

Prediction of labor market marginalization in psychiatric patients following internet-delivered cognitive behavioral therapy

Olly Kravchenko¹, Matthew Halvorsen^{1,2}, Julia Bäckman¹, Viktor Kaldo^{1,3}, James J Crowley^{1,2}, Ralf Kuja-Halkola⁴, Christian Rück¹, John Wallert^{1,4}

1. Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, & Stockholm Health Care Services, Region Stockholm, Stockholm, Sweden
2. Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, USA
3. Department of Psychology, Faculty of Health and Life Sciences, Linnaeus University, Växjö, Sweden
4. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Corresponding author: olly.kravchenko@ki.se

TABLE OF CONTENTS

S1. LISA database registers.....	1
S2. Polygenic risk scores.....	2
Table S1. Complete list of predictors used in full models.....	3
S3. Modeling approaches and hyperparameter tuning.....	12
Table S2. Polygenic score associations with labor market marginalization.....	13
Table S3. Nested cross-validation performance in full cohort (n = 2630).....	15
Table S4. Nested cross-validation performance in PGS sub-cohort (n = 2182)	16
Table S5. Holdout test set performance in full cohort (n = 2630) (all metrics).....	17
Table S6. Holdout test set performance in PGS sub-cohort (n = 2182) (all metrics).....	18
Fig. S1. Confusion matrices of the holdout test set performance.....	19
Fig. S2. ROC curves of models with and without PGS in the holdout test set (n = 435).....	20
Fig. S3. Calibration curves of the recalibrated models evaluated on the holdout test set.....	21
Fig. S4. Global predictor importances per model.....	22
Table S7. Normalized global SHAP importances per model.....	24

S1. LISA database registers

The Longitudinal Integrated Database for Health Insurance and Labour Market Studies (LISA) [1] is managed by Statistics Sweden (SCB) and combines data from the following national registers:

1. The Register-based Employment Statistics
2. The Total Population Register
3. The Education Register
4. The Register of Income and Taxation
5. The Swedish Public Employment Service
6. The Swedish Social Insurance Agency

S2. Polygenic risk scores

Work- and disability-related traits selected from the UK Biobank GWAS Atlas [2]*

Trait	UK Biobank field ID
1 Long-standing illness, disability or infirmity	f.2188.0.0_logistic
2 Current employment status: In paid employment or self-employed	6142_1_logistic
3 Other serious medical condition/disability (diagnosed by doctor)	f.2473.0.0_logistic
4 Household income	f.738.0.0_res
5 Time employed in main current job	f.757.0.0_res
6 Disability living allowance	6146_2_logistic
7 Family history of certain disabilities and chronic diseases (leading to disablement)	41204_Z82_logistic
8 Current employment status: Unable to work because of sickness or disability	6142_4_logistic

**Despite phenotypic overlap, the eight PGS were only weakly correlated and thus all were retained for the analysis*

PGS for psychiatric disorders and traits with references to respective GWAS

1. Depression [3]
2. Anxiety disorders [4]
3. Panic disorder [5]
4. Obsessive-compulsive disorder [6]
5. Tourette's syndrome and other tic disorders [7]
6. Post-traumatic stress disorder [8]
7. Attention deficit hyperactivity disorder [9]
8. Autism spectrum disorder [10]
9. Bipolar disorder [11]
10. Anorexia nervosa [12]
11. Schizophrenia [13]
12. Multiple substance use disorders [14]
13. Cross-disorder PGS [15]
14. Intelligence [16]
15. Educational attainment [17]

Table S1. Complete list of predictors used in full models

Predictor type	Data source	Predictor	Definition	Missingness in full cohort (n = 2630)	Comment
Socio-demographic	Swedish personal number	Age	Number of years	0	
Socio-demographic	Swedish personal number	Sex	Biological sex (Male = 0 / Female = 1)	0	
Socio-demographic	IPSY & LISA	Marital status	1. Married - Officially married - De facto married (in a committed relationship): 'married/living together/living apart together' 2. Unmarried - Single - Separated - Widowed	0	By default, used self-reported at IPSY due to better construct validity. For 6 patients with missing values, used register-based, derived from LISA.
Socio-demographic	IPSY & LISA	Parental status	Has children = 1 / No children = 0	6	
Socio-demographic	LISA	Single parent	Yes / No	0	Derived from Family status in LISA (codes 210 (single father with child <18) and 230 (single mother with child <18)). For 6 patients with missing values, derived from relevant LISA variables pertaining to number of children and family type.
Socio-demographic	IPSY & LISA	Educational attainment	University (Unfinished university and Finished university) / No university (<7 years, 7-9 years, Unfinished gymnasium, Vocational school, Gymnasium)	0	By default, used self-reported at IPSY. For 6 patients with missing values, used register-based, derived from LISA.

Socio-demographic	LISA & IPSY	Employment status	1. Employed - Employed / gainfully employed - Not employed / not gainfully employed according to the definition, but with an employer income statement or self-employment/business income during the year 2. Unemployed Not employed / not gainfully employed, without an employer income statement and without self-employment/business income during the year	0	For 6 patients with missing values, used self-reported
Socio-demographic	LISA	Immigrant 1st gen	Immigrant status. Country of birth other than Sweden: Yes / No	0	
Socio-demographic	LISA	Immigrant 2nd gen	Parental immigrant status. Mother's or father's country of birth other than Sweden: Yes / No	196	
Socio-demographic	LISA	Income quintile	Annual disposable household income adjusted for family composition. Divided into quintiles.	5	
Socio-demographic	LISA	Income trajectory	Percentage change in annual income from the year before treatment (t-1) compared to the previous year (t-2): - Growth (Increase of 20% or more) - Stable (Change within $\pm 20\%$) - Reduction (Decrease of 20% or more)	22	
Socio-demographic	LISA	Receipt of financial benefits 1 year before	Calendar year before treatment start year Any of the following: - Rehabilitation compensation - Unemployment benefit - Compensation in connection with a labor market policy program - Sickness/work injury benefit - Social assistance	5	- Rehabilitation compensation (rehabiliteringsersättning) is paid out during the sickness if an individual undergoes vocational training - Unemployment benefit (arbetslöshetsersättning) applies to individuals who are currently unemployed but have worked previously

			- Child care allowance - Sickness / activity compensation Yes / No		- Compensation in connection with a labor market policy program by the Swedish Public Employment Service (åtgärdssysselsättning) - Sickness/work injury benefit (sjukpenning/arbetsskadesjukpenning) is paid out by the Swedish Social Insurance Agency in case of prolonged sickness (over 2 weeks) - Social assistance (ekonomiskt bidrag) is provided to households that experience economic hardships and require support to pay fixed expenses - Child care allowance (vårdbidrag) is financial assistance to parents with children who have special needs - Sickness compensation (sjukersättning)/activity compensation (aktivitetersättning), previously referred to as early retirement, is compensation for individuals who are prevented from working full-time ever in the future due to serious sickness, most commonly a psychiatric disorder
Socio-demographic	LISA	Receipt of student income 1 year before	Yes / No Calendar year before treatment start year	5	
Prior LMM	LISA	Sickness absence >90 days 1 year before	Prolonged sickness absence (LMM component at baseline) in the calendar year preceding the treatment start year	5	Sickness/work injury benefit (sjukpenning/arbetsskadesjukpenning) is paid out by the Swedish Social Insurance Agency in case of prolonged sickness (over 2 weeks)
Prior LMM	LISA	Prior 5-year sickness absence days	Total number of days receiving sickness benefits within the 5 calendar years preceding the treatment start year (sum of days for each of the 5 years)	120	Log-transformed to address skewness. Sickness/work injury benefit (sjukpenning/arbetsskadesjukpenning) is paid out by the Swedish Social Insurance Agency in case of prolonged sickness (over 2 weeks). If any of the 5 preceding years is NA, the 5-year sum is NA.

