

Multimodal Graph Neural Networks in Healthcare: A Review of Fusion Strategies Across Biomedical Domains

Maria Vaida^{1,*}, Ziyuan Huang²

¹ Harrisburg University of Science and Technology, Harrisburg, PA, USA

² UMass Chan Medical School, Worcester, MA, USA

Correspondence*:

Corresponding Author

mvaida@harrisburgu.edu

2 ABSTRACT

3 Graph Neural Networks (GNNs) have transformed multimodal healthcare data integration by
4 capturing complex, non-Euclidean relationships across diverse sources such as electronic health
5 records, medical imaging, genomic profiles, and clinical notes. This review synthesizes GNN
6 applications in healthcare, highlighting their impact on clinical decision-making through multimodal
7 integration, advanced fusion strategies, and attention mechanisms. Key applications include drug
8 interaction and discovery, cancer detection and prognosis, clinical status prediction, infectious
9 disease modeling, genomics, and the diagnosis of mental health and neurological disorders.
10 Various GNN architectures demonstrate consistent applications in modeling both intra- and
11 intermodal relationships. GNN architectures, such as Graph Convolutional Networks and Graph
12 Attention Networks, are integrated with Convolutional Neural Networks (CNNs), transformer-
13 based models, temporal encoders, and optimization algorithms to facilitate robust multimodal
14 integration. Early, intermediate, late, and hybrid fusion strategies, enhanced by attention
15 mechanisms like multi-head attention, enable dynamic prioritization of critical relationships,
16 improving accuracy and interpretability. However, challenges remain, including data heterogeneity,
17 computational demands, and the need for greater interpretability. Addressing these challenges
18 presents opportunities to advance GNN adoption in medicine through scalable, transparent GNN
19 models.

20 **Keywords:** Graph Neural Networks, Multimodal Fusion, Healthcare Applications, Biomedical Data Integration, Attention Mechanisms,
21 Drug Discovery, Cancer Prognosis, Neurological Disorders, Electronic Health Records, Epidemiology Forecasting, Clinical Prediction,
22 Genomics Analysis

INTRODUCTION

23 Graphs serve as fundamental mathematical structures for representing and analyzing the complex
24 relationships inherent in multimodal datasets. In the healthcare domain, nodes in a graph can represent
25 medical entities such as patients, diseases, genes, proteins, medications, and healthcare providers, while
26 edges capture the associations or interactions among them Paul et al. (2024). Node and edge features
27 may incorporate additional attributes, including patient demographic details, disease states, medical notes,
28 or medication properties Li et al. (2023a). Traditional machine learning and deep learning techniques,

29 designed primarily for Euclidean data, often struggle to accommodate the non-Euclidean nature of relational
30 medical data. GNNs address this limitation by extending deep neural networks to graph-structured data by
31 aggregating and propagating information from neighboring nodes to learn high-order interactions through
32 methods such as contrastive, generative, and explainable GNNs Kumar et al. (2023b); Sefer (2025b,a); Cetin
33 and Sefer (2025). This enables GNNs to generate graph-level representations that capture the structural and
34 semantic complexity of medical data Lee et al. (2024a); Kumar et al. (2023b). GNNs have proven effective
35 in a wide range of healthcare applications, from disease diagnosis and comorbidity prediction to patient
36 referral optimization and emotional intelligence modeling in clinical settings Sangeetha et al. (2024); Pablo
37 et al. (2024); Wang (2022); Xu et al. (2024).

38 Healthcare data is inherently diverse and often available in multiple modalities, including structured
39 data like EHRs, unstructured data like clinical notes, and complex forms like medical images (MRI,
40 CT, PET, EEG, MEG), chemical, laboratory, temporal, and genomic data. Integrating and analyzing
41 these heterogeneous data sources is crucial for a holistic understanding of disease and patient conditions.
42 Multimodal learning, which aims to leverage complementary information from different modalities, is a
43 logical tool for incorporating these disparate data sources Waqas et al. (2024); Stahlschmidt et al. (2022);
44 Teoh et al. (2024); Dumyn et al. (2024). GNNs are particularly well suited for multimodal healthcare
45 applications, as they can model the intricate relationships within and between these diverse data streams
46 and can be fused together with other deep learning or machine learning models Dawn et al. (2024); Paul
47 et al. (2024); Johnson et al. (2024).

48 This paper provides a review of the recent applications of GNNs in healthcare, with a specific focus on
49 approaches that incorporate multimodal data. We structure the review by grouping applications into key
50 themes: pharmacology, oncology, epidemiology, neuropsychiatry, clinical risk prediction, and genomics.
51 By examining the methodologies, findings, and challenges within each area, this review aims to offer a
52 comprehensive overview of the current landscape and potential future directions for GNNs in computational
53 healthcare.

54 We defined the scope in advance to include primary studies that (1) apply a graph neural network to a
55 biomedical or clinical task and (2) integrate at least two data modalities or combine graph learning with
56 other encoders within an explicit fusion scheme. We searched PubMed, Google Scholar, and arXiv for
57 studies published between January 2020 and August 2025 using combinations of graph-learning terms (e.g.,
58 GNN, GCN, GraphSAGE, GAT, heterogeneous graph), multimodality terms (e.g., multimodal, fusion),
59 and health-domain terms (e.g., clinical, oncology, pharmacology, genomics). Titles and abstracts were
60 screened against predefined inclusion and exclusion criteria, followed by full-text assessment. We included
61 studies that reported the fusion strategy and described the architectural components used; single-modality
62 GNNs, non-health domains, and papers lacking full text were excluded. The search identified 121 records,
63 of which 85 studies met the eligibility criteria and were included in the review. Because reporting practices
64 and evaluation metrics vary widely across domains, we used descriptive synthesis rather than quantitative
65 meta-analysis. Complete search strings and eligibility details are provided in the Supplementary Tables
66 (S1, S2, and S3).

1 PHARMACOLOGY

67 Pharmacology-focused multimodal GNN frameworks unify molecular, biological, and clinical signals
68 under predominantly intermediate, attention-aware fusion, with early fusion used when EHR/image
69 or graph features are concatenated prior to graph convolutions (Table 2). Heterogeneous graphs

70 (drugs–targets–diseases–genes–adverse events), patient/population graphs, meta-path encoders with
71 explainable decoders, and attention are common graph modeling approaches Gao et al. (2025b); Huang et al.
72 (2023); Zhou et al. (2024); Dawn et al. (2024). Drug-drug interaction models integrate drug–protein–disease
73 multiplexes with multi-head attention, temporal or GNN/DNN pipelines, and graph transformers Yu et al.
74 (2023); Gan et al. (2023); Al-Rabeah and Lakizadeh (2022); OMKUMARCHANDRAUMAKANTHAM
75 and PATHAK (2024); Wang et al. (2024a); Xiong et al. (2023). Drug–target affinity prediction tasks fuse
76 molecular graphs with knowledge-graph embeddings and attention modules Yella et al. (2022); Zhang
77 et al. (2023b); Xiang et al. (2025). Drug repurposing leverages knowledge-graph VAEs/GraphSAGE over
78 drug databases to prioritize candidates, while adversarial designs extend to adverse events prediction
79 and drug recommendations Hsieh et al. (2020, 2021); Artiñano-Muñoz et al. (2024); Lin et al. (2023);
80 Abdeddaiem et al. (2025). Time series and causal structure are explicit in models that learn temporal edges
81 or motif-level constraints (e.g., CT-GNN/MDTCKGNN) and in prescription prediction with time-aware
82 modules (T-LSTM) Kalla et al. (2023); Liu et al. (2020). Vision-centric tasks (pill classification) add
83 ConvNet/RPN with graph topology learning. Protein localization alteration and colonization-risk models
84 adapt GraphSAGE/GCN/GAT to dynamic clinical graphs Nguyen et al. (2023); Wang et al. (2023a);
85 Gouareb et al. (2023).

86 GraphSAGE/GCN/GAT/RGCN provide the backbone of drug-related multimodal GNN approaches,
87 with attention (often multi-head) capturing neighbor weighting and modality selection. VGAE/GAN
88 variants aid representation learning and robustness Yu et al. (2023); Wang et al. (2024a); Xiang et al.
89 (2025); Abdeddaiem et al. (2025). Datasets span FAERS, SIDER/OFFSIDES/TWOSIDES, DrugBank,
90 KEGG, STRING, CCLE/GDSC, KIBA/DAVIS, RepoDB, and MIMIC-III/MIMIC-IV, enabling cross-
91 domain evaluation from molecules to bedside Gao et al. (2025b); Dawn et al. (2024); Al-Rabeah and
92 Lakizadeh (2022); Yella et al. (2022); Zhang et al. (2023b); Liu et al. (2020). Recent surveys have argued
93 that multimodal, knowledge-graph-aware, and temporally grounded GNNs tend to improve property
94 prediction, DDI/ADE surveillance, repurposing, and recommendation while enhancing mechanistic insight
95 and scalability Paul et al. (2024); A.Tabatabaei et al. (2025); Yao et al. (2024); Wang et al. (2024c); Li et al.
96 (2023a).

2 ONCOLOGY

97 Oncology-focused multimodal graph frameworks fuse histopathology, radiology, omics, and clinical
98 covariates to support diagnosis, risk stratification, and treatment planning tasks (Table 3). Most systems
99 pair modality-specific encoders, such as (CNNs/ViTs or radiomics for images, text encoders for reports,
100 and pathway/interaction graphs for omics, with graph layers under intermediate fusion, frequently using
101 attention for weighting Kulandaivelu et al. (2024); Kim et al. (2023); Alzoubi et al. (2024); Pratap Joshi et al.
102 (2025); Yan et al. (2024); Gowri et al. (2024). Population graphs connect patients via various similarity
103 measures in imaging and clinical embeddings (head and neck, ovarian cancers), while pathways and
104 knowledge graphs encode gene–gene or entity relations for subtype and survival modeling Peng et al.
105 (2024); Ghantasala et al. (2024); Li et al. (2023b). Lesser used, late fusion is applied when independently
106 learned patient–gene bipartite embeddings are aligned for survival (MGNN) Gao et al. (2022), whereas early
107 fusion concatenates raw/image features before graph reasoning in lung and federated liver cancer models
108 Li et al. (2023b); Moharana et al. (2025). Beyond core oncology tasks, misinformation detection integrates
109 text encoders with R-GCN over medical knowledge graphs under early fusion Cui et al. (2020). These
110 architectures standardize heterogeneous inputs, learn structure-aware patient and pathway representations,

111 and improve generalization via similarity graphs and attention-based aggregation across modalities and
112 fusion types Li et al. (2023a); Paul et al. (2024); Waqas et al. (2024).

3 NEUROPSYCHIATRY

113 Multimodal GNN frameworks extended to neurological domains have been applied to conditions such as
114 Alzheimer's disease, Parkinson's disease, depression, autism spectrum disorder, Schizophrenia, and even
115 emotion recognition and sentiment analysis by integrating diverse linguistic, genomic, behavioral, imaging,
116 and physiological data Teoh et al. (2024); Zhang et al. (2023a); Xu et al. (2024); Sangeetha et al. (2024);
117 Khemani et al. (2024).

118 Neuropsychiatry multimodal GNN pipelines unify imaging (fMRI/sMRI/DTI/PET), electrophysiology
119 (EEG), speech/text, and omics within subject or population-level graphs (Table 4). A common approach
120 in Alzheimer's disease prediction integrates imaging-driven fusion with cross-attention Transformers
121 (CsAGP, GCNCS), dual hypergraphs (DHFWSL), multiplex subject graphs (HetMed), and hypergraph
122 attention fusion (HCNN-MAFN) Tang et al. (2023a); Luo et al. (2024); Kim et al. (2023); Kumar et al.
123 (2023a); Lee et al. (2024b). Parkinson's studies pair connectomic encoders with omics via attention
124 (JOIN-GCLA) and patient-similarity graphs (AdaMedGraph) Chan et al. (2022); Lian et al. (2023).
125 Autism Spectrum Disorder models treat rs-fMRI as signals on DTI graphs (M-GCN) to intermediate
126 spatio-temporal/demographic fusion (IFC-GNN) and VAE-aligned Transformer/Graph-U-Net encoders
127 (MM-GTUNets) D'Souza et al. (2021); Wang et al. (2024b); Cai et al. (2025). For Major Depressive
128 Disorder, interview-centric systems employ heterogeneous attention over audio-video-text (AVS-GNN,
129 DSE-HGAT), while imaging/population approaches (LGMF-GNN, FC-HGNN, Ensemble GNN) couple
130 local ROI graphs to global subject graphs Li et al. (2025, 2024); Liu et al. (2024); Gu et al. (2025);
131 Venkatapathy et al. (2023); Lee et al. (2024a). Schizophrenia pipelines tend to model EEG channel-graphs
132 and dual-branch DTI attention networks integrating FA/FN features Jiang et al. (2023); Gao et al. (2025a).
133 Attention weights filter population graphs based on their similarity, and learn multi-scale spatial-temporal
134 patterns by combining CNN/Transformer encoders with GNN message passing inside the fusion stack.

