Deep Learning for Alzheimer's Disease Drug Repurposing using Knowledge Graph and Multi-level Evidence

Kang-Lin Hsieh,[1]* German Plascencia-Villa,[2]* Ko-Hong Lin,[1]* George Perry,[2] Xiaoqian Jiang,[1] Yejin Kim[1]

*Contributed equally

[1]Center for Secure Artificial Intelligence for Healthcare, School of Biomedical Informatics, University of Texas Health Science Center at Houston

[2]Department of Neuroscience, Developmental and Regenerative Biology, University of Texas at San Antonio

Footnotes

[1]KH, GPV, KL contributed equally to this work.

[2]To whom correspondence may be addressed. Email: Yejin.Kim@uth.tmc.edu

[3]Author Contributions: XJ and YK provided motivation for the study and contributed to the idea development; KH, KL, and YK developed the computational method; KH and KL performed computational experiments; GPV and GP performed biological experiments; KH, KL, and YK wrote the initial manuscript; and KH, GPV, KL, GP, XJ, and YK edited the manuscript.

[4]Conflict of interest: The authors have no conflict of interest to declare.

Keywords: Alzheimer's Disease, drug repurposing, multi-task learning, systems biology, network medicine

ABSTRACT

Developing drugs for treating Alzheimer's disease (AD) has been extremely challenging and costly due to limited knowledge on underlying biological mechanisms and therapeutic targets. Repurposing drugs or their combination has shown potential in accelerating drug development due to the reduced drug toxicity while targeting multiple pathologies. To address the challenge in AD drug development, we developed a multi-task machine learning pipeline to integrate a comprehensive knowledge graph on biological/pharmacological interactions and multi-level evidence on drug efficacy, to identify repurposable drugs and their combination candidates. We developed and computationally validated a heterogeneous graph representation model with transfer learning from universal biomedical databases and with joint optimization with AD risk genes. Using the drug embedding from the heterogeneous graph representation model, we ranked drug candidates based on evidence from post-treatment transcriptomic patterns, mechanistic efficacy in preclinical models, population-based treatment effect, and Phase II/III clinical trials. We experimentally validated the top-ranked candidates in neuronal cells, identifying drug combinations with efficacy in reducing oxidative stress and safety in maintaining neuronal viability and morphology. Our neuronal response experiments confirmed several biologically efficacious drug combinations (e.g., Galantamine + Mifepristone). This pipeline showed that harmonizing heterogeneous and complementary data/knowledge, including human interactome, transcriptome patterns, experimental efficacy, and real-world patient data shed light on the drug development of complex diseases.

INTRODUCTION

Developing drugs for treating Alzheimer's disease (AD) has been extremely challenging and costly. The total cost of developing new AD drugs, including failures, is estimated at $5.7 billion - seven times more than the cost of developing cancer medicines.(1) The most recent FDA approval of Aducanumab is an amyloid beta-directed monoclonal antibody, which aims to treat patients by reducing the buildup of β-amyloid.(2) Despite showing some promising efficacy in removing β-amyloid, only patients with mild stage and amyloid burden can be benefited from the expensive treatment (estimated to cost $56,000 per year).(3) To develop potential treatments in a timely and cost-effective manner, drug repurposing has shown good potential due to the reduced risk of drug toxicity of previously approved drugs. What is even more promising is to study combinatorial drug repositioning, which has demonstrated their potential for synergistically treating complicated diseases, including cancer,(4) diabetes,(5) metabolic syndrome,(6) cardiovascular disease,(7) thanks to their ability to target multiple pathologies.

Most current AD drug repurposing investigates various perspectives, including drug perturbation transcriptome profiles, network pharmacology, and treatment effects in real-world patient data. Transcriptomic-based strategy compares drug-induced gene expression with gene expression in AD specimen,(8–10) which captures integrated molecular changes in different AD pathology. Network pharmacology is another approach that represents a drug's multi-target capacity in a human interaction network.(11–13) This method aims to identify hidden interaction among drugs and proteins (or disease targets) by estimating the proximity between entities in the network. Alternatively, the real-world data approach leverages large sets of heterogeneous patients' data to obtain off-label indication via counterfactual treatment effect estimation.(14, 15) Among these drug repurposing strategies, each strategy captures different aspects in multimodality and multiscale of the AD treatment landscape. These different approaches provide disparate but complementary evidence into potential drugs.(16) Success in a single aspect, however, cannot guarantee clinical effectiveness due to AD's heterogeneous pathogenesis and associated comorbid conditions. Consequently, there is a critical need to integrate diverse evidence of

transcriptomic observations, heterogeneous biological interactions, preclinical experimental results, and real-world observation to target multiple perspectives in AD drug development.

Thus, our objective is to develop a multi-task learning pipeline for AD drug repurposing and drug combination prediction. We first built an AD knowledge graph that integrates the heterogeneous human interactome of drugs, genes, pathways, and functional ontology directly or indirectly related to AD. A knowledge graph representation learning was applied to encapsulate the entities and heterogeneous interactions into embedding. Next, repurposable drug candidates were identified based on multi-level evidence from drug perturbation transcriptome profiles, clinical trial, and preclinical trial history, as well as population-based treatment effect estimation. We then mechanistically validated the efficacy and safety of our computationally predicted drug candidates using mouse brain cell lines. Our cell-based validation showed that our proposed framework can predict highly promising repurposable drug combinations.

RESULTS

AD knowledge graph representation

We integrated various sources of information from several human-curated databases to build a comprehensive AD knowledge graph using drugs, genes, pathways, and gene ontology, together with edges for drug-target interaction, gene-gene interactions, drug-drug structural similarity, pathway-gene, GO-gene, and GO-drug associations (Fig. 1a, Method 1). Using the AD knowledge graph, we encapsulated the molecular interactions into representation (embedding) using the knowledge graph representation method we designed (Method 2). The AD knowledge graph we built is specific to AD-related molecular interactions. Underlying AD etiology is an open research area and current knowledge of AD-related interactions is likely incomplete. To address the incomplete knowledge of AD, we utilized transfer learning from the universal biological database representation (Method 3). Additionally, we encouraged the gene embedding in the knowledge graph to distinguish high-risk AD-related genes, by jointly optimizing the target classification task and the graph representation (Method 4).

After training the model to learn heterogeneous relationships in the AD knowledge graph, we evaluated whether the embeddings faithfully condense the AD knowledge. We investigated whether the node embedding can restore the local interactions (Method 5, Table 1) and visually examined the global position of the node embedding via UMAP plot (Fig. 2).(20) Overall, the interaction (edge) prediction accuracy of embeddings with transfer learning and fine-tuning achieved AUROC=0.991 and AUPRC=0.988, which is higher than DRKG's universal embedding (AUROC=0.565, AUPRC=0.648) and AD knowledge graph representation without transfer learning (AUROC=0.959, AUPRC=0.958) (Table 1). This higher accuracy with transfer learning implies that the universal pharmacological interactions encompassing the AD-specific molecular interactions can complement the AD knowledge graph representation. By ablating one edge type and measuring the edge prediction accuracy (Method 5), we found that the drug-target or drug-pathway interactions are the most unique interactions that are hard to infer if they are excluded (Table S1). For the AD-related gene prediction task (Method 4), the gene embedding was well trained to distinguish AD-related genes from other remaining genes (Table S2). Visualizing the embedding via UMAP, we found

that the overall node distribution is generally dispersed throughout the plot rather than clustered with the same node type. Undergoing clinical trial drugs are aggregated in three main groups, each with similar therapeutic or pharmacological categories, implying the node embedding well reflects the contextual information. In all, the high accuracy in missing interaction prediction and AD-related gene prediction, together with global distribution in UMAP plot, implies that the node embedding faithfully captures the local and global network topology of the AD knowledge graph regardless of edge types, with the help of universal embedding that provides complementary information to restore the masked AD knowledge. This implied the potential applicability of our derived embeddings for drug repurposing tasks.

Identifying drug candidates using transcriptomic, mechanistic, and epidemiological evidence

After we synthesized the heterogeneous AD knowledge into the embedding, we identified repurposable drugs using the drug embedding that encapsulates molecular interactions and pharmacological interactions. Our strategy is to prioritize drugs with multi-level evidence on the drug's efficacy, such as transcriptomic reversed patterns, mechanistic efficacy, and population-based treatment effects (Method 7, Table S3), using multi-task learning (Method 6). As a result, the ranking model accuracy was AUROC=0.910 and AUPRC=0.472 to predict drugs with at least one of the drug efficacy labels. We presented the top-ranked drugs in Table S4 (precision at top 600=0.42) and highlighted the top 10 in Table 2.

