Global Research journal of Natural Science and Technology (Grjnst)
Volume: 02 - Issue 3 (2024)
p.ISSN: 2790-7643 e.ISSN: 2790-7651
www.grjnst.com

Unraveling Neurodegeneration: Bioinformatics Approaches to Diagnosis and Biomarker Discovery

Sohaib Shakeel

Department of Zoology, University of Gujrat, Gujrat, Pakistan

sohaibshakeel974@gmail.com

Hajra Qaiser

Department of Zoology, University of Gujrat, Gujrat, Pakistan

hajraqaiser686@gmail.com

Hafiza Pakeeza Abid

Department of Zoology, University of Gujrat, Gujrat, Pakistan

hafizapakeezaabid@gmail.com

Hamna Saleem

Department of Zoology, University of Gujrat, Gujrat, Pakistan

hamnasaleem257@gmail.com

Shamsa Kausar

Department of Zoology, University of Kotli AK, Pakistan

shamsakousar93@gmail.com

Abstract

Neurodegenerative diseases (NDDs), including Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis, are progressive disorders characterized by irreversible neuronal loss and complex, multifactorial pathologies. Early and accurate diagnosis remains a critical challenge due to overlapping clinical symptoms and the absence of sensitive, disease-specific biomarkers. Bioinformatics has emerged as a transformative discipline, enabling the integration and analysis of high-throughput multi-omics data genomics, transcriptomics, proteomics, metabolomics, and epigenomics alongside neuroimaging features to uncover novel molecular signatures and diagnostic targets. Systems biology and network-based approaches facilitate the identification of disrupted pathways and gene-disease associations, while machine learning and artificial intelligence enhance predictive modeling and patient

stratification. This review synthesizes current advancements in bioinformatics-driven diagnostics for NDDs, highlighting tools for biomarker discovery, imaging-omics integration, and AI-powered decision support systems. We also address ongoing challenges, including data heterogeneity, standardization barriers, and ethical concerns, and propose strategic directions such as real-time bio sensing, federated learning, and unified neuroinformatics platforms. Collectively, this review underscores the pivotal role of bioinformatics in revolutionizing NDD diagnosis and advancing the vision of precision neurology.

Keywords: Neurodegeneration, Bioinformatics, Biomarker Discovery, Diagnosis

1. Introduction

Neurodegenerative diseases (NDDs), including Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS), represent a diverse group of chronic, progressive neurological disorders that are characterized by the selective degeneration of neurons in the central nervous system. These diseases manifest as cognitive and/or motor impairments, leading to a profound loss of independence and, ultimately, premature mortality (DeTure & Dickson, 2019). With aging populations and increasing life expectancy, the global prevalence of NDDs is projected to rise sharply in the coming decades, posing a growing public health and socioeconomic challenge. Despite decades of research, early and accurate diagnosis of NDDs remains a major clinical hurdle. The overlapping clinical features among different NDDs, such as memory loss, behavioural changes, and motor dysfunction, often lead to diagnostic uncertainty, especially during the prodromal or preclinical stages (Jack et al., 2018). Traditional diagnostic tools including neuropsychological testing, imaging techniques, and fluid biomarkers typically detect disease only after significant neurodegeneration has already occurred. For instance, PD is usually diagnosed after the emergence of motor symptoms, by which point over 60% of dopaminergic neurons in the substantia nigra may have already degenerated (Wired, 2024). Consequently, the window for effective disease-modifying interventions is severely limited. One of the critical unmet needs in the field is the identification of reliable early-stage and pre-symptomatic biomarkers that can enable differential diagnosis, monitor disease progression, and guide personalized therapeutic strategies. In this context, bioinformatics has emerged as a transformative discipline, bridging computational science and biology to interpret large-scale,

high-dimensional data generated through next-generation omics technologies. Platforms such as genomics, transcriptomics, proteomics, metabolomics, and epigenomics provide deep insights into the molecular landscape of NDDs, but it is through bioinformatics that these complex data are mined for biologically and clinically meaningful patterns (Torkamani et al., 2018). Recent advances in systems biology and network-based analyses have further strengthened the role of bioinformatics in deciphering disease mechanisms. For example, network biology approaches can reveal convergent molecular pathways and shared genetic risk loci among multiple NDDs, offering new avenues for cross-disease biomarker discovery and therapeutic repurposing (Del-Aguila et al., 2023). Tools such as DisGeNET, KEGG PATHWAY , and Neuro-D-Net facilitate the exploration of gene-disease associations, protein interaction networks, and disease-specific signaling cascades. These resources allow researchers to construct integrative molecular maps that underpin neurodegenerative processes and to identify key regulatory hubs that may serve as potential diagnostic targets.

