Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Biological Determinants of Adverse Drug Reaction Predictability in Drug–Gene Networks

Mihir Sunil Chachar

Ollscoil na Gaillimhe – University of Galway

Haixuan Yang

haixuan.yang@universityofgalway.ie

Ollscoil na Gaillimhe – University of Galway

Research Article

Keywords:

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11719277/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Biological Determinants of Adverse Drug Reaction Predictability in Drug–Gene Networks

Mihir Sunil Chachar[1] and Haixuan Yang[1*]

[1]School of Mathematical & Statistical Sciences, University of Galway, Galway, H91TK33, Ireland.

*Corresponding author(s). E-mail(s): haixuan.yang@universityofgalway.ie;
Contributing authors: m.chachar1@universityofgalway.ie;

Abstract

Adverse drug reactions (ADRs) are a major cause of drug attrition and withdrawal, yet computational models compete on accuracy without asking why some ADRs are more predictable than others. Rather than proposing a more accurate predictor, this study uses an interpretable network model to investigate why individual ADRs differ in predictability. We build a drug similarity network from the drug–gene interaction profiles of 536 drugs and predict 5,599 side effects by harmonic function-based label propagation under five-fold cross-validation. Performance (AUROC 0.670, AUPRC 0.210) exceeded that of a k-nearest-neighbour baseline on the same similarity matrix and fell to chance when the network was shuffled. Because AUPRC depends strongly on ADR frequency, AUROC measured predictability for the 970 ADRs reported for at least 10 drugs; 76% were predicted significantly better than chance. Predictability was strongly associated with target coherence, defined as the extent to which drugs associated with the same ADR are more similar to one another in their gene targets than they are to other drugs (Spearman's $\rho = 0.73$). The most predictable ADRs were class effects with over-represented receptor genes, such as coronary arteriospasm (HTR1B, HTR1D), and tardive dyskinesia (DRD2, HTR2A); the least predictable were non-specific reactions of pharmacologically unrelated drugs. Pathway enrichment was found for 88% of the 100 best predicted ADRs, but only 26% of the 100 worst, mainly in neurotransmitter receptor signalling. These results show that ADR predictability reflects the biological coherence of the drugs associated with an ADR, particularly their shared targets.

Keywords: Adverse drug reactions, Drug–gene interactions, Network propagation, Harmonic function, Target coherence, Pathway enrichment

1 Introduction

Adverse drug reactions (ADRs) remain one of the major causes of patient morbidity, mortality, prolonged hospitalisation, and escalating healthcare costs worldwide. Many drugs fail to reach the market despite rigorous preclinical and clinical trials because of adverse reactions, and others are withdrawn or receive regulatory safety warnings after post-marketing surveillance [1]. Many reactions become evident only after widespread use of the drug, owing to population heterogeneity, limited trial sizes, genetic diversity, polypharmacy and long-term drug exposure. ADR prediction models based on artificial intelligence, network biology and large-scale biomedical databases now complement traditional pharmacovigilance by enabling prediction before clinical observation [2, 3].

Early approaches relied on similarity-based inference, using chemical similarity, molecular fingerprints, therapeutic classifications, and known side-effect profiles, assuming that structurally similar compounds show similar pharmacological behaviour [1, 4]. Common ADRs were predicted reasonably well from similarity matrices, but these methods did not capture the molecular mechanisms responsible for the reactions [5]. Later studies integrated multiple biological similarity measures into comprehensive prediction frameworks, improving accuracy but remaining largely feature-driven, without explicitly modelling the biological relationships governing drug action [1, 6].

Deep learning enabled models to learn hierarchical representations directly from molecular structures and high-dimensional biological data. Neural networks, graph-based transformers and embedding methods have improved performance across multiple ADR benchmark datasets [7]. Integrating drug-induced gene expression profiles with chemical structure information improved ADR prediction, as did molecular topology and higher-order knowledge graphs [8–10]. Incorporating diverse biological entities such as proteins, pathways, transformers, enzymes, diseases, and patients into unified graph architectures helped to capture the complex interactions underlying ADRs [11, 12], and transformer-based graph models and pathway-aware graph neural networks model dependencies between molecular features and biological pathways [2, 13, 14].

Despite these advances, several limitations remain. First, most studies use metrics such as AUROC and AUPRC to evaluate and optimise predictive performance rather than to understand the biological mechanisms responsible for differences in prediction across individual ADRs, so current approaches cannot fully explain why certain ADRs are consistently predicted with greater confidence than others [11, 12]. Second, many state-of-the-art graph neural network (GNN) models require large training datasets, extensive parameter optimisation and complex heterogeneous graph architectures, which reduce transparency and limit biological interpretability. Third, pathway analyses are often performed only as post hoc validation and rarely inform the central interpretation of model performance. This leaves a gap between computational prediction and mechanistic understanding of ADR biology [2, 13, 15–17].

Biological evidence suggests that ADRs are network-based phenomena rather than isolated events: the therapeutic and toxicological effects of a drug arise from its target proteins and genes, whose downstream effects propagate across signalling pathways, regulatory networks and cellular processes [6, 18, 19]. Genes encoding drug targets

carry substantial predictive information through genetic interactions, target gene phenotypes and protein interaction networks. Disrupting highly connected hubs triggers widespread biological consequences, whereas disrupting isolated or multi-functional genes produces heterogeneous and less predictable results [2, 20, 21]. This supports the hypothesis that drug–gene interaction networks fundamentally influence predictability.

Graph-based harmonic propagation offers a principled and interpretable framework for semi-supervised prediction, using Gaussian random fields and harmonic functions to propagate labels across a weighted network while preserving known annotations [22]. The approach is deterministic and governed by graph topology and minimum-energy optimisation, in contrast to deep learning. Laplacian-based feature selection and graph regularisation have since strengthened its application to drug–gene interaction networks [15, 23]. However, previous studies have not systematically explored whether the biological organisation of these networks affects ADR predictability, leaving open the question of whether predictable ADRs are governed by coherent biological pathways while less predictable ADRs reflect fragmented mechanisms. Rather than developing a more complex model to maximise predictive accuracy, this study uses an interpretable network-propagation framework to investigate why some ADRs are intrinsically more predictable from drug–gene relationships than others. We combine drug–gene interaction networks with harmonic function propagation to predict and rank ADRs, relate the ranking to the coherence of drug targets, and interpret it through the target genes over-represented among the drugs associated with each ADR and through the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and Gene Ontology (GO) enrichment analysis. The objective is therefore not a state-of-the-art ADR predictor, but an explanation of the biological determinants of ADR predictability.

2 Methods

The overall computational framework is shown in Fig. 1. The workflow consists of (i) construction of drug–side effect and drug–gene interaction matrices, (ii) generation of a weighted drug similarity network, (iii) harmonic function-based semi-supervised prediction of ADRs with five-fold cross-validation, (iv) quantification of the predictability of each ADR, and (v) biological interpretation of the most and least predictable ADRs through their over-represented target genes and KEGG and GO enrichment analysis.

2.1 Drug, gene and ADR data integration

Drug–gene interaction data were extracted from the Drug–Gene Interaction Database (DGIdb) and transformed into a binary matrix in which rows represent drugs ($n = 777$), and columns represent genes ($n = 2022$); any reported interaction between a drug and a gene was coded as one. Drug–side effect data were obtained from Zhong et al. [4], who processed them from SIDER; the curated drug-ADR associations were compiled into an independent binary matrix (1020 drugs, 5599 side effects). A standard drug identifier was used in both datasets to ensure consistent integration. Duplicates and incomplete records were removed during preprocessing, and ADRs were defined by the Medical Dictionary for Regulatory Activities (MedDRA) preferred terms. All