Identifying drug combination

As indicated by the complexity of the AD interactome network, using single drugs to treat AD might result in limited effects. To improve treatment efficacy, we identified potential drug combinations from the top 600 ranked drugs and existing FDA-approved AD drugs. The FDA has approved four drugs for the treatment of symptoms and memory loss in AD – rivastigmine, galantamine, donepezil, memantine, and memantine combined with donepezil. These drugs have two main mechanisms of action, as cholinesterase (AChE) inhibitors or *N*-methyl D-aspartate (NMDA) antagonists. Hippocampal cholinergic hypofunction is closely related to the cognitive deficits of Alzheimer's disease. These pharmacological treatments are

intended to maintain or stabilize neuronal function, management of behavioral symptoms, and slow down the rate of memory loss by regulating neurotransmitters. In general, AChE inhibitors are recommended for the treatment of mild-to-moderate AD, whereas NMDA antagonist is prescribed to treat moderate-to-severe AD.(32) Unfortunately, after an expected period of daily administration of these AChE inhibitors and NMDA antagonist drugs, some side effects are experienced by AD patients, including nausea, vomiting, diarrhea, indigestion, muscle cramps, fatigue, loss of appetite, weight loss, dizziness, headache, and confusion. Here, we proposed to discover alternative and effective treatments for AD by combining the FDA-approved AD drugs with the top-ranked drugs identified in our computational model.

To identify drug combinations with synergistic interactions without degradation in safety, we leveraged the "Complementary Exposure pattern" (Method 8) indicating that a drug combination is therapeutically effective if the targets of the drugs hit the disease module without overlap.(33) We presented the identified drug combinations in Table S5.

In vitro validation on safety, tolerability, and neuronal responses of drug combinations

We mechanistically validated the top-ranked drug combinations by measuring cytotoxicity and oxidative stress on hippocampal neurons (Method 9, Fig. 1f). We identified significant efficacy in multiple drug combinations involving existing and novel drug candidates for AD.

Memantine functions as an antagonist of glutamatergic NMDA receptors, type 3 serotonin (5HT3) receptor, and nicotinic acetylcholine receptor (nAChRs), allowing neuronal availability of neurotransmitters and improving cognition and behavior in moderate-to-severe AD patients. In combination with Tolcapone, an FDA-approved catechol-O-methyltransferase inhibitor used in the management of Parkinson's disease to increase levels of peripheral dopamine, we observed safety and tolerability in cholinergic neurons during the dose- and time-course treatments. The production of reactive oxygen species (ROS) was reduced by 5.6±2.2% with Tolca, and in the presence of Aβ, only the combined dosage of Mem+Tolca was able to reestablish oxidative stress to basal levels, whereas the single Mem treatment presented a 5.5±2.7% increase (Fig. 3). Memantine is usually administered in combination with donepezil, but in combination with

Tolcapone, the drugs could have significant benefits on global and functional outcomes in AD subjects.(34, 35)

Donepezil, an AChE inhibitor, allows an increase in acetylcholine in cholinergic neurons. This compound has been used for 25 years to delay cognitive decline and symptoms of AD. High-rank drugs with potential activity tested included octyl-methoxycinnamate, nebivolol, desogestrel, and salicylic acid. The combination of Don+octyl-methoxycinnamate showed high tolerability and safety. Functionally, donepezil effectively reduced the amount of ROS by 6.82-7.98%, but its combination with octyl-met showed a slight increased by 4.3-9.1% (Fig. 3), this could be related to the limited solubility of this active compound and that it is usually employed in topical formations.

Galantamine is a weak AChE inhibitor and allosteric potentiator of acetylcholine receptors, intended to treat the progression of symptoms, memory loss, and thinking ability in mild-to-moderate AD. Gal prevents acetylcholine processing and stimulation of its release by nicotinic receptors. When galantamine was combined with mifepristone, caffeine, or diclofenac, it caused a reduction in oxidative stress responses by 5.8-6.3%. These neuroprotective effects were more evident in the neurons dosed with Aβ, observing a reduction of ROS produced of 17.7% with gal+mifep, 13.5% with gal+caffeine, and 2% with gal+diclofenac, whereas single galantamine treatment presented a 5.5% increase in neuronal oxidative stress (Fig. 3). The neuroprotective effect of mifepristone is related to preventing stress-induced responses, protecting neurons from undergoing programmed cell death (apoptosis) and increasing AMPA receptor expression,(36) also mifepristone may promote rapid repair of the synaptic alterations in the hippocampus.(37) In contrast, caffeine activates noradrenaline neurons and the releasing of dopamine, and counteracting the development of cognitive impairment as an antioxidant.(38) Diclofenac is an anti-inflammatory drug targeting voltage-dependent K^+ channels, with neuroprotective effects in neurons.(39) In pilot studies, diclofenac showed positive effects to reduce the risk of AD in veterans.(40)

Rivastigmine is an AChE and butyrylcholinesterase inhibitor, used to treat memory loss and to improve thinking abilities in AD patients. The tolerability and safety of rivastigmine or in combination with lithium, vitamin E, and diclofenac was tested, observing that these drugs were well-tolerated by cholinergic neurons.

Neuronal responses indicated that these active compounds effectively reduced oxidative stress in neurons by 4.8-6.2%. In the presence of Aβ, the rivastigmine, riv+lithium, and riv+vit E were able to regulate ROS with similar levels for all these treatments, whereas riv+diclof reduced by 6.8% the amount of intracellular ROS (Fig. 3). At the neuronal level, lithium can reduce excitatory neurotransmitters (dopamine and glutamate) and increase GABA.(41) Lithium formulation has been used as a treatment to prevent cognitive impairment in dementia and is currently under clinical trials (Phase III). The results indicate that long-term administration of lithium attenuated the cognitive and functional decline in mild cognitive impairment and increased CSF-AD biomarkers.(42) Vitamin E (tocopherol) is a neuroprotectant with antioxidant properties,(43) this compound is under clinical trials (Phase II/III) to determine effects on clinical progression of AD, the formulations tested include pure vit E or in combination with selenium or memantine.

DISCUSSION

The objective of this study was to provide proof-of-concept hypotheses on drug combinations to treat AD by knowledge graph embedding of heterogeneous interactome and multi-task learning of complementary evidence on drug efficacy. Our AD knowledge graph embedding boosted by pretrained universal pharmacological embedding integrates not only the human interactomes but also millions of interactions in public biomedical literature. Our multi-task learning of AD-related genes and multi-level drug efficacy closes the gap between the scattered knowledge in AD research.

Our methodology tackles important challenges to harmonize fragmented multimodal data. Recently, several prior works have used the graph neural network approach for drug development, either for general purpose,(44, 45) for rare diseases,(46) or for specific diseases including COVID-19(47–49) and AD.(50) Our methodological strength over the prior approach is i) transfer learning of universal embedding to integrate broader modality to address the incomplete knowledge, ii) multi-task learning for multimodal interactions, AD-related genes, and multi-level drug efficacy to mitigate the lack of established therapeutic targets, and iii) *in vitro* validation of the computational model via hippocampal neuron cells.

Drug combinations that we validated are less toxic (FDA-approved drugs) and economically viable (generic drugs). AD is multifactorial and the use of multi-target therapeutic interventions addressing several molecular targets seems to be an alternative approach to modify the course of AD progression.(51) Different clinical studies demonstrate that a combination therapy is more efficient than monotherapy, especially in mild-to-moderate AD to slow down the rate of cognitive impairment. Our integrative approach has identified promising drug combinations adjuvant to existing FDA-approved AD drugs. All drug combinations we identified are available in generic formulations, making them economically viable and available as an alternative to patented formulations with a high monthly cost (>500 USD).