4 EPIDEMIOLOGY

135 Recent epidemic-forecasting and COVID-19 outcome models fuse temporal sequence encoders with
136 structure-aware GNNs (Table 5). For population-level spread, architectures stack temporal CNN/DNN
137 modules with attention-based GNN layers to capture local and global transmission patterns (MSGNN,
138 EpiGNN) and augment signals with LLM-derived social media features or dual topologies to improve
139 influenza forecasts (MGLEP, Dual-Topo-STGCN) Qiu et al. (2024); Xie et al. (2022); Tran et al. (2024);
140 Luo et al. (2025). Within hospitals, contact graphs linking patients and healthcare workers use GraphSAGE
141 and attention to model hospital-acquired infection transmission Gouareb et al. (2023). For COVID-19
142 prognosis, multimodal pipelines use attention to fuse CT-derived features with KNN population graphs
143 Keicher et al. (2023), while edge-flexible GCNN frameworks integrate imaging, tabular, and temporal
144 signals (CNN/LSTM and population GNN) to allow post-training edge adaptability Tariq et al. (2023,
145 2025). These models emphasize spatiotemporal message passing, attention for weighting neighbors and
146 signals, and adaptable graph construction to handle dynamic data.

5 CLINICAL

147 EHR-based multimodal graph frameworks aim to support clinical prediction and treatment planning through
148 merging diverse medical data modalities Li et al. (2022); Xu et al. (2024). When combined with knowledge
149 graphs, these models offer flexibility in terms of both inputs and prediction tasks Nye (2023); Rajabi and
150 Kafaie (2022). Most models integrate structured EHR (diagnoses, procedures, meds, labs, vitals) with at
151 least one unstructured or high-dimensional stream, be it clinical notes, medical images (CXR, fundus),
152 genomics, or wearable/sensor data, often via CNNs for imaging, TF-IDF/BioBERT for text, and temporal
153 trajectory layers for labs/vitals AL-Sabri et al. (2024); Tang et al. (2023b); Zedadra et al. (2025); Pablo et al.
154 (2024); Wang et al. (2025). The graph connectivity tends to be modeled as patient–patient similarity graphs,
155 knowledge graphs linking encounters to conditions, and heterogeneous graphs (e.g., sensor and metapath
156 views) (Table 6). Dynamic network edges implemented in conjunction with learned message-passing
157 connectivity from static KGs allow graphs to adapt to new information without the need for retraining Liu
158 et al. (2021); Valls et al. (2023); Gao et al. (2024); Wang et al. (2025); Christos Maroudis et al. (2025).

159 In terms of multimodal fusion strategies, the majority of models start with modality-specific encoders
160 (CNNs for images, BiGRU/LSTM/Transformers for sequences/text), which are then integrated into
161 GNN backbones (GraphSAGE, GNN/GAT, heterogeneous GNN), with attention used both for cross-
162 modal weighting and within graph layers AL-Sabri et al. (2024); Tang et al. (2023b); Begum (2024);
163 Boschi et al. (2024); Ghanvatkar and Rajan (2023). Temporal structure can be modeled at the node level
164 (RNN/Transformer encoders per patient), edge level (temporal embeddings that define adaptive edges),
165 and graph level (dynamic GNNs that rebuild neighborhoods by top-k similarity each step). Dissentangled
166 dynamic attention separates invariant vs. shifting patterns and fairness-aware designs Tang et al. (2023b);
167 Zhang et al. (2024); Christos Maroudis et al. (2025).

168 MIMIC-III and MIMIC-IV are two of the most used datasets for mortality and length-of-stay prediction,
169 as well as readmission, sepsis trajectory modeling, and heart-disease graphs, integrated with similarity-
170 based measures, temporal encoders, dynamic graph update strategies, and privacy-preserving architectures
171 AL-Sabri et al. (2024); Tang et al. (2023b); Ghanvatkar and Rajan (2023); Christos Maroudis et al. (2025);
172 Begum (2024). Imaging-heavy models join population graphs with CNN/radiomics for tasks such as
173 ophthalmology and DR screening (APPOS, MESSIDOR) Gao et al. (2024); Zedadra et al. (2025), while
174 sensor-centric pipelines exploit heterogeneous sensor-and-knowledge graphs Wang et al. (2025). SHARE,
175 Synthea, and ANIC datasets support multitask longitudinal modeling, ER triage, and out-of-distribution
176 ICU biomarker forecasting Boschi et al. (2024); Valls et al. (2023); Zhang et al. (2024).

177 By uniting EHR, imaging, genomic, temporal, and sensor-derived information within attention-based
178 graph representations, diagnostics and prognostic models capture both the relational and temporal
179 complexities inherent in patient care Oss Boll et al. (2024). Their reliance on attention-based fusion
180 and invariant pattern learning reflects a shift toward systems capable of modeling data heterogeneity and
181 distribution shifts, resulting in scalable and generalizable clinical decision-support systems.

6 GENOMICS

182 Across lncRNA–miRNA interaction prediction, GNN models implement sequence-aware fusion with
183 attention, built over heterogeneous similarity graphs (Table 7). Modalities and features typically
184 combine primary sequence (k-mers), similarity networks (sequence/functional/disease), and structural
185 or physicochemical descriptors into unified node–edge representations Wang et al. (2023b, 2022); Wang
186 and Chen (2023); Zhang et al. (2022). Sequence embeddings are often initialized via unsupervised

objectives (e.g., k-mer Doc2Vec) before graph learning, then refined with inductive backbones such as GraphSAGE and attention layers to weight informative neighbors Wang et al. (2023b); Zhang et al. (2022). Heterogeneous/bipartite graphs integrate lncRNA–miRNA and miRNA–disease with similarity measures, structured probabilistic layers, or multi-channel attention Wang et al. (2022); Wang and Chen (2023). Datasets such as LncACTdb, LNCipedia, miRBased, ncRNAsNP, and HMDD are integrated into pretrained sequence embeddings, heterogeneous similarity graphs, and attention-based GNNs to improve link prediction fidelity and mechanistic interpretability of gene expression.

DISCUSSION

Healthcare data is inherently multimodal, and integrating information from different sources can provide a more comprehensive view of a patient's health status or disease characteristics. Graph Neural Networks facilitate this by providing a framework to model relationships between and within each modality. The strengths of GNNs lie in their integration with other deep learning models by taking advantage of advanced fusion strategies, particularly those employing attention mechanisms. GNN integrations with CNNs, RNNs, autoencoders, language transformers, machine learning classification or regression models, and optimization algorithms facilitate multimodal data preprocessing and merging, as illustrated in the workflow of fusion types in Figure 1.

Across research areas and prediction tasks, intermediate fusion is the prevailing design (Figure 2 Panels A and B, and Figure 3 Panel B). In epidemic forecasting, temporal encoders fuse data via attention-based graph layers to capture local and global spread Qiu et al. (2024); Xie et al. (2022); Tran et al. (2024); Luo et al. (2025). Hospital-acquired infection models combine contact graphs with attention inside the graph pipeline Gouareb et al. (2023). COVID-19 outcome prediction uses intermediate fusion that joins CT features with population graphs with adaptable edges Keicher et al. (2023); Tariq et al. (2023, 2025). Clinical prediction and operations also favor intermediate fusion, where modality-specific encoders precede GraphSAGE, GCN, GAT, or heterogeneous GNN layers AL-Sabri et al. (2024); Tang et al. (2023b); Begum (2024); Boschi et al. (2024); Valls et al. (2023); Zhang et al. (2024); Christos Maroudis et al. (2025). Oncology mostly follows the same pattern, with late fusion used when independent embeddings are aligned after training and early fusion used when features are concatenated before graph reasoning Gao et al. (2022); Li et al. (2023b); Moharana et al. (2025); Alzoubi et al. (2024); Pratap Joshi et al. (2025); Peng et al. (2024); Yan et al. (2024). Gene expression studies implement sequence-aware intermediate fusion that mixes pretrained sequence embeddings with similarity graphs and attention Wang et al. (2023b, 2022); Wang and Chen (2023); Zhang et al. (2022).

Across the 85 studies reviewed, intermediate fusion accounts for 81% of models (n=69), with the highest use in neuropsychiatry (83%) and pharmacology (74%), and attention layers are present in over 60% of systems. Early fusion constitutes 15% (n=13), largely in oncology for raw feature concatenation. Late fusion appears in 1% (n=1) for embedding alignment in genomics and hybrid fusion in 2% (n=2), both in neuropsychiatry. Intermediate fusion is associated with the strongest outcomes, with top models reaching mean AUC values near 0.95 and accuracies near 0.92 (Table 8). Early fusion supports simpler feature integration with broader performance ranges (AUC 0.84–0.99), while late fusion suits alignment-driven tasks such as MGNN, where modality-specific embeddings are correlated only after independent training (AUC 0.98). Intermediate fusion consistently yields the most discriminative models, including Alzheimer's systems achieving AUC values up to 1.00, consistent with prior analyses of multimodal GNNs Paul et al. (2024); Li et al. (2023a).

228 In terms of datasets, population-level forecasting relies on datasets such as JHU CSSE, ILINet, OxCGRT,
229 and social media signals Qiu et al. (2024); Xie et al. (2022); Tran et al. (2024); Luo et al. (2025). Clinical
230 prediction is often validated on MIMIC III and MIMIC IV for mortality, readmission, sepsis, and length of
231 stay, and on institutional cohorts for triage and dynamic biomarker prediction AL-Sabri et al. (2024); Tang
232 et al. (2023b); Ghanvatkar and Rajan (2023); Christos Maroudis et al. (2025); Begum (2024); Zhang et al.
233 (2024). Imaging-heavy ophthalmology and retinal screening use APTOS and MESSIDOR and report gains
234 when CNN features are integrated into patient similarity or knowledge graphs Gao et al. (2024); Zedadra
235 et al. (2025). Oncology combines TCIA archive and disease-specific collections for radiology, whole slide
236 pathology, and multi-omic cohorts for survival modeling Peng et al. (2024); Alzoubi et al. (2024); Yan
237 et al. (2024); Gao et al. (2022). Gene regulatory and interaction studies rely on LncACTdb, LNCipedia,
238 miRBase, ncRNAsNP, HMDD, and GENCODE, which support sequence pretraining and heterogeneous
239 graph construction Wang et al. (2023b, 2022); Wang and Chen (2023); Zhang et al. (2022).

240 The most prevalent layer types include GraphSAGE, GCN, GAT, and heterogeneous GNNs. Temporal
241 encoders at the node level include LSTM, GRU, and temporal GNNs. Attention is used to weight neighbors
242 and modalities. In epidemic forecasting, temporal encoders feed attention-based graph layers Qiu et al.
243 (2024); Xie et al. (2022); Tran et al. (2024); Luo et al. (2025). In clinical prediction, GraphSAGE and
244 heterogeneous GNNs are combined with BiGRU or Transformer text encoders and time-aware designs
245 AL-Sabri et al. (2024); Tang et al. (2023b); Begum (2024); Boschi et al. (2024). In oncology, attention
246 GNNs integrate imaging and omics Alzoubi et al. (2024); Peng et al. (2024); Yan et al. (2024). Gene
247 interaction models pair GraphSAGE with Doc2Vec k-mer embeddings, CRF layers, and multi-channel
248 attention Wang et al. (2023b, 2022); Wang and Chen (2023); Zhang et al. (2022). Alzheimer's, COVID-19
249 Outcomes, and Drug-Target Prediction exhibit the highest layer type diversity, with 90%, 70%, and 60% of
250 the models respectively combining multiple layer types, reflecting their complex multimodal requirements,
251 as illustrated in the varied fusion strategies of Figure 3, Panel C. GNN + attention has the highest prevalence
252 across included studies (63%), with CNN/Conv following closely with an incidence of 40% across studies,
253 particularly in tasks like Alzheimer's and COVID-19 outcomes.

254 Forecasting tasks tend to model spatiotemporal data using intermediate fusion that aligns mobility and
255 case signals with graph dynamics Qiu et al. (2024); Xie et al. (2022); Tran et al. (2024); Luo et al. (2025).
256 Operational and clinical tasks embed structured EHR, notes, images, and vitals with modality-specific
257 encoders, which are fused in graph layers with attention AL-Sabri et al. (2024); Tang et al. (2023b);
258 Valls et al. (2023); Ghanvatkar and Rajan (2023); Zhang et al. (2024); Christos Maroudis et al. (2025).
259 Neuropsychiatric tasks combine temporal encoders with imaging, electrophysiology, language, and omics
260 within subject or population graphs with attention mechanisms Cai et al. (2025); Liu et al. (2024); Li et al.
261 (2024). Temporal encoders concentrate on time-dependent problems, including epidemic forecasting (75%)
262 and COVID-19 outcomes (67%). A large overlap between attention mechanisms and temporal encoders
263 has been observed in epidemic forecasting (75% attention; 75% temporal), ICU length of stay, ovarian
264 cancer, prescription prediction, sepsis trajectory modeling, and neurodegenerative disease (Figure 2, Panel
265 D and Figure 3, Panel A).

