Modeling Path Importance for Effective Alzheimer's Disease Drug Repurposing

Shunian Xiang[1*], Patrick J. Lawrence[1*], Bo Peng[2], ChienWei Chiang[1], PhD, Dokyoon Kim[3], PhD, Li Shen[3], PhD, and Xia Ning[1,2,4†], PhD

[1]*Biomedical Informatics Department, The Ohio State University, Columbus, OH 43210, USA*
[2]*Computer Science and Engineering Department, The Ohio State University, Columbus, OH 43210, USA*
[3]*Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA 19104 USA*
[4]*Translational Data Analytics Institute, The Ohio State University, Columbus, OH 43210, USA*
Co-first author; authors contributed equally to this work
[†]*E-mail: ning.104@osu.edu*

Recently, drug repurposing has emerged as an effective and resource-efficient paradigm for `AD` drug discovery. Among various methods for drug repurposing, network-based methods have shown promising results as they are capable of leveraging complex networks that integrate multiple interaction types, such as protein-protein interactions, to more effectively identify candidate drugs. However, existing approaches typically assume paths of the same length in the network have equal importance in identifying the therapeutic effect of drugs. Other domains have found that same length paths do not necessarily have the same importance. Thus, relying on this assumption may be deleterious to drug repurposing attempts. In this work, we propose `MPI` (Modeling Path Importance), a novel network-based method for `AD` drug repurposing. `MPI` is unique in that it prioritizes important paths via learned node embeddings, which can effectively capture a network's rich structural information. Thus, leveraging learned embeddings allows `MPI` to effectively differentiate the importance among paths. We evaluate `MPI` against a commonly used baseline method that identifies anti-`AD` drug candidates primarily based on the shortest paths between drugs and `AD` in the network. We observe that among the top-50 ranked drugs, `MPI` prioritizes 20.0% more drugs with anti-`AD` evidence compared to the baseline. Finally, Cox proportional-hazard models produced from insurance claims data aid us in identifying the use of etodolac, nicotine, and `BBB`-crossing `ACE-INHs` as having a reduced risk of `AD`, suggesting such drugs may be viable candidates for repurposing and should be explored further in future studies.

Keywords: Alzheimer's Disease; Drug Repurposing; Machine Learning.

1. Introduction

Alzheimer's Disease, denoted `AD`, is a progressive neurodegenerative disorder that accounts for 60%-70% of dementia cases and affects more than 50 million people worldwide today.[1,2] Given the large number of affected individuals and `AD`'s life-threatening nature,[3] extensive resources have been dedicated to developing `AD`-modifying drugs. Since 2003, inefficacy or

toxicity has accounted for a 95+% failure rate among candidates evaluated for `AD` treatment.[4,5] Furthermore, none of the current US Food and Drug Administration (FDA)-approved `AD` drugs are curative; they only slow disease progression. Because of the immense resources required to conduct clinical trials,[6] the numerous failed clinical trials have necessitated the development of a more resource-efficient method for `AD` drug discovery. In the last decade, the identification of new therapeutic indications for existing FDA-approved drugs, referred to as drug repurposing,[7] has emerged as an effective and resource-efficient paradigm for drug discovery.[8] This is an attractive option as the toxicity, pharmacokinetics, and pharmacodynamics of FDA-approved drugs have already been thoroughly investigated by previous clinical trials.[7,9]

Recently, the curation of comprehensive drug databases has enabled the development of computational methods for `AD` drug repurposing.[10–14] Among all the methods, network-based methods have shown promising results and emerged as a popular approach.[13,15,16] Network-based methods utilize comprehensive protein-protein, drug-target, and `AD`-protein interactions to effectively reveal potential therapeutic effects of drugs on `AD`. Though promising, existing methods[13] measure the therapeutic effects of drugs on `AD` primarily using count and length of the paths connecting drug nodes and the `AD` node in the network. Paths of the same length are considered equivalently effective at identifying the therapeutic effect of drugs by these methods. However, in other domains, paths of the same length have been shown to exhibit substantially different levels of importance.[17,18] As such, assuming equal length paths have equal importance could be detrimental to effective drug repurposing for `AD`.

In this work, we propose a novel method to conduct drug repurposing for `AD`, `MPI` (Modeling Path Importance), to address this limitation. Similar to existing methods,[13,14] `MPI` leverages the interactions between drugs and `AD` via proteins as indications of the potential therapeutic effects of drugs on `AD`. Based on the interactions, `MPI` introduces a scoring function to score and rank drugs for their anti-`AD` effectiveness. `MPI` is unique in that it learns node embeddings[19] and prioritizes important paths via these learned embeddings. Recent work[20] has shown that the learned node embeddings can effectively capture the rich structure information within a network. Thus, scoring paths using node embeddings allows `MPI` to utilize the network structure information to better prioritize paths for effective `AD` drug repurposing. Specifically, in this study, `MPI` leverages `DeepWalk`,[21] a widely used network learning approach, to generate node embeddings. Edges are scored using a normalized dot product between the learned node embeddings; paths and drugs are scored by multiplying individual edge scores. Note that because `MPI` serves as a general framework, other network learning approaches, such as `Node2Vec`[20] and graph neural networks,[22] could also be easily incorporated to generate node embeddings.

In this study, we construct a network to conduct drug repurposing for `AD` by combining protein-protein interactions (PPIs), drug-target interactions (DTIs), and `AD`-protein interactions (APIs) from multiple data sources. To investigate the effectiveness of `MPI`, we compare `MPI` against a commonly utilized network-based drug repurposing method for `AD`,[13,23] denoted as `BSL`, using our network. Our experimental results demonstrate that among the top-50 ranked drugs, `MPI` prioritizes 20% more drugs with anti-`AD` evidence compared to `BSL`. We examine published literature and analyze insurance claims meta data to evaluate the evidence of anti-`AD`

activity among `MPI`'s top prioritized candidates. The results of our evaluation find consensus between published experimental results and our own analysis for a few drug candidates. Notably, angiotensin converting enzyme inhibitors (`ACE-INHs`) represent a class of drugs that should be further explored for their anti-AD properties. Moreover, other drugs, such as nicotine, that enhance the brains response to acetylcholine and reduce cholinergic atrophy should be examined as well. Conversely, we find that, relative to other evaluated drugs, long-term use of trihexyphenidyl increases the risk of `AD`. This was corroborated by previously published *in vivo* experiments.[24] Finally, we find etodolac to confer the lowest risk of developing `AD` among all cyclooxygenase inhibitors (`COX-INHs`) in our network. Altogether, these findings suggest that `MPI` may be a viable option with respect to identifying repurposing candidates to treat `AD`.

2. Materials and Methods

2.1. *Network construction*

PPIs, DTIs and APIs have shown utility for `AD` drug repurposing.[13] As such, we construct our network using these interactions. Below, we describe our process for compiling the PPIs, DTIs and APIs used to construct our network from public data sources. In total, our network has 327,924, 2,854, and 230 edges corresponding to PPIs, DTIs, and APIs. These edges connect one `AD` node, 18,527 protein nodes, and 386 drug nodes.

