

Supplementary Material

Timing-Aware Inpatient Drug-Drug Interaction Detection and Interpretable Prioritization for Pharmacist Review at Discharge: A MIMIC-IV Study

This supplement contains the artifacts needed to re-derive the detection layer and the complete metric tables summarized in the main text. Machine-readable versions of every table accompany this document as CSV files, and the detection logic is supplied as a runnable SQL file (S3_detection_logic.sql). None of these files contains patient-level data, so none is restricted by the MIMIC-IV data-use agreement.

Contents. S1 medication mapping table. S2 nine-rule specification. S3 detection and feature-aggregation SQL. S4 full per-model metrics. S5 bootstrap confidence intervals. S6 age 65 and older sensitivity analysis. S7 exploratory 30-day readmission analysis. S8 medication-count operating rules. S9 overlap-threshold sensitivity. S10 full decision-threshold grid. S11 calibration bins. S12 full coefficient table. S13 persistence of flagged combinations to discharge. S14 effect of the co-activity requirement against a co-listing comparator.

Supplementary Table S1. The exact-match medication mapping table (39 entries, 29 generic names, 10 classes). A medication order maps only when its lowercased source string equals a source name exactly, which is why mapped medication burden runs below the full medication count. Shown in two column groups to save space; the machine-readable file is S1_mapping_table.csv.

Source name	Generic	Class	Source name	Generic	Class
warfarin	warfarin	warfarin	fentanyl	fentanyl	opioid
coumadin	warfarin	warfarin	oxycodone	oxycodone	opioid
apixaban	apixaban	doac	sertraline	sertraline	ssri
eliquis	apixaban	doac	zoloft	sertraline	ssri
rivaroxaban	rivaroxaban	doac	fluoxetine	fluoxetine	ssri
xarelto	rivaroxaban	doac	prozac	fluoxetine	ssri
dabigatran	dabigatran	doac	citalopram	citalopram	ssri
pradaxa	dabigatran	doac	escitalopram	escitalopram	ssri
edoxaban	edoxaban	doac	paroxetine	paroxetine	ssri
savaysa	edoxaban	doac	amiodarone	amiodarone	antiarrhythmic
aspirin	aspirin	antiplatelet	trimethoprim-sulfamethoxazole	trimethoprim-sulfamethoxazole	antibiotic
ibuprofen	ibuprofen	nsaid	tmp-smx	trimethoprim-sulfamethoxazole	antibiotic
naproxen	naproxen	nsaid	bactrim	trimethoprim-sulfamethoxazole	antibiotic
ketorolac	ketorolac	nsaid	sepra	trimethoprim-sulfamethoxazole	antibiotic
lorazepam	lorazepam	benzodiazepine	diphenhydramine	diphenhydramine	anticholinergic
diazepam	diazepam	benzodiazepine	hydroxyzine	hydroxyzine	anticholinergic
alprazolam	alprazolam	benzodiazepine	meclizine	meclizine	anticholinergic
clonazepam	clonazepam	benzodiazepine	dicyclomine	dicyclomine	anticholinergic
morphine	morphine	opioid	hyoscyamine	hyoscyamine	anticholinergic
hydromorphone	hydromorphone	opioid			

Supplementary Table S2. The nine-rule specification, given verbatim as implemented. The rule type states whether each side matches a generic name or a therapeutic class. Seven rules cite one or more specific tables of the 2023 AGS Beers Criteria. R006 and R008 rest on clinical consensus, so neither corresponds to a numbered Beers table entry. The machine-readable file is S2_ruleset.csv and includes the rationale text.