Prior LMM	LISA	Unemployment >180 days 1 year before	Prolonged unemployment (LMM component at baseline) in the calendar year preceding the treatment start year	5	
Prior LMM	LISA	Prior 5-year unemployment days	Total number of days in unemployment within the 5 calendar years preceding the treatment start year (sum of days for each of the 5 years)	120	Log-transformed to address skewness. If any of the 5 preceding years is NA, the 5-year sum is NA
Prior LMM	LISA	Prior 5-year disability pension	Receipt of disability pension in any of the 5 calendar years preceding the treatment start year: Yes / No (LMM component at baseline)	4	Disability pension in Sweden comprises sickness compensation (sjukersättning) and activity compensation (aktivitetsersättning) for individuals who are prevented from working full-time ever in the future due to serious sickness, most commonly a psychiatric disorder
Baseline clinical characteristics	IPSY	Pre-treatment symptom severity	MADRS-S score for patients with depression, PDSS-SR for panic disorder, and LSAS-SR for social anxiety disorder before treatment start. Respective scales harmonized to a 0-100 scale and pooled into a single variable.	3	
Baseline clinical characteristics	IPSY	EQ-5D sum score	- No difficulties = 0 - Some difficulties = 1 - Significant difficulties = 2	115	EQ-5D score covers 5 dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression), each with 3 levels coded 0-2: - No difficulties = 0 - Some difficulties = 1 - Significant difficulties = 2 EQ-5D sum score was computed by summing item levels across the five dimensions, yielding a theoretical range of 0-10. Observed range in the sample: 0-7
Baseline clinical characteristics	IPSY	CGI-S score	CGI-S, Clinical Global Impressions, Severity of Illness, clinician-rated. Numeric score.	264	1 = Normal—not at all ill, symptoms of disorder not present past seven days 2 = Borderline mentally ill—subtle or suspected pathology 3 = Mildly ill—clearly established symptoms with minimal, if any, distress or difficulty in

					social and occupational function 4 = Moderately ill—overt symptoms causing noticeable, but modest, functional impairment or distress; symptom level may warrant medication 5 = Markedly ill—intrusive symptoms that distinctly impair social/occupational function or cause intrusive levels of distress 6 = Severely ill—disruptive pathology, behavior and function are frequently influenced by symptoms, may require assistance from others 7 = Among the most extremely ill patients—pathology drastically interferes in many life functions; may be hospitalized Observed range in the sample: 2-6
Baseline clinical characteristics	IPSY	Insomnia	Insomnia Severity Index score, dichotomized: - Insomnia >14 - No insomnia <=14	83	
Baseline clinical characteristics	IPSY	AUDIT score	Alcohol Use Disorders Identification Test. Numeric score.	96	
Baseline clinical characteristics	IPSY	DUDIT score	Drug Use Disorders Identification Test. Numeric score.	7	
Baseline clinical characteristics	IPSY & LISA	Psychiatric comorbidities	Comorbidity (presence of any of the 17 listed psychiatric disorders) / No comorbidity	0	By default, used self-reported at IPSY, but for 98 patients with missing values, used register-based comorbidity (presence of a psychiatric diagnosis in NPR or VAL, other than primary diagnosis, before treatment)
Baseline clinical characteristics	IPSY	Family history of psychopathology	Family history (presence of any of the 37 listed psychiatric disorders among close relatives) / No family history	100	
Baseline clinical characteristics	IPSY	Suicide risk	Clinician-rated suicide risk: Low (0) / Medium (1)	101	For 19 patients with missing values, used the suicide risk value from the clinician-rated MADRS

Baseline clinical characteristics	IPSY	Suicide attempts	Self-reported previous suicide attempts: Yes / No	226	
Baseline clinical characteristics	NPR	Inpatient psychiatric care	Ever admitted to inpatient psychiatric care before treatment: Yes / No	0	Constructed from NPR Slutenvård (inpatient care), by filtering MVO (specialization) starting with 9 (Psychiatry)
Baseline clinical characteristics	NPR & VAL	Prior F1 diagnosis	Mental and behavioral disorders due to psychoactive substance use: Yes / No	0	Received any time before treatment. Duplicates between NPR and VAL are excluded.
Baseline clinical characteristics	NPR & VAL	Prior F3 diagnosis	Mood [affective] disorders ever before: Yes / No	0	Bipolar disorder is excluded and treated as a separate variable to avoid conflation with depressive disorders targeted by the studied treatment. The depression diagnosis that a patient receives to start the ICBT treatment is transferred to the NPR. Hence, to not capture the diagnosis related to this very healthcare episode, all records from the month preceding treatment are excluded. Duplicates between NPR and VAL are excluded.
Baseline clinical characteristics	NPR & VAL	Prior F4 diagnosis	Anxiety, dissociative, stress-related, somatoform and other nonpsychotic mental disorders ever before: Yes / No	0	The panic disorder / social anxiety diagnosis that a patient receives to start the ICBT treatment is transferred to the NPR. Hence, to not capture the diagnosis related to this very healthcare episode, all records from the month preceding treatment are excluded. Duplicates between NPR and VAL are excluded.
Baseline clinical characteristics	NPR & VAL	Prior F5 diagnosis	Behavioral syndromes associated with physiological disturbances and physical factors ever before: Yes / No	0	Received any time before treatment. Duplicates between NPR and VAL are excluded.
Baseline clinical characteristics	NPR & VAL	Prior F6 diagnosis	Disorders of adult personality and behavior ever before: Yes / No	0	Received any time before treatment. Duplicates between NPR and VAL are excluded.

Baseline clinical characteristics	NPR & VAL	ADHD diagnosis	Attention deficit hyperactivity disorder: Yes / No	0	Due to neurodevelopmental nature and persistent lifelong manifestation, not limited to diagnoses received before ICBT treatment. Duplicates between NPR and VAL are excluded.
Baseline clinical characteristics	NPR & VAL	ASD diagnosis	Autism spectrum disorder: Yes / No	0	Due to neurodevelopmental nature and persistent lifelong manifestation, not limited to diagnoses received before ICBT treatment. Duplicates between NPR and VAL are excluded.
Baseline clinical characteristics	NPR & VAL	Prior psychiatric visits	Number of records of primary or outpatient care visits related to a psychiatric diagnosis	0	For F3 and F4 diagnoses, all records from the month preceding treatment are excluded. Duplicates between NPR and VAL are excluded.
Baseline clinical characteristics	NPR & VAL	Years since the first psychiatric diagnosis	Number of years	0	Days since any first psychiatric diagnosis, recalculated in years. Serves as a proxy for age of onset of psychiatric morbidity.
Baseline clinical characteristics	PDR	Prior dispensations of psychotropic medication	Number of dispensations	0	
Baseline clinical characteristics	PDR	Prior use of antidepressants	Yes / No	0	For F3 and F4 diagnoses, all records from the month preceding treatment are excluded.
Baseline clinical characteristics	PDR	Prior use of anxiolytics	Yes / No	0	For F3 and F4 diagnoses, all records from the month preceding treatment are excluded.
Baseline clinical characteristics	PDR	Prior use of hypnotics and sedatives	Yes / No	0	For F3 and F4 diagnoses, all records from the month preceding treatment are excluded.
Baseline clinical characteristics	PDR	Prior use of antipsychotics	Yes / No	0	For F3 and F4 diagnoses, all records from the month preceding treatment are excluded.
Baseline clinical characteristics	IPSY & PDR	Current use of psychotropic medication	Self-reported taking any psychotropic medication at the moment of treatment start: Yes / No	0	For 440 patients with missing values, used records of psychotropic drugs dispensed within 30 days before treatment start in PDR

Baseline clinical characteristics	VAL	Musculoskeletal diagnosis	Diagnosis of the musculoskeletal system and connective tissue one year before treatment start: Yes/ No	0	Data on all healthcare appointments from primary care, extracted from the central Stockholm County Council data warehouse (VAL). Diagnostic codes from ICD-10-CM chapter 13: M00-M99. Reason for inclusion: Diseases of the musculoskeletal system are the second most common cause of sickness absence [18] and disability pension [19] in Sweden, have a high rate of comorbidity and shared genetic liability with mental disorders [20, 21], and have shown synergistic effect with mental disorders on sickness absence [22] and disability [23].
ICBT treatment course	IPSY	ICBT target diagnosis: Depression	Patient is undergoing ICBT treatment for depression: Yes / No	0	
ICBT treatment course	IPSY	ICBT target diagnosis: Panic disorder	Patient is undergoing ICBT treatment for panic disorder: Yes / No	0	
ICBT treatment course	IPSY	ICBT target diagnosis: Social anxiety	Patient is undergoing ICBT treatment for social anxiety: Yes / No	0	
ICBT treatment course	IPSY	Login duration in hours	Total time logged in to the treatment platform	1	Converted from seconds to hours. Log-transformed to address outliers, e.g., a patient who left the browser tab open overnight (stale session)
ICBT treatment course	IPSY	Completed modules	Total number of completed ICBT modules	0	Used as a standard metric of treatment adherence
ICBT treatment course	IPSY	Message length in characters	Total length of messages sent by patient in characters	0	Text-based therapy modules and communication with the therapist. Log-transformed to address skewness.
ICBT treatment course	IPSY	Non-beneficial attrition	Non-beneficial attrition was defined as early treatment discontinuation without	0	To reliably identify those patients who most likely left the intervention too early to

			documented treatment response (see Treatment response section). Coded as 0: - Completers (patients who completed all 10 modules) - Beneficial attrition (patients who quit early (<10 modules) BUT achieved treatment responder status (defined in Treatment response section)) Coded as 1: - Non-beneficial attrition (patients who completed <10 modules AND did not achieve treatment responder status) - Unknown (patients who completed <10 modules AND have missing treatment responder status)		experience benefits, we followed the methodology in Zantvoort et al. (2024) [24] to construct the variable that takes not only attrition but also treatment response into account.
ICBT treatment course	IPSY	Treatment response	Clinically meaningful improvement following ICBT treatment, coded as ANY of the following criteria was met: 1. Reduction of symptoms from pre- to post-treatment - Depression: $\geq 40\%$ (MADRS-S) - Panic disorder: $\geq 40\%$ (PDSS-SR) - Social anxiety: $\geq 30\%$ (LSAS-SR) 2. Remission - Depression: ≤ 10 (MADRS-S) - Panic disorder: ≤ 7 (PDSS-SR) - Social anxiety: ≤ 35 (LSAS-SR)	517	

ICBT, Internet-delivered cognitive behavior therapy; IPSY, Internet psychiatry clinic; MADRS-S, Montgomery–Åsberg Depression Rating Scale, self-rating version; PDSS-SR, Panic Disorder Severity Scale, self-report; LSAS-SR, Liebowitz Social Anxiety Scale, self-report; NPR, National patient register; LISA, The longitudinal integrated database for health insurance and labor market studies; PDR, Prescribed drug register; ADHD, Attention deficit hyperactivity disorder; ASD, Autism spectrum disorder; CBT, Cognitive behavioral therapy; LMM, Labor market marginalization

S3. Modeling approaches and hyperparameter tuning

1. *Elastic net penalized logistic regression (EN)* is an easily interpretable regularized classification method that combines L1 (LASSO) and L2 (Ridge) penalties to improve model performance by handling multicollinearity and performing feature selection. It shrinks irrelevant feature coefficients to zero, creating sparse models and improving performance in high-dimensional settings with correlated features.