2.1.1. *Protein-protein interactions (PPIs)*

Following Chen et al.,[13] we include a comprehensive list of human PPIs consisting of 327,924 interactions. This list aggregates a total of 21 bioinformatics and systems biology databases with combinations of five types of experimental evidence. We refer the audience of interest to Chen et al.[13] for a detailed description of the databases.

2.1.2. *Drug-target interactions (DTIs)*

We assemble drug-target interactions and bioactivity data from 4 commonly used databases (each downloaded in November 2022): the ChEMBL database[25] (v31), the binding database,[26] the therapeutic target database,[27] and the IUPHAR/BPS guide to pharmacology database.[28] We retain the drug-target interactions that satisfy the following inclusion criteria: 1) binding affinities, including K_i, K_d, IC_{50}, or EC_{50}, must be less than or equal to 10 μM; 2) protein targets and their respective proteins must have a unique UniProt[29] accession number; 3) protein targets must be marked as reviewed in the UniProt database; 4) protein targets must be present in *homo sapiens*.

Additionally, we retain drugs for which we have sufficient sample size to conduct quantitative analysis using MarketScan[30] insurance claims meta data (see Section 2.4). Specifically, included drugs have at least 100 patients with their first dose at least 2 years prior to an `AD` diagnosis (dx). Additionally, these drugs must have at least 15 patients who eventually received an `AD` dx. Applying these filters yielded 2,854 edges connecting 386 FDA-approved drugs to 548 protein targets.

2.1.3. `AD`-*protein interactions (APIs)*

The `AD`-associated proteins included in the network were identified from multiple sources. 54 β-amyloid-related proteins and 27 tauopathy-related proteins were obtained from Cheng et al..[13] The authors identified proteins that satisfied at least one of the following criteria: 1) the proteins are validated in large-scale amyloid or tauopathy genome-wide association studies; 2) *in vivo* experimental models exhibit evidence that knockdown or overexpression of the protein leads to `AD`-like amyloid or tau pathology. We also include 93 unique late-onset `AD` common risk proteins identified by 7 large-scale genetic studies .[31–37] We further incorporate a set of 118 `AD`-associated proteins introduced in at least 2 out of the 6 following databases (each was downloaded in November 2022): the online Mendelian inheritance in man database,[38] the comparative toxicogenomics database,[39] the HuGE navigator database,[40] the DisGeNET database[41] (v7.0), the ClinVar database[42] and the Open Targets database[43] (v22.09). In total, our network is comprised of 230 unique, `AD`-associated proteins. Each of the `AD`-associated proteins are connected to a single `AD` node with each edge between a protein and the `AD` node representing an API in our network.

2.2. *Modeling path importance for* `AD` *drug repurposing*

In this work, we denote the constructed network as G. Each node in G is denoted as v_i. Specifically, drug nodes, protein nodes and the `AD` node are v_i^d, v_i^g, and v_i^a , respectively. Note that the index, i, does not apply to the `AD` node as there are not multiple in our network. Each edge that connects node v_i to node v_j is denoted as e_{ij}. Each path is denoted as p_m, and the set of edges involved in a path is denoted $\mathbb{E}_{p_m}$. Below, we denote matrices, scalars and row vectors using uppercase, lowercase, and bold lowercase letters, respectively.

In `MPI`, we leveraged `DeepWalk`,[21] a widely used node embedding approach, to learn embeddings for each node in G. First, for each node v_i in the network, we conduct 256 random walks originating from this node, and terminating once the path length reaches 128. `DeepWalk` is then trained by sliding a window of length 10 over the generated paths. Nodes within the same window are forced to have similar embeddings following the objective function defined in the original paper.[21] Node embeddings for `MPI` are produced such that they have 128 dimensions.

After generating node embeddings, we score edge, e_{ij}, using a normalized dot product of the embedding of v_i^x ($x=d$, g or a) and v_j^y ($y=d$, g or a) as follows:

$$w_{ij} = \frac{\exp(\mathbf{v}_i^x \mathbf{v}_j^{y\top})}{\sum_{k \in \mathbb{V}} \exp(\mathbf{v}_i^x \mathbf{v}_k^{y\top})}, \tag{1}$$

where w_{ij} is the score of the edge e_{ij}; $\mathbf{v}_i^x$ and $\mathbf{v}_j^y$ is the learned embedding of node v_i^x and v_j^y, respectively; $\exp(\cdot)$ is the exponential function; and $\mathbb{V}$ is the set of all the nodes in the network. Note that, in Equation 1, only one of v_i^x and v_j^y could be the `AD` node. These edge scores are calculated with node embeddings which implicitly capture the rich structural information within the network. Thus, compared to existing methods, `MPI` can better leverage a network's structural information for `AD` drug repurposing. We calculate the score for each

path by multiplying the scores of its individual edges as follows:

$$s_{p_m} = \prod_{e_{ij} \in \mathbb{E}_{p_m}} w_{ij}, \tag{2}$$

where s_{p_m} is the score of the path p_m; and $\mathbb{E}_p$ is the set of all the edges in the path p. The score for each drug (i.e., $s_{v_i^d}$) is then defined as the summation of the scores from all 3-hop or shorter paths that originate from the `AD` node and terminate at the drug node.

2.3. Baseline method

To evaluate the performance of `MPI`, we compare `MPI` against a network-based method recently developed by Cheng et al.,[13] denoted as `BSL`. `BSL` scores drugs based on the shortest distance between the drug targets and the `AD`-associated proteins (Section 2.1.3) in the network. Specifically, we denote $\mathbb{T}(i)$ as the set of protein targets associated with a given drug v_i^d, and denote $\mathbb{P}$ as the set of `AD`-associated proteins. The proximity between these two sets is calculated as the average shortest distance between elements in $\mathbb{T}(i)$ and $\mathbb{P}$ as follows:

$$r(\mathbb{T}(i), \mathbb{P}) = \frac{1}{|\mathbb{T}(i)| + |\mathbb{P}|} \left(\sum_{v_j \in \mathbb{T}(i)} \min_{v_k \in \mathbb{P}} d(v_j, v_k) + \sum_{v_k \in \mathbb{P}} \min_{v_j \in \mathbb{T}(i)} d(v_k, v_j) \right), \tag{3}$$

where $r(\mathbb{T}(i), \mathbb{P})$ is the proximity between these two sets; $|\mathbb{T}(i)|$ and $|\mathbb{P}|$ is the size of $\mathbb{T}(i)$ and $\mathbb{P}$, respectively; and $\min_{v_k \in \mathbb{P}} d(v_j, v_k)$ is the shortest distance between v_j and any elements in $\mathbb{P}$. Subsequently, we conduct a permutation test to assess the statistical significance of the calculated proximity. The resulting z-score from this test is used as the score of drug v_i.[13] In `BSL`, a lower drug score implies a higher potential for effective `AD` treatment.

