A review of neuroimaging-based data-driven approach for Alzheimer’s disease heterogeneity analysis

https://doi.org/10.1515/revneuro-2023-0033
Received March 21, 2023; accepted June 18, 2023; published online July 10, 2023

Abstract: Alzheimer’s disease (AD) is a complex form of dementia and due to its high phenotypic variability, its diagnosis and monitoring can be quite challenging. Biomarkers play a crucial role in AD diagnosis and monitoring, but interpreting these biomarkers can be problematic due to their spatial and temporal heterogeneity. Therefore, researchers are increasingly turning to imaging-based biomarkers that employ data-driven computational approaches to examine the heterogeneity of AD. In this comprehensive review article, we aim to provide health professionals with a comprehensive view of past applications of data-driven computational approaches in studying AD heterogeneity and planning future research directions. We first define and offer basic insights into different categories of heterogeneity analysis, including spatial heterogeneity, temporal heterogeneity, and spatial-temporal heterogeneity. Then, we scrutinize 22 articles relating to spatial heterogeneity, 14 articles relating to temporal heterogeneity, and five articles relating to spatial-temporal heterogeneity, highlighting the strengths and limitations of these strategies. Furthermore, we discuss the importance of understanding spatial heterogeneity in AD subtypes and their clinical manifestations, temporalspatial heterogeneity analysis for AD, and the emerging role of omics data integration in advancing personalized diagnosis and treatment for AD patients. By emphasizing the significance of understanding AD heterogeneity, we hope to stimulate further research in this field to facilitate the development of personalized interventions for AD patients.

Keywords: Alzheimer’s disease; data-driven; disease progression model; heterogeneity; neuroimaging

1 Introduction

Alzheimer’s disease (AD) is a degenerative disease of the central nervous system characterized by an insidious onset and relatively slow progression (Alzheimer’s Association. 2019). Clinically diagnosed AD patients exhibit learning and memory impairment, language issues, and cognitive deficits that eventually lead to death (Dubois et al. 2007). The pathobiology of AD is complex and multifactorial, with hallmark features including the accumulation of amyloid-beta (Aβ) plaques, the formation of neurofibrillary tangles (NFTs) composed of hyperphosphorylated tau proteins, and neurodegeneration leading to brain shrinkage (Rajasekhar et al. 2015). AD has recently been shifted from a clinical-pathologic entity to a clinical-biological entity (Dubois et al. 2014), reflecting advances in the understanding of the underlying pathophysiology of the disease. Biomarkers are essential tools for diagnosing and tracking AD, but their utility may be complicated by spatial and temporal heterogeneity. Spatial heterogeneity refers to the distinct biomarker patterns that occur within the brain, such as the identification of three distinct AD subtypes based on the relative density of NFTs in the hippocampus and cortex, as demonstrated in an autopsy study (Murray et al. 2011). These subtypes include typical AD (TAD), hippocampal-sparing AD (HSAD), and limbic-predominant AD (LPAD). HSAD, for example, shows higher NFT counts in the association cortex and lower NFT counts in the hippocampus compared to TAD, whereas LPAD exhibits the opposite pattern. These findings challenge the notion that AD is a single disease entity and provide evidence for the existence of distinct AD subtypes. Temporal heterogeneity is also present in AD research, as subjects may be at varying stages of the disease. These challenges underscore the need for a comprehensive and integrated approach to biomarker analysis in AD diagnosis and monitoring. Various methods are available
to assess AD biomarkers, including cerebrospinal fluid (CSF) analysis, blood tests, neuropsychological testing, and neuroimaging. Neuroimaging techniques, such as magnetic resonance imaging (MRI) and [18F]-fluorodeoxyglucose (FDG) positron emission tomography (PET), allow for non-invasive monitoring of AD biomarkers, both structurally and functionally. MRI can detect changes in brain volume and shape, while amyloid or tau PET can measure the accumulation of Aβ or tau proteins, which are hallmarks of AD pathology. However, the complexity of heterogeneity analysis for neuroimaging biomarkers arises from interactions between different biomarkers and brain regions. Therefore, it is crucial to develop methods that can account for such heterogeneity to improve the accuracy of AD diagnosis and monitoring.

Following the proposal of the spatial subtype concept, gray matter atrophy patterns from antemortem MRI have been found to correlate with NFT pathology in different subtypes (Whitwell et al. 2012), suggesting the potential for MRI to predict pathological subtypes in vivo. However, hypothesis-driven studies grouping MRI scans into predetermined subgroups have limitations and may not fully capture the spatial heterogeneity of the disease, as observed in several studies (Byun et al. 2015; Risacher et al. 2017). In contrast, data-driven neuroimaging studies utilizing machine learning algorithms ranging from simple unsupervised models to more complex semi-supervised algorithms have shown promise in identifying AD subtypes with atrophy patterns resembling neuropathological subtypes as well as more diverse atrophy patterns. It is noteworthy that the AD subtype approach assumes patients within the same subtype are in the same disease stage.

In addition to identifying AD subtypes, it is also crucial to develop methods that can track disease progression using biomarkers. Stages-only models, such as regression against the disease stage and data-driven disease progression modeling, have been developed, but early studies employing a hypothesis-driven approach (Dickerson et al. 2009; Jack et al. 2010; Scahill et al. 2002) have limitations as they relied on clinical symptom staging for their classifications. To overcome these limitations, data-driven approaches using machine learning algorithms have emerged, representing significant advancements in the study of AD progression. One such example is the algorithm developed by Fonteijn et al. (2011), which uses the event-based model (EBM) and identifies biomarker progression patterns that are independent of preconceived clinical stages. This approach enables a deeper understanding of the temporal progression of biomarkers across various conditions and provides a more comprehensive view of disease progression. However, despite these advancements, these models assume that one subtype exists for all patients and that they follow a similar disease trajectory.

Recent studies (Young et al. 2018) have attempted to combine the two types of models to estimate AD subtypes and progression simultaneously, as both heterogeneities have non-negligible limitations. The combination of subtype and temporal heterogeneity models enables a more accurate and detailed understanding of the disease’s complexity.

Several review articles have recently been published, highlighting the novel findings of disease subtypes obtained through data-driven investigations of AD heterogeneity (Blanken et al. 2019; Zhang et al. 2021). However, there has been a relative lack of attention given to the comparison and evolution of data-driven approaches in AD heterogeneity analysis. Therefore, a systematic review of the three primary data-driven approaches has been conducted to improve disease prediction in individuals and provide a useful reference for future AD heterogeneity studies. Our literature search strategy is detailed in Section 2, while Section 3 provides an overview of the studies that employed various data-driven methodologies to study AD heterogeneity. These studies were divided into three categories: spatial heterogeneity, temporal heterogeneity, and spatial-temporal heterogeneity, as depicted in Figure 1. The subtypes generated by spatial heterogeneity, the possibility of different biomarker sequences in temporal heterogeneity, the recent advancements in spatial-temporal heterogeneity analysis for AD, and the emerging role of omics data integration in advancing personalized diagnosis and treatment for AD patients are discussed in the final section. Overall, this review article serves as a supplement to the existing literature, enhances our understanding of the current state of AD heterogeneity analysis, and aims to improve disease prediction in individuals while providing a valuable resource for future research in this field.

2 Search and review methodology

A comprehensive search for related literature was performed, utilizing pertinent articles indexed in the PubMed, Web of Science, and ScienceDirect databases. The inclusion criteria for the publications were restricted to those published between January 1, 2011, and December 15, 2022. As the data-driven approach can be segregated into three primary categories, a systematic search was conducted via three independent pathways, as depicted in Figure 2.