Our study has several limitations. Instead of a mechanism-based approach that identifies therapeutic targets first and identifies drugs interacting with them, our approach relies on historical or data-driven evidence, such as estimated drug efficacy (transcriptomic reversed patterns, population-based treatment effect) and existing efficacy (preclinical or clinical trials). If this data-driven drug efficacy is biased, our model's drug ranking can be also biased. For example, drugs targeting amyloid beta are much more studied than drugs targeting other pathologies, thus our model has a potential risk to trap in the circularity of failed drugs targeting amyloid beta. Population-based treatment effects in real-world data can partly mitigate the bias in historical research data toward amyloid-beta. Another limitation is that the identified drug combinations are based on neither pharmacokinetics/pharmacodynamics nor biological experimental data on synergistic drug combinations in treating AD. Instead, our approach was to use the complementary exposure pattern, which oversimplifies the synergistic mechanism of actions.

Despite the limitations, our drug repurposing strategy has a broad impact not only on AD but also on complex diseases, in which complex mechanisms of actions and therapeutic targets are an open question. The well-known deficiency of knowledge and insufficiency of data in AD research, when studied separately, prevent us from capturing the complex AD's heterogeneous pathogenesis. Our pipeline integrates complementary information at multiple stages (such as transcriptomic patterns, preclinical efficacy, clinical partial effectiveness, and population-based treatment effect) by jointly optimization to

better identify promising drug combinations. Hypotheses generated from this integrative approach may shed light on complex disease drug development.

MATERIALS AND METHODS

1. Build AD knowledge graph

We built the AD knowledge graph from various information sources. We queried "Alzheimer's Disease" and downloaded AD-related multimodal interactions from the Comparative Toxicogenomics Database (CTD), a curated knowledge base with genes, drugs, GO, and pathways.(52) CTD compiles various interactions through text mining and manual curations by domain experts, then infers associations based on the relationships between the entities.(52, 53)

AD-related genes. We derived high-confident AD-related genes from Agora(54) and CTD. Agora's nominated gene list, contributed by researchers from the National Institute on Aging's Accelerating Medicines Partnership in Alzheimer's Disease (AMP-AD) consortium, was identified using computational analyses of high-dimensional genomic, proteomic, and/or metabolomic assays in human samples. CTD's gene list, on the other hand, was curated from literature to either have a biomarker, have a therapeutic target, or have indirect interactions with AD. In this study, we extracted 743 AD-related genes. Those genes were further connected by pathways and related gene ontologies.

Drugs and genes. We collected precompiled drug-gene interactions from CTD, including 6,543 unique FDA-approved drugs or experimental compounds, as well as 16,997 genes that interact with those drugs. 119 drugs have direct evidence reporting either the drugs' relevance to AD etiology (e.g., Streptozocin) or treatment (e.g., Donepezil). The remaining drugs are from literature reporting the potential efficacy of the drugs. 117 genes have direct evidence reporting either the gene's relevance to AD etiology (e.g., APOE) or treatment (e.g., LEP). The remaining genes are from literature reporting the potential relevance of the genes to AD. The 110,637 drug and target gene interactions are from literature with cell-based or animal-based evidence.

Gene and gene. We retrieved gene-gene interaction from the STRING database.(55) STRING has seven types of evidence for protein-protein interaction: biological experiment, pathway inference, text mining, gene co-expression, neighborhood genes in chromosomes, fusion, and co-occurrence.(55) We selected

123,738 protein-protein interactions that have >95% combined confidence scores across the seven points of evidence.

Drug and drug structural similarity. A drug with a similar structure usually accompanies similar targets. To best represent each compound's chemical structure, we utilized four types of fingerprints: atom-pair fingerprints,(56) MACCS fingerprint,(57) Morgan/Circular,(58) and Topological-torsion fingerprints.(59) Sørensen–Dice similarity coefficients were computed between every two drugs (or compounds) using RDkit (version 2021.03.4).(60) For each type of fingerprint, we calculated the *z*-score of the Sørensen–Dice coefficient to obtain pairwise compound similarity and defined compound pairs with *z*-score>=3 to be structurally similar. In total, we determined 37,669 compound pairs that are structurally similar to each other.

Pathway and AD-related genes. For each AD-related gene, CTD linked them with their involved pathways from Reactome(61) and KEGG(62) databases. Various pathways of a multifactorial disease like AD can display the diverse etiology and mechanisms of the disease. Querying "Alzheimer's Disease" from CTD, we extracted 1,778 gene-pathway interactions containing 678 unique pathways.

Pathway and drugs. Among 6,543 drugs of interest, their drug targets from CTD were used to query related Reactome(61) and KEGG(62) pathways and computed enriched drug-pathway associations using Enrichr.(63) For each drug, enriched pathways for all its target genes were identified using a threshold of p-value < 0.05. Drugs and their enriched pathways were connected as drug-pathway interactions. There are 110,637 drug-pathway interactions, including 1,622 pathways in our knowledge graph.

GO and AD-related genes. Gene ontologies are structured and controlled vocabularies that represent each gene's participating biological processes, molecular functions, and cellular components. We extracted gene ontologies for AD-related genes from CTD to construct gene-GO edges. A total of 7,368 gene-GO interactions were collected between 89 AD risk genes and 2,965 unique GO terms.

GO and drugs. CTD also inferred associations between GO and certain drugs under Alzheimer's Disease. Drug-GO interactions were established based on either a combination of curated drug-GO interactions and

drug-AD interactions, or a gene-GO annotation in AD, or both. These expanded our AD knowledge graph with 6,817 drug-GO interactions, including 2,377 unique GO terms and 108 unique drugs.

2. Knowledge graph representation

Graph neural network (GNN) is one field of deep neural networks that derive a vectorized representation of nodes, edges, or whole graphs. Adopting GNN into the biomedical network facilitates the integration of multimodal and complex relationships. The graph node embedding can preserve the node's local role and global position in the graph via iterative and nonlinear message passing and aggregation. It learns the structural properties of the neighborhood and the graph's overall topological structure (64), thus allowing us not only to restore the known interactions but also infer unknown interactions. The knowledge graph representation that contains such putative interactions can help us find new indications of drugs.(65) Recently GNN has demonstrated a great advance in predicting hidden interactions (e.g., PPIs, drug-drug adverse interactions, and drug-target interactions) and the discovery of new molecules (45, 66). GNN is also used to derive representation from a graph with multiple types of relations and nodes (i.e., heterogeneous network or knowledge graph). The Drug Repurposing Knowledge Graph (DRKG) uses the knowledge graph representation to capture whole topological relations from seven biomedical databases (49)(67). The knowledge graph includes 97,238 nodes (from 13 node types including gene, molecular function, pathway, disease, symptom, anatomy, cellular component, compounds, side effect, ATC, and pharmacologic class) and 5,874,261 interactions (from 107 edge types) from DrugBank,(68) HetioNet,(69) GNBR,(67) STRING,(70) and IntAct.(71) Their representation offers a general and universal embedding of these entities, which can further enrich other domain-specific knowledge representations.

In this study, we developed the knowledge graph representation model by customizing the deep variational graph neural autoencoders (VGAE) to incorporate multiple types of relations or edges (Fig. 1b).(17) The self-supervised graph autoencoder method encodes the nodes into a latent vector (embedding) and reconstructs the given graph structure (i.e., graph adjacency matrix) with the encoded latent vector. The variational graph approach can encapsulate the complex network into a probabilistic distribution (rather

than deterministic vector representation) for nodes considering the uncertainties of our knowledge for AD; therefore, it alleviates the overfitting/underfitting issues due to partial knowledge. We customized the VGAE to incorporate the different types of edge by setting different weight matrices for each type of edge. This multi-relational approach respects the different propagation for each interaction type. Since our objective is to derive node embedding that reflects the overall graph topology, the multi-relational VGAE model was trained to reconstruct the missing interaction using the node embeddings as an autoencoding manner. For the optimization, in the message-passing step, each node (entity)'s embedding is iteratively updated by aggregating the neighbors embedding, in which the aggregation function is a mean of the neighbor's features, concatenation with current embedding, and a single layer of a neural network on the concatenated one. We set the embedding size as 128 after several trials. We used PyTorch Geometric for implementation (72). The model structure was (1×400) → Graph convolution to (1×256) → RELU → Dropout → Summation of multiple edge types → Batch norm → Graph convolution to 1×128 (mean) and 1×128 (variance). We randomly split the knowledge graph edge (i.e., positive edges) into 90% for training and 10 % for testing. We randomly generated the same number of fake or negative edges from which the positive edges will be predicted.

