S.T. Yau High School Science Award

Research Report

The Team

Name of team member: Mythreya Dharani
School: Bergen County Academies
City, Country: Paramus, NJ, USA

Title of Research Report

Identifying robust metabolomics signatures for multi-scale stratification of Alzheimer's disease patients into clinically relevant subgroups with AutoSGI

Date

August 24, 2025

Identifying robust metabolomics signatures for multi-scale stratification of Alzheimer's disease patients into clinically relevant subgroups with AutoSGI

Mythreya Dharani

Abstract

Recent studies have established the existence of widespread metabolic alteration in Alzheimer's disease (AD). One way to further analyze possible metabolic mechanisms affecting AD severity is with patient subgroup identification. However, the inherent heterogeneity of the disease, along with the high-throughput profiling of patient samples, makes identifying metabolites that will effectively stratify patients into subgroups difficult. To address this issue, we propose AutoSGI, a data driven algorithm that (1) finds metabolomics signatures carrying concentrated biological information and (2) uses these signatures to cluster AD patients into multi-scale subgroups with distinct disease severity. We test whether AutoSGI can effectively identify patient subgroups using metabolomics measurements from 500 postmortem brain tissue samples. AutoSGI is able to identify subgroups with significant differences in neuropathological and cognitive outcomes defining disease stage. We demonstrate these subgroups display high stability and that in general, AutoSGI can outperform other baseline methods of identification. We also implicate multiple groups of metabolites contained in the identified signatures with advanced AD progression. Some groups are structurally similar like tryptophan, tyrosine, and phenylalanine, the aromatic acids. Other groups functionally interact, like phospholipid intermediates glycerophosphocholine and glycerophosphoethanolamine, as well as the neurotoxic guanidinoacetate and excitatory transmitter glutamate. Ultimately, our findings help elucidate metabolism's role in AD and can support treatment of patients affected by the disease in a more individualized manner.

Keywords: *Alzheimer's disease, metabolism, subgroup identification, machine learning, precision medicine, AutoSGI*

Acknowledgements

I would like to thank Dr. Richa Batra and Dr. Jan Krumsiek for their amazing mentorship with this research project over the last two years. I would also like to thank Dr. Mustafa Buyukozkan for his timely statistical suggestions and guidance. Finally, I would like to thank Dr. Matthias Arnold and Dr. Rima F. Kaddurah-Daouk for their domain-specific insights and collaboration.

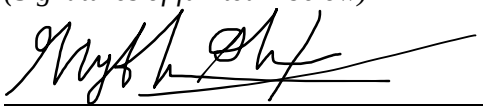
Commitments on Academic Honesty and Integrity

We hereby declare that we

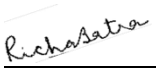
1. are fully committed to the principle of honesty, integrity and fair play throughout the competition.
2. actually perform the research work ourselves and thus truly understand the content of the work.
3. observe the common standard of academic integrity adopted by most journals and degree theses.
4. have declared all the assistance and contribution we have received from any personnel, agency, institution, etc. for the research work.
5. undertake to avoid getting in touch with assessment panel members in a way that may lead to direct or indirect conflict of interest.
6. undertake to avoid any interaction with assessment panel members that would undermine the neutrality of the panel member and fairness of the assessment process.
7. observe the safety regulations of the laboratory(ies) where we conduct the experiment(s), if applicable.
8. observe all rules and regulations of the competition.
9. agree that the decision of YHSA is final in all matters related to the competition.

We understand and agree that failure to honour the above commitments may lead to disqualification from the competition and/or removal of reward, if applicable; that any unethical deeds, if found, will be disclosed to the school principal of team member(s) and relevant parties if deemed necessary; and that the decision of YHSA is final and no appeal will be accepted.

(Signatures of full team below)



Name of team member: Mythreya Dharani



Name of supervising teacher: Richa Batra

Table of Contents

Please note that this paper follows an Introduction-Results/Discussion-Methods format.

We believe this will best convey this project's contributions to the growing body of computational neuroscience research.

]

1. Introduction

As of 2019, dementia affects approximately 57 million people globally. This number is estimated to increase almost three-fold by 2050 (GBD 2019 Dementia Forecasting Collaborators, 2022). Alzheimer's disease (AD) accounts for 60-80% of these cases, with around 6 million American cases alone ("2020 Alzheimer's Disease Facts and Figures," 2020). Though there is a vested interest in finding a treatment for AD, it remains difficult to do so because of how it varies in its presentation across patients. For example, the rate of cognitive decline can be much faster for atypical AD patients. Further, while most AD patients do have a late age of onset, some develop it before turning 65, and there are differences in the exact onset generally. These interpatient complexities are also present when trying to understand the etiology of AD and what defines the disease (Avelar-Pereira et al., 2023; Kauppi et al., 2020; Rollo et al., 2016; Shade et al., 2024; Sirkis et al., 2022).

One way to tackle the complex heterogeneity of AD is by identifying patient subgroups. Fundamentally, the motivation is that the manifestation of certain clinical phenotypes in patients are driven by their specific biomolecular changes. Therefore, by finding patients who are similar to each other biologically, we can consequently find subgroups with more clinical homogeneity. This can help guide treatment decisions on a more precise level but also provides insight into the biological mechanisms of AD.

Many subgroup identification methods have been developed based on these ideas. One common unsupervised approach is clustering the patients into a discrete number of subgroups. Consensus clustering has been used in this manner for patients with breast cancer, lung adenocarcinoma, hepatocellular carcinoma, and pulmonary fibrosis (Horr & Buechler, 2021; Jiang et al., 2024; M. Zhang et al., 2024; N. Zhang et al., 2021). Supervised algorithms have also been utilized, such as an XGBoost-based patient clustering in regards to COVID-19 status and a Bayesian framework to find network-based melanoma subgroups (Cooper et al., 2021; Qin et al., 2024).

In this study, we focus on deriving AD patient subgroups that differ by current disease severity. We previously created a subgroup identification tool known as SGI, or SubGroup Identification

(Buyukozkan et al., 2022). SGI first performs unsupervised hierarchical clustering using the biomolecular measurements of given patient data. Each split point of the resulting hierarchy stratifies a subgroup of patients into two smaller ones. SGI then tests for disease stage differences (based on phenotypes) between the two smaller groups that form at every split point to find which stratifications are meaningful. Since stratification is conducted at every level of the hierarchy, SGI can find patients who are able to be characterized at a more granular level without sacrificing accuracy for others. This offers a distinct advantage over traditional approaches described above which place each patient in a single subgroup.

One limitation of SGI for our objective is that it will assume all features equally dictate AD disease stage, since they are given the same weight when clustering. However, this is not necessarily the case; when analyzing heterogeneous diseases with high-throughput features, finding informative feature subsets is often necessary. This has been observed in contexts like prediction (Qiang et al., 2024), single phenotype association (Batra et al., 2023), and most importantly with SGI, where using only acylcarnitines profiles of AD patients revealed cognitively distinct subgroups at multiple scales (Arnold et al., 2024).

To this end, we developed AutoSGI, a tool for data-driven feature selection followed by multi-scale patient clustering (SGI). AutoSGI first groups features into feature sets at multiple scales. For each feature set, AutoSGI performs multi-scale subgroup identification on patient measurements restricted to those features. Using postmortem brain metabolic profiles of a large cohort with AD and aging brains, AutoSGI is able to identify feature sets that are biologically relevant in the context of disease progression. These features then stratify patients into distinct AD stage subgroups at multiple granularities, increasing clinical applicability compared to traditional methods.

2. Results & Discussion

2.1 Cohort overview

For this study, we analyzed metabolomics measurements of Alzhemer's patients, collected from post-mortem brain samples in the ROS/MAP cohort (Batra et al., 2023). These patients were enrolled in one of two studies conducted by the Rush Alzheimer's Disease Center, the Religious Order Study (ROS) or the Rush Memory and Aging Project (MAP). Metabolites were sampled from the dorsolateral prefrontal cortex and profiled with an untargeted metabolomics platform, Metabolon Inc. Preprocessing procedures for the raw untargeted metabolomics measurements are described in the methodology section (**Methods 3.1**). Along with the metabolic profiles themselves, several clinical outcomes of the ROS/MAP patients were utilized, including potential confounding variables like age, sex, and APOE status, as well as disease stage indicators like severity of neurofibrillary tangle and neuritic plaque pathology. The clinical outcomes used in this analysis are listed in Table 1.

Table 1. Description of clinical variables used in this study, along with variable type and name.

Clinical Variable Name	Description
anye4	APOE4 ε4 allele carrier
age_death	Age of death
educ	Years of education
bmi	Body-mass index
msex	Patient sex
pmi	Post-mortem interval until brain sample was collected
braaksc	Braak score; measure of neurofibrillary tangle severity
ceradsc	CERAD score; measure of neuritic plaque severity
cogdx	Cognitive diagnosis

2.2 Metabolic alterations are a hallmark of AD and may associate with disease stage

In this study, we identify ROS/MAP patient subgroups that differ by current disease severity solely based on their metabolic profiles in neurologically relevant regions. We utilize metabolic measurements because of evidence establishing the widespread metabolic alterations that occur in AD. A panel of 8-9

plasma metabolites predicted the risk of AD and vascular AD onset (Qiang et al., 2024). 33 blood metabolites of different sub-molecular classes were significantly dysregulated when comparing dementia patients to cognitively normal ones (Teruya et al., 2021). Most importantly, metabolites involved in cholesterol metabolism, osmoregulation, and glycolysis have been found to be significantly associated with AD traits in the ROS/MAP cohort (Batra et al., 2023). The analysis of metabolite and AD stage interplay in this study was limited. We aim to fill this knowledge gap with a comprehensive investigation using AutoSGI as a basis.