The first pathway entailed the use of the keywords ‘Alzheimer’s Disease’, ‘imaging’, ‘subtype’, ‘heterogeneous’, and ‘heterogeneity’. The literature search garnered 784

In the third pathway, the search terms ‘Alzheimer’s Disease’, ‘subtype’, ‘heterogeneous’, ‘heterogeneity’, ‘stage’, ‘event-
based model, ‘progression model’, ‘trajectory model’, ‘pathway model’ and ‘imaging’ were employed in combinations. This pathway yielded 93 articles in PubMed, 191 articles in Web of Science, and 103 articles in ScienceDirect. To establish eligibility for analysis, articles were screened according to predefined criteria. Studies that did not meet the following criteria were excluded from further analysis: 1) studies with non-human subjects; 2) unavailable full-texts; 3) non-English studies; 4) review papers or papers that were not journal articles; 5) biological, neuropsychological, or pharmacological research only; 6) absence of neuroimaging features as algorithm input features; 7) studies pertaining to dementia or neurological diseases other than AD; and 8) non-data-driven study methods. Ultimately, only 41 articles (22 for spatial heterogeneity, 14 for temporal heterogeneity, and 5 for spatial-temporal heterogeneity) were deemed suitable for further analysis.

3 Data-driven methods for AD heterogeneity studies

3.1 Spatial heterogeneity method

In recent years, the use of data-driven computational methods to identify AD subtypes from large-scale AD data has gained increased popularity. These methods fall into two broad categories: “unsupervised” and “semi-supervised” algorithms. The unsupervised algorithms do not require prior knowledge of disease subtypes, and they can reveal patterns in the data that may not be immediately apparent. In contrast, the semi-supervised algorithms incorporate some prior knowledge into the analysis, typically in the form of a control group, and can improve the accuracy of subtype identification. Both types of algorithms have shown promise in AD heterogeneity analysis and have the potential to improve our understanding of this complex disease.

3.1.1 Unsupervised method

Table 1 highlights the literature pertaining to the utilization of spatial heterogeneity in unsupervised methods. Unsupervised methods utilize neuroimaging biomarkers to identify clusters without any pre-defined labels. Clustering methods, which are frequently employed in data-driven analyses, are used to identify subtypes of AD. The two main types of clustering methods are hierarchical clustering and partitional clustering. In the hierarchy of clustering, the hierarchy of clusters is presented through a process of splitting and combining clusters at different levels. Conversely, partitioning clustering creates and evaluates partitions based on a specific set of criteria.

Agglomerative hierarchical clustering (AHC) is one of the most widely used hierarchical clustering methods in AD subtyping. AHC constructs a hierarchy of clusters, where each AD patient starts as a singleton cluster, and clusters are sequentially combined based on their distance to other clusters (Day and Edelsbrunner 1984). AHC also allows for arbitrary selection of the number of clusters, which is beneficial when the optimal number of clusters is unknown.

In various studies, AHC has been applied with different linkage techniques, such as Ward’s linkage, complete linkage, single linkage, and average linkage, to calculate the criteria for the merging of clusters. For instance, Hwang et al. (2016) and Noh et al. (2014) have exhibited the merit of implementing AHC utilizing Ward’s linkage to analyze cortical thickness measurements in the context of AD patients. Similarly, Kärkkäinen et al. (2020) adopted the complete linkage technique to analyze gray matter density maps of AD patients and successfully identified four well-defined subtypes of AD. Additionally, Ossenkoppele et al. (2020) extensively measured cortical thickness parameters in the entire occipital/parietal cortex and frontal cortex, as well as the total intracranial volume-weighted hippocampal volumes. They employed a two-step clustering algorithm on these variables, which first utilized a sequential clustering approach with a modified cluster feature tree and a model-based distance criterion. In the second step, the AHC method was used with the preclusters generated from the first step as inputs, which successfully identified four subtypes of AD.

It is noteworthy that functional imaging techniques, such as PET, hold the potential to provide complementary information to the structural features of gray matter. For example, Jeon et al. (2019) and Levin et al. (2021) used cortical thickness and functional features from different kinds of PET as input features for AHC analysis with Ward’s linkage clustering. These features proved useful in heterogeneity analysis, leading to the identification of three different AD subtypes. Tau neurofibrillary tangles are the primary driver of downstream neurodegeneration and cognitive decline in AD. However, there exists significant variability in the tau-neurodegeneration relationship (Duara and Barker 2022), which may indicate atrophy deviation from the expected tau level in a specific area. Das and his colleagues (2021) calculated the tau-neurodegeneration mismatch metric and established six clusters using AHC with Ward’s linkage. The utilization of the random forest similarity measure provides vital information on both the linear and nonlinear interactions between variables. Poullakis and his collaborators (2018) employed this measure to extract the similarity
matrix from the cortical and subcortical volume features. They subsequently reduced the dimensional representation of this matrix using classical multidimensional scaling. By utilizing AHC and employing the average linkage algorithm, they identified two typical diffuse atrophy subtypes and three atypical AD subtypes. These various subtypes exhibit significantly different atrophy patterns, demographic, clinical, and cognitive characteristics, as well as variable rates of cognitive decline. Overall, the application of AHC in conjunction with appropriate linkage techniques emerges as a crucial tool in clustering analysis to distinguish between AD subtypes.

Hierarchical clustering offers the advantage of creating a hierarchy of clusters, but it may be sensitive to noise and outliers. Partitioning clustering, on the other hand, assigns data to different clusters without using a hierarchical structure, making it more straightforward and scalable. However, it usually requires a fixed number of clusters to be specified (Saxena et al. 2017). Non-negative matrix factorization (NMF) is a method that decomposes a feature matrix into two matrices, representing feature representation and potential clusters. Ten Kate et al. (2018) employed NMF to identify four subtypes of AD patients. The Louvain method is a partitional clustering technique that optimizes modularity and can be used to identify communities in a network. It is particularly useful for analyzing complex systems with many interdependent components.

Table 1: Summary of unsupervised spatial heterogeneity methods.