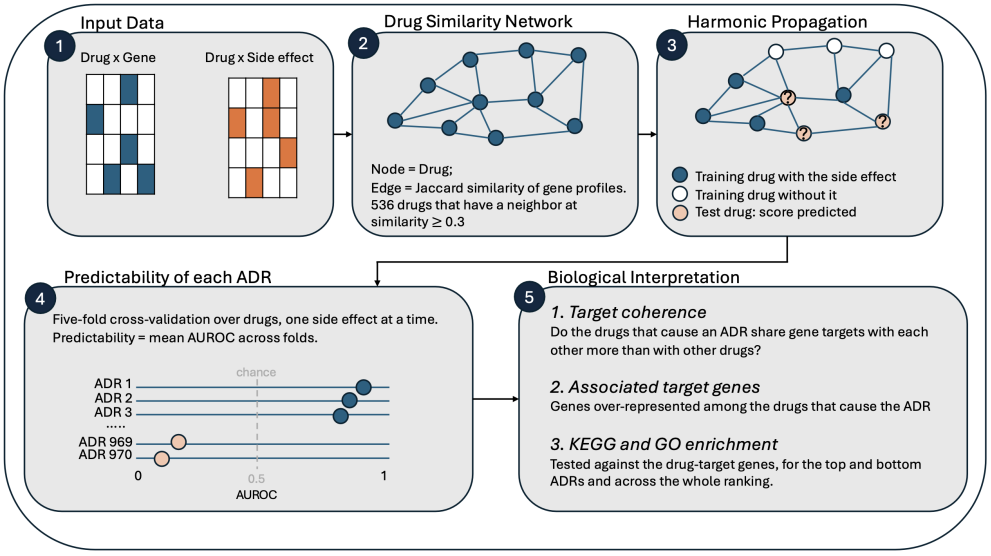

Fig. 1 Workflow of the study. (1) Binary drug–gene and drug–side effect matrices are assembled. (2) Drugs are connected in a network weighted by Jaccard similarity of their gene interaction profiles. (3) For each side effect, harmonic propagation predicts a score for the test drugs from the labels of the training drugs. (4) Five-fold cross-validation gives the mean AUROC of each ADR, by which the ADRs are ranked. (5) The ranking is interpreted through target coherence, the target genes over-represented among the drugs associated with each ADR, and KEGG and GO enrichment.

777 drugs of the drug–gene matrix were also present in the side-effect matrix and formed the starting set for network construction.

2.2 Construction of the drug similarity network

Because the interaction profiles are binary, the Jaccard coefficient [24, 25] was used to quantify drug–drug similarity. For two drugs i and j,

$$J(i,j) = \frac{|A_i \cap A_j|}{|A_i \cup A_j|}, \tag{1}$$

where A_i and A_j denote the corresponding binary feature sets. Pairwise similarity was calculated independently from drug–gene interactions (drug–gene similarity) and from drug–ADR annotations (drug–side effect similarity).

To limit the sparsity, a drug was retained only if its drug–gene similarity to at least one other drug was ≥ 0.3, which left 536 of the 777 drugs (69%). Drugs below this threshold share too few targets with any other drug for similarity-based inference. The threshold of 0.3 was therefore used to select drugs with sufficient target-profile similarity for network-based analysis, rather than to remove lower-weighted edges during the label propagation. Among the retained drugs, 2,416 pairs have a similarity ≥ 0.3; these edges come from 41 components, the largest containing 407 drugs. Once the 536

drugs had been selected, all non-zero pairwise similarities among them were retained as propagation weights: the weighted matrix $\mathbf{W} = [W_ij]$ contains the drug–gene similarity of every pair of retained drugs (36,726 non-zero pairs, forming three components), with self-similarity set to zero. The degree of each drug node is $d_i = \sum_j w_{ij}$, and these values form the diagonal degree matrix $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_n)$. Because the threshold determines which drugs are analysed, the analysis was repeated for thresholds between 0.1 and 0.5 to establish whether the principal conclusions depend on this choice (Section 3.1).

2.3 Network propagation using the harmonic function

The unnormalised graph Laplacian is $\mathbf{L} = \mathbf{D} - \mathbf{W}$. Splitting the drugs into labelled (training, l) and unlabelled (test, u) groups, the Laplacian and the label vector for a given ADR can be written in block form as

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{ll} & \mathbf{L}_{lu} \\ \mathbf{L}_{ul} & \mathbf{L}_{uu} \end{bmatrix}, \qquad \mathbf{f} = \begin{bmatrix} \mathbf{f}_l \\ \mathbf{f}_u \end{bmatrix}, \tag{2}$$

where $\mathbf{f}_l$ contains the known ADR annotations and $\mathbf{f}_u$ the unknown scores to be predicted. The harmonic solution minimises the graph energy [22]

$$E(f) = \frac{1}{2} \sum_{i,j} w_{ij} \, (f_i - f_j)^2, \tag{3}$$

so that highly similar drugs are forced to share similar adverse reaction scores. Setting the derivative of the energy with respect to $\mathbf{f}_u$ to zero gives the linear system $\mathbf{L}_{uu}\mathbf{f}_u = -\mathbf{L}_{ul}\mathbf{f}_l$. With a small regularisation term $\epsilon = 10^{-6}$, which keeps the system solvable when a test drug has no path to any labelled drug, the closed-form prediction (harmonic scores) is

$$\mathbf{f}_u = (\mathbf{L}_{uu} + \epsilon\mathbf{I})^{-1}(-\mathbf{L}_{ul}\mathbf{f}_l), \tag{4}$$

where $\mathbf{I}$ is the identity matrix.

2.4 Cross-validation and performance evaluation

The framework was evaluated using five-fold cross-validation over drugs. In each fold, approximately 80% of drugs (428–429) were used for training and the remaining 107–108 for testing, such that every drug was tested exactly once and no test-set drug appeared in the corresponding training set. In a given fold, an ADR was evaluated if at least one training drug and at least one test drug carried it and at least one test drug did not.

Performance was assessed using the area under the receiver operating characteristic curve (AUROC), the area under the precision–recall curve (AUPRC, computed as average precision) and the maximum F-measure (F_{max}). Harmonic scores were rounded to 10 decimal places before the metrics were computed, so that scores that differ only through floating-point error are treated as ties. AUROC and AUPRC

were computed per ADR and fold and then averaged over all evaluated ADR–fold combinations.

F_{max}, the standard measure in the evaluation of protein function prediction [26, 27], avoids choosing a single decision threshold. Harmonic scores and k-NN scores both lie between 0 and 1, so the same range of thresholds were applied to both without normalisation. For a threshold t, a side effect was predicted for a test drug if its score was $\geq t$, and

$$\mathrm{pr}(t) = \frac{1}{m(t)} \sum_{i=1}^{m(t)} \frac{\mathrm{TP}_i(t)}{\mathrm{TP}_i(t) + \mathrm{FP}_i(t)}, \qquad \mathrm{rc}(t) = \frac{1}{n} \sum_{i=1}^{n} \frac{\mathrm{TP}_i(t)}{\mathrm{TP}_i(t) + \mathrm{FN}_i(t)}, \tag{5}$$

where precision is average over the $m(t)$ test drugs with at least one predicted side effect and recall over all n test drugs. Then

$$F_{\max} = \max_t \frac{2\,\mathrm{pr}(t)\,\mathrm{rc}(t)}{\mathrm{pr}(t) + \mathrm{rc}(t)}, \tag{6}$$

where t from 0.01 to 1.00 in steps of 0.01. $F_{\max}$ was computed for each fold and averaged. To check that choosing t on the test fold did not inflate the result, F was also computed for each fold at the threshold that was best on the other four folds.

A k-NN classifier was used as a baseline. It uses the same weighted matrix $\mathbf{W}$ and scores a side effect for a test drug by the proportion of its k most similar training drugs that carry that side effect. k-NN was chosen as the baseline because it uses the same drug–gene similarity matrix as harmonic propagation, so that the comparison assesses the contribution of network propagation without confounding differences in input features. The baseline was evaluated on the same folds for $k = 5, 10, 15$ and 20, and the value with the highest mean AUPRC ($k = 10$) is reported. As a negative control, the drug labels of $\mathbf{W}$ were randomly permuted, which preserves the structure of the network but removes the correspondence between a drug's neighbours and its side effects, and the cross-validation was repeated.

2.5 Predictability of individual ADRs

The chance level of AUPRC equals the proportion of positive drugs [28], so AUPRC values are not comparable between ADRs of different frequency without any normalisation. The predictability of an ADR was therefore defined as its mean AUROC across the five folds, which was substantially less dependent on ADR frequency on this dataset (Section 3.2). ADRs annotated to an identical set of drugs were merged into a single entry. To obtain stable estimates, an ADR was ranked only if at least 10 drugs carried it, at least 10 drugs did not, and it could be evaluated in all five folds.