266 Attention mechanisms and modality-specific encoders such as CNNs, RNNs, and graph layers that retain
267 spatial, temporal, and relational structure correspond to higher predictive reliability across biomedical
268 settings (Table 8). Attention-based intermediate fusion appears in most high-performing systems, particularly
269 in tasks requiring integration of structured molecular features, clinical text, and imaging. Architectures
270 combining GraphSAGE, GCN, or heterogeneous GNN layers with temporal or vision encoders achieve

271 the strongest AUC and accuracy ranges in genomics, neuropsychiatry, and oncology. Domains with well-
272 defined structural priors, such as ncRNA–miRNA prediction and drug–drug interaction modeling, show
273 tighter performance bounds, whereas models operating on heterogeneous EHR or epidemiological data
274 exhibit broader variability.

275 This review has several limitations. Marked heterogeneity in cohorts and nomenclature limits cross-study
276 comparability and meta-analytic potential. Our harmonized taxonomy (early/intermediate/late fusion,
277 layer families) may introduce classification error for mixed or sparsely described architectures, and many
278 abstractions rely on self-reported methods without code or full graph-construction details. External validity
279 is often weak, since numerous studies lack external validation. Widely used datasets (e.g., MIMIC, ADNI,
280 ABIDE, and public KGs) may carry sampling biases that may hinder generalization. Finally, we did
281 not apply a formal risk-of-bias tool or rerun models, as the main scope of this review is to build an
282 understanding of how multimodal medical data is being integrated in GNNs.

CONCLUSION

283 GNNs offer a robust framework for modeling complex relationships across diverse modalities such
284 as electronic health records, medical imaging, genomic profiles, and clinical notes. By synthesizing
285 advancements in drug discovery, cancer detection, mental health diagnosis, epidemiology, clinical risk
286 prediction, and gene expression analysis, this review has highlighted GNNs' ability to enhance clinical
287 decision-making by leveraging graph-structured representations to capture intricate relationships among
288 patients, diseases, drugs, imaging, text, and biological entities. The integration of GNNs with deep
289 learning models, such as CNN, LSTM, RNN, dimensionality reduction, machine learning, and optimization
290 algorithms, enhances their ability to process diverse data modalities. Multiple fusion strategies, such as early,
291 intermediate, late, and hybrid, are employed to fuse multimodal data into a unified prediction framework.
292 However, data heterogeneity across modalities, varying in structure and noise levels, complicates graph
293 construction and fusion, while resource-intensive computations pose scalability issues. Interpretability and
294 causality are essential for clinical adoption, with attention-based mechanisms offering partial solutions but
295 requiring further development. Real-world use of multimodal GNNs also faces regulatory and operational
296 barriers. Many models rely on complex graph-construction choices and stochastic training procedures that
297 limit reproducibility across institutions, while the absence of standardized evaluation criteria complicates
298 regulatory review. Deployment requires attention to data governance, privacy compliance, and integration
299 with existing clinical workflows. Ensuring model generalizability across diverse datasets, addressing data
300 availability, and complying with ethical, privacy, and security regulations are additional constraints that are
301 yet to be fully addressed.

302 Several research directions follow from the patterns identified in this review. First, causal GNNs are
303 needed to disentangle mechanistic relations from observational correlations in multimodal biomedical
304 graphs, particularly for tasks such as treatment effect modeling, disease progression, and drug interaction
305 inference. Second, privacy-preserving federated graph learning is essential for cross-institutional
306 multimodal datasets. Third, the field lacks standardized explainability benchmarks for subgraph attribution,
307 modality-specific contribution, and stability under perturbation, which would allow systematic comparison
308 across fusion architectures. Lastly, future benchmarks should evaluate fusion strategies under controlled
309 data heterogeneity to determine when early, late, or hybrid designs offer measurable advantages to ensure
310 that multimodal GNNs are mechanistically informative, privacy-aligned, and reproducible at clinical scale.

CONFLICT OF INTEREST STATEMENT

311 The authors declare that the research was conducted in the absence of any commercial or financial
 312 relationships that could be construed as a potential conflict of interest.

FUNDING

313 The authors declare that no financial support was received for the research, authorship, and publication of
 314 this article.

Table 1 List of Acronyms

Acronym	Definition
ABIDE	Autism Brain Imaging Data Exchange
ADNI	Alzheimer's Disease Neuroimaging Initiative
ADHD-200	ADHD-200 Consortium neuroimaging dataset
ANIC	Australian National Intensive Care dataset (as cited)
APOTOS	Asia Pacific Tele-Ophthalmology Society diabetic retinopathy dataset
CBIS-DDSM	Curated Breast Imaging Subset of DDSM
CBHS	Commonwealth Bank Health Society (private insurer cohort)
CCLE	Cancer Cell Line Encyclopedia
CDC ILI ILINet	CDC Influenza-Like Illness / Outpatient ILI Surveillance Network
CMMMD	Chinese Mammography Database (as cited)
CTD	Comparative Toxicogenomics Database
DAIC-WOZ	Distress Analysis Interview Corpus—Wizard of Oz
DAUH	Dong-A University Hospital (as cited)
DAVIS	Kinase inhibitor binding benchmark (Davis et al.)
DDSM	Digital Database for Screening Mammography
DISNET	Disease Networks knowledge base
EUH	Emory University Hospital
FAERS	FDA Adverse Event Reporting System
GENCODE	Comprehensive gene annotation resource
GEO	Gene Expression Omnibus
GDSC	Genomics of Drug Sensitivity in Cancer
HCP	Human Connectome Project
HMDD	Human microRNA Disease Database
iCTCF	International COVID-19 CT dataset (as cited)
IQ-OTHNCCD	IQ-OTH/NCCD lung cancer imaging datasets
JHU CSSE	Johns Hopkins University CSSE COVID-19 repository
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIBA	Kinase Inhibitor BioActivity benchmark
KRI	Korea Research Institute COVID-19 cohort (as cited)
LncACTdb	Long Non-coding RNA–Associated Competing Endogenous RNA Database
LNCipedia	Long Non-Coding RNA knowledge base
MESSIDOR-2	Retinal fundus dataset for DR screening

Continued on next page

Acronym	Definition
miRBase	microRNA sequence database
MIMIC-IIIMIMIC-IV	Medical Information Mart for Intensive Care v3 / v4
MODMA	Multimodal Depression Dataset
MSBB	Mount Sinai Brain Bank
ncRNASNP	Non-coding RNA Single Nucleotide Polymorphisms database
OASIS	Open Access Series of Imaging Studies
OCD (Ovarian)	Ovarian Cancer Dataset
OxCGRT	Oxford COVID-19 Government Response Tracker
PDBP	Parkinson's Disease Biomarkers Program
PLCO	Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial
PPMI	Parkinson's Progression Markers Initiative
RepoDB	Drug repurposing database
REST-meta-MDD	REST-meta-MDD Consortium dataset
ROSMAP	Religious Orders Study and Memory and Aging Project
SEER	Surveillance, Epidemiology, and End Results (NCI)
SPAIN-COVID	Spain COVID epidemiological dataset (as cited)
STITCH	Search Tool for Interactions of Chemicals
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
Synthea	Synthetic patient EHR generator
TCIA	The Cancer Imaging Archive
TWOSIDES	Large drug–drug interaction side-effect dataset
WSI	Whole-Slide Images (pathology)
DTI (imaging)	Diffusion Tensor Imaging (distinct from Drug–Target Interaction)
EEG	Electroencephalography
fMRI, sMRI	Functional, Structural Magnetic Resonance Imaging
PET	Positron Emission Tomography
WSI	WholeSlide Images (pathology)
ADE	Adverse Drug Event
ASD	Autism Spectrum Disorder
DDI	Drug–Drug Interaction
DTI (task)	Drug–Target Interaction (disambiguated from imaging DTI)
DR	Diabetic Retinopathy
HAI	Healthcare-Associated Infection
ICU	Intensive Care Unit
LOS	Length of Stay
MDD	Major Depressive Disorder
BERT	Bidirectional Encoder Representations from Transformers
BioBERT	Biomedical BERT
BiGRU	Bidirectional Gated Recurrent Unit
BiLSTM	Bidirectional Long Short-Term Memory
CAL	Content-Aware Layer (paper-specific)
CNN	Convolutional Neural Network

Continued on next page

Acronym	Definition
CRF	Conditional Random Field
DNPGF	Dual-Nonlocal Pyramid Graph Filter (paper-specific)
GAT	Graph Attention Network
GCN	Graph Convolutional Network
GCNN	Graph Convolutional Neural Network (generic)
GGNN	Gated Graph Neural Network
GIN, GINConv	Graph Isomorphism Network / convolutional layer
GNN	Graph Neural Network
GNNRAI	GNN with Region-Aware Integration (paper-specific)
GraphSAGE	Graph Sample and Aggregate
Graph Transformer	Transformer architecture on graphs
GTN	Graph Transformer Network (define in text; some papers vary)
HGAT	Heterogeneous Graph Attention Network
HCNN-MAFN	Hypergraph CNN with Multimodal Attention Fusion Network (paper-specific)
HGNN	Heterogeneous Graph Neural Network
HeteroGCN	Heterogeneous Graph Convolutional Network
ICA	Independent Component Analysis
KNN	k-Nearest Neighbors
LAYOUTLM	Document layout-aware Transformer
LINE	Large-scale Information Network Embedding
LLM	Large Language Model
LSTM	Long Short-Term Memory
MacBERT	Chinese BERT variant
MLP	Multi-Layer Perceptron
RGCN	Relational Graph Convolutional Network
RNN	Recurrent Neural Network
RPN	Region Proposal Network
SCL	Semantic Convolutional Layer (paper-specific)
SDNE	Structural Deep Network Embedding
ST-GNN / STGCN	Spatio-Temporal GNN / Spatio-Temporal GCN
Transformer	Self-attention neural network
U-Net	U-shaped convolutional encoder-decoder
VAE	Variational Autoencoder
VGAE	Variational Graph Autoencoder
ViT	Vision Transformer
VSA	Variational Spatial Attention (paper-specific)
Early, Intermediate, Late Fusion	Fusion timing categories used in this review
GAN	Generative Adversarial Network
GAT (attention)	Graph attention mechanism/layer
KG	Knowledge Graph
KGE	Knowledge Graph Embedding

Continued on next page

Acronym	Definition
Q-Learning (RL)	Reinforcement Learning Q-learning
RL	Reinforcement Learning
ROI	Region of Interest

DATA AVAILABILITY STATEMENT

315 Code and data are publicly available at <https://doi.org/10.5281/zenodo.17202525>.

REFERENCES

- 316 Abdeddaiem, O., Zaatra, A., Bessodok, A., and Yanes, N. (2025). Adversarial graph neural network
317 for medication recommendation (agmr). In *International COnference on Decision Aid and Artificial
318 Intelligence (ICODAI 2024)*. 47–61
- 319 Al-Rabeah, M. H. and Lakizadeh, A. (2022). Prediction of drug-drug interaction events using graph neural
320 networks based feature extraction. *Scientific Reports* 12, 15590. doi:10.1038/s41598-022-19999-4
- 321 AL-Sabri, R., Gao, J., Chen, J., Oloulade, B. M., Wu, Z., Abdullah, M., et al. (2024). M3GNAS:
322 Multi-modal Multi-view Graph Neural Architecture Search for Medical Outcome Predictions. In *2024
323 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (Lisbon, Portugal: IEEE),
324 1783–1788. doi:10.1109/BIBM62325.2024.10821927
- 325 Alzoubi, I., Zhang, L., Zheng, Y., Loh, C., Wang, X., and Graeber, M. B. (2024). PathoGraph: An
326 Attention-Based Graph Neural Network Capable of Prognostication Based on CD276 Labelling of
327 Malignant Glioma Cells. *Cancers* 16, 750. doi:10.3390/cancers16040750
- 328 Artiñano-Muñoz, R., Prieto-Santamaría, L., Pérez-Pérez, A., and Rodríguez-González, A. (2024).
329 DRAGON: Drug Repurposing via Graph Neural Networks with Drug and Protein Embeddings as
330 Features. In *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*
331 (Guadalajara, Mexico: IEEE), 170–175. doi:10.1109/CBMS61543.2024.00036
- 332 A.Tabatabaei, A., Mahdavi, M. E., Beiranvand, E., Gharaghani, S., and Adibi, P. (2025). Graph Neural
333 Network-Based Approaches to Drug Repurposing: A Comprehensive Survey doi:10.31224/4410
- 334 Begum, U. S. (2024). Federated And Multi-Modal Learning Algorithms for Healthcare and Cross-Domain
335 Analytics. *PatternIQ Mining* 1. doi:10.70023/sahd/241104
- 336 Boll, H. O., Amirahmadi, A., Soliman, A., Byttner, S., and Recamonde-Mendoza, M. (2024). Graph
337 Neural Networks for Heart Failure Prediction on an EHR-Based Patient Similarity Graph doi:10.48550/
338 arXiv.2411.19742. ArXiv:2411.19742 [cs]
- 339 Boschi, T., Bonin, F., Ordonez-Hurtado, R., Rousseau, C., Pascale, A., and Dinsmore, J. (2024). Functional
340 Graph Convolutional Networks: A unified multi-task and multi-modal learning framework to facilitate
341 health and social-care insights. In *Proceedings of the Thirty-Third International Joint Conference on
342 Artificial Intelligence*. 7188–7196. doi:10.24963/ijcai.2024/795. ArXiv:2403.10158 [cs]
- 343 Cai, H., Huang, X., Liu, Z., Liao, W., Dai, H., Wu, Z., et al. (2023). Exploring multimodal approaches
344 for alzheimer’s disease detection using patient speech transcript and audio data. *arXiv preprint
345 arXiv:2307.02514*
- 346 Cai, L., Zeng, W., Chen, H., Zhang, H., Li, Y., Feng, Y., et al. (2025). MM-GTUNets: Unified Multi-Modal
347 Graph Deep Learning for Brain Disorders Prediction doi:10.48550/arXiv.2406.14455. ArXiv:2406.14455
348 [cs]
- 349 Cetin, S. and Sefer, E. (2025). A graphlet-based explanation generator for graph neural networks over
350 biological datasets. *Current Bioinformatics* 20, 840–851

