

Sparsity is all you need: rethinking biologically informed neural networks

Isabella Caranzano¹, Corrado Pancotti², Cesare Rollo^{3,4}, Flavio Sartori¹, Pietro Liò⁵, Piero Fariselli^{1,*}, Tiziana Sanavia^{1,*}

¹AI and Computational Biomedicine Unit, Department of Medical Sciences, University of Torino, Via Santena 19, 10126 Torino, Italy

²Helmholtz Zentrum München, Helmholtz AI Central Unit, Ingolstädter Landstraße 1, 85764 Oberschleißheim-Neuherberg, Germany

³Department of Computer Science, University of Copenhagen, Universitetsparken 1, 2100 Copenhagen, Denmark

⁴Center for Health Data Science, Department of Public Health, University of Copenhagen, Øster Søgade 16B, 10.1, 1355 Copenhagen, Denmark

⁵Department of Computer Science and Technology, University of Cambridge, William Gates Building, 15 JJ Thomson Ave, CB3 0FD Cambridge, United Kingdom

*Corresponding authors. Piero Fariselli, E-mail: piero.fariselli@unito.it; Tiziana Sanavia, AI and Computational Biomedicine Unit, Department of Medical Sciences, University of Torino, Via Santena 19, 10126 Torino, Italy. E-mail: tiziana.sanavia@unito.it.

Abstract

Biologically informed neural networks are increasingly adopted in bioinformatics under the premise that embedding biological knowledge into model architectures yields more accurate and interpretable predictions. This approach has driven a growing literature of pathway-informed models aiming to move beyond black-box learning by explicitly encoding biological structure. However, it remains unclear whether these models exploit biological knowledge or instead benefit from a different inductive bias. Here, we systematically investigate this question across 29 state-of-the-art pathway-informed neural networks by explicitly decoupling biological annotations from network architecture. For each evaluable model, we implement a structure-matched randomization protocol, in which pathway annotations are replaced with random associations while preserving sparsity and architectural constraints, allowing for a direct comparison under controlled conditions. Across multiple prediction tasks, datasets, and evaluation metrics, the randomized models consistently match or outperform their biologically informed counterparts. Moreover, pathway-informed models show no systematic advantage in interpretability: randomized models recover disease-associated biomarkers with comparable accuracy and yield highly correlated feature rankings. Our results reveal that the performance gains commonly attributed to biological pathway integration arise predominantly from sparsity-induced regularization rather than from biological knowledge itself. We provide a general evaluation workflow to test whether biological priors contribute predictive information beyond sparsity, offering practical guidance for the development of biology-aware neural networks. The code implementing the proposed methodology is available on GitHub at https://github.com/comphiomed-unito/Pathway_Randomization.

Keywords biologically informed neural networks, pathway-based deep learning, model randomization analysis, multi-omics integration, cancer genomics

Introduction

Genes and molecular entities organize into pathways, biological processes, and regulatory mechanisms. This structured knowledge can guide neural network design in bioinformatics and biomedical applications. Encoding this structure directly into neural network architectures appears to offer a way to move beyond black-box prediction, combining strong predictive performance with built-in biological interpretability [1]. This premise has driven the rapid adoption of biologically informed neural networks (BINNs), in which curated annotations from databases such as Kyoto Encyclopedia of Genes and Genomes (KEGG) [2] and Reactome [3] constrain model architectures or input representations. Across multi-omics applications, BINNs are increasingly used for classification, regression, and survival analysis,

often with the explicit claim that embedding biological knowledge improves both generalization and interpretability [4].

Early approaches introduced a single pathway layer into multilayer perceptrons [5–8], combined with sparse coding mechanisms such as dropout and gene–pathway pruning based on curated annotations [9–14]. Subsequent works extended this paradigm by incorporating biological information across multiple layers, modeling pathway interactions, or representing pathways as independent subnetworks [15–19]. More recent models incorporate attention mechanisms, transformers or variational autoencoders to further increase representational capacity while retaining pathway structure [20–22]. Across these designs, pathway annotations shape network topology, ensuring that functionally related entities share connections

Received: April 10, 2026. **Revised:** June 14, 2026. **Accepted:** July 7, 2026

© The Author(s) 2026. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

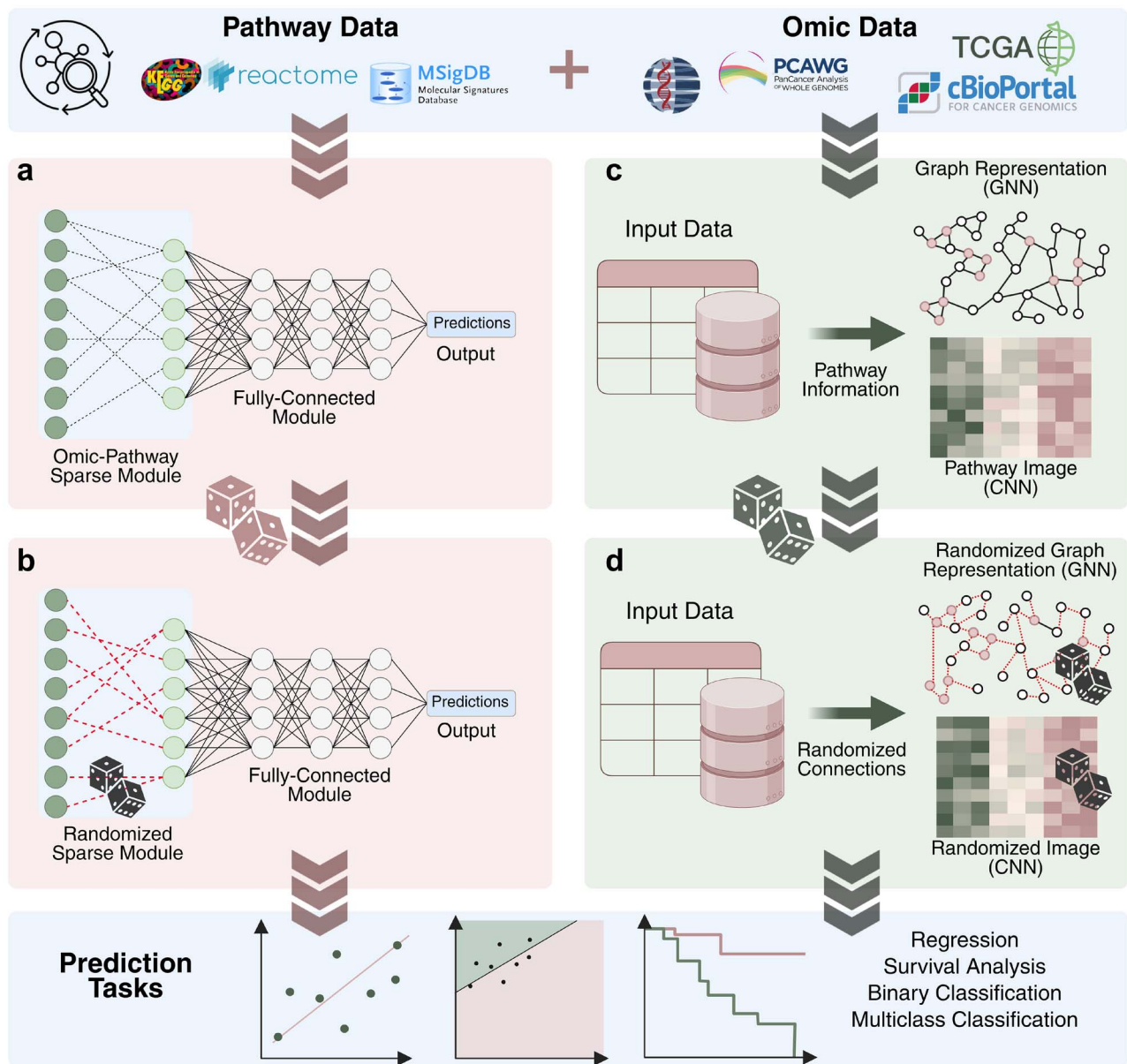


Figure 1. Schematic representation of pathway integration approaches in neural networks for omics data and their relative randomization. Pathway information can be incorporated in two ways (panels a and c): (a) A neural network utilizing pathway information by enforcing structured connections, introducing sparsity in the model. (b) A structure-matched randomized counterpart in which connections are assigned independently of pathway annotations while preserving network sparsity. (c) A data transformation strategy that incorporates pathway information to convert tabular omics data into graphs or images. (d) A randomized data transformation approach that generates graphs or images through a randomization procedure rather than predefined pathway structures.

while pruning interactions based on curated knowledge. Alternative strategies exploit pathway information through data transformation, enabling architectures designed for non-tabular data. Graph neural networks (GNNs) represent genes or pathways as nodes connected according to pathway-specific relationships [23, 24], or model pathway–pathway interactions via graph convolutional or attention-based layers [25–29]. Hybrid approaches combine GNNs with pathway aggregation modules [30], or integrate interaction networks through shared GNN layers and meta-graph representations for downstream prediction [31] (a schematic representation is shown in Fig. 1).

Despite their diversity, these methods share a key feature: pathway annotations act primarily as a sparsity prior, reducing the number of trainable parameters and constraining the space of representable functions. From a learning-theoretic perspective, this design reflects compositional sparsity, where complex functions are decomposed into interactions over limited input subsets. This principle underlies the success of convolutional, graph, and transformer neural networks, which leverage structured sparsity to mitigate the curse of dimensionality [32–36].

