

UNIVERSITÉ DE LILLE  
**FACULTÉ DE MÉDECINE HENRI WAREMBOURG**  
Année : 2023

THÈSE POUR LE DIPLÔME D'ÉTAT  
DE DOCTEUR EN MÉDECINE

**Elaboration d'un réseau de neurones convolutifs 3D pour classer  
les TEP-FDG entre maladie d'Alzheimer, démence frontotemporale  
et cognitivement normal**

Présentée et soutenue publiquement le 2 mars 2023  
au Pôle Recherche  
par **Antoine ROGEAU**

---

**JURY**

**Président :**

**Monsieur le Professeur Damien HUGLO**

**Assesseurs :**

**Monsieur le Professeur Franck SEMAH**

**Monsieur le Professeur Thibaud LEBOUVIER**

**Monsieur le Docteur Renaud LOPES**

**Directeur de thèse :**

**Monsieur le Docteur Florent HIVES**

---

## **AVERTISSEMENT**

**La Faculté n'entend donner aucune approbation aux opinions émises dans les thèses : celles-ci sont propres à leurs auteurs.**



# TABLE DES MATIERES

ABREVIATIONS.....	5
<b>I. INTRODUCTION .....</b>	<b>6</b>
1) GENERALITES .....	6
2) MALADIE D'ALZHEIMER .....	7
3) SPECTRE DES DEMENCES FRONTOTEMPORALES.....	10
4) LA TEP-FDG ET SON INTERET EN NEURODEGENERESCENCE .....	12
5) APPRENTISSAGE PROFOND ET RESEAUX DE NEURONES CONVOLUTIFS .....	14
Objectifs de cette étude .....	19
<b>II. ARTICLE .....</b>	<b>20</b>
ABSTRACT .....	20
INTRODUCTION .....	21
METHODS .....	23
Subjects.....	23
Data acquisition .....	24
Image preprocessing.....	25
Datasets and data augmentation.....	26
Neural network building .....	26
Clinical Interpretation .....	27
Network attention .....	27
Network extrapolability .....	28
RESULTS.....	30
Demographics .....	30
Model training .....	31
Model evaluation .....	33
Clinical interpretation .....	35
Network attention .....	35
Network extrapolability .....	35
DISCUSSION .....	37
Strengths.....	39
Limitations .....	40
CONCLUSION .....	41
ACKNOWLEDGEMENTS .....	42
<b>III. DISCUSSION .....</b>	<b>44</b>
1) SUGGESTIONS POUR LE DÉVELOPPEMENT D'UN FUTUR LOGICIEL D'IA APPLICABLE.....	44
2) FUTURE PLACE DE LA MÉDECINE NUCLÉAIRE EN NEUROLOGIE COGNITIVE.....	45
3) VERS UNE FIN DE L'INTERPRÉTATION VISUELLE ? .....	48
<b>IV. CONCLUSION.....</b>	<b>50</b>
BIBLIOGRAPHY.....	51
SUPPLEMENTARY INFORMATION.....	58

## ABREVIATIONS

### Françaises

<b>AMM</b>	autorisation de mise sur le marché
<b>APP</b>	aphasie primaire progressive
<b>CNIL</b>	Commission Nationale de l'Informatique et des Libertés
<b>CS</b>	cognitivement sain
<b>DCL</b>	démence à corps de Lewy
<b>DFT</b>	démence frontotemporale
<b>DLFT</b>	dégénérescence lobaire frontotemporale
<b>FDG</b>	fluorodéoxyglucose
<b>IRM</b>	imagerie par résonance magnétique
<b>LCS</b>	liquide cérebrospinal
<b>MA</b>	maladie d'Alzheimer
<b>PA</b>	potentiel d'action
<b>TEP</b>	tomographie par émission de positons

### English

<b>ACC</b>	anterior cingulate cortex
<b>AD</b>	Alzheimer's disease
<b>ADNI</b>	Alzheimer's Disease Neuroimaging Initiative
<b>AI</b>	artificial intelligence
<b>CN</b>	cognitively normal
<b>CNN</b>	convolutional neural network
<b>CT</b>	computed tomography
<b>DALY</b>	disability-adjusted life year
<b>DLB</b>	dementia with Lewy bodies
<b>EANM</b>	European Association of Nuclear Medicine
<b>FDA</b>	Food and Drug Administration
<b>FTD</b>	frontotemporal dementia
<b>FTLDNI</b>	Frontotemporal Lobar Degeneration Neuroimaging Initiative
<b>GE</b>	General Electric
<b>ICBM</b>	International Consortium of Brain Mapping
<b>IWG</b>	International Working Group
<b>LIIFE</b>	Lille In vivo Imaging and Functional Exploration
<b>LUT</b>	lookup table
<b>MCI</b>	mild cognitive impairment
<b>MMSE</b>	mini-mental state examination
<b>MRI</b>	magnetic resonance imaging
<b>NIA-AA</b>	National Institute on Aging and Alzheimer's Association
<b>PART</b>	primary age-related tauopathy
<b>PCC</b>	posterior cingulate cortex
<b>PET</b>	positron emission tomography
<b>PPA</b>	primary progressive aphasia
<b>ReLU</b>	rectified linear unit
<b>ROC</b>	receiver operating characteristic
<b>SD</b>	standard deviation
<b>SPM12</b>	Statistical Parametric Mapping 12
<b>SV2A</b>	synaptic vesicle glycoprotein 2A
<b>TDP-43</b>	TAR DNA-binding protein 43
<b>ULille</b>	University of Lille

# I. INTRODUCTION

---

## 1) Généralités

Le syndrome démentiel est défini par une dégradation des fonctions cognitives à l'origine d'une perte d'autonomie, en l'absence d'un syndrome confusionnel ou d'un autre trouble mental pouvant expliquer à lui seul le tableau clinique<sup>1</sup>.

Les démences augmentent en incidence en lien avec le vieillissement de la population et leur prévalence mondiale est prévue d'être multipliée par 3 d'ici 2050<sup>2</sup>. Du fait de leur retentissement sévère sur la qualité de vie du patient et de leur impact financier, avec des dépenses en moyenne 184 500 dollars plus élevées comparées à un individu sans démence<sup>3</sup>, la société devra apprendre à vivre avec cette véritable épidémie de troubles cognitifs.

D'un point de vue médical, il est judicieux de rappeler la différence entre diagnostic syndromique et maladie. Afin de pouvoir établir une classification nosologique juste, on définit la 'maladie' telle que la résultante de processus biochimiques ou biologiques à la différence du 'syndrome' défini à partir des symptômes ou d'une certaine présentation radiologique. Il existe une confusion pour certaines entités lorsque les classifications clinique, radiologique et anatomo-pathologique s'entre-courent. Par exemple, il est reporté au mieux 56% de corrélation dans les séries de démence corticobasale entre clinique et anatomopathologie<sup>4</sup>.

Ainsi, il est important de définir la maladie neurodégénérative sur les observations anatomopathologiques, ceci d'autant plus avec l'arrivée de thérapies ciblées anti-amyloïde telles que l'aducanumab déjà approuvé par la *Food and Drug Administration*

(FDA)<sup>5</sup>, et le lecanemab dont les récents résultats de phase III démontrent un déclin cognitif moindre sur 18 mois comparé à un placebo chez les sujets présentant une maladie d'Alzheimer (MA) non évoluée<sup>6, 7</sup>. Des thérapies anti-tau sont également en phase d'essai<sup>8, 9</sup>.

## 2) Maladie d'Alzheimer

La MA est la cause la plus fréquente de démence et touche 5,4 millions de personnes aux Etats-Unis seuls<sup>10</sup>. Elle conduit typiquement à un tableau clinique où les troubles mnésiques sont au premier plan. Cependant, il existe des formes aphasiques, visuelles que sont l'aphasie primaire logopénique, l'atrophie corticale postérieure qui sont, dans la majorité des cas, secondaires à des MA, ou comportementales désignées comme le variant frontal de la MA<sup>11</sup>. Quatre-vingt-dix pourcents des cas sont à début tardif, et bien que l'origine soit considérée multifactorielle, l'hérédité est estimée entre 58-79%<sup>12</sup>.

Sa définition a évolué à plusieurs reprises, et maintenant les critères de la *National Institute on Aging and Alzheimer's Association* (NIA-AA) la définit comme l'association d'une tauopathie et d'une amyloïdopathie avec la notion de MA préclinique (encore débattue) où la cognition reste préservée, et de syndrome clinique de MA quand les biomarqueurs sont absents en la présence d'un syndrome amnésique avec trouble de l'encodage tel que dans la forme typique de MA dite amnésique<sup>13</sup>.

Ainsi sur le plan anatomopathologique, ses deux caractéristiques sont l'hyperphosphorylation de la protéine tau et le dépôt de plaques amyloïdes. Il était initialement pensé que la mort neuronale était secondaire aux plaques amyloïdes,

avec un possible effet médiateur de la protéine tau, selon l'hypothèse de la cascade amyloïde<sup>14</sup>. Les biomarqueurs amyloïdes sont détectables parfois des décennies avant l'apparition des premiers symptômes et précèdent les biomarqueurs tau (figure 1). Aussi, l'accumulation de peptides amyloïdes perturbe l'équilibre glutamate/GABA en faveur du premier, conduisant à une hyperexcitabilité et la formation d'un cercle vicieux avec augmentation de la production de peptides A $\beta$  secondaire à l'hyperexcitabilité<sup>8</sup>. Cependant, l'hypothèse d'un rôle prédominant de la protéine amyloïde fut remis en cause par la présence de plaques amyloïdes chez de nombreux sujets âgés cognitivement sains, et par les résultats controversés de la phase III de l'aducanumab<sup>15, 16</sup>.

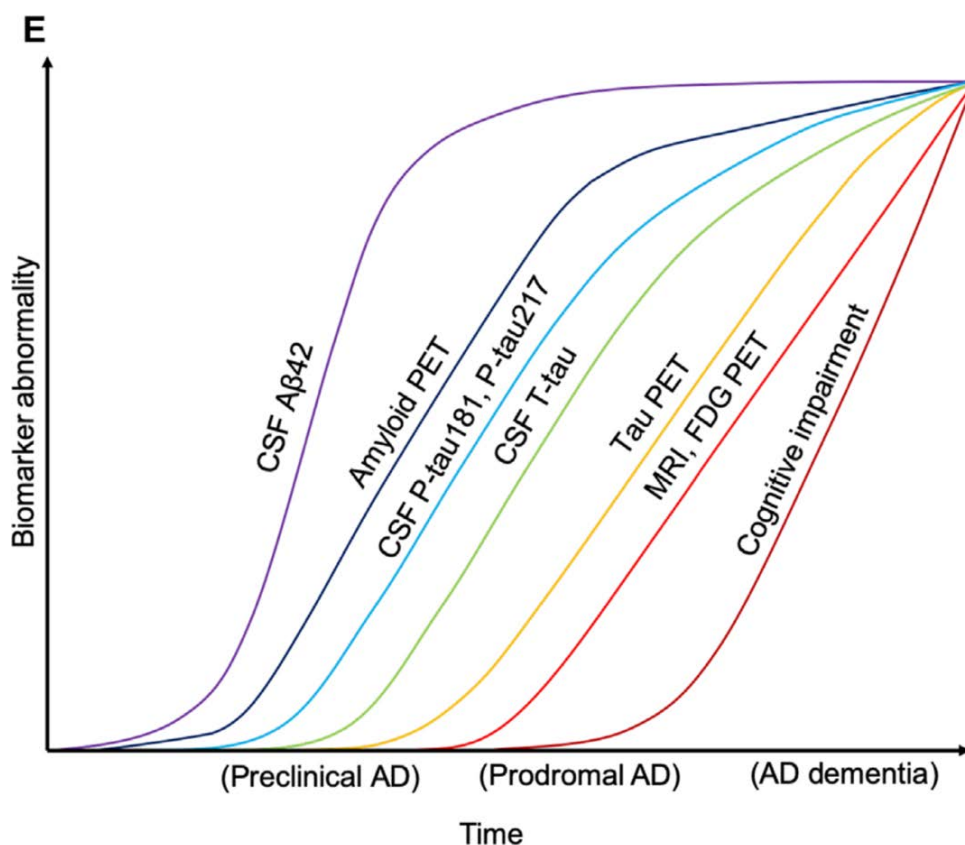


Figure 1 : Pathologies tau et amyloïde apparaissent des années avant le déclin cognitif et pourraient permettre un diagnostic précoce. Extrait de Mattsson-Carlgren *et al* (2020)<sup>17</sup>.



L'état de la science s'est ensuite tourné vers une toxicité primordiale des neurofibrilles de protéine tau, notamment suggérée par une meilleure corrélation entre symptomatologie, densité synaptique et pathologie tau (figure 2) et la présence d'amas de protéine tau dans d'autres maladies neurodégénératives<sup>18, 19</sup>. La MA est une tauopathie 3R/4R, avec des isoformes de protéine tau ayant 3 ou 4 domaines répétés pour la fixation des microtubules. Les agrégats apparaissent en premier dans le cortex entorhinal et l'hippocampe, puis dans le cortex associatif. Une hypothèse avancée pour la propagation de la pathologie tau serait celle d'une propagation trans-synaptique *prion-like*<sup>20</sup>. Néanmoins, l'identification récente de *primary age-related tauopathies (PART)*, aux neurofibrilles tau également 3R/4R en l'absence de pathologie amyloïde et associées à des symptômes moins sévères, a conduit à l'hypothèse d'une synergie entre toxicité tau et pathologie amyloïde<sup>21</sup>.

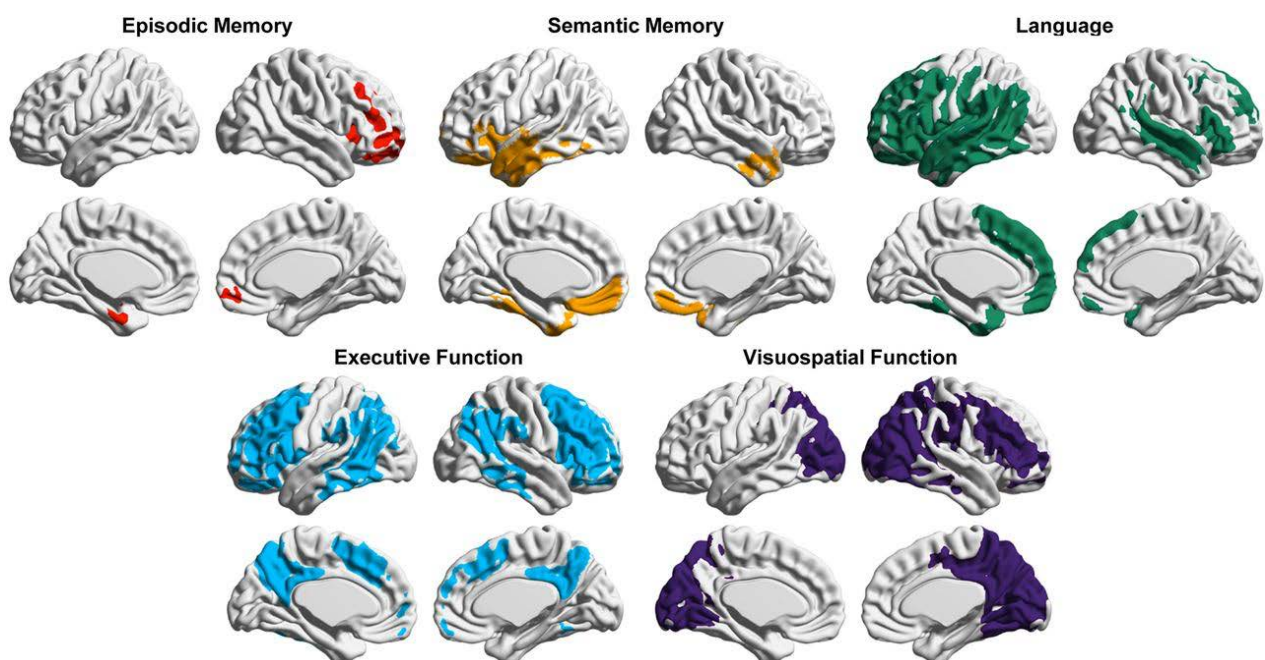


Figure 2 : Corrélation entre accumulation du [18F]-flortaucipir et l'atteinte de fonctions cognitives de certains domaines. Voxels significatifs à *peak-level* non-corrigé  $p < 0,001$  et *cluster-level Family-wise error* corrigé  $< 0,05$ . De Tanner et Rabinovici<sup>22</sup>.