We used grid search to tune L1 ratios in the range {0.1, 0.5, 1.0} and regularization strength (C values) in the range {0.1, 0.5, 1, 5, 10}.

2. *Random Forest (RF)* is a non-linear tree-based classification method that allows interaction effects.

We used grid search to tune number of trees (n_estimators) {100, 300, 500}, maximum tree depth (max_depth) {None, 5, 10, 20}, minimum samples to split (min_samples_split) {2, 10, 50}, minimum samples per leaf (min_samples_leaf) {1, 2, 4, 10}, and features per split (max_features) {"sqrt", "log2"}.

3. *XGBoost* is a gradient boosting classification method that can be beneficial if effects are somewhat additive but non-linear and in the presence of complex interactions.

For hyperparameter tuning, we compared randomized search, grid search, and Bayesian optimization across three search spaces. Because all three strategies yielded essentially identical discrimination performance, we selected randomized search on the basis of slightly lower log loss. We tuned the number of gradient-boosted trees (n_estimators) {200, 400, 600, 800}, learning rate (learning_rate) as 10 evenly spaced values between 0.01 and 0.15, maximum tree depth (max_depth) {2, 3, 4, 5}, minimum child weight (min_child_weight) {1, 3, 5}, subsample ratio of training instances (subsample) {0.6, 0.8, 1.0}, subsample ratio of columns when constructing each tree (colsample_bytree) {0.6, 0.8, 1.0}, L1 regularization term on weights (reg_alpha) {0, 0.5, 1.0}, and L2 regularization term on weights (reg_lambda) {1, 5, 10}.

Table S2. Polygenic score associations with labor market marginalization one year after ICBT

Trait	UK Biobank field ID	OR [95% CI]	p-value
Educational attainment		0.83 [0.71, 0.96]	0.013*
Depression		1.18 [1.02, 1.38]	0.029*
Long-standing illness, disability or infirmity	f.2188.0.0_logistic	1.15 [0.98, 1.33]	0.080
Multiple substance use disorders		1.13 [0.97, 1.32]	0.109
Current employment status: In paid employment or self-employed	6142_1_logistic	0.88 [0.76, 1.03]	0.117
Obsessive-compulsive disorder		1.10 [0.94, 1.28]	0.224
Other serious medical condition/disability (diagnosed by doctor)	f.2473.0.0_logistic	1.10 [0.94, 1.27]	0.242
Anxiety disorders		1.08 [0.93, 1.26]	0.295
Panic disorder		1.07 [0.93, 1.24]	0.335
Household income	f.738.0.0_res	0.95 [0.82, 1.12]	0.554
Tourette's syndrome and other tic disorders		1.04 [0.90, 1.22]	0.576
Autism spectrum disorder		0.96 [0.82, 1.12]	0.588
Cross-disorder PGS		1.04 [0.90, 1.21]	0.592
Time employed in main current job	f.757.0.0_res	0.96 [0.83, 1.12]	0.611
Attention deficit hyperactivity disorder		1.04 [0.89, 1.21]	0.612
Post-traumatic stress disorder		1.04 [0.89, 1.21]	0.620
Intelligence		0.97 [0.83, 1.13]	0.687
Disability living allowance	6146_2_logistic	1.03 [0.88, 1.20]	0.719
Anorexia nervosa		1.03 [0.88, 1.19]	0.740

Bipolar disorder		1.02 [0.87, 1.18]	0.841
Family history of certain disabilities and chronic diseases (leading to disablement)	41204_Z82_logistic	0.99 [0.85, 1.15]	0.888
Current employment status: Unable to work because of sickness or disability	6142_4_logistic	0.99 [0.85, 1.15]	0.918
Schizophrenia		1.00 [0.86, 1.17]	0.959

ICBT, Internet-delivered cognitive behavior therapy; PGS, Polygenic scores; OR, Odds ratio; CI, Confidence interval

**Significance level $p < .05$*

Table S3. Nested cross-validation performance in full cohort ($n = 2630$)

	Baseline EN	Full EN	Full RF	Full XGBoost
AUC [95% CI]	0.756 [0.719, 0.794]	0.755 [0.717, 0.793]	0.751 [0.707, 0.794]	0.748 [0.706, 0.789]
BACC [95% CI]	0.684 [0.641, 0.726]	0.675 [0.633, 0.718]	0.645 [0.594, 0.696]	0.671 [0.632, 0.709]
F1 [95% CI]	0.302 [0.267, 0.336]	0.296 [0.261, 0.332]	0.314 [0.234, 0.394]	0.313 [0.274, 0.353]
Log Loss [95% CI]	0.566 [0.545, 0.588]	0.560 [0.539, 0.582]	0.422 [0.394, 0.450]	0.501 [0.482, 0.519]
MCC [95% CI]	0.235 [0.183, 0.287]	0.226 [0.174, 0.279]	0.260 [0.179, 0.341]	0.239 [0.187, 0.291]

EN, Elastic net; RF, Random forest; AUC, Area under the receiver operating characteristic curve; CI, Confidence interval; BACC, Balanced accuracy; MCC, Matthews correlation coefficient

Nested CV outer fold performance is reported with 95% CI to account for multiple-imputation uncertainty (Rubin's rules)

Table S4. Nested cross-validation performance in PGS sub-cohort ($n = 2182$)

	Elastic Net		Random Forest		XGBoost	
	With PGS	No PGS	With PGS	No PGS	With PGS	No PGS
AUC [95% CI]*	0.734 [0.708, 0.760]	0.744 [0.722, 0.767]	0.758 [0.736, 0.780]	0.753 [0.739, 0.768]	0.750 [0.728, 0.771]	0.762 [0.742, 0.781]
BACC [95% CI]	0.682 [0.657, 0.706]	0.677 [0.646, 0.708]	0.603 [0.568, 0.639]	0.667 [0.652, 0.682]	0.651 [0.621, 0.681]	0.682 [0.656, 0.707]
F1 [95% CI]	0.288 [0.266, 0.309]	0.284 [0.259, 0.308]	0.261 [0.180, 0.342]	0.348 [0.325, 0.372]	0.310 [0.274, 0.345]	0.316 [0.289, 0.343]
Log Loss [95% CI]	0.561 [0.536, 0.586]	0.569 [0.546, 0.592]	0.376 [0.351, 0.400]	0.444 [0.421, 0.468]	0.414 [0.385, 0.444]	0.489 [0.476, 0.502]
MCC [95% CI]	0.228 [0.197, 0.258]	0.222 [0.184, 0.259]	0.219 [0.143, 0.294]	0.284 [0.257, 0.310]	0.243 [0.203, 0.282]	0.251 [0.216, 0.286]

PGS, Polygenic scores; AUC, Area under the receiver operating characteristic curve; CI, Confidence interval; BACC, Balanced accuracy; MCC, Matthews correlation coefficient

*Nested CV outer fold performance is reported with 95% CI to account for multiple-imputation uncertainty (Rubin's rules)

Table S5. Holdout test set performance in full cohort ($n = 2630$) (all metrics)

	Baseline EN	Full EN	Full RF	Full XGBoost
AUC [95% CI]*	0.750 [0.665, 0.829]	0.754 [0.670, 0.832]	0.765 [0.673, 0.848]	0.768 [0.680, 0.850]
BACC [95% CI]	0.603 [0.533, 0.682]	0.631 [0.554, 0.714]	0.650 [0.569, 0.733]	0.650 [0.566, 0.736]
F1 [95% CI]	0.238 [0.121, 0.353]	0.255 [0.145, 0.365]	0.286 [0.165, 0.400]	0.286 [0.163, 0.404]
Log Loss [95% CI]	0.231 [0.191, 0.271]	0.229 [0.190, 0.269]	0.225 [0.185, 0.266]	0.223 [0.183, 0.265]
MCC [95% CI]	0.177 [0.056, 0.305]	0.195 [0.076, 0.317]	0.230 [0.102, 0.355]	0.230 [0.100, 0.355]
Threshold**	0.227	0.199	0.199	0.191
Sensitivity	0.286	0.371	0.400	0.400
Specificity	0.920	0.890	0.900	0.900
PPV	0.204	0.194	0.222	0.222
NPV	0.947	0.952	0.955	0.955