3. Transfer learning from a large drug knowledge graph

We boosted the node embedding's biological relevance in the AD knowledge graph with millions of interactions from public biomedical literature by transfer learning. Transfer learning is to transfer knowledge from previously learned universal models to domain-specific models. By transferring and fine-tuning the universal knowledge, the AD network embedding can respect universal pharmacological interactions while prioritizing AD-related interactions. Injecting this universal knowledge into the AD domain is critically important in addressing the uncertainty in AD knowledge. We used the DRKG's pretrained node embedding as initial node features in the AD knowledge graph. A recent attempt (e.g., DRKG network(49)) has been made to obtain the universal embedding of 15 million pharmacological entities considering 39 different types of interactions among genes and compounds from seven large

biomedical databases including DrugBank, HetioNet, GNBR, STRING, IntAct, and DGLdb.(44) We retrieved the pretrained embedding for drugs, genes, pathways, and functional ontology from DRKG. We set the initial node feature of the AD knowledge graph as the pre-trained embedding of DRKG and fine-tuned them to the AD knowledge graph. When nodes in the knowledge graph were not matched to entities in the universal embedding, we learned the embedding as parameters (e.g., one hot to embedding vectors).

4. Multi-task learning to distinguish high-confidence AD-related genes from other genes

One of the most important tasks in AD drug development is identifying driver genes related to AD. We encouraged the knowledge graph representation on genes to discriminate the AD-related genes out of the 16,997 genes in the graph. We obtained 743 high-confidence AD-related genes from Agora(54) and CTD. We hypothesized that one can predict AD-related genes using gene interaction with other genes, GO, and pathways. We created a node classification task to predict whether a gene is one of the AD-related genes using the gene representation in the AD knowledge graph. The node classification consists of two graph convolution layers (GraphSAGE) that sample a node's adjacent neighborhoods and aggregate their representation. The dimensions of those two convolution layers were Graph convolution (128, 64) → RELU → Dropout → Graph convolution (64, 1). We split the gene nodes into 90% for training and 10% for testing. As the labels are not balanced (only 743 AD-related genes out of 16,997 genes), we randomly split the gene nodes into 90% for training and 10% for testing while maintaining the ratio of AD gene and non-AD gene. This AD-related gene classification task was jointly optimized with the relational VGAE so that the AD knowledge graph representation can distinguish AD-related genes from all other genes. We used a multi-task optimization technique that resolves conflict among multiple tasks' gradients by normal vector projections,(73) to achieve a better local optimum.

5. Validating the quality of knowledge graph representation

We evaluated the quality of AD knowledge graph representation by checking whether the learned node embedding can restore the known edges. The edge prediction task is to predict whether an edge exists

between two nodes using the two nodes' embedding. The edge prediction task's accuracy will be high if the node embedding encapsulates the edge information faithfully. For training and test, we randomly set aside 10% of the edges for evaluation. We randomly sampled negative or fake edges that do not exist in the knowledge graph (within the same node types) and predicted whether a given edge is real or fake.(17) We measured the edge prediction accuracy by each edge type, to check whether any edge types were neglected or underrepresented in the node embedding. The model was trained using all edge types and evaluated separately for each edge type. We also ablated one edge type and observed how one edge type contributed to predicting all the edge types (Table S1). This experiment allowed us to determine which interactomes are unique and hard to be replaced or implicitly inferred by other edge types. The edge prediction accuracy was measured as the area under the receiver operating curve (AUROC) and the area under the precision-recall curve (AUPRC). For the node classification task to distinguish between AD-related genes and remaining genes, we also measured the classification accuracy using the AUROC and AUPRC.

6. Transcriptomic, mechanistic, and epidemiological evidence for drug efficacy

We built a pipeline to prioritize drugs that are similar to drugs with historical evidence. Particularly, we focused on drugs with multiple historical evidence. AD is a complex disease and its effective therapeutic targets are largely unknown, thus drugs interacting with a few targets can have limited efficacy.(74) A holistic evidence may alternatively capture the efficacy in the systemic and complex disease. We believe the most desirable drug candidates would have multiple levels of complementary evidence to support their efficacy, such as post-treatment transcriptomic patterns, mechanistic efficacy in preclinical animal models, and epidemiological treatment effects in real-world patient data. We identified each evidence as follow:

Transcriptomic reversed patterns: We examined reversed patterns of the AD model's genetic signature and drug-induced gene signature using gene set enrichment analysis. We leveraged Gareth et al. 's iPSC-derived cortical neurons studies(9) and identified 149 drugs with significant transcriptomic reversed patterns. Among the 149 drugs, 76 drugs were matched to our 6,543 drug candidates.

Mechanistic efficacy in preclinical models: Alzheimer's Disease Preclinical Efficacy Database (AlzPED) accumulated pre-clinical evidence for effective AD drugs. We identified 338 drugs with significant efficacy from 791 animal-based studies targeting various pathologies including amyloid-beta, tau, and neuroinflammation. Among the 338 drugs, 189 drugs were matched to our 6,543 drug candidates.

Population-based treatment effect: We estimated drugs' population-based treatment effect as real-world evidence of drug efficacy. Healthcare claim data is a unique resource to investigate treatment effects of prescribed drugs on reducing AD onset risk during the preclinical stage of AD. We used Optum's de-identified Clinformatics® Data Mart subscribed by UTHealth. This billing-purpose claim data is inherently noisy and sparse, thus requires careful data preparation (e.g., setting observation windows to exclude subjects who are not old enough to have AD onset, identifying AD onset via billing-purpose diagnosis and medication codes, grouping high-dimensional diagnosis codes into clinical comorbidities as confounding variables). We followed the data preparation process in our previous work(75) based on target trials.(76) We calculated the average treatment effect among treated (ATT) using inverse probability of treatment weighting (IPTW).(77, 78) See details at Supplementary Method 1. We identified 126 drugs with positive treatment effects (i.e., ATT<0 to reduce AD onset). Among the 126 drugs, 101 drugs were matched to our 6,543 drug candidates.

Clinical trials failed at Phase II/III: Many drugs in the failed clinical trials passed several phases and showed non-trivial response rates (e.g., good response on certain subpopulations). We hypothesize individual drugs (focusing on different targets) from previously failed clinical trials respond differently on sub-populations, and therefore, cocktails of complementary drugs might cover a larger population. Also, because drug testing in clinical trials is partially efficacious in treating AD, a drug that is similar to these trial drugs (and their combinations) can have potential efficacy too. We collected 294 drugs from 2,675 interventional clinical trials for AD from ClinicalTrial.gov. Among the 294 drugs, 210 drugs were matched to our 6,543 drug candidates.

7. Prioritizing repurposable drugs with multiple evidence and multi-task learning

We extracted four positive labels from extensive calculation and search (Method 7). We applied the multi-task learning(73) to jointly build a drug ranking model that ranks drugs high if the drugs satisfy multiple evidence. We designed a ranking model with neural networks and Bayesian pairwise ranking loss.(79, 80) The positive samples to prioritize were the drugs with positive evidence for each task; the unlabeled or unknown samples were the remaining drugs in the AD knowledge graph. The architecture was two fully connected layers (with the size of 128→128→1) with residual connection, nonlinear activation (ReLU), dropout, batch norm in the middle, and the optimization loss (Bayesian pairwise ranking loss). The model was trained to minimize four loss functions for each task. We validated the ranking model by evaluating whether the positive drugs (under clinical trials) are ranked high. We measured the accuracy of the drug ranking model using the area under the receiver operating curve (AUROC) and area under the precision-recall curve (AUPRC) of top k drugs with 50% training and 50% test cross-validation. We purposely set the portion of the training set lower because the clinical trials are not our sole "gold standard" to prioritize drugs.

8. Drug combination search

We identified efficacious drug combinations from top-ranked drugs. Our approach is to leverage the network of high-confidence drugs' targets and AD risk genes. Our hypothesis was that "a drug combination is therapeutically effective only if the targets of the drugs both hit the disease module, but they target a separate neighborhood (Complementary Exposure pattern)",(33) which has been successfully applied to hypertension(33) and COVID-19.(48) We searched the drug combinations within the top-ranked drugs. We counted the number of genes in the AD module that a drug combination hits, where the drug combination's targets are disjoint.