Aujourd'hui, l'hypothèse avancée est celle d'interactions entre pathologie tau et pathologie amyloïde avec, bien qu'une toxicité essentiellement d'origine tau, un effet catalyseur des plaques amyloïdes sur la propagation de la pathologie tau<sup>23</sup>. Il est aussi pensé que le système immunitaire joue un rôle avec possiblement un effet médiateur de la microglie activée<sup>24</sup>. L'hypothèse *dual pathway* émettait l'idée que les deux pathologies pouvaient être initiées par des mécanismes communs sans interaction entre elles<sup>25</sup>. Celle-ci pourrait être adaptée en remplaçant la nature indépendante par une nature synergique entre tau et amyloïde et en reprenant l'idée d'une origine commune. A titre d'exemple, l'insulinorésistance aggrave l'accumulation de tau et la formation de plaques amyloïde<sup>26</sup>.

### 3) Spectre des démences frontotemporales

La première description de démence frontotemporale (DFT) fut rapportée par Arnold Pick en 1892, mais c'est Alois Alzheimer qui observa des caractéristiques anatomopathologiques telles que des corps d'inclusion argentophiles et donna le nom de "maladie de Pick". Après une période où toutes les démences étaient considérées d'origine vasculaire, les équipes de Lund et de Manchester, dans les années 1980, posèrent les fondements de la classification sémiologique toujours utilisée ce jour, décrivant trois entités cliniques : la démence frontotemporale, l'aphasie primaire et la démence sémantique. Depuis, cette classification a été revue, et nous parlons aujourd'hui de DFT de variant comportemental, d'aphasie primaire progressive (APP) aphasique ou agrammaticale et d'APP sémantique en suivant les critères de Rascovsky et de Gorno-Tempini<sup>27, 28</sup>.

Sur un plan anatomopathologique, la dégénérescence lobaire frontotemporale (DLFT) montre également l'accumulation anormale d'une protéine, notamment tau dans 40% des cas, mais plus communément *TAR DNA-binding protein 43* (TDP-43) dans 50% des cas, de protéines de la famille FET dans 5-10% des cas, et les cas restants comprennent des situations où la protéine n'a pas encore été identifiée ou des cas sans inclusion de protéine comme les DLFT-ni (*no inclusion*)<sup>29, 30</sup>. Ainsi, avec l'approche thérapeutique actuelle d'élimination du surplus de protéine anormale, il est vital de différencier MA et DLFT. D'autant plus que même les DLFT-tau restent différentes de la MA, puisqu'elles ne présentent pas de pathologie amyloïde et leur isoforme tau sont de type 3R ou 4R, contrairement à la MA qui associe 3R et 4R. Il est à noter qu'aujourd'hui le terme de maladie de Pick est réservé à une DLFT-tau 3R.

De manière intéressante, il existe une corrélation avec les formes cliniques précédemment décrites et les caractéristiques de DLFT (figure 3). La DFT de variant comportemental peut être secondaire à n'importe quel type de DLFT, alors que le variant sémantique est secondaire essentiellement à une DLFT-TDP-43 et le variant agrammatical est secondaire de manière égale à une DLFT-TDP-43 ou tau<sup>31</sup>.

A l'instar de la MA, les critères établis ne comprennent pas de biomarqueurs biologiques pour le diagnostic de DFT, et la DLFT n'est souvent prouvée que lors de l'autopsie. Ainsi, le développement de l'imagerie métabolique afin d'identifier le processus responsable est essentiel. L'agrément aux collectivités vient d'être obtenu pour le flutémétamol, le flortaucipir devrait arriver sur le marché bientôt et la recherche continue pour identifier des ligands TDP-43 et FET<sup>32</sup>.

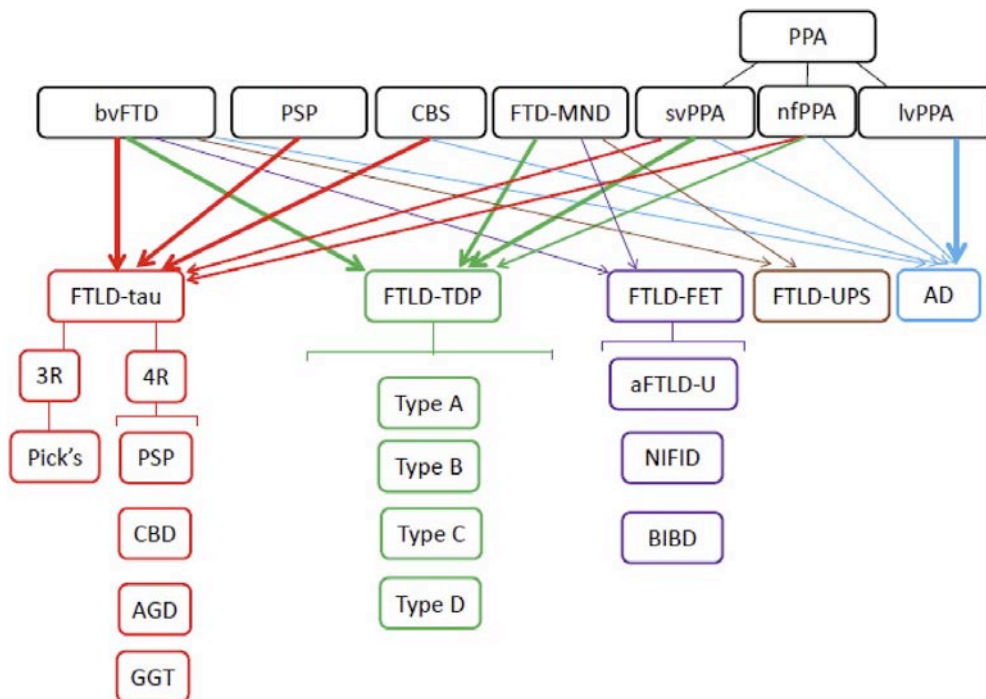


Figure 3 : Correspondance clinico-anatomopathologique du spectre des DFT. Il est à noter que la MA peut être responsable de formes cliniques comportementales, sémantiques ou agrammaticales de DFT. Extrait d'Olney *et al* (2017)<sup>31</sup>.

#### 4) La TEP-FDG et son intérêt en neurodégénérescence

La tomographie par émission de positons (TEP) au fluorodéoxyglucose (FDG) a été initialement développée pour l'étude du métabolisme cérébral. Elle est basée sur le couplage du glucose au fluor-18. Le fluor-18 émet un positon qui s'annihile avec un électron dans la matière et émet deux photons d'annihilation antiparallèles qui sont détectés par la caméra TEP. Le fluorodéoxyglucose entre, comme le glucose, dans le cycle de Krebs, mais s'accumule après l'étape de phosphorylation par l'hexokinase II. Cette accumulation intracellulaire permet l'étude du métabolisme neuronal et glial et aboutit à l'observation de *patterns* différents selon les pathologies neurodégénératives.

Les critères diagnostiques des DFT requièrent l'observation d'un *pattern* d'imagerie typique en imagerie par résonance magnétique (IRM) ou en TEP pour poser un diagnostic de 'DFT probable'. Pour la MA, puisqu'il existe davantage de biomarqueurs, l'imagerie n'est pas indispensable mais fait partie des critères diagnostiques mineurs de l'*International Working Group (IWG)* par Dubois *et al*, notamment si la clinique ou les biomarqueurs du liquide cébrospinal (LCS) sont peu aidants<sup>33</sup>. De manière intéressante, cette dernière révision des critères diagnostiques place la TEP-FDG avant l'IRM dans l'arsenal diagnostique. Il a d'ailleurs été montré que de nombreux sujets mal diagnostiqués malgré des explorations neurologiques en IRM bénéficiaient d'une amélioration dans la précision diagnostique avec la TEP-FDG (figure 4).

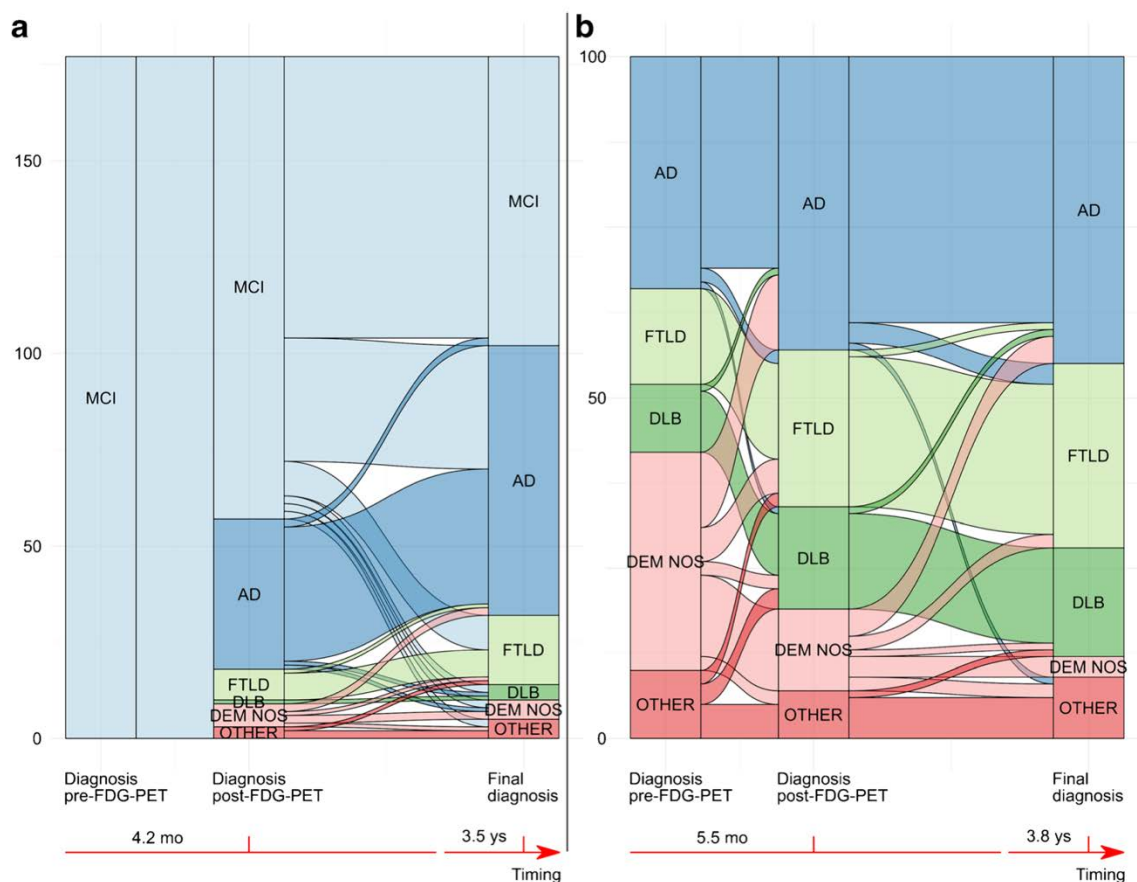


Figure 4 : Diagrammes fluviaux de reclassification diagnostique après TEP-FDG, (A) de sujets initialement diagnostiqués comme *mild cognitive impairment* (MCI), et (B) de sujets diagnostiqués avec une démence de nature incertaine. De Perini *et al* (2021).

Puisqu'il est raisonnable d'imaginer que l'hypométabolisme précède la dégénérescence, la TEP-FDG peut aider à un diagnostic plus précoce, comme le montre la sensibilité plus élevée comparativement à l'IRM, à 90% vs 70% dans une série de DFT comportementales<sup>34</sup>, bien que ceci soit limité par une spécificité plus faible due à une confusion avec les troubles neuropsychiatriques. Dans les cas où le doute diagnostique persiste entre maladie neurodégénérative et trouble neuropsychiatrique, une deuxième TEP-FDG à un an permet de réduire l'incertitude diagnostique de 80% à 34% avec un quart des cas changeant de diagnostic<sup>35</sup>.

#### 5) Apprentissage profond et réseaux de neurones convolutifs

L'intelligence artificielle (IA) est un vaste domaine qui inclut les méthodes de *Machine learning* ou apprentissage par machine. Celles-ci comprennent les méthodes de *Deep learning* ou d'apprentissage profond, représentées essentiellement par les réseaux de neurones.

Le réseau de neurones est inspiré par le modèle du perceptron imaginé en 1957 par Frank Rosenblatt<sup>36</sup>. L'équation 1 résume l'opération effectuée par le perceptron, les valeurs  $i$  (pour *input*) sont multipliées par les poids  $w$  puis additionnées. La fonction  $f$  est appliquée à cette valeur, et il en résulte la sortie  $o$ .

$$o = f\left(\sum_{k=1}^n i_k \cdot w_k\right) \quad (1)$$

De manière un peu analogue au neurone biologique et au potentiel d'action (PA), la fonction d'activation est appliquée à chaque neurone d'un réseau. Il existe différentes fonctions d'activation, mais la plus populaire de par sa simplicité et son efficacité est

le fonction Unité Linéaire Rectifiée (*ReLU*). Celle-ci retourne 0 lorsque l'entrée est négative, ou la valeur elle-même lorsque l'entrée est positive (figure 5).

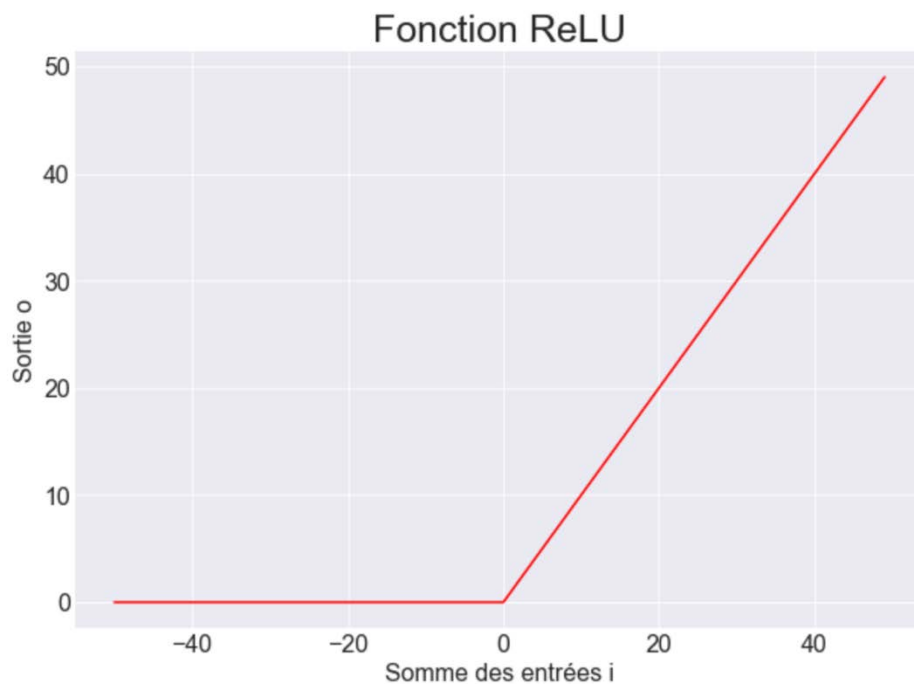


Figure 5 : Courbe de la fonction Unité Linéaire Rectifiée. Dessinée par l'auteur.

