



HAL
open science

Développements en radiomique pour une meilleure caractérisation du gliome infiltrant du tronc cérébral à partir d'imagerie par résonance magnétique

Jessica Goya Outi

► **To cite this version:**

Jessica Goya Outi. Développements en radiomique pour une meilleure caractérisation du gliome infiltrant du tronc cérébral à partir d'imagerie par résonance magnétique. Imagerie médicale. Université Paris Saclay (COMUE), 2019. Français. NNT : 2019SACLS219 . tel-02412711

HAL Id: tel-02412711

<https://theses.hal.science/tel-02412711v1>

Submitted on 15 Dec 2019

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.

Développements en radiomique pour une meilleure caractérisation du gliome infiltrant du tronc cérébral à partir d'imagerie par résonance magnétique

Thèse de doctorat de l'Université Paris-Saclay
préparée à Université Paris-Sud

Ecole doctorale n°575 Electrical, Optical, Bio: PHYSICS AND ENGINEERING (EOBE)
Spécialité de doctorat : Imagerie et Physique Médicale

Thèse présentée et soutenue à Orsay, le 25/09/2019, par

JESSICA GOYA OUTI

Composition du Jury :

Nicolas Passat Professeur, Université de Reims Champagne-Ardenne	Président
Carole Lartzien Directrice de recherche, CNRS (INSA CREATIS Lyon)	Rapporteur
Sébastien Benzekry Chargé de recherche, INRIA (TEAM MONC Bordeaux)	Rapporteur
Jacques Grill MD, PhD, Institut Gustave Roussy (Université Paris-Sud)	Examineur
Nathalie Boddaert Professeur, AP-HP (Hôpital Necker)	Examinatrice
Frédérique Frouin Chargée de recherche, INSERM (IMIV)	Directrice de thèse

Remerciements

Je commence par m'excuser car les remerciements sont certainement incomplets.

Je tiens à remercier :

- ma directrice de thèse, Frédérique Frouin, pour sa disponibilité, gentillesse, patience et implication dans mon travail ;
- Alexis Pister, pour son travail de stage sur l'automatisation de la standardisation d'intensités des images d'IRM ;
- le radiologue Raphael Calmon, pour m'avoir confié ses données et d'avoir consacré du temps à m'expliquer des concepts importants ;
- la directrice du laboratoire IMIV, Irène Buvat, pour toutes les corrections de mes travaux et les bons « insights » ;
- mes collègues Fanny Orhac et Cristophe Nioche pour leur soutien scientifique et moral ;
- les personnes de Neurospin, Vincent Frouin, Cathy Philippe et Hélène Urien, pour leur travail de stockage de données d'imagerie ;
- les personnes de l'Hôpital Necker, Natalie Boddaert et Stéphanie Puget, car grâce à leur bonne prise en charge des patients, ils ont pu intégrer ce travail de thèse ;
- les personnes de Gustave Roussy, Jacques Grill, Lila Saidoun, David Castel, Marie-Anne Dembily, pour avoir consacré du temps à me confier les données d'imagerie et les informations moléculaires des patients ;
- les membres du jury, Carole Lartizien, Sébastien Benzekry, Nicolas Passat, et encore une autre fois, Nathalie Boddaert et Jacques Grill pour avoir accepté d'évaluer mon travail de thèse, pour le temps consacré à la lecture et pour les corrections constructives.

Table des matières

Table des matières	v
Introduction	vii
1 La radiomique en IRM pour la neuro-oncologie	1
1.1 Présentation de la radiomique en IRM pour la neuro-oncologie	2
1.2 Étapes de la radiomique	14
1.3 Références	34
2 Le gliome infiltrant du tronc cérébral	45
2.1 Caractéristiques cliniques	46
2.2 Biopsie et sous-types moléculaires	46
2.3 GITC selon la classification de l’OMS	48
2.4 Approches thérapeutiques	48
2.5 Rôle de l’IRM	49
2.6 Radiomique pour la caractérisation du GITC	51
2.7 Bases de données de patients atteint de GITC	52
2.8 Références	55
3 La standardisation des images IRM cérébrales	61
3.1 État de l’art	62
3.2 La méthode de standardisation proposée	62
3.3 Définition d’un pipeline de contrôle de qualité	66
3.4 Application des méthodes de standardisation aux cohortes GITC	69
3.5 Application des méthodes de standardisation sur une cohorte indépendante présentant des lésions cérébrales	77
3.6 Discussion et conclusion	80
3.7 Références	88
4 Etudes préliminaires en radiomique	91
4.1 Questions méthodologiques posées pour le calcul des indices radiomiques en IRM	92
4.2 Impact de la discrétisation des niveaux de gris sur les indices radiomiques	92
4.3 Impact du volume de la région d’intérêt sur les indices radiomiques	106
4.4 Proposition d’une méthode de calcul local d’indices radiomiques	112
4.5 Discussion et conclusion	116
4.6 Références	118
5 Études radiomiques appliquées aux prédictions des mutations du GITC	121
5.1 Analyses uni-variées des mutations génétiques associées au GITC	123
5.2 Analyses multi-variées pour prédire les mutations génétiques associées au GITC	134
5.3 La méthode multi-modèles proposée	151
5.4 Discussion et conclusion	162
5.5 Références	167

6 Études radiomiques appliquées aux analyses de survie des patients atteints de GITC	169
6.1 Analyses uni-variées de la survie globale des patients atteints de GITC	170
6.2 Analyses multi-variées de la survie globale des patients GITC	180
6.3 Discussion et conclusion	186
6.4 Références	189
Conclusion et perspectives	191
Liste de publications	195
A Annexes	I
A.1 Tableaux annexes	I
A.2 Documents annexes	XVI
B Liste des acronymes	XXXVII
C Glossaire	XXXIX
D Liste des symboles	XLI

Introduction

Introduction générale

L'imagerie médicale est devenue indispensable en cancérologie car elle permet d'analyser une lésion dans sa globalité de façon non-invasive et elle peut être répétée plusieurs fois au cours du traitement. Elle s'avère être ainsi un complément à la biopsie, pouvant fournir des informations sur l'environnement de la tumeur, voire sur l'ensemble des lésions lorsque le cancer est à un stade avancé.

La radiomique est une discipline récente qui repose sur l'hypothèse que certaines informations pertinentes qui ne sont pas facilement repérables visuellement sur les images peuvent être trouvées en calculant une grande quantité d'indices quantitatifs. Cela permettrait de mieux caractériser une tumeur pour mettre en place une médecine personnalisée visant à proposer pour chaque patient un traitement spécifique, dans l'objectif d'améliorer la réponse à la thérapie et de prolonger la survie globale ou la survie sans progression.

Le gliome infiltrant du tronc cérébral (GITC) est une tumeur pédiatrique rare, très agressive et non-opérable. Aucun traitement ne s'est montré efficace jusqu'à présent, sauf la radiothérapie qui permet une réduction momentanée des symptômes liés à la pathologie. Néanmoins, seulement 10% des patients survivent plus de deux ans après le diagnostic. Avec la réalisation sans complication majeure de biopsies du GITC, à partir de 2007, différents sous-types moléculaires du GITC ont été décrits. Ainsi, des protocoles de recherche clinique sont actuellement conduits pour tester de nouvelles thérapies ciblées. Néanmoins, la biopsie est locale, invasive et ne peut pas être répétée. Ainsi, le principal outil de caractérisation du GITC est l'imagerie par résonance magnétique (IRM) multi-modale.

A notre connaissance, aucune étude utilisant la radiomique pour caractériser les patients atteints de GITC n'a été publiée jusqu'à présent. Cette approche permettant une analyse systématique de différents paramètres nous est apparue très pertinente pour aider à la compréhension de la maladie en explorant systématiquement les informations présentes dans les images d'IRM.

Pour initier le travail, nous avons eu accès à des données d'imagerie acquises depuis 2009 par les équipes de Gustave Roussy et de l'Hôpital Necker. Très récemment, nous avons eu accès à des données issues du programme hospitalier de recherche clinique (PHRC) BIOMEDE. Les données d'imagerie de diagnostic, réalisées avant la biopsie et le début du traitement, ont été acquises suivant le même protocole d'acquisition. De façon générale, quatre modalités d'IRM structurelle étaient disponibles pour l'analyse.

Objectifs de la thèse

L'objectif de cette thèse a été de mettre en place des techniques d'analyse radiomique adaptées aux patients atteints de GITC. Deux aspects ont été analysés :

- la capacité des modèles radiomiques à prédire les principales mutations génétiques connues associées aux GITC à partir de l'acquisition réalisée au moment du diagnostic ;
- la capacité des modèles radiomiques à prédire la survie globale des patients à partir de l'acquisition réalisée au moment du diagnostic.

Le contexte des données, en cours d'inclusion, et le cas de données manquantes ont été considérés pour répondre au mieux au contexte clinique.

Structure du document

Le manuscrit est organisé en six chapitres.

Le chapitre 1 introduit la radiomique et précise son utilisation dans le contexte de l'IRM en neuro-oncologie. Les étapes nécessaires pour la réalisation d'une étude radiomique, qui incluent le pré-traitement des images, le calcul d'indices d'imagerie et les méthodes d'apprentissage automatique, sont exposés.

Le chapitre 2 présente les principales caractéristiques cliniques, radiologiques et moléculaires du GITC et présente l'état de l'art relatif à l'apport de l'imagerie pour le diagnostic et la prise en charge des patients. Ce chapitre détaille également les bases de données des patients atteints de GITC utilisées dans cette thèse.

Le chapitre 3 considère la standardisation des intensités des images d'IRM cérébrales, car de façon générale les intensités en IRM ne sont pas exprimées en unité standard et les gammes de valeurs sont très variables entre les acquisitions. Une méthode de standardisation proposée dans la littérature a été adaptée, appliquée et analysée sur les images de nos bases de données.

Le chapitre 4 expose les études méthodologiques réalisées concernant le calcul des indices de texture. A l'issue des analyses réalisées dans ce chapitre, des recommandations sont fournies pour le calcul de ces indices.

Le chapitre 5 présente les analyses radiomiques réalisées pour la prédiction des mutations génétiques associées au GITC en s'appuyant sur les recommandations établies précédemment. Le pipeline de sélection d'indices que nous avons développé est décrit et les résultats de classification de différentes méthodes d'apprentissage automatique sont détaillés. Dans un deuxième temps, une méthodologie adaptée à l'analyse des bases de données d'imagerie multi-modales à données manquantes est proposée pour prendre en compte le contexte des données disponibles pour cette thèse. L'application de cette méthodologie est illustrée sur un cas d'étude.

Le chapitre 6 présente les analyses radiomiques réalisées pour la prédiction de la survie globale des patients atteints de GITC.

Enfin, une discussion générale relative aux résultats obtenus dans la thèse ainsi que les perspectives pour l'analyse radiomique des patients atteints de GITC concluent le manuscrit.

Chapitre 1

La radiomique en IRM pour la neuro-oncologie

« L'essentiel est invisible pour les yeux »

Antoine de Saint-Exupéry



Ce chapitre est une introduction à la radiomique focalisée sur l'IRM pour la neuro-oncologie. L'état de l'art de la radiomique en IRM en neuro-oncologie est discuté avec une revue des publications dans ce domaine. Ce chapitre donne aussi les étapes nécessaires à la mise en œuvre de la radiomique qui incluent différents domaines : le traitement d'images, la vision par ordinateur, l'apprentissage automatique et les statistiques. Enfin, une discussion sur l'apprentissage profond dans le contexte de la radiomique est présentée.

Sommaire

1.1 Présentation de la radiomique en IRM pour la neuro-oncologie	2
1.1.1 Généralités sur l'IRM cérébrale	2
1.1.2 Généralités sur les tumeurs du système nerveux central	6
1.1.3 Applications de la radiomique en IRM pour la neuro-oncologie	8
1.2 Étapes de la radiomique	14
1.2.1 Standardisation d'images	15
1.2.2 Segmentation des tumeurs	15
1.2.3 Indices radiomiques	16
1.2.4 Sélection d'indices	20
1.2.5 Méthodes d'apprentissage automatique	22
1.2.6 Études de comparaison de méthodes de sélection d'indices et de modèles d'apprentissage supervisé	27
1.2.7 Métriques d'évaluation des modèles	28
1.2.8 L'apprentissage profond	33
1.3 Références	34

1.1 Présentation de la radiomique en IRM pour la neuro-oncologie

La radiomique, suivant le terme introduit par [LAMBIN et collab. \[2012\]](#), permet d'explorer les tumeurs par l'extraction d'une grande quantité d'indices quantitatifs des images médicales. L'hypothèse est que ces indices peuvent synthétiser des informations pertinentes qui ne sont pas facilement repérables par les radiologues et cliniciens. Les images médicales utilisées peuvent être d'origine anatomique, fonctionnelle ou moléculaire (figure 1.1). Cette discipline est définie notamment en oncologie. Toutes les sources d'information d'imagerie ainsi que des informations cliniques peuvent être combinées afin de mieux caractériser une tumeur. Le but est de mettre en place une médecine personnalisée visant de proposer pour chaque patient un traitement spécifique et ainsi améliorer sa réponse à la thérapie et prolonger sa durée de survie.

En oncologie, les intérêts majeurs des analyses radiomiques sont liés à :

- la non-invasivité de l'imagerie
- la représentation de la tumeur dans sa globalité, différemment des biopsies ponctuelles qui représentent seulement une partie restreinte de la tumeur
- la possibilité d'analyser la tumeur à plusieurs reprises et de suivre ainsi sa progression ou sa régression au fil du temps

Les promesses de la radiomique sont très ambitieuses et attractives. Depuis son introduction en 2012, le nombre de publications faisant appel à la radiomique a augmenté exponentiellement (figure 1.2).

La modalité d'imagerie la plus utilisée parmi toutes les publications "radiomique" est l'[imagerie par résonance magnétique \(IRM\)](#) (350 publications, avec la recherche ("*magnetic resonance imaging*" OR "*mri*" OR "*mr*" OR ("*magnetic*" AND "*resonance*" AND "*imaging*")) AND *radiomics* dans PubMed), suivi de la [tomodensitométrie \(TDM\)](#) (327 publications, avec la recherche ("*ct*" OR ("*Computed*" AND "*Tomographie*") OR "*Computed Tomographie*") AND "*radiomics*" dans PubMed) et de la [tomographie par émission de positons \(TEP\)](#) (157 publications, avec la recherche *pet* OR "*proton emission tomography*" OR ("*proton*" AND "*emission*" AND "*tomography*") AND "*radiomics*" dans PubMed).

L'[IRM](#) est la technique d'imagerie la plus répandue pour imager le cerveau. L'hétérogénéité des tumeurs est bien représentée sur les images [IRM](#), qui traduisent différentes propriétés physiopathologiques. Au 2 mai 2019, 89 publications liées à la radiomique pour la neuro-oncologie en [IRM](#) ont été trouvées avec la recherche ("*magnetic resonance imaging*" OR ("*magnetic*" AND "*resonance*" AND "*imaging*") OR "*magnetic resonance imaging*" OR "*mri*" OR "*mr*") AND "*radiomics*" AND ("*neuro*" AND "*oncology*") OR "*neuro oncology*" OR ("*brain*" AND "*tumor*") OR "*glioma*" OR "*glioblastoma*") dans PubMed.

Comme cette thèse porte sur la radiomique pour les images d'[IRM](#) de patients atteints de [gliome infiltrant du tronc cérébral \(GITC\)](#), une tumeur cérébrale, ce chapitre détaille la radiomique spécifiquement dans le contexte de l'[IRM](#) en neuro-oncologie.

1.1.1 Généralités sur l'IRM cérébrale

L'[IRM](#) s'appuie sur les propriétés magnétiques des noyaux d'hydrogène. Rapidement, un champ magnétique uniforme élevé est appliqué et permet d'aligner les protons des molécules d'eau dans les tissus. Ensuite une perturbation du champ magnétique local par radio-fréquence est générée pour désaligner localement les protons. Ces protons retournent à leur position d'équilibre progressivement en libérant des ondes de radio-fréquence. Après un certain temps ([temps d'écho \(TE\)](#)), le signal est mesuré. Une transformée de Fourier est utilisée pour convertir les informations de fréquence en niveaux d'intensité, qui se traduit par des niveaux de gris des voxels de l'image. Les intensités de chaque tissu sur l'image sont différentes car les temps de relaxation longitudinal et transversal (T1 et T2) des protons de différents tissus sont différents. En faisant varier le [temps de répétition \(TR\)](#) entre les perturbations du champ magnétique et le [TE](#), différents types d'images peuvent être créés.

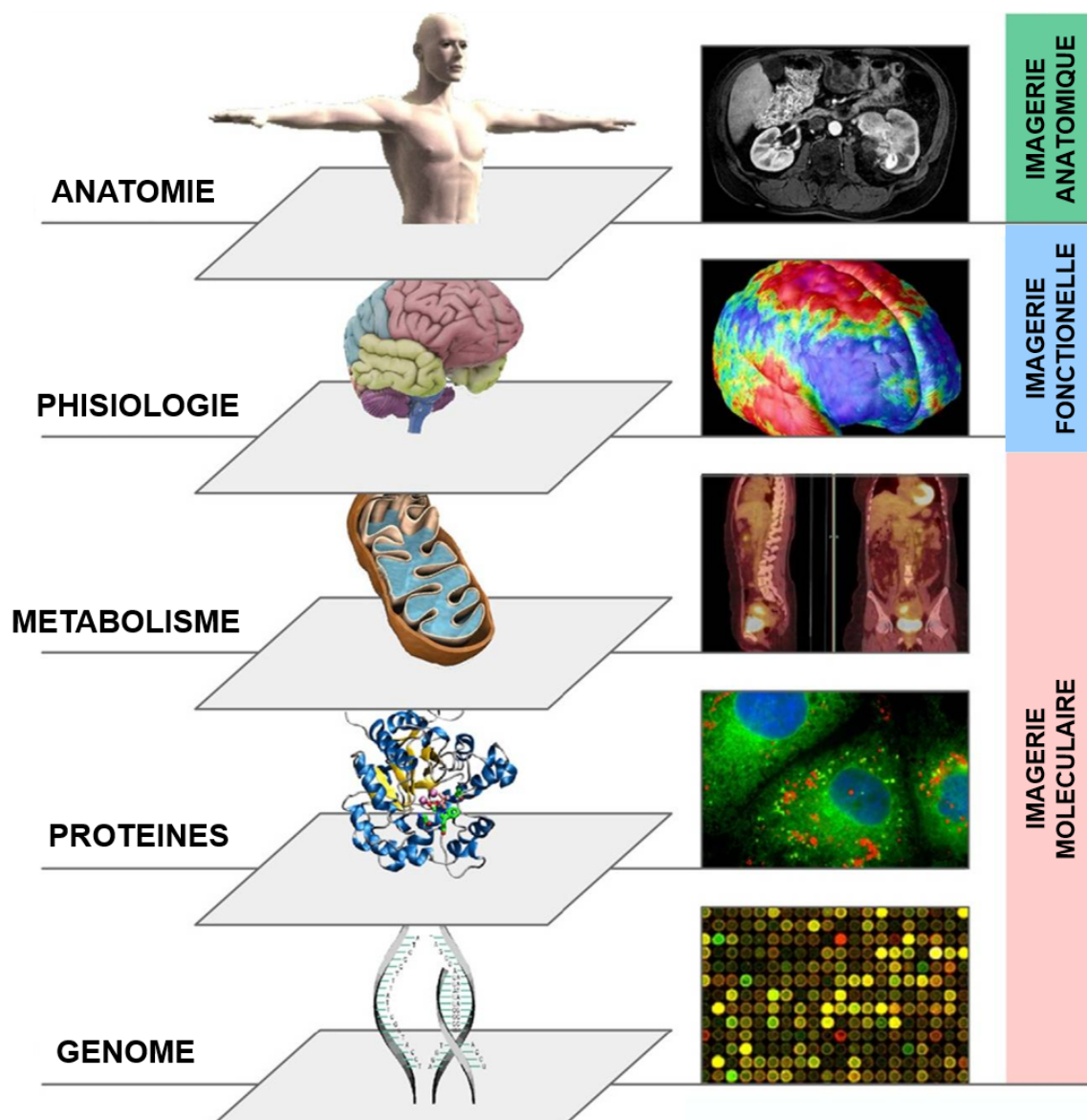


FIGURE 1.1 – Les différents niveaux d'imagerie médicale. Source : Philippe Lambin (LAMBIN et collab. [2012])

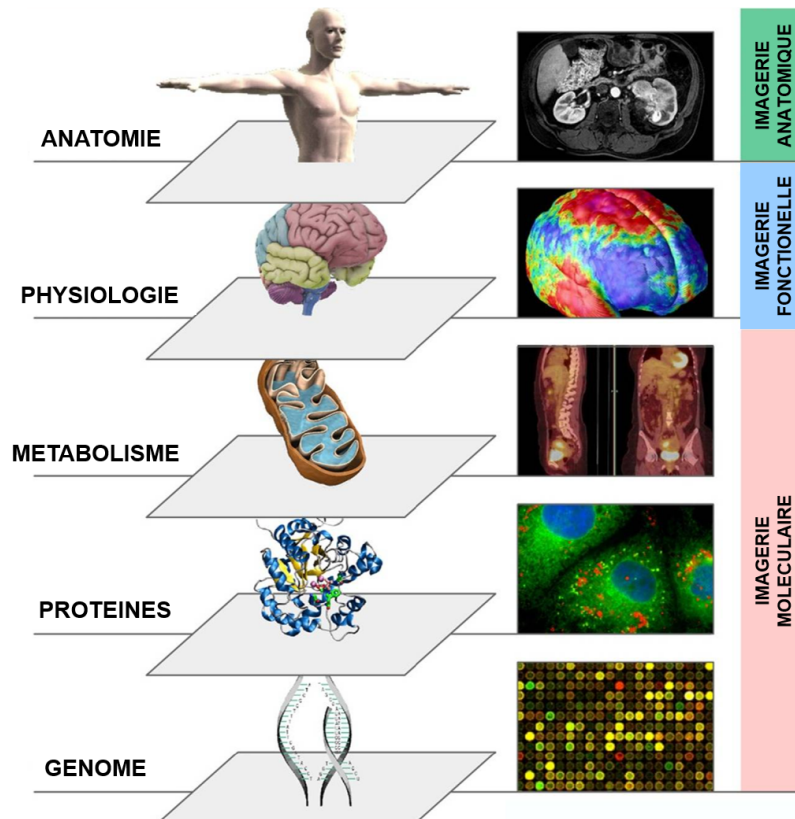


FIGURE 1.2 – Nombre de publications avec le mot clé "radiomics" dans PubMed. Interrogation faite le 2 Mai 2019.

L'IRM conventionnelle

Les séquences d'IRM les plus répandues sont les séquences pondérées en T1 et les séquences pondérées en T2, *T1-weighted* (T1w) et *T2-weighted* (T2w). Pour la séquence T1w les TR et TE utilisés sont courts et les intensités de l'image dépendent essentiellement des temps T1 des tissus de l'image, tandis que pour la séquence T2w les TR et TE utilisés sont longs et les intensités dépendent des temps T2 des tissus. Les séquences T1w peuvent aussi être réalisées suite à l'injection intraveineuse d'un produit de contraste (*T1-weighted après injection de produit de contraste* (T1c)), les plus utilisés sont à base de Gadolinium (Gd-DTPA et Gd-DOTA). Ce produit paramagnétique réduit le temps de relaxation longitudinal T1 et rehausse ainsi spécialement les structures vasculaires. Une autre séquence très utilisée en IRM cérébrale est la séquence *fluid-attenuated inversion recovery* (FLAIR) qui est similaire à la séquence T2w mais utilise des TR et TE très longs de façon à réduire le signal du liquide céphalorachidien (LCR).

Sur les images cérébrales, les différentes séquences sont facilement identifiables comme le montre la figure 1.3. La matière blanche (MB) est claire sur les séquences T1w et T1c, et gris foncé sur les séquences T2w et FLAIR. La matière grise (MG) est gris foncée sur les séquences T1w et T1c, et claire sur les séquences T2w et FLAIR. Le LCR est noir sur les séquences T1w, T1c et FLAIR et très intense sur les séquences T2w. De plus, sur les images T1c les vaisseaux sont très intenses.

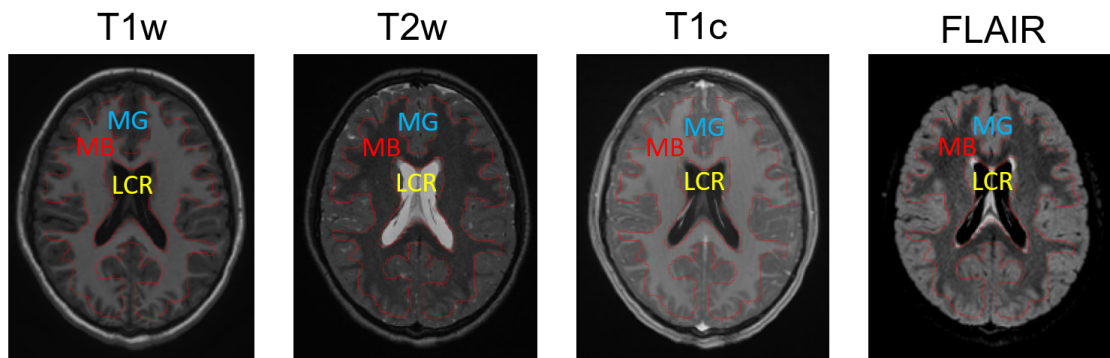


FIGURE 1.3 – Comparaison des séquences IRM : T1w, T2w, T1c et FLAIR

Les séquences **IRM** peuvent ainsi apporter différentes informations sur les tumeurs. Par exemple, les zones rehaussées sur les images **T1c** après l'injection de produit de contraste sont généralement des régions perfusées avec une densité cellulaire importante et des régions où il y a une rupture de la barrière hémato-encéphalique. Avec ce type d'imagerie, les régions nécrosées peuvent être différenciées des régions tumorales viables. D'un autre côté, les images **FLAIR** reflètent bien la présence d'œdème tumoral lié au processus d'inflammation.

L'IRM de perfusion

Les techniques d'imagerie de perfusion ont comme objectif l'étude *in vivo* de la vascularisation du tissu. Or, les tumeurs développent généralement leur propre réseau vasculaire pour apporter les nutriments et l'oxygène, ainsi que pour assurer l'élimination des déchets métaboliques. De cette façon, la formation des vaisseaux sanguins tumoraux est considérée comme une étape essentielle pour qu'une tumeur augmente de taille (**FOLKMAN [1971]**).

Pour les images de perfusion, les techniques d'**IRM** les plus utilisées pour l'étude des tumeurs cérébrales sont : de "premier passage" appelées *dynamic susceptibility contrast (DSC)*, de "rehaussement dynamique" appelées *dynamic contrast enhancement (DCE)* et de marquage de spins appelées *arterial spin labelling (ASL)* (**GUNTHER [2014]**).

Les techniques **DSC** sont capables d'estimer le volume sanguin cérébral (*cerebral blood volume (CBV)*), le flux sanguin cérébral (*cerebral blood flow (CBF)*) et le temps moyen de transit (*mean transit time (MTT)*) du produit de contraste dans le réseau capillaire.

Les techniques **DCE** se basent sur les séquences pondérées en T1 et sont capables d'estimer le volume plasmatique, le volume de l'espace extravasculaire et extracellulaire du tissu, le taux de transfert du traceur dépendant de la perméabilité vasculaire et le ratio de reflux du contraste du tissu vers les vaisseaux.

Finalement, l'**ASL** marque les spins des protons de l'eau situés à l'intérieur des vaisseaux. C'est la plus récente des techniques à être appliquée en pratique clinique. Des acquisitions à intervalles différents entre le marquage et la lecture sont nécessaires pour mesurer d'autres paramètres que le **CBF**, comme le temps d'arrivée du bolus, le temps d'arrivée au pic et le temps d'échange.

L'IRM de diffusion

Les images de diffusion en **IRM**, quant à elles, mesurent l'atténuation du signal liée au mouvement brownien de l'eau (le mouvement microscopique aléatoire des molécules lié à la chaleur du tissu). La diffusion de l'eau est approchée par le coefficient de diffusion apparent (*apparent diffusion coefficient (ADC)*).

Comme les séquences de diffusion ne sont pas sensibles au mouvement brownien lorsqu'il est perpendiculaire à la direction du gradient de champ magnétique appliqué, pour obtenir une

image sensible à l'amplitude du mouvement brownien, mais indépendante de sa direction, l'acquisition est répétée de façon à ce que les gradients soient appliqués sur au moins trois directions orthogonales. Les images résultantes sont moyennées en une image isotropique pondérée en diffusion (*diffusion weighted images* (DWI)) (LUYPAERT et collab. [2001]).

Les techniques appelées *diffusion tensor images* (DTI) appliquent des gradients avec un nombre suffisant (≥ 6 , en pratique généralement entre 15 et 30) de directions pour mesurer l'anisotropie de diffusion due à certains obstacles à la diffusion de l'eau qui possèdent une orientation spatiale (ex. les axones, les fibres musculaires, etc.), favorisant la diffusion dans une direction particulière (MUKHERJEE et collab. [2008]).

La spectroscopie

La spectroscopie par résonance magnétique mesure la concentration des métabolites chimiques dans un volume choisi de tissu cérébral. La réduction du pic de *N-acétylaspartate* (NAA) et du rapport NAA / Créatine (Cr) observé dans les tumeurs cérébrales indique une diminution de la viabilité et du nombre de neurones. Une augmentation du pic de *Choline* (Cho), ou une augmentation des rapports Cho / NAA ou Cho / Cr est associée à une augmentation du renouvellement des membranes cellulaires et à une densité cellulaire plus élevée, en raison de la prolifération des cellules tumorales. La concentration de choline est liée au grade des tumeurs et à leur densité cellulaire (BRANDÃO et CASTILLO [2016]).

1.1.2 Généralités sur les tumeurs du système nerveux central

Selon l'*organisation mondiale de la santé* (OMS), il existe plus de 120 types de tumeur du système nerveux central (LOUIS et collab. [2016]). La classification de l'OMS de 2016 intègre le phénotype et le génotype des tumeurs envisageant la définition d'entités nosologiques correspondant à un groupe homogène de lésions. Ces définitions plus étroites des tumeurs visent l'amélioration de la prise en charge des patients et la détermination plus précise du pronostic et de la réponse au traitement. Le diagnostic d'une tumeur se fait grâce à 4 niveaux d'information (figure 1.4) :

1. les aspects histologiques permettant d'établir le type histologique et le grade de malignité;
2. les résultats de tests moléculaires pertinents à la classification des lésions dans une des entités nosologiques;
3. le diagnostic intégré à partir de l'information des deux examens précédents;
4. les altérations moléculaires pouvant donner un pronostic sans influence sur le diagnostic nosologique.

La figure 1.5 illustre l'exemple du chemin de classification des gliomes infiltrants. Les gliomes infiltrants présentent un modèle de croissance infiltrant. Les cellules tumorales (une seule cellule ou un petit groupe de cellules) envahissent le parenchyme cérébral. La capacité d'une tumeur de diffuser à partir d'une seule cellule qui migre sur de longues distances est une caractéristique typique et exclusive des gliomes (HUTTNER [2017]).

Dans un premier temps, l'apparence histologique des gliomes infiltrants est analysée : astrocytaire, oligodendrogiale ou oligoastrocytaire. Le système de gradation des tumeurs, variant de I (moins maligne) à IV (plus maligne), repose sur des critères histologiques incluant densité cellulaire, atypies nucléaires, prolifération de la vascularisation et nécrose. Les gliomes infiltrants incluent les tumeurs de grade II – gliomes de bas-grade, Grade III – tumeurs anaplasique; et Grade IV – *glioblastome multiforme* (GBM) et gliomes infiltrants pédiatriques (dont le *GITC*).

Dans un deuxième niveau, les gliomes infiltrants présentent fréquemment la mutation de l'enzyme *isocitrate déhydrogénase* (IDH). En général, la présence de la mutation de l'enzyme IDH est associée à un meilleur pronostic (REIFENBERGER et collab. [2016]). Les mutations de l'IDH sur les astrocytomes de grades II et III sont souvent accompagnées d'autres mutations, comme par exemple une perte de l'activité nucléaire de l'enzyme ATRX (un régulateur de la transcription de

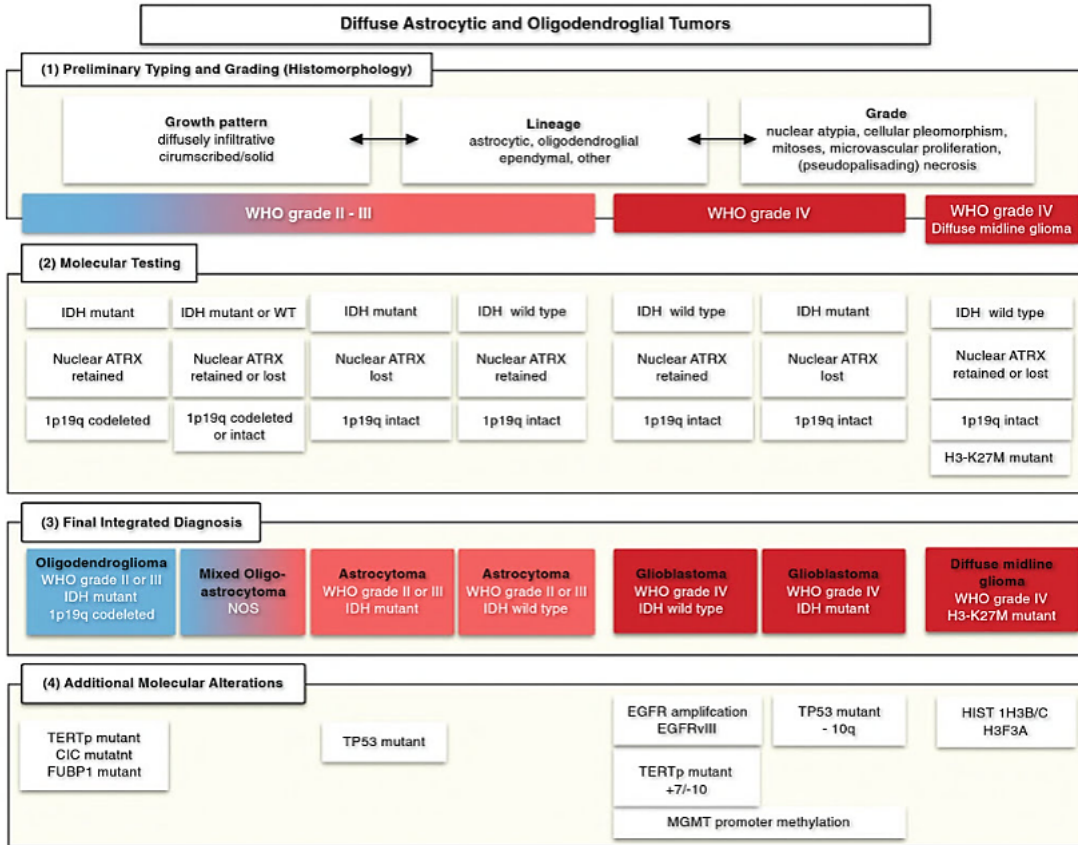


FIGURE 1.4 – Structure en 4 niveaux prévue pour la classification des tumeurs du système nerveux central selon l’OMS en 2016. Source : Anita Huttner (HUTTNER [2017])

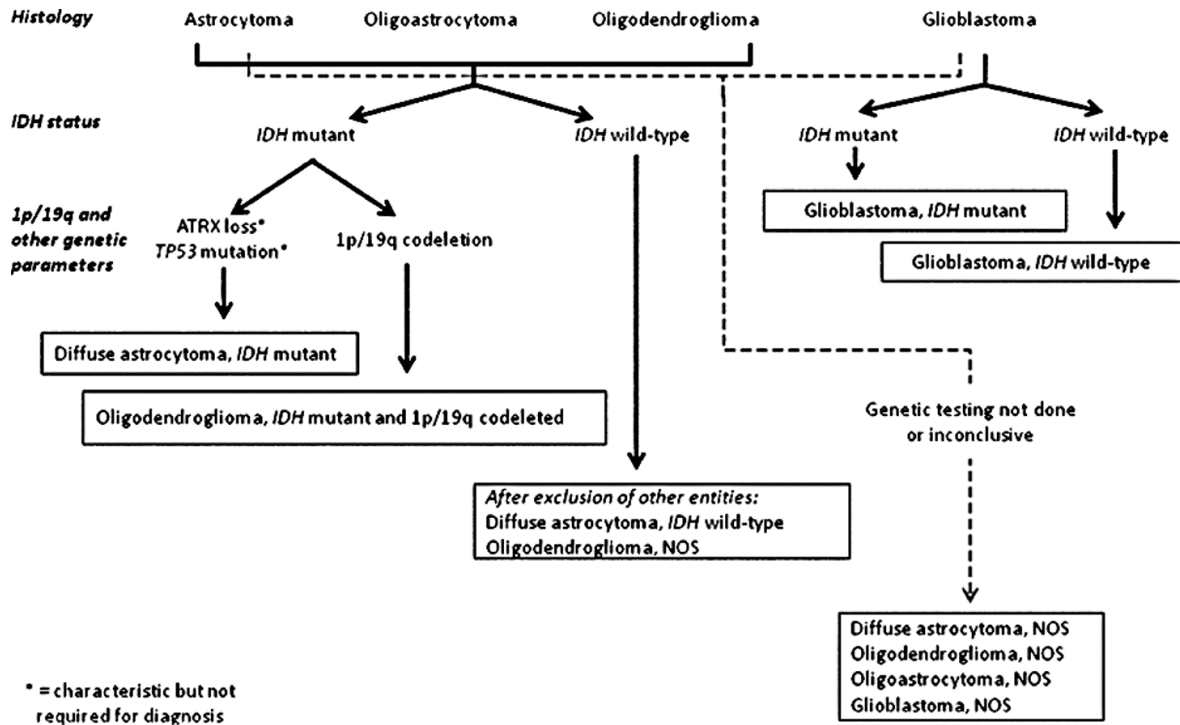


FIGURE 1.5 – Chemin de classification des gliomes infiltrants selon l’OMS 2016. Source : David N. Louis (LOUIS et collab. [2016])

l'ADN) et / ou des mutations de la TP53 impliquées dans le contrôle de la mort cellulaire en présence de stress ou de dégradation génétique (CALMON [2017]). D'un autre côté, la combinaison de la mutation de l'enzyme IDH associée à une codélétion des chromosomes 1p/9q définit les oligodendrogliomes et les oligodendrogliomes anaplasiques. De plus, plusieurs corrélations ont déjà été établies, montrant que la perte 1p/9q semble ne pas co-exister avec une mutation de la protéine p53 (TP53) ou une amplification récepteur du facteur de croissance épidermique (EGFR).

Finalement, dans un quatrième niveau, la présence de méthylation du promoteur du gène méthylguanine méthyltransférase (MGMT), par exemple, peut guider le choix d'un traitement, car ce gène est marqueur de bon pronostic des GBM traités par chimiothérapie, spécialement la chimiothérapie alkylante (THON et collab. [2013]).

1.1.3 Applications de la radiomique en IRM pour la neuro-oncologie

L'IRM est un outil important de diagnostic et la modalité d'imagerie de préférence en neuro-oncologie, car non-ionisante pour le patient et la qualité des images est élevée. De plus, avec les différentes séquences, l'hétérogénéité des tumeurs est bien représentée sur les images d'IRM et traduisent les différentes propriétés physiopathologiques.

État de l'art

Les tableaux 1.1, 1.2, 1.3, 1.4 et 1.5 présentent le type de tumeur du système nerveux central, le nombre de tumeurs incluses dans l'étude, le sujet d'investigation et les séquences d'IRM utilisées dans les 89 publications évoquant des études radiomiques en IRM pour la neuro-oncologie. Dans ces tableaux, le symbole "?" est utilisé quand l'information n'a pas été trouvée et l'acronyme "nc" quand l'information ne convient pas à l'article.

Environ la moitié des articles (41) traitent de GBM, le nombre de sujets est assez faible (en moyenne 142), et les séquences d'IRM les plus utilisées sont les séquences conventionnelles (T1w, T2w, T1c et FLAIR). Concernant les sujets abordés, ils se divisent majoritairement en trois groupes :

- **prédiction de réponses à un traitement et analyse de survie (28 études)** : s'intéressent majoritairement à la prédiction de la survie globale et/ou de la survie sans progression, mais aussi à l'infiltration des tumeurs et la pseudoprogession.
- **prédiction de l'histologie ou du grade de la tumeur (17 études)** : s'intéressent surtout à la distinction des gliomes de bas grade et de haut grade, et à la différenciation entre GBM et d'autres types de tumeurs comme les lymphomes et les métastases. Le but est notamment une meilleure prise en charge des patients.
- **prédiction des informations moléculaires associées à la tumeur (25 études)** : s'intéressent à la prédiction de mutation IDH, de codélétion 1p/9q, d'expression EGFR, de sécrétion de Periostin, de méthylation MGMT, de mutation ATRX et de mutation TP53. Ces informations moléculaires peuvent guider le choix des thérapies.

La liste ci-dessus représente un bon panorama des principales interrogations cliniques auxquelles peut répondre la radiomique. Ces questions sont, en fait, abordées depuis longtemps, avant l'introduction du terme "radiomique" (DIEHN et collab. [2008]; GANESHAN et collab. [2010]; JOHNSON et collab. [2006]; NIE et collab. [2008]). L'introduction de ce terme est, néanmoins pertinente car la qualité de l'imagerie médicale, la standardisation des protocoles, la capacité de calcul et de stockage des ressources informatiques, ainsi que l'efficacité des méthodes d'apprentissage automatique ne cessent d'augmenter depuis les dix dernières années. Cette évolution peut permettre la création de modèles à la fois complexes et performants.

A présent, en mai 2019, il n'existe aucune étude sur la radiomique pour l'étude des patients du GTC (recherche dans PubMed avec ((diffuse AND midline AND "glioma") OR ("diffuse" AND "intrinsic" AND "pontine" AND "glioma") OR "DIPG") AND radiomics), gliome pédiatrique étudié dans cette thèse et qui sera présenté dans le chapitre 2.

Référence	Type de tumeur du SNC	Nombre de tumeurs	Question	Séquences
LEE et collab. [2015]	GBM	65	1. Survie globale, 2. expression EGFR	T1c, FLAIR
GROSSMANN et collab. [2016]	GBM	142	1. Sous-types moléculaires, 2. Survie globale	T1c, FLAIR
KOTROTSOU et collab. [2016]	Tout	nc	Revue générale	nc
COLEN et collab. [2016]	Tout	nc	Revue : Classification OMS 2016 dans le contexte de la radiomique	nc
MCGARRY et collab. [2016]	GBM	81	Survie globale	DWI, T1w, T1c, FLAIR
KICKINGEREDER et collab. [2016]	GBM	172	1. Survie globale, 2. Survie sans progression	T1w, T1c, FLAIR
HSU et collab. [2016]	GBM	nc	Revue générale	nc
PRASANNA et collab. [2016]	Tout	42	Nécrose vs. progression	T1w
PRASANNA et collab. [2017]	GBM	65	Survie > 18 mois vs. Survie < 7 mois	T1w, T2w, T1c, FLAIR
YU et collab. [2017a]	Gliomes de bas grade	92	mutation IDH1	?
YU et collab. [2017b]	Gliomes de bas grade	110	mutation IDH1	FLAIR
LOPEZ et collab. [2017]	GBM	17	Relations avec le volume métabolique tumoral et volume-cible clinique	Spectroscopie, T1c, FLAIR
ABROL et collab. [2017]	Tout	nc	Revue générale	nc
GALLDIKS et LANGEN [2017]	Tout	nc	Revue générale	TEP-IRM
HYARE et collab. [2017]	GBM	nc	Revue générale	nc
LEE et collab. [2017]	GBM	45	Stabilité des indices vis-à-vis de la segmentation	T1c, FLAIR
GROSSMANN et collab. [2017]	GBM	291	1. Survie globale, 2. Survie sans progression	T1w, T2w, T1c, FLAIR
LI et collab. [2017b]	Gliomes de bas grade	151	mutation IDH1	T1c, FLAIR

TABLEAU 1.1 – Première partie : État de l'art des études radiomiques en IRM en neuro-oncologie.

Référence	Type de tumeur du SNC	Nombre de tumeurs	Question	Séquences
LI et collab. [2018b]	Gliomes de bas grade	270	Expression EGFR	?
LEHRER et collab. [2017]	Gliomes de bas grade	57	Profil protéomique	T1w, T2w, T1c, FLAIR
SHUKLA et collab. [2017]	GBM	nc	Revue générale	nc
LAO et collab. [2017]	GBM	112	Survie globale	T1w, T2w, T1c, FLAIR
LI et collab. [2017c]	GBM	96	Age < 45 ans vs. âge ≥ 45 ans	T1c
XI et collab. [2018]	GBM	98	méthylation du promoteur du gène MGMT	T1w, T2w, T1c
KICKINGEREDER et collab. [2017]	GBM	181	1. Survie globale, 2. Survie sans progression	T1w, T2w, T1c, FLAIR
LIU et collab. [2017]	Gliomes	?	mutation IDH1	?
CHO et PARK [2017]	Gliomes	300	Gliome de bas grade vs. gliome de haut grade	T1w, T2w, T1c, FLAIR
WU et collab. [2017]	Gliomes	129	Segmentation de tumeurs	T2w
FETIT et collab. [2018]	Tumeur cérébrales pédiatriques	134	Médulloblastome vs. astrocytomes pilocytiques vs. épendymomes	T1w, T2w
LI et collab. [2017a]	GBM	92	Survie globale	T1w, T2w, T1c, FLAIR
CHEN et collab. [2018b]	Lymphome et GBM	96	Lymphome vs. GBM	T1c
SHOFTY et collab. [2018]	Gliomes de bas grade	47	Codélation 1p/19q	T1w, T1c, FLAIR
ARTZI et collab. [2018]	Gliomes de haut grade et métastases	102	Tissu tumoral vs. Tissu non-tumoral	T1w, T2w, T1c, FLAIR, DTI, DCE
ZHANG et collab. [2018]	Gliomes de bas grade	103	1. Mutation IDH1, 2. Mutation TP53	T1w, T2w, T1c, FLAIR
LI et collab. [2018a]	Gliomes de bas grade	186	Mutation ATRX	T1w, T2w, T1c, FLAIR

TABLEAU 1.2 – Deuxième partie : État de l’art des études radiomiques en IRM en neuro-oncologie.

Référence	Type de tumeur du SNC	Nombre de tumeurs	Question	Séquences
KANG et collab. [2018]	Lymphome et GBM	112	Lymphome vs. GBM	Diffusion et anatomique (?)
PEEKEN et collab. [2018]	GBM	189	1. Survie globale, 2. Survie sans progression	T1w, T2w, T1c, FLAIR, Diffusion
RATHORE et collab. [2018]	GBM	90	Infiltration	T1w, T2w, T1c, FLAIR, DTI, DSC
KICKINGEREDER et ANDRONESI [2018]	Tout	nc	Revue : spectroscopie, mutation IDH1, réponse aux thérapies, survie	nc
LI et collab. [2018d]	GBM	193	Méthylation du promoteur du gène MGMT	?
TIAN et collab. [2018]	Gliomes	153	1. Gliome de bas grade vs. gliome de haut grade, 2. Grade III vs. grade IV	T1w, T2w, T1c, DWI, ASL
VERDUIN et collab. [2018]	GBM	nc	Revue : réponse aux traitements	nc
SUH et collab. [2018]	Lymphome et GBM	77	Lymphome vs. GBM	T2w, T1c, FLAIR
PARK et KIM [2018]	Tout	nc	Revue générale	nc
DIAMANDIS et collab. [2018]	Gliomes	65	1. Codélation 1p/19q, 2. Mutation IDH1/2	Spectroscopie
LU et collab. [2018]	Gliomes	284	1. Codélation 1p/19q, 2. Mutation IDH1/2	T1w, T2w, T1c, FLAIR
CHEN et collab. [2018a]	Gliomes	274	Gliome de bas grade vs. gliome de haut grade	T1w, T2w, T1c, FLAIR
COQUERY et collab. [2018]	GBM	19 (pré-clinique)	Réponse aux thérapies	T2w, Diffusion, ASL
LOHMANN et collab. [2018]	Gliomes de haut grade	nc	Revue : TEP-IRM	nc
YANG et collab. [2018]	Gliomes	20	Stabilité des indices vis-à-vis l'acquisition et reconstruction d'images	T1w, T2w
CHADDAD et collab. [2018]	GBM	40	1. Survie globale, 2. Survie sans progression	T1w, FLAIR
LY et GERSTNER [2018]	Tout	nc	Revue générale	nc
WANG et collab. [2018]	Gliomes	12	1. Survie globale, 2. Réponse aux traitements	T1w, T2w, DCE, DWI
LIU et collab. [2018c]	Gliomes de bas grade	286	Epilepsie	T2w

TABLEAU 1.3 – Troisième partie : État de l'art des études radiomiques en IRM en neuro-oncologie.

Référence	Type de tumeur du SNC	Nombre de tumeurs	Question	Séquences
WEI et collab. [2019]	Astrocytomes	136	Méthylation du promoteur du gène MGMT	T1c, FLAIR, ADC
ZINN et collab. [2018]	GBM	133	Expression periostin	T1c, FLAIR
ARITA et collab. [2018]	Gliomes	199	1. Mutation IDH, 2. Promoteur TERT	T1w, T2w, T1c, FLAIR
HAN et collab. [2018]	Gliomes de bas grade	277	Codélation 1p/19q	?
KIM et collab. [2018a]	GBM	61	Pseudoprogression vs. Progression	T1c, FLAIR, ADC, CBV
DITMER et collab. [2018]	Gliomes	94	Gliome de bas grade vs. Gliome de haut grade	T1c
LIU et collab. [2018a]	Gliomes de haut grade	68	Survie globale	fonctionnelle
REN et collab. [2019]	Gliomes de bas grade	57	1. Mutation IDH1, 2. Expression ATRX	T2w, FLAIR, ASL, DWI
KIM et collab. [2018b]	Lymphome et GBM	143	Lymphome vs. GBM	T2w, T1c, DWI
WANG et collab. [2019]	Gliomes	85	Grade	T2w, T1c, ADC
BAE et collab. [2018]	GBM	217	1. Survie globale, 2. Survie sans progression	T1w, T2w, FLAIR, DTI
KUTHURU et collab. [2018]	Gliomes de bas grade	108	1. Codélation 1p/19q, 2. Mutation IDH1	T1w, T2w, T1c, FLAIR
SU et collab. [2019]	Gliomes	220	1. Grade, 2. Prolifération	T2w, T1c, FLAIR, DWI, ASL
LOTAN et collab. [2019]	Gliomes	nc	Revue générale	nc
LIU et collab. [2018b]	Gliomes de bas grade	300	Survie sans progression	T2w
WU et collab. [2018]	Lymphome et GBM	?	Lymphome vs. GBM	T1c
KIM et collab. [2019]	GBM	93	Survie globale	ADC
DONG et collab. [2018]	Pilocytic astrocytomes et GBM	66	Pilocytic astrocytomes vs. GBM	T1c
LI et collab. [2018c]	GBM	225	Mutation IDH	T1w, T2w, T1c, FLAIR

TABLEAU 1.4 – Quatrième partie : État de l’art des études radiomiques en IRM en neuro-oncologie.

Référence	Type de tumeur du SNC	Nombre de tumeurs	Question	Séquences
SANGHANI et collab. [2018]	GBM	163	Survie globale	?
YIN et collab. [2018]	Gliomes de haut grade et métastases	86	Gliome de haut-grade vs. métastases	FLAIR
KNIEP et collab. [2019]	Métastases	658	Origine des métastases	T1w, T1c, FLAIR
ARTZI et collab. [2019]	GBM et métastases	439	GBM vs. Sous-types de métastases (sein, poumon, autres)	T1c
CHEN et collab. [2019b]	GBM	127	Survie globale	T1c
RUDIE et collab. [2019]	Tout	nc	Revue générale	nc
NIE et collab. [2019]	Gliomes de haut grade	68	Survie globale	T1c, DTI, fonctionnelle
WU et collab. [2019]	Gliomes infiltrants	126	Mutation IDH	T1w, T2w, T1c, FLAIR
TIXIER et collab. [2019]	GBM	159	Survie globale	T1w, T1c, FLAIR
DELLA SETA et collab. [2019]	Métastases	48	Survie globale	T1c
ZHANG et collab. [2019]	GBM	105	Survie globale	T1w, T2w, T1c, FLAIR
VAMVAKAS et collab. [2019]	Gliomes	40	Gliome de bas grade vs. gliome de haut grade	T1w, T2w, T1c, FLAIR, DTI
YUN et collab. [2019]	Lymphome et GBM	195	Lymphome vs. GBM	T1c, Diffusion
LIAO et collab.	GBM	137	Survie globale	FLAIR
PRASANNA et collab. [2019]	Gliomes	331	1. Segmentation de tumeurs, 2. Survie globale	T2w, T1c, FLAIR
CHADDAD et collab. [2019]	GBM	nc	Revue générale	nc

TABLEAU 1.5 – Cinquième partie : État de l’art des études radiomiques en IRM en neuro-oncologie.

Principales découvertes

Les résultats de publications radiomiques en IRM en neuro-oncologie sont en général très prometteurs. La radiomique s'avère ainsi, une source de biomarqueurs non-invasifs, calculés à partir de l'IRM, pour la prédiction de caractéristiques histologiques, moléculaires et pronostiques des tumeurs cérébrales. Quelques exemples de résultats sur les principaux sujets d'investigation sont présentés dans ce paragraphe.

CHEN et collab. [2018b] ont réussi à classifier les lymphomes et les GBM avec une exactitude de 0,991 en validation croisée et de 0,982 en cohorte de test indépendante en utilisant des indices radiomiques calculés sur les images T1c. En utilisant une cohorte plus large et en ajoutant des indices dérivées des images T2w et DWI, KIM et collab. [2018b] ont obtenu une aire sous la courbe ROC (AUC) de 0,956 en cohorte de test indépendante.

L'étude de CHO et PARK [2017] utilisant 300 patients a différencié les gliomes de bas grade des gliomes de haut grade avec une exactitude de 0,9 en utilisant seulement les séquences d'IRM conventionnelles. Plus tard, TIAN et collab. [2018] ont réalisé la même tâche avec moins de patients (153) mais en ajoutant des informations extraites des images DWI et ASL, aboutissant à une exactitude de 0,968.

La plupart des études cherchant à prédire la mutation IDH ont obtenu des résultats prometteurs. WU et collab. [2019] ont trouvé un modèle, utilisant des indices radiomiques issus d'IRM conventionnelle pour les gliomes infiltrants qui prédit la mutation IDH avec une exactitude de 0,89 en test. De même, LI et collab. [2018c] ont atteint une exactitude de 0,97 pour les patients atteints de GBM et ZHANG et collab. [2018], une exactitude de 0,80 pour les patients atteints de gliomes de bas grade. Pour le cas de gliomes de bas grade, REN et collab. [2019] ont atteint récemment une exactitude de 0,95 en ajoutant des informations de perfusion.

Concernant la prédiction de survie, CHEN et collab. [2019b] ont récemment conclu qu'un modèle radiomique combinant des informations cliniques avec des indices calculés sur des images T1c peut prédire les patients survivant plus de 12 mois après le diagnostic avec une exactitude de 0,851 sur une cohorte de test indépendante. Pour les gliomes de bas grade, LIU et collab. [2018b] ont proposé une signature composée de 9 indices radiomiques calculés sur les images T2w pour la prédiction de survie sans progression. L'indice de concordance (indice C) obtenu est de 0,823 en cohorte de test indépendante.

Ainsi les développements en radiomique pour les tumeurs cérébrales avancent rapidement, toutes les publications étant comprises entre les années 2015 et 2019. Les résultats sont encourageants, cependant les techniques de mise en œuvre de la radiomique peuvent varier beaucoup entre les différentes études. Une clarification et homogénéisation des méthodologies adoptées sont envisagées pour mieux comparer les résultats et transférer les modèles à différentes cohortes. Le paragraphe suivant présente les principales étapes de la mise en œuvre de la radiomique.

1.2 Étapes de la radiomique

Les études radiomiques regroupent différentes disciplines, telles que les techniques d'acquisition des images, les statistiques, la vision par ordinateur et l'apprentissage automatique. Les étapes classiques d'une étude radiomique sont illustrées sur la figure 1.6 et sont détaillées dans la suite du paragraphe.

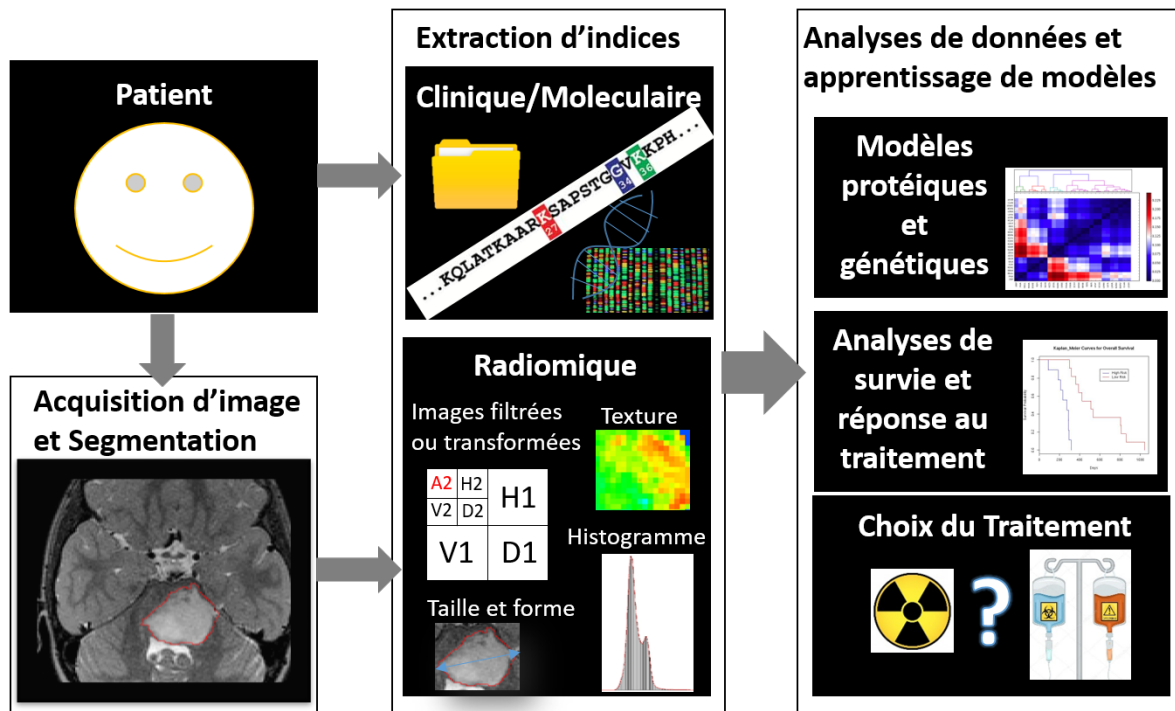


FIGURE 1.6 – Schéma des principales étapes d'une étude radiomique.

1.2.1 Standardisation d'images

Les techniques d'IRM utilisées en clinique montrent une large variation des paramètres d'acquisition comme les antennes utilisées pour l'acquisition, la résolution spatiale, l'administration d'agents de contraste, le type de séquence (temps d'écho, temps de répétition, entre autres) et les algorithmes de reconstruction de chaque constructeur. Tout cela peut influencer sur les intensités, le bruit et la texture des images. Ainsi, les indices radiomiques calculés sur les images peuvent refléter les variations des paramètres d'acquisition mêlées aux caractéristiques biologiques des tissus. Pour contourner ces variations, des efforts doivent être mis en place pour standardiser les protocoles d'acquisition au sein de la même institution, mais aussi de différentes institutions.

De plus, pour le cas particulier de l'IRM, les intensités des images sont arbitraires et elles varient même sur deux volumes d'images acquises pour un même patient, dans la même position, avec le même scanner et le même protocole d'acquisition. Quelques techniques de standardisation d'intensités des images IRM ont été proposées et seront discutées dans le chapitre 3.

1.2.2 Segmentation des tumeurs

La segmentation des régions tumorales est un autre aspect important des études radiomiques, car cette étape impacte principalement les indices de taille et de forme des tumeurs, mais aussi les indices de texture. Plusieurs méthodes automatiques et semi-automatiques de segmentation ont été proposées dans la littérature en vue de réduire le temps nécessaire au contourage et la charge de travail des opérateurs, augmenter la reproductibilité et la précision des tracés. Actuellement, les méthodes les plus performantes dans les compétitions internationales de segmentation (comme le MICCAI BraTS) sont les méthodes d'apprentissage profond. Néanmoins la vérité terrain des tracés est toujours la segmentation manuelle faite par un radiologue expérimenté. Cette vérité présente une variabilité inter et intra-opérateur rarement prise en compte dans ce type d'études.

Il n'existe pas de méthode de segmentation idéale et des études de stabilité et reproductibilité des indices radiomiques sont recommandées afin de sélectionner les indices les plus robustes.

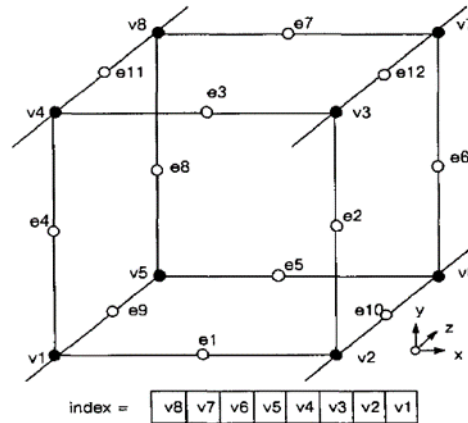


FIGURE 1.7 – Le nombre binaire associé à chaque cube par les valeurs des 8 sommet (V1, ..., V8). Les points (e1, .. e12) sont les possibles sommets des triangles de la maille de surface de la région. Source : William Lorensen (LORENSEN et CLINE [1987])

1.2.3 Indices radiomiques

Les indices quantitatifs extraits des régions délimitées dans les images, appelés indices radiomiques sont généralement :

- des indices de taille et forme de la tumeur ;
- des indices d'intensité de premier-ordre ;
- des indices de texture, calculés à partir des matrices de relations entre les voxels ;
- des indices d'ordre plus élevé calculés par exemple à partir des images transformées en ondelettes (SELVARAJ et GANESAN [2003]), ou filtrées par les convolutions de Laplace avec des noyaux gaussiens (SOTAK et BOYER [1989])

Un des problèmes de la radiomique est le manque de reproductibilité des indices, différentes équations mathématiques ayant été proposées pour chaque indice. Afin de standardiser les définitions des indices, une initiative internationale a été créée, la *image biomarker standardisation initiative* (IBSI) (ZWANENBURG et collab. [2016]).

Les indices de taille et de forme sont indépendants des valeurs des intensités dans l'image et peuvent ainsi être calculés sur un masque de segmentation de la tumeur. L'indice le plus simple est le volume (V) qui est estimé par le nombre de voxels inclus dans la région multiplié par le volume d'un voxel. Pour le calcul des autres indices de forme, l'IBSI recommande que les régions segmentées soient approximées par un maillage triangulaire pour limiter l'effet de volume partiel. Les sommets des triangles sont définis par l'algorithme introduit par LORENSEN et CLINE [1987], les *marching cubes* où un cube évolue sur la région et pour chaque position du cube, chacun des 8 sommets est mis à 0 s'il est hors de la région segmentée ou à 1 s'il est dans la région segmentée. Ainsi, si on définit un nombre binaire par les valeurs de chaque sommet juxtaposées, il y a $2^8 = 256$ valeurs possibles (comme le montre la figure 1.7), donc, la surface de la région peut intersecter le cube de 256 façons différentes. La taille du cube est choisie en fonction de la précision souhaitée. Sur la figure 1.7, les points notés (e1, ..., e12) sont les 12 sommets possibles des triangles. La figure 1.8 montre quelques exemples.

Quelques exemples d'indices de forme sont calculés à partir de cette représentation sont indiqués :

- l'aire de la surface (équation 1.1) est estimée comme étant égale à la somme des aires de chaque triangle i ($i = 1, 2, \dots, N_f$) de sommets a_i, b_i et c_i ,

$$A = \sum_{i=1}^{N_f} \frac{1}{2} |a_i b_i \times a_i c_i|; \quad (1.1)$$

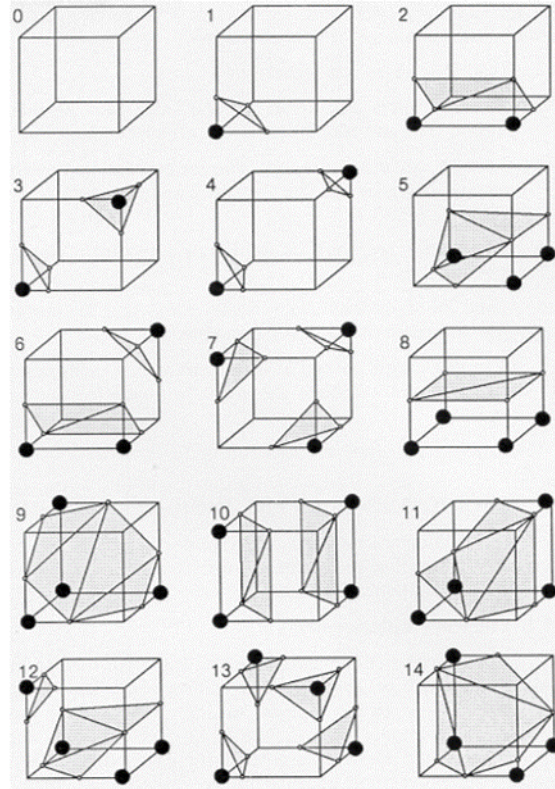


FIGURE 1.8 – Exemples de triangles définis par les valeurs des sommets des cubes. Les sommets à l’intérieur de la surface sont en gras.

- la sphéricité (équation 1.2) mesure la rondeur de la région par rapport à une sphère qui a la plus petite surface possible pour un volume déterminé. La valeur de la sphéricité varie ainsi entre 0 et 1 (1 indique une sphère parfaite),

$$sphéricité = \frac{\sqrt[3]{36\pi V^2}}{A}; \quad (1.2)$$

- le diamètre maximal est la distance maximale entre deux sommets parmi tous les sommets des triangles de la surface.

Les indices de premier ordre reposent sur les valeurs d’intensité des voxels et généralement ne demandent pas la discrétisation des valeurs. Quelques exemples sont la valeur moyenne d’intensité, la valeur minimale, maximale, médiane, l’écart-type, le kurtosis, entre autres. Les formulations mathématiques sont les formulations classiques des statistiques.

Les indices de deuxième ordre, ou indices de texture, expriment l’hétérogénéité spatiale en comparant les valeurs des intensités des voxels avec celles de leurs voisins. Pour cela, les images sont transformées en matrices qui expriment les caractéristiques recherchées et sont utilisées pour calculer des indices. Des exemples de telles matrices sont la **matrice de cooccurrence des niveaux de gris (GLCM)** (HARALICK et collab. [1973]), la **gray level run-length matrix (GLRLM)** (GALLOWAY [1975]), la **gray level size zone matrix (GLSZM)** (THIBAUT et collab. [2009]), entre autres. Avant de calculer ces matrices, les intensités sont ré-échantillonnées afin de réduire le nombre de niveaux de gris possibles (NN) pour réduire le temps de calcul et l’impact du bruit.

La **GLCM** est la matrice de probabilité de deuxième ordre de NN lignes et NN colonnes. La valeur attribuée à la position (i, j) est le nombre d’occurrences d’un voxel d’intensité i à une distance δ dans la direction θ d’un voxel d’intensité j ($GLCM(i, j|\delta, \theta)$). Pour $\delta = 1$, θ peut avoir 13 valeurs différentes dans une configuration 3D, et 4 valeurs différentes dans une configuration 2D. Par ailleurs, $\delta = 1$ est la valeur la plus couramment utilisée. Un exemple en 2D est illustré sur la figure 1.9. Pour fusionner les indices de texture de la **GLCM**, calculés pour chaque direction θ , six

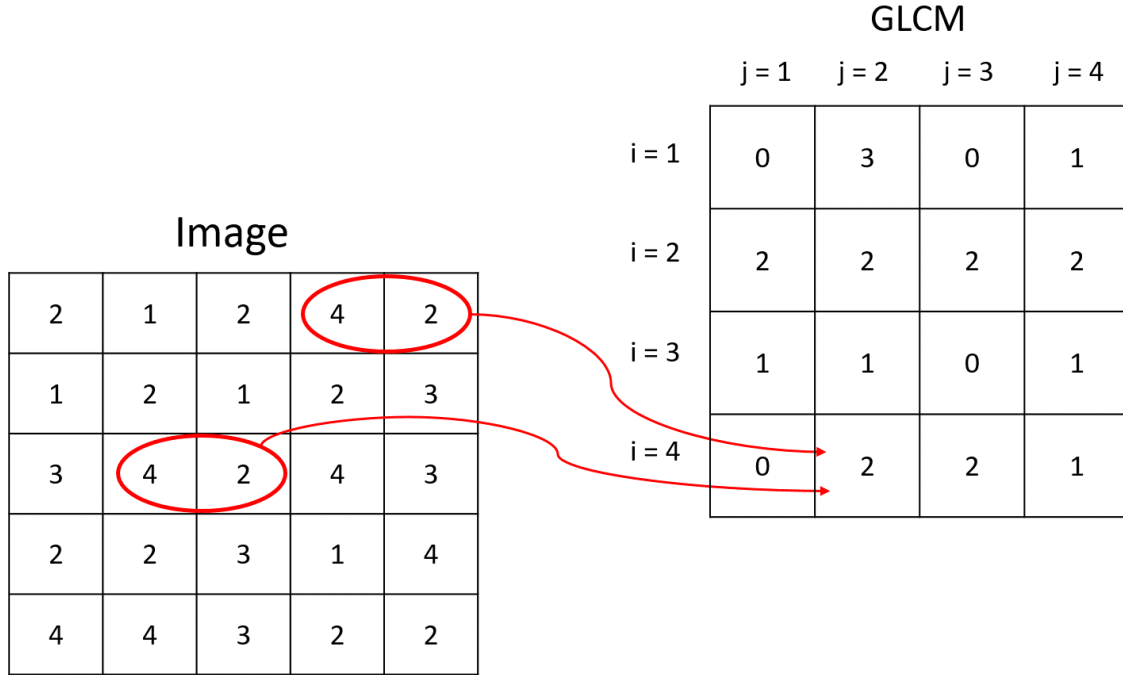


FIGURE 1.9 – Exemple d’une image 5x5 rééchantillonnée avec $NN = 4$ niveaux d’intensité et de la matrice GLCM associée pour $\delta = 1$ et $\theta = 0^\circ$ (direction horizontale).

méthodes ont été proposées selon l’**IBSI**. Pour le logiciel utilisé dans cette thèse, Pyradiomics (**VAN GRIETHUYSEN et collab. [2017]**), les indices sont calculés pour chacune des 13 directions et ensuite la moyenne des 13 valeurs obtenues est faite. Ainsi, aucune direction n’est privilégiée.

Les indices de texture quantifient les co-occurrences d’intensités avec différentes pondérations. Quelques exemples sont présentés (formulations recommandées par **IBSI**), en considérant que les probabilités jointes $p(i, j)$ sont égales à la valeur de la $GLCM(i, j)$ divisée par la somme de toutes les valeurs de la GLCM :

- l’entropie (équation 1.3) mesure la variabilité d’intensité dans les voisinages des voxels (ϵ est un paramètre de régularisation),

$$entropie = - \sum_{i=1}^{NN} \sum_{j=1}^{NN} p(i, j) \log_2(p(i, j) + \epsilon); \quad (1.3)$$

- le contraste (équation 1.4) mesure la variation locale d’intensité. Une valeur de contraste élevée implique que les voxels voisins ont des différences importantes de valeurs d’intensité,

$$contraste = \sum_{i=1}^{NN} \sum_{j=1}^{NN} (i - j)^2 p(i, j); \quad (1.4)$$

- l’intensité moyenne (équation 1.5) est la somme pondérée des probabilités jointes,

$$moyenne = \sum_{i=1}^{NN} \sum_{j=1}^{NN} p(i, j) i. \quad (1.5)$$

La **GLRLM** caractérise les longueurs de plage de niveaux de gris qui sont mesurées par la longueur en voxels. La valeur de la position (i, j) pour un angle θ est le nombre de longueurs de plage de valeur d’intensité i et de longueur j ($GLRLM(i, j|\theta)$). Ainsi les dimensions de la matrice **GLRLM** sont NN et la dimension de l’image dans la direction θ . Un exemple 2D est illustré sur la figure 1.10. Les considérations et choix pour la fusion d’indices calculés à différentes valeurs de θ sont faits pour la matrice **GLRLM** de façon identique à ceux faits pour la matrice **GLCM**.

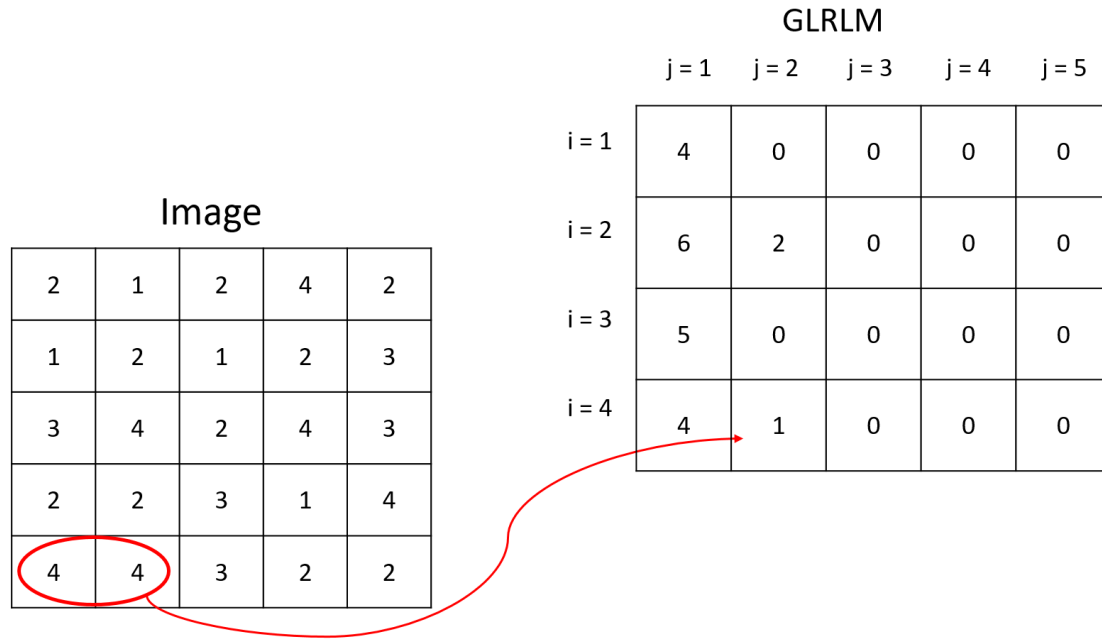


FIGURE 1.10 – Exemple d’une image 5x5 rééchantillonnée avec $NN = 4$ niveaux d’intensité et de la matrice GLRLM associée pour $\theta = 0^\circ$ (direction horizontale).

Quelques exemples d’indice qui peuvent être calculés à partir de la matrice **GLRLM** sont présentés (formulations recommandées par **IBSI**), où r_i est la somme des valeurs de la matrice **GLRLM** sur la ligne i , r_j est la somme des valeurs de la matrice **GLRLM** sur la colonne j , et N_r est le nombre de longueurs possibles :

- Le *short run emphasis (SRE)* (équation 1.6) met l’accent sur les plages courtes de niveaux de gris,

$$SRE = \frac{\sum_{j=1}^{N_r} \frac{r_j}{j^2}}{N_r}; \quad (1.6)$$

- Le *long run emphasis (LRE)* (équation 1.7) met l’accent sur les plages longues de niveaux de gris,

$$LRE = \frac{\sum_{j=1}^{N_r} j^2 r_j}{N_r}; \quad (1.7)$$

- Le *gray level non-uniformity (GLN)* (équation 1.8) mesure la similarité des intensités de gris dans l’image. Une valeur faible de GLN reflète une distribution plus équilibrée des valeurs d’intensité,

$$GLN = \frac{\sum_{i=1}^{NN} r_i}{N_r}. \quad (1.8)$$

La matrice **GLSZM** quantifie les régions de voxels connectés de même intensité. A la différence des matrices **GLCM** et **GLRLM**, cette matrice ne dépend pas de l’angle traduisant la direction selon laquelle est fait le calcul. Sa taille est de NN par le nombre maximal de voxels connectés. La valeur de la position (i, j) de la matrice **GLSZM** est le nombre de zones d’intensité i et taille j . Un exemple est illustré sur la figure 1.11.

Les indices de la **GLSZM** sont équivalents aux indices de la **GLRLM**, caractérisant les zones à la place des plages de longueur selon une direction donnée. Tous ces indices peuvent être calculés de façon différente par les différents logiciels proposés. **NIOCHE et collab. [2018]** ont détaillé ces

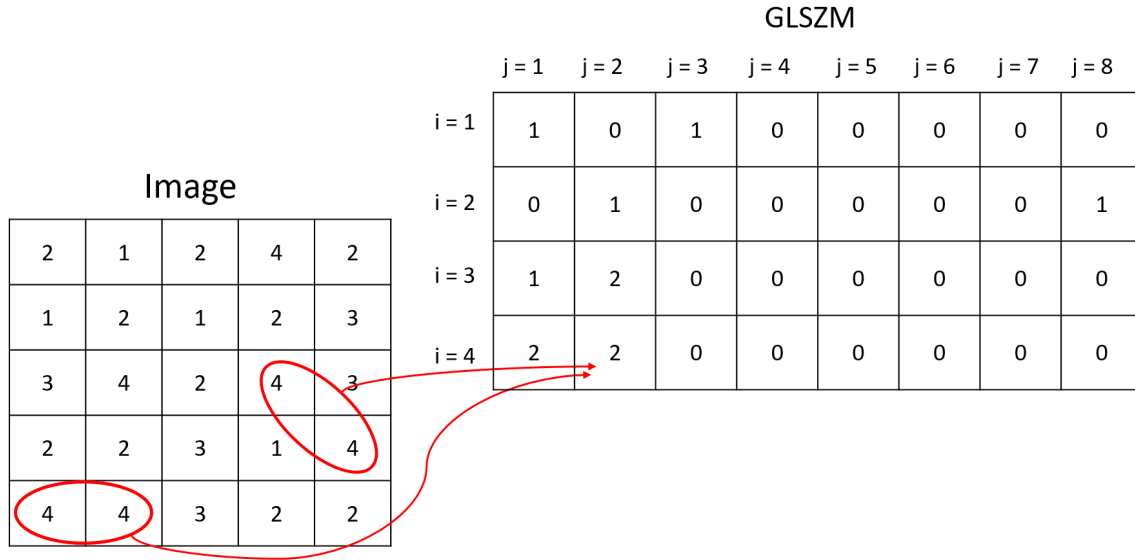


FIGURE 1.11 – Exemple d’une image 5x5 rééchantillonnée avec NN = 4 niveaux d’intensité et de la matrice GLSZM associée.

différences, notamment pour PyRadiomics et LIFEx (NIOCHE et collab. [2018]). Dans ce travail de thèse, la librairie PyRadiomics a été utilisée pour les calculs des indices. Les formules des indices utilisées par PyRadiomics sont présentées en annexe A.2.1.

Finalement, les indices de premier-ordre et de texture peuvent être aussi calculés sur des images transformées ou filtrées comme par exemple sur les images transformées en ondelette (autant de fois désirées) et sur la convolution de filtres laplaciens de noyau gaussien. Les images transformées en ondelette sont la décomposition de l’image en huit images (toutes les combinaisons d’application de filtre passe-bas ou passe-haut dans chaque direction). Les images filtrées par convolution de filtres laplaciens de noyau gaussien rehaussent les changements des valeurs de gris et sont sensibles aux contours.

1.2.4 Sélection d’indices

BELLMAN et COLLECTION [1961] ont démontré que la valeur prédictive d’un modèle peut décroître avec l’ajout d’indices si le nombre de sujets est maintenu constant (phénomène de malédiction de la dimensionnalité, aussi appelé effet de Hughes). Les analyses radiomiques utilisent des centaines (ou des milliers) d’indices radiomiques, parmi eux, se trouvent des indices intercorrélés et des indices qui n’ont aucune valeur prédictive. Ainsi, une étape importante du traitement des données est la sélection d’indices pour réduire la dimension du problème. De plus, les modèles constitués d’un faible nombre d’indices sont plus simples à expliciter

Dans la littérature, les méthodes de sélection d’indices sont divisées en trois catégories : les méthodes *Filter*, les méthodes *Wrapper* et les méthodes *Embedded*. Les méthodes *Wrapper* et *Embedded* sont dépendantes de la méthode d’apprentissage alors que les méthodes *Filter* ne le sont pas (PARMAR et collab. [2015]).

Méthodes *Wrapper*

Les méthodes *Wrapper* (KOHAVI et JOHN [1997]) sont des méthodes de recherche exhaustive. Elles cherchent dans tout l’espace d’indices, un sous-groupe d’indices pertinents et non-redondants. La performance (normalement en validation croisée) d’une méthode d’apprentissage particulière est utilisée pour mesurer la pertinence du sous-groupe d’indices. Ces méthodes sont coûteuses en ressources informatiques et peuvent sélectionner un sous-groupe d’indices spéci-

fique pour une méthode d'apprentissage, par conséquent, peu généralisable.

Deux exemples de méthodes *Wrapper* sont la méthode d'élimination progressive (en anglais, *backward elimination*) et la méthode de sélection ascendante (en anglais, *forward selection*). Dans l'élimination progressive, un modèle est créé avec tous les indices et à chaque itération l'indice qui fait décroître le moins la performance du modèle est éliminé, tandis que dans la sélection ascendante, un modèle est créé avec l'indice le plus prédictif et à chaque itération l'indice qui fait croître le plus la performance du modèle est ajouté.

Méthodes *Embedded*

Les méthodes *Embedded* intègrent la sélection d'indices dans le processus d'apprentissage et sont plus efficaces sur le plan informatique que les méthodes *Wrapper*. Cependant, ces méthodes sont aussi peu généralisables.

Les approches d'arbres de décision et les forêts aléatoires (BREIMAN [2001]), utilisent une méthode *Embedded* en sélectionnant un indice dans chaque étape du processus de croissance de l'arbre pour diviser l'ensemble d'échantillons en sous-ensembles plus petits. Le processus de génération de l'arbre de décision est ainsi également le processus de sélection des fonctionnalités.

Pour les régressions linéaires, les régularisations Ridge et *Least Absolute Shrinkage and Selection Operator* (LASSO) (TIBSHIRANI [1996]) sont deux exemples de méthodes *Embedded*.

La régression Ridge (HOERL et KENNARD [1970]) utilise la régularisation L2 pour pénaliser la régression quand les coefficients ($\beta = \beta_1, \dots, \beta_j$) du modèle sont en cours d'apprentissage. Un paramètre λ_R est utilisé pour régler l'intensité de la pénalisation. Avec $\lambda_R = 0$, aucune régularisation n'est faite, et au fur et à mesure que λ_R croît, les coefficients liés aux indices les moins contributifs sont davantage réduits. Même si les paramètres peuvent être réduits à presque zéro, la régression Ridge n'élimine aucun indice. Le problème à minimiser avec une régularisation L2 est rappelé par l'équation 1.9.

$$\hat{\beta}_{Ridge} = \underset{\beta \in \mathbb{R}}{\operatorname{argmin}} \|y - X\beta\|_2^2 + \lambda_R \|\beta\|_2^2 \quad (1.9)$$

Au contraire, le LASSO utilise la régularisation L1 pour pénaliser le modèle en imposant d'avoir un certain nombre de coefficients à zéro (équation 1.10). Ainsi, LASSO élimine les indices qui ne contribuent pas beaucoup au modèle lors de la création du modèle (détermination des coefficients).

$$\hat{B}_{LASSO} = \underset{B \in \mathbb{R}}{\operatorname{argmin}} \|y - XB\|_2^2 + \lambda_L \|B\|_1 \quad (1.10)$$

Finalement, une combinaison des deux régularisations peut être réalisée en faisant varier les valeurs de λ_L et λ_R de l'équation 1.11. Cette régularisation est appelée *Elastic Net*.

$$\hat{B}_{eNet} = \underset{B \in \mathbb{R}}{\operatorname{argmin}} \|y - XB\|_2^2 + \lambda_L \|B\|_1 + \lambda_R \|\beta\|_2^2 \quad (1.11)$$

Méthodes Filtre

Les méthodes Filtre sont des méthodes de classification d'indices basées sur des critères heuristiques, indépendants du modèle d'apprentissage. Ces méthodes sont efficaces sur le plan informatique et généralisables. Le critère de sélection est ainsi l'élément déterminant des méthodes Filtre de sélection d'indices. Les critères de sélection peuvent être divisés en deux catégories : les critères uni-variés et les critères multi-variés.

Dans le cas des critères uni-variés, uniquement la pertinence des indices est considérée en ignorant leur redondance, alors que les critères multi-variés étudient l'interaction des indices, pondérant la pertinence des indices avec leur redondance.