2.4. Validation using MarketScan database

We use MarketScan medicare supplemental database from 2012–2021 to evaluate drug impact on `AD` onset via Cox proportional-hazard models.[30] The MarketScan database includes data for over 8 million unique individuals and is comprised of demographic information, administrative information, diagnoses, procedures, and pharmacy records. International Classification of Disease (ICD)-9/ICD-10 codes denote diagnoses and National Drug Codes (NDCs) record pharmacy claims. We use the ICD-9/ICD-10 codes listed in Supplementary Table S4♠ to define `AD` and comorbidities, which are included as covariates in Cox proportional-hazard models. We conduct our analysis over 1,632,218 unique individuals who were at least 65 years by 2022 and possessed a minimum of five years insurance enrollment prior of first `AD` diagnosis. Drugs from our constructed network are mapped to NDC codes by partial matching of generic names from MartketScan redbook. We only include patients who took or started taking a drug at least two years prior to `AD` diagnosis to mitigate the possibility that patients starting a drug already had `AD` given that `AD` is difficult to diagnosis.
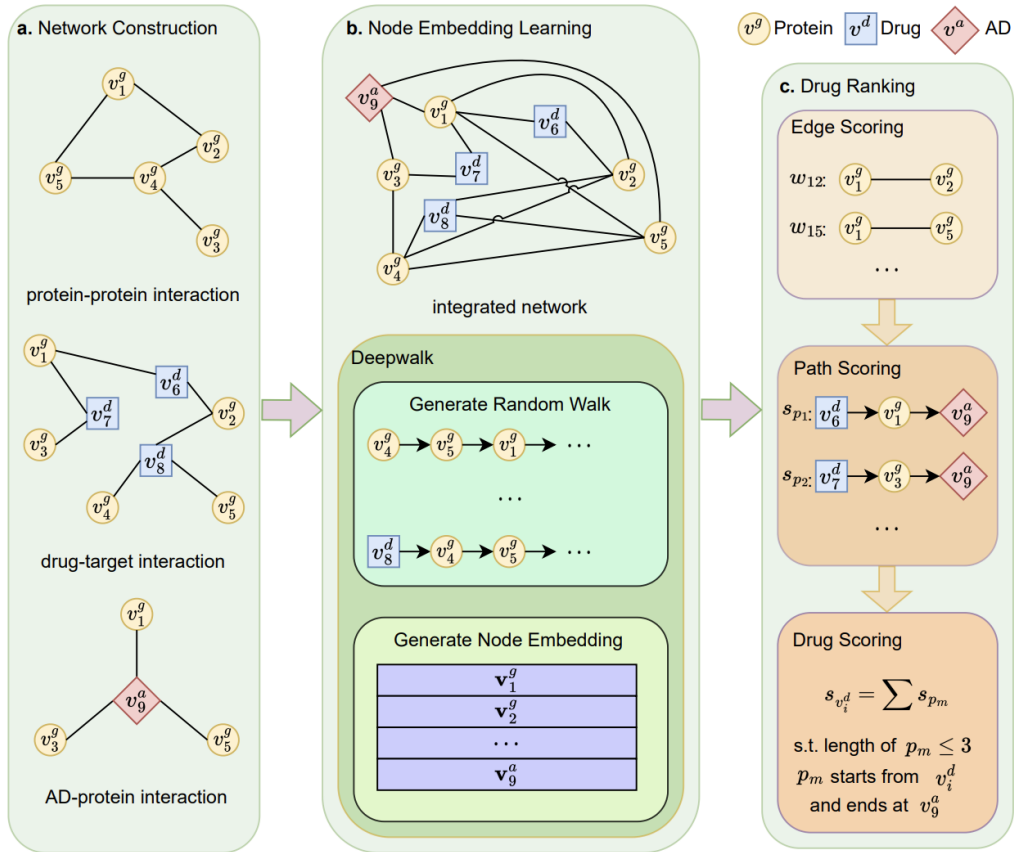
Fig. 1: Figure 1a shows the network construction process in `MPI`. Figure 1b shows the `DeepWalk`-based node embedding generation in `MPI`. Figure 1c shows the edge, path and drug scoring in `MPI`.

3. Results

3.1. `MPI` *for* `AD` *drug repurposing*

In this study, we curate a network consisting PPIs, DTIs and AGIs and propose a novel network-based method, `MPI`, for `AD` drug repurposing. We propose `MPI` with the following intuitions: 1) proteins that associated with `AD` are localized in the corresponding disease module within the comprehensive human PPI network; 2) the drug target(s) for a disease may also be targeted for other diseases (e.g., `AD`) owing to common functional targets and pathways elucidated by PPIs; 3) if a drug node is linked to the `AD` node through the paths of drug targets and `AD`-associated proteins in the PPI, the drug may have a treatment effect on `AD`.

We implement `MPI` using the following steps: 1) integrate `AD`-protein interactions, drug-target interactions and protein-protein interactions to generate a comprehensive network (Figure 1a), 2) employ `DeepWalk` to learn node embeddings which capture the structural information within the network (Figure 1b), and 3) score edges, paths and drugs based on the learned

---

♠Supplementary material and code can be found here: `https://github.com/ninglab/MPI`

embeddings to leverage the structural information for better `AD` drug repurposing (Figure 1c). Then we identify plausible treatment candidates from the top-ranked drugs using a literature search of the published evidence. We collected 327,924 PPIs from 21 bioinformatics and systems biology databases (Section 2.1.1). We also collected 2,854 DTIs from 4 commonly used databases (Section 2.1.2), and 230 comprehensive APIs from multiple resources (Section 2.1.3). By aggregating all the interactions, we construct a drug-protein-`AD` network comprised of 386 drug nodes, 18,527 protein nodes, 1 `AD` node, and 331,008 edges. More details about the network construction are available in Section 2. To the best of our knowledge, `MPI` is the first method which effectively repurposes drug candidates for `AD` treatment by prioritizing paths between drug nodes and the `AD` node using learned node embeddings.

