



PREDICTION OF ALZHEIMER'S MICRO RNA – DISEASE CAUSAL ASSOCIATION USING GRAPH NEURAL NETWORKS

Soundharya Manogaran,¹ Pradeep Kumar Yadalam,² Ramya Ramados,¹
Carlos M. Ardila^{2,3}

¹ Department of Oral Pathology & Oral Biology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India.

²Department of Periodontics, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, SIMATS, Saveetha University, Chennai, Tamil Nadu, India.

³Department of Basic Sciences, Biomedical Stomatology Research Group, Faculty of Dentistry, Universidad de Antioquia U de A, Medellín, Colombia.

Received: 02/28/2025
Accepted: 03/06/2025

CORRESPONDENCE: Carlos M. Ardila, Calle 70 No. 52-21, Medellín, Colombia and Pradeep Kumar Yadalam, Saveetha University, Chennai, Tamil Nadu, India;

EMAIL: martin.ardila@udea.edu.co / pradeepkumar.sdc@saveetha.com



ABSTRACT

Introduction: Alzheimer's disease (AD) is a neurodegenerative disorder characterized by the presence of tau protein tangles and beta-amyloid plaques, which disrupt brain function and lead to neuronal loss. **Objective:** To elucidate the molecular mechanisms underlying AD and identify novel therapeutic targets, focusing on establishing causal links between microRNAs and AD, critical for early detection and disease monitoring. **Methods:** The study utilized the HMDD V4.0 database to examine a dataset concerning causality in microRNA-disease associations, comprising microRNA expression profiles and clinical information from AD patients and healthy individuals. The dataset was subjected to graph neural networks to predict causal interactions between microRNAs and AD. **Results:** The graph neural network model demonstrated robust performance in predicting Alzheimer's disease, achieving a precision of 69.44%, recall of 83.33%, and F1 score of 75.75%, indicating its potential in identifying biomarkers or therapeutic targets using microRNA data. **Conclusion:** The graph neural network model for predicting causal associations between Alzheimer's disease and microRNAs shows promise. However, further investigation is warranted regarding dataset balance, adjustments to classification thresholds, incorporation of additional features, and exploration of advanced model architectures.

KEYWORDS: Computational Biology; Alzheimer's Disease; MicroRNA; Bioinformatics



PREDICCIÓN DE LA ASOCIACIÓN CAUSAL MICRO ARN-ENFERMEDAD DE ALZHEIMER UTILIZANDO REDES NEURONALES DE GRAFOS

RESUMEN

Introducción: La enfermedad de Alzheimer (EA) es un trastorno neurodegenerativo caracterizado por la presencia de ovillos de proteína tau y placas beta-amiloideas, que interrumpen la función cerebral y conducen a la pérdida neuronal. **Objetivo:** Dilucidar los mecanismos moleculares subyacentes a la EA e identificar nuevas dianas terapéuticas, con un enfoque en el establecimiento de vínculos causales entre los microARNs y la EA, críticos para la detección temprana y el monitoreo de la enfermedad. **Métodos:** El estudio utilizó la base de datos HMDD V4.0 para examinar un conjunto de datos sobre la causalidad en las asociaciones microARN-enfermedad, que comprende perfiles de expresión de microARN e información clínica de pacientes con EA e individuos sanos. El conjunto de datos se sometió a redes neuronales de grafos para predecir las interacciones causales entre los microARNs y la EA. **Resultados:** El modelo de red neuronal de grafos demostró un rendimiento robusto en la predicción de la enfermedad de Alzheimer, alcanzando una precisión del 69.44%, una recuperación del 83.33% y una puntuación F1 del 75.75%, lo que indica su potencial en la identificación de biomarcadores o dianas terapéuticas utilizando datos de microARN. **Conclusión:** El modelo de red neuronal de grafos para la predicción de asociaciones causales entre la enfermedad de Alzheimer y los



microARNs muestra ser prometedor. Sin embargo, se justifica una mayor investigación sobre el equilibrio del conjunto de datos, los ajustes a los umbrales de clasificación, la incorporación de características adicionales y la exploración de arquitecturas de modelos avanzados.