2.3 Distinct feature sets of metabolites result in distinct patient subgroup hierarchies

We first analyzed whether feature selection was needed to identify multi-scale subgroups in the context of AD. We find that the distribution of metabolite-metabolite correlations derived from ROS/MAP patient measurements has a mean of 0.0037 and a standard deviation of 0.21 (**Fig 1a**). Because the distribution of correlations are closely bound near 0, the metabolites are generally independent of each other. This means there is high heterogeneity within the biomolecules being measured by ROS/MAP, and certain metabolites will not contribute information that is relevant for subgroup identification. Further, it is also likely that different metabolite sets will also have unique biological information that impacts clustering. We confirm this hypothesis by generating 1000 random metabolite sets. For each feature set, we cluster patients using only those features, in a similar fashion as SGI and AutoSGI. We thus produce 1000 total patient hierarchies. As expected, when measuring their pairwise similarity with cophenetic correlation, there is only a moderate mean correlation of 0.41 and standard deviation of 0.081 (**Fig 1b**). Deliberate selection of clinically relevant feature sets to appropriately is therefore crucial for appropriate stratification of patients into subgroups.

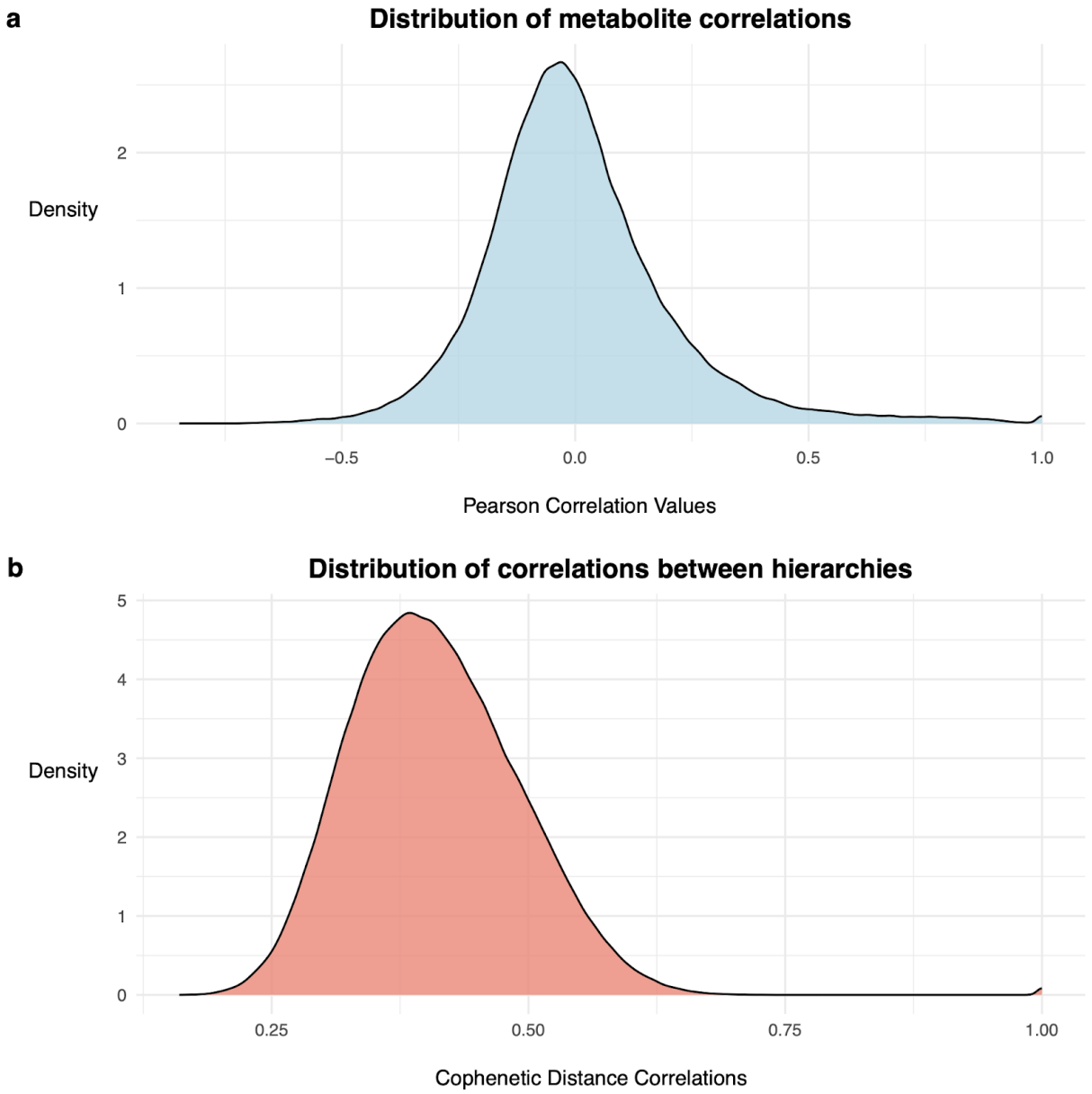


Figure 1: Clustering of heterogeneous metabolites sets results in unique patient hierarchies. a. Density plot of metabolite-metabolite correlations using ROS/MAP patient data. Range is -1 to 1, which is the Pearson correlation range. **b.** Density plot of correlations between different multi-scale patient hierarchies identified with hierarchical clustering of random metabolite feature sets. Range is 0 to 1, which is the normal cophenetic correlation range.

2.4 Subgroups in distinct AD stages can be retrieved using AutoSGI

We applied AutoSGI to the ROS/MAP cohort with the aim of finding multi-scale subgroups that are in different stages of AD. The full methodology of AutoSGI is important to understand the results, and is described in **Methods 3.3**. Essentially, AutoSGI finds different sets of metabolites to hierarchically

cluster the patients into several hierarchies. In a hierarchy driven by a specific feature set, at every branch point in that hierarchy, patients are stratified into two smaller subgroups. AutoSGI finds whether this split is clinically relevant by statistically testing the difference in disease state indicators between the two subgroups. The hierarchies that best split patients into subgroups in unique disease stages are analyzed to determine how they characterize AD severity.

We use three common AD biomarkers to measure the disease state a patient is in. These clinical variables, occasionally referred to as outcomes, are cognitive diagnosis (cogdx), Braak score (braaksc), and CERAD score (ceradsc). We chose these outcomes as they capture both physical neuropathology (braaksc with tau, ceradsc with amyloid) and cognitive function (cogdx) of a patient. We also include all the covariates as outcomes in the analysis to see if it is the driving factor for the disease stage outcome differences.

We found a total of 666 potentially relevant patient hierarchies, of which 87 had at least one subgroup pair which were significantly different (FDR corrected p-value < 0.05) for an outcome. We manually reviewed the remaining hierarchies and analyzed two of the most promising ones.

Higher levels of aromatic amino acids are associated with advanced Alzheimer's disease stages. The first hierarchy was clustered based on three metabolites, tryptophan, phenylalanine, and tyrosine. This hierarchy split patients at the top into two subgroups, labeled 2 and 3 (**Figure 2a**). These subgroups had significantly different levels of all three disease stage outcomes. In general, subgroup 3 contains healthier individuals in the earlier stages of AD. For example, utilizing the CERAD measure of neuritic plaques as a diagnostic measure, twice as many patients in subgroup 2 compared to subgroup 3 had a definite AD diagnosis (score of 4). Conversely, more than two times the patients in subgroup 3 had a definite "no AD" diagnosis (score of 1) than in subgroup 2 (**Figure 2d**). We observe similar results for Braak score and cognitive diagnosis (**Figure 2b-c**). Note we reverse CERAD scores in this analysis so higher scores indicate more AD progression.

Plotting the metabolic profiles of the patients below the hierarchy, we then noticed higher levels of all three metabolites in subgroup 2 compared to subgroup 3. We performed two-sample t-tests on all the

metabolite levels and confirmed that the subgroup differences were in fact significant ($p < 0.05$). Exact *p*-values for these tests are recorded in **Supplementary Table 3**; boxplots of the metabolites separated by subgroup are displayed in **Figures 2e-g.** We subsequently looked for biological mechanisms these metabolites have been reported to be involved in, since our results point to their higher activity driving a progressed AD state. First of all, phenylalanine, tyrosine, and tryptophan were found to be the only members of the aromatic amino acid group and thus have closely related functions. Specifically, in the context of neurological dysregulation, phenylalanine and tryptophan have been upregulated in human and mouse brains with AD (Nilsen et al., 2014; Xu et al., 2016). It has also been suggested that tryptophan degradation is part of an alternative pathway that forms quinolinic acid, linked to neurodegeneration or oxidative damage in AD patients (Griffin & Bradshaw, 2017). Both findings are supported by the fact that the three aromatic amino acids are able to find patient subgroups with differing disease stages. Moreover, tyrosine has been studied to a considerably smaller degree than the other two metabolites, yet remains informative here; thus, we implicate it as a potential focus for future investigations.

