

Université de Lille  
Ecole graduée de Biologie Santé de Lille

**THESE DE DOCTORAT**

*En Informatique et Intelligence Artificielle*

**Conception d'espaces latents interprétables pour la  
médecine de précision : applications en neuroimagerie  
multimodale**

Par Dorian Manouvriez

Soutenance prévue le 9 décembre 2025

---

**Composition du jury**

---

**Augustin Lecler** *rapporleur*

Professeur des Universités - Praticien Hospitalier, Université Paris Cité

**Benjamin Lemasson** *rapporleur*

Chargé de Recherche Inserm, Institut de Neurosciences de Grenoble, U1216

**Carole Frindel** *examinatrice*

Maîtresse de conférences, INSA Lyon

**Gaël Chételat** *examinatrice*

Directrice de recherche, INSERM UA20 Neuropresage, Université de Caen

**Jean-Pierre Pruvo** *président*

Professeur des Universités - Praticien Hospitalier, Université de Lille

**Renaud Lopes** *directeur de thèse*

Maitre de Conférences des Universités - Praticien Hospitalier, Université de Lille

**Grégory Kuchcinski** *co-directeur de thèse*

Professeur des Universités - Praticien Hospitalier, Université de Lille

## Communications et publications

### Article publié

Manouvriez, D., Kuchcinski, G., Roca, V., Sillaire, A., Bertoux, M., Delbeuck, X., Pruvost, J.P., Lecerf, S., Pasquier, F., Lebouvier, T., & others (2025). BrainAGE latent representation clustering is associated with longitudinal disease progression in early-onset Alzheimer's disease. *Journal of Neuroradiology*, 52(5), 101365.

### Article en cours de révision

PatientSpace: An interpretable graph-based latent space for multimodal neuroimaging biomarker learning in Alzheimer's Disease and Frontotemporal Dementia

### Posters et communications orales en conférences

#### Posters

- PatientSpace: An interpretable graph -based latent space for multimodal neuroimaging biomarker learning in Alzheimer's Disease and Frontotemporal Dementia (**IABM2024, Grenoble, France - IABM2025, Nice, France – ESMRMB 2025, Marseille, France**)

#### Communications orales

Lopes, R., Manouvriez, D., Kuchcinski, G., Lecerf, S., Rogeau, A., Lahousse, H., Bertoux, M., Lebouvier, T., & Semah, F. (2025). PatientSpace: apprentissage de marqueurs d'imagerie par IA générative pour les démences de type Alzheimer et apparentées. Médecine Nucléaire, 49(2), 98.

- PatientSpace: An interpretable graph -based latent space for multimodal neuroimaging biomarker learning in Alzheimer's Disease and Frontotemporal Dementia (**ECCN 2025, Madrid, Espagne – JFMN 2025, Montrouge, France**)

### Formation

Au cours de ma thèse, j'ai eu l'opportunité d'effectuer un stage de quatre mois (mars à juin 2024) au sein de l'équipe MANIFOLD de l'University College London, sous la supervision du Pr. James H. Cole et de la Dr. Mariam Zabihi. J'ai également contribué à l'enseignement des statistiques auprès d'étudiants de deuxième année en orthophonie. Par ailleurs, j'ai eu la possibilité d'encadrer un étudiant de master pendant un stage de trois semaines dans le cadre de ses études de médecine, afin de lui faire découvrir le monde de la recherche.

## **Remerciements**

A M. Benjamin Lemasson et M. Augustin Lecler, je vous remercie d'avoir accepté d'être rapporteurs de cette thèse. J'espère que vous y trouverez de l'intérêt et que sa lecture saura retenir votre attention.

A Mme Carole Frindel, je vous remercie d'avoir accepté de participer à cette soutenance en tant qu'examinatrice de mon travail de thèse.

A Mme Gaël Chételat, je vous remercie pour votre implication lors de mes comités de suivi de thèse. Les échanges que nous avons eus ont été particulièrement constructifs et ont contribué de manière significative à l'avancée de mes travaux. Je vous suis également reconnaissant d'avoir accepté le rôle d'examinatrice pour cette soutenance.

A M. Jean-Pierre Pruvost, je vous remercie pour votre soutien tout au long de ces trois années de thèse. Je vous suis également reconnaissant d'avoir permis l'obtention du contrat CIFRE qui a rendu ce travail possible. Enfin, je vous remercie de présider le jury de cette soutenance.

A Renaud Lopes, je te remercie pour la confiance que tu m'as accordée en me confiant ce projet de thèse. Tes remarques, tes conseils et ton expérience dans la recherche ont été essentiels non seulement à la réussite de ce travail, mais aussi à mon développement en tant que jeune chercheur.

A Grégory Kuchcinski, je te remercie pour ton investissement tout au long de cette thèse, ainsi que pour ton expertise scientifique et radiologique. Ton accompagnement m'a permis d'approfondir ma compréhension de mon sujet de thèse et des interactions entre recherche et pratique clinique.

Je remercie aussi tous les membres de l'équipe du Lille Invivo Imaging and Functional Exploration (LIIFE) avec lesquels j'ai eu grand plaisir à évoluer tout au long de ces trois années : Vincent, Barnabé, Cécile, Romain, Julien, Amal, Sabine, Maxime, Julien, Jean-Baptiste.

J'adresse également mes remerciements à toutes les personnes qui ont contribué, de près ou de loin, à ce projet et ont permis son avancement : Simon Lecerf, Thibaud Lebouvier, Maxime Bertoux, Bastien Leguellec, Martin Bretzner, Lucas Shorten.

Enfin, je souhaite exprimer ma gratitude envers les Plateformes Lilloises en Biologie Santé (PLBS), ainsi qu'envers la société Siemens Healthcare SAS pour le financement de cette thèse.

## Table des matières

<b>RESUME .....</b>	<b>8</b>
<b>ABSTRACT .....</b>	<b>10</b>
<b>INTRODUCTION .....</b>	<b>12</b>
1.    LES MODELES NORMATIFS .....	12
1.1.    Définition .....	12
2.    LES MODELES NORMATIFS EN NEUROIMAGERIE .....	13
2.1.    Intérêt .....	13
2.2.    Les modèles normatifs en neuroimagerie : aspects pratiques .....	14
3.    VIEILLISSEMENT CEREBRAL ET MODELE NORMATIF : L'APPROCHE BRAINAGE .....	15
3.1.    Définition .....	15
3.2.    Intérêt du BrainAGE .....	15
4.    A LA RECHERCHE DE PATIENTS SIMILAIRES .....	16
4.1.    Extraire des similarités.....	16
4.2.    Sous typage de maladie .....	18
4.3.    Aide au diagnostic et pronostic.....	18
4.4.    Combler les données manquantes avec des patients de références.....	18
4.5.    Limitations .....	18
5.    OBJECTIF DE LA THESE .....	19
<b>CHAPITRE I L'ESPACE LATENT, UNE REPRESENTATION REDUITE DES DONNEES MEDICALES .....</b>	<b>21</b>
I.1. <b>ESPACE LATENT .....</b>	<b>21</b>
I.1.1. <i>Espaces latents : définition et cadre conceptuel</i> .....	21
I.1.2. <i>Réduction de dimension et variables latentes</i> .....	21
I.1.3. <i>Apprentissage profond et exploitation des espaces latents</i> .....	21
I.2. <b>METHODE D'APPRENTISSAGE ET IMPACT SUR L'ESPACE LATENT.....</b>	<b>22</b>
I.2.1. <i>Méthodes d'apprentissage</i> .....	22
I.2.2. <i>Apprentissage multimodal</i> .....	26
I.2.3. <i>Conclusion</i> .....	27
I.3. <b>STRUCTURATION DE L'ESPACE LATENT.....</b>	<b>27</b>
I.3.1. <i>Autoencodeur variationnel</i> .....	28
I.3.2. <i>Extensions du VAE et disentanglement learning</i> .....	29
I.3.3. <i>Favoriser l'interprétation : intégrer des critères de similarités</i> .....	32
I.4. <b>LES ESPACES LATENTS EN NEUROIMAGERIE .....</b>	<b>33</b>
I.4.1. <i>Harmonisation de séquence IRM</i> .....	33
I.4.2. <i>Désenchevêtrement de pattern de maladie</i> .....	33
I.4.3. <i>Etude de progression de maladie</i> .....	34
I.4.4. <i>Association compte rendu image</i> .....	34
I.5. <b>RAPPEL DES OBJECTIFS DE THESE .....</b>	<b>37</b>
<b>CHAPITRE II ARTICLE « BRAINAGE LATENT REPRESENTATION CLUSTERING IS ASSOCIATED WITH LONGITUDINAL DISEASE PROGRESSION IN EARLY ONSET ALZHEIMER'S DISEASE ....</b>	<b>38</b>
II.1. <b>INTRODUCTION .....</b>	<b>39</b>
II.2. <b>MATERIELS ET METHODES.....</b>	<b>40</b>
II.2.1. <i>Population EOAD</i> .....	40

<i>II.2.2. Acquisition et prétraitement des IRM</i> .....	40
<i>II.2.3. Modèle BrainAGE</i> .....	40
<i>II.2.4. Algorithme de clustering</i> .....	43
<i>II.2.5. Évaluation neuropsychologique</i> .....	43
<i>II.2.6. Analyses volumétriques et morphométriques</i> .....	43
<i>II.2.7. Analyses statistiques</i> .....	44
<b>II.3. RESULTATS</b> .....	<b>45</b>
<i>II.3.1. Participants</i> .....	45
<i>II.3.2. Clustering</i> .....	46
<i>II.3.3. Différences démographiques, biologiques et cliniques au départ</i> .....	47
<i>II.3.4. Différences neuropsychologiques et de volumes cérébraux à l'inclusion</i> .....	48
<i>II.3.5. Différences entre clusters dans l'analyse longitudinale</i> .....	50
<b>II.4. DISCUSSION</b> .....	<b>50</b>
<b>II.5. CONCLUSION</b> .....	<b>53</b>
<b>II.6 ANNEXES</b> .....	<b>54</b>

## **CHAPITRE III PATIENTSPACE : UN ESPACE LATENT INTERPRETABLE POUR L'APPRENTISSAGE DE BIOMARQUEURS DE NEUROIMAGERIE DANS LA DEMENCE FRONTO-TEMPORALE ..... 57**

<b>III.1. INTRODUCTION</b> .....	<b>58</b>
<b>III.2. MATERIELS ET METHODES</b> .....	<b>59</b>
<i>III.2.1. Base de données</i> .....	59
<i>III.2.2. Prétraitement des images</i> .....	59
<i>III.2.3. Dataset</i> .....	59
<i>III.2.4. Modèle de deep learning</i> .....	59
<i>III.2.5. Représentation par graphe de l'espace latent – PatientSpace</i> .....	61
<i>III.2.6. Expérimentations</i> .....	62
<b>III.3. RESULTATS</b> .....	<b>63</b>
<i>III.3.1. Démographie</i> .....	63
<i>III.3.2. Interprétation du PatientSpace</i> .....	64
<b>III.4. DISCUSSION</b> .....	<b>68</b>
<i>III.4.1. Méthodologie</i> .....	68
<i>III.4.2. Performance classification diagnostique</i> .....	68
<i>III.4.3. Sous-types et interprétation des clusters</i> .....	68
<i>III.4.3. Interprétation au niveau individuel</i> .....	69
<i>III.4.4. Limites et perspectives</i> .....	69
<b>III.5. CONCLUSION</b> .....	<b>69</b>
<b>III.6. ANNEXES</b> .....	<b>70</b>
<i>III.6. S1. Dataset externe</i> .....	70
<i>III.6. S2. Détails du modèle</i> .....	70
<i>III.6. S3. Voxel based Morphometry</i> .....	71

## **CHAPITRE IV ARTICLE “PATIENTSPACE: AN INTERPRETABLE GRAPH -BASED LATENT SPACE FOR MULTIMODAL NEUROIMAGING BIOMARKER LEARNING IN ALZHEIMER’S DISEASE AND FRONTOTEMPORAL DEMENTIA” ..... 72**

<b>IV.1. INTRODUCTION</b> .....	<b>73</b>
<b>IV.2. TRAVAUX CONNEXES</b> .....	<b>74</b>
<i>IV.2.1. Apprentissage profond basé sur la neuroimagerie pour la MA et la DFT : de la classification aux modèles génératifs multimodaux</i> .....	74
<i>IV.2.2. Vers des espaces latents interprétables en neuroimagerie</i> .....	75

<b>IV.3. MATERIELS ET METHODES.....</b>	<b>75</b>
IV.3.1. Participants et acquisitions des données .....	75
IV.3.2. Prétraitement des images.....	76
IV.3.3. Ensemble des données.....	76
IV.3.4. Construction et extraction d'un espace latent structuré .....	77
IV.3.5. Expérimentations.....	80
IV.3.6. Etudes d'ablation.....	82
<b>IV.4. RESULTATS.....</b>	<b>82</b>
IV.4.1. Données démographiques .....	82
IV.4.2. Interprétation du PatientSpace.....	84
IV.4.3. Comparaison avec les méthodes de référence.....	89
IV.4.4. Performances de classification du diagnostic sur un ensemble de données externe.....	90
IV.4.5. Analyse de survie des MCI (MEMENTO Dataset) .....	91
<b>IV.5. DISCUSSION .....</b>	<b>94</b>
IV.5.1. Avancées méthodologiques .....	94
IV.5.2. Performances de la classification diagnostique.....	95
IV.5.3. Sous-types de maladies et interprétation des clusters .....	96
IV.5.4. Interprétabilité au niveau du patient.....	97
IV.5.5. Trajectoires longitudinales des sujets MCI.....	97
IV.5.6. Limites et perspectives.....	98
<b>IV.6. CONCLUSION .....</b>	<b>98</b>
<b>IV.7. APPENDIX.....</b>	<b>99</b>
IV.7. A1 : Résultats complémentaires sur les variantes du PatientSpace.....	99
<b>IV.8. ANNEXES.....</b>	<b>102</b>
IV.8. S.1. Acquisition des données .....	102
IV.8. S.2. Ensemble des données .....	103
IV.8. S.3. Détails supplémentaires du modèle .....	104
IV.8. S.4. Organisation du PatientSpace.....	105
IV.8. S.5. Voxel based Morphometry.....	106

## **CHAPITRE V PREDICTION DU PRONOSTIC FONCTIONNEL APRES AVC ISCHEMIQUE DANS LES HAUTS-DE-FRANCE GRACE A UNE APPROCHE VISION-LANGAGE..... 107**

<b>V.1. INTRODUCTION.....</b>	<b>108</b>
<b>V.2. MATERIELS ET METHODES.....</b>	<b>109</b>
V.2.1. Ethique .....	109
V.2.2. Participants .....	109
V.2.3. Prétraitement des données .....	109
V.2.4. Dataset .....	110
V.2.5. Modèles multimodaux et baseline .....	110
V.2.5.1. Modèle Baseline.....	111
V.2.5.2. Modèles multimodaux et interprétabilité .....	111
V.2.6. Entraînement et évaluation .....	113
V.2.7. Tests statistiques .....	113
<b>V.3. RESULTATS.....</b>	<b>113</b>
V.3.1. Démographie .....	113
V.3.2. Performance classification.....	114
V.3.2. Interaction multimodale .....	116
<b>V.4. DISCUSSION .....</b>	<b>117</b>
V.4.1. Performance du modèle.....	117

V.4.2. Interprétabilité du modèle .....	117
V.4.3. Limites.....	118
<b>V.5. CONCLUSION .....</b>	<b>118</b>
<b>VI. DISCUSSION ET CONCLUSION.....</b>	<b>119</b>
<b>VI.1. STRUCTURATION EFFICACE DES ESPACES LATENTS .....</b>	<b>119</b>
VI.1.1. Structuration par apprentissage multitâche.....	120
VI.1.2. L'âge, un facteur structurant complexe.....	120
<b>VI.2. DONNEES MULTIMODALES ET INTEGRATION LATENTE .....</b>	<b>121</b>
VI.2.1. De l'unimodal au multimodal .....	121
VI.2.2. Enrichir les modalités par contexte.....	122
<b>VI.3. SIMILARITE ENTRE PATIENTS .....</b>	<b>122</b>
<b>VI.4. CONCEVOIR DES ESPACES LATENTS INTERPRETABLES.....</b>	<b>123</b>
<b>VI.5. PERSPECTIVES AUTOUR DES TRAVAUX DE THESE .....</b>	<b>124</b>
<b>VI.6. CONCLUSION .....</b>	<b>126</b>
<b>ANNEXE LISTE DES FIGURES .....</b>	<b>127</b>
<b>ANNEXE LISTE DES TABLEAUX .....</b>	<b>130</b>
<b>ANNEXE LISTE DES ABREVIATIONS .....</b>	<b>132</b>
<b>REFERENCES .....</b>	<b>134</b>

# Résumé

---

L'essor du deep learning et l'accès croissant à des volumes massifs de données médicales — diagnostics, traitements, informations génétiques, antécédents cliniques, imagerie ou encore comptes rendus textuels — transforment profondément la recherche biomédicale. Cette abondance ouvre des perspectives inédites pour la médecine de précision : identifier des patients présentant des profils similaires, anticiper l'évolution des pathologies et adapter les prises en charge de façon individualisée. Néanmoins, exploiter efficacement ces données reste un défi majeur. Leur hétérogénéité et leur complexité exigent de concevoir des représentations capables de refléter fidèlement les similarités et dissimilarités entre individus, tout en restant exploitables dans un cadre clinique.

Dans ce contexte, les espaces latents offrent un cadre méthodologique particulièrement pertinent. En projetant des données médicales complexes dans des espaces de dimension réduite, ils permettent de condenser l'information tout en préservant sa structure essentielle. Le développement d'espaces latents structurés et interprétables, adaptés au domaine médical, constitue le cœur de cette thèse. De plus, comme les données médicales sont par nature hétérogènes, l'intégration progressive de la multimodalité en représente un prolongement naturel.

La première partie introduit les fondements conceptuels des espaces latents et les principales méthodes permettant de les construire, d'abord dans un cadre unimodal puis multimodal. Une attention particulière est portée à leur structuration, une condition essentielle pour concilier fidélité, généralisation et interprétation clinique. Une brève revue des applications existantes en neuroimagerie vient compléter cette mise en contexte.

La deuxième partie s'appuie sur le modèle BrainAGE afin de démontrer qu'un espace latent appris de manière supervisée peut encoder des dimensions cliniquement pertinentes. Cette approche a permis d'identifier des sous-groupes de patients dans la maladie d'Alzheimer à début précoce, indépendamment des caractéristiques cliniques conventionnelles, mettant en évidence le potentiel des représentations latentes à saisir la complexité des dimensions phénotypiques.

Les troisième et quatrième parties présentent le PatientSpace, un cadre méthodologique destiné à construire des espaces latents interprétables dans le contexte des maladies neurodégénératives. Appliqué initialement à l'imagerie par résonance magnétique (IRM) de patients atteints de démences fronto-temporales (DFT), le PatientSpace a permis d'identifier des sous-groupes cohérents tout en fournissant une explicabilité clinique des regroupements. L'approche a ensuite été étendue à plusieurs démences, notamment Alzheimer et DFT, et enrichie d'une dimension multimodale en intégrant à la fois des données IRM et de tomographie par émission de positons (TEP). Ces travaux ont démontré la robustesse de la méthode et sa capacité à capturer la variabilité interindividuelle tout en restant exploitable dans un cadre clinique.

Enfin, la cinquième partie explore un cadre multimodal plus hétérogène, appliqué à la prédiction du pronostic fonctionnel à trois mois après un AVC chez des patients candidats à une thrombectomie. En combinant des séquences IRM (FLAIR et DWI), des comptes rendus radiologiques et des données cliniques, les espaces latents ont été structurés de manière à mettre en évidence le rôle et les

interactions de chaque modalité dans la prédition du mRS à 3 mois. L'intégration de ces sources via des mécanismes d'attention et de gating a non seulement amélioré la performance prédictive, mais aussi permis de quantifier l'apport spécifique de chaque modalité, fournissant ainsi une interprétation directement exploitable en pratique clinique.

Ces travaux apportent des avancées dans l'exploitation et la structuration des espaces latents en contexte médical, ouvrant la voie à des représentations interprétables et pertinentes pour la pratique clinique, et favorisant le développement de nouvelles approches de diagnostic, de pronostic et de médecine personnalisée.

# Abstract

---

The rise of deep learning and the increasing availability of large-scale medical data—diagnoses, treatments, genetic information, clinical histories, imaging, and textual reports—are profoundly transforming biomedical research. This abundance opens unprecedented opportunities for precision medicine: identifying patients with similar profiles, anticipating disease progression, and tailoring individualized care. However, effectively leveraging these data remains a major challenge. Their heterogeneity and complexity require the design of representations capable of accurately reflecting similarities and differences between individuals, while remaining clinically actionable.

In this context, latent spaces provide a particularly relevant methodological framework. By projecting complex medical data into lower-dimensional spaces, they allow information to be condensed while preserving its essential structure. The development of structured and interpretable latent spaces, specifically adapted to the medical domain, constitutes the core of this doctoral research. Moreover, as medical data are inherently heterogeneous, the progressive integration of multimodal sources represents a natural extension of this approach.

The first part introduces the conceptual foundations of latent spaces and the main methods for their construction, initially in a unimodal and then in a multimodal framework. Particular attention is given to their structuring, which is essential to balance fidelity, generalization, and clinical interpretability. A brief review of existing applications in neuroimaging complements this context.

The second part builds on the BrainAGE model to demonstrate that a supervisedly learned latent space can encode clinically relevant dimensions. This approach enabled the identification of patient subgroups in early onset Alzheimer's disease independently of conventional clinical features, highlighting the potential of latent representations to capture the complexity of phenotypic dimensions.

The third and fourth parts present PatientSpace, a methodological framework for constructing interpretable latent spaces in the context of neurodegenerative diseases. Initially applied to magnetic resonance imaging (MRI) of patients with frontotemporal dementia (FTD), PatientSpace enabled the identification of coherent subgroups while providing clinical interpretability of the clusters. The approach was then extended to multiple dementias, including Alzheimer's and FTD, and enriched with a multimodal dimension by integrating both MRI and positron emission tomography (PET) data. These studies demonstrated the robustness of the method and its ability to capture inter-individual variability while remaining clinically interpretable.

Finally, the fifth part explores a more heterogeneous multimodal setting, applied to predicting three-month functional outcomes after stroke in patients eligible for thrombectomy. By combining MRI sequences (FLAIR and DWI), radiology reports, and clinical data, latent spaces were structured to highlight the role and interactions of each modality in predicting the three-month modified Rankin Scale (mRS). Integration of these sources through attention and gating mechanisms not only improved predictive performance but also allowed quantification of the specific contribution of each modality, providing interpretations directly actionable in clinical practice.

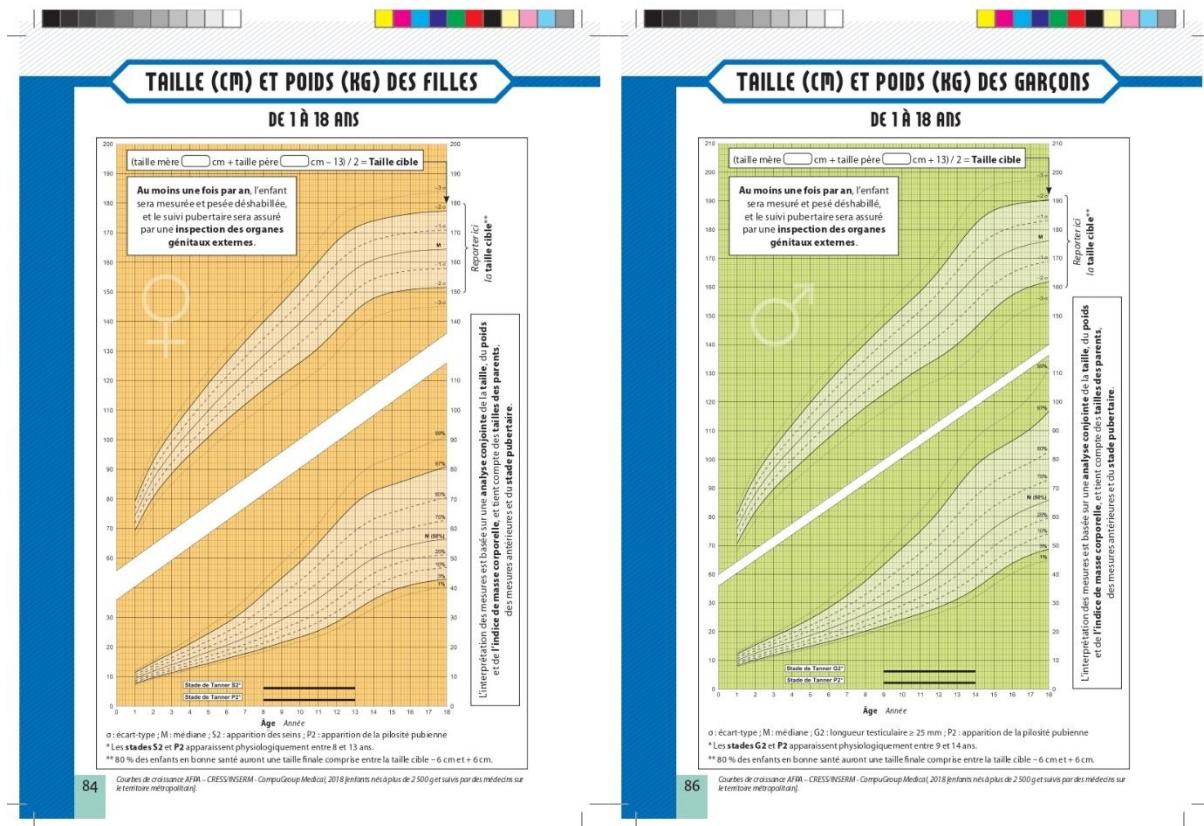
Overall, these works advance the exploitation and structuring of latent spaces in medical contexts, paving the way for interpretable and clinically relevant representations, and supporting the development of novel approaches for diagnosis, prognosis, and personalized medicine.

# Introduction

## 1. Les modèles normatifs

### 1.1. Définition

Les modèles normatifs constituent une approche puissante pour analyser et interpréter les différences individuelles au sein d'une population. Contrairement aux méthodes traditionnelles fondées sur des moyennes de groupe, ces modèles visent à capturer la variabilité naturelle des caractéristiques biologiques, comportementales ou cliniques dans une population de référence. Cette « norme » ainsi définie sert ensuite de point de comparaison pour évaluer dans quelle mesure un individu s'en écarte, en permettant d'identifier des profils atypiques ou pathologiques (Figure 1). (Rutherford et al., 2022).



1. Figure 1.

Les courbes de taille et de poids des enfants constituent des références normatives. L'écart observé par rapport à la population de référence permet de quantifier des critères tels que la maigreur, le surpoids ou l'obésité (image provenant de <https://cress-umr1153.fr/fr/courbes-de-croissance-de-reference-du-carnet-de-sante/> 05/08/2025)

L'intérêt principal de cette approche est qu'elle permet de détecter des variations propres à un individu, ouvrant ainsi la voie à une médecine plus personnalisée et précise (Rutherford et al.,

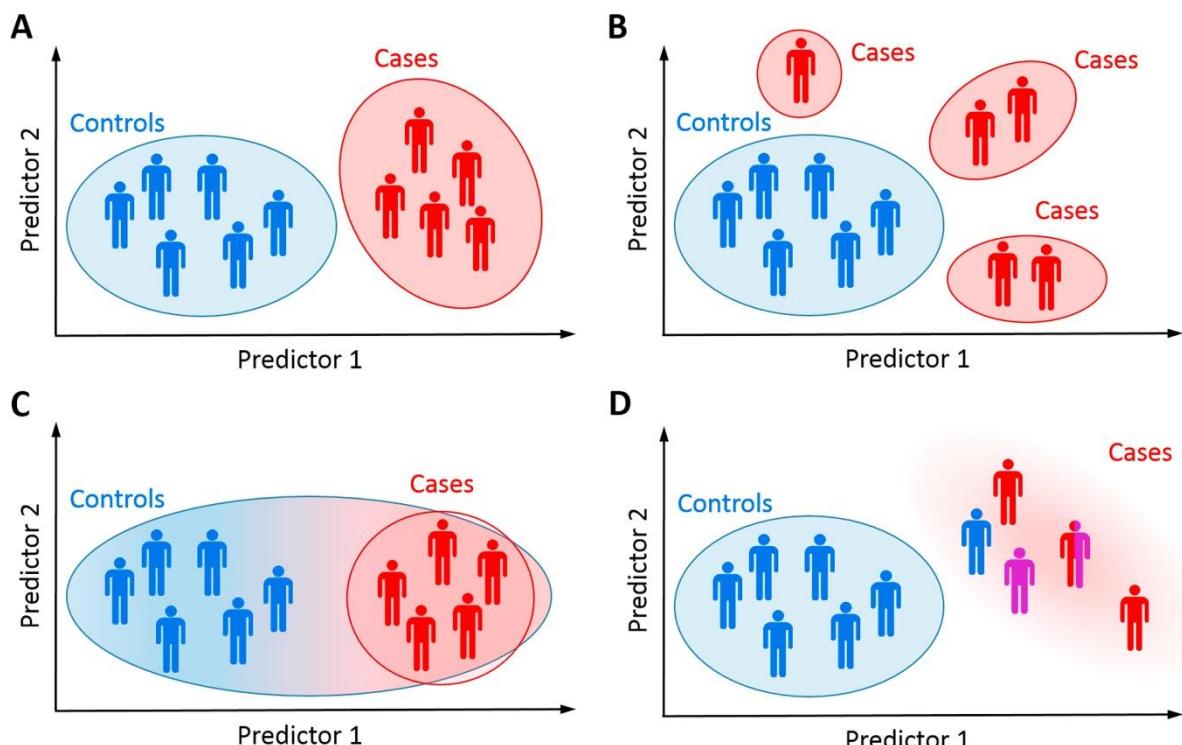
2022). Dans un premier temps, un modèle de référence est élaboré à partir de données issues d'une population saine, ce qui permet de définir des profils attendus. Dans un second temps, les écarts individuels par rapport à cette norme sont quantifiés. Cette approche méthodologique a démontré son efficacité dans le domaine médical, où elle s'avère particulièrement pertinente pour l'étude de la complexité et de l'hétérogénéité des pathologies (Kang et al., 2024; Pinaya et al., 2021; Tabbal et al., 2025).

## 2. Les modèles normatifs en neuroimagerie

### 2.1. Intérêt

Bien que les modèles normatifs soient couramment utilisés dans certains domaines médicaux, tels que la pédiatrie avec les courbes de taille et de poids, ils commencent seulement à être adoptés en neuroimagerie, notamment pour l'étude du développement cérébral, du vieillissement et des démences (Bozek et al., 2023).

L'un des principaux atouts des modèles normatifs réside dans leur capacité à intégrer l'hétérogénéité naturelle de la population au sein des cohortes cliniques. Ils permettent ainsi d'obtenir un diagnostic individualisé, centré sur le patient plutôt que sur un groupe. L'analyse ne repose plus sur la moyenne d'un groupe de patients, mais exclusivement sur l'échelle individuelle, ce qui permet de s'affranchir de l'hétérogénéité de la population cible. Cet aspect est particulièrement pertinent dans le domaine médical, où l'hétérogénéité constitue un facteur déterminant (Marquand et al., 2016).



2. Figure 2

Intérêt des modèles normatifs (image et légende issues de (Marquand et al., 2016)). L'approche classique cas-témoin suppose que les cas et les témoins forment chacun un groupe bien défini (A). Cette hypothèse peut être raisonnable dans certains contextes, mais de nombreuses autres configurations sont possibles en pratique. La population clinique peut être constituée de plusieurs sous-groupes présentant chacun une pathologie distincte

(B) ; la variabilité liée à la maladie peut s'inscrire au sein même de la variabilité observée chez les sujets sains (C) ; ou encore, le groupe clinique peut apparaître diffus et hétérogène en raison d'erreurs diagnostiques, de comorbidités ou de l'agrégation de pathologies différentes (D).

Dans le scénario idéal, l'approche cas-témoins permet de distinguer clairement deux groupes (Figure 2A). Toutefois, la réalité est souvent bien différente, en raison des variations propres à la population générale mais aussi de l'hétérogénéité au sein du groupe clinique étudié. Plus largement, les approches normatives offrent la possibilité de caractériser finement les différences interindividuelles, ce qui favorise une compréhension plus nuancée de l'expression des pathologies et permet un suivi plus personnalisé des patients.

Elles s'inscrivent pleinement dans le cadre de la médecine de précision, dont l'objectif est d'adapter les traitements au profil spécifique de chaque individu (Marquand et al., 2019).

Un autre avantage réside dans leur accessibilité en termes d'interprétabilité : il suffit en effet de définir des seuils pour quantifier les variations et, par conséquent, interpréter les résultats.

## 2.2. Les modèles normatifs en neuroimagerie : aspects pratiques

Deux grandes approches de modèles normatifs peuvent être distinguées.

La première repose sur des caractéristiques dérivées directement des données d'intérêt, comme les volumes cérébraux ou les épaisseurs corticales. La seconde utilise l'imagerie complète afin de générer un espace latent normatif.

### 2.2.1. Modèle normatif par région d'intérêt

Dans cette approche, le cerveau est représenté à travers des mesures régionales (volumes, épaisseurs corticales, etc.). L'objectif est de prédire ces mesures à partir de covariables telles que l'âge, le sexe ou encore des attributs cognitifs et psychiatriques. La prédiction repose sur des modèles capables d'estimer une distribution, tels que les processus gaussiens (Gaussian Process Regression, GPR), la régression linéaire bayésienne, la régression hiérarchique bayésienne ou encore les modèles additifs généralisés (GAMLSS) (Bozek et al., 2023; Dinga et al., 2021; Fraza et al., 2021; Marquand et al., 2016).

### 2.2.2. Modèle normatif sur l'imagerie

De manière analogue, mais sans passer par des caractéristiques régionales, il est possible de modéliser directement les distributions issues de l'imagerie complète (unimodale ou multimodale). Ces approches s'appuient notamment sur les autoencodeurs variationnels (VAE), ainsi que sur des architectures plus complexes telles que les produits, les mélanges ou les produits de mélanges d'experts (Pinaya et al., 2018). Dans ce cadre, les scores de déviation sont généralement calculés sous la forme :

$$d = \frac{x - \mu}{\sigma}$$

où  $x$  est la valeur observée,  $\mu$  la moyenne prédictive et  $\sigma$  l'écart-type estimé.

### 2.2.3. Exemple d'application

Dans le cadre de la maladie d'Alzheimer, des modèles normatifs entraînés sur plusieurs milliers de sujets sains issus de la UK Biobank ont permis de détecter de manière sensible les altérations cérébrales chez des patients présentant un trouble cognitif léger (MCI), stade précoce de la maladie, ainsi que chez des patients atteints d'Alzheimer à un stade plus avancé. Ces modèles ont montré que

les individus malades s'écartent significativement de la norme, l'ampleur de cette déviation reflétant la sévérité des altérations cérébrales (Verdi et al., 2024).

### **3. Vieillissement cérébral et modèle normatif : l'approche BrainAGE**

#### **3.1. Définition**

Le modèle Brain Age Gap Estimation (BrainAGE) est conçu pour prédire l'âge chronologique d'un individu appartenant à une population de référence à partir de variables spécifiques, telles que des IRM cérébrales de sujets sains ou encore des volumes cérébraux régionaux. L'écart entre l'âge cérébral estimé et l'âge réel, appelé Predicted Age Difference (PAD), fournit un indicateur quantitatif de la déviation par rapport à un vieillissement cérébral typique.

Le modèle BrainAGE part du principe que, chez les individus en bonne santé, l'âge cérébral estimé à partir des images correspond à l'âge chronologique réel. Bien que développé pour le cerveau, ce concept peut être étendu à d'autres organes (Cole et al., 2019, 2017; Franke and Gaser, 2019; Seitz-Holland et al., 2024).

#### **3.2. Intérêt du BrainAGE**

L'intérêt principal du modèle BrainAGE ne réside pas uniquement dans sa capacité à estimer l'âge cérébral chez des sujets sains, mais surtout dans son application à des populations cliniques, notamment les patients atteints de démence ou d'autres pathologies cérébrales telles que les maladies neurologiques et psychiatriques (Dias et al., 2025; Franke and Gaser, 2019). Plusieurs études ont en effet montré que, chez ces patients, l'âge cérébral prédit par le modèle est en moyenne plus élevé que l'âge chronologique. Cette surestimation traduit une accélération du vieillissement cérébral liée à la pathologie (Smith et al., 2019).

En pratique, le BrainAGE apprend à exploiter des biomarqueurs associés à l'âge, tels que la diminution progressive de la matière grise ou l'augmentation du volume de liquide cérebrospinal. Lorsqu'il est appliqué à des patients présentant une atrophie cérébrale anormale, comme c'est le cas dans les démences, ces caractéristiques pathologiques sont interprétées comme un vieillissement prématué, conduisant à une surestimation de l'âge cérébral. Le BrainAGE constitue alors un indicateur quantitatif de la déviation par rapport à un vieillissement cérébral typique. Ce score a été utilisé comme biomarqueur potentiel pour évaluer la sévérité ou la progression de différents types de démence, offrant ainsi une mesure continue et objectivable à partir de l'imagerie cérébrale (Abeyasinghe et al., 2025).

Au-delà de la simple mesure d'une déviation individuelle, l'approche BrainAGE permet de comparer les profils de différents patients. Des individus présentant un même BrainAGE ou des schémas d'atrophie similaires peuvent être regroupés, révélant ainsi des trajectoires communes de vieillissement cérébral. Cette capacité à mettre en évidence les similarités et différences entre patients contribue directement à une médecine plus personnalisée, en facilitant l'identification de sous-groupes cliniquement pertinents.

Dans cette optique, la recherche de patients similaires — c'est-à-dire partageant des caractéristiques cliniques, biologiques ou structurelles — s'inscrit comme un prolongement naturel de la modélisation normative.

## 4. A la recherche de patients similaires

Avec l'émergence du deep learning et du big data, un volume massif de données spécifiques aux patients — incluant diagnostics, traitements, informations génétiques ou encore antécédents médicaux — est devenu accessible. Cette explosion de données ouvre la voie à une amélioration significative de la prise en charge médicale, via des outils d'aide à la décision, des modèles prédictifs de progression de la maladie, ou encore des analyses comparatives de l'efficacité des traitements. Dans ce contexte, l'identification de similarités ou de dissimilarités entre patients devient une ressource précieuse pour optimiser les parcours de soins (Parimbelli et al., 2018).

### 4.1. Extraire des similarités

Cette section vise à présenter de manière non exhaustive les méthodes et algorithmes permettant d'identifier des similarités entre patients.

#### 4.1.1. Distance, dissimilarité et similarité

Mathématiquement, une distance est une application  $d$  vérifiant les propriétés suivantes

$$\begin{aligned} \forall(x,y) \in E^2, d(x,y) = 0 &\Leftrightarrow x = y & (P1) \\ \forall(x,y) \in E^2, d(x,y) &= d(y,x) & (P2) \\ \forall(x,y,z) \in E^3, d(x,y) &\leq d(x,z) + d(z,y) & (P3) \\ \forall(x,y) \in E^2, d(x,y) &\geq 0 & (P4) \end{aligned}$$

Cette distance permet de mesurer la dissimilarité entre deux entités  $x$  et  $y$  :

Si  $d(x,y) = 0$ , alors  $x$  et  $y$  sont identiques

Plus  $d(x,y)$  est grande, plus  $x$  et  $y$  sont différents

Une mesure de dissimilarité peut être vue comme une distance relaxée, où l'inégalité triangulaire (P3) n'est pas nécessairement respectée. L'idée reste similaire : deux objets sont identiques si leur dissimilarité est nulle, et plus la dissimilarité est élevée, plus les objets sont différents.

Par extension, une mesure de similarité évalue la ressemblance entre deux objets. Pour une similarité  $s$ , on a typiquement les propriétés suivantes

$$\begin{aligned} \forall(x,y) \in E^2, s(x,y) &\leq s(x,x) \\ \forall(x,y) \in E^2, s(x,y) &= s(y,x) \end{aligned}$$

La positivité de la similarité

$$\forall(x,y) \in E^2, s(x,y) \geq 0$$

est souvent conservée pour faciliter l'interprétation mais elle n'est pas obligatoire.

Par exemple, la similarité cosinus

$$s(x,y) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

peut prendre des valeurs entre  $[-1; 1]$ .

Il est possible de passer d'une mesure de similarité à une mesure de dissimilarité, et inversement, ce qui permet de relier les deux notions de manière simple et cohérente.

Ainsi, si  $s$  est la similarité cosinus, alors

$$d(x, y) = 1 - s(x, y)$$

définit une mesure de dissimilarité.

Inversement, si  $d$  est une mesure de dissimilarité, alors

$$s(x, y) = \frac{1}{d(x, y) + \epsilon}, \quad \text{ou} \quad s(x, y) = \exp\left(-\frac{d(x, y)}{\gamma}\right),$$

avec  $\epsilon, \gamma > 0$ , définissent une mesure de similarité.

Parmi les fonctions respectant ces propriétés, les distances classiques comme la distance euclidienne, la distance de Manhattan ou la distance Mahalanobis sont couramment utilisées pour mesurer des dissimilarités, tandis que des mesures telles que la similarité cosinus sont fréquemment employées pour évaluer des similarités.

En pratique, si l'on considère deux vecteurs représentant l'épaisseur corticale de deux patients  $x$  et  $y$ , la similarité ou dissimilarité entre ces patients peut être estimée par la distance euclidienne entre ces vecteurs.

#### 4.1.2. Algorithme de clustering

Les mesures de similarité permettent de comparer des patients de manière individuelle, mais elles ne fournissent pas directement de structure globale au sein de la population. Les algorithmes de clustering ont pour objectif de regrouper les patients en fonction de leurs profils, en minimisant certaines fonctions de coût ou en maximisant des critères de cohésion. Parmi les méthodes classiques, on peut citer :

- K-means, qui cherche à partitionner les patients en groupes en minimisant la variance intra-cluster (Lloyd, 1982);
- Clustering hiérarchique, qui construit un arbre hiérarchique de regroupement basé sur des distances successives (Murtagh and Contreras, 2012);
- Algorithmes de détection de communautés (par exemple Louvain), particulièrement adaptés aux structures en réseau (Blondel et al., 2008);
- Modèles de mélange gaussien, qui supposent que les données proviennent d'une combinaison de distributions gaussiennes et permettent d'estimer des sous-groupes probabilistes.

Ces algorithmes permettent d'extraire des sous-groupes de patients partageant des caractéristiques similaires dans l'espace des vecteurs de données, qu'il s'agisse d'indices cliniques, d'épaisseurs corticales ou d'autres mesures neuroimageries. L'utilisation d'une mesure de similarité adaptée ou d'un algorithme de clustering approprié constitue ainsi une étape clé pour l'identification de profils patients homogènes et l'étude de la variabilité interindividuelle.

Par exemple, une étude a généré des vecteurs d'atrophie basés sur des modèles normatifs et les a stratifiés à l'aide d'un modèle de mélange gaussien pour identifier des sous-groupes au sein d'une population de patients atteints de la maladie d'Alzheimer (Kang et al., 2024). Dans une autre approche, la théorie des graphes a été utilisée pour extraire des communautés de patients partageant des épaisseurs corticales cérébrales similaires (Park et al., 2017).

#### **4.2. Sous typage de maladie**

Les similarités inter-patients peuvent notamment être exploitées pour identifier des sous-types de pathologies. Dans le cas des maladies neurodégénératives, et en particulier d'Alzheimer, certaines études (Kang et al., 2024; Poulakis et al., 2022) ont mis en évidence des clusters de patients présentant des phénotypes d'atrophie distincts. Appliqués à une population MCI, ces sous-types révélaient des vitesses différentes de progression vers la démence. D'autres travaux, comme celui de Ferreira et al. (Ferreira et al., 2020), ont montré que la maladie d'Alzheimer pouvait être décrite selon deux axes : la sévérité de la maladie et sa typicité. Identifier des patients similaires permettrait alors de mieux positionner un individu au sein de ces axes, et de mieux comprendre comment il s'inscrit dans un spectre clinique déjà connu.

#### **4.3. Aide au diagnostic et pronostic**

Un autre domaine où la similarité entre patients peut jouer un rôle central est celui du diagnostic et du pronostic. Les modèles d'intelligence artificielle montrent déjà une forte efficacité pour la classification diagnostique ou la prédiction de trajectoires de maladie (Arbabshirani et al., 2017; Rathore et al., 2017). Toutefois, ils souffrent souvent du problème de "boîte noire", rendant leur interprétation difficile pour les cliniciens. L'approche par similarité, fondée sur l'identification de patients proches dans l'espace latent, apporte une solution plus interprétable : elle permet de raisonner par analogie, en comparant un patient à d'autres aux trajectoires connues. Cela peut contribuer à renforcer la confiance des praticiens et à rendre la décision médicale plus transparente.

#### **4.4. Combler les données manquantes avec des patients de références**

Enfin, dans le domaine médical, il est fréquent de faire face à des données manquantes. Les critères de similarité peuvent ici aussi jouer un rôle important : en identifiant des patients de référence — dont les données sont complètes — il devient possible de compenser les données manquantes d'un patient incomplet en s'appuyant sur les profils des plus proches voisins (Beretta and Santaniello, 2016; Jazayeri et al., 2020). Cette approche ouvre la voie à des analyses robustes, même en présence d'informations partielles.

#### **4.5. Limitations**

Malgré ses bénéfices, l'approche fondée sur la similarité entre patients comporte plusieurs limites importantes.

La qualité et l'homogénéité des données (scanners IRM, données cliniques, génétiques, historiques) restent cruciales : des variations de protocole ou des biais d'échantillonnage peuvent altérer la définition de la similarité et biaiser les regroupements.

En termes d'imputation, le recours à des profils de patients similaires pour combler des données manquantes peut introduire des erreurs ou amplifier les biais existants si la similarité n'est pas rigoureusement définie (Azur et al., 2011).

Par ailleurs, l'interprétation clinique de ces similarités impose une prudence : deux patients perçus comme similaires selon certains critères d'imagerie peuvent diverger sur d'autres dimensions (comorbidités, traitement), limitant la robustesse des conclusions.

La généralisation constitue un autre défi majeur : les clusters identifiés dans une cohorte spécifique peuvent ne pas se reproduire dans d'autres populations, particulièrement si l'échantillon est restreint ou peu diversifié (Matta et al., 2024).

## 5. Objectif de la thèse

Cette thèse vise à développer des méthodes pour apprendre des espaces latents structurés et interprétables, qu'ils soient unimodaux ou multimodaux, dans le contexte de l'analyse neuroimagerie. L'objectif principal est d'étudier comment ces espaces peuvent être organisés pour capturer les similarités et dissimilarités entre patients, en tenant compte de la complexité des données cliniques, démographiques et d'imagerie. Tout au long de ce manuscrit, la structuration de l'espace latent sera au cœur des travaux. En effet, cette structuration influence directement la manière dont on identifie des patients similaires, ce qui peut orienter le diagnostic, le pronostic ou encore le phénotypage pathologique. Cette structuration est cependant délicate : une organisation trop naïve de l'espace latent peut conduire à des regroupements peu informatifs (par exemple des patients du même âge ou avec le même diagnostic, sans que leurs profils cliniques ou d'imagerie ne soient véritablement proches), tandis qu'une structuration trop complexe peut nuire à la généralisation des modèles ou à leur interprétabilité clinique. De plus, l'intégration ou non de certaines modalités (cliniques, démographiques, ou multimodales en imagerie) a un impact direct sur la nature des similarités induites dans l'espace latent. Une réflexion approfondie est donc nécessaire sur la manière de construire ces représentations pour qu'elles soient à la fois fidèles, robustes et exploitables en pratique clinique. C'est autour de ces enjeux — structuration pertinente d'espaces latents et mesure de similarité interindividuelle — que s'articulent les contributions de cette thèse.

Le manuscrit suit alors le plan suivant :

1. **Chapitre 1 : L'espace latent, une représentation réduite de données médicales.** Ce chapitre introductif présente les fondements conceptuels des espaces latents, leur intérêt pour représenter de manière compacte et interprétable des données médicales complexes, ainsi que les enjeux liés à leur structuration.
2. **Chapitre 2. Article « BrainAGE latent representation clustering is associated with longitudinal disease progression in early-onset Alzheimer's disease ».** Nous explorons ici la capacité d'un espace latent appris en supervision à capturer des informations au-delà de la tâche ciblée (prédiction de l'âge), à travers l'exemple du modèle BrainAGE. Nous montrons que cet espace encode des dimensions cliniquement pertinentes, révélant une structuration implicite utile à la caractérisation des pathologies.
3. **Chapitre 3. PatientSpace : un espace latent structuré.** Nous développons un espace latent structuré et interprétable, construit à partir d'une architecture de VAE conditionnel prenant en entrée l'âge et les imageries cérébrales. Cet espace sert à caractériser les démences fronto-temporales (DFT). Pour en améliorer la structuration et l'interprétation, nous

introduisons un critère de similarité fondé sur une approche contrastive, permettant d'organiser les représentations latentes. Cette organisation autorise ensuite l'analyse fine des profils d'imagerie et des données cliniques des patients, en s'appuyant sur la méthode des plus proches voisins.

4. **Chapitre 4. Article “PatientSpace: An interpretable graph -based latent space for multimodal neuroimaging biomarker learning in Alzheimer’s Disease and Frontotemporal Dementia”.** Nous étendons l’approche précédente à un cadre multimodal combinant IRM T1w et FDG-TEP. L’âge du patient conditionne désormais le prior du VAE, introduisant une régularisation temporelle de l’espace latent, de manière à refléter la progression de la maladie. Le critère de similarité est maintenu mais formulé selon une approche non contrastive. Ce PatientSpace multimodal permet ainsi d’explorer la similarité multimodale entre patients et d’enrichir l’explicabilité clinique des regroupements.
5. **Chapitre 5. Prédiction du pronostic fonctionnel après AVC ischémique dans les Hauts-de-France grâce à une approche vision-langage.** Dans la continuité des chapitres précédents, nous élargissons la multimodalité à des données hétérogènes en combinant imagerie, texte et données cliniques. Les séquences IRM (FLAIR et DWI), les comptes rendus radiologiques et les données cliniques tabulaires sont intégrés dans un modèle d’attention et de fusion pour prédire le devenir fonctionnel à trois mois (mRS) chez des patients victimes d’un AVC et candidats à une thrombectomie mécanique. Cette étude constitue une première étape exploratoire vers une intégration explicite et robuste de modalités multiples dans un cadre prédictif.
6. **Discussion générale et perspectives autour des travaux de thèse**

# Chapitre I

## L'espace latent, une représentation réduite des données médicales

---

### I.1. Espace latent

#### I.1.1. Espaces latents : définition et cadre conceptuel

Dans les systèmes d'apprentissage automatique, comprendre et manipuler les représentations internes des données est un enjeu central. C'est dans ce cadre qu'émerge la notion d'espace latent : un espace vectoriel de dimension inférieure, non directement observable, dans lequel les données sont projetées afin de révéler des structures sous-jacentes plus simples ou plus pertinentes. Ce concept joue un rôle fondamental, tant dans les méthodes classiques de réduction de dimension que dans les architectures profondes modernes.

Cette section retrace l'évolution du concept d'espace latent, des techniques linéaires initiales aux représentations apprises automatiquement par des réseaux neuronaux.

#### I.1.2. Réduction de dimension et variables latentes

Dans de nombreux domaines de l'apprentissage automatique, les données brutes sont souvent de très haute dimension. A titre d'exemple, une séquence 3D T1 en écho de gradient génère une image de plusieurs millions de voxels. Travailler directement dans ces espaces pose plusieurs défis : un coût computationnel élevé, la présence de bruit, ainsi que des difficultés d'interprétation. Une approche courante pour remédier à ces problèmes consiste à projeter les données dans un espace de dimension réduite qui conserve l'information essentielle : c'est ce que l'on appelle un espace latent. L'un des premiers algorithmes de réduction de dimension est l'analyse en composantes principales (ACP) (Hotelling, 1933) qui propose projection linéaire maximisant la variance des données. Cette technique permet de représenter l'information principale dans un espace réduit facilitant ainsi l'exploration et la visualisation (Gewers et al., 2021; Jolliffe, 2002).

La notion de variable latente s'est ensuite développée dans les modèles probabilistes, tels que les mixtures de Gaussiennes (GMM). Dans ce cadre, on suppose que les données sont générées à partir d'une combinaison de distributions gaussiennes. Les variables latentes correspondent alors à des facteurs cachés : les paramètres des composantes des gaussiennes (la moyenne et la matrice de covariance) et l'appartenance probabiliste de chaque observation à chaque gaussienne (Reynolds, 2015). Un autre exemple majeur est celui des machines de Boltzmann restreintes (RBM), qui ont joué un rôle pionnier dans le développement des autoencodeurs profonds, dont nous reviendrons dessus. Ces modèles générèrent des représentations latentes en apprenant, à partir des données observées, des structures cachées gouvernées par une distribution probabiliste particulière (Hinton, 2012; Larochelle and Bengio, 2008).

#### I.1.3. Apprentissage profond et exploitation des espaces latents

Comme mentionné précédemment, travailler directement dans l'espace brut des données n'est pas toujours optimal. Les méthodes classiques de réduction de dimension sont efficaces pour des

données tabulaires, elles s'avèrent moins adaptées aux modalités complexes telles que l'image, le texte ou l'audio. Cette limitation tient notamment à la structure intrinsèquement non-linéaire de ces données, ainsi qu'à la forte dépendance locale et contextuelle entre leurs éléments.

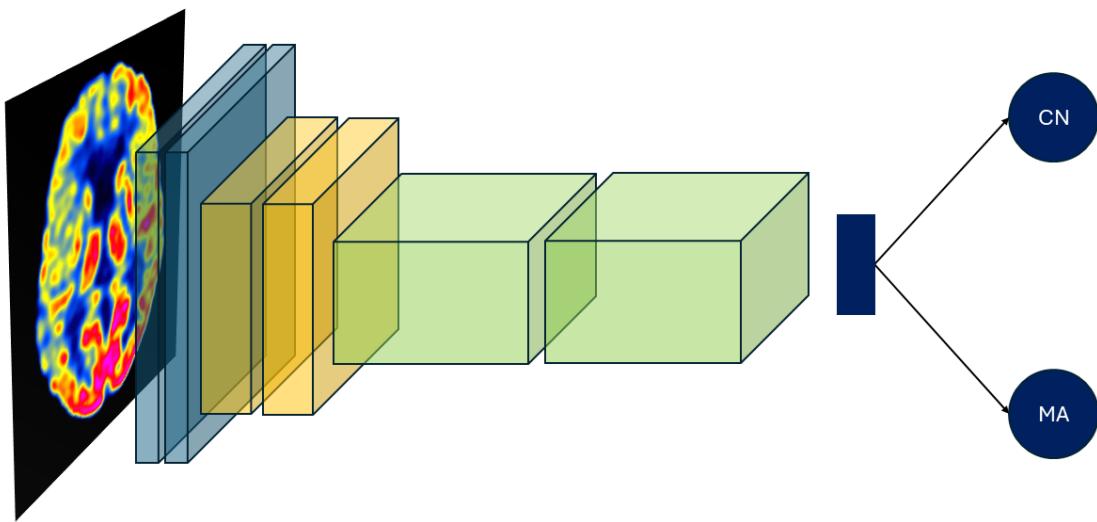
L'émergence de l'apprentissage profond a permis de surmonter ces limites. Grâce à des architectures adaptées, ces méthodes peuvent apprendre automatiquement des représentations latentes pertinentes. Par exemple, les réseaux de neurones convolutifs (CNN) sont capables d'extraire hiérarchiquement des caractéristiques visuelles pertinentes à partir d'images. Chaque couche du réseau transforme progressivement les données en représentations de plus en plus riches et compressées. Ces représentations apprises ont conduit à des avancées majeures, à la fois pour l'analyse, la compréhension et la génération de données complexes.

## I.2. Méthode d'apprentissage et impact sur l'espace latent

La construction d'un espace latent peut s'appuyer sur trois grandes familles d'apprentissage. Chacune de ces approches influence la qualité, la structure et la richesse des représentations latentes générées.

### I.2.1. Méthodes d'apprentissage

#### I.2.1.1. Apprentissage supervisé



3. Figure I-1. Apprentissage supervisé.

Une imagerie TEP est traitée par un réseau de neurones convolutif (CNN) afin de prédire, à partir de l'espace latent, si le patient est sain (CN) ou atteint de la maladie d'Alzheimer (MA).

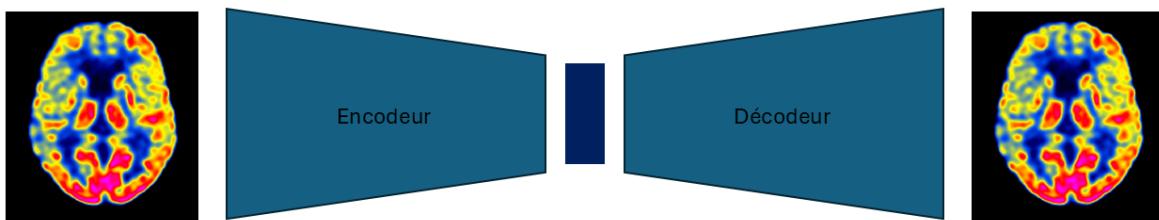
L'apprentissage supervisé (Figure I-1) consiste à apprendre une fonction de prédiction à partir d'un jeu de données annoté. Cette approche est couramment utilisée pour des tâches de classification ou de régression. Dans ce cadre, l'espace latent issu d'un apprentissage supervisé est optimisé pour discriminer les différentes classes, en extrayant des caractéristiques spécifiques à chaque catégorie plutôt que des attributs globaux.

Un exemple pertinent en neuroimagerie concerne la classification des démences à partir d'images cérébrales (Figure I-1). Par exemple, Rogeau et al (Rogeau et al., 2024) utilisent l'imagerie TEP-FDG pour distinguer les sujets sains, les patients atteints de la maladie d'Alzheimer et ceux souffrant de

démence fronto-temporale. L'espace latent appris permettait de faire émerger des patterns spécifiques à chaque catégorie par une méthode d'occlusion.

Cet apprentissage est relativement simple à mettre en œuvre, néanmoins, mais il présente plusieurs limites : il nécessite un jeu de données richement annoté, ce qui représente un coût important, et l'espace latent ainsi construit est souvent étroitement lié à la tâche supervisée, ce qui limite sa capacité de généralisation à d'autres types d'analyses.

### I.2.1.2. Apprentissage non supervisé



4. Figure I-2. Apprentissage non supervisé.

Une image TEP est encodée dans un espace latent par un encodeur, puis cet espace latent est décodé par un décodeur afin de régénérer l'image TEP d'origine.

Contrairement à l'apprentissage supervisé, l'apprentissage non-supervisé (Figure I-2) ne requiert pas d'annotations. Dans ce contexte, les architectures les plus courantes sont les autoencodeurs, qui apprennent à encoder une donnée dans un espace latent puis à la reconstruire à partir de cette représentation (Hinton and Salakhutdinov, 2006). Cette approche permet de capturer des structures inhérentes aux données sans supervision explicite, et d'extraire des caractéristiques à la fois globales et spécifiques. Les variantes probabilistes, telles que les autoencodeurs variationnels (VAE), permettent d'obtenir en outre un espace latent structuré, comme nous le verrons dans la section suivante.

