

Online Resource 1: Supplementary Information for “PE-Tab2Conv: Clinically structured convolutional representation learning for preeclampsia risk modelling from routine electronic health records”

Sindhu Rakasiwi^{1,2,3,4*}, Abu Salam^{1,2,3,4*}, Ika Novita Dewi^{1,2,3,4},
Danang Wahyu Utomo^{1,2,3,4}, Dhita Aulia Octaviani⁵,
Sasha-Gay Wright⁶

^{1*}Faculty of Computer Science, Universitas Dian Nuswantoro, Imam
Bonjol Street No. 207, Semarang, Central Java 50131, Indonesia.

²Intelligent Distributed Surveillance and Security (IDSS), Universitas
Dian Nuswantoro, Semarang, Central Java 50131, Indonesia.

³Dinus Research Group for AI in Medical Science (DREAMS),
Universitas Dian Nuswantoro, Semarang, Central Java 50131, Indonesia.

⁴OncoDoc AI Research Laboratory, Universitas Dian Nuswantoro,
Semarang, Central Java 50131, Indonesia.

⁵Department of Midwifery, Poltekkes Kemenkes Semarang, Semarang,
Central Java 50268, Indonesia.

⁶Biomedical Engineering, Faculty of Engineering, The University of the
West Indies, Mona Campus, 4 St. Johns Close, Kingston 7, Jamaica.

*Corresponding author(s). E-mail(s): sindhu.rakasiwi@dsn.dinus.ac.id;
abu.salam@dsn.dinus.ac.id;

Journal: Network Modeling Analysis in Health Informatics and Bioinformatics

Contents: Supplementary Methods; Supplementary Tables S1–S7; Supplementary
Figures S1–S4; Supplementary Notes S1–S2

Supplementary Information

This Supplementary Information provides additional methodological details, experimental registries, robustness summaries, threshold-analysis output registries, prediction-output registries, and supplementary figures supporting the main manuscript. The main manuscript reports the primary temporal validation, layout ablation, cross-site positive-case transportability, and interpretability results. Supplementary Tables S1–S7, Supplementary Figures S1–S4, and Notes S1–S2 correspond to the supplementary items cited in the main text.

Supplementary Methods

S1. Experimental design registry

The experimental workflow was organized to preserve the separation between model development, locked temporal validation, and cross-site positive-case transportability. Table S1 summarizes the main experimental scenarios, their inputs, validation design, and intended methodological purpose.

Table S1: Experimental scenarios and validation roles

Scen. ID	Representation	Feature space	Development input	Validation / assessment input	Purpose
S1	Original tabular baseline	Core	RS Adhyatma 2021–2022 encoded core features	Locked RS Adhyatma 2023 encoded core features	Estimate baseline temporal performance using portable predictors shared across hospitals.
S2	Original tabular baseline	Extended	RS Adhyatma 2021–2022 encoded extended features	Locked RS Adhyatma 2023 encoded extended features	Estimate full-information temporal performance using additional obstetric-history and family-history predictors available within RS Adhyatma.
S3	PE-Tab2Conv end-to-end CNN	Core	RS Adhyatma 2021–2022 core $4 \times 6 \times 1$ matrices	Locked RS Adhyatma 2023 core $4 \times 6 \times 1$ matrices	Evaluate direct CNN-based prediction using the portable core matrix representation.
S4	PE-Tab2Conv end-to-end CNN	Extended	RS Adhyatma 2021–2022 extended $4 \times 6 \times 1$ matrices	Locked RS Adhyatma 2023 extended $4 \times 6 \times 1$ matrices	Evaluate direct CNN-based prediction using the full-information extended matrix representation.

Scen. ID	Representation	Feature space	Development input	Validation / assessment input	Purpose
S5	PE-Tab2Conv embedding + downstream ML	Core	CNN embeddings extracted from RS Adhyatma 2021–2022 core matrices	CNN embeddings extracted from locked RS Adhyatma 2023 core matrices	Evaluate whether compact convolutional embeddings improve or remain competitive with original tabular core baselines.
S6	PE-Tab2Conv embedding + downstream ML	Extended	CNN embeddings extracted from RS Adhyatma 2021–2022 extended matrices	CNN embeddings extracted from locked RS Adhyatma 2023 extended matrices	Evaluate whether compact convolutional embeddings improve or remain competitive with original tabular extended baselines.
S7	Layout ablation	Core and extended	Semantic, random, correlation-based, mutual-information, and information-gain layouts generated using development data only	Locked RS Adhyatma 2023 temporal validation	Test whether semantic clinical feature ordering provides a performance advantage beyond arbitrary or data-driven feature placement.
S8	Cross-site positive-case trans-portability	Core only	Locked portable core model trained on RS Adhyatma development data	RS Adhyatma non-PE, RS Adhyatma PE, and RS Wongsongoro PE-only records	Assess whether the portable core model preserves elevated risk signal in external PE-positive cases without claiming balanced external validation.
S9	Interpretability analysis	Core and extended	Selected original tabular and PE-Tab2Conv models	Feature-level, domain-level, and matrix-level attribution summaries	Assess whether model behavior aligns with clinically meaningful maternal, obstetric, family-history, socioeconomic, and comorbidity domains.

S2. Model family and hyperparameter search registry

Model development and hyperparameter tuning were performed using the RS Adhyatma 2021–2022 development cohort only. The locked RS Adhyatma 2023 temporal validation cohort was not used for hyperparameter tuning, early stopping, threshold selection, feature ranking, layout selection, reference-seed selection, or recalibration. Average precision was used as the internal cross-validation objective because the

endpoint was imbalanced and AUPRC was the primary temporal validation metric. Table S2 reports the prespecified search spaces, CNN training settings, evaluation parameters, and reference-seed selection policy retained by the analysis pipeline.

Table S2: Model families, hyperparameter search spaces, and locked evaluation settings

Component	Parameter(s)	Search space / fixed setting	Role
Logistic regression	Regularization strength; penalty	$C \in \{0.01, 0.1, 1, 10\}$; penalty $\in \{L1, L2\}$	Internal model selection by cross-validated average precision.
Support vector machine	C ; kernel; gamma	$C \in \{0.1, 1, 10\}$; kernel $\in \{\text{linear}, \text{RBF}\}$; gamma = scale	Internal model selection by cross-validated average precision.
Random forest	Number of trees; maximum depth; minimum samples per leaf	$n_{\text{estimators}} \in \{300, 500\}$; max depth $\in \{\text{None}, 3, 5\}$; min samples leaf $\in \{1, 3, 5\}$	Internal model selection by cross-validated average precision.
XGBoost	Number of estimators; maximum depth; learning rate; row and column subsampling	$n_{\text{estimators}} \in \{200, 400\}$; depth $\in \{2, 3\}$; learning rate $\in \{0.03, 0.05, 0.1\}$; subsample and colsample $\in \{0.8, 1.0\}$	Internal model selection by cross-validated average precision.
PE-Tab2Conv CNN	Convolutional filters; kernel; embedding; dropout; learning rate	Filters = 8 and 16; kernel = 2×2 ; embedding = 32; dropout = 0.10; learning rate = 0.001	Fixed lightweight feature-extractor architecture used across core and extended matrices.
CNN training	Batch size; maximum epochs; early stopping; reduce-LR patience	Batch size = 32; maximum epochs = 200; early-stopping patience = 20; reduce-LR patience = 8	Development-cohort training and convergence control.
Stochastic robustness	Training seeds	42, 101, 202, 303, 404, 505, 606, 707, 808, 909	Multi-seed assessment of embedding initialization.
Internal selection	Cross-validation	5-fold stratified cross-validation; scoring = average precision	Applied only within the development cohort.
Reference-seed policy	Core / Extended	Core seed = 909; Extended seed = 505	Seeds selected from internal-validation AUPRC, not from locked temporal-test performance.
Threshold analysis	Decision thresholds	0.50, 0.30, and 0.20	Prespecified operating-point evaluation on the locked temporal cohort.