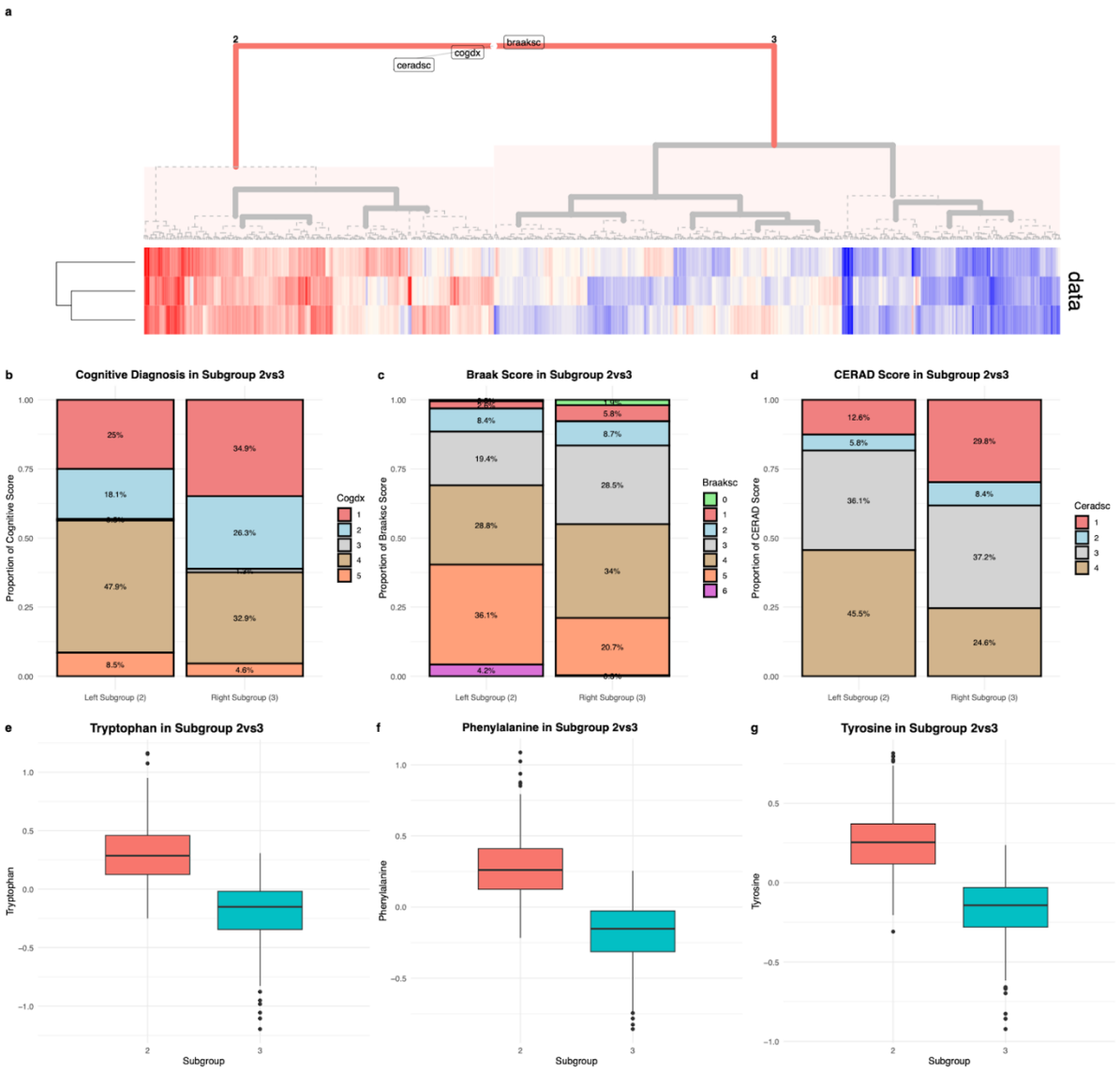


Figure 2: AutoSGI hierarchy 1 stratifies patients into two distinct subgroups in different stages of cognitive and neuropathological health. a. Hierarchy depicting patient subgroups (2 and 3). Split of patients into subgroups denoted by the red branch, where numbers on either side are the identification numbers of the subgroups. Outcomes marked at the branch split point have significant differences in patient subgroups on either side. Heatmap at bottom of plot represents the metabolomics measurements of the clustered patients. **b-d.** Boxplots of subgroups 2 and 3 for each ordinal disease stage outcome measure (higher scores correspond to later AD stage characteristics). The left subgroup on the first branch split, subgroup 2, is thus less healthy than subgroup 3. **e-g**. Numeric boxplots plotting the metabolites used for clustering for the two subgroups, which also have significantly higher levels in subgroup 2.

Higher levels of neurotoxic guanidinoacetate and excitatory neurotransmitter glutamate are associated with advanced AD stages. The second hierarchy was clustered with a panel of seven metabolites identified by AutoSGI: N-acetylglycine, 2-aminoadipate, guanidinoacetate, glutamate, glycerophosphoethanolamine, X-25020, and glycerophosphorylcholine. All patients were placed in subgroup 2 or 3, which had significant differences in cognitive diagnosis (**Fig 3a-b**). Subgroup 2 in particular had lower proportions of late-stage AD signs. Around 75% of subgroup 2 patients had no cognitive impairment (score of 1) or only mild cognitive impairment (score of 2-3). In contrast, the right subgroup in the hierarchy, subgroup 3, was generally less healthy with less than half of patients having cognitive diagnosis of 3 and lower. However, subgroup 3 was also stratified further into subgroup 4 and 5 shown by the blue branch split (**Fig 3a**). Notably, we observed that patients in subgroup 3 specifically could be stratified by outcomes both neuropathological (Braak score) and cognitive (cognitive diagnosis). Though subgroup 3 originally indicated it had unhealthier patients overall, at a more granular scale, AutoSGI was able to find it had comparatively healthier and unhealthier patients within it. In particular, subgroup 4 had more patients in earlier AD stages and subgroup 5 had more patients in later AD stages (**Fig 3c-d**).

We then analyzed metabolic differences in a similar fashion as with hierarchy 1; all the seven metabolites had significantly higher levels in subgroup 3 (**Supplementary Table 4**), indicating their presence may contribute to unhealthier neurological states. For brevity, we only show boxplots for glutamate and guanidinoacetate levels **(Fig 3a-c, Fig 3e-f)**. The same observation made for the initial stratification holds for the second split of subgroup 3 into subgroup 4 and 5: extremely high metabolic levels were present in the unhealthier subgroup 4 and moderately high metabolic levels were present in the healthier subgroup 5 (**Fig 3g-h**). Interestingly, subgroup 5 is healthier than subgroup 4, but is comparatively less healthy compared to subgroup 2 at the top of the hierarchy. This means the continuous spectrum of low metabolic levels to high metabolic levels generally correspond to the most healthy to least healthy patients. With a single-level subgroup identification method, this would not have been identified, highlighting the unique ability of AutoSGI to successfully find patient stratifications at additional scales.

Lastly, we examined the feature set for this hierarchy consisting of the seven metabolites previously mentioned. These metabolites belong to several functional classes, including glycerophosphocholine and glycerophosphoethanolamine, which play a major role in phospholipid metabolism and had higher concentrations in AD brain regions up to 150% and 52%, respectively, corroborating our results (Bluszt-ajn et al., 1990). Some were also previously determined to individually associate with AD; 2-amino-adipate had significantly different levels in autosomal dominant AD and sporadic AD brain samples (Novotny et al., 2023). Finally, interactions between metabolites were discovered earlier. For example, the uptake of glutamate into astrocytes and other cells within the brain has been found to be reduced by guanidinoacetate (Marques et al., 2019). This may seem to contradict the high levels of both metabolites in ROS/MAP patients with progressed AD outcomes. However, this could be attributed to the bulk metabolomics used in our study which lacks cellular resolution. Astrocyte-specific glutamate levels may be lower but in this case are not captured as the signal from all the cells are measured together. Single-cell metabolomics can be used in future studies to elucidate this relationship and the cell-type specific abundance of these metabolites.

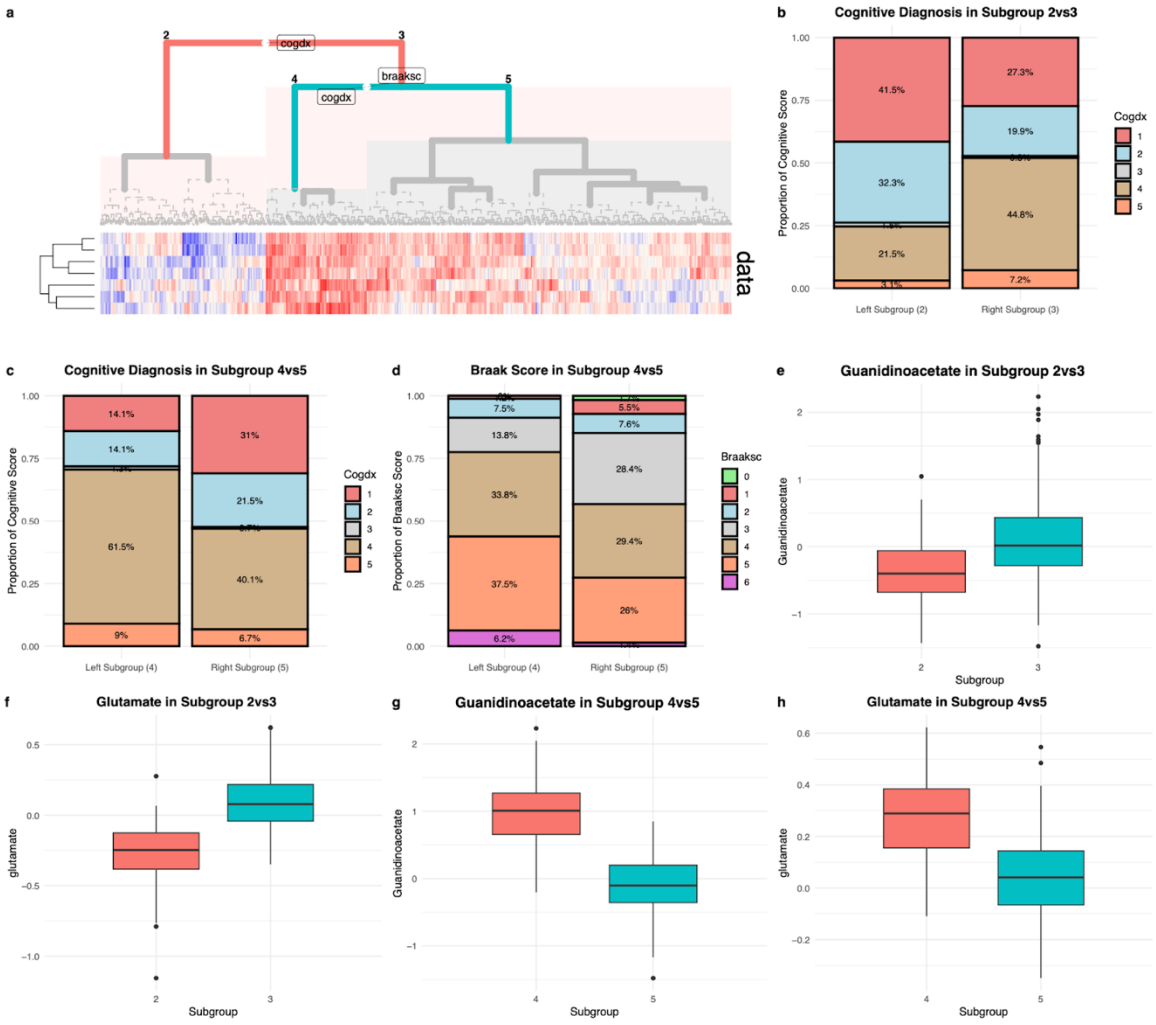


Figure 3: AutoSGI hierarchy 2 stratifies patients into subgroups at different scales, each in unique stages of neuropathological and cognitive health. a. Hierarchy depicting patient subgroups at various scales. This includes subgroups 2, 3, 4, and 5, where the latter two originate from subgroup 3 (or more granularly stratified). Subgroup splits of patients denoted by the red and blue branches, with numbers on either side of the branches indicating the subgroup identification number. Clinical outcomes marked at the branch split point to significant differences in that outcome for the patient subgroups on either side.