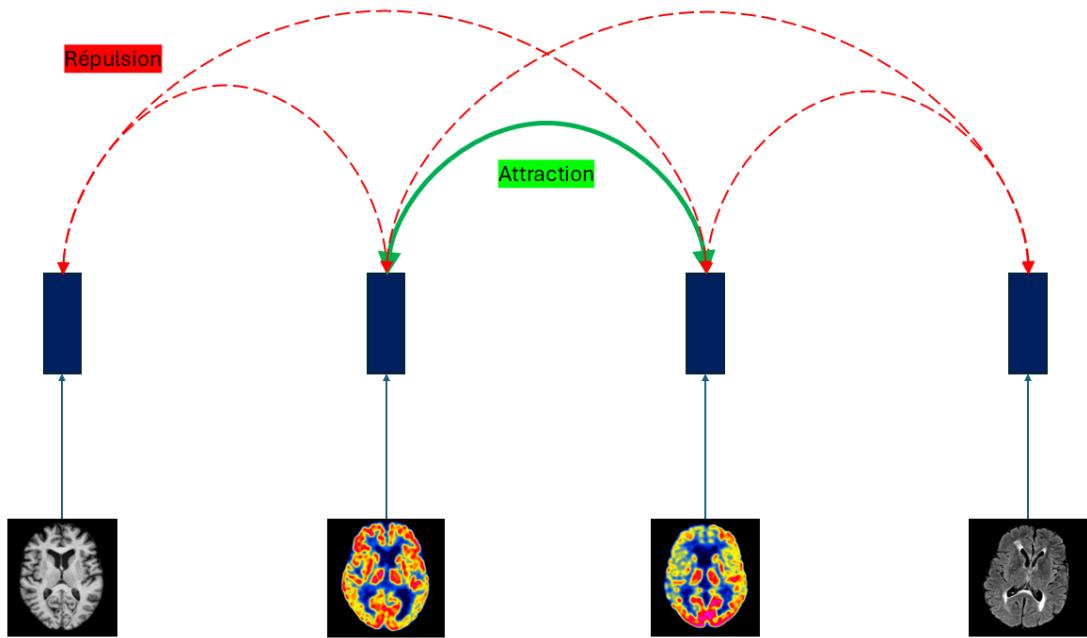
Dans le domaine de la neuroimagerie, les autoencodeurs sont utilisés pour extraire des représentations compactes d'images cérébrales, facilitant par la suite des tâches telles que la détection d'anomalie. Luo et al (Luo et al., 2023) entraînent un autoencodeur sur une base de sujets sains puis ils l'appliquent sur une base de sujet avec AVC pour segmenter les lésions. Toutefois, bien que ces méthodes ne nécessitent pas d'annotations, les espaces latents obtenus sont souvent difficiles à interpréter directement et peuvent capturer des caractéristiques non informatives à certaines tâches sous-jacentes.

### I.2.1.3. Apprentissage auto-supervisé

L'apprentissage auto-supervisé consiste à définir, à partir des données elles-mêmes, des tâches auxiliaires permettant d'entraîner un modèle sans supervision explicite. Cette stratégie exploite la structure intrinsèque des données pour guider l'apprentissage. Deux grandes approches se distinguent :

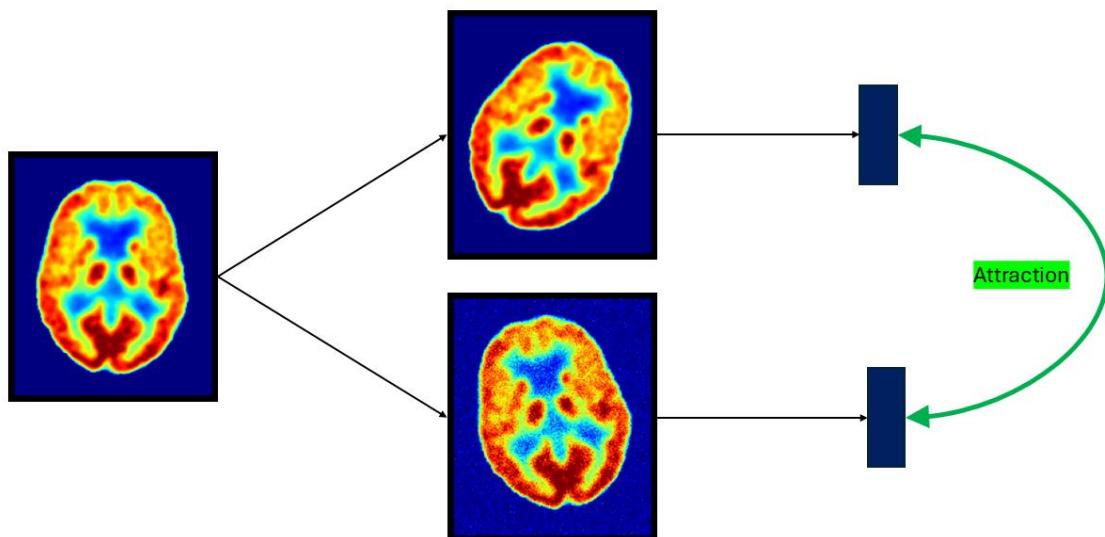
- Le contrastive learning consiste à rapprocher, dans l'espace latent, les représentations de données similaires tout en éloignant celles correspondant à des données différentes. (Figure I-3)

- Le non-contrastive learning vise à apprendre des représentations stables à partir de différentes augmentations artificielles d'une même donnée, sans recourir à des exemples négatifs explicites. (Figure I-4).



5. Figure I-3. Contrastive Learning.

Chaque modalité d'imagerie est encodée dans un espace latent, puis les représentations issues d'imageries similaires sont rapprochées tandis que celles provenant d'imageries dissimilaires sont éloignées



6. Figure I-4. Non-contrastive learning.

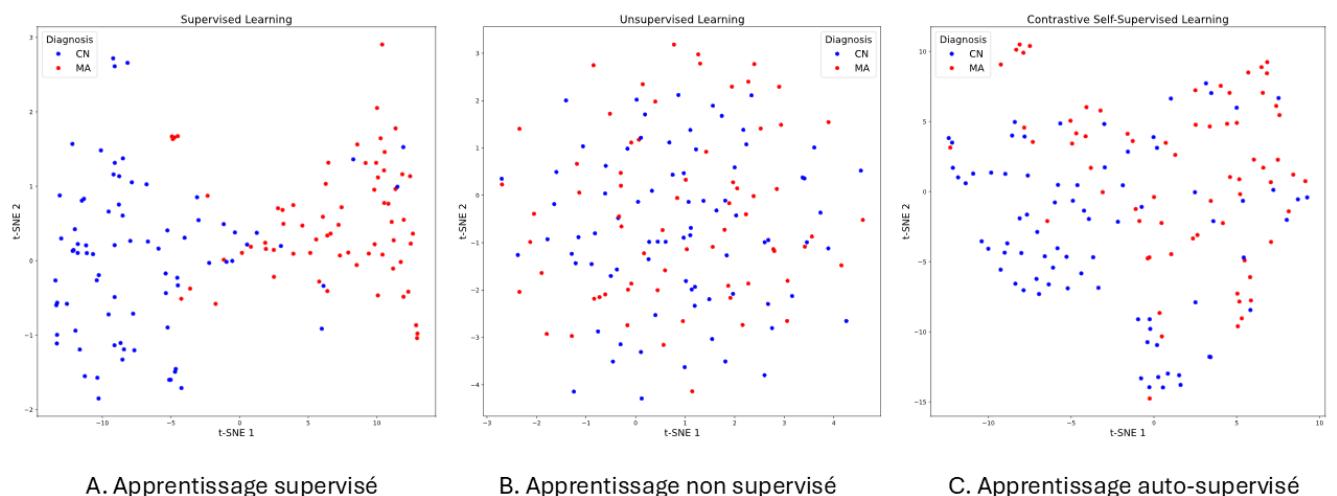
Deux images TEP transformées sont générées à partir de l'imagerie d'origine, puis leurs espaces latents sont rapprochés.

Ces méthodes sont couramment utilisées pour le pré-entraînement des modèles dits de fondation. Toutefois, elles requièrent généralement des bases de données de grande taille — en particulier dans le contrastive learning, où un nombre important d'exemples négatifs est nécessaire — et

présentent une optimisation complexe. Dans le cas du non-contrastive learning, ce problème peut se traduire par un effondrement de l'espace latent, où toutes les représentations convergent vers une représentation unique.

#### I.2.1.4. Visualisation des espaces latents et synthèse des méthodes

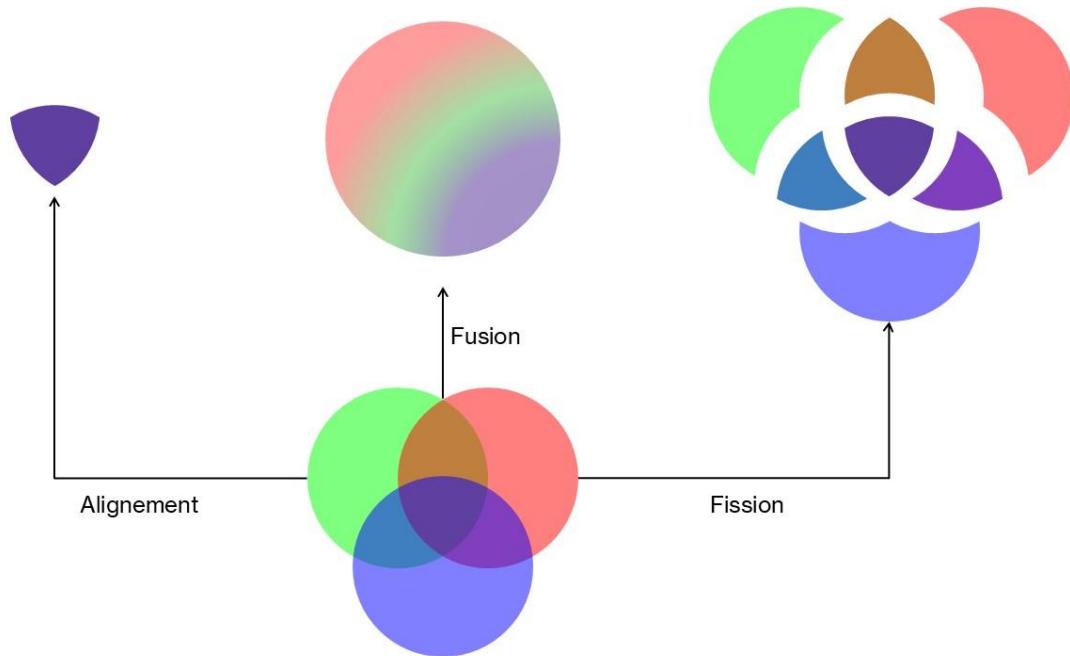
Nous avons généré des espaces latents pour chaque approche afin de comparer les résultats obtenus. Les méthodes comprenaient : une approche supervisée pour classifier le diagnostic MA ou CN d'un patient, un autoencodeur pour la méthode non supervisée, et une approche auto-supervisée combinant contrastive learning et non-contrastive learning. Dans cette dernière, deux transformations étaient générées à partir d'une même imagerie ; leurs espaces latents étaient rapprochés entre eux et éloignés de ceux des autres imageries.



7. Figure I-5. Espaces latents.  
Espace latent obtenu après apprentissage supervisé (en haut), non supervisé (milieu) et auto-supervisé (en bas). Chaque point représente un patient, et la couleur indique le diagnostic : contrôle (CN, bleu) ou maladie d'Alzheimer (MA, rouge).

Comme indiqué précédemment, l'espace latent généré par l'apprentissage supervisé est principalement discriminatif : les deux groupes sont clairement séparés (Figure I-5 A), contrairement à l'apprentissage non supervisé (Figure I-5 B), qui ne révèle aucune structure apparente et présente un mélange des données. L'apprentissage auto-supervisé produit un espace latent similaire à celui de l'apprentissage non supervisé, mais avec une distinction entre les diagnostics, comme dans l'apprentissage supervisé, sans que le modèle ait été explicitement entraîné pour différencier MA et CN (Figure I-5 C).

## I.2.2. Apprentissage multimodal



8. Figure I-6. Apprentissage multimodal.

Les trois cercles en bas représentent trois modalités différentes. Lors de l’alignement, un espace latent est extrait en ne conservant que l’information commune à chaque modalité. Lors de la fusion, l’espace latent extrait contient l’ensemble des informations disponibles, sans distinguer ce qui est spécifique ou commun aux modalités. Enfin, lors de la fission, les données sont décomposées en espaces latents spécifiques à chaque modalité et en espaces latents partagés entre modalités.

L’apprentissage multimodal vise à intégrer différentes sources de données – comme le texte, les images, l’audio ou encore les tableaux – dans un espace latent commun. Ce paradigme est particulièrement pertinent en médecine, où les données sont naturellement multimodales : imagerie cérébrale (IRM T1, FLAIR, TEP), comptes-rendus cliniques en texte libre, ou encore des variables démographiques et cliniques structurées (âge, sexe, scores cognitifs, etc.). Cette section présente les principales méthodes de génération d’un espace latent à partir de données multimodales.

Trois stratégies sont distinguées et illustrées dans la Figure I-6.

### I.2.2.1. Fusion de modalité

La première approche repose sur la fusion des modalités, où les informations provenant de différentes sources sont combinées dans un espace latent commun. Cette fusion s’effectue souvent directement dans l’espace latent via des opérations telles que la concaténation ou la somme des représentations encodées. Punjabi et al (Punjabi et al., 2019) proposent une fusion de données TEP et IRM dans un espace latent commun, dans le but de prédire le diagnostic clinique avec davantage de précision.

### I.2.2.2. Alignement de modalité

Une seconde stratégie repose sur l’alignement des modalités, visant à rapprocher les représentations latentes issues de modalités différentes lorsqu’elles décrivent une même entité. Chaque modalité est alors projetée indépendamment dans l’espace latent, mais de façon à refléter des facteurs sous-jacents communs. Cette approche s’inspire des principes de l’apprentissage auto-

supervisé. Le modèle MERLIN exploite l’alignement modalité-image/texte pour générer des comptes rendus à partir d’imagerie médicale, et organiser les données selon leur similarité, facilitant ainsi une stratification automatique des patients (Blankemeier et al., 2024).

#### I.2.2.3 Fission de modalité

Enfin, une troisième approche dite de fission des modalités consiste à structurer l'espace latent en plusieurs sous-espaces : certains spécifiques à chaque modalité, d'autres partagés entre elles. Cette organisation permet de préserver à la fois les informations propres à chaque source de données et celles utiles à leur analyse conjointe.

#### I.2.2.4. Complexité des données multimodales

Bien que l'intégration de données multimodales soit prometteuse dans le domaine médical — où la majorité des données sont effectivement de nature multimodale — elle soulève de nombreux défis méthodologiques. Le choix de l'approche (fusion, alignement ou fission) dépend étroitement de la nature des modalités et de leur degré de recouvrement. L'alignement est efficace pour extraire des représentations partagées, mais incapable de capter des informations spécifiques à chaque modalité. Si les modalités ne présentent aucun chevauchement, l'apprentissage d'un espace latent aligné devient impossible. La fusion est attrayante par sa capacité à intégrer à la fois des caractéristiques spécifiques et communes. Toutefois, cette approche est difficile à interpréter (effet boîte noire), et une modalité dominante peut occulter les autres. Enfin, la fission ressemble à la fusion, mais avec une structuration explicite entre les parties spécifiques et partagées. Cette méthode permet une meilleure interprétabilité, mais devient rapidement coûteuse en calcul (le nombre de sous-espaces latents croît exponentiellement avec le nombre de modalités :  $2^n - 1$  dans le pire des cas).

Enfin, on peut aussi évoquer le problème du multimodal collapse dans lequel une ou plusieurs modalités sont prédominantes par rapport aux restes, incitant le modèle à oublier certaines modalités (Huang et al., 2022).

### I.2.3. Conclusion

Nous avons évoqué dans cette sous-section comment extraire un espace latent. Nous avons aussi vu que les stratégies d'apprentissage ainsi que l'intégration de données multimodales ou non pouvaient influencer la nature de l'espace latent extrait. Dans le domaine médical, deux critères sont cruciaux : l'interprétabilité (comprendre la structure de l'espace) et l'informativité (retrouver les caractéristiques clés des populations étudiées, comme une atrophie hippocampique chez les patients Alzheimer). Ces exigences rendent la structuration de l'espace latent particulièrement délicate.

## I.3. Structuration de l'espace latent

Dans cette sous-partie, nous examinons les principales approches permettant de structurer un espace latent de manière à le rendre à la fois informatif et interprétable. Nous commencerons par présenter le fonctionnement des autoencodeurs variationnels (VAE), puis nous aborderons les principes du *disentanglement learning*, qui visent à isoler les facteurs de variation sous-jacents dans l'espace latent. Enfin, nous décrirons des stratégies destinées à renforcer l'interprétabilité des représentations latentes, notamment par l'introduction de critères de similarité ou de consistance.

### I.3.1. Autoencodeur variationnel

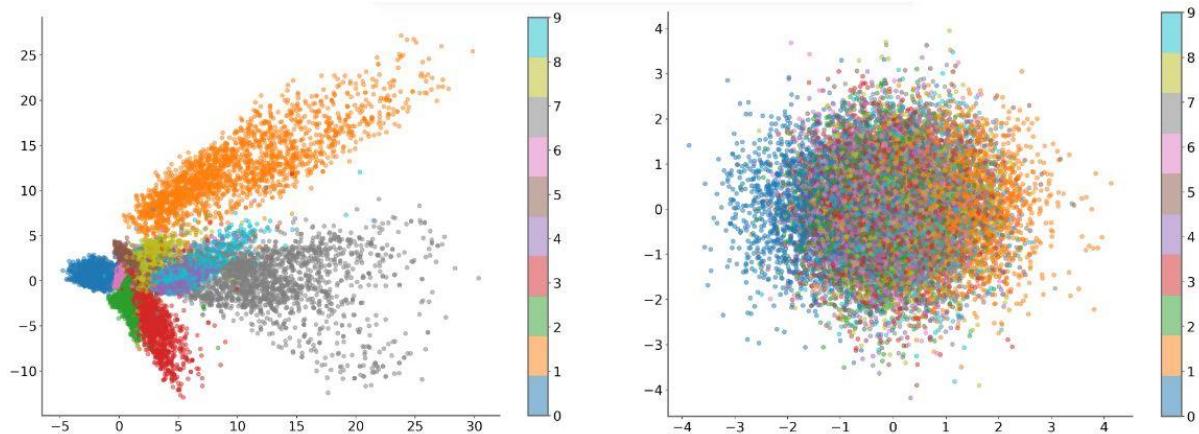
Le VAE est un modèle génératif probabiliste dans lequel chaque donnée n'est plus représentée par un point fixe dans l'espace latent, mais par une distribution aléatoire contrainte par un *prior*, en général une gaussienne multivariée centrée réduite. Cette formulation induit un espace latent continu et régularisé (Kingma and Welling, 2019, 2013).

#### I.3.1.1. Apprentissage et fonction de coût

L'apprentissage des VAE repose sur la minimisation de la fonction ELBO (Evidence Lower Bound) (Eq. I-1) qui combine deux termes : (1) la vraisemblance de reconstruction, notée  $E_{q(z|x)}[\ln p(x|z)]$ , qui mesure la capacité du modèle à reconstruire fidèlement les données ; (2) la divergence de Kullback-Leibler (KL) entre la distribution latente apprise et le prior.

$$ELBO = -E_{q(z|x)}[\ln p(x|z)] + KL(q(z|x) \parallel p(z)) \quad (\text{Eq.I-1})$$

En ajustant la pondération entre ces deux composantes, on contrôle le compromis entre qualité de reconstruction et régularisation de l'espace latent (Figure I-7).



9. Figure I-7. Impact de la pondération.

Influence de la pondération entre les deux termes de la fonction de perte. À gauche, la reconstruction est priorisée. À droite, une régularisation plus forte constraint l'espace latent à suivre une distribution gaussienne standard en deux dimensions. (Image issue de FIDLE : <https://cloud.univ-grenoble-alpes.fr/index.php/s/wxCztjYBbQ6zwD6?dir=undefined&openfile=959404079> 05/08/2025)

#### I.3.1.2. Rôle et choix du prior

Le choix du prior est un élément déterminant dans la structuration de l'espace latent. Si l'on utilise classiquement une gaussienne centrée réduite, plusieurs travaux ont montré qu'un prior plus flexible peut améliorer la qualité des représentations latentes. Le posterior agrégé (aggregated posterior) – une mixture des distributions  $q(z|x)$  obtenues pour différentes observations – a ainsi été proposé comme une approximation optimale dans certains cas (Tomczak and Welling, 2018). Plus récemment, des approches ont proposé d'apprendre directement le *prior* à partir des données, en le rendant adaptatif ou conditionné (Ilse et al., 2020; Klushyn et al., 2019; Sohn et al., 2015). Dans ces cas, le *prior* dépend d'informations supplémentaires, comme des variables discrètes (diagnostic) ou continues (âge), injectées dans la distribution  $p(z|y)$  (Ilse et al., 2020; Zhao et al., 2019a).

### I.3.1.3. Information contenue dans l'espace latent

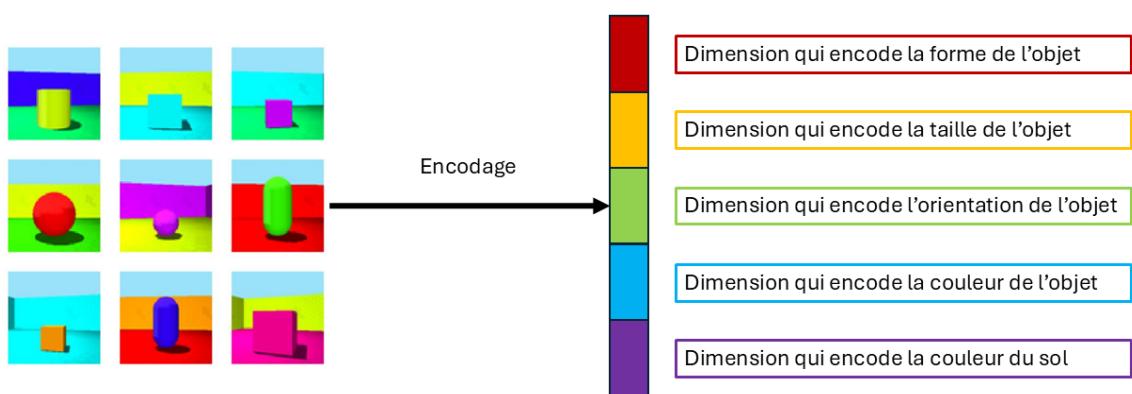
Un défi majeur dans l'entraînement des VAE est l'appauvrissement de la représentation latente. Une régularisation trop forte, via le terme de divergence KL, peut entraîner un posterior collapse, c'est-à-dire une situation où les représentations latentes deviennent non informatives (Alemi et al., 2018; Dieng et al., 2019). Pour éviter ce phénomène, plusieurs stratégies ont été proposées, consistant à découpler temporairement l'apprentissage de l'encodage et la régularisation de l'espace latent. Ces approches visent à maintenir un équilibre entre expressivité des représentations et qualité de reconstruction (Fu et al., 2019; Klushyn et al., 2019; Rezende and Viola, 2018).

### I.3.2. Extensions du VAE et disentanglement learning

Dans cette sous-partie, nous présentons certaines extensions du VAE. D'autres extensions existent évidemment, mais celles-ci ont été utilisées dans le cadre des travaux et doivent donc être introduites. Avant cela, il convient d'étudier la notion de disentanglement learning (apprentissage désenchevêtré).

#### I.3.2.1. Disentanglement learning : principe et définition

Le disentanglement learning (apprentissage désenchevêtré) a pour objectif de structurer l'espace latent afin que chaque dimension encode un facteur de variation distinct et indépendant, reflétant une caractéristique spécifique de la donnée (Figure I-8).



10. Figure I-8. Disentanglement learning.

Chaque dimension de l'espace latent contrôle une caractéristique des données initiales. (Image de gauche issue du dataset 3D Shapes : <https://github.com/google-deepmind/3d-shapes/tree/master> 05/08/2025).

#### I.3.2.2. $\beta$ -VAE et biais inductifs

Le  $\beta$ -VAE constitue une extension du VAE classique, dans laquelle un facteur  $\beta > 1$  est introduit dans la fonction de perte ELBO (Eq. I-2) (Higgins et al., 2022). Ce paramètre permet d'accentuer la régularisation en forçant chaque dimension de l'espace latent à suivre une distribution normale centrée réduite, avec une covariance diagonale (souvent l'identité). Cette contrainte favorise une séparation explicite des dimensions latentes, en imposant une variance diagonale et donc une indépendance entre elles, ce qui rend l'espace latent plus structuré et facilite un éventuel désenchevêtement.

$$ELBO = -E_{q(z|x)}[\ln p(x|z)] + \beta KL(q(z|x) || p(z)) \quad (\text{Eq. I-2})$$

Toutefois, il a été démontré qu'un désenchevêtrement totalement non supervisé est théoriquement impossible sans hypothèses fortes, soit sur la distribution des données, soit sur l'architecture du modèle (Locatello et al., 2019). L'introduction de biais inductifs est donc devenue une stratégie incontournable. Ces biais peuvent être incorporés à différents niveaux : dans la structure du modèle (ex. séparation explicite de sous-espaces), dans la fonction de coût (ex. pénalisation spécifique), ou dans le jeu de données (ex. en contrôlant les sources de variation). Dans le cas des autoencodeurs variationnels (ou même dans la conception des espaces latents en général), on suppose une indépendance conditionnelle entre différents facteurs (espaces latents), chacun étant guidé par un attribut particulier (Eq. I-3).

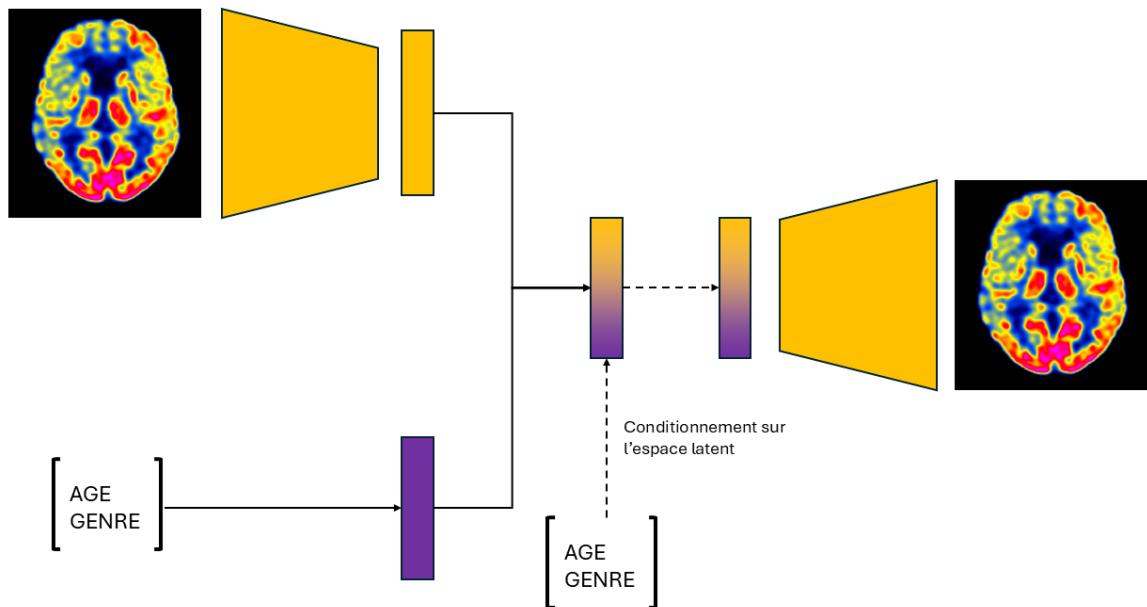
$$p(z_{AGE}, z_{diagnostic} | AGE, diagnostic) = p(z_{AGE} | AGE) p(z_{diagnostic} | diagnostic) \quad (\text{Eq. I-3})$$

Dans l'exemple ci-dessus, l'espace latent  $z_{AGE}$  est explicitement contrôlé par l'âge, tandis que  $z_{diagnostic}$  est contrôlé par le diagnostic. De plus, aucun des deux ne peut a priori contrôler le facteur de l'autre.

### I.3.2.3. VAE conditionnel

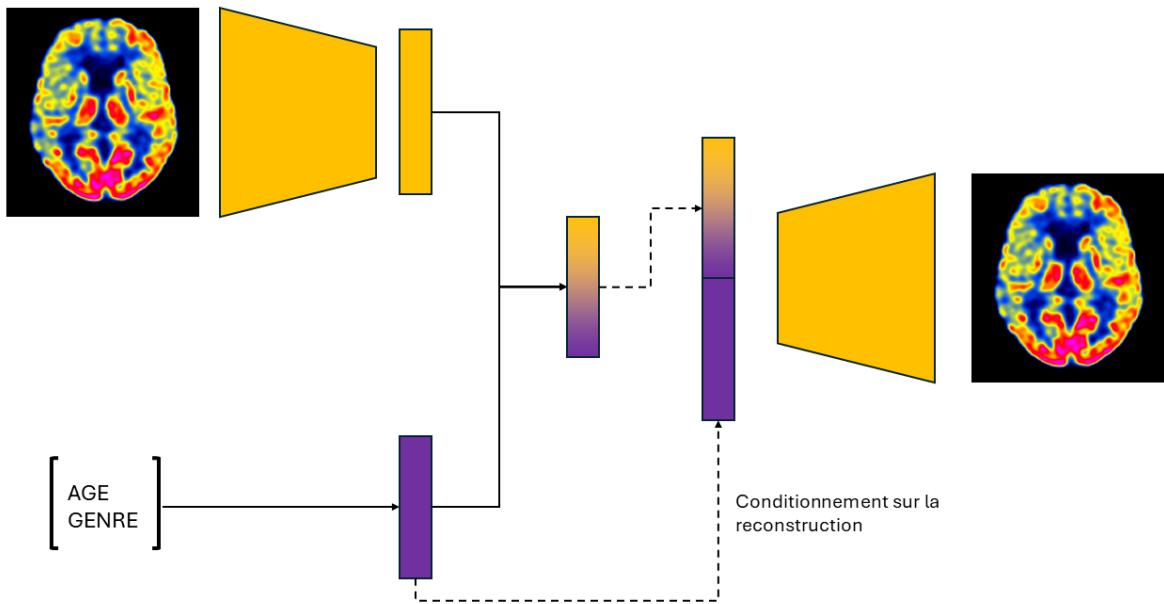
Dans le VAE conditionnel, l'espace latent (Figure I-9 et Eq. I-4) ou la reconstruction (Figure I-10 et Eq. I-5) est conditionné(e) par des variables auxiliaires. Ces variables peuvent être utilisées lors de l'encodage de l'espace latent, mais leur utilisation n'est pas obligatoire.

Dans les exemples présentés, elles sont utilisées pour l'encodage, ce qui permet d'établir une analogie avec les méthodes multimodales présentées dans la section I.2.2. Contrairement au VAE classique dont l'objectif est d'approximer  $p(x)$ , celui du VAE conditionnel est d'approximer  $p(x | y)$  où  $y$  correspond à des variables auxiliaires.



11. Figure I-9. VAE conditionnel sur l'espace latent.

Dans cet exemple, les variables cliniques et l'imagerie sont encodées afin de générer un espace latent commun. Les variables cliniques sont ensuite réutilisées pour conditionner cet espace latent commun.



12. Figure I-10. VAE conditionnel sur la reconstruction.

Dans cet exemple, les variables cliniques et l'imagerie sont encodées afin de générer un espace latent commun. Les variables cliniques sont ensuite réutilisées pour conditionner la reconstruction de l'imagerie.

$$ELBO_1 = -E_{q(z|x,y)}[\ln p(x|z)] + \beta KL(q(z|x,y) || p(z|y)) \quad (\text{Eq.I-4})$$

$$ELBO_2 = -E_{q(z|x,y)}[\ln p(x|z,y)] + \beta KL(q(z|x,y)||p(z)) \quad (\text{Eq.I-5})$$

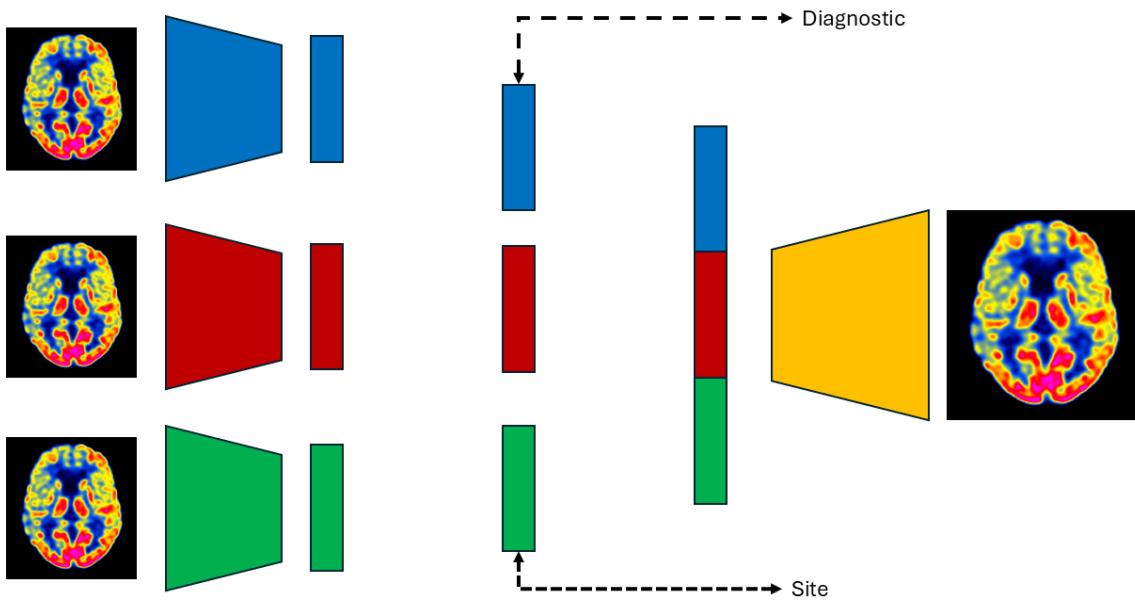
Au cours de cette thèse, nous étudierons l'utilisation de ces deux méthodes de conditionnement et leur impact sur l'espace latent, ainsi que les situations dans lesquelles les données auxiliaires doivent être incluses en entrée du modèle.

#### I.3.2.4. Domain Invariant Variational Autoencoder (DIVA)

Comme vu précédemment avec le  $\beta$ -VAE, l'introduction de biais inductifs constitue une stratégie incontournable pour désenchevêtrer l'espace latent. En supposant des espaces latents désenchevêtrés et guidés par des critères supervisés, les probabilités entre les encodeurs sont considérées comme indépendantes, de même que les probabilités entre les priors.

L'idée centrale est de définir explicitement les facteurs de variation que l'on souhaite voir émerger, ce qui est à la base de l'architecture DIVA (Ilse et al., 2020). Les auteurs divisent l'espace latent en trois sous-espaces distincts : l'un induit par la classe de la donnée, un autre par le site d'acquisition, et un dernier dédié aux variations résiduelles. Cette structuration permet d'isoler les sources de variation connues de celles qui ne sont pas pertinentes pour la tâche ciblée (Figure I-11).

On note également que l'architecture DIVA intègre le concept de conditionnement de l'espace latent présenté précédemment.



13. Figure I-11. Modèle DIVA.

Trois espaces latents indépendants sont générés par trois encodeurs. Deux de ces espaces sont guidés par la classification et le conditionnement sur le site et le diagnostic, tandis qu'un espace latent est régularisé par une gaussienne centrée réduite classique. Les trois espaces latents sont ensuite utilisés pour décoder l'image d'origine.

La méthode DIVA présente plusieurs avantages. Grâce à l'architecture VAE, le modèle bénéficie d'une régularisation naturelle, liée à la continuité de l'espace latent ainsi qu'à l'aspect multitâche. Le critère de disentanglement est également intégré en séparant explicitement les sous-espaces latents. De plus, le modèle renforce le conditionnement des espaces latents en ajoutant une tâche de supervision. Toutefois, cette approche présente un inconvénient majeur : elle nécessite un nombre croissant d'encodeurs lorsque l'on souhaite introduire davantage de critères désenchevêtrés, ce qui augmente considérablement la dimension de l'espace latent ainsi que les coûts mémoire et computationnels.

### I.3.3. Favoriser l'interprétation : intégrer des critères de similarités

En complément des contraintes probabilistes ou architecturales, il est possible de renforcer l'interprétabilité de l'espace latent en y intégrant des critères de similarité. Inspirées des approches auto-supervisées, ces régularisations visent à organiser les représentations latentes en fonction de relations sémantiques, morphologiques ou cliniques — en particulier dans les contextes médicaux. L'intégration d'un critère de consistance, appliqué entre des données brutes ou des versions transformées de ces données, permet d'ajuster la structure de l'espace latent en rapprochant les représentations de données partageant des caractéristiques communes, et en éloignant celles qui sont dissemblables. Dans une première formulation, ce critère imposait que la distribution latente d'une donnée transformée soit incluse dans celle de la donnée d'origine (Sinha and Dieng, 2022). Ce mécanisme contribue à stabiliser les représentations tout en préservant leur expressivité. Néanmoins, une pondération excessive de ce critère peut altérer la structure naturelle des données, en réduisant des variations informatives ou en rigidifiant l'espace latent. Il est donc essentiel de calibrer finement ce type de régularisation. De manière plus générale, ces critères de consistance,

tout comme d'autres formes d'auto-supervision, peuvent être définis soit de façon explicite, par exemple à partir d'un diagnostic connu (supervision partielle), soit de manière implicite, en s'appuyant sur la structure interne des données elles-mêmes.

## I.4. Les espaces latents en neuroimagerie

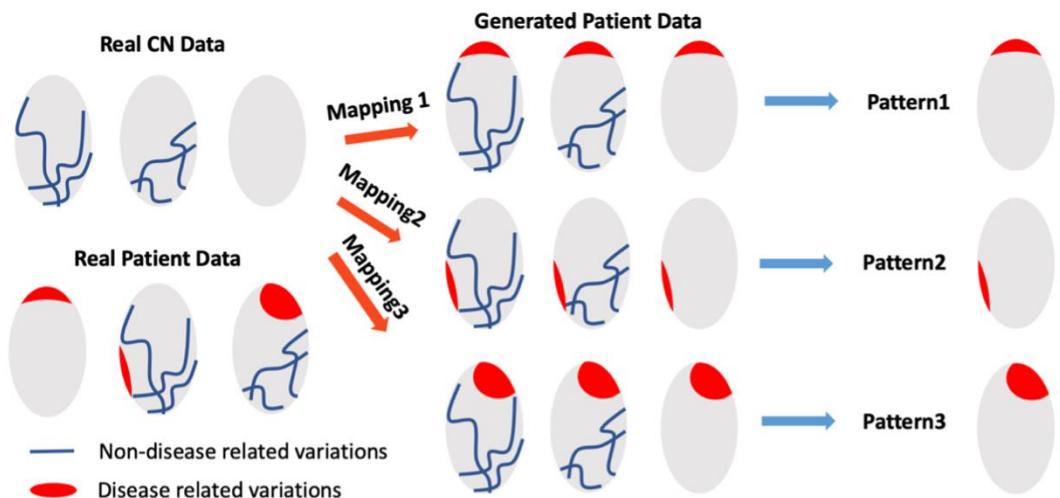
Les espaces latents sont devenus un outil central en neuroimagerie pour traiter la complexité et la diversité des données cérébrales. Ils permettent de capturer les principales sources de variation d'intérêt. Plusieurs grandes catégories d'applications émergent dans la littérature.

### I.4.1. Harmonisation de séquence IRM

L'une des utilisations les plus répandues concerne l'harmonisation inter-sites. Les images cérébrales acquises dans différents centres ou sur des scanners distincts sont sujettes à des variations techniques qui peuvent biaiser les analyses. Les modèles d'apprentissage utilisent alors deux espaces latents : l'un dédié au contenu biologique de l'image, l'autre au style, représentant les spécificités liées au site d'acquisition. Cette séparation permet de neutraliser les effets non biologiques tout en préservant les informations d'intérêt (Caldera et al., 2025; Liu and Yap, 2024; Zuo et al., 2023).

### I.4.2. Désenchevêtrement de pattern de maladie

Un second champ d'application concerne le désenchevêtrement des facteurs pathologiques dans les données d'imagerie cérébrale. L'objectif est d'isoler les composantes associées à la maladie de celles reflétant des variations normales (âge, sexe, variabilité interindividuelle). Une stratégie fréquente consiste à utiliser une base de sujets sains comme référence. Les différences observées chez les patients permettent alors de dégager des signatures pathologiques spécifiques (Figure I-12).



14. Figure I-12. SMILE-GAN.

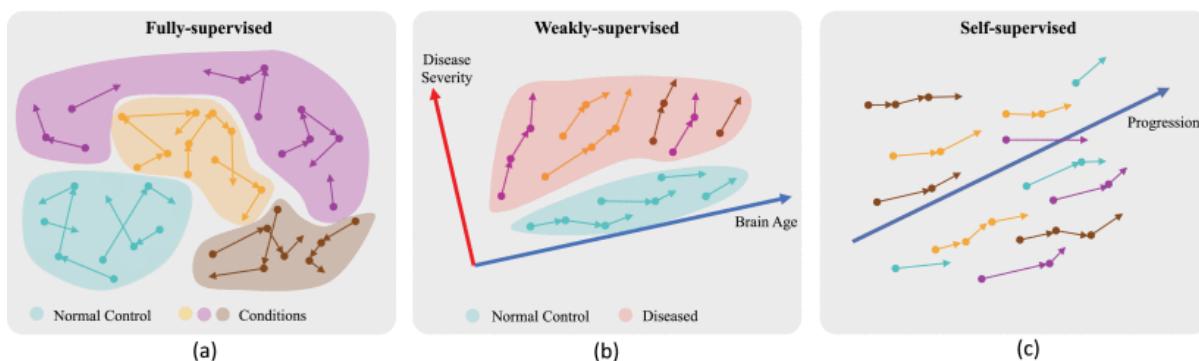
Illustration du disentanglement learning appliquée au pattern de maladie. (Image issue de SMILE-GAN <https://github.com/zhijian-yang/SmileGAN/blob/main/datasets/Smile-GAN.png>). Le modèle apprend à séparer les variations liées à la pathologie de celles observées dans une population saine.

SMILE-GAN utilise des réseaux antagonistes génératifs (GANs) pour combiner des images de sujets sains et des informations issues d'un cluster pathologique, générant ainsi une image artificielle correspondant à un patient malade (Yang et al., 2021). Dans une approche similaire, SepVAE adopte

une approche VAE utilisant deux encodeurs distincts : l'un apprend un espace latent partagé entre sujets sains et malades, tandis que l'autre capture spécifiquement les variations liées à la pathologie (Louiset et al., 2024). Enfin, Kang et al. (Kang et al., 2024) adoptent une approche indirecte : elle utilise des vecteurs d'atrophie dérivés d'un modèle normatif, puis applique un apprentissage auto-supervisé pour construire un espace latent structuré. Ce dernier permet d'identifier quatre sous-types d'atrophie : diffus, pariétal, temporo-médial et minimal.

### I.4.3. Etude de progression de maladie

Les espaces latents offrent également un cadre pertinent pour modéliser la progression des maladies neurodégénératives. L'évolution temporelle de la pathologie peut ainsi être décrite comme une trajectoire au sein de cet espace, à partir de données d'imagerie — longitudinales ou transversales — enrichies de variables cliniques. Comme mentionné précédemment, le modèle BrainAGE a pour objectif d'estimer l'âge chronologique d'un individu sain à partir de son imagerie cérébrale. Or, il a été montré que les patients atteints de pathologies neurodégénératives présentent fréquemment un âge cérébral prédict supérieur à leur âge réel. En outre, plusieurs études ont mis en évidence que l'espace latent exploité pour l'estimation de l'âge contient également des informations cliniques pertinentes, allant au-delà de la seule prédiction de l'âge (Cole et al., 2017; Leonardsen et al., 2022). D'autres modèles cherchent à dissocier vieillissement normal et progression pathologique. L'étude de Ouyang et al (Ouyang et al., 2022) propose de structurer l'espace latent selon deux axes indépendants : l'un pour le vieillissement naturel et l'autre reflétant la sévérité de la pathologie (Figure I-13). En s'appuyant sur des hypothèses fortes (indépendance des deux processus, constance du vieillissement), ces modèles apprennent à projeter chaque sujet dans un espace permettant une lecture conjointe et séparée de ces dynamiques. Cette structuration repose souvent sur des VAE et l'analyse de cohortes longitudinales.



15. Figure I-13. Vieillissement naturel et progression de la maladie.

Illustration des hypothèses de modélisation du vieillissement normal et de la progression pathologique (Figure I-13.b), affectant la structure cérébrale de manière indépendante. (Image issue de (Ouyang et al., 2022)).

### I.4.4. Association compte rendu image

Dans l'association des images aux comptes rendus, l'espace latent généré correspond à un espace latent commun aux deux modalités.

Dans un premier temps, une brève introduction sera consacrée aux modèles de traitement du texte, puis sera présentée l'utilisation des Large Language Models (LLMs) dans les comptes rendus ainsi que dans la construction des espaces latents. Cette section n'a pas pour objectif de présenter l'ensemble des modèles de langage, mais simplement d'en exposer les concepts généraux.

#### I.4.4.1. Large Language Models (LLMs)

Depuis quelques années, les LLMs constituent l'état de l'art en traitement du langage naturel. Cette section présente de manière générale leur fonctionnement, en suivant les principales étapes : la conversion du texte en tokens, le rôle du mécanisme d'attention, l'utilisation du token spécial [CLS], et enfin la prise en compte de l'ordre des tokens.

##### Du texte au vecteur

La première étape consiste à transformer le texte en une représentation numérique exploitable par un modèle de deep learning. Schématiquement, le processus se déroule en trois temps : (1) extraction des tokens à partir du texte ; (2) association de chaque token à un ID unique dans le vocabulaire ; (3) récupération du vecteur numérique correspondant à cet ID dans la table d'embeddings, utilisé ensuite par le modèle. Un token peut représenter un mot, une sous-partie de mot, un groupe de mots, voire un ensemble de caractères.

##### Les couches de Self Attention

Le cœur des LLMs repose sur les couches Transformers (Vaswani et al., 2023), dont le mécanisme central est l'attention. Celui-ci permet de modéliser les relations entre tokens et d'enrichir leur représentation contextuelle.

Ce processus repose sur trois composantes : (1) la query (Q), qui définit ce que l'on cherche dans la séquence ; (2) la key (K), qui permet d'identifier l'information pertinente ; et (3) la value (V), qui correspond à la valeur extraite lorsque la correspondance entre query et key est trouvée.

Le score d'attention entre une query Q et une key K est obtenu par un produit scalaire entre Q et K, normalisé par la racine carrée de la dimension  $d_k$  de la key, puis passé par une fonction softmax (Eq. I-6). Ce score indique la force de la relation : une valeur proche de zéro traduit une absence de lien, tandis qu'une valeur élevée reflète une dépendance forte. Les valeurs (V) sont ensuite pondérées par ces scores afin de produire la représentation finale des tokens.

Enfin, pour capturer simultanément plusieurs types de relations contextuelles, le mécanisme est décliné en plusieurs têtes d'attention indépendantes (multi-head attention), dont les résultats sont concaténés. Cela permet au modèle d'extraire différents points de vue sur la même séquence textuelle et d'obtenir une représentation plus riche et expressive.

$$\text{Attention} = \text{softmax}\left(\frac{QK}{\sqrt{d_k}}\right) \quad (\text{Eq. I-6})$$

##### Comprendre la séquence entière : le token [CLS]

Le traitement d'une séquence par un LLM génère une matrice de taille ( $N\_tokens, D\_embeddings$ ), où chaque ligne correspond à l'embedding contextuel d'un token. Cependant, pour des tâches comme la classification, il faut obtenir une représentation unique de la séquence entière.

C'est le rôle du token [CLS] (classification token). Placé en début de texte, il est entraîné pour agréger l'information issue de tous les tokens via les couches de Transformers. À la sortie, son embedding constitue une représentation condensée de la séquence, directement exploitable pour des tâches supervisées telles que la classification de texte. Plus largement, il peut aussi être utilisé dans d'autres contextes nécessitant un espace latent global, par exemple pour l'association entre imagerie et compte rendu.

### Préserver l'ordre des tokens

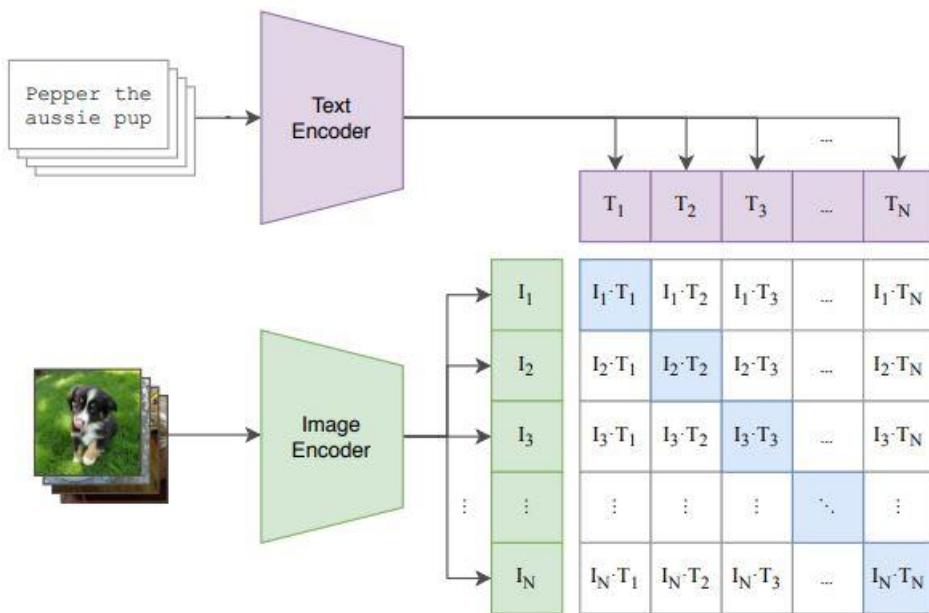
Un dernier élément essentiel est la prise en compte de l'ordre des tokens, indispensable pour comprendre correctement le texte. Comme les Transformers ne possèdent pas de notion intrinsèque de séquentialité, il faut ajouter un encodage positionnel aux embeddings. Cet encodage peut être : fixe et déterministe, comme dans l'architecture originale des Transformers, où les positions sont codées par des combinaisons sinusoïdales de différentes fréquences, ou appris conjointement avec les embeddings. Ces vecteurs de position sont additionnés aux embeddings des tokens avant leur passage dans les couches de self-attention.

### I.4.4.2. Extension des transformers aux images

Les modèles Transformers, initialement développés pour le traitement du texte dans les LLMs, ont ensuite été généralisés à d'autres types de séquences, notamment les images. On parle alors de Vision Transformers (ViTs)(Dosovitskiy et al., 2020). Dans ce cas, l'image est découpée en patchs de taille fixe, chacun étant considéré comme un token. Ces tokens visuels sont ensuite traités par le modèle de la même manière que les tokens textuels dans les LLMs.

### I.4.4.3. Association texte-image

Les méthodes d'apprentissage texte-image s'appuient sur des principes proches de ceux du contrastive learning. L'idée est de rapprocher une image et son texte associé tout en éloignant les autres associations possibles (Figure I-15)



16. Figure I-14. Association texte-image.

Image provenant du modèle CLIP (Radford et al., 2021). Dans ce cas, l'objectif est de contraindre la matrice d'association à se rapprocher d'une matrice identité, de façon à aligner chaque image uniquement avec son texte associé.

En neuroimagerie, une voie de recherche particulièrement prometteuse consiste à apprendre des espaces latents partagés entre imagerie médicale et texte clinique. Ces approches reposent sur l'alignement multimodal et visent à construire une représentation conjointe des données visuelles

(IRM, TEP, etc.) et textuelles (comptes rendus, observations cliniques). Ce type d'apprentissage est au cœur du développement des modèles de fondation biomédicaux, à l'instar des modèles multimodaux vision-langage utilisés dans d'autres domaines. L'objectif est de tirer parti de cette représentation commune afin d'améliorer des tâches telles que le diagnostic assisté, le suivi longitudinal des patients ou encore leur stratification (Blankemeier et al., 2024).

## I.5. Rappel des objectifs de thèse

Comme précisé dans l'introduction, cette thèse a pour objectif de développer des méthodes d'apprentissage d'espaces latents structurés, qu'ils soient unimodaux ou multimodaux, dans le contexte de l'analyse en neuroimagerie.

Le chapitre II explore dans quelle mesure un espace latent appris en supervision peut extraire non seulement les informations directement liées à la tâche cible, mais également des dimensions plus générales, révélant ainsi une structuration implicite et informative au-delà de la supervision explicite.

Le chapitre III présente la conception d'un espace latent structuré et interprétable reposant sur une architecture de type DIVA, intégrant l'âge et l'imagerie cérébrale comme variables d'entrée. Cet espace est exploité pour analyser les démences fronto-temporales (DFT) à travers une approche par plus proches voisins, permettant d'explorer conjointement les profils d'imagerie et les caractéristiques cliniques des patients.

Le chapitre IV étend cette approche à un cadre multimodal en combinant IRM T1w et FDG-TEP. L'espace latent y est structuré selon une logique différente, l'âge du patient servant à conditionner le prior du VAE afin d'introduire une régularisation temporelle et de refléter la progression de la maladie. Cette méthode est appliquée à des cohortes comprenant des témoins cognitivement normaux, des patients atteints de la maladie d'Alzheimer et des patients atteints de DFT.

Enfin, le chapitre V constitue une ouverture prospective. Il explore l'extension de la multimodalité à des sources d'information plus hétérogènes, incluant des séquences IRM FLAIR et DWI, les comptes rendus radiologiques et les données cliniques tabulaires. Ces modalités sont intégrées dans un modèle multimodal de type vision-langage, où des mécanismes d'attention assurent la fusion des informations. L'objectif est de prédire le devenir fonctionnel à trois mois chez des patients victimes d'un AVC ischémique et traités par thrombectomie mécanique.

## Chapitre II

# Article « BrainAGE latent representation clustering is associated with longitudinal disease progression in early onset Alzheimer's disease

---

Le chapitre I a posé les bases théoriques des espaces latents, en détaillant leurs propriétés et leur intérêt pour l'analyse de données complexes. Dans le chapitre II, nous passons de la théorie à l'application : un modèle BrainAGE pré-entraîné sur une population saine est utilisé pour extraire un espace latent sur des patients atteints de la maladie d'Alzheimer à début précoce. Cette approche illustre comment les espaces latents peuvent capturer la complexité des dimensions phénotypiques et permet d'identifier des sous-groupes de patients par clustering, montrant ainsi le potentiel concret de ces représentations pour la compréhension de la maladie.

## II.1. Introduction

La maladie d'Alzheimer (MA) est un trouble neurodégénératif complexe et hétérogène, caractérisé par une variabilité interindividuelle significative dans son apparition, sa progression et ses manifestations cliniques. Cette hétérogénéité est évidente dans l'éventail de symptômes, les rythmes de déclin cognitif et les caractéristiques neuropathologiques observés chez les patients. Comprendre cette hétérogénéité est crucial pour améliorer la précision des outils diagnostiques dans la MA et développer des approches thérapeutiques plus personnalisées (Devi and Scheltens, 2018).

Parmi les facteurs contribuant à cette hétérogénéité, l'âge d'apparition joue un rôle critique. Les patients atteints de la forme précoce de la MA (EOAD, *Early-Onset Alzheimer's Disease*) sont plus susceptibles de connaître une progression plus rapide (Mendez, 2017) et présentent souvent des caractéristiques cliniques atypiques, généralement sous forme de symptômes non mnésiques (Koedam et al., 2010). Cependant, stratifier des patients atteints d'EOAD en fonction de la progression de leur maladie demeure difficile, en grande partie en raison de la résilience individuelle face au développement neuropathologique (réserve cérébrale) et de la capacité à compenser le déclin cognitif (réserve cognitive).

Ainsi, l'établissement de biomarqueurs cliniquement disponibles permettant d'identifier les individus présentant des vitesses de progression similaires est d'une importance capitale pour la médecine clinique et de précision. L'imagerie par résonance magnétique (IRM) pourrait constituer un outil puissant pour identifier l'atrophie neuroanatomique étroitement associée à la distribution de la neuropathologie tau (Theriault et al., 2021; Xia et al., 2017). L'atrophie corticale globale est plus prononcée dans l'EOAD comparée à la MA à début tardif (Aziz et al., 2017; Migliaccio et al., 2015). Mais des profils d'atrophie focale ont été associés à des sous-types biologiques distincts de la maladie (Ossenkoppele et al., 2015a). Une évaluation approfondie de la structure cérébrale est donc nécessaire, ce qui peut être réalisé grâce à l'application de techniques d'intelligence artificielle (IA). L'IA, y compris l'apprentissage automatique et l'apprentissage profond, a la capacité de détecter des motifs complexes et subtils dans les images IRM.

Récemment, plusieurs modèles d'IA, regroupés sous le terme de BrainAGE (Brain Age Gap Estimation), ont été développés pour évaluer les écarts de la structure cérébrale par rapport à la trajectoire normale du vieillissement chez les individus cognitivement sains (Franke and Gaser, 2019; Guo et al., 2024). Ces méthodes consistent à entraîner un algorithme à prédire l'âge d'un individu sain à partir des images IRM de son cerveau. Le score BrainAGE, défini comme la différence entre l'âge cérébral estimé et l'âge chronologique, peut refléter la résilience ou la vulnérabilité du cerveau. Un score plus jeune peut indiquer une plus grande résilience face au déclin cognitif et à la progression de la maladie, tandis qu'un score plus âgé pourrait suggérer une résilience réduite et un risque accru de troubles cognitifs, reliant directement le modèle à la fois à la résilience cérébrale et à la santé cognitive. Le score BrainAGE, appris à partir de modèles d'apprentissage profond, encode une série de caractéristiques hiérarchiques au travers de couches convolutionnelles. Ces caractéristiques capturent et condensent l'information essentielle des images IRM pour estimer l'âge du sujet. Les caractéristiques condensées obtenues forment une représentation encodée des images IRM. En raison de sa haute dimension, cette représentation encodée peut également contenir des informations supplémentaires pour le phénotypage cérébral individuel au-delà de l'estimation de l'âge (Leonardsen et al., 2022).

Dans une étude précédente, nous avons démontré que le score BrainAGE était associé à la sévérité du déficit cognitif chez les patients atteints d'EOAD, et dans une moindre mesure, à leurs phénotypes neuropsychologiques (Gautherot et al., 2021). L'objectif de cette étude était de développer un cadre d'IA basé sur le modèle BrainAGE et l'IRM structurelle afin de réaliser une stratification des patients atteints d'EOAD. Nous faisons l'hypothèse que les représentations encodées apprises par le modèle BrainAGE constituent un biomarqueur de la résilience cérébrale face à la progression de la maladie, et que l'application d'un algorithme de regroupement à ces représentations permettra d'identifier des groupes de patients présentant des rythmes distincts de progression de la maladie, indépendamment de leur phénotype.