Les réseaux de neurones sont basés sur le modèle du perceptron, puisque chaque neurone contient l'opération décrite ci-dessus. Les réseaux de neurones les plus simples sont entièrement connectés et faits de couches dites denses, contenant des centaines ou des milliers de neurones où toutes les entrées  $i$  sont fournies à un neurone qui communique sa sortie  $o$  à tous les neurones de la couche suivante. Il s'en suit une prédiction et une comparaison de cette prédiction à la valeur réelle. La différence entre les deux, correspondant à la fonction coût, permet le calcul d'une dérivé, ou gradient, qui est ensuite rétropropagée sur l'ensemble du réseau afin de mettre à jour les poids. L'ensemble de cette procédure correspond à une époque (figure 6).

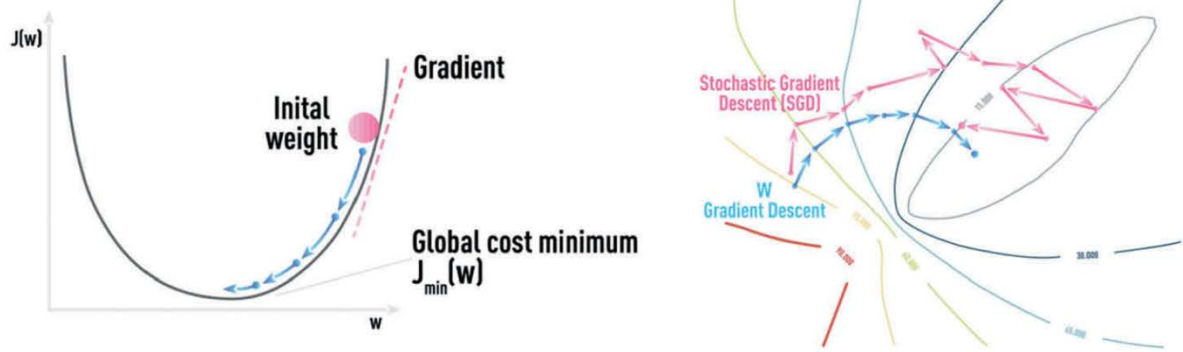


Figure 6 : Calcul du gradient à partir de la fonction coût et rétropropagation afin de diminuer la fonction coût, correspondant à la diminution de la différence entre prédictions et valeurs réelles, gauche en 1 dimension, et droite en 2 dimensions. Extrait de Gavrishchaka *et al* (2019)<sup>37</sup>.

En *computer vision* ou vision par ordinateur, il a été observé que l'opération de convolution était particulièrement utile afin d'extraire les caractéristiques de l'image. Le premier réseau de neurones convolutifs (CNN) fut développé en 1989 par Yann LeCun afin de reconnaître les chiffres<sup>38</sup>, et a été utilisé pendant de nombreuses années par les services postaux américains.

Les CNN multicouches sont inspirés par l'organisation stratifiée de la rétine et du cortex visuel. A titre d'exemple, le principe de sélectivité d'orientation basé sur le fait qu'un neurone déclenche un PA en fonction de l'orientation d'une ligne, est retrouvé dans les CNN<sup>39, 40</sup>, ceci grâce à l'opération de convolution qui permet d'extraire les caractéristiques des formes vues par le réseau. Ensuite, l'analogie peut être faite entre cellules simples, complexes et hypercomplexes du cortex qui déclenchent des PA si l'objet dans le champ visuel répond exactement à leurs critères (de complexité plus élevée) et la profondeur des couches de neurones dans un réseau 2D où la première



couche détecte les lignes, la seconde les figures, la troisième les visages ou les corps etc (figure 7)<sup>41</sup>.

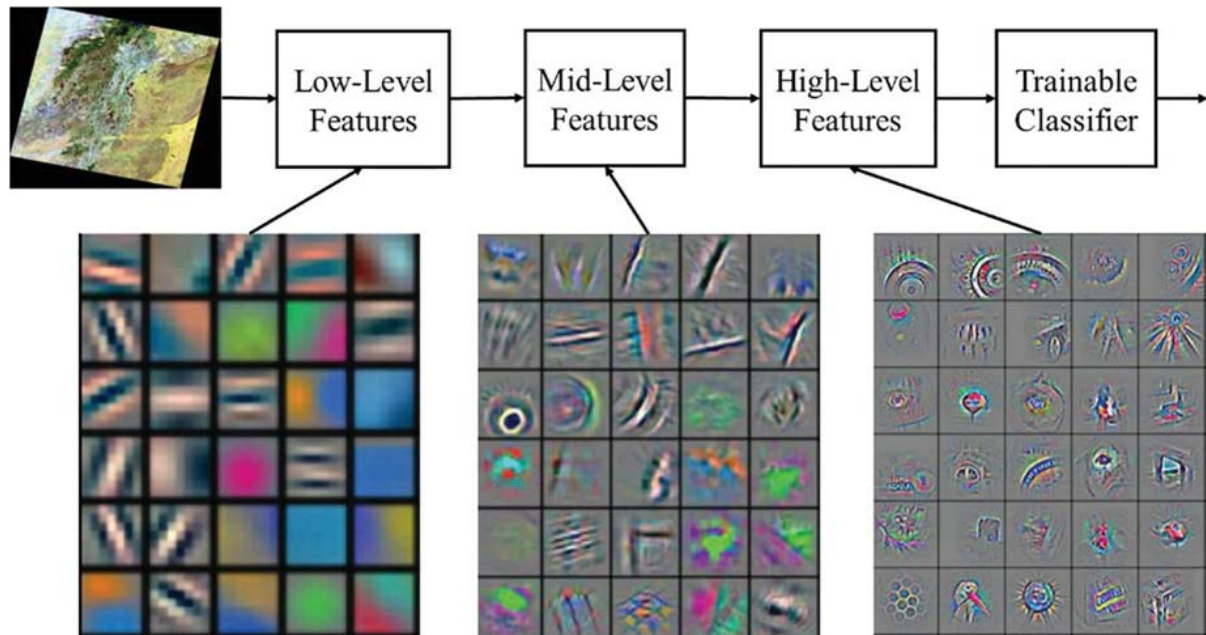


Figure 7 : Les couches superficielles d'un réseau reconnaissent les formes simples, et la complexité des formes reconnues augmente avec la profondeur des couches. De Traoré *et al* (2018)<sup>42</sup>.

Parmi les CNN, il existe certaines architectures plus populaires que d'autres, notamment ResNet, AlexNet ou Inception. Simonyan et Zisserman ont proposé l'architecture VGG16<sup>43</sup>, qui s'est classée seconde à la *ImageNet Large Scale Visual Recognition Competition* en 2014 totalisant une précision de 92,7% sur 100 000 images classées en 1000 catégories<sup>44</sup>. L'architecture VGG a fait ses preuves à multiples reprises dans la classification d'imagerie médicale, et nous reprenons dans le présent travail une architecture inspirée par le VGG16 qui a déjà fait ses preuves dans la classification de TEP-FDG cérébrales (figure 8)<sup>45</sup>.

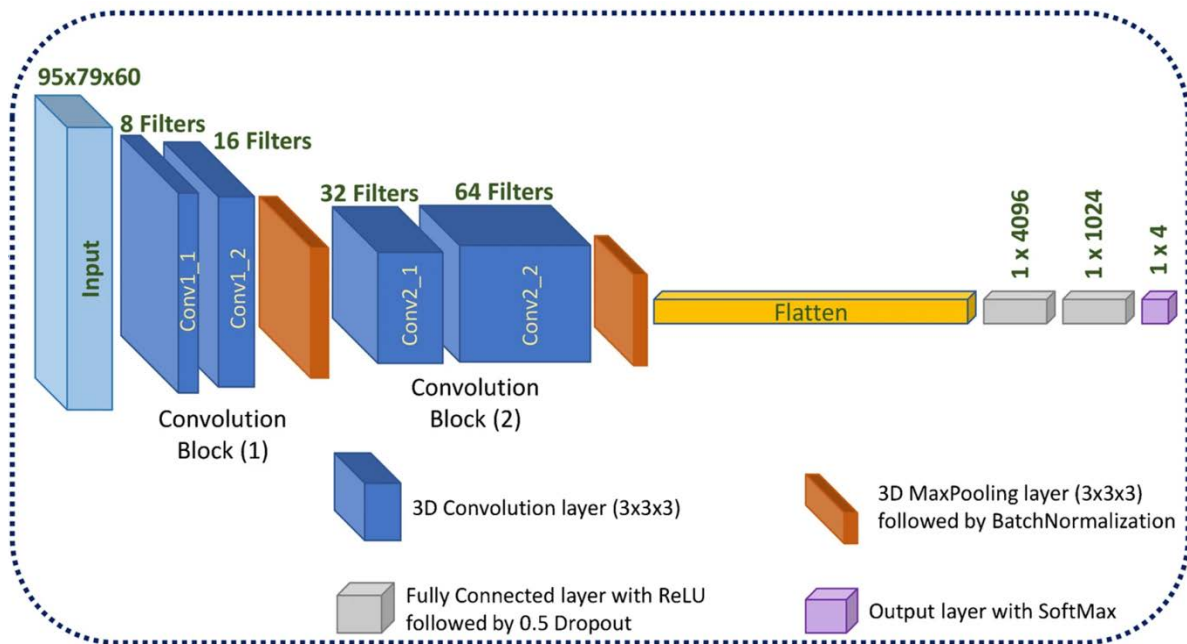


Figure 8 : Réseau convolutif adapté du réseau VGG16 développé par le *Visual Geometry Group* de l'Université d'Oxford, et personnalisé à la classification de TEP-FDG cérébrales en 4 catégories. Extrait d'Etminani *et al* (2022)<sup>45</sup>.

## **Objectifs de cette étude**

L'IA a récemment démontré d'excellentes performances dans la classification de TEP-FDG cérébrales de patients ayant une MA, une démence à corps de Lewy (DCL), un MCI ou de sujets sains<sup>45</sup>. Cependant, les performances pour discriminer MA vs DFT n'ont jamais été évaluées en imagerie métabolique. Ce besoin est renforcé par l'existence de variants frontaux de la MA pouvant être confondus avec des DFT et l'arrivée de thérapies ciblées pour lesquelles un diagnostic certain sera nécessaire.

Nous avons recueilli les TEP-FDG de patients ayant reçu un diagnostic probable de MA ou de DFT, toutes formes cliniques confondues incluant les APP et les formes génétiques, dans le service du centre mémoire du CHU de Lille. Afin de pouvoir ajouter une troisième catégorie diagnostique de sujets cognitivement sains (CS), et de favoriser l'extrapolation du réseau de neurones, nous avons recueilli les TEP-FDG de patients et de sujets sains de deux bases de données américaines de sujets avec une MA, une DFT ou CS.

L'analyse principale porte sur la comparaison des performances de ce réseau de neurones vs l'interprétation clinique des médecins nucléaires sur un jeu de 59 TEP-FDG gardées à part et n'ayant pas servi à l'entraînement du réseau. Les analyses secondaires comprennent une épreuve d'occlusion afin de déterminer les régions cérébrales guidant la classification du réseau de neurones et le test sur deux autres catégories de sujets présentant une DCL ou un trouble anxieux afin d'évaluer son extrapolation.

Les résultats et une partie de la discussion de ce travail sont présentés sous forme d'article.

## II. ARTICLE

---

### A 3D NEURAL NETWORK TO CLASSIFY SUBJECTS AS ALZHEIMER'S DISEASE, FRONTOTEMPORAL DEMENTIA OR HEALTHY CONTROLS USING BRAIN 18F-FDG PET

#### ABSTRACT

**INTRODUCTION:** With the arrival of disease-modifying drugs, neurodegenerative diseases will require an accurate diagnosis for optimal treatment. Convolutional neural networks are powerful deep learning techniques that can provide great help to physicians in image analysis. Deep learning classification of Alzheimer's disease (AD), frontotemporal dementia (FTD) or cognitively normal (CN) subjects based on brain glucose metabolism is therefore of interest but has not been previously investigated.

**METHODS:** Retrospective [18F]-FDG-PET scans of 199 AD, 192 FTD and 200 CN subjects were collected from our local database, Alzheimer's disease and frontotemporal lobar degeneration neuroimaging initiative databases. Training of a 3D VGG16-like neural network was performed on 90% of scans using data augmentation and cross-validation. Performance was compared to physicians' clinical interpretation in the 10% independent test set. Regions determining classification were identified in an occlusion experiment.

**RESULTS:** Test dataset subjects were age- and sex-matched. The model achieved an overall 89.8% accuracy in predicting the class of test scans. Areas under the ROC curves were 93.3% for AD, 95.3% for FTD, and 99.9% for CN. The physicians' consensus showed a 69.5% accuracy, and there was substantial agreement between them ( $\kappa = 0.61$ , 95% CI: 0.49-0.73). Anterior cingulate, medial temporal and dorsolateral prefrontal cortex were common regions to AD or FTD classification.

**CONCLUSION:** To our knowledge, this is the first study to introduce a neural network able to discriminate AD and FTD based on [18F]-FDG PET scans, and to isolate CN subjects with high accuracy. This adds to the growing literature of deep learning in nuclear medicine and suggests future incorporation of these techniques in clinical practice.

## INTRODUCTION

Alzheimer's disease (AD) is a debilitating condition that affects millions of people worldwide. It is the fourth leading cause of disability-adjusted life-year (DALY) in older adults and its prevalence is set to increase dramatically due to population aging<sup>2, 46</sup>. This reiterates the urgent need for effective disease-modifying drugs to treat neurodegenerative disorders.

An accurate diagnosis is paramount to the development of effective treatments because of the variety of pathological mechanisms at play in different neurodegenerative diseases. This is all the more challenging in diseases where clinical pictures overlap such as AD and the frontotemporal dementia (FTD) spectrum, given the fact that gold standard diagnosis relies on pathological findings, only possible after the patient's death, and the importance of an early diagnosis for optimal effects of disease-modifying drugs.

Current diagnostic criteria rely on a compatible clinical presentation, cerebrospinal fluid biomarkers and magnetic resonance imaging (MRI) or positron emission tomography (PET) findings<sup>27, 28, 47</sup>. Despite a moderate inter-observer variability in interpretation<sup>48</sup>, brain imaging thereby holds a crucial position in the diagnosis of cognitive disorders. Therefore, this calls for the development of more accurate and objective interpretation tools to diagnose patients with precision, but also to design clinical trials wherein inclusion/exclusion criteria are met with certainty.

Artificial intelligence (AI) techniques have been successfully applied in radiology, using advanced algorithms and big data to make predictions and improve diagnosis. These

techniques have become increasingly popular in the field of neuroimaging, especially in the area of structural imaging<sup>49-51</sup>. However, metabolic imaging can detect neuronal and glial glucose metabolism disturbance that may occur before structural changes in the brain<sup>52</sup>. For example, glucose metabolism alterations have been observed in ApoE4 carriers, who show metabolic abnormalities in posterior associative regions but do not display cognitive issues yet<sup>53</sup>. Furthermore, FDG-PET has been shown to help accurately reclassify subjects with cognitive disturbance who were previously misdiagnosed despite MRI scanning<sup>54</sup>. Altogether, this makes a strong case for applying AI techniques to metabolic imaging. A remarkable example of which is a recent work that classified with high accuracy [18F]-FDG-PET scans into dementia with Lewy bodies (DLB), Alzheimer's spectrum or cognitively normal using a 3D convolutional neural network (CNN)<sup>45</sup>.

The current study aims to demonstrate how a 3D CNN can facilitate the diagnosis between AD, FTD or normal aging using [18F]-FDG-PET. We hypothesize that this neural network will perform at least as well as expert-based classification. To test this, we will compare the network's performance with that of nuclear medicine physicians. Additionally, we will assess the network's ability to classify scans of dementia with Lewy bodies (DLB) and anxiety disorders. We predict that DLB scans will be predominantly labeled as AD and anxiety disorder scans as cognitively normal (CN). Finally, we will use an occlusion map to evaluate the network's attention to each category.

## METHODS

### Subjects

For an overall view of methods, see figure 1. We retrospectively collected [18F]-FDG-PET brain scans from the University of Lille (ULille) Memory Clinic database, the Alzheimer's Disease Neuroimaging Initiative (ADNI) 2 and 3 database, and the Frontotemporal Lobar Degeneration Neuroimaging Initiative (FTLDNI) database.

