



HAL
open science

Nouveaux biomarqueurs d'imagerie pour la prise en charge des lymphomes primitifs du système nerveux central : études en TEP-IRM au 18F-FDG et à la 18F-FLUDARABINE

Laura Rozenblum

► To cite this version:

Laura Rozenblum. Nouveaux biomarqueurs d'imagerie pour la prise en charge des lymphomes primitifs du système nerveux central : études en TEP-IRM au 18F-FDG et à la 18F-FLUDARABINE. Neurobiologie. Sorbonne Université, 2024. Français. ⟨NNT : 2024SORUS150⟩. ⟨tel-05015130⟩

HAL Id: tel-05015130

<https://theses.hal.science/tel-05015130v1>

Submitted on 1 Apr 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Sorbonne Université

Ecole doctorale 158

Laboratoire d'Imagerie Biomédicale / Equipe 4 – Connectivité Neurale et Plasticité

Evaluation de nouveaux biomarqueurs d'imagerie TEP-IRM au 18F-FDG et à la 18F-Fludarabine dans les lymphomes primitifs du système nerveux central

*New Biomarkers For Diagnosis And Prognosis Of Primary Cerebral Lymphoma
A PET-MRI Study Using 18F-FDG And 18F-Fludarabine*

Par Laura ROZENBLUM

Thèse de doctorat de Sciences

Présentée et soutenue publiquement le 03 Juillet 2024

Devant un jury composé de :

Mme Laure Fournier	Professeur des Universités - Praticien Hospitalier	Université Paris Cité	Rapporteur
Mme Laetitia Vercellino	Professeur des Universités - Praticien Hospitalier	Université Paris Cité	Rapporteur
Mr Georges El Fakhri	Professor	Yale University	Examineur
Mr Eric Guedj	Professeur des Universités - Praticien Hospitalier	Aix-Marseille Université	Examineur
Mr Khê Hoang-Xuan	Professeur des Universités - Praticien Hospitalier	Sorbonne Université	Président du jury
Mme Aurélie Kas	Professeur des Universités - Praticien Hospitalier	Sorbonne Université	Directrice de thèse

Résumé en français

Le lymphome primitif du système nerveux central (LPSNC) est une pathologie au pronostic sombre, dont l'incidence est en hausse, notamment chez les patients âgés et immunocompétents. La démarche thérapeutique pour les patients atteints de LPSNC s'appuie sur l'évaluation de deux indicateurs pronostiques établis : l'âge et le performans status. Néanmoins, la nécessité d'améliorer les stratégies thérapeutiques actuelles, centrées sur une poly-chimiothérapie d'induction à base de méthotrexate à haute dose suivie potentiellement par un traitement d'entretien, a suscité un intérêt croissant dans l'identification de nouveaux biomarqueurs. Cette thèse se positionne dans ce contexte de recherche et d'innovation dans les LPSNC, promue par le réseau LOC (lymphome oculo-cérébral) en France, et visant à améliorer la stratification pronostique et le suivi thérapeutique chez ces patients.

L'imagerie par résonance magnétique (IRM) cérébrale constitue aujourd'hui le standard de vérité dans le suivi des LPSNC. Cependant, plusieurs études récentes ont mis en évidence ses limites, en particulier devant le pourcentage élevé de patients récidivants rapidement malgré une réponse jugée complète sur l'évaluation de fin de traitement. Notre recherche explore le potentiel de la tomographie par émission de positon (TEP) au 18F-Fluorodeoxyglucose (FDG) combinée à l'IRM comme outil innovant dans la prise en charge des patients atteints de LPSNC.

Cette recherche s'est articulée autour de trois axes principaux: d'une part la recherche de biomarqueurs TEP-IRM pronostiques de la réponse de fin de traitement d'induction, d'autre part l'évaluation de cette technique comme outil d'évaluation précoce de la réponse sous chimiothérapie et enfin l'exploitation de méthodes d'intelligence artificielle pour affiner la compréhension physiopathologique via des analyses radiomiques poussées, et élaborer un modèle prédictif en apprentissage profond, généralisable à large échelle.

Ce travail s'appuie sur les données prospectives longitudinales de la cohorte Localyse, une étude ancillaire de l'essai multicentrique de phase III BLOCAGE-01, qui étudie l'intérêt d'une chimiothérapie de maintenance chez les patients âgés, immunocompétents, atteints de LPSNC. Les résultats de ce travail doctoral ont permis l'identification de deux nouveaux biomarqueurs pronostiques en TEP FDG dans les LPSNC et ont démontré la supériorité potentielle de la TEP interim pour évaluer la réponse thérapeutique après deux cycles de chimiothérapie en comparaison avec l'IRM interim. Intégrant les avancées les plus récentes en matière d'analyse radiomique et d'apprentissage profond, cette thèse s'inscrit dans le développement d'outils robustes et novateurs en IA appliquée à la neuro-oncologie. Enfin, elle examine le potentiel de la 18F-Fludarabine, un radiotracer ciblant spécifiquement les lymphocytes B, qui pourrait offrir de nouvelles perspectives dans la prise en charge des patients atteints de LPSNC dans le futur.

En conclusion, nos travaux contribuent au développement de nouveaux outils facilitant la gestion des patients atteints de LPSNC et ouvrent la voie au développement de stratégies thérapeutiques guidées par la TEP.

Mots-clés

Lymphome primitif du système nerveux central, TEP, IRM, radiomique, apprentissage automatique, apprentissage profond

Abstract

Primary central nervous system lymphoma (PCNSL) is a malignancy with a dismal prognosis, increasingly prevalent, especially among elderly and immunocompetent patients. The therapeutic approach for PCNSL relies on evaluating established prognostic indicators: patient age and performance status. However, there is a growing need to enhance current therapeutic strategies, which are centered on high-dose methotrexate-based induction polychemotherapy, potentially followed by maintenance treatment. This has encouraged interest in identifying new biomarkers. This thesis positions itself within this research and innovation framework in PCNSL, supported by the LOC network (Lymphomes Oculo-Cerebraux) in France, aiming to improve prognostic stratification and therapeutic monitoring for these patients.

Currently, cerebral magnetic resonance imaging (MRI) is the gold standard for PCNSL follow-up. However, recent studies have highlighted its limitations, notably the high percentage of patients relapsing rapidly despite a complete response on end-of-treatment evaluation. Our research investigates the potential of positron emission tomography (PET) with ¹⁸F-fluorodeoxyglucose (FDG) combined with MRI as an innovative tool in the management of PCNSL patients.

The study is structured around three main axes: firstly, exploring PET-MRI FDG as a prognostic biomarker for end-of-induction treatment response; secondly, its utility as a tool for early evaluation of the chemotherapy response; and finally, exploiting artificial intelligence methods to refine the pathophysiological understanding through advanced radiomic analyses, and to develop a predictive deep-learning model scalable for broad application.

This work relied on the longitudinal prospective data from the Localyse cohort, an ancillary study of the phase III multicentric BLOCAGE-01 trial, investigating the utility of maintenance chemotherapy in aged, immunocompetent patients. The results of this doctoral research identified two novel prognostic biomarkers in FDG-PET for PCNSL and demonstrated the potential superiority of interim PET over interim MRI in assessing the therapeutic response post two chemotherapy cycles. Incorporating the latest advances in radiomic analysis and deep learning, this thesis contributes to the development of robust and innovative tools in AI applied to neuro-oncology. Lastly, it examines the potential of ¹⁸F-Fludarabine, a radiotracer specifically targeting B lymphocytes, which could offer new perspectives in the future management of PCNSL patients.

In conclusion, our work facilitates the development of new tools for managing PCNSL patients, paving the way for treatment strategies guided by PET findings.

Keywords

Primary central nervous system lymphoma, PET, MRI, radiomics, machine-learning, deep-learning

Table des matières

Remerciements	Erreur ! Signet non défini.
Résumé en français	2
Abstract	3
Table des matières	4
Table des figures	6
Liste des tableaux	7
Liste des annexes	7
Glossaire	8
Introduction générale	10
Liste des publications et communications	13
Chapitre 1. Introduction sur le lymphome primitif du système nerveux central	14
I. Epidémiologie	14
II. Diagnostic	15
Diagnostic clinique.....	15
Diagnostic radiologique	15
Confirmation histologique.....	22
Bilan d’extension.....	23
Chapitre 2. Le rôle de la TEP	24
I. Principes physiques	24
II. 18F-Fluorodeoxyglucose	26
III. Quantification	26
IV. Intérêt de la TEP-IRM	28
V. Article 1.	31
Chapitre 3. Biomarqueurs pronostiques du LPSNC	50
I. Données de survie	50
II. Biomarqueurs pronostiques	50
Cliniques.....	50
Biologiques.....	52
Imagerie.....	52
III. Article 2	56
Chapitre 4. Prise en charge et évaluation thérapeutique du LPSNC	73
I. Traitement	73
II. Evaluation thérapeutique	74
Critères IRM.....	74
Critères en TEP	76
III. Article 3	79
Chapitre 5. Analyse radiomique en TEP	96
I. Introduction générale à la radiomique	96
Définition de la question et estimation des biais potentiels	99
Pré-traitement des images	100
Sélection des paramètres radiomiques	109
Modélisation statistique.....	114
Evaluation des performances	116
Chapitre 6. Deep learning et TEP	118
I. Etat de l’art IA et imagerie médicale	118
II. Workflow en deep learning	119
Augmentation des données.....	119

Auto-encodeur	120
Segmenteur	124
SWIN transformer	125
Optuna	126
Gradcam	126
III. Article 4	128
Chapitre 7. TEP-IRM à la 18F-Fludarabine	144
Chapitre 8. Conclusion et perspectives	150
Bibliographie	151
Annexes	178

Table des figures

FIGURE 1.1 Principe des gradients de diffusion	17
FIGURE 1.2 Calcul de l'ADC	17
FIGURE 1.3 Courbe représentant le pourcentage de perte du signal en fonction du temps de passage du bolus de produit de contraste.....	19
FIGURE 1.4 Image typique de LPSNC en IRM multimodale.....	20
FIGURE 1.5 Coupe histologique d'un LPSNC.....	23
FIGURE 2.1 Devenir dans la matière d'un positon émis par le 18F-FDG	24
FIGURE 2.2 Types de LOR pouvant être enregistrées par un imageur TEP	25
FIGURE 3.1 Représentation du score MILAS	54
FIGURE 4.1 Score de Deauville et méthode de calcul du ΔSUV_{max}	77
FIGURE 5.1 Comparaison de la distribution des données, après réduction de dimension par ACP, sans harmonisation, avec une harmonisation par ComBat et avec une harmonisation par M-ComBat	102
FIGURE 5.2 Exemples de masques sur l'IRM injecté et la TEP FDG	103
FIGURE 5.3 Discrétisation en intensité selon la méthode relative ou absolue.....	105
FIGURE 5.4 Représentation schématique des paramètres radiomiques de premier ordre.....	107
FIGURE 5.5 Représentation schématique des paramètres radiomiques de second et troisième ordre	108
FIGURE 5.6 Exemple de sélection de paramètres radiomiques par méthode univariée.....	110
FIGURE 5.7 Exemples de mélange aléatoire des voxels au sein d'une VOI.....	111
FIGURE 5.8 Application de la méthode du mélange des voxels pour sélectionner les paramètres radiomiques robustes en TEP	112
FIGURE 5.9 Corrélogramme entre les radiomiques retenus comme robustes et les marqueurs cliniques et semi-quantitatifs	113
FIGURE 5.10 Exemple d'architecture d'un MLP.....	115
FIGURE 6.1 Schéma d'architecture d'un auto-encodeur.....	120
FIGURE 6.2 Représentation mathématique de la fonction ReLu	121
FIGURE 6.3 Exemple de courbes de loss	122
FIGURE 6.4 Exemple d'architecture d'un réseau de neurones convolutifs.....	123
FIGURE 6.5 Schéma d'architecture d'un segmenteur	124
FIGURE 6.6 Représentation simplifiée des fenêtres de vision dans un Swin transformer et dans un vision transformer (ViT)	126
FIGURE 6.7 Exemple d'application du Grad-CAM chez un patient atteint de LPSNC.....	127
FIGURE 7.1 Représentation de la molécule de 18F-Fludarabine	145
FIGURE 7.2 Images TEP de modèles murins de LPNSC et GBM en 18F-Fludarabine et 18F-FDG	146

Liste des tableaux

Tableau 3.1 Scores pronostiques IELSG et MSKCC	51
Tableau 3.2 Données de la littérature sur les biomarqueurs pronostiques en IRM	53
Tableau 4.1 Score IPCG [6]	75
Tableau 4.2 Données de la littérature sur l'intérêt de la TEP dans le suivi des patients atteints de LPSNC	77
Tableau 5. Données de la littérature sur l'intérêt des radiomiques dans le LPSNC	97

Liste des annexes

Annexe 1. Echelle de Karnofsky	167
Annexe 2. Résumé du protocole LOCALYSE	168
Annexe 3. Score METRICS (Kocak, 2024)	173
Annexe 4. Score Radiomics Quality Score (Lambin, 2017)	175
Annexe 5. Checklist CLAIM (Mongan, 2020)	177
Annexe 6. Checklist mi-CLAIM (Norgeot, 2020)	178
Annexe 7. Checklist FUTURE-AI (Lekadir, 2023)	179

Glossaire

ACC : Accuracy

ACP : Analyse en composante principale

ADC : Apparent diffusion coefficient

ADW : Advantage workstation

ASCT : Autologous stem cell transplantation

ASL : Arterial spin labelling

AUC : Area under the receiver operating characteristic curve

BAT : Bolus arrival time

BHE : Barrière hémato-encéphalique

BTIP : Brain Tumor Imaging Protocol

BTK : Bruton tyrosine kinase

CBTRUS : Central Brain Tumor Registry of The United States

CBF : Cerebral blood flow

CBV : Cerebral blood volume

Cho : Choline

CNN : Convolutional neural network

Cr : Creatine

DL : Deep learning

EANO : European association of neuro-oncology

FDG : Fluorodeoxyglucose

FLAIR : Fluid attenuated inversion recovery

GBM : Glioblastome

GLCM : Grey level co-occurrence matrix

GLRLM : Grey level run length matrix

GLSZM : Grey level size zone matrix

GPU : Graphics processing unit

IA : Intelligence artificielle

IBSI : Imaging Biomarker Standardization Initiative

ICANS : Immune effector cell-associated neurotoxicity syndrome

IL : Interleukin

INCa : Institut National du Cancer

IPCG : International primary central nervous system lymphoma collaborative group

IRM : Imagerie par résonance magnétique

KPS : Karnofsky Performance Status

LBDGC : Lymphome B diffus à grandes cellules

LCR : Liquide céphalo-rachidien

LOC : Lymphome Oculo-Cérébral

LOR : Line of response

LPSNC : Lymphome primitif du système nerveux central

MILAS : Metabolic imaging lymphoma aggressiveness scale

ML : Machine learning

MLP : Multilayer perceptron
MTT : Mean transit time
MTV : Metabolic tumor volume
MTX : Methotrexate
NAA : N-acetyl-aspartate
NGDLM : Neighboring grey level dependence matrix
NGTDM : Neighboring grey tone difference matrix
OSEM : Ordered subset expectation maximization
PERCIST : PET response criteria in solid tumors
PS : Performans status
ROC : Receiver operating characteristic
SG : Survie globale
SNC : Système nerveux central
SRM : Spectrométrie par résonance magnétique
SSP : Survie sans progression
SUV : Standardized uptake value
SVM : Support vector machine
TDM : Tomodensitométrie
TE : Temps d'écho
TEP : Tomographie par émission de positon
TLG : Total lesion glycolysis
TSE : Turbo spin echo
TTP : Time to peak
VOI : Volume of interest
VPN : Valeur prédictive négative
VPP : Valeur prédictive positive
WHO : World health organization

Introduction Générale

Le lymphome primitif du système nerveux central (LPSNC) est un lymphome non hodgkinien agressif, dont l'incidence est en augmentation depuis deux décennies, notamment chez les patients âgés et immunocompétents [1–3]. Malgré de nombreux essais thérapeutiques en cours, le pronostic de cette maladie reste sombre. Une étude récente sur 1000 patients avec LPSNC *de novo* a montré un taux de mortalité à 25% dans les six mois suivant le diagnostic chez les patients âgés [1]. A l'heure actuelle, les seuls facteurs pronostiques validés dans la littérature sont l'âge et le statut fonctionnel [4].

En France, l'ensemble des patients atteints d'un LPSNC est recensé par le réseau Lymphome Oculo-Cérébral. Ce réseau national s'articule autour d'un centre expert national sur les sites de la Pitié-Salpêtrière (Pr. Hoang-Xuan) et de Curie Saint-Cloud (Dr. Soussain) et de vingt centres experts régionaux. Sa mission est de promouvoir les essais thérapeutiques et la recherche dans les LPSNC. Aujourd'hui, que ce soit en clinique ou dans les essais thérapeutiques, l'examen de référence pour le diagnostic des LPSNC est l'imagerie par résonance magnétique (IRM) cérébrale [5]. Au cours du suivi, les critères de réponse, appelés critères IPCG car issus du groupe collaboratif international sur le LPSNC (International PCNSL Collaborative Group), reposent également en grande partie sur l'IRM cérébrale [6]. Cependant, plusieurs articles récents ont soulevé la question de la fiabilité de cet examen, en raison de la survenue de récurrences précoces chez des patients pourtant jugés en réponse complète sur l'IRM [7,8]. Van der Meulen et al. ont également montré sur une cohorte prospective de près de 200 patients qu'il n'y avait pas de différence significative en termes de survie sans progression (SSP) ou de survie globale (SG) entre les patients en réponse complète et ceux en réponse partielle sur l'IRM de fin de traitement [8].

En 2021, l'IPCG a intégré la tomographie par émission de positons (TEP) au 18-Fluorodeoxyglucose (FDG) dans les protocoles d'imagerie des LPSNC, notamment pour écarter une atteinte extra-cérébrale, critère excluant le diagnostic de LPSNC [5]. Bien que son application soit actuellement restreinte à la TEP corps-entier, la TEP-FDG cérébrale montre un potentiel en tant que marqueur pronostique et outil d'évaluation thérapeutique dans les LPSNC. Des études préliminaires suggèrent des paramètres TEP pronostiques et soulignent son utilité pour l'évaluation thérapeutique en milieu et en fin de traitement, suivant ainsi les pratiques établies pour les lymphomes systémiques [9–16].

Cependant, ces études antérieures, majoritairement rétrospectives, se basaient sur un nombre restreint de patients et montraient des résultats hétérogènes [9,13–15,17]. L'objectif principal de cette thèse était d'examiner, de façon prospective et dans un centre expert, l'intérêt de la TEP-FDG en tant que biomarqueur d'imagerie pour les LPSNC sur un large échantillon de patients. En parallèle, l'objectif secondaire visait à comparer les données de la TEP avec celles issues de l'IRM multimodale, incluant diffusion, perfusion et spectroscopie, pour dresser une représentation précise et complète des particularités morfo-métaboliques des LPSNC.

Pour ce faire, mon travail s'est appuyé sur les données du protocole LOCALYSE (PI : Pr A. KAS), étude soutenue par le réseau LOC et ancillaire de l'essai prospectif multicentrique BLOCAGE-01 (PI : Pr K. Hoang-Xuan) qui visait à évaluer chez les patients âgés souffrant d'un LPSNC, l'intérêt d'une stratégie de maintenance versus surveillance après réponse complète à une chimiothérapie de première ligne à base de méthotrexate à haute dose. Dans le cadre du protocole LOCALYSE, les patients ont bénéficié à la phase d'induction de la chimiothérapie identique pour tous, de 3 TEP-IRM cérébrale au FDG, acquises dans le service de médecine nucléaire de la Pitié-Salpêtrière, au diagnostic, à mi-traitement (2 cycles) et en fin de chimiothérapie d'induction (fin de traitement). L'objectif de ma thèse était de répondre à trois questions principales :

1. La TEP-IRM au FDG pré-thérapeutique est-elle prédictive de la réponse de fin de traitement et pronostique de la survie ?
2. L'évaluation précoce en TEP-IRM au FDG, permet-elle une stratification fiable des patients selon leur risque de progression ou de décès précoce ?
3. Des approches d'intelligence artificielle (*machine learning/deep learning*) permettent-elles de corroborer la pertinence de la TEP-IRM initiale comme biomarqueur et d'identifier des algorithmes reproductibles à large échelle pour classifier les patients selon l'agressivité tumorale au diagnostic ?

Le travail en lien avec cette dernière question s'est fait dans le cadre de ma mobilité à Boston, au sein du *Gordon Center for Medical Imaging*, sous la direction du Professeur G. El Fakhri et le co-encadrement de Thibault Marin.

Dans le premier chapitre de cette thèse, nous aborderons la pathologie du LPSNC, en portant une attention particulière aux performances diagnostiques de l'IRM et aux caractéristiques observables. Nous soulignerons l'importance clinique de l'identification de biomarqueurs

d'imagerie novateurs visant une prise en charge personnalisée et une surveillance optimisée pour les patients présentant un haut risque de récurrence ou de progression de la maladie.

Le second chapitre sera dédié à la TEP, incluant un bref exposé des principes physiques sous-jacents, suivi d'une revue des données actuelles de la littérature concernant cette technique d'imagerie dans le contexte des LPSNC. Le troisième chapitre traitera des biomarqueurs pronostiques à l'étude à ce jour, qu'ils soient cliniques, biologiques ou d'imagerie. Le quatrième chapitre présentera une synthèse de l'état actuel des traitements disponibles pour les patients âgés immunocompétents atteints de LPSNC. Les chapitres 5 et 6 exposeront les méthodologies liées à l'intelligence artificielle (IA), notamment les techniques de machine learning (ML) basées sur l'analyse radiomique et les méthodes de deep learning (DL) appliquées directement aux images de TEP-IRM. Au fil des chapitres, nous intégrerons les publications associées à chaque thématique. Les recherches en cours en lien avec l'intelligence artificielle, qui touchent à leur terme, seront exposées à travers une description approfondie des avancées méthodologiques et une présentation des résultats préliminaires obtenus.

En conclusion et en guise de perspectives, nous ouvrirons la discussion sur l'utilisation future de la ^{18}F -Fludarabine, un nouveau radiotraceur en cours d'exploration pour son avidité potentielle plus forte pour les lymphocytes B [18–24]. L'étude pilote Fludaloc, démarrée début 2024 et devant inclure huit patients, sera présentée et les enjeux et attentes liés à ce nouveau radiotraceur seront discutés.

Liste des publications et communications

Articles publiés dans des revues à comité de lecture

Rozenblum L, Houillier C, Soussain C, Bertaux M, Choquet S, Galanaud D, Hoang-Xuan K, Kas A. Role of Positron Emission Tomography in Primary Central Nervous System Lymphoma. *Cancers (Basel)*. 2022 Aug 23;14(17):4071

Rozenblum L, Galanaud D, Houillier C, Soussain C, Baptiste A, Belin L, Edeline V, Naggara P, Soret M, Causse-Lemercier V, Willems L, Choquet S, Ursu R, Hoang-Xuan K, Kas A. [18F]FDG PET-MRI provides survival biomarkers in primary central nervous system lymphoma in the elderly: an ancillary study from the BLOCAGE trial of the LOC network. *Eur J Nucl Med Mol Imaging*. 2023 Oct;50(12):3684-3696.

Rozenblum L, Houillier C, Baptiste A, Soussain C, Edeline V, Naggara P, Soret M, Causse-Lemercier V, Willems L, Choquet S, Ursu R, Galanaud D, Belin L, Hoang-Xuan K, Kas A. Interim FDG-PET improves treatment failure prediction in primary central nervous system Lymphoma: a LOC network prospective multicentric study. *Neuro Oncol*. 2024

Communications orales

Rozenblum L, Houillier C, Baptiste A, Soussain C, Edeline V, Naggara P, Soret M, Causse-Lemercier V, Willems L, Choquet S, Ursu R, Galanaud D, Belin L, Hoang-Xuan K, Kas A. Etude prospective multicentrique du réseau LOC : TEP interim dans le lymphome primitif du système nerveux central. *JFMN* 2024

Rozenblum L, Prognostic imaging biomarkers of Primary Central Nervous System Lymphoma in the new era of Artificial Intelligence, 2024, Symposium des Lauréats de l'Institut Servier, Paris, France

Communications affichées

Rozenblum L, Galanaud D, Houillier C, Soussain C, Baptiste A, Belin L, et al. [18F-FDG PET improves prediction of treatment failure in Primary Central Nervous System Lymphoma : results from a multicentric prospective study. *SNMMI Annual Meeting*, 2023

Rozenblum L, Houillier C, Baptiste A, Soussain C, Edeline V, Naggara P, Soret M, Causse-Lemercier V, Willems L, Choquet S, Ursu R, Galanaud D, Belin L, Hoang-Xuan K, Kas A. 18F-FDG PET-MRI provides new outcome biomarkers in primary central nervous system lymphoma: results of a French multicentric prospective study. *SNMMI Annual Meeting*, 2023

Chapitre 1. Introduction sur le lymphome primitif du système nerveux central

I. Epidémiologie

Le lymphome primitif du système nerveux central (SNC), pathologie rare du SNC, se caractérise par sa localisation exclusive au cerveau, à la moelle épinière, aux leptoméniges et aux yeux. La classification actualisée de 2021 des tumeurs cérébrales regroupe sous cette appellation le lymphome primitif B diffus à grandes cellules (LBDGC) du SNC, le lymphome du SNC lié à une immunodépression, la granulomatose lymphomatoïde, ainsi que le lymphome B intravasculaire [25]. Cette pathologie constitue une entité à part, distincte des lymphomes systémiques avec manifestation cérébrale.

Le LPSNC représente environ 4 % des tumeurs du SNC et entre 4 % et 6 % des lymphomes extra-ganglionnaire. Son incidence, estimée entre 0,3 et 0,6 pour 100 000 personnes, a connu une augmentation notable ces cinquante dernières années, en particulier parmi les patients immunocompétents et ceux âgés de plus de 60 ans [1,2]. En effet, l'incidence chez les septuagénaires et octogénaires a été relevée à 4,32 pour 100 000 personnes aux États-Unis selon le CBTRUS (Central Brain Tumor Registry of The United States) [26]. L'âge médian au diagnostic est aujourd'hui de 65 ans, avec un pic d'incidence chez les patients de plus de 75 ans. Une légère prédominance masculine a été également mise en évidence au diagnostic [27,28].

Inversement, depuis l'introduction des trithérapies antivirales dans les années 1990, l'incidence du LPSNC associée à l'immunodépression a considérablement diminué, passant de 5 cas pour 1 000 personnes entre 1991 et 1994 à 0,32 cas pour 1 000 personnes après 1999 [29]. Le LPSNC est aussi détectable chez les patients immunodéprimés à la suite d'une transplantation d'organe et est souvent lié à la présence d'une infection par le virus Epstein-Barr dans ces cas.

En France, les patients atteints de LPSNC sont recensés par le réseau LOC (Lymphome-Oculo-Cérébral), un réseau expert national, créé en 2011 et financé par l'Institut National du Cancer (INCa) (Prof. K. Hoang-Xuan, Dr. C. Soussain). L'objectif du réseau LOC est d'améliorer la prise en charge des patients atteints de LPSNC, de constituer des bases de données cliniques et de tissu histologique et de promouvoir de nouveaux essais thérapeutiques. Le réseau est constitué d'un centre expert en Ile de France, à l'Hôpital de la Pitié-Salpêtrière à Paris et à

l'Institut Curie à Saint Cloud, et de vingt centres experts régionaux répartis dans toute la France. Dans le cadre de cette thèse, la recherche se portera uniquement sur le LPSNC du patient âgé de plus de 60 ans et immunocompétent.

II. Diagnostic

Diagnostic clinique

Dans 92 % des cas, le LPSNC se manifeste dans le parenchyme cérébral, entraînant un déficit neurologique focal qui dépend de la localisation tumorale, et une altération des fonctions cognitives. Contrairement au lymphome systémique, les signes généraux tels que la fièvre et les sueurs nocturnes, sont rarement présents. Des manifestations neuropsychiatriques telles que la dépression, l'ataxie ou les hallucinations, peuvent parfois prédominer et ainsi retarder le diagnostic, particulièrement chez les patients âgés où une démence de type maladie d'Alzheimer peut être évoquée en première hypothèse [28]. Les symptômes d'hypertension intracrânienne, tels que les céphalées, les nausées ou les vomissements, sont également fréquents. En revanche, les crises d'épilepsie sont moins fréquentes que dans les gliomes [1,3,30].

En présence d'une atteinte leptoméningée, les nerfs crâniens VI et VII sont les plus fréquemment touchés. Les atteintes oculaires sont observées dans 25 % des cas, affectant principalement le segment postérieur de l'œil, en particulier le vitré et la rétine, souvent de manière bilatérale. Bien que ces atteintes soient généralement asymptomatiques, elles peuvent conduire à une diminution de l'acuité visuelle, à une vision floue ou à la perception de corps flottants [31]. L'atteinte de la moelle épinière reste exceptionnelle (moins de 1 % des cas) et peut se manifester sous la forme d'une myélopathie subaiguë.

Diagnostic radiologique

Lorsqu'une pathologie cérébrale est suspectée, l'imagerie par résonance magnétique (IRM) cérébrale constitue l'examen de première intention. Cet examen clé offre une caractérisation précise de la lésion et définit son étendue. Les séquences IRM pertinentes au diagnostic sont celles recommandées par les standards internationaux, en particulier ceux publiés par Barajas et al. en 2021 [5]. L'adhésion à ces directives facilite la standardisation de la procédure diagnostique et est essentielle pour le suivi thérapeutique. Le protocole IRM optimal s'effectue sur un appareil 3 Tesla, incluant une séquence T1 pré et post-injection de 0.1 mmol/kg de

gadolinium, ainsi que des séquences TSE T2W (Turbo Spin Echo T2 weighted) et T2 FLAIR (Fluid attenuated inversion recovery). Classiquement, le LPSNC se manifeste par une lésion en hyposignal T1, iso ou hyposignal T2, et prend fortement le contraste après injection de gadolinium, en raison de la rupture de la barrière hémato-encéphalique. Les lésions apparaissent souvent sous forme d'une masse unique (60 à 80 % des cas), de localisation préférentielle supratentorielle (80 %), et péri-ventriculaire profonde [32,33]. Une localisation dans le lobe frontal est rapportée dans 20 à 43 % des cas et les ganglions de la base sont atteints dans 13 à 20 % des cas [34]. Un autre signe distinctif est la prise de contraste linéaire périvasculaire, indicative de l'invasion angiotropique où les cellules tumorales s'alignent de manière centripète autour des espaces de Virchow-Robin, formant un motif d'infiltration en manchon. Les cellules tumorales du LPSNC peuvent aussi infiltrer la région médiane et se propager vers le côté opposé à travers le corps calleux, donnant à la lésion une apparence en aile de papillon, caractéristique [33]. En revanche, la présence de composantes hémorragiques ou calcifiées est rarement observée, et contrairement au glioblastome, son principal diagnostic différentiel, la nécrose est généralement absente.

La séquence T2 FLAIR est utile pour l'évaluation de l'œdème vasogénique. Il apparaît le plus souvent étendu au-delà de la zone de prise de contraste, mais reste modéré sans effet de masse, contrairement à ce qui est observé dans les lésions gliales de haut grade ou les métastases. Par ailleurs, des cas de LPSNC avec un hypersignal isolé sur la séquence T2 FLAIR sans prise de contraste associée, ont également été rapportés dans la littérature mais cette présentation reste rare [35,36].

Les techniques d'IRM multimodale, telles que les séquences de diffusion, de perfusion ou de spectroscopie peuvent également être utiles au diagnostic de LPSNC. L'imagerie de diffusion a pour objectif de révéler les mouvements microscopiques de l'eau dans les tissus cérébraux. Elle repose sur l'idée que plus l'eau peut se déplacer librement dans une région, plus l'agitation moléculaire sera intense, ce qui se traduit par une diffusion élevée. À l'inverse, lorsque des obstacles entravent ces mouvements, la diffusion est réduite. Pour obtenir une séquence de diffusion, deux gradients de diffusion sont appliqués de part et d'autre de l'impulsion de radiofréquence de 180° dans une séquence de type EPI-SE (écho-planar spin écho) : d'abord un gradient $b=0$, puis un gradient b compris entre 500 et 3000 s/mm^2 .

Pour les protons immobiles, le déphasage provoqué par le premier gradient est entièrement compensé par le deuxième ; ainsi le signal des molécules d'eau immobiles n'est pas atténué (signal élevé). En revanche, les protons mobiles subissent un déphasage plus rapide lors de l'application du premier gradient, et ce déphasage n'est pas compensé par le deuxième gradient, entraînant une atténuation du signal (Figure 1.1). Cette atténuation est d'autant plus marquée que les mouvements moléculaires sont rapides.

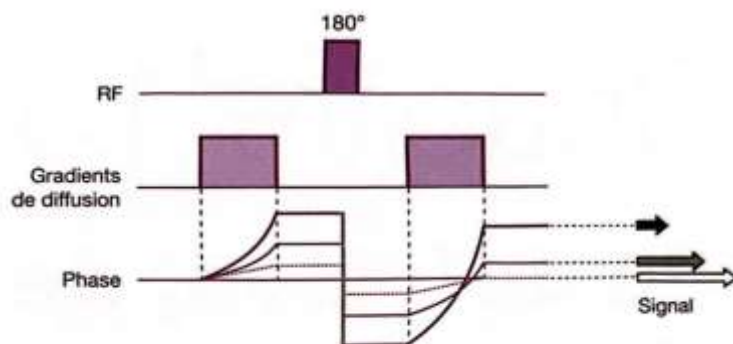


FIGURE 1.1 Principe des gradients de diffusion

D'après Comprendre l'IRM, B Kastler, Chapitre 15

Un autre paramètre mesurable en imagerie de diffusion est l'ADC (Coefficient de Diffusion Apparent en anglais). L'ADC permet une évaluation plus précise des anomalies de diffusion. Pour calculer l'ADC de différents tissus biologiques, une acquisition sans gradient de diffusion ($b=0$) est réalisée pour obtenir un signal de référence S_0 , suivie d'une acquisition comportant des gradients de diffusion à une valeur b déterminée, produisant un signal S . La valeur de l'ADC est alors calculée à partir de la relation : $\log(S/S_0) = -b \cdot \text{ADC}$ (Figure 1.2).

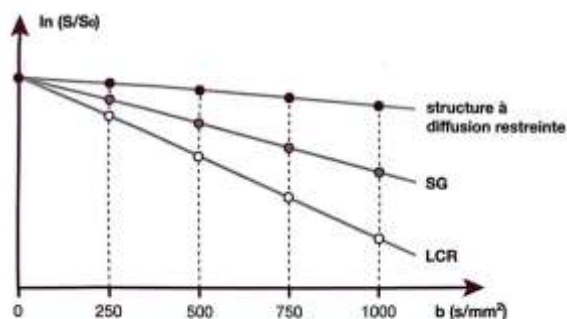


FIGURE 1.2 Calcul de l'ADC

D'après Comprendre l'IRM, B Kastler, Chapitre 15

Pour les tumeurs cérébrales, les recommandations récentes du BTIP (Brain Tumor Imaging Protocol) préconisent l'utilisation de trois valeurs de gradient b pour le calcul de l'ADC : 0, 500 et 1000 s/mm² [37].

Les LPSNC se distinguent par leur hypercellularité importante et un rapport nucléocytoplasmique élevé, ce qui se traduit par une forte restriction de la diffusion des molécules d'eau. En conséquence, les LPSNC apparaissent en hypersignal sur les séquences de diffusion et présentent des valeurs d'ADC relativement basses, généralement comprises entre 0,7 et 0,9.10⁻³ mm²/s. Toutefois, dans environ 10% des cas, il est possible d'observer des LPSNC sans restriction notable de la diffusion [38,39].

L'imagerie de perfusion est une technique très utile en neuro-oncologie, qui permet d'évaluer les variations du volume sanguin dans les tissus pathologiques. Son objectif est d'analyser la microcirculation dans une région spécifique en utilisant des traceurs, qu'ils soient d'origine exogène ou endogène. Lorsque l'imagerie de perfusion utilise des traceurs exogènes, elle repose sur l'utilisation de produits de contraste contenant généralement des chélates de gadolinium. Ces produits restent confinés dans les vaisseaux sanguins et ne franchissent pas la barrière hémato-encéphalique intacte. Lorsque le gadolinium se trouve dans les capillaires de la microvascularisation, il induit une diminution du signal sur la séquence de perfusion pondérée en T2*, par effet de susceptibilité magnétique (perfusion dynamique de contraste de susceptibilité magnétique ou DSC – dynamic susceptibility contrast).

Pour obtenir des informations sur la dynamique du signal liée au passage du produit de contraste, il est nécessaire de réaliser plusieurs acquisitions dans la région d'intérêt pendant l'injection du produit de contraste. Cette injection doit se faire en bolus pour obtenir une concentration élevée de gadolinium dans les vaisseaux. Les données obtenues permettent de calculer la variation d'intensité du signal en fonction du temps pour chaque voxel. Les résultats peuvent être représentés sous forme d'une courbe, avec le temps de passage du produit de contraste en abscisse et le pourcentage de perte de signal en ordonnée. La courbe obtenue présente une forme caractéristique, débutant par une ligne rectiligne représentant une intensité de signal normale (ligne de base), suivie d'une baisse du signal lors de l'arrivée du bolus de gadolinium, atteignant ensuite un pic correspondant à la concentration maximale du produit de contraste, pour enfin revenir à la ligne de base. À partir de cette courbe, plusieurs paramètres peuvent être calculés, notamment le temps d'arrivée du bolus (ou Bolus arrival time - BAT), le

temps d'arrivée du pic (ou Time to peak – TTP), le temps de transit moyen (Mean transit time – MTT) et l'amplitude du pic (correspondant à la perte maximale d'intensité de signal) (Figure 1.3). Le volume sanguin cérébral relatif (ou relative cerebral blood volume – rCBV) est déterminé par l'aire sous la courbe de décroissance du signal et exprimé en ml/100g de tissu. Une hypoperfusion se traduira par une diminution de cette aire. Le débit sanguin cérébral relatif (ou relative cerebral flow- rCBF) peut quant à lui être calculé par la relation suivante : $CBF = CBV / MTT$, et s'exprime en ml/100g de tissu/min.

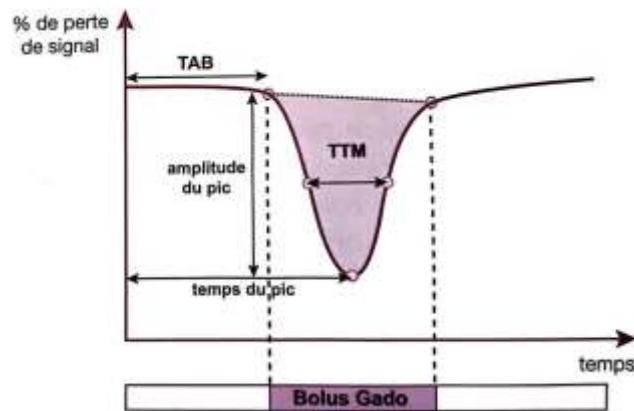


FIGURE 1.3 Courbe représentant le pourcentage de perte du signal en fonction du temps de passage du bolus de produit de contraste.

D'après Comprendre l'IRM, B Kastler, Chapitre 15

En raison de son tropisme périvasculaire, le LPSNC ne présente qu'une élévation très modérée du rCBV [40]. Sa courbe de perfusion présente un schéma temporel distinctif, caractérisé par un retour au-dessus de la ligne de base, souvent de manière horizontale [41] (Figure 1.4). Cela reflète la rupture de la barrière hémato-encéphalique et la diffusion du produit de contraste dans l'espace interstitiel.

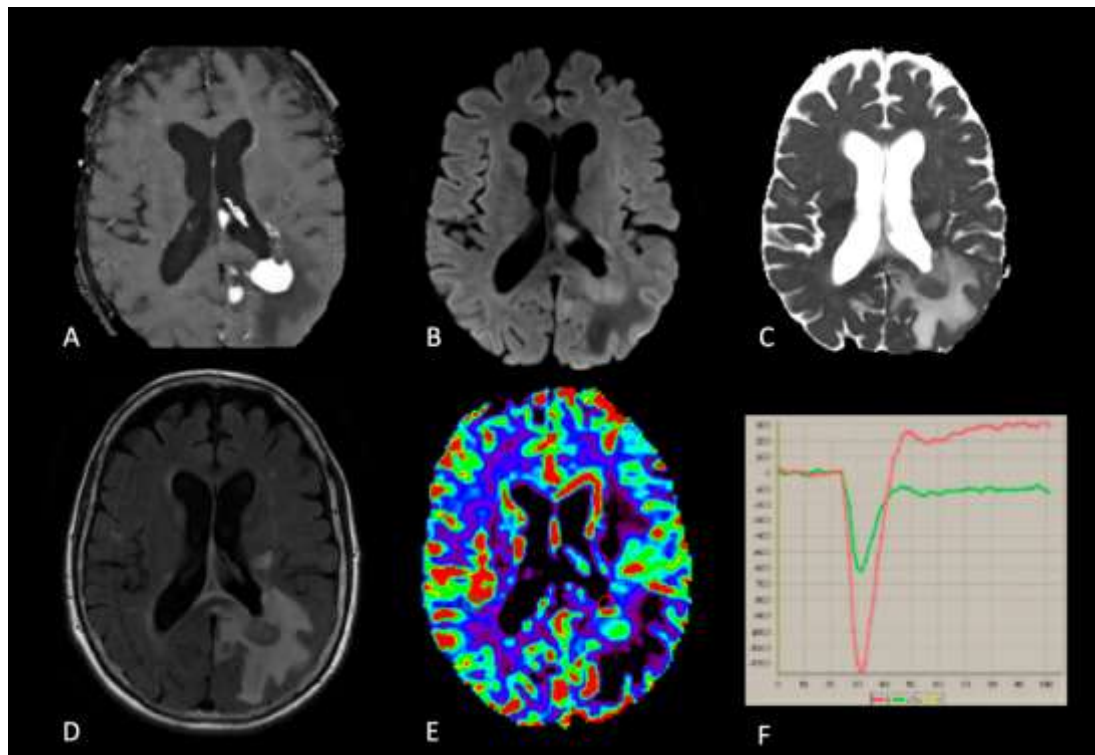


FIGURE 1.4 Image typique de LPSNC en IRM multimodale

(Amabilité Dr Lucia Nichelli)

Lésion temporale péri-ventriculaire gauche prenant intensément le contraste après injection de gadolinium, avec un aspect en « boule de neige » (A), avec un hypersignal sur la séquence de diffusion (B) traduisant une restriction de la mobilité des molécules d'eau, un hyposignal de la lésion et un hypersignal de l'œdème vasogénique sur la séquence T2 (C). Sur la séquence FLAIR, l'œdème apparaît en hypersignal, d'étendue relativement limitée (D). La carte paramétrique du rCBV ne montre pas d'hyperperfusion significative (E). En revanche, il existe un aspect typique de la courbe de perfusion (ligne rouge) avec un retour au-dessus de la ligne de base (ligne verte correspondant au tissu sain) (F).

Une autre méthode de perfusion, la perfusion DCE (dynamic contrast enhanced) analyse l'accumulation du gadolinium dans les espaces extracellulaires en séquence T1. L'analyse des images pondérées en T1 est soit qualitative soit semi-quantitative, notamment selon le modèle de Tofts, qui permet d'extraire trois principaux paramètres : la constante de transfert volumique du plasma sanguin vers l'espace extravasculaire (K_{trans}), le volume plasmatique (V_p) et le volume de l'espace extravasculaire extracellulaire par unité de volume tissulaire (V_e) [42]. Le K_{trans} reflète en partie la perméabilité des vaisseaux tumoraux et peut servir de marqueur de substitution de la néoangiogénèse tumorale.

Il est important de noter que dans l'étude Localyse, qui sera discutée plus en détail ultérieurement, les cartes paramétriques de rCBV ont été générées à l'aide du logiciel ADW (Advantage Workstation, General Electric). Ce logiciel effectue un prétraitement qui inclut

une correction pour le "leakage", ou fuite du produit de contraste. Ce phénomène se produit lorsque le gadolinium s'échappe du compartiment vasculaire vers le compartiment extravasculaire. Il est particulièrement prévalant dans les cas de LPSNC, en raison de la rupture de la barrière hémato-encéphalique (BHE) et de l'augmentation de la perméabilité vasculaire. Cette fuite peut provoquer une diminution du signal T2*, ce qui, sans correction, pourrait mener à une sous-estimation des valeurs de rCBV [43].

Une autre technique de perfusion cérébrale est connue sous le nom d'Arterial Spin Labeling (ASL). Cette technique consiste à utiliser des traceurs endogènes, c'est-à-dire des protons d'hydrogène "marqués". Dans cette procédure, une impulsion de radiofréquence est appliquée aux spins sanguins en mouvement avant qu'ils n'atteignent la zone du cerveau ciblée pour l'imagerie. Cette action produit une altération du signal IRM dans la région d'intérêt, qui est comparé au signal de référence, acquis sans le marquage des spins. Cette comparaison permet de déduire le flux sanguin cérébral. Peu d'études ont investigué le rôle de l'ASL dans les LPSNC. Notamment, seule son utilité pour le diagnostic différentiel entre pathologie lymphoïde et lésion gliale a pour le moment été mis en avant [44].

Enfin, la spectrométrie par résonance magnétique (SRM) ou spectroscopie, permet de donner des informations sur le contenu chimique d'une structure donnée. Elle fonctionne en détectant les différents métabolites présents dans un volume défini en se basant sur les variations uniques de leurs fréquences de résonance, qui sont ensuite analysées en domaine fréquentiel à la suite d'une transformation de Fourier. Dans les études présentées, la technique de spectrométrie monovoxel a été privilégiée. Cette méthode implique l'application de trois impulsions radiofréquences sélectives suivant les axes x, y et z pour isoler et acquérir deux spectres distincts : l'un après un temps d'écho (TE) long, l'autre après un TE court. Dans le cerveau sain, le spectre obtenu à partir d'un TE long révèle les métabolites à relaxation longue, comme le N-acétyl-aspartate (NAA), marqueur de l'intégrité neuronale ; la créatine (Cr), indicateur du métabolisme énergétique ; et la choline (Cho), associée à l'activité des membranes cellulaires. À l'inverse, le spectre à TE court permet de détecter des composés à relaxation courte, tels que le glutamate (Glu), signe de stress cellulaire, et le myo-inositol (mI), qui est lié à l'activation gliale. Dans les cas typiques de LPSNC, la SRM révèle généralement une augmentation du pic de choline et de lactate (indicateur de souffrance cellulaire anoxique), ainsi qu'une réduction significative du NAA. Le pic de lipides observé est habituellement moins prononcé que celui observé en cas de métastases cérébrales.

En conclusion, l'IRM cérébrale est l'examen de choix lorsqu'il existe une suspicion de lymphome cérébral. Comme mentionné précédemment, les principaux diagnostics différentiels sont le glioblastome et les métastases cérébrales. Il est important de souligner que l'administration de corticoïdes peut altérer l'apparence des lésions sur les images IRM [45]. De ce fait, il est fortement conseillé de réaliser l'IRM avant l'initiation de tout traitement cytotoxique altérant la perméabilité de la membrane, pour assurer la plus grande précision diagnostique.

Confirmation histologique

Pour confirmer un diagnostic de LPSNC, l'examen de référence demeure l'analyse histologique (Figure 1.5). Cette dernière peut être issue d'une biopsie stéréotaxique ou d'une vitrectomie si une atteinte oculaire est associée. L'analyse cytologique du liquide céphalo-rachidien (LCR) représente également une option diagnostique. Il est important que ces prélèvements soient effectués avant le début de tout traitement cytotoxique pour éviter des résultats faussement négatifs. Toutefois, l'obtention de ces échantillons peut s'avérer compliquée, en particulier chez les patients âgés et présentant de multiples comorbidités. Ainsi, de nouveaux biomarqueurs détectables dans le LCR et le vitré ont été proposés comme alternatives potentielles pour le diagnostic de LPSNC. Parmi ces indicateurs, une augmentation des niveaux d'interleukine 10 (IL-10) et un ratio élevé IL-10 / IL-6 se sont révélés être des marqueurs prometteurs [4,44,45]. En effet, l'IL-10 possède des propriétés anti-inflammatoires, responsables d'une forte activation immunitaire qui promeut la prolifération cellulaire y compris tumorale, et inhibe la mort cellulaire par apoptose. Certaines études ont retenu un seuil d'IL-10 à 2-3 pg/mL et un ratio IL10/IL6 >0.94 pour retenir le diagnostic de LPSNC [48–51].

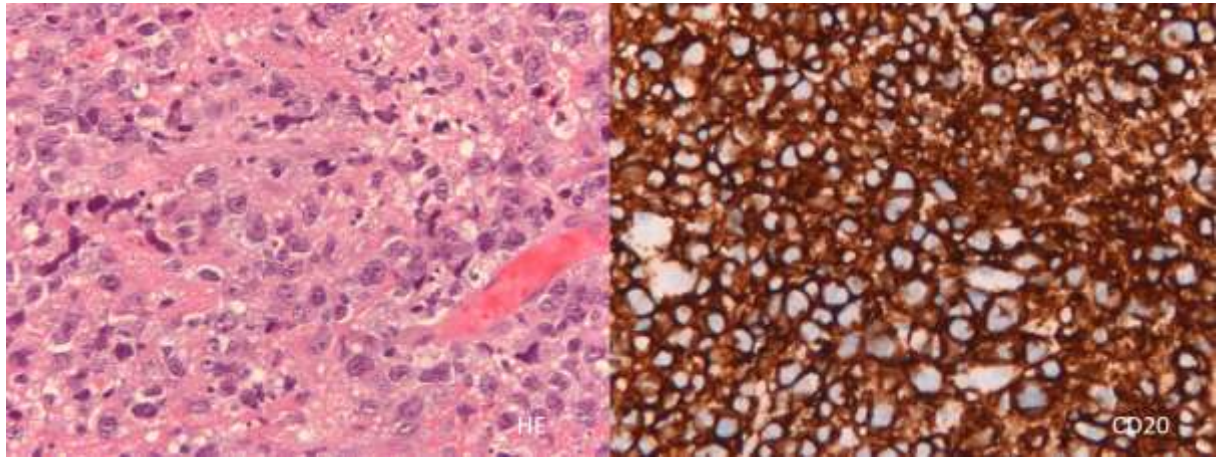


FIGURE 1.5 Coupe histologique d'un LPSNC

(Amabilité Dr. Suzanne Tran)

A gauche, coloration hématoxyline-éosine montrant une prolifération diffuse de grandes cellules. À droite, une immunocoloration pour le CD20 met en évidence une expression uniforme et dense de cet antigène de surface des lymphocytes B sur les cellules tumorales, ce qui confirme la lignée B du lymphome

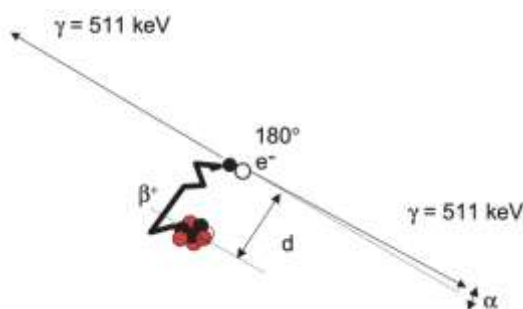
Bilan d'extension

Dans le cadre du lymphome cérébral, la distinction entre atteinte primitive et secondaire à une atteinte systémique est primordiale car déterminante pour la stratégie thérapeutique. Un scanner thoraco-abdomino-pelvien et une biopsie de moelle osseuse doivent ainsi être réalisés pour éliminer toute atteinte systémique. Suivant les recommandations établies en 2021 par le groupe d'experts sur les LPSNC (IPCG), une tomographie par émission de positon au 18F-Fluorodeoxyglucose (TEP-FDG) est recommandée pour compléter le bilan d'extension [5]. Des études récentes ont montré que la TEP sur l'ensemble du corps (de mi-tête à mi-cuisses) peut révéler entre 4 et 8 % de cas d'atteintes systémiques non détectées par les méthodes d'imagerie classique [5,52,53]. Le chapitre suivant abordera en détail cette modalité d'imagerie et discutera de son rôle potentiel diagnostique et pronostique dans la prise en charge du LPSNC.

Chapitre 2. Le rôle de la TEP

I. Principes physiques

La médecine nucléaire, qui inclut l'imagerie TEP et la scintigraphie conventionnelle, est une modalité d'imagerie fonctionnelle, permettant d'imager des processus pathologiques *in vivo* en ciblant des molécules d'intérêt à l'aide d'un radiopharmaceutique. Celui-ci est constitué par un vecteur biologique dirigé vers le processus étudié et d'un isotope radioactif. Dans le cadre de l'imagerie TEP, le radioisotope possède un excès de protons et revient à son état stable par désintégration β^+ . Lors de la désintégration, un proton (p) du noyau va être converti en neutron (n) entraînant l'émission d'une particule β^+ , le positon (β^+), et d'un neutrino électronique (ν_e) : $p^+ \rightarrow n + \beta^+ + \nu_e$. Après son émission, le positon ne parcourt qu'une courte distance, le libre parcours moyen, de l'ordre de quelques millimètres, avant de rencontrer un électron avec lequel il va s'annihiler. Cette annihilation va produire deux photons gamma de 511 keV qui sont émis en coïncidence, c'est à dire à 180 degré l'un de l'autre [54] (Figure 2.1).



(D'après Dreuille et al. EMC. 2002)

FIGURE 2.1 Devenir dans la matière d'un positon émis par le 18F-FDG

Chaque paire de photons va ensuite être détectée par la couronne de scintillateurs placée autour du sujet, sous forme d'une ligne de réponse (LOR). Pour obtenir une image fidèle aux processus *in vivo*, il est essentiel de détecter uniquement les vraies coïncidences, c'est-à-dire celles résultant d'une annihilation générant deux photons n'ayant subi aucune interaction avant leur détection. Toutefois, la détection inclut également des LOR non idéales qui peuvent dégrader la qualité de l'image, notamment les coïncidences diffusées et les coïncidences fortuites [55] (Figure 2.2). Les coïncidences diffusées proviennent de photons qui ont subi une ou plusieurs interactions par diffusion Compton avant leur détection tandis que les coïncidences fortuites

proviennent de mise en en coïncidence de deux photons ne provenant pas de la même annihilation. Pour pallier à ces biais physiques, des corrections mathématiques sont intégrées dans la reconstruction des images et sont indispensables avant l'analyse par le médecin nucléaire.

Une autre limitation physique à considérer est l'effet de volume partiel, qui conduit à une sous-évaluation de l'intensité de fixation des structures de petite taille, typiquement inférieures à deux ou trois fois la résolution spatiale du système d'imagerie. Sur le plan mathématique, l'intensité perçue en un point donné de l'image est le résultat de la convolution de la fonction de réponse du système à ce point spécifique. Cela entraîne le déversement de l'intensité d'un voxel sur ses voisins, générant une corrélation locale entre les voxels qui modifie la texture réelle de l'image. Aujourd'hui, cette limite est corrigée par l'implémentation de méthode de correction de la fonction de dispersion ponctuelle (ou PSF - point spread function) [56–58].

Du fait de la détection en coïncidences, le système de détection est organisé en une couronne placée autour du patient. Chaque sous-ensemble de la couronne est composé d'un scintillateur, sensible à la détection des photons de 511keV et de photomultiplicateurs, qui vont convertir les photons en signal électrique et l'amplifier pour le rendre exploitable par le système informatique.

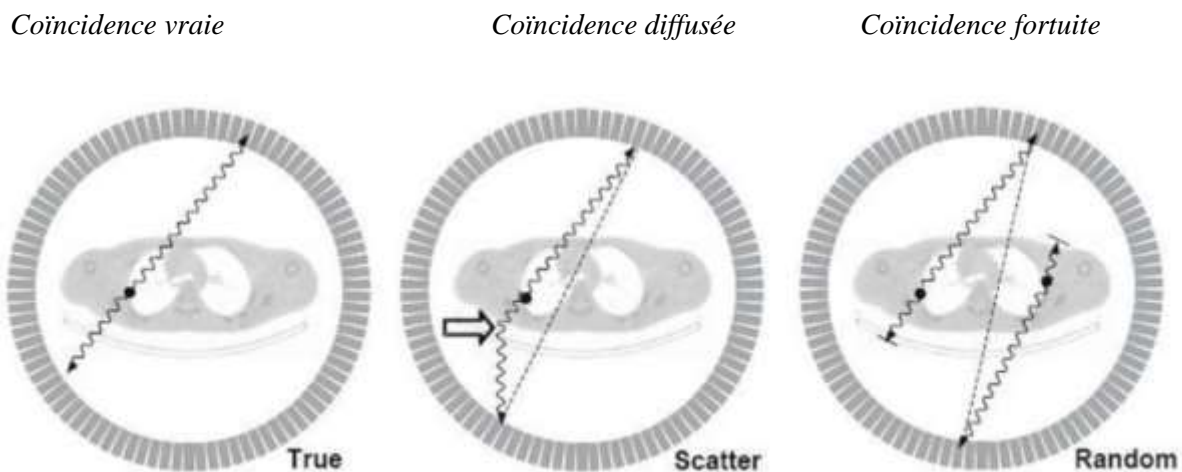


FIGURE 2.2 Types de LOR pouvant être enregistrées par un imageur TEP
(*True coïncidence vraie, scatter : coïncidence diffusée, random : coïncidence fortuite*)

II. 18F-Fluorodeoxyglucose

Dans le domaine de l'imagerie par TEP, en particulier en oncologie, le radiotraceur le plus couramment utilisé est le 18F-Fluorodeoxyglucose (FDG). Le FDG est un analogue structural du glucose, dans lequel le groupe hydroxyle en position 2 est remplacé par un atome de fluor 18. Semblable au glucose, le FDG transite rapidement des vaisseaux sanguins vers l'espace interstitiel, puis est transporté à l'intérieur de la cellule via les transporteurs de glucose GLUT1 et GLUT3. Une fois dans la cellule, il est phosphorylé en FDG-6-phosphate par l'enzyme hexokinase. Contrairement au glucose-6-phosphate qui est soit isomérisé en fructose-6-phosphate dans la voie glycolytique, soit oxydé en 6-phosphogluconolactone dans la voie des pentoses phosphates, le 18F-FDG-6-P ne peut subir de catabolisme supplémentaire, à cause de l'absence d'un atome d'oxygène en position C-2. Par conséquent, le FDG est piégé à l'intérieur de la cellule et s'y accumule progressivement [59,60].

L'intérêt de ce traceur est particulièrement marqué en cancérologie, car il a été établi que les cellules cancéreuses produisent de l'énergie principalement par glycolyse anaérobie plutôt que par phosphorylation oxydative mitochondriale, un phénomène connu sous le nom d'effet Warburg. C'est cette caractéristique qui explique pourquoi une large variété de tumeurs montre une forte captation du FDG [61].

Par ailleurs, il est à noter que la demi-vie physique du fluor-18 est de 110 minutes permettant une relative souplesse dans la gestion de ce produit (pas de nécessité de cyclotron sur place) et une exposition limitée aux radiations [62].

III. Quantification

L'interprétation des images TEP en routine clinique comprend souvent une première étape d'analyse visuelle destinée à repérer les régions de fixation anormale par rapport à la fixation physiologique des organes sains. Toutefois, cette analyse est généralement complétée par une évaluation quantitative, qui réduit la dépendance à l'observateur et améliore l'évaluation dans le suivi longitudinal. L'index de mesure privilégié est la valeur de fixation normalisée, ou standardized uptake value (SUV), calculée selon la formule ci-dessous :

$$SUV = \frac{\text{Fixation (Bq/mL)}}{\text{activité (Bq)/volume (mL)}}$$

L'estimation de la valeur du SUV repose sur l'hypothèse que la densité corporelle du patient est équivalente à celle de l'eau, soit 1 g/mL, ce qui fait du SUV une mesure dimensionnelle. Cependant, il est essentiel d'utiliser le SUV avec prudence car plusieurs facteurs peuvent influencer son calcul, y compris les protocoles spécifiques d'acquisition et de reconstruction des images [63–65].

Différents paramètres semi-quantitatifs sont habituellement mesurés en TEP :

- le SUV_{max} : représente la valeur maximale de captation dans un voxel unique de la ROI. Ce paramètre peut être sensible au bruit de l'image, à la résolution spatiale et aux paramètres d'acquisition (durée d'acquisition, algorithme de reconstruction et corrections appliquées).
- le SUV_{mean} : correspond à la valeur moyenne du SUV dans une ROI.
- le SUV_{peak} : représente la valeur moyenne à l'intérieur d'une sphère de 1cm³ dans la région la plus fortement hypermétabolique à l'intérieur de la ROI, selon la définition retenue par les critères de réponse en TEP pour les tumeurs solides, PERCIST [66].
- le MTV (metabolic tumor volume) : correspond au volume métabolique au sein de la ROI et dépend du seuillage prédéfini. Pour cette thèse, nous avons choisi d'adopter un seuil de 41% du SUV_{max}, en accord avec les directives de l'EANM pour l'imagerie des tumeurs oncologiques présentant une absorption élevée et uniforme, ce qui correspond aux caractéristiques des LPSNC [67]. Cette approche est depuis longtemps établie dans le suivi des lymphomes systémiques, notamment le lymphome de Burkitt et le lymphome de Hodgkin [68,69]. Bien que des discussions récentes aient questionné la pertinence de ce seuil de 41% du SUV_{max}, notamment pour la segmentation automatisée qui mesure la masse tumorale totale et peut rencontrer des tissus tumoraux hétérogènes, il est important de préciser que ces considérations ne concernent pas le LPSNC [70]. En effet, pour le LPSNC où les anomalies sont principalement cérébrales, l'évaluation de la charge tumorale totale est moins complexe que pour les lymphomes disséminés.
- le TLG (total lesion glycolysis) : concept introduit en 1999 par Larson et al, et qui définit le produit du SUV_{mean} et du MTV [71].

IV. Intérêt de la TEP-IRM

En médecine nucléaire, l'utilisation de caméras hybrides combinant l'imagerie fonctionnelle par TEP et l'imagerie anatomique par TDM a été rapidement adoptée depuis la fin des années 90. L'intégration de l'IRM à la TEP, bien que conceptualisée à la même époque, a été confrontée à de plus grands défis techniques, le principal étant la nécessité de faire fonctionner les détecteurs TEP en présence d'un champ magnétique intense et d'ondes de radiofréquences émises pendant les acquisitions IRM, sans interférence entre les deux modalités. De plus, l'espace limité à l'intérieur des tunnels d'IRM pour insérer les détecteurs de TEP complique encore cette intégration.

Des prototypes pour petits animaux ont d'abord été développés ; puis, depuis 2011, des systèmes TEP-IRM pour le corps entier ont été commercialisés pour un usage clinique. Les premières tentatives par General Electric (GE) Healthcare avec le système Tri-modality PET/CT-MR, et par Philips avec le système Ingenuity TF PET/MR, impliquaient des acquisitions séquentielles. Dans le système de GE, la table d'examen était transférée entre deux pièces distinctes, tandis que dans celui de Philips, elle était conçue pour pivoter entre les deux modalités. Cependant, ces configurations ne permettaient pas un recalage parfait des images, même avec le patient immobilisé. Ces modèles ont depuis été retirés du marché.

Le premier système TEP-IRM intégré a été commercialisé par Siemens sous le nom de Biograph mMR et installé à l'Université de Munich en 2011. Plus récemment, GE a développé le système SIGNA PET/MR qui est installé à l'hôpital de la Pitié-Salpêtrière. Ce système utilise des détecteurs TEP à base de semi-conducteurs au silicium (SiPM), plutôt que des photodiodes à avalanche, ce qui permet l'utilisation de la technologie de temps de vol (Time Of Flight, TOF) [72]. Celle-ci mesure les différences de temps d'arrivée des photons de 511keV en coïncidence, permettant ainsi une localisation plus précise des événements d'annihilation. Sur la caméra GE, le détecteur TEP est inséré dans un tunnel IRM de 3 Tesla, avec un champ de vue maximal de 50 cm x 50 cm x 50 cm.

L'alliance de la TEP et de l'IRM assure une correspondance spatiale parfaite des images fonctionnelles et anatomiques, ce qui améliore considérablement la précision diagnostique et la comparabilité des données en recherche. En effet, bien que le cerveau soit un organe rigide, les logiciels de recalage offrent rarement une satisfaction totale lorsque l'IRM est réalisée séparément. Grâce à sa haute résolution spatiale, la TEP-IRM est particulièrement efficace pour

évaluer les petites structures anatomiques et détecter les changements subtils induits par le traitement. Elle réduit aussi l'exposition aux radiations ionisantes et optimise le temps de prise en charge du patient.

Outre son utilité dans le diagnostic des maladies neurodégénératives, la TEP-IRM est fréquemment employée pour les patients atteints de cancers pelviens, notamment gynécologiques et prostatiques, ainsi que pour les cancers ORL et en neuro-oncologie [73,74]. Plusieurs études ont investigué l'imagerie des gliomes en TEP-IRM et ont démontré que les deux techniques fournissent des informations complémentaires. En effet, les anomalies détectées par TEP et IRM ne coïncident pas systématiquement, suggérant l'existence de processus physiopathologiques distincts. Par ailleurs, des résultats prometteurs ont été publiés concernant la prédiction moléculaire *in vivo*, la planification de traitement en radiothérapie ou encore le suivi thérapeutique des gliomes, grâce à la TEP-IRM [75–78]

La TEP-IRM se distingue du TEP-CT par l'utilisation de techniques de correction d'atténuation différentes. Alors que le TEP-CT utilise les unités Hounsfield pour générer une carte d'atténuation paramétrique, la TEP-IRM, n'ayant pas accès à ces données, compense par l'utilisation de méthodes de segmentation, en particulier avec les séquences Dixon. Ces dernières permettent une classification tissulaire en quatre catégories (eau, graisse, air, poumon) et assignent à chaque catégorie un coefficient d'atténuation prédéfini [79]. Néanmoins, elles ne segmentent pas l'os, rendant nécessaire l'emploi d'un atlas pour la correction de l'atténuation osseuse par recalage élastique. Cette approche est controversée en raison des variations osseuses individuelles non prises en compte et des biais systématiques dans la quantification des intensités de fixation, mis en avant notamment par l'étude de Sgard et al. [80]. Actuellement, une méthode utilisant des séquences IRM à temps d'écho (TE) très court est privilégiée. La méthode ZTE (zero TE) avec un TE de l'ordre de quelques microsecondes permet de détecter l'os cortical en IRM, créant un 'pseudo-CT' et il a été démontré que cette technique réduisait considérablement les erreurs de mesure, tant régionales qu'inter-individuelles. Actuellement, des techniques basées sur des méthodes d'intelligence artificielle sont également à l'étude pour perfectionner cette correction d'atténuation [81].

Il est à noter que, pour la recherche présentée dans cette thèse, la méthode Dixon combinée à un atlas a été retenue, car les premiers patients inclus en 2016 ne disposaient que de cette technique et sa continuité a été maintenue pour assurer l'homogénéité des données.

Un autre défi actuel en TEP-IRM concerne la correction des mouvements. En imagerie cérébrale, les mouvements de la tête peuvent significativement altérer la qualité des images. Plusieurs techniques sont actuellement en développement pour améliorer la correction des mouvements, incluant des méthodes de suivi optique externe mais également des approches basées sur les données IRM, telles que l'utilisation de séquences BOLD (blood oxygenation level-dependent) ou écho planaire dynamiques [74].

V. Article 1.

La revue ci-après présente le rôle de la TEP dans la prise en charge du LPSNC.

Intérêt de la TEP dans les lymphomes primitifs du SNC. Cancers, 2022

Points clés :

- La TEP-FDG corps entier est recommandée selon les guidelines IPCG de 2021 pour confirmer l'origine primitive du lymphome cérébral.
- Le LPSNC montre un hypermétabolisme du FDG nettement supérieur, environ 2 à 3 fois celui du tissu cérébral sain adjacent.
- Aucune corrélation significative n'a été établie entre l'intensité du métabolisme FDG dans le LPSNC et le volume tumoral.
- Les effets des stéroïdes sur la réduction du métabolisme du FDG en TEP restent sujets à discussion.
- Pour le diagnostic différentiel avec le glioblastome multiforme chez les patients immunocompétents, l'association d'une forte fixation en FDG et d'une faible perfusion en ASL tend à indiquer un LPSNC.
- Chez les patients immunodéprimés, la TEP au FDG aide au diagnostic différentiel entre le LPSNC et la toxoplasmose, ce dernier présentant une fixation du FDG très faible.
- Différents paramètres pronostiques ont été explorés en TEP, dont le SUV_{max}, le MILAS (metabolic imaging lymphoma aggressiveness scale ; un score visuel basé sur une référence cérébelleuse, présenté plus en détail au chapitre III, partie II), le MTV et le TLG ; cependant, il n'existe pas de consensus sur leur performance entre les études.
- Des études ont examiné l'utilité de la TEP comme outil d'évaluation thérapeutique, mais il n'y a pas d'accord sur les critères spécifiques ou le calendrier optimal pour le suivi.
- Au-delà du FDG, d'autres traceurs sont à l'étude, tels que la ¹¹C-Méthionine, le ⁶⁸Ga-Pentixafor ou la ¹⁸F-Fludarabine.

Review

Role of Positron Emission Tomography in Primary Central Nervous System Lymphoma

Laura Rozenblum ^{1,2,*} , Caroline Houillier ³, Carole Soussain ⁴, Marc Bertaux ⁵ , Sylvain Choquet ⁶, Damien Galanaud ^{2,7}, Khê Hoang-Xuan ³ and Aurélie Kas ^{1,2}

¹ Service de Médecine Nucléaire, Hôpitaux Universitaires Pitié-Salpêtrière Charles Foix, AP-HP, Sorbonne Université, 75013 Paris, France

² Laboratoire d'Imagerie Biomédicale, INSERM U1146, 75006 Paris, France

³ ICM, et IHU, Service de Neurologie 2, Hôpitaux Universitaires Pitié-Salpêtrière Charles Foix, AP-HP, Sorbonne Université, 75013 Paris, France

⁴ Service d'Hématologie, Institut Curie, Site Saint-Cloud and INSERM U932 Institut Curie, Université PSL, 75005 Paris, France

⁵ Service de Médecine Nucléaire, Hôpital Foch, 92150 Suresnes, France

⁶ Service d'Hématologie Clinique, Hôpitaux Universitaires Pitié-Salpêtrière Charles Foix, AP-HP, Sorbonne Université, 75013 Paris, France

⁷ Service de Neuroradiologie, Hôpitaux Universitaires Pitié-Salpêtrière Charles Foix, AP-HP, Sorbonne Université, 75013 Paris, France

* Correspondence: laura.rozenblum@gmail.com

Simple Summary: Primary central nervous system lymphoma (PCNSL) is a rare but highly aggressive lymphoma with increasing incidence in immunocompetent patients. To date, the only established biomarkers for survival are age and functional status. Currently, the magnetic resonance imaging (MRI) criteria of the International Collaborative Group on Primary Central Nervous System Lymphoma are the only ones recommended for follow-up. However, early occurrence of recurrence after treatment in patients with a complete response on MRI raises the question of its performance in assessing residual disease. While the use of ¹⁸F-fluorodeoxyglucose body positron emission tomography for identification of systemic disease has been established and can be pivotal in patient treatment decisions, the role of brain PET scan is less clear. Here we review the potential role of PET in the management of patients with PCNSL, both at diagnosis and for follow-up under treatment.

Abstract: The incidence of primary central nervous system lymphoma has increased over the past two decades in immunocompetent patients and the prognosis remains poor. A diagnosis and complete evaluation of the patient is needed without delay, but histologic evaluation is not always available and PCNSL can mimic a variety of brain lesions on MRI. In this article, we review the potential role of ¹⁸F-FDG PET for the diagnosis of PCNSL in immunocompetent and immunocompromised patients. Its contribution to systemic assessment at the time of diagnosis has been well established by expert societies over the past decade. In addition, ¹⁸F-FDG provides valuable information for differential diagnosis and outcome prediction. The literature also shows the potential role of ¹⁸F-FDG as a therapeutic evaluation tool during the treatment and the end of the treatment. Finally, we present several new radiotracers that may have a potential role in the management of PCNSL in the future.