EN, Elastic net; RF, Random forest; AUC, Area under the receiver operating characteristic curve; CI, Confidence interval; BACC, Balanced accuracy; MCC, Matthews correlation coefficient; PPV, Positive predictive value; NPV, Negative predictive value

*Holdout test set performance is reported with 95% bootstrap CI to quantify sampling variability

**Thresholds chosen on the training set to maximize MCC using cross-fitted calibrated probabilities and applied unchanged to the holdout test set

Table S6. Holdout test set performance in PGS sub-cohort ($n = 2182$) (all metrics)

	Elastic Net		Random Forest		XGBoost	
	With PGS	No PGS	With PGS	No PGS	With PGS	No PGS
AUC [95% CI]	0.737 [0.643, 0.823]	0.762 [0.672, 0.846]	0.737 [0.640, 0.825]	0.773 [0.680, 0.857]	0.734 [0.638, 0.824]	0.772 [0.681, 0.854]
BACC [95% CI]	0.689 [0.609, 0.768]	0.680 [0.595, 0.773]	0.645 [0.554, 0.739]	0.685 [0.597, 0.774]	0.632 [0.539, 0.725]	0.655 [0.567, 0.746]
F1 [95% CI]	0.235 [0.157, 0.316]	0.291 [0.183, 0.403]	0.255 [0.145, 0.362]	0.302 [0.186, 0.409]	0.231 [0.130, 0.331]	0.295 [0.169, 0.416]
Log Loss [95% CI]	0.237 [0.188, 0.290]	0.233 [0.186, 0.284]	0.234 [0.185, 0.287]	0.229 [0.183, 0.279]	0.234 [0.184, 0.290]	0.227 [0.180, 0.279]
MCC [95% CI]	0.203 [0.112, 0.292]	0.240 [0.123, 0.364]	0.194 [0.069, 0.315]	0.252 [0.129, 0.372]	0.167 [0.047, 0.283]	0.237 [0.101, 0.372]
Threshold*	0.094	0.151	0.135	0.163	0.130	0.186
Sensitivity	0.710	0.516	0.452	0.516	0.452	0.419
Specificity	0.668	0.844	0.839	0.854	0.812	0.891
PPV	0.141	0.203	0.177	0.213	0.156	0.228
NPV	0.968	0.958	0.952	0.958	0.951	0.952

PGS, Polygenic scores; EN, Elastic net; RF, Random forest; AUC, Area under the receiver operating characteristic curve; CI, Confidence interval; BACC, Balanced accuracy; MCC, Matthews correlation coefficient; PPV, Positive predictive value; NPV, Negative predictive value

Holdout test set performance is reported with 95% bootstrap CI to quantify sampling variability

*Thresholds chosen on the training set to maximize MCC using cross-fitted calibrated probabilities and applied unchanged to the holdout test set

Fig. S1. Confusion matrices of the holdout test set performance

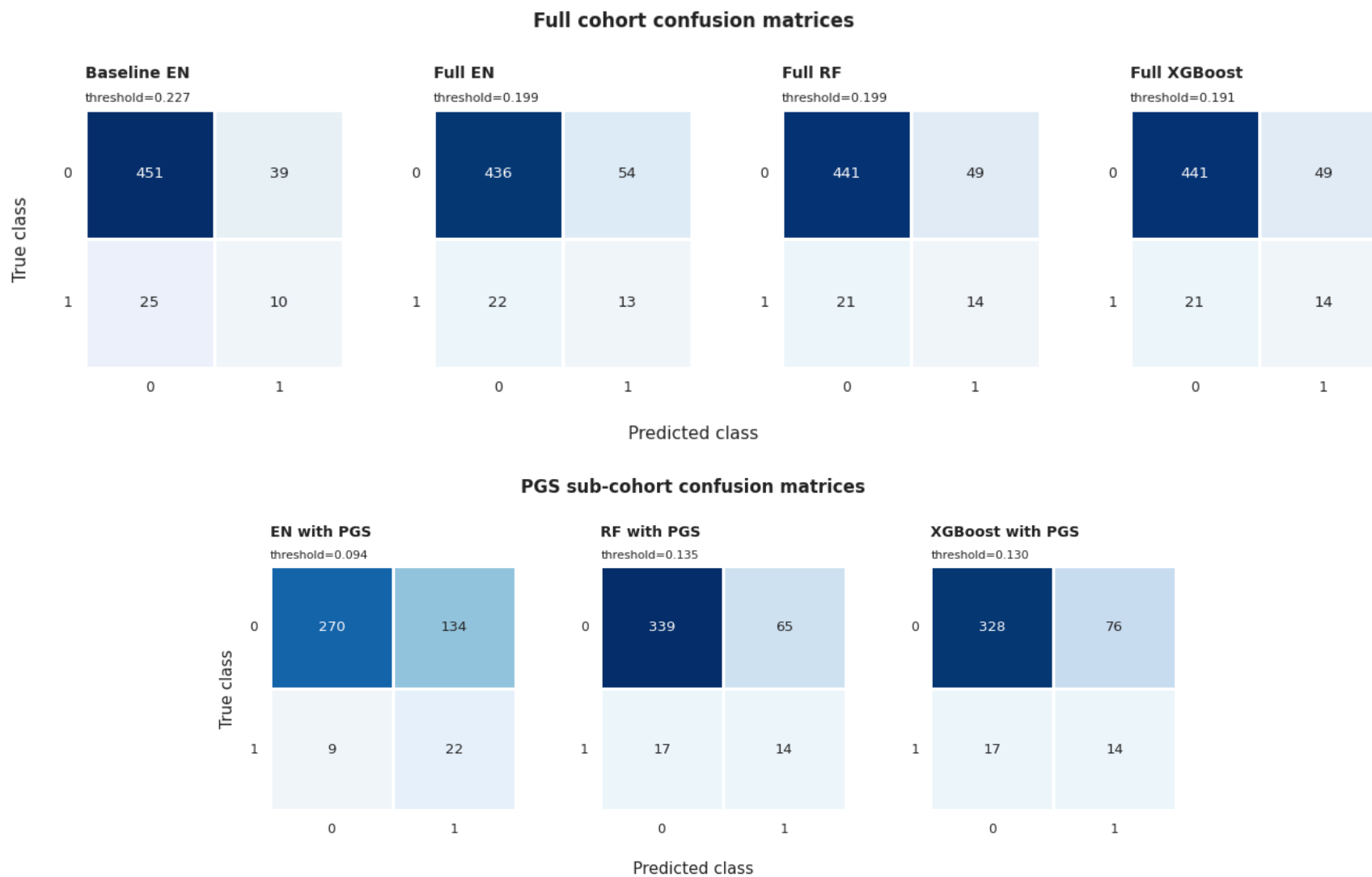


Fig. S2. ROC curves of models with and without PGS in the holdout test set ($n = 435$)

