Purpose

Ten randomly−selected associations plus two Effect (E) examples were checked against the published scientific literature, to judge whether the expert cascade classification and the unsupervised cluster assignment are biologically accurate. EPIC: E = Effect (drug causes the ADR), P = Predisposition, I = Indication (confounding by indication), C = inConclusive (low−evidence residual).

−−−

Part A −− 10 random associations

1. simvastatin x varicose veins (I83.1) x DUSP7 −− Expert E vs unsupervised P / P / P / I

The expert cascade calls this Effect (drug causes the event) on FAERS=STRONG, tissue=enriched, N_INDEPENDENT=3. The literature points the opposite way: statins suppress varicose remodeling −− HMG−CoA reductase inhibitors block varicose vein development in mice, atorvastatin/rosuvastatin inhibit wall−stress AP−1 activity in venous smooth−muscle cells, and simvastatin has been trialled for venous ulcers. The curator agrees: "not an indication; shared vascular risk factors (age, obesity, sedentary lifestyle); weak comorbidity." DUSP7 (a dual−specificity phosphatase) has no venous mechanism.

This is the clearest case where the expert classification is likely wrong: labelling a protective/confounded association as a drug−caused Effect inverts the biology. The unsupervised methods, pulling it away from Effect toward Predisposition (3/4) or Indication (HDBSCAN), show appropriate skepticism −− though none lands on the most defensible label, inConclusive/comorbidity (shared metabolic risk driving both statin use and varicose veins). Verdict: expert E is biologically backwards; true class is confounding (C).

2. atenolol x psoriasis (L40.9) x MICAL3 −− Expert C vs unsupervised P / P / P / P

Beta−blockers including atenolol are historically among the drugs most strongly tied to drug−induced/aggravated psoriasis (~20% of psoriasis patients; calcium/cAMP keratinocyte mechanism), with atenolol prominent in adverse−event reports (Cohen BJD 2008; Brauchli BJD 2008, OR 1.31). However, a 2022+ multi−method reanalysis (NHANES + FDA) found no significant signal, arguing the label may be inappropriate. So this is a real−but−contested drug−induced effect.

The expert calls it Comorbidity −− but if atenolol genuinely induces psoriasis, that is an Effect, not coincidental co−occurrence. The cascade is internally inconsistent: the curator quotes strong causal evidence yet the tier lands on the weakest causal class. All four unsupervised methods say Predisposition −− closer to "drug−related" than the expert's Comorbidity. Verdict: expert undercalls a plausible drug−induced ADR; unsupervised P is more defensible, though the honest answer is "contested." MICAL3 is almost certainly incidental.

3. lisinopril x oesophagitis (K20) x GNG13 −− Expert C vs unsupervised P / P / P / P

ACE−inhibitor cough is a well−known class effect that can clinically mimic GERD/oesophagitis −− the "diagnostic confounding" the curator flags. There is no mechanism by which lisinopril causes true oesophageal inflammation, and GNG13 (a G−protein gamma−subunit in taste/neural tissue) has no oesophageal role; the signal is

weak (FAERS weak, tissue low, N_INDEPENDENT=1).

Here the expert Comorbidity/inConclusive call is the more defensible one −− it correctly reads the association as confounding rather than a real drug mechanism. The unanimous unsupervised "Predisposition" reflects the systematic bias: the rare−variant axis (GNG13 is a rare variant) pulls these signals toward P regardless of biology. Verdict: expert right, unsupervised over−calling.

4. amlodipine x pain, unspecified (R52.9) x NPAS3 −− Expert P vs unsupervised P / P / P / I

The outcome is the problem: "unspecified pain" is maximally non−specific, and NPAS3 is a neurodevelopmental transcription factor (brain−expressed) with no link to amlodipine or pain. The Predisposition call was driven by tissue="specific" (a brain−expression artifact), with otherwise weak evidence.

Neither classification is trustworthy because the association is near−noise −− a vague outcome plus an incidental rare variant. Expert P and three unsupervised P agree, but on thin grounds; HDBSCAN's dissent to Indication is no better supported. Verdict: low confidence on both sides; should be flagged uninterpretable.

5. atorvastatin x hypertrophic skin disorders (L91.8) x GALNT12 −− Expert C vs unsupervised P / P / P / P

L91.8 (keloids and other hypertrophic skin disorders) has no established statin association, and GALNT12 (a GalNAc−transferase, colorectal−cancer gene) has no skin−hypertrophy mechanism. Evidence is weak throughout; the curator notes "no indication."

The expert Comorbidity/inConclusive is appropriate −− it reads a weak, mechanism−less association as low−confidence co−occurrence. The unanimous unsupervised Predisposition is again the rare−variant pull misfiring. Verdict: expert correct, unsupervised systematically over−calling P.

