**Project title: Multimodal Integration of Proteomics and Neuroimaging for Advancing Biomarker Discovery in Alzheimer’s Disease**

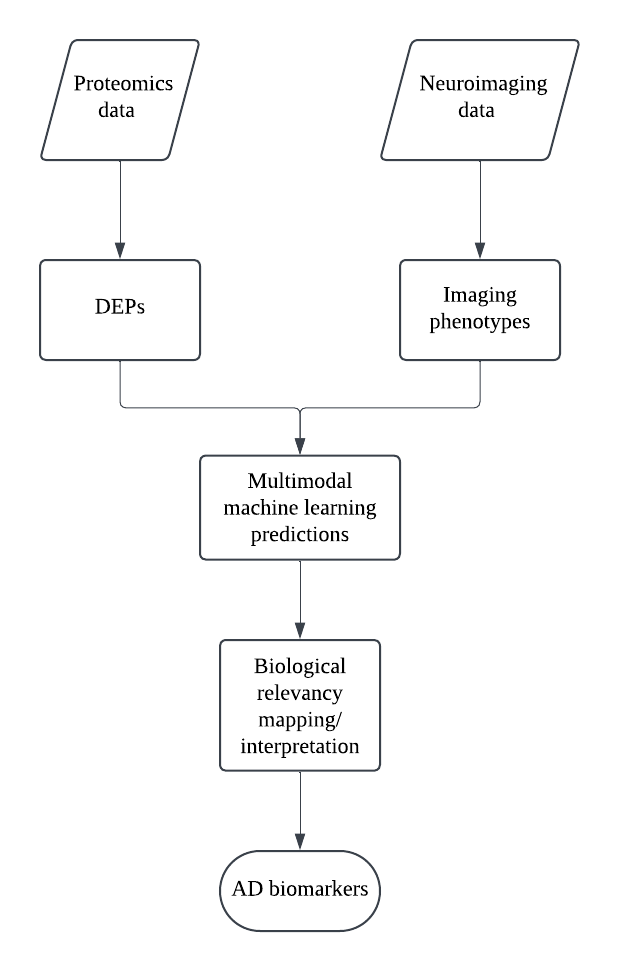
|  |
| --- |
| **A5. Provide an executive summary of the research (**max. 300 words).  This project aims to investigate the multimodal integration of proteomics data with brain imaging models in Alzheimer’s disease (AD) to enhance the understanding of disease mechanisms and improve diagnostic accuracy. Advances in proteomics and neuroimaging have greatly improved our understanding of neurodegenerative conditions like AD. Proteomics, the large-scale study of proteins and their functions, has provided insights into molecular processes, with specific proteomic signatures identified as biomarkers for early diagnosis and disease progression. Simultaneously, neuroimaging techniques, such as MRI and PET, offer high-resolution, non-invasive assessments of brain structure and function, with imaging phenotypes correlating to underlying biological processes. Recent studies have demonstrated the potential of integrating proteomic data with neuroimaging metrics through computational approaches, revealing significant associations that deepen our understanding of disease mechanisms.  Building on this growing body of work, this project will leverage advanced machine learning techniques to conduct integrative analyses of proteomic profiles and neuroimaging data, aiming to identify key biomarkers associated with AD progression. The integration of these different types of data modalities is expected to uncover novel insights into the underlying biological processes of AD, facilitating early detection and treatment strategies. An interdisciplinary team of proteomics, neuroimaging, and data science will employ advanced computational methods to develop predictive models that link specific proteomic biomarkers to imaging phenotypes.  The anticipated outcomes include the identification of potential biomarkers for early diagnosis, improved understanding of AD pathology, and the establishment of a framework for future research in multimodal data integration. Funding is sought specifically to support the employment of a research assistant to contribute to this transformative initiative, aiming to significantly advance AD research and contribute to more effective interventions and improved patient outcomes. |