3.2. *Comparing anti-`AD` evidence of `MPI`'s and `BSL`'s top-50 drugs*

We compare the top-50 drugs prioritized by `MPI` and `BSL` to evaluate their capacity for repurposing drugs to treat `AD`. Specifically, we score and rank all 386 drug nodes in our network using `MPI` and `BSL`. The complete rankings are reported in Supplementary Table S3. We then perform a literature search to evaluate the anti-`AD` evidence of the top-50 ranked drugs for both `MPI` and `BSL`. We define anti-`AD` evidence as any published experimental result(s), which demonstrate a drug either protects against the development of `AD` or ameliorates aberrant cellular phenotypes caused by `AD`. We present `MPI`'s and `BSL`'s top-10 drugs and their anti-`AD` evidence in Table 1 and Table 2, respectively. The complete rankings for the top-50 drugs and their anti-`AD` evidence is available in Supplementary Tables S1 and S2. Based on the significance of the anti-`AD` evidence, we categorized drugs into the following 6 types in decreasing order of significance: 1) drugs which are FDA-approved for `AD` treatment (approved); 2) drugs that have demonstrated anti-`AD` effects in completed clinical trials or are under investigation in `AD` clinical trials (clinical); 3) drugs which have demonstrated anti-`AD` effects in *in vivo* experiments (*in vivo*); 4) drugs which have
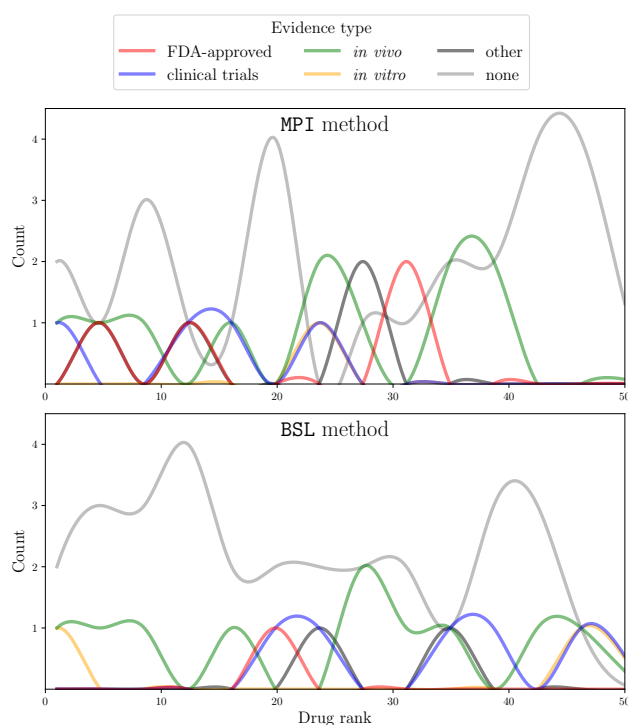
(a) Distribution of drug rank by evidence type for `BSL` and `MPI` methods.
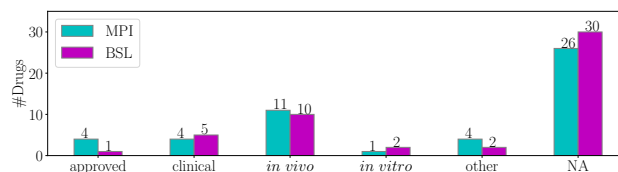
(b) Drug counts for each type of evidence among `MPI`'s and `BSL`'s top-50 drugs.

Fig. 2: Evaluation of drug rank distributions: `MPI` and `BSL`.

demonstrated anti-`AD` effects in *in vitro* experiments (*in vitro*); 5) drugs which show anti-`AD` effects in observational studies, cohort studies or analyses in insurance data (other); 6) drugs that either do not have the above 5 types of evidence or have been demonstrated ineffective or damaging for `AD` (NA). We present the distribution of the top-50 drugs from `MPI` and `BSL` over the different types of evidence in Figure 2a and the counts of each evidence type in Figure 2b. In Figure 2a, we observe more drugs with evidence ranked highly by `MPI` compared to `BSL`. This is supported by Figure 2b, which confirms that `MPI` identified more evidential anti-`AD` drugs compared to `BSL` in the top-50 ranked drugs. Specifically, among the top-50 ranked drugs, `MPI` prioritized 24 evidential anti-`AD` drugs while `BSL` only prioritized 20 evidential anti-`AD` drugs, demonstrating an improvement of 20%. Figures 2a and 2b also show `MPI` outperforms `BSL` in prioritizing drugs with significant evidence. `MPI` prioritizes all the 4 FDA-approved anti-`AD` drugs (e.g., galantamine, rivastigmine, donepezil and memantine) in our network among the top-50. In contrast, `BSL` prioritizes only a single FDA-approved anti-`AD` drug (donepezil) among the top-50.

We also observe in Table 1 and Table 2 that `MPI` is more effective than `BSL` at prioritizing anti-`AD` drugs among the very top (top-10) of the ranking list. That is, among the top-10 drugs, 6 drugs from `MPI` have anti-`AD` evidence including the FDA-approved `AD` drug galantamine, while only 4 drugs from `BSL` are evidential. As presented in Section 2, compared to `BSL`, `MPI` learns node embeddings to capture the rich structural information within the network, and leverage the structural information to better identify anti-`AD` drugs. The superior performance of `MPI` over `BSL` demonstrates the effectiveness of leveraging the network structural information to conduct repurposing to identify candidates for `AD` treatment. We also notice that both `MPI` and `BSL` prioritize 17 drugs in concordance within their top-50 drug lists. Among the 17 drugs, 5 drugs demonstrate anti-`AD` evidence: donepezil is an FDA-approved anti-`AD` drug; nicotine and rasagiline have clinical anti-`AD` evidence; and fluvoxamine and fluoxetine have *in vivo* anti-`AD` evidence. The drugs nicotine, rasagiline, fluvoxamine, and fluoxetine could be promising repurposing candidates. We leave the investigation of these drugs to future research.

3.3. *Identifying repurposing candidates with anti-`AD` activity*

In order to identify plausible candidates for repurposing, we produce Cox proportional-hazard models (see Section 2.4) to ascertain whether there is consensus between the MarketScan insurance data and the `AD`-related evidence we found for top ranked candidates prioritized by `MPI`. Specifically, we use hazard ratios (`HR`) to identify whether any evidential drug elicited reduced the risk of `AD` diagnosis among patients who took the drug compared against those that did not. We present each drug's `HR` with their significance levels in Supplementary Table S5b; the `HR` for each drug's covariates (sex, age, and additional common comorbidities) are reported in Supplementary Table S5c-t. A `HR` below 1 indicates that a drug has a protective effect, while a `HR` above 1 indicates that a drug has a damaging effect. Figure 3a plots Kaplan-Meier (`KM`) survival curves. These plots depict a patient's likelihood of being diagnosed with `AD` following long-term use of either an individual prescribed drug or a drug with a given mechanism of action (MOA). For MOAs, we group highly-prioritized drugs with published

Table 1: Top-10 Drugs from `MPI`

Drug	MOA	Indication	Anti-`AD`	Evidence
varenicline	`AChR-Ag`	smoking cessation	N	-
fosinopril	`ACE-INH`	hypertension	Y	*in vivo*[44]
nicotine	`AChR-Ag`	smoking cessation	Y	clinical[45]
nizatidine	histamine receptor antagonist	duodenal ulcer disease	N	-
piroxicam	`COX-INH`	osteoarthritis	Y	other[46,47]
meloxicam	`COX-INH`	osteoarthritis	Y	*in vivo*[48–50]
galantamine	`AChE-INH`	Alzheimer's disease	Y	approved
bromfenac	`COX-INH`	inflammation	N	-
etodolac	`COX-INH`	osteoarthritis	Y	*in vivo*[51]
pyridostigmine	`AChE-INH`	myasthenia gravis	N	-

In this table, the column "Drug" shows the identified top-10 ranked drugs; the column "MOA" shows the mechanism of action of each drug; the column "Indication" presents the indication of each drug; the column "Anti-`AD`" indicates if the drug has evidenced anti-`AD` effects; and the column "Evidence" presents the type of the evidence. In this table, `ACE-INH` represents the angiotensin converting enzyme inhibitor; `COX-INH` represents the cyclooxygenase inhibitor; `AChE-INH` represents the acetylcholinesterase inhibitor; and `AChR-Ag` represents the acetylcholine receptor agonist.