BINNs can therefore be viewed as a specific representation of this principle. By grouping genes into pathways or organizing pathways

into hierarchical modules, structured sparsity is imposed on feature interactions, using biological knowledge as a mechanism for implementing a compositional structure [37–40]. However, architectural constraints shape the hypothesis space without uniquely determining internal representations. As a result, these internal representations may deviate from their intended biological meaning, even when predictions remain accurate. This raises a fundamental and largely unexplored question: do BINNs exploit biological knowledge beyond the sparsity they impose, or do they succeed primarily because biological annotations impose structured sparsity, regardless of their biological meaning? Answering this requires more than comparing BINNs to their fully connected baselines. It requires isolating biological annotations from architectural effects by introducing appropriate null models that preserve sparsity and structure while removing biological meaning.

In this work, we systematically evaluate 29 state-of-the-art pathway-informed neural networks by explicitly decoupling biological annotations from architectural structure. For each model for which an experimental evaluation was feasible, we built a randomized counterpart in which pathway annotations were replaced with random associations that preserved sparsity and structural constraints. This framework allows us to test a fundamental hypothesis: if biological pathway knowledge contributes predictive information beyond sparsity, then disrupting that knowledge while preserving structure should degrade performance and interpretability. Conversely, if the randomized models perform comparably, the contribution of biological annotations must be reconsidered. Finally, we propose a structured evaluation protocol, in the form of a workflow, to assess when and whether the integration of pathway information is beneficial. This methodology is applicable across different BINN architectures, prediction tasks, and data modalities, providing a robust benchmark for systematically comparing biologically informed models against their structure-matched randomized counterparts.

Results

State-of-the-art pathway-informed approaches in deep learning

We reviewed state-of-the-art pathway-informed neural network studies and examined their methods, assumptions, and reported results. We then applied the proposed evaluation protocol to systematically compare each model with its structure-matched randomized counterpart. We selected models with publicly available code to ensure reproducible comparison. Table 1 summarizes recent BINNs that integrate pathway annotations either into the model architecture or into the organization of the input data. The models included at each analysis level are described in detail in Supplementary Table S1. As shown in Fig. 2, models address tasks ranging from classification to survival analysis and regression. Tables S2–S5 in Supplementary Data summarize feature space and sample statistics. Most models rely on Reactome [3] and KEGG [2], with additional resources including PID [41], BioCarta [42], MSigDB [43], GO BP [44], ConsensusPathDB [45], WikiPathways [46], and Enrichr-KG [47]. Some models exploit gene–pathway associations as priors, while others include gene–gene and pathway–pathway interactions.

Benchmarking biological priors against structure-matched null models

For each method reported in Table 1, we compared the original pathway-informed model with a structure-matched randomized counterpart in which biological priors were replaced by random associations while preserving network sparsity and architectural constraints. Twenty independent train/test splits were evaluated for each model.

Across all prediction tasks, randomized models achieved performance statistically indistinguishable from biologically informed counterparts (Fig. 3; Table 2). In addition, for models such as MPVNN, BINN, pathDNN, and APNet, randomized versions significantly outperformed the original pathway-informed models (paired Wilcoxon signed-rank test P -values $< .01$).

These findings were consistent across evaluation metrics (AUC, accuracy, C-index, and F1 score), datasets, and repeated randomization trials (Supplementary Fig. S1). Moreover, no significant differences were observed between the performance distribution obtained from a single randomization and those obtained across 30 independent randomizations (Kolmogorov–Smirnov test, all $P > .05$), indicating that the results are not driven by favorable random seeds.

Pathway-informed neural networks differ substantially in computational cost. Across the evaluated models, training times ranged from seconds to multiple days per run, with the most complex architectures, such as graph-based and transformer-based models, requiring orders of magnitude more computation (Table 2).

The role of sparsity as a structural prior in BINNs

The previous analysis showed that biologically inspired neural networks perform equivalently or worse than randomized counterparts. By construction, these randomized networks preserved exactly the same level of sparsity found in their biologically informed counterparts.

We then investigated whether the level of sparsity introduced by the pathway annotations might be optimal for model performance. To this end, we compared randomized neural networks at different sparsity levels around those induced by the biological pathway annotations.

Tables S6–S9 in Supplementary Data illustrate pathway-induced sparsity levels exploited by the original implementation of each biologically informed model. Overall, the pathway-induced sparsity ranged between 59.2% observed with the miRNA-Pathway information used by DeepKEGG (Supplementary Table S8) and 99.99% in the pathway–pathway network exploited by P-NET (Supplementary Table S9).

In Fig. 4, we report the results for the five neural networks that can be feasibly trained and tested under different conditions: BINN, DeepKEGG, PASNet, PathCNN, and PINNet. Considering the biological information exploited by these tools, most models operated at high sparsity levels. Excluding the application of DeepKEGG to miRNA-based pathway annotations (achieving a sparsity between 59.2% and 66.9% across different tumors, as shown in Supplementary Table S8), the other methods employed networks constrained by sparsity levels ranging between 96.8% (BINN) and 99.5% (PINNet). Comparing the original models with randomized counterparts across different sparsity levels, we found that the sparsity levels induced by pathway annotations yielded performance comparable to, or significantly

Table 1 Overview of deep learning models integrating pathway information for various prediction tasks. Models are categorized by publication year, journal, prediction task, pathway source, code availability, data availability, and input data type.

Model	Year	Journal	Prediction task	Pathway Source	Code	Data	Data Type
PASNet [9]	2018	BMC Bioinformatics	Binary classification long-term VS Short-term survival	Reactozme	✓	✓	Gene expression
CoxPASNet [10]	2018	IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2018	Survival analysis	Reactome	✓	✓	Gene expression
MiNet [11]	2019	ISBRA 2019	Survival analysis	Reactome	✓	✓	Gene expression, CNV, DNA methylation
pathDNN [12]	2020	Journal of Chemical Information and Modeling	Drug sensitivity prediction	KEGG	✓	✓	Gene expression, drug targets
Multi-scale NN [5]	2020	Plos one	Prediction of disease, pathway, and gene associations	Reactome	✓	✓	Gene expression
GCN-MAE [23]	2020	Bioinformatics	Cancer subtype classification	KEGG	Code not available	X	Gene expression
P-NET [15]	2021	Nature	Cancer state prediction	Reactome	✓	✓	Mutations, CNA
PathCNN [48]	2021	Bioinformatics	Binary classification long-term versus Short-term survival	KEGG	✓	NB: Only processed data	Gene expression, DNA methylation, CNV
PathDeep [6]	2021	International Journal of Molecular Sciences	Classification cancer versus normal tissue	MSigDB	✓	NB: Only toy dataset available	Gene expression
PathGNN [24]	2022	BMC Bioinformatics	Binary classification long-term versus Short-term survival	Reactome	✓	✓	Gene expression, clinical data
MPVNN [7]	2022	Bioinformatics	Survival analysis	Unknown	✓	✓	Gene expression
GCS-Net [16]	2022	Journal of Oncology	Binary classification long-term versus Short-term survival	Reactome	Code not available	X	CNV, Somatic mutations, clinical data
REDDA [29]	2022	Computers in Biology and Medicine	Drug-disease association prediction	KEGG	✓	✓	Drugs Informations, Proteomics, Gene Expression
ReGeNNe [49]	2023	Bioinformatics	Classification (kidney stage, kidney vs liver, binary survival for ovarian)	PID, BioCarta, Reactome	✓	X	Gene expression
BINN [17]	2023	Nature Communications	Phenotypes classification	Reactome	✓	✓	Proteomic Data
PINNet [18]	2023	Frontiers in Aging Neuroscience	Alzheimer disease classification	KEGG, GO BP	✓	✓	Gene expression
PGLCN [25]	2023	Computational and Structural Biotechnology Journal	Tumor mutation burden prediction	Reactome	X	Github with empty files, not usable	Gene expression, CNV, Methylation
PathExpSurv [13]	2023	BMC Bioinformatics	Survival analysis	KEGG	✓	✓	Gene Expression
EMGNN [31]	2023	Bioinformatics	Cancer gene prediction	Consensus PathDB	✓	X	Mutations, CNA, Methylation, Gene expression
DeepKEGG [20]	2024	Briefings in Bioinformatics	Cancer recurrence prediction - Binary classification	KEGG	✓	✓	mRNA expression, SNV, miRNA
GraphPath [28]	2024	Bioinformatics	Cancer status classification	KEGG	✓	✓	CNA, Mutation
Pathformer [21]	2024	Bioinformatics	Disease diagnosis and prognosis	KEGG, PID, Reactome, BioCarta	✓	✓	Gene expression (or multimodal)
Autosurv [22]	2024	Precision Oncology	Survival analysis	Reactome	✓	✓	Gene expression, miRNA
CRESCENT [26]	2024	IEEE Journal of Biomedical and Health Informatics	Cancer survival analysis	Consensus-PathDB	Code not available	X	Gene expression

(continued)

Table 1 Continued.