Rule	Type	Left	Right	Min age	Severity	Evidence
R001	generic+class	warfarin	nsaid	none	major	AGS Beers 2023 Table 2
R002	generic+generic	warfarin	aspirin	none	major	AGS Beers 2023 Table 2
R003	generic+class	warfarin	ssri	none	major	AGS Beers 2023 Table 5
R004	generic+generic	warfarin	amiodarone	none	major	AGS Beers 2023 Table 5
R005	generic+generic	warfarin	trimethoprim-sulfamethoxazole	none	major	AGS Beers 2023 Table 5
R006	class+generic	doac	aspirin	none	moderate	Clinical consensus on combined antithrombotic therapy
R007	class+class	opioid	benzodiazepine	none	major	AGS Beers 2023 Table 5
R008	class+class	nsaid	nsaid	none	minor	Clinical consensus on duplicate therapy
R009	class+class	anticholinergic	anticholinergic	65	moderate	AGS Beers 2023 Tables 2,3,5

Supplementary Table S4. Full per-model metrics on the patient-disjoint held-out test set (71,935 admissions, 6,220 with a major rule match) for every feature set, model, and operating point. This is the complete version of main-text Table 3. The machine-readable file S4_full_model_metrics.csv also carries accuracy, negative predictive value, and the full confusion matrix for each row.

Feature set	Model	Operating point	Thr.	ROC AUC	AP	Prec.	Recall	F1	Spec.
core_3	logistic regression	default 0.5	0.500	0.860	0.331	0.248	0.768	0.375	0.780
core_3	logistic regression	train f1 optimal	0.564	0.860	0.331	0.265	0.686	0.382	0.820
core_3	logistic regression	train recall 0.80	0.474	0.860	0.331	0.242	0.799	0.371	0.763
core_3	random forest	default 0.5	0.500	0.850	0.298	0.254	0.648	0.364	0.819
core_3	random forest	train f1 optimal	0.670	0.850	0.298	0.298	0.428	0.351	0.904
core_3	random forest	train recall 0.80	0.592	0.850	0.298	0.273	0.538	0.362	0.865
core_3	xgboost	default 0.5	0.500	0.871	0.351	0.215	0.896	0.347	0.690
core_3	xgboost	train f1 optimal	0.738	0.871	0.351	0.316	0.557	0.403	0.886
core_3	xgboost	train recall 0.80	0.625	0.871	0.351	0.251	0.793	0.381	0.775
expanded	logistic regression	default 0.5	0.500	0.872	0.360	0.267	0.781	0.398	0.797
expanded	logistic regression	train f1 optimal	0.671	0.872	0.360	0.316	0.600	0.414	0.877
expanded	logistic regression	train recall 0.80	0.487	0.872	0.360	0.264	0.797	0.397	0.790
expanded	random forest	default 0.5	0.500	0.886	0.382	0.320	0.662	0.431	0.867
expanded	random forest	train f1 optimal	0.656	0.886	0.382	0.410	0.409	0.410	0.944
expanded	random forest	train recall 0.80	0.633	0.886	0.382	0.391	0.445	0.416	0.934
expanded	xgboost	default 0.5	0.500	0.888	0.395	0.237	0.898	0.375	0.726
expanded	xgboost	train f1 optimal	0.753	0.888	0.395	0.354	0.567	0.436	0.902
expanded	xgboost	train recall 0.80	0.619	0.888	0.395	0.280	0.791	0.414	0.808
n_drugs_only	logistic regression	default 0.5	0.500	0.829	0.256	0.251	0.609	0.355	0.828
n_drugs_only	logistic regression	train f1 optimal	0.738	0.829	0.256	0.251	0.609	0.355	0.828
n_drugs_only	logistic regression	train recall 0.80	0.407	0.829	0.256	0.164	1.000	0.281	0.517
n_drugs_only	random forest	default 0.5	0.500	0.829	0.256	0.164	1.000	0.281	0.517
n_drugs_only	random forest	train f1 optimal	0.734	0.829	0.256	0.251	0.609	0.355	0.828
n_drugs_only	random forest	train recall 0.80	0.560	0.829	0.256	0.164	1.000	0.281	0.517
n_drugs_only	xgboost	default 0.5	0.500	0.829	0.256	0.164	1.000	0.281	0.517
n_drugs_only	xgboost	train f1 optimal	0.734	0.829	0.256	0.251	0.609	0.355	0.828
n_drugs_only	xgboost	train recall 0.80	0.560	0.829	0.256	0.164	1.000	0.281	0.517
none	always negative	predict all negative			0.086	n/a	0.000	0.000	1.000

Supplementary Table S5. Bootstrap 95 percent confidence intervals from 1000 stratified resamples of the held-out test set. ROC AUC and average precision are given for every feature set and model. The threshold-based metrics at 0.5 are given for the expanded logistic model, with recall also given for expanded XGBoost. The final rows give the paired expanded versus n_drugs_only differences.

Model or comparison	Metric	Estimate	CI lower	CI upper
n_drugs_only, logistic regression	roc auc	0.829	0.825	0.833
n_drugs_only, logistic regression	avg precision	0.256	0.249	0.263
n_drugs_only, random forest	roc auc	0.829	0.825	0.833
n_drugs_only, random forest	avg precision	0.256	0.249	0.263
n_drugs_only, xgboost	roc auc	0.829	0.825	0.833
n_drugs_only, xgboost	avg precision	0.256	0.249	0.263
core_3, logistic regression	roc auc	0.860	0.857	0.864
core_3, logistic regression	avg precision	0.331	0.322	0.342
core_3, random forest	roc auc	0.850	0.847	0.854
core_3, random forest	avg precision	0.298	0.290	0.307
core_3, xgboost	roc auc	0.871	0.868	0.875
core_3, xgboost	avg precision	0.351	0.341	0.361
expanded, logistic regression	roc auc	0.872	0.869	0.876
expanded, logistic regression	avg precision	0.360	0.350	0.371
expanded, logistic regression	precision@0.5	0.267	0.263	0.271
expanded, logistic regression	recall@0.5	0.781	0.771	0.791
expanded, logistic regression	f1@0.5	0.398	0.392	0.403
expanded, logistic regression	specificity@0.5	0.797	0.794	0.800
expanded, random forest	roc auc	0.886	0.882	0.889
expanded, random forest	avg precision	0.382	0.372	0.393
expanded, xgboost	roc auc	0.888	0.885	0.891
expanded, xgboost	avg precision	0.395	0.384	0.406
expanded, xgboost	recall@0.5	0.898	0.891	0.905
expanded_vs_n_drugs_only, logistic regression	delta roc auc	0.043	0.041	0.046
expanded_vs_n_drugs_only, logistic regression	delta avg precision	0.104	0.099	0.109

Supplementary Table S6. Sensitivity analysis restricted to admissions with age at least 65 (cohort n = 145,533; cohort major-match rate 12.3 percent). Every row below is computed on the patient-disjoint held-out split of that subgroup. Bootstrap confidence intervals for this subgroup are in S6b_age65_bootstrap_ci.csv.