ADNI ([adni.loni.usc.edu](http://adni.loni.usc.edu)) started in 2003 as a public-private partnership. Its Principal Investigator was Michael W. Weiner, MD. The primary goal of ADNI has been to test whether imaging/biological markers, clinical and neuropsychological assessment can help to appraise the course of mild cognitive impairment (MCI) and early AD. For up-to-date information, see [www.adni-info.org](http://www.adni-info.org). FTLDNI was launched in 2010, funded by the National Institute of Aging, and led by Principal Investigator Howard Rosen, MD, University of California, San Francisco. These data are the result of collaborative efforts from 3 North American sites. For up-to-date information, visit <http://memory.ucsf.edu/research/studies/nifd>.

To prevent the neural network to learn acquisition procedure specificities, we aimed to build a multicentric dataset with around 200 scans per category. The AD category included 100 AD scans acquired between 2015 and 2021 randomly selected from the ULille database, 16 of which had a certain diagnosis of AD, and 84 had a probable diagnosis with typical CSF biomarkers, as well as 104 ADNI 2 & 3 scans with a "high confidence" diagnosis. The FTD category included FTLDNI database scans (N = 59), 4 of which did not have a diagnosis and were excluded, as well as all FTD scans acquired at ULille between 2015 and 2021 (N = 137), out of which 16 had a certain

diagnosis and 121 had a probable diagnosis of semantic variant of PPA, agrammatical variant of PPA or behavioral variant of FTD. The CN category included all control scans from FTLDNI (N = 36) and 164 randomly selected baseline scans from ADNI 2 & 3 (supplementary figure 1).

### **Data acquisition**

PET-scans from ULille database were acquired on a hybrid 4-ring Biograph mCT-Flow PET/CT with 20-slice CT and 4 x 4 mm<sup>2</sup> lutetium oxyorthosilicate crystals (Siemens Medical Solutions USA, Inc., Molecular Imaging, Knoxville, TN, USA). Mean tracer dose was 177 MBq (SD = 19 MBq). Thirty minutes post injection, a low-dose CT scan of the brain was acquired for attenuation correction of the PET data, and 10-minute emission images were subsequently acquired. The PET data were reconstructed iteratively using an ordered subset expectation maximization algorithm with 8 iterations and 21 subsets. The reconstruction process included decay, random and scatter corrections and 2-mm full width at half-maximum Gaussian kernel smoothing. For each PET examination, the reconstructed images comprised a series of 109 axial slices with the following parameters: field of view = 408 × 408 × 221.3 mm<sup>3</sup>, matrix = 400 × 400 × 109, and voxel size = 1.02 × 1.02 × 2.03 mm<sup>3</sup>.

For ADNI acquisition the protocol can be found at [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). FDG tracer dose was 185 MBq (+/- 10%), and between 30 and 60 min after injection 6 dynamic 3D scans of 5-min frames were acquired. A low-dose CT scan was acquired for attenuation correction, or for PET-only scanners an attenuation correction scan was acquired using rod sources. Every scan was reviewed by the PET quality control team.



All FTLDNI scans were acquired at the Mayo Clinic center on a GE Discovery RX PET/CT scanner. Participants were injected with 185 MBq (+/- 10%) of FDG and acquisition started 30 minutes later, consisting of six 5-minute dynamic frames. A CT-scan, obtained prior to injection of FDG, was used for attenuation correction, and reconstruction used 3D filtered-back projection technique<sup>55</sup>.

### **Image preprocessing**

Preprocessing was done using MATLAB R2020a (MathWorks, Natick, MA, USA) & the Statistical Parametric Mapping 12 (SPM12; Wellcome Trust Centre for Neuroimaging, London, UK, <http://www.fil.ion.ucl.ac.uk/spm>). PET scans from ADNI and FTLDNI databases were downloaded in NIFTI format. Each of them consisted of 6 NIFTI files, which were realigned and averaged. PET scans from the ULille database were downloaded in DICOM format and subsequently converted to NIFTI with `dcm2niix`<sup>56</sup>. We spatially normalized each scan using default parameters of SPM12 normalization function and the International Consortium of Brain Mapping (ICBM) template. Gray matter, white matter, cerebrospinal fluid, bone, soft tissue and air probability maps were estimated using default parameters of SPM12 segmentation function. Voxels with a probability > 0.7 of being gray matter, white matter or cerebrospinal fluid were kept, and voxels assumed to be bone, soft tissue or air were discarded. Feature-wise normalization was done through dividing each voxel by the maximum of the 3D scan it belonged to, so that all voxel values were between [0, 1]. Every scan was inspected in MRICron (<https://www.nitrc.org/projects/mricron>) to ensure correct preprocessing. After visual inspection, 5 ADNI AD scans were excluded because of negative value voxels, which may have conflicted with the use of the ReLU

function. Nineteen FTLDNI FTD, 6 ULille FTD, 1 ULille AD and 1 ADNI CN scans (N = 27, 4,6%) were noticed to have segmentation issues, and were reprocessed changing the default clean up parameter from light cleanup to no cleanup. Preprocessing resulted in 3D volumes of 95 x 79 x 79 voxels (2 mm isotropic). Since top and bottom layers contain cerebral regions considered not relevant for classification (primary sensorimotor cortex, cerebellum), they were removed, and final volume dimensions were 95 x 79 x 60.

### **Datasets and data augmentation**

The whole dataset contained 591 preprocessed [18F]-FDG PET scans. Training and testing datasets were built on a 90% - 10% basis, resulting in 532 scans in the training dataset and 59 scans in the testing dataset.

To prevent overfitting of the model, we augmented the training dataset on-the-fly using tensorflow loader. Flipping along the sagittal plan, and +/- 10° random rotations across the 3 plans were performed in 50% of cases to limit computational cost. Ten-percent translations across the 3 axes were performed for each scan, +/- 8, +/- 10 and +/- 6 voxels along the x, y and z axis, respectively.

### **Neural network building**

Training was done at the Lille In Vivo Imaging and Functional Exploration (LIIFE) research lab at Lille University Hospitals, on a computer with Linux Ubuntu 20.04 operating system, 12 CPU Intel® Xeon® W-2133 3.60GHz for a total of 102GB of memory, and an NVIDIA Quadro RTX 6000 with 24GB of memory. To facilitate the selection of efficient hyperparameters, we used Bayesian optimization through a large

range of hyperparameters for 200 iterations over the native training dataset using 5-fold cross validation, 150 epochs and an early stopping callback with patience = 10 monitoring validation loss. After examination of the Bayesian search, we retained a batch size of 6, Adagrad optimizer, a 0.5 dropout rate, a 0.0005 learning rate, and we performed end-to-end training on the augmented dataset for 150 epochs keeping an early stopping callback with patience = 20 based on validation loss and saving only the best model based on validation accuracy. Finally, the model with the highest validation accuracy was retrained on the whole non-augmented dataset with stochastic gradient descent with 0.0001 learning rate and 0.9 momentum for 50 epochs.

### **Clinical Interpretation**

Clinical interpretation of the testing set was performed by 2 French board-accredited nuclear medicine physicians, and a resident in nuclear medicine with respectively 12, 9 years and 3 years of experience in the field of nuclear medicine. Non-normalized native brain volumes were visualized using ITK-SNAP ([www.itksnap.org](http://www.itksnap.org))<sup>57</sup>, without any clinical information. Due to physicians' preferences, a French rainbow lookup table (LUT) was used for visualization (supplementary figure 2). Fleiss' kappa coefficient was calculated to evaluate the interobserver agreement. Scans for which there was diagnostic disagreement were reviewed in order to reach consensual agreement.

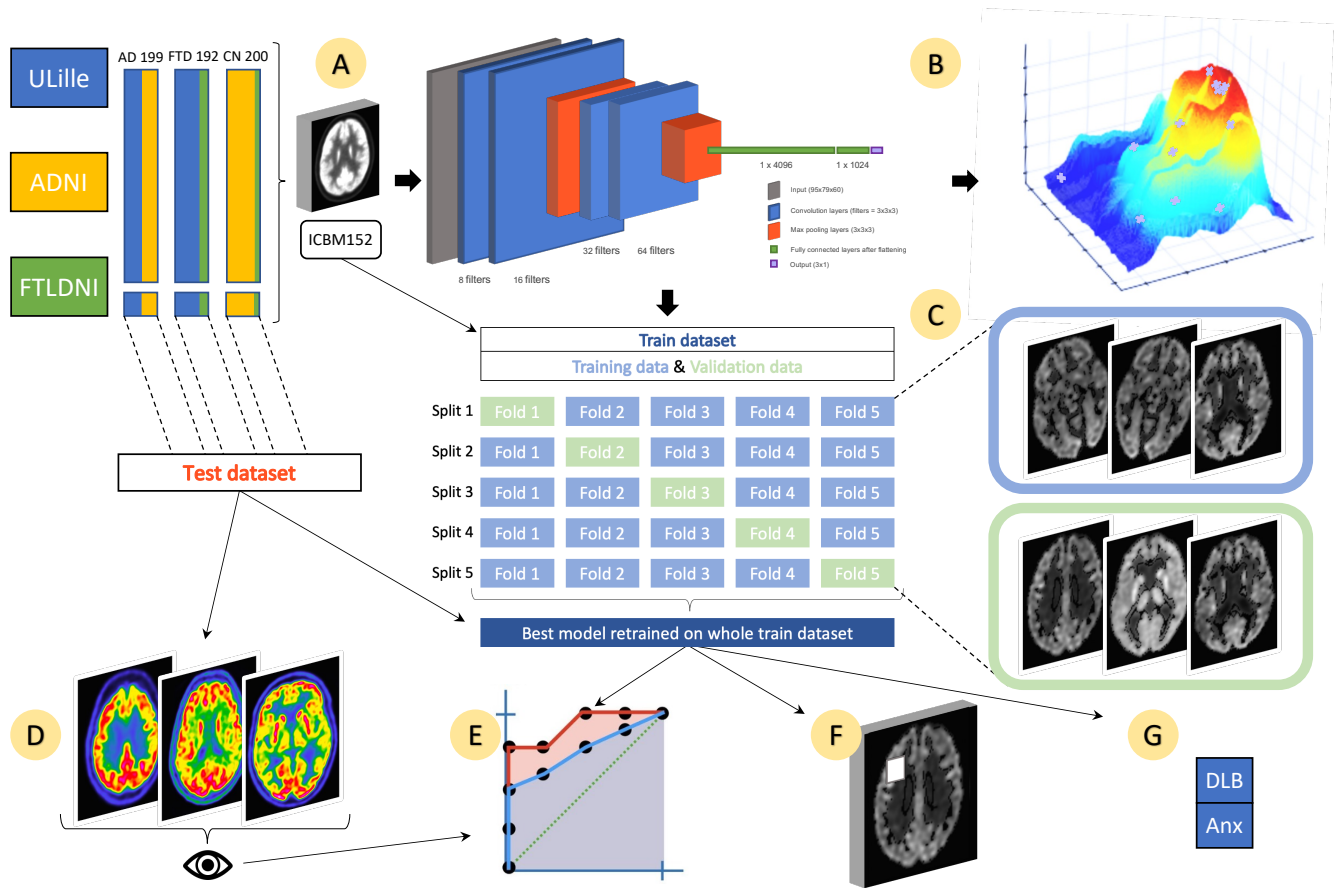
### **Network attention**

Building on the assumption that the cross-entropy response will substantially increase when voxels relevant for classification are occluded, we performed an occlusion experiment over the non-augmented training dataset<sup>58</sup>. A 5-voxel cube was removed

from each normalized scan with a stride of 2 for all 3 directions. The cross-entropy variations for each scan are plotted as a function of the position of the occlusion cube, generating 38 x 46 x 29 voxel volumes. These volumes were normalized through dividing by their max, resized to 75 x 91 x 56 voxel volumes through third order spline interpolation, and averaged over each category. Since voxels outside the brain should not influence classification, we inferred the threshold for significant influence of prediction is represented by the minimum voxel outside the brain, and therefore all heatmap voxels equal or superior to this value were removed. To show regions exclusively common to neurodegenerative classification, a fourth map was generated through intersection of AD and FTD maps and exclusion of voxels belonging to the CN map. Finally, heatmaps are displayed over each category's averaged brain, and for the intersection map over the overall averaged brain.

### **Network extrapolability**

With the purpose of testing the network on significantly different subjects, 20 DLB scans and 16 anxiety disorder scans were included from the ULille database. Although brain glucose metabolism alterations are reported in anxiety disorders, FDG-PET is not recommended by the European Association of Nuclear Medicine (EANM) guidelines in psychiatric disorders<sup>59</sup>, probably because of the lack of evident hypometabolism pattern, and in our clinical experience anxiety disorders abnormalities remain more subtle compared to other psychiatric disorders. Nineteen DLB scans had a probable diagnosis and 1 had a certain diagnosis as per the consensus diagnostic criteria<sup>60</sup>, and anxiety disorder was a diagnosis of exclusion. Image preprocessing was done identically to scans from training/testing datasets.



**Figure 1: Overview of methods.** (A) Dataset of 199 AD, 192 FTD and 200 CN scans from ULille, ADNI and FTLDNI databases split between train & test sets on a 90%-10% basis. Scans featurewise & spatially normalized to ICBM152 template (B) Network architecture and Bayes search to select hyperparameters. (C) Training on augmented data through fivefold cross-validation, and best model kept for retraining on whole train dataset. (D) Clinical interpretation by 3 physicians of the native scans. (E) Performance comparison in the independent test set between model and physicians' consensual agreement. (F) Occlusion experiment in train set. (G) Predictions of the model of DLB and Anx subjects. AD, Alzheimer's disease; ADNI, Alzheimer's disease neuroimaging initiative; Anx, anxiety disorder; CN, cognitively normal; DLB, dementia with Lewy bodies; FTD, frontotemporal dementia; FTLDNI, frontotemporal lobar degeneration neuroimaging initiative; ICBM, International Consortium for Brain Mapping; ULille, University of Lille.

## RESULTS

The presented study follows the framework of the reference methodology MR004 introduced by the *Commission Nationale de l'Informatique et des Libertés* (CNIL), and we verified ULille subjects did not oppose the use of their clinical data for research purposes.

### Demographics

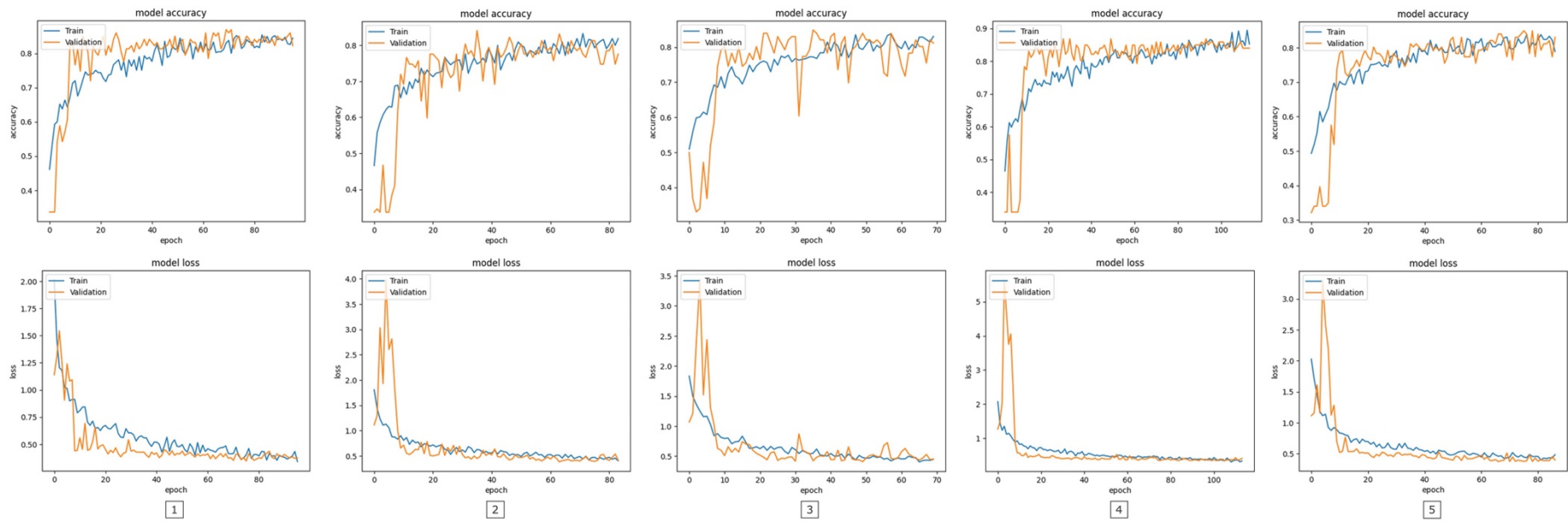
Subject demographics are summarized in table 1. There were no significant differences between classes of the test dataset in terms of sex or age ( $p > 0.05$ ). CN subjects showed a significantly higher MMSE than AD or FTD subjects ( $p < 0.0001$  each), and AD subjects also showed better scores than FTD subjects ( $p < 0.0001$ ).