As discussed earlier in **Methods 3.3**, when statistically testing for differences between subgroups we do not correct for most covariates since they were not significantly correlated with any single metabolite. Only the post-mortem interval (pmi) till patient brain sample collection is regressed out of the metabolomics data as it was significantly associated. Yet, none of the covariates were significantly different between the subgroup pairs with distinct AD disease stages. In fact, they were usually highly

insignificant. Taken together, AutoSGI is able to find feature subsets that utilize underlying metabolic information to directly differentiate subgroups by disease stage, rather than by indirect association.

We list the FDR-corrected *p*-values calculated by the statistical tests between all analyzed subgroups and outcomes in **Supplementary Tables 1-2**.

2.5 AutoSGI stratified patients into robust subgroups with clinical and biological relevance

The patient subgroups identified in hierarchies 1 and 2 by AutoSGI had potentially meaningful clinical differences, but the robustness of the results were not yet assessed. Most metabolites that were utilized for patient clustering were not measured in other large-scale AD studies like ADNI. Therefore, to estimate robustness, we relied on calculating stability metrics within the ROS/MAP cohort itself. For each analyzed subgroup, the dataset used by AutoSGI to find it was altered 1000 times via resampling. The proportion of the original subgroup that was retained every iteration was computed with Jaccard similarity (Hennig, 2007).

In the first hierarchy, subgroups 2 and 3 had average stability scores of 0.79 and 0.7, respectively. Though subgroup 3 has a large standard deviation making it less stable, this means it is likely both subgroups are well defined (Hennig, 2007; Henning, 2008). Strikingly, the stability of the subgroups in the second hierarchy parallel the clinical homogeneity they contain (**Fig 4a-b**). Subgroup 2 has a moderately high mean stability of 0.67, and we observe that at smaller scales there are no more relevant clinical differences between patients. Subgroup 3 on the other hand has a very low mean of 0.51, indicating the metabolic profiles of its patients are poorly resolved against the other subgroup at the same scale, subgroup 2. Notably however, when an additional stratification on subgroup 3 is performed by AutoSGI, the stability of the new subgroups, 4 and 5, dramatically increase to 0.64 and 0.68 **(Fig 4c-f)**; this same stratification also revealed new disease stage differences within subgroup 3. Consequently, these results support the argument for separating patients in a robust, biologically intrinsic manner to delineate phenotypic differences.

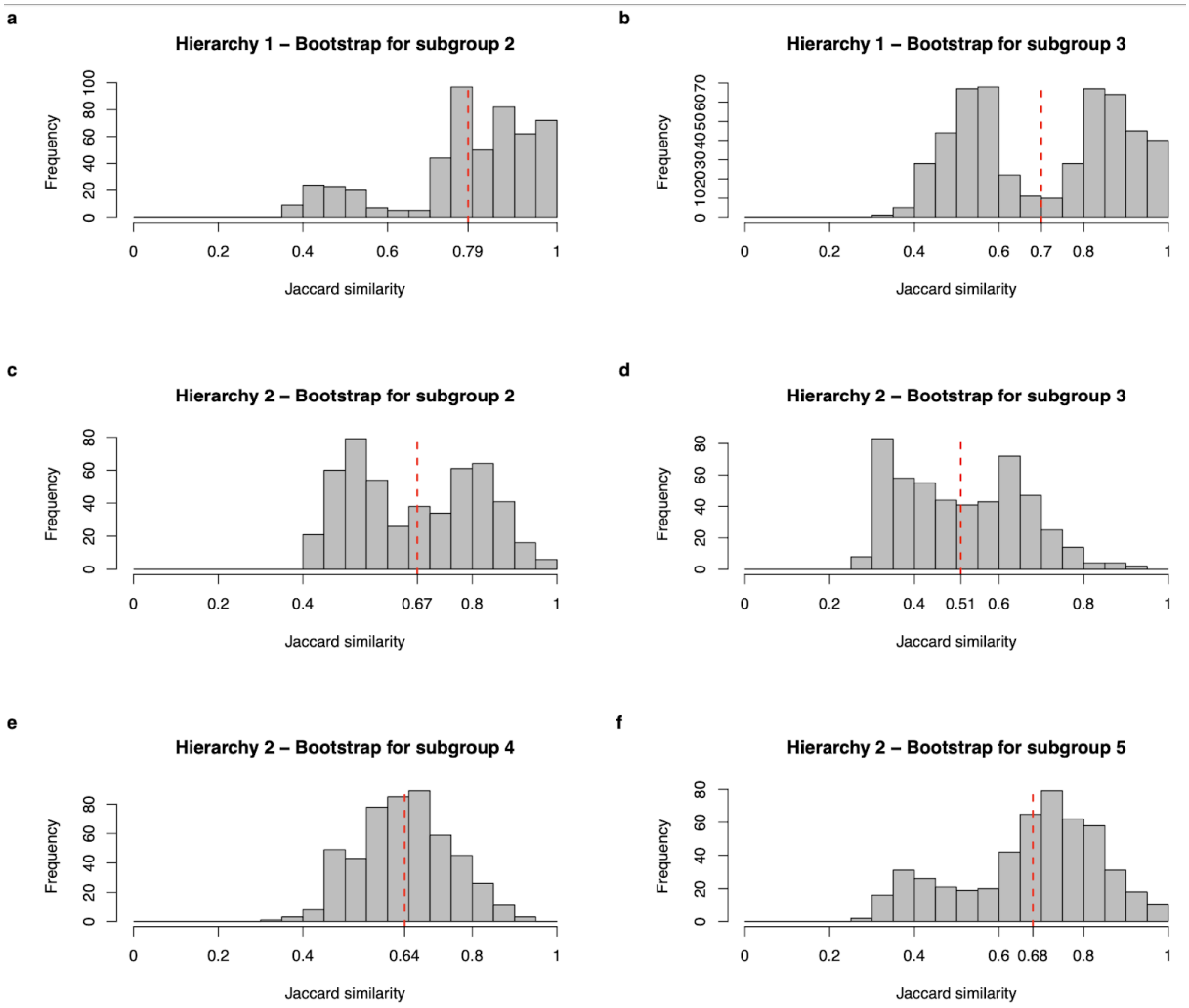
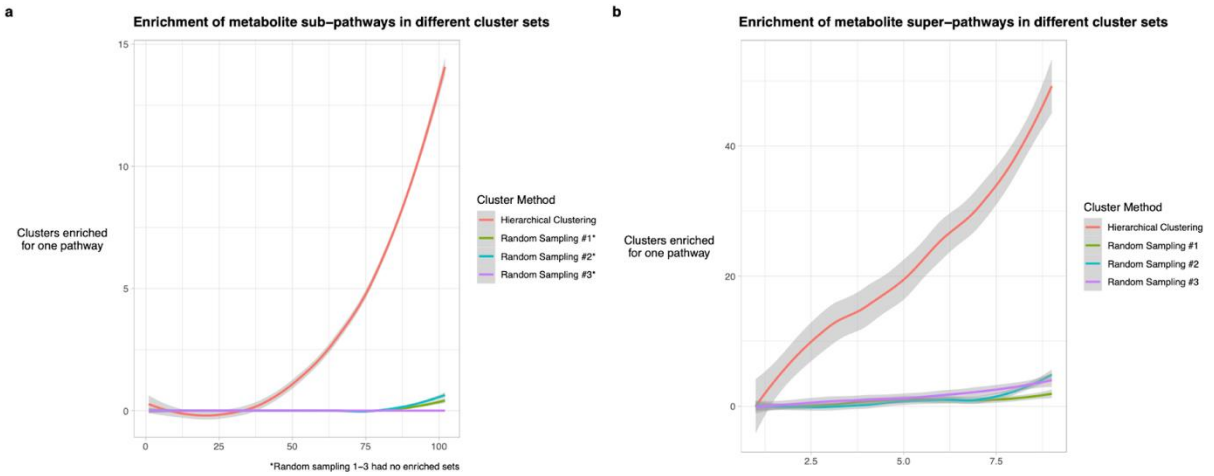


Figure 4: **Both selected AutoSGI hierarchies able to stratify patients into stable clusters. a-b**. Jaccard similarity (stability metric) of patient subgroups 2 and 3 in AutoSGI hierarchy 1 across 1000 bootstraps and plotted as a histogram. Means are indicated with dashed red lines. **c-f**. Jaccard similarity of subgroups in AutoSGI hierarchy 2, plotted in a similar fashion as a-b. Plots c-d focus on patient subgroups 2 and 3. Plots e-f similarly focus on patient subgroups 4 and 5, which are stratified from subgroup 3.