Model	Year	Journal	Prediction task	Pathway Source	Code	Data	Data Type
DeepBINN [19]	2024	2024 11th IEEE Swiss Conference on Data Science (SDS)	Septic acute kidney injury phenotype classification	MSigDB	X	X	Proteomics
Multilevel-GNN [30]	2024	Briefings in Bioinformatics	Tumor risk prediction	KEGG	✓	✓	Gene expression, CNV, Methylation, Clinical data
APNet [14]	2025	Bioinformatics	Prediction of COVID-19 severity	Enrichr-KG, KEGG,ref3, GO BP, WikiPathways 2021	✓	✓	Bulk plasma proteomics, Single-cell RNA sequencing
HallmarkGraph [27]	2025	Bioinformatics	Hierarchical tumor subtypes classification	Reactome	✓	✓	Gene expression
KNET [8]	2025	Computational Biology and Chemistry	Drug response prediction	KEGG	Code not available	X	Gene expression, Mutations, CNV

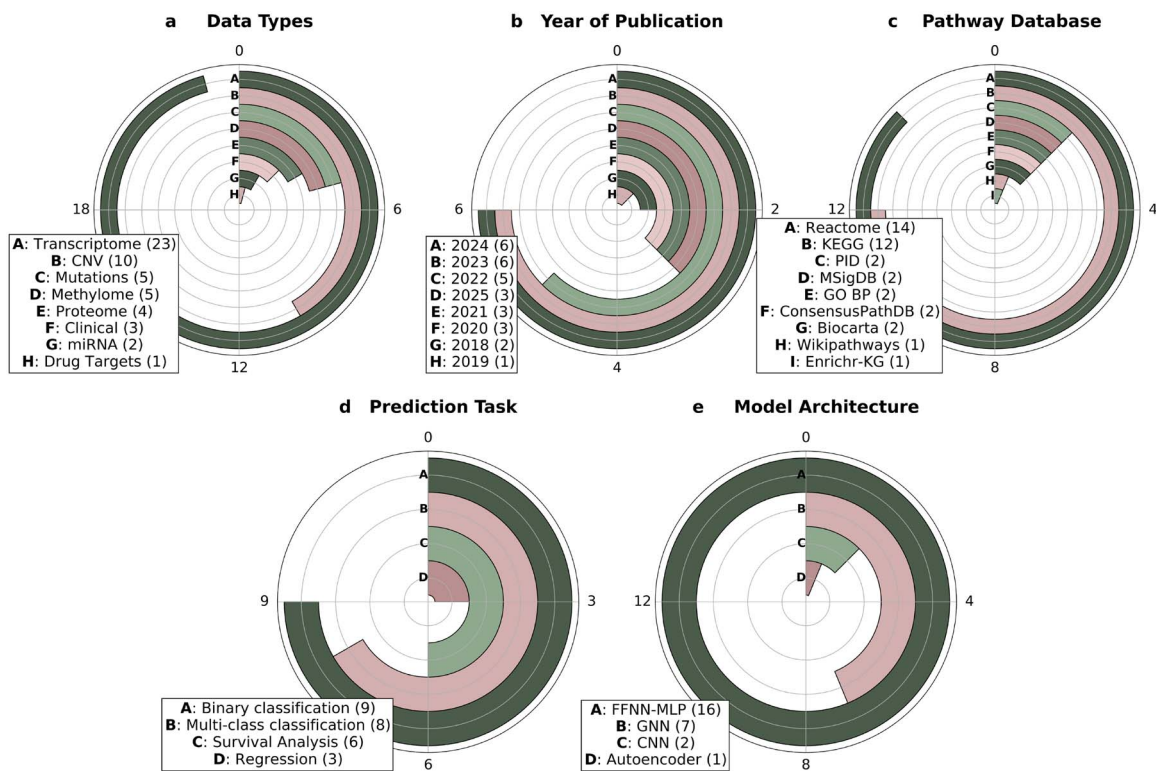


Figure 2. Circular bar plots summarizing characteristics of deep learning models that integrate pathway information. The plots show distributions for (a) input data types used, (b) publication year, (c) pathway database sources, (d) prediction tasks, and (e) model architectures (FFNN-MLP: feed-forward neural network - multi-layer perceptron, GNN: graph neural network, CNN: convolutional neural network, AE: autoencoders). Each segment's length corresponds to the count of models within each category.

worse than, the best-performing sparsity levels identified using randomized networks across the 60%–99% sparsity range. Statistically significant improvements favoring sparsity induced by randomization were observed only for BINN and DeepKEGG (maximum P -value 1.4×10^{-7}). These findings suggest that the sparsity induced by biological annotations was suboptimal compared to alternative sparsity levels achieved through randomization, indicating that pathway annotations primarily act as a hard-coded sparsity prior and that its biological origin does not necessarily provide an optimal inductive

bias. To complement the sparsity-level analysis, we further investigated whether the number of pathways included in the model affected predictive performance (see Supplementary Fig. S2). Specifically, we evaluated the same five representative models while progressively varying the fraction of retained pathways from 0.10 to 1.00. For each pathway fraction, pathway-informed models were compared with structure-matched randomized counterparts that preserved the same number of pathways and structural constraints. Overall, increasing the number of pathways did not lead to a consistent improvement in

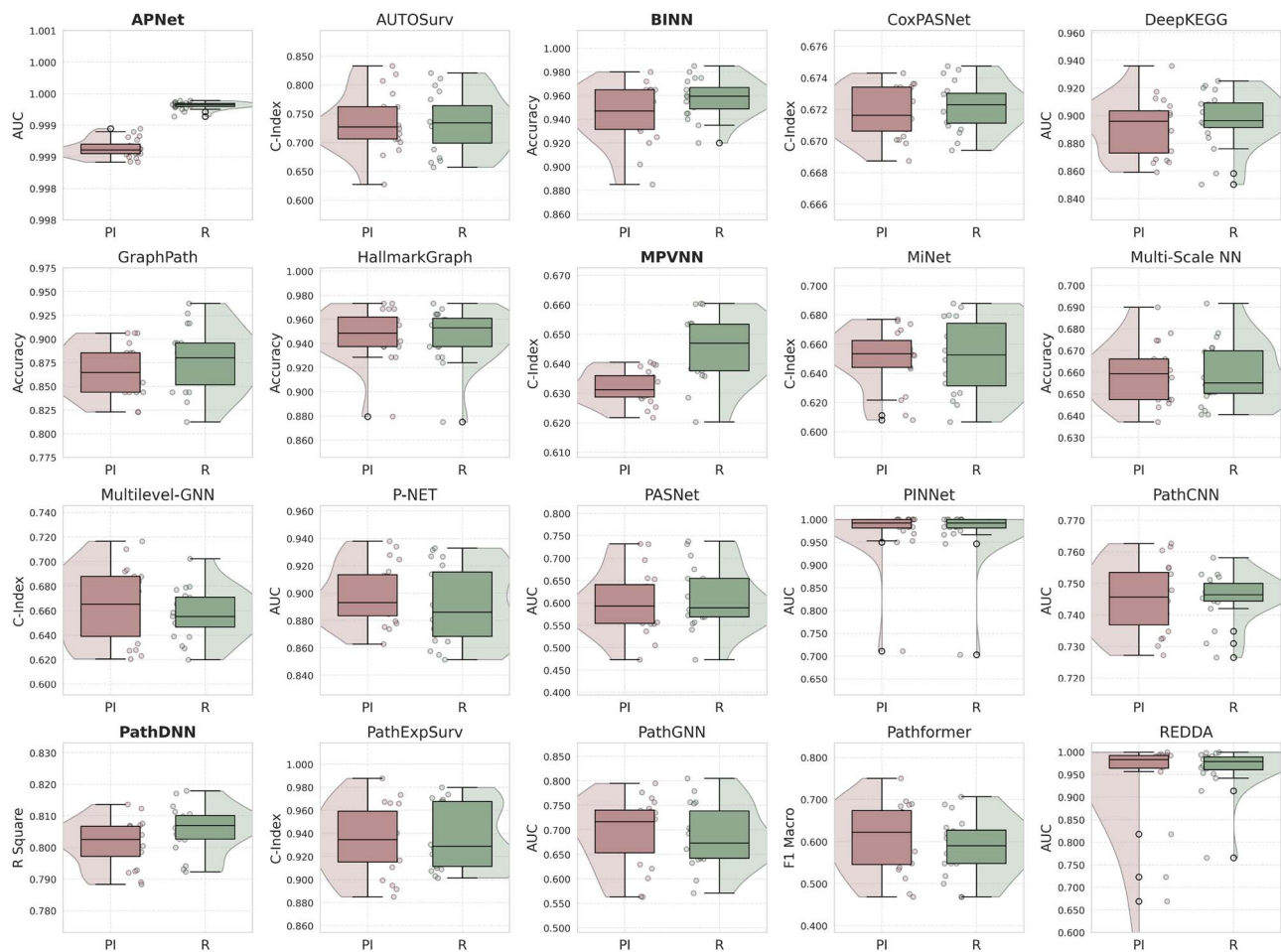


Figure 3. Model performance comparison across accuracy, AUC, C-index, F1 macro, and R-square metrics using violin plots. Models are grouped as pathway-informed (pink) and Randomized (green). The width reflects the distribution of scores, with central lines for median values and box plots indicating interquartile ranges. Models for which randomized versions significantly outperform pathway-informed versions are bolded in the x-axis labels. The results for the MPVNN, PathExpSurv, and DeepKEGG models represent average outcomes across different tumor types considered (detailed findings for each specific tumor type are provided in the [Supplementary Data](#)).

predictive performance. In most cases, performance reached a plateau before the full set of pathways was included, and randomized models remained comparable to their pathway-informed counterparts.

Comparison of biological information extracted by pathway-informed models and their randomized counterparts

A key motivation for BINNs is biological interpretability. To assess whether biological priors improve the identification of disease-relevant features or whether the interpretability can instead arise independently of biological annotations, we examined whether structure-matched randomized models are still able to identify disease-relevant biomarkers. We focused on four representative pathway-informed architectures (PINNet, DeepKEGG, BINN, and PASNet) for which feature-level interpretability analyses are feasible and comparable. In the original studies, interpretability was evaluated using heterogeneous and sometimes non-reproducible criteria, often without providing a consistent ranking of features. To enable a fair comparison with the randomized counterparts, which cannot support pathway-level interpretation by construction, we adopted feature importance as a common interpretability proxy. This

approach allows us to test whether biological pathway information is necessary to prioritize disease-associated features over unrelated ones. Disease-feature associations were obtained from widely used curated resources, including DisGeNet [50] and GeDiPNet [51] (see [Fig. 5](#) caption for the model-specific disease contexts and resources: Sepsis/BINN, Liver Hepatocellular Carcinoma (LIHC)/DeepKEGG, Glioblastoma Multiforme (GBM)/PASNet, and Alzheimer's Disease (AD)/PINNet).