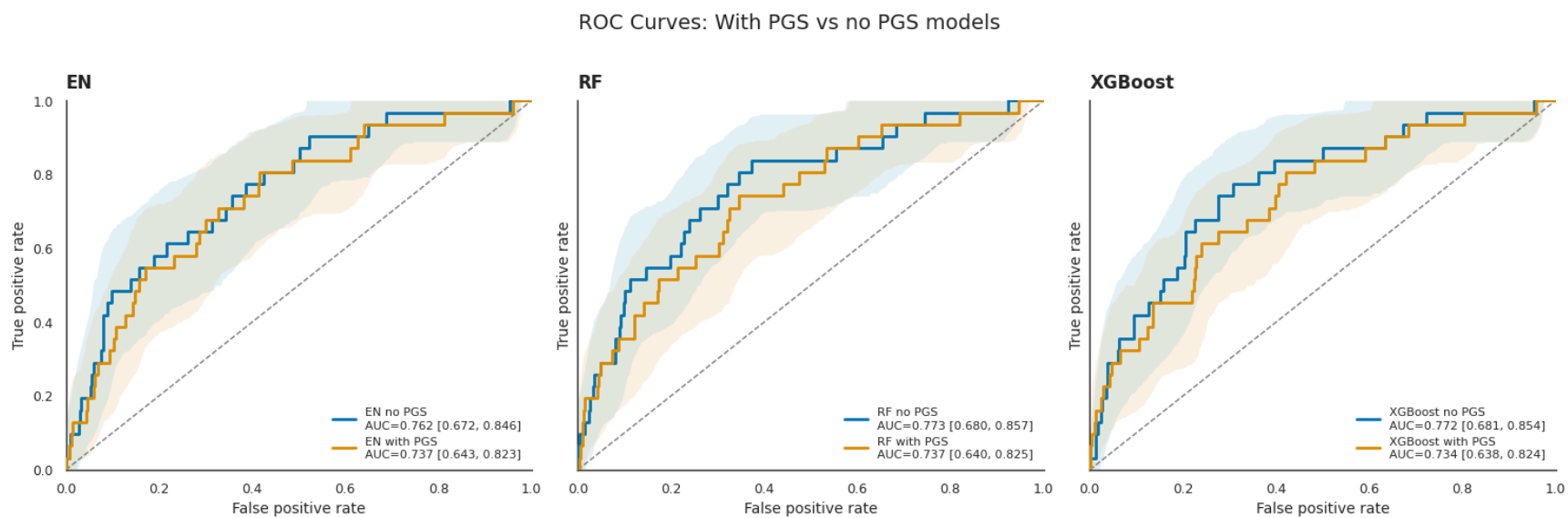
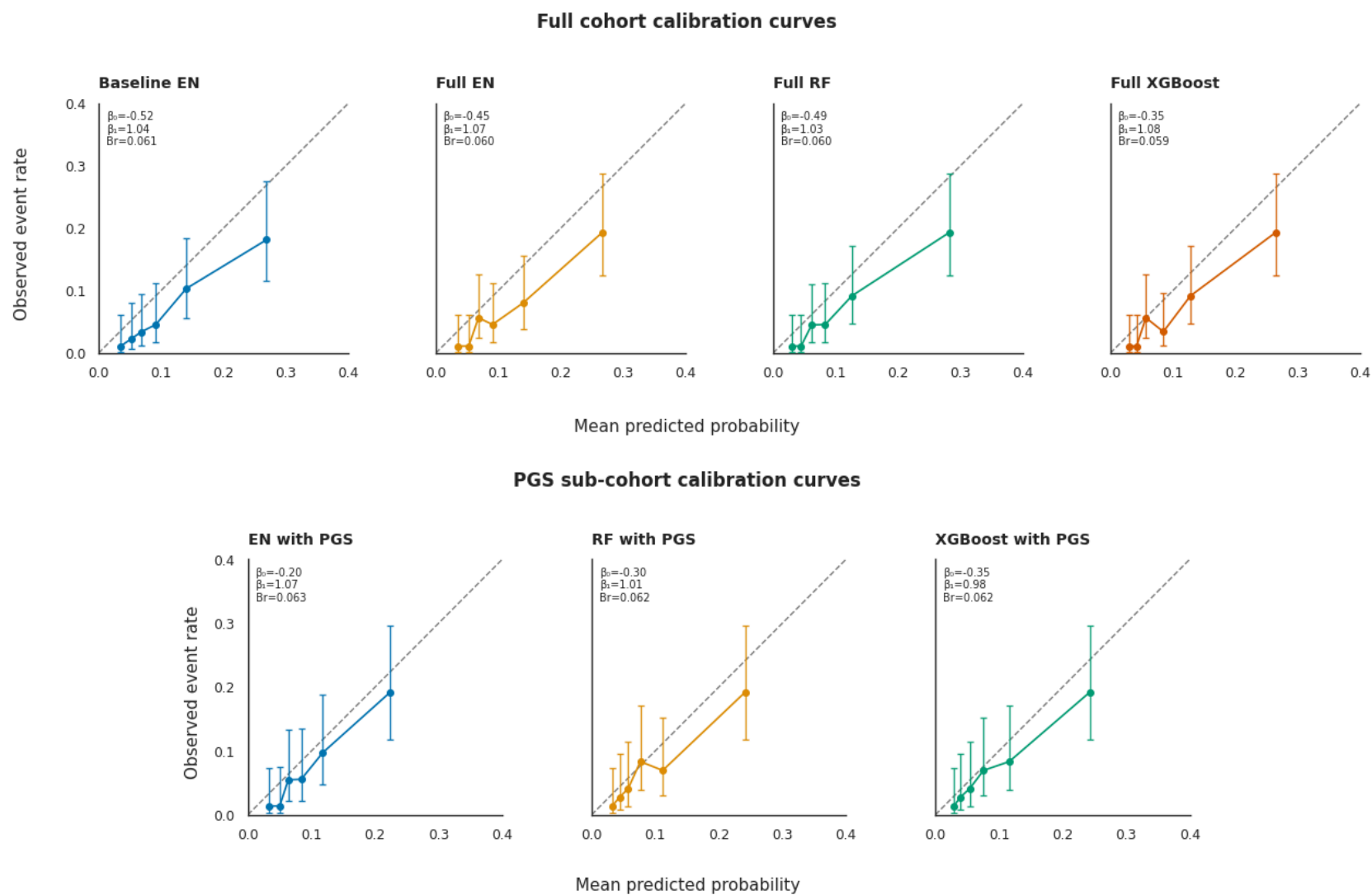
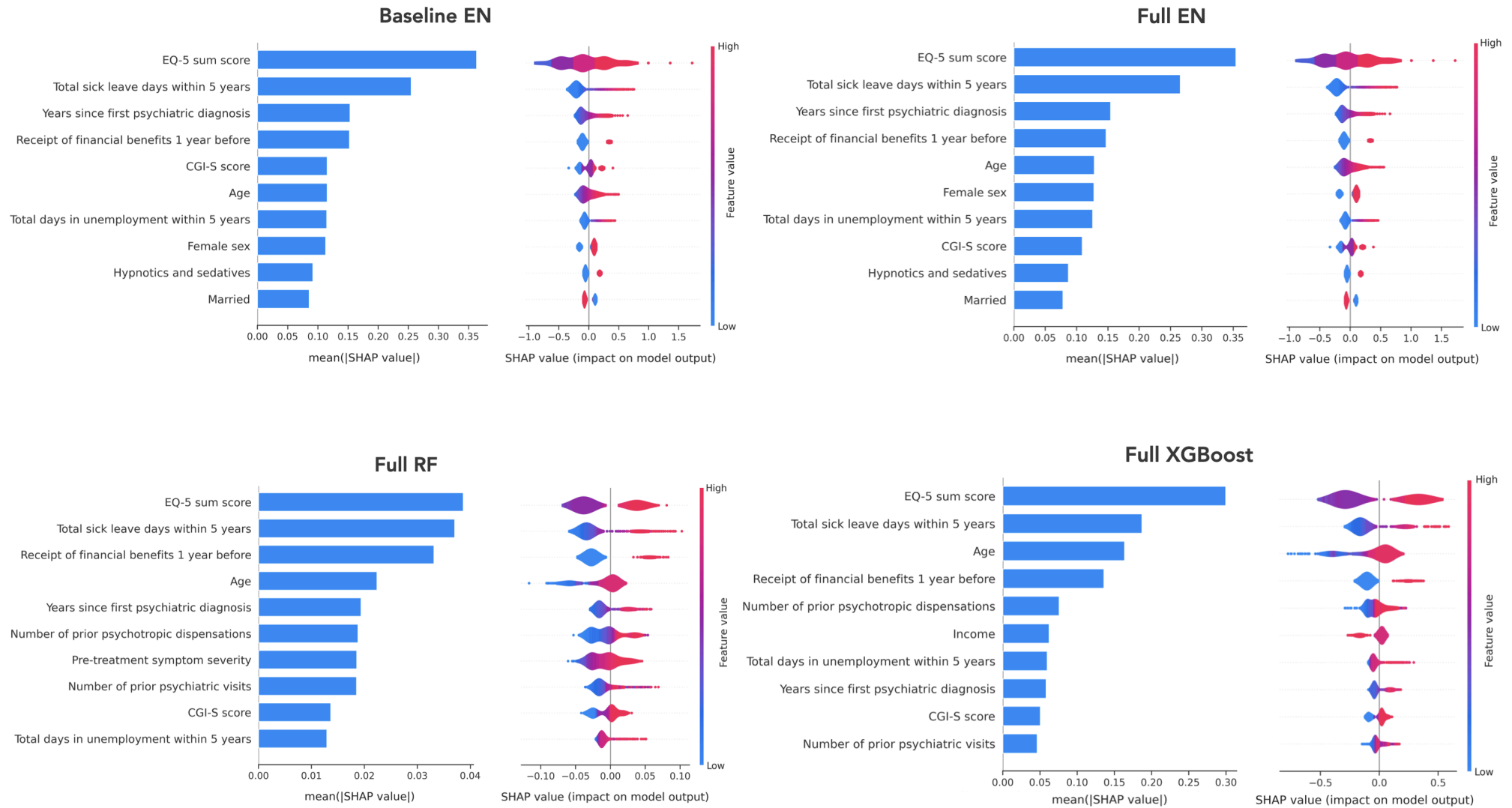