## II.2. Matériels et méthodes

### II.2.1. Population EOAD

Dans la présente étude, des données acquises prospectivement ont été analysées rétrospectivement. Les patients atteints d'EOAD ont été recrutés et examinés au Centre Mémoire de Ressources et de Recherche du Centre Hospitalier Universitaire de Lille, France. Ces participants faisaient partie de la cohorte COMAJ (Maureille et al., 2017). Les critères d'inclusion étaient les suivants : apparition des symptômes à 60 ans ou plus tôt, et respect des critères NIA-AA (McKhann et al., 2011). Ces évaluations comprenaient antécédents médicaux, examens neurologiques et neuropsychologiques, génotypage, imagerie par résonance magnétique (IRM) et imagerie au 18TEP FDG, suivi de progression de la maladie ainsi que dosage des biomarqueurs du liquide céphalo-rachidien (LCR). Les patients ont été suivis annuellement pendant une période de 6 ans, avec des évaluations cliniques, neuropsychologiques et par IRM. Pour les besoins de la présente étude, les participants présentant une mutation génétique ou une forme familiale ont été exclus. L'étude a reçu l'approbation du comité d'éthique local (CPP Ile-de-France VI Groupe Hospitalier Pitié-Salpêtrière ; référence 110-05). Un consentement éclairé écrit a été obtenu auprès de tous les participants.

### II.2.2. Acquisition et prétraitement des IRM

Des images pondérées T1 tridimensionnelles (3D T1) en écho de gradient ont été acquises sur un scanner IRM 3T (Achieva, Philips, Best, Pays-Bas), utilisant une antenne tête à 8 canaux et une antenne corps entier pour la transmission (champ de vue =  $256 \times 256 \times 160 \text{ mm}^3$ , taille de voxel isotropique  $1 \times 1 \times 1 \text{ mm}^3$ , TR = 9,9 ms, TE = 4,6 ms et angle de bascule =  $8^\circ$ ) pour l'évaluation initiale et le suivi. Les images ont été corrigées des effets d'inhomogénéité du champ magnétique et le crâne a été retiré à l'aide du logiciel VolBrain (Manjón, 2016). Ensuite, les images 3D-T1 prétraitées ont été enregistrées linéairement dans l'espace MNI et rééchantillonnées à  $1 \text{ mm}^3$  à l'aide du logiciel SPM (<https://www.fil.ion.ucl.ac.uk/spm/software/>). Une normalisation d'intensité a été effectuée par normalisation min-max.

### II.2.3. Modèle BrainAGE

Nous avons utilisé le modèle BrainAGE pré-entraîné décrit par Gautherot et al.(Gautherot et al., 2021) comme extracteur de caractéristiques (Figure II-1), développé sur 3 227 IRM-T1 provenant de 2 065 participants sains. Une approbation éthique a été obtenue pour tous les sites de collecte de données, et un consentement éclairé a été obtenu auprès de chaque participant. Tous les individus ont été confirmés exempts de maladies neurologiques ou psychiatriques conformément aux

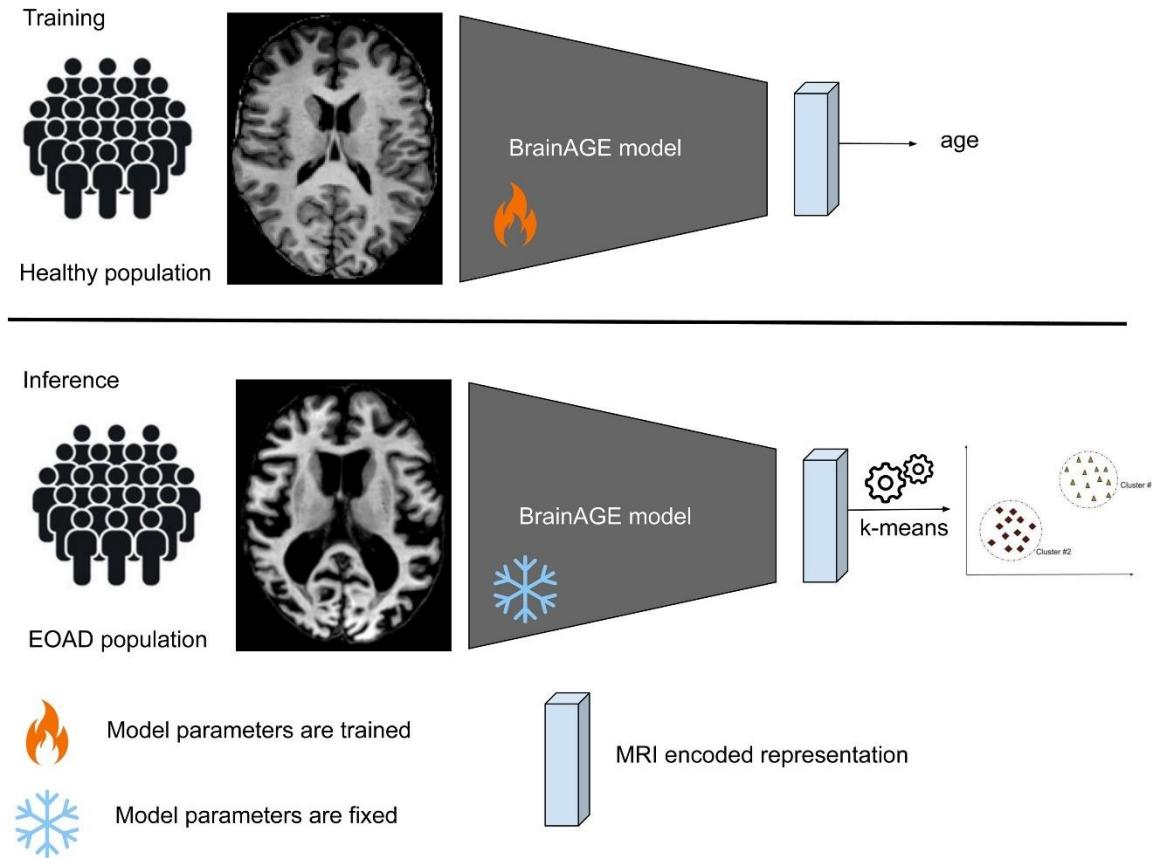
protocoles locaux des études. Les détails de la base de données utilisée pour entraîner le modèle BrainAGE sont fournis dans le Tableau II-1. Des détails supplémentaires comparant la population EOAD et la base d’entraînement sont présentés dans la Figure II-S2 et le Tableau II-S1.

Etude	N	Age, années	Age étendue	Femme, n (%)	SCANNER	Website
IXI	181	$46.87 \pm 16.73$	20-81	94 (51%)	PHILIPS	<a href="https://brain-development.org/">https://brain-development.org/</a>
HCP	1783	$28.76 \pm 3.7$	22-37	968 (54%)	SIEMENS	<a href="https://www.humanconnectome.org/">https://www.humanconnectome.org/</a>
COBRE	238	$38.18 \pm 11.53$	18-65	64 (6%)	SIEMENS	<a href="https://www.mrn.org/common/cobre-phase-3">https://www.mrn.org/common/cobre-phase-3</a>
MCIC	264	$33.60 \pm 12.21$	18-60	80 (30%)	SIEMENS	<a href="https://www.nitrc.org/projects/mcic/">https://www.nitrc.org/projects/mcic/</a>
NmorphCH	141	$31.37 \pm 8.42$	20-46	66 (46%)	SIEMENS	<a href="http://schizconnect.org/">http://schizconnect.org/</a>
NKIRS	620	$42.62 \pm 18.27$	18-85	406 (65%)	SIEMENS	<a href="http://fcon_1000.projects.nitrc.org/indi/enhanced/">http://fcon_1000.projects.nitrc.org/indi/enhanced/</a>

*1. Tableau II-1 Informations sur les sujets de la base d’entraînement du modèle BrainAGE.*

*L’âge est indiqué en moyenne  $\pm$  écart-type. L’intervalle d’âge est indiqué en valeurs minimale et maximale. Les variables catégorielles sont rapportées sous la forme nombre sur le nombre total (pourcentage).*

Le modèle BrainAGE était basé sur une architecture de réseau de neurones convolutifs tridimensionnels (3D-CNN) (Cole et al., 2017; Gautherot et al., 2021). L’architecture CNN se composait de cinq blocs répétés suivis d’une couche flatten et d’une couche linéaire pour prédire l’âge cérébral. Chaque bloc comprenait une convolution 3D, une activation ReLU, une convolution 3D, une normalisation par lot 3D, une activation ReLU et une opération de max-pooling.



17. Figure II-1. Pipeline de la méthode.

Haut : le modèle BrainAGE est entraîné sur une population saine afin de générer une représentation encodée de l'IRM capable de prédire l'âge.

Bas : le modèle BrainAGE est utilisé pour extraire la représentation encodée de l'IRM de chaque patient de la population EOAD, représentation ensuite soumise à un regroupement pour identifier des sous-groupes distincts d'EOAD.

La donnée entrée était une IRM pondérée T1 3D prétraitée avec des dimensions de  $182 \times 218 \times 182$  voxels. L'entraînement du modèle a été optimisé à l'aide de l'algorithme de descente de gradient stochastique (SGD) avec un taux d'apprentissage de 0,001, un momentum de 0,1 et un décroissement du taux d'apprentissage de 5e-05 (Sutskever et al., 2013). La fonction de coût était l'erreur absolue moyenne (MAE), et l'entraînement a été réalisé sur 150 époques avec une taille de lot de 8. Un arrêt précoce a été appliqué en fonction de la MAE de validation afin d'éviter le surapprentissage. Des techniques d'augmentation des données, incluant des translations et des rotations, ont été appliquées pendant l'entraînement pour limiter le surapprentissage et pour améliorer les performances, comme cela a été empiriquement observé (Shorten and Khoshgoftaar, 2019).

L'âge cérébral estimé par le modèle BrainAGE a été corrigé du phénomène de régression vers la moyenne (Barnett et al., 2005). Ce phénomène peut induire un biais systématique — dans notre cas, une surestimation de l'âge des participants les plus jeunes et une sous-estimation de celui des plus âgés. Pour pallier ce problème, nous avons ajusté l'âge prédit en utilisant le modèle de correction suivant (Liang et al., 2019) :

$$\text{regressed age predicted} = \text{intercept} + \alpha \times \text{age} + \text{error} \quad (\text{Eq.II-1.})$$

Comme notre analyse utilisait le modèle BrainAGE développé par Gautherot et al. (Gautherot et al., 2021), nous avons repris leur paramétrisation. La valeur de  $\alpha$  a été sélectionnée par validation croisée k-fold sur leur jeu de test, utilisé exclusivement pour ajuster les hyperparamètres du modèle BrainAGE et a été estimée à  $\alpha = 0,13$ .

#### **II.2.4. Algorithme de clustering**

La Figure II-1 illustre le pipeline utilisé pour l'analyse de clustering. Les représentations encodées de l'IRM de la population EOAD ont été extraites en utilisant le modèle BrainAGE pré-entraîné sur une cohorte saine. Ces représentations ont ensuite été utilisées comme entrée pour un algorithme de regroupement k-means (Lloyd, 1982). L'algorithme k-means a été exécuté avec la bibliothèque Python scikit-learn (Pedregosa et al., 2011), avec une initialisation k-means++ (Arthur and Vassilvitskii, 2007). Pour surmonter la haute dimension de la représentation encodée (19 200 caractéristiques), le k-means a été exécuté 150 fois et nous avons retenu le regroupement ayant la plus faible somme des carrés intra-classe. Le nombre optimal de clusters a été déterminé à l'aide de quatre métriques : le Gap Statistic, le score de Silhouette, les valeurs de Calinski-Harabasz et de Davies-Bouldin (Petersen et al., 2024).

#### **II.2.5. Évaluation neuropsychologique**

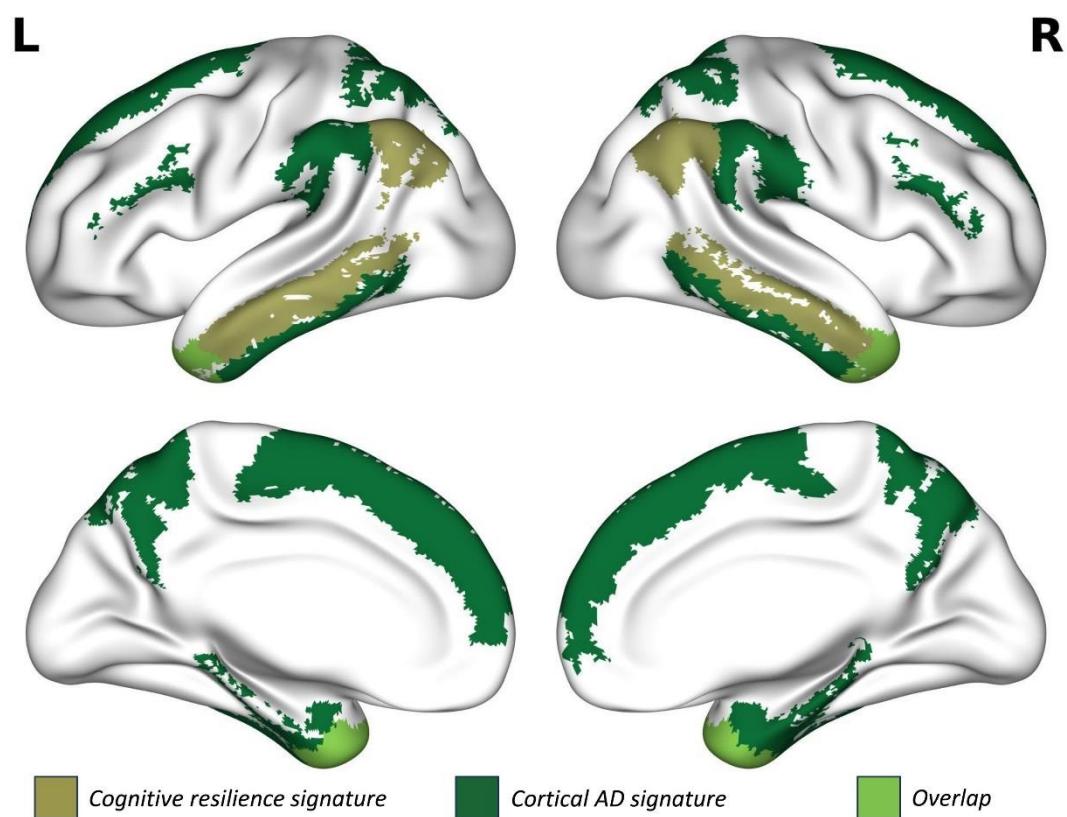
Afin de prendre en compte la sévérité de la maladie, le déclin cognitif et les phénotypes cliniques des participants au sein des clusters ont été collectés. Le score CDR-SoB (Clinical Dementia Rating – Sum of Boxes) a été utilisé comme marqueur de sévérité de la maladie. Si le CDR-SoB n'était pas disponible mais que le CDR avait été évalué, une valeur de CDR-SoB était attribuée à partir du score CDR correspondant, conformément aux recommandations d'O'Bryant et al. (O'Bryant et al., 2008b). Le score MMSE (Mini-Mental State Examination) a également été utilisé comme marqueur de sévérité de la maladie (Folstein et al., 1985). Si le score MMSE était supérieur à 10, une évaluation neuropsychologique détaillée était réalisée, couvrant quatre grands domaines cognitifs : la mémoire épisodique (évaluée avec le Visual Association Test (VAT) (Lindeboom et al., 2002)); le langage (test de dénomination DO80 avec 80 images); la fonction visuospatiale (test Beery VMI de Beery-Buktenica (Lim et al., 2015)); et les fonctions exécutives (fluences verbales catégorielles (animaux) et phonémiques (lettre P)). Concernant les phénotypes cliniques, des réunions rétrospectives multidisciplinaires ont été organisées pour déterminer l'atteinte cognitive initiale et la plus marquée des patients, sur la base de (i) l'anamnèse rapportée par l'aide au moment du diagnostic et (ii) le profil neuropsychologique à l'inclusion. Conformément aux critères NIA-AA modifiés pour la maladie d'Alzheimer probable (McKhann et al., 2011), les déficits cognitifs initiaux et prédominants (après l'amnésie) concernaient le langage, les fonctions visuospatiales ou exécutives. Par conséquent, dans cette étude, les phénotypes non mnésiques ont été catégorisés selon la présence de troubles du langage, des fonctions visuospatiales ou exécutives, correspondant respectivement à la variante logopénique de l'aphasie primaire progressive (Gorno-Tempini et al., 2011), à l'atrophie corticale postérieure (Crutch et al., 2017), et au variant comportemental/dysexécutif de la MA (Ossenkopp et al., 2015b).

#### **II.2.6. Analyses volumétriques et morphométriques**

Afin d'évaluer les facteurs neuroanatomiques appris par notre approche de regroupement basée sur le modèle BrainAGE, plusieurs analyses volumétriques et morphométriques cérébrales ont été effectuées. Les volumes cérébraux (substance grise, hippocampe, substance blanche et LCR) ont été

estimés à l'aide du logiciel VolBrain (Manjón and Coupé, 2016) et normalisés par le volume intracrânien.

De manière analogue aux phénotypes cognitifs (mnésique et non mnésique), nous avons défini des phénotypes d'atrophie cérébrale basés sur le ratio volume hippocampique/volume cortical (HV:CTV). Le phénotype Limbic-predominant (LP) correspondait à un HV:CTV inférieur au 25<sup>e</sup> percentile, le phénotype Hippocampal sparing (HpSp) à un HV:CTV supérieur au 75<sup>e</sup> percentile, et le Typical AD (tAD) correspondait à un HV:CTV situé entre le 25<sup>e</sup> et le 75<sup>e</sup> percentile (Kuchcinski et al., 2023; Risacher et al., 2017). La signature corticale de la MA (Dickerson et al., 2009; Dickerson and Wolk, 2012) et la signature de résilience cognitive (Arenaza-Urquijo et al., 2019) ont été évaluées en moyennant l'épaisseur corticale de régions d'intérêt spécifiques issues de l'atlas de Destrieux (Destrieux et al., 2010) (Figure II-2). L'épaisseur corticale a été estimée avec le logiciel Freesurfer (v.7.1.0, <https://surfer.nmr.mgh.harvard.edu/>).



18. Figure II-2. Régions d'intérêt pour les signatures corticales de l'AD et cognitives  
Signature corticale de la MA : gyrus frontal supérieur et inférieur – gyrus supramarginal – gyrus pariétal supérieur – gyrus précuneus – gyrus parahippocampique – gyrus temporal inférieur  
Signature de résilience cognitive : gyrus cingulaire antérieur – gyrus angulaire – gyrus temporal moyen  
Chevauchement entre la signature AD et la signature de résilience : pôle temporal

### II.2.7. Analyses statistiques

Les analyses statistiques des données démographiques, biologiques et cliniques ont été menées afin de différencier les caractéristiques des patients au sein de chaque cluster. Ces analyses ont été effectuées à l'aide du logiciel R (v.4.2.2, <https://www.r-project.org/>). Le seuil de significativité statistique a été fixé à  $p < 0,05$ , corrigé pour comparaisons multiples à l'aide du False Discovery Rate (FDR).

#### **II.2.7.1. Analyse à l'inclusion**

Afin de comprendre les bases biologiques des clusters BrainAGE, nous avons évalué les différences intergroupes à l'inclusion concernant les données démographiques et biologiques entre les clusters en utilisant une régression linéaire pour les variables continues et une régression logistique pour les variables catégorielles. Les variables significatives ont été incluses comme covariables dans un modèle de régression linéaire pour comparer les différences intergroupes des scores neuropsychologiques et des mesures volumétriques en IRM.

#### **II.2.7.2. Analyse longitudinale**

Pour estimer la valeur pronostique des clusters BrainAGE, nous avons conduit une analyse longitudinale des marqueurs de sévérité de la maladie, ajustée sur la sévérité à l'inclusion. Les différences entre clusters pour les scores neuropsychologiques et les mesures volumétriques en IRM ont été examinées à l'aide de modèles linéaires mixtes univariés (LME). Cette approche offre flexibilité et puissance pour analyser des données longitudinales, en tenant compte des données manquantes, de temps de suivi non uniformes et en incluant des participants ne disposant que d'une seule mesure temporelle, afin de caractériser des différences spécifiques à la population. Les modèles LME ont été ajustés avec (i) des effets fixes, incluant l'âge chronologique à l'inclusion, le sexe, le groupe de cluster, l'interaction groupe × temps de suivi et (ii) des effets aléatoires limités à l'intercept, car l'objectif était d'étudier les effets des clusters entre les groupes plutôt que sur les patients individuellement.

La valeur prédictive de la représentation encodée de BrainAGE a été comparée à celle de la différence entre âge prédit et âge chronologique (PAD) et de la réserve cérébrale (BR). Les analyses longitudinales ont été répétées avec les groupes PAD et BR. Le PAD a été calculé comme la différence entre l'âge chronologique et l'âge cérébral prédit corrigé (Gautherot et al., 2021; Leonardsen et al., 2022). La BR a été estimée à l'aide du volume intracrânien (ICV) (Stern et al., 2019; Sumowski et al., 2013). Les groupes ont été créés en fonction de la médiane de chaque biomarqueur. Nous avons ensuite construit un vecteur de caractéristiques manuellement pour chaque sujet afin de fournir un autre point de comparaison — au-delà de la simple stratification de la population par PAD ou BR — pour notre méthode d'IA. Ce vecteur incluait les volumes normalisés de substance grise, de substance blanche, de LCR, de l'hippocampe, ainsi que l'âge du sujet. Nous avons ensuite appliqué l'algorithme de clustering k-means afin de regrouper les sujets en deux clusters distincts sur la base de ces caractéristiques.

### **II.3. Résultats**

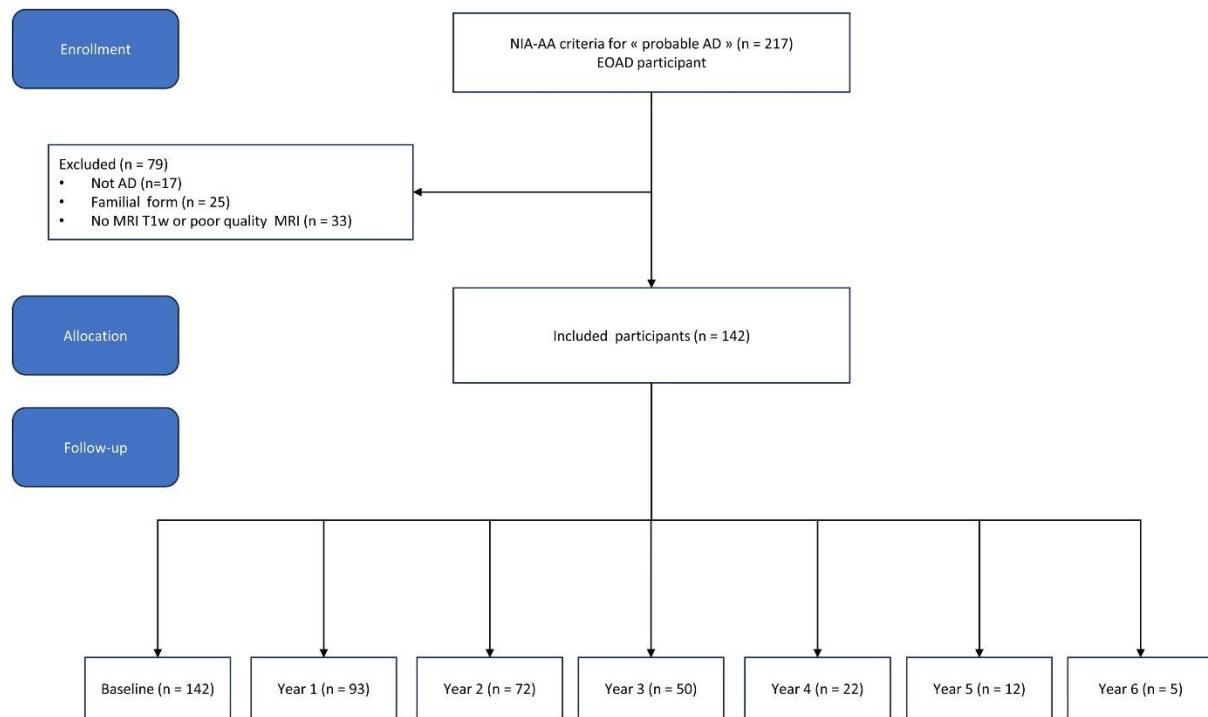
#### **II.3.1. Participants**

De 2009 à 2017, un total de 217 participants répondant aux critères de MA probable ont été inclus dans cette étude. Sur la base de la disponibilité et de la qualité des données IRM, 142 participants sans mutation génétique ni forme familiale ont été retenus à l'inclusion (âge moyen  $59,33 \pm 4,00$  ans, niveau d'éducation moyen  $9,88 \pm 2,94$  ans ; 57 %).

À l'inclusion, tous les participants avaient un score MMSE. Les autres évaluations neuropsychologiques variaient, avec des données disponibles pour 136 participants pour le CDR-SoB et 97 participants pour le test de Beery. De plus, 124 participants disposaient d'analyses de

biomarqueurs du LCR et 132 participants avaient un génotype APOE ε4. Globalement, les participants présentaient une sévérité de maladie modérée et un déficit cognitif modéré (score médian CDR-SoB 6,00 ; MMSE moyen  $16,45 \pm 6,52$ ) avec un âge au début des symptômes de  $54,28 \pm 3,55$  ans et une durée de la maladie de  $5,11 \pm 2,78$  ans. Parmi les participants, 50 % présentaient un phénotype non mnésique.

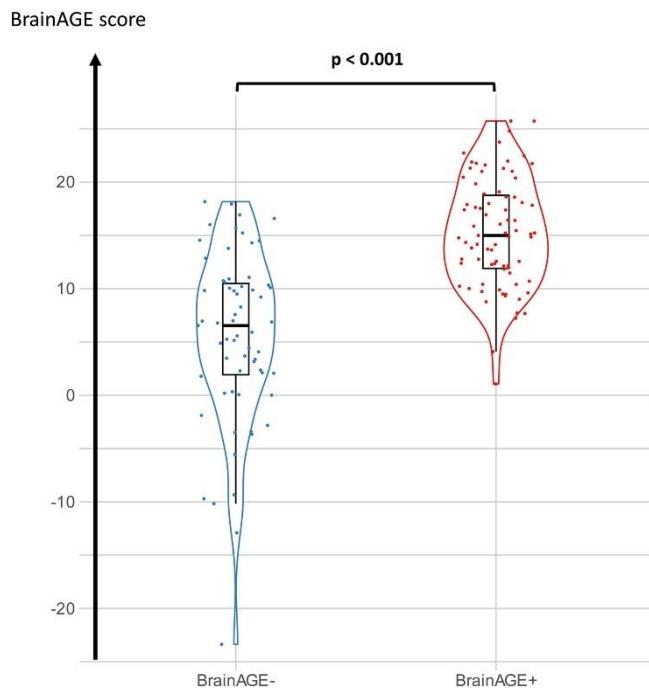
Le nombre de participants a diminué au fil du temps, atteignant 5 participants à l'année 6 (Figure II-3).



19. Figure II-3. Diagramme de flux de la population

### II.3.2. Clustering

Le clustering a été réalisé uniquement à l'inclusion, et les patients ont conservé leur affectation de cluster tout au long de l'analyse longitudinale. La solution optimale basée sur nos métriques variait entre deux et trois clusters (Figure II-S1). Cependant, la solution à trois clusters a donné un cluster ne contenant que trois patients. Nous avons donc opté pour la solution à deux clusters.



20. Figure II-4. Distribution des scores BrainAGE entre les clusters.

L'abscisse correspond au cluster associé à la distribution. L'ordonnée correspond au score BrainAGE (âge cérébral corrigé – âge chronologique).  $p\text{-value} < 0.001$ .

Pour assurer la pertinence des clusters et des représentations encodées par le modèle, nous avons comparé les scores BrainAGE des patients dans chaque cluster (Figure II-4). Le cluster #1 comprenait 63 sujets avec un âge cérébral prédict moyen de  $63,44 \pm 7,21$  ans, tandis que le cluster #2 comprenait 79 sujets avec un âge cérébral prédict moyen de  $75,66 \pm 4,38$  ans. Les scores BrainAGE différaient entre les clusters ( $5,44 \pm 8,13$  ans pour le cluster #1 et  $15,25 \pm 5,11$  ans pour le cluster #2,  $p < 0,001$ ) ; sur la base de ces résultats, nous avons désigné le cluster #1 comme BrainAGE- et le cluster #2 comme BrainAGE+.

### II.3.3. Différences démographiques, biologiques et cliniques au départ

Les différences démographiques, biologiques et cliniques entre clusters sont présentées dans le tableau II-2. Les clusters ne présentaient des différences que pour l'âge et le sexe, qui ont donc été inclus comme covariables dans les analyses statistiques suivantes. Les patients du cluster BrainAGE+ étaient plus âgés et plus susceptibles d'être de sexe masculin.

Variable	N	BrainAGE-	BrainAGE+	FDR corrected p-value	$f^2$ effect
<i>Démographie</i>					
Age, année	142	$57.99 \pm 4.33$	$60.42 \pm 3.36$	<b>0.001</b>	0.10
Education, année	129	$9.98 \pm 3.12$ (# de manquant, n=7)	$9.79 \pm 2.81$ (# de manquant, n=6)	0.72	0.001
Durée de la maladie,	139	$4.63 \pm 2.81$ (# de	$5.55 \pm 2.73$ (# de	0.07	0.03

		manquant, n=2)	manquant, n=1)		
		Odd Ratio			
Femme, n	142	45/63 (71%)	37/79 (46%)	<b>0.007</b>	2.83
<i>Biologie</i>				<i>f<sup>2</sup> effect</i>	
Total TAU, pg/mL	125	799.08 ± 451.67 (# de manquant, n=6)	724.71 ± 387.32 (# de manquant, n=11)	0.32	0.008
p-TAU / Aβ42	126	0.23 ± 0.14 (# de manquant, n = 7)	0.21 ± 0.14 (# de manquant, n = 9)	0.64	0.003
				Odd Ratio	
APOE ε4	138	34/60 (57%) (# de manquant, n=3)	39/78 (50%) (# de manquant, n=1)	0.49	0.77

2. Tableau II-2. Différences démographiques et biologiques à l'inclusion.

Les variables continues sont présentées sous forme de moyenne ± d'écart-type. Les variables catégorielles sont indiquées comme le nombre sur le nombre total (pourcentage). APOE ε4, au moins un allèle APOE ε4. Le nombre de valeurs manquantes est présenté sous la forme # de valeurs manquantes. Les valeurs p significatives corrigées du FDR sont en gras. Les valeurs p sont corrigées séparément entre les parties démographies et biologies.

### II.3.4. Différences neuropsychologiques et de volumes cérébraux à l'inclusion

Les différences neuropsychologiques, de volumes et d'épaisseur corticale sont présentées dans le tableau II-3.

Variable	BrainAGE-	BrainAGE+	FDR corrected p-value	<i>f<sup>2</sup> effect</i>
<i>Scores cognitifs</i>				
MMSE	19.32 ± 4.62	14.14 ± 6.93	<b>&lt; 0.001</b>	0.16
CDR-SoB (médiane -IQR)	5 - [3.5 - 6.62]	7 - [5 - 11]	<b>0.002</b>	0.09
VAT	3.52 ± 3.78	3.96 ± 3.84	0.35	0.01
DO80	71.02 ± 9.50	64.71 ± 14.66	<b>0.01</b>	0.08
Beery	19.06 ± 4.01	18.08 ± 4.20	0.27	0.03
Letter fluency	11.52 ± 7.92	8.23 ± 6.28	0.07	0.05
Categorical fluency	14.77 ± 6.45	11.28 ± 7.09	<b>0.004</b>	0.10
<i>Volumes cérébraux (pourcentage de volume)</i>				
Substance grise	0.35 ± 0.03	0.32 ± 0.02	<b>&lt; 0.001</b>	0.36
Substance blanche	0.28 ± 0.02	0.27 ± 0.03	0.96	0.004

Hippocampe	<b>4.77e-03 ± 6.65e-04</b>	<b>4.40e-03 ± 7.44e-04</b>	0.40	0.01
LCR	0.26 ± 0.03	0.30 ± 0.04	<b>&lt; 0.001</b>	0.21
<i>Phénotype cognitif</i>				<b>Odd Ratio</b>
Amnésique	33/60 (55%)	38/77 (49%)	0.51	0.80
<i>Phénotype d'atrophie</i>				
Typical AD, n	35/63 (55%)	35/79 (46%)		REF
Hippocampal sparing, n	14/63 (22%)	22/79 (27%)	0.76	1.95
Limbic predominant, n	14/63 (22%)	22/79 (27%)	0.95	0.95
<i>Épaisseur corticale totale de la surface (mm)</i>				<b>f<sup>2</sup> effect</b>
cortical AD signature	2.31 ± 0.11	2.20 ± 0.12	<b>&lt; 0.001</b>	0.20
cortical resilience signature	2.43 ± 0.13	2.29 ± 0.12	<b>&lt; 0.001</b>	0.23

3. Tableau II-3. Différences neuropsychologiques, volumétriques cérébrales, d'atrophie et de phénotypes cliniques à l'inclusion.

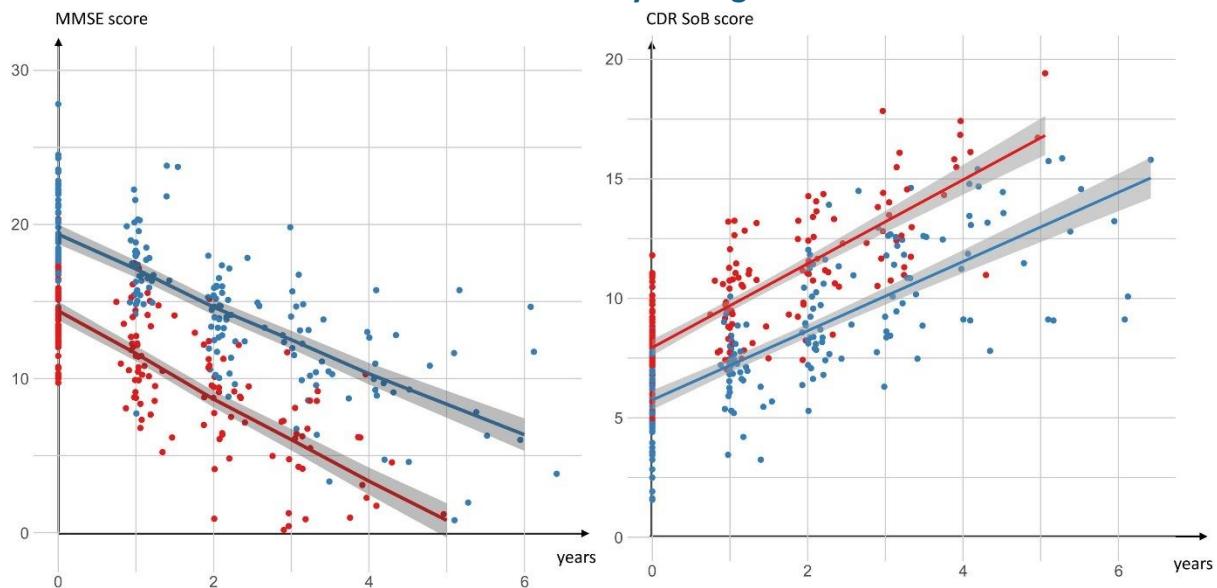
Les variables continues sont présentées sous forme de moyenne ± d'écart-type. Les variables catégorielles sont indiquées comme le nombre sur le nombre total (pourcentage). Les valeurs p significatives sont en gras. IQR : Écart interquartile.

Le cluster BrainAGE+ avait des scores MMSE plus faibles (p-value corrigée FDR < 0,001) et des scores CDR-SoB plus élevés (p-value corrigée FDR < 0,001) par rapport au cluster BrainAGE-.

Concernant les tests neuropsychologiques détaillés, le cluster BrainAGE+ présentait des performances inférieures dans les fonctions exécutives (test de fluence catégorielle, p-value corrigée FDR = 0,004) et le langage (test DO80, p-value corrigée FDR = 0,01). De plus, le cluster BrainAGE+ présentait une moindre quantité de substance grise et des volumes de LCR plus importants.

Le type de cluster n'était associé à aucun phénotype cognitif (mnésique vs non mnésique) ni à un pattern d'atrophie (typical AD vs hippocampal sparing vs limbic predominant). Cependant, une épaisseur corticale plus faible dans les zones correspondant à la signature corticale AD et à la signature de résilience cognitive a été observée dans le cluster BrainAGE+.

### II.3.5. Différences entre clusters dans l'analyse longitudinale



21. Figure II-5. Évolution cognitive par cluster.

Le cluster BrainAGE+ est représenté en rouge, le BrainAGE- en bleu. Les p-values pour la différence de pente sont respectivement 0.02 (MMSE) et 0.03 (CDR-SoB).

L'analyse longitudinale a été conduite sur tous les scores cognitifs et volumes cérébraux, ajustée pour le sexe et l'âge à l'inclusion. Les sujets du cluster BrainAGE+ ont présenté un déclin cognitif plus rapide sur le MMSE et une progression plus rapide du CDR-SoB (Figure II-5). Le volume de LCR a augmenté plus rapidement dans le cluster BrainAGE+, le volume de l'hippocampe a diminué plus rapidement dans le cluster BrainAGE-, tandis que les autres volumes cérébraux n'ont pas montré de différences dans les taux de progression (Tableau II-4).

	BrainAGE-	BrainAGE+	p-value
<b>Scores de sévérité</b>			
MMSE, pts / année	-2.35 +/- 0.15	-3.02 +/- 0.25	<b>0.02</b>
CDR-SoB, pts / année	1.58 +/- 0.10	1.99 +/- 0.16	<b>0.03</b>
<b>Volumes cérébraux</b>			
Substance grise, % / année	-8.37e-03 ± 5.68e-04	-6.52e-03 ± 9.33e-04	0.14
Substance blanche, % / année	-7.88e-03 ± 9.89e-04	-4.56e-03 ± 1.61e-03	0.16
Hippocampe, % / année	-1.76e-04 ± 8.18e-06	-1.42e-04 ± 1.35e-05	<b>0.03</b>
LCR, % / année	1.72e-02 ± 1.28e-03	1.16e-02 ± 2.09e-03	<b>0.02</b>

4. Tableau II-4. Taux de déclin entre les clusters en termes de sévérité et de volumes cérébraux.

Contrairement aux clusters BrainAGE, les groupes PAD et BR ou les clusters basés sur le vecteur créé à la main n'ont montré aucune différence dans la variation du MMSE ou du CDR-SoB (Tableaux II-S2, II-S3 et II-S4).

### II.4. Discussion

Dans cette étude, nous avons appliqué une approche de clustering sur la représentation encodée du BrainAGE dans une population atteinte d'EOAD. Deux clusters ont été identifiés et caractérisés par des différences significatives du BrainAGE. Les patients classés BrainAGE+ étaient plus âgés et plus

souvent de sexe masculin. Les évaluations neuropsychologiques ont révélé une moindre efficacité cognitive globale, une sévérité de la maladie plus élevée et une performance réduite aux tests de dénomination et de fluence verbale. L'analyse IRM a montré une diminution du volume hippocampique chez ces patients. Les analyses longitudinales ont démontré un déclin plus marqué dans le cluster BrainAGE+.

Des études antérieures appliquant le BrainAGE à la maladie d'Alzheimer à début tardif (LOAD) ont systématiquement montré un âge cérébral prédict plus élevé par rapport aux sujets sains (Franke et al., 2010). Chez les patients présentant un déficit cognitif léger (MCI), le BrainAGE s'est révélé être un prédicteur précis de la conversion vers la maladie d'Alzheimer, chaque année supplémentaire de BrainAGE étant corrélée à un risque accru de 10 % de développer la maladie sur une période de suivi de 36 mois (Gaser et al., 2013). Des analyses longitudinales ont également montré une accélération du BrainAGE chez les patients MCI converters, ainsi que chez les patients MA (Löwe et al., 2016). À ce jour, les analyses du BrainAGE chez les patients à début précoce restent limitées. Dans une étude précédente, nous avions observé une augmentation du BrainAGE dans l'EOAD et son accroissement significatif au cours du suivi longitudinal (Gautherot et al., 2021). Sa corrélation avec le déclin cognitif longitudinal n'avait pas été évaluée auparavant.

La nouveauté de la présente étude réside dans l'exploitation de la représentation encodée du BrainAGE, spécifiquement la sortie de la dernière couche flatten du modèle. Conformément aux recherches antérieures mettant en évidence l'utilité des représentations internes dans les modèles de BrainAGE pour les tâches de classification (Leonardsen et al., 2022), nos résultats montrent que le clustering basé sur la représentation encodée du BrainAGE est associé au déclin cognitif longitudinal. Comparées à d'autres approches – telles que le PAD, les représentations absolues ou le vecteur créé à la main – ces représentations dérivées du BrainAGE capturent plus efficacement l'hétérogénéité spatiale des patterns d'atrophie cérébrale.

Nos résultats fournissent des preuves supplémentaires en faveur de la pertinence biologique des méthodes BrainAGE. Comparée à la LOAD, l'EOAD se caractérise par un phénotype hétérogène avec une prévalence plus élevée de présentations atypiques. Selon Ferreira et al. (Ferreira et al., 2020), le spectre de la MA peut être décrit selon deux axes principaux : l'un relatif à la typicité et l'autre à la sévérité. Dans notre étude, la représentation encodée du BrainAGE était associée à des biomarqueurs multimodaux de sévérité, incluant l'atrophie corticale de la matière grise et l'altération cognitive globale. Cependant, les clusters BrainAGE ne différaient pas significativement en termes de présentation clinique (mnésique vs non-mnésique) ni de patterns d'atrophie IRM (hippocampal sparing vs limbic predominant vs typical AD).

Fait intéressant, les clusters différaient non seulement par leurs caractéristiques à l'inclusion, mais aussi par leur valeur prédictive du déclin cognitif sur une période de suivi de six ans. Par exemple, la diminution du MMSE et l'augmentation du CDR-SoB étaient respectivement de 0,7 point/an et 0,4 pts/an plus rapides dans le cluster BrainAGE+. En revanche, l'augmentation longitudinale du volume de LCR était plus rapide dans le cluster BrainAGE-. Cette apparente divergence peut s'expliquer par un effet plafond chez les patients BrainAGE+ présentant déjà une atrophie cérébrale sévère.

Le déclin cognitif dans la maladie d'Alzheimer est principalement influencé par la charge en amyloïde et tau (Cody et al., 2024; Zhou et al., 2024). Pour une quantité donnée de lésions neuropathologiques, la sévérité clinique est modulée par la réserve cognitive et cérébrale. Dans

cette étude, nous n'avons trouvé aucune association entre le niveau d'éducation, proxy de la réserve cognitive, et la représentation encodée du BrainAGE. Dans l'ensemble, nos résultats suggèrent que la représentation encodée du BrainAGE pourrait servir de biomarqueur de la réserve cérébrale. La réserve cérébrale fait référence à la capacité du cerveau à tolérer des dommages neuropathologiques tout en maintenant sa structure (Stern et al., 2023). Elle est influencée par divers facteurs, incluant la génétique, l'éducation, l'activité sociale, cognitive et physique, ainsi que les facteurs liés au mode de vie.

Conformément au modèle de différenciation renforcée proposé par Bocancea et al. (Bocancea et al., 2023), les différences à l'inclusion des performances cognitives entre clusters sont amplifiées au fil du temps. Cet effet protecteur dans le cluster BrainAGE- peut s'expliquer par deux mécanismes. Premièrement, les différences à l'inclusion de BrainAGE peuvent refléter une combinaison de facteurs génétiques et développementaux contribuant à une structure cérébrale plus résiliente. Deuxièmement, l'effet protecteur supplémentaire sur le taux de déclin cognitif durant le suivi longitudinal peut être porté par des mécanismes compensatoires plus actifs, incluant le recrutement accru de réseaux neuronaux alternatifs, la régulation à la hausse des facteurs neurotrophiques ou un remodelage synaptique plus efficace (Aron et al., 2022).

Cette hypothèse est renforcée par les associations à l'inclusion avec les facteurs démographiques (âge plus jeune, sexe féminin) et les résultats de l'imagerie cérébrale (signature corticale de « résilience cérébrale ») déjà liées à la réserve cérébrale (Ossenkoppele et al., 2020). L'effet protecteur du sexe féminin pourrait impliquer plusieurs mécanismes, incluant une atténuation des altérations liées à l'âge de l'expression génique énergétique, une activation renforcée du système immunitaire chez les femmes par rapport aux hommes, ainsi que des réponses inflammatoires spécifiques au sexe face aux lésions neuropathologiques dues à des déficiences en hormones stéroïdes sexuelles (Zhu et al., 2021). Un âge plus jeune a également été associé à une plus grande résilience cérébrale, attribuable à une moindre prévalence de comorbidités cérébrales telles que les maladies cérébrovasculaires et à des mécanismes de réparation neuronale plus efficaces (Kirkwood, 2005).

Ces résultats ont des implications pour l'application clinique du BrainAGE. Si le BrainAGE reflète de manière fiable la réserve cérébrale, il pourrait constituer un outil précieux pour la stratification des patients dans les essais cliniques, ainsi que pour la planification personnalisée du traitement et du suivi. L'identification précoce des patients ayant une réserve cérébrale faible pourrait permettre des interventions ciblées, ralentissant potentiellement le déclin cognitif. Ceci est particulièrement critique dans l'EOAD, où de nombreux patients sont encore actifs professionnellement, soulignant la nécessité de stratégies de prise en charge proactives et efficaces.

La représentation encodée de l'IRM par le BrainAGE présente des avantages distincts répondant aux principales limitations de la prédiction de maladie basée sur l'imagerie. Une force majeure réside dans son indépendance vis-à-vis de la maladie. Alors que les modèles deep learning ont montré leur utilité pour prédire la progression de la maladie à partir de données IRM et TEP (Bringas et al., 2020; Li et al., 2019; Singh et al., 2017), ils sont généralement entraînés pour des catégories diagnostiques spécifiques. Ce design spécifique limite leur généralisabilité et requiert des annotations expertes étendues pour assurer la qualité des labels, ce qui freine leur applicabilité à grande échelle et inter-maladies. En revanche, BrainAGE fournit un biomarqueur transdiagnostique et indépendant de la

maladie (Abeyasinghe et al., 2025; Eickhoff et al., 2021; Gaser et al., 2013) , adapté aux populations cliniques hétérogènes et aux études à large échelle sans nécessiter d'ajustement spécifique.

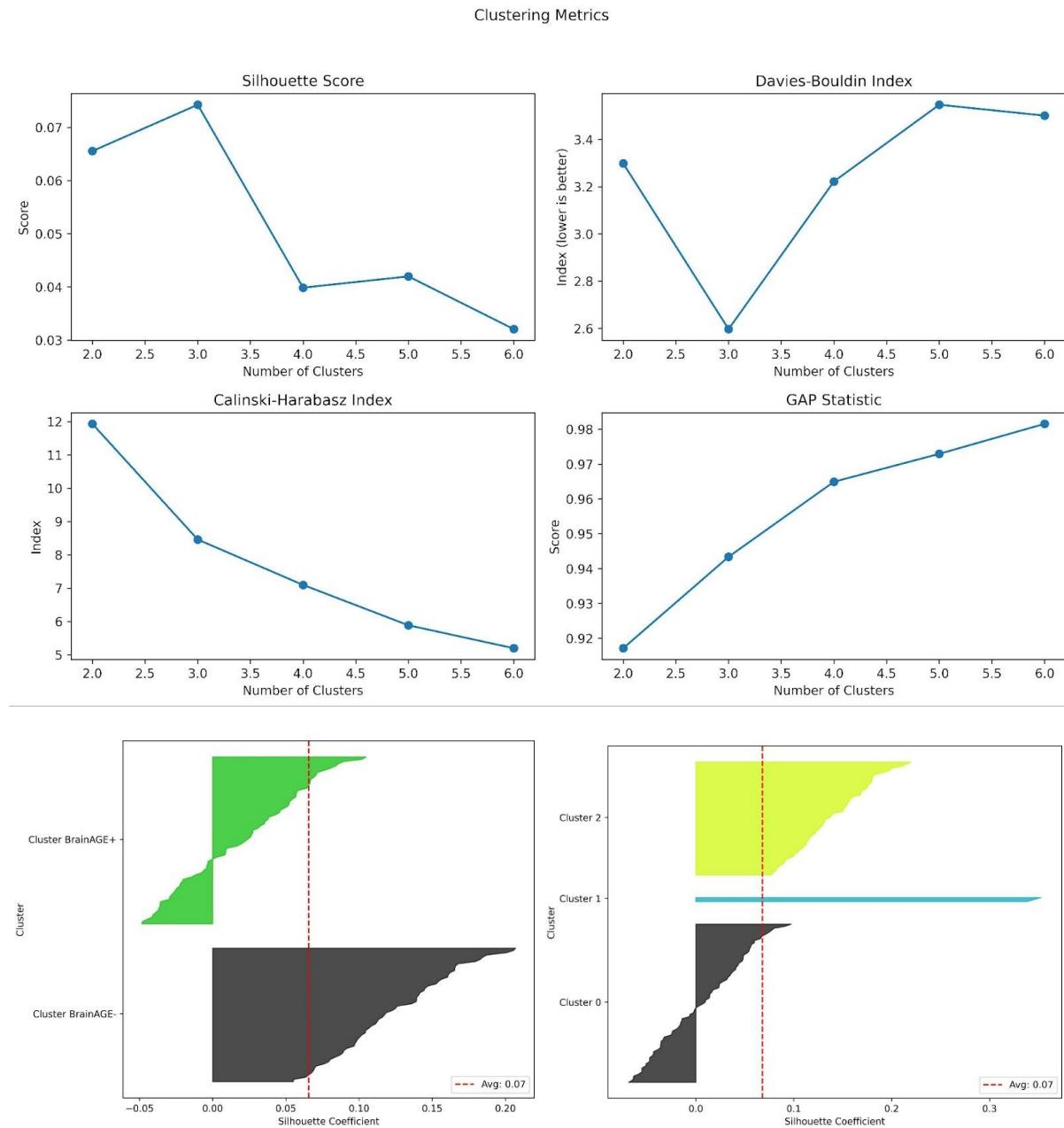
Une autre force majeure de BrainAGE réside dans sa capacité à extraire automatiquement des caractéristiques pertinentes et robustes à partir des données IRM. Contrairement aux approches traditionnelles basées sur des caractéristiques extraites à la main, qui peuvent passer à côté de patterns subtils mais cliniquement significatifs, BrainAGE exploite le deep learning pour capturer des marqueurs complexes et de haute dimension associés au vieillissement cérébral et à la sévérité de la maladie.

Notre étude présente certaines limites. Premièrement, étant une étude monocentrique, la reproductibilité de nos résultats dans d'autres centres reste à tester (Pruvo et al., 2025). La conduite d'études multicentriques renforcerait la robustesse de nos conclusions. Deuxièmement, notre population était principalement constituée de patients EOAD modérés à sévères. Il est donc essentiel d'étendre notre méthode aux patients EOAD à un stade précoce pour évaluer son applicabilité à la sélection de patients pour les essais cliniques. Enfin, bien que l'on puisse argumenter que nos clusters représentent des variations de stade de maladie plutôt que de réserve cérébrale, notre analyse longitudinale – ajustée sur les différences à l'inclusion – démontre que l'attribution au cluster influence indépendamment le déclin cognitif, indépendamment de la sévérité à l'inclusion. Enfin, notre modèle de deep learning BrainAGE repose uniquement sur des IRM-T1w, capturant l'information structurelle. L'exploration d'une approche neuroimagerie multimodale serait précieuse, différentes modalités pouvant fournir des informations complémentaires (Cole, 2020).

## II.5. Conclusion

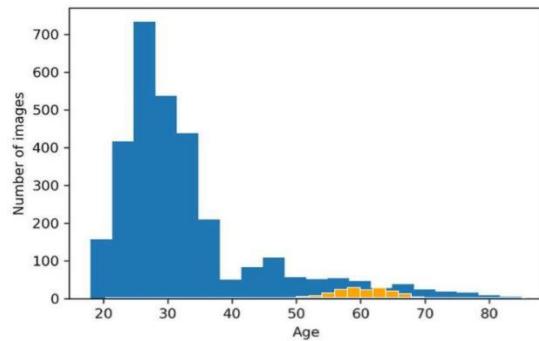
Dans cette étude, nous avons montré que la représentation encodée de BrainAGE peut distinguer des groupes de patients caractérisés par des taux distincts de déclin cognitif, indépendamment de leur phénotype. Ces résultats soulignent le potentiel d'utilisation de BrainAGE comme biomarqueur pour mieux comprendre et gérer l'EOAD.

## II.6 Annexes



22. Figure II-S1. Sélection du nombre optimal de clusters.

En haut : Évaluation de la qualité du cluster sur différents nombres de cluster à l'aide de quatre métriques de performance. En bas : les tracés de silhouette pour les solutions à deux et trois groupes sont illustrés. Pour les deux, un cluster reflète la continuité de la représentation encodée par IRM utilisée pour le clustering, comme l'indiquent les valeurs de silhouette négatives, suggérant que certains sujets se trouvent près de la limite entre les clusters.



23. Figure II-S2. Répartition par âge de l'ensemble d'entraînement du BrainAGE (bleu) et de la population EOAD (orange).

	BrainAGE entraînement	EOAD population
Age étendue	18 - 81	49 - 68
Femme, n (%)	1678 (52%)	81 (57%)
SCANNER, n (%)	SIEMENS 181 (6%)	0 (0%) 142 (100%)

5. Tableau II-S1. Caractéristiques de l'ensemble des d'entraînement du BrainAGE et de la population EOAD.

	BR-	BR+	FDR corrected p-value
<i>Scores de sévérité</i>			
MMSE, pts/années	-3.05 ± 0.24	-2.44 ± 0.28	0.11
CDR-SoB, pts/années	1.81 ± 0.15	1.69 ± 0.18	1.
<i>Volumes cérébraux</i>			
Matière grise, %/ans	-7.16e-03 ± 8.73e-04	-7.94e-03 ± 1.02e-03	1.
Matière blanche, %/années	-2.70e-03 ± 1.44e-04	-8.73e-03 ± 1.70e-03	<b>0.001</b>
Hippocampe, %/années	-1.56e-04 ± 1.28e-05	-1.66e-04 ± 1.49e-05	1.
LCR, %/années	9.41e-03 ± 1.88e-03	1.74e-02 ± 2.20e-03	<b>0.001</b>

6. Tableau II-S2. Taux de déclin entre les groupes basés sur le BR en fonction de la sévérité et du volume cérébral.

	PAD-	PAD+	FDR corrected p-value

---

### Scores de sévérité

MMSE, pts/années	$-2.44 \pm 0.16$	$-2.80 \pm 0.26$	0.50
CDR-SoB, pts/années	$1.53 \pm 0.1$	$2.03 \pm 0.16$	0.99

### Volumes cérébraux

Matière gris, %/ans	$-8.06e-03 \pm 5.73e-04$	$-6.92e-03 \pm 9.35e-03$	0.67
Matière blanche, %/années	$-7.39e-03 \pm 9.92e-04$	$-5.21e-03 \pm 1.61e-03$	0.71
Hippocampe, %/années	$-1.73e-04 \pm 8.26e-06$	$-1.46e-04 \pm 1.35e-05$	0.17
LCR, %/années	$1.64e-03 \pm 1.29e-03$	$1.26e-02 \pm 2.10e-03$	0.21

7. Tableau II-S3. Taux de déclin entre les groupes basés sur le PAD en fonction de la gravité et du volume cérébral.

	Cluster#1	Cluster#2	FDR corrected p-value
<b>Scores de sévérité</b>			
MMSE, pts/années	$-2.79 \pm 0.35$	$-2.55 \pm 0.13$	0.95
CDR-SoB, pts/années	$1.60 \pm 0.23$	$1.74 \pm 0.08$	1.

8. Tableau II-S4. Taux de déclin entre les clusters de vecteurs "handcrafted" en fonction de la sévérité.  
Les vecteurs sont composés de volumes normalisés en matière grise, de matière blanche normalisée, de LCR normalisée, d'hippocampe normalisée et d'âge.

# Chapitre III

## PatientSpace : un espace latent interprétable pour l'apprentissage de biomarqueurs de neuroimagerie dans la démence fronto-temporale

---

Dans le chapitre II, nous avons montré que les espaces latents apprennent des structures inhérentes aux données, comme en témoigne la capacité de BrainAGE à différencier des sous-groupes de patients dans une population EOAD. Dans le chapitre III, nous présentons le PatientSpace, un espace latent structuré et interprétable construit à partir d'IRM T1w et de données cliniques. Cet espace organise les sujets de manière à séparer les patients atteints de démences fronto-temporales (DFT) des témoins cognitivement normaux (CN), tout en rapprochant les individus présentant des profils neuroanatomiques similaires, offrant ainsi une représentation cliniquement informative et interprétable.

### **III.1. Introduction**

Les démences fronto-temporales (DFT) constituent un groupe hétérogène de maladies neurodégénératives rares, caractérisées par une atteinte progressive des lobes frontaux et temporaux. Trois phénotypes cliniques majeurs sont classiquement décrits : le variant comportemental (bvFTD), l'aphasie progressive non fluente (PNFA) et le variant sémantique (SV) (Ljubenkov and Miller, 2016).

Le bvFTD, le plus fréquent, représente environ 50 % des cas et se manifeste principalement par des perturbations comportementales de type dysexécutif incluant désinhibition, apathie, inertie et perte d'empathie (Rascovsky et al., 2011). Le PNFA est marqué par des troubles praxiques de la parole, un agrammatisme, une expression non fluente et des difficultés syntaxiques, tandis que le SV se traduit par un discours fluent mais vidé de sens, associé à une perte progressive des connaissances lexicales et conceptuelles (Gorno-Tempini et al., 2004).

Malgré des critères diagnostiques établis, la distinction entre sous-types reste difficile en pratique clinique, en raison du chevauchement progressif des symptômes au cours de l'évolution de la maladie. Ainsi, un patient bvFTD peut présenter une diminution de la fluence verbale, conduisant à une confusion avec un PNFA. De plus, des symptômes similaires s'observent dans d'autres pathologies neurodégénératives telles que la paralysie supra-nucléaire progressive, les syndromes cortico-basal ou encore la maladie d'Alzheimer (Battista and Gallucci, 2017; Ljubenkov and Miller, 2016).

L'imagerie cérébrale constitue un outil central pour le diagnostic des DFT. L'imagerie par résonance magnétique (IRM) permet d'associer des profils d'atrophie cérébrale aux présentations cliniques : le bvFTD se caractérise par une atrophie fronto-temporale diffuse, le SV par une atrophie asymétrique du lobe temporal antérieur, et le PNFA par une atrophie frontale inférieure gauche avec atteinte de l'insula antérieure (Peet et al., 2021). Cependant, la variabilité entre centres et observateurs demeure un obstacle majeur. Une étude pionnière a mis en évidence un faible accord inter-observateurs ( $\kappa$  moyen = 0,34) lors de l'évaluation qualitative de l'atrophie à l'IRM (Scheltens et al., 1997). Dans ce contexte, des méthodes d'analyse automatisées apparaissent comme une approche prometteuse pour améliorer la reproductibilité et la sensibilité du diagnostic.

Les méthodes d'apprentissage automatique, et plus particulièrement l'apprentissage profond, ont montré un fort potentiel pour la classification des maladies neurodégénératives, y compris la DFT (Kim et al., 2019; Metz et al., 2025). Toutefois, la majorité des modèles reposent sur des architectures complexes de type réseaux de neurones profonds, souvent qualifiées de « boîtes noires », difficiles à interpréter et donc limitées pour un usage clinique (Holzinger et al., 2019). Pour pallier cette limite, de nouvelles approches visent à construire des espaces latents interprétables, capables de représenter l'hétérogénéité des patients de manière intelligible. Cette stratégie a récemment été appliquée à la maladie d'Alzheimer, avec la création d'espaces latents permettant d'expliquer la variabilité des profils d'atrophie cérébrale (Kang et al., 2024).

Dans cette étude, nous présentons le PatientSpace, un espace latent structuré et interprétable, construit à partir d'IRM T1w et de données cliniques de patients atteints ou non de DFT. Cet espace organise les sujets selon deux axes principaux : le premier distingue les patients DFT des témoins cognitivement normaux (CN), tandis que le second rapproche les individus présentant des profils neuroanatomiques similaires.