In BINN, PASNet, and PINNet, disease-associated features received significantly higher importance scores in the randomized version, whereas this distinction was not significant in the corresponding pathway-informed versions of BINN and PASNet ([Fig. 5](#)). For PINNet, significant separation between AD-related and unrelated genes was observed in both the pathway-informed and randomized settings, consistent with the original study and disease–gene annotations from the AlzGene database [53]. Finally, for DeepKEGG, using the same attribution procedure as in the original work, neither the pathway-informed nor the randomized model showed a significant difference in importance scores between LIHC-related and unrelated features. Restricting the analysis to the top 100 ranked features, both models identified a comparable number of LIHC-associated genes (11 and 13, respectively). Moreover, feature rankings between the pathway-informed and randomized models were moderately to strongly

Table 2 Table summarizing the performance comparison between *pathway-informed* and *randomized* versions of various deep learning models across different evaluation metrics. Each model's performance is reported in terms of its specific metric (e.g. AUC, C-index, accuracy, R-squared), alongside the corresponding mean $\pm$ SD values. The table also includes the *execution time* for each model, with a legend denoting the time required for 20 runs, categorized as follows: + represents seconds, ++ represents minutes, +++ represents hours, and ++++ represents days. For certain models, the performance is further divided into *Omic-Pathway Network (OP)*, *Pathway-Pathway Network (PP)*, or a combination of both (*OP + PP*), to reflect the different configurations evaluated. Bolded values indicate cases where the randomized version outperformed the pathway-informed version. Results for MPVNN, PathExpSurv, and DeepKEGG correspond to averages across the different tumor types evaluated (detailed results are provided in Supplementary Tables S12–S14).

Model	Metric	Pathway-informed model	Randomized model	Execution time
PASNet	AUC	0.600 $\pm$ 0.067	0.608 $\pm$ 0.065	++
CoxPASNet	C-Index	0.672 $\pm$ 0.002	0.672 $\pm$ 0.002	++
MiNet	C-Index	0.650 $\pm$ 0.020	0.652 $\pm$ 0.025	+++
pathDNN	R^2	0.801 $\pm$ 0.007	0.806 $\pm$ 0.007	+++
Multi-scale NN	Accuracy	0.660 $\pm$ 0.013	0.659 $\pm$ 0.014	+++
P-NET	AUC	0.899 $\pm$ 0.021	OP 0.896 $\pm$ 0.024 PP 0.887 $\pm$ 0.025 OP + PP 0.892 $\pm$ 0.027	++
PathCNN	AUC	0.745 $\pm$ 0.011	0.746 $\pm$ 0.007	++
PathGNN	AUC	0.693 $\pm$ 0.067	0.687 $\pm$ 0.060	++++
MPVNN	C-Index	0.632 $\pm$ 0.081	0.645 $\pm$ 0.086	+++
REDDA	AUC	0.899 $\pm$ 0.221	0.964 $\pm$ 0.050	+++
BINN	Accuracy	0.944 $\pm$ 0.023	OP 0.958 $\pm$ 0.016 OP + PP 0.937 $\pm$ 0.017	++
PINNet	AUC	0.974 $\pm$ 0.062	0.974 $\pm$ 0.064	+
PathExpSurv	C-Index	0.934 $\pm$ 0.029	0.936 $\pm$ 0.027	+++
DeepKEGG	AUC	0.892 $\pm$ 0.088	0.897 $\pm$ 0.090	++
Autosurv	C-Index	0.734 $\pm$ 0.048	0.732 $\pm$ 0.048	+++
Multilevel-GNN	C-Index	0.665 $\pm$ 0.029	0.657 $\pm$ 0.020	+++
GraphPath	Accuracy	0.867 $\pm$ 0.026	PP 0.878 $\pm$ 0.032	++++
Pathformer	F1 Macro	0.609 $\pm$ 0.077	OP 0.614 $\pm$ 0.071 OP + PP 0.587 $\pm$ 0.067	++++
APNet	AUC	0.999 $\pm$ 0.000	1.000 $\pm$ 0.000	++
HallmarkGraph	Accuracy	0.948 $\pm$ 0.021	0.947 $\pm$ 0.022	++

correlated (Spearman's $\rho = 0.4$ for DeepKEGG and $\rho = 0.56$ for BINN; Fig. S3 in Supplementary Data), indicating substantial alignment in attribution patterns despite the absence of biological pathway annotations.

Overall, these results suggest that the ability to identify disease-relevant biomarkers does not systematically depend on biological pathway annotations, but can emerge from sparsity-induced architectural constraints alone. The strong agreement between feature importance rankings obtained from pathway-informed and randomized models challenges the assumption that pathway annotations are essential for guiding feature selection. Instead, feature-level interpretability appears to reflect the structured inductive bias imposed by the model architecture rather than the biological information encoded by pathway annotations.

Discussion

In this study, we show that the empirical success of BINNs is largely explained by structural sparsity imposed by biological priors rather than by the biological semantics of pathway annotations, and we provide a practical evaluation protocol to test their contribution. Across multiple learning scenarios, datasets and metrics, models in which pathway information was randomized while preserving architectural constraints consistently matched or, in several cases, outperformed their biologically informed counterparts (Fig. 3). These findings indicate that the predictive gains commonly attributed to biological pathways do not require biologically meaningful connectivity.

Our results further demonstrate that randomized sparsification alone reproduces the performance benefits associated with pathway integration. Repeated randomization experiments on BINN, DeepKEGG, PASNet, PathCNN, and PINNet confirmed that model behavior is not sensitive to specific random seeds (Supplementary Fig. S1). Moreover, the sparsity levels induced by pathway annotations often do not coincide with those yielding optimal performance (Fig. 4). In models such as BINN and DeepKEGG, the best predictive accuracy was achieved at sparsity levels differing from those imposed by biological pathways, highlighting that pathway-derived sparsity does not necessarily constitute an optimal inductive bias. Collectively, these results suggest that sparsity should be treated as a tunable modeling choice rather than as a fixed property dictated by biological priors.

Several factors may explain why pathway integration provides no consistent benefits beyond sparsity (Fig. 6). First, curated pathway annotations cover only a subset of gene products, potentially excluding predictive features. Second, pathway-based sparsity can lead to 'superposed internal representations,' in which units combine unrelated signals, rather than encoding biologically coherent modules [54]. Such representations may support accurate prediction while deviating substantially from the intended biological interpretation. Third, relevant genes may be overlooked because they are absent from pathway databases.

In addition, pathway resources provide static representations of inherently dynamic processes. Pathway activity varies across cell types, disease states and environments, yet current models typically rely on fixed annotations. Extensive pathway overlap further

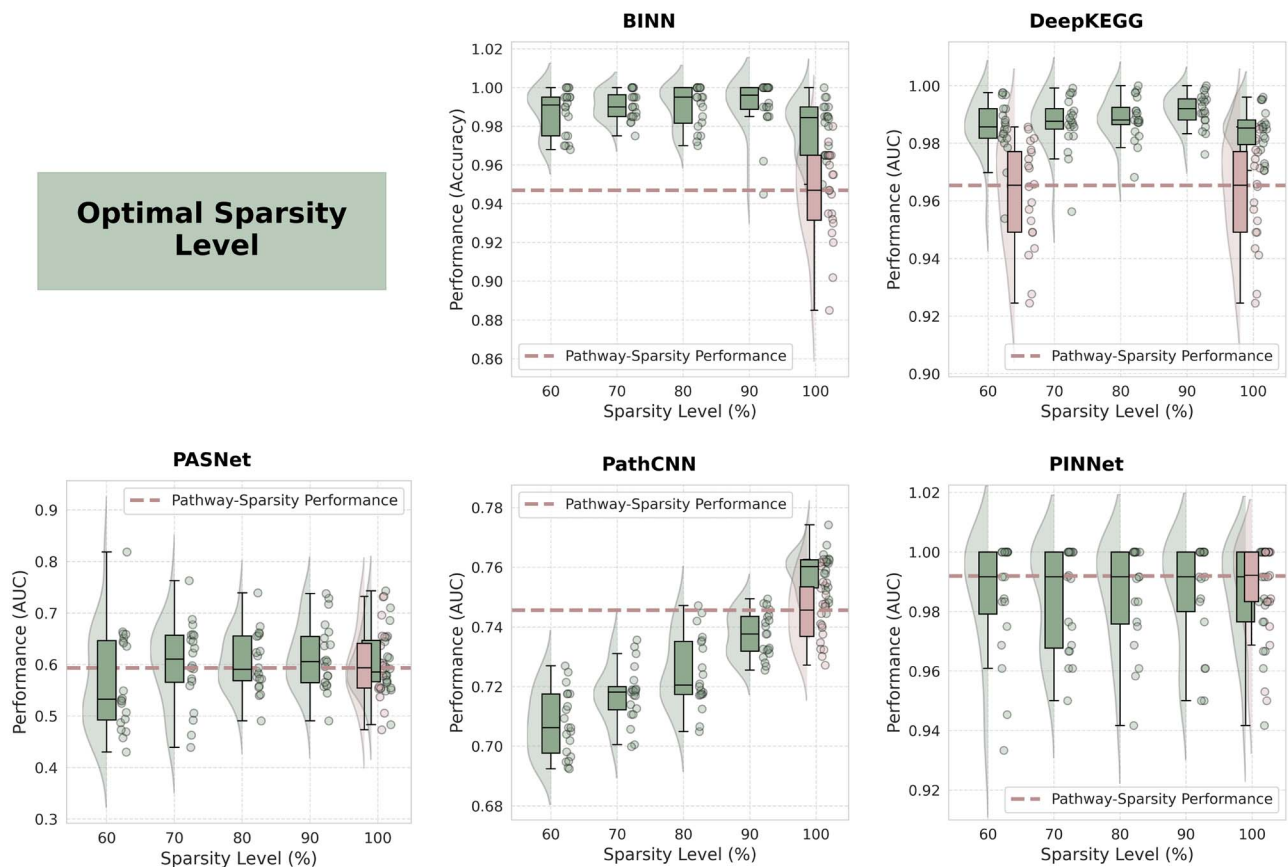


Figure 4. Effect of sparsity level on predictive performance. Green boxplots represent the performance (measured as accuracy or AUC) of each model—BINN, DeepKEGG, PASNet, PathCNN, and PINNet—across varying sparsity levels (60%–99%). Pink boxplots indicate performance at the sparsity level induced by pathway information. For DeepKEGG, the pink boxplots are repeated, as the pathway-induced sparsity level varies across omics, ranging from 63.7% for miRNAs to 98.9% for mRNAs. In general, boxplots illustrate the distribution of performance across runs, while violin plots provide density estimates. The dashed pink line indicates the performance achieved at the pathway-induced sparsity level. Pathway-Induced sparsity levels for all models are reported in [Tables S6, S8, and S9 in the Supplementary Data](#).

complicates interpretation, as redundancy among pathways can blur distinctions between modules. These limitations suggest that the lack of benefit observed here reflects not a failure of biology *per se*, but a mismatch between its representation and how learning objectives exploit it.