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Dataset</th>
<th>Features</th>
<th>AD subtypes</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hwang et al. (2016)</td>
<td>ADNI-2 (AD: 77)</td>
<td>Cortical thickness</td>
<td>Medial temporal; parietal dominant; diffuse</td>
<td>AHC with Ward’s linkage</td>
</tr>
<tr>
<td>Das et al. (2021)</td>
<td>ADNI (MCI: 80; AD: 57)</td>
<td>Tau-neurodegeneration mismatch metric</td>
<td>Six clusters</td>
<td>AHC with Ward’s linkage</td>
</tr>
<tr>
<td>Jeon et al. (2019)</td>
<td>SMC, GUGMC (AD: 83)</td>
<td>Cortical thickness, functional features</td>
<td>Medial-temporal dominant; parietal-dominant; diffuse</td>
<td>AHC with Ward’s linkage</td>
</tr>
<tr>
<td>Ten Kate et al. (2018)</td>
<td>ADC-d (AD: 299); ADC-v (AD:181); ADNI-1/2/GO (AD:227)</td>
<td>Gray matter volumes</td>
<td>Medial-temporal predominant atrophy; parieto-occipital atrophy; mild atrophy; diffuse cortical atrophy</td>
<td>NMF</td>
</tr>
<tr>
<td>Kärkkäinen et al. (2020)</td>
<td>ADNI-1/2/GO (MCI: 751; AD: 282)</td>
<td>Gray matter density</td>
<td>Typical; cortically-driven atypical; limbic-predominant; early-onset</td>
<td>AHC with complete linkage</td>
</tr>
<tr>
<td>Levin et al. (2021)</td>
<td>ADNI-1/2/3/GO (AD: 177)</td>
<td>Voxel-wise FDG-PET</td>
<td>Typical; limbic-predominant; cortical-predominant</td>
<td>AHC with Ward’s linkage</td>
</tr>
<tr>
<td>Noh et al. (2014)</td>
<td>SMC (AD: 152)</td>
<td>Cortical thickness</td>
<td>Bilateral medial temporal-dominant atrophy; parietal-dominant; diffuse atrophy</td>
<td>AHC</td>
</tr>
<tr>
<td>Ossenkoppele et al. (2020)</td>
<td>MDCGSH, Swedish BioFINDER study, UCSF AD Research Center (MCI: 83; AD: 177)</td>
<td>Cortical thickness, total intracranial and hippocampal volumes</td>
<td>Typical; limbic-predominant; hippocampal-sparing; mild atrophy</td>
<td>AHC</td>
</tr>
<tr>
<td>Park et al. (2017)</td>
<td>SMC (AD: 225)</td>
<td>Cortical thickness</td>
<td>Parietal-predominant; medial temporal-predominant; diffuse</td>
<td>Louvain method</td>
</tr>
<tr>
<td>Poulakis et al. (2018)</td>
<td>AddNeuroMed, ADNI-1 (AD: 299)</td>
<td>Cortical and subcortical volume</td>
<td>Minimal atrophy; limbic-predominant; hippocampal-sparing; diffuse 1; diffuse 2</td>
<td>AHC with average linkage</td>
</tr>
<tr>
<td>Poulakis et al. (2020)</td>
<td>ADNI-1 (AD: 72)</td>
<td>Cortical thickness and subcortical volume</td>
<td>Typical diffuse; minimal atrophy; hippocampal-sparing</td>
<td>MMGLMM</td>
</tr>
<tr>
<td>Poulakis et al. (2022)</td>
<td>ADNI, J-ADNI, AIBL, AddNeuroMed (AD: 891)</td>
<td>Cortical thickness, subcortical volume</td>
<td>Diffuse atrophy; limbic-predominant atrophy; limbic-predominant atrophy+; hippocampal-sparing atrophy; minimal atrophy</td>
<td>Bayesian clustering</td>
</tr>
<tr>
<td>Sui et al. (2018)</td>
<td>ADNI-2/GO (MCI: 134; AD: 48)</td>
<td>White matter impairment factors</td>
<td>Temporal-frontal impairment factor; parietal factor; long fiber bundle factor</td>
<td>LDA</td>
</tr>
<tr>
<td>Sun et al. (2019)</td>
<td>ADNI-2/GO (MCI: 454; AD: 149)</td>
<td>Gray matter density</td>
<td>Medial temporal lobe-memory; lateral temporal-language; posterior cortical-executive</td>
<td>MMLDA</td>
</tr>
<tr>
<td>Toledo et al. (2022)</td>
<td>ADNI (AD: 282)</td>
<td>Tau PET and gray matter volume</td>
<td>Four tau (I–IV) clusters; diffuse; hippocampal-sparing; limbic-predominant</td>
<td>RCC</td>
</tr>
<tr>
<td>Zhang et al. (2016)</td>
<td>ADNI-1 (MCI: 394; AD: 188)</td>
<td>Gray matter density</td>
<td>Temporal atrophy; subcortical atrophy; cortical atrophy</td>
<td>LDA</td>
</tr>
</tbody>
</table>

Databases: Alzheimer’s Disease Neuroimaging Initiative (ADNI); Samsung Medical Center (SMC); Amsterdam Dementia Cohort (discovery/validation dataset) (AD(d/v)); Gachon University Gil Medical Center (GUGMC); Memory Disorder Clinic of Gangnam Severance Hospital (MDCGSH); University of California San Francisco (UCSF); Japanese Alzheimer’s Disease Neuroimaging Initiative (J-ADNI); Australian Imaging, Biomarkers and Lifestyle study (AIBL).
in a greedy manner by integrating groups into a single node and conducting modularity clustering on condensed graphs (Blondel et al. 2008). Cluster validity metrics are utilized to evaluate the quality of cluster partitions and modularity. Park et al. (2017) employed the Louvain method to identify AD subtypes by calculating the cortical atrophy pattern for each patient and constructing a similarity matrix based on the correlation. Higher modularity values suggest a stronger intra-module connection and weaker inter-module links, indicative of better partitions. Recently, Toledo et al. (2022) used robust collaborative clustering (RCC) to determine three subtypes. RCC uses the matrix trivialization technique to cluster topics and features into heterogeneous groups simultaneously, minimizing the influence of noise and outliers and enhancing robustness.

Bayesian model-based approaches have garnered attention in neuroimaging research due to their ability to handle uncertainty and model complexity. One such approach is Latent Dirichlet Allocation (LDA), a generative three-layer Bayesian mixture model that has been extensively used in text analysis to discover latent topics automatically in a collection of text documents (Blei et al. 2003). In neuroimaging studies, LDA has been applied to recognize distinct latent factors in overlapping patterns in AD patients. Research has shown that AD patients exhibit multiple latent atrophy patterns rather than just one atrophy subtype. To identify these patterns, AD patients were regarded as text documents, atrophy factors as topics, and atrophy-affected voxels as dictionary words. Each patient was then modeled as an unorganized collection of words, each of which had one or more atrophy factors. Each of those atrophy factors was represented as a probability distribution across the atrophy voxels. Correspondingly, each patient exhibited one or more latent atrophy factors, each linked to different degrees of atrophy across the brain. Through the utilization of LDA, researchers were able to identify three latent gray matter density atrophy factors, namely the temporal atrophy factor, the subcortical atrophy factor, and the cortical atrophy factor (Zhang et al. 2016). Additionally, white matter impairment plays a significant role in the pathogenesis of AD (Agosta et al. 2011). Therefore, the LDA method proves to be a suitable framework for studying heterogeneity and identifying latent white matter impairment factors. Sui et al. (2018) leveraged the LDA method to access whole-brain white matter skeleton images and uncovered three latent factors: the temporal-frontal impairment factor, the parietal factor, and the long fiber bundle factor. Expanding upon the LDA method, researchers (Sun et al. 2019) have also used multimodality LDA (MMLDA) to uncover potential factors in both atrophy patterns and cognitive deficits. In one study, MMLDA was used to identify three atrophy-cognitive factors in AD patients, demonstrating the potential of this method for uncovering complex disease mechanisms. Another Bayesian model-based approach proposed by Poulakis et al. (2020) is the multivariate mixture of generalized mixed-effect models (MMGLMM), which is a generalized linear method for clustering optimization using Bayesian methods. This approach works on both cross-sectional and longitudinal data and accounts for unequal visits per individual. Based on cortical thickness and subcortical volume, three different atrophy patterns were identified. Later, a study (Poulakis et al. 2022) combined MMGLMM with Bayesian inference to explore the heterogeneity of longitudinal brain data in AD patients, resulting in the identification of five different subtypes and deducing the order of atrophy regions in conjunction with their longitudinal data. Overall, Bayesian model-based approaches, including LDA, MMLDA, and MMGLMM, are potent tools for uncovering sophisticated patterns in neuroimaging data and have the potential to advance our understanding of AD subtypes.

3.1.2 Semi-supervised method

Semi-supervised learning is an important field in artificial intelligence that seeks to use both supervised and unsupervised learning techniques to maximize the effectiveness of a given learning algorithm. This technique differs from traditional unsupervised learning methods, which typically focus on singling out one primary imaging feature or pattern. Instead, semi-supervised learning meticulously leverages both labeled and unlabeled data to construct a baseline classifier that enables the distinction between control and AD groups. As an added step, this classifier is further updated in an unsupervised manner, leading to the identification of multiple imaging patterns that have a close association with AD. Ultimately, this holistic analysis of data is instrumental in better comprehending the disease. Researchers have taken notice of this promising method and demonstrated a growing inclination toward using semi-supervised learning to uncover significant developments in heterogeneous data. A glance at Table 2 shows compelling results that attest to the efficacy of this technique in predicting and understanding diseases.