**Table 1: Data demographics.** \* Some subjects had several scans acquired at different timepoints. \*\* All subjects with a certain diagnosis belonged to ULille. All ADNI subjects had a probable diagnosis (NINCDS-ADRDA criteria). Diagnostic certitude is not reported in FTLDNI. \*\*\* Two AD subjects and 12 FTD subjects did not have a MMSE score within a year of scanning. AD, Alzheimer's disease; ADNI, Alzheimer's disease neuroimaging initiative; CN, cognitively normal; FTD, frontotemporal dementia; FTLDNI, frontotemporal lobar degeneration neuroimaging initiative; MCI, mild cognitive impairment; SD, standard deviation; ULille, University of Lille.

Train set									
No. subjects*	AD			FTD				CN	Total
No. scans	170			126				155	451
Diagnosis	179			173				180	532
	Probable with biomarkers	Clinically probable	Certain**	BV	NFV	SV	Certain**	Converted later to MCI/dementia	
	42.5%	49.7%	7.8%	79.2%	5.8%	5.8%	9.2%	0	
Source									
ULille	50.3%			71.1%				0	40%
ADNI	49.7%			0				81.7%	44.4%
FTLDNI	0			28.9%				18.3%	15.6%
Clinical data									
Age [mean (SD)]	68.2 (10.7)			68.4 (8.7)				71.5 (7.9)	69.4 (9.3)
Female Sex	53.6%			37%				56.7%	49.5%
MMSE [mean (SD)]	19.3 (6.2)			21.9 (6.5)				29.1 (1.2)	
Test set									
No. subjects*	AD			FTD				CN	Total
No. scans	20			18				20	58
Diagnosis	20			19				20	59
	Probable with biomarkers	Clinically probable	Certain**	BV	NFV	SV	Certain**	Converted later to MCI/dementia	
	50%	50%	0	94.7%	5.3%	0	0	16.1%	
Source									
ULille	50%			73.7%				0	40.7%
ADNI	50%			0				85%	45.8%
FTLDNI	0			26.3%				15%	13.5%
Clinical data									
Age [mean (SD)]	69.9 (7.5)			68.6 (6.8)				72.1 (7.2)	70.3 (7.2)
Female Sex	35%			36.8%				50%	40.7%
MMSE [mean (SD)]	22.8 (2.9)			12.4 (10.6)				28.9 (1.0)	
Overall dataset									
No. subjects*	AD			FTD				CN	Total
No. scans	190			136				172	498
Diagnosis	199			192				200	591
	Probable with biomarkers	Clinically probable	Certain**	BV	NFV	SV	Certain**	Converted later to MCI/dementia	
	43.2%	49.8%	7%	80.7%	5.7%	5.2%	8.3%	14.5%	
Source									
ULille	50.3%			71.4%				0	40.1%
ADNI	49.7%			0				82%	44.5%
FTLDNI	0			28.6%				18%	15.4%
Clinical data									
Age [mean (SD)]	68.4 (10.4)			68.4 (8.5)				71.6 (7.8)	69.5 (9.1)
Female Sex	51.8%			37%				56%	48.4%
MMSE [mean (SD)]***	19.5 (6.2)			21.7 (6.8)				29.1 (1.2)	

## Model training

Training of 200 models using Bayesian optimization was completed in 49 hours and 31 minutes. The results of the 20 best performing models can be found at supplementary table 1. Fivefold cross-validation training of the final model with data augmentation on training set took 1 hour and 58 minutes to complete. Model 1 was chosen for retraining on the whole train dataset, as it demonstrated the highest validation accuracy of 91.5% (figure 2).



**Figure 2: Model training using 5-fold cross validation.** Top graphs show train and validation accuracies. Bottom graphs show train and validation losses. Model 1 had the highest validation accuracy and was selected for further training.



## Model evaluation

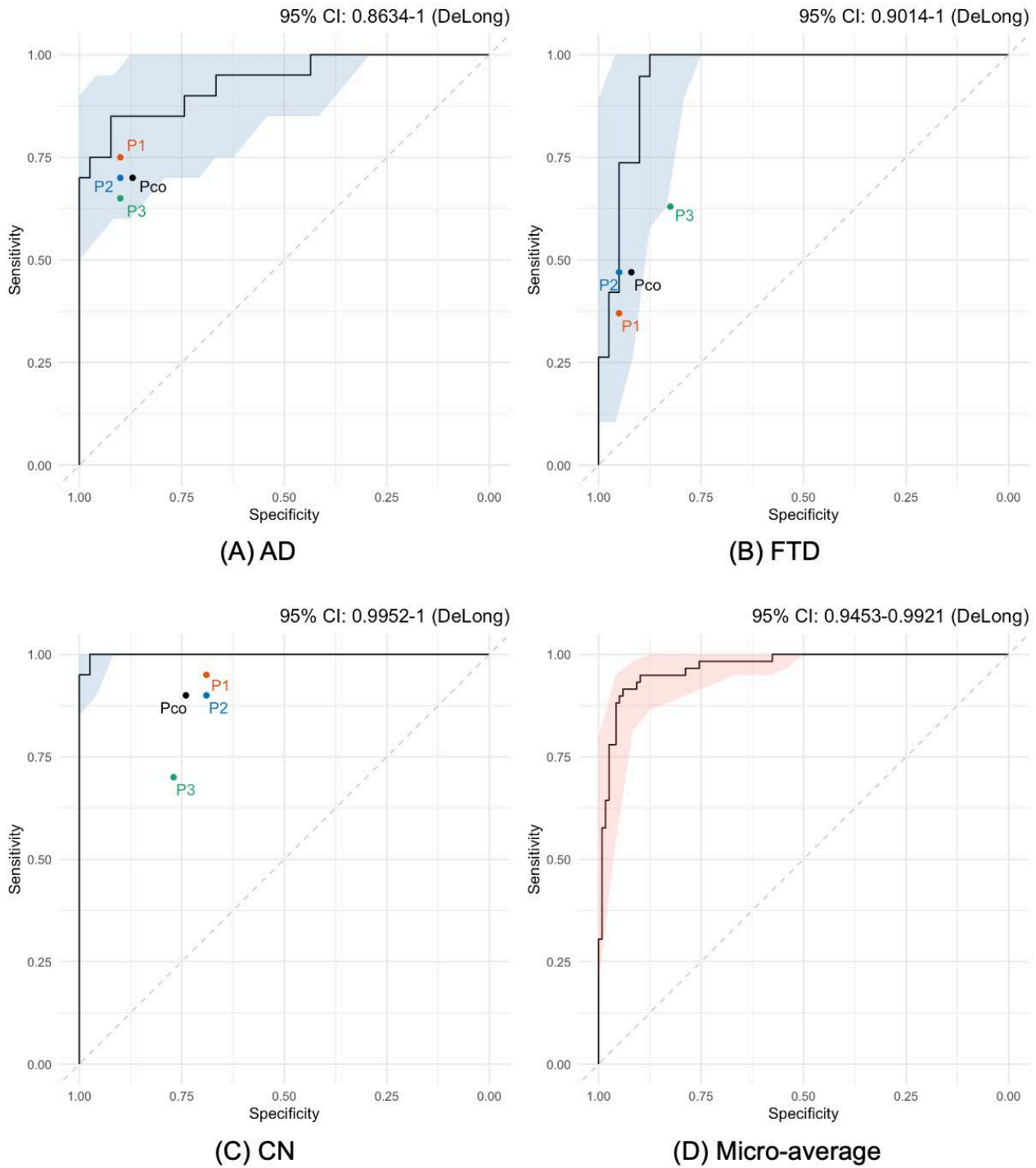
The overall accuracy on the test set of the final model was 89.7%. Model label classification is described in detail in table 2. The AUC for predictions were 93.3% for AD, 95.3% for FTD and 99.9% for CN (figure 3). The model performed best in detecting CN with 100% sensitivity (all 20 CN cases were detected), 97% specificity (38 out of 39 non-CN cases were correctly ruled out by the model), 95% precision (20 out of 21 cases labeled as CN were correctly classified), and F1 score 98%. Detailed metrics for AD and FTD can be found in table 3.

**Table 2: Confusion matrix.** AD, Alzheimer's disease; CN, cognitively normal; FTD, frontotemporal dementia.

		Actual labels				Total
		AD	FTD	CN	Total	
Model labels	AD	15	1	0	16	
	FTD	4	18	0	22	
	CN	1	0	20	21	
	Total	20	19	20	59	
Physician labels	AD	14	3	2	19	
	FTD	3	9	0	12	
	CN	3	7	18	28	
	Total	20	19	20	59	

**Table 3: Metrics.** AD, Alzheimer's disease; CN, cognitively normal; FTD, frontotemporal dementia; Se, sensitivity; Sp, specificity.

		Metrics				
		Se	Sp	Precision	F1-score	Support
Model	AD	75% (15/20)	97% (38/39)	94% (15/16)	83%	20
	FTD	95% (18/19)	90% (36/40)	82% (18/22)	88%	19
	CN	100% (20/20)	97% (38/39)	95% (20/21)	98%	20
Physician	AD	70% (14/20)	87% (34/39)	74% (14/19)	72%	20
	FTD	47% (9/19)	92% (37/40)	75% (9/12)	58%	19
	CN	90% (18/20)	74% (29/39)	64% (18/28)	75%	20



**Figure 3: ROC curves.** AD, Alzheimer's disease; CN, cognitively normal; FTD, frontotemporal; P1, physician 1; P2, physician 2; P3, physician 3; Pco, consensus between physicians; ROC, receiver operating characteristic.

## **Clinical interpretation**

Across the testing dataset, there was substantial agreement between physicians, as evidenced by a Fleiss' kappa coefficient of 0.61 (95% CI: 0.49-0.73). Consensual physician labeling (table 2) had a 69.5% accuracy (N = 41). The highest specificity was observed for FTD at 92.5%, but this contrasted with a low sensitivity of 47%. Comparatively, CN showed the highest sensitivity of 90% (table 3). Further information on labeling and detailed metrics per physician can be found in supplementary tables 2 and 3.

## **Network attention**

The most prominent regions in driving AD classification were found to be the medial temporal, lateral temporal, dorsolateral frontal, parieto-occipital, anterior cingulate cortex (ACC), and particularly the precuneus and posterior cingulate cortex (PCC). The FTD heatmap showed that mostly anterior regions, such as the dorsolateral and mesial frontal regions, and to a lesser extent the medial temporal and cerebellum regions, were driving classification. The CN heatmap showed a more diffuse pattern, with the orbitofrontal, associative visual areas, and insular regions being the most significant. Finally, regions common to AD or FTD classification, but not found in CN classification, were the anterior cingulate, medial temporal, and dorsolateral prefrontal regions (figure 4).

## **Network extrapolability**

Demographics of subjects with DLB or anxiety disorder can be found in supplementary table 3. Thirteen DLB scans were classified as FTD, 7 as AD and 0 as CN. Four anxiety disorder scans were classified as CN, 2 as AD and 10 as FTD.



**Figure 4: Occlusion heatmaps.** Top-left: Alzheimer's disease, top-right: frontotemporal dementia, bottom-left cognitively normal, bottom-right: intersection between Alzheimer's disease and frontotemporal dementia with exclusion of cognitively normal voxels. Color-bars show the variation in prediction certitude.

## DISCUSSION

The future of nuclear medicine in neurodegenerative diseases may be multitracer-based with the emergence of amyloid and tau PET, and multiscan-based to evaluate the efficacy of disease-modifying drugs in decreasing the protein load and the evolution of brain metabolism over time. In this study, we introduce a model able to classify [18F]-FDG-PET scans into AD, FTD or CN subjects with high accuracy, doing better than a consensus reached by physicians specialized in nuclear medicine in all metrics. Such a tool would be of great assistance to physicians to break free from intra- and inter-observer variability in clinical strategies involving repeated PET scans.

We extend the previous work from Etminani *et al* by presenting a CNN model also highly efficient in differentiating AD from FTD-spectrum diseases<sup>45</sup>. They reported in their work 96.4%, 71.4%, 96.2% and 94.7% AUC for AD, MCI, DLB and CN, respectively. Our model tends to show slightly better performances, but this should be put into perspective with the fact we did not include MCI subjects, which, understandably, was the most misclassified category in their work. Regarding clinical interpretation, while the consensus accuracy is similar at 69.5% in the present study (compared to 57% in Etminani *et al* with an additional category), our Fleiss' kappa of 0.61 (compared to 0.19) indicates stronger agreement. This low variance in physician performance reflects a shared understanding of the test set and suggests that it is unlikely for another physician to outperform the model. The strength of our model may be that it shows very high accuracy in isolating CN subjects, accurately classifying all CN subjects, and erroneously labeling just 1 AD subject as CN. This is interesting, since a first application for AI in clinical practice might be to rapidly screen brain scans and presents them to the physician as high-probability to be normal/abnormal to speed

up analysis and improve confidence, as shown in other fields<sup>61, 62</sup>. We also show deep learning models can easily be trained at the local scale and almost ready-to-use, given the vast amount of usable data in a hospital. Furthermore, deep learning classification of AD, FTD and CN has already been presented using structural MRI and demonstrated a 91.8% overall accuracy<sup>49</sup>. An ultimate approach in the future might be to combine PET and MRI data to further improve accuracy.

In the occlusion experiment, the PCC was found to be the most important region by large to drive AD classification. This is consistent with the traditional hypometabolism pattern described in visual interpretation, but in clinical practice the emphasis tends to be put on other posterior associative areas such as temporoparietal junctions, and therefore it reiterates the need to scrupulously examine the PCC when facing a possible diagnosis of AD. The PCC is an important area for cognition showing decreased [18F]-FDG uptake in early AD<sup>63</sup> and also, although inconsistently, decreased uptake in FTD<sup>64</sup>. Contrarily to the aforementioned work, in which the PCC was a common driver for classification into AD, MCI or DLB<sup>45</sup>, herein the PCC was not found to guide classification towards FTD, and consequently should not be considered as a hallmark of neurodegenerative disease. As expected, we found an anteroposterior gradient for regions driving FTD classification, but it came as a surprise that the cerebellum was also a decisive region. Following this result, we visualized FTD scans and it did seem many scans showed a certain degree of cerebellar hypometabolism. The cerebellum is debated to show altered cerebellar glucose metabolism in C9orf72 mutations (N = 10 in the train set) vs non-C9orf72 FTD<sup>65, 66</sup>. However, it is considered intact in most neurodegenerative disorders and recommended as a reference region for intensity normalization by the EANM<sup>59</sup>. This

calls for further investigation, and it may be interesting in the future to do a voxelwise analysis between non-C9orf72 FTD and AD using a cerebellar mask. Finally, it seemed mostly anterior regions such as the ACC and the dorsolateral prefrontal cortex, or mesial temporal regions were found to be common to AD or FTD classification and absent in CN classification.

In testing extrapolability to other neuropsychiatric disorders, we found that all DLB subjects were classified as AD or FTD. This highlights the coherence of our model in successfully ruling out subjects with another neurodegenerative disease as healthy subjects, which would be the most prejudicious situation in a clinical setting. This is also in line with the findings of Etminani *et al*, where their dimensionality reduction visualization clearly separated CN and DLB. On the other hand, whereas we expected most subjects with an anxiety disorder to be classified as CN, 75% were classified as AD or FTD. Although abnormal metabolic patterns are not traditionally described in psychiatric disorders, significant hypometabolism has already been demonstrated predominantly in the frontal cortex<sup>67</sup>, including in anxiety disorders<sup>68, 69</sup>. Therefore, it should come as no surprise the present model labeled these subjects as disease subjects, and especially as FTD.

### **Strengths**

As shown in figure 2, we successfully prevented overfitting during the cross-validated training using several regularization methods such as data augmentation, batch normalization, dropout and early stopping. This, we believe, shows that the network could learn robust features during training that would allow its extrapolation to new brain scans with a different acquisition protocol. Similarly, its training over a large

multicentric dataset, well-balanced across the different categories would also facilitate utilization in different conditions. Using a large test set (N = 59), we are also confident in the accuracy of its performance.

Another strength is that we could include subjects who received a certain diagnosis (N = 30). This is of interest since there is a large overlap in radiological presentations of neurodegenerative diseases<sup>31</sup>, and protein-targeting drugs will require an unequivocal diagnosis. Except subjects from FTL/DNI for which diagnostic certainty was not reported, the rest of our subject had a probable diagnosis. Longitudinal analyses have shown individuals fulfilling FTD probable criteria will continue to do so over time or move to the certain category after postmortem analysis<sup>70</sup>. This is, however, to put into perspective with the fact that most often neurodegenerative pathologies overlap, as pure AD is thought to only represent 30% of cases and in 30% of cases associates with TDP-43 pathology<sup>11</sup>.

## **Limitations**

Unfortunately, a same center could not provide subjects to all 3 categories, which would assure the network did not learn any acquisition specificities. For example, the CN category did not include any ULille subjects. This is because our database, built with clinical data, does not include any healthy control subjects. This could also be viewed as a wider limit to the development of AI in clinical practice since healthy controls are not found among hospital patients in usual conditions.

It could be argued that subjects having several scans may have led to inflated performances of the model. However, we do not think this is a serious issue since



these are different scans with different acquisition and radiotracer activity and thereby different voxel values. Moreover, except 1 subject who had 1-month interval scans, there was a minimum of 6 months between scans for other subjects which is enough to see progressive metabolic changes<sup>71</sup>.

Probably because of regions showing severe hypometabolism, SPM segmentation basic parameters first considered large cortical areas as not gray matter, and these scans (N = 27) required another segmentation treatment. This is a minor issue in the current study, but it reflects what may happen when developing an AI pipeline. Even if AI is successfully incorporated in imaging departments, this highlights the need for quality checks to ensure coherent results.

## **CONCLUSION**

In the current study, we demonstrated the ability of a tailor-made 3D CNN to accurately classify [18F]-FDG PET-scans between AD, FTD or CN subjects. We showed its performance was better than traditional clinical interpretation by experienced physicians. These results add to the growing field of AI in metabolic imaging and suggest our clinical practice may change in the nearby future integrating these tools. To our knowledge, this is the first work using deep learning techniques to classify AD vs. FTD subjects based on brain glucose metabolism.

### *Code availability*

Scripts are publicly available on GitHub ([https://github.com/rogeauA/NeuralNet\\_AD-FTD\\_FDGPET](https://github.com/rogeauA/NeuralNet_AD-FTD_FDGPET)).

## **ACKNOWLEDGEMENTS**

We do not have any conflicts of interest to declare.

Public data collection and sharing was funded by ADNI (National Institutes of Health Grant U01 AG024904), DOD ADNI (Department of Defense award number W81XWH-12-2-0012), and FTLDNI (National Institutes of Health Grant R01 AG032306).

ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. Canadian sites receive funds from the Canadian Institutes of Health Research. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education.

ADNI is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California, and FTLDNI is coordinated by the University of

California, San Francisco, Memory and Aging Center. ADNI and FTLDNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

### III. DISCUSSION

---

Ce travail démontre la faisabilité d'entraîner un réseau de neurones qui surpasse les performances diagnostiques de 3 médecins nucléaires dans l'étude de TEP cérébrales au [18F]-FDG. Nous montrons également que l'attention de ce réseau de neurones est rationnelle vis-à-vis des anomalies de métabolisme connues, et que son extrapolation à d'autres catégories diagnostiques donne des résultats pertinents. L'objet de cette discussion finale sera de discuter plusieurs pistes de développement futur.

#### 1) Suggestions pour le développement d'un futur logiciel d'IA applicable

Bien que nous démontrions dans ce travail que des logiciels d'IA peuvent facilement être conçus à l'échelle locale, il serait important dans le futur que leur développement soit multicentrique. D'une part, utiliser un logiciel d'IA permet de s'affranchir de la variabilité inter-observateur, c'est pourquoi la multiplication des logiciels n'est pas désirable puisque cela réintroduirait une variabilité inter-logiciel, et ainsi l'uniformisation des pratiques permettrait de faciliter le dialogue avec le clinicien. D'autre part, la coopération multicentrique permettrait d'élargir considérablement les bases de données, notamment dans les maladies neurodégénératives où le diagnostic de certitude est posé par l'autopsie, dont la pratique a diminué avec un taux estimé à 2-16%<sup>72</sup>, il est ainsi essentiel de pouvoir partager ces données. Finalement, comme vu précédemment, le caractère multicentrique d'une base de données offre une variété de paramètres d'acquisition et aboutit à une meilleure extrapolation.

En lien avec le travail présenté, le développement futur le plus évident serait d'intégrer les autres catégories diagnostiques de maladies neurodégénératives. Afin d'optimiser

l'apprentissage d'un nouveau réseau, les méthodes d'apprentissage par transfert ou *transfer learning* seraient les plus appropriées. Le principe de celles-ci est de reprendre un modèle déjà entraîné, d'empêcher les premières couches, identifiant les formes les plus simples, de modifier leurs poids, et d'ajuster le modèle en le ré-entraînant sur de nouvelles données (figure 9). Par rapport à un modèle entraîné *de novo*, de meilleures performances ont été montrées dans la classification de cancer colorectal alors que le modèle de base avait été pré-entraîné sur des données très différentes<sup>73</sup>.

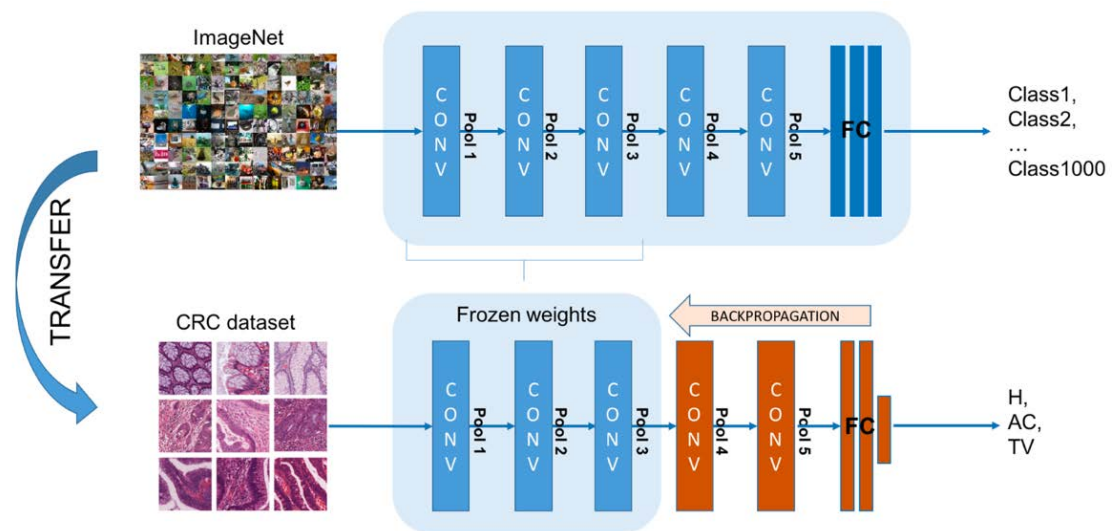


Figure 9 : Méthode d'apprentissage par transfert. Extrait de Ponzio *et al* (2018)<sup>73</sup>.

## 2) Future place de la médecine nucléaire en neurologie cognitive

Bien que la TEP au FDG soit un outil diagnostique important, tout comme l'IRM morphologique ou même la clinique, son principe ne permettra jamais de diagnostiquer avec certitude une maladie neurodégénérative telle que définie dans le paragraphe 1 de l'introduction.

Cependant, nous sommes probablement au début de l'âge d'or de l'imagerie moléculaire dans l'étude du cerveau, en neurologie et en psychiatrie<sup>74</sup>. En neurodégénératif, le développement le plus évident, ayant déjà débuté, est l'apparition de radiotraceurs spécifiques à la pathologie de chaque entité (figure 10).

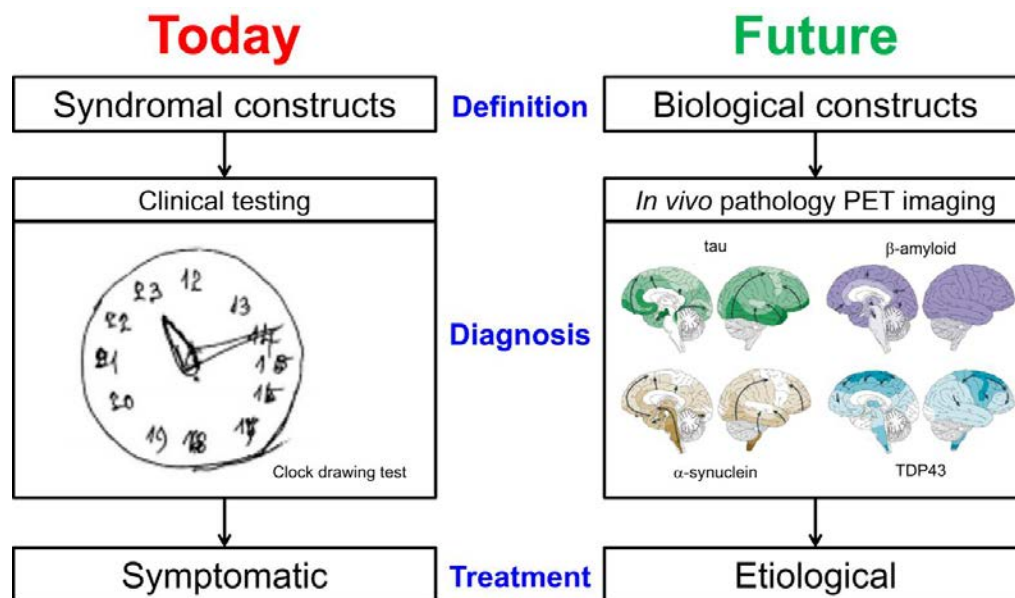


Figure 10 : Schéma du changement de paradigme dans la définition, le diagnostic et le traitement des maladies neurodégénératives. Extrait de Weber *et al* (2020)<sup>75</sup>.

De nouveaux traceurs étudiant la neurodégénérescence de manière générale verront aussi le jour, par exemple, les traceurs de densité synaptique ciblant la *synaptic vesicle glycoprotein 2A* (SV2A). Il a été récemment montré que leur sensibilité pouvait être meilleure que celle de la TEP au FDG dans les formes précoces de maladie d'Huntington ou de l'IRM volumétrique dans les formes précoces de MA<sup>76, 77</sup>. Ceux-ci auront un rôle évident à jouer dans les essais cliniques de nouveaux traitements, mais également en clinique comme pour décider un changement thérapeutique ou une intensification d'un traitement en cours.

Les améliorations techniques, comme la TEP corps entier, favoriseront aussi l'essor de nouvelles applications. L'acquisition d'images dynamiques, grâce à celle-ci, permettra de se dispenser d'une mesure invasive de la fonction artérielle d'entrée<sup>78</sup>. Ainsi, l'estimation de la constante  $K_i$ , mieux précise que la *standard uptake value* (SUV) qui est influencée par de nombreux facteurs, sera simplifiée. Elle permettra également l'obtention de multiples paramètres permettant d'évaluer la vitesse d'un traceur pour traverser la barrière hémato-encéphalique, pour se fixer sur son récepteur et pour être éliminé du cerveau, potentiellement éclairant la physiopathologie des maladies<sup>79</sup>.

Enfin, il est discutable qu'une partie de ces informations soit redondante avec les tests biologiques, notamment avec les biomarqueurs LCS de tau, de peptide amyloïde ou de neurofilaments. Sans compter la dimension spatiale supplémentaire à l'imagerie, il est possible d'imaginer qu'ensemble ils permettront le développement de stratégies diagnostiques par les neurologues comme déjà le cas dans d'autres pathologies (figure 11).

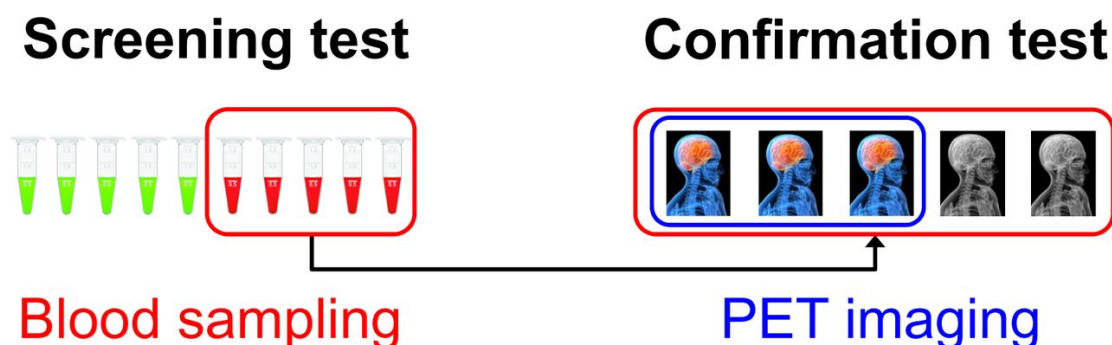


Figure 11 : Possible future place des examens d'imagerie moléculaire vis-à-vis des tests biologiques. Extrait de Barthel *et al* (2022)<sup>74</sup>.

### 3) Vers une fin de l'interprétation visuelle ?

Déjà présents dans certaines applications, la prochaine décennie continuera à voir l'apparition d'algorithmes d'IA afin d'optimiser le temps des radiologues et des médecins nucléaires, avec des logiciels de segmentation automatique<sup>80</sup>, de reconstruction d'images<sup>81</sup>, et de gestion des patients<sup>82</sup>. Ainsi l'interprétation d'images semble encore la propriété des médecins imageurs. Pourtant, en 2022, l'Union Européenne a accordé le marquage CE à l'outil ChestLink produit par l'entreprise Oxipit<sup>83</sup>. Celui-ci détecte avec une sensibilité de 99.8% les radiographies thoraciques normales, génère un compte-rendu dans le même temps, et dans le cadre d'un examen anormal, peut suggérer certains diagnostics.

D'un point de vue personnel, je pense que dans un futur plus lointain l'interprétation visuelle est certainement amenée à changer et probablement à diminuer en proportion d'examens réalisés (peut-être pas en absolu si le nombre d'examens augmente exponentiellement). Les performances obtenues par un logiciel d'IA et le temps d'interprétation nettement plus court apporteront de manière ultime un bénéfice au patient et seront demandés par une politique d'efficience. Les progrès informatiques tels que les calculateurs quantiques permettront un réel essor du *big data*. Sans pouvoir prédire une réelle temporalité, les examens à réponse binaire tels que la scintigraphie splénique seront les premiers impactés, puis viendront les examens où l'information reste limitée avec un diagnostic multi-classe tels que la TEP cérébrale, puis les derniers seront ceux avec une grande quantité d'information telle que la TDM-TAP. Est-ce pour autant la fin des médecins imageurs ? Non, puisque le contrôle qualité restera indispensable et pour les cas où l'algorithme est incertain de son interprétation, le médecin devra continuer à regarder les images afin de trancher.