Whether an ADR was predicted better than chance was tested on the out-of-fold scores of all drugs with a one-sided Mann–Whitney U test (scores of drugs with the ADR greater than scores of drugs without it). P-values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure, and FDR < 0.05 was considered significant.

To relate predictability to the organisation of drug targets, we computed for each ADR the mean drug–gene similarity between pairs of drugs that carry the ADR (within-ADR similarity) and the mean similarity between drugs that carry it and drugs that do not (between-group similarity). Their difference is referred to as *target coherence*. Associations with mean AUROC were assessed using Spearman's rank correlation.

2.6 Target genes and functional enrichment

The 10 ADRs with the highest and the 10 with the lowest mean AUROC were examined in detail. For each ADR, every gene targeted by at least two of its drugs was tested for over-representation among the drugs that carry the ADR relative to all 536 drugs (hypergeometric test, Benjamini–Hochberg correction across the genes tested for that ADR). Genes with FDR < 0.05 are referred to as the associated target genes of the ADR.

Pathway and functional enrichment was tested for every ADR in the two groups that had at least three associated target genes. KEGG pathway and GO Biological Process gene sets were obtained from the Enrichr libraries KEGG_2026 and GO_Biological_Process_2026. Over-representation of each gene set among the associated target genes was assessed with a hypergeometric test in which the background was the 1,017 genes that interact with at least one of the 536 drugs. This background was used because every gene in the dataset is a drug target, and a genome-wide background would overstate enrichment. Gene sets with fewer than five background genes were not tested, which left 246 KEGG pathways and 1,407 GO terms. P-values were adjusted with the Benjamini–Hochberg procedure within each library and ADR, and an adjusted p-value < 0.05 was considered significant.

To examine enrichment beyond these two groups of 10, the same test was applied to every ranked ADR with at least three associated target genes. The best and worst predicted ADRs were then compared at three group sizes (10, 100 and half of the ranking, 485 ADRs each) with respect to proportion of ADRs that had at least one significant term; ADRs with too few associated genes to be tested were counted as having none. For each gene set that was significant for at least five ADRs, we compared the number of ADRs enriched for it in the top and bottom groups (Fisher's exact test) and, without any cut-off, the AUROC of the ADRs enriched for it with that of all other ranked ADRs (two-sided Mann–Whitney U test). These p-values were adjusted with the Benjamini–Hochberg procedure across gene sets within each library.

All analyses were performed in Python in the Google Colab environment, using NumPy, SciPy, scikit-learn, NetworkX and Matplotlib.

3 Results

3.1 Overall performance and robustness

The analysed network comprised 536 drugs and 5,599 side effects. Across the five folds, 2,752 side effects could be evaluated in at least one fold, giving 10,146 side effect–fold evaluations. Harmonic propagation achieved a mean AUROC of 0.670 (range across

folds 0.662–0.685) and a mean AUPRC of 0.210 (0.192–0.233) (Table 1, Fig. 2). On average 6.4% of the test drugs were positive for an evaluated side effect. $F_{\max}$ was 0.466 (0.455–0.476), reached at thresholds between 0.26 and 0.30; with the threshold chosen on the other four folds it was 0.464.

Compared with the k-NN baseline, harmonic propagation achieved higher AUROC (0.670 versus 0.627) and AUPRC (0.210 versus 0.171), whereas $F_{\max}$ was similar (0.466 versus 0.459) (Table 1). $F_{\max}$ is computed per drug over all of its side effects and is therefore dominated by common ADRs: the 341 side effects reported for at least 10% of the drugs account for 63% of all drug–side effect associations. The advantage of propagation thus lies mainly in ranking the drugs for each individual ADR, which AUROC and AUPRC measure. When the drug labels of the network were permuted, performance fell to an AUROC of 0.512, an AUPRC of 0.107 and an $F_{\max}$ of 0.419, showing that the predictive signal comes from the correspondence between drug–gene similarity and side-effect profiles. This correspondence is modest at the level of drug pairs; across all 143,380 pairs of analysed drugs, side-effect similarity increased only weakly with drug–gene similarity (Spearman's $\rho = 0.18$).

Table 1 Mean performance across five cross-validation folds for harmonic propagation and the k-NN baseline ($k = 10$)

Metric	Harmonic propagation	k-NN baseline	Difference
AUROC	0.670	0.627	0.043
AUPRC	0.210	0.171	0.039
$F_{\max}$	0.466	0.459	0.007

Note: AUROC and AUPRC are means over all evaluated side effect–fold combinations; $F_{\max}$ is the mean of the five folds (best threshold 0.26–0.30 for harmonic propagation and 0.21–0.31 for k-NN). Differences were calculated from unrounded values.

The similarity threshold determines which drugs enter the analysis. Lower thresholds retained more drugs and produced a more connected graph, and higher thresholds fragmented it (Table 2). Mean AUROC ranged from 0.633 to 0.692 and was highest at thresholds of 0.1–0.2, whereas mean AUPRC increased with the threshold. Because each threshold is evaluated on a different set of drugs these values are not strictly comparable, but they show that the principal conclusions are not specific to the threshold of 0.3.

3.2 Predictability of individual ADRs

Performance differed widely between side effects, and the two summary metrics behaved differently with respect to how common a side effect is (Fig. 3). Mean AUPRC rose steadily with the number of drugs carrying the side effect, approaching 0.9 for side effects reported for most drugs, and took extreme values of 0 or 1 for side effects with only two to four drugs. Mean AUROC showed no such trend. AUROC was therefore used to compare side effects, restricted to those with at least 10 drugs on each side that were evaluable in every fold.

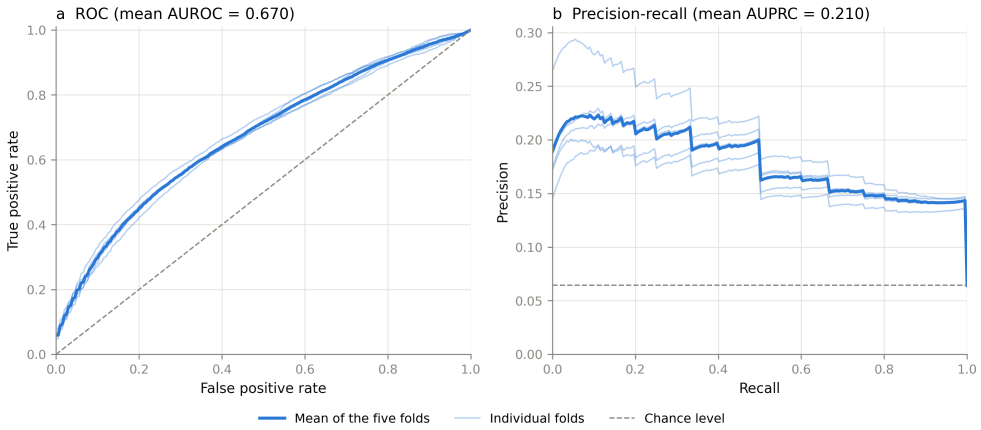

Fig. 2 Performance curves across five-fold cross-validation. **(a)** Mean ROC curve of each fold (thin lines) and of all folds (thick line); the dashed diagonal represents a random classifier. **(b)** Mean precision–recall curves; the dashed line marks the mean proportion of positive test drugs

Table 2 Effect of the drug–gene similarity threshold on the network and on the performance of harmonic propagation

Threshold	Drugs (%)	Edges	Components	Largest	AUROC	AUPRC
0.1	758 (98)	20,388	5	747	0.689	0.179
0.2	694 (89)	6,396	16	634	0.692	0.201
0.3	536 (69)	2,416	41	407	0.670	0.210
0.4	401 (52)	1,116	56	160	0.654	0.228
0.5	308 (40)	686	56	32	0.633	0.243

Note: Drugs, number retained (percentage of the 777 drugs). Edges, components and largest (size of the largest component) refer to the graph of retained drugs with edges at or above the threshold. Performance is the mean over all evaluable side effect–fold combinations.