A central motivation for pathway-informed models is interpretability. However, our analysis reveals that interpretability outcomes fail to meaningfully distinguish biologically informed models from their randomized counterparts. Across models, both pathway-informed and randomized networks identified disease-associated features with comparable importance ([Fig. 5](#)), and feature importance rankings were often strongly correlated ([Supplementary Fig. S3](#)). Feature ablation experiments further demonstrated that removing highly predictive features did not uncover latent advantages of pathway information ([Supplementary Fig. S4](#)), indicating that biological priors are not merely masked by dominant biomarkers. These results suggest that feature-level interpretability metrics primarily reflect architectural inductive bias, rather than the biological correctness of pathway annotations.

Despite the promise of intrinsic explainability in BINNs, often contrasted with potentially unstable post hoc attribution methods [1], evaluation of explanation quality remains largely ad hoc. Most existing studies report only a small subset of top-ranked pathways or features, without assessing reproducibility across data splits or external

cohorts. Moreover, interpretability analyses that perturb individual pathway nodes in isolation do not reflect the distributed and nonlinear nature of neural network representations. As emphasized in a recent work, explainability should be treated as a first-class design objective rather than as an afterthought [55].

Taken together, our findings do not imply that biological knowledge or pathways are intrinsically irrelevant or dispensable in predictive modeling. Rather, they suggest that, when pathway annotations are used mainly as fixed architectural masks or static sparse connectivity patterns, their contribution cannot be readily disentangled from the regularization induced by sparsity. Future work may therefore explore biologically informed constraints that more directly influence the learning objective, such as context-dependent pathway activity, mechanistic consistency penalties, or experimentally grounded interaction priors, rather than relying solely on architectural masking.

More broadly, our results highlight the need for rigorous, structure-matched null models when evaluating biologically informed architectures. To this end, we introduce a general benchmarking workflow ([Fig. 7](#)) that enables systematic comparison between pathway-informed models and randomized counterparts across different pathway-informed architectures and prediction tasks. Such comparison is essential to determine whether improvements arise from biological insight or from generic inductive biases. Future biologically informed models should therefore demonstrate benefits beyond



Figure 5. Comparative feature importance of disease-associated features across pathway-informed and randomized models. Boxplots display the distribution of feature importance scores for disease-associated features (colored green) and unrelated ones (colored pink), across four models and disease contexts: (a) BINN - Sepsis, (b) DeepKEGG - LIHC, (c) PASNet - GBM, and (d) PINNet - AD. For each disease context, we compared pathway-informed models to their randomized counterparts, evaluating the ability of each model to assign higher importance to condition-specific features. Significant differences (red asterisks) between related and unrelated features scores were observed for the randomized BINN model ($P = .027$, panel a), the randomized PASNet model ($P = .040$, panel c), and both pathway-informed and randomized versions of the PINNet model for AD ($P < .001$, panel d). For clarity, we report only the results based on mRNA features for the DeepKEGG model, which integrates multiple data modalities. Analyses performed on SNV and miRNA features yielded consistent findings, with no statistically significant differences between pathway-informed and randomized models in their ability to prioritize biologically relevant features. Features of interest related to the analyzed diseases were obtained from curated databases such as DisGeNet [50], GeDiPNet [51], GeneCards [52], and AlzGene [53].

sparsity through appropriate controls and falsifiable benchmarks. Only then can claims of biological interpretability be grounded in biological contribution rather than architectural illusion.

Materials and methods

We outline practical guidelines for integrating pathway information into omics-based predictive models, following the schema displayed in Fig. 7. We formalize these guidelines as a step-by-step evaluation protocol. This workflow shows how to (i) combine pathway and omics data, (ii) encode these associations into graph representations, (iii) embed the resulting structures into neural network architectures, and (iv) benchmark performance against randomized baselines. These guidelines offer a clear framework that summarizes the following Methods section, helping researchers identify whether performance gains come from biological priors or from sparsity-induced regularization. This protocol is general and can be applied

to any pathway-informed architecture or omics-based predictive task.

Randomization procedure

Prior biological knowledge was encoded as either a bipartite graph, connecting features (e.g. genes) to pathways or a simple graph, connecting pathways to one another. Details on graph structures and metrics are reported in [Supplementary Data](#). Considering feature-pathway associations (the same procedure can also be applied to pathway-pathway associations), let A be a binary matrix of dimensions $N \times M$:

$$A = [a_{ij}], \quad \text{with } a_{ij} \in \{0, 1\}, \quad (1)$$

where N and M represent the number of features (e.g. genes, SNVs, miRNAs etc.) and the number of pathways, respectively. The matrix entries a_{ij} take values 1 (association) or 0 (no association).

Pathway vs. Randomized Models: Why Performance Aligns

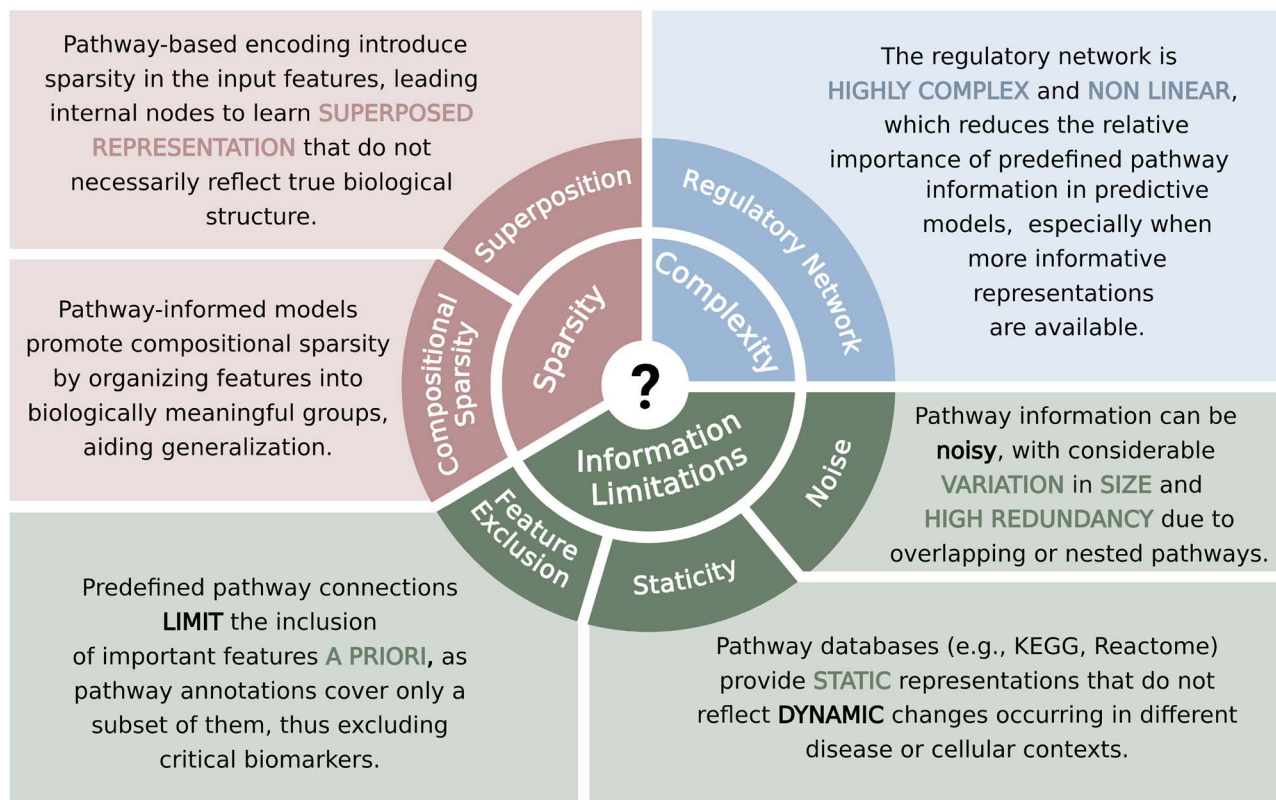


Figure 6. Hypothetical causes for the alignment in performance between pathway-informed and randomized models. Despite integrating biological knowledge, randomized models often perform comparably or better with respect to models incorporating pathway information. This figure summarizes several hypothetical factors that may contribute to explain this phenomenon.

The total number of connections in the matrix is defined as:

$$C = \sum_{i=1}^N \sum_{j=1}^M a_{ij}. \quad (2)$$

The number of features associated with each pathway j is given by:

$$d_j = \sum_{i=1}^N a_{ij}, \quad \forall j \in \{1, \dots, M\}. \quad (3)$$

The pathway randomization procedure generates a null model by permuting the associations a_{ij} according to the pathway annotations. The permutation preserves the following constraints:

1. **Preservation of the total number of associations:**

$$\sum_{i=1}^N \sum_{j=1}^M a'_{ij} = C, \quad (4)$$

where $A' = [a'_{ij}]$ is the matrix resulting after randomization.