Même si l'IA reste une technologie nouvelle, ce type de changement, lié à l'automatisation des tâches, a déjà eu lieu dans d'autres professions : le pharmacien d'officine ne fabrique que rarement des médicaments, le médecin biologiste ne compte plus lui-même les cellules sur une lame de microscope. Finalement, il semble essentiel que les médecins imageurs embrassent ces développements et qu'ils maîtrisent dans le futur les notions d'IA afin de prévenir une accapitation par les entreprises *high-tech*.

## **IV. CONCLUSION**

---

Ce travail démontre la capacité d'un réseau de neurones 3D à classifier avec précision les TEP au [18F]-FDG entre MA, DFT ou CS. Ses performances sont meilleures qu'un consensus atteint par 3 médecins nucléaires. Les cortex cingulaire antérieur, préfrontal dorsolatéral et temporal mésial sont des régions communes à la classification en MA ou DFT et ne guidant pas la classification en CS. Extrapolant ce modèle à d'autres catégories diagnostiques, aucun sujet DCL n'était classé en CS et la plupart des sujets avec un trouble anxieux étaient classés en tant que DFT. Ces résultats apparaissent dans un contexte d'engouement grandissant pour l'IA et suggèrent que de telles méthodes entreront le quotidien des médecins imageurs dans le futur proche.

## BIBLIOGRAPHY

1. American Psychiatric Association. Neurocognitive Disorders. *Diagnostic and Statistical Manual of Mental Disorders*, 5th, text rev. edn: Washington, DC, 2022.
2. Collaborators GBDDF. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *Lancet Public Health* 2022; **7**(2): e105-e125.
3. Jutkowitz E, Kane RL, Gaugler JE, MacLehose RF, Dowd B, Kuntz KM. Societal and Family Lifetime Cost of Dementia: Implications for Policy. *J Am Geriatr Soc* 2017; **65**(10): 2169-2175.
4. Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B *et al.* Criteria for the diagnosis of corticobasal degeneration. *Neurology* 2013; **80**(5): 496-503.
5. Dhillon S. Aducanumab: First Approval. *Drugs* 2021; **81**(12): 1437-1443.
6. Tahami Monfared AA, Tafazzoli A, Ye W, Chavan A, Zhang Q. Long-Term Health Outcomes of Lecanemab in Patients with Early Alzheimer's Disease Using Simulation Modeling. *Neurol Ther* 2022; **11**(2): 863-880.
7. van Dyck CH, Swanson CJ, Aisen P, Bateman RJ, Chen C, Gee M *et al.* Lecanemab in Early Alzheimer's Disease. *N Engl J Med* 2022.
8. Jeremic D, Jimenez-Diaz L, Navarro-Lopez JD. Past, present and future of therapeutic strategies against amyloid-beta peptides in Alzheimer's disease: a systematic review. *Ageing Res Rev* 2021; **72**: 101496.
9. Congdon EE, Sigurdsson EM. Tau-targeting therapies for Alzheimer disease. *Nat Rev Neurol* 2018; **14**(7): 399-415.
10. Association As. 2016 Alzheimer's disease facts and figures. *Alzheimers Dement* 2016; **12**(4): 459-509.
11. Villain N, Dubois B. Alzheimer's Disease Including Focal Presentations. *Semin Neurol* 2019; **39**(2): 213-226.
12. Sims R, Hill M, Williams J. The multiplex model of the genetics of Alzheimer's disease. *Nat Neurosci* 2020; **23**(3): 311-322.
13. Jack CR, Jr., Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB *et al.* NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* 2018; **14**(4): 535-562.
14. Ittner LM, Gotz J. Amyloid-beta and tau--a toxic pas de deux in Alzheimer's disease. *Nat Rev Neurosci* 2011; **12**(2): 65-72.

15. Roberts RO, Aakre JA, Kremers WK, Vassilaki M, Knopman DS, Mielke MM *et al.* Prevalence and Outcomes of Amyloid Positivity Among Persons Without Dementia in a Longitudinal, Population-Based Setting. *JAMA Neurol* 2018; **75**(8): 970-979.
16. Salloway S, Chalkias S, Barkhof F, Burkett P, Barakos J, Purcell D *et al.* Amyloid-Related Imaging Abnormalities in 2 Phase 3 Studies Evaluating Aducanumab in Patients With Early Alzheimer Disease. *JAMA Neurol* 2022; **79**(1): 13-21.
17. Mattsson-Carlgrén N, Andersson E, Janelidze S, Ossenkoppele R, Insel P, Strandberg O *et al.* Abeta deposition is associated with increases in soluble and phosphorylated tau that precede a positive Tau PET in Alzheimer's disease. *Sci Adv* 2020; **6**(16): eaaz2387.
18. Bejanin A, Schonhaut DR, La Joie R, Kramer JH, Baker SL, Sosa N *et al.* Tau pathology and neurodegeneration contribute to cognitive impairment in Alzheimer's disease. *Brain* 2017; **140**(12): 3286-3300.
19. Vanderlinden G, Ceccarini J, Vande Casteele T, Michiels L, Lemmens R, Triau E *et al.* Spatial decrease of synaptic density in amnesic mild cognitive impairment follows the tau build-up pattern. *Mol Psychiatry* 2022.
20. Mudher A, Colin M, Dujardin S, Medina M, Dewachter I, Alavi Naini SM *et al.* What is the evidence that tau pathology spreads through prion-like propagation? *Acta Neuropathol Commun* 2017; **5**(1): 99.
21. Crary JF, Trojanowski JQ, Schneider JA, Abisambra JF, Abner EL, Alafuzoff I *et al.* Primary age-related tauopathy (PART): a common pathology associated with human aging. *Acta Neuropathol* 2014; **128**(6): 755-766.
22. Tanner JA, Rabinovici GD. Relationship Between Tau and Cognition in the Evolution of Alzheimer's Disease: New Insights from Tau PET. *J Nucl Med* 2021; **62**(5): 612-613.
23. Pooler AM, Polydoro M, Maury EA, Nicholls SB, Reddy SM, Wegmann S *et al.* Amyloid accelerates tau propagation and toxicity in a model of early Alzheimer's disease. *Acta Neuropathol Commun* 2015; **3**: 14.
24. Clayton KA. Amyloid plaque deposition accelerates tau propagation via activation of microglia in a humanized APP mouse model. Boston University 2021.
25. Small SA, Duff K. Linking Abeta and tau in late-onset Alzheimer's disease: a dual pathway hypothesis. *Neuron* 2008; **60**(4): 534-542.
26. Mullins RJ, Diehl TC, Chia CW, Kapogiannis D. Insulin Resistance as a Link between Amyloid-Beta and Tau Pathologies in Alzheimer's Disease. *Front Aging Neurosci* 2017; **9**: 118.

27. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J *et al.* Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 2011; **134**(Pt 9): 2456-2477.
28. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF *et al.* Classification of primary progressive aphasia and its variants. *Neurology* 2011; **76**(11): 1006-1014.
29. Mackenzie IR, Neumann M, Bigio EH, Cairns NJ, Alafuzoff I, Kril J *et al.* Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: an update. *Acta Neuropathol* 2010; **119**(1): 1-4.
30. Mackenzie IR, Neumann M. Molecular neuropathology of frontotemporal dementia: insights into disease mechanisms from postmortem studies. *J Neurochem* 2016; **138 Suppl 1**: 54-70.
31. Olney NT, Spina S, Miller BL. Frontotemporal Dementia. *Neurol Clin* 2017; **35**(2): 339-374.
32. Knight A, Morrone C, Varlow C, Yu H, McQuade P, Vasdev N. In vitro evaluation of Tau-PET Radioligands for imaging TDP-43 in ALS. Soc Nuclear Med2022.
33. Dubois B, Villain N, Frisoni GB, Rabinovici GD, Sabbagh M, Cappa S *et al.* Clinical diagnosis of Alzheimer's disease: recommendations of the International Working Group. *Lancet Neurol* 2021; **20**(6): 484-496.
34. Vijverberg EG, Wattjes MP, Dols A, Krudop WA, Moller C, Peters A *et al.* Diagnostic Accuracy of MRI and Additional [18F]FDG-PET for Behavioral Variant Frontotemporal Dementia in Patients with Late Onset Behavioral Changes. *J Alzheimers Dis* 2016; **53**(4): 1287-1297.
35. Bergeron D, Beaugregard JM, Guimond J, Fortin MP, Houde M, Poulin S *et al.* Clinical Impact of a Second FDG-PET in Atypical/Unclear Dementia Syndromes. *J Alzheimers Dis* 2016; **49**(3): 695-705.
36. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review* 1958; **65**(6): 386.
37. Gavrishchaka V, Senyukova O, Koepke M. Synergy of physics-based reasoning and machine learning in biomedical applications: Towards unlimited deep learning with limited data. *Advances in Physics: X* 2019; **4**: 1582361.
38. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W *et al.* Backpropagation applied to handwritten zip code recognition. *Neural computation* 1989; **1**(4): 541-551.
39. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 1962; **160**: 106-154.

40. Levick WR. Receptive fields and trigger features of ganglion cells in the visual streak of the rabbits retina. *J Physiol* 1967; **188**(3): 285-307.
41. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**(7553): 436-444.
42. Traore BB, Kamsu-Foguem B, Tangara F. Deep convolution neural network for image recognition. *Ecological Informatics* 2018; **48**: 257-268.
43. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:14091556* 2014.
44. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S *et al.* Imagenet large scale visual recognition challenge. *International journal of computer vision* 2015; **115**(3): 211-252.
45. Etminani K, Soliman A, Davidsson A, Chang JR, Martinez-Sanchis B, Byttner S *et al.* A 3D deep learning model to predict the diagnosis of dementia with Lewy bodies, Alzheimer's disease, and mild cognitive impairment using brain 18F-FDG PET. *Eur J Nucl Med Mol Imaging* 2022; **49**(2): 563-584.
46. Diseases GBD, Injuries C. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020; **396**(10258): 1204-1222.
47. Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K *et al.* Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol* 2014; **13**(6): 614-629.
48. Brucher N, Mandegaran R, Filleron T, Wagner T. Measurement of inter- and intra-observer variability in the routine clinical interpretation of brain 18-FDG PET-CT. *Ann Nucl Med* 2015; **29**(3): 233-239.
49. Hu J, Qing Z, Liu R, Zhang X, Lv P, Wang M *et al.* Deep Learning-Based Classification and Voxel-Based Visualization of Frontotemporal Dementia and Alzheimer's Disease. *Front Neurosci* 2020; **14**: 626154.
50. Nemoto K, Sakaguchi H, Kasai W, Hotta M, Kamei R, Noguchi T *et al.* Differentiating Dementia with Lewy Bodies and Alzheimer's Disease by Deep Learning to Structural MRI. *J Neuroimaging* 2021; **31**(3): 579-587.
51. Jo T, Nho K, Saykin AJ. Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data. *Front Aging Neurosci* 2019; **11**: 220.
52. Jack CR, Jr., Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW *et al.* Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol* 2010; **9**(1): 119-128.

53. Reiman EM, Chen K, Alexander GE, Caselli RJ, Bandy D, Osborne D *et al.* Functional brain abnormalities in young adults at genetic risk for late-onset Alzheimer's dementia. *Proc Natl Acad Sci U S A* 2004; **101**(1): 284-289.
54. Perini G, Rodriguez-Vieitez E, Kadir A, Sala A, Savitcheva I, Nordberg A. Clinical impact of (18)F-FDG-PET among memory clinic patients with uncertain diagnosis. *Eur J Nucl Med Mol Imaging* 2021; **48**(2): 612-622.
55. Bejanin A, Tammewar G, Marx G, Cobigo Y, Iaccarino L, Kornak J *et al.* Longitudinal structural and metabolic changes in frontotemporal dementia. *Neurology* 2020; **95**(2): e140-e154.
56. Li X, Morgan P, Ashburner J, Smith J, Rorden C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods* 2016; **264**.
57. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC *et al.* User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 2006; **31**(3): 1116-1128.
58. Fleet D, Pajdla T, Schiele B, Tuytelaars T, (eds). Visualizing and Understanding Convolutional Networks. *Proceedings of the Computer Vision – ECCV 2014; 2014// 2014*; Cham. Springer International Publishing.
59. Guedj E, Varrone A, Boellaard R, Albert NL, Barthel H, van Berckel B *et al.* EANM procedure guidelines for brain PET imaging using [(18)F]FDG, version 3. *Eur J Nucl Med Mol Imaging* 2022; **49**(2): 632-651.
60. McKeith IG, Boeve BF, Dickson DW, Halliday G, Taylor JP, Weintraub D *et al.* Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology* 2017; **89**(1): 88-100.
61. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H *et al.* International evaluation of an AI system for breast cancer screening. *Nature* 2020; **577**(7788): 89-94.
62. Hao S, Liu C, Li N, Wu Y, Li D, Gao Q *et al.* Clinical evaluation of AI-assisted screening for diabetic retinopathy in rural areas of midwest China. *PLoS One* 2022; **17**(10): e0275983.
63. Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, Kuhl DE. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol* 1997; **42**(1): 85-94.
64. Scheltens NME, van der Weijden K, Adriaanse SM, van Assema D, Oomen PP, Krudop WA *et al.* Hypometabolism of the posterior cingulate cortex is not restricted to Alzheimer's disease. *Neuroimage Clin* 2018; **19**: 625-632.

65. Castelnovo V, Caminiti SP, Riva N, Magnani G, Silani V, Perani D. Heterogeneous brain FDG-PET metabolic patterns in patients with C9orf72 mutation. *Neurol Sci* 2019; **40**(3): 515-521.
66. Diehl-Schmid J, Licata A, Goldhardt O, Forstl H, Yakushew I, Otto M *et al.* FDG-PET underscores the key role of the thalamus in frontotemporal lobar degeneration caused by C9ORF72 mutations. *Transl Psychiatry* 2019; **9**(1): 54.
67. Townsend L, Pillinger T, Selvaggi P, Veronese M, Turkheimer F, Howes O. Brain glucose metabolism in schizophrenia: a systematic review and meta-analysis of (18)FDG-PET studies in schizophrenia. *Psychol Med* 2022: 1-18.
68. Baeken C, Wu GR, De Raedt R. Dorsomedial frontal cortical metabolic differences of comorbid generalized anxiety disorder in refractory major depression: A [(18)F] FDG PET brain imaging study. *J Affect Disord* 2018; **227**: 550-553.
69. Kang EH, Park JE, Lee KH, Cho YS, Kim JJ, Yu BH. Regional brain metabolism and treatment response in panic disorder patients: an [18F]FDG-PET study. *Neuropsychobiology* 2012; **66**(2): 106-111.
70. Devenney E, Bartley L, Hoon C, O'Callaghan C, Kumfor F, Hornberger M *et al.* Progression in Behavioral Variant Frontotemporal Dementia: A Longitudinal Study. *JAMA Neurol* 2015; **72**(12): 1501-1509.
71. Forster S, Buschert VC, Buchholz HG, Teipel SJ, Friese U, Zach C *et al.* Effects of a 6-month cognitive intervention program on brain metabolism in amnesic mild cognitive impairment and mild Alzheimer's disease. *J Alzheimers Dis* 2011; **25**(4): 695-706.
72. Trombert\* V, Huber P, Michel J-P. Autopsie ou la mort au service de la vie. *InfoKara* 2005; **53**(3): 77-80.
73. Colorectal Cancer Classification using Deep Convolutional Networks. An Experimental Study. *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOIMAGING 2018)*2018. SciTePress.
74. Barthel H, Villemagne VL, Drzezga A. Future Directions in Molecular Imaging of Neurodegenerative Disorders. *J Nucl Med* 2022; **63**(Suppl 1): 68S-74S.
75. Weber WA, Czernin J, Anderson CJ, Badawi RD, Barthel H, Bengel F *et al.* The Future of Nuclear Medicine, Molecular Imaging, and Theranostics. *J Nucl Med* 2020; **61**(Suppl 2): 263S-272S.
76. Delva A, Michiels L, Koole M, Van Laere K, Vandenberghe W. Synaptic Damage and Its Clinical Correlates in People With Early Huntington Disease: A PET Study. *Neurology* 2022; **98**(1): e83-e94.