One notable methodology in this area is CHIMERA, which was developed by Dong et al. (2016). The approach taken in this framework is to combine multiple regularized clustering methods with the aim of modeling AD pathology. The study participants were divided into two groups: a pathological group, which included patients with MCI and AD, and a normal control group. CHIMERA employs an innovative probability distribution mapping technique that transforms the reference distribution into patient-specific
The double-cyclic optimization procedure was employed to clustering using the max-margin multiple SVM classification and clustering (MAGIC), which employed an unsupervised representation learning algorithm called orthogonal projection non-negative matrix factorization (opNMF) to extract multi-scale, biologically interpretable feature representations from the entire brain. After extracting high-dimensional imaging data to a low-dimensional representation via opNMF, the HYDRA algorithm was utilized to create a convex polytope for simultaneous classification and clustering using the max-margin multiple SVM classifiers. The double-cyclic optimization procedure was employed to merge the information from multi-scale characteristics for consistent inter-scale clustering solutions. Finally, the resulting multi-scale clustering solutions were combined using consensus clustering to create the final subtypes.

The spatial heterogeneity methods categorization is displayed in Figure 3. The ongoing exploration and application of semi-supervised learning methods within AD research have significant potential to improve the accuracy of diagnosis, which would ultimately result in improved outcomes for patients.

### 3.2 AD subtypes in spatial heterogeneity studies

Murray et al. (2011) have identified three major subtypes of AD through postmortem examination: TAD (approximately 75%), HSAD, and LPAD (approximately 25%).

TAD is the most widespread subtype of AD, accounting for the majority of cases. It is characterized by diffuse atrophy of the cerebral cortex, which can be observed through structural MRI. Several studies have referred to this subtype as the “diffuse” or “widespread” AD subtype (Dong et al. 2017; Hwang et al. 2016; Jeon et al. 2019; Noh et al. 2014; Park et al. 2017; Poulakis et al. 2018, 2020; Ten Kate et al. 2018; Varol et al. 2017; Zhang et al. 2021b). In contrast, LPAD is distinguished by the presence of tangled neurofibrillary tangles in the medial temporal lobe (MTL) (Murray et al. 2011). Several studies (Hwang et al. 2016; Jeon et al. 2019; Noh et al. 2014; Park et al. 2017; Sun et al. 2019; Ten Kate et al. 2018; Varol et al. 2017; Zhang et al. 2016) have directly labeled LPAD as an “MTL” or “MT” (medial temporal) AD subtype. HSAD, on the other hand, is characterized by brain atrophy mainly in the temporal, parietal, and frontal lobes while sparing the MTL. Additionally, this subtype is occasionally referred to as the “cortical atrophy only” AD subtype (Byun et al. 2015). Some other studies include “parietal” (Hwang et al. 2016;
Noh et al. 2014; Park et al. 2017) or “posterior cortical” (Sun et al. 2019) to describe this subtype. As for minimal atrophy AD (MAD), it has an atrophy pattern that resembles normal aging (Mizuno et al. 2000) and is frequently labeled as “mild” (Dong et al. 2017; Ten Kate et al. 2018) or “spared” (Byun et al. 2015) in comparison to other AD subtypes. Beyond the four primary subtypes of AD, emerging research efforts have identified new subtypes that warrant close examination. Levin et al. (2021) have identified a “cortical-predominant” subtype, which stands out due to a younger onset of the disease and severe executive dysfunction. Furthermore, Sun et al. (2019) have identified an atrophy factor that is connected with lateral temporal lobe atrophy, implying that language deficits may constitute a noteworthy feature of this subtype. These discoveries underscore the significance of comprehending the distinct subtypes of AD and their respective clinical features, as they may have implications for streamlined diagnosis, targeted treatment options, and disease management.

3.3 Temporal heterogeneity based on different biomarkers

Given the continuous nature of AD’s progression, there has been growing interest in exploring data-driven approaches to examine how AD biomarker changes occur throughout various stages of disease progression. Unlike the previous section that classified AD patients into different subtypes, the temporal heterogeneity divided AD patients into various stages of an evolving AD process. The literature that utilized temporal heterogeneity approaches is compiled in Table 3. Among the most frequently utilized models for predicting AD progression is the EBM. This model was first presented by Fonteinj et al. (2011) and has subsequently been employed in numerous studies to predict the biomarker ordering and staging of participants as they move from a healthy state to the full spectrum of the disease.

The EBM constitutes a powerful probabilistic framework for modeling disease progression focused on biomarkers. It captures changes in biomarker status as discrete “events,” tracking the shift from a “normal” to an “abnormal” state. The EBM incorporates a comprehensive list of such events, denoted as $E_i$ and ordered according to a sequence $S$, accompanied by the corresponding set of biomarkers $X$. This model leverages this information to identify the most probable event sequence and quantify the connectivity between different biomarkers.

The EBM’s crucial aspect is the estimation of the likelihood function $P(X|S)$ in disease progression simulation. This calculation is based on individual measurements and is essential in identifying changing biomarkers and their sequence. The likelihood function $P(X|S)$ integrates two likelihoods, indicating the probability of the biomarker measurement with or without an event’s occurrence. This model assumes a uniform prior probability. In order to estimate the posterior distribution, Bayes’ theorem is employed, employing the Markov Chain Monte Carlo (MCMC) sampling process to generate the most probable event sequence list. The event order is established, and the positional variance diagram is used to comprehend the range of event sequencing across various patients. Fonteinj et al. (2012) have implemented a variant of EBM on longitudinal familial AD data, incorporating a slightly different characteristic event sequence in conjunction with positional variance diagrams based on clinical status and...
regional atrophy. Rather than the Gaussian mixture model, a uniform model was employed in modeling P(X|S).

EBM offers a significant edge through its statistical mechanism, which enables the acquisition of knowledge on normal and abnormal biomarker value distributions from cross-sectional or short-term longitudinal data without prior staging or cut points. However, in distinguishing between early-onset familial AD and late-onset sporadic AD, a significant challenge persists, particularly for patients with late-onset AD who may have received a misdiagnosis. A key limitation of the original EBM is its assumption that all patients follow a single sequence of events leading to AD. To resolve this issue, Young et al. proposed two modifications. The first modification proposed (Young et al. 2014) employs a combination of two normal distributions to model each biomarker separately. This modification enables the inclusion of asymptomatic AD cases and misdiagnosed patients in EBM analysis, thereby preventing their contamination of control subjects. The second modification proposed (Young et al. 2015a) loosens the core assumption of EBM of a single event sequence for all subjects. Instead, it employs the generalized Mallows model that parameterizes variance in the primary event sequence to enable subjects to deviate from it. The Dirichlet process mixture of generalized Mal- lows models permits subgroups of subjects to adhere to distinct event sequences, with each subject having an applicable variance. This modification significantly enhances the EBM's accuracy in identifying the most probable order of events leading to AD.

In machine learning, generative models and discriminative models focus on modeling joint probability and conditional probability, respectively. The original EBM (Fonteijn et al. 2011) and its subsequent variations (Oxtoby et al. 2018; Young et al. 2014) belong to the family of generative models with a focus on maximizing likelihood P(X|S). On the other hand, discriminative models, such as the discriminative event-based model (DEBM) conceptualized by Venkatraghavan et al. (2019), estimate a subject's event order based on the posterior probability of a single biomarker becoming abnormal. This model utilizes a generalized Mallows model.

