CS5891 – Network Analysis in Healthcare – Final Report

Disease-Drug Interactions Using Network Analysis

Abstract:

We present a network analysis of disease-drug associations to explore novel drug repurposing opportunities and to deepen our understanding of disease progression mechanisms. This method is particularly relevant for addressing prevalent diseases and is critical in the context of rare diseases like Alzheimer's, which currently lack effective treatment options. Our approach involves repurposing existing drugs which act as alternatives that offer similar efficacy. We conducted a comprehensive comparison of candidate drugs against a spectrum of drugs, including those approved, failed, or in phases I, II, and III of clinical trials for Alzheimer's. This comparison was based on structural similarity, quantified using the Tanimoto coefficient. Additionally, we calculated node-level metrics within the network to identify key influential nodes. Following the network construction, we executed a clustering analysis. This enabled us to categorize drugs based on their CoDRes values and identify the most structurally similar drugs for potential repurposing. The study culminates in the ranking of the top 10 candidate drugs for repurposing, providing a valuable resource for future therapeutic development.

Introduction:

Disease-drug interactions is crucial for effective healthcare management. These interactions help us identify how drugs affect a specific disease, enabling the development of more effective and targeted treatments. By understanding the interaction between a drug and a disease, we can predict the drug's efficacy, potential side effects, and how it might alter the disease's progression. This knowledge is vital for personalizing medicine, ensuring that patients receive the most suitable treatment for their condition. Additionally, it aids in the discovery of new uses for existing drugs, potentially leading to quicker and more cost-effective treatment options.

Alzheimer's disease (AD) is a chronic, progressive neurodegenerative disorder primarily characterized by cognitive impairment and dementia, eventually leading to premature death. It's the most common cause of dementia, accounting for up to 75% of cases. AD patients often experience behavioral changes, sleep disturbances, and loss of bodily functions. The disease's pathophysiology involves the accumulation of amyloid-β plaques and tau neurofibrillary tangles in the brain, contributing to neuronal loss and neurodegeneration. Current treatments for AD are limited and primarily aim to slow its progression. Recent developments in treatment offer some hope, but their real-world effectiveness is yet to be fully validated.

Our primary objective has been inspired by the paper *Network-based stage-specific drug repurposing for Alzheimer's disease,* which proposes novel method for identifying potential therapies for Alzheimer's disease by repurposing existing drugs. This approach involves ranking drugs based on their structural similarities to existing treatments and evaluating their ability to penetrate the blood-brain barrier. The method identifies 26 candidate drugs for different stages of Alzheimer's disease, offering a promising strategy for drug discovery in neurodegenerative disease.

Network analysis offers a new approach to understanding and treating elusive diseases like Alzheimer's by exploring disease-drug interactions. It focuses on drug repurposing, enhancing our knowledge of potential treatments. The importance of drug repurposing for conditions like Alzheimer's, where effective treatments are limited, lies in finding new applications for existing drugs. This approach goes beyond mere reuse; it revitalizes the battle against long-standing medical challenges.

Our approach involves a detailed comparison of a wide range of drugs, including those approved, failed, or in various stages of clinical trials for Alzheimer's. We focus on the structural similarity between drugs, measured by the Tanimoto coefficient, a common metric in cheminformatics for comparing molecular structures. By setting a Tanimoto coefficient threshold of 0.5, we select the most similar drugs to form our main network. Within this network, we analyze node-level metrics such as betweenness, closeness, PageRank, and eigenvector centralities to understand the network's structure and pinpoint key drugs that could be effective against Alzheimer's. This network-based approach advances beyond traditional drug discovery methods by utilizing connections and interaction patterns to uncover new possibilities for drug repurposing.

The network analysis culminates in a clustering exercise crucial for categorizing drugs using their CoDRes (Coefficient of Drug Repurposing), a metric measuring a drug's potential for Alzheimer's repurposing based on structural similarities with other drugs. This clustering groups structurally similar drugs and paves the way for highlighting the top 10 most promising candidates for repurposing.

Materials and Methods

Data:

From the Bio-SNAP database of Disease-drug associations, 26 candidate drugs were chosen for analysis in the context of Alzheimer's disease. This database, which can be accessed at Bio-SNAP Database, includes the MESHID of diseases along with the chemical IDs of drugs. In total, it lists 124 drugs that are currently being tested for Alzheimer's. The BioSNAP database for disease-drug associations is comprehensive, encompassing 5,535 disease nodes, 1,662 drug nodes, and 466,656 edges.

The Alzheimer's Association database was the source for information on FDA-approved drugs for Alzheimer's. This database, detailed at Alzheimer's Association, lists 8 drugs that have received FDA approval for the treatment of Alzheimer's disease. To identify drugs that failed in clinical trials, the study referred to PubMed Central, specifically reviewing an article that examined over 2,700 clinical trials for Alzheimer's drugs. This paper, available at PubMed Central, provided insights into the drugs that did not meet efficacy criteria for Alzheimer's treatment.

All clinical trial phases (I, II, and III) were considered for validation against the information from the Alzheimer's Association Journal. A total of 143 drugs were identified through this source: 31 from Phase III, 82 from Phase II, and 30 from Phase I trials. More details can be found at Alzheimer's Association Journal. The CoDReS scores validate each candidate drug against a disease, accessible at CoDReS Database, was utilized for calculating the CoDReS value which provides a structural and functional score.

Data processing:

For our analysis, we utilized 26 candidate drugs. After gathering all the drugs from each category, we obtained their SMILE scores from DrugBank to calculate the Tanimoto coefficient. Below are examples of some of the SMILE scores.