## **III.2. Matériels et méthodes**

### **III.2.1. Base de données**

Les données analysées proviennent de la cohorte Frontotemporal Lobar Degeneration Neuroimaging Initiative (FLTDNI, <http://memory.uscf.edu/research/studies/nifd>). Cette étude longitudinale recueille des données cliniques, biologiques et d'imagerie cérébrale de patients atteints de DFT ainsi que des CN. Dans ce travail, nous avons utilisé les IRM pondérées en T1, ainsi que les variables démographiques associées (âge, sexe, niveau d'éducation). Les scores cognitifs et cliniques étaient également disponibles, incluant le clinical dementia rating (CDR), et le CDR sum of boxes (CDR-SB). Les critères d'inclusion reposaient sur les recommandations établies par Rascovsky et al. (Rascovsky et al., 2011) pour le variant comportemental, et par Gorno-Tempini et al. (Gorno-Tempini et al., 2004) pour les formes aphasiques. Les témoins sains ne présentaient ni lésions cérébrales ni antécédents de maladies neurodégénératives.

### **III.2.2. Prétraitement des images**

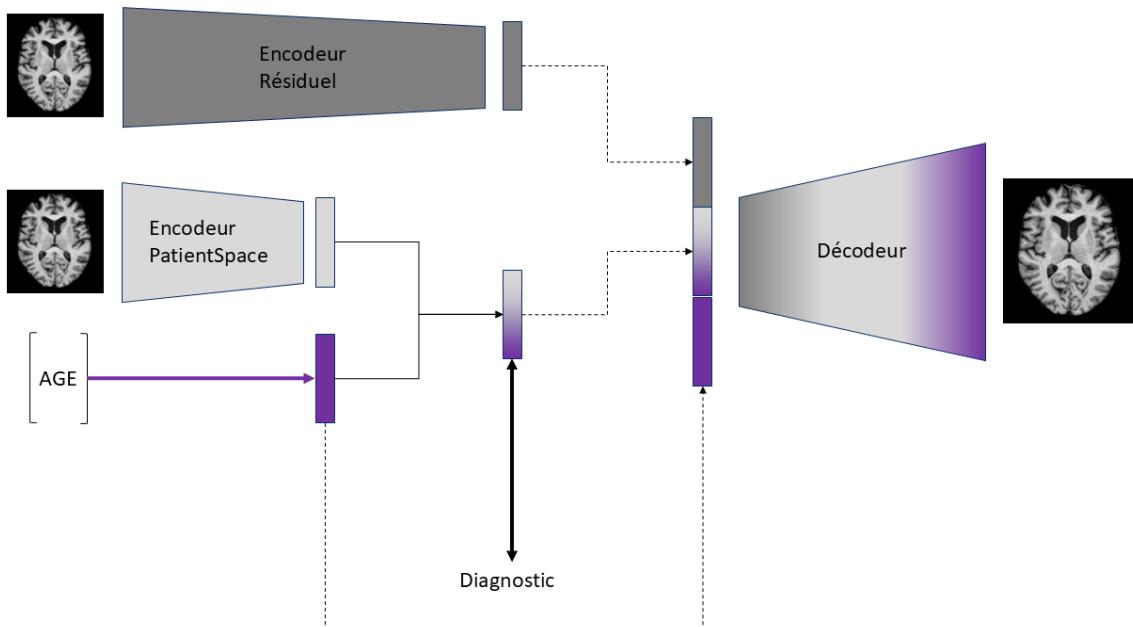
Les images ont été prétraitées suivant le pipeline IGUANE (Roca et al., 2025). Le prétraitement comprenait (1) extraction du cerveau grâce à HD-BET (Isensee et al., 2019); (2) une correction d'inhomogénéité de champ avec l'algorithme N4ITK (Tustison et al., 2010); (3) un recalage dans l'espace MNI de 1mm<sup>3</sup> avec FSL-FLIRT (Jenkinson et al., 2002); (4) un recadrage de taille 160 x 192 x 160 voxels; (5) une normalisation d'intensité par IGUANE sur la base de référence SALD (Wei et al., 2018) afin de corriger les effets sites puis (6) une normalisation z-score intra crânienne. Un contrôle qualité a été effectué sur l'ensemble des données afin d'exclure les images de mauvaise qualité.

### **III.2.3. Dataset**

Afin d'entraîner et d'évaluer notre modèle, les données ont été divisées en trois sous-ensembles : un ensemble d'entraînement (60 %), un ensemble de validation (20 %) et un ensemble de test (20 %). Pour garantir l'indépendance des ensembles, les IRM issues d'un même patient ont toujours été regroupées dans un même sous-ensemble. Une stratification a été appliquée afin d'assurer un équilibre entre les groupes cliniques, les tranches d'âge et les scores cognitifs (CDR-SB). Cette stratégie visait à garantir la représentativité des sous-ensembles et à limiter les biais lors de l'entraînement.

### **III.2.4. Modèle de deep learning**

Une visualisation globale du modèle est illustrée Figure III-1.



24. Figure III-1. Architecture du modèle.

Le modèle se compose en 2 encodeurs : un unimodal prenant en entrée seulement l'image et un encodeur prenant l'âge et l'image en entrée. L'espace latent associé à l'encodeur résiduel n'est utilisé que pour la reconstruction. L'espace latent dérivé de l'âge et de l'IRM est utilisé à la fois pour la reconstruction et la prédiction du diagnostic. Enfin le vecteur associé à l'âge est utilisé pour générer l'espace latent du PatientSpace et la reconstruction de l'image.

### Architecture

Le modèle proposé repose sur une architecture inspirée du Domain Invariant Variational autoencoder (DIVA) (Ilse et al., 2020), enrichie par une régularisation consistante inspirée des travaux de Sinha et Dieng (Sinha and Dieng, 2022). L'objectif était d'apprendre un espace latent structuré par le diagnostic et contraint par un critère de similarité, de façon à capturer à la fois la séparation entre patients sains et DFT, et la proximité morphologique entre profils neuroanatomiques similaires. Le modèle intégrait deux encodeurs : un encodeur combinant l'IRM et l'âge conçu pour capturer les caractéristiques directement liées à la maladie, et un encodeur unimodal basé uniquement sur l'IRM destiné à extraire les caractéristiques morphologiques indépendantes de la maladie. Chaque encodeur unimodal adoptait la même architecture, composée de blocs résiduels 3D (He et al., 2016) intégrant des convolutions 3D, une normalisation de groupe (GN) (Wu and He, 2018)— choisie pour ses performances optimales avec de petits lots — et une fonction d'activation ReLU. Le sous-échantillonnage reposait sur des convolutions à pas 2, doublant le nombre de filtres, suivies d'une activation ReLU. La sortie de chaque encodeur était flatten puis transformée par des couches linéaires pour produire les représentations latentes, conformément à Kingma and Welling (Kingma and Welling, 2019).

Le décodeur reprenait la structure des encodeurs : un suréchantillonnage trilinéaire 3D doublait d'abord les dimensions spatiales, suivi de convolutions 3D, puis d'un bloc résiduel 3D. La dernière couche de chaque décodeur appliquait une convolution 3D avec un noyau de  $1 \times 1 \times 1$ , tandis que toutes les autres convolutions utilisaient des noyaux de  $3 \times 3 \times 3$ .

### Fonction de perte.

L'entraînement reposait sur une fonction de perte multitâche, combinant reconstruction, classification et régularisation. La perte globale était définie comme suit (Eq III-1) :

$$L = \lambda_{VAE} L_{VAE} + \lambda_{CE} L_{CE} + \lambda_{SIM} \beta_c L_{SIM} \quad (\text{Eq III-1})$$

Les différentes fonctions de perte étaient spécifiées de la manière suivante.

La perte du VAE reprenait la formulation de la  $\beta$ -VAE introduite par Ilse et al (Ilse et al., 2020) :

$$\begin{aligned} L_{VAE} = & (IRM - \widehat{IRM})^2 + \beta_c D_{KL}(Z_{IRM} || N(0, I)) \\ & + \beta_c D_{KL}(Z_{AGE+IRM} || p(Z_{AGE+IRM} | diagnostic)) \end{aligned} \quad (\text{Eq. III-2})$$

Ici,  $IRM$  désignait l'image initiale et  $\widehat{IRM}$  l'image reconstruite,  $N(0, I)$  correspondait à une distribution gaussienne multivariée centrée et de covariance identité.  $p(Z_{AGE+IRM} | diagnostic)$  représentait le prior conditionnel appris en fonction du diagnostic. Le facteur  $\beta_c$  pondérait la divergence de KL selon une règle cyclique (Fu et al., 2019).

La perte de classification reposait sur l'entropie croisée. Enfin, la perte de régularisation minimisait la similarité entre la représentation multimodale de l'image originale  $Z_{ORIG}$  et celle de la version transformée  $Z_{AUG}$ , tout en maximisant la dissimilarité avec des représentations issues d'images différentes  $Z_{DIFF}$ . Inspirée de Yeh et al (Yeh et al., 2022), cette perte était implémentée comme suit :

$$L_{SIM} = -\frac{S(Z_{ORIG}, Z_{AUG})}{\tau} + \log \sum_{Z_{DIFF}} \exp\left(\frac{S(Z_{ORIG}, Z_{DIFF})}{\tau}\right) \quad (\text{Eq. III-3})$$

Avec  $S(a, b) = -\frac{1}{2}(D_{KL}(a || b) + D_{KL}(b || a))$  la mesure de similarité entre a et b, et  $D_{KL}$  la divergence de KullBack-Leibler entre a et b.

### Procédure d'entraînement.

La procédure d'entraînement s'inspirait de celle décrite par Sinha and Dieng (Sinha and Dieng, 2022), appliquant la fonction de perte du VAE aussi bien aux images originales qu'aux images transformées. Pour réduire le risque de surapprentissage, la classification n'était effectuée que sur les données transformées. Afin d'obtenir une représentation latente plus représentative, une méthode cyclique était utilisée pour pondérer simultanément la perte de KL et la perte de similarité. Les poids de ces pertes augmentaient progressivement de 0 à 1 sur 13 époques, puis restaient constants jusqu'à la 25<sup>e</sup> époque, avant d'être réinitialisés à 0, le cycle se répétant ensuite. (Figure III-S1).

Les méthodes d'augmentation de données reposaient sur la bibliothèque MONAI (Cardoso et al., 2022) et incluaient : une translation aléatoire ( $\pm 10$  voxels sur chaque axe), une rotation aléatoire ( $\pm 0,4$  radians) et un zoom aléatoire (facteur d'échelle compris entre 0,9 et 1). L'augmentation par retournement était exclue afin de préserver les asymétries potentielles liées aux phénotypes de la maladie. De plus, un bruit gaussien était ajouté aux entrées, qui étaient ensuite débruitées lors de la reconstruction, conformément à des travaux ayant montré que cette stratégie améliore la robustesse des caractéristiques extraites (Vincent et al., 2008).

### III.2.5. Représentation par graphe de l'espace latent – PatientSpace

À partir des représentations latentes issues de l'encodeur prenant en entrée l'imagerie et l'âge, nous avons construit une représentation sous forme de graphe appelé PatientSpace, définie comme  $G =$

$(V, E, A)$  où  $V$  correspond à l'ensemble des nœuds représentant les patients,  $E$  à l'ensemble des arêtes connectant ces nœuds, et  $A$  à la matrice d'adjacence pondérée. Chaque nœud  $v_i \in V$  correspond à un patient, et la pondération des arêtes encode la similarité morphologique entre patients dans l'espace latent.

La mesure de similarité reposait sur la mesure de dissimilarité utilisée dans la perte de consistance décrite dans la section précédente ( $L_{SIM}$ ). Pour chaque paire de sujets  $(i, j)$ , la pondération  $a_{ij}$  de la matrice d'adjacence était définie de la manière suivante :

$$a_{ij} = \begin{cases} \frac{1}{d(i,j)} / \sum_{k \in N(i)} \frac{1}{d(i,k)}, & \text{si } j \text{ est un voisin de } i \\ 0, & \text{sinon} \end{cases} \quad (\text{Eq III-4})$$

Où  $d(i,j)$  représente la mesure de dissimilarité entre  $v_i$  et  $v_j$  définie comme

$$d(i,j) = \frac{1}{2}(D_{KL}(v_i \parallel v_j) + D_{KL}(v_j \parallel v_i)) \quad (\text{Eq.III.5})$$

Et  $N(i)$  représente les voisins de  $v_i$ . Le nombre K optimal est détaillé section III.6. S2.2.

## III.2.6. Expérimentations

### III.2.6.1. Analyse du graphe

Deux analyses ont été menées pour étudier les propriétés structurelles du PatientSpace. La première s'est concentrée sur la topologie globale en identifiant les clusters dans le graphe. La seconde a mis l'accent sur les connaissances au niveau local en effectuant une analyse de voisinage au niveau du patient individuel.

**Clustering du PatientSpace.** Le PatientSpace a été regroupé à l'aide d'un clustering agglomératif utilisant le critère de Ward, en préservant la structure du graphique. Chaque regroupement a ensuite été décrit en fonction du statut diagnostique (DFT ou CN), de la distribution des âges, du sexe, du niveau d'éducation ainsi que du score cognitif CDR-SB corrigé pour ces covariables. Les comparaisons statistiques s'appuyaient sur un test du Khi-deux pour les variables catégorielles et un test de Kruskal-Wallis pour les variables continues, avec une correction pour test multiple par la méthode du False discovery rate proposée par Benjamini et Hochberg (BH-FDR) (Benjamini and Hochberg, 1995). Afin de relier ces regroupements aux substrats neuroanatomiques, une analyse morphométrique basée sur la Voxel-Based Morphometry (VBM) a été réalisée, permettant d'identifier les patterns d'atrophie significatifs associés à chaque cluster (section III.6. S3).

**PatientSpace au niveau individuel.** Le PatientSpace a aussi été examiné au niveau du patient en prédisant le diagnostic, le phénotype ainsi que les CDR-SB dans la base de test. Pour tous les patients non vus, la probabilité d'être assigné au diagnostique  $D$  était estimé par :

$$\operatorname{argmax}_D P(\widehat{DX}_i = D) = \sum_j a_{ij} \mathbf{1}(DX_j = D) \quad (\text{Eq. III- 6})$$

Où  $DX_j$  est le diagnostique du sujet  $v_j$ .

Le CDR-SB a été estimé de la même manière que le diagnostic. Les scores CDR-SB ont ensuite été catégorisés selon le degré de démence comme suit : pas de démence (CDR-SB = 0) ; légère démence (CDR-SB entre 0.5 et 4) ; démence moyenne (CDR-SB entre [4.5 – 9]) ; démence modérée (CDR-SB entre [9.5 – 15.5]) et démence sévère (CDR-SB entre [16 – 18]) (O'Bryant et al., 2008a).

Pour tous les patients non vus, la probabilité d'être assigné au CDR-SB  $c$  a été estimée par :

$$\operatorname{argmax}_c P(\widehat{\text{CDRSB}}_i = c) = \sum_j a_{ij} \mathbf{1}(\text{CDRSB}_j = c) \quad (\text{Eq. III. 7})$$

Les résultats obtenus à partir des approches basées sur les graphes sont désignés sous le nom PatientGraph et ceux obtenus par le classifieur natif du réseau sous le nom DL-classifier dans les sections suivantes.

### III.2.6.2. Classification

Enfin, la robustesse du modèle a été évaluée en comparant les deux approches de classification : d'une part le classifieur de deep learning utilisant uniquement la sortie du réseau, et d'autre part l'approche PatientGraph, qui exploite les similarités locales dans l'espace latent. Les performances ont été mesurées par l'aire sous la courbe ROC (AUC), la sensibilité et la spécificité, avec des intervalles de confiance à 95 %. Les analyses ont été réalisées sur l'ensemble de test interne, distinct des données d'entraînement (désigné comme INTERNE), ainsi que sur une base externe de patients DFT provenant du Centre Mémoire de Lille et du National Alzheimer's Coordinating Center (NACC, <https://naccdata.org/>), désignée comme EXTERNE. Une description détaillée de ces populations est fournie en section III.6. S1.

## III.3. Résultats

### III.3.1. Démographie

La démographie des ensembles d'entraînement, de validation et de test est résumée dans le Tableau III-1. L'âge moyen différait significativement entre les groupes ( $p < 0,0001$ ), les sujets BV étant plus jeunes que les CN, PNFA et SV. Le niveau d'éducation était plus élevé chez les sujets CN que chez les patients ( $p = 0,0002$ ). Les scores CDR-SB reflétaient une gradation claire de sévérité cognitive : CN = 0, PNFA = 1,5, SV = 4,3 et BV = 6,5.

Ensemble d'entraînement					
	CN	BV	PNFA	SV	Total
<b>Nombre d'images</b>	203	126	69	80	536
<b>Site</b>					
<b>UCSF</b>	100 %	66,7 %	92,8 %	90 %	83,8 %
<b>MGH</b>	0 %	7,9 %	7,2 %	7,5 %	4,1 %
<b>MAYO</b>	0 %	25,4 %	0	2,5 %	12,1 %
<b>Age, années</b>	65,5 (7,9)	62,7 (5,9)	67,2 (7,2)	65,6 (6,1)	<b><math>p &lt; 0,0001</math> PNFA = CON = SV &gt; BV</b>
<b>Sexe féminin</b>	53,7 %	30,2 %	56,5 %	35 %	
<b>Éducation</b>	17,8 (1,8)	16,7 (2,7)	16,7 (2,7)	16,4 (2,6)	<b><math>p = 0,0002</math> CON &gt; BV = PNFA = SV</b>
<b>CDR-SB</b>	0 (0–0)	6,5 (4,5–10)	1,5 (0,5–3)	4,3 (2,5–7)	<b><math>p &lt; 0,0001</math> CON &lt; PNFA &lt; SV &lt; BV</b>
Ensemble de validation					
<b>Nombre d'images</b>	61 (41%)	37 (25%)	25 (17%)	25 (17%)	148
<b>Site</b>					
<b>UCSF</b>	100 %	70,3 %	100 %	88 %	90,5 %
<b>MGH</b>	0 %	0 %	0 %	12 %	2 %
<b>MAYO</b>	0 %	29,7 %	0 %	0 %	7,5 %
<b>Age, années</b>	65,25 (8,4)	61,9 (6,6)	71,6 (7,3)	62,3 (6,5)	<b><math>p &lt; 0,0001</math> PNFA &gt; CON &gt; BV; PNFA &gt; SV</b>
<b>Sexe féminin</b>	68,8 %	40,5 %	52 %	36 %	
<b>Éducation</b>	17,5 (1,4)	13,78 (2,0)	15,36 (2,5)	17,9 (1,8)	<b><math>p &lt; 0,0001</math> CON &gt; PNFA &gt; BV; SV &gt; PNFA &gt; BV</b>

<b>CDR-SB</b>	0 (0–0)	6 (5–10,3)	1,5 (0,5–3)	4 (2,6–5,7)	<b>p &lt; 0,0001 BV &gt; PNFA &gt; CON; SV &gt; CON</b>
<b>Ensemble de test (INTERNE)</b>					
<b>Nombre d'images</b>	75	40	18	23	156
<b>Site</b>					
<b>UCSF</b>	100 %	85 %	83,3 %	82,6 %	91,7 %
<b>MGH</b>	0 %	0 %	16,7 %	17,4 %	4,5 %
<b>MAYO</b>	0 %	15 %	0 %	0 %	3,8 %
<b>Age, années</b>	65,7 (7,1)	63,0 (5,9)	73,8 (5,6)	61,44 (5,3)	<b>p &lt; 0,0001 CON &lt; PNFA</b>
<b>Sexe féminin</b>	65,3 %	27,5 %	44,4 %	47,8 %	
<b>Éducation</b>	16,7 (1,9)	14,9 (2,7)	15,1 (2,5)	15,6 (3,1)	<b>p = 0,0006 CON &gt; BV</b>
<b>CDR-SB</b>	0 (0–0)	8 (5,5–12)	4,5 (3–8,5)	4,5 (3,1–6,3)	<b>p &lt; 0,0001 CON &lt; SV &lt; BV; CON &lt; PNFA</b>

9. Tableau III-1. Démographie des bases entraînement, validation et test.

### III.3.2. Interprétation du PatientSpace

#### III.3.2.1. PatientSpace clusters

La Figure III-2 illustre le PatientSpace, composé de six clusters : deux regroupant des sujets CN et quatre regroupant des patients atteints de démence. Les caractéristiques démographiques et cliniques de ces clusters sont résumées dans le Tableau III-2.

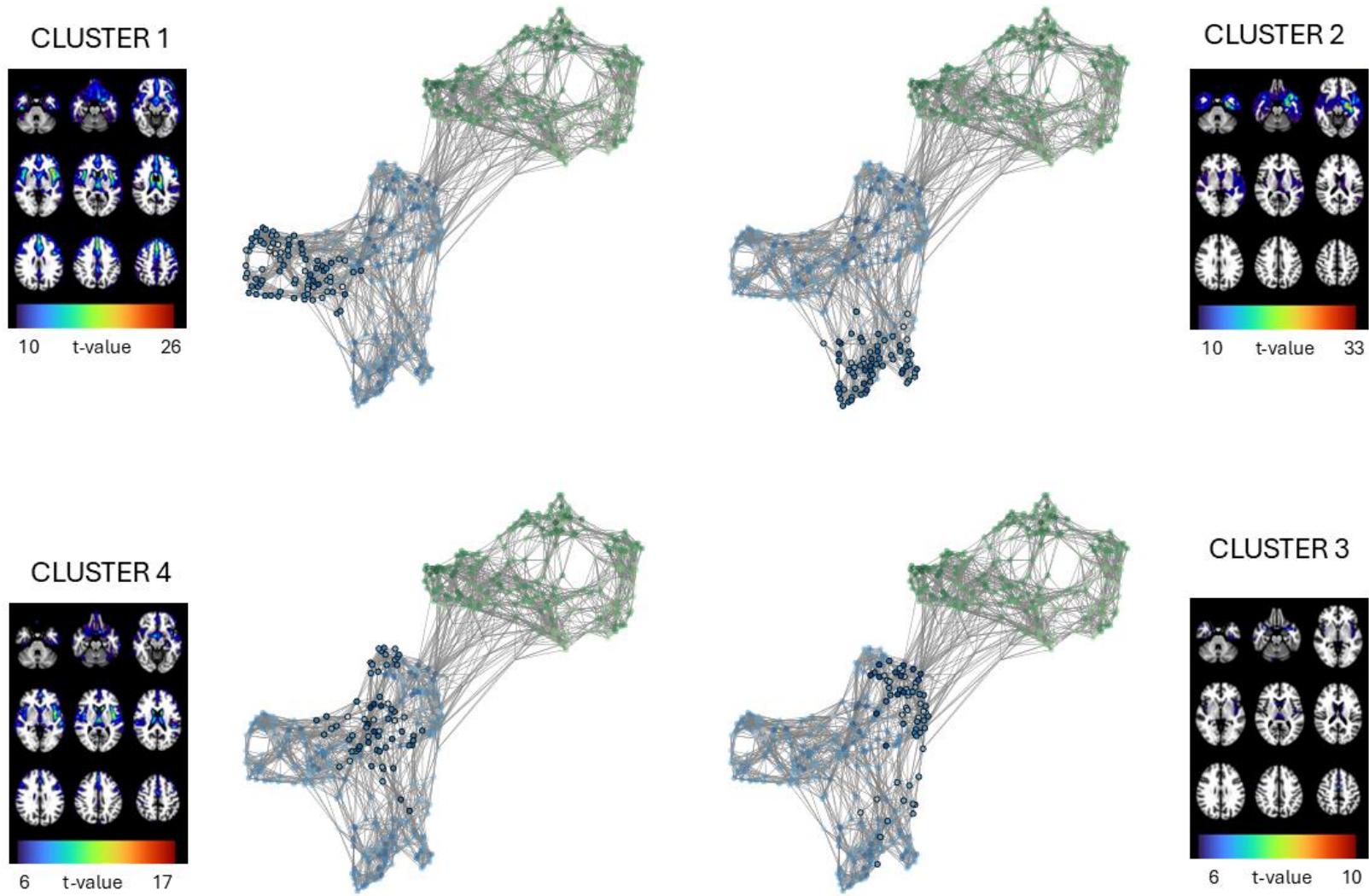
Le cluster 1 rassemblait principalement des patients bvFTD (88,9 %), tandis que le cluster 2 était constitué en grande majorité de patients SV (85,9 %). Les clusters 3 et 4, en revanche, présentaient des profils plus hétérogènes, incluant des patients bvFTD, PNFA et une minorité de SV. Les sujets du cluster 2 étaient significativement plus âgés (âge moyen : 66,1 ans) que ceux du cluster 3 (63,9 ans ; p = 0,04) et du cluster 1 (63,5 ans ; p = 0,005). Concernant la sévérité clinique, le cluster 1 affichait un score médian de CDR-SB de 6,8, significativement plus élevé (p < 0,001) que ceux du cluster 2 (4,5), du cluster 3 (3,0) et du cluster 4 (3,5). Le score du cluster 2 restait toutefois significativement plus élevé que celui du cluster 3 (p = 0,015). Aucune différence significative n'a été observée concernant le niveau d'éducation entre les clusters.

L'analyse des données de neuroimagerie a permis de dégager quatre profils distincts. Le cluster 1 présentait une atrophie diffuse, particulièrement marquée au niveau des cornes frontales et de la partie centrale des ventricules latéraux. Les gyri insulaires étaient fortement atteints, avec une atteinte plus prononcée à gauche. Une atrophie frontale asymétrique, prédominante à gauche, concernait également les gyri frontaux inférieurs, moyens et médians. Le cluster 2 montrait un profil caractéristique des SV, avec une atrophie localisée et asymétrique, prédominant sur l'hémisphère gauche. L'atteinte concernait principalement le lobe temporal, avec une atteinte particulièrement marquée de l'uncus, du gyrus parahippocampique et du gyrus fusiforme gauche. Le gyrus temporal supérieur gauche et l'insula gauche étaient également sévèrement atteints, tandis que le lobe temporal droit et les ventricules latéraux présentaient une atteinte plus modérée. Le cluster 3 se distinguait par une atrophie globalement faible en comparaison avec les autres clusters, localisée principalement à l'insula gauche et à la portion antérieure du gyrus temporal supérieur. Une dilatation symétrique des cornes frontales des ventricules latéraux était également observée. Enfin, le cluster 4 se caractérisait par une atrophie très diffuse mais moins sévère que celle observée dans les clusters 1 et 2, avec une atteinte plus marquée au niveau des cornes frontales et de la partie centrale des ventricules latéraux, associée à une atteinte modérée des lobes frontaux et temporaux.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total
Effectifs	81	64	63	65	273
<b>Phénotypes</b>					
BV	72 (88,9 %)	4 (6,3 %)	24 (38,1 %)	26 (40 %)	126 (46,1 %)
PNFA	9 (11,1 %)	5 (7,8 %)	28 (44,4 %)	27 (41,5 %)	69 (25,3 %)
SV	0 (0 %)	55 (85,9 %)	11 (17,5 %)	12 (18,5 %)	78 (28,6 %)
<b>Données cliniques</b>					
Age [moyenne (écart-type)]	63,5 (6,1)	66,1 (5,9)	63,9 (7,6)	65,4 (6,5)	<b>p = 0,02</b> <b>C3 &lt; C2</b> <b>C1 &lt; C2</b>
Genre féminin	30,9 %	42,2 %	55,5 %	27,7 %	
Éducation [moyenne (écart-type)]	16,8 (2,7)	16,5 (2,7)	17,1 (2,3)	16,0 (2,9)	<b>p = 0,18</b>
CDR-SB [médiane (quartiles)]	6,8 (4,9 – 10,6)	4,5 (2 – 7)	3 (1,1 – 4,9)	3,5 (1,5 – 7)	<b>p &lt; 0,0001</b> <b>C3 &lt; C2 &lt; C1</b> <b>C4 &lt; C1</b>

**10. Tableau III-2. Caractéristiques des clusters**

L'âge et le niveau d'éducation sont exprimés en moyenne (écart-type), la proportion de femmes est indiquée en pourcentage, et le CDR-SB est présenté en médiane avec l'intervalle interquartile (Q1-Q3).



25. Figure III-2. Visualisation du PatientSpace et des clusters.

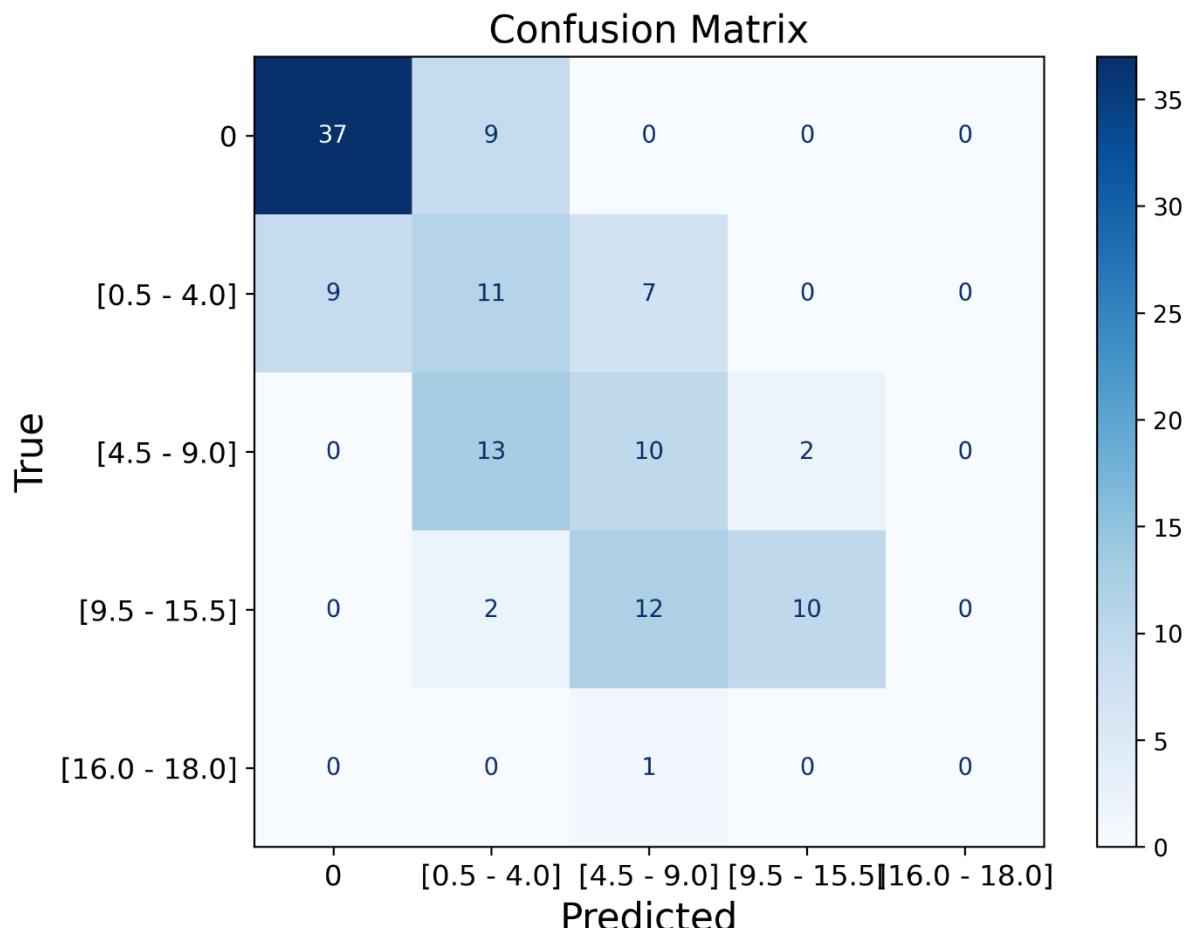
### III.3.2.2. Caractérisation individuelle et classification

		AUC	Sens	Spec
INTERNE	DL-classifier	0.931 [0.890 - 0.972]	0.889 [0.820 - 0.957]	0.920 [0.859 - 0.981]
	PatientGraph	0.905 [0.856 - 0.953]	0.889 [0.820 - 0.957]	0.840 [0.757 - 0.923]
EXTERNE	DL-classifier	0.892 [0.848 - 0.935]	0.705 [0.621 - 0.790]	0.847 [0.780 - 0.914]
	PatientGraph	0.834 [0.780 - 0.887]	0.786 [0.710 - 0.862]	0.829 [0.759 - 0.899]

11. Tableau III-3. Performances des modèles sur les bases INTERNE et EXTERNE.

Les résultats sont rapportés en valeur [IC95%].

Les résultats sur la base interne et la base externe sont rapportés dans le tableau III-3. Sur la base interne, DL-classifier et PatientGraph ont présenté des performances comparables, avec une sensibilité similaire, bien que la spécificité ait été légèrement plus élevée pour DL-classifier. Sur la base externe, les performances générales ont été légèrement réduites pour les deux modèles, DL-classifier ayant conservé une meilleure spécificité tandis que PatientGraph a montré une sensibilité légèrement supérieure. Globalement, les résultats ont indiqué une bonne cohérence des modèles, mais une généralisation plus limitée sur la base externe.



26. Figure III-3. Prédiction des CDR-SB par la méthode PatientGraph (base INTERNE).

Les résultats de la prédiction du CDR-SB sont présentés dans la Figure III-3. Dans la majorité des cas, les prédictions des classes du CDR-SB correspondaient à la classe réelle ou à une classe adjacente. Il est important de souligner qu'aucun cas de démence sévère n'a été classé comme absence ou forme légère de démence ; réciproquement, aucun sujet indemne ou atteint d'une démence légère n'a été classé en démence sévère.

## III. 4. Discussion

Dans cette étude, nous avons introduit le PatientSpace, une représentation par graphe d'un espace latent capable de différencier les phénotypes de neuroimagerie et de caractériser la sévérité de la DFT. Construit à partir du modèle DIVA (Ilse et al., 2020) entraîné sur des IRM T1w et l'âge des patients, le PatientSpace permet une analyse interprétable des patients, en s'appuyant sur le voisinage des sujets dans l'espace latent.

### III.4.1. Méthodologie

Le PatientSpace repose sur le modèle DIVA, enrichi par plusieurs innovations méthodologiques. La structuration de l'espace latent a nécessité l'introduction de biais inductifs, notamment la classification supervisée et un critère de similarité, afin de contraindre la représentation latente (Locatello et al., 2019). La séparation entre sujets sains et DFT a permis de donner une première structure à l'espace, mais ne suffisait pas à capturer la diversité des sous-types cliniques de la DFT, caractérisés par des patterns d'imagerie hétérogènes (Peet et al., 2021). Pour renforcer l'interprétabilité, nous avons ajouté une perte de consistance, rendant l'espace latent plus robuste aux variations phénotypiques et permettant la construction d'un graphe où la proximité reflète une ressemblance biologique.

### III.4.2. Performance classification diagnostique

La projection de nouveaux sujets dans l'espace latent, combinée à une classification par voisinage, a permis d'obtenir des performances raisonnables. La généralisation sur la base externe a été globalement satisfaisante, bien qu'une baisse de l'AUC et de la sensibilité ait été observée. Cette limite peut s'expliquer par deux facteurs principaux. Premièrement, contrairement aux approches contrastives classiques nécessitant un grand nombre d'exemples négatifs (Yeh et al., 2022), notre modèle a été entraîné avec un nombre restreint, ce qui a probablement limité sa capacité à apprendre des représentations discriminantes robustes. Deuxièmement, l'hétérogénéité intrinsèque des DFT entre bases de données — liée aux différences cliniques, pathologiques, génétiques, mais aussi aux protocoles d'acquisition et critères diagnostiques — complique la généralisation des modèles (Ma et al., 2024). Ainsi, un modèle entraîné sur une base donnée risque de ne pas couvrir toute la variabilité de la maladie, limitant sa transférabilité externe.

### III.4.3. Sous-types et interprétation des clusters

L'analyse du PatientSpace a révélé quatre clusters de patients atteints de démence, caractérisés par des profils cliniques et biologiques distincts. Deux d'entre eux correspondaient majoritairement à des phénotypes connus : le cluster 1 regroupait principalement des patients bvFTD (88,9 %), plus jeunes mais présentant des scores cliniques élevés, tandis que le cluster 2 comprenait surtout des patients SV (85,9 %), d'âge intermédiaire avec des CDR-SB modérés. Les profils d'atrophie associés à ces clusters étaient conformes aux descriptions de la littérature : atteinte frontale prédominante

chez les bvFTD et atteinte temporelle asymétrique, surtout gauche, chez les SV (Rohrer et al., 2015; Whitwell, 2019).

Les clusters 3 et 4 étaient plus hétérogènes, reflétant la diversité clinique et biologique des DFT. L'atrophie y était plus modérée et plus diffuse, touchant notamment les ventricules et les régions insulaires. Ces clusters pourraient correspondre à des formes atypiques ou précoce de la maladie, ou encore à des profils pour lesquels l'IRM ne révèle pas encore d'atrophie spécifique. Leur positionnement, à proximité des sujets sains ou entre les clusters typiques, suggère que le PatientSpace capture potentiellement la dynamique de progression de la maladie, bien que cette hypothèse doive encore être confirmée par des données longitudinales.

### **III.4.3. Interprétation au niveau individuel**

Au-delà des clusters, le PatientSpace a également permis une caractérisation individuelle des patients, tant en termes de diagnostic que de sévérité. L'analyse des voisinages dans le graphe a montré que la prédiction du CDR-SB était correcte dans la majorité des cas, les erreurs étant limitées à des classes adjacentes et sans inversion extrême (par exemple, une démence sévère prédite comme légère). Cette capacité à représenter les caractéristiques individuelles confère au modèle une meilleure interprétabilité et une pertinence clinique renforcée.

### **III.4.4. Limites et perspectives**

Plusieurs limites doivent être soulignées. Premièrement, l'absence de suivi longitudinal empêche de confirmer la progression des patients entre clusters, rendant leur positionnement dans le PatientSpace encore hypothétique. Deuxièmement, malgré l'harmonisation des données multicentriques, des divergences persistent entre les centres (protocoles d'acquisition, critères cliniques, caractéristiques démographiques), ce qui peut limiter la généralisation du modèle. Troisièmement, la taille de la base de données reste modeste et certaines informations cruciales, notamment histopathologiques et génétiques, n'étaient pas disponibles pour tous les sujets. L'intégration de ces données pourrait améliorer la robustesse du modèle en capturant la variabilité biologique et phénotypique des DFT.

Malgré ces limites, le PatientSpace constitue une approche interprétable et flexible, adaptée à l'analyse à la fois individuelle et à l'échelle des groupes. Les perspectives incluent l'optimisation de la fonction de perte afin d'exploiter davantage d'exemples négatifs dans le calcul de similarité, l'intégration de données longitudinales et multimodales (imagerie, génétique, histopathologie), ainsi que le développement de méthodes plus robustes pour prendre en compte l'hétérogénéité des DFT, dans le but d'améliorer la généralisation et la précision clinique du modèle.

## **III.5. Conclusion**

En conclusion, le PatientSpace offre un espace latent interprétable et cliniquement pertinent pour l'analyse des imageries cérébrales chez les patients atteints de DFT. En combinant l'apprentissage profond et analyse basée sur un graphe de similarité, il permet à la fois la découverte de sous-types à l'échelle de la population et une interprétation individualisée des patients, avec des applications potentielles en diagnostic et suivi de la sévérité de la maladie.

### III. 6. Annexes

#### III.6. S1. Dataset externe

Base externe

	DFT	CN	Total
Nombre de sujets	112	111	223
Âge (Moyenne, écart type)	69,4 (9)	62,9 (8,7)	<b>p &lt; 0,0001</b> DFT > CN

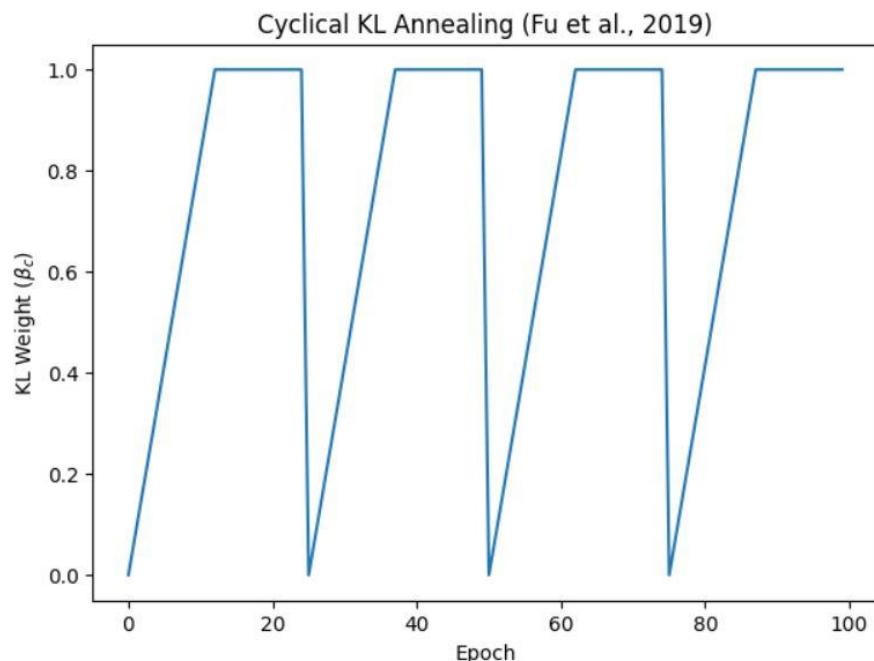
12. Tableau III-S1. Démographie de la base de données externe.

Les participants provenaient majoritairement de la base de Lille (74 %), et dans une moindre mesure de la base NACC (26 %). Les scores CDR-SB n'étaient pas disponibles pour les sujets issus de la base de Lille. Concernant la répartition des phénotypes, la majorité des patients présentaient une forme bvFTD (57 %), tandis que les autres correspondaient à des phénotypes PNFA (4 %), SV (4 %), ou restaient non précisés (35 %).

#### III.6. S2. Détails du modèle

##### III.6. S2.1. Entrainement

Le facteur  $\beta_c$  suivait une règle de mise à jour cyclique, comme illustré en figure III-S1. Les coefficients de pondération  $\lambda$  ont été fixés respectivement à 0,01 pour la fonction de perte de reconstruction et à 100 pour la fonction de perte de classification. L'optimisation a été réalisée à l'aide de l'algorithme AdamW (Loshchilov and Hutter, 2019), avec un taux d'apprentissage initial de  $10^{-4}$  et un paramètre de weight decay de  $10^{-5}$ .



27. Figure III-S1. Règle cyclique du facteur  $\beta_c$ .

### **III.6. S2.2. Nombre de voisins optimal**

Le nombre optimal de voisins K a été déterminé en projetant les sujets de validation dans le PatientSpace et en choisissant la valeur de K qui maximisait l'AUC (K = 10).

### **III.6. S3. Voxel based Morphometry**

Afin d'analyser les profils d'atrophie entre les différents clusters, les IRM structurelles ont été prétraitées à l'aide du logiciel SPM (Statistical Parametric Mapping, <https://www.fil.ion.ucl.ac.uk/spm/>). Dans un premier temps, les volumes ont été segmentés en matière grise, matière blanche et LCR. Les images de matière grise ont ensuite été registrées dans l'espace MNI, puis lissées à l'aide d'un noyau gaussien de 8 mm<sup>3</sup>.

L'analyse statistique a été menée à l'aide de tests t bilatéraux à deux échantillons, en intégrant comme covariables l'âge, le volume intracrânien total, ainsi que le sexe des participants. Ces comparaisons ont permis d'opposer les clusters de patients à des groupes témoins appariés selon l'âge et le sexe. Les résultats ont été corrigés pour comparaisons multiples par le Family-Wise Error rate (FWE), avec un seuil de significativité fixé à 0,05.

## Chapitre IV

# Article “PatientSpace: An interpretable graph -based latent space for multimodal neuroimaging biomarker learning in Alzheimer’s Disease and Frontotemporal Dementia”

---

Dans le chapitre III, nous avons présenté le PatientSpace, un espace latent structuré et interprétable construit à partir d’IRM T1w et de données cliniques, qui organise les sujets de manière à séparer les patients atteints de DFT des CN, tout en rapprochant les individus présentant des profils neuroanatomiques similaires, offrant ainsi une représentation cliniquement informative. Dans le chapitre IV, nous étendons ce cadre à une approche multimodale, intégrant IRM T1w et TEP-FDG, en considérant plusieurs diagnostics simultanément (CN, DFT et Alzheimer) et en conditionnant l’espace latent à l’âge afin de capturer une dimension longitudinale reflétant la progression de la maladie, permettant ainsi de représenter de manière cliniquement informative la variabilité interindividuelle.

## **IV.1. Introduction**

La maladie d'Alzheimer (MA) est un trouble neurodégénératif complexe et hétérogène, présentant une variabilité dans la présentation clinique, la génétique et les caractéristiques en neuroimagerie (Aziz et al., 2017; Ferreira et al., 2020; Habes et al., 2020; Koedam et al., 2010). Cette hétérogénéité se manifeste à tous les stades de la maladie, depuis le déficit cognitif léger (MCI) prodromal jusqu'à la MA avancée (Poulakis et al., 2022), ainsi qu'en fonction de l'âge d'apparition, les formes à début précoce (EOAD) présentant plus fréquemment des formes non amnésiques impliquant les fonctions exécutives, visuelles ou langagières (Koedam et al., 2010; Kuchcinski et al., 2023; Ossenkoppele et al., 2015a).

Au-delà de la MA, la démence frontotemporale (DFT) présente des syndromes cliniques et des caractéristiques neuroimagerie qui se chevauchent, compliquant le diagnostic différentiel (Musa et al., 2020). La DFT comprend des sous-types hétérogènes : la variante comportementale (bvDFT), caractérisée par une atrophie et un hypométabolisme frontaux (Habes et al., 2020), et des variantes langagières telles que la variante sémantique (SV) et l'aphasie progressive non fluente (PNFA). Certains sous-types se chevauchent anatomiquement avec la MA dans les régions temporales ou pariétales (Koenig et al., 2018), et certaines formes cliniques imitent la MA non amnésique, rendant nécessaire un cadre uniifié pour démêler l'hétérogénéité MA/DFT.

La neuroimagerie a été centrale pour caractériser cette hétérogénéité. L'IRM structurelle identifie l'atrophie corticale et sous-corticale comme conséquence tardive de la maladie, tandis que la TEP-FDG détecte le déclin métabolique précoce. Les deux ont révélé des profils anormaux distincts associés aux sous-types cliniques (Ossenkoppele et al., 2015b). Par exemple, l'IRM structurelle a permis de délimiter les phénotypes hippocampal-sparing, limbic-predominant et MA typique (Kuchcinski et al., 2023; Risacher et al., 2017) et des analyses volumétriques basées sur des régions d'intérêt (ROI) ont été utilisées pour distinguer les individus MA, MCI et cognitivement normaux (CN) (Ferreira et al., 2020; Poulakis et al., 2018). De même, les caractéristiques volumétriques et l'épaisseur corticale ont été employées pour sous-typier la DFT (Ma et al., 2024; Peet et al., 2021). Toutefois, ces approches reposent sur des ROI prédéfinies et risquent de négliger des signatures pathologiques plus subtiles et distribuées à travers le cerveau.

L'intelligence artificielle (IA) permet des analyses à l'échelle du cerveau entier, capturant des motifs complexes au-delà des ROI prédéfinies. Les méthodes d'apprentissage profond ont montré leur potentiel pour classifier les sujets MA, DFT et CN à partir de la TEP-FDG (Rogeau et al., 2024) ou de cartes de probabilité d'atrophie basées sur l'IRM (Nguyen et al., 2023). Cependant, les modèles de classification manquent généralement d'interprétabilité et ne capturent pas l'hétérogénéité de la maladie ni les phénotypes atypiques.

La modélisation de l'espace latent répond à ce problème en réduisant les données neuroimagerie de haute dimension en représentations compactes conservant les principales caractéristiques pathologiques. Des études antérieures ont utilisé le clustering pour définir des sous-types de MA (Kang et al., 2024; Park et al., 2017; Poulakis et al., 2018) ou des réseaux antagonistes génératifs pour dériver des représentations liées à la maladie (Louiset et al., 2024; Yang et al., 2022, 2020). Bien que ces méthodes aient révélé des sous-types pertinents de MA, la plupart se concentrent exclusivement sur la MA, reposent sur une seule modalité et capturent principalement des motifs structurels globaux, offrant une interprétabilité limitée au niveau du patient. L'IA générative

présente un avantage supplémentaire en permettant de synthétiser des représentations liées à la maladie et de démêler la variation pathologique de la variabilité saine, mais reste sous-explorée dans la recherche multimodale MA/DFT.

Pour dépasser ces limitations, nous proposons le PatientSpace, un cadre basé sur une extension d'un autoencodeur variationnel (VAE) domaine invariant, intégrant l'IRM T1 et la TEP-FDG dans un espace latent structuré en graphe et interprétable. Le PatientSpace est conçu pour capturer à la fois les altérations anatomiques et métaboliques, et est organisé selon trois dimensions : (1) la classification diagnostique en groupes MA, DFT et CN ; (2) l'âge, reflétant la progression de la maladie et la variabilité interindividuelle (Aziz et al., 2017; Koedam et al., 2010) ; et (3) une mesure de similarité basée sur la neuroimagerie qui positionne les sujets présentant des profils comparables à proximité dans l'espace latent. En combinant la modélisation générative et la représentation basée sur graphe, le PatientSpace permet une analyse unifiée de l'hétérogénéité MA et DFT, démêlant les sous-types de la maladie tout en fournissant des informations interprétables au niveau du patient.

## IV.2. Travaux connexes

### IV.2.1. Apprentissage profond basé sur la neuroimagerie pour la MA et la DFT : de la classification aux modèles génératifs multimodaux

Étant donné le chevauchement des caractéristiques cliniques et neuroimagerie de la MA et de la DFT, de nombreuses études ont exploré des méthodes automatisées de classification et de sous-typage à partir de la neuroimagerie. Ces travaux fournissent un contexte important pour notre étude et mettent en évidence à la fois les progrès réalisés et les limites encore présentes dans l'utilisation de l'IA pour étudier l'hétérogénéité des maladies.

Les approches d'apprentissage profond ont été largement appliquées à la classification de sujets MA, DFT et CN à partir de données neuroimagerie unimodales. Par exemple, Rogeau et al. (Rogeau et al., 2024) ont utilisé un réseau de neurones convolutif 3D pour classifier MA, DFT et CN à partir de TEP-FDG, tandis que Nguyen et al. (Nguyen et al., 2023) se sont appuyés sur des cartes de probabilité d'atrophie dérivées de l'IRM pour la classification. Bien que ces modèles unimodaux soient efficaces, ils sont principalement conçus pour la discrimination diagnostique et offrent peu d'informations sur l'hétérogénéité de la maladie ou sur des profils de neuroimagerie atypiques.

Pour aller au-delà de la simple classification, des stratégies de regroupement basées sur les mesures d'atrophie corticale ou les caractéristiques volumétriques ont été appliquées pour définir des sous-types de MA (Kang et al., 2024; Park et al., 2017; Poulakis et al., 2018). Des cadres génératifs tels que SMILE-GAN (Yang et al., 2022, 2020) et SepVAE (Louiset et al., 2024) permettent en outre de démêler les variations liées à la maladie de celles observées chez les sujets sains en générant des représentations de type MA à partir de sujets CN. Malgré leur potentiel, la plupart de ces approches restent limitées à la MA et aux analyses unimodales, négligeant des informations multimodales complémentaires. Pour pallier cette limitation, Antelmi et al. (Antelmi et al., 2019) ont introduit un autoencodeur variationnel multicanal intégrant IRM T1, TEP-FDG et TEP-amyoïde, démontrant l'intérêt de combiner données structurelles et métaboliques pour le sous-typage de la MA. Cependant, les approches génératives multimodales ont rarement été étendues à la DFT, et peu

offrent des espaces latents structurés et interprétables capturant l'hétérogénéité à la fois dans les populations MA et DFT.

#### **IV.2.2. Vers des espaces latents interprétables en neuroimagerie**

Les VAE sont puissants pour extraire des représentations latentes, mais leur interprétabilité reste limitée. En neuroimagerie, cette interprétabilité est cruciale car elle permet de relier le comportement du modèle aux mécanismes de la maladie et d'améliorer la confiance clinique. Pour relever ce défi, le concept de représentations latentes démêlées a été introduit, où chaque dimension latente correspond idéalement à un facteur génératif distinct de variation (Bengio et al., 2013; Burgess et al., 2018).

Le démêlage est souvent obtenu en introduisant une tâche de prédiction auxiliaire ou en conditionnant les sous-espaces latents sur des variables d'intérêt, alignant ainsi certaines dimensions avec des sources de variation ciblées. Cette stratégie a montré des applications prometteuses en neuroimagerie. Par exemple, Zhao et al. (Zhao et al., 2019b) ont conditionné l'espace latent sur l'âge chronologique, permettant des reconstructions capturant les modifications liées à l'âge sur l'IRM. De même, Liu et Yap (Liu et Yap, 2024) ont séparé la variation d'imagerie liée au site des signaux biologiques, facilitant l'harmonisation des données. D'autres approches ont visé à isoler la variabilité pathologique : Louiset et al. (Louiset et al., 2024) ont séparé les sous-espaces latents sains et pathologiques, tandis que Kang et al. (Kang et al., 2024) ont combiné démêlage et regroupement pour révéler des sous-groupes pertinents liés à la maladie.

Ces méthodes exploitent souvent les facteurs démêlés pour manipuler les images, par exemple en fixant ou supprimant une variable lors de la reconstruction afin d'harmoniser les données ou visualiser des effets spécifiques (âge, pathologie) (Ainsworth et al., 2018; Burgess et al., 2018; Li et al., 2019). Cependant, l'interprétabilité de ces cadres reste principalement liée à des facteurs de variation globaux (âge, site, pathologie) plutôt qu'aux différences au niveau du patient.

Notre travail est motivé par cette limitation. Nous visons à construire des espaces latents démêlés où la variation est guidée non seulement par les sources issues de l'imagerie, mais aussi par la similarité entre patients. Ce changement rapproche l'interprétabilité du niveau clinique, permettant d'explorer les différences individuelles entre patients plutôt que de se limiter aux facteurs de variation au niveau populationnel.

### **IV.3. Matériels et méthodes**

#### **IV.3.1. Participants et acquisitions des données**

Les données ont été compilées à partir de bases publiques et de bases privées locales. Le jeu de données provient de bases publiques accessibles : l'Alzheimer's Disease Neuroimaging Initiative (ADNI, <https://adni.loni.usc.edu/>), l'étude longitudinale sur la dégénérescence lobaire frontotemporale (NIFD, <http://memory.ucsf.edu/research/studies/nifd>) ainsi que de bases privées locales du CHU de Lille.

Les participants ADNI comprenaient des individus issus des protocoles ADNI1, ADNI2, ADNIGO et ADNI3. Les participants devaient disposer à la fois d'une IRM 3D T1 et d'une TEP-FDG réalisées lors

de la même session et être diagnostiqués soit CN, soit MA. Les participants présentant un trouble cognitif léger (MCI) dans la cohorte ADNI ont été exclus de l'étude.

De la même manière, les participants de la cohorte NIFD devaient disposer d'une IRM T1w et d'une TEP-FDG réalisées lors de la même session, et être diagnostiqués comme CN ou DFT. Le diagnostic de DFT reposait sur les critères de Rascovsky et al. (Rascovsky et al., 2011), complétés par les critères de Gorno-Tempini et al. pour les variantes langagières (Gorno-Tempini et al., 2004).

Le jeu de données du CHU de Lille comprenait des participants diagnostiqués avec la MA ou la DFT, suivis au Centre de Mémoire de Lille entre 2010 et 2020. Les participants MA ont été inclus sur la base de la disponibilité d'une IRM T1 3D et d'une TEP-FDG réalisées dans un intervalle de 6 mois, ainsi que d'un diagnostic probable de MA selon les critères NINCDS/ADRDA (McKhann et al., 1984), conformes aux critères utilisés dans la cohorte ADNI. Alternativement, les participants présentant un MCI dû à la MA avec probabilité intermédiaire, selon les critères NIA-AA de 2011 (Albert et al., 2011) ont été inclus, sans restriction sur le score MMSE au moment de l'IRM. Tous les patients ont bénéficié d'une IRM 3D T1 et d'une TEP-FDG, généralement lors de la première année de suivi, dans le cadre de leur évaluation clinique pour confirmer le diagnostic de MA. Les examens supplémentaires réalisés ultérieurement, pour des suivis de recherche, le suivi thérapeutique ou l'évaluation de la progression de l'atrophie cérébrale, notamment chez les patients MCI inclus, ont également été considérés.

Les sujets DFT ont été inclus sur la base d'un diagnostic probable de DFT selon les critères de Rascovsky (Rascovsky et al., 2011), sans restriction sur le score MMSE au moment de l'IRM. Chaque patient a bénéficié d'une TEP-FDG cérébrale au cours de la première année de suivi, car l'hypométabolisme frontal ou temporal antérieur constitue un critère diagnostique de DFT probable. Les examens supplémentaires réalisés dans le temps ont été inclus pour des raisons similaires à celles du groupe MA (recherche, suivi thérapeutique, suivi de l'atrophie frontale).

Le détail des informations relatives à chaque IRM T1 et TEP-FDG est fourni en section IV.8. S1. Un accord écrit a été obtenu de tous les participants du CHU de Lille.

### **IV.3.2. Prétraitement des images**

Les images T1w ont été prétraitées comme suit : (i) correction d'inhomogénéité de champ avec N4ITK (ii) extraction du cerveau avec SynthStrip ([Hoopes et al., 2022](#)), (iii) registration affine vers l'espace MNI de 1 mm<sup>3</sup> à l'aide d'outils de normalisation avancée (ANTs, <https://github.com/ANTsX/ANTs>), (iv) recadrage à 160 x 192 x 160 voxels, et (v) normalisation du signal cérébral par z-score.

Les images TEP-FDG ont été prétraitées comme suit : (i) Moyennage des images dynamiques avec FreeSurfer si non déjà effectué (<https://surfer.nmr.mgh.harvard.edu/>), (ii) extraction du cerveau avec SynthStrip, (iii) registration linéaire vers l'IRM T1 avec ANTs, (iv) registration affine vers l'espace MNI de 1 mm<sup>3</sup> à l'aide de la transformation T1w à MNI, (v) recadrage à 160 x 192 x 160 voxels et (v) normalisation du signal cérébral par z-score.