### Table 3: Summary of temporal heterogeneity approaches.

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Dataset</th>
<th>Biomarkers</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archetti et al. (2019)</td>
<td>Seven independent cohorts (MCI: 1424; AD: 743)</td>
<td>Cognitive test scores, CSF and imaging biomarkers</td>
<td>EBM variant and DEBM</td>
</tr>
<tr>
<td>Fonteijn et al. (2011)</td>
<td>NHNN (AD: 9)</td>
<td>Clinical status and regional atrophy</td>
<td>Original EBM</td>
</tr>
<tr>
<td>Fonteijn et al. (2012)</td>
<td>NHNN (AD: 9)</td>
<td>Clinical status and regional atrophy</td>
<td>EBM variant</td>
</tr>
<tr>
<td>Golriz Khatami et al. (2022)</td>
<td>Ten independent AD cohort (MCI: 851; AD: 380)</td>
<td>Cognitive test scores, CSF and imaging biomarkers</td>
<td>EBM variant and a rank aggregation algorithm</td>
</tr>
<tr>
<td>Marinescu et al. (2019)</td>
<td>ADNI (MCI: 235; AD: 81), DRC (PCA: 23; TAD: 20)</td>
<td>Imaging biomarkers</td>
<td>DIVE</td>
</tr>
<tr>
<td>Oxtoby et al. (2018)</td>
<td>ADNI (MCI: 338)</td>
<td>Cognitive test scores, CSF and imaging biomarkers</td>
<td>EBM variant and a non-parametric differential equation model</td>
</tr>
<tr>
<td>Schiratti et al. (2015)</td>
<td>ADNI-1/2/GO (MCI: 753; AD: 248)</td>
<td>ADAS-Cog scores and imaging biomarkers</td>
<td>Mixed-effects models</td>
</tr>
<tr>
<td>Schmidt-Richberg et al. (2015)</td>
<td>ADNI-1/2/GO (MCI: 184; AD: 88)</td>
<td>Cognitive test scores and imaging biomarkers</td>
<td>Quantile regression</td>
</tr>
<tr>
<td>Venkatraghavan et al. (2019)</td>
<td>ADNI-1/2/GO (MCI: 235; AD: 342)</td>
<td>Cognitive test scores, CSF and imaging biomarkers</td>
<td>DEBM</td>
</tr>
<tr>
<td>Venkatraghavan et al. (2021a)</td>
<td>ADNI-1/2/GO (MCI: 235; AD: 342)</td>
<td>Cognitive test scores, APOE, CSF and imaging biomarkers</td>
<td>Co-init DEBM</td>
</tr>
<tr>
<td>Venkatraghavan et al. (2021b)</td>
<td>ADNI-1/2/GO (MCI: 732; AD: 223); RS (MCI: 2807; AD: 25)</td>
<td>APOE and imaging biomarkers</td>
<td>Co-init DEBM</td>
</tr>
<tr>
<td>Young et al. (2014)</td>
<td>ADNI (MCI: 228; AD: 81)</td>
<td>Cognitive test scores, CSF and regional atrophy</td>
<td>EBM variant</td>
</tr>
<tr>
<td>Young et al. (2015a)</td>
<td>ADNI (MCI: 149; AD: 98)</td>
<td>Cognitive test scores, CSF and imaging biomarkers</td>
<td>EBM variant</td>
</tr>
<tr>
<td>Young et al. (2015b)</td>
<td>ADNI-1 (MCI: 398; AD: 192)</td>
<td>Cognitive test scores, CSF and imaging biomarkers</td>
<td>EBM variant and DEM</td>
</tr>
</tbody>
</table>

Databases: National Hospital for Neurology and Neurosurgery (NHNN); Dominantly Inherited Alzheimer Network (DIAN); Rotterdam Study (RS); Dementia Research Centre (DRC); Posterior Cortical Atrophy (PCA). Methods: differential equation model (DEM); Data-driven Inference of Vertexwise Evolution (DIVE).
to describe the distribution of the event order and gauges the distance between orderings using Kendall’s Tau distance. In tandem, event centers are used, and their relative distance between events is ascertained to construct a timeline for AD progression and stage the subjects. This approach enables more effective handling of large datasets and improves stability in the ordering of abnormal events. To apply DEBM to different population groups, Venkatraghavan et al. (2021a) divided it into group-specific and group-aspecific components. The group-aspecific section utilizes the entire dataset for training, while only data from a particular group is utilized to train the group-specific parts.

Beyond EBM-based models, alternative approaches have been proposed for estimating disease progression in AD. For instance, Schiratti et al. (2015) introduced a mixed-effects model incorporating time reparameterization to estimate disease progression. This approach leverages a time shift parameter instead of the intercept of random effects to measure the onset of disease, whether it occurs early or late. Acceleration factors are also employed to account for the variability in the pace of disease progression. Another approach utilized is quantile regression, which approximates clinical biomarkers’ trajectory and individual-level disease rate progression (Schmidt-Richberg et al. 2015). Furthermore, Marinescu et al. (2019) proposed an image-based disease progression model with single-vertex resolution that integrates unsupervised learning and disease progression modeling. This approach utilizes an expectation-maximization framework that iteratively assigns each vertex to a cluster, predicts the trajectory in each cluster, and determines the disease progression score for each subject until convergence is achieved.

### 3.4 Spatial-temporal heterogeneity method

Most previous studies investigating AD heterogeneity have focused either on spatial heterogeneity or temporal heterogeneity. Spatial heterogeneity suggests that different patient subgroups exhibit distinct features even at the same disease stage, while temporal heterogeneity posits that all patients appear different at various stages of disease progression. These narrow assumptions limit the generalizability of the findings. To address these limitations, Young et al. (2018) proposed an unsupervised algorithm named subtype and stage inference (SuStaIn) to uncover both cross-sectional heterogeneity and temporal stages of AD. SuStaIn integrates traditional cluster analysis with EBM. In contrast to the EBM approach that views events as a rapid transition from “normal” to “abnormal” (Fonteijn et al. 2012), the progression trajectory of biomarkers for each subtype of SuStaIn is described by a linear Z-score model, where events are represented as a continuous and linear change from one Z-score to another. When applied to a large collection of cross-sectional MRIs, it identified three AD subtypes, each reflecting unique neuropathological mechanisms and clinical presentations of AD. SuStaIn also provides a framework for identifying disease stages and tracking disease progression over time. The SuStaIn model has shown great potential in uncovering AD subtypes and temporal stages. However, it is crucial to assess whether SuStaIn is a robust and generalizable model that can be applied to various datasets. In the study by Archetti et al. (2021), several datasets were utilized to train the SuStaIn model, with subsequent assessments of its robustness and generalizability across various independent datasets. The results indicated that SuStaIn provides a dependable framework capable of accurately identifying subtypes and stages of AD within a variety of datasets. Collij et al. (2022) further utilized SuStaIn to analyze Aβ-protein accumulation subtypes based on amyloid-PET data, identifying three subtypes named the frontal, apical, and occipital subtypes. The outcomes revealed the reliability of SuStaIn as a method for identifying and characterizing distinct AD subtypes based on different biomarkers. Additionally, AD is often distinguished by the propagation of tau pathology throughout the cerebral cortex. SuStaIn also offers a framework for identifying disease stages and tracking disease progression over time. In a recent study undertaken by Vogel and colleagues (2021), the SuStaIn model was employed in conjunction with tau-PET scans to identify distinct spatiotemporal trajectories of tau pathology in individuals with AD. This innovative approach successfully identified four stable subtypes of tau pathology, which were effectively replicated across a separate sample that utilized a dissimilar radiotracer. Previous research already accepted a significant correlation between the degree of cerebral asymmetry and key indicators of AD progression, including disease severity, cognitive outcomes, and neuropsychiatric symptoms (Bruen et al. 2008; Frings et al. 2015; Geroldiet al. 2000; Wolf et al. 2001). The findings of Vogel et al.’s (2021) investigation revealed an intriguing trend concerning the relationship between laterality and SuStaIn stage, with the latter producing an increasing trend of cerebral lateralization. However, it should be noted that despite this trend, a considerable proportion of individuals still exhibited reverse lateralization compared against the group average for their subtype. As a result, these findings may suggest a more complex interrelationship between lateralization and subtyping.