Feature set	Model	Operating point	ROC AUC	AP	Prec.	Recall	F1
core_3	logistic regression	default 0.5	0.855	0.387	0.318	0.691	0.435
core_3	logistic regression	train f1 optimal	0.855	0.387	0.308	0.733	0.434
core_3	logistic regression	train recall 0.80	0.855	0.387	0.290	0.810	0.427
core_3	random forest	default 0.5	0.840	0.343	0.293	0.702	0.413
core_3	random forest	train f1 optimal	0.840	0.343	0.335	0.497	0.400
core_3	random forest	train recall 0.80	0.840	0.343	0.315	0.580	0.408
core_3	xgboost	default 0.5	0.863	0.404	0.254	0.960	0.402
core_3	xgboost	train f1 optimal	0.863	0.404	0.331	0.707	0.451
core_3	xgboost	train recall 0.80	0.863	0.404	0.305	0.802	0.442
expanded	logistic regression	default 0.5	0.869	0.420	0.332	0.769	0.464
expanded	logistic regression	train f1 optimal	0.869	0.420	0.331	0.770	0.463
expanded	logistic regression	train recall 0.80	0.869	0.420	0.323	0.805	0.461
expanded	random forest	default 0.5	0.875	0.433	0.352	0.723	0.474
expanded	random forest	train f1 optimal	0.875	0.433	0.437	0.478	0.457
expanded	random forest	train recall 0.80	0.875	0.433	0.433	0.484	0.457
expanded	xgboost	default 0.5	0.882	0.454	0.288	0.914	0.437
expanded	xgboost	train f1 optimal	0.882	0.454	0.372	0.710	0.488
expanded	xgboost	train recall 0.80	0.882	0.454	0.339	0.793	0.475
n_drugs_only	logistic regression	default 0.5	0.839	0.332	0.347	0.557	0.427
n_drugs_only	logistic regression	train f1 optimal	0.839	0.332	0.347	0.557	0.427
n_drugs_only	logistic regression	train recall 0.80	0.839	0.332	0.237	1.000	0.383
n_drugs_only	random forest	default 0.5	0.839	0.332	0.237	1.000	0.383
n_drugs_only	random forest	train f1 optimal	0.839	0.332	0.347	0.557	0.427
n_drugs_only	random forest	train recall 0.80	0.839	0.332	0.237	1.000	0.383
n_drugs_only	xgboost	default 0.5	0.839	0.332	0.237	1.000	0.383
n_drugs_only	xgboost	train f1 optimal	0.839	0.332	0.347	0.557	0.427
n_drugs_only	xgboost	train recall 0.80	0.839	0.332	0.237	1.000	0.383
none	always negative	predict all negative		0.119	n/a	0.000	0.000

Supplementary Table S7. Exploratory analysis of all-cause 30-day readmission (cohort readmission rate 20.3 percent, and 20.4 percent in the held-out test split, which is the value the always-negative average precision reproduces). The DDI summary features are permitted for this target because they are known by discharge. Discrimination is modest and the DDI features add a ROC AUC gain of 0.002 computed before rounding (0.6143 against 0.6127), so this endpoint is reported as exploratory only. Coefficients are in S7b_readmission_coefficients.csv.

Feature set	Model	Operating point	ROC AUC	AP	Prec.	Recall	F1
core_3	logistic regression	default 0.5	0.569	0.245	0.251	0.382	0.303
core_3	logistic regression	train f1 optimal	0.569	0.245	0.217	0.865	0.347
core_3	logistic regression	train recall 0.80	0.569	0.245	0.221	0.801	0.347
core_3	random forest	default 0.5	0.533	0.225	0.227	0.354	0.277
core_3	random forest	train f1 optimal	0.533	0.225	0.227	0.363	0.279
core_3	random forest	train recall 0.80	0.533	0.225	0.220	0.508	0.307
core_3	xgboost	default 0.5	0.591	0.258	0.253	0.550	0.346
core_3	xgboost	train f1 optimal	0.591	0.258	0.236	0.699	0.353
core_3	xgboost	train recall 0.80	0.591	0.258	0.226	0.800	0.352
expanded_no_ddi	logistic regression	default 0.5	0.613	0.267	0.267	0.587	0.367
expanded_no_ddi	logistic regression	train f1 optimal	0.613	0.267	0.256	0.680	0.372
expanded_no_ddi	logistic regression	train recall 0.80	0.613	0.267	0.239	0.798	0.368
expanded_no_ddi	random forest	default 0.5	0.626	0.297	0.320	0.291	0.305
expanded_no_ddi	random forest	train f1 optimal	0.626	0.297	0.324	0.268	0.293
expanded_no_ddi	random forest	train recall 0.80	0.626	0.297	0.319	0.297	0.308
expanded_no_ddi	xgboost	default 0.5	0.638	0.293	0.273	0.634	0.381
expanded_no_ddi	xgboost	train f1 optimal	0.638	0.293	0.267	0.665	0.381
expanded_no_ddi	xgboost	train recall 0.80	0.638	0.293	0.247	0.799	0.378
expanded_with_ddi	logistic regression	default 0.5	0.614	0.267	0.268	0.593	0.369
expanded_with_ddi	logistic regression	train f1 optimal	0.614	0.267	0.248	0.734	0.371
expanded_with_ddi	logistic regression	train recall 0.80	0.614	0.267	0.240	0.801	0.369
expanded_with_ddi	random forest	default 0.5	0.630	0.300	0.316	0.303	0.309
expanded_with_ddi	random forest	train f1 optimal	0.630	0.300	0.327	0.254	0.286
expanded_with_ddi	random forest	train recall 0.80	0.630	0.300	0.315	0.313	0.314
expanded_with_ddi	xgboost	default 0.5	0.639	0.294	0.272	0.633	0.381
expanded_with_ddi	xgboost	train f1 optimal	0.639	0.294	0.262	0.700	0.381
expanded_with_ddi	xgboost	train recall 0.80	0.639	0.294	0.248	0.798	0.378
none	always negative	predict all negative		0.204	n/a	0.000	0.000