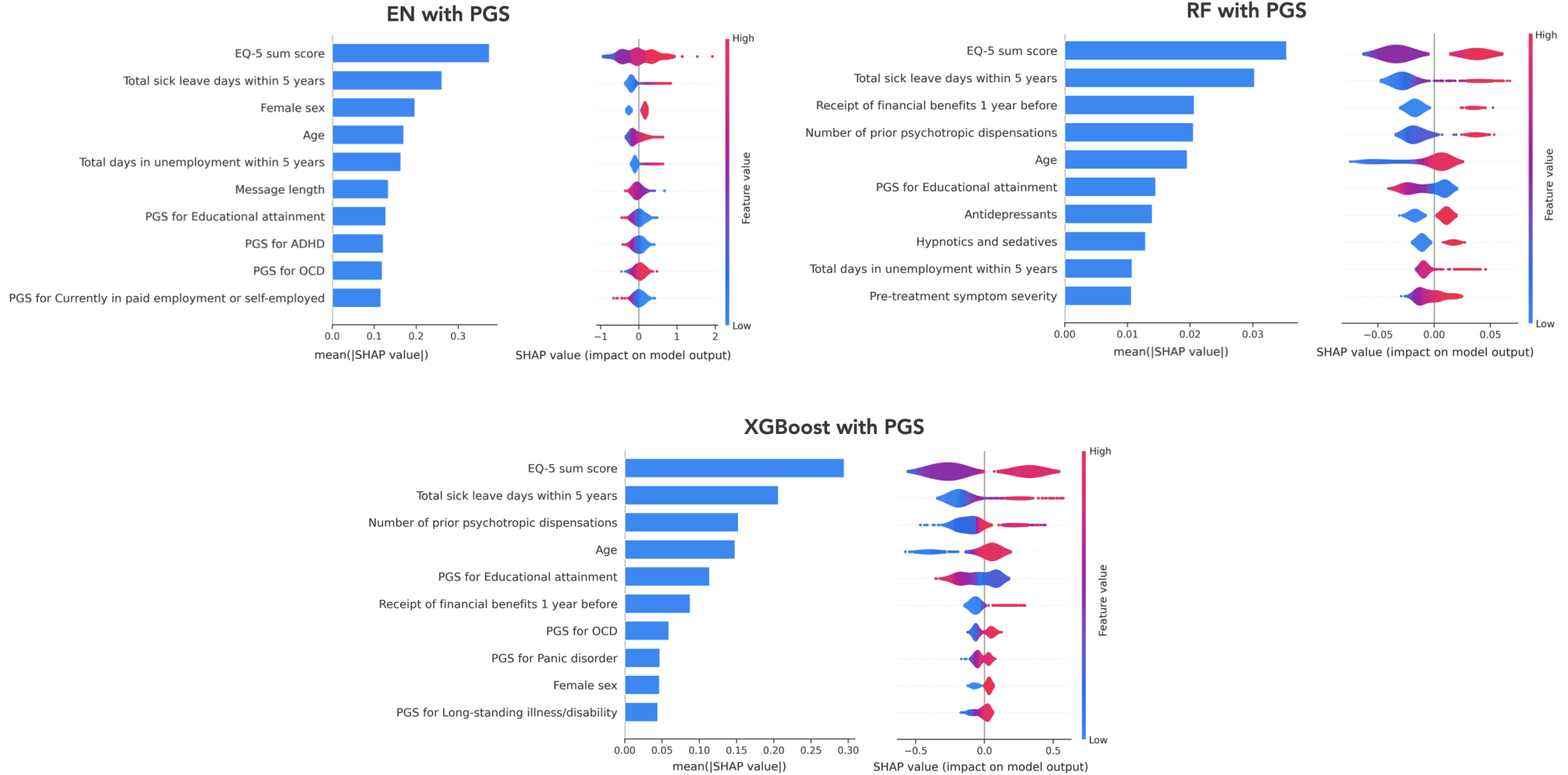
Fig. S3. Calibration curves of the recalibrated models evaluated on the holdout test set



Quantile binning with equal counts of samples (~72-88 per bin). Vertical error bars are Wilson 95% CIs that reflect binomial uncertainty for the observed event rate in each bin. Axes were restricted to 0-0.4 because predicted probabilities did not exceed ~0.35 in the holdout test set.

Fig. S4. Global predictor importances per model





Both bar plots and violin plots of the top-10 most contributing predictors are ranked in decreasing order according to mean absolute SHAP value (i.e. impact on prediction). Bar plots are based on absolute SHAP values, regardless of the direction of effect. On violin plots, the spread (width) indicates distribution of SHAP impacts. The color gradient encodes predictor values, from high (red) to low (blue). Horizontal position represents direction of impact: positive SHAP values indicate a predictor importance in favor of the positive class prediction (LMM), while negative SHAP values are in favor of the negative class prediction (non-LMM). Larger absolute SHAP values indicate a larger impact on the model output.

Table S7. Normalized global SHAP importances per model

Model	Predictor	mean(SHAP)*	SHAP (%)**
BASELINE EN	EQ-5D sum score	0.362991	17.2
	Prior 5-year sickness absence days	0.254548	12.1
	Years since first psychiatric diagnosis	0.153073	7.2
	Receipt of financial benefits 1 year before	0.152032	7.2
	CGI-S score	0.115213	5.5
	Age	0.115178	5.5
	Prior 5-year unemployment days	0.114572	5.4
	Female sex	0.112832	5.3
	Prior hypnotics and sedatives	0.091439	4.3
	Married	0.085458	4.0
	Receipt of student income 1 year before	0.075824	3.6
	Prior anxiolytics	0.066602	3.2
	Current psychotropic medication	0.052827	2.5
	Prior F3 diagnosis (excl. BPAD)	0.045791	2.2
	Immigrant 2nd generation	0.043618	2.1
	Income	0.041373	2.0
	ADHD	0.038635	1.8
	Target diagnosis for ICBT: Depression	0.038338	1.8
	Family history of psychopathology	0.036030	1.7
	Income trajectory	0.032292	1.5
	Suicide risk	0.030569	1.4

	Prior psychiatric visits	0.029619	1.4
	Prior psychotropic dispensations	0.012568	0.6
	Prior antidepressants	0.010395	0.5
FULL EN	EQ-5D sum score	0.354115	15.7
	Prior 5-year sickness absence days	0.265277	11.8
	Years since first psychiatric diagnosis	0.154018	6.8
	Receipt of financial benefits 1 year before	0.146663	6.5
	Age	0.128009	5.7
	Female sex	0.127511	5.6
	Prior 5-year unemployment days	0.125381	5.6
	CGI-S score	0.108668	4.8
	Prior hypnotics and sedatives	0.086572	3.8
	Married	0.077919	3.5
	Receipt of student income 1 year before	0.073813	3.3
	Prior anxiolytics	0.071054	3.1
	Treatment response	0.069921	3.1
	Current psychotropic medication	0.053092	2.4
	Income trajectory	0.051750	2.3
	Prior F3 diagnosis (excl. BPAD)	0.050919	2.3
	Target diagnosis for ICBT: Depression	0.042750	1.9
	Income	0.042318	1.9
	Immigrant 2nd generation	0.041668	1.8
	Suicide risk	0.040183	1.8

	Family history of psychopathology	0.037499	1.7
	ADHD	0.035424	1.6
	Prior psychiatric visits	0.033312	1.5
	Message length	0.027953	1.2
	Prior psychotropic dispensations	0.008439	0.4
	Prior antidepressants	0.003143	0.1
FULL RF	EQ-5D sum score	0.038648	9.2
	Prior 5-year sickness absence days	0.037033	8.8
	Receipt of financial benefits 1 year before	0.033116	7.9
	Age	0.022332	5.3
	Years since first psychiatric diagnosis	0.019334	4.6
	Prior psychotropic dispensations	0.018763	4.5
	Pre-treatment symptom severity	0.018544	4.4
	Prior psychiatric visits	0.018499	4.4
	CGI-S score	0.01361	3.2
	Prior 5-year unemployment days	0.012857	3.1
	AUDIT score	0.010125	2.4
	Prior antidepressants	0.009762	2.3
	Message length	0.009651	2.3
	Prior anxiolytics	0.009272	2.2
	Treatment response	0.009188	2.2
	Sickness absence >90 days 1 year before	0.009037	2.2
	Prior hypnotics and sedatives	0.009017	2.1

Receipt of student income 1 year before	0.008705	2.1
Income	0.008694	2.1
Current psychotropic medication	0.00852	2.0
Suicide risk	0.007984	1.9
Female sex	0.007942	1.9
Login duration	0.007714	1.8
Completed modules	0.006558	1.6
Psychiatric comorbidities	0.006265	1.5
Prior F4 diagnosis	0.004704	1.1
Immigrant 2nd generation	0.004446	1.1
ADHD	0.004421	1.1
Family history of psychopathology	0.004276	1.0
Inpatient psychiatric care	0.003754	0.9
Insomnia	0.003661	0.9
Married	0.0034	0.8
Prior F5 diagnosis	0.003272	0.8
Musculoskeletal diagnosis	0.002764	0.7
Income trajectory	0.002665	0.6
Target diagnosis for ICBT: Social anxiety disorder	0.002342	0.6
Prior F3 diagnosis (excl. BPAD)	0.0021	0.5
DUDIT score	0.001765	0.4
University education	0.001593	0.4
Target diagnosis for ICBT: Panic disorder	0.001576	0.4
Target diagnosis for ICBT: Depression	0.00154	0.4