Drug Name	**Category**	**SMILE**
Maprotiline	Candidate Drug	CNCCCC12CCC(C3=CC=CC=C13)C1=CC=CC=C21
Donepezil	Approved	COC1=C(OC)C=C2C(=O)C(CC3CCN(CC4=CC=CC=C4)CC3)CC2=C1
Tarenflurbil	Failed	C[C@@H](C(O)=O)C1=CC(F)=C(C=C1)C1=CC=CC=C1
Dabigatran	Phase I	CN1C(CNC2=CC=C(C=C2)C(N)=N)=NC2=C1C=CC(=C2)C(=O)N(CCC(O)=O)C1=NC=CC=C1
Bromocriptine	Phase II	[H][C@@]12CCCN1C(=O)[C@H](CC(C)C)N1C(=O)[C@](NC(=O)[C@H]3CN(C)[C@]4([H])CC5=C(Br)NC6=CC=CC(=C56)C4=C3)(O[C@@]21O)C(C)C

Table 1. SMILE values

Methodology:

Upon retrieving the SMILE scores for all drugs, we compared each candidate drug with every drug from each category using the Tanimoto Coefficient. The Tanimoto similarity is particularly useful for recognizing drugs that are structurally more similar to approved drugs or those in clinical trials and less similar to previously failed drugs, aiding in the effective selection of promising candidates for Alzheimer's. It is a ratio representing the shared features between two molecules relative to their total features, ranges from 0 to 1. In our model, drugs are depicted as nodes, and their interconnections, or edges, are assigned weights based on this coefficient. Approved drugs are given a weight of 1, failed drugs -1, and drugs in different trial phases receive incremental weights (0.2 for Phase I, 0.4 for Phase II, 0.6 for Phase III). These weights, indicative of the drug’s stage in development, are then integrated with the pre-existing connection weights in the network.

Drug Candidate	Drug_Category	Comparison_Dru	Tanimoto_similarity	Weights
Maprotiline	Approved	Donepezil	0.10435091	0.10435091
Maprotiline	Approved	Galantamine	0.338363469	0.338363469
Maprotiline	Approved	Rivastigmine	0.442754479	0.442754479
Maprotiline	Approved	Memantine	0.028820868	0.028820868
Maprotiline	Approved	Brexpiprazole	0.051395529	0.051395529
Maprotiline	Phase 1	Allopregnanolone	0.829319946	0.165863989
Xanthine	Phase 1	Allopregnanolone	0.561831055	0.112366211
Olanzapine	Phase 1	Allopregnanolone	0.592566357	0.118513271
Maprotiline	Failed	Lanabecestat	0.222188526	-0.222188526
Maprotiline	Failed	Verubecestat	0.228215401	-0.228215401
Maprotiline	Failed	Azeliragon	0.341675942	-0.341675942
Staurosporine	Phase 1	Allopregnanolone	0.558464628	0.111692926
Tenamfetamine	Phase 1	Allopregnanolone	0.620755272	0.124151054
Maprotiline	Failed	Tarenflurbil	0.360203689	-0.360203689
Dantrolene	Phase 1	Allopregnanolone	0.91913884	0.183827768
Gabapentin	Phase 1	Allopregnanolone	0.517126703	0.103425341
Maprotiline	Phase 1	Dabigatran	0.467157393	0.093431479
Pregabalin	Phase 1	Allopregnanolone	0.864425483	0.172885097
Maprotiline	Phase 1	Emtricitabine	0.457967683	0.091593537
Topiramate	Phase 1	Allopregnanolone	0.804497908	0.160899582

Tanimoto Calculation

We developed a comprehensive network representing various drug associations. After calculating the Tanimoto coefficient, a threshold of 0.5 is applied to eliminate low-matching pairs. This filtered data is then used to construct a network for assessing drug repurposing. Centrality metrics like eigenvector, betweenness, closeness, and degree centrality are computed to evaluate the network's structure. We use CoDReS (Composite Drug Reranking Scoring) which is a web-based tool that enhances drug repurposing decisions by integrating an initial drug ranking with functional and structural scores, and by applying structural similarity clustering to suggest potential drug candidates for further validation.

The CoDReS database appears as below, we need to input in our drug list file along with any weight we want to associated with it. It renders a table with all the scores considering the drug list as our input file and validates against Alzheimer's disease. A sample snapshot is shown below. Exemplar drugs are also highlighted which show high efficiency against the drug.

We applied K-means clustering to cluster the drugs based on their similar structures, which enabled the identification and ranking of the most promising candidates in each cluster for the purpose of Alzheimer's disease repurposing analysis.

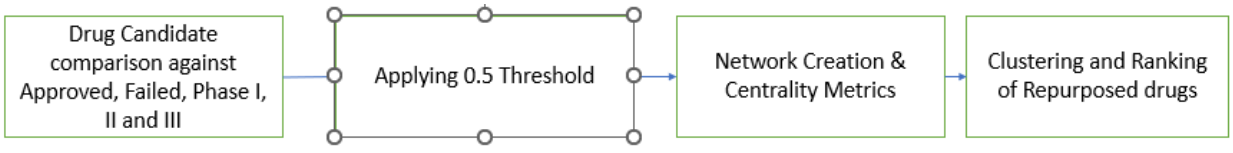

Fig 1. Methodology

Results:

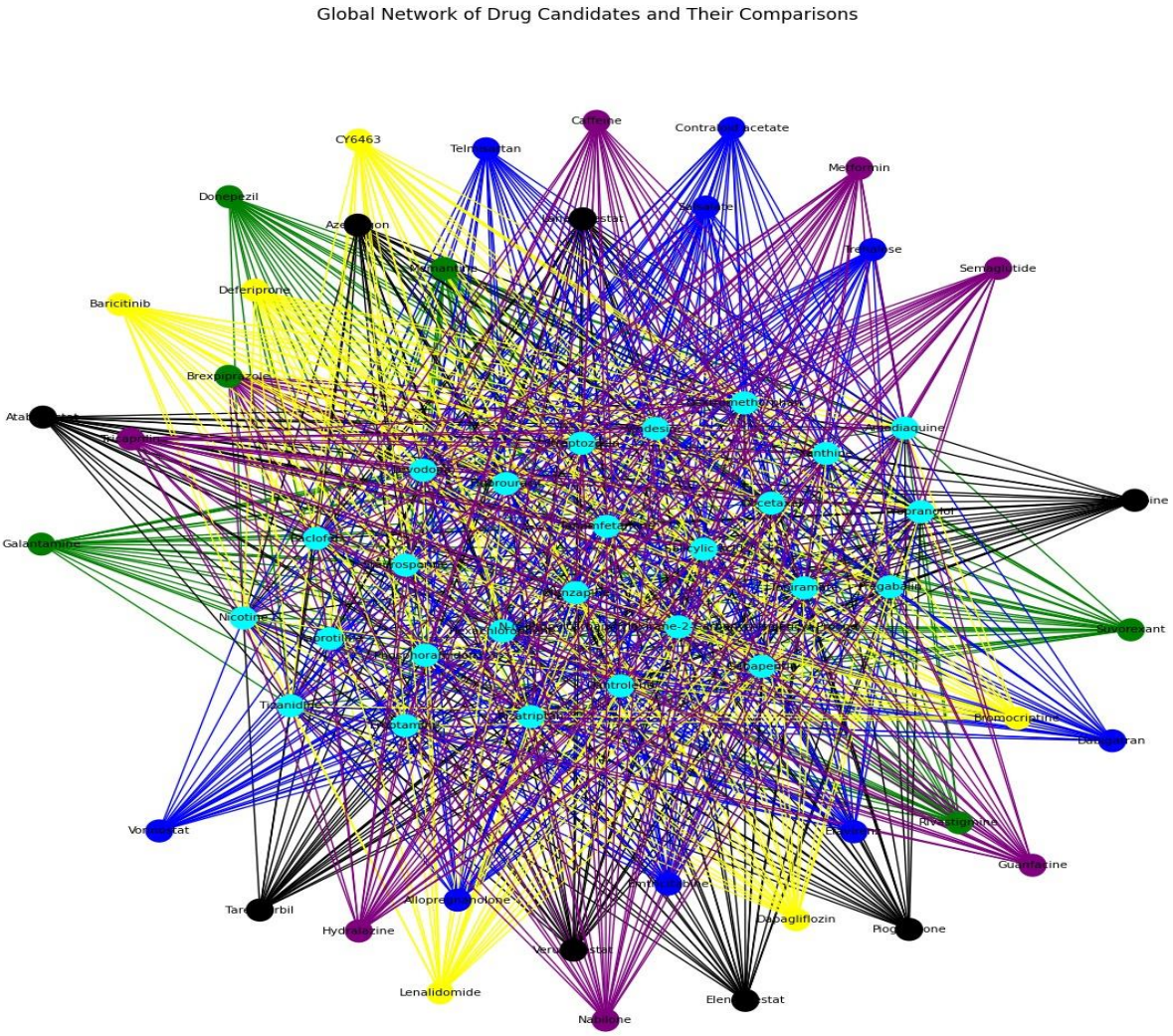

Fig 2. Global Network

Fig 2 illustrates a comprehensive network encompassing all associations between the drug candidates and their comparative drugs. This network consists of 62 nodes and 936 edges in total.

Post applying a threshold of 0.5 Tanimoto coefficient, our primary network would look as *Fig 3*.

```
Number of nodes: 62
Number of edges: 371
```

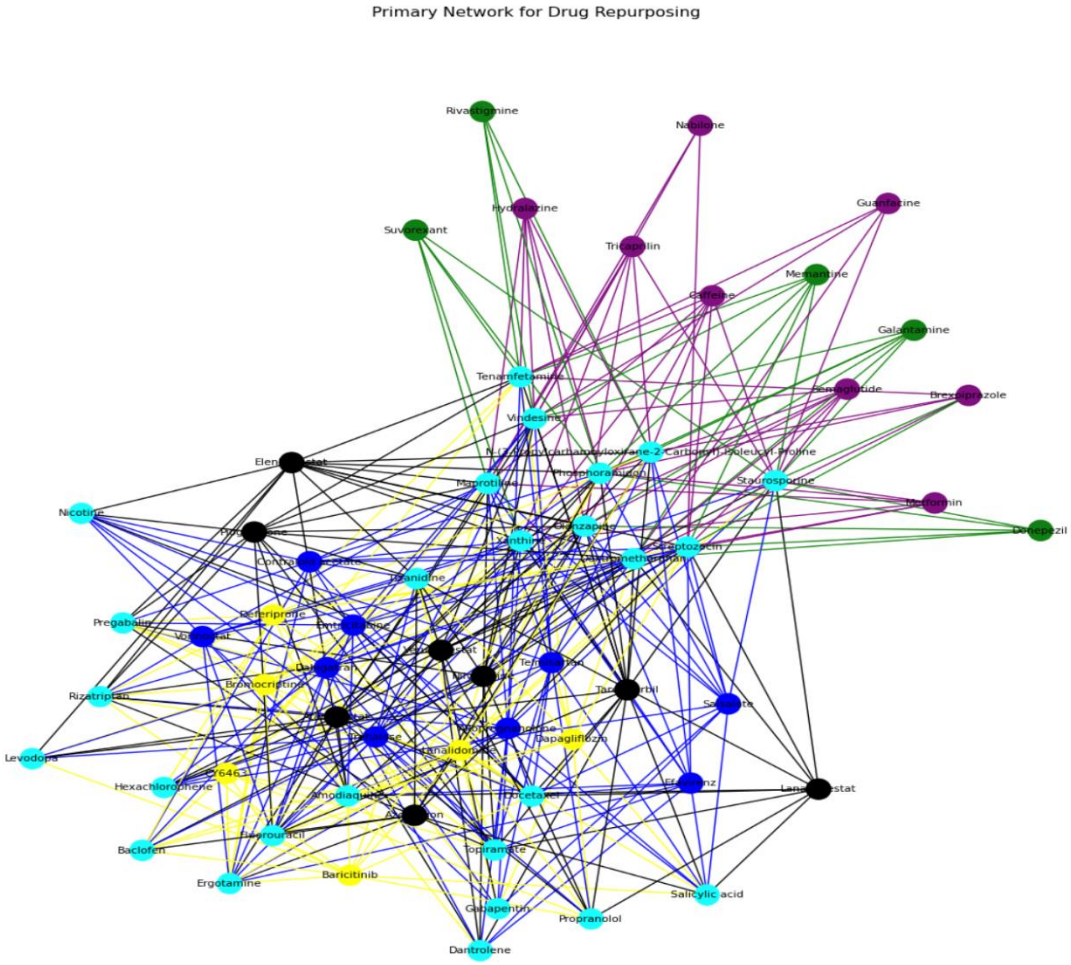

Fig 3. Primary Network

Fig 3 was created for drug repurposing comprises 62 nodes and 371 edges. Following its construction, we proceeded to compute various centrality measures, as detailed below.