Component	Parameter(s)	Search space / fixed setting	Role
Bootstrap analysis	Resamples; random state	2,000 resamples; random state = 2026	Confidence-interval estimation on the locked temporal cohort.

S3. PE-Tab2Conv per-seed temporal validation results

PE-Tab2Conv embeddings were evaluated across 10 prespecified random seeds to quantify sensitivity to stochastic CNN initialization. For every seed, the CNN feature extractor was trained using the RS Adhyatma 2021–2022 development cohort, downstream classifiers were selected using development-cohort cross-validation, and final metrics were computed once on the locked RS Adhyatma 2023 temporal validation cohort. The reference seeds used for the principal comparison in the main manuscript were selected from internal-validation AUPRC: seed 909 for the core feature space and seed 505 for the extended feature space. Table S3 reports the complete seed-specific temporal validation results for PE-Tab2Conv embedding models.

Table S3: Seed-specific temporal validation results for PE-Tab2Conv embedding models

Feature set	Model	Seed	AUROC	AUPRC	Brier	F1	Run role
Core	LR	42	0.595	0.198	0.204	0.364	–
Core	SVM	42	0.636	0.234	0.073	0.000	–
Core	RF	42	0.680	0.212	0.164	0.333	–
Core	XGBoost	42	0.621	0.186	0.170	0.263	–
Core	LR	101	0.638	0.243	0.191	0.280	–
Core	SVM	101	0.658	0.241	0.074	0.000	–
Core	RF	101	0.670	0.225	0.162	0.303	–
Core	XGBoost	101	0.587	0.167	0.172	0.261	–
Core	LR	202	0.656	0.278	0.242	0.265	–
Core	SVM	202	0.651	0.263	0.079	0.000	–
Core	RF	202	0.685	0.215	0.162	0.304	–
Core	XGBoost	202	0.601	0.196	0.175	0.261	–
Core	LR	303	0.659	0.208	0.181	0.275	–
Core	SVM	303	0.727	0.235	0.073	0.000	–
Core	RF	303	0.686	0.222	0.159	0.368	–
Core	XGBoost	303	0.641	0.208	0.169	0.286	–
Core	LR	404	0.639	0.250	0.242	0.273	–
Core	SVM	404	0.628	0.241	0.079	0.000	–
Core	RF	404	0.689	0.226	0.164	0.301	–
Core	XGBoost	404	0.678	0.216	0.171	0.274	–
Core	LR	505	0.629	0.254	0.190	0.360	–
Core	SVM	505	0.637	0.277	0.073	0.000	–
Core	RF	505	0.678	0.223	0.162	0.376	–
Core	XGBoost	505	0.621	0.176	0.170	0.266	–
Core	LR	606	0.623	0.229	0.217	0.305	–
Core	SVM	606	0.660	0.225	0.075	0.000	–
Core	RF	606	0.693	0.224	0.163	0.349	–
Core	XGBoost	606	0.626	0.207	0.168	0.277	–
Core	LR	707	0.581	0.205	0.222	0.376	–
Core	SVM	707	0.683	0.256	0.073	0.000	–
Core	RF	707	0.665	0.219	0.160	0.357	–
Core	XGBoost	707	0.610	0.186	0.164	0.277	–
Core	LR	808	0.684	0.237	0.199	0.361	–

Feature set	Model	Seed	AUROC	AUPRC	Brier	F1	Run role
Core	SVM	808	0.676	0.235	0.073	0.000	–
Core	RF	808	0.672	0.214	0.162	0.321	–
Core	XGBoost	808	0.641	0.188	0.169	0.273	–
Core	LR	909	0.643	0.251	0.184	0.273	Reference
Core	SVM	909	0.649	0.256	0.074	0.000	Reference
Core	RF	909	0.663	0.210	0.163	0.329	Reference
Core	XGBoost	909	0.639	0.192	0.169	0.273	Reference
Extended	LR	42	0.915	0.825	0.044	0.760	–
Extended	SVM	42	0.913	0.808	0.033	0.744	–
Extended	RF	42	0.921	0.829	0.039	0.763	–
Extended	XGBoost	42	0.920	0.828	0.038	0.755	–
Extended	LR	101	0.924	0.823	0.060	0.716	–
Extended	SVM	101	0.923	0.803	0.030	0.787	–
Extended	RF	101	0.914	0.820	0.038	0.763	–
Extended	XGBoost	101	0.917	0.832	0.040	0.768	–
Extended	LR	202	0.918	0.775	0.049	0.729	–
Extended	SVM	202	0.918	0.782	0.028	0.824	–
Extended	RF	202	0.917	0.823	0.037	0.771	–
Extended	XGBoost	202	0.922	0.826	0.040	0.760	–
Extended	LR	303	0.924	0.834	0.167	0.731	–
Extended	SVM	303	0.924	0.834	0.027	0.785	–
Extended	RF	303	0.916	0.819	0.037	0.771	–
Extended	XGBoost	303	0.872	0.753	0.046	0.702	–
Extended	LR	404	0.916	0.823	0.048	0.738	–
Extended	SVM	404	0.917	0.823	0.029	0.800	–
Extended	RF	404	0.920	0.822	0.038	0.750	–
Extended	XGBoost	404	0.927	0.827	0.038	0.776	–
Extended	LR	505	0.922	0.817	0.047	0.716	Reference
Extended	SVM	505	0.929	0.806	0.026	0.810	Reference
Extended	RF	505	0.918	0.814	0.036	0.761	Reference
Extended	XGBoost	505	0.886	0.768	0.040	0.753	Reference
Extended	LR	606	0.922	0.789	0.048	0.717	–
Extended	SVM	606	0.923	0.795	0.038	0.545	–
Extended	RF	606	0.917	0.816	0.039	0.758	–
Extended	XGBoost	606	0.921	0.822	0.040	0.760	–
Extended	LR	707	0.924	0.794	0.065	0.716	–
Extended	SVM	707	0.924	0.795	0.027	0.824	–
Extended	RF	707	0.916	0.816	0.037	0.758	–
Extended	XGBoost	707	0.915	0.810	0.040	0.752	–
Extended	LR	808	0.926	0.805	0.047	0.738	–
Extended	SVM	808	0.917	0.797	0.029	0.765	–
Extended	RF	808	0.921	0.804	0.037	0.760	–
Extended	XGBoost	808	0.921	0.813	0.041	0.740	–
Extended	LR	909	0.916	0.802	0.192	0.763	–
Extended	SVM	909	0.916	0.801	0.031	0.737	–
Extended	RF	909	0.920	0.805	0.037	0.776	–
Extended	XGBoost	909	0.909	0.804	0.038	0.771	–

S4. Bootstrap confidence intervals

Bootstrap confidence intervals were calculated using 2,000 resamples of the locked RS Adhyatma 2023 temporal validation cohort. Resampling was performed on the locked prediction outputs without model refitting. Table S4 reports the complete bootstrap

uncertainty registry available for the principal original-tabular and PE-Tab2Conv configurations. These intervals characterize uncertainty within the temporal validation cohort and do not represent external validation intervals.