Despite its usefulness, the SuStaIn method has a limitation in that it cannot accommodate discrete ordinal data. To address this issue, Young et al. (2021) extended SuStaIn to
handle a variety of discrete ordinal data. The extension of SuStaIn for ordinal data involves substituting the data likelihood of the scored events model for the data likelihood of each subtype in the SuStaIn model. Their results demonstrated that the extended SuStaIn model improved subtyping and staging of individuals, and precise uncertainty estimates on discrete scored data. These findings highlight the potential of SuStaIn as a promising tool for identifying AD subtypes and providing stages based on various biomarkers, including discrete ordinal data.

4 Discussion

4.1 Spatial heterogeneity in AD subtypes and its clinical manifestations

Advancements in AD research have revealed that AD is not a singular disease, but rather constitutes a range of subtypes that display distinct pathological and clinical features and exhibit spatial heterogeneity. A multitude of investigations have identified several comparable AD neuropathological subtypes at the group level (Dong et al. 2017; Hwang et al. 2016; Jeon et al. 2019; Levin et al. 2021; Noh et al. 2014; Ossenkoppele et al. 2020; Park et al. 2017; Poulakis et al. 2020; Ten Kate et al. 2018; Varol et al. 2017). These subtypes exhibit varying characteristics regarding genotype, demographics, clinical presentation, and pathological features, with some resembling TAD subtypes, LPAD subtypes, HSAD subtypes, and MAD subtypes. It is noteworthy that these subtypes differ notably in their brain neuropathology and distribution, as well as in their impact on clinical presentation and prognosis. Therefore, comprehending the spatial heterogeneity of AD subtypes is imperative, as it can support advancements in research and improve clinical management and treatment.

AD is a complex neurodegenerative disorder that develops from the interaction of multiple factors, including Aβ and tau-related neuropathologies, as well as variable heterogeneous and environmental traits. These factors contribute to the development of unique structural, cognitive, and functional brain impairments among AD subtypes. The clinical manifestation and progression of AD result from the balance between risk and protective factors (Ferreira et al. 2020; Jellinger 2020). Furthermore, the diverse spatial manifestations and underlying pathological patterns among AD subtypes contribute to variations in their clinicopathological presentations (Jellinger 2020). Protective factors primarily encompass cognitive reserve and related concepts of brain resilience/resistance. In contrast, risk factors are predominantly related to gender, APOE, and age. Demographic factors, including age, gender, and education, are commonly used in current subtype studies, with age (Hebert et al. 2013; Riedel et al. 2016) and female gender (Farrer et al. 1997) being the major risk factors for AD. Patients with low levels of education exhibit clinical symptoms of AD earlier than those with high levels, indicating the crucial role of cognitive reserve in delaying the onset of the disease (Persson et al. 2017). The APOE ε4 allele is widely regarded as a highly significant and extensively reported risk factor for the development of AD neuropathology. APOE ε4 is associated with long-fiber tract factors, the elevation of plasma cholesterol, and contributes to the development of cardiovascular disease, neurodegenerative disease, and cognitive deterioration (Mahley et al. 2009; Zehnder et al. 2009). Studies have found that APOE ε4 increases Aβ deposition and Aβ oligomer formation and consistently shows higher tau, phosphorylated tau (p-tau), and tau/Aβ42 ratios (Mattsson et al. 2009), indicating its crucial role in the development and progression of AD.

Current research indicates that subtypes of AD exhibit distinctive clinical and neuropathological profiles. Patients with TAD are typically older (Dong et al. 2017; Ferreira et al. 2018; Hwang et al. 2016; Noh et al. 2014; Varol et al. 2017) or only slightly younger than those with LPAD (Ossenkoppele et al. 2020; Risacher et al. 2017; Ten Kate et al. 2018). They have a higher proportion of the APOE ε4 allele and abnormal levels of Aβ42 (Dong et al. 2017; Ferreira et al. 2017; Ossenkoppele et al. 2020; Ten Kate et al. 2018; Varol et al. 2017; Zhang et al. 2016). LPAD patients are mainly female, older, have lower levels of education (Dong et al. 2017; Ferreira et al. 2017; Jeon et al. 2019; Noh et al. 2014; Ossenkoppele et al. 2020; Park et al. 2017), higher levels of Aβ42 and lower levels of total tau and p-tau (Byun et al. 2015; Ferreira et al. 2017; Poulakis et al. 2018; Varol et al. 2017). HSAD patients, on the other hand, tend to be predominantly male, younger, and have a higher level of education (Ferreira et al. 2017; Ossenkoppele et al. 2020; Poulakis et al. 2018). MAD patients have an earlier age of onset and the lowest level of education (Ferreira et al. 2017; Poulakis et al. 2018; Zhang et al. 2021b), with little brain atrophy despite being diagnosed with AD, a polarized proportion of APOE ε4 carriers, and abnormal levels of Aβ and tau (Dong et al. 2017; Ferreira et al. 2017; Persson et al. 2017; Poulakis et al. 2018).

A comprehensive understanding of AD subtypes is crucial for developing targeted therapies and improving patient outcomes. In a systematic review and meta-analysis conducted by Ferreira et al. (2020), it was concluded that AD heterogeneity is primarily determined by two major aspects: typicity and severity. Typicity refers to the extent to which a patient’s symptoms align with those of TAD. Two
subtypes, LPAD and HSAD, fall within the AD spectrum but are distinct from TAD. LPAD and HSAD exhibit classic clinical and neuroimaging features of AD, but with different ages of onset, gender distributions, and educational levels. On the other hand, severity distinguishes MAD from TAD. Severity refers to the extent and rate of brain atrophy and cognitive decline. MAD is the least severe subtype and presents minimal brain atrophy despite a diagnosis of AD. In contrast, TAD is the most severe subtype, characterized by rapid cognitive decline and significant brain atrophy. These two aspects of typicality and severity play a crucial role in determining whether a person fits a particular AD subtype based on protective factors, risk factors, and concurrent non-AD brain pathologies. By providing a framework for understanding and characterizing AD subtypes, the distinction between typicality and severity is valuable for studying and treating the disease.

Researchers have employed a variety of neuroimaging modalities and imaging features to better comprehend and diagnose the different subtypes of AD. It should be noted that distinct imaging techniques capture different aspects of AD and reveal unique characteristics of the disease. One such study by Levin et al. (2021) utilized FDG-PET to measure metabolic changes in the brain and identified three hypometabolic AD subtypes: TAD, LPAD, and a relatively rare “cortical-predominant” subtype. A study by Sui et al. (2018) used an LDA approach to analyze whole brain white matter skeleton images and identified three latent factors in AD: a temporal-frontal impairment factor, a parietal factor, and a long fiber bundle factor. While each of these neuroimaging modalities provides unique insights into the disease, combining different neuroimaging techniques can provide a comprehensive understanding of AD subtypes and their underlying pathologies. For example, Jeon et al. (2019) categorized AD patients into three subtypes based on structural MRI, tau PET, and amyloid PET using AHC. The subtypes showed different cortical atrophy and tau deposition patterns, but no significant variation in amyloid deposition. In another recent study, Toledo et al. (2022) used both gray matter atrophy and tau PET to obtain separate clusters, with partial overlap detected between the two. Of note, the tau clusters displayed a stronger correlation with clinical progression. This finding may be attributed to the fundamental role of tau neurofibrillary tangles in spurring downstream neurodegeneration and cognitive decline in AD. Nevertheless, the relationship between tau burden and neurodegeneration can be regionally heterogeneous in select areas, resulting in unexpected atrophy levels in specific brain regions. To address this issue, Das et al. (2021) utilized AHC to group AD patients into six distinct clusters, modeling the variability in the tau-neurodegeneration relationship explicitly. This approach can identify alternative neurodegenerative contributors and provide a more comprehensive understanding of AD pathophysiology. In addition, several studies have utilized various data driven models and feature sets to investigate AD subtypes. For example, a recent study conducted by Zhang et al. (2022) examined the application of an identical subtype approach to voxel-based and surface-based morphological measures. The findings of this study suggested that comparable AD subtypes at the individual level could be achieved using either of these measures. In contrast, a study by Mohanty et al. (2020) compared the effectiveness of five different methods of subtyping AD. The study revealed that while each approach identified subtypes with similar demographic and clinical characteristics at the group level, there was a low level of individual overlap among the different methods. This finding emphasizes the need for further research to identify the most effective AD subtyping methods and to understand the underlying reasons for inconsistencies.