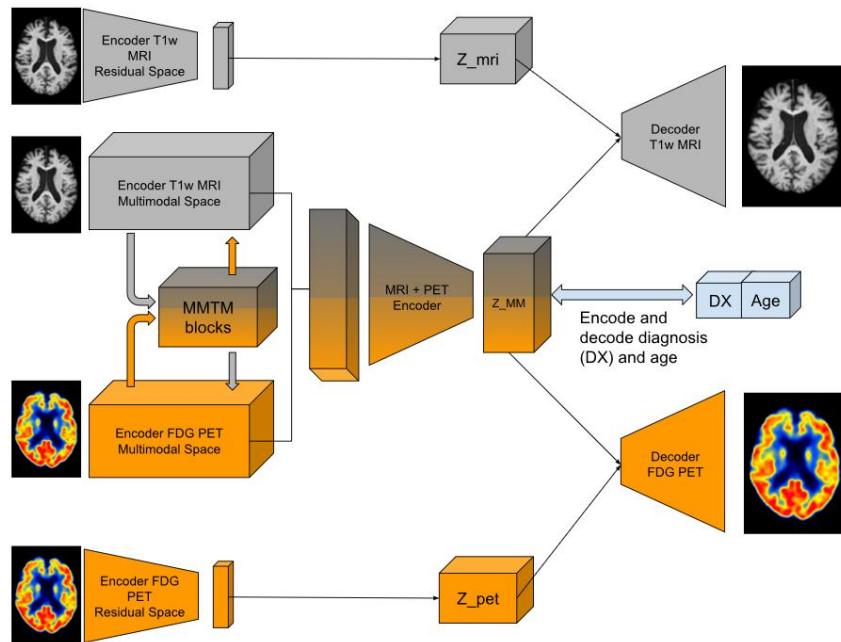
### **IV.3.3. Ensemble des données**

Le jeu de données complet comprenait 1 522 paires de TEP-FDG et d'IRM T1 3D prétraitées. Les ensembles d'entraînement, de validation et de test ont été séparés aléatoirement selon une

répartition 60 % – 20 % – 20 %, stratifiée par diagnostic et distribution d'âge. Cette répartition a abouti à 946 paires de scans pour l'ensemble d'entraînement, 283 paires pour l'ensemble de validation et 293 paires pour l'ensemble de test. Pour chaque participant, l'âge, le sexe, le diagnostic et le score MMSE ont été enregistrés au moment de l'acquisition de l'IRM ( $\pm 6$  mois).

#### IV.3.4. Construction et extraction d'un espace latent structuré

##### IV.3.4.1. Vue d'ensemble de la méthode



28. Figure IV-1. Architecture réseau.

Le modèle comprend quatre encodeurs : deux encodeurs unimodaux résiduels, qui génèrent des représentations latentes unimodales résiduelles ( $Z_{mri}$  et  $Z_{pet}$ ), et deux encodeurs unimodaux supplémentaires enrichis par des blocs MMTM. Les sorties des encodeurs pilotés par les blocs MMTM sont concaténées et projetées dans un espace latent multimodal partagé ( $Z_{MM}$ ). Cette représentation multimodale, conditionnée par l'âge et le diagnostic, est utilisée conjointement avec les espaces latents unimodaux pour décoder chaque modalité séparément.

**Architecture.** Le modèle utilisé pour construire l'espace latent est une extension du DIVA (Domain Invariant Variational Autoencoder) (Ilse et al., 2020), enrichie par une régularisation consistante (Sinha and Dieng, 2022). L'objectif est d'apprendre un espace latent multimodal  $Z_{MM}$ , structuré par diagnostic, âge et régularisation consistante, ainsi que deux espaces latents résiduels spécifiques aux modalités ( $Z_{MRI}$  et  $Z_{PET}$ ). Le modèle comprend trois encodeurs (Figure IV-1) : deux encodeurs unimodaux générant  $Z_{MRI}$  et  $Z_{PET}$ , et un encodeur multimodal produisant  $Z_{MM}$ . Chaque modalité est ensuite reconstruite par des décodeurs indépendants : la reconstruction IRM est réalisée à partir de  $(Z_{MRI}, Z_{MM})$ , tandis que la reconstruction TEP-FDG utilise  $(Z_{PET}, Z_{MM})$ .

**Encodeur multimodal.** Notre modèle étend l'architecture décrite par Ilse et al. (Ilse et al., 2020) à un contexte multimodal. Deux encodeurs de caractéristiques spécifiques à chaque modalité sont couplés à un Multimodal Transfer Module (MMTM) (Vaezi Joze et al., 2020). Les blocs MMTM sont intégrés dans les trois dernières couches de chaque encodeur, permettant de capturer des caractéristiques spécifiques à chaque modalité et d'aligner les informations partagées. La fusion des

caractéristiques s'effectue ensuite par concaténation des canaux, avant un passage dans un bloc résiduel 3D (He et al., 2016) qui réduit le nombre de canaux de moitié.

Tous les encodeurs unimodaux suivent le même design : des blocs résiduels 3D comprenant des convolutions 3D, une GroupNormalization (GN) (Wu and He, 2018), choisie pour ses performances sur de petits lots, et une fonction d'activation ReLU. Le sous-échantillonnage est effectué par convolutions avec stride 2, doublant le nombre de canaux, suivi d'une activation ReLU. La sortie de chaque encodeur est aplatie puis passée à travers des couches linéaires pour générer  $\mu$  et  $\log \sigma^2$  (Kingma and Welling, 2019).

**Décodeurs.** Les décodeurs reproduisent la structure des encodeurs : un upsampling trilinéaire 3D double les dimensions spatiales, suivi de convolutions 3D, puis d'un bloc résiduel 3D. La dernière couche de chaque décodeur utilise des convolutions 3D avec un noyau  $1 \times 1 \times 1$ , tandis que toutes les autres convolutions utilisent un noyau  $3 \times 3 \times 3$ .

**Fonction de perte.** Le modèle est entraîné avec une fonction de perte composite combinant plusieurs objectifs : reconstruction neuroimagerie, classification diagnostique, prédiction de l'âge et régularisation de l'espace latent. Selon l'approche d'apprentissage multitâche (Caruana, 1993), la perte globale  $L_{overall}$  est définie comme :

$$L_{overall} = \lambda_1 w_1 L_1 + \lambda_2 w_2 L_2 + \lambda_3 w_3 L_3 + \lambda_4 w_4 L_4 \quad (\text{Eq.IV-1})$$

Chaque  $L_t$  correspond à la perte d'une tâche spécifique. Les constantes  $\lambda_t$  contrôlent l'échelle des gradients, les différences de magnitude et la priorité des tâches. Les poids dynamiques  $w_t$  ajustent de manière adaptative la contribution de chaque tâche en fonction de sa vitesse d'apprentissage (Liu et al., 2019) :

$$w_t(e) = (\text{nombre de tâche})(\exp\left(\frac{L_t(e-1)}{2 L_t(e-2)}\right)) / \left[\sum_T \exp\left(\frac{L_T(e-1)}{2 L_T(e-2)}\right)\right] \quad (\text{Eq IV-2})$$

Où,  $e$  désigne l'époque d'entraînement actuelle,  $L_t(e-1)$  et  $L_t(e-2)$  sont les pertes spécifiques à la tâche pour les deux époques précédentes, et  $T$  désigne l'ensemble de toutes les tâches.

Les définitions spécifiques de chaque terme de perte  $L_t$  sont les suivantes. :

$L_1$  constitue une extension multimodale de la perte  $\beta$ -VAE du modèle DIVA (Ilse et al., 2020) :

$$\begin{aligned} L_1 = & (X_{MRI} - \hat{X}_{MRI})^2 + \beta D_{KL}[Z_{MRI} || N(0, I)] \\ & + (X_{PET} - \hat{X}_{PET})^2 + \beta D_{KL}[Z_{PET} || N(0, I)] \\ & + \beta D_{KL}[Z_{MM} || p(Z_{MM} | age, diagnosis)] \end{aligned} \quad (\text{Eq IV-3})$$

où  $X$  est l'entrée pour chaque modalité,  $\hat{X}$  sa reconstruction,  $N(0, I)$  la gaussienne multivariée de moyenne nulle et covariance identité, et  $\beta$  un coefficient positif équilibrant la perte KL et la perte de reconstruction.  $p(Z_{MM} | age, diagnosis)$  correspond au prior conditionnel appris à partir de l'âge et du diagnostic pour récupérer  $Z_{MM}$ .

$L_2$ , la perte de classification diagnostique, utilise l'entropie croisée pondérée pour gérer le déséquilibre entre classes : les poids  $\alpha_D$  assignés à chaque diagnostic  $D$  sont définis comme  $\alpha_D = \frac{N_{total}}{N_D \times C}$ , où  $N_{total}$  est le nombre total de scans,  $N_D$  est le nombre de scans pour le diagnostic  $D$  et  $C$  est le nombre total de classes de diagnostic.

Pour la perte de prédiction de l'âge  $L_3$ , l'erreur quadratique moyenne (EQM) a été utilisée.

La perte de consistance  $L_4$  étend la régularisation consistante introduite par Sinha et Dieng en encourageant l'alignement entre l'espace latent multimodal dérivé de l'image originale  $Z_{MM}^{ORIG}$  et celui calculé à partir de sa version transformée aléatoirement  $Z_{MM}^{AUG}$  – obtenue par translation, rotation, zoom et/ou ajout de bruit gaussien. Elle est définie comme une divergence de Kullback-Leibler (KL) symétrique :

$$L_4 = \frac{1}{2} [D_{KL}(Z_{MM}^{ORIG} || Z_{MM}^{AUG}) + D_{KL}(Z_{MM}^{AUG} || Z_{MM}^{ORIG})] \quad (\text{Eq. IV-4})$$

où  $D_{KL}$  désigne la divergence KL. Contrairement à la formulation originale, la perte est symétrisée pour garantir la robustesse de l'espace latent multimodal. Des détails supplémentaires sur chaque composante de perte et la procédure d'entraînement sont fournis en matériel supplémentaire (section IV.8. S2).

**Procédure d'entraînement.** La procédure suit celle décrite Sinha et Dieng (Sinha and Dieng, 2022), en appliquant la perte VAE aux images originales  $X^{ORIG}$  et transformées  $X^{AUG}$ . Pour limiter le surapprentissage, la prédiction de l'âge et du diagnostic n'a été effectuée que sur  $X^{AUG}$ . Conformément aux recommandations d'Alemi et al. (Alemi et al., 2018) pour extraire des représentations latentes significatives, une pondération constante élevée a été attribuée à la perte de classification ( $L_2$ ). Les poids associés aux pertes de régularisation de l'espace latent ( $L_1$  et  $L_4$ ) ont été augmentés linéairement au cours des dix premières époques –  $\beta$  est passé de 0,1 à 1 pour  $L_1$  et  $\lambda_4$  de 0,2 à 2 pour  $L_4$ . Des détails supplémentaires sur la procédure d'entraînement sont donnés à la section IV.8. S3.1.

**Augmentation des données** Nous avons appliqué des augmentations aux images IRM et TEP-FDG via la bibliothèque MONAI (Cardoso et al., 2022). Les transformations comprenaient : translation aléatoire ( $\pm 10$  voxels sur chaque axe), rotation aléatoire ( $\pm 0,4$  radians), et zoom aléatoire (facteur 0,9–1). Les flips étaient exclus afin de préserver les asymétries potentielles des profils pathologiques (Madhavan et al., 2013; Whitwell, 2019). Du bruit gaussien ( $\mu = 0, \sigma \sim U([0, 0.15])$ ) a été ajouté, et les entrées ont été débruitées pendant la reconstruction, conformément à des travaux montrant que cette approche améliore la robustesse des caractéristiques (Vincent et al., 2008). Les images IRM et TEP d'un même sujet ont subi des augmentations identiques, chacune appliquée avec une probabilité de 0,95.

**Sélection du modèle.** Le modèle final a été choisi sur la base de la perte minimale combinée pour la prédiction d'âge, la classification et la régularisation consistante sur l'ensemble de validation.

#### IV.3.4.2. Représentation par graphe de l'espace latent – PatientSpace

Nous avons construit une représentation sous forme de graphe appelé PatientSpace, définie comme  $G = (V, E, A)$  où  $V$  correspond à l'ensemble des nœuds représentant les patients,  $E$  à l'ensemble des arêtes connectant ces nœuds, et  $A$  à la matrice d'adjacence pondérée. Chaque nœud  $v_i \in V$  correspond à un patient, et la pondération des arêtes encode la similarité morphologique entre patients dans l'espace latent.

Pour chaque paire de sujets  $(i, j)$ , la pondération  $a_{ij}$  de la matrice d'adjacence était définie de la manière suivante :

$$a_{ij} = \begin{cases} \frac{1}{d(i,j)} / \sum_{k \in N(i)} \frac{1}{d(i,k)}, & \text{si } j \text{ est un voisin de } i \\ 0, & \text{Sinon} \end{cases} \quad (\text{Eq IV-5})$$

où,  $d(i, j)$  représente la mesure de la dissimilarité (perte de consistance  $L_4$ ) entre  $i$  et  $j$ .

$N(i)$  représente les  $K$  voisins de  $i$  (voir la section IV.8. S3.2 pour le nombre optimal de voisins  $K$ )

#### IV.3.5. Expérimentations

##### IV.3.5.1. Analyse basée sur le graphe

Deux analyses ont été menées pour étudier les propriétés structurelles du PatientSpace. La première s'est concentrée sur la topologie globale en identifiant des clusters au sein du graphe. La seconde a mis l'accent sur les informations locales en réalisant une analyse de voisinage au niveau de chaque patient.

**Clustering dans le PatientSpace.** Le PatientSpace a été clusterisé en utilisant un algorithme de clustering agglomératif avec le critère de Ward, tout en préservant la structure du graphe. Les clusters ont été analysés selon : (i) le nombre de sujets CN, AD et FTD, (ii) la répartition selon l'âge et le sexe, (iii) les volumes cérébraux et le ratio de valeur normalisée d'absorption standardisée (SUVR), et le score MMSE ajusté pour l'âge et le sexe. Les comparaisons statistiques ont été réalisées à l'aide du test du khi-deux pour les variables catégorielles et du test de Kruskal-Wallis pour les variables continues, avec correction pour comparaisons multiples via la méthode du faux taux de découverte de Benjamini-Hochberg (BH-FDR) (Benjamini et Hochberg, 1995). Les profils d'atrophie et de métabolisme des clusters de maladies ont été comparés via une approche de Voxel-Based Morphometry (VBM) (voir données supplémentaires IV.8. S5).

**Analyse de PatientSpace au niveau individuel.** Le PatientSpace a également été examiné au niveau du patient en prédisant le diagnostic, les volumes cérébraux, le SUVR et le MMSE chez les sujets de l'ensemble test. Pour chaque sujet non vu  $i$  de l'ensemble test, la probabilité d'être assigné au diagnostic  $D$  a été estimée par

$$\operatorname{argmax}_D P(\widehat{DX}_i = D) = \sum_j a_{ij} 1(DX_j = D) \quad (\text{Eq IV-6})$$

où  $DX_j$  est le diagnostic du sujet  $v_j$ .

L'aire sous la courbe (AUC), la spécificité et la sensibilité ont été utilisées pour évaluer la performance de classification diagnostique.

Les volumes cérébraux et le SUVR ont été prédits pour des régions d'intérêt (ROI) spécifiques associées à la MA et la DFT. Les ROI liées à la MA incluaient l'hippocampe, le précunéus, le gyrus parahippocampique, le cortex entorhinal, le cortex cingulaire, le gyrus temporal moyen latéral, le gyrus frontal et le gyrus angulaire (Fennema-Notestine et al., 2009; Landau et al., 2011). Les ROI liées à la DFT comprenaient le cortex insulaire, l'insula antérieure, le pôle temporal, le gyrus temporal moyen, le gyrus frontal moyen, le gyrus frontal supérieur et le gyrus frontal inférieur (Peet et al., 2021; Risacher and Saykin, 2013; Whitwell, 2019). Nous avons également inclus dans notre analyse la matière grise (GM), de la matière blanche (WM), le liquide céphalo-rachidien (LCR), ainsi que les régions frontales et temporales. Les volumes cérébraux ont été extraits avec AssemblyNet (Coupé et al., 2020), et normalisés par le volume intracrânien, tandis que le SUVR a été calculé en divisant la SUV régionale par la SUV totale.

Pour chaque sujet  $i$  de l'ensemble de test, non vu lors de l'entraînement, le volume cérébral ou le SUVR estimé  $\hat{y}$  a été calculé comme la moyenne pondérée des mesures  $y$  correspondantes des sujets voisins, selon la formule suivante :

$$\hat{y}_i = \sum_j a_{ij} y_j \quad (\text{Eq IV-7})$$

Le coefficient de corrélation de Pearson a été calculé entre les valeurs réelles et prédites. La force de la corrélation a été interprétée comme : faible ( $r < 0.3$ ) ; modérée ( $0.3 \leq r < 0.6$ ) ; et forte ( $r \geq 0.6$ ) (Akoglu, 2018).

Le MMSE a été estimé de manière similaire au diagnostic. Les scores MMSE ont été catégorisés selon les stades de démence : absence ou suspicion de démence (MMSE 27–30), démence légère (MMSE 20–26), démence modérée (MMSE 10–19) et démence sévère (MMSE < 10) (Perneczky et al., 2006). Pour chaque sujet  $i$  de l'ensemble de test, non vu lors de l'entraînement, la probabilité d'obtenir un score MMSE  $m$  a été estimée à l'aide de :

$$\operatorname{argmax}_m P(\widehat{\text{MMSE}}_i = m) = \sum_j a_{ij} 1(\text{MMSE}_j = m) \quad (\text{Eq IV-8})$$

Les performances MMSE ont été évaluées à l'aide de la précision équilibrée, de la régression ordinaire, ainsi que de la MAE et de la RMSE, optimisées pour les données déséquilibrées (Baccianella et al., 2009).

Les résultats obtenus à partir des approches basées sur les graphes sont désignés sous le nom PatientGraph et ceux obtenus par le classifieur natif du réseau sous le nom DL-classifier dans les sections suivantes.

#### **IV.3.5.2. Validation externe**

**Classification** Nous avons utilisé la base de données SOCRATES pour évaluer les performances de classification. Une description détaillée de la population étudiée est fournie en section IV.8. S2.2.

**Analyse de survie MCI** Nous avons utilisé le jeu de données MEMENTO (numéro d'autorisation 2010 A01394 35) pour étudier la conversion des MCI et suivre la progression de la maladie. Ce jeu

multicentrique comprenait des individus présentant un MCI, indépendamment de l'étiologie sous-jacente, telle que la MA, la DFT ou d'autres pathologies. Une description détaillée de la population est fournie en section IV.8. S2.2.

L'analyse de survie a été utilisée pour étudier la probabilité de conversion des MCI en démence. La première session des sujets MCI de la cohorte MEMENTO a été projetée dans le PatientSpace et assignée à un cluster  $c$  selon :

$$\operatorname{argmax}_c P(\widehat{\text{Cluster}}_i = c) = \sum_j a_{ij} \mathbf{1}(\text{Cluster}_j = c) \quad (\text{Eq IV-9})$$

Une analyse de survie de Kaplan–Meier a ensuite été appliquée pour examiner les taux de conversion selon les clusters. Les différences statistiques entre distributions de survie ont été évaluées via des tests log-rank par paires (Fleming and Harrington, 1981) avec correction BH-FDR pour comparaisons multiples (Benjamini and Hochberg, 1995). Pour garantir la robustesse des tests, seuls les clusters contenant plus de dix sujets ont été conservés. Les trajectoires longitudinales des sujets MCI converters ont été suivies via une matrice de transition. À chaque point temporel où les données IRM et TEP-FDG étaient disponibles, l'appartenance au cluster a été réassignée selon l'équation l'Eq. IV-9.

#### IV.3.5.3. Comparaison avec des méthodes de référence

Pour évaluer les performances du modèle, nous avons comparé notre module de classification deep learning (DL-Classifier) et notre approche PatientGraph à deux méthodes d'apprentissage profond sur l'ensemble de test : (i) PET-CNN, un classifieur basé sur la TEP-FDG réimplémenté à partir de Rogeau et al. (Rogeau et al., 2024) (ii) AssemblyNet-AD-FTD, un classifieur basé sur l'IRM T1 décrit par Nguyen et al. (Nguyen et al., 2023). Nous avons réalisé une classification multiclass et en avons dérivé une classification CN vs démence (DEM) à partir des résultats multiclass pour examiner les erreurs de classification.

#### IV.3.6. Etudes d'ablation

Pour évaluer l'impact de notre architecture deep learning et de l'organisation de l'espace latent, cinq variantes ont été mises en place. Deux variantes unimodales ont été implémentées afin d'évaluer l'influence de la multimodalité sur l'espace latent, appelées respectivement modèle TEP et modèle IRM. Pour les autres variantes, l'architecture multimodale a été conservée, mais certaines composantes ont été supprimées successivement : (1) l'âge, (2) la perte de consistance, et (3) le diagnostic. Ces modèles sont respectivement désignés age ablated, consistency ablated et vanilla model.

### IV.4. Résultats

#### IV.4.1. Données démographiques

Les caractéristiques démographiques des sujets sont résumées dans le tableau IV-1. L'âge était significativement plus élevé chez les sujets CN, tandis qu'aucune différence significative n'a été observée entre les groupes MA et DFT dans l'ensemble des jeux de données. La répartition selon le sexe ne montrait pas de différence significative dans les ensembles d'entraînement et de validation, mais une différence significative était observée entre les groupes MA et DFT dans l'ensemble de

test. Les scores de MMSE différaient significativement dans tous les ensembles de données, les sujets CN obtenant les scores les plus élevés, suivis des DFT, puis des MA.

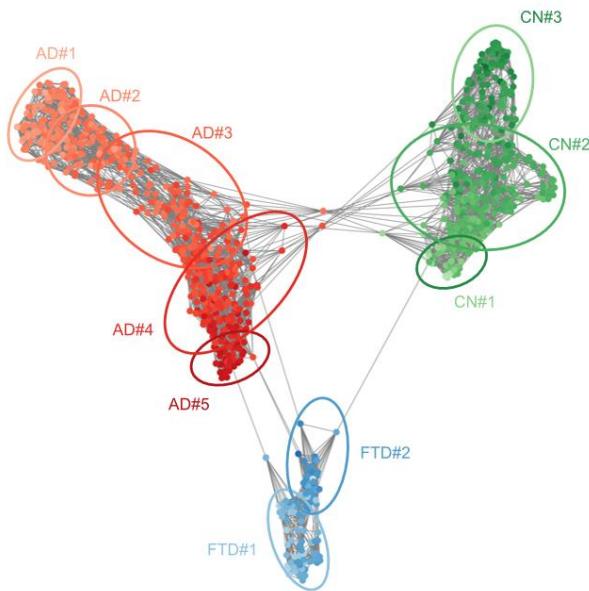
	CN	MA	DFT	Valeur p
<i>Ensemble d'entraînement</i>				
Nombre de scans	336	504	106	-
Âge, années	74.20 ± 7.76	68.31 ± 10.33	65.81 ± 9.19	< 0,001 MA < CN DFT < CN
Femmes, n (%)	169 (50%)	270 (54%)	45 (43%)	0.11
MMSE	29 (28 - 30)	20 (15 - 23)	25 (21 - 27)	< 0,001 MA < DFT < CN
<i>Ensemble de validation</i>				
Nombre de scans	89	155	39	-
Âge, années	73,64 ± 6,77	67,68 ± 8,97	68,22 ± 7,06	< 0,001 MA < CN DFT < CN
Femmes, n (%)	41 (46%)	58 (37%)	15 (38%)	0.40
MMSE	30 (29-30)	20 (16.5 - 23)	25 (22 - 28)	< 0,001 MA < DFT < CN
<i>Ensemble de test</i>				
Nombre de scans	97	150	46	-
Âge, années	74,14 ± 6,47	68.18 ± 9.34	68.76 ± 7.34	< 0,001 MA < CN DFT < CN
Femmes, n (%)	60 (62%)	68 (45%)	11 (24%)	0.03
MMSE	29.5 (29 - 30)	21 (16 - 24)	24 (17.5 - 27)	< 0,001 MA < DFT < CN

### 13. Tableau IV-1. Caractéristiques démographiques et cliniques des données.

L'âge est rapporté sous forme de moyenne ± écart type, le sexe féminin comme nombre (proportion), et le MMSE comme médiane avec intervalle interquartile (Q1–Q3). Les comparaisons statistiques ont été réalisées à l'aide de tests de Kruskal-Wallis pour l'âge et le MMSE, et d'un test du chi-carré pour le sexe. Une valeur de  $p \leq 0,05$  était considérée comme statistiquement significative.

## IV.4.2. Interprétation du PatientSpace

### IV.4.2.1. PatientSpace clusters



29. Figure IV-2. Graphe du PatientSpace

Chaque nœud correspond à un sujet et chaque arête à une connexion binaire ( $1$  si  $a_{ij} > 0$ ;  $0$  sinon). Les sujets MA sont représentés en rouge, les DFT en bleu et les CN en vert. Les clusters extraits par classification hiérarchique agglomérative sont encerclés et nommés.

La figure IV-2 présente le PatientSpace, qui se compose de dix clusters : trois clusters CN, deux clusters DFT et cinq clusters MA. Les caractéristiques des clusters sont présentées dans le tableau IV-2. Ces clusters présentaient des variations dans la distribution de l'âge. Les clusters DFT ne montraient pas de différences notables selon le sexe, contrairement aux clusters MA, le cluster AD#1 comprenant la proportion la plus élevée de femmes et le cluster AD#5 la plus faible. De plus, AD#1 et AD#2 présentaient des scores MMSE significativement plus bas que les autres clusters MA. Les sujets DFT partageaient davantage de liens avec les individus MA, en particulier ceux des clusters AD#3 et AD#4, tout en ne montrant qu'une seule interaction avec les sujets CN. À l'inverse, les sujets CN présentaient une plus grande similarité avec les sujets AD#1, AD#2 et AD#3. De manière intéressante, le cluster AD#5 apparaissait isolé, sans interaction observée avec les CN ou les DFT.

	Âge, années	Femmes, n (%)	MMSE
AD#1	$56.07 \pm 3.95$	64 (75%)	15 (9 - 19)
AD#2	$61,22 \pm 4,29$	65 (52%)	16 (11 - 20)
AD#3	$68.66 \pm 6.00$	59 (53%)	21.5 (18.25 - 23.75)

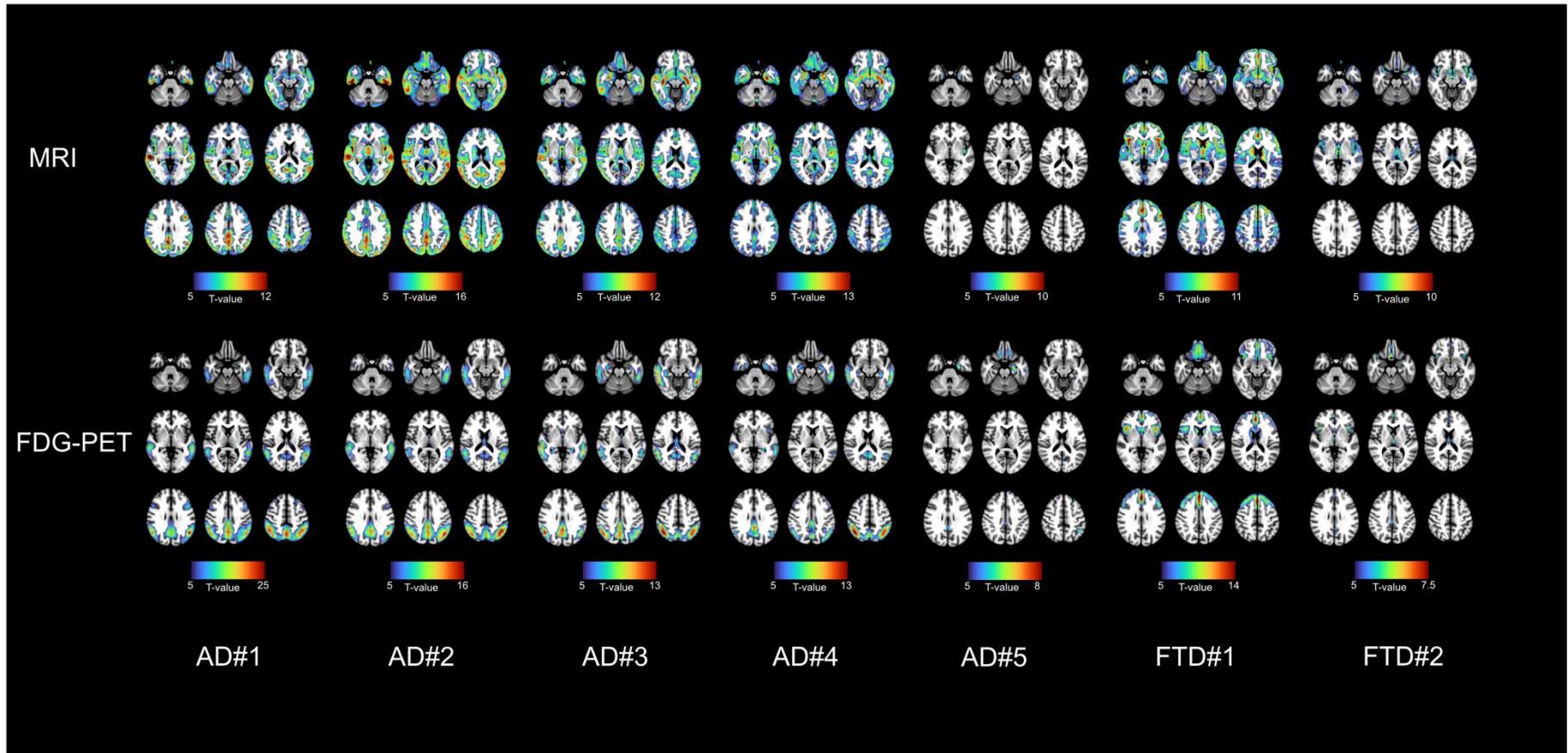
AD#4	$76.74 \pm 4.29$	71 (50%)	23 (20 - 25)
AD#5	$85,37 \pm 3,83$	11 (27%)	22 (20 - 24)
FTD#1	$61.61 \pm 7.25$	23 (38%)	25 (20 - 27)
FTD#2	$71.29 \pm 8.59$	22 (47%)	25 (21 - 27.75)
CN#1	$62,63 \pm 7,19$	23 (60%)	30 (29 - 30)
CN#2	$72.95 \pm 5.11$	106 (53%)	29 (28 - 30)
CN#3	$81.25 \pm 5.34$	40 (41%)	29 (28 - 30)

*14. Tableau IV-2. Caractéristiques des clusters.*

*L'âge est exprimé en moyenne  $\pm$  écart-type, la proportion de femmes est indiquée en effectif (pourcentage) et le MMSE est présenté en médiane avec l'intervalle interquartile (Q1-Q3).*

#### IV.4.2.2. Modèles spécifiques d'atrophie et d'hypométabolisme

Des profils distincts d'atrophie et d'hypométabolisme ont été observés selon les clusters pathologiques (figure IV-3). Les clusters AD#1, AD#2, AD#3 et AD#4 présentaient les caractéristiques classiques de la MA, comprenant un hypométabolisme dans les lobes temporaux et le précuneus, associé à une atrophie des régions temporales postérieures et médiales ainsi que de l'hippocampe. Bien que ces clusters partageaient globalement des profils spatiaux similaires, l'hypométabolisme de la TEP-FDG était plus marqué dans AD#1 et AD#2. Fait notable, AD#1 et AD#2 ne montraient pas d'hypométabolisme hippocampique. Par ailleurs, le cluster AD#2 présentait une atteinte postérieure plus prononcée. À l'inverse, le cluster AD#5 montrait peu ou pas d'hypométabolisme détectable et une atrophie minimale, limitée au lobe temporal gauche. Le cluster FTD#1 présentait un hypométabolisme caractéristique des lobes frontaux et du noyau caudé, ainsi qu'une atrophie typique des régions frontales et temporales. En revanche, le cluster FTD#2 ne présentait quasiment pas d'hypométabolisme détectable dans les régions frontales, avec une atrophie principalement localisée dans les régions temporales et l'insula.



**30. Figure IV-3. Atrophie et hypométabolisme dans les clusters.**

Les rangées représentent les modalités d'imagerie, tandis que les colonnes correspondent aux clusters de démence.

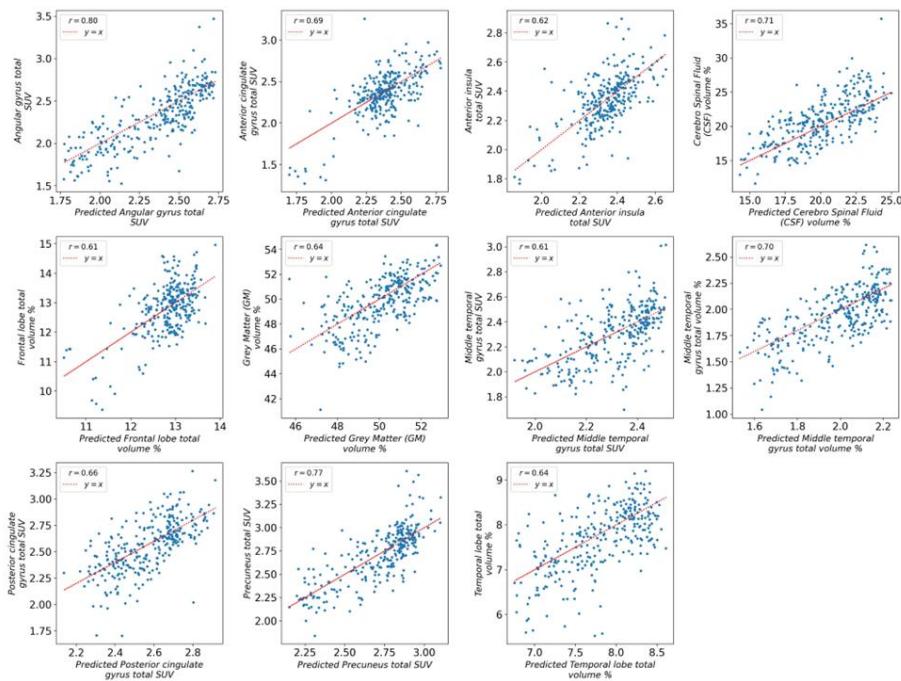


#### IV.4.2.3. Caractérisation individuelle de PatientSpace

Le tableau IV-S3 présente les coefficients de corrélation de Pearson entre les volumes cérébraux et les valeurs SUVR prédits et ceux de référence. La figure IV-4 illustre, à des fins de visualisation, les régions présentant les corrélations les plus fortes ( $r \geq 0,6$ ). Les mesures prédictives et de référence de la structure et du métabolisme cérébral étaient généralement bien corrélées, mais l'intensité de la corrélation variait selon les régions. Pour les volumes cérébraux globaux, les corrélations étaient systématiquement fortes, en particulier pour le LCR ( $r = 0,71$ ) et la substance grise ( $r = 0,64$ ). Dans les ROI associées à la MA, des corrélations modérées à fortes étaient observées dans l'hippocampe, le précuneus, le cortex cingulaire antérieur et le gyrus angulaire, tandis que des associations plus faibles apparaissaient dans les cortex entorhinal et cingulaire postérieur. Dans les ROI associées à la DFT, l'insula et le gyrus temporal moyen présentaient de fortes corrélations, alors que les gyri frontaux supérieur et inférieur n'affichaient que de faibles relations.

Concernant les SUVR, les corrélations étaient globalement plus élevées, avec des valeurs importantes dans le précuneus ( $r = 0,77$ ) et le gyrus angulaire ( $r = 0,80$ ), suggérant une prédiction fiable dans des régions clés de la MA, tandis que des corrélations modérées persistaient dans l'hippocampe et le cortex entorhinal. Pour les SUVR associés à la DFT, de fortes corrélations étaient également observées dans l'insula antérieure et le gyrus temporal moyen, mais seulement modérées dans les régions frontales et temporales inférieures.

La classification du MMSE atteignait une exactitude équilibrée de 0,56, une RMSE de 0,85 et une MAE de 0,54. Dans la plupart des cas, les prédictions de classe MMSE se situaient dans la catégorie correcte ou dans une catégorie adjacente (figure IV-5). De manière importante, aucun cas de démence sévère n'a été mal classé comme absence/doute de démence, et les cas avec absence/doute de démence n'ont été que rarement prédits comme modérés, et jamais comme sévères.



31. Figure IV-4. Analyse de corrélation de Pearson entre les volumes cérébraux prédits et la réalité terrain et les valeurs SUVR.

	Severe dementia	Moderate dementia	Mild dementia	No/Questionable dementia
True MMSE class	7	8	7	0
	Severe dementia	Moderate dementia	Mild dementia	No/Questionable dementia
Severe dementia	5	21	24	1
Moderate dementia	4	6	68	21
Mild dementia	0	1	20	100
No/Questionable dementia				
Predicted MMSE class	Severe dementia	Moderate dementia	Mild dementia	No/Questionable dementia

32. Figure IV-5. Prédiction du MMSE dans les classes de démence.

#### IV.4.3. Comparaison avec les méthodes de référence

Les résultats pour les deux tâches sont rapportés dans les tableaux IV-3 et IV-4. Dans le cadre multiclass, DL-Classifier et PatientGraph ont présenté des performances globalement compétitives, avec une sensibilité légèrement plus faible pour la DFT par rapport aux méthodes de pointe, et une AUC légèrement plus basse avec PatientGraph. De même, leur spécificité pour la MA et les CN était légèrement inférieure à celle du PET-CNN. Toutefois, dans la tâche de classification CN vs DEM, les deux méthodes ont montré des performances comparables à celles du PET-CNN. Comparées à AssemblyNet-AD-FTD, nos méthodes ont surpassé ce dernier sur toutes les autres métriques, à l'exception de la sensibilité et de l'AUC chez les patients DFT.

	<b>DL-Classifier</b>	<b>PatientGraph</b>	<b>PET-CNN</b>	<b>AssemblyNet-AD-FTD</b>
<b>AUC</b>	0.99 [0.98 – 1]	0.98 [0.96 – 0.99]	0.99 [0.98 – 1]	0.96 [0.94 – 0.98]
<b>Sens</b>	0.93 [0.89 – 0.96]	0.92 [0.88 – 0.96]	0.95 [0.92 – 0.98]	0.91 [0.87 – 0.95]
<b>Spec</b>	0.99 [0.97 – 1]	0.99 [0.97 – 1]	0.95 [0.90 – 0.99]	0.87 [0.80 – 0.93]

15. Tableau IV-3. Performances de classification binaire.

Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité.

#### CN vs AD vs FTD

		<b>DL-Classifier</b>	<b>PatientGraph</b>	<b>PET-CNN</b>	<b>AssemblyNet-AD-FTD</b>
<b>CN</b>	<b>AUC</b>	0.99 [0.98–1]	0.98 [0.97–1]	0.99 [0.97–1]	0.96 [0.94–0.99]
	<i>Sens</i>	0.99 [0.97–1]	0.99 [0.97–1]	0.95 [0.90–0.99]	0.87 [0.80–0.93]
	<i>Spec.</i>	0.93 [0.89–0.96]	0.92 [0.88–0.96]	0.95 [0.92–0.98]	0.91 [0.87–0.95]
<b>MA</b>	<b>AUC</b>	0.97 [0.95–0.99]	0.95 [0.92–0.97]	0.97 [0.95–0.99]	0.94 [0.92–0.97]
	<i>Sens</i>	0.91 [0.86–0.95]	0.91 [0.87–0.96]	0.91 [0.86–0.95]	0.82 [0.76–0.88]
	<i>Spec.</i>	0.92 [0.87–0.96]	0.92 [0.87–0.96]	0.93 [0.89–0.97]	0.87 [0.81–0.92]
<b>DFT</b>	<b>AUC</b>	0.95 [0.91–0.99]	0.89 [0.82–0.95]	0.97 [0.93–1]	0.95 [0.91–0.99]
	<i>Sens</i>	0.70 [0.56–0.83]	0.67 [0.54–0.81]	0.78 [0.66–0.90]	0.78 [0.66–0.90]
	<i>Spec.</i>	0.99 [0.99–1]	0.99 [0.99–1]	0.96 [0.93–0.98]	0.94 [0.91–0.97]

16. Tableau IV-4. Performances de classification multiclasse.

Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité.

#### IV.4.4. Performances de classification du diagnostic sur un ensemble de données externe

Les méthodes DL-Classifier et PatientGraph ont obtenu de bonnes performances sur le jeu de données SOCRATES (Tableaux IV-5 et IV-6). Globalement, leurs résultats étaient comparables, avec certaines différences : PatientGraph surpassait légèrement le DL-Classifier dans la prédiction des groupes pathologiques, tandis que le DL-Classifier montrait de meilleures performances pour les cas témoins. Les performances sur le jeu de données externe étaient cohérentes avec celles observées sur le jeu de test interne (Tableaux IV-3 et IV-4). Comparées aux références précédentes, nos méthodes ont maintenu des valeurs élevées d'AUC, de sensibilité et de spécificité, confirmant leur robustesse à travers les différents jeux de données.

---

**CN vs DEM**

	<b>DL-Classifier</b>	<b>PatientGraph</b>
<b>AUC</b>	0.98 [0.94 – 1]	0.98 [0.94 – 1]
<b>Sens</b>	0.93 [0.84 – 1.]	0.98[0.93 – 1.]
<b>Spec</b>	0.95 [0.85 – 0.1.]	0.80 [0.61 – 0.973]

**17. Tableau IV-5. Performances de classification binaire sur le jeu de données SOCRATES.**

Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité.

---

**CN vs AD vs FTD**

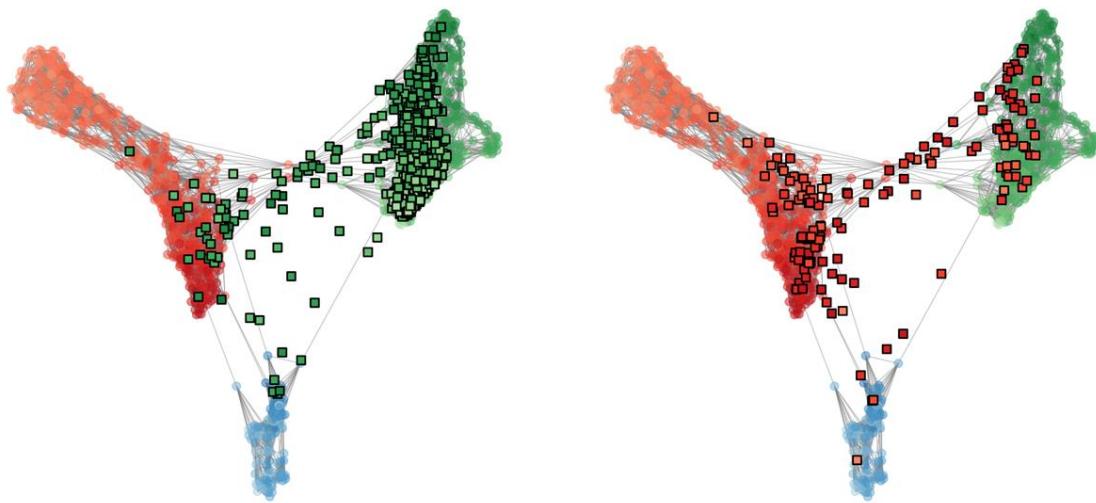
		<b>DL-Classifier</b>	<b>PatientGraph</b>
<b>CN</b>	<i>AUC</i>	0.98 [0.93 – 1]	0.98 [0.93 – 1]
	<i>Sens</i>	0.95 [0.85 – 1]	0.79 [0.60 – 0.97]
	<i>Spec.</i>	0.93 [0.85 – 1]	0.98 [0.93 – 1]
<b>MA</b>	<i>AUC</i>	0.97 [0.95- 0.99]	0.97 [0.91 – 1]
	<i>Sens</i>	0.85 [0.69 – 1]	0.9 [0.77 – 1]
	<i>Spec.</i>	0.97 [0.93 – 1]	0.92 [0.84 – 1]
<b>DFT</b>	<i>AUC</i>	0.95 [0.91 – 0.99]	0.99 [0.95 – 1.]
	<i>Sens</i>	0.90 [0.77 – 1]	0.95 [0.85 – 1]
	<i>Spec.</i>	0.95 [0.88 – 1]	0.92 [0.84 – 1]

**18. Tableau IV-6. Performances de classification multiclasse sur le jeu de données SOCRATES.**

Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité.

## IV.4.5. Analyse de survie des MCI (MEMENTO Dataset)

### IV.4.5.1. Analyse de survie

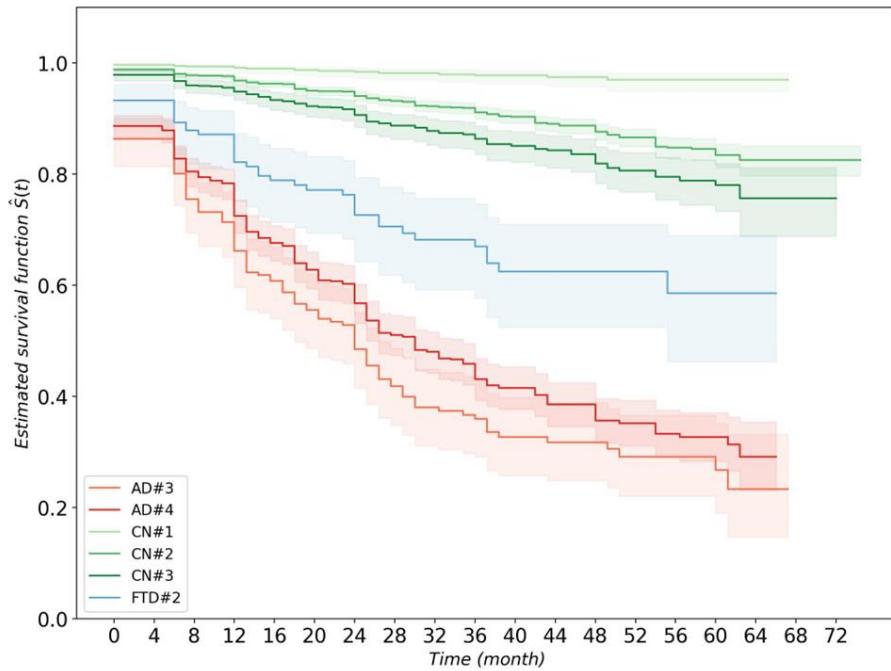


33. Figure IV-6. Projection des MCI dans le PatientSpace.

Pour une meilleure visualisation, les MCI stables (à gauche) ont été séparés des converters (à droite). Les carrés verts représentent les MCI stables, tandis que les carrés rouges représentent les MCI converters. Les cercles correspondent aux sujets du PatientSpace : vert pour CN, rouge pour MA et bleu pour DFT.

Lors de la session initiale, 61 % (86 sur 141) des MCI converters étaient projetés dans des clusters liés à la démence, tandis que 89 % (471 sur 532) des MCI stables étaient projetés dans les clusters CN (Figure IV-6). Fait intéressant, aucun MCI n'était attribué au cluster AD#1 ; un seul converteur était projeté dans AD#2, six dans AD#5, et un dans FTD#1. Ces clusters ont été exclus des analyses de survie ultérieures en raison de leur taille.

L'analyse de survie de la cohorte de MCI (Figure IV-7) a révélé des trajectoires de conversion significativement différentes entre les clusters ( $p < 0,001$ ). Parmi les clusters CN, les MCI initialement projetés dans CN#1 présentaient le taux de progression le plus faible vers la démence, suivis de CN#2, tandis que CN#3 montrait un risque intermédiaire, modéré. En revanche, les clusters liés à la démence présentaient des taux de conversion nettement plus élevés. AD#3 montrait la progression la plus rapide vers la démence, suivi de AD#4, tandis que FTD#2 présentait le taux de conversion le plus bas parmi les clusters liés à la démence.

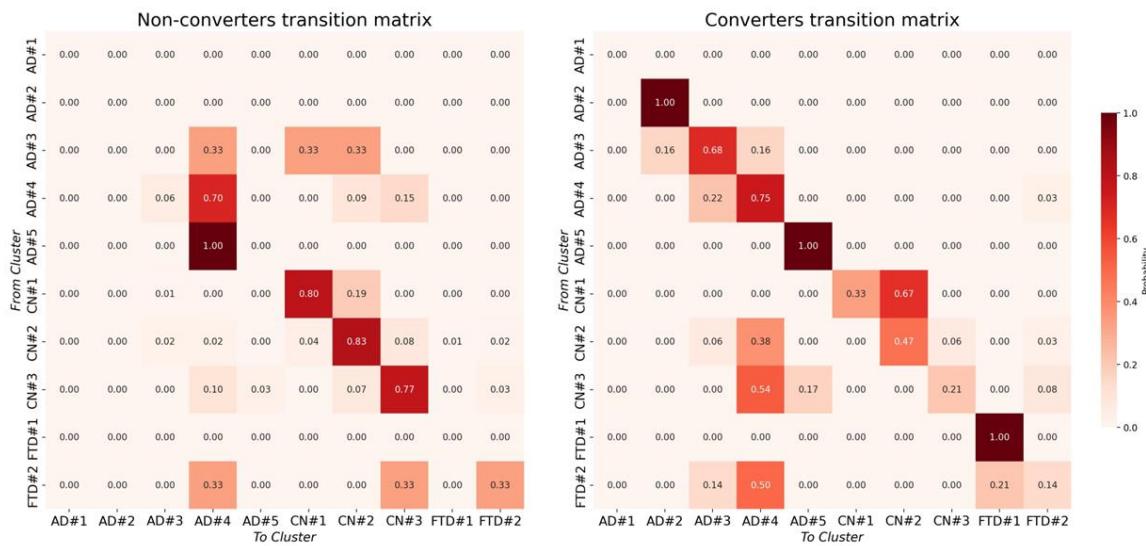


34. Figure IV-7. Analyse de Kaplan-Meier pour la conversion vers la démence chez les sujets MCI.

#### IV.4.5.2. Évolution des MCI converters

La matrice de transition pour les MCI stables montrait qu'ils demeuraient majoritairement au sein des clusters CN (Figure IV-8). Parmi les sujets MCI stables projetés dans AD#3, 66 % revenaient vers un cluster CN, alors que seulement 24 % de ceux du cluster AD#4 montraient une telle réversion. Enfin, les MCI stables associés à FTD#2 présentaient une probabilité de 33 % de réversion vers un cluster CN.

En revanche, la matrice de transition des MCI converters montrait qu'aucun ne revenait des clusters liés à la démence vers les clusters CN (Figure IV-7). De plus, les sujets MCI du cluster CN#1 ne pouvaient que passer au cluster CN#2 avec une probabilité de 67 %, ou rester dans CN#1. Ensuite, ceux projetés dans CN#2 présentaient une probabilité de 47 % de transition vers un cluster lié à la démence, tandis que ceux dans CN#3 avaient une probabilité de 79 %.



35. Figure IV-8. Matrices de transition pour les MCI stables (gauche) et converters (droite).

Les valeurs correspondaient aux probabilités de transition pour les sujets MCI converters et les sujets MCI stables se déplaçant entre les clusters au fil du temps.

## IV.5. Discussion

Nous avons introduit le PatientSpace, un cadre de représentation latente basé sur un graphe, conçu pour décomposer l'hétérogénéité des maladies neurodégénératives à partir de données de neuroimagerie multimodale. Fondé sur un VAE structuré entraîné à partir de l'IRM T1w et la TEP FDG, le PatientSpace permet une analyse interprétable des patients individuellement, via des comparaisons de voisinage dans l'espace latent. Le modèle intègre les diagnostics, l'âge ainsi qu'un nouveau terme de régularisation consistante, structurant l'espace latent pour une meilleure pertinence biologique et clinique. Nos résultats montrent que le PatientSpace capture l'hétérogénéité de la maladie chez les individus MA, DFT et CN, offrant des informations complémentaires sur les sous-types de la maladie, sa sévérité et sa progression.

### IV.5.1. Avancées méthodologiques

Le PatientSpace s'appuie sur les techniques existantes de modélisation en espace latent et les étend en intégrant plusieurs innovations clés.

Au cœur du modèle se trouve un VAE multimodal qui intègre à la fois l'IRM T1w et la TEP FDG, capturant ainsi les signatures anatomiques et métaboliques complémentaires de la neurodégénérescence. Plutôt que de décomposer explicitement l'espace latent en composants partagés et spécifiques à chaque modalité (Lee and Pavlovic, 2021), notre architecture multitâche peut implicitement privilégier une modalité par rapport à l'autre. Si ce choix peut retarder l'apprentissage des caractéristiques provenant de la modalité moins priorisée, il présente l'avantage de capturer le décalage temporel entre les marqueurs de la TEP FDG et de l'IRM T1 de la neurodégénérescence. Sur le plan clinique, il est bien établi que la TEP FDG peut détecter des changements neurodégénératifs plus précocement dans le cours de la maladie, tandis que l'IRM

réflète principalement l'atrophie structurelle et la sévérité de la maladie aux stades avancés (Del Sole et al., 2017).

Cependant, obtenir un espace latent à la fois structuré et désentrelacé ne peut se faire uniquement par apprentissage non supervisé. Des biais inductifs supplémentaires, tels que la supervision, sont nécessaires pour contraindre la représentation latente (Locatello et al., 2019). À cette fin, l'âge et le diagnostic ont été inclus comme signaux de supervision structurant l'espace latent selon des axes cliniquement pertinents. Le diagnostic permet de distinguer les différentes maladies neurodégénératives, tandis que l'âge capture la variabilité naturelle et la progression de la maladie. Ceci est particulièrement important étant donné la distinction clinique entre formes à début précoce et formes à début tardif des maladies (Rossor et al., 2010; Seath et al., 2024; Tellechea et al., 2018)

Néanmoins, se reposer uniquement sur des contraintes supervisées peut limiter l'interprétabilité du modèle et ne pas saisir l'hétérogénéité connue de maladies comme la MA et la DFT, qui présentent des phénotypes de neuroimagerie divers et des variations dépendantes du stade (Dubois et al., 2023; Koenig et al., 2018; Musa et al., 2020; Ossenkoppele et al., 2015b). Pour y remédier, un terme de régularisation consistante a été introduit afin d'améliorer la robustesse et la stabilité des représentations latentes face à la variabilité phénotypique dérivée de l'imagerie. Cette régularisation sert de proxy pour la similarité des neuroimageries, permettant la construction d'un graphe où la proximité reflète la ressemblance biologique.

Parmi tous les composants, la régularisation consistante a eu l'impact le plus important sur les performances du modèle (Tableaux IV-A1 et IV-A2). Cette approche se distingue de la plupart des modèles existants, qui se concentrent uniquement sur la classification ou le clustering non supervisé, sans fournir d'interprétabilité explicite ni de pertinence clinique.

#### **IV.5.2. Performances de la classification diagnostique**

En projetant de nouveaux sujets dans le PatientSpace et en s'appuyant uniquement sur les informations provenant des cas voisins, notre modèle a atteint des performances comparables à celles des méthodes de référence ainsi qu'à la tête de classification du modèle VAE. La généralisation a été validée à l'aide des images SOCRATES, acquises par TEP-IRM sur un site externe. Si les résultats étaient compétitifs pour la classification CN versus démence, le PatientSpace a montré une sensibilité réduite pour différencier MA et DFT. Plusieurs facteurs expliquent probablement cette limitation.

Premièrement, les cas de DFT étaient sous-représentés (13 % des données), limitant la capacité du modèle à capturer l'ensemble des profils liés à la DFT. Selon la littérature (Ghosh and Lippa, 2015), la DFT comprend trois sous-types principaux : le variant comportementale (bvFTD), la forme la plus fréquente, ainsi que le variant sémantique (SV) et le variant aphasic non fluente progressive (PNFA), représentant ensemble environ 45 % de tous les cas de DFT (SD : 20–25 %, PNFA : ~25 %). En revanche, seuls 17 des 106 patients DFT de notre cohorte (16 %) présentaient un profil SV ou PNFA, indiquant une sous-représentation importante de ces sous-types.

Deuxièmement, le jeu de données multicentrique ne contenait pas la représentation complète des diagnostics à chaque site, ce qui peut introduire des biais spécifiques aux sites. Troisièmement, l'architecture de notre modèle n'imposait pas de séparation stricte de l'espace latent entre les

catégories diagnostiques, ce qui a pu préserver des recouvrements cohérents avec l'hétérogénéité clinique observée en pratique.

Pour pallier le déséquilibre de classification, les modèles futurs pourraient utiliser des classifieurs hiérarchiques, séparant d'abord CN et démence, puis distinguant la MA de la DFT. Cette approche en deux étapes a montré son intérêt pour des tâches similaires (Kim et al., 2019). D'autres stratégies possibles incluent la repondération de la perte (Cui et al., 2019) ou le hard-sample mining (Lin et al., 2018). Cependant, les chevauchements dans l'espace latent peuvent également refléter le continuum biologique entre des présentations atypiques de la MA et de la DFT, en particulier pour les formes précoces ou non amnésiques (Harciarek and Jodzio, 2005; Mendez, 2006).

#### **IV.5.3. Sous-types de maladies et interprétation des clusters**

Le clustering agglomératif du graphe latent, construit sur le jeu d'entraînement, a révélé 7 clusters liés à la démence présentant une signification clinique et biologique. Parmi ceux-ci, cinq clusters MA distincts ont émergé. Les clusters AD#1 et AD#2 correspondaient à des EOAD (MA à début précoce), tandis qu'AD#4 et AD#5 représentaient des LOAD (MA à début tardif). AD#3 apparaissait comme un cluster intermédiaire entre EOAD et LOAD. Notamment, AD#1 et AD#2 présentaient les scores MMSE les plus bas, cohérents avec la sévérité plus importante typiquement observée dans l'EOAD (Koedam et al., 2010).

Deux principaux profils d'atrophie ont été identifiés : un profil d'atrophie typique (AD#1 – AD#4) et un profil d'atrophie minimale (AD#5), en accord avec des résultats précédents (Kang et al., 2024; Poulakis et al., 2018). L'hypométabolisme de la TEP-FDG a confirmé ces profils connus (Levin et al., 2021) : AD#1, AD#2, AD#3 et AD#4 présentaient un hypométabolisme typique, tandis qu'AD#5 montrait un hypométabolisme minimal.

Contrairement aux études antérieures ayant identifié des profils atypiques de MA, nos clusters MA présentaient majoritairement une atrophie temporo-pariétale et un hypométabolisme typiques. Plusieurs facteurs peuvent expliquer cette observation. Premièrement, nous avons utilisé uniquement des données de neuroimagerie, alors que les travaux précédents incorporaient souvent des données cognitives, génétiques ou extraites de l'imagerie. De plus, notre approche repose sur des données multimodales, alors que d'autres études se concentraient sur une seule modalité. Deuxièmement, notre modèle se focalise sur l'apprentissage de représentations structurées et non sur l'optimisation du clustering. En outre, nous avons inclus plusieurs populations pathologiques, améliorant la pertinence clinique et reflétant mieux les situations diagnostiques réelles. Troisièmement, l'âge chronologique a été explicitement modélisé plutôt que régressé, ce qui a pu orienter l'espace latent selon la variabilité liée à l'âge. Enfin, l'inclusion de données issues du centre de la mémoire de Lille a introduit un spectre plus large de sévérité et de comorbidités par rapport aux cohortes de recherche comme ADNI.

Au-delà de la MA, deux clusters distincts de DFT ont été identifiés. L'un présentait une atrophie et un hypométabolisme fronto-temporaux typiques de la bvFTD ; l'autre montrait une atteinte prédominante des lobes temporaux et incluait la majorité des cas SV et PNFA. Ces résultats sont cohérents avec les phénotypes de neuroimagerie connus de la DFT (Peet et al., 2021) : la bvFTD affecte principalement les lobes frontaux, tandis que la SV et la PNFA impliquent les régions temporales et insulaires.

Dans l'ensemble, ces résultats soulignent la capacité du PatientSpace à retrouver les phénotypes de neuroimagerie bien établis de la MA et de la DFT à partir de l'imagerie seule, offrant un outil précieux pour soutenir le diagnostic différentiel dans des présentations cliniques complexes ou ambiguës.

#### **IV.5.4. Interprétabilité au niveau du patient**

En analysant les voisinages locaux du graphe, nous avons montré que le PatientSpace permet une inférence des biomarqueurs structurels et métaboliques au niveau du patient. Les valeurs prédites et de référence pour les volumes cérébraux régionaux et les SUVR étaient modérément à fortement corrélées dans les principales régions d'intérêt. Ces résultats indiquent que le PatientSpace capture non seulement les motifs au niveau du groupe, mais permet également une interprétation individualisée, cruciale pour un déploiement clinique. La précision des prédictions était la plus élevée pour les volumes globaux et les régions temporo-pariétales, tandis que des corrélations plus faibles dans les régions frontales et limbiques suggèrent certaines limites de robustesse du modèle.