Table S4: Bootstrap confidence intervals for principal temporal validation models

Representation	Feature set	Model	Seed	Metric	Mean	95% CI
Original tabular	Extended	XGBoost	–	AUROC	0.911	0.840–0.968
Original tabular	Extended	XGBoost	–	AUPRC	0.802	0.682–0.898
Original tabular	Extended	XGBoost	–	Brier score	0.047	0.035–0.059
Original tabular	Extended	XGBoost	–	sensitivity	0.806	0.684–0.913
Original tabular	Extended	XGBoost	–	specificity	0.959	0.940–0.976
Original tabular	Extended	XGBoost	–	PPV	0.651	0.521–0.780
Original tabular	Extended	XGBoost	–	NPV	0.981	0.968–0.992
Original tabular	Extended	XGBoost	–	F1-score	0.718	0.611–0.813
Original tabular	Extended	RF	–	AUROC	0.908	0.837–0.963
Original tabular	Extended	RF	–	AUPRC	0.733	0.580–0.864
Original tabular	Extended	RF	–	Brier score	0.045	0.035–0.057
Original tabular	Extended	RF	–	sensitivity	0.719	0.579–0.839
Original tabular	Extended	RF	–	specificity	0.975	0.961–0.989
Original tabular	Extended	RF	–	PPV	0.735	0.600–0.864
Original tabular	Extended	RF	–	NPV	0.973	0.958–0.988
Original tabular	Extended	RF	–	F1-score	0.725	0.610–0.824
Original tabular	Extended	LR	–	AUROC	0.907	0.833–0.963
Original tabular	Extended	LR	–	AUPRC	0.750	0.594–0.878
Original tabular	Extended	LR	–	Brier score	0.053	0.043–0.064
Original tabular	Extended	LR	–	sensitivity	0.806	0.684–0.913
Original tabular	Extended	LR	–	specificity	0.965	0.947–0.980
Original tabular	Extended	LR	–	PPV	0.687	0.558–0.810
Original tabular	Extended	LR	–	NPV	0.981	0.969–0.992
Original tabular	Extended	LR	–	F1-score	0.740	0.635–0.830
PE-Tab2Conv embedding	Extended	LR	505	AUROC	0.923	0.858–0.974
PE-Tab2Conv embedding	Extended	LR	505	AUPRC	0.820	0.695–0.917
PE-Tab2Conv embedding	Extended	LR	505	Brier score	0.047	0.036–0.059
PE-Tab2Conv embedding	Extended	LR	505	sensitivity	0.848	0.740–0.941
PE-Tab2Conv embedding	Extended	LR	505	specificity	0.951	0.930–0.969
PE-Tab2Conv embedding	Extended	LR	505	PPV	0.620	0.493–0.745
PE-Tab2Conv embedding	Extended	LR	505	NPV	0.985	0.973–0.994
PE-Tab2Conv embedding	Extended	LR	505	F1-score	0.714	0.609–0.808
PE-Tab2Conv embedding	Extended	SVM	505	AUROC	0.929	0.867–0.977
PE-Tab2Conv embedding	Extended	SVM	505	AUPRC	0.810	0.678–0.910
PE-Tab2Conv embedding	Extended	SVM	505	Brier score	0.026	0.016–0.038
PE-Tab2Conv embedding	Extended	SVM	505	sensitivity	0.739	0.605–0.860

Representation	Feature set	Model	Seed	Metric	Mean	95% CI
PE-Tab2Conv embedding	Extended	SVM	505	specificity	0.992	0.983–0.998
PE-Tab2Conv embedding	Extended	SVM	505	PPV	0.895	0.786–0.976
PE-Tab2Conv embedding	Extended	SVM	505	NPV	0.976	0.961–0.988
PE-Tab2Conv embedding	Extended	SVM	505	F1-score	0.808	0.709–0.889
PE-Tab2Conv embedding	Extended	RF	505	AUROC	0.918	0.850–0.973
PE-Tab2Conv embedding	Extended	RF	505	AUPRC	0.816	0.700–0.906
PE-Tab2Conv embedding	Extended	RF	505	Brier score	0.035	0.024–0.048
PE-Tab2Conv embedding	Extended	RF	505	sensitivity	0.762	0.632–0.878
PE-Tab2Conv embedding	Extended	RF	505	specificity	0.978	0.963–0.990
PE-Tab2Conv embedding	Extended	RF	505	PPV	0.763	0.633–0.884
PE-Tab2Conv embedding	Extended	RF	505	NPV	0.977	0.963–0.990
PE-Tab2Conv embedding	Extended	RF	505	F1-score	0.760	0.651–0.851
PE-Tab2Conv embedding	Core	LR	909	AUROC	0.643	0.537–0.746
PE-Tab2Conv embedding	Core	LR	909	AUPRC	0.267	0.148–0.401
PE-Tab2Conv embedding	Core	LR	909	Brier score	0.184	0.172–0.196
PE-Tab2Conv embedding	Core	LR	909	sensitivity	0.478	0.333–0.629
PE-Tab2Conv embedding	Core	LR	909	specificity	0.808	0.773–0.845
PE-Tab2Conv embedding	Core	LR	909	PPV	0.191	0.119–0.267
PE-Tab2Conv embedding	Core	LR	909	NPV	0.942	0.920–0.964
PE-Tab2Conv embedding	Core	LR	909	F1-score	0.272	0.177–0.366
PE-Tab2Conv embedding	Core	SVM	909	AUROC	0.649	0.545–0.751
PE-Tab2Conv embedding	Core	SVM	909	AUPRC	0.272	0.153–0.408
PE-Tab2Conv embedding	Core	SVM	909	Brier score	0.074	0.055–0.094
PE-Tab2Conv embedding	Core	SVM	909	sensitivity	0.000	0.000–0.000
PE-Tab2Conv embedding	Core	SVM	909	specificity	1.000	1.000–1.000
PE-Tab2Conv embedding	Core	SVM	909	PPV	N.A.	N.A.–N.A.
PE-Tab2Conv embedding	Core	SVM	909	NPV	0.913	0.889–0.936
PE-Tab2Conv embedding	Core	SVM	909	F1-score	0.000	0.000–0.000

Representation	Feature set	Model	Seed	Metric	Mean	95% CI
PE-Tab2Conv CNN	Extended	CNN	505	AUROC	0.925	0.861–0.974
PE-Tab2Conv CNN	Extended	CNN	505	AUPRC	0.819	0.695–0.915
PE-Tab2Conv CNN	Extended	CNN	505	Brier score	0.043	0.031–0.055
PE-Tab2Conv CNN	Extended	CNN	505	sensitivity	0.848	0.740–0.941
PE-Tab2Conv CNN	Extended	CNN	505	specificity	0.951	0.930–0.969
PE-Tab2Conv CNN	Extended	CNN	505	PPV	0.620	0.493–0.745
PE-Tab2Conv CNN	Extended	CNN	505	NPV	0.985	0.973–0.994
PE-Tab2Conv CNN	Extended	CNN	505	F1-score	0.714	0.609–0.808

S5. Layout ablation and complete random-layout registry

The layout-ablation experiment evaluated whether the semantic clinical layout provided a performance advantage beyond arbitrary feature placement. Each candidate layout underwent experiment-specific CNN retraining on the development cohort, embedding extraction, downstream Logistic Regression refitting, and locked temporal evaluation. In the core feature space, the semantic layout AUPRC was 0.195, compared with a random-layout mean of 0.221 and 95th percentile of 0.257 ($p_{\text{emp}} = 0.931$). In the extended feature space, the semantic layout AUPRC was 0.765, compared with a random-layout mean of 0.744 and 95th percentile of 0.810 ($p_{\text{emp}} = 0.297$).