Supplementary Table S8. Ranking the worklist by mapped medication count alone, evaluated as an operating rule on the held-out test set. Because `n_drugs` is a small integer, a one-feature logistic model on it reduces exactly to these thresholds, so these rows are the complete set of operating points available to a medication-count sort. Main-text Table 5 compares them against the expanded logistic model at matched recall.

Operating rule	Alerts	Cohort flagged (%)	True positives	Recall	Prec.	Alerts per true positive
<code>n_drugs >= 2</code>	37,980	52.8	6,220	1.000	0.164	6.11
<code>n_drugs >= 3</code>	15,103	21.0	3,789	0.609	0.251	3.99
<code>n_drugs >= 4</code>	5,042	7.0	1,728	0.278	0.343	2.92
<code>n_drugs >= 5</code>	1,489	2.1	661	0.106	0.444	2.25
<code>n_drugs >= 6</code>	355	0.5	185	0.030	0.521	1.92
<code>n_drugs >= 7</code>	91	0.1	54	0.009	0.593	1.69
<code>n_drugs >= 8</code>	18	0.0	10	0.002	0.556	1.80
<code>n_drugs >= 9</code>	4	0.0	1	0.000	0.250	4.00

Supplementary Table S9. Sensitivity of the detection counts to the co-activity threshold. Each row re-runs the detection layer at a different minimum administration-window overlap, counting co-active order-line pairs. The two-hour row reproduces the counts reported in the main text. A zero-hour threshold still requires a strictly positive overlap, so it remains stricter than the co-listing comparator in Supplementary Table S14. The final column gives the change in the number of admissions carrying a major match relative to the two-hour setting.

Overlap thresh- old (h)	Major pairs	Moderate pairs	Minor pairs	Admissions major	with Change vs 2 h (%)
0	189,797	10,302	16,849	34,617	+11.1
1	145,955	10,302	16,849	31,477	+1.0
2	141,189	10,016	16,575	31,169	+0.0
4	133,252	9,593	16,039	30,514	-2.1
8	116,519	8,947	15,162	28,584	-8.3

Supplementary Table S10. The full decision-threshold grid for the expanded logistic model on the held-out test set, including the three named operating points. This is the grid exported to drive the dashboard threshold slider.