9. In vitro validation on safety, tolerability, drug combinations

Neuronal culture. Hippocampal neurons (HT-22) were cultured in DMEM (Sigma-Aldrich #D6429) supplemented with 10% FBS, and 1X Antibiotic-Antimycotic (#30-004-CI, Corning). Neurons were

subcultured every 3-4 days at approximately 75-80% of confluence and kept in a T-75 flask in a 5% CO2 incubator at 37°C. For bioassays, the neurons were detached with 0.25% trypsin-EDTA (Sigma-Aldrich #T3924), pelleted at 300g, and diluted in fresh DMEM. Cells were counted with LUNA-II automated cell counter (Logos) and seeded in 96-well-plates at 5000 cells/well the day before. Morphology was monitored with an inverted optical microscope (Olympus CKX31).

Drug combinations. All the drugs were ACS grade or higher purity, prepared fresh immediately prior to use in the bioassays. Stock solutions of the drugs were prepared by dissolving in DMSO (Sigma-Aldrich #) at 50 mM. Each active compound was then diluted with DMEM to 200, 100, 50, and 25 μM. As a control, we employed DMEM with DMSO.

Cytotoxicity and oxidative stress responses. Cell Counting Kit-8 (Dojindo, #CK04) was used to determine cell viability, proliferation, and cytotoxicity, measuring the absorbance of water-soluble tetrazolium salt (reduced WST-8) at 450 nm. Neurons were subcultured in 96-well plates and dosed with the active drugs diluted in DMEM, incubated at 37°C, and absorbance measured after 24, 48, and 72 h of treatment (SynergyLX multi-mode plate reader, Biotek). Each treatment was performed in triplicates. Oxidative stress responses of dosed neurons were determined with Fluorometric Intracellular ROS kit (Sigma-Aldrich #MAK142). Cells were subcultured into Costa-Assay (Corning #3904) clear flat bottom, tissue culture treated, black-coated 96-well plates. Neurons were dosed with active drugs (as described), incubated for 24, 48, or 72 h, and then fluorescence of intracellular ROS was determined following the protocol of the kit using DCFDA as a fluorogenic sensor (lex=650/lem=675nm) (SynergyLX multi-mode plate reader, Biotek). The same protocol was followed using 10 μM $A\beta_{42}$ (Anaspec, #AS24224) in combination with active drugs at 50 μM to quantify intracellular ROS in neurons. Each treatment was performed in triplicates.

Imaging. Neurons were analyzed in fluorescence, bright-field, phase contrast, live-cell, and Z-stack imaging with Digital Imaging System Celena-S (Logos). Micrographs were analyzed with Fiji-Image J.

Statistical analysis. Data acquired from bioassays were normalized to control for each data set, the normalized average is presented as mean ± standard error of the mean (SEM). SigmaPlot was used for the analysis.

Acknowledgments

KLH was supported by CPRIT RP140113 (Computational Cancer Biology Training Program Fellowship from Gulf Coast Consortia). XJ is CPRIT Scholar in Cancer Research (RR180012), and he was supported in part by Christopher Sarofim Family Professorship, UT Stars award, UTHealth startup, the National Institute of Health (NIH) under award number R01AG066749. YK was supported in part by UTHealth startup and national Institute of Health (NIH) under award number R01AG066749.

Data Availability

All study data are publicly available from CTD, STRING, Agora, ClinicalTrial.gov, and AlzPED. Experimental results are available at supplementary materials. The code is available at https://github.com/freshnemo/AD-KG.

Supplementary materials

Table S1. Ablation study on edge prediction.

Table S2. Node classification task to predict therapeutic targets out of all gene nodes.

Table S3. Drugs with statistically significant treatment effects in reducing AD onset in Optum claim data.

Table S4. Top 100 drugs predicted by multitask learning model and multi-level evidence.

Table S5. Drug combinations that satisfy the complementary exposure pattern from the top 30 drugs.

Table S6. Subject's demographics for population-based treatment effect estimation.

Supplementary method 1. Population-based treatment effect

REFERENCES

1. J. Cummings, C. Reiber, P. Kumar, The price of progress: Funding and financing Alzheimer's disease drug development. *Alzheimers. Dement.* **4**, 330–343 (2018).

2. Office of the Commissioner, FDA Grants Accelerated Approval for Alzheimer's Drug (2021) (June 14, 2021).

3. W. K. Fleming, C. R. Brown, W. H. Shrank, Costly New Alzheimer Disease Medications on the

Horizon—Financing Alternatives for Medicare. *JAMA Health Forum* **1**, e201148 (2020).

4. Yejin Kim, Shuyu Zheng, Jing Tang, W. Jim Zheng, Zhao Li, Xiaoqian Jiang, Anti-cancer Drug Synergy Prediction in Understudied Tissues using Transfer Learning. *J. Am. Med. Inform. Assoc.*

5. A. Cahn, W. T. Cefalu, Clinical Considerations for Use of Initial Combination Therapy in Type 2 Diabetes. *Diabetes Care* **39 Suppl 2**, S137–45 (2016).

6. G. Bakris, *et al.*, Differences in glucose tolerance between fixed-dose antihypertensive drug combinations in people with metabolic syndrome. *Diabetes Care* **29** (2006).

7. Working Group on the Summit on Combination Therapy for CVD, Combination pharmacotherapy to prevent cardiovascular disease: present status and challenges. *Eur. Heart J.* **35**, 353–364 (2013).

8. J. C. Siavelis, M. M. Bourdakou, E. I. Athanasiadis, G. M. Spyrou, K. S. Nikita, Bioinformatics methods in drug repurposing for Alzheimer's disease. *Brief. Bioinform.* **17** (2016).

9. G. Williams, *et al.*, Drug repurposing for Alzheimer's disease based on transcriptional profiling of human iPSC-derived cortical neurons. *Transl. Psychiatry* **9**, 220 (2019).

10. M. Sirota, *et al.*, Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci. Transl. Med.* **3**, 96ra77 (2011).

11. T.-H. Pham, Y. Qiu, J. Zeng, L. Xie, P. Zhang, A deep learning framework for high-throughput mechanism-driven phenotype compound screening. *bioRxiv* (2020) https:/doi.org/10.1101/2020.07.19.211235.

12. L. Huang, *et al.*, Driver network as a biomarker: systematic integration and network modeling of multi-omics data to derive driver signaling pathways for drug combination prediction. *Bioinformatics* **35**, 3709–3717 (2019).

13. Y. Chen, R. Xu, Context-sensitive network analysis identifies food metabolites associated with Alzheimer's disease: an exploratory study. *BMC Medical Genomics* **12** (2019).

14. N. Geifman, R. D. Brinton, R. E. Kennedy, L. S. Schneider, A. J. Butte, Evidence for benefit of statins to modify cognitive decline and risk in Alzheimer's disease. *Alzheimer's Research & Therapy* **9** (2017).

15. J. M. Zissimopoulos, D. Barthold, R. D. Brinton, G. Joyce, Sex and Race Differences in the Association Between Statin Use and the Incidence of Alzheimer Disease. *JAMA Neurol.* **74**, 225–232 (2017).

16. F. Cheng, *et al.*, Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.* **9**, 2691 (2018).

17. T. N. Kipf, M. Welling, Variational Graph Auto-Encoders (2016) (July 16, 2020).

18. M. Schlichtkrull, *et al.*, Modeling Relational Data with Graph Convolutional Networks (2017) (July 21, 2020).

19. Y. Zhou, *et al.*, Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery* **6**, 1–18 (2020).

20. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction (2018) (October 15, 2021).

21. K. Rockwood, J. Mintzer, L. Truyen, T. Wessel, D. Wilkinson, Effects of a flexible galantamine dose in Alzheimer's disease: a randomised, controlled trial. *J. Neurol. Neurosurg. Psychiatry* **71**, 589–595 (2001).

22. Z. M. Schrire, *et al.*, Feasibility of 3-month melatonin supplementation for brain oxidative stress and sleep in mild cognitive impairment: protocol for a randomised, placebo-controlled study. *BMJ Open* **11**, e041500 (2021).

23. E. Mhillaj, *et al.*, Celecoxib Exerts Neuroprotective Effects in β-Amyloid-Treated SH-SY5Y Cells Through the Regulation of Heme Oxygenase-1: Novel Insights for an Old Drug. *Front Cell Dev Biol* **8**, 561179 (2020).