The semantic-layout AUPRC of 0.765 is specific to this standardized layout-ablation experiment and is therefore distinct from the principal PE-Tab2Conv embedding + LR result (AUPRC 0.817) reported in the main manuscript. Table S5 provides the complete 200-run random-layout registry used to derive the empirical distribution summaries (100 core and 100 extended layouts).

Table S5: Complete random-layout temporal validation registry for PE-Tab2Conv ablation

Feature set	Random layout ID	CNN seed	AUROC	AUPRC	Brier
Core	random_seed_1000	42	0.621	0.196	0.240
Core	random_seed_1001	42	0.663	0.235	0.205
Core	random_seed_1002	42	0.698	0.209	0.189
Core	random_seed_1003	42	0.644	0.249	0.186
Core	random_seed_1004	42	0.658	0.223	0.207
Core	random_seed_1005	42	0.645	0.212	0.194
Core	random_seed_1006	42	0.624	0.209	0.208
Core	random_seed_1007	42	0.711	0.231	0.206
Core	random_seed_1008	42	0.635	0.224	0.216
Core	random_seed_1009	42	0.700	0.294	0.184
Core	random_seed_1010	42	0.652	0.219	0.204
Core	random_seed_1011	42	0.643	0.222	0.247
Core	random_seed_1012	42	0.618	0.238	0.205
Core	random_seed_1013	42	0.707	0.215	0.212
Core	random_seed_1014	42	0.702	0.238	0.211
Core	random_seed_1015	42	0.587	0.205	0.228
Core	random_seed_1016	42	0.663	0.249	0.210
Core	random_seed_1017	42	0.644	0.209	0.188
Core	random_seed_1018	42	0.700	0.227	0.208
Core	random_seed_1019	42	0.585	0.170	0.250
Core	random_seed_1020	42	0.652	0.205	0.208

Feature set	Random layout ID	CNN seed	AUROC	AUPRC	Brier
Core	random_seed_1021	42	0.676	0.225	0.208
Core	random_seed_1022	42	0.707	0.222	0.210
Core	random_seed_1023	42	0.605	0.211	0.221
Core	random_seed_1024	42	0.675	0.221	0.224
Core	random_seed_1025	42	0.663	0.242	0.181
Core	random_seed_1026	42	0.652	0.196	0.212
Core	random_seed_1027	42	0.631	0.247	0.250
Core	random_seed_1028	42	0.714	0.212	0.227
Core	random_seed_1029	42	0.701	0.245	0.206
Core	random_seed_1030	42	0.664	0.219	0.227
Core	random_seed_1031	42	0.681	0.216	0.207
Core	random_seed_1032	42	0.667	0.227	0.208
Core	random_seed_1033	42	0.639	0.257	0.235
Core	random_seed_1034	42	0.624	0.221	0.221
Core	random_seed_1035	42	0.651	0.222	0.195
Core	random_seed_1036	42	0.681	0.269	0.208
Core	random_seed_1037	42	0.682	0.224	0.182
Core	random_seed_1038	42	0.538	0.169	0.228
Core	random_seed_1039	42	0.643	0.204	0.182
Core	random_seed_1040	42	0.642	0.237	0.183
Core	random_seed_1041	42	0.709	0.232	0.205
Core	random_seed_1042	42	0.687	0.260	0.183
Core	random_seed_1043	42	0.660	0.220	0.183
Core	random_seed_1044	42	0.645	0.195	0.181
Core	random_seed_1045	42	0.698	0.212	0.225
Core	random_seed_1046	42	0.709	0.226	0.218
Core	random_seed_1047	42	0.679	0.219	0.187
Core	random_seed_1048	42	0.655	0.223	0.185
Core	random_seed_1049	42	0.627	0.199	0.209
Core	random_seed_1050	42	0.684	0.224	0.222
Core	random_seed_1051	42	0.711	0.255	0.183
Core	random_seed_1052	42	0.621	0.206	0.188
Core	random_seed_1053	42	0.712	0.240	0.178
Core	random_seed_1054	42	0.661	0.208	0.186
Core	random_seed_1055	42	0.680	0.217	0.180
Core	random_seed_1056	42	0.689	0.232	0.229
Core	random_seed_1057	42	0.664	0.217	0.210
Core	random_seed_1058	42	0.625	0.217	0.207
Core	random_seed_1059	42	0.697	0.222	0.215
Core	random_seed_1060	42	0.597	0.210	0.226
Core	random_seed_1061	42	0.664	0.214	0.180
Core	random_seed_1062	42	0.663	0.223	0.184
Core	random_seed_1063	42	0.723	0.315	0.215
Core	random_seed_1064	42	0.706	0.212	0.207
Core	random_seed_1065	42	0.652	0.213	0.209
Core	random_seed_1066	42	0.477	0.132	0.250
Core	random_seed_1067	42	0.613	0.198	0.237
Core	random_seed_1068	42	0.642	0.219	0.190
Core	random_seed_1069	42	0.648	0.203	0.190
Core	random_seed_1070	42	0.671	0.261	0.197
Core	random_seed_1071	42	0.653	0.209	0.196
Core	random_seed_1072	42	0.628	0.220	0.176
Core	random_seed_1073	42	0.618	0.201	0.250
Core	random_seed_1074	42	0.694	0.254	0.207
Core	random_seed_1075	42	0.633	0.234	0.208
Core	random_seed_1076	42	0.629	0.192	0.199
Core	random_seed_1077	42	0.613	0.189	0.216