In parallel, the convergence of bioinformatics with artificial intelligence (AI) and machine learning (ML) has opened new frontiers in predictive modeling and diagnostic tool development. Machine learning algorithms excel at identifying subtle, non-linear relationships in omics and clinical data, enabling the creation of robust, data-driven models for disease prediction and classification (Lu et al., 2022). Moreover, explainable AI (XAI) frameworks now offer the interpretability required for clinical adoption, fostering trust and transparency in diagnostic decision-making (Tonekaboni et al., 2019). Beyond molecular data, neuroimaging informatics has become an integral component of the diagnostic pipeline. Advanced image processing tools such as FreeSurfer, FSL, and SPM enable detailed analysis of brain structure and function, providing quantitative imaging biomarkers that can be correlated with molecular signatures to improve diagnostic specificity (Reimand et al., 2022). The integration of imaging data with multi-omics information through bioinformatics platforms supports a multimodal diagnostic approach, which is increasingly recognized as essential for capturing the full complexity of NDDs.

2. Omics Technologies and Data Integration

The integration of omics technologies through bioinformatics has transformed our understanding of neurodegenerative diseases (NDDs), providing a multilayered view of the molecular mechanisms involved in disease initiation and progression. Genomics plays a

fundamental role in identifying genetic variants associated with disease susceptibility. Techniques such as genome-wide association studies (GWAS), whole genome sequencing (WGS), and whole exome sequencing (WES) have uncovered numerous risk loci for disorders like Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS). Notably, GWAS has revealed variants in genes such as *APOE*, *MAPT*, and *TREM2* in AD, and *LRRK2* and *SNCA* in PD (Jansen et al., 2019). Bioinformatics tools such as GATK, ANNOVAR, and VEP are widely used for variant calling, annotation, and prioritization, enabling researchers to identify pathogenic variants with greater confidence (McKenna et al., 2018). Transcriptomics complements genomic data by measuring gene expression levels and uncovering regulatory dynamics in diseased tissues. Technologies such as RNA sequencing (RNA-seq) and microarrays have enabled the discovery of differentially expressed genes and alternative splicing events across NDDs. Tools like DESeq2, edgeR, and STAR aligner are commonly used to quantify transcript levels and detect aberrant transcription patterns that may underlie disease processes (Dobin et al., 2013). For instance, transcriptomic profiling in AD brain tissue has shown upregulation of inflammatory cytokines and downregulation of synaptic genes, implicating immune dysregulation and neuronal signaling deficits in disease pathology (Mathys et al., 2019). Proteomics and metabolomics offer further insights into disease phenotypes by characterizing changes in protein abundance and metabolic pathways. Mass spectrometry-based proteomics has identified panels of protein biomarkers in cerebrospinal fluid (CSF) that reflect tau pathology, synaptic dysfunction, and neuroinflammation (Brinkmalm et al., 2023). Moreover, protein interaction networks generated using tools like STRING and Cytoscape can highlight hub proteins that may serve as therapeutic targets. In parallel, metabolomic profiling has revealed disruptions in glucose metabolism, lipid signaling, and neurotransmitter pathways in NDDs, particularly in AD and PD, suggesting metabolic dysregulation as a core feature of disease (Wilkins & Trushina, 2018). Epigenomics adds another crucial layer by capturing modifications to chromatin structure that regulate gene expression without altering DNA sequence. DNA methylation profiling has identified epigenetic marks associated with aging and neurodegeneration, particularly in the promoters of immune and neuronal genes (Smith et al., 2021). Non-coding RNAs, including microRNAs and long non-coding RNAs, also

contribute to post-transcriptional gene regulation in the diseased brain. Tools such as Bismark, MethyKit, and EpiDISH are extensively used to analyze methylation and epigenomic datasets.