24. N. Pomara, P. M. Doraiswamy, H. Tun, S. Ferris, Mifepristone (RU 486) for Alzheimer's disease. *Neurology* **58**, 1436 (2002).

25. J. W. Ashford, Treatment of Alzheimer's Disease: Trazodone, Sleep, Serotonin, Norepinephrine, and Future Directions. *J. Alzheimers. Dis.* **67**, 923–930 (2019).

26. D. Van Dam, K. Coen, P. P. De Deyn, Ibuprofen modifies cognitive disease progression in an Alzheimer's mouse model. *J. Psychopharmacol.* **24**, 383–388 (2010).

27. G. Tong, *et al.*, Effects of single doses of avagacestat (BMS-708163) on cerebrospinal fluid Aβ levels in healthy young men. *Clin. Drug Investig.* **32**, 761–769 (2012).

28. C. R. Hopkins, ACS chemical neuroscience molecule spotlight on BMS-708163. *ACS Chem. Neurosci.* **3**, 149–150 (2012).

29. L. Escribano, *et al.*, Rosiglitazone reverses memory decline and hippocampal glucocorticoid receptor down-regulation in an Alzheimer's disease mouse model. *Biochem. Biophys. Res. Commun.* **379**, 406–410 (2009).

30. S. C. Hartsel, T. R. Weiland, Amphotericin B binds to amyloid fibrils and delays their formation: a therapeutic mechanism? *Biochemistry* **42**, 6228–6233 (2003).

31. B. Winblad, *et al.*, Phenserine efficacy in Alzheimer's disease. *J. Alzheimers. Dis.* **22**, 1201–1208 (2010).

32. G. Plascencia-Villa, G. Perry, Alzheimer's Disease Pharmacology. *Reference Module in Biomedical Sciences* (2021) https:/doi.org/10.1016/b978-0-12-820472-6.00018-9.

33. F. Cheng, I. A. Kovács, A.-L. Barabási, Network-based prediction of drug combinations. *Nat. Commun.* **10**, 1197 (2019).

34. R. Fremont, *et al.*, Tolcapone Treatment for Cognitive and Behavioral Symptoms in Behavioral Variant Frontotemporal Dementia: A Placebo-Controlled Crossover Study. *J. Alzheimers. Dis.* **75**, 1391–1403 (2020).

35. J. A. Apud, *et al.*, Tolcapone improves cognition and cortical information processing in normal human subjects. *Neuropsychopharmacology* **32**, 1011–1020 (2007).

36. A. M. Ghoumari, *et al.*, Neuroprotective effect of mifepristone involves neuron depolarization. *FASEB J.* **20**, 1377–1386 (2006).

37. L.-M. Wu, *et al.*, Mifepristone repairs region-dependent alteration of synapsin I in hippocampus in rat model of depression. *Neuropsychopharmacology* **32**, 2500–2510 (2007).

38. P. Londzin, M. Zamora, B. Kąkol, A. Taborek, J. Folwarczna, Potential of Caffeine in Alzheimer's Disease—A Review of Experimental Studies. *Nutrients* **13**, 537 (2021).

39. S. Naeem, R. Najam, S. S. Khan, T. Mirza, B. Sikandar, Neuroprotective effect of diclofenac on chlorpromazine induced catalepsy in rats. *Metab. Brain Dis.* **34**, 1191–1199 (2019).

40. O. Stuve, R. A. Weideman, D. M. McMahan, D. A. Jacob, B. B. Little, Diclofenac reduces the risk of Alzheimer's disease: a pilot analysis of NSAIDs in two US veteran populations. *Therapeutic Advances in Neurological Disorders* **13**, 175628642093567 (2020).

41. G. S. Malhi, M. Tanious, P. Das, C. M. Coulston, M. Berk, Potential mechanisms of action of lithium in bipolar disorder. Current understanding. *CNS Drugs* **27**, 135–153 (2013).

42. O. V. Forlenza, M. Radanovic, L. L. Talib, W. F. Gattaz, Clinical and biological effects of long-term lithium treatment in older adults with amnestic mild cognitive impairment: randomised clinical trial. *Br. J. Psychiatry* **215**, 668–674 (2019).

43. C. K. Sen, S. Khanna, S. Roy, Tocotrienol: the natural vitamin E to defend the nervous system? *Ann. N. Y. Acad. Sci.* **1031**, 127–142 (2004).

44. gnn4dr, gnn4dr/DRKG. *GitHub* (July 15, 2020).

45. S. K. Mohamed, V. Nováček, A. Nounu, Discovering protein drug targets using knowledge graph embeddings. *Bioinformatics* **36**, 603–610 (2020).

46. D. N. Sosa, *et al.*, A Literature-Based Knowledge Graph Embedding Method for Identifying Drug Repurposing Opportunities in Rare Diseases. *Pac. Symp. Biocomput.* **25** (2020).

47. D. M. Gysi, *et al.*, Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021).

48. K. Hsieh, *et al.*, Drug Repurposing for COVID-19 using Graph Neural Network with Genetic, Mechanistic, and Epidemiological Validation. *Research Square* https:/doi.org/10.21203/rs.3.rs-114758/v1 (September 22, 2021).

49. X. Zeng, *et al.*, Repurpose Open Data to Discover Therapeutics for COVID-19 using Deep Learning (2020) (July 23, 2020).

50. E. Sügis, *et al.*, HENA, heterogeneous network-based data set for Alzheimer's disease. *Scientific Data* **6**, 1–18 (2019).

51. M. T. Kabir, *et al.*, Combination Drug Therapy for the Management of Alzheimer's Disease. *International Journal of Molecular Sciences* **21**, 3272 (2020).

52. A. P. Davis, *et al.*, The Comparative Toxicogenomics Database: update 2017. *Nucleic Acids Res.* **45**, D972–D978 (2017).

53. A. P. Davis, *et al.*, Comparative Toxicogenomics Database: a knowledgebase and discovery tool for chemical-gene-disease networks. *Nucleic Acids Res.* **37**, D786–92 (2009).

54. Agora (August 31, 2021).

55. D. Szklarczyk, *et al.*, The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362 (2017).

56. R. E. Carhart, D. H. Smith, R. Venkataraghavan, Atom pairs as molecular features in structure-activity studies: definition and applications. *Journal of Chemical Information and Modeling* **25**, 64–73 (1985).

57. J. L. Durant, B. A. Leland, D. R. Henry, J. G. Nourse, Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inf. Comput. Sci.* **42**, 1273–1280 (2002).

58. H. L. Morgan, The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *Journal of Chemical Documentation* **5**, 107–113 (1965).

59. R. Nilakantan, N. Bauman, J. Scott Dixon, R. Venkataraghavan, Topological torsion: a new molecular descriptor for SAR applications. Comparison with other descriptors. *Journal of Chemical Information and Modeling* **27**, 82–85 (1987).

60. G. Landrum, RDKit (September 22, 2021).

61. A. Fabregat, *et al.*, The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).

62. M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, K. Morishima, KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).

63. A. Subramanian, *et al.*, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

64. W. L. Hamilton, R. Ying, J. Leskovec, Inductive Representation Learning on Large Graphs (2017) (July 16, 2020).

65. T. Gaudelet, *et al.*, Utilizing graph machine learning within drug discovery and development. *Brief. Bioinform.* (2021) https:/doi.org/10.1093/bib/bbab159 (August 31, 2021).

66. M. Zitnik, M. Agrawal, J. Leskovec, Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics* **34**, i457–i466 (2018).

67. B. Percha, R. B. Altman, A global network of biomedical relationships derived from text. *Bioinformatics* **34**, 2614–2624 (2018).

68. V. Law, *et al.*, DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.* **42**, D1091–7 (2014).

69. D. S. Himmelstein, *et al.*, Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife* **6** (2017).

70. D. Szklarczyk, *et al.*, STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2018).

71. IntAct, IntAct (June 29, 2021).

72. M. Fey, J. E. Lenssen, Fast Graph Representation Learning with PyTorch Geometric (2019) (August 6, 2020).

73. T. Yu, *et al.*, Gradient Surgery for Multi-Task Learning (2020) (August 31, 2021).

74. M. J. Oset-Gasque, J. Marco-Contelles, Alzheimer's Disease, the "One-Molecule, One-Target" Paradigm, and the Multitarget Directed Ligand Approach. *ACS Chemical Neuroscience* **9**, 401–403 (2018).