2.6 AutoSGI identifies feature sets which are enriched at various pathway levels

We also demonstrate that AutoSGI patient hierarchies are interpretable via the feature sets that define them. Specifically, we can examine what biological pathways are significantly enriched in these feature sets, which consequently allows for a mechanistic investigation of AD. We first determined enrichment of metabolite super-pathways and sub-pathways in the AutoSGI feature sets via Fisher's exact test. We then compare the enrichment levels obtained to ones resulting from 3 random ways of sampling features

(**Methods 3.4**). We found that after correction, no feature sets identified by random sampling had enrichment in sub-pathways. On the other hand, or the top 25% most overrepresented sub-pathways, AutoSGI had at least 5 feature sets enriched in those pathways (**Fig 5a**). AutoSGI also had superior biological alignment for super-pathways (**Fig 5b**). Moreover, we noticed that pathways which AutoSGI feature sets are enriched in have been found to be involved in important AD traits outside of ROS/MAP; this includes pathways like hexoslyceramides (Akyol et al., 2021; Dehghan et al., 2022) and long chain saturated fatty acids (Fan et al., 2023), which are the most highly enriched sub-pathways **(Fig 5c)**.



c

Most Over–Represented Sub Pathways
Hexosylceramides (HCER)
Long Chain Saturated Fatty Acid
Long Chain Polyunsaturated Fatty Acid (n3 and n6)
Sphingomyelins
Lysophospholipid
Long Chain Monounsaturated Fatty Acid
Leucine, Isoleucine and Valine Metabolism
Diacylglycerol
Fructose, Mannose and Galactose Metabolism
Fatty Acid Metabolism (Acyl Carnitine, Polyunsaturated)

d

Most Over–Represented Super Pathways
Lipid
Amino Acid
Carbohydrate
Nucleotide
Cofactors and Vitamins
Peptide
Energy
Xenobiotics
NA

Figure 5: AutoSGI feature sets are significantly enriched in biological pathways at several scales. **a-b**. Number of feature sets enriched in each sub-pathway/super-pathway for each selection method, from least to greatest. The scatterplot was smoothed into a line using locally estimated scatterplot smoothing (LOESS). **c-d**. Sub-pathways and super-pathways most represented in AutoSGI feature sets (top 10 and top 9, respectively).

2.6 AutoSGI outperforms baseline multi-scale clustering

Finally, we measure the performance of AutoSGI against baseline approaches on a quantitative level. We propose a new metric to measure efficacy of an individual patient hierarchy across several outcomes inspired by similar scores for single outcome and single-scale subgroup identification (Rappoport & Shamir, 2019; Xie et al., 2024). Briefly, the score combines the strength of the clinical difference between each of the smaller two subgroups that are stratified in the patient hierarchy, across all stratifications (**Methods 3.7**). Therefore, since AutoSGI finds 666 hierarchies, the same number of scores will be calculated. We compare the top 25 scores of AutoSGI hierarchies with 3 baseline methods, which randomly select feature sets in different ways (**Methods 3.7**). AutoSGI significantly outperforms the baseline approaches numerically, with the highest average score.

It is important to note that AutoSGI best performs in the first 15 hierarchies before the scores decline, whereas sampling from all possible feature sets (Random Sampling 3) had a relatively constant score for each hierarchy, eventually performing better in the last 10 hierarchies. This observation holds true statistically as well. The top 25 AutoSGI scores are significantly greater compared to the first and second random sampling method ($p < 0.05$, Mann-Whitney U test). Only the top 10 AutoSGI scores retain significance against the third random sampling method. This may be due to the similarity-based feature selection that AutoSGI conducts. Here, the best possible scale for biological information to arise is interrogated at every level. Therefore, a minority of cases will have this natural separation where there is an improvement in patient clustering. However, this does not diminish from AutoSGI's overall success, since the best performing hierarchies (and subgroups) are the ones important in a clinical setting.

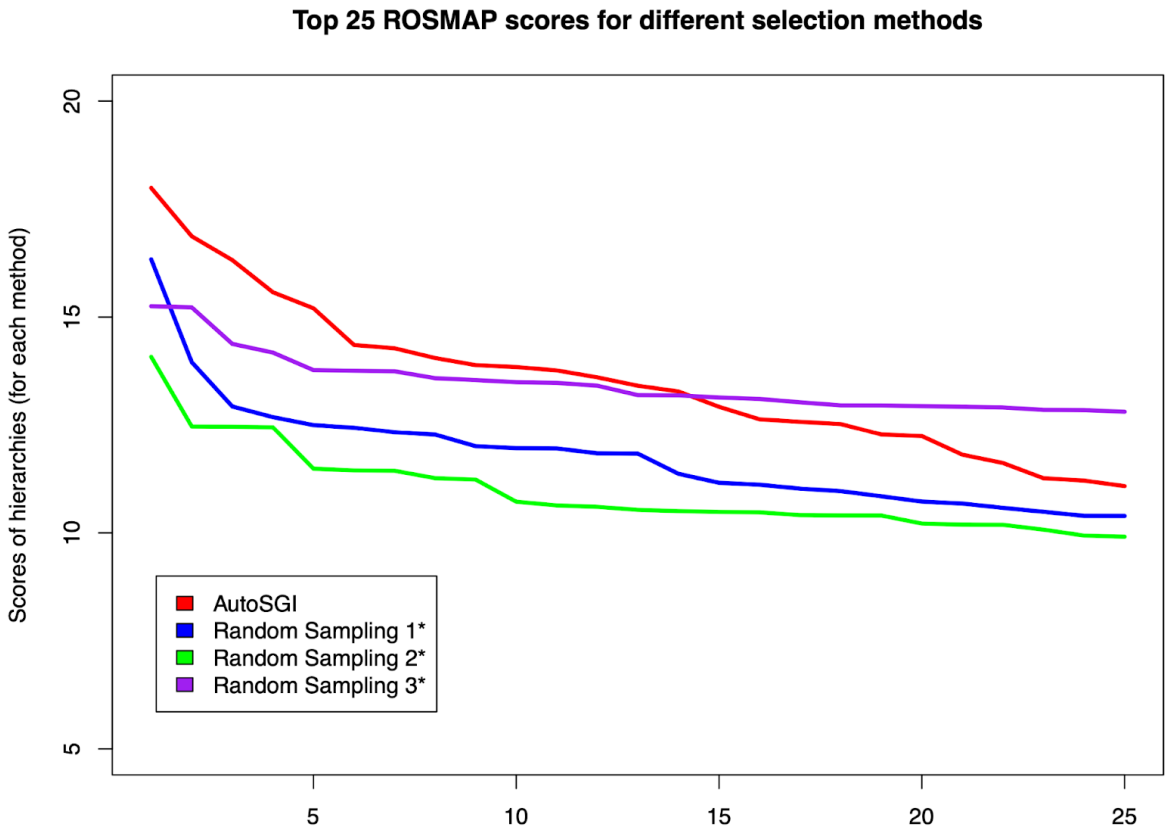


Figure 6: **AutoSGI outperforms other methods for stratifying patients in a clinically relevant manner.** Top 25 scores for all 4 methods (AutoSGI and 3 random approaches), sorted in reverse order and plotted as a line graph.

2.7. **Conclusion & Future Directions**

In this study, we developed AutoSGI, a novel subgroup identification tool that places 500 AD patients in the ROS/MAP cohort into subgroups with different disease stages. We evaluated the clinical utility and robustness of AutoSGI with several analyses. Here, we discuss evidence to support our argument that AutoSGI can effectively leverage metabolic levels of postmortem brain samples to identify subgroups at multiple scales.

We analyzed two hierarchies produced by AutoSGI. The first hierarchy split patients into two subgroups, labeled 2 and 3. Subgroup 2 patients had significantly higher AD progression, with more severe neuropathological and cognitive indicators. The second hierarchy split patients at several scales; there was an initial split of ROS/MAP patients into subgroups 2 and 3, along with an additional stratification of subgroup 3 into subgroups 4 and 5. We observed that patients in subgroup 2 had significantly less AD

severity in this case. Although other clinical distinctions were unable to be made for these patients, subgroup 3 could be divided into unhealthier and unhealthier subgroups (subgroup 4 and 5, respectively). Given that subgroup 4 still exhibited more advanced phenotypes than subgroup 2, it can be inferred that AutoSGI identified low, intermediate, and high severity disease states within each subgroup. When analyzing the stability of these three subgroups, it was also much higher compared to the two subgroups split at the first level of the hierarchy. This supports the idea that by splitting patients at every part of a hierarchy, we are able to find natural points of strong biological separation (or stability), ultimately leading to phenotypic separation as well.

We found that metabolite feature selection via hierarchical clustering results in alignment with known biological pathways. In particular, many metabolites were enriched at a sub pathway level and a super pathway level. Pathways which were overrepresented included structurally defined groups, such as long chain saturated fatty acids and long chain polyunsaturated fatty acids, and functionally defined groups like metabolites involved in fructose, mannose, and galactose metabolism. The metabolites used for clustering in the 2 hierarchies followed this trend. The first hierarchy was structurally based as it used the three aromatic acids, tyrosine, phenylalanine, and tryptophan. These aromatic amino acids are associated with advanced AD stages. On the other hand, the seven metabolites in the second pathway were more functionally defined, composed of several different metabolic pathways. Glycerophosphocholine and glycerophosphoethanolamine are two of these metabolites and are phospholipid catabolic intermediates (Blusztajn et al., 1990). Guanidinoacetate, which is neurotoxic and glutamate, an excitatory neurotransmitter, interact and are associated with AD severity as well. However, there was still a presence of structural similarity; another two of the metabolites, glutamate and 2-aminoadipate, are homologs (Young & Ajami, 2000). Taken together, AutoSGI was able to select metabolites with more concentrated biological information compared to the heterogeneity we initially observed. This helped improve performance of the method in general, as seen by our comparison against baseline methods.