6. ramipril x disorientation (R41.0) x GJB6 −− Expert P vs unsupervised P / P / P / I

This has a genuine pathway: ACE inhibitors can cause hyponatremia/SIADH, and severe hyponatremia (Na+ <120 mEq/L) produces confusion/disorientation in the elderly −− documented in case reports where confusion resolved on drug withdrawal. So ramipril can cause disorientation, but indirectly via electrolytes, not through GJB6 (connexin−30, a deafness/skin gene). Signal moderately strong (FAERS=STRONG, tissue enriched, N_INDEPENDENT=2).

The drug−>ADR link is real, making the expert Predisposition plausible (arguably edging to Effect). But the gene−level claim is spurious −− the causal chain is drug−>hyponatremia−>confusion, independent of GJB6. Verdict: drug−ADR plausible (P defensible), gene incidental; strong FAERS justifies taking the signal seriously while distrusting the variant.

7. atorvastatin x fatty liver (K76.0) x FBLIM1 −− Expert I vs unsupervised P / P / P / I

Statins are not formally indicated for NAFLD but are safe and beneficial (atorvastatin reduces aminotransferases, ameliorates NASH; recommended for the accompanying dyslipidemia). NAFLD is a strong metabolic−syndrome marker, so statin users are enriched for fatty liver: confounding by (metabolic) indication.

The expert Indication call is the right read −− it identifies

confounding−by−indication even if "indication" is slightly loose (statins aren't indicated for NAFLD specifically). The unsupervised split (3xP, HDBSCAN I) shows the methods catching the rare−variant signal (FBLIM1, incidental) rather than the confounding structure. Verdict: expert I more defensible.

8. ramipril x acute myocardial infarction (I21.9) x KLF12 −− Expert I vs unsupervised I / I / I / I [OK]

A textbook positive control. ACE inhibitors post−MI are a Class I ESC indication (AIRE, HOPE, SAVE −− prevention of ventricular remodeling); ramipril is prescribed because of the MI. The curator labels it "direct indication"; KLF12 is incidental.

Expert and all four unsupervised methods agree on Indication −− the cleanest result in the set. When the truth is unambiguous confounding−by−indication, the rule cascade and the data−driven clustering converge. Verdict: correct and unanimous −− a reassuring validation that the pipeline gets clear cases right.

9. clopidogrel x prosthetic−device complications (T82.8) x HPSE2 −− Expert I vs unsupervised I / P / P / I

Dual antiplatelet therapy (clopidogrel + aspirin) after coronary stenting is Class I ESC −− clopidogrel is given because the patient has the device, the curator's "maximum confounding." FAERS is STRONG (heavy co−reporting), HPSE2 (heparanase−2, urinary/bladder gene) incidental.

The expert Indication call is correct. The unsupervised methods are split (k−means + HDBSCAN agree I; GMM + LCA drift to P) −− the GMM/LCA dissent is the rare−variant axis mistaking a confounded indication for predisposition. Verdict: expert right; exposes a real weakness −− strong confounding−by−indication can be misread as Predisposition when the variant is rare.

10. ramipril x psoriasis (L40.9) x GLP1R −− Expert C vs unsupervised P / P / P / I

The most biologically interesting −− both drug and gene have real literature. ACE inhibitors are associated with psoriasis (meta−analysis pooled OR 1.52, 95% CI 1.16−2.00; MR + pharmacovigilance). Independently, GLP1R has genuine psoriasis biology: GLP−1R expression is increased in psoriatic plaques, and genetic proxies of GLP1R expression are associated with psoriasis/psoriatic−arthritis risk. Unlike the other nine genes, GLP1R is not obviously incidental.

Yet the expert labels this Comorbidity/inConclusive and the statistics are thin (FAERS weak, N_INDEPENDENT=0 −− no replication). A fascinating tension: the biology is more compelling than the statistics. The expert undercalls it (driven by weak replication); the unsupervised methods scatter without recognizing the GLP1R−psoriasis link, because that biology isn't in their feature space. Verdict: both classifiers miss the interesting signal −− a case where manual, biology−aware review outperforms either automated classifier; deserves follow−up despite weak replication.

−−−

Part B −− two Effect (E) examples

11. citalopram x osteoporosis (M81.9) x DLEU1 −− Expert E vs unsupervised P / P / P / I

This is a genuine drug−induced effect. SSRI−associated bone loss is well documented: a meta−analytic fracture RR ~ 1.6, decreased bone mineral density, and

FDA/EMA warnings. The mechanism is biologically specific −− serotonin transporters are present on osteoblasts, osteoclasts and osteocytes; SSRIs impair osteoclast function peripherally but trigger a central serotonin−dependent sympathetic rise that increases bone resorption, with a net effect of bone loss. Notably, the gene DLEU1 is a documented regulator of osteoclastogenesis −− so unlike most atlas genes, this one has a plausible bone mechanism, and the curator flags the "tissue rescue" by literature.