2. **Preservation of the number of features per pathway:**

$$\sum_{i=1}^N a'_{ij} = d_j, \quad \forall j \in \{1, \dots, M\}. \quad (5)$$

3. **Uniform sampling of connections:** The reassignment of the connections is performed uniformly among all possible configurations satisfying the above constraints, ensuring that no additional structural bias is introduced.

The randomization operation can be performed through a uniform permutation of the connections while maintaining the above constraints. The algorithm for this process is:

1. Extract a list of all existing 1's in matrix A along with their respective indices (i, j) .
2. Shuffle this list uniformly.
3. Redistribute the 1's in the matrix A' while ensuring that each column j maintains the same number of connections d_j as in the original matrix.

This procedure preserves the sparsity structure of the original matrix.

In the approach illustrated in Fig. 1, panel (a), neuron connections within neural networks were replaced with random ones, while maintaining the same number of connections per neuron. This preserves the architectural sparsity induced by pathway annotations. Similarly, in the modality shown in Fig. 1, panel (c), randomization involved transforming tabular data into structured data by substituting the original pathway priors. Specifically, in GNNs, this was achieved by introducing random connections among the nodes in the input graphs. For CNNs, the randomization step consisted of constructing a

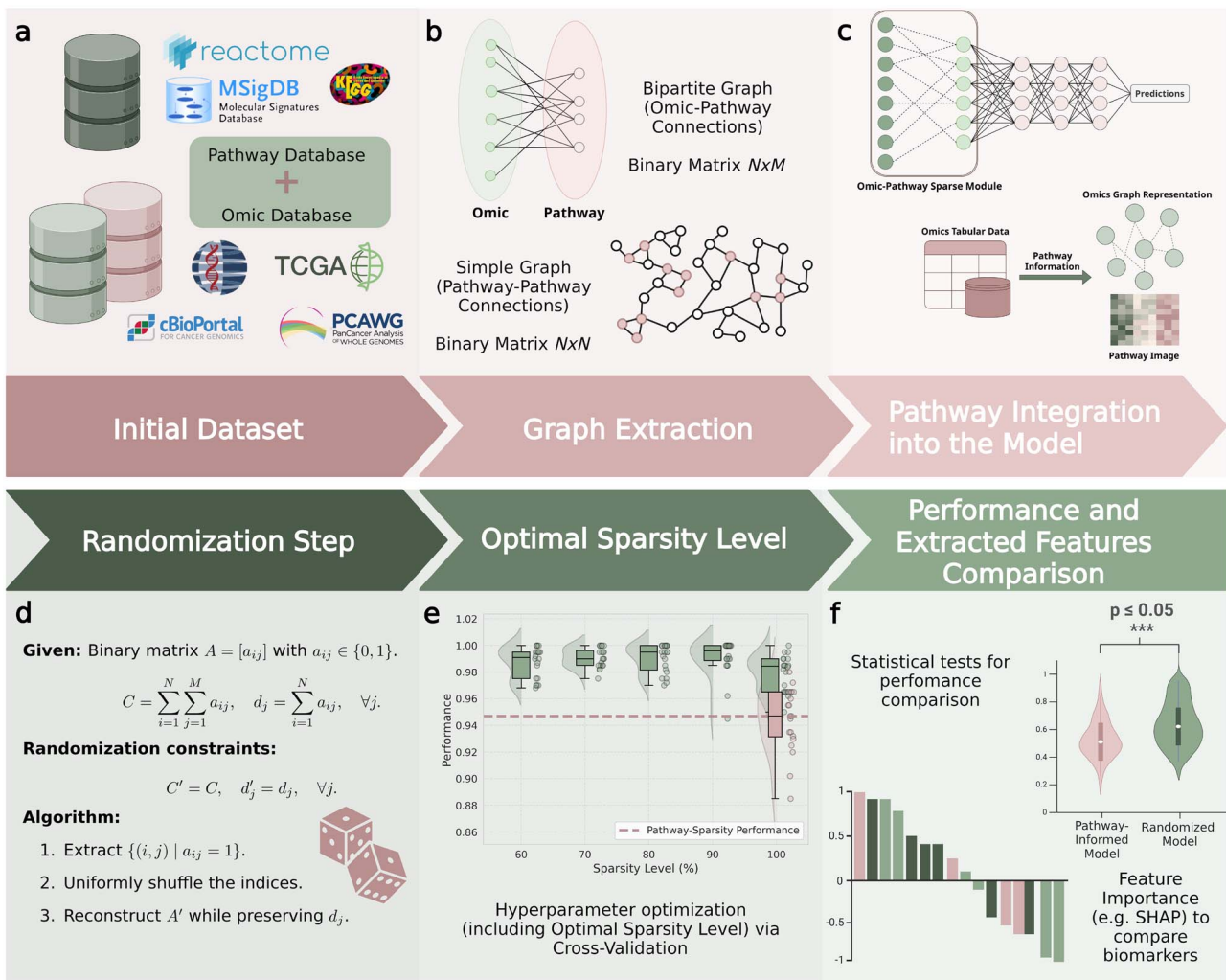


Figure 7. Guidelines for integrating pathway information into predictive models with rigorous benchmarking. This figure outlines a principled workflow for incorporating biological pathway knowledge into omics-based predictive models while ensuring robust validation against randomized baselines. (a) Datasets from pathway (e.g. Reactome, KEGG) and omics sources (e.g. TCGA, PCAWG) are combined to build a bipartite graph linking omic features to pathways or a simple graph linking pathways to each other. (b) The graph is encoded as a binary matrix either representing feature-to-pathway or pathway-to-pathway associations. (c) The graph structure is embedded into the model via a sparse omic-pathway module that enriches standard omics data with biologically informed connectivity or by modifying the structure of the input data (e.g. in GNN- and CNN-based models). (d) To assess the added value of true biological structure, a randomization step permutes pathway connections while preserving degree distributions, ensuring a fair comparison. (e) Optional optimization step: Optimize the sparsity of the omic-pathway graph to achieve better predictive performance. This is done using a cross-validation framework. In this step, the original degree distribution constraint is relaxed, allowing for a more flexible exploration of graph structures that may enhance model accuracy. (f) Statistical analyses and feature attribution methods (e.g. SHAP) are employed to compare model performance and feature relevance between biologically informed and randomized counterparts. This workflow enables rigorous validation of pathway integration, ensuring that observed improvements are due to meaningful biological priors.

“pathway image” by assigning random omics-related entities (e.g. genes, SNVs, miRNAs, etc.) to each pathway. In both cases, the number of connections in the network or the number of omics-related entities per pathway was preserved to maintain the same level of sparsity that was achieved through biological priors. This ensures that the randomization process mirrors the structural characteristics of the original models, preserving the sparsity effects while eliminating the biological information encoded by the pathway annotations.

Hyperparameter selection

After randomizing the model structures, both pathway-informed models and their randomized counterparts were trained and

evaluated to compare predictive performance. For each model, 20 independent 80/20 train/test splits were generated, with stratification according to task-specific labels when required. To ensure reproducibility, the exact seeds used to generate the 20 train/test splits and the 30 randomization-stability trials are reported in [Supplementary Table S11](#). When hyperparameter values were explicitly provided by the original authors, these values were used for both model versions. When original hyperparameters were unavailable or incomplete, hyperparameters were selected using five-fold cross-validation on the training set only. The best-performing hyperparameter configuration across validation folds was then retrained on the full training set and evaluated on the held-out test set. The same optimization protocol and search space were applied

to pathway-informed and randomized versions, and the validation metric matched the primary evaluation metric of the corresponding task. When early stopping was implemented in the original model code, we retained the original criterion; otherwise, no additional early stopping procedure was introduced. Further details are reported in [Supplementary Table S10](#).

Extended analysis of pathway-informed models

In addition to comparing predictive performance between randomized and pathway-informed models, further analyses were performed on PINNet, the fastest model to run, along with four other models: BINN, DeepKEGG, PASNet, and PathCNN.

Randomization trials

The randomization procedure described above was repeated 30 times, modifying the seed for the randomization functions, ensuring that each trial sampled a different set of connections among the C possible ones. For each randomization, the model was run 20 times. This was done to ensure that the results obtained with a single randomization were not due to a particularly favorable random seed.

Optimal sparsity level

We tested whether the biological contribution of pathway annotations lies not in the particular connections they encode, but in the overall level of sparsity they impose on the neural network. To this end, we first constructed a sparse neural network in which sparsity was defined by the number of pathway-derived connections, C , and compared its performance with that of a fully connected model. We then varied the number of retained connections between 60% and 99% of all possible connections, i.e.

$$C' = (1 - k) \cdot NM, \quad \text{with } k \in [0.6, 0.99] \quad (6)$$

Randomization was performed by uniformly sampling from all possible connections while ensuring that the total number of connections is equal to the desired sparsity level C' . In contrast to the previous randomization procedure, the number of connections assigned to each pathway was not constrained, allowing connections to be distributed freely across pathways.

The sparsity thresholds were chosen by varying around the biological sparsity induced by the pathways. As shown in [Supplementary Tables S6 and S8](#), the most common level of pathway-induced sparsity was approximately 97%–99%, while for miRNA in the DeepKEGG model, it was around 59%–66%. Statistical comparisons were applied to determine whether pathway-informed sparsity provided an advantage over arbitrary levels of sparsity.

In addition, to evaluate whether the number of pathways affected predictive performance, we performed a pathway-fraction analysis on BINN, DeepKEGG, PASNet, PathCNN, and PINNet by retaining 10%, 25%, 50%, 75%, and 100% of the available pathways. For each fraction, pathway-informed models were compared with randomized counterparts preserving the same number of retained pathways and structural constraints, using the same 20 train/test splits adopted in the main analysis.