La pertinence des indices est une mesure de l'association de l'indice avec la variable cible (à prédire), tandis que la redondance des indices est la redondance présente dans un indice particulier par rapport à l'ensemble des indices déjà sélectionnés.

La pertinence des indices peut être quantifiée par différentes métriques, comme par exemple le test de Wilcoxon et l'index de Gini, deux critères uni-variés.

Le test de Wilcoxon est une méthode non paramétrique basée sur les rangs pour la comparaison des médianes de population des deux classes. Une valeur élevée de la statistique du test de Wilcoxon suggère des médianes différentes pour les deux classes. Si N est le nombre d'échantillons, n_m est le nombre d'échantillons de la classe m , r_{mi} est le rang de l'échantillon i de la classe m , μ_{rm} est la valeur moyenne des rangs des échantillons de la classe m , et μ_r est la valeur moyenne des rangs de tous les échantillons, le score de Wilcoxon est défini par l'équation 1.12 pour le cas d'une classification binaire.

$$W = (N - 1) \frac{\sum_{m=1}^2 n_m (\mu_{rm} - \mu_r)^2}{\sum_{m=1}^2 \sum_{i=1}^{n_m} (r_{mi} - \mu_r)^2} \quad (1.12)$$

L'index de Gini défini par l'équation 1.13 pour les classifications binaires, où p_i est la probabilité de la classe i . Une valeur faible d'index de Gini implique une bonne pertinence.

$$G(p_i) = p_i(1 - p_i) \quad (1.13)$$

Un exemple de critère multi-varié est le *minimum redundancy maximum relevance (mRMR)* (PENG et collab. [2005]), qui est une somme pondérée de la pertinence des indices et de leur redondance. Le score mRMR est calculé par l'équation 1.14.

$$\text{MRMR} = I(X_k; Y) - \frac{1}{\|S\|} \sum_{X_j \in S} I(X_k; X_j) \quad (1.14)$$

où $I(X_k; Y)$ est l'information mutuelle (COVER et THOMAS [2006]) entre l'indice X_k et la variable à prédire Y , le deuxième terme ($\sum_{X_j \in S} I(X_k; X_j)$) correspond à la redondance entre les indices X_k et X_j et S est le groupe d'indices sélectionnés.

1.2.5 Méthodes d'apprentissage automatique

En apprentissage automatique, les problèmes de régression sont séparés des problèmes de classification : les problèmes de prédiction d'une variable quantitative sont des problèmes de régression tandis que les problèmes de prédiction d'une variable qualitative sont des problèmes de classification. Pour la mise en œuvre de la radiomique dans un problème de classification ou de régression, une méthode supervisée capable d'apprendre une "réponse" à partir des données disponibles pour bien prédire les "réponses" associées à un nouveau jeu de données est nécessaire. Pour cela, plusieurs méthodes d'apprentissage automatique de classification et de régression provenant de différents domaines de l'informatique et des statistiques ont été proposées.

Méthodes d'apprentissage automatique pour les classifications binaires

Les classifications binaires, dont les réponses sont "oui" (1) ou "non" (0), sont très fréquentes dans les études radiomiques, comme la classification de patients présentant ou ne présentant pas la mutation d'un certain gène, des patients répondant ou ne répondant pas à un traitement, entre autres. La plupart des méthodes d'apprentissage automatiques proposées sont bien adaptées aux études de classification binaire. Parmi elles, se trouvent la régression logistique, la méthode des k plus proches voisins, la machine à vecteurs de support et les forêts aléatoires. Ces méthodes sont largement utilisées dans la littérature mais correspondent à des principes différents.

La régression logistique (en anglais, *Logistic regression (Logit)*) La régression logistique est une méthode linéaire de classification utilisée en statistiques, introduite initialement par BERKSON [1944]. Dans le cadre de la régression logistique binaire, la variable à prédire (0,1) (variable expliquée) est nommée y et $X = (x_1, x_2, \dots, x_j)$ définit le vecteur des J variables prédictives (variables

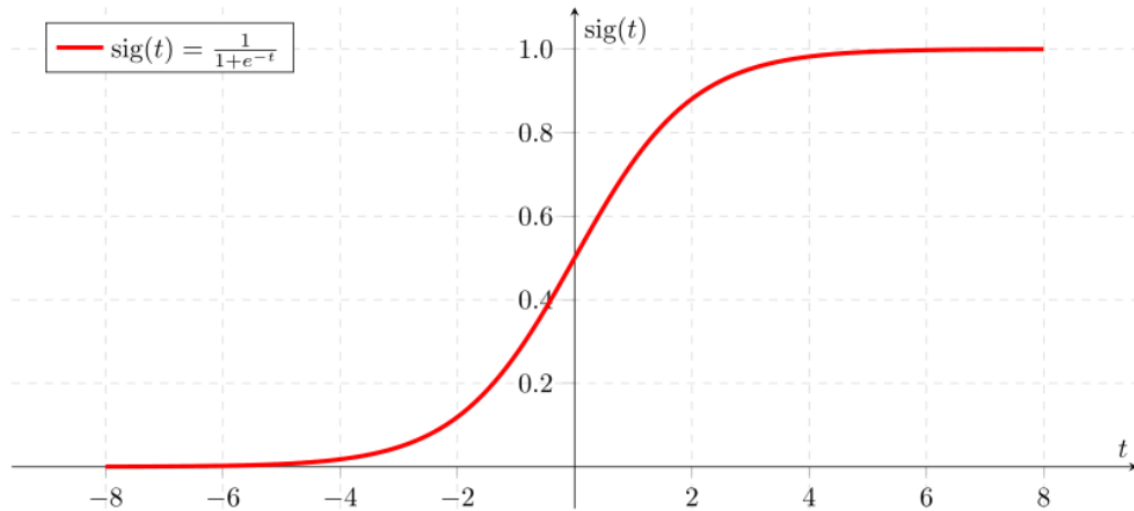


FIGURE 1.12 – Fonction Sigmoid (sig(t)). Source : *Caglar Subasi*

explicatives). La régression logistique vise à estimer les paramètres b de l'équation 1.15, où $P(1|X)$ est la probabilité *a posteriori* de $y = 1$ sachant les valeurs des variables X .

$$\ln\left(\frac{P(1|X)}{1 - P(1|X)}\right) = b_0 + b_1 x_1 + \dots + b_j x_j \quad (1.15)$$

Les paramètres b sont estimés par la maximisation de la fonction de vraisemblance (équation 1.16) de l'échantillon d'entraînement de taille n .

$$L(X, Y, b) = \prod_{i=1}^n P(1|X_i)^{y_i} (1 - P(1|X_i))^{1-y_i} \quad (1.16)$$

La vraisemblance décrit la plausibilité d'une valeur des paramètres d'un modèle, étant donné l'observation d'un certain nombre de réalisations d'une variable aléatoire.

Une fois les paramètres b établis, la fonction logistique est définie par l'équation 1.17. Elle se présente la forme d'une fonction sigmoïde (figure 1.12).

$$P(1|X, b) = \frac{1}{1 + e^{-(b_0 + b_1 x_1 + \dots + b_j x_j)}} \quad (1.17)$$

Une nouvelle entrée X_n est classifiée à 1 si $P(1|X_n, b) \geq 0,5$ et classifiée à 0 si $P(1|X_n, b) < 0,5$.

La régression logistique est adaptée pour les classifications binaires quand les variables prédictives ne sont pas trop corrélées entre elles. **TABACHNICK et FIDELL [2013]** indique que quand les coefficients de corrélation entre les variables sont inférieurs à 0,9, cette condition est satisfaite. Une autre condition de succès de cette méthode de classification est l'élimination des variables non-corrélées à la variable à prédire.

Les k plus proches voisins (en anglais, *k Nearest Neighbors* (KNN)) La méthode des k plus proches voisins (**ALTMAN [1992]**) est une méthode d'apprentissage supervisée non-paramétrique qui classe chaque nouvelle entrée en fonction de ses k plus proches voisins, dans l'espace des indices identifiés par l'apprentissage. La nouvelle entrée est classée dans la classe d'appartenance la plus fréquente de ses k plus proches voisins. La figure 1.13 montre un exemple de classification d'une nouvelle entrée dans le cas de 2 indices. Si $k = 3$, la nouvelle entrée est classée comme "rouge", tandis que si $k = 5$, la nouvelle entrée est classée comme "bleu". Une pondération de l'influence contributive des voisinages peut être aussi appliquée de façon que les plus proches voisins contribuent plus au choix de classification que les voisins plus éloignés.

La méthode des k plus proches voisins est adaptée une fois que les indices sont sélectionnés en fonction de leur relation avec la variable à prédire.

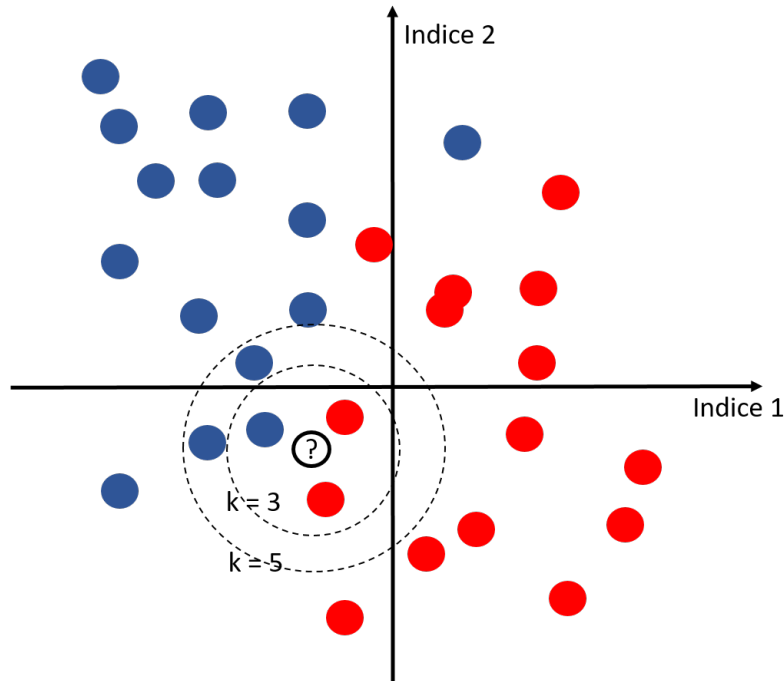


FIGURE 1.13 – Exemple de classification KNN avec 2 indices descriptifs, 2 classes (rouge et bleu) et $k = 3$ ou $k = 5$.

Les machines à vecteurs de support (en anglais, *Support Vectors Machine (SVM)*) La méthode des machines à vecteurs de support (CORTES et VAPNIK [1995]) cherche à définir la fonction h qui est une combinaison linéaire des variables d'entrée x (équation 1.18), où $h(x) = 0$ est défini comme l'hyperplan séparateur des deux classes considérées ($y = 1$ si $h(x) \geq 0$ et $y = 0$ si $h(x) < 0$).

$$h(x) = w^T x + b \quad (1.18)$$

Les valeurs de w et b caractérisant l'hyperplan dans l'espace des indices sont définies par des vecteurs de support, qui sont les individus les plus proches de l'hyperplan séparateur. L'équation finale est estimée par la maximisation de la marge (la distance) entre les vecteurs de support et l'hyperplan. La figure 1.14 illustre l'hyperplan dans l'espace des indices x_1 et x_2 . Les vecteurs de support sont les deux échantillons bleu et un échantillon vert sur les lignes pointillées (les marges).

Quand les classes ne sont pas séparables par un hyperplan linéaire dans l'espace d'indices, l'équation 1.18 peut être généralisée (équation 1.19) avec l'utilisation d'une fonction $\phi(X)$ qui projette les échantillons dans un espace de dimension supérieure où la séparation linéaire est possible (figure 1.15).

$$h(x) = w^T \phi(x) + b \quad (1.19)$$

Cette astuce est appelé *kernel trick*.

Une nouvelle entrée est ainsi projetée dans l'espace d'indices défini dans l'apprentissage (ré-décrit par $\phi(X)$ ou pas) et sa localisation par rapport à l'hyperplan séparateur définit sa classification.

Les forêts aléatoires (en anglais, *Random Forests (RF)*) Les forêts aléatoires (BREIMAN [2001]) sont des combinaisons d'arbres de décisions qui prennent comme entrée des sous-ensembles aléatoires des données d'entraînement tirées avec remplacement (*bootstrap*) de telle façon que

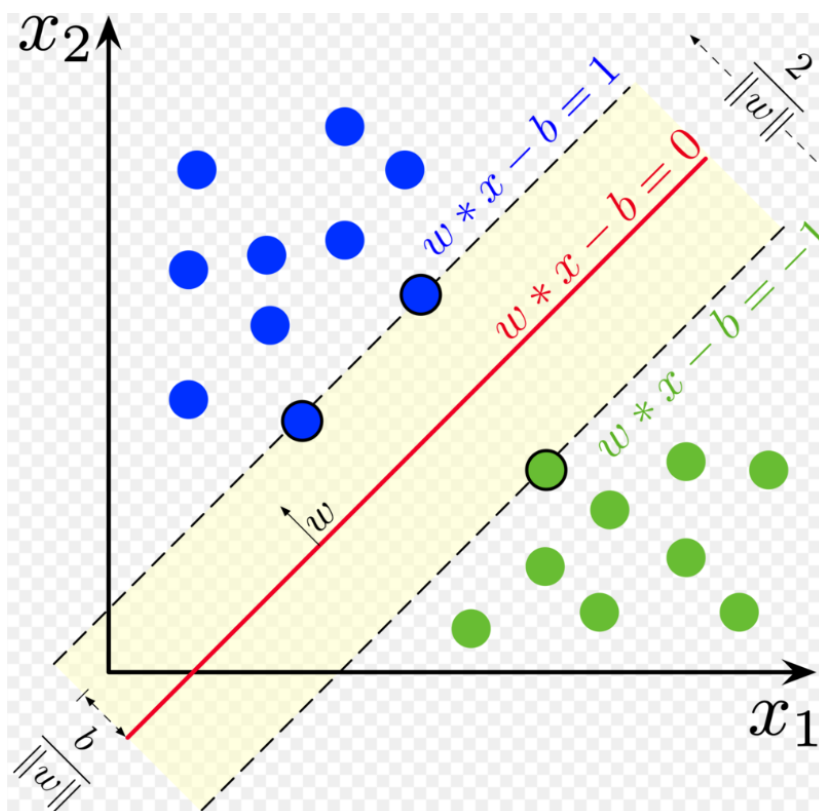


FIGURE 1.14 – Hyperplan et marges pour une SVM formée avec des échantillons de deux classes. Les échantillons en marge sont appelés les vecteurs de support. Source : [Wikipedia](#)

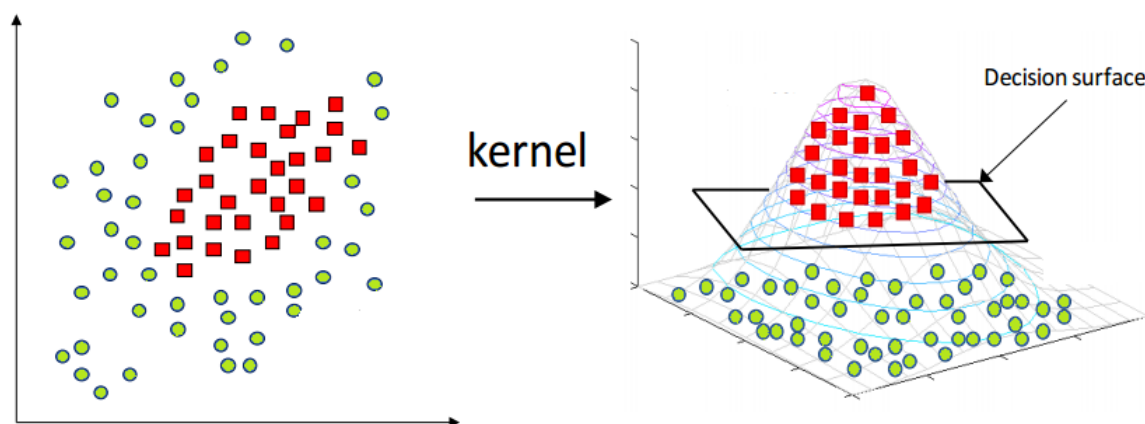


FIGURE 1.15 – Exemple de transformation d'un problème non-linéairement séparable. Source : [Rashmi Jain](#)

les sous-ensembles ont la même distribution et le même nombre de sujets que les données d'entraînement.

Pour chaque arbre, le nœud racine sépare les entrées en deux groupes, en appliquant un seuil d'un des indices et l'opération est répétée pour les nœuds suivants. Le critère le plus fréquemment utilisé pour choisir l'indice et le seuil est l'index de Gini. Chacun des nœuds de l'arbre définit l'indice à prendre en compte et le seuil de répartition de manière à représenter la plus grande réduction d'impureté possible entre le nœud "parent" et les deux nœuds "enfants". Les séparations s'arrêtent quand le critère d'impureté minimale souhaité pour les nœuds finaux (les feuilles des arbres) est atteint, par exemple une impureté de zéro quand toutes les feuilles contiennent seulement des éléments d'une même classe.

La classification finale d'une forêt aléatoire est faite en moyennant les votes de toutes les arbres de la forêt ce qui améliore la précision, augmente la résistance au bruit et contrôle le sur-apprentissage. La figure 1.16 montre un schéma d'une forêt aléatoire de N arbres.

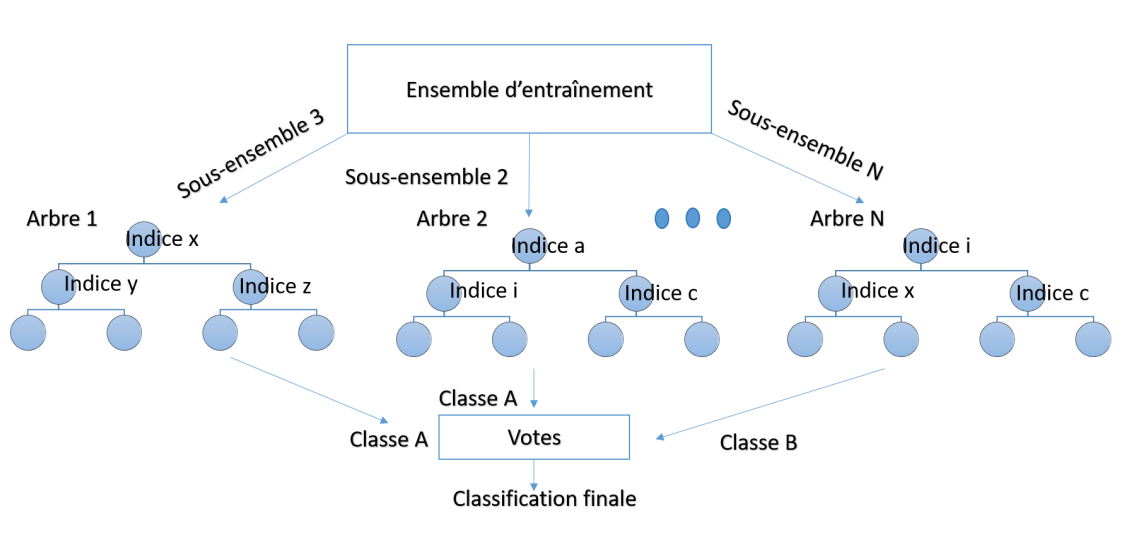


FIGURE 1.16 – Schéma d'une forêt aléatoire.

Méthodes d'apprentissage automatique pour une régression

Les problèmes de régression cherchent à prédire une variable quantitative à partir d'autres (normalement plusieurs autres) variables explicatives. Les méthodes de régression peuvent être paramétriques ou non-paramétriques (MANSKI [1991]).

Dans les régressions paramétriques, le but est d'optimiser les paramètres d'une fonction prédéterminée pour expliquer la variable à prédire par les variables explicatives. La fonction peut être linéaire (régression linéaire), polynomiale (régression polynomiale), entre autres.

Dans la régression non paramétrique la fonction d'estimation, ne prend pas de forme prédéterminée. Elle exige des tailles d'échantillons plus importantes que celles de la régression basée sur des modèles paramétriques car les données doivent fournir la structure du modèle en plus de l'estimation du modèle.

Méthodes d'apprentissage automatique pour les analyses de survie

Un autre type d'étude très fréquent en radiomique est l'analyse de survie. La donnée de survie est le temps écoulé à partir d'un moment précis avant qu'un événement ne survienne. Par exemple, le temps entre la date du diagnostic du patient jusqu'à son décès, ou jusqu'à une récurrence ou une aggravation de son état. Si au moment où le temps d'observation s'achève, le patient est toujours vivant, ou si par n'importe quelle raison, il est perdu de vue, le temps de survie est dit 'censuré'. Dans ce cas, les méthodes statistiques traditionnelles de régression ne peuvent pas

être appliquées. Deux exemples de méthodes d'analyse de survie classiques dans la littérature sont les modèles de Cox à risques proportionnels et plus récemment, les modèles de forêt aléatoire de survie.

Modèle de Cox à risques proportionnels La régression de Cox est utilisée pour la construction d'un modèle de survie qui prédit le temps écoulé avant qu'un événement ne survienne. Le modèle de Cox exprime la fonction de risque instantané de décès γ en fonction du temps t et des covariables X_1, \dots, X_n selon l'équation 1.20.

$$\gamma(t, X_1, \dots, X_n) = \gamma_0(t) \exp\left(\sum_{i=1}^n \beta_i X_i\right) \quad (1.20)$$

Le paramètre γ_0 est appelé le risque de base et correspond au risque instantané de décès lorsque toutes les covariables sont nulles. Le paramètre $\gamma(t, X_1, \dots, X_n)$ correspond au risque instantané de décès étant donné que le patient est vivant au instant t .

Ce modèle de risque proportionnel fait l'hypothèse que le rapport entre le risque instantané de décès d'un patient et le risque instantané de décès d'un autre patient est constant pour tout temps, t . Cette condition peut être testée par le test de Schoenfeld qui vérifie la pente d'une régression linéaire des résidus en fonction du temps. Si elle est différente de zéro, la condition de risque proportionnels n'est pas vérifiée.

Les méthodes de régularisation Ridge et de sélection LASSO sont souvent combinées avec la régression de Cox.

Modèle de forêt aléatoire de survie avec sélection d'indices par profondeur minimale Les forêts aléatoires de survie sont des forêts de régression adaptés aux données censurées de survie (ISHWARAN et collab. [2008]). Le critère de répartition des patients dans chaque nœud est la maximisation du test du logRank. Le test du logRank est le test standard pour vérifier si deux courbes de survie sont différentes. Lorsqu'il est significatif, il permet de rejeter l'hypothèse que les deux courbes sont superposées. Il analyse si, globalement au niveau de chaque décès, la distance entre les deux courbes est plus grande que ce que pourrait expliquer le hasard.

Une sélection d'indices peut être intégrée à la méthode (ISHWARAN et collab. [2010]). La prédictivité de chaque indice est estimée par ce qu'on appelle la profondeur minimale de l'indice. Un sous-arbre maximal correspondant à un indice est le plus grand sous-arbre qui a comme racine un nœud qui divise les sujets par cet indice. La profondeur minimale d'un indice est la distance minimale entre la racine de l'arbre à la racine d'un sous-arbre maximal de l'indice. Une profondeur minimale plus petite implique que l'indice est plus prédictif. Par exemple, les sous-arbres maximaux pour l'indice x_1 sont en bleu sur la figure 1.17. Il y en a exactement deux. La profondeur minimale pour x_1 est égale à un. La moyenne des profondeurs minimales de tous les indices est utilisée comme seuil pour définir si un indice donné est sélectionné ou pas.

1.2.6 Études de comparaison de méthodes de sélection d'indices et de modèles d'apprentissage supervisé

Dans les 89 études relatives à l'utilisation de la radiomique en IRM pour la neuro-oncologie, différentes méthodes ont été utilisées pour la sélection d'indices et l'apprentissage automatique. Il n'y a pas de consensus sur la meilleure méthode; le choix dépend de la base de données et de la question posée.

Avant d'avancer à la prochaine étape de la radiomique (l'évaluation des modèles), deux études assez représentatives méritent d'être analysées.

L'étude de PARMAR et collab. [2015] a comparé 14 méthodes de type Filtre pour la sélection d'indices, classiquement utilisées dans la littérature, combinées avec 12 méthodes d'apprentissage supervisé (figure 1.18) pour prédire la survie de patients avec cancer du poumon, en séparant ceux qui ont vécu moins (0) ou plus (1) de deux ans après le diagnostic. Les modèles ont été définis avec une cohorte de 310 patients, en utilisant 440 indices radiomiques calculées sur des images

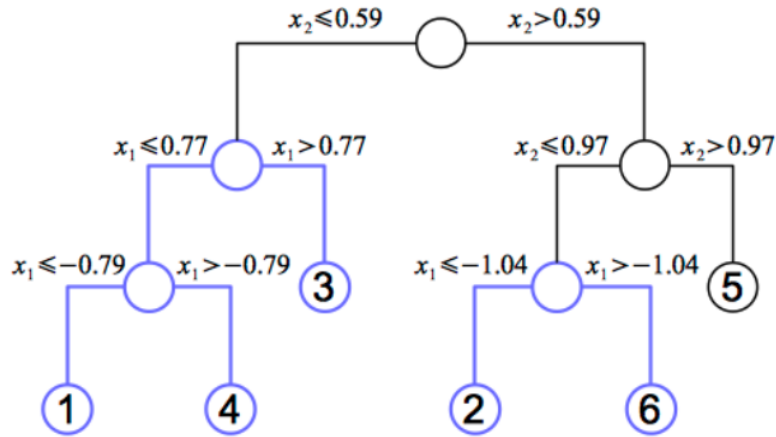


FIGURE 1.17 – Exemple de sous-arbre maximal. Source : *Udaya Kogalur et Hemant Ishwaran*

de TDM. Les résultats de test sur 154 patients d’une cohorte indépendante ont été les meilleurs avec la méthode de sélection d’indices basée sur le test de Wilcoxon et une méthode de classification faisant appel à une forêt aléatoire. Néanmoins, les auteurs ont utilisé des paramètres fixes pour toutes les méthodes d’apprentissage et différents choix de paramètres auraient peut-être pu conduire à des résultats différents.

D’un autre côté, l’étude de (LEGER et collab. [2017]) a comparé 12 méthodes de sélection d’indices combinées avec 11 méthodes d’apprentissage différentes (tableau 1.6) pour les analyses de survie de patients atteints de cancers des voies aérodigestives supérieures. Un total de 213 patients a permis l’entraînement des modèles avec 1610 indices calculés sur des images de TDM et les données de 80 patients indépendants ont été inclus dans la base de test. Ils ont conclu que plusieurs méthodes sont performantes et donnent des résultats équivalents. Néanmoins, l’analyse de différentes méthodes peut réduire les effets aléatoires qui peuvent se produire en sélectionnant une seule méthode de prédiction.

Même si la modalité d’imagerie utilisée n’a pas été l’IRM, ces études illustrent la préoccupation des auteurs vis-à-vis à la grande quantité de choix de méthodes de sélection d’indices et d’apprentissage automatique disponibles pour la radiomique. De plus, ces études suggèrent qu’il n’y a pas des méthodes globalement plus performantes, et que différentes techniques peuvent être testées en parallèle pour résoudre un même problème.

1.2.7 Métriques d’évaluation des modèles

Différentes métriques sont utilisées pour évaluer les résultats de classification et de régression.

Évaluation des modèles de classification

La métrique plus simple est l’exactitude, calculée comme étant le nombre d’éléments bien classifiés sur le total d’éléments. L’exactitude varie entre 0 et 1. Cette métrique n’est pas pertinente dans le cas de données non équilibrées, car sur un jeu de données de 10 patients ayant un cancer et 90 ne l’ayant pas, un modèle trivial répondant que tous les patients sont en bonne santé aura une exactitude de 0,9.

D’autres métriques simples qui caractérisent la performance du modèle sont le taux de vrai positif (TVP), taux de faux positif (TFP), valeur prédictive positive (VPP), valeur prédictive négative (VPN), la sensibilité (se), la spécificité (sp), la précision (pr) et le rappel (rp). Les définitions sont rappelées sur la figure 1.19.

Classification method acronym	Classification method name	Feature Selection method acronym	Feature selection method name
Nnet	Neural network	RELF	Relief
DT	Decision Tree	FSCR	Fisher score
BST	Boosting	GINI	Gini index
BY	Bayesian	CHSQ	Chi-square score
BAG	Bagging	JMI	Joint mutual information
RF	Random Forset	CIFE	Conditional infomax feature extraction
MARS	Multi adaptive regression splines	DISR	Double input symmetric relevance
SVM	Support vector machines	MIM	Mutual information maximization
DA	Discriminant analysis	CMIM	Conditional mutual information maximization
NN	Neirest neighbour	ICAP	Interaction capping
GLM	Generalized linear models	TSCR	T-test score
PLSR	Partial least squares and prinicipal componenet regression	MRMR	Minimum redundancy maximum relevance
—	—	MIFS	Mutual information feature selection
—	—	WLCX	Wilcoxon

FIGURE 1.18 – Les méthodes de sélection d'indices et de classification utilisés par Parmar. *Source : Chintan Parmar (PARMAR et collab. [2015])*

Sélection d'indice	Modèles d'apprentissage automatique
Corrélation de Pearson	Cox
Corrélation de Spearman	NET-Cox
Mutual information maximization (MIM)	Cox - LASSO
Mutual information feature selection (MIFS)	Cox - Elastic Net
Performance uni-variée du modèle Cox	Boosting Trees - Cox (BT-Cox)
Performance multi-variée du modèle Cox	Boosting Trees - C-index (BT-Cindex)
Random forest minimal depth (RF-MD)	Boosting gradient linear models - Cox (BGLM-Cox)
Random forest variable importance (RF-VI)	Boosting gradient linear models - C-index (BGLM-Cindex)
Random forest based on maximally selected rank statistics variable importance (MSR-RFVI)	Random survival Forest (RSF)
Random forest based on permutation variable importance (PVI-RF)	
Random forest using maximally selected rank statistics (MSR-RF)	
Sélection aléatoire	

TABLEAU 1.6 – Les méthodes de sélection d'indices et d'apprentissage automatique utilisées par **LEGER et collab.** [2017]

		Vrai		
		1	0	
Estimé	1	Vrai Positif (VP)	Faux Positif (FP)	$VPP = pr = \frac{VP}{VP+FP}$
	0	Faux Négatif (FN)	Vrai Négatif (VN)	
		$se = \frac{VP}{VP+FN}$	$sp = \frac{VN}{VN+FP}$	
		$TVP = rp = se = \frac{VP}{VP+FN}$	$TFP = 1 - sp = \frac{FP}{VN+FP}$	

FIGURE 1.19 – Définitions de taux de vrai positif (TVP), taux de faux positif (TFP), valeur prédictive positive (VPP), valeur prédictive négative (VPN), sensibilité (se), spécificité (sp), précision (pr) et rappel (rp).

Pour évaluer les performances d'un modèle en faisant varier le seuil de classification, la courbe **fonction d'efficacité du récepteur (courbe ROC)** peut être tracée, cette courbe trace les sensibilités en fonction de $1 - \text{spécificité}$. L'**aire sous la courbe ROC (AUC)** fournit une mesure agrégée des performances pour tous les seuils de classification possibles. Toutefois, l'**AUC** n'évalue pas de façon objective la classification finale choisie par le modèle, elle indique une mesure de la probabilité pour que le modèle classe un exemple positif aléatoire au-dessus d'un exemple négatif aléatoire.

Par exemple, pour un modèle de forêt aléatoire qui calcule la probabilité de chaque élément d'appartenir à la classe 1 par la moyenne des réponses de chaque arbre, la réponse finale sera 1 si plus de la moitié des arbres classifient l'élément comme 1 (probabilité $> 0,5$). Si on définit l'**AUC** associée à la classification en faisant varier le seuil de probabilité on peut obtenir des résultats difficiles à interpréter car l'**AUC** ne prend que la probabilité en compte. Ainsi, avec un jeu de données présentant 50% d'éléments 0 et 50% d'éléments 1, si tous les éléments sont prédits comme 0 mais tous les éléments 1 ont été estimés par la forêt aléatoire comme ayant une probabilité 0,4 d'être 1 alors que les autres ont été estimés par la forêt aléatoire comme ayant une probabilité 0,2 d'être 1, l'**AUC** considère que ce modèle de forêt aléatoire est parfait, avec une valeur de **AUC** = 1,0.

L'**AUC** reste toutefois pertinente pour l'évaluation uni-variée, c'est-à-dire, pour vérifier si un indice prédit la réponse en séparant les éléments par des seuils sur ses valeurs.

Deux autres métriques qui évaluent également la classification finale donnée par un modèle sont l'indice de Youden et le score F1.

L'indice de Youden est défini par l'équation 1.21, prenant en compte la sensibilité et la spécificité du modèle, et varie entre -1 et 1 (la méthode est parfaite).

$$\text{Youden} = sp + se - 1 \quad (1.21)$$

D'un autre côté, le score F1 est la moyenne harmonique (donnée par l'équation 1.22) de la précision et du rappel, il varie entre 0 et 1 (la méthode est parfaite). Le score F1-pondéré est la moyenne pondérée des scores F1 de chaque classe ($y = 0$ et $y = 1$) par le nombre d'échantillons de chaque classe ($N_{y=0} = VN + FP$ et $N_{y=1} = VP + FN$).

$$\text{F1-pondéré} = \frac{N_{y=0} \left(\frac{2prrp}{pr+rp} \right)_{y=0} + N_{y=1} \left(\frac{2prrp}{pr+rp} \right)_{y=1}}{N_{y=0} + N_{y=1}} \quad (1.22)$$

Pour illustrer comment ces deux indices de performance évaluent les modèles différemment, le tableau 1.7 montre des exemples de résultats de classification avec les valeurs de l'indice de Youden et du score F1-pondéré associées sur un exemple déséquilibré.

$\mathbf{y} = [1, 0, 1, 0, 1, 1, 1, 1, 0, 1]$	F1-pondéré	Youden
$y_0 = [1, 0, 1, 0, 1, 1, 1, 1, 0, 1]$	1,00 (1)	1,00 (1)
$y_1 = [1, 1, 1, 0, 1, 1, 1, 1, 0, 1]$	0,89 (3)	0,67 (4)
$y_2 = [0, 0, 1, 0, 1, 1, 1, 1, 0, 1]$	0,90 (2)	0,86 (2)
$y_3 = [1, 1, 1, 1, 1, 1, 1, 1, 0, 1]$	0,76 (6)	0,33 (7)
$y_4 = [0, 0, 0, 0, 1, 1, 1, 1, 0, 1]$	0,81 (4)	0,71 (3)
$y_5 = [0, 1, 1, 0, 1, 1, 1, 1, 0, 1]$	0,80 (5)	0,52 (5)
$y_6 = [0, 1, 0, 0, 1, 1, 1, 1, 0, 1]$	0,71 (7)	0,38 (6)
$y_7 = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0]$	0,14 (9)	0,00 (8)
$y_8 = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1]$	0,58 (8)	0,00 (8)
$y_9 = [0, 1, 0, 1, 0, 0, 0, 0, 1, 0]$	0,00 (10)	-1,00 (9)

TABLEAU 1.7 – Analyse des comportements de l’indice de Youden et du score F1-pondéré pour dix exemples de résultats de prédiction (y_0, \dots, y_9) de la variable y . Les prédictions erronées sont affichées en rouge. Le rang des scores de chaque prédiction est entre parenthèses.

Dans l’exemple du tableau 1.7, les réponses \mathbf{y} , 1 et 0, vraies ont une proportion de 7 : 3. En comparant les scores pour la classification y_1 et y_2 , on peut noter que les deux métriques considèrent une erreur de prédiction de la classe minoritaire (0) comme plus importante. Néanmoins, l’indice de Youden considère que la mauvaise prédiction de deux éléments de la classe majoritaire (y_4) est moins importante qu’une seule mauvaise prédiction de la classe minoritaire (y_1), tandis que le score F1-pondéré considère le contraire. En comparant les scores de y_7 et y_8 , l’indice de Youden retourne la même valeur lorsque toutes les prédictions sont dans une seule classe alors que le score F1-pondéré considère le fait que prédire tous les éléments comme la classe majoritaire est moins pire que les prédire tous comme la classe minoritaire. En conclusion, on ne peut pas établir qu’une mesure d’évaluation est meilleure qu’une autre, il faut juste avoir en tête ce que chacune traduit.

Évaluation des modèles de régression et de survie

L’**indice C** d’Harrel est l’équivalent de l’**AUC** pour les régressions. Une valeur de l’**indice C** égale à 0,5 correspond à une variable non-informative et une valeur de l’**indice C** égale à 1 correspond à une association parfaite de la variable avec les données de survie. L’**indice C** est essentiellement une mesure de corrélation de rangs, motivée par le tau de Kendall pour les données de survie censurées. Pour calculer l’**indice C** entre deux variables X et Y , toutes les paires d’éléments (x_i, y_i) , (x_j, y_j) sont comparées pour $i < j$. Elles sont classifiées comme paires cohérentes si $x_i > x_j$ et $y_i > y_j$ et les paires sont appelées comparables si $y_i > y_j$. L’**indice C** est finalement calculé par l’équation 1.23.

$$Indice\ C = \frac{nb.\ paires\ coherentes}{nb.\ paires\ comparables} \quad (1.23)$$

Le ratio de vraisemblance compare le modèle estimé et le modèle aléatoire, à $\beta = 0$ pour un modèle de Cox par exemple, avec l’hypothèse nulle d’équivalence des deux modèles.

Pour l’évaluation de la survie, les courbes de Kaplan-Meier sont l’estimation du maximum de vraisemblance non-paramétrique de la probabilité de survie à l’instant t , $S(t)$ (équation 1.24), pour n_i décès et d_i patients censurés jusqu’à l’instant t_i .

$$\hat{S}(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i} \quad (1.24)$$

Une mesure de comparaison entre deux groupes des temps de survie est le **rapport de hasards (RH)** (en anglais, *hazard ratio*). Le **RH** entre deux groupes est simplement la division des valeurs γ (risque instantané de décès) de ces deux groupes. Dans le cas de la régression de Cox, il est égal à $exp(\beta)$. L’intervalle de confiance à 95% est exprimé par $[exp(\beta - 1,96et(\beta)); exp(\beta + 1,96et(\beta))]$,

où $et(\beta)$ est l'écart-type des paramètres β . Un RH égal à 1 signifie qu'il n'y a pas de différence de hasards entre les deux groupes, tandis qu'un RH égal à 2 par exemple signifie qu'un groupe a deux fois plus de chances de décès que l'autre. Pour l'évaluation d'un indice continu, la valeur du RH traduit le rapport de hasards entre les éléments présentant une unité de différence de l'indice continu.

1.2.8 L'apprentissage profond

L'apprentissage profond en imagerie est une technique d'apprentissage automatique définissant un réseau de neurones à plusieurs couches pour caractériser l'image à différents niveaux d'abstraction. Ces techniques attirent massivement l'attention des chercheurs depuis 2011, quand plusieurs compétitions en imagerie ont été remportées par des équipes utilisant les réseaux de neurones de convolution profonds (LECUN et collab. [1998]). En 2011, ces techniques ont eu, pour la première fois, une performance supérieure à l'homme dans le concours de reconnaissance de formes visuelles (CIREŞAN et collab. [2011]). En 2012, la compétition ISBI de segmentation de structures neuronales sur des images issues de microscope électronique a été remportée par la même équipe (CIREŞAN et collab. [2012]) et la compétition ImageNet de classification de 1000 classes différentes d'images a été remportée avec une grosse avance par KRIZHEVSKY et collab. [2012] qui ont utilisé un système similaire. L'équipe de Cireşan a remporté les challenges ICPR 2012 et le challenge MICCAI 2013 pour la détection de cancer dans une grande base de données d'imagerie (CIREŞAN et collab. [2013]). Depuis l'obtention de ces performances élevées, nous sommes entrés dans ce qu'on appelle la révolution de l'apprentissage profond.

L'apprentissage profond ne nécessite pas l'extraction d'indices descriptifs car il tire les informations pertinentes directement de l'image. Cela veut dire que les étapes de segmentation et d'extraction d'indices sont globalisées et entièrement automatiques, ce qui le distingue des techniques conventionnelles d'apprentissage automatique. D'un côté, cette approche est avantageuse car moins dépendante vis-à-vis d'une étape intermédiaire et plus simple à appliquer. Cependant, les résultats ne peuvent pas être interprétés simplement, et il est difficile d'expliquer comment le réseau a trouvé la réponse. De plus, l'aspect apprentissage profond et la complexité des modèles nécessitent une base de données de dimension importante, notamment quand les images sont tri-dimensionnelles.

Dans le domaine médical, la quantité de données étiquetées est assez limitée, ce qui complique la mise en place de réseaux de neurones profonds. Une alternative est l'apprentissage par transfert, qui permet d'utiliser un réseau de neurones conçu auparavant pour répondre à un problème spécifique comme base du modèle qui répondra un nouveau problème en ajoutant quelques couches supplémentaires d'apprentissage. L'idée est inspirée de la façon d'apprendre du cerveau humain qui utilise ses connaissances dans un certain domaine pour apprendre plus vite dans un autre. Par exemple une personne qui parle espagnol peut transférer ses connaissances pour apprendre plus vite l'italien.

Dans le cas de l'imagerie 2D, des réseaux construits avec les images de ImageNet (14 millions d'images avec 20.000 catégories) peuvent être potentiellement utilisés pour créer un classifieur en imagerie médicale comme l'étude récente de YUAN et collab. [2019] l'a montré. Les auteurs ont utilisé l'apprentissage par transfert pour différencier des images IRM 2D (T2w et cartes de diffusion ADC) de cancer de la prostate entre bas-grade et haut-grade atteignant une exactitude de 87%. D'un autre côté, les réseaux entraînés avec ces bases de données ne peuvent être utilisés que pour les coupes 2D d'images 3D. Une étude très récente (CHEN et collab. [2019a]) a créé le réseau Med3D en utilisant des images de huit bases de données d'imagerie médicale (IRM et TDM) 3D publiques avec des étiquettes de segmentation. L'utilisation d'apprentissage par transfert à partir de Med3D a abouti à des résultats plus performants que les modèles d'apprentissage profonds créés à partir de zéro pour une étude de segmentation du poumon (80 images IRM et TDM) ainsi que pour une tâche de classification de la malignité de nodules pulmonaires (1.010 images TDM), ce qui montre la flexibilité du réseau Med3D. Néanmoins, le nombre de patients utilisés pour affiner les modèles varie entre une centaine et des milliers, ce qui reste très élevé pour les études de

maladies rares.

Ainsi, les techniques d'apprentissage profond sont très prometteuses mais plus difficiles à interpréter et l'application de ces techniques nécessite encore un nombre important de données en apprentissage.

Conclusion

La radiomique est un sujet de recherche assez récent et très prometteur. Avec les avancées technologiques, une quantité importante de données peut être traitée et des informations quantitatives issues des images médicales qui ne sont pas repérables par les radiologues peuvent être utilisées pour découvrir le lien entre informations génétiques et images ou prédire la réponse à un traitement. L'IRM est la technique la plus référencée dans les publications en radiomique. En neuro-oncologie, les résultats publiés sont encourageants, notamment pour la classification des tumeurs en bas grade et haut grade, pour la prédiction de la survie des patients et pour la prédiction de certaines mutations génétiques importantes pour la prise en charge des patients.

Concernant les aspects méthodologiques de la radiomique, les indices radiomiques traduisent des informations de forme, d'intensité et de texture des tumeurs. Plusieurs mises en œuvre des études radiomiques existent, notamment pour la sélection d'indices et l'apprentissage automatique. Il n'existe pas de consensus sur les méthodes globalement les plus performantes; ainsi différentes techniques peuvent être testées en parallèle pour résoudre un même problème. L'avancement des techniques d'apprentissage profond met au défi les techniques d'apprentissage conventionnelles grâce à des résultats très performants. Cependant, ces techniques plus complexes exigent aussi plus de données d'entraînement et donnent des résultats moins interprétables, ce qui peut limiter leur utilisation dans le domaine de l'imagerie médicale.

1.3 Références

- ABROL, S., A. KOTROTSOU, A. SALEM, P. O. ZINN et R. R. COLLEN. 2017, «Radiomic Phenotyping in Brain Cancer to Unravel Hidden Information in Medical Images», *Top Magn Reson Imaging*, vol. 26, n° 1, p. 43–53. [9](#)
- ALTMAN, N. 1992, «An introduction to kernel and nearest-neighbor nonparametric regression», *The American Statistician*, vol. 46, n° 3, doi :10.1080/00031305.1992.10475879, p. 175–185. URL <https://www.tandfonline.com/doi/abs/10.1080/00031305.1992.10475879>. [23](#)
- ARITA, H., M. KINOSHITA, A. KAWAGUCHI, M. TAKAHASHI, Y. NARITA, Y. TERAKAWA, N. TSUYUGUCHI, Y. OKITA, M. NONAKA, S. MORIUCHI, M. TAKAGAKI, Y. FUJIMOTO, J. FUKAI, S. IZUMOTO, K. ISHIBASHI, Y. NAKAJIMA, T. SHOFUDA, D. KANEMATSU, E. YOSHIOKA, Y. KODAMA, M. MANO, K. MORI, K. ICHIMURA et Y. KANEMURA. 2018, «Lesion location implemented magnetic resonance imaging radiomics for predicting IDH and TERT promoter mutations in grade II/III gliomas», *Sci Rep*, vol. 8, n° 1, p. 11 773. [12](#)
- ARTZI, M., I. BRESSLER et D. BEN BASHAT. 2019, «Differentiation between glioblastoma, brain metastasis and subtypes using radiomics analysis», *J Magn Reson Imaging*. [13](#)
- ARTZI, M., G. LIBERMAN, D. T. BLUMENTHAL, O. AIZENSTEIN, F. BOKSTEIN et D. BEN BASHAT. 2018, «Differentiation between vasogenic edema and infiltrative tumor in patients with high-grade gliomas using texture patch-based analysis», *J Magn Reson Imaging*. [10](#)

- BAE, S., Y. S. CHOI, S. S. AHN, J. H. CHANG, S.-G. KANG, E. H. KIM, S. H. KIM et S.-K. LEE. 2018, «Radiomic mri phenotyping of glioblastoma : Improving survival prediction», *Radiology*, vol. 289, n° 3, doi :10.1148/radiol.2018180200, p. 797–806. URL <https://doi.org/10.1148/radiol.2018180200>, pMID : 30252635. 12
- BELLMAN, R. et K. M. R. COLLECTION. 1961, *Adaptive Control Processes : A Guided Tour*, Princeton Legacy Library, Princeton University Press. URL <https://books.google.fr/books?id=POAmAAAAMAAJ>. 20
- BERKSON, J. 1944, «Application of the logistic function to bio-assay», *Journal of the American Statistical Association*, vol. 39, n° 227, doi :10.1080/01621459.1944.10500699, p. 357–365. URL <https://doi.org/10.1080/01621459.1944.10500699>. 22
- BRANDÃO, L. A. et M. CASTILLO. 2016, «Adult brain tumors : Clinical applications of magnetic resonance spectroscopy», *Magnetic Resonance Imaging Clinics of North America*, vol. 24, n° 4, doi :<https://doi.org/10.1016/j.mric.2016.07.005>, p. 781 – 809, ISSN 1064-9689. URL <http://www.sciencedirect.com/science/article/pii/S106496891630054X>, imaging of Brain Tumors. 6
- BREIMAN, L. 2001, «Random forests», *Machine Learning*, vol. 45, n° 1, doi :10.1023/A:1010933404324, p. 5–32, ISSN 1573-0565. URL <https://doi.org/10.1023/A:1010933404324>. 21, 24
- CALMON, R. 2017, *Serial multimodal MRI of diffuse intrinsic pontine in children*, Theses, Université Pierre et Marie Curie - Paris VI. URL <https://tel.archives-ouvertes.fr/tel-01793106>. 8
- CHADDAD, A., M. J. KUCHARCZYK, P. DANIEL, S. SABRI, B. J. JEAN-CLAUDE, T. NIAZI et B. ABDULKARIM. 2019, «Radiomics in glioblastoma : Current status and challenges facing clinical implementation», *Frontiers in Oncology*, vol. 9, doi :10.3389/fonc.2019.00374, p. 374, ISSN 2234-943X. URL <https://www.frontiersin.org/article/10.3389/fonc.2019.00374>. 13
- CHADDAD, A., S. SABRI, T. NIAZI et B. ABDULKARIM. 2018, «Prediction of survival with multi-scale radiomic analysis in glioblastoma patients», *Med Biol Eng Comput*, vol. 56, n° 12, p. 2287–2300. 11
- CHEN, S., K. MA et Y. ZHENG. 2019a, «Med3d : Transfer learning for 3d medical image analysis», *CoRR*, vol. abs/1904.00625. URL <http://arxiv.org/abs/1904.00625>. 33
- CHEN, W., B. LIU, S. PENG, J. SUN et X. QIAO. 2018a, «Computer-Aided Grading of Gliomas Combining Automatic Segmentation and Radiomics», *Int J Biomed Imaging*, vol. 2018, p. 2512 037. 11
- CHEN, X., M. FANG, D. DONG, L. LIU, X. XU, X. WEI, X. JIANG, L. QIN et Z. LIU. 2019b, «Development and Validation of a MRI-Based Radiomics Prognostic Classifier in Patients with Primary Glioblastoma Multiforme», *Acad Radiol*. 13, 14
- CHEN, Y., Z. LI, G. WU, J. YU, Y. WANG, X. LV, X. JU et Z. CHEN. 2018b, «Primary central nervous system lymphoma and glioblastoma differentiation based on conventional magnetic resonance imaging by high-throughput SIFT features», *Int. J. Neurosci.*, vol. 128, n° 7, p. 608–618. 10, 14
- CHO, H.-H. et H. PARK. 2017, «Classification of low-grade and high-grade glioma using multimodal image radiomics features», p. 3081–3084, doi :10.1109/EMBC.2017.8037508. 10, 14
- CIREŞAN, D., U. MEIER, J. MASCI et J. SCHMIDHUBER. 2011, «A committee of neural networks for traffic sign classification», dans *The 2011 International Joint Conference on Neural Networks*, ISSN 2161-4407, p. 1918–1921, doi :10.1109/IJCNN.2011.6033458. 33

- CIREŞAN, D., A. GIUSTI, L. M. GAMBARDELLA et J. SCHMIDHUBER. 2012, «Deep neural networks segment neuronal membranes in electron microscopy images», dans *Advances in Neural Information Processing Systems 25*, édité par F. Pereira, C. J. C. Burges, L. Bottou et K. Q. Weinberger, Curran Associates, Inc., p. 2843–2851. [33](#)
- CIREŞAN, D. C., A. GIUSTI, L. M. GAMBARDELLA et J. SCHMIDHUBER. 2013, «Mitosis detection in breast cancer histology images with deep neural networks», dans *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, édité par K. Mori, I. Sakuma, Y. Sato, C. Barillot et N. Navab, Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-642-40763-5, p. 411–418. [33](#)
- COLEN, R. R., I. HASSAN, N. ELSHAFAEY et P. O. ZINN. 2016, «Shedding Light on the 2016 World Health Organization Classification of Tumors of the Central Nervous System in the Era of Radiomics and Radiogenomics», *Magn Reson Imaging Clin N Am*, vol. 24, n° 4, p. 741–749. [9](#)
- COQUERY, N., R. SERDUC, C. REMY, E. L. BARBIER et B. LEMASSON. 2018, «Cluster versus ROI analysis to assess combined antiangiogenic therapy and radiotherapy in the F98 rat-glioma model», *NMR Biomed*, vol. 31, n° 8, p. e3933. [11](#)
- CORTES, C. et V. VAPNIK. 1995, «Support-vector networks», *Machine Learning*, vol. 20, n° 3, doi :10.1007/BF00994018, p. 273–297, ISSN 1573-0565. URL <https://doi.org/10.1007/BF00994018>. [24](#)
- COVER, T. M. et J. A. THOMAS. 2006, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*, Wiley-Interscience, New York, NY, USA, ISBN 0471241954. [22](#)
- DELLA SETA, M., F. COLLETTINI, J. CHAPIRO, A. ANGELIDIS, F. ENGELING, B. HAMM et D. KAUL. 2019, «A 3D quantitative imaging biomarker in pre-treatment MRI predicts overall survival after stereotactic radiation therapy of patients with a singular brain metastasis», *Acta Radiol*, p. 284185119831692. [13](#)
- DIAMANDIS, E., C. P. S. GABRIEL, U. WURTEMBERGER, K. GUGGENBERGER, H. URBACH, O. STASZEWSKI, S. LASSMANN, O. SCHNELL, J. GRAUVOGEL, I. MADER et D. H. HEILAND. 2018, «MR-spectroscopic imaging of glial tumors in the spotlight of the 2016 WHO classification», *J. Neurooncol.*, vol. 139, n° 2, p. 431–440. [11](#)
- DIEHN, M., C. NARDINI, D. S WANG, S. MCGOVERN, M. JAYARAMAN, Y. LIANG, K. ALDAPE, S. CHA et M. D KUO. 2008, «Identification of noninvasive imaging surrogates for brain tumor gene-expression modules», *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, doi :10.1073/pnas.0801279105, p. 5213–8. [8](#)
- DITMER, A., B. ZHANG, T. SHUJAAT, A. PAVLINA, N. LUIBRAND, M. GASKILL-SHIPLEY et A. VAGAL. 2018, «Diagnostic accuracy of MRI texture analysis for grading gliomas», *J. Neurooncol.*, vol. 140, n° 3, p. 583–589. [12](#)
- DONG, F., Q. LI, D. XU, W. XIU, Q. ZENG, X. ZHU, F. XU, B. JIANG et M. ZHANG. 2018, «Differentiation between pilocytic astrocytoma and glioblastoma : a decision tree model using contrast-enhanced magnetic resonance imaging-derived quantitative radiomic features», *Eur Radiol*. [12](#)
- FETIT, A. E., J. NOVAK, D. RODRIGUEZ, D. P. AUER, C. A. CLARK, R. G. GRUNDY, A. C. PEET et T. N. ARVANITIS. 2018, «Radiomics in paediatric neuro-oncology : A multicentre study on mri texture analysis», *NMR in Biomedicine*, vol. 31, n° 1, doi :10.1002/nbm.3781, p. e3781. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3781>, e3781 NBM-17-0092.R1. [10](#)
- FOLKMAN, J. 1971, «Tumor angiogenesis : Therapeutic implications», *New England Journal of Medicine*, vol. 285, n° 21, doi :10.1056/NEJM197111182852108, p. 1182–1186. URL <https://doi.org/10.1056/NEJM197111182852108>, pMID : 4938153. [5](#)

- GALLDIKS, N. et K. J. LANGEN. 2017, «Amino acid PET in neuro-oncology : applications in the clinic», *Expert Rev Anticancer Ther*, vol. 17, n° 5, p. 395–397. [9](#)
- GALLOWAY, M. M. 1975, «Texture analysis using gray level run lengths», *Computer Graphics and Image Processing*, vol. 4, n° 2, doi :[https://doi.org/10.1016/S0146-664X\(75\)80008-6](https://doi.org/10.1016/S0146-664X(75)80008-6), p. 172 – 179, ISSN 0146-664X. URL <http://www.sciencedirect.com/science/article/pii/S0146664X75800086>. [17](#)
- GANESHAN, B., S. ABALEKE, R. YOUNG, C. CHATWIN et K. MILES. 2010, «Texture analysis of non-small cell lung cancer on unenhanced computed tomography : Initial evidence for a relationship with tumour glucose metabolism and stage», *Cancer imaging : the official publication of the International Cancer Imaging Society*, vol. 10, doi :10.1102/1470-7330.2010.0021, p. 137–43. [8](#)
- GROSSMANN, P., D. A. GUTMAN, W. D. DUNN, C. A. HOLDER et H. J. AERTS. 2016, «Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in Glioblastoma», *BMC Cancer*, vol. 16, p. 611. [9](#)
- GROSSMANN, P., V. NARAYAN, K. CHANG, R. RAHMAN, L. ABREY, D. A. REARDON, L. H. SCHWARTZ, P. Y. WEN, B. M. ALEXANDER, R. HUANG et H. J. W. L. AERTS. 2017, «Quantitative imaging biomarkers for risk stratification of patients with recurrent glioblastoma treated with bevacizumab», *Neuro-Oncology*, vol. 19, n° 12, doi :10.1093/neuonc/nox092, p. 1688–1697, ISSN 1522-8517. URL <https://doi.org/10.1093/neuonc/nox092>. [9](#)
- GUNTHER, M. 2014, «Perfusion imaging», *Journal of Magnetic Resonance Imaging*, vol. 40, n° 2, doi :10.1002/jmri.24382, p. 269–279. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.24382>. [5](#)
- HAN, Y., Z. XIE, Y. ZANG, S. ZHANG, D. GU, M. ZHOU, O. GEVAERT, J. WEI, C. LI, H. CHEN, J. DU, Z. LIU, D. DONG, J. TIAN et D. ZHOU. 2018, «Non-invasive genotype prediction of chromosome 1p/19q co-deletion by development and validation of an MRI-based radiomics signature in lower-grade gliomas», *J. Neurooncol.*, vol. 140, n° 2, p. 297–306. [12](#)
- HARALICK, R., K. SHANMUGAM et I. DINSTEN. 1973, «Texture features for image classification», *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, n° 6. [17](#)
- HOERL, A. E. et R. W. KENNARD. 1970, «Ridge regression : Biased estimation for nonorthogonal problems», *Technometrics*, vol. 12, n° 1, doi :10.1080/00401706.1970.10488634, p. 55–67. URL <https://www.tandfonline.com/doi/abs/10.1080/00401706.1970.10488634>. [21](#)
- HSU, W., S. EL-SADEN et R. K. TAIRA. 2016, «Medical Imaging Informatics», *Adv. Exp. Med. Biol.*, vol. 939, p. 167–224. [9](#)
- HUTTNER, A. 2017, *Molecular Neuropathology and the Ontogeny of Malignant Gliomas*, p. 15 – 29, doi :10.1007/978-3-319-49864-5_2. [6, 7](#)
- HYARE, H., S. THUST et J. REES. 2017, «Advanced MRI Techniques in the Monitoring of Treatment of Gliomas», *Curr Treat Options Neurol*, vol. 19, n° 3, p. 11. [9](#)
- ISHWARAN, H., U. KOGALUR, E. GORODESKI, A. MINN et M. LAUER. 2010, «High-dimensional variable selection for survival data», *Journal of the American Statistical Association*, vol. 105, n° 489, doi :10.1198/jasa.2009.tm08622, p. 205–217, ISSN 0162-1459. [27](#)
- ISHWARAN, H., U. B. KOGALUR, E. H. BLACKSTONE et M. S. LAUER. 2008, «Random survival forests», *Ann. Appl. Stat.*, vol. 2, n° 3, doi :10.1214/08-AOAS169, p. 841–860. URL <https://doi.org/10.1214/08-AOAS169>. [27](#)

- JOHNSON, W. E., C. LI et A. RABINOVIC. 2006, «Adjusting batch effects in microarray expression data using empirical Bayes methods», *Biostatistics*, vol. 8, n° 1, doi :10.1093/biostatistics/kxj037, p. 118–127, ISSN 1465-4644. URL <https://dx.doi.org/10.1093/biostatistics/kxj037>. 8
- KANG, D., J. E. PARK, Y. H. KIM, J. H. KIM, J. Y. OH, J. KIM, Y. KIM, S. T. KIM et H. S. KIM. 2018, «Diffusion radiomics as a diagnostic model for atypical manifestation of primary central nervous system lymphoma : development and multicenter external validation», *Neuro-oncology*, vol. 20, n° 9, p. 1251–1261. 11
- KICKINGEREDER, P. et O. C. ANDRONESI. 2018, «Radiomics, Metabolic, and Molecular MRI for Brain Tumors», *Semin Neurol*, vol. 38, n° 1, p. 32–40. 11
- KICKINGEREDER, P., M. GOTZ, J. MUSCHELLI, A. WICK, U. NEUBERGER, R. T. SHINOHARA, M. SILL, M. NOWOSIELSKI, H. P. SCHLEMMER, A. RADBRUCH, W. WICK, M. BENDSZUS, K. H. MAIER-HEIN et D. BONEKAMP. 2016, «Large-scale Radiomic Profiling of Recurrent Glioblastoma Identifies an Imaging Predictor for Stratifying Anti-Angiogenic Treatment Response», *Clin. Cancer Res.*, vol. 22, n° 23, p. 5765–5771. 9
- KICKINGEREDER, P., U. NEUBERGER, D. BONEKAMP, P. L. PIECHOTTA, M. GÖTZ, A. WICK, M. SILL, A. KRATZ, R. T. SHINOHARA, D. T. W. JONES, A. RADBRUCH, J. MUSCHELLI, A. UNTERBERG, J. DEBUS, H.-P. SCHLEMMER, C. HEROLD-MENDE, S. PFISTER, A. VON DEIMLING, W. WICK, D. CAPPER, K. H. MAIER-HEIN et M. BENDSZUS. 2017, «Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma», *Neuro-Oncology*, vol. 20, n° 6, doi :10.1093/neuonc/nox188, p. 848–857, ISSN 1522-8517. URL <https://doi.org/10.1093/neuonc/nox188>. 10
- KIM, B. S., S. T. KIM, J. H. KIM, H. J. SEOL, D. H. NAM, H. J. SHIN, J. I. LEE et D. S. KONG. 2019, «Apparent Diffusion Coefficient as a Predictive Biomarker for Survival in Patients with Treatment-Naïve Glioblastoma Using Quantitative Multiparametric Magnetic Resonance Profiling», *World Neurosurg*, vol. 122, p. e812–e820. 12
- KIM, J. Y., J. E. PARK, Y. JO, W. H. SHIM, S. J. NAM, J. H. KIM, R. E. YOO, S. H. CHOI et H. S. KIM. 2018a, «Incorporating diffusion- and perfusion-weighted MRI into a radiomics model improves diagnostic performance for pseudoprogression in glioblastoma patients», *Neuro-oncology*. 12
- KIM, Y., H. H. CHO, S. T. KIM, H. PARK, D. NAM et D. S. KONG. 2018b, «Radiomics features to distinguish glioblastoma from primary central nervous system lymphoma on multi-parametric MRI», *Neuroradiology*, vol. 60, n° 12, p. 1297–1305. 12, 14
- KNIEP, H. C., F. MADESTA, T. SCHNEIDER, U. HANNING, M. H. SCHONFELD, G. SCHON, J. FIEHLER, T. GAUER, R. WERNER et S. GELLISSEN. 2019, «Radiomics of Brain MRI : Utility in Prediction of Metastatic Tumor Type», *Radiology*, vol. 290, n° 2, p. 479–487. 13
- KOHAVI, R. et G. H. JOHN. 1997, «Wrappers for feature subset selection», *Artificial Intelligence*, vol. 97, n° 1, doi :[https://doi.org/10.1016/S0004-3702\(97\)00043-X](https://doi.org/10.1016/S0004-3702(97)00043-X), p. 273 – 324, ISSN 0004-3702. URL <http://www.sciencedirect.com/science/article/pii/S000437029700043X>, relevance. 20
- KOTROTSOU, A., P. O. ZINN et R. R. COLEN. 2016, «Radiomics in Brain Tumors : An Emerging Technique for Characterization of Tumor Environment», *Magn Reson Imaging Clin N Am*, vol. 24, n° 4, p. 719–729. 9
- KRIZHEVSKY, A., I. SUTSKEVER et G. E. HINTON. 2012, «Imagenet classification with deep convolutional neural networks», *Neural Information Processing Systems*, vol. 25, doi :10.1145/3065386. 33

- KUTHURU, S., W. DEADERICK, H. BAI, C. SU, T. VU, V. MONGA et A. RAO. 2018, «A Visually Interpretable, Dictionary-Based Approach to Imaging-Genomic Modeling, With Low-Grade Glioma as a Case Study», *Cancer Inform*, vol. 17, p. 1176935118802 796. [12](#)
- LAMBIN, P., E. RIOS-VELAZQUEZ, R. LEIJENAAR, S. CARVALHO, R. G. VAN STIPHOUT, P. GRANTON, C. M. ZEGERS, R. GILLIES, R. BOELLARD, A. DEKKER et H. J. AERTS. 2012, «Radiomics : Extracting more information from medical images using advanced feature analysis», *European Journal of Cancer*, vol. 48, n° 4, doi :<https://doi.org/10.1016/j.ejca.2011.11.036>, p. 441 – 446, ISSN 0959-8049. URL <http://www.sciencedirect.com/science/article/pii/S0959804911009993>. [2, 3](#)
- LAO, J., Y. CHEN, Z. C. LI, Q. LI, J. ZHANG, J. LIU et G. ZHAI. 2017, «A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme», *Sci Rep*, vol. 7, n° 1, p. 10353. [10](#)
- LECUN, Y., L. BOTTOU, Y. BENGIO et P. HAFFNER. 1998, «Gradient-based learning applied to document recognition», dans *Proceedings of the IEEE*, p. 2278–2324. [33](#)
- LEE, J., S. NARANG, J. J. MARTINEZ, G. RAO et A. RAO. 2015, «Associating spatial diversity features of radiologically defined tumor habitats with epidermal growth factor receptor driver status and 12-month survival in glioblastoma : methods and preliminary investigation», *J Med Imaging (Bellingham)*, vol. 2, n° 4, p. 041 006. [9](#)
- LEE, M., B. WOO, M. D. KUO, N. JAMSHIDI et J. H. KIM. 2017, «Quality of Radiomic Features in Glioblastoma Multiforme : Impact of Semi-Automated Tumor Segmentation Software», *Korean J Radiol*, vol. 18, n° 3, p. 498–509. [9](#)
- LEGER, S., A. ZWANENBURG, K. PILZ, F. LOHAUS, A. LINGE, K. ZÖPHEL, J. KOTZERKE, A. SCHREIBER, I. TINHOFFER, V. BUDACH, A. SAK, M. STUSCHKE, P. BALERMPAS, C. RÖDEL, U. GANSWINDT, C. BELKA, S. PIGORSCH, U.-P. D. S. COMBS, D. MÖNNICH et C. RICHTER. 2017, «A comparative study of machine learning methods for time-to-event survival data for radiomics risk modeling», *Scientific Reports*, vol. 7, doi :[10.1038/s41598-017-13448-3](https://doi.org/10.1038/s41598-017-13448-3). [28, 30](#)
- LEHRER, M., A. BHADRA, V. RAVIKUMAR, J. Y. CHEN, M. WINTERMARK, S. N. HWANG, C. A. HOLDER, E. P. HUANG, B. FEVRIER-SULLIVAN, J. B. FREYMAN et A. RAO. 2017, «Multiple-response regression analysis links magnetic resonance imaging features to de-regulated protein expression and pathway activity in lower grade glioma», *Oncoscience*, vol. 4, n° 5-6, p. 57–66. [10](#)
- LI, Q., H. BAI, Y. CHEN, Q. SUN, L. LIU, S. ZHOU, G. WANG, C. LIANG et Z. C. LI. 2017a, «A Fully-Automatic Multiparametric Radiomics Model : Towards Reproducible and Prognostic Imaging Signature for Prediction of Overall Survival in Glioblastoma Multiforme», *Sci Rep*, vol. 7, n° 1, p. 14331. [10](#)
- LI, Y., X. LIU, Z. QIAN, Z. SUN, K. XU, K. WANG, X. FAN, Z. ZHANG, S. LI, Y. WANG et T. JIANG. 2018a, «Genotype prediction of ATRX mutation in lower-grade gliomas using an MRI radiomics signature», *Eur Radiol*, vol. 28, n° 7, p. 2960–2968. [10](#)
- LI, Y., X. LIU, K. XU, Z. QIAN, K. WANG, X. FAN, S. LI, Y. WANG et T. JIANG. 2018b, «MRI features can predict EGFR expression in lower grade gliomas : A voxel-based radiomic analysis», *Eur Radiol*, vol. 28, n° 1, p. 356–362. [10](#)
- LI, Z., H. BAI, Q. SUN, Y. ZHAO, Y.-C. LV, J. ZHOU, C. LIANG, Y.-S. CHEN, D. LIANG et H. ZHENG. 2018c, «Multiregional radiomics profiling from multiparametric mri : Identifying an imaging predictor of idh1 mutation status in glioblastoma», *Cancer Medicine*, vol. 7, doi :[10.1002/cam4.1863](https://doi.org/10.1002/cam4.1863), p. 1–11. [12, 14](#)

- LI, Z., Y. WANG, J. YU, Y. GUO et W. CAO. 2017b, «Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma», *Sci Rep*, vol. 7, n° 1, p. 5467. [9](#)
- LI, Z., Y. WANG, J. YU, Y. GUO et Q. ZHANG. 2017c, «Age groups related glioblastoma study based on radiomics approach», *Comput Assist Surg (Abingdon)*, vol. 22, n° sup1, p. 18–25. [10](#)
- LI, Z. C., H. BAI, Q. SUN, Q. LI, L. LIU, Y. ZOU, Y. CHEN, C. LIANG et H. ZHENG. 2018d, «Multiregional radiomics features from multiparametric MRI for prediction of MGMT methylation status in glioblastoma multiforme : A multicentre study», *Eur Radiol*, vol. 28, n° 9, p. 3640–3650. [11](#)
- LIAO, X., B. CAI, B. TIAN, Y. LUO, W. SONG et Y. LI. «Machine-learning based radiogenomics analysis of mri features and metagenes in glioblastoma multiforme patients with different survival time», *Journal of Cellular and Molecular Medicine*, vol. 23, n° 6, doi :10.1111/jcmm.14328, p. 4375–4385. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/jcmm.14328>. [13](#)
- LIU, L., H. ZHANG, J. WU, Z. YU, X. CHEN, I. REKIK, Q. WANG, J. LU et D. SHEN. 2018a, «Overall survival time prediction for high-grade glioma patients based on large-scale brain functional networks», *Brain Imaging Behav*. [12](#)
- LIU, T., G. WU, J. YU, Y. GUO, Y. WANG, Z. SHI et L. CHEN. 2017, «A mrmrsrc feature selection method for radiomics approach», p. 616–619, doi :10.1109/EMBC.2017.8036900. [10](#)
- LIU, X., Y. LI, Z. QIAN, Z. SUN, K. XU, K. WANG, S. LIU, X. FAN, S. LI, Z. ZHANG, T. JIANG et Y. WANG. 2018b, «A radiomic signature as a non-invasive predictor of progression-free survival in patients with lower-grade gliomas», *NeuroImage : Clinical*, vol. 20, doi :<https://doi.org/10.1016/j.nicl.2018.10.014>, p. 1070 – 1077, ISSN 2213-1582. URL <http://www.sciencedirect.com/science/article/pii/S2213158218303243>. [12](#), [14](#)
- LIU, Z., Y. WANG, X. LIU, Y. DU, Z. TANG, K. WANG, J. WEI, D. DONG, Y. ZANG, J. DAI, T. JIANG et J. TIAN. 2018c, «Radiomics analysis allows for precise prediction of epilepsy in patients with low-grade gliomas», *Neuroimage Clin*, vol. 19, p. 271–278. [11](#)
- LOHMANN, P., M. KOCHER, J. STEGER et N. GALLDIKS. 2018, «Radiomics derived from amino-acid PET and conventional MRI in patients with high-grade gliomas», *Q J Nucl Med Mol Imaging*, vol. 62, n° 3, p. 272–280. [11](#)
- LOPEZ, C. J., N. NAGORNAYA, N. A. PARRA, D. KWON, F. ISHKANIAN, A. M. MARKOE, A. MAUDSLEY et R. STOYANOVA. 2017, «Association of Radiomics and Metabolic Tumor Volumes in Radiation Treatment of Glioblastoma Multiforme», *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 97, n° 3, p. 586–595. [9](#)
- LORENSEN, W. E. et H. E. CLINE. 1987, «Marching cubes : A high resolution 3d surface construction algorithm», *SIGGRAPH Comput. Graph.*, vol. 21, n° 4, doi :10.1145/37402.37422, p. 163–169, ISSN 0097-8930. URL <http://doi.acm.org/10.1145/37402.37422>. [16](#)
- LOTAN, E., R. JAIN, N. RAZAVIAN, G. M. FATTERPEKAR et Y. W. LUI. 2019, «State of the Art : Machine Learning Applications in Glioma Imaging», *AJR Am J Roentgenol*, vol. 212, n° 1, p. 26–37. [12](#)
- LOUIS, D. N., A. PERRY, G. REIFENBERGER, A. VON DEIMLING, D. FIGARELLA-BRANGER, W. K. CAVENEE, H. OHGAKI, O. D. WIESTLER, P. KLEIHUES et D. W. ELLISON. 2016, «The 2016 world health organization classification of tumors of the central nervous system : a summary», *Acta Neuropathologica*, vol. 131, n° 6, doi :10.1007/s00401-016-1545-1, p. 803–820, ISSN 1432-0533. URL <https://doi.org/10.1007/s00401-016-1545-1>. [6](#), [7](#)
- LU, C. F., F. T. HSU, K. L. HSIEH, Y. J. KAO, S. J. CHENG, J. B. HSU, P. H. TSAI, R. J. CHEN, C. C. HUANG, Y. YEN et C. Y. CHEN. 2018, «Machine Learning-Based Radiomics for Molecular Subtyping of Gliomas», *Clin. Cancer Res.*, vol. 24, n° 18, p. 4429–4436. [11](#)

- LUYPAERT, R., S. BOUJRAF, S. SOURBRON et M. OSTEAX. 2001, «Diffusion and perfusion mri : basic physics», *European Journal of Radiology*, vol. 38, n° 1, doi :[https://doi.org/10.1016/S0720-048X\(01\)00286-8](https://doi.org/10.1016/S0720-048X(01)00286-8), p. 19 – 27, ISSN 0720-048X. URL <http://www.sciencedirect.com/science/article/pii/S0720048X01002868>. 6
- LY, K. I. et E. R. GERSTNER. 2018, «The Role of Advanced Brain Tumor Imaging in the Care of Patients with Central Nervous System Malignancies», *Curr Treat Options Oncol*, vol. 19, n° 8, p. 40. 11
- MANSKI, C. F. 1991, «Regression», *Journal of Economic Literature*, vol. 29, n° 1, p. 34–50, ISSN 00220515. URL <http://www.jstor.org/stable/2727353>. 26
- MCGARRY, S. D., S. L. HURRELL, A. L. KACZMAROWSKI, E. J. COCHRAN, J. CONNELLY, S. D. RAND, K. M. SCHMAINDA et P. S. LAVIOLETTE. 2016, «Magnetic Resonance Imaging-Based Radiomic Profiles Predict Patient Prognosis in Newly Diagnosed Glioblastoma Before Therapy», *Tomography*, vol. 2, n° 3, p. 223–228. 9
- MUKHERJEE, P., J. BERMAN, S. CHUNG, C. HESS et R. HENRY. 2008, «Diffusion tensor mr imaging and fiber tractography : Theoretic underpinnings», *American Journal of Neuroradiology*, vol. 29, n° 4, doi :10.3174/ajnr.A1051, p. 632–641, ISSN 0195-6108. URL <http://www.ajnr.org/content/29/4/632>. 6
- NIE, D., J. LU, H. ZHANG, E. ADELI, J. WANG, Z. YU, L. LIU, Q. WANG, J. WU et D. SHEN. 2019, «Multi-Channel 3D Deep Feature Learning for Survival Time Prediction of Brain Tumor Patients Using Multi-Modal Neuroimages», *Sci Rep*, vol. 9, n° 1, p. 1103. 13
- NIE, K., J.-H. CHEN, H. J. YU, Y. CHU, O. NALCIOGLU et M.-Y. SU. 2008, «Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast mri», *Academic Radiology*, vol. 15, n° 12, doi :<https://doi.org/10.1016/j.acra.2008.06.005>, p. 1513 – 1525, ISSN 1076-6332. URL <http://www.sciencedirect.com/science/article/pii/S1076633208003991>. 8
- NIOCHE, C., F. ORLHAC, S. BOUGHDAD, S. REUZE, J. GOYA-OUTI, C. ROBERT, C. PELLOT-BARAKAT, M. SOUSSAN, F. ERIQUE FROUIN et I. BUVAT. 2018, «Lifex : A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity», *Cancer Research*, vol. 78, n° 16, doi :10.1158/0008-5472.CAN-18-0125, p. 4786–4789, ISSN 15387445. URL <http://cancerres.aacrjournals.org/>. 19, 20
- PARK, J. E. et H. S. KIM. 2018, «Radiomics as a Quantitative Imaging Biomarker : Practical Considerations and the Current Standpoint in Neuro-oncologic Studies», *Nucl Med Mol Imaging*, vol. 52, n° 2, p. 99–108. 11
- PARMAR, C., P. GROSSMANN, J. BUSSINK, P. LAMBIN et H. AERTS. 2015, «Machine learning methods for quantitative radiomic biomarkers», *Scientific reports*, vol. 5, doi :10.1038/srep13087, p. 13 087. 20, 27, 29
- PEEKEN, J. C., J. HESSE, B. HALLER, K. A. KESSEL, F. NUSSLIN et S. E. COMBS. 2018, «Semantic imaging features predict disease progression and survival in glioblastoma multiforme patients», *Strahlenther Onkol*, vol. 194, n° 6, p. 580–590. 11
- PENG, H., F. LONG et C. DING. 2005, «Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, n° 8, doi :10.1109/TPAMI.2005.159, p. 1226–1238, ISSN 0162-8828. 22
- PRASANNA, P., A. KARNAWAT, M. ISMAIL, A. MADABHUSHI et P. TIWARI. 2019, «Radiomics-based convolutional neural network for brain tumor segmentation on multiparametric magnetic resonance imaging», *J Med Imaging (Bellingham)*, vol. 6, n° 2, p. 024 005. 13

- PRASANNA, P., J. PATEL, S. PARTOVI, A. MADABHUSHI et P. TIWARI. 2017, «Radiomic features from the peritumoral brain parenchyma on treatment-naïve multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme : Preliminary findings», *Eur Radiol*, vol. 27, n° 10, p. 4188–4197. [9](#)
- PRASANNA, P., P. TIWARI et A. MADABHUSHI. 2016, «Co-occurrence of Local Anisotropic Gradient Orientations (CoLLAGe) : A new radiomics descriptor», *Sci Rep*, vol. 6, p. 37 241. [9](#)
- RATHORE, S., H. AKBARI, J. DOSHI, G. SHUKLA, M. ROZYCKI, M. BILELLO, R. LUSTIG et C. DAVATZIKOS. 2018, «Radiomic signature of infiltration in peritumoral edema predicts subsequent recurrence in glioblastoma : implications for personalized radiotherapy planning», *J Med Imaging (Bellingham)*, vol. 5, n° 2, p. 021 219. [11](#)
- REIFENBERGER, G., H.-G. WIRSCHING, C. B. KNOBBE-THOMSEN et M. WELLER. 2016, «Advances in the molecular genetics of gliomas — implications for classification and therapy», *Nature Reviews Clinical Oncology*, vol. 14, p. 434–452. [6](#)
- REN, Y., X. ZHANG, W. RUI, H. PANG, T. QIU, J. WANG, Q. XIE, T. JIN, H. ZHANG, H. CHEN, Y. ZHANG, H. LU, Z. YAO, J. ZHANG et X. FENG. 2019, «Noninvasive Prediction of IDH1 Mutation and ATRX Expression Loss in Low-Grade Gliomas Using Multiparametric MR Radiomic Features», *J Magn Reson Imaging*, vol. 49, n° 3, p. 808–817. [12](#), [14](#)
- RUDIE, J. D., A. M. RAUSCHECKER, R. N. BRYAN, C. DAVATZIKOS et S. MOHAN. 2019, «Emerging Applications of Artificial Intelligence in Neuro-Oncology», *Radiology*, vol. 290, n° 3, p. 607–618. [13](#)
- SANGHANI, P., B. T. ANG, N. K. K. KING et H. REN. 2018, «Overall survival prediction in glioblastoma multiforme patients from volumetric, shape and texture features using machine learning», *Surg Oncol*, vol. 27, n° 4, p. 709–714. [13](#)
- SELVARAJ, A. et L. GANESAN. 2003, «Texture classification using wavelet transform», *Pattern Recognition Letters*, vol. 24, doi :10.1016/S0167-8655(02)00390-2, p. 1513–1521. [16](#)
- SHOFTY, B., M. ARTZI, D. BEN BASHAT, G. LIBERMAN, O. HAIM, A. KASHANIAN, F. BOKSTEIN, D. T. BLUMENTHAL, Z. RAM et T. SHAHAR. 2018, «Mri radiomics analysis of molecular alterations in low-grade gliomas», *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, n° 4, doi :10.1007/s11548-017-1691-5, p. 563–571, ISSN 1861-6429. URL <https://doi.org/10.1007/s11548-017-1691-5>. [10](#)
- SHUKLA, G., G. S. ALEXANDER, S. BAKAS, R. NIKAM, K. TALEKAR, J. D. PALMER et W. SHI. 2017, «Advanced magnetic resonance imaging in glioblastoma : a review», *Chin Clin Oncol*, vol. 6, n° 4, p. 40. [10](#)
- SOTAK, G. et K. BOYER. 1989, «The laplacian-of-gaussian kernel : A formal analysis and design procedure for fast, accurate convolution and full-frame output», *Computer Vision, Graphics, and Image Processing*, vol. 48, n° 2, doi :https://doi.org/10.1016/S0734-189X(89)80036-2, p. 147 – 189, ISSN 0734-189X. URL <http://www.sciencedirect.com/science/article/pii/S0734189X89800362>. [16](#)
- SU, C., J. JIANG, S. ZHANG, J. SHI, K. XU, N. SHEN, J. ZHANG, L. LI, L. ZHAO, J. ZHANG, Y. QIN, Y. LIU et W. ZHU. 2019, «Radiomics based on multicontrast MRI can precisely differentiate among glioma subtypes and predict tumour-proliferative behaviour», *Eur Radiol*, vol. 29, n° 4, p. 1986–1996. [12](#)
- SUH, H. B., Y. S. CHOI, S. BAE, S. S. AHN, J. H. CHANG, S. G. KANG, E. H. KIM, S. H. KIM et S. K. LEE. 2018, «Primary central nervous system lymphoma and atypical glioblastoma : Differentiation using radiomics approach», *Eur Radiol*, vol. 28, n° 9, p. 3832–3839. [11](#)

- TABACHNICK, B. G. et L. S. FIDELL. 2013, 6^e éd., Pearson, ISBN 978-0205849574. [23](#)
- THIBAUT, G., B. FERTIL, C. L. NAVARRO, S. PEREIRA, P. CAU, N. LÉVY, J. SEQUEIRA et J.-L. MARI. 2009, «Texture indexes and gray level size zone matrix. Application to cell nuclei classification», dans *10th International Conference on Pattern Recognition and Information Processing, PRIP 2009*, Minsk, Belarus, p. 140–145. URL <https://hal.archives-ouvertes.fr/hal-01499715>. [17](#)
- THON, N., F. KRETH et S. KRETH. 2013, «Personalized treatment strategies in glioblastoma : Mgmt promoter methylation status», *OncoTargets and Therapy*, doi :10.2147/OTT.S50208. [8](#)
- TIAN, Q., L. F. YAN, X. ZHANG, X. ZHANG, Y. C. HU, Y. HAN, Z. C. LIU, H. Y. NAN, Q. SUN, Y. Z. SUN, Y. YANG, Y. YU, J. ZHANG, B. HU, G. XIAO, P. CHEN, S. TIAN, J. XU, W. WANG et G. B. CUI. 2018, «Radiomics strategy for glioma grading using texture features from multiparametric MRI», *J Magn Reson Imaging*, vol. 48, n^o 6, p. 1518–1528. [11](#), [14](#)
- TIBSHIRANI, R. 1996, «Regression shrinkage and selection via the lasso», *Journal of the Royal Statistical Society : Series B (Methodological)*, vol. 58, n^o 1, doi :10.1111/j.2517-6161.1996.tb02080.x, p. 267–288. URL <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1996.tb02080.x>. [21](#)
- TIXIER, F., H. UM, D. BERMUDEZ, A. IYER, A. APTE, M. S. GRAHAM, K. S. NEVEL, J. DEASY, R. J. YOUNG et H. VEERARAGHAVAN. 2019, «Preoperative mri-radiomics features improve prediction of survival in glioblastoma patients over mgmt methylation status alone», *Oncotarget*, vol. 10, doi :10.18632/oncotarget.26578. [13](#)
- VAMVAKAS, A., S. C. WILLIAMS, K. THEODOROU, E. KAPSALAKI, K. FOUNTAS, C. KAPPAS, K. VASSIOU et I. TSOUGOS. 2019, «Imaging biomarker analysis of advanced multiparametric MRI for glioma grading», *Phys Med*, vol. 60, p. 188–198. [13](#)
- VAN GRIETHUYSEN, J. J., A. FEDOROV, C. PARMAR, A. HOSNY, N. AUCOIN, V. NARAYAN, R. G. BEETS-TAN, J. C. FILLION-ROBIN, S. PIEPER et H. J. AERTS. 2017, «Computational radiomics system to decode the radiographic phenotype», *Cancer Research*, vol. 77, n^o 21, doi :10.1158/0008-5472.CAN-17-0339, p. e104–e107, ISSN 15387445. [18](#)
- VERDUIN, M., I. COMPTER, D. STEIJVERS, A. A. POSTMA, D. B. P. EEKERS, M. M. ANTEN, L. ACKERMANS, M. TER LAAN, R. T. H. LEIJENAAR, T. VAN DE WEIJER, V. C. G. TJAN-HEIJNEN, A. HOEBEN et M. VOOIJS. 2018, «Noninvasive Glioblastoma Testing : Multimodal Approach to Monitoring and Predicting Treatment Response», *Dis. Markers*, vol. 2018, p. 2908609. [11](#)
- WANG, C., W. SUN, J. KIRKPATRICK, Z. CHANG et F. F. YIN. 2018, «Assessment of concurrent stereotactic radiosurgery and bevacizumab treatment of recurrent malignant gliomas using multimodality MRI imaging and radiomics analysis», *J Radiosurg SBRT*, vol. 5, n^o 3, p. 171–181. [11](#)
- WANG, Q., Q. LI, R. MI, H. YE, H. ZHANG, B. CHEN, Y. LI, G. HUANG et J. XIA. 2019, «Radiomics Nomogram Building From Multiparametric MRI to Predict Grade in Patients With Glioma : A Cohort Study», *J Magn Reson Imaging*, vol. 49, n^o 3, p. 825–833. [12](#)
- WEI, J., G. YANG, X. HAO, D. GU, Y. TAN, X. WANG, D. DONG, S. ZHANG, L. WANG, H. ZHANG et J. TIAN. 2019, «A multi-sequence and habitat-based MRI radiomics signature for preoperative prediction of MGMT promoter methylation in astrocytomas with prognostic implication», *Eur Radiol*, vol. 29, n^o 2, p. 877–888. [12](#)
- WU, G., Z. LI, Y. WANG, J. YU, Y. CHEN et Z. CHEN. 2018, «[Primary central nervous system lymphoma and glioblastoma image differentiation based on sparse representation system]», *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*, vol. 35, n^o 5, p. 754–760. [12](#)

- WU, S., J. MENG, Q. YU, P. LI et S. FU. 2019, «Radiomics-based machine learning methods for isocitrate dehydrogenase genotype prediction of diffuse gliomas», *J. Cancer Res. Clin. Oncol.*, vol. 145, n° 3, p. 543–550. [13](#), [14](#)
- WU, Y. P., Y. S. LIN, W. G. WU, C. YANG, J. Q. GU, Y. BAI et M. Y. WANG. 2017, «Semiautomatic Segmentation of Glioma on Mobile Devices», *J Healthc Eng*, vol. 2017, p. 8054939. [10](#)
- XI, Y. B., F. GUO, Z. L. XU, C. LI, W. WEI, P. TIAN, T. T. LIU, L. LIU, G. CHEN, J. YE, G. CHENG, L. B. CUI, H. J. ZHANG, W. QIN et H. YIN. 2018, «Radiomics signature : A potential biomarker for the prediction of MGMT promoter methylation in glioblastoma», *J Magn Reson Imaging*, vol. 47, n° 5, p. 1380–1387. [10](#)
- YANG, F., N. DOGAN, R. STOYANOVA et J. C. FORD. 2018, «Evaluation of radiomic texture feature error due to MRI acquisition and reconstruction : A simulation study utilizing ground truth», *Phys Med*, vol. 50, p. 26–36. [11](#)
- YIN, H. L., D. B. LI, Y. JIANG, S. H. LI, Y. CHEN et G. W. LIN. 2018, «[High-throughput texture analysis in the distinction of single metastatic brain tumors from high-grade gliomas]», *Zhonghua Zhong Liu Za Zhi*, vol. 40, n° 11, p. 841–846. [13](#)
- YU, J., Z. SHI, C. JI, Y. LIAN, Y. WANG, L. CHEN et Y. MAO. 2017a, «Anatomical location differences between mutated and wild-type isocitrate dehydrogenase 1 in low-grade gliomas», *Int. J. Neurosci.*, vol. 127, n° 10, p. 873–880. [9](#)
- YU, J., Z. SHI, Y. LIAN, Z. LI, T. LIU, Y. GAO, Y. WANG, L. CHEN et Y. MAO. 2017b, «Noninvasive IDH1 mutation estimation based on a quantitative radiomics approach for grade II glioma», *Eur Radiol*, vol. 27, n° 8, p. 3509–3522. [9](#)
- YUAN, Y., W. QIN, M. BUYOUNOUSKI, B. IBRAGIMOV, S. HANCOCK, B. HAN et L. XING. 2019, «Prostate cancer classification with multiparametric mri transfer learning model», *Medical Physics*, vol. 46, n° 2, doi :10.1002/mp.13367, p. 756–765. URL <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.13367>. [33](#)
- YUN, J., J. E. PARK, H. LEE, S. HAM, N. KIM et H. S. KIM. 2019, «Radiomic features and multilayer perceptron network classifier : a robust MRI classification strategy for distinguishing glioblastoma from primary central nervous system lymphoma», *Sci Rep*, vol. 9, n° 1, p. 5746. [13](#)
- ZHANG, X., H. LU, Q. TIAN, N. FENG, L. YIN, X. XU, P. DU et Y. LIU. 2019, «A radiomics nomogram based on multiparametric MRI might stratify glioblastoma patients according to survival», *Eur Radiol*. [13](#)
- ZHANG, X., Q. TIAN, L. WANG, Y. LIU, B. LI, Z. LIANG, P. GAO, K. ZHENG, B. ZHAO et H. LU. 2018, «Radiomics strategy for molecular subtype stratification of lower-grade glioma : Detecting idh and tp53 mutations based on multimodal mri : Radiomics : Molecular stratification of lrgg», *Journal of Magnetic Resonance Imaging*, vol. 48, doi :10.1002/jmri.25960. [10](#), [14](#)
- ZINN, P. O., S. K. SINGH, A. KOTROTSOU, I. HASSAN, G. THOMAS, M. M. LUEDI, A. ELAKKAD, N. EL-SHAFFEY, T. IDRIS, J. MOSLEY, J. GUMIN, G. N. FULLER, J. F. DE GROOT, V. BALADANDAYUTHAPANI, E. P. SULMAN, A. J. KUMAR, R. SAWAYA, F. F. LANG, D. PIWNICA-WORMS et R. R. COLEN. 2018, «A Coclinical Radiogenomic Validation Study : Conserved Magnetic Resonance Radiomic Appearance of Periostin-Expressing Glioblastoma in Patients and Xenograft Models», *Clin. Cancer Res.*, vol. 24, n° 24, p. 6288–6299. [12](#)
- ZWANENBURG, A., S. LEGER, M. VALLIÈRES et S. LÖCK. 2016, «Image biomarker standardisation initiative - feature definitions», *CoRR*, vol. abs/1612.07003. URL <http://arxiv.org/abs/1612.07003>. [16](#)

Chapitre 2

Le gliome infiltrant du tronc cérébral

*« Il y a en chacun de nous des
calculs que nous nommons
espérance »*

Platon



Ce chapitre est une introduction au gliome infiltrant du tronc cérébral (GITC). Il présente les caractéristiques de cette tumeur pédiatrique en clinique, en radiologie et en génétique. Il les situe dans les classifications proposées par l'organisation mondiale de la santé. Le chapitre donne un bref état de l'art des traitements proposés à présent, ainsi que l'état de l'art des découvertes permettant d'associer des informations radiologiques avec le pronostic des patients. Ce chapitre fait aussi un bilan des études qui se rapprochent de la radiomique dans le cadre du GITC. Enfin, il finit par la présentation des deux bases de données de patients GITC qui ont été utilisées pour ce travail de thèse.