The integration of multi-omics data represents a powerful approach for capturing the complexity of NDDs. Data fusion techniques and machine learning models, including random forests and neural networks, are increasingly used to integrate heterogeneous data types and identify molecular signatures predictive of disease states (Zitnik et al., 2019). Network-based methods, such as similarity network fusion (SNF) and multi-layered network modeling, facilitate the correlation of genomic, transcriptomic, proteomic, and metabolomic data to identify common regulatory modules and prioritize biomarker candidates. These integrative strategies not only enhance diagnostic accuracy but also inform precision medicine efforts aimed at individualized therapeutic interventions.

3. Biomarker Discovery Using Bioinformatics

The discovery of reliable biomarkers is a critical goal in the study of neurodegenerative diseases (NDDs), as they offer potential for early diagnosis, prognosis, and therapeutic monitoring. Biomarkers can be classified into three primary categories: **diagnostic biomarkers**, which facilitate early detection; **prognostic biomarkers**, which predict disease progression; and **predictive biomarkers**, which assess response to specific therapies (Hampel et al., 2018). Traditional biomarker discovery has relied heavily on experimental methods; however, the growing availability of multi-omics datasets and computational tools has shifted the paradigm towards **bioinformatics-driven biomarker identification**.

Computational methods, including differential gene expression analysis, network-based approaches, and meta-analysis of publicly available datasets, have enabled the identification of candidate biomarkers across different biological layers genomic, transcriptomic, proteomic, and metabolomics. For example, tools such as GEO2R, limma, and WGCNA allow researchers to explore co-expression networks and identify gene modules associated with disease phenotypes (Langfelder & Horvath, 2019). In AD, transcriptomic analyses using these methods have highlighted *APP*, *BACE1*, and *MAPT* as consistently dysregulated genes (Mostafavi et al., 2018). Similarly, in ALS, computational studies have identified *TARDBP* and *FUS* as key biomarkers through integrative bioinformatics pipelines (Liscic et al., 2020).

The integration of **artificial intelligence (AI)** and **machine learning (ML)** has further revolutionized biomarker screening. ML algorithms such as support vector machines (SVM), random forests, and deep neural networks can learn from complex, high-dimensional datasets to classify disease states and prioritize candidate biomarkers with high accuracy (Libbrecht & Noble, 2015). Recent studies have demonstrated the power of ML in distinguishing early-stage AD patients from cognitively normal individuals using transcriptomic and imaging data (Ansart et al., 2021). In PD, AI-driven models have been used to identify blood-based gene expression signatures that differentiate between disease subtypes and predict disease severity (Glaab & Schneider, 2021). These tools not only increase the efficiency of biomarker discovery but also provide predictive models that may guide clinical decision-making.

Several **validated bioinformatics-derived biomarkers** have already made their way into clinical and preclinical pipelines. For instance, *amyloid-beta 42 (Aβ42)* and *phosphorylated tau (p-tau),* initially identified through proteomic screening and validated by large-scale cohorts like ADNI, are now core CSF biomarkers for Alzheimer's diagnosis (Hansson et al., 2018). In PD, α-synuclein levels in CSF and specific methylation changes in *SNCA* gene promoters have shown diagnostic potential. For ALS, neurofilament light chain (NfL) has emerged as a robust biomarker identified through proteogenomic integration approaches (Benatar et al., 2022).

In conclusion, bioinformatics has drastically improved the landscape of biomarker discovery in NDDs by integrating complex datasets, applying sophisticated computational algorithms, and leveraging AI/ML to accelerate the identification and validation of reliable biomarkers. Continued advances in data integration and model interpretability will be essential in translating these discoveries into clinical practice.

4. Neuroimaging Informatics and Data Mining

Neuroimaging informatics has emerged as a vital discipline in the study of neurodegenerative diseases (NDDs), offering the ability to visualize structural, functional, and molecular changes in the brain. The integration of neuroimaging techniques such as magnetic resonance imaging (MRI) and positron emission tomography (PET) with high-throughput omics data (genomics, transcriptomics, and proteomics) allows for a systems-level understanding of disease progression. This combined approach, often referred to as **imaging-omics**, enables the identification of molecular correlates of neuroanatomical abnormalities and provides