Table 2: Top-10 Drugs from `BSL`

Drug	MOA	Indication	Anti-`AD`	Evidence
tetracycline	bacterial 30S ribosomal subunit inhibitor	respiratory tract infections	Y	*in vitro*[52]
selegiline	monoamine oxidase inhibitor	Parkinson's Disease	N	-
ceftriaxone	bacterial cell wall synthesis inhibitor	gonorrhea	Y	*in vivo*[53]
ibuprofen	`COX-INH`	headache	N	-
levobunolol	adrenergic receptor antagonist	glaucoma	N	-
ketoprofen	`COX-INH`	rheumatoid arthritis	N	-
carbidopa	aromatic L-amino acid decarboxylase inhibitor	Parkinson's Disease	N	-
sulindac	`COX-INH`	osteoarthritis	Y	*in vivo*[54]
biotin	vitamin B	supplement	Y	*in vivo*[55]
lansoprazole	ATPase inhibitor	heartburn	N	-

In this table, the column "Drug" shows the identified top-10 ranked drugs; the column "MOA" shows the mechanism of action of each drug; the column "Indication" presents the indication of each drug; the column "Anti-`AD`" indicates if the drug has evidenced anti-`AD` effects; and the column "Evidence" presents the type of the evidence. In this table, `ACE-INH` represents the angiotensin converting enzyme inhibitor; `COX-INH` represents the cyclooxygenase inhibitor; `AChE-INH` represents the acetylcholinesterase inhibitor; and `AChR-Ag` represents the acetylcholine receptor agonist.

evidence of anti-`AD` activity (see Table 1). Bupropion (`HR` = 1.04; non-significant) was included as a negative control as clinical trials found the drug had no significant effect on cognition in AD patients.[56] Trihexyphenydil (`HR` = 1.71; $\alpha < 0.001$) was included as a positive control for

damaging effects due to the evidence documented in Supplementary Table S1. The `COX-INHs` group includes the following drugs: piroxicam, meloxicam, etodolac, and flurbiprofen. The `ACE-INHs` group includes the following drugs: fosinopril, trandolapril, and lisinopril. Note that we only include blood brain barrier (`BBB`) crossing `ACE-INHs` in this group as non-`BBB`-crossing `ACE-INHs` have exhibited very limited effects on `AD`.[57] We also include time-to-event analysis for 4 of `BSL`'s top prioritized drugs (See Supplementary Figure S1). Unlike `MPI`, we observe only one of `BSL`'s drugs (sulindac) with reduced time-to-event compared to bupropion; however, this difference is not significant.
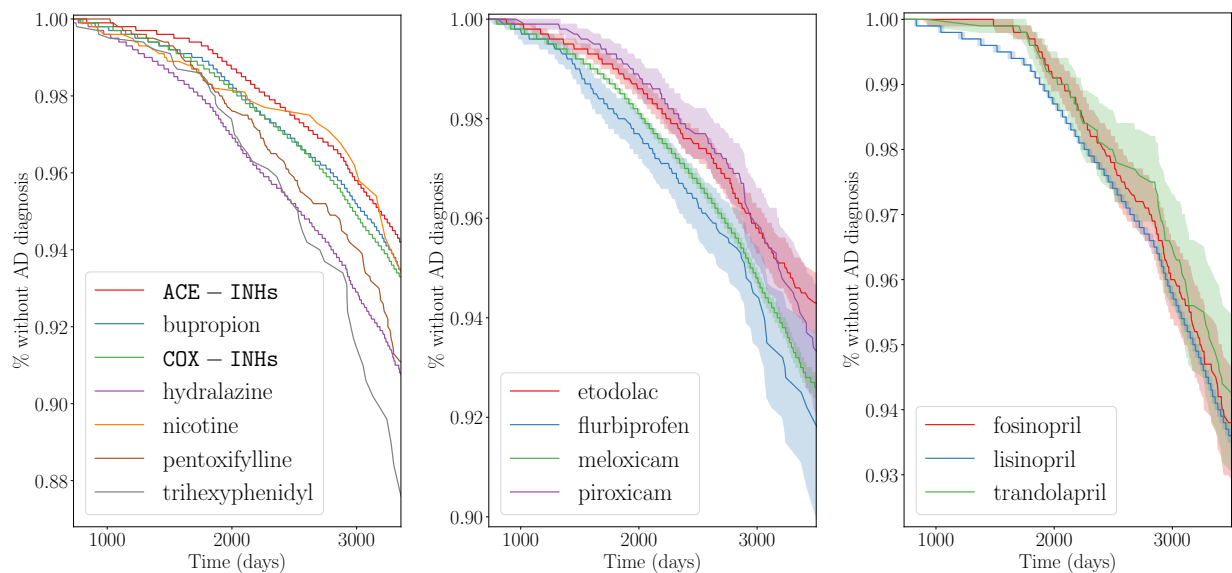
(a) Drugs and MOAs with published anti-`AD` evidence

(b) `COX-INHs`. Shaded regions represent 95% confidence intervals.

(c) `ACE-INHs`. Shaded regions represent 95% confidence intervals.

Fig. 3: Unadjusted Kaplan−Meier plots for cox proportional-hazard models

3.4. *Analyzing the MOAs of* `MPI`*'s top-50 drugs*

To identify groups of drugs whose anti-`AD` properties should be further examined and explored, we examine the top-50 drugs prioritized by `MPI` for any common MOAs. We find that `COX-INHs` and `ACE-INHs` are the most common MOAs prioritized by `MPI`. Both `COX-INHs` and `ACE-INHs` have published evidence of anti-`AD` activity. That said, experimental results suggest that long-term administration of `COX-INHs` may only have protective properties, reducing the risk of `AD` onset.[46] Moreover, meloxicam ($HR = 0.86$; $\alpha < 0.05$), has even shown therapeutic potential, reversing cognitive decline via inhibition of neuronal apoptosis.[48,49] However, in Figure 3a, we observe that `COX-INHs` as a class do not yield reduced risk of `AD` compared to the negative control. That said, we find etodolac significantly reduces the risk of `AD` ($HR = 0.78$; $\alpha < 0.001$) compared to other `COX-INHs`, including flurbiprofen ($HR = 0.95$; non-significant) (Figure 3b). This suggests that only certain `COX-INHs`, such as etodolac, may elicit protective effects against `AD` onset. Importantly, this may be a result of differences in target as etodolac targets *COX2*, while flurbiprofen targets *COX1*. On the other hand, `ACE-INHs` were found to also protect

against `AD` onset in Figure 3a. Specifically, we evaluate only `ACE-INHs` that cross the blood brain barrier (`BBB`) as previous insurance claims metadata analyses have indicated those that do not cross the `BBB` have no effect on `AD`.[57] To see if any of the `BBB` crossing `ACE-INHs` have a greater protective effect that others, we produce a `KM` plot for fosinopril, lisinopril, and trandolapril (Figure 3c). Unlike for `COX-INHs`, `ACE-INHs` do not elicit any significant by-drug difference in `AD` onset as illustrated in Figure 3c. While `MPI` prioritized four `BBB` crossing (`BBBx`) and four non-`BBBx` `ACE-INHs` in the top-50, the `BBBx` `ACE-INHs` had a lower average rank compared to the non-`BBBx` `ACE-INHs` (15 and 19, respectively).

