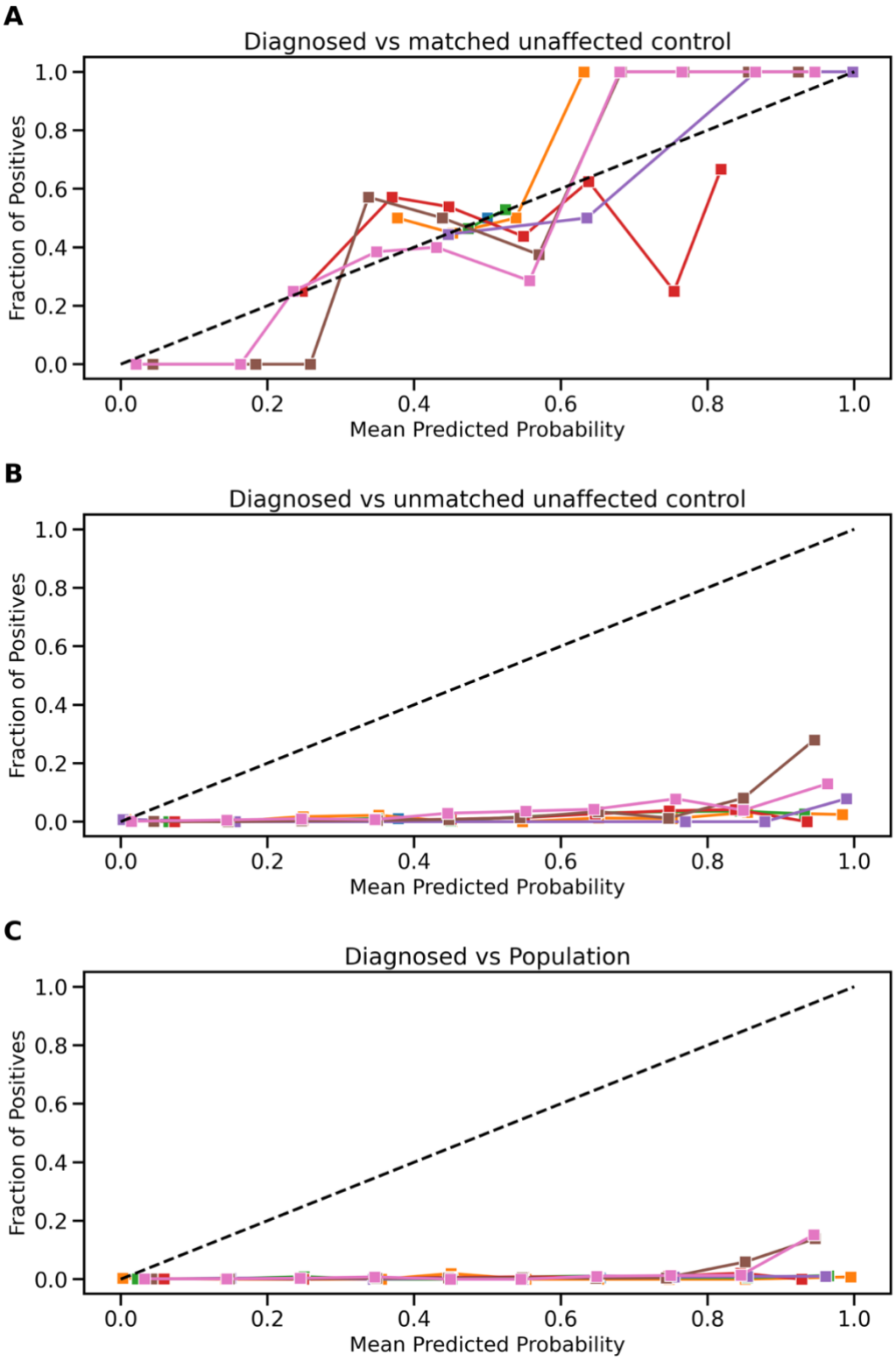
Supplemental Figure 17: Correlation of model predictions

The pearson correlation between the mean predicted probabilities of the different models to identify prodromal cases from the population on the test sets is shown. The color indicates the pearson r coefficient with darker red meaning higher correlation and the number indicates the associated p-value (two-sided). 0.05 Bonferroni corrected significance is reached at 7.14 $x10^{-3}$