Sommaire

2.1	Caractéristiques cliniques	46
2.2	Biopsie et sous-types moléculaires	46
2.3	GITC selon la classification de l'OMS	48
2.4	Approches thérapeutiques	48
2.5	Rôle de l'IRM	49
2.5.1	Images anatomiques	49
2.5.2	Images de perfusion	49
2.5.3	Images de diffusion	50
2.5.4	Spectroscopie	50
2.6	Radiomique pour la caractérisation du GITC	51
2.7	Bases de données de patients atteint de GITC	52
2.7.1	PREBIOMEDE	52
2.7.2	BIOMEDE	54
2.8	Références	55

2.1 Caractéristiques cliniques

Le **GITC** est un gliome de haut-grade affectant surtout les enfants. L'âge moyen de diagnostic se situe entre 5 et 8 ans. Environ 90% des patients décèdent moins de 2 ans après le diagnostic. Les trois symptômes principaux qui conduisent à la découverte de cette pathologie sont l'apparition inexplicable d'ataxie, les signes cliniques de l'atteinte des voies longues et de l'atteinte d'une paire crânienne (**CALMON [2017]**). Ils se traduisent par des symptômes cliniques communément signalés : mouvements oculaires anormaux ou limités, diplopie, sourire asymétrique, maladresse, difficulté à marcher, perte d'équilibre et faiblesse. Le délai entre les premiers signes et le diagnostic est généralement inférieur à 3 mois. Cependant, comme tous ces symptômes ne sont pas spécifiques du **GITC**, la présentation clinique seule n'est pas pathognomonique.

L'**IRM** est la modalité d'imagerie médicale de référence pour diagnostiquer le **GITC** qui se présente comme une lésion expansive centrée sur la protubérance et envahissant plus de 50% de la protubérance. Le **GITC** présente un signal élevé dans les séquences pondérées en **T2w** et **FLAIR** et un signal faible dans les séquences pondérées en **T1w**. Après injection de produit de contraste, le rehaussement est faible ou absent, parfois une cocarde est visible. La figure 2.1 montre des coupes d'images **T1w**, **T2w**, **T1c** et **FLAIR** d'un patient atteint de **GITC**. Ainsi, le diagnostic du **GITC** a été longtemps fait uniquement à partir des aspects cliniques et radiologiques.

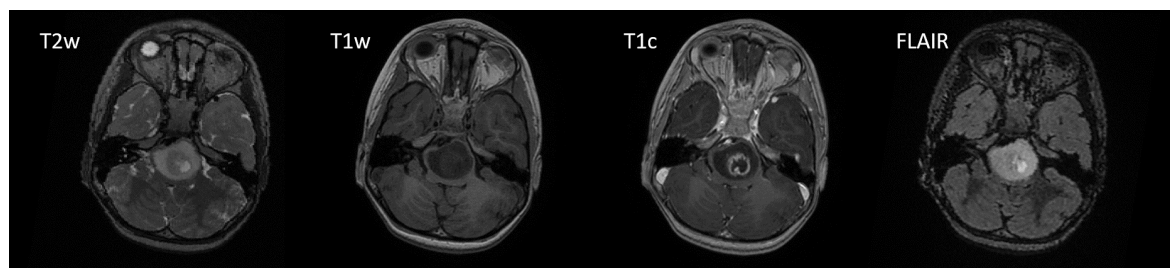


FIGURE 2.1 – Exemple d'un GITC sur les séquences IRM pondérées en T2, T1, T1 après contraste et FLAIR

2.2 Biopsie et sous-types moléculaires

Jusqu'à il y a une dizaine d'années, pour les cas de **GITC** pour lesquels l'imagerie étaient typique, la biopsie était considérée comme un geste non pertinent et dangereux. La biopsie était admise seulement dans les cas radiologiquement ou cliniquement atypiques (**ALBRIGHT et collab. [1993]**; **CARTMILL et PUNT [1999]**).

Néanmoins, l'étude de **ROUJEAU et collab. [2007]** a démontré que la biopsie du tronc peut être menée en sécurité si elle est réalisée par un neurochirurgien pédiatrique expérimenté. Ainsi, à partir de 2007, la tendance à effectuer des biopsies du tronc au moment du diagnostic a crû. Grâce à l'analyse génomique de ces biopsies, il a été trouvé que la mutation récurrente K27M des gènes de l'histone H3 caractérise jusqu'à 90% des **GITC** (**CASTEL et collab. [2015]**; **JONES et collab. [2016]**; **WU et collab. [2012]**). Ces "oncohistones" originellement identifiés dans les gliomes de haut-grade pédiatriques, sont désormais également identifiés pour 2 types de tumeurs osseuses et pour certains types très rares de leucémies (figure 2.2).

Dans les **GITC**, les "oncohistones" affectent les gènes de la variante H3.3 (H3F3A), et les gènes de la variante H3.1 (HIST1H3B/C). Il existe également un cas de **GITC** présentant une mutation du gène HIST2H3C responsable de la variante d'histone H3.2 (**CASTEL et collab. [2015]**). Les mutations H3K27M sont souvent associées avec d'autres mutations comme la mutation du gène TP53 ou celle du gène ACVR1.

Les patients du sous-groupe H3.1K27M sont plus jeunes et sont associés à un meilleur pronostic. Le tableau 2.1 montre les profils des trois sous-groupes de patients **GITC** communément admis : H3.1K27M, H3.3K27M et H3 K27M *wild-type* (WT).

La mutation du gène ACVR1, découverte plus récemment, est apparemment très associée au

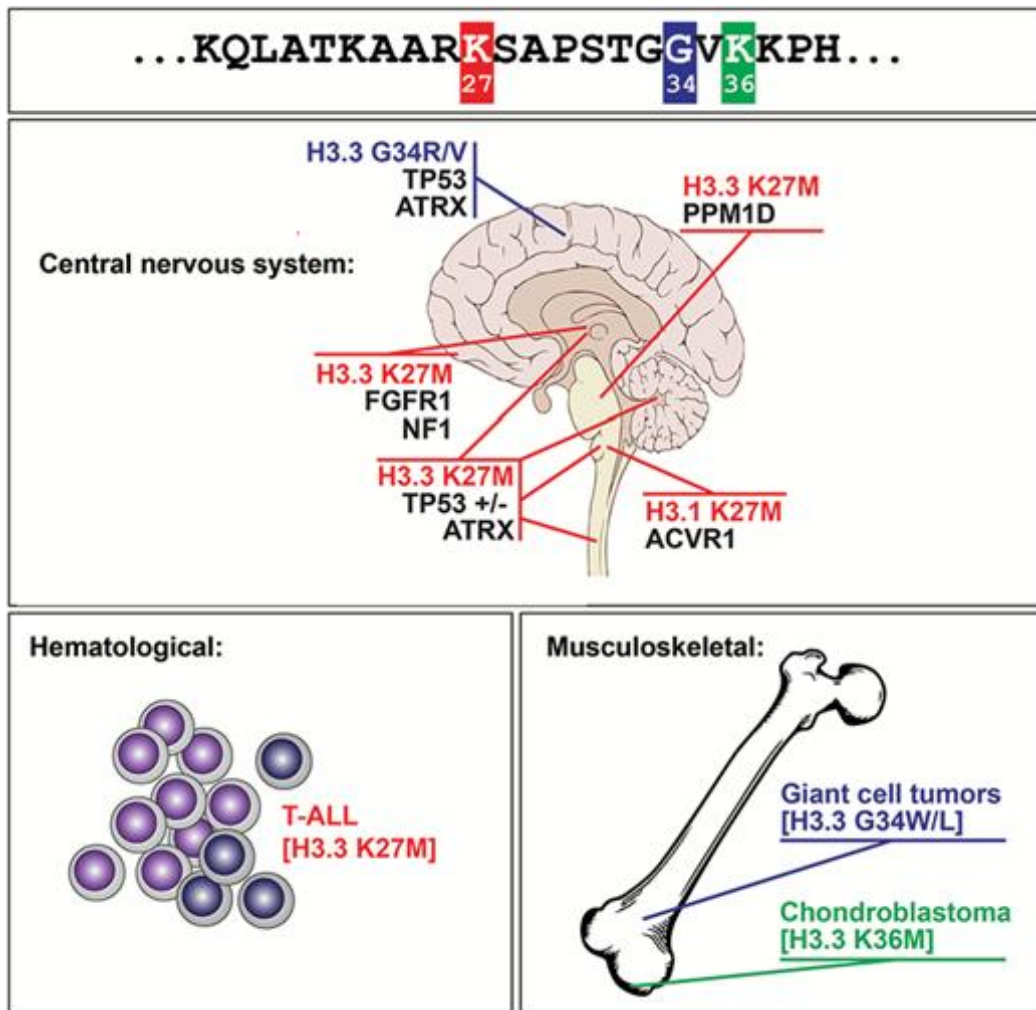


FIGURE 2.2 – Les mutations de "oncohistones" connues dans les maladies chez l'homme. Source : Kenneth J. Cohen (COHEN et collab. [2017])

	H3.3K27M muté	H3.1K27M muté	H3K27M wild-type
Proportion	60%	30%	10%
Médiane de l'âge au diagnostic	7,5 ans	5,5 ans	6,0 ans
Réponse à la radiothérapie	Mauvaise	Bonne	Mauvaise
Médiane de la survie globale	9 mois	15 mois	10 mois
Mutations fréquemment associées	TP53, PPM1D, voie mTOR (RTK)	ACVR1, voie mTor (RTK)	ACVR1
Présence de métastases leptoméningées	Environ 50%	Rares	Indéterminée
Altération du nombre de copies	Amplification : PDGFRA	Gain des chromosomes 1p et 2	Gain Nmyc et ID2

TABEAU 2.1 – Profils des trois principaux sous-groupes de patients atteints de GITC (COHEN et collab. [2017]; MACKAY et collab. [2017])

sous-groupe H3.1K27M, présentant un meilleur pronostic (BUCZKOWICZ et collab. [2014]; TAYLOR et collab. [2014]; WU et collab. [2014]). Cependant, selon l'étude de CASTEL et collab. [2015], l'impact de la mutation H3.1K27M est plus fort que celui de la mutation ACVR1. D'un autre côté, des altérations liées à la mutation de la protéine p53 ou de la protéine PPMD1, ont été fortement associées au sous-groupe H3.3K27M, présentant un plus mauvais pronostic.

Toutes ces informations moléculaires ont impacté la classification des tumeurs cérébrales de l'OMS et peuvent aujourd'hui guider la stratification de différentes thérapies.

2.3 GITC selon la classification de l'OMS

En 2016, l'OMS a actualisé la classification des tumeurs cérébrales qui se base conjointement sur le phénotype histologique et le génotype moléculaire (LOUIS et collab. [2016]).

Dans cette classification, les tumeurs cérébrales pédiatriques sont séparées des tumeurs cérébrales adultes. Les gliomes infiltrants pédiatriques de haut grade, diffèrent des tumeurs chez l'adulte par trois aspects : le nombre de copies d'ADN, les types de mutations rencontrées, et les expressions de gène (CALMON [2017]).

Ainsi, au premier niveau, les GITC sont classifiés histomorphologiquement comme des gliomes infiltrants de grade IV. Au deuxième niveau, les analyses moléculaires contribuent au diagnostic final car les mutations IDH sont rares chez l'enfant alors que la mutation K27M des histones H3 est propre à l'enfant et à l'adolescent. Le GITC est ainsi inclus dans un groupe spécifique et unique de la classification des tumeurs du système nerveux central de l'OMS sous le nom de "*Diffuse midline glioma, H3-K27M-mutant*" (LOUIS et collab. [2016]) en combinant les informations des deux premiers niveaux. Finalement, des informations moléculaires spécifiques peuvent affiner le choix de prise en charge et le pronostic du patient, comme l'information sur la variante de l'histone H3-K27M-mutant (H3.1 ou H3.3) ou sur la mutation du gène TP53.

2.4 Approches thérapeutiques

La radiothérapie est le seul traitement qui a fait preuve d'efficacité en prolongeant le temps de survie globale des patients atteints de GITC. Par conséquent la radiothérapie est le traitement standard des GITC (COHEN et collab. [2017]). Une dose totale de 54-60 Gy est délivrée sur 6 semaines, divisée en fractions quotidiennes de 1,8-2,0 Gy. Ce traitement produit des améliorations transitoires dans les fonctions neurologiques et une augmentation du temps de survie sans progression (KEBUDI et CAKIR [2013]).

Des nouvelles découvertes sur la physiopathologie moléculaire des GITC ouvrent la possibilité de traitements ciblés. Il est peu probable que les GITC soient curables avec un seul médicament, mais plutôt avec une combinaison de traitements visant différentes cibles (COHEN et collab. [2017]).

Des modèles pré-cliniques ont été mis en place, ce qui a permis plusieurs tests *in vitro* et *in vivo*. Même si le GITC est résistant à la majorité des chimiothérapies classiques (GRASSO et collab. [2015]), quelques agents se sont montrés efficaces *in vitro* (GRASSO et collab. [2015]; NAGARAJA et collab. [2017]; PIUNTI et collab. [2017]; ZHANG et collab. [2017]). Cependant, l'efficacité de la plupart de ces agents est limitée par la faible pénétration des drogues au delà de la barrière hémato-encéphalique. A ce jour, différentes drogues sont en cours de test dans le projet BIOMEDE (programme hospitalier de recherche clinique (PHRC) commencé en 2014) sur des cohortes de patients GITC, ayant différentes cibles. Dans le projet BIOMEDE 2.0 qui suivra en 2020 le programme BIOMEDE, quatre drogues (dasatinib, erlotinib, évérolimus et mébendazole) seront testées en combinaison avec la radiothérapie. Pour les patients ayant des informations moléculaires à partir de la biopsie de diagnostic, ces informations seront utilisées pour le choix de la drogue. Les patients sans informations moléculaires seront randomisés pour le choix de la drogue.

D'autres thérapies sont en cours d'essai comme le ONC201 qui a été testé sur une patiente GITC qui n'avait pas répondu à la radiothérapie. Après le début du traitement avec ONC201, le

volume de la tumeur a réduit et la patiente est toujours vivante 22 mois après le diagnostic, avec une amélioration clinique considérable (HALL et collab. [2019]).

2.5 Rôle de l'IRM

L'imagerie permet l'analyse de la tumeur complète et de suivre dans le temps les réponses des patients aux traitements. Comme déjà dit au chapitre précédent, l'IRM est utilisée pour le diagnostic des patients GITC. L'IRM, avec les différentes séquences disponibles, permet une caractérisation de la tumeur par des images anatomiques pondérées en T1 et T2, des images de perfusion, de diffusion ainsi que par spectroscopie, mesurant la concentration de différents métabolites chimiques.

2.5.1 Images anatomiques

L'analyse de la tumeur sur les images anatomiques est faite surtout pour le diagnostic et la suivi des tumeurs et se base sur des critères de localisation, volume et forme. En 1990, MACDONALD et collab. [1990] ont proposé des critères de classification en neuro-oncologie des tumeurs gliales malignes sus-tentorielles qui se basent sur la modification de la taille des lésions mesurées sur une coupe transversale qui présente la plus grande surface rehaussée par le contraste. Les deux diamètres perpendiculaires de la composante tumorale rehaussée sont mesurés, leur produit est calculé et comparé aux valeurs antérieures. Selon la variation des valeurs du produit des diamètres combinés à des critères cliniques, quatre catégories de réponses sont possibles :

- Réponse complète (CR) lorsque il n'y a plus de tumeur qui se rehausse, en l'absence de corticoïdes et chez un patient cliniquement stable ou dont l'état s'est amélioré;
- Réponse partielle (PR) lorsque le produit des diamètres a réduit de plus de 50%, sans augmentation des corticoïdes et chez un patient cliniquement stable ou dont l'état s'est amélioré;
- Progression (PD) lorsque il y a une augmentation de plus de 25% du produit des diamètres perpendiculaires ou l'apparition d'un nouveau foyer tumoral, ou la dégradation de l'état clinique sans diminution des corticoïdes;
- Maladie stable (SD) pour les autres cas.

Les critères de Macdonald ont été largement adoptés dans les études de neuro-oncologie et en pratique clinique. Cependant, ils ne contiennent ni de méthode d'évaluation des tumeurs sans composante rehaussée, ni les définitions des lésions mesurables et non-mesurables ELLINGSON et collab. [2017]. Ainsi, le groupe d'étude *Response Assessment in Neuro-Oncology* (RANO), un groupe multidisciplinaire international dont l'objectif est d'améliorer les critères de réponse et de définir des paramètres pour les études de neuro-oncologie HUANG et WEN [2016], a publié des critères révisés pour l'évaluation de la réponse en neuro-oncologie en 2010 WEN et collab. [2010].

Les critères RANO introduisent la possibilité d'une progression tumorale sans rehaussement, très commun en neuro-oncologie, lorsque la composante T2w ou FLAIR augmente sans qu'un facteur de comorbidité pouvant expliquer l'augmentation ne soit présent.

2.5.2 Images de perfusion

Pour aller plus loin, JASPAN et collab. [2016] ont proposé une extension des critères RANO pour l'évaluation des gliomes pédiatriques de haut-grade. Ils proposent l'intégration des images de perfusion et de diffusion, tout comme l'identification de la pseudo-progression des tumeurs, un phénomène après radiothérapie mimant les signes de progression tumorale, mais de résolution spontanée (DE WIT et collab. [2004]).

DANGOULOFF-ROS et collab. [2016] ont montré que les tumeurs de haut grade présentent des valeurs plus élevées de flux sanguin comparées aux tumeurs de bas grade. Dans cette étude, le flux

était mesuré par des méthodes de perfusion sans injection de produit de contraste, par marquage des spins artériels (ASL). En proposant un seuil de 50 mL/min par 100g de tissu, ils ont obtenu une sensibilité de 65% et une spécificité de 94% pour les tumeurs de la fosse postérieure, y compris le GITC. Les études de [SEDLACIK et collab. \[2013\]](#) et [CALMON et collab. \[2017a\]](#) ont montré que le flux sanguin tumoral augmente pendant la radiothérapie puis diminue progressivement et que la survie des patients atteints de GITC est associée à l'importance de l'augmentation de la perfusion. Une augmentation plus importante du flux sanguin après radiothérapie a été observée chez les patients présentant une pseudo-progression par [CALMON et collab. \[2017b\]](#).

2.5.3 Images de diffusion

L'hypothèse que les tumeurs ont une densité cellulaire et un rapport noyau-cytosol élevé, augmentant la résistance à la diffusion a été confirmée par de nombreuses études. Ainsi, les gliomes de haut grade présentent généralement des valeurs d'ADC plus basses que les gliomes de bas grade ([EIDEL et collab. \[2016\]](#); [ELLINGSON et collab. \[2010\]](#); [GUO et collab. \[2002\]](#)).

Les valeurs d'ADC des GITC mesurées avant traitement semblent positivement associées à la survie des patients ([CHEN et collab. \[2010\]](#)). Selon une étude de [LOBER et collab. \[2014\]](#), les patients porteurs d'un gliome infiltrant du tronc cérébral, dont la valeur d'ADC de la tumeur est au-dessus du seuil de $1,3 \mu m^2/ms$ auraient une survie médiane bien supérieure à celle des patients dont l'ADC de la tumeur est en dessous de ce seuil. En revanche, [CARCELLER et collab. \[2019\]](#) ont conclu récemment que l'ADC après radiothérapie n'est pas une valeur prédictive pour les patients atteints de gliomes de haut-grade et ceux atteints de GITC.

La forme des histogrammes des valeurs d'ADC est aussi corrélée à la survie des patients. Une survie plus longue est associée aux lésions dont l'histogramme des valeurs d'ADC présentent une distribution bimodale. De plus, les histogrammes présentant des coefficients de dissymétrie plus élevés sont associés à des survies plus courtes ([POUSSAINT et collab. \[2015\]](#)).

En outre, la technique de tenseur de diffusion (DTI) a été utilisée pour démontrer que dans le cas des GITC les fibres axonales traversant le tronc cérébral sont déplacées. L'altération des métriques de tenseur de diffusion semble indiquer une extension longitudinale des tumeurs plus précocement que l'altération de signal liée à l'imagerie morphologique. [CESCHIN et collab. \[2015\]](#) ont publié une étude pour distinguer la pseudo-progression de la vraie progression chez 21 patients traités par vaccin à base de peptides, avec des résultats prometteurs.

2.5.4 Spectroscopie

Il n'existe pas beaucoup d'études sur la spectroscopie des patients GITC et elles sont majoritairement anciens et contradictoires entre elles. [LAPRIE et collab. \[2005\]](#) ont montré que la réponse à la radiothérapie est associée à une réduction du rapport Cho / NAA et à la disparition des pics de lipides et de lactate. Néanmoins, [THAKUR et collab. \[2006\]](#) ont réalisé une analyse longitudinale de deux patients n'ayant pas présenté de réduction du rapport Cho / NAA après radiothérapie, et une apparition de lactate et de lipides, alors que ces patients présentaient une amélioration clinique et une réduction du volume tumoral. En revanche [LÖBEL et collab. \[2016\]](#) n'ont trouvé aucun changement significatif du rapport Cho / NAA du tissu tumoral au fil du temps.

Par ailleurs, [PANIGRAHY et collab. \[2008\]](#) suggèrent que le ratio Cho / NAA peut détecter la progression de la tumeur plusieurs semaines avant l'imagerie morphologique et [STEFFEN-SMITH et collab. \[2011\]](#) indiquent que l'index Cho / NAA au moment du diagnostic peut prédire la survie des patients.

Finalement, la présence de citrate dans des tumeurs cérébrales pédiatriques, notamment dans les GITC est décrite par [SEYMOUR et collab. \[2008\]](#). Le citrate est un métabolite intermédiaire du cycle de Krebs. Selon [YEOM et collab. \[2015\]](#) sa concentration est un marqueur d'hypoxie tumorale et inversement corrélée à la perfusion tissulaire.

	Risque intermédiaire vs. risque standard	Risque élevé vs. risque standard	Risque élevé vs. risque intermédiaire
Rapport de hasards (95% intervalle de confiance); valeur-p	1,29 (0,9–1,68); 0,20	1,67 (1,26–2,08); 0,014	1,29 (1,00–1,58); 0,09

TABEAU 2.2 – Résultats de test de validation externe obtenus par [VELDHUIJZEN VAN ZANTEN et collab. \[2017\]](#) avec la signature de survie conçue par [JANSEN et collab. \[2014\]](#).

2.6 Radiomique pour la caractérisation du GITC

Aucune étude n'a été trouvée par une recherche dans PubMed sur la radiomique pour les patients atteints de [GITC](#) en mai 2019. On peut supposer que cela est dû au fait que le terme "*radiomics*" est très récent (2012) de même que la nouvelle définition du [GITC](#) (2016). De plus, le [GITC](#) étant une maladie très rare, la constitution de bases de données avec des protocoles d'imagerie homogènes s'avère compliquée.

Comme cité dans les paragraphes précédents, les paramètres de volume structural, de perfusion et de diffusion sont en cours d'investigation pour la prédiction de la malignité des tumeurs, de la réponse au traitement et du temps de survie. Ce type d'études s'approche des études radiomiques, car elles essaient de prédire le pronostic des patients ayant des tumeurs, avec des informations radiologiques. Cependant, elles n'utilisent pas un nombre élevé d'indices ni des indices d'imagerie décrivant l'hétérogénéité tumorale, caractéristique fréquemment retrouvée pour les études radiomiques.

En cherchant à caractériser l'hétérogénéité tumorale, l'étude de [HARWARD et collab. \[2018\]](#) a créé un indice d'hétérogénéité calculé par le rapport entre l'écart-type des intensités des voxels dans la tumeur et l'écart-type des intensités des voxels du corps calleux (considéré comme région de référence). Ils ont conclu que pour les 17 patients [GITC](#) inclus dans l'étude, les images [T2w](#) représentent mieux l'hétérogénéité tumorale que les images [T1c](#), [FLAIR](#) et [ADC](#), ce qui peut caractériser les foyer d'hypercellularité. Cependant, aucune relation n'a été trouvée entre cet indice et la survie des patients. Encore une fois, cette étude s'approche d'une étude radiomique mais utilise seulement un indice.

Plus récemment, [COLAFATI et collab. \[2019\]](#) ont montré avec une cohorte de 29 patients atteints de [GITC](#) que la présence des complications du nerf crânien, identifiées sur des images d'[IRM](#) conventionnelles, est associée à un mauvais pronostic.

S'approchant davantage de la radiomique, en 2014, [JANSEN et collab. \[2014\]](#) ont proposé une signature pour la prédiction de la survie des patients atteints de [GITC](#). Les patients ont été séparés en trois groupes : risque standard, risque intermédiaire et risque élevé. La signature a été estimée avec 316 patients [GITC](#) est composée de trois paramètres cliniques favorables (âge, durée des symptômes, traitement par chimiothérapie) et d'un indice radiologique qualitatif défavorable (présence d'un anneau de rehaussement sur les images [T1c](#) au moment du diagnostic, comme sur l'illustration de la figure 2.1). Plus récemment, [VELDHUIJZEN VAN ZANTEN et collab. \[2017\]](#) ont validé la signature avec 249 patients indépendants. Ils ont obtenu des résultats de [RH](#) corrects (tableau 2.2). Dans ces études, le nombre de patients important a été possible car l'utilisation d'un indice qualitatif d'imagerie permet l'utilisation d'images de différentes qualités faisant appel à différents paramètres d'acquisition et provenant de différents protocoles.

Enfin, [HARGRAVE et collab. \[2008\]](#) ont réalisé une analyse rétrospective de 39 patients avec 13 indices radiologiques et ils ont conclu que l'imagerie anatomique n'est pas capable de prédire la survie des patients. Néanmoins, les indices utilisés étaient plutôt qualitatifs, de type binaire ou catégoriel, comme présence de nécrose (oui ou non), hétérogène en [T1w](#) (oui ou non).

La compilation de ces résultats suggère que les indices radiologiques simples ne sont pas des bons prédicteurs de la survie; cependant, si on les combine avec des paramètres cliniques, la performance peut s'améliorer. Pour donner une continuité à ces études, la pertinence de la radio-

mique pour la prédiction du pronostic des patients doit être investiguée. L'hypothèse associée est que les indices quantitatifs, de premier-ordre ou de texture, portent des informations pertinentes, non repérables par les indices radiologiques simples testés précédemment.

Or, comme montré dans le chapitre 1, la moitié des études radiomiques en IRM en neuro-oncologie s'intéressent aux GBM chez l'adulte, et obtiennent des résultats prometteurs. Comme les GITC s'approchent des GBM, cela suggère la pertinence de réaliser des études radiomiques pour les patients atteints de GITC.

Comme pour le GBM, les questions d'intérêt de la radiomique pour les patients atteints de GITC sont :

- la détermination d'informations génétiques et moléculaires, notamment la mutation de l'histone H3,
- la prédiction de la réponse au traitement, notamment la pseudo-progression,
- la survie globale ou sans progression des patients.

Récemment LOBON-IGLESIAS et collab. [2018] se sont intéressés au temps de survie après la rechute des patients atteints de GITC pour choisir ceux qui seraient éligibles pour tester différentes thérapies après rechute. Dans cette étude, les auteurs concluent que l'indépendance aux stéroïdes, le statut de la mutation de l'histone H3 et le score de Lansky devraient être considérés pour prédire l'éligibilité des patients. Aucun paramètre radiologique n'a été inclus.

2.7 Bases de données de patients atteint de GITC

Deux bases de données ont été utilisées dans l'ensemble de ce travail de thèse : PREBIOMEDE et une partie de l'étude BIOMEDE. Les familles des patients ont autorisé l'utilisation, pour la recherche, des données cliniques et radiologiques acquises pendant la prise en charge. Les consentements ont été obtenus en accord avec le *Institutional Review Board (IRB)* de *Gustave Roussy*. Les images ont été acquises au format *Digital imaging and communications in medicine (DICOM)* et ont été systématiquement anonymisées et transformées en format *Neuroimaging Informatics Technology Initiative (NIFTI)* pour tout de ce travail de thèse.

2.7.1 PREBIOMEDE

La base de données PREBIOMEDE n'est pas à strictement parler liée à une étude clinique. Elle concerne les patients diagnostiqués à l'*hôpital Necker enfants malades* entre les années 2009 et 2014. L'accès à cette base de données a été progressivement étendu pendant la durée de la thèse. Un résumé des principales caractéristiques des patients inclus est disponible dans le tableau 2.3. La majorité des patients a été soumise à un protocole d'imagerie qui inclut des images T1w, T2w, T1c et FLAIR au moment du diagnostic. Un prélèvement tumoral a été réalisé par biopsie pour déterminer les mutations génétiques associées avec la tumeur (pour certains patients, la biopsie n'a pas été possible ou n'a pas apporté de résultats concluants). Ensuite, plusieurs examens de suivi par imagerie ont eu lieu après la biopsie et au long du traitement. Les informations détaillées sont fournies en annexe, dans le tableau A.1.

Les paramètres d'acquisition d'imagerie pour cette cohorte figurent dans le tableau 2.4.

Cohorte étoile : L'ensemble d'images des deux premières sessions (examens diagnostic et après biopsie) pour les 30 patients (en gras sur le tableau A.1, 17 filles et 13 garçons de 3 à 15 ans) est appelé *cohorte étoile* dans ce document. Tous ces patients présentaient les 4 modalités d'IRM structurale au moment de l'examen "diagnostic" et au moins une de ces modalités dans le premier examen de suivi. Toutes les acquisitions IRM ont été réalisées à l'hôpital Necker avec un imageur 1,5 Tesla (*Signa HDxt/Discovery MR450, GE Medical Systems, Milwaukee, WI, USA*) avec quasiment les mêmes paramètres pour toutes les acquisitions (tableau 2.5).

TABLEAU 2.3 – Résumé de caractéristiques des patients avec GITC de la cohorte PREBIOMEDE.

Nombre de patients	48
Filles / garçons	28 / 20
Âge moyen au moment du diagnostic	7,0 ± 3,2 ans
H3.1 / H3.3 / H3.2 ou WT / Inconnu	12 / 26 / 4 / 6
ACVR1 / WT / Inconnu	9 / 30 / 9
TP53 / WT / Inconnu	22 / 17 / 9
Survie globale moyenne ¹	409 ± 240 jours
Nombre de sessions d'imagerie	296
Nombre d'images T1w	253
Nombre d'images T1c	244
Nombre d'images T2w	274
Nombre d'images FLAIR	241

¹ En ne considérant que les 46 patients pour lesquels la date de décès est connue.

TABLEAU 2.4 – Paramètres d'acquisition d'imagerie pour les images de la cohorte PREBIOMEDE.

	T1w (253 images)	T1c (244 images)	T2w (274 images)	FLAIR (241 images)
TE (ms)	2,5–20,0	2,0–11,58	20,0–136,7	86,00–388,00
TR (ms)	7,0–1.800	7,1–1.930	880,0–8.000	8.400–9.702
Taille du pixel (mm)	0,39–0,94	0,23–0,94	0,23–0,98	0,43–1,20
Ép. de coupe (mm)	1,4–6,5	0,5–4,6	0,5–6,5	0,5–6,5
Champ magnétique	241 à 1,5T / 11 à 3T	236 à 1,5T / 8 à 3T	263 à 1,5T / 10 à 3T	231 à 1,5T / 10 à 3T
Centre	Plusieurs	Plusieurs	Plusieurs	Plusieurs

TABLEAU 2.5 – Paramètres d'acquisition d'imagerie pour les sessions avant et après biopsie des 30 patients de la (cohorte étoile).

	T1w ¹ (50 images)	T1c ² (48 images)	T2w (57 images)	FLAIR ³ (41 images)
TE (ms)	3,6–5,4	3,6–6,8	93,0–115,1	129,9–172,8
TR (ms)	12,0–12,7	12,0–15,2	2.500	8.400–9.002
Flip angle (°)	13	13	90	90
Plan	Axial	Axial	Sagittal	Sagittal
Taille du pixel (mm)	0,43–0,47	0,43–0,47	0,43–0,49	0,43–0,53
Épaisseur de coupe (mm)	1,4–1,6	1,2–1,6	1,0	1,2–1,4
Champ magnétique	1,5T	1,5T	1,5T	1,5T
Centre	Necker	Necker	Necker	Necker

¹ Sauf pour la deuxième session du patient sub-9944, où TE = 10,2 ms, TR = 500 ms, plan d'acquisition sagittal et épaisseur de coupe 1,0 mm.

² Sauf pour la première session du patient sub-5951, où TE = 10,1 ms, TR = 600 ms et plan d'acquisition sagittal.

³ Sauf pour la première session du patient sub-5272, où le plan d'acquisition est sagittal et l'épaisseur de coupe de 4,0 mm.

2.7.2 BIOMEDE

La base de données appelée BIOMEDE dans cette thèse concerne une partie des patients du PHRC National Cancer 2011 BIOMEDE. Le PHRC est une étude multi-centrique s'étalant entre les années 2014 et 2019. Nous avons eu accès à une partie de cette base en février 2019. Un résumé des principales caractéristiques des patients disponibles pour ce travail de thèse est présenté dans le tableau 2.6. Le protocole d'imagerie utilisé était identique à celui utilisé pour PREBIOMEDE. Les différences entre PREBIOMEDE et BIOMEDE sont liées aux thérapies, dans lesquelles, la radiothérapie est combinée avec trois différentes thérapies ciblées : dasatinib, erlotinib et temsirolimus. Les informations détaillées sont données en annexe, dans le tableau A.2.

TABLEAU 2.6 – Résumé de caractéristiques des patients avec GITC de la cohorte BIOMEDE disponibles pour la thèse (en février 2019).

Nombre de patients	38
Filles / garçons	12 / 26
Âge moyen au diagnostic	9,0 ± 3,8 ans
H3.1 / H3.3 / H3.2 ou WT / Inconnu	5 / 20 / 1 / 12
ACVR1 / WT / Inconnu	7 / 21 / 10
TP53 / WT / Inconnu	14 / 12 / 12
Survie globale moyenne ¹	323 ± 142 jours
Nombre de sessions d'imagerie	176
Nombre d'images T1w	129
Nombre d'images T1c	132
Nombre d'images T2w	162
Nombre d'images FLAIR	131

¹ En ne considérant que les 28 patients pour lesquels la date de décès est connue.

Les paramètres d'acquisition d'imagerie pour cette cohorte sont présentés dans le tableau 2.7.

TABLEAU 2.7 – Paramètres d'acquisition d'imagerie pour l'ensemble des images de la cohorte BIOMEDE disponibles en février 2019.

	T1w (129 images)	T1c(132 images)	T2w (162 images)	FLAIR (131 images)
TE (ms)	3,0–17,8	2,1–16,0	30,0–150,0	88,0–310,0
TR (ms)	6,9–789	6,1–2.000	2.400–12.000	4.800–11.000
Taille du pixel (mm)	0,43 – 1,00	0,38 – 0,94	0,33 – 1,88	0,39 – 1,13
Ép. de coupe (mm)	0,50 – 7,00	0,30 – 3,0	0,50 – 3,50	0,50 – 6,50
Champ magnétique	36 1,5T / 93 3T	39 1,5T / 93 3T	60 1,5T / 102 3T	38 1,5T / 93 3T
Centre	Plusieurs	Plusieurs	Plusieurs	Plusieurs

Pour conclure, les deux bases de données PREBIOMEDE et BIOMEDE présentaient des données manquantes sur les quatre modalités structurales d'IRM au moment du diagnostic. Des images de perfusion et de diffusion étaient aussi disponibles pour la majorité des examens. Finalement, en 2020, le nouveau PHRC BIOMEDE 2.0 sera ouvert et introduira deux nouveaux médicaments dans la prise en charge thérapeutique des patients atteints du GITC.

Conclusion

Le gliome infiltrant du tronc cérébral (GITC) est une tumeur pédiatrique rare et létale. Pendant des années, le diagnostic a reposé uniquement sur des symptômes cliniques et l'observation de la tumeur en IRM. Depuis 2007, avec des biopsies plus systématiques au moment du diagnostic, la caractérisation moléculaire du GITC a été possible. Ainsi, la présence fréquente de la mutation K27M des gènes des histones H3 a été révélée. De plus, il a été montré que les variantes des mutations sont liées à des pronostics différents. A cette date, la radiothérapie est le seul traitement qui ait fait preuve d'efficacité pour augmenter le temps de survie globale des patients. Néanmoins différents agents chimiothérapeutiques sont en cours d'investigation. Avec différentes modalités d'IRM, il a été montré qu'une réduction importante du volume lésionnel après la radiothérapie et des valeurs d'ADC plus élevées avant traitement sont associées à un pronostic plus favorable. En outre l'augmentation plus importante du flux sanguin après radiothérapie a été observée chez les patients présentant une pseudo-progression. A ce jour, aucune étude radiomique sur les GITC n'a été publiée, si ce n'est des études utilisant des paramètres cliniques et d'imagerie simples ou catégoriels. Finalement, deux bases de données de patients GITC obtenues de façon successive faisant appel à des protocoles d'acquisition similaires ont été explorées pour ce travail de thèse.

2.8 Références

- ALBRIGHT, A. L., R. J. PACKER, R. ZIMMERMAN, L. B. RORKE, J. BOYETT et G. D. HAMMOND. 1993, «Magnetic Resonance Scans Should Replace Biopsies for the Diagnosis of Diffuse Brain Stem Gliomas : A Report from the Children's Cancer Group», *Neurosurgery*, vol. 33, n° 6, doi :10.1097/00006123-199312000-00010, p. 1026–1030, ISSN 0148-396X. URL <https://doi.org/10.1097/00006123-199312000-00010>. 46
- BUCZKOWICZ, P., C. HOEMAN, P. RAKOPOULOS, S. PAJOVIC, L. LETOURNEAU, M. DZAMBA, A. MORRISON, P. LEWIS, E. BOUFFET, U. BARTELS, J. ZUCCARO, S. AGNIHOTRI, S. RYALL, M. BARSZCZYK, Y. CHORNENKYY, M. BOURGEY, G. BOURQUE, A. MONTPETIT, F. CORDERO et C. HAWKINS. 2014, «Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating acvr1 mutations», *Nature genetics*, vol. 46, doi :10.1038/ng.2936. 48
- CALMON, R. 2017, *Serial multimodal MRI of diffuse intrinsic pontine in children*, Theses, Université Pierre et Marie Curie - Paris VI. URL <https://tel.archives-ouvertes.fr/tel-01793106>. 46, 48
- CALMON, R., S. PUGET, P. VARLET, K. BECCARIA, T. BLAUWBLOMME, D. GREVENT, C. SAINTE-ROSE, D. CASTEL, C. DUFOUR, F. DHERMAIN, S. BOLLE, A. SAITOVITCH, M. ZILBOVICIUS, F. BRUNELLE, J. GRILL et N. BODDAERT. 2017a, «Multimodal magnetic resonance imaging of treatment-induced changes to diffuse infiltrating pontine gliomas in children and correlation to patient progression-free survival», *International Journal of Radiation Oncology*Biophysics*Physics*, vol. 99, n° 2, doi :<https://doi.org/10.1016/j.ijrobp.2017.04.007>, p. 476 – 485, ISSN 0360-3016. URL <http://www.sciencedirect.com/science/article/pii/S0360301617308052>. 50
- CALMON, R., S. PUGET, P. VARLET, V. DANGOULOFF-ROS, T. BLAUWBLOMME, K. BECCARIA, D. GREVENT, C. SAINTE-ROSE, D. CASTEL, M.-A. DEBILY, C. DUFOUR, S. BOLLE, F. DHERMAIN, A. SAITOVITCH, M. ZILBOVICIUS, F. BRUNELLE, J. GRILL et N. BODDAERT. 2017b, «Cerebral blood flow changes after radiation therapy identifies pseudoprogression in diffuse intrinsic pontine gliomas», *Neuro-Oncology*, vol. 20, n° 7, doi :10.1093/neuonc/nox227, p. 994–1002, ISSN 1522-8517. URL <https://doi.org/10.1093/neuonc/nox227>. 50
- CARCELLER, F., N. P. JEROME, L. A. FOWKES, K. KHABRA, A. MACKINNON, F. BAUTISTA, L. V. MARSHALL, S. VAIDYA, H. MANDEVILLE, V. MORGAN, M. O. LEACH et D.-M. KOH. 2019, «Post-

- radiotherapy apparent diffusion coefficient (adc) in children and young adults with high-grade gliomas and diffuse intrinsic pontine gliomas», *Pediatric Hematology and Oncology*, vol. 0, n° 0, doi :10.1080/08880018.2019.1592267, p. 1–10. URL <https://doi.org/10.1080/08880018.2019.1592267>, PMID : 30978130. 50
- CARTMILL, M. et J. PUNT. 1999, «Diffuse brain stem glioma. a review of stereotactic biopsies», *Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery*, vol. 15, doi :10.1007/s003810050379, p. 235–7; discussion 238. 46
- CASTEL, D., C. PHILIPPE, R. CALMON, L. LE DRET, N. TRUFFAUX, N. BODDAERT, M. PAGÈS, K. R. TAYLOR, P. SAULNIER, L. LACROIX, A. MACKAY, C. JONES, C. SAINTE-ROSE, T. BLAUWBLOMME, F. ANDREIUOLO, S. PUGET, J. GRILL, P. VARLET et M.-A. DEBILY. 2015, «Histone h3f3a and hist1h3b k27m mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes», *Acta Neuropathologica*, vol. 130, n° 6, doi : 10.1007/s00401-015-1478-0, p. 815–827, ISSN 1432-0533. URL <https://doi.org/10.1007/s00401-015-1478-0>. 46, 48
- CESCHIN, R., B. KURLAND, S. ABBERBOCK, B. ELLINGSON, H. OKADA, R. JAKACKI, I. POLLACK et A. PANIGRAHY. 2015, «Parametric response mapping of apparent diffusion coefficient as an imaging biomarker to distinguish pseudoprogression from true tumor progression in peptide-based vaccine therapy for pediatric diffuse intrinsic pontine glioma», *American Journal of Neuroradiology*, vol. 36, n° 11, doi :10.3174/ajnr.A4428, p. 2170–2176, ISSN 0195-6108. URL <http://www.ajnr.org/content/36/11/2170>. 50
- CHEN, H., A. PANIGRAHY, G. DHALL, J. FINLAY, M. NELSON et S. BLUML. 2010, «Apparent diffusion and fractional anisotropy of diffuse intrinsic brain stem gliomas», *American Journal of Neuroradiology*, vol. 31, n° 10, doi :10.3174/ajnr.A2179, p. 1879–1885, ISSN 0195-6108. URL <http://www.ajnr.org/content/31/10/1879>. 50
- COHEN, K. J., N. JABADO et J. GRILL. 2017, «Diffuse intrinsic pontine gliomas - Current management and new biologic insights. Is there a glimmer of hope?», *Neuro-Oncology*, vol. 19, n° 8, doi :10.1093/neuonc/nox021, p. 11025 – 1034, ISSN 15235866. 47, 48
- COLAFATI, G. S., I. P. VOICU, C. CARDUCCI, M. CAULO, M. VINCI, F. DIOMEDI-CAMASSEI, P. MERLI, A. CARAI, E. MIELE, A. CACCHIONE, P. TOMÀ, F. LOCATELLI et A. MASTRONUZZI. 2019, «Direct involvement of cranial nerve v at diagnosis in patients with diffuse intrinsic pontine glioma : A potential magnetic resonance predictor of short-term survival», *Frontiers in Oncology*, vol. 9, doi :10.3389/fonc.2019.00204, p. 204, ISSN 2234-943X. URL <https://www.frontiersin.org/article/10.3389/fonc.2019.00204>. 51
- DANGOULOFF-ROS, V., C. DEROULERS, F. FOISSAC, M. BADOUAL, E. SHOTAR, D. GRÉVENT, R. CALMON, M. PAGÈS, J. GRILL, C. DUFOUR, T. BLAUWBLOMME, S. PUGET, M. ZERAH, C. SAINTE-ROSE, F. BRUNELLE, P. VARLET et N. BODDAERT. 2016, «Arterial spin labeling to predict brain tumor grading in children : Correlations between histopathologic vascular density and perfusion mr imaging», *Radiology*, vol. 281, n° 2, doi :10.1148/radiol.2016152228, p. 553–566. URL <https://doi.org/10.1148/radiol.2016152228>, PMID : 27257950. 49
- EIDEL, O., J.-O. NEUMANN, S. BURTH, P. J. KIESLICH, C. JUNGK, F. SAHM, P. KICKINGEREDER, K. KIENING, A. UNTERBERG, W. WICK, H.-P. SCHLEMMER, M. BENDSZUS et A. RADBRUCH. 2016, «Automatic analysis of cellularity in glioblastoma and correlation with adc using trajectory analysis and automatic nuclei counting», *PLOS ONE*, vol. 11, n° 7, doi :10.1371/journal.pone.0160250, p. 1–10. URL <https://doi.org/10.1371/journal.pone.0160250>. 50
- ELLINGSON, B. M., C. CHUNG, W. B. POPE, J. L. BOXERMAN et T. J. KAUFMANN. 2017, «Pseudoprogression, radionecrosis, inflammation or true tumor progression? challenges associated with

- glioblastoma response assessment in an evolving therapeutic landscape», *Journal of Neuro-Oncology*, vol. 134, n° 3, doi :10.1007/s11060-017-2375-2, p. 495–504, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-017-2375-2>. 49
- ELLINGSON, B. M., M. G. MALKIN, S. D. RAND, J. M. CONNELLY, C. QUINSEY, P. S. LAVIOLETTE, D. P. BEDEKAR et K. M. SCHMAINDA. 2010, «Validation of functional diffusion maps (fdms) as a biomarker for human glioma cellularity», *Journal of Magnetic Resonance Imaging*, vol. 31, n° 3, doi :10.1002/jmri.22068, p. 538–548. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.22068>. 50
- GRASSO, C. S., Y. TANG, N. TRUFFAUX, N. E. BERLOW, L. LIU, M.-A. DEBILY, M. J. QUIST, L. E. DAVIS, E. C. HUANG, P. J. WOO, A. PONNUSWAMI, S. CHEN, T. B. JOHUNG, W. SUN, M. KOGISO, Y. DU, L. QI, Y. HUANG, M. HUTT-CABEZAS, K. E. WARREN, L. LE DRET, P. S. MELTZER, H. MAO, M. QUEZADO, D. G. VAN VUURDEN, J. ABRAHAM, M. FOULADI, M. N. SVALINA, N. WANG, C. HAWKINS, J. NAZARIAN, M. M. ALONSO, E. H. RAABE, E. HULLEMAN, P. T. SPELLMAN, X.-N. LI, C. KELLER, R. PAL, J. GRILL et M. MONJE. 2015, «Functionally defined therapeutic targets in diffuse intrinsic pontine glioma», *Nature medicine*, vol. 21, n° 6, doi:10.1038/nm.3855, p. 555–559, ISSN 1078-8956. URL <http://europepmc.org/articles/PMC4862411>. 48
- GUO, A. C., T. J. CUMMINGS, R. C. DASH et J. M. PROVENZALE. 2002, «Lymphomas and high-grade astrocytomas : Comparison of water diffusibility and histologic characteristics», *Radiology*, vol. 224, n° 1, doi :10.1148/radiol.2241010637, p. 177–183. URL <https://doi.org/10.1148/radiol.2241010637>, pMID : 12091680. 50
- HALL, M. D., Y. ODIA, J. E. ALLEN, R. TARAPORE, Z. KHATIB, T. N. NIAZI, D. DAGHISTANI, L. SCHALOP, A. S. CHI, W. OSTER et M. P. MEHTA. 2019, «First clinical experience with drd2/3 antagonist onc201 in h3 k27m-mutant pediatric diffuse intrinsic pontine glioma : a case report», *Journal of Neurosurgery : Pediatrics PED*, vol. 23, n° 6, p. 719 – 725. URL <https://thejns.org/pediatrics/view/journals/j-neurosurg-pediatr/23/6/article-p719.xml>. 49
- HARGRAVE, D., N. CHUANG et E. BOUFFET. 2008, «Conventional mri cannot predict survival in childhood diffuse intrinsic pontine glioma», *Journal of Neuro-Oncology*, vol. 86, n° 3, doi : 10.1007/s11060-007-9473-5, p. 313–319, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-007-9473-5>. 51
- HARWARD, S., S. HARRISON FARBER, M. MALINZAK, O. BECHER et E. M. THOMPSON. 2018, «T2-weighted images are superior to other mr image types for the determination of diffuse intrinsic pontine glioma intratumoral heterogeneity», *Child's Nervous System*, vol. 34, n° 3, doi : 10.1007/s00381-017-3659-8, p. 449–455, ISSN 1433-0350. URL <https://doi.org/10.1007/s00381-017-3659-8>. 51
- HUANG, R. Y. et P. Y. WEN. 2016, «Response assessment in neuro-oncology criteria and clinical endpoints», *Magnetic Resonance Imaging Clinics of North America*, vol. 24, n° 4, doi :<https://doi.org/10.1016/j.mric.2016.06.003>, p. 705 – 718, ISSN 1064-9689. URL <http://www.sciencedirect.com/science/article/pii/S1064968916300435>, imaging of Brain Tumors. 49
- JANSEN, M. H., S. E. VELDHUIJZEN VAN ZANTEN, E. SANCHEZ ALIAGA, M. W. HEYMANS, M. WARMUTH-METZ, D. HARGRAVE, E. J. VAN DER HOEVEN, C. E. GIDDING, E. S. DE BONT, O. S. ESHGHI, R. REDDINGIUS, C. M. PEETERS, A. Y. SCHOUTEN-VAN MEETEREN, R. H. GOOSKENS, B. GRANZEN, G. M. PAARDEKOOPER, G. O. JANSSENS, D. P. NOSKE, F. BARKHOF, C. M. KRAMM, W. P. VANDERTOP, G. J. KASPERS et D. G. VAN VUURDEN. 2014, «Survival prediction model of children with diffuse intrinsic pontine glioma based on clinical and radiological criteria», *Neuro-Oncology*, vol. 17, n° 1, doi:10.1093/neuonc/nou104, p. 160–166, ISSN 1522-8517. URL <https://doi.org/10.1093/neuonc/nou104>. 51

- JASPAN, T., P. MORGAN, M. WARMUTH-METZ, E. SANCHEZ ALIAGA, D. WARREN, R. CALMON, J. GRILL, D. HARGRAVE, J. GARCIA et G. ZAHLMANN. 2016, «Response assessment in pediatric neuro-oncology : Implementation and expansion of the rano criteria in a randomized phase ii trial of pediatric patients with newly diagnosed high-grade gliomas», *American Journal of Neuroradiology*, vol. 37, n° 9, doi :10.3174/ajnr.A4782, p. 1581–1587, ISSN 0195-6108. URL <http://www.ajnr.org/content/37/9/1581>. 49
- JONES, C., M. A. KARAJANNIS, D. T. W. JONES, M. W. KIERAN, M. MONJE, S. J. BAKER, O. J. BECHER, Y.-J. CHO, N. GUPTA, C. HAWKINS, D. HARGRAVE, D. A. HAAS-KOGAN, N. JABADO, X.-N. LI, S. MUELLER, T. NICOLAIDES, R. J. PACKER, A. I. PERSSON, J. J. PHILLIPS, E. F. SIMONDS, J. M. STAFFORD, Y. TANG, S. M. PFISTER et W. A. WEISS. 2016, «Pediatric high-grade glioma : biologically and clinically in need of new thinking», *Neuro-Oncology*, vol. 19, n° 2, doi :10.1093/neuonc/nov101, p. 153–161, ISSN 1522-8517. URL <https://doi.org/10.1093/neuonc/nov101>. 46
- KEBUDI, R. et F. B. CAKIR. 2013, «Management of diffuse pontine gliomas in children : Recent developments», *Pediatric Drugs*, vol. 15, n° 5, doi :10.1007/s40272-013-0033-5, p. 351–362, ISSN 1179-2019. URL <https://doi.org/10.1007/s40272-013-0033-5>. 48
- LAPRIE, A., A. PIRZKALL, D. A. HAAS-KOGAN, S. CHA, A. BANERJEE, T. P. LE, Y. LU, S. NELSON et T. R. MCKNIGHT. 2005, «Longitudinal multivoxel mr spectroscopy study of pediatric diffuse brainstem gliomas treated with radiotherapy», *International Journal of Radiation Oncology*Biophysics*, vol. 62, n° 1, doi :<https://doi.org/10.1016/j.ijrobp.2004.09.027>, p. 20 – 31, ISSN 0360-3016. URL <http://www.sciencedirect.com/science/article/pii/S0360301604026653>. 50
- LÖBEL, U., S. HWANG, A. EDWARDS, Y. LI, X. LI, A. BRONISCHER et Z. PATAY. 2016, «Discrepant longitudinal volumetric and metabolic evolution of diffuse intrinsic pontine gliomas during treatment : implications for current response assessment strategies», *Neuroradiology*, vol. 58, n° 10, doi :10.1007/s00234-016-1724-8, p. 1027–1034, ISSN 1432-1920. URL <https://doi.org/10.1007/s00234-016-1724-8>. 50
- LOBER, R. M., Y.-J. CHO, Y. TANG, P. D. BARNES, M. S. EDWARDS, H. VOGEL, P. G. FISHER, M. MONJE et K. W. YEOM. 2014, «Diffusion-weighted mri derived apparent diffusion coefficient identifies prognostically distinct subgroups of pediatric diffuse intrinsic pontine glioma», *Journal of Neuro-Oncology*, vol. 117, n° 1, doi :10.1007/s11060-014-1375-8, p. 175–182, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-014-1375-8>. 50
- LOBON-IGLESIAS, M. J., G. GIRAUD, D. CASTEL, C. PHILIPPE, M. A. DEBILY, C. BRIANDET, F. FOUYSSAC, E. DE CARLI, C. DUFOUR, D. VALTEAU-COUANET, C. SAINTE-ROSE, T. BLAUWBLOMME, K. BECCARIA, M. ZERAH, S. PUGET, R. CALMON, N. BODDAERT, S. BOLLE, P. VARLET et J. GRILL. 2018, «Diffuse intrinsic pontine gliomas (dipg) at recurrence : is there a window to test new therapies in some patients?», *Journal of Neuro-Oncology*, vol. 137, n° 1, doi : 10.1007/s11060-017-2702-7, p. 111–118, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-017-2702-7>. 52
- LOUIS, D. N., A. PERRY, G. REIFENBERGER, A. VON DEIMLING, D. FIGARELLA-BRANGER, W. K. CAVENEY, H. OHGAKI, O. D. WIESTLER, P. KLEIHUES et D. W. ELLISON. 2016, «The 2016 world health organization classification of tumors of the central nervous system : a summary», *Acta Neuropathologica*, vol. 131, n° 6, doi :10.1007/s00401-016-1545-1, p. 803–820, ISSN 1432-0533. URL <https://doi.org/10.1007/s00401-016-1545-1>. 48
- MACDONALD, D. R., T. L. CASCINO, S. C. SCHOLD et J. G. CAIRNCROSS. 1990, «Response criteria for phase ii studies of supratentorial malignant glioma.», *Journal of Clinical Oncology*, vol. 8, n° 7, doi:10.1200/JCO.1990.8.7.1277, p. 1277–1280. URL <https://doi.org/10.1200/JCO.1990.8.7.1277>, PMID : 2358840. 49

- MACKAY, A., A. BURFORD, D. CARVALHO, E. IZQUIERDO, J. FAZAL-SALOM, K. R. TAYLOR, L. BJERKE, M. CLARKE, M. VINCI, M. NANDHABALAN, S. TEMELSO, S. POPOV, V. MOLINARI, P. RAMAN, A. J. WAANDERS, H. J. HAN, S. GUPTA, L. MARSHALL, S. ZACHAROULIS, S. VAIDYA, H. C. MANDEVILLE, L. R. BRIDGES, A. J. MARTIN, S. AL-SARRAJ, C. CHANDLER, H.-K. NG, X. LI, K. MU, S. TRABELSI, D. H.-B. BRAHIM, A. N. KISLJAKOV, D. M. KONOVALOV, A. S. MOORE, A. M. CARCABOSO, M. SUNOL, C. DE TORRES, O. CRUZ, J. MORA, L. I. SHATS, J. N. STAVALE, L. T. BIDINOTTO, R. M. REIS, N. ENTZ-WERLE, M. FARRELL, J. CRYAN, D. CRIMMINS, J. CAIRD, J. PEARS, M. MONJE, M.-A. DEBILY, D. CASTEL, J. GRILL, C. HAWKINS, H. NIKBAKHT, N. JABADO, S. J. BAKER, S. M. PFISTER, D. T. W. JONES, M. FOULADI, A. O. VON BUEREN, M. BAUDIS, A. RESNICK et C. JONES. 2017, «Integrated molecular meta-analysis of 1,000 pediatric high-grade and diffuse intrinsic pontine glioma», *Cancer cell*, vol. 32, n° 4, doi :10.1016/j.ccell.2017.08.017, p. 520—537.e5, ISSN 1535-6108. URL <http://europepmc.org/articles/PMC5637314>. 47
- NAGARAJA, S., N. A. VITANZA, P. J. WOO, K. R. TAYLOR, F. LIU, L. ZHANG, M. LI, W. MENG, A. PONNUSWAMI, W. SUN, J. MA, E. HULLEMAN, T. SWIGUT, J. WYSOCKA, Y. TANG et M. MONJE. 2017, «Transcriptional dependencies in diffuse intrinsic pontine glioma», *Cancer Cell*, vol. 31, n° 5, doi :https://doi.org/10.1016/j.ccell.2017.03.011, p. 635 – 652.e6, ISSN 1535-6108. URL <http://www.sciencedirect.com/science/article/pii/S1535610817301071>. 48
- PANIGRAHY, A., J. NELSON, MARVIN D., J. L. FINLAY, R. SPOSTO, M. D. KRIEGER, F. H. GILLES et S. BLÜML. 2008, «Metabolism of diffuse intrinsic brainstem gliomas in children», *Neuro-Oncology*, vol. 10, n° 1, doi :10.1215/15228517-2007-042, p. 32–44, ISSN 1522-8517. URL <https://doi.org/10.1215/15228517-2007-042>. 50
- PIUNTI, A., R. HASHIZUME, M. MORGAN, E. T BARTOM, C. HORBINSKI, S. MARSHALL, E. RENDLEMAN, Q. MA, Y.-H. TAKAHASHI, A. WOODFIN, A. MISHARIN, N. A ABSHIRU, R. R LULLA, A. M SARATSI, N. KELLEHER, C. JAMES et A. SHILATIFARD. 2017, «Therapeutic targeting of polycomb and bet bromodomain proteins in diffuse intrinsic pontine gliomas», *Nature Medicine*, vol. 23, doi :10.1038/nm.4296. 48
- POUSSAINT, T. Y., S. VAJAPPEYAM, K. I. RICCI, A. PANIGRAHY, M. KOCAK, L. E. KUN, J. M. BOYETT, I. F. POLLACK et M. FOULADI. 2015, «Apparent diffusion coefficient histogram metrics correlate with survival in diffuse intrinsic pontine glioma : a report from the Pediatric Brain Tumor Consortium», *Neuro-Oncology*, vol. 18, n° 5, doi :10.1093/neuonc/nov256, p. 725–734, ISSN 1522-8517. URL <https://doi.org/10.1093/neuonc/nov256>. 50
- ROUJEAU, T., G. MACHADO, M. R. GARNETT, C. MIQUEL, S. PUGET, B. GEOERGER, J. GRILL, N. BODDAERT, F. D. ROCCO, M. ZERAH et C. SAINTE-ROSE. 2007, «Stereotactic biopsy of diffuse pontine lesions in children», *Journal of Neurosurgery : Pediatrics PED*, vol. 107, n° 1, p. 1. 46
- SEDLACIK, J., A. WINCHELL, M. KOCAK, R. LOEFFLER, A. BRONISCHER et C. HILLENBRAND. 2013, «Mr imaging assessment of tumor perfusion and 3d segmented volume at baseline, during treatment, and at tumor progression in children with newly diagnosed diffuse intrinsic pontine glioma», *American Journal of Neuroradiology*, vol. 34, n° 7, doi :10.3174/ajnr.A3421, p. 1450–1455, ISSN 0195-6108. URL <http://www.ajnr.org/content/34/7/1450>. 50
- SEYMOUR, Z., A. PANIGRAHY, J. FINLAY, M. NELSON et S. BLÜML. 2008, «Citrate in pediatric cns tumors?», *American Journal of Neuroradiology*, vol. 29, n° 5, doi :10.3174/ajnr.A1018, p. 1006–1011, ISSN 0195-6108. URL <http://www.ajnr.org/content/29/5/1006>. 50
- STEFFEN-SMITH, E. A., J. H. SHIH, S. J. HIPPI, R. BENT et K. E. WARREN. 2011, «Proton magnetic resonance spectroscopy predicts survival in children with diffuse intrinsic pontine glioma», *Journal of Neuro-Oncology*, vol. 105, n° 2, doi :10.1007/s11060-011-0601-x, p. 365–373, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-011-0601-x>. 50

- TAYLOR, K. R., A. MACKAY, N. TRUFFAUX, Y. BUTTERFIELD, O. MOROZOVA, C. PHILIPPE, D. CASTEL, C. S. GRASSO, M. VINCI, D. CARVALHO, A. M. CARCABOSO, C. DE TORRES, O. CRUZ, J. MORA, N. ENTZ-WERLE, W. J. INGRAM, M. MONJE, D. HARGRAVE, A. N. BULLOCK, S. PUGET, S. YIP, C. JONES et J. GRILL. 2014, «Recurrent activating *acvr1* mutations in diffuse intrinsic pontine glioma», *Nature genetics*, vol. 46, n° 5, doi :10.1038/ng.2925, p. 457—461, ISSN 1061-4036. URL <http://europepmc.org/articles/PMC4018681>. 48
- THAKUR, S., S. KARIMI, I. DUNKEL, J. KOUTCHER et W. HUANG. 2006, «Longitudinal mr spectroscopic imaging of pediatric diffuse pontine tumors to assess tumor aggression and progression», *American Journal of Neuroradiology*, vol. 27, n° 4, p. 806–809, ISSN 0195-6108. URL <http://www.ajnr.org/content/27/4/806>. 50
- WEN, P. Y., D. R. MACDONALD, D. A. REARDON, T. F. CLOUGHESY, A. G. SORENSEN, E. GALANIS, J. DEGROOT, W. WICK, M. R. GILBERT, A. B. LASSMAN, C. TSIEN, T. MIKKELSEN, E. T. WONG, M. C. CHAMBERLAIN, R. STUPP, K. R. LAMBORN, M. A. VOGELBAUM, M. J. VAN DEN BENT et S. M. CHANG. 2010, «Updated response assessment criteria for high-grade gliomas : Response assessment in neuro-oncology working group», *Journal of Clinical Oncology*, vol. 28, n° 11, doi :10.1200/JCO.2009.26.3541, p. 1963–1972. URL <https://doi.org/10.1200/JCO.2009.26.3541>, pMID : 20231676. 49
- DE WIT, M. C., H. G. DE BRUIN, W. EIJKENBOOM, P. A. SILLEVIS SMITT et M. J. VAN DEN BENT. 2004, «Immediate post-radiotherapy changes in malignant glioma can mimic tumor progression», *Neurology*, vol. 63, n° 3, doi :10.1212/01.WNL.0000133398.11870.9A, p. 535–537, ISSN 0028-3878. URL <https://n.neurology.org/content/63/3/535>. 49
- WU, G., A. BRONISCHER, T. MCEACHRON, C. LU, B. PAUGH, J. BECKSFORT, C. QU, L. DING, R. HUETHER, M. PARKER, J. ZHANG, A. GAJJAR, M. A DYER, C. MULLIGHAN, R. GILBERTSON, E. MARDIS, R. WILSON, J. R DOWNING, D. ELLISON et S. J BAKER. 2012, «Somatic histone h3 alterations in paediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas», *Nature genetics*, vol. 44, doi :10.1038/ng.1102, p. 251–3. 46
- WU, G., A. DIAZ, B. PAUGH, S. RANKIN, B. JU, Y. LI, X. ZHU, C. QU, X. CHEN, J. ZHANG, J. EASTON, M. EDMONSON, X. MA, C. LU, P. NAGAHAWATTE, E. HEDLUND, M. RUSCH, S. POUNDS, T. LIN, A. ONAR-THOMAS, R. HUETHER, R. KRIWACKI, M. PARKER, P. GUPTA, J. BECKSFORT, L. WEI, H. MULDER, K. BOGGS, B. VADODARIA, D. YERGEAU, J. RUSSELL, K. OCHOA, R. FULTON, L. FULTON, C. JONES, F. BOOP, A. BRONISCHER, C. WETMORE, A. GAJJAR, L. DING, E. MARDIS, R. WILSON, M. TAYLOR, J. DOWNING, D. ELLISON, J. ZHANG et S. BAKER. 2014, «The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma.», *Nature Genetics*, vol. 46, n° 5, doi :10.1038/ng.2938, p. 444–450. 48
- YEOM, K. W., R. M. LOBER, M. D. NELSON, A. PANIGRAHY et S. BLÜML. 2015, «Citrate concentrations increase with hypoperfusion in pediatric diffuse intrinsic pontine glioma», *Journal of Neuro-Oncology*, vol. 122, n° 2, doi :10.1007/s11060-015-1726-0, p. 383–389, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-015-1726-0>. 50
- VELDHUIJZEN VAN ZANTEN, S. E. M., A. LANE, M. W. HEYMANS, J. BAUGH, B. CHANEY, L. M. HOFFMAN, R. DOUGHMAN, M. H. A. JANSEN, E. SANCHEZ, W. P. VANDERTOP, G. J. L. KASPERS, D. G. VAN VUURDEN, M. FOULADI, B. V. JONES et J. LEACH. 2017, «External validation of the diffuse intrinsic pontine glioma survival prediction model : a collaborative report from the international dipg registry and the siop dipg registry», *Journal of Neuro-Oncology*, vol. 134, n° 1, doi : 10.1007/s11060-017-2514-9, p. 231–240, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-017-2514-9>. 51
- ZHANG, Y., W. DONG, J. ZHU, L. WANG, X. WU et H. SHAN. 2017, «Combination of *ezh2* inhibitor and *bet* inhibitor for treatment of diffuse intrinsic pontine glioma», *Cell and Bioscience*, vol. 7, doi :10.1186/s13578-017-0184-0. 48

Chapitre 3

La standardisation des images IRM cérébrales

« On ne compare pas des pommes avec des oranges. »

Proverbe Québécois



Ce chapitre considère la standardisation des intensités des images d'IRM cérébrales, un pré-traitement essentiel pour les comparaisons quantitatives des images, par conséquent pour la radiomique. Pour commencer, un bref état de l'art des méthodes existantes est présenté, ensuite, la méthode *hybrid White Stripe* et les adaptations adéquates que nous avons réalisées sont décrites. Le pipeline automatique de standardisation et le contrôle de qualité automatisé mis en place sont ensuite présentés. Les méthodes proposées ont été appliquées à deux cohortes de patients atteints de GTC et à une cohorte de patients présentant différentes lésions cérébrales acquises sur deux machines différentes à quelques semaines d'intervalle à champ magnétique d'intensités 1,5T et 3T. Les résultats présentés à la fin de ce chapitre montrent l'efficacité de la standardisation dans les tissus tumoraux et sains du cerveau.

Sommaire

3.1 État de l'art	62
3.2 La méthode de standardisation proposée	62
3.3 Définition d'un pipeline de contrôle de qualité	66
3.4 Application des méthodes de standardisation aux cohortes GTC	69
3.5 Application des méthodes de standardisation sur une cohorte indépendante présentant des lésions cérébrales	77
3.6 Discussion et conclusion	80
3.7 Références	88

3.1 État de l'art

Différemment de la TEP ou du scanner, les intensités en IRM ne sont pas exprimées en unité standard. A cause des non-homogénéités du champ magnétique, des excitations de radio-fréquence et des méthodes de reconstruction d'image, entre autres, les gammes de valeurs sont très variables entre plusieurs images acquises par des séquences similaires sur des machines différentes, voire sur une seule machine à des temps différents. Étant donné que les intensités d'images brutes ne sont pas équivalentes pour un même patient et entre patients, la standardisation des intensités est indispensable pour la réalisation de tâches de comparaison intra et inter-sujet. C'est pourquoi quelques équipes se sont intéressées au pré-traitement d'intensités d'images IRM brutes. Ce pré-traitement consiste généralement en deux étapes : (1) la correction du biais lié à l'inhomogénéité du champ magnétique et (2) la standardisation des intensités. Pour l'étape (1), l'algorithme *open-source* N4ITK (TUSTISON et collab. [2010]) est consensuellement utilisé tandis que pour l'étape (2), différentes solutions rapportées ci-dessous ont été proposées.

CHRISTENSEN [2003] a proposé une méthode qui utilise les dérivées d'ordre pair de l'histogramme simulé d'intensités du cerveau pour estimer une valeur caractéristique de la matière blanche dans plusieurs modalités d'IRM. Cette valeur est utilisée pour normaliser les intensités. MADABHUSHI et UDUPA [2006] ont proposé deux méthodes d'apprentissage qui utilisent des paramètres dérivés de régions importantes du cerveau automatiquement repérées pour standardiser l'image de façon non-linéaire. La méthode de mise en correspondance d'histogramme proposée par NYŪL et collab. [2000] et affinée par SHAH et collab. [2011] pour la sclérose en plaques, utilise une transformation linéaire par morceaux, en déciles d'intensité plus précisément, pour trouver une correspondance entre l'histogramme d'intensité de chaque sujet et un histogramme de référence. Même si cette dernière méthode est largement utilisée pour la sclérose en plaques, la présence de tumeurs volumineuses peut être gênante et d'importantes variations entre différents sujets peuvent être supprimées.

La méthode *hybrid white stripe* (hWS) de SHINOHARA et collab. [2014] applique une transformation de type *z-score* sur tout le cerveau avec des paramètres provenant de la matière blanche d'apparence "normale". Les principaux avantages de cette méthode sont : (a) sa simplicité – pas besoin de segmentation précise de tissu, la réalisation d'une transformation linéaire, (b) sa robustesse aux anomalies présentées dans le cerveau, (c) la possibilité de combiner la multi-modalité pour augmenter sa précision et (d) la facilité d'ajout de nouveaux patients dans la base de données sans la nécessité de retraiter les images déjà présentes.

Plus récemment, FORTIN et collab. [2016] ont proposé une approche pour supprimer les variations d'intensité non-désirées issues des méthodes de standardisation comme le hWS en les modélisant à travers les sujets de l'étude à partir des incohérences d'intensités dans un tissu de référence, le LCR, le *Removal of Artificial Voxel Effect by Linear regression* (RAVEL). Même si dans cette dernière étude, RAVEL a amélioré la standardisation des intensités dans la MG après hWS, cette approche requiert des conditions assez fortes : (a) toutes les images doivent être recalées non-linéairement pour une analyse voxel-à-voxel sur tous les patients, (b) la méthode nécessite la segmentation de tissus du cerveau, en particulier le LCR, (c) il est impossible d'ajouter des données à la base sans devoir refaire la correction pour toutes les images.

3.2 La méthode de standardisation proposée

Dans cette thèse, j'ai choisi la méthode hWS pour standardiser les images en raison de sa simplicité et de sa pertinence dans le traitement des images issues des bases de données présentées dans la section 2.7. L'utilisation des quatre modalités d'IRM structurelle présentes dans les acquisitions augmente la précision de la méthode.

Le processus de standardisation d'intensités est présenté sur la figure 3.1. Les non-homogénéités du champ magnétique ont été corrigées par l'algorithme N4ITK (TUSTISON et collab. [2010]). Pour chaque patient, les différentes images ont été recalées par une transformation rigide aux

images de références de chaque patient avec FSL-FLIRT (JENKINSON et collab. [2002]). Les étapes pour le choix de l'image de référence et pour le recalage sont :

1. Pour chaque session, l'image de référence de la session est choisie comme étant l'image T2w si elle a une épaisseur de coupe inférieure à 2 mm et si le volume n'est pas tronqué. Sinon, l'image de la session avec la plus petite épaisseur de coupe est choisie comme référence.
2. La session de référence est choisie comme la session la plus ancienne dont l'image de référence a une épaisseur de coupe inférieure à 1,8 mm, si aucune session ne correspond à ce critère, la session la plus ancienne est choisie.
3. Pour chaque session, toutes les images sont recalées rigidement sur l'image de référence de la session.
4. Pour chaque patient, toutes les images de références des sessions sont recalées rigidement sur l'image de référence de la session de référence.
5. Les transformations sont appliquées en cascade résultant en toutes les images de chaque patient recalées sur l'image de référence de la session de référence, ce qui est appelé repère patient.

Les intensités de chaque modalité ont été standardisées à travers tous les patients et sessions d'imagerie avec le hWS, une méthode qui se base sur la distribution d'intensités de la matière blanche normale et ne nécessite pas de segmentation fine des tissus.

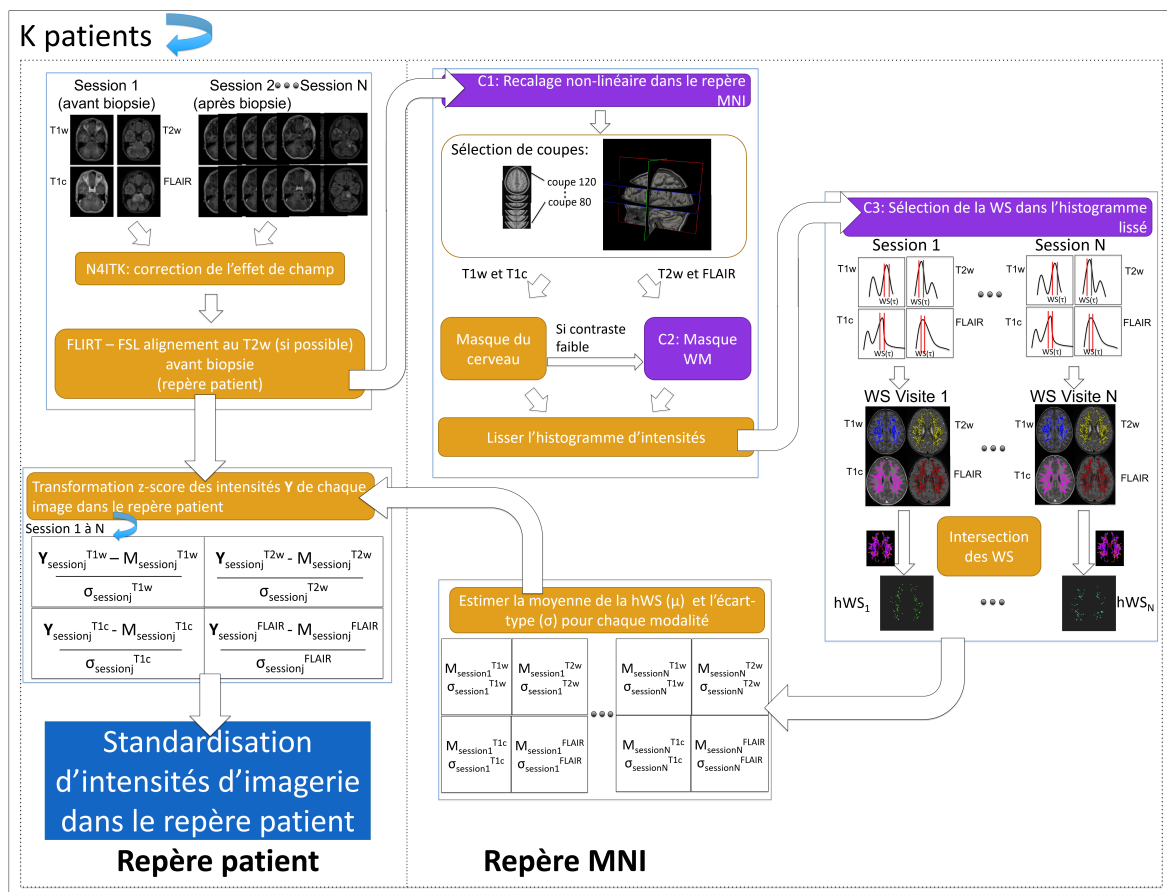


FIGURE 3.1 – Processus de standardisation d'intensité IRM. Comparé à la méthode hWS, les modifications introduites sont en violet et nommées changements C1, C2 et C3.

En utilisant le paquet du logiciel R 'WhiteStripe, Version 1.1.1', quelques modifications ont été introduites pour mieux prendre en considération les spécificités de nos données. Dans la méthode hWS, à partir de l'histogramme lissé par *penalized spline smoother* RUPPERT et collab. [2003] ($f_{ij}^m(x)$) des intensités (X_{ij}^m) des voxels compris dans les coupes 80 à 120 (V_{cpij}) d'une image de

référence dans le repère *Montreal Neurological Institute* (MNI), le mode $\mu_{i,j}^{m*}$ – pour le patient i , la session j , et la modalité m – est estimé comme étant le dernier pic de $f_{ij}^m(x)$ pour les images pondérées en T1 (T1w et T1c) et comme le pic le plus élevé pour les images pondérées en T2 (T2w et FLAIR). L'utilisation des coupes 80 à 120 est très bien adaptée pour les patients GTC puisqu'elles ne contiennent pas de tumeur, généralement. Cependant, dans notre étude, les patients sont des enfants et le recalage rigide ou affine dans le repère MNI, n'étant pas suffisant, a été remplacé par une transformation difféomorphique (AVANTS et collab. [2008]) avec le paquet du logiciel ANTs pour le langage de programmation R (ANTsR) (changement C1 sur la figure 3.1). En outre, dans certaines images T2w, comme montré dans la figure 3.2, le pic le plus haut de l'histogramme peut correspondre à la MG et non à la MB. Donc, le changement C2 de la figure 3.1 a été introduit. Un masque (W_{wm03}) a été défini à partir de la carte de probabilité d'appartenance à la MB dans le repère MNI disponible dans le logiciel SPM12 (ASHBURNER et FRISTON [2005]), en appliquant un seuil à une probabilité d'appartenance à la MB plus grande que 0,3. Cette contrainte garantit que le plus haut pic de l'histogramme correspond bien à la MB. Ce masque a été utilisé aussi pour les images T1w dès que certaines images montraient très peu de contraste entre la MB et la MG, résultant en un seul pic d'intensité pour les deux tissus. Il faut noter qu'aucune segmentation spécifique n'a été utilisée et la méthode est donc indépendante vis-à-vis d'une segmentation fine.