PALABRAS CLAVE: Biología computacional; Enfermedad de Alzheimer; microARN; Bioinformática.

INTRODUCTION

Alzheimer's disease (AD) is a complex neurodegenerative disorder that diminishes identity and autonomy through its impact on behavior, memory, and cognitive abilities. Discovered over a century ago, AD is characterized by the presence of tau protein tangles and beta-amyloid plaques, which disrupt brain function and lead to neuronal death (1). Despite its pervasive occurrence and devastating consequences, effective

treatments remain elusive. As the global population ages, AD is poised to exert a significant socioeconomic burden on caregivers and healthcare systems. Advancements in genetics, neuroscience, and technology are driving ongoing progress in AD research, offering hope for mitigating its profound impact through improved early detection, enhanced diagnostic precision, and personalized interventions (1, 2).



MicroRNAs (miRNAs), small RNA molecules, modulate gene expression by binding to target messenger RNA (mRNA) molecules, either facilitating or inhibiting their translation (3, 4). Dysregulation of miRNAs significantly influences the pathogenesis of Alzheimer's disease (AD). Research suggests that specific miRNAs may regulate genes and processes associated with AD pathophysiology, such as the formation and clearance of tau protein tangles, amyloid-beta ($A\beta$) plaque deposition, neuroinflammation, synaptic dysfunction, and neuronal apoptosis. Dysregulated miRNAs have been identified in various tissues and biofluids of AD patients, indicating their potential utility as biomarkers for prognosis, diagnosis, and treatment response

assessment. Understanding the causal relationships between miRNAs and AD pathophysiology holds promise for early disease detection, personalized therapeutic interventions, and cognitive decline prevention (5, 6).

Graph neural networks (GNNs) represent a powerful machine-learning paradigm capable of capturing complex relationships inherent in graph-structured data (7, 8). By amalgamating principles from graph theory and neural networks, GNNs can directly process graph-structured data, exploiting its inherent structure and connectivity. In Alzheimer's disease research, graphical neural networks (GNNs) are utilized to predict causal relationships involving microRNAs (miRNAs). This approach



shows potential for unraveling the underlying pathophysiology of AD and identifying viable therapeutic targets (9, 10). GNNs can integrate multi-omics data into a unified graph representation, facilitating the learning of intricate dependencies and patterns. This enables the prediction of associations between miRNAs and diseases, elucidating the causal mechanisms underlying AD pathophysiology. The GNN-based methodology encompasses data integration, graph construction, feature encoding, graph convolution, prediction, and interpretation, offering insights into the molecular mechanisms driving AD and aiding in the development of targeted therapeutic interventions. Moreover, interpretable models can provide valuable insights into the biological mechanisms

and regulatory networks underpinning AD pathophysiology, paving the way for personalized precision medicine approaches.

Aberrant miRNA expression is implicated in various diseases, and deep learning methods like ADPMDA demonstrate efficient prediction of miRNA-disease associations, achieving a mean AUC value of 94.75% in experiments (10-12). GraphTar, a novel miRNA target prediction method, utilizes a graph-based representation and word2vec encoding for RNA sequences, surpassing existing approaches, albeit necessitating expanded datasets for further exploration (13).

This study aims to elucidate the molecular mechanisms of Alzheimer's disease (AD) and identify novel



therapeutic targets, with a specific focus on delineating causal associations between microRNAs (miRNAs) and AD, essential for early disease detection and monitoring. Identification of specific miRNAs holds promise for understanding disease pathogenesis and facilitating targeted interventions. Graph neural networks (GNNs) can predict causal associations between miRNAs and diseases, thereby enhancing the ability to identify miRNAs causally linked to AD. MiRNA-based therapeutics, an emerging field in AD research, have the potential to restore normal gene expression patterns and attenuate or reverse disease progression.

Our study aims to predict miRNA-disease causal associations using graph neural

network-based approaches in Alzheimer's disease.

Materials and Methods

Dataset Preparation

The study utilized HMDD V4.0 (14) to analyze the microRNA-disease association causality dataset, comprising microRNA, disease, and causality information. Nodes representing microRNAs, diseases, and causality were designated, and graph neural networks were employed to predict edge interactions as causality between microRNAs and diseases, specifically Alzheimer's disease. This methodology facilitated effective analysis and the discovery of potential causal connections within the biomedical context.