Efforts to identify AD subtypes have yielded inconsistent results, posing a significant challenge for achieving consistent subtype identification. One possible explanation for the inconsistency in AD subtype identification at both the population and individual levels is the intrinsic determinism of AD subtypes. Subtypes of AD may differ in their underlying genetic, environmental, or pathological factors, leading to variations in subtype identification. Additionally, distinct methods of subtype identification may place different emphases on the degrees of typicality and severity. These may lead to variations in subtype identification when compared across different methods. To address these inconsistencies, further research is needed to identify and validate the most effective methods for subtype identification. This research should incorporate rigorous and systematic analyses, accounting for the heterogeneity of the disease across different subpopulations. By elucidating the underlying causes of inconsistencies, researchers can develop more accurate and comprehensive approaches for subtyping AD, leading to improved diagnosis, treatment, and ultimately patient outcomes.

4.2 Biomarkers abnormal ordering and AD stages

The A/T/N framework plays a pivotal role in AD diagnosis and staging, identifying three specific biomarkers closely linked to AD pathology and providing essential diagnostic and disease staging evidence (Jack et al. 2018). Categorizing participants into four sections based on biomarker status, the A/T/N framework emphasizes the significance of
pathophysiological evidence from the three biomarkers for defining and staging AD. However, it is acknowledged that the clinical grading defined by the A/T/N framework may not be enough for accurate grading of disease progression. As a result, researchers have proposed an extension of the A/T/N framework: temporal heterogeneity, employing machine learning algorithms to detect variations in the abnormal ordering of biomarkers and distinguish different AD stages. The temporal heterogeneity approach offers a more precise method of assigning individuals into possible disease stages based on their biomarkers, granting a more detailed understanding of AD progression. By characterizing the temporal heterogeneity of the biomarker profile, it is possible to track a person's disease progression more accurately, leading to a more efficient evaluation of potential interventions' effectiveness.

Scientists aim to model the temporal heterogeneity of AD by analyzing the sequence of abnormalities in various biomarkers, including CSF, imaging, and cognitive-psychological scores, within disease progression. Through employing an EBM-based methodology, several studies (Venkatraghavan et al. 2021a; Young et al. 2014, 2015b) have identified similar abnormal patterns, with CSF biomarkers appearing abnormal before imaging biomarkers, and cognitive scores following suit. However, other studies (Archetti et al. 2019; Golriz Khatami et al. 2022; Oxtoby et al. 2018; Young et al. 2015a) present a different ordering of events. The variability in results may be attributed to multiple factors. Firstly, inconsistencies in AD severity estimates can stem from differing staging criteria used by various algorithms. For example, Archetti et al. (2019) compared two data-driven computational models for temporal heterogeneity and discovered varied event ordering. Both approaches indicated that the CSF biomarker became abnormal first, followed by cognitive scores, while imaging biomarkers showed significant regional atrophy at later stages of AD. However, the first aberrant biomarker in the EBM sequence was p-tau, while the DEBM sequence began with the $A\beta_{1-42}$ and $A\beta_{1-42}/p$-tau ratios. Furthermore, the models showed variations in the abnormal patterns of gray matter atrophy throughout the temporal lobes. In the DEBM approach, the abnormalities started at the hippocampus, entorhinal cortex, fusiform, and mid-temporal areas, moving to the precuneus region. In contrast, the EBM model showed that the primary abnormal imaging biomarker was the ventricles. The observed variations in the results might result from the different estimates of the Gaussian mixture model used in the two algorithms, as well as from the DEBM algorithm's built-in smoothing effect. Secondly, the order in which abnormal events occur may also be influenced by the subject cohorts used in different studies. Young et al. (2014) reported that $A\beta_{1-42}$, p-tau, and total tau were the abnormal CSF biomarker sequences for amyloid-positive or APOE-positive participants. However, in a larger population, total tau and p-tau were found to be detected earlier than $A\beta_{1-42}$. Recently, Golriz Khatami et al. (2022) adopted a derived EBM approach to compare 10 independent datasets. The results showed that event sequences were relatively consistent across most datasets, with CSF A$\beta$ being the first to change, followed by tauopathy, cognitive scores, FDG-PET, and ultimately abnormalities in specific brain regions. Nonetheless, minor differences were observed between cohorts, with the MRI regional event order exhibiting the most variation. It is essential to note that varying recruitment standards in different cohorts may lead to a high proportion of participants with a particular AD subtype or a cohort with AD patients with heterogeneous brain disease, which may explain the observed variations in the event sequences. These findings underscore the need for standardization in recruitment criteria and diagnostic processes to improve the reliability and reproducibility of findings in AD research.

### 4.3 Developments in spatial-temporal heterogeneity analysis for AD

Despite being important tools in the study of AD, either spatial or temporal heterogeneity has relatively limited abilities to explain individual progression and subtype differences. In many studies, the ADNI dataset and subjects in the early or middle stages of the disease were used, thus limiting the differences observed in disease progression. While the results of spatial heterogeneity may be influenced to some extent, the degree of influence is limited. On the other hand, the temporal heterogeneity results reflect the average timeline of biomarker progression in different subtypes. The variations in the subtype distribution of AD can impact the outcomes of both spatial and temporal heterogeneity analyses, with the latter being more susceptible to bias. These limitations call for the development of more advanced spatial and temporal heterogeneity methods to provide better insights into AD progression and subtyping. By integrating both spatial and temporal heterogeneity analyses, researchers can gain a deeper understanding of the progression of disease stages in different subtypes (Young et al. 2018).

Spatial-temporal heterogeneity methods are deemed the theoretical ideal for analyzing AD progression and subtype differences. Although one known spatial-temporal heterogeneity method, developed by Young et al. (2018), has displayed exceptional performance in various experiments, considerable work continues to be required in this
The joint optimization of temporal and spatial heterogeneity issues is a crucial aspect of advancing knowledge on AD progression and subtyping. To achieve this goal, the successful outcomes of temporal and spatial heterogeneity methods must be utilized.