	Immigrant 1st generation	0.001449	0.3
	Children	0.001422	0.3
	Single parent	0.001172	0.3
	Employed	0.001156	0.3
	Non-beneficial attrition	0.001139	0.3
	Prior antipsychotics	0.001041	0.2
	Suicide attempts	0.000637	0.2
	Prior F1 diagnosis	0.000633	0.2
	Prior F6 diagnosis	0.000616	0.1
	Unemployment >180 days 1 year before	0.00061	0.1
	ASD	0.000152	0.0
	Prior 5-year disability pension	0.000051	0.0
Full XGBoost	EQ-5D sum score	0.299292	18.1
	Prior 5-year sickness absence days	0.186822	11.3
	Age	0.163399	9.9
	Receipt of financial benefits 1 year before	0.135691	8.2
	Prior psychotropic dispensations	0.075544	4.6
	Income	0.061929	3.8
	Prior 5-year unemployment days	0.05944	3.6
	Years since first psychiatric diagnosis	0.058099	3.5
	CGI-S score	0.050172	3.0
	Prior psychiatric visits	0.045852	2.8
	Female sex	0.045388	2.8

Treatment response	0.039946	2.4
Prior anxiolytics	0.039023	2.4
Pre-treatment symptom severity	0.037307	2.3
Message length	0.033821	2.1
Completed modules	0.032774	2.0
AUDIT score	0.031563	1.9
Prior hypnotics and sedatives	0.030486	1.8
Receipt of student income 1 year before	0.020636	1.3
Login duration	0.019354	1.2
Family history of psychopathology	0.018919	1.1
ADHD	0.017024	1.0
Psychiatric comorbidities	0.014349	0.9
Immigrant 2nd generation	0.013674	0.8
Suicide risk	0.012216	0.7
Income trajectory	0.01155	0.7
Insomnia	0.010949	0.7
Prior antidepressants	0.009578	0.6
Current psychotropic medication	0.009295	0.6
Prior F4 diagnosis	0.009161	0.6
Sickness absence >90 days 1 year before	0.008835	0.5
Target diagnosis for ICBT: Panic disorder	0.008109	0.5
University education	0.006819	0.4
Employed	0.004902	0.3
Married	0.003915	0.2

	Musculoskeletal diagnosis	0.003119	0.2
	Prior F5 diagnosis	0.002545	0.2
	Target diagnosis for ICBT: Social anxiety disorder	0.002465	0.1
	Non-beneficial attrition	0.002354	0.1
	Suicide attempts	0.002285	0.1
	Inpatient psychiatric care	0.002041	0.1
	Target diagnosis for ICBT: Depression	0.001983	0.1
	Single parent	0.001186	0.1
	Prior F3 diagnosis (excl. BPAD)	0.001149	0.1
	Prior antipsychotics	0.000961	0.1
	Immigrant 1st generation	0.000946	0.1
	Children	0.000714	0.0
	Prior F6 diagnosis	0.000527	0.0
	DUDIT score	0.000496	0.0
	Prior F1 diagnosis	0.000259	0.0
	Unemployment >180 days 1 year before	0.00017	0.0
EN with PGS	EQ-5D sum score	0.374067	11.0
	Prior 5-year sickness absence days	0.260918	7.7
	Female sex	0.196502	5.8
	Age	0.170084	5.0
	Prior 5-year unemployment days	0.162818	4.8
	Message length	0.133352	3.9
	PGS for Educational attainment	0.127032	3.7

PGS for ADHD	0.121257	3.6
PGS for OCD	0.118581	3.5
PGS for Currently in paid employment or self-employed	0.115741	3.4
CGI-S score	0.111647	3.3
Income	0.110139	3.2
PGS for Long-standing illness/disability	0.108677	3.2
Years since first psychiatric diagnosis	0.105528	3.1
Login duration	0.100083	2.9
Prior antidepressants	0.089822	2.6
Target diagnosis for ICBT: Depression	0.086607	2.5
Receipt of financial benefits 1 year before	0.082651	2.4
PGS for Unable to work because of sickness/disability	0.070411	2.1
Non-beneficial attrition	0.06704	2.0
Treatment response	0.059106	1.7
Married	0.053587	1.6
Receipt of student income 1 year before	0.052654	1.5
PGS for Anorexia nervosa	0.047195	1.4
Prior psychiatric visits	0.04618	1.4
PGS for Panic disorder	0.042398	1.2
Prior hypnotics and sedatives	0.041992	1.2
PGS for Time employed in main current job	0.040483	1.2
PGS for Schizophrenia	0.039491	1.2
PGS for Tourette's & tic disorders	0.030486	0.9
Family history of psychopathology	0.02898	0.9

	Immigrant 2nd generation	0.028225	0.8
	Current psychotropic medication	0.027767	0.8
	Prior psychotropic dispensations	0.025407	0.7
	PGS for Other serious medical condition/disability	0.024753	0.7
	Psychiatric comorbidities	0.021558	0.6
	Inpatient psychiatric care	0.019213	0.6
	ADHD	0.017874	0.5
	AUDIT score	0.012495	0.4
	PGS for Bipolar disorder	0.009918	0.3
	PGS for Household income	0.008782	0.3
	Income trajectory	0.005295	0.2
	Suicide risk	0.002194	0.1
	Insomnia	0.000908	0.0
	DUDIT score	0.000322	0.0
	PGS Cross-disorder	0.000268	0.0
RF with PGS	EQ-5D sum score	0.035356	8.5
	Prior 5-year sickness absence days	0.030217	7.3
	Receipt of financial benefits 1 year before	0.020637	5.0
	Prior psychotropic dispensations	0.020507	4.9
	Age	0.019524	4.7
	PGS for Educational attainment	0.014484	3.5
	Prior antidepressants	0.01393	3.3
	Prior hypnotics and sedatives	0.012851	3.1

Prior 5-year unemployment days	0.010704	2.6
Pre-treatment symptom severity	0.010617	2.5
Years since first psychiatric diagnosis	0.010063	2.4
Prior anxiolytics	0.01005	2.4
Prior psychiatric visits	0.009664	2.3
PGS for Currently in paid employment or self-employed	0.008329	2.0
PGS for OCD	0.008132	2.0
PGS for Long-standing illness/disability	0.007573	1.8
Current psychotropic medication	0.007382	1.8
Sickness absence >90 days 1 year before	0.006781	1.6
CGI-S score	0.006568	1.6
PGS for Anxiety disorders	0.00641	1.5
Message length	0.005891	1.4
PGS for Panic disorder	0.005724	1.4
Income	0.005546	1.3
Female sex	0.005491	1.3
AUDIT score	0.005477	1.3
PGS for Schizophrenia	0.005377	1.3
Login duration	0.005293	1.3
PGS for Bipolar disorder	0.005221	1.3
PGS for Household income	0.004827	1.2
PGS for Tourette's & tic disorders	0.004769	1.1
PGS for Intelligence	0.004766	1.1
Treatment response	0.004396	1.1

PGS for SUD	0.004113	1.0
PGS for Other serious medical condition/disability	0.004105	1.0
PGS for Family history of disabilities	0.00396	1.0
PGS for ADHD	0.003874	0.9
PGS Cross-disorder	0.003853	0.9
PGS for Time employed in main current job	0.003849	0.9
PGS for Unable to work because of sickness/disability	0.003756	0.9
Prior F3 diagnosis (excl. BPAD)	0.003665	0.9
PGS for Depression	0.00363	0.9
PGS for ASD	0.00354	0.8
PGS for Disability living allowance	0.003263	0.8
Immigrant 2nd generation	0.003241	0.8
Suicide risk	0.00322	0.8
Psychiatric comorbidities	0.002997	0.7
PGS for Anorexia nervosa	0.002931	0.7
PGS for PTSD	0.002845	0.7
Musculoskeletal diagnosis	0.002779	0.7
Receipt of student income 1 year before	0.00268	0.6
Completed modules	0.002665	0.6
Inpatient psychiatric care	0.00226	0.5
Prior F4 diagnosis	0.0021	0.5
Non-beneficial attrition	0.001983	0.5
Children	0.001974	0.5
Family history of psychopathology	0.001704	0.4

	ADHD	0.001466	0.4
	Income trajectory	0.001136	0.3
	Insomnia	0.001124	0.3
	Married	0.000814	0.2
	Target diagnosis for ICBT: Depression	0.000552	0.1
	Target diagnosis for ICBT: Panic disorder	0.000498	0.1
	Target diagnosis for ICBT: Social anxiety disorder	0.000495	0.1
	Prior F5 diagnosis	0.000444	0.1
	University education	0.000407	0.1
	Single parent	0.000382	0.1
	Prior F1 diagnosis	0.000378	0.1
	DUDIT score	0.000338	0.1
	Suicide attempts	0.000316	0.1
	Prior antipsychotics	0.000239	0.1
	Unemployment >180 days 1 year before	0.000179	0.0
	Employed	0.00016	0.0
	Immigrant 1st generation	0.000145	0.0
	Prior F6 diagnosis	0.000053	0.0
XGBoost with PGS	EQ-5D sum score	0.294368	15.5
	Prior 5-year sickness absence days	0.205799	10.8
	Prior psychotropic dispensations	0.152337	8.0
	Age	0.14775	7.8
	PGS for Educational attainment	0.113673	6.0