A dataset sourced from HMDD V4.0 (14) was employed, incorporating microRNA expression profiles and associated clinical data from individuals with Alzheimer's disease and healthy controls. Preprocessing steps were implemented to clean and normalize the microRNA expression data while integrating it with clinical information. The Alzheimer's disease dataset, inclusive of microRNA, disease, causality (categorized as "no" and "yes"), and primary literature references (PMIDs), was extracted from the downloaded dataset. Nodes were allocated for microRNA, disease, and causality, serving as edges, and node features were assigned and labeled. These data were then subjected to graph neural networks (GNNs) with the aim of

predicting the edge interaction of causality.

This study employs GNNs to investigate the intricate relationships between microRNAs and Alzheimer's disease, with the goal of predicting their causality using structured data. This comprehension is pivotal for understanding molecular mechanisms and identifying potential therapeutic targets.

Graph Neural Network Architecture

GNNs are deep learning models designed to utilize graph-structured data for tasks such as link prediction and node classification. They comprise multiple layers processing node and edge features, facilitating information propagation across the graph. Techniques like



message passing enable nodes to capture higher-order dependencies and refine hidden states. GNNs commonly utilize the Adam optimizer for optimization. The architecture of a GNN encompasses multiple layers of graph convolutional operations, facilitating information propagation between nodes. The output of the final layer is inputted into a classification layer to predict the causality of microRNAs in Alzheimer's disease.

In addition to the number of layers and units, other hyperparameters that can be tuned include the learning rate, dropout rate, and regularization strength. The learning rate dictates the step size at which the model updates its parameters during training. A high learning rate may cause the model to overshoot optimal values, while a low learning rate may

decelerate the learning process. The dropout rate serves as a regularization technique, randomly setting a fraction of input units to 0 during training to prevent overfitting. The regularization strength governs the balance between the model's capability to fit the training data and its generalization to unseen data.

The hyperparameter selection process typically involves grid search and model evaluation. The model exhibiting the best performance on the validation set is chosen for further evaluation on the test set. By fine-tuning the hyperparameters, the performance of the GNN model can be optimized, leading to enhanced predictions of microRNA causality in Alzheimer's disease.



Results

The graph neural network model accurately predicted Alzheimer's disease, achieving a precision of 69.44%, recall of 83.33%, and an F1 score of 75.75%. This methodology demonstrates promise in the identification of potential biomarkers or therapeutic targets for Alzheimer's disease utilizing microRNA data.

The precision and recall metrics of the model are pivotal for discerning genuine

associations. Precision, quantifying the ratio of true positive predictions to total positive predictions, attains an accuracy of 69.44%. Recall, or sensitivity, achieves an accuracy of 83.33%, signifying successful identification of positive associations. The F1-score, serving as a harmonic mean of precision and recall, offers a balanced assessment of these metrics, indicating commendable performance.