In the SuStaIn analysis, the subtypes were achieved by unsupervised clustering. Unsupervised learning methods are associated with limitations such as vulnerability to noise and outliers, sensitivity to initial conditions, difficulty handling imbalanced data, and subjectivity. To address these concerns, researchers have proposed various solutions, including k-means++ (Aggarwal and Singh 2019) and spectral clustering (Varol et al. 2017), to mitigate sensitivity to initial conditions. To achieve better management of noisy and outlier-laden datasets, using robust clustering algorithms such as density-based spatial clustering of applications with noise (DBSCAN) (Ram et al. 2010) and ordering points to identify cluster structure (OPTICS) (Kanagala and Krishnaiah 2016) can be beneficial. Additionally, techniques such as oversampling and undersampling have been proposed to address imbalanced data. In contrast to unsupervised learning, semi-supervised methods can considerably improve the effectiveness of clustering in AD subtype identification as they reduce subjectivity by incorporating a control group. The incorporation of semi-supervised learning into spatial-temporal analysis can significantly enhance the efficacy and accuracy of AD subtyping. In the SuStaIn analysis, an earlier EBM approach (Fontein et al. 2011; Young et al. 2014) was used. However, this method faces scalability issues with many biomarkers. Several modifications and developments to the EBM have been introduced to enhance its efficacy. For instance, using the generalized Mallows model and Dirichlet process mixture of generalized Mallows models (Young et al. 2015a), they addressed the assumption of a single order of events for all subjects. Furthermore, new methods such as DEBM (Venkatraghavan et al. 2019) have been developed, which estimates the event order of a subject directly based on the posterior probability of a single biomarker becoming abnormal and estimates subgroups with different event sequences. The joint optimization of temporal and spatial heterogeneity issues requires building upon the successes of both temporal and spatial heterogeneity methods. Continued effort towards refining existing techniques and proposing new, scalable methods will lead to more accurate insights into AD progression and subtyping.

The spatial-temporal method combines both spatial and temporal heterogeneity but possesses some inherent weaknesses of both approaches. One of the main limitations is that the approach assumes that the trajectory of biomarkers increases monotonically as the disease progresses at the population level. This assumption can result in individuals being placed in a lower disease stage than their baseline in follow-up visits, which may lead to discontinuous or non-monotonic stage transitions. Another weakness arises from the number of biomarkers and test scores being used, which determine the number of stages, potentially resulting in redundant models. Furthermore, because longitudinal data is challenging to collect, cross-sectional designs are preferred by many researchers, even though such studies are often limited by between- and within-study variations, which interfere with drawing firm causal pathways. Despite these limitations, the spatial-temporal method has shown great promise as a powerful tool for AD heterogeneity research, potentially accelerating the development of AD-modifying therapies and enabling better personalized treatment options.

### 4.4 From omics to personalized treatments: the road to effective AD interventions

The current state of AD drug trials remains challenging and complex, despite significant efforts from pharmaceutical companies and researchers. Recent years have seen an increase in attention and funding devoted to AD research, leading to a promising landscape of current and ongoing drug trials (Cummings et al. 2023). However, to date, the US Food and Drug Administration (FDA) and European counterparts have only approved a limited number of drugs for the treatment of AD, underlining the significant unmet need for innovative and effective treatments. Specifically, the FDA approved two new drugs, aducanumab and lecanemab, in 2021 and 2023, respectively. Aducanumab and lecanemab are both fully human monoclonal antibodies that have been designed to specifically bind to aggregated forms of Aβ. These drugs have been developed with the aim of promoting the clearance of Aβ aggregates in the brain (Mintun et al. 2021; van Dyck et al. 2023). Clinical studies have shown that these drugs provide some relief to AD patients, but their efficacy is often limited in merely alleviating symptoms, instead of reversing disease pathology (Lane et al. 2011; Mintun et al. 2021; van Dyck et al. 2023). The development of effective drugs for AD has been hindered by a reliance on a one-size-fits-all approach in current drug trials (Devi and Scheltens 2018; Sohn et al. 2018). The variable responsiveness to treatments, due to multiple pathological mechanisms contributing to the disease to varying degrees, highlights the need for diverse therapeutic approaches, including personalized and stratified medicine.

To facilitate the successful implementation of precision medicine in AD, thorough analysis of the disease's
heterogeneity using multi-dimensional omics data and neuroimaging technology is required (Wang et al. 2021; Zhao et al. 2022). Neuroimaging modalities capture macroscopic differences in brain structure and function, while multi-omics data allow for microscopic investigations of various molecular signatures of AD. Combining the study of various omics disciplines, such as genomics, transcriptomics, proteomics, metabolomics, and epigenomics, with neuroimaging yields complementary information that sheds light on AD's heterogeneous mechanisms throughout the disease course. Genomics investigates an individual's genetic variants and their susceptibility to developing AD or related conditions (Argueta et al. 2022). For example, a particular brain atrophy pattern in a given AD subtype could be studied for genetic variants that give rise to a similar pattern. Transcriptomics examines gene expression patterns to identify AD-related biomarkers that differ among individuals (Iturria-Medina et al. 2022). Proteomics and metabolomics, on the other hand, explore changes in protein and metabolite levels in response to various stimuli or disease states, providing insight into the molecular mechanisms underlying AD's heterogeneity (Donnelly et al. 2019; Navas-Carrillo et al. 2021). Epigenomics analyzes modifications to chromatin structure or DNA methylation patterns, yielding crucial epigenetic information on the AD pathology at an individualized level (Alagiakrishnan et al. 2012). When these omics studies are combined with comprehensive neuroimaging techniques, researchers gain a more complete understanding of the multiple pathological mechanisms that contribute to AD's heterogeneity. This integrated approach can lead to improved classification of AD into subtypes and enable the identification of potential therapeutic targets unique to each subtype.

The integration of neuroimaging data with multi-dimensional omics data provides researchers with an unprecedented opportunity to comprehensively study the heterogeneous mechanisms underlying AD, paving the way for the development of personalized therapeutic interventions. However, the derivation of meaningful information from these large-scale, heterogeneous datasets poses significant computational challenges. Various analytical approaches, such as principal component analysis, canonical correlation analysis, and multivariate partial least square analysis, have been employed to dimensionally reduce significant data sets and elucidate overall correlations. A recent study (Lorenzini et al. 2022) has demonstrated the successful application of multivariate partial least square analysis in identifying regional associations between white matter hyperintensities and amyloid burden and their joint influence on cognitive processes. However, despite these promising advances, the successful translation of insights gained from integrated neuroimaging and omics studies into effective AD interventions necessitates overcoming significant computational, analytical, and interpretive challenges. Key amongst these is the integration and normalization of heterogeneous datasets, statistical standardization, and data-driven tools. As computational neurobiology rapidly advances, researchers need to develop reproducible, high-throughput, user-friendly, and effective platforms for studying integrated data to unlock the full potential of these datasets.

5 Conclusions

As the prevalence of AD rises, it has become a critical global public health issue and a significant area of research interest. One important aspect of this disease is its heterogeneity. This study aims to delve into data-driven approaches to investigate AD heterogeneity and provide novel insights into this critical issue. This review paper offers a detailed analysis of three data-driven methods for exploring AD heterogeneity, namely spatial heterogeneity, temporal heterogeneity, and spatial-temporal heterogeneity. These approaches are explained in-depth, and their respective strengths and limitations are explored to provide a comprehensive overview of the available techniques. Moreover, this study identifies new perspectives and proposes innovative ideas for future research. The research value of this study lies in the detailed exploration of the various methods for investigating AD heterogeneity. Researchers can make informed decisions on which method(s) to use in their studies by relying on the comprehensive analysis of the strengths and limitations of these methods presented in this paper. Furthermore, by providing novel insights, we identify new research directions, providing an impetus for further research in the field of AD heterogeneity. Overall, this review paper contributes to the body of knowledge on AD heterogeneity by providing an innovative perspective on the subject. It offers important implications for clinical practice, disease prevention, and public health policies by expanding our understanding of AD heterogeneity.

Author contributions: Lingyu Liu: Contributed to the literature search and writing of the main manuscript. Lan Lin: Contributed to the revision of the manuscript. Wenjie Kang, Shen Sun and Shucai Wu: Contributed to the idea of the manuscript.

Research funding: This research was financially supported by grants from National Natural Science Foundation of
China (81971683), Natural Science Foundation of Beijing Municipality (L182010).

Conflict of interest statement: The authors have no competing interests to declare that are relevant to the content of this article.

References


