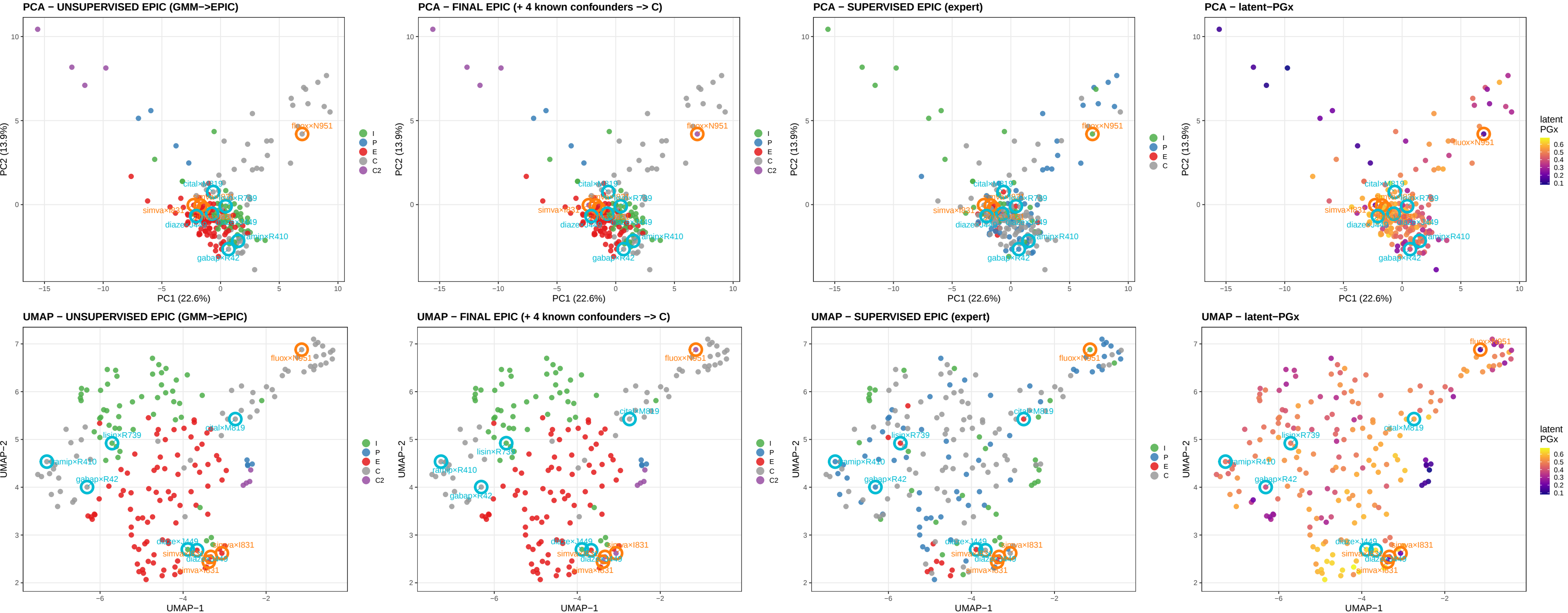


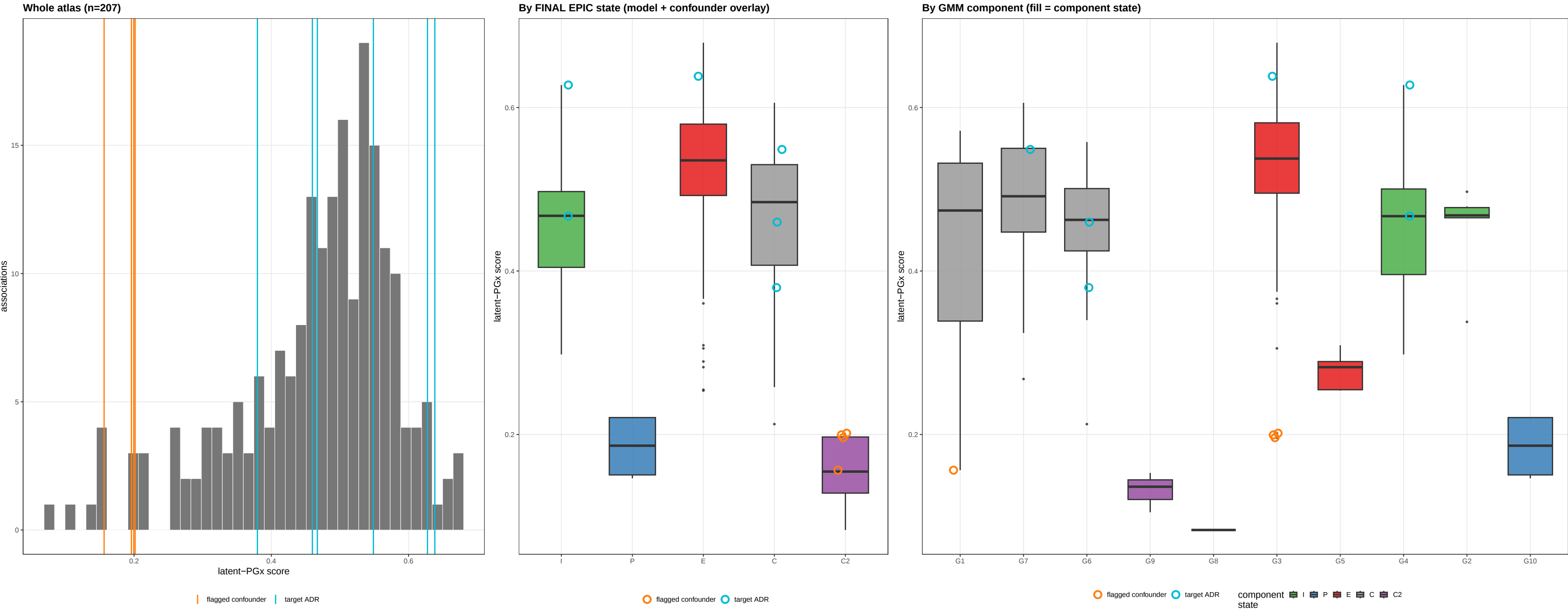
ABM Atlas – 207 assoc.: unsupervised EPIC (GMM->EPIC), + known-confounder overlay, vs expert supervised, 44-feature model

Left: each GMM component mapped to one EPIC state (arg-max of its axes) – no expert labels. Centre-left: FINAL = model + external age/reverse-pharmacology confounder annotation forcing 4 rows to C2. Centre-right: expert cascade. Right: latent-PGx. Cyan = 6 wanted ADF



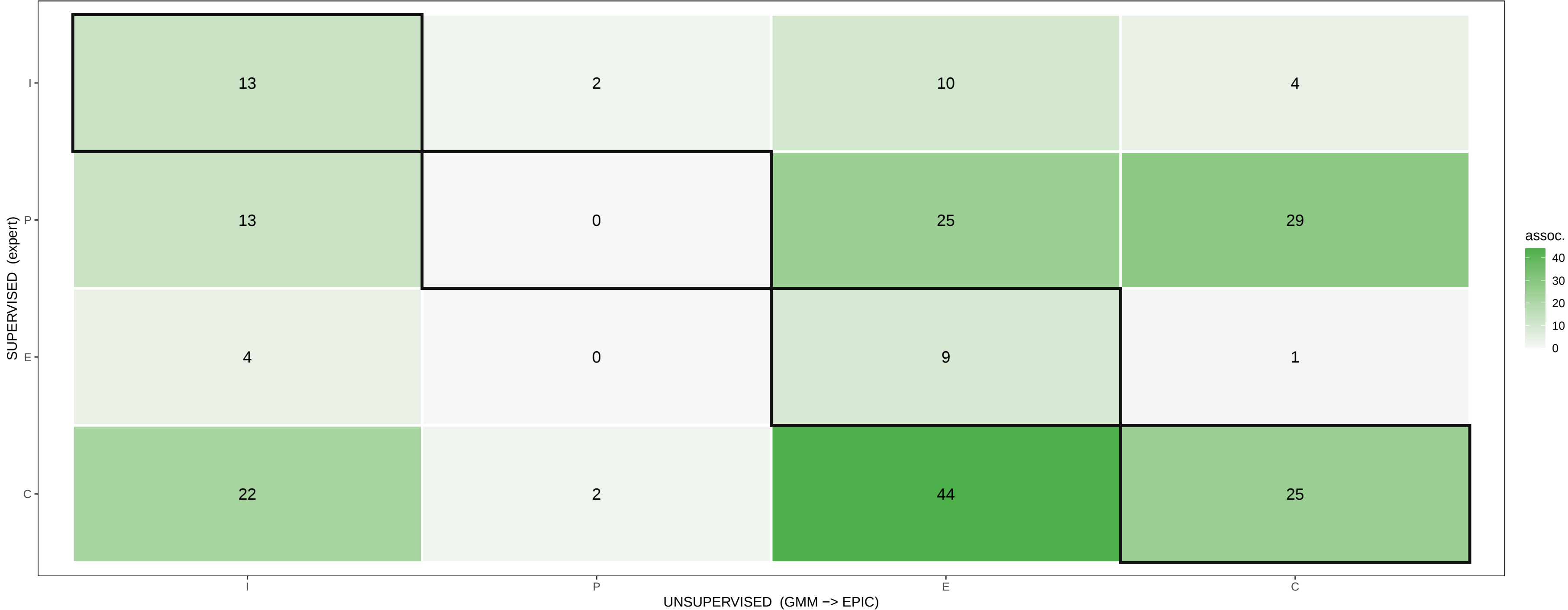
latent-PGx prioritisation score: whole atlas, by EPIC state, by GMM component