Feature set	Random layout ID	CNN seed	AUROC	AUPRC	Brier
Core	random_seed_1078	42	0.651	0.227	0.185
Core	random_seed_1079	42	0.698	0.236	0.210
Core	random_seed_1080	42	0.628	0.199	0.208
Core	random_seed_1081	42	0.637	0.218	0.202
Core	random_seed_1082	42	0.622	0.224	0.217
Core	random_seed_1083	42	0.666	0.239	0.210
Core	random_seed_1084	42	0.678	0.250	0.190
Core	random_seed_1085	42	0.715	0.233	0.197
Core	random_seed_1086	42	0.719	0.231	0.248
Core	random_seed_1087	42	0.660	0.214	0.213
Core	random_seed_1088	42	0.672	0.249	0.208
Core	random_seed_1089	42	0.653	0.230	0.175
Core	random_seed_1090	42	0.663	0.250	0.223
Core	random_seed_1091	42	0.631	0.199	0.197
Core	random_seed_1092	42	0.640	0.210	0.180
Core	random_seed_1093	42	0.652	0.205	0.207
Core	random_seed_1094	42	0.646	0.205	0.184
Core	random_seed_1095	42	0.663	0.245	0.211
Core	random_seed_1096	42	0.621	0.200	0.234
Core	random_seed_1097	42	0.596	0.221	0.220
Core	random_seed_1098	42	0.640	0.201	0.193
Core	random_seed_1099	42	0.600	0.162	0.226
Extended	random_seed_1000	42	0.931	0.749	0.047
Extended	random_seed_1001	42	0.917	0.751	0.047
Extended	random_seed_1002	42	0.897	0.716	0.048
Extended	random_seed_1003	42	0.900	0.774	0.048
Extended	random_seed_1004	42	0.908	0.713	0.052
Extended	random_seed_1005	42	0.924	0.812	0.048
Extended	random_seed_1006	42	0.913	0.721	0.055
Extended	random_seed_1007	42	0.922	0.738	0.069
Extended	random_seed_1008	42	0.880	0.715	0.090
Extended	random_seed_1009	42	0.912	0.737	0.051
Extended	random_seed_1010	42	0.921	0.775	0.048
Extended	random_seed_1011	42	0.916	0.758	0.046
Extended	random_seed_1012	42	0.921	0.733	0.042
Extended	random_seed_1013	42	0.924	0.743	0.048
Extended	random_seed_1014	42	0.917	0.716	0.047
Extended	random_seed_1015	42	0.906	0.801	0.047
Extended	random_seed_1016	42	0.913	0.735	0.053
Extended	random_seed_1017	42	0.918	0.789	0.047
Extended	random_seed_1018	42	0.880	0.679	0.115
Extended	random_seed_1019	42	0.917	0.710	0.046
Extended	random_seed_1020	42	0.922	0.783	0.051
Extended	random_seed_1021	42	0.916	0.747	0.061
Extended	random_seed_1022	42	0.923	0.758	0.047
Extended	random_seed_1023	42	0.908	0.718	0.204
Extended	random_seed_1024	42	0.914	0.789	0.047
Extended	random_seed_1025	42	0.909	0.748	0.060
Extended	random_seed_1026	42	0.917	0.788	0.052
Extended	random_seed_1027	42	0.903	0.663	0.064
Extended	random_seed_1028	42	0.913	0.696	0.175
Extended	random_seed_1029	42	0.925	0.731	0.047
Extended	random_seed_1030	42	0.924	0.839	0.041
Extended	random_seed_1031	42	0.917	0.756	0.046
Extended	random_seed_1032	42	0.912	0.764	0.046
Extended	random_seed_1033	42	0.911	0.750	0.064
Extended	random_seed_1034	42	0.909	0.694	0.054

Feature set	Random layout ID	CNN seed	AUROC	AUPRC	Brier
Extended	random_seed_1035	42	0.906	0.719	0.060
Extended	random_seed_1036	42	0.915	0.755	0.047
Extended	random_seed_1037	42	0.920	0.813	0.041
Extended	random_seed_1038	42	0.919	0.763	0.061
Extended	random_seed_1039	42	0.905	0.679	0.051
Extended	random_seed_1040	42	0.919	0.777	0.064
Extended	random_seed_1041	42	0.909	0.799	0.043
Extended	random_seed_1042	42	0.917	0.738	0.054
Extended	random_seed_1043	42	0.916	0.692	0.052
Extended	random_seed_1044	42	0.890	0.718	0.055
Extended	random_seed_1045	42	0.906	0.731	0.181
Extended	random_seed_1046	42	0.904	0.747	0.062
Extended	random_seed_1047	42	0.915	0.753	0.045
Extended	random_seed_1048	42	0.883	0.704	0.067
Extended	random_seed_1049	42	0.896	0.773	0.066
Extended	random_seed_1050	42	0.905	0.729	0.046
Extended	random_seed_1051	42	0.897	0.718	0.071
Extended	random_seed_1052	42	0.921	0.767	0.062
Extended	random_seed_1053	42	0.920	0.757	0.049
Extended	random_seed_1054	42	0.871	0.703	0.087
Extended	random_seed_1055	42	0.892	0.649	0.118
Extended	random_seed_1056	42	0.913	0.758	0.066
Extended	random_seed_1057	42	0.917	0.802	0.048
Extended	random_seed_1058	42	0.914	0.793	0.044
Extended	random_seed_1059	42	0.913	0.733	0.076
Extended	random_seed_1060	42	0.909	0.738	0.065
Extended	random_seed_1061	42	0.915	0.765	0.079
Extended	random_seed_1062	42	0.906	0.719	0.058
Extended	random_seed_1063	42	0.883	0.731	0.191
Extended	random_seed_1064	42	0.921	0.723	0.064
Extended	random_seed_1065	42	0.921	0.707	0.048
Extended	random_seed_1066	42	0.900	0.728	0.063
Extended	random_seed_1067	42	0.920	0.748	0.052
Extended	random_seed_1068	42	0.911	0.681	0.180
Extended	random_seed_1069	42	0.913	0.813	0.047
Extended	random_seed_1070	42	0.899	0.776	0.067
Extended	random_seed_1071	42	0.919	0.761	0.045
Extended	random_seed_1072	42	0.907	0.714	0.049
Extended	random_seed_1073	42	0.867	0.699	0.085
Extended	random_seed_1074	42	0.921	0.744	0.047
Extended	random_seed_1075	42	0.906	0.761	0.042
Extended	random_seed_1076	42	0.889	0.717	0.074
Extended	random_seed_1077	42	0.920	0.785	0.049
Extended	random_seed_1078	42	0.915	0.752	0.064
Extended	random_seed_1079	42	0.912	0.774	0.045
Extended	random_seed_1080	42	0.905	0.733	0.073
Extended	random_seed_1081	42	0.904	0.667	0.076
Extended	random_seed_1082	42	0.923	0.718	0.052
Extended	random_seed_1083	42	0.911	0.787	0.046
Extended	random_seed_1084	42	0.874	0.702	0.063
Extended	random_seed_1085	42	0.915	0.780	0.046
Extended	random_seed_1086	42	0.915	0.770	0.045
Extended	random_seed_1087	42	0.869	0.717	0.187
Extended	random_seed_1088	42	0.908	0.779	0.042
Extended	random_seed_1089	42	0.914	0.783	0.050
Extended	random_seed_1090	42	0.904	0.733	0.044
Extended	random_seed_1091	42	0.917	0.772	0.061

Feature set	Random layout ID	CNN seed	AUROC	AUPRC	Brier
Extended	random_seed_1092	42	0.909	0.765	0.064
Extended	random_seed_1093	42	0.889	0.714	0.060
Extended	random_seed_1094	42	0.916	0.810	0.045
Extended	random_seed_1095	42	0.897	0.705	0.072
Extended	random_seed_1096	42	0.866	0.652	0.243
Extended	random_seed_1097	42	0.918	0.814	0.049
Extended	random_seed_1098	42	0.903	0.751	0.049
Extended	random_seed_1099	42	0.911	0.792	0.061

S6. Full threshold-based operating characteristics

Threshold-dependent operating characteristics were evaluated at the prespecified cut-offs of 0.50, 0.30, and 0.20 on the locked RS Adhyatma 2023 temporal validation cohort. Table S6 reports the complete threshold registry for original-tabular models, reference-seed PE-Tab2Conv embedding models, and end-to-end CNN models. No threshold was selected using the locked temporal cohort; the three cutoffs were used only to characterize operating behavior.