- 351 Chan, Y. H., Wang, C., Soh, W. K., and Rajapakse, J. C. (2022). Combining Neuroimaging and Omics
352 Datasets for Disease Classification Using Graph Neural Networks. *Frontiers in Neuroscience* 16, 866666.
353 doi:10.3389/fnins.2022.866666
- 354 Christos Maroudis, A., Karathanasopoulou, K., Stylianides, C. C., Dimitrakopoulos, G., and Panayides,
355 A. S. (2025). Fairness-Aware Graph Neural Networks for ICU Length of Stay Prediction in IoT-Enabled
356 Environments. *IEEE Access* 13, 64516–64533. doi:10.1109/ACCESS.2025.3560180
- 357 Cui, L., Seo, H., Tabar, M., Ma, F., Wang, S., and Lee, D. (2020). DETERRENT: Knowledge Guided
358 Graph Attention Network for Detecting Healthcare Misinformation. In *Proceedings of the 26th ACM*
359 *SIGKDD International Conference on Knowledge Discovery & Data Mining* (Virtual Event CA USA:
360 ACM), 492–502. doi:10.1145/3394486.3403092
- 361 Dawn, S., Chakraborty, M., Maulik, U., and Bandyopadhyay, S. (2024). Adverse Drug Event Prediction
362 with a Multi-Layer Heterogeneous Graph Neural Network Architecture. In *2024 IEEE 21st India Council*
363 *International Conference (INDICON)* (Kharagpur, India: IEEE), 1–6. doi:10.1109/INDICON63790.
364 2024.10958347
- 365 D’Souza, N. S., Wang, H., Giovannini, A., Foncubierta-Rodriguez, A., Beck, K. L., Boyko, O., et al.
366 (2023). MaxCorrMGNN: A Multi-Graph Neural Network Framework for Generalized Multimodal
367 Fusion of Medical Data for Outcome Prediction doi:10.48550/arXiv.2307.07093
- 368 Dumyn, I., Basystiuk, O., and Dumyn, A. (2024). Graph-based approaches for multimodal medical data
369 processing. In *Proceedings of the 7th International Conference on Informatics & Data-Driven Medicine*
370 *Birmingham, United Kingdom*. 349–362
- 371 D’Souza, N. S., Nebel, M. B., Crocetti, D., and Robin, J. (2021). M-GCN: A Multimodal
372 Graph Convolutional Network to Integrate Functional and Structural Connectomics Data to Predict
373 Multidimensional Phenotypic Characterizations
- 374 Gan, Y., Liu, W., Xu, G., Yan, C., and Zou, G. (2023). Dmfddi: deep multimodal fusion for drug–drug
375 interaction prediction. *Briefings in Bioinformatics* 24, bbad397
- 376 Gao, J., Lyu, T., Xiong, F., Wang, J., Ke, W., and Li, Z. (2022). Predicting the Survival of Cancer Patients
377 With Multimodal Graph Neural Network. *IEEE/ACM Transactions on Computational Biology and*
378 *Bioinformatics* 19, 699–709. doi:10.1109/TCBB.2021.3083566
- 379 Gao, J., Tang, H., Wang, Z., Li, Y., Luo, N., Song, M., et al. (2025a). Graph Neural Networks and
380 Multimodal DTI Features for Schizophrenia Classification: Insights from Brain Network Analysis and
381 Gene Expression. *Neuroscience Bulletin* doi:10.1007/s12264-025-01385-5
- 382 Gao, W., Rong, F., Shao, L., Deng, Z., Xiao, D., Zhang, R., et al. (2024). Enhancing ophthalmology
383 medical record management with multi-modal knowledge graphs. *Scientific Reports* 14, 23221. doi:10.
384 1038/s41598-024-73316-9
- 385 Gao, Y., Zhang, X., Sun, Z., Chandak, P., Bu, J., and Wang, H. (2025b). Precision Adverse Drug
386 Reactions Prediction with Heterogeneous Graph Neural Network. *Advanced Science* 12, 2404671.
387 doi:10.1002/advs.202404671
- 388 Ghantasala, G. S. P., Dilip, K., Vidyullatha, P., Allabun, S., Alqahtani, M. S., Othman, M., et al. (2024).
389 Enhanced ovarian cancer survival prediction using temporal analysis and graph neural networks. *BMC*
390 *Medical Informatics and Decision Making* 24, 299. doi:10.1186/s12911-024-02665-2
- 391 Ghanvatkar, S. and Rajan, V. (2023). Graph-Based Patient Representation for Multimodal Clinical Data:
392 Addressing Data Heterogeneity doi:10.1101/2023.12.07.23299673
- 393 Gouareb, R., Bornet, A., Proios, D., Pereira, S. G., and Teodoro, D. (2023). Detection of Patients at Risk
394 of Multidrug-Resistant Enterobacteriaceae Infection Using Graph Neural Networks: A Retrospective
395 Study. *Health Data Science* 3, 0099. doi:10.34133/hds.0099

- 396 Gowri, B. S., M, S., Gehlot, Y., and Varma, V. C. (2024). Deep Fusion of Vision Transformers, Graph
397 Neural Networks, and LayoutLM for Enhanced Multimodal Detection of Lung Cancer: A Novel
398 Approach in Computational Oncology. In *2024 International Conference on Computing and Data
399 Science (ICCDs)* (Chennai, India: IEEE), 1–6. doi:10.1109/ICCDs60734.2024.10560446
- 400 Gu, Y., Peng, S., Li, Y., Gao, L., and Dong, Y. (2025). FC-HGNN: A heterogeneous graph neural network
401 based on brain functional connectivity for mental disorder identification. *Information Fusion* 113,
402 102619. doi:10.1016/j.inffus.2024.102619
- 403 Hsieh, K., Wang, Y., Chen, L., Zhao, Z., Savitz, S., Jiang, X., et al. (2021). Drug repurposing for
404 COVID-19 using graph neural network and harmonizing multiple evidence. *Scientific Reports* 11, 23179.
405 doi:10.1038/s41598-021-02353-5
- 406 Hsieh, K.-L., Wang, Y., Chen, L., Zhao, Z., Savitz, S., Jiang, X., et al. (2020). Drug Repurposing for
407 COVID-19 using Graph Neural Network with Genetic, Mechanistic, and Epidemiological Validation
408 doi:10.21203/rs.3.rs-114758/v1
- 409 Huang, T., Lin, K.-H., Machado-Vieira, R., Soares, J. C., Jiang, X., and Kim, Y. (2023). Explainable drug
410 side effect prediction via biologically informed graph neural network doi:10.1101/2023.05.26.23290615
- 411 Jiang, H., Chen, P., Sun, Z., Liang, C., Xue, R., Zhao, L., et al. (2023). Assisting schizophrenia
412 diagnosis using clinical electroencephalography and interpretable graph neural networks: a real-world
413 and cross-site study. *Neuropsychopharmacology* 48, 1920–1930. doi:10.1038/s41386-023-01658-5
- 414 Johnson, R., Li, M. M., Noori, A., Queen, O., and Zitnik, M. (2024). Graph Artificial
415 Intelligence in Medicine. *Annual Review of Biomedical Data Science* 7, 345–368. doi:10.1146/
416 annurev-biodatasci-110723-024625
- 417 Kalla, A., Mukhopadhyay, S., Ralte, Z., and Kar, I. (2023). Exploring the impact of motif-driven causal
418 temporal analysis using graph neural network in improving large language model performance for
419 pharmacovigilance. In *2023 9th International Conference on Advanced Computing and Communication
420 Systems (ICACCS)* (IEEE), vol. 1, 1769–1776
- 421 Keicher, M., Burwinkel, H., Bani-Harouni, D., Paschali, M., Czempiel, T., Burian, E., et al. (2023).
422 Multimodal graph attention network for COVID-19 outcome prediction. *Scientific Reports* 13, 19539.
423 doi:10.1038/s41598-023-46625-8
- 424 Khemani, B., Malave, S., Patil, S., Shilotri, N., Varma, S., Vishwakarma, V., et al. (2024). Sentimatrix:
425 sentiment analysis using GNN in healthcare. *International Journal of Information Technology* 16,
426 5213–5219. doi:10.1007/s41870-024-02142-z
- 427 Kim, S., Lee, N., Lee, J., Hyun, D., and Park, C. (2023). Heterogeneous Graph Learning for Multi-Modal
428 Medical Data Analysis. *Proceedings of the AAAI Conference on Artificial Intelligence* 37, 5141–5150.
429 doi:10.1609/aaai.v37i4.25643
- 430 Kulandaivelu, G., Taluja, A., Gawas, M., and Kumar Nath, R. (2024). Automated breast cancer
431 diagnosis optimized with higher-order attribute-enhancing heterogeneous graph neural networks using
432 mammogram images. *Biomedical Signal Processing and Control* 97, 106659. doi:10.1016/j.bspc.2024.
433 106659
- 434 Kumar, A., Nashte, A., Porwal Amit, R., and Choudhary, C. (2023a). Hypergraph Neural Networks
435 with Attention-based Fusion for Multimodal Medical Data Integration and Analysis. In *2023 Seventh
436 International Conference on Image Information Processing (ICIIP)* (Solan, India: IEEE), 628–633.
437 doi:10.1109/ICIIP61524.2023.10537751
- 438 Kumar, R., Verma, D., Raj, J. R. F., Rao, A. L. N., Chari, S. L., and Khan, A. K. (2023b). Graph
439 Convolutional Networks For Disease Mapping and Classification in Healthcare. In *2023 International
440 Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIHI)* (Raipur, India:

- 441 IEEE), 1–7. doi:10.1109/ICAIH157871.2023.10489220
442 Lee, D.-J., Shin, D.-H., Son, Y.-H., Han, J.-W., Oh, J.-H., Kim, D.-H., et al. (2024a). Spectral Graph
443 Neural Network-Based Multi-Atlas Brain Network Fusion for Major Depressive Disorder Diagnosis.
444 *IEEE Journal of Biomedical and Health Informatics* 28, 2967–2978. doi:10.1109/JBHI.2024.3366662
445 Lee, G.-B., Jeong, Y.-J., Kang, D.-Y., Yun, H.-J., and Yoon, M. (2024b). Multimodal feature fusion-based
446 graph convolutional networks for Alzheimer's disease stage classification using F-18 florbetaben brain
447 PET images and clinical indicators. *PLOS ONE* 19, e0315809. doi:10.1371/journal.pone.0315809
448 Li, M., Sun, X., and Wang, M. (2024). Detecting Depression With Heterogeneous Graph Neural Network
449 in Clinical Interview Transcript. *IEEE Transactions on Computational Social Systems* 11, 1315–1324.
450 doi:10.1109/TCSS.2023.3263056
451 Li, M. M., Huang, K., and Zitnik, M. (2022). Graph representation learning in biomedicine and healthcare.
452 *Nature Biomedical Engineering* 6, 1353–1369. doi:10.1038/s41551-022-00942-x
453 Li, R., Yuan, X., Radfar, M., Marendy, P., Ni, W., O'Brien, T. J., et al. (2023a). Graph Signal Processing,
454 Graph Neural Network and Graph Learning on Biological Data: A Systematic Review. *IEEE Reviews in
455 Biomedical Engineering* 16, 109–135. doi:10.1109/RBME.2021.3122522
456 Li, R., Zhou, L., Wang, Y., Shan, F., Chen, X., and Liu, L. (2023b). A graph neural network model for
457 the diagnosis of lung adenocarcinoma based on multimodal features and an edge-generation network.
458 *Quantitative Imaging in Medicine and Surgery* 13, 5333–5348. doi:10.21037/qims-23-2
459 Li, Y., Sun, C., and Dong, Y. (2025). A Novel Audio-Visual Multimodal Semi-Supervised Model Based
460 on Graph Neural Networks for Depression Detection. In *ICASSP 2025 - 2025 IEEE International
461 Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Hyderabad, India: IEEE), 1–5.
462 doi:10.1109/ICASSP49660.2025.10888673
463 Lian, J., Luo, X., Shan, C., Han, D., Vardhanabuti, V., and Li, D. (2023). AdaMedGraph: Adaboosting
464 Graph Neural Networks for Personalized Medicine doi:10.48550/arXiv.2311.14304. ArXiv:2311.14304
465 [cs]
466 Lin, K.-H., Hsieh, K.-L., Jiang, X., and Kim, Y. (2023). Integrating Comorbidity Knowledge for
467 Alzheimer's Disease Drug Repurposing using Multi-task Graph Neural Network
468 Liu, S., Li, T., Ding, H., Tang, B., Wang, X., Chen, Q., et al. (2020). A hybrid method of recurrent neural
469 network and graph neural network for next-period prescription prediction. *International Journal of
470 Machine Learning and Cybernetics* 11, 2849–2856. doi:10.1007/s13042-020-01155-x
471 Liu, S., Zhou, J., Zhu, X., Zhang, Y., Zhou, X., Zhang, S., et al. (2024). An objective quantitative diagnosis
472 of depression using a local-to-global multimodal fusion graph neural network. *Patterns* 5, 101081.
473 doi:10.1016/j.patter.2024.101081
474 Liu, W., Yin, L., Wang, C., Liu, F., and Ni, Z. (2021). Multitask Healthcare Management Recommendation
475 System Leveraging Knowledge Graph. *Journal of Healthcare Engineering* 2021, 1–12. doi:10.1155/
476 2021/1233483
477 Luo, J., Wang, X., Fan, X., He, Y., Du, X., Chen, Y.-Q., et al. (2025). A novel graph neural network based
478 approach for influenza-like illness nowcasting: exploring the interplay of temporal, geographical, and
479 functional spatial features. *BMC Public Health* 25, 408. doi:10.1186/s12889-025-21618-6
480 Luo, Y., Chen, H., Yin, T., Horng, S.-J., and Li, T. (2024). Dual hypergraphs with feature weighted
481 and latent space learning for the diagnosis of Alzheimer's disease. *Information Fusion* 112, 102546.
482 doi:10.1016/j.inffus.2024.102546
483 Moharana, S. K., Sethi, N., and Punuri, S. B. (2025). Advancing Liver Disease Prediction with Multi-Modal
484 Graph Neural Networks and Federated Meta-Learning. In *2025 International Conference on Intelligent
485 Computing and Control Systems (ICICCS)* (Erode, India: IEEE), 692–697. doi:10.1109/ICICCS65191.

- 486 2025.10985240
- 487 Nguyen, A. D., Pham, H. H., Trung, H. T., Nguyen, Q. V. H., Truong, T. N., and Nguyen, P. L. (2023). High
488 accurate and explainable multi-pill detection framework with graph neural network-assisted multimodal
489 data fusion. *PLOS ONE* 18, e0291865. doi:10.1371/journal.pone.0291865
- 490 Nye, L. (2023). Digital Twins for Patient Care via Knowledge Graphs and Closed-Form Continuous-Time
491 Liquid Neural Networks
- 492 OMKUMARCHANDRAUMAKANTHAM, S. and PATHAK, V. (2024). Detecting side effects of adverse
493 drug reactions through drug-drug interactions using graph neural networks and self-supervised learning
- 494 Oss Boll, H., Amirahmadi, A., Ghazani, M. M., Morais, W. O. D., Freitas, E. P. D., Soliman, A., et al.
495 (2024). Graph neural networks for clinical risk prediction based on electronic health records: A survey.
496 *Journal of Biomedical Informatics* 151, 104616. doi:10.1016/j.jbi.2024.104616
- 497 Pablo, J., Gonzaliam, M., and Safaei, M. (2024). Graph neural networks for modeling disease relationships:
498 A framework for multi-disease diagnostics and comorbidity prediction
- 499 Paul, S. G., Saha, A., Hasan, M. Z., Noori, S. R. H., and Moustafa, A. (2024). A Systematic Review
500 of Graph Neural Network in Healthcare-Based Applications: Recent Advances, Trends, and Future
501 Directions. *IEEE Access* 12, 15145–15170. doi:10.1109/ACCESS.2024.3354809
- 502 Peng, J., Peng, L., Zhou, Z., Han, X., Xu, H., Lu, L., et al. (2024). Multi-Level fusion graph neural
503 network: Application to PET and CT imaging for risk stratification of head and neck cancer. *Biomedical
504 Signal Processing and Control* 92, 106137. doi:10.1016/j.bspc.2024.106137
- 505 Pratap Joshi, K., Gowda, V. B., Bidare Divakarachari, P., Siddappa Parameshwarappa, P., and Patra, R. K.
506 (2025). VSA-GCNN: Attention Guided Graph Neural Networks for Brain Tumor Segmentation and
507 Classification. *Big Data and Cognitive Computing* 9, 29. doi:10.3390/bdcc9020029
- 508 Qiu, M., Tan, Z., and Bao, B.-k. (2024). MSGNN: Multi-scale Spatio-temporal Graph Neural Network
509 for Epidemic Forecasting. *Data Mining and Knowledge Discovery* 38, 2348–2376. doi:10.1007/
510 s10618-024-01035-w. ArXiv:2308.15840 [cs]
- 511 Rajabi, E. and Kafaie, S. (2022). Knowledge Graphs and Explainable AI in Healthcare. *Information* 13,
512 459. doi:10.3390/info13100459
- 513 Sangeetha, S. K. B., Immanuel, R. R., Mathivanan, S. K., Cho, J., and Easwaramoorthy, S. V. (2024).
514 An Empirical Analysis of Multimodal Affective Computing Approaches for Advancing Emotional
515 Intelligence in Artificial Intelligence for Healthcare. *IEEE Access* 12, 114416–114434. doi:10.1109/
516 ACCESS.2024.3444494
- 517 Sefer, E. (2025a). Anomaly detection via graph contrastive learning. In *Proceedings of the 2025 SIAM
518 International Conference on Data Mining (SDM)* (SIAM), 51–60
- 519 Sefer, E. (2025b). Drgat: Predicting drug responses via diffusion-based graph attention network. *Journal
520 of Computational Biology* 32, 330–350
- 521 Stahlschmidt, S. R., Ulfenborg, B., and Synnergren, J. (2022). Multimodal deep learning for biomedical
522 data fusion: a review. *Briefings in Bioinformatics* 23, bbab569. doi:10.1093/bib/bbab569
- 523 Tang, C., Wei, M., Sun, J., Wang, S., Zhang, Y., Initiative, A. D. N., et al. (2023a). Csagp: detecting
524 alzheimer's disease from multimodal images via dual-transformer with cross-attention and graph pooling.
525 *Journal of King Saud University-Computer and Information Sciences* 35, 101618
- 526 Tang, S., Tariq, A., Dunnmon, J., Sharma, U., Elugunti, P., Rubin, D., et al. (2023b). Multimodal
527 spatiotemporal graph neural networks for improved prediction of 30-day all-cause hospital readmission
- 528 Tariq, A., Kaur, G., Su, L., Gichoya, J., Patel, B., and Banerjee, I. (2023). Generalizable Model Design for
529 Clinical Event Prediction using Graph Neural Networks doi:10.1101/2023.03.22.23287599

- 530 Tariq, A., Kaur, G., Su, L., Gichoya, J., Patel, B., and Banerjee, I. (2025). Adaptable graph neural networks
531 design to support generalizability for clinical event prediction. *Journal of Biomedical Informatics* 163,
532 104794. doi:10.1016/j.jbi.2025.104794
- 533 Teoh, J. R., Dong, J., Zuo, X., Lai, K. W., Hasikin, K., and Wu, X. (2024). Advancing healthcare through
534 multimodal data fusion: a comprehensive review of techniques and applications. *PeerJ Computer Science*
535 10, e2298. doi:10.7717/peerj-cs.2298
- 536 Tran, K.-T., Hy, T. S., Jiang, L., and Vu, X.-S. (2024). MGLEP: Multimodal Graph Learning for Modeling
537 Emerging Pandemics with Big Data. *Scientific Reports* 14, 16377. doi:10.1038/s41598-024-67146-y
- 538 Tripathy, R. K., Frohock, Z., Wang, H., Cary, G. A., Keegan, S., Carter, G. W., et al. (2025). Effective
539 integration of multi-omics with prior knowledge to identify biomarkers via explainable graph neural
540 networks. *npj Systems Biology and Applications* 11, 43. doi:10.1038/s41540-025-00519-9
- 541 Valls, V., Zayats, M., and Pascale, A. (2023). Information Flow in Graph Neural Networks: A Clinical
542 Triage Use Case doi:10.48550/arXiv.2309.06081
- 543 Venkatapathy, S., Votinov, M., Wagels, L., Kim, S., Lee, M., Habel, U., et al. (2023). Ensemble graph
544 neural network model for classification of major depressive disorder using whole-brain functional
545 connectivity. *Frontiers in Psychiatry* 14, 1125339. doi:10.3389/fpsyg.2023.1125339
- 546 Vijay Anand, R., Shanmuga Priyan, T., Guru Brahmam, M., Balusamy, B., and Benedetto, F. (2024).
547 IMNMAGN: Integrative Multimodal Approach for Enhanced Detection of Neurodegenerative Diseases
548 Using Fusion of Multidomain Analysis With Graph Networks. *IEEE Access* 12, 73095–73112. doi:10.
549 1109/ACCESS.2024.3403860
- 550 Wang, G., Guo, Z., You, G., Xu, M., Cao, C., and Hu, X. (2024a). MMDDI-MGPFF: Multi-Modal Drug
551 Representation Learning with Molecular Graph and Pharmacological Feature Fusion for Drug-Drug
552 Interaction Event Prediction. In *2024 IEEE International Conference on Bioinformatics and Biomedicine*
553 (*BIBM*) (Lisbon, Portugal: IEEE), 467–470. doi:10.1109/BIBM62325.2024.10822838
- 554 Wang, H., Qiu, X., Li, B., Tan, X., and Huang, J. (2025). Multimodal heterogeneous graph fusion for
555 automated obstructive sleep apnea-hypopnea syndrome diagnosis. *Complex & Intelligent Systems* 11, 44.
556 doi:10.1007/s40747-024-01648-0
- 557 Wang, R.-H., Luo, T., Zhang, H.-L., and Du, P.-F. (2023a). Pla-gnn: computational inference of protein
558 subcellular location alterations under drug treatments with deep graph neural networks. *Computers in*
559 *Biology and Medicine* 157, 106775
- 560 Wang, W. and Chen, H. (2023). Predicting miRNA-disease associations based on lncRNA-miRNA
561 interactions and graph convolution networks. *Briefings in Bioinformatics* 24, bbac495. doi:10.1093/bib/
562 bbac495
- 563 Wang, W., Zhang, L., Sun, J., Zhao, Q., and Shuai, J. (2022). Predicting the potential human
564 lncRNA-miRNA interactions based on graph convolution network with conditional random field.
565 *Briefings in Bioinformatics* 23, bbac463. doi:10.1093/bib/bbac463
- 566 Wang, X., Zhang, X., Chen, Y., and Yang, X. (2024b). IFC-GNN: Combining interactions of functional
567 connectivity with multimodal graph neural networks for ASD brain disorder analysis. *Alexandria*
568 *Engineering Journal* 98, 44–55. doi:10.1016/j.aej.2024.04.023
- 569 Wang, Y. (2022). Investigating US Healthcare Referral System with Graph Neural Networks. In *2022 4th*
570 *International Conference on Frontiers Technology of Information and Computer (ICFTIC)* (Qingdao,
571 China: IEEE), 767–773. doi:10.1109/ICFTIC57696.2022.10075286
- 572 Wang, Y., Hou, W., Sheng, N., Zhao, Z., Liu, J., Huang, L., et al. (2024c). Graph pooling in graph
573 neural networks: methods and their applications in omics studies. *Artificial Intelligence Review* 57, 294.
574 doi:10.1007/s10462-024-10918-9