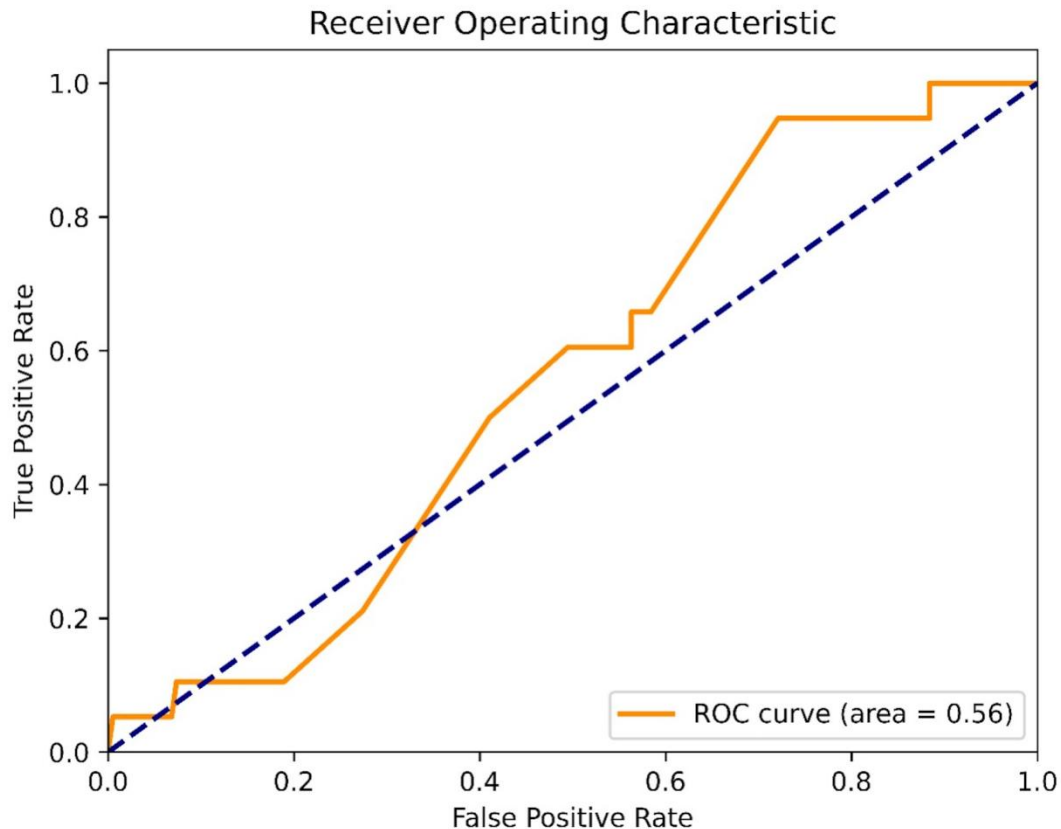


Figure 1 illustrates the ROC curve predicting the causal association between microRNAs and Alzheimer's disease. It charts the True Positive Rate (TPR) against the False Positive Rate (1-specificity) across various threshold settings. The AUC value of 0.56 indicates that the model performs slightly better

than random guessing but lacks robustness in distinguishing between positive and negative cases. The proximity of the ROC curve to the top-left corner signifies superior model performance.