Par ailleurs, la prédiction du stade MMSE avec le PatientSpace a atteint des performances acceptables, la majorité des cas étant classés dans le niveau de sévérité correct ou adjacent. L'absence de classifications grossières (par exemple, une démence sévère prédite comme CN) renforce encore la fiabilité de l'approche.

#### **IV.5.5. Trajectoires longitudinales des sujets MCI**

La projection des sujets MCI de la cohorte externe MEMENTO dans le PatientSpace a révélé de fortes associations entre l'appartenance aux clusters et les taux de conversion vers la démence.

Notamment, aucun MCI n'a été projeté dans le cluster AD#1, un seul dans AD#2, six dans AD#5 et un dans FTD#1. Cette distribution reflète l'effet de l'âge et de la sévérité de la maladie : les clusters AD#1 et AD#2 sont principalement composés d'individus plus jeunes présentant une démence modérée à sévère, les plaçant en dehors de l'âge et du stade typiques des patients MCI. De même, le faible nombre de MCI projetés dans FTD#1 provient probablement à la fois de différences d'âge et de discordances neuroimageries – FTD#1 présente un schéma caractéristique d'atrophie et d'hypométabolisme bvFTD, tandis que les MCI de MEMENTO présentent principalement des profils prodromaux de la MA.

Les MCI projetés dans les clusters liés à la démence avaient tendance à progresser plus rapidement, tandis que ceux mappés aux clusters CN restaient généralement stables. Ces taux de conversion étaient cohérents avec les profils de neuroimagerie sous-jacents d'atrophie et d'hypométabolisme caractéristiques de chaque cluster. Par exemple, le cluster FTD#2, qui présentait le moins d'atrophie et d'hypométabolisme parmi les clusters liés à la démence, avait également le taux de conversion le plus faible. Inversement, bien que AD#3 et AD#4 présentent des niveaux similaires d'atrophie et d'hypométabolisme, AD#3 était associé à une progression plus rapide, probablement en raison de sa proportion plus élevée de patients EOAD issus du CHU de Lille, cohérente avec les rapports de déclin plus rapide chez l'EOAD (Seath et al., 2024; Tellechea et al., 2018).

L'analyse des matrices de transition a également montré que l'appartenance aux clusters était stable dans le temps, la majorité des patients suivant des trajectoires monotones des clusters CN vers les clusters liés à la démence. Aucune réversion n'a été observée, reflétant la nature progressive de la neurodégénérescence. Ces résultats sont en accord avec des travaux précédents (Misra et al., 2009;

Pagani et al., 2017), confirmant que les MCI converters présentent souvent déjà des modifications cérébrales détectables en termes d'atrophie ou de métabolisme.

Dans l'ensemble, ces résultats soutiennent l'utilisation du PatientSpace pour le staging et le suivi de la progression de la maladie dans les populations prodromales.

#### **IV.5.6. Limites et perspectives**

Notre approche requiert actuellement des données multimodales complètes pour chaque sujet, limitant son applicabilité dans des contextes réels où les données manquantes sont fréquentes. Les travaux futurs devraient adapter le PatientSpace pour gérer les modalités partiellement manquantes, par exemple via l'imputation de modalités ou le modality dropout (Wu et al., 2024).

Malheureusement, aucun centre unique n'a fourni des sujets couvrant les trois catégories diagnostiques, ce qui aurait permis de minimiser le risque d'apprentissage de biais spécifiques aux sites. Par exemple, le CHU de Lille n'incluait aucun sujet CN (Tableau IV-S1), car notre base de données, construite à partir de données cliniques, ne contient pas de témoins sains. Cela reflète un défi plus large dans l'application de l'IA aux données cliniques : les témoins sains sont généralement sous-représentés ou absents dans les cohortes hospitalières. Dans ce contexte, il est important de noter que l'augmentation des données, bien que nécessaire pour accroître la diversité des exemples d'entraînement, peut théoriquement accentuer les biais inter-classes (Balestrieri et al., 2022).

L'utilisation de l'IRM T1w et de la TEP-FDG comme principaux biomarqueurs peut être remise en question, d'autres techniques d'imagerie et biomarqueurs pouvant également aider à distinguer les maladies neurodégénératives. Par exemple, l'imagerie par tenseur de diffusion (DTI) et l'IRM fonctionnelle en repos ont montré des résultats prometteurs pour différencier ces pathologies (Agosta et al., 2017; Goveas et al., 2015; Young et al., 2020). L'imagerie TEP amyloïde et tau a également montré de bonnes performances diagnostiques (Groot et al., 2024; Ioannou et al., 2025; Rabinovici et al., 2011; Vandenbergh et al., 2013). Les analyses de LCR et les biomarqueurs sanguins pourraient aussi être envisagés. Cependant, le prélèvement du LCR nécessite une ponction lombaire invasive, les biomarqueurs sanguins ne sont pas encore largement utilisés en pratique clinique, la TEP amyloïde coûte environ deux fois plus cher que la TEP-FDG (Contador et al., 2023; Teunissen et al., 2022), et la TEP tau montre une sensibilité limitée aux stades précoces, avec de nombreux sujets MCI ou certains patients MA amyloïde-positifs présentant des scans négatifs (Ossenkoppele and Hansson, 2021). Enfin, l'IRM T1w et la TEP-FDG restent des outils de référence standard pour le diagnostic de démence (Young et al., 2020). La TEP-FDG a pris de l'importance dans les diagnostics cliniques récents (Chételat et al., 2020), et l'IRM T1w est la séquence la plus couramment utilisée dans les essais cliniques pour suspicion de neurodégénérescence et constitue la référence pour le staging de la maladie (Del Sole et al., 2017; Young et al., 2020).

Notre étude s'est concentrée exclusivement sur la MA et la DFT. Étendre le modèle à d'autres syndromes neurodégénératifs, tels que la maladie à corps de Lewy ou les troubles parkinsoniens, pourrait améliorer son utilité en tant qu'outil de biomarqueurs de neuroimagerie à usage général.

#### **IV.6. Conclusion**

En conclusion, le PatientSpace offre un espace latent interprétable et cliniquement pertinent pour l'analyse des données de neuroimagerie multimodale dans les maladies neurodégénératives. En

combinant l'apprentissage profond et l'analyse basée sur les graphes, il permet à la fois la découverte de sous-types au niveau de la population et l'interprétabilité au niveau du patient, avec des applications potentielles en diagnostic, pronostic et suivi de la progression de la maladie.

## IV.7. Appendix

### IV.7. A1 : Résultats complémentaires sur les variantes du PatientSpace

CN vs DEM		IRM+TEP	IRM	TEP	Consistency ablated	Age ablated	Vanilla
Classifier DL	AUC	0.99 [0.98 – 1.00]	0.96 [0.94 – 0.98]	0.99 [0.98 – 1.00]	0.98 [0.97 – 0.99]	0.99 [0.98 – 1.00]	0.99 [0.98 – 1.00]
	Sens	0.93 [0.89 – 0.96]	0.85 [0.80 – 0.90]	0.95 [0.93 – 0.98]	0.89 [0.84 – 0.93]	0.92 [0.88 – 0.96]	0.96 [0.93 – 0.99]
	Spec.	0.99 [0.97 – 1.00]	0.95 [0.90 – 0.99]	0.97 [0.94 – 1.00]	0.99 [0.97 – 1.00]	0.97 [0.94 – 1.00]	0.97 [0.94 – 1.00]
PatientGraph	AUC	0.98 [0.96 – 0.99]	0.96 [0.94 – 0.98]	0.98 [0.97 – 1.00]	0.88 [0.85 – 0.92]	0.96 [0.94 – 0.98]	0.95 [0.93 – 0.98]
	Sens	0.92 [0.88 – 0.96]	0.85 [0.80 – 0.90]	0.93 [0.89 – 0.97]	0.90 [0.86 – 0.94]	0.89 [0.85 – 0.94]	0.85 [0.80 – 0.90]
	Spec.	0.99 [0.97 – 1.00]	0.95 [0.90 – 0.99]	0.98 [0.95 – 1.00]	0.76 [0.68 – 0.85]	0.93 [0.88 – 0.98]	0.91 [0.85 – 0.97]

19. Tableau IV-A1. Performances de classification binaire.

Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité

*CN vs AD vs DFT*

		Diagnostique	IRM+TEP	IRM	TEP	Consistency ablated	Age ablated	Vanilla
<b>DL-Classifier</b>	CN	AUC	0.99 [0.98–1.00]	0.96 [0.93–0.99]	0.99 [0.97–1.00]	0.98 [0.96–1.00]	0.99 [0.98–1.00]	0.99 [0.98–1.00]
		Sens	0.99 [0.97–1.00]	0.95 [0.90–0.99]	0.97 [0.94–1.00]	0.99 [0.97–1.00]	0.97 [0.94–1.00]	0.97 [0.94–1.00]
		Spec	0.93 [0.89–0.96]	0.85 [0.80–0.90]	0.95 [0.93–0.98]	0.89 [0.84–0.93]	0.92 [0.88–0.96]	0.96 [0.93–0.99]
	MA	AUC	0.97 [0.95–0.99]	0.90 [0.86–0.94]	0.95 [0.93–0.98]	0.95 [0.92–0.97]	0.95 [0.93–0.98]	0.97 [0.95–0.99]
		Sens	0.91 [0.86–0.95]	0.77 [0.71–0.84]	0.90 [0.85–0.95]	0.79 [0.72–0.85]	0.86 [0.80–0.92]	0.94 [0.90–0.98]
		Spec	0.92 [0.87–0.96]	0.90 [0.85–0.95]	0.89 [0.84–0.94]	0.93 [0.89–0.97]	0.91 [0.86–0.96]	0.90 [0.85–0.95]
	DFT	AUC	0.95 [0.91–0.99]	0.88 [0.82–0.95]	0.92 [0.86–0.97]	0.91 [0.85–0.97]	0.92 [0.87–0.98]	0.96 [0.92–1.00]
		Sens	0.70 [0.56–0.83]	0.63 [0.49–0.77]	0.65 [0.52–0.79]	0.74 [0.61–0.87]	0.61 [0.47–0.75]	0.65 [0.52–0.79]
		Spec	0.99 [0.99–1.00]	0.95 [0.93–0.98]	0.96 [0.94–0.99]	0.95 [0.92–0.98]	0.95 [0.92–0.98]	1.00 [1.00–1.00]
<b>PatientGraph</b>	CN	AUC	0.98 [0.97–1.00]	0.94 [0.91–0.98]	0.98 [0.96–1.00]	0.88 [0.84–0.93]	0.96 [0.94–0.99]	0.95 [0.92–0.98]
		Sens	0.99 [0.97–1.00]	0.88 [0.81–0.94]	0.98 [0.95–1.00]	0.76 [0.68–0.85]	0.93 [0.88–0.98]	0.91 [0.85–0.97]
		Spec	0.92 [0.88–0.96]	0.88 [0.83–0.92]	0.93 [0.89–0.97]	0.90 [0.86–0.94]	0.89 [0.85–0.94]	0.85 [0.80–0.90]
	MA	AUC	0.95 [0.92–0.97]	0.87 [0.83–0.91]	0.95 [0.92–0.97]	0.87 [0.83–0.91]	0.94 [0.91–0.97]	0.89 [0.86–0.93]
		Sens	0.91 [0.87–0.96]	0.83 [0.77–0.89]	0.92 [0.88–0.96]	0.89 [0.84–0.94]	0.89 [0.84–0.94]	0.83 [0.77–0.89]

	Spec	0.92 [0.87–0.96]	0.80 [0.74–0.87]	0.89 [0.84–0.94]	0.77 [0.70–0.84]	0.92 [0.87–0.96]	0.77 [0.70–0.84]
DFT	AUC	0.89 [0.82–0.95]	0.79 [0.71–0.87]	0.88 [0.82–0.95]	0.88 [0.82–0.95]	0.89 [0.82–0.95]	0.88 [0.82–0.95]
	Sens	0.67 [0.54–0.81]	0.48 [0.33–0.62]	0.61 [0.47–0.75]	0.54 [0.40–0.69]	0.61 [0.47–0.75]	0.37 [0.23–0.51]
	Spec	0.99 [0.99–1.00]	0.96 [0.94–0.98]	0.99 [0.98–1.00]	0.97 [0.95–0.99]	0.97 [0.95–0.99]	1.00 [1.00–1.00]

20. Tableau IV-A2. Performances de classification multiclasses.

Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité.

## **IV.8. Annexes**

### **IV.8. S.1. Acquisition des données**

#### **IV.8. S.1.1. TEP-FDG**

Les scans TEP-FDG issus de la base de données du Centre Hospitalier Universitaire de Lille ont été acquis sur un PET/CT hybride Biograph mCT-Flow à 4 anneaux (Siemens) avec un scanner CT de 20 coupes et des cristaux de lutétium oxyorthosilicate de  $4 \times 4 \text{ mm}^2$ . La dose moyenne de FDG était de 177 MBq (SD = 19 MBq). Après 30 minutes, un CT à faible dose a été réalisé pour la correction d'atténuation, suivi d'une acquisition PET de 10 minutes. Les données ont été reconstruites à l'aide d'un algorithme OSEM (8 itérations, 21 sous-ensembles) avec corrections de décroissance, aléatoires et de diffusion, et un lissage gaussien de 2 mm. Les images finales consistaient en 109 coupes axiales (FOV =  $408 \times 408 \times 221,3 \text{ mm}^3$ , matrice =  $400 \times 400 \times 109$ , taille de voxel =  $1,02 \times 1,02 \times 2,03 \text{ mm}^3$ ).

Pour l'acquisition d'ADNI, le protocole peut être trouvé à l'adresse suivante [https://adni.loni.usc.edu/wp-content/uploads/2010/05/ADNI2\\_PET\\_Tech\\_Manual\\_0142011.pdf](https://adni.loni.usc.edu/wp-content/uploads/2010/05/ADNI2_PET_Tech_Manual_0142011.pdf). La dose de traceur FDG était de 185 MBq ( $\pm 10\%$ ), et entre 30 et 60 minutes après l'injection, six acquisitions dynamiques 3D de 5 minutes chacune ont été réalisées. Un scan CT à faible dose a été acquis pour la correction d'atténuation, ou, pour les scanners TEP seuls, un scan de correction d'atténuation a été réalisé à l'aide de sources en tige.

Tous les scans FTLDNI ont été acquis au centre Mayo Clinic sur un scanner PET/CT GE Discovery RX. Les participants ont reçu une injection de 185 MBq ( $\pm 10\%$ ) de FDG, et l'acquisition a commencé 30 minutes plus tard, consistant en six frames dynamiques de 5 minutes. Un scan CT, obtenu avant l'injection de FDG, a été utilisé pour la correction d'atténuation, et la reconstruction a été réalisée selon la technique de rétroprojection filtrée 3D.

Toutes les images ont été examinées par un expert (Antoine Rogeau) pour le contrôle de la qualité visuelle.

#### **IV.8. S.1.2. IRM pondérée T1**

Les scans IRM T1w du Centre Hospitalier Universitaire de Lille ont été acquis sur un scanner 3T (Achieva Philips, Best, Pays-Bas) avec les paramètres suivants : TR/TE = 9,9 ms / 4,6 ms, angle de bascule =  $8^\circ$ , taille de voxel =  $0,94 \times 0,94 \times 1,2 \text{ mm}^3$ .

Pour l'acquisition de l'ADNI, le protocole de chaque IRM peut être trouvé à l'adresse suivante : <https://adni.loni.usc.edu/data-samples/adni-data/neuroimaging/mri/mri-scanner-protocols/>.

Les IRM T1w ont été acquises sur des scanners General Electric Healthcare, Philips Medical Systems et Siemens Medical Solutions à 3T pour les protocoles ADNI2/GO/3 et à 1,5T pour le protocole ADNI1.

Toutes les images ont été examinées par un expert (Grégory Kuchcinski) pour le contrôle de la qualité visuelle.

#### **IV.8. S.1.3. Différences entre les sites**

Le tableau IV-S1 présente la répartition des diagnostics entre les différents sites.

	CN	MA	DFT
ADNI	463	379	0
NIFD	59	0	87
CHU de Lille	0	430	104

21. Tableau IV-S1. Répartition des diagnostics entre les sites.

## IV.8. S.2. Ensemble des données externes

### IV.8. S.2.1. Jeu de données MEMENTO

Le jeu de données MEMENTO (<https://portal.dementiasplatform.uk/>) a été utilisé pour évaluer la conversion des MCI et l'évolution de la maladie au sein du PatientSpace. Les critères d'inclusion étaient les suivants : être âgé de 18 ans ou plus ; présenter au moins un léger déficit cognitif, défini comme un score inférieur d'une déviation standard à la moyenne (par rapport aux normes d'âge et d'éducation) dans un ou plusieurs domaines cognitifs (cette déviation devant être identifiée dans les six mois précédent la date d'inclusion) OU avoir une plainte cognitive isolée, quelle qu'en soit la durée, si l'âge est de 60 ans ou plus (sans déficit cognitif tel que défini ci-dessus) ; CDR  $\leq 0,5$  et absence de démence ; acuité visuelle et auditive suffisante pour passer les tests neuropsychologiques ; avoir signé un consentement éclairé ; et être affilié à un régime d'assurance maladie.

Il est à noter qu'aucun des participants issus de la cohorte MCI n'a été inclus dans l'entraînement du modèle.

#### Jeu de données MEMENTO

	Ligne de base
Diagnostic	MCI
N balayages	673
Âge, années	70.24 $\pm$ 9.10
Femmes, %	381 (57%)
MMSE	28 (27 - 29)
Converters, %	141 (21%)

22. Tableau IV-S2. Caractéristiques démographiques de l'ensemble de données MEMENTO à l'inclusion.

#### IV.8. S.2.2. Jeu de données SOCRATES

Le jeu de données SOCRATES a été utilisé pour évaluer la généralisation de nos méthodes. Tous les sujets ont été recrutés dans un centre expert universitaire tertiaire de mémoire (Institut de la Mémoire et de la Maladie d'Alzheimer – Département de Neurologie, Hôpital Universitaire Pitié-Salpêtrière) à partir de la cohorte SOCRATES, visant à suivre un groupe de patients diagnostiqués avec la maladie d'Alzheimer (MA) ou des troubles apparentés, âgés de 40 à 80 ans et présentant un score au Mini-Mental State Examination (MMSE) supérieur à 10 au moment de l'inclusion.

Les critères d'exclusion comprenaient la démence vasculaire (selon les critères NINDS-AIREN) ou la démence mixte (avec un score de Hachinski > 4), toute pathologie neurologique concomitante significative, un cancer actif, des antécédents de cancer cérébral, une maladie métastatique cérébrale ou une irradiation cérébrale antérieure.

Le diagnostic de la MA et des troubles apparentés a été établi par une évaluation multidisciplinaire incluant un examen neurologique, un bilan neuropsychologique complet, ainsi que des analyses biologiques, génétiques et des neuroimageries.

Les patients ont passé l'examen PET-MR entre 2015 et 2021. Toutes les données PET et IRM ont été acquises simultanément à l'aide d'un système intégré PET/IRM (3T SIGNA PET/MRI, GE Healthcare, Milwaukee, États-Unis) 30 minutes après l'injection intraveineuse de 2 MBq/kg de 18F-FDG. Les images PET 18F-FDG ont été acquises pendant 15 minutes. Les images PET ont été acquises et reconstruites de manière itérative en utilisant l'algorithme OSEM (Ordered Subsets Expectation Maximization) avec 4 itérations et 28 sous-ensembles, et corrigées des effets de volume partiel. La résolution finale des images était de  $1,17 \times 1,17 \times 2,78 \text{ mm}^3$ . Pour la correction de l'atténuation des photons, une méthode basée sur un atlas a été mise en œuvre dans le système PET/IRM, en utilisant soit l'approche VPFXS, soit QCFX. L'acquisition IRM comprenait au moins des images IRM 3D pondérées en T1.

	CN	MA	DFT
N balayages	19	20	20
Âge, années	$65.45 \pm 9.16$	$69,42 \pm 7,66$	$65.14 \pm 9.78$
Femmes, %	50%	45%	45%
MMSE	-	20 (15 – 23)	24.5 (21 – 27)

23. Tableau IV-S3. Caractéristiques démographiques de l'ensemble de données SOCRATES.

#### IV.8. S.3. Détails supplémentaires du modèle

##### IV.8. S.3.1. Entrainement

Au cours des deux premières époques, les poids  $w_t$  ont été fixés à 1.

Les autres poids de régularisation ont été fixés comme suit  $\lambda_1 = 0,01, \lambda_2 = 300, \lambda_3 = 1$ .

Pour  $\lambda_4$ , son impact a été progressivement augmenté de 0,2 à 2 au cours des dix premières époques. Nous avons utilisé l'optimiseur AdamW (Loshchilov et Hutter, 2019) avec un taux d'apprentissage de  $10^{-4}$  et un coefficient de régularisation (weight decay) de  $10^{-5}$ .

#### IV.8. S.3.2. Nombre optimal de voisins

Le nombre optimal de voisins K a été déterminé en projetant les sujets de validation dans PatientSpace et en sélectionnant la valeur de K qui maximisait la précision équilibrée de la classification et l'AUC moyenne macro.

#### IV.8. S.4. Organisation du PatientSpace

<i>ROI</i>	<i>Corrélation de Pearson r</i>	<i>IC-95 %</i>	<i>Force de corrélation</i>
<b>Volumes globaux du cerveau</b>			
GM	<b>0.64</b>	[0.57 - 0.70]	Fort
WM	<u>0.56</u>	[0.48 - 0.70]	Modéré à Fort
LCR	<b>0.71</b>	[0.64 - 0.76]	Fort
Temporal Lobe	<b>0.64</b>	[0.56 - 0.70]	Fort
Frontal Lobe	<b>0.61</b>	[0.53 - 0.68]	Fort
<b>Volumes de retour sur investissement AD</b>			
Hippocampe	<u>0.54</u>	[0.45 - 0.61]	Modéré à Fort
Précuneus	<u>0.53</u>	[0.45 - 0.61]	Modéré à Fort
Gyrus parahippocampique	0.41	[0.31 - 0.50]	Modéré
Cortex entorhinal	0.38	[0.27 - 0.47]	Faible - Modéré
Gyrus cingulaire postérieur	0.37	[0.27 - 0.47]	Faible - Modéré
Gyrus cingulaire antérieur	<u>0.53</u>	[0.44 - 0.64]	Modéré à Fort
Gyrus angulaire	<b>0.60</b>	[0.52 - 0.67]	Fort
Gyrus frontal moyen	0.51	[0.42 - 0.59]	Modéré
<b>Volumes de retours sur investissement DFT</b>			
Cortex insulaire	<u>0.57</u>	[0.48 - 0.64]	Modéré à Fort
Insula antérieure	<b>0.60</b>	[0.52 - 0.67]	Fort
Gyrus frontal supérieur	0.28	[0.17 - 0.38]	Faible
Gyrus frontal inf.	0.26	[0.15 - 0.37]	Faible
Pôle temporel	<u>0.51</u>	[0.48 - 0.60]	Modéré à Fort
Gyrus temporal moyen	<b>0.70</b>	[0.63 - 0.75]	Fort
Gyrus temporal sup.	0.44	[0.35 - 0.53]	Modéré
Inf. gyrus temporal	0.44	[0.37 - 0.55]	Modéré
<b>AD ROIs SUVR</b>			
Hippocampe	0.46	[0.36 - 0.54]	Modéré
Précuneus	<b>0.77</b>	[0.72 - 0.81]	Fort
Gyrus parahippocampique	<u>0.51</u>	[0.43 - 0.60]	Modéré à Fort
Cortex entorhinal	0.49	[0.40 - 0.58]	Modéré
Gyrus cingulaire postérieur	<b>0.66</b>	[0.59 - 0.72]	Fort
Gyrus cingulaire antérieur	<b>0.69</b>	[0.63 - 0.75]	Fort
Gyrus angulaire	<b>0.80</b>	[0.75 - 0.83]	Fort

Gyrus frontal moyen	<b>0.52</b>	[0.43 - 0.60]	Modéré à Fort
<b>VUS DFT ROIs</b>			
Insula antérieure	<b>0.62</b>	[0.55 - 0.69]	Fort
Gyrus frontal supérieur	0.43	[0.33 - 0.52]	Modéré
Gyrus frontal inf.	0.46	[0.36 - 0.54]	Modéré
Pôle temporel	0.49	[0.39 - 0.57]	Modéré
Gyrus temporal moyen	<b>0.61</b>	[0.54 - 0.68]	Fort
Gyrus temporal sup.	<b>0.53</b>	[0.45 - 0.61]	Modéré à Fort
Inf. gyrus temporal	0.44	[0.34 - 0.53]	Modéré

24. Tableau IV-S4. Corrélation de Pearson entre les ROIs des MA et des DFT.

Les corrélations les plus fortes sont en gras. Modéré - Les corrélations modérées sont soulignées.

#### IV.8. S.5. Voxel based Morphometry

Afin d'analyser les profils d'atrophie entre les différents clusters, les IRM structurelles ont été prétraitées à l'aide du logiciel SPM (Statistical Parametric Mapping, <https://www.fil.ion.ucl.ac.uk/spm/>). Dans un premier temps, les volumes ont été segmentés en matière grise, matière blanche et LCR. Les images de matière grise ont ensuite été registrées dans l'espace MNI, puis lissées à l'aide d'un noyau gaussien de 8 mm<sup>3</sup>.

L'analyse statistique a été menée à l'aide de tests t bilatéraux à deux échantillons, en intégrant comme covariables l'âge, le volume intracrânien total, ainsi que le sexe des participants. Ces comparaisons ont permis d'opposer les clusters de patients à des groupes témoins appariés selon l'âge et le sexe. Les résultats ont été corrigés pour comparaisons multiples par le Family-Wise Error rate (FWE), avec un seuil de significativité fixé à 0,05.

# Chapitre V

## Prédiction du pronostic fonctionnel après AVC ischémique dans les Hauts-de-France grâce à une approche vision-langage

---

Dans le chapitre IV, nous avons développé le PatientSpace dans une approche multimodale, combinant IRM T1w et TEP-FDG pour distinguer plusieurs diagnostics (CN, DFT et Alzheimer). Cet espace latent structuré et interprétable organise les patients selon leur similarité multimodale (IRM et TEP), permettant d'explorer la variabilité interindividuelle tout en facilitant l'explicabilité clinique. Néanmoins, le PatientSpace présente certaines limites : il repose uniquement sur des modalités similaires (imagerie) et ne permet pas d'identifier quelle modalité contribue à la prédiction diagnostique, ni de préciser la manière dont chacune influe sur cette prédiction. Le chapitre V aborde cette question de l'interprétabilité multimodale à travers une étude préliminaire intégrant quatre modalités hétérogènes — IRM FLAIR, IRM de diffusion (DWI), comptes rendus radiologiques et données cliniques — dans le cadre de la prédiction du pronostic fonctionnel après accident vasculaire cérébral (AVC), à l'aide d'une approche vision-langage. Ce travail constitue une première étape exploratoire vers un projet plus ambitieux, ce qui justifie que les résultats et la discussion ne soient présentés ici qu'à un niveau préliminaire.

## V.1.Introduction

Les accidents vasculaires cérébraux (AVC) constituent l'une des principales causes de mortalité et de handicap à long terme dans le monde (Katan and Luft, 2018). On distingue deux grands types d'AVC : les AVC ischémiques, qui représentent environ 85 % des cas, et les AVC hémorragiques, qui en représentent environ 15 % (Campbell et al., 2019). Le traitement aigu des AVC ischémiques repose principalement sur deux stratégies de revascularisation : la thrombolyse intraveineuse et la thrombectomie mécanique, qui améliorent significativement la survie et l'autonomie fonctionnelle des patients (Powers et al., 2018). Cependant, malgré ces avancées, un grand nombre de patients ne parviennent pas à retrouver une qualité de vie similaire à celle précédant l'AVC. Identifier précocement les patients susceptibles de présenter un bon ou un mauvais devenir fonctionnel constitue donc un enjeu majeur pour la médecine de précision appliquée aux AVC.

L'évaluation du devenir fonctionnel repose classiquement sur le *modified Rankin Scale* (mRS), une échelle ordinaire en sept niveaux allant de l'absence de symptôme (0) au décès (6) (Broderick et al., 2017). Dans les études cliniques et en recherche, il est fréquent de dichotomiser ce score afin de distinguer un bon rétablissement ( $mRS \leq 2$ ) d'un mauvais rétablissement ( $mRS > 2$ ), généralement à trois mois après l'AVC (Weisscher et al., 2008). Cette approche standardisée a favorisé le développement de nombreuses méthodes d'IA.

De nombreux travaux ont ainsi exploré la prédition du mRS à 3 mois à partir de différents types de données. Les facteurs influençant le rétablissement sont multiples : la sévérité initiale de l'AVC, sa localisation, mais aussi des variables cliniques telles que l'âge, le sexe, le score NIHSS (National Institutes of Health Stroke Scale) gradant la sévérité des symptômes de 0 à 42 (Chalos et al., 2020), ou encore des comorbidités comme le tabagisme et l'alcoolisme. Des approches d'apprentissage automatique ont montré des performances prometteuses. Par exemple, Liu et al. (Liu et al., 2023) ont combiné données cliniques et d'imagerie acquises entre 1 et 7 jours post-AVC dans une approche d'ensemble pour prédire le mRS à 3 mois sous forme de régression. Borsos et al (Borsos et al., 2024) ont proposé un modèle de fusion injectant les variables cliniques directement dans le pipeline de traitement de l'imagerie de perfusion. D'autres travaux, enfin, ont montré qu'il était possible de se baser uniquement sur l'imagerie post-AVC pour segmenter la lésion et prédire le pronostic fonctionnel (Nishi et al., 2020).

Cependant, la majorité des algorithmes pronostiques existants exploitent des données collectées pendant la phase aiguë de l'AVC (entre 1 et 7 jours après l'événement), notamment à partir d'imageries de suivi. Si ces informations permettent de mieux caractériser l'étendue finale de la lésion et d'étudier les mécanismes de récupération, elles arrivent trop tard pour influencer la stratégie thérapeutique initiale. À l'inverse, la phase hyperaiguë (les premières heures suivant l'AVC) représente une fenêtre critique où les décisions de revascularisation (thrombolyse intraveineuse, thrombectomie mécanique) doivent être prises rapidement (Jiang et al., 2025). Dans ce contexte, la mise à disposition d'outils pronostiques précoce est essentielle pour guider la prise en charge et anticiper l'évolution fonctionnelle.

Certaines modalités d'imagerie jouent un rôle clé à ce stade. L'IRM de diffusion (DWI) constitue la séquence la plus sensible pour détecter précocement une lésion ischémique, et reste aujourd'hui le gold standard du diagnostic positif. L'IRM FLAIR (Fluid Attenuated Inversion Recovery) complète cette information en fournissant des indices sur l'ancienneté des lésions, notamment via l'approche

du mismatch DWI–FLAIR utilisée pour sélectionner les patients éligibles à une thrombolyse lorsque l'heure de début des symptômes est inconnue (G et al., 2011). Elle permet également d'évaluer l'état de santé cérébrale global, en particulier l'atrophie, qui peut influencer le devenir fonctionnel.

Les comptes rendus médicaux de neuroimagerie apportent enfin une dimension interprétative précieuse. Ils synthétisent l'expertise du neuroradiologue en décrivant des éléments cliniquement déterminants tels que la présence d'une occlusion de gros vaisseau, l'existence de microsaignements, la trophicité cérébrale ou encore la topographie exacte des lésions. Ces données « enrichies par l'expert » représentent une source d'information complémentaire aux images brutes.

L'objectif de cette étude est donc d'exploiter conjointement, par un modèle vision-langage, plusieurs sources d'information disponibles à la phase aigüe — l'imagerie FLAIR et DWI, les comptes rendus médicaux issus des examens de neuroimagerie, ainsi que les données cliniques — afin de construire un espace latent multimodal interprétable pour la prédiction du mRS à 3 mois.

## V.2. Matériels et méthodes

### V.2.1. Ethique

L'étude a été approuvée par le Comité de protection des personnes Nord-Ouest IV, qui l'a classée comme observationnelle le 9 mars 2010. Le Comité de protection des informations personnelles des patients a validé le protocole le 21 décembre 2010 (référence n° 10.677). Par ailleurs, le service Délégué à la protection des données (DPO) du CHU de Lille a attesté de la conformité des modalités de mise en œuvre du projet avec la réglementation applicable en matière de protection des données personnelles, notamment le Règlement général sur la protection des données (RGPD, UE 2016/679).

### V.2.2. Participants

Tous les patients admis pour une thrombectomie mécanique consécutive à un AVC ischémique entre janvier 2015 et décembre 2021 ont été inclus dans l'étude. Les patients ont été pris en charge conformément aux recommandations internationales en vigueur au moment de leur admission (Powers et al., 2018).

### V.2.3. Prétraitement des données

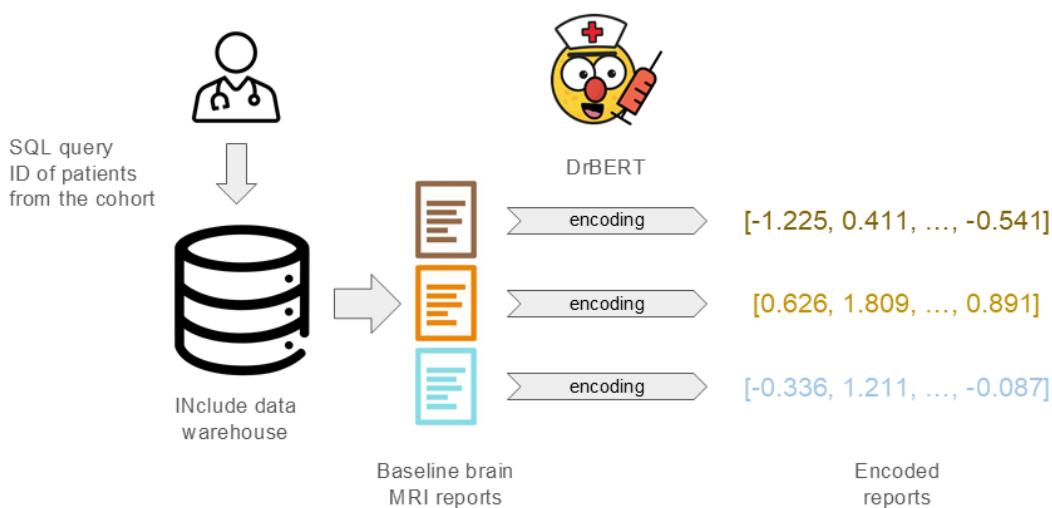
#### V.2.3.1. Données cliniques

En plus du score mRS 3 mois après l'AVC, des variables cliniques à la phase aigüe de l'AVC, c'est-à-dire avant le traitement par thrombectomie, ont été collectées : l'âge et le score NIHSS comme variables continues ; le sexe, l'hypertension, le diabète, l'hypercholestérolémie, le tabagisme et la consommation d'alcool comme variables catégorielles ; et le score mRS comme variable ordinaire. Ces variables ont été normalisées et encodées afin d'être directement exploitables par les modèles. Les variables continues ont été centrées et réduites à l'aide d'un *z-score*. Les variables catégorielles ont été transformées par un encodage *one-hot*. La variable ordinaire a été encodée de façon croissante pour respecter son caractère graduel. L'ensemble de ces variables avant thrombectomie a été concaténé en un vecteur clinique unique. Tous les prétraitements ont été réalisés avec la bibliothèque scikit-learn (Pedregosa et al., 2011).

### V.2.3.2. Données imageries

L'IRM acquise à la phase aigüe pour le diagnostic de l'AVC a été récupérée pour l'ensemble des patients. Les imageries FLAIR et DWI ont été prétraitées de la manière suivante : (1) extraction du cerveau avec SynthStrip (Hoopes et al., 2022); (2) correction d'inhomogénéité de champ avec N4ITK (Tustison et al., 2010); (3) alignement rigide sur un cerveau de la base de données pour éviter les gros changements d'orientation entre les patients avec fsl flirt (Jenkinson et al., 2002); et (4) normalisation des intensités par z-score intra cerveau.

### V.2.3.3. Données textuelles



36. Figure V-1. Pré-traitements des données textuelles.

Les comptes rendus radiologiques correspondants à l'IRM ont été extraits de l'entrepôt de données de santé du CHU de Lille (INCLUDE) (Figure V-1). Pour limiter le bruit, seule la section « Résultats » de chaque rapport a été identifiée à l'aide d'un regex et conservée. Les textes ont ensuite été tokenisés à l'aide du modèle DrBERT, un modèle de langage pré-entraîné sur des données biomédicales françaises (Labrak et al., 2023).

### V.2.4. Dataset

L'ensemble de données incluait 719 patients, chacun disposant d'une imagerie FLAIR, d'une imagerie DWI, d'un compte rendu radiologique et des variables cliniques tabulaires. Les patients ont été répartis de manière aléatoire en deux sous-ensembles : 80 % pour l'entraînement et 20 % pour le test, en veillant à maintenir la stratification selon les valeurs du mRS à 3 mois. Pour l'optimisation des hyperparamètres, une validation croisée à 5 plis (5-fold cross-validation) a été appliquée sur la base d'entraînement. Le modèle final a ensuite été réentraîné sur l'ensemble de la base d'entraînement avec les meilleurs paramètres retenus.

### V.2.5. Modèles multimodaux et baseline

Dans cette étude, trois modèles multimodaux de type *deep learning* ont été implémentés et comparés entre eux, en plus d'un modèle baseline reposant uniquement sur les données cliniques.

### V.2.5.1. Modèle Baseline

Le modèle de référence utilisait uniquement les variables cliniques suivantes : âge, score NIHSS avant intervention, sexe, hypertension, diabète, hypercholestérolémie, tabagisme, consommation d'alcool et mRS avant intervention.

Ce modèle correspondait à une machine à vecteurs de support (SVM) (Schölkopf and Smola, 2001) implémentée avec la bibliothèque scikit-learn (Pedregosa et al., 2011). Le noyau utilisé était gaussien (RBF), avec un paramètre  $\gamma = 1/(nb\ données \times var(données))$ .

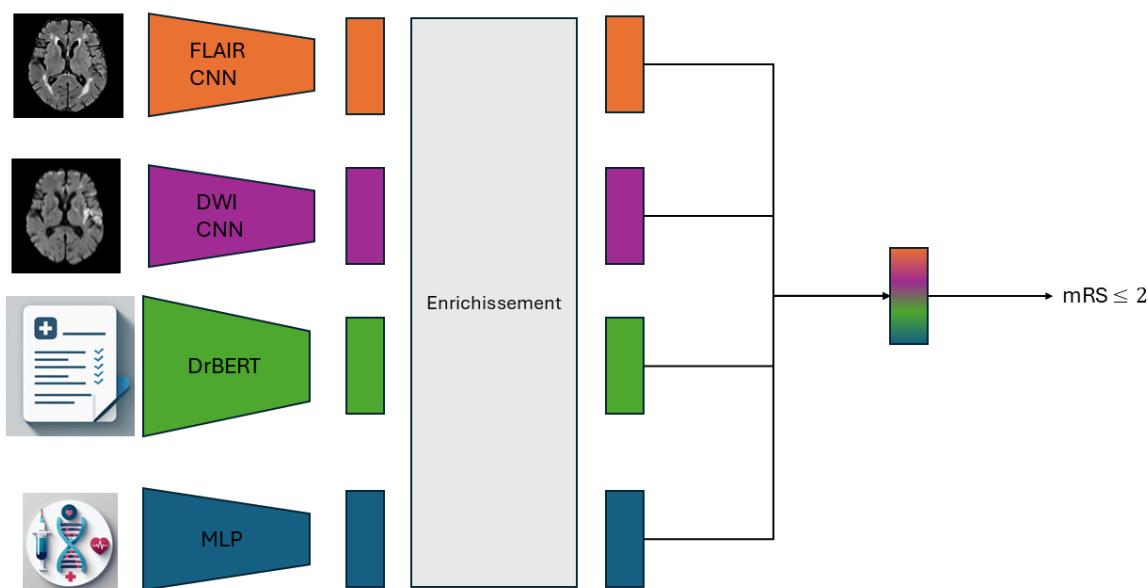
### V.2.5.2. Modèles multimodaux et interprétabilité

#### V.2.5.2.1. Visualisation du modèle

Le modèle multimodal complet intégrait les quatre modalités disponibles : imagerie FLAIR, imagerie DWI, comptes rendus radiologiques et données cliniques tabulaires (Figure V-2).

Les images FLAIR et DWI étaient encodées par deux CNN pour encoder les images, les textes étaient encodés par DrBERT (Labrak et al., 2023) et enfin, les données cliniques étaient transformées par un encodeur simple basé sur des couches entièrement connectées.

Trois variantes de modèles ont été testée : texte + clinique (référencé comme texte), Imagerie (DWI, FLAIR) + clinique (référencé comme imagerie) et enfin le modèle complet incluant toutes les modalités (référencé comme texte + image).



37. Figure V-2. Visualisation du modèle multimodal.  
MLP : perceptron multicouche.

#### V.2.5.2.1. Enrichir les modalités par contexte

Comme illustré dans la figure V-2, chaque modalité était initialement encodée indépendamment des autres, puis projetée dans un espace latent commun. Cette approche présente toutefois une limite : avant la fusion, une modalité ne pouvait pas exploiter l'information issue des autres, ce qui risquait de produire des représentations partielles et peu informatives. Cette difficulté est particulièrement marquée dans le cas des AVC : il est fréquent que certaines modalités révèlent une lésion alors

qu'elle demeure invisible sur d'autres (Chen and Ni, 2012; Kim et al., 2025).

Pour pallier cette limitation, nous avons introduit un mécanisme d'enrichissement intermodal reposant sur une combinaison de cross-attention et de residual gate mechanism (Savarese and Figueiredo, 2019; Wei et al., 2020).

Formellement, soit  $X_{cible}$  la représentation latente d'une modalité cible, et  $X_{context}$  l'ensemble des autres modalités. L'enrichissement était défini par :

$$X_{enrichi} = X_{cible} + \text{gate}([X_{cible}, \text{CrossAtt}(X_{cible}, X_{context})] \times \text{CrossAtt}(X_{cible}, X_{context})) \quad (\text{Eq. V-1})$$

Où l'opérateur  $\text{gate}(\cdot)$  contrôle l'intégration adaptative des informations croisées.

Concrètement, la cross-attention extrait des informations contextuelles pertinentes à partir des modalités voisines, tandis que la  $\text{gate}$  module leur intégration : si le contexte est peu informatif ou bruité, son poids est réduit ; à l'inverse, s'il apporte une information complémentaire, son intégration est favorisée.

#### V.2.5.2.2. Fusionner les modalités

Une fois enrichies, les modalités étaient fusionnées dans un espace latent commun en vue de la classification. Pour cela, nous avons utilisé une méthode d'attention inspirée du multiple instance learning (Ilse et al., 2018).

Chaque modalité  $m$  (FLAIR, DWI, texte, clinique) était représentée par un vecteur latent  $z_m$ . La fusion était réalisée selon la formule :

$$z = \sum_m \alpha_m z_m \quad (\text{Eq. V-2})$$

où  $\alpha_m$  correspond au poids d'attention attribué à la modalité  $m$  vérifiant  $\sum_m \alpha_m = 1$ . Ces poids permettaient d'interpréter la contribution relative de chaque modalité dans la décision finale du modèle. Ainsi, le processus de fusion ne reposait pas uniquement sur une concaténation brute, mais sur une pondération adaptative et interprétable des modalités.

#### V.2.5.2.3. Classification binaire et ordinaire

L'objectif de la tâche était de prédire le devenir fonctionnel à 3 mois après l'AVC, évalué par le score mRS. Afin d'explorer différentes formulations du problème, deux types de classification ont été considérés : (1) Classification binaire : distinction entre bon pronostic ( $mRS \leq 2$ ) et mauvais pronostic ( $mRS > 2$ ). Cette dichotomisation correspondait à la pratique la plus répandue dans les essais cliniques et dans les études pronostiques ; (2) Classification ordinaire : prédiction directe du score mRS comme une variable ordinaire à 7 niveaux (0 à 6). Cette approche permettait de préserver l'information contenue dans la structure hiérarchique du score, évitant la perte de granularité inhérente à la dichotomisation.

Pour la classification binaire, la couche de sortie du réseau consistait en deux neurones, optimisé par une cross-entropy. Pour la classification ordinaire, nous avons implémenté une couche ordinaire proposée par Cao et al. (Cao et al., 2020). Dans ce schéma, le problème est reformulé en une suite de comparaisons binaires de type  $P(mRS > k)$  avec  $k \in \{0, \dots, 5\}$ . Chaque seuil  $k$  correspondait ainsi à un classifieur binaire spécifique, ce qui permettait de respecter l'ordre naturel des catégories.

## V.2.6. Entraînement et évaluation

Les modèles ont été entraînés sur la base d'entraînement avec une validation croisée à 5 plis afin d'optimiser les hyperparamètres (taux d'apprentissage, poids de régularisation, dimension des couches latentes). Pour l'optimisation, nous avons utilisé l'optimiseur AdamW avec un learning rate initial de 1e-4 et un weight decay de 1e-5 (Loshchilov and Hutter, 2019).

La fonction de perte combinait les tâches binaires et ordinaires pour favoriser la régularisation et l'apprentissage multitâche. Plus précisément, la perte finale correspondait à la somme pondérée de la binary cross-entropy pour la classification binaire et de la loss ordinaire pour la classification ordinaire. Cette approche permettait au modèle de capturer à la fois la séparation grossière entre bon et mauvais pronostic et la structure graduelle du score mRS.

Pour évaluer les performances, nous avons utilisé les métriques suivantes : AUC (Area Under the Curve) de la courbe ROC pour les classifications binaire et ordinaire et Sensibilité et spécificité, calculées avec des intervalles de confiance à 95 % sur les prédictions de la base de test.

Enfin, pour l'interprétabilité du modèle, les poids d'attention issus du mécanisme MIL et du module d'enrichissement cross-modal ont été analysés. Cela a permis de visualiser l'influence relative de chaque modalité dans la décision finale, tant au niveau individuel qu'au niveau global du cohort. Ces analyses offrent une perspective fine sur le rôle complémentaire des données cliniques, textuelles et d'imagerie dans la prédiction du mRS à 3 mois.

## V.2.7. Tests statistiques

L'ensemble des tests statistiques a été réalisé de la manière suivante : les variables continues étaient comparées par kruskal and wallis, les variables catégorielles et ordinaires par un test khi deux. Les tests ont été corrigés par la méthode Benjamini Hochberg le seuil de significativité était fixé à  $p \leq 0.05$  (Benjamini and Hochberg, 1995).

# V.3. Résultats

## V.3.1. Démographie

La démographie des patients est résumée dans le Tableau V-1. Les valeurs de l'âge et du score NIHSS ainsi que le mRS avant intervention différaient significativement entre les groupes à bon pronostic ( $mRS \leq 2$ ) et mauvais pronostic ( $mRS > 2$ ). Les autres variables cliniques, comme le cholestérol ou le sexe, ne présentaient pas de différences significatives dans la base de test.

<b>Base entraînement</b>	<i>mRS &gt; 2</i>	<i>mRS ≤ 2</i>	p-value
<b>Nb scans</b>	355	220	-
<b>Age, années</b>	$72.95 \pm 14.64$	$62.06 \pm 15.06$	<b>&lt; 0.001</b>
<b>NIHSS</b>	$16.99 \pm 7.85$	$11.82 \pm 6.83$	<b>&lt; 0.001</b>
<b>Femme, n (%)</b>	203 (57.2%)	97 (44.1%)	<b>&lt; 0.001</b>
<b>Hypertension artérielle</b>	257 (72.4%)	111 (50.5%)	<b>&lt; 0.001</b>
<b>diabète</b>	94 (26.5%)	34 (15.5%)	<b>0.003</b>
<b>Cholesterole</b>	161 (45.4%)	84 (38.2%)	0.11

<b>Fumeur</b>	137 (38.6%)	111 (50.5%)	<b>0.007</b>
<b>Alcoolique</b>	42 (11.8%)	42 (19.1%)	<b>0.02</b>
<b>mRS avant intervention</b>	0	165 (46.5%)	181 (82.3%)
1	54 (15.2%)	24 (10.9%)	<b>&lt; 0.001</b>
2	43 (12.1%)	15 (6.8%)	
3	75 (21.1%)	0 (0.0%)	
4	17 (4.8%)	0 (0.0%)	
5	1 (0.3%)	0 (0.0%)	
<b>Base de test</b>			
<b>Nb scans</b>	89	55	
<b>Age, années</b>	73.13 ± 14.15	66.09 ± 14.36	<b>0.005</b>
<b>NIHSS</b>	19.07 ± 8.32	11.76 ± 7.11	<b>&lt; 0.001</b>
<b>Femme, n (%)</b>	46 (51.7%)	24 (43.6%)	0.52
<b>Hypertension artérielle</b>	68 (76.4%)	34 (61.8%)	0.15
<b>diabète</b>	17 (19.1%)	5 (9.1%)	0.24
<b>Cholesterole</b>	40 (44.9%)	23 (41.8%)	0.84
<b>Fumeur</b>	31 (34.8%)	25 (45.5%)	0.35
<b>Alcoolique</b>	7 (7.9%)	6 (10.9%)	0.81
<b>mRS avant intervention</b>	0	39 (43.8%)	46 (83.6%)
1	15 (16.9%)	6 (10.9%)	<b>&lt; 0.001</b>
2	14 (15.7%)	3 (5.5%)	
3	17 (19.1%)	0 (0.0%)	
4	4 (4.5%)	0 (0.0%)	
5	0 (0.0%)	0 (0.0%)	

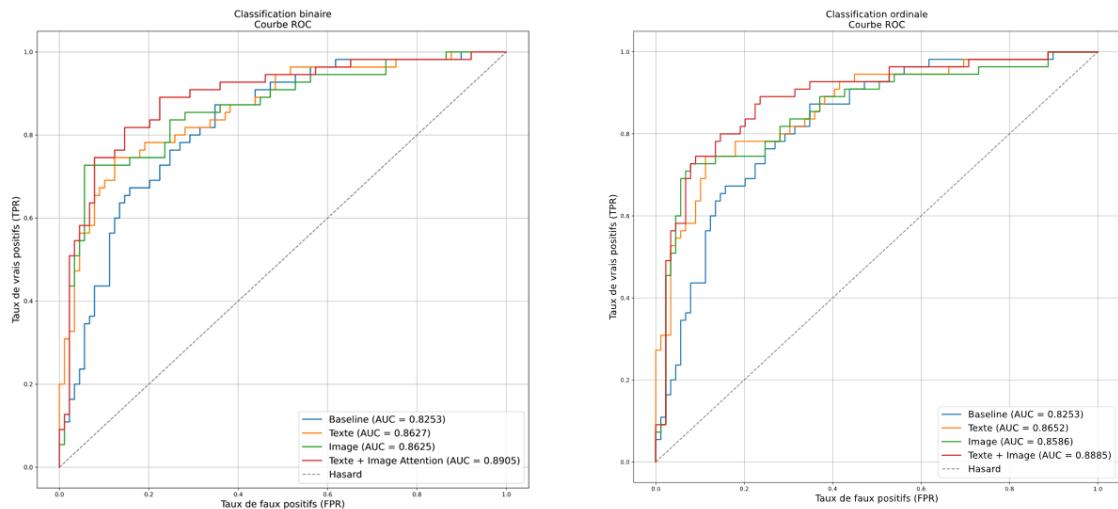
25. Tableau V-1. Caractéristiques démographiques de la base d'entraînement et de test.

### V.3.2. Performance classification

Classification binaire				
	Baseline	Texte + Clinique	Image + Clinique	
<b>AUC</b>	0.8356 [0.7632- 0.9081]	0.8627 [0.7957- 0.9297]	0.8625 [0.7955- 0.9295]	0.8905 [0.8300- 0.9510]
<b>Sens</b>	0.6545 [0.5289- 0.7802]	0.8182 [0.7162- 0.9201]	0.7273 [0.6096- 0.8450]	0.8364 [0.7386- 0.9341]
<b>Spec</b>	0.8539 [0.7806- 0.9273]	0.7079 [0.6134- 0.8023]	0.9438 [0.8960- 0.9917]	0.7865 [0.7014- 0.8716]
Classification ordinaire				
<b>AUC</b>		0.8652 [0.7987- 0.9316]	0.8586 [0.7908- 0.9265]	0.8885 [0.8275- 0.9494]
<b>Sens</b>		0.5455 [0.4139- 0.6771]	0.5091 [0.3770- 0.6412]	0.7455 [0.6303- 0.8606]
<b>Spec</b>		0.9438 [0.8960- 0.9917]	0.9663 [0.9288- 1.0038]	0.9101 [0.8507- 0.9695]

26. Tableau V-2. Performance classification binaire et ordinale.

Les résultats sont de la forme valeur [IC95%]. AUC : Aire sous la courbe, Sens : Sensibilité, Spec : Spécificité.

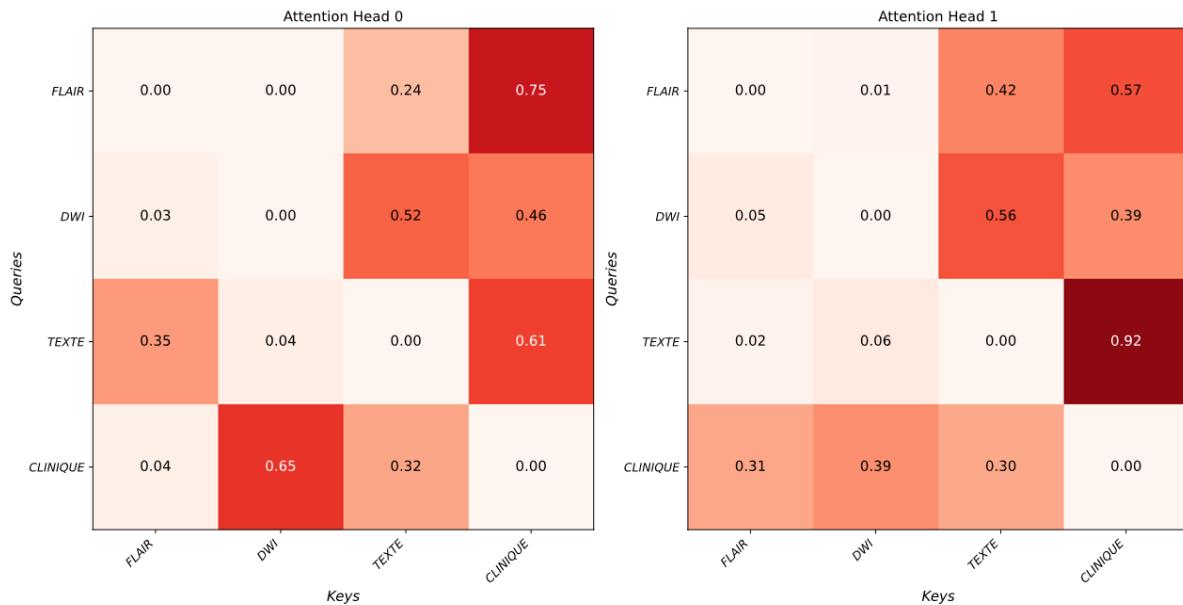


38. Figure V-3. Courbe ROC.

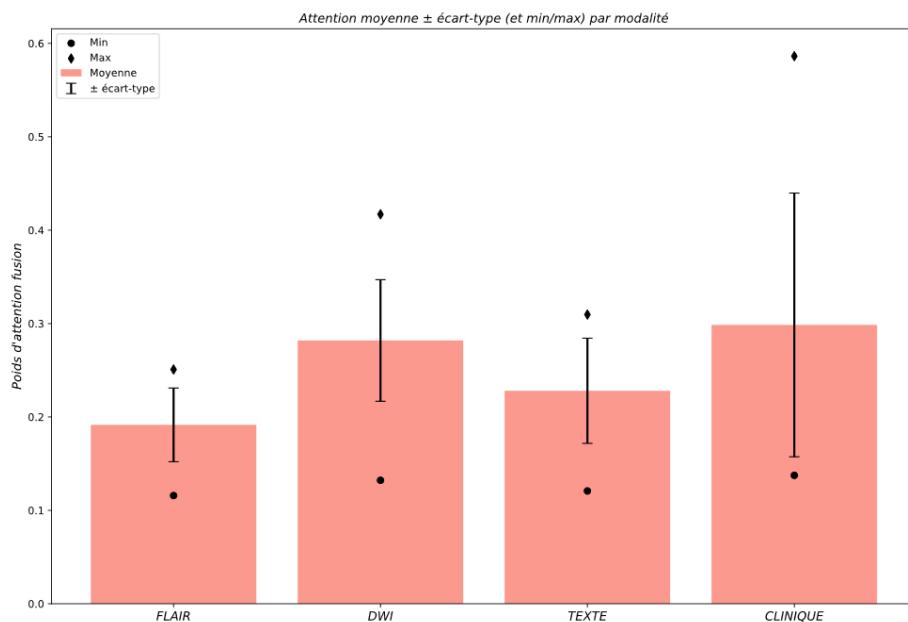
A gauche classification binaire et à droite classification ordinaire. La courbe bleue correspond au modèle baseline, la courbe orange au modèle clinique et textuel, la courbe verte au modèle clinique et imagerie et enfin la courbe rouge au modèle complet.

Les performances des modèles de prédiction du mRS à 3 mois sont présentées dans le Tableau V-2 et la Figure V-3, pour les classifications binaire et ordinaire. Les résultats montrent que le modèle baseline, utilisant uniquement les données cliniques, présentait des performances correctes mais limitées. L'ajout des données textuelles issues des comptes rendus médicaux ou des imageries améliorait légèrement les résultats comparés à ceux du modèle baseline. Le modèle complet, intégrant toutes les modalités (clinique, texte et images), fournissait les meilleures performances globales. En comparant les deux types de classification, binaire et ordinaire, les résultats étaient globalement cohérents.

### V.3.2. Interaction multimodale



39. Figure V-4. Visualisation de l'attention crossmodal.  
En ordonnée, les modalités queries. En abscisse, les modalités keys.



40. Figure V-5. Visualisation de l'attention multimodale pour la fusion.

La figure V-4 montre les poids d'attention pour l'enrichissement cross-modal. En ordonnée se trouvaient les modalités *query* et en abscisse les modalités *key*. On observait que le texte et les données cliniques influençaient fortement l'enrichissement des images, tandis que l'influence des images entre elles était faible. Pour la modalité clinique, une tête d'attention était surtout influencée par la DWI et le texte, tandis que l'autre était relativement uniforme vis-à-vis des autres

modalités. Pour la modalité texte, une tête se concentrait principalement sur la clinique, l'autre incluait également une faible influence de la FLAIR.

La figure V-5 présente la fusion des modalités dans l'espace latent avant classification. En moyenne, la modalité clinique et la DWI avaient les poids d'attention les plus élevés, suivies des comptes rendus médicaux, puis de l'imagerie FLAIR.

## V.4. Discussion

Dans cette étude, nous avons développé un modèle multimodal pour prédire le score mRS à 3 mois après un AVC ischémique, en intégrant imagerie (FLAIR, DWI), données cliniques et comptes rendus médicaux. L'objectif était de construire un espace latent commun interprétable, permettant de combiner des sources hétérogènes et d'améliorer la précision prédictive du pronostic fonctionnel.