Receipt of financial benefits 1 year before	0.087535	4.6
PGS for OCD	0.059065	3.1
PGS for Panic disorder	0.046771	2.5
Female sex	0.046273	2.4
PGS for Long-standing illness/disability	0.043949	2.3
Prior 5-year unemployment days	0.043809	2.3
Income	0.040398	2.1
Message length	0.038041	2.0
PGS for Depression	0.030235	1.6
PGS for Currently in paid employment or self-employed	0.029077	1.5
Years since first psychiatric diagnosis	0.027997	1.5
Current psychotropic medication	0.027799	1.5
PGS for ADHD	0.026699	1.4
Treatment response	0.025271	1.3
Pre-treatment symptom severity	0.025016	1.3
CGI-S score	0.022324	1.2
PGS for Bipolar disorder	0.021316	1.1
PGS for Disability living allowance	0.020411	1.1
PGS for Anxiety disorders	0.019716	1.0
Completed modules	0.018526	1.0
PGS for Household income	0.017873	0.9
PGS for Schizophrenia	0.016183	0.9
Login duration	0.01573	0.8
Prior anxiolytics	0.01536	0.8

PGS for Intelligence	0.015267	0.8
Prior psychiatric visits	0.015069	0.8
Prior hypnotics and sedatives	0.014826	0.8
PGS Cross-disorder	0.014162	0.7
PGS for Other serious medical condition/disability	0.014	0.7
PGS for Unable to work because of sickness/disability	0.012806	0.7
PGS for Anorexia nervosa	0.010529	0.6
Psychiatric comorbidities	0.010238	0.5
Prior antidepressants	0.009809	0.5
PGS for ASD	0.009642	0.5
PGS for Family history of disabilities	0.009615	0.5
PGS for Time employed in main current job	0.009266	0.5
AUDIT score	0.008437	0.4
PGS for Tourette's & tic disorders	0.007818	0.4
PGS for PTSD	0.005809	0.3
PGS for SUD	0.005415	0.3
Family history of psychopathology	0.004133	0.2
Insomnia	0.003931	0.2
Receipt of student income 1 year before	0.003738	0.2
Suicide risk	0.003734	0.2
Sickness absence >90 days 1 year before	0.003249	0.2
ADHD	0.002707	0.1
Non-beneficial attrition	0.002637	0.1
Immigrant 2nd generation	0.002591	0.1

Target diagnosis for ICBT: Panic disorder	0.002479	0.1
Target diagnosis for ICBT: Social anxiety disorder	0.002466	0.1
Prior F3 diagnosis (excl. BPAD)	0.002134	0.1
Inpatient psychiatric care	0.001936	0.1
Unemployment >180 days 1 year before	0.001778	0.1
Income trajectory	0.001679	0.1
Employed	0.001508	0.1
Prior F4 diagnosis	0.001256	0.1
Immigrant 1st generation	0.000937	0.0
Target diagnosis for ICBT: Depression	0.000728	0.0
Prior F5 diagnosis	0.000624	0.0
Married	0.000621	0.0
Suicide attempts	0.000554	0.0
DUDIT score	0.000528	0.0
Prior antipsychotics	0.000493	0.0
Prior F6 diagnosis	0.000248	0.0
University education	0.00024	0.0
ASD	0.000181	0.0
Single parent	0.00017	0.0
Musculoskeletal diagnosis	0.000129	0.0

**Mean(|SHAP|) values are reported in the native output space used by each explainer (LinearExplainer for EN and TreeExplainer for RF and XGBoost). Therefore, raw mean(|SHAP|) magnitudes are not directly comparable across model types.*

***Normalized global SHAP importances for each algorithm, reported as mean absolute SHAP and as percentage of the total mean absolute SHAP across predictors included in the model. The table displays only predictors with non-zero mean(|SHAP|), hence SHAP% values may not sum to exactly 100%.*

References

1. Ludvigsson JF, Svedberg P, Olén O, et al (2019) The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research. *Eur J Epidemiol* 34:423–437. <https://doi.org/10.1007/s10654-019-00511-8>
2. Watanabe K, Stringer S, Frei O, et al (2019) A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet* 51:1339–1348. <https://doi.org/10.1038/s41588-019-0481-0>
3. Adams MJ, Streit F, Meng X, et al (2025) Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. *Cell* 188:640–652.e9. <https://doi.org/10.1016/j.cell.2024.12.002>
4. Otowa T, Hek K, Lee M, et al (2016) Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry* 21:1391–1399. <https://doi.org/10.1038/mp.2015.197>
5. Forstner AJ, Awasthi S, Wolf C, et al (2021) Genome-wide association study of panic disorder reveals genetic overlap with neuroticism and depression. *Mol Psychiatry* 26:4179–4190. <https://doi.org/10.1038/s41380-019-0590-2>
6. International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OC GAS) (2018) Revealing the complex genetic architecture of obsessive–compulsive disorder using meta-analysis. *Mol Psychiatry* 23:1181–1188. <https://doi.org/10.1038/mp.2017.154>
7. Yu D, Sul JH, Tsetsos F, et al (2019) Interrogating the Genetic Determinants of Tourette’s Syndrome and Other Tic Disorders Through Genome-Wide Association Studies. *Am J Psychiatry* 176:217–227. <https://doi.org/10.1176/appi.ajp.2018.18070857>
8. Nievergelt CM, Maihofer AX, Atkinson EG, et al (2024) Genome-wide association analyses identify 95 risk loci and provide insights into the neurobiology of post-traumatic stress disorder. *Nat Genet* 56:792–808. <https://doi.org/10.1038/s41588-024-01707-9>
9. Demontis D, Walters GB, Athanasiadis G, et al (2023) Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat Genet* 55:198–208. <https://doi.org/10.1038/s41588-022-01285-8>

10. Grove J, Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium, BUPGEN, et al (2019) Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* 51:431–444. <https://doi.org/10.1038/s41588-019-0344-8>
11. O’Connell KS, Koromina M, Van Der Veen T, et al (2025) Genomics yields biological and phenotypic insights into bipolar disorder. *Nature* 639:968–975. <https://doi.org/10.1038/s41586-024-08468-9>
12. Watson HJ, Anorexia Nervosa Genetics Initiative, Eating Disorders Working Group of the Psychiatric Genomics Consortium, et al (2019) Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat Genet* 51:1207–1214. <https://doi.org/10.1038/s41588-019-0439-2>
13. Trubetskoy V, Pardiñas AF, Qi T, et al (2022) Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* 604:502–508. <https://doi.org/10.1038/s41586-022-04434-5>
14. Hatoum AS, Colbert SMC, Johnson EC, et al (2023) Multivariate genome-wide association meta-analysis of over 1 million subjects identifies loci underlying multiple substance use disorders. *Nat Ment Health* 1:210–223. <https://doi.org/10.1038/s44220-023-00034-y>
15. Lee PH, Anttila V, Won H, et al (2019) Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* 179:1469-1482.e11. <https://doi.org/10.1016/j.cell.2019.11.020>
16. Savage JE, Jansen PR, Stringer S, et al (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* 50:912–919. <https://doi.org/10.1038/s41588-018-0152-6>
17. Okbay A, Wu Y, Wang N, et al (2022) Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat Genet* 54:437–449. <https://doi.org/10.1038/s41588-022-01016-z>
18. Försäkringskassan (2025) Social Insurance in Figures 2025. <https://www.forsakringskassan.se/download/18.242206fc198acf662d68d/1757339678327/social-insurance-in-figures-2025.pdf>. Accessed 4 Feb 2026
19. Försäkringskassan Statistikdatabas. <https://www.forsakringskassan.se/statistik-och-analys/statistikdatabas#!/sjuk/sa-bestand-diagnos>. Accessed 30 Jan 2026

20. Mather L, Kärkkäinen S, Narusyte J, et al (2020) Sick leave due to back pain, common mental disorders and disability pension: Common genetic liability. *Eur J Pain* 24:1892–1901. <https://doi.org/10.1002/ejp.1635>
21. Kärkkäinen S, Bergström J, Ropponen A, et al (2023) Sickness absence transitions among Swedish twins with back, neck or shoulder pain and common mental disorders applying a multi-state approach. *Sci Rep* 13:10520. <https://doi.org/10.1038/s41598-023-37572-5>
22. Melkevik O, Clausen T, Pedersen J, et al (2018) Comorbid symptoms of depression and musculoskeletal pain and risk of long term sickness absence. *BMC Public Health* 18:981. <https://doi.org/10.1186/s12889-018-5740-y>
23. Dorner TE, Alexanderson K, Svedberg P, et al (2016) Synergistic effect between back pain and common mental disorders and the risk of future disability pension: a nationwide study from Sweden. *Psychol Med* 46:425–436. <https://doi.org/10.1017/S003329171500197X>
24. Zantvoort K, Hentati Isacsson N, Funk B, Kaldo V (2024) Dataset size versus homogeneity: A machine learning study on pooling intervention data in e-mental health dropout predictions. *Digit Health* 10:20552076241248920. <https://doi.org/10.1177/20552076241248920>