There are several limitations of this study which we hope to address in future work. First, we only identify patient subgroups using brain metabolomics information. However, this may not provide a complete picture of the complex biomolecular systems in AD. There is an increasing amount of evidence

that multi-omics analyses better characterize neurological mechanisms that lead to AD and affect its severity. For instance, a transcriptomics and proteomics based approach in human cortical samples defined a potential timeline of "transcriptional, translational and post-translational alterations" relevant to neurofibrillary pathology (Marttinen et al., 2019). Similarly, a cohort of patients had multi-omic profiling of their cerebrospinal fluid, or CSF (Clark et al., 2021). It was found that multiple lipids, proteins, and metabolites contributed to inter-patient variance, showing the importance of including all biological modalities. AutoSGI lends itself well to these kinds of studies since it can be extended to multi-omics datasets with integrative hierarchical clustering (Schweickart et al., 2024; X. Zhang et al., 2022) Applying AutoSGI to other omics modalities in the ROS/MAP cohort like proteomics, epigenetics, and transcriptomics, as well as metabolomics, would be one natural direction to explore.

Second, the subgroups we found were part of a single-cohort study, albeit a large one. As a result, determining if the feature sets used for clustering are effective on other cohorts would increase the generalizability of our results. We did not conduct such an analysis in this paper because metabolites measured in ROS/MAP were usually not measured in other common AD cohorts like ADNI (St John-Williams et al., 2017). A recent set of 342 brain samples from the Mayo Clinic Brain Bank may allow for such an analysis since it measures a group of metabolites that overlap more with the ROS/MAP data. Alternatively, we are currently testing the conversion of individual metabolite measurements in the ROS/MAP cohort to metabolic sub-pathway enrichment scores. Feature sets with this transformed data would theoretically consist of sets of pathways that interact or are similar. This approach is promising for three reasons. Crucially, sub-pathway annotations can be standardized across cohorts more easily compared to individual metabolite names. Pathway enrichment scores have also improved cohort transferability in other scenarios because of the use of multiple features to stabilize signals (Blum et al., 2022; Tang et al., 2022). Furthermore, the interpretability of the results could potentially increase since we use pathway enrichment scores for features; any subgroups identified with phenotypic differences can then be explained by these pathways, which are constructed using prior biological knowledge.

Finally, AutoSGI produced many potentially relevant hierarchies but required manual selection to identify the ones promising enough for analysis. We did create a metric to score random patient hierarchies

against AutoSGI ones, but we found that it is useful for general comparison between sets of hierarchies, and not individual ones. Ranking hierarchies requires consolidation of many factors like the number of significant splits, the separability of metabolite levels between subgroups, and the levels of subgroup stratification, which is currently not feasible. We rely on an initial filtering like the one used in this study (leaving 87 hierarchies out of a total of 666) to reduce the number of hierarchies that need to be manually considered instead.

In conclusion, we present the first metabolomics-guided investigation that stratifies AD patients into multiple levels of subgroups using data-driven feature selection. At the core of our analyses was the development of the AutoSGI framework. We demonstrated AutoSGI's unique ability to simultaneously (1) assess a patient's disease state for clinical intervention and (2) find metabolic signatures that provide insight into AD mechanisms. Future work will focus on multi-omic and multi-cohort applications. Overall, this research will contribute to an evolving knowledge base regarding AD etiology, directing individualized treatment in the future.

3. Methods

3.1 Data Preprocessing

ROS/MAP metabolomics data was collected as described in **Results 2.1**. We preprocessed the raw untargeted metabolomics measurements in a similar manner as Batra et al. in the original analysis of the data (Batra et al., 2023). Briefly, 667 metabolites with at least 75% values present (measured in 375 patients or more) were kept. The metabolite variance was minimized using probabilistic quotient normalization and transformed with a logarithmic scale. Outliers were then detected and a pipeline for imputation was performed. Finally, the distributions of metabolite values were brought to a comparable scale with standardization.

3.2 Determining importance of feature sets for clustering

We first measured the diversity of the metabolites measured in the ROS/MAP dataset by calculating all metabolite-metabolite correlations (with Pearson's correlation coefficient). Smaller correlation magnitude as a whole would indicate that the metabolites were biologically heterogeneous. This would presumably also impact multi-scale patient clustering, the focus in this study. We thus assessed whether using a subset of metabolites to create patient hierarchies would change the multi-scale clusters produced. Accordingly, we randomly sampled the 667 ROS/MAP metabolites for 1000 repetitions, generating a new feature set each time. Then, we performed hierarchical clustering with Euclidean distance as a dissimilarity metric and Ward's method as a linkage approach. Between each pair of these patient hierarchies, we calculated the cophenetic correlation. This metric is a common quantitative measure of how similar two hierarchies are by comparing where each pair of patients are merged in either one. The distribution of cophenetic correlations across all pairs of metabolite set derived hierarchies would indicate appropriate metabolite selection for patient clustering is important when the mean is far from 1 or the variance is large.

3.3 AutoSGI

Here, we describe the AutoSGI algorithm, which is the central contribution of this paper. An overview of the AutoSGI pipeline is displayed below (**Fig 7**). As mentioned previously, we extend SGI

(Buyukozkan et al., 2022), our original approach for subgroup identification, to build the AutoSGI framework (**Introduction**). In particular, SGI would find patient subgroups in the ROS/MAP cohort using hierarchical clustering of all their feature measurements. But, this could lead to non-clinically relevant subgroups due to the presence of features which are not informative of the disease. AutoSGI addresses this problem by subsetting the features and testing each set to identify feature sets that inform AD severity. Specifically, we argue that features with similar measurements in the patient cohort will be biologically similar and are more likely to jointly influence related phenotypes. Multiple studies have observed this general idea (Schweickart et al., 2024; Qiu et al., 2023). With this motivation in mind, AutoSGI performs secondary hierarchical clustering on the 667 features themselves. The clusters that form are cut at every level, producing 666 feature sets.

For each of these identified feature sets, AutoSGI applies a similar methodology as SGI for multi-scale subgroup identification. In brief, for a specific feature set, AutoSGI hierarchically clusters the patients with Euclidean distance and Ward linkage. We refer to the patient clusters altogether as a hierarchy. Patients in a larger subgroup in a hierarchy are split into two smaller ones at every branch point; these two smaller subgroups are called subgroup pairs. We test these subgroups for differences in Braak scores, CERAD scores and clinical diagnosis outcomes (disease stage indicators) using ordinal regression from the *rms* package. Since there are many statistical tests being conducted, we apply false discovery rate (FDR)-based multiple testing correction on the resulting *p*-values from each test across all the hierarchies.

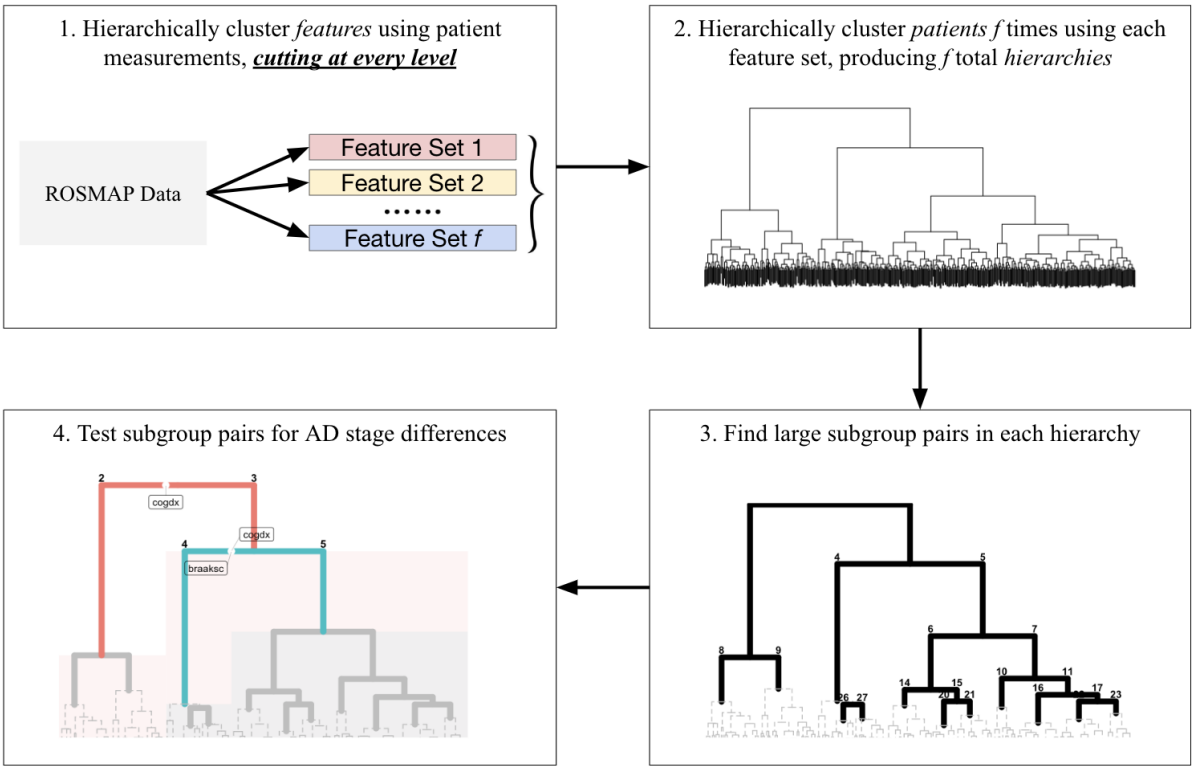


Figure 7: **Overview of the AutoSGI pipeline.** AutoSGI follows the 4 major steps displayed in this figure to ultimately produce the ROS/MAP patient hierarchies. These hierarchies contain subgroup pairs at multiple scales that are tested for significant differences in disease stage biomarkers, prompting further analysis in some cases.