insight into disease mechanisms (Reimand et al., 2022). In particular, **imaging-genomics** the intersection of genetic data with neuroimaging phenotypes has shown considerable promise in linking gene variants to regional brain changes. Genome-wide association studies (GWAS) and polygenic risk scores have been correlated with neuroimaging metrics such as cortical thickness and white matter integrity in Alzheimer's and Parkinson's disease cohorts (Hu et al., 2023). For example, recent studies have confirmed that APOE ε4 carriers exhibit reduced hippocampal volume and increased amyloid deposition even in preclinical stages, demonstrating the potential of combining PET/MRI imaging with genetic markers for early diagnosis (Shang et al., 2023). Several bioinformatics tools and platforms facilitate this multimodal data analysis. **FreeSurfer** remains a widely used tool for cortical and subcortical segmentation, producing quantifiable imaging phenotypes that can be integrated with genomic data (Fischl et al., 2021). **SPM (Statistical Parametric Mapping)** supports voxel-wise statistical analyses of brain imaging data, and is commonly used in conjunction with **BrainNet Viewer** to visualize brain networks and their association with transcriptomic signatures (Xia et al., 2022). Additionally, machine learning-based pipelines like **NeuroMiner** and **ImagingGenetics** have enabled the discovery of imaging-genomic biomarkers using large datasets such as ADNI and UK Biobank (Bhalla et al., 2023).

5.Machine Learning and Artificial Intelligence in Diagnosis

The advent of **machine learning (ML)** and **artificial intelligence (AI)** has revolutionized the landscape of biomedical diagnostics, especially in the context of complex and multifactorial diseases such as neurodegenerative disorders (NDDs). Traditional diagnostic techniques, although valuable, often rely heavily on clinical symptoms that emerge only in the advanced stages of disease. These symptoms can be highly variable, overlapping across diseases like Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS), making early and differential diagnosis particularly challenging. AI-driven models offer a transformative solution by learning from large-scale, heterogeneous datasets ranging from omics (genomics, transcriptomics, proteomics) to clinical records and neuroimaging data thus enabling earlier, more accurate, and reproducible diagnostic outcomes (Kim et al., 2023). At the heart of these advances are **supervised ML algorithms** such as support vector machines (SVM), random forests, k-nearest neighbors (k-NN), and ensemble learning

techniques. These methods have been extensively applied to classify disease stages, predict progression, and identify risk profiles from high-dimensional data. For instance, studies have demonstrated the effective use of SVMs and logistic regression models to predict the conversion from mild cognitive impairment (MCI) to Alzheimer's disease based on a combination of cerebrospinal fluid (CSF) biomarkers, neuroimaging features, and cognitive assessments (Huang et al., 2021; Choi et al., 2023). Feature selection and dimensionality reduction techniques such as principal component analysis (PCA) and t-SNE are often applied in these pipelines to optimize model performance and interpretability.

In recent years, **deep learning (DL)**, a subfield of ML, has emerged as a particularly powerful tool in neurodegenerative research. Convolutional neural networks (CNNs), for example, are widely used for the automated analysis of brain imaging data such as structural MRI, functional MRI (fMRI), and PET scans. These models are capable of learning hierarchical spatial features directly from raw images, allowing for fine-grained detection of disease-specific patterns that may not be discernible to the human eye. CNN-based approaches have achieved classification accuracies of over 90% in distinguishing AD patients from cognitively normal individuals (Lu et al., 2022). Hybrid models that combine CNNs with recurrent neural networks (RNNs) or attention mechanisms have further improved the detection of temporal patterns in longitudinal datasets, offering insights into disease progression.

Another emerging area involves **graph neural networks (GNNs)** and spatiotemporal models that incorporate brain connectivity data and network-level features. These models treat the brain as a graph of interconnected nodes (regions) and have shown superior performance in capturing global and local alterations in brain structure and function, particularly in early or preclinical stages of NDDs (Kim et al., 2023).