After merging side effects with identical drug profiles (2,596 distinct profiles), 970 ADRs met these criteria, with between 10 and 475 drugs each (median 34). Among them, the number of drugs was strongly correlated with AUPRC (Spearman's $\rho = 0.69$) but only weakly with AUROC ($\rho = -0.10$). Mean AUROC ranged from 0.280 to 0.964 (median 0.672, interquartile range 0.615–0.744). In total, 742 of the 970 ADRs (76%) were predicted significantly better than chance (FDR < 0.05), and 36 had a mean AUROC below 0.5.

3.3 Predictability reflects the coherence of drug targets

The central result of this study is the strong association between predictability and target coherence (Spearman's $\rho = 0.73$, $p = 2.0 \times 10^{-159}$; Fig. 4). The association was driven mainly by the similarity among the drugs that carry the ADR ($\rho = 0.64$) and much less by their similarity to other drugs ($\rho = 0.20$). In the quartile of

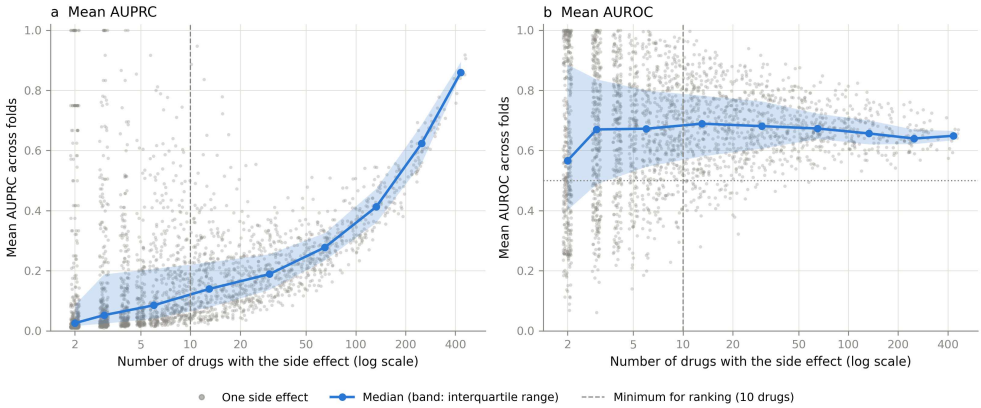

Fig. 3 Performance against the number of drugs that carry each side effect. **a** Mean AUPRC. **b** Mean AUROC. Each dot is one side effect; the line and band show the median and interquartile range within frequency bins. The vertical dashed line marks the minimum of 10 drugs required for ranking; the dotted line in **b** marks chance

ADRs with the highest AUROC, the median within-ADR similarity was 0.06, about twice the between-group similarity (0.032), and 99% of ADRs were predicted significantly better than chance. In the lowest quartile, within-ADR similarity (0.032) was indistinguishable from between-group similarity (0.029), and 26% were significant.

Because target coherence is calculated from the same drug–gene similarities used for propagation, its association with predictive performance is not statistically independent of the model. The significance of this result therefore lies not in demonstrating an independent predictor of model performance, but in quantifying the extent to which ADR predictability reflects the organisation of drugs in target space. The gene- and pathway- level analyses below, which use the identity of the shared genes rather than the similarity values, provide the biological interpretation of this relationship.

The same pattern was seen at the level of individual genes. At least one associated target gene was found for 881 of the 970 ADRs, and the number of associated genes increased with AUROC ($\rho = 0.52$): the median was 33 genes in the highest quartile and 4 in the lowest, where 24% of ADRs had none.

3.4 The most and least predictable ADRs

The 10 most predictable ADRs had mean AUROC values between 0.910 and 0.964 and were all significant (Table 3, Fig. 5). Each was linked to a recognisable group of drugs and to a set of associated target genes. Coronary arteriospasm was reported for seven triptans, two ergot alkaloids and esmolol, and its associated genes were serotonin receptors (HTR1D, HTR1B, HTR1E, HTR1F). Serotonin syndrome was associated with HTR1B, HTR1A and the serotonin transporter gene SLC6A4, tardive dyskinesia with serotonin and dopamine receptor genes (HTR2A, DRD2), and skin striae, reported almost exclusively for corticosteroids, with the glucocorticoid receptor gene NR3C1. Not every association pointed to a mechanism: for superinfection, reported

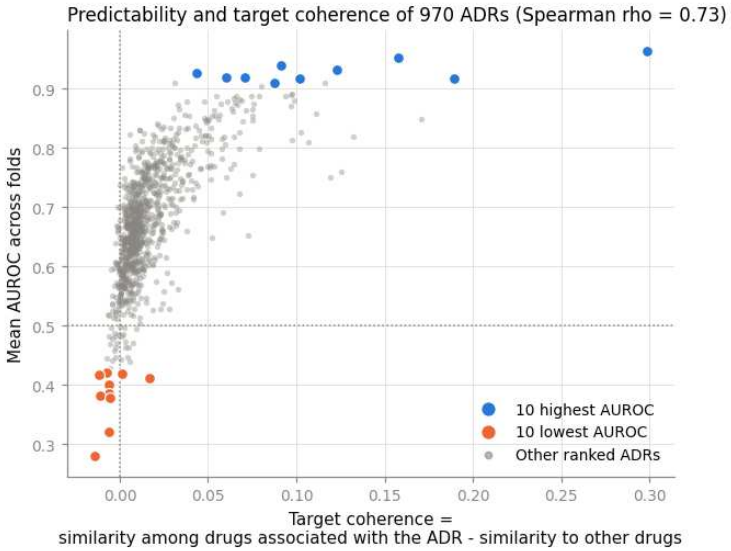

Fig. 4 Predictability against target coherence for the 970 ranked ADRs. Target coherence is the mean drug–gene similarity among the drugs that carry an ADR minus their mean similarity to drugs that do not. The 10 ADRs with the highest and lowest mean AUROC are highlighted. Dotted lines mark chance-level AUROC and zero coherence

for antibacterial agents, the associated genes were transporters and metabolic genes (ABCC2, ABCC4, G6PD, NR1I2) rather than targets related to the reaction itself. The genes of the remaining ADRs are given in Table 3.

The 10 least predictable ADRs had mean AUROC values between 0.280 and 0.421, and none was predicted better than chance (Table 4, Fig. 5). Their within-ADR similarity (0.006–0.058) was close to or below their between-group similarity, six of the ten had no associated target gene, and the drugs associated with them came from unrelated pharmacological classes. They included non-specific reactions (face swelling, eye swelling, altered or bitter taste, skin infection), congenital conditions and lung carcinoma.

3.5 Pathway and Gene Ontology enrichment

Enrichment could be tested for all 10 of the most predictable ADRs but for only two of the 10 least predictable, because the other eight had fewer than three associated target genes (Table 5).

For the ADRs linked to monoamine pharmacology, the enriched terms matched the associated receptors. The genes associated with coronary arteriospasm and serotonin syndrome were enriched for the KEGG serotonergic synapse pathway and for GO terms describing G protein-coupled serotonin receptor signalling, and those associated with tardive dyskinesia for neuroactive ligand signalling and the serotonergic synapse. The genes associated with superinfection were enriched for ABC transporters and xenobiotic transport.

Table 3 The 10 most predictable ADRs and their associated target genes

ADR	Drugs	AUROC	S_w	S_b	Genes	Leading associated genes
Arteriospasm coronary	10	0.964 (0.076)	0.311	0.012	7	HTR1D (8/19), HTR1E (7/13), HTR1B (8/24)
Parotid swelling	12	0.953 (0.041)	0.206	0.048	19	SLC6A2 (9/22), SLC6A4 (9/27), DRD1 (10/49)
Serotonin syndrome	25	0.939 (0.061)	0.121	0.030	24	HTR1B (14/24), SLC6A4 (13/27), HTR1A (13/30)
Protrusion tongue	13	0.932 (0.060)	0.165	0.042	25	DRD2 (10/42), ATXN2 (7/14), DRD1 (10/49)
Venous thromboembolism	15	0.927 (0.052)	0.069	0.026	34	MC4R (5/6), RABEP1 (4/5), SH2B1 (4/5)
Superinfection	11	0.919 (0.059)	0.097	0.037	10	ABCC2 (5/19), ABCC4 (3/8), G6PD (5/36)
Myelosuppression	12	0.919 (0.075)	0.102	0.031	69	TUBA1B (6/7), TUBA1C (6/7), TUBA3D (6/7)
Skin striae	16	0.917 (0.068)	0.213	0.024	12	ANXA1 (11/12), NR3C1 (13/22), AR (9/70)
Tardive dyskinesia	31	0.917 (0.047)	0.139	0.037	47	HTR2A (22/41), DRD2 (20/42), HRH1 (17/42)
Drug tolerance	10	0.910 (0.028)	0.123	0.036	19	CYP3A5 (8/54), FKBP5 (4/5), SLC6A4 (6/27)

Note: Drugs, number of drugs carrying the ADR. AUROC, mean (standard deviation) across five folds. S_w, mean drug–gene similarity among the drugs carrying the ADR; S_b, mean similarity between these drugs and all other drugs. Genes, number of associated target genes (FDR < 0.05). For each leading gene, the number of the ADR's drugs that interact with it is given over the number of all 536 drugs that interact with it.

Elsewhere, enrichment was less informative. The leading KEGG terms for myelosuppression were all driven by the same tubulin genes, and the 63 pathways significant for venous thromboembolism were almost all cancer signalling pathways that share the growth factor receptor and kinase genes targeted by the antineoplastic drugs in this group. No term was significant for skin striae; the closest was nuclear receptor-mediated steroid hormone signalling (adjusted $p = 0.06$). In the bottom group, the significant terms for xerophthalmia came from three UDP-glucuronosyltransferase genes shared by two drugs, and those for lung carcinoma from tubulin genes shared by docetaxel and vinorelbine; neither related to the reaction.

3.6 Pathway enrichment across the whole ranking

Of the 970 ranked ADRs, 802 (83%) had at least three associated target genes and could be tested, and 552 (57%) and 578 (60%) had at least one significant KEGG pathway or GO term, respectively. Both proportions fell steadily along the ranking (Fig. 6): in the best predicted tenth, 99% of ADRs could be tested and 88% had a significant KEGG pathway, against 54% and 26% in the worst predicted tenth.

The difference between well and poorly predicted ADRs was present at every group size (Table 6). Among the 100 best predicted ADRs, 88% had a significant KEGG pathway, compared with 26% of the 100 worst predicted; for the two halves of the

Table 4 The 10 least predictable ADRs and their associated target genes

ADR	Drugs	AUROC	S_w	S_b	Genes	Leading associated genes
Xerophthalmia	15	0.421 (0.184)	0.020	0.028	3	UGT1A10 (2/2), UGT1A7 (2/2), UGT1A8 (2/2)
Taste bitter	17	0.419 (0.141)	0.034	0.033	0	None
Carcinoma of lung	10	0.416 (0.102)	0.006	0.017	22	XRCC1 (2/2), CASP7 (2/3), RRM1 (2/3)
Taste altered	12	0.411 (0.128)	0.058	0.041	1	CYP19A1 (3/9)
Skin infection	10	0.401 (0.149)	0.024	0.030	0	None
Eye swelling	20	0.386 (0.217)	0.022	0.029	0	None
Congenital anomaly[1]	13	0.383 (0.301)	0.015	0.026	2	ABCC1 (3/13), RAPGEF4 (2/4)
Paresis	21	0.379 (0.054)	0.024	0.029	0	None
Swelling face	22	0.322 (0.066)	0.027	0.033	0	None
Congenital eye disorder	13	0.280 (0.095)	0.011	0.025	0	None

Note: Columns as in Table 3.

[1] Annotated to the same drugs as developmental delay.

ranking the proportions were 76% and 38%. Part of this difference reflects the number of associated genes: the worst predicted ADRs had a median of three or fewer, so that many of them could not be tested at all.

At the level of individual gene sets, 163 KEGG pathways and 285 GO terms were significant for at least five ADRs. ADRs enriched for 25 of these KEGG pathways and 145 of these GO terms had a higher AUROC than the remaining ADRs, and no gene set was associated with a lower AUROC. The same direction was found at the group cut-offs: 5 KEGG pathways and 30 GO terms were more frequent among the 100 best than among the 100 worst predicted ADRs, and 20 and 90, respectively, in the top than in the bottom half; none was more frequent in a bottom group. With groups of 10, no individual gene set reached significance.

The gene sets more strongly associated with predictability describe neurotransmitter receptor signalling (Table 7). Neuroactive ligand signalling was enriched for 200 ADRs, including 57 of the 100 best predicted and 6 of the 100 worst predicted, and the serotonergic synapse for 154 ADRs (37 and 2). The leading GO terms were chemical synaptic transmission and serotonin, dopamine and histamine receptor signalling. Further gene sets associated with predictability included drug transporters (ABC transporters, folate transport) and, less strongly, cardiac muscle contraction and action potential.

4 Discussion

4.1 Predictability as a property of the ADR

Previous studies have shown that biological information improves ADR prediction [2, 7, 8, 11–14]. We asked a different question: why are some ADRs predicted more

Table 5 KEGG pathway and GO Biological Process enrichment of the associated target genes of the most and least predictable ADRs

ADR	Genes	KEGG	Leading KEGG pathway	GO	Leading GO Biological Process term
Most predictable ADRs					
Arteriospasm coronary	7	5	Neuroactive ligand signaling (7/104; 1.5×10^{-5})	6	Adenylate cyclase-inhibiting serotonin receptor signaling pathway (5/7; 6.9×10^{-8})
Parotid swelling	19	2	Neuroactive ligand signaling (9/104; 8.3×10^{-3})	1	G protein-coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger (7/36; 2.0×10^{-3})
Serotonin syndrome	24	4	Serotonergic synapse (12/56; 9.4×10^{-8})	5	G protein-coupled serotonin receptor signaling pathway (8/11; 5.9×10^{-9})
Protrusion tongue	25	0	None	22	Phospholipase C-activating dopamine receptor signaling pathway (5/6; 4.9×10^{-5})
Venous thromboembolism	34	63	Endometrial cancer (9/24; 2.0×10^{-6})	21	Epidermal growth factor receptor signaling pathway (6/15; 3.4×10^{-3})
Superinfection	10	2	ABC transporters (3/14; 2.9×10^{-2})	3	Leukotriene transport (3/6; 1.9×10^{-2})
Myelosuppression	69	20	Motor proteins (15/18; 1.1×10^{-13})	6	Generation of neurons (10/22; 5.1×10^{-4})
Skin striae	12	0	None	0	None
Tardive dyskinesia	47	3	Neuroactive ligand signaling (21/104; 7.9×10^{-8})	18	G protein-coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger (15/36; 3.5×10^{-9})
Drug tolerance	19	2	Neuroactive ligand signaling (9/104; 8.3×10^{-3})	2	G protein-coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger (6/36; 1.9×10^{-2})
Least predictable ADRs					
Xerophthalmia	3	11	Ascorbate and aldarate metabolism (3/11; 1.0×10^{-4})	16	Flavone metabolic process (3/9; 3.0×10^{-4})
Carcinoma of lung	22	15	Motor proteins (13/18; 5.4×10^{-18})	2	Generation of neurons (7/22; 1.5×10^{-4})

Note: Genes, number of associated target genes tested. KEGG and GO, number of significant terms (adjusted $p < 0.05$) among 246 KEGG pathways and 1,407 GO terms. For the leading term, the number of associated genes in the term is given over the number of background genes in the term, followed by the adjusted p-value. The other eight of the least predictable ADRs had fewer than three associated target genes and were not tested.