Model interpretability and feature importance

Finally, we investigated the interpretability of the models to assess whether, even in the absence of pathway information, the randomized model could still identify relevant biomarkers for the

disease under study. Feature importance was estimated using the interpretability methods described in the original studies. For BINN, interpretability analyses were performed using SHAP (SHapley Additive exPlanations). In order to take into account the node connectivity and to avoid possible biases due to highly connected nodes, the resulting SHAP values were adjusted using the logarithm of the number of nodes in each node's reachable subgraph. Feature importance in PINNet was calculated through Deep SHAP, using the DeepExplainer SHAP package in Python. The obtained SHAP values were then aggregated across different cross-validation folds to obtain the final attribution scores, normalized as z-scores. For DeepKEGG, a simplified version of the DeepLIFT method was used. In this approach, the contribution of each feature to the model predictions is computed by multiplying the gradient of the output with respect to the input by the difference between the actual outputs and a reference activation (which was set to zero in this study). The feature importance was then obtained by aggregating across all samples to assess the overall relevance. In PASNet, since no interpretability module was provided in the corresponding GitHub repository of the model, a basic permutation importance approach was employed. Specifically, each feature was separately permuted, while keeping all the others fixed. Drops in performance were measured to estimate the features' relevance to the model output. Features were then ranked in descending order of performance impact. To assess the biological relevance of the model-derived feature rankings, we then compared the importance scores of disease-related features, derived from curated databases, with those of unrelated ones, for both pathway-informed models and their randomized counterparts. For each method, we considered the list of features associated with the specific disease investigated in the original study. Specifically, for BINN, Sepsis-related features ($n = 249$) were obtained from GeneCards and DisGeNet. For DeepKEGG, LIHC-related genes ($n = 428$) were retrieved from GeDiPNet. PASNet used 183 genes linked to Glioblastoma multiforme from GeDiPNet [51] and DisGeNet [50]. Finally, for PINNet (AD), 681 Alzheimer's-related genes were retrieved from the AlzGene database [53].

Feature ablation study

A feature ablation study was conducted on the models, assessing whether removing highly informative features for the prediction task would highlight the role of the pathways. This analysis tested whether the influence of the pathways was somehow being overshadowed by the contribution of highly predictive features. To identify and discard highly discriminative features from each dataset, we employed a Mann-Whitney U test-based approach. After each set of features was removed, the performance of the pathway-informed and randomized model versions was compared again.

Statistical analysis

Comparisons of predictive performance between biologically informed models and their randomized counterparts were performed using the Wilcoxon signed-rank test. When analyses encompassed multiple cancer types, P -values were combined using Fisher's combined probability test. Differences in feature importance between disease-associated and non-disease-associated features were assessed using the Mann-Whitney U test. Spearman's rank correlation coefficient was used to assess the concordance of feature rankings between biologically informed models and their randomized counterparts. P -values $< .05$ were considered statistically significant.

Computational resources and limitations of evaluated methods

All prediction analyses were executed on an NVIDIA GeForce RTX 4070 Max-Q GPU with 8 GB of memory. For models with higher memory demands (e.g. Autosurv, GraphPath, and Pathformer), a Tesla V100 SXM2 GPU with 32 GB of memory was utilized.

Unfortunately, several models could not be evaluated due to specific limitations. It was not possible to perform the performance comparison for models GCN-MAE and GCS-Net due to the unavailability of the code in their respective GitHub repositories. Additionally, models PathDeep, ReGeNNe, and PGLCN could not be included in the analysis because the necessary data for making predictions were not available.

Finally, we excluded the PathCNN model from the interpretability analyses since they were carried out using Gradient-weighted Class Activation Mapping (Grad-CAM) methods and focused on the pathway images provided as input to the model, thereby identifying entire pathways as important features. In such a setting, randomization fundamentally changes the semantic meaning of the pathway images, making any comparison of the most important features meaningless since, in the randomized model, those features no longer correspond to actual biological pathways.

Key points

- We propose a structure-matched randomization protocol to isolate the contribution of biological pathway annotations from the effects of sparsity in neural networks.
- Across the experimentally evaluable pathway-informed models, randomized counterparts achieve comparable or superior predictive performance.
- Biological pathway annotations do not confer a systematic advantage in interpretability, as randomized models recover disease-associated biomarkers with comparable effectiveness and correlated feature rankings.
- We provide a general benchmarking workflow demonstrating that predictive performance and interpretability can arise from structural constraints alone.

Acknowledgements

We thank the Fondazione Compagnia di San Paolo for supporting the ON-AIR project, of which this work is a part. I.C. acknowledges support from the University of Torino under the PhD programme in Complex Systems for Quantitative Biomedicine. T.S. acknowledges support from the PRIN project "Investigating the role of NF-YA isoform/lncRNA axis in mesoderm specification" (Grant ID: 20224TWKNJ). We also thank the Internationalization" program and PNRR M4C2 HPC—1.4 "CENTRI NAZIONALI"—Spoke 8.

Author contributions

Conceptualization: I.C., P.F. and T.S.; Formal Analysis, Methodology, Validation and Visualization: I.C.; Supervision: P.F. and T.S.; Writing—original draft: I.C.; Writing—review & editing: I.C., T.S., P.F., P.L., C.P., C.R. and F.S. All authors read and approved the final manuscript.

Supplementary material

[Supplementary material](#) is available at *Briefings in Bioinformatics* online.

Conflicts of interest

The authors declare that they have no competing interests.

Funding

This work was supported by the Italian Ministry of University and Research through the "Grant for Internationalization" program.

Data availability

The code used for the pathway connections randomization procedure can be found at the link: https://github.com/compmiomed-unito/Pathway_Randomization. This repository provides tools for pathway randomization in neural networks for omics data analysis, including functions to shuffle pathway connections while preserving specific constraints (e.g. desired sparsity levels). Code and datasets used to train the specific models were obtained from their respective repositories. A list of the models along with the links to their repositories can be found in the [Supplementary data](#).

References

1. Selby DA, Sprang M, Ewald J *et al.* Beyond the black box with biologically informed neural networks. *Nat Rev Genet* 2025;**26**:371–2. <https://doi.org/10.1038/s41576-025-00826-1>
2. Kanehisa M, Furumichi M, Sato Y *et al.* KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res* 2023;**51**:D587–92, 10. <https://doi.org/10.1093/nar/gkac963>
3. Milacic M, Beavers D, Conley P *et al.* The reactome pathway knowledgebase 2024. *Nucleic Acids Res* 2023;**52**:D672–8. <https://doi.org/10.1093/nar/gkad1025>
4. Selby DA, Jakhmola R, Sprang M *et al.* Visible neural networks for multi-omics integration: a critical review. *Front Artif Intell* 2025;**8**:1595291. <https://doi.org/10.3389/frai.2025.1595291>
5. Gaudet T, Malod-Dognin N, Sánchez-Valle J *et al.* Unveiling new disease, pathway, and gene associations via multi-scale neural network. *PLoS One* 2020;**15**:e0231059. <https://doi.org/10.1371/journal.pone.0231059>
6. Park S, Huang E, Ahn T. Classification and functional analysis between cancer and normal tissues using explainable pathway deep learning through RNA-sequencing gene expression. *Int J Mol Sci* 2021;**22**:11531. <https://doi.org/10.3390/ijms222111531>
7. Roy GG, Geard N, Verspoor K *et al.* MPVNN: mutated pathway visible neural network architecture for interpretable prediction of cancer-specific survival risk. *Bioinformatics* 2022;**38**:5026–32. <https://doi.org/10.1093/bioinformatics/btac636>
8. Ran M, Zhang S-L, Tam KY. Identifying meaningful drug response biomarkers from public pharmacogenomic datasets with biologically informed interpretable neural networks. *Comput Biol Chem* 2025;**120**:108669. <https://doi.org/10.1016/j.compbiolchem.2025.108669>
9. Hao J, Kim Y, Kim TK *et al.* PASNet: pathway-associated sparse deep neural network for prognosis prediction from high-throughput data. *BMC Bioinformatics* 2018;**19**:510. <https://doi.org/10.1186/s12859-018-2500-z>
10. Hao J, Kim Y, Mallavarapu T *et al.* Cox-PASNet: an artificial neural network for predicting prognosis in cancer patients based on pathway-associated sparse deep neural networks. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 381–386, Piscataway, NJ, USA: IEEE, 2018. <https://doi.org/10.1109/BIBM.2018.8621345>
11. Hao J, Masum M, Oh JH *et al.* Gene- and pathway-based deep neural network for multi-omics data integration to predict cancer