Another important distinction between `COX-INHs` and `ACE-INHs` is that `ACE-INHs` have been shown to have some ameliorative potential; whereas, `COX-INHs` have only shown protective effects. In fact, fosninpril and lisinopril (ranked 2nd and 24th by `MPI`, respectively) was found to reduce cognitive decline in animal models of `AD`.[44,58] In Figure 3a, we find that `BBBx` `ACE-INHs` consistently exhibit decreased risk of `AD` relative to our negative control drug, bupropion. Additionally, there does not appear to be a significant difference between any of the `BBBx` `ACE-INHs` with respect to their protection against `AD`, indicating that they are possibly all viable candidates for repurposing. This is in agreement with other published evidence that has identified `BBBx` `ACE-INHs` as having protective effect on `AD` development. Interestingly, `MPI` prioritized 133.6% more `COX-INHs` and 700.9% more `ACE-INHs` than `BSL` in the top-50 from all such drugs in our network. `MPI`'s ability to prioritize more drugs from MOAs with known anti-AD activity suggests that it may be a more viable option when identifying candidates for drug repurposing.

`MPI` also highly prioritizes drugs that increase the brain's response to acetylcholine, either by reducing its degradation (acetylcholinesterase inhibitors, `AChE-INHs`) or by stimulating its receptors (acetylcholine receptor agonists, `AChR-Ags`). This is important as acetylcholine's (`ACh`) synaptic bioavailability is an important contributor to `AD` progression. That is, there is evidence that cholenergic atrophy and `ACh` deficiency is linked with cognitive decline in `AD` patients.[59] Moveover, many of the current FDA-approved drugs indicated to slow `AD` progression target this mechanism of disease progression via `AChE-INHs` (e.g., donepezil, rivastigmine, and galantamine). `AChR-Ags`, also enhances `ACh` signaling. Such drugs, such as nicotine, accomplish this by increasing the response of `ACh` receptors located on the post-synaptic neuron. Interestingly, nicotine, was found to significantly improve cognition in patients with mild cognitive impairment, which is a precursor to `AD`.[45] We also find long-term nicotine use to have a protective effect (`HR` $= 0.532$; $\alpha < 0.001$), with respect to `AD` onset. In Figure 3a, we observe similar risk of developing `AD` to `ACE-INHs` after six to seven years. Conversely, we find evidence that long-term use of trihexyphenidyl, which reduces the activity of `ACh` receptors, is associated with `AD`-like neurodegeneration in rats.[24] This is corroborated by Figure 3a, where we observe the highest risk of `AD` elicited by trihexyphenidyl. More than eight years on trihexyphenidyl was associated with a substantial increase in the risk of `AD` relative to the other drugs evaluated in Figure 3a. These findings confirm that `ACh` signaling is closely linked with `AD` progression. As such, exploring other drugs and drug classes which either increase `ACh` synaptic bioavailability or enhance neuronal response to `ACh` should be further examined for anti-`AD` activity.

4. Discussion

In this work, we propose a novel network-based, `AD`-specific drug repurposing approach called `MPI`. `MPI` improves upon prior network-based methods by leveraging node embeddings learned via `DeepWalk` to prioritize `AD`-associated paths. Moreover, the use of learned embeddings allows `MPI` to more effectively capture a network's rich topology than previous approaches, such as `BSL`. In a direct comparison, we find that 20% more of `MPI`'s highly prioritized drug candidates (top-50) have published anti-AD evidence compared to `BSL`'s highly prioritized drug candidates. In addition to evidence in literature, we leverage insurance claims data to produce Cox proportional-hazard models. Among all the drugs we evaluate, these models identified `BBBx` `ACE-INHs` as having the lowest risk of `AD`. Similarly, etodolac was found to have the lowest risk of `AD` among the four `COX-INHs` we evaluated (Figure 3b), indicating that this drug in particular may have protective effect despite the class as a whole not exhibiting a significantly reduced risk of `AD` compared to our negative control (Figure 3a). Additionally, `MPI` highly prioritizes drugs that target the cholinergic system. Each of the approved `AD` drugs in our dataset that are also `AChE-INHs` are prioritized in the top-50 by `MPI`. `MPI` also highly prioritizes nicotine, an `AChR-Ags`. This prioritization is supported by both literature and our Cox models, which suggest nicotine is associated with reduced risk of `AD`. Altogether, the results presented in this work highlight etodolac, nicotine, and `ACE-INHs` as viable candidates for repurposing to treat `AD` and, as such, deserve further examination in future studies.

Despite its promising results, `MPI` exhibits a few limitations. The PPI network we construct is a simplification of molecular pathways. Like many other network-based approaches, `MPI` does not consider loops nor the directionality of PPI as these can be difficult for models to learn. In our context, this means that highly ranked candidates are only likely to be in close proximity to `AD`-related genes. To improve drug prioritization, models must be capable of identifying drugs that are both upstream of and in close proximity to these `AD`-related genes. In future studies, we will leverage directed interactions either by hard coding them or learning them. One way directionality might be learned is through the use of multi-omics data. Examining how changes to genomic and epigenomic profiles affect gene expression could facilitate learning where genes are in pathways. Furthermore, by leveraging multi-omics data, we may be able to provide more personalized drug recommendations.

5. Acknowledgments

References

1. M. V. F. Silva, C. d. M. G. Loures, L. C. V. Alves, L. C. de Souza, K. B. G. Borges and M. d. G. Carvalho, Alzheimer's disease: risk factors and potentially protective measures, <u>Journal of biomedical science</u> **26**, 1 (2019).
2. E. Passeri, K. Elkhoury, M. Morsink, K. Broersen, M. Linder, A. Tamayol, C. Malaplate,

F. T. Yen and E. Arab-Tehrany, Alzheimer's disease: Treatment strategies and their limitations, International journal of molecular sciences **23**, p. 13954 (2022).

3. T. Athar, K. Al Balushi and S. A. Khan, Recent advances on drug development and emerging therapeutic agents for alzheimer's disease, Molecular biology reports **48**, 5629 (2021).

4. T.-W. Yu, H.-Y. Lane and C.-H. Lin, Novel therapeutic approaches for alzheimer's disease: An updated review, International journal of molecular sciences **22**, p. 8208 (2021).

5. C. K. Kim, Y. R. Lee, L. Ong, M. Gold, A. Kalali and J. Sarkar, Alzheimer's disease: Key insights from two decades of clinical trial failures, Journal of Alzheimer's Disease **87**, 83 (2022).