Cyan circles = the 6 wanted target ADRs. Orange circles = the 4 unwanted confounders. E is the prioritised state (E vs rest, one-sided Mann-Whitney $p = 3.6e-07$).



Unsupervised (GMM->EPIC) vs supervised (expert) EPIC state

Rows = expert cascade, columns = GMM cluster mapped to EPIC. Boxed = agreement (23% of rows). No expert labels entered the unsupervised column.



GMM components named from their parameters (state = arg-max of the four EPIC axes)

comp	EPIC	name	n	latent	margin	decisive	on_BNF	pct_before	after/before	LOG10P_ICD10	LOG10P_Rx	beta_sup	repl_z	pleio	case_CT	FAERS
G1	C	Carrier-depleted (weak confounder)	20	0.47	0.20	True	0	5.71	14.17	0.84	0.17	2.91	2.09	1	0.14	0.88
G7	C	Sub-threshold predisposition	20	0.49	0.41	True	0	10.66	7.75	0.96	0.61	2.73	1.74	1	1.00	1.23
G6	C	Unreplicated drug-triggered signal	19	0.46	0.05	False	0	0.41	50.75	0.44	0.47	3.68	0.56	1	0.00	1.41
G9	C2	Coronary-risk locus (LPA, 9p21)	3	0.14	2.17	True	1	1.39	41.69	28.87	4.74	0.08	2.80	16	188.00	2.10
G8	C2	Prescription-driving locus (APOE)	1	0.08	4.61	True	0	10.68	7.73	99.47	35.16	0.13	3.56	15	13.00	NA
G3	E	Replicated drug-triggered signal	83	0.54	0.45	True	0	1.79	35.62	0.43	0.29	3.36	2.22	1	2.00	0.78
G5	E	Drug-triggered, no pharmacovigilance support	5	0.28	0.18	True	0	0.00	21.49	0.40	0.33	3.66	2.41	1	2.00	0.46
G4	I	Pre-existing condition (timing only)	45	0.47	0.32	True	0	15.97	5.01	0.48	0.27	3.61	1.59	1	0.00	0.82
G2	I	On-label indication (BNF-confirmed)	7	0.47	2.98	True	1	15.77	5.08	2.00	0.30	1.76	2.62	1	4.00	2.53
G10	P	Major-effect disease locus (Factor V Leiden)	4	0.19	0.11	True	0	15.03	5.83	101.62	0.55	0.22	3.47	5	8.00	0.05

G1 (C) – arg-max C but margin only 0.03; median A1_CASE_CT 0.14 – mapping is low-confidence
G7 (C) – same shape as G10 but LOG10P_ICD10 1.0
G6 (C) – FAERS 1.41 but repl_z 0.56 and zero test carriers
G9 (C2) – variant causes a comorbid disease of the treated population; pleiotropy 16
G8 (C2) – variant causes the drug's own indication; LOG10P_prescribed 35.2
G3 (E) – repl_z 2.22, after/before 35.6, beta_superiority 3.36
G5 (E) – repl_z 2.41 but FAERS 0.46; latent sinks to 0.216
G4 (I) – on_BNF 0.00 but pct_before 16.0% – off-label / protopathic
G2 (I) – on_BNF 1.00, FAERS 2.53; 7/7 expert agreement
G10 (P) – LOG10P_ICD10 101.6, beta_superiority 0.22 – the drug adds nothing