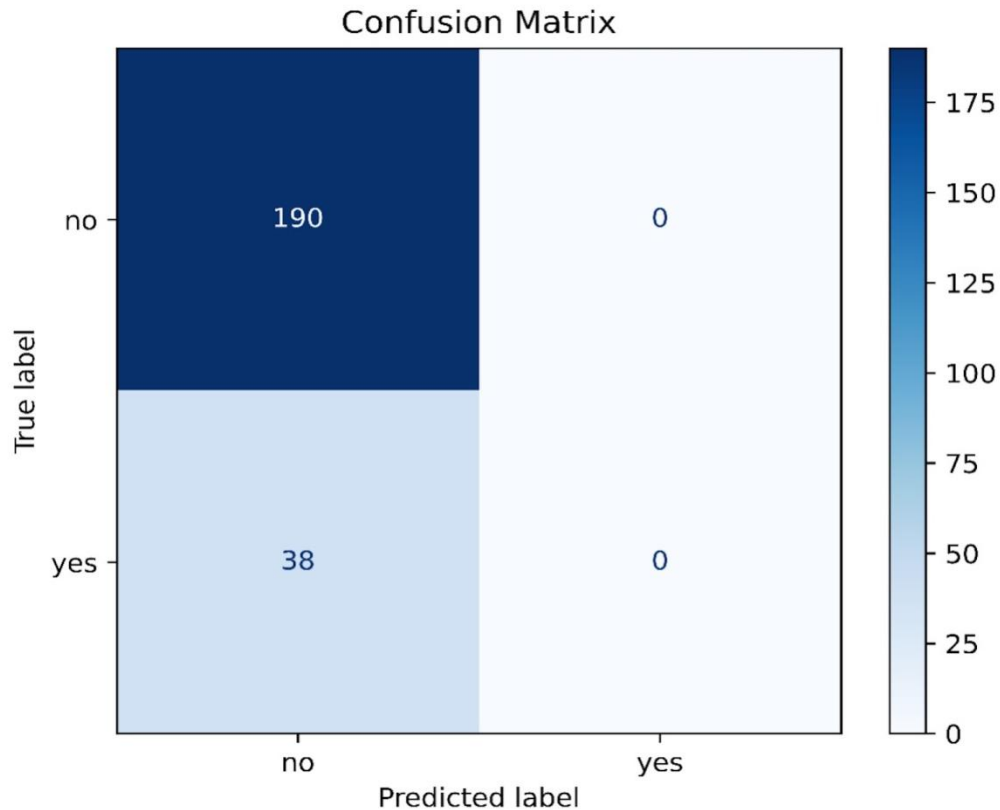


Figure 2 depicts the confusion matrix of the graph neural network model for predicting causal associations with Alzheimer's microRNA disease. It accurately recognized 190 instances of "no" cases but did not identify any "yes" cases, resulting in 38 false negatives and zero true positives. The model effectively

discerned 190 instances lacking a causal association between a microRNA and Alzheimer's disease but did not detect any instances exhibiting a causal association. While proficient in identifying cases without a causal association, the model struggled to accurately identify cases with a causal association, highlighting a

significant limitation in its ability to predict microRNA-disease causal

associations in the context of Alzheimer's disease.

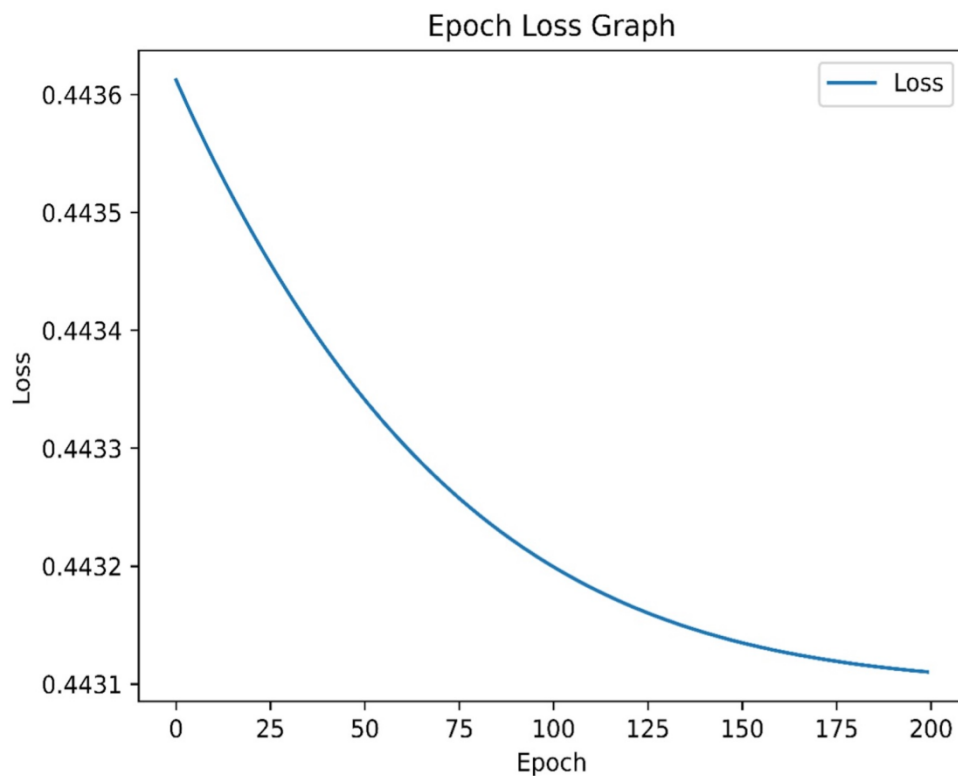


Figure 3 illustrates the Epoch Loss Graph, demonstrating the performance of the neural network model in predicting causal associations between Alzheimer's microRNA disease and training epochs.

The model's loss initiates at 0.4436 and steadily diminishes with increasing epochs, suggesting effective learning from the training data. By epoch 200, the



loss decreases to 0.4431, signifying enhanced accuracy.