- survival outcomes. In: Cai Z, Skums P, Li M (eds), *Bioinformatics Research and Applications*, 113–24. Cham: Springer International Publishing, 2019 https://doi.org/10.1007/978-3-030-20242-2_10.
12. Deng L, Cai Y, Zhang W *et al*. Pathway-guided deep neural network toward interpretable and predictive modeling of drug sensitivity. *J Chem Inf Model* 2020;**60**:4497–505. <https://doi.org/10.1021/acs.jcim.0c00331>
 13. Hou Z, Leng J, Yu J *et al*. PathExpSurv: pathway expansion for explainable survival analysis and disease gene discovery. *BMC Bioinformatics* 2023;**24**:434. <https://doi.org/10.1186/s12859-023-05535-2>
 14. Gavrilidis GI, Vasileiou V, Dimitsaki S *et al*. APnet, an explainable sparse deep learning model to discover differentially active drivers of severe Covid-19. *Bioinformatics* 2025;**41**:btaf063. <https://doi.org/10.1093/bioinformatics/btaf063>
 15. Elmarakeby HA, Hwang J, Arafeh R *et al*. Biologically informed deep neural network for prostate cancer discovery. *Nature* 2021;**598**:348–52. <https://doi.org/10.1038/s41586-021-03922-4>
 16. Hu J, Yu W, Dai Y *et al*. A deep neural network for gastric cancer prognosis prediction based on biological information pathways. *J Oncol* 2022;**2022**:1–9. <https://doi.org/10.1155/2022/2965166>
 17. Hartman E, Scott AM, Karlsson C *et al*. Interpreting biologically informed neural networks for enhanced proteomic biomarker discovery and pathway analysis. *Nat Commun* 2023;**14**:5359. <https://doi.org/10.1038/s41467-023-41146-4>
 18. Kim H, Lee H. PINNet: a deep neural network with pathway prior knowledge for Alzheimer's disease. *Front Aging Neurosci* 2023;**15**:1126156. <https://doi.org/10.3389/fnagi.2023.1126156>
 19. Meirer J, Wittwer LD, Revol V *et al*. DeepBINN: a tailored biologically-informed neural network for robust biomarker identification. In: *2024 11th IEEE Swiss Conference on Data Science (SDS)*, Piscataway, NJ, USA: IEEE, 2024. <https://doi.org/10.1109/SDS60720.2024.00044>
 20. Lan W, Liao H, Chen Q *et al*. DeepKEGG: a multi-omics data integration framework with biological insights for cancer recurrence prediction and biomarker discovery. *Brief Bioinform* 2024;**25**:bbae185. <https://doi.org/10.1093/bib/bbae185>
 21. Liu X, Tao Y, Cai Z *et al*. Pathformer: a biological pathway informed transformer for disease diagnosis and prognosis using multi-omics data. *Bioinformatics* 2024;**40**:btae316. <https://doi.org/10.1093/bioinformatics/btae316>
 22. Jiang L, Xu C, Bai Y *et al*. Autosurv: interpretable deep learning framework for cancer survival analysis incorporating clinical and multi-omics data. *NPJ Precis Oncol* 2024;**8**:4. <https://doi.org/10.1038/s41698-023-00494-6>
 23. Lee S, Lim S, Lee T *et al*. Cancer subtype classification and modeling by pathway attention and propagation. *Bioinformatics* 2020;**36**:3818–24. <https://doi.org/10.1093/bioinformatics/btaa203>
 24. Liang B, Gong H, Lu L *et al*. Risk stratification and pathway analysis based on graph neural network and interpretable algorithm. *BMC Bioinformatics* 2022;**23**:394. <https://doi.org/10.1186/s12859-022-04950-1>
 25. Liu C, Wan AH, Liang H *et al*. Biological informed graph neural network for tumor mutation burden prediction and immunotherapy-related pathway analysis in gastric cancer. *Comput Struct Biotechnol J* 2023;**21**:4540–51. <https://doi.org/10.1016/j.csbj.2023.09.021>
 26. Cai H, Liao Y, Zhu L *et al*. Improving cancer survival prediction via graph convolutional neural network learning on protein-protein interaction networks. *IEEE J Biomed Health Inform* 2024;**28**:1134–43. <https://doi.org/10.1109/JBHI.2023.3332640>
 27. Zhang Q, Liu F, Lai X. Hallmarkgraph: a cancer hallmark informed graph neural network for classifying hierarchical tumor subtypes. *Bioinformatics* 2025;**41**:btaf444. <https://doi.org/10.1093/bioinformatics/btaf444>
 28. Ma T, Wang J. Graphpath: a graph attention model for molecular stratification with interpretability based on the pathway–pathway interaction network. *Bioinformatics* 2024;**40**:btae165. <https://doi.org/10.1093/bioinformatics/btae165>
 29. Yaowen G, Zheng S, Yin Q *et al*. REDDA: integrating multiple biological relations to heterogeneous graph neural network for drug-disease association prediction. *Comput Biol Med* 2022;**150**:106127. <https://doi.org/10.1016/j.compbimed.2022.106127>
 30. Yan H, Weng D, Dongguo Li YG *et al*. Prior knowledge-guided multilevel graph neural network for tumor risk prediction and interpretation via multi-omics data integration. *Brief Bioinform* 2024;**25**:bbae184. <https://doi.org/10.1093/bib/bbae184>
 31. Chatzianastasis M, Vazirgiannis M, Zhang Z. Explainable multilayer graph neural network for cancer gene prediction. *Bioinformatics* 2023;**39**:btad643. <https://doi.org/10.1093/bioinformatics/btad643>
 32. Wen W, Wu C, Wang Y *et al*. Learning structured sparsity in deep neural networks. In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Vol. **30**, 2082–2090, Red Hook, NY: Curran Associates Inc., 2016. <https://doi.org/10.5555/3157096.3157329>
 33. Hastie T, Tibshirani R, Wainwright M. *Statistical Learning with Sparsity: The Lasso and Generalizations*. Boca Raton, FL: Chapman & Hall, CRC Press, 2015. <https://doi.org/10.5555/2834535>
 34. Poggio T. How deep sparse networks avoid the curse of dimensionality: efficiently computable functions are compositionally sparse. In: *Technical Report CBMM Memo 118*. Cambridge, MA: Center for Brains, Minds and Machines (CBMM), 2022. <https://hdl.handle.net/1721.1/145776>
 35. Hoefler T, Alistarh D, Ben-Nun T *et al*. Sparsity in deep learning: pruning and growth for efficient inference and training in neural networks. *J Mach Learn Res*, 2021;**22**:1–124. <https://doi.org/10.5555/3546258.3546499>
 36. Poggio T, Fraser M. Compositional sparsity of learnable functions. *Bull Am Math Soc*, 2024;**61**:438–456. <https://doi.org/10.1090/bull/1820>
 37. Bach F. Structured sparsity-inducing norms through submodular functions. *arXiv preprint*, arXiv:1008.4220, 2010. <https://arxiv.org/abs/1008.4220>
 38. Zhang X, Zhao J. Group variable selection via group sparse neural network. *Comput Stat Data Anal* 2024;**192**:107911. <https://doi.org/10.1016/j.csda.2023.107911>
 39. Yoon J, Hwang SJ. Combined group and exclusive sparsity for deep neural networks. In: *Proceedings of the 34th International Conference on Machine Learning*, vol. 70 of *Proceedings of Machine Learning Research*. Precup D, Teh YW (eds), 3958–66. Cambridge, MA, USA: PMLR, 2017.
 40. Scardapane S, Comminiello D, Hussain A *et al*. Group sparse regularization for deep neural networks. *Neurocomputing* 2017;**241**:81–9. <https://doi.org/10.1016/j.neucom.2017.02.029>
 41. Schaefer CF, Anthony K, Krupa S *et al*. PID: the pathway interaction database. *Nucleic Acids Res* 2009;**37**:D674–9. <https://doi.org/10.1093/nar/gkn653>
 42. Nishimura D. Biocarta. *Biotech Software & Internet Report*, 2004. <https://doi.org/10.1089/152791601750294344>
 43. Liberzon A, Subramanian A, Pinchback R *et al*. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 2011;**27**:1739–40. <https://doi.org/10.1093/bioinformatics/btr260>

44. Ashburner M, Ball CA, Blake JA *et al.* Gene ontology: tool for the unification of biology. *Nat Genet* 2000;**25**:25–9. <https://doi.org/10.1038/75556>
45. Kamburov A, Herwig R. ConsensusPathDB 2022: molecular interactions update as a resource for network biology. *Nucleic Acids Res* 2021;**50**:2021. <https://doi.org/10.1093/nar/gkab1128>
46. Agrawal A, Balci H, Hanspers K *et al.* Wikipathways 2024: next generation pathway database. *Nucleic Acids Res* 2023;**52**:D679–89. <https://doi.org/10.1093/nar/gkad960>
47. Evangelista JE, Xie Z, Marino GB *et al.* Enrichr-KG: bridging enrichment analysis across multiple libraries. *Nucleic Acids Res* 2023;**51**:W168–W179. <https://doi.org/10.1093/nar/gkad393>
48. Oh JH, Choi W, Ko E *et al.* PathCNN: interpretable convolutional neural networks for survival prediction and pathway analysis applied to glioblastoma. *Bioinformatics* 2021;**37**:i443–i450. <https://doi.org/10.1093/bioinformatics/btab285>
49. Sharma D, Xu W. ReGeNNe: genetic pathway-based deep neural network using canonical correlation regularizer for disease prediction. *Bioinformatics* 2023;**39**:btad679. <https://doi.org/10.1093/bioinformatics/btad679>
50. Piñero J, Ramírez-Angueta JM, Saüch-Pitarch J *et al.* The disgenet knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res* 2020;**48**:D845–D855. <https://doi.org/10.1093/nar/gkz1021>
51. Kundu I, Sharma M, Barai RS *et al.* GeDiPNet: online resource of curated gene-disease associations for polypharmacological targets discovery. *Genes Dis* 2022;**10**:647–649. <https://doi.org/10.1016/j.gendis.2022.05.034>
52. Safran M, Rosen N, Twik M *et al.* *The GeneCards Suite*. Singapore: Springer Nature Singapore, 2021. 27–56. ISBN 978-981-16-5812-9. https://doi.org/10.1007/978-981-16-5812-9_2
53. Bertram L, McQueen M, Mullin K *et al.* Systematic meta-analyses of Alzheimer disease genetic association studies: the alzgene database. *Nat Genet* 2007;**39**:17–23. <https://doi.org/10.1038/ng1934>
54. Elhage N, Hume T, Olsson C *et al.* Toy models of superposition. *arXiv preprint arXiv:2209.10652*, 2022. <https://arxiv.org/abs/2209.10652>
55. Beckh K, Müller S, Jakobs M *et al.* Harnessing prior knowledge for explainable machine learning: an overview. In: *2023 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, 450–63. Piscataway, NJ, USA: IEEE, 2023. <https://doi.org/10.1109/SaTML54575.2023.00038>