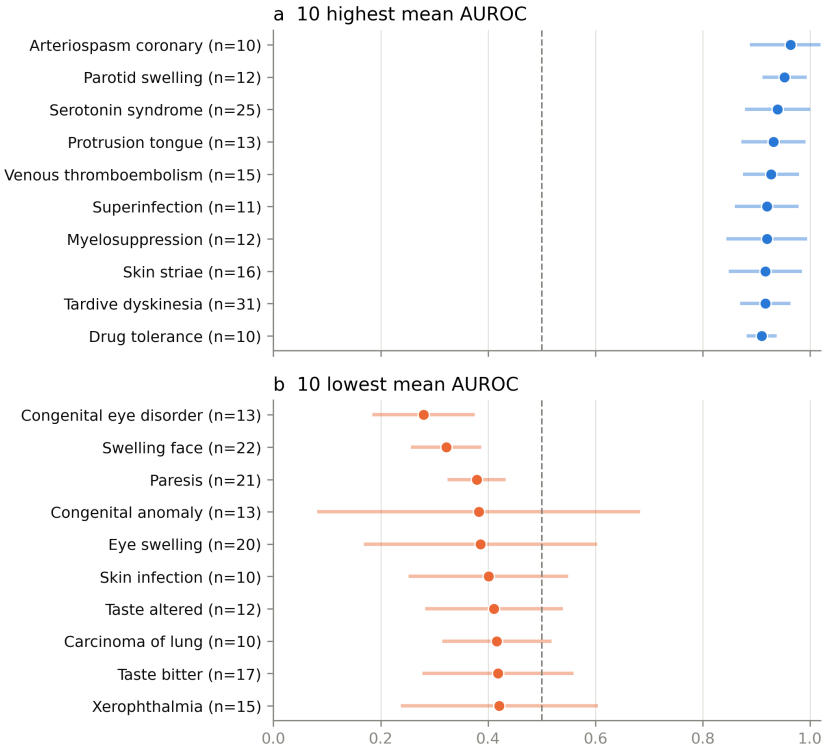

Fig. 5 The most and least predictable ADRs. **(a)** The 10 ADRs with the highest mean AUROC. **(b)** The 10 ADRs with the lowest mean AUROC. Points show the mean across five folds and bars the standard deviation; n is the number of drugs carrying the ADR and the dashed line marks chance

reliably than others? The purpose of using the harmonic model was therefore interpretability rather than state-of-the-art predictive performance; the predicted score of a drug is determined only by the labels of drugs with similar gene interaction profiles [22]. Overall performance was moderate (AUROC 0.670) but stable across folds and abolished when the network was shuffled. The comparison with k-NN, which uses the same target-similarity representation, was intended not to establish state-of-the-art performance but to test whether propagation extracts additional information from it; it did so for the ranking of drugs within each ADR, whereas the drug-level $F_{\max}$, dominated by common ADRs, was similar for both methods. Behind the average, predictability ranged widely between ADRs.

Measuring this variation requires care. AUPRC is the appropriate summary for the overall, highly imbalanced task [28], but its chance level equals the frequency of the ADR, and in our data a ranking by AUPRC mostly reflects frequency. Once a minimum number of drugs was required, AUROC was not materially associated with ADR frequency, and it was the basis of all comparisons between ADRs. Studies comparing prediction performance across side effects should report a measure that depends little on prevalence, together with a minimum number of positive drugs.

Table 6 Pathway enrichment among the best and worst predicted ADRs at three group sizes

Group	Median AUROC	Median genes	Testable (%)	KEGG (%)	GO (%)
Top 10	0.923	21.5	100	80	90
Bottom 10	0.394	0	20	20	20
Top 100	0.848	32.5	99	88	87
Bottom 100	0.526	3	54	26	25
Top half	0.744	25	95	76	77
Bottom half	0.615	6	70	38	42

Note: Median genes, median number of associated target genes. Testable, ADRs with at least three associated target genes. KEGG and GO, ADRs with at least one significant KEGG pathway or GO Biological Process term, as a percentage of all ADRs in the group. Each half contains 485 ADRs. For the comparison of top and bottom groups (Fisher's exact test), $p = 0.023$ (KEGG) and $p = 0.005$ (GO) for groups of 10, and $p < 0.001$ for all other comparisons, including the proportion testable at every group size.

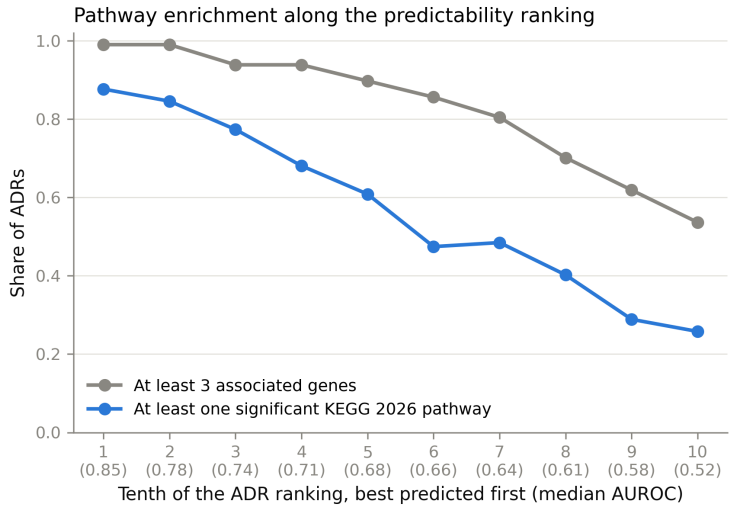

Fig. 6 Pathway enrichment along the predictability ranking. The 970 ranked ADRs are divided into tenths, from the best predicted to the worst predicted; the median AUROC of each tenth is given parentheses. Lines show the proportion of ADRs with at least three associated target genes and with at least one significant KEGG pathway

4.2 Coherent targets make an ADR

The central finding of this study is the strong relationship between ADR predictability and target coherence (Fig. 4). ADRs were predicted well when the drugs associated with them resemble one another in their gene interaction profiles more than they resemble other drugs, and this relationship was strong ($\rho = 0.73$). As target coherence is derived from the similarities that the model uses, this relationship measures how strongly the organisation of drugs in target space constrains predictability rather

Table 7 Gene sets most strongly associated with ADR predictability

Gene set	ADRs	$AUROC_{in}$	$AUROC_{out}$	100	Half
KEGG pathways					
Neuroactive ligand signaling	200	0.749	0.659	57/6[1]	159/41[1]
Neuroactive ligand-receptor interaction	120	0.752	0.664	35/6[1]	98/22[1]
Serotonergic synapse	154	0.738	0.663	37/2[1]	113/41[1]
Taste transduction[2]	140	0.727	0.664	24/4[1]	104/36[1]
Folate transport and metabolism	29	0.764	0.670	8/0	26/3[1]
Nicotine addiction	57	0.729	0.669	12/3	48/9[1]
ABC transporters	25	0.760	0.670	6/0	23/2[1]
Retrograde endocannabinoid signaling	64	0.727	0.668	14/3	53/11[1]
GO Biological Process terms					
Chemical synaptic transmission	141	0.767	0.661	44/4[1]	123/18[1]
Anterograde trans-synaptic signaling	172	0.751	0.660	47/2[1]	140/32[1]
G protein-coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger	144	0.751	0.663	44/3[1]	112/32[1]
G protein-coupled serotonin receptor signaling pathway	188	0.739	0.661	43/3[1]	137/51[1]
Response to histamine	84	0.766	0.666	30/3[1]	72/12[1]
Phospholipase C-activating dopamine receptor signaling pathway	59	0.777	0.668	23/0[1]	54/5[1]

Note: The eight KEGG pathways and six GO terms with the smallest p-values in the comparison of AUROC between ADRs enriched for the gene set and all other ranked ADRs (adjusted $p < 0.001$ for all). ADRs, number of ranked ADRs for which the gene set was significant. $AUROC_{in}$ and $AUROC_{out}$, median AUROC of these ADRs and of the remaining ADRs. The last two columns give the number of ADRs enriched for the gene set among the 100 best and 100 worst predicted ADRs (100) and in the top and bottom half of the ranking (Half), as top/bottom.

[1] Adjusted $p < 0.05$ for the difference between the top and bottom group.

[2] Significant through serotonin receptor genes that are annotated to this pathway.

than providing an independent predictor (Section 3.3). The most predictable ADRs were recognisable class effects: vasospasm with serotonin 5-HT_1 receptor agonists, extrapyramidal reactions with dopamine receptor antagonists, serotonin syndrome with serotonergic agents and skin striae with glucocorticoid receptor agonists. Many of the leading associated genes encode G protein-coupled receptors or nuclear receptors, target families in which many drugs act on the same few proteins [29–32]. Such shared targets place the drugs in the same neighbourhood of the network, where label propagation is effective [33–35].