Discussion

Alzheimer's disease stands as a neurodegenerative condition characterized by cognitive decline and memory impairment (15,16), believed to be influenced by both genetic and environmental factors. Recent investigations have underscored the role of microRNAs (miRNAs) in the development and progression of Alzheimer's disease. Dysregulated miRNAs have been pinpointed in Alzheimer's disease, targeting genes crucial for neuronal survival, synaptic plasticity, and inflammation (17-19). A recent study employed protein-protein

interaction networks and miRNA-gene interactions to link Alzheimer's disease with miRNAs, identifying 257 novel AD-related miRNAs via a one-class SVM semi-clustering method, yielding higher AUC values and reliable outcomes (2,20).

Previous research utilized a network of miRNA-target interactions (MTIs) relevant to Alzheimer's disease (AD), sourced from data in miRTarBase. The network comprised seven MTI subnetworks and 12 MTI pairs, incorporating nodes such as APP, BACE1, NCSTN, SIRT, SP1, and specific miRNAs like miR-9, miR-16, and miR-181c. This study scrutinized the interactions of miRNA targets with proteins and their enrichment for AD-



associated miRNAs, elucidating their role in gene expression regulation (20,21). Nonetheless, consensus on validation strength remains elusive, and our study endeavors to explore microRNA disease causation prediction to uncover novel targets.

MicroRNAs (miRNAs) play pivotal roles in post-transcriptional gene regulation, with implications in various diseases including cancer. Graph neural networks (GNNs) have been leveraged in drug screening for small molecules targeting miR-21, predicting potential drugs based on structural similarities and showcasing potential for discovering novel therapeutic agents. GATMDA, a computational framework (9,10), employs

a graph attention network to identify miRNA-disease associations, achieving high performance with an average AUC of 0.9566 in five-fold cross-validation.

Recent investigations highlight the potential of graph neural networks (GNNs) in predicting Alzheimer's disease risk. These robust machine learning models can capture intricate biomarker relationships, offering insights into disease progression and facilitating more accurate prognosis and earlier interventions. Incorporating time-varying information into the GNN framework could further enhance prediction accuracy (22-24). The development of explainable GNN models can furnish personalized explanations for early Alzheimer's disease diagnosis, empowering clinicians to make informed decisions and tailor treatment



strategies. The graph neural network model adeptly predicts Alzheimer's microRNA disease causal associations, achieving an accuracy of 83.33%, precision of 69.44%, recall of 83.33%, and an F1-score of 75.76%, rendering it a promising tool for unraveling complex relationships.

The accuracy of a model is influenced by its proficiency in predicting cases with no causal association accurately. However, a more nuanced analysis is provided by the ROC curve and AUC value. The AUC value of 0.56 suggests the model performs slightly better than random guessing but struggles to discriminate between positive and negative cases. This disparity may arise from class imbalance,

false positives, and negative costs. Therefore, the model's subpar performance in the ROC curve implies inadequate identification of cases with causal associations. The graph neural network model's accuracy in predicting Alzheimer's microRNA disease (25-27) causal associations holds promise, yet there exist limitations and potential avenues for enhancement (28,29). Dataset imbalance, where one class outweighs the other in terms of samples, can bias the model towards predicting the majority class, resulting in higher accuracy but poorer performance in identifying cases with causal associations (30,31). The model may also prioritize minimizing false positives at the expense of higher false negatives, leading to a lower AUC value. Future endeavors encompass



dataset balancing, classification threshold adjustments, and exploration of additional features or advanced GNN architectures. Extending the application to other diseases could yield valuable insights into molecular mechanisms, though limitations include data availability, heterogeneity, lack of experimental validation, and interpretability.

Conclusion

The graph neural network model for predicting causal associations between Alzheimer's disease and microRNAs exhibits promise; however, certain limitations need to be addressed. Future research endeavors should prioritize addressing dataset imbalance, fine-tuning classification thresholds, incorporating

additional features, and exploring advanced model architectures. These efforts hold potential for enhancing the accuracy and robustness of the predictive model, thereby advancing our understanding of Alzheimer's disease pathophysiology and identifying novel therapeutic targets.

Conflict of Interest

The authors have no conflicts of interest to declare

Authors contribution

Pradeep Kumar Yadalam, Soundharya Manogaran, Ramya Ramadoss and Carlos M. Ardila contributed to the conception,



analysis, interpretation of data, and drafting of the manuscript.