Degree Centrality

The highest Degree centrality is Dextromethorphan and lowest Degree centrality is Nabilone with the below measures. Fig 4 represents degree centrality network.

Drug Name	**Degree Centrality Measure**
Dextromethorphan(Highest)	0.3770491803278689
Nabilone(Lowest)	0.04918032786885246

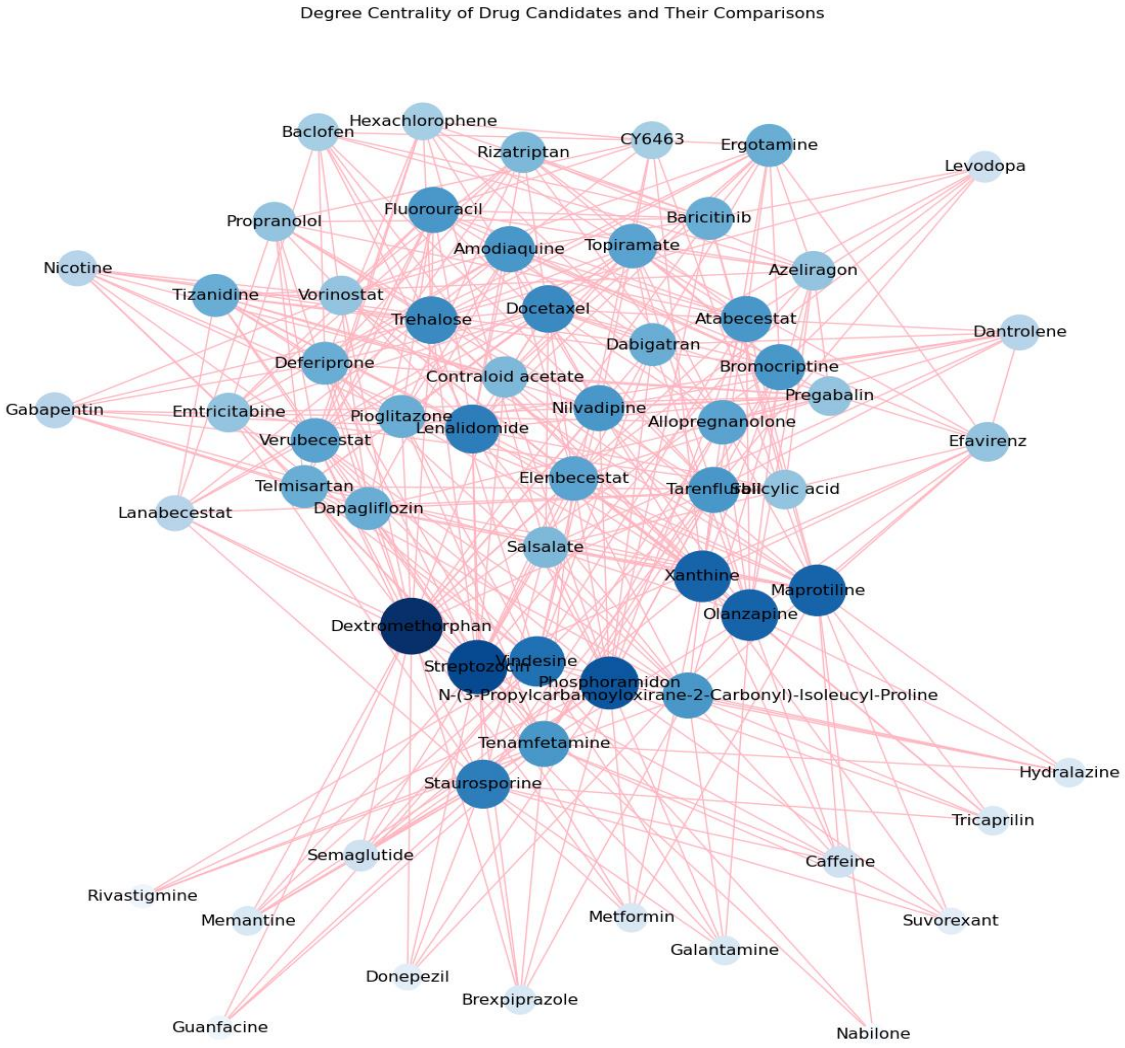

Fig 4. Degree Centrality

Betweenness Centrality

The highest Betweenness centrality is Dextromethorphan and lowest Betweenness centrality is Nabilone with the below measures. Fig 5 represents Betweenness centrality network.

Drug Name	**Betweenness Centrality Measure**
Dextromethorphan(Highest)	0.0697237538230456
Nabilone(Lowest)	0.0007479645617631779