The least predictable ADRs lacked this structure. They were reported for drugs from unrelated classes that shared few or no targets, and most had no gene over-represented among their drugs. Several are non-specific clinical terms, such as face

or eye swelling and taste disturbance, which can arise through many mechanisms or may reflect the treated disease or the reporting process rather than a pharmacological action. For these ADRs, target-based similarity carries no usable signal, and more complex models trained on the same information would be expected to face the same limit.

A shared annotation does not have to be mechanistic. For superinfection, the over-represented genes were transporters and metabolic genes of antibacterial agents, whose own targets are not human genes. Enzymes and transporters that handle chemically diverse drugs can therefore create similarity that predicts an ADR without explaining it [36–38], and highly connected genes of this kind should be interpreted with caution [39–41]. Likewise, where the shared genes belong to a large gene family, such as the tubulins, or to the kinase targets of anticancer drugs, many pathways become significant without identifying a mechanism for the reaction. The number of enriched pathways is therefore not a measure of biological coherence, and enrichment is informative only when it is tested against the genes that drugs can actually target.

Across the whole ranking, well predicted ADRs were far more likely to show pathway enrichment, whether the 10, the 100 or half of the ADRs at each end were compared. Poorly predicted ADRs were marked by the absence of enrichment rather than by a different biology. The pathways that distinguished predictable ADRs were dominated by serotonin, dopamine and histamine receptor signalling and synaptic transmission. In this dataset, therefore, a predictable ADR is to a large extent one associated with drugs that act on monoamine receptors, a densely annotated and heavily shared group of targets. Whether the same holds for target classes that are less well represented in drug–gene databases remains to be shown.

4.3 Limitations and future directions

The study has several limitations. Both data sources are curated databases and are subject to annotation bias: well-studied drugs have more recorded gene interactions and more recorded side effects, and the absence of an annotation does not establish the absence of an effect. Only 536 of 777 drugs passed the similarity threshold, and although the results were stable across thresholds, drugs with unique target profiles were not evaluated. ADRs with fewer than 10 drugs, which form the majority of MedDRA terms in that dataset, could not be ranked reliably. MedDRA terms overlap; identical drug profiles were merged, but closely related terms remain as separate entries. For the same reason, and because one drug contributes to many ADRs, the ADRs are not independent observations, and the p-values of the comparisons across ADRs, in particular those for individual pathways, are likely to be too small. The 36 ADRs with a mean AUROC below 0.5 suggest that harmonic scores are influenced by how well connected a drug is, which was not corrected for. Finally, the analysis used a single cross-validation partition and one type of similarity.

Future work should test whether the relationship between target coherence and predictability holds for other models and for similarities derived from independent data, such as tissue-specific or multi-omics networks [42–44], and should examine whether

pharmacokinetic and pharmacodynamic gene annotations contribute differently to prediction. Graph neural networks that retain the transparency of label propagation are a natural extension [45].

5 Conclusion

Most predictive models treat ADR prediction as a purely computational problem. This work instead used an interpretable network model to ask why some ADRs are consistently easier to predict than others. With harmonic propagation over a drug–gene similarity network, 76% of 970 ADRs were predicted better than chance, and their predictability, measured by AUROC, which showed little association with ADR frequency in our data, was closely tied to how coherent the targets of the associated drugs are. Class effects mediated by shared receptors were predicted with high accuracy and mapped to the corresponding signalling pathways, whereas non-specific reactions reported for pharmacologically unrelated drugs were predicted no better than chance and showed little or no pathway enrichment. Differences in predictability therefore reflects, at least in part, the biological organisation underlying the ADR rather than model performance alone. In practice, target coherence can be computed before model training and may help identify ADRs for which target-based predictions are likely to be more or less reliable.

Supplementary information. Supplementary Figs. S1–S4 show side–effect similarity against drug–gene similarity for all drug pairs, the drug similarity network with the training and testing drugs of one fold, AUROC against AUPRC for all side effects, and fold-level performance of harmonic propagation and the k-NN baseline. Supplementary Table S1 lists, for every KEGG pathway and GO term that was significant for at least five ADRs, its association with predictability.

Acknowledgements. The authors thank Dr Gregory Wheeler for help during the course of this research, the University of Galway for providing the computational resources and academic environment necessary to conduct this work, and the MSc programme staff for their support.

Declarations

- **Funding:** No funding was received for conducting this study.
- **Competing interests:** The authors declare that they have no competing interests.
- **Ethics approval and consent to participate:** Not applicable.
- **Consent for publication:** Not applicable.
- **Data availability:** The datasets used in this study are available at https://github.com/Mihir2728/Drug-Gene-SE-NM-Manuscript.git.
- **Materials availability:** Not applicable.
- **Code availability:** The analysis code (Python notebook) that reproduces the results and figures is available at https://github.com/Mihir2728/Drug-Gene-SE-NM-Manuscript.git. Note that we used AI to modify and reassemble the codes in Google Colab.

- **Author contribution:** M.S.C.: conceptualisation, data curation, methodology, formal analysis, investigation, visualisation and writing of the original draft. H.Y.: conceptualisation, supervision, methodology guidance and critical review.

References

[1] Pandit, A.A., Dubey, S.A.: A comprehensive review on adverse drug reactions (ADRs) detection and prediction models. In: 2021 13th International Conference on Computational Intelligence and Communication Networks (CICN), pp. 123–127. IEEE, Lima, Peru (2021). https://doi.org/10.1109/CICN51697.2021.9574639

[2] Funari, A., Fiscon, G., Paci, P.: Network medicine and systems pharmacology approaches to predicting adverse drug effects. British Journal of Pharmacology **183**(8), 1627–1639 (2026) https://doi.org/10.1111/bph.17330

[3] Basile, A.O., Yahi, A., Tatonetti, N.P.: Artificial intelligence for drug toxicity and safety. Trends in Pharmacological Sciences **40**(9), 624–635 (2019) https://doi.org/10.1016/j.tips.2019.07.005

[4] Zhong, Y., Seoighe, C., Yang, H.: Non-negative matrix factorization combined with kernel regression for the prediction of adverse drug reaction profiles. Bioinformatics Advances **4**(1), vbae009 (2024) https://doi.org/10.1093/bioadv/vbae009

[5] Shi, W., Yang, H., Xie, L., Yin, X.-X., Zhang, Y.: A review of machine learning-based methods for predicting drug–target interactions. Health Information Science and Systems **12**(1), 30 (2024) https://doi.org/10.1007/s13755-024-00287-6

[6] Seo, S., Lee, T., Kim, M.-h., Yoon, Y.: Prediction of side effects using comprehensive similarity measures. BioMed Research International **2020**(1), 1357630 (2020) https://doi.org/10.1155/2020/1357630

[7] Yadav, M., Ahlawat, P., Singh, V.: A graph-based transformer neural network for multi-label ADR prediction. Arabian Journal for Science and Engineering **50**(10), 7323–7341 (2025) https://doi.org/10.1007/s13369-024-09342-6

[8] Uner, O.C., Kuru, H.I., Cinbis, R.G., Tastan, O., Cicek, A.E.: DeepSide: a deep learning approach for drug side effect prediction. IEEE/ACM Transactions on Computational Biology and Bioinformatics **20**(1), 330–339 (2023) https://doi.org/10.1109/TCBB.2022.3141103

[9] Sharmin, A., Mahmud, B.U.: Multi-relational knowledge graph for drug repurposing and side-effect burden prediction using gene–drug–disease associations. BioChem **6**(2), 13 (2026) https://doi.org/10.3390/biochem6020013

[10] Pancino, N., Perron, Y., Bongini, P., Scarselli, F.: Drug side effect prediction with

deep learning molecular embedding in a graph-of-graphs domain. Mathematics **10**(23), 4550 (2022) https://doi.org/10.3390/math10234550

[11] Huang, T., Lin, K.-H., Machado-Vieira, R., Soares, J.C., Jiang, X., Kim, Y.: Explainable drug side effect prediction via biologically informed graph neural network. Preprint at https://doi.org/10.1101/2023.05.26.23290615 (2023)

[12] Gao, Y., Zhang, X., Sun, Z., Chandak, P., Bu, J., Wang, H.: Precision adverse drug reactions prediction with heterogeneous graph neural network. Advanced Science **12**(4), 2404671 (2025) https://doi.org/10.1002/advs.202404671