Keywords: PCNS lymphoma; ¹⁸F-FDG PET; outcome; therapeutic evaluation



Citation: Rozenblum, L.; Houillier, C.; Soussain, C.; Bertaux, M.; Choquet, S.; Galanaud, D.; Hoang-Xuan, K.; Kas, A. Role of Positron Emission Tomography in Primary Central Nervous System Lymphoma. *Cancers* **2022**, *14*, 4071. <https://doi.org/10.3390/cancers14174071>

Academic Editor: David Wong

Received: 23 June 2022

Accepted: 19 August 2022

Published: 23 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Epidemiology

Primary central nervous system lymphoma (PCNSL) is a highly aggressive non-Hodgkin's lymphoma confined to the central nervous system (CNS), including the brain, leptomeninges, eyes, and spinal cord. PCNSL can develop in both immunosuppressed (IS) (HIV, organ transplant, immunosuppressive agents) and immunocompetent (IC) patients.

The diagnostic confirmation is obtained from histopathology after stereotactic brain biopsy, or from vitreous or cerebrospinal fluid (CSF) cytology. In IC patients, PCNSL are mostly diffuse large B-cell lymphomas (>90%) [1,2]. In IS patients, brain lymphoma disease is essentially related to cytokine dysregulation that results from opportunistic infections with Epstein–Barr virus [3]. PCNSL incidence in IS patients greatly exceeds (about 1000 times higher) that of the general population. In IC patients, PCNSL is rare, accounting for approximately 6% of new malignant primary CNS tumors, but its incidence has increased over the past two decades, most notably in the elderly [4–6]. The first line treatment relies on an induction polychemotherapy based on high-dose methotrexate follow-up by a consolidation treatment with whole-brain radiotherapy or autologous stem cell transplantation in younger patients [7]. Despite advances in the management of this disease, median overall survival remains poor in the elderly population (15 months) [8,9]. In a French cohort of 1000 patients with newly diagnosed PCNSL, 25% of the elderly patients died within the first 6 months after diagnosis [9].

2. Brain Imaging Features at Diagnosis

2.1. Magnetic Resonance Imaging (MRI)

MRI is the modality of choice for the initial assessment and the follow-up of PCNSL. At diagnosis, the European Association of Neuro-Oncology (EANO) and International PCNSL Collaborative Group (IPCG) guidelines recommend T2 weighted/fluid-attenuated inversion recovery sequence (T2/FLAIR) and T1-weighted sequences pre- and post-contrast injection [10,11].

In IC, half of the lesions (50–70%) are solitary, most often located in the periventricular area (e.g., thalamus, corpus callosum and basal ganglia). The lesions usually appear isointense on T1-weighted images, slightly hypointense on T2-weighted images, and show intense homogeneous contrast enhancement without necrosis. Due to their high cellularity, brain lymphomas restrict the normal random movement of water molecules in the brain tissue, corresponding to “restricted diffusion.” This results in a hypersignal on diffusion-weighted imaging (DWI) and low apparent diffusion coefficient (ADC) values on ADC parametric maps. Vasogenic surrounding edema is usually present on T2-weighted images but classically with little mass effect relative to size [12]. On dynamic-susceptibility contrast (DSC) perfusion images, PCNSL does not show florid neoangiogenesis but a characteristic type of curve with a significant increase in signal intensity above the baseline due to massive leakage of contrast media into the interstitial space [13]. Spectroscopy can measure several different metabolites that reflect specific cellular and biochemical processes to differentiate tumor types [14]. However, PCNSL spectroscopy is not very specific and shows increased choline and decreased N-acetylaspartate (NAA), creatine (Cr), and myoinositol, which can be observed in several types of brain tumors. Nevertheless, a prominent lipid peak on spectroscopy in brain tumors without central necrosis seems to be a feature of malignant brain lymphoma [15].

Few studies have also reported atypical presentation such as disseminated lesions or T2-hyperintense lesions with variable diffusion restriction and no contrast enhancement [16–19] (Figure 1). In a retrospective analysis of 127 IC patient scans, Bataille et al. reported nine cases of exclusively cortical lesions mimicking a meningioma [20]. Less frequent presentations involve the eyes, leptomeninges, spine, or vascular system [21]. Only 1% of PCNSL present with spinal involvement, which typically appears as a solitary infiltrative lesion with multifocal areas of contrast enhancement. On the other hand, intravascular lymphoma is responsible for blood slowing that can lead to endoluminal thrombosis or microinfarcts.

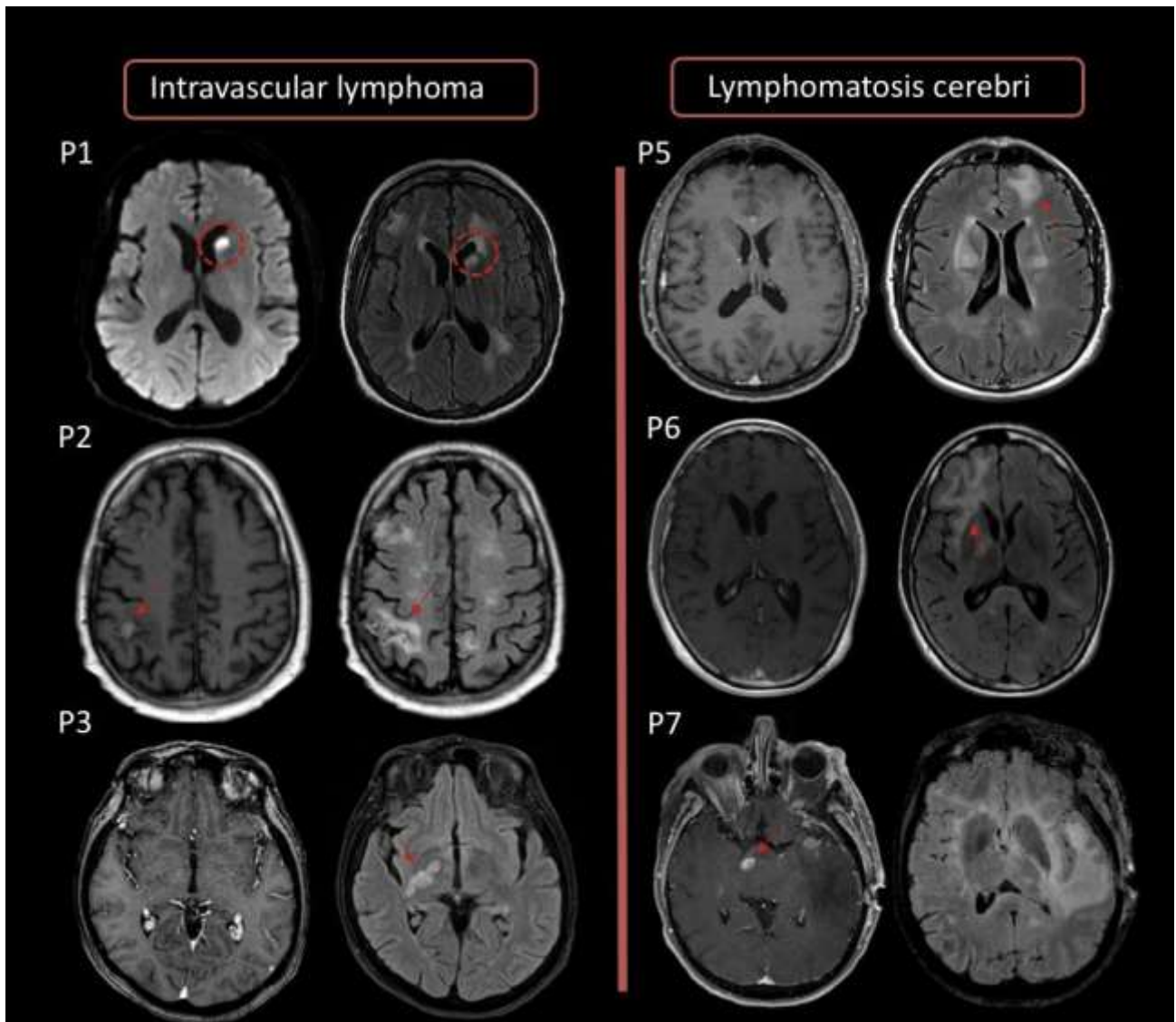


Figure 1. Case examples of MRI presentations of intravascular lymphoma (Patients 1, 2, and 3) and cerebral lymphomatosis (Patients 5, 6, and 7). **Left:** intravascular lymphoma presents as cortical, subcortical, or periventricular lesions with FLAIR hypersignal (2nd column) and little or no gadolinium-weighted T1+ enhancement. The diffusion-weighted image of the first patient (P1) shows a high hypersignal of the head of the caudate nucleus mimicking a stroke (red dotted circle). **Right:** lesions of cerebral lymphomatosis appear as infiltrating, confluent lesions with FLAIR hypersignal, in the deep and subcortical white matter (red arrows), without mass effect (4th column). T1+ gadolinium-weighted images (3rd column) show no enhancement, with the exception of the last patient (bottom right) who had both typical homogeneous enhanced lesions and an infiltrating FLAIR hypersignal.

On MRI, this subtype may appear as diffuse leukoencephalopathy or multiple small infarcts [21]. In IS patients, lesions are often multifocal, with nodular or peripheral ringlike enhancement reported in up to 75% of cases.

Although these imaging features are suggestive for PCNSL, they are not pathognomonic and several differential diagnoses may be discussed in case of suspected brain lymphoma. Therefore, confirmation of the diagnosis of PCNSL by histology or cytology is essential before treatment can be started and, if clinically possible, biopsy should be done

prior to steroids as they decrease the sensitivity of the histopathological diagnosis [10]. In IC patients, the differential diagnosis includes other types of brain tumors (mainly high-grade gliomas, brain metastases), whereas progressive multifocal leukoencephalopathy and cerebral toxoplasmosis can be discussed in IS patients (Table 1) [12,22].

2.2. ^{18}F -FDG Positron Emission Tomography (PET)

2.2.1. Features at Diagnosis

Positron emission tomography is a molecular imaging that allows the visualization of biological targets and the detection of physiopathological processes in vivo, by targeting molecules of interest with a radiopharmaceutical radiolabeled with a short-lived radionuclide such as carbon-11 and fluorine-18. Today, ^{18}F -FDG is the workhorse radiopharmaceutical used in patients with brain conditions or cancer. As has been shown for systemic lymphomas, PCNSL exhibits elevated glucose metabolism and a high avidity for ^{18}F -FDG (2–3 times higher than of the healthy grey matter) due to very high cellular density with an accelerated glycolytic metabolism [19,23] (Figure 2).

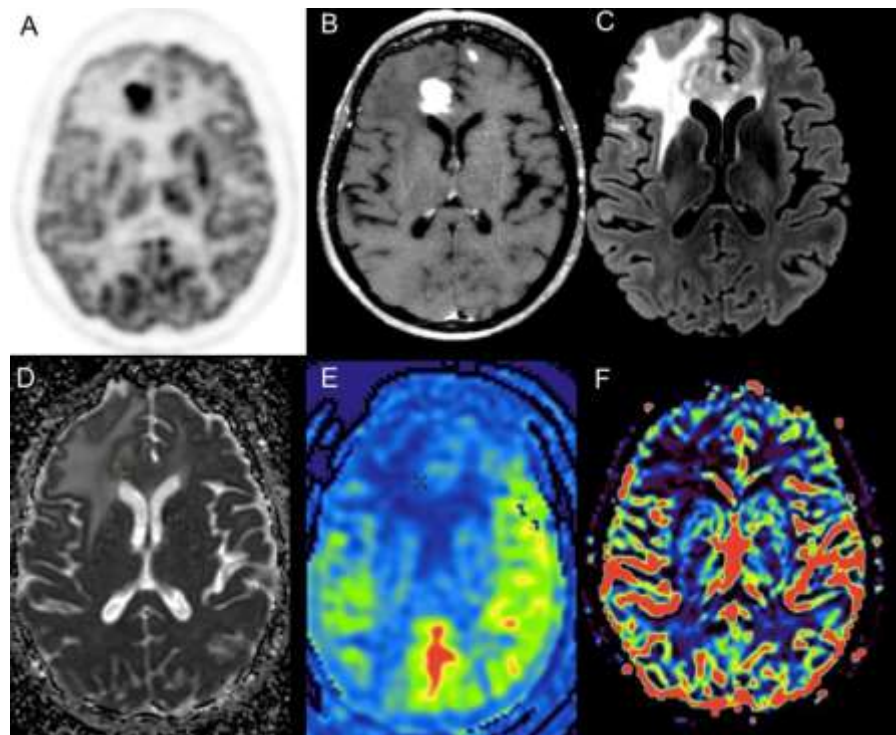


Figure 2. Primary central nervous system lymphoma in a 72-year-old female patient with typical PET and MRI features: high and homogeneous ^{18}F -FDG PET uptake (SUVmax = 27, T/ N = 2.1) (A); “snowball” pattern with intense and homogeneous enhancement after gadolinium enhancement (B); extensive hypersignal on T2-FLAIR weighted images, respecting perilesional edema and respecting the cortical ribbon (C). Parametric ADC map show only moderate diffusion restriction (D), and no hyperperfusion is detectable either on arterial spin labelling (E), or dynamic susceptibility contrast (F, rCBV parametric map) perfusion sequences.

In a recent meta-analysis including 486 patients from 22 studies, pooled maximum standardized uptake value (SUVmax) in PCNSL ranged from 8.4 to 27.8 with a mean of 18.1 (95% CI, 16.0–20.1) compared to 10.4 for inflammatory or malignant brain diseases other than PCNSL [24]. However, as a SUVmax cut-off value might be influenced by the model of scanner, imaging protocol or injected dose, the ratio of tumor to a contralateral mirror region of unaffected brain tissue (T/ N) is preferentially analyzed. Gupta et al. showed

that a T/ N ratios cut-off at 1.66 led to a PPV and NPV of 90% (95% CI: 79–96%) and 69% (95% CI: 54–81%) for differentiating CNS lymphomas from non-lymphomas diseases [25].

A limitation that has been raised for PCNSL detectability with ^{18}F -FDG-PET is the potential impact of corticosteroids [26–28]. Corticosteroids can induce apoptosis of lymphoma cells very rapidly and it has been recognized that pre-administration of corticosteroids can significantly shrink the size of the tumor and even induce tumor disappearance on MRI and histological evaluation [29]. However, the impact of steroids on ^{18}F -FDG uptake remains unclear. In a small cohort, Rosenfeld et al. first described a difference in SUV between steroid-treated ($n = 7$) and untreated cases ($n = 3$) [28]. Yamagashi et al. confirmed a negative correlation between T/ N and cumulative corticosteroid dose in a retrospective study of 19 patients. However, these differences did not reach statistical significance in either study [28,30].

Of note, ^{18}F -FDG avidity appears to be independent of tumor size [31]. In addition, atypical MRI presentation forms of PCNSL (disseminated lesions with faint enhancement, lack of enhancement, and ring-shaped enhanced lesions) have been described with significantly lower uptake values than typical PCNSL [32] (Figures 3 and 4). Recently, Kim et al. suggested that low ^{18}F -FDG avidity was more frequently related to negativity for MUM1 expression, a protein known to be a crucial regulator of B-cell development and tumorigenesis [33].

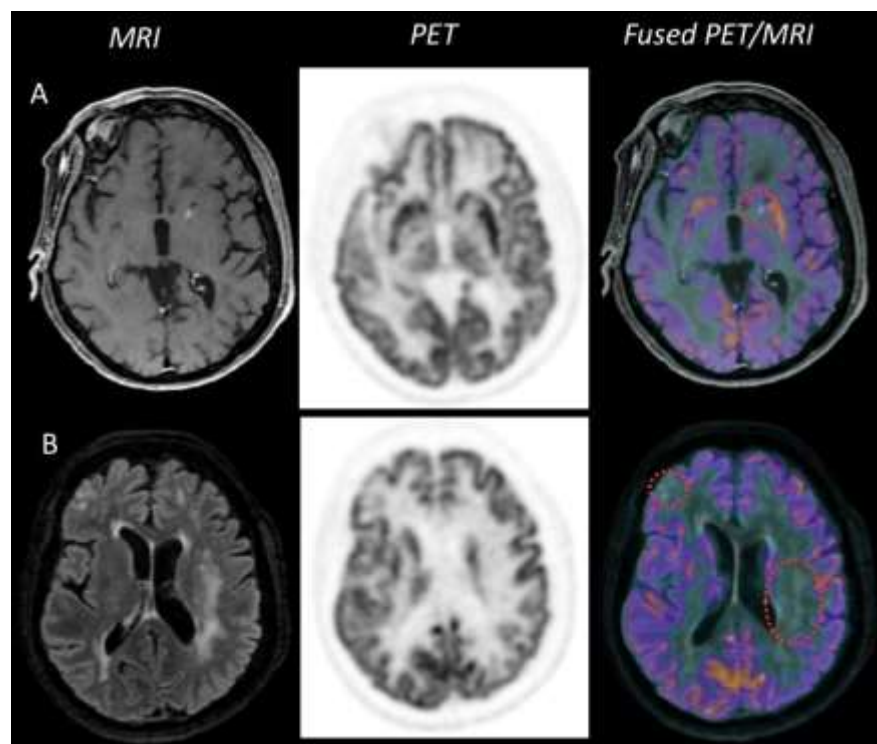


Figure 3. Example case of PET-MRI images of an intravascular lymphoma. Gadolinium-weighted T1+ imaging (A) showed a small enhanced punctiform lesion in the medial part of the left striatum, with no metabolic uptake on ^{18}F -FDG PET images. The cortical and subcortical lesions in FLAIR hypersignal (B) did not show uptake with ^{18}F -FDG either.

In IC patients, the main differential diagnosis of PCNSL is high-grade glioma (HGG), the most prevalent and aggressive primary malignancy of the brain parenchyma. To distinguish between these two entities, brain imaging techniques are valuable tools: on the one hand, to help clinicians decide between stereotactic biopsy in case of suspected lymphoma or extensive surgery in case of suspected HGG; on the other hand, to quickly decide on therapeutic management if the patient is recused for invasive procedure.

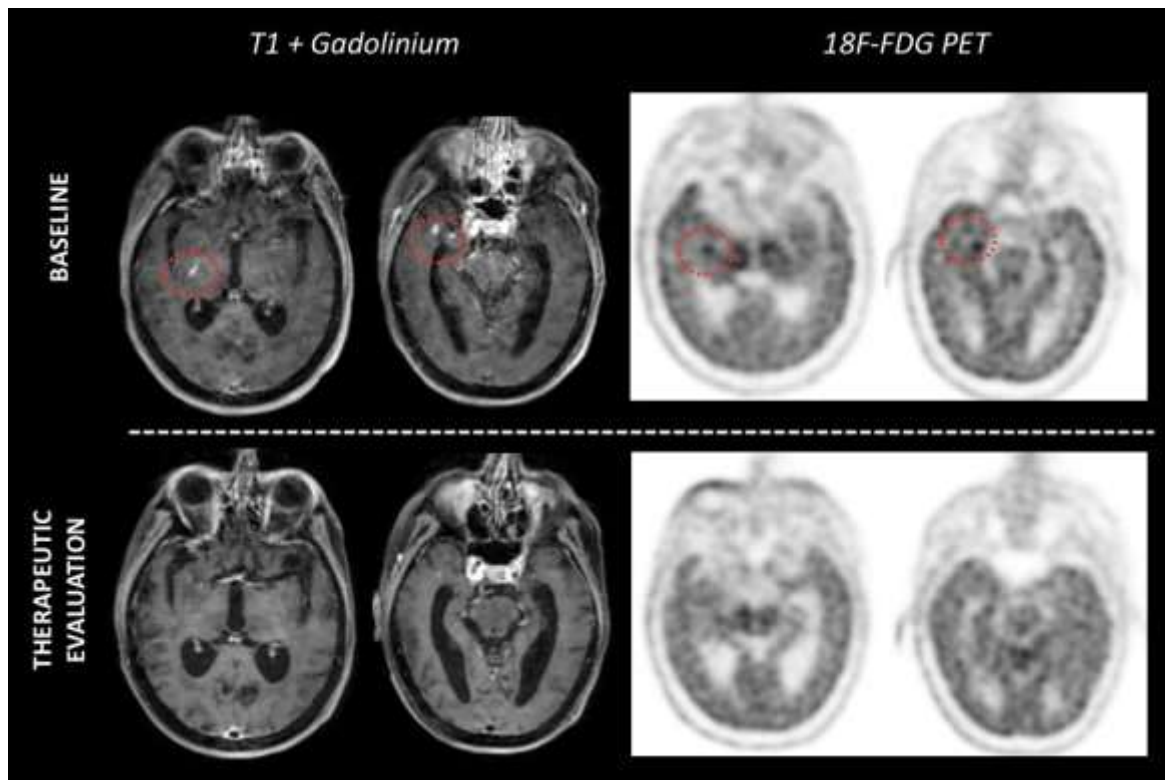


Figure 4. Example case of PET-MRI images of a lymphomatosis cerebri before and after Methotrexate therapy. Baseline images (**top line**) showed small nodular enhanced lesions anterior to the temporal horn of the right ventricle and a linear contrast-enhanced lesion at the posterior arm of the right internal capsule. ^{18}F -FDG PET images revealed strong uptake consistent with these lesions. After three cycles of methotrexate treatment (**bottom line**), MRI images showed a partial response, while ^{18}F -FDG PET images had already normalized.

However, the MRI presentation of HGG can mimic that of PCNSL and vice versa: a “homogeneous pattern” with a “butterfly pattern through the corpus callosum” can be observed in true HGG whereas confirmed PCNSL in IC patients may show annular enhancement with central necrosis, normally seen in HGG [12,16]. Literature has shown good diagnostic performance of ^{18}F -FDG PET-CT in this setting, as PCNSL is known to have higher ^{18}F -FDG uptake than HGG [34,35]. Some studies have proposed SUV_{\max} cut-off ranging from 12 to 15 to confirm the diagnosis of PCNSL [27,36]. A recent study of 65 immunocompetent patients confirmed that T/ N was significantly lower in HGG than in PCNSL. Sensitivity and specificity of T/ N vary respectively between 90–100% and 75–87%, depending on the threshold used by the different teams [34,37,38]. In contrast, no added value was reported for metabolic tumor volume or total lesion glycolysis [37]. In addition, the combination of high T/ N (cutoff value of 2.23) on ^{18}F -FDG PET and low maximal tumor blood flow on ASL (cutoff value of 2.07) has been described to improve discrimination between PCNSL and HGG (accuracy 99.1%, sensitivity 95%, and specificity 96.4%) [34].

To further improve radiological diagnosis, machine learning has recently emerged as an important tool. Petersen et al. conducted a meta-analysis of 23 papers, including 3 papers on ^{18}F -FDG, that published machine learning-based classification algorithms to distinguish between HGG and PCNSL [39]. The machine learning models appear to have high accuracy as the algorithms were able to replicate the results of a senior subspecialty-trained radiologist. Using radiomic approach, Kong et al. recently reported AUC as high as 0.971–0.998 for the selected features [40].

In IS patients and especially in HIV-infected patients, CNS lymphoma and toxoplasmosis are the two most common brain disease. While the distinction between these two entities is critical because of the dramatic change in patient management, their MRI features may overlap. Although considered the gold standard, histopathologic confirmation is rarely available and diagnosis relies on an anti-toxoplasma treatment test that may delay initiation of adequate chemotherapy in true CNS lymphoma. Single-photon emission computed tomography (SPECT) using thallium-201 has historically been used for differential diagnosis between tumor and non-neoplastic lesions, because thallium-201 accumulates in actively dividing cells via a transport promoted by an adenosine triphosphate cell membrane active pump but not in infectious process [41]. In a meta-analysis of 667 patients from 18 studies, Yang et al. found a respective pooled sensitivity and specificity for thallium-201 SPECT of 92% and 84% [42].

^{18}F -FDG has also been extensively evaluated [43–48]. All studies revealed high accuracy of this technique with sensitivities close to 100% and specificities ranging from 75% to 100% i.e., high metabolic activity for CNS lymphoma versus normal to reduced intralésional metabolic activity relative to adjacent brain parenchyma for cerebral toxoplasma [43–45,48] (Figure 5).

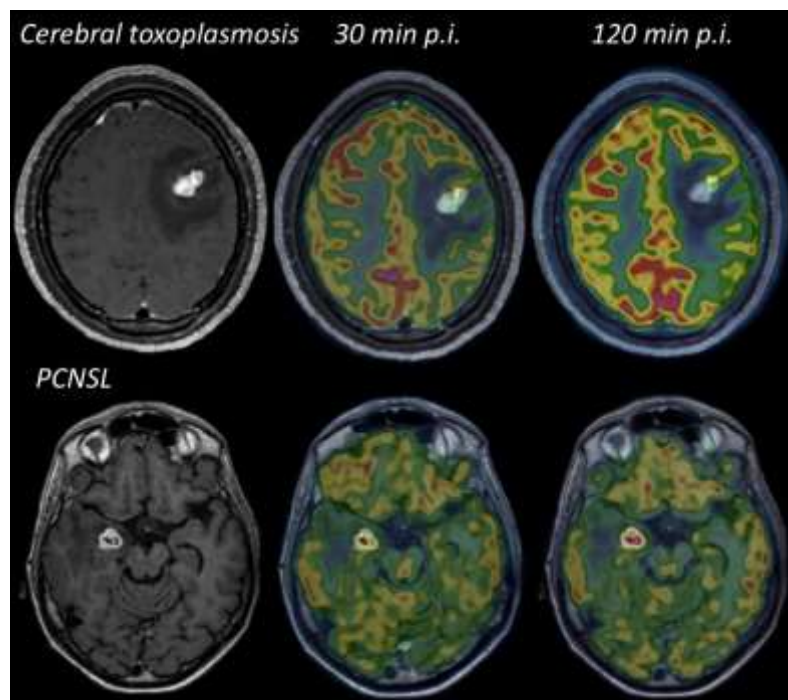


Figure 5. Illustrative case of the value of delayed acquisition of ^{18}F -FDG PET to differentiate infection from malignancy in an IS patient. The first patient was a 19-year-old female with AIDS disease, while the second patient was undergoing immunosuppressive therapy for renal transplantation. The ^{18}F -FDG PET scan was performed for both patients at 30 and 120 min after injection (p.i.). In the first case, tracer uptake was very low at 30 min p.i. and decreased with time (SUVmax = 4.7 and 3.2; T/ N = 0.3 and 0.2). The patient was diagnosed with cerebral toxoplasmosis by serological evaluation. In the second case, the lesion showed a progressive increase in metabolic activity over time compared with normal tissue (SUVmax = 5.9 and 11.2; T/ N = 1.5 and 1.7) and brain biopsy confirmed brain lymphoma.

In summary, while ^{18}F -FDG brain PET appears to be highly reliable for the differential diagnosis of brain lesions in IS patients, its potential role in IC patients needs to be further analyzed, especially in the new era of artificial intelligence algorithm applied to PET.

Table 1. MRI and PET characteristics of PCNSL and its main differential diagnoses [16,49–53].

Biophysical Features	MRI						PET			
	T1	T1+Gd	T2W/FLAIR	DWI	SWI Detection of Hemosiderin, Ferritin and Calcium	ADC Cellular Restriction	Spectroscopy Detection of Brain Metabolites	rCBV Microvascular Blood Volume	¹⁸ F-FDG Glycemic metabolism	¹¹ C-MET Amino-Acid Analog
PCNSL IC	Iso or hypointense Uni or multifocal Profound or periventricular	Homogeneously enhancing parenchymal mass	Tumor is in hyposignal compared to the hypersignal of the adjacent edema FLAIR	Hyperintense	Rare microhemorrhage and calcification	Very low (lower than GBM)	Cho/Cr ↗ Lipid ↗ NAA ↘ (less marked than in glioblastoma)	Low typical perfusion curve returning above the baseline	High uptake	High uptake
Glioblastoma	Hypo to isointense Central hemispheric white matter	Heterogeneous enhancement with central necrosis	Hyperintense surrounded by vasogenic edema	Hyperintense in the solid portion	Frequent small hemorrhages	Low	Cho/Cr ↗ Lactate ↗ Lipids ↗ NAA ↘	High	High uptake (but less pronounced than in PSCNSL)	High uptake
PCNSL ID	Iso or hypointense Multifocal lesions Basal ganglia and corpus callosum	Nodular, or ringlike patterns	Hypointense to slightly hyperintense Small amount of edema and mass effect	Variable	Spontaneous hemorrhage more frequent than for PCNSL IC	Usually rADC <1.6 but overlapping ratio with topoplasmia lesions	Choline ↗ NAA ↘	Low	High uptake	High uptake
Toxoplasmosis	Iso to hypointense Multifocal Basal ganglia, corticomedullary junction of the cerebral hemispheres	Ringlike or nodular enhancement patterns	Low hypointense to hyperintense Large edema and mass effect	Variable	Occasional hemorrhages	Usually rADC >1.6 but overlapping ratio with PCNSL	Choline mild ↗ NAA ↘ Lactate ↗ Lipids ↗	Low	Hypometabolic	Low to high uptake
Progressive Multifocal Leukoencephalopathy	Hypointense Multifocal and asymmetric involvement Subcortical white matter and centrum semi ovale	Non enhancing or, rarely, mildly enhancing	Hyperintense Multiple small punctate lesions outside the main PML lesions	Central core with low signal surrounded by a rim of high signal intensity	Typical leukocortical band/rim	Variable/Low ADC reflects active lesions	Choline ↗ NAA ↘ Lactate ↗ Lipids ↗	Low	Low uptake	Low to high uptake

Gd: gadolinium, FLAIR: Fluid-attenuated inversion recovery, DWI: diffusion-weighted images, SWI: susceptibility-weighted images, ADC: apparent diffusion coefficient, rCBV: relative cerebral blood volume, For spectroscopy: choline (Cho): marker of cell membrane synthesis, lactate: anaerobic glycolysis, N-acetylaspartate (NAA): neuronal viability, Creatinine (Cr): cell density, Lipids: cell necrosis, PML: progressive multifocal leukoencephalopathy.

2.2.2. Outcome Prediction

PCNSL has poor prognosis. To date, the only established prognostic factors are age and functional status at diagnosis [9]. For therapeutic trials, two multipoint scoring scales including clinico-biological factors have been proposed, the International Extranodal Lymphoma Study Group (IELSG) scale, and the Memorial Sloan-Kettering Cancer Center (MSKCC) scale [54,55]. Recently, other biological factors or molecular characteristics of tumors have been identified. For example, it has been reported that high protein levels of interleukin (IL)-10 at diagnosis or co-expression of MYC/BCL2 in tumors correlate with low overall survival (OS) and progression-free survival (PFS) [56,57]. In addition, many prognostic factors have been proposed in MRI, including multifocal lesions, high tumor volume, or subtentorial involvement, but with controversial results [58,59]. Some also analyzed more advanced parameters, i.e., diffusion and perfusion MRI sequences, with promising results for ADC parameters or blood plasma volume (Vp) and volume transfer constant (Ktrans) [60,61]. However, none of these results have been corroborated by studies with large patient cohorts.

As mentioned previously, CNS lymphoma is characterized by high ^{18}F -FDG uptake. Several studies have explored the role of intensity of ^{18}F -FDG uptake or hypermetabolic volume on pretreatment PET, as prognostic markers. Varied indexes have been proposed but they have not been always corroborated (Table 2). A poorer survival was associated with high SUVmax (>12) in a series of 17 IC patients [62]. Kasenda et al. developed a visual scale based on the tumor SUVmax normalized to the normal cerebellar uptake (MILAS for Metabolic Imaging Lymphoma Aggressiveness Scale). They found a negative correlation between increased MILAS (greater than 3 corresponding i.e., ^{18}F -FDG uptake of the tumor 200% higher than that of the cerebellum) and PFS (54.7 vs. 3.8 months) [63]. Other studies have emphasized the potential prognostic role of total lesion glycolysis (TLG) or metabolic tumor volume (MTV) [64,65]. For example, PFS and OS were significantly shorter in patients with $\text{MTV} \geq 9.8 \text{ cm}^3$ and $\text{TLG} \geq 94 \text{ g/mL} \times \text{cm}^3$ in a cohort of 52 patients [65]. More recently, Krebs et al. suggest that adding SUVmax values and volumes of all lesions (up to five) may predict response at the end of treatment and are related to the risk of disease progression or death [66]. In conclusion, high ^{18}F -FDG uptake and large metabolic tumor volume seem to be good predictors of poor prognosis; however due to the lack of consensus and the limited number of patients in each study, neither of these parameters is yet applicable in daily routine.

Table 2. ^{18}F -FDG PET Biomarkers of poor PFS.

Biomarkers	Threshold	Analysis	HR	95% CI	<i>p</i>	REF
SUVmax	≥ 12	Univariate			<0.05	Kawai [62]
MILAS	>3 i.e.	Multivariate	1.46	1.10–1.94	0.010	Kasenda [63]
MTV	$\geq 9.8 \text{ cm}^3$	Multivariate	5.35	1.89–12.8	0.037	Albano [65]
TLG	≥ 94	Multivariate	4.54	1.37–11.6	0.045	Albano [65]
	>227	Univariate			0.02	Okoyucu [64]
Sum SUVmax	-	Multivariate	1.09	1.04–1.14	<0.001	Krebs [66]

3. Role of ^{18}F -FDG PET in the Management of PCNSL

3.1. Systemic Assessment

Identification of systemic disease in suspected brain lymphoma is pivotal to treatment decisions, as chemotherapy regimens vary between systemic lymphoma with CNS involvement and PCNSL, which requires additional methotrexate therapy. Historically, staging evaluation included CT scan, bone marrow biopsy, and testicular ultrasound in men [67].

First studies showed low additional value of body ^{18}F -FDG PET to detect systemic involvement, found in only 2% of patients [68]. With the emergence of new PET-CT sys-

tems, literature has shown that body PET could detect systemic disease in up to 10% of patients [25,69]. Several studies have reported superior diagnostic performance of whole-body ^{18}F -FDG PET-CT compared to whole-body CT (4.9% vs. 2.5%) [31,68–70]. Advantage of PET is to identify malignant foci outside the “chest-abdomen-pelvis” (CAP) area, investigated by CT [71] (Figure 6). A recent meta-analysis on 1040 patients, reported systemic lymphoma lesions in upper and lower limbs bones, lymph nodes, adrenal gland, soft tissue, spleen, liver, jejunum, testicle, thyroid gland, and nose [72]. Although the increase in false-positive ^{18}F -FDG PET results has been blamed for unnecessary interventions, PET also enables detection of incidental secondary malignancies (1.5% to 3.1%) with significant therapeutic impact [72].

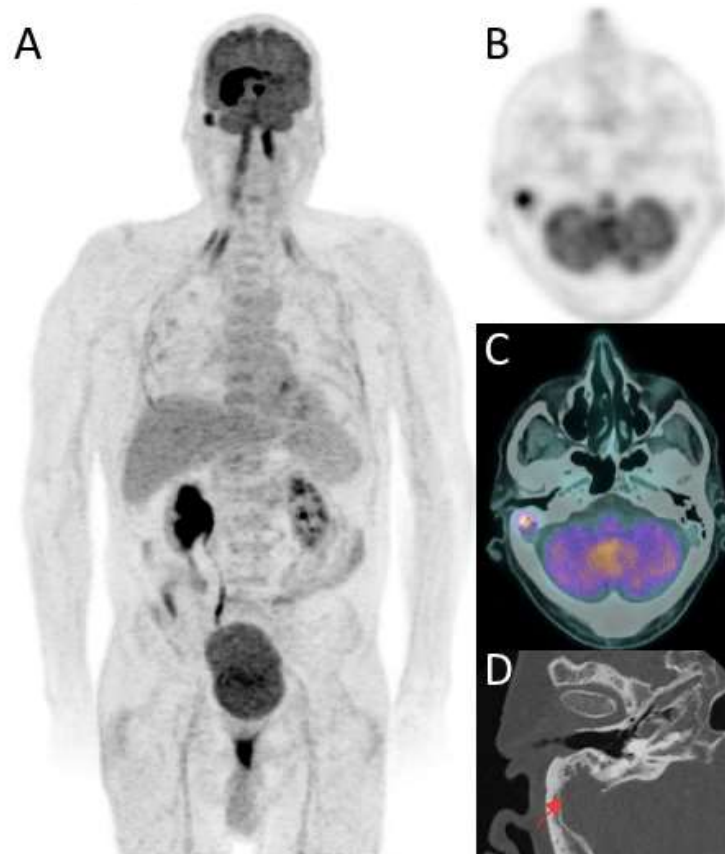


Figure 6. ^{18}F -FDG PET/CT reveals systemic involvement in the right mastoiditis in a patient with right temporal and paraventricular brain lesion highly avid for ^{18}F -FDG. **A:** Maximum intensity projection (MIP) of ^{18}F -FDG PET images. PET images showed an hypermetabolic right mastoid process filling: SUVmax at 8.8 (**B:** axial PET images, **C:** axial PET/CT fusion). Imaging was completed by a dedicated thin-section CT scan which revealed irregular focal bone lysis of the right mastoid (**D**). Histological examination confirmed a B-cell lymphoma proliferation.

Today, the role of whole-body ^{18}F -FDG-PET in the initial diagnostic of PCNSL to rule out systemic involvement has been set in stone by several expert societies and PET should be prescribed as first-line when its realization does not compromise prompt disease management [10,11,67,73].

3.2. Therapeutic Evaluation

Over the past decades, several therapeutic trials have proposed new treatment regimens for patients with PCNSL [74]. Currently, multidrug therapy including high-dose methotrexate is a standard of care and is combined with whole-brain radiotherapy or autol-

ogous stem-cell transplantation in young patients. Because of the high risk of neurotoxicity, whole brain radiotherapy is no longer recommended for patients older than 60 years [75].

Today, treatment follow-up is based on the International PCNSL Collaborative Group (IPCG) radiographic response criteria, which recommends gadolinium-enhanced MRI performed approximately every 2 months during active treatment [67]. However, this classification includes an unconfirmed complete response, because of the inability of MRI to distinguish residual tumor nidus from biopsy-related scar tissue [67,76]. In addition, the early onset of recurrence after treatment in patients with a complete response on MRI raises the question of the performance of MRI in assessing residual disease. Recently, Van der Meulen et al. showed that there were no significant differences in PFS and OS between patients with a complete or partial MRI response at the end of protocol treatment in a phase III randomized controlled trial of 199 patients with PCNSL [77].

While the role of interim ^{18}F -FDG PET has been well established in aggressive systemic lymphomas, few studies have examined its value in PCNSL. Birsen ($n = 25$ patients), Jo ($n = 66$), and Palmedo ($n = 8$) et al. agree on the prognostic impact of interim PET on the end-of-treatment response with a high negative predictive value (94.74% to 100%) [78–80]. In addition, they showed that PET was an earlier biomarker of response than morphological MRI. However, it is still unclear when PET should be performed to stratify patients according to their prognosis between interim and final evaluation [78–82]. Further prospective trials in large cohorts are therefore needed to establish PET in the therapeutic management of patients with PCNSL. A prospective study exploring ^{18}F -FDG PET in treatment responses in CNS lymphoma is currently underway in France (NCT03582254) to address the role of ^{18}F -FDG PET in the management of CNS lymphoma during treatment (Figure 7).

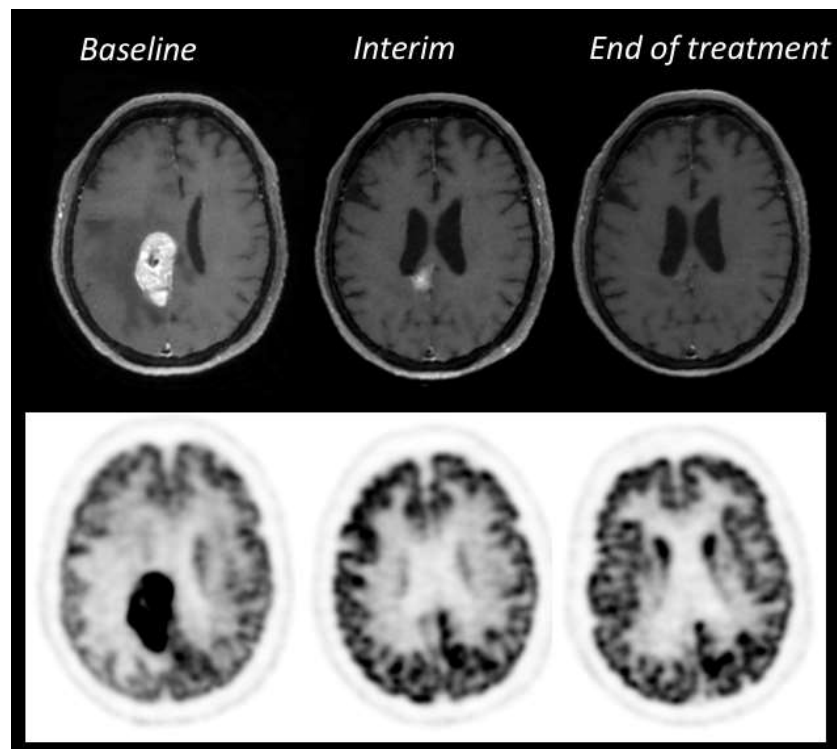


Figure 7. Therapeutic ^{18}F -FDG PET-MRI evaluation of PCNSL in an immunocompetent patient treated with R-MPVA chemotherapy (Rituximab, Methotrexate, Procarbazine, Vincristine, Cytarabine). Two months after the beginning of the treatment, MRI showed a partial response (81% decrease of the enhanced MRI volume on 3DT1-gadolinium sequence) while PET was already negative. Follow-up showed a complete response on both MRI and PET at the end of treatment and the patient was disease free after 3 years. This case illustrates the potential role of interim PET as an early prognostic biomarker.

Recently, the role of immune evasion and suppression of the immune microenvironment has been revealed as a key element of the pathogenesis of PCNSL. Therefore, new therapeutic trials using immunotherapy agents are currently underway [83]. On MRI, these agents may be responsible for an effect called pseudoprogression, related to changes in BBB permeability: contrast enhancement may increase or even appear under treatment while there is no real tumor progression, and follow-up shows a normalization of these images without therapeutic changes [11]. While these phenomena have been extensively studied in glial pathologies on PET, no study has yet reported relevant elements on this subject, to our knowledge in PCNSL (Figure 8).

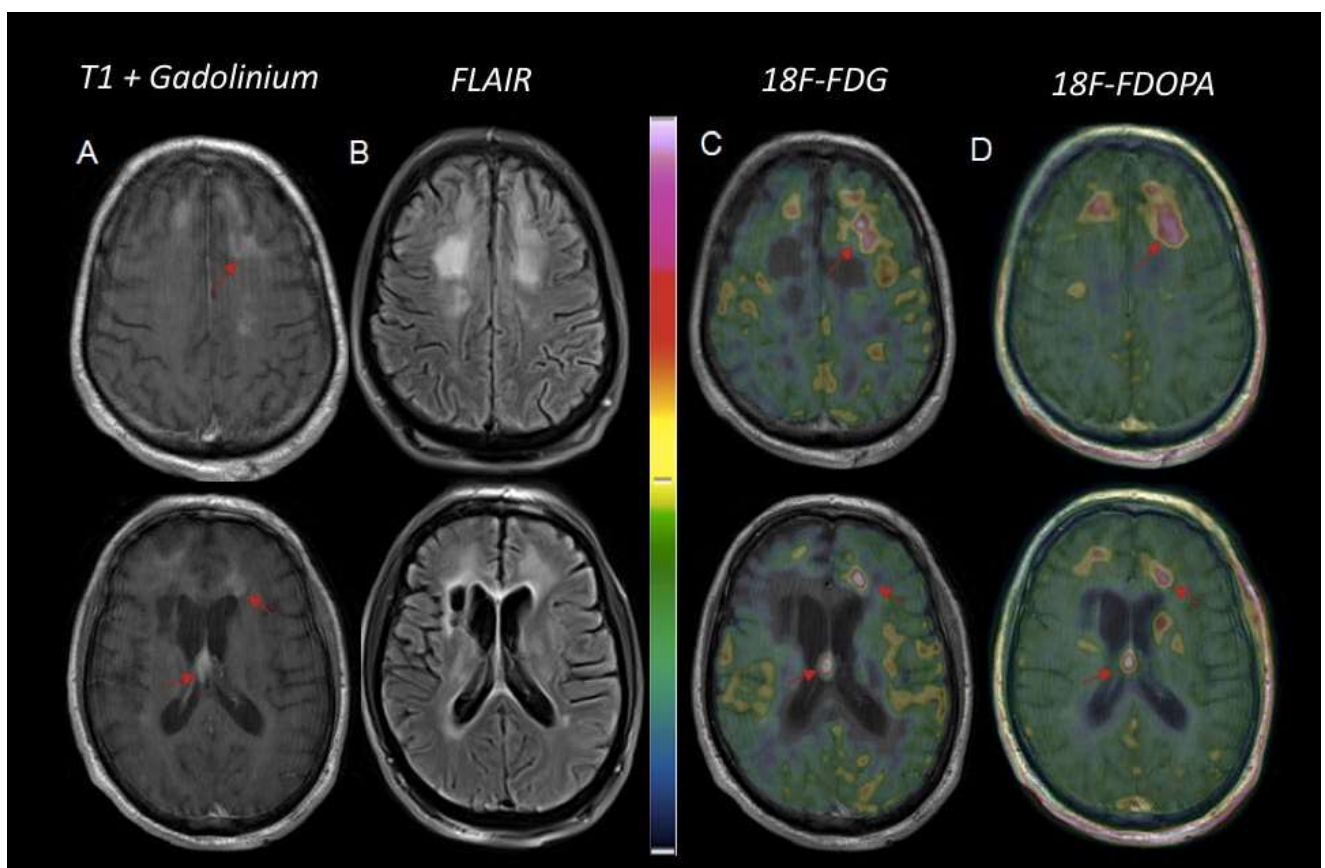


Figure 8. Illustrative case of suspicion of pseudoprogression versus progression in a 57-year-old man with IC PCNSL treated with immunotherapy (Pembrolizumab). MRI images showed suspicious bi-frontal and interventricular contrast-enhanced lesions on post gadolinium T1-weighted MRI, (A) and a T2-FLAIR hypersignal (B). Two enhanced-lesions located in the septum pellucidum and in front of the left frontal horn of the lateral ventricle, extending through the surrounding white matter, were highly hypermetabolic on ^{18}F -FDG PET (C: axial PET) and showed strong uptake on ^{18}F -FDOPA PET (D: axial PET): tumoral to normal contralateral tissue ratio (T/ N) = 3.9 and 1.9 respectively, with ^{18}F -FDG versus 2.7 and 2.2 with ^{18}F -FDOPA. A third enhanced-lesion in the right frontal lobe showed milder uptake both on ^{18}F -FDG and ^{18}F -FDOPA PET (T/ N = 1.5 and 1.8, respectively), but was considered suspicious due to its localization in white matter. Follow-up MRI confirmed the diagnosis of PCNSL relapse.

4. Others PET Tracers in PCNSL

Others PET tracers have been evaluated for characterization and therapeutic evaluation of PCNSL. The most analyzed is ^{11}C -Methionine (^{11}C -MET), an amino acid tracer correlated with cell proliferation activity, which benefits from a low uptake in normal brain compared to ^{18}F -FDG. In pre-therapeutic evaluation, the literature has shown good

agreement between ^{18}F -FDG and methionine tracers but without additional diagnostic or prognostic value for the latter [84,85]. ^{11}C -MET could also be useful for early detection of non-responders, as it has been described that a high uptake level at the interim assessment (after four cycles of methotrexate) was associated with decreased free survival [86]. Few studies on small cohort have suggested a potential additional role of ^{11}C -MET, for example to stratify patients according to the immunohistochemical subtypes or for differential diagnostic between CNSL and HGG based on kinetic analysis [87,88]. However, the main limitation of ^{11}C -MET is the need for an on-site cyclotron due to the short 20-min half-life of this radiotracer.

In addition to amino acid analogues, other PET tracers studied for CNS lymphomas are ^{68}Ga -Pentixafor and ^{18}F -Fludarabine. ^{68}Ga -Pentixafor specifically targets CXCR4 receptors that are overexpressed by lymphoma, leukemia, and myeloma [89]. In two small pilot studies, Starzer et al. ($n = 7$) and Herhaus et al. ($n = 8$) showed 100% accuracy for detection of PCNSL between MRI and ^{68}Ga -Pentixafor, with high contrast between lesions and normal brain background uptake. In addition, lesions with lower uptake of the CXCR4 tracer were associated with a better response to standard methotrexate-based therapy, suggesting a potential role for this new tracer for therapeutic follow-up applications [89,90]. However, the additional value of ^{68}Ga -Pentixafor remains questionable, especially for atypical PCNSL, because this tracer is dependent on the breakdown of the blood-brain barrier. In preclinical studies, Hovanishyan et al. showed that ^{18}F -Fludarabine (^{18}F -FDB), a PET radiotracer version of the drug fludarabine that binds equilibrium nucleoside transporters (ENT1), could distinguish glioblastoma from CNS lymphomas, without being biased by activated macrophages and other inflammatory cells [91]. ^{18}F -FDB showed high uptake in CNS lymphomas, while the tracer was rapidly cleared from glioma cells. These results have recently been confirmed in a first-in man study using PET with ^{18}F -FDB [92]. Early results on ^{18}F -FDB also showed that the tracer uptake could be independent of BBB leakage, which should allow metabolic characterization of atypical PCNSL [91,93].

5. Conclusions

^{18}F -FDG is the most widely used radiotracer in this disease because of its high uptake by lymphoid cells and its easy and inexpensive access.

Its role in systemic assessment at initial diagnosis has been widely established by expert society over the past decade. Several studies have evaluated its prognostic role during follow-up, but at present, the available data are from small, mostly retrospective series. Furthermore, in the new era of machine learning, future biomarkers are likely to be a combination of parameters from different imaging techniques, and nuclear physicians and neuroradiologists will need to cooperate closely to have a real impact on clinicians' management decisions.

In addition, several other radiotracers have been proposed, including ^{11}C -MET, ^{68}Ga -Pentixafor, and ^{18}F -Fludarabine. Further prospective studies are needed to confirm their potential role in the management of PCNSL.

Author Contributions: Conceptualization, L.R. and A.K.; methodology, L.R.; software, L.R. and C.H.; validation, A.K., C.H. and D.G.; formal analysis, L.R.; investigation, L.R.; resources, L.R. and C.H.; data curation, L.R.; writing—original draft preparation, L.R.; writing—review and editing, A.K., C.S., D.G., K.H.-X., S.C., C.H., M.B.; visualization, A.K.; supervision, A.K.; project administration, A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Miller, D.C.; Hochberg, F.H.; Harris, N.L.; Gruber, M.L.; Louis, D.N.; Cohen, H. Pathology with Clinical Correlations of Primary Central Nervous System Non-Hodgkin's Lymphoma. The Massachusetts General Hospital Experience 1958–1989. *Cancer* **1994**, *74*, 1383–1397. [[CrossRef](#)]
2. Camilleri-Broët, S.; Martin, A.; Moreau, A.; Angonin, R.; Hénin, D.; Gontier, M.F.; Rousselet, M.C.; Caulet-Maugendre, S.; Cuillière, P.; Lefrancq, T.; et al. Primary Central Nervous System Lymphomas in 72 Immunocompetent Patients: Pathologic Findings and Clinical Correlations. Groupe Ouest Est d'étude Des Leucémies et Autres Maladies Du Sang (GOELAMS). *Am. J. Clin. Pathol.* **1998**, *110*, 607–612. [[CrossRef](#)] [[PubMed](#)]
3. Grogg, K.L.; Miller, R.F.; Dogan, A. HIV Infection and Lymphoma. *J. Clin. Pathol.* **2007**, *60*, 1365–1372. [[CrossRef](#)] [[PubMed](#)]
4. Mendez, J.S.; Ostrom, Q.T.; Gittleman, H.; Kruchko, C.; DeAngelis, L.M.; Barnholtz-Sloan, J.S.; Grommes, C. The Elderly Left Behind—Changes in Survival Trends of Primary Central Nervous System Lymphoma over the Past 4 Decades. *Neuro-Oncology* **2018**, *20*, 687–694. [[CrossRef](#)] [[PubMed](#)]
5. Villano, J.L.; Koshy, M.; Shaikh, H.; Dolecek, T.A.; McCarthy, B.J. Age, Gender, and Racial Differences in Incidence and Survival in Primary CNS Lymphoma. *Br. J. Cancer* **2011**, *105*, 1414–1418. [[CrossRef](#)] [[PubMed](#)]
6. Ostrom, Q.T.; Cioffi, G.; Gittleman, H.; Patil, N.; Waite, K.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012–2016. *Neuro-Oncology* **2019**, *21*, v1–v100. [[CrossRef](#)] [[PubMed](#)]
7. Bhagavathi, S.; Wilson, J.D. Primary Central Nervous System Lymphoma. *Arch. Pathol. Lab. Med.* **2008**, *132*, 1830–1834. [[CrossRef](#)]
8. Morales-Martinez, A.; Lozano-Sanchez, F.; Duran-Peña, A.; Hoang-Xuan, K.; Houillier, C. Primary Central Nervous System Lymphoma in Elderly Patients: Management and Perspectives. *Cancers* **2021**, *13*, 3479. [[CrossRef](#)]
9. Houillier, C.; Soussain, C.; Ghesquière, H.; Soubeyran, P.; Chinot, O.; Taillandier, L.; Lamy, T.; Choquet, S.; Ahle, G.; Damaj, G.; et al. Management and Outcome of Primary CNS Lymphoma in the Modern Era: An LOC Network Study. *Neurology* **2020**, *94*, e1027–e1039. [[CrossRef](#)]
10. Hoang-Xuan, K.; Bessell, E.; Bromberg, J.; Hottinger, A.F.; Preusser, M.; Rudà, R.; Schlegel, U.; Siegal, T.; Soussain, C.; Abacioglu, U.; et al. Diagnosis and Treatment of Primary CNS Lymphoma in Immunocompetent Patients: Guidelines from the European Association for Neuro-Oncology. *Lancet Oncol.* **2015**, *16*, e322–e332. [[CrossRef](#)]
11. Barajas, R.F.; Politi, L.S.; Anzalone, N.; Schöder, H.; Fox, C.P.; Boxerman, J.L.; Kaufmann, T.J.; Quarles, C.C.; Ellingson, B.M.; Auer, D.; et al. Consensus Recommendations for MRI and PET Imaging of Primary Central Nervous System Lymphoma: Guideline Statement from the International Primary CNS Lymphoma Collaborative Group (IPCG). *Neuro-Oncology* **2021**, *23*, 1056–1071. [[CrossRef](#)] [[PubMed](#)]
12. Chiavazza, C.; Pellerino, A.; Ferrio, F.; Cistaro, A.; Soffietti, R.; Rudà, R. Primary CNS Lymphomas: Challenges in Diagnosis and Monitoring. *BioMed Res. Int.* **2018**, *2018*, 3606970. [[CrossRef](#)] [[PubMed](#)]
13. Hartmann, M.; Heiland, S.; Harting, I.; Tronnier, V.M.; Sommer, C.; Ludwig, R.; Sartor, K. Distinguishing of Primary Cerebral Lymphoma from High-Grade Glioma with Perfusion-Weighted Magnetic Resonance Imaging. *Neurosci. Lett.* **2003**, *338*, 119–122. [[CrossRef](#)]
14. Fink, J.R.; Muzi, M.; Peck, M.; Krohn, K.A. Multimodality Brain Tumor Imaging: MR Imaging, PET, and PET/MR Imaging. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **2015**, *56*, 1554–1561. [[CrossRef](#)]
15. Yamasaki, F.; Takayasu, T.; Nosaka, R.; Amatya, V.J.; Doskaliyev, A.; Akiyama, Y.; Tominaga, A.; Takeshima, Y.; Sugiyama, K.; Kurisu, K. Magnetic Resonance Spectroscopy Detection of High Lipid Levels in Intraaxial Tumors without Central Necrosis: A Characteristic of Malignant Lymphoma. *J. Neurosurg.* **2015**, *122*, 1370–1379. [[CrossRef](#)]
16. Haldorsen, I.S.; Kråkenes, J.; Krossnes, B.K.; Mella, O.; Espeland, A. CT and MR Imaging Features of Primary Central Nervous System Lymphoma in Norway, 1989–2003. *AJNR Am. J. Neuroradiol.* **2009**, *30*, 744–751. [[CrossRef](#)]
17. Küker, W.; Nägele, T.; Korfel, A.; Heckl, S.; Thiel, E.; Bamberg, M.; Weller, M.; Herrlinger, U. Primary Central Nervous System Lymphomas (PCNSL): MRI Features at Presentation in 100 Patients. *J. Neurooncol.* **2005**, *72*, 169–177. [[CrossRef](#)]
18. Terae, S.; Ogata, A. Nonenhancing Primary Central Nervous System Lymphoma. *Neuroradiology* **1996**, *38*, 34–37. [[CrossRef](#)]
19. Nishiyama, Y.; Yamamoto, Y.; Monden, T.; Sasakawa, Y.; Kawai, N.; Satoh, K.; Ohkawa, M. Diagnostic Value of Kinetic Analysis Using Dynamic FDG PET in Immunocompetent Patients with Primary CNS Lymphoma. *Eur. J. Nucl. Med. Mol. Imaging* **2007**, *34*, 78–86. [[CrossRef](#)]
20. Bataille, B.; Delwail, V.; Menet, E.; Vandermarcq, P.; Ingrand, P.; Wager, M.; Guy, G.; Lapierre, F. Primary Intracerebral Malignant Lymphoma: Report of 248 Cases. *J. Neurosurg.* **2000**, *92*, 261–266. [[CrossRef](#)]
21. Krebs, S.; Barasch, J.G.; Young, R.J.; Grommes, C.; Schöder, H. Positron Emission Tomography and Magnetic Resonance Imaging in Primary Central Nervous System Lymphoma—a Narrative Review. *Ann. Lymphoma* **2021**, *5*, 15. [[CrossRef](#)]
22. Brandsma, D.; Bromberg, J.E.C. Primary CNS Lymphoma in HIV Infection. *Handb. Clin. Neurol.* **2018**, *152*, 177–186. [[CrossRef](#)]
23. Kawai, N.; Nishiyama, Y.; Miyake, K.; Tamiya, T.; Nagao, S. Evaluation of Tumor FDG Transport and Metabolism in Primary Central Nervous System Lymphoma Using [18F]Fluorodeoxyglucose (FDG) Positron Emission Tomography (PET) Kinetic Analysis. *Ann. Nucl. Med.* **2005**, *19*, 685–690. [[CrossRef](#)]

24. Gupta, T.; Manjali, J.J.; Kannan, S.; Purandare, N.; Rangarajan, V. Diagnostic Performance of Pretreatment 18F-Fluorodeoxyglucose Positron Emission Tomography With or Without Computed Tomography in Patients With Primary Central Nervous System Lymphoma: Updated Systematic Review and Diagnostic Test Accuracy Meta-Analyses. *Clin. Lymphoma Myeloma Leuk.* **2021**, *21*, 497–507. [[CrossRef](#)]
25. Gupta, M.; Gupta, T.; Purandare, N.; Rangarajan, V.; Puranik, A.; Moiyadi, A.; Shetty, P.; Epari, S.; Sahay, A.; Mahajan, A.; et al. Utility of Fluoro-Deoxy-Glucose Positron Emission Tomography/Computed Tomography in the Diagnostic and Staging Evaluation of Patients with Primary CNS Lymphoma. *CNS Oncol.* **2019**, *8*, CNS46. [[CrossRef](#)]
26. Roelcke, U.; Leenders, K.L. Positron Emission Tomography in Patients with Primary CNS Lymphomas. *J. Neurooncol.* **1999**, *43*, 231–236. [[CrossRef](#)]
27. Kosaka, N.; Tsuchida, T.; Uematsu, H.; Kimura, H.; Okazawa, H.; Itoh, H. 18F-FDG PET of Common Enhancing Malignant Brain Tumors. *AJR Am. J. Roentgenol.* **2008**, *190*, W365–W369. [[CrossRef](#)]
28. Rosenfeld, S.S.; Hoffman, J.M.; Coleman, R.E.; Glantz, M.J.; Hanson, M.W.; Schold, S.C. Studies of Primary Central Nervous System Lymphoma with Fluorine-18-Fluorodeoxyglucose Positron Emission Tomography. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **1992**, *33*, 532–536.
29. Baraniskin, A.; Deckert, M.; Schulte-Altendorneburg, G.; Schlegel, U.; Schroers, R. Current Strategies in the Diagnosis of Diffuse Large B-Cell Lymphoma of the Central Nervous System. *Br. J. Haematol.* **2012**, *156*, 421–432. [[CrossRef](#)]
30. Yamaguchi, S.; Hirata, K.; Kobayashi, H.; Shiga, T.; Manabe, O.; Kobayashi, K.; Motegi, H.; Terasaka, S.; Houkin, K. The Diagnostic Role of (18)F-FDG PET for Primary Central Nervous System Lymphoma. *Ann. Nucl. Med.* **2014**, *28*, 603–609. [[CrossRef](#)]
31. Albano, D.; Bosio, G.; Bertoli, M.; Giubbini, R.; Bertagna, F. 18F-FDG PET/CT in Primary Brain Lymphoma. *J. Neurooncol.* **2018**, *136*, 577–583. [[CrossRef](#)]
32. Kawai, N.; Okubo, S.; Miyake, K.; Maeda, Y.; Yamamoto, Y.; Nishiyama, Y.; Tamiya, T. Use of PET in the Diagnosis of Primary CNS Lymphoma in Patients with Atypical MR Findings. *Ann. Nucl. Med.* **2010**, *24*, 335–343. [[CrossRef](#)]
33. Kim, H.O.; Kim, J.S.; Kim, S.-O.; Chae, S.Y.; Oh, S.J.; Seo, M.; Lee, S.H.; Oh, J.S.; Ryu, J.-S.; Huh, J.-R.; et al. Clinicopathological Characteristics of Primary Central Nervous System Lymphoma with Low 18F-Fluorodeoxyglucose Uptake on Brain Positron Emission Tomography. *Medicine* **2020**, *99*, e20140. [[CrossRef](#)]
34. Hatakeyama, J.; Ono, T.; Takahashi, M.; Oda, M.; Shimizu, H. Differentiating between Primary Central Nervous System Lymphoma and Glioblastoma: The Diagnostic Value of Combining 18F-Fluorodeoxyglucose Positron Emission Tomography with Arterial Spin Labeling. *Neurol. Med. Chir.* **2021**, *61*, 367–375. [[CrossRef](#)]
35. Yamashita, K.; Yoshiura, T.; Hiwatashi, A.; Togao, O.; Yoshimoto, K.; Suzuki, S.O.; Abe, K.; Kikuchi, K.; Maruoka, Y.; Mizoguchi, M.; et al. Differentiating Primary CNS Lymphoma from Glioblastoma Multiforme: Assessment Using Arterial Spin Labeling, Diffusion-Weighted Imaging, and ¹⁸F-Fluorodeoxyglucose Positron Emission Tomography. *Neuroradiology* **2013**, *55*, 135–143. [[CrossRef](#)]
36. Makino, K.; Hirai, T.; Nakamura, H.; Murakami, R.; Kitajima, M.; Shigematsu, Y.; Nakashima, R.; Shiraiishi, S.; Uetani, H.; Iwashita, K.; et al. Does Adding FDG-PET to MRI Improve the Differentiation between Primary Cerebral Lymphoma and Glioblastoma? Observer Performance Study. *Ann. Nucl. Med.* **2011**, *25*, 432–438. [[CrossRef](#)]
37. Zhou, W.; Wen, J.; Hua, F.; Xu, W.; Lu, X.; Yin, B.; Geng, D.; Guan, Y. 18F-FDG PET/CT in Immunocompetent Patients with Primary Central Nervous System Lymphoma: Differentiation from Glioblastoma and Correlation with DWI. *Eur. J. Radiol.* **2018**, *104*, 26–32. [[CrossRef](#)]
38. Zhang, S.; Wang, J.; Wang, K.; Li, X.; Zhao, X.; Chen, Q.; Zhang, W.; Ai, L. Differentiation of High-Grade Glioma and Primary Central Nervous System Lymphoma: Multiparametric Imaging of the Enhancing Tumor and Peritumoral Regions Based on Hybrid 18F-FDG PET/MRI. *Eur. J. Radiol.* **2022**, *150*, 110235. [[CrossRef](#)]
39. Cassinelli Petersen, G.I.; Shatalov, J.; Verma, T.; Brim, W.R.; Subramanian, H.; Brackett, A.; Bahar, R.C.; Merkaç, S.; Zeevi, T.; Staib, L.H.; et al. Machine Learning in Differentiating Gliomas from Primary CNS Lymphomas: A Systematic Review, Reporting Quality, and Risk of Bias Assessment. *AJNR Am. J. Neuroradiol.* **2022**, *43*, 526–533. [[CrossRef](#)]
40. Kong, Z.; Jiang, C.; Zhu, R.; Feng, S.; Wang, Y.; Li, J.; Chen, W.; Liu, P.; Zhao, D.; Ma, W.; et al. 18F-FDG-PET-Based Radiomics Features to Distinguish Primary Central Nervous System Lymphoma from Glioblastoma. *NeuroImage Clin.* **2019**, *23*, 101912. [[CrossRef](#)]
41. O'Malley, J.P.; Ziessman, H.A.; Kumar, P.N.; Harkness, B.A.; Tall, J.G.; Pierce, P.F. Diagnosis of Intracranial Lymphoma in Patients with AIDS: Value of 201TI Single-Photon Emission Computed Tomography. *AJR Am. J. Roentgenol.* **1994**, *163*, 417–421. [[CrossRef](#)] [[PubMed](#)]
42. Yang, M.; Sun, J.; Bai, H.X.; Tao, Y.; Tang, X.; States, L.J.; Zhang, Z.; Zhou, J.; Farwell, M.D.; Zhang, P.; et al. Diagnostic Accuracy of SPECT, PET, and MRS for Primary Central Nervous System Lymphoma in HIV Patients: A Systematic Review and Meta-Analysis. *Medicine* **2017**, *96*, e6676. [[CrossRef](#)] [[PubMed](#)]
43. Pierce, M.A.; Johnson, M.D.; Maciunas, R.J.; Murray, M.J.; Allen, G.S.; Harbison, M.A.; Creasy, J.L.; Kessler, R.M. Evaluating Contrast-Enhancing Brain Lesions in Patients with AIDS by Using Positron Emission Tomography. *Ann. Intern. Med.* **1995**, *123*, 594–598. [[CrossRef](#)] [[PubMed](#)]
44. Heald, A.E.; Hoffman, J.M.; Bartlett, J.A.; Waskin, H.A. Differentiation of Central Nervous System Lesions in AIDS Patients Using Positron Emission Tomography (PET). *Int. J. STD AIDS* **1996**, *7*, 337–346. [[CrossRef](#)] [[PubMed](#)]

45. Villringer, K.; Jäger, H.; Dichgans, M.; Ziegler, S.; Poppinger, J.; Herz, M.; Kruschke, C.; Minoshima, S.; Pfister, H.W.; Schwaiger, M. Differential Diagnosis of CNS Lesions in AIDS Patients by FDG-PET. *J. Comput. Assist. Tomogr.* **1995**, *19*, 532–536. [[CrossRef](#)] [[PubMed](#)]
46. Hoffman, J.M.; Waskin, H.A.; Schifter, T.; Hanson, M.W.; Gray, L.; Rosenfeld, S.; Coleman, R.E. FDG-PET in Differentiating Lymphoma from Nonmalignant Central Nervous System Lesions in Patients with AIDS. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **1993**, *34*, 567–575.
47. Marcus, C.; Feizi, P.; Hogg, J.; Summerfield, H.; Castellani, R.; Sriwastava, S.; Marano, G.D. Imaging in Differentiating Cerebral Toxoplasmosis and Primary CNS Lymphoma With Special Focus on FDG PET/CT. *AJR Am. J. Roentgenol.* **2021**, *216*, 157–164. [[CrossRef](#)] [[PubMed](#)]
48. Westwood, T.D.; Hogan, C.; Julyan, P.J.; Coutts, G.; Bonington, S.; Carrington, B.; Taylor, B.; Khoo, S.; Bonington, A. Utility of FDG-PETCT and Magnetic Resonance Spectroscopy in Differentiating between Cerebral Lymphoma and Non-Malignant CNS Lesions in HIV-Infected Patients. *Eur. J. Radiol.* **2013**, *82*, e374–e379. [[CrossRef](#)]
49. Dibble, E.H.; Boxerman, J.L.; Baird, G.L.; Donahue, J.E.; Rogg, J.M. Toxoplasmosis versus Lymphoma: Cerebral Lesion Characterization Using DSC-MRI Revisited. *Clin. Neurol. Neurosurg.* **2017**, *152*, 84–89. [[CrossRef](#)]
50. Bakshi, R. Neuroimaging of HIV and AIDS Related Illnesses: A Review. *Front. Biosci.* **2004**, *9*, 632. [[CrossRef](#)]
51. Schroeder, P.C.; Post, M.J.D.; Oschatz, E.; Stadler, A.; Bruce-Gregorios, J.; Thurnher, M.M. Analysis of the Utility of Diffusion-Weighted MRI and Apparent Diffusion Coefficient Values in Distinguishing Central Nervous System Toxoplasmosis from Lymphoma. *Neuroradiology* **2006**, *48*, 715–720. [[CrossRef](#)]
52. Floriano, V.H.; Torres, U.S.; Spotti, A.R.; Ferraz-Filho, J.R.L.; Tognola, W.A. The Role of Dynamic Susceptibility Contrast-Enhanced Perfusion MR Imaging in Differentiating between Infectious and Neoplastic Focal Brain Lesions: Results from a Cohort of 100 Consecutive Patients. *PLoS ONE* **2013**, *8*, e81509. [[CrossRef](#)]
53. Baldassari, L.E.; Wattjes, M.P.; Cortese, I.C.M.; Gass, A.; Metz, I.; Yousry, T.; Reich, D.S.; Richert, N. The Neuroradiology of Progressive Multifocal Leukoencephalopathy: A Clinical Trial Perspective. *Brain* **2022**, *145*, 426–440. [[CrossRef](#)]
54. Abrey, L.E.; Ben-Porat, L.; Panageas, K.S.; Yahalom, J.; Berkey, B.; Curran, W.; Schultz, C.; Leibel, S.; Nelson, D.; Mehta, M.; et al. Primary Central Nervous System Lymphoma: The Memorial Sloan-Kettering Cancer Center Prognostic Model. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2006**, *24*, 5711–5715. [[CrossRef](#)]
55. Ferreri, A.J.M.; Blay, J.-Y.; Reni, M.; Pasini, F.; Spina, M.; Ambrosetti, A.; Calderoni, A.; Rossi, A.; Vavassori, V.; Conconi, A.; et al. Prognostic Scoring System for Primary CNS Lymphomas: The International Extranodal Lymphoma Study Group Experience. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2003**, *21*, 266–272. [[CrossRef](#)]
56. Yuan, Y.; Ding, T.; Wang, S.; Chen, H.; Mao, Y.; Chen, T. Current and Emerging Therapies for Primary Central Nervous System Lymphoma. *Biomark. Res.* **2021**, *9*, 32. [[CrossRef](#)]
57. Sasayama, T.; Nakamizo, S.; Nishihara, M.; Kawamura, A.; Tanaka, H.; Mizukawa, K.; Miyake, S.; Taniguchi, M.; Hosoda, K.; Kohmura, E. Cerebrospinal Fluid Interleukin-10 Is a Potentially Useful Biomarker in Immunocompetent Primary Central Nervous System Lymphoma (PCNSL). *Neuro-Oncology* **2012**, *14*, 368–380. [[CrossRef](#)]
58. Yuan, X.-G.; Huang, Y.-R.; Yu, T.; Xu, Y.; Liang, Y.; Zhang, X.-H.; Sun, C.-R.; Zhao, X.-Y. Primary Central Nervous System Lymphoma in China: A Single-Center Retrospective Analysis of 167 Cases. *Ann. Hematol.* **2020**, *99*, 93–104. [[CrossRef](#)]
59. Tabouret, E.; Houillier, C.; Martin-Duverneuil, N.; Blonski, M.; Soussain, C.; Ghesquière, H.; Houot, R.; Larrieu, D.; Soubeyran, P.; Gressin, R.; et al. Patterns of Response and Relapse in Primary CNS Lymphomas after First-Line Chemotherapy: Imaging Analysis of the ANOCEF-GOELAMS Prospective Randomized Trial. *Neuro-Oncology* **2017**, *19*, 422–429. [[CrossRef](#)]
60. Wieduwilt, M.J.; Valles, F.; Issa, S.; Behler, C.M.; Hwang, J.; McDermott, M.; Treseler, P.; O'Brien, J.; Shuman, M.A.; Cha, S.; et al. Immunochemotherapy with Intensive Consolidation for Primary CNS Lymphoma: A Pilot Study and Prognostic Assessment by Diffusion-Weighted MRI. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2012**, *18*, 1146–1155. [[CrossRef](#)]
61. Hatzoglou, V.; Oh, J.H.; Buck, O.; Lin, X.; Lee, M.; Shukla-Dave, A.; Young, R.J.; Peck, K.K.; Vachha, B.; Holodny, A.I.; et al. Pretreatment Dynamic Contrast-Enhanced MRI Biomarkers Correlate with Progression-Free Survival in Primary Central Nervous System Lymphoma. *J. Neurooncol.* **2018**, *140*, 351–358. [[CrossRef](#)] [[PubMed](#)]
62. Kawai, N.; Zhen, H.-N.; Miyake, K.; Yamamoto, Y.; Nishiyama, Y.; Tamiya, T. Prognostic Value of Pretreatment 18F-FDG PET in Patients with Primary Central Nervous System Lymphoma: SUV-Based Assessment. *J. Neurooncol.* **2010**, *100*, 225–232. [[CrossRef](#)] [[PubMed](#)]
63. Kasenda, B.; Haug, V.; Schorb, E.; Fritsch, K.; Finke, J.; Mix, M.; Hader, C.; Weber, W.A.; Illerhaus, G.; Meyer, P.T. 18F-FDG PET Is an Independent Outcome Predictor in Primary Central Nervous System Lymphoma. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **2013**, *54*, 184–191. [[CrossRef](#)] [[PubMed](#)]
64. Okuyucu, K.; Alagoz, E.; Ince, S.; Ozaydin, S.; Arslan, N. Can Metabolic Tumor Parameters on Primary Staging 18F-FDG PET/CT Aid in Risk Stratification of Primary Central Nervous System Lymphomas for Patient Management as a Prognostic Model? *Rev. Espanola Med. Nucl. E Imagen Mol.* **2018**, *37*, 9–14. [[CrossRef](#)]
65. Albano, D.; Bertoli, M.; Battistotti, M.; Rodella, C.; Statuto, M.; Giubbini, R.; Bertagna, F. Prognostic Role of Pretreatment 18F-FDG PET/CT in Primary Brain Lymphoma. *Ann. Nucl. Med.* **2018**, *32*, 532–541. [[CrossRef](#)]
66. Krebs, S.; Mauguen, A.; Yildirim, O.; Hatzoglou, V.; Francis, J.H.; Schaff, L.R.; Mellinghoff, I.K.; Schöder, H.; Grommes, C. Prognostic Value of [18F]FDG PET/CT in Patients with CNS Lymphoma Receiving Ibrutinib-Based Therapies. *Eur. J. Nucl. Med. Mol. Imaging* **2021**, *48*, 3940–3950. [[CrossRef](#)]

67. Abrey, L.E.; Batchelor, T.T.; Ferreri, A.J.M.; Gospodarowicz, M.; Pulczynski, E.J.; Zucca, E.; Smith, J.R.; Korfel, A.; Soussain, C.; DeAngelis, L.M.; et al. Report of an International Workshop to Standardize Baseline Evaluation and Response Criteria for Primary CNS Lymphoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2005**, *23*, 5034–5043. [[CrossRef](#)]
68. Suh, C.H.; Kim, H.S.; Park, J.E.; Jung, S.C.; Choi, C.G.; Kim, S.J. Primary Central Nervous System Lymphoma: Diagnostic Yield of Whole-Body CT and FDG PET/CT for Initial Systemic Imaging. *Radiology* **2019**, *292*, 440–446. [[CrossRef](#)]
69. Bertaux, M.; Houillier, C.; Edeline, V.; Habert, M.-O.; Mokhtari, K.; Giron, A.; Bergeret, S.; Hoang-Xuan, K.; Cassoux, N.; Touitou, V.; et al. Use of FDG-PET/CT for Systemic Assessment of Suspected Primary Central Nervous System Lymphoma: A LOC Study. *J. Neurooncol.* **2020**, *148*, 343–352. [[CrossRef](#)]
70. Malani, R.; Bhatia, A.; Wolfe, J.; Grommes, C. Staging Identifies Non-CNS Malignancies in a Large Cohort with Newly Diagnosed Lymphomatous Brain Lesions. *Leuk. Lymphoma* **2019**, *60*, 2278–2282. [[CrossRef](#)]
71. Mohile, N.A.; DeAngelis, L.M.; Abrey, L.E. The Utility of Body FDG PET in Staging Primary Central Nervous System Lymphoma. *Neuro-Oncol.* **2008**, *10*, 223–228. [[CrossRef](#)]
72. Park, H.Y.; Suh, C.H.; Huang, R.Y.; Guenette, J.P.; Kim, H.S. Diagnostic Yield of Body CT and Whole-Body FDG PET/CT for Initial Systemic Staging in Patients With Suspected Primary CNS Lymphoma: A Systematic Review and Meta-Analysis. *AJR Am. J. Roentgenol.* **2021**, *216*, 1172–1182. [[CrossRef](#)]
73. Fox, C.P.; Phillips, E.H.; Smith, J.; Linton, K.; Gallop-Evans, E.; Hemmaway, C.; Auer, D.P.; Fuller, C.; Davies, A.J.; McKay, P.; et al. Guidelines for the Diagnosis and Management of Primary Central Nervous System Diffuse Large B-Cell Lymphoma. *Br. J. Haematol.* **2019**, *184*, 348–363. [[CrossRef](#)]
74. Seidel, S.; Nilius-Eliliwi, V.; Kowalski, T.; Vangala, D.B.; Schlegel, U.; Schroers, R. High-Dose Chemotherapy with Autologous Hematopoietic Stem Cell Transplantation in Relapsed or Refractory Primary CNS Lymphoma: A Retrospective Monocentric Analysis of Long-Term Outcome, Prognostic Factors, and Toxicity. *Cancers* **2022**, *14*, 2100. [[CrossRef](#)]
75. Grommes, C.; DeAngelis, L.M. Primary CNS Lymphoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2017**, *35*, 2410–2418. [[CrossRef](#)]
76. Maza, S.; Buchert, R.; Brenner, W.; Munz, D.L.; Thiel, E.; Korfel, A.; Kiewe, P. Brain and Whole-Body FDG-PET in Diagnosis, Treatment Monitoring and Long-Term Follow-up of Primary CNS Lymphoma. *Radiol. Oncol.* **2013**, *47*, 103–110. [[CrossRef](#)]
77. Van der Meulen, M.; Postma, A.A.; Smits, M.; Bakunina, K.; Minnema, M.C.; Seute, T.; Cull, G.; Enting, R.H.; van der Poel, M.; Stevens, W.B.C.; et al. Extent of Radiological Response Does Not Reflect Survival in Primary Central Nervous System Lymphoma. *Neuro-Oncology Adv.* **2021**, *3*, vdab007. [[CrossRef](#)]
78. Jo, J.-C.; Yoon, D.H.; Kim, S.; Lee, K.; Kang, E.H.; Park, J.S.; Ryu, J.-S.; Huh, J.; Park, C.-S.; Kim, J.H.; et al. Interim 18F-FDG PET/CT May Not Predict the Outcome in Primary Central Nervous System Lymphoma Patients Treated with Sequential Treatment with Methotrexate and Cytarabine. *Ann. Hematol.* **2017**, *96*, 1509–1515. [[CrossRef](#)]
79. Birsén, R.; Blanc, E.; Willems, L.; Burróni, B.; Legoff, M.; Le Ray, E.; Pilorge, S.; Salah, S.; Quentin, A.; Deau, B.; et al. Prognostic Value of Early 18F-FDG PET Scanning Evaluation in Immunocompetent Primary CNS Lymphoma Patients. *Oncotarget* **2018**, *9*, 16822–16831. [[CrossRef](#)]
80. Palmedo, H.; Urbach, H.; Bender, H.; Schlegel, U.; Schmidt-Wolf, I.G.H.; Matthies, A.; Linnebank, M.; Joe, A.; Bucerius, J.; Biersack, H.-J.; et al. FDG-PET in Immunocompetent Patients with Primary Central Nervous System Lymphoma: Correlation with MRI and Clinical Follow-Up. *Eur. J. Nucl. Med. Mol. Imaging* **2006**, *33*, 164–168. [[CrossRef](#)]
81. Mercadal, S.; Cortés-Romera, M.; Vélez, P.; Climent, F.; Gámez, C.; González-Barca, E. Positron emission tomography combined with computed tomography in the initial evaluation and response assessment in primary central nervous system lymphoma. *Med. Clin.* **2015**, *144*, 503–506. [[CrossRef](#)]
82. Kawai, N.; Miyake, K.; Yamamoto, Y.; Nishiyama, Y.; Tamiya, T. 18F-FDG PET in the Diagnosis and Treatment of Primary Central Nervous System Lymphoma. *BioMed Res. Int.* **2013**, *2013*, 247152. [[CrossRef](#)]
83. Tun, H.W.; Johnston, P.B.; DeAngelis, L.M.; Atherton, P.J.; Pederson, L.D.; Koenig, P.A.; Reeder, C.B.; Omuro, A.M.P.; Schiff, D.; O'Neill, B.; et al. Phase 1 Study of Pomalidomide and Dexamethasone for Relapsed/Refractory Primary CNS or Vitreoretinal Lymphoma. *Blood* **2018**, *132*, 2240–2248. [[CrossRef](#)]
84. Jang, S.J.; Lee, K.-H.; Lee, J.Y.; Choi, J.Y.; Kim, B.-T.; Kim, S.J.; Kim, W.S. (11)C-Methionine PET/CT and MRI of Primary Central Nervous System Diffuse Large B-Cell Lymphoma before and after High-Dose Methotrexate. *Clin. Nucl. Med.* **2012**, *37*, e241–e244. [[CrossRef](#)]
85. Kawase, Y.; Yamamoto, Y.; Kameyama, R.; Kawai, N.; Kudomi, N.; Nishiyama, Y. Comparison of 11C-Methionine PET and 18F-FDG PET in Patients with Primary Central Nervous System Lymphoma. *Mol. Imaging Biol.* **2011**, *13*, 1284–1289. [[CrossRef](#)]
86. Ahn, S.-Y.; Kwon, S.Y.; Jung, S.-H.; Ahn, J.-S.; Yoo, S.W.; Min, J.-J.; Bom, H.-S.; Ki, S.Y.; Kim, H.-J.; Lee, J.-J.; et al. Prognostic Significance of Interim 11C-Methionine PET/CT in Primary Central Nervous System Lymphoma. *Clin. Nucl. Med.* **2018**, *43*, e259–e264. [[CrossRef](#)]
87. Inoue, A.; Ohnishi, T.; Kohno, S.; Matsumoto, S.; Nishikawa, M.; Ohue, S.; Ozaki, S.; Suehiro, S.; Kurata, M.; Fukushima, M.; et al. Prognostic Significance of Immunohistochemical Subtypes Based on the Stage of B-Cell Differentiation in Primary CNS Lymphoma. *Int. J. Clin. Exp. Pathol.* **2019**, *12*, 1457–1467. [[CrossRef](#)]

88. Nomura, Y.; Asano, Y.; Shinoda, J.; Yano, H.; Ikegame, Y.; Kawasaki, T.; Nakayama, N.; Maruyama, T.; Muragaki, Y.; Iwama, T. Characteristics of Time-Activity Curves Obtained from Dynamic 11C-Methionine PET in Common Primary Brain Tumors. *J. Neurooncol.* **2018**, *138*, 649–658. [[CrossRef](#)]
89. Starzer, A.M.; Berghoff, A.S.; Traub-Weidinger, T.; Haug, A.R.; Widhalm, G.; Hacker, M.; Rausch, I.; Preusser, M.; Mayerhoefer, M.E. Assessment of Central Nervous System Lymphoma Based on CXCR4 Expression In Vivo Using 68Ga-Pentixafor PET/MRI. *Clin. Nucl. Med.* **2021**, *46*, 16–20. [[CrossRef](#)]
90. Herhaus, P.; Lipkova, J.; Lammer, F.; Yakushev, I.; Vag, T.; Slotta-Huspenina, J.; Habringer, S.; Lapa, C.; Pukrop, T.; Hellwig, D.; et al. CXCR4-Targeted PET Imaging of Central Nervous System B-Cell Lymphoma. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **2020**, *61*, 1765–1771. [[CrossRef](#)]
91. Hovhannisyanyan, N.; Fillesoye, F.; Guillouet, S.; Ibazizene, M.; Toutain, J.; Gourand, F.; Valable, S.; Plancoulaine, B.; Barré, L. [¹⁸F]Fludarabine-PET as a Promising Tool for Differentiating CNS Lymphoma and Glioblastoma: Comparative Analysis with [¹⁸F]FDG in Human Xenograft Models. *Theranostics* **2018**, *8*, 4563–4573. [[CrossRef](#)] [[PubMed](#)]
92. Postnov, A.; Toutain, J.; Pronin, I.; Valable, S.; Gourand, F.; Kalaeva, D.; Vikhrova, N.; Pyzhik, E.; Guillouet, S.; Kobayakov, G.; et al. First-in-Man Noninvasive Initial Diagnostic Approach of Primary CNS Lymphoma Versus Glioblastoma Using PET With 18F-Fludarabine and l-[Methyl-11C]Methionine. *Clin. Nucl. Med.* **2022**, *47*, 699–706. [[CrossRef](#)] [[PubMed](#)]
93. Barré, L.; Hovhannisyanyan, N.; Bodet-Milin, C.; Kraeber-Bodéré, F.; Damaj, G. [¹⁸F]-Fludarabine for Hematological Malignancies. *Front. Med.* **2019**, *6*, 77. [[CrossRef](#)] [[PubMed](#)]

Chapitre 3. Biomarqueurs pronostiques du LPSNC

I. Données de survie

En dépit des avancées thérapeutiques notables, le LPSNC reste une affection avec un pronostic particulièrement défavorable [82]. Les chiffres issus du réseau LOC, regroupant une cohorte de 1002 patients traités entre 2011 et 2016, indiquent une médiane de survie globale de 25 mois et un taux de survie sans progression à 5 ans s'élevant à 38% [1,83]. Ce pronostic est encore plus alarmant chez les patients âgés de plus de 65 ans, avec un taux de mortalité précoce, c'est-à-dire dans les 6 mois suivant le diagnostic, atteignant les 25%. Bien que le LPSNC soit couramment lié à une infection par le virus d'Epstein-Barr (EBV) chez les individus immunodéprimés, les facteurs de risque demeurent largement méconnus pour les lymphomes cérébraux affectant les sujets immunocompétents.

II. Biomarqueurs pronostiques

Cliniques

Dans le contexte du LPSNC, les deux facteurs pronostiques reconnus sont l'âge et le performance status (PS). Les données de la littérature fixent généralement le seuil à 60 ans pour parler de LPSNC du sujet âgé ; un seuil au-delà duquel le risque de toxicités cérébrales liées au traitement de consolidation est accru [83]. Toutefois, des études récentes ont révélé un avantage à poursuivre le traitement de consolidation par autogreffe de cellules souches chez des patients âgés rigoureusement sélectionnés, repoussant ainsi la limite de la catégorie « âgée » jusqu'à 70 ans [84]. Les diverses options thérapeutiques pour le LPSNC seront exposées en détail au chapitre 4. Concernant le statut fonctionnel, une méta-analyse incluant 783 patients immunocompétents de plus de 60 ans a démontré qu'un score Karnofsky Performance Status (KPS) supérieur ou égal à 70 était prédictif d'une meilleure survie globale, indépendamment de l'âge ou du traitement choisi [85] (cf. Annexe I). Ces deux paramètres (âge et KPS) sont intégrés dans deux modèles de scores pronostiques cliniques reconnus : celui de l'International Extra nodal Lymphoma Study Group (IELSG) et celui du Memorial Sloan-Kettering Cancer Center (MSKCC). Ces deux scores sont présentés ci-dessous [27,86] : (tableau 3.1)

IELSG score	Variable	Favorable feature (value 0)	Unfavorable feature (value 1)
	Age (years)	<60	>60
	ECOG PS	0-1	>1
	LDH serum level	normal	elevated
	CSF protein level	normal	elevated
	Involvement of deep regions of the CNS	no	yes

MSKCC score		Good Risk	Intermediate Risk	High Risk
	Age	<50	≥ 50	≥ 70
ECOG PS			≥ 70	<70

Tableau 3.1 Scores pronostiques IELSG et MSKCC. *D'après Ferreri, JCO, 2003*
 CNS : Central nervous system, ECOG : Eastern Cooperative Oncology Group, PS : performance status, LDH : lactate dehydrogenase, CSF : cerebrospinal fluid, MSKCC : Memorial Sloan-Kettering Cancer Center

Récemment, le score 3F a ré-émergé comme un outil prometteur pour affiner la stratification pronostique des patients bénéficiant de nouvelles thérapies de consolidation, influençant à la fois la SSP et la SG. Le score 3F, développé initialement par Jang et al., se compose de trois paramètres: l'âge (supérieur à 50 ans), l'état de performance selon l'ECOG (score ECOG PS ≥ 2) et la lymphopénie (compte absolu de lymphocytes $\leq 875/\mu\text{L}$) [87,88].

Le rôle du genre reste quant à lui controversé ; la plupart des séries ne retrouvant pas d'association avec la survie [1,27,89–91]. Enfin, le statut cognitif au moment du diagnostic apparaît également comme un facteur prédictif indépendant significatif dans le LPSNC. Plusieurs études multicentriques ont établi un lien entre le déficit cognitif, évalué par l'échelle MMSE (mini-mental state examination) ou la MDRS (mattis dementia rating scale), et un pronostic défavorable [92,93].

Biologiques

Historiquement, les marqueurs biologiques tels que la lactate déshydrogénase (LDH) et les protéines de colony stimulating factors (CSF) ont été intégrés dans le score IELSG pour leur rôle pronostique présumé dans le LPSNC. Cependant, les études ultérieures ont rapporté des résultats variables, en partie à cause de l'absence de seuils bien définis. L'attention se porte désormais sur l'interleukine-10 (IL-10) et le ratio IL-10/IL-6, tant au diagnostic qu'en tant qu'indicateurs pronostiques durant l'évolution de la maladie [4,94]. L'émergence du profilage génétique a permis l'identification de mutations moléculaires dans le LPSNC, influençant les voies de signalisation tumorale. Certaines de ces altérations, comme BCL2, MYD et PD1, ont été associées au devenir clinique, mais ces découvertes, principalement issues d'études rétrospectives de faible envergure, nécessitent confirmation à une échelle plus vaste. [95,96].

Par ailleurs, l'implication de l'inflammation dans la progression tumorale a récemment pris de l'importance dans la recherche en oncologie, l'inflammation jouant un rôle dans la croissance tumorale, en modelant le microenvironnement et en favorisant la prolifération, la survie et la migration des cellules cancéreuses. Dans ce contexte, l'élévation de la protéine C-réactive (CRP) a été identifiée comme un marqueur pronostique potentiel dans le LPSNC [97]. Enfin, une étude multicentrique publiée en 2023 a présenté un nouveau modèle basé sur une échelle à deux facteurs, incluant l'ECOG et le taux d'albumine qui permettrait de stratifier la survie à 5 ans des patients de manière significative [98]. Bien que les mécanismes conduisant à une baisse du taux d'albumine chez les patients atteints de cancer restent encore mal compris, plusieurs hypothèses ont été proposées : une diminution de la production d'albumine liée à une réduction de la production de cytokines IL-6, une fuite vers les espaces extravasculaires due à l'augmentation de la perméabilité microvasculaire en réponse à une élévation des niveaux de facteur de nécrose tumorale (TNF), ou encore une dégradation accrue associée à la réponse inflammatoire systémique contre la tumeur [98].

Imagerie

En IRM, plusieurs biomarqueurs ont été évalués, tels que la taille des lésions, leur nombre et localisation, comme nous le détaillons dans le tableau ci-dessous :

Study	PMID	Nb	Biomarker of treatment failure	Biomarker of poor outcome
Anatomical Characteristics				
Ferreri, JCO, 2003	12525518	378		deep brain involvement (OS)
Ghesquière, Hematol Oncol, 2013	22886543	91		deep brain involvement (OS)
Tabouret, Neuro-Oncol, 2017	27994065	85		Infra-tentorial location (OS) Large volume (>11.2 cc) (OS)
Jahr, Brain Behav, 2018	29541540	79		Deep brain involvement (OS) Large volume (>64.9mm) (OS)
Diffusion biomarkers				
Barajas, AJNR, 2010	19729544	18		$ADC_{mean} < 692.10^{-6} \text{ mm}^2/\text{s}$ (PFS and OS)
Wieduwilt, Clin Cancer Res, 2012	22228634	31		$ADC_{min} < 384.10^{-6} \text{ mm}^2/\text{s}$ (PFS and OS)
Zhang, Clin Radiol, 2016	27341986	28		↘ ADC (poor PFS)
Huang, AJNR, 2016	27390318	35	↘ ADCmin	
Chong, Cancers, 2019	31597366	93		↘ ADC (poor OS)
Baek, Blood Res, 2020	32408414	67		↘ ADC (poor OS)
Cornell, J Pers Med, 2023	37511795	38	Large volume, ↘ ADC_{min} , ADC_{mean}	
Chien, Front Neurol, 2023	37638189	52	↘ ADC ($680.10^{-6} \text{ mm}^2/\text{s}$) and ADCr ($970.10^{-6} \text{ mm}^2/\text{s}$)	↘ ADC and ADCr (poor PFS)
Perfusion biomarkers				
Valles, AJNR, 2013	22936096	25		$ADC_{min} < 384.10^{-6}$ - $rCBV_{mean} < 1.43$ (PFS and OS)
Chung, Medicine, 2016	26871782	48	Heterogeneous perfusion	
Hatzoglou, Journal of Neuro-Oncol, 2018	30073640	18		↘ Vp_{mean} and K^{trans}_{mean} (PFS)
Fu, Eur Radiol, 2021	32997181	56		↗ K^{trans} (PFS)
Luo, J Clin Neurosci, 2022	34929639	35		Low perfusion (PFS)

Tableau 3.2 Données de la littérature sur les biomarqueurs pronostiques en IRM

Certaines études ont établi un lien entre la présence de lésions multifocales et une diminution de la survie [99,100], tandis que d'autres études ont identifié la localisation profonde ou sous-tentorielle des lésions comme étant un indicateur de pronostic défavorable. Il a également été démontré que l'augmentation du volume de la lésion rehaussée sur la séquence T1 avec injection de gadolinium, était un marqueur de mauvais pronostic [101,102]. Par exemple, l'étude de Tabouret et al. a observé que les patients avec un volume tumoral supérieur à 11.4cc avait une SSP moindre [101]. Récemment, l'intérêt s'est porté sur les biomarqueurs d'imagerie multimodale, tels que l'ADC ou les indicateurs de perfusion. Les résultats des différentes études montrent qu'une baisse de l'ADC est associée à une survie réduite [103].

En ce qui concerne les indicateurs de perfusion, les résultats sont plus variables. Hatzoglou et al. ont souligné l'importance du volume plasmatique sanguin (Vp) et de la constante de transfert de volume (K^{trans}), notant un pronostic d'autant plus défavorable que ces valeurs sont basses [104]. Toutefois, ces conclusions n'ont pas été confirmées par d'autres études.

En médecine nucléaire, quelques études ont investigué les paramètres TEP-FDG comme biomarqueurs pronostiques des LPSNC mais l'utilisation en pratique de cette technique, en dehors du bilan d'extension, reste non standardisée. En 2010, Kawai et al. ont suggéré une corrélation entre le SUVmax et les SSP et SG mais les travaux ultérieurs n'ont pas corroboré ces résultats [105]. En 2013, l'étude de Kasenda et al. a introduit le score MILAS comme un nouvel outil pronostique. Le MILAS est une échelle de couleur constituée de dix gradations établies en fonction de l'activité métabolique physiologique du cervelet, qui est définie en blanc et correspond à une intensité de 10 à 20 % du SUVmax. Les zones présentant un métabolisme inférieur à celui du cervelet sont représentées en noir. Les zones avec un métabolisme supérieur à celui du cervelet sont différenciées en huit gradations en fonction de leur intensité : de 20 à 30 % du SUVmax en bleu clair ; de 30 à 40 % du SUVmax en bleu foncé ; de 40 à 50 % du SUVmax en vert clair ; de 50 à 60 % du SUVmax en vert foncé ; de 60 à 70 % du SUVmax en orange ; de 70 à 80 % du SUVmax en rouge ; de 80 à 90 % du SUVmax en violet clair et de 90 à 100 % du SUVmax en violet foncé [13]. (Figure 3.1).

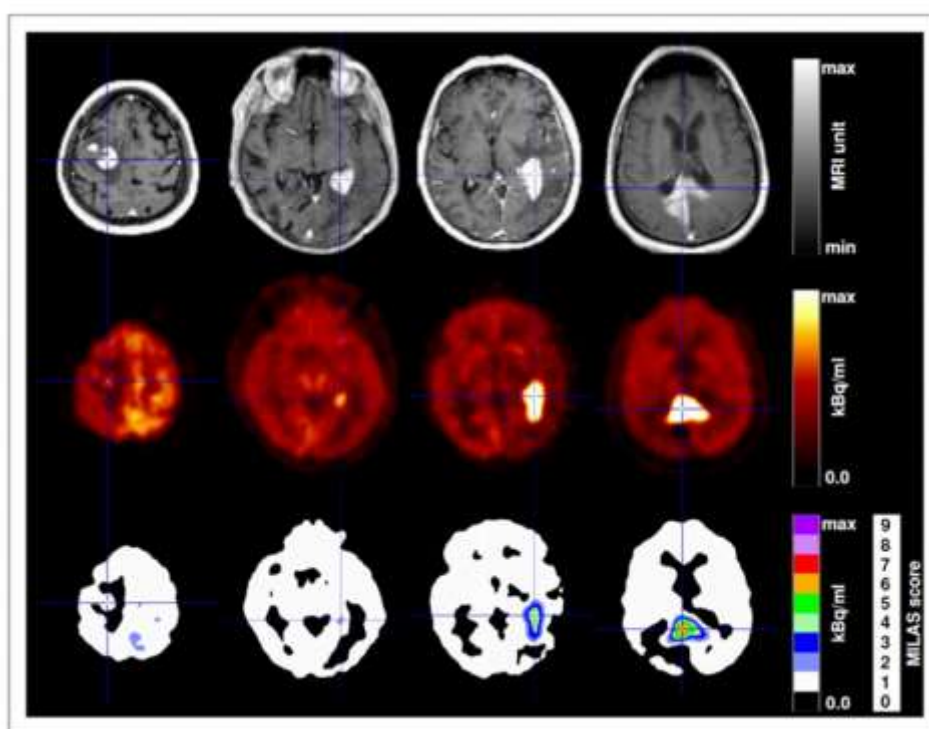


FIGURE 3.1 Représentation du score MILAS

IRM avec injection de produit de contraste (1ère rangée) et TEP-FDG (2ème rangée) de quatre patients atteints de LPSNC. La troisième rangée illustre l'évaluation semi-quantitative MILAS, qui emploie une échelle colorimétrique pour quantifier la captation du FDG par rapport à une zone de référence dans le cervelet. D'après Kasenda, JNM, 2013 [13]

Dans une étude portant sur 52 patients atteints de LPSNC et ayant bénéficié d'une TEP-FDG en baseline, Albano et al., n'ont observé aucun effet pronostique lié au SUVmax. Toutefois, l'étude a révélé que les patients présentant un MTV ≥ 9.8 cm³ et un TLG ≥ 94 g.cm³ avaient une SSP et une SG plus courtes [15]. Par ailleurs, une étude prospective incluant 53 patients traités par chimiothérapie à base d'Ibrutinib, menée par Krebs et al. a indiqué que le SUVmax, le MTV et le TLG pourraient être des marqueurs prédictifs de la réponse en fin de traitement [9]. Dans cette étude, les auteurs ont également introduit une nouvelle méthode qui consistait à calculer les biomarqueurs sumSUVmax, sumMTV et sumTLG en additionnant les valeurs de ces paramètres mesurées pour chaque lésion, jusqu'à un maximum de cinq lésions cibles. Ces marqueurs étaient corrélés à la réponse en fin de traitement, mais un sumSUVmax élevé était le seul indicateur d'une SSP défavorable [9].

Face à un corpus limité d'études, majoritairement rétrospectives et portant sur de petits effectifs, l'étude LOCALYZE se positionne comme un pivot dans l'évaluation de l'utilité du TEP FDG conjointement à l'IRM multimodale, dans la prise en charge du lymphome du système nerveux central primaire. Ancillaire à l'essai BLOCAG-01, qui avait pour but d'évaluer l'intérêt d'une chimiothérapie de maintenance, associant rituximab, méthotrexate et temozolomide, face à une simple stratégie de surveillance, dans l'optique d'améliorer la survie ; LOCALYZE bénéficie d'une TEP-IRM au FDG avant traitement (TEP baseline) ainsi qu'un suivi standardisé pour tous les patients inclus. Un résumé détaillé du protocole Localyse est disponible en annexe.

III. Article 2.

Nouveaux biomarqueurs pronostiques TEP-IRM dans le lymphome primitif du système nerveux central chez le sujet âgé : une étude ancillaire de l'essai BLOCAGE, soutenu par le réseau LOC

Objectifs

Les objectifs de notre étude étaient d'une part, de décrire les caractéristiques morpho-métaboliques avant traitement des LPSNC *de novo*, en TEP et en IRM, dans une cohorte de patients âgés immunocompétents ; d'autre part, d'évaluer la valeur prédictive de la TEP-IRM baseline sur la réponse en fin de chimiothérapie d'induction et enfin d'évaluer la valeur pronostique de la TEP-IRM baseline sur la SG des patients.

Méthodes

Cette recherche prospective, multicentrique (IP : A. KAS) est une étude ancillaire de l'essai de phase III BLOCAGE-01 (IP : K. Hoang-Xuan). Les patients inclus étaient immunocompétents, âgés de plus de 60 ans et avaient un diagnostic histologiquement prouvé de LPSNC. Les participants ont bénéficié d'une TEP-IRM au FDG pré-thérapeutique, à mi-parcours de la phase d'induction par R-MPVA et à la fin de ce traitement. Les réponses en fin de traitement ont été évaluées par des neuro-oncologues et neuroradiologues experts, en utilisant les critères IPCG. La SSP et la SG ont été calculées depuis la date du diagnostic jusqu' à la progression de la maladie ou le décès du patient, respectivement.

En termes d'analyse d'images, une lésion cible a été définie en cas de lésions multiples, en retenant celle ayant le SUVmax le plus élevé. Les paramètres TEP évalués étaient le SUVmax, le SUVmean et le SUVpeak de la lésion ainsi que le ratio de ces mesures au côté controlatéral et au cervelet sains. Le score MILAS a également été calculé pour chaque patient, en suivant la méthodologie décrite par Kasenda et al [13]. Le MTV a été établi avec un seuil fixé à 41% du SUVmax, et le TLG a également été calculé. En suivant la méthodologie de Krebs et al., nous avons aussi déterminé le sumSUVmax et le sumMTV en additionnant respectivement les valeurs de SUVmax et de MTV des lésions cérébrales [9]. L'analyse IRM, réalisée avec le logiciel Advantage Workstation de GE, a permis d'extraire les volumes en séquences T1 avec gadolinium et en T2 FLAIR. Des cartes paramétriques ont été générées pour évaluer les paramètres d'ADC, de CBV et de CBF.

Résultats

65 patients ont été inclus dans l'étude, parmi lesquels 54 ont pu bénéficier d'une TEP-IRM initiale et d'une évaluation complète en fin de traitement. L'âge moyen des participants était de 71.5 ans. À la fin de la chimiothérapie d'induction, 37 patients (69%) étaient considérés en réponse (RC, RCu ou RP), tandis que 17 patients (31%) étaient classés comme ayant une progression de la maladie.

L'imagerie a révélé que la majorité des lésions de LPSNC était supratentorielle (88 %) et profonde (80 %), avec une répartition homogène entre localisation unique (48 %) et multifocale (52 %). En IRM, les lésions montraient une prise de contraste intense dans 85 % des cas, associée à une restriction de diffusion et une perfusion modérée. En TEP, l'avidité pour le FDG était très élevée, environ 3 à 5 fois supérieure à celle du parenchyme sain controlatéral, sans impact détecté de la prise antérieure de corticoïdes. Un fort hypermétabolisme et un volume métabolique élevé sur la TEP baseline étaient significativement corrélés à des valeurs faibles d'ADCmean et élevées de CBF sur l'IRM baseline.

La SSP médiane était de 29,6 mois (IC à 95% [9,9 ; NA]) avec un suivi médian de 37 mois (IC à 95% [33,3 ; 47,4]). Un SUVmean cérébelleux élevé et un sumMTV bas à la TEP pré-thérapeutique se sont révélés être des prédicteurs significatifs d'une réponse complète en fin de traitement. Un volume tumoral élevé, mesuré en TEP ou sur l'IRM-T1 avec gadolinium, était un prédicteur significatif d'une moins bonne SG, tandis qu'un score MILAS élevé au diagnostic était associé à un risque accru de décès précoce (moins de 1 an).

Discussion

Bien que les seuls facteurs pronostiques actuellement bien établis soient le KPS et l'âge, ces indicateurs sont insuffisants pour guider de manière optimale la prise en charge thérapeutique des patients atteints de LPSNC. Notre étude a permis d'évaluer de manière synchronisée l'imagerie par TEP et IRM, révélant une prise de contraste intense et homogène des lésions et un hypermétabolisme intense, en accord avec la littérature existante. Nous avons également observé une restriction notable de la diffusion, probablement due au caractère fortement cellulaire de ces lésions

Nos résultats ont également révélé que le SUVmean cérébelleux constituait un marqueur prédictif de la réponse en fin de traitement et de la survie. Le métabolisme cérébelleux avait déjà été identifié comme un potentiel facteur pronostique dans le lymphome systémique dans deux études récentes [106,107]. Notre étude n'a pas mis en évidence de corrélation entre le

SUVmean cérébelleux et le volume, l'intensité ou la localisation de la lésion, possiblement du fait d'un échantillon de patients limité. De plus, le SUVmean du parenchyme sain n'était pas associé à la réponse de fin de traitement ou à la survie. Le cervelet est traditionnellement associé aux fonctions motrices mais joue également un rôle dans les fonctions cognitives grâce à ses connexions avec des réseaux neuronaux impliqués dans les fonctions exécutives et le mode par défaut [108–110]. Dans notre cohorte, nous n'avons pas retrouvé de corrélation entre le métabolisme cérébelleux et le KPS au diagnostic. Une étude du lien entre le SUVmean du cervelet et le score neuropsychologique au diagnostic pourrait faire l'objet d'étude intéressante dans le futur.

Dans notre étude, le volume tumoral mesuré en TEP ou sur l'IRM T1 avec gadolinium était un indicateur de mauvaise réponse à la chimiothérapie d'induction et de survie réduite. Ces observations sont cohérentes avec celles de l'étude d'Albano et al., qui avait montré sur une cohorte de 52 patients que ceux présentant un MTV élevé avaient une SSP et une SG diminuées [15]. Enfin, nos résultats confirment également ceux de Kasenda et al. concernant le score MILAS comme indicateur d'une mauvaise SG [13].

En conclusion, notre étude est l'une des plus large à explorer les biomarqueurs d'imagerie chez les patients atteints de LPSNC et la première à bénéficier d'une imagerie combinée TEP-IRM. Des recherches supplémentaires sont nécessaires pour évaluer le rôle potentiel de la TEP-IRM au FDG pour l'évaluation thérapeutique.



[18F]FDG PET-MRI provides survival biomarkers in primary central nervous system lymphoma in the elderly: an ancillary study from the BLOCAGE trial of the LOC network

Laura Rozenblum^{1,2} · Damien Galanaud^{2,3} · Caroline Houillier⁴ · Carole Soussain⁵ · Amandine Baptiste⁶ · Lisa Belin⁶ · Véronique Edeline⁷ · Philippe Naggara¹ · Marine Soret^{1,2} · Valérie Causse-Lemercier¹ · Lise Willems⁸ · Sylvain Choquet⁹ · Renata Ursu¹⁰ · Khê Hoang-Xuan⁴ · Aurélie Kas^{1,2}

Received: 21 April 2023 / Accepted: 5 July 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Purpose Primary central nervous system lymphoma (PCNSL) incidence is rising among elderly patients, presenting challenges due to poor prognosis and treatment-related toxicity risks. This study explores the potential of combining [18F] fluorodeoxyglucose ([18F]FDG) PET scans and multimodal MRI for improving management in elderly patients with de novo PCNSL.

Methods Immunocompetent patients over 60 years with de novo PCNSL were prospectively enrolled in a multicentric study between January 2016 and April 2021. Patients underwent brain [18F]FDG PET-MRI before receiving high-dose methotrexate-based chemotherapy. Relationships between extracted PET (metabolic tumor volume (MTV), sum of MTV for up to five lesions (sumMTV), metabolic imaging lymphoma aggressiveness score (MILAS)) and MRI parameters (tumor contrast-enhancement size, cerebral blood volume (CBV), cerebral blood flow (CBF), apparent diffusion coefficient (ADC)) and treatment response and outcomes were analyzed.

Results Of 54 newly diagnosed diffuse large B-cell PCNSL patients, 52 had positive PET and MRI with highly [18F]FDG-avid and contrast-enhanced disease (SUV_{max}: 27.7 [22.8–36]). High [18F]FDG uptake and metabolic volume were significantly associated with low ADC_{mean} values and high CBF at baseline. Among patients, 69% achieved an objective response at the end of induction therapy, while 17 were progressive. Higher cerebellar SUV_{mean} and lower sumMTV at diagnosis were significant predictors of complete response: 6.4 [5.7–7.7] vs 5.4 [4.5–6.6] ($p=0.04$) and 5.5 [2.1–13.3] vs 15.9 [4.2–19.5] ($p=0.01$), respectively. Two-year overall survival (OS) was 71%, with a median progression-free survival (PFS) of 29.6 months and a median follow-up of 37 months. Larger tumor volumes on PET or enhanced T1-weighted MRI were significant predictors of poorer OS, while a high MILAS score at diagnosis was associated with early death (< 1 year).

Conclusion Baseline cerebellar metabolism and sumMTV may predict response to end of chemotherapy in PCNSL. Tumor volume and MILAS at baseline are strong prognostic factors.

Keywords PET/MRI · [18F]FDG · Primary central nervous system lymphoma · Outcome analysis

Khê Hoang-Xuan and Aurélie Kas contributed equally to this work.

Trial: LOCALYZE, NCT02313389, ancillary of phase III clinical trial BLOCAGE01 (Registered July 10, 2018 – retrospectively registered); <https://clinicaltrials.gov/ct2/show/NCT03582254?term=LOCALYZE&draw=2&rank=1>.

Extended author information available on the last page of the article

Introduction

Primary central nervous system lymphoma (PCNSL) is a subtype of non-Hodgkin's lymphoma that affects the brain, eyes, and leptomeninges. It is the third-most common malignant primary brain tumor after glioblastoma and diffuse astrocytoma [1]. In recent years, the incidence of PCNSL has increased in immunocompetent patients and has received growing attention, resulting in numerous research efforts and publications [2]. The increase in incidence has been particularly marked in older patients, but despite improved

survival rates in younger patients, the prognosis for the elderly remains poor [3, 4]. To date, molecular biology and prognostic biomarkers of PCNSL are not well understood. Age and performance status, incorporated in the International Extranodal Lymphoma Study Group and Memorial Sloan-Kettering Cancer Center scores, are the only recognized survival biomarkers [5, 6].

Standard treatment for PCNSL is based on high-dose methotrexate (HD-MTX) induction chemotherapy for patients over 60 years and on a combination of HD-MTX-based chemotherapy and high-dose chemotherapy with autologous stem cell transplantation for younger patients [7, 8]. Response to first-line therapy is a key predictor of prognosis and survival, making it a reliable surrogate endpoint for assessing response before starting chemotherapy and tailoring treatment to the aggressiveness of the disease [9]. Currently, the size of contrast-enhanced tumor in magnetic resonance imaging (MRI) is the gold standard for assessing response to treatment for PCNSL [10]. Although studies have suggested the prognostic value of markers for tumor cellularity and perfusion, such as apparent diffusion coefficient (ADC) and relative regional cerebral blood volume (rCBV) measured prior to treatment, these results are based on series with a low number of subjects and need further validation [11–13].

[18F]Fluoro-deoxyglucose positron emission tomography ([18F]FDG PET) has been widely used in oncology imaging, and its utility has been firmly established in the management of systemic lymphoma [14]. The use of whole-body [18F]FDG PET to rule out systemic disease in PCNSL has been established and plays a crucial role in patient treatment decisions [15]. However, the role of brain PET scans in PCNSL remains less clear. Several brain [18F]FDG PET survival markers have shown potential, but results are inconsistent across studies. For instance, some studies have shown that higher [18F]FDG uptake of brain lesions, whether measured through the sum of the maximum standardized uptake value (SUV_{max}) or visually represented by a 10-step color scale, has an inverse relationship with progression-free survival (PFS). Other studies have reported shorter survival in patients with larger metabolic tumor volume (MTV) or total lesion glycolysis (TLG) [16–18]. Given the significant variability in prognosis observed in patients with PCNSL, and the limited predictive value of age and performance status in predicting chemoresistance or the risk of relapse following chemotherapy, the use of imaging-based prognostic parameters shows promise in guiding treatment decisions, such as considering autograft consolidation in elderly patients, as well as selecting patients for novel therapeutic strategies.

This study (LOCALYZE, NCT03582254) is part of a phase III randomized controlled trial (BLOCAGE01 NCT02313389) that aimed to evaluate the benefits of maintenance chemotherapy (rituximab, methotrexate,

temozolomide) versus observation, to prolong complete remission (CR) with acceptable toxicity in elderly patients with PCNSL after standard induction chemotherapy. This ancillary study was conducted (i) concurrently with the BLOCAGE01 study, (ii) during the induction chemotherapy regimen—identical for all patients—and (iii) included a brain [18F]FDG PET-MRI at baseline. Our objectives were (i) to compare the PET and MRI findings at diagnosis in elderly immunocompetent patients with *newly diagnosed* PCNSL; (ii) evaluate the predictive value of baseline [18F]FDG PET-MRI on the response at the end of induction therapy in a homogeneously-treated group of patients; (iii) evaluate the prognostic value of baseline [18F]FDG PET-MRI on patients' survival.

Materials and methods

Clinical study

Patients included in the BLOCAGE01 study were eligible to this prospective ancillary study (LOCALYZE) and recruited by five centers from the french “Lymphomes Oculo-Cerebraux” Group (LOC) (Neuro-Oncology department in Saint-Louis Hospital, APHP, Paris; Haematology department in Institut Curie, Saint Cloud, Neuro-Oncology and Haematology departments in Pitié-Salpêtrière Hospital; APHP, Paris and Haematology department in Cochin Hospital; AP-HP, Paris) between January 2016 and April 2021. The inclusion criteria included (i) newly diagnosed PCNSL, confirmed either by brain histology or cytology in the cerebrospinal fluid (CSF) or vitrectomy; (ii) age of 60 years or older; (iii) Karnofsky Performance Status (KPS) greater than 40; and (iv) no evidence of systemic non-Hodgkin lymphoma as demonstrated by a contrast-enhanced CT and/or whole-body [18F]FDG PET/CT scan. Exclusion criteria included (i) any other active malignancy (with the exception of basal cell carcinoma of the skin and cervical carcinoma in situ), (ii) pre-existing immunodeficiency (from HIV or organ transplant treatment), (iii) low-grade histology, (iv) isolated primary vitreoretinal lymphoma, (v) central nervous system (CNS) relapse of a systemic lymphoma, and (vi) prior treatment for PCNSL. At the start of chemotherapy and registration into LOCALYZE, patients underwent a pretherapeutic evaluation including ophthalmologic examination, lumbar puncture with Interleukine-10 bio-assay, bone marrow analysis, brain MRI, and [18F]FDG PET-CT to rule out systemic involvement.

Patients included in LOCALYZE underwent a baseline [18F]FDG PET-MRI in a median of 3.5 [1–5] days prior to the start of a high-dose methotrexate-based chemotherapy (HD-MTX-based CT). The chemotherapy regimen consisted of four 28-day cycles of rituximab (375 mg/m², D1),

methotrexate (3.5 g/m², D1 and D15), procarbazine (100 mg/m², D1 to D7), and vincristine (1.4 mg/m², D1 and D15) (R-MPV) followed by a cycle of cytarabine (3 g/m², D1, D2).

Therapeutic response was assessed at end of therapy by board-certified neurooncologists and neuroradiologist (CH, KHX, DG) using International PCNSL Collaborative Group criteria (IPCG), based on MR imaging, corticosteroid use, and CSF cytology and slit lamp examination in case of CSF or ocular involvement at baseline [19, 20]. The ICPG criteria classify treatment response into five categories: complete response (CR), unconfirmed complete response (CRu), partial response (PR), stable disease, and progressive disease (PD). CR requires complete disappearance of all signs of the disease on gadolinium-enhanced MRI as well as on CSF cytology and ocular examination. CRu includes patients who meet CR criteria but continue to require corticosteroids or show minor abnormalities. PR involves a 50% decrease in the primary brain lesion, while PD involves a 25% increase in the lesion or new disease sites [20]. Only patients who obtained a CR or CRu were subsequently randomized into two groups in the BLOCAGE01 study: maintenance arm with rituximab, methotrexate, and temozolomide or observation arm.

Progression-free survival (PFS) was defined as the time between the diagnosis and the progression of the disease or the death of the patient. Overall survival (OS) was calculated as the time between the diagnosis and the death of the patient whatever the cause. The cutoff date for PFS and OS event was the 26 October 2022.

All participants provided informed consent for their participation in the study. The study was approved by the protection of persons committee (2014–002597-37) and the French National Agency for the Safety of Medicines and Health Products (ANSM), and abided by the principles of the Declaration of Helsinki.

[18F]FDG PET-MRI imaging protocol

All PET/MR studies were performed with hybrid PET/MRI system (Signa PET/MR 3 T General Electric Healthcare) at the Pitié-Salpêtrière Hospital's nuclear medicine department. Brain PET images were acquired 90 min after the injection of 2 MBq/kg of [18F]FDG and lasted for 20 min. Patients fasted for at least 4 h and blood glucose was measured before the exam could be performed. PET reconstruction with OSEM-3D included time-of-flight and spatial resolution modeling (6 iterations, 28 subsets, 6.4 mm transaxial post-reconstruction Gaussian filter). Reconstructed voxel size was 1.2 × 1.2 × 2.8 mm. MRI included 3D susceptibility-weighted imaging (SWAN), 3D fluid attenuated inversion recovery (FLAIR), T2* weighted perfusion, diffusion, monovoxel spectroscopy, 3D arterial spin labelling (ASL)

sequences, and 3D T1 weighted spin-echo sequence before and after injection of 0.2 ml/kg of gadoterate dimeglumine (Dotarem®). Four-tissue-class (soft tissue, fat, lung, and air) attenuation correction maps were calculated from the 2-point Dixon sequences, supplemented by a single atlas to capture bone information. Each patient underwent an interim PET-MRI (after 2 cycles of induction therapy) and a final PET-MRI (at end of therapy), in addition to the baseline PET-MRI performed at diagnosis. The results presented in this article focus on the significance of the baseline PET-MRI.

[18F]FDG PET-MRI images analysis

Images were analyzed by two board-certified nuclear medicine specialists, who were blinded to patients' follow-up, using the General Electric Healthcare Advantage Workstation (L.R., A.K.). For patients with multiple lesions on baseline PET-MRI, the target lesion was defined as the lesion with the highest SUV_{max}. Quantified parameters extracted from PET included [18F]FDG-uptake (SUV_{mean}, SUV_{peak} and SUV_{max}) in the lesions as well as in normal contralateral brain (using a mirror region-of interest [ROI]) to calculate tumor-to-background ratio (TBR), and normal cerebellum (using a spherical volume-of-interest [VOI] of 1.5 cm³ positioned on cerebellar hemisphere unaffected by diaschisis or tumor infiltration) to calculate metabolic imaging lymphoma aggressiveness score (MILAS), respectively [18]. The MILAS score, as described by Kasenda et al., is a custom-made 10-step scale that reflects the uptake of [18F]FDG in comparison to the cerebellum, which is designated as grade 1. Any [18F]FDG uptake below that of the cerebellum is coded as grade 0 (black), while uptakes exceeding cerebellar levels are represented in eight distinct gradations based on intensity, which range from grades 2 to 9. In this study, we employed this numerical grading scale from 0 to 9 as a replacement for the original color representation.

Metabolic volume of the target lesion was established using a 41% SUV_{max} threshold and TLG calculated by multiplying the MTV and SUV_{mean} of the lesion. The sum of SUV_{max} (sumSUV_{max}) and the sum of MTV (sumMTV) using up to five target lesions were calculated as a measure of patient tumor burden [17].

For MRI analysis, the 3D tumor was delineated using a semi-automated volume on the T1-weighted enhanced sequence. Apparent diffusion coefficient (ADC), relative cerebral blood volume (rCBV), and cerebral blood flow (CBF) maps were generated with General Electric Healthcare Advantage Workstation for diffusion and perfusion sequence analysis [21, 22]. A 50 mm² ROI was manually drawn on the most significant area of the tumor, avoiding the cystic, necrotic, and hemorrhagic components. A contralateral ROI within the normal-appearing white matter was used for standardization (relative ADC (rADC) and relative CBF

(rCBF)). The size of the FLAIR hypersignal was calculated using the sum of the cranio-caudal, antero-posterior, and transversal axes.

Statistical analysis

Continuous variables are expressed as medians [Q1–Q3] and categorical variables as percentages. Comparisons of patients were undertaken according to their response status group at the end of therapy. “Responders” include patients with complete IPCG response (CR), unconfirmed complete response (CRu), and partial response (PR) in opposition to “Non-responders” (i.e., patients in progression at end of therapy). Patients with missing primary endpoint were excluded from analysis, reasons of missing primary endpoint were described. Continuous variables were compared between groups using non-parametric Wilcoxon rank sum test. Categorical variables were compared with the Fisher’s exact test. For the exploration of correlations between PET and MRI parameters, non-parametric Spearman correlation coefficients were reported. OS function was estimated using Kaplan–Meier method and median OS with 95% confidence interval were provided. Associations between clinical factors and PET parameters with OS was assessed through univariate and multivariate Cox proportional hazard models. Hazard ratios (HR) are presented with 95% confidence intervals (95% CI). Log-linearity was explored for continuous variables. Schoenfeld residuals plots were used to check assumption of proportional hazards. In case of non-respect, a time-varying effect variable was introduced; thus, it was reported on distinct time intervals in order to fulfill the hypothesis overall. Final multivariate model was established using a

backward selection based on likelihood ratio test, where clinical pertinent covariates (sex, age, KPS) were imposed in the model. All tests were two-sided and used a significance level of 0.05. Analyses were computed using R software version 4.1.1 (<https://cran.r-project.org/>).

Results

Clinical study

In our study, out of the total 65 participants, 59 patients underwent a baseline PET-MRI, carried out 90 ± 4 min post-injection of 140 (125–250) MBq of [^{18}F]FDG. The blood glucose level was uniformly recorded at 1.1 ± 0.3 g/L for all these patients. IPCG or clinical response were missing for 5 patients (1 refusal and 4 premature death unrelated to tumor progression within 1 month of diagnosis); thus, 54 patients were analyzed. The flowchart of the study is presented in Fig. 1, and the baseline characteristics of the patient cohort are summarized in Table 1. The median age of the patients was 71.5 years (66–77 years), and the median KPS was 70 (range: 50–100). Of the 43 patients who underwent ophthalmological examination, 3 (7%) were diagnosed with lymphomatous ocular involvement, and 12 out of 41 (28%) with a lumbar puncture had CSF involvement. Diagnostic confirmation was obtained from brain biopsy/surgery in 50 patients and from CSF analysis in 4 patients, with a diagnosis of diffuse large B-cell lymphoma in all cases. Baseline CSF IL-10 assay results were available for 42 out of 54 patients, with a median value of 17.5 pg/mL (range: 0–1731 pg/mL; lower limit of sensitivity: 2.5 pg/mL). Of

Fig. 1 Flowchart. CR, complete response; Cru, unconfirmed complete response; PR, partial response; PD, progressive disease

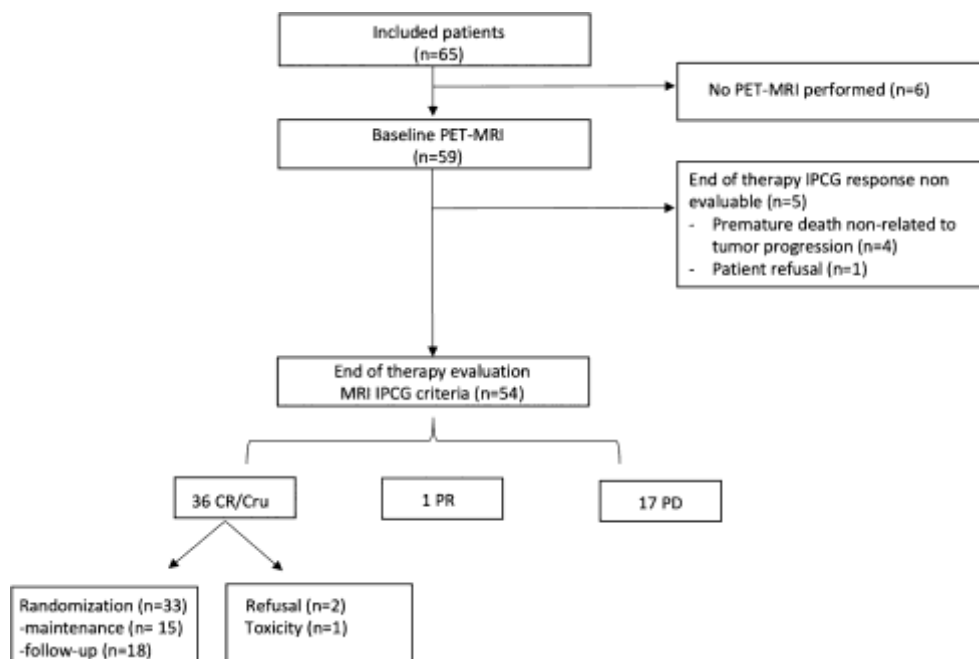


Table 1 Patients and lesion characteristics

	All patients (<i>n</i> = 54)	Responders (<i>n</i> = 37)	Non-responders (<i>n</i> = 17)	<i>p</i>
Sex, <i>n</i> (%)				0.23
Male	21 (38.9%)	12 (32.4%)	9 (53%)	
Female	33 (61.1%)	25 (67.6%)	8 (47%)	
Age (years)				0.81
Median [Q1–Q3]	71.5 [66–77]	72 [64–77]	71 [66–80]	
Karnofsky Performance Status				0.34
Median (IQR)	70 [60–80]	70 [60–80]	60 [60–70]	
Randomization, <i>n</i> (%)				
No	21 (39%)			
Yes—observation arm	18 (33.3%)			
Yes—maintenance arm	15 (27.7%)			
Hydrocortisone dose	(<i>n</i> = 50)	(<i>n</i> = 36)	(<i>n</i> = 14)	0.03
Median [Q1–Q3]	240 [10–300]	240 [70–340]	80 [0–240]	
Treatment completion, <i>n</i> (%)				< 0.0001
4 to 5 RMPV and cytarabine	40 (74.1%)	34 (91.9%)	6 (35.3%)	
< 4 RMPV	14 (25.9%)	3 (8.1%)	11 (64.7%)	
Initial location, <i>n</i> (%)				
Supra-tentorial	37 (68.5%)	25 (67.6%)	12 (70.6%)	0.15
Infra-tentorial	6 (11.1%)	6 (16.2%)	0 (0%)	
Both	11 (20.4%)	6 (16.2%)	5 (29.4%)	
Superficial	11 (20.4%)	7 (18.9%)	4 (23.5%)	0.73
Deep	43 (79.6%)	30 (81.1%)	13 (76.5%)	

the 54 patients, 38 received corticosteroids (median dose of hydrocortisone: 240 mg daily) at the time of PET-MR imaging.

Twenty-six percent of the patients (14/54) did not complete induction chemotherapy due to methotrexate-related renal toxicity (*n* = 5) or progression (*n* = 9). Response to treatment was assessed by MRI at the end of induction therapy using the IPCG criteria: 69% of patients (*n* = 37) were found to have an MRI objective response, including 35 CR, one CRu, and one PR. These patients were considered the responder group, while the non-responder group consisted of 17 (31%) patients with disease progression on MRI or as established by clinical assessment.

At end of therapy, 61% of patients (*n* = 33/54) were randomized in the BLOCAGE trial, with 15 in the maintenance arm and 18 in the follow-up arm.

PET-MRI PCNSL features at diagnosis

PCNSL displayed contrast enhancement in 96% of cases on MRI with 26 (48%) having unifocal lesions and 28 (52%) patients having multifocal disease, with up to 13 solid non-confluent lesions. Forty-four patients (85%) had a homogeneously enhancing lesion, two (4%) had a ring-shaped lesion, and six (11%) had heterogeneous enhancement. The majority of the lesions were located in

supra-tentorial areas (69%), but eleven patients (20%) had tumors infiltrating both supra- and infra-tentorial regions. Most of the lesions were found in deep brain structures, including the thalamus (*n* = 7), striatum (*n* = 13), corpus callosum (*n* = 15), and brainstem (*n* = 10). Lesions showed high diffusion restriction, with median ADCmean and rADCmean values of 679 mm²/s [594–775] and 0.89 [0.76–1], respectively. Spectroscopy analysis revealed a high choline/creatine ratio of 1.7 [1.4–2.2] (Fig. 2). Perfusion was low on dynamic susceptibility contrast (DSC) sequence with median rCBVmax values of 1.57 [1–2.37]. Hyperperfusion was detectable in 41 patients on the CBF parametric map of the ASL sequence.

Fifty-two of the 54 analyzed patients had PET-positive lesions at the time of diagnosis, with the remaining two having a residual sub-centimetric lesion not detected following surgical biopsy. At baseline, MRI scans detected more lesions than [18F]FDG PET in 44% of cases, primarily due to the confluent aspect of multiple lesions on the PET scans. However, no lesions were missed in a patient-based analysis, apart from the two patients with partial resection. PCNSL were found to be highly avid for [18F]FDG, with median values for SUVmax and TBRmax (ratio of SUVmax of the lesion on SUVmean of the contralateral healthy parenchyma) 27.7 [22.8–36] and 5.3 [3.8–7.9], respectively. Of the three patients with ocular involvement, none was

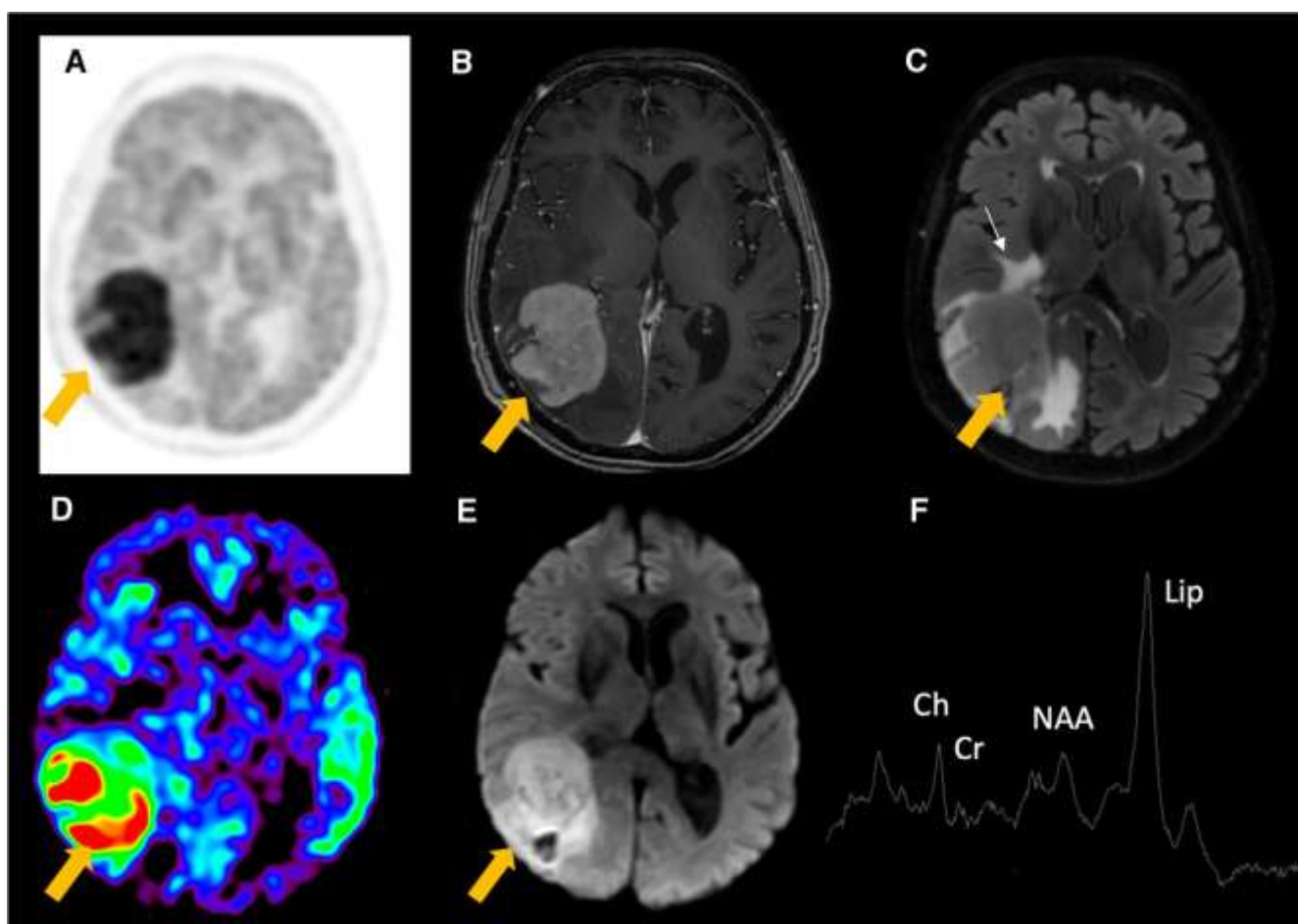


Fig. 2 Primary central nervous system lymphoma in a 76-year-old female patient with typical PET and MRI features: high ^{18}F -FDG avidity ($\text{SUV}_{\text{max}}=44$, $\text{TBR}_{\text{max}}=13$) (A). “Snowball” pattern with intense and homogeneous enhancement on T1-weighted gadolinium-enhanced sequence (B). Axial T2 FLAIR image showed hypointense lesion mass (yellow arrow) and hyperintense surrounding edema

(white arrow) (C). Interestingly, CBF ASL parametric map showed high hyperperfusion (D). Diffusion sequence showed high hyper-signal related to the hypercellularity of the lesion (E). Spectroscopy acquired at short time echo of $\text{TE}=35$ ms in the lesion, showed a very high lipid peak, elevated choline, and important rate of choline/creatinine (F)

detectable on PET. Baseline PET and MRI parameters are detailed in Table 2.

Tumor cellularity, indicated by low ADC_{mean} values, was inversely correlated with ^{18}F -FDG uptake ($\rho = -0.59$ for the ratio of tumor SUV_{peak} to contralateral normal brain SUV_{max}) and volume ($\rho = -0.54$ for MTV and -0.53 for sumMTV). In contrast, perfusion as assessed by CBF on ASL was positively correlated with ^{18}F -FDG uptake ($\rho = 0.61$ for the ratio of tumor SUV_{peak} to contralateral normal brain SUV_{max}). The correlations between these parameters are presented in Fig. 3. Interestingly, our results show that the dose of corticosteroid therapy at the time of the baseline ^{18}F -FDG PET did not exhibit a linear correlation with ^{18}F -FDG avidity or MRI parameters in our study cohort. The coefficients for the relationships between corticoid therapy dose and

these parameters were $\rho = 0.26$ for SUV_{max} , $\rho = 0.01$ for MTV, $\rho = -0.02$ for ADC_{mean} , and $\rho = 0.15$ for contrast-enhancement volume of the target lesion.

Association of baseline ^{18}F -FDG PET-MRI parameters with treatment outcome

Results showed that both high SUV_{mean} of the normal cerebellum and low sumMTV were associated with being in the good-responder group (6.4 [5.7–7.5] vs 5.4 [4.6–6.5], $p = 0.04$ and 5 [1.8–10.8] vs 17.1 [5.3–19.9], $p = 0.01$, respectively). The study also found a trend of a lower $\text{rADC}_{\text{mean}}$ being related to the non-responder group, but statistical significance was not reached. Details can be found in Table 2 and Fig. 4.

Table 2 PET-MRI features according to end-of-induction chemotherapy responses

	All patients (n = 54)	Responders (n = 37)	Non-responders (n = 17)	P
PET parameters (median, Q1–Q3)				
SUVmax	27.7 [22.8–36]	27.8 [20.2–38.4]	26.9 [24.4–30.9]	0.70
TBRmax	5.3 [3.8–7.9]	5.0 [3.3–7.3]	5.7 [4.1–8.2]	0.46
SumSUVmax	37.2 [25.9–58.2]	36.7 [24.3–58.8]	38.1 [26.9–58.1]	0.70
SUVmean cerebellum	6.1 [5–7.2]	6.4 [5.7–7.5]	5.4 [4.6–6.6]	0.04
MILAS score	7 [6–9]	7 [5–9]	9 [7–9]	0.15
MTV of the target volume (n = 52)	5 [2.0–13.3]	3.3 [1.7–7.5]	9.5 [3.2–17.5]	0.04
TLG of the target volume (n = 52)	95 [30.4–249.3]	72.2 [24–135.9]	144 [58.8–263]	0.09
SumMTV (n = 52)	6.3 [2.2–16.0]	5 [1.8–10.8]	17.1 [5.3–19.9]	0.01
MRI parameters (median, Q1–Q3)				
ADCmean (n = 49)	679 [594–775]	699 [637–778]	643 [553–774]	0.18
rADCmean (n = 48)	0.89 [0.76–1]	0.92 [0.83–1]	0.77 [0.7–0.9]	0.06
Enhanced target lesion volume	6.72 [2.5–15.9]	4.7 [2.2–12.2]	11.8 [4–29.4]	0.10
Choline/Creatine ratio (n = 37)	1.7 [1.4–2.2]	1.7 [1.4–2.2]	1.6 [1.3–1.8]	0.44
CBFmax (ASL)	61 [42–81]	61 [38–80]	63 [48–88]	0.31
rCBFmax (ASL) (n = 49)	1.49 [0.95–1.86]	1.38 [0.91–1.85]	1.55 [1.47–1.96]	0.25
FLAIR size	149 [91–202]	138 [78–202]	181 [134–202]	0.38
Biological parameter (median, Q1–Q3)				
IL-10 (n = 42)	17.5 [6–91.5]	16 [6–63]	36 [8–105]	0.44

Bold p values < 0.05, Q1: first quartile and Q3: third quartile

Association of baseline [18F]FDG PET-MRI parameters with survival

Median PFS was 29.6 months (95%CI [9.9;NA]) with a median follow-up of 37 months (95%CI [33.3;47.4]).

Nineteen parameters extracted from MRI and PET tested were tested in univariate analyses. Tumor burden measured by sumMTV on PET and volume of the contrast-enhancement of the target lesion on MRI were significantly associated with poorer PFS (Table 3). A lower rADCmean was

Fig. 3 Correlogram of PET and MRI features of baseline exam. Non-parametric Spearman correlation coefficients (rho) are reported. Positive correlations are in red, negative correlations are in blue, and variables highly correlated are in the darker shades. rCBVmax, relative cerebral blood volume (maximal value); rCBFmax, relative cerebral blood flow (maximal value); rADCmean, relative apparent diffusion coefficient (mean value); TBRmax, tumor-to-background ratio (SUVmax tumoral/SUVmean contralateral)

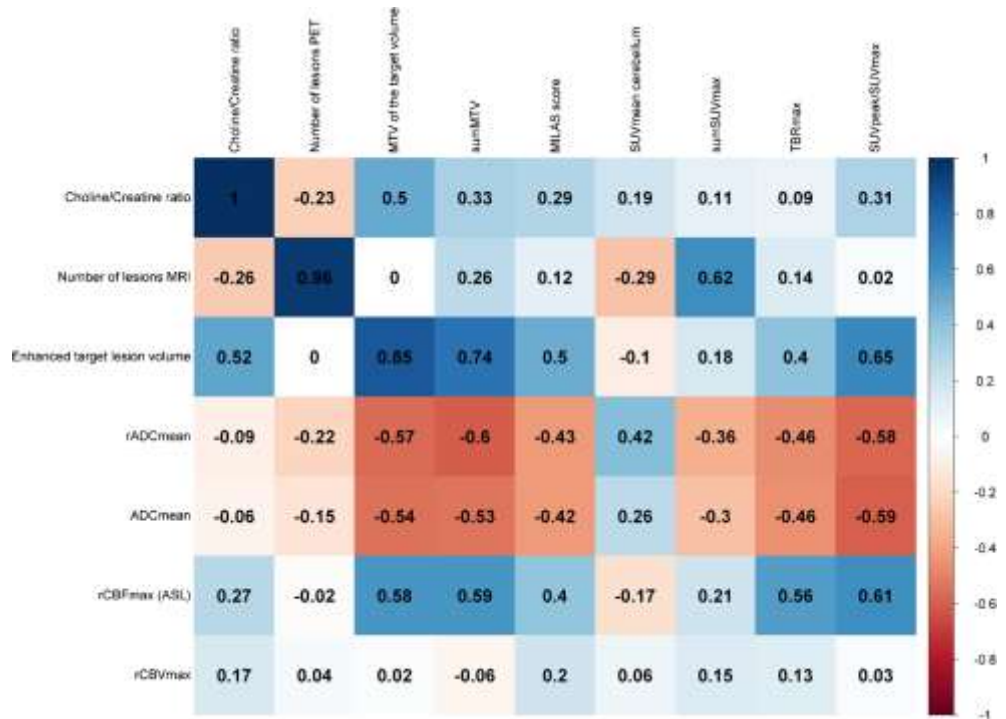
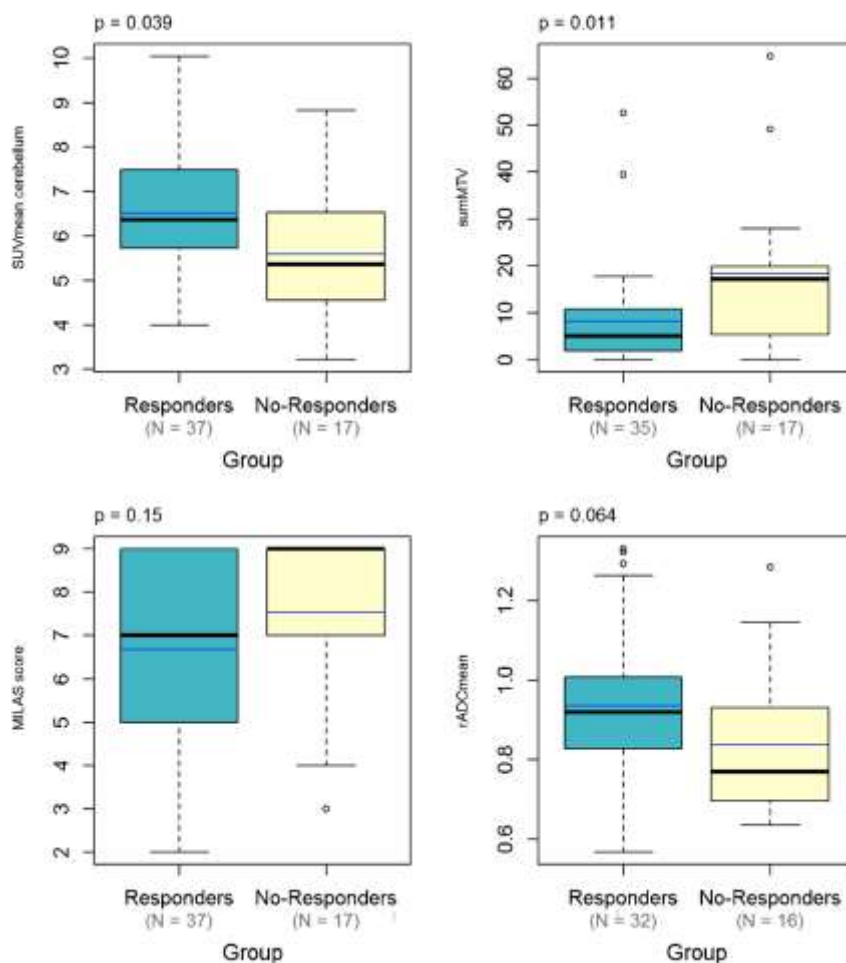


Fig. 4 Distribution of PET and MRI parameters according to the response assessment at end of therapy. Bold lines show the medians, dark blue lines show the means; box limits indicate the 25th and 75th percentiles. Wilcoxon tests *p* value are reported



also significantly associated with poorer PFS. Once adjusted on clinical factors (gender, age, and KPS), sumMTV on PET, volume contrast enhancement of the target lesion on PET and low rADCmean on MRI remain independent significant prognostic factors of poorer PFS (Table 4). In contrast, our

study did not demonstrate a significant association between perfusion parameters (DSC or ASL) and disease prognosis. SUVmean cerebellum seemed to be a time-varying effect covariate with lower SUVmean cerebellum associated with poorer PFS in late progressions in multivariate analysis.

Table 3 Prognostic PET-MRI biomarkers of PFS on univariate analysis

PFS	Univariate (n=54)	
	HR (95% CI)	<i>p</i>
Gender, male (ref female)	1.66 [0.78,3.52]	0.183
Age ≥ 70 (ref < 70 years)	1.23 [0.57,2.66]	0.604
KPS ≥ 70 (ref < 70)	0.84 [0.39,1.81]	0.661
SUVmean cerebellum (increase unit=1)	0.61 [0.41,0.93]*	0.060
	1.08 [0.73,1.61]**	
SumMTV (increase unit=10)	1.36 [1.08,1.72]	0.006
MILAS score	1.04 [0.87;1.25]	0.642
Enhanced target lesion volume (increase unit=10)	1.28 [1.06,1.53]	0.006
rADCmean > 0.885 (ref ≤ 0.885)	0.36 [0.16,0.81]	0.01

*This hazard ratio measured the early effect on one point increase of SUVmean cerebellum on PFS before 6 months of follow-up

**This hazard ratio measured the late effect on one point increase of SUVmean cerebellum on PFS after 6 months of follow-up

Bold *p* values < 0.05

Table 4 Prognostic PET-MRI biomarkers of PFS on multivariate analysis

PFS	Multivariate analysis (PET), <i>n</i> = 52		Multivariate analysis (MRI), <i>n</i> = 46	
	HR (95% CI)	<i>p</i>	HR (95% CI)	<i>p</i>
Gender M (ref F)	2.22 [0.91,5.41]	0.08	1.99 [0.84,4.68]	0.1
Age ≥ 70 (ref < 70 years)	2.61 [1.01,6.89]	0.042	1.12 [0.46,2.74]	0.8
KPS ≥ 70 (ref < 70)	1.11 [0.49,2.52]	0.8	0.90 [0.39,2.07]	0.8
SUVmean cerebellum by increase unit = 1)	0.93 [0.68,1.29]* 0.62 [0.44, 0.88]**	<0.001		
SumMTV (increase unit = 10)	1.46 [1.14,1.86]	0.008		
rADCmean (ref < 0.885)			0.37 [0.16,0.87]	0.018

*This hazard ratio measured the early effect on one point increase of SUVmean cerebellum on PFS before 6 months of follow-up

**This hazard ratio measured the late effect on one point increase of SUVmean cerebellum on PFS after 6 months of follow-up

Bold *p* values < 0.05

The OS at 6 months was 76% [67;89.6], 73% [62;86.3] at 1 year, and 71% [59.1;84.5] at 2 years. In univariate analyses, biomarkers of poorer OS were high sumMTV (HR = 1.66, 95%CI 1.27–2.17 ($p < 0.001$)) and high volume of the enhanced target lesion (HR = 1.52, 95%CI 1.24–1.87 ($p < 0.001$)) (Supplemental Table 1). MILAS score was found to be a time-varying effect covariate with higher score associated to poorer immediate survival (< 1 year): 2.16 [1.11;4.19], $p = 0.02$. The sumMTV and MILAS remained independent PET prognostic factors for OS in multivariate analysis including the clinical variables (gender, age, KPS), with a C-index of 0.84 [0.76–0.92]. Enhanced target lesion volume was found to be an independent MRI prognostic factor for OS in multivariate analysis including the clinical variables (gender, age, KPS), with a C-index of 0.77 [0.65–0.90] (Supplemental Table 2).

In this study, clinical variables such as age, gender, and KPS were not found to have a significant impact on OS in univariate analysis, but age (and gender in PET model) became significant predictors of OS in final multivariate models.

Discussion

This prospective study assessed the value of simultaneous [18F]FDG PET-MRI to provide prognostic biomarkers for PCNSL in immunocompetent elderly patients homogeneously treated with an HD-MTX-based CT regimen at first-line. Currently, age and KPS are the only well-established clinical markers, but they are not sufficient in guiding the management of PCNSL in elderly patients, which has a poor prognosis [23]. The use of hybrid PET-MR system in this study provided the opportunity to compare multimodality data with optimal temporal synchronization and spatial co-registration between both modalities [24]. We found that larger tumor volume on PET-MRI and lower [18F]FDG

uptake in the normal cerebellum were associated with non-response to HD-MTX-based CT. Additionally, the study confirmed the impact of MILAS on OS once adjusted on clinical factor. Association between high sum MTV and poorer prognostic (both PFS and OS) was also confirmed in our study, the largest cohort published so far.

In our cohort, PCNSL presented as homogeneously enhanced lesions, with the majority involving the supratentorial brain parenchyma, a pattern in line with previous reports of immunocompetent patients [25–27]. We also reported low diffusivity, a hallmark of PCNSL due to its hypercellularity [28, 29]. This study also conformed the high avidity of PCNSL for [18F]FDG in immunocompetent patients, with lesions metabolism five times higher than that of the healthy tissue. This is slightly higher than that reported in previous literature [10]. The high sensitivity of the PET-MRI system equipped new generation silicon photomultiplier detectors, the quantification of delayed PET images acquired between 90 and 110 min post-injection (to increase tumor contrast) and different methods for quantifying the ratios between studies (using either SUVmean or SUVmax as for uptake reference) may explained these differences with previous published data and limit direct comparison with prior studies [30]. Rapid regression of primary brain lymphomas under corticosteroid therapy has been reported in a number of cases on MRI [31]. For [18F]FDG PET, studies on corticosteroid's impact have been conducted on small sample sizes with contradictory results and interestingly, our study showed no impact of corticosteroids on PET measurements.

We found an inverse correlation between tumor [18F]FDG uptake and volume on baseline PET and the ADC value on MRI. Previous studies have described the high cellularity of PCNSL on diffusion-weighted imaging but only Zhou et al. previously reported an inverse correlation between ADC and uptake intensity in a cohort of 28

patients [31–33]. In our study, we also observed a positive correlation between cerebral blood flow measured using ASL perfusion and [18F]FDG PET uptake, which to our knowledge, has not been reported in the literature. PCNSL typically lacks neoangiogenesis, as reflected by moderately increased relative cerebral blood volume (rCBV) values in T2* dynamic susceptibility contrast perfusion studies [34]. Our study found a low rCBVmax ratio (2.6 ± 3.9), in the range of values previously reported in the literature [31]. However, ASL is a perfusion technique that allows absolute quantification of CBF without the use of contrast media and which may provide additional diagnostic value in brain tumors [35].

Our results showed that a lower cerebellar metabolism, as measured by SUVmean, was associated with nonresponse to induction HD-MTX-based CT, poorer PFS, and poorer OS. Hanaoka et al. previously demonstrated that [18F]FDG is diverted from the brain to the tumor tissue in systemic lymphoma and showed that SUVmax of the cerebral cortex and the cerebellum were significantly different between the lymphoma and control subjects and correlated with tumor volume [36]. To evaluate the possible competition between the tumor and normal brain FDG uptake, we analyzed the correlation of cerebellum SUVmean with PET volume, intensity, and tumor location, but no significant relationships was found despite a lack of power may not be excluded. Furthermore, the SUVmax of the normal contralateral brain region was found to not significantly differentiate between non responders and responders. In systemic lymphoma, the mean SUV of the cerebellum has been shown to correlate with tumor burden, making it a useful indicator for tumor burden measurement. To further explore the relationship between healthy tissue metabolism and survival, univariate models incorporating healthy areas such as cerebellar SUVmean or normal liver SUVmean have been studied, with findings indicating that these parameters correlate with survival [37, 38].

In our study of patients with CNS lymphoma, we also found that higher volume on PET or enhanced T1-weighted MRI sequence was a second independent biomarker of poor response to induction therapy and poor survival. Although tumor burden is well established as a prognostic factor in systemic lymphomas, data in the CNS setting is limited [39–42]. Our findings align with previous research by Albano et al., who reported in a study of 52 patients that patients with larger MTV had shorter PFS and OS [16].

Various studies have investigated the association between tumor [18F]FDG avidity and response to treatment or survival in lymphoma, with some studies reporting a positive association and others reporting that only TLG provides prognostic information [17, 18, 43–45]. In our cohort of 54 patients, the SUVmax failed to predict response to end of therapy. However, the MILAS score emerged as a prognostic

biomarker of OS. This score is a simple visual assessment system that uses the cerebellar [18F]FDG uptake as a reference to semi-quantitatively evaluate the relative metabolism of the tumor compared to the reference region [18]. However, it should be noted that the MILAS score is based on a cerebellar uptake ratio and therefore may reflect the dominant role of normal cerebellar uptake on survival prognosis.

With regards to the utilization of multimodal MRI, our study revealed that ADCmean was the only parameter which correlated with overall survival, with a lower ADCmean being linked to a higher risk of death. This result aligns with previous studies on primary central nervous system lymphoma (PCNSL) [9, 12, 13, 46]. Barajas found an inverse relationship between ADC values and tumor cell density, which was determined to be a significant factor in determining the outcome [13]. In contrast, our study showed that perfusion parameters derived from DSC or ASL did not emerge as prognostic biomarkers, unlike previous findings, using dynamic-contrast-enhanced perfusion, that indicated a lower baseline volume transfer constant (K^{trans}) was associated with a poorer prognosis [9, 11, 12]. Valles et al. hypothesized that a low tumor rCBV in PCNSL may reflect a lack of tumor angiogenesis and a reduction in the number of vessels capable of delivering intravenous methotrexate to the tumor, which could account for the worse prognosis in these patients [12]. Further research is required to confirm the prognostic role of perfusion and its kinetic evolution under treatment.

Our study is limited by its relatively small sample size of patients. Nevertheless, this is the largest prospective study investigating the role of PET imaging in PCNSL so far. It benefited from a prospective design and the utilization of a new hybrid PET-MRI system, which facilitated precise registration of both modalities. Of note, we have included two patients who had undergone partial excisions at baseline, which could have potentially influenced the outcome analysis. However, it's important to note that surgery is typically not recommended for patients with PCNSL and its impact on the prognosis remains uncertain. Our analysis yielded statistically consistent results, irrespective of the inclusion or exclusion of these two specific patients. The only exception was the rADC, which showed a trend towards significance for PFS in the multivariate analysis ($p=0.07$), as opposed to a statistically significant result ($p=0.018$) observed otherwise.

Another potential limitation of the study is that patients received different treatment regimens after induction chemotherapy, but the first 5 months of therapy was consistent for all participants and our primary endpoint was the response at the end of therapy. Furthermore, identification of prognostic factors will always be contaminated by subsequent lines of treatments. The data analysis for the BLOCAGE 01 clinical trial is scheduled to be conducted

in 2024, providing further insights into the study's findings. Interestingly, the KPS did not emerge as a prognostic factor in our cohort. This finding could potentially be attributed to the choice of our study's KPS threshold set at 70% for the survival analysis. However, it is worth noting that in line with previous literature a KPS threshold of $\geq 80\%$ emerged as a significant predictor of outcome in our study [47, 48]. Several other non-invasive baseline biomarkers have shown potential impact on outcome prediction, such as IL-10 or circulating tumor DNA, and future studies should further investigate potential combined biomarkers (biological and imaging) to enhance the accuracy of patient stratification [49, 50].

Conclusion

Our study is one of the largest to investigate imaging biomarkers in patients with PCNSL and the first to benefit from a PET-MRI scan at baseline. Our findings highlight the significance of tumor volume and cerebellum uptake as key indicators of treatment response. Furthermore, our results show that baseline tumor volume and MILAS score have an impact on overall survival. Further studies are needed to evaluate the potential role of [18F]FDG PET-MRI in therapeutic assessment.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00259-023-06334-w>.

Acknowledgements We would like to thank Dr. Inna Gertsenshteyn for proofreading our manuscript before submission. We would like to thank the Assistance publique—hôpitaux de Paris for the support given to this clinical study (PHRC LOCALYZE).

Author contribution Professors Kas and Galanaud, as well as Drs. Houillier, Rozenblum, V. Causse Lemerrier, and M. Soret, contributed to the study's conception and design. Dr. Rozenblum and Prof. Galanaud performed the MRI images data collection, while P. Naggara, Dr. Rozenblum, and Prof. Kas performed the PET images data collection. Drs. Soussain, Ursu, Willems, and Houillier, as well as Prof. Choquet and Hoang-Xuan, recruited the patients. A. Baptiste and L. Belin, in collaboration with Dr. Rozenblum, performed the statistical analysis. All authors reviewed and approved the final version of the manuscript.

Data Availability Original data will be made available upon reasonable request by e-mail.

Declarations

Ethics approval The study was approved by the protection of persons committee (2014-002597-37) and the French National Agency for the Safety of Medicines and Health Products (ANSM), and abided by the principles of the Declaration of Helsinki.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent for publication The authors affirm that human research participants provided informed consent for publication of the images in Fig. 2.

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

References

1. Lauw MIS, Lucas C-HG, Ohgami RS, Wen KW. Primary central nervous system lymphomas: a diagnostic overview of key histomorphologic, immunophenotypic, and genetic features. *Diagnostics*. 2020;10:1076.
2. Bauchet L, Rigau V, Mathieu-Daudé H, Figarella-Branger D, Hugues D, Palusseau L, et al. French brain tumor data bank: methodology and first results on 10,000 cases. *J Neurooncol*. 2007;84:189–99.
3. Houillier C, Soussain C, Ghesquière H, Soubeyran P, Chinot O, Taillandier L, et al. Management and outcome of primary CNS lymphoma in the modern era: An LOC network study. *Neurology*. 2020;94:e1027–39.
4. Morales-Martinez A, Lozano-Sanchez F, Duran-Peña A, Hoang-Xuan K, Houillier C. Primary central nervous system lymphoma in elderly patients: management and perspectives. *Cancers (Basel)*. 2021;13(14):3479. <https://doi.org/10.3390/cancers13143479>.
5. Ferreri AJM, Blay J-Y, Reni M, Pasini F, Spina M, Ambrosetti A, et al. Prognostic scoring system for primary CNS lymphomas: the International Extranodal Lymphoma Study Group experience. *J Clin Oncol*. 2003;21:266–72.
6. Abrey LE, Ben-Porat L, Panageas KS, Yahalom J, Berkey B, Curran W, et al. Primary central nervous system lymphoma: the Memorial Sloan-Kettering Cancer Center prognostic model. *J Clin Oncol*. 2006;24:5711–5.
7. Hernández-Verdín I, Kirasic E, Wienand K, Mokhtari K, Eimer S, Loiseau H, Rousseau A, Paillassa J, Ahle G, Lerintu F, Uro-Coste E, Oberic L, Figarella-Branger D, Chinot O, Gauchotte G, Taillandier L, Marolleau JP, Polivka M, Adam C, Ursu R, Schmitt A, Barillot N, Nichelli L, Lozano-Sánchez F, Ibañez-Julιά MJ, Peyre M, Mathon B, Abada Y, Charlotte F, Davi F, Stewart C, de Reyniès A, Choquet S, Soussain C, Houillier C, Chapuy B, Hoang-Xuan K, Alentorn A. Molecular and clinical diversity in primary central nervous system lymphoma. *Ann Oncol*. 2023;34(2):186–99. <https://doi.org/10.1016/j.annonc.2022.11.002>.
8. Hoang-Xuan K, Deckert M, Ferreri AJM, Furtner J, Gallego Perez-Larraya J, Henriksson R, et al. European Association of Neuro-Oncology (EANO) guidelines for treatment of primary central nervous system lymphoma (PCNSL). *Neuro Oncol*. 2023;25:37–53.
9. Fu F, Sun X, Li Y, Liu Y, Shan Y, Ji N, et al. Dynamic contrast-enhanced magnetic resonance imaging biomarkers predict chemotherapeutic responses and survival in primary central-nervous-system lymphoma. *Eur Radiol*. 2021;31:1863–71.
10. Gupta T, Manjali JJ, Kannan S, Purandare N, Rangarajan V. Diagnostic performance of pretreatment 18F-fluorodeoxyglucose positron emission tomography with or without computed tomography in patients with primary central nervous system lymphoma: updated systematic review and diagnostic test accuracy meta-analyses. *Clin Lymphoma Myeloma Leuk*. 2021;21:497–507.
11. Hatzoglou V, Oh JH, Buck O, Lin X, Lee M, Shukla-Dave A, et al. Pretreatment dynamic contrast-enhanced MRI biomarkers correlate with progression-free survival in primary central nervous system lymphoma. *J Neurooncol*. 2018;140:351–8.


12. Valles FE, Perez-Valles CL, Regalado S, Barajas RF, Rubenstein JL, Cha S. Combined diffusion and perfusion MR imaging as biomarkers of prognosis in immunocompetent patients with primary central nervous system lymphoma. *AJNR Am J Neuroradiol*. 2013;34:35–40.
13. Barajas RF, Rubenstein JL, Chang JS, Hwang J, Cha S. Diffusion-weighted MR imaging derived apparent diffusion coefficient is predictive of clinical outcome in primary central nervous system lymphoma. *AJNR Am J Neuroradiol*. 2010;31:60–6.
14. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. 2014;32:3059–68.
15. Gupta M, Gupta T, Purandare N, Rangarajan V, Puranik A, Moiyadi A, et al. Utility of flouro-deoxy-glucose positron emission tomography/computed tomography in the diagnostic and staging evaluation of patients with primary CNS lymphoma. *CNS Oncol*. 2019;8:CNS46.
16. Albano D, Bertoli M, Battistotti M, Rodella C, Statuto M, Giubini R, et al. Prognostic role of pretreatment 18F-FDG PET/CT in primary brain lymphoma. *Ann Nucl Med*. 2018;32:532–41.
17. Krebs S, Mauguen A, Yildirim O, Hatzoglou V, Francis JH, Schaff LR, et al. Prognostic value of [18F]FDG PET/CT in patients with CNS lymphoma receiving ibrutinib-based therapies. *Eur J Nucl Med Mol Imaging*. 2021;48:3940–50.
18. Kasenda B, Haug V, Schorb E, Fritsch K, Finke J, Mix M, et al. 18F-FDG PET is an independent outcome predictor in primary central nervous system lymphoma. *J Nucl Med*. 2013;54:184–91.
19. Barajas RF, Politi LS, Anzalone N, Schöder H, Fox CP, Boxerman JL, et al. Consensus recommendations for MRI and PET imaging of primary central nervous system lymphoma: guideline statement from the International Primary CNS Lymphoma Collaborative Group (IPCG). *Neuro Oncol*. 2021;23:1056–71.
20. Abrey LE, Batchelor TT, Ferreri AJM, Gospodarowicz M, Pulczynski EJ, Zucca E, et al. Report of an international workshop to standardize baseline evaluation and response criteria for primary CNS lymphoma. *J Clin Oncol*. 2005;23:5034–43.
21. Rosen BR, Belliveau JW, Vevea JM, Brady TJ. Perfusion imaging with NMR contrast agents. *Magn Reson Med*. 1990;14:249–65.
22. Ostergaard L, Weisskoff RM, Chesler DA, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis. *Magn Reson Med*. 1996;36:715–25.
23. Rozenblum L, Houillier C, Soussain C, Bertaux M, Choquet S, Galanaud D, et al. Role of positron emission tomography in primary central nervous system lymphoma. *Cancers (Basel)*. 2022;14:4071.
24. Pyatigorskaya N, Habert M-O, Rozenblum L. Contribution of PET-MRI in brain diseases in clinical practice. *Curr Opin Neurol*. 2020;33:430–8.
25. Mansour A, Qandeel M, Abdel-Razeq H, Abu Ali HA. MR imaging features of intracranial primary CNS lymphoma in immune competent patients. *Cancer Imaging*. 2014;14:22.
26. Slone HW, Blake JJ, Shah R, Guttikonda S, Bourekas EC. CT and MRI findings of intracranial lymphoma. *AJR Am J Roentgenol*. 2005;184:1679–85.
27. Mohile NA, Abrey LE. Primary central nervous system lymphoma. *Neurol Clin*. 2007;25(1193–207):xi.
28. Haldorsen IS, Espeland A, Larsson E-M. Central nervous system lymphoma: characteristic findings on traditional and advanced imaging. *AJNR Am J Neuroradiol*. 2011;32:984–92.
29. Zacharia TT, Law M, Naidich TP, Leeds NE. Central nervous system lymphoma characterization by diffusion-weighted imaging and MR spectroscopy. *J Neuroimaging*. 2008;18:411–7.
30. Soret M, Maisonobe J-A, Desarnaud S, Bergeret S, Causse-Lemercier V, Berenbaum A, et al. Ultra-low-dose in brain 18F-FDG PET/MRI in clinical settings. *Sci Rep*. 2022;12:15341.
31. Voultsinou D, Mantatzis M, Gerukis T, Heva A, Birbilis T, Prassopoulos P. Magnetic resonance imaging patterns in central nervous system lymphomas: a pictorial review. *Clin Imaging*. 2021;78:1–7.
32. Chiavazza C, Pellerino A, Ferrio F, Cistaro A, Soffietti R, Rudà R. Primary CNS Lymphomas: challenges in Diagnosis and Monitoring. *Biomed Res Int*. 2018;2018:3606970.
33. Zhou W, Wen J, Hua F, Xu W, Lu X, Yin B, et al. 18F-FDG PET/CT in immunocompetent patients with primary central nervous system lymphoma: differentiation from glioblastoma and correlation with DWI. *Eur J Radiol*. 2018;104:26–32.
34. Blasel S, Vorwerk R, Kiyose M, Mittelbronn M, Brunnberg U, Ackermann H, et al. New MR perfusion features in primary central nervous system lymphomas: pattern and prognostic impact. *J Neurol*. 2018;265:647–58.
35. You G, Wu H, Lei B, Wan X, Chen S, Zheng N. Diagnostic accuracy of arterial spin labeling in differentiating between primary central nervous system lymphoma and high-grade glioma: a systematic review and meta-analysis. *Expert Rev Anticancer Ther*. 2022;22:763–71.
36. Hanaoka K, Hosono M, Shimono T, Usami K, Komeya Y, Tsuchiya N, et al. Decreased brain FDG uptake in patients with extensive non-Hodgkin's lymphoma lesions. *Ann Nucl Med*. 2010;24:707–11.
37. Morland D, Triumbari EKA, Maiolo E, Cuccaro A, Treglia G, Hohaus S, et al. Healthy organs uptake on baseline 18F-FDG PET/CT as an alternative to total metabolic tumor volume to predict event-free survival in classical Hodgkin's lymphoma. *Front Med (Lausanne)*. 2022;9:913866.
38. Godard F, Durot E, Durot C, Hoeffel C, Delmer A, Morland D. Cerebellum/liver index in pretherapeutic 18F-FDG PET/CT as a predictive marker of progression-free survival in follicular lymphoma treated by immunochemotherapy and rituximab maintenance. *Medicine*. 2022;101:e28791.
39. Vercellino L, Cottreau A-S, Casasnovas O, Tilly H, Feugier P, Chartier L, et al. High total metabolic tumor volume at baseline predicts survival independent of response to therapy. *Blood*. 2020;135:1396–405.
40. Ceriani L, Gritti G, Cascione L, Piroso MC, Polino A, Ruberto T, Stathis A, Bruno A, Moccia AA, Giovannella L, Hayoz S, Schär S, Dirnhofer S, Rambaldi A, Martinelli G, Mamot C, Zucca E. SAKK38/07 study: integration of baseline metabolic heterogeneity and metabolic tumor volume in DLBCL prognostic model. *Blood Adv*. 2020;4(6):1082–92. <https://doi.org/10.1182/bloodadvances.2019001201>.
41. Moskowitz AJ, Schöder H, Gavane S, Thoren KL, Fleisher M, Yahalom J, et al. Prognostic significance of baseline metabolic tumor volume in relapsed and refractory Hodgkin lymphoma. *Blood*. 2017;130:2196–203.
42. Huntoon K, Makary MS, Shah VS, Aquino A, Pandya V, Giglio P, Slone HW, Elder JB. Pretreatment findings on magnetic resonance imaging in primary central nervous system lymphoma may predict overall survival duration. *Neuroradiol J*. 2023. <https://doi.org/10.1177/19714009231154681>.
43. Okuyucu K, Alagoz E, Ince S, Ozaydin S, Arslan N. Can metabolic tumor parameters on primary staging 18F-FDG PET/CT aid in risk stratification of primary central nervous system lymphomas for patient management as a prognostic model? *Rev Esp Med Nucl Imagen Mol (Engl Ed)*. 2018;37:9–14.
44. Inoue A, Ohnishi T, Kohno S, Matsumoto S, Nishikawa M, Ohue S, et al. Prognostic significance of immunohistochemical subtypes based on the stage of B-cell differentiation in primary CNS lymphoma. *Int J Clin Exp Pathol*. 2019;12:1457–67.

45. Kawai N, Okubo S, Miyake K, Maeda Y, Yamamoto Y, Nishiyama Y, et al. Use of PET in the diagnosis of primary CNS lymphoma in patients with atypical MR findings. *Ann Nucl Med*. 2010;24:335–43.
46. Zhang D, Hu L-B, Henning TD, Ravarani EM, Zou L-G, Feng X-Y, et al. MRI findings of primary CNS lymphoma in 26 immunocompetent patients. *Korean J Radiol*. 2010;11:269–77.
47. Curry LD, Munker R, Li N, Yan D, Pryor P, Nozad S, et al. Performance status, comorbidities, and cycles of methotrexate exert the greatest influence on outcomes of primary and secondary CNS lymphomas: the Lexington experience. *Ann Hematol*. 2023;102:141–54.
48. Wei L, Gao Y, Prochazka KT, Liu R, Wang L, Liu B, et al. A novel prognostic model based on pretreatment serum albumin and ECOG PS for primary CNS lymphoma: an international, multicenter study. *J Neurooncol* [Internet]. 2023 [cited 2023 Jun 2]; Available from: <https://link.springer.com/https://doi.org/10.1007/s11060-023-04337-z>
49. Mutter JA, Alig SK, Esfahani MS, Lauer EM, Mitschke J, Kurtz DM, Kühn J, Bleul S, Olsen M, Liu CL, Jin MC, Macaulay CW, Neidert N, Volk T, Eisenblaetter M, Rauer S, Heiland DH, Finke J, Duyster J, Wehrle J, Prinz M, Illerhaus G, Reinacher PC, Schorb E, Diehn M, Alizadeh AA, Scherer F. Circulating Tumor DNA Profiling for Detection, Risk Stratification, and Classification of Brain Lymphomas. *J Clin Oncol*. 2023;41(9):1684–94. <https://doi.org/10.1200/JCO.22.00826>.
50. Sasayama T, Nakamizo S, Nishihara M, Kawamura A, Tanaka H, Mizukawa K, et al. Cerebrospinal fluid interleukin-10 is a potentially useful biomarker in immunocompetent primary central nervous system lymphoma (PCNSL). *Neuro Oncol*. 2012;14:368–80.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Laura Rozenblum^{1,2}  · Damien Galanaud^{2,3} · Caroline Houillier⁴ · Carole Soussain⁵ · Amandine Baptiste⁶ · Lisa Belin⁶ · Véronique Edeline⁷ · Philippe Naggara¹ · Marine Soret^{1,2} · Valérie Causse-Lemercier¹ · Lise Willems⁸ · Sylvain Choquet⁹ · Renata Ursu¹⁰ · Khê Hoang-Xuan⁴ · Aurélie Kas^{1,2}

✉ Laura Rozenblum
Laura.rozenblum@gmail.com

¹ Department of Nuclear Medicine, Groupe Hospitalier Pitié-Salpêtrière, APHP, Sorbonne Université, Paris, France

² Sorbonne Université, INSERM, CNRS, Laboratoire d'Imagerie Biomédicale, LIB, Paris, France

³ Department of Neuroradiology, Groupe Hospitalier Pitié-Salpêtrière, APHP, Sorbonne Université, Paris, France

⁴ Department of Neurology 2 Mazarin, APHP, Hôpitaux Universitaires La Pitié Salpêtrière - Charles Foix, Sorbonne Université, Inserm, CNRS, UMR S 1127, Institut du Cerveau, ICM, Paris, France

⁵ Department of Hematology, Institut Curie, Site Saint-Cloud and INSERM U932 Institut Curie, Université PSL, 75005 Paris, France

⁶ Department of Public Health, Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie Et de Santé Publique, AP-HP, Hôpitaux Universitaires Pitié-Salpêtrière - Charles Foix, Paris, France

⁷ Department of Nuclear Medicine, Institut Curie, St-Cloud, France

⁸ Department of Hematology, Cochin Hospital, APHP, Paris, France

⁹ Department of Hematology, Groupe Hospitalier Pitié-Salpêtrière, APHP, Sorbonne Université, Paris, France

¹⁰ Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpital Saint-Louis, Service de Neurologie, Paris, France

Supplemental Table 1. Prognostic PET-MRI biomarkers of overall survival on univariate analysis

	Univariate	
	HR (95% CI)	p
Overall survival		
Sex M (ref = F)	1.61 [0.60,4.28]	0.339
Age >= 70 (ref < 70 years)	3.13 [0.89,10.98]	0.060
KPS >= 70 (ref < 70)	0.59 [0.22,1.59]	0.294
SUVmean cerebellum	0.69 [0.47,1.02]	0.060
SumMTV (increase unit =10)	1.66 [1.27,2.17]	< 0.001
MILAS score (by 1-point increase)	2.16 [1.11;4.19] *	0.020
	0.93 [0.63;1.36]**	
Enhanced target lesion volume (increase unit =10)	1.52 [1.24,1.87]	< 0.001
ADCmean (increase unit =100)	0.65 [0.41,1.03]	0.063

*This hazard ratio measured the early effect on one-point increase of MILAS score on OS before 1 year of follow-up.

**This hazard ratio measured the late effect on one-point increase of MILAS score on OS after 1 year of follow-up.

Supplemental Table 2. Prognostic PET-MRI biomarkers of overall survival on multivariate analysis

	Multivariate analysis (PET) n=52		Multivariate analysis (MRI) n=47	
	HR (95% CI)	p	HR (95% CI)	p
Overall survival				
Sex M (ref = F)	3.26 [1.03,10.29]	0.044	2.93 [0.96,8.91]	0.06
Age (ref < 70 years)	7.97 [1.76,36.11]	0.002	6.73 [1.34,33.73]	0.007
KPS (ref < 70)	0.68 [0.23, 1.99]	0.48	0.89 [0.30,2.60]	0.83
SumMTV (increase unit =10)	1.47 [1.04,2.07]	0.036		
MILAS score (increase unit=1)	2.20 [0.98;4.93] *	0.049		
	0.90 [0.55, 1.45]**			
Enhanced target lesion volume (increase unit =10)			1.66 [1.30,2.13]	< 0.001

*This hazard ratio measured the early effect on one-point increase of MILAS score on OS before 1 year of follow-up.

**This hazard ratio measured the late effect on one-point increase of MILAS score on OS after 1 year of follow-up.

Chapitre 4. Prise en charge et évaluation thérapeutique du LPSNC

I. Traitement

Malgré d'importantes avancées thérapeutiques, le pronostic du LPSNC reste défavorable. La chirurgie n'est pas préconisée en raison de la nature infiltrante des lésions et de leur localisation souvent profonde, ce qui complique la procédure opératoire. De plus, les données de la littérature ne montrent pas de bénéfice de la chirurgie en termes de survie par rapport à d'autres approches [84,111].

Un autre obstacle significatif à l'efficacité des traitements dans le LPSNC est la barrière hémato-encéphalique (BHE), qui limite la biodisponibilité intracérébrale de nombreux agents thérapeutiques. En 2023, la Société Européenne de Neuro-oncologie (EANO) a publié de nouvelles recommandations pour la prise en charge du LPSNC [84]. Le traitement du LPSNC se divise en phases d'induction et de consolidation. Pour l'induction, le traitement de référence repose sur une chimiothérapie intraveineuse à base de méthotrexate à haute dose, qui traverse efficacement la BHE [1,83]. Ce traitement peut être combiné avec des poly-chimiothérapies, telles que l'association procarbazine, vincristine et cytarabine dans le protocole MPVA, ou l'aracytine et le thiotepa, dans le cadre du schéma MATRix [112–117]. Plusieurs études ont montré que l'ajout du Rituximab au protocole MPV-A (protocole R-MPVA) avait permis une amélioration du taux de rémission complète [112,118,119]. Cependant, ces résultats ont été remis en question par des études plus récentes, ne montrant qu'un bénéfice plus mitigé à l'utilisation du Rituximab [120,121]. Ainsi les recommandations EANO 2023 restent prudentes quant à l'utilisation de cet anticorps [122].

La radiothérapie cérébrale *in toto*, traditionnellement recommandée chez les sujets jeunes comme traitement de consolidation, est associée à un risque élevé de neurotoxicité [123–125]. Elle est donc déconseillée chez les patients âgés. Des traitements de consolidation par chimiothérapie sont à l'étude, comme dans l'essai de phase III BLOCAGE-01 (PI : K. Hoang-Xuan), qui combine Rituximab, Méthotrexate et Temozolomide. La greffe de cellules souches autologues (ASCT) constitue une option pour les patients jeunes, en bonne condition organique, comme l'ont suggéré les essais IELSG-32 et PRECIS [126,127]. Un essai américain de phase II est en cours pour évaluer le Nivolumab en traitement de maintenance chez les patients avec

ADN circulant persistant dans le liquide céphalo-rachidien (LCR) après la chimiothérapie d'induction (NCT04401774).

En cas de rechute, les options thérapeutiques actuelles montrent un bénéfice limité. L'étude AcSé sur le Pembrolizumab a rapporté un taux de réponse de 26% et une SSP de 2.6 mois [128]. D'autres traitements ciblés, comme les inhibiteurs de la tyrosine kinase de Bruton (BTK), explorés dans l'essai de phase II GBR121, sont également envisagés comme thérapies potentielles [129–131]. Parallèlement, le traitement par cellules CAR-T émerge comme une option prometteuse, même si les données préliminaires restent limitées et que des préoccupations subsistent concernant leur toxicité neurologique [132,133].

II. Evaluation thérapeutique

Critères IRM

L'évaluation thérapeutique des LPSNC repose sur les critères du Groupe International pour les Lymphomes Primaires du SNC (IPCG) (Tableau 4.1). Ces critères sont principalement basés sur l'imagerie IRM T1 après injection de produit de contraste mais intègrent également une composante clinico-biologique avec un examen ophtalmologique, la prise en compte de l'administration de stéroïdes, ainsi qu'une analyse cytologique du liquide céphalorachidien (LCR). La quantification de la charge tumorale en IRM s'effectue conventionnellement en mesurant les deux plus grands diamètres de la lésion rehaussée par le contraste [6]. Cette approche soulève des préoccupations en matière de fiabilité, en raison de la variabilité inter-observateurs et des fluctuations intrinsèques liées aux paramètres d'acquisition, à la dose de gadolinium administrée au moment de l'injection. Dans l'objectif d'harmoniser les pratiques radiologiques et de pallier ces fluctuations, Barajas et al. ont récemment proposé des recommandations sur le protocole IRM spécifique à l'évaluation des LPSNC [5]. Néanmoins, la problématique des LPSNC atypiques, qui ne captent pas le contraste, demeure quant à leur évaluation par cette méthode [134].

	IRM	Stéroïdes	Ophtalmologie	LCR
RC (rémission complète)	Pas de prise de contraste	Arrêtés depuis au moins 2 semaines (sauf si indiqués pour une autre pathologie)	Normal → Pas d'évaluation ophtalmologique si OPH1 négative	Pas de cellule tumorale → Pas de PL si PL1 négative
Rcu (rémission complète incertaine)	RC avec prise de contraste < 3 mm au niveau de la biopsie ou sur site avec microsaignement <i>Si pas d'évolution dans le temps = RC</i>	RC radiologique mais stéroïdes non arrêtés	Anomalies mineures mais non spécifiques d'une origine tumorale <i>Si pas d'évolution = RC</i>	Pas de cellule tumorale
RP (rémission partielle)	Diminution de plus de 50% du volume des prises de contraste		Diminution de l'infiltrat du vitré, de la rétine ou du nerf optique mais persistance	Cellules malignes ou suspectes persistantes mais diminution de 50% du volume des prises de contraste cérébrales
Stabilité	Réponse inférieure à la réponse partielle mais sans critère de progression			
Progression	Augmentation de plus de 25% du volume des prises de contrastes Apparition de nouvelles lésions (SNC ou ailleurs)		Progression des lésions (vitré, rétine, nerf optique)	Apparition de cellules tumorales

Tableau 4.1 Score IPCG [6]

Actuellement, la séquence T2-FLAIR n'est pas incluse dans les critères standard d'évaluation des LPSNC. Estephan et al., dans une étude impliquant 47 patients, ont observé que 62% des patients présentaient une augmentation de l'hypersignal FLAIR pendant le traitement, avec des modifications d'autant plus précoces et importantes que le patient été traité par Rituximab [135]. En revanche, le lien entre la modification du signal FLAIR et la survie n'était pas évalué et cette question reste ouverte pour les prochaines études.

Depuis quelques années la pertinence des critères IPCG pour évaluer la réponse des LPSNC a été remise en question. L'étude de Van der Meulen et al. a par exemple montré que les patients en réponse complète et ceux en partielle en fin de traitement n'avaient pas de différence de SSP ou SG, dans un essai randomisé de phase III impliquant près de 200 patients [8]. De même, Cornell et al. en 2023 ont indiqué que la réponse interim en IRM n'était pas prédictive de la réponse de fin de traitement, avec 33 % des patients en réponse précoce (RC ou RP) progressant en fin de traitement [136].

Ces constatations ont suscité un intérêt pour l'identification d'outils alternatifs aptes à distinguer plus efficacement les patients répondant au traitement de ceux n'y répondant pas. À ce jour, il n'existe pas de données concrètes pour affirmer si les paramètres de diffusion ou de perfusion peuvent constituer des indicateurs fiables de réponse thérapeutique. Des marqueurs biologiques apparaissent comme une avenue prometteuse ; notamment, les travaux de Zhang et al. ont mis en évidence que les variations du niveau d'interleukine 10 surpassent la performance de l'IRM T1 post-contrast pour prédire les récurrences dans le LPSNC [137].

La TEP-FDG, quant à elle, se profile comme un outil potentiel d'évaluation, particulièrement pour l'estimation précoce de la réponse, s'alignant sur les données reconnues pour les lymphomes systémiques [14,16,17,138].

Critères en TEP

La TEP au 18F-FDG constitue actuellement un examen de routine dans l'évaluation thérapeutique des tumeurs solides et des hémopathies lymphoïdes. Pour le lymphome B diffus à grandes cellules (LBDCG) systémique, elle occupe une position clé pour définir la stratégie thérapeutique. Les deux outils principaux pour l'évaluation de la réponse thérapeutique sont l'échelle visuelle du score de Deauville, qui compare l'intensité de la fixation tumorale à celle du foie et du médiastin, et la variation du SUVmax entre les lésions les plus hypermétaboliques sur deux examens successifs [139–143] (Figure 4.1). Ces outils sont utilisés en clinique pour l'évaluation en fin de traitement et de manière précoce, après deux cycles de traitement, par une TEP dite interim. L'intérêt pronostique de la TEP interim est reconnue dans la littérature, et son utilité dans les essais d'intensification ou de réduction de la thérapie a également été prouvée [144–146]

score 1 : pas d'hyperfixation
 score 2 : fixation ≤ médiastin
 score 3 : fixation > médiastin mais ≤ foie
 score 4 : fixation > foie
 score 5 : fixation >> foie et/ou progression

$$\Delta\text{SUV} = 100 \times \frac{\text{SUV}_{T_0} - \text{SUV}_{T_2}}{\text{SUV}_{T_0}}$$

Score de Deauville

Calcul du $\Delta\text{SUV}_{\text{max}}$

FIGURE 4.1 Score de Deauville et méthode de calcul du $\Delta\text{SUV}_{\text{max}}$

Dans le contexte des LPSNC, les données issues de la littérature sont nettement moins abondantes. Divers travaux ont étudié le rôle de la TEP au 18F-FDG dans l'évaluation thérapeutique des LPSNC, tant en cours de traitement (interim) qu'en fin de traitement [12,14,17,138,147,148]. Cependant, la plupart de ces études sont rétrospectives et portent sur un nombre limité de patients. La cohorte la plus large à ce jour a inclus 66 patients évalués rétrospectivement, révélant que seul le résultat de la TEP en fin de traitement était corrélé à la survie, avec une survie sans progression (SSP) de 36,2 mois pour les TEP négatives, c'est-à-dire sans fixation résiduelle, contre 3,9 mois pour les TEP positives [14]. Nous présentons ci-dessous un résumé des études examinant la TEP comme outil d'évaluation thérapeutique (Tableau 4.2).

	Study	PMID	Nb	Study type	Timing	Treatment	Results
18F-FDG	Birsen et al, Oncotarget, 2018	29682187	25	retrospective	Interim	2 Rituximab, MTX, Temozolomide	2y-PFS : 72% for PET ₋ vs 33% for PET ₊
			66		Interim	4 MTX	No impact
	Jo et al, Ann Hematol, 2017	28725988	66	retrospective	Final	4 MTX + 2 Cytarabine	PFS : 36.2 m for PET ₋ vs 3.9 m for PET ₊
	De Bonilla Damia, Rev esp, 2017	28438519	7	retrospective	Final	BAM(5), BEAM(1), R-BAM(1)	PET/MRI discordance for 3 out of 7 patients
	Mercadal et al, Medicina Clinica, 2015	25433778	7	retrospective	Final	MTX	2y-OS : 100% for PET ₋ vs 38% for PET ₊
	Maza et al, Radiol. Oncol, 2013	23801905	8	prospective	Final	6 MTX	PET/MRI discordance for 2 out of 6 patients
	Palmedo et al, EJMNM, 2006	16220304	8	retrospective	Interim after first cycle (5) / Final (3)	6 MTX + Cytarabine	PET/MRI discordance for 3 out of 6 patients
11C-MET	Jang et al, CNM, 2012	22955084	4	Clinical cases	Final	9 MTX	PET/MRI concordance for all patient
	Ahn et al, CNM, 2008	29877882	26	prospective	Interim	4 MTX	TBR and MTV as prognostic biomarkers of PFS

Tableau 4.2 Données de la littérature sur l'intérêt de la TEP dans le suivi des patients atteints de LPSNC

Il est notable que les critères de positivité de la TEP présentent de légères variations selon les différentes études. La majorité se base sur une évaluation visuelle, avec comme critère principal une captation du traceur supérieure à celle du cortex cérébral normal [14,17,138,147]. Birsén et al. précise que le cortex normal doit être évalué du côté opposé à la lésion, tandis que Jo et al. prennent en compte l'ensemble du cortex sain, y compris la captation physiologique au niveau du cervelet [14,147]. L'étude de Palmedo et al. introduit une nuance entre une fixation suspecte (accumulation de FDG semblable à celle du cortex) et une fixation hautement suspecte (accumulation de FDG excédant celle du cortex) [12]. Les recherches de Maza et al., quant à elles, s'appuient sur un indice quantitatif, considérant un rapport de captation supérieur à 1.15 par rapport au côté controlatéral comme significativement élevé [148].

III. Article 3

Etude prospective multicentrique du réseau LOC : TEP interim dans le lymphome primitif du système nerveux central

Objectif

L'objectif de notre étude était d'évaluer l'intérêt de la TEP-IRM interim au FDG chez des patients de plus de 60 ans atteints de LPSNC *de novo*, comme indicateur de la réponse en fin de traitement d'induction et de la survie.

Méthodes

La cohorte, identique à celle présentée précédemment (cf. Article 2), a bénéficié de trois TEP-IRM au 18F-FDG : en pré-thérapeutique, en interim (après 2 cycles) et en fin de chimiothérapie d'induction. L'évaluation de fin de traitement s'est basée sur les critères IPCG et l'examen clinique. L'analyse de la TEP incluait une évaluation visuelle classifiant les images en positif ou négatif selon la présence ou non d'une fixation résiduelle supérieure à celle d'une zone cérébrale miroir controlatérale, et une analyse quantitative mesurant les variations des indices TEP déjà évalués pré-thérapeutiquement : $\Delta\text{SUV}_{\text{max}1-2}$, $\Delta\text{SUV}_{\text{mean}1-2}$, et $\Delta\text{SUV}_{\text{peak}1-2}$, ΔMTV_{1-2} . En IRM, les patients étaient classés en fonction de leur réponse en deux groupes : les bons répondeurs (RC et RCu) et les mauvais répondeurs (SD, PR ou PG). Les variations des paramètres IRM ont aussi été collectées : ΔADC_{1-2} , ΔCBV_{1-2} , ΔVol_{1-2} . L'évaluation de la SSP et de la SG a été réalisée de la date du deuxième cycle de traitement jusqu'à la progression ou le décès respectivement.

Résultats

Sur 48 TEP interim réalisées, 9 étaient positives et concordantes avec une RP sur l'IRM interim. Parmi ces patients, 8 sur 9 (89 %) ont montré une progression de la maladie en fin de traitement. Trente-huit patients présentaient une TEP interim négative, parmi lesquels 20 avaient également une RC sur l'IRM interim. Parmi eux, 90 % étaient toujours en réponse en fin de traitement. La valeur prédictive positive de la TEP interim pour prédire la réponse de fin de traitement d'induction était nettement supérieure à celle de l'IRM interim : 89% (IC à 95% [61–100]) contre 44% (IC à 95% [26–63]) ($p < 0.001$).

Discussion

Le traitement des LPSNC repose sur une chimiothérapie d'induction suivie éventuellement d'un traitement de maintenance. Actuellement, il manque de données pour déterminer quels patients bénéficieraient d'un traitement de maintenance. En IRM, la réponse est souvent incomplète après 2 cycles et il est délicat de stratifier ces patients entre bons et mauvais répondeurs. Notre étude a démontré que la TEP interim permettait de classer les patients entre bons et mauvais répondeurs, de manière plus fiable que l'IRM interim. Bien que quelques études dans la littérature aient examiné le rôle de la TEP dans l'évaluation thérapeutique, les résultats étaient discordants. L'étude de Birsan et al. avait montré de bonnes performances pour la VPN pour la TEP interim : 94.74%, IC à 95% [75.61%–99.05%], tandis que l'étude de Jo et al. n'avait pas mis en évidence d'impact pronostique de la TEP interim [14,147].

Notre étude confirme que la TEP interim au FDG est un outil d'évaluation thérapeutique fiable et robuste pour le LPSNC, avec un effet « on/off » similaire à celui identifié dans le lymphome de Hodgkin. L'analyse semi-quantitative des variations des paramètres TEP et IRM n'a pas ajouté de valeur aux performances. Notre étude ouvre la voie à une prise en charge plus personnalisée des patients, et a déjà modifié les pratiques cliniques locales à la Pitié-Salpêtrière. Les patients en RP sur l'IRM interim sont désormais systématiquement évalués par TEP FDG et surveillés méticuleusement si la TEP interim est positive. Le futur de ces recherches pourrait mener à des essais thérapeutiques guidés par la TEP.

Neuro-Oncology

XX(XX), 1–10, 2024 | <https://doi.org/10.1093/neuonc/noae029> | Advance Access date 16 February 2024

1

Interim FDG-PET improves treatment failure prediction in primary central nervous system lymphoma: An LOC network prospective multicentric study

Laura Rozenblum^{*}, Caroline Houillier, Amandine Baptiste, Carole Soussain, Véronique Edeline, Philippe Naggara, Marine Soret, Valérie Causse-Lemercier, Lise Willems, Sylvain Choquet, Renata Ursu, Damien Galanaud, Lisa Belin[†], Khê Hoang-Xuan[†], and Aurélie Kas[†]

All author affiliations are listed at the end of the article

Corresponding author: Laura Rozenblum, MD, Department of Nuclear Medicine, Groupe Hospitalier Pitié-Salpêtrière, Assistance Publique - Hôpitaux de Paris, 47-83 boulevard de l'hôpital, 75013 Paris, France (laura.rozenblum@aphp.fr).

[†]These authors contributed equally to this work.

Abstract

Background. The purpose of our study was to assess the predictive and prognostic role of 2-18F-fluoro-2-deoxy-D-glucose (FDG) positron emission tomography (PET)/MRI during high-dose methotrexate-based chemotherapy (HD-MBC) in de novo primary central nervous system lymphoma (PCNSL) patients aged 60 and above.

Methods. This prospective multicentric ancillary study included 65 immunocompetent patients who received induction HD-MBC as part of the BLOCAGE01 phase III trial. FDG-PET/MRI were acquired at baseline, post 2 cycles (PET/MRI2), and posttreatment (PET/MRI3). FDG-PET response was dichotomized with “positive” indicating persistent tumor uptake higher than the contralateral mirroring brain region. Performances of FDG-PET and International PCNSL Collaborative Group criteria in predicting induction response, progression-free survival (PFS), and overall survival (OS) were compared.

Results. Of the 48 PET2 scans performed, 9 were positive and aligned with a partial response (PR) on MRI2. Among these, 8 (89%) progressed by the end of the induction phase. In contrast, 35/39 (90%) of PET2-negative patients achieved complete response (CR). Among the 18 discordant responses at interim (PET_{CR}/MRI_{PR}), 83% ultimately achieved CR. Eighty-seven percent of the PET2-negative patients were disease free at 6 months versus 11% of the PET2-positive patients ($P < .001$). The MRI2 response did not significantly differentiate patients based on their PFS, regardless of whether they were in CR or PR. Both PET2 and MRI2 independently predicted OS in multivariate analysis, with PET2 showing a stronger association.

Conclusions. Our study highlights the potential of interim FDG-PET for early management of PCNSL patients. Response-driven treatment based on PET2 may guide future clinical trials.

Trial. LOCALYZE, NCT03582254, ancillary of phase III clinical trial BLOCAGE01, NCT02313389 (Registered July 10, 2018—retrospectively registered) <https://clinicaltrials.gov/ct2/show/NCT03582254?term=LOCALYZE&draw=2&rank=1>.

Key Points

- 2-18F-Fluoro-2-deoxy-D-glucose (FDG)-positron emission tomography (PET) provides early, reliable differentiation of primary central nervous system lymphoma patient responses.
- Interim FDG-PET robustly stratifies 1-year progression-free survival, outperforming MRI in cases of partial responses.
- FDG-PET response's predictive power suggests a pivotal role in future clinical trials.

Importance of the Study

Management of primary central nervous system lymphomas (PCNSL) in the elderly is a critical challenge in neuro-oncology. Current tools for early risk stratification are insufficiently accurate, leading to suboptimal treatment strategies. This study highlights the potential of interim 2-18F-fluoro-2-deoxy-D-glucose (FDG)-positron emission tomography (PET) to differentiate between PCNSL patients with good and poor treatment responses. Interim FDG-PET may outperform interim morphological MRI in case of partial response, providing

a more effective prediction of progression-free survival and response at the end of induction therapy.

In our study, the predictive ability of interim FDG-PET imaging was also confirmed by multivariate analysis. These results reinforce the role of FDG-PET in PCNSL management and may serve as a foundation for future FDG-PET-guided clinical trials. The integration of FDG-PET imaging into routine practice holds promise for refining risk-adapted treatment strategies and improving outcomes in PCNSL.

Primary central nervous system lymphoma (PCNSL) is a rare form of non-Hodgkin lymphoma that is confined to the brain, eyes, and cerebrospinal fluid (CSF), without any systemic spread.¹ The treatment for PCNSL is based on high-dose methotrexate-based chemotherapy (HD-MBC) for elderly patients and HD-MBC associated or not with stem cell auto-transplantation as consolidation for younger and fit patients.²⁻⁷ The prognosis of PCNSL is highly variable, particularly in the elderly, and clinical markers such as age or performance status are insufficient to predict risks of chemoresistance or relapse after chemotherapy.⁸ The International Primary CNS Lymphoma Collaborative Group (IPCG) currently recommends the use of criteria based mainly on brain MRI for response assessment.⁹

Tabouret et al.¹⁰ have described a correlation between objective response on T1 contrast-enhanced MRI at the end of treatment and overall survival (OS). However, several studies have reported early detection of recurrence in patients who were in complete response (CR) on end-of-treatment MRI, thereby questioning its efficacy in assessing residual disease.^{11,12} Van der Meulen et al.¹³, in a randomized controlled trial involving 199 PCNSL patients, found no significant differences in progression-free survival (PFS) or OS between patients with CR or partial response (PR) on MRI at the end of first-line therapy. Hence, although MRI is a valuable tool in PCNSL assessment, there is a need for improved assessment methods, in particular for interim evaluation.

Positron emission tomography (PET) has been validated as a reliable tool for the early evaluation of therapeutic responses in 18F-fluoro-deoxyglucose (FDG)-avid systemic lymphomas.^{14,15} Several methods have been developed for assessing therapeutic responses in FDG-avid systemic lymphomas, such as the Deauville score (a visual score based on comparing glucose uptake to the liver and mediastinum) and the Δ SUVmax approach (percentage change in the SUVmax of the target lesion between baseline and interim FDG-PET staging).¹⁶ Unfavorable results in systemic lymphoma are defined as uptake visually superior to the liver for the Deauville score and as a $\leq 66\%$ relative SUVmax reduction for the Δ SUVmax approach.¹⁷

As we have previously demonstrated, PCNSL tumor exhibits high FDG avidity at baseline with a ratio to unaffected cortex ranging from 3 to 5.^{18,19} Few studies have explored FDG-PET for therapeutic assessment in newly

diagnosed PCNSL.²⁰⁻²⁷ These studies were mostly retrospective, with different time points and discordant conclusions. Birsén et al.,²⁴ in a cohort of 25 patients, found that a negative interim FDG-PET performed after 2 therapy cycles was associated with longer PFS. Conversely, Jo et al.²² found no effect with interim FDG-PET (performed after 4 therapy cycles) but identified the prognostic value of the FDG-PET at the end of first-line therapy in a group of 66 patients. Considering potential toxicities and an unfavorable risk-benefit ratio with consolidation treatment, robust FDG-PET criteria for risk-tailored adapted treatment hold promise.

The LOCALYZE study (NCT03582254) is an ancillary investigation of the phase III randomized controlled trial BLOCAGE01 (NCT02313389). BLOCAGE01 aims to assess maintenance chemotherapy benefits versus observation in elderly newly diagnosed PCNSL patients, extending CR duration with acceptable toxicity after HD-MBC induction chemotherapy. LOCALYZE, conducted concurrently with BLOCAGE01 during induction chemotherapy, involves baseline, interim, and end-of-therapy brain FDG-PET/MRI. Our objectives were 2-fold: (i) to ascertain the predictive value of the interim FDG-PET/MRI (PET/MRI2) on the response to first-line treatment and (ii) to evaluate the prognostic implications of PET/MRI2 on PFS and OS.

Materials and Methods

Study Design

Immunocompetent patients over 60 years old with newly diagnosed PCNSL included in phase III clinical trial BLOCAGE01 and eligible for this ancillary prospectively multicentric study (LOCALYZE NCT02313389) were included between January 2016 and April 2021. Patients were recruited by 5 centers from the French "Lymphomes Oculo-Cerebraux" (LOC) network: Neuro-Oncology department in Saint-Louis Hospital, AP-HP, Paris; Haematology department in Institut Curie, Saint Cloud, Neuro-Oncology and Haematology departments in Pitié-Salpêtrière Hospital; AP-HP, Paris and Haematology department in Cochin Hospital; AP-HP, Paris. Main inclusion criteria consisted of (i) newly diagnosed PCNSL, (ii) confirmed either

by brain histology or cytology in the CSF or vitrectomy, (iii) age of 60 years or older, (iv) Karnofsky Performance Status (KPS) greater than 40, and (v) no evidence of systemic non-Hodgkin lymphoma as demonstrated by a contrast-enhanced CT and/or whole-body FDG-PET/CT scan. Main exclusion criteria included (i) any other active malignancy (with the exception of basal cell carcinoma of the skin and cervical carcinoma in situ); (ii) pre-existing immunodeficiency (from HIV or organ transplant treatment); (iii) low-grade histology; (iv) isolated primary vitreoretinal lymphoma; (v) CNS relapse of a systemic lymphoma, and (vi) prior treatment for PCNSL. Patients included in LOCALYZE underwent a baseline FDG-PET/MRI within 21 days prior to the start of an HD-MBC, an interim FDG-PET/MRI after 2 cycles, and an end-of-therapy FDG-PET/MRI after completion of the chemotherapy. The chemotherapy regimen consisted of four 28-day cycles of rituximab (R) (375 mg/m², D1), methotrexate (M) (3.5 g/m², D1 and D15), procarbazine (P) (100 mg/m², D1–D7), and vincristine (V) (1.4 mg/m², D1 and D15) followed by a cycle of cytarabine (A) (3 g/m², D1 and D2; R-MPVA).

Therapeutic efficacy was assessed at the end of the treatment regimen by board-certified experts in neuro-oncology and neuroradiology using the IPCG framework. This diagnostic scheme incorporates parameters such as gadolinium-enhanced MRI findings, corticosteroid use, CSF cytology, and ocular assessment. Outcomes are categorized into CR, unconfirmed complete response (CRu), PR, stable disease, and progressive disease (PD). CR necessitates the unequivocal absence of disease markers across all tests, while CRu pertains to cases meeting CR criteria yet continuing corticosteroid treatment or presenting with minor anomalies. PR entails a 50% reduction in the primary cerebral lesion, and PD is characterized by a 25% enlargement in lesion size or the appearance of novel disease sites.

Within the scope of the BLOCAGE01 trial, those attaining CR or CRu statuses were subsequently randomized into 2 distinct arms. One arm underwent a 7-month maintenance treatment involving rituximab, methotrexate, and temozolomide, while the other was subjected to observational management. Clinical prognostic indices such as gender, age, and KPS were recorded.

The PFS was defined as the time from the end of the second cycle of chemotherapy to progression, relapse, or death from any cause. The OS was defined as the time from the end of the second cycle of chemotherapy to death from any cause. The cutoff date for PFS and OS events was October 26, 2022.

All participants provided informed consent for their participation in the study, which was approved by the Protection of Person's Committee (2014-002597-37) and the French National Agency for the Safety of Medicines and Health Products (ANSM). The study was conducted in accordance with the principles of the Declaration of Helsinki.

FDG-PET/MRI Protocol

The imaging studies for this study were conducted at the nuclear medicine department of the Pitié-Salpêtrière Hospital using a hybrid 3T PET/MRI system (Signa PET/MR General Electric Healthcare). Static FDG-PET images

were acquired 90 minutes postinjection of 2 MBq/kg (125–250 MBq) of FDG over 20 minutes. Patients fasted for at least 4 hours and blood glucose was measured before the exam could be performed. MRI included 3D susceptibility-weighted imaging (SWAN), 3D fluid-attenuated inversion recovery (FLAIR), T2*-weighted perfusion, diffusion, monovoxel spectroscopy, 3D arterial spin labeling (ASL) sequences, and 3DT1-weighted spin-echo sequence before and after injection of 0.2 mL/kg of gadoterate dimeglumine (Dotarem; Supplementary Figure 1).

FDG-PET/MRI Analysis

MRI assessments were conducted independently from FDG-PET findings and interpreted by a board-certified neuroradiologist (DG) in accordance with IPCG criteria. Using the General Electric Healthcare Advantage Workstation (ADW) software, we generated maps for apparent diffusion coefficient (ADC), cerebral blood volume (CBV), and cerebral blood flow (CBF).^{28,29} The preprocessing of perfusion data included automatic arterial input function generation, deconvolution, and leakage correction. We also performed quantitative measurements on multiparametric MRI images through a region-of-interest (ROI) analysis. For both diffusion and perfusion metrics, a 50 mm² ROI was manually outlined on the most conspicuous tumor area, intentionally excluding any cystic, necrotic, or hemorrhagic components, as previously reported in the literature.^{28–30} A contralateral ROI within normal-appearing white matter served as a standardization reference for calculating relative ADC (rADC), CBV (rCBV), and CBF (rCBF). Percentage changes between baseline and interim examinations were computed for ADC, and CBF based on the ASL sequence, denoted as ΔADC_{1-2} and ΔCBF_{1-2} , respectively.

FDG-PET scans underwent both visual and semiquantitative analyses using the ADW software. For the former, independent evaluations were carried out by 2 board-certified nuclear medicine specialists (L.R. and A.K.) using a dichotomous rating scale. An FDG-PET scan was classified as positive if it exhibited pathologic tracer uptake in the brain that exceeded the intensity observed in the corresponding mirroring unaffected contralateral cortex, whether white or gray matter based on the specific location. Conversely, an FDG-PET scan was categorized as negative if it showed no evidence of pathologic uptake.

In the semiquantitative analysis, the target lesion was identified as the one displaying the highest maximum standard uptake value (SUVmax). We recorded the percentage changes (Δ) in SUVmax, SUVmean, and SUVpeak for this lesion between the baseline and subsequent interim examinations, denoted as $\Delta\text{SUVmax}_{1-2}$, $\Delta\text{SUVmean}_{1-2}$, and $\Delta\text{SUVpeak}_{1-2}$.

Concurrently, assessments were made in the contralateral brain hemisphere that was uninvolved, using a mirror ROI to calculate the tumor-to-background ratio. Specifically, the SUVmax in the tumor was compared to both the SUVmax and the SUVmean in the contralateral cortex, generating TBRmax and TBRmean, respectively.

Measurements in an unaffected region of the cerebellar hemisphere were also taken to compute the tumor-to-cerebellum ratio: specifically, the SUVmax of the lesion

was compared divided by to the SUVmean of the cerebellum. The metabolic volume for the target lesion was determined using a threshold of 41% of the SUVmax, both at baseline and in subsequent FDG-PET scans, in accordance with the EANM procedure guidelines for tumor imaging for 18F-FDG-avid and homogeneous tumors, independently of the MRI evaluation.³¹ Percentage volume change was documented as (ΔMTV_{1-2}).

We used the same target lesion on the enhanced MRI sequence to assess changes in the volume of the enhancing target lesion, denoted as ΔVol_{1-2} .

In follow-up assessments, patients were grouped into 2 categories: "Good responders" and "Poor responders." For MRI, the Good responders group included CR and CRu while the Poor responders group included the patients with SD, PR, or progression. For PET, Good responders group had negative scans, indicating complete normalization of any prior pathological uptake. Poor responders had positive FDG-PET scans, indicating either PR, stable disease, or progression. A PR was marked by reduced lesion metabolism, yet with residual activity exceeding that of the contralateral mirror cortex. Stable disease showed no change in the target lesion, while progression was defined by new foci, extension, or increased lesion metabolism. It should be emphasized that treatment decisions were exclusively guided by the IPCG criteria upon completion of the induction therapy, and were not affected by interim or final FDG-PET assessment findings.

Statistical Analysis

Continuous variables are expressed as medians [Q1–Q3] and categorical variables as percentages. PFS and OS functions were estimated using the Kaplan–Meier method, and median PFS with 95% CI was provided. Associations between clinical factors and FDG-PET parameters with PFS or OS were assessed through univariate and multivariate Cox proportional hazard models. Hazard ratios (HR) are presented with 95% CI. Log-linearity was explored for continuous variables. Schoenfeld residual plots were used to check the assumption of proportional hazards. A final multivariate model was established using a backward selection based on the likelihood ratio test, where clinical pertinent covariates (sex, age, and KPS) were imposed in the model. For age and KPS, cutoffs were defined based on clinical knowledge. For all other continuous predictors, log-linearity was explored, and in the case of nonrespect of this assumption, variables were dichotomized at the median value (observed on the dataset). Sensitivity, specificity, PPV and NPV of interim FDG-PET, and MRI were calculated considering the response group at the end of treatment as the gold standard of response (where CR or CRu were defined as negative). Corresponding 95% CIs were estimated using bootstrap on 5000 iterations. A Mc Nemar paired test was used to compare specificities of interim PET and MRI on patients who were evaluated for both exams. All tests were 2-tailed and used a significance level of 0.05. In order to take into account the large number of comparisons explored across the analysis, Benjamin–Hochberg method was used to correct *P*-values testing the prognostic effect of parameter variation between examinations 1 and 2.

Analyses were computed using R software version 4.1.1 (<https://cran.r-project.org/>).

Results

Patients Characteristics

A total of 65 patients were included in the LOCALYZE study, of which 59 underwent baseline FDG-PET/MRI (PET/MRI1). Clinical assessments based on IPCG criteria were unavailable for 5 patients: 1 declined further evaluation and 4 died due to causes unrelated to tumor progression within a month following diagnosis. Therefore, the analysis was conducted on a total of 54 patients. The median age was 71.5 years (range 60–85) and the median KPS was 70% (range 50%–100%). All observed lesions were identified as diffuse large B-cell lymphoma (Table 1). FDG-PET was positive for 52 out of 54 patients at baseline (FDG-PET1) with a median SUVmax of 27.7 [22.8–36] and a median TBRmax of 5.3 [3.8–7.9]. Two patients exhibited a negative FDG-PET1, which was attributed to prior surgical biopsies that left subcentimetric partial residues smaller than the resolution capacity of the PET camera. T1-weighted sequence after contrast showed high contrast enhancement for all lesions. Baseline characteristics of FDG-PET/MRI and demographic data comparing response groups have been described and can be found in our previous work.¹⁸

Out of 54 patients, 48 (89%) had an FDG-PET/MRI2 after the second chemotherapy cycle (median 28 days, IQR [23–29]). Two patients withdrew and 4 discontinued treatment before the second R-MPVA cycle (3 due to toxicities and 1 to PD). Of the total cohort, 43 patients (80%) had the final FDG-PET/MRI (PET/MRI3) after the last chemotherapy cycle (median 28 days, IQR [23–34]). Five patients were excluded due to disease progression before completing the final R-MPVA cycle. See Figure 1 and Supplementary Figure 2 for the study flow chart and visual FDG-PET/MRI assessment results.

Following completion of induction chemotherapy, 33/54 (61%) of the patients were randomized in the phase III BLOCAGE01 trial. Among them, 18 patients (55%) were assigned to the observation arm, while 15 patients (45%) were assigned to the maintenance arm.

Among the 48 patients with an FDG-PET/MRI interim evaluation, 34 were in CR/CRu, 1 in PR, and 12 had PD at the end of induction therapy (Time 3) according to IPCG criteria. None of the patients had stable disease. Of the 48 patients, 1 was excluded from the interim MRI evaluation due to the lack of contrast-enhanced imaging, resulting in 47 patients being analyzed at this stage.

Results of the Interim FDG-PET/MRI Based on Visual Assessment

Concordant Poor responder patients.—Nine of 47 patients (19%) were in PR on FDG-PET2 with concordant PR on MRI2. Eight of these patients (89%) experienced disease progression between T2 and T3 and 1 achieved a CR.

Table 1. Patients Characteristics

	All Patients (n = 54)
Sex, n (%)	
Male	21 (38.9%)
Female	33 (61.1%)
Age (years)	
Median [Q1-Q3]	71.5 [66-77]
Karnofsky Performance Status	
Median (range)	70 (60-80)
Diagnostic	
Lumbar puncture	5 (9.3%)
Biopsy	47 (87.0%)
Surgery	2 (3.7%)
Randomization, n (%)	
No	21 (39%)
Yes—observation arm	18 (33.3%)
Yes—maintenance arm	15 (27.7%)
Treatment completion, n (%)	
4-5 RMPV and cytarabine	40 (74.1%)
<4 RMPV	14 (25.9%)

Concordant Good Responders patients.—Thirty-eight of 47 patients (81%) were FDG-PET2 negative, including 20 patients with a concordant CR on MRI2. Among them, 18 (90%) maintained a CR on the final FDG-PET/MRI assessment (PET/MRI3), while 2 (10%) experienced treatment failure between T2 and T3.

Discordant Responders patients on PET and MRI.—Among the 38 FDG-PET2 negative patients, 18 showed a discordant PR on MRI2. Of these, the majority (15/18, 83%) achieved a CR at the end of induction MRI (MRI3). Two patients (11%) had a PD at T3, and 1 patient (6%) remained in a discordant PET3_{CR}/MRI3_{PR} state but achieved a CR after a 47-month follow-up.

It is worth noting that the patient with an available FDG-PET but an unanalyzable MRI achieved CR on his FDG-PET2 scan and maintained this CR status through a 29-month follow-up.

The specificity of a positive interim exam for predicting end-of-therapy response was higher for FDG-PET2 compared to MRI2: 97% (95% CI [90-100]) versus 55% (95% CI [37-72]) ($P < .001$) (Figure 2). The negative predictive values were close for both exams (85% (95% CI [72-95]) for MRI2 and 90% (95% CI [75-100]) for FDG-PET2), but the positive predictive value was higher for FDG-PET2 over MRI2: 89% (95% CI [61-100]) versus 44% (95% CI [26-63]) ($P < .001$) (Supplementary Table 1). There was only one false-positive patient reported on FDG-PET2 (Poor responder), but the FDG uptake for this patient was minimal, with a TBRmax of

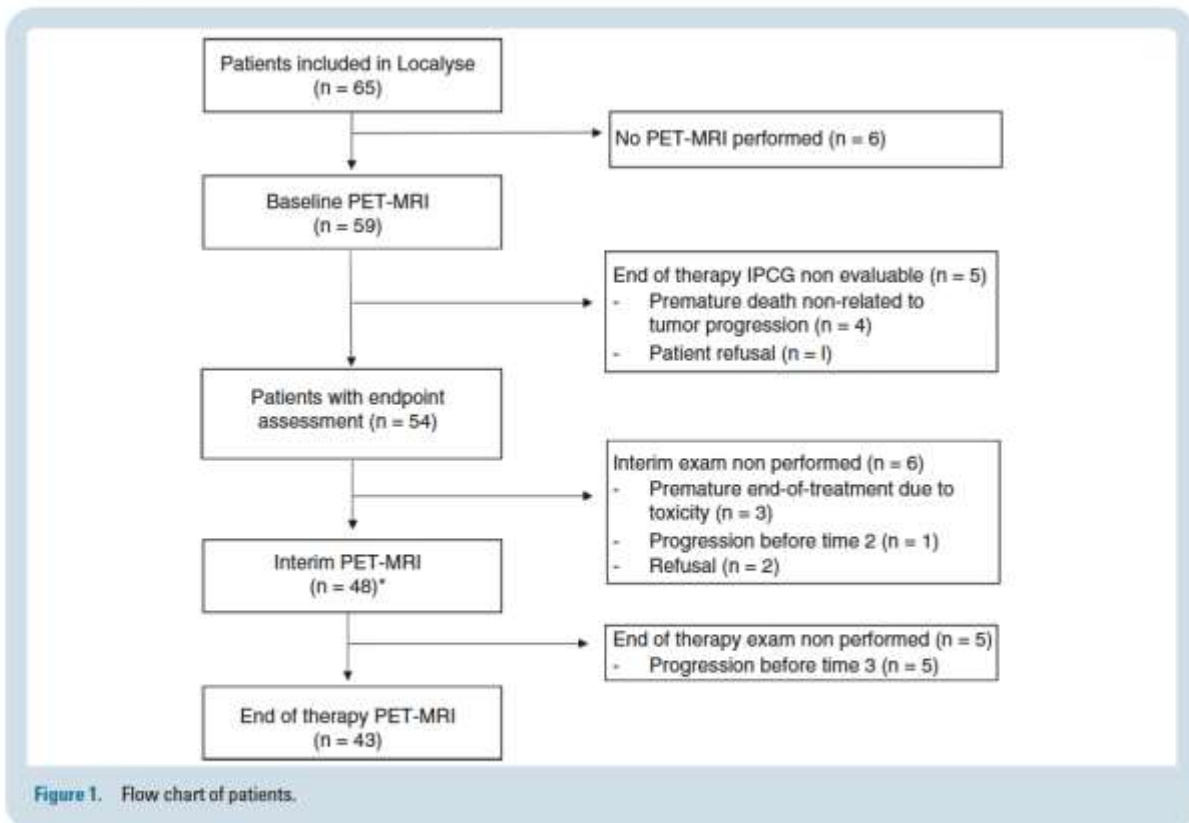


Figure 1. Flow chart of patients.

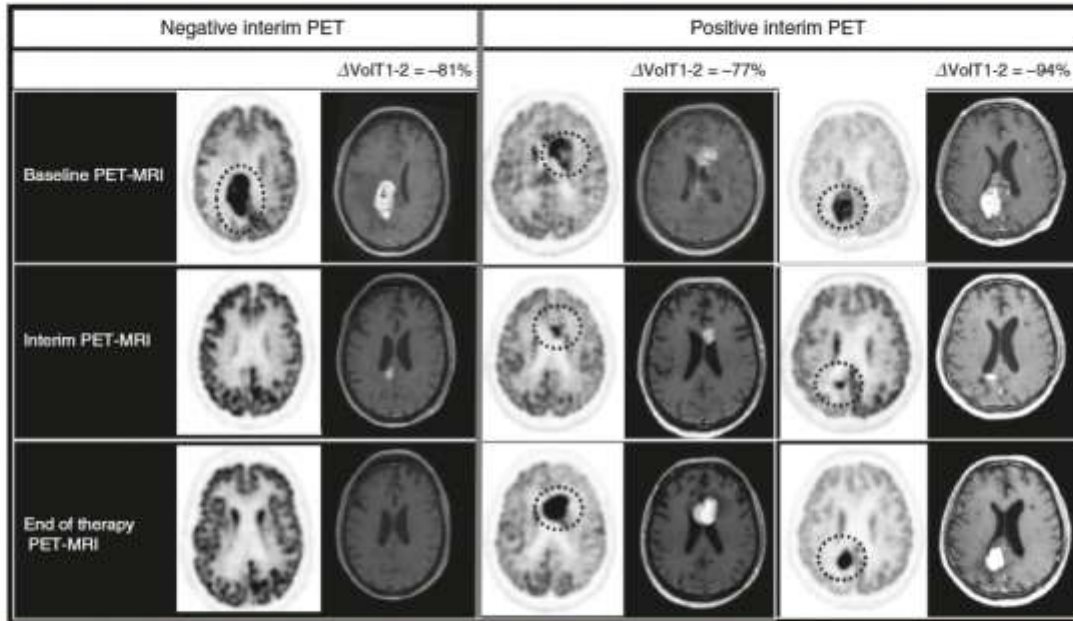


Figure 2 Representative cases. Negative interim PET patient: 79-year-old woman with PCNSL. Contrast-enhanced T1-weighted MRI showed a PR on interim assessment ($\Delta\text{Vol}_{1-2} = -81\%$) but interim FDG-PET showed a complete metabolic response. The end-of-therapy assessment was normal for both FDG-PET and MRI, and the patient was still alive at the end of the follow-up period. Positive interim PET patient: first column; 70-year-old woman with PCNSL. Contrast-enhanced T1-weighted MRI showed a PR on interim assessment ($\Delta\text{Vol}_{1-2} = -77\%$) concordant with a PR on FDG-PET ($\Delta\text{SUV}_{\text{max}1-2} = -12\%$). End-of-therapy evaluation showed a concordant progression and the patient died of lymphoma 3 months later. The second column displays the image of a 64-year-old man with a significant reduction in tumor volume on interim MRI ($\Delta\text{Vol}_{1-2} = -91\%$), indicating a good PR. However, his interim PET scan shows positive findings, and the patient experienced disease progression 3 months after the interim assessment.

1.06. This value is significantly below the median of positive FDG-PET2, which stands at 3.0 [2.4–6.7] (Supplementary Table 2).

Prognostic Value of Interim FDG-PET/MRI Based on Visual Assessment

The median follow-up from T2 was 34 months (95% CI, 26.7–44.4), and the median PFS was 31.6 months (95% CI, 15–NA). Twelve progressions were observed between T2 and T3 and 10 additional progressions were noted in the concordant PET_{3CR}/MRI_{3CR} group after the completion of induction therapy, with a median time to progression of 11 months. Throughout the follow-up period, a total of 20 patients died, with lymphoma progression accounting for 14 cases, toxicity for 1 case, and other comorbidities for the remaining cases.

Eighty-seven percent (CI: 77%–98%) of the PET2-negative patients were disease free at 6 months versus 11% (CI: 2%–70%) of the PET2-positive patients (Logrank $P < .001$). The PFS of patients with CR on MRI2 did not significantly differ from patients with PR on MRI2 (Figure 3). Notably, the PFS of patients with discordant responses at T2 (PET2CR/MRI2PR) was similar to that of patients with concordant CR (Supplementary Figure 3). The predictive value of FDG-PET2 remained significant independently of patient

demographics, including sex, age, and KPS. Multivariate analysis further confirmed that CR on FDG-PET2 served as an independent predictor of improved PFS, with an HR of 10.36 (95% CI, 3.82–28.07, $P < .001$; Table 2). Of note, steroid therapy was not included in the multivariate model, as we previously demonstrated that this therapy had no statistically significant impact on 18F-FDG avidity at baseline.¹⁸

Regarding OS prediction, both FDG-PET2 and MRI2 responses were significantly associated with patient outcomes, in univariate and multivariate analyses. However, associations with OS were higher with FDG-PET compared to MRI: HR was 87.76 [10.37–742.56], $P < .001$ for FDG-PET2 and 11.40 [1.46–88.81], $P = .003$ for MRI2 (Supplementary Table 3; Supplementary Figure 4).

Quantitative Analysis of FDG-PET/MRI

The change in SUVmax value between the first and second assessments ($\Delta\text{SUV}_{\text{max}1-2}$) was significantly higher in the Good responders group compared to the Poor responders group: -76.1% (IQR -61.6% to -82%) versus -36.4% (IQR -7.2% to -74%), $P = .021$. The change in contrast-enhanced volume (ΔVol_{1-2}) was also significantly higher in the Good responders group compared to the Poor responders group ($P = .0001$), although both groups demonstrated a high level of decrease in

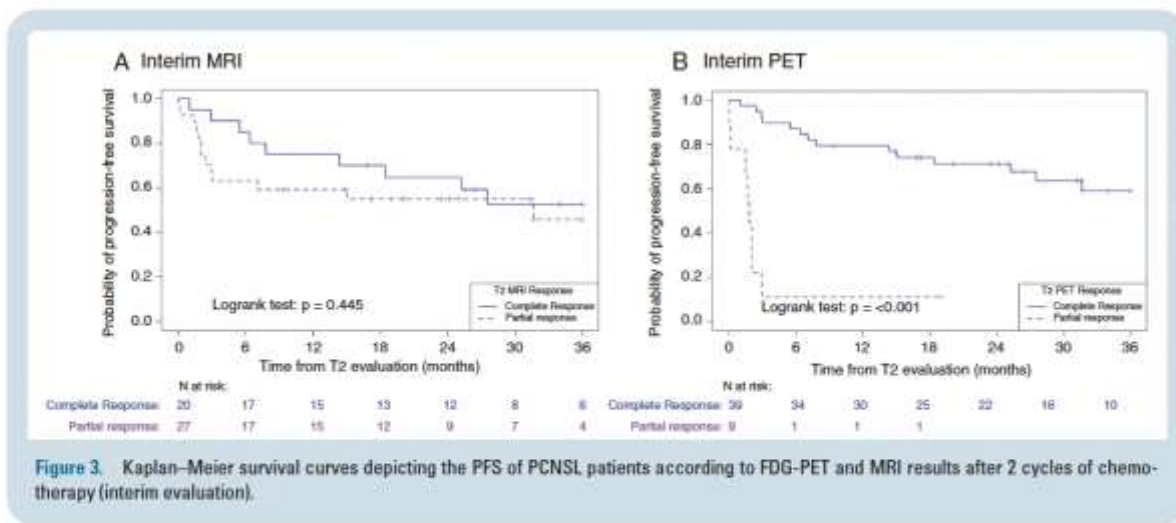


Table 2. Multivariate Analysis of Clinical and FDG-PET/MRI Scans Based on Visual Analysis on PFS

PFS	Multivariate Analysis (PET, n = 48)		Multivariate Analysis (MRI, n = 47)	
	HR (95% CI)	P	HR (95% CI)	P
Male (ref F)	1.67 [0.60, 4.64]	.33	1.69 [0.65, 4.36]	.29
Age ≥ 70 (ref < 70 years)	1.38 [0.50, 3.80]	.53	1.37 [0.52, 3.56]	.52
KPS ≥ 70 (ref < 70)	0.76 [0.32, 1.79]	.54	0.88 [0.37, 2.08]	.77
INTERIM PET response PR (ref CR)	10.36 [3.82, 28.07]	<.001		
INTERIM MRI response PR (ref CR)			1.33 [0.54, 3.23]	.53

Bold P values <.05.

enhanced lesion: -96.18% (IQR -82.62% to -100%) in the Poor responders group versus -100% in the Good responders group. No significant difference was observed in the decrease of contrast-enhanced volume between the concordant FDG-PET_{2PR}/MRI_{2PR} and the discordant FDG-PET_{2CR}/MRI_{2PR} groups: -77% (IQR -60% to -96%) versus -84% (IQR -79% to -95%), respectively (P = .54). The analyses of visual and semiquantitative parameters from PET-positive examinations across the 3 time points are presented in [Supplementary Table 2](#).

The MTV variation of the target lesion between T1 and T2 (Δ MTV₁₋₂) was the only FDG-PET/MRI parameter significantly correlated with treatment failure and OS ([Supplementary Table 4](#)). No significant changes were observed in diffusion or perfusion MRI parameters that could serve as prognostic indicators of outcome ([Table 3](#)).

Discussion

Despite advances in managing older patients, the prognosis for elderly PCNSL patients remains poor. The standard of care includes HD-MBC with or without consolidation treatment like maintenance chemotherapy or

autologous stem cell transplantation for fit patients.^{4,5} However, this approach is debated due to its high risk of toxicities. Established prognostic factors like age or functional status are insufficient to guide physicians in predetermining which patients would benefit from consolidation treatment.

The clinical efficacy of MRI, as per IPCG criteria, has shown limitations, particularly in its predictive capacity. It often fails to identify patients at risk of relapse despite showing a CR or PR on T1 contrast-enhanced sequences at interim assessment. It also struggles to predict disease-free status postinduction therapy.¹³ Indeed, the use of T1 contrast-enhanced weighted MRI as a response marker is challenging due to variations in sequence parameters, timing, and contrast dosage.³²⁻³⁴ Multimodal parametric MRI, incorporating diffusion and perfusion sequences, has been investigated as a prognostic tool in PCNSL at baseline, but not yet for therapeutic assessment.³⁵ In our study, we found no significant correlation between variations in diffusion or perfusion markers and patient outcomes.

This indicates the need for more robust tools in the management of PCNSL patients. Our study emphasizes the potential role of early FDG-PET, conducted after 2 chemotherapy cycles, in predicting patient responses. In newly diagnosed elderly patients with PCNSL, we demonstrated

Table 3. Prognostic Interim FDG-PET/MRI Biomarkers of PFS on Univariate Analysis

	Univariate (n = 54)	
	HR (95% CI)	P
PFS		
Male (ref F)	1.41 [0.62, 3.21]	.41
Age ≥ 70 (ref < 70 years)	0.95 [0.42, 2.14]	.90
KPS ≥ 70 (ref < 70)	0.87 [0.38, 2]	.75
ΔSUVmax ₁₋₂ (increase unit = 10)	0.89 [0.81, 0.98]	.26*
ΔMTV ₁₋₂ (ref < 100)	0.13 [0.05, 0.37]	<.001*
ΔVol ₁₋₂ enhanced lesion (ref < 98.3%)	0.86 [0.36, 2.07]	1*
ΔrADCmean ₁₋₂ (increase unit = 10)	1 [0.91, 1.09]	1*
ΔADCmin ₁₋₂ (increase unit = 10)	0.95 [0.87, 1.03]	1*
ΔrCBFmax ₁₋₂ (increase unit = 10)	0.99 [0.87, 1.11]	1*
ΔrCBVmax ₁₋₂ (increase unit = 10)	0.98 [0.94, 1.01]	1*

KPS, Karnofsky performans status; Vol., volume.

*P-values reported were corrected using Benjamini–Hochberg to take into account multiple comparisons computed in the analysis.

that interim FDG-PET scans were more effective than interim MRI scans in cases of PR or CR unconfirmed, and clinical parameters at distinguishing between Good responders and Poor responders. Moreover, interim FDG-PET's prognostic impact on PFS and OS was independent of baseline clinical factors like age, gender, and KPS.

Few studies have examined the role of FDG-PET in PCNSL assessment, often limited by small sample sizes and retrospective designs.²²⁻²⁷ Our findings emphasize that interim FDG-PET offers a more timely and reliable assessment compared to MRI in the evaluation of PCNSL patients. Many patients show partial PR on interim MRI even when their early FDG-PET is negative, with the majority ultimately achieving CR by the end of treatment.²⁹ For example, Birsén et al., in a 25-patient cohort, found interim PET scans, following 2 rituximab, methotrexate, and temozolomide therapy cycles, strongly correlated with end-of-therapy MRI responses (PPV of 66.67%, 95% CI, 33.34%–88.89%, and an NPV of 94.74%, 95% CI, 75.61%–99.05%). Jo et al.,²² in a 66-patient cohort, determined that end-of-therapy FDG-PET could stratify patients by PFS, but interim results had no significant impact. The interim FDG-PET in their study was conducted later than in ours and in Birsén's work (after the fourth methotrexate cycle instead of the second). This timing suggests that an early interim FDG-PET might facilitate the detection of distinct "on/off" responses, as observed in studies of Hodgkin's lymphoma.³⁸ On the other hand, it could be hypothesized that changes observed through MRI may lag behind the actual tumor response. In our patient cohort, we observed that the persistence of residual disease on interim FDG-PET, as indicated by higher uptake than in the contralateral healthy brain, was correlated with a suboptimal response at the end of therapy and with the overall prognosis. This finding implies that our dichotomous scale might be employed as conveniently as the Deauville score for systemic lymphoma in future clinical trials involving PCNSL.

It is significant to note that in our study, 8/9 patients (89%) exhibited a pattern of local progression. In these cases, the tumor, which had previously been classified as PR or CR at the interim assessment, reemerged at the same location during the end-of-treatment evaluation. While previous studies have mainly described patterns of relapse occurring at distinct sites, they did not specifically look at patients who showed early progression but rather focused on the recurrence following initial line therapy.^{10,37,38}

Our study's limitations include a relatively small patient sample size, due to the disease's rarity. However, our cohort, with predefined evaluation time points and uniform induction chemotherapy, is among the largest PET studies reported. The study's strengths encompass its prospective design and the use of an innovative hybrid PET/MRI system. This advanced technology facilitates concurrent analysis of underlying pathophysiological mechanisms by allowing for simultaneous study of metabolic uptake, as well as diffusion and perfusion patterns.³⁹ A further limitation is the variation in postinduction PET-MRI management strategies, with some patients receiving maintenance therapy and others only monitored. Subgroup analyses are pending until the BLOCAGE01 trial concludes (primary analysis planned for 2024). Despite these limitations, our study's key finding is that interim FDG-PET imaging could serve as an easy, early, and independent biomarker of PCNSL progression. These findings, however, require further validation in a larger patient cohort.

Other imaging modalities, such as 11C-methionine and 18F-fludarabine, are being investigated for their usefulness in managing PCNSL. However, data on their utility remain limited, and further studies are required to ascertain their role in this disease.⁴⁰⁻⁴² Circulating tumor DNA has also been found to be independent predictors of aggressive lymphomas and may play a role in therapeutic assessment in the future.⁴³

To the best of our knowledge, our study is the first to provide prospective evidence supporting the role of early FDG-PET evaluation in predicting chemotherapy response

and outcomes in elderly PCNSL patients. This approach, which can be readily incorporated into clinical practice, holds promise as a key tool in the management of PCNSL. Our findings suggest that close MRI follow-up should be promptly considered for patients with positive interim FDG-PET results. While our study underscores the potential of PET in personalized medicine strategies for PCNSL, further research is needed to fully substantiate its prediction role.

Supplementary material

Supplementary material is available online at *Neuro-Oncology* (<https://academic.oup.com/neuro-oncology>).

Keywords

early assessment | FDG | PET/MRI | primary central nervous system lymphoma | prognosis

Funding

This work was funded by Assistance Publique—Hôpitaux de Paris: PHRC LOCALYZE.

Acknowledgments

We would like to thank Dr. Inna Gertsenshteyn for proofreading our manuscript before submission.

Conflict of interest statement

None declared.

Authorship statement

A.K. and D.G., as well as C.H., L.R., V.C.-L., and M.S., contributed to the study's conception and design. L.R. and D.G. performed the MRI image data collection, while P.N., L.R., and A.K. performed the PET image data collection. C.S., R.U., L.W., and C.H., as well as S.C. and K.H.-X., recruited the patients. A.B. and L.B., in collaboration with L.R., performed the statistical analysis. All authors reviewed and approved the final version of the manuscript.

Data availability

Original data will be made available upon reasonable request by e-mail.

Affiliations

Department of Nuclear Medicine, Groupe Hospitalier Pitié-Salpêtrière, Assistance Publique—Hôpitaux de Paris (AP-HP), Sorbonne Université, Paris, France (L.R., P.N., M.S., V.C.-L., A.K.); INSERM, CNRS, Laboratoire d'Imagerie Biomédicale, Sorbonne Université, Paris, France (L.R., M.S., D.G., A.K.); Department of Neurology 2 Mazarin, AP-HP, Hôpitaux Universitaires La Pitié Salpêtrière—Charles Foix, Inserm, CNRS, Institut du Cerveau, Sorbonne Université, Paris, France (C.H., K.H.-X.); Department of Public Health, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, AP-HP, Hôpitaux Universitaires Pitié-Salpêtrière—Charles Foix, Sorbonne Université, Paris, France (A.B., L.B.); Department of Haematology, Institut Curie, Site Saint-Cloud and INSERM U932 Institut Curie, Université PSL, Paris, France (C.S.); Department of Nuclear Medicine, Institut Curie, St-Cloud, France (V.E.); Department of Haematology, Cochin Hospital, AP-HP, Paris (L.W.); Department of Haematology, Groupe Hospitalier Pitié-Salpêtrière, AP-HP, Sorbonne Université, Paris, France (S.C.); Department of Neurology, AP-HP, Hôpital Saint-Louis, Paris, France (R.U.); Department of Neuroradiology, Groupe Hospitalier Pitié-Salpêtrière, AP-HP, Sorbonne Université, Paris, France (D.G.)

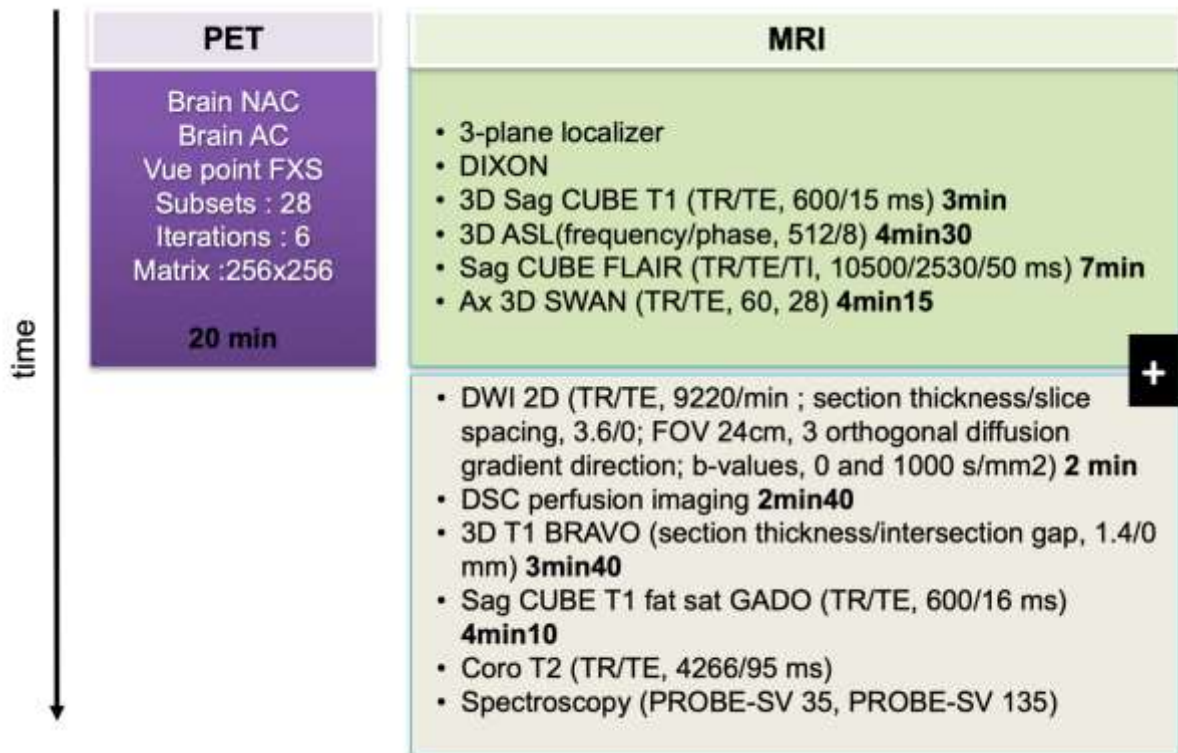
References

1. Lauw MIS, Lucas CHG, Ohgami RS, Wen KW. Primary central nervous system lymphomas: a diagnostic overview of key histomorphologic, immunophenotypic, and genetic features. *Diagnostics (Basel)*. 2020;10(12):1076.
2. Thiel E, Korfel A, Martus P, et al. High-dose methotrexate with or without whole brain radiotherapy for primary CNS lymphoma (G-PCNSL-SG-1): a phase 3, randomised, non-inferiority trial. *Lancet Oncol*. 2010;11(11):1036–1047.
3. Morales-Martinez A, Lozano-Sanchez F, Duran-Peña A, Hoang-Xuan K, Houillier C. Primary central nervous system lymphoma in elderly patients: management and perspectives. *Cancers (Basel)*. 2021;13(14):3479.
4. Houillier C, Soussain C, Ghesquière H, et al. Management and outcome of primary CNS lymphoma in the modern era: an LOC network study. *Neurology*. 2020;94(10):e1027–e1039.
5. Hoang-Xuan K, Deckert M, Ferreri AJM, et al. European Association of Neuro-Oncology (EANO) guidelines for treatment of primary central nervous system lymphoma (PCNSL). *Neuro Oncol*. 2023;25(1):37–53.
6. Hernández-Verdin I, Kirasic E, Wienand K, et al. Molecular and clinical diversity in primary central nervous system lymphoma. *Ann Oncol*. 2023;34(2):186–199.S0923753422047329.
7. Rachdi A, Hernandez-Tost H, Herzi D, et al. Recent advances in the diagnosis and the treatment of primary CNS lymphoma. *Rev Neurol (Paris)*. 2023;179(5):481–489.
8. Morales-Martinez A, Nichelli L, Hernandez-Verdin I, et al. Prognostic factors in primary central nervous system lymphoma. *Curr Opin Oncol*. 2022;34(6):676–684.
9. Barajas RF, Politi LS, Anzalone N, et al. Consensus recommendations for MRI and PET imaging of primary central nervous system lymphoma: guideline statement from the International Primary CNS Lymphoma Collaborative Group (IPCG). *Neuro Oncol*. 2021;23(7):1056–1071.

10. Tabouret E, Houillier C, Martin-Duverneuil N, et al. Patterns of response and relapse in primary CNS lymphomas after first-line chemotherapy: imaging analysis of the ANOCEF-GOELAMS prospective randomized trial. *Neuro Oncol.* 2017;19(3):422–429.
11. Abroy LE, Ben-Porat L, Panageas KS, et al. Primary central nervous system lymphoma: the Memorial Sloan-Kettering Cancer Center prognostic model. *J Clin Oncol.* 2006;24(36):5711–5715.
12. Siddiqi T, Wang X, Blanchard MS, et al. CD19-directed CAR T-cell therapy for treatment of primary CNS lymphoma. *Blood Adv.* 2021;5(20):4059–4063.
13. van der Meulen M, Postma AA, Smits M, et al. Extent of radiological response does not reflect survival in primary central nervous system lymphoma. *Neurooncol Adv.* 2021;3(1):vdab007.
14. Cheson BD, Fisher RI, Barrington SF, et al.; Alliance, Australasian Leukaemia and Lymphoma Group. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol.* 2014;32(27):3059–3068.
15. Rekowski J, Hüttmann A, Schmitz C, et al. Interim PET evaluation in diffuse large B-cell lymphoma using published recommendations: comparison of the Deauville 5-point scale and the Δ SUVmax method. *J Nucl Med.* 2021;62(1):37–42.
16. Meignan M, Gallamini A, Meignan M, Gallamini A, Haioun C. Report on the First International Workshop on Interim-PET-Scan in Lymphoma. *Leuk Lymphoma.* 2009;50(8):1257–1260.
17. Lin C, Itti E, Haioun C, et al. Early 18F-FDG PET for prediction of prognosis in patients with diffuse large B-cell lymphoma: SUV-based assessment versus visual analysis. *J Nucl Med.* 2007;48(10):1626–1632.
18. Rozenblum L, Galanaud D, Houillier C, et al. [18F]FDG PET-MRI provides survival biomarkers in primary central nervous system lymphoma in the elderly: an ancillary study from the BLOCAGE trial of the LOC network. *Eur J Nucl Med Mol Imaging.* 2023;50(12):3684–3696.
19. Rozenblum L, Houillier C, Soussain C, et al. Role of positron emission tomography in primary central nervous system lymphoma. *Cancers (Basel).* 2022;14(17):4071.
20. Rozenblum L, Bertaux M, Bielle F, et al. 18F-FDOPA PET/CT findings in a patient with primary cerebral amyloidoma. *Clin Nucl Med.* 2020;45(4):e206–e207.
21. Albano D, Bosio G, Bertoli M, Giubbini R, Bertagna F. 18F-FDG PET/CT in primary brain lymphoma. *J Neurooncol.* 2018;136(3):577–583.
22. Jo JC, Yoon DH, Kim S, et al. Interim 18F-FDG PET/CT may not predict the outcome in primary central nervous system lymphoma patients treated with sequential treatment with methotrexate and cytarabine. *Ann Hematol.* 2017;96(9):1509–1515.
23. Mercadal S, Cortés-Romera M, Vélez P, et al. Positron emission tomography combined with computed tomography in the initial evaluation and response assessment in primary central nervous system lymphoma. *Med Clin (Barc).* 2015;144(11):503–506.
24. Birsan R, Blanc E, Willems L, et al. Prognostic value of early 18F-FDG PET scanning evaluation in immunocompetent primary CNS lymphoma patients. *Oncotarget.* 2018;9(24):16822–16831.
25. Maza S, Buchert R, Brenner W, et al. Brain and whole-body FDG-PET in diagnosis, treatment monitoring and long-term follow-up of primary CNS lymphoma. *Radiol Oncol.* 2013;47(2):103–110.
26. de-Bonilla-Damiá A, Fernández-López R, Capote-Huelva FJ, et al. Role of 18F-FDG PET/CT in primary brain lymphoma. *Rev Esp Med Nucl Imagen Mol.* 2017;36(5):298–303.
27. Palmedo H, Urbach H, Bender H, et al. FDG-PET in immunocompetent patients with primary central nervous system lymphoma: correlation with MRI and clinical follow-up. *Eur J Nucl Med Mol Imaging.* 2006;33(2):164–168.
28. Rosen BR, Belliveau JW, Vevea JM, Brady TJ. Perfusion imaging with NMR contrast agents. *Magn Reson Med.* 1990;14(2):249–265.
29. Ostergaard L, Weisskoff RM, Chesler DA, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: mathematical approach and statistical analysis. *Magn Reson Med.* 1996;36(5):715–725.
30. Valles FE, Perez-Valles CL, Regalado S, et al. Combined diffusion and perfusion MR imaging as biomarkers of prognosis in immunocompetent patients with primary central nervous system lymphoma. *AJNR Am J Neuroradiol.* 2013;34(1):35–40.
31. Boellaard R, Delgado-Bolton R, Oyen WJG, et al.; European Association of Nuclear Medicine (EANM). FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging.* 2015;42(2):328–354.
32. Ellingson BM, Bendszus M, Boxerman J, et al.; Jumpstarting Brain Tumor Drug Development Coalition Imaging Standardization Steering Committee. Consensus recommendations for a standardized brain tumor imaging protocol in clinical trials. *Neuro Oncol.* 2015;17(9):1188–1198.
33. Kaufmann TJ, Smits M, Boxerman J, et al. Consensus recommendations for a standardized brain tumor imaging protocol for clinical trials in brain metastases. *Neuro Oncol.* 2020;22(6):757–772.
34. Furutani K, Harada M, Mawlan M, Nishitani H. Difference in enhancement between spin echo and 3-dimensional fast spoiled gradient recalled acquisition in steady state magnetic resonance imaging of brain metastasis at 3-T magnetic resonance imaging. *J Comput Assist Tomogr.* 2008;32(2):313–319.
35. Ambady P, Hu LS, Politi LS, Anzalone N, Barajas RF. Primary central nervous system lymphoma: advances in MRI and PET imaging. *Ann Lymphoma.* 2021;5:27.
36. Trotman J, Barrington SF. The role of PET in first-line treatment of Hodgkin lymphoma. *Lancet Haematol.* 2021;8(1):e67–e79.
37. Ambady P, Fu R, Netto JP, et al. Patterns of relapse in primary central nervous system lymphoma: inferences regarding the role of the neurovascular unit and monoclonal antibodies in treating occult CNS disease. *Fluids Barriers CNS.* 2017;14(1):16.
38. Schulte-Altedorneburg G, Heuser L, Pels H. MRI patterns in recurrence of primary CNS lymphoma in immunocompetent patients. *Eur J Radiol.* 2012;81(9):2380–2385.
39. Pyatigorskaya N, Habert MD, Rozenblum L. Contribution of PET-MRI in brain diseases in clinical practice. *Curr Opin Neurol.* 2020;33(4):430–438.
40. Postnov A, Toutain J, Pronin I, et al. First-in-man noninvasive initial diagnostic approach of primary CNS lymphoma versus glioblastoma using PET with 18F-fludarabine and I-[methyl-11C]methionine. *Clin Nucl Med.* 2022;47(8):699–706.
41. Jang SJ, Lee KH, Lee JY, et al. (11C)-methionine PET/CT and MRI of primary central nervous system diffuse large B-cell lymphoma before and after high-dose methotrexate. *Clin Nucl Med.* 2012;37(10):e241–e244.
42. Ahn SY, Kwon SY, Jung SH, et al. Prognostic significance of interim 11C-methionine PET/CT in primary central nervous system lymphoma. *Clin Nucl Med.* 2018;43(8):e259–e264.
43. Zhai Y, Zhou X, Wang X. Novel insights into the biomarkers and therapies for primary central nervous system lymphoma. *Ther Adv Med Oncol.* 2022;14:175883592210937.

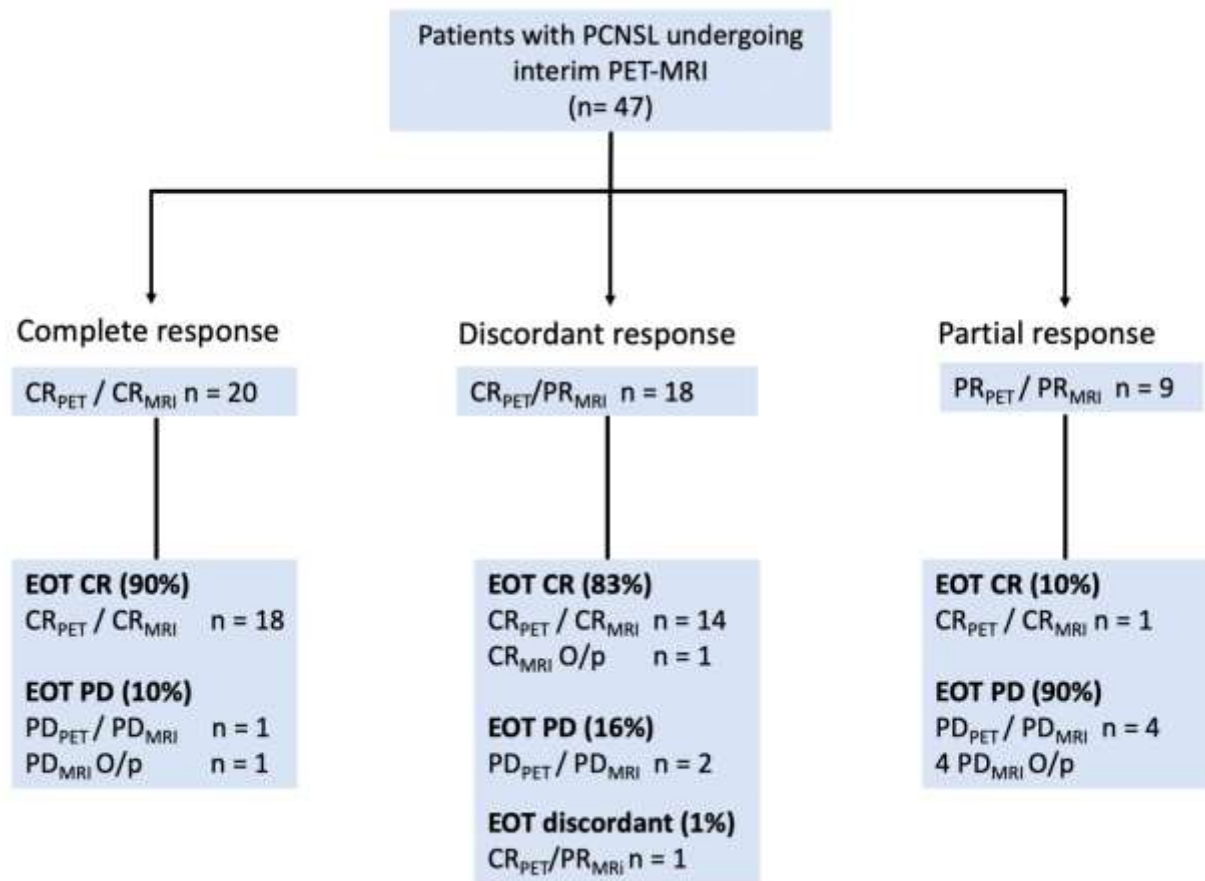
Supplemental Data

Supplemental Figure 1. PET-MRI protocol



NAC: Non-attenuation corrected, AC: Attenuation corrected, 3D: three dimensions, 2D: two dimensions, ASL: Arterial spin labelling, FLAIR: Fluid Attenuation Inversion Recovery, Swan: Susceptibility weighted angiography, DWI: Diffusion-weighted imaging, DSC: Dynamic susceptibility contrast, BRAVO: Brain volume imaging, Fat sat gado: Fat saturation gadolinium, Coro T2: Coronal T2, Probe-sv: Proton brain exam-single voxel

Supplemental Figure 2. Visual assessment at interim and EOT PET-MRI

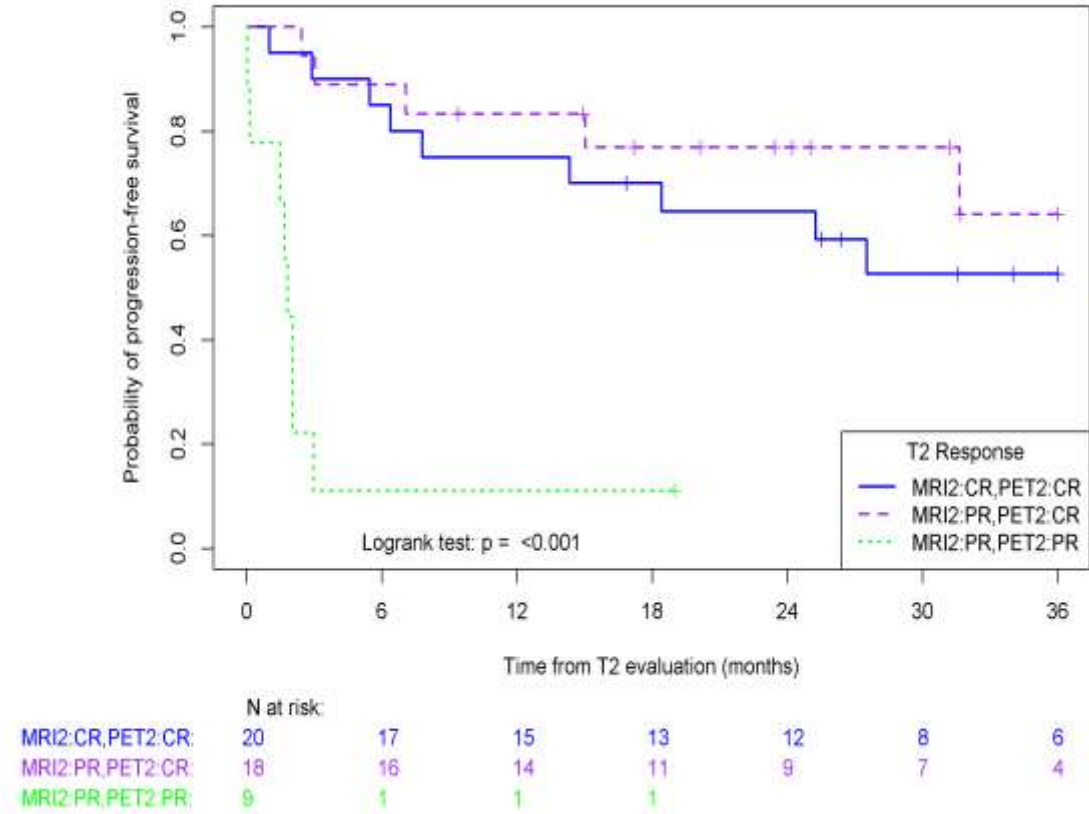


CR : complete response; PR : partial response, PG : progression, EOT : end of therapy

O/p : out of protocol MRI (no PET performed)

Of note, 2 patients did not have an evaluable interim PET-MRI (one not performed, one performed without enhancement) but had an EOT PET-MRI

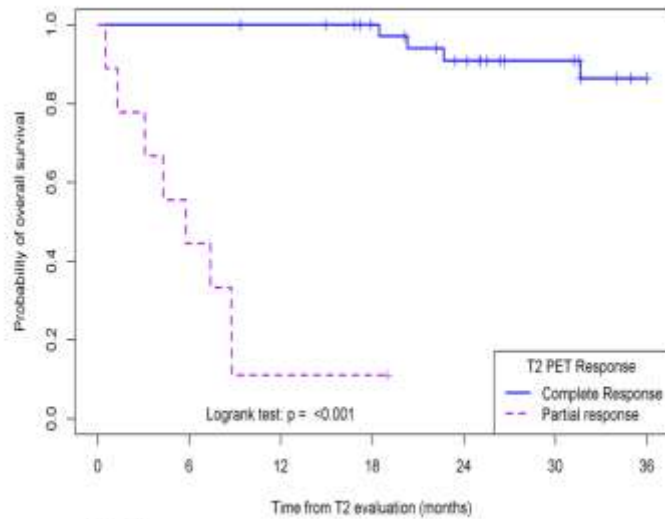
Supplemental Figure 3. Kaplan-Meier survival curves depicting the progression-free survival of PCNSL patients at 36 months according to combined FDG-PET and MRI results after 2 cycles of chemotherapy (interim evaluation)



T2 : Time 2 (interim evaluation), CR : complete response, PR : partial response

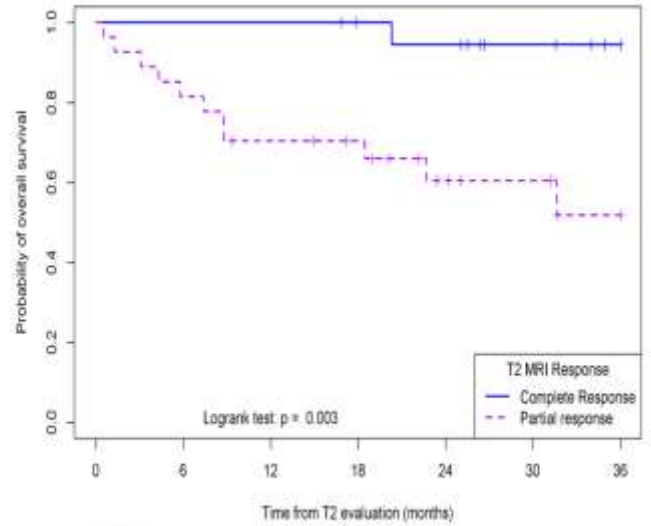
Supplemental figure 4. Kaplan-Meier survival curves depicting the overall survival of PCNSL patients at 36 months according to FDG-PET and MRI results after 2 cycles of chemotherapy (interim evaluation)

A. Interim PET



	0	6	12	18	24	30	36
Complete Response	39	39	38	34	28	22	15
Partial response	9	4	1	1			

B. Interim MRI



	0	6	12	18	24	30	36
Complete Response	20	20	20	18	17	13	10
Partial response	27	22	18	15	10	8	5

Supplemental Table 1. Performances metrics of interim PET and MRI.

	Indicator	Estimation %	95% CI Bootstrap		Indicator	Estimation %	95% CI Bootstrap
	Interim MRI	Sensitivity	85.71		[64.29-100]	Interim PET	Sensitivity
Specificity		54.55	[37.14-71.79]	Specificity	97.06		[90.33-100]
PPV		44.44	[25.81-62.96]	PPV	88.89		[61.54-100]
NPV		90.00	[75-100]	NPV	84.62		[72.5-94.87]

PPV: Positive Predictive Value, NPV: Negative Predictive Value, CI: Confidence Interval.

Supplemental Table 2. Visual and semi-quantitative analyses of PET-positive exams at baseline, interim and end of therapy

Exam	N	Visual analysis		SUVmax Median [IQR]	TBRmax Median [IQR]	MTV Median [IQR]	TLG Median [IQR]
		Good Responders	Poor Responders				
PET1	54			27.7 [22.8-36]	5.3 [3.8-7.9]	5 [2.1-13.3]	95 [30.4-249.3]
PET2	48	39	9	19.8 [11.1-23.2]	3.0 [2.4-6.7]	2.6 [0.7-3.4]	37 [3.8-70.1]
PET3	43	36	7	12.5 [7.8-33.1]	2.8 [2.1-4.1]	5.5 [3.1-13.2]	80.6 [23.5-270.5]

n: number of patients analyzed, SUV: Standard Uptake Value, TBR: Tumor-to-Background Ratio, MTV: Metabolic Tumor Volume, TLG: Total Lesion Glycolysis, IQR: Interquartile Range.

Supplemental Table 3. Multivariate analysis of clinical and PET-MRI scans based on visual analysis on OS

OS	Multivariate analysis (PET) n=48		Multivariate analysis (MRI) n=47	
	HR (95% CI)	p	HR (95% CI)	p
Male (ref F)	2.56 [0.64,10.25]	0.19	1.61 [0.47,5.54]	0.45
Age (ref < 70 years)	6.48 [1.14,36.95]	0.02	1.77 [0.46,6.89]	0.39
KPS (ref < 70)	0.37 [0.10,1.36]	0.12	0.76 [0.23,2.58]	0.66
INTERIM PET response (ref CR)	129.17[13.35,1249.74]	<0.001		
INTERIM MRI response (red CR)			9.47 [1.16,77.13]	0.007

Supplemental Table 4. Univariate analysis of clinical and PET-MRI scans on OS

OS	Univariate analysis n=48 for PET, n=47 for MRI	
	HR (95% CI)	p
Male (ref F)	1.21 [0.39,3.83]	0.74
Age (ref < 70 years)	2.31 [0.62,8.54]	0.20
KPS (ref < 70)	0.50 [[0.16,1.57]	0.22
INTERIM PET response (ref Complete Response)	87.76 [10.37,742.56]	<0.001
INTERIM MRI response (red Complete Response)	11.40 [1.46,88.81]	0.003
ΔSUVmax ₁₋₂ (increase unit =10)	0.90 [0.79,1.02]	0.81*
ΔMTV ₁₋₂ (ref< 100)	0.03 [0.01,0.16]	<0.001*
ΔVol. ₁₋₂ enhanced lesion (ref:[33.9,98.3])	0.26 [0.07,0.96]	0.30*
ΔrADCmean ₁₋₂ (increase unit =10)	1.01 [0.90,1.13]	1.00*
ΔADCmin ₁₋₂ (increase unit =10)	1.01 [0.90,1.14]	1.00*
ΔCBFmax ₁₋₂ (ref:[-267,18.2])	0.96 [0.28,3.34]	1.00*
ΔCBVmax ₁₋₂ (increase unit =10)	0.97[0.93,1.01]	0.90

* p-values reported were corrected using Benjamini-Hochberg to take into account multiple comparisons computed in the analysis.

Chapitre 5. Analyse radiomique en TEP

I. Introduction générale à la radiomique

Le domaine de la radiomique, émergé en 2012 [149], s'est épanoui sur la révélation percutante de Gillies en 2016 : "Les images sont plus que des images, elles sont des données" [150]. L'essor académique de ce concept se mesure à l'évolution exponentielle de la littérature scientifique (moins de 10 articles en 2014, environ 250 en 2018, et près de 2000 en 2020) [151].

L'objectif d'une analyse radiomique est d'analyser l'image médicale au-delà du visuel en disséquant la matrice de l'image à l'échelle du voxel, extrayant des caractéristiques de texture spatiales complexes, inappréciables à l'œil humain. L'explication de l'engouement pour ce type d'analyse, repose sur des découvertes génomiques montrant que le degré d'hétérogénéité de la tumeur est un des marqueurs pronostiques déterminant en oncologie [152–155]. Par ailleurs, il a été montré une forte corrélation entre les caractéristiques radiomiques et les indices d'hétérogénéité au niveau cellulaire. Ainsi, l'analyse des paramètres radiomiques d'une image de cancer permettrait de déduire l'hétérogénéité tumorale *in-vivo*, avec pour avantage d'étudier l'ensemble de la tumeur ; au contraire d'une analyse histologique, qui par définition n'étudie qu'un échantillon du volume tumoral. De nombreuses études en cancérologie ont ainsi montré des « signatures radiomiques » associées à un profil d'agressivité de tumeurs [62,156,157].

En neuro-oncologie, domaine où la médecine personnalisée progresse rapidement, la radiomique a confirmé son potentiel. Des investigations récentes ont mis en lumière ses applications novatrices, notamment dans l'évaluation des gliomes [158–160]. Plusieurs équipes ont exploré le potentiel de la radiomique pour la détection de mutations spécifiques, telles que celle de l'isocitrate déshydrogénase (IDH), via des signatures radiomiques distinctives, et pour différencier la radionécrose de la progression tumorale chez les patients atteints de glioblastome traités par radiothérapie [161–163]. Par ailleurs, quelques études ont également investigué l'existence de signatures radiomiques prédictives de la survie dans les gliomes [164,165]. Cependant, les études portant sur le LPSNC sont encore rares. La plupart se sont axées sur le diagnostic différentiel entre les glioblastomes multiformes (GBM) et les LPSNC, pathologies aux caractéristiques radiologiques en IRM parfois similaires, ce qui complexifie le diagnostic, en particulier chez les patients âgés aux comorbidités multiples, chez qui une biopsie s'avère souvent impraticable [166].

Peu d'études ont jusqu'à maintenant exploré l'intérêt de la radiomique pour prédire la réponse de fin de traitement ou la survie des patients atteints de LPSNC. Les rares disponibles portent exclusivement sur l'IRM. Le tableau présenté ci-dessous rapporte quelques études de radiomique dans les LPSNC (Tableau 5.1). L'étude menée par Chen et al. a analysé rétrospectivement les paramètres radiomiques chez 52 patients, montrant que de faibles valeurs de GLCM-contrast et de GLCM-dissimilarity, associées à une valeur élevée de GLCM-homogeneity, étaient liées à une survie globale réduite [167]. En effet, une valeur élevée de GLCM-homogeneity suggère une hétérogénéité accrue des niveaux de gris sur la matrice de cooccurrence, indiquant potentiellement la présence de néovascularisation ou de zones d'hypoxie/nécrose, des facteurs de mauvais pronostic connus par ailleurs. D'autres études, comme celles de She et al. et de Destito et al., ont tenté de développer des modèles basés sur les paramètres radiomiques, en utilisant des méthodologies de machine learning ou de deep learning mais sans sélection préalable de caractéristiques [168,169]. De son côté, l'étude récente de Wang et al. a comparé plusieurs modèles dont un radiomique (issu de l'IRM T1 avec injection), un clinique, et un combinant des paramètres radiomiques et biologiques. Ce dernier a montré les meilleures performances, surpassant le modèle basé uniquement sur les critères cliniques [170].

Study	Training/testing	External validation	Aim	Modality	Classifier	Results
Wang, J. Neurooncol, 2024	131/30	no	Prediction of end of treatment response	CE-T1 and bio-clinical data	Logistic regression Support Vector Machine	AUC = 0.86 (95% CI: 0.68, 0.96) for radiomic model
She, Int J Comput Assist Radiol Surg, 2023	56/0	no	Prediction of OS	T1, T2, CE-T1 and bio-clinical data	Logistic regression Support Vector Machine	AUC = 0.81 ± 0.03 for radiomic model using CET1 : on cross-validation
Destito, Bioengineering, 2023	56/0 for OS 47/0 for PFS	no	Prediction of one-year OS and PFS Compare MRI normalization	T1, T2, CE-T1	Extra Tree Classifier Support Vector Machine Logistic regression Random Forest kNeighbors	AUC = 0.86 ± 0.09 for radiomic model predicting OS (on training) AUC = 0.84 ± 0.13 for radiomic model predicting PFS (on training)
Chen, 2019, Contrast Media & Molecular Imaging	52/0	no	Prediction of OS	T1	No use of classifier	OS rate worse in case of low GLCM-Contrast low GLCM-Dissimilarity high GLCM-Homogeneity

Tableau 5.1 Données de la littérature sur l'intérêt des radiomiques dans le LPSNC

Alors que la littérature scientifique sur la radiomique connaît une croissance rapide, l'intégration clinique des signatures radiomiques reste aujourd'hui limitée. Cette situation découle principalement d'une méthodologie de recherche insuffisamment rigoureuse et de l'absence de normes d'interprétation solides pour les reviewers et éditeurs.

Les défis incluent l'omission de cohorte de validation, le manque d'évaluation externe à large échelle, et une compréhension limitée de la pertinence biologique des signatures radiomiques [171]. Dans le contexte de la TEP, la dépendance des paramètres radiomiques aux modalités d'acquisition et de reconstruction est un facteur critique ; la validation de la fiabilité des indices avant leur intégration dans les signatures radiomiques est donc essentielle [172].

Pour surmonter ces obstacles, des initiatives comme l'Imaging Biomarker Standardization Initiative (IBSI), le Radiomics Quality Score (RQS) ou encore le score METRICS proposé par Kocak et al. en 2024 ont été développées, visant à uniformiser la procédure d'extraction des biomarqueurs et l'approche méthodologique [173–175]. En médecine nucléaire, les recommandations conjointes de l'EANM et du SNMMI en 2023 constituent un progrès notable pour l'adoption des radiomiques [151].

Actuellement, la communauté scientifique explore deux stratégies distinctes concernant la méthodologie en radiomique. La première s'appuie sur l'extraction de paramètres analytiques standardisés et leur évaluation via des techniques d'apprentissage automatique. La seconde préfère l'extraction directe d'informations à partir des images par des réseaux de neurones profonds, menant à des tâches de classification répondant à la question posée. En d'autres termes, les données d'entrée (*input*) seront différentes : paramètres quantitatifs dans le premier cas et images brutes dans le second ; mais la prédiction recherchée (*output*) sera la même.

Le troisième volet de ma thèse a porté sur la comparaison de ces deux méthodes : l'analyse radiomique basée sur l'apprentissage automatique et l'apprentissage profond, appliquée aux LPSNC chez le patient âgé immunocompétent en TEP-IRM. Les chapitres suivants décriront avec précision notre protocole de recherche, mettant en lumière l'adhésion aux recommandations actuelles et les stratégies mises en œuvre pour contourner tout biais potentiel. Les résultats préliminaires obtenus seront aussi discutés en détail.

Workflow en radiomique

Définition de la question et estimation des biais potentiels

Le premier point d'attention dans une analyse radiomique est la définition de la problématique de recherche. La sélection des sujets doit non seulement être rigoureuse mais également attentive aux biais potentiels. Un modèle nouvellement créé sera performant au sein d'une population définie, répondant aux critères d'inclusion de l'étude et pourra ne pas être aussi efficient dans un autre type de population (sexe, âge ou type d'établissement de soins différents). Pour renforcer la capacité de généralisation d'un modèle radiomique, un large éventail de données est primordial. Néanmoins, la multitude des variables radiomiques extractibles et la complexité intrinsèque des analyses multimodales d'imagerie exigent une surveillance méticuleuse pour éviter le risque d'erreurs statistiques découlant des comparaisons multiples [151].

Dans les domaines du machine learning et du deep learning, il est essentiel de différencier deux sous-ensembles de données pour valider un modèle. Les données de training servent à la construction du modèle. Ces données peuvent être elles-mêmes être subdiviser pour comporter un sous-ensemble de validation, utilisées pour ajuster les hyperparamètres, qui sont des paramètres de configuration du modèle pouvant influencer significativement la performance de l'algorithme d'apprentissage. Le deuxième sous-ensemble, constitué des données de test qui n'ont été utilisées ni pour la construction ni pour l'ajustement du modèle, est destiné à l'évaluation finale de la performance. Idéalement, ces données doivent être issues d'un contexte différent, comme une période temporelle distincte, des caméras différentes, ou encore mieux, d'une population recueillie dans un autre centre.

Notre recherche s'est concentrée sur l'évaluation radiomique des LPSNC, une pathologie dont la rareté impose des contraintes sur l'effectif de patients pouvant être étudiés. Malgré cette limitation, nous avons adopté une méthodologie rigoureuse, suivant les dernières recommandations des sociétés savantes [151,173]. La population pour la phase d'entraînement (training) provenait de la cohorte LOCALYZE, que nous avons détaillée antérieurement. Pour le testing, nous avons sélectionné 23 patients indépendants, atteints de LPSNC confirmé histologiquement, tous âgés de plus de 60 ans, et ayant eu une TEP-CT cérébrale (mCT Flow de Siemens) dans le service de médecine nucléaire de l'hôpital de la Pitié-Salpêtrière, entre 2019 et 2023.

Pré-traitement des images

La première phase de notre recherche a été l'élaboration de modèles d'apprentissage automatique, basés sur des données radiomiques, préalablement extraites d'un volume d'intérêt. Nous détaillerons ici les étapes nécessaires à cette analyse.

1. Correction des inhomogénéités de champs magnétique

Les inhomogénéités à basse fréquence dans le champ magnétique d'une IRM peuvent entraîner des distorsions de signal, affectant la fidélité des paramètres de texture, notamment dans les régions périphériques par rapport au centre de l'aimant. Bien que leur impact sur l'interprétation clinique soit mineur, il est impératif de les corriger pour une analyse radiomique précise des paramètres de texture de 1^{er} et 2nd ordre. La méthode N4, reconnue pour sa fiabilité dans divers contextes anatomiques, a été sélectionnée pour sa validation dans des études multicentriques [176,177].

2. Normalisation des images en intensité

Cette étape est essentielle dans le prétraitement radiomique, car elle uniformise les valeurs des pixels sur l'ensemble des images. En particulier, en IRM, les valeurs de voxels ne reposent pas sur une échelle standardisée mais sur des valeurs arbitraires qui dépendent des caractéristiques de la caméra, du champ magnétique, et des paramètres d'acquisition, et peuvent varier d'un patient à l'autre, même s'ils ont été acquis sur la même machine.

En IRM, les principales méthodes de normalisation rapportées dans la littérature sont :

- le Z-score : la normalisation par Z-score ajuste les valeurs de l'image ($I(x)$) en soustrayant l'intensité moyenne du signal (μ) et divisant par l'écart type de tous les voxels du cerveau (σ)

$$I_{norm}(x) = \frac{I(x) - \mu}{\sigma}$$

Facile et rapide à mettre en œuvre, cette technique s'est révélée particulièrement performante, comme en témoigne l'analyse réalisée par Destito et al., où elle a été identifiée comme la plus à même de générer un volume important de caractéristiques radiomiques fiables [169].

- la méthode WhiteStripe est spécifique au cerveau et permet de redéfinir la valeur des pixels en fonction des valeurs normales observées dans la substance blanche (ws). Pour cela elle normalise l'intensité du signal en soustrayant I_x par la valeur moyenne de la substance blanche (μ_{ws}) et divisant par l'écart type de la valeur des voxels de la substance blanche (σ_{ws}).

Cette méthode implique en revanche qu'il n'y ait pas d'atteinte dans la substance blanche.

$$I_{ws}(x) = \frac{(I(x) - \mu_{ws})}{\sigma_{ws}}$$

- Min-Max : Cette approche standardise l'image en recalibrant l'éventail des valeurs à l'intervalle [0, 1], via l'équation ci-dessous, $\min(x)$ et $\max(x)$ correspondant respectivement aux valeurs d'intensité les plus faibles et les plus élevées dans l'image.

$$I_{norm}(x) = \frac{I(x) - \min(x)}{\max(x) - \min(x)}$$

- La méthode proposée par Nyul et al. [178,179], appelée normalisation d'histogramme par correspondance, procède d'abord à l'établissement d'un histogramme standard à partir d'un jeu d'images de référence, lequel inclut des points de repère à des percentiles spécifiques. Une fois cet histogramme standard défini, il sert à la normalisation des images test. Cela se fait en alignant linéairement les valeurs des images test avec celles de l'histogramme de référence, par tranches déterminées par ces percentiles.

Dans le cas des images TEP, la conversion des données consiste simplement à transformer les images de Bq.mL-1 en SUV, selon la formule explicitée plus haut dans cette thèse.

Pour la normalisation en intensité, que ce soit en TEP ou en IRM, il est également possible d'opter pour une harmonisation *a posteriori* des données numériques extraites. La méthode ComBat, largement reconnue et adoptée suite aux travaux de Fortin et al. et Orhac et al. [180–183], se distingue par sa simplicité d'emploi et son efficacité de calcul. Cette technique nécessite cependant que chaque lot (ou batch) comprenne au minimum 20 patients examinés avec le même équipement d'imagerie et les mêmes paramètres d'acquisition et de reconstruction. Elle a toutefois pour effet de faire converger les données vers une moyenne arbitraire qui n'est représentative d'aucun des lots d'origine. Pour contourner ce biais, une variante de cette méthode, nommée M-ComBat, a été introduite par Stein et al. [184]. M-ComBat ajuste les données vers un lot de référence spécifique, plutôt que vers un consensus généralisé (Figure 5.1).

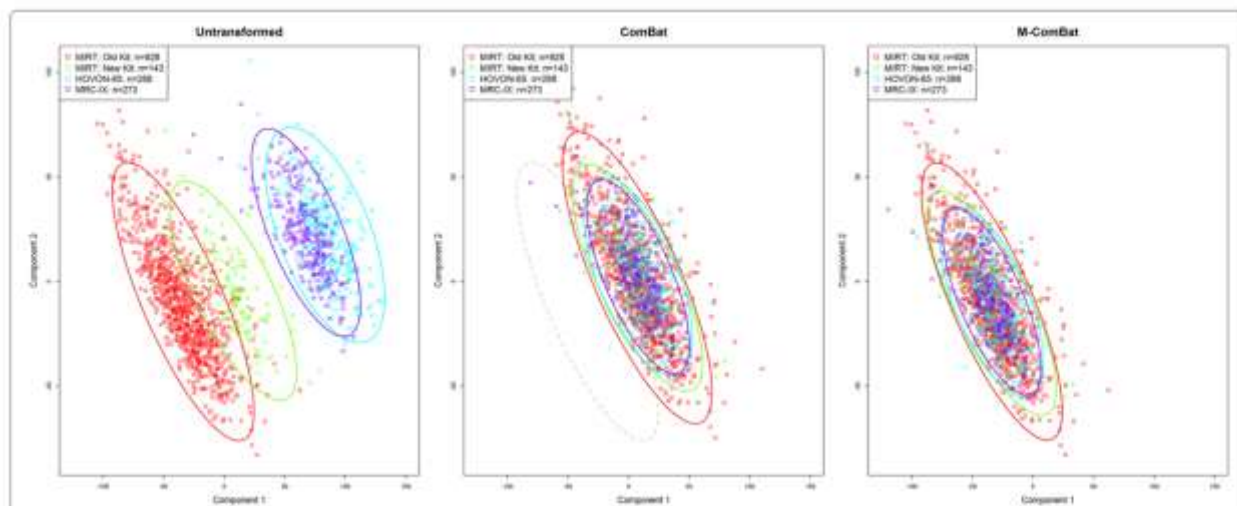


FIGURE 5.1 Comparaison de la distribution des données, après réduction de dimension par ACP, sans harmonisation, avec une harmonisation par ComBat et avec une harmonisation par M-ComBat.

D'après Stein et al. [184] Tandis que l'harmonisation avec ComBat réajuste les données vers un espace défini de manière arbitraire, M-ComBat, quant à elle, oriente les données vers l'espace d'une référence prédéterminée (ici MIRT : Old Kit)

Dans notre étude, une harmonisation a posteriori des données n'a pas été nécessaire, compte tenu du choix d'une population de training parfaitement homogène en termes d'acquisition et paramètres de reconstruction (population issue du protocole prospectif Localyze).

3. Segmentation

La segmentation du volume d'intérêt (VOI) constitue une autre étape importante du prétraitement en radiomique, visant à isoler avec précision la zone d'étude sur l'image. Bien que l'IBSI ne prescrive pas de méthode de segmentation spécifique, elle insiste sur l'importance de cette étape comme étant une source significative de variabilité dans les études radiomiques [174]. Conformément aux recommandations de la Société Européenne d'Informatique d'Imagerie Médicale, les procédés de segmentation, qu'ils soient manuels ou semi-automatiques, doivent être rigoureusement documentés [173].

Les directives communes de l'EANM/SNMIMI de 2023 soulignent plusieurs principes essentiels pour la segmentation en TEP [151]:

1. Privilégier une méthode de segmentation spécifique plutôt que sensible pour éviter les erreurs de quantification dues à l'effet de volume partiel sur les bords des images.

2. Nécessiter l'intervention d'un expert pour affiner les segmentations automatiques, par exemple celles établies sur des seuils tels que 41 % du SUVmax.
3. Favoriser les méthodes semi-automatiques pour garantir la reproductibilité des résultats.

De plus, des techniques de segmentation avancées utilisant des réseaux de neurones profonds sont en développement et pourraient offrir des avancées significatives en termes de reproductibilité et de gain de temps de traitement. Des recherches appliquées aux lésions gliales et aux métastases cérébrales en neuro-oncologie montrent des avancées prometteuses [185–188]. Toutefois, actuellement, aucun algorithme de segmentation basé sur ces technologies n'a été validé dans la littérature ou recommandé pour utilisation systématique dans la radiomique.

Pour notre étude TEP, nous avons adopté un seuillage à 41% du SUVmax, conformément aux lignes directrices EANM pour l'imagerie tumorale [67]. En IRM, nous avons privilégié une segmentation semi-automatique de la prise de contraste, ajustée manuellement. Pour les deux modalités, la segmentation a été réalisée via le logiciel LifeX [189] (Figure 5.2).

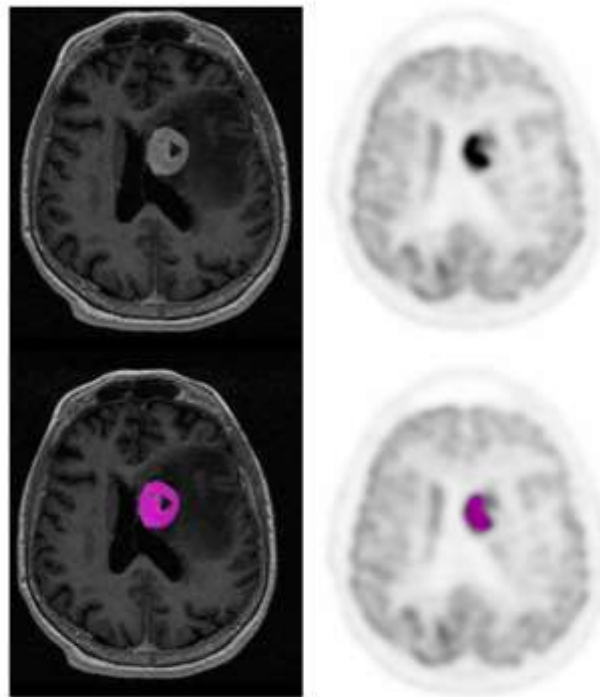


FIGURE 5.2 Exemples de masques sur l'IRM injecté et la TEP FDG

Les masques ont été réalisés grâce au logiciel LifeX, sur les séquences IRM 3DT1BRAVO avec injection de produit de contraste et TEP avec correction d'atténuation.

4. Ré-échantillonnage spatial

Les algorithmes d'extraction radiomique supposent que les voxels sont isotropes, c'est-à-dire uniformes dans les trois dimensions de l'espace. Lorsque cette uniformité fait défaut, un rééchantillonnage spatial devient nécessaire. De plus, certains attributs radiomiques sont influencés par la taille du voxel, rendant essentielle l'uniformisation de la taille des voxels pour permettre la comparabilité des images au sein d'une même étude.

Le rééchantillonnage spatial fait appel à des algorithmes d'interpolation tels que :

- Interpolation au plus proche voisin: Chaque voxel dans l'image rééchantillonnée reçoit la valeur du voxel le plus proche dans l'image originale. Cette méthode est rapide mais peut créer un effet d'escalier, où des blocs de pixels adjacents ont des valeurs identiques, ce qui donne un aspect artificiel à la texture.

- Interpolation trilinéaire: Elle calcule la valeur d'un voxel en utilisant une moyenne pondérée des valeurs des huit voxels les plus proches dans l'image originale, basée sur leur distance au voxel cible. Cette technique lisse les transitions entre les voxels et préserve mieux les gradients d'intensité et les textures fines, réduisant ainsi les artefacts.

- Interpolation tricubique: Une méthode plus sophistiquée qui prend en compte les valeurs des voxels environnants pour calculer les valeurs des voxels dans l'image rééchantillonnée, non seulement à partir des huit voxels les plus proches (comme dans l'interpolation trilinéaire) mais aussi de leurs voisins, en appliquant une fonction cubique. Cela permet une transition encore plus douce des intensités afin de mieux capturer les contrastes, mais peut conduire à l'introduction de valeurs extrêmes ou non réalistes, telles que des valeurs négatives [190,191].

Dans notre recherche, nous avons opté pour une interpolation trilinéaire, conformément aux recommandations de l'IBSI, pour rééchantillonner les images TEP et IRM à une résolution homogène de $2 \times 2 \times 2 \text{ mm}^3$, équilibrant la précision et la préservation de la qualité des textures originales.

5. Discrétisation

La discrétisation, étape finale du prétraitement d'images en tomographie par émission de positons (TEP), consiste à classer les intensités de signal des voxels en catégories distinctes, ou "bins", situés entre le SUV minimum et le SUV maximum de la lésion. Cette classification crée

des intervalles réguliers sur l'histogramme d'intensité, facilitant le calcul des attributs radiomiques. La discrétisation s'articule autour de deux variables principales : le nombre et la largeur des bins. Les approches de discrétisation se divisent en deux catégories :

1. La discrétisation à nombre de bins fixe, ou discrétisation relative, assigne une valeur discrète à chaque voxel basée sur sa position relative dans l'éventail des intensités allant du SUVmin au SUVmax, puis multiplie cette valeur par un nombre prédéterminé de bins. Cette technique peut conduire à des tailles de bins variables entre patients.

2. La discrétisation à taille de bin fixe, aussi appelée discrétisation absolue, calcule la valeur discrète d'un voxel par rapport à l'intensité minimale dans la zone d'intérêt et divise cette différence par une largeur de bin constante prédéfinie. Cette méthode génère un nombre de bins qui peut varier d'un patient à l'autre, contrairement à la discrétisation relative.

Une illustration issue de la publication d'Orlhac et al. en 2021 compare visuellement ces deux techniques de discrétisation (Figure 5.3).

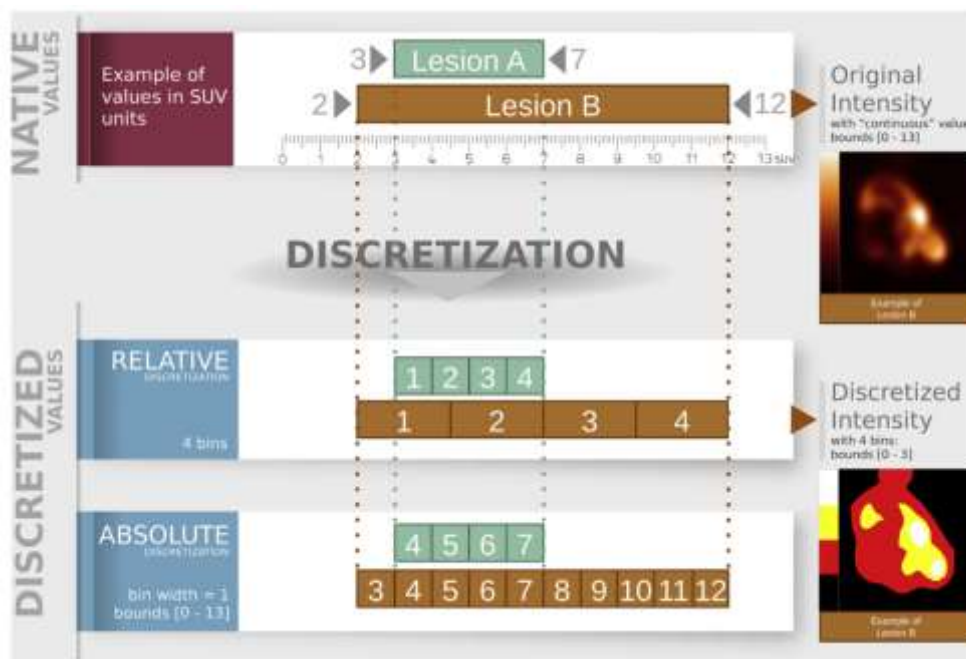


FIGURE 5.3. Discrétisation en intensité selon la méthode relative ou absolue.

D'après Orhac et al. [192] Ici, les auteurs illustrent l'étape de discrétisation pour 2 lésions (A et B), dont les intensités varient entre SUV=3 et SUV=7 pour la lésion A et entre SUV=2 et SUV=12 pour la lésion B, avant discrétisation. La discrétisation relative définit un nombre de bin fixe (ici 4) alors que la discrétisation absolue définit la largeur du bin (ici 1)

Selon les recommandations de l'IBSI, la discrétisation absolue doit être privilégiée en TEP, avec une borne minimale à 0 pour tous les patients. Dans notre étude, centrée sur des lésions avec une forte avidité pour le FDG, où le SUV maximal a atteint 64 chez l'un des sujets, nous avons choisi une discrétisation absolue. Les largeurs de bin ont été fixées à 0.3, aboutissant à un total de 214 bins (64 divisé par 0.3) pour chaque patient.

Extraction des paramètres radiomiques

Après le prétraitement, l'extraction des caractéristiques dans la zone d'intérêt (VOI) implique le calcul de plusieurs centaines de paramètres. Initialement, le manque d'uniformité dans la dénomination et la définition des caractéristiques radiomiques entravait l'interprétation et la comparabilité des recherches. C'est pour surmonter cette difficulté que l'IBSI a été créée. Fruit d'une collaboration multicentrique, l'IBSI vise à standardiser la nomenclature et les définitions des attributs radiomiques, et à établir des ensembles de données de référence pour la vérification et l'étalonnage des outils d'extraction radiomique. Aujourd'hui, plusieurs logiciels alignés sur les recommandations de l'IBSI, comme LIFEx, 3DSlicer, Matlab et PyRadiomics, sont disponibles pour l'extraction des caractéristiques radiomiques [193].

Les caractéristiques radiomiques sont regroupées en catégories distinctes [194]:

Indices de forme : Ces mesures, incluant volume, surface, périmètre et diamètre, évaluent la géométrie de la zone d'intérêt. La sphéricité en particulier, compare la géométrie de la zone d'intérêt à une forme sphérique.

Indices d'intensité et de premier ordre : Ces indices conventionnels décrivent l'intensité au sein du volume d'intérêt, avec des paramètres tels que la valeur minimale, maximale, moyenne, et l'écart-type. Les indices de premier ordre dérivent de l'histogramme des niveaux de gris, incluant notamment le skewness, qui mesure l'asymétrie de la distribution des intensités, et le kurtosis, qui évalue la concentration des valeurs autour de la moyenne et détecte les valeurs extrêmes dans une distribution normale (Figure 5.4).

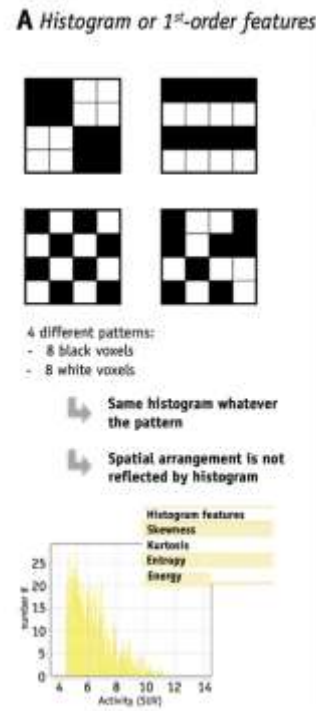


FIGURE 5.4 Représentation schématique des paramètres radiomiques de premier ordre.

D'après Reuzé et al. [195] Les caractéristiques d'histogramme ou de premier ordre reflètent uniquement la distribution de l'intensité des voxels. Pour différents motifs contenant le même nombre de pixels noirs et blancs, mais pas au même endroit, l'histogramme est identique. Par conséquent, les valeurs des caractéristiques basées sur l'histogramme (par exemple, l'asymétrie, le kurtosis, l'énergie) sont égales.

Indices de deuxième et troisième ordre :

Les indices de deuxième et troisième ordre approfondissent l'analyse radiomique en examinant les relations spatiales entre les pixels ou voxels, permettant une évaluation détaillée de l'hétérogénéité texturale de l'image :

- Matrice de cooccurrence des niveaux de gris (GLCM) : La GLCM est une matrice statistique qui quantifie la fréquence à laquelle des combinaisons de niveaux de gris apparaissent côte à côte dans une image, selon une direction et une distance spécifiées. Cela permet d'évaluer des attributs tels que le contraste (différence de luminance rendant une structure visible), la corrélation (comment les niveaux de gris varient par rapport à leur moyenne dans l'image), ou l'homogénéité (uniformité des distributions des niveaux de gris). [196]

- Matrice de longueur de parcours de niveau de gris (GLRLM) : La GLRLM évalue la longueur des séquences de voxels ayant la même valeur de niveau de gris dans une direction donnée. Elle permet de calculer des paramètres tels que la longueur de parcours courte ou longue, reflétant la prévalence de textures fines ou grossières respectivement (Figure 5.5).

- Matrice de taille de zone de niveau de gris (GLSZM): Cette matrice compte le nombre de zones contiguës pour chaque niveau de gris.
- Matrice de différence de tonalité de gris de voisinage (NGTDM) : La NGTDM mesure les variations de l'intensité entre un pixel ou voxel et ses voisins immédiats. Elle est utile pour estimer la complexité de la texture, car une image avec des variations locales importantes indiquera une texture plus complexe.
- Matrice de dépendance de niveau de gris de voisinage (NGLDM) : Cette matrice évalue si les pixels ou voxels environnants sont considérés comme connectés au pixel central, en fonction d'une plage définie de différences d'intensités [195].

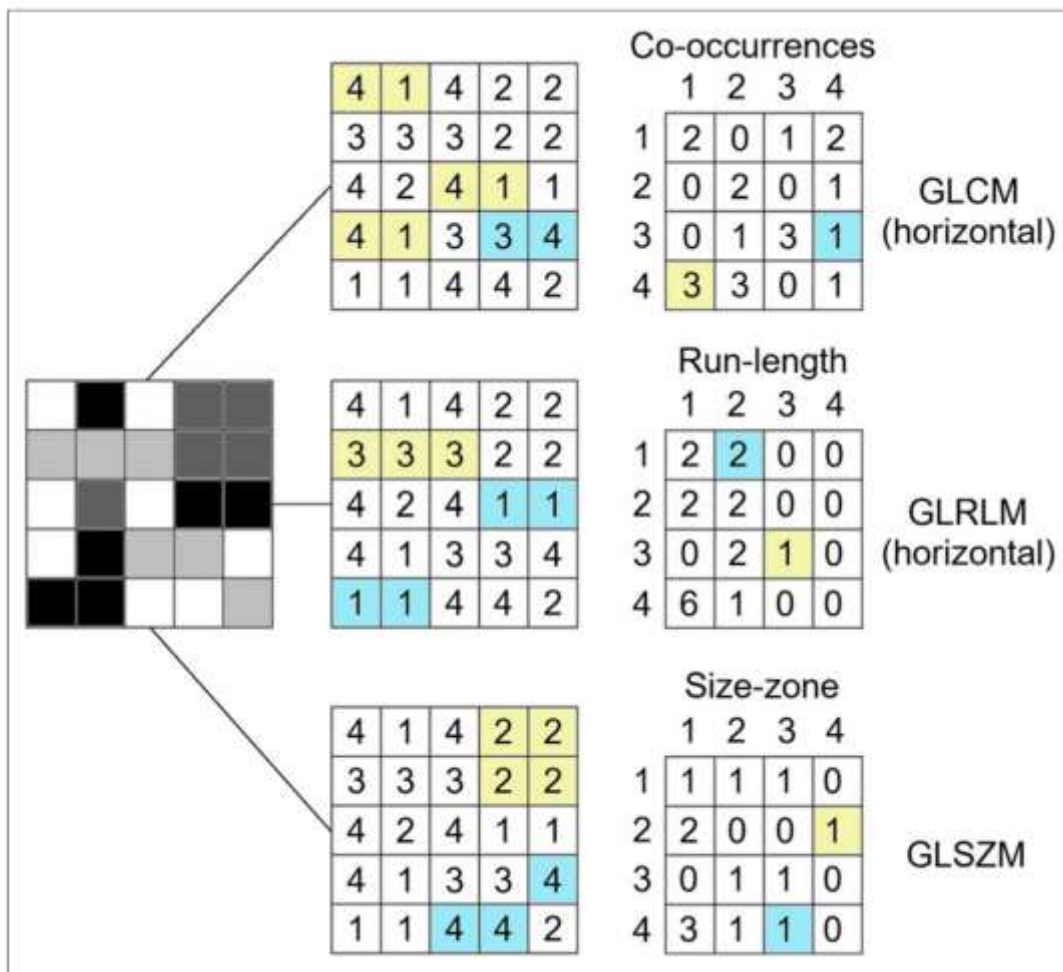


FIGURE 5.5 Représentation schématique des paramètres radiométriques de second et troisième ordre

D'après Mayerhoefer, 2020 [194], Figure 2. Calcul des caractéristiques de texture radiométriques. Tandis que la GLCM se base sur des paires de pixels, la GLRLM repose sur des séquences, et la GLSZM sur des zones de pixels voisins ayant le même niveau de gris.

Dans le cadre de ce travail, nous avons extrait les paramètres radiomiques en utilisant le package PyRadiomics de Python, à partir des VOI des séquences 3DT1BRAVO et PET. Pour chaque type d'image, nous avons extrait 107 paramètres, répartis comme suit :

- 18 paramètres de statistiques de premier ordre,
- 14 paramètres liés à la forme et à la taille,
- 75 paramètres de texture dérivés des matrices GLCM, GLRLM, GLSZM et NGLDM.

Sélection des paramètres radiomiques

Parmi les nombreux paramètres extraits, tous ne contribuent pas également à l'élaboration d'un modèle pronostique. Certains paramètres peuvent se révéler non pertinents, trop corrélés entre eux, ou redondants par rapport aux données cliniques ou autres mesures semi-quantitatives d'imagerie. Or, la valeur d'une signature radiomique réside dans sa capacité à apporter des informations supplémentaires au-delà des méthodes simplifiées. Les modèles radiomiques qui intègrent différentes modalités d'imagerie se doivent également de démontrer une performance accrue par rapport à ceux basés sur une seule modalité [173].

L'augmentation excessive de la dimensionnalité, par le biais d'un nombre élevé de modalités d'imagerie ou de paramètres sélectionnés, peut mener à un surajustement, également connu sous le terme d'overfitting [197]. Ce phénomène se produit lorsque le modèle s'ajuste trop précisément aux particularités des données d'entraînement, au point d'incorporer leurs anomalies spécifiques (bruit), ce qui affecte sa capacité à bien fonctionner avec de nouvelles données. Pour contenir la dimensionnalité, il est essentiel de procéder à une sélection rigoureuse des caractéristiques, qui peut être orientée par des critères méthodologiques ou cliniques. Dans notre étude, l'objectif était de comparer des modèles basés sur des données exclusivement cliniques, IRM puis TEP, avant d'explorer un modèle combiné potentiel.

Pour intégrer les paramètres radiomiques dans un modèle prédictif, il existe deux approches principales.

- La première, guidée par la clinique, repose sur une hypothèse forte concernant les paramètres qui seraient pertinents pour répondre à une question spécifique. Cette méthode présuppose que la question a déjà été explorée dans la littérature et qu'une justification logique peut être utilisée pour guider le choix des paramètres radiomiques. Néanmoins, cette approche peut s'avérer complexe, car il est rare que les études rapportent des résultats cohérents et une explication physiopathologique pour le choix des caractéristiques est rarement fournie. Il est également

possible de guider ce choix en évaluant individuellement chaque paramètre de manière univariée, en sélectionnant ceux qui montrent une association significative avec le résultat d'intérêt, tel que la réponse au traitement. Cette sélection peut être effectuée à l'aide de tests paramétriques, tels que le test de Student, ou non-paramétriques, comme le test de Wilcoxon, (Figure 5.6) [198,199]. Cependant, cette méthode a pour inconvénient de ne considérer chaque caractéristique que de manière isolée, sans tenir compte des interactions potentielles entre elles qui pourraient être prédictives de façon significative.

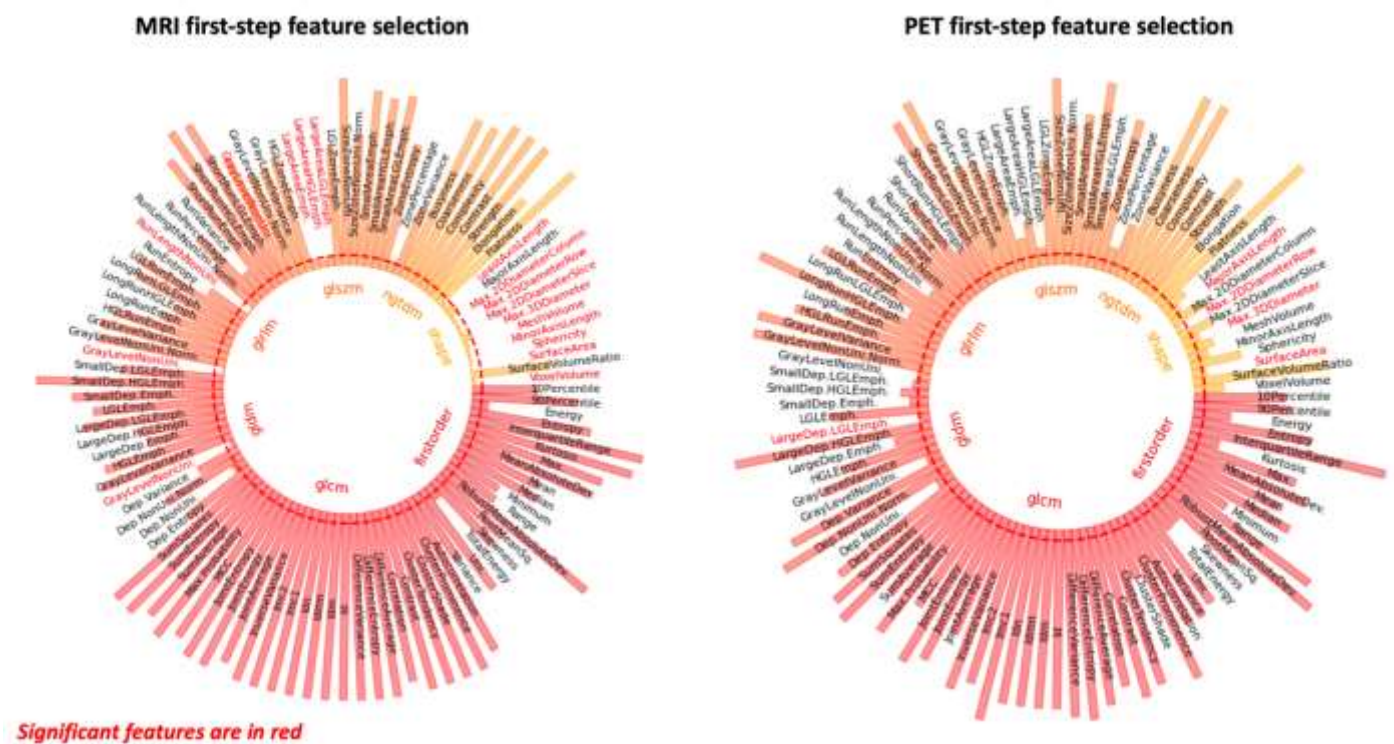


FIGURE 5.6 Exemple de sélection de paramètres radiomiques par méthode univariée

Le cercle rouge correspond au seuil de sélection pour les features associées à la réponse de fin de traitement (test de Wilcoxon, avec $p < 0.05$)

- L'approche méthodologique pour la sélection des caractéristiques peut être divisée en deux catégories. La première, connue sous le nom d'approche *wrapper*, intègre la sélection des caractéristiques dans le modèle d'apprentissage automatique, avec l'objectif de conserver un nombre réduit de paramètres spécifiquement adaptés pour optimiser les performances du classificateur [200]. Le LASSO, proposé par Tibshirani en 1996, est un algorithme de régression particulièrement célèbre dans ce cadre, qui a la particularité d'attribuer des coefficients nuls à certaines caractéristiques, les excluant ainsi du modèle final [201]. L'Elastic Net, une extension du LASSO, combine ses pénalités de régularisation avec celles de la régression Ridge, ce qui est bénéfique en

présence d'attributs corrélés, car il peut inclure ou exclure ces groupes de manière cohérente, renforçant ainsi la robustesse et la généralisabilité du modèle [202]. La deuxième catégorie, l'approche *par filtre*, évalue la robustesse de chaque caractéristique indépendamment de son utilité dans la classification [203]. Cette méthode est née de la constatation que certains paramètres radiomiques ne sont pas stables et montrent des variations significatives, même dans des conditions d'acquisition identiques ou entre différentes caméras (TEP-CT et TEP-IRM) [195,204]. Pour être cliniquement utiles et généralisables, les caractéristiques doivent être reproductibles et répétables. En 2019, Pfaehler et al. ont suggéré une méthode en trois étapes pour sélectionner ces caractéristiques, en particulier pour les paramètres de second ordre qui sont plus sensibles au bruit [172]:

1. Randomisation des voxels : cette étape consiste à mélanger aléatoirement les valeurs des voxels dans la VOI, puis à recalculer les caractéristiques radiomiques pour ces nouvelles distributions (Figure 5.7). Les caractéristiques de la distribution originale des voxels sont ensuite comparées à celles issues des distributions aléatoires. Si les valeurs diffèrent significativement, les caractéristiques sont considérées comme représentatives de la texture réelle de l'image. Une caractéristique est considérée comme aléatoire si sa valeur se situe dans l'intervalle de confiance à 95 %, basé sur les distributions aléatoires pour chaque tumeur individuelle. Pour être incluse dans les analyses suivantes, une caractéristique doit afficher une valeur hors de cet intervalle pour 90 % des patients.

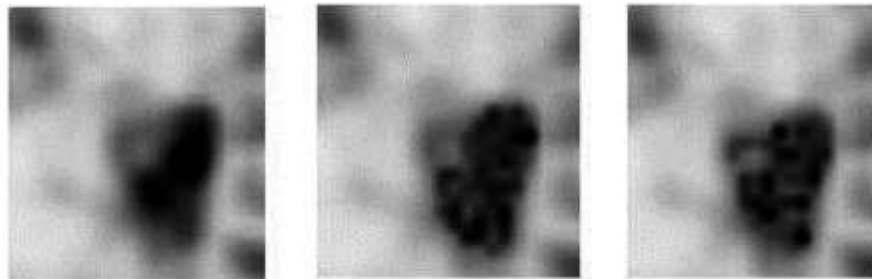


FIGURE 5.7 Exemples de mélange aléatoire des voxels au sein d'une VOI

D'après Pfaehler et al. [172] Tumeur originale (à gauche) et deux exemples (au centre et à droite) de la même tumeur après avoir mélangé aléatoirement les valeurs d'intensité dans la VOI.

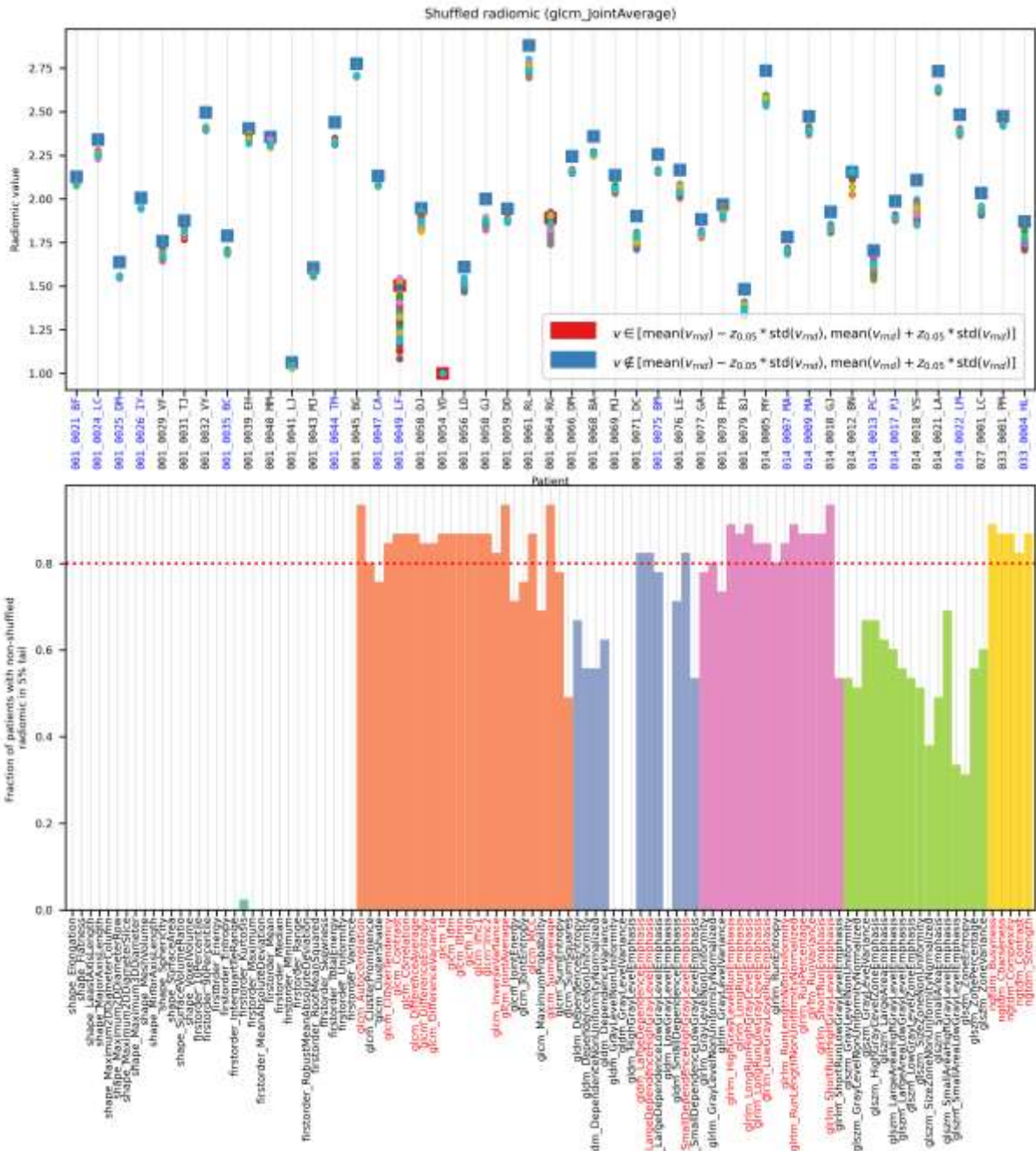


FIGURE 5.8 Application de la méthode du mélange des voxels pour sélectionner les paramètres radiomiques robustes en TEP

Le premier graphique représente la distribution d'une radiomique donnée (ici g lcm_JointAverage) dans l'ensemble de la population de training. Le carré bleu représente la valeur de cette radiomique en TEP et les points de couleurs représentent les valeurs de cette radiomique issues des 50 images avec mélanges aléatoires des voxels. Un carré rouge correspond au fait que chez un patient donné, la valeur de la radiomique initiale est comprise dans l'intervalle de confiance à 95% de la distribution des valeurs des 50 images mélangées aléatoirement. Pour être considérée robuste, une radiomique doit avoir un carré bleu chez plus de 80% des patients de la cohorte de training. Le graphique du bas représente

l'ensemble des radiomiques de second et troisième ordre étudiés, avec en rouge, les paramètres radiomiques retenus selon ce seuil de 80%.

2. Corrélation avec les métriques conventionnelles : comme nous l'avons vu plus haut, une évaluation de la corrélation entre les paramètres radiomiques et les métriques PET semi-quantitatives standards est nécessaire afin de déterminer leur réelle valeur ajoutée.

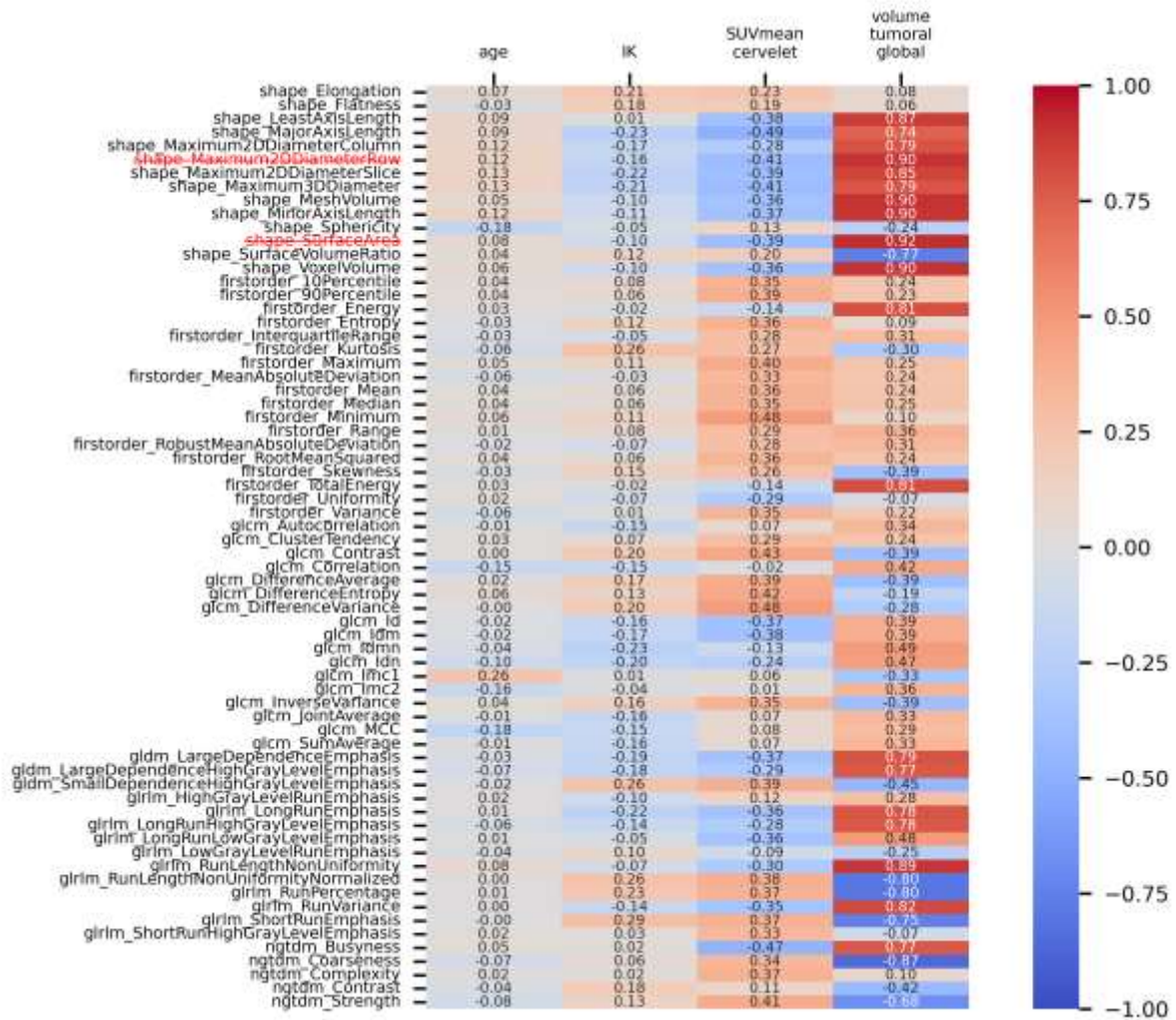


FIGURE 5.9 Corrélogramme entre les radiomiques retenus comme robustes et les marqueurs cliniques et semi-quantitatifs

Pour notre étude, nous avons étudié la corrélation entre les radiomiques (forme, premier ordre et radiomiques robustes de second et troisième ordres) et les données pronostiques cliniques connus, à savoir l'âge et le KPS ; ainsi qu'avec deux biomarqueurs TEP semi-quantitatifs mis en évidence dans notre étude précédente (SUVmean du cervelet et le sumMTV représentant la charge tumorale totale intracérébrale). Nous avons éliminé les paramètres ayant un ρ de Spearman > 0.9.

3. Corrélation mutuelle entre paramètres radiomiques : enfin, il est important que les paramètres sélectionnés reflètent des aspects différents de la tumeur et ne soient donc pas fortement corrélés entre eux. Par conséquent, si deux paramètres ou plus présentent une forte corrélation, celui qui montre la plus faible corrélation avec les métriques conventionnelles est conservé pour la constitution du modèle.

Modélisation statistique

La modélisation s'attache à dériver, à partir d'un ensemble de données d'apprentissage, une fonction de classification visant à minimiser l'erreur de prédiction. Cette fonction repose sur des hyperparamètres, qui sont affinés afin d'optimiser les performances du modèle. Les hyperparamètres incluent, par exemple, le nombre d'itérations, l'algorithme d'optimisation ou le taux d'apprentissage (« learning rate »). Pour la sélection des hyperparamètres, la population d'apprentissage est souvent divisée en sous-ensembles d'entraînement et de validation. La cross-validation génère ces sous-ensembles de validation à plusieurs reprises (les « folds »), chaque fold représentant une portion différente de l'ensemble d'entraînement. Les performances du modèle sont ainsi évaluées de manière itérative, chaque fold servant alternativement à l'entraînement et à la validation, avec ajustement progressif des hyperparamètres.

À l'heure actuelle, de multiples algorithmes de machine learning sont accessibles sans qu'un consensus ne se dégage sur la supériorité de l'un par rapport à l'autre. Il est ainsi conseillé de tester plusieurs algorithmes pour apprécier leurs performances sur un jeu de données spécifique. Notre étude a impliqué l'utilisation de six modèles distincts pour l'évaluation et la comparaison de leurs capacités prédictives.

1. **L'analyse discriminante linéaire (LDA – linear discriminant analysis)** vise à déterminer une combinaison linéaire de variables qui sépare au mieux les différentes classes entre elles.
2. **L'analyse en composantes principales (PCA – principal component analysis)** réorganise les variables originales en un nouvel ensemble de variables indépendantes, les composantes principales. Ces dernières définissent les axes selon lesquels les données présentent la plus grande variance : la première composante principale représente la plus grande variance des données, la seconde la majeure partie de la variance restante, tout en étant orthogonale à la première, et ainsi de suite. Cette

transformation des données en un espace de dimensionnalité réduite facilite leur interprétation, tout en préservant l'essentiel de l'information qui distingue les lésions.

3. **La forêt aléatoire (RF – random forest)** [205,206] crée un modèle prédictif en combinant de multiples arbres de décision. Chaque arbre apporte son vote pour la classification finale, qui se base sur la majorité des votes, renforçant ainsi la fiabilité de la prédiction.
4. **Les machines à vecteurs de support (SVM – support vector machine)** [207] cherchent l'hyperplan qui sépare le mieux les classes dans l'espace des caractéristiques. Pour gérer les relations complexes entre les variables, les SVM projettent les données dans un espace de plus grande dimension où une séparation linéaire est possible, notamment à l'aide de fonctions noyaux telles que les noyaux polynomiaux ou gaussiens, rendant les SVM particulièrement adaptées à l'analyse de données de grande dimension.
5. **Le perceptron multicouche (MLP – multilayer perceptron)** est un réseau de neurones de petite taille comprenant une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. Chaque neurone dans les couches cachées possède un poids w , un biais b et une fonction d'activation σ , qui introduit de la non-linéarité dans les relations inter-neuronales, et est connecté à tous les neurones des couches adjacentes. Dans un MLP, les données circulent dans une seule direction, de l'entrée vers la sortie. L'apprentissage des poids se fait via un processus de rétropropagation, minimisant progressivement l'erreur par ajustement des poids selon leur influence sur cette dernière (Figure 5.10).

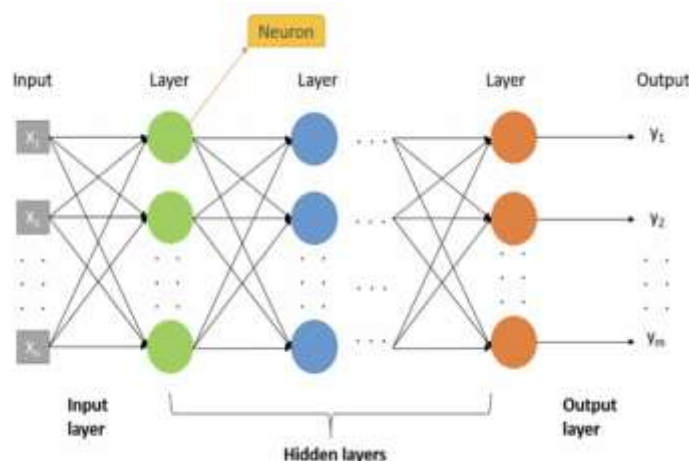


FIGURE 5.10 Exemple d'architecture d'un MLP

6. **AdaBoost** [208] est une technique d'ensemble de boosting qui se distingue des méthodes telles que Random Forest en cela qu'elle construit un modèle fort à partir de plusieurs modèles faibles (weak learners) de manière séquentielle. Au contraire du random forest qui utilise des arbres de décisions indépendants, l'Adaboost utilise des algorithmes de faibles performances non indépendants, ou chaque algorithme est entraîné pour corriger les erreurs de l'algorithme précédent.

Evaluation des performances

Pour garantir une évaluation précise d'un modèle prédictif, il est important de le tester sur un ensemble de données de test indépendant et non utilisé durant l'entraînement. Idéalement, cet ensemble devrait provenir d'une source externe, par exemple un autre établissement ou une autre période de collecte de données, afin de tester la généralisabilité du modèle. Si cela n'est pas possible, on peut utiliser une partie de l'ensemble de données initial, séparée avant toute analyse en respectant un ratio de séparation de 70/30 ou 80/20 % des patients [209]. Dans ce cas, la séparation des données doit être aléatoire mais reproductible, ce qui peut être fait en fixant une "seed" (graine) pour la génération de nombres aléatoires. Il faut aussi veiller à ne pas déséquilibrer la proportion de la variable cible dans le groupe de test, en utilisant une stratification par exemple. De plus, la séparation doit se faire par patient et non pas coupes d'images pour éviter de disperser les données d'un même patient à la fois dans la population d'entraînement et de test.

L'utilisation de la validation croisée imbriquée (nested cross-validation) permet de tester la fiabilité des résultats du modèle en les évaluant plusieurs fois avec des groupes de test variables (seed différente). Cette technique aide à prévenir une performance dépendante d'une division particulière des données, un écueil fréquent dans les études publiées en radiomique.

Néanmoins, cette méthode comporte des pièges, tels que le risque de "leakage" ou fuite d'informations si les caractéristiques choisies pour le modèle sont issues du même ensemble de données que celui utilisé pour le test. De même, si l'on réévalue et sélectionne les caractéristiques à chaque nouvelle division des données, il se peut que le modèle final perde de sa capacité à révéler les vrais indicateurs radiomiques, car ils sont traités comme des variables ajustables du modèle. Cela peut obscurcir la signification biologique des résultats.

Dans notre étude, les performances des modèles ont été évaluées à partir de courbes ROC (receiver operating characteristic). Des métriques telles que l'aire sous la courbe ROC (AUC), la précision (ACC), la sensibilité (SEN), la spécificité (SPE), la valeur prédictive positive (VPP) et la valeur prédictive négative (VPN) ont été calculées pour mesurer les performances du modèle sur la population test.

Chapitre 6. Deep learning et TEP

I. Etat de l'art IA et imagerie médicale

L'intelligence artificielle (IA), initiée par Alan Turing, englobe les systèmes d'automatisation avancée mimant les capacités intellectuelles humaines [210]. L'IA a progressé par étapes successives, et l'arrivée des ordinateurs personnels, couplée à l'innovation algorithmique, a donné naissance au machine learning (ML) dans les années 80. Avec l'avènement des unités de traitement graphique (GPU), le deep learning (DL) a pris son essor dans les années 2010, s'appuyant sur des réseaux de neurones profonds pour analyser et interpréter des données complexes [211]. À la différence du ML, les réseaux de DL peuvent opérer de manière non supervisée, c'est-à-dire sans annotations (*labels*) préalables, extrayant de façon autonome les caractéristiques pertinentes directement des images.

En radiologie et médecine nucléaire, l'application du DL s'est étendue rapidement, avec des implications significatives dans l'amélioration des flux de travail, des processus de quantification ou encore de la qualité des images, par des techniques avancées de réduction du bruit ou de correction de mouvements [212,213]. En neuro-oncologie, bien que les approches radiomiques restent prédominantes, l'orientation vers le DL gagne du terrain [151]. Chelliah et al. ont par exemple mis au point un modèle exploitant les images IRM T1 et T2 post-radiothérapie chez des patients atteints de GBM afin de déterminer avec précision et fiabilité la survie des patients [125]. Dans le contexte du LPSNC, l'étude de She et al. est la seule à l'heure actuelle à s'être intéressée à la prédiction de SG à partir d'une méthodologie de DL basée sur l'IRM [168]. Dans cette étude, les chercheurs ont amélioré leur algorithme de DL en le pré-entraînant sur une base de données publique et renommée, BraTS2020, qui compile des images de GBM. Cette première étape de pré-entraînement a permis d'adapter plus finement l'algorithme aux données de leur propre cohorte composée de 56 patients atteints de LPSNC. Les performances obtenues se sont révélées prometteuses, avec une aire sous la courbe ROC de 0.81 (± 0.03) et une précision de 0.87 (± 0.07). Il convient de souligner que ces résultats proviennent d'une validation croisée interne plutôt que d'une évaluation sur une population de test indépendante.

Contrairement au machine learning et à la radiomique, il existe un déficit de directives pour l'emploi des méthodologies en IA. Des initiatives comme CLAIM (checklist for artificial intelligence in medical imaging), Minimum Information - CLAIM et FUTURE-AI ont été lancées afin de soutenir la rigueur scientifique et l'intégrité éthique dans les études d'IA appliquées à l'imagerie médicale [214–216]. CLAIM insiste sur l'importance de décrire minutieusement la méthodologie de modélisation, les cohortes utilisées et la gestion des données manquantes [214]. MI-CLAIM offre une liste de vérifications aidant auteurs et évaluateurs à assurer la rigueur d'une étude en IA, facilitant ainsi une réplication rapide des recherches médicales [215]. Enfin, le consortium FUTURE-AI, rassemblant 118 experts internationaux de 51 pays, cherche à établir des directives pour une IA fiable en santé, s'appuyant sur six lignes directrices : l'équité, l'universalité, la traçabilité, l'utilisabilité, la robustesse et l'explicabilité.

II. Workflow en deep learning

Dans le domaine du deep learning, les réseaux de neurones convolutifs (CNN) ont été introduits en 1998 par LeCun et al. et ont longtemps été privilégiés pour le traitement d'images [217]. Les CNN utilisent des filtres convolutifs pour extraire les caractéristiques locales des images, permettant une analyse efficace pour des tâches telles que la reconnaissance d'objets. Plus récemment, les avancées en IA se sont orientées vers des architectures avancées telles que les réseaux U-Net et les Vision Transformers (ViT). Le U-Net, un sous-type de CNN, est particulièrement efficace pour la segmentation d'images et le débruitage, grâce à leur structure permettant de traiter et de conserver des informations spatiales à différentes échelles [218–220]. Les Vision Transformers se distinguent par leur mécanisme d'auto-attention, essentiel à l'apprentissage de motifs complexes et à la compréhension des relations spatiales à l'échelle de l'image [221]. Cette fonctionnalité augmente considérablement leur efficacité, car elle permet aux modèles de se concentrer sur les zones les plus informatives pour une tâche spécifique.

Augmentation des données

En DL, l'accès à un volume conséquent de données de qualité et annotées est recommandé pour développer un algorithme performant. Cependant, cela s'avère particulièrement ardu pour des pathologies rares, telles que le LPSNC. Les techniques d'augmentation artificielle de données peuvent pallier, au moins partiellement, à cette limitation. Elles enrichissent le jeu de données

d'entraînement par l'intégration de copies modifiées des données existantes, via des manipulations telles que des rotations, des retournements ou des translations. Une méthode complémentaire consiste à diviser les images en patches, c'est-à-dire en portions d'une image 3D réduites en tranches 2D, possiblement avec chevauchement des portions entre elles. Chaque patch est classifié individuellement, puis les résultats sont agrégés au niveau du patient pour évaluer les performances du modèle. Des études récentes, notamment sur la maladie d'Alzheimer, ont mis en évidence les bénéfices de l'utilisation de patches pour l'amélioration des performances par rapport à une analyse globale au niveau du patient [222].

Auto-encodeur

L'auto-encodeur est un modèle d'apprentissage entraîné de manière non supervisée dont le but est de réduire la dimensionnalité de l'image avant la tâche de classification. Il est composé de trois principaux éléments : un encodeur, un espace latent ou code et un décodeur. Cette architecture est conçue pour encoder les données en représentation réduite dans l'espace latent, qui doit préserver les aspects les plus importants des données originales de sorte à reconstruire l'*output* aussi fidèlement que possible à l'*input*. L'intérêt d'un auto-encodeur est de procéder à la tâche de classification à partir de l'espace latent, qui représente les données les plus pertinentes de l'image initiale (Figure 6.1).

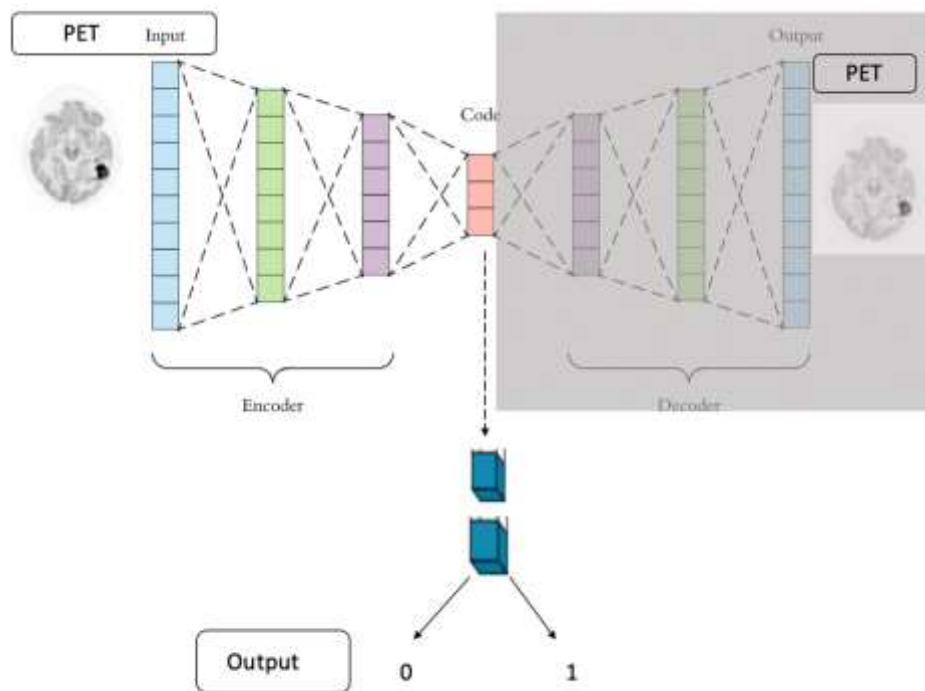


FIGURE 6.1 Schéma d'architecture d'un auto-encodeur

Comme dans un réseau neuronal classique, les poids du modèle sont d'abord assignés aléatoirement puis affinés par rétropropagation en fonction des performances du dernier nœud de l'algorithme dans la tâche de classification étudiée. Pour une tâche de classification binaire en 0 ou 1, on utilisera une fonction de perte de cross-entropy binaire.

Nous détaillerons ici les paramètres impliqués dans un auto-encodeur :

- le nombre de couches de convolution : la convolution constitue la pierre angulaire des réseaux de neurones convolutifs. Cette technique mathématique est utilisée pour traiter les signaux et les images en appliquant un filtre, ou noyau, qui parcourt l'image et effectue une opération de multiplication et de sommation locale des valeurs de pixels. Les couches initiales d'un CNN tendent à détecter des caractéristiques simples, comme les bords et les textures, tandis que les couches plus profondes sont capables de reconnaître des aspects plus abstraits et complexes.

En pratique, chaque couche de convolution est suivie d'une fonction d'activation. En DL la fonction Rectified Linear Unit (ReLU) est la fonction d'activation la plus couramment utilisée (Figure 6.2).

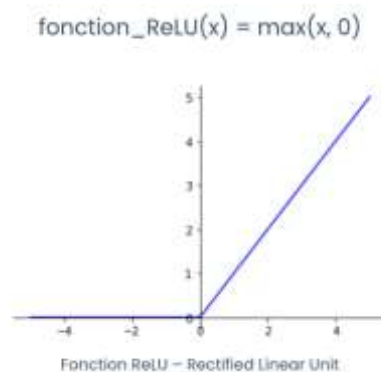


FIGURE 6.2 Représentation mathématique de la fonction ReLU

Elle permet d'appliquer un filtre en sortie de couche, permettant d'ignorer les aspects non pertinents de l'image. Pour autant, un excès de convolutions dans un modèle peut entraîner un surajustement. Cela se manifeste, comme nous l'avons déjà souligné, par des performances élevées sur les données d'entraînement mais une faible capacité de généralisation sur de nouvelles données. Pour contrer l'overfitting, plusieurs stratégies peuvent être adoptées :

- **La régularisation** : L1 (LASSO) ou L2 (Ridge) vise à atténuer l'influence des poids les plus importants pour favoriser un modèle plus équilibré.

- **La normalisation par lots** (batch normalization) : Elle applique une normalisation après chaque couche plutôt que sur l'ensemble du jeu de données, pour limiter la variance intra-données et favoriser la stabilité du modèle.

- **Le dropout** : Il consiste à désactiver aléatoirement des neurones, ce qui force le modèle à ne pas trop dépendre de certaines caractéristiques et à mieux généraliser.

- **L'arrêt prématuré** (early stopping) : Cette technique de régularisation consiste à arrêter l'entraînement dès que le loss de validation atteint un plateau ou commence à augmenter (Figure 6.3). Le loss correspond à la mesure de l'erreur qu'un algorithme d'apprentissage automatique cherche à minimiser au cours de l'entraînement.

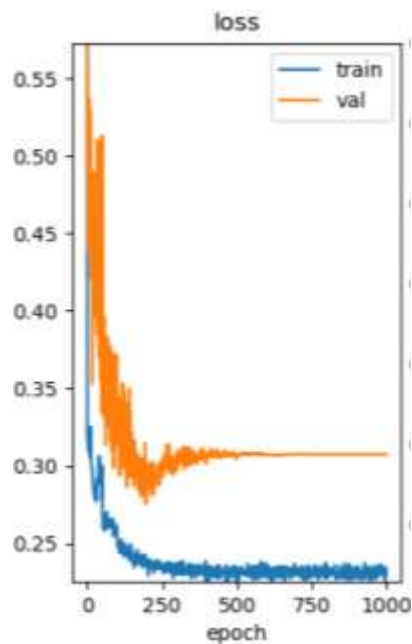


FIGURE 6.3 Exemple de courbes de loss

Exemple de courbes de loss sur la population de training et de validation à mesure des epochs.

Les autres paramètres de l'auto-encodeur sont :

- **Maxpool** : Après chaque convolution, la couche de maxpooling agit comme un mécanisme de sous-échantillonnage. Elle réduit la dimensionnalité spatiale des cartes de caractéristiques (feature maps), permettant ainsi d'exploiter des corrélations spatiales plus lointaines. En sélectionnant la valeur maximale d'une fenêtre spécifique sur la carte de caractéristiques, le max pooling conserve les caractéristiques prédominantes.

- **Flatten** : La couche flatten convertit les cartes de caractéristiques multidimensionnelles en un vecteur unidimensionnel. Cette transformation est essentielle pour passer des représentations spatiales traitées par le CNN à une forme compatible avec les couches denses (fully connected layers), qui modélisent les relations linéaires et non linéaires pour exécuter des tâches telles que la classification ou la régression (Figure 6.4).

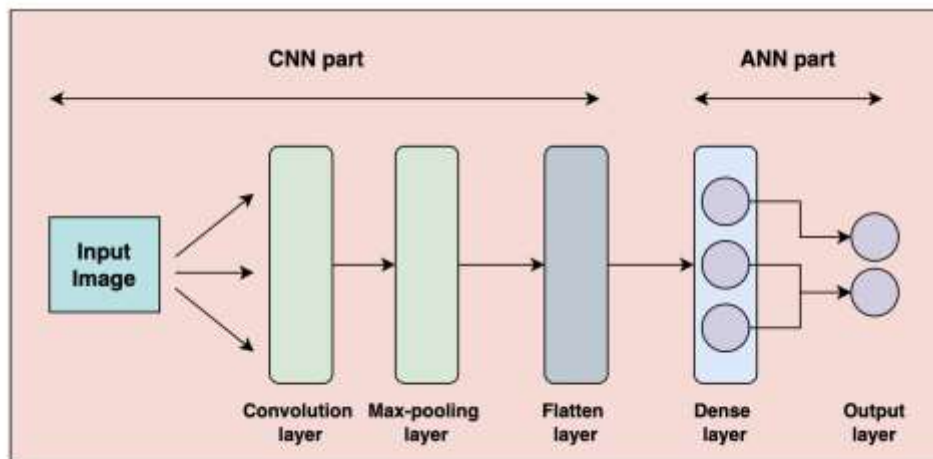


FIGURE 6.4 Exemple d'architecture d'un réseau de neurones convolutifs

La figure montre l'enchaînement des couches de convolution et de maxpooling. La couche de flatten permet la transition entre le CNN et l'ANN (artificial neural network) qui comprend la couche dense, connectée à la couche de sortie.

- **Couches denses** : Les couches denses, également connues sous le nom de couches entièrement connectées, sont des éléments fondamentaux dans un CNN. Dans un CNN les couches denses sont les couches prédictives du modèle, elles se trouvent à la fin du réseau, prennent en entrée la représentation vectorielle de l'image et tentent de résoudre le problème associé à celle-ci (par exemple une classification) .

- **Learning rate** : le learning rate est un hyperparamètre qui détermine la mise à jour des poids du modèle pendant l'entraînement. Une valeur trop grande du learning rate peut amener le modèle à diverger, une valeur trop petite peut allonger le temps d'entraînement inutilement. Le learning rate peut être fixe ou ajusté au fil du temps (LR Schedule) pour affiner l'apprentissage à mesure que le modèle progresse.

- **Epochs** : Une epoch représente une itération complète de l'ensemble d'entraînement à travers le réseau de neurones pendant le processus d'apprentissage. Plus le nombre d'epochs est élevé, plus le modèle a de chances d'apprendre à partir des données. Cependant, un nombre trop élevé d'epochs peut conduire à de l'overfitting, tandis qu'un nombre insuffisant pourrait ne pas permettre au modèle d'apprendre suffisamment de l'ensemble des données.

Segmenteur

En imagerie médicale, l'utilisation du DL à des fins de segmentation est une branche majeure de l'intelligence artificielle. En effet, que ce soit à des fins thérapeutiques pour faciliter la radiothérapie ou quantitatifs pour mesurer la charge tumorale rapidement, les algorithmes de segmentation constituent un vaste champs d'investigation [220,223,224]. Dans le cadre de notre travail, nous avons étudié un algorithme de segmentation comme première étape de l'algorithme de classification pour orienter le réseau vers la région d'intérêt à explorer (ici la tumeur en TEP ou en IRM T1 avec injection de produit de contraste). Ainsi, contrairement à l'auto-encodeur qui ne nécessite en données d'entrée que l'image brute, le segmenter nécessite également un contour de la tumeur, appelé masque, pour toutes les données d'entraînement (Figure 6.5).

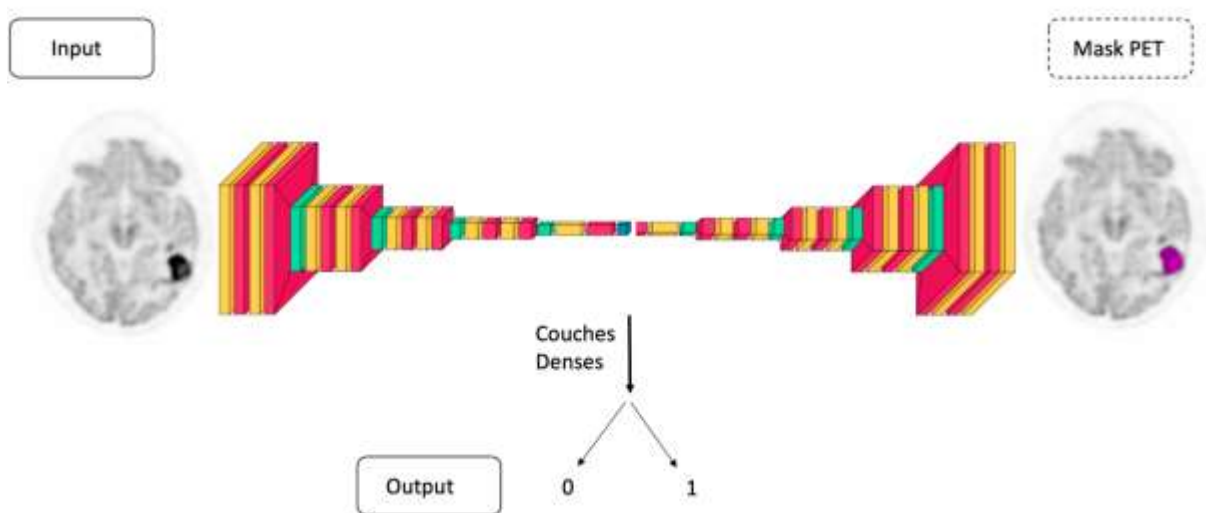


FIGURE 6.5 Schéma d'architecture d'un segmenteur

L'utilisation de réseaux segmenteurs présente l'avantage de focaliser l'algorithme sur la pathologie concernée, mais avec le risque d'occulter des informations pertinentes du reste de l'image. Cet aspect est particulièrement important en oncologie, où l'impact du microenvironnement tumoral sur la réponse aux traitements est de plus en plus reconnu comme étant significatif [225]. Dans notre première étude, nous avons notamment mis en évidence l'impact de l'activité métabolique du cervelet sain sur la réponse au traitement à la fin d'une chimiothérapie d'induction.

SWIN transformer

Le dernier type de réseau que nous avons étudié est un SWIN transformer (Figure 6.6). Sa méthodologie, décrite par Liu et al. en 2021 est une transposition des principes des transformateurs développés pour le traitement du langage, tels que ceux utilisés dans les modèles pré-entraînés génératifs (GPT), à la vision par ordinateur [221]. À l'instar des modèles GPT-langage qui emploient des jetons de mots ou *token* comme éléments de base, le SWIN Transformer utilise des fragments d'image de petite taille fixe, appelés patches. Contrairement aux CNN traditionnels, le SWIN Transformer permet d'analyser les images à grande échelle, en tenant compte du contexte global. Contrairement aux CNNs qui opèrent à une échelle fixe, le Swin Transformer divise l'image en fenêtres, qui sont ensuite subdivisées en patches. Il calcule les relations entre les patches à l'intérieur de chaque fenêtre, puis établit des connexions entre les couches via des fenêtres chevauchantes, intégrant progressivement tous les patches dans le processus. Cela permet une meilleure représentation des caractéristiques, améliorant la performance des tâches visuelles sans la complexité computationnelle croissante des CNNs.

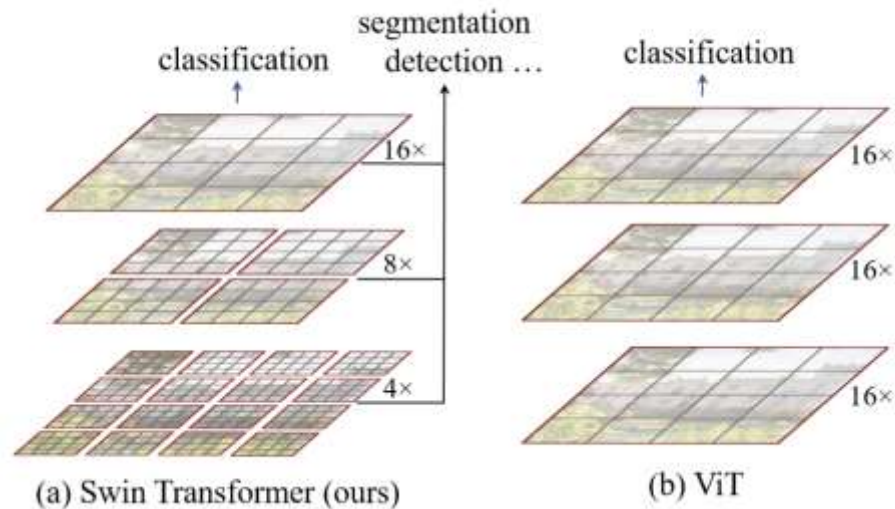


FIGURE 6.6 Représentation simplifiée des fenêtres de vision dans un Swin transformer (a) et dans un vision transformer (ViT) (b)

D'après Liu et al. 2021[221]

Optuna

La sélection des hyperparamètres est essentielle pour optimiser les performances du modèle mais peut être une procédure longue et ardue. Traditionnellement, cette sélection se fait de manière aléatoire ou via une "grid search", où chaque combinaison possible est évaluée. Toutefois, ces approches peuvent être inefficaces et consommer beaucoup de temps. Optuna est une solution logicielle, offerte sous Python, qui offre une alternative avancée avec son algorithme d'optimisation bayésienne [226]. Elle permet de naviguer de manière plus intelligente dans l'espace des hyperparamètres, et son interface en ligne permet le suivi en temps réel ainsi qu'une interprétabilité complète des étapes d'optimisation.

Grad-CAM

L'application du DL en imagerie médicale, tout en offrant des avancées substantielles, soulève des questions méthodologiques et éthiques importantes. Les recommandations récentes en matière de ML et DL mettent l'accent sur le besoin d'algorithmes explicables, une démarche connue sous le nom de XAI (eXplainable AI) [173,214,216]. L'objectif du XAI est de fournir une transparence quant aux décisions algorithmiques et d'identifier leurs éventuelles failles. Reyes et al. en 2020 ont notamment discuté de ces enjeux dans le contexte de la radiologie, suggérant l'utilisation de cartes de visualisation pour interpréter les décisions des modèles de DL [227].

Dans le cadre de cette thèse, nous avons utilisé le gradient-weighted class activation mapping (Grad-CAM), une technique de XAI qui permet de visualiser l'importance attribuée à chaque pixel dans la prédiction d'un modèle [228,229]08/07/2024 12:17:00. L'idée fondamentale du Grad-CAM est de produire des cartes de chaleur qui mettent en évidence les régions les plus importantes pour la prédiction d'une classe spécifique. Au lieu de calculer le gradient par rapport à la couche d'entrée, le Grad-CAM utilise les gradients de la sortie par rapport à une des couches profondes pour déterminer l'importance des neurones dans cette couche (Figure 6.7). Ces gradients sont ensuite combinés avec les activations des neurones pour produire une carte de chaleur localisée indiquant les régions significatives pour la classification. Cette approche favorise une meilleure interprétation et confiance dans les décisions prises par les modèles complexes, ce qui est particulièrement important dans les applications médicales où les enjeux sont élevés. Cette méthode offre également l'avantage d'examiner et de discuter des erreurs de classification effectuées par le réseau.

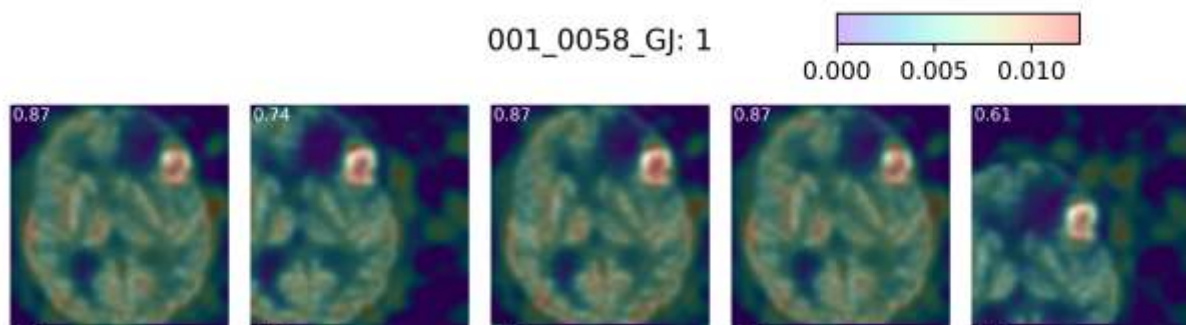


FIGURE 6.7 Exemple d'application du Grad-CAM chez un patient atteint de LPSNC

Cet exemple illustre l'utilisation de Grad-CAM pour la visualisation, où l'échelle chromatique indique l'importance de chaque pixel dans la décision de l'algorithme. Les zones en rouge marquent les pixels les plus influents sur le résultat final de la classification.

III. Article 4

Comparing radiomic machine learning and deep learning approaches for predicting treatment failure in elderly immunocompetent patients with primary central nervous system lymphoma: a PET-MRI based study.

Introduction

Primary Central Nervous System Lymphoma (PCNSL) is an uncommon yet aggressive variant of non-Hodgkin lymphoma, affecting the brain, spinal cord, meninges, or vitreoretinal space, without systemic involvement [230]. Despite representing only 4% of primary brain tumors, there has been a notable rise in PCNSL incidence, especially among immunocompetent and elderly patients [1,28,231]. The typical treatment for PCNSL includes induction with high-dose MTX-based chemotherapy, followed by consolidation, which relies on autologous stem cell transplantation in younger patients. However, this approach is subject to ongoing debate in older patients due to the potential risk of brain toxicities [84].

Despite the advent of new therapeutic regimens, the prognosis of PCNSL remains grim, with 15-25% of patients exhibiting no response to chemotherapy and 25-50% relapsing post-initial therapy [28]. The identification of patients who could potentially benefit from consolidation therapy is a pressing unmet need in the management of elderly PCNSL patients. Currently, age and performance status are the only clinical biomarkers established as prognostic factors, yet they are insufficient for enabling a robust, tailored treatment approach [27,28,86].

Magnetic resonance imaging (MRI) is the gold standard for PCNSL patient evaluation, but few MRI parameters have demonstrated efficacy in patient stratification [104,232–234]. Some studies have underscored the prognostic role of tumor location, with deep or infratentorial localizations associated with poorer overall survival [27,101]. Others have suggested the potential of Apparent Diffusion Coefficient (ADC) as a prognostic marker, with high diffusion restriction correlating with poor outcomes [103,234–238]. The role of perfusion in PCNSL prognosis remains unclear due to contradictory results [99,104,232,233,239,240].

Positron emission tomography (PET), an in-vivo molecular technique, has also been explored as a prognostic biomarker for PCNSL [9,13,16,241]. Our recent work demonstrated that the SUV_{mean} of the normal cerebellum and the total volume measured on PET are potential

biomarkers of poor response following the first line of therapy [242]. However, these findings require further validation on external cohorts.

Radiomics, an emerging field that extracts numerous quantitative features from medical images, provides insights into lesion properties such as size, shape, and heterogeneity [151,243]. In neuro-oncology, it has shown promise in tumor characterization, especially in predicting molecular characteristics in gliomas [160,162]. In PCNSL, radiomics has been primarily studied for differential diagnosis between PCNSL, GBM, and Mets [244–246]. Few studies have investigated the role of radiomics in outcome prediction for PCNSL, and to the best of our knowledge, none have focused on PET [169,247]. Destito et al. analyzed 80 de novo PCNSL patients and demonstrated that radiomic features extracted from T1 weighted sequences predict overall survival (OS) and progression-free survival (PFS) better than clinical features alone when using machine learning (ML) models [169]. Chen et al., using non-machine learning methods, showed in a retrospective study of 52 patients that texture features from contrast-enhanced MRI could serve as prognostic biomarkers for PCNSL patients [247].

Conversely, deep learning (DL), a branch of artificial intelligence that autonomously identifies features from raw images to directly model clinical outcomes, is gaining significant traction in medical imaging. Convolutional neural networks (CNNs) have been particularly effective in computer vision, classifying images across various domains [222]. She et al. have employed this architecture in patients with PCNSL and demonstrated accurate quantification of molecular tumor burden and robust prediction of risk, survivability, and treatment response [168]. Recently, innovative frameworks such as the SWIN transformer have been introduced, with application in medical tasks like segmentation and classification [248,249].

The aim of our study was to conduct a comparative analysis of radiomic-based ML versus DL models for predicting treatment failure in elderly immunocompetent patients with Primary Central Nervous System Lymphoma, using Positron Emission Tomography-Magnetic Resonance Imaging (PET-MRI) datasets.

Methods

Clinical Dataset

This retrospective study included immunocompetent patients aged 60 and above, diagnosed with PCNSL, confirmed by histopathology or elevated cerebrospinal fluid markers. PET scans

were conducted at the Department of Nuclear Medicine of Pitié-Salpêtrière Hospital, 90 minutes post-injection of 2MBq/kg of ¹⁸F-FDG. Patients were required to fast for at least 4 hours prior to the scan, and blood glucose levels were monitored before the procedure. Following the baseline PET, patients underwent first-line chemotherapy based on high-dose methotrexate treatment. The training set consisted of patients scanned on the TOF PET-MRI system (Signa PET/MR 3T General Electric Healthcare) between 2016 and 2021, who were enrolled in the prospective multicentric french Localyze study (NCT02313389)[242]. The OSEM-3D reconstruction involved 6 iterations, 28 subsets, and a 6.4 mm transaxial post-reconstruction Gaussian filter. Four-tissue-class (soft tissue, fat, lung, and air) attenuation correction maps were calculated from the 2-point Dixon sequences, supplemented by a single atlas to capture bone information. The external cohort comprised consecutive patients, diagnosed between 2019 and 2023 imaged on a different camera (mCT Flow Siemens) at the same center, and not included in the Localyze study. For this camera, the OSEM-3D images were reconstructed with 8 iterations, 21 subsets, and no post-reconstruction filter. PET images were corrected for attenuation using the acquired low-dose CT data. T1 contrast-enhanced weighted sequence of MRI performed within a month of the PET were used for these patients. PET and MRI images were semi-automatically co-registered using the ADW server software (GE Healthcare). Demographic, clinical, and qualitative imaging variables were systematically gathered. Patients from the external dataset presenting with incomplete records or insufficient follow-up data were excluded from the analysis. The primary endpoint was to predict the response at the end of the first-line therapy, assessed by a neuro-radiologist (DG) and a neuro-oncologist (CH), according to the International Primary Central Nervous System Lymphoma Collaborative Group (IPCG) criteria. The task was modeled as a binary classification problem, with the group “responders” including the complete response (CR), complete response unconfirmed (Cru) and partial response (PR) and the group “non-responders” including the stable disease (SD) and progressive disease (PD).

All participants provided informed consent for their participation in the study. The study was approved by the protection of persons committee (2014–002597-37) and the French National Agency for the Safety of Medicines and Health Products (ANSM), and abided by the principles of the Declaration of Helsinki. Data analysis were performed in collaboration with the Gordon Center for Medical Imaging, Massachusetts General Hospital and approved by the Institutional Review Board (IRB 2022P002011).

Image Pre-processing

PET images were standardized using the SUV normalization factor and rescaled to a consistent voxel size of $2 \times 2 \times 2 \text{ mm}^3$, with subsequent data normalization implemented using a fixed bin width of SUV 0.3 [192]. MRI scans underwent bias field correction utilizing the N4ITK MRI bias correction module in Python, followed by a Z-score standardization, and rescaling of the voxel size to $2 \times 2 \times 2 \text{ mm}^3$ for uniformity across imaging modalities [169,176]. Tumor delineation was meticulously carried out on axial PET and T1 contrast-enhanced MR images by a nuclear medicine specialist with seven years' expertise in PET-MRI interpretation (LR), using the LifeX software (<http://www.lifexsoft.org>) [189]. Segmentation of FDG-avid lesions was conducted with a set threshold at 41% of the SUVmax, and a semi-automatic 3D ROI was delineated for contrast-enhanced lesions, with manual refinement performed slice by slice to ensure precision.

Radiomic Study and feature selection

Prior to the radiomic analysis, two conventional PET metrics were computed: the SUVmean of the cerebellum, acquired using a spherical volume-of-interest (VOI) of 1.5 cm^3 on a cerebellar hemisphere not affected by diaschisis or tumor infiltration, and the sum of the metabolic tumor volume (sumMTV) across up to five target lesions, serving as a quantifier of patient tumor burden [242]. For the extraction of radiomic features, we used the Python package PyRadiomics. The algorithms and computational formulas, alongside the entire suite of radiomic features, are detailed on the PyRadiomics documentation (<https://pyradiomics.readthedocs.io/en/latest/features.html>) [193]. A comprehensive set of 107 features per imaging modality were extracted, encompassing 18 first-order statistical parameters, 14 shape descriptors, and 75 textural features, all adhering to the current standards set by the Image Biomarker Standardization Initiative [174].

Considering the extensive number of potential features against the backdrop of the sample size and the inter-correlation of variables, we employed a systematic feature selection process. Initially, the consistency of textural radiomic features (RFs) was assessed excluding first-order and shape features, using a method outlined by Pfaehler et al. [172]. For each RF, signal within the ROI was randomized 50 times, and the RF was recalculated. An RF was deemed unreliable if its true value fell within the 95% confidence interval (CI) generated by this random permutation. An RF was retained if its true value stood outside the 95% CI for at least 90% of the cohort. Subsequently, RFs exhibiting a strong correlation (Spearman correlation > 0.9) with either of the two established PET metrics were excluded. Lastly, features demonstrating

excessive mutual correlation (Spearman correlation > 0.7) were also removed to ensure the independence and relevance of the final feature set.

Machine learning pipeline

Six classifiers were investigated to evaluate the primary endpoint: two linear, two non-linear and two ensemble classifiers. The linear classifiers used were linear discriminant analysis (LDA) and principal component analysis (PCA). The non-linear classifiers used were a support vector machine (SVM) and a multi-layer perceptron (MLP) [206,207]. The ensemble classifiers used were random forest (RF) and boosting of classification trees with AdaBoost [208]. Each dataset was tested with the six classifiers. The external testing cohort was not involved in the process of feature selection and model training. The models were implemented with the Python package Scikit-learn. For models with tuning parameters, the best hyper-parameters were tuned using cross-validation on the training data to avoid bias.

Deep learning pipeline

In the development of our deep learning pipeline, each image matrix (256x256) was partitioned into overlapping patches. These patches were treated as distinct samples during the model training phase, while a per-patient basis was employed for the testing procedure. We used a SWIN transformer architecture for our model, leveraging its ability to capture intricate patterns through self-attention mechanisms and scalable receptive fields [221]. The optimization of the model was carried out using Optuna, a versatile framework that employs the Hyperband method, allowing for an efficient search through various combinations of hyperparameters (including patch size, dropout rate, and learning rate) [226]. To minimize overfitting and ensure the generalizability of the model, a binary cross-entropy loss function was implemented for network training. Moreover, we incorporated data augmentation techniques such as image mirroring and translation during training to enhance the model's robustness against variations in imaging data.

The performance of the models was evaluated on both the training and testing cohorts using Receiver Operating Characteristic (ROC) curve analysis. Metrics such as the area under the ROC curve (AUC), accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), and negative predictive value (NPV) were computed to gauge the diagnostic efficacy of the model when applied to the testing cohort. Subsequently, 95% confidence

intervals (CIs) were determined using a non-parametric bootstrap method, which was iterated 5000 times to generate a bootstrap distribution of the results.

Results

Patient Characteristics

The training population included patients from a previously described prospective Localize cohort [242]. Out of 54 patients who underwent baseline PET-MRI, six were excluded due to insufficient contrast uptake size (<8mm) on MRI for proper radiomic evaluation. Additionally, two patients were excluded for having T1-weighted MRI contrast-enhanced acquisitions in sagittal mode with fat saturation but not 3DT1BRAVO acquisitions, to maintain data uniformity for radiomic extraction. The testing cohort consisted of 23 patients with histologically confirmed PCNSL, aged over 60, who underwent brain PET-CT at Pitié-Salpêtrière (Siemens mCT Flow) between 2019 and September 2023. Each patient also received a contrast-enhanced T1-3D MRI within the month following the PET. For 21 of them, MRI was performed on a 1.5T machine (GE Healthcare) at Pitié-Salpêtrière, and for the remaining two, the examinations were conducted externally but retained the same imaging sequence. Figure 1 provides a flowchart detailing the selection process for both the training (n=46) and external validation (n=23) cohorts.

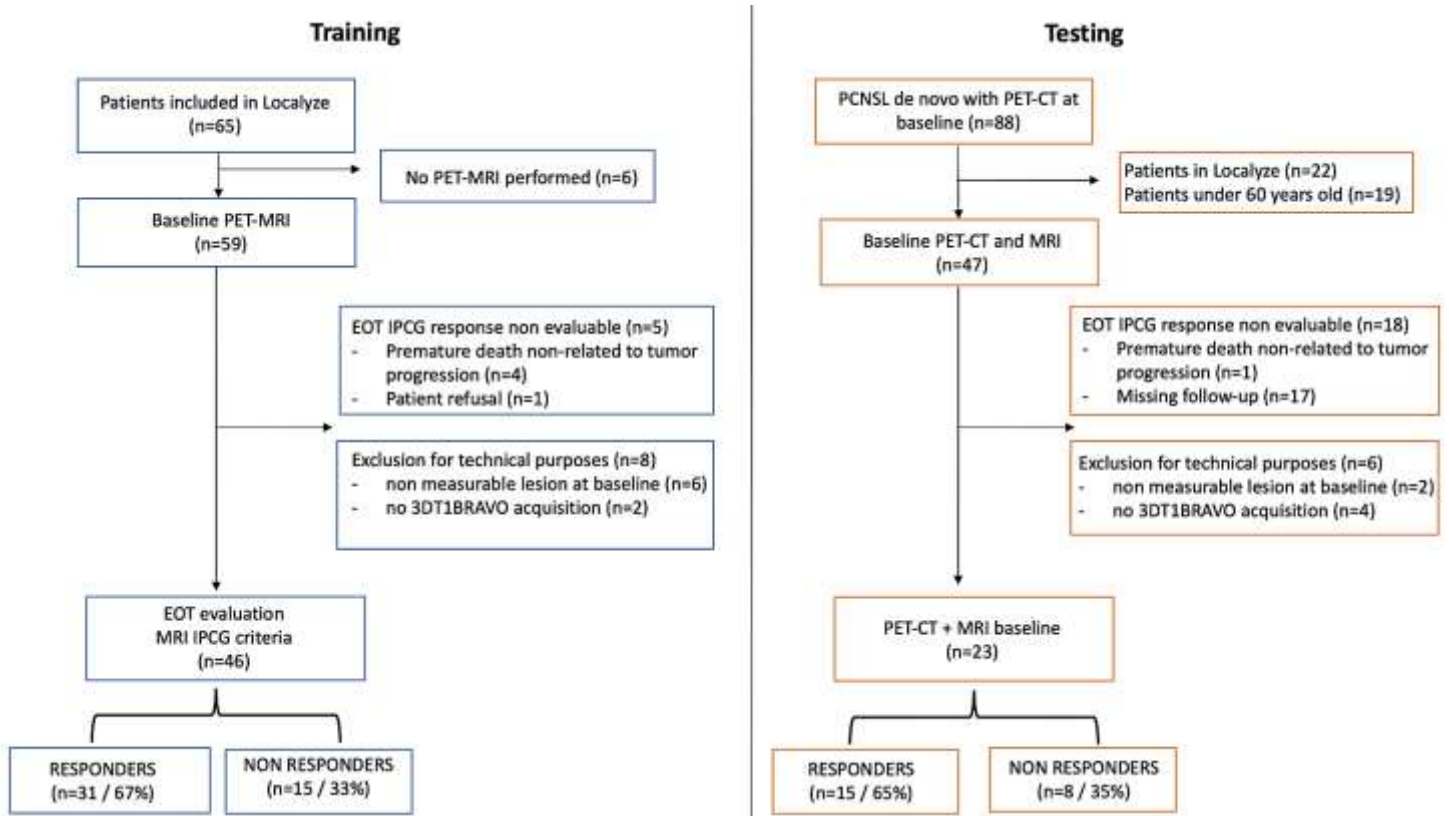


Figure 1. Flow chart of the population

EOT : End of treatment, IPCG: International PCNSL collaborative group

Key prognostic factors for primary central nervous system lymphoma (PCNSL), namely age and functional status as per the Eastern Cooperative Oncology Group (ECOG) scale, showed no significant differences between the two cohorts. Semi-quantitative PET parameters identified in our prior study also did not differ statistically across the cohorts. All patients in our study had diffuse large B-cell lymphoma upon histological examination. Participants in the Localize study uniformly received the same induction treatment per the BLOCAGE01 protocol, consisting of 4 or 5 cycles of R-MPVA chemotherapy. In the validation cohort, 17 patients underwent this identical regimen, while 4 received methotrexate monotherapy and 2 others, in addition to R-MPVA, were administered Ibrutinib and lenalidomide, respectively. Within the training set, 15 (33%) patients showed disease progression (non-responders) during the first-line therapy. In the testing set, 8 (35%) were classified as non-responders following the evaluation at the end of induction therapy. Patient demographics and lesion characteristics are outlined in Table 1.

	Training cohort (n=46)	Testing population (n=23)	p
Sex –female, n (%)	28 (61%)	9 (31%)	0.08
Age (years) Med [min-max]	71.5 [60-85]	71 [61-85]	0.67
ECOG, ≤1, n (%)	26 (57%)	12 (51%)	0.16
Initial location, n (%)			0.51
Supra-tentorial	30 (67%)	18 (78%)	
Infra-tentorial	5 (11%)	1 (4%)	
Both	10 (22%)	4 (17%)	
Superficial	8 (18%)	9 (39%)	0.06
Deep	37 (82%)	14 (61%)	
Multifocal	24 (52%)	11 (48%)	0.73
sumMTV, mean ± SD	7.4 [0.6-64.7]	4.34 [0.5-32.8]	0.07
SUVmean cerebellum Med [min-max]	6.1 [3.2-10]	5.4 [2.6-12]	0.21
SUVmax Med [min-max]	28.2 [9.9-64.2]	19.7 [8.3-63.4]	0.003

Table 1. Characteristics of the patients

ECOG : Eastern Cooperative Oncology Group

The workflow for the radiomic analysis is depicted in Figure 2.

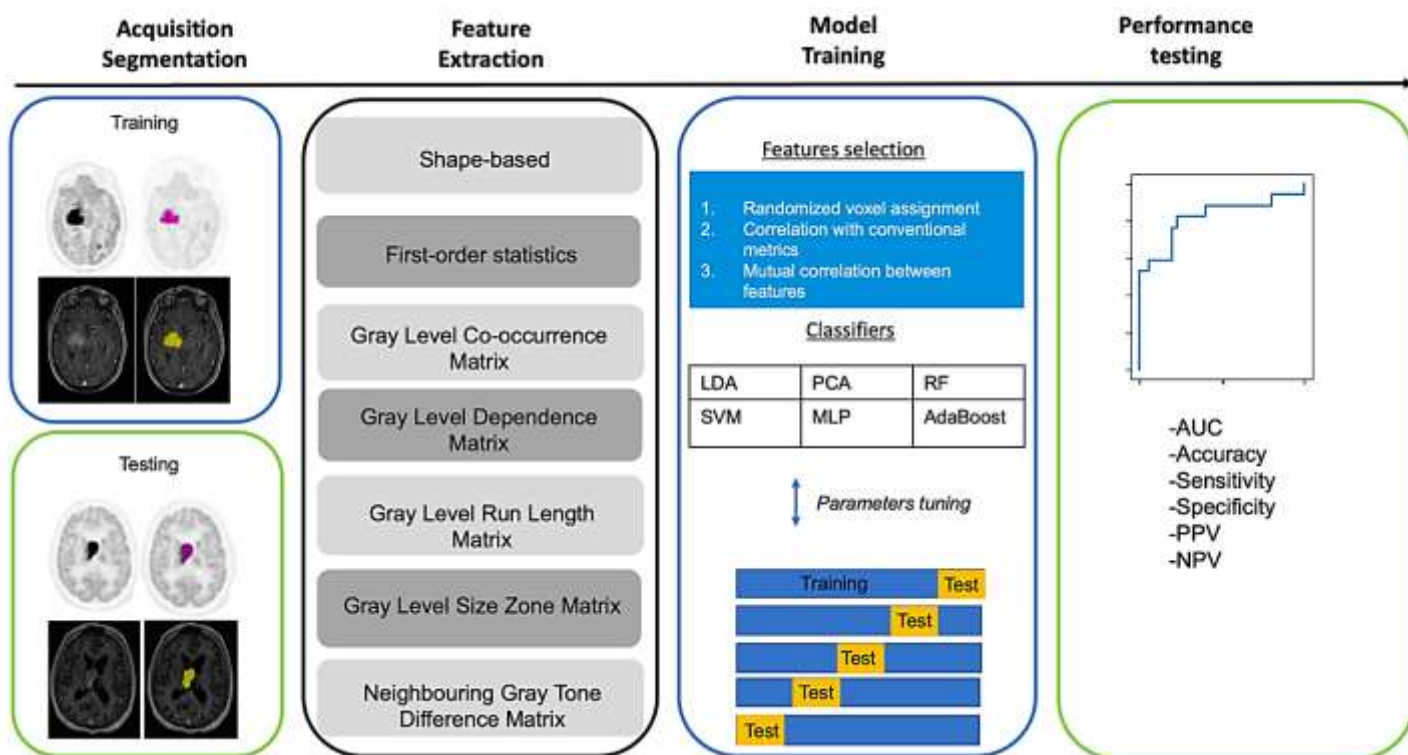


Figure 2. Workflow of the radiomic analysis

LDA : Linear discriminant analysis, PCA :Principal component analysis, RF : Random forest, SVM : Support vector machine, MLP : Multi-layer perceptron, AUC :Aare under the curve, PPV : Positive predictive value, NPV : Negative predictive value

Preliminary results from radiomics-based and deep-learning models

Our initial data analysis was based on a methodology that deviated slightly from the one previously stated. We began by combining patients from both the PET-MRI and PET-CT cohorts, later separating them into a training cohort with 80% of the patients and a testing cohort with the remaining 20% (no nested cross-validation). A stratification based on the end-of-treatment response was also performed to ensure group homogeneity. For the radiomics analysis, PET and MRI parameters were initially selected through a LASSO process, retaining the following indicators: in MRI - Sphericity, GLSZM (LargeAreaLowGrayLevelEmphasis), GLSZM (ZonePercentage), GLDM (LargeDependenceHighGrayLevelEmphasis); and in PET – shape (Maximum2DDiameterRow), GLRLM (LongRunLowGrayLevelEmphasis), and SUVmean_cerebellum. These criteria led to the performances presented below for the six ML classifiers (Table 2):

Classifiers		LDA	PCA	RF	SVM	MLP	Adaboost
Training (n=46)	ACC	0.78	0.77	0.96	0.84	0.83	1
	AUC	0.73	0.68	0.94	0.78	0.77	1
Testing (n=23)	ACC	0.85	0.81	0.76	0.76	0.85	0.76
	AUC	0.78	0.71	0.68	0.71	0.82	0.71

Table 2. Preliminary results of the radiomic analysis

LDA : Linear discriminant analysis, PCA :Principal component analysis, RF : Random forest, SVM : Support vector machine, MLP : Multi-layer perceptron

For the DL analysis, we evaluated three architectures using only PET data as input: an autoencoder, a segmenter, and a SWIN transformer, the latter showing the most promising results. An AUC of 0.76 was achieved for the training cohort and 0.80 for the testing cohort by adjusting hyperparameters using the Optuna software. The details of the hyperparameters selected for this model are recorded in the table below (Table 3):

Hyper Parameter	Value
Patch Size	[1,128,128]
Patch Overlap	[0,96,96]
Dropout Rate	0.05122812384099595
Number of Heads	8
Embedding Dimension	32
Window Size	4
Shift Size	1
Learning Rate Schedule	Reduction of the learning rate by 50% every 50 epochs
Total Epochs	200
Batch Size	32

Table 3. Hyperparameters for the PET deep-learning model

Discussion

The development of novel methodologies in AI has paved the way for significant advancements in medical tools, particularly in medical imaging. Among these, radiomics analysis through ML algorithms or DL approaches hold the promise for a deeper understanding of the physiopathological processes underlying tumors, as well as an early assessment of each tumor's aggression potential and its response to specific treatments. However, current literature often lacks methodological rigor, with few studies adhering to the guidelines recommended by learned societies concerning AI methodologies. Our study examines the efficacy of ML and DL methodologies in predicting treatment outcomes for elderly immunocompetent patients with PCNSL, employing a stringent methodological framework [151,214]. Our preliminary results suggest that these techniques have the potential to develop reliable and scalable algorithms, which could lead to personalized treatment options for patients by detecting the aggressive potential of tumors from baseline imaging.

In the field of neuro-oncology, research on ML and DL has mainly been directed at distinguishing between GBM and PCNSL, with fewer studies addressing their ability to predict treatment outcomes or patient survival [164,166,168]. Furthermore, while radiomic analyses have progressed, they often emphasize the accuracy of radiomic signatures without exploring their biological relevance. Yet, it is recognized that tumours with greater intratumoral heterogeneity typically have a poorer prognosis, possibly due to intrinsic aggressive nature or resistance to treatment [250]. In PCNSL, Chen et al.'s work highlighted that a homogeneity

index (GLCMHomogeneity) on T1 with gadolinium MRI was significantly associated with better OS [247]. This association could reflect the pathological processes of the tumor, such as neovascularization and hypoxia, leading to necrosis on MRI; often considered an indicator of poor prognosis. Interestingly, Destito et al.'s study on radiomics in PCNSL found that three of our six model parameters were identical, even though they were extracted from T1-weighted MRIs without contrast [169]. These included the maximum 2D diameter in the sagittal plane, the large area low grey level emphasis, and the zone percentage. Furthermore, our model emphasized sphericity as a biomarker indicative of a positive response. This is consistent with findings by Chung et al., who, in a cohort of 50 immunocompetent PCNSL patients, observed that a diffuse round DCE-MRI pattern was significantly more associated with complete response (CR) compared to a non-diffuse multifocal pattern [99].

Applying DL in rare conditions like PCNSL poses challenges due to limited data availability. She et al. demonstrated the potential of a ResNet model with transfer learning to predict OS in a small PCNSL group of 56 patients [168]. This approach uses a model previously trained on extensive datasets in related fields, specifically gliomas, and adapts it for use with PCNSL.

Our study has several strengths. First, rigorous methodologies were used to extract the radiomic features and develop the DL model, aligning with current guidelines [149,173,214–216]. Secondly, we used a genuine external validation cohort, acquired using a different camera and over a distinct time window. Thirdly, to reduce the inherent "black box" effect of artificial intelligence, we closely analyzed the identified radiomics and applied the Grad-CAM method to elucidate the network and cases of failure [228]. Nevertheless, our study has limitations. The case number was restricted due to the rarity of the disease. In addition, patients were recruited by a tertiary center, inducing a potential selection bias. We centered on the end-of-induction treatment response over progression-free or overall survival, which could carry broader implications. Yet, this focus diminishes the confounding impact of inconsistent maintenance treatments, particularly relevant in the older PCNSL population. Finally, expanding our model to include clinical-biological and imaging data could potentially enrich the predictive algorithm, but at the expense of increasing the risk of statistical overextension.

In conclusion, our research opens a promising avenue in applying AI to neuro-oncology imaging. The present approach combines FDG-PET and MRI data with radiomic and deep learning methods in compliance with current guidelines. Further studies with larger prospective cohorts are necessary to substantiate our findings.

References

1. Lauw MIS, Lucas C-HG, Ohgami RS, Wen KW. Primary Central Nervous System Lymphomas: A Diagnostic Overview of Key Histomorphologic, Immunophenotypic, and Genetic Features. *Diagnostics*. 2020;10:1076.
2. Houillier C, Soussain C, Ghesquières H, Soubeyran P, Chinot O, Taillandier L, et al. Management and outcome of primary CNS lymphoma in the modern era: An LOC network study. *Neurology*. 2020;94:e1027–39.
3. Morales-Martinez A, Lozano-Sanchez F, Duran-Peña A, Hoang-Xuan K, Houillier C. Primary Central Nervous System Lymphoma in Elderly Patients: Management and Perspectives. *Cancers (Basel)*. 2021;13.
4. Ferreri AJM, Calimeri T, Cwynarski K, Dietrich J, Grommes C, Hoang-Xuan K, et al. Primary central nervous system lymphoma. *Nat Rev Dis Primers*. 2023;9:29.
5. Hoang-Xuan K, Deckert M, Ferreri AJM, Furtner J, Gallego Perez-Larraya J, Henriksson R, et al. European Association of Neuro-Oncology (EANO) guidelines for treatment of primary central nervous system lymphoma (PCNSL). *Neuro Oncol*. 2023;25:37–53.
6. Abrey LE, Ben-Porat L, Panageas KS, Yahalom J, Berkey B, Curran W, et al. Primary central nervous system lymphoma: the Memorial Sloan-Kettering Cancer Center prognostic model. *J Clin Oncol*. 2006;24:5711–5.
7. Ferreri AJM, Blay J-Y, Reni M, Pasini F, Spina M, Ambrosetti A, et al. Prognostic scoring system for primary CNS lymphomas: the International Extranodal Lymphoma Study Group experience. *J Clin Oncol*. 2003;21:266–72.
8. Fu F, Sun X, Li Y, Liu Y, Shan Y, Ji N, et al. Dynamic contrast-enhanced magnetic resonance imaging biomarkers predict chemotherapeutic responses and survival in primary central-nervous-system lymphoma. *Eur Radiol*. 2021;31:1863–71.
9. Hatzoglou V, Oh JH, Buck O, Lin X, Lee M, Shukla-Dave A, et al. Pretreatment dynamic contrast-enhanced MRI biomarkers correlate with progression-free survival in primary central nervous system lymphoma. *J Neurooncol*. 2018;140:351–8.
10. Valles FE, Perez-Valles CL, Regalado S, Barajas RF, Rubenstein JL, Cha S. Combined diffusion and perfusion MR imaging as biomarkers of prognosis in immunocompetent patients with primary central nervous system lymphoma. *AJNR Am J Neuroradiol*. 2013;34:35–40.
11. Barajas RF, Rubenstein JL, Chang JS, Hwang J, Cha S. Diffusion-weighted MR imaging derived apparent diffusion coefficient is predictive of clinical outcome in primary central nervous system lymphoma. *AJNR Am J Neuroradiol*. 2010;31:60–6.
12. Tabouret E, Houillier C, Martin-Duverneuil N, Blonski M, Soussain C, Ghesquières H, et al. Patterns of response and relapse in primary CNS lymphomas after first-line chemotherapy: imaging analysis of the ANOCEF-GOELAMS prospective randomized trial. *Neuro Oncol*. 2017;19:422–9.
13. Wieduwilt MJ, Valles F, Issa S, Behler CM, Hwang J, McDermott M, et al. Immunochemotherapy with intensive consolidation for primary CNS lymphoma: a pilot study and prognostic assessment by diffusion-weighted MRI. *Clin Cancer Res*. 2012;18:1146–55.

14. Chong I, Ostrom Q, Khan B, Dandachi D, Garg N, Kotrotsou A, et al. Whole Tumor Histogram Analysis Using DW MRI in Primary Central Nervous System Lymphoma Correlates with Tumor Biomarkers and Outcome. *Cancers (Basel)*. 2019;11:1506.
15. Baek DW, Cho HJ, Bae JH, Sohn SK, Moon JH. Apparent diffusion coefficient as a valuable quantitative parameter for predicting clinical outcomes in patients with newly diagnosed primary CNS lymphoma. *Blood Res*. 2020;55:99–106.
16. Huang W-Y, Wen J-B, Wu G, Yin B, Li J-J, Geng D-Y. Diffusion-Weighted Imaging for Predicting and Monitoring Primary Central Nervous System Lymphoma Treatment Response. *AJNR Am J Neuroradiol*. 2016;37:2010–8.
17. Zhang Y, Zhang Q, Wang X-X, Deng X-F, Zhu Y-Z. Value of pretherapeutic DWI in evaluating prognosis and therapeutic effect in immunocompetent patients with primary central nervous system lymphoma given high-dose methotrexate-based chemotherapy: ADC-based assessment. *Clin Radiol*. 2016;71:1018–29.
18. Chung SR, Choi YJ, Kim HS, Park JE, Shim WH, Kim SJ. Tumor Vascular Permeability Pattern Is Associated With Complete Response in Immunocompetent Patients With Newly Diagnosed Primary Central Nervous System Lymphoma: Retrospective Cohort Study. *Medicine (Baltimore)*. 2016;95:e2624.
19. Blasel S, Vorwerk R, Kiyose M, Mittelbronn M, Brunnberg U, Ackermann H, et al. New MR perfusion features in primary central nervous system lymphomas: pattern and prognostic impact. *J Neurol*. 2018;265:647–58.
20. Luo X, Zhu Y, Zhang Y, Zhang Q, Wang X, Deng X. Parameters of MR perfusion-weighted imaging predict the response and prognosis to high-dose methotrexate-based chemotherapy in immunocompetent patients with primary central nervous system lymphoma. *J Clin Neurosci*. 2022;95:151–8.
21. Gupta T, Manjali JJ, Kannan S, Purandare N, Rangarajan V. Diagnostic Performance of Pretreatment 18F-Fluorodeoxyglucose Positron Emission Tomography With or Without Computed Tomography in Patients With Primary Central Nervous System Lymphoma: Updated Systematic Review and Diagnostic Test Accuracy Meta-analyses. *Clin Lymphoma Myeloma Leuk*. 2021;21:497–507.
22. Albano D, Bosio G, Bertoli M, Giubbini R, Bertagna F. 18F-FDG PET/CT in primary brain lymphoma. *J Neurooncol*. 2018;136:577–83.
23. Krebs S, Mauguen A, Yildirim O, Hatzoglou V, Francis JH, Schaff LR, et al. Prognostic value of [18F]FDG PET/CT in patients with CNS lymphoma receiving ibrutinib-based therapies. *Eur J Nucl Med Mol Imaging*. 2021;48:3940–50.
24. Kasenda B, Haug V, Schorb E, Fritsch K, Finke J, Mix M, et al. 18F-FDG PET is an independent outcome predictor in primary central nervous system lymphoma. *J Nucl Med*. 2013;54:184–91.
25. Rozenblum L, Galanaud D, Houillier C, Soussain C, Baptiste A, Belin L, et al. [18F]FDG PET-MRI provides survival biomarkers in primary central nervous system lymphoma in the elderly: an ancillary study from the BLOCAGE trial of the LOC network. *Eur J Nucl Med Mol Imaging*. 2023;
26. Hatt M, Krizsan AK, Rahmim A, Bradshaw TJ, Costa PF, Forgacs A, et al. Joint EANM/SNMMI guideline on radiomics in nuclear medicine : Jointly supported by the EANM Physics Committee and the SNMMI Physics, Instrumentation and Data Sciences Council. *Eur J Nucl Med Mol Imaging*. 2023;50:352–75.
27. Acharya UR, Hagiwara Y, Sudarshan VK, Chan WY, Ng KH. Towards precision medicine: from quantitative imaging to radiomics. *J Zhejiang Univ Sci B*. 2018;19:6–24.

28. Lohmann P, Lerche C, Bauer EK, Steger J, Stoffels G, Blau T, et al. Predicting IDH genotype in gliomas using FET PET radiomics. *Sci Rep.* 2018;8:13328.
29. Zhu M, Li S, Kuang Y, Hill VB, Heimberger AB, Zhai L, et al. Artificial intelligence in the radiomic analysis of glioblastomas: A review, taxonomy, and perspective. *Front Oncol.* 2022;12:924245.
30. Cao X, Tan D, Liu Z, Liao M, Kan Y, Yao R, et al. Differentiating solitary brain metastases from glioblastoma by radiomics features derived from MRI and 18F-FDG-PET and the combined application of multiple models. *Sci Rep.* 2022;12:5722.
31. Guha A, Goda JS, Dasgupta A, Mahajan A, Halder S, Gawde J, et al. Classifying primary central nervous system lymphoma from glioblastoma using deep learning and radiomics based machine learning approach - a systematic review and meta-analysis. *Front Oncol.* 2022;12:884173.
32. Kong Z, Jiang C, Zhu R, Feng S, Wang Y, Li J, et al. 18F-FDG-PET-based radiomics features to distinguish primary central nervous system lymphoma from glioblastoma. *Neuroimage Clin.* 2019;23:101912.
33. Destito M, Marzullo A, Leone R, Zaffino P, Steffanoni S, Erbella F, et al. Radiomics-Based Machine Learning Model for Predicting Overall and Progression-Free Survival in Rare Cancer: A Case Study for Primary CNS Lymphoma Patients. *Bioengineering (Basel).* 2023;10:285.
34. Chen C, Zhuo H, Wei X, Ma X. Contrast-Enhanced MRI Texture Parameters as Potential Prognostic Factors for Primary Central Nervous System Lymphoma Patients Receiving High-Dose Methotrexate-Based Chemotherapy. *Contrast Media Mol Imaging.* 2019;2019:5481491.
35. Cheng D, Liu M, Fu J, Wang Y. Classification of MR brain images by combination of multi-CNNs for AD diagnosis. In: Falco CM, Jiang X, editors. Hong Kong, China; 2017 [cited 2024 Apr 8]. p. 1042042. Available from: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2281808>
36. She Z, Marzullo A, Destito M, Spadea MF, Leone R, Anzalone N, et al. Deep learning-based overall survival prediction model in patients with rare cancer: a case study for primary central nervous system lymphoma. *Int J Comput Assist Radiol Surg.* 2023;18:1849–56.
37. Li B, Xu Y, Wang Y, Zhang B. DECTNet: Dual Encoder Network combined convolution and Transformer architecture for medical image segmentation. *PLoS One.* 2024;19:e0301019.
38. Song B, Kc DR, Yang RY, Li S, Zhang C, Liang R. Classification of Mobile-Based Oral Cancer Images Using the Vision Transformer and the Swin Transformer. *Cancers (Basel).* 2024;16:987.
39. Orlhac F, Nioche C, Klyuzhin I, Rahmim A, Buvat I. Radiomics in PET Imaging:: A Practical Guide for Newcomers. *PET Clin.* 2021;16:597–612.
40. Vovk U, Pernus F, Likar B. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Trans Med Imaging.* 2007;26:405–21.
41. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEx: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res.* 2018;78:4786–9.
42. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 2017;77:e104–7.

43. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020;295:328–38.
44. Pfaehler E, Beukinga RJ, de Jong JR, Slart RHJA, Slump CH, Dierckx RAJO, et al. Repeatability of 18 F-FDG PET radiomic features: A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method. *Med Phys*. 2019;46:665–78.
45. Gao X, Chu C, Li Y, Lu P, Wang W, Liu W, et al. The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from (18)F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur J Radiol*. 2015;84:312–7.
46. Desbordes P, Ruan S, Modzelewski R, Pineau P, Vauclin S, Gouel P, et al. Predictive value of initial FDG-PET features for treatment response and survival in esophageal cancer patients treated with chemo-radiation therapy using a random forest classifier. Ahmad A, editor. *PLoS ONE*. 2017;12:e0173208.
47. Schapire RE. Explaining AdaBoost. In: Schölkopf B, Luo Z, Vovk V, editors. *Empirical Inference* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013 [cited 2024 Apr 12]. p. 37–52. Available from: https://link.springer.com/10.1007/978-3-642-41136-6_5
48. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. 2021 [cited 2024 Apr 12]; Available from: <https://arxiv.org/abs/2103.14030>
49. Lacerda P, Barros B, Albuquerque C, Conci A. Hyperparameter Optimization for COVID-19 Pneumonia Diagnosis Based on Chest CT. *Sensors (Basel)*. 2021;21:2174.
50. Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiology: Artificial Intelligence*. 2020;2:e200029.
51. Wiltgen T, Fleischmann DF, Kaiser L, Holzgreve A, Corradini S, Landry G, et al. 18F-FET PET radiomics-based survival prediction in glioblastoma patients receiving radio(chemo)therapy. *Radiat Oncol*. 2022;17:198.
52. Garaba A, Aslam N, Ponzio F, Panciani PP, Brinjikji W, Fontanella M, et al. Radiomics for differentiation of gliomas from primary central nervous system lymphomas: a systematic review and meta-analysis. *Front Oncol*. 2024;14:1291861.
53. Davnall F, Yip CSP, Ljungqvist G, Selmi M, Ng F, Sanghera B, et al. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights Imaging*. 2012;3:573–89.
54. Kocak B, Akinci D'Antonoli T, Mercaldo N, Alberich-Bayarri A, Baessler B, Ambrosini I, et al. METHodological RadiomICs Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMII. *Insights Imaging*. 2024;15:8.
55. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48:441–6.
56. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med*. 2020;26:1320–4.

57. Lekadir K, Feragen A, Fofanah AJ, Frangi AF, Buyx A, Emelie A, et al. FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. 2023 [cited 2024 Apr 12]; Available from: <https://arxiv.org/abs/2309.12325>
58. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. 2017 IEEE International Conference on Computer Vision (ICCV) [Internet]. Venice: IEEE; 2017 [cited 2024 Apr 10]. p. 618–26. Available from: <http://ieeexplore.ieee.org/document/8237336/>

Chapitre 7. TEP-IRM à la 18F-Fludarabine

En dehors du FDG, d'autres radiotraceurs TEP ont été investigués dans la littérature pour l'imagerie des LPSNC. Dans une petite cohorte de 26 patients, l'équipe de Ahn et al. a par exemple souligné l'intérêt de deux marqueurs TEP à la 11C-Méthionine (le ratio tumeur/cortex controlatéral et le MTV) comme outils pronostiques de la survie globale [251]. D'autres séries de cas ont appuyé ces résultats, en montrant que la 11C-Méthionine permettait d'aider au diagnostic de la maladie résiduelle en cas de réponse complète non confirmée (CRu) [252,253]. En revanche, les études comparatives entre le FDG et la Méthionine n'ont pas mis en évidence de supériorité de la Méthionine dans le LPSNC [254]. En particulier, une étude récente a montré que le FDG surpassait la Méthionine dans la distinction entre LPSNC et GBM (pas de différence significative du ratio tumeur sur tissu sain en Méthionine versus T/N=3.02±1.32 pour les LPSNC contre 1.49 ± 0.76 pour les GBM en FDG, p<0.001) [255]. De plus, la Méthionine présente des difficultés techniques pour son utilisation en clinique, du fait de la brève demi-vie radioactive du Carbone-11(20 minutes).

Récemment, un autre radiotraceur, la 18F-Fludarabine, a été investiguée, notamment en préclinique (Figure 7.1). La Fludarabine est un analogue de l'adénosine, utilisé en oncologie dans le traitement des leucémies lymphoïdes chroniques ou des lymphomes indolents [256]. La captation de la Fludarabine par les cellules lymphocytaires est assurée par deux types de transporteurs : les ENTs (equilibrative nucleoside transporter) et les CNTs (concentrative nucleoside transporter) [257]. Une fois dans la cellule, la molécule est phosphorylée plusieurs fois, ce qui aboutit à la formation d'un métabolite triphosphate (F-ara-ATP), responsable des effets cytotoxiques dans la cellule par inhibition de la synthèse de l'ADN. Grâce à la substitution d'un atome de fluor-19 par un atome de fluor-18 ; un radiopharmaceutique fluoré a été développé pour l'imagerie TEP, initialement dans le laboratoire Cyceron, en France. Du fait des doses extrêmement faibles injectés de 18F-Fludarabine pour l'imagerie médicale, aucun effet cytotoxique n'est attendu avec le radiopharmaceutique TEP.

Nom chimique : [^{18}F]-Fludarabine
Structure : $\text{C}_{10}\text{H}_{12}^{18}\text{FN}_5\text{O}_4$
Poids moléculaire = 285,23 g.mol $^{-1}$

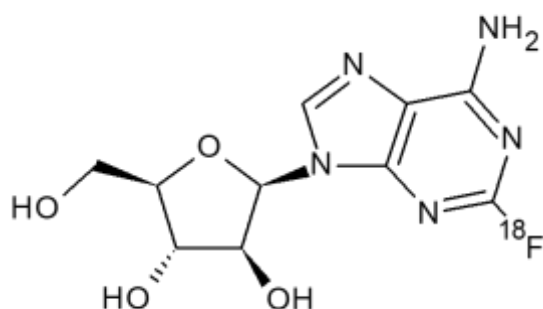


FIGURE 7.1. Représentation de la molécule de 18F-Fludarabine

Comparativement au FDG, la Fludarabine présenterait une affinité plus forte pour les lymphocytes B [258]. En outre, son absence de captation par le parenchyme cérébral sain, promet un meilleur contraste signal-bruit en imagerie TEP. Les recherches précliniques initiales sur la 18F-fludarabine, impliquant des souris saines et d'autres porteuses de lymphomes xénotransplantés, ont révélé une captation spécifique au tissu lymphoïde. En 2018, une étude sur des modèles murins xénotransplantés de lymphome cérébral (n=9) et de glioblastomes (n=4) a renforcé ces résultats en montrant que le ratio de captation tumeur/tissu sain était significativement plus élevé dans le lymphome intracérébral et corrélé avec la présence du CD79, un marqueur des cellules B malignes [21] (Figure 7.2). Ce radiopharmaceutique pourrait donc être un outil de diagnostic alternatif à la biopsie pour les patients avec un diagnostic différentiel incertain entre GBM et LPSNC, notamment ceux ne pouvant pas subir de biopsie en raison de comorbidités importantes ou d'une localisation difficilement accessible.

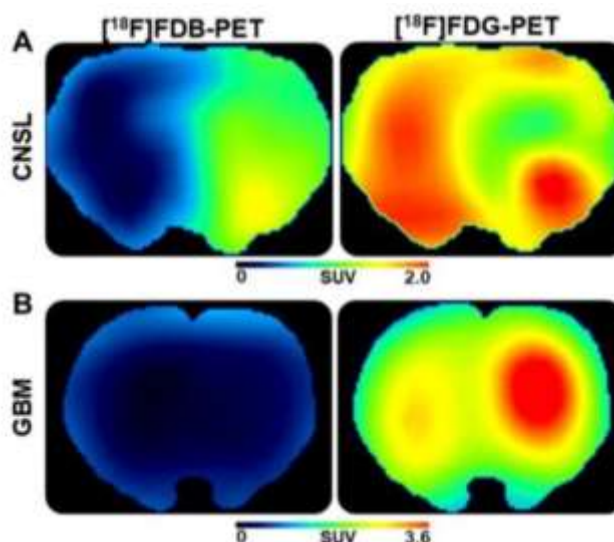


FIGURE 7.2. Images TEP de modèles murins de LPSNC et GBM en 18F-Fludarabine et 18F-FDG

D'après Hovanishian, theranostic, 2018 [21]

Imagerie en coupe de souris xenogreffées avec un modèle de LPSNC (A) ou de GBM (B). La première colonne montre une imagerie TEP petit-animal en 18F-Fludarabine (fixation intense du LPSNC et nulle du GBM) tandis que la deuxième colonne montre une imagerie en 18F-FDG (fixation similaire des deux types de tumeurs).

Chez l'humain, l'utilisation de ce radiotracer a été rapportée principalement dans un contexte systémique. Une étude préliminaire menée auprès de cinq patients souffrant de leucémie lymphoïde chronique de type B a révélé une adéquation entre la captation de 18F-Fludarabine en TEP-TDM et la localisation clinique et tomodynamométrique des sites tumoraux (ganglions, amygdales, rate) [20].

Une seule étude a exploré l'emploi de la 18F-Fludarabine en intracérébral chez l'homme [18]. Réalisée par Postnov et al., cette recherche visait à affiner la distinction diagnostique entre le LPSNC et le GBM. Les 31 participants ont reçu une dose de 4MBq/kg de 18F-Fludarabine, suivie de trois acquisitions TEP de 20 minutes chacune : la première à 10 minutes post-injection, la seconde à 40 minutes, et la dernière à 100 minutes. L'analyse a révélé des profils de captation distincts pour les deux pathologies : une diminution de la fixation après un pic initial (wash-out) pour les GBM, et une courbe d'intensité progressivement croissante pour les LPSNC.

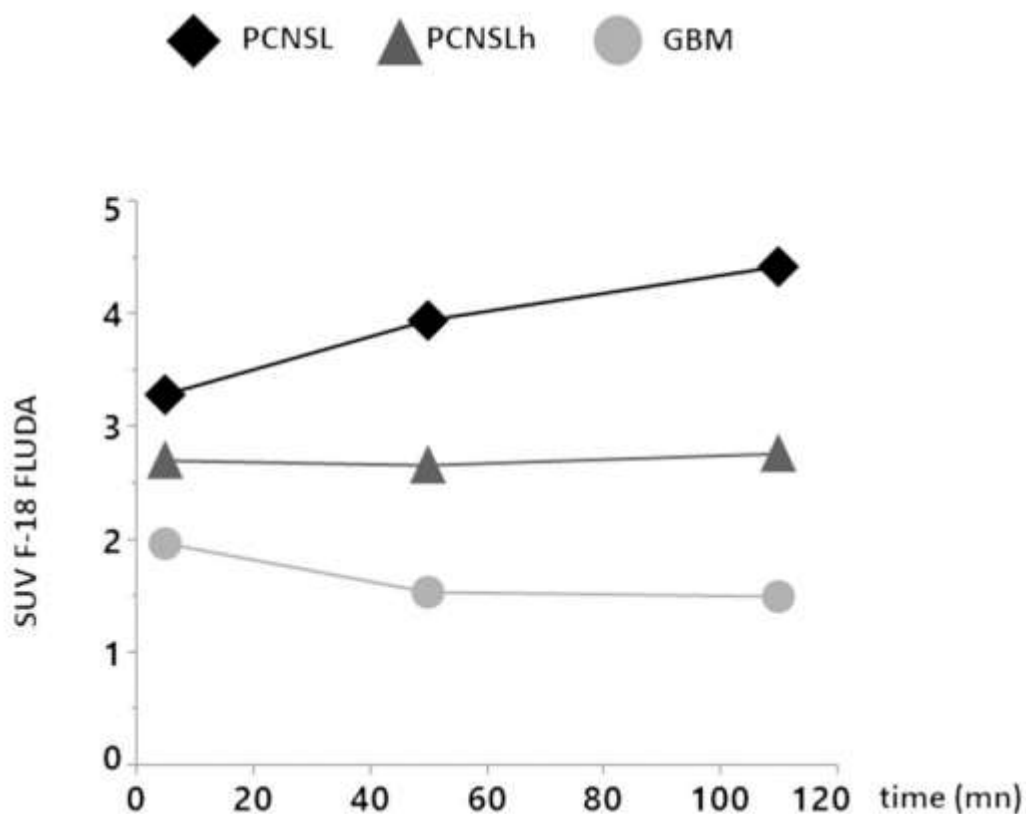


FIGURE 7.3. Evolution cinétique du ratio tumeur sur tissu sain controlatéral en 18F-Fludarabine

D'après Postnov et al. CNM, 2022; PCNSL : primary central nervous system lymphoma, PCNSLh : PCNSL treated with corticosteroids, GBM : glioblastoma

La principale contrainte de cette étude résidait dans le fait que 10 des 18 patients atteints de LPSNC étaient sous corticoïdes lors de l'évaluation TEP. Pour ces patients, les paramètres cinétiques et semi-quantitatifs sur l'acquisition tardive (entre 100-120 minutes) ne permettaient pas de distinguer de manière significative le LPSNC du GBM. Or, les patients atteints de LPSNC sont souvent en mauvais état général au moment du diagnostic et nécessitent un traitement rapide par corticoïdes, ce qui rend difficile la réalisation d'une acquisition TEP avant l'instauration de ce traitement. Une autre limite repose sur le fait que les paramètres retenus par les auteurs comme les plus significatifs, sont ceux de la troisième acquisition, soit à plus d'une heure trente post-injection. Là encore, une application clinique semble compromise car l'état général des patients ne permet pas une attente si longue.

En conséquence, une analyse réalisée sur un intervalle de temps plus court semble plus adaptée pour définir des paramètres TEP à la 18F-Fludarabine, qui seraient facilement transposables en

contexte clinique. C'est dans cette perspective que l'étude pilote Fludaloc (PI : Prof. A. Kas), actuellement en cours à la Pitié-Salpêtrière et soutenue par le réseau LOC, a été conçue.

Cette étude vise à caractériser la distribution cérébrale, la cinétique tissulaire, et l'intensité de la fixation de la 18F-Fludarabine en TEP-IRM chez huit patients présentant des lésions de LPSNC, prouvées histologiquement, avant tout traitement. L'objectif secondaire est de comparer les données de la TEP-Fludarabine avec celles de la TEP-FDG et de l'IRM multimodalité. Chaque patient bénéficie d'une injection de 4 MBq/kg de 18F-Fludarabine et d'une acquisition dynamique des images de 0 à 60 minutes. Initialement prévue pour débiter au début de ma thèse, cette étude a été retardée en raison de difficultés de radio synthèse du radiopharmaceutique. Les premières inclusions ont débuté en mai 2024 et à ce jour, deux patients ont bénéficié d'une TEP-IRM à la 18F-Fludarabine.

Pour l'analyse des images, les paramètres suivants seront colligés : le SUVmax de la lésion, le ratio du SUVmax de la lésion à la captation du tissu sain mesurée soit sur une zone miroir controlatérale (TBRmax), soit sur le cervelet, ainsi que le MTV et le TLG. Nous examinerons également la cinétique de captation de la 18F-Fludarabine de T0 à 60 minutes post-injection. L'IRM sera analysée de manière semi-quantitative, incluant la mesure du volume sur les séquences T1 avec injection de produit de contraste et FLAIR, ainsi que le calcul des paramètres de diffusion (ADC), de perfusion (rCBV et rCBF sur la perfusion DSC et rCBF sur l'ASL) et de spectroscopie. Enfin, nous identifierons les zones de divergence (mismatch) en TEP et IRM pour évaluer si elles présentent des caractéristiques tumorales différentes des zones parfaitement matchées.

Ci-dessous sont présentés les images statiques et l'analyse cinétique du premier patient.

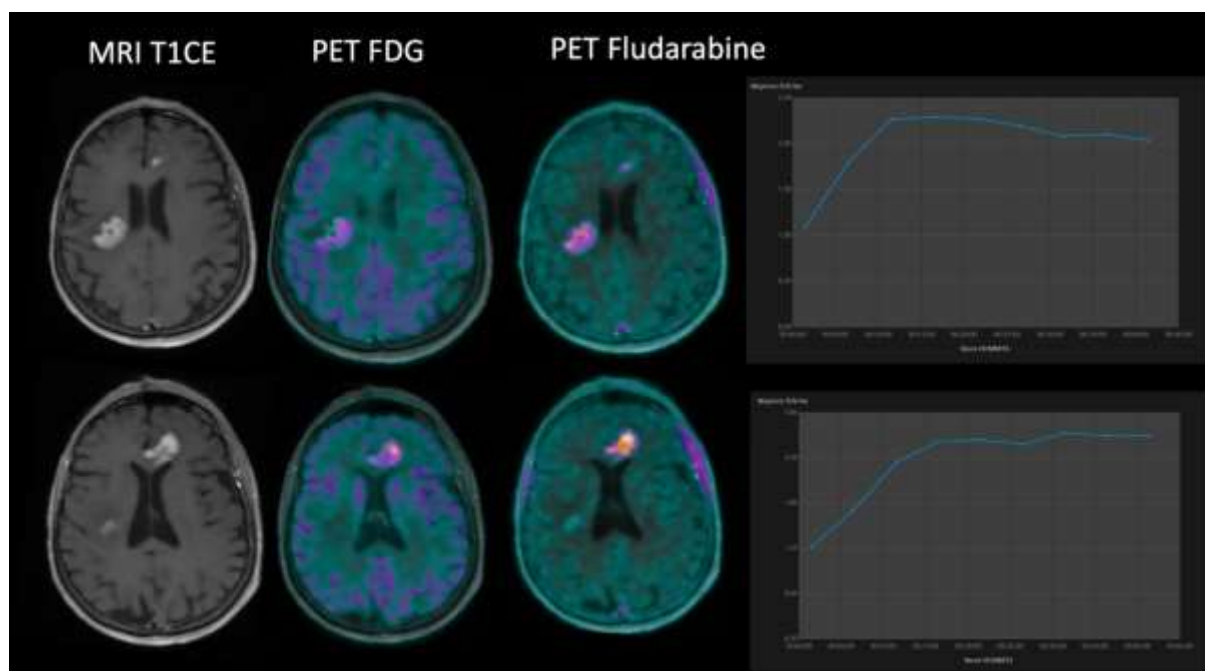


FIGURE 7.4. Analyse préliminaire des premières images TEP-IRM à la 18F-Fludarabine

Patiente de 68 ans avec localisation cérébrale d'un LBDGC, sans atteinte ophtalmologique. L'IRM révèle deux lésions prenant intensément le contraste, l'une en fronto-pariétale droite de 23 mm avec large plage d'œdème et l'autre en paramédiane gauche de 27mm avec léger effet de masse et œdème périphérique. La TEP révèle un hypermétabolisme des deux lésions avec un ratio du métabolisme tumoral au parenchyme sain controlatéral élevé : SUVmax=11,4, TBRmax=1,4 pour la lésion fronto-pariétale droite et SUVmax=12,7, TBRmax = 1.9 pour la lésion frontale gauche en TEP-FDG. Sur la TEP-Fludarabine : SUVmax=3,3, TBRmax=2, 5 pour la lésion fronto-pariétale droite et SUVmax=4,4, TBRmax=4 pour la lésion frontale gauche. Les courbes cinétiques montrent une captation croissante du traceur au cours du temps pour les deux lésions, avec toutefois un aspect de plateau plus marqué dès la 10ème minute pour la lésion fronto-pariétale droite

La durée prévue de l'étude Fludaloc est de 12 mois. En parallèle, une autre étude, GLIOLOC (IP : Prof. A Kas), est actuellement en attente de financement. Cette recherche aura pour but de comparer la cinétique et l'intensité de captation de la 18F-Fludarabine en TEP-IRM chez des patients diagnostiqués avec LPSNC (n=10) et GBM (n=10). L'objectif à long terme est de développer des études utilisant la 18F-Fludarabine comme moyen d'évaluation thérapeutique pour les LPSNC. Nos résultats présentés précédemment ont en effet indiqué que le métabolisme en FDG est moindre sous traitement comparativement aux mesures baseline, (Article 3, Supplemental Table 2) et un traceur plus spécifique permettrait donc d'augmenter l'efficacité du diagnostic en cas de récurrence.

Chapitre 8. Conclusion et perspectives

En conclusion, les résultats présentés dans cette thèse confirment l'intérêt de la TEP au 18F-FDG comme outil d'imagerie dans la prise en charge des patients atteints de LPSNC. Deux marqueurs de la TEP baseline apparaissent comme prédictifs de la réponse au traitement : le volume tumoral global et la fixation physiologique sur le cervelet sain. Par ailleurs, nous avons démontré la supériorité de la TEP sur l'IRM dans les cas de réponse partielle sur cette dernière au moment de l'évaluation interim, comme outil de stratification des patients entre « bons » ou « mauvais » répondeurs. Ces découvertes encouragent la validation de ces index dans de plus larges études multicentriques avec une population plus diversifiée. Pour ce faire, nous avons établi une méthodologie robuste d'analyse en *machine-learning* et *deep-learning* dont les modèles pourraient être utilisés à large échelle.

Nos travaux ouvrent la voie à l'intégration de la TEP FDG dans les recommandations de prise en charge des LPSNC. Les résultats de ces recherches ont déjà eu un impact clinique, puisque les cliniciens adoptent maintenant un rythme de surveillance IRM plus rapproché pour les patients avec réponse défavorable sur la TEP interim. Dans le futur, les nouveaux essais thérapeutiques pourraient adapter la prise en charge en fonction du groupe de réponse TEP-FDG. Par ailleurs, une étude pilote investiguant la 18F-Fludarabine est actuellement en cours pour la caractérisation des LPSNC au diagnostic. Ce radiotracer pourrait également être utilisé pour le diagnostic différentiel entre GBM et LPSNC ou encore pour l'évaluation thérapeutique, de manière plus sensible que le FDG.

Bibliographie

1. Houillier C, Soussain C, Ghesquières H, Soubeyran P, Chinot O, Taillandier L, et al. Management and outcome of primary CNS lymphoma in the modern era: An LOC network study. *Neurology*. 2020;94:e1027–39.
2. Bauchet L, Rigau V, Mathieu-Daudé H, Figarella-Branger D, Hugues D, Palusseau L, et al. French brain tumor data bank: Methodology and first results on 10,000 cases. *J Neurooncol*. 2007;84:189–99.
3. Hoang-Xuan K, Bessell E, Bromberg J, Hottinger AF, Preusser M, Rudà R, et al. Diagnosis and treatment of primary CNS lymphoma in immunocompetent patients: guidelines from the European Association for Neuro-Oncology. *Lancet Oncol*. 2015;16:e322–332.
4. Morales-Martinez A, Nichelli L, Hernandez-Verdin I, Houillier C, Alentorn A, Hoang-Xuan K. Prognostic factors in primary central nervous system lymphoma. *Curr Opin Oncol*. 2022;34:676–84.
5. Barajas RF, Politi LS, Anzalone N, Schöder H, Fox CP, Boxerman JL, et al. Consensus recommendations for MRI and PET imaging of primary central nervous system lymphoma: guideline statement from the International Primary CNS Lymphoma Collaborative Group (IPCG). *Neuro Oncol*. 2021;23:1056–71.
6. Abrey LE, Batchelor TT, Ferreri AJM, Gospodarowicz M, Pulczynski EJ, Zucca E, et al. Report of an international workshop to standardize baseline evaluation and response criteria for primary CNS lymphoma. *J Clin Oncol*. 2005;23:5034–43.
7. Pels H, Juergens A, Schirgens I, Glasmacher A, Schulz H, Engert A, et al. Early complete response during chemotherapy predicts favorable outcome in patients with primary CNS lymphoma. *Neuro-Oncology*. 2010;12:720–4.
8. van der Meulen M, Postma AA, Smits M, Bakunina K, Minnema MC, Seute T, et al. Extent of radiological response does not reflect survival in primary central nervous system lymphoma. *Neurooncol Adv*. 2021;3:vdab007.
9. Krebs S, Mauguén A, Yildirim O, Hatzoglou V, Francis JH, Schaff LR, et al. Prognostic value of [18F]FDG PET/CT in patients with CNS lymphoma receiving ibrutinib-based therapies. *Eur J Nucl Med Mol Imaging*. 2021;48:3940–50.
10. Krebs S, Barasch JG, Young RJ, Grommes C, Schöder H. Positron emission tomography and magnetic resonance imaging in primary central nervous system lymphoma—a narrative review. *Ann Lymphoma*. 2021;5.
11. Rozenblum L, Houillier C, Soussain C, Bertaux M, Choquet S, Galanaud D, et al. Role of Positron Emission Tomography in Primary Central Nervous System Lymphoma. *Cancers (Basel)*. 2022;14:4071.
12. Palmedo H, Urbach H, Bender H, Schlegel U, Schmidt-Wolf IGH, Matthies A, et al. FDG-PET in immunocompetent patients with primary central nervous system lymphoma: correlation with MRI and clinical follow-up. *Eur J Nucl Med Mol Imaging*. 2006;33:164–8.
13. Kasenda B, Haug V, Schorb E, Fritsch K, Finke J, Mix M, et al. 18F-FDG PET is an independent outcome predictor in primary central nervous system lymphoma. *J Nucl Med*. 2013;54:184–91.
14. Jo J-C, Yoon DH, Kim S, Lee K, Kang EH, Park JS, et al. Interim 18F-FDG PET/CT may not predict the outcome in primary central nervous system lymphoma patients treated with sequential treatment with methotrexate and cytarabine. *Ann Hematol*. 2017;96:1509–15.
15. Albano D, Bertoli M, Battistotti M, Rodella C, Statuto M, Giubbini R, et al. Prognostic role of pretreatment 18F-FDG PET/CT in primary brain lymphoma. *Ann Nucl Med*. 2018;32:532–41.
16. Albano D, Bosio G, Bertoli M, Giubbini R, Bertagna F. 18F-FDG PET/CT in primary

brain lymphoma. *J Neurooncol.* 2018;136:577–83.

17. de-Bonilla-Damiá Á, Fernández-López R, Capote-Huelva FJ, de la Cruz-Vicente F, Egea-Guerrero JJ, Borrego-Dorado I. Role of 18F-FDG PET/CT in primary brain lymphoma. *Rev Esp Med Nucl Imagen Mol.* 2017;36:298–303.

18. Postnov A, Toutain J, Pronin I, Valable S, Gourand F, Kalaeva D, et al. First-in-Man Noninvasive Initial Diagnostic Approach of Primary CNS Lymphoma Versus Glioblastoma Using PET With 18F-Fludarabine and L-[methyl-11C]Methionine. *Clin Nucl Med.* 2022;

19. Barré L, Hovhannisyán N, Bodet-Milin C, Kraeber-Bodéré F, Damaj G. [18F]-Fludarabine for Hematological Malignancies. *Front Med (Lausanne).* 2019;6:77.

20. Chantepie S, Hovhannisyán N, Guillouet S, Pelage J-P, Ibazizene M, Bodet-Milin C, et al. 18F-Fludarabine PET for Lymphoma Imaging: First-in-Humans Study on DLBCL and CLL Patients. *J Nucl Med.* 2018;59:1380–5.

21. Hovhannisyán N, Fillesoye F, Guillouet S, Ibazizene M, Toutain J, Gourand F, et al. [18F]Fludarabine-PET as a promising tool for differentiating CNS lymphoma and glioblastoma: Comparative analysis with [18F]FDG in human xenograft models. *Theranostics.* 2018;8:4563–73.

22. Hovhannisyán N, Dhilly M, Fidalgo M, Fillesoye F, Guillouet S, Sola B, et al. [18F]Fludarabine-PET in a murine model of multiple myeloma. *PLoS One.* 2017;12:e0177125.

23. Hovhannisyán N, Dhilly M, Guillouet S, Leporrier M, Barré L. Comparative Analysis between [(18)F]Fludarabine-PET and [(18)F]FDG-PET in a Murine Model of Inflammation. *Mol Pharm.* 2016;13:2136–9.

24. Hovhannisyán N, Guillouet S, Fillesoye F, Dhilly M, Patin D, Galateau F, et al. Evaluation of the specificity of [(18)F]fludarabine PET/CT in a xenograft model of follicular lymphoma: comparison with [(18)F]FDG and impact of rituximab therapy. *EJNMMI Res.* 2015;5:23.

25. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol.* 2021;23:1231–51.

26. Mendez JS, Ostrom QT, Gittleman H, Kruchko C, DeAngelis LM, Barnholtz-Sloan JS, et al. The elderly left behind—changes in survival trends of primary central nervous system lymphoma over the past 4 decades. *Neuro Oncol.* 2018;20:687–94.

27. Ferreri AJM, Blay J-Y, Reni M, Pasini F, Spina M, Ambrosetti A, et al. Prognostic scoring system for primary CNS lymphomas: the International Extranodal Lymphoma Study Group experience. *J Clin Oncol.* 2003;21:266–72.

28. Ferreri AJM, Calimeri T, Cwynarski K, Dietrich J, Grommes C, Hoang-Xuan K, et al. Primary central nervous system lymphoma. *Nat Rev Dis Primers.* 2023;9:29.

29. Wolf T, Brodt H-R, Fichtlscherer S, Mantzsch K, Hoelzer D, Helm EB, et al. Changing incidence and prognostic factors of survival in AIDS-related non-Hodgkin's lymphoma in the era of highly active antiretroviral therapy (HAART). *Leuk Lymphoma.* 2005;46:207–15.

30. Shao L, Xu C, Wu H, Jamal M, Pan S, Li S, et al. Recent Progress on Primary Central Nervous System Lymphoma-From Bench to Bedside. *Front Oncol.* 2021;11:689843.

31. Schaff LR, Mellingshoff IK. Glioblastoma and Other Primary Brain Malignancies in Adults: A Review. *JAMA.* 2023;329:574–87.

32. Küker W, Nägele T, Korfel A, Heckl S, Thiel E, Bamberg M, et al. Primary central nervous system lymphomas (PCNSL): MRI features at presentation in 100 patients. *J Neurooncol.* 2005;72:169–77.

33. Cheng G, Zhang J. Imaging features (CT, MRI, MRS, and PET/CT) of primary central nervous system lymphoma in immunocompetent patients. *Neurol Sci.* 2019;40:535–42.

34. Capasso R, Negro A, Russo C, Zeccolini F, Muto G, Caranci F, et al. Conventional and Advanced MRI Techniques in the Evaluation of Primary CNS Lymphoma. *Semin Ultrasound*

CT MR. 2023;44:126–35.

35. Haldorsen IS, Kråkenes J, Krossnes BK, Mella O, Espeland A. CT and MR imaging features of primary central nervous system lymphoma in Norway, 1989-2003. *AJNR Am J Neuroradiol.* 2009;30:744–51.
36. Haldorsen IS, Espeland A, Larsson E-M. Central nervous system lymphoma: characteristic findings on traditional and advanced imaging. *AJNR Am J Neuroradiol.* 2011;32:984–92.
37. Ellingson BM, Bendszus M, Boxerman J, Barboriak D, Erickson BJ, Smits M, et al. Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials. *Neuro Oncol.* 2015;17:1188–98.
38. Horger M, Fenchel M, Nägele T, Moehle R, Claussen CD, Beschoner R, et al. Water Diffusivity: Comparison of Primary CNS Lymphoma and Astrocytic Tumor Infiltrating the Corpus Callosum. *American Journal of Roentgenology.* 2009;193:1384–7.
39. Zacharia TT, Law M, Naidich TP, Leeds NE. Central nervous system lymphoma characterization by diffusion-weighted imaging and MR spectroscopy. *J Neuroimaging.* 2008;18:411–7.
40. Talybov R, Beylerli O, Mochalov V, Prokopenko A, Ilyasova T, Trofimova T, et al. Multiparametric MR Imaging Features of Primary CNS Lymphomas. *Front Surg.* 2022;9:887249.
41. Blasel S, Jurcoane A, Bähr O, Weise L, Harter PN, Hattingen E. MR perfusion in and around the contrast-enhancement of primary CNS lymphomas. *J Neurooncol.* 2013;114:127–34.
42. Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp MV, et al. Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusible tracer: standardized quantities and symbols. *J Magn Reson Imaging.* 1999;10:223–32.
43. Kim YS, Choi SH, Yoo R-E, Kang KM, Yun TJ, Kim J-H, et al. Leakage correction improves prognosis prediction of dynamic susceptibility contrast perfusion MRI in primary central nervous system lymphoma. *Sci Rep.* 2018;8:456.
44. You G, Wu H, Lei B, Wan X, Chen S, Zheng N. Diagnostic accuracy of arterial spin labeling in differentiating between primary central nervous system lymphoma and high-grade glioma: a systematic review and meta-analysis. *Expert Review of Anticancer Therapy.* 2022;22:763–71.
45. Ambady P, Hu LS, Politi LS, Anzalone N, Barajas RF. Primary central nervous system lymphoma: advances in MRI and PET imaging. *Ann Lymphoma.* 2021;5:27.
46. Geng M, Song Y, Xiao H, Wu Z, Deng X, Chen C, et al. Clinical significance of interleukin-10 concentration in the cerebrospinal fluid of patients with primary central nervous system lymphoma. *Oncol Lett.* 2021;21:2.
47. Song Y, Zhang W, Zhang L, Wu W, Zhang Y, Han X, et al. Cerebrospinal Fluid IL-10 and IL-10/IL-6 as Accurate Diagnostic Biomarkers for Primary Central Nervous System Large B-cell Lymphoma. *Sci Rep.* 2016;6:38671.
48. O’Garra A, Stapleton G, Dhar V, Pearce M, Schumacher J, Rugo H, et al. Production of cytokines by mouse B cells: B lymphomas and normal B cells produce interleukin 10. *Int Immunol.* 1990;2:821–32.
49. Moore KW, de Waal Malefyt R, Coffman RL, O’Garra A. Interleukin-10 and the interleukin-10 receptor. *Annu Rev Immunol.* 2001;19:683–765.
50. Ferreri AJM, Calimeri T, Lopedote P, Francaviglia I, Daverio R, Iacona C, et al. MYD88 L265P mutation and interleukin-10 detection in cerebrospinal fluid are highly specific discriminating markers in patients with primary central nervous system lymphoma: results from a prospective study. *Br J Haematol.* 2021;193:497–505.
51. Sasagawa Y, Akai T, Tachibana O, Iizuka H. Diagnostic value of interleukin-10 in cerebrospinal fluid for diffuse large B-cell lymphoma of the central nervous system. *J*

Neurooncol. 2015;121:177–83.

52. Bertaux M, Houillier C, Edeline V, Habert M-O, Mokhtari K, Giron A, et al. Use of FDG-PET/CT for systemic assessment of suspected primary central nervous system lymphoma: a LOC study. *J Neurooncol.* 2020;148:343–52.

53. Gupta M, Gupta T, Purandare N, Rangarajan V, Puranik A, Moiyadi A, et al. Utility of flouro-deoxy-glucose positron emission tomography/computed tomography in the diagnostic and staging evaluation of patients with primary CNS lymphoma. *CNS Oncol.* 2019;8:CNS46.

54. Bailey DL, Townsend DW, Valk PE, Maisey MN, editors. *Positron Emission Tomography: Basic Sciences* [Internet]. London: Springer London; 2005 [cited 2024 Apr 10]. Available from: <https://link.springer.com/10.1007/b136169>

55. Badawi RD, Miller MP, Bailey DL, Marsden PK. Randoms variance reduction in 3D PET. *Phys Med Biol.* 1999;44:941–54.

56. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med.* 2007;48:932–45.

57. Erlandsson K, Dickson J, Arridge S, Atkinson D, Ourselin S, Hutton BF. MR Imaging-Guided Partial Volume Correction of PET Data in PET/MR Imaging. *PET Clin.* 2016;11:161–77.

58. Erlandsson K, Buvat I, Pretorius PH, Thomas BA, Hutton BF. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys Med Biol.* 2012;57:R119-159.

59. Kelloff GJ, Hoffman JM, Johnson B, Scher HI, Siegel BA, Cheng EY, et al. Progress and Promise of FDG-PET Imaging for Cancer Patient Management and Oncologic Drug Development. *Clinical Cancer Research.* 2005;11:2785–808.

60. Nelson CA, Wang JQ, Leav I, Crane PD. The interaction among glucose transport, hexokinase, and glucose-6-phosphatase with respect to 3H-2-deoxyglucose retention in murine tumor models. *Nuclear Medicine and Biology.* 1996;23:533–41.

61. Warburg O, Wind F, Negelein E. THE METABOLISM OF TUMORS IN THE BODY. *J Gen Physiol.* 1927;8:519–30.

62. Conti M, Eriksson L. Physics of pure and non-pure positron emitters for PET: a review and a discussion. *EJNMMI Phys.* 2016;3:8.

63. Buvat I. Les limites du SUV. *Médecine Nucléaire.* 2007;31:165–72.

64. Boellaard R. Standards for PET Image Acquisition and Quantitative Data Analysis. *J Nucl Med.* 2009;50:11S-20S.

65. Adams MC, Turkington TG, Wilson JM, Wong TZ. A Systematic Review of the Factors Affecting Accuracy of SUV Measurements. *American Journal of Roentgenology.* 2010;195:310–20.

66. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors. *J Nucl Med.* 2009;50:122S-150S.

67. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging.* 2015;42:328–54.

68. Albano D, Bosio G, Pagani C, Re A, Tucci A, Giubbini R, et al. Prognostic role of baseline 18F-FDG PET/CT metabolic parameters in Burkitt lymphoma. *Eur J Nucl Med Mol Imaging.* 2019;46:87–96.

69. Albano D, Mazzoletti A, Spallino M, Muzi C, Zilioli VR, Pagani C, et al. Prognostic role of baseline 18F-FDG PET/CT metabolic parameters in elderly HL: a two-center experience in 123 patients. *Ann Hematol.* 2020;99:1321–30.

70. Barrington SF, Zwezerijnen BGJC, de Vet HCW, Heymans MW, Mikhaeel NG, Burggraaff CN, et al. Automated Segmentation of Baseline Metabolic Total Tumor Burden in Diffuse Large B-Cell Lymphoma: Which Method Is Most Successful? A Study on Behalf of the PETRA Consortium. *J Nucl Med.* 2021;62:332–7.

71. Larson SM, Erdi Y, Akhurst T, Mazumdar M, Macapinlac HA, Finn RD, et al. Tumor Treatment Response Based on Visual and Quantitative Changes in Global Tumor Glycolysis Using PET-FDG Imaging. The Visual Response Score and the Change in Total Lesion Glycolysis. *Clin Positron Imaging*. 1999;2:159–71.
72. Wehner J, Weissler B, Dueppenbecker P, Gebhardt P, Schug D, Ruetten W, et al. PET/MRI insert using digital SiPMs: Investigation of MR-compatibility. *Nucl Instrum Methods Phys Res A*. 2014;734:116–21.
73. Pyatigorskaya N, Habert M-O, Rozenblum L. Contribution of PET-MRI in brain diseases in clinical practice. *Curr Opin Neurol*. 2020;33:430–8.
74. Kas A, Rozenblum L, Pyatigorskaya N. Clinical Value of Hybrid PET/MR Imaging: Brain Imaging Using PET/MR Imaging. *Magn Reson Imaging Clin N Am*. 2023;31:591–604.
75. Göttler J, Lukas M, Kluge A, Kaczmarz S, Gempt J, Ringel F, et al. Intra-lesional spatial correlation of static and dynamic FET-PET parameters with MRI-based cerebral blood volume in patients with untreated glioma. *Eur J Nucl Med Mol Imaging*. 2017;44:392–7.
76. Pellerin A, Khalifé M, Sanson M, Rozenblum-Beddok L, Bertaux M, Soret M, et al. Simultaneously acquired PET and ASL imaging biomarkers may be helpful in differentiating progression from pseudo-progression in treated gliomas. *Eur Radiol*. 2021;31:7395–405.
77. Verger A, Filss CP, Lohmann P, Stoffels G, Sabel M, Wittsack HJ, et al. Comparison of 18F-FET PET and perfusion-weighted MRI for glioma grading: a hybrid PET/MR study. *Eur J Nucl Med Mol Imaging*. 2017;44:2257–65.
78. Girard A, Le Reste P-J, Metais A, Carsin Nicol B, Chiforeanu DC, Bannier E, et al. Combining 18F-DOPA PET and MRI with perfusion-weighted imaging improves delineation of high-grade subregions in enhancing and non-enhancing gliomas prior treatment: a biopsy-controlled study. *J Neurooncol*. 2021;155:287–95.
79. Berker Y, Franke J, Salomon A, Palmowski M, Donker HCW, Temur Y, et al. MRI-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/Dixon MRI sequence. *J Nucl Med*. 2012;53:796–804.
80. Sgard B, Khalifé M, Bouchut A, Fernandez B, Soret M, Giron A, et al. ZTE MR-based attenuation correction in brain FDG-PET/MR: performance in patients with cognitive impairment. *Eur Radiol*. 2020;30:1770–9.
81. Blanc-Durand P, Khalife M, Sgard B, Kaushik S, Soret M, Tiss A, et al. Attenuation correction using 3D deep convolutional neural network for brain 18F-FDG PET/MR: Comparison with Atlas, ZTE and CT based attenuation correction. *PLoS One*. 2019;14:e0223141.
82. Grommes C, Rubenstein JL, DeAngelis LM, Ferreri AJM, Batchelor TT. Comprehensive approach to diagnosis and treatment of newly diagnosed primary CNS lymphoma. *Neuro Oncol*. 2019;21:296–305.
83. Morales-Martinez A, Lozano-Sanchez F, Duran-Peña A, Hoang-Xuan K, Houillier C. Primary Central Nervous System Lymphoma in Elderly Patients: Management and Perspectives. *Cancers (Basel)*. 2021;13:3479.
84. Hoang-Xuan K, Deckert M, Ferreri AJM, Furtner J, Gallego Perez-Larraya J, Henriksson R, et al. European Association of Neuro-Oncology (EANO) guidelines for treatment of primary central nervous system lymphoma (PCNSL). *Neuro Oncol*. 2023;25:37–53.
85. Kasenda B, Ferreri AJM, Marturano E, Forst D, Bromberg J, Ghesquieres H, et al. First-line treatment and outcome of elderly patients with primary central nervous system lymphoma (PCNSL)--a systematic review and individual patient data meta-analysis. *Ann Oncol*. 2015;26:1305–13.
86. Abrey LE, Ben-Porat L, Panageas KS, Yahalom J, Berkey B, Curran W, et al. Primary central nervous system lymphoma: the Memorial Sloan-Kettering Cancer Center prognostic model. *J Clin Oncol*. 2006;24:5711–5.
87. Jang JE, Kim YR, Kim S-J, Cho H, Chung H, Lee JY, et al. A new prognostic model

using absolute lymphocyte count in patients with primary central nervous system lymphoma. *Eur J Cancer*. 2016;57:127–35.

88. Zeremski V, Adolph L, Beer S, Berisha M, Jacobs B, Kahl C, et al. Relevance of different prognostic scores in primary CNS lymphoma in the era of intensified treatment regimens: A retrospective, multicenter analysis of 174 patients. *Eur J Haematol*. 2024;

89. Thiel E, Korfel A, Martus P, Kanz L, Griesinger F, Rauch M, et al. High-dose methotrexate with or without whole brain radiotherapy for primary CNS lymphoma (G-PCNSL-SG-1): a phase 3, randomised, non-inferiority trial. *Lancet Oncol*. 2010;11:1036–47.

90. Liu J, Wang Y, Liu Y, Liu Z, Cui Q, Ji N, et al. Immunohistochemical profile and prognostic significance in primary central nervous system lymphoma: Analysis of 89 cases. *Oncol Lett*. 2017;14:5505–12.

91. Kreher S, Strehlow F, Martus P, Roth P, Hertenstein B, Röth A, et al. Prognostic impact of intraocular involvement in primary CNS lymphoma: experience from the G-PCNSL-SG1 trial. *Ann Hematol*. 2015;94:409–14.

92. Omuro A, Chinot O, Taillandier L, Ghesquieres H, Soussain C, Delwail V, et al. Methotrexate and temozolomide versus methotrexate, procarbazine, vincristine, and cytarabine for primary CNS lymphoma in an elderly population: an intergroup ANOCEF-GOELAMS randomised phase 2 trial. *Lancet Haematol*. 2015;2:e251-259.

93. van der Meulen M, Dirven L, Bakunina K, van den Bent MJ, Issa S, Doorduijn JK, et al. MMSE is an independent prognostic factor for survival in primary central nervous system lymphoma. *J Neurooncol*. 2021;152:357–62.

94. Sasayama T, Nakamizo S, Nishihara M, Kawamura A, Tanaka H, Mizukawa K, et al. Cerebrospinal fluid interleukin-10 is a potentially useful biomarker in immunocompetent primary central nervous system lymphoma (PCNSL). *Neuro Oncol*. 2012;14:368–80.

95. Ge L, Lu S, Xu L, Yan H. MYC, BCL2, and BCL6 expression as prognostic indicators in primary central nervous system lymphoma: A systematic review and meta-analysis. *Clin Neurol Neurosurg*. 2021;208:106838.

96. Mo SS, Cleveland J, Rubenstein JL. Primary CNS lymphoma: update on molecular pathogenesis and therapy. *Leuk Lymphoma*. 2023;64:57–65.

97. Zuo J, Lei T, Zhong S, Zhou J, Liu R, Wu C, et al. C-reactive protein levels, the prognostic nutritional index, and the lactate dehydrogenase-to-lymphocyte ratio are important prognostic factors in primary central nervous system lymphoma: a single-center study of 223 patients. *Neurosurg Rev*. 2023;47:17.

98. Wei L, Gao Y, Prochazka KT, Liu R, Wang L, Liu B, et al. A novel prognostic model based on pretreatment serum albumin and ECOG PS for primary CNS lymphoma: an international, multi-center study. *J Neurooncol* [Internet]. 2023 [cited 2023 Jun 2]; Available from: <https://link.springer.com/10.1007/s11060-023-04337-z>

99. Chung SR, Choi YJ, Kim HS, Park JE, Shim WH, Kim SJ. Tumor Vascular Permeability Pattern Is Associated With Complete Response in Immunocompetent Patients With Newly Diagnosed Primary Central Nervous System Lymphoma: Retrospective Cohort Study. *Medicine (Baltimore)*. 2016;95:e2624.

100. Yuan X-G, Huang Y-R, Yu T, Xu Y, Liang Y, Zhang X-H, et al. Primary central nervous system lymphoma in China: a single-center retrospective analysis of 167 cases. *Ann Hematol*. 2020;99:93–104.