Table S6: Full threshold-based operating characteristics on the locked temporal validation cohort

Representation	Feature set	Model	Seed	Thresh.	Sens.	Spec.	PPV	NPV	F1	High-risk, n (%)
Original tabular	Core	LR	–	0.50	0.370	0.934	0.347	0.940	0.358	49 (9.2%)
Original tabular	Core	LR	–	0.30	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
Original tabular	Core	LR	–	0.20	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
Original tabular	Core	RF	–	0.50	0.326	0.953	0.395	0.937	0.357	38 (7.2%)
Original tabular	Core	RF	–	0.30	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
Original tabular	Core	RF	–	0.20	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
Original tabular	Core	SVM	–	0.50	0.000	1.000	N.A.	0.913	0.000	0 (0.0%)
Original tabular	Core	SVM	–	0.30	0.022	1.000	1.000	0.915	0.043	1 (0.2%)
Original tabular	Core	SVM	–	0.20	0.304	0.961	0.424	0.936	0.354	33 (6.2%)
Original tabular	Core	XGBoost	–	0.50	0.348	0.946	0.381	0.939	0.364	42 (7.9%)
Original tabular	Core	XGBoost	–	0.30	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
Original tabular	Core	XGBoost	–	0.20	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
PE-Tab2Conv embedding	Core	LR	909	0.50	0.478	0.808	0.191	0.942	0.273	115 (21.7%)
PE-Tab2Conv embedding	Core	LR	909	0.30	0.978	0.074	0.091	0.973	0.167	494 (93.0%)
PE-Tab2Conv embedding	Core	LR	909	0.20	1.000	0.004	0.087	1.000	0.160	529 (99.6%)
PE-Tab2Conv embedding	Core	RF	909	0.50	0.283	0.959	0.394	0.934	0.329	33 (6.2%)
PE-Tab2Conv embedding	Core	RF	909	0.30	0.957	0.041	0.086	0.909	0.159	509 (95.9%)
PE-Tab2Conv embedding	Core	RF	909	0.20	0.978	0.002	0.085	0.500	0.157	529 (99.6%)
PE-Tab2Conv embedding	Core	SVM	909	0.50	0.000	1.000	N.A.	0.913	0.000	0 (0.0%)
PE-Tab2Conv embedding	Core	SVM	909	0.30	0.000	0.998	0.000	0.913	0.000	1 (0.2%)
PE-Tab2Conv embedding	Core	SVM	909	0.20	0.304	0.963	0.438	0.936	0.359	32 (6.0%)
PE-Tab2Conv embedding	Core	XGBoost	909	0.50	0.413	0.847	0.204	0.938	0.273	93 (17.5%)

Representation	Feature set	Model	Seed	Thresh.	Sens.	Spec.	PPV	NPV	F1	High-risk, n (%)
PE-Tab2Conv embedding	Core	XGBoost	909	0.30	0.913	0.068	0.085	0.892	0.156	494 (93.0%)
PE-Tab2Conv embedding	Core	XGBoost	909	0.20	0.935	0.054	0.086	0.897	0.157	502 (94.5%)
PE-Tab2Conv CNN	Core	CNN	909	0.50	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
PE-Tab2Conv CNN	Core	CNN	909	0.30	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
PE-Tab2Conv CNN	Core	CNN	909	0.20	1.000	0.000	0.087	N.A.	0.159	531 (100.0%)
Original tabular	Extended	LR	–	0.50	0.804	0.965	0.685	0.981	0.740	54 (10.2%)
Original tabular	Extended	LR	–	0.30	0.804	0.955	0.627	0.981	0.705	59 (11.1%)
Original tabular	Extended	LR	–	0.20	0.870	0.858	0.367	0.986	0.516	109 (20.5%)
Original tabular	Extended	RF	–	0.50	0.717	0.975	0.733	0.973	0.725	45 (8.5%)
Original tabular	Extended	RF	–	0.30	0.783	0.967	0.692	0.979	0.735	52 (9.8%)
Original tabular	Extended	RF	–	0.20	0.826	0.953	0.623	0.983	0.710	61 (11.5%)
Original tabular	Extended	SVM	–	0.50	0.587	0.996	0.931	0.962	0.720	29 (5.5%)
Original tabular	Extended	SVM	–	0.30	0.739	0.984	0.810	0.975	0.773	42 (7.9%)
Original tabular	Extended	SVM	–	0.20	0.761	0.979	0.778	0.977	0.769	45 (8.5%)
Original tabular	Extended	XGBoost	–	0.50	0.804	0.959	0.649	0.981	0.718	57 (10.7%)
Original tabular	Extended	XGBoost	–	0.30	0.804	0.944	0.578	0.981	0.673	64 (12.1%)
Original tabular	Extended	XGBoost	–	0.20	0.870	0.860	0.370	0.986	0.519	108 (20.3%)
PE-Tab2Conv embedding	Extended	LR	505	0.50	0.848	0.951	0.619	0.985	0.716	63 (11.9%)
PE-Tab2Conv embedding	Extended	LR	505	0.30	0.848	0.951	0.619	0.985	0.716	63 (11.9%)
PE-Tab2Conv embedding	Extended	LR	505	0.20	0.848	0.938	0.565	0.985	0.678	69 (13.0%)
PE-Tab2Conv embedding	Extended	RF	505	0.50	0.761	0.977	0.761	0.977	0.761	46 (8.7%)
PE-Tab2Conv embedding	Extended	RF	505	0.30	0.783	0.967	0.692	0.979	0.735	52 (9.8%)
PE-Tab2Conv embedding	Extended	RF	505	0.20	0.848	0.948	0.609	0.985	0.709	64 (12.1%)
PE-Tab2Conv embedding	Extended	SVM	505	0.50	0.739	0.992	0.895	0.976	0.810	38 (7.2%)
PE-Tab2Conv embedding	Extended	SVM	505	0.30	0.804	0.988	0.860	0.982	0.831	43 (8.1%)

Representation	Feature set	Model	Seed	Thresh.	Sens.	Spec.	PPV	NPV	F1	High-risk, n (%)
PE-Tab2Conv embedding	Extended	SVM	505	0.20	0.826	0.963	0.679	0.983	0.745	56 (10.5%)
PE-Tab2Conv embedding	Extended	XGBoost	505	0.50	0.761	0.975	0.745	0.977	0.753	47 (8.9%)
PE-Tab2Conv embedding	Extended	XGBoost	505	0.30	0.783	0.973	0.735	0.979	0.758	49 (9.2%)
PE-Tab2Conv embedding	Extended	XGBoost	505	0.20	0.783	0.971	0.720	0.979	0.750	50 (9.4%)
PE-Tab2Conv CNN	Extended	CNN	505	0.50	0.848	0.951	0.619	0.985	0.716	63 (11.9%)
PE-Tab2Conv CNN	Extended	CNN	505	0.30	0.848	0.951	0.619	0.985	0.716	63 (11.9%)
PE-Tab2Conv CNN	Extended	CNN	505	0.20	0.848	0.942	0.582	0.985	0.690	67 (12.6%)

S7. Full cross-site prediction and transportability summary

Cross-site analysis was restricted to the portable core feature space because prior preeclampsia history and family history of hypertension or preeclampsia were structurally unavailable in RS Wongsonegoro. Table [S7](#) reports the complete cohort-level prediction summary for the portable core models applied to RS Adhyatma non-PE records, RS Adhyatma PE records, and the RS Wongsonegoro PE-only cohort. Because the external site contained no non-PE controls, these results quantify positive-case transportability and cannot be used to estimate external specificity, AUROC, calibration, or false-positive burden.