6. J. L. Cummings, D. P. Goldman, N. R. Simmons-Stern and E. Ponton, The costs of developing treatments for alzheimer's disease: A retrospective exploration, Alzheimer's & Dementia **18**, 469 (2022).

7. S. Pushpakom, F. Iorio, P. A. Eyers, K. J. Escott, S. Hopper, A. Wells, A. Doig, T. Guilliams, J. Latimer, C. McNamee et al., Drug repurposing: progress, challenges and recommendations, Nature reviews Drug discovery **18**, 41 (2019).

8. P. Zhan, B. Yu and L. Ouyang, Drug repurposing: An effective strategy to accelerate contemporary drug discovery, Drug discovery today **27**, p. 1785 (2022).

9. C. G. Begley, M. Ashton, J. Baell, M. Bettess, M. P. Brown, B. Carter, W. N. Charman, C. Davis, S. Fisher, I. Frazer et al., Drug repurposing: Misconceptions, challenges, and opportunities for academic researchers, Science Translational Medicine **13**, p. eabd5524 (2021).

10. K. Park, A review of computational drug repurposing, Translational and clinical pharmacology **27**, 59 (2019).

11. D. N. Sosa and R. B. Altman, Contexts and contradictions: a roadmap for computational drug repurposing with knowledge inference, Briefings in Bioinformatics **23**, p. bbac268 (2022).

12. D. Morselli Gysi, Í. Do Valle, M. Zitnik, A. Ameli, X. Gan, O. Varol, S. D. Ghiassian, J. Patten, R. A. Davey, J. Loscalzo et al., Network medicine framework for identifying drug-repurposing opportunities for covid-19, Proceedings of the National Academy of Sciences **118**, p. e2025581118 (2021).

13. F. Cheng, R. J. Desai, D. E. Handy, R. Wang, S. Schneeweiss, A.-L. Barabasi and J. Loscalzo, Network-based approach to prediction and population-based validation of in silico drug repurposing, Nature communications **9**, p. 2691 (2018).

14. L. Cai, C. Lu, J. Xu, Y. Meng, P. Wang, X. Fu, X. Zeng and Y. Su, Drug repositioning based on the heterogeneous information fusion graph convolutional network, Briefings in bioinformatics **22**, p. bbab319 (2021).

15. K. Savva, M. Zachariou, M. M. Bourdakou, N. Dietis and G. M. Spyrou, Network-based stage-specific drug repurposing for alzheimer's disease, Computational and Structural Biotechnology Journal **20**, 1427 (2022).

16. J. Fang, P. Zhang, Q. Wang, Y. Zhou, C.-W. Chiang, R. Chen, B. Zhang, B. Li, S. J. Lewis, A. A. Pieper et al., Network-based translation of gwas findings to pathobiology and drug repurposing for alzheimer's disease, MedRxiv , 2020 (2020).

17. Y. Sun, J. Han, X. Yan, P. S. Yu and T. Wu, Pathsim: Meta path-based top-k similarity search in heterogeneous information networks, Proceedings of the VLDB Endowment **4**, 992 (2011).

18. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò and Y. Bengio, Graph attention networks (2018).

19. S. Abu-El-Haija, B. Perozzi, R. Al-Rfou and A. A. Alemi, Watch your step: Learning node embeddings via graph attention, Advances in neural information processing systems **31** (2018).

20. A. Grover and J. Leskovec, Node2vec: Scalable feature learning for networks, p. 855–864 (2016).

21. B. Perozzi, R. Al-Rfou and S. Skiena, Deepwalk: Online learning of social representations, p. 701–710 (2014).

22. T. N. Kipf and M. Welling, Semi-supervised classification with graph convolutional networks,

arXiv preprint arXiv:1609.02907 (2016).

23. J. Fang, A. A. Pieper, R. Nussinov, G. Lee, L. Bekris, J. B. Leverenz, J. Cummings and F. Cheng, Harnessing endophenotypes and network medicine for alzheimer's drug repurposing, Medicinal research reviews **40**, 2386 (2020).

24. Y. Huang, Z. Zhao, X. Wei, Y. Zheng, J. Yu, J. Zheng and L. Wang, Long-term trihexyphenidyl exposure alters neuroimmune response and inflammation in aging rat: relevance to age and alzheimer's disease, Journal of Neuroinflammation **13**, 1 (2016).

25. D. Mendez, A. Gaulton, A. P. Bento, J. Chambers, M. De Veij, E. Félix, M. P. Magariños, J. F. Mosquera, P. Mutowo, M. Nowotka et al., Chembl: towards direct deposition of bioassay data, Nucleic acids research **47**, D930 (2019).

26. T. Liu, Y. Lin, X. Wen, R. N. Jorissen and M. K. Gilson, Bindingdb: a web-accessible database of experimentally determined protein–ligand binding affinities, Nucleic acids research **35**, D198 (2007).

27. X. Chen, Z. L. Ji and Y. Z. Chen, Ttd: therapeutic target database, Nucleic acids research **30**, 412 (2002).

28. S. D. Harding, J. F. Armstrong, E. Faccenda, C. Southan, S. P. Alexander, A. P. Davenport, A. J. Pawson, M. Spedding, J. A. Davies and NC-IUPHAR, The iuphar/bps guide to pharmacology in 2022: curating pharmacology for covid-19, malaria and antibacterials, Nucleic Acids Research **50**, D1282 (2022).

29. R. Apweiler, A. Bairoch, C. H. Wu, W. C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane et al., Uniprot: the universal protein knowledgebase, Nucleic acids research **32**, D115 (2004).

30. Stanford Center for Population Health Sciences, MarketScan Medicare Supplemental (February 2023).

31. J.-C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, G. Jun, A. L. DeStefano, J. C. Bis, G. W. Beecham et al., Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for alzheimer's disease, Nature genetics **45**, 1452 (2013).

32. R. E. Marioni, S. E. Harris, Q. Zhang, A. F. McRae, S. P. Hagenaars, W. D. Hill, G. Davies, C. W. Ritchie, C. R. Gale, J. M. Starr et al., Gwas on family history of alzheimer's disease, Translational psychiatry **8**, p. 99 (2018).

33. I. E. Jansen, J. E. Savage, K. Watanabe, J. Bryois, D. M. Williams, S. Steinberg, J. Sealock, I. K. Karlsson, S. Hägg, L. Athanasiu et al., Genome-wide meta-analysis identifies new loci and functional pathways influencing alzheimer's disease risk, Nature genetics **51**, 404 (2019).

34. B. W. Kunkle, B. Grenier-Boley, R. Sims, J. C. Bis, V. Damotte, A. C. Naj, A. Boland, M. Vronskaya, S. J. Van Der Lee, A. Amlie-Wolf et al., Genetic meta-analysis of diagnosed alzheimer's disease identifies new risk loci and implicates aβ, tau, immunity and lipid processing, Nature genetics **51**, 414 (2019).

35. I. De Rojas, S. Moreno-Grau, N. Tesi, B. Grenier-Boley, V. Andrade, I. E. Jansen, N. L. Pedersen, N. Stringa, A. Zettergren, I. Hernández et al., Common variants in alzheimer's disease and risk stratification by polygenic risk scores, Nature communications **12**, p. 3417 (2021).