Pradeep Kumar Yadalam, Soundharya Manogaran, Ramya Ramadoss and Carlos M. Ardila: Conceptualization

Pradeep Kumar Yadalam, and Carlos M. Ardila: Methodology

Pradeep Kumar Yadalam, Soundharya Manogaran, Ramya Ramadoss and Carlos M. Ardila: Data curation

Pradeep Kumar Yadalam, and Carlos M. Ardila: Writing- Original draft preparation.

Pradeep Kumar Yadalam, Soundharya Manogaran, Ramya Ramadoss and Carlos M. Ardila: Visualization

Pradeep Kumar Yadalam, and Carlos M. Ardila: Investigation.

Pradeep Kumar Yadalam, Soundharya Manogaran, Ramya Ramadoss and Carlos M. Ardila: Validation and supervision.

Pradeep Kumar Yadalam, and Carlos M. Ardila: Writing- Reviewing and Editing.

Ethics approval: Not applicable.

REFERENCES

1. Wei H, Zhu Z, Xu Y, Lin L, Chen Q, Liu Y, et al. Microglia-derived exosomes selective sorted by YB-1 alleviate nerve damage and cognitive outcome in Alzheimer's disease. *J Transl Med.* 2024;22(1):466.
2. Li JZ, Ramalingam N, Li S. Targeting epigenetic mechanisms



in amyloid- β -mediated Alzheimer's pathophysiology: unveiling therapeutic potential. *Neural Regen Res.* 2025;20(1):54–66.

3. Jadhav SP. MicroRNAs in microglia: deciphering their role in neurodegenerative diseases. *Front Cell Neurosci.* 2024;18:1391537.

4. Singh M, Agarwal V, Panoram P, Jindal D, Agarwal S, Rai SN, et al. A Comprehensive Review and Androgen Deprivation Therapy and Its Impact on Alzheimer's Disease Risk in Older Men with Prostate Cancer. *Degener Neurol Neuromuscul Dis.* 2024;14:33–46.

5. Piergiorgio RM, Vasconcelos ATR de, Santos-Rebouças CB. Understanding the (epi)genetic dysregulation in Parkinson's disease through an integrative

brain competitive endogenous RNA network. *Mech Ageing Dev.* 2024;219:111942.

6. Turk A, Kunej T, Peterlin B. MicroRNA-Target Interaction Regulatory Network in Alzheimer's Disease. *J Pers Med.* 2021 Dec;11(12): 1275

7. Haga CL, Yang XD, Gheit IS, Phinney DG. Graph neural networks for the identification of novel inhibitors of a small RNA. *SLAS Discov.* 2023;28(8):402–9.

8. Dong B, Sun W, Xu D, Wang G, Zhang T. DAEMDA: A Method with Dual-Channel Attention Encoding for miRNA-Disease Association Prediction. *Biomolecules.* 2023;13(10):1514

9. Yang Y, Sun Y, Li F, Guan B, Liu JX, Shang J. MGCNRF:



Prediction of Disease-Related miRNAs Based on Multiple Graph Convolutional Networks and Random Forest. *IEEE Trans Neural Netw Learn Syst.* 2024;35(11):15701-15709.

10. Jia C, Wang F, Xing B, Li S, Zhao Y, Li Y, et al. DGAMDA: Predicting miRNA-disease association based on dynamic graph attention network. *Int J Numer Method Biomed Eng.* 2024;40(5):e3809.

11. Yan C, Duan G, Li N, Zhang L, Wu FX, Wang J. PDMDA: predicting deep-level miRNA-disease associations with graph neural networks and sequence features. *Bioinformatics.* 2022;38(8):2226–34.

12. Ning Q, Zhao Y, Gao J, Chen C, Li X, Li T, et al. AMHMDA: attention aware multi-view similarity networks and hypergraph learning for miRNA-disease associations identification. *Brief Bioinform.* 2023;24(2):bbad094

13. Przybyszewski J, Malawski M, Licholai S. GraphTar: applying word2vec and graph neural networks to miRNA target prediction. *BMC Bioinformatics.* 2023;24(1):436.