Table S7: Full cross-site prediction and positive-case transportability summary for portable core models

Representation	Model	Cohort group	n	Mean	Median	Q1–Q3	High-risk 0.50	High-risk 0.30	High-risk 0.20
Original tabular	LR	RS Adhyatma Non-PE	1558	0.402	0.356	0.3514–0.3564	115 (7.4%)	1558 (100.0%)	1558 (100.0%)
Original tabular	LR	RS Adhyatma PE	113	0.583	0.493	0.3514–0.8493	48 (42.5%)	113 (100.0%)	113 (100.0%)
Original tabular	LR	RS Wongsonegoro PE	72	0.812	0.936	0.5570–0.9908	63 (87.5%)	72 (100.0%)	72 (100.0%)
Original tabular	RF	RS Adhyatma Non-PE	1558	0.412	0.393	0.3863–0.3928	90 (5.8%)	1558 (100.0%)	1558 (100.0%)
Original tabular	RF	RS Adhyatma PE	113	0.532	0.437	0.3863–0.7534	44 (38.9%)	113 (100.0%)	113 (100.0%)
Original tabular	RF	RS Wongsonegoro PE	72	0.540	0.605	0.3928–0.6637	43 (59.7%)	71 (98.6%)	72 (100.0%)
Original tabular	SVM	RS Adhyatma Non-PE	1558	0.053	0.038	0.0382–0.0388	0 (0.0%)	1 (0.1%)	84 (5.4%)
Original tabular	SVM	RS Adhyatma PE	113	0.126	0.053	0.0382–0.2844	0 (0.0%)	1 (0.9%)	39 (34.5%)
Original tabular	SVM	RS Wongsonegoro PE	72	0.147	0.096	0.0587–0.2202	0 (0.0%)	13 (18.1%)	19 (26.4%)
Original tabular	XGBoost	RS Adhyatma Non-PE	1558	0.408	0.366	0.3657–0.3825	98 (6.3%)	1558 (100.0%)	1558 (100.0%)
Original tabular	XGBoost	RS Adhyatma PE	113	0.557	0.481	0.3659–0.8431	46 (40.7%)	113 (100.0%)	113 (100.0%)
Original tabular	XGBoost	RS Wongsonegoro PE	72	0.691	0.823	0.4308–0.8526	50 (69.4%)	72 (100.0%)	72 (100.0%)
PE-Tab2Conv embedding	LR	RS Adhyatma Non-PE	1558	0.408	0.376	0.3522–0.3949	297 (19.1%)	1458 (93.6%)	1556 (99.9%)
PE-Tab2Conv embedding	LR	RS Adhyatma PE	113	0.572	0.503	0.3522–0.8516	58 (51.3%)	110 (97.3%)	113 (100.0%)
PE-Tab2Conv embedding	LR	RS Wongsonegoro PE	72	0.675	0.796	0.4437–0.8666	52 (72.2%)	68 (94.4%)	72 (100.0%)
PE-Tab2Conv embedding	RF	RS Adhyatma Non-PE	1558	0.378	0.338	0.3250–0.3748	80 (5.1%)	1508 (96.8%)	1552 (99.6%)
PE-Tab2Conv embedding	RF	RS Adhyatma PE	113	0.540	0.486	0.3381–0.7595	43 (38.1%)	111 (98.2%)	112 (99.1%)

Representation	Model	Cohort group	n	Mean	Median	Q1–Q3	High-risk 0.50	High-risk 0.30	High-risk 0.20
PE-Tab2Conv embedding	RF	RS Wongsonegoro PE	72	0.546	0.560	0.3747–0.6255	42 (58.3%)	72 (100.0%)	72 (100.0%)
PE-Tab2Conv embedding	SVM	RS Adhyatma Non-PE	1558	0.050	0.038	0.0382–0.0393	0 (0.0%)	1 (0.1%)	82 (5.3%)
PE-Tab2Conv embedding	SVM	RS Adhyatma PE	113	0.106	0.047	0.0382–0.2076	0 (0.0%)	1 (0.9%)	42 (37.2%)
PE-Tab2Conv embedding	SVM	RS Wongsonegoro PE	72	0.138	0.149	0.0468–0.2110	0 (0.0%)	2 (2.8%)	27 (37.5%)
PE-Tab2Conv embedding	XGBoost	RS Adhyatma Non-PE	1558	0.379	0.343	0.3348–0.3632	248 (15.9%)	1473 (94.5%)	1487 (95.4%)
PE-Tab2Conv embedding	XGBoost	RS Adhyatma PE	113	0.565	0.519	0.3433–0.8772	58 (51.3%)	109 (96.5%)	110 (97.3%)
PE-Tab2Conv embedding	XGBoost	RS Wongsonegoro PE	72	0.505	0.408	0.3433–0.7316	26 (36.1%)	58 (80.6%)	69 (95.8%)
PE-Tab2Conv CNN	CNN	RS Adhyatma Non-PE	1558	0.530	0.535	0.5155–0.5346	1558 (100.0%)	1558 (100.0%)	1558 (100.0%)
PE-Tab2Conv CNN	CNN	RS Adhyatma PE	113	0.543	0.536	0.5215–0.5647	113 (100.0%)	113 (100.0%)	113 (100.0%)
PE-Tab2Conv CNN	CNN	RS Wongsonegoro PE	72	0.548	0.549	0.5338–0.5614	72 (100.0%)	72 (100.0%)	72 (100.0%)

Supplementary Figures

Supplementary Figure S1. Example PE-Tab2Conv matrix visual check

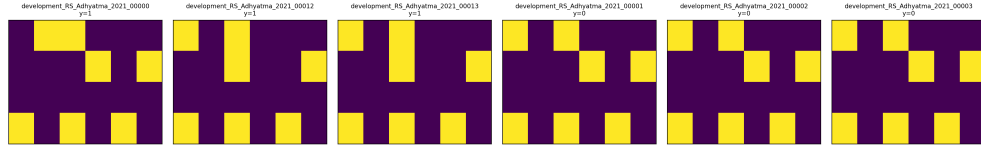


Fig. S1 Visual check of sample PE-Tab2Conv matrices. Example matrices were selected from the development cohort to verify tensor construction, matrix orientation, and clinically guided feature placement before CNN training

Supplementary Figure S2. Training history for PE-Tab2Conv CNN

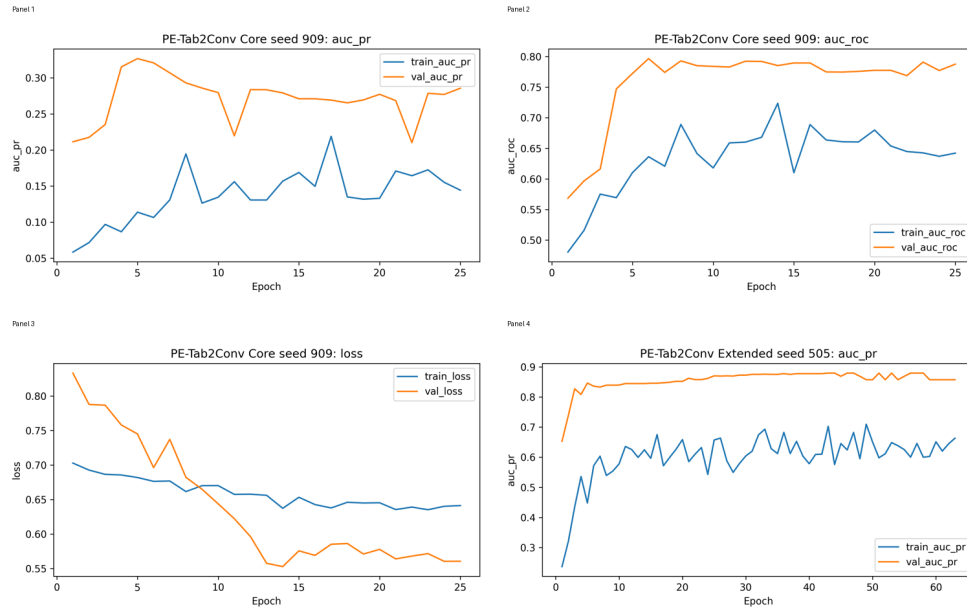


Fig. S2 Representative training history for the PE-Tab2Conv CNN feature extractor. Training and internal validation curves were inspected to monitor convergence and overfitting during development-cohort training