3.4 Examining disease stage differences in subgroups identified by AutoSGI

We used AutoSGI in the ROS/MAP cohort to find 666 distinct patient hierarchies each with multi-scale subgroups. We considered 87 hierarchies which split the patients at the first level of the tree into two subgroups with significant differences in any disease stage outcome, filtering out the rest. This was a stringent enough threshold to allow for evaluation of the remaining hierarchies within a reasonable amount of time. However, the rigor of the rule being used for filtering should be adjusted based on the number of biomolecular features being measured for each patient, since this will correspond with the number of hierarchies AutoSGI identifies.

We selected 2 final hierarchies for downstream analysis. In each hierarchy, the outcome differences between all relevant subgroup pairs were plotted as an ordinal boxplot. The metabolomics measurements that drove that clustering were plotted as a numeric boxplot.

3.5 Evaluating stability of patient subgroups

The stability of the subgroups identified by AutoSGI in the two hierarchies we chose was examined next. We define a subgroup *S* as more stable if it is resilient to perturbations in the ROS/MAP measurements. That is, the overall subgroup structure should appear in a new hierarchy constructed from data without major changes. We utilize a previously proposed algorithm for this purpose as it is applicable for general cluster robustness scoring (Hennig, 2007; Hennig, 2008). The R package *fpc* implements this procedure (*clusterboot*) and was used in this analysis (*R: Clusterwise Cluster Stability Assessment by Resampling*, 2024.). Essentially, the score of *S* describing its stability is calculated by first resampling ROS/MAP metabolomics measurements without replacement for 1000 bootstrap replications. The hierarchical clustering process that AutoSGI performs is conducted for each iteration. The score is then the maximum Jaccard index between the subgroups of the new tree at the same scale as *S* in the original tree, and *S* itself. We plot the stability scores over the bootstrapped replications for the subgroups deemed relevant (**Methods 3.4**) with a histogram and compute their individual means.

3.6 Measuring biological alignment of AutoSGI feature sets

AutoSGI performs dual hierarchical clustering for both features and patients as detailed in section 2.3. We wanted to see if the 666 candidate feature sets AutoSGI identifies in this way aligns with prior biological knowledge, specifically metabolic pathways. To this end, we collected available metabolite annotations at a super-pathway (9 total) and sub-pathway level (102 total). For every pathway, we performed Fisher's exact test on each feature set to determine if there was significant enrichment in that pathway. Bonferroni correction was used to account for multiple tests. Finally, to confirm that any enrichment is not due to the size of the feature sets chosen but rather the similarity-based clustering itself, we compared the levels of pathway overrepresentation to 3 random selection methods. These methods also select 666 feature sets, either by (1) randomly sampling sizes from the distribution of AutoSGI feature set sizes, and then selecting feature sets with those fixed sizes; (2) sampling sizes in correspondence (one to one) with the AutoSGI feature set sizes and then selecting feature sets with that those sizes; or (3) randomly selecting a feature set from all possible feature sets.

3.7 Quantitatively benchmarking AutoSGI

To conclude the analyses, we developed a quantitative metric to benchmark the performance of AutoSGI in identifying clinically relevant subgroups. Previous scores have been developed for single-scale subgroup identification for a single outcome using the *p*-value from a two-sample statistical test (Rappoport & Shamir, 2019; Xie et al., 2024). We extend this metric to scoring multi-scale subgroup identification over several outcomes by applying it across every subgroup pair and clinical outcome.

Specifically, the score defined for a single multi-scale subgroup identification result follows:

$$\sum_{s=1}^{n} \sum_{c=1}^{k} -\log_{10}(p_{s,c}), \text{ over all } n \text{ subgroup pairs and } k \text{ clinical outcomes.}$$

Here, the $p_{s,c}$ refers to the corrected *p*-value obtained from testing for a difference in clinical outcome *c* on subgroup pair *s*. Note that this score is not necessarily an accurate metric to rank individual hierarchies in terms of effective subgroup identification. This is because it is hard to explicitly capture the various choices a researcher is making in selecting the most promising hierarchies. Rather, it can measure the efficacy of AutoSGI compared to baseline approaches with a global fashion.

These baseline approaches (3 in total) use the same subgroup identification scheme as AutoSGI, but with the feature sets as described in section 2.6. Thus, a total of 666 scores will be generated for each selection method. We compare the scores of the top 25 hierarchies from each approach.

4. Supplement

4.1 Supplementary Table 1. Corrected *p*-values from statistical tests for outcome differences in hierarchy 1. Significant p-values less than 0.05 are highlighted in gray.

Subgroup Pair	Clinical Variable Name	FDR corrected *p*-value
2vs3	anye4	1.00000000000
2vs3	cogdx	0.02648850202
2vs3	age_death	0.97683522830
2vs3	educ	1.00000000000
2vs3	msex	0.11863005582
2vs3	braaksc	0.00153476574
2vs3	ceradsc	0.00007591131
2vs3	pmi	1.00000000000
2vs3	bmi	1.00000000000

4.2 Supplementary Table 2. Corrected *p*-values from statistical tests for outcome differences in hierarchy 2. Significant p-values less than 0.05 are highlighted in gray.

Subgroup Pair	Clinical Variable Name	FDR corrected *p*-value
2vs3	anye4	1.000000000
2vs3	cogdx	0.001344414
2vs3	age_death	0.972782985
2vs3	educ	1.000000000
2vs3	msex	1.000000000
2vs3	braaksc	0.444417072
2vs3	ceradsc	0.052884063
2vs3	pmi	1.000000000
2vs3	bmi	1.000000000
4vs5	anye4	1.000000000
4vs5	cogdx	0.04738144
4vs5	age_death	0.94826898
4vs5	educ	1.00000000
4vs5	msex	0.87429648
4vs5	braaksc	0.02310150
4vs5	ceradsc	0.06850478
4vs5	pmi	1.00000000
4vs5	bmi	1.00000000

4.3 Supplementary Table 3. *p*-values from statistical tests for metabolite differences in sub-groups analyzed from hierarchy 1. Significant p-values less than 0.05 are highlighted in gray.

Subgroup Pair	Metabolite Variable Name	*p*-value
2vs3	phenylalanine	6.708134e-74
2vs3	tryptophan	1.643506e-66
2vs3	tyrosine	2.359875e-76

4.4 Supplementary Table 4. *p*-values from statistical tests for metabolite differences in sub-groups analyzed from hierarchy 2. Significant p-values less than 0.05 are highlighted in gray.

Subgroup Pair	Metabolite Variable Name	*p*-value
2vs3	N-acetylglycine	1.806655e-16
2vs3	2-aminoadipate	1.232717e-26
2vs3	guanidinoacetate	1.194706e-19
2vs3	glutamate	1.912826e-43
2vs3	glycerophosphoethanolamine	1.327059e-29
2vs3	X-25020	6.891636e-52
2vs3	glycerophosphorylcholine	1.194620e-34
4vs5	N-acetylglycine	3.942103e-21
4vs5	2-aminoadipate	1.142067e-29
4vs5	guanidinoacetate	3.626545e-31
4vs5	glutamate	1.633228e-22
4vs5	glycerophosphoethanolamine	8.668945e-20
4vs5	X-25020	1.028341e-11
4vs5	glycerophosphorylcholine	4.701493e-24

4.5 Supplementary Code. Please find all code needed to replicate these analyses, as well as the general AutoSGI package code, attached as a zip file with this project submission.

5. References

2020 Alzheimer's disease facts and figures. (2020). *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *16*(3), 391–460.

Akyol, S., Ugur, Z., Yilmaz, A., Ustun, I., Gorti, S. K. K., Oh, K., McGuinness, B., Passmore, P., Kehoe, P. G., Maddens, M. E., Green, B. D., & Graham, S. F. (2021). Lipid profiling of Alzheimer's disease brain highlights enrichment in glycerol(phospho)lipid, and sphingolipid metabolism. *Cells (Basel, Switzerland)*, *10*(10), 2591.

Arnold, M., Buyukozkan, M., Doraiswamy, P. M., Nho, K., Wu, T., Gudnason, V., Launer, L. J., Wang-Sattler, R., Adamski, J., De Jager, P. L., Ertekin-Taner, N., Bennett, D. A., Saykin, A. J., Peters, A., Suhre, K., Kaddurah-Daouk, R., Kastenmüller, G., & Krumsiek, J. (2024). Individual bioenergetic capacity as a potential source of resilience to Alzheimer's disease. In *Neurology* (No. medrxiv;2024.01.23.23297820v1). medRxiv. https://www.medrxiv.org/content/10.1101/2024.01.23.23297820v1.full

Avelar-Pereira, B., Belloy, M. E., O'Hara, R., Hosseini, S. M. H., & Alzheimer's Disease Neuroimaging Initiative. (2023). Decoding the heterogeneity of Alzheimer's disease diagnosis and progression using multilayer networks. *Molecular Psychiatry*, *28*(6), 2423–2432.

Batra, R., Arnold, M., Wörheide, M. A., Allen, M., Wang, X., Blach, C., Levey, A. I., Seyfried, N. T., Ertekin-Taner, N., Bennett, D. A., Kastenmüller, G., Kaddurah-Daouk, R. F., Krumsiek, J., & Alzheimer's Disease Metabolomics Consortium (ADMC). (2023). The landscape of metabolic brain alterations in Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *19*(3), 980–998.

Blum, B. C., Lin, W., Lawton, M. L., Liu, Q., Kwan, J., Turcinovic, I., Hekman, R., Hu, P., & Emili, A. (2022). Multiomic metabolic enrichment network analysis reveals metabolite-protein physical interaction subnetworks altered in cancer. *Molecular & Cellular Proteomics: MCP*, *21*(1), 100189.

Blusztajn, J. K., Lopez Gonzalez-Coviella, I., Logue, M., Growdon, J. H., & Wurtman, R. J. (1990). Levels of phospholipid catabolic intermediates, glycerophosphocholine and glycerophosphoethanolamine, are elevated in brains of Alzheimer's disease but not of Down's syndrome patients. *Brain Research*, *536*(1-2), 240–244.

Buyukozkan, M., Suhre, K., & Krumsiek, J. (2022). SGI: automatic clinical subgroup identification in omics datasets. *Bioinformatics (Oxford, England)*, *38*(2), 573–576.