**Part B: Detailed proposal**

|  |
| --- |
| **B1. Research background including Hypothesis/Research questions and Literature reviews**  Alzheimer's Disease (AD) is a progressive neurodegenerative disorder marked by complex pathological processes, including amyloid-beta accumulation, tau pathology, and neuroinflammation. While significant advancements have been made in understanding these mechanisms, there remains a pressing need for early diagnostic markers and integrative models that bridge molecular and structural data to enhance disease characterization.  Proteomics, the large-scale study of proteins, has emerged as a powerful tool for identifying molecular signatures associated with AD. Studies by Desaire et al. (2022) and Wang et al. (2023) have demonstrated that specific protein profiles can serve as biomarkers for early detection, disease progression, and therapeutic monitoring. Concurrently, advancements in neuroimaging techniques, such as MRI and PET, have provided high-resolution insights into structural and functional brain changes, enabling non-invasive assessments of disease progression (Kim et al., 2022; Singh and Kumar, 2024).  Recent research started to explore the integration of proteomics and neuroimaging data to provide a more comprehensive understanding of AD. Casanova et al. (2024) utilized machine learning algorithms to reveal significant associations between proteomic profiles and neuroimaging phenotypes, highlighting the potential of interdisciplinary approaches to uncover novel disease mechanisms. Moreover, However, there remains a gap in leveraging these integrated datasets to develop predictive models for AD diagnosis and progression.  While multi-modal integration has gained attention, the majority of existing studies have focused on combining neuroimaging with genetic data, such as DNA variants and SNPs (Meng et al., 2022; Oh et al., 24; Vilkaite et al., 2024). Among the limited studies that have integrated proteomics within multimodal frameworks, the work by Nazeri et al. (2014) stands out as an early example, even though it was conducted over a decade ago and using simple analytics methods like t-test, logistic regression and Parallel Independent Component Analysis (PICA).  Additionally, although amyloid-beta and tau proteins are well-studied in AD research, broader proteomic profiles (e.g., signalling proteins, inflammatory markers, metabolic enzymes) remain underexplored in relation to neuroimaging phenotypes. This creates a critical knowledge gap in understanding how diverse protein alterations map onto structural and functional brain changes in AD.  We hypothesize that specific proteomic signatures are significantly correlated with neuroimaging phenotypes, and that these relationships can be uncovered using advanced computational modelling to improve predictive accuracy. This study aims to address current limitations by integrating proteomics and neuroimaging data using machine learning techniques, thereby advancing our understanding of AD pathophysiology and contributing to the development of early diagnostic tools and personalized therapeutic strategies.  Reference:  Casanova, R., Walker, K. A., Justice, J. N., Anderson, A., Duggan, M. R., Cordon, J., ... & Hughes, T. M. (2024). Associations of plasma proteomics and age-related outcomes with brain age in a diverse cohort. *GeroScience*, 1-13.  Desaire, H., Stepler, K. E., & Robinson, R. A. (2022). Exposing the brain proteomic signatures of Alzheimer’s disease in diverse racial groups: Leveraging multiple data sets and machine learning. *Journal of proteome research, 21*(4), 1095-1104.  Kim, J. S., Han, J. W., Bae, J. B., Moon, D. G., Shin, J., Kong, J. E., ... & Kim, K. W. (2022). Deep learning-based diagnosis of Alzheimer’s disease using brain magnetic resonance images: An empirical study. *Scientific Reports, 12*(1), 18007.  Meng, X., Liu, J., Fan, X., Bian, C., Wei, Q., Wang, Z., ... & Jiao, Z. (2022). Multi-modal neuroimaging neural network-based feature detection for diagnosis of Alzheimer’s disease. *Frontiers in Aging Neuroscience, 14*, 911220.  Oh, S., Kim, S., Lee, J. E., Park, B. Y., Won, J. H., & Park, H. (2024). Multimodal analysis of disease onset in Alzheimer’s disease using Connectome, Molecular, and genetics data. *NeuroImage: Clinical, 43*, 103660.  Singh, A., & Kumar, R. (2024). Brain MRI Image Analysis for Alzheimer’s Disease (AD) Prediction Using Deep Learning Approaches. *SN Computer Science, 5*(1), 160.  Vilkaite, Gabriele, Jacob Vogel, and Niklas Mattsson-Carlgren. "Integrating amyloid and tau imaging with proteomics and genomics in Alzheimer’s disease." *Cell Reports Medicine 5.9* (2024).  Wang, Y., Sun, Y., Wang, Y., Jia, S., Qiao, Y., Zhou, Z., ... & Peng, D. (2023). Identification of novel diagnostic panel for mild cognitive impairment and Alzheimer’s disease: findings based on urine proteomics and machine learning. *Alzheimer's Research & Therapy, 15*(1), 191. |