Supplementary Figure S3. High-risk proportion by group and threshold for original-tabular logistic regression

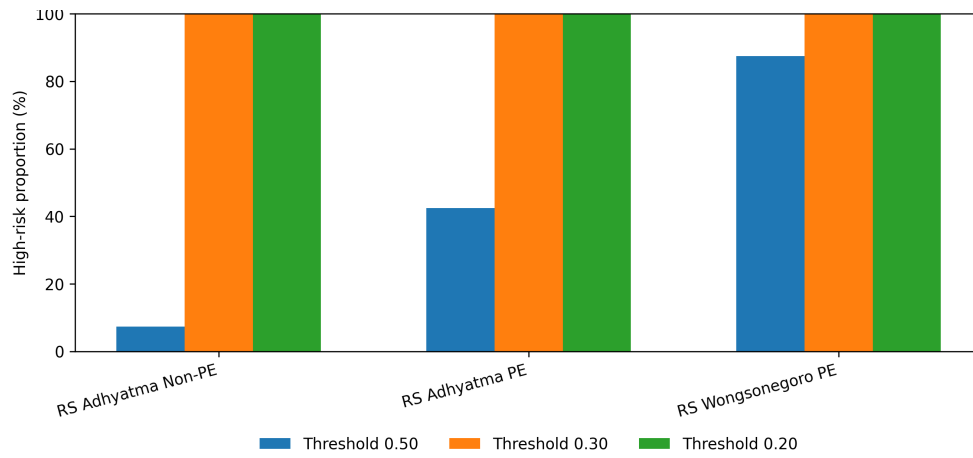
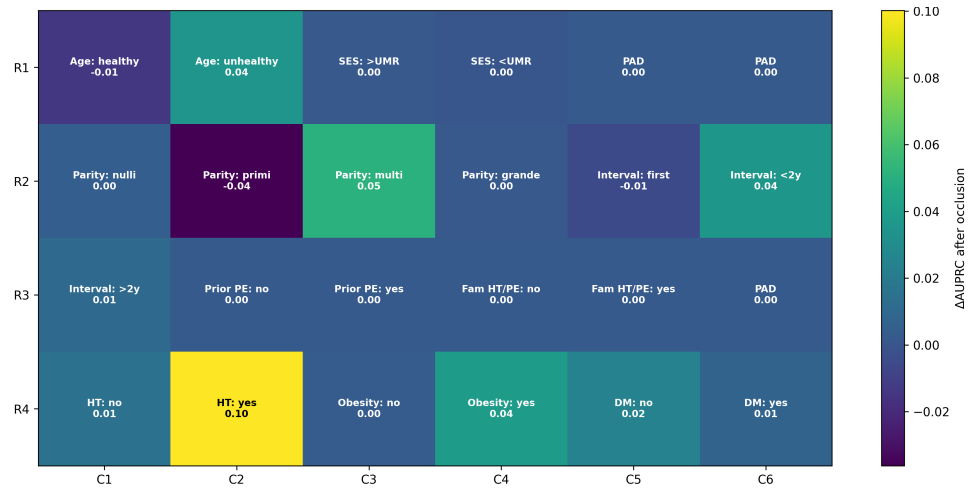
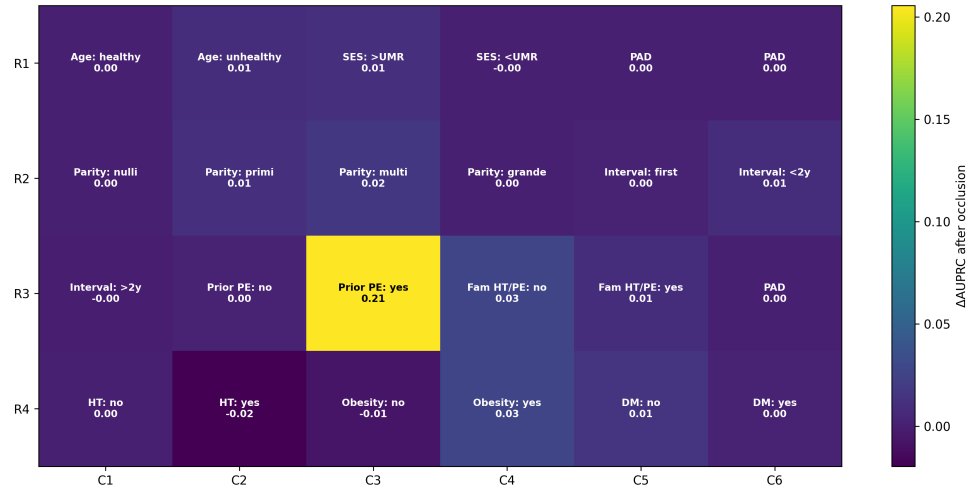


Fig. S3 High-risk proportion by cohort group and decision threshold for the portable core original-tabular logistic regression model in the cross-site positive-case transportability analysis. Thresholds of 0.50, 0.30, and 0.20 were evaluated for RS Adhyatma non-PE records, RS Adhyatma PE records, and RS Wongsonegoro PE records

Supplementary Figure S4. Matrix occlusion heatmaps



(a) Core PE-Tab2Conv representation



(b) Extended PE-Tab2Conv representation

Fig. S4 Matrix occlusion heatmaps for the core and extended PE-Tab2Conv representations. Cell-level occlusion was used to summarize the change in AUPRC after masking each matrix position

Supplementary Notes

Note S1. Interpretation of positive-case transportability

The RS Wongsonegoro cohort contained PE cases only and no non-PE controls. Therefore, the cross-site analysis was not designed to estimate external AUROC, specificity, calibration, decision-curve behavior, or false-positive burden. Instead, it was interpreted as a positive-case transportability assessment: the locked portable core model was applied unchanged to an external PE-only cohort to determine whether predicted risk remained elevated relative to internal non-PE records. This distinction is important because balanced external validation requires both positive and negative outcome groups from the external site.

Note S2. Interpretation of semantic layout ablation

The semantic PE-Tab2Conv layout preserved clinical interpretability by organizing matrix regions into maternal baseline, socioeconomic, obstetric-history, family-history, and comorbidity domains. However, the layout-ablation analysis showed that the semantic layout was not statistically superior to random layouts. The semantic layout should therefore be interpreted as clinically meaningful and competitive, but not as a proven optimal spatial arrangement.

The semantic-layout AUPRC of 0.765 represents the standardized layout-ablation experiment, in which each candidate layout underwent experiment-specific CNN retraining, embedding extraction, and downstream Logistic Regression refitting. It is therefore distinct from the principal PE-Tab2Conv embedding + LR result (AUPRC 0.817) reported in the main manuscript. Overall, the findings indicate that performance gains in the extended setting are more plausibly associated with the combined contribution of richer clinical predictor availability and convolutional representation learning than with semantic spatial ordering alone.