The expert Effect call is biologically correct −− citalopram causes osteoporosis through an established serotonin−bone pathway. Strikingly, all four unsupervised methods miss it, calling Predisposition (3/4) or Indication (HDBSCAN). This is the concrete manifestation of the structural finding that Effect does not separate as a cluster: even a literature−validated, mechanism−supported drug−induced effect gets absorbed into the rare−variant Predisposition axis. Verdict: expert E correct; unsupervised methods systematically fail to recover Effect.

12. diazepam x COPD (J44.9) x PDE1C −− Expert E vs unsupervised P / P / P / I

Benzodiazepines in COPD are a clinically important, well−established hazard: benzodiazepine use is associated with a ~45% increased risk of COPD exacerbations, respiratory failure, and increased mortality, via GABA−A−mediated central respiratory depression (hypoxemia, reduced respiratory drive, hypercapnia). PDE1C (a phosphodiesterase) is expressed in airway/vascular smooth muscle, giving the gene−level signal more plausibility than most. The curator notes the dual reality: benzodiazepines are relatively contraindicated in COPD (respiratory depression), yet heavily used in this anxious/dyspneic population −− so both a real ADR and confounded co−prescription are present.

The expert Effect call is defensible −− diazepam mechanistically worsens respiratory outcomes in COPD. Again, all four unsupervised methods miss it (P/P/P/I), confirming that the data−driven approach cannot recover the Effect class even when the causality is clinically established. Verdict: expert E correct (with a confounding caveat from co−prescription); unsupervised methods fail to recover Effect, as in example 11.

−−−

Honest advice −− patterns across all 12

1. Gene−level calls are mostly noise. Of twelve genes, only three have real biology (GLP1R−psoriasis, DLEU1−osteoclastogenesis, PDE1C−airway). The other nine are incidental rare variants. This is the single biggest caveat for the atlas: it treats gene−level GWAS hits as meaningful, but at the individual−association level most are spurious. The real, checkable signal is almost always the drug−>ADR epidemiology, not the specific variant.

2. The expert cascade's Effect tier is largely sound. Two of the three Effect cases examined (SSRI/osteoporosis, benzodiazepine/COPD) are genuine, literature−validated drug−induced effects with plausible mechanisms; only one (statin/varicose) was biologically backwards. The cascade also reads confounding−by−indication correctly (#7, #8, #9). Its main weakness is being conservative on drug−induced ADRs −− calling real/contested effects (atenolol & ramipril psoriasis, ramipril disorientation) "Comorbidity."

3. The unsupervised methods cannot recover Effect, and over−call Predisposition. In both genuine Effect cases (#11, #12) all four methods said P or I, never E.

And for confounded indications with rare variants (#3, #5, #9) they over−call P.
The rare−variant standard−error axis dominates their behaviour, biasing them toward
Predisposition regardless of biology. This is the concrete, literature−grounded
version of the T4−recovery failure documented quantitatively elsewhere.

4. Neither classifier is ground truth −− use as a screen. Treat the
classification (either version) as a prioritization layer, not a verdict. Every
association destined for the manuscript needs the manual triage performed here:
(a) is the drug−>ADR link real, protective, or confounded? (b) is the gene plausible
or incidental? The cases worth manual follow−up are the disagreements (#1, #2,
#10) and the biology−rich outliers (#10 GLP1R, #11 DLEU1) −− not the unanimous ones.

−−−

Sources

− Beta−blocker psoriasis: Azzouz, Br J Clin Pharmacol 2022 (doi 10.1111/bcp.15330); reanalysis (no signal),
HMP Global; mechanisms, PMC7398737.
− ACE inhibitor psoriasis: meta−analysis OR 1.52, Sci Rep 2021 (s41598−021−89490−z); MR + pharmacovigilance,
Expert Rev Clin Pharmacol 2024 (PubMed 38078460).
− Statins / NAFLD: PMC5998729; atorvastatin in NAFLD, PubMed 38304133.
− ACEi hyponatremia / confusion: PMC6368929.
− Statins / venous disease: HMG−CoA inhibitors suppress varicose remodeling, PubMed 26908399; IntechOpen
review.
− GLP1R / psoriasis: increased in plaques, PubMed 23362875; genetic proxies, medRxiv 2025.10.12.25337838.
− SSRI / bone loss: serotonin reuptake inhibitors and bone, PMC6372777; central/peripheral mechanism,
PMC5053870; meta−analysis, PMC9568413.
− Benzodiazepines / COPD: ~45% exacerbation risk, Eur Respir J 44(2):332; Ann Am Thorac Soc PMC6344455.