- 575 Wang, Z., Bao, R., Wu, Y., Liu, G., Yang, L., Zhan, L., et al. (2024d). A Self-guided Multimodal Approach
576 to Enhancing Graph Representation Learning for Alzheimer's Diseases doi:10.48550/arXiv.2412.06212.
577 ArXiv:2412.06212 [cs]
- 578 Wang, Z., Liang, S., Liu, S., Meng, Z., Wang, J., and Liang, S. (2023b). Sequence pre-training-based graph
579 neural network for predicting lncRNA-miRNA associations. *Briefings in Bioinformatics* 24, bbad317.
580 doi:10.1093/bib/bbad317
- 581 Waqas, A., Tripathi, A., Ramachandran, R. P., Stewart, P. A., and Rasool, G. (2024). Multimodal data
582 integration for oncology in the era of deep neural networks: a review. *Frontiers in Artificial Intelligence*
583 7, 1408843. doi:10.3389/frai.2024.1408843
- 584 Xiang, Y., Li, X., Gao, Q., Xia, J., and Yue, Z. (2025). ExplainMIX: Explaining Drug Response Prediction
585 in Directed Graph Neural Networks with Multi-Omics Fusion. *IEEE Journal of Biomedical and Health*
586 *Informatics*, 1–11doi:10.1109/JBHI.2025.3550353
- 587 Xie, F., Zhang, Z., Li, L., Zhou, B., and Tan, Y. (2022). EpiGNN: Exploring Spatial Transmission with
588 Graph Neural Network for Regional Epidemic Forecasting doi:10.48550/arXiv.2208.11517
- 589 Xing, T., Dou, Y., Chen, X., Zhou, J., Xie, X., and Peng, S. (2024). An adaptive multi-graph neural
590 network with multimodal feature fusion learning for MDD detection. *Scientific Reports* 14, 28400.
591 doi:10.1038/s41598-024-79981-0
- 592 Xiong, Z., Liu, S., Huang, F., Wang, Z., Liu, X., Zhang, Z., et al. (2023). Multi-Relational Contrastive
593 Learning Graph Neural Network for Drug-Drug Interaction Event Prediction. *Proceedings of the AAAI*
594 *Conference on Artificial Intelligence* 37, 5339–5347. doi:10.1609/aaai.v37i4.25665
- 595 Xu, X., Li, J., Zhu, Z., Zhao, L., Wang, H., Song, C., et al. (2024). A Comprehensive Review on
596 Synergy of Multi-Modal Data and AI Technologies in Medical Diagnosis. *Bioengineering* 11, 219.
597 doi:10.3390/bioengineering11030219
- 598 Yan, Y., He, S., Yu, Z., Yuan, J., Liu, Z., and Chen, Y. (2024). Investigation of Customized Medical
599 Decision Algorithms Utilizing Graph Neural Networks. In *2024 IEEE 2nd International Conference*
600 *on Sensors, Electronics and Computer Engineering (ICSECE)* (Jinzhou, China: IEEE), 1238–1245.
601 doi:10.1109/ICSECE61636.2024.10729331
- 602 Yao, R., Shen, Z., Xu, X., Ling, G., Xiang, R., Song, T., et al. (2024). Knowledge mapping of graph neural
603 networks for drug discovery: a bibliometric and visualized analysis. *Frontiers in Pharmacology* 15,
604 1393415. doi:10.3389/fphar.2024.1393415
- 605 Yella, J. K., Ghandikota, S. K., and Jegga, A. G. (2022). GraMDTA: Multimodal Graph Neural Networks
606 for Predicting Drug-Target Associations
- 607 Yu, H., Li, K., Dong, W., Song, S., Gao, C., and Shi, J. (2023). Attention-based cross domain graph
608 neural network for prediction of drug–drug interactions. *Briefings in Bioinformatics* 24, bbad155.
609 doi:10.1093/bib/bbad155
- 610 Zedadra, A., Zedadra, O., Yassine Salah-Salah, M., and Guerrieri, A. (2025). Graph-Aware Multimodal
611 Deep Learning for Classification of Diabetic Retinopathy Images. *IEEE Access* 13, 74799–74810.
612 doi:10.1109/ACCESS.2025.3564529
- 613 Zhang, H., Wang, Y., Pan, Z., Sun, X., Mou, M., Zhang, B., et al. (2022). ncRNAInter: a novel strategy
614 based on graph neural network to discover interactions between lncRNA and miRNA. *Briefings in*
615 *Bioinformatics* 23, bbac411. doi:10.1093/bib/bbac411
- 616 Zhang, S., Yang, J., Zhang, Y., Zhong, J., Hu, W., Li, C., et al. (2023a). The Combination of a Graph
617 Neural Network Technique and Brain Imaging to Diagnose Neurological Disorders: A Review and
618 Outlook. *Brain Sciences* 13, 1462. doi:10.3390/brainsci13101462

- 619 Zhang, S., Yang, K., Liu, Z., Lai, X., Yang, Z., Zeng, J., et al. (2023b). Drugai: a multi-view deep
620 learning model for predicting drug–target activating/inhibiting mechanisms. *Briefings in bioinformatics*
621 24, bbac526
- 622 Zhang, Z., Li, X., Teng, F., Lin, N., Zhu, X., Wang, X., et al. (2024). Out-of-Distribution Generalized
623 Dynamic Graph Neural Network for Human Albumin Prediction doi:10.48550/arXiv.2311.15545.
624 ArXiv:2311.15545 [cs]
- 625 Zhou, F., Khushi, M., Brett, J., and Uddin, S. (2024). Graph neural network-based subgraph analysis
626 for predicting adverse drug events. *Computers in Biology and Medicine* 183, 109282. doi:10.1016/j.
627 compbiomed.2024.109282

Table 2. Graph-based models across pharmacology-related tasks.

Task	Model	Fusion	Dataset	Layers	AUC	F1	Accuracy	
Adverse Drug Event Gao et al. (2025b)	PreciseADR	Early	FAERS adverse drug events, demographics, diseases, drugs	HGNN	0.54–0.84	NR	NR	
Adverse Drug Event Kalla et al. (2023)	MDTCKGNN	Intermediate	ade_corpus_v2; PHEE	HGNN with attention	NR	NR	NR	
Adverse Drug Event Zhou et al. (2024)	Patient-centric GNN	Early	Australian CBHS	GraphSAGE with attention	0.88–0.96	0.53–0.90	0.89–0.94	
Adverse Drug Event Huang et al. (2023)	HHAN-DSI	Intermediate	SIDER, OFFSIDE, GO	HGNN + TransE	0.84–0.94	NR	NR	
Adverse Drug Event Dawn et al. (2024)	MI-GNN	Intermediate	Decagon DDI	HGNN	0.81–0.95	NR	NR	
Colonization Risk Prediction Gouareb et al. (2023)	Spatiotemporal model	Early	MIMIC-III	GCN, GAT, GraphSAGE	0.89–0.96	NR	0.03–0.96	
Drug Repurposing Artiñano-Muñoz et al. (2024)	DRAGON	Intermediate	DISNET	GraphSAGE	0.91–0.95	NR	NR	
Drug Repurposing Hsieh et al. (2020)	COVID-19 KG	Intermediate	Literature, CTD	VGAE, GraphSAGE	0.78–0.90	NR	NR	
Drug Repurposing Hsieh et al. (2021)	SARS-CoV2 KG	Intermediate	Literature, CTD	VGAE, GraphSAGE	0.77–0.90	NR	NR	
Drug Repurposing Abdeddaiem et al. (2025)	AGMR	Early	MIMIC-III	GNN, GAN	NR	0.85–0.88	0.86–0.88	
Drug Repurposing Lin et al. (2023)	AD Drug Repurposing	Early	STRING, CTD, GO	GraphSAGE	0.84–0.99	NR	NR	
Drug–Drug Interaction Gan et al. (2023)	DMFDDI	Intermediate	Zhang, ChCh-Mine, DeepDDI	Temporal HGNN	0.95–0.99	0.93–0.97	NR	
Drug–Drug Interaction Yu et al. (2023)	ACDGNN	Intermediate	Gene/disease/pathway KG	HGNN with attention	0.71–0.99	0.67–0.94	0.67–0.97	
Drug–Drug Interaction Al-Rabeah and Lakizadeh (2022)	GNN-DDI	Intermediate	DrugBank, KEGG	HGNN	0.99–1.00	0.41–0.86	0.67–0.92	
Drug–Drug Interaction OMKUMARACHANDRAUMAKANTHAM and PATHAK (2024)	DeepSide	Intermediate	TwoSides; DrugBank	GraphSAGE with attention	NR	0.83–0.99	0.77–0.99	
Drug–Drug Interaction Wang et al. (2024a)	MMDDI-MGPFF	Intermediate	DrugBank	GINConv with attention	NR	0.96	0.88	
Drug–Drug Interaction Xiong et al. (2023)	MRCGNN	Intermediate	Deng, Ryu datasets	TrimNet + GNN	NR	0.78–0.89	0.89–0.90	
Drug–Target Prediction Zhang et al. (2023b)	DrugAI	Intermediate	DrugBank	AttentiveFP, DeepWalk, SDNE	LINE, node2vec,	0.88–0.97	0.87–0.89	0.85–0.93
Drug–Target Prediction Xiang et al. (2025)	ExplainMIX	Intermediate	CCLE, PubChem	GDSC, RGCN	0.00–1.0	0.73–0.97	NR	
Drug–Target Prediction Yella et al. (2022)	GraMDTA	Intermediate	DrugBank, RepoDB, DisGeNET	CNN, GraphSAGE with attention	0.88–0.92	0.69–0.80	NR	
Pill Classification Nguyen et al. (2023)	PGPNet	Intermediate	User-captured pill images	ConvNet, RPN, GTN	NR	NR	0.70–0.90	
Prescription Prediction Liu et al. (2020)	RGNN	Intermediate	MIMIC-III	T-LSTM, Temporal GNN	0.82–0.84	NR	NR	
Protein Localization Wang et al. (2023a)	Alteration PLA-GNN	Early	GEO	GraphSAGE	NR	NR	0.410–0.41	

Note. NR = Not reported by the original study.

Table 3. Graph-based models across oncology-related tasks.

Task	Model	Fusion	Dataset	Layers	AUC	F1	Accuracy
Breast Cancer (Gao et al. (2022))	MGNN	Late	2,500 breast cancer patients with gene expression, CNA, and clinical data	Temporal bipartite graphs, CCA	0.98	NR	0.95
Breast Cancer (Kulandaivelu et al. (2024))	ABCD-CAHGNN-MI	Intermediate	DDSM and CBIS-DDSM	DNPGF, GLCM, SCL, CAL, GNN with attention	QOLCT, 0.94–0.97	0.95–0.97	0.97–0.98
Breast Cancer (Kim et al. (2023))	HetMed	Intermediate	Duke-Breast and CMMMD	ResNet, CNN, GCN with attention	NR	0.70–0.86	NR
Glioma (Alzoubi et al. (2024))	PathoFusion	Intermediate	WSI pathology images	CNN, GCN with attention	NR	NR	0.83–0.85
Glioma (Pratap Joshi et al. (2025))	VSA-GCNN	Intermediate	BraTS 2019, 2020, 2021	AlexNet, VSA, GCN with attention	NR	0.97–0.98	0.92–1.00
Lung cancer (Li et al. (2023b))	Lung adenocarcinoma multiclassification model	Early	Zhongshan Hospital; Shanghai Public Health Clinical Center	CNN, GIN, GNN with attention	0.92–0.95	NR	0.87–0.95
Head & Neck Cancer (Peng et al. (2024))	MLF-GNN	Intermediate	TCIA	GraphSAGE, GNN with attention	NR	NR	0.85–0.94
Liver Cancer (Moharana et al. (2025))	FML-LDP	Early	Clinical, demographic, genetic, and imaging data	CNN, GNN with attention, federated meta-learning	NR	0.85–0.93	0.92–0.97
Ovarian Cancer (Ghantasala et al. (2024))	Temporal Analysis GNN	Intermediate +	OCD and SEER	NCI RNN, GNN with attention	0.60–0.82	0.56–0.78	0.56–0.79
Skin Cancer (Yan et al. (2024))	MSF-CNN	Intermediate	ISIC dataset	CNN, GNN with attention	0.66–0.76	0.55–0.63	0.77–0.82
Multi-Cancer Detection (Gowri et al. (2024))	Vision Transformers + GNNs + LayoutLM	Intermediate	IQ-OTH/NCCD Lung Cancer Dataset; PLCO Lung Dataset	ViTs, GNN, LayoutLM	NR	NR	NR
Oncology Misinformation Detection (Cui et al. (2020))	DETERRENT	Early	KnowLife, Healthline, ScienceDaily, NIH, MNT, Mayo Clinic, Cleveland Clinic, WebMD	BiGRU, RGCN with attention	0.54–0.83	0.28–0.67	0.44–0.70

Note. NR = Not reported by the original study.

Table 4. Graph-based models across neuro/psychiatric tasks.