Pour chaque image la *white stripe* (WS) est définie comme étant l'ensemble des voxels dans le repère MNI ayant une intensité proche de $\mu_{i,j}^{m*}$. Un paramètre τ est utilisé pour définir cette proximité; il représente une valeur de tolérance autour de la valeur de la fonction de distribution cumulative de X_{ij}^m correspondant aux voxels Vcp_{ij} masqués par W_{wm03} (ou pas selon les critères définis plus haut) pour $x = \mu_{i,j}^{m*}$, plus précisément :

1. La valeur de la fonction de distribution cumulative en $\mu_{i,j}^{m*}$ est définie par $F_{ij}^m(\mu_{i,j}^{m*})$:

$$F_{ij}^m(\mu_{i,j}^{m*}) = \int_{-\inf}^{\mu_{i,j}^{m*}} f_{ij}^m(x) dx \quad (3.1)$$

2. Une plage de valeurs, R_{ij}^m , autour de $F_{ij}^m(\mu_{i,j}^{m*})$ est définie en fonction du paramètre τ :

$$R_{ij}^m = [Rmin_{ij}^m; Rmax_{ij}^m] = [F_{ij}^m(\mu_{i,j}^{m*}) - \tau; F_{ij}^m(\mu_{i,j}^{m*}) + \tau] \quad (3.2)$$

3. Les intensités correspondantes, I_{ij}^m sont définies :

$$I_{ij}^m = [F_{ij}^{m-1}(Rmin_{ij}^m); F_{ij}^{m-1}(Rmax_{ij}^m)] \quad (3.3)$$

4. Finalement, tous les voxels ($v_{i,j}^m$) de l'image de modalité m , du patient i , session j qui présentent une valeur d'intensité y dans l'intervalle I_{ij}^m sont les voxels de la $WS_{i,j}^m$:

$$WS_{i,j}^m = \{v \in v_{i,j}^m | y(v) \in I_{ij}^m\} \quad (3.4)$$

Quand les images sont masquées par la MB, le changement C3 sur la figure 3.1 est nécessaire. Le paramètre τ a été modifié pour sélectionner un nombre de voxels WS globalement similaire pour chaque modalité. La valeur par défaut de $\tau=0.05$ a été utilisée quand le masque de MB n'était pas nécessaire alors que la valeur $\tau=0.08$ a été utilisée sinon pour garder une cohérence dans le nombre de voxels retenus (figure 3.3). La hWS du patient i , session j a été définie comme l'intersection de WS de toutes les modalités IRM acquises dans la session j (M_j) (équation 3.5).

$$hWS_{i,j} = \bigcap_{m \in M_j} WS_{i,j}^m \quad (3.5)$$

Chaque image de la session j est enfin normalisée – dans le repère du sujet en question –, par la soustraction de la valeur moyenne des voxels de la hWS et la division par leur écart-type (équation 3.6).

$$Yst_{i,j}^m = \frac{Y_{i,j}^m - moyenne(y^m(hWS_{i,j}))}{ecart(y^m(hWS_{i,j}))} \quad (3.6)$$

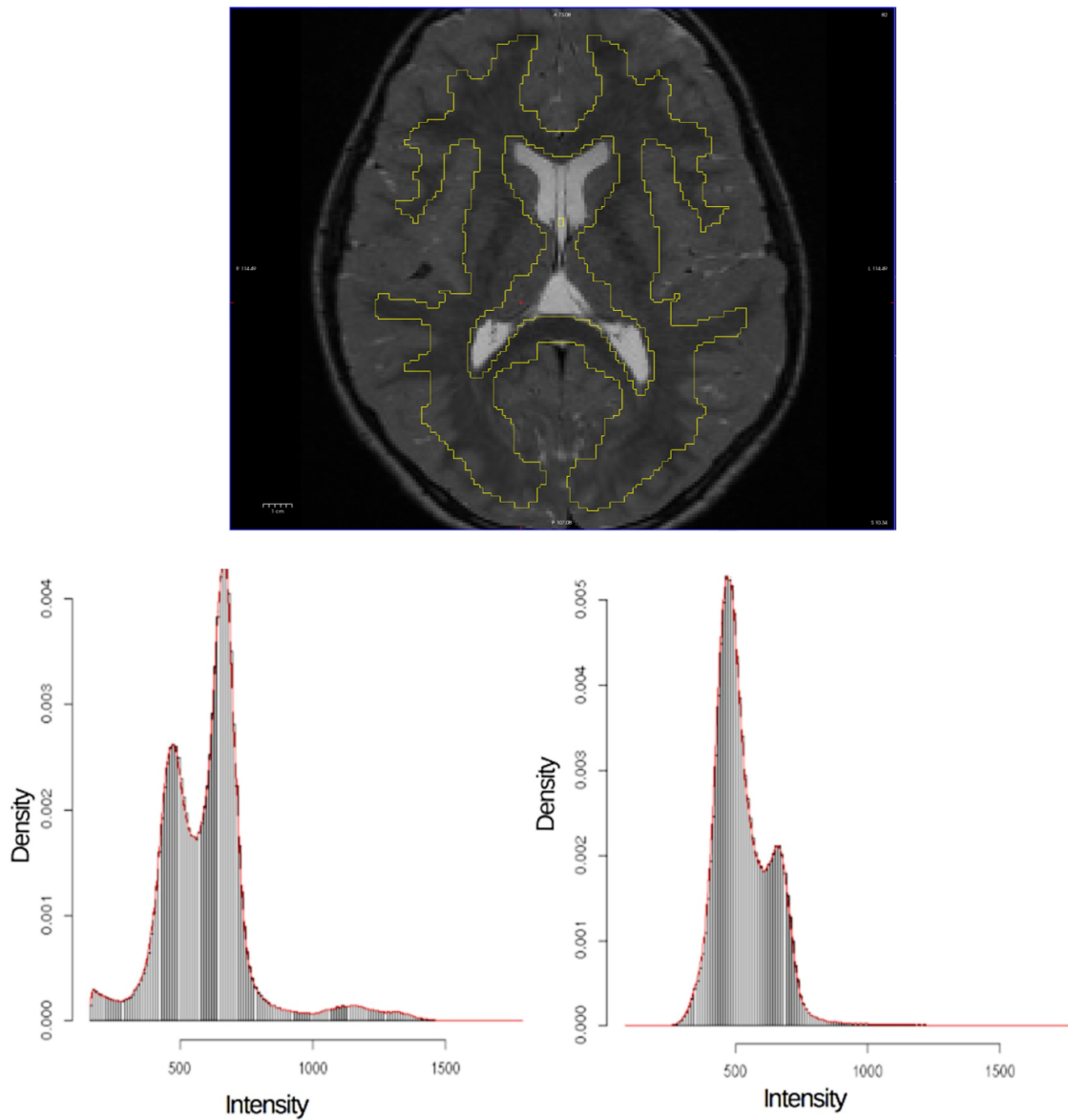


FIGURE 3.2 – Patient PREBIOMEDE sub-5272, première session : histogramme à 200 bins des coupes 80 à 120 (repère MNI) des intensités T2w sans (en bas à gauche) et avec (en bas à droite) l’application d’un masque de MB. En rouge, les courbes des histogrammes lissées utilisées par la méthode hWS pour trouver le pic. L’image de la coupe 82 (repère MNI) du patient correspondant avec le contour du masque de MB est montrée en haut.

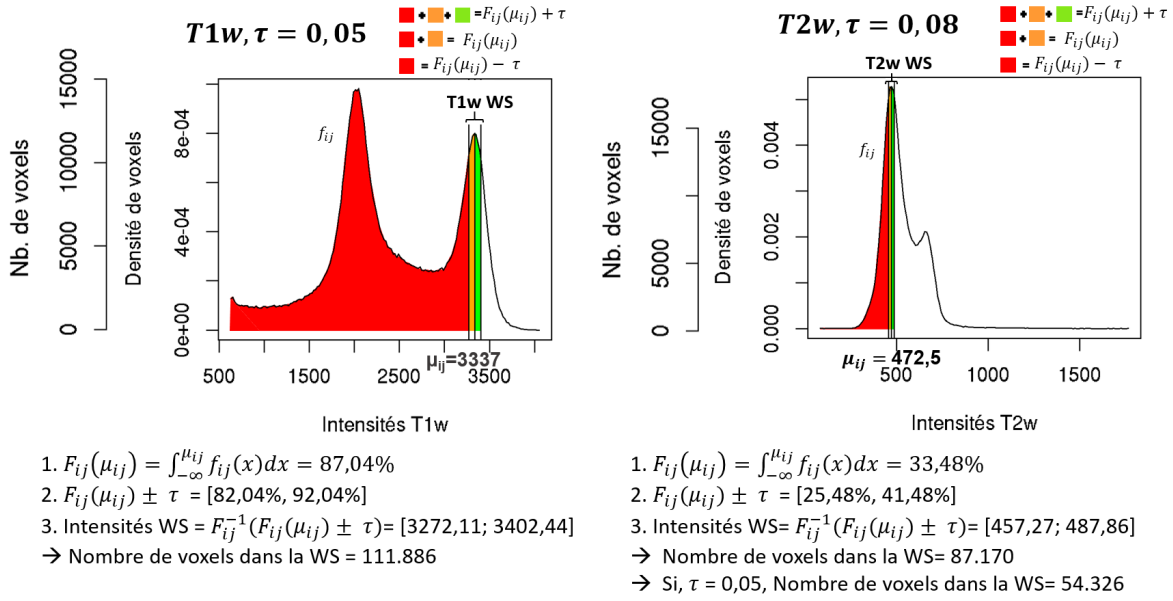


FIGURE 3.3 – Patient PREBIOMEDE sub-5272, première session : Distribution des intensités des voxels des coupes 80 à 120 dans le repère MNI : à gauche, pour l’image T1w après masquage du cerveau; à droite, pour l’image T2w après masquage de la matière blanche. Le choix de la valeur de τ est différent dans ces deux cas pour que le nombre de voxels dans les deux WS soit quasiment le même.

3.3 Définition d’un pipeline de contrôle de qualité

Dans cette étude les patients ont été inclus au fur et à mesure; donc il était important d’automatiser toute la procédure de normalisation en vue d’optimiser l’analyse à chaque nouvelle arrivée d’images. Pour cela il s’est avéré essentiel de créer un pipeline de **contrôle automatique de qualité (CQ)** pour vérifier si les images étaient bien standardisées.

Avec l’arrivée de données BIOMEDE et l’ajout des différentes visites des patients PREBIOMEDE, une procédure homogène de stockage de données a été conçue en partenariat avec Neurospin (figure 3.4). Les images DICOM pseudonymisées ont été anonymisées, triées (puisqu’il y avait plusieurs séquences d’images T1w, T2w correspondant à différents essais à des images acquises suivant différentes orientations et à des images à épaisseur de coupe variable) et organisées selon les recommandations *brain imaging data structure (BIDS)*. Les images ont été finalement publiées via un serveur *open-source* CubicWeb qui permet l’accès externe contrôlé (GRIGIS et colab. [2017]).

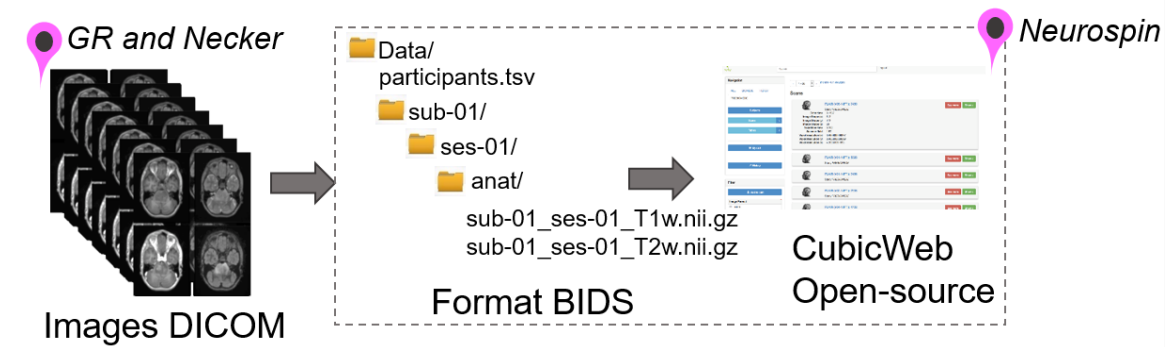


FIGURE 3.4 – Pipeline du stockage de données à Neurospin.

La première cohorte, la **cohorte étoile**, a été utilisée comme cohorte de référence, car pour ces données, toutes les étapes du pipeline de standardisation ont été validées visuellement et indi-

viduellement. Le nombre de voxels des **WS** était autour de 100.000 pour chaque modalité ce qui revient à environ 2.000 voxels dans la **hWS** quand les quatre modalités sont analysées. Pour la visualisation des résultats, des masques de **MB** et **MG** ont été définis avec la carte de probabilité d'appartenance à la **MB** et à la **MG** dans le repère **MNI** disponible dans le logiciel SPM12, en appliquant un seuil de probabilité d'appartenance plus grande que 0,9 pour la **MB** et 0,75 pour la **MG**. De plus, les coupes 31 à 51 dans le repère **MNI** ont été exclues pour éviter le tronc cérébral et donc la tumeur. Après standardisation, les histogrammes d'intensité de la **MB** devraient être centrés à 0, avec un écart-type assez similaire entre les patients. En revanche, les histogrammes d'intensité de la matière grise devraient être centrés à une valeur éloignée de 0 et variables suivant les différentes modalités. Chaque image a été ensuite vérifiée pour garantir que le recalage était correct et que les résultats suggérés par les histogrammes d'intensité correspondaient bien à ceux d'une image standardisée.

Ensuite, pour chaque nouvelle session acquise, des critères de **CQ** ont été établis. La **distance de Hellinger (DH)** **NIKULIN [2001]** a été utilisée dans le contrôle de qualité avec le paquet R 'statip'. Cette distance, introduite par Ernst Hellinger en 1909, est utilisée pour quantifier la similarité entre deux distributions continues P et Q. Elle varie entre 0 et 1. La valeur est 1 quand P est à 0 pour tous les valeurs où Q n'est pas à 0 ou vice-versa. Le carré de la **DH** entre deux distributions P et Q est défini comme indiqué par l'équation 3.7 :

$$DH^2(P, Q) = \frac{1}{2} \int (\sqrt{dP} - \sqrt{dQ})^2 \quad (3.7)$$

Les critères de qualité ont été définis comme suit :

1. la **DH** entre la distribution d'intensité de la **MB** du nouveau patient (fMB_N^m) et la distribution moyenne d'intensités de la **MB** de tous les patients de la **cohorte étoile** (fMB_{moy}^{m*}) est-elle plus petite que la plus grande distance entre la distribution individuelle d'un patient i (fMB_i^{m*}) de la **cohorte étoile** à cette distribution moyenne? (Figure 3.5)

$$DH(fMB_N^m, fMB_{moy}^{m*}) < \max_{i \in [1;30]} DH(fMB_i^{m*}, fMB_{moy}^{m*}) \quad (3.8)$$

2. La distribution des intensités de la **MG** est-elle centrée assez éloignée de zéro? Nous avons considéré que la valeur d'intensité correspondant au pic de l'histogramme en **MG** doit être plus petite que -5 pour les images **T1w** et **T1c** et plus grande que 5 pour les images **T2w** et **FLAIR** (Figure 3.6). Quand l'image ne présente pas assez de contraste entre la **MG** et la **MB** ou lorsque des voxels de la **WS** contiennent de la **MG**, l'histogramme d'intensités de la **MG** est centrée sur zéro après standardisation.

Le choix des seuils -5 et 5 pour le critère de la **MG** est arbitraire; l'idée était de garantir que les voxels sélectionnés pour la **hWS** n'appartenaient pas à la **MG**.

Les images respectant les 2 critères de qualité sont jugées comme bien standardisées par le **CQ**.

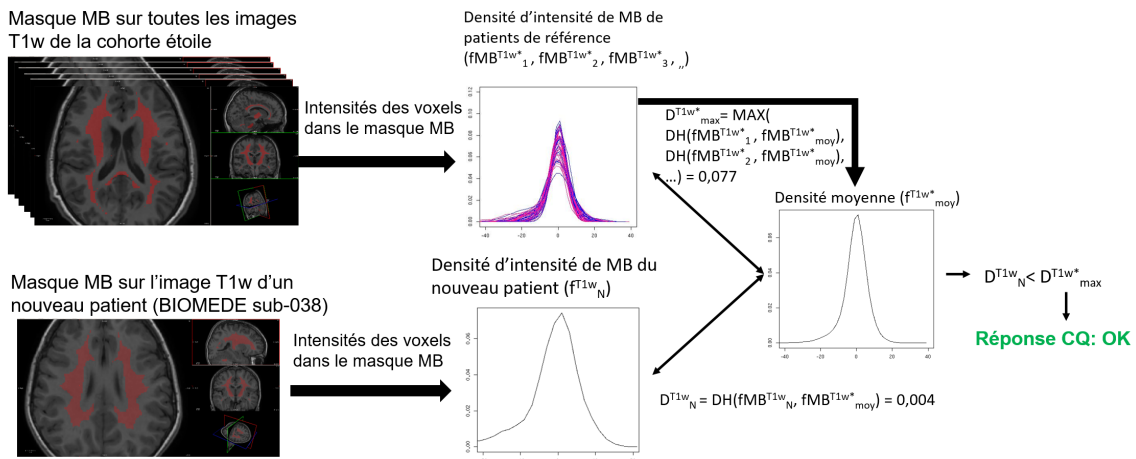


FIGURE 3.5 – Critère de qualité de la standardisation pour la MB. Exemple pour une image T1w d'un patient BIOMEDE (sub-038, première session).

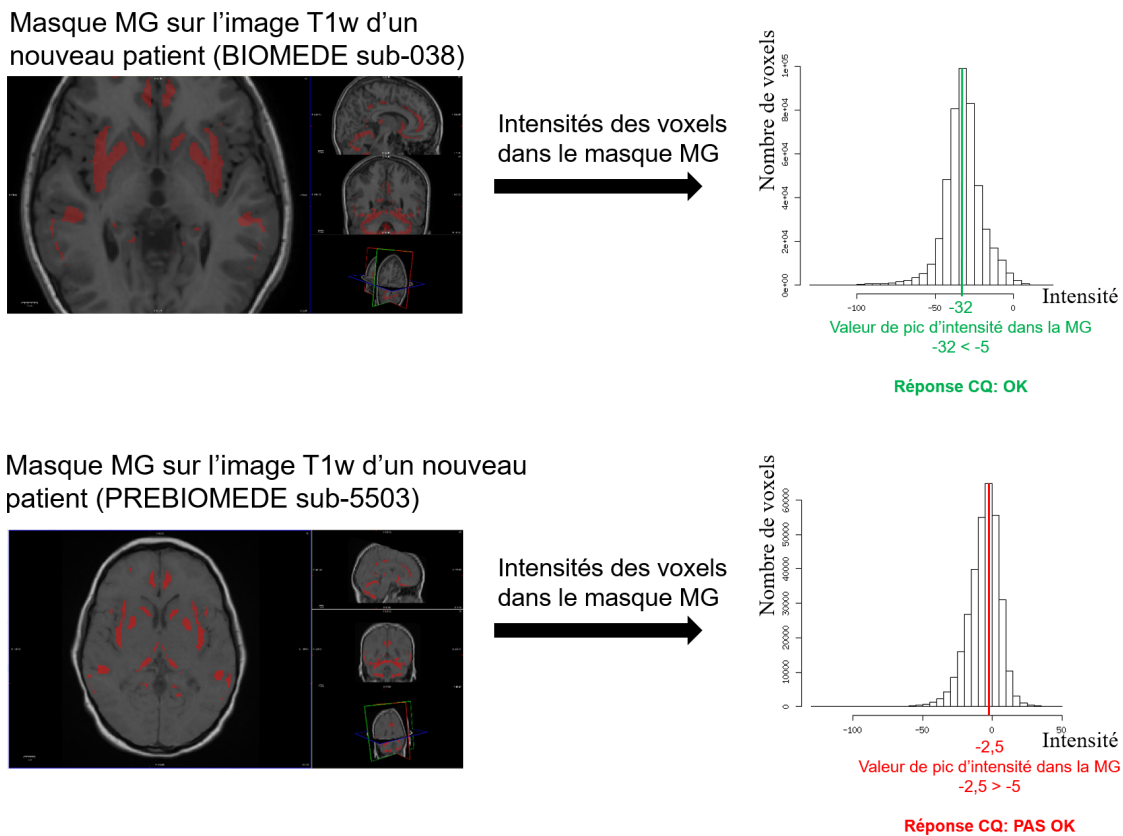


FIGURE 3.6 – Critère de qualité de la standardisation pour la MG. Exemple pour deux images T1w de deux patients : en haut, la première session d'imagerie du patient BIOMEDE sub-038 qui passe le contrôle de qualité et en bas, la première session d'imagerie du patient PREBIOMEDE sub-5503 qui ne passe pas le contrôle de qualité.

3.4 Application des méthodes de standardisation aux cohortes GITC

Dans un premier temps, les images de la **cohorte étoile** ont été soumises au pipeline de standardisation. Pour chaque session d'un patient donné, la **hWS** a été définie comme l'intersection des **WS** de toutes les modalités d'**IRM** présentes (entre une et quatre).

Les courbes de distribution d'intensités dans la **MB** et la **MG** avant et après standardisation sont présentées sur la figure 3.7 pour les quatre modalités. Les valeurs de la **DH** moyenne entre toute combinaison de paire de courbes de chaque modalité sont affichées sur le graphe correspondant. Pour toutes les modalités, après standardisation, la **DH** moyenne a diminué d'un facteur supérieur à 10 dans la **MB** et supérieur à 2,5 dans la **MG**. Les valeurs d'intensités de la **MG** correspondant au pic de l'histogramme pour les images de la **cohorte étoile** sont présentées sur la figure 3.8.

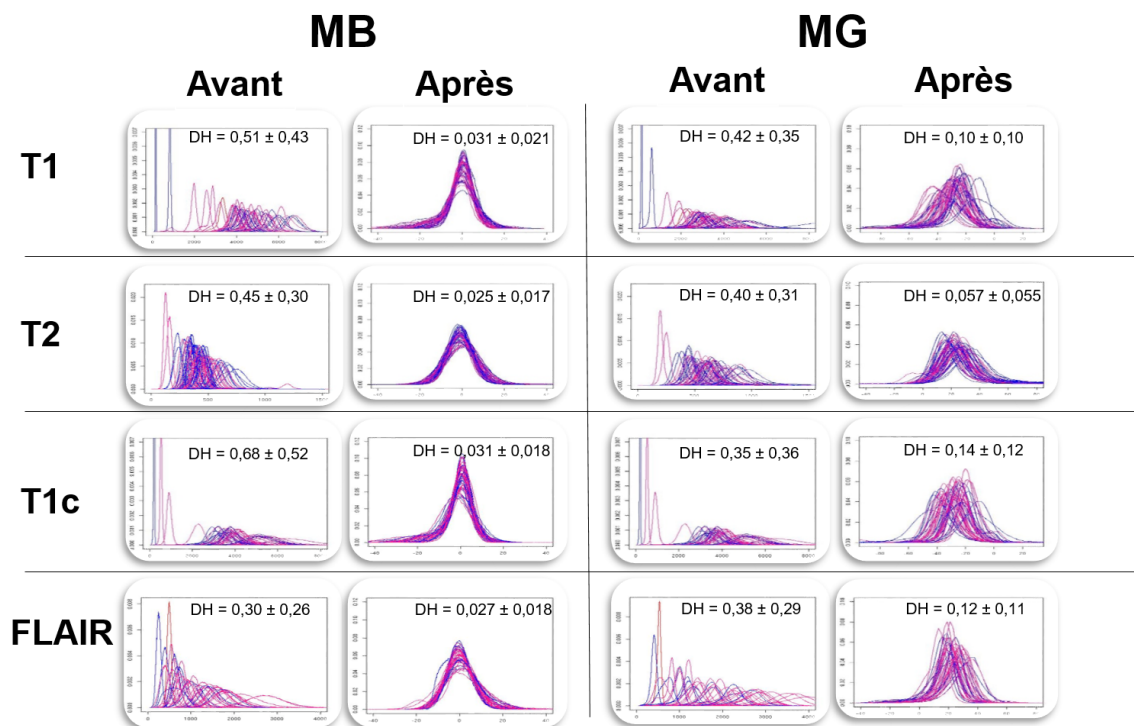


FIGURE 3.7 – Histogrammes d'intensités de la **MB** et de la **MG** avant et après standardisation pour les images **T1w**, **T2w**, **T1c** et **FLAIR** de la **cohorte étoile**. Les courbes bleues correspondent à des acquisitions avant biopsie et les courbes roses à des acquisitions après biopsie. Pour chaque modalité d'**IRM**, la **DH** moyenne des courbes deux-à-deux et l'écart-type associé sont affichés.

Les distances ($DH(fMB_i^m, fMB_{moy}^{m*})$) de chaque courbe de distributions d'intensités de la **cohorte étoile** à la distribution moyenne sont présentées sur la figure 3.9. Les images **PREBIOMEDE** (1012 images) et **BIOMEDE** (554 images) ont été soumises aux pipelines de standardisation et de **CQ**. En excluant 4 images **PREBIOMEDE** et 9 images **BIOMEDE** – qui étaient, en fait, des scanners du tronc ou des membres inférieurs qui n'avaient pas été correctement triés et bloquaient le processus – 90,6% des images **PREBIOMEDE** et 84,8% des images **BIOMEDE** ont été jugées correctement standardisées selon les critères **CQ**. Les histogrammes de distribution d'intensité de ces images sont présentés sur la figure 3.10 pour **PREBIOMEDE** et sur la figure 3.11 pour **BIOMEDE**. La figure 3.12 montre la répartition des distances ($DH(hist_N, hist_{moy}^*)$) pour chacun des 1008 volumes d'images **PREBIOMEDE** à la distribution moyenne de référence. La figure 3.13 présente les distances de chacun des 545 volumes d'images **BIOMEDE** à la distribution moyenne de référence. Les valeurs de pic de l'histogramme d'intensités de la **MG** sont présentées sur la figure 3.14 pour **PREBIOMEDE** et sur la figure 3.15 pour **BIOMEDE**. Les images de la **cohorte étoile** sont parmi les 1008 images **PREBIOMEDE** testées car nous avons vérifié que le stockage d'images à Neurospin

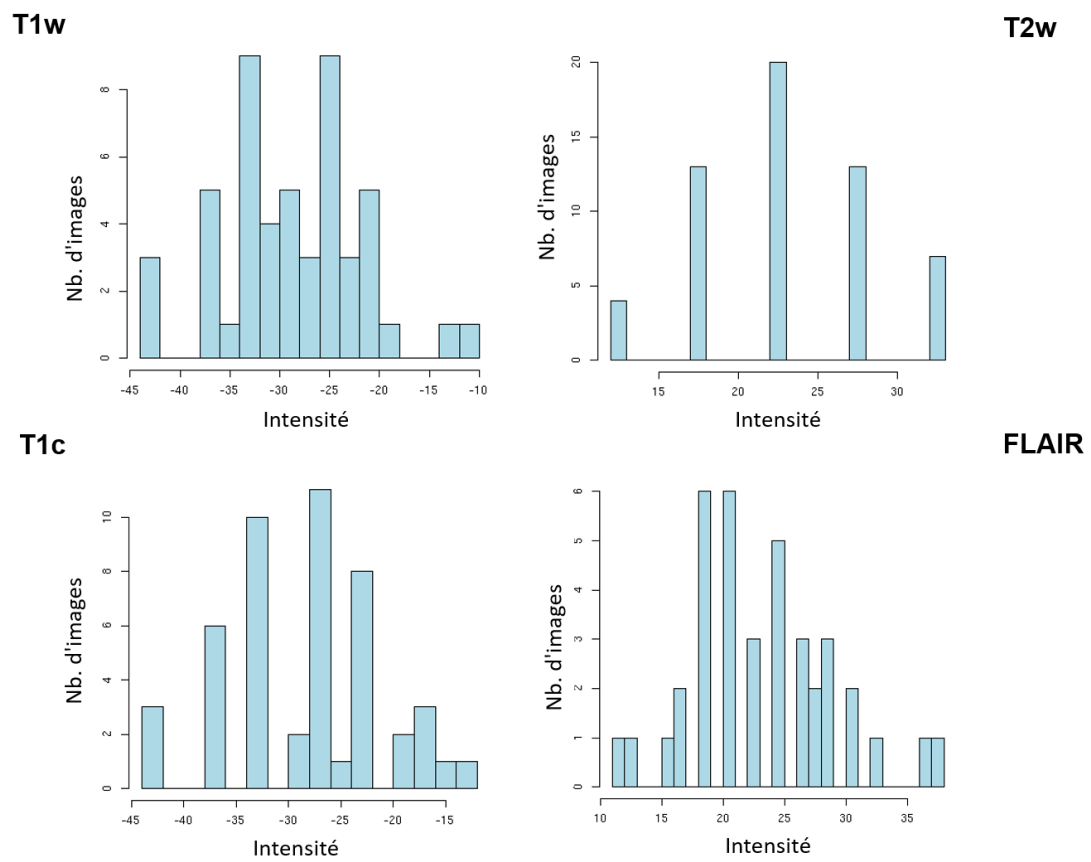


FIGURE 3.8 – Pic de l’histogramme d’intensités après standardisation de la MG des images T1w, T2w, T1c et FLAIR de la cohorte étoile.

avec les conditions d’anonymisation et de mise en forme ne changeaient pas drastiquement les résultats du pipeline .

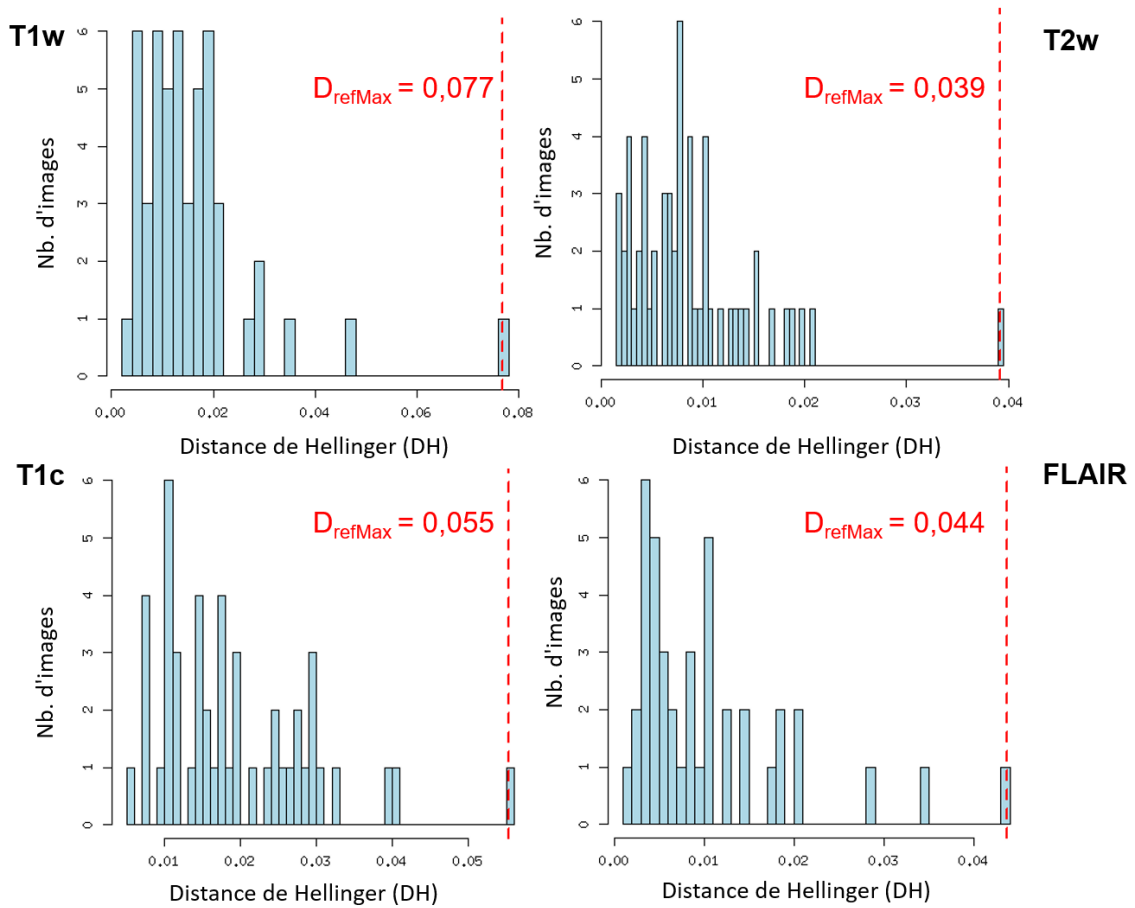


FIGURE 3.9 – Distance de Hellinger des distributions d’intensités de la MB des images de la cohorte étoile à la distribution moyenne de référence pour les images T1w, T2w, T1c et FLAIR. La distance maximale, utilisée comme seuil pour le critère de contrôle de qualité est affichée en rouge.

Pour résumer, le tableau 3.1 montre le nombre d’images de la cohorte PREBIOMEDE et le tableau 3.2 montre le nombre d’images de la cohorte BIOMEDE qui n’ont pas passé le CQ par modalité.

TABLEAU 3.1 – Nombre d’images PREBIOMEDE qui n’ont pas passé le contrôle qualité par modalité et par critère du contrôle de qualité non-respecté.

	Critère MB seul	Critère MG seul	Critères MB et MG	Total
T1w	6	5	3	14
T2w	24	1	8	33
T1c	14	0	4	18
FLAIR	16	3	11	30
Total	60	9	26	95

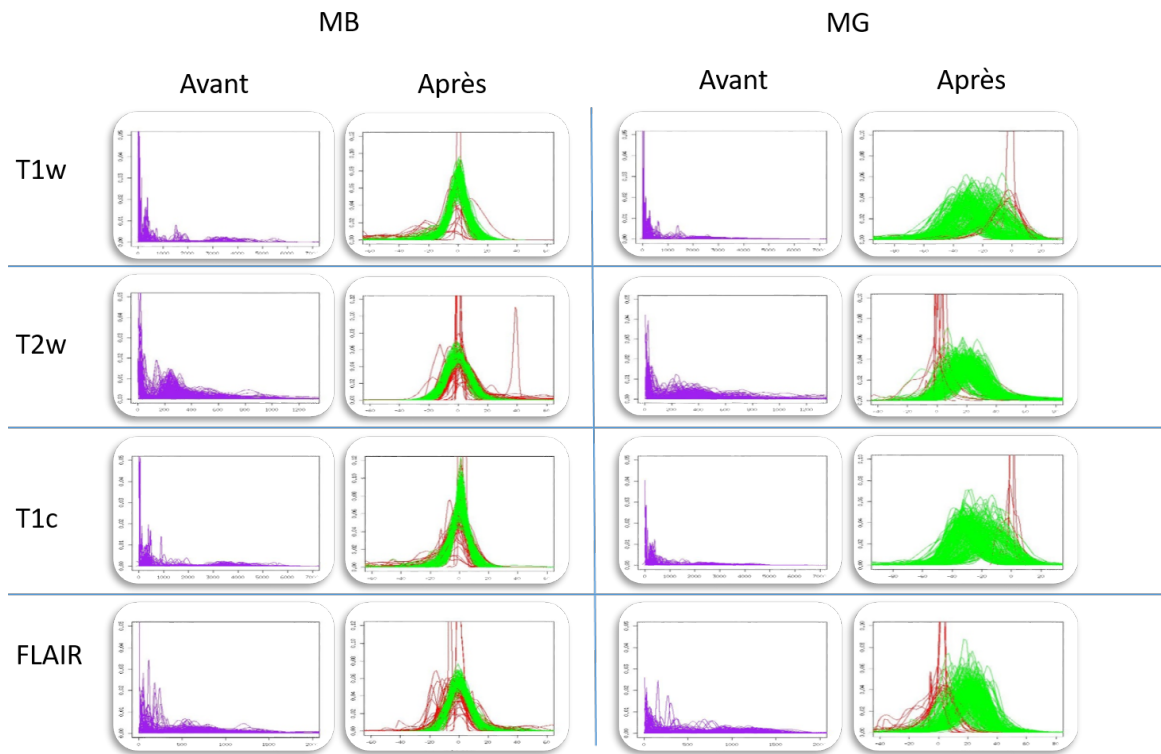


FIGURE 3.10 – Histogrammes d’intensités de la MB et de la MG avant et après standardisation pour les images T1w, T2w, T1c et FLAIR de la cohorte PREBIOMEDE. Chaque courbe verte correspond à une image qui a passé le contrôle automatique de qualité et chaque courbe rouge à une image qui n’a pas passé le contrôle de qualité.

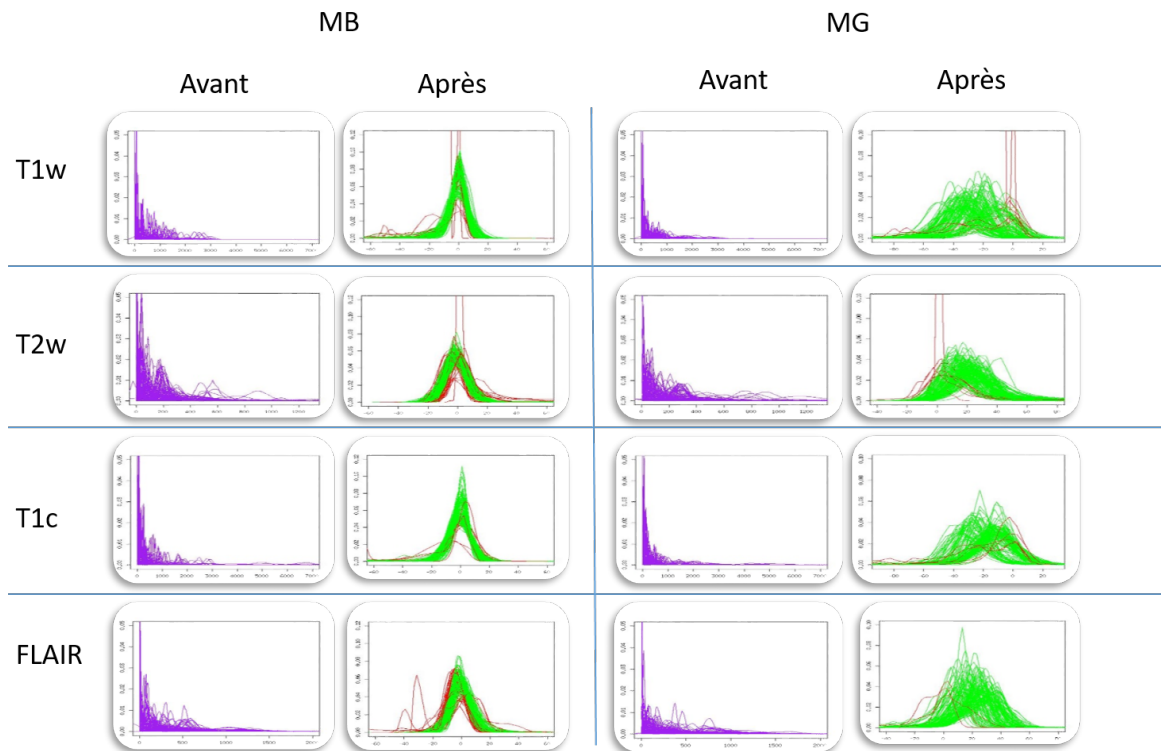


FIGURE 3.11 – Histogrammes d’intensités de la MB et de la MG avant et après standardisation pour les images T1w, T2w, T1c et FLAIR de la cohorte BIOMEDE. Chaque courbe verte correspond à une image qui a passé le contrôle de qualité et chaque courbe rouge à une image qui n’a pas passé le contrôle automatique de qualité.

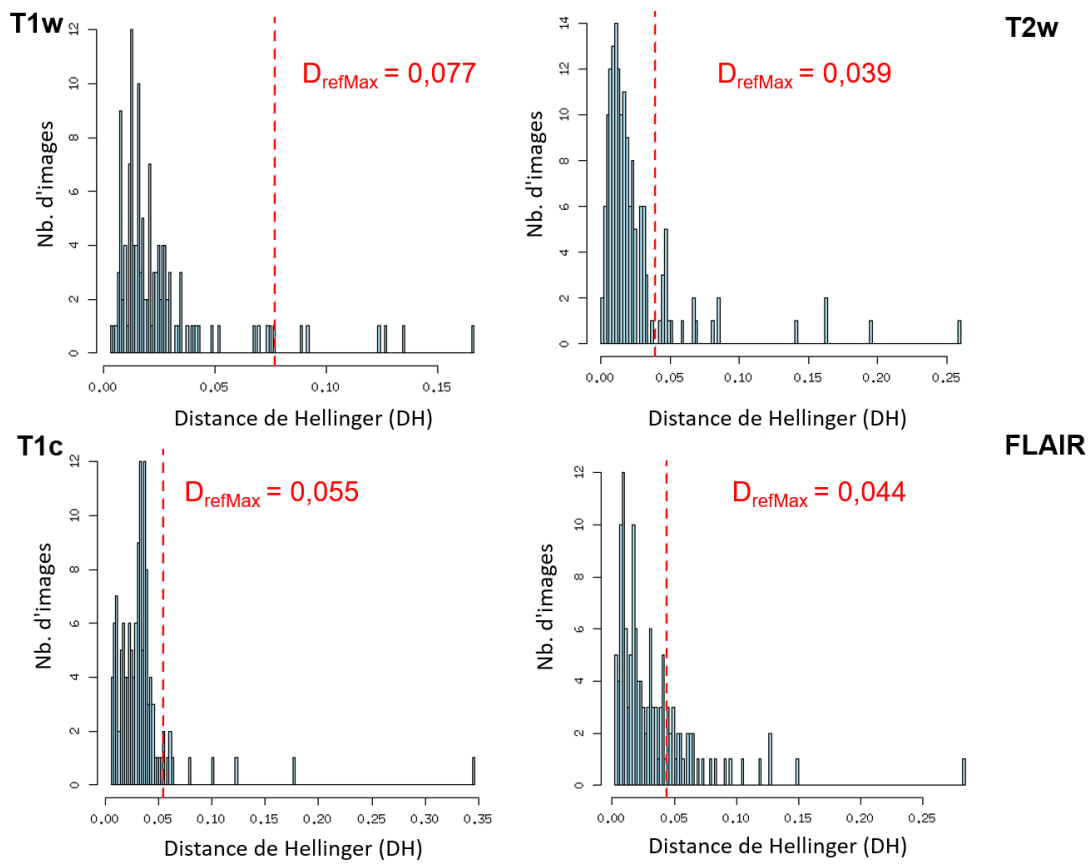


FIGURE 3.12 – Distribution des distances de Hellinger de patients PREBIOMEDE pour les intensités de la MB des images T1w, T2w, T1c et FLAIR. Les images qui présentent une DH plus grande que le seuil affiché en rouge ne passent pas le contrôle de qualité.

TABLEAU 3.2 – Nombre d'images BIOMEDE qui n'ont pas passé le contrôle qualité par modalité et par critère du contrôle de qualité non-respecté.

	Critère MB seul	Critère MG seul	Critères MB et MG	Total
T1w	5	4	3	12
T2w	18	2	5	25
T1c	8	4	1	13
FLAIR	30	1	2	33
Total	61	11	11	83

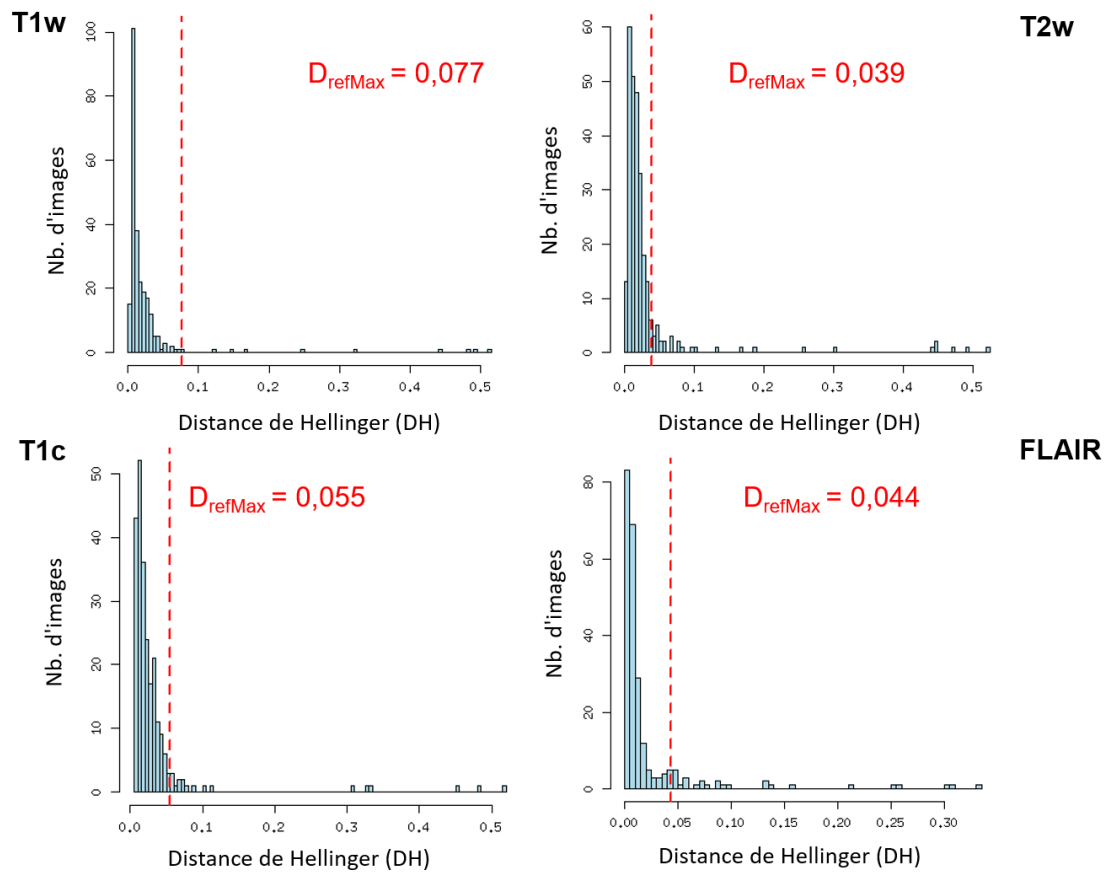


FIGURE 3.13 – Distribution des distances de Hellinger de patients BIOMEDE pour les intensités de la MB des images T1w, T2w, T1c et FLAIR. Les images qui présentent une DH plus grande que le seuil affiché en rouge ne passent pas le contrôle de qualité.

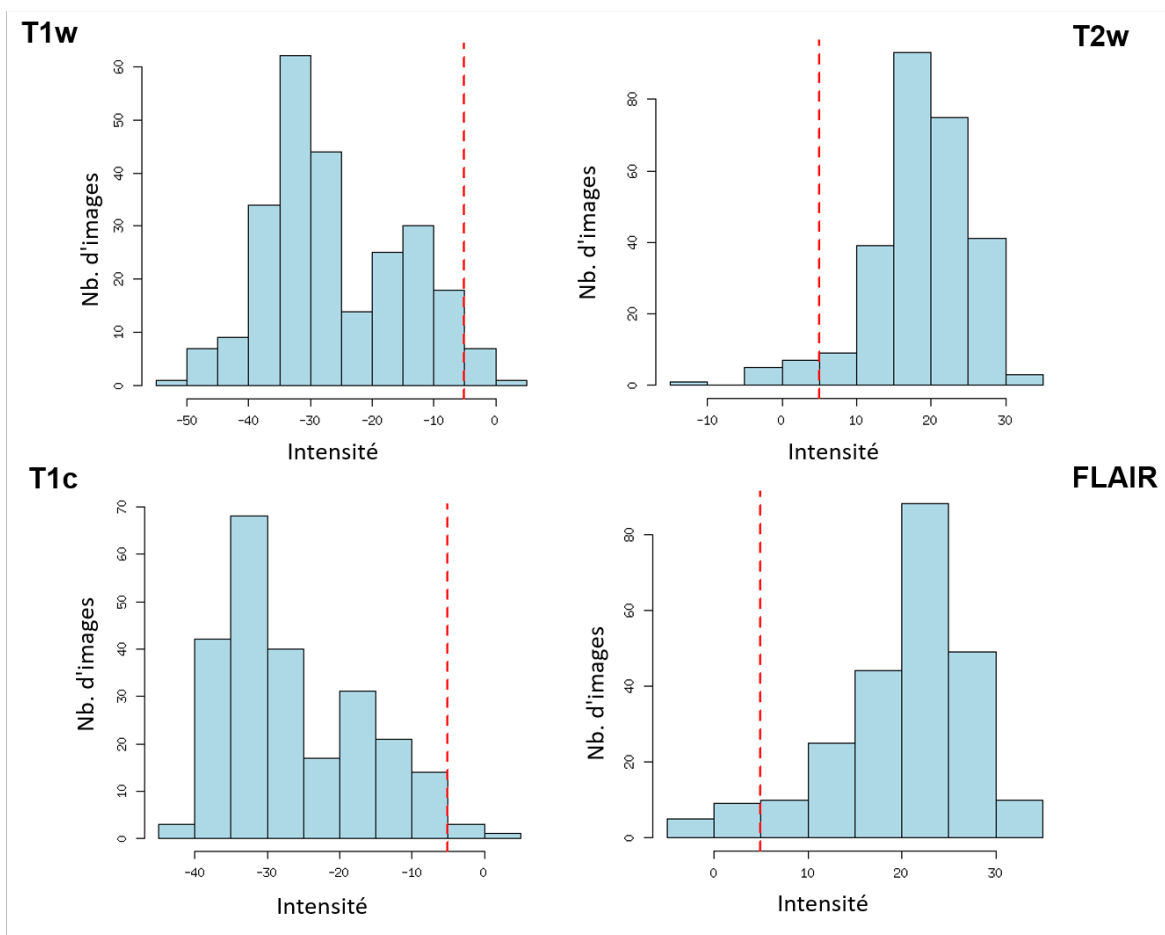


FIGURE 3.14 – Pic de l’histogramme d’intensités de la MG de patients PREBIOMEDE pour les images T1w, T2w, T1c et FLAIR. Les images qui présentent le pic des intensités en MG entre -5 et 5 (affichés en rouge) ne passent pas le contrôle de qualité.

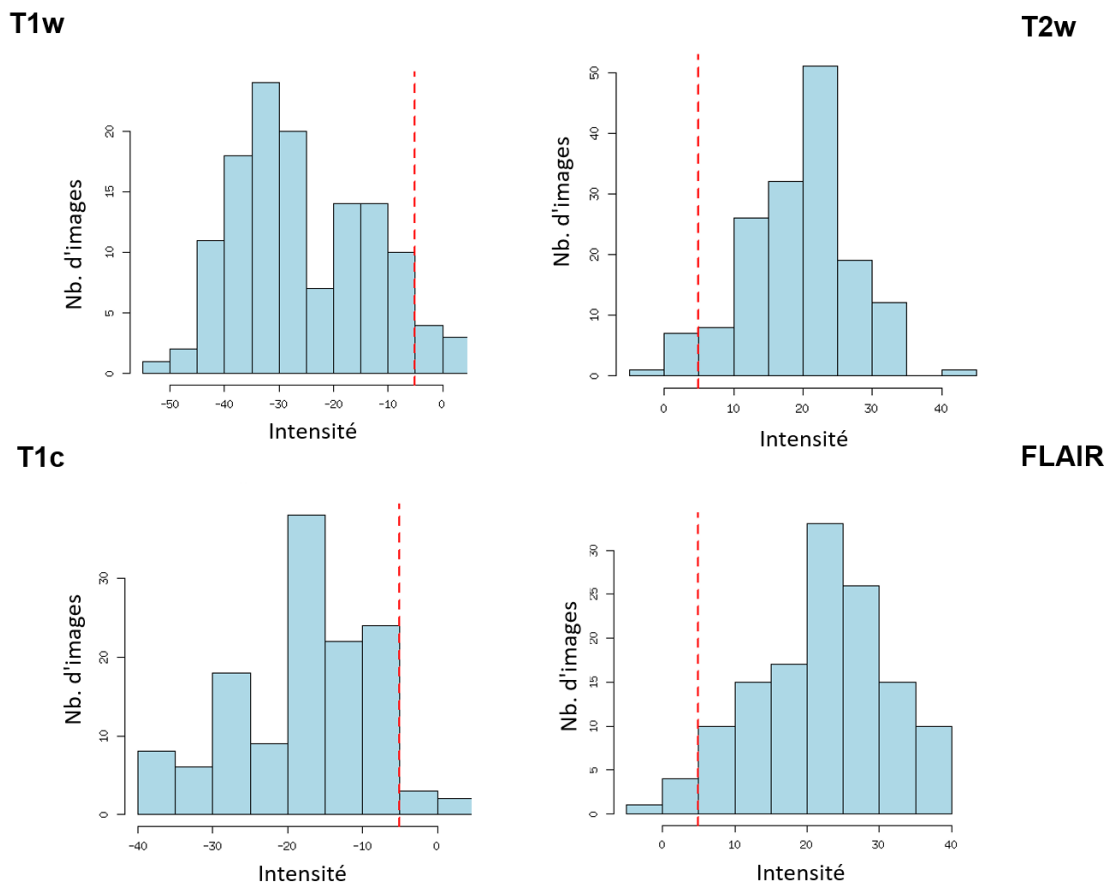


FIGURE 3.15 – Pic de l’histogramme d’intensités de la **MG** de patients BIOMEDE pour les images **T1w**, **T2w**, **T1c** et **FLAIR**. Les images qui présentent le pic des intensités en **MG** entre -5 et 5 (affichés en rouge) ne passent pas le contrôle de qualité.

3.5 Application des méthodes de standardisation sur une cohorte indépendante présentant des lésions cérébrales

Dans cette base de lésion cérébrales (LC cf. description dans le tableau A.3 en annexe), 20 patients avec des lésions cérébrales ont été imagés dans une IRM 1,5T puis dans une IRM 3T dans un intervalle de temps de 2 à 80 jours sans avoir eu de traitement, de chirurgie et sans évolution visible des lésions. Chaque patient avait de une à trois lésions segmentées par deux radiologues (A et L). Parmi les 20 patients, 19 ont des images FLAIR avec les deux machines (1,5T et 3T) et 15 ont des images T1c avec les deux machines, totalisant 68 images. Toutes les images d'un même patient ont été recalées sur l'image T1c à 1,5T. L'intérêt de cette base de données est de vérifier si la standardisation d'intensités reste valable pour des lésions dans différentes régions cérébrales et si les régions tumorales peuvent être comparées d'un examen à l'autre. Avec les bases de données BIOMEDE et PREBIOMEDE, il n'était pas possible de comparer les régions tumorales car les tumeurs évoluent possiblement d'une session à l'autre pour un même patient.

Pour le CQ portant sur la MB et la MG, les voxels correspondant à la tumeur ont été retirés avec le masque tracé par le radiologue A sur les images 1,5T. Parmi les 68 images, 15 n'ont pas passé le CQ. Huit d'entre elles avaient une DH à l'histogramme moyen de référence plus petite que 0,077 (la distance maximale en T1w de la cohorte étoile). La figure 3.16 montre les distances de chacune des 68 images de la base LC à la distribution moyenne de référence, les valeurs moyennes d'intensités de la MG pour ces images sont présentées sur la figure 3.17 et les histogrammes de distribution d'intensité sont présentés sur la figure 3.18. Finalement, la synthèse du nombre d'images qui ne passent pas les critères CQ est présentée dans le tableau 3.3.

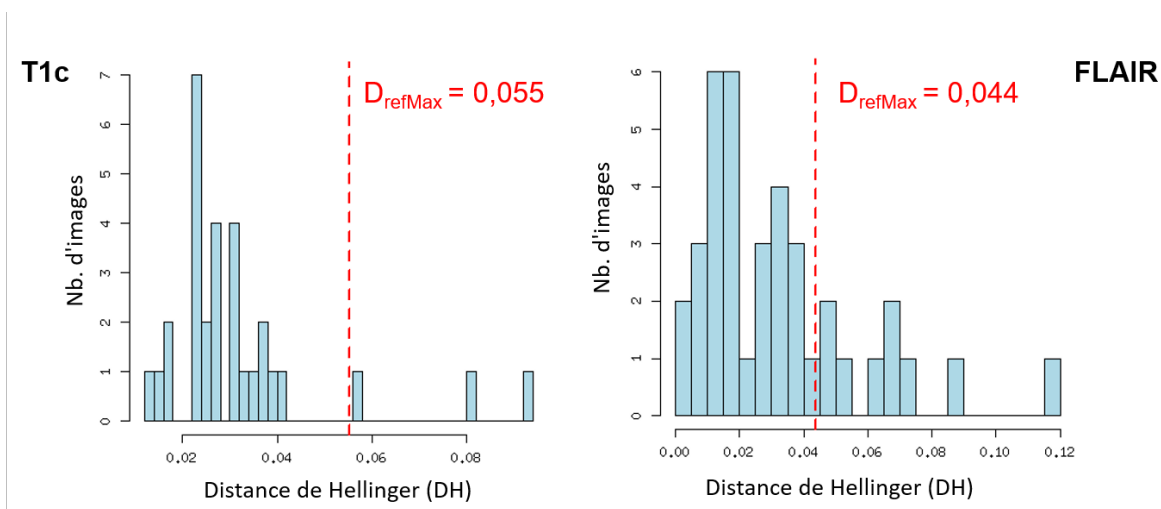


FIGURE 3.16 – Distance de Hellinger de patients de la base LC pour les intensités de la MB des images T1c et FLAIR. Les images qui présentent une DH plus grande que le seuil affiché en rouge ne passent pas le critère de qualité.

TABEAU 3.3 – Nombre d'images LC qui n'ont pas passé le contrôle qualité par modalité et par type de critère.

	Critère MB seul	Critère MG seul	Critères MB et MG	Total
T1c	2	2	1	5
FLAIR	9	1	0	10
Total	11	3	1	15

La figure 3.19 compare les valeurs d'intensité moyenne sur les deux examens réalisés à 1,5T et 3T dans les tumeurs avant et après standardisation d'intensité en regroupant les segmentations faites par les deux experts A et L. Les observations aberrantes en T1c sont identifiées sur la figure.

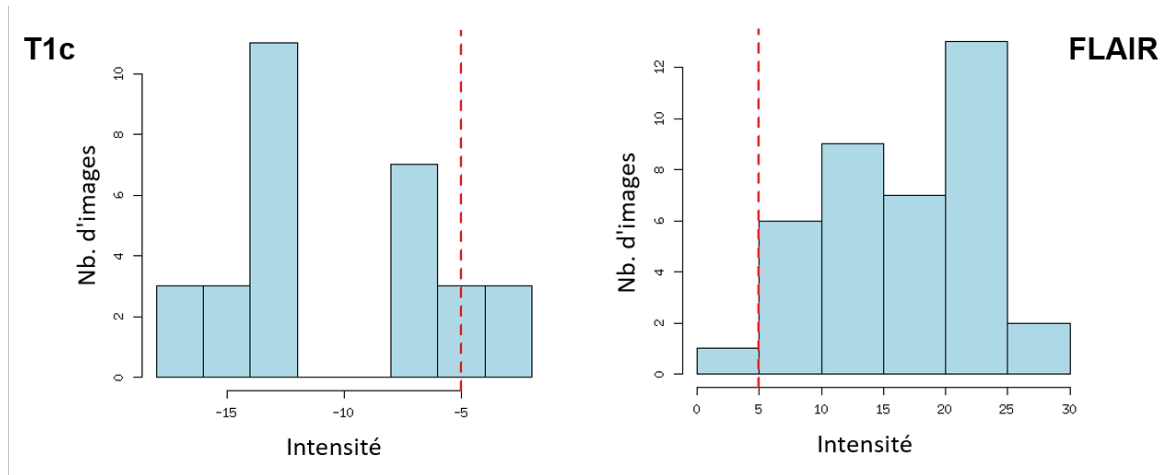


FIGURE 3.17 – Centre de l’histogramme d’intensités de la **MG** des patients de la base LC pour les images **T1c** et **FLAIR**. Les images qui présentent un pic d’intensité en **MG** qui ne respecte par les seuils affichés en rouge ne passent pas le critère de qualité.

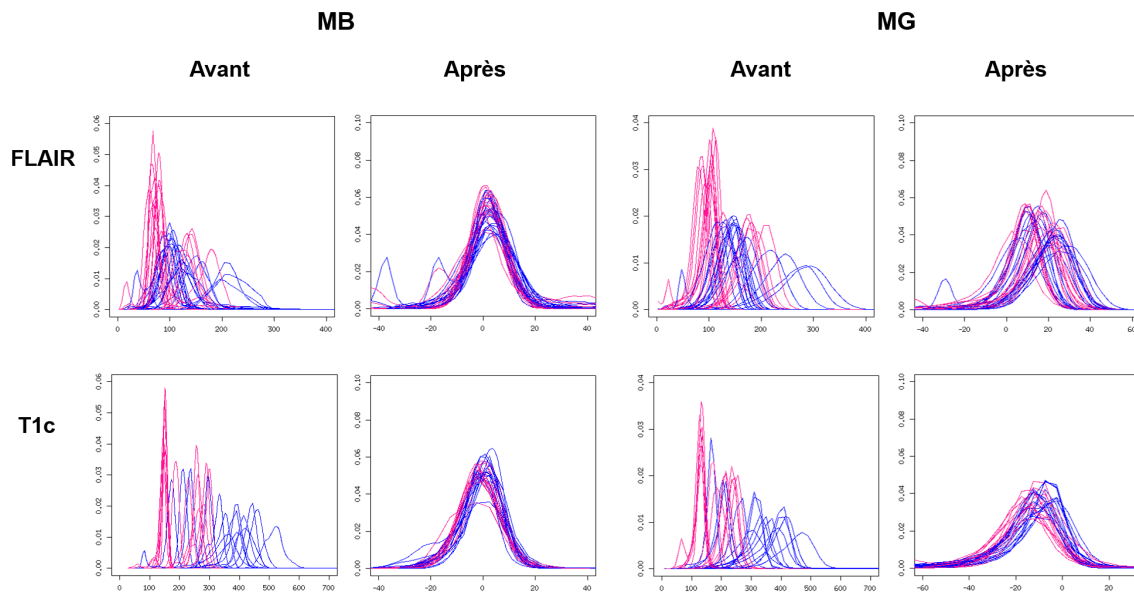


FIGURE 3.18 – Histogrammes d’intensités de la **MB** et de la **MG** avant et après standardisation pour les images **T1c** et **FLAIR** de la cohorte LC. Chaque courbe rose correspond à une image provenant de la machine 1,5T et chaque courbe bleu à une image provenant de la machine 3T.

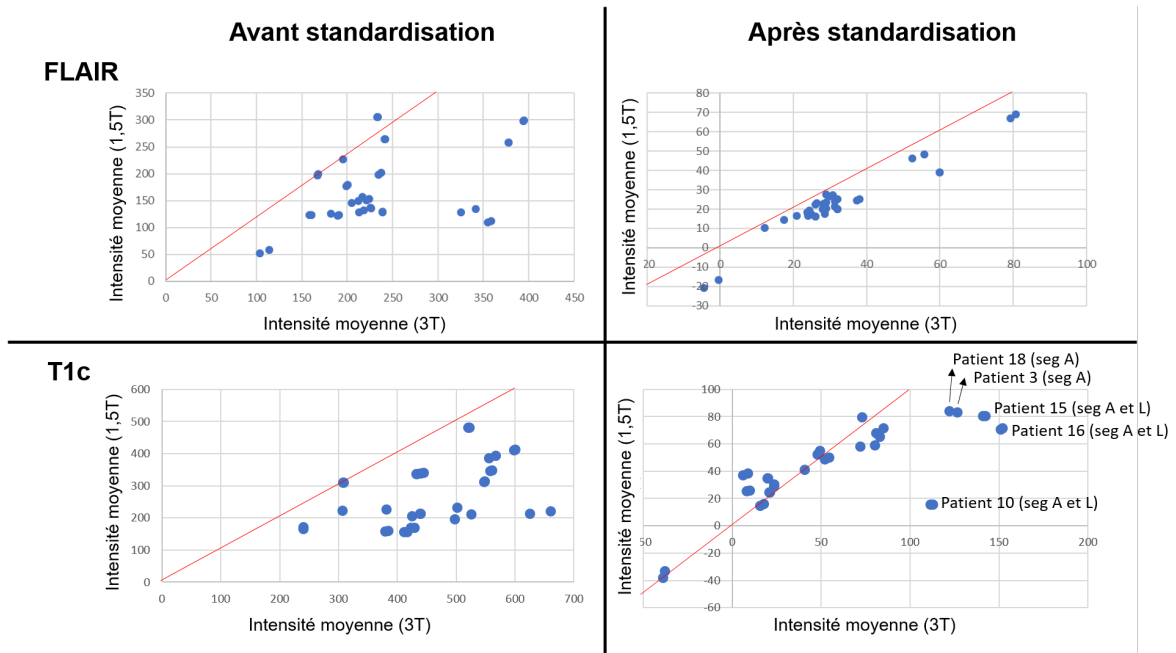


FIGURE 3.19 – Chaque point représente les valeurs moyennes d’intensité en 1,5T et 3T avant et après standardisation en **FLAIR** (première ligne) et **T1c** (deuxième ligne). La ligne rouge correspond à la droite identité. Les patients qui présentent des résultats aberrants après standardisation sont identifiés ainsi que la segmentation (A ou L) correspondant au point.

La distance moyenne de chaque point à la droite identité a été réduite de $57,17 \pm 41,29$ à $6,03 \pm 3,41$ pour **FLAIR** et de $131,65 \pm 74,13$ à $16,99 \pm 20,14$ pour **T1c**. La figure 3.20 présente les boîtes à moustaches des distances de points à la ligne d’identité avant et après standardisation par **hWS** pour les images **T1c** et **FLAIR**.

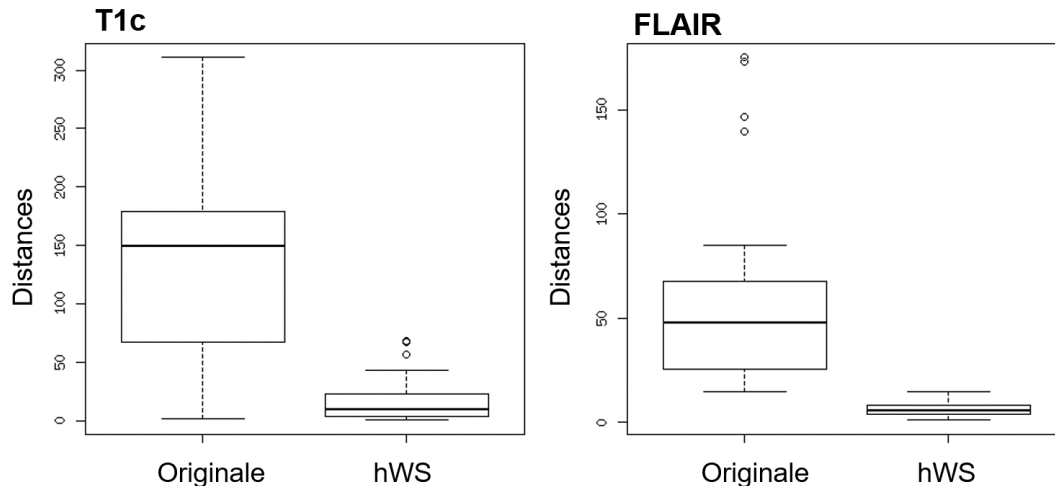


FIGURE 3.20 – Boîtes à moustaches des distances de chaque point (moyenne d'intensité sur l'examen à 3T versus moyenne d'intensité sur l'examen à 1,5T) à la ligne d'identité avant et après standardisation. A gauche pour les images T1c et à droite pour les images FLAIR.

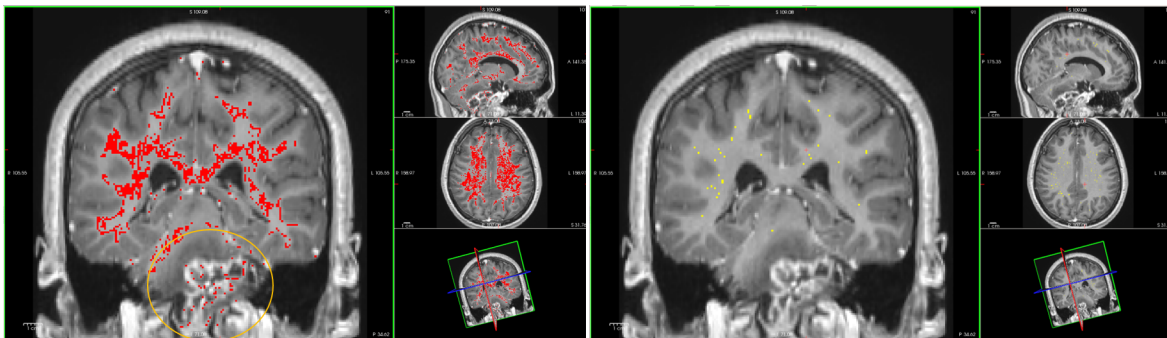
3.6 Discussion et conclusion

Les méthodes proposées dans ce chapitre ont été conçues pour surmonter les incohérences provenant des intensités arbitraires de l'IRM pour les tâches de comparaison quantitative d'imagerie intra et inter-patient. La méthode de standardisation d'intensité d'IRM, hWS que j'ai modifiée, ainsi que le CQ pour chaque nouvelle image normalisée ont été présentés et analysés.

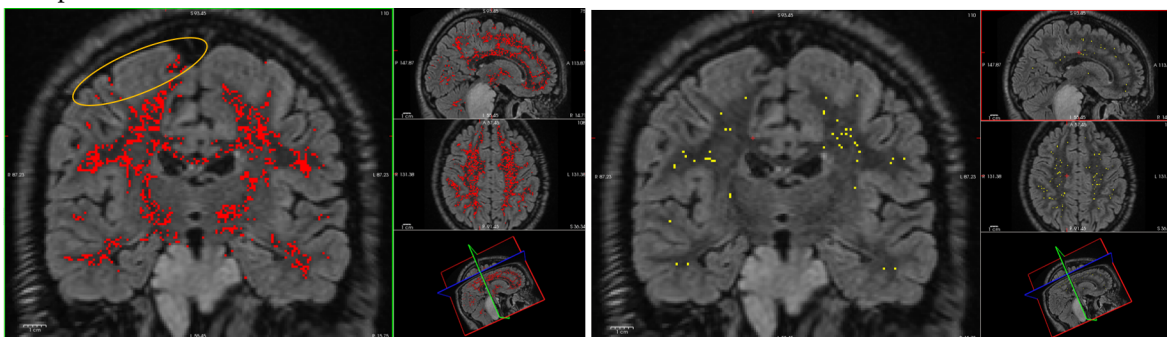
Cohorte Étoile : Les 196 premières images (cohorte étoile) correspondant aux sessions de diagnostic et de premier suivi post-biopsie de 30 patients de la cohorte PREBIOMEDE ont été normalisées avec succès. Avant standardisation, les histogrammes d'intensité de la MB et de la MG varient entre les patients, entre les sessions, même quand les images ont été acquises avec la même machine et des paramètres d'acquisition similaires. Après la standardisation hWS, les histogrammes d'intensité de la MB sont quasiment superposables, avec une réduction drastique de la DH moyenne entre les patients. Même si la réduction de la DH dans la MG a été moins importante, elle reste toujours réelle. Une cause possible à cette moindre réduction de la DH est le fait que la MG est plus compliquée à segmenter que la MB et qu'un nombre important de voxels qui ne correspondent pas uniquement à de la MG ont pu être utilisés pour la création des histogrammes de distribution d'intensités (Figure 3.7). La méthode proposée est intéressante quand la base de données est continuellement augmentée parce que chaque ensemble d'images d'une session d'un patient ne dépend que de lui-même pour être standardisé, différemment d'autres méthodes, comme la mise en correspondance d'histogrammes NYÛL et collab. [2000] et l'approche RAVEL FORTIN et collab. [2016] qui nécessitent l'utilisation de la base de données complète.

Un autre intérêt de la méthode est de s'appuyer sur la multi-modalité pour augmenter la robustesse des voxels sélectionnés grâce à l'approche hWS qui prend en compte pour chaque session d'imagerie toutes les modalités disponibles. Par exemple sur la figure 3.21a, quelques voxels très perfusés ont été sélectionnés avec WS dans l'image T1c (dans les cercle jaune) et ils sont éliminés avec l'intersection de différentes WS. Sur la figure 3.21b, une situation similaire avec la WS en FLAIR est présentée où quelques voxels de la MG sont sélectionnés puis éliminés dans la hWS. Néanmoins, les intensités moyennes et les écart-types sont assez similaires dans les WS de chaque modalité et dans la hWS (Tableau 3.4), ce qui montre que la standardisation d'intensités avec une WS uni-modalité est finalement quasi similaire à celle obtenue par la hWS. Cela indique aussi que les sessions des patients qui ont été imagés par une seule modalité d'IRM structurelle ne posent

pas de problème pour la standardisation d'intensités. Cependant, l'utilisation de la **hWS** reste préférable, car elle augmente la robustesse de voxels sélectionnés comme appartenant à la **MB** d'apparence normale.



(a) Patient PREBIOMEDE sub-7931, session de diagnostic : les voxels de la **WS** de T1c en rouge (à gauche) et les voxels de la **hWS** en jaune (à droite) sur l'image T1c. Des voxels dans la tumeur (dans le cercle jaune) sont présents dans la **WS** de T1c et éliminés de la **hWS**.



(b) Patient PREBIOMEDE sub-8044, session de diagnostic : les voxels de la **WS** de FLAIR en rouge (à gauche) et les voxels de la **hWS** en jaune (à droite) sur l'image FLAIR. Des voxels dans la **MG** (dans le cercle jaune) sont présents dans la **WS** de FLAIR et éliminés de la **hWS**.

FIGURE 3.21 – Comparaison des voxels dans une **WS** uni-modalité et des voxels inclus dans la **hWS**.

TABLEAU 3.4 – Valeurs des pics des intensités (écart-type) obtenus pour chaque **WS** et pour la **hWS** pour les patients présentés sur la figure 3.21. Le nombre de voxels (Nb. de vx) de chaque **WS** est également indiqué.

	T1w		T1c		T2w		FLAIR	
	WS	hWS	WS	hWS	WS	hWS	WS	hWS
sub-7931	6307 (71)	6316 (71)	9193 (105)	9193 (103)	346 (6)	346 (6)	2793 (68)	2793 (68)
Nb. de vx	109.681	2.468	11.969	2.468	84.222	2.468	88.913	2.468
sub-8044	4748 (59)	4757(58)	6393 (74)	6390 (74)	465 (8)	464 (8)	1635 (51)	1636 (52)
Nb. de vx	104.605	3.637	100.162	3.637	87.800	3.637	84976	3.637

Cohortes PREBIOMEDE et BIOMEDE : Pour l'inclusion, le stockage, la normalisation et la vérification de nouvelles sessions d'imagerie, un pipeline automatique de contrôle de qualité a été développé. Le pipeline mis en place permet la standardisation automatique de chaque nouvelle image ajoutée et la validation de cette standardisation. Ce processus est de très grande importance pour ce projet dans le contexte des protocoles cliniques toujours en cours d'inclusion. En effet l'inclusion progressive de nouveaux sujets pour une maladie très rare permet d'augmenter la puissance statistique des modèles et la robustesse de résultats d'apprentissage. Les résultats obtenus avec les cohortes PREBIOMEDE et BIOMEDE sont prometteurs quant à la robustesse de la méthode de standardisation car 88,5% de 1554 images ont été standardisées correctement de façon automatique.

Les 11,5% (178 images) d'images qui n'ont pas été jugées comme correctement standardisées ont été vérifiées visuellement. Pour 35 images, la **DH** était assez proche du seuil pour la modalité correspondante et l'analyse visuelle ne justifiait pas un re-traitement ou une exclusion de ces images. Le seuil du critère de qualité de standardisation pourrait être adapté dans le futur pour réduire le nombre de faux positifs. Dans un premier temps, ils ont été définis à partir de 30 sujets.

Concernant les 143 autres volumes d'image, les sources d'erreur ont été :

- 40 images de faible qualité, qui incluent : des images bruitées comme illustré sur la figure 3.22, des volumes d'image tronquées, ou des images à faible résolution spatiale;
- 3 images où la non-uniformité du champ magnétique n'a pas été bien corrigée par l'algorithme N4ITK;
- 16 volumes d'images pour lesquels le recalage diffeomorphique vers le repère **MNI** n'a pas été satisfaisant.
- 29 volumes d'image pour lesquels le recalage rigide FSL-FLIRT n'a pas été satisfaisant.
- 11 volumes d'image avec des masses tumorales ou d'oedème dans la partie supérieure du cerveau.
- 34 volumes d'image pour lesquels le patient présentait des conditions anatomiques particulières comme l'élargissement des ventricules cérébraux (figure 3.24).

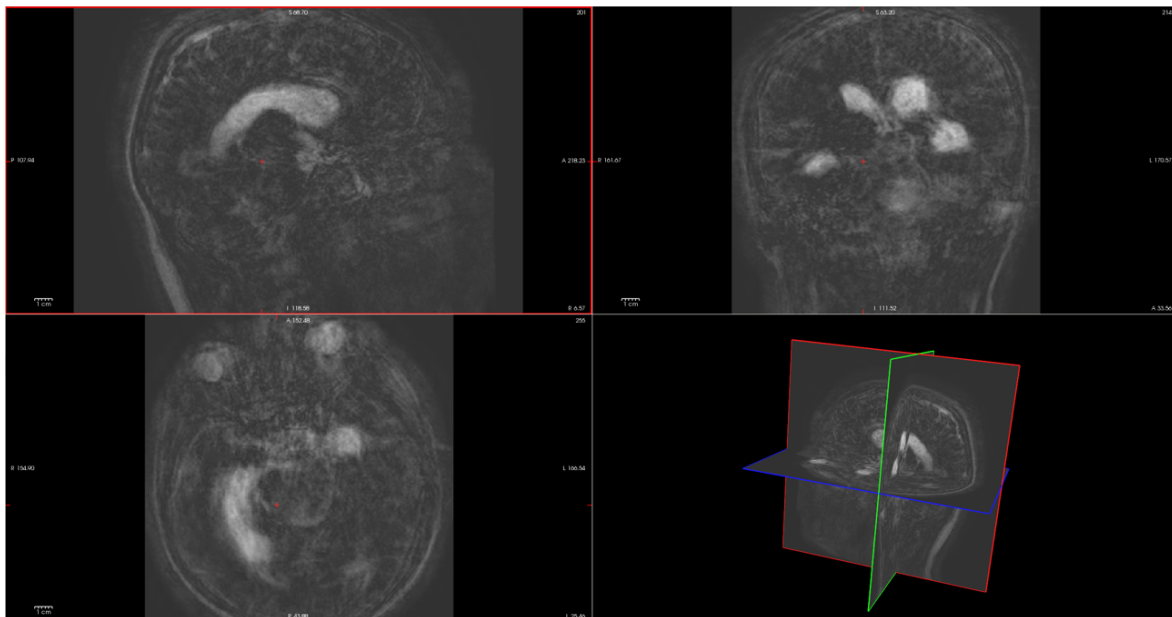


FIGURE 3.22 – Patient BIOMEDE sub-114, troisième session : Exemple d'image T2w très bruitée. La **DH** à l'histogramme d'intensité moyen de référence en **MB** est 0,16 et le pic de son histogramme d'intensités dans la **MG** vaut 2,5.

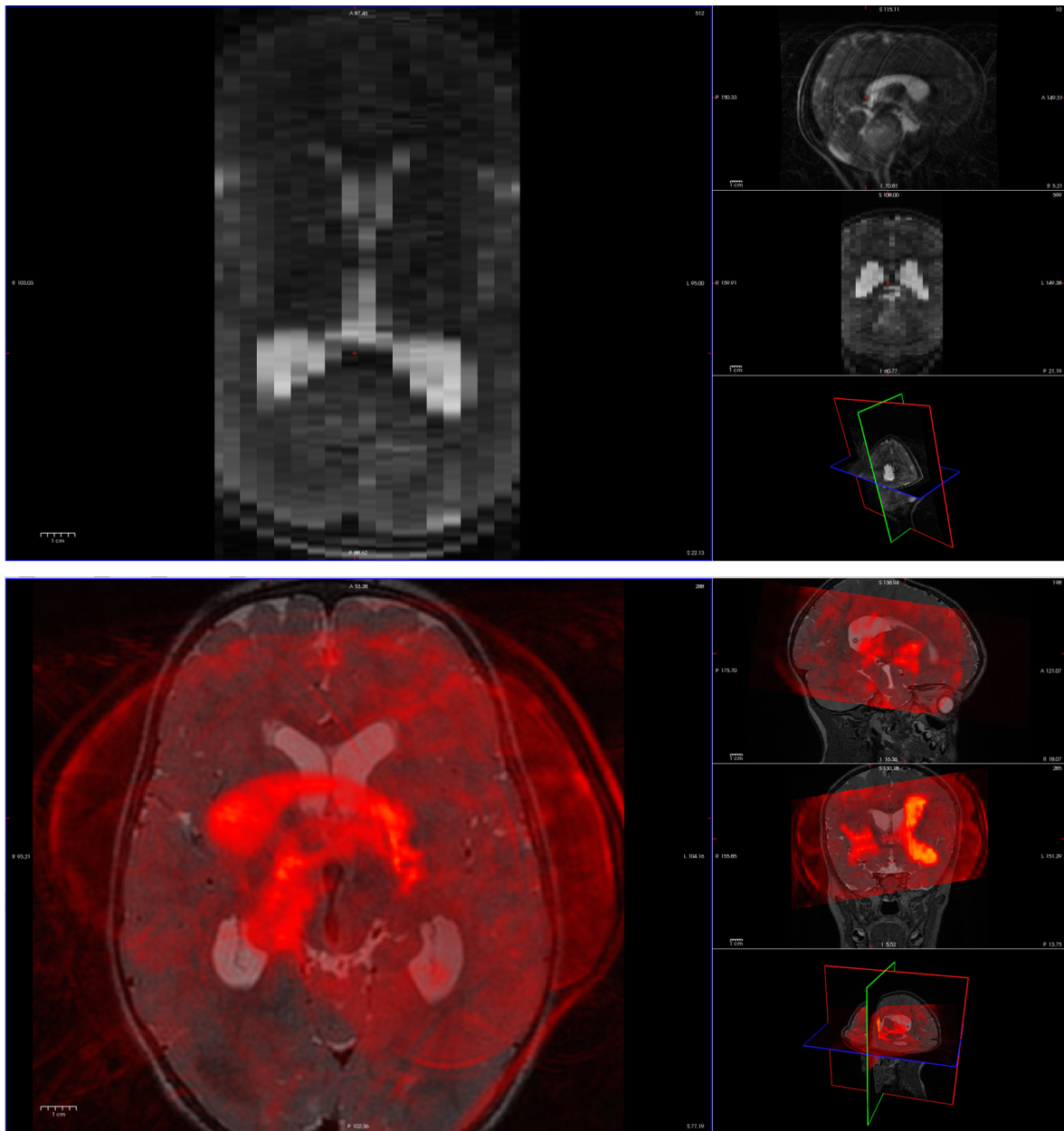


FIGURE 3.23 – Patient PREBIOMEDE sub-7905, première session : exemple d’image T2w originalement tronquée et de basse résolution (en haut) qui n’est pas bien recalée sur l’image de référence T2w de la deuxième session. L’image mal recalée est en rouge sur l’image de référence en échelle de gris sur le bas de la figure. La **DH** à l’histogramme d’intensité de référence en **MB** est égale à 0,08 et le pic de son histogramme d’intensités dans la **MG** vaut -11.

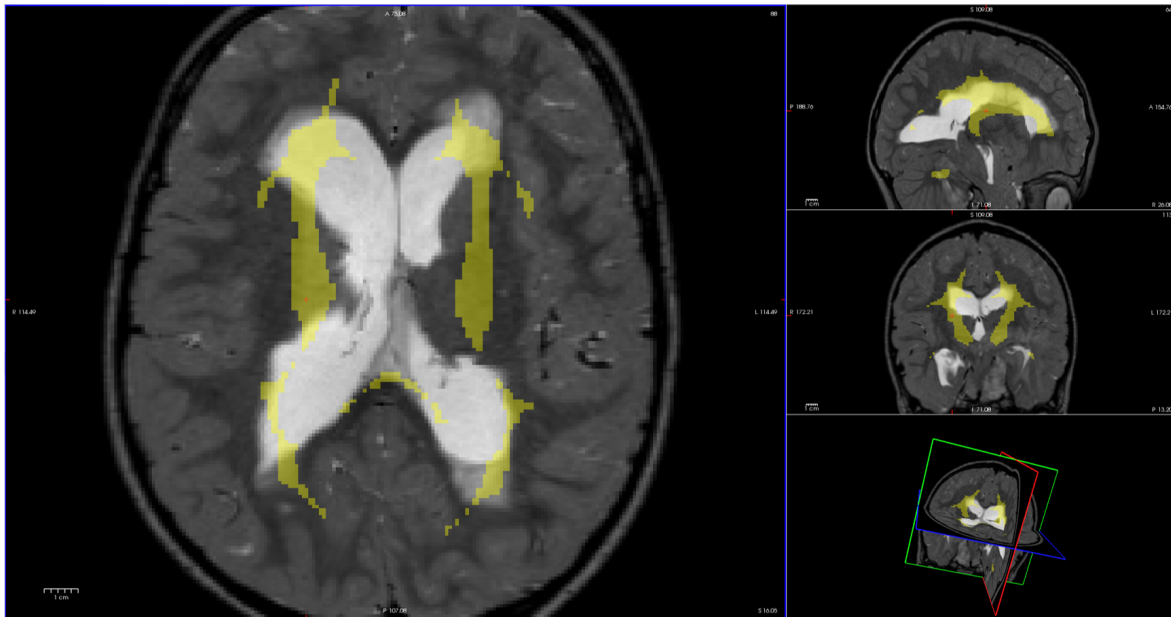


FIGURE 3.24 – Patient PREBIOMEDE sub-1756, dixième session : exemple d’image T2w avec des ventricules élargis. La DH à l’histogramme d’intensité moyenne de référence en MB est 0,26 et le pic de son histogramme d’intensités dans la MG vaut 12,5. Les masques de MB ne sont alors pas correctement estimés.