| **B2. Objectives of the research:**   1. To develop an analytical framework that combines proteomic data with neuroimaging in facilitating a comprehensive understanding of the biological processes underlying neurodegenerative diseases. 2. To apply advanced computational methods, such as machine learning and deep learning algorithms, in constructing predictive models that correlate specific proteomic signatures with neuroimaging phenotypes. 3. To identify specific proteomic biomarkers that are predictive in neuroimaging phenotypes associated with neurodegenerative diseases, contributing to the discovery of potential biomarkers for early diagnosis and disease progression. |
| **B3. Methodology**  The proposed study aims to develop an integrated framework that combines proteomic data with neuroimaging to enhance the understanding of Alzheimer’s Disease. The conceptual framework is built upon the hypothesis that specific proteins are correlated with neuroimaging phenotypes, and these correlations can provide insights into disease mechanisms and biomarkers for early diagnosis.  Study Design:  1. Data Collection:  Proteomic Analysis: Proteomics dataset will be obtained from public database such as ProteomeXchange (https://www.proteomexchange.org/) and Alzheimer’s Disease Neuroimaging Initiative (ADNI).  Neuroimaging: Neuroimaging dataset of AD patients will be obtained from public database such as the Alzheimer’s Disease Neuroimaging Initiative (ADNI).  The datasets retrieved from ProteomeXchange and ADNI will be used for model pre-training. For accurate and meaningful correlation between protein expression and neuroimaging biomarkers, testing will be performed using data from the same patient cohort, where both proteomics and neuroimaging profiles are available. Such datasets can be accessed through ADNI.  2. Data Pre-processing:  Normalize protein expression values, missing value imputation and neuroimage pre-processing. This will involve pre-processing both datasets to ensure compatibility and performing statistical analyses to identify significant correlations.  3. Feature extraction:  Extract key features such as Differential Expressed Proteins (DEPs) and neuroimaging phenotypes for both modalities.  4. Data Integration and Analysis:  The study will employ machine learning algorithms to integrate proteomic and neuroimaging data. Predictive models will be constructed to correlate proteomic biomarkers with neuroimaging phenotypes.  5. Biological interpretation and relation:  To map the result from the proposed multimodal to the biological perspective to ensure the relativeness of the outcome with AD domain.  6. Outcome:  The primary outcomes will include the identification of novel biomarkers and the development of predictive models that enhance diagnostic accuracy for neurodegenerative diseases.  Please refer to the enclosed appendix for the methodology flowchart, Gantt chart and milestones. |

Appendix

**Flow chart**



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Gantt Chart** | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Item** | **Activity** | **Year 1** | | | | | | | **Year 2** | | | | | | | | | | |
| **6** | **7** | **8** | **9** | **10** | **11** | **12** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| **1** | Literature study |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | Data Collection |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Data pre-processing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Datasets integration and prelimininary analysis |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | Model development and validation |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 | Results interpretation and case studies |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 | Report writing and presentation |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Milestones** | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Item** | **Activity** | **Year 1** | | | | | | | **Year 2** | | | | | | | | | | |
| **6** | **7** | **8** | **9** | **10** | **11** | **12** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| **1** | Completion literature study |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | Completion of data Collection |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Completion of data pre-processing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Completion of datasets integration and prelimininary analysis |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | Completion of model development and validation |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 | Completion of results interpretation and case studies |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 | Completion of final report |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |