Supplementary Table S1. Baseline comparison between the training and internal validation datasets

Characteristic	Category/statistic	Training (*n* = 9,424), *n* (%) or median (*IQR*)	Internal validation (*n* = 4,038), *n* (%) or median (*IQR*)	Z/χ^2	*P* value	Bonferroni *P*
BMI	Median (*IQR*)	22.90 (21.30, 24.50)	22.90 (21.30, 24.50)	0.211	0.833	1.000
ASD score	Median (*IQR*)	19.00 (19.00, 23.00)	19.00 (19.00, 23.00)	2.306	0.021	0.317
SDS score	Median (*IQR*)	40.00 (31.25, 60.00)	41.25 (31.25, 61.25)	-1.856	0.063	0.952
SAS score	Median (*IQR*)	35.00 (28.75, 43.75)	35.00 (28.75, 43.75)	-1.018	0.309	1.000
Age, y	≤ 25	5,497 (58.3)	2,328 (57.7)	1.649	0.439	1.000
	26–30	2,582 (27.4)	1,149 (28.5)			
	≥ 31	1,345 (14.3)	561 (13.9)			
Sex	Female	218 (2.3)	101 (2.5)	0.432	0.552	1.000
	Male	9,206 (97.7)	3,937 (97.5)			
Marital status	Unmarried	6,757 (71.7)	2,932 (72.6)	1.530	0.465	1.000
	Married	2,621 (27.8)	1,090 (27.0)			
	Other	46 (0.5)	16 (0.4)			
Ethnicity	Minority ethnicity	959 (10.2)	389 (9.6)	0.924	0.352	1.000
	Han ethnicity	8,465 (89.8)	3,649 (90.4)			
Education	High school or less	2,426 (25.7)	1,073 (26.6)	1.012	0.325	1.000
	College or above	6,998 (74.3)	2,965 (73.4)			
Monthly family income per capita, CNY	≤ 2,999	2,493 (26.5)	992 (24.6)	7.212	0.065	1.000
	3,000–4,999	2,983 (31.7)	1,268 (31.4)			
	5,000–7,999	1,934 (20.5)	859 (21.3)			
	≥ 8,000	2,014 (21.4)	919 (22.8)			
Recent trauma exposure	No	9,165 (97.3)	3,936 (97.5)	0.535	0.501	1.000
	Yes	259 (2.7)	102 (2.5)			
Previous mental illness	No	9,419 (99.9)	4,031 (99.8)	Fisher exact/ simulated	0.052	0.785
	Yes	5 (0.1)	7 (0.2)			
Smoking status	No	5,137 (54.5)	2,171 (53.8)	0.633	0.437	1.000
	Yes	4,287 (45.5)	1,867 (46.2)			
Passive smoking	No	8,196 (87.0)	3,491 (86.5)	0.657	0.434	1.000
	Yes	1,228 (13.0)	547 (13.5)			
Alcohol use	No	8,985 (95.3)	3,867 (95.8)	1.172	0.300	1.000
	Yes	439 (4.7)	171 (4.2)			
Family history of mental illness	No	9,414 (99.9)	4,029 (99.8)	2.735	0.161	1.000
	Yes	10 (0.1)	9 (0.2)			

Note. BMI, body mass index; ASD, acute stress disorder; SDS, self-rating depression scale; SAS, self-rating anxiety scale.

Supplementary Table S2. Baseline comparison between the derivation dataset and the external validation dataset