101. Tabouret E, Houillier C, Martin-Duverneuil N, Blonski M, Soussain C, Ghesquière H, et al. Patterns of response and relapse in primary CNS lymphomas after first-line chemotherapy: imaging analysis of the ANOCEF-GOELAMS prospective randomized trial. *Neuro Oncol*. 2017;19:422–9.

102. Jahr G, Broi MD, Holte H, Beiske K, Meling TR. Evaluation of Memorial Sloan-Kettering Cancer Center and International Extranodal Lymphoma Study Group prognostic scoring systems to predict Overall Survival in intracranial Primary CNS lymphoma. *Brain and*

Behavior. 2018;8:e00928.

103. Wieduwilt MJ, Valles F, Issa S, Behler CM, Hwang J, McDermott M, et al. Immunochemotherapy with intensive consolidation for primary CNS lymphoma: a pilot study and prognostic assessment by diffusion-weighted MRI. *Clin Cancer Res.* 2012;18:1146–55.
104. Hatzoglou V, Oh JH, Buck O, Lin X, Lee M, Shukla-Dave A, et al. Pretreatment dynamic contrast-enhanced MRI biomarkers correlate with progression-free survival in primary central nervous system lymphoma. *J Neurooncol.* 2018;140:351–8.
105. Kawai N, Zhen H-N, Miyake K, Yamamoto Y, Nishiyama Y, Tamiya T. Prognostic value of pretreatment 18F-FDG PET in patients with primary central nervous system lymphoma: SUV-based assessment. *J Neurooncol.* 2010;100:225–32.
106. Godard F, Durot E, Durot C, Hoeffel C, Delmer A, Morland D. Cerebellum/liver index in pretherapeutic 18F-FDG PET/CT as a predictive marker of progression-free survival in follicular lymphoma treated by immunochemotherapy and rituximab maintenance. *Medicine.* 2022;101:e28791.
107. Morland D, Zizi G, Godard F, Gauchy A-C, Durot C, Hoeffel C, et al. 18F-FDG cerebellum/liver index as a prognostic factor for progression-free survival in diffuse large B-cell lymphoma. *Ann Nucl Med.* 2021;35:785–93.
108. Bodranghien F, Bastian A, Casali C, Hallett M, Louis ED, Manto M, et al. Consensus Paper: Revisiting the Symptoms and Signs of Cerebellar Syndrome. *Cerebellum.* 2016;15:369–91.
109. Buckner RL. The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. *Neuron.* 2013;80:807–15.
110. Manni E, Petrosini L. Luciani's work on the cerebellum a century later. *Trends Neurosci.* 1997;20:112–6.
111. Jahr G, Da Broi M, Holte H, Beiske K, Meling TR. The role of surgery in intracranial PCNSL. *Neurosurg Rev.* 2018;41:1037–44.
112. Houillier C, Ghesquière H, Chabrot C, Soussain C, Ahle G, Choquet S, et al. Rituximab, methotrexate, procarbazine, vincristine and intensified cytarabine consolidation for primary central nervous system lymphoma (PCNSL) in the elderly: a LOC network study. *J Neurooncol.* 2017;133:315–20.
113. Ferreri AJM, Cwynarski K, Pulczynski E, Fox CP, Schorb E, Celico C, et al. Long-term efficacy, safety and neurotoxicity of MATRix regimen followed by autologous transplant in primary CNS lymphoma: 7-year results of the IELSG32 randomized trial. *Leukemia.* 2022;36:1870–8.
114. Schorb E, Fox CP, Kasenda B, Linton K, Martinez-Calle N, Calimeri T, et al. Induction therapy with the MATRix regimen in patients with newly diagnosed primary diffuse large B-cell lymphoma of the central nervous system - an international study of feasibility and efficacy in routine clinical practice. *Br J Haematol.* 2020;189:879–87.
115. Calimeri T, Steffanoni S, Gagliardi F, Chiara A, Ferreri AJM. How we treat primary central nervous system lymphoma. *ESMO Open.* 2021;6:100213.
116. Abrey LE, Yahalom J, DeAngelis LM. Treatment for primary CNS lymphoma: the next step. *J Clin Oncol.* 2000;18:3144–50.
117. Gavrilovic IT, Hormigo A, Yahalom J, DeAngelis LM, Abrey LE. Long-term follow-up of high-dose methotrexate-based therapy with and without whole brain irradiation for newly diagnosed primary CNS lymphoma. *J Clin Oncol.* 2006;24:4570–4.
118. Shah GD, Yahalom J, Correa DD, Lai RK, Raizer JJ, Schiff D, et al. Combined immunochemotherapy with reduced whole-brain radiotherapy for newly diagnosed primary CNS lymphoma. *J Clin Oncol.* 2007;25:4730–5.
119. Morris PG, Correa DD, Yahalom J, Raizer JJ, Schiff D, Grant B, et al. Rituximab, methotrexate, procarbazine, and vincristine followed by consolidation reduced-dose whole-brain radiotherapy and cytarabine in newly diagnosed primary CNS lymphoma: final results

and long-term outcome. *J Clin Oncol*. 2013;31:3971–9.

120. Bromberg JEC, Issa S, Bakunina K, Minnema MC, Seute T, Durian M, et al. Rituximab in patients with primary CNS lymphoma (HOVON 105/ALLG NHL 24): a randomised, open-label, phase 3 intergroup study. *Lancet Oncol*. 2019;20:216–28.

121. Schmitt AM, Herbrand AK, Fox CP, Bakunina K, Bromberg JEC, Cwynarski K, et al. Rituximab in primary central nervous system lymphoma-A systematic review and meta-analysis. *Hematol Oncol*. 2019;37:548–57.

122. Zhang Y, Liu Z, Gao C, Bian H, Ma Y, Jing F, et al. Role of Rituximab in Treatment of Patients With Primary Central Nervous System Lymphoma (PCNSL): A Systematic Review and Meta-Analysis. *Clin Lymphoma Myeloma Leuk*. 2023;23:733–41.

123. Volpini ME, Song J, Samant R, MacDonald D, Nair VJ. Cranial Radiation Therapy as Salvage in the Treatment of Relapsed Primary CNS Lymphoma. *Curr Oncol*. 2022;29:8160–70.

124. Thomas A, Houillier C, Antoni D, Hoang-Xuan K, Soussain C, Jacob J, et al. Radiotherapy for newly diagnosed primary central nervous system lymphoma: role and perspective. *Rep Pract Oncol Radiother*. 2023;28:271–85.

125. Lesueur P, Damaj G, Hoang-Xuan K, Roland V, Schmitt A, Chinot O, et al. Reduced-dose WBRT as consolidation treatment for patients with primary CNS lymphoma: an LOC network study. *Blood Adv*. 2022;6:4807–15.

126. Houillier C, Taillandier L, Dureau S, Lamy T, Laadhari M, Chinot O, et al. Radiotherapy or Autologous Stem-Cell Transplantation for Primary CNS Lymphoma in Patients 60 Years of Age and Younger: Results of the Intergroup ANOCEF-GOELAMS Randomized Phase II PRECIS Study. *J Clin Oncol*. 2019;37:823–33.

127. Ferreri AJM, Cwynarski K, Pulczynski E, Fox CP, Schorb E, La Rosée P, et al. Whole-brain radiotherapy or autologous stem-cell transplantation as consolidation strategies after high-dose methotrexate-based chemoimmunotherapy in patients with primary CNS lymphoma: results of the second randomisation of the International Extranodal Lymphoma Study Group-32 phase 2 trial. *Lancet Haematol*. 2017;4:e510–23.

128. Hoang-Xuan K, Houot R, Soussain C, Blonski M, Schmitt A, Delwail V, et al. First Results of the Acsé Pembrolizumab Phase II in the Primary CNS Lymphoma (PCNSL) Cohort. *Blood*. 2020;136:15–6.

129. Rachdi A, Hernandez-Tost H, Herzi D, Morales-Martinez A, Hernández-Verdin I, Houillier C, et al. Recent advances in the diagnosis and the treatment of primary CNS lymphoma. *Rev Neurol (Paris)*. 2023;179:481–9.

130. Schaff L, Nayak L, Grommes C. Bruton’s tyrosine kinase (BTK) inhibitors for the treatment of primary central nervous system lymphoma (PCNSL): current progress and latest advances. *Leuk Lymphoma*. 2024;1–13.

131. Zou R, Zhou X, Liu H, Wang P, Xia F, Kang L, et al. Long-term Complete Remission of Decitabine-Primed Tandem CD19/CD22 CAR-T Therapy with PD-1 and BTK Inhibitors Maintenance in a Refractory Primary Central Nervous System Lymphoma Patient. *Cancer Res Treat*. 2023;55:1363–8.

132. Cook MR, Dorris CS, Makambi KH, Luo Y, Munshi PN, Donato M, et al. Toxicity and efficacy of CAR T-cell therapy in primary and secondary CNS lymphoma: a meta-analysis of 128 patients. *Blood Adv*. 2023;7:32–9.

133. Choquet S, Soussain C, Azar N, Morel V, Metz C, Ursu R, et al. CAR T-cell therapy induces a high rate of prolonged remission in relapsed primary CNS lymphoma: Real-life results of the LOC network. *Am J Hematol*. 2024;

134. Lin X, Khan IRA, Seet YHC, Lee HY, Yu W-Y. Atypical radiological findings of primary central nervous system lymphoma. *Neuroradiology*. 2020;62:669–76.

135. Estephan F, Ye X, Dzaye O, Wagner-Johnston N, Swinnen L, Gladstone DE, et al. White matter changes in primary central nervous system lymphoma patients treated with

- high-dose methotrexate with or without rituximab. *J Neurooncol.* 2019;145:461–6.
136. Cornell I, Al Busaidi A, Wastling S, Anjari M, Cwynarski K, Fox CP, et al. Early MRI Predictors of Relapse in Primary Central Nervous System Lymphoma Treated with MATRix Immunochemotherapy. *J Pers Med.* 2023;13:1182.
137. Zhang Y, Zou D, Yin J, Zhang L, Zhang X, Wang W, et al. Changes in cerebrospinal fluid interleukin-10 levels display better performance in predicting disease relapse than conventional magnetic resonance imaging in primary central nervous system lymphoma. *BMC Cancer.* 2021;21:183.
138. Mercadal S, Cortés-Romera M, Vélez P, Climent F, Gámez C, González-Barca E. [Positron emission tomography combined with computed tomography in the initial evaluation and response assessment in primary central nervous system lymphoma]. *Med Clin (Barc).* 2015;144:503–6.
139. Meignan M, Gallamini A, Meignan M, Gallamini A, Haioun C. Report on the First International Workshop on Interim-PET-Scan in Lymphoma. *Leuk Lymphoma.* 2009;50:1257–60.
140. Gallamini A, Hutchings M, Rigacci L, Specht L, Merli F, Hansen M, et al. Early interim 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. *J Clin Oncol.* 2007;25:3746–52.
141. Itti E, Meignan M, Berriolo-Riedinger A, Biggi A, Cashen AF, Véra P, et al. An international confirmatory study of the prognostic value of early PET/CT in diffuse large B-cell lymphoma: comparison between Deauville criteria and Δ SUVmax. *Eur J Nucl Med Mol Imaging.* 2013;40:1312–20.
142. Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Müller SP, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J Clin Oncol.* 2014;32:3048–58.
143. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol.* 2014;32:3059–68.
144. Burggraaff CN, Eertink JJ, Lugtenburg PJ, Hoekstra OS, Arens AIJ, de Keizer B, et al. 18F-FDG PET Improves Baseline Clinical Predictors of Response in Diffuse Large B-Cell Lymphoma: The HOVON-84 Study. *J Nucl Med.* 2022;63:1001–7.
145. Allen SG, Rosen BS, Aryal M, Cao Y, Schipper MJ, Wong KK, et al. Initial Feasibility and Acute Toxicity Outcomes From a Phase 2 Trial of 18F-Fluorodeoxyglucose Positron Emission Tomography Response-Based De-escalated Definitive Chemoradiotherapy for p16+ Oropharynx Cancer: A Planned Interim Analysis. *Int J Radiat Oncol Biol Phys.* 2023;117:171–80.
146. Press OW, Li H, Schöder H, Straus DJ, Moskowitz CH, LeBlanc M, et al. US Intergroup Trial of Response-Adapted Therapy for Stage III to IV Hodgkin Lymphoma Using Early Interim Fluorodeoxyglucose-Positron Emission Tomography Imaging: Southwest Oncology Group S0816. *J Clin Oncol.* 2016;34:2020–7.
147. Birsén R, Blanc E, Willems L, Burroni B, Legoff M, Le Ray E, et al. Prognostic value of early 18F-FDG PET scanning evaluation in immunocompetent primary CNS lymphoma patients. *Oncotarget.* 2018;9:16822–31.
148. Maza S, Buchert R, Brenner W, Munz DL, Thiel E, Korfel A, et al. Brain and whole-body FDG-PET in diagnosis, treatment monitoring and long-term follow-up of primary CNS lymphoma. *Radiol Oncol.* 2013;47:103–10.
149. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48:441–6.

150. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016;278:563–77.
151. Hatt M, Krizsan AK, Rahmim A, Bradshaw TJ, Costa PF, Forgacs A, et al. Joint EANM/SNMMI guideline on radiomics in nuclear medicine : Jointly supported by the EANM Physics Committee and the SNMMI Physics, Instrumentation and Data Sciences Council. *Eur J Nucl Med Mol Imaging*. 2023;50:352–75.
152. Yang F, Wang Y, Li Q, Cao L, Sun Z, Jin J, et al. Intratumor heterogeneity predicts metastasis of triple-negative breast cancer. *Carcinogenesis*. 2017;38:900–9.
153. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*. 2013;501:338–45.
154. Liu J, Dang H, Wang XW. The significance of intertumor and intratumor heterogeneity in liver cancer. *Exp Mol Med*. 2018;50:e416.
155. Morris LGT, Riaz N, Desrichard A, Şenbabaoğlu Y, Hakimi AA, Makarov V, et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget*. 2016;7:10051–63.
156. Gong X-Q, Tao Y-Y, Wu Y, Liu N, Yu X, Wang R, et al. Progress of MRI Radiomics in Hepatocellular Carcinoma. *Front Oncol*. 2021;11:698373.
157. Delli Pizzi A, Chiarelli AM, Chiacchiaretta P, d’Annibale M, Croce P, Rosa C, et al. MRI-based clinical-radiomics model predicts tumor response before treatment in locally advanced rectal cancer. *Sci Rep*. 2021;11:5379.
158. Lohmann P, Galldiks N, Kocher M, Heinzl A, Filss CP, Stegmayr C, et al. Radiomics in neuro-oncology: Basics, workflow, and applications. *Methods*. 2021;188:112–21.
159. Lohmann P, Franceschi E, Vollmuth P, Dhermain F, Weller M, Preusser M, et al. Radiomics in neuro-oncological clinical trials. *Lancet Digit Health*. 2022;4:e841–9.
160. Zhu M, Li S, Kuang Y, Hill VB, Heimberger AB, Zhai L, et al. Artificial intelligence in the radiomic analysis of glioblastomas: A review, taxonomy, and perspective. *Front Oncol*. 2022;12:924245.
161. Zhou H, Chang K, Bai HX, Xiao B, Su C, Bi WL, et al. Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas. *J Neurooncol*. 2019;142:299–307.
162. Lohmann P, Lerche C, Bauer EK, Steger J, Stoffels G, Blau T, et al. Predicting IDH genotype in gliomas using FET PET radiomics. *Sci Rep*. 2018;8:13328.
163. Peng L, Parekh V, Huang P, Lin DD, Sheikh K, Baker B, et al. Distinguishing True Progression From Radionecrosis After Stereotactic Radiation Therapy for Brain Metastases With Machine Learning and Radiomics. *Int J Radiat Oncol Biol Phys*. 2018;102:1236–43.
164. Wiltgen T, Fleischmann DF, Kaiser L, Holzgreve A, Corradini S, Landry G, et al. 18F-FET PET radiomics-based survival prediction in glioblastoma patients receiving radio(chemo)therapy. *Radiat Oncol*. 2022;17:198.
165. Li G, Li L, Li Y, Qian Z, Wu F, He Y, et al. An MRI radiomics approach to predict survival and tumour-infiltrating macrophages in gliomas. *Brain*. 2022;145:1151–61.
166. Garaba A, Aslam N, Ponzio F, Panciani PP, Brinjikji W, Fontanella M, et al. Radiomics for differentiation of gliomas from primary central nervous system lymphomas: a systematic review and meta-analysis. *Front Oncol*. 2024;14:1291861.
167. Chen C, Zheng A, Ou X, Wang J, Ma X. Comparison of Radiomics-Based Machine-Learning Classifiers in Diagnosis of Glioblastoma From Primary Central Nervous System Lymphoma. *Front Oncol*. 2020;10:1151.
168. She Z, Marzullo A, Destito M, Spadea MF, Leone R, Anzalone N, et al. Deep learning-based overall survival prediction model in patients with rare cancer: a case study for primary central nervous system lymphoma. *Int J Comput Assist Radiol Surg*. 2023;18:1849–56.
169. Destito M, Marzullo A, Leone R, Zaffino P, Steffanoni S, Erbella F, et al. Radiomics-Based Machine Learning Model for Predicting Overall and Progression-Free Survival in Rare

- Cancer: A Case Study for Primary CNS Lymphoma Patients. *Bioengineering* (Basel). 2023;10:285.
170. Wang X, Zhao L, Wang S, Zhao X, Chen L, Sun X, et al. Utility of contrast-enhanced MRI radiomics features combined with clinical indicators for predicting induction chemotherapy response in primary central nervous system lymphoma. *J Neurooncol*. 2024;166:451–60.
171. Liu Y, Chen P-HC, Krause J, Peng L. How to Read Articles That Use Machine Learning: Users' Guides to the Medical Literature. *JAMA*. 2019;322:1806.
172. Pfaehler E, Beukinga RJ, de Jong JR, Slart RHJA, Slump CH, Dierckx RAJO, et al. Repeatability of 18 F-FDG PET radiomic features: A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method. *Med Phys*. 2019;46:665–78.
173. Kocak B, Akinci D'Antonoli T, Mercaldo N, Alberich-Bayarri A, Baessler B, Ambrosini I, et al. METHodological RadiomICs Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMII. *Insights Imaging*. 2024;15:8.
174. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020;295:328–38.
175. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14:749–62.
176. Vovk U, Pernus F, Likar B. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Trans Med Imaging*. 2007;26:405–21.
177. Nguyen AA-T, Onishi N, Carmona-Bozo J, Li W, Kornak J, Newitt DC, et al. Post-Processing Bias Field Inhomogeneity Correction for Assessing Background Parenchymal Enhancement on Breast MRI as a Quantitative Marker of Treatment Response. *Tomography*. 2022;8:891–904.
178. Nyúl LG, Udupa JK. On standardizing the MR image intensity scale. *Magn Reson Med*. 1999;42:1072–81.
179. Shah M, Xiao Y, Subbanna N, Francis S, Arnold DL, Collins DL, et al. Evaluating intensity normalization on MRIs of human brain with multiple sclerosis. *Med Image Anal*. 2011;15:267–82.
180. Fortin J-P, Parker D, Tunç B, Watanabe T, Elliott MA, Ruparel K, et al. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage*. 2017;161:149–70.
181. Orhac F, Lecler A, Savatovski J, Goya-Outi J, Nioche C, Charbonneau F, et al. How can we combat multicenter variability in MR radiomics? Validation of a correction procedure. *Eur Radiol*. 2021;31:2272–80.
182. Orhac F, Eertink JJ, Cottreau A-S, Zijlstra JM, Thieblemont C, Meignan M, et al. A Guide to ComBat Harmonization of Imaging Biomarkers in Multicenter Studies. *J Nucl Med*. 2022;63:172–9.
183. Orhac F, Boughdad S, Philippe C, Stalla-Bourdillon H, Nioche C, Champion L, et al. A Postreconstruction Harmonization Method for Multicenter Radiomic Studies in PET. *J Nucl Med*. 2018;59:1321–8.
184. Stein CK, Qu P, Epstein J, Buros A, Rosenthal A, Crowley J, et al. Removing batch effects from purified plasma cell gene expression microarrays with modified ComBat. *BMC Bioinformatics*. 2015;16:63.
185. Soltaninejad M, Yang G, Lambrou T, Allinson N, Jones TL, Barrick TR, et al. Supervised learning based multimodal MRI brain tumour segmentation using texture features from supervoxels. *Comput Methods Programs Biomed*. 2018;157:69–84.
186. Deng W, Shi Q, Luo K, Yang Y, Ning N. Brain Tumor Segmentation Based on Improved Convolutional Neural Network in Combination with Non-quantifiable Local Texture Feature. *J Med Syst*. 2019;43:152.

187. Prasanna P, Karnawat A, Ismail M, Madabhushi A, Tiwari P. Radiomics-based convolutional neural network for brain tumor segmentation on multiparametric magnetic resonance imaging. *J Med Imaging (Bellingham)*. 2019;6:024005.
188. van der Voort SR, Incekara F, Wijnenga MMJ, Kapsas G, Gahrman R, Schouten JW, et al. Combined molecular subtyping, grading, and segmentation of glioma using multi-task deep learning. *Neuro Oncol*. 2023;25:279–89.
189. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEEx: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res*. 2018;78:4786–9.
190. Thévenaz P, Blu T, Unser M. Image Interpolation and Resampling. *Handbook of Medical Imaging* [Internet]. Elsevier; 2000 [cited 2024 Apr 12]. p. 393–420. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780120777907500308>
191. Lekien F, Marsden J. Tricubic interpolation in three dimensions. *Int J Numer Meth Engng*. 2005;63:455–71.
192. Orlhac F, Nioche C, Klyuzhin I, Rahmim A, Buvat I. Radiomics in PET Imaging: A Practical Guide for Newcomers. *PET Clin*. 2021;16:597–612.
193. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res*. 2017;77:e104–7.
194. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to Radiomics. *J Nucl Med*. 2020;61:488–95.
195. Reuzé S, Schernberg A, Orlhac F, Sun R, Chargari C, Dercle L, et al. Radiomics in Nuclear Medicine Applied to Radiation Therapy: Methods, Pitfalls, and Challenges. *Int J Radiat Oncol Biol Phys*. 2018;102:1117–42.
196. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. *IEEE Trans Syst, Man, Cybern*. 1973;SMC-3:610–21.
197. Guyon I, Elisseeff A. An Introduction to Feature Extraction. In: Guyon I, Nikravesh M, Gunn S, Zadeh LA, editors. *Feature Extraction* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006 [cited 2024 Apr 16]. p. 1–25. Available from: http://link.springer.com/10.1007/978-3-540-35488-8_1
198. Jafari P, Azuaje F. An assessment of recently published gene expression data analyses: reporting experimental design and statistical factors. *BMC Med Inform Decis Mak*. 2006;6:27.
199. Thomas JG, Olson JM, Tapscott SJ, Zhao LP. An Efficient and Robust Statistical Modeling Approach to Discover Differentially Expressed Genes Using Genomic Expression Profiles. *Genome Res*. 2001;11:1227–36.
200. Kohavi R, John GH. Wrappers for feature subset selection. *Artificial Intelligence*. 1997;97:273–324.
201. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1996;58:267–88.
202. Zou H, Hastie T. Regularization and Variable Selection Via the Elastic Net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 2005;67:301–20.
203. Balagurunathan Y, Kumar V, Gu Y, Kim J, Wang H, Liu Y, et al. Test-retest reproducibility analysis of lung CT image features. *J Digit Imaging*. 2014;27:805–23.
204. Vuong D, Tanadini-Lang S, Huellner MW, Veit-Haibach P, Unkelbach J, Andratschke N, et al. Interchangeability of radiomic features between [18F]- FDG PET / CT and [18F]- FDG PET / MR. *Medical Physics*. 2019;46:1677–85.
205. Roman-Jimenez G, Acosta O, Leseur J, Devillers A, Der Sarkissian H, Guzman L, et al. Random forests to predict tumor recurrence following cervical cancer therapy using pre- and per-treatment 18F-FDG PET parameters. *Annu Int Conf IEEE Eng Med Biol Soc*. 2016;2016:2444–7.

206. Desbordes P, Ruan S, Modzelewski R, Pineau P, Vauclin S, Gouel P, et al. Predictive value of initial FDG-PET features for treatment response and survival in esophageal cancer patients treated with chemo-radiation therapy using a random forest classifier. Ahmad A, editor. *PLoS ONE*. 2017;12:e0173208.
207. Gao X, Chu C, Li Y, Lu P, Wang W, Liu W, et al. The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from (18)F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur J Radiol*. 2015;84:312–7.
208. Schapire RE. Explaining AdaBoost. In: Schölkopf B, Luo Z, Vovk V, editors. *Empirical Inference* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013 [cited 2024 Apr 12]. p. 37–52. Available from: https://link.springer.com/10.1007/978-3-642-41136-6_5
209. Ge G, Zhang J. Feature selection methods and predictive models in CT lung cancer radiomics. *J Appl Clin Med Phys*. 2023;24:e13869.
210. Turing AM. I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*. 1950;LIX:433–60.
211. Sarker IH. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput Sci*. 2021;2:420.
212. Nensa F, Demircioglu A, Rischpler C. Artificial Intelligence in Nuclear Medicine. *J Nucl Med*. 2019;60:29S-37S.
213. Visvikis D, Lambin P, Beuschau Mauridsen K, Hustinx R, Lassmann M, Rischpler C, et al. Application of artificial intelligence in nuclear medicine and molecular imaging: a review of current status and future perspectives for clinical translation. *Eur J Nucl Med Mol Imaging*. 2022;49:4452–63.
214. Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiology: Artificial Intelligence*. 2020;2:e200029.
215. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med*. 2020;26:1320–4.
216. Lekadir K, Feragen A, Fofanah AJ, Frangi AF, Buyx A, Emelie A, et al. FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. 2023 [cited 2024 Apr 12]; Available from: <https://arxiv.org/abs/2309.12325>
217. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998;86:2278–324.
218. Gong K, Johnson K, El Fakhri G, Li Q, Pan T. PET image denoising based on denoising diffusion probabilistic model. *Eur J Nucl Med Mol Imaging*. 2024;51:358–68.
219. Gong K, Han P, El Fakhri G, Ma C, Li Q. Arterial spin labeling MR image denoising and reconstruction using unsupervised deep learning. *NMR Biomed*. 2022;35:e4224.
220. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* [Internet]. Cham: Springer International Publishing; 2015 [cited 2024 Apr 12]. p. 234–41. Available from: http://link.springer.com/10.1007/978-3-319-24574-4_28
221. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. 2021 [cited 2024 Apr 12]; Available from: <https://arxiv.org/abs/2103.14030>
222. Cheng D, Liu M, Fu J, Wang Y. Classification of MR brain images by combination of multi-CNNs for AD diagnosis. In: Falco CM, Jiang X, editors. Hong Kong, China; 2017 [cited 2024 Apr 8]. p. 1042042. Available from: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2281808>
223. Blanc-Durand P, Jégou S, Kanoun S, Berriolo-Riedinger A, Bodet-Milin C, Kraeber-

- Bodéré 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;48:1362–70.
224. Marin T, Zhuo Y, Lahoud RM, Tian F, Ma X, Xing F, et al. Deep learning-based GTV contouring modeling inter- and intra- observer variability in sarcomas. *Radiother Oncol*. 2022;167:269–76.
225. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature*. 2013;501:346–54.
226. Lacerda P, Barros B, Albuquerque C, Conci A. Hyperparameter Optimization for COVID-19 Pneumonia Diagnosis Based on Chest CT. *Sensors (Basel)*. 2021;21:2174.
227. Reyes M, Meier R, Pereira S, Silva CA, Dahlweid F-M, Tengg-Kobligk HV, et al. On the Interpretability of Artificial Intelligence in Radiology: Challenges and Opportunities. *Radiology: Artificial Intelligence*. 2020;2:e190043.
228. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. 2017 IEEE International Conference on Computer Vision (ICCV) [Internet]. Venice: IEEE; 2017 [cited 2024 Apr 10]. p. 618–26. Available from: <http://ieeexplore.ieee.org/document/8237336/>
229. Zhang H, Ogasawara K. Grad-CAM-Based Explainable Artificial Intelligence Related to Medical Text Processing. *Bioengineering*. 2023;10:1070.
230. Lauw MIS, Lucas C-HG, Ohgami RS, Wen KW. Primary Central Nervous System Lymphomas: A Diagnostic Overview of Key Histomorphologic, Immunophenotypic, and Genetic Features. *Diagnostics*. 2020;10:1076.
231. Morales-Martinez A, Lozano-Sanchez F, Duran-Peña A, Hoang-Xuan K, Houillier C. Primary Central Nervous System Lymphoma in Elderly Patients: Management and Perspectives. *Cancers (Basel)*. 2021;13.
232. Fu F, Sun X, Li Y, Liu Y, Shan Y, Ji N, et al. Dynamic contrast-enhanced magnetic resonance imaging biomarkers predict chemotherapeutic responses and survival in primary central-nervous-system lymphoma. *Eur Radiol*. 2021;31:1863–71.
233. Valles FE, Perez-Valles CL, Regalado S, Barajas RF, Rubenstein JL, Cha S. Combined diffusion and perfusion MR imaging as biomarkers of prognosis in immunocompetent patients with primary central nervous system lymphoma. *AJNR Am J Neuroradiol*. 2013;34:35–40.
234. Barajas RF, Rubenstein JL, Chang JS, Hwang J, Cha S. Diffusion-weighted MR imaging derived apparent diffusion coefficient is predictive of clinical outcome in primary central nervous system lymphoma. *AJNR Am J Neuroradiol*. 2010;31:60–6.
235. Chong I, Ostrom Q, Khan B, Dandachi D, Garg N, Kotrotsou A, et al. Whole Tumor Histogram Analysis Using DW MRI in Primary Central Nervous System Lymphoma Correlates with Tumor Biomarkers and Outcome. *Cancers (Basel)*. 2019;11:1506.
236. Baek DW, Cho HJ, Bae JH, Sohn SK, Moon JH. Apparent diffusion coefficient as a valuable quantitative parameter for predicting clinical outcomes in patients with newly diagnosed primary CNS lymphoma. *Blood Res*. 2020;55:99–106.
237. Huang W-Y, Wen J-B, Wu G, Yin B, Li J-J, Geng D-Y. Diffusion-Weighted Imaging for Predicting and Monitoring Primary Central Nervous System Lymphoma Treatment Response. *AJNR Am J Neuroradiol*. 2016;37:2010–8.
238. Zhang Y, Zhang Q, Wang X-X, Deng X-F, Zhu Y-Z. Value of pretherapeutic DWI in evaluating prognosis and therapeutic effect in immunocompetent patients with primary central nervous system lymphoma given high-dose methotrexate-based chemotherapy: ADC-based assessment. *Clin Radiol*. 2016;71:1018–29.
239. Blasel S, Vorwerk R, Kiyose M, Mittelbronn M, Brunnberg U, Ackermann H, et al. New MR perfusion features in primary central nervous system lymphomas: pattern and prognostic impact. *J Neurol*. 2018;265:647–58.

240. Luo X, Zhu Y, Zhang Y, Zhang Q, Wang X, Deng X. Parameters of MR perfusion-weighted imaging predict the response and prognosis to high-dose methotrexate-based chemotherapy in immunocompetent patients with primary central nervous system lymphoma. *J Clin Neurosci*. 2022;95:151–8.
241. Gupta T, Manjali JJ, Kannan S, Purandare N, Rangarajan V. Diagnostic Performance of Pretreatment 18F-Fluorodeoxyglucose Positron Emission Tomography With or Without Computed Tomography in Patients With Primary Central Nervous System Lymphoma: Updated Systematic Review and Diagnostic Test Accuracy Meta-analyses. *Clin Lymphoma Myeloma Leuk*. 2021;21:497–507.
242. Rozenblum L, Galanaud D, Houillier C, Soussain C, Baptiste A, Belin L, et al. [18F]FDG PET-MRI provides survival biomarkers in primary central nervous system lymphoma in the elderly: an ancillary study from the BLOCAGE trial of the LOC network. *Eur J Nucl Med Mol Imaging*. 2023;
243. Acharya UR, Hagiwara Y, Sudarshan VK, Chan WY, Ng KH. Towards precision medicine: from quantitative imaging to radiomics. *J Zhejiang Univ Sci B*. 2018;19:6–24.
244. Cao X, Tan D, Liu Z, Liao M, Kan Y, Yao R, et al. Differentiating solitary brain metastases from glioblastoma by radiomics features derived from MRI and 18F-FDG-PET and the combined application of multiple models. *Sci Rep*. 2022;12:5722.
245. Guha A, Goda JS, Dasgupta A, Mahajan A, Halder S, Gawde J, et al. Classifying primary central nervous system lymphoma from glioblastoma using deep learning and radiomics based machine learning approach - a systematic review and meta-analysis. *Front Oncol*. 2022;12:884173.
246. Kong Z, Jiang C, Zhu R, Feng S, Wang Y, Li J, et al. 18F-FDG-PET-based radiomics features to distinguish primary central nervous system lymphoma from glioblastoma. *Neuroimage Clin*. 2019;23:101912.
247. Chen C, Zhuo H, Wei X, Ma X. Contrast-Enhanced MRI Texture Parameters as Potential Prognostic Factors for Primary Central Nervous System Lymphoma Patients Receiving High-Dose Methotrexate-Based Chemotherapy. *Contrast Media Mol Imaging*. 2019;2019:5481491.
248. Li B, Xu Y, Wang Y, Zhang B. DECTNet: Dual Encoder Network combined convolution and Transformer architecture for medical image segmentation. *PLoS One*. 2024;19:e0301019.
249. Song B, Kc DR, Yang RY, Li S, Zhang C, Liang R. Classification of Mobile-Based Oral Cancer Images Using the Vision Transformer and the Swin Transformer. *Cancers (Basel)*. 2024;16:987.
250. Davnall F, Yip CSP, Ljungqvist G, Selmi M, Ng F, Sanghera B, et al. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights Imaging*. 2012;3:573–89.
251. Ahn S-Y, Kwon SY, Jung S-H, Ahn J-S, Yoo SW, Min J-J, et al. Prognostic Significance of Interim 11C-Methionine PET/CT in Primary Central Nervous System Lymphoma. *Clin Nucl Med*. 2018;43:e259–64.
252. Jang SJ, Lee K-H, Lee JY, Choi JY, Kim B-T, Kim SJ, et al. (11)C-methionine PET/CT and MRI of primary central nervous system diffuse large B-cell lymphoma before and after high-dose methotrexate. *Clin Nucl Med*. 2012;37:e241-244.
253. Miyakita Y, Ohno M, Takahashi M, Kurihara H, Katai H, Narita Y. Usefulness of carbon-11-labeled methionine positron-emission tomography for assessing the treatment response of primary central nervous system lymphoma. *Japanese Journal of Clinical Oncology*. 2020;50:512–8.
254. Kawase Y, Yamamoto Y, Kameyama R, Kawai N, Kudomi N, Nishiyama Y. Comparison of 11C-methionine PET and 18F-FDG PET in patients with primary central nervous system lymphoma. *Mol Imaging Biol*. 2011;13:1284–9.

255. Norikane T, Mitamura K, Yamamoto Y, Manabe Y, Murao M, Arai-Okuda H, et al. Comparative evaluation of ¹¹C-methionine and ¹⁸F-fluorodeoxyglucose positron emission tomography for distinguishing between primary central nervous system lymphoma and isocitrate dehydrogenase-wildtype glioblastoma. *J Neurooncol.* 2024;166:195–201.
256. Lenz G, Hiddemann W, Dreyling M. The role of fludarabine in the treatment of follicular and mantle cell lymphoma. *Cancer.* 2004;101:883–93.
257. Pastor-Anglada M, Pérez-Torras S. Nucleoside transporter proteins as biomarkers of drug responsiveness and drug targets. *Front Pharmacol.* 2015;6:13.
258. Barrueco JR, Jacobsen DM, Chang CH, Brockman RW, Sirotinak FM. Proposed mechanism of therapeutic selectivity for 9-beta-D-arabinofuranosyl-2-fluoroadenine against murine leukemia based upon lower capacities for transport and phosphorylation in proliferative intestinal epithelium compared to tumor cells. *Cancer Res.* 1987;47:700–6.

Annexe 1. Echelle de Karnofsky

Définition	%	Critères
Capable de mener une activité normale et de travailler ; pas besoin de soins particuliers	100	Normal ; pas de plaintes ; pas d'évidence de maladie
	90	Capable d'une activité normale ; signes ou symptômes mineurs en relation avec la maladie
	80	Activité normale avec effort ; signes ou symptômes de la maladie
Incapable de travailler ; capable de vivre à domicile et de subvenir à la plupart de ses besoins	70	Capacité de subvenir à ses besoins ; incapable d'avoir une activité normale et professionnelle active
	60	Requiert une assistance occasionnelle mais est capable de subvenir à la plupart de ses besoins
	50	Requiert une assistance et des soins médicaux fréquents
Incapable de subvenir à ses besoins ; requiert un équivalent de soins institutionnels ou hospitaliers	40	Invalide ; requiert des soins et une assistance importants
	30	Sévèrement invalide ; hospitalisation indiquée bien que le décès ne soit pas imminent
	20	Extrêmement malade ; hospitalisation nécessaire ; traitement actif de soutien nécessaire
	10	Mourant ; mort imminente
	0	Décédé

Karnofsky D. The clinical evaluation of chemotherapeutic agents in cancer. Columbia University Press, New-York 1949 :191-205

Annexe 2. Résumé du protocole LOCALYSE

LOCALYSE Study (NCT03582254) ancillary of BLOCAGE01 (NCT02313389)

Selection of patients

Inclusion Criteria:

- Newly diagnosed primary cerebral lymphoma
- Age ≥ 60 years
- Pathology proven diagnosis or positive cytology of the Cerebrospinal Fluid or vitreous
- Karnofsky Performance Status ≥ 40
- No evidence of systemic Non-Hodgkin lymphomas (body CT scan, bone marrow biopsy)
- Adequate haematological, renal and hepatic function
- Calculated creatinine clearance > 40 ml/min

Exclusion Criteria:

- Uncontrolled diabetes with fasting glycaemia > 200 mg/dL
- Sensitivity to active substance in ^{18}F -FDG
- Calculated creatinine clearance < 40 ml/min
- No contraindication to MRI (magnetic resonance imaging)

Schema and treatment plan

Treatment (related to BLOCAGE01)

1/ Induction Chemotherapy:

4 cycles of 28 days of R-MPV: Rituximab I.V. 375 mg/m^2 : Day 1 (D1) Methotrexate (MTX) I.V. 3.5 g/m^2 : D2 and D15, Procarbazine P.O. 100 mg/m^2 : D2 to D8, Vincristine I.V. 1.4 mg/m^2 : D2 and D15, Methylprednisolone at least at the dose of 60 mg/day from D2 to D6

**(Rituximab and other molecules can be administered on the same day if an 8-hour interval between injections is observed)*

Addition of a 5th cycle of R-MTX for patients in partial response only after 4 cycles of R-MPV: Rituximab I.V. 375 mg/m^2 : D1, MTX iv 3.5 g/m^2 : D2 and D15

Consolidation with R-cytarabine (A): 1 cycle Cytarabine I.V. $3 \text{ g/m}^2/\text{day}$ for 2 days, starting 2 weeks after the last infusion of MTX: D1 and D2 Rituximab IV 375 mg/m^2 : D1 with prophylaxis by Pegfilgrastim after the Cytarabine injection

2/ Evaluation of the response to induction chemotherapy according to the IPCG criteria

3/ Randomization only for patients in complete response

Arm 1: Monitoring

Arm 2: Maintenance chemotherapy (R-MT Protocol)

R-MT for 7 cycles of 28 days: Rituximab IV 375 mg/m²: Day 1 (D1) for cycles 1 and 4

MTX IV 3.5 g/m²: Day 2 (D2)*, Temozolomide P.O. 150 mg/m²: Day 2 (D2)* to Day 6 (D6)

**Chemotherapy on D2 applies only to cycles 1 and 4 (C1 and C4) since Rituximab is administered on D1*

Images (related to LOCALYSE)

The examination will be carried out on a Signa 3T GE Healthcare hybrid PET-MRI camera, located in the Nuclear Medicine Department of the Pitié-Salpêtrière Hospital Group.

Patients will be asked to fast for 4-6 hours before the examination. The FDG injection will be performed if the capillary blood glucose level, checked prior to the injection, is less than or equal to 2 g/l. A dose of 2 MBq/Kg (minimum: 125 MBq – maximum: 250 MBq) will be injected directly intravenously. Participants will be left to rest in a reclining position in a quiet, unlit room for at least 80 minutes post-injection. The prepared activity, measurement time, injection time, injected volume, and residual activity will be recorded to determine the actual injected activity. A cerebral acquisition in list mode will be performed on the hybrid PET-MRI camera. It will begin 90 ± 10 minutes after injection and last 30 minutes. The images will be reconstructed and corrected for physical phenomena according to the manufacturer's recommendations. They will be expressed in Standard Uptake Value (SUV).

18F-FDG PET-MRI will be performed :

- prior to initiation of chemotherapy (PET1),
- after two chemotherapy cycles (PET2),
- at the end of the first-line chemotherapy regimens (PET3).

Rules for dose modification

N/A in LOCALYSE

In BLOCAGE01 :

- In the case of hematological toxicity of grade ≥ 3 , the dose of chemotherapeutic agents would be reduced by 30%.
- In the case of acute kidney failure of grade ≥ 2 , resumption of MTX can only be proposed if it returns to normal and with a reduced dose of 50%. If the reduced-dose treatment is well tolerated on a renal basis, a resumption of MTX at full dose may be considered in subsequent courses.
- In the event of the appearance of a progressive multifocal leukoencephalopathy, or severe skin reactions such as Lyell's syndrome and Stevens-Johnson syndrome, treatment with Rituximab would be permanently stopped.

Measurement of treatment effect including response criteria, definitions of response and survival, and methods of measurement

End of treatment assessment according to the International PCNSL Collaborative

Group criteria :

- Evaluation of Cerebral Lesions

The lesion volume (component enhanced by the injection of contrast medium) will be assessed by measuring the three largest orthogonal diameters on brain MRI.

- Ophthalmological Evaluation

The ophthalmological examination will systematically include a fundoscopic examination and a slit-lamp examination, as well as fluorescein angiography if necessary.

A dosage of IL10 and IL6 in the anterior chamber of the eye is recommended in case of involvement or if there is doubt about ophthalmic involvement. If the dosage of IL10 and IL6 is abnormal in the initial examination, it will be repeated at the end of chemotherapy treatment.

- Evaluation of Cerebrospinalfluid (CSF)

A minimum of 3 ml, ideally 10 ml of CSF, will be sent for cytology. The measurement of protein levels should only be done on CSF obtained by lumbar puncture, as normal protein levels are lower in ventricular puncture. If possible, a dosage of β 2-microglobulin, lymphocyte immunophenotyping by flow cytometry, a study of the rearrangement of heavy chains of immunoglobulins for clonality research, and a dosage of cytokines IL10 and IL6 will also be performed.

Definition of response :

- Treatment Response Rate is defined according to the IPCG criteria

Response	Brain Imaging	Corticosteroid Dose	Eye Examination	CSF Cytology
CR	No contrast enhancement	None	Normal	Negative
CRu	No contrast enhancement	Any	Normal	Negative
	Minimal abnormality	Any	Minor RPE abnormality	Negative
PR	50% decrease in enhancing tumor	Irrelevant	Minor RPE abnormality or normal	Negative
	No contrast enhancement	Irrelevant	Decrease in vitreous cells or retinal infiltrate	Persistent or suspicious
PD	25% increase in lesion	Irrelevant	Recurrent or new ocular disease	Recurrent or positive
	Any new site of disease: CNS or systemic			

Abbreviations: CR, complete response; CRu, unconfirmed complete response; RPE, retinal pigment epithelium; PR, partial response; PD, progressive disease.

from Abrey et al. JCO, 2005

Definition of survival :

- Progression-Free Survival (Primary Endpoint) is defined as the time elapsed between the date of completion of the PET scan at the end of chemotherapy and the date of tumor progression or death. Subjects alive without progression at the date of last follow-up will be censored at that date. Tumor progression is defined according to the IPCG criteria.

- Overall Survival is defined as the time elapsed between the date of completion of the PET scan at the end of chemotherapy and the date of death, regardless of the cause. Subjects alive at the date of last follow-up will be censored at that date.

Reasons for early cessation of trial therapy

In LOCAYZE : N/A

In BLOCAGE01 :

The sponsor AP-HP or the Competent Authority (ANSM) may prematurely interrupt all or part of the research temporarily or permanently in the following situations:

- First, in the case of unexpected serious adverse reactions (SUSARs) in a treatment arm or an imbalance of serious adverse reactions between the 2 treatment arms, requiring a re-evaluation of the risk/benefit ratio of the research.
- Similarly, unforeseen events or new information related to the product, in light of which the research or clinical program's objectives will likely not be met, may lead the sponsor AP-HP or the Competent Authority (ANSM) to prematurely interrupt the research.
- The sponsor AP-HP reserves the right to permanently suspend inclusions at any time if it is found that the inclusion objectives are not met.

In the event of termination of the research, subjects included in the research must be followed until the end of their participation, as provided by the protocol.

In the case of premature termination of the research, the decision and justification are transmitted by the sponsor AP-HP to the Competent Authority (ANSM) and to the CPP (ethic committee) within 15 days.

Objectives and entire statistical section (including endpoints)

Primary Objective: Evaluate the predictive value of FDG-PET performed at the end of R-MPV-A chemotherapy, on the PFS of immunocompetent patients aged 60 and over, suffering from high-grade histological PCNSL, newly diagnosed and histologically proven.

The secondary objectives focus on evaluating:

- Predictive and Prognostic Value of pre-therapeutic and interim FDG-PET (after 2 cycles of chemotherapy).
- Predictive and Prognostic Value of diffusion and perfusion MRI
- Added Value of the combination of several imaging biomarkers on the prediction of PFS including lesion volumes, metabolite ratios in spectroscopy, relative cerebral blood flow, and magnetic susceptibility indices.
- Predictive Values of initial and interim PET-MRI on end-of-treatment MRI results
- Prognostic Value of pre-therapeutic and interim PET-MRI on PFS and OS at 6 months from the first PET-MRI examination.
- Added Prognostic Value of the combination of several imaging and biology biomarkers (PET parameters, diffusion and perfusion MRI, and IL6 and IL10 sampling in the CSF) on PFS and OS at 6 months from the first PET-MRI examination.
- Prognostic Value of end-of-treatment PET-MRI on PFS and OS at 6 months from the third PET-MRI examination.

Description of planned statistical methods including the schedule of planned interim analyses

- Main Analysis

The analysis will be conducted in patients alive without progression at the end of chemotherapy. PFS rates in the PET-negative and PET-positive groups will be compared using the Log-Rank test.

-Secondary Analyses

- To estimate the contribution of end-of-chemotherapy PET compared to other potential predictors of PFS from the end of treatment, an analysis adjusted for these factors will be implemented using a Cox model. The factors studied will be the end-of-treatment PET (positive versus negative), treatment, Karnofsky index at the end of treatment, age (70 years or older or younger), response type (complete, partial, or stable according to IPCG criteria), quantitative and semi-quantitative parameters measured in diffusion and perfusion MRI, and the change in SUV between inclusion and end of treatment. The same analysis will be conducted to study the predictive value of PET on OS. This analysis will concern patients alive without progression at the end of treatment.
- Treatment responses defined from the consensus criteria of the IPCG and defined based on the results of interim and end-of-chemotherapy PET will be presented. Proportions of concordant and discordant cases between the two evaluation methods will be estimated.
- Correlations between the time evolutions (pre-therapeutic, interim, and end-of-treatment) of the different quantitative biomarkers of PET-MRI will be studied using a mixed regression model.
- The predictive effect on PFS of pre-therapeutic PET will be studied using a Cox model. The analysis will cover all patients included in the study. The factors studied will be pre-therapeutic PET, Karnofsky index at inclusion, age (70 years or older or younger), and quantitative parameters measured in diffusion and perfusion MRI.
- The predictive values of initial and interim PET-MRI on the results of end-of-treatment PET-MRI will be studied (sensitivity, specificity, negative and positive predictive values).
- The prognostic value of pre-therapeutic and interim PET-MRI on PFS and OS at 6 months from the first PET-MRI examination will be assessed using Cox models.
- The added prognostic value of the combination of several imaging and biology biomarkers (PET parameters, diffusion and perfusion MRI, and IL6 and IL10 sampling data in CSF) on PFS and OS at 6 months from the first PET-MRI examination will be assessed using Cox models.

The hypotheses are based on the PFS rates observed in the CNS LYMPHOMA study. This study included the same population that is planned to be included in the BLOCAGE trial. The 2-year PFS rate is estimated at 69% for patients for whom the end-of-treatment PET scan will be negative and 30% for patients with a positive PET scan. Furthermore, it is expected that 30% of patients will have a positive PET scan and 70% a negative PET scan.

For an inclusion period of 4 years and 4 months, and a minimum follow-up duration of two years, 40 evaluable patients at the end of treatment are needed to ensure the study has a power of 80%, with a type I error risk of 5%. It is anticipated that 28% of patients will have died or progressed before the end of treatment (data from CNS LYMPHOMA). The inclusion of 66 patients will allow for 40 evaluable patients at the end of treatment.

Annexe 3. Score METRICS (Kocak, 2024)

16/04/2024 16:57

METRICS

METRICS Tool v1.0

Please fill out all conditions first for relevant sections and then all active items to calculate METRICS score.

Please note that default option is "No".

? Stands for explanation of items and conditions.

C Stands for conditional items or sections.

Items/Conditions	Definitions	Weights	Options
Study Design			
Item#1	? Adherence to radiomics and/or machine learning-specific checklists or guidelines	0.0368	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#2	? Eligibility criteria that describe a representative study population	0.0735	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#3	? High-quality reference standard with a clear definition	0.0919	<input type="radio"/> Yes <input checked="" type="radio"/> No
Imaging Data			
Item#4	? Multi-center	0.0438	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#5	? Clinical translatability of the imaging data source for radiomics analysis	0.0292	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#6	? Imaging protocol with acquisition parameters	0.0438	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#7	? The interval between imaging used and reference standard	0.0292	<input type="radio"/> Yes <input checked="" type="radio"/> No
Segmentation C			
Condition#1	? Does the study include segmentation?		<input type="radio"/> Yes <input checked="" type="radio"/> No
Condition#2	? Does the study include fully automated segmentation?		<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#8	? Transparent description of segmentation methodology	0.0337	<input type="radio"/> Yes <input type="radio"/> No
Item#9	? Formal evaluation of fully automated segmentation C	0.0225	<input type="radio"/> Yes <input type="radio"/> No
Item#10	? Test set segmentation masks produced by a single reader or automated tool	0.0112	<input type="radio"/> Yes <input type="radio"/> No
Image Processing and Feature Extraction			
Condition#3	? Does the study include hand-crafted feature extraction?		<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#11	? Appropriate use of image preprocessing techniques with transparent description	0.0622	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#12	? Use of standardized feature extraction software C	0.0311	<input type="radio"/> Yes <input type="radio"/> No
Item#13	? Transparent reporting of feature extraction parameters, otherwise providing a default configuration statement	0.0415	<input type="radio"/> Yes <input checked="" type="radio"/> No
Feature Processing			
Condition#4	? Does the study include tabular data?		<input type="radio"/> Yes <input checked="" type="radio"/> No
Condition#5	? Does the study include end-to-end deep learning?		<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#14	? Removal of non-robust features C	0.0200	<input type="radio"/> Yes <input type="radio"/> No
Item#15	? Removal of redundant features C	0.0200	<input type="radio"/> Yes <input type="radio"/> No
Item#16	? Appropriateness of dimensionality compared to data size C	0.0300	<input type="radio"/> Yes <input type="radio"/> No
Item#17	? Robustness assessment of end-to-end deep learning pipelines C	0.0200	<input type="radio"/> Yes <input type="radio"/> No
Preparation for Modeling			
Item#18	? Proper data partitioning process	0.0599	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#19	? Handling of confounding factors	0.0300	<input type="radio"/> Yes <input checked="" type="radio"/> No
Metrics and Comparison			
Item#20	? Use of appropriate performance evaluation metrics for task	0.0352	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#21	? Consideration of uncertainty	0.0234	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#22	? Calibration assessment	0.0176	<input type="radio"/> Yes <input checked="" type="radio"/> No
Item#23	? Use of uni-parametric imaging or proof of its inferiority	0.0117	<input type="radio"/> Yes <input checked="" type="radio"/> No

<https://metricscore.github.io/metrics/METRICS.html>

1/2

16/04/2024 16:57

METRICS

Item#24	<input type="checkbox"/> ?	Comparison with a non-radiomic approach or proof of added clinical value	0.0293	<input type="radio"/> Yes	<input checked="" type="radio"/> No
Item#25	<input type="checkbox"/> ?	Comparison with simple or classical statistical models	0.0176	<input type="radio"/> Yes	<input checked="" type="radio"/> No
Testing					
Item#26	<input type="checkbox"/> ?	Internal testing	0.0375	<input type="radio"/> Yes	<input checked="" type="radio"/> No
Item#27	<input type="checkbox"/> ?	External testing	0.0749	<input type="radio"/> Yes	<input checked="" type="radio"/> No
Open Science					
Item#28	<input type="checkbox"/> ?	Data availability	0.0075	<input type="radio"/> Yes	<input checked="" type="radio"/> No
Item#29	<input type="checkbox"/> ?	Code availability	0.0075	<input type="radio"/> Yes	<input checked="" type="radio"/> No
Item#30	<input type="checkbox"/> ?	Model availability	0.0075	<input type="radio"/> Yes	<input checked="" type="radio"/> No
			Total METRICS score:	0.0%	
			<input type="checkbox"/> ?	Quality category:	
			<input type="checkbox"/> ?	Publication ID:	

If you publish any work which uses this tool, please cite the following publication:

Kocak B, Akinci D, Antonoli T, Mercaldo N, et al. METHodological Radiomics Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMI1. Insights Imaging. 2024;15(1):8. Published 2024 Jan 17. doi:10.1186/s13244-023-01572-w

Annexe 4. Score Radiomics Quality Score (Lambin, 2017)

Image protocol quality - well-documented image protocols (for example, contrast, slice thickness, energy, etc.) and/or usage of public image protocols allow reproducibility/replicability	<input type="checkbox"/> protocols well documented <input type="checkbox"/> public protocol used <input type="checkbox"/> none
Multiple segmentations - possible actions are: segmentation by different physicians/algorithms/software, perturbing segmentations by (random) noise, segmentation at different breathing cycles. Analyse feature robustness to segmentation variabilities	<input type="radio"/> yes <input checked="" type="radio"/> no
Phantom study on all scanners - detect inter-scanner differences and vendor-dependent features. Analyse feature robustness to these sources of variability	<input type="radio"/> yes <input checked="" type="radio"/> no
Imaging at multiple time points - collect images of individuals at additional time points. Analyse feature robustness to temporal variabilities (for example, organ movement, organ expansion/shrinkage)	<input type="radio"/> yes <input checked="" type="radio"/> no
Feature reduction or adjustment for multiple testing - decreases the risk of overfitting. Overfitting is inevitable if the number of features exceeds the number of samples. Consider feature robustness when selecting features	<input type="radio"/> Either measure is implemented <input checked="" type="radio"/> Neither measure is implemented
Multivariable analysis with non radiomics features (for example, EGFR mutation) - is expected to provide a more holistic model. Permits correlating/inferencing between radiomics and non radiomics features	<input type="radio"/> yes <input checked="" type="radio"/> no
Detect and discuss biological correlates - demonstration of phenotypic differences (possibly associated with underlying gene-protein expression patterns) deepens understanding of radiomics and biology	<input type="radio"/> yes <input checked="" type="radio"/> no
Cut-off analyses - determine risk groups by either the median, a previously published cut-off or report a continuous risk variable. Reduces the risk of reporting overly optimistic results	<input type="radio"/> yes <input checked="" type="radio"/> no
Discrimination statistics - report discrimination statistics (for example, C-statistic, ROC curve, AUC) and their statistical significance (for example, p-values, confidence intervals). One can also apply resampling method (for example, bootstrapping, cross-validation)	<input type="checkbox"/> a discrimination statistic and its statistical significance are reported <input type="checkbox"/> a resampling method technique is also applied <input type="checkbox"/> none

<p>Calibration statistics - report calibration statistics (for example, Calibration-in-the-large/slope, calibration plots) and their statistical significance (for example, P-values, confidence intervals). One can also apply resampling method (for example, bootstrapping, cross-validation)</p>	<p><input type="checkbox"/> a calibration statistic and its statistical significance are reported</p> <p><input type="checkbox"/> a resampling method technique is applied</p> <p><input type="checkbox"/> none</p>
<p>Prospective study registered in a trial database - provides the highest level of evidence supporting the clinical validity and usefulness of the radiomics biomarker</p>	<p><input type="radio"/> yes</p> <p><input checked="" type="radio"/> no</p>
<p>Validation - the validation is performed without retraining and without adaptation of the cut-off value, provides crucial information with regard to credible clinical performance</p>	<p><input checked="" type="checkbox"/> No validation</p> <p><input type="checkbox"/> validation is based on a dataset from the same institute</p> <p><input type="checkbox"/> validation is based on a dataset from another institute</p> <p><input type="checkbox"/> validation is based on two datasets from two distinct institutes</p> <p><input type="checkbox"/> the study validates a previously published signature</p> <p><input type="checkbox"/> validation is based on three or more datasets from distinct institutes</p>
<p>Comparison to 'gold standard' - assess the extent to which the model agrees with/is superior to the current 'gold standard' method (for example, TNM-staging for survival prediction). This comparison shows the added value of radiomics</p>	<p><input type="radio"/> yes</p> <p><input checked="" type="radio"/> no</p>
<p>Potential clinical utility - report on the current and potential application of the model in a clinical setting (for example, decision curve analysis).</p>	<p><input type="radio"/> yes</p> <p><input checked="" type="radio"/> no</p>
<p>Cost-effectiveness analysis - report on the cost-effectiveness of the clinical application (for example, QALYs generated)</p>	<p><input type="radio"/> yes</p> <p><input checked="" type="radio"/> no</p>
<p>Open science and data - make code and data publicly available. Open science facilitates knowledge transfer and reproducibility of the study</p>	<p><input type="checkbox"/> scans are open source</p> <p><input type="checkbox"/> region of interest segmentations are open source</p> <p><input type="checkbox"/> the code is open sourced</p> <p><input type="checkbox"/> radiomics features are calculated on a set of representative ROIs and the calculated features and representative ROIs are open source</p>

Annexe 5. Checklist CLAIM (Mongan, 2020)

Checklist for Artificial Intelligence in Medical Imaging

Checklist for Artificial Intelligence in Medical Imaging (CLAIM)		
Section/Topic	No.	Item
TITLE or ABSTRACT		
	1	Identification as a study of AI methodology, specifying the category of technology used (eg, deep learning)
ABSTRACT		
	2	Structured summary of study design, methods, results, and conclusions
INTRODUCTION		
	3	Scientific and clinical background, including the intended use and clinical role of the AI approach
	4	Study objectives and hypotheses
METHODS		
Study Design	5	Prospective or retrospective study
	6	Study goal, such as model creation, exploratory study, feasibility study, noninferiority trial
Data	7	Data sources
	8	Eligibility criteria: how, where, and when potentially eligible participants or studies were identified (eg, symptoms, results from previous tests, inclusion in registry, patient-care setting, location, dates)
	9	Data preprocessing steps
	10	Selection of data subsets, if applicable
	11	Definitions of data elements, with references to common data elements
	12	De-identification methods
	13	How missing data were handled
Ground Truth	14	Definition of ground truth reference standard, in sufficient detail to allow replication
	15	Rationale for choosing the reference standard (if alternatives exist)
	16	Source of ground truth annotations; qualifications and preparation of annotators
	17	Annotation tools
	18	Measurement of inter- and intrarater variability; methods to mitigate variability and/or resolve discrepancies
Data Partitions	19	Intended sample size and how it was determined
	20	How data were assigned to partitions; specify proportions
	21	Level at which partitions are disjoint (eg, image, study, patient, institution)
Model	22	Detailed description of model, including inputs, outputs, all intermediate layers and connections
	23	Software libraries, frameworks, and packages
	24	Initialization of model parameters (eg, randomization, transfer learning)
Training	25	Details of training approach, including data augmentation, hyperparameters, number of models trained
	26	Method of selecting the final model
	27	Ensembling techniques, if applicable
Evaluation	28	Metrics of model performance
	29	Statistical measures of significance and uncertainty (eg, confidence intervals)
	30	Robustness or sensitivity analysis
	31	Methods for explainability or interpretability (eg, saliency maps) and how they were validated
	32	Validation or testing on external data
RESULTS		
Data	33	Flow of participants or cases, using a diagram to indicate inclusion and exclusion
	34	Demographic and clinical characteristics of cases in each partition
Model performance	35	Performance metrics for optimal model(s) on all data partitions
	36	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)
	37	Failure analysis of incorrectly classified cases
DISCUSSION		
	38	Study limitations, including potential bias, statistical uncertainty, and generalizability
	39	Implications for practice, including the intended use and/or clinical role
OTHER INFORMATION		
	40	Registration number and name of registry
	41	Where the full study protocol can be accessed
	42	Sources of funding and other support; role of funders

Annexe 6. Checklist mi-CLAIM (Norgeot, 2020)

Table 1 The MI-CLAIM checklist		
Before paper submission		
Study design (Part 1)	Completed: page number	Notes if not completed
The clinical problem in which the model will be employed is clearly detailed in the paper.	<input type="checkbox"/>	
The research question is clearly stated.	<input type="checkbox"/>	
The characteristics of the cohorts (training and test sets) are detailed in the text.	<input type="checkbox"/>	
The cohorts (training and test sets) are shown to be representative of real-world clinical settings.	<input type="checkbox"/>	
The state-of-the-art solution used as a baseline for comparison has been identified and detailed.	<input type="checkbox"/>	
Data and optimization (Parts 2, 3)	Completed: page number	Notes if not completed
The origin of the data is described and the original format is detailed in the paper.	<input type="checkbox"/>	
Transformations of the data before it is applied to the proposed model are described.	<input type="checkbox"/>	
The independence between training and test sets has been proven in the paper.	<input type="checkbox"/>	
Details on the models that were evaluated and the code developed to select the best model are provided.	<input type="checkbox"/>	
Is the input data type structured or unstructured?	<input type="checkbox"/> Structured <input type="checkbox"/> Unstructured	
Model performance (Part 4)	Completed: page number	Notes if not completed
The primary metric selected to evaluate algorithm performance (e.g., AUC, F-score, etc.), including the justification for selection, has been clearly stated.	<input type="checkbox"/>	
The primary metric selected to evaluate the clinical utility of the model (e.g., PPV, NNT, etc.), including the justification for selection, has been clearly stated.	<input type="checkbox"/>	
The performance comparison between baseline and proposed model is presented with the appropriate statistical significance.	<input type="checkbox"/>	
Model examination (Part 5)	Completed: page number	Notes if not completed
Examination technique 1*	<input type="checkbox"/>	
Examination technique 2*	<input type="checkbox"/>	
A discussion of the relevance of the examination results with respect to model/algorithm performance is presented.	<input type="checkbox"/>	
A discussion of the feasibility and significance of model interpretability at the case level if examination methods are uninterpretable is presented.	<input type="checkbox"/>	
A discussion of the reliability and robustness of the model as the underlying data distribution shifts is included.	<input type="checkbox"/>	
Reproducibility (Part 6): choose appropriate tier of transparency	Notes	
Tier 1: complete sharing of the code	<input type="checkbox"/>	
Tier 2: allow a third party to evaluate the code for accuracy/fairness; share the results of this evaluation	<input type="checkbox"/>	
Tier 3: release of a virtual machine (binary) for running the code on new data without sharing its details	<input type="checkbox"/>	
Tier 4: no sharing	<input type="checkbox"/>	

PPV, positive predictive value; NNT, numbers needed to treat. *Common examination approaches based on study type: for studies involving exclusively structured data, coefficients and sensitivity analysis are often appropriate; for studies involving unstructured data in the domains of image analysis or natural language processing, saliency maps (or equivalents) and sensitivity analyses are often appropriate.

Annexe 7. Checklist FUTURE-AI (Lekadir, 2023)

33

Table 1: Overview of the FUTURE-AI checklist organised per principle and ranked according to the AI stages in which the respective checklist item is applicable. The stages consist of (1) clinical conceptualisation, (2) end-user requirement gathering, (3) technical design, (4) data collection and preparation, (5) AI implementation and optimisation, (6) AI evaluation, and (7) AI deployment and monitoring.

No	Stage	Title	Question(s)
Fairness			
1.	1,2	Multi-disciplinarity	Did you design your AI algorithm with a diverse team of stakeholders? Did you collect requirements from a diverse set of end-users?
2.	1,2,4	Definition of fairness	Did you define fairness for your specific imaging application? Did you ask clinicians about hidden sources of data imbalance?
3.	4	Metadata labelling	At data collection, did you record key metadata variables on individuals and groups?
4.	5,6	Estimation of data (im)balance	Did you inspect and ensure the diversity of the training and evaluation data?
5.	4-6	Multi-centre datasets	Did you train and assess your algorithm for multi-centre imaging samples? Does your algorithm maintain accuracy across radiology units and geographical locations? In particular, is it applicable in centres with reduced imaging quality (e.g. in resource-limited countries)?
6.	5-7	Transparency of fairness	Did you document the data characteristics, including existing (im)balances?
7.	6	Fairness evaluation and metrics	Did you thoroughly evaluate the fairness of your AI algorithm? Did you use a suitable dataset and dedicated metrics? Do you have a mechanism for continuous evaluation of your algorithm's fairness?
8.	5,6	Fairness corrective measures	If you identified sources of bias in the data, did you implement mitigation measures?
9.	6	Continuous monitoring of fairness	Do you have a mechanism for continuous testing of your algorithm's fairness over its lifetime?
10.	6	Information and training on fairness	Did you prepare information and training material for the radiologists and clinicians, to inform on potential biases and maximise fairness during the algorithm's use?
Universality			
11.	1,2	Definition of clinical task	Did you use a universal definition of the clinical task?
12.	2,3,5	Software standardisation	Did you design and implement the imaging AI solution using proven libraries and framework standards that readily allow for extension and maintenance?
13.	3,4	Image annotation standardisation	Did you annotate your dataset in an objective, reproducible and standardised way?
14.	3,4	Variation of quantified biomarkers	Are the methods you used for feature quantification compliant with consensus provided by standards initiatives?
15.	6	Evaluation metric selection and reporting	Did you use universal, transparent, comparable, and reproducible criteria and metrics for your model's performance assessment?
16.	6	Reference dataset evaluation	Did you evaluate your model on at least one open access benchmark dataset that is representative of your model's task and expected real-world data exposure after deployment?
17.	6	Reporting standards compliance	Did you adhere to a standardised reporting guideline when assessing and communicating the design and findings of your study?
Traceability			
18.	1-3	Model scope	Did you agree with the clinicians/radiologists a precise definition of the model's scope? Did you precisely define the model's intended use, the input imaging modalities, the necessary steps to provide the input to the AI model, the reference ground truth if any, the intended output and the use case scenarios? Did you check for any known limitations of the diagnostic/prognostic problem faced?
19.	4	Data provenance	Did you prepare a complete documentation of the imaging dataset you used? Did you include the relevant DICOM tags? Did you structurally list the related clinical/genomic/pathology data?
20.	4,5	Data localization and distribution	Did you annotate the location of data over the network? Did you analyse dataset statistics with respect to the capability to represent the phenomenon at hand over the various clinical sites? Did you quantify missing values and any gaps or known biases?
21.	4,5	Data-preparation documentation	Did you keep track in a structured manner of the whole pre-processing pipeline of imaging and related data? Did you specify input/output, nature, prerequisites and requirements of your pre-processing and data preparation methods?
22.	3-5	Specification of clinical references	Did you include a clear description of the radiological/clinical standards or biomarkers used as reference? Did you include a complete record of the segmentation process, if any?
23.	5	Training recording	Did you record the details of the training process? Did you included a careful description of imaging and non-imaging features?
24.	6	Validation documentation	Did you document your validation process and the model selection approach agreed with clinicians?
25.	5-7	Final model details	Did you detail the characteristics of the final model released?
26.	7	Traceability tool	Did you equipped your model with a traceability tool? Did you manage the dynamics of your model?
27.	1-7	AI Model passport	Did you prepare a full metadata record of all the pieces of information of your model?
28.	5-7	Accountability and risk specification	Did you make a risk analysis for your model? Did you prepare a tool to keep track of the usage of your model?

Table 1: Overview of the FUTURE-AI checklist organised per principle and ranked according to the AI stages in which the respective checklist item is applicable. The stages consist of (1) clinical conceptualisation, (2) end-user requirement gathering, (3) technical design, (4) data collection and preparation, (5) AI implementation and optimisation, (6) AI evaluation, and (7) AI deployment and monitoring. (cont'd)

No	Stage	Title	Question(s)
Usability			
29.	1	User engagement	Did you engage users in the design and development of the AI tool?
30.	2	Requirements definition	Did you compile end-user requirements?
31.	3	User interfaces	Did you design appropriate user interfaces?
32.	5	Usable explainability	Did you implement any type of explainability that will be usable and actionable by the radiologist?
33.	6	Usability testing	Did you design an appropriate usability study?
34.	6	In-silico validation	Did you consider an in-silico validation of usability?
35.	6	Usability metrics	Did you define the appropriate usability metrics for evaluation?
36.	6,7	Clinical Integration	Did you evaluate the usability of your tool after integration in the clinical workflows of the clinical sites?
37.	6,7	Training material	Did you provide end-users with resources to learn to adopt and appropriately work with your tool?
38.	5-7	Usability monitoring	Did you implement monitoring mechanisms to assess changes in user needs and re-evaluate the appropriateness of the AI solution though time?
Robustness			
39.	4	Image harmonisation	Did you implement any image harmonisation solutions to account for image heterogeneity?
40.	4	Feature harmonisation	Did you perform any feature harmonisation study before developing your predictive models? Did you assess, minimise, and report the variation across features?
41.	4	Intra- and inter-observer variability	Did you perform any intra- and inter-observer annotation studies?
42.	4	Quality control	Did you use any quality control tools to identify abnormal deviations or artefacts in images?
43.	4	Phantoms	Did you use phantoms to harmonise patient images and/or measurements?
44.	4,5	Data augmentation for model training	Did you use data augmentation techniques to improve training of AI models?
45.	5,6	Training on heterogeneous data	Did you train and evaluate your tools with heterogeneous datasets from multiple clinical centres, vendors, and protocols?
46.	5,6	Uncertainty estimation	Did you report any kind of model uncertainty beyond the classifier's discriminant or confidence score?
47.	6,7	Equity in accessibility	Did you optimise your tool with images from resource-limited settings in low-to-middle countries?
Explainability			
48.	1,2	Clinical requirements on explainability	Did you consult with the clinicians to determine which explainability methods suit them? Did you intuitively present the different explanation methods to the clinicians and did they develop a clear understanding of them?
49.	1-5	Incorporation of clinical concepts	Did you consider using clinical annotations and clinical concepts as parameters of the AI algorithms or neural networks to explicitly introduce a level of clinical interpretability?
50.	1,3,5	Multiple explanation methods	Did you explore multiple and complementary explainability methods?
51.	6	Identifying explainable imaging biomarkers	To increase clinical value, did you evaluate if the explainability methods enable to identify imaging features or structures that can serve as imaging biomarkers? Did you determine if the identified imaging biomarkers are previously known?
52.	6	Quantitative evaluation of explainability	Did you use some quantitative evaluation tests to determine if the explanations are robust and trustworthy?
53.	6	Qualitative evaluation of explainability	Did you perform some qualitative evaluation tests with clinicians?
54.	6	Robustness of explainability against adversarial attacks	Did you evaluate robustness to adversarial attacks, by assessing if the explanations remain consistent when the input images are subjected to small input perturbations and noise?
55.	6,7	Explainability in clinical practice	Did you evaluate clinical effects of explainability methods by performing a human-AI study, where clinician performs clinical tasks using the AI tool with & without explanations? Did you identify any resulting bias introduced by explainability methods?