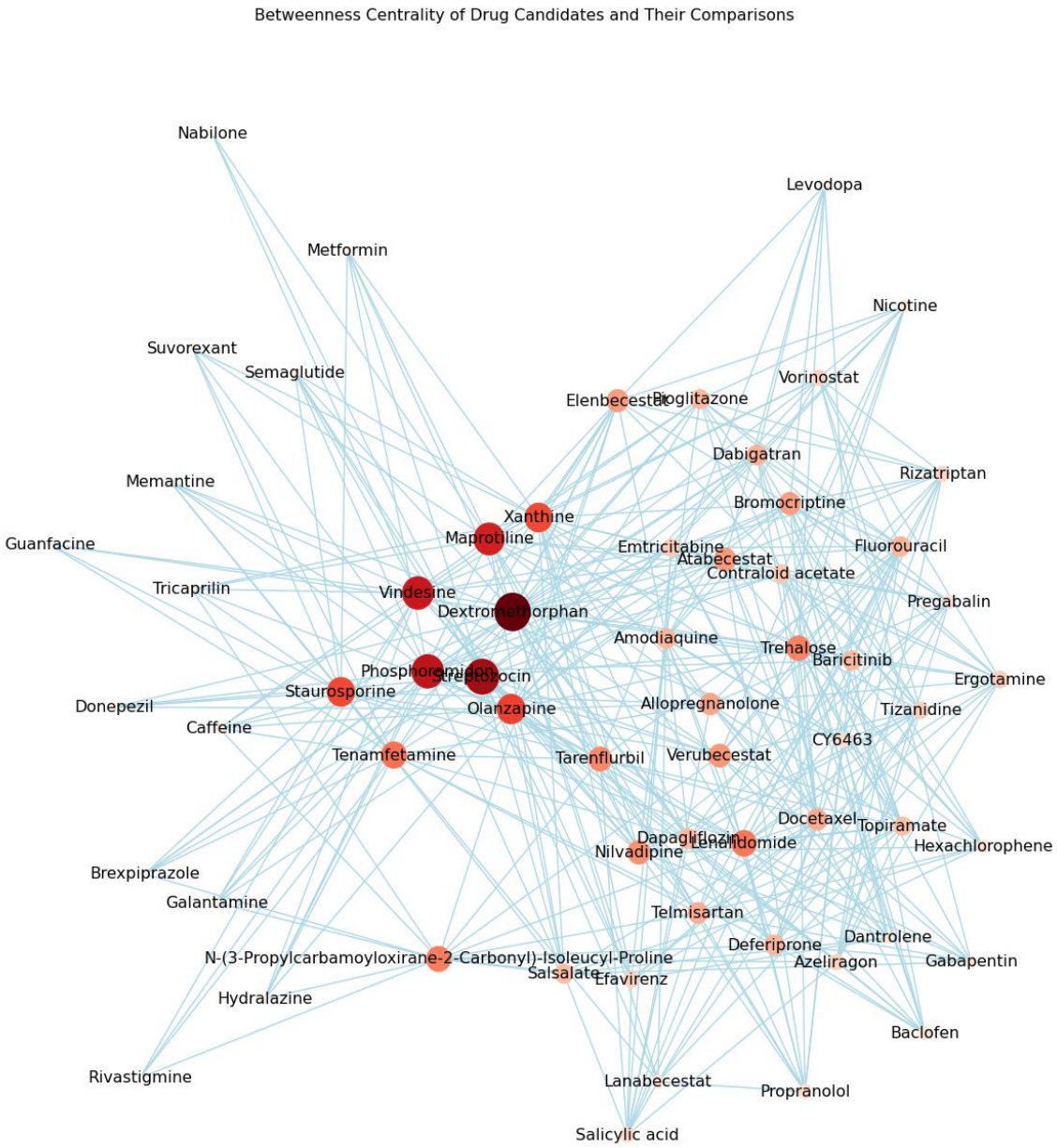

Fig 5. **Betweenness Centrality**

Closeness Centrality

The highest Closeness centrality is Dextromethorphan and lowest Closeness centrality is Nabilone with the below measures. Fig 6 represents closeness centrality network.

Drug Name	**Closeness Centrality Measure**
Dextromethorphan(Highest)	0.5446428571428571
Rivastigmine(Lowest)	0.4236111111111111

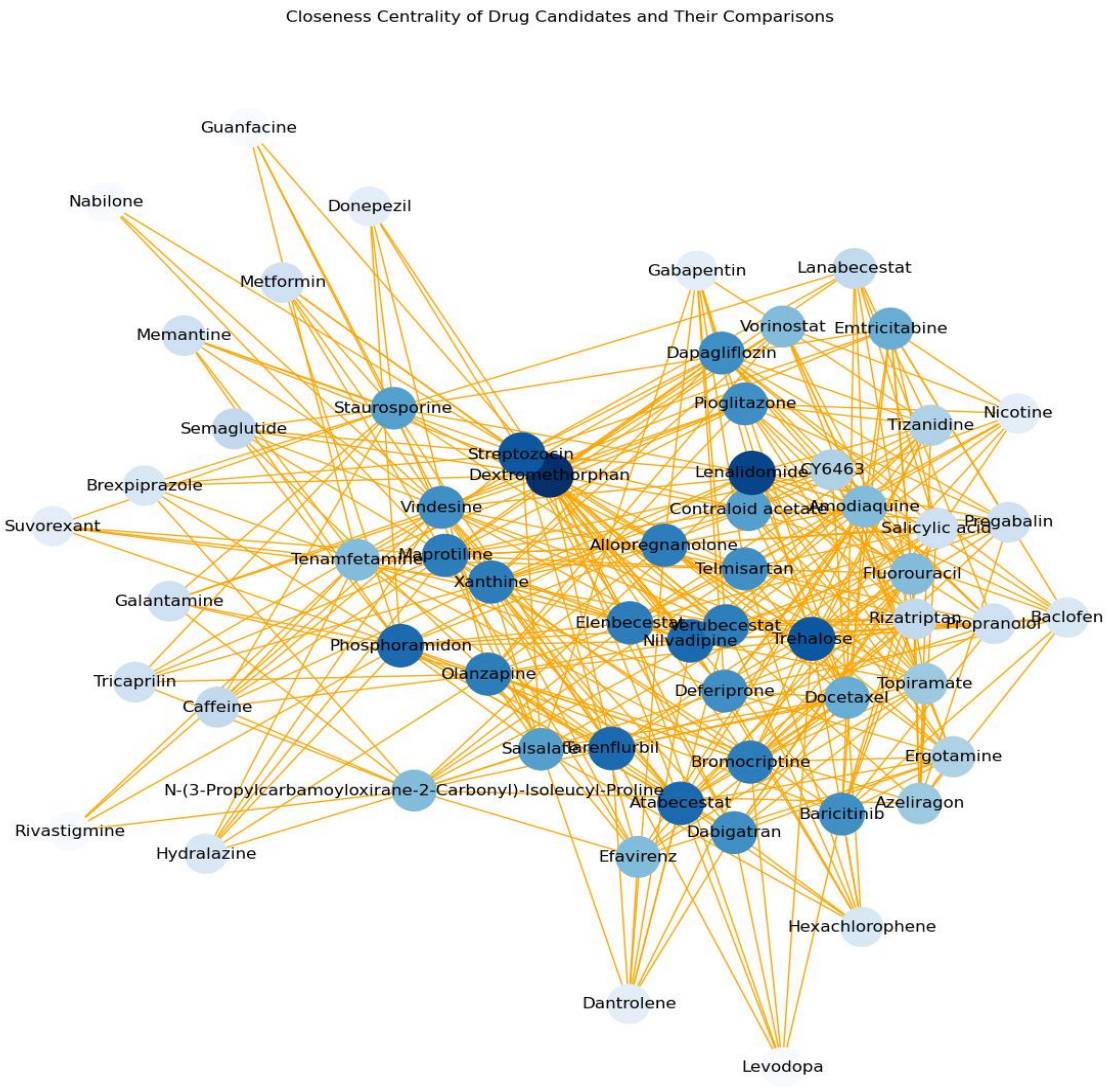

Fig 6. Closeness Centrality

Eigen Vector Centrality

The highest Eigen Vector centrality is Dextromethorphan and lowest Eigen Vector centrality is Nabilone with the below measures. Fig 7 represents Eigen Vector Centrality network.