Task	Model	Fusion	Dataset	Layers	AUC	F1	Accuracy	
Alzheimer's (Cai et al. (2023))	AD-GNN	Intermediate	Augmented Pitt Cookie-Theft dataset	BERT, GraphSAGE, BiLSTM, GGNN	NR	NR	0.77–0.85	
Alzheimer's (Wang et al. (2024d))	Knowledge-infused MM-GNN	Intermediate	OASIS; ADNI-D	LLMs, GNN	0.46–0.67	0.46–0.68	0.55–0.82	
Alzheimer's (Tang et al. (2023a))	CsAGP	Intermediate	ADNI (ADNI1/GO and ADNI2)	CNN, Transformers, with attention	Vision GNN	0.99–1.00	NR	0.94–0.99
Alzheimer's (Luo et al. (2024))	DHFWLSL	Intermediate	ADNI (ADNI1/GO and ADNI2)	Dual HGNN with Laplacian regularization	NR	0.42–0.93	0.51–0.94	
Alzheimer's (Kim et al. (2023))	HetMed	Intermediate	ADNI	HGNN, CNN (ResNet), GNN with attention	NR	0.70–0.86	NR	
Alzheimer's (Kumar et al. (2023a))	HCNN-MAFN	Intermediate	ADNI	HGNN with attention	0.98–0.99	0.94–0.96	0.94–0.96	
Alzheimer's (Lee et al. (2024b))	GCNCS	Intermediate	ADNI and DAUH	CNN, GNN	0.92–0.97	0.90–0.99	0.88–0.94	
Alzheimer's (Tripathy et al. (2025))	GNNRAI	Intermediate	ROSMAP, MSBB, Mayo	GNN with attention	0.95–1.00	0.95–1.00	0.76–1.00	
Parkinson's (Lian et al. (2023))	AdaMedGraph	Early	PPMI and PDBP	GNN	0.65–0.76	NR	NR	
Parkinson's (Chan et al. (2022))	JOIN-GCLA	Intermediate	PPMI	GNN with attention	NR	NR	0.90–1.00	
Neurodegenerative (Vijay Anand et al. (2024))	IMNMAGN	Intermediate	BioGPS BrainLat	and ICA, Analysis, Beamforming, CNN, GNN with attention	0.95–0.97	NR	0.91–0.97	
Autism (ASD) (D'Souza et al. (2021))	M-GCN	Early	HCP and KKI	GCN	NR	NR	NR	
Autism (ASD) (Wang et al. (2024b))	IFC-GNN	Intermediate	ABIDE I	Temporal GNN	NR	NR	0.64–0.81	
Autism (ASD) (Cai et al. (2025))	MM-GTUNets	Intermediate	ABIDE I ADHD000	and VAE Learning,	0.88–0.91	NR	0.82–0.83	
Major Depressive Disorder (Li et al. (2025))	AVS-GNN	Intermediate	DAIC-WOZ DVlog	and LSTM, GNN, MLP	NR	0.74–0.88	0.75–0.86	
Major Depressive Disorder (Xing et al. (2024))	EMO-GCN	Intermediate	MODMA	GraphSAGE, GNN with attention	NR	0.89–0.96	0.90–0.97	
Major Depressive Disorder (Liu et al. (2024))	LGMF-GNN	Intermediate	SRPBS and REST- meta-MDD	BiGRU, Snowball GNN	0.73–0.81	0.65–0.91	0.70–0.79	
Major Depressive Disorder (Venkatapathy et al. (2023))	Ensemble GNN	Intermediate	REST-meta-MDD	GNN with attention and GraphSAGE	0.71–0.77	NR	0.70–0.72	
Major Depressive Disorder (Gu et al. (2025))	FC-HGNN	Intermediate	ABIDE and REST- meta-MDD	GNN with attention	0.95–1.00	0.93–1.00	0.92–1.00	
Major Depressive Disorder (Lee et al. (2024a))	Spectral GNN	Early/Late	REST-meta-MDD	Spectral GNNs	0.66–0.74	NR	0.67–0.73	
Major Depressive Disorder (Li et al. (2024))	DSE-HGAT	Intermediate	DAIC-WOZ	BiLSTM, GNN with attention	NR	0.79	NR	
Schizophrenia (Jiang et al. (2023))	Multimodal GNN for EEG	Early/ Intermediate	Chengdu, Hangzhou, Moscow datasets	GNN	0.70–0.85	NR	0.70–0.88	
Schizophrenia (Gao et al. (2025a))	GNN and Multimodal DTI	Intermediate	7 sites across China	GNN with attention	NR	0.74–0.76	0.71–0.74	

Note. NR = Not reported by the original study.

Table 5. Graph-based models for epidemic forecasting and outcomes.

Task	Model	Fusion	Dataset	Layers	AUC	F1	Accuracy
Epidemic forecasting (Qiu et al. (2024))	MSGNN	Intermediate	JHU CSSE	Temporal CNN, GNN with attention	NR	NR	NR
Epidemic forecasting (Xie et al. (2022))	EpiGNN	Intermediate	COVID Prefectures, ILINet, ILI, and GNN with attention JHU-CSSE, Spain-COVID	CNN, AutoregressiveDNN, and GNN with attention	NR	NR	NR
Epidemic forecasting (Tran et al. (2024))	MGLEP	Intermediate	JHU CSSE, OxCGRT, COVID-19 Twitter chatter	BertTweet, RNN, GNN with attention	NR	NR	NR
Epidemic forecasting (Luo et al. (2025))	Dual-Topo-STGCN	Intermediate	CDC surveillance	ILI RNN, GNN	NR	NR	NR
HAI transmission (Gouareb et al. (2023))	MDRE-TransGraph	Early	MIMIC-III	GNN with attention, GraphSAGE	0.89–0.96	NR	0.84–0.97
COVID-19 outcomes (Keicher et al. (2023))	Multimodal GAT	Intermediate	iCTCF and KRI	U-Net and KNN-based, GNN with attention	0.57–0.77	0.18–0.78	0.73–0.74
COVID-19 outcomes (Tariq et al. (2023))	GCNN for clinical event prediction	Intermediate	COVID-19 Emory University Hospital (EUH)	CNN, GNN, LSTM	0.50–0.91	NR	NR
COVID-19 outcomes (Tariq et al. (2025))	Adaptable GCNN for clinical event prediction	Intermediate	COVID-19 Emory University Hospital (EUH)	DenseNet-121, GraphSAGE, LSTM	0.58–0.92	NR	NR

Note. NR = Not reported by the original study.

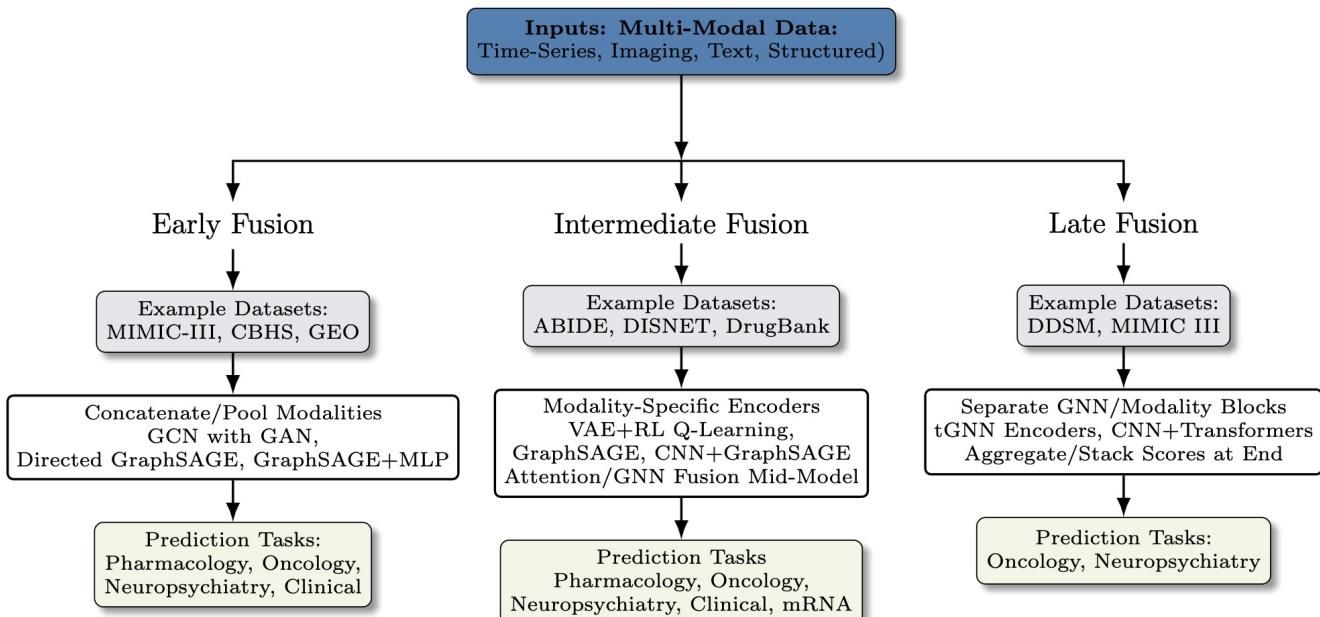


Figure 1. Conceptual workflow of multimodal fusion strategies. Early, intermediate, and late fusion integrate heterogeneous inputs for downstream prediction tasks. In early fusion, modalities are concatenated or pooled up front and passed to a unified encoder. In intermediate fusion, each modality is first processed by a modality-specific encoder, and features are combined mid-model via attention/GNN layers. In late fusion, separate modality/GNN branches are trained, and their scores are combined only at the decision stage. Prediction layers are dominated by fully connected layers, multiple-layer perceptrons, or machine learning classifiers.

Table 6. Graph-based models for clinical prediction, pathways, and hospital operations.

Task	Model	Fusion	Dataset	Layers	AUC	F1	Accuracy
Tuberculosis (D'Souza et al. (2023))	MaxCorr-MGNN	Intermediate	Tuberculosis Data Exploration Portal	Hirschfeld–Gebelein–Rényi maximal correlation and GNN	0.77–0.78	NR	NR
Care pathway prediction (Liu et al. (2021))	Multitask Healthcare Management System	Intermediate	600,000 multimodal samples (structured, images)	CNN (ResNet), GNN, NR Word2Vec, RNN	NR	NR	NR
Clinical risk prediction (AL-Sabri et al. (2024))	M3GNAS	Intermediate	MIMIC-III	BiGRU, BioBERT, GNN with attention	0.70–0.91	NR	NR
Hospital readmission (Tang et al. (2023b))	MM-STGNN	Intermediate	MIMIC-IV; 9,958 admissions / 44,084 radiographs / 9,162 patients	GraphSAGE, RNN, GNN	0.58–0.91	NR	NR
Federated diagnosis (Begum (2024))	FH-MMA	Intermediate	MIMIC-III	CNN, Transformers, and GNN with attention	NR	NR	0.93–0.95
Multitask longitudinal modeling (Boschi et al. (2024))	funGCN	Intermediate	SHARE and synthetic dataset	GNN	NR	NR	0.58–0.93
Clinical triage (Valls et al. (2023))	Masked-Connectivity Triage GNN	Intermediate	Synthea	GNN, KG, Temporal GNN	NR	NR	0.40–0.85
Comorbidity prediction (Pablo et al. (2024))	Multitask Comorbidity GCN	Intermediate	Imaging genomics clinical notes	+ CNN, BERT, GNN +	0.96	0.93	0.95
Sleep apnea diagnosis (Wang et al. (2025))	HeteroGCFNet	Intermediate	OSAHS	BiLSTM, GNN with attention	NR	0.80–0.84	0.84–0.88
Sepsis trajectory modeling (Ghanvatkar and Rajan (2023))	Dynamic Clinician-in-the-Loop GNN	Intermediate	MIMIC-IV	GNN, Temporal HGNN with attention	0.74	0.36	0.87
Ophthalmology auxiliary diagnosis (Gao et al. (2024))	CGAT-ADM	Intermediate	Ophthalmic EMRs (Beijing Tongren Hospital)	BERT, metapath2vec, GNN with attention	NR	NR	NR
Diabetic retinopathy (Zedadra et al. (2025))	DRdiag	Intermediate	APTOPS 2019; MESSIDOR0	CNN, GNN	NR	0.96	0.96–0.98
Heart disease (Boll et al. (2024))	Patient-KNN Graph	Early	MIMIC-III	GraphSAGE, KNN, Graph Transformers, GNN with attention	0.75–0.79	0.47–0.53	0.70–0.80
ICU albumin prediction (Zhang et al. (2024))	DyG-HAP	Intermediate	ANIC	Disentangled dynamic graph with attention	NR	NR	NR
ICU length of stay (Christos Maroudis et al. (2025))	Fairness-Aware Dynamic ST-GNN	Intermediate	MIMIC-IV	LSTM, GNN with attention	0.82–0.91	NR	NR

Note. NR = Not reported by the original study.