75. Y. Ling, P. Upadhyaya, L. Chen, X. Jiang, Y. Kim, Heterogeneous Treatment Effect Estimation using machine learning for Healthcare application: tutorial and benchmark (2021) (October 18, 2021).

76. M. A. Hernán, J. M. Robins, Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am. J. Epidemiol.* **183**, 758–764 (2016).

77. J. Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge University Press, 2000).

78. J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Elsevier, 2014).

79. Y. Kim, K. Kim, C. Park, H. Yu, Sequential and Diverse Recommendation with Long Tail. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence* (2019) https:/doi.org/10.24963/ijcai.2019/380.

80. S. Rendle, C. Freudenthaler, Z. Gantner, L. Schmidt-Thieme, BPR: Bayesian Personalized Ranking from Implicit Feedback (2012) (July 16, 2020).

81. Hirano, K., Imbens, G. W. & Ridder, G. Efficient estimation of average treatment effects using the estimated propensity score. Econometrica 71, 1161–1189 (2003).

FIGURES AND TABLES

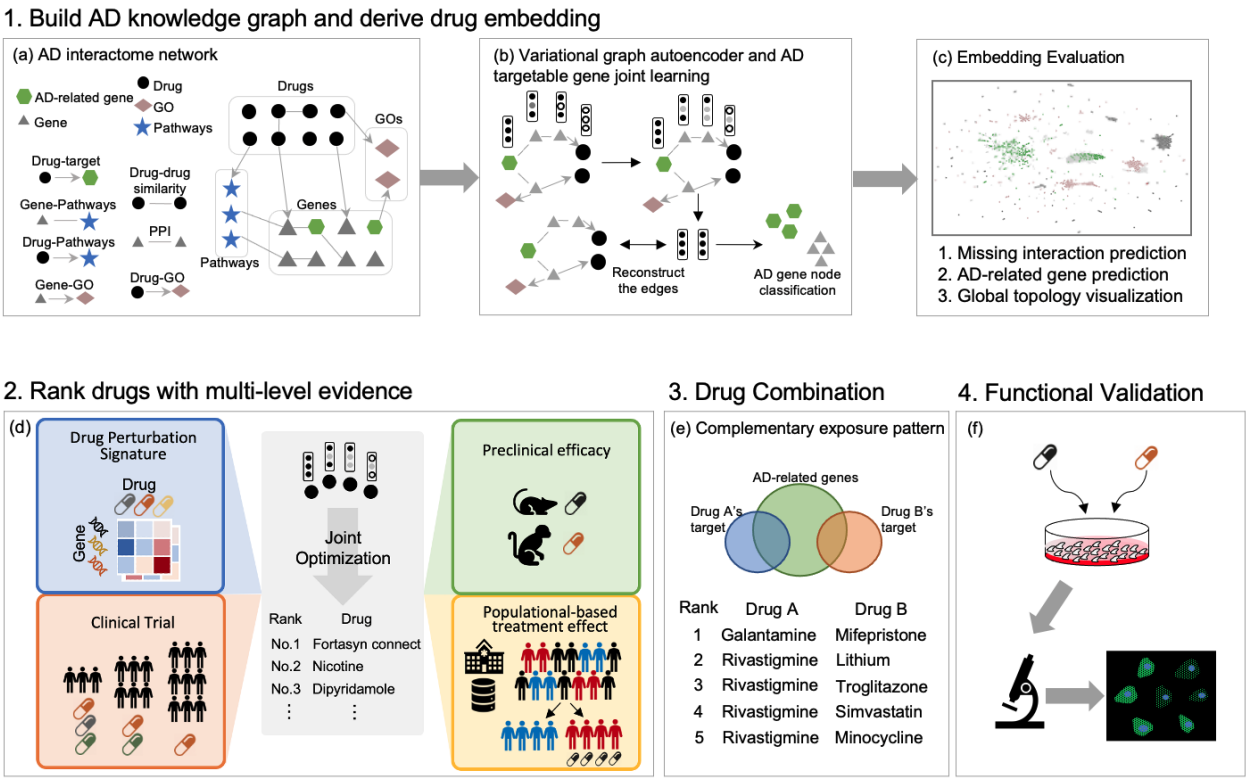

Figure 1. Study workflow. (a) Build the AD knowledge graph with nodes (drugs, genes, pathways, and gene ontology or GO) and edges (drug-target interaction, drug-drug structural similarity, gene-gene interaction, gene-pathway association, gene-GO association, and drug-GO association). (b) Derive the node embeddings using the multi-relational and variational graph autoencoder.(17, 18) (c) Evaluate our derived embeddings' performance on missing interaction prediction, AD-related gene prediction, and global topology visualization. (d) Rank drug candidates using multi-level evidence in drug perturbation signatures, preclinical efficacies, clinical trials, and population-based treatment effects. (e) Search drug combinations satisfying complementary exposure patterns(19) using the high-ranked drug candidates. (f) Validate drug combinations using oxidative stress assays on HT22 Mouse Hippocampal Neuronal Cell Line.

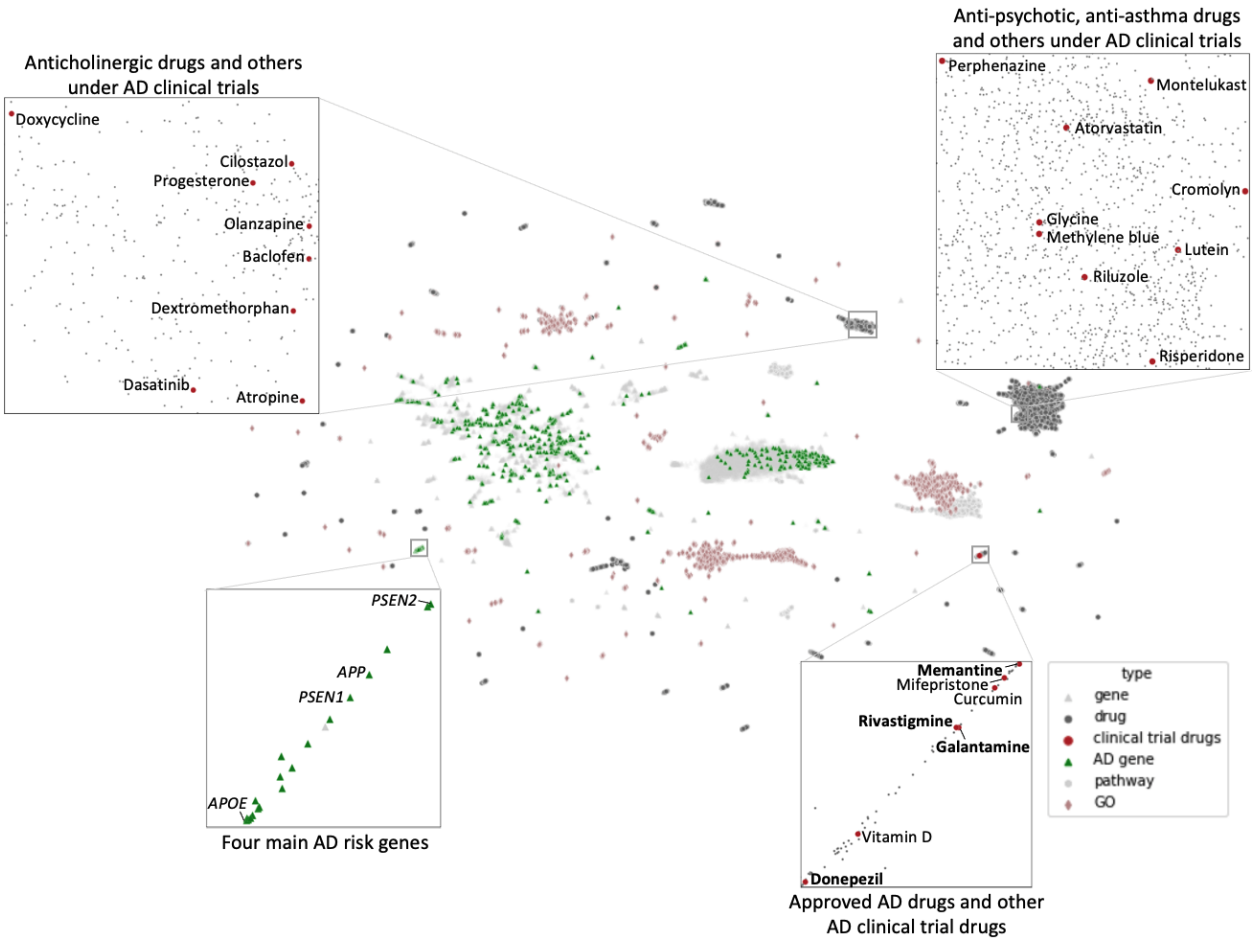

Figure 2. **UMAP plot for lower dimensional projection of node embedding.** Nodes with similar embeddings are adjacent in the UMAP plot. Four current FDA-approved AD drugs were indicated in bold text. Four main AD risk genes (green triangle) and undergoing clinical trial drug (red circle) clusters were highlighted in subplots. Genes, which are in gray triangles, were mixed with drugs, which are in the black rounds. Gene ontologies (pink diamond) and pathways (light grey circle) are adjacent to genes and drugs.