77. Mecca AP, Chen MK, O'Dell RS, Naganawa M, Toyonaga T, Godek TA *et al.* In vivo measurement of widespread synaptic loss in Alzheimer's disease with SV2A PET. *Alzheimers Dement* 2020; **16**(7): 974-982.
78. Wang Y, Li E, Cherry SR, Wang G. Total-Body PET Kinetic Modeling and Potential Opportunities Using Deep Learning. *PET Clin* 2021; **16**(4): 613-625.
79. Wang G, Nardo L, Parikh M, Abdelhafez YG, Li E, Spencer BA *et al.* Total-Body PET Multiparametric Imaging of Cancer Using a Voxelwise Strategy of Compartmental Modeling. *J Nucl Med* 2022; **63**(8): 1274-1281.
80. Blanc-Durand P, Jegou S, Kanoun S, Berriolo-Riedinger A, Bodet-Milin C, Kraeber-Bodere F *et al.* Fully automatic segmentation of diffuse large B cell lymphoma lesions on 3D FDG-PET/CT for total metabolic tumour volume prediction using a convolutional neural network. *Eur J Nucl Med Mol Imaging* 2021; **48**(5): 1362-1370.
81. Reader AJ, Schramm G. Artificial Intelligence for PET Image Reconstruction. *J Nucl Med* 2021; **62**(10): 1330-1333.
82. Dawoodbhoy FM, Delaney J, Cecula P, Yu J, Peacock I, Tan J *et al.* AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units. *Heliyon* 2021; **7**(5): e06993.
83. Overman D. Oxipit Awarded CE Mark for Autonomous AI Medical Imaging Application. *AXIS Imaging News* 2022.

## **SUPPLEMENTARY INFORMATION**

FIGURE 1

**Flow chart of data.**

FIGURE 2

**French rainbow color scale used for visual interpretation.**

TABLE 1

**Hyperparameters search with Bayesian optimization.**

TABLE 2

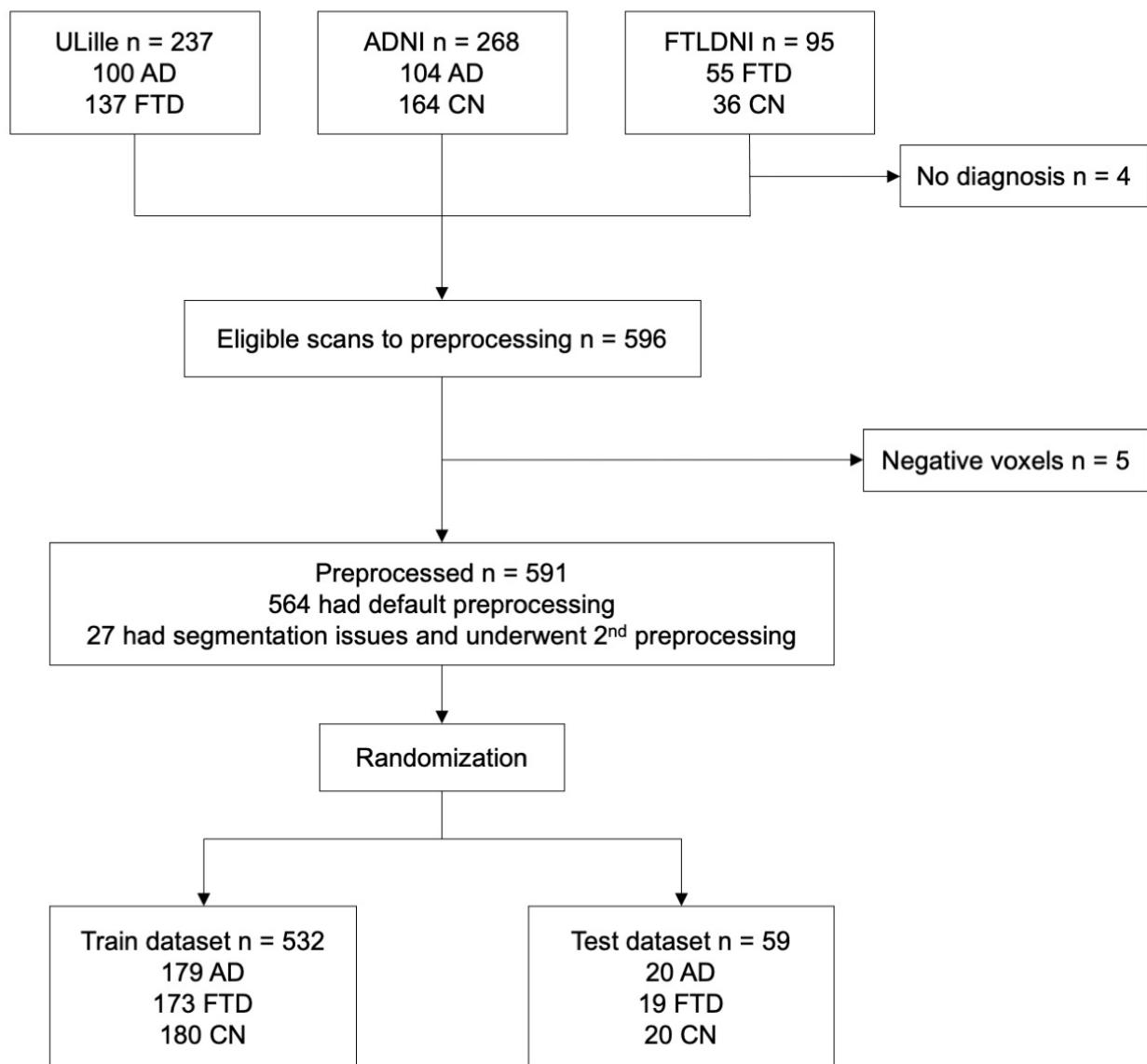
**Detailed labeling by each physician.**

TABLE 3

**Detailed metrics for each physician.**

TABLE 3

**Demographics of dementia with Lewy bodies and anxiety disorder subjects.**



**Supplementary figure 1: Flow chart of data.** AD, Alzheimer's disease; ADNI, Alzheimer's disease neuroimaging initiative; CN, cognitively normal; FTD, frontotemporal dementia; FTLDNI, frontotemporal lobar degeneration neuroimaging initiative; ULille, University of Lille.



**Supplementary figure 2: French rainbow color scale used for visual interpretation.**

**Supplementary table 1: Hyperparameters search with Bayesian optimization.** Only the 20 best-performing models are shown.

mean_fit_time	std_fit_time	mean_score_time	std_score_time	batch_size	dropout	initializer	learning_rate	optimizer	split0_score	split1_score	split2_score	split3_score	split4_score	mean_score	std_score
164.637	22.386	0.5815	0.0073	8	0.6	glorot_uniform	0.0006	adagrad	0.8491	0.8774	0.8302	0.7736	0.7905	0.8241	0.0379
191.512	31.320	0.5731	0.0046	10	0.5	glorot_normal	0.0003	adagrad	0.7736	0.8962	0.8019	0.7736	0.8667	0.8224	0.0502
166.454	6.715	0.5708	0.0055	10	0.5	glorot_normal	0.0006	adagrad	0.8019	0.8679	0.8208	0.8396	0.7810	0.8222	0.0300
141.529	6.620	0.6003	0.0059	6	0.5	glorot_normal	0.006	adagrad	0.8491	0.8962	0.7358	0.8208	0.7905	0.8185	0.0540
175.964	28.099	0.5762	0.0070	10	0.5	glorot_normal	0.003	adagrad	0.8491	0.9057	0.7547	0.8019	0.7810	0.8185	0.0535
174.273	14.316	0.5727	0.0018	10	0.5	glorot_normal	0.0006	adagrad	0.8774	0.8491	0.7736	0.8113	0.7810	0.8185	0.0397
190.927	29.448	0.5677	0.0126	10	0.5	glorot_normal	0.0003	adagrad	0.7830	0.8396	0.9057	0.7264	0.8286	0.8167	0.0598
192.951	37.382	0.5806	0.0097	8	0.6	glorot_normal	0.0006	adagrad	0.8302	0.7642	0.8491	0.7642	0.8667	0.8148	0.0430
181.685	17.503	0.5725	0.0080	10	0.6	glorot_uniform	0.0006	adagrad	0.7642	0.8302	0.7925	0.8585	0.8286	0.8148	0.0329
184.557	28.476	0.5777	0.0063	10	0.5	glorot_normal	0.0003	adagrad	0.8019	0.8113	0.8585	0.8113	0.7905	0.8147	0.0232
168.728	20.628	0.5674	0.0099	10	0.5	glorot_normal	0.0006	adagrad	0.8302	0.7925	0.7736	0.8208	0.8476	0.8129	0.0266
211.120	30.585	0.5689	0.0090	10	0.6	glorot_uniform	0.0003	adagrad	0.8491	0.8208	0.7736	0.8396	0.7810	0.8128	0.0305
181.984	21.508	0.5677	0.0127	10	0.5	glorot_normal	0.0003	adagrad	0.7358	0.8019	0.7642	0.8679	0.8857	0.8111	0.0579
164.019	6.774	0.5664	0.0033	10	0.5	glorot_normal	0.0003	adagrad	0.7925	0.8113	0.8113	0.8208	0.8190	0.8110	0.0100
141.302	26.013	0.6003	0.0055	6	0.6	he_normal	0.0006	adagrad	0.7358	0.8302	0.8585	0.7453	0.8762	0.8092	0.0580
178.887	31.089	0.5842	0.0064	8	0.5	glorot_normal	0.0003	adagrad	0.8113	0.7170	0.8396	0.8208	0.8571	0.8092	0.0487
210.884	49.196	0.5662	0.0069	10	0.5	glorot_normal	0.0003	adagrad	0.8019	0.7830	0.8208	0.8302	0.8000	0.8072	0.0166
216.587	39.726	0.5863	0.0116	8	0.6	glorot_uniform	0.0003	adagrad	0.8113	0.7736	0.7830	0.7736	0.8857	0.8054	0.0425
183.847	29.367	0.5907	0.0143	8	0.6	glorot_uniform	0.006	adagrad	0.7547	0.8208	0.8491	0.7642	0.8381	0.8054	0.0387
194.857	18.087	0.5730	0.0193	10	0.6	glorot_normal	0.003	adagrad	0.8302	0.8113	0.7736	0.8113	0.8000	0.8053	0.0186

**Supplementary table 2: Detailed labeling by each physician.** AD, Alzheimer's disease; CN, cognitively normal; FTD, frontotemporal dementia.

		<b>Actual labels</b>				
		AD	FTD	CN	Total	
<b>Physician 1 labels</b>	AD	15	3	1	19	
	FTD	2	7	0	9	
	CN	3	9	19	31	
	Total	20	19	20	59	
<b>Physician 2 labels</b>	AD	14	2	2	18	
	FTD	2	9	0	11	
	CN	4	8	18	30	
	Total	20	19	20	59	
<b>Physician 3 labels</b>	AD	13	2	2	17	
	FTD	3	12	4	19	
	CN	4	5	14	23	
	Total	20	19	20	59	

**Supplementary table 3: Detailed metrics for each physician.** AD, Alzheimer's disease; CN, cognitively normal; FTD, frontotemporal dementia; Se, sensitivity; Sp, specificity.

		<b>Metrics</b>					
		Se	Sp	Precision	F1-score	Support	
<b>Physician 1</b>	AD		75%	90%	79%	77%	20
	FTD		37%	95%	78%	50%	19
	CN		95%	69%	61%	75%	20
		Se	Sp	Precision	F1-score	Support	
<b>Physician 2</b>	AD		70%	90%	78%	74%	20
	FTD		47%	95%	82%	60%	19
	CN		90%	69%	60%	72%	20
		Se	Sp	Precision	F1-score	Support	
<b>Physician 3</b>	AD		65%	90%	76%	70%	20
	FTD		63%	82.50%	63%	63%	19
	CN		70%	77%	61%	65%	20

**Supplementary table 4: Demographics of dementia with Lewy bodies and anxiety disorder subjects.** \*One DLB subjects and 3 anxiety disorder subjects did not have a MMSE score within a year. DLB, dementia with Lewy bodies; MMSE, mini-mental state examination; SD, standard deviation.

	<b>DLB</b>		<b>Anxiety disorder</b>
No. subjects	20		16
No. scans	20		16
Diagnosis	Probable 95%	Certain 5%	
Source			
ULille	100%		100%
Clinical data			
Age [mean (SD)]	73.5 (8.5)		62.7 (8.8)
Female Sex	25%		75%
MMSE* [mean (SD)]	21.8 (4.5)		26.6 (1.9)





**AUTEUR : Nom : ROGEAU**

**Prénom : Antoine**

**Date de soutenance : 2 mars 2023**

**Titre de la thèse :**

Elaboration d'un réseau de neurones convolutifs 3D pour classer les TEP-FDG entre maladie d'Alzheimer, démence frontotemporale et cognitivement normal

**Thèse - Médecine - Lille 2023**

**Cadre de classement : Médecine nucléaire**

**DES + FST/option : Médecine nucléaire**

**Mots-clés : maladie d'Alzheimer, démence frontotemporale, IA, réseau de neurones**

**Résumé :**

**Introduction :** Avec l'arrivée de traitements influençant l'évolution des maladies neurodégénératives, il sera essentiel de diagnostiquer avec précision celles-ci. Les réseaux de neurones convolutifs sont de puissantes techniques d'apprentissage profond pouvant aider les médecins imageurs dans l'interprétation. La classification par apprentissage profond de sujets avec une maladie d'Alzheimer (MA), une démence frontotemporale (DFT) ou cognitivement sains (CS), basée sur le métabolisme cérébral du glucose, n'a jamais été présentée dans la littérature et ainsi mérite considération.

**Méthodes :** Les TEP au [18F]-FDG de 199 MA, 192 DFT et 200 CS sujets ont été collectées rétrospectivement de notre base de données locale et des bases de données *Alzheimer's disease* et *frontotemporal lobar degeneration neuroimaging initiative*. L'entraînement d'un réseau 3D inspiré de l'architecture VGG16 a été effectué sur 90% des données avec augmentation de données et validation croisée. Les performances ont été comparées à l'interprétation clinique sur les 10% de données restantes. Les régions guidant la classification ont été identifiées dans une expérience d'occlusion.

**Résultats :** Les sujets des données tests étaient appariés en âge et en sexe. Le modèle montrait une précision de 89,8% à prédire les classes des données tests. Les aires sous la courbe ROC étaient de 93,3% pour la MA, 95,3% pour la DFT et 99,9% pour les CS. Le consensus atteint par les médecins avait une précision de 69.5%, et la concordance était importante entre eux ( $\kappa = 0.61$ , 95% CI : 0.49-0.73). Les cortex cingulaire antérieur, temporal mésial et préfrontal dorsolatéral étaient des régions communes à la classification en MA ou en DFT.

**Discussion :** A notre connaissance, il s'agit de la première étude à présenter un réseau de neurones capable de discriminer MA et DFT sur la base de la TEP au [18F]-FDG, et d'isoler les sujets CS avec haute précision. Cela ajoute à la littérature grandissante de l'apprentissage profond en médecine nucléaire et suggère l'incorporation future de ces techniques en pratique clinique.

**Composition du Jury :**

**Président : Professeur Damien HUGLO**

**Assesseurs : Professeur Franck SEMAH, Professeur Thibaud LEBOUVIER & Docteur Renaud LOPES**

**Directeur de thèse : Docteur Florent HIVES**