L’utilisation des images de faible qualité ou avec un effet de non-uniformité du champ magnétique n’est, en général, pas recommandée pour les analyses quantitatives.

Les recalages diffeomorphiques sont par défaut réalisés après l’application d’un masque sur le cerveau. Dans les cas qui n’ont pas fonctionné, l’image de la tête entière pour le calcul de la transformation de recalage peut permettre d’améliorer la qualité du recalage.

Les recalage rigides qui ont échoué peuvent être corrigés par un recalage affine avec 9 degrés de liberté. Le pic d’intensité de la MG à -10 repéré pour une image PREBIOMEDE T2w présentée sur la figure 3.14 correspond à une image mal recalée par rapport à l’image T2w de référence. Il se trouve qu’elle était originalement déjà tronquée et de basse résolution spatiale (voir figure 3.23), elle est détectée comme mal standardisée par le critère sur la MB.

Enfin, pour la plupart des volumes d’image correspondant aux deux dernières sources d’erreur citées (ventricules et tumeur), le problème est lié à la segmentation de la matière blanche pour la vérification des intensités; cependant l’image était néanmoins bien normalisée et pouvait être gardée.

Lésions Cérébrales : La méthode de standardisation a été testée sur une base de données d’IRM cérébrales chez des adultes atteints de lésions cérébrales. Pour les 15 images qui n’ont pas passé le CQ, l’inspection visuelle a montré des ventricules dilatés, ce qui biaise l’estimation du masque de MB dans le repère patient et l’histogramme des intensités associées à la MB. D’autre part, en excluant les lésions segmentées, certaines régions anormales restent dans les voxels analysés (figure 3.25), d’où les anomalies dans l’histogramme d’intensité de la MB. Les 4 images qui ne passent pas le CQ selon le critère de la MG ont, en effet, un très faible contraste entre MB et MG.

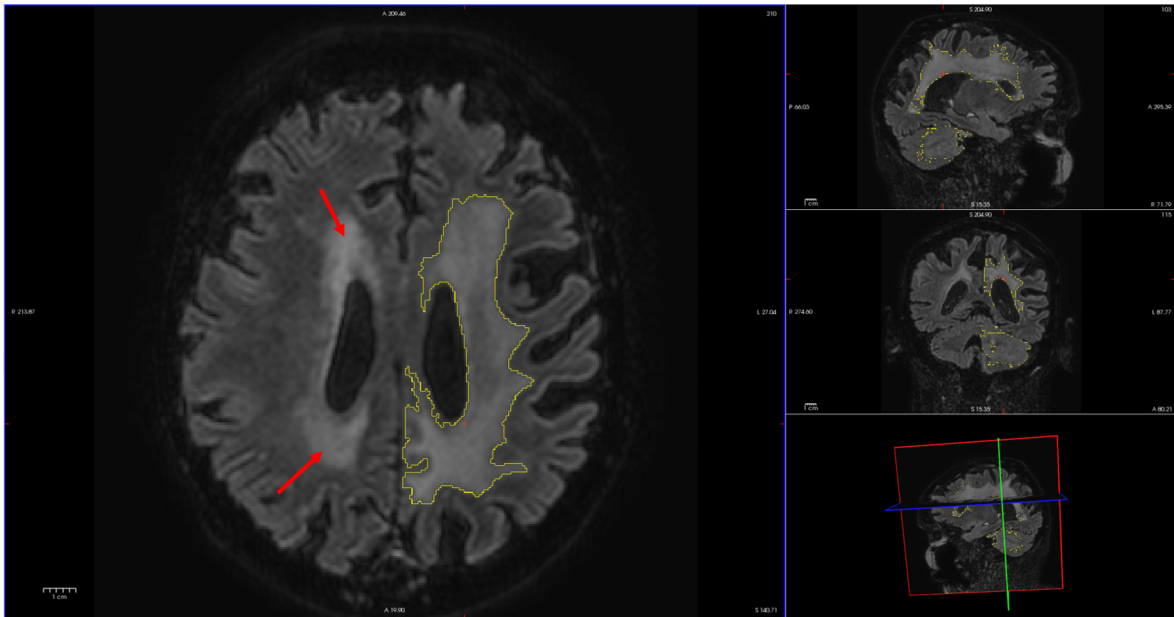


FIGURE 3.25 – Patient LC 14, examen à 1,5T : exemple d'image FLAIR avec des anomalies non-segmentées (flèches rouges). La *DH* à l'histogramme d'intensité moyenne de référence en *MB* est 0,07 et le pic de l'histogramme des intensités de la *MG* est égal à 9.

Avec cette cohorte LC nous avons observé que la standardisation a des effets positifs dans la région tumorale aussi. En considérant que les tumeurs n'ont pas évolué entre l'examen réalisé à 1,5T et celui réalisé à 3T, les points des graphes de la figure 3.19 devraient être sur la ligne identité. Or, après la standardisation d'intensités la distance à la ligne identité a diminué de 89,5% pour *FLAIR* et de 87,1% pour *T1c*. En réalité quelques tumeurs ne sont pas complètement identiques dans les deux examens et certains écarts à la ligne d'identité sont attendus. Les cas aberrants sont trouvés sur les images *T1c*. Les images réalisées à 1,5T et 3T pour les patients 10, 16, 3, 18 et 15 sont affichées respectivement sur les figures 3.26, 3.27, 3.28, 3.29 et 3.30 après normalisation avec le tracé du masque de la tumeur segmenté par A.

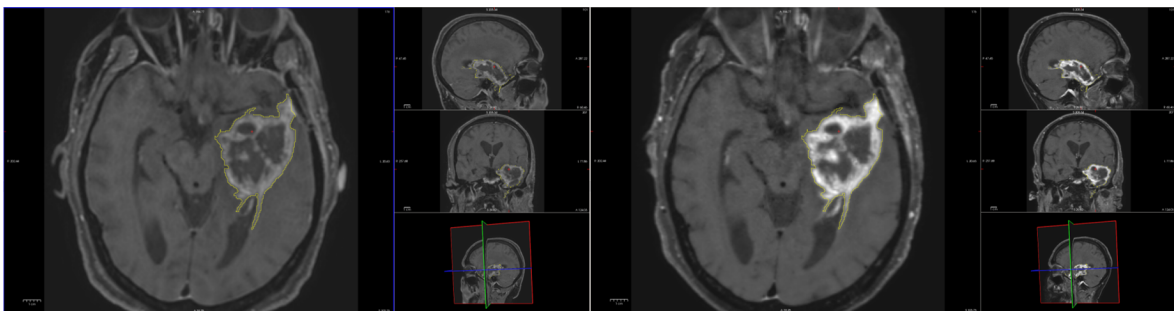


FIGURE 3.26 – Patient LC 10, 1,5T (à gauche) et 3T (à droite) : exemple d'images *T1c* qui présentent des valeurs d'intensités moyennes différentes entre les examens à 1,5T et à 3T. La segmentation de la tumeur faite par A sur l'image 1,5T est tracée en jaune. Ces deux images n'ont pas passé le CQ.

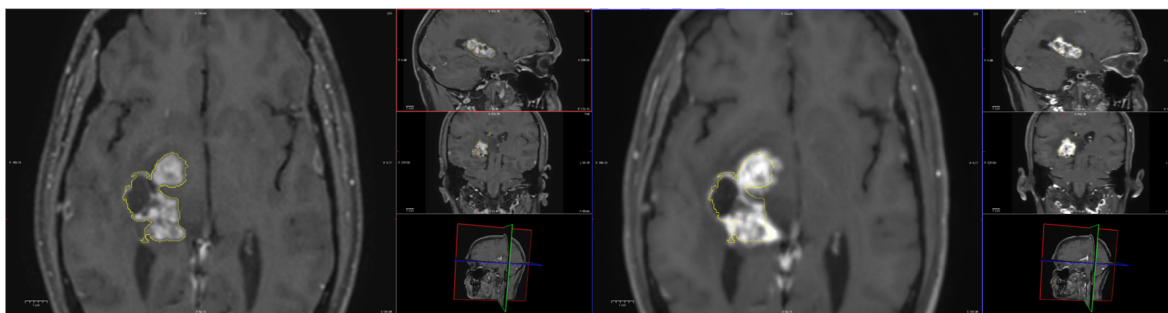


FIGURE 3.27 – Patient LC 16, 1,5T (à gauche) et 3T (à droite) : exemple d’images T1c qui présentent des valeurs d’intensités moyennes différentes entre les examens à 1,5T et à 3T. La segmentation de la tumeur faite par A sur l’image 1,5T est tracée en jaune. L’image à 3T n’a pas passé le CQ.

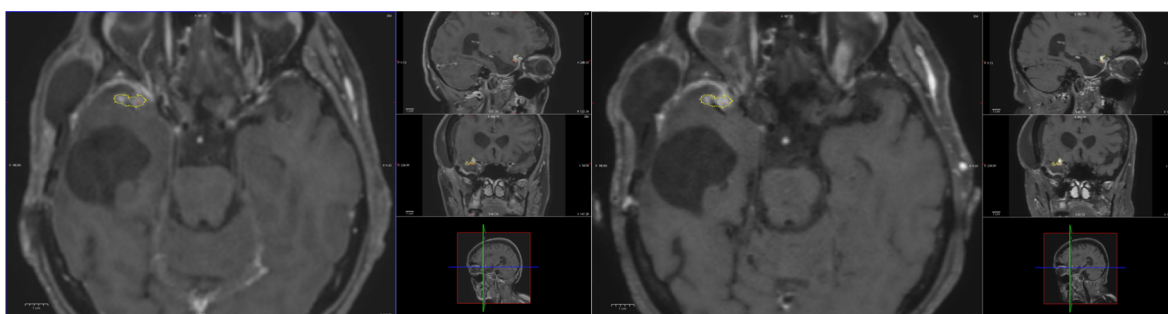


FIGURE 3.28 – Patient LC 3, 1,5T (à gauche) et 3T (à droite) : exemple d’images T1c qui présentent des valeurs d’intensités moyennes différentes entre les examens à 1,5T et à 3T. La segmentation de la tumeur faite par A sur l’image 1,5T est tracée en jaune.

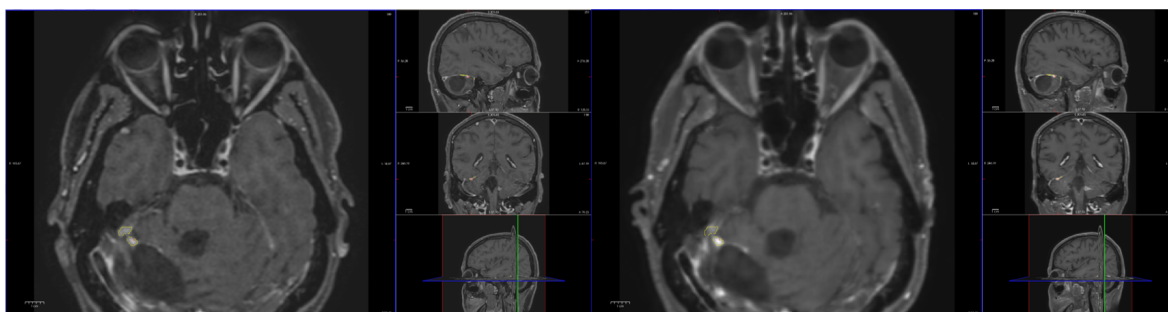


FIGURE 3.29 – Patient LC 18, 1,5T (à gauche) et 3T (à droite) : exemple d’images T1c qui présentent des valeurs d’intensités moyennes différentes entre les examens à 1,5T et à 3T. La segmentation de la tumeur faite par A sur l’image 1,5T est tracée en jaune.

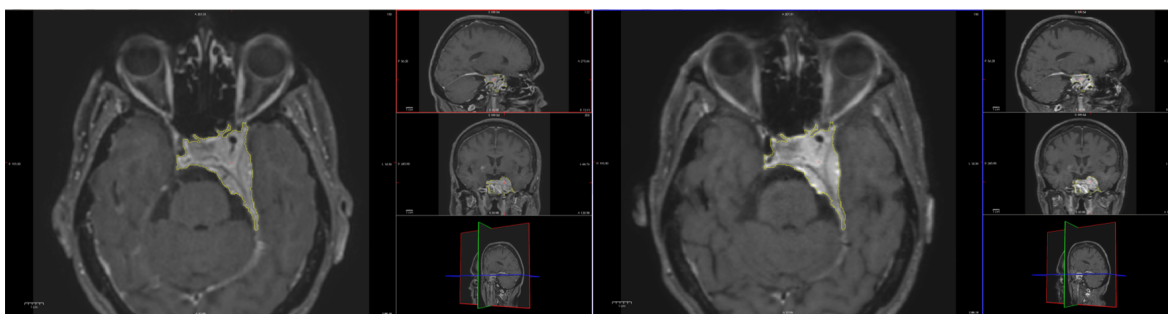


FIGURE 3.30 – Patient LC 15, 1,5T (à gauche) et 3T (à droite) : exemple d’images T1c qui présentent des valeurs d’intensités moyennes différentes entre les examens à 1,5T et à 3T. La segmentation de la tumeur faite par A sur l’image 1,5T est tracée en jaune.

Les images T1c du patient 10 à 1,5T et 3T n'ont pas passé le CQ (DH = 0,093 et DH = 0,081) et ces 2 images font partie des 7 images qui présentent des DH supérieures à 0,077. L'image T1c du patient 16 à 3T n'a pas passé le CQ de MB (DH = 0,057) et par les images, on peut vérifier que la prise de contraste a évolué entre les deux examens. Pour les patients 3 et 18, seulement une des segmentations de la tumeur était très éloignée de la droite identité et en observant les images, il apparaît que les tumeurs sont très peu volumineuses; des petits changements de prise de contraste ou des petites imperfections du recalage ou de la segmentation peuvent justifier les écarts trouvés. Sur les images du patient 15, il est possible de noter que la prise de contraste a augmenté en 3T même si les deux images ont passé le contrôle de qualité.

Comme seulement les images T1c présentent des différences très significatives dans les lésions, on peut légitimement faire l'hypothèse que la prise de contraste peut être variable avec de petites variations de temps écoulé après l'injection du produit ou de quantité de produit injecté et des variations liées à l'IRM 3T. Même si les tumeurs en FLAIR ne présentent pas de cas aberrants, on observe un décalage global entre les intensités à 1,5T et celles à 3T. Les raisons de ce comportement n'ont pas été complètement clarifiées. Une normalisation des intensités en utilisant l'approche ComBat (JOHNSON et collab. [2006]; ORLHAC et collab. [2018]) a été réalisée après le pipeline de standardisation et les valeurs moyennes dans la tumeur se sont rapprochées de la ligne d'identité (figure 3.31). Sur la même figure, on voit le résultat en appliquant ComBat sur les valeurs moyennes des images non-standardisées. Ainsi, la combinaison de deux méthodes donne les meilleurs résultats. Les distances moyennes des points à la ligne d'identité pour chaque cas sont présentés dans le tableau 3.5 et les boîtes à moustaches associées sur la figure 3.32 .

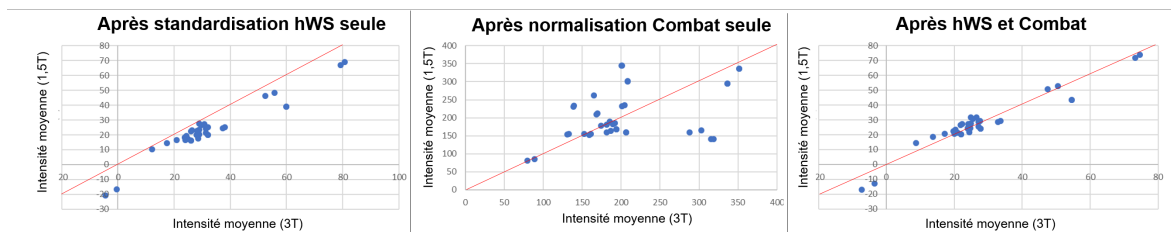


FIGURE 3.31 – Chaque point représente les valeurs moyennes d'intensité en 1,5T et 3T sans ComBat avec standardisation hWS (à gauche), normalisées par ComBat sans standardisation hWS (au milieu) et normalisées par ComBat et standardisation d'intensités hWS au préalable (à droite) pour les images FLAIR.

TABEAU 3.5 – Distances moyennes (\pm écart-type) des points (intensité à 3T, intensité à 1,5T) à la ligne d'identité avant standardisation (image originale), avec standardisation hWS, avec Combat et avec hWS et Combat.

	Originale	hWS	ComBat	hWS et ComBat
T1c	131,65 \pm 74,13	16,99 \pm 20,14	58,76 \pm 42,63	14,58 \pm 13,58
FLAIR	57,17 \pm 41,29	6,03 \pm 3,41	36,62 \pm 37,52	2,55 \pm 0,08

Des études précédentes ont montré que l'utilisation de ComBat seule pour les indices d'imagerie TEP est suffisante (ORLHAC et collab. [2018]). Par contre, en IRM, la standardisation des intensités des images est nécessaire même si les patients viennent du même centre, à cause des intensités arbitraires. Même si dans ce cas des LC, la normalisation ComBat en plus de la standardisation hWS optimise les résultats, nous ne l'avons pas appliquée sur les premières cohortes GITC car les images utilisées dans cette thèse viennent quasiment toutes d'un même centre avec le même imageur à 1,5T. En outre, nous n'avons pas suffisamment de cas venant d'autres centres pour réaliser une normalisation ComBat valable. De plus, dans le cas des lésions cérébrales des mêmes patients, les phénotypes des populations imagées à 1,5T et 3T étaient similaires, ce qui est très pertinent pour ComBat. Ce n'est pas le cas pour les cohortes GITC.

D'un autre côté, la méthode RAVEL citée dans l'état de l'art de ce chapitre n'a pas été utilisée car l'intérêt de recalcr les cerveaux de patients – ayant entre 0 à 30 ans – pour une analyse voxel-

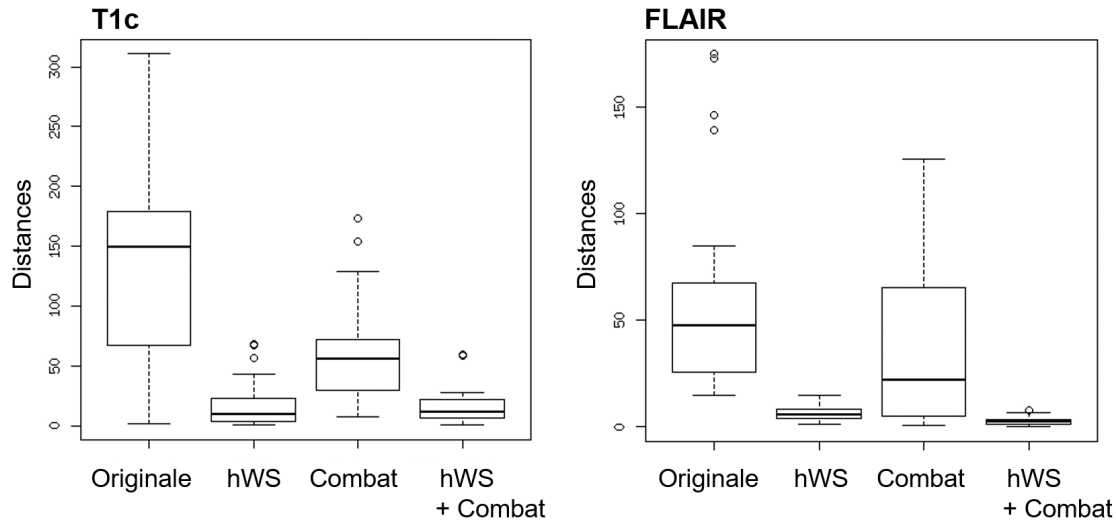


FIGURE 3.32 – Boîtes à moustaches des distances de chaque point (intensité à 3T, intensité à 1,5T) à la ligne d'identité avant standardisation (image originale), avec standardisation hWS, avec Combat et avec hWS et Combat. À gauche pour les images T1c et à droite pour les images FLAIR.

par-voxel ne nous a pas semblé pertinent. De plus, cette approche ne se prête pas facilement à l'ajout continu de patients. Elle pourra néanmoins être testée, une fois l'étude BIOMEDE terminée.

Finalement, les études de ce chapitre ont inspiré des travaux en cours dans le laboratoire IMIV pour la normalisation des images d'IRM du sein et du poumon. Le pipeline proposé a été adapté pour la région pectorale, en choisissant la graisse comme tissu de référence à la place de la MB du cerveau.

Conclusion

Le pipeline d'analyse automatique proposé dans ce chapitre comprenant une méthode de standardisation adaptée de la méthode hWS et un contrôle de qualité s'est montré efficace pour les examens multi-temps et multi-centriques, de patients des cohortes GITC, PREBIOMEDE et BIOMEDE, et pour une cohorte avec d'autres lésions cérébrales. Son efficacité a été démontrée dans la MB, la MG et dans les régions tumorales. Ce pipeline est d'une grande importance pour pouvoir comparer de façon quantitative différents patients. L'ajout d'images à la base de données permettant l'introduction de nouvelles sessions d'imagerie ou de nouveaux patients a été simplifié et automatisé. Ainsi, les études à venir sur le GITC pourront bénéficier de ce pipeline.

3.7 Références

- ASHBURNER, J. et K. J. FRISTON. 2005, «Unified segmentation», *NeuroImage*, vol. 26, n° 3, doi :10.1016/j.neuroimage.2005.02.018, p. 839–851, ISSN 10538119. URL <http://linkinghub.elsevier.com/retrieve/pii/S1053811905001102>. 64
- AVANTS, B. B., C. L. EPSTEIN, M. GROSSMAN et J. C. GEE. 2008, «Symmetric diffeomorphic image registration with cross-correlation : Evaluating automated labeling of elderly and neurodegenerative brain», *Medical Image Analysis*, vol. 12, n° 1, doi :10.1016/j.media.2007.06.004, p. 26–41, ISSN 13618415. 64

- CHRISTENSEN, J. D. 2003, «Normalization of brain magnetic resonance images using histogram even-order derivative analysis», *Magnetic Resonance Imaging*, vol. 21, n° 7, doi :10.1016/S0730-725X(03)00102-4, p. 817–820, ISSN 0730725X. URL <http://www.ncbi.nlm.nih.gov/pubmed/14559347>. 62
- FORTIN, J. P., E. M. SWEENEY, J. MUSCHELLI, C. M. CRAINICEANU et R. T. SHINOHARA. 2016, «Removing inter-subject technical variability in magnetic resonance imaging studies», *NeuroImage*, doi :10.1016/j.neuroimage.2016.02.036, ISSN 10959572. 62, 80
- GRIGIS, A., D. GOYARD, R. CHERBONNIER, T. GAREAU, D. MARCUS, D. J. JUST et A. GRIGIS. 2017, «Neuroimaging, Genetics, and Clinical Data Sharing in Python Using the CubicWeb Framework», *Frontiers in Neuroinformatics*, vol. 11, n° March, doi :10.3389/fninf.2017.00018, p. 1–11. 66
- JENKINSON, M., P. BANNISTER, M. BRADY et S. SMITH. 2002, «Improved optimization for the robust and accurate linear registration and motion correction of brain images», *NeuroImage*, vol. 17, n° 2, doi :10.1016/S1053-8119(02)91132-8, p. 825–841, ISSN 10538119. 63
- JOHNSON, W. E., C. LI et A. RABINOVIC. 2006, «Adjusting batch effects in microarray expression data using empirical Bayes methods», *Biostatistics*, vol. 8, n° 1, doi :10.1093/biostatistics/kxj037, p. 118–127, ISSN 1465-4644. URL <https://dx.doi.org/10.1093/biostatistics/kxj037>. 87
- MADABHUSHI, A. et J. K. UDUPA. 2006, «New methods of MR image intensity standardization via generalized scale», *Medical Physics*, vol. 33, doi :10.1118/1.2335487, p. 3426–3434, ISSN 00942405. 62
- NIKULIN, M. S. 2001, *Hellinger distance estimation*, vol. 151, 1994^e éd., Springer, ISBN 978-1556080104, 1–5 p.. 67
- NYÚL, L. G., J. K. UDUPA et X. ZHANG. 2000, «New variants of a method of MRI scale standardization», *IEEE Transactions on Medical Imaging*, vol. 19, n° 2, doi :10.1109/42.836373, p. 143–150, ISSN 02780062. URL <http://ieeexplore.ieee.org/document/836373/>. 62, 80
- ORLHAC, F., S. BOUGHDAD, C. PHILIPPE, H. STALLA-BOURDILLON, C. NIOCHE, L. CHAMPION, M. SOUSSAN, F. FROUIN, V. FROUIN et I. BUVAT. 2018, «A postreconstruction harmonization method for multicenter radiomic studies in pet», *Journal of Nuclear Medicine*, vol. 59, doi : 10.2967/jnumed.117.199935, p. jnumed.117.199935. 87
- RUPPERT, D., M. P. WAND et R. J. CARROLL. 2003, *Semiparametric Regression*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, doi :10.1017/CBO9780511755453. 63
- SHAH, M., Y. XIAO, N. SUBBANNA, S. FRANCIS, D. L. ARNOLD, D. L. COLLINS et T. ARBEL. 2011, «Evaluating intensity normalization on MRIs of human brain with multiple sclerosis», *Medical Image Analysis*, vol. 15, n° 2, doi :10.1016/J.MEDIA.2010.12.003, p. 267–282, ISSN 1361-8415. URL <https://www.sciencedirect.com/science/article/pii/S1361841510001337?via=ihub>. 62
- SHINOHARA, R. T., E. M. SWEENEY, J. GOLDSMITH, N. SHIEE, F. J. MATEEN, P. A. CALABRESI, S. JARSO, D. L. PHAM, D. S. REICH et C. M. CRAINICEANU. 2014, «Statistical normalization techniques for magnetic resonance imaging», *NeuroImage : Clinical*, vol. 6, doi :10.1016/j.nicl.2014.08.008, p. 9–19, ISSN 22131582. 62
- TUSTISON, N. J., B. B. AVANTS, P. A. COOK, Y. ZHENG, A. EGAN, P. A. YUSHKEVICH et J. C. GEE. 2010, «N4ITK : Improved N3 bias correction», *IEEE Transactions on Medical Imaging*, vol. 29, n° 6, doi :10.1109/TMI.2010.2046908, p. 1310–1320, ISSN 02780062. 62

Chapitre 4

Etudes préliminaires en radiomique

« Ceux qui ne marchent que fort lentement peuvent avancer beaucoup davantage, s'ils suivent toujours le droit chemin, que ne font ceux qui courent, et qui s'en éloignent. »

René Descartes



Ce chapitre considère les choix méthodologiques pour l'extraction d'indices pour les études en radiomique. Nous nous sommes intéressés à analyser, en premier, la performance de trois types de discrétisation des niveaux de gris des images pour le calcul des indices de texture. La méthode pour les études radiomiques en IRM a pu être choisie. Ensuite, j'ai analysé les corrélations des indices de texture avec le volume de la région dans laquelle ils sont calculés. En conséquence des résultats obtenus, ce chapitre présente une méthode de calcul d'indices radiomiques qui, tout en caractérisant les tumeurs globalement, ne sont pas biaisés par le volume de la région considérée.

Sommaire

4.1	Questions méthodologiques posées pour le calcul des indices radiomiques en IRM	92
4.2	Impact de la discrétisation des niveaux de gris sur les indices radiomiques	92
4.2.1	Comparaison de trois méthodes de discrétisation	95
4.2.2	Choix de différents paramètres pour chaque méthode	96
4.2.3	Lien entre inspection visuelle et indices de texture	98
4.3	Impact du volume de la région d'intérêt sur les indices radiomiques	106
4.3.1	Corrélation entre indices radiomiques et volume de la région d'intérêt	107
4.3.2	Impact du volume d'intérêt sur les indices radiomiques de régions similaires	107
4.4	Proposition d'une méthode de calcul local d'indices radiomiques	112
4.4.1	Calcul d'indices radiomiques locaux dans la tumeur	112
4.4.2	Comparaison entre les indices radiomiques locaux et globaux pour caractériser les régions tumorales	112
4.5	Discussion et conclusion	116
4.6	Références	118

4.1 Questions méthodologiques posées pour le calcul des indices radiomiques en IRM

Dans le domaine prometteur de la radiomique, les indices de texture en imagerie peuvent être calculés de plusieurs façons. Le manque d'attention porté aux choix méthodologiques conduit à des indices qui ne reflètent pas bien l'hétérogénéité tumorale ou qui ne sont pas comparables entre différents patients. En IRM, le premier pas pour arriver à des indices d'imagerie comparables est la standardisation d'intensités décrite dans le chapitre 3. En plus de cette normalisation, d'autres éléments de calcul peuvent avoir aussi un grand impact sur les caractéristiques extraites.

Certaines études en TEP comme celles d'ORLHAC et collab. [2014] et de HATT et collab. [2015] ont montré la dépendance de quelques indices radiomiques aux volumes des tumeurs dans lesquels ils sont calculés. Cette dépendance a déjà conduit à des conclusions qui doivent être nuancées. Par exemple, AERTS et collab. [2014] ont trouvé une signature radiomique qui inclut l'indice *Gray Level Non-Uniformity* des images tomographiques pour prédire du pronostic des patients de cancer du poumon. Or, cet indice est fortement corrélé au volume tumoral et la bonne performance de la signature pouvait être simplement expliquée par le volume tumoral (VALLIERES et collab. [2018]). Récemment, LARROZA et collab. [2016] ont recommandé l'utilisation de régions de même taille pour analyser toutes les tumeurs d'une étude en IRM 2D quand les tailles des tumeurs varient beaucoup. Cependant, à notre connaissance cette problématique de corrélation avec le volume n'a pas été approfondie en IRM 3D.

Un autre choix méthodologique d'importance est celui des différentes méthodes de discrétisation de niveaux de gris pour le calcul des indices de texture de deuxième ordre. En TEP, des méthodes telles que la discrétisation absolue et la discrétisation relative ont été comparées (LEIJENAAR et collab. [2015]; ORLHAC et collab. [2015]). En IRM, MOLINA et collab. [2017] ont montré que les indices de texture calculés avec différents intervalles d'intensité n'étaient pas comparables entre-eux. Ils concluent que pour comparer ces indices, il faut qu'ils soient tous calculés avec la même discrétisation. Néanmoins, aucune étude en IRM n'a comparé les différentes méthodes de discrétisation entre elles pour indiquer la meilleure à choisir.

Pour les études méthodologiques d'extraction d'indices radiomiques présentés dans ce chapitre, la cohorte étoile avec les images standardisées a été utilisée.

4.2 Impact de la discrétisation des niveaux de gris sur les indices radiomiques

Pour choisir la meilleure méthode de discrétisation des niveaux de gris d'une image IRM, la région dans laquelle la biopsie a été réalisée a été utilisée. Étant donné que les différentes modalités d'IRM de la cohorte étoile sont recalées entre elles par patients, pour faciliter l'identification de la région de la biopsie, les images T2w (ou T1w) des sessions après biopsie ont été soustraites des images avant biopsie de la modalité correspondante (figure 4.1). Comme les biopsies sont faites avec une aiguille qui prélève environ une sphère de rayon 5 mm, une région sphérique de 4,5 mm de rayon a été définie autour de l'endroit de la biopsie pour chacun des patients de la cohorte étoile avec le logiciel LIFEx (NIOCHE et collab. [2018]) qui propose la création de régions sphériques avec une interface interactive. Ces régions sont notées $VOI_{biopsie}$. Ce même tracé a été utilisé sur toutes les modalités d'imagerie structurelle du même patient.

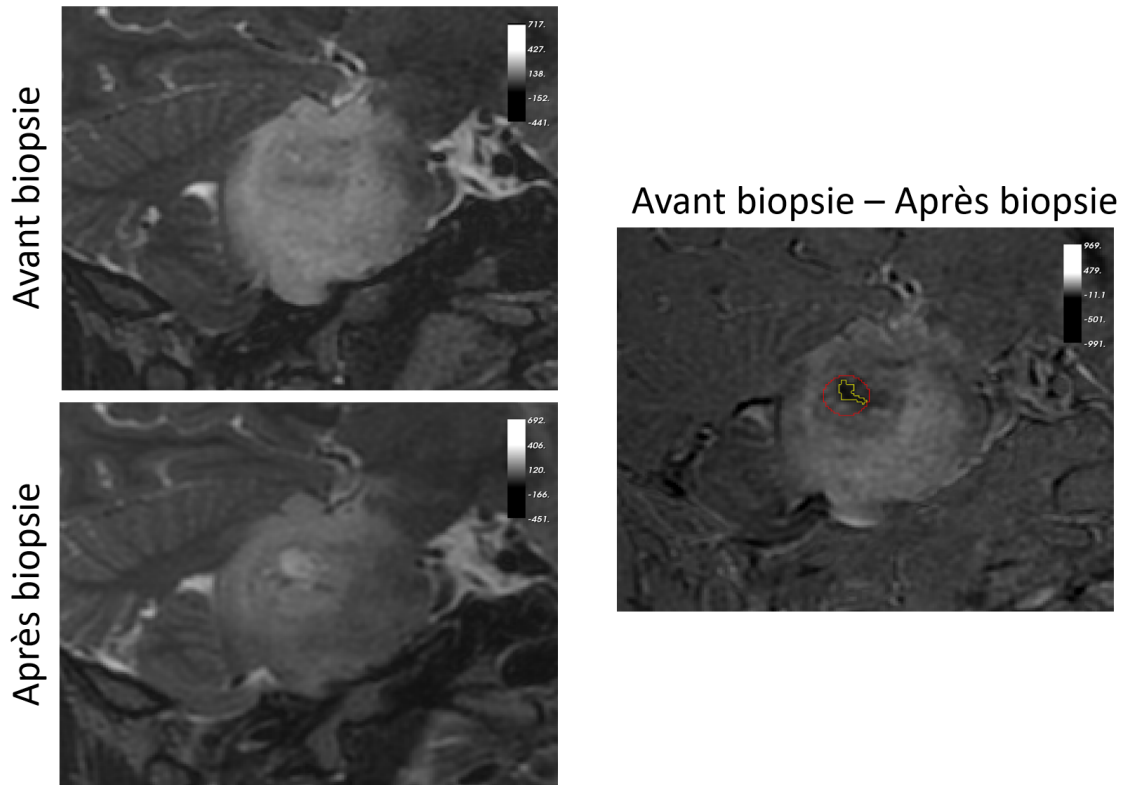


FIGURE 4.1 – Patient PREBIOMEDE sub-5951 : Région de plus grosse différence dans l’image de soustraction des examens avant et après biopsie en jaune. En rouge, le contour de la sphère de rayon 4,5 mm ($VOI_{biopsie}$).

Toutes ces images ont été ré-échantillonnées à des voxels isotropes de 1mm x 1mm x 1mm avant le calcul des matrices de texture pour que toutes les directions de l’image (x, y et z) aient la même importance dans les indices de texture. Avec l’aide de la librairie PyRadiomics version 1.1.1 (VAN GRIETHUYSEN et collab. [2017]), 60 indices radiomiques de texture (28 de la matrice GLCM, 16 de la GLSZM et 16 de la GLRLM) ont été calculés dans la $VOI_{biopsie}$ sur les images T1w, T2w, T1c et FLAIR de diagnostic des patients de la cohorte étoile. Le tableau 4.1 liste tous les indices de textures analysés et la documentation de PyRadiomics, en annexe A.2.1, décrit les équations mathématiques pour les calculer.

Dans cette étude, trois types de discrétisations, qui ont déjà été utilisées précédemment dans des études en TEP et en IRM, ont été comparés :

1. $d_{rel, f_{LN}}$: bornes relatives et largeur de niveau de gris fixe. Pour chaque $VOI_{biopsie}$, les valeurs minimales et maximales des intensités au sein de la région définissent les bornes de discrétisation. La même largeur d’intensité pour la discrétisation est utilisée pour tous les patients (LEIJENAAR et collab. [2015]).
2. $d_{rel, f_{NN}}$: bornes relatives et nombre de niveaux de gris fixe. Pour chaque $VOI_{biopsie}$, les valeurs minimale et maximale des intensités au sein de la région définissent les bornes de discrétisation. Le même nombre de niveaux de gris de discrétisation est utilisé pour tous les patients (MICHOUX et collab. [2015], YANG et collab. [2015], MOLINA et collab. [2016], KICKINGEREDER et collab. [2016]).
3. $d_{abs, f_{NN}}$: bornes absolues et largeur de niveau de gris fixe. Le minimum des valeurs des 10^{ème} percentiles des 30 patients et le maximum des valeurs des 90^{ème} percentiles ont défini les bornes de discrétisation pour tous les patients. Le même nombre de niveaux de gris de discrétisation est utilisé pour tous les patients, ce qui implique que la même largeur de niveau de gris de discrétisation a été aussi utilisée pour tous les patients (ORLHAC et collab. [2015]).

TABEAU 4.1 – Liste des indices de texture de deuxième ordre calculés avec PyRadiomics.

GLCM	GLRLM	GLSZM
SumVariance	ShortRunLowGrayLevelEmphasis	GrayLevelVariance
Homogeneity1	GrayLevelVariance	LowIntensityLargeAreaEmphasis
Homogeneity2	LowGrayLevelRunEmphasis	HighIntensitySmallAreaEmphasis
ClusterShade	GrayLevelNonUniformityNormalized	SmallAreaEmphasis
MaximumProbability	RunVariance	LargeAreaEmmphasis
Idmn	GrayLevelNonUniformity	ZoneVariance
SumVariance2	LongRunEmphasis	SizeZoneVariabilityNormalized
Contrast	ShortRunHighGrayLevelEmphasis	LowIntensitySmallAreaEmphasis
DifferenceEntropy	RunLenghtNonUniformity	HighIntensityEmphasis
InverseVariance	ShortRunEmphasis	IntensityVariabilityNormalized
Entropy	LongRunHighGrayLevelEmphasis	ZonePercentage
Dissimilarity	RunPercentage	LowIntensityEmphasis
DifferenceVariance	LongRunLowGrayLevelEmphasis	SizeZoneVariability
Idn	RunEntropy	IntensityVariability
Idm	HighGrayLevelRunEmphasis	ZoneEntropy
Correlation	RunLengthNonUniformityNormalized	HighIntensityLargeAreaEmphasis
AutoCorrelation		
SumEntropy		
AverageIntensity		
Energy		
SumSquares		
ClusterProminence		
SumAverage		
Imc2		
Imc1		
DifferenceAverage		
Id		
ClusterTendency		

Ces trois types de discrétisation sont illustrés sur la figure 4.2 qui montre deux régions simulées (ROI1 et ROI2). Pour chaque type de discrétisation, les valeurs des intensités de borne, la largeur du niveau de gris et le nombre de niveaux de gris, ainsi que les images discrétisées correspondantes sont présentés.

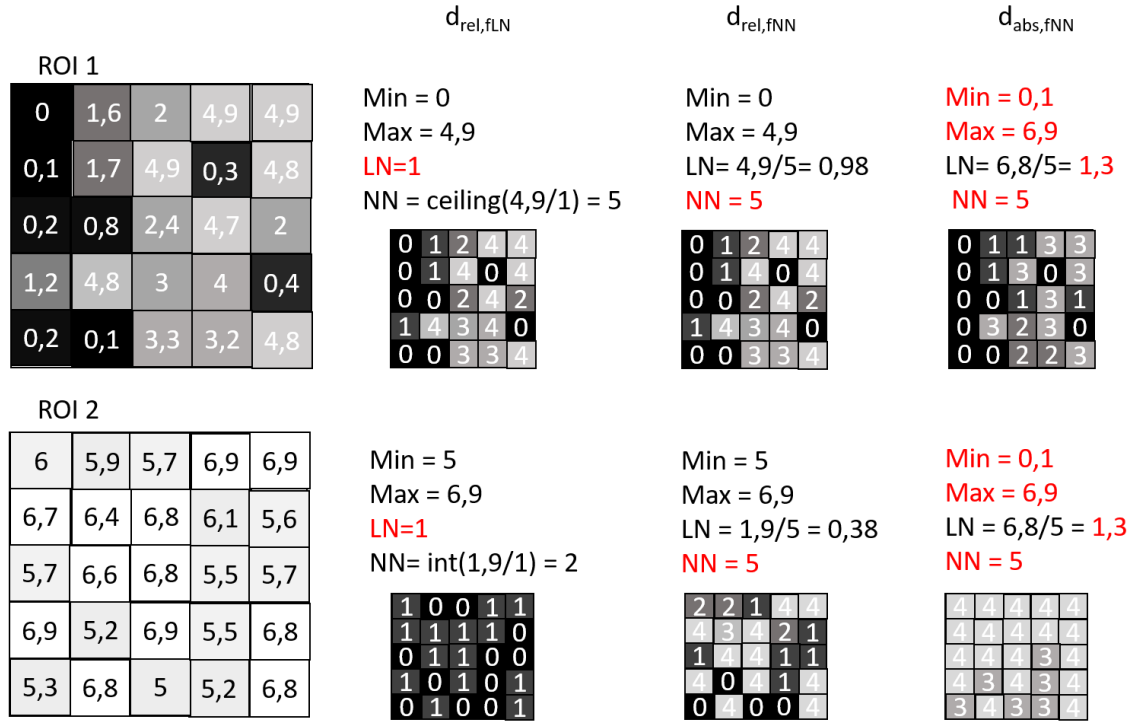


FIGURE 4.2 – Illustration des trois types de discrétisation proposés avec deux ROIs différentes. Les paramètres qui sont fixés pour toutes les ROIs de l'étude sont en rouge.

Les analyses ont été divisées en trois parties. D'abord, nous avons comparé les trois méthodes de discrétisation entre elles en ordonnant les patients par les valeurs des indices radiomiques extraits avec chaque méthode. Ensuite, en faisant varier le nombre et la largeur de niveau de gris à fixer nous avons comparé comment le classement des patients varie. Pour conclure, les classements obtenus par chacune des méthodes de discrétisation ont été vérifiés avec l'inspection visuelle faite par trois experts.

4.2.1 Comparaison de trois méthodes de discrétisation

Pour comparer les trois méthodes de discrétisation, le nombre de niveaux de gris (NN) a été fixé à 64 pour les discrétisations $d_{rel,fNN}$ et $d_{abs,fNN}$ et la largeur des niveaux de gris (LN) a été fixée à 2 pour la discrétisation $d_{rel,fLN}$.

Pour chacun des 240 (60 x 4 modalités IRM) indices extraits de la $VOI_{biopsie}$, les rangs des n=30 patients ont été définis selon les valeurs croissantes des indices. La corrélation de Spearman (r) a été évaluée entre les rangs de patients calculés pour toute combinaison de paire de méthodes ($d_{rel,fLN}$, $d_{rel,fNN}$ et $d_{abs,fNN}$), pour chaque indice radiomique. Plus grande est la valeur de r, plus similaires sont les rangs des patients calculés par les deux méthodes comparées. La corrélation de Spearman (r) est rappelée par l'équation 4.1, où R_{d1} (R_{d2}) est le rang des patients obtenu par la méthode de discrétisation d1 (d2) et $\sigma_{R_{d1}}$ ($\sigma_{R_{d2}}$) est son écart-type. La valeur de r varie de -1 à 1.

$$r = \frac{cov(R_{d1}, R_{d2})}{\sigma_{R_{d1}} \sigma_{R_{d2}}} \quad (4.1)$$

La figure 4.3 résume les combinaisons étudiées.

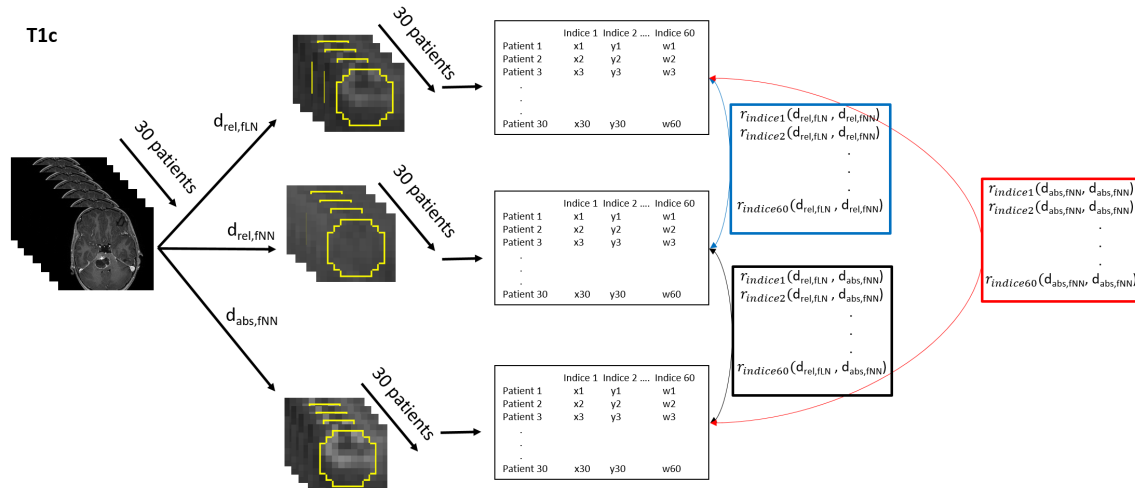


FIGURE 4.3 – Schéma de la méthode utilisée pour vérifier la cohérence des indices radiomiques avec les trois méthodes de discrétisation. Sur le schéma, le processus est illustré pour les images T1c, mais il est répété pour les 4 modalités d'IRM

Les rangs des patients étaient fortement corrélés ($|r| > 0,7$) pour 20 de 240 indices entre les discrétisations $d_{rel,fLN}$ et $d_{rel,fNN}$, pour 188 de 240 indices entre $d_{rel,fLN}$ et $d_{abs,fNN}$, et pour 9 de 240 indices entre $d_{rel,fNN}$ et $d_{abs,fNN}$. La figure 4.4 montre les $|r|$ obtenus dans les trois configurations pour chaque indice radiomique. Il ressort de cette étude que les méthodes $d_{rel,fLN}$ et $d_{abs,fNN}$ donnent des résultats très similaires tandis que la méthode $d_{rel,fNN}$ donne des résultats différents.

4.2.2 Choix de différents paramètres pour chaque méthode

Pour comparer les trois méthodes de discrétisation comme présenté au paragraphe précédent, le nombre de niveaux de gris a été fixé ($NN = 64$) ainsi que la largeur des niveaux de gris ($LN = 2$). Afin de vérifier la robustesse de chaque méthode selon le choix de ces paramètres, pour la discrétisation $d_{rel,fLN}$, quatre autres valeurs de LN ont été testées ($LN \in \{0,75; 1; 3; 4\}$) et les indices radiomiques résultants ont été comparés aux indices obtenus avec $LN = 2$ en utilisant l'approche reposant sur la corrélation de Spearman. Ensuite, pour les discrétisations $d_{rel,fNN}$ et $d_{abs,fNN}$, quatre autres valeurs de NN ont été testées ($NN \in \{8; 16; 32; 128\}$) et les indices radiomiques résultants ont été comparés aux indices obtenus avec $NN = 64$ avec la corrélation de Spearman (équation 4.1). La figure 4.5 résume toutes les combinaisons pour lesquelles l'étude a été menée.

La figure 4.6 montre les valeurs absolues de r pour chaque indice avec la discrétisation $d_{rel,fLN}$ et différents choix de paramètres. On peut noter que les rangs sont fortement corrélés, à l'exception de quatre indices de la matrice GLSZM dans les quatre modalités d'IRM, pour lesquels les rangs dépendent du choix de LN.

Les valeurs absolues de r pour chaque indice en fonction de NN avec la discrétisation $d_{rel,fNN}$ sont présentées sur la figure 4.7, avec la discrétisation $d_{abs,fNN}$ sur la figure 4.8. Les résultats sont plutôt cohérents entre les différentes valeurs de NN; cependant, les valeurs les plus petites de NN (8 et 16) conduisent à différents classements de patients pour quelques indices. Dans la majorité des cas, ces indices sont les mêmes pour toutes les modalités IRM et surtout pour ceux qui proviennent de la GLSZM.

Globalement, la cohérence des valeurs des indices par rapport aux choix de paramètres de chaque méthode de discrétisation est très élevée. Pour la discrétisation $d_{rel,fLN}$, il faut porter plus d'attention à certains indices de la matrice GLSZM. D'un autre côté, pour les discrétisations $d_{rel,fNN}$ et $d_{abs,fNN}$, les valeurs de $|r|$ sont faibles pour $NN = 8$ et $NN = 16$; cependant, on peut considérer que ces nombres sont trop faibles pour traduire l'hétérogénéité des images.

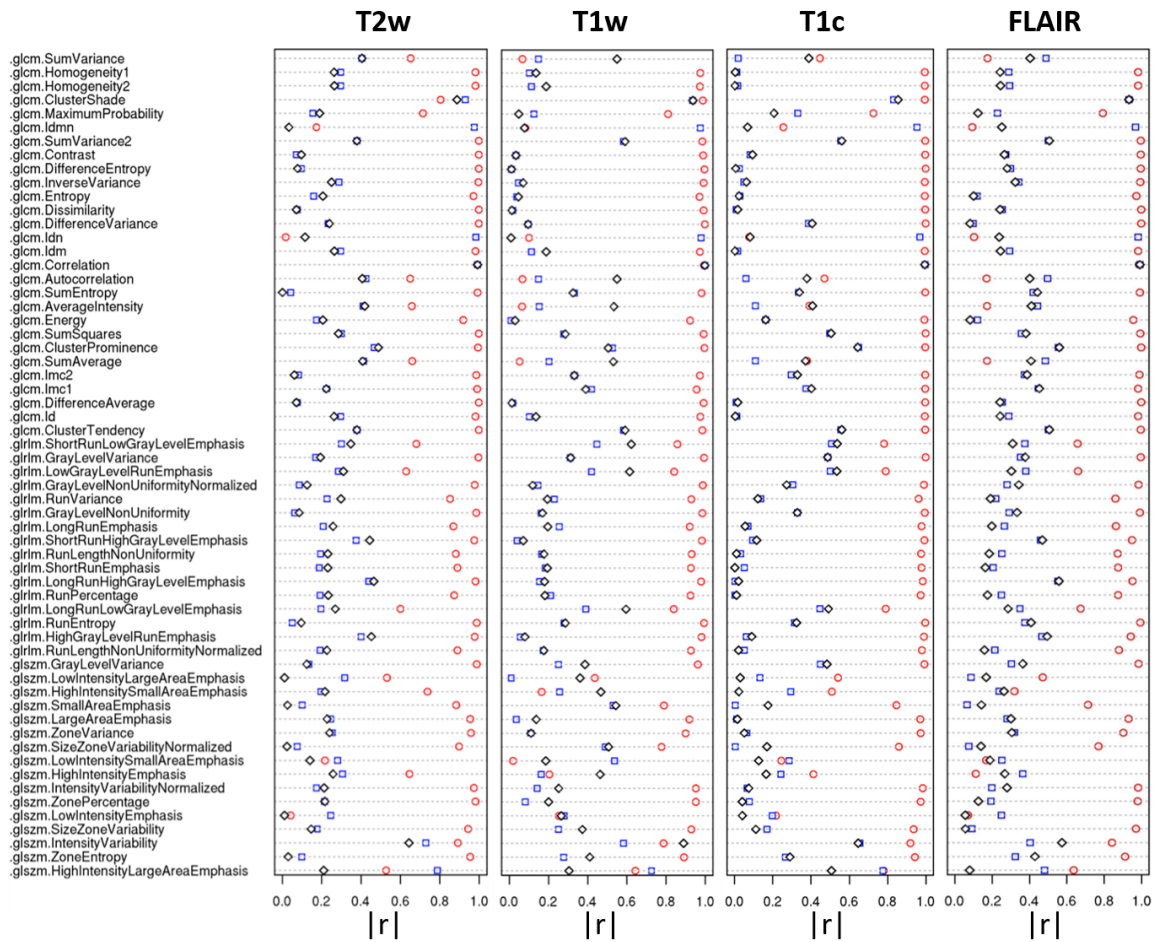


FIGURE 4.4 – Coefficient de corrélation de Spearman (en valeur absolue $|r|$) pour chacun des 240 indices radiomiques entre les discrétisations $d_{rel, fLN}$ et $d_{rel, fNN}$ en bleu, $d_{rel, fLN}$ et $d_{abs, fNN}$ en rouge et entre $d_{rel, fNN}$ et $d_{abs, fNN}$ en noir.

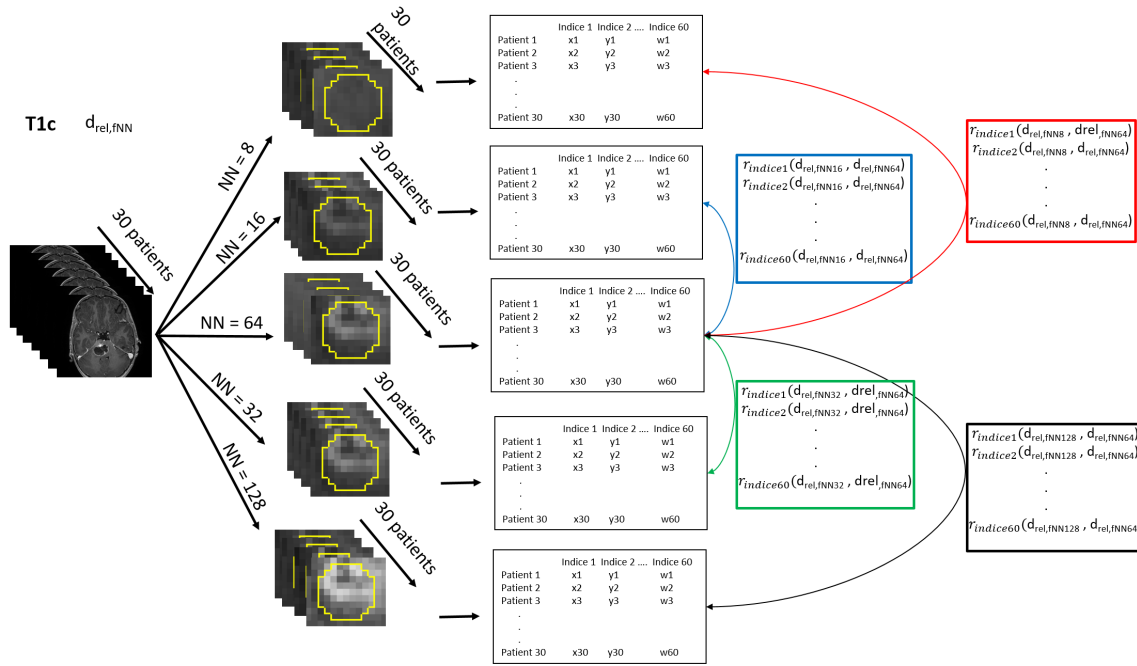


FIGURE 4.5 – Schéma des r calculés pour vérifier la cohérence des indices radiomiques en fonction du choix des paramètres NN et LN . Sur le schéma, le processus pour les images T1c et la discrétisation $d_{rel,fNN}$ est présenté, mais la même approche est considérée pour les 4 modalités d'IRM et les 3 types de discrétisations.

4.2.3 Lien entre inspection visuelle et indices de texture

Après avoir vérifié que les méthodes $d_{rel,fLN}$ et $d_{abs,fNN}$ sont cohérentes entre elles, à la différence de la méthode $d_{rel,fNN}$, une étude sur la cohérence des indices dérivés par chaque méthode avec l'inspection visuelle de trois experts a été menée. Pour cette étude, $LN = 2$ a été fixée pour $d_{rel,fLN}$ et $NN = 64$ pour $d_{rel,fNN}$ et $d_{abs,fNN}$.

Comme les indices de textures sont calculés pour quantifier l'hétérogénéité d'une tumeur (ORLHAC et collab. [2015]), trois experts ont étiqueté comme 1 = hétérogène ou 0 = homogène chaque $VOI_{biopsie}$ dans chaque modalité d'IRM indépendamment (30 x 4 modalités = 120 étiquettes par expert). Ils avaient accès à trois images 2D correspondant aux trois axes centraux (axial, sagittal et coronal) de la $VOI_{biopsie}$ dans chaque modalité, comme illustré sur la figure 4.9. Les images étaient représentées avec la même échelle de couleur pour tous les sujets. Chaque expert a travaillé indépendamment.

Pour chaque indice radiomique, les étiquettes données par chaque expert séparément ont été utilisées pour calculer les fonctions d'efficacité du récepteur (courbes ROC) traçant la sensibilité versus $(1 - \text{spécificité})$. Pour vérifier la capacité de chaque indice à caractériser l'hétérogénéité tumorale en accord avec une appréciation visuelle, l'AUC a été calculée pour chacun des indices radiomiques entre chaque étiquetage des experts et chaque méthode de discrétisation. Comme certains indices augmentent avec l'hétérogénéité (comme l'Entropie) et d'autres diminuent (comme l'Homogénéité), l'étiquette 'positive' pour le calcul de l'AUC a été définie comme celle pour laquelle l'AUC est plus grande que 0,5 pour au moins deux méthodes de discrétisation.

Le tableau 4.2 montre les AUC moyennes des 240 indices radiomiques obtenues pour chaque expert et chaque méthode de discrétisation. Les méthodes $d_{rel,fLN}$ et $d_{abs,fNN}$ sont cohérentes avec les étiquettes des trois experts, ce qui n'est pas toujours le cas pour $d_{rel,fNN}$.

Plus en détail, les figures 4.11, 4.12 et 4.13 montrent les AUC de chaque indice radiomique pour les trois méthodes de discrétisation pour les experts 1, 2 et 3 respectivement.

L'indice radiomique $T1c\ GLCM\ Homogeneity1$, par exemple, présente une AUC moyenne pour les trois experts de 0,90 pour $d_{rel,fLN}$ et pour $d_{abs,fNN}$ et de 0,43 pour $d_{rel,fNN}$. Dans le cas de cet indice, le rang 1 correspond au patient avec l'homogénéité la plus faible, donc, le plus hétérogène.

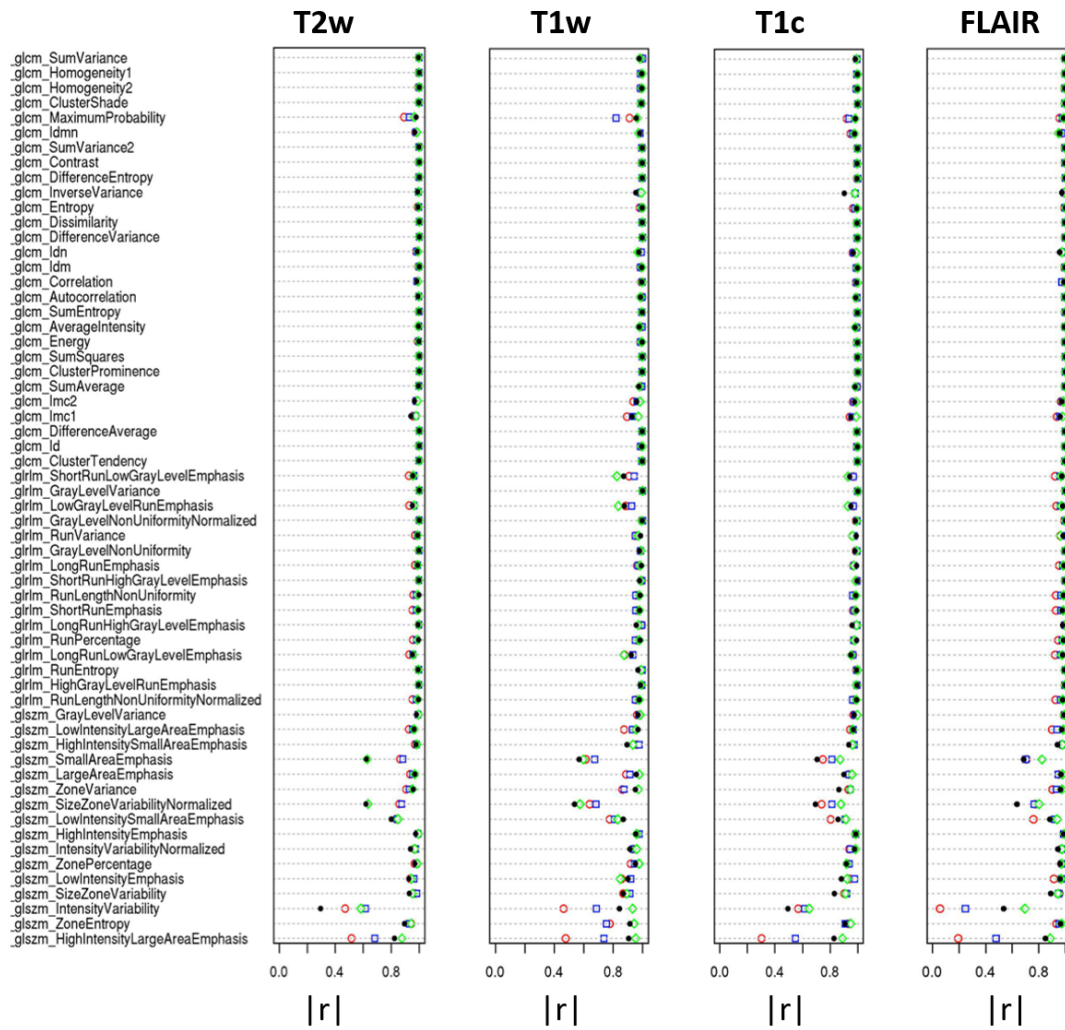


FIGURE 4.6 – Coefficients de corrélation de Spearman (en valeur absolue $|r|$) pour chacun des 240 indices radiomiques de la discrétisation $d_{rel,fLN}$ avec $LN = 2$ versus : $LN = 0,75$ en rouge, $LN = 1$ en bleu, $LN = 3$ en vert et $LN = 4$ en noir.

TABEAU 4.2 – AUC moyennes des 240 indices radiomiques obtenus pour les différentes méthodes de discrétisation en fonction des trois experts.

	$d_{rel,fLN}$	$d_{rel,fNN}$	$d_{abs,fNN}$
Expert 1	0,87	0,63	0,83
Expert 2	0,83	0,61	0,79
Expert 3	0,76	0,59	0,74

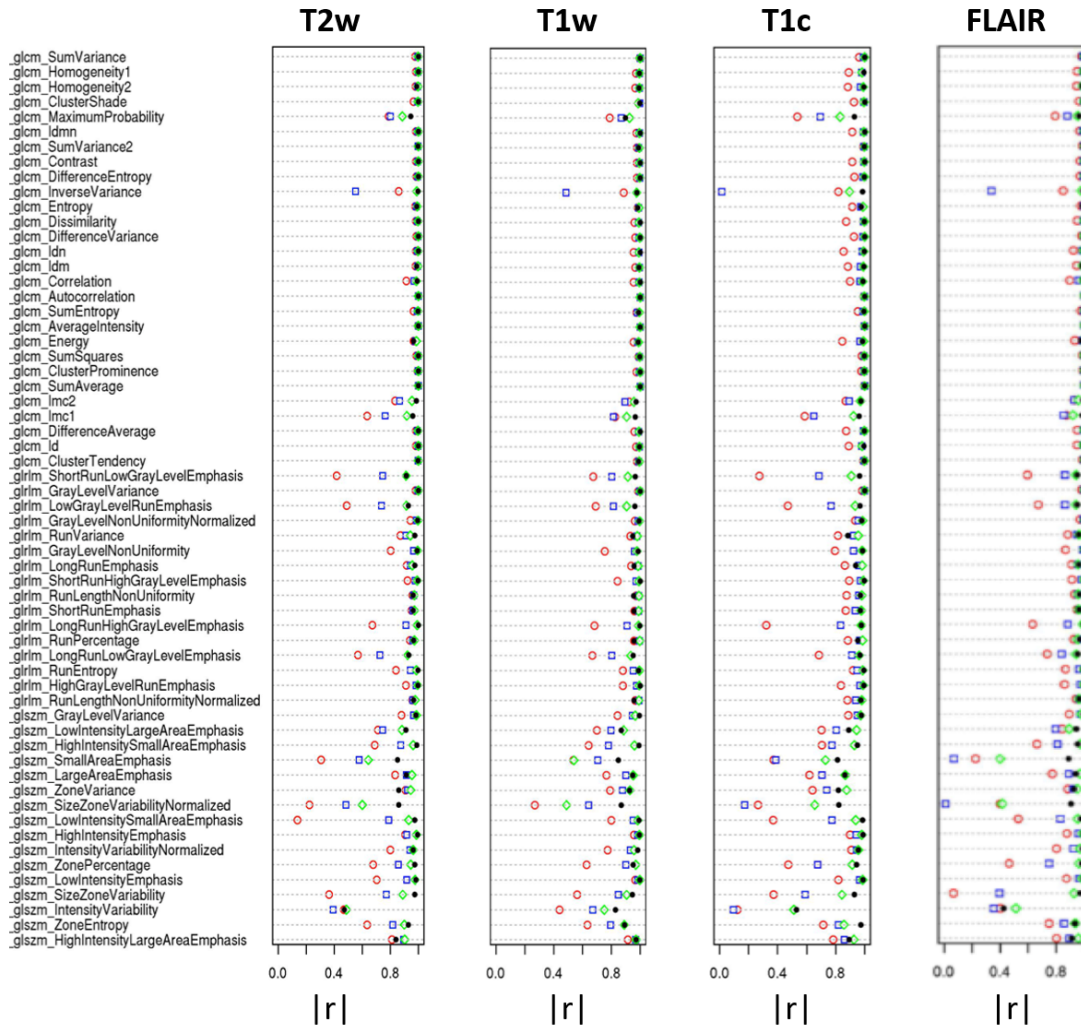


FIGURE 4.7 – Coefficients de corrélation de Spearman (en valeur absolue $|r|$) pour chacun des 240 indices radiomiques de la discrétisation $d_{rel,fNN}$ avec $NN = 64$ versus : $NN = 8$ en rouge, $LN = 16$ en bleu, $LN = 32$ en vert et $LN = 128$ en noir.

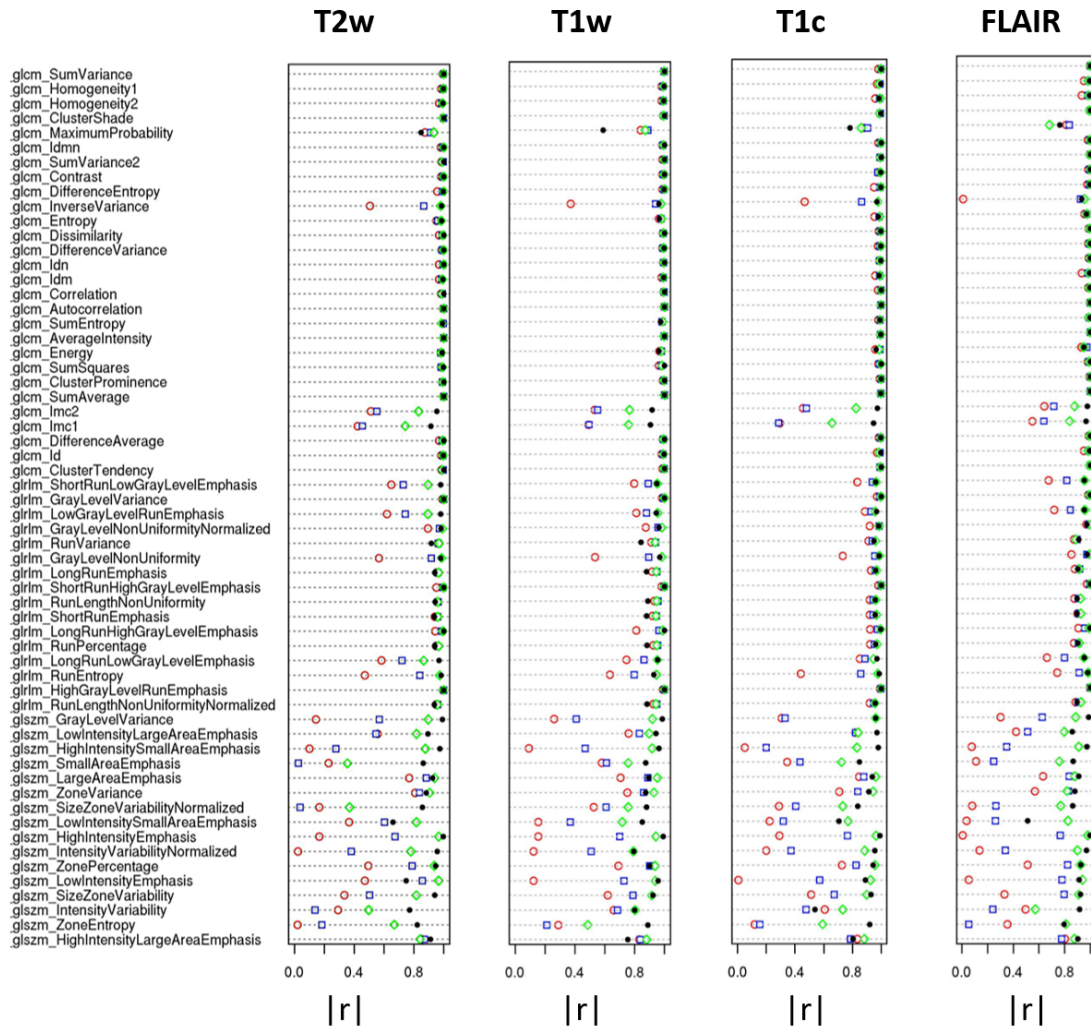


FIGURE 4.8 – Coefficients de corrélation de Spearman (en valeur absolue $|r|$) pour chacun des 240 indices radiomiques de la discrétisation $d_{abs, f_{NN}}$ avec $NN = 64$ versus : $NN = 8$ en rouge, $LN = 16$ en bleu, $LN = 32$ en vert et $LN = 128$ en noir.

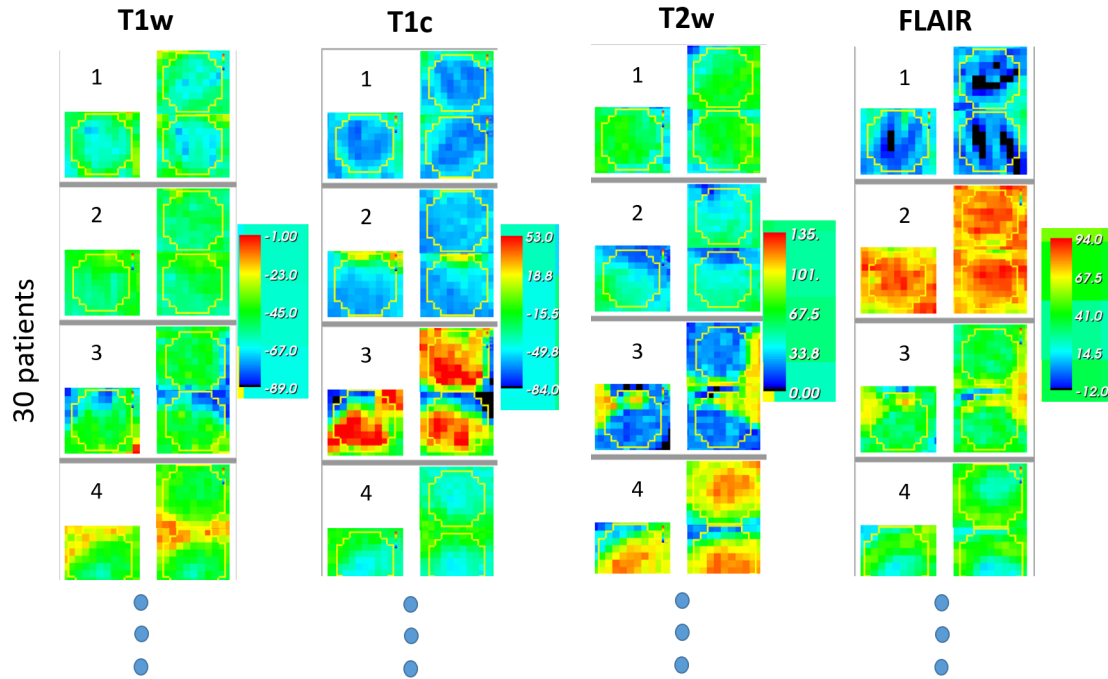


FIGURE 4.9 – Exemple d’images fournies à chaque expert pour étiqueter la région (délimitée par le cercle jaune) comme hétérogène ou homogène. Les coupes centrales de chacun des trois axes sont données en échelle de couleur.

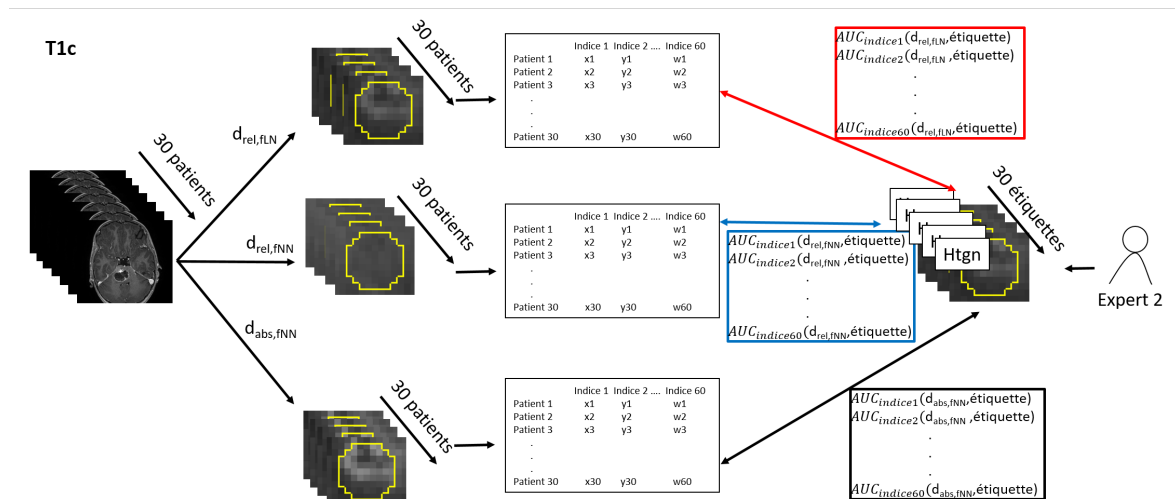


FIGURE 4.10 – Schéma des AUC calculées pour vérifier la cohérence des indices radiomique avec l’inspection visuelle. Sur le schéma, le processus pour les images T1c et l’expert 2 sont présentés, mais la même approche est considérée pour les 4 modalités d’IRM et les 3 experts.

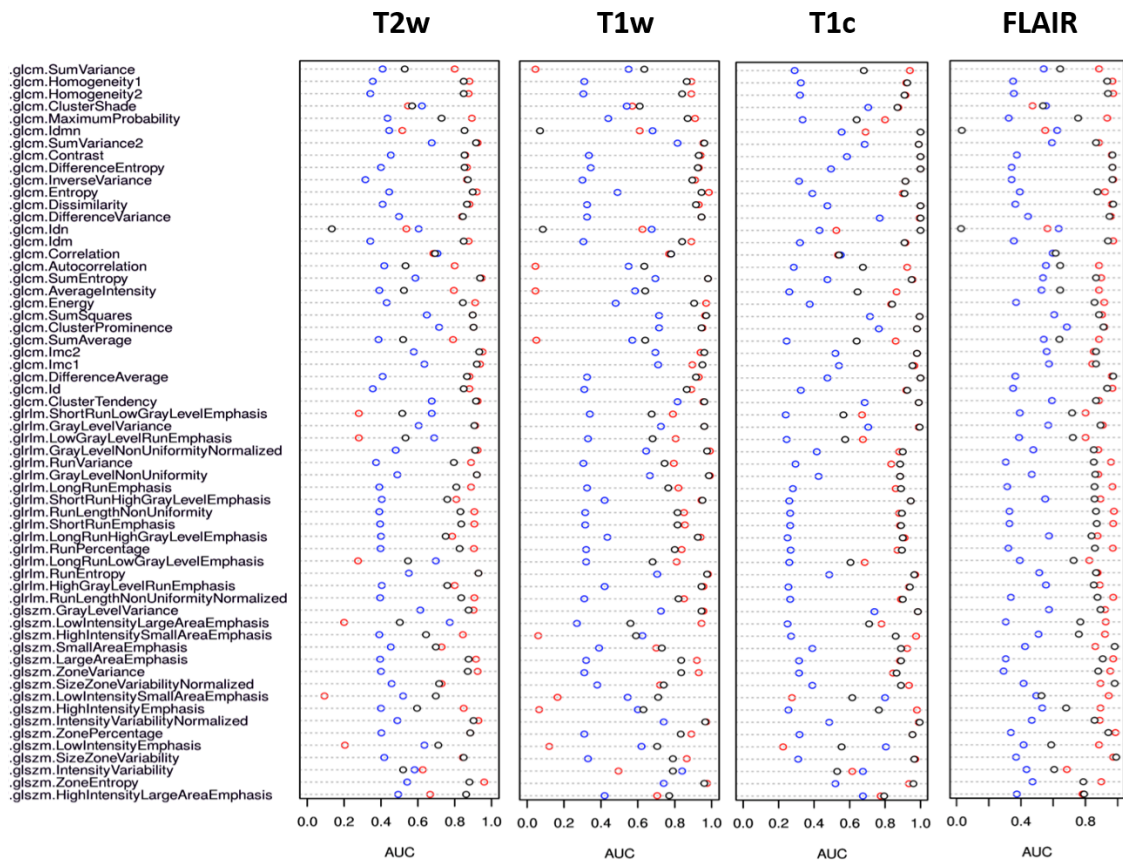


FIGURE 4.11 – AUC obtenues pour chacun des 240 indices radiomiques en fonction des étiquettes de l'expert 1 avec les trois types de discrétisation $d_{rel, fLN}$ en rouge, $d_{rel, fNN}$ en bleu et $d_{abs, fNN}$ en noir.

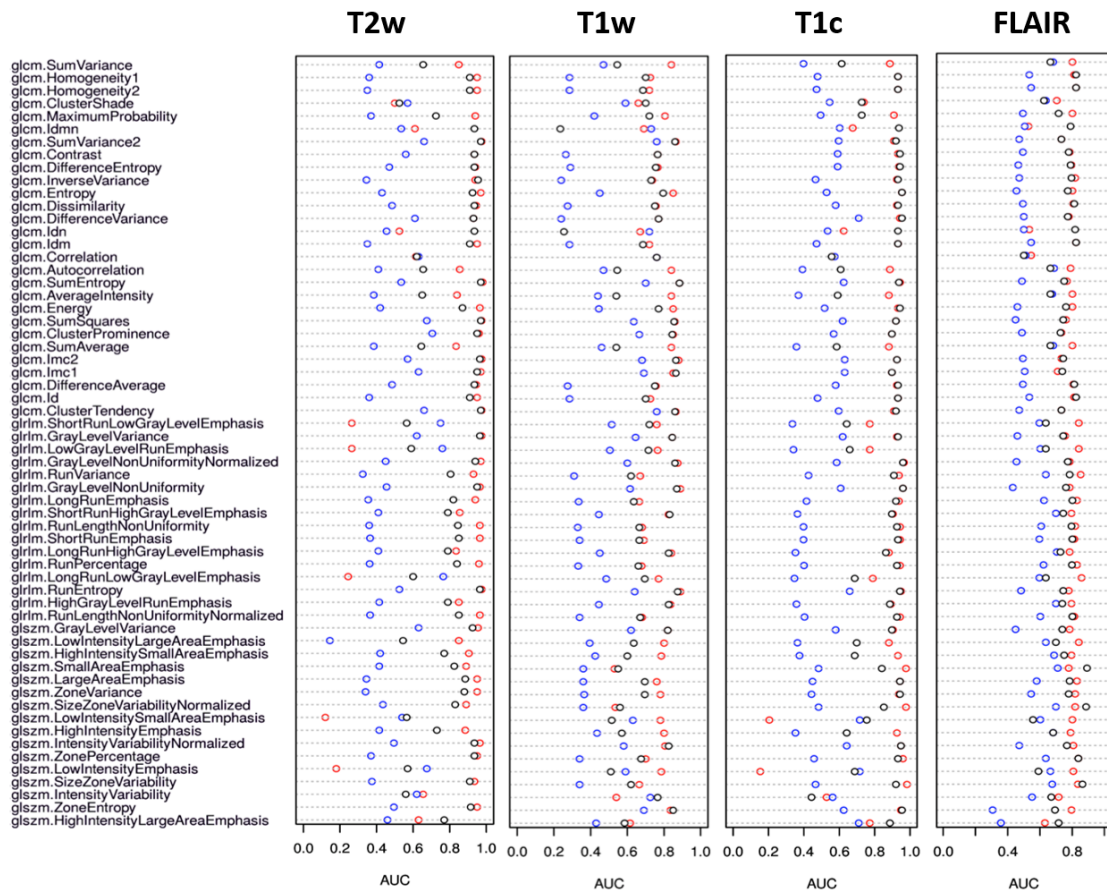


FIGURE 4.12 – AUC obtenues pour chacun des 240 indices radiomiques en fonction des étiquettes de l'expert 2 avec les trois types de discrétisation $d_{rel, f_{LN}}$ en rouge, $d_{rel, f_{NN}}$ en bleu et $d_{abs, f_{NN}}$ en noir.

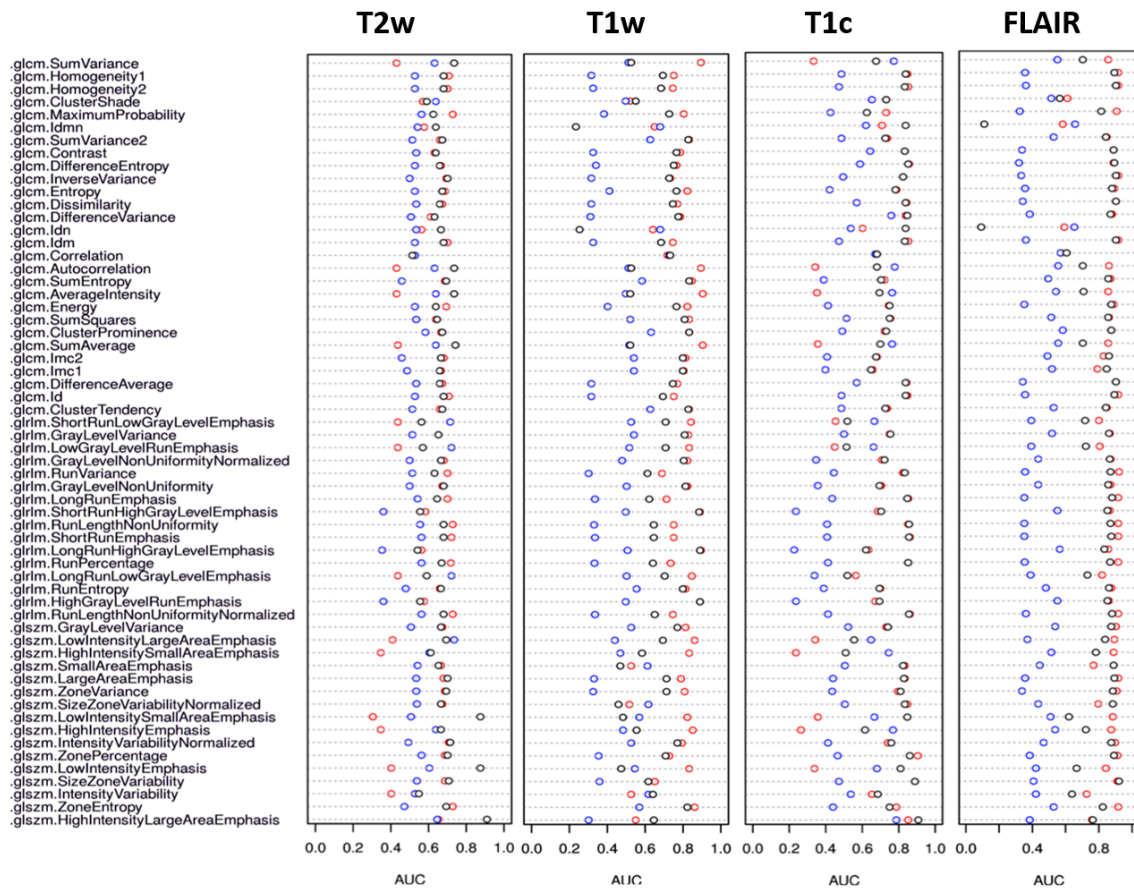


FIGURE 4.13 – AUC obtenues pour chacun des 240 indices radiomiques en fonction des étiquettes de l'expert 3 avec les trois types de discrétisation $d_{rel, fLN}$ en rouge, $d_{rel, fNN}$ en bleu et $d_{abs, fNN}$ en noir.