14. Cui C, Zhong B, Fan R, Cui Q. HMDD v4.0: a database for experimentally supported human microRNA-disease associations. *Nucleic Acids Res.* 2024;52(D1):D1327–32.

15. Zhang L, Lu Q, Chang C. Epigenetics in Health and Disease.



Adv Exp Med Biol. 2020;1253:3–55.

16. Le TD, Liu L, Tsykin A, Goodall GJ, Liu B, Sun BY, et al. Inferring microRNA–mRNA causal regulatory relationships from expression data. *Bioinformatics*. 2013;29(6):765–71.

17. Li G, Fang T, Zhang Y, Liang C, Xiao Q, Luo J. Predicting miRNA-disease associations based on graph attention network with multi-source information. *BMC Bioinformatics*. 2022;23(1):244.

18. Budakoti M, Panwar AS, Molpa D, Singh RK, Büsselberg D, Mishra AP, et al. Micro-RNA: The darkhorse of cancer. *Cell Signal*. 2021;83:109995.

19. Yao D, Li B, Zhan X, Zhan X, Yu L. GCNFORMER: graph convolutional network and transformer for predicting lncRNA-disease associations. *BMC Bioinformatics*. 2024;25(1):5

20. Li R, Yao S, Wei F, Chen M, Zhong Y, Zou C, et al. Downregulation of miR-181c-5p in Alzheimer’s disease weakens the response of microglia to A β phagocytosis. *Sci Rep*. 2024 ;14(1):11487.

21. Lahiri DK, Maloney B, Wang R, White FA, Sambamurti K, Greig NH, et al. The seeds of its regulation: natural antisense transcripts as single-gene control switches in neurodegenerative disorders. *Ageing Res Rev*. 2024; 99:102336



22. Fu Y, Yang R, Zhang L, Fu X. HGECD: A Heterogeneous Graph Embedding Model for CircRNA-Disease Association Prediction. *IEEE J Biomed Health Inform.* 2023;27(10):5177–86.

23. Xie GB, Yu JR, Lin ZY, Gu GS, Chen RB, Xu HJ, et al. Prediction of miRNA-disease associations based on strengthened hypergraph convolutional autoencoder. *Comput Biol Chem.* 2024;108:107992.

24. Ai N, Liang Y, Yuan H, Ouyang D, Xie S, Liu X. GDCL-NcDA: identifying non-coding RNA-disease associations via contrastive learning between deep graph learning and deep matrix factorization. *BMC Genomics.* 2023;24(1):424.

25. Hu H, Zhao H, Zhong T, Dong X, Wang L, Han P, et al. Adaptive deep propagation graph neural network for predicting miRNA-disease associations. *Brief Funct Genomics.* 2023;22(5):453–62.

26. Ma Z, Kuang Z, Deng L. NGCICM: A Novel Deep Learning-Based Method for Predicting circRNA-miRNA Interactions. *IEEE/ACM Trans Comput Biol Bioinform.* 2023;20(5):3080–92.

27. Niu Z, Gao X, Xia Z, Zhao S, Sun H, Wang H, et al. Prediction of small molecule drug-miRNA associations based on GNNs and CNNs. *Front Genet.* 2023;14:1201934

28. Luo H, Liang H, Liu H, Fan Z, Wei Y, Yao X, et al. TEMINET: A



Co-Informative and Trustworthy Multi-Omics Integration Network for Diagnostic Prediction. *Int J Mol Sci.* 2024; 25(3):1655

Heterogeneous Graph Attention Networks. *Front Genet.* 2021;12:727744.

29. Zhong Y, Peng Y, Lin Y, Chen D, Zhang H, Zheng W, et al. MODILM: towards better complex diseases classification using a novel multi-omics data integration learning model. *BMC Med Inform Decis Mak.* 2023;23(1):82.

30. Tian Z, Han C, Xu L, Teng Z, Song W. MGCNSS: miRNA-disease association prediction with multi-layer graph convolution and distance-based negative sample selection strategy. *Brief Bioinform.* 2024;25(3):bbae168

31. Ji C, Wang Y, Ni J, Zheng C, Su Y. Predicting miRNA-Disease Associations Based on