Drug Name	**EigenVector Centrality Measure**
Dextromethorphan(Highest)	0.20613630681512882
Nabilone(Lowest)	0.03611315336729385

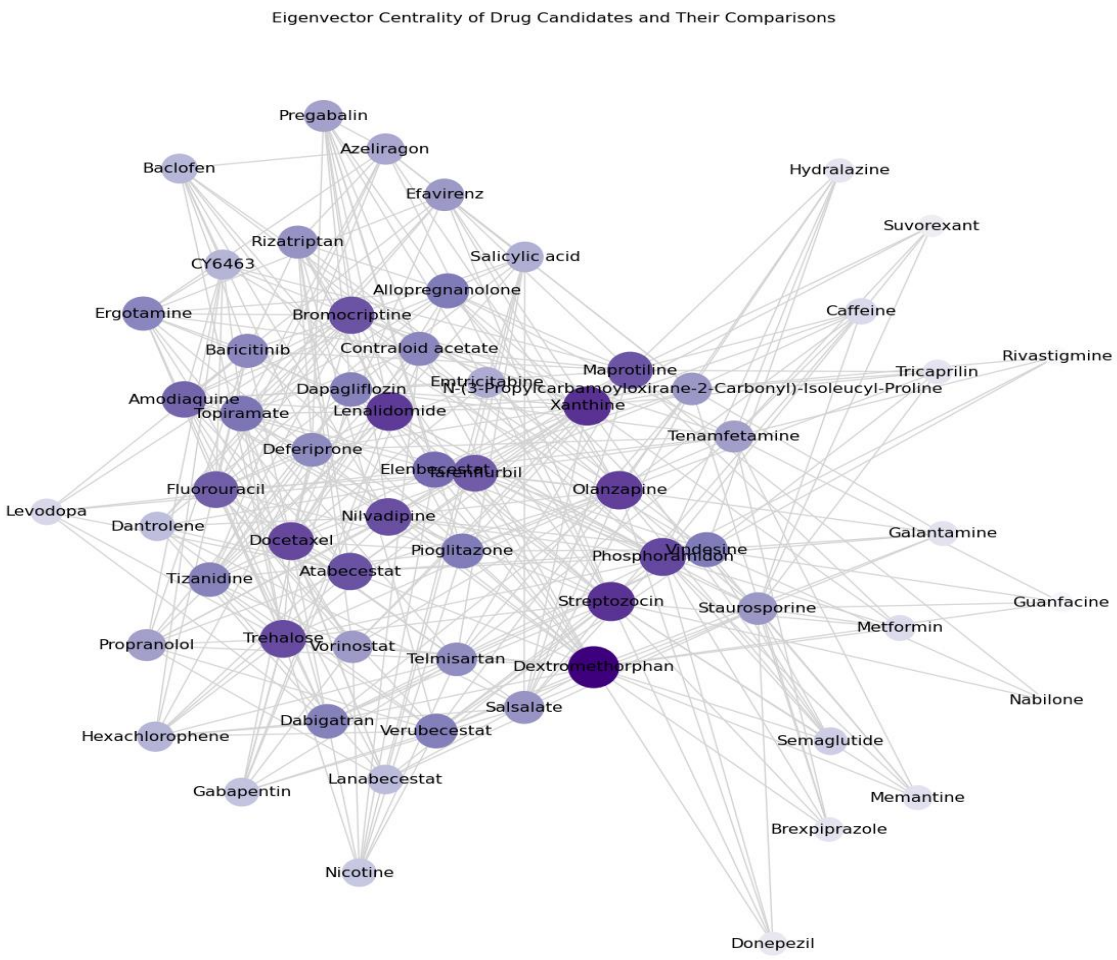

Fig 7. Eigen Vector Centrality

PageRank Centrality

The highest PageRank centrality is Dextromethorphan and lowest PageRank centrality is Nabilone with the below measures. Fig 8 represents Eigen Vector Centrality network.

Drug Name	**PageRank Centrality Measure**
Dextromethorphan(Highest)	0.029635837985969197
Nabilone(Lowest)	0.005833638824394775

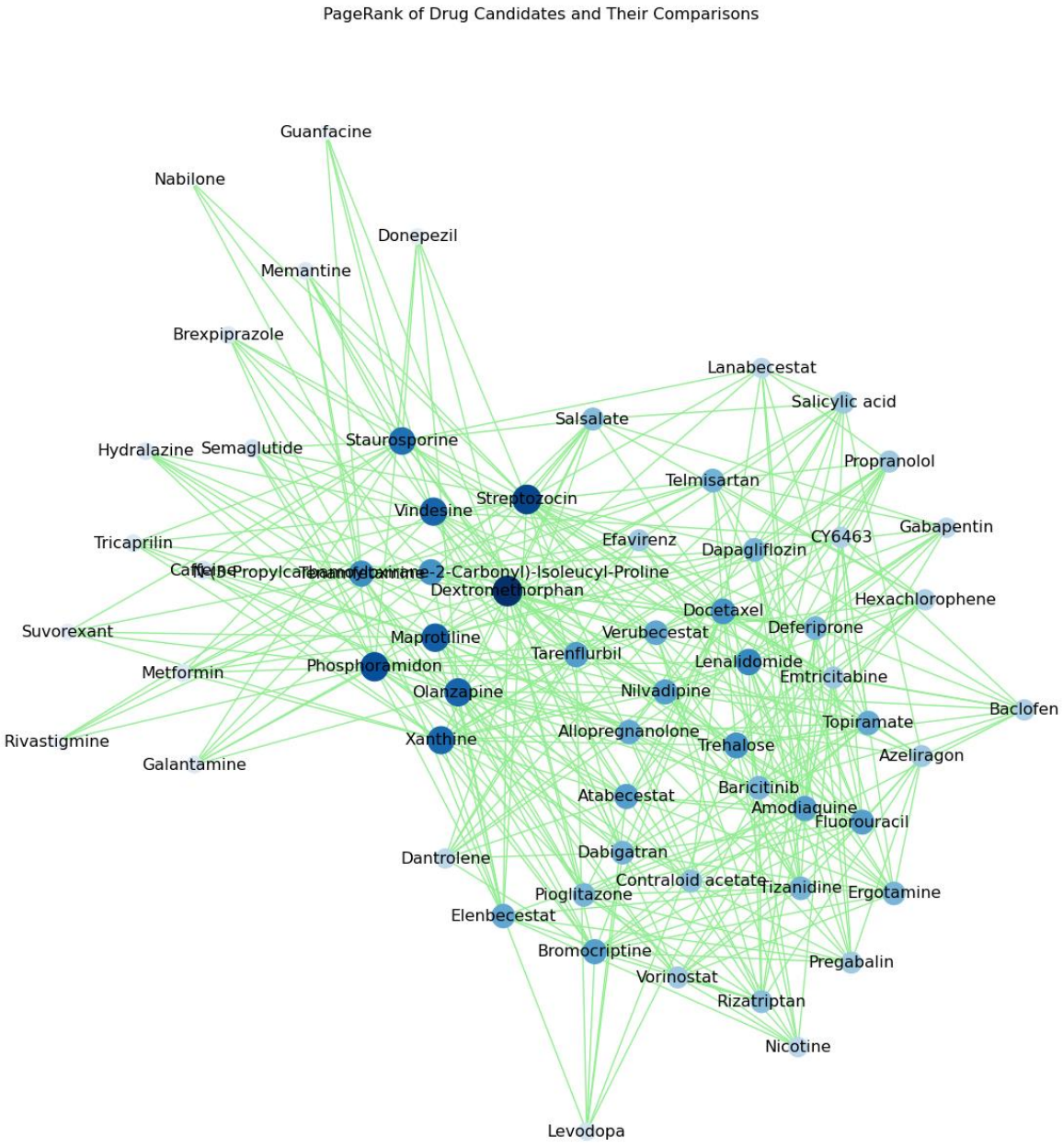

Fig 8. PageRank Centrality

Clustering

We employed K-means clustering to group the drugs according to their CoDReS values. We chose K-means for its ability to efficiently categorize data into distinct clusters.

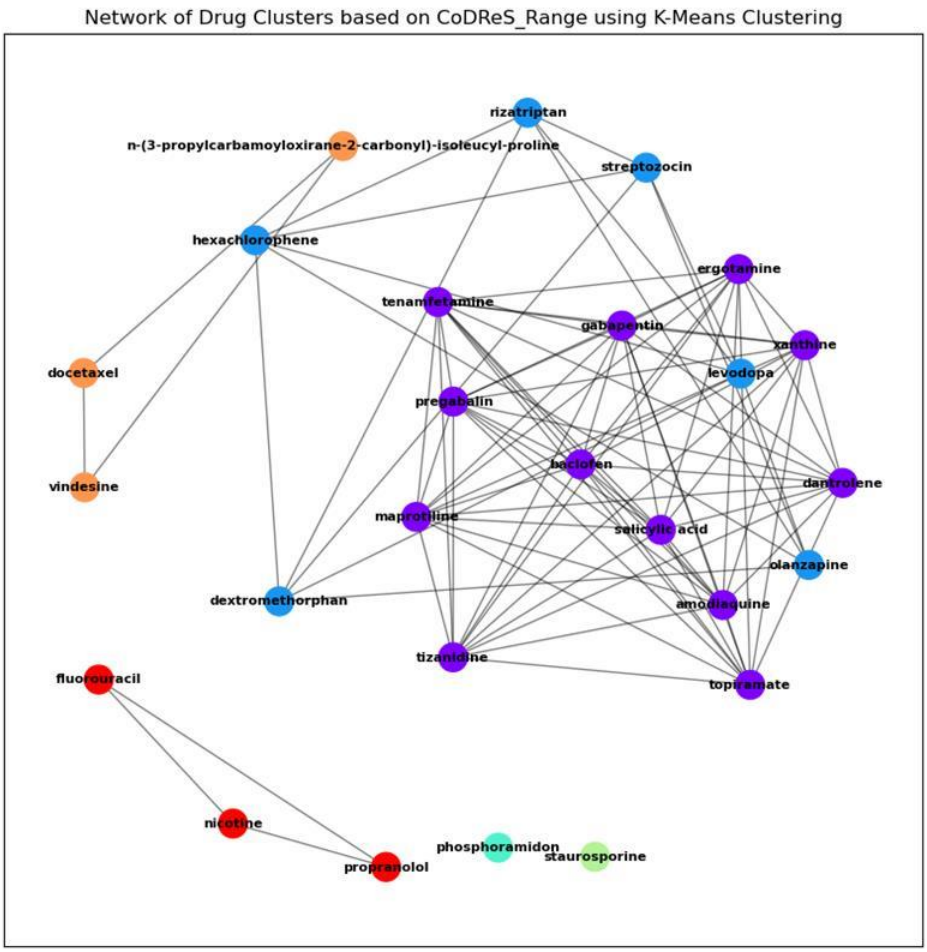

Fig 9. Repurposed Drugs Clustering

From the clusters created, we identified that almost 12 drugs share the similarity coefficient with other drugs treated for Alzheimer's which has been listed below.