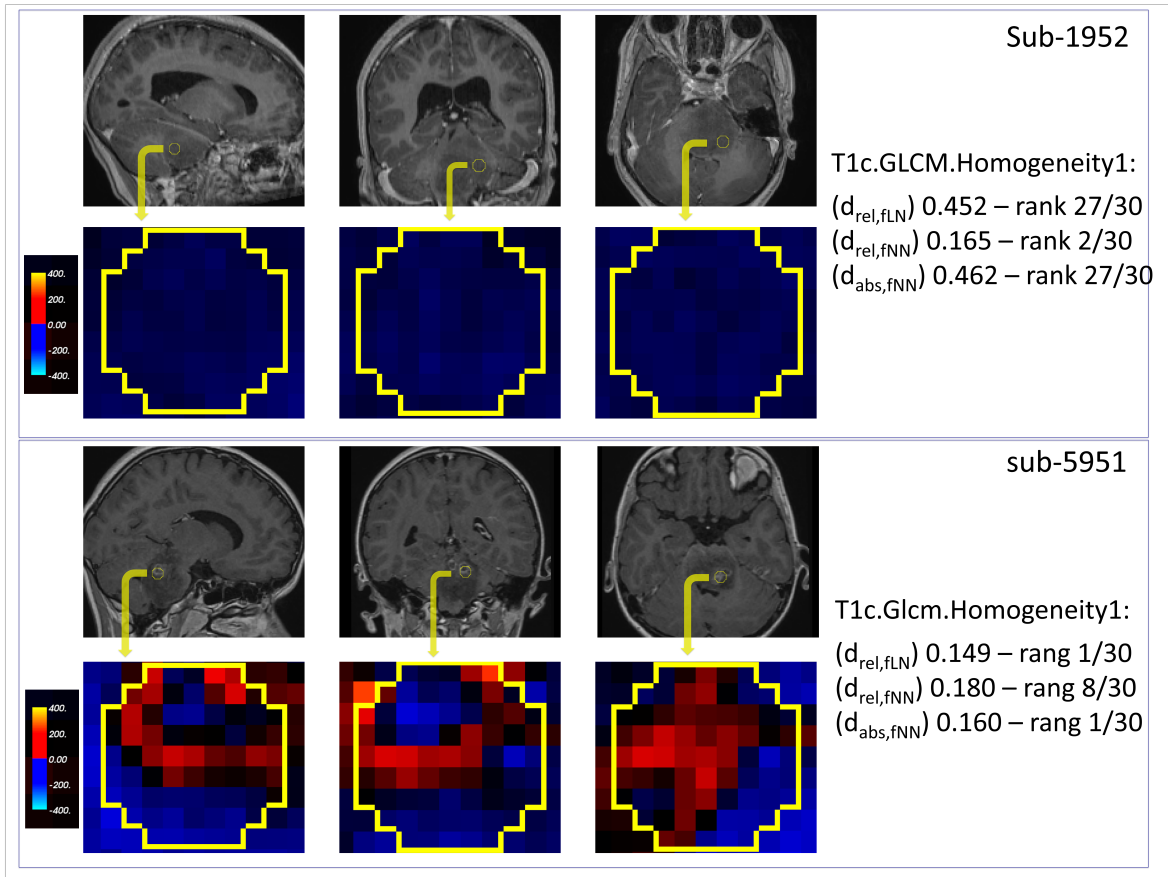


FIGURE 4.14 – Les coupes centrales sagittale, coronale et axiale des $VOI_{biopsie}$ des patients sub-1952 (en haut) et sub-5951 (en bas), visite de diagnostic, avec les régions correspondantes zoomées et colorées en échelle de couleur. La même échelle a été utilisée pour chaque image. Les valeurs de $T1c\ GLCM\ Homogeneity1$ pour les trois méthodes de discrétisation sont présentées à droite, avec les rangs des patients. Plus la valeur d'homogénéité est élevée, plus la tumeur est homogène.

Sur la figure 4.14, les trois coupes centrales des images $T1c$ de deux patients sont présentées : le patient sub-1952 de rang 27 pour $d_{rel, fLN}$ et pour $d_{abs, fNN}$, et de rang 2 pour $d_{rel, fNN}$ par rapport à l'indice $T1c\ GLCM\ Homogeneity1$; et le patient sub-5951 de rang 1 pour $d_{rel, fLN}$ et pour $d_{abs, fNN}$, et de rang 8 pour $d_{rel, fNN}$ par rapport à l'indice $T1c\ GLCM\ Homogeneity1$. L'image de la $VOI_{biopsie}$ du premier patient a été classée comme homogène par les 3 experts, alors que l'image de la $VOI_{biopsie}$ du second patient a été classée comme hétérogène par les 3 experts.

Ainsi, les méthodes de discrétisation $d_{rel, fLN}$ et $d_{abs, fNN}$ sont plus cohérentes avec l'inspection visuelle que la méthode $d_{rel, fNN}$.

4.3 Impact du volume de la région d'intérêt sur les indices radiomiques

Pour évaluer l'impact du volume de la région dans laquelle les indices radiomiques sont calculés sur les valeurs des indices, une région de la MB non atteinte par la tumeur a été analysée. Nous avons décidé de choisir cette région de MB de façon à ce qu'elle représente la même zone pour tous les patients. Deux régions sphériques concentriques de $3mm$ et $5mm$ de rayon ont été dessinées dans la MB du lobe frontal droit du cerveau de référence MNI de voxel isotrope $1\ mm^3$. Ensuite, avec l'aide des transformations diffeomorphiques qui ont été calculées entre chaque patient et le repère MNI dans l'étape de standardisation des intensités, présentée dans le chapitre 3, ces deux régions sphériques ont été reportées sur les images de chaque patient. Elles sont notées VOI_{MNI3} et VOI_{MNI5} . Les volumes de ces régions d'intérêt dans les repères patients, varient entre 39 et $424\ mm^3$ selon le rayon original de la sphère et la taille du cerveau de chaque patient.

Additionnellement, deux régions sphériques de rayon 4,5 mm ont été tracées directement dans le repère de chaque patient : une région concentrique avec VOI_{MNI3} et VOI_{MNI5} du côté droit du cerveau, VOI_D , et une autre dans le côté gauche du cerveau en miroir des autres VOIs, VOI_G . Toutes ces régions sont illustrées sur la figure 4.15.

Les 60 indices radiomiques du tableau 4.1 ont été calculés dans chacune des quatre régions d'intérêt dans les 120 (4 x 30) images de diagnostic avec la discrétisation $d_{rel, fLN}$ (LN = 2). Comme pour l'étude des méthodes de discrétisation, toutes ces images standardisées ont été ré-échantillonnées en prenant un voxel de 1 mm x 1 mm x 1 mm.

Deux études différentes ont été réalisées :

1. La corrélation entre chaque indice radiomique issu des régions VOI_{MNI3} et VOI_{MNI5} et les volumes a été quantifiée;
2. Les indices issus des VOI_{MNI5} et VOI_G , des VOI_{MNI5} et VOI_R , et des VOI_D et VOI_G ont été comparés pour tester s'ils présentaient des différences statistiquement significatives.

4.3.1 Corrélation entre indices radiomiques et volume de la région d'intérêt

La corrélation de Spearman (r) a été calculée entre chaque indice issu des régions VOI_{MNI3} et VOI_{MNI5} et les volumes des régions pour les 30 sujets. Ainsi, 46 des 240 indices présentent une forte corrélation avec le volume ($|r| > 0,7$). Le tableau 4.3 présente les valeurs de r de chaque indice. Parmi les 60 indices, neuf sont systématiquement corrélés au volume pour toutes les modalités : *GLCM IMC1*, *GLCM IMC2*, *GLRLM GrayLevelNonUniformity*, *GLRLM RunLengthNonUniformity*, *GLSZMHighIntensityLargeAreaEmphasis*, *GLSZM IntensityVariability*, *GLSZM LargeAreaEmphasis*, *GLSZM SizeZoneVariability*, *GLSZM ZoneVariance*. Ces 9 paramètres sont en gras dans le tableau 4.3.

La figure 4.16 illustre trois types d'indices, qui présentent des relations différentes avec le volume : *FLAIR GLCM RunLengthNonUniformity* ($r = 0,99$), *T2w GLCM Entropy* ($r = 0,82$, sa corrélation décroît avec l'augmentation de volume), et *T1w Homogeneity1* ($r = 0,03$).

4.3.2 Impact du volume d'intérêt sur les indices radiomiques de régions similaires

Cette deuxième étude s'appuie sur le principe que les régions VOI_{MNI5} , VOI_G et VOI_D sont similaires et devraient présenter des indices radiomiques équivalents. Si des différences statistiques significatives sont trouvées entre les indices radiomiques calculés dans VOI_{MNI5} et VOI_G , et entre les indices calculés dans VOI_{MNI5} et VOI_R , et pas entre les indices calculés dans VOI_D et VOI_G , ces différences sont probablement une conséquence de l'effet volume et éventuelle déformation du volume liée aux recalages de VOI_{MNI5} (entre 120 mm^3 et 424 mm^3).

Le test de la somme des rangs de Wilcoxon suivi d'une correction pour les tests-multiples de Benjamini-Hochberg a comparé les indices issus des VOI_{MNI5} et VOI_G , des VOI_{MNI5} et VOI_R , et des VOI_D et VOI_G . Les différences sont considérées comme significatives pour les valeurs-p inférieures à 0,05. Aucune différence statistiquement significative n'a été trouvée entre les indices calculés dans les régions VOI_D et VOI_G , toutes les valeurs-p étant supérieures à 0,84. En revanche, 81 indices sont statistiquement significativement différents quand calculés dans les régions VOI_{MNI5} et VOI_D et 41 quand calculés dans les régions VOI_{MNI5} et VOI_G . Le tableau 4.4 liste les indices qui présentent des différences statistiquement significatives. Des indices fortement corrélés au volume figurent en gras. Parmi les 81 (respectivement 41) indices, 34 (respectivement 29) sont fortement corrélés au volume. Trente-huit parmi les 41 indices sont inclus dans les 81 indices, les 3 autres sont soulignés en rouge.

La figure 4.17 présente les graphes des trois indices présentés sur la figure 4.16 calculés dans les régions VOI_D et VOI_G .

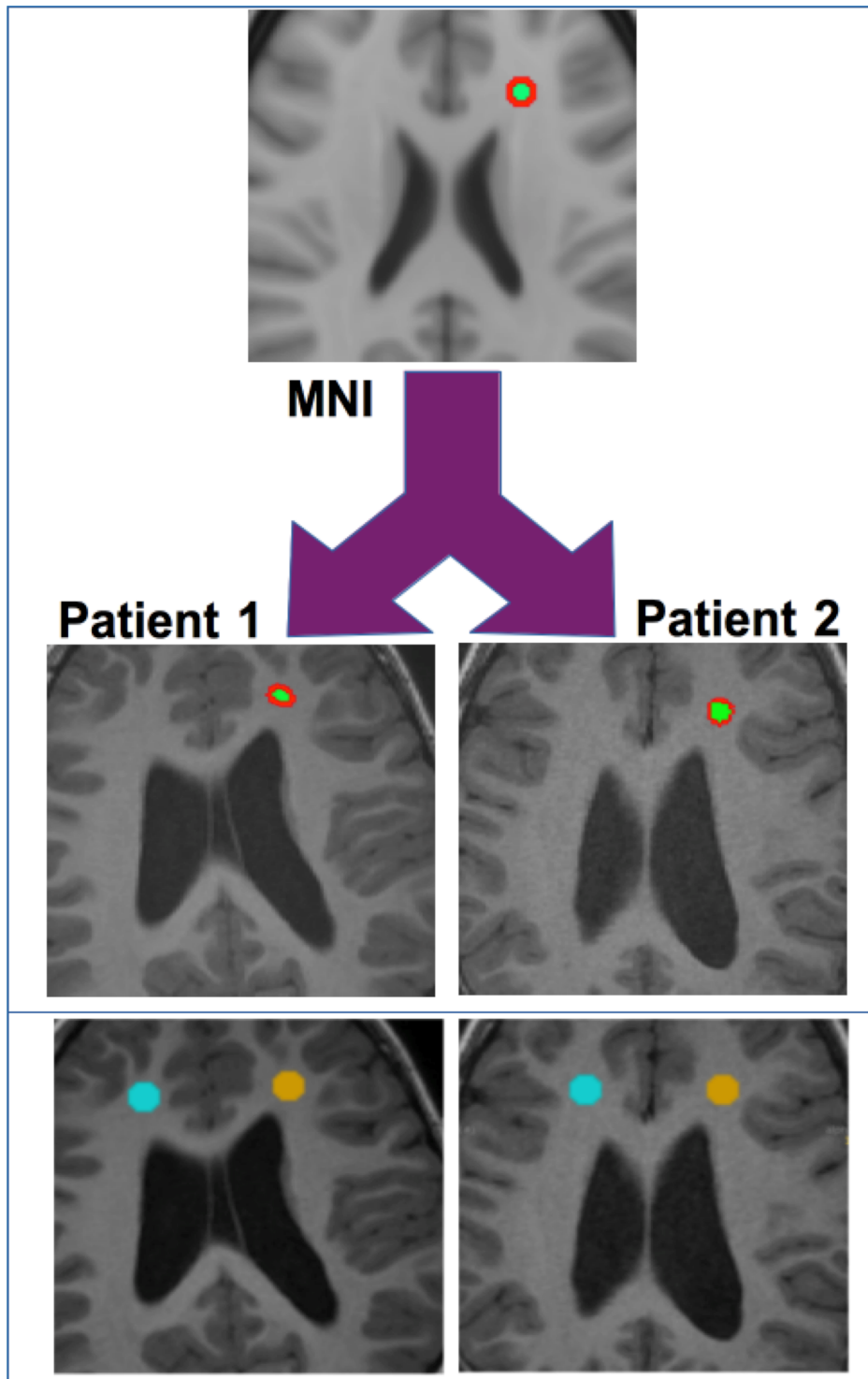


FIGURE 4.15 – Les sphères concentriques de rayon 3 mm (en vert) et 5 mm (en rouge) ont été d’abord dessinées dans le cerveau MNI (en haut). Ces régions ont été repositionnées (au milieu) dans le repère de chaque patient (exemple pour 2 patients différents), VOI_{MNI3} (en vert) et VOI_{MNI5} (en rouge). Les sphères de 4.5 mm de rayon ont été tracés dans le repère de chaque patient : VOI_D en jaune et VOI_G en bleu cyan (en bas).

TABLEAU 4.3 – Valeurs de corrélation de Spearman (r) entre les valeurs de chaque indice radiomique et le volume des VOIs. Les indices qui présentent une forte corrélation dans les 4 modalités sont en gras.

	T1w r	T2w r	T1c r	FLAIR r
glcm.SumVariance	0.430134	0.421603	0.419991	0.288080
glcm.Homogeneity1	0.027816	0.048685	0.066469	0.122185
glcm.Homogeneity2	0.028538	0.050324	0.066358	0.123658
glcm.ClusterShade	-0.101177	0.007253	0.125325	0.148472
glcm.MaximumProbability	-0.356523	-0.740778	-0.328512	-0.642102
glcm.Idmn	0.635210	0.709766	0.692037	0.651606
glcm.SumVariance2	0.186431	0.228030	0.178762	0.093007
glcm.Contrast	0.037848	0.054187	0.046267	-0.062635
glcm.DifferenceEntropy	0.206466	0.424437	0.218721	0.244397
glcm.InverseVariance	0.039348	-0.045517	0.033235	0.141470
glcm.Entropy	0.491962	0.817668	0.522029	0.732414
glcm.Dissimilarity	0.014228	0.029455	0.003724	-0.110430
glcm.DifferenceVariance	0.073055	0.125436	0.115682	0.006280
glcm.Idn	0.591388	0.713740	0.702152	0.619538
glcm.Idm	0.028538	0.050324	0.066358	0.123658
glcm.Correlation	0.405624	0.377781	0.307004	0.310533
glcm.Autocorrelation	0.479513	0.474095	0.466119	0.340461
glcm.SumEntropy	0.460978	0.619038	0.462007	0.570381
glcm.AverageIntensity	0.476290	0.463479	0.451253	0.348659
glcm.Energy	-0.407431	-0.747503	-0.431634	-0.648743
glcm.SumSquares	0.142331	0.143443	0.109680	0.035680
glcm.ClusterProminence	0.207189	0.255929	0.223000	0.114765
glcm.SumAverage	0.476290	0.463479	0.451253	0.348659
glcm.lmc2	-0.798244	-0.862212	-0.761814	-0.775485
glcm.lmc1	0.886860	0.923624	0.842510	0.878691
glcm.DifferenceAverage	0.014228	0.029455	0.003724	-0.110430
glcm.Id	0.027816	0.048685	0.066469	0.122185
glcm.ClusterTendency	0.186431	0.228030	0.178762	0.093007
glrlm.ShortRunLowGrayLevelEmphasis	-0.660998	-0.663888	-0.591750	-0.596029
glrlm.GrayLevelVariance	0.143693	0.161366	0.159143	0.021341
glrlm.LowGrayLevelRunEmphasis	-0.630931	-0.649132	-0.566490	-0.579968
glrlm.GrayLevelNonUniformityNormalized	-0.226140	-0.285913	-0.168952	-0.121045
glrlm.RunVariance	0.517027	0.582107	0.576077	0.462701
glrlm.GrayLevelNonUniformity	0.963556	0.967668	0.959665	0.942437
glrlm.LongRunEmphasis	0.480708	0.559210	0.496520	0.455921
glrlm.ShortRunHighGrayLevelEmphasis	0.402012	0.433996	0.394982	0.251150
glrlm.RunLengthNonUniformity	0.980201	0.993984	0.985258	0.991733
glrlm.ShortRunEmphasis	-0.429411	-0.540369	-0.451308	-0.442055
glrlm.LongRunHighGrayLevelEmphasis	0.602420	0.542009	0.600170	0.439887
glrlm.RunPercentage	-0.475234	-0.571996	-0.517007	-0.470655
glrlm.LongRunLowGrayLevelEmphasis	-0.494741	-0.597585	-0.451503	-0.480097
glrlm.RunEntropy	0.663416	0.645992	0.664055	0.531922
glrlm.HighGrayLevelRunEmphasis	0.440860	0.440499	0.441082	0.287302
glrlm.RunLengthNonUniformityNormalized	-0.447223	-0.569908	-0.491657	-0.452698
glszm.GrayLevelVariance	0.516388	0.493268	0.530171	0.335654
glszm.LowIntensityLargeAreaEmphasis	0.507440	0.369083	0.492740	0.421325
glszm.HighIntensitySmallAreaEmphasis	0.472733	0.479208	0.464869	0.268768
glszm.SmallAreaEmphasis	0.276715	0.266294	0.401039	-0.040348
glszm.LargeAreaEmphasis	0.740444	0.749976	0.792909	0.733358
glszm.ZoneVariance	0.790963	0.806608	0.841732	0.764148
glszm.SizeZoneVariabilityNormalized	0.201020	0.224556	0.237651	-0.130410
glszm.LowIntensitySmallAreaEmphasis	-0.442916	-0.553874	-0.138302	-0.503022
glszm.HighIntensityEmphasis	0.501966	0.481347	0.449141	0.319036
glszm.IntensityVariabilityNormalized	-0.201476	-0.517625	-0.344616	-0.328058
glszm.ZonePercentage	-0.533569	-0.564148	-0.616420	-0.580968
glszm.LowIntensityEmphasis	-0.493546	-0.589082	-0.472233	-0.452725
glszm.SizeZoneVariability	0.700083	0.839176	0.761696	0.721965
glszm.IntensityVariability	0.854475	0.951356	0.883352	0.909145
glszm.ZoneEntropy	0.642352	0.870215	0.724778	0.842177
glszm.HighIntensityLargeAreaEmphasis	0.864324	0.912759	0.901171	0.898948

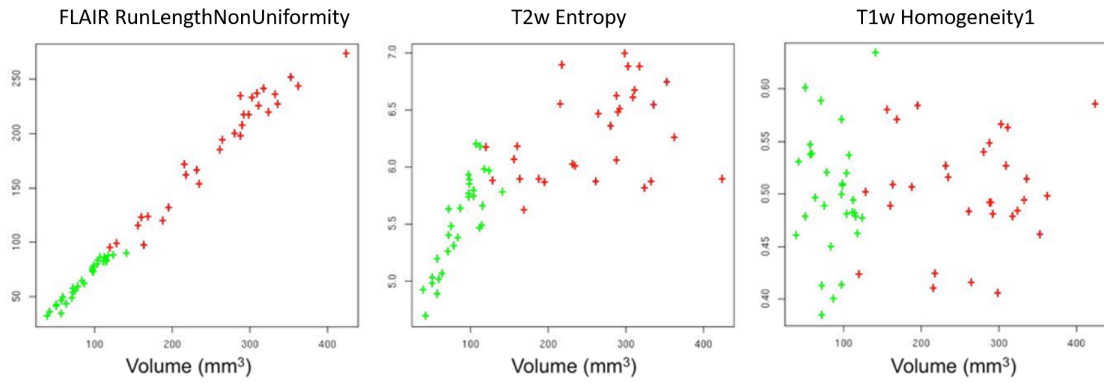


FIGURE 4.16 – Graphes montrant les valeurs de trois indices différents par rapport au volume. Les valeurs extraites des VOI_{MNI3} sont en vert et celles extraites des VOI_{MNI5} sont en rouge.

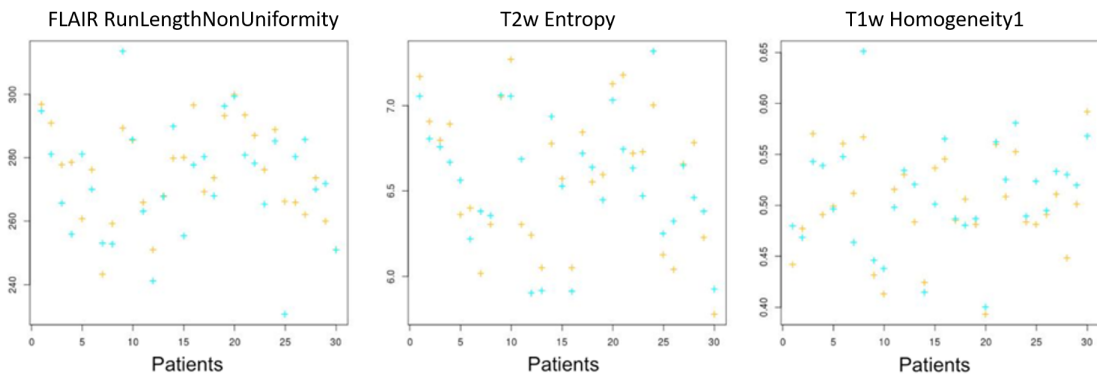


FIGURE 4.17 – Valeurs de trois indices de texture pour chacun des 30 patients. Les valeurs extraites de VOI_D sont en jaune et celles extraites de VOI_G sont en bleu cyan.

TABLEAU 4.4 – Liste des 41 indices calculés dans les régions VOI_{MNI5} et VOI_G et des 81 indices statistiquement significativement différents quand calculés dans VOI_{MNI5} et VOI_D . Les indices fortement corrélés au volume figurent en gras. Les trois indices compris dans la série des 41 et pas dans la série des 81 sont soulignés en rouge.

	Matrice de texture	Indice de texture (VOI_{MNI5} vs. VOI_G)	Indice de texture (VOI_{MNI5} vs. ROI_D)
	GLCM	Correlation, Idmn, Idn	Correlation, Idmn, Idn, ClusterShade, ClusterProminence, AverageIntensity, Autocorrelation, SumVariance, SumAverage
T1w	GLRLM	GrayLevelNonUniformity , LongRunHighGrayLevelEmphasis, LowGrayLevelRunEmphasis, RunLengthNonUniformity , ShortRunLowGrayLevelEmphasis	GrayLevelNonUniformity , LongRunHighGrayLevelEmphasis, LowGrayLevelRunEmphasis, RunLengthNonUniformity , ShortRunLowGrayLevelEmphasis, HighGrayLevelRunEmphasis, LongRunLowGrayLevelEmphasis, RunEntropy, ShortRunHighGrayLevelEmphasis
	GLSZM	HighIntensityLargeAreaEmphasis , LargeAreaEmphasis , ZoneVariance	HighIntensityLargeAreaEmphasis , LargeAreaEmphasis , ZoneVariance , GrayLevelVariance, HighIntensityEmphasis, HighIntensitySmallAreaEmphasis, LowIntensityEmphasis, LowIntensitySmallAreaEmphasis, SizeZoneVariability, ZoneEntropy
	GLCM	<u>Entropy</u> , <u>Imc1</u> , <u>Imc2</u>	Entropy , Imc1 , Idn , Idnm , MaximumProbability
T2w	GLRLM	GrayLevelNonUniformity , RunLengthNonUniformity	GrayLevelNonUniformity , RunLengthNonUniformity , RunEntropy
	GLSZM	HighIntensityLargeAreaEmphasis , IntensityVariability , LargeAreaEmphasis , SizeZoneVariability , ZoneEntropy , ZoneVariance	HighIntensityLargeAreaEmphasis , IntensityVariability , SizeZoneVariability , ZoneEntropy , ZoneVariance , GrayLevelVariance , IntensityVariabilityNormalized
	GLCM	Correlation, Idmn, Idn	Correlation, Idmn, Idn , ClusterShade, ClusterProminence, AverageIntensity, Autocorrelation, SumAverage, SumVariance
T1c	GLRLM	GrayLevelNonUniformity , LongRunHighGrayLevelEmphasis, LowGrayLevelRunEmphasis, RunLengthNonUniformity	GrayLevelNonUniformity , LongRunHighGrayLevelEmphasis, LowGrayLevelRunEmphasis, RunLengthNonUniformity , ShortRunLowGrayLevelEmphasis, ShortRunHighGrayLevelEmphasis, RunEntropy, LongRunLowGrayLevelEmphasis, HighGrayLevelRunEmphasis
	GLSZM	GrayLevelVariance, HighIntensityLargeAreaEmphasis , HighIntensitySmallAreaEmphasis, LargeAreaEmphasis , SizeZoneVariability , ZoneVariance	GrayLevelVariance, HighIntensityLargeAreaEmphasis , HighIntensitySmallAreaEmphasis, LargeAreaEmphasis , SizeZoneVariability , ZoneVariance , ZoneEntropy , LowIntensityEmphasis, IntensityVariabilityNormalized, HighIntensityEmphasis
	GLCM	<u>Imc1</u>	Idn , Idmn
FLAIR	GLRLM	GrayLevelNonUniformity , RunLengthNonUniformity	GrayLevelNonUniformity , RunLengthNonUniformity , RunEntropy
	GLSZM	HighIntensityLargeAreaEmphasis , IntensityVariability , SizeZoneVariability	HighIntensityLargeAreaEmphasis , IntensityVariability , SizeZoneVariability , ZoneEntropy , GrayLevelVariance

4.4 Proposition d'une méthode de calcul local d'indices radiomiques

Comme plusieurs indices radiomiques se sont avérés très impactés par le volume, nous avons développé une méthode qui permettrait d'analyser l'hétérogénéité tumorale par tous les indices de texture sans avoir à subir l'impact du volume. J'ai fait l'hypothèse que parmi ces indices il y en aurait certains qui refléteraient des informations importantes pour le **GITC**, une fois libérés de leur corrélation avec le volume. C'est pour cela que j'ai proposé la méthode de calcul d'indices locaux.

4.4.1 Calcul d'indices radiomiques locaux dans la tumeur

Comme déjà suggéré par **LARROZA et collab. [2016]**, l'utilisation de régions d'intérêt de volume constant est une option pour supprimer l'effet volume dans les valeurs des indices de texture. Cependant, le choix du volume constant à être utilisé dans toutes les tumeurs n'est pas évident, car les tailles des tumeurs varient beaucoup entre les patients. L'utilisation d'un même volume pour tous les patients aurait pour conséquence des régions d'intérêt qui ne couvrent pas un pourcentage suffisant de la tumeur ou qui la dépassent largement.

Pour analyser les tumeurs dans les images **IRM**, des régions sphériques ont été tracées avec **LIFEx (NIOCHE et collab. [2018])** dans la région tumorale, incluant le maximum possible de la tumeur, VOI_{Tumeur} . Ensuite, les indices radiomiques sont calculés dans une sphère de rayon constant 5mm, VOI_c , dont le centre se déplace avec un pas d'un voxel dans la VOI_{Tumeur} et, pour chaque position, les indices radiomiques sont calculés dans la VOI_c . Chaque voxel de la VOI_{Tumeur} est donc caractérisé par les valeurs des indices calculés dans la VOI_c centrée en ce voxel. Par conséquent, une carte de distribution de chaque indice est créée dans la VOI_{Tumeur} sauf sur les bords. Finalement, la moyenne de chaque carte d'indice a été estimée et utilisée comme indice radiomique local caractérisant la VOI_{Tumeur} . La figure 4.18 montre les étapes de calcul des indices radiomiques locaux ainsi que le calcul des indices radiomiques globaux, qui sont plus fréquemment utilisés dans la littérature mais qui sont impactés par les volumes différents des régions tumorales.

Des exemples de cartes d'indices radiomiques sont présentés sur la figure 4.19 avec la carte de valeurs d'intensités pour illustrer les différentes informations obtenues pour chaque indice.

4.4.2 Comparaison entre les indices radiomiques locaux et globaux pour caractériser les régions tumorales

Les coefficients de corrélation de Spearman entre les indices radiomiques et le volume ont été calculés pour comparer les deux approches (locale et globale). Aucun indice de texture local ne présente une corrélation avec le volume supérieure à 0,35, alors que 12 indices de texture globaux

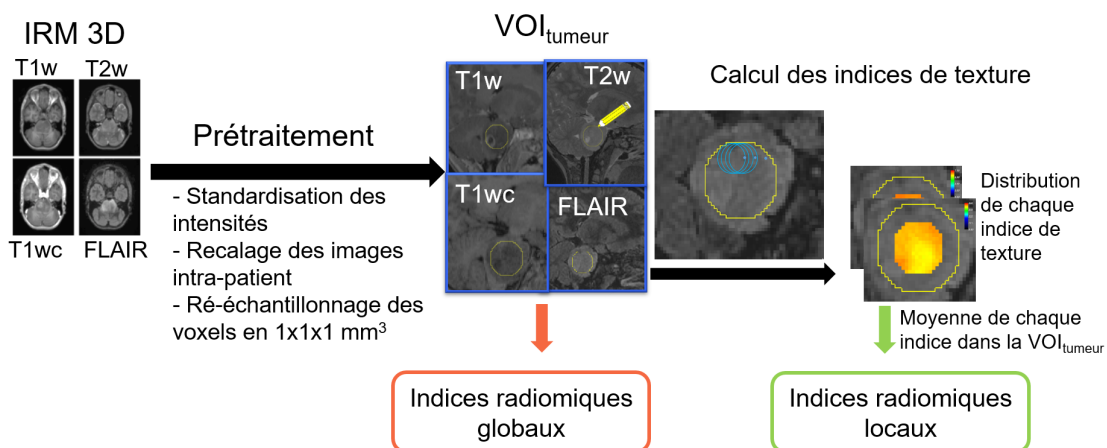


FIGURE 4.18 – Calcul des indices radiomiques globaux et locaux dans la tumeur.

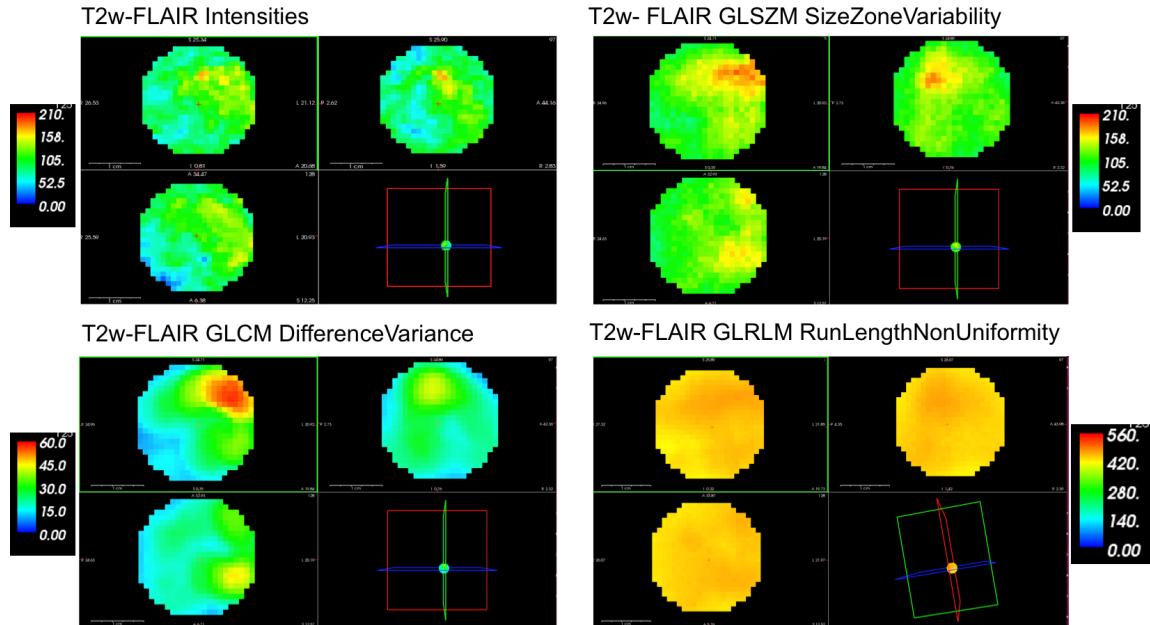


FIGURE 4.19 – Patient PREBIOMEDE sub-3593, session de diagnostic : coupes coronale, sagittale and axiale des intensités T2w-FLAIR après standardisation (en haut à gauche) et de trois indices radiomiques dans VOI_{tumeur} .

sont fortement corrélés au volume ($|r| > 0,7$). Le graphe de la figure 4.20 montre les valeurs de r pour chaque indice radiomique calculé dans la VOI_{tumeur} avec les deux approches.

Pour aller plus loin, la figure 4.21 montre les coefficients de corrélation de Spearman entre chaque indice calculé globalement versus localement sur les 30 patients. On observe que les versions globale et locale de la majorité (199 sur 240) des indices sont corrélés ($|r| > 0,7$).

La figure 4.22 compare les valeurs calculées globalement et localement des trois indices présentés sur la figure 4.16 et qui présentent des comportements différents avec le volume. Il est possible de noter que les valeurs locales et globales de l'indice *GLCM Homogeneity*, qui n'est pas corrélé au volume, sont très similaires alors que pour les indices *GLRLM RunLengthNonUniformity* et *GLCM Entropy*, elles sont différentes. Pour le *GLRLM RunLengthNonUniformity*, la corrélation est très faible.

L'indice *GLRLM RunLengthNonUniformity* représente la similarité des longueurs des séries des voxels d'intensités similaires. Une valeur faible de l'indice reflète des longueurs plus homogènes dans l'image. La figure 4.23 présente les images *FLAIR* de deux patients, les volumes de VOI_{tumeur} et les valeurs de l'indice *GLRLM RunLengthNonUniformity* calculées globalement et localement. L'exemple montre que les rangs des patients selon cet indice est plus cohérent quand ils sont calculés localement. En effet, le rang plus faible de *GLRLM RunLengthNonUniformity* local du patient sub-8044 reflète bien une tumeur plus homogène que celle du patient sub-6046 et qui présente plusieurs séries longues d'intensités similaires.

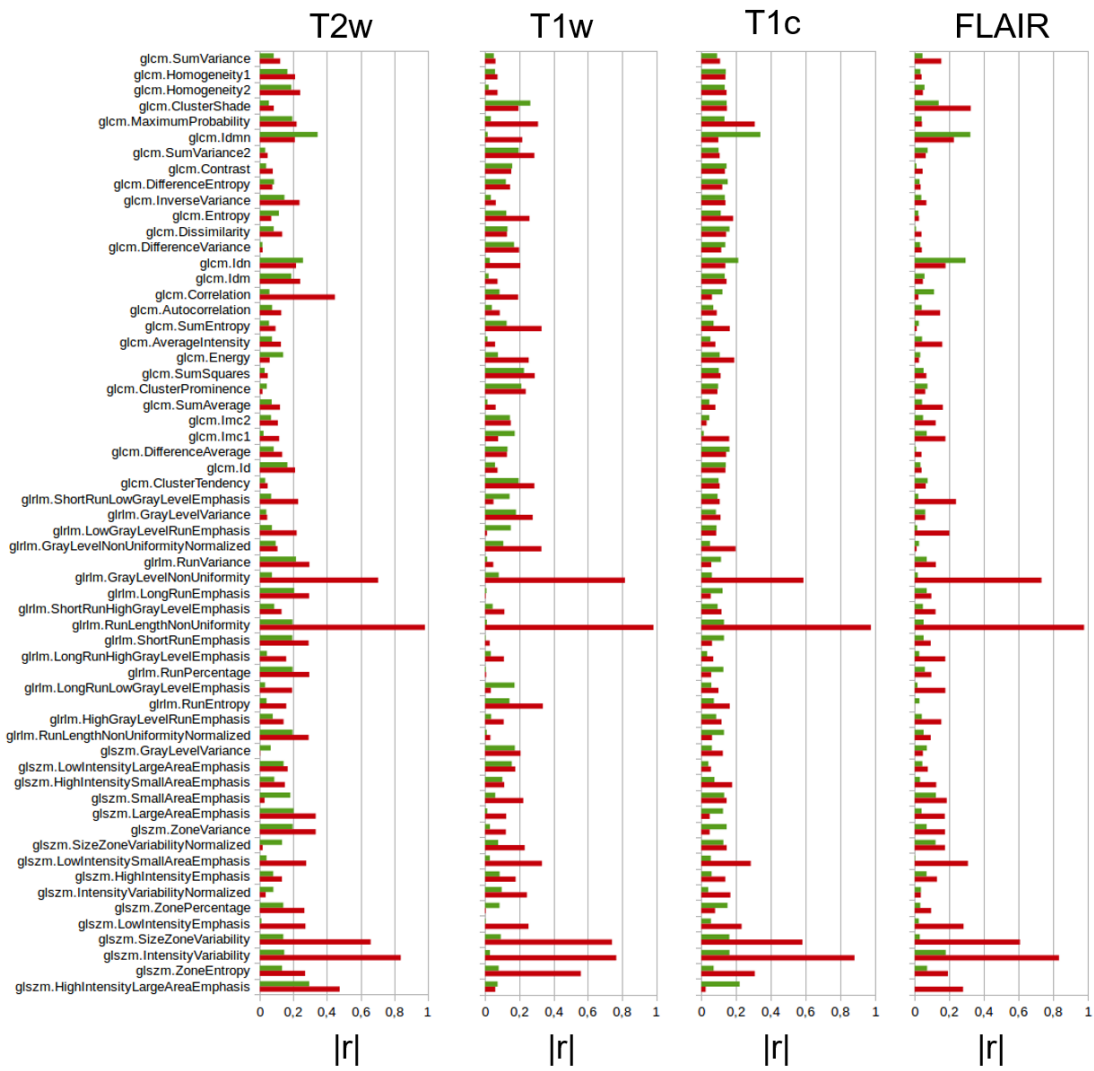


FIGURE 4.20 – Valeurs absolues des coefficients de corrélation de Spearman ($|r|$) entre chaque indice radiomique et le volume de VOI_{tumeur} correspondant. Les valeurs correspondant aux indices locaux sont en vert et celles correspondant aux indices globaux en rouge.

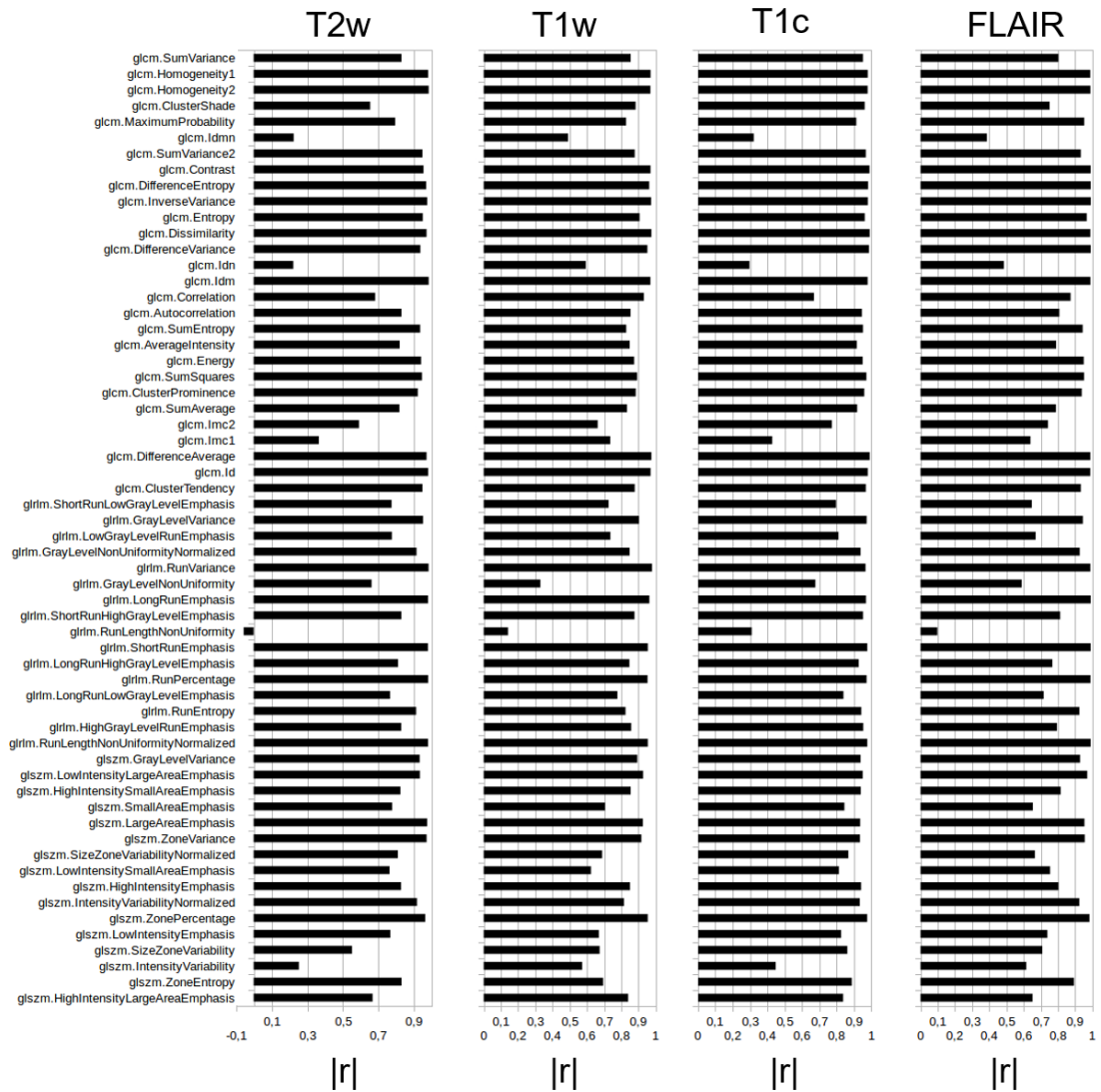


FIGURE 4.21 – Valeurs des coefficients de corrélation de Spearman (r) entre chaque indice radiomique local et l'indice radiomique correspondant global.

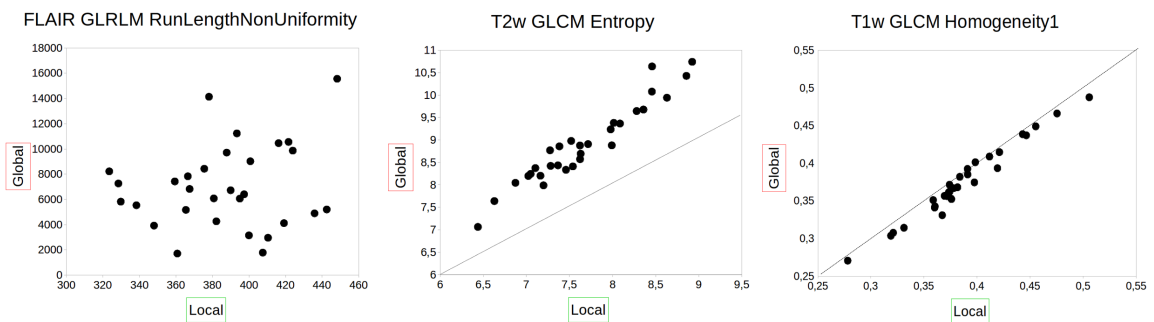


FIGURE 4.22 – Les valeurs de trois différents indices calculés localement versus globalement dans la VOI_{tumeur} . Les lignes pointillées sur les graphes du milieu et à droite sont les lignes d'identité ($x=y$).

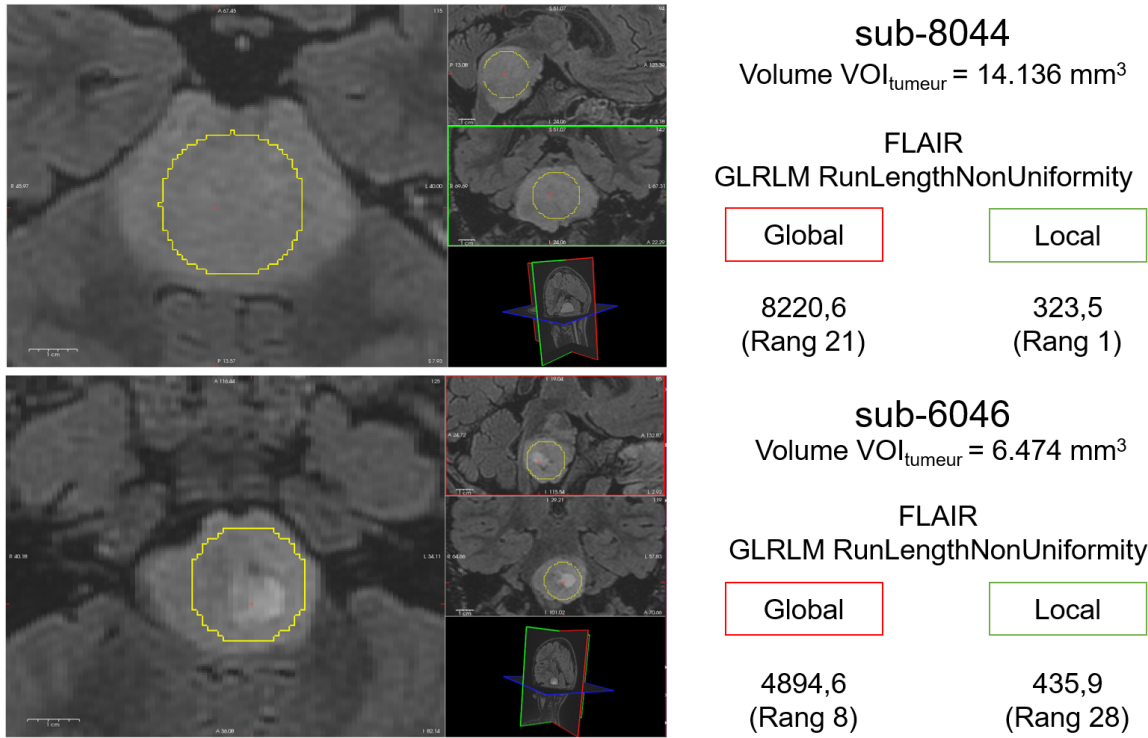


FIGURE 4.23 – Exemple d’images FLAIR de deux patients et les valeurs de $GLRLM RunLengthNonUniformity$ calculées dans la VOI_{tumeur} (en jaune) par les méthodes globale et locale. Les volumes des VOI_{tumeur} sont aussi affichés.

4.5 Discussion et conclusion

Dans ce chapitre, des études préliminaires conduites sur les indices radiomiques ont été présentées. Elles représentent une étape fondamentale pour évaluer l’apport de la radiomique pour mieux caractériser les patients [GITC](#).

Discretisation : L’étude des méthodes de discretisation a montré que les méthodes $d_{rel,fLN}$ et $d_{abs,fNN}$ présentent des résultats globalement similaires. Pour les deux méthodes, la largeur de l’intervalle sur lequel les niveaux de gris sont regroupés est la même pour tous les patients. Le fait de fixer les minimum et maximum des valeurs d’intensité ne modifie pas beaucoup les rangs des patients selon leurs indices radiomiques. Par contre, les rangs de patients issus de la méthode $d_{rel,fNN}$ sont très différents des autres méthodes, même si le même nombre de niveaux de gris a été fixé ($NN=64$) pour $d_{abs,fNN}$ et $d_{rel,fNN}$. Ces résultats sont en accord avec les résultats trouvés par [ORLHAC et collab. \[2015\]](#) en [TEP](#) qui ont comparé les méthodes $d_{abs,fNN}$ et $d_{rel,fNN}$ mais aussi avec les résultats de [LEIJENAAR et collab. \[2015\]](#) qui ont comparé $d_{rel,fNN}$ et $d_{rel,fLN}$. En plus, les classements des patients en utilisant les indices radiomiques calculés avec $d_{rel,fLN}$ et $d_{abs,fNN}$ sont cohérents avec l’évaluation visuelle de 3 experts, alors que lorsqu’ils sont calculés avec $d_{rel,fNN}$, ils ne sont pas cohérents avec l’impression visuelle comme rapporté par [ORLHAC et collab. \[2015\]](#).

Pour la comparaison des trois méthodes, les paramètres nombre de niveaux de gris et largeur de l’intervalle des niveaux de gris ont été fixés, mais des résultats similaires ont été obtenus avec l’utilisation de différentes valeurs pour ces paramètres. Tous les résultats suggèrent ainsi que l’important pour une comparaison inter-patient cohérente avec l’impression visuelle est de fixer la largeur des niveaux de gris. Pour la suite de cette thèse, la discretisation $d_{rel,fLN}$ sera utilisée, car elle obtient les meilleurs résultats (meilleure cohérence avec les trois experts) et n’exige pas l’étape du choix de valeurs d’intensité maximale et minimale de la discretisation comme le de-

mande $d_{abs,fNN}$.

Les analyses des méthodes de discrétisation ont été menées avec une sphère de rayon = 4,5 mm, la $VOI_{biopsie}$, qui inclut 389 voxels pour les images ré-échantillonnées à une taille de voxel égale à 1 mm x 1 mm x 1 mm. Même si la taille de la $VOI_{biopsie}$ est petite, elle est suffisante pour avoir des résultats fiables. Par exemple, [ORLHAC et collab. \[2014\]](#) ont suggéré une taille minimale de $4 \times 4 \times 4 = 64$ voxels, car certains indices comme ceux provenant de la matrice $GLRLM$ se basent sur des séries de voxels voisins dans les 3 directions (x,y,z) et des séries de moins de 4 voxels ne seraient pas expressives. Les résultats présentés sur la figure 4.21 en comparant les résultats de texture obtenus de façon globale et de façon locale avec des sphères de rayon 5mm confirment aussi que des régions aussi petites reflètent des informations pertinentes.

Finalement, [DURON et collab. \[2019\]](#) ont récemment comparé les discrétisations $d_{rel,fLN}$ et $d_{rel,fNN}$ en IRM pour 74 patients de lésion des glandes lacrymales et 30 patientes de cancer du sein. Ils ont conclu que la discrétisation $d_{rel,fLN}$ est plus reproductible et ont également conseillé son utilisation pour les études radiomiques en IRM .

Volume de la ROI : Ce chapitre a montré que le volume de la région d'intérêt utilisée pour calculer les indices radiomiques peut biaiser les valeurs de ces indices. Des résultats similaires ont été obtenus par [SUKIO et collab. \[2015\]](#) pour des régions 2D en IRM .

Une conséquence de cette dépendance au volume est la difficulté de mener des études comparatives. Nous avons montré des différences statistiquement significatives entre de nombreux indices radiomiques calculés dans VOI_{MNI5} et les régions VOI_L et VOI_R qui représentent des régions "équivalentes" situées dans la MB . Or, ces différences ne sont pas trouvées entre VOI_L et VOI_R qui ont la même taille; ainsi ces différences sont très probablement une conséquence des différentes tailles de VOI_{MNI5} et non pas des différences au sein de la MB .

Les résultats de ce chapitre, montrent que la comparaison des valeurs d'indices radiomiques doit être réalisée avec attention quand les régions d'intérêt ont des volumes différents. Ainsi, l'étude de [MICHOUX et collab. \[2015\]](#) a proposé un modèle radiomique qui inclut l'indice $GLRLM_{GrayLevelNonUniformity}$. Comme cet indice est fortement corrélé au volume, ce modèle montre probablement un impact du volume.

Il faut souligner que les corrélations avec le volume ont été analysées dans un cas extrême : l'analyse a été faite dans la MB qui ne présente quasiment pas de variations entre les patients. Dans ce cas, les biais induits par le volume peuvent être beaucoup plus prononcés que dans le cas plus usuel en radiomique qui est la comparaison de tumeurs qui ont non seulement des volumes différents mais aussi des caractéristiques d'hétérogénéité variables. De plus, les régions utilisées pour l'étude sont petites et nous avons montré que certains indices, comme le $GLCM_{Entropy}$, présentent une corrélation plus forte avec le volume quand les régions d'intérêt sont petites. Ce résultat peut être aussi retrouvé en regardant l'analyse faite avec les régions tumorales (figure 4.20). En effet, seulement douze indices se sont montrés fortement corrélés au volume en comparaison aux 46 indices trouvés avec l'étude sur les petites régions de la MB . Ce résultat est dû aux fortes variations d'hétérogénéité que les tumeurs ont entre elles (hétérogénéité qui n'existe pas dans la MB) et aux tailles des régions beaucoup plus importantes (environ 10.000 voxels). Néanmoins, dans cette étape d'investigation, il est bien d'analyser les cas extrêmes pour être averti des biais potentiels et ne pas aboutir à des conclusions erronées quant à l'interprétation des résultats.

Indices radiomiques locaux : En fonction des résultats montrant l'influence du volume sur les paramètres radiomiques, nous avons introduit une méthode de calcul local pour pouvoir analyser les différentes tumeurs $GITC$, qui ont des tailles variables. Cette méthode calcule des indices qui ont, en effet, une très faible corrélation avec le volume tumoral ($|r| < 0,35$, figure 4.21).

Les résultats de la figure 4.21 confirment aussi que les indices locaux traduisent bien l'hétérogénéité de la VOI_{Tumeur} puisqu'ils sont très corrélés aux indices globaux pour lesquelles les différences de volume des VOI_{Tumeur} n'ont pas un impact important. Dans ce cas, les indices locaux et globaux traduisent l'hétérogénéité tumorale.

De plus, ces indices locaux renvoient des valeurs plus cohérentes avec le caractère hétérogène des régions dans les cas d'indices très corrélés au volume, comme le *GLRLM RunLengthNonUniformity*, les effets de l'hétérogénéité tumorale sont masqués par l'effet du volume quand cet indice est calculé de façon globale.

Conclusion

Les analyses menées dans ce chapitre ont montré l'importance de bien choisir la méthode de calcul d'indices radiomiques avant de conduire des analyses comparatives. Notamment, la discrétisation de niveaux de gris avec une largeur constante pour tous les patients est essentielle pour une comparaison appropriée. Les bornes minimale et maximale d'intensité peuvent être relatives à la région d'intérêt (et pas absolues) pour faciliter les calculs et maintenir des comparaisons cohérentes. En outre, le volume de la région d'intérêt peut biaiser plusieurs indices radiomiques. Pour pallier ce problème, une méthode locale d'extraction d'indices a été proposée pour évaluer la tumeur dans sa globalité en calculant les indices dans des régions de taille constante. Ces indices locaux ne sont pas corrélés au volume de la région et ont une relation plus cohérente avec leur définition.

4.6 Références

- AERTS, H., E. RIOS VELAZQUEZ, R. LEIJENAAR, C. PARMAR, P. GROSSMANN, S. CAVALHO, J. BUSSINK, R. MONSHOUWER, B. HAIBE-KAINS, D. RIETVELD, F. HOEBERS, M. M. RIETBERGEN, C. LEEMANS, A. DEKKER, J. QUACKENBUSH, R. GILLIES et P. LAMBIN. 2014, «Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach», *Nature communications*, vol. 5, doi :10.1038/ncomms5006, p. 4006. 92
- DURON, L., D. BALVAY, S. VANDE PERRE, A. BOUCHOUICHA, J. SAVATOVSKY, J.-C. SADIK, I. THOMASSIN-NAGGARA, L. FOURNIER et A. LECLER. 2019, «Gray-level discretization impacts reproducible MRI radiomics texture features», *Plos One*, vol. 14, n° 3, doi :10.1371/journal.pone.0213459, p. e0213459, ISSN 1932-6203. URL <http://dx.plos.org/10.1371/journal.pone.0213459>. 117
- HATT, M., M. MAJDOUB, M. VALLIÈRES, F. TIXIER, C. CHEZE-LE REST, D. GROHEUX, E. HINDIÉ, A. MARTINEAU, O. PRADIER, R. HUSTINX, R. PERDRISOT, R. GUILLEVIN, I. EL NAQA et D. VISVIKIS. 2015, «18F-FDG PET Uptake Characterization Through Texture Analysis : Investigating the Complementary Nature of Heterogeneity and Functional Tumor Volume in a Multi-Cancer Site Patient Cohort», *Journal of Nuclear Medicine*, vol. 56, n° 1, doi :10.2967/jnumed.114.144055, p. 38–44, ISSN 1535-5667. URL <http://jnm.snmjournals.org/content/56/1/38.abstractN2>. 92
- KICKINGEREDER, P., S. BURTH, A. WICK, M. GÖTZ, O. EIDEL, H.-P. SCHLEMMER, K. H. MAIERHEIN, W. WICK, M. BENDSZUS, A. RADBRUCH et D. BONEKAMP. 2016, «Radiomic Profiling of Glioblastoma : Identifying an Imaging Predictor of Patient Survival with Improved Performance over Established Clinical and Radiologic Risk Models», *Radiology*, vol. 280, n° 3, doi :10.1148/radiol.2016160845, p. 880–889, ISSN 0033-8419. URL <http://pubs.rsna.org/doi/10.1148/radiol.2016160845>. 93
- LARROZA, A., V. BODÍ et D. MORATAL. 2016, «Texture Analysis in Magnetic Resonance Imaging : Review and Considerations for Future Applications», dans *Assessment of Cellular and Organ Function and Dysfunction using Direct and Derived MRI Methodologies*, InTech, ISBN 978-953-51-2723-9, p. 75–106, doi :10.5772/64641. 92, 112

- LEIJENAAR, R. T. H., G. NALBANTOV, S. CARVALHO, W. J. C. VAN ELMPT, E. G. C. TROOST, R. BOELLAARD, H. J. W. L. AERTS, R. J. GILLIES et P. LAMBIN. 2015, «The effect of SUV discretization in quantitative FDG-PET Radiomics : the need for standardized methodology in tumor texture analysis.», *Scientific reports*, vol. 5, n° August, doi :10.1038/srep11075, p. 11 075, ISSN 2045-2322. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4525145&tool=pmcentrez&rendertype=abstract>. 92, 93, 116
- MICHOUX, N., S. VAN DEN BROECK, L. LACOSTE, L. FELLAH, C. GALANT, M. BERLIÈRE et I. LECONTE. 2015, «Texture analysis on MR images helps predicting non-response to NAC in breast cancer.», *BMC cancer*, vol. 15, doi :10.1186/s12885-015-1563-8, p. 574, ISSN 1471-2407. URL <http://www.ncbi.nlm.nih.gov/pubmed/26243303><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4526309>. 93, 117
- MOLINA, D., J. PÉREZ-BETETA, B. LUQUE, E. ARREGUI, M. CALVO, J. M. BORRÁS, C. LÓPEZ, J. MARTINO, C. VELASQUEZ, B. ASENJO, M. BENAVIDES, I. HERRUZO, A. MARTÍNEZ-GONZÁLEZ, L. PÉREZ-ROMASANTA, E. ARANA et V. M. PÉREZ-GARCÍA. 2016, «Tumour heterogeneity in glioblastoma assessed by MRI texture analysis : A potential marker of survival», *British Journal of Radiology*, vol. 89, n° 1064, doi :10.1259/bjr.20160242, p. 20160 242, ISSN 00071285. URL <http://www.ncbi.nlm.nih.gov/pubmed/27319577><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5124892>. 93
- MOLINA, D., J. PÉREZ-BETETA, A. MARTÍNEZ-GONZÁLEZ, J. MARTINO, C. VELASQUEZ, E. ARANA et V. M. PÉREZ-GARCÍA. 2017, «Lack of robustness of textural measures obtained from 3D brain tumor MRIs impose a need for standardization», *PLoS ONE*, vol. 12, n° 6, doi :10.1371/journal.pone.0178843, p. e0178 843, ISSN 19326203. 92
- NIOCHE, C., F. ORLHAC, S. BOUGHDAD, S. REUZE, J. GOYA-OUTI, C. ROBERT, C. PELLOT-BARAKAT, M. SOUSSAN, F. ERIQUE FROUIN et I. BUVAT. 2018, «Lifex : A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity», *Cancer Research*, vol. 78, n° 16, doi :10.1158/0008-5472.CAN-18-0125, p. 4786–4789, ISSN 15387445. URL <http://cancerres.aacrjournals.org/>. 92, 112
- ORLHAC, F., M. SOUSSAN, K. CHOUAHNIA, E. MARTINOD et I. BUVAT. 2015, «18F-FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer», *PLoS ONE*, vol. 10, n° 12, doi :10.1371/journal.pone.0145063, p. e0145 063, ISSN 19326203. 92, 93, 98, 116
- ORLHAC, F., M. SOUSSAN, J.-A. MAISONOBE, C. A. GARCIA, B. VANDERLINDEN et I. BUVAT. 2014, «Tumor Texture Analysis in 18F-FDG PET : Relationships Between Texture Parameters, Histogram Indices, Standardized Uptake Values, Metabolic Volumes, and Total Lesion Glycolysis», *Journal of Nuclear Medicine*, vol. 55, n° 3, doi :10.2967/jnumed.113.129858, p. 414–422, ISSN 0161-5505. URL <http://jnm.snmjournals.org/content/55/3/414.abstract><http://jnm.snmjournals.org/cgi/doi/10.2967/jnumed.113.129858>. 92, 117
- SIKIO, M., K. K. HOLLI-HELENIUS, P. RYYMIN, P. DASTIDAR, H. ESKOLA et L. HARRISON. 2015, «The effect of region of interest size on textural parameters», dans *9th International Symposium on Image and Signal Processing and Analysis, ISPA 2015*, ISBN 9781467380324, p. 149–153, doi : 10.1109/ISPA.2015.7306049. 117
- VALLIERES, M., D. VISVIKIS et M. HATT. 2018, «Dependency of a validated radiomics signature on tumor volume and potential corrections», *Journal of Nuclear Medicine*, vol. 59 (supplement 1), p. 640. 92
- VAN GRIETHUYSEN, J. J., A. FEDOROV, C. PARMAR, A. HOSNY, N. AUCCOIN, V. NARAYAN, R. G. BEETS-TAN, J. C. FILLION-ROBIN, S. PIEPER et H. J. AERTS. 2017, «Computational radiomics system to decode the radiographic phenotype», *Cancer Research*, vol. 77, n° 21, doi :10.1158/0008-5472.CAN-17-0339, p. e104–e107, ISSN 15387445. 93

YANG, D., G. RAO, J. MARTINEZ, A. VEERARAGHAVAN et A. RAO. 2015, «Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma», *Medical Physics*, vol. 42, n° 11, doi :10.1118/1.4934373, p. 6725–6735, ISSN 2473-4209. URL <http://onlinelibrary.wiley.com/doi/10.1118/1.4934373/abstract>. 93

Chapitre 5

Études radiomiques appliquées aux prédictions des mutations du GITC

« Regretter ce que l'on n'a pas, c'est gâcher ce que l'on a »

Ken Keyes



Ce chapitre considère les analyses radiomiques réalisées pour la prédiction des mutations génétiques associées au GITC. Les images de diagnostic ont été utilisées pour évaluer la pertinence de la radiomique pour la classification de trois mutations : les mutations des histones H3.1 ou H3.3; la présence ou absence de mutation sur le gène TP53; la présence ou absence de mutation sur le gène ACVR1. L'étude commence par une analyse uni-variée réalisée pour chaque indice (indices cliniques, indices d'imagerie locaux et indices d'imagerie globaux) suivie de la procédure de sélection d'indices proposée dans ce chapitre. Les indices sélectionnés sont ensuite utilisés pour définir des modèles radiomiques multi-variés en appliquant différentes techniques d'apprentissage automatique. Ce chapitre finit par la proposition d'une nouvelle méthodologie, multi-modèles, pour réaliser les analyses radiomiques avec des cohortes de patients de faible effectif, avec données manquantes et qui sont progressivement augmentées. La méthode est présentée puis illustrée sur la prédiction des mutations H3.1 versus H3.3 sur les patients atteints de GITC.

Sommaire

5.1 Analyses uni-variées des mutations génétiques associées au GITC	123
5.1.1 Mutation de l'histone H3	123
5.1.2 Mutation du gène TP53	127
5.1.3 Mutation du gène ACVR1	131
5.1.4 Synthèse des résultats des analyses uni-variées pour les trois mutations . . .	134
5.2 Analyses multi-variées pour prédire les mutations génétiques associées au GITC	134
5.2.1 Pipeline de sélection d'indices radiomiques proposé	135
5.2.2 Modèles d'apprentissage automatique utilisés	139
5.2.3 Application des modèles de classification pour prédire la mutation de l'histone H3	140
5.2.4 Application des modèles de classification pour la mutation du gène TP53	145
5.2.5 Application des modèles de classification pour la mutation ACVR1	148
5.3 La méthode multi-modèles proposée	151

5.3.1 Définition de la méthode	151
5.3.2 Cas d'usage de la méthode multi-modèles sur les patients GITC	152
5.4 Discussion et conclusion	162
5.5 Références	167

5.1 Analyses uni-variées des mutations génétiques associées au GITC

Les mutations présentes dans les GITC sont associées à des pronostics différents et guident actuellement la définition du traitement pour chaque patient. Elles sont à présent identifiées par une chirurgie de biopsie mais il arrive que les patients refusent de s’y soumettre ou que les résultats des analyses de biopsies ne soient pas concluants. Il est donc de grand intérêt que l’on puisse les prédire de façon non-invasive par un modèle radiomique qui utilise les images de diagnostic du protocole d’IRM multi-modale défini pour ces patients.

Les mutations étudiées sont les mutations de l’histone H3, du gène TP53 et du gène ACVR1, présentées dans le chapitre 2. Comme déjà mentionné, ces trois mutations sont très corrélées. Cela se vérifie sur la cohorte PREBIOMEDE, comme le montre la figure 5.1 donnant les matrices de confusion des mutations comparées deux-à-deux. Néanmoins, comme la corrélation n’est pas parfaite, les analyses de ce chapitre ont été réalisées pour chaque mutation séparément.

Avant de créer les modèles prédictifs pour chaque mutation, nous avons analysé les degrés de prédiction indice par indice. Pour avoir des puissances statistiques comparables pour tous les indices, les analyses d’imagerie ont été faites sur les patients PREBIOMEDE qui avaient bénéficié des quatre modalités d’IRM structurelles (T1w, T2w, T1c, FLAIR) à la session de diagnostic (avant la biopsie). Ainsi, tous les indices ont été analysés avec le même nombre de patients pour chaque étude. Ces études concernent les 38 patients PREBIOMEDE soulignés en rouge dans le tableau A.1 en annexe, tous les patients de la cohorte étoile ayant été inclus dans cette étude. Cet ensemble de 38 patients est nommé cohorte radiomique 4Mod.

Trois indices cliniques ont été évalués : l’âge au diagnostic, le sexe et le volume de la VOI_{Tumeur} . Les 19 indices d’imagerie de premier ordre (voir le tableau 5.1) ont été ajoutés aux indices de texture (tableau 4.1) pour les quatre modalités d’IRM, totalisant 316 indices globaux et 316 indices locaux par patient. Les formules mathématiques utilisées par PyRadiomics v1.1.1 figurent en annexe (A.2.1).

5.1.1 Mutation de l’histone H3

Pour évaluer la prédiction de la mutation H3, seuls les 30 patients de la cohorte radiomique 4Mod qui présentaient la mutation H3.1 (8 patients) ou H3.3 (22 patients) ont été inclus dans l’étude.

Indices cliniques

Les pouvoirs discriminants du volume et de l’âge, qui sont des variables continues, ont été évalués par l’AUC, alors que celui du sexe des patients, qui est une variable catégorielle, a été évalué par un test du Chi-2, l’hypothèse nulle étant que la mutation H3 était indépendante du sexe. La figure 5.2 montre les relations de ces trois indices cliniques avec la mutation H3 par des boîtes à moustaches pour l’âge et le volume et par une représentation en barres pour le sexe. Le principal résultat est que la valeur de l’AUC de l’âge au diagnostic pour la discrimination de la mutation H3 est élevée.

	ACVR1 - WT	ACVR1		TP53 - WT	TP53		ACVR1 - WT	ACVR1
H3.1	3	8	H3.1	9	2	TP53 -WT	8	7
H3.3	23	0	H3.3	6	17	TP53	18	1

FIGURE 5.1 – Matrices de confusion des mutations H3, TP53 et ACVR1 comparées 2 à 2 pour les 34 patients PREBIOMEDE présentant les mutations de type H3.1 ou H3.3.

TABEAU 5.1 – Liste des indices de premier ordre (appelés *firstorder* par *PyRadiomics*)

Indices de premier ordre (firstorder)
InterquartileRange
Skewness
Uniformity
MeanAbsoluteDeviation
Energy
RobustMeanAbsoluteDeviation
Median
TotalEnergy
Maximum
RootMeanSquared
90Percentile
Minimum
Entropy
StandardDeviation
Range
Variance
10Percentile
Kurtosis
Mean

30 patients de la cohorte radiomique 4Mod avec la mutation H3.1 ou H3.3 :

8 H3.1 vs. 22 H3.3

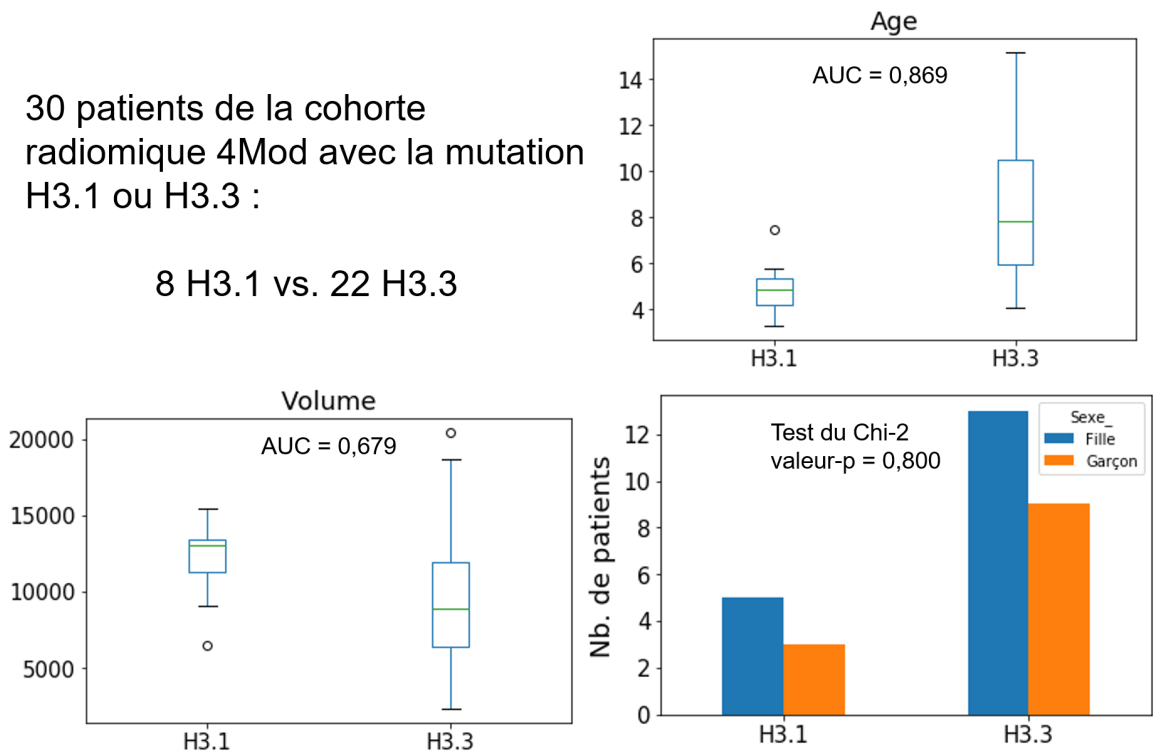


FIGURE 5.2 – Analyse des trois paramètres cliniques pour l'évaluation de la mutation H3 pour les patients de la cohorte radiomique 4Mod. Le volume de la VOI_{tumeur} ainsi que l'âge au diagnostic sont représentés par des boîtes à moustaches et les AUCs sont indiquées. Le sexe du patient est représenté par des barres et la valeur-p du test du Chi-2 est effectué.

Ensuite, de façon exploratoire, pour augmenter la puissance statistique des résultats, et ne comparer que les indices cliniques entre eux, tous les patients de la cohorte PREBIOMEDE pour

lesquels les informations cliniques étaient disponibles ont été analysés. Au total, 42 patients ont été inclus dans cette analyse : 12 patients avec une mutation H3.1 et 30 patients avec une mutation en H3.3. Les résultats sont présentés sur la figure 5.3. On peut noter que l'ajout des 12 patients diminue légèrement les différences entre les deux groupes (H3.1 et H3.3). Cependant, l'âge au diagnostic reste un indice avec une AUC élevée.

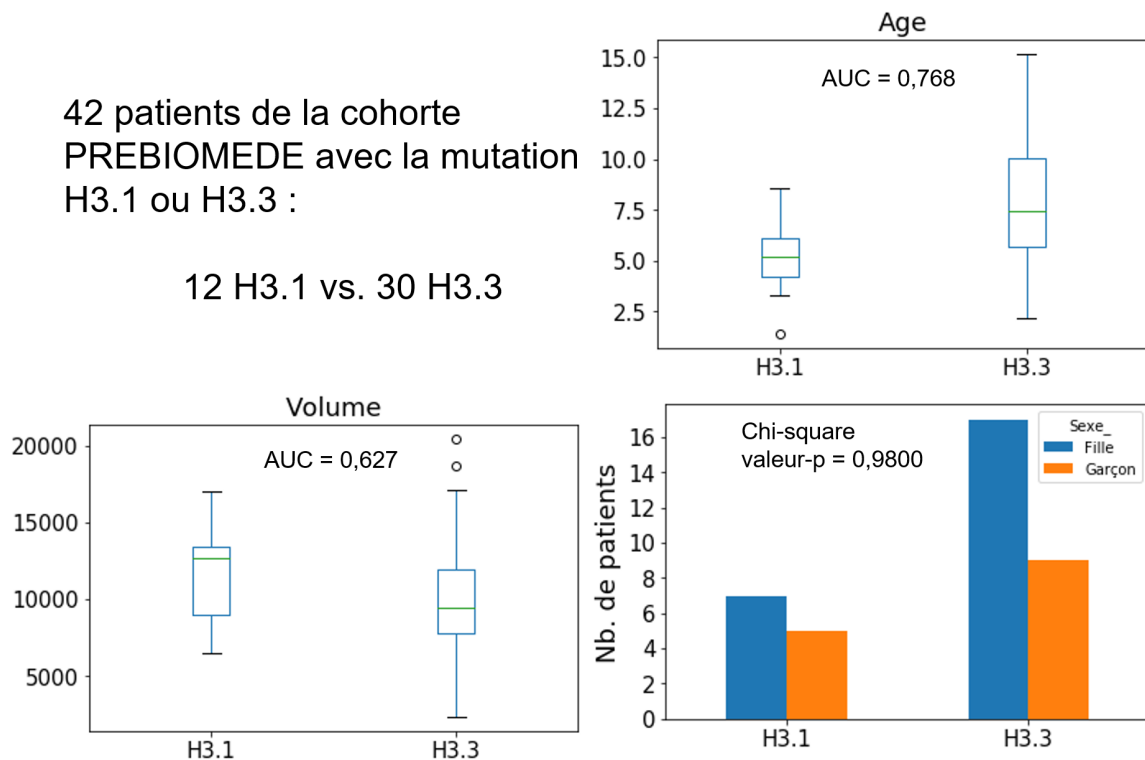


FIGURE 5.3 – Analyse des trois paramètres cliniques pour l'évaluation de la mutation H3 pour les patients de la cohorte PREBIOMEDE. Le volume de la VOI_{tumeur} ainsi que l'âge au diagnostic sont représentés par des boîtes à moustaches et les AUCs sont indiquées. Le sexe du patient est représenté par des barres et la valeur-p du test du Chi-2 est effectué.

Finalement, la même analyse a été faite pour les patients de notre cohorte BIOMEDE. La figure 5.4 présente les résultats. Dans cette cohorte la distribution de classes est plus déséquilibrée (5 patients H3.1 et 20 patients H3.3) et l'âge au diagnostic présente une AUC extrêmement élevée.

25 patients de la cohorte BIOMEDE avec la mutation H3.1 ou H3.3 :

5 H3.1 vs. 20 H3.3

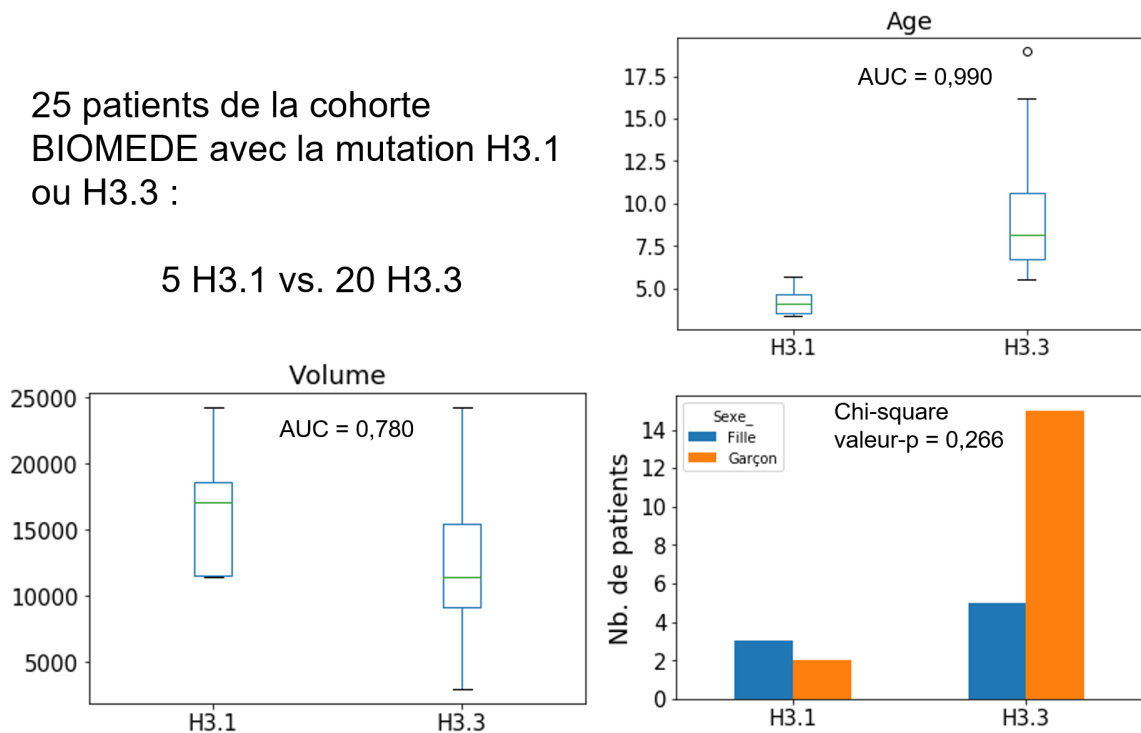


FIGURE 5.4 – Analyse des trois paramètres cliniques pour l'évaluation de la mutation H3 pour les patients de la cohorte BIOMEDE. Le volume de la VOI_{tumeur} ainsi que l'âge au diagnostic sont représentés par des boîtes à moustaches et les AUCs sont indiquées. Le sexe du patient est représenté par des barres et la valeur-p du test du Chi-2 est effectué.

L'analyse des trois cohortes montre que les valeurs des **AUC** peuvent varier beaucoup étant donné le faible effectif présent dans ces cohortes. L'**AUC** de l'âge varie entre 0,77 et 0,99 et celle du volume entre 0,63 et 0,78. Néanmoins, il est possible de mettre en évidence des tendances dont l'âge au diagnostic qui est toujours un bon critère pour prédire des mutations H3.1 ou H3.3. Ce résultat a été mis en évidence par d'autres études ([CASTEL et collab. \[2015\]](#); [COHEN et collab. \[2017\]](#); [MACKAY et collab. \[2017\]](#)).

Indices d'imagerie

Les indices d'imagerie ont été évalués individuellement par leur **AUC**, comme montré dans le tableau [A.4](#) pour les 30 patients de la [cohorte radiomique 4Mod](#). Parmi les indices calculés en **FLAIR**, 33 indices locaux et 17 globaux présentent une **AUC** supérieure ou égale à 0,75. Parmi les indices calculés en **T1c**, cinq indices locaux et six globaux présentent une **AUC** supérieure ou égale à 0,75. Parmi les indices calculés en **T1w**, dix indices locaux et sept globaux présentent une **AUC** supérieure ou égale à 0,75. Parmi les indices calculés en **T2w**, aucun indice local et 3 globaux présentent une **AUC** supérieure ou égale à 0,75.

Ainsi, la modalité **FLAIR** présente le plus grand nombre d'indices avec une **AUC** supérieure à 0,75. Cependant, les sept indices qui présentent une **AUC** supérieure ou égale à l'**AUC** de l'âge au diagnostic pour cette même cohorte sont trouvés pour les modalités **T1w** et **T1c** et sont affichés en vert dans le tableau en annexe [A.4](#). Ces indices sont tous de premier ordre. Trois exemples de ces indices qui présentent une **AUC** élevée sont illustrés sur la figure [5.5](#).

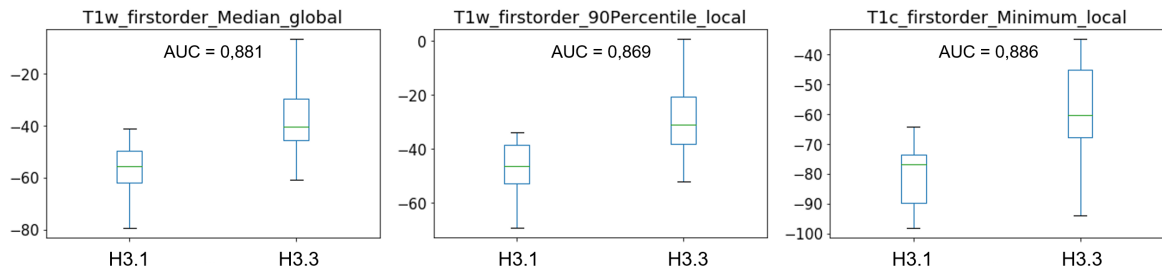


FIGURE 5.5 – Boîte à moustaches de trois indices d’imagerie qui présentent une AUC plus élevée que l’âge au diagnostic pour prédire la mutation H3.

Comme déjà indiqué, il faut plus se fier aux tendances qu’aux valeurs des AUC. Les indices qui présentent l’AUC plus élevée que l’âge au diagnostic ont été signalés à titre indicatif. A ce stade, on peut seulement affirmer que les indices à l’AUC élevée ont tendance à pouvoir différencier les mutations H3.1 des mutations H3.3.

5.1.2 Mutation du gène TP53

Pour évaluer la prédiction de la mutation du gène TP53, les 33 patients de la [cohorte radiomique 4Mod](#) qui présentaient l’information de la présence de la mutation TP53 ont été inclus. Au total, 13 patients sont TP53-WT (pas de mutation du gène TP53) et 20 présentaient une mutation du gène TP53.

Indices cliniques

Les pouvoirs discriminants de chaque indice clinique pour la mutation du gène TP53 ont été évalués comme pour la mutation H3. La figure 5.6 présente les relations de ces trois indices cliniques avec la mutation TP53 par les boîtes à moustaches de l’âge au diagnostic et du volume et par une représentation en barres pour le sexe. Aucun indice clinique ne semble être prédictif. Les boîtes à moustaches suggèrent cependant une différence d’âge et de volume entre les patients mutés en TP53 et TP53-WT.