### V.4.1. Performance du modèle

Le modèle complet, intégrant toutes les modalités (FLAIR, DWI, comptes rendus radiologiques et variables cliniques), a surpassé les approches unimodales et bimodales, confirmant que la complémentarité des modalités est essentielle pour capturer la complexité des trajectoires post-AVC. Nos résultats confirment que la fusion multimodale permet de produire des représentations plus riches et plus robustes, en cohérence avec des travaux récents (Jung et al., 2024; Shurab et al., 2024). Un aspect original de notre approche réside dans l'intégration des comptes rendus médicaux, une modalité rarement exploitée dans les modèles de pronostic post-AVC. En effet, la plupart des travaux se concentrent sur l'imagerie et les données cliniques tabulaires (Borsos et al., 2024; Isensee et al., 2019; Liu et al., 2023; Zihni et al., 2020), alors que les textes contiennent des informations contextuelles complémentaires susceptibles d'affiner la prédiction (occlusion d'un gros vaisseau, présence de microsaignements, trophicité cérébrale, etc.). Les comptes rendus d'imagerie sont pourtant une donnée accessible en pratique clinique, où chaque imagerie pour AVC fait l'objet d'un compte rendu rédigé par un radiologue. Nos résultats suggèrent que ces données textuelles brutes apportent un signal pertinent, renforçant la robustesse du modèle lorsqu'elles sont combinées aux autres modalités.

### V.4.2. Interprétabilité du modèle

L'une des forces majeures de notre approche résidait dans l'interprétabilité du modèle. Les mécanismes d'attention appliqués à l'enrichissement cross-modal et à la fusion des modalités ont permis de visualiser la contribution relative de chaque source d'information à la décision finale pour chaque patient. Ces analyses ont montré quelles modalités étaient les plus influentes dans différentes situations, permettant de comprendre comment les informations cliniques, textuelles et d'imagerie interagissaient.

Le mécanisme de *residual gating* a renforcé cette interprétabilité en modulant l'intégration des informations contextuelles provenant des autres modalités. Ce dispositif a permis au modèle de filtrer les informations non pertinentes tout en conservant les apports significatifs pour enrichir la représentation d'une modalité cible.

Grâce à cette approche, il a été possible d'identifier que la DWI et les variables cliniques contribuaient généralement le plus à la prédiction du mRS, tandis que les comptes rendus médicaux et la FLAIR apportaient des informations complémentaires selon les cas. Ces observations confirment que l'apprentissage multimodal, associé à des mécanismes d'attention et de gating, offre

un niveau d’interprétabilité exploitable pour l’analyse patient-centrée, tout en maintenant de bonnes performances prédictives

#### **V.4.3. Limites**

Malgré les performances prometteuses et l’interprétabilité offerte par notre approche, plusieurs limites ont été identifiées. Premièrement, la taille de la cohorte était relativement modeste, avec 719 patients, ce qui peut limiter la généralisation du modèle à d’autres populations ou centres hospitaliers.

Deuxièmement, certaines modalités pouvaient contenir du bruit ou des informations redondantes. Par exemple, la FLAIR a montré un poids d’attention global plus faible et moins consistant, ce qui suggère que sa contribution était parfois limitée par rapport à d’autres modalités plus discriminantes.

Troisièmement, bien que le mécanisme d’attention et de gating aient amélioré l’interprétabilité, ils ne permettaient pas de capturer toutes les interactions complexes possibles entre les modalités. Les corrélations non linéaires plus subtiles ou les combinaisons de variables rares pouvaient rester sous-représentées dans l’espace latent.

Enfin, le modèle prédisait le mRS à 3 mois à partir de données collectées principalement à un temps précis avant l’intervention. Les événements intermédiaires post-AVC, tels que les complications ou les variations de traitement, n’étaient pas pris en compte, ce qui peut limiter la précision des prédictions dans des contextes cliniques plus dynamiques. L’intégration d’informations post-opératoires, comme le succès de la recanalisation ou d’autres événements cliniques, pourrait améliorer la qualité des prédictions et la pertinence du modèle pour le pronostic fonctionnel (Jiang et al., 2025).

### **V.5. Conclusion**

Cette étude a démontré que l’intégration multimodale de l’imagerie (FLAIR et DWI), des comptes rendus médicaux et des données cliniques permettait de prédire le score mRS à 3 mois avec de bonnes performances. Notre modèle, basé sur un espace latent commun et des mécanismes d’attention, offre des prédictions robustes et interprétables. Les résultats mettent en évidence la valeur prédominante de la DWI et des données cliniques, avec un apport complémentaire des textes et de la FLAIR. Ces avancées ouvrent la voie à des travaux futurs visant à améliorer la généralisation et la robustesse des modèles dans des contextes cliniques variés.

# VI. Discussion et conclusion

---

Au fil de ce manuscrit, nous avons exploré la conception et l'exploitation d'espaces latents appliqués à l'imagerie cérébrale et aux maladies neurologiques. L'hypothèse directrice de cette thèse est que les espaces latents, s'ils sont conçus et structurés de manière adéquate, peuvent constituer une interface privilégiée entre l'intelligence artificielle et la pratique clinique.

Les premiers travaux (chapitre II) ont montré, à travers l'exemple du BrainAGE, qu'une représentation latente pouvait constituer un biomarqueur pertinent, capable de différencier des sous-groupes de patients atteints de maladie d'Alzheimer indépendamment de leur phénotype clinique. Cette première étape illustre le potentiel des représentations latentes à saisir des dimensions phénotypiques, difficiles à capturer par les approches cliniques conventionnelles.

Nous avons ensuite proposé le PatientSpace (chapitres III et IV), un cadre méthodologique combinant apprentissage supervisé, non supervisé et auto-supervisé. L'originalité de PatientSpace réside dans sa capacité à concilier généralisation, robustesse et explicabilité, répondant ainsi à l'une des critiques majeures de l'IA en médecine : l'opacité des modèles. Sur la DFT (chapitre III), le PatientSpace a permis d'identifier des sous-groupes cohérents tout en rendant les décisions algorithmiques interprétables. Sur la MA et les DFT (chapitre IV), cette approche a été étendue à une fusion multimodale intra-imagerie (TEP et IRM) par concaténation des espaces latents. Bien que "simple", cette stratégie a démontré la pertinence de la combinaison de modalités de neuroimagerie pour mieux caractériser la variabilité interindividuelle, tout en confirmant la robustesse du PatientSpace dans un contexte clinique plus complexe.

Enfin, cette logique a été prolongée dans le chapitre V vers une multimodalité plus hétérogène (séquences IRM FLAIR et DWI, données cliniques et comptes rendus médicaux), intégrée dans un espace latent commun par des mécanismes d'attention et de gating. Cette approche a permis non seulement d'améliorer la performance prédictive pour le pronostic fonctionnel post-AVC, mais aussi de pondérer explicitement la contribution de chaque modalité, offrant ainsi une interprétation directement exploitable par le clinicien.

Ces travaux tracent ainsi une progression méthodologique et conceptuelle claire : de l'utilisation d'un biomarqueur latent unique (BrainAGE), à la structuration d'espaces interprétables pour des pathologies neurodégénératives (PatientSpace), jusqu'à l'intégration multimodale hétérogène et interprétable pour le pronostic de l'AVC. Au-delà de la performance brute des modèles, l'apport principal de cette thèse réside dans la démonstration que les espaces latents peuvent être structurés, multimodaux et interprétables, ouvrant la voie à une meilleure intégration des approches d'IA dans la recherche et la pratique clinique.

Dans la suite de cette discussion, nous aborderons successivement : (i) la structuration efficace des espaces latents, (ii) l'intégration de données multimodales, (iii) la définition et l'exploitation de la similarité inter-patients, avant de conclure sur les perspectives ouvertes par ces travaux.

## VI.1. Structuration efficace des espaces latents

Comme nous l'avons montré au cours de ce manuscrit, les espaces latents constituent des outils puissants pour réduire la dimensionnalité et extraire des représentations pertinentes de données

médicales complexes. Cependant, la nature de ces représentations dépend étroitement du paradigme d'apprentissage employé. Les approches non supervisées et auto-supervisées permettent de capturer des représentations riches et globales, susceptibles de révéler des structures émergentes dans les données (Le-Khac et al., 2020). Leur valeur clinique immédiate demeure toutefois limitée, car elles ne sont pas directement orientées vers une tâche médicale spécifique. À l'inverse, les approches supervisées privilégient des caractéristiques discriminantes alignées sur une tâche définie (par ex. classification diagnostique), mais risquent de réduire la richesse de la représentation.

La combinaison de ces approches apparaît donc essentielle pour obtenir des représentations à la fois expressives et cliniquement exploitables. Nos travaux ont montré que cet équilibre pouvait être atteint grâce à des stratégies adaptées, notamment via l'apprentissage multitâche et l'intégration explicite de variables structurantes telles que l'âge.

### **VI.1.1. Structuration par apprentissage multitâche**

Dans les chapitres III et IV, nous avons exploré la voie du multitask learning (Caruana, 1993) avec la conception du PatientSpace, en intégrant simultanément apprentissage supervisé, non supervisé et auto-supervisé. Le principe est de partager une représentation commune à plusieurs tâches, dans l'espoir d'améliorer la généralisation et la robustesse (Ruder, 2017). Nos résultats sur la DFT (chapitre III) et sur Alzheimer/DFT (chapitre IV) confirment que ce paradigme permet d'extraire des représentations cohérentes, capables de capturer à la fois des dimensions diagnostiques et des structures émergentes non spécifiées a priori (par exemple la sévérité de la maladie). Toutefois, ces approches posent des difficultés bien connues : le negative transfer, où des tâches faiblement liées se perturbent mutuellement (Standley et al., 2020), ou encore la compétition entre tâches, où certaines dominent l'entraînement au détriment des autres (Crawshaw, 2020). Dans le domaine médical, ces limites sont particulièrement critiques, car un modèle trop généraliste peut perdre en précision et compromettre la fiabilité des décisions cliniques. À l'inverse, un modèle trop spécialisé peut manquer de généralisation. Plusieurs stratégies ont été proposées pour pallier ces limites, qu'il s'agisse de modifier l'architecture des réseaux (Crawshaw, 2020; Guo et al., 2018; Ma et al., 2018), d'ajuster la fonction de perte ou réguler la rétropropagation du gradient (Chen et al., 2018; Cipolla et al., 2018; Liu et al., 2019; Yu et al., 2020).

Dans ce contexte, le PatientSpace illustre l'intérêt d'un apprentissage multitâche adapté aux spécificités médicales : il permet de concilier généralisation et pertinence clinique, tout en fournissant une représentation explicable et exploitable par le clinicien.

### **VI.1.2. L'âge, un facteur structurant complexe**

Nos travaux ont également mis en lumière la complexité de l'âge comme covariable. Dans le PatientSpace, nous l'avons intégré de plusieurs manières :

- comme variable d'entrée (chapitre III), permettant de normaliser partiellement les représentations vis-à-vis de son influence
- comme contrainte latente ou variable prédictive (chapitre IV), favorisant une structuration longitudinale de l'espace
- comme biomarqueur dérivé à travers le BrainAGE (chapitre II), qui a montré une corrélation robuste avec la sévérité clinique.

Cette diversité d'approches reflète la double nature de l'âge : facteur de progression pathologique (deux patients au même stade clinique peuvent présenter des profils distincts selon l'âge), mais aussi facteur discriminant (formes précoces d'Alzheimer versus formes tardives, impliquant des patterns neuroanatomiques différents ; (Koedam et al., 2010)).

Un espace latent adapté aux maladies neurodégénératives doit donc répondre à trois critères : (1) distinguer les formes cliniques connues (différences d'atrophie, d'hypométabolisme, etc.) ; (2) permettre un suivi longitudinal patient-spécifique, notamment pour le diagnostic précoce et le suivi des MCI ; et (3) intégrer des métriques dérivées comme le BrainAGE, afin de relier âge biologique, structure cérébrale et progression pathologique.

Ainsi, la structuration des espaces latents ne se résume pas à une simple réduction de dimension, mais nécessite un design méthodologique réfléchi, intégrant à la fois des contraintes d'apprentissage et des covariables cliniquement pertinentes.

## VI.2. Données multimodales et intégration latente

Les données médicales sont intrinsèquement multimodales, combinant imagerie, variables cliniques, scores cognitifs ou encore biomarqueurs biologiques. Leur intégration est donc essentielle pour exploiter pleinement leur potentiel et construire des modèles à la fois performants et interprétables. Comme souligné par Baltrusaitis et al. (Baltrusaitis et al., 2019), une exploitation réfléchie de la multimodalité permet d'améliorer la qualité des représentations, d'accroître la robustesse des modèles et, en contexte médical, de rapprocher les prédictions algorithmiques des réalités cliniques.

### VI.2.1. De l'unimodal au multimodal

Dans la littérature, trois stratégies principales d'intégration multimodale sont décrites. La fusion précoce consiste à concaténer toutes les modalités dès l'entrée du réseau. Cette approche présente l'avantage de permettre l'apprentissage immédiat des interactions entre modalités, mais elle impose que toutes les données soient compatibles dès le départ, ce qui peut être problématique lorsque des modalités très hétérogènes, comme du texte et de l'imagerie, sont combinées. De plus, la fusion précoce peut limiter la capacité du réseau à extraire des caractéristiques propres à chaque modalité. La fusion intermédiaire consiste, quant à elle, à traiter chaque modalité indépendamment avant de les combiner dans un espace latent commun. Cette méthode préserve les informations spécifiques à chaque modalité tout en permettant, après enrichissement, de capturer des interactions complexes. Elle est particulièrement adaptée à des architectures capables de gérer des modalités hétérogènes, bien qu'elle nécessite une architecture distincte pour chaque type de donnée, ce qui peut augmenter les besoins en mémoire et en calcul.

Enfin, la fusion tardive (late fusion) agrège les sorties finales de chaque modalité pour produire une prédiction unique. Cette approche, simple à mettre en œuvre et inspirée de l'ensemble learning, présente l'inconvénient de ne pas exploiter les interactions intermodales, ce qui peut limiter les performances dans des contextes où la complémentarité des données est cruciale.

Dans ce manuscrit, nous avons d'abord exploré une stratégie de fusion intermédiaire par concaténation (chapitre IV), combinant les espaces latents issus de l'IRM et de la TEP. Cette approche a permis d'obtenir de bonnes performances et de confirmer l'intérêt d'exploiter la complémentarité entre modalités d'imagerie pour mieux caractériser la variabilité interindividuelle

dans les maladies d'Alzheimer et de la DFT. Toutefois, elle présente une limite majeure : elle ne fournit aucune information sur la contribution respective des modalités, ce qui restreint son interprétabilité et donc son exploitation clinique. Le chapitre V a marqué une avancée en proposant une multimodalité plus hétérogène (IRM FLAIR et DWI, données cliniques, comptes rendus médicaux) intégrée grâce à des mécanismes d'attention et de gating. Cette approche a permis non seulement d'améliorer la performance prédictive du pronostic post-AVC, mais aussi de pondérer explicitement l'apport de chaque modalité. Les analyses ont montré que la DWI et les variables cliniques étaient souvent les plus déterminantes, tandis que la FLAIR et les comptes rendus textuels apportaient une information complémentaire utile dans certains cas. Cette capacité à identifier les modalités prédominantes ou secondaires constitue un atout majeur pour l'interprétabilité et rapproche le modèle d'une utilisation clinique effective.

### **VI.2.2. Enrichir les modalités par contexte**

Un défi central de la multimodalité est que chaque modalité porte sa propre échelle, sa granularité et parfois son propre bruit. Fusionner directement ces représentations peut conduire à des espaces latents déséquilibrés, où certaines modalités dominent artificiellement.

Pour pallier cette limite, nous avons proposé un mécanisme d'enrichissement cross-modal (chapitre V), où chaque modalité est augmentée par l'information issue des autres, grâce à une attention croisée résiduelle modulée par une gate. Celle-ci agit comme un filtre, autorisant uniquement les informations pertinentes tout en limitant l'impact du bruit. Cette approche présente un double intérêt : elle préserve l'identité de chaque modalité et capture les interactions complexes entre elles, tout en fournissant un mécanisme interprétable sur l'origine de l'information.

Dans le cas de l'AVC, cette stratégie s'est révélée particulièrement adaptée : certaines lésions n'étaient visibles que sur la DWI, tandis que d'autres informations cliniques ou textuelles apportaient un contexte complémentaire. L'attention croisée a permis de mettre en évidence ces complémentarités, tout en évitant la dilution du signal spécifique à chaque modalité.

Ainsi, l'intégration multimodale dans un espace latent ne doit pas être envisagée comme une simple étape de concaténation, mais comme une véritable opération de négociation d'information entre modalités. La conception de mécanismes d'enrichissement et de pondération constitue une condition essentielle pour construire des espaces latents robustes, interprétables et cliniquement pertinents.

### **VI.3. Similarité entre patients**

Un des fils conducteurs de cette thèse est la conception d'espaces latents permettant de représenter la similarité entre patients. En effet, la médecine de précision repose sur la capacité à identifier des sous-groupes homogènes de patients, que ce soit pour adapter le traitement, affiner le pronostic ou comprendre la progression des maladies. Dans ce contexte, les espaces latents offrent une opportunité unique : ils projettent les individus dans un espace continu où la proximité reflète un degré de similarité.

Néanmoins, la définition de cette similarité est loin d'être triviale. Dans le domaine médical, deux patients peuvent être similaires sur le plan anatomique (patterns d'atrophie comparables en IRM), tout en divergeant fortement sur le plan fonctionnel ou biologique. À l'inverse, deux profils cliniques

proches peuvent reposer sur des mécanismes pathologiques très différents. Chaque modalité porte donc sa propre structure de similarité, qui n'est pas nécessairement réductible à celle des autres.

Nos travaux illustrent cette tension entre préservation et fusion des similarités. Dans le chapitre III, nous avons montré qu'un espace latent construit à partir de l'IRM permet de définir une similarité phénotypique robuste, cohérente avec la littérature clinique sur la DFT. Cette approche unimodale garantit une interprétation claire : la proximité entre patients correspond directement à des patrons d'atrophie similaires. Le chapitre IV a exploré une étape supplémentaire en combinant deux modalités d'imagerie (IRM et TEP). Cette fusion a permis de mieux caractériser la variabilité interindividuelle, mais la notion de similarité a changé : elle ne reflète plus exactement celle de chaque modalité prise séparément, mais correspond à une mesure combinée au sein de l'espace latent intégré. Ce résultat souligne un point essentiel : la fusion multimodale ne préserve pas les similarités unimodales, mais en construit une nouvelle, parfois divergente, dont l'interprétation doit être soigneusement analysée.

Ces résultats plaident pour une approche méthodologique prudente : plutôt que de chercher à imposer une définition unique de la similarité, il est préférable de préserver d'abord les similarités propres à chaque modalité. Cette stratégie présente plusieurs avantages. Elle permet de conserver l'identité unimodale du patient, garantissant que les informations spécifiques à chaque type de données ne soient pas diluées dans une fusion précoce. Elle ouvre également la voie à une stratification plus fine : par exemple, il devient possible d'identifier des patients présentant un pattern pathologique dans une modalité donnée (comme la TEP) mais pas dans une autre (comme l'IRM), ce qui pourrait refléter des différences de stade ou de progression de la maladie. Ce raisonnement est particulièrement pertinent dans les maladies neurodégénératives, où certaines modalités — la TEP, par exemple — détectent des anomalies plus précocement que d'autres, comme l'IRM structurelle (Del Sole et al., 2017). Une fois ces similarités unimodales établies, leur combinaison peut ensuite être explorée à travers des mécanismes explicites et interprétables, permettant de construire des représentations intégrées tout en conservant traçabilité, flexibilité et pertinence clinique.

#### **VI.4. Concevoir des espaces latents interprétables**

À l'issue des sections précédentes, il apparaît que la structuration des espaces latents, l'intégration multimodale et la définition de la similarité entre patient constituent des axes complémentaires mais interconnectés. Cette section se propose de synthétiser ces éléments pour proposer une démarche méthodologique cohérente permettant de concevoir un espace latent à la fois structuré, multimodal et interprétable, adapté aux besoins cliniques et aux contraintes des données médicales. La construction d'un tel espace repose sur une progression méthodologique claire similaire à la conception du PatientSpace, que nous présentons ici sous la forme d'une approche en plusieurs étapes :

##### **1. Étape initiale – apprentissage des informations générales**

Le modèle est d'abord entraîné pour capturer des informations générales, par exemple en distinguant les sujets sains des patients atteints de démence. La reconstruction des données originales sert de régularisation, garantissant que l'espace latent conserve une

représentation fidèle de chaque modalité d'entrée. Cette étape établit des fondations discriminatives stables et interprétables.

## 2. Étape intermédiaire – affinage par contraintes complexes

L'espace latent est ensuite affiné par des contraintes plus complexes, telles que la structuration par similarité/dissimilarité et/ou la structuration longitudinale (via l'intégration de l'âge ou des intervalles temporels entre examens successifs).

## 3. Étape de stabilisation – préservation de la robustesse et de l'équilibre

Pour éviter que certaines contraintes dominent l'apprentissage ou que le modèle oublie les attributs fondamentaux, une stratégie cyclique est utilisée (Fu et al., 2019). Cette méthode réactive périodiquement les objectifs initiaux, garantissant un équilibre entre structuration avancée et conservation des informations générales.

Lors du passage à la multimodalité, chaque latent unimodal est combiné via des mécanismes d'attention et de gating, comme présenté dans le chapitre V. Cette intégration permet de maintenir la similarité entre patient pour chaque modalité, quantifier l'influence relative de chaque modalité pour enrichir la représentation finale et enfin offrir une interprétation transparente des contributions de chaque type de données à la tâche clinique considérée. En combinant ces étapes, il devient possible de construire un espace latent structuré, multimodal et interprétable, capable à la fois d'identifier des patients similaires dans chaque modalité, de capturer les interactions entre modalités, et de fournir des informations exploitables par les cliniciens pour le diagnostic, le suivi ou le pronostic.

Avant d'aborder les perspectives, il convient de souligner deux recommandations méthodologiques issues de cette thèse.

Tout d'abord, pour préserver l'identité et l'unicité des représentations latentes, il est recommandé d'utiliser une approche de similarité contrastive. Celle-ci permet d'éviter que les latents ne s'effondrent vers un vecteur trop générique et de garantir que chaque patient conserve une représentation discriminante, à la fois pour les analyses unimodales et pour l'intégration multimodale.

Ensuite, l'âge, en tant que covariable critique dans les maladies neurodégénératives, doit être utilisé de manière adaptée au contexte : il peut conditionner l'espace latent lorsqu'un suivi longitudinal est nécessaire, afin de structurer les trajectoires individuelles, ou simplement être introduit comme variable d'entrée lorsque l'objectif n'est pas de modéliser explicitement la progression dans le temps.

Ces recommandations méthodologiques servent de fondement à la discussion des perspectives, qui explorent les extensions possibles du PatientSpace et les améliorations à apporter à l'intégration multimodale et à l'interprétabilité clinique.

## VI.5. Perspectives autour des travaux de thèse

Comme discuté dans les chapitres III et IV, le PatientSpace apparaît comme un outil prometteur pour extraire des informations pertinentes à partir de données d'imagerie, notamment en ce qui concerne les phénotypes, la sévérité et la dimension longitudinale des patients.

Une première piste d'amélioration consisterait à renforcer la robustesse des représentations

unimodales. Il s'agirait d'explorer des latents interprétables, capables de capturer des patterns spécifiques à chaque modalité avant leur intégration multimodale. Cette approche permettrait de préserver l'identité unimodale du patient, facilitant ainsi la création de nouvelles stratifications : par exemple, identifier des patients présentant un pattern pathologique dans une modalité mais pas dans une autre.

Par ailleurs, l'utilisation de mesures de similarité contrastives pourrait être plus adaptée que les approches non-contrastives employées dans nos travaux du chapitre IV. En effet, les patterns extraits semblaient fortement corrélés à l'âge des patients, probablement parce que la similarité reposait uniquement sur des exemples positifs. Cette situation pouvait projeter deux patients phénotypiquement différents mais d'âge similaire dans un espace latent proche. Une approche contrastive, intégrant des exemples négatifs explicites, permettrait de limiter ce biais et de favoriser des représentations discriminantes, comme observé dans le chapitre III.

Une autre perspective majeure concerne l'extension multimodale du PatientSpace, en s'appuyant sur les méthodes introduites dans le chapitre V. Cela permettrait non seulement d'identifier la contribution relative de chaque modalité à la prédiction d'un diagnostic, mais aussi de mieux comprendre pourquoi certaines modalités apportent plus d'information que d'autres. Par exemple, la TEP peut fournir des signaux précoces de pathologie avant qu'ils ne soient visibles en IRM. L'intégration de mécanismes d'attention appliqués au sein d'architectures profondes offrirait une voie intéressante pour quantifier et interpréter ces contributions, dépassant ainsi la simple concaténation des latents modaux.

Les travaux du chapitre V ouvrent également la possibilité de combiner les approches multimodales avec le PatientSpace pour d'autres pathologies, comme l'AVC ischémique. Cela permettrait de projeter les patients dans l'espace latent et d'explorer différents profils de manière interprétable, à l'échelle individuelle. Des méthodes de clustering ou de stratification pourraient alors être appliquées pour identifier des sous-groupes, par exemple des patients répondant favorablement à la recanalisation et d'autres moins réactifs. Cette approche rejoint les efforts actuels visant à identifier des biomarqueurs d'imagerie prédictifs du succès de la recanalisation et du pronostic fonctionnel post-AVC (Gaviria and Eltayeb Hamid, 2024).

Enfin, un axe d'amélioration concerne le mécanisme d'attention lui-même. Dans nos travaux, nous avons utilisé une variante inspirée du Multiple Instance Learning, reposant sur l'hypothèse qu'au moins une instance est discriminante pour établir un diagnostic. Or, dans un contexte clinique complexe, cette hypothèse est réductrice : ce n'est pas une instance isolée, mais souvent une combinaison de plusieurs observations (au sein d'une modalité ou entre modalités) qui apporte une information déterminante. Repenser le mécanisme d'attention — par exemple en le focalisant sur certaines dimensions spécifiques plutôt que sur l'ensemble du vecteur latent, ou en explorant des variantes comme la cross-attention — pourrait améliorer la finesse de l'interprétation et la qualité des prédictions.

En résumé, les perspectives de cette thèse s'articulent autour du renforcement des espaces latents unimodaux, similarité contrastive, extension multimodale et optimisation des mécanismes d'attention, visant à construire des modèles plus robustes, interprétables et applicables à la pratique clinique.

## **VI.6. Conclusion**

Cette thèse a exploré la conception et l'exploitation d'espaces latents structurés pour l'analyse de données médicales complexes. Le chapitre I a présenté les concepts fondamentaux des espaces latents et les méthodes de création, posant les bases théoriques pour les travaux suivants. Dans le chapitre II, nous avons montré que le BrainAGE constitue un biomarqueur robuste pour différencier des sous-groupes de patients et mieux comprendre le déclin cognitif dans la maladie d'Alzheimer. Les chapitres III et IV ont introduit le PatientSpace, un espace latent interprétable capable de capturer des phénotypes cliniques et des trajectoires longitudinales. L'aspect multimodal a été progressivement intégré afin d'évaluer son apport dans la structuration et l'interprétabilité de PatientSpace.

Le chapitre V a été consacré à l'évaluation du traitement multimodal hétérogène, intégrant imagerie, données cliniques et textuelles, en utilisant attention et gating pour améliorer la performance prédictive et l'interprétabilité au niveau patient.

Ces travaux démontrent que la structuration latente, combinée à une intégration multimodale réfléchie, permet de créer des représentations à la fois robustes, interprétables et cliniquement pertinentes, ouvrant la voie à des applications en suivi longitudinal, stratification patient et prédiction de pronostic.

# ANNEXE

## Liste des Figures

---

1. Figure 1. Les courbes de taille et de poids des enfants constituent des références normatives. L'écart observé par rapport à la population de référence permet de quantifier des critères tels que la maigreur, le surpoids ou l'obésité (image provenant de <a href="https://cress-umr1153.fr/fr/courbes-de-croissance-de-reference-du-carnet-de-sante/">https://cress-umr1153.fr/fr/courbes-de-croissance-de-reference-du-carnet-de-sante/</a> ) .....	12
2. Figure 2 Intérêt des modèles normatifs (image et légende issues de (Marquand et al., 2016)). L'approche classique cas-témoins suppose que les cas et les témoins forment chacun un groupe bien défini (A). Cette hypothèse peut être raisonnable dans certains contextes, mais de nombreuses autres configurations sont possibles en pratique. La population clinique peut être constituée de plusieurs sous-groupes présentant chacun une pathologie distincte (B) ; la variabilité liée à la maladie peut s'inscrire au sein même de la variabilité observée chez les sujets sains (C) ; ou encore, le groupe clinique peut apparaître diffus et hétérogène en raison d'erreurs diagnostiques, de comorbidités ou de l'agrégation de pathologies différentes (D).....	13
3. Figure I-1. Apprentissage supervisé. Une imagerie TEP est traitée par un réseau de neurones convolutif (CNN) afin de prédire, à partir de l'espace latent, si le patient est sain (CN) ou atteint de la maladie d'Alzheimer (MA).....	22
4. Figure I-2. Apprentissage non supervisé. Une image TEP est encodée dans un espace latent par un encodeur, puis cet espace latent est décodé par un décodeur afin de régénérer l'image TEP d'origine. .....	23
5. Figure I-3. Contrastive Learning. Chaque modalité d'imagerie est encodée dans un espace latent, puis les représentations issues d'imageries similaires sont rapprochées tandis que celles provenant d'imageries dissimilaires sont éloignées .....	24
6. Figure I-4. Non-contrastive learning. Deux images TEP transformées sont générées à partir de l'imagerie d'origine, puis leurs espaces latents sont rapprochés.....	24
7. Figure I-5. Espaces latents. Espace latent obtenu après apprentissage supervisé (en haut), non supervisé (milieu) et auto-supervisé (en bas). Chaque point représente un patient, et la couleur indique le diagnostic : contrôle (CN, violet) ou maladie d'Alzheimer (MA, jaune). .....	25
8. Figure I-6. Apprentissage multimodal. Les trois cercles en bas représentent trois modalités différentes. Lors de l'alignement, un espace latent est extrait en ne conservant que l'information commune à chaque modalité. Lors de la fusion, l'espace latent extrait contient l'ensemble des informations disponibles, sans distinguer ce qui est spécifique ou commun aux modalités. Enfin, lors de la fission, les données sont décomposées en espaces latents spécifiques à chaque modalité et en espaces latents partagés entre modalités. ....	26
9. Figure I-7. Impact de la pondération. Influence de la pondération entre les deux termes de la fonction de perte. À gauche, la reconstruction est priorisée. À droite, une régularisation plus forte contraint l'espace latent à suivre une distribution gaussienne standard en deux dimensions. (Image issue de FIDLE : <a href="https://cloud.univ-grenoble-alpes.fr/index.php/s/wxCztjYBbQ6zwd6?dir=undefined&amp;openfile=959404079">https://cloud.univ-grenoble-alpes.fr/index.php/s/wxCztjYBbQ6zwd6?dir=undefined&amp;openfile=959404079</a> ).....	28
10. Figure I-8. Disentanglement learning. Chaque dimension de l'espace latent contrôle une caractéristique des données initiales. (Image de gauche issue du dataset 3D Shapes : <a href="https://github.com/google-deepmind/3d-shapes/tree/master">https://github.com/google-deepmind/3d-shapes/tree/master</a> ). ..	29
11. Figure I-9. VAE conditionnel sur l'espace latent. Dans cet exemple, les variables cliniques et l'imagerie sont encodées afin de générer un espace latent commun. Les variables cliniques sont ensuite réutilisées pour conditionner cet espace latent commun. .....	30
12. Figure I-10. VAE conditionnel sur la reconstruction. Dans cet exemple, les variables cliniques et l'imagerie sont encodées afin de générer un espace latent commun. Les variables cliniques sont ensuite réutilisées pour conditionner la reconstruction de l'imagerie.....	31
13. Figure I-11. Modèle DIVA. Trois espaces latents indépendants sont générés par trois encodeurs. Deux de ces espaces sont guidés par la classification et le conditionnement sur le site et le diagnostic, tandis qu'un espace latent est régularisé par une gaussienne centrée réduite classique. Les trois espaces latents sont ensuite utilisés pour décoder l'image d'origine. ....	32
14. Figure I-12. SMILE-GAN. Illustration du disentanglement learning appliquée au pattern de maladie. (Image issue de SMILE-GAN <a href="https://github.com/zhijian-yang/SmileGAN/blob/main/datasets/Smile-GAN.png">https://github.com/zhijian-yang/SmileGAN/blob/main/datasets/Smile-GAN.png</a> ). Le modèle apprend à séparer les variations liées à la pathologie de celles observées dans une population saine. .....	33
15. Figure I-13. Vieillissement naturel et progression de la maladie. Illustration des hypothèses de modélisation du vieillissement normal et de la progression pathologique (Figure I-13.b), affectant la structure cérébrale de manière indépendante. (Image issue de (Ouyang et al., 2022)). .....	34

16. Figure I-14. Association texte-image. Image provenant du modèle CLIP (Radford et al., 2021). Dans ce cas, l'objectif est de contraindre la matrice d'association à se rapprocher d'une matrice identité, de façon à aligner chaque image uniquement avec son texte associé.....	36
17. Figure II-1. Pipeline de la méthode. En haut : Le modèle BrainAGE est entraîné sur une population saine pour générer une représentation encodée par l'IRM capable de prédire l'âge. En bas : Le modèle BrainAGE est utilisé pour extraire une représentation codée par IRM de chaque patient dans la population EOAD, qui est ensuite clusterisé pour identifier des groupes EOAD distincts.....	42
18. Figure II-2. Régions d'intérêt pour la signature corticale et la signature cognitive. Cortical AD : gyrus supérieur et inférieur, gyrus supramarginal, gyrus precuneus, gyrus parahippocampal, temporal inférieur Signature cognitive : gyrus cingulate antérieur, gyrus angulaire, gyrus milieu temporal Superposition (overlap) : pole temporal.....	44
19. Figure II-3. Diagramme de flux de la population.....	46
20. Figure II-4. Distribution des scores BrainAGE entre les clusters. L'abscisse correspond au cluster associé à la distribution. L'ordonnée correspond au score BrainAGE (brai nage corrigé – age chronologique). p-value < 0.001.....	47
21. Figure II-5. Évolution cognitive par cluster. Le cluster BrainAGE+ est représenté en rouge, le BrainAGE- en bleu. Les p-values pour l'écart entre les pentes sont respectivement 0.02 (MMSE) et 0.03 (CDR-SoB).....	50
22. Figure II-S1. Sélection du nombre optimal de clusters. En haut : Évaluation de la qualité du cluster sur différents nombres de cluster à l'aide de quatre métriques de performance. En bas : les tracés de silhouette pour les solutions à deux et trois groupes sont illustrés. Pour les deux, un cluster reflète la continuité de la représentation encodée par IRM utilisée pour le clustering, comme l'indiquent les valeurs de silhouette négatives, suggérant que certains sujets se trouvent près de la limite entre les clusters.....	54
23. Figure II-S2. Répartition par âge de l'ensemble d'entraînement du BrainAGE (bleu) et de la population EOAD (orange).....	55
24. Figure III-1. Architecture du modèle. Le modèle se compose en 2 encodeurs : un unimodal prenant en entrée seulement l'image et un encodeur prenant l'âge et l'image en entrée. L'espace latent associé à l'encodeur résiduel n'est utilisé que pour la reconstruction. L'espace latent dérivé de l'âge et de l'IRM est utilisé à la fois pour la reconstruction et la prédition du diagnostic. Enfin le vecteur associé à l'âge est utilisé pour générer l'espace latent du PatientSpace et la reconstruction de l'image.....	60
26. Figure III-2. Visualisation du PatientSpace et des clusters.....	66
27. Figure III-3. Prédiction des CDR-SB par la méthode PatientGraph (base INTERNE).....	67
28. Figure III-S1. Règle cyclique du facteur $\beta_c$ .....	70
29. Figure IV-1. Architecture réseau. Le modèle comprend quatre encodeurs : deux encodeurs unimodaux résiduels, qui génèrent des représentations latentes unimodales résiduelles ( $Z_{mri}$ et $Z_{pet}$ ), et deux encodeurs unimodaux supplémentaires enrichis par des blocs MMTM. Les sorties des encodeurs pilotés par les blocs MMTM sont concaténées et projetées dans un espace latent multimodal partagé ( $Z_{MM}$ ). Cette représentation multimodale, conditionnée par l'âge et le diagnostic, est utilisée conjointement avec les espaces latents unimodaux pour décoder chaque modalité séparément. ....	77
30. Figure IV-2. Graphe du PatientSpace .....	84
31. Figure IV-3. Atrophie et hypométabolisme dans les clusters. Les rangées représentent les modalités d'imagerie, tandis que les colonnes correspondent aux clusters de démence. ....	86
32. Figure IV-4. Analyse de corrélation de Pearson entre les volumes cérébraux prédits et la réalité terrain et les valeurs SUVR. ....	88
33. Figure IV-5. Prédiction du MMSE dans les classes de démence.....	89
34. Figure IV-6. Projection des MCI dans le PatientSpace. Pour une meilleure visualisation, les MCI stables (à gauche) ont été séparés des converters (à droite). Les carrés verts représentent les MCI stables, tandis que les carrés rouges représentent les MCI converters. Les cercles correspondent aux sujets du PatientSpace : vert pour CN, rouge pour MA et bleu pour DFT. ....	92
35. Figure IV-7. Analyse de Kaplan-Meier pour la conversion de la démence chez les sujets MCI.....	93
36. Figure IV-8. Matrices de transition pour les MCI stables (gauche) et converters (droite). Les valeurs correspondaient aux probabilités de transition pour les sujets MCI converters et les sujets MCI stables se déplaçant entre les clusters au fil du temps. ....	94
37. Figure V-1. Pré-traitements des données textuelles.....	110
38. Figure V-2. Visualisation du modèle multimodal. MLP : perceptron multicouche.....	111
39. Figure V-3. Courbe ROC. A gauche classification binaire et à droite classification ordinaire. La courbe bleue correspond au modèle baseline, la courbe orange au modèle clinique et textuel, la courbe verte au modèle clinique et imagerie et enfin la courbe rouge au modèle complet.....	115
40. Figure V-4. Visualtion de l'attention crossmodal. En ordonnée, les modalités queries. En abscisse, les modalités keys. ....	116

**41.** Figure V-5. Visualisation de l'attention multimodale pour la fusion.....116

# ANNEXE

## Liste des Tableaux

---

1. Tableau II-1. Caractéristiques des sujets pour l'ensemble des données d'entraînement du modèle BrainAGE. L'âge est exprimé en moyenne ± écart-type. La tranche d'âge est exprimée en min-max. Les variables catégorielles sont indiquées comme le nombre sur le nombre total (pourcentage). .....	41
2. Tableau II-2. Différences démographiques et biologiques à l'inclusion. Les variables continues sont présentées sous forme de moyenne ± d'écart-type. Les variables catégorielles sont indiquées comme le nombre sur le nombre total (pourcentage). APOE ε4, au moins un allèle APOE ε4. Le nombre de valeurs manquantes est présenté sous la forme # de valeurs manquantes. Les valeurs p significatives corrigées du FDR sont en gras. Les valeurs p sont corrigées séparément entre les parties démographies et biologies.....	48
3. Tableau II-3. Différences neuropsychologiques, volumétriques cérébrales, d'atrophie et de phénotypes cliniques à l'inclusion. Les variables continues sont présentées sous forme de moyenne ± d'écart-type. Les variables catégorielles sont indiquées comme le nombre sur le nombre total (pourcentage). Les valeurs p significatives sont en gras. IQR : Écart interquartile.....	49
4. Tableau II-4. Taux de déclin entre les clusters en termes de sévérité et de volumes cérébraux. ....	50
5. Tableau II-S1. Caractéristiques de l'ensemble des d'entraînement du BrainAGE et de la population EOAD.....	55
6. Tableau II-S2. Taux de déclin entre les groupes basés sur le BR en fonction de la sévérité et du volume cérébral. ....	55
7. Tableau II-S3. Taux de déclin entre les groupes basés sur le PAD en fonction de la gravité et du volume cérébral.....	56
8. Tableau II-S4. Taux de déclin entre les clusters de vecteurs "fabriqués à la main" en fonction de la sévérité. Les vecteurs sont composés de volumes normalisés en matière grise, de matière blanche normalisée, de LCR normalisée, d'hippocampe normalisée et d'âge. ....	56
9. Tableau III-1. Démographie des bases entraînement, validation et test. ....	64
10. Tableau III-2. Caractéristiques des clusters L'âge et le niveau d'éducation sont exprimés en moyenne (écart-type), la proportion de femmes est indiquée en pourcentage, et le CDR-SB est présenté en médiane avec l'intervalle interquartile (Q1-Q3). .....	65
11. Tableau III-3. Performances des modèles sur les bases INTERNE et EXTERNE. Les résultats sont rapportés en valeur [IC95%].....	67
12. Tableau III-S1. Démographie de la base de données externe. ....	70
13. Tableau IV-1. Caractéristiques démographiques et cliniques de l'ensemble de données. L'âge est exprimé en moyenne ± écart-type, le sexe féminin en nombre (proportion) et le MMSE en médiane avec écart interquartile (Q1-Q3). Les comparaisons statistiques ont été effectuées à l'aide du test de Kruskal-Wallis pour l'âge et le MMSE, et du test du khi carré pour le sexe. Une valeur de p ≤ 0,05 a été considérée comme statistiquement significative. ....	83
14. Tableau IV-2. Caractéristiques des clusters. L'âge est exprimé en moyenne ± écart-type, la proportion de femmes est indiquée en effectif (pourcentage) et le MMSE est présenté en médiane avec l'intervalle interquartile (Q1-Q3). ....	85
15. Tableau IV-3. Performances de classification binaire. Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité. ....	90
16. Tableau IV-4. Performances de classification multiclasse. Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité. ....	90
17. Tableau IV-5. Performances de classification binaire sur le jeu de données SOCRATES. Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité. ....	91
18. Tableau IV-6. Performances de classification multiclasse sur le jeu de données SOCRATES. Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité. ....	91
19. Tableau IV-A1. Performances de classification binaire. Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité .....	99
20. Tableau IV-A2. Performances de classification multiclases. Les résultats sont présentés sous forme de valeur [intervalle de confiance 95 %]. AUC : Aire sous la courbe ; Sens. : Sensibilité ; Spec : Spécificité. ....	101
21. Tableau IV-S1. Répartition des diagnostics entre les sites. ....	103
22. Tableau IV-S2. Caractéristiques démographiques de l'ensemble de données MEMENTO à l'inclusion. ....	104
23. Tableau IV-S3. Caractéristiques démographiques de l'ensemble de données SOCRATES. ....	104

24. Tableau IV-S4. Corrélation de Pearson entre les ROIs des MA et des DFT. Les corrélations les plus fortes sont en gras. Modéré - Les corrélations modérées sont soulignées.	106
25. Tableau V-1. Caractéristiques démographiques de la base d'entraînement et de test.	114
26. Tableau V-2. Performance classification binaire et ordinaire. Les résultats sont de la forme valeur [IC95%]. AUC : Aire sous la courbe, Sens : Sensibilité, Spec : Spécificité.	114

# ANNEXE

## Liste des abréviations

---

AUC : aire sous la courbe ROC  
AVC : accident vasculaire cérébral  
BH-FDR : Benjamini et Hochberg False Discovery Rate  
BR : réserve cérébrale  
bvFTD : variant comportemental de la DFT  
CDR : Clinical Dementia Rating  
CDR-SoB : Clinical Dementia Rating Scale Sum of Boxes  
CNN : réseaux de neurones convolutifs  
DEM : démence  
DIVA : Domain Invariant Variational autoencoder  
DTI : imagerie par tenseur de diffusion  
ELBO : evidence lower bound  
EQM : erreur quadratique moyenne  
EOAD : Early-Onset Alzheimer's Disease  
FDR : False Discovery Rate  
FLAIR : Fluid Attenuated Inversion Recovery  
FWE : Family-Wise Error rate  
GAMLSS : modèles additifs généralisés  
GANs : réseaux antagonistes génératifs  
GM : matière grise  
GMM : mixtures de Gaussiennes  
HpSp : préservation hippocampique  
ICV : volume intracrânien  
IA : intelligence artificielle  
KL : Kullback-Leibler  
LCR : liquide cérébrospinal  
LOAD : maladie d'Alzheimer à début tardif  
LLM : Large language models  
LME : modèles linéaires à effets mixtes univariés  
LP : limbique prédominant  
MCI : trouble cognitif léger  
MA : maladie d'Alzheimer  
MAE : erreur absolue moyenne  
MMSE : Mini-Mental State Examination  
MMTM : module de transfert multimodal  
MRI / IRM : imagerie par résonance magnétique  
mRS : modified Rankin Scale  
OSEM : Ordered Subsets Expectation Maximization  
PAD : Predicted Age Difference

PET / TEP : émission de positons  
PNFA : aphasic progressive non fluente  
RBM : machines de Boltzmann restreintes  
ROI : régions d'intérêt  
SGD : descente de gradient stochastique  
SVM : machine à vecteurs de support  
SUV : Standard Uptake Value  
SUVR : Standard Uptake Value Ratio  
SV : variant sémantique  
TE : temps d'écho  
TR : temps de répétition  
tAD : forme typique de la MA  
VAE : autoencodeurs variationnels  
ViTs : Vision Transformers  
VBM : Voxel-Based Morphometry  
VAT : Visual Association Test  
WM : matière blanche

# Références

---

- Abeyasinghe, P.M., Cole, J.H., Razi, A., Poudel, G.R., Paulsen, J.S., Tabrizi, S.J., Long, J.D., Georgiou-Karistianis, N., 2025. Brain Age as a New Measure of Disease Stratification in Huntington's Disease. *Mov. Disord. Off. J. Mov. Disord. Soc.* 40, 627–641.  
<https://doi.org/10.1002/mds.30109>
- Agosta, F., Galantucci, S., Filippi, M., 2017. Advanced magnetic resonance imaging of neurodegenerative diseases. *Neurol. Sci. Off. J. Ital. Neurol. Soc. Ital. Soc. Clin. Neurophysiol.* 38, 41–51. <https://doi.org/10.1007/s10072-016-2764-x>
- Ainsworth, S.K., Foti, N.J., Lee, A.K.C., Fox, E.B., 2018. oi-VAE: Output Interpretable VAEs for Nonlinear Group Factor Analysis, in: Proceedings of the 35th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 119–128.
- Akoglu, H., 2018. User's guide to correlation coefficients. *Turk. J. Emerg. Med.* 18, 91–93.  
<https://doi.org/10.1016/j.tjem.2018.08.001>
- Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M.C., Thies, B., Phelps, C.H., 2011. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement. J. Alzheimers Assoc.* 7, 270–279.  
<https://doi.org/10.1016/j.jalz.2011.03.008>
- Alemi, A.A., Poole, B., Fischer, I., Dillon, J.V., Saurous, R.A., Murphy, K., 2018. Fixing a Broken ELBO.  
<https://doi.org/10.48550/arXiv.1711.00464>
- Antelmi, L., Ayache, N., Robert, P., Lorenzi, M., 2019. Sparse Multi-Channel Variational Autoencoder for the Joint Analysis of Heterogeneous Data, in: Proceedings of the 36th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 302–311.
- Arbabshirani, M.R., Plis, S., Sui, J., Calhoun, V.D., 2017. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *NeuroImage, Individual Subject Prediction* 145, 137–165. <https://doi.org/10.1016/j.neuroimage.2016.02.079>
- Arenaza-Urquijo, E.M., Przybelski, S.A., Lesnick, T.L., Graff-Radford, J., Machulda, M.M., Knopman, D.S., Schwarz, C.G., Lowe, V.J., Mielke, M.M., Petersen, R.C., Jack, C.R., Jr., Vemuri, P., 2019. The metabolic brain signature of cognitive resilience in the 80+: beyond Alzheimer pathologies. *Brain* 142, 1134–1147. <https://doi.org/10.1093/brain/awz037>
- Aron, L., Zullo, J., Yankner, B.A., 2022. The adaptive aging brain. *Curr. Opin. Neurobiol.* 72, 91–100.  
<https://doi.org/10.1016/j.conb.2021.09.009>
- Arthur, D., Vassilvitskii, S., 2007. k-means++: the advantages of careful seeding, in: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '07. Society for Industrial and Applied Mathematics, USA, pp. 1027–1035.
- Aziz, A.-L., Giusiano, B., Joubert, S., Duprat, L., Didic, M., Gueriot, C., Koric, L., Boucraut, J., Felician, O., Ranjeva, J.-P., Guedj, E., Ceccaldi, M., 2017. Difference in imaging biomarkers of neurodegeneration between early and late-onset amnestic Alzheimer's disease. *Neurobiol. Aging* 54, 22–30. <https://doi.org/10.1016/j.neurobiolaging.2017.02.010>
- Azur, M.J., Stuart, E.A., Frangakis, C., Leaf, P.J., 2011. Multiple imputation by chained equations: what is it and how does it work? *Int. J. Methods Psychiatr. Res.* 20, 40–49.  
<https://doi.org/10.1002/mpr.329>
- Baccianella, S., Esuli, A., Sebastiani, F., 2009. Evaluation Measures for Ordinal Regression, in: 2009 Ninth International Conference on Intelligent Systems Design and Applications. Presented at the 2009 Ninth International Conference on Intelligent Systems Design and Applications, pp. 283–287. <https://doi.org/10.1109/ISDA.2009.230>

- Balestrieri, R., Bottou, L., LeCun, Y., 2022. The Effects of Regularization and Data Augmentation are Class Dependent. <https://doi.org/10.48550/arXiv.2204.03632>
- Baltrusaitis, T., Ahuja, C., Morency, L.-P., 2019. Multimodal Machine Learning: A Survey and Taxonomy. *IEEE Trans Pattern Anal Mach Intell* 41, 423–443.  
<https://doi.org/10.1109/TPAMI.2018.2798607>
- Barnett, A.G., van der Pols, J.C., Dobson, A.J., 2005. Regression to the mean: what it is and how to deal with it. *Int. J. Epidemiol.* 34, 215–220. <https://doi.org/10.1093/ije/dyh299>
- Battista, M.E.D., Gallucci, M., 2017. Overlap between frontotemporal dementia and Alzheimer's disease: case report. *Geriatr. Care* 3. <https://doi.org/10.4081/gc.2017.6866>
- Bengio, Y., Courville, A., Vincent, P., 2013. Representation Learning: A Review and New Perspectives. *IEEE Trans Pattern Anal Mach Intell* 35, 1798–1828. <https://doi.org/10.1109/TPAMI.2013.50>
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300.  
<https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Beretta, L., Santaniello, A., 2016. Nearest neighbor imputation algorithms: a critical evaluation. *BMC Med. Inform. Decis. Mak.* 16, 74. <https://doi.org/10.1186/s12911-016-0318-z>
- Blankemeier, L., Cohen, J.P., Kumar, A., Van Veen, D., Gardezi, S.J.S., Paschali, M., Chen, Z., Delbrouck, J.-B., Reis, E., Truyts, C., Bluethgen, C., Jensen, M.E.K., Ostmeier, S., Varma, M., Valanarasu, J.M.J., Fang, Z., Huo, Z., Nabulsi, Z., Ardila, D., Weng, W.-H., Amaro, E., Ahuja, N., Fries, J., Shah, N.H., Johnston, A., Boutin, R.D., Wentland, A., Langlotz, C.P., Hom, J., Gatidis, S., Chaudhari, A.S., 2024. Merlin: A Vision Language Foundation Model for 3D Computed Tomography. *Res. Sq. rs.3.rs-4546309*. <https://doi.org/10.21203/rs.3.rs-4546309/v1>
- Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* 2008, P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Bocancea, D.I., Svensson, A.L., van Loenhoud, A.C., Groot, C., Barkhof, F., Strandberg, O., Smith, R., La Joie, R., Rosen, H.J., Pontecorvo, M.J., Rabinovici, G.D., van der Flier, W.M., Hansson, O., Ossenkoppele, R., 2023. Determinants of cognitive and brain resilience to tau pathology: a longitudinal analysis. *Brain* 146, 3719–3734. <https://doi.org/10.1093/brain/awad100>
- Borsos, B., Allaart, C.G., van Halteren, A., 2024. Predicting stroke outcome: A case for multimodal deep learning methods with tabular and CT Perfusion data. *Artif. Intell. Med.* 147, 102719. <https://doi.org/10.1016/j.artmed.2023.102719>
- Bozek, J., Griffanti, L., Lau, S., Jenkinson, M., 2023. Normative models for neuroimaging markers: Impact of model selection, sample size and evaluation criteria. *Neuroimage* 268, 119864. <https://doi.org/10.1016/j.neuroimage.2023.119864>
- Bringas, S., Salomón, S., Duque, R., Lage, C., Montaña, J.L., 2020. Alzheimer's Disease stage identification using deep learning models. *J. Biomed. Inform.* 109, 103514. <https://doi.org/10.1016/j.jbi.2020.103514>
- Broderick, J.P., Adeoye, O., Elm, J., 2017. The Evolution of the Modified Rankin Scale and Its Use in Future Stroke Trials. *Stroke* 48, 2007–2012.  
<https://doi.org/10.1161/STROKEAHA.117.017866>
- Burgess, C.P., Higgins, I., Pal, A., Matthey, L., Watters, N., Desjardins, G., Lerchner, A., 2018. Understanding disentangling in \$\beta\$-VAE. <https://doi.org/10.48550/arXiv.1804.03599>
- Caldera, L., Cavinato, L., Cirone, A., Cama, I., Garbarino, S., Lodi, R., Tagliavini, F., Nigri, A., Francesco, S.D., Cappozzo, A., Piana, M., Ieva, F., 2025. DISARM++: Beyond scanner-free harmonization. <https://doi.org/10.48550/arXiv.2505.03715>
- Campbell, B.C.V., De Silva, D.A., Macleod, M.R., Coutts, S.B., Schwamm, L.H., Davis, S.M., Donnan, G.A., 2019. Ischaemic stroke. *Nat. Rev. Dis. Primer* 5, 70. <https://doi.org/10.1038/s41572-019-0118-8>

- Cao, W., Mirjalili, V., Raschka, S., 2020. Rank consistent ordinal regression for neural networks with application to age estimation. *Pattern Recognit. Lett.* 140, 325–331.  
<https://doi.org/10.1016/j.patrec.2020.11.008>
- Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., Nath, V., He, Y., Xu, Z., Hatamizadeh, A., Myronenko, A., Zhu, W., Liu, Y., Zheng, M., Tang, Y., Yang, I., Zephyr, M., Hashemian, B., Alle, S., Darestani, M.Z., Budd, C., Modat, M., Vercauteren, T., Wang, G., Li, Y., Hu, Y., Fu, Y., Gorman, B., Johnson, H., Genereaux, B., Erdal, B.S., Gupta, V., Diaz-Pinto, A., Dourson, A., Maier-Hein, L., Jaeger, P.F., Baumgartner, M., Kalpathy-Cramer, J., Flores, M., Kirby, J., Cooper, L.A.D., Roth, H.R., Xu, D., Bericat, D., Floca, R., Zhou, S.K., Shuaib, H., Farahani, K., Maier-Hein, K.H., Aylward, S., Dogra, P., Ourselin, S., Feng, A., 2022. MONAI: An open-source framework for deep learning in healthcare.  
<https://doi.org/10.48550/arXiv.2211.02701>
- Caruana, R.A., 1993. Multitask Learning: A Knowledge-Based Source of Inductive Bias. Elsevier, pp. 41–48. <https://doi.org/10.1016/B978-1-55860-307-3.50012-5>
- Chalos, V., van der Ende, N.A.M., Lingsma, H.F., Mulder, M.J.H.L., Venema, E., Dijkland, S.A., Berkhemer, O.A., Yoo, A.J., Broderick, J.P., Palesch, Y.Y., Yeatts, S.D., Roos, Y.B.W.E.M., van Oostenbrugge, R.J., van Zwam, W.H., Majolie, C.B.L.M., van der Lugt, A., Roozenbeek, B., Dippel, D.W.J., 2020. National Institutes of Health Stroke Scale. *Stroke* 51, 282–290.  
<https://doi.org/10.1161/STROKEAHA.119.026791>
- Chen, F., Ni, Y.-C., 2012. Magnetic resonance diffusion-perfusion mismatch in acute ischemic stroke: An update. *World J. Radiol.* 4, 63–74. <https://doi.org/10.4329/wjr.v4.i3.63>
- Chen, Z., Badrinarayanan, V., Lee, C.-Y., Rabinovich, A., 2018. GradNorm: Gradient Normalization for Adaptive Loss Balancing in Deep Multitask Networks, in: Proceedings of the 35th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 794–803.
- Chételat, G., Arbizu, J., Barthel, H., Garibotto, V., Law, I., Morbelli, S., van de Giessen, E., Agosta, F., Barkhof, F., Brooks, D.J., Carrillo, M.C., Dubois, B., Fjell, A.M., Frisoni, G.B., Hansson, O., Herholz, K., Hutton, B.F., Jack, C.R., Lammertsma, A.A., Landau, S.M., Minoshima, S., Nobili, F., Nordberg, A., Ossenkoppele, R., Oyen, W.J.G., Perani, D., Rabinovici, G.D., Scheltens, P., Villemagne, V.L., Zetterberg, H., Drzezga, A., 2020. Amyloid-PET and 18F-FDG-PET in the diagnostic investigation of Alzheimer's disease and other dementias. *Lancet Neurol.* 19, 951–962. [https://doi.org/10.1016/S1474-4422\(20\)30314-8](https://doi.org/10.1016/S1474-4422(20)30314-8)
- Cipolla, R., Gal, Y., Kendall, A., 2018. Multi-task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Presented at the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Salt Lake City, UT, USA, pp. 7482–7491.  
<https://doi.org/10.1109/CVPR.2018.00781>
- Cody, K.A., Langhough, R.E., Zammit, M.D., Clark, L., Chin, N., Christian, B.T., Betthauser, T.J., Johnson, S.C., 2024. Characterizing brain tau and cognitive decline along the amyloid timeline in Alzheimer's disease. *Brain J. Neurol.* 147, 2144–2157.  
<https://doi.org/10.1093/brain/awae116>
- Cole, J.H., 2020. Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiol. Aging* 92, 34–42.  
<https://doi.org/10.1016/j.neurobiolaging.2020.03.014>
- Cole, J.H., Marioni, R.E., Harris, S.E., Deary, I.J., 2019. Brain age and other bodily “ages”: implications for neuropsychiatry. *Mol. Psychiatry* 24, 266–281. <https://doi.org/10.1038/s41380-018-0098-1>
- Cole, J.H., Poudel, R.P.K., Tsagkrasoulis, D., Caan, M.W.A., Steves, C., Spector, T.D., Montana, G., 2017. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage* 163, 115–124.  
<https://doi.org/10.1016/j.neuroimage.2017.07.059>

- Contador, J., Vargas-Martínez, A.M., Sánchez-Valle, R., Trapero-Bertran, M., Lladó, A., 2023. Cost-effectiveness of Alzheimer's disease CSF biomarkers and amyloid-PET in early-onset cognitive impairment diagnosis. *Eur. Arch. Psychiatry Clin. Neurosci.* 273, 243–252. <https://doi.org/10.1007/s00406-022-01439-z>
- Coupé, P., Mansencal, B., Clément, M., Giraud, R., Denis de Senneville, B., Ta, V.-T., Lepetit, V., Manjon, J.V., 2020. AssemblyNet: A large ensemble of CNNs for 3D whole brain MRI segmentation. *NeuroImage* 219, 117026. <https://doi.org/10.1016/j.neuroimage.2020.117026>
- Crawshaw, M., 2020. Multi-Task Learning with Deep Neural Networks: A Survey. <https://doi.org/10.48550/arXiv.2009.09796>
- Crutch, S.J., Schott, J.M., Rabinovici, G.D., Murray, M., Snowden, J.S., van der Flier, W.M., Dickerson, B.C., Vandenberghe, R., Ahmed, S., Bak, T.H., Boeve, B.F., Butler, C., Cappa, S.F., Ceccaldi, M., de Souza, L.C., Dubois, B., Felician, O., Galasko, D., Graff-Radford, J., Graff-Radford, N.R., Hof, P.R., Krolak-Salmon, P., Lehmann, M., Magnin, E., Mendez, M.F., Nestor, P.J., Onyike, C.U., Pelak, V.S., Pijnenburg, Y., Primaitivo, S., Rossor, M.N., Ryan, N.S., Scheltens, P., Shakespeare, T.J., Suárez González, A., Tang-Wai, D.F., Yong, K.X.X., Carrillo, M., Fox, N.C., Alzheimer's Association ISTAART Atypical Alzheimer's Disease and Associated Syndromes Professional Interest Area, 2017. Consensus classification of posterior cortical atrophy. *Alzheimers Dement.* 13, 870–884. <https://doi.org/10.1016/j.jalz.2017.01.014>
- Cui, Y., Jia, M., Lin, T.-Y., Song, Y., Belongie, S., 2019. Class-Balanced Loss Based on Effective Number of Samples. <https://doi.org/10.48550/arXiv.1901.05555>
- Del Sole, A., Malaspina, S., Biasina, A.M., 2017. Magnetic resonance imaging and positron emission tomography in the diagnosis of neurodegenerative dementias. *Funct. Neurol.* 31, 205–215. <https://doi.org/10.11138/FNeur/2016.31.4.205>
- Destrieux, C., Fischl, B., Dale, A., Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage* 53, 1–15. <https://doi.org/10.1016/j.neuroimage.2010.06.010>
- Devi, G., Scheltens, P., 2018. Heterogeneity of Alzheimer's disease: consequence for drug trials? *Alzheimers Res. Ther.* 10, 122. <https://doi.org/10.1186/s13195-018-0455-y>
- Dias, M.F., Duarte, J.V., de Carvalho, P., Castelo-Branco, M., 2025. Unravelling pathological ageing with brain age gap estimation in Alzheimer's disease, diabetes and schizophrenia. *Brain Commun.* 7, fcaf109. <https://doi.org/10.1093/braincomms/fcaf109>
- Dickerson, B.C., Bakkour, A., Salat, D.H., Feczkó, E., Pacheco, J., Greve, D.N., Grodstein, F., Wright, C.I., Blacker, D., Rosas, H.D., Sperling, R.A., Atri, A., Growdon, J.H., Hyman, B.T., Morris, J.C., Fischl, B., Buckner, R.L., 2009. The Cortical Signature of Alzheimer's Disease: Regionally Specific Cortical Thinning Relates to Symptom Severity in Very Mild to Mild AD Dementia and is Detectable in Asymptomatic Amyloid-Positive Individuals. *Cereb. Cortex* 19, 497–510. <https://doi.org/10.1093/cercor/bhn113>
- Dickerson, B.C., Wolk, D.A., 2012. MRI cortical thickness biomarker predicts AD-like CSF and cognitive decline in normal adults. *Neurology* 78, 84–90. <https://doi.org/10.1212/WNL.0b013e31823efc6c>
- Dieng, A.B., Kim, Y., Rush, A.M., Blei, D.M., 2019. Avoiding Latent Variable Collapse with Generative Skip Models, in: Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics. Presented at the The 22nd International Conference on Artificial Intelligence and Statistics, PMLR, pp. 2397–2405.
- Dingga, R., Fraza, C.J., Bayer, J.M.M., Kia, S.M., Beckmann, C.F., Marquand, A.F., 2021. Normative modeling of neuroimaging data using generalized additive models of location scale and shape. <https://doi.org/10.1101/2021.06.14.448106>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2020. An Image is Worth 16x16

Words: Transformers for Image Recognition at Scale. Presented at the International Conference on Learning Representations.