Beyond individual model development, AI is also being embedded into **diagnostic decision support systems (DDSS)**. These systems are designed to assist clinicians by integrating diverse and complex data types including electronic health records (EHRs), genetic profiles, imaging data, and even wearable sensor outputs into a unified framework that suggests differential diagnoses, flags high-risk individuals, and recommends follow-up evaluations. Importantly, explainable AI (XAI) techniques are being incorporated into these tools to

enhance transparency and build clinician trust by providing interpretable results and highlighting key decision-driving features (Nassif et al., 2023).Moreover, the growing availability of open-access databases such as the Alzheimer's Disease Neuroimaging Initiative (ADNI), Parkinson's Progression Markers Initiative (PPMI), and UK Biobank has provided an unprecedented volume of standardized data for model training, validation, and benchmarking. These resources have facilitated collaborative, reproducible research and accelerated the deployment of AI-driven diagnostics in clinical settings.

6. Challenges and Limitations

Despite the rapid progress in bioinformatics and its application to the diagnosis of neurodegenerative diseases (NDDs), several formidable challenges and limitations continue to hinder its full clinical integration. One of the most prominent issues is the variability and inconsistency in data quality across different omics platforms and neuroimaging modalities. Datasets often originate from diverse populations, experimental conditions, and technologies such as varying sequencing platforms or imaging protocols leading to batch effects, missing values, and normalization biases that can severely compromise data interpretation and model performance (Johnson et al., 2023). Moreover, a lack of universal data standards and poor metadata annotation further limit the reproducibility and interoperability of bioinformatics workflows, complicating large-scale multi-center studies or meta-analyses. The promise of early diagnosis through biomarker discovery is also significantly constrained by the limited number of clinically validated biomarkers. While numerous computational studies have proposed potential biomarkers using transcriptomics, proteomics, and metabolomics data, only a few such as cerebrospinal fluid (CSF) amyloid-beta and tau proteins for Alzheimer's disease have met the stringent criteria for clinical application. Many studies are limited by small cohort sizes, inadequate replication, and population biases, reducing the reliability and generalizability of proposed biomarkers (Rabinovici et al., 2023). Additionally, the integration of artificial intelligence (AI) and machine learning (ML) techniques, although powerful, introduces challenges of its own. Most deep learning models, including convolutional neural networks (CNNs) applied to brain scans or multi-omics datasets, operate as "black boxes" with limited transparency, making it difficult for clinicians to understand or trust their decision-making processes (Arrieta et al., 2020). This lack of explain ability

reduces the clinical acceptance of these tools, especially in high-stakes diagnostic scenarios. Ethical and legal considerations also play a critical role in limiting the application of bioinformatics in NDDs. The handling of sensitive patient data such as genomic sequences, brain images, and longitudinal health records raises significant privacy concerns, particularly regarding data ownership, informed consent, and the potential for re-identification even in anonymized datasets (Price & Cohen, 2019). These concerns are compounded by the risk of algorithmic bias, where predictive models may perform disproportionately poorly in underrepresented demographic groups, exacerbating existing health disparities. Finally, there exists a notable gap between computational discovery and clinical practice. Many bioinformatics tools and AI models are developed in controlled research environments with curated datasets, but fail to maintain their accuracy and utility in real-world clinical settings, where data may be incomplete, inconsistent, or confounded by comorbid conditions (Topol, 2023). Bridging this translational divide requires not only interdisciplinary collaboration between bioinformaticians, clinicians, and regulatory bodies but also robust infrastructure to support data integration, decision support systems, and continuous validation in diverse patient populations. Without addressing these multifaceted challenges, the translation of bioinformatics-driven discoveries into meaningful clinical interventions for neurodegenerative diseases will remain a formidable task.

7. Future Directions

As the field of neurodegenerative disease (NDD) research continues to evolve, several innovative and interdisciplinary directions are shaping the future of bioinformatics-driven diagnostics and interventions. One transformative avenue is the **integration of wearable technology and real-time health monitoring**, which enables the continuous capture of physiological, behavioural, and environmental data in naturalistic settings. Wearable devices such as smartwatches, motion sensors, EEG headbands, and voice-enabled digital biomarkers can detect subtle motor, cognitive, and sleep pattern changes that may signal the early onset of conditions like Parkinson’s disease or Alzheimer’s disease (Lipsmeier et al., 2023). Bioinformatics platforms that incorporate this data can support continuous disease monitoring, digital phenotyping, and timely therapeutic intervention, moving diagnosis beyond clinic-based snapshots into dynamic, patient-centered care.