33 patients de la cohorte radiomique 4Mod avec l'information sur la mutation du gène TP53 :

13 TP53-WT vs. 20 TP53

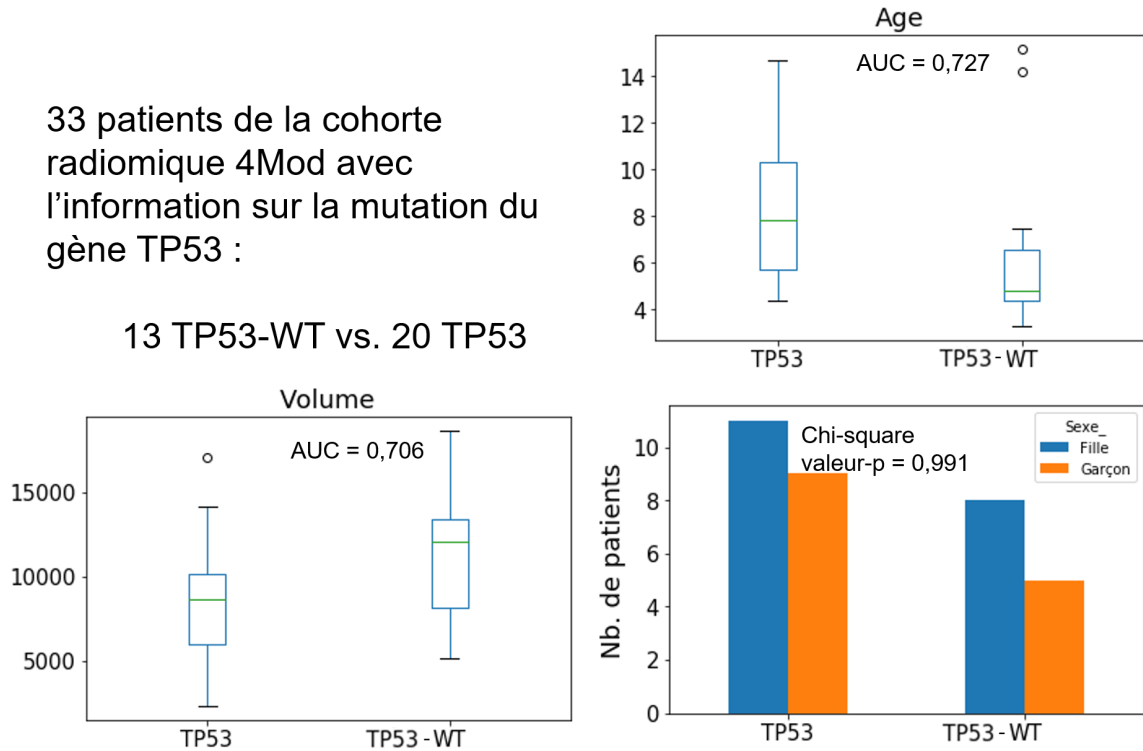


FIGURE 5.6 – Analyse des trois paramètres cliniques pour l'évaluation de la mutation TP53 pour les 33 patients de la cohorte radiomique 4Mod. Le volume de la VOI_{tumeur} ainsi que l'âge au diagnostic sont représentés par des boîtes à moustaches et les AUCs sont indiquées. Le sexe du patient est représenté par des barres et la valeur-p du test du Chi-2 est effectué.

Ensuite, uniquement en mode exploratoire, pour augmenter la puissance statistique des résultats et ne comparer que les indices cliniques entre eux, tous les patients de la cohorte PRE-BIOMEDE pour lesquels les informations cliniques étaient disponibles ont été analysés. Au total, 39 patients ont été inclus dans cette analyse : 17 patients TP53-WT et 22 patients présentant une mutation du gène TP53. Les résultats sont présentés sur la figure 5.7.

39 patients de la cohorte
PREBIOMEDE avec l'information
sur la mutation du gène TP53 :

17 TP53-WT vs. 22 TP53

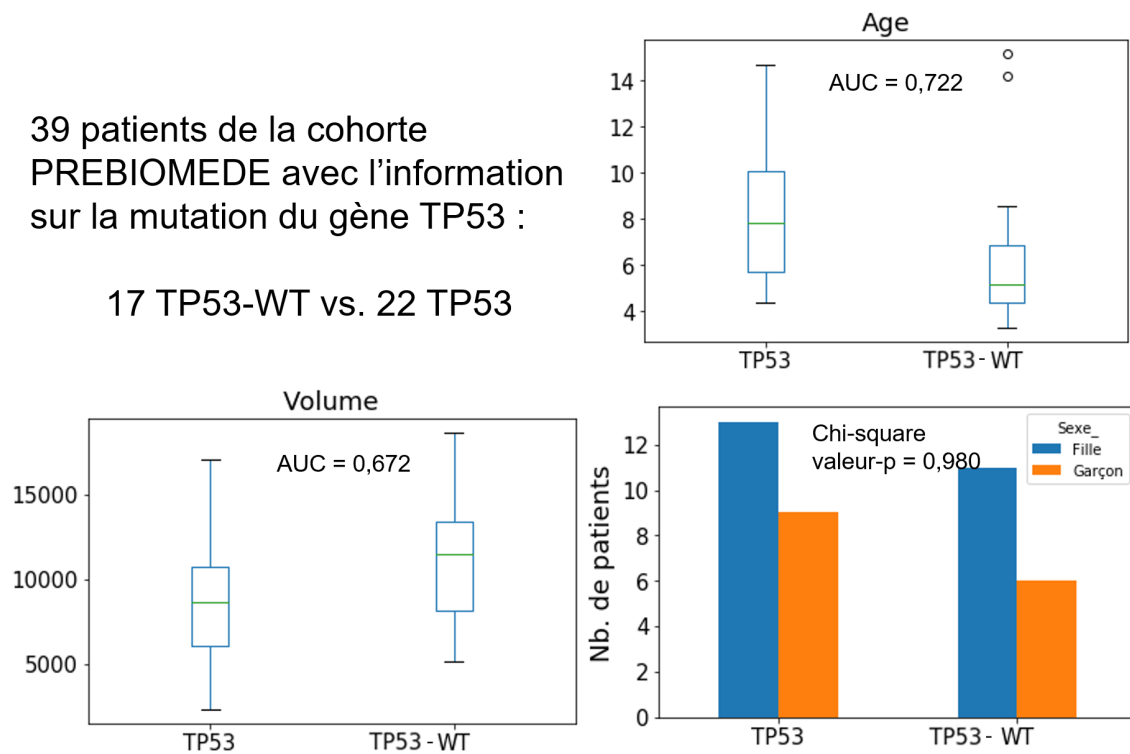


FIGURE 5.7 – Analyse des trois paramètres cliniques pour l'évaluation de la mutation du gène TP53 pour les 39 patients de la cohorte PREBIOMEDE. Le volume de la VOI_{tumeur} ainsi que l'âge au diagnostic sont représentés par des boîtes à moustaches et les AUCs sont indiquées. Le sexe du patient est représenté par des barres et la valeur-p du test du Chi-2 est effectué.

Finalement, la même analyse a été faite pour les patients de la cohorte BIOMEDE. La figure 5.8 présente les résultats. Dans cette cohorte il y a 12 patients WT et 14 patients présentant une mutation du gène TP53. Pour cette cohorte, l'âge présente une AUC très élevée.

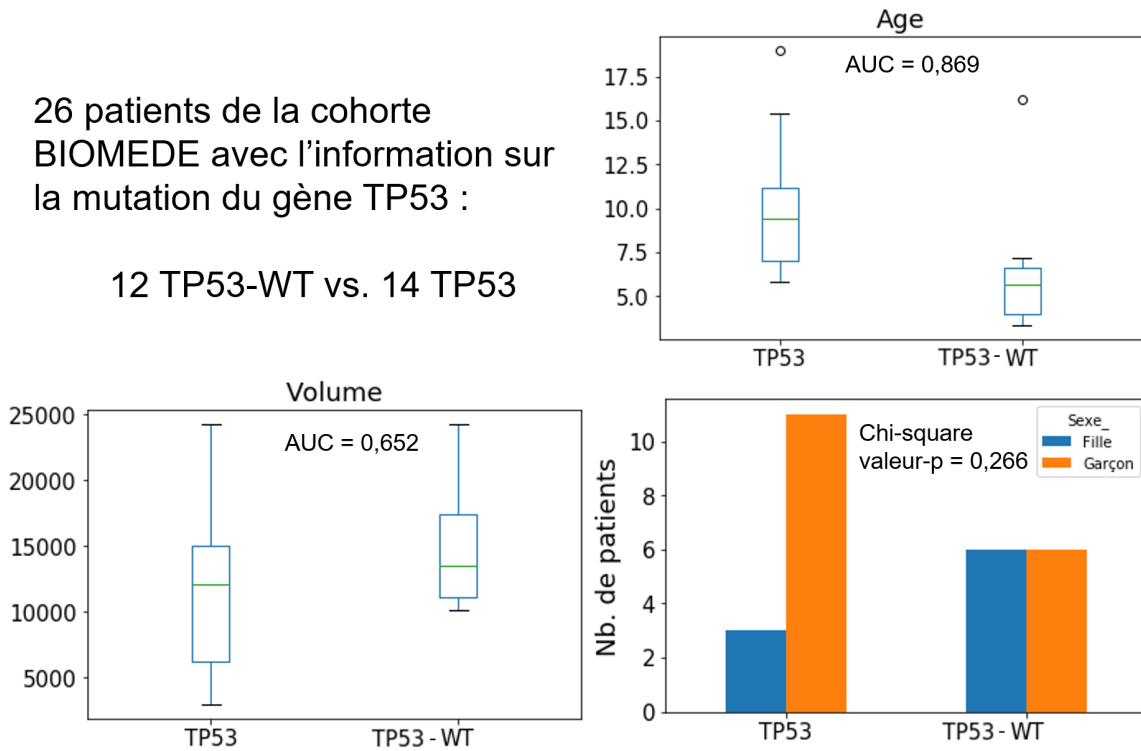


FIGURE 5.8 – Analyse des trois paramètres cliniques pour l'évaluation de la mutation du gène pour les 26 patients de la cohorte BIOMEDE. Le volume de la VOI_{tumeur} ainsi que l'âge au diagnostic sont représentés par des boîtes à moustaches et les AUCs sont indiquées. Le sexe du patient est représenté par des barres et la valeur-p du test du Chi-2 est effectué.

L'âge est ainsi l'indice clinique qui permet le mieux de bien différencier les patients présentant une mutation du gène TP53 et les patients TP53-WT.

Indices d'imagerie

Les indices d'imagerie ont été évalués individuellement par leur AUC, comme synthétisé dans le tableau A.5 pour les 33 patients de la cohorte radiomique 4Mod. Parmi les indices calculés en FLAIR, 33 indices locaux et 22 globaux présentent une AUC supérieure ou égale à 0,75. Parmi les indices calculés en T1c, deux indices locaux et deux globaux présentent une AUC supérieure ou égale à 0,75. Parmi les indices calculés en T1w, aucun indice local et un global présente une AUC supérieure ou égale à 0,75. Finalement, parmi les indices calculés en T2w, deux indices locaux et 20 indices globaux présentent une AUC supérieure ou égale à 0,75. La modalité FLAIR présente le plus grand nombre d'indices avec une AUC supérieure à 0,75. Pour cette cohorte radiomique 4Mod, tous ces indices présentent une AUC supérieure à celle de l'âge au diagnostic (0,727).

Les six indices ayant une AUC supérieure à 0,85 sont représentés en vert dans le tableau en annexe A.5. Parmi ces indices, seul, FLAIR glszm SizeZoneVariability présente une corrélation de Spearman supérieure à 0,7 avec le volume. Le volume ne semble pas être un excellent prédicteur de la mutation du gène TP53. Les boîtes à moustaches des trois indices qui présentent des valeurs d'AUC élevées sont présentés sur la figure 5.9.

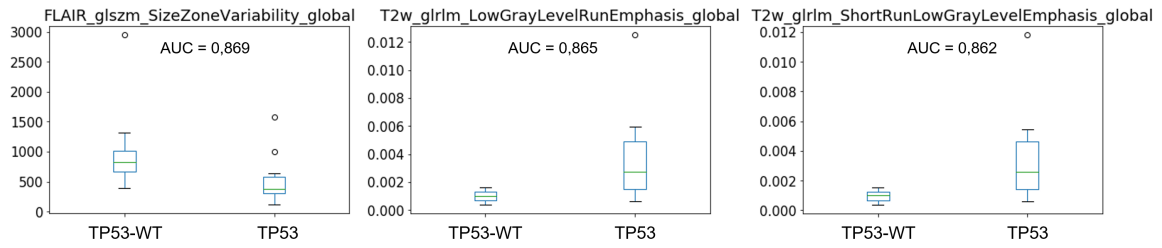


FIGURE 5.9 – Boîtes à moustaches des trois indices d’imagerie qui présentent des AUCs élevées pour prédire la mutation du gène TP53.

5.1.3 Mutation du gène ACVR1

Pour évaluer la prédiction de la mutation du gène ACVR1, les 33 patients de la **cohorte radiomique 4Mod** qui présentaient l’information de la présence de la mutation du gène ACVR1 ont été inclus (voir tableau A.1). Au total, 26 patients étaient ACVR1-WT (pas de mutation du gène ACVR1) et 7 présentaient une mutation du gène ACVR1.

Indices cliniques

Les pouvoirs discriminants de chaque indice clinique pour la prédiction de la mutation du gène ACVR1 ont été évalués comme pour les mutations H3 et TP53. La figure 5.10 montre les relations de ces trois indices cliniques avec la mutation du gène ACVR1. Aucun indice clinique n’est fortement prédictif de cette mutation.

33 patients de la cohorte radiomique 4Mod avec l’information sur la mutation du gène ACVR1 :

26 ACVR1-WT vs 7 ACVR1

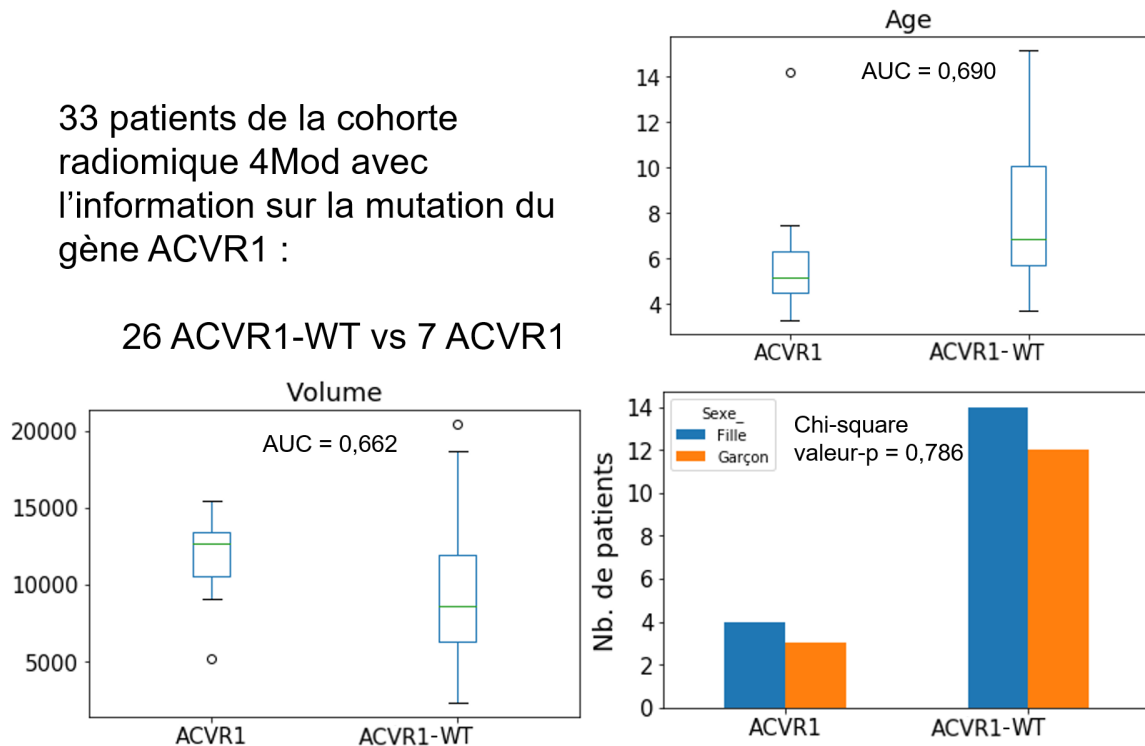


FIGURE 5.10 – Analyse des trois paramètres cliniques pour l’évaluation de la mutation du gène ACVR1 pour les 33 patients de la cohorte radiomique 4Mod.

Ensuite, tous les patients de la cohorte PREBIOMEDE pour lesquels les informations cliniques étaient disponibles ont été analysés. Au total, 39 patients ont été inclus dans cette analyse : 30 patients ACVR1-WT et 9 patients présentant une mutation du gène ACVR1. Les résultats sont présentés

tés sur la figure 5.11. Les indices cliniques ne semblent toujours pas être d'excellents prédicteurs de la mutation du gène ACVR1.

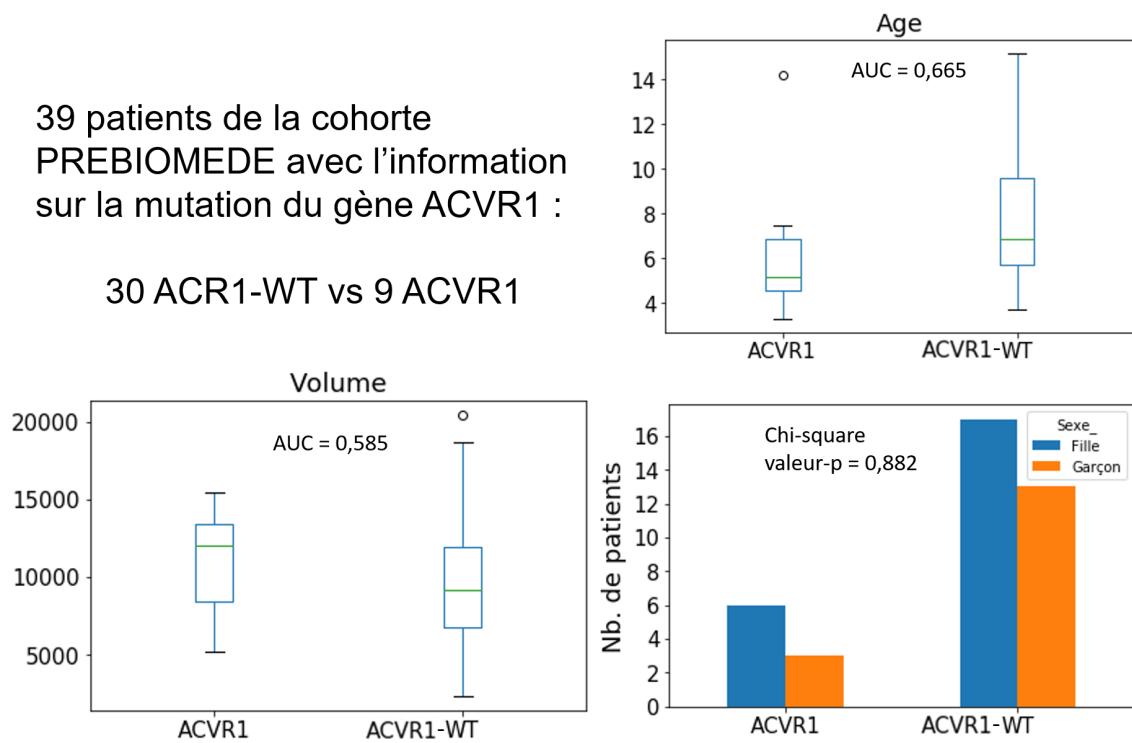


FIGURE 5.11 – Analyse des trois paramètres cliniques pour l'évaluation de la mutation du gène ACVR1 pour les 39 patients de la cohorte PREBIOMEDE.

Finalement, la même analyse a été réalisée pour les patients de la cohorte BIOMEDE. La figure 5.12 présente les résultats. Dans cette cohorte il y a 21 patients ACVR1-WT et 7 patients présentant une mutation du gène ACVR1. Cette cohorte montre que les patients présentant une mutation du gène ACVR1 ont tendance à être plus jeunes.

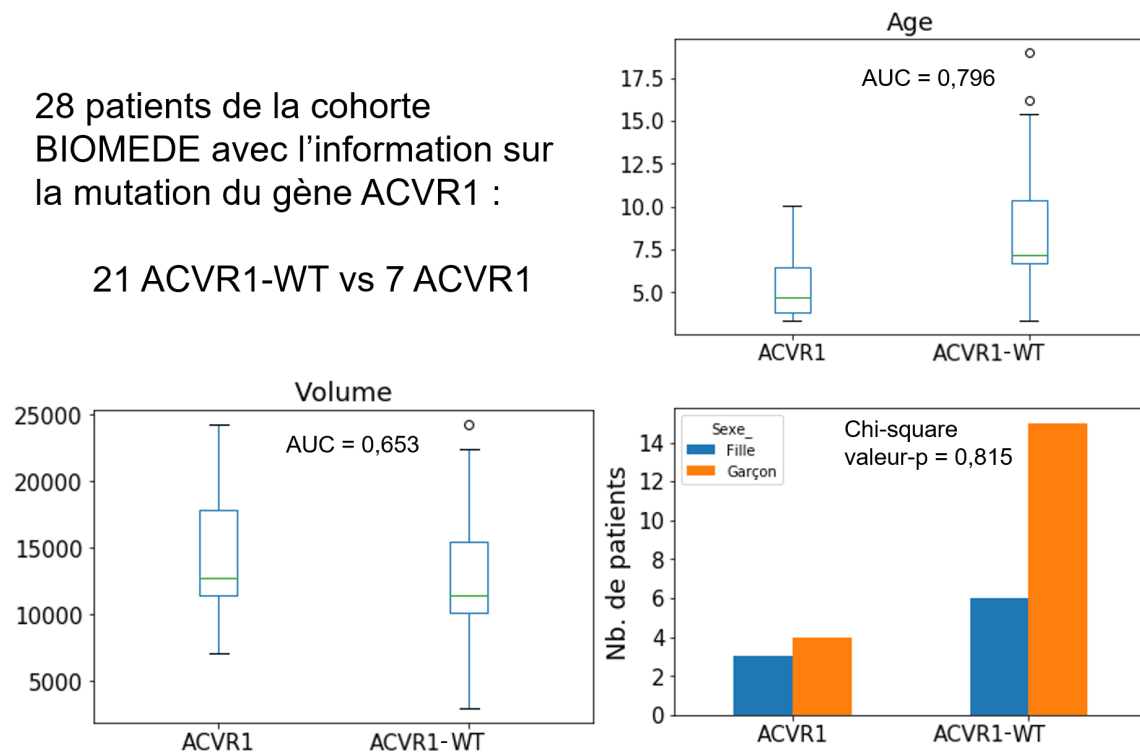


FIGURE 5.12 – Analyse des trois paramètres cliniques pour l'évaluation de la mutation du gène ACVR1 pour les 28 patients de la cohorte BIOMEDE

Indices d'imagerie

Les indices d'imagerie ont été évalués individuellement. L'AUC de la prédiction de la mutation du gène ACVR1 est présentée sur le tableau en annexe A.6 pour les 33 patients de la *cohorte radiomique 4Mod*. Parmi les indices calculés en FLAIR, un indice local et trois globaux présentent une AUC supérieure ou égale à 0,75. Parmi les indices calculés en T1c, deux indices locaux et deux globaux présentent une AUC supérieure ou égale à 0,75. Parmi les indices calculés en T1w, dix indices locaux et huit globaux présentent une AUC supérieure ou égale à 0,75. Parmi les indices calculés en T2w, aucun indice, ni local ni global ne présente une AUC supérieure ou égale à 0,75. Ainsi, la modalité T1w présente le plus grand nombre d'indices avec une AUC supérieure à 0,75. Pour cette *cohorte radiomique 4Mod*, ces indices présentent une AUC supérieure à celle de l'âge au diagnostic (0,690).

Parmi ces indices, trois présentent une AUC supérieure ou égale à 0,85. Ils sont affichés en vert dans le tableau A.6. Ces indices sont tous du premier ordre. Les boîtes à moustaches de ces trois indices sont présentées sur la figure 5.13.

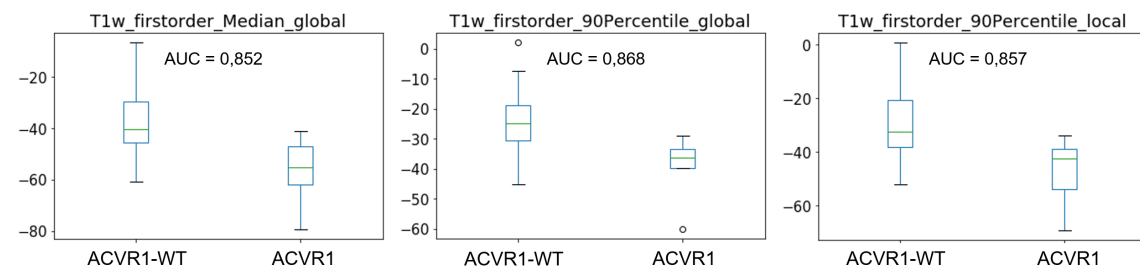


FIGURE 5.13 – Boîte à moustaches de trois indices d'imagerie qui présentent des AUCs élevées pour la prédiction de la mutation du gène ACVR1.

5.1.4 Synthèse des résultats des analyses uni-variées pour les trois mutations

Le tableau 5.2 résume les résultats obtenus avec les analyses uni-variées pour la classification des patients en fonction des trois mutations observées.

Mutation	Clinique	Imagerie
H3.1 vs. H3.3	Âge + Sexe - Volume +-	T1w - (firstorder +) T2w - T1c - (firstorder +) FLAIR +
TP53 vs. TP53-WT	Âge + Sexe - Volume +-	T1w - T2w + T1c - FLAIR +
ACVR1 vs. ACVR1-WT	Âge +- Sexe - Volume -	T1w - (firstorder +) T2w - T1c - FLAIR -

TABLEAU 5.2 – Résumé des résultats des analyses uni-variées. Les indices sont qualifiés comme de très bons prédicteurs (+), bons prédicteurs (+-) et des faibles prédicteurs (-).

5.2 Analyses multi-variées pour prédire les mutations génétiques associées au GITC

Après avoir analysé les tendances des indices cliniques et d'imagerie individuellement pour la classification des mutations, nous avons conduit des analyses multi-variées qui sont au cœur de la radiomique.

L'analyse radiomique multi-variée des maladies rares comme le **GITC** se trouve confrontée à un problème de faible effectif et de grande dimension (nombre d'indices caractérisant les individus très supérieure au nombre d'individus), ce qui peut conduire à des erreurs et à du sur-apprentissage.

La figure 5.14 montre un schéma des études multi-variées conduites pour chaque mutation. Les méthodes seront décrites avec plus de détails dans les paragraphes suivants, mais en bref, trois types de données d'entrée ont été utilisées : seulement les indices cliniques, seulement les indices d'imagerie ou les deux types d'indices combinés. Les analyses avec les indices globaux et locaux ont été faites séparément.

La première étape de l'approche multi-variée proposée dans cette thèse consiste à une procédure de sélection d'indices pertinents. Ensuite, les données sont sur-échantillonnées pour prendre en compte le déséquilibre des effectifs dans chaque classe (8 patients H3.1 versus 22 patients H3.3, par exemple) et les valeurs des indices sont normalisées par z-score. Finalement des modèles d'apprentissage automatique ont été mis en œuvre pour la classification des mutations associées à chaque patient.

Comme le nombre de sujets est réduit, toutes les analyses de performances menées sur la **cohorte radiomique 4Mod** ont été faites avec une **validation croisée avec exclusion d'un élément (LOOCV)** qui est plus adaptée aux petites cohortes que la validation croisée par pli. Les deux étapes : sélection d'indice et apprentissage de modèles sont à l'intérieur de la boucle de **LOOCV**.

Il faut rappeler que les analyses de prédiction des mutations ACVR1 et TP53 incluent 33 patients de la **cohorte radiomique 4Mod** alors que l'analyse de la mutation H3 inclut 30 patients de la **cohorte radiomique 4Mod**. Les 30 ou 33 plis de **LOOCV** sont répétés 30 fois pour prendre en considération le hasard introduit par certaines méthodes d'apprentissage utilisées et le sur-échantillonnage pour rééquilibrer les effectifs.

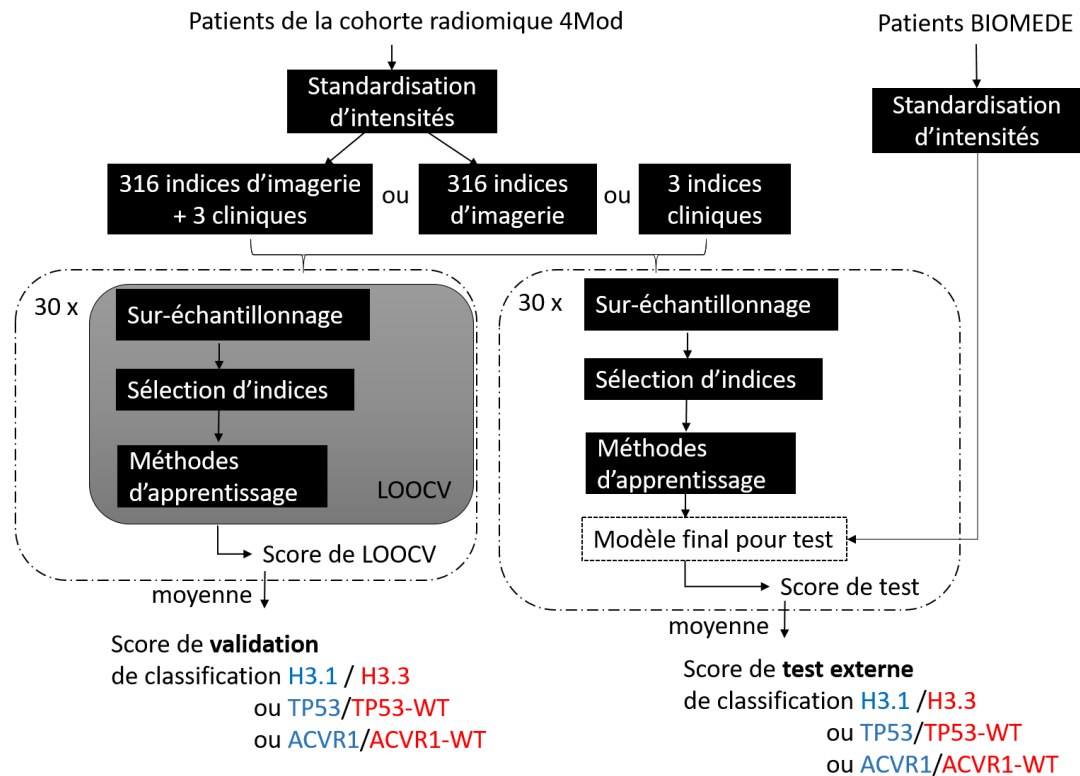


FIGURE 5.14 – Schéma des études multi-variées conduites pour chaque mutation et pour les ensembles d'indices d'imagerie globaux et locaux séparément.

Pour les raisons d'accès différé aux bases de données, les patients de la cohorte BIOMEDE ont été utilisés comme base de test de validation externe (indépendante, jamais utilisée dans la conception des modèles). Ainsi, les patients de la cohorte BIOMEDE sont classés avec les modèles conçus avec les patients de la [cohorte radiomique 4Mod](#). Comme pour la boucle de LOOCV, la conception des modèles finaux et le test externe sont répétés 30 fois pour prendre en considération le hasard introduit dans la conception des modèles par certaines méthodes d'apprentissage utilisées et le sur-échantillonnage.

5.2.1 Pipeline de sélection d'indices radiomiques proposé

Comme indiqué dans le chapitre 1 de cette thèse, il existe plusieurs méthodes pour réduire la dimension d'un problème de classification par la sélection des indices les plus pertinents. En imagerie, une étape importante est de sélectionner les indices qui sont robustes aux petits décalages de segmentation, car les délimitations des tumeurs tracées par plusieurs radiologues, ou par des méthodes automatiques et même par un seul radiologue à plusieurs reprises peuvent présenter des variabilités de plusieurs voxels. Dans les cas de classification supervisée, la pertinence d'un indice peut être aussi évaluée par sa performance individuelle de classification. Finalement, il faut aussi réduire la redondance entre indices. C'est ainsi qu'une sélection des variables en trois étapes a été proposée :

1. Élimination des indices sensibles à la segmentation
2. Élimination des indices de faible performance
3. Regroupement d'indices fortement corrélés

Ces trois étapes sont explicitées dans la suite du document. Elles ont été appliquées pour les études de classification des trois types de mutations H3, TP53 et ACVR1.

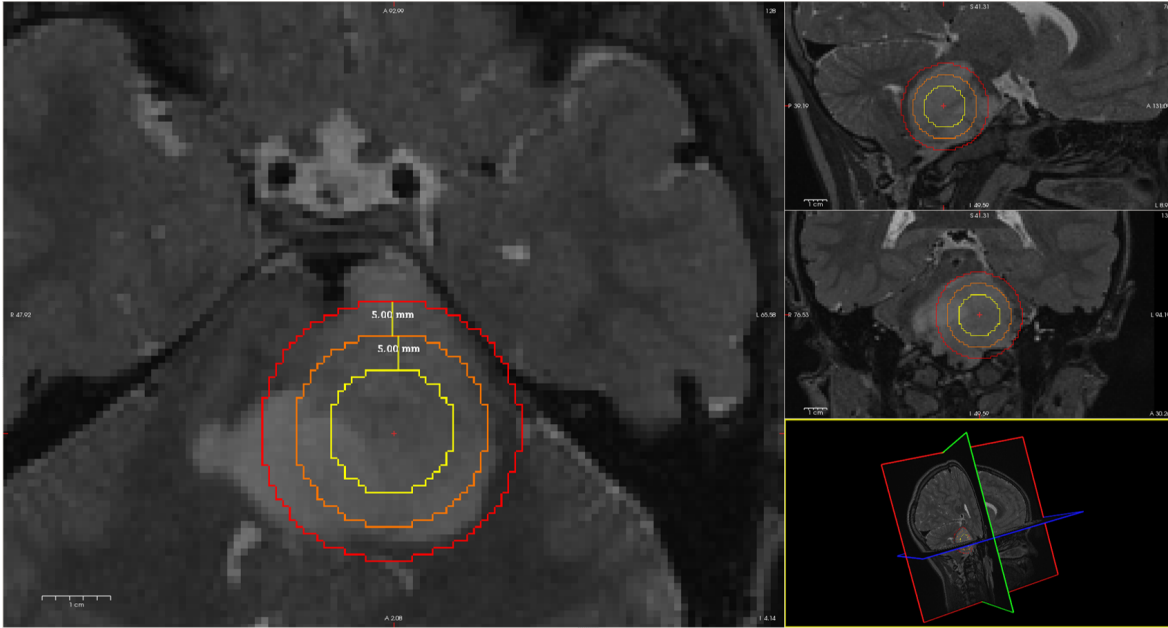


FIGURE 5.15 – Patient PREBIOMEDE sub-3037 : La VOI_{Tumeur} en orange, la VOI_{Tdil} en rouge et la VOI_{Tero} en jaune.

Robustesse à la segmentation

Pour notre étude, les tumeurs ont été approximées par des sphères (VOI_{Tumeur}). La sélection d'indices robustes à la segmentation s'est avérée cruciale, car ces sphères peuvent être très éloignées de la segmentation précise de la tumeur.

Pour cela, deux régions supplémentaires par tumeur ont été tracées : une version dilatée de la VOI_{Tumeur} (VOI_{Tdil}) et une version érodée (VOI_{Tero}). Les deux opérations ont été faites avec le paquet Python *scipy.ndimage* en utilisant comme structure d'érosion et dilatation une sphère de rayon 5 mm (figure 5.15).

Les 316 indices d'imagerie ont été calculés dans les deux régions supplémentaires pour les patients de la *cohorte radiomique 4Mod* de façon globale et locale et ont été comparés. Comme il existe plusieurs versions du *coefficient de corrélation intra-classe (CCI)*, il faut préciser que le critère "*ICC(2,1)*" (Koo et Li [2016]), défini par l'équation 5.1, a été utilisé.

$$ICC(2, 1) = \frac{MQ_L - MQ_C}{MQ_L + (k - 1)MQ_E + \frac{k}{n}(MQ_C - MQ_E)} \quad (5.1)$$

où MQ_L est la moyenne quadratique des valeurs selon les patients, MQ_C est la moyenne quadratique des valeurs selon les segmentations, MQ_E désigne la moyenne quadratique de l'erreur résiduelle entre les valeurs calculées dans les différentes régions, k est le nombre de segmentations et n est le nombre de patients. Cette définition est adaptée au problème de robustesse aux tracés de segmentation traité ici :

1. *Modèle* : elle considère que les trois segmentations sont des classes tirées aléatoirement d'un ensemble plus large et tente donc de généraliser pour différentes segmentations
2. *Définition* : elle considère que les valeurs des indices données par une seule définition de segmentation seront utilisées (VOI_{Tumeur}) et non pas la moyenne des valeurs des trois segmentations
3. *Type* : elle évalue l'accord absolu entre les classes (les erreurs systématiques et résiduelles comptent)

Dans l'article de Koo et Li [2016], il est suggéré qu'une valeur de *CCI* plus petite que 0,50 traduit un faible accord, entre 0,5 et 0,75 un accord modéré, entre 0,75 et 0,9 un bon accord et

supérieure à 0,9 un excellent accord. Les indices qui présentent un CCI supérieur que 0,9 sont gardés. Les tumeurs n'étant pas segmentées de façon précise, un seuil élevé a été choisi en vue de garantir que les indices calculés dans la VOI_{Tumeur} ne soient pas impactés par le tracé de la région et traduisent les caractéristiques de texture de la tumeur.

Pour illustration, lorsque les 38 patients de la [cohorte radiomique 4Mod](#) sont utilisés, les valeurs de CCI pour tous les indices sont présentés dans le [tableau 5.3](#). Les étapes de sélection d'indices sont appliquées pour les sous-échantillons de chaque pli de la LOOCV indépendamment, mais les résultats obtenus pour les différents plis dans chaque étude suivent les tendances des résultats présentés dans le [tableau 5.3](#). Ce critère de robustesse aide notamment à éliminer les indices fortement corrélés au volume. Ainsi, de nombreux indices de texture globaux, qui dépendent plus fortement du volume que les indices locaux, présentent une valeur faible de CCI.

Performance individuelle

Pour les problèmes de classification en fonction des mutations, les performances individuelles ont été évaluées par l'AUC de chaque indice. Le seuil pour définir les indices à garder est un compromis entre éliminer les indices qui peuvent dégrader les performances du modèle et garder les indices qui peuvent améliorer les performances du modèle même si leur performance individuelle n'est pas excellente.

Pour illustration, comme présenté dans l'analyse indice par indice, pour un seuil d'une AUC minimale de 0,75, par exemple, en utilisant la [cohorte radiomique 4Mod](#) : pour la mutation de l'histone H3, 48 indices locaux et 33 indices globaux passent le critère; pour la mutation du gène TP53, 37 indices locaux et 45 indices globaux passent le critère; pour la mutation du gène ACVR1, 12 indices locaux et 13 indices globaux passent le critère. Lorsque ce critère est appliqué dans les plis de la LOOCV, les mêmes tendances sont suivies.

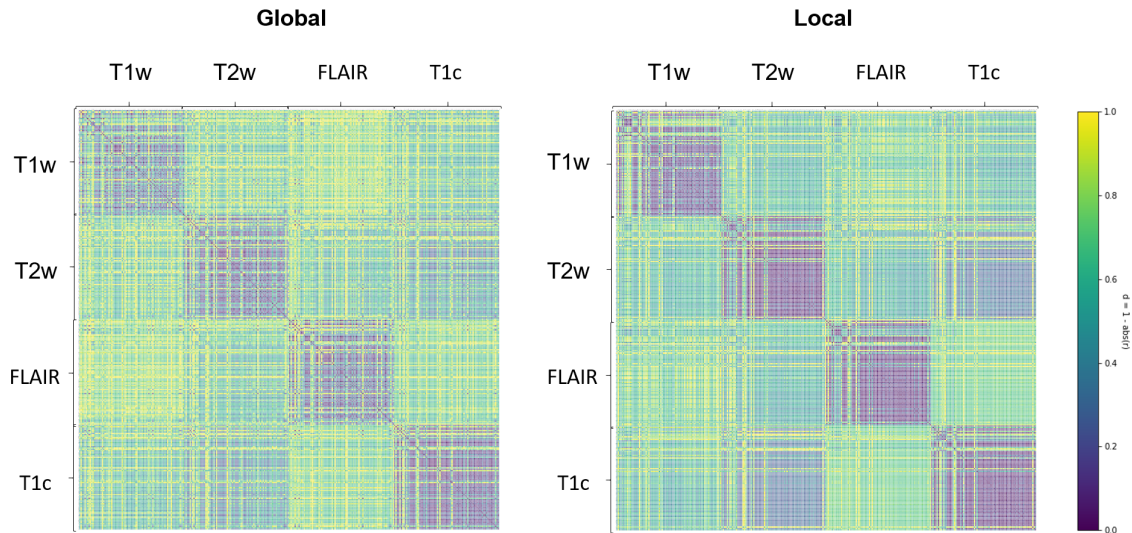
Redondance des indices

Enfin, les indices redondants sont éliminés par regroupement hiérarchique avec $d = 1 - abs(r)$ comme mesure de dissimilarité, où r est l'indice de corrélation de Spearman. Les groupes (en anglais *cluster*) d'indices ont été définis de sorte que tous les éléments du groupe aient $d \leq 0,15$, c'est à dire, que tous les éléments du groupe soient corrélés deux-à-deux avec $|r| > 0,85$. Le seuil est aussi un compromis entre éliminer les indices fortement corrélés et garder les indices qui peuvent être complémentaires. L'indice qui présente l'AUC la plus élevée de chaque groupe est gardé.

En utilisant les 38 patients de la [cohorte radiomique 4Mod](#), les cartes en couleurs présentent les valeurs de $d = 1 - abs(r)$ pour tous les 316 paires d'indices locaux et globaux présentées sur la [figure 5.16](#). On peut remarquer par exemple, qu'il y a une grande similarité entre plusieurs indices calculés dans la même modalité, et donc, les groupes sont souvent formés intra-modalité.

TABLEAU 5.3 – CCI des indices d'imagerie. Toutes les valeurs supérieures ou égales à 0,90 sont affichées en vert.

	T1w		T2w		T1c		FLAIR	
	global	local	global	local	global	local	global	local
firstorder.InterquartileRange	0,70	0,81	0,66	0,94	0,80	0,97	0,82	0,96
firstorder.Skewness	0,78	0,84	0,57	0,85	0,40	0,87	0,54	0,82
firstorder.Uniformity	0,75	0,91	0,89	0,95	0,86	0,96	0,87	0,96
firstorder.MeanAbsoluteDeviation	0,72	0,82	0,77	0,95	0,90	0,98	0,77	0,96
firstorder.Energy	0,76	0,98	0,76	0,98	0,76	0,96	0,76	0,98
firstorder.RobustMeanAbsoluteDeviation	0,69	0,81	0,69	0,95	0,83	0,97	0,81	0,96
firstorder.Median	0,96	0,98	0,98	0,98	0,94	0,94	0,97	0,98
firstorder.TotalEnergy	0,76	0,98	0,76	0,98	0,76	0,96	0,76	0,98
firstorder.Maximum	0,87	0,91	0,79	0,99	0,70	0,98	0,99	0,97
firstorder.RootMeanSquared	0,95	0,98	0,96	0,98	0,93	0,96	0,95	0,98
firstorder.90Percentile	0,71	0,95	0,98	0,98	0,89	0,97	0,98	0,98
firstorder.Minimum	0,84	0,99	0,44	0,96	0,89	0,99	0,39	0,93
firstorder.Entropy	0,80	0,90	0,89	0,95	0,92	0,98	0,89	0,96
firstorder.StandardDeviation	0,74	0,83	0,81	0,95	0,93	0,98	0,75	0,95
firstorder.Range	0,82	0,88	0,66	0,96	0,78	0,99	0,76	0,95
firstorder.Variance	0,66	0,71	0,74	0,96	0,91	0,98	0,59	0,95
firstorder.10Percentile	0,99	0,99	0,80	0,98	0,98	0,98	0,72	0,97
firstorder.Kurtosis	0,81	0,57	0,36	0,77	0,00	0,79	0,11	0,62
firstorder.Mean	0,95	0,98	0,96	0,98	0,93	0,96	0,94	0,98
glcm.SumVariance	0,33	0,77	0,52	0,89	0,91	0,98	0,41	0,85
glcm.Homogeneity1	0,94	0,97	0,94	0,96	0,94	0,98	0,95	0,97
glcm.Homogeneity2	0,94	0,97	0,95	0,96	0,94	0,98	0,96	0,98
glcm.ClusterShade	0,21	0,38	0,42	0,49	0,87	0,93	0,24	0,49
glcm.MaximumProbability	0,82	0,95	0,91	0,97	0,87	0,96	0,92	0,97
glcm.Idmn	0,64	0,89	0,39	0,75	0,65	0,86	0,38	0,82
glcm.SumVariance2	0,67	0,68	0,78	0,96	0,91	0,98	0,64	0,95
glcm.Contrast	0,75	0,93	0,79	0,95	0,92	0,96	0,73	0,95
glcm.DifferenceEntropy	0,90	0,97	0,90	0,96	0,95	0,98	0,90	0,97
glcm.InverseVariance	0,94	0,98	0,95	0,96	0,94	0,98	0,96	0,98
glcm.Entropy	0,85	0,94	0,91	0,96	0,94	0,98	0,92	0,97
glcm.Dissimilarity	0,90	0,96	0,89	0,95	0,93	0,97	0,89	0,96
glcm.DifferenceVariance	0,59	0,90	0,72	0,96	0,96	0,98	0,59	0,95
glcm.Idn	0,70	0,90	0,38	0,73	0,48	0,83	0,37	0,80
glcm.Idm	0,94	0,97	0,95	0,96	0,94	0,98	0,96	0,98
glcm.Correlation	0,75	0,87	0,60	0,77	0,64	0,87	0,69	0,91
glcm.Autocorrelation	0,35	0,79	0,53	0,89	0,91	0,98	0,42	0,86
glcm.SumEntropy	0,80	0,89	0,90	0,95	0,92	0,97	0,90	0,96
glcm.AverageIntensity	0,58	0,90	0,60	0,92	0,91	0,98	0,45	0,91
glcm.Energy	0,76	0,94	0,92	0,97	0,85	0,96	0,87	0,97
glcm.SumSquares	0,67	0,73	0,78	0,96	0,91	0,97	0,64	0,95
glcm.ClusterProminence	0,54	0,34	0,44	0,85	0,92	0,96	0,15	0,73
glcm.SumAverage	0,58	0,89	0,60	0,92	0,91	0,98	0,45	0,91
glcm.Imc2	0,75	0,85	0,78	0,93	0,70	0,94	0,73	0,93
glcm.Imc1	0,87	0,82	0,76	0,92	0,75	0,97	0,83	0,93
glcm.DifferenceAverage	0,90	0,96	0,89	0,95	0,93	0,97	0,89	0,96
glcm.Id	0,94	0,97	0,94	0,96	0,94	0,98	0,95	0,97
glcm.ClusterTendency	0,67	0,68	0,78	0,96	0,91	0,98	0,64	0,95
grlm.ShortRunLowGrayLevelEmphasis	0,62	0,96	0,50	0,94	0,79	0,98	0,46	0,92
grlm.GrayLevelVariance	0,66	0,71	0,73	0,96	0,92	0,98	0,59	0,94
grlm.LowGrayLevelRunEmphasis	0,62	0,96	0,51	0,94	0,79	0,98	0,45	0,92
grlm.GrayLevelNonUniformityNormalized	0,75	0,90	0,89	0,95	0,86	0,96	0,86	0,96
grlm.RunVariance	0,95	0,97	0,97	0,98	0,96	0,98	0,97	0,98
grlm.GrayLevelNonUniformity	0,78	0,90	0,82	0,95	0,84	0,96	0,85	0,96
grlm.LongRunEmphasis	0,94	0,97	0,97	0,97	0,96	0,98	0,97	0,98
grlm.ShortRunHighGrayLevelEmphasis	0,38	0,80	0,54	0,90	0,92	0,98	0,43	0,87
grlm.RunLengthNonUniformity	0,75	0,97	0,75	0,97	0,76	0,98	0,75	0,98
grlm.ShortRunEmphasis	0,94	0,97	0,96	0,97	0,95	0,98	0,97	0,98
grlm.LongRunHighGrayLevelEmphasis	0,33	0,79	0,50	0,89	0,89	0,98	0,37	0,85
grlm.RunPercentage	0,95	0,97	0,97	0,97	0,95	0,98	0,97	0,98
grlm.LongRunLowGrayLevelEmphasis	0,62	0,96	0,53	0,95	0,81	0,98	0,45	0,93
grlm.RunEntropy	0,78	0,87	0,87	0,95	0,91	0,98	0,86	0,96
grlm.HighGrayLevelRunEmphasis	0,37	0,79	0,53	0,90	0,91	0,98	0,42	0,87
grlm.RunLengthNonUniformityNormalized	0,94	0,97	0,96	0,97	0,95	0,98	0,97	0,98
glszm.GrayLevelVariance	0,65	0,74	0,69	0,96	0,94	0,98	0,52	0,93
glszm.LowIntensityLargeAreaEmphasis	0,68	0,91	0,84	0,96	0,89	0,96	0,56	0,94
glszm.HighIntensitySmallAreaEmphasis	0,42	0,80	0,60	0,92	0,95	0,99	0,45	0,89
glszm.SmallAreaEmphasis	0,75	0,96	0,75	0,94	0,90	0,98	0,67	0,95
glszm.LargeAreaEmphasis	0,88	0,91	0,87	0,97	0,89	0,95	0,96	0,97
glszm.ZoneVariance	0,87	0,91	0,86	0,98	0,88	0,94	0,96	0,97
glszm.SizeZoneVariabilityNormalized	0,75	0,96	0,76	0,94	0,91	0,98	0,66	0,95
glszm.LowIntensitySmallAreaEmphasis	0,59	0,94	0,50	0,96	0,68	0,97	0,56	0,93
glszm.HighIntensityEmphasis	0,42	0,79	0,56	0,91	0,94	0,99	0,44	0,88
glszm.IntensityVariabilityNormalized	0,78	0,88	0,84	0,95	0,85	0,96	0,71	0,96
glszm.ZonePercentage	0,94	0,97	0,94	0,96	0,94	0,98	0,94	0,98
glszm.LowIntensityEmphasis	0,59	0,94	0,51	0,96	0,71	0,98	0,54	0,93
glszm.SizeZoneVariability	0,62	0,97	0,70	0,95	0,80	0,98	0,64	0,97
glszm.IntensityVariability	0,82	0,97	0,84	0,86	0,79	0,97	0,87	0,94
glszm.ZoneEntropy	0,82	0,90	0,88	0,96	0,87	0,97	0,82	0,96
glszm.HighIntensityLargeAreaEmphasis	0,38	0,93	0,40	0,73	0,50	0,95	0,49	0,86


 FIGURE 5.16 – Matrices de dissimilarité ($d = 1 - |r|$) des indices d'imagerie locaux et globaux.

5.2.2 Modèles d'apprentissage automatique utilisés

Pour les N plis de **LOOCV** les indices ont été standardisés avec la soustraction de la moyenne et la division par l'écart-type de chaque paramètre en utilisant les $N-1$ patients d'entraînement. Les mêmes valeurs des paramètres de standardisation ont été utilisées pour standardiser les indices du patient de validation. Comme dans les trois cas de classification des mutations, les proportions de patients porteurs des différentes mutations ne sont pas équilibrés (ACVR1 : 26 vs. 7; TP53 : 20 vs. 13; H3 : 22 vs. 8), les patients d'entraînement des classes minoritaires ont été sur-échantillonnés avec la technique *Synthetic minority Over-sampling Technique (SMOTE)* (CHAWLA et collab. [2002]) à chaque pli de **LOOCV**. L'avantage de **SMOTE** par rapport à la technique de sur-échantillonnage conventionnelle est qu'il simule des nouveaux échantillons au lieu de copier des échantillons existants.

L'analyse de la performance des classifications a été faite par l'indice de Youden et le score F1-pondéré en répétant la **LOOCV** 30 fois pour estimer aussi les écart-types des scores de prédiction dus aux aspects aléatoires de la méthode de sur-échantillonnage **SMOTE**, des régressions logistiques et des forêts aléatoires.

Comme il n'existe pas de consensus sur la meilleure méthode d'apprentissage, car leurs performances dépendent de la base de données, nous avons choisi de comparer les performances de quatre méthodes différentes largement employées dans la littérature pour les modèles de classifications : la régression logistique, les k plus proches voisins, la machine à vecteurs de support et les forêts aléatoires. Toutes les méthodes ont été employées en utilisant le paquet logiciel *scikit-learn* en Python.

La régression logistique (Logit) : Dans cette étude, pour la minimisation de la vraisemblance la pénalisation L2 a été utilisée et le paramètre de pénalité C a été fixé à 1.

Les k plus proches voisins (KNN) : Dans cette étude, la distance euclidienne et $k = 3$ avec des poids uniformes ont été utilisés.

Les machines à vecteurs de support (SVM) : Dans cette étude, un noyau linéaire et le paramètre de pénalité $C = 1$ ont été utilisés.

Les forêts aléatoires (RF) : Dans cette étude, le critère de répartition des sujets à chaque nœud d'un arbre a été l'index de Gini. Pour les classifications des trois mutations associées au **GITC**, les

forêts aléatoires étaient composées de 100 arbres, l'arbre étant construit jusqu'à que les nœuds terminaux soient purs, le nombre d'indices considérés à chaque nœud pour être le séparateur est la racine carré du nombre total d'indices, les patients considérés à chaque arbre sont tirés avec remplacement ("*bootstrap*").

5.2.3 Application des modèles de classification pour prédire la mutation de l'histone H3

Sélection d'indice

Les étapes de la sélection d'indices sont, comme déjà annoncé précédemment, faites dans chaque pli de la **LOOCV**. Avec les 30 patients présentant soit une mutation H3.1 soit une mutation H3.3 de la **cohorte radiomique 4Mod**, en moyenne, 7 ± 1 indices d'imagerie locaux et 4 ± 1 indices globaux sont sélectionnés après les trois étapes de la procédure de sélection. Les indices sélectionnés dans chaque pli de la **LOOCV** sont souvent les mêmes. Cinq indices locaux (*T1w firstorder Minimum*, *T1w firstorder 90Percentile*, *FLAIR glrlm LongRunLowGrayLevelEmphasis*, *T1c firstorder Minimum*, *T1c glcm ClusterShade*) et trois globaux (*T1c firstorder 10Percentile*, *FLAIR glcm Homogeneity1*, *T1w firstorder Median*) sont sélectionnés dans plus de 25 plis sur 30. Parmi les trois indices cliniques, l'âge au diagnostic a été sélectionné dans tous les plis, le volume et le sexe jamais.

Les indices les plus souvent sélectionnés correspondent bien aux appréciations qualitatives des images faites par les radiologues. Les tumeurs correspondant à une mutation H3.1 présentent souvent une intensité plus faible dans les images **T1w** (*T1w firstorder Minimum*, *T1w firstorder 90Percentile*, *T1w firstorder Median*) et une région moins homogène et plus intense dans les images **FLAIR** (*FLAIR glrlm LongRunLowGrayLevelEmphasis*, *FLAIR glcm Homogeneity1*). De plus, les tumeurs présentant une mutation H3.1 présentent plus souvent des nécroses centrales et des zones périphériques intenses en forme d'anneau dans les images **T1c** (*T1c firstorder Minimum*, *T1c glcm ClusterShade*, *T1c firstorder 10Percentile*). Quelques exemples d'images **FLAIR** et **T1c** de patients issus des deux groupes sont présentés sur la figure 5.17 où les différences des caractéristiques sélectionnées peuvent être bien observées.

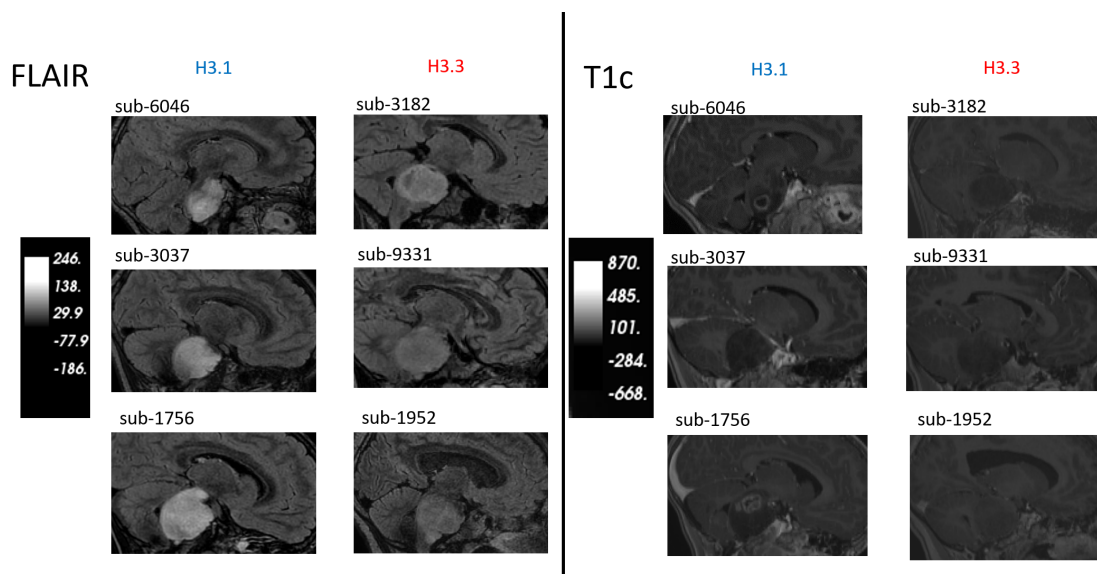


FIGURE 5.17 – Exemple d'images T1c et FLAIR de patients de la cohorte radiomique 4Mod présentant soit une mutation H3.1 soit une mutation H3.3.

La figure 5.18 présente le nombre moyen d'indices d'imagerie sélectionnés par pli de la **LOOCV** dans chaque étape de la procédure de sélection d'indice (et l'écart-type correspondant) pour les deux types d'indices, locaux et globaux. On peut noter, comme déjà indiqué, que plus d'indices

globaux sont éliminés par l'étape 1. Cela est cohérent car les indices corrélés au volume ont tendance à changer entre les différentes définitions de VOI_{tumeur} .

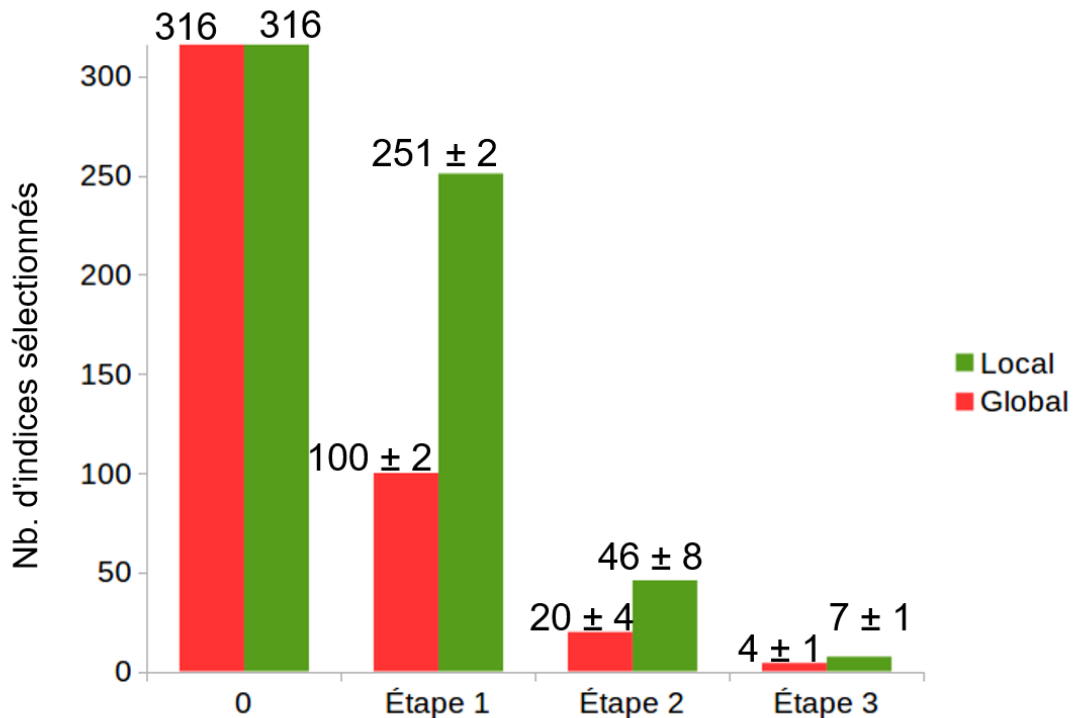


FIGURE 5.18 – Nombre moyen d'indices sélectionnés à chaque étape de la procédure de sélection d'indices pour les 30 patients de la cohorte 4Mod qui présentent une mutation de l'histone H3 (variantes H3.1 ou H3.3).

Classification (validation)

Le tableau 5.4 présente les évaluations des quatre modèles selon la sensibilité, la spécificité, l'index de Youden et le score F1-pondéré. Les valeurs moyennes et les écart-types des scores de 30 répétitions de LOOCV sont affichés.

La figure 5.19 montre les scores F1-pondérés de LOOCV pour chaque modèle de classification H3.1 vs H3.3 avec les indices d'imagerie, locaux ou globaux, conservés par les trois étapes de sélection d'indices. Chaque point représente une des 30 répétitions de la procédure LOOCV. La moyenne et l'écart-type sont aussi affichés. La valeur de F1-pondéré si tous les patients de la cohorte radiomique 4Mod étaient classifiés comme H3.3 (la classe la plus fréquente) serait égale à 0,62. Cette valeur est tracée sur la figure pour faciliter l'interprétation des résultats.

TABLEAU 5.4 – Résultats de validation obtenus avec les différents modèles utilisant seulement les indices cliniques, d'imagerie (local ou global) et combinés (imagerie et clinique).

		RF	KNN	SVM	Logit
Clinique	Se	0,74 ± 0,04	0,64 ± 0,10	0,61 ± 0,05	0,66 ± 0,03
	Sp	0,75 ± 0,08	0,75 ± 0,04	0,88 ± 0,00	0,88 ± 0,00
	Youden	0,49 ± 0,10	0,38 ± 0,10	0,49 ± 0,05	0,54 ± 0,03
	F1-pond	0,74 ± 0,04	0,68 ± 0,07	0,70 ± 0,04	0,73 ± 0,02
Global					
Imagerie	Se	0,82 ± 0,06	0,76 ± 0,08	0,75 ± 0,04	0,76 ± 0,04
	Sp	0,50 ± 0,06	0,50 ± 0,02	0,69 ± 0,09	0,75 ± 0,00
	Youden	0,32 ± 0,09	0,25 ± 0,07	0,44 ± 0,11	0,51 ± 0,04
	F1-pond	0,74 ± 0,04	0,70 ± 0,05	0,74 ± 0,04	0,77 ± 0,03
Combiné	Se	0,84 ± 0,03	0,88 ± 0,02	0,88 ± 0,03	0,83 ± 0,03
	Sp	0,57 ± 0,09	0,68 ± 0,08	0,63 ± 0,02	0,74 ± 0,04
	Youden	0,41 ± 0,09	0,57 ± 0,08	0,51 ± 0,04	0,57 ± 0,05
	F1-pond	0,77 ± 0,03	0,83 ± 0,03	0,81 ± 0,02	0,81 ± 0,02
Local					
Imagerie	Se	0,83 ± 0,02	0,71 ± 0,03	0,76 ± 0,04	0,74 ± 0,02
	Sp	0,59 ± 0,06	0,57 ± 0,08	0,49 ± 0,10	0,62 ± 0,07
	Youden	0,42 ± 0,07	0,28 ± 0,08	0,25 ± 0,10	0,35 ± 0,07
	F1-pond	0,77 ± 0,02	0,68 ± 0,03	0,69 ± 0,04	0,72 ± 0,02
Combiné	Se	0,85 ± 0,02	0,82 ± 0,02	0,85 ± 0,03	0,84 ± 0,02
	Sp	0,58 ± 0,08	0,69 ± 0,09	0,63 ± 0,00	0,64 ± 0,04
	Youden	0,43 ± 0,09	0,51 ± 0,10	0,48 ± 0,03	0,47 ± 0,05
	F1-pond	0,78 ± 0,03	0,79 ± 0,03	0,81 ± 0,01	0,79 ± 0,02

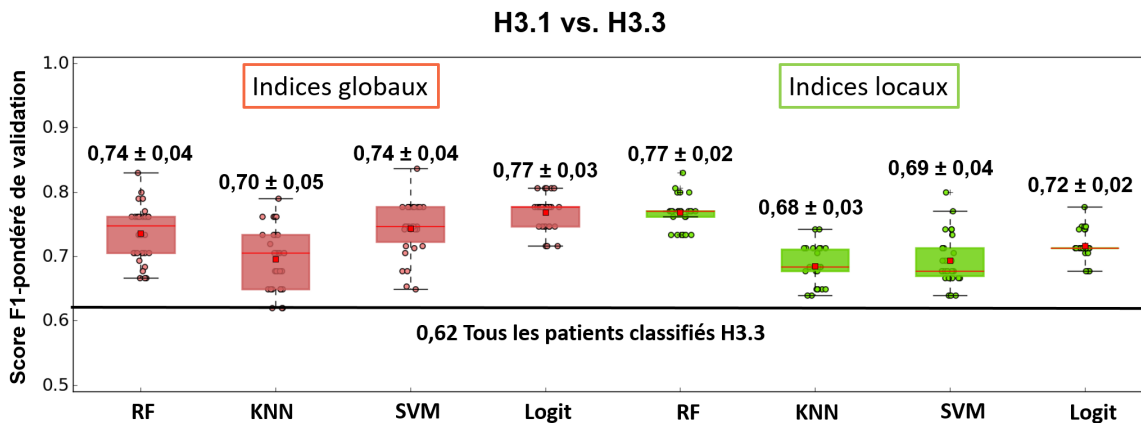


FIGURE 5.19 – Les boîtes à moustaches des valeurs des scores F1-pondérés obtenus pour les 30 répétitions de la LOOCV. La moyenne et l'écart-type des scores F1-pondérés de validation estimés sont donnés pour les 4 modèles pour les indices d'imagerie sélectionnés. Le score d'un modèle trivial, classant tous les patients comme H3.3 est affiché.

Tous les modèles sont plus performants que le modèle trivial de classification où tous les patients seraient classés comme H3.3. Différents modèles donnent des résultats quasi équivalents. La figure 5.20 montre que les modèles combinant les indices d'imagerie et les indices cliniques ont des performances systématiquement meilleures à celles des modèles cliniques et d'imagerie seuls pour les quatre méthodes d'apprentissage en évaluant avec le score F1-pondéré.

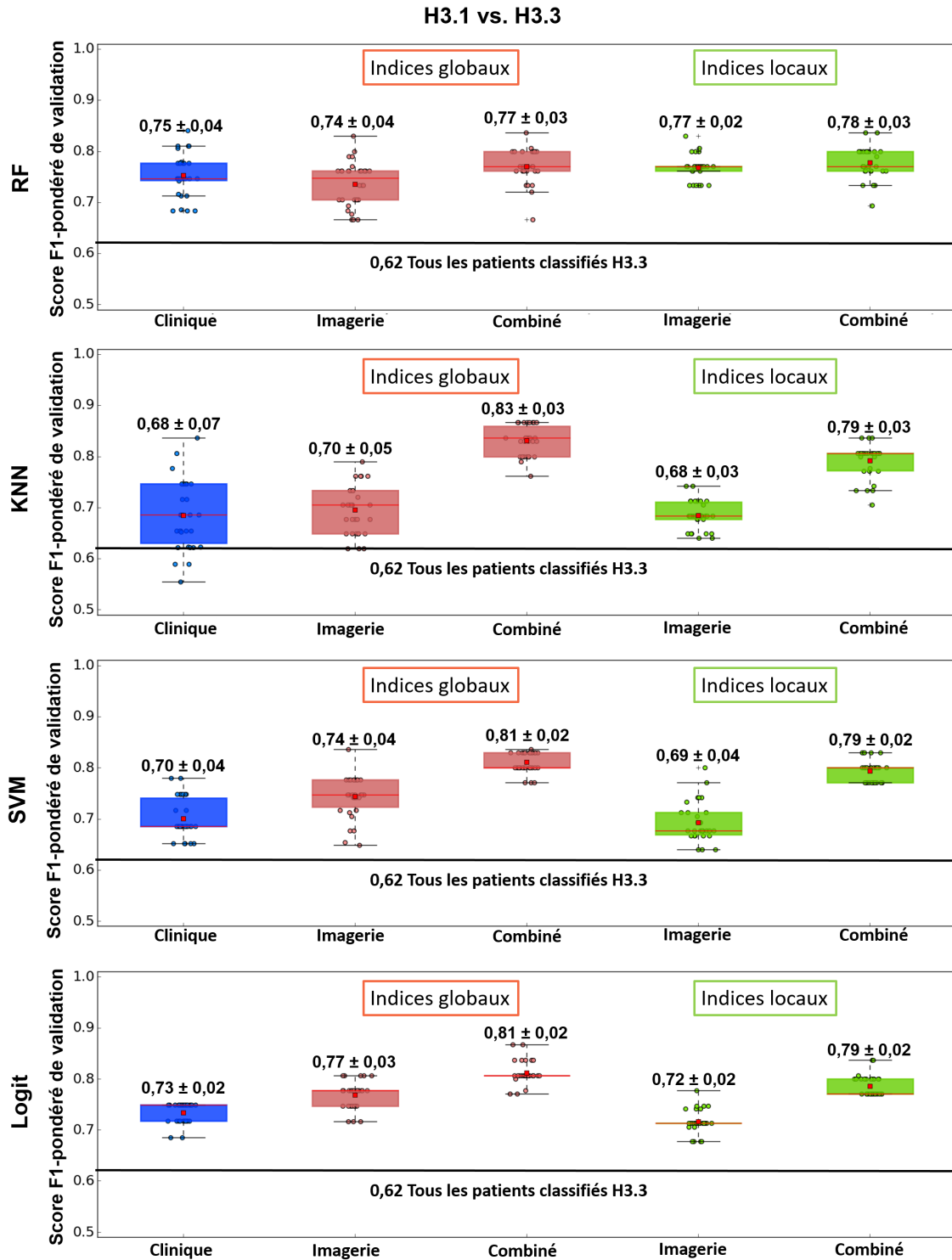


FIGURE 5.20 – Les boîtes à moustaches des valeurs des scores F1-pondérés obtenus pour les 30 répétitions de la LOOCV. La moyenne et l'écart-type des scores F1-pondérés de validation estimés sont donnés pour les 4 modèles pour les indices cliniques seuls, les indices d'imagerie sélectionnés seuls et la combinaison des indices cliniques et d'imagerie. Le score d'un modèle trivial, classant tous les patients comme H3.3 est affiché.

Ce comportement de modèles combinés est moins remarquable pour les indices de Youden car chaque cas H3.1 bien classé augmente plus la valeur de l'indice de Youden que le score F1-pondéré tandis que chaque cas H3.3 mal classé détériore beaucoup moins l'indice de Youden que le score F1-pondéré. De cette façon, même si la spécificité du modèle clinique est toujours plus

faible que les spécificités des modèles d'imagerie et combiné (voir tableau 5.4), l'indice de Youden peut-être supérieur (par exemple, pour les modèles RF) car la sensibilité est plus élevée. Avec la cohorte utilisé, chaque patient H3.1 bien classé augmente la sensibilité de 0,125.

Classification (test)

Un score de test se base sur des patients jamais vus par le modèle. Ainsi, pour créer les modèles, nous avons utilisé tous les patients de la [cohorte radiomique 4Mod](#) et pour les tester, les 17 patients BIOMEDE qui étaient imagés avant biopsie par les modalités retenues dans le modèle final et pour lesquels le statut de la mutation de l'histone H3 était connu. Le test a été répété 30 fois pour prendre en compte les variations aléatoires de [SMOTE](#), des forêts aléatoires et des régressions logistiques.

Les indices d'imagerie globaux et cliniques sélectionnés dans un modèle combiné, qui sont globalement les plus performants en validation, sont *FLAIR GLCM Homogeneity*, *T1w firstorder Median*, *T1c firstorder 10Percentile* et l'âge au diagnostic. Les quatre méthodes d'apprentissage automatique ont été utilisées. Le même a été fait en utilisant les indices locaux. Les indices locaux sélectionnés dans le modèle combiné sont *T1c firstorder Minimum*, *T1w firstorder 90Percentile*, *T1w firstorder Maximum*, *T1c glszm SizeZoneVariabilityNormalized*, *T1c glcm ClusterShade*, *FLAIR glrlm LongRunLowGrayLevelEmphasis*, *FLAIR glszm LowIntensityEmphasis* et l'âge au diagnostic. Donc, pour les deux types de modèles les indices les plus fréquemment sélectionnés dans les couches de [LOOCV](#) sont sélectionnés dans le modèle final, les images **T2w** ne sont pas nécessaires. Parmi les patients BIOMEDE, 17 présentent des images **T1w**, **T1c** et **FLAIR** de diagnostic et étaient mutés H3.1 (4) ou H3.3 (13). Le tableau 5.5 rappelle les scores de validation (cf. tableau 5.4) et donne les scores de test.

TABLEAU 5.5 – Résultats de test des modèles H3 combinés (global et local).

		RF	KNN	SVM	Logit
Global					
Validation (LOOCV)	Se	0,84 ± 0,03	0,88 ± 0,02	0,88 ± 0,03	0,83 ± 0,03
	Sp	0,57 ± 0,09	0,68 ± 0,08	0,63 ± 0,02	0,74 ± 0,04
	Youden	0,41 ± 0,09	0,57 ± 0,08	0,51 ± 0,04	0,57 ± 0,05
	F1-pond	0,77 ± 0,03	0,83 ± 0,03	0,81 ± 0,02	0,81 ± 0,02
Test BIOMEDE	Se	0,78 ± 0,09	0,79 ± 0,03	0,92 ± 0,01	0,92 ± 0,01
	Sp	0,48 ± 0,27	0,27 ± 0,17	0,18 ± 0,17	0,18 ± 0,17
	Youden	0,25 ± 0,32	0,05 ± 0,16	0,10 ± 0,17	0,10 ± 0,17
	F1-pond	0,71 ± 0,11	0,66 ± 0,04	0,70 ± 0,07	0,70 ± 0,07
Local					
Validation (LOOCV)	Se	0,85 ± 0,02	0,82 ± 0,02	0,85 ± 0,03	0,84 ± 0,02
	Sp	0,58 ± 0,08	0,69 ± 0,09	0,63 ± 0,00	0,64 ± 0,04
	Youden	0,43 ± 0,09	0,51 ± 0,10	0,48 ± 0,03	0,47 ± 0,05
	F1-pond	0,78 ± 0,03	0,79 ± 0,03	0,81 ± 0,01	0,79 ± 0,02
Test BIOMEDE	Se	0,68 ± 0,08	0,83 ± 0,05	0,74 ± 0,05	0,75 ± 0,04
	Sp	0,19 ± 0,20	0,27 ± 0,09	0,10 ± 0,12	0,13 ± 0,12
	Youden	-0,13 ± 0,22	0,09 ± 0,12	-0,16 ± 0,10	-0,12 ± 0,12
	F1-pond	0,57 ± 0,07	0,68 ± 0,05	0,58 ± 0,03	0,60 ± 0,04

Dans les scores de test, les modèles ont des difficultés à détecter les 4 patients H3.1, en général seulement un patient sur quatre est bien identifié comme muté en H3.1 (spécificité de test égale à environ 0,25) alors que les modèles sont bien adaptés à identifier les patients d'entraînement mutés en H3.1 (sensibilité d'entraînement égale à environ 1,00). Seules les forêts aléatoires avec les indices globaux arrivent à bien classer en moyenne deux patients mutés H3.1. D'autre part, presque tous les patients mutés H3.3 sont bien classés (les valeurs de sensibilité sont élevées). Les

modèles utilisant les indices d'imagerie globaux présentent en général de meilleurs résultats. La valeur du score F1-pondéré pour tous les patients classés comme mutés en H3.3 pour cette cohorte est égal à 0,66; donc, tous les modèles globaux sont équivalents ou meilleurs que le modèle trivial.

La distribution des patients d'entraînement et de test selon les indices sélectionnés pour le modèle global est présentée sur la figure 5.21. Les valeurs des indices affichées sont normalisées par la moyenne et l'écart-type des valeurs de la *cohorte radiomique 4Mod*. Les patients d'entraînement sont représentés par des losanges et les patients de test par des cercles et ils sont colorés en fonction des mutations H3.1 (en rouge) et H3.3 (en bleu).

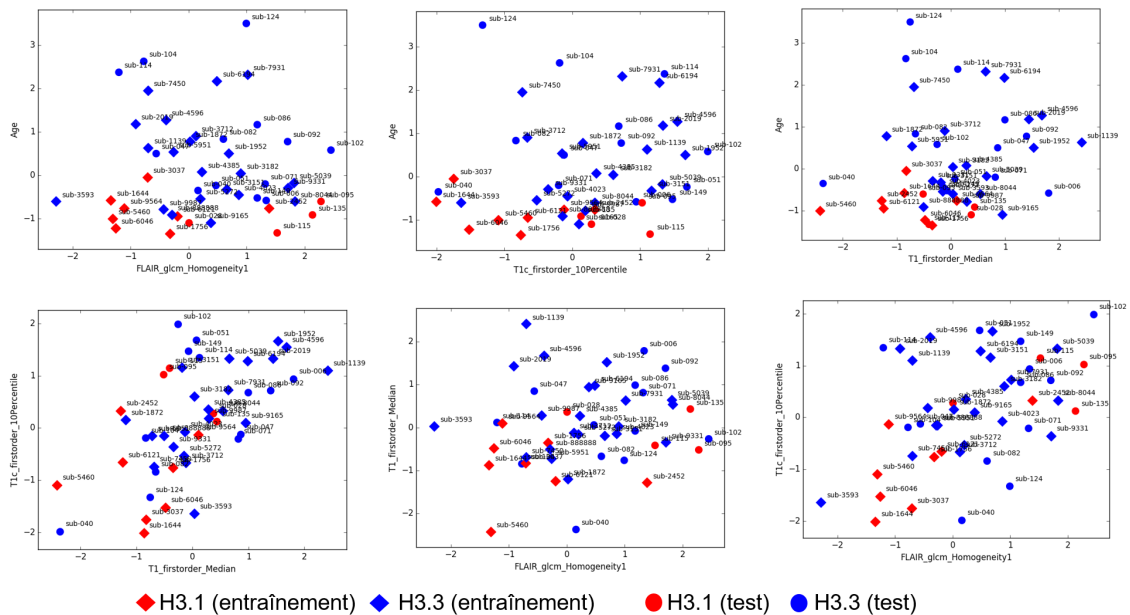


FIGURE 5.21 – Distribution des patients d'entraînement et de test selon les indices sélectionnés pour le modèle global de la mutation H3. Les valeurs des indices sont normalisés. Les patients d'entraînement sont représentés par des losanges et les patients de test par des cercles et ils sont colorés par leur mutation H3.1 (en rouge) et H3.3 (en bleu).

Les patients de la *cohorte radiomique 4Mod* présentant une mutation H3.1 ont des valeurs plus faibles pour les quatre indices (*FLAIR GLCM Homogeneity*, *T1w firstorder Median*, *T1c firstorder 10Percentile* et âge au diagnostic) alors que les quatre patients BIOMEDE présentant une mutation H3.1 se mélangent avec les autres patients présentant une mutation H3.3 pour les indices d'imagerie en *T1w* et en *T1c*. De plus, les trois patients sub-095, sub-115 et sub-135 présentant une mutation H3.1 ont une valeur élevée de *FLAIR GLCM Homogeneity* et ont ainsi un comportement inversé par rapport à la description qui était faible avec les patients de la base d'apprentissage.

Ainsi, les relations entre les indices d'imagerie et la mutation H3 trouvées pour les patients de la *cohorte radiomique 4Mod* ne se sont pas très bien reproduites pour la cohorte de test BIOMEDE que nous avons eue à analyser.

5.2.4 Application des modèles de classification pour la mutation du gène TP53

Sélection d'indice

Avec les 33 patients de la *cohorte radiomique 4Mod* pour lesquels l'information sur la mutation en TP53 était disponible, en moyenne, 5 ± 1 indices d'imagerie locaux et 3 ± 1 indices globaux sont sélectionnés après les trois étapes de la procédure de sélection. Deux indices locaux (*FLAIR firstorder Median* et *T1c glszm HighIntensityLargeAreaEmphasis*) et trois globaux (*T1c firstorder 10Percentile*, *FLAIR firstorder Median*, *T1c firstorder Median*) sont sélectionnés dans plus de 27

plis sur 33. Les indices qui présentaient des bonnes performances individuelles ont été éliminés par le critère de robustesse à la segmentation. En plus, parmi les trois indices cliniques, l'âge au diagnostic passe le seuil de 0,75 de AUC 3 fois dans les 33 plis, et le volume et le sexe jamais.

La figure 5.22 montre le nombre moyen d'indices d'imagerie sélectionnés par couche de la LOOCV dans chaque étape de la procédure de sélection d'indice (et l'écart type correspondant) pour les deux types d'indices, locaux et globaux. Une quantité importante d'indices globaux sont éliminés par l'étape 1.

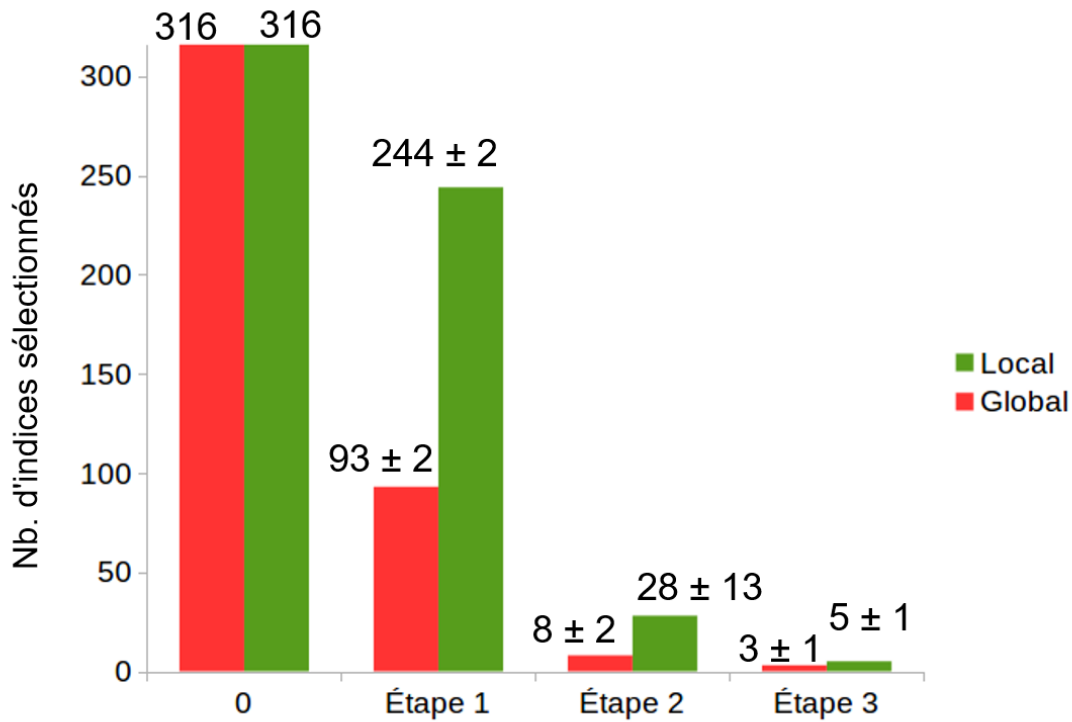


FIGURE 5.22 – Nombre moyen d'indices sélectionnés à chaque étape de la procédure de sélection d'indice pour les 33 patients de la cohorte 4Mod avec l'information sur la mutation TP53.

Classification (validation)

Comme les indices cliniques ne satisfont pas le critère de performance individuelle, nous n'avons testé ni les modèles cliniques, ni les modèles combinés. Ainsi, le tableau 5.6 montre uniquement les résultats de validation en LOOCV des modèles d'imagerie.

TABLEAU 5.6 – Résultats de validation des modèles d'imagerie (global et local).

		RF	KNN	SVM	Logit
Global					
Imagerie	Se	0,63 ± 0,04	0,63 ± 0,04	0,63 ± 0,03	0,63 ± 0,02
	Sp	0,46 ± 0,06	0,59 ± 0,08	0,68 ± 0,03	0,68 ± 0,03
	Youden	0,09 ± 0,08	0,23 ± 0,10	0,30 ± 0,05	0,31 ± 0,03
	F1-pond	0,56 ± 0,04	0,62 ± 0,05	0,65 ± 0,02	0,65 ± 0,02
Local					
Imagerie	Se	0,56 ± 0,03	0,51 ± 0,06	0,46 ± 0,03	0,56 ± 0,04
	Sp	0,48 ± 0,08	0,63 ± 0,08	0,68 ± 0,03	0,69 ± 0,00
	Youden	0,04 ± 0,08	0,14 ± 0,10	0,14 ± 0,04	0,25 ± 0,04
	F1-pond	0,53 ± 0,03	0,56 ± 0,05	0,55 ± 0,02	0,61 ± 0,02

La figure 5.23 montre les scores F1-pondérés de LOOCV pour chaque modèle de classification TP53 vs. TP53-WT avec les indices d'imagerie, locaux ou globaux, sélectionnés par les trois étapes de sélection d'indices. Chaque point représente une des 30 répétitions de la LOOCV. La moyenne et l'écart-type sont aussi affichés. La valeur de F1-pondéré si tous les patients de la cohorte radiomique 4Mod étaient classés comme TP53 (la classe la plus fréquente) serait 0,52. Cette valeur est tracée sur la figure pour faciliter l'analyse de qualité des modèles.