- Dubois, B., von Arnim, C.A.F., Burnie, N., Bozeat, S., Cummings, J., 2023. Biomarkers in Alzheimer's disease: role in early and differential diagnosis and recognition of atypical variants. *Alzheimers Res. Ther.* 15, 175. <https://doi.org/10.1186/s13195-023-01314-6>
- Eickhoff, C.R., Hoffstaedter, F., Caspers, J., Reetz, K., Mathys, C., Dogan, I., Amunts, K., Schnitzler, A., Eickhoff, S.B., 2021. Advanced brain ageing in Parkinson's disease is related to disease duration and individual impairment. *Brain Commun.* 3, fcab191. <https://doi.org/10.1093/braincomms/fcab191>
- Fennema-Notestine, C., Hagler Jr., D.J., McEvoy, L.K., Fleisher, A.S., Wu, E.H., Karow, D.S., Dale, A.M., 2009. Structural MRI biomarkers for preclinical and mild Alzheimer's disease. *Hum. Brain Mapp.* 30, 3238–3253. <https://doi.org/10.1002/hbm.20744>
- Ferreira, D., Nordberg, A., Westman, E., 2020. Biological subtypes of Alzheimer disease: A systematic review and meta-analysis. *Neurology* 94, 436. <https://doi.org/10.1212/WNL.0000000000009058>
- Fleming, T.R., Harrington, D.P., 1981. A class of hypothesis tests for one and two sample censored survival data. *Commun. Stat. - Theory Methods* 10, 763–794. <https://doi.org/10.1080/03610928108828073>
- Folstein, M., Anthony, J.C., Parhad, I., Duffy, B., Gruenberg, E.M., 1985. The Meaning of Cognitive Impairment in the Elderly. *J. Am. Geriatr. Soc.* 33, 228–235. <https://doi.org/10.1111/j.1532-5415.1985.tb07109.x>
- Franke, K., Gaser, C., 2019. Ten Years of BrainAGE as a Neuroimaging Biomarker of Brain Aging: What Insights Have We Gained? *Front. Neurol.* 10.
- Franke, K., Ziegler, G., Klöppel, S., Gaser, C., Alzheimer's Disease Neuroimaging Initiative, 2010. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. *NeuroImage* 50, 883–892. <https://doi.org/10.1016/j.neuroimage.2010.01.005>
- Fraza, C.J., Dinga, R., Beckmann, C.F., Marquand, A.F., 2021. Warped Bayesian linear regression for normative modelling of big data. *NeuroImage* 245, 118715. <https://doi.org/10.1016/j.neuroimage.2021.118715>
- Fu, H., Li, C., Liu, X., Gao, J., Celikyilmaz, A., Carin, L., 2019. Cyclical Annealing Schedule: A Simple Approach to Mitigating KL Vanishing, in: Burstein, J., Doran, C., Solorio, T. (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Presented at the NAACL-HLT 2019, Association for Computational Linguistics, Minneapolis, Minnesota, pp. 240–250. <https://doi.org/10.18653/v1/N19-1021>
- G, T., B, C., M, Ebinger, Q, H., T, T., O, W., Js, K., L, B., Oc, S, S, W, S, C, A, T, Nd, F, I, G, M, R, T, E, M, K, M, Endres, Dw, K, V, D, Ag, S, Ds, L, Jb, F, J, F, C, G, 2011. DWI-FLAIR mismatch for the identification of patients with acute ischaemic stroke within 4·5 h of symptom onset (PRE-FLAIR): a multicentre observational study. *Lancet Neurol.* 10. [https://doi.org/10.1016/S1474-4422\(11\)70192-2](https://doi.org/10.1016/S1474-4422(11)70192-2)
- Gaser, C., Franke, K., Klöppel, S., Koutsouleris, N., Sauer, H., Alzheimer's Disease Neuroimaging Initiative, 2013. BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease. *PLoS One* 8, e67346. <https://doi.org/10.1371/journal.pone.0067346>
- Gautherot, M., Kuchcinski, G., Bordier, C., Sillaire, A.R., Delbeuck, X., Leroy, M., Leclerc, X., Pruvost, J.-P., Pasquier, F., Lopes, R., 2021. Longitudinal Analysis of Brain-Predicted Age in Amnestic and Non-amnestic Sporadic Early-Onset Alzheimer's Disease. *Front. Aging Neurosci.* 13.
- Gaviria, E., Eltayeb Hamid, A.H., 2024. Neuroimaging biomarkers for predicting stroke outcomes: A systematic review. *Health Sci. Rep.* 7, e2221. <https://doi.org/10.1002/hsr2.2221>

- Gewers, F.L., Ferreira, G.R., Arruda, H.F.D., Silva, F.N., Comin, C.H., Amancio, D.R., Costa, L.D.F., 2021. Principal Component Analysis: A Natural Approach to Data Exploration. *ACM Comput Surv* 54, 70:1-70:34. <https://doi.org/10.1145/3447755>
- Ghosh, S., Lippa, C.F., 2015. Clinical Subtypes of Frontotemporal Dementia. *Am. J. Alzheimers Dis. Other Demen.* 30, 653–661. <https://doi.org/10.1177/1533317513494442>
- Gorno-Tempini, M.L., Dronkers, N.F., Rankin, K.P., Ogar, J.M., Phengrasamy, L., Rosen, H.J., Johnson, J.K., Weiner, M.W., Miller, B.L., 2004. Cognition and anatomy in three variants of primary progressive aphasia. *Ann. Neurol.* 55, 335–346. <https://doi.org/10.1002/ana.10825>
- Gorno-Tempini, M.L., Hillis, A.E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S.F., Ogar, J.M., Rohrer, J.D., Black, S., Boeve, B.F., Manes, F., Dronkers, N.F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B.L., Knopman, D.S., Hodges, J.R., Mesulam, M.M., Grossman, M., 2011. Classification of primary progressive aphasia and its variants. *Neurology* 76, 1006–1014. <https://doi.org/10.1212/WNL.0b013e31821103e6>
- Goveas, J., O'Dwyer, L., Mascalchi, M., Cosottini, M., Diciotti, S., De Santis, S., Passamonti, L., Tessa, C., Toschi, N., Giannelli, M., 2015. Diffusion-MRI in neurodegenerative disorders. *Magn. Reson. Imaging* 33, 853–876. <https://doi.org/10.1016/j.mri.2015.04.006>
- Groot, C., Smith, R., Collij, L.E., Mastenbroek, S.E., Stomrud, E., Binette, A.P., Leuzy, A., Palmqvist, S., Mattsson-Carlsgren, N., Strandberg, O., Cho, H., Lyoo, C.H., Frisoni, G.B., Peretti, D.E., Garibotto, V., La Joie, R., Soleimani-Meigooni, D.N., Rabinovici, G., Ossenkoppele, R., Hansson, O., 2024. Tau Positron Emission Tomography for Predicting Dementia in Individuals With Mild Cognitive Impairment. *JAMA Neurol.* 81, 845–856. <https://doi.org/10.1001/jamaneurol.2024.1612>
- Guo, M., Haque, A., Huang, D.-A., Yeung, S., Fei-Fei, L., 2018. Dynamic Task Prioritization for Multitask Learning, in: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), *Computer Vision – ECCV 2018, Lecture Notes in Computer Science*. Springer International Publishing, Cham, pp. 282–299. [https://doi.org/10.1007/978-3-030-01270-0\\_17](https://doi.org/10.1007/978-3-030-01270-0_17)
- Guo, X., Ding, Y., Xu, W., Wang, D., Yu, H., Lin, Y., Chang, S., Zhang, Q., Zhang, Y., 2024. Predicting brain age gap with radiomics and automl: A Promising approach for age-Related brain degeneration biomarkers. *J. Neuroradiol.* 51, 265–273. <https://doi.org/10.1016/j.neurad.2023.09.002>
- Habes, M., Grothe, M.J., Tunc, B., McMillan, C., Wolk, D.A., Davatzikos, C., 2020. Disentangling Heterogeneity in Alzheimer's Disease and Related Dementias Using Data-Driven Methods. *Biol. Psychiatry* 88, 70–82. <https://doi.org/10.1016/j.biopsych.2020.01.016>
- Harciarek, M., Jodzio, K., 2005. Neuropsychological differences between frontotemporal dementia and Alzheimer's disease: a review. *Neuropsychol. Rev.* 15, 131–145. <https://doi.org/10.1007/s11065-005-7093-4>
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Identity Mappings in Deep Residual Networks.
- Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Mohamed, S., Lerchner, A., 2022. beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework. Presented at the International Conference on Learning Representations.
- Hinton, G.E., 2012. A Practical Guide to Training Restricted Boltzmann Machines, in: Montavon, G., Orr, G.B., Müller, K.-R. (Eds.), *Neural Networks: Tricks of the Trade: Second Edition*. Springer, Berlin, Heidelberg, pp. 599–619. [https://doi.org/10.1007/978-3-642-35289-8\\_32](https://doi.org/10.1007/978-3-642-35289-8_32)
- Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the Dimensionality of Data with Neural Networks. *Science* 313, 504–507. <https://doi.org/10.1126/science.1127647>
- Holzinger, A., Langs, G., Denk, H., Zatloukal, K., Müller, H., 2019. Causability and explainability of artificial intelligence in medicine. *WIREs Data Min. Knowl. Discov.* 9, e1312. <https://doi.org/10.1002/widm.1312>
- Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M., 2022. SynthStrip: skull-stripping for any brain image. *NeuroImage* 260, 119474. <https://doi.org/10.1016/j.neuroimage.2022.119474>

- Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24, 417–441. <https://doi.org/10.1037/h0071325>
- Huang, Y., Lin, J., Zhou, C., Yang, H., Huang, L., 2022. Modality Competition: What Makes Joint Training of Multi-modal Network Fail in Deep Learning? (Provably), in: Proceedings of the 39th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 9226–9259.
- Ilse, M., Tomczak, J., Welling, M., 2018. Attention-based Deep Multiple Instance Learning, in: Proceedings of the 35th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 2127–2136.
- Ilse, M., Tomczak, J.M., Louizos, C., Welling, M., 2020. DIVA: Domain Invariant Variational Autoencoders, in: Proceedings of the Third Conference on Medical Imaging with Deep Learning. Presented at the Medical Imaging with Deep Learning, PMLR, pp. 322–348.
- Ioannou, K., Bucci, M., Tzortzakis, A., Savitcheva, I., Nordberg, A., Chiotis, K., 2025. Tau PET positivity predicts clinically relevant cognitive decline driven by Alzheimer's disease compared to comorbid cases; proof of concept in the ADNI study. *Mol. Psychiatry* 30, 587–599. <https://doi.org/10.1038/s41380-024-02672-9>
- Isensee, F., Schell, M., Pflueger, I., Brugnara, G., Bonekamp, D., Neuberger, U., Wick, A., Schlemmer, H.-P., Heiland, S., Wick, W., Bendszus, M., Maier-Hein, K.H., Kickingereder, P., 2019. Automated brain extraction of multisequence MRI using artificial neural networks. *Hum. Brain Mapp.* 40, 4952–4964. <https://doi.org/10.1002/hbm.24750>
- Jazayeri, A., Liang, O.S., Yang, C.C., 2020. Imputation of Missing Data in Electronic Health Records Based on Patients' Similarities. *J. Healthc. Inform. Res.* 4, 295–307. <https://doi.org/10.1007/s41666-020-00073-5>
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* 17, 825–841. <https://doi.org/10.1006/nimg.2002.1132>
- Jiang, B., Pham, N., van Staalduin, E.K., Liu, Y., Nazari-Farsani, S., Sanaat, A., van Voorst, H., Fettahoglu, A., Kim, D., Ouyang, J., Kumar, A., Srivatsan, A., Hussein, R., Lansberg, M.G., Boada, F., Zaharchuk, G., 2025. Deep Learning Applications in Imaging of Acute Ischemic Stroke: A Systematic Review and Narrative Summary. *Radiology* 315, e240775. <https://doi.org/10.1148/radiol.240775>
- Jolliffe, I.T. (Ed.), 2002. Principal Component Analysis and Factor Analysis, in: Principal Component Analysis. Springer, New York, NY, pp. 150–166. [https://doi.org/10.1007/0-387-22440-8\\_7](https://doi.org/10.1007/0-387-22440-8_7)
- Jung, H.-S., Lee, E.-J., Chang, D.-I., Cho, H.J., Lee, J., Cha, J.-K., Park, M.-S., Yu, K.H., Jung, J.-M., Ahn, S.H., Kim, D.-E., Lee, J.H., Hong, K.-S., Sohn, S.-I., Park, K.-P., Kwon, S.U., Kim, J.S., Chang, J.Y., Kim, B.J., Kang, D.-W., KOSNI Investigators, 2024. A Multimodal Ensemble Deep Learning Model for Functional Outcome Prognosis of Stroke Patients. *J. Stroke* 26, 312–320. <https://doi.org/10.5853/jos.2023.03426>
- Kang, S., Kim, S.-W., Seong, J.-K., 2024. Disentangling brain atrophy heterogeneity in Alzheimer's disease: A deep self-supervised approach with interpretable latent space. *NeuroImage* 297, 120737. <https://doi.org/10.1016/j.neuroimage.2024.120737>
- Katan, M., Luft, A., 2018. Global Burden of Stroke. *Semin. Neurol.* 38, 208–211. <https://doi.org/10.1055/s-0038-1649503>
- Kim, J.P., Kim, J., Park, Y.H., Park, S.B., Lee, J.S., Yoo, S., Kim, E.-J., Kim, H.J., Na, D.L., Brown, J.A., Lockhart, S.N., Seo, S.W., Seong, J.-K., 2019. Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. *NeuroImage Clin.* 23, 101811. <https://doi.org/10.1016/j.nicl.2019.101811>
- Kim, P.J., Kim, D., Lee, J., Kim, H.C., Seo, J.H., Lee, S.Y., Kwon, D.H., Park, H., Yoo, J., Park, S., 2025. Deep learning-based classification of diffusion-weighted imaging-fluid-attenuated inversion recovery mismatch. *Sci. Rep.* 15, 5924. <https://doi.org/10.1038/s41598-025-90214-w>

- Kingma, D.P., Welling, M., 2019. An Introduction to Variational Autoencoders. *Found. Trends® Mach. Learn.* 12, 307–392. <https://doi.org/10.1561/2200000056>
- Kingma, D.P., Welling, M., 2013. Auto-Encoding Variational Bayes.
- Kirkwood, T.B.L., 2005. Understanding the Odd Science of Aging. *Cell* 120, 437–447. <https://doi.org/10.1016/j.cell.2005.01.027>
- Klushyn, A., Chen, N., Kurle, R., Cseke, B., van der Smagt, P., 2019. Learning Hierarchical Priors in VAEs. <https://doi.org/10.48550/arXiv.1905.04982>
- Koedam, E.L.G.E., Lauffer, V., van der Vlies, A.E., van der Flier, W.M., Scheltens, P., Pijnenburg, Y.A.L., 2010. Early-versus late-onset Alzheimer's disease: more than age alone. *J. Alzheimers Dis. JAD* 19, 1401–1408. <https://doi.org/10.3233/JAD-2010-1337>
- Koenig, A.M., Nobuhara, C.K., Williams, V.J., Arnold, S.E., 2018. Biomarkers in Alzheimer's, Frontotemporal, Lewy Body, and Vascular Dementias. *Focus J. Life Long Learn. Psychiatry* 16, 164–172. <https://doi.org/10.1176/appi.focus.20170048>
- Kuchcinski, G., Patin, L., Lopes, R., Leroy, M., Delbeuck, X., Rollin-Sillaire, A., Lebouvier, T., Wang, Y., Spincemaille, P., Tourdias, T., Hacein-Bey, L., Devos, D., Pasquier, F., Leclerc, X., Pruvost, J.-P., Verclytte, S., 2023. Quantitative susceptibility mapping demonstrates different patterns of iron overload in subtypes of early-onset Alzheimer's disease. *Eur. Radiol.* 33, 184–195. <https://doi.org/10.1007/s00330-022-09014-9>
- Labrak, Y., Bazoge, A., Dufour, R., Rovier, M., Morin, E., Daille, B., Gourraud, P.-A., 2023. DrBERT: A Robust Pre-trained Model in French for Biomedical and Clinical domains. <https://doi.org/10.48550/arXiv.2304.00958>
- Landau, S.M., Harvey, D., Madison, C.M., Koeppe, R.A., Reiman, E.M., Foster, N.L., Weiner, M.W., Jagust, W.J., 2011. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. *Neurobiol. Aging* 32, 1207–1218. <https://doi.org/10.1016/j.neurobiolaging.2009.07.002>
- Larochelle, H., Bengio, Y., 2008. Classification using discriminative restricted Boltzmann machines, in: Proceedings of the 25th International Conference on Machine Learning, ICML '08. Association for Computing Machinery, New York, NY, USA, pp. 536–543. <https://doi.org/10.1145/1390156.1390224>
- Lee, M., Pavlovic, V., 2021. Private-Shared Disentangled Multimodal VAE for Learning of Latent Representations, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Presented at the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, Nashville, TN, USA, pp. 1692–1700. <https://doi.org/10.1109/CVPRW53098.2021.00185>
- Le-Khac, P.H., Healy, G., Smeaton, A.F., 2020. Contrastive Representation Learning: A Framework and Review. *IEEE Access* 8, 193907–193934. <https://doi.org/10.1109/ACCESS.2020.3031549>
- Leonardsen, E.H., Peng, H., Kaufmann, T., Agartz, I., Andreassen, O.A., Celius, E.G., Espeseth, T., Harbo, H.F., Høgestøl, E.A., Lange, A.-M. de, Marquand, A.F., Vidal-Piñeiro, D., Roe, J.M., Selbæk, G., Sørensen, Ø., Smith, S.M., Westlye, L.T., Wolfers, T., Wang, Y., 2022. Deep neural networks learn general and clinically relevant representations of the ageing brain. *NeuroImage* 256, 119210. <https://doi.org/10.1016/j.neuroimage.2022.119210>
- Levin, F., Ferreira, D., Lange, C., Dyrba, M., Westman, E., Buchert, R., Teipel, S.J., Grothe, M.J., for the Alzheimer's Disease Neuroimaging Initiative, 2021. Data-driven FDG-PET subtypes of Alzheimer's disease-related neurodegeneration. *Alzheimers Res. Ther.* 13, 49. <https://doi.org/10.1186/s13195-021-00785-9>
- Li, H., Habes, M., Wolk, D.A., Fan, Y., 2019. A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal MRI. *Alzheimers Dement. J. Alzheimers Assoc.* 15, 1059–1070. <https://doi.org/10.1016/j.jalz.2019.02.007>
- Liang, H., Zhang, F., Niu, X., 2019. Investigating systematic bias in brain age estimation with application to post-traumatic stress disorders. *Hum. Brain Mapp.* 40, 3143–3152. <https://doi.org/10.1002/hbm.24588>

- Lim, C.Y., Tan, P.C., Koh, C., Koh, E., Guo, H., Yusoff, N.D., See, C.Q., Tan, T., 2015. Beery-Buktenica Developmental Test of Visual-Motor Integration (Beery-VMII): lessons from exploration of cultural variations in visual-motor integration performance of preschoolers. *Child Care Health Dev.* 41, 213–221. <https://doi.org/10.1111/cch.12190>
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2018. Focal Loss for Dense Object Detection. <https://doi.org/10.48550/arXiv.1708.02002>
- Lindeboom, J., Schmand, B., Tulner, L., Walstra, G., Jonker, C., 2002. Visual association test to detect early dementia of the Alzheimer type. *J. Neurol. Neurosurg. Psychiatry* 73, 126–133. <https://doi.org/10.1136/jnnp.73.2.126>
- Liu, S., Johns, E., Davison, A.J., 2019. End-To-End Multi-Task Learning With Attention, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1871–1880. <https://doi.org/10.1109/CVPR.2019.00197>
- Liu, S., Yap, P.-T., 2024. Learning multi-site harmonization of magnetic resonance images without traveling human phantoms. *Commun. Eng.* 3, 6. <https://doi.org/10.1038/s44172-023-00140-w>
- Liu, Y., Yu, Y., Ouyang, J., Jiang, B., Yang, G., Ostmeier, S., Wintermark, M., Michel, P., Liebeskind, D.S., Lansberg, M.G., Albers, G.W., Zaharchuk, G., 2023. Functional Outcome Prediction in Acute Ischemic Stroke Using a Fused Imaging and Clinical Deep Learning Model. *Stroke* 54, 2316–2327. <https://doi.org/10.1161/STROKEAHA.123.044072>
- Ljubenkov, P.A., Miller, B.L., 2016. A Clinical Guide to Frontotemporal Dementias. *Focus J. Life Long Learn. Psychiatry* 14, 448–464. <https://doi.org/10.1176/appi.focus.20160018>
- Lloyd, S., 1982. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* 28, 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
- Locatello, F., Bauer, S., Lucic, M., Raetsch, G., Gelly, S., Schölkopf, B., Bachem, O., 2019. Challenging Common Assumptions in the Unsupervised Learning of Disentangled Representations, in: Proceedings of the 36th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 4114–4124.
- Loshchilov, I., Hutter, F., 2019. Decoupled Weight Decay Regularization. <https://doi.org/10.48550/arXiv.1711.05101>
- Louiset, R., Duchesnay, E., Antoine, G., Dufumier, B., Gori, P., 2024. SepVAE: a contrastive VAE to separate pathological patterns from healthy ones, in: Proceedings of The 7nd International Conference on Medical Imaging with Deep Learning. Presented at the Medical Imaging with Deep Learning, PMLR, pp. 918–936.
- Löwe, L.C., Gaser, C., Franke, K., Alzheimer's Disease Neuroimaging Initiative, 2016. The Effect of the APOE Genotype on Individual BrainAGE in Normal Aging, Mild Cognitive Impairment, and Alzheimer's Disease. *PloS One* 11, e0157514. <https://doi.org/10.1371/journal.pone.0157514>
- Luo, G., Xie, W., Gao, R., Zheng, T., Chen, L., Sun, H., 2023. Unsupervised anomaly detection in brain MRI: Learning abstract distribution from massive healthy brains. *Comput. Biol. Med.* 154, 106610. <https://doi.org/10.1016/j.combiomed.2023.106610>
- Ma, D., Stocks, J., Rosen, H., Kantarci, K., Lockhart, S.N., Bateman, J.R., Craft, S., Gurcan, M.N., Popuri, K., Beg, M.F., Wang, L., on behalf of the ALLFTD consortium, 2024. Differential diagnosis of frontotemporal dementia subtypes with explainable deep learning on structural MRI. *Front. Neurosci.* 18. <https://doi.org/10.3389/fnins.2024.1331677>
- Ma, J., Zhao, Z., Yi, X., Chen, J., Hong, L., Chi, E.H., 2018. Modeling Task Relationships in Multi-task Learning with Multi-gate Mixture-of-Experts, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Presented at the KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, London United Kingdom, pp. 1930–1939. <https://doi.org/10.1145/3219819.3220007>
- Madhavan, A., Whitwell, J.L., Weigand, S.D., Duffy, J.R., Strand, E.A., Machulda, M.M., Tosakulwong, N., Senjem, M.L., Gunter, J.L., Lowe, V.J., Petersen, R.C., Clifford R Jack, J., Josephs, K.A.,

2013. FDG PET and MRI in Logopenic Primary Progressive Aphasia versus Dementia of the Alzheimer's Type. *PLoS ONE* 8, e62471. <https://doi.org/10.1371/journal.pone.0062471>
- Manjón, J.V., Coupé, P., 2016. volBrain: An Online MRI Brain Volumetry System. *Front. Neuroinformatics* 10.
- Marquand, A.F., Kia, S.M., Zabihi, M., Wolfers, T., Buitelaar, J.K., Beckmann, C.F., 2019. Conceptualizing mental disorders as deviations from normative functioning. *Mol. Psychiatry* 24, 1415–1424. <https://doi.org/10.1038/s41380-019-0441-1>
- Marquand, A.F., Rezek, I., Buitelaar, J., Beckmann, C.F., 2016. Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. *Biol. Psychiatry* 80, 552–561. <https://doi.org/10.1016/j.biopsych.2015.12.023>
- Matta, S., Lamard, M., Zhang, P., Le Guilcher, A., Borderie, L., Cochener, B., Quellec, G., 2024. A systematic review of generalization research in medical image classification. *Comput. Biol. Med.* 183, 109256. <https://doi.org/10.1016/j.combiomed.2024.109256>
- Maureille, A., Adeline, R.-S., Chen, Y., Dubois, B., Didier, H., Pasquier, F., 2017. Suivi à deux ans de la cohorte de malades Alzheimer jeunes (COMAJ) à Lille. *Rev. Neurol. (Paris)*, Journées de Neurologie de langue française 2017 173, S39. <https://doi.org/10.1016/j.neurol.2017.01.008>
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M., 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 34, 939–944. <https://doi.org/10.1212/wnl.34.7.939>
- McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack Jr., C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morris, J.C., Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub, S., Phelps, C.H., 2011. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 263–269. <https://doi.org/10.1016/j.jalz.2011.03.005>
- Mendez, M.F., 2017. Early-Onset Alzheimer's Disease. *Neurol. Clin.* 35, 263–281. <https://doi.org/10.1016/j.ncl.2017.01.005>
- Mendez, M.F., 2006. The accurate diagnosis of early-onset dementia. *Int. J. Psychiatry Med.* 36, 401–412. <https://doi.org/10.2190/Q6J4-R143-P630-KW41>
- Metz, A., Zeighami, Y., Ducharme, S., Villeneuve, S., Dadar, M., 2025. Frontotemporal dementia subtyping using machine learning, multivariate statistics and neuroimaging. *Brain Commun.* 7, fcaf065. <https://doi.org/10.1093/braincomms/fcaf065>
- Migliaccio, R., Agosta, F., Possin, K.L., Canu, E., Filippi, M., Rabinovici, G.D., Rosen, H.J., Miller, B.L., Gorno-Tempini, M.L., 2015. Mapping the Progression of Atrophy in Early- and Late-Onset Alzheimer's Disease. *J. Alzheimers Dis. JAD* 46, 351–364. <https://doi.org/10.3233/JAD-142292>
- Misra, C., Fan, Y., Davatzikos, C., 2009. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI. *NeuroImage* 44, 1415–1422. <https://doi.org/10.1016/j.neuroimage.2008.10.031>
- Murtagh, F., Contreras, P., 2012. Algorithms for hierarchical clustering: an overview. *WIREs Data Min. Knowl. Discov.* 2, 86–97. <https://doi.org/10.1002/widm.53>
- Musa, G., Slachevsky, A., Muñoz-Neira, C., Méndez, C., Villagra, R., González-Billault, C., Ibáñez, A., Hornberger, M., Lillo, P., 2020. Alzheimer's Disease or Behavioral Variant Frontotemporal Dementia? Review of Key Points Toward an Accurate Clinical and Neuropsychological Diagnosis. *J. Alzheimers Dis. JAD* 73, 833–848. <https://doi.org/10.3233/JAD-190924>
- Nguyen, H.-D., Clément, M., Planche, V., Mansencal, B., Coupé, P., 2023. Deep grading for MRI-based differential diagnosis of Alzheimer's disease and Frontotemporal dementia. *Artif. Intell. Med.* 144, 102636. <https://doi.org/10.1016/j.artmed.2023.102636>
- Nishi, H., Oishi, N., Ishii, A., Ono, I., Ogura, T., Sunohara, T., Chihara, H., Fukumitsu, R., Okawa, M., Yamana, N., Imamura, H., Sadamasa, N., Hatano, T., Nakahara, I., Sakai, N., Miyamoto, S.,

2020. Deep Learning-Derived High-Level Neuroimaging Features Predict Clinical Outcomes for Large Vessel Occlusion. *Stroke* 51, 1484–1492.  
<https://doi.org/10.1161/STROKEAHA.119.028101>
- O'Bryant, S.E., Waring, S.C., Cullum, C.M., Hall, J., Lacritz, L., Massman, P.J., Lupo, P.J., Reisch, J.S., Doody, R., 2008a. Staging Dementia Using Clinical Dementia Rating Scale Sum of Boxes Scores. *Arch. Neurol.* 65, 1091–1095. <https://doi.org/10.1001/archneur.65.8.1091>
- O'Bryant, S.E., Waring, S.C., Cullum, C.M., Hall, J., Lacritz, L., Massman, P.J., Lupo, P.J., Reisch, J.S., Doody, R., Texas Alzheimer's Research Consortium, 2008b. Staging Dementia Using Clinical Dementia Rating Scale Sum of Boxes Scores: A Texas Alzheimer's Research Consortium Study. *Arch. Neurol.* 65, 1091–1095. <https://doi.org/10.1001/archneur.65.8.1091>
- Ossenkoppele, R., Cohn-Sheehy, B.I., La Joie, R., Vogel, J.W., Möller, C., Lehmann, M., van Berckel, B.N.M., Seeley, W.W., Pijnenburg, Y.A., Gorno-Tempini, M.L., Kramer, J.H., Barkhof, F., Rosen, H.J., van der Flier, W.M., Jagust, W.J., Miller, B.L., Scheltens, P., Rabinovici, G.D., 2015a. Atrophy patterns in early clinical stages across distinct phenotypes of Alzheimer's disease. *Hum. Brain Mapp.* 36, 4421–4437. <https://doi.org/10.1002/hbm.22927>
- Ossenkoppele, R., Hansson, O., 2021. Towards clinical application of tau PET tracers for diagnosing dementia due to Alzheimer's disease. *Alzheimers Dement.* 17, 1998–2008.  
<https://doi.org/10.1002/alz.12356>
- Ossenkoppele, R., Lyoo, C.H., Jester-Broms, J., Sudre, C.H., Cho, H., Ryu, Y.H., Choi, J.Y., Smith, R., Strandberg, O., Palmqvist, S., Kramer, J., Boxer, A.L., Gorno-Tempini, M.L., Miller, B.L., La Joie, R., Rabinovici, G.D., Hansson, O., 2020. Assessment of Demographic, Genetic, and Imaging Variables Associated With Brain Resilience and Cognitive Resilience to Pathological Tau in Patients With Alzheimer Disease. *JAMA Neurol.* 77, 632–642.  
<https://doi.org/10.1001/jamaneurol.2019.5154>
- Ossenkoppele, R., Pijnenburg, Y.A.L., Perry, D.C., Cohn-Sheehy, B.I., Scheltens, N.M.E., Vogel, J.W., Kramer, J.H., van der Vlies, A.E., Joie, R.L., Rosen, H.J., van der Flier, W.M., Grinberg, L.T., Rozemuller, A.J., Huang, E.J., van Berckel, B.N.M., Miller, B.L., Barkhof, F., Jagust, W.J., Scheltens, P., Seeley, W.W., Rabinovici, G.D., 2015b. The behavioural/dysexecutive variant of Alzheimer's disease: clinical, neuroimaging and pathological features. *Brain* 138, 2732–2749. <https://doi.org/10.1093/brain/awv191>
- Ouyang, J., Zhao, Q., Adeli, E., Zaharchuk, G., Pohl, K.M., 2022. Disentangling Normal Aging from Severity of Disease via Weak Supervision on Longitudinal MRI. *IEEE Trans. Med. Imaging* 41, 2558–2569. <https://doi.org/10.1109/TMI.2022.3166131>
- Pagani, M., Nobili, F., Morbelli, S., Arnaldi, D., Giuliani, A., Öberg, J., Girtler, N., Brugnolo, A., Picco, A., Bauckneht, M., Piva, R., Chincarini, A., Sambuceti, G., Jonsson, C., De Carli, F., 2017. Early identification of MCI converting to AD: a FDG PET study. *Eur. J. Nucl. Med. Mol. Imaging* 44, 2042–2052. <https://doi.org/10.1007/s00259-017-3761-x>
- Parimbelli, E., Marini, S., Sacchi, L., Bellazzi, R., 2018. Patient similarity for precision medicine: A systematic review. *J. Biomed. Inform.* 83, 87–96. <https://doi.org/10.1016/j.jbi.2018.06.001>
- Park, J.-Y., Na, H.K., Kim, S., Kim, H., Kim, H.J., Seo, S.W., Na, D.L., Han, C.E., Seong, J.-K., 2017. Robust Identification of Alzheimer's Disease subtypes based on cortical atrophy patterns. *Sci. Rep.* 7, 43270. <https://doi.org/10.1038/srep43270>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Peet, B.T., Spina, S., Mundada, N., La Joie, R., 2021. Neuroimaging in Frontotemporal Dementia: Heterogeneity and Relationships with Underlying Neuropathology. *Neurotherapeutics* 18, 728–752. <https://doi.org/10.1007/s13311-021-01101-x>
- Perneczky, R., Wagenpfeil, S., Komossa, K., Grimmer, T., Diehl, J., Kurz, A., 2006. Mapping scores onto stages: mini-mental state examination and clinical dementia rating. *Am. J. Geriatr.*

- Psychiatry Off. J. Am. Assoc. Geriatr. Psychiatry 14, 139–144.  
<https://doi.org/10.1097/01.JGP.0000192478.82189.a8>
- Petersen, K.K., Nallapu, B.T., Lipton, R.B., Grober, E., Ezzati, A., 2024. MRI-guided clustering of patients with mild dementia due to Alzheimer's disease using self-organizing maps. Neuroimage Rep. 4, 100227. <https://doi.org/10.1016/j.ynirp.2024.100227>
- Pinaya, W.H.L., Mechelli, A., Sato, J.R., 2018. Using deep autoencoders to identify abnormal brain structural patterns in neuropsychiatric disorders: A large-scale multi-sample study. Hum. Brain Mapp. 40, 944–954. <https://doi.org/10.1002/hbm.24423>
- Pinaya, W.H.L., Scarpazza, C., Garcia-Dias, R., Vieira, S., Baecker, L., F da Costa, P., Redolfi, A., Frisoni, G.B., Pievani, M., Calhoun, V.D., Sato, J.R., Mechelli, A., 2021. Using normative modelling to detect disease progression in mild cognitive impairment and Alzheimer's disease in a cross-sectional multi-cohort study. Sci. Rep. 11, 15746. <https://doi.org/10.1038/s41598-021-95098-0>
- Poulakis, K., Pereira, J.B., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Lovestone, S., Simmons, A., Wahlund, L.-O., Westman, E., 2018. Heterogeneous patterns of brain atrophy in Alzheimer's disease. Neurobiol. Aging 65, 98–108.  
<https://doi.org/10.1016/j.neurobiolaging.2018.01.009>
- Poulakis, K., Pereira, J.B., Muehlboeck, J.-S., Wahlund, L.-O., Smedby, Ö., Volpe, G., Masters, C.L., Ames, D., Niimi, Y., Iwatsubo, T., Ferreira, D., Westman, E., 2022. Multi-cohort and longitudinal Bayesian clustering study of stage and subtype in Alzheimer's disease. Nat. Commun. 13, 4566. <https://doi.org/10.1038/s41467-022-32202-6>
- Powers, W.J., Rabinstein, A.A., Ackerson, T., Adeoye, O.M., Bambakidis, N.C., Becker, K., Biller, J., Brown, M., Demaerschalk, B.M., Hoh, B., Jauch, E.C., Kidwell, C.S., Leslie-Mazwi, T.M., Ovbiagele, B., Scott, P.A., Sheth, K.N., Southerland, A.M., Summers, D.V., Tirschwell, D.L., American Heart Association Stroke Council, 2018. 2018 Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. Stroke 49, e46–e110.  
<https://doi.org/10.1161/STR.0000000000000158>
- Privo, J.P., Kuchcinski, G., Bretzner, M., Krystal, S., Dumont, J., Guellec, B.L., Hacein-Bey, L., Lopes, R., 2025. The ARIANES initiative: A vision for a structured regional MRI network in the Hauts-de-France. J. Neuroradiol. 52, 101238. <https://doi.org/10.1016/j.neurad.2024.101238>
- Punjabi, A., Martersteck, A., Wang, Y., Parrish, T.B., Katsaggelos, A.K., Initiative, and the A.D.N., 2019. Neuroimaging modality fusion in Alzheimer's classification using convolutional neural networks. PLOS ONE 14, e0225759. <https://doi.org/10.1371/journal.pone.0225759>
- Rabinovici, G.D., Rosen, H.J., Alkalay, A., Kornak, J., Furst, A.J., Agarwal, N., Mormino, E.C., O'Neil, J.P., Janabi, M., Karydas, A., Growdon, M.E., Jang, J.Y., Huang, E.J., Dearmond, S.J., Trojanowski, J.Q., Grinberg, L.T., Gorno-Tempini, M.L., Seeley, W.W., Miller, B.L., Jagust, W.J., 2011. Amyloid vs FDG-PET in the differential diagnosis of AD and FTLD. Neurology 77, 2034–2042. <https://doi.org/10.1212/WNL.0b013e31823b9c5e>
- Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I., 2021. Learning Transferable Visual Models From Natural Language Supervision, in: Proceedings of the 38th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 8748–8763.
- Rascovsky, K., Hodges, J.R., Knopman, D., Mendez, M.F., Kramer, J.H., Neuhaus, J., van Swieten, J.C., Seelaar, H., Doppler, E.G.P., Onyike, C.U., Hillis, A.E., Josephs, K.A., Boeve, B.F., Kertesz, A., Seeley, W.W., Rankin, K.P., Johnson, J.K., Gorno-Tempini, M.-L., Rosen, H., Prioleau-Latham, C.E., Lee, A., Kipps, C.M., Lillo, P., Piguet, O., Rohrer, J.D., Rossor, M.N., Warren, J.D., Fox, N.C., Galasko, D., Salmon, D.P., Black, S.E., Mesulam, M., Weintraub, S., Dickerson, B.C., Diehl-Schmid, J., Pasquier, F., Deramecourt, V., Lebert, F., Pijnenburg, Y., Chow, T.W., Manes, F., Grafman, J., Cappa, S.F., Freedman, M., Grossman, M., Miller, B.L., 2011. Sensitivity of

- revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain J. Neurol.* 134, 2456–2477. <https://doi.org/10.1093/brain/awr179>
- Rathore, S., Habes, M., Iftikhar, M.A., Shacklett, A., Davatzikos, C., 2017. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage* 155, 530–548. <https://doi.org/10.1016/j.neuroimage.2017.03.057>
- Reynolds, D., 2015. Gaussian Mixture Models, in: Encyclopedia of Biometrics. Springer, Boston, MA, pp. 827–832. [https://doi.org/10.1007/978-1-4899-7488-4\\_196](https://doi.org/10.1007/978-1-4899-7488-4_196)
- Rezende, D.J., Viola, F., 2018. Taming VAEs. <https://doi.org/10.48550/arXiv.1810.00597>
- Risacher, S., Saykin, A., 2013. Neuroimaging Biomarkers of Neurodegenerative Diseases and Dementia. *Semin. Neurol.* 33, 386–416. <https://doi.org/10.1055/s-0033-1359312>
- Risacher, S.L., Anderson, W.H., Charil, A., Castelluccio, P.F., Shcherbinin, S., Saykin, A.J., Schwarz, A.J., Alzheimer's Disease Neuroimaging Initiative, 2017. Alzheimer disease brain atrophy subtypes are associated with cognition and rate of decline. *Neurology* 89, 2176–2186. <https://doi.org/10.1212/WNL.0000000000004670>
- Roca, V., Kuchcinski, G., Pruvost, J.-P., Manouvriez, D., Lopes, R., 2025. IGUANE: A 3D generalizable CycleGAN for multicenter harmonization of brain MR images. *Med. Image Anal.* 99, 103388. <https://doi.org/10.1016/j.media.2024.103388>
- Rogreau, A., Hives, F., Bordier, C., Lahousse, H., Roca, V., Lebouvier, T., Pasquier, F., Huglo, D., Semah, F., Lopes, R., 2024. A 3D convolutional neural network to classify subjects as Alzheimer's disease, frontotemporal dementia or healthy controls using brain 18F-FDG PET. *NeuroImage* 288, 120530. <https://doi.org/10.1016/j.neuroimage.2024.120530>
- Rohrer, J.D., Nicholas, J.M., Cash, D.M., Swieten, J. van, Doppler, E., Jiskoot, L., Minkelen, R. van, Rombouts, S.A., Cardoso, M.J., Clegg, S., Espak, M., Mead, S., Thomas, D.L., Vita, E.D., Masellis, M., Black, S.E., Freedman, M., Keren, R., MacIntosh, B.J., Rogava, E., Tang-Wai, D., Tartaglia, M.C., Laforce, R., Tagliavini, F., Tiraboschi, P., Redaelli, V., Prioni, S., Grisoli, M., Borroni, B., Padovani, A., Galimberti, D., Scarpini, E., Arighi, A., Fumagalli, G., Rowe, J.B., Coyle-Gilchrist, I., Graff, C., Fallström, M., Jelic, V., Ståhlbom, A.K., Andersson, C., Thonberg, H., Lilius, L., Frisoni, G.B., Binetti, G., Pievani, M., Bocchetta, M., Benussi, L., Ghidoni, R., Finger, E., Sorbi, S., Nacmias, B., Lombardi, G., Polito, C., Warren, J.D., Ourselin, S., Fox, N.C., Rossor, M.N., 2015. Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the Genetic Frontotemporal dementia Initiative (GENFI) study: a cross-sectional analysis. *Lancet Neurol.* 14, 253–262. [https://doi.org/10.1016/S1474-4422\(14\)70324-2](https://doi.org/10.1016/S1474-4422(14)70324-2)
- Rossor, M.N., Fox, N.C., Mummary, C.J., Schott, J.M., Warren, J.D., 2010. The diagnosis of young-onset dementia. *Lancet Neurol.* 9, 793–806. [https://doi.org/10.1016/S1474-4422\(10\)70159-9](https://doi.org/10.1016/S1474-4422(10)70159-9)
- Ruder, S., 2017. An Overview of Multi-Task Learning in Deep Neural Networks. <https://doi.org/10.48550/arXiv.1706.05098>
- Rutherford, S., Kia, S.M., Wolfers, T., Fraza, C., Zabihi, M., Dinga, R., Berthet, P., Worker, A., Verdi, S., Ruhe, H.G., Beckmann, C.F., Marquand, A.F., 2022. The normative modeling framework for computational psychiatry. *Nat. Protoc.* 17, 1711–1734. <https://doi.org/10.1038/s41596-022-00696-5>
- Savarese, P.H.P., Figueiredo, D.R., 2019. Residual Gates: A Simple Mechanism for Improved Network Optimization.
- Scheltens, P., Pasquier, F., Weerts, J.G., Barkhof, F., Leys, D., 1997. Qualitative assessment of cerebral atrophy on MRI: inter- and intra-observer reproducibility in dementia and normal aging. *Eur. Neurol.* 37, 95–99. <https://doi.org/10.1159/000117417>
- Schölkopf, B., Smola, A.J., 2001. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. The MIT Press. <https://doi.org/10.7551/mitpress/4175.001.0001>

- Seath, P., Macedo-Orrego, L.E., Velayudhan, L., 2024. Clinical characteristics of early-onset versus late-onset Alzheimer's disease: a systematic review and meta-analysis. *Int. Psychogeriatr.* 36, 1093–1109. <https://doi.org/10.1017/S1041610223000509>
- Seitz-Holland, J., Haas, S.S., Penzel, N., Reichenberg, A., Pasternak, O., 2024. BrainAGE, brain health, and mental disorders: A systematic review. *Neurosci. Biobehav. Rev.* 159, 105581. <https://doi.org/10.1016/j.neubiorev.2024.105581>
- Shorten, C., Khoshgoftaar, T.M., 2019. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* 6, 60. <https://doi.org/10.1186/s40537-019-0197-0>
- Shurrob, S., Guerra-Manzanares, A., Magid, A., Piechowski-Jozwiak, B., Atashzar, S.F., Shamout, F.E., 2024. Multimodal Machine Learning for Stroke Prognosis and Diagnosis: A Systematic Review. *IEEE J. Biomed. Health Inform.* 28, 6958–6973. <https://doi.org/10.1109/JBHI.2024.3448238>
- Singh, S., Srivastava, A., Mi, L., Caselli, R.J., Chen, K., Goradia, D., Reiman, E.M., Wang, Y., 2017. Deep Learning based Classification of FDG-PET Data for Alzheimers Disease Categories. *Proc. SPIE-- Int. Soc. Opt. Eng.* 10572, 105720J. <https://doi.org/10.1117/12.2294537>
- Sinha, S., Dieng, A.B., 2022. Consistency Regularization for Variational Auto-Encoders. <https://doi.org/10.48550/arXiv.2105.14859>
- Smith, S.M., Vidaurre, D., Alfaro-Almagro, F., Nichols, T.E., Miller, K.L., 2019. Estimation of brain age delta from brain imaging. *NeuroImage* 200, 528–539. <https://doi.org/10.1016/j.neuroimage.2019.06.017>
- Sohn, K., Lee, H., Yan, X., 2015. Learning Structured Output Representation using Deep Conditional Generative Models, in: *Advances in Neural Information Processing Systems*. Curran Associates, Inc.
- Standley, T., Zamir, A., Chen, D., Guibas, L., Malik, J., Savarese, S., 2020. Which tasks should be learned together in multi-task learning?, in: *Proceedings of the 37th International Conference on Machine Learning, ICML'20*. JMLR.org, pp. 9120–9132.
- Stern, Y., Albert, M., Barnes, C.A., Cabeza, R., Pascual-Leone, A., Rapp, P.R., 2023. A framework for concepts of reserve and resilience in aging. *Neurobiol. Aging* 124, 100–103. <https://doi.org/10.1016/j.neurobiolaging.2022.10.015>
- Stern, Y., Barnes, C.A., Grady, C., Jones, R.N., Raz, N., 2019. Brain Reserve, Cognitive Reserve, Compensation, and Maintenance: Operationalization, Validity, and Mechanisms of Cognitive Resilience. *Neurobiol. Aging* 83, 124–129. <https://doi.org/10.1016/j.neurobiolaging.2019.03.022>
- Sumowski, J.F., Rocca, M.A., Leavitt, V.M., Riccitelli, G., Comi, G., DeLuca, J., Filippi, M., 2013. Brain reserve and cognitive reserve in multiple sclerosis. *Neurology* 80, 2186–2193. <https://doi.org/10.1212/WNL.0b013e318296e98b>
- Sutskever, I., Martens, J., Dahl, G., Hinton, G., 2013. On the importance of initialization and momentum in deep learning, in: *Proceedings of the 30th International Conference on Machine Learning*. Presented at the International Conference on Machine Learning, PMLR, pp. 1139–1147.
- Tabbal, J., Ebadi, A., Mheich, A., Kabbara, A., Güntekin, B., Yener, G., Paban, V., Gschwandtner, U., Fuhr, P., Verin, M., Babiloni, C., Allouch, S., Hassan, M., 2025. Characterizing the heterogeneity of neurodegenerative diseases through EEG normative modeling. *Npj Park. Dis.* 11, 117. <https://doi.org/10.1038/s41531-025-00957-6>
- Tellechea, P., Pujol, N., Esteve-Belloch, P., Echeveste, B., García-Eulate, M.R., Arbizu, J., Riverol, M., 2018. Early- and late-onset Alzheimer disease: Are they the same entity? *Neurologia* 33, 244–253. <https://doi.org/10.1016/j.nrl.2015.08.002>
- Teunissen, C.E., Verberk, I.M.W., Thijssen, E.H., Vermunt, L., Hansson, O., Zetterberg, H., van der Flier, W.M., Mielke, M.M., del Campo, M., 2022. Blood-based biomarkers for Alzheimer's disease: towards clinical implementation. *Lancet Neurol.* 21, 66–77. [https://doi.org/10.1016/S1474-4422\(21\)00361-6](https://doi.org/10.1016/S1474-4422(21)00361-6)

- Therriault, J., Pascoal, T.A., Savard, M., Benedet, A.L., Chamoun, M., Tissot, C., Lussier, F., Kang, M.S., Thomas, E., Terada, T., Rej, S., Massarweh, G., Nasreddine, Z., Vitali, P., Soucy, J.-P., Saha-Chaudhuri, P., Gauthier, S., Rosa-Neto, P., 2021. Topographic Distribution of Amyloid- $\beta$ , Tau, and Atrophy in Patients With Behavioral/Dysexecutive Alzheimer Disease. *Neurology* 96, e81–e92. <https://doi.org/10.1212/WNL.00000000000011081>
- Tomczak, J.M., Welling, M., 2018. VAE with a VampPrior. <https://doi.org/10.48550/arXiv.1705.07120>
- Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>
- Vaezi Joze, H.R., Shaban, A., Iuzzolino, M.L., Koishida, K., 2020. MMTM: Multimodal Transfer Module for CNN Fusion, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Seattle, WA, USA, pp. 13286–13296. <https://doi.org/10.1109/CVPR42600.2020.01330>
- Vandenbergh, R., Adamczuk, K., Dupont, P., Laere, K.V., Chételat, G., 2013. Amyloid PET in clinical practice: Its place in the multidimensional space of Alzheimer's disease. *NeuroImage Clin.* 2, 497–511. <https://doi.org/10.1016/j.nicl.2013.03.014>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2023. Attention Is All You Need. <https://doi.org/10.48550/arXiv.1706.03762>
- Verdi, S., Rutherford, S., Fraza, C., Tosun, D., Altmann, A., Raket, L.L., Schott, J.M., Marquand, A.F., Cole, J.H., Alzheimer's Disease Neuroimaging Initiative, 2024. Personalizing progressive changes to brain structure in Alzheimer's disease using normative modeling. *Alzheimers Dement. J. Alzheimers Assoc.* 20, 6998–7012. <https://doi.org/10.1002/alz.14174>
- Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.-A., 2008. Extracting and composing robust features with denoising autoencoders, in: Proceedings of the 25th International Conference on Machine Learning - ICML '08. Presented at the the 25th international conference, ACM Press, Helsinki, Finland, pp. 1096–1103. <https://doi.org/10.1145/1390156.1390294>
- Wei, D., Zhuang, K., Ai, L., Chen, Q., Yang, W., Liu, W., Wang, K., Sun, J., Qiu, J., 2018. Structural and functional brain scans from the cross-sectional Southwest University adult lifespan dataset. *Sci. Data* 5, 180134. <https://doi.org/10.1038/sdata.2018.134>
- Wei, X., Zhang, T., Li, Y., Zhang, Y., Wu, F., 2020. Multi-Modality Cross Attention Network for Image and Sentence Matching, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 10938–10947. <https://doi.org/10.1109/CVPR42600.2020.01095>
- Weisscher, N., Vermeulen, M., Roos, Y.B., de Haan, R.J., 2008. What should be defined as good outcome in stroke trials; a modified Rankin score of 0-1 or 0-2? *J. Neurol.* 255, 867–874. <https://doi.org/10.1007/s00415-008-0796-8>
- Whitwell, J.L., 2019. Neuroimaging across the FTD spectrum. *Prog. Mol. Biol. Transl. Sci.* 165, 187–223. <https://doi.org/10.1016/bs.pmbts.2019.05.009>
- Wu, R., Wang, H., Chen, H.-T., Carneiro, G., 2024. Deep Multimodal Learning with Missing Modality: A Survey. <https://doi.org/10.48550/arXiv.2409.07825>
- Wu, Y., He, K., 2018. Group Normalization. Presented at the Proceedings of the European Conference on Computer Vision (ECCV), pp. 3–19.
- Xia, C., Makaretz, S.J., Caso, C., McGinnis, S., Gomperts, S.N., Sepulcre, J., Gomez-Isla, T., Hyman, B.T., Schultz, A., Vasdev, N., Johnson, K.A., Dickerson, B.C., 2017. Association of In Vivo [18F]AV-1451 Tau PET Imaging Results With Cortical Atrophy and Symptoms in Typical and Atypical Alzheimer Disease. *JAMA Neurol.* 74, 427–436. <https://doi.org/10.1001/jamaneurol.2016.5755>
- Yang, Z., Nasrallah, I.M., Shou, H., Wen, J., Doshi, J., Habes, M., Erus, G., Abdulkadir, A., Resnick, S.M., Albert, M.S., Maruff, P., Fripp, J., Morris, J.C., Wolk, D.A., Davatzikos, C., 2021. A deep

- learning framework identifies dimensional representations of Alzheimer's Disease from brain structure. *Nat. Commun.* 12, 7065. <https://doi.org/10.1038/s41467-021-26703-z>
- Yang, Z., Wen, J., Davatzikos, C., 2022. Surreal-GAN:Semi-Supervised Representation Learning via GAN for uncovering heterogeneous disease-related imaging patterns.  
<https://doi.org/10.48550/arXiv.2205.04523>
- Yang, Z., Wen, J., Davatzikos, C., 2020. Smile-GANs: Semi-supervised clustering via GANs for dissecting brain disease heterogeneity from medical images.  
<https://doi.org/10.48550/arXiv.2006.15255>
- Yeh, C.-H., Hong, C.-Y., Hsu, Y.-C., Liu, T.-L., Chen, Y., LeCun, Y., 2022. Decoupled Contrastive Learning. <https://doi.org/10.48550/arXiv.2110.06848>
- Young, P.N.E., Estarellas, M., Coomans, E., Srikrishna, M., Beaumont, H., Maass, A., Venkataraman, A.V., Lissaman, R., Jiménez, D., Betts, M.J., McGlinchey, E., Berron, D., O'Connor, A., Fox, N.C., Pereira, J.B., Jagust, W., Carter, S.F., Paterson, R.W., Schöll, M., 2020. Imaging biomarkers in neurodegeneration: current and future practices. *Alzheimers Res. Ther.* 12, 49. <https://doi.org/10.1186/s13195-020-00612-7>
- Yu, T., Kumar, S., Gupta, A., Levine, S., Hausman, K., Finn, C., 2020. Gradient Surgery for Multi-Task Learning. <https://doi.org/10.48550/arXiv.2001.06782>
- Zhao, Q., Adeli, E., Honnorat, N., Leng, T., Pohl, K.M., 2019a. Variational AutoEncoder For Regression: Application to Brain Aging Analysis. *Med. Image Comput. Comput.-Assist. Interv. MICCAI Int. Conf. Med. Image Comput. Comput.-Assist. Interv.* 11765, 823–831.  
[https://doi.org/10.1007/978-3-030-32245-8\\_91](https://doi.org/10.1007/978-3-030-32245-8_91)
- Zhao, Q., Adeli, E., Honnorat, N., Leng, T., Pohl, K.M., 2019b. Variational AutoEncoder For Regression: Application to Brain Aging Analysis. *Med. Image Comput. Comput.-Assist. Interv. MICCAI Int. Conf. Med. Image Comput. Comput.-Assist. Interv.* 11765, 823–831.  
[https://doi.org/10.1007/978-3-030-32245-8\\_91](https://doi.org/10.1007/978-3-030-32245-8_91)
- Zhou, L., Butler, T.A., Wang, X.H., Xi, K., Tanzi, E.B., Glodzik, L., Chiang, G.C., de Leon, M.J., Li, Y., 2024. Multimodal assessment of brain fluid clearance is associated with amyloid-beta deposition in humans. *J. Neuroradiol.* 51, 101164.  
<https://doi.org/10.1016/j.neurad.2023.10.009>
- Zhu, D., Montagne, A., Zhao, Z., 2021. Alzheimer's pathogenic mechanisms and underlying sex difference. *Cell. Mol. Life Sci.* 78, 4907–4920. <https://doi.org/10.1007/s00018-021-03830-w>
- Zihni, E., Madai, V., Khalil, A., Galinovic, I., Fiebach, J., Kelleher, J., Frey, D., Livne, M., 2020. Multimodal Fusion Strategies for Outcome Prediction in Stroke:, in: Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies. Presented at the 13th International Conference on Health Informatics, SCITEPRESS - Science and Technology Publications, Valletta, Malta, pp. 421–428.  
<https://doi.org/10.5220/0008957304210428>
- Zuo, L., Liu, Y., Xue, Y., Dewey, B.E., Remedios, S.W., Hays, S.P., Bilgel, M., Mowry, E.M., Newsome, S.D., Calabresi, P.A., Resnick, S.M., Prince, J.L., Carass, A., 2023. HACA3: A unified approach for multi-site MR image harmonization. *Comput. Med. Imaging Graph.* 109, 102285.  
<https://doi.org/10.1016/j.compmedimag.2023.102285>