Simultaneously, the shift toward **precision medicine and personalized diagnosis** is becoming increasingly feasible with advances in omics technologies and integrative analytics. Combining genomics, transcriptomics, epigenomics, and metabolomics with patient-specific clinical and lifestyle data allows for stratification of individuals by genetic risk, disease subtypes, and likely treatment responses (Hampel et al., 2021). Machine learning models are being developed to generate individualized risk profiles and progression trajectories, laying the groundwork for personalized treatment algorithms that could significantly improve outcomes in heterogeneous diseases like Alzheimer's and ALS (Fazal et al., 2022).

Artificial intelligence (AI) itself continues to advance rapidly, with new learning paradigms like **federated learning** offering scalable, privacy-preserving solutions to the problem of data silos in healthcare. In federated learning, models are trained collaboratively across multiple institutions without exchanging raw patient data preserving privacy while harnessing the power of large, diverse, and geographically distributed datasets (Rieke et al., 2020). This approach is especially valuable for NDDs, where data scarcity and population bias have historically limited the development of generalizable diagnostic models.

Finally, there is a pressing need for the development of **comprehensive, interoperable neuroinformatics platforms** capable of integrating multimodal data including imaging, omics, clinical records, and wearable data within unified analytical ecosystems. Such platforms should support not only data storage and access but also advanced visualization, AI-powered analytics, and collaborative research features. Ongoing efforts like the European Human Brain Project and the NIH BRAIN Initiative are already laying the groundwork for these infrastructures (Amunts et al., 2022). These systems will be essential in bridging the translational gap between bench and bedside by enabling scalable biomarker discovery, real-world clinical decision support, and the deployment of personalized therapeutic strategies across healthcare systems.

8. Recommendations

To fully realize the diagnostic and translational potential of bioinformatics in neurodegenerative diseases (NDDs), several strategic actions are necessary. Standardizing data acquisition protocols across multi-omics and neuroimaging platforms is essential to

reduce variability and enhance reproducibility. The implementation of unified data formats, metadata standards, and interoperable databases will facilitate more reliable multi-center studies and integrative analyses. Equally important is increasing the scale and diversity of patient cohorts through international collaborations and the inclusion of underrepresented populations, which will enhance the generalizability and equity of biomarker discovery. Bridging the translational gap between computational modelling and clinical practice is also critical; this involves co-developing diagnostic tools with clinician input, validating them in real-world settings, and embedding them within electronic health record (EHR) systems for seamless clinical integration. Furthermore, bioinformatics pipelines should leverage explainable artificial intelligence (XAI) to promote transparency, interpretability, and trust among healthcare providers and patients. Strengthening ethical safeguards is imperative, especially concerning data privacy, security, and algorithmic fairness in the handling of sensitive genomic and neuroimaging data. Finally, sustained investment in interdisciplinary education and training programs will be vital to equip researchers, clinicians, and data scientists with the skills required to co-develop and utilize next-generation neuroinformatics tools. These coordinated efforts will be instrumental in advancing precision diagnostics, accelerating therapeutic innovation, and ultimately improving clinical outcomes for individuals affected by NDDs

9. Conclusion

Bioinformatics has revolutionized the diagnosis and understanding of neurodegenerative diseases by enabling the integration of multi-omics data, neuroimaging, and machine learning to uncover novel biomarkers and build predictive models for early detection and personalized care. These advances promise improved disease classification and targeted therapies; however, challenges such as data heterogeneity, limited clinical translation, and ethical concerns persist. Addressing these issues through standardization, validation, and interdisciplinary collaboration is essential to fully realize the clinical potential of bioinformatics. As the field evolves, it will play an increasingly vital role in advancing precision medicine and improving outcomes for patients with neurodegenerative diseases.