Characteristic	Category/ statistic	Derivation dataset (*n* = 13,462), *n* (%) or median (*IQR*)	External validation (*n* = 8,757), *n* (%) or median (*IQR*)	Z/χ^2	*P* value	Bonferroni *P*
BMI	Median (*IQR*)	22.90 (21.30, 24.50)	22.60 (21.20, 24.20)	4.574	< 0.001	< 0.001
ASD score	Median (*IQR*)	19.00 (19.00, 23.00)	19.00 (19.00, 21.00)	11.008	< 0.001	< 0.001
SDS score	Median (*IQR*)	40.00 (31.25, 60.00)	37.50 (28.75, 55.00)	11.001	< 0.001	< 0.001
SAS score	Median (*IQR*)	35.00 (28.75, 43.75)	36.25 (28.75, 43.75)	-6.073	< 0.001	< 0.001
Age, y	≤ 25	7,825 (58.1)	6,197 (70.8)	365.233	< 0.001	< 0.001
	26-30	3,731 (27.7)	1,727 (19.7)			
	≥ 31	1,906 (14.2)	833 (9.5)			
Sex	Female	319 (2.4)	120 (1.4)	27.356	< 0.001	< 0.001
	Male	13,143 (97.6)	8,637 (98.6)			
Marital status	Unmarried	9,689 (72.0)	7,123 (81.3)	253.685	< 0.001	< 0.001
	Married	3,711 (27.6)	1,601 (18.3)			
	Other	62 (0.5)	33 (0.4)			
Ethnicity	Minority ethnicity	1,348 (10.0)	632 (7.2)	51.109	< 0.001	< 0.001
	Han ethnicity	12,114 (90.0)	8,125 (92.8)			
Education	High school or less	3,499 (26.0)	3,668 (41.9)	613.435	< 0.001	< 0.001
	College or above	9,963 (74.0)	5,089 (58.1)			
Monthly family income per capita, CNY	≤ 2,999	3,485 (25.9)	2,548 (29.1)	74.203	< 0.001	< 0.001
	3,000–4,999	4,251 (31.6)	2,825 (32.3)			
	5,000–7,999	2,793 (20.7)	1,866 (21.3)			
	≥ 8,000	2,933 (21.8)	1,518 (17.3)			
Recent trauma exposure	No	13,101 (97.3)	8,476 (96.8)	5.256	0.024	0.365
	Yes	361 (2.7)	281 (3.2)			
Previous mental illness	No	13450 (99.9)	8,750 (99.9)	0.053	1.000	1.000
	Yes	12 (0.1)	7 (0.1)			
Smoking status	No	7,308 (54.3)	5,106 (58.3)	34.802	< 0.001	< 0.001
	Yes	6,154 (45.7)	3,651 (41.7)			
Passive smoking	No	11,687 (86.8)	7,981 (91.1)	97.599	< 0.001	< 0.001
	Yes	1,775 (13.2)	776 (8.9)			
Alcohol use	No	12,852 (95.5)	8,262 (94.3)	14.117	< 0.001	0.003
	Yes	610 (4.5)	495 (5.7)			
Family history of mental illness	No	13,443 (99.9)	8,743 (99.8)	0.126	0.860	1.000
	Yes	19 (0.1)	14 (0.2)			

Note. BMI, body mass index; ASD, acute stress disorder; SDS, self-rating depression scale; SAS, self-rating anxiety scale.

Supplementary Table S3. Multi-indicator ranking of PCA-based models in external validation (Youden threshold; top five models)

Model	AUC	Sensitivity	Specificity	PPV	MCC	Brier	Sum of ranking differences	Pareto front
XGBoost	0.8931	0.6596	0.9788	0.2520	0.3981	0.0078	28.5000	Yes
Random Forest	0.8903	0.7447	0.9249	0.0971	0.2510	0.0127	39.0000	Yes
GBM	0.8929	0.6915	0.9462	0.1224	0.2753	0.0253	47.0000	Yes
Linear SVM	0.8894	0.7234	0.9222	0.0916	0.2389	0.0520	48.5000	Yes
Kernel SVM	0.8773	0.7128	0.9298	0.0993	0.2483	0.0439	51.5000	Yes

Note. AUC, area under the curve; PPV, positive predictive value; MCC, Matthews correlation coefficient.

Supplementary Table S4. Extended external-validation metrics for PCA-based models

Model	Balanced accuracy	F1-score	MCC	Brier score	AUPRC
Linear SVM	0.8228	0.1627	0.2389	0.0520	0.4581
XGBoost	0.8192	0.3647	0.3981	0.0078	0.4389
Multilayer Perceptron	0.8063	0.2206	0.2807	0.0153	0.3468
Random Forest	0.8348	0.1718	0.2510	0.0127	0.3498
GBM	0.8189	0.2080	0.2753	0.0253	0.4144
Kernel SVM	0.8213	0.1743	0.2483	0.0439	0.4015
AdaBoost	0.8065	0.1307	0.2049	0.0667	0.3911
Logistic Regression	0.8069	0.2652	0.3160	0.0796	0.4248
Decision Tree	0.7897	0.0965	0.1672	0.0606	0.1817

Note. AUC, accuracy, sensitivity, specificity, PPV, and NPV are reported in the main performance tables and are not repeated here. Balanced accuracy, F1-score, and MCC are threshold-based metrics calculated using the Youden threshold. Brier score and AUPRC were calculated from predicted probabilities. MCC, Matthews correlation coefficient; AUPRC, area under the precision-recall curve; PPV, positive predictive value; NPV, negative predictive value.

Supplementary Table S5. DeLong tests for AUC comparisons in the internal and external validation datasets

Validation dataset	Route	Reference model	Compared model	Reference AUC	Compared AUC	DeLong *P*	Holm-adjusted *P*
Internal validation	PCA	Linear SVM	Logistic Regression	0.9829	0.9368	< 0.001	< 0.001
Internal validation	PCA	Linear SVM	Decision Tree	0.9829	0.9247	< 0.001	0.0032
Internal validation	PCA	Linear SVM	AdaBoost	0.9829	0.9619	< 0.001	0.0038
Internal validation	PCA	Linear SVM	Kernel SVM	0.9829	0.9611	0.0232	0.1162
Internal validation	PCA	Linear SVM	GBM	0.9829	0.9694	0.0859	0.3435
Internal validation	PCA	Linear SVM	Multilayer Perceptron	0.9829	0.9662	0.1464	0.4392
Internal validation	PCA	Linear SVM	Random Forest	0.9829	0.9787	0.3378	0.6755
Internal validation	PCA	Linear SVM	XGBoost	0.9829	0.9805	0.6254	0.6755
Internal validation	FA	Random Forest	Linear SVM	0.9514	0.7614	< 0.001	< 0.001
Internal validation	FA	Random Forest	Logistic Regression	0.9514	0.7637	< 0.001	< 0.001
Internal validation	FA	Random Forest	Kernel SVM	0.9514	0.7741	< 0.001	< 0.001
Internal validation	FA	Random Forest	Decision Tree	0.9514	0.8677	< 0.001	< 0.001
Internal validation	FA	Random Forest	AdaBoost	0.9514	0.9078	0.0017	0.0068
Internal validation	FA	Random Forest	Multilayer Perceptron	0.9514	0.9321	0.0507	0.1521
Internal validation	FA	Random Forest	GBM	0.9514	0.9297	0.0673	0.1521
Internal validation	FA	Random Forest	XGBoost	0.9514	0.9450	0.4295	0.4295
External validation	PCA	XGBoost	GBM	0.8931	0.8929	0.9772	1.0000

External validation	PCA	XGBoost	Random Forest	0.8931	0.8903	0.8354	1.0000
External validation	PCA	XGBoost	Linear SVM	0.8931	0.8894	0.7615	1.0000
External validation	PCA	XGBoost	AdaBoost	0.8931	0.8780	0.0331	0.2320
External validation	PCA	XGBoost	Kernel SVM	0.8931	0.8773	0.1355	0.5418
External validation	PCA	XGBoost	Logistic Regression	0.8931	0.8745	0.0903	0.4514
External validation	PCA	XGBoost	Decision Tree	0.8931	0.8568	0.0517	0.3101
External validation	PCA	XGBoost	Multilayer Perceptron	0.8931	0.8390	0.0170	0.1364
External validation	FA	XGBoost	AdaBoost	0.8832	0.8761	0.5783	0.7743
External validation	FA	XGBoost	GBM	0.8832	0.8705	0.2683	0.7743
External validation	FA	XGBoost	Decision Tree	0.8832	0.8591	0.0743	0.2974
External validation	FA	XGBoost	Multilayer Perceptron	0.8832	0.8590	0.2581	0.7743
External validation	FA	XGBoost	Random Forest	0.8832	0.8425	0.0191	0.1149
External validation	FA	XGBoost	Kernel SVM	0.8832	0.8404	0.0221	0.1149
External validation	FA	XGBoost	Linear SVM	0.8832	0.8286	0.0042	0.0333
External validation	FA	XGBoost	Logistic Regression	0.8832	0.8281	0.0042	0.0333

Note. AUC, area under the curve.

Supplementary Table S6. Calibration-related indices and Brier scores for PCA-based models in external validation

Model	AUC	Brier	Calibration intercept	Calibration slope	HL statistic	HL *P*
XGBoost	0.8931	0.0078	-0.3141	0.9136	32.2369	< 0.001
GBM	0.8929	0.0253	-2.9943	0.8697	704.1785	< 0.001
Random Forest	0.8903	0.0127	-1.7633	0.6225	220,000,000,203.1940	< 0.001
Linear SVM	0.8894	0.0520	-3.5395	1.2141	2,052.6096	< 0.001
AdaBoost	0.8780	0.0667	-3.8376	0.9948	2,373.7012	< 0.001
Kernel SVM	0.8773	0.0439	-3.8041	0.7432	1,429.9853	< 0.001
Logistic Regression	0.8745	0.0796	-4.1218	0.9534	3,042.0861	< 0.001
Decision Tree	0.8568	0.0606	-4.1532	0.6966	2,295.7735	< 0.001
Multilayer Perceptron	0.8390	0.0153	-3.5138	0.3020	410,000,007,541.2940	< 0.001

Note. Brier score assesses prediction error, with lower values indicating smaller prediction error. Calibration intercept and slope describe agreement between predicted probabilities and observed outcomes. HL, Hosmer-Lemeshow goodness-of-fit test; AUC, area under the curve.

Supplementary Table S7. Calibration-related indices and Brier scores for factor-analysis-based models in external validation

Model	AUC	Brier	Calibration intercept	Calibration slope	HL statistic	HL *P*
XGBoost	0.8832	0.0084	-0.4733	1.0226	23.8954	0.0024
AdaBoost	0.8761	0.0766	-3.8303	1.2714	2,955.6499	< 0.001
GBM	0.8705	0.0480	-3.5508	0.9423	1,698.0018	< 0.001
Decision Tree	0.8591	0.0854	-4.3002	0.8536	4,231.8335	< 0.001
Multilayer Perceptron	0.8590	0.0617	-4.1292	0.7526	2,297.3993	< 0.001
Random Forest	0.8425	0.0204	-2.2345	0.4775	450,000,000,440.6190	< 0.001
Kernel SVM	0.8404	0.1627	-4.4488	1.1490	5,941.9054	< 0.001
Linear SVM	0.8286	0.1946	-4.4667	1.8217	7,137.8754	< 0.001
Logistic Regression	0.8281	0.1967	-4.4786	1.7924	7,248.9639	< 0.001

Note. Brier score assesses prediction error, with lower values indicating smaller prediction error. Calibration intercept and slope describe agreement between predicted probabilities and observed outcomes. HL, Hosmer-Lemeshow goodness-of-fit test; AUC, area under the curve.

Supplementary Table S8. Decision-curve analysis net-benefit summary in the external validation dataset