Table 7. Graph-based models for ncRNA–miRNA interaction prediction.

Task	Model	Fusion	Dataset	Layers	AUC	F1	Accuracy
lncRNA–miRNA interaction (Wang et al. (2023b))	SPGNN	Intermediate	LncACTdb; LNCipedia; miRBase	3.0; k-mer GraphSAGE, with attention	Doc2Vec, GNN	0.84	0.75–0.76 NR
lncRNA–miRNA interaction (Wang et al. (2022))	GCNCRF	Intermediate	lncRNAsNP2; LncACTdb; LNCipedia; miRBase	3.0; Fields, attention	Conditional Random GNN with	0.88–0.95	0.13–0.14 0.97–0.98
lncRNA–miRNA interaction (Wang and Chen (2023))	MAGCN	Intermediate	ncRNAsNP v2.0; HMDD v3.0	CNN, attention	GNN with	0.90	0.50–0.51 0.94
ncRNA–miRNA interaction (Zhang et al. (2022))	ncRNAInter	Intermediate	lncRNAsNP; miRBase v22.1; GENCODE v38	GraphSAGE with neighbor sampling	0.97–0.99	0.93–0.96	0.93–0.96

Note. NR = Not reported by the original study.

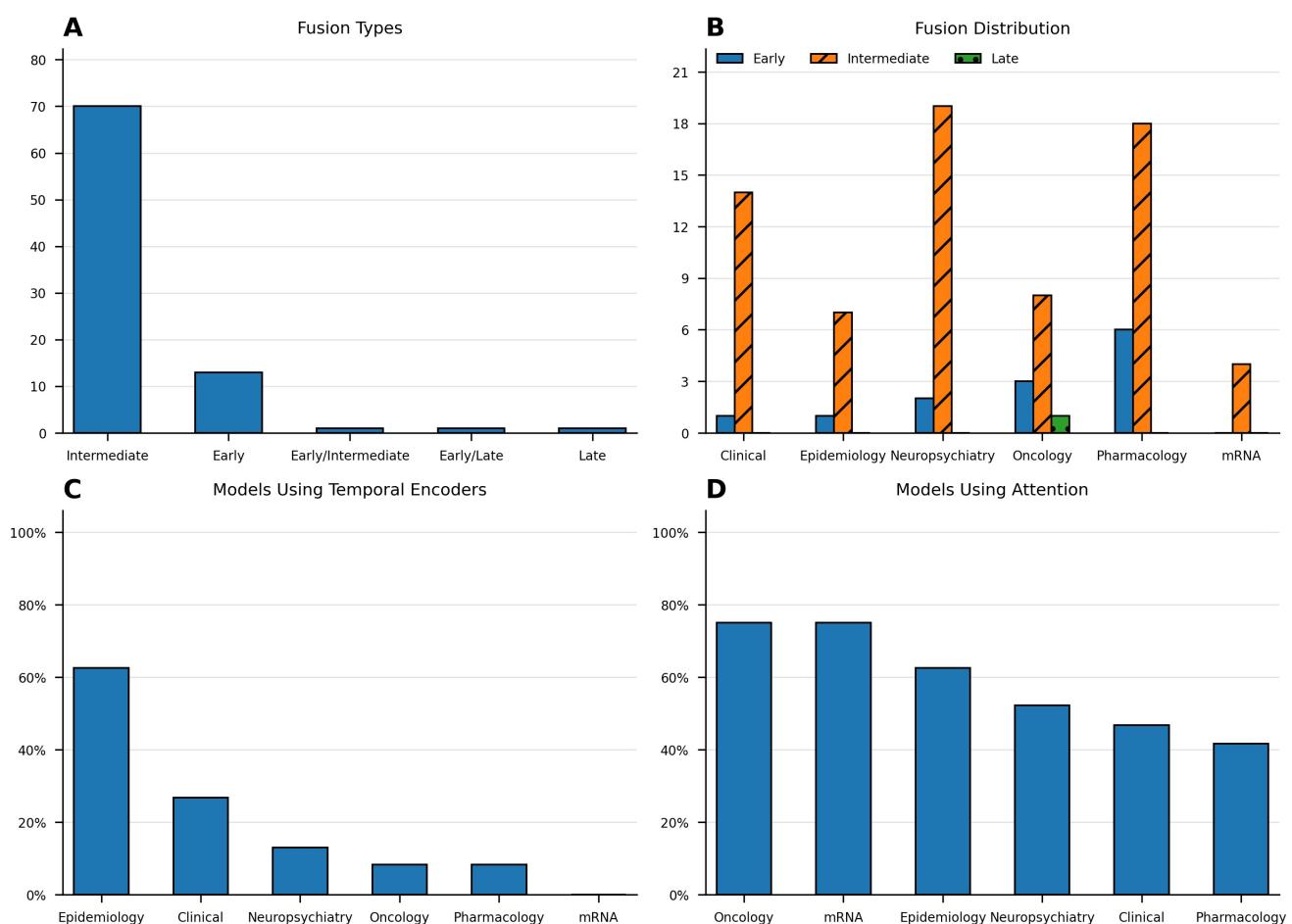


Figure 2. Multimodal fusion strategies and encoder usage across research areas. (A) Overall distribution of fusion strategies across all models. (B) Fusion distribution by area. (C) Share of models that include a temporal encoder by area. Share of models that include an attention mechanism by area.

Table 8. Summary comparison of top-performing multimodal GNN models across biomedical domains, selected based on highest AUC, accuracy, and F1 scores, highlighting architectures, datasets, fusion types, and performance outcomes to identify effective strategies.

Domain	Task	Model	Fusion	Dataset	Architecture (Layers)	Performance Outcomes
Pharmacology	Drug–Drug Interaction	ACDGNN (Yu et al., 2023)	Intermediate	Gene, disease, pathway KG	HGNN with attention	AUC: 0.99–1.00; F1: 0.41–0.86; Acc: 0.67–0.92
Pharmacology	Drug Repurposing	AD Drug Repurposing (Lin et al.)	Early	STRING, GO, CTD	GraphSAGE	AUC: 0.95–0.99; F1: 0.93–0.97; Acc: NR
Oncology	Glioma	VSA-GCNN (Pratap Joshi et al., 2025)	Intermediate	BraTS 2019/2020/2021	AlexNet, VSA, GCN with attention	AUC: NR; F1: 0.97–0.98; Acc: 0.92–1.00
Oncology	Breast Cancer	MGNN (Gao et al., 2022)	Late	2,500 patients (gene expression, CNA, clinical data)	Temporal GNN on bipartite graphs; CCA fusion	AUC: 0.98; F1: NR; Acc: 0.95
Neuropsychiatry	Major Depressive Disorder	FC-HGNN (Gu et al., 2025)	Intermediate	ABIDE; REST-meta-MDD	GNN with attention	AUC: 0.95–1.00; F1: 0.93–1.00; Acc: 0.92–1.00
Neuropsychiatry	Alzheimer's Disease	CsAGP (Tang et al., 2023)	Intermediate	ADNI1/GO; ADNI2	CNN; Vision Transformers; GNN with attention	AUC: 0.99–1.00; F1: NR; Acc: 0.94–0.99
Epidemiology	HAI Transmission	MDRE-TransGraph (Gouareb et al., 2023)	Early	MIMIC-III	GNN with attention; GraphSAGE	AUC: 0.89–0.96; F1: NR; Acc: 0.84–0.97
Epidemiology	COVID-19 Outcomes	Adaptable GCNN (Tariq et al., 2025)	Intermediate	EUH COVID-19 Cohort	DenseNet-121; GraphSAGE; LSTM	AUC: 0.58–0.92; F1: NR; Acc: NR
Clinical	Comorbidity Prediction	Multitask Comorbidity GCN	Intermediate	Imaging genomics + clinical notes	+ CNN; BERT; GNN	AUC: 0.96; F1: 0.93; Acc: 0.95
Clinical	Diabetic Retinopathy	DRdiag (Zedadra et al., 2025)	Intermediate	APOTOS 2019; MESSIDOR	CNN; GNN	AUC: NR; F1: 0.96; Acc: 0.96–0.98
Genomics	ncRNA–miRNA Interaction	ncRNAInter (Zhang et al., 2022)	Intermediate	lncRNAsNP2; miRBase v22.1; GENCODE v38	GraphSAGE with neighbor sampling	AUC: 0.97–0.99; F1: 0.93–0.96; Acc: 0.93–0.96
Genomics	lncRNA–miRNA Interaction	GCNCRF (Wang et al., 2022)	Intermediate	lncRNAsNP2; LncACTdb 3.0; LNCipedia; miRBase	CRF + GNN with attention	AUC: 0.88–0.95; F1: 0.13–0.14; Acc: 0.97–0.98

Note. NR = Not reported by the original study.

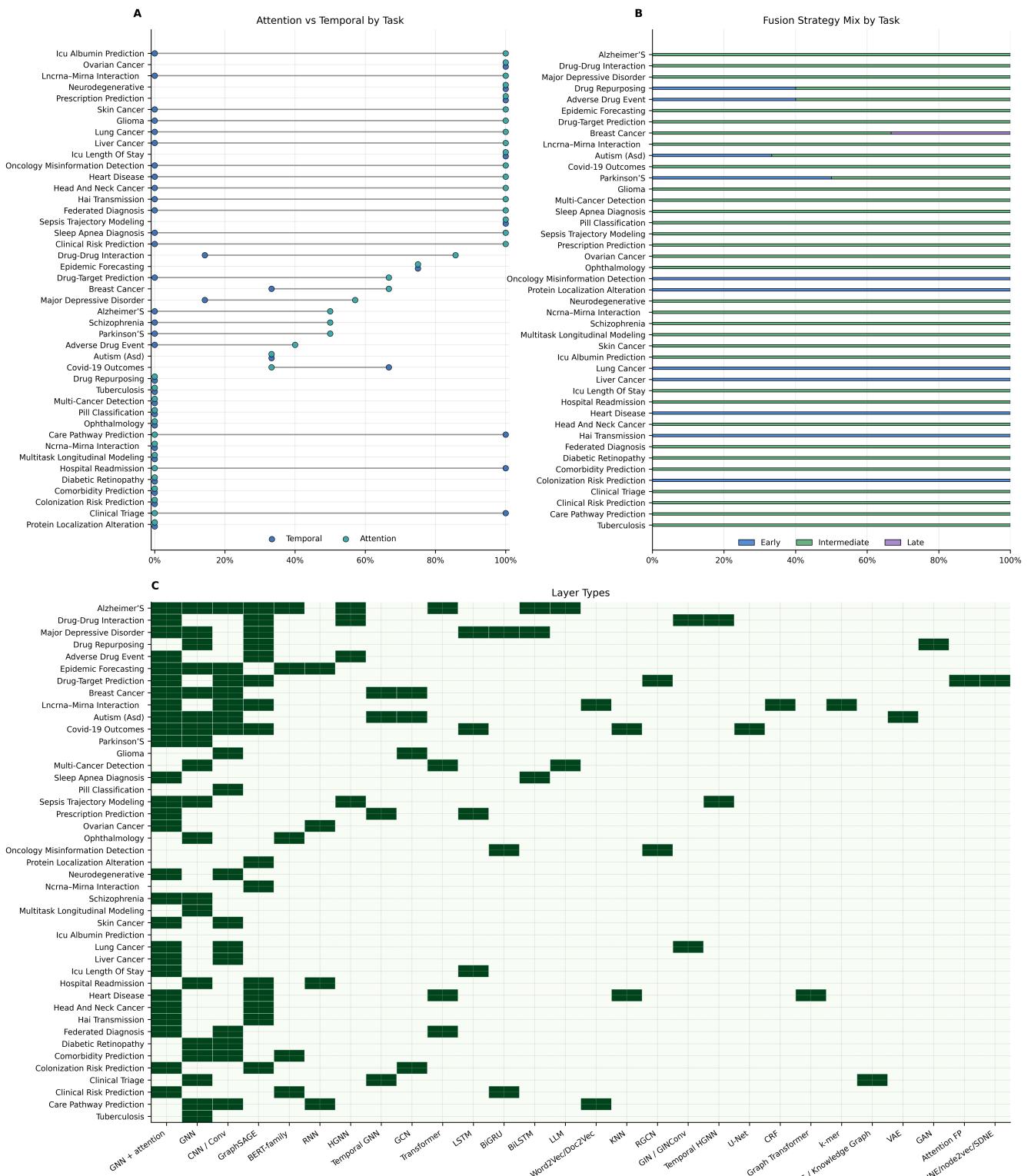


Figure 3. Architectural patterns across tasks. (A) Gap chart comparing the share of models using attention versus temporal encoders for the top tasks. (B) Normalized (100%) stacked bars showing the fusion strategy mix. Values are the proportion of models per task that use each fusion scheme. (C) Heatmap of layer types extracted from model descriptions.