References

- Amunts, K., et al. (2022). The Human Brain Project: Creating a European research infrastructure to decode the human brain. *Neuron*, 110(16), 2573–2585.
- Fazal, M. I., Patel, M. E., Toth, P., & Matyas, B. (2022). Machine learning and deep learning applications in precision medicine for neurodegenerative diseases. *Journal of Personalized Medicine*, 12(7), 1045.
- Hampel, H., Cummings, J., Blennow, K., et al. (2021). Developing the ATX(N) classification system for precision medicine in Alzheimer's disease. *Nature Reviews Neurology*, 17(9), 580–589.
- Lipsmeier, F., et al. (2023). Remote monitoring of Parkinson's disease progression using smartphone-based assessments: a 24-month observational study. *NPJ Parkinson's Disease*, 9(1), 21.
- Mirelman, A., et al. (2022). Digital health for monitoring mobility in Parkinson's disease: a multinational longitudinal study. *NPJ Digital Medicine*, 5(1), 126.
- Arrieta, A. B., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.
- Johnson, S. C., et al. (2023). Harmonizing big data for neurodegeneration research: Strategies and tools. *Frontiers in Neuroscience*, 17, 1152742.
- Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. *Nature Medicine*, 25(1), 37–43.
- Rabinovici, G. D., et al. (2023). Biomarker validation in Alzheimer's disease: Challenges and strategies. *Nature Reviews Neurology*, 19(1), 15–30.
- Topol, E. J. (2023). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 29, 1090–1101.
- Wang, L., et al. (2022). Harmonization strategies for integrating heterogeneous omics data in neurodegeneration research. *Briefings in Bioinformatics*, 23(4), bbac236.
- Woo, C. W., et al. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*, 20(3), 365–377

- Choi, H., et al. (2023). A machine learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease using multimodal biomarkers. *Scientific Reports*, 13, 1089.
- Huang, Y., et al. (2021). Machine learning-based prediction of Alzheimer's disease using neuroimaging and genetic data. *Frontiers in Aging Neuroscience*, 13, 682110.
- Kim, H., et al. (2023). Graph neural network-based deep learning model for early diagnosis of Alzheimer's disease using multimodal imaging data. *NeuroImage*, 273, 120062.
- Li, F., et al. (2022). An interpretable machine learning framework for early prediction of Parkinson's disease using genomic and clinical data. *Briefings in Bioinformatics*, 23(1), bbab408.
- Lu, D., et al. (2022). Deep learning for early prediction of Alzheimer's disease progression using brain imaging. *Translational Psychiatry*, 12, 146.
- Nassif, H., et al. (2023). An AI-powered clinical decision support system for differential diagnosis of neurodegenerative disorders. *NPJ Digital Medicine*, 6, 72
- Bhalla, M., et al. (2023). Deep learning-based multimodal fusion for early Alzheimer's disease diagnosis using imaging and genetics data. *NeuroImage: Clinical*, 38, 103408.
- Fischl, B., et al. (2021). Advances in FreeSurfer: Recent improvements in automated brain MRI analysis. *NeuroImage*, 238, 118273.
- Hu, J., et al. (2023). Genomic associations of brain imaging phenotypes reveal shared biology in neurodegenerative diseases. *Nature Communications*, 14(1), 1359.
- Mueller, S. G., et al. (2022). Longitudinal imaging-genetics study of hippocampal volume and APOE genotype in aging. *Alzheimer's Research & Therapy*, 14, 81.
- Reimand, J., et al. (2022). Systems-level integration of imaging and omics in brain disorders. *Nature Reviews Neuroscience*, 23(9), 589–606.
- Shang, J., et al. (2023). Linking genetic risk and neuroimaging changes across Alzheimer's disease continuum using machine learning. *Molecular Psychiatry*.
- Wang, L., et al. (2023). Omics and neuroimaging integration for early detection of Alzheimer's disease. *Translational Psychiatry*, 13, 206.

- Wen, J., et al. (2021). Multi-scale brain imaging genetic analysis in 19,000 UK Biobank participants. *Nature Communications*, 12, 4918.
- Xia, M., Wang, J., & He, Y. (2022). BrainNet Viewer: updated features for visualization of brain connectomics. *Frontiers in Neuroscience*, 16, 876251.
- Zhao, Y., et al. (2022). NeuroMiner: A machine learning framework for multimodal biomarker discovery in neurodegenerative diseases. *Bioinformatics*, 38(15), 3768–3775.
- Ansart, M., Epelbaum, S., Bassignana, G., et al. (2021). Predicting cognitive scores with graph neural networks in Alzheimer's disease. *Frontiers in Aging Neuroscience, 13*, 734807.
- Benatar, M., Wuu, J., Andersen, P. M., et al. (2022). Neurofilament light: A candidate biomarker of presymptomatic ALS and phenoconversion. *Annals of Neurology, 91*(2), 276–285.
- Glaab, E., & Schneider, R. (2021). Machine learning for precision medicine in neurodegenerative diseases. *Trends in Molecular Medicine, 27*(4), 347–362.
- Hampel, H., O'Bryant, S. E., Molinuevo, J. L., et al. (2018). Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic. *Nature Reviews Neurology, 14*(11), 639–652.
- Hansson, O., Seibyl, J., Stomrud, E., et al. (2018). CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimer's & Dementia, 14*(11), 1470–1481.
- Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics, 9*, 559.
- Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics, 16*(6), 321–332.
- Liscic, R. M., Bagaric, D., Crovella, S., & Cecchi, C. (2020). Neurofilament light chain as a biomarker for diagnosis, disease progression, and prognosis of amyotrophic lateral sclerosis. *Journal of Cellular and Molecular Medicine, 24*(9), 5201–5207.