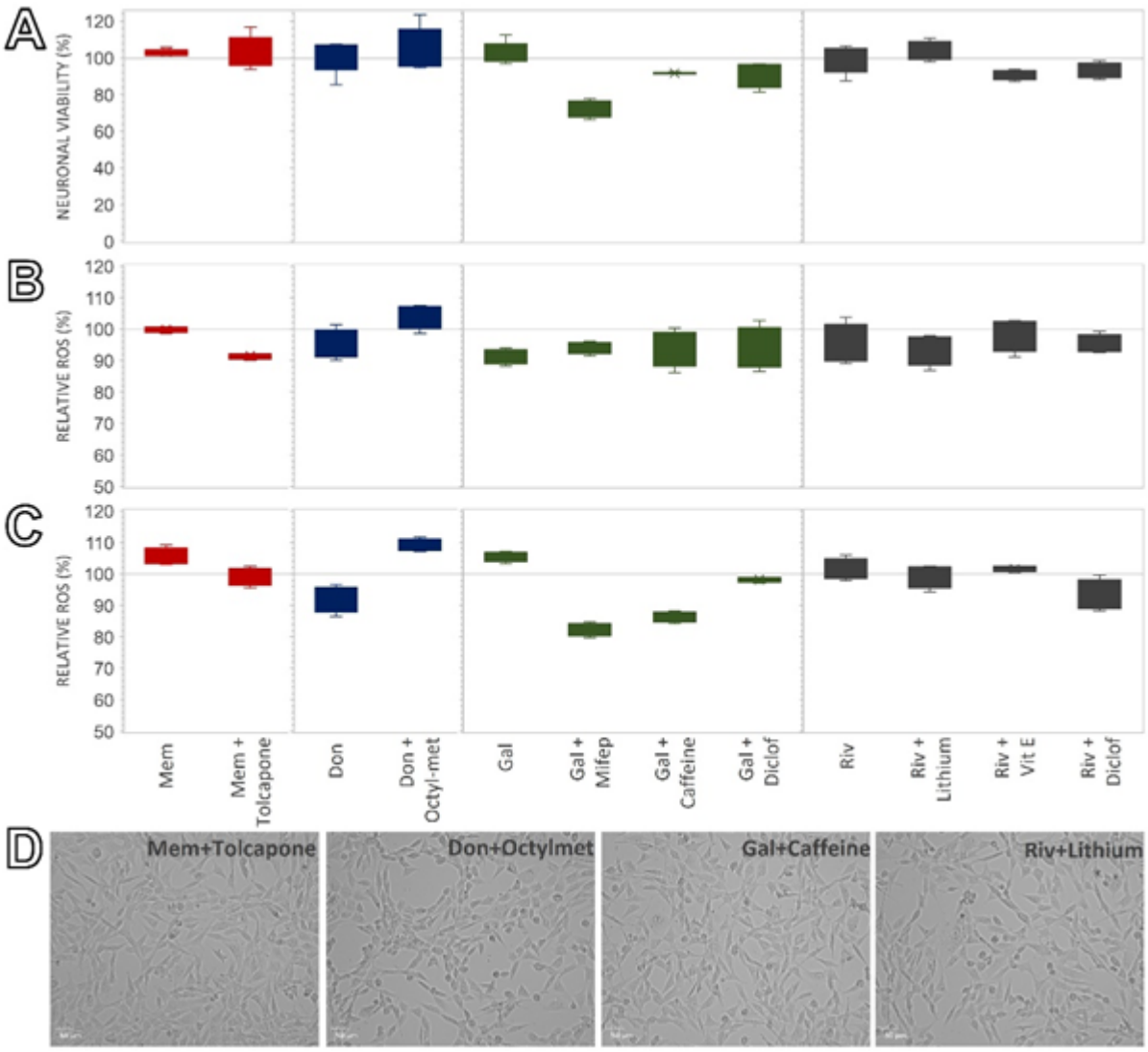

Figure 3. Neuronal responses to drug combinations. (A) Neuronal viability after 24 h with 50 µM of drug combinations. (B) Quantification of intracellular oxidative stress (DCFDA fluorescence) after 24 h with 50µM of drug A + 50µM of drug B. (C) Quantification of intracellular oxidative stress (DCFDA fluorescence) of neurons dosed with 10 µM Aβ, after 24 h with 50µM of drug A + 50µM of drug B. All data sets were normalized to neurons dosed with vehicle (DMSO) and the control represented as a horizontal line at 100% across all the treatments. (D) Live-cell brightfield imaging of neurons dosed for 24 h with 50µM of drug A + 50µM of drug B.

Table 1. Edge prediction accuracy using AD knowledge graph's node embedding. To examine which edge type is difficult/easy to contain in the node embedding, we measured the edge prediction accuracy for each edge type separately. The accuracy values for all edge types were consistent without any edge types being under-fitted. For AD knowledge graph representation without transfer learning, we initialized the node embedding with random initialization.

Which edge type to predict?	Pretrained universal embedding		AD knowledge graph representation without transfer learning		AD knowledge graph representation with transfer learning	
	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
All edge types	0.565	0.648	0.959	0.958	0.991	0.988
Drug - gene	0.528	0.489	0.863	0.795	0.960	0.941
Gene - gene	0.987	0.985	0.936	0.887	0.991	0.984
Drug - GO	0.604	0.554	0.657	0.578	0.889	0.895
Gene - GO	0.613	0.593	0.925	0.912	0.985	0.974
Drug - pathway	0.282	0.376	0.989	0.988	0.998	0.995
Drug - drug structural similarity	0.897	0.867	0.981	0.985	0.994	0.990
Gene - pathway	0.397	0.459	0.590	0.639	0.956	0.947

Table 2. Selected promising drugs with supporting evidence and literature. +: positive evidence, blank: not investigated, NA: not available.

Name	Pharmacological classes	Primary indications	Main targets	Multi-level evidence				Reference
				Clinical trials	Transcriptomic relevance	Population-based treatment effect	Preclinical trial	
Galantamine	Acetylcholinesterase inhibitor	Mild to moderate AD	ACHE, CHRNA7, BCHE	+	+		+	21
Melatonin	Acetamides	Insomnia	MTNR1A, MTNR1B, ESR1	+	+		+	22
Celecoxib	NSAID	Osteoarthritis, Rheumatoid arthritis	PTGS2, PDPK1, CDH11	+	+		+	23
Mifepristone	Hormone modulators	Early pregnancy termination	PGR, NR3C1, KLK3	+				24
Trazodone	Antidepressants	Major depressive disorder	HTR2A, HTR2C, SLC6A4	+		+		25
Ibuprofen	NSAID	Pain, Osteoarthritis, Rheumatoid arthritis	COX-1, COX-2, PTGS2	+		+	+	26
Avagacestat (BMS-708163)	Anti-amyloid	NA	γ-secretase	+				27, 28
Rosiglitazone	Thiazolidinedione	Type 2 diabetes	PPARG, ACSL4, PPARA	+	+		+	29
Amphotericin B	Anti-infectives	Fungal infections	Ergosterol					30
Physostigmine	Anticholinergic	Glaucoma	ACHE, CHRNA4, CHRNB2				+	31