Threshold	Prec.	Recall	Spec.	Alerts	Alerts per true positive	Cohort (%)	flagged
0.500	0.267	0.781	0.797	18,208	3.75	25.3	
0.671	0.316	0.600	0.877	11,786	3.16	16.4	
0.487	0.264	0.797	0.790	18,764	3.78	26.1	
0.050	0.096	1.000	0.113	64,506	10.37	89.7	
0.100	0.116	0.999	0.278	53,671	8.64	74.6	
0.150	0.139	0.994	0.418	44,439	7.19	61.8	
0.200	0.159	0.982	0.508	38,425	6.29	53.4	
0.250	0.177	0.963	0.575	33,903	5.66	47.1	
0.300	0.195	0.938	0.634	29,859	5.12	41.5	
0.350	0.214	0.904	0.686	26,253	4.67	36.5	
0.400	0.236	0.872	0.732	23,017	4.24	32.0	
0.450	0.253	0.830	0.768	20,393	3.95	28.3	
0.500	0.267	0.781	0.797	18,208	3.75	25.3	
0.550	0.281	0.738	0.821	16,336	3.56	22.7	
0.600	0.294	0.687	0.844	14,551	3.41	20.2	
0.650	0.310	0.629	0.867	12,635	3.23	17.6	
0.700	0.325	0.561	0.890	10,729	3.07	14.9	
0.750	0.346	0.488	0.913	8,768	2.89	12.2	
0.800	0.366	0.421	0.931	7,158	2.73	10.0	
0.850	0.391	0.359	0.947	5,710	2.55	7.9	
0.900	0.431	0.277	0.965	4,000	2.32	5.6	
0.950	0.485	0.167	0.983	2,140	2.06	3.0	

Supplementary Table S11. Reliability-curve bins for the expanded logistic model, the numeric form of main-text Figure 2B. Observed frequency sits below mean predicted probability in every bin, which is the upward shift produced by class weighting. The Brier score is 0.151.

Bin	Mean predicted probability	Observed frequency
0	0.033	0.000
1	0.066	0.000
2	0.099	0.002
3	0.136	0.005
4	0.190	0.016
5	0.269	0.040
6	0.368	0.066
7	0.509	0.147
8	0.698	0.225
9	0.908	0.365

Supplementary Table S12. All fourteen expanded logistic-regression coefficients and odds ratios to six decimal places, including the two missingness indicators. Features were not standardized, so each odds ratio is per one unit of that feature in its native units and coefficient magnitudes are not comparable across features.

Feature	Coefficient	Odds ratio
n_drugs	1.519371	4.569349
has_heart_failure	0.882444	2.416799
missingindicator_albumin_median	0.376502	1.457178
has_ckd	0.279891	1.322985
has_diabetes	0.087667	1.091624
creatinine_median	0.062099	1.064067
age	0.030269	1.030732
n_diagnoses	-0.000008	0.999992
los_hours	-0.000311	0.999689
albumin_median	-0.041242	0.959597
has_liver_disease	-0.148532	0.861972
has_copd	-0.158881	0.853098
has_dementia	-0.479749	0.618939
missingindicator_creatinine_median	-0.560855	0.570721

Supplementary Table S13. Persistence of flagged combinations to discharge. For each severity, the number of admissions carrying at least one rule match, and the number in which at least one flagged pair was still co-active during the final 24 hours of the stay. MIMIC-IV records the ordered stop time, which is frequently closed administratively at discharge, so these figures are approximate.

Severity	Admissions with match	Admissions with pair active in final 24 h	Percent
major	31,169	23,585	75.7
moderate	7,699	6,486	84.2
minor	9,148	2,900	31.7

Supplementary Table S14. Effect of the co-activity requirement against a co-listing comparator. The comparator applies the same ruleset to any pair of mapped medications appearing in an admission, with no timing condition. Each distinct drug pair is counted once per admission on both sides so the comparison is like for like.

Severity	Co-listed pairs	Timing-aware pairs	Suppressed (%)	Co-listed admissions	ad- Timing-aware admissions	Fewer (%)
major	59,222	39,495	33.3	42,900	31,169	27.3
moderate	9,737	7,779	20.1	9,588	7,699	19.7
minor	15,355	9,173	40.3	15,049	9,148	39.2