Drug Name	**CoDReS**
Baclofen	0.74
Topiramate	0.74
Salicylic acid	0.72
Gabapentin	0.72
Pregabalin	0.71
Tenamfetamine	0.71
Maprotiline	0.71
Amodiaquine	0.71
Dantrolene	0.70

Xanthine	0.70

Table 2. Cluster 1

Drug Name	CoDReS
Streptozocin	0.84
Levodopa	0.84
Dextromethorphan	0.83
Rizatriptan	0.81
Hexachlorophene	0.80
Olanzapine	0.79

Table 3. Cluster 2

Drug Name	CoDReS
Propanolol	0.93
Nicotine	0.90
Flurouracil	0.85

Table 4. Cluster 6

Discussion

The global network illustrates a bipartite network consisting of two distinct sets of nodes representing drug entities and their corresponding categories or targets. The 62 nodes are connected by 938 edges based on Tanimoto coefficients, which indicates similarity or relationship strength without interconnections within the same node group. This structure is key for analyzing potential drug-target interactions and therapeutic correlations.

The primary network zeroes in on potential drug repurposing with an emphasis on significant similarity, featuring only connections where the Tanimoto coefficient is above 0.5. This filter reduces the network to 371 edges, highlighting stronger, more relevant drug interactions and discarding those with weaker similarity. Such a threshold ensures the network reflects more potent chemical resemblances, vital for spotlighting repurposing opportunities among the 62 drugs shown, while excluding less pertinent, weaker chemical analogies to streamline the identification process.

Now, lets discuss about the centrality measures.

The degree network visualizes the degree centrality of drug candidates, with Dextromethorphan having the highest centrality at 0.377, indicating it has numerous connections to other drugs, which may suggest its potential for repurposing. In contrast, Nabilone has the lowest centrality at 0.049, showing it has fewer connections, possibly indicating a more niche application. Degree centrality is pivotal for identifying key drugs that could be significant in drug repurposing. The betweenness centrality network illustrates how often a node acts as a bridge along the shortest path between two other nodes. Dextromethorphan exhibits the highest betweenness centrality at 0.0697, indicating it plays a crucial role in the network,

potentially connecting various therapeutic areas or drug interactions. Conversely, Nabilone has the lowest centrality at approximately 0.00075, suggesting it is less central to the network's connectivity.

The closeness centrality network represents a measure of how close a node is to all other nodes in the network, which can indicate a drug's accessibility. Dextromethorphan stands out with the highest closeness centrality score of approximately 0.545, suggesting that it is relatively central within the network and may be quickly reachable or influential over other nodes. Rivastigmine, not Nabilone as initially indicated, has the lowest closeness centrality at approximately 0.424, indicating it is less central and potentially less influential in the network. Eigenvector centrality measures a node's influence based on the principle that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. Here, Dextromethorphan has the highest eigenvector centrality score of approximately 0.206, suggesting it is a highly influential drug within the network, potentially due to strong connections to other significant nodes. On the other end, Nabilone has the lowest score, around 0.036, indicating it has fewer or weaker influential connections.

PageRank centrality measure is akin to the algorithm used by Google for ranking web pages. In this context, it ranks drug candidates based on their connectivity within the network. Dextromethorphan emerges with the highest PageRank centrality score of approximately 0.0296, signaling it as a key node within this network, potentially indicating its broader prominence in drug interactions. Nabilone has the lowest score of about 0.0058, suggesting it has a more peripheral role in this network.

As mentioned above, CoDReS quantifies the potential for a given drug to be repurposed based on its similarity to drugs already used for a particular condition. In the context of Alzheimer's disease, drugs with a high CoDReS value are likely to share significant biochemical or pharmacological properties with medications currently prescribed for this condition. Utilizing K-means clustering, a widely recognized method for its efficiency in sorting data into meaningful clusters, we classified drugs by their CoDReS values. This approach enabled the identification of drugs that, though not originally intended for Alzheimer's treatment, share significant similarities with Alzheimer’s drugs, suggesting potential new therapeutic uses.

Cluster 1 in the provided network categorizes drugs with CoDReS values that are relatively high, indicating a strong similarity with drugs used for treating Alzheimer's. This cluster includes drugs like Baclofen and Topiramate, each with a CoDReS of 0.74. Cluster 2 comprises drugs like Streptozocin and Levodopa, with CoDReS values also on the higher side (0.84 and 0.84 respectively), indicating their potential for repurposing. Cluster 6 features drugs with a high CoDReS average value of 0.90, indicating a very strong similarity with drugs used to treat Alzheimer's disease. This cluster includes Propranolol, Nicotine, and Fluorouracil, with individual CoDReS values of 0.93, 0.90, and 0.85 respectively.

The implications for drug repurposing are significant here. Given their high CoDReS values, these drugs could potentially be repurposed to treat Alzheimer’s. For instance, Propranolol's high CoDReS value suggests that, although traditionally used as a beta-blocker for cardiovascular issues, it may influence pathways relevant to Alzheimer's treatment. Nicotine has been researched for its neuroprotective effects, which might be leveraged for Alzheimer’s therapy. Fluorouracil, a chemotherapy agent, could offer insights into novel therapeutic pathways due to its biochemical properties. If these drugs can engage targets in the Alzheimer's disease process that are not addressed by current treatments, they may offer new therapeutic strategies, potentially with the added benefit of an already well-characterized safety profile. This could accelerate their clinical application for Alzheimer's compared to developing new drugs from scratch.

As part of future work, further validation of these repurposed drugs is necessary in order to confirm its efficiency fighting against Alzheimer's.

Conclusion:

Network analysis has identified several promising candidates for drug repurposing in the treatment of Alzheimer's disease, leveraging CoDReS values and centrality metrics. By focusing on drugs with structural similarities to those used in Alzheimer's treatment, we've pinpointed potential alternatives that may offer similar efficacy. This demonstrates the power of network analysis in drug repurposing, providing a data-driven foundation for future therapeutic development and a hopeful pathway for enhancing Alzheimer's care.

References:

- Network-based stage-specific drug repurposing for Alzheimer's disease
- Alzheimer's Disease: Key Insights from Two Decades of Clinical Trial Failures
- Reasons for Failed Trials of Disease-Modifying Treatments for Alzheimer Disease and Their Contribution in Recent Research
- A Web Tool for Ranking Candidate Drugs Against a Selected Disease Based on a Combination of Functional and Structural Criteria
- CoDReS
- DrugBank
- ChatGPT