[13] Chrysolite, R.S.P., Sushmethaa, M., Muthulakshmi, V.: Graph neural network-based prediction of drug side effects via gene expression and pathway interaction modeling. In: 2025 2nd International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), pp. 1–6. IEEE, Chennai, India (2025). https://doi.org/10.1109/ICECONF65644.2025.11379573

[14] Kim, E., Nam, H.: DeSIDE-DDI: interpretable prediction of drug–drug interactions using drug-induced gene expressions. Journal of Cheminformatics **14**(1), 9 (2022) https://doi.org/10.1186/s13321-022-00589-5

[15] Carvajal Rico, J., Alaeddini, A., Faruqui, S.H.A., Fisher-Hoch, S.P., Mccormick, J.B.: A Laplacian regularized graph neural network for predictive modeling of multiple chronic conditions. Computer Methods and Programs in Biomedicine **247**, 108058 (2024) https://doi.org/10.1016/j.cmpb.2024.108058

[16] Beam, C.: Resolving power: a general approach to compare the distinguishing ability of threshold-free evaluation metrics. Machine Learning **114**(1), 9 (2025) https://doi.org/10.1007/s10994-024-06723-8

[17] Zhang, R., Wen, H., Lin, Z., Li, B., Zhou, X.: Artificial intelligence-driven drug toxicity prediction: advances, challenges, and future directions. Toxics **13**(7), 525 (2025) https://doi.org/10.3390/toxics13070525

[18] Qian, S., Liang, S., Yu, H.: Leveraging genetic interactions for adverse drug-drug interaction prediction. PLOS Computational Biology **15**(5), e1007068 (2019) https://doi.org/10.1371/journal.pcbi.1007068

[19] Park, J., Lee, S., Kim, K., Jung, J., Lee, D.: Large-scale prediction of adverse drug reactions-related proteins with network embedding. Bioinformatics **39**(1), btac843 (2023) https://doi.org/10.1093/bioinformatics/btac843

[20] Nguyen, P.A., Born, D.A., Deaton, A.M., Nioi, P., Ward, L.D.: Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. Nature Communications **10**(1), 1579 (2019) https://doi.org/10.1038/s41467-019-09407-3

[21] Lin, S.-F., Xiao, K.-T., Huang, Y.-T., Chiu, C.-C., Soo, V.-W.: Analysis of adverse drug reactions using drug and drug target interactions and graph-based methods. Artificial Intelligence in Medicine **48**(2-3), 161–166 (2010) https://doi.org/10.1016/j.artmed.2009.11.002

[22] Zhu, X., Ghahramani, Z., Lafferty, J.: Semi-supervised learning using Gaussian fields and harmonic functions. In: Proceedings of the 20th International Conference on Machine Learning (ICML-03), pp. 912–919 (2003)

[23] Xie, M., Lei, X., Zhong, J., Ouyang, J., Li, G.: Drug response prediction using graph representation learning and Laplacian feature selection. BMC Bioinformatics **23**(S8), 532 (2022) https://doi.org/10.1186/s12859-022-05080-4

[24] Jaccard, P.: The distribution of the flora in the alpine zone. New Phytologist **11**(2), 37–50 (1912) https://doi.org/10.1111/j.1469-8137.1912.tb05611.x

[25] Rogers, D.J., Tanimoto, T.T.: A computer program for classifying plants. Science **132**(3434), 1115–1118 (1960) https://doi.org/10.1126/science.132.3434.1115

[26] Radivojac, P., Clark, W.T., Oron, T.R., Schnoes, A.M., Wittkop, T., Sokolov, A., Graim, K., Funk, C., Verspoor, K., Ben-Hur, A., Pandey, G., *et al.*: A large-scale evaluation of computational protein function prediction. Nature Methods **10**(3), 221–227 (2013) https://doi.org/10.1038/nmeth.2340

[27] Torres, M., Yang, H., Romero, A.E., Paccanaro, A.: Protein function prediction for newly sequenced organisms. Nature Machine Intelligence **3**(12), 1050–1060 (2021)

[28] Saito, T., Rehmsmeier, M.: The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE **10**(3), e0118432 (2015) https://doi.org/10.1371/journal.pone.0118432

[29] Yıldırım, M.A., Goh, K.-I., Cusick, M.E., Barabási, A.-L., Vidal, M.: Drug–target network. Nature Biotechnology **25**(10), 1119–1126 (2007) https://doi.org/10.1038/nbt1338

[30] Hopkins, A.L.: Network pharmacology: the next paradigm in drug discovery. Nature Chemical Biology **4**(11), 682–690 (2008) https://doi.org/10.1038/nchembio.118

[31] Hauser, A.S., Attwood, M.M., Rask-Andersen, M., Schiöth, H.B., Gloriam, D.E.: Trends in GPCR drug discovery: new agents, targets and indications. Nature Reviews Drug Discovery **16**(12), 829–842 (2017) https://doi.org/10.1038/nrd.2017.178

[32] Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., Evans, R.M.: The nuclear

receptor superfamily: the second decade. Cell **83**(6), 835–839 (1995) https://doi.org/10.1016/0092-8674(95)90199-X

[33] Zhou, D., Bousquet, O., Lal, T.N., Weston, J., Schölkopf, B.: Learning with local and global consistency. In: Advances in Neural Information Processing Systems 16, pp. 321–328 (2004)

[34] Barabási, A.-L., Gulbahce, N., Loscalzo, J.: Network medicine: a network-based approach to human disease. Nature Reviews Genetics **12**(1), 56–68 (2011) https://doi.org/10.1038/nrg2918

[35] Guney, E., Menche, J., Vidal, M., Barábasi, A.-L.: Network-based in silico drug efficacy screening. Nature Communications **7**(1), 10331 (2016) https://doi.org/10.1038/ncomms10331

[36] Wilkinson, G.R.: Drug metabolism and variability among patients in drug response. New England Journal of Medicine **352**(21), 2211–2221 (2005) https://doi.org/10.1056/NEJMra032424

[37] Ingelman-Sundberg, M.: Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. Trends in Pharmacological Sciences **25**(4), 193–200 (2004) https://doi.org/10.1016/j.tips.2004.02.007

[38] Zanger, U.M., Schwab, M.: Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacology & Therapeutics **138**(1), 103–141 (2013) https://doi.org/10.1016/j.pharmthera.2012.12.007

[39] Jeong, H., Mason, S.P., Barabási, A.-L., Oltvai, Z.N.: Lethality and centrality in protein networks. Nature **411**(6833), 41–42 (2001) https://doi.org/10.1038/35075138

[40] Barabási, A.-L., Oltvai, Z.N.: Network biology: understanding the cell's functional organization. Nature Reviews Genetics **5**(2), 101–113 (2004) https://doi.org/10.1038/nrg1272

[41] Goh, K.-I., Cusick, M.E., Valle, D., Childs, B., Vidal, M., Barabási, A.-L.: The human disease network. Proceedings of the National Academy of Sciences **104**(21), 8685–8690 (2007) https://doi.org/10.1073/pnas.0701361104

[42] Lage, K., Karlberg, E.O., Størling, Z.M., Ólason, P.Í., Pedersen, A.G., Rigina, O., Hinsby, A.M., Tümer, Z., Pociot, F., Tommerup, N., Moreau, Y., Brunak, S.: A human phenome-interactome network of protein complexes implicated in genetic disorders. Nature Biotechnology **25**(3), 309–316 (2007) https://doi.org/10.1038/nbt1295

[43] Greene, C.S., Krishnan, A., Wong, A.K., Ricciotti, E., Zelaya, R.A., Himmelstein,

D.S., Zhang, R., Hartmann, B.M., Zaslavsky, E., Sealfon, S.C., Chasman, D.I., FitzGerald, G.A., Dolinski, K., Grosser, T., Troyanskaya, O.G.: Understanding multicellular function and disease with human tissue-specific networks. Nature Genetics **47**(6), 569–576 (2015) https://doi.org/10.1038/ng.3259

[44] Hasin, Y., Seldin, M., Lusis, A.: Multi-omics approaches to disease. Genome Biology **18**(1), 83 (2017) https://doi.org/10.1186/s13059-017-1215-1

[45] Zitnik, M., Agrawal, M., Leskovec, J.: Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics **34**(13), i457–i466 (2018) https://doi.org/10.1093/bioinformatics/bty294