Route	Model	NB at 1%	NB at 3%	NB at 5%	NB at 10%	NB at 20%	Max NB (1%–20%)
PCA	XGBoost	0.0071	0.0062	0.0056	0.0049	0.0036	0.0071
PCA	GBM	0.0036	-0.0028	-0.0058	-0.0110	-0.0139	0.0036
PCA	Random Forest	0.0057	0.0029	0.0012	-0.0008	-0.0018	0.0057
PCA	Linear SVM	0.0007	-0.0191	-0.0352	-0.0582	-0.0714	0.0007
PCA	AdaBoost	0.0029	-0.0085	-0.0178	-0.0375	-0.0653	0.0029
PCA	Kernel SVM	0.0032	-0.0052	-0.0107	-0.0187	-0.0300	0.0032
PCA	Logistic Regression	0.0013	-0.0124	-0.0229	-0.0434	-0.0726	0.0013
PCA	Decision Tree	0.0010	-0.0068	-0.0070	-0.0133	-0.0278	0.0010
PCA	Multilayer Perceptron	0.0067	0.0059	0.0053	0.0041	0.0019	0.0067
FA	XGBoost	0.0063	0.0059	0.0051	0.0042	0.0027	0.0063
FA	AdaBoost	0.0007	-0.0198	-0.0395	-0.0717	-0.0990	0.0007
FA	GBM	0.0011	-0.0124	-0.0195	-0.0307	-0.0339	0.0011
FA	Decision Tree	0.0008	-0.0197	-0.0395	-0.0941	-0.0697	0.0008
FA	Multilayer Perceptron	0.0023	-0.0094	-0.0187	-0.0325	-0.0462	0.0023
FA	Random Forest	0.0042	0.0002	-0.0023	-0.0055	-0.0100	0.0042
FA	Kernel SVM	0.0007	-0.0199	-0.0413	-0.0990	-0.2105	0.0007
FA	Linear SVM	0.0007	-0.0199	-0.0413	-0.0991	-0.1992	0.0007
FA	Logistic Regression	0.0007	-0.0199	-0.0413	-0.0991	-0.1997	0.0007

Note. NB, net benefit. Decision-curve analysis was used to compare the potential clinical net benefit of each model across threshold probabilities relative to default strategies such as treating all or treating none. Because the external validation dataset had low PTSD prevalence and the model was intended for screening, the low-to-moderate threshold range of 1%–20% was summarised.

Supplementary Table S9. Extended external-validation metrics for factor-analysis-based models

Model	Balanced accuracy	F1-score	MCC	Brier score	AUPRC
XGBoost	0.8317	0.1781	0.2553	0.0084	0.3529
Random Forest	0.8200	0.1307	0.2095	0.0204	0.2243
Multilayer Perceptron	0.7944	0.1462	0.2145	0.0617	0.3937
GBM	0.8149	0.1398	0.2160	0.0480	0.2844
AdaBoost	0.8025	0.1224	0.1959	0.0766	0.3402
Decision Tree	0.8203	0.1545	0.2309	0.0854	0.0793
Kernel SVM	0.7601	0.1395	0.1966	0.1627	0.2654
Logistic Regression	0.7121	0.0597	0.1057	0.1967	0.2647
Linear SVM	0.7407	0.0982	0.1529	0.1946	0.2778

Note. AUC, accuracy, sensitivity, specificity, PPV, and NPV are reported in the main performance tables and are not repeated here. Balanced accuracy, F1-score, and MCC are threshold-based metrics calculated using the Youden threshold. Brier score and AUPRC were calculated from predicted probabilities. MCC, Matthews correlation coefficient; AUPRC, area under the precision-recall curve; PPV, positive predictive value; NPV, negative predictive value.

Supplementary Table S10. Bootstrap stability assessment of original-variable SHAP importance rankings

Predictor	Original mean |SHAP|	Original rank	Bootstrap mean |SHAP|	Bootstrap SD	95% *CI* lower	95% *CI* upper	Top-8 frequency	Median rank	Mean rank
ASD score	0.892272	1	0.892232	0.000979	0.890292	0.894134	1.000	1	1
SAS score	0.032839	2	0.032834	0.000388	0.032115	0.033615	1.000	2	2
SDS score	0.014386	3	0.014378	0.000111	0.014168	0.014602	1.000	3	3
Passive smoking	0.005760	4	0.005760	0.000057	0.005647	0.005868	1.000	4	4
Monthly income	0.005429	5	0.005431	0.000040	0.005354	0.005509	1.000	5	5
Smoking	0.004563	6	0.004563	0.000069	0.004436	0.004706	1.000	6	6
BMI	0.002986	7	0.002988	0.000102	0.002790	0.003198	1.000	7	7.235
Age	0.002905	8	0.002904	0.000047	0.002820	0.003002	1.000	8	7.765

Note. ASD, acute stress disorder; SDS, self-rating depression scale; SAS, self-rating anxiety scale; BMI, body mass index; *CI*, confidence interval.