- Del-Aguila, J. L., Li, Z., Fernandez, M. V., et al. (2023). Computational analysis of five neurodegenerative diseases reveals shared and specific genetic loci. *Frontiers in Neuroscience, 17*, 10651457.
- Guo, W., Zhang, Z., Wei, Y., & Chen, H. (2021). Emerging roles of mitochondrial dysfunction in neurodegenerative diseases. *Biomedicine & Pharmacotherapy, 133*, 110968.
- Hwang, D., Lee, I. Y., Yoo, H., Gehlenborg, N., Cho, J. H., Petritis, B., ... & Hood, L. (2009). A systems approach to prion disease. *Molecular Systems Biology, 5*, 252.
- Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., & Tanabe, M. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research, 45*(D1), D353–D361.
- Kumar, D., Ambasta, R. K., & Kumar, P. (2017). Dissecting the molecular mechanisms of neurodegenerative diseases through network biology. *Frontiers in Neuroscience, 11*, 660.
- Piñero, J., Bravo, À., Queralt-Rosinach, N., Gutiérrez-Sacristán, A., Deu-Pons, J., Centeno, E., ... & Furlong, L. I. (2017). DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Research, 45*(D1), D833–D839.
- Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, 13(10), R87.
- Blauwendraat, C., Nalls, M. A., & Singleton, A. B. (2020). The genetic architecture of Parkinson's disease. *The Lancet Neurology*, 19(2), 170–178.
- Brinkmalm, G., et al. (2023). Cerebrospinal fluid biomarkers for Alzheimer’s and Parkinson’s diseases: state of the art and future perspectives. *Nature Reviews Neurology*, 19, 195–212.
- Dobin, A., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15–21.

- Jansen, I. E., et al. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature Genetics*, 51(3), 404–413.
- Krueger, F., & Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*, 27(11), 1571–1572.
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550.
- Mathys, H., et al. (2019). Single-cell transcriptomic analysis of Alzheimer’s disease. *Nature*, 570(7761), 332–337.
- McKenna, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303.
- Smith, A. R., Smith, R. G., Burrage, J., Troakes, C., Al-Sarraj, S., Kalaria, R. N., ... & Mill, J. (2021). Epigenetic changes in Alzheimer’s disease: DNA methylation, histone modifications and microRNA regulation. *Trends in Neurosciences*, 44(9), 747–765.
- Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16), e164.
- Wilkins, J. M., & Trushina, E. (2018). Application of metabolomics in Alzheimer’s disease. *Frontiers in Neurology*, 8, 719.
- Zitnik, M., Nguyen, F., Wang, B., Leskovec, J., Goldenberg, A., & Hoffman, M. M. (2019). Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. *Information Fusion*, 50, 71–91.
- Jin, B., Fei, G., Sang, S., & Zhong, C. (2023). Identification of biomarkers differentiating Alzheimer’s disease from other neurodegenerative diseases by integrated bioinformatic analysis and machine-learning strategies. *Frontiers in Molecular Neuroscience*, 16, 1152279.
- Selvam, S. (2024). Biomarkers in neurodegenerative diseases: a broad overview. *Exploration of Neuroprotective Therapy*, 4, 119–147. Winchester, L. M., Harshfield,

E. L., Shi, L., et al. (2023). Artificial intelligence for biomarker discovery in Alzheimer's disease and dementia. *Alzheimer's & Dementia*, 19(12), 5860–5871.