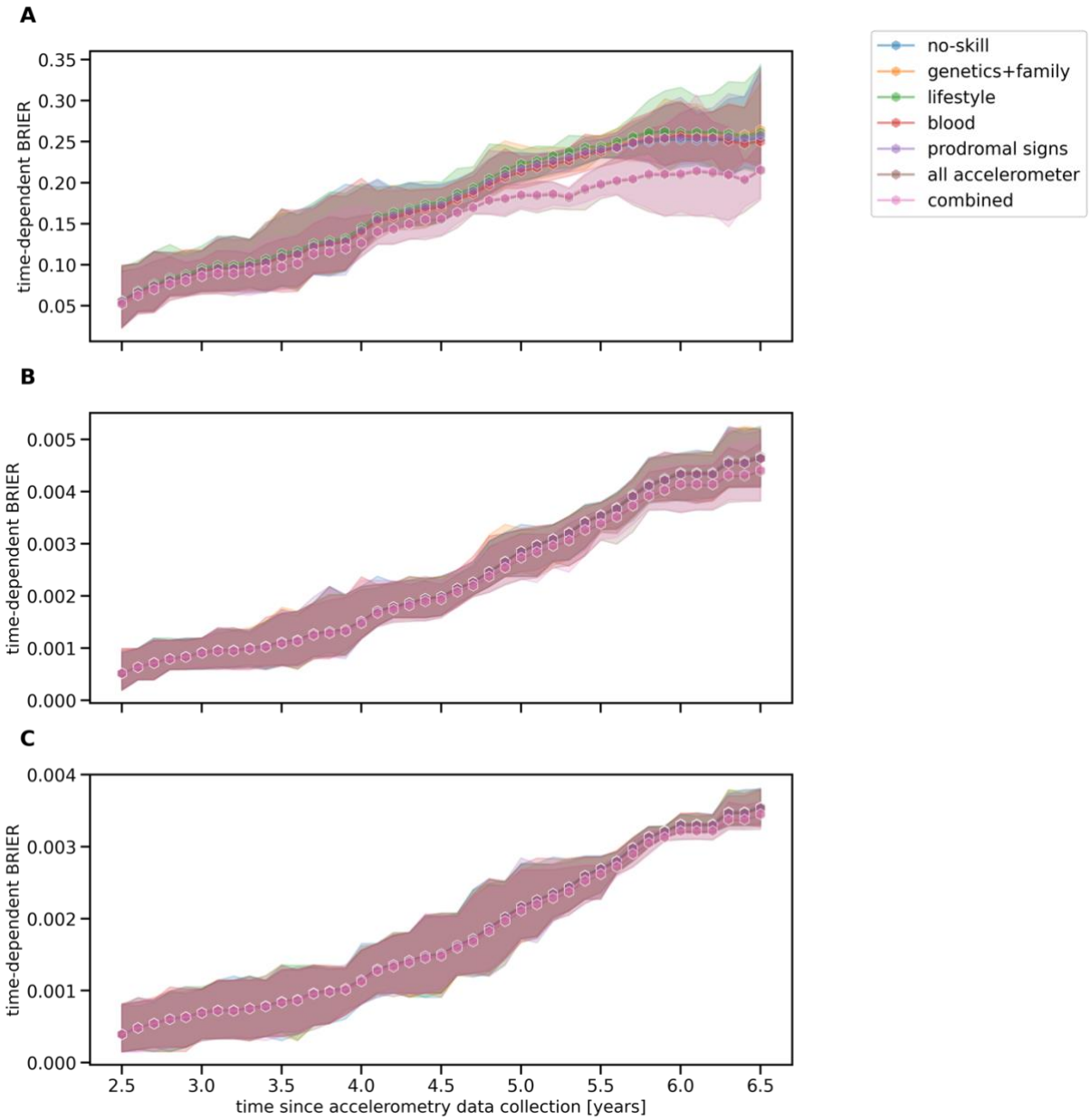
Supplemental Figure 18: Correlation of predictors

The pearson correlations between the features of the different modalities in the unaffected controls are shown. The color indicates the pearson r coefficient with darker red meaning higher correlation.



Supplemental Figure 19: Calibration of the diagnosed models

The calibration plots are shown for each single modality model and the combined one. We show this for all three control groups: [A] matched unaffected controls, [B] unmatched unaffected controls, and [C] general population. The curves show the mean predicted probability binned into ten bins against the true fraction of positives/prevalence. The dashed black line indicates perfect calibration.



Supplemental Figure 20: Brier score for survival models

A performance evaluation of the survival models is provided in a time-dependent manner. The mean brier score of the random survival forests is plotted for several evaluation time-points (years since data collection) together with Bonferroni-adjusted 95% confidence across the five outer cross validation folds interval for seven years since accelerometry data collection. We show this for [A] a control group made up of matched unaffected controls, [B] a control group including all unaffected controls, and [C] a control group representing the general population.

References:

Wu, Y., Byrne, E. M., Zheng, Z., Kemper, K. E., Yengo, L., Mallett, A. J., Yang, J., Visscher, P. M., & Wray, N. R. (2019). Genome-wide association study of medication-use and associated disease in the UK Biobank. *Nat Commun*, *10*(1), 1891. https://doi.org/10.1038/s41467-019-09572-5

TRIPOD Checklist: Prediction Model Development

Section/Topic	Item	Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	1
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	2
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	/
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	17-19
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	17-19
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	17
	5b	Describe eligibility criteria for participants.	17
	5c	Give details of treatments received, if relevant.	Sup. 1
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	17
	6b	Report any actions to blind assessment of the outcome to be predicted.	/
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	18-20
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	/
Sample size	8	Explain how the study size was arrived at.	20
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	20
Statistical analysis methods	10a	Describe how predictors were handled in the analyses.	21
	10b	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	20-21
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	21
Risk groups	11	Provide details on how risk groups were created, if done.	/
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	Sup Fig 20
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	Sup Table 8&9
Model development	14a	Specify the number of participants and outcome events in each analysis.	22
	14b	If done, report the unadjusted association between each candidate predictor and outcome.	/
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Sup Figure 16-22
	15b	Explain how to the use the prediction model.	17
Model performance	16	Report performance measures (with CIs) for the prediction model.	8-11 (sup table 5-7), calibration Sup Fig 27&28, Sup Table 18), page 24
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	15
Interpretation	19b	Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence.	14-16
Implications	20	Discuss the potential clinical use of the model and implications for future research.	16-17
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	24
Funding	22	Give the source of funding and the role of the funders for the present study.	24

We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.