Clark, C., Dayon, L., Masoodi, M., Bowman, G. L., & Popp, J. (2021). An integrative multi-omics approach reveals new central nervous system pathway alterations in Alzheimer's disease. *Alzheimer's Research & Therapy*, *13*(1). https://doi.org/10.1186/s13195-021-00814-7

Cooper, A., Doyle, O., & Bourke, A. (2021). Supervised clustering for subgroup discovery: An application to COVID-19 symptomatology. In *Communications in Computer and Information Science* (pp. 408–422). Springer International Publishing.

Dehghan, A., Pinto, R. C., Karaman, I., Huang, J., Durainayagam, B. R., Ghanbari, M., Nazeer, A., Zhong, Q., Liggi, S., Whiley, L., Mustafa, R., Kivipelto, M., Solomon, A., Ngandu, T., Kanekiyo, T., Aikawa, T., Radulescu, C. I., Barnes, S. J., Graça, G., … Elliott, P. (2022). Metabolome-wide association study on ABCA7 indicates a role of ceramide metabolism in Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(43), e2206083119.

Fan, L., Borenstein, A. R., Wang, S., Nho, K., Zhu, X., Wen, W., Huang, X., Mortimer, J. A., Shrubsole, M. J., Dai, Q., & Alzheimer's Disease Neuroimaging Initiative. (2023). Associations of circulating saturated long-chain fatty acids with risk of mild cognitive impairment and Alzheimer's disease in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort. *EBioMedicine*, *97*(104818), 104818.

GBD 2019 Dementia Forecasting Collaborators. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *The Lancet. Public Health*, *7*(2), e105–e125.

Griffin, J. W. D., & Bradshaw, P. C. (2017). Amino Acid Catabolism in Alzheimer's Disease Brain: Friend or Foe? *Oxidative Medicine and Cellular Longevity*, *2017*(1). https://doi.org/10.1155/2017/5472792

Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis*, *52*(1), 258–271.

Hennig, C. (2008). Dissolution point and isolation robustness: Robustness criteria for general cluster analysis methods. *Journal of Multivariate Analysis*, *99*(6), 1154–1176.

Horr, C., & Buechler, S. A. (2021). Breast Cancer Consensus Subtypes: A system for subtyping breast cancer tumors based on gene expression. *NPJ Breast Cancer*, *7*(1), 136.

Jiang, J., Qian, B., Guo, Y., & He, Z. (2024). Identification of subgroups and development of prognostic risk models along the glycolysis-cholesterol synthesis axis in lung adenocarcinoma. *Scientific Reports*, *14*(1), 14704.

Kauppi, K., Rönnlund, M., Nordin Adolfsson, A., Pudas, S., & Adolfsson, R. (2020). Effects of polygenic risk for Alzheimer's disease on rate of cognitive decline in normal aging. *Translational Psychiatry*, *10*(1), 250.

Marques, E. P., Ferreira, F. S., Santos, T. M., Prezzi, C. A., Martins, L. A. M., Bobermin, L. D., Quincozes-Santos, A., & Wyse, A. T. S. (2019). Cross-talk between guanidinoacetate neurotoxicity, memory and possible neuroprotective role of creatine. *Biochimica et Biophysica Acta. Molecular Basis of Disease*, *1865*(11), 165529.

Marttinen, M., Paananen, J., Neme, A., Mitra, V., Takalo, M., Natunen, T., Paldanius, K. M. A., Mäkinen, P., Bremang, M., Kurki, M. I., Rauramaa, T., Leinonen, V., Soininen, H., Haapasalo, A., Pike, I., & Hiltunen, M. (2019). A multiomic approach to characterize the temporal sequence in Alzheimer's disease-related pathology. *Neurobiology of Disease*, *124*, 454–468.

Nilsen, L. H., Witter, M. P., & Sonnewald, U. (2014). Neuronal and astrocytic metabolism in a transgenic rat model of Alzheimer's disease. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *34*(5), 906–914.

Novotny, B. C., Fernandez, M. V., Wang, C., Budde, J. P., Bergmann, K., Eteleeb, A. M., Bradley, J., Webster, C., Ebl, C., Norton, J., Gentsch, J., Dube, U., Wang, F., Morris, J. C., Bateman, R. J., Perrin, R. J., McDade, E., Xiong, C., Chhatwal, J., … Harari. (2023). Metabolomic and lipidomic signatures in autosomal dominant and late-onset Alzheimer's disease brains. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *19*(5), 1785–1799.

Qiang, Y.-X., You, J., He, X.-Y., Guo, Y., Deng, Y.-T., Gao, P.-Y., Wu, X.-R., Feng, J.-F., Cheng, W., & Yu, J.-T. (2024). Plasma metabolic profiles predict future dementia and dementia subtypes: a prospective analysis of 274,160 participants. *Alzheimer's Research & Therapy*, *16*(1). https://doi.org/10.1186/s13195-023-01379-3

Qin, X., Liu, X., Ma, S., & Wu, M. (2024). Supervised Bayesian joint graphical model for simultaneous network estimation and subgroup identification. In *arXiv [stat.ME]*. arXiv. http://arxiv.org/abs/2403.19994

Qiu, S., Cai, Y., Yao, H., Lin, C., Xie, Y., Tang, S., & Zhang, A. (2023). Small molecule metabolites: discovery of biomarkers and therapeutic targets. *Signal Transduction and Targeted Therapy*, *8*(1), 132.

Rappoport, N., & Shamir, R. (2019). NEMO: cancer subtyping by integration of partial multi-omic data. *Bioinformatics (Oxford, England)*, *35*(18), 3348–3356.

R: Clusterwise cluster stability assessment by resampling. (n.d.). Retrieved October 31, 2024, from https://search.r-project.org/CRAN/refmans/fpc/html/clusterboot.html

Rollo, J. L., Banihashemi, N., Vafaee, F., Crawford, J. W., Kuncic, Z., & Holsinger, R. M. D. (2016). Unraveling the mechanistic complexity of Alzheimer's disease through systems biology. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *12*(6), 708–718.

Schweickart, A., Chetnik, K., Batra, R., Kaddurah-Daouk, R., Suhre, K., Halama, A., & Krumsiek, J. (2024). AutoFocus: a hierarchical framework to explore multi-omic disease associations spanning multiple scales of biomolecular interaction. *Communications Biology*, *7*(1), 1094.

Shade, L. M. P., Katsumata, Y., Abner, E. L., Aung, K. Z., Claas, S. A., Qiao, Q., Heberle, B. A., Brandon, J. A., Page, M. L., Hohman, T. J., Mukherjee, S., Mayeux, R. P., Farrer, L. A., Schellenberg, G. D., Haines, J. L., Kukull, W. A., Nho, K., Saykin, A. J., Bennett, D. A., … Fardo, D.

W. (2024). GWAS of multiple neuropathology endophenotypes identifies new risk loci and provides insights into the genetic risk of dementia. *Nature Genetics*, 1–15.

Sirkis, D. W., Bonham, L. W., Johnson, T. P., La Joie, R., & Yokoyama, J. S. (2022). Dissecting the clinical heterogeneity of early-onset Alzheimer's disease. *Molecular Psychiatry*, *27*(6), 2674–2688.

St John-Williams, L., Blach, C., Toledo, J. B., Rotroff, D. M., Kim, S., Klavins, K., Baillie, R., Han, X., Mahmoudiandehkordi, S., Jack, J., Massaro, T. J., Lucas, J. E., Louie, G., Motsinger-Reif, A. A., Risacher, S. L., Alzheimer's Disease Neuroimaging Initiative, Alzheimer's Disease Metabolomics Consortium, Saykin, A. J., Kastenmüller, G., … Kaddurah-Daouk, R. (2017). Targeted metabolomics and medication classification data from participants in the ADNI1 cohort. *Scientific Data*, *4*, 170140.

Tang, Y.-C., Powell, R. T., & Gottlieb, A. (2022). Molecular pathways enhance drug response prediction using transfer learning from cell lines to tumors and patient-derived xenografts. *Scientific Reports*, *12*(1), 16109.

Teruya, T., Chen, Y.-J., Kondoh, H., Fukuji, Y., & Yanagida, M. (2021). Whole-blood metabolomics of dementia patients reveal classes of disease-linked metabolites. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(37), e2022857118.

Xie, M., Kuang, Y., Song, M., & Bao, E. (2024). Subtype-MGTP: a cancer subtype identification framework based on multi-omics translation. *Bioinformatics (Oxford, England)*, *40*(6), btae360.

Xu, J., Begley, P., Church, S. J., Patassini, S., Hollywood, K. A., Jüllig, M., Curtis, M. A., Waldvogel, H. J., Faull, R. L. M., Unwin, R. D., & Cooper, G. J. S. (2016). Graded perturbations of metabolism in multiple regions of human brain in Alzheimer's disease: Snapshot of a pervasive metabolic disorder. *Biochimica et Biophysica Acta*, *1862*(6), 1084–1092.

Young, V. R., & Ajami, A. M. (2000). Glutamate: an amino acid of particular distinction. *The Journal of Nutrition*, *130*(4S Suppl), 892S – 900S.

Zhang, M., Zhang, S., Guo, W., & He, Y. (2024). Novel molecular hepatocellular carcinoma subtypes and RiskScore utilizing apoptosis-related genes. *Scientific Reports*, *14*(1), 3913.

Zhang, N., Guo, Y., Wu, C., Jiang, B., & Wang, Y. (2021). Identification of the molecular subgroups in idiopathic Pulmonary Fibrosis by gene expression profiles. *Computational and Mathematical Methods in Medicine*, *2021*, 7922594.

Zhang, X., Zhou, Z., Xu, H., & Liu, C.-T. (2022). Integrative clustering methods for multi-omics data. *Wiley Interdisciplinary Reviews. Computational Statistics*, *14*(3). https://doi.org/10.1002/wics.1553