36. D. P. Wightman, I. E. Jansen, J. E. Savage, A. A. Shadrin, S. Bahrami, D. Holland, A. Rongve, S. Børte, B. S. Winsvold, O. K. Drange et al., A genome-wide association study with 1,126,563 individuals identifies new risk loci for alzheimer's disease, Nature genetics **53**, 1276 (2021).

37. C. Bellenguez, F. Küçükali, I. E. Jansen, L. Kleineidam, S. Moreno-Grau, N. Amin, A. C. Naj, R. Campos-Martin, B. Grenier-Boley, V. Andrade et al., New insights into the genetic etiology of alzheimer's disease and related dementias, Nature genetics **54**, 412 (2022).

38. A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini and V. A. McKusick, Online mendelian inheritance in man (omim), a knowledgebase of human genes and genetic disorders, Nucleic acids research **33**, D514 (2005).

39. A. P. Davis, C. G. Murphy, C. A. Saraceni-Richards, M. C. Rosenstein, T. C. Wiegers and C. J. Mattingly, Comparative toxicogenomics database: a knowledgebase and discovery tool for chemical–gene–disease networks, Nucleic acids research **37**, D786 (2009).
40. W. Yu, M. Gwinn, M. Clyne, A. Yesupriya and M. J. Khoury, A navigator for human genome epidemiology, Nature genetics **40**, 124 (2008).
41. J. Piñero, À. Bravo, N. Queralt-Rosinach, A. Gutiérrez-Sacristán, J. Deu-Pons, E. Centeno, J. García-García, F. Sanz and L. I. Furlong, Disgenet: a comprehensive platform integrating information on human disease-associated genes and variants, Nucleic acids research , p. gkw943 (2016).
42. M. J. Landrum, J. M. Lee, M. Benson, G. R. Brown, C. Chao, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, W. Jang et al., Clinvar: improving access to variant interpretations and supporting evidence, Nucleic acids research **46**, D1062 (2018).
43. D. Ochoa, A. Hercules, M. Carmona, D. Suveges, J. Baker, C. Malangone, I. Lopez, A. Miranda, C. Cruz-Castillo, L. Fumis et al., The next-generation open targets platform: reimagined, redesigned, rebuilt, Nucleic acids research **51**, D1353 (2023).
44. D. Deb, K. Bairy, V. Nayak, M. Rao et al., Comparative effect of lisinopril and fosinopril in mitigating learning and memory deficit in scopolamine-induced amnesic rats, Advances in Pharmacological and Pharmaceutical Sciences **2015** (2015).
45. P. Newhouse, K. Kellar, P. Aisen, H. White, K. Wesnes, E. Coderre, A. Pfaff, H. Wilkins, D. Howard and E. Levin, Nicotine treatment of mild cognitive impairment: a 6-month double-blind pilot clinical trial, Neurology **78**, 91 (2012).
46. C. Zhang, Y. Wang, D. Wang, J. Zhang and F. Zhang, Nsaid exposure and risk of alzheimer's disease: an updated meta-analysis from cohort studies, Frontiers in aging neuroscience **10**, p. 83 (2018).
47. B. P. Imbimbo, V. Solfrizzi and F. Panza, Are nsaids useful to treat alzheimer's disease or mild cognitive impairment?, Frontiers in aging neuroscience **2**, p. 1517 (2010).
48. F. R. Ianiski, C. B. Alves, C. F. Ferreira, V. C. Rech, L. Savegnago, E. A. Wilhelm and C. Luchese, Meloxicam-loaded nanocapsules as an alternative to improve memory decline in an alzheimer's disease model in mice: involvement of na+, k+-atpase, Metabolic brain disease **31**, 793 (2016).
49. P. Guan, D. Zhu and P. Wang, Meloxicam inhibits apoptosis in neurons by deactivating tumor necrosis factor receptor superfamily member 25, leading to the decreased cleavage of dna fragmentation factor subunit α in alzheimer's disease, Molecular Neurobiology **60**, 395 (2023).
50. F. R. Ianiski, C. B. Alves, A. C. G. Souza, S. Pinton, S. S. Roman, C. R. Rhoden, M. P. Alves and C. Luchese, Protective effect of meloxicam-loaded nanocapsules against amyloid-β peptide-induced damage in mice, Behavioural Brain Research **230**, 100 (2012).
51. K. H. Elfakhri, I. M. Abdallah, A. D. Brannen and A. Kaddoumi, Multi-faceted therapeutic strategy for treatment of alzheimer's disease by concurrent administration of etodolac and α-tocopherol, Neurobiology of disease **125**, 123 (2019).
52. G. Forloni, L. Colombo, L. Girola, F. Tagliavini and M. Salmona, Anti-amyloidogenic activity of tetracyclines: studies in vitro, FEBS letters **487**, 404 (2001).
53. M. A. Tikhonova, T. G. Amstislavskaya, Y.-J. Ho, A. A. Akopyan, M. V. Tenditnik, M. V. Ovsyukova, A. A. Bashirzade, N. I. Dubrovina and L. I. Aftanas, Neuroprotective effects of ceftriaxone involve the reduction of aβ burden and neuroinflammatory response in a mouse model of alzheimer's disease, Frontiers in Neuroscience **15**, p. 736786 (2021).
54. J. P. Modi, H. Prentice and J.-Y. Wu, Sulindac for stroke treatment: neuroprotective mechanism and therapy, Neural Regeneration Research **9**, p. 2023 (2014).
55. K. M. Lohr, B. Frost, C. Scherzer and M. B. Feany, Biotin rescues mitochondrial dysfunction and neurotoxicity in a tauopathy model, Proceedings of the National Academy of Sciences **117**,

33608 (2020).

56. F. Maier, A. Spottke, J.-P. Bach, C. Bartels, K. Buerger, R. Dodel, A. Fellgiebel, K. Fliessbach, L. Frölich, L. Hausner et al., Bupropion for the treatment of apathy in alzheimer disease: a randomized clinical trial, JAMA network open **3**, e206027 (2020).

57. M. Ouk, C.-Y. Wu, J. S. Rabin, A. Jackson, J. D. Edwards, J. Ramirez, M. Masellis, R. H. Swartz, N. Herrmann, K. L. Lanctot et al., The use of angiotensin-converting enzyme inhibitors vs. angiotensin receptor blockers and cognitive decline in alzheimer's disease: the importance of blood-brain barrier penetration and apoe ε4 carrier status, Alzheimer's Research & Therapy **13**, 1 (2021).

58. J. Thomas, H. Smith, C. A. Smith, L. Coward, G. Gorman, M. De Luca and P. Jumbo-Lucioni, The angiotensin-converting enzyme inhibitor lisinopril mitigates memory and motor deficits in a drosophila model of alzheimer's disease, Pathophysiology **28**, 307 (2021).

59. R. Knowles, Denitrification, Microbiological reviews **46**, 43 (1982).