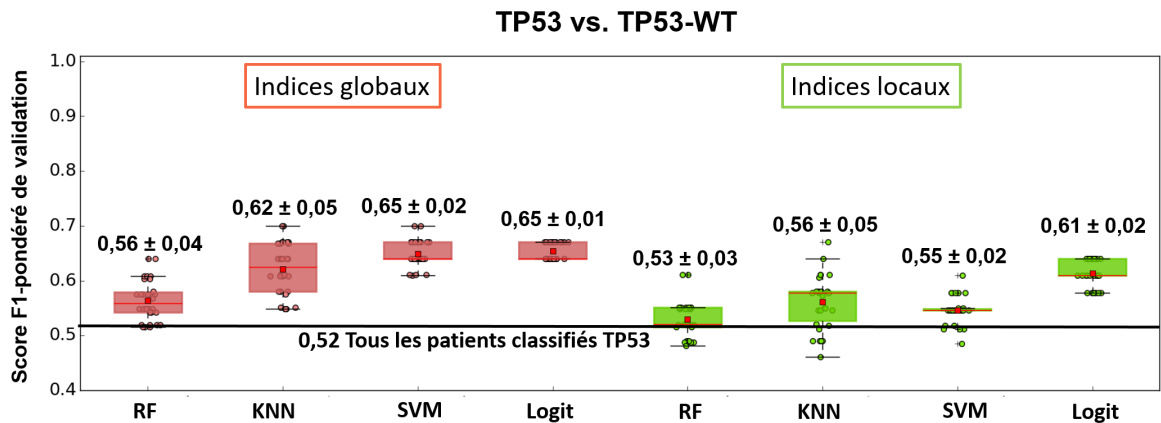


FIGURE 5.23 – Les boîtes à moustaches des valeurs des F1-pondérés des 30 répétitions de la LOOCV. La moyenne et l'écart-type des scores F1-pondérés de validation pour tous les modèles de classification de TP53 et TP53-WT estimés avec les indices d'imagerie sélectionnés. Le score d'un modèle trivial, classant tous les patients comme présentant une mutation du gène TP53 est affiché.

On peut observer que la plupart des modèles ne sont pas très performants, principalement les modèles utilisant les indices locaux. La régression logistique semble avoir une performance acceptable pour les deux types d'indices. Les modèles SVM et KNN ont des performances acceptables pour les indices globaux.

Pour la régression logistique, un modèle clinique avec l'indice clinique présentant l'AUC la plus élevée (même si inférieure à 0,75), et un modèle combiné avec ce même indice clinique et les indices d'imagerie sélectionnés ont été testés et les scores F1-pondérés résultants sont présentés sur la figure 5.24. Le modèle clinique a une performance très faible et la performance du modèle combiné est dégradée par les indices cliniques quand comparé aux modèles d'imagerie seuls.

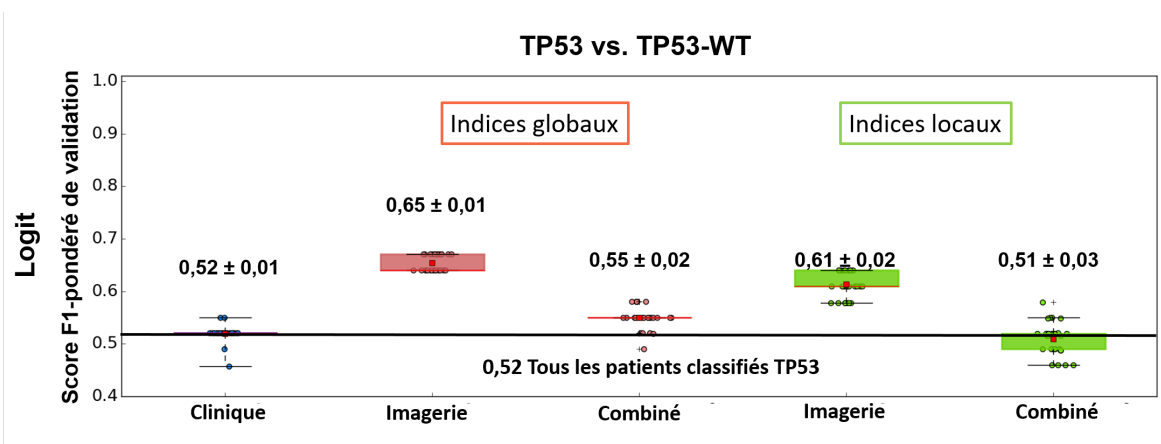


FIGURE 5.24 – Les boîtes à moustaches des valeurs des F1-pondérés des 30 répétitions de la LOOCV. La moyenne et l'écart-type de score F1-pondéré de validation pour les modèles de régression logistique (logit) de TP53 vs. TP53-WT estimés avec les indices d'imagerie sélectionnés seuls, les indices cliniques et la combinaison des indices cliniques et d'imagerie. Le score d'un modèle trivial, classant tous les patients comme TP53 est affiché.

Classification (test)

Le modèle est construit en utilisant les indices locaux sélectionnés *FLAIR glszm SizeZoneVariability*, *T2w glrlm ShortRunLowGrayLevelEmphasis*, *T1c glszm HighIntensityLargeAreaEmphasis*, *FLAIR firstorder 90Percentile*, *FLAIR firstorder Median* et *T1c firstorder 10Percentile*. Les indices sélectionnés dans le modèle d'imagerie avec les indices globaux sont *FLAIR firstorder Median*, *T1c firstorder 10Percentile* et *T1c firstorder Median*. Les indices fréquemment sélectionnés dans les plis de LOOCV sont inclus. Tous les patients qui avaient l'information sur la mutation TP53 et au moins T1c et FLAIR en image de diagnostic ont été inclus dans la cohorte de test du modèle global (10 TP53 et 9 TP53-WT). Ceux qui présentaient au moins T2w, T1c et FLAIR ont été inclus dans la cohorte de test du modèle local (9 TP53 et 8 TP53-WT).

Le tableau 5.7 présente les résultats de test des modèles d'indices d'imagerie locaux et globaux comparés aux résultats de validation en LOOCV. Comme la proportion des patients (TP53 et TP53-WT) est équilibrée, le score F1-pondéré banal vaut environ 0,5. Les modèles d'indices d'imagerie globaux sont légèrement meilleurs que les modèles d'indices d'imagerie locaux mais aucun modèle n'est performant en test (score F1-pondéré de l'ordre de 0,55 et indice de Youden de l'ordre de 0,08).

TABLEAU 5.7 – Résultats de test des modèles TP53 d'imagerie (global et local).

		RF	KNN	SVM	Logit
Global					
Validation (LOOCV)	Se	0,63 ± 0,04	0,63 ± 0,04	0,63 ± 0,03	0,63 ± 0,02
	Sp	0,46 ± 0,06	0,59 ± 0,08	0,68 ± 0,03	0,68 ± 0,03
	Youden	0,09 ± 0,08	0,23 ± 0,10	0,30 ± 0,05	0,31 ± 0,03
	F1-pond	0,56 ± 0,04	0,62 ± 0,05	0,65 ± 0,02	0,65 ± 0,02
Test BIOMEDE	Se	0,74 ± 0,08	0,76 ± 0,07	0,70 ± 0,00	0,70 ± 0,00
	Sp	0,38 ± 0,06	0,33 ± 0,00	0,37 ± 0,05	0,41 ± 0,07
	Youden	0,12 ± 0,09	0,09 ± 0,07	0,07 ± 0,05	0,11 ± 0,07
	F1-pond	0,55 ± 0,04	0,54 ± 0,03	0,53 ± 0,03	0,55 ± 0,04
Local					
Validation (LOOCV)	Se	0,56 ± 0,03	0,51 ± 0,06	0,46 ± 0,03	0,56 ± 0,04
	Sp	0,48 ± 0,08	0,63 ± 0,08	0,68 ± 0,03	0,69 ± 0,00
	Youden	0,04 ± 0,08	0,14 ± 0,10	0,14 ± 0,04	0,25 ± 0,04
	F1-pond	0,53 ± 0,03	0,56 ± 0,05	0,55 ± 0,02	0,61 ± 0,02
Test BIOMEDE	Se	0,50 ± 0,16	0,62 ± 0,09	0,55 ± 0,06	0,59 ± 0,05
	Sp	0,43 ± 0,18	0,41 ± 0,10	0,52 ± 0,09	0,49 ± 0,03
	Youden	0,07 ± 0,16	0,03 ± 0,10	0,07 ± 0,06	0,08 ± 0,05
	F1-pond	0,46 ± 0,08	0,52 ± 0,05	0,54 ± 0,03	0,54 ± 0,03

5.2.5 Application des modèles de classification pour la mutation ACVR1

Sélection d'indice

Avec les 33 patients de la [cohorte radiomique 4Mod](#) pour lesquels l'information sur la mutation en ACVR1 était disponible, en moyenne, 3±1 indices d'imagerie locaux et 2±1 indices globaux sont sélectionnés après les trois étapes de la procédure de sélection. Deux indices locaux (*T1w firstorder Maximum*, *T1w firstorder 90Percentile*) et un indice global (*T1w firstorder Median*) sont sélectionnés dans au moins 27 plis sur 33. En plus, parmi les trois indices cliniques, l'âge au diagnostic passe le seuil de 0,75 de AUC une seule fois dans les 33 couches, le volume et le sexe jamais.

La figure 5.25 montre le nombre moyen d'indices d'imagerie sélectionnés par pli de la LOOCV dans chaque étape de la procédure de sélection d'indice (et l'écart type correspondant) pour les

deux types d'indices, locaux et globaux. Une quantité importante d'indices globaux est éliminée par l'étape 1, comme déjà souligné par les autres études.

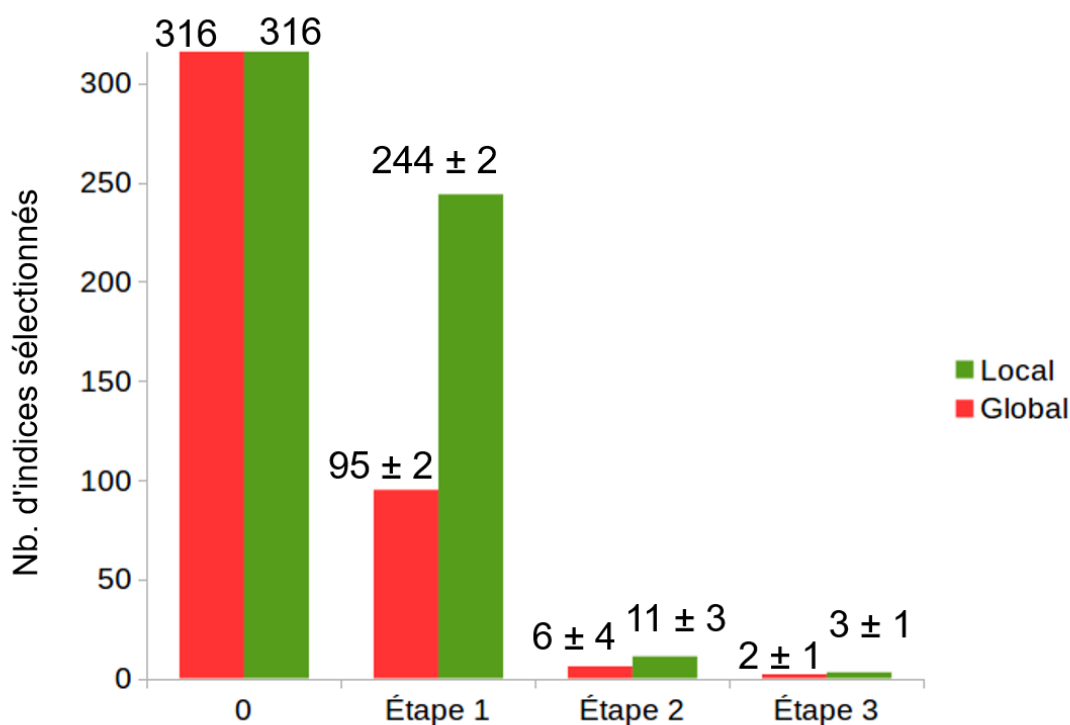


FIGURE 5.25 – Nombre moyen d'indices sélectionnés à chaque étape de la procédure de sélection d'indices pour les 33 patients de la cohorte 4Mod avec l'information sur la mutation du gène ACVR1.

Classification (validation)

Comme pour la mutation du gène TP53, les indices cliniques ne satisfont pas le critère de performance individuelle. Ainsi, le tableau 5.8 résume les résultats de LOOCV uniquement pour les modèles d'imagerie.

TABEAU 5.8 – Résultats de validation des modèles d'imagerie (global et local) pour la mutation du gène ACVR1.

		RF	KNN	SVM	Logit
		Global			
Imagerie	Se	0,43 ± 0,11	0,44 ± 0,10	0,51 ± 0,07	0,53 ± 0,07
	Sp	0,64 ± 0,05	0,64 ± 0,04	0,66 ± 0,06	0,64 ± 0,05
	Youden	0,07 ± 0,12	0,08 ± 0,11	0,18 ± 0,09	0,17 ± 0,08
	F1-pond	0,63 ± 0,04	0,63 ± 0,04	0,66 ± 0,04	0,65 ± 0,03
		Local			
Imagerie	Se	0,20 ± 0,08	0,44 ± 0,08	0,31 ± 0,05	0,38 ± 0,09
	Sp	0,66 ± 0,02	0,53 ± 0,04	0,62 ± 0,05	0,64 ± 0,03
	Youden	-0,14 ± 0,08	-0,03 ± 0,08	-0,07 ± 0,07	0,01 ± 0,09
	F1-pond	0,59 ± 0,02	0,56 ± 0,03	0,59 ± 0,03	0,61 ± 0,03

La figure 5.26 montre les scores F1-pondérés de LOOCV pour chaque modèle de classification ACVR1 vs ACVR1-WT avec les indices d'imagerie, locaux ou globaux, sélectionnés par les trois étapes de sélection d'indices. Chaque point représente une des 30 répétitions de la LOOCV. La

moyenne et l'écart-type sont aussi affichés. La valeur de F1-pondéré si tous les patients de la **cohorte radiomique 4Mod** étaient classifiés comme ACVR1-WT (la classe la plus fréquente) serait 0,69. Cette valeur est tracé sur la figure pour faciliter l'analyse de qualité des modèles.

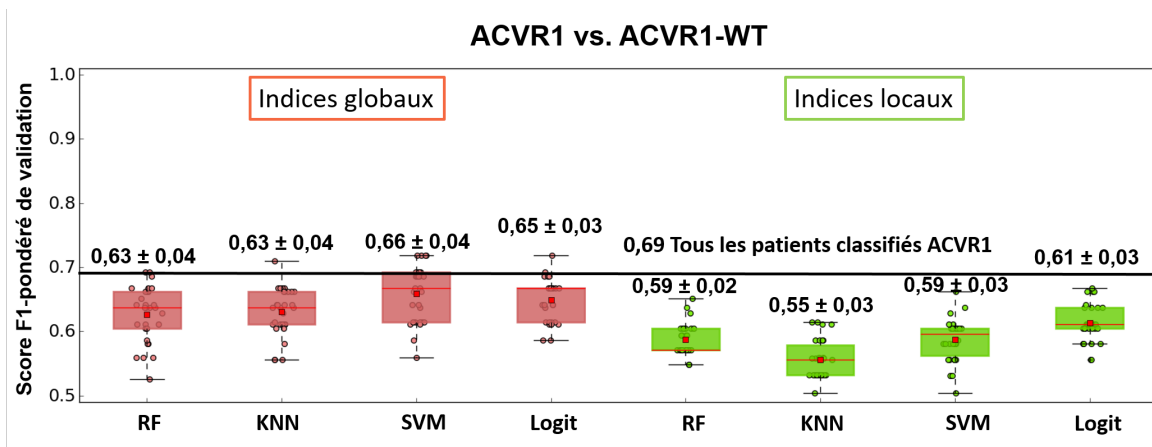


FIGURE 5.26 – Les boîtes à moustaches des valeurs des scores F1-pondérés des 30 répétitions de la LOOCV. La moyenne et l'écart-type des scores F1-pondérés de validation pour tous les modèles de classification de ACVR1 et ACVR1-WT estimés avec les indices d'imagerie sélectionnés. Le score d'un modèle trivial, classant tous les patients comme ACVR1 est affiché.

La classification de la mutation du gène ACVR1 est rendue plus difficile par le nombre réduit de patients présentant une mutation du gène ACVR1, insuffisant pour caractériser le groupe. Aucun modèle ne présente une bonne performance de validation.

Si on force la création d'un modèle clinique et d'un modèle combiné avec l'indice clinique le plus performant, comme il a été fait pour la mutation TP53, on observe également une dégradation du modèle combiné en comparaison au modèle d'imagerie seul comme montré sur la figure 5.27 pour l'exemple des modèles SVM.

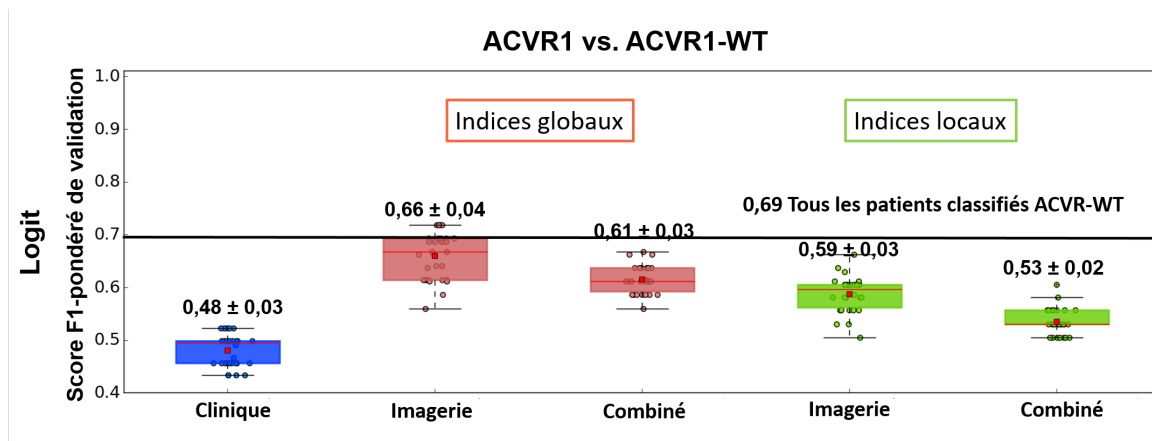


FIGURE 5.27 – Les boîtes à moustaches des valeurs des scores F1-pondérés des 30 répétitions de la LOOCV. La moyenne et l'écart-type de score F1-pondéré de validation pour les modèles de machine à vecteur support (SVM) de ACVR1 vs. ACVR1-WT estimés avec les indices d'imagerie sélectionnés seuls, les indices cliniques et la combinaison des indices cliniques et d'imagerie. Le score d'un modèle trivial, classant tous les patients comme ACVR1-WT est affiché.

Classification (test)

En utilisant les indices globaux, les indices sélectionnés dans un modèle purement d'imagerie sont *T1w firstorder Median*, *T1c firstorder 10Percentile* et *T1c firstorder Median*. Les indice sélectionnés dans le modèle d'imagerie sont *T1w firstorder 90Percentile*, *T1w firstorder Maximum* et

T1c firstorder Minimum. Donc, tous les patients qui avait l'information sur la mutation ACVR1 et au moins **T1c** et **T1w** en image au diagnostic ont été inclus dans la cohorte de test des modèles global et local (5 ACVR1 et 14 ACVR1-WT).

Les scores de test sont présentés dans le tableau 5.9. Le score F1-pondéré si tous les patients sont classifiés comme ACVR1-WT est de 0,63. Les performances en test sont équivalentes aux performances en validation en **LOOCV**. Aucun modèle n'est performant en test.

TABLEAU 5.9 – Résultats de test des modèles ACVR1 combinés (global et local).

		RF	KNN	SVM	Logit
Global					
Validation (LOOCV)	Se	0,43 ± 0,11	0,44 ± 0,10	0,51 ± 0,07	0,53 ± 0,07
	Sp	0,64 ± 0,05	0,64 ± 0,04	0,66 ± 0,06	0,64 ± 0,05
	Youden	0,07 ± 0,12	0,08 ± 0,11	0,18 ± 0,09	0,17 ± 0,08
	F1-pond	0,63 ± 0,04	0,63 ± 0,04	0,66 ± 0,04	0,65 ± 0,03
Test BIOMEDE	Se	0,03 ± 0,07	0,02 ± 0,06	0,35 ± 0,17	0,22 ± 0,07
	Sp	0,72 ± 0,06	0,75 ± 0,04	0,72 ± 0,06	0,73 ± 0,03
	Youden	-0,25 ± 0,10	-0,23 ± 0,06	0,07 ± 0,12	-0,05 ± 0,06
	F1-pond	0,52 ± 0,04	0,53 ± 0,02	0,63 ± 0,03	0,59 ± 0,02
Local					
Validation (LOOCV)	Se	0,20 ± 0,08	0,44 ± 0,08	0,31 ± 0,05	0,38 ± 0,09
	Sp	0,66 ± 0,02	0,53 ± 0,04	0,62 ± 0,05	0,64 ± 0,03
	Youden	-0,14 ± 0,08	-0,03 ± 0,08	-0,07 ± 0,07	0,01 ± 0,09
	F1-pond	0,59 ± 0,02	0,56 ± 0,03	0,59 ± 0,03	0,61 ± 0,03
Test BIOMEDE	Se	0,15 ± 0,10	0,20 ± 0,00	0,42 ± 0,13	0,48 ± 0,10
	Sp	0,74 ± 0,03	0,75 ± 0,04	0,66 ± 0,06	0,66 ± 0,03
	Youden	-0,11 ± 0,12	-0,05 ± 0,04	0,08 ± 0,10	0,14 ± 0,09
	F1-pond	0,58 ± 0,05	0,60 ± 0,02	0,61 ± 0,03	0,63 ± 0,03

5.3 La méthode multi-modèles proposée

Dans les études des tumeurs rares comme le **GITC** le nombre de patients est réduit et souvent dans la mise en œuvre des protocoles cliniques, certaines modalités d'imagerie ne sont pas acquises ou sont de qualité insuffisante et ne peuvent pas être prises en compte de façon fiable. Pour cette thèse, le nombre de patients disponibles est faible mais la base de données d'imagerie augmente progressivement avec l'intégration des données du **PHRC BIOMEDE**. Ainsi, nous avons voulu créer une méthodologie capable d'estimer des modèles prédictifs des mutations qui puissent être graduellement améliorés avec l'intégration de chaque nouveau patient. De plus, j'ai voulu inclure dans la construction des modèles tous les patients diagnostiqués avec **GITC**, car ils sont très rares. Il s'agit d'éviter, par exemple, de ne pas insérer dans la construction du modèle un nouveau patient parce qu'il n'a été imagé que par la modalité **T1w**. Cette approche reviendrait à gaspiller de l'information. De manière équivalente, il est souhaitable de pouvoir donner la meilleure estimation possible de la mutation de chaque nouveau patient pour lequel elle n'a pas été définie, indépendamment des modalités d'imagerie disponibles. C'est la raison pour laquelle j'ai proposé une méthodologie multi-modèles.

5.3.1 Définition de la méthode

Pour établir la meilleure prédiction possible, je propose la constitution de 16 modèles (M01, ..., M16) correspondant à toutes les combinaisons possibles d'information disponible (cf. tableau 5.10). Chaque modèle est construit avec tous les patients qui disposent des informations nécessaires pour le modèle en question. Chaque modèle utilise la méthodologie présentée dans la pre-

mière étude de ce chapitre, c'est-à-dire l'extraction de 79 indices par modalité d'imagerie (sur les images standardisées) et de trois indices cliniques suivie des trois étapes de sélection d'indices, de l'estimation d'un modèle d'apprentissage automatique et de l'analyse de qualité par l'indice de Youden et le score F1-pondéré mesurés par LOOCV.

Tout patient ayant l'information relative à la mutation associée à sa tumeur peut être inclus dans la création des modèles pour augmenter la puissance statistique de prédiction de ces modèles. Chaque nouveau patient dont la mutation associée à sa tumeur est inconnue peut bénéficier des modèles pour lesquels il a les informations nécessaires pour prédire la mutation.

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16
Clinique	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
79 indices T1w		X				X	X	X				X	X	X		X
79 indices T2w			X			X			X	X		X	X		X	X
79 indices FLAIR				X			X		X		X	X		X	X	X
79 indices T1c					X			X		X	X		X	X	X	X

TABLEAU 5.10 – Les informations nécessaires pour la conception des 16 modèles (M01,...,M16) sont indiquées par un 'X'.

Le choix final de prédiction de la mutation du patient est, à présent, la prédiction du "meilleur" modèle que le nouveau patient peut utiliser. Comme les modèles ont des performances différentes, la prédiction de la mutation est accompagnée du score de LOOCV du modèle utilisé pour informer sur la confiance espérée dans la prédiction, ainsi que les indices pris en compte pour faire cette prédiction, et ce de façon à atténuer l'effet "boîte noire" des analyses radiomiques.

Dans la suite de ce chapitre, cette méthode est illustrée pour la prédiction de la mutation H3.1 versus H3.3, en utilisant les indices d'imagerie globaux et la régression logistique.

5.3.2 Cas d'usage de la méthode multi-modèles sur les patients GITC

L'application de la méthodologie multi-modèles pour la prédiction de la mutation H3.1 versus H3.3 a été faite en cinq étapes :

1. Estimation de la qualité des 16 modèles par les scores de LOOCV avec l'ensemble des patients de la cohorte PREBIOMEDE disponibles (entre 30 et 38 par modèle) ;
2. Création de modèles finaux en utilisant tous les patients PREBIOMEDE pour la prédiction des mutations ;
3. Analyse de la performance de test avec les patients BIOMEDE ;
4. Estimation de qualité des modèles "mis-à-jour" par les scores de LOOCV regroupant les patients PREBIOMEDE et BIOMEDE ;
5. Création de modèles finaux "mis-à-jour" en utilisant tous les patients PREBIOMEDE et BIOMEDE pour la prédiction des mutations.

Les étapes 1 à 3 sont schématisées sur la figure 5.28. Les étapes 4 et 5 sont l'équivalent des étapes 1 et 2 en fusionnant les cohortes PREBIOMEDE et BIOMEDE en données d'entrée. Parmi les patients PREBIOMEDE, 38 présentent une mutation du type H3.1 ou H3.3 et parmi les patients BIOMEDE disponibles, 25 présentent une mutation H3.1 ou H3.3.

Estimation de la qualité des modèles avec la base de données PREBIOMEDE

Dans un premier temps, conformément au recueil des données disponibles pour la première partie de la thèse, seuls les patients PREBIOMEDE sont disponibles. Le tableau 5.11 montre le nombre de patients PREBIOMEDE utilisé pour la construction et l'analyse de chaque modèle ainsi que la distribution de ces patients selon la mutation de l'histone H3 (H3.1 versus H3.3).

Les performances moyennes et écart-types en LOOCV de chaque modèle selon la sensibilité, la spécificité, l'indice de Youden et le score F1-pondéré sont présentées sur la figure 5.29. Le nombre moyen et l'écart-type (dans les couches de LOOCV) d'indice sélectionnés pour chaque modèle

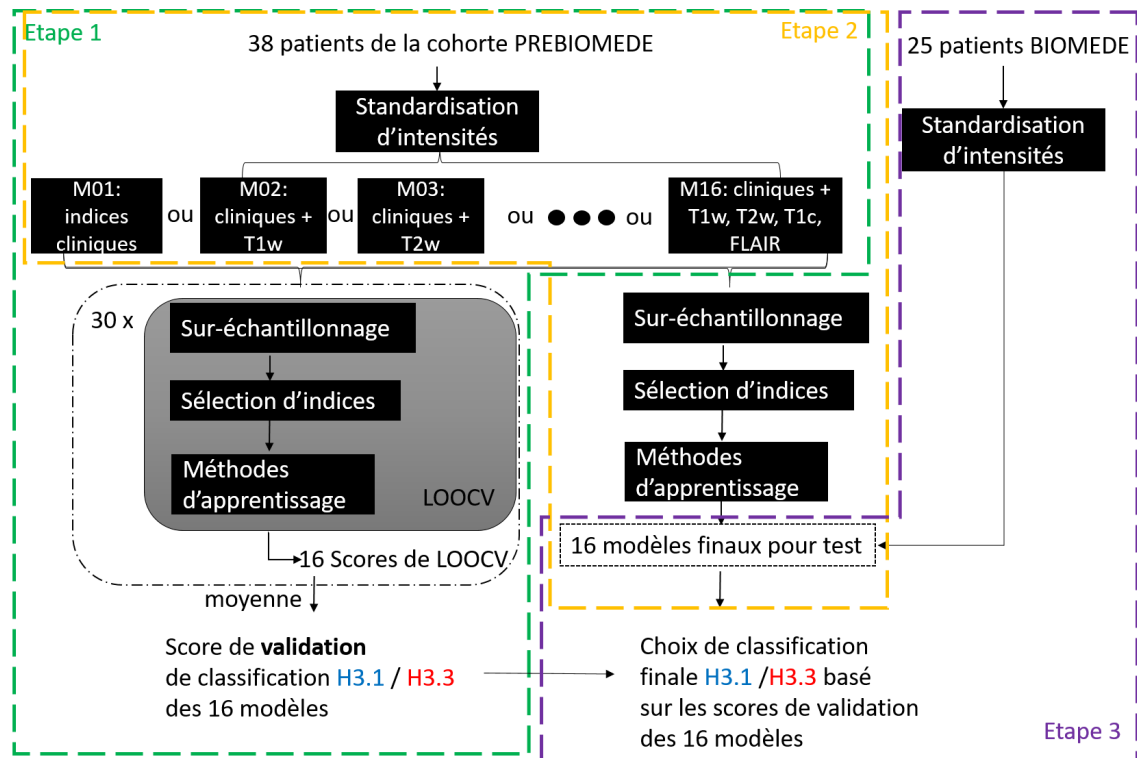


FIGURE 5.28 – Schéma des étapes 1 à 3 de l'analyse multi-modèles pour la mutation H3.

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16
Nb. Patients	38	33	32	31	32	31	32	31	31	31	30	31	31	30	30	30
H3.1	12	9	9	8	9	8	9	8	8	9	8	8	9	8	8	8
H3.3	26	24	23	23	23	23	23	23	23	22	22	23	22	22	22	22

TABLEAU 5.11 – Nombre de patients PREBIOMEDE présentant les informations nécessaires pour la création de chacun des 16 modèles.

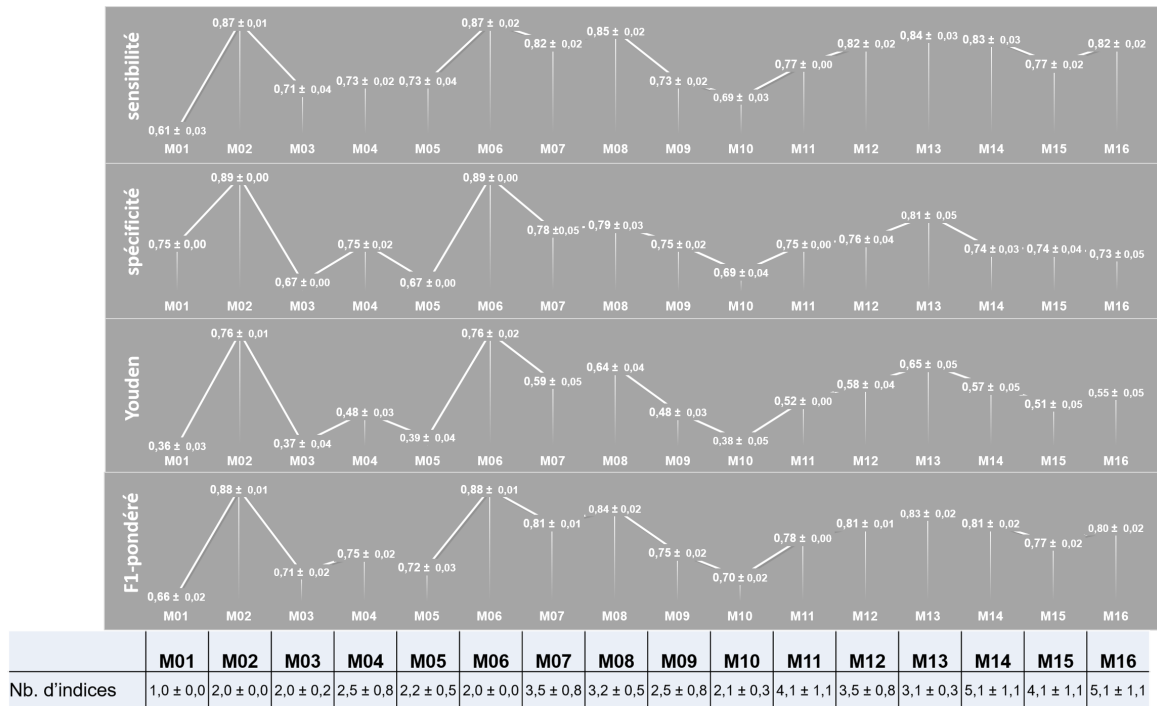


FIGURE 5.29 – Performances (moyennes ± écart-type) des 30 répétitions en LOOCV de chaque modèle (M01 à M16) selon la spécificité, la sensibilité, l'indice de Youden et le score F1-pondéré. Le nombre moyen ± l'écart-type d'indices sélectionnés dans chaque modèle est affiché dans le tableau du bas.

est également indiqué. Les rangs des modèles selon l'indice de Youden et le score F1-pondéré sont très similaires. Le modèle M02 ainsi que le modèle M06, présentent les meilleurs indices de Youden et de scores F1-pondérés, de sensibilité et de spécificité.

Ainsi, avec la cohorte PREBIOMEDE, le modèle qui utilise la modalité **IRM T1w** et les indices cliniques présente la meilleure performance en validation.

Modèles conçus avec les informations disponibles pour les patients PREBIOMEDE

Une fois que l'estimation de la performance de chaque modèle est connue, les 16 modèles finaux pour tester les nouveaux patients peuvent être définis. Seulement cinq indices d'imagerie et l'âge au diagnostic ont été sélectionnés pour les 16 modèles, ils sont présentés dans le tableau 5.12 avec un code d'identification. Le tableau 5.13 présente les indices sélectionnés pour chaque modèle.

Id1	T1w firstorder Median
Id2	T2w glcm Entropy
Id3	FLAIR glcm Homogeneity1
Id4	FLAIR firstorder 90Percentile
Id5	T1c firstorder 10Percentile

TABLEAU 5.12 – Les cinq indices d'imagerie qui ont été sélectionnés dans au moins un des 16 modèles créés avec les 38 patients PREBIOMEDE.

Les modèles qui utilisent les mêmes indices ne sont pas identiques car un nombre différent de patients a été utilisé pour estimer chaque modèle (cf. tableau 5.11) et il existe de petites variations liées à la part aléatoire introduite par le ré-échantillonnage **SMOTE** et la régression logistique.

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16
Clinique	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age
T1w	Id1					Id1	Id1	Id1				Id1	Id1	Id1		Id1
T2w			Id2													
FLAIR				Id3, Id4			Id3, Id4		Id3, Id4		Id3	Id3, Id4		Id3	Id3	Id3
T1c					Id5			Id5		Id5	Id5		Id5	Id5	Id5	Id5

TABLEAU 5.13 – Les indices sélectionnés dans chacun de 16 modèles définis à partir des patients PREBIO-MEDE. Les codes d’identification des indices sont explicités dans le tableau 5.12.

Performance des modèles sur les patients BIOMEDE

Les 25 patients BIOMEDE ayant une mutation du type H3.1 ou H3.3 ont été utilisés comme nouveaux patients pour tester les modèles. Le tableau 5.14 présente le nombre de patients BIOMEDE utilisés pour tester chaque modèle ainsi que la distribution des deux mutations de l’histone H3. Seuls les indices sélectionnés sont pris en compte pour que les patients à tester soient retenus. Ainsi, pour le modèle M06, par exemple, qui utilise les images T1w et T2w mais qui sélectionne seulement un indice T1w, les 21 patients ayant au moins la modalité T1w sont testés (même si seuls 19 patients avaient les deux modalités).

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16
Nb. Patients	25	21	22	20	21	21	19	18	20	21	18	19	18	17	18	17
H3.1	5	4	4	4	5	4	4	4	4	5	4	4	4	4	4	4
H3.3	20	17	18	16	16	17	15	14	16	16	14	15	14	13	14	13

TABLEAU 5.14 – Nombre de patients BIOMEDE ayant les informations nécessaires pour tester chacun des 16 modèles finaux.

La création des modèles finaux et le test ont été répétés 30 fois pour tenir compte du côté aléatoire du ré-échantillonnage et de la régression logistique. Les valeurs moyennes de la spécificité, la sensibilité, l’indice de Youden et le score F1-pondéré pour chaque modèle sont présentées sur la figure 5.30. On peut noter que les scores ne varient pas beaucoup dans les différentes répétitions (faible écart-type). Un exemple de sortie de prédiction faite pour chaque patient BIOMEDE par chaque modèle est montré sur le tableau 5.15.

Patient	Vrai H3	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16
sub-006	H3.3	H3.1	H3.3	H3.1	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-009	H3.3	H3.1	-	-	H3.3	H3.1	-	-	-	-	-	H3.1	-	-	-	-	-
sub-016	H3.1	H3.1	-	-	-	H3.1	-	-	-	-	-	-	-	-	-	-	-
sub-028	H3.1	H3.1	H3.3	H3.1	H3.1	H3.1	H3.3	H3.3	H3.3	H3.1	H3.1	H3.1	H3.3	H3.3	H3.3	H3.1	H3.3
sub-038	H3.3	H3.3	H3.3	H3.1	H3.1	-	H3.3	H3.3	-	H3.1	-	-	H3.3	-	-	-	-
sub-040	H3.3	H3.3	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1
sub-047	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-051	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-058	H3.3	H3.3	-	H3.3	-	H3.3	-	-	-	-	H3.3	-	-	-	-	-	-
sub-071	H3.3	H3.3	H3.3	-	H3.3	H3.1	-	H3.3	H3.3	-	-	H3.3	-	-	H3.3	-	-
sub-080	H3.1	H3.3	H3.1	H3.3	H3.3	H3.3	H3.1	H3.3	H3.1	H3.3	H3.3	H3.3	H3.3	H3.1	H3.3	H3.3	H3.3
sub-082	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-084	H3.3	H3.3	H3.3	H3.3	-	H3.3	H3.3	-	H3.3	-	H3.3	-	-	H3.3	-	-	-
sub-086	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-092	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-095	H3.1	H3.1	H3.1	H3.3	H3.3	H3.3	H3.1	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-100	H3.3	H3.3	H3.3	H3.3	-	-	H3.3	-	-	-	-	-	-	-	-	-	-
sub-102	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-104	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-110	H3.3	H3.3	H3.1	H3.3	H3.3	-	H3.1	H3.3	-	H3.3	-	-	H3.3	-	-	-	-
sub-114	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-115	H3.1	H3.1	H3.1	H3.1	H3.3	H3.3	H3.1	H3.3	H3.3	H3.3	H3.1	H3.3	H3.3	H3.1	H3.3	H3.3	H3.3
sub-124	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-135	H3.1	H3.1	H3.3	H3.1	H3.3	H3.1	H3.3	H3.3	H3.3	H3.3	H3.1	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-149	H3.3	H3.1	H3.3	H3.1	H3.1	H3.3	H3.3	H3.3	H3.3	H3.1	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3

TABLEAU 5.15 – Prédiction de la mutation H3 faite pour chaque patient BIOMEDE par chaque modèle de l’approche multi-modèles possible. Les modèles ont été appris avec les patients PREBIO-MEDE.

Le modèle M01 qui se base sur l’âge au diagnostic est très performant pour les patients BIO-

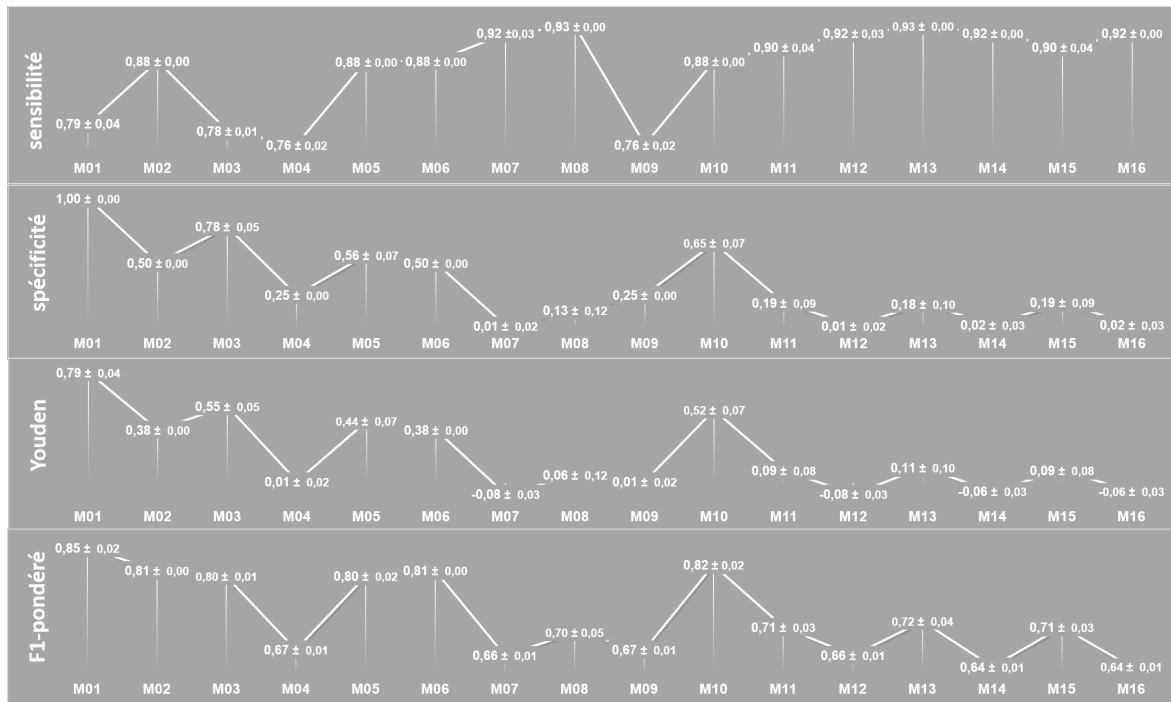


FIGURE 5.30 – Performances moyennes \pm écart-type de 30 répétitions de test avec les patients BIOMEDE de chaque modèle (M01 à M16) final conçu avec les patients PREBIOMEDE, établies en fonction de la sensibilité, la spécificité, l'indice de Youden et le score F1-pondéré.

MEDE (sensibilité = $0,79 \pm 0,04$, spécificité = $1,00 \pm 0,00$, Youden = $0,79 \pm 0,04$, F1-pondéré = $0,85 \pm 0,02$) les modèles M02 et M06, qui sont les deux meilleurs en LOOCV pour les patients PRE-BIOMEDE, sont aussi performants en test sur les patients BIOMEDE (sensibilité = $0,88 \pm 0,00$, spécificité = $0,50 \pm 0,00$, Youden = $0,38 \pm 0,00$, F1-pondéré = $0,81 \pm 0,00$).

Performance de la prédiction finale : Pour chaque nouveau patient, même si sa mutation peut être prédite avec plusieurs modèles, il faut opter pour une réponse finale. Nous nous sommes fondée sur les scores F1-pondérés de LOOCV pour classer les modèles M01 à M16. Les résultats présentés sur la figure 5.30 ne sont pas utilisés car on simule une situation de prédiction avec la vraie classification inconnue. Ainsi, pour chaque nouveau patient, la réponse finale est la réponse du modèle le plus performant en validation qu'il peut utiliser sur la base du score F1-pondéré. S'il y a plus d'un modèle retenu, le modèle qui a été conçu avec le plus grand nombre de patients est choisi.

Pour simuler une situation d'arrivée de nouveaux patients, les réponses d'une des 30 répétitions (qui ne varient presque pas) sont utilisées pour créer le tableau 5.16 qui montre les prédictions finales pour les patients BIOMEDE, la mutation réelle et les indices qui ont été utilisés pour faire la prédiction. De plus les caractéristiques du modèle utilisé pour la prédiction sont indiquées : le rang du modèle parmi les 16, le score de F1-pondéré et le nombre de patients utilisés pour concevoir le modèle. La performance de la prédiction finale est résumée dans le tableau 5.17 : **sensibilité = 0,85, spécificité = 0,60, Youden = 0,45, F1-pondéré = 0,81**. Ces résultats sont meilleurs que les résultats de test obtenus avec une approche traditionnelle de la régression logistique du modèle combiné d'indices cliniques et d'indices d'imagerie globaux montrés précédemment dans le tableau 5.5 (sensibilité = $0,92 \pm 0,01$, spécificité = $0,18 \pm 0,17$, Youden = $0,10 \pm 0,17$, F1-pondéré = $0,70 \pm 0,07$).

La distribution des patients BIOMEDE dans l'espace associé aux quatre modèles qui ont été utilisés pour la prédiction finale (M02, M03, M05 et M11) est présentée sur la figure 5.31. La couleur de fond est la prédiction du modèle Logit dans l'espace des indices (H3.1 en rouge, H3.3 en bleu). Chaque cercle représente un patient et il est coloré selon la mutation associée à ce patient. Les

Patient	Préd H3	Vrai H3	Modèle, Rang : F1-pond	Indices utilisés	H3.1 / H3.3
sub-006	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-009	H3.1	H3.3	M11, Rang 9 : 0,78	Age, Id3, Id5	8 / 22
sub-016	H3.1	H3.1	M05, Rang 13 : 0,72	Age, Id5	9 / 23
sub-028	H3.3	H3.1	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-038	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-040	H3.1	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-047	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-051	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-058	H3.3	H3.3	M05, Rang 13 : 0,72	Age, Id5	9 / 23
sub-071	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-082	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-084	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-086	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-092	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-095	H3.1	H3.1	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-100	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-102	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-104	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-110	H3.1	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-114	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-115	H3.1	H3.1	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-124	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-135	H3.3	H3.1	M02, Rang 1 : 0,88	Age, Id1	9 / 24
sub-143	H3.3	H3.3	M03, Rang 14 : 0,71	Age, Id2	9 / 23
sub-149	H3.3	H3.3	M02, Rang 1 : 0,88	Age, Id1	9 / 24

TABLEAU 5.16 – Les prédictions finales pour les patients BIOMEDE comparées à leur statut réel et les indices qui ont été utilisés pour faire la prédiction ainsi que les caractéristiques du modèle qui a été utilisé pour sa prédiction (son rang parmi les 16 modèles, le score de F1-pondéré et le nombre de patients utilisés pour concevoir le modèle).

Sensibilité	Spécificité	Youden	F1-pond
0,85	0,60	0,45	0,81

TABLEAU 5.17 – Performance de la prédiction finale de la mutation H3 pour les patients BIOMEDE, en utilisant l'approche multi-modèles.

H3.1 / H3.3

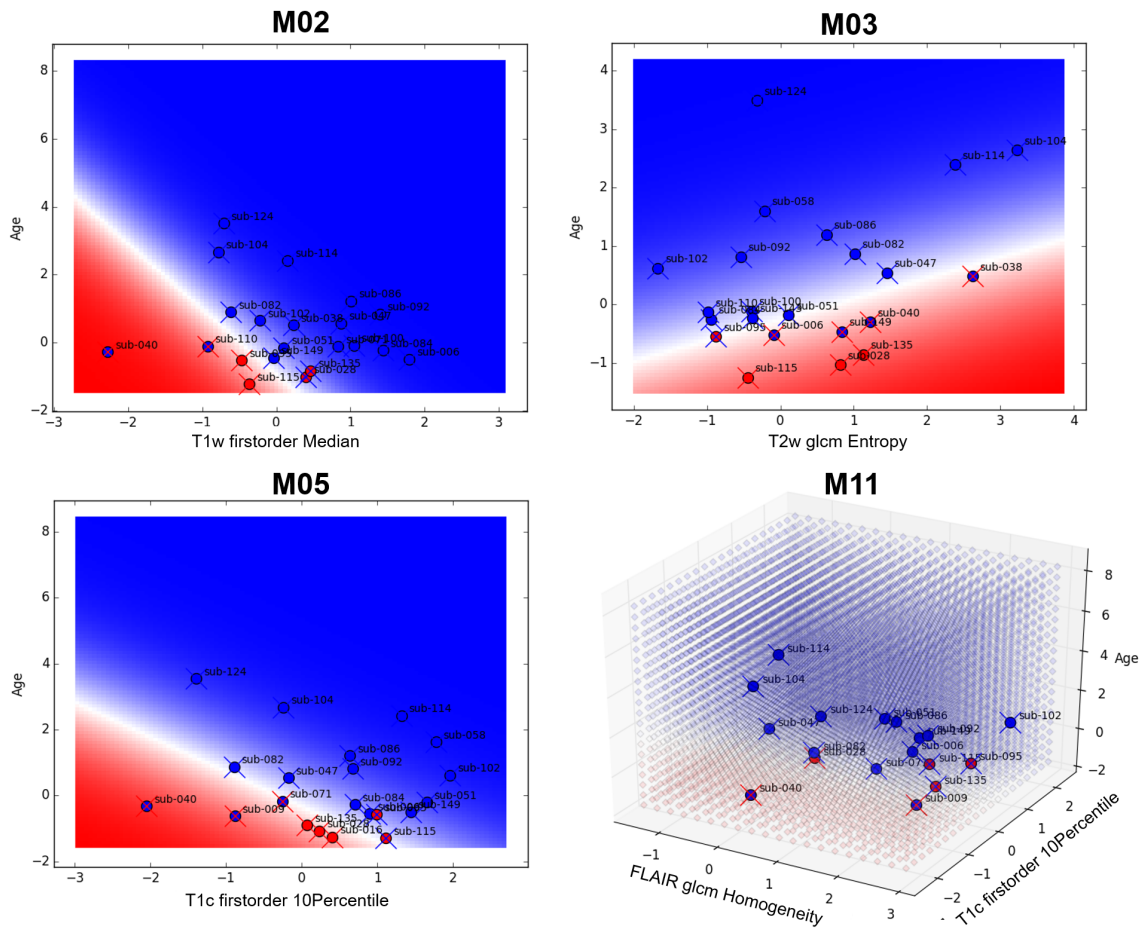


FIGURE 5.31 – Distributions des patients BIOMEDE dans l'espace des modèles M02, M03, M05 et M11. La couleur de fond est la prédiction du modèle Logit dans l'espace des indices. Chaque cercle représente un patient et il est coloré selon la mutation associée à ce patient. Les 'X' sont colorés par la prédiction du modèle.

'X' sont colorés par la prédiction du modèle. Ces graphes nous permettent d'évaluer aussi si les patients mal-classés sont éloignés ou très proches de la ligne de séparation des classes.

Conception des modèles et estimation de leur qualité en fusionnant les bases PREBIOMEDE avec BIOMEDE

Pour continuer sur notre cas d'usage, une fois que les données sur les patients PREBIOMEDE et BIOMEDE ont été disponibles, tous les patients GITC ayant une mutation du type H3.1 ou H3.3 disponibles pour cette thèse ont été rassemblés pour estimer les 16 modèles et leurs performances en validation. Les nombres de patients utilisés pour chaque modèle sont donc la somme des tableaux 5.11 et 5.14 et sont résumés dans le tableau 5.18. Cette étape doit être répétée à chaque fois que des nouveaux patients sont inclus afin d'avoir des modèles à jour.

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16
Nb. Patients	63	54	54	51	53	52	51	49	51	52	48	50	49	47	48	47
H3.1	17	13	13	12	14	12	13	12	12	14	12	12	13	12	12	12
H3.3	46	41	41	39	39	40	38	37	39	38	36	38	36	35	36	35

TABLEAU 5.18 – Nombre de patients PREBIOMEDE et BIOMEDE avec les informations nécessaires pour créer chaque modèle.

La figure 5.32 présente la performance de chaque modèle obtenue en fusionnant les bases en

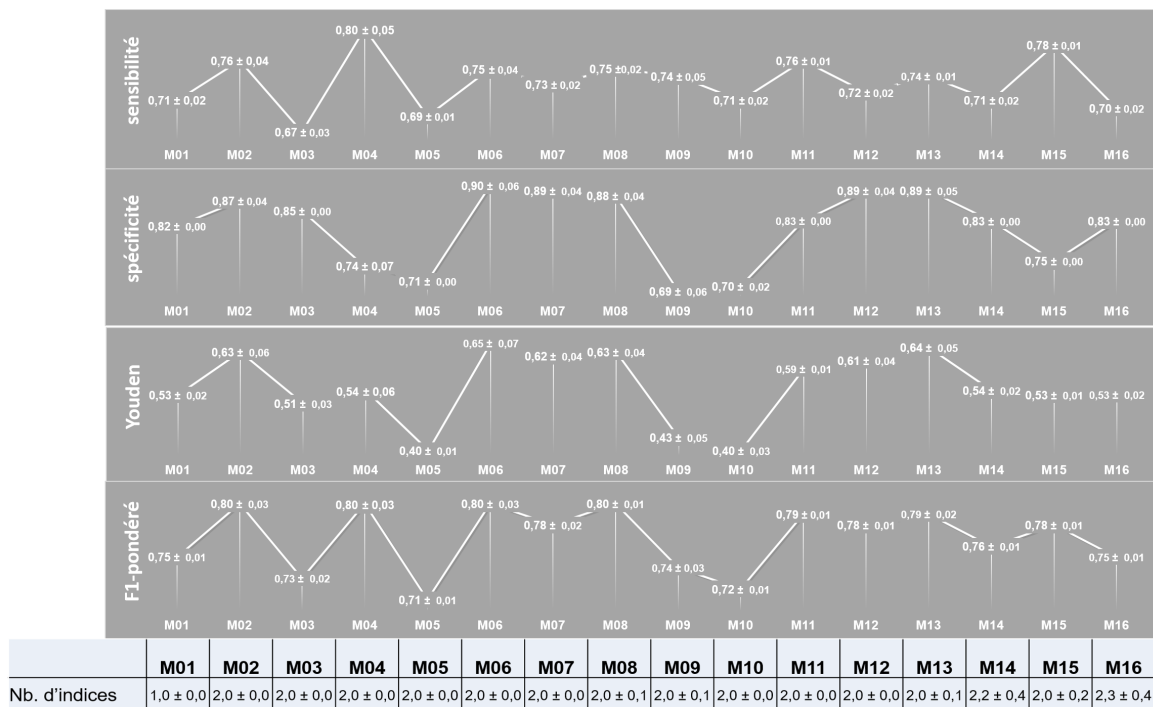


FIGURE 5.32 – Performances moyennes ± écart-type de 30 répétitions en LOOCV de chaque modèle (M01 à M16) estimé en regroupant les patients PREBIOMEDE et BIOMEDE selon la spécificité, la sensibilité, l'indice de Youden et le score F1-pondéré. Le nombre moyen ± l'écart-type d'indices sélectionnés dans chaque modèle est affiché dans le tableau du bas.

fonction de la sensibilité, la spécificité, l'indice de Youden et le score F1-pondéré.

On voit une cohérence entre les modèles performants et non performants créés avec les patients PREBIOMEDE seuls et fusionnés avec les patients BIOMEDE. Dans les deux cas, les modèles M02 et M06 ont de bonnes performances et les modèles M03, M05 et M10 sont parmi les moins performants.

Modèles conçus en fusionnant les bases PREBIOMEDE et BIOMEDE

Finalement, en utilisant tous les patients, les modèles à jour utilisent les indices indiqués dans le tableau 5.19. Plus précisément, les indices utilisés dans chaque modèle sont présentés dans le tableau 5.20.

Id1	T1w firstorder Median
Id3	FLAIR glcm Homogeneity1
Id6	T2w firstorder 90Percentile
Id7	T1c glcm DifferenceEntropy
Id8	T1c glcm Entropy
Id9	FLAIR glcm Homogeneity2

TABEAU 5.19 – Les six indices d'imagerie qui ont été sélectionnés dans au moins un des 16 modèles créés avec la fusion des bases PREBIOMEDE et BIOMEDE. Les codes d'identification des indices qui sont aussi sélectionnés dans les modèles conçus avec la base PREBIOMEDE ont été gardés.

Avec l'ajout de patients, la valeur médiane en T1w et l'homogénéité en FLAIR restent sélectionnés dans les modèles. Globalement, les modèles avec la fusion des deux bases sélectionnent moins d'indices que les modèles estimés avec la base PREBIOMEDE. La moitié des modèles utilisent les mêmes indices (Âge au diagnostic et T1 firstorder Median).

Pour finir avec ce cas d'usage, le tableau 5.21 présente les prédictions des mutations des pa-

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16
Clinique	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age	Age
T1w		Id1				Id1	Id1	Id1				Id1	Id1	Id1		Id1
T2w			Id6													
FLAIR				Id3					Id9							
T1c					Id7					Id8	Id7				Id7	

TABLEAU 5.20 – Les indices sélectionnés dans chacun de 16 modèles constitués en fusionnant les bases PREBIOMEDE et BIOMEDE. Les noms des indices sont présentés dans le tableau 5.19

tients PREBIOMEDE et BIOMEDE qui n’ont pas subi la chirurgie de biopsie ou qui présentent des résultats non concluants. Ces prédictions pourraient à terme guider le choix de traitement et devront être améliorées avec de nouveaux patients pour être plus sûres. A présent, le modèle n’est pas parfaitement adapté car parmi ces patients il peut exister des patients WT, qui ne sont pas considérés. Les prédictions faites par chaque modèle sont présentées dans le tableau 5.22.

Patient	Préd H3	Modèle, Rang : F1-pond	Sur le modèle	
			Indices utilisés	H3.1 / H3.3
sub-7494	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-5196	H3.3	M11, Rang 6 : 0,79	Age, Id7	12 / 36
sub-6450	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-4437	H3.1	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-5503	H3.1	M01, Rang 11 : 0,75	Age	17 / 46
sub-1829	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-037	H3.1	M11, Rang 6 : 0,79	Age, Id7	12 / 36
sub-073	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-080	H3.1	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-108	H3.1	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-160	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-162	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-163	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-166	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-178	H3.1	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-190	H3.1	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-218	H3.3	M02 , Rang 1 : 0,80	Age, Id1	13 / 41
sub-219	H3.1	M02 , Rang 1 : 0,80	Age, Id1	13 / 41

TABLEAU 5.21 – Les prédictions finales pour les patients PREBIOMEDE et BIOMEDE pour lesquels on ne connaît pas la vraie mutation H3 et les indices qui ont été utilisés pour faire la prédiction ainsi que les caractéristiques du modèle qui a été utilisé pour sa prédiction (son rang parmi les 16 modèles, le score de F1-pondéré et le nombre de patients utilisés pour concevoir le modèle).

La figure 5.33 montre les distributions des patients pour lesquels le statut de la mutation H3 est inconnu dans l’espace des trois modèles qui ont été utilisés pour la prédiction finale (M01 ,M02 et M11). La couleur de fond est la prédiction du modèle Logit dans l’espace des indices pour les modèles M02 et M11 (mutés en H3.1 en rouge et mutés en H3.3 en bleu). Pour le modèle M01 qui utilise seulement l’âge au diagnostic, la courbe représente la probabilité estimée par le classifieur Logit que le patient soit muté en H3.3. Les ‘X’ sont colorés par la prédiction du modèle. Ces graphes nous permettent d’évaluer aussi si le patient est éloigné ou très proche de la ligne de séparation des classes.

On peut remarquer que le patient *sub-5503* qui a utilisé le modèle M01 a été classifié comme muté en H3.1 et il a, en effet, une probabilité très faible d’être muté en H3.3. D’un autre côté, le patient *sub-5196* qui a utilisé le modèle M11 a été classifié comme muté en H3.3 mais il est sur la ligne de séparation des classes.

Dès que des patients supplémentaires seront inclus, ceux qui ont les résultats des analyses

Patient	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16
sub-7494	H3.1	H3.3	H3.1	H3.3	H3.1	H3.3	H3.3	H3.3	H3.1	H3.1	H3.1	H3.3	H3.3	H3.3	H3.1	H3.3
sub-5196	H3.1	-	H3.1	-	H3.1	-	-	-	-	H3.1	H3.3	-	-	-	H3.1	-
sub-6450	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-4437	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1
sub-5503	H3.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
sub-1829	H3.1	H3.3	H3.1	H3.1	H3.1	H3.3	H3.3	H3.3	H3.1	H3.1	H3.1	H3.3	H3.3	H3.3	H3.1	H3.3
sub-037	H3.1	-	H3.1	-	H3.1	-	-	-	-	H3.1	H3.1	-	-	-	H3.1	-
sub-073	H3.3	H3.3	H3.3	H3.3	-	H3.3	H3.3	H3.3	H3.3	-	-	H3.3	H3.3	H3.3	-	H3.3
sub-080	H3.3	H3.1	H3.3	H3.3	H3.3	H3.1	H3.1	H3.1	H3.3	H3.3	H3.3	H3.1	H3.1	H3.1	H3.3	H3.1
sub-108	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1	H3.1
sub-160	H3.3	H3.3	H3.3	-	H3.3	H3.3	H3.3	H3.3	-	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-162	H3.3	H3.3	H3.3	-	H3.3	H3.3	H3.3	H3.3	-	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-163	H3.3	H3.3	-	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-166	H3.3	H3.3	H3.3	-	H3.3	H3.3	H3.3	H3.3	-	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3
sub-178	H3.1	H3.1	H3.1	H3.1	-	H3.1	H3.1	H3.1	H3.1	-	-	H3.1	H3.1	H3.1	-	H3.1
sub-190	H3.3	H3.1	H3.3	H3.3	-	H3.1	H3.1	H3.1	H3.3	-	-	H3.1	H3.1	H3.1	-	H3.1
sub-218	H3.1	H3.3	H3.1	-	H3.1	H3.3	H3.3	H3.3	-	H3.1	H3.1	H3.3	H3.3	H3.3	H3.1	H3.3
sub-219	H3.3	H3.1	-	H3.3	H3.3	H3.3	H3.3	H3.1	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3	H3.3

TABLEAU 5.22 – Prédiction de la mutation H3 faite pour chaque patient PREBIOMEDE et BIOMEDE par chaque modèle de l’approche multi-modèles estimé avec les 63 patients des bases PREBIOMEDE et BIOMEDE mutés en H3.1 ou H3.3.

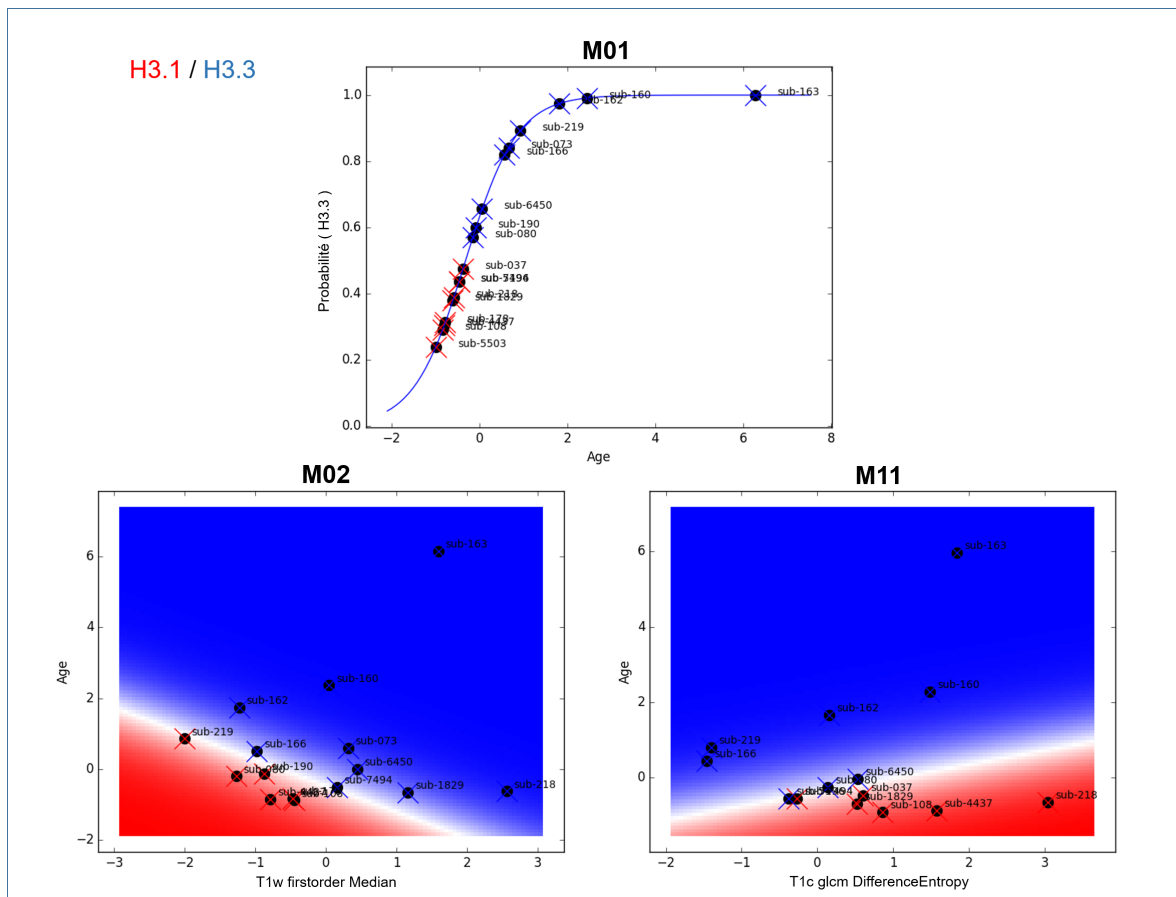


FIGURE 5.33 – Distributions des patients PREBIOMEDE et BIOMEDE pour lesquels on ne connaît pas la mutation H3 dans l’espace des modèles M01, M02 et M11. La couleur de fond est la prédiction du modèle Logit dans l’espace des indices pour les modèles M02 et M11. Pour le modèle M01, la courbe représente la probabilité estimée par le classifieur Logit que le patient soit muté en H3.3. Les ‘X’ sont colorés par la prédiction du modèle.

histologiques seront introduits dans la conception des modèles pour réaliser leur mise à jour. Les modèles à jour seront ensuite utilisés pour prédire la mutation des patients qui n'ont pas cette information.

5.4 Discussion et conclusion

Dans la première étude de ce chapitre, des analyses radiomiques ont été conduites avec les images de diagnostic et les indices cliniques pour répondre à trois tâches distinctes de classification :

1. Classification des patients présentant la mutation H3.1 ou H3.3
2. Classification des patients présentant une mutation du gène TP53 ou ne la présentant pas (TP53-WT)
3. Classification des patients présentant une mutation du gène ACVR1 ou ne la présentant pas (ACVR-WT)

La deuxième étude de ce chapitre a proposé une méthodologie multi-modèles adaptée à la prise en charge des patients GITC. Le principe est de pouvoir inclure tous les patients dans les études radiomiques, même ceux qui n'ont pas suivi le protocole d'imagerie complet. L'application de la méthodologie a été illustrée sur la prédiction de la mutation H3.1 et H3.3.

Les expérimentations de la première étude ont été réalisées avec un sous-ensemble de la cohorte PREBIOMEDE, la [cohorte radiomique 4Mod](#), pour laquelle les patients présentaient les quatre types d'imagerie structurale (T1w, T2w, T1c et FLAIR) au moment du diagnostic. Pour des questions d'inclusion progressive des données, les modèles ont été testés sur les patients de la cohorte BIOMEDE qui présentaient les modalités d'imagerie demandées par les modèles.

Avec l'analyse individuelle des indices cliniques, il a été possible de constater la variation importante des valeurs de AUC selon la cohorte analysée. En effet, avec un faible nombre de sujets, la précision statistique des résultats est faible et seulement des indices très caractéristiques du problème de classification étudié peuvent être identifiés. C'est le cas de l'âge au moment du diagnostic pour l'étude de la mutation H3.

Classification des patients présentant les mutations des histones H3.1 et H3.3 : Pour cette tâche de classification, l'analyse uni-variée a révélé que plusieurs indices sont potentiellement efficaces pour différencier les patients qui présentent une mutation H3.1 de ceux présentant une mutation H3.3 par les valeurs de AUC calculées pour chaque indice séparément. Parmi les sept indices qui présentent une AUC supérieure à l'âge au diagnostic pour la classification des mutations H3, il existe une prévalence d'indices d'imagerie de premier-ordre, donc plus simples. Il se trouve que les deux indices locaux en T1c (10Percentile et Minimum) sont corrélés avec un r de Spearman de 0,90 d'après les données de cette cohorte, les deux indices locaux en T1w (Maximum et 90Percentile) sont corrélés avec un r de Spearman de 0,92. Finalement, les trois indices globaux de T1w (Mean, Median, RootMeanSquare) sont tous corrélés deux-à-deux avec un r de Spearman supérieur à 0,99. Cela veut dire que les sept indices traduisent trois types d'informations différentes et non sept.

Les résultats des modèles multi-variés sont cohérents avec les modèles uni-variés, car parmi les indices sélectionnés il y a toujours un indice représentant chacun des trois groupes d'indices informatifs ainsi que des indices de la modalité FLAIR qui présentent également des AUC élevées. Les modèles utilisant les indices globaux et les indices locaux reposent respectivement sur environ quatre et sept indices d'imagerie et les résultats de validation en LOOCV suggèrent que les modèles combinant ces indices d'imagerie avec l'âge au diagnostic sont les plus performants. Aucune différence significative réelle n'a été trouvée entre les résultats de différentes méthodes d'apprentissage automatique mises-en-œuvre.

Il n'existe pas forcément de correspondance entre la performance en LOOCV et celle en test. Par exemple, en comparant les modèles combinés d'indices simples d'imagerie globaux et cli-

niques, la méthode de forêt aléatoire est la moins performante en validation mais la plus performante en test pour l'indice de Youden et pour le score F1-pondéré. Néanmoins, les résultats de test ne sont pas aussi bons que les résultats de validation surtout en raison de la faible spécificité (capacité de détecter les patients mutés en H3.1). Nous avons exploré quatre pistes de solution pour comprendre ces différences de résultats :

1. Standardisation des intensités : une possible cause serait l'insuffisance de la standardisation d'intensités proposée dans le chapitre 3 pour corriger les variations indésirables des images des cohortes PREBIOMEDE et BIOMEDE. En effet ces cohortes proviennent d'années différentes (entre 2009 et 2014 versus entre 2014 et 2019) où la qualité des images a pu changer. Cependant, les résultats de la normalisation des images BIOMEDE ont été validés en ayant comme modèle des images PREBIOMEDE et toutes les images faites sur l'examen avant biopsie ont passé les critères de validation.
2. Sélection d'indices : la deuxième hypothèse serait la pertinence de l'étape de sélection d'indices. Cependant, les indices sélectionnés sont cohérents avec les analyses qualitatives faites par les radiologues, ils sont aussi cohérents entre les expérimentations (analyse individuelle, les couches de la LOOCV et le modèle final). En plus, j'ai testé une autre méthode de sélection, le mRMR (PENG et collab. [2005]) cité dans le chapitre 1. Le tableau 5.23 montre les résultats de validation et de test des modèles SVM combinés (imagerie + clinique) précédés soit de la sélection mRMR de 5 indices, soit du processus de trois étapes proposé dans ce chapitre. Les résultats sont équivalents.
3. Modèles d'apprentissage automatique : la troisième hypothèse est la mauvaise mise en œuvre des modèles d'apprentissage. Cependant, des méthodes simples d'apprentissage ont été choisies et les résultats sont cohérents entre les différentes approches. D'autres paramètres pour les quatre méthodes d'apprentissage automatique de classification ont été testés par une recherche systématique dans une grille dans l'espace des hyperparamètres pour estimer les paramètres optimaux en validation croisée interne à cinq couches, mais cela ajoute de la complexité à l'algorithme sans améliorer excessivement les résultats. Les résultats figurent dans le tableau 5.24 pour les modèles combinant les indices cliniques avec les indices d'imagerie globaux. Les hyperparamètres qui ont été testés sont présentés dans le tableau 5.25, les noms reprenant ceux de *scikit-learn*.
4. Faible nombre de sujets : la dernière hypothèse est que le nombre de sujets soit trop faible. Le nombre de patients utilisés pour l'entraînement, spécialement le nombre de patients présentant une mutation du type H3.1, peut être insuffisant pour caractériser le groupe. De plus, le nombre de patients utilisés pour le test est insuffisant pour évaluer le modèle et cette base de données présente un déséquilibre plus important entre les deux groupes (4 H3.1 et 13 H3.3) que la base de données d'entraînement (8 H3.1 et 22 H3.3). D'après les résultats, cette hypothèse de trop faible échantillon est la plus plausible.

Cette hypothèse de faible nombre de patients se renforce par les résultats des analyses individuelles des indices cliniques présentés au début de ce chapitre. Les valeurs des AUC ne sont pas stables entre les cohortes. Ainsi, plus de patients sont nécessaires pour que l'on puisse affirmer avec certitude les indices liés à chaque type de mutation et repérer les cas aberrants. Une étude récemment conduite dans le laboratoire IMIV (Dirand et collab. [2019], article soumis) a cherché à vérifier les performances de modèles radiomiques dans les conditions de petits échantillons. Ils ont adopté une stratégie de sous-échantillonnage de larges cohortes (environ 400 patients) ayant prouvé la valeur prédictive de certains indices radiomiques. Ils ont trouvé que dans les conditions de faible effectif (par exemple 40 patients : 30 de la classe 1 versus 10 de la classe 2), 80% des indices se montrant prédictifs en validation croisée n'ont aucune valeur prédictive dans une cohorte de test externe.

En vue des résultats, il reste plausible qu'il y ait suffisamment d'information dans l'imagerie pour la prédiction de la mutation de l'histone H3, mais le faible effectif des cohortes ne permet pas d'identifier des différences qui existent sur certains indices selon les deux types de mutations,

Méthode de sélection		3 étapes	mRMR
Global			
Validation (LOOCV)	Se	0,88 ± 0,03	0,71 ± 0,09
	Sp	0,63 ± 0,02	0,45 ± 0,18
	Youden	0,51 ± 0,04	0,16 ± 0,19
	F1-pond	0,81 ± 0,02	0,65 ± 0,08
Test BIOMEDE	Se	0,92 ± 0,01	0,68 ± 0,12
	Sp	0,18 ± 0,17	0,38 ± 0,24
	Youden	0,10 ± 0,17	0,05 ± 0,26
	F1-pond	0,70 ± 0,07	0,61 ± 0,10
Local			
Validation (LOOCV)	Se	0,76 ± 0,04	0,81 ± 0,06
	Sp	0,49 ± 0,10	0,48 ± 0,13
	Youden	0,25 ± 0,10	0,29 ± 0,15
	F1-pond	0,69 ± 0,04	0,72 ± 0,06
Test BIOMEDE	Se	0,74 ± 0,05	0,71 ± 0,12
	Sp	0,10 ± 0,12	0,30 ± 0,24
	Youden	-0,16 ± 0,10	0,01 ± 0,28
	F1-pond	0,58 ± 0,03	0,61 ± 0,10

TABLEAU 5.23 – Résultats de validation et de test des modèles SVM combinés (imagerie + clinique) précédés soit de la sélection de variable mRMR à cinq indices soit du processus en trois étapes proposé dans ce chapitre.

		RF	KNN	SVM	Logit
Hyperparamètres retenus pour les résultats présentés dans ce chapitre					
Validation (LOOCV)	Se	0,84 ± 0,03	0,88 ± 0,02	0,88 ± 0,03	0,83 ± 0,03
	Sp	0,57 ± 0,09	0,68 ± 0,08	0,63 ± 0,02	0,74 ± 0,04
	Youden	0,41 ± 0,09	0,57 ± 0,08	0,51 ± 0,04	0,57 ± 0,05
	F1-pond	0,77 ± 0,03	0,83 ± 0,03	0,81 ± 0,02	0,81 ± 0,02
Test BIOMEDE	Se	0,78 ± 0,09	0,79 ± 0,03	0,92 ± 0,01	0,92 ± 0,01
	Sp	0,48 ± 0,27	0,27 ± 0,17	0,18 ± 0,17	0,18 ± 0,17
	Youden	0,25 ± 0,32	0,05 ± 0,16	0,10 ± 0,17	0,10 ± 0,17
	F1-pond	0,71 ± 0,11	0,66 ± 0,04	0,70 ± 0,07	0,70 ± 0,07
Recherche par grille					
Validation (LOOCV)	Se	0,80 ± 0,05	0,84 ± 0,02	0,83 ± 0,05	0,85 ± 0,04
	Sp	0,54 ± 0,13	0,63 ± 0,08	0,68 ± 0,07	0,63 ± 0,00
	Youden	0,34 ± 0,13	0,46 ± 0,09	0,51 ± 0,08	0,47 ± 0,04
	F1-pond	0,73 ± 0,05	0,78 ± 0,03	0,80 ± 0,03	0,79 ± 0,02
Test BIOMEDE	Se	0,83 ± 0,08	0,80 ± 0,05	0,86 ± 0,05	0,87 ± 0,04
	Sp	0,31 ± 0,25	0,39 ± 0,17	0,27 ± 0,13	0,33 ± 0,15
	Youden	0,14 ± 0,25	0,19 ± 0,16	0,13 ± 0,14	0,20 ± 0,16
	F1-pond	0,69 ± 0,08	0,70 ± 0,05	0,70 ± 0,05	0,73 ± 0,05

TABLEAU 5.24 – Comparaison des résultats obtenus avec les hyperparamètres choisis pour cette thèse et par une recherche par grille pour les quatre méthodes d'apprentissage des modèles combinant les indices cliniques avec les indices d'imagerie globaux.

Méthode	Paramètres
RF	"n_estimators" : [3, 10, 50, 100] "max_depth" : [3, None] "max_features" : [1, None, 'auto'] "min_samples_split" : [2, 3, 6] "bootstrap" : [True, False] "criterion" : ["gini", "entropy"]
KNN	"n_neighbors" : [1, 2, 3, 4, 5] "weights" : ['uniform', 'distance'] "p" : [1,2,3]
SVM	"kernel" : ['linear', 'rbf'] "C" : [0,1, 1, 10]
Logit	"C" : [0,001, 0,01, 0,1, 1, 10, 100, 1000] "penalty" : ["l1", "l2"]

TABLEAU 5.25 – Paramètres utilisées pour la recherche par grille de chaque méthode d'apprentissage.

identifiant plutôt les différences spécifiques de la cohorte d'entraînement. En effet, dans les problèmes réels il est presque impossible de trouver des indices qui caractérisent parfaitement la variable à prédire et il existe souvent des données aberrantes. Par conséquent, pour qu'une étude à faible effectif soit couronnée de succès, il faut qu'il existe des indices assez fortement liés à la variable à prédire qui puissent être mis en évidence avec ces sujets et qu'il n'y ait pas trop de cas aberrants.

Finalement, dans le cas de la classification des patients selon leur mutation de de l'histone H3, il existe deux autres classes plus rares, le H3-WT et une mutation H3.2 qui n'ont pas été inclus dans le modèle car le nombre de sujets concernés était beaucoup trop faible. Ainsi, un patient qui ne présente pas de données histologiques sera prédit comme H3.1 ou H3.3 même s'il est WT. Avec une augmentation suffisante du nombre d'effectifs, au moins les patients H3-WT pourront être inclus dans le modèle de prédiction.

Prédiction de la présence de la mutation du gène TP53 : Pour cette mutation, les indices qui présentent l'AUC la plus élevée pour les patients de la [cohorte radiomique 4Mod](#) sont calculés en [T2w](#) et [FLAIR](#) mais ils sont tous éliminés par le critère de sélection de robustesse à la segmentation. A la différence des résultats obtenus pour la mutation de l'histone H3, aucun indice clinique n'est fortement lié à la mutation du gène TP53 dans la [cohorte radiomique 4Mod](#). Aucun modèle n'a présenté de résultats satisfaisants en validation ni en test, même si les performances sont meilleures que le hasard. L'ajout forcé d'un indice clinique qui n'avait pas été sélectionné dans les modèles combinés a dégradé les performances, ce qui suggère la pertinence de la méthode de sélection d'indices.

Même si la présence de la mutation TP53 est fortement liée à la mutation H3.3, les résultats de validation et de test ne sont pas concluants. Le faible nombre de sujets ne permet pas d'affirmer que les relations trouvées entre les indices radiomiques et la mutation H3 pour les patients de la [cohorte radiomique 4Mod](#) soient aussi trouvées entre les indices radiomiques et la mutation du gène TP53. Les hypothèses relatives à ces résultats restent les mêmes que celles énoncées pour expliquer les difficultés à prédire la mutation H3.3.

Prédiction de la présence de la mutation gène ACVR1 : L'étude de la mutation ACVR1 se montrait plus compliquée dès le départ avec l'analyse uni-variée des indices. Très peu d'indices présentent une AUC élevée et le déséquilibre entre les classes ACVR1 et ACVR1-WT est très important. Les indices qui présentent les valeurs de AUC les plus élevées passent les critères de sélection et sont présents dans les modèles multi-variés. Cependant, les résultats de validation et de test sont moins performants qu'un modèle trivial (score F1-pondéré < 0,69). Comme pour la prédiction de

la mutation du gène TP53, l'ajout forcé d'un indice clinique qui n'avait pas été sélectionné dans les modèles combinés a dégradé les performances, ce qui renforce la pertinence de la méthode de sélection d'indices.

Les hypothèses pour expliquer la cause de ces résultats sont les mêmes que celles énoncées pour la mutation H3.3. Le faible nombre de patients ne permet pas non plus que les relations trouvées entre les indices radiomiques et la mutation H3 pour les patients de la [cohorte radiomique 4Mod](#) soient aussi trouvées entre les indices radiomiques et la mutation ACVR1 qui est fortement associée à la mutation H3.1.

Classification des patients mutés suivant les mutations H3.1 et H3.3 avec la méthode multi-modèles : Dans la deuxième partie de ce chapitre, la méthodologie multi-modèles a été appliquée sur l'étude de la prédiction de la mutation de l'histone H3.1 versus H3.3, en considérant seulement les indices d'imagerie du type global et la méthode de régression logistique.

L'illustration d'estimation des modèles avec les patients de la base PREBIOMEDE et de test avec les patients de la base BIOMEDE a abouti à des résultats plus performants que les résultats de la première partie du chapitre. De plus, il est maintenant possible de réaliser une prédiction pour tous les patients de la base BIOMEDE.

Il existe des méthodes proposées dans la littérature pour gérer les valeurs manquantes afin de pouvoir inclure tous les patients. Parmi ces méthodes d'imputation il existe la répétition des valeurs les plus fréquentes, l'insertion de la moyenne des valeurs ou des méthodes plus complexes d'arbres ou de plus proches voisins par exemple ([BERTSIMAS et collab. \[2018\]](#)). Cependant, comme ici les données manquantes sont des images qui impliquent 79 indices et comme le nombre de patients est faible, la proportion des données artificielles par rapport aux données réelles serait trop élevée, ce qui ajouterait, encore plus d'incertitude au modèle. Pour cette raison l'imputation n'a pas été retenue.

Dans le cas d'usage présenté dans ce chapitre, le choix final de prédiction de la mutation a été fait par le meilleur modèle (selon le score F1-pondéré) qu'il pouvait utiliser. D'autres possibilités pourraient être étudiées comme faire une moyenne pondérée des réponses en prenant en compte la qualité de chaque modèle (son score de [LOOCV](#)) mais aussi la distance du patient classifié aux limites de séparation des classes dans l'espace d'indices du modèle. Cela serait certainement plus robuste et devrait être évalué, une fois que le nombre de patients sera assez élevé et les performances de [LOOCV](#) plus proches des performances de test. En effet une étude récente menée dans le laboratoire IMIV (Dirand et collab. [2019], article soumis) a montré que l'amélioration des conditions de construction et évaluation des modèles, notamment avec l'ajout d'échantillons implique une meilleure performance en terme de l'indice de Youden de validation ainsi qu'un rapprochement des performances de validation et de test externe.

Avec ces conditions de faible nombre de patients, les modèles les plus performants utilisent seulement 2 indices. L'ajout d'indices peut tromper le modèle ou conduire à un sur-apprentissage. Cela ne veut pas dire qu'il n'y a pas d'information dans les indices, mais plutôt que les différences entre les deux groupes de patients ne sont pas assez stables pour les patients présentés en test. L'ajout de davantage de patients dans la modélisation, ce qui peut être envisagé à l'issue de l'étude BIOMEDE, peut changer les résultats en arrivant à des modèles plus complexes. C'est pour cette raison que nous considérons qu'il faut continuer à analyser tous les modèles et toutes les modalités d'imagerie disponibles.

En outre, nous savons que les valeurs de score de performance des modèles M01 à M16 ne sont pas strictement comparables car le nombre de patients utilisés dans chaque modèle varie. Cependant les résultats obtenus avec la fusion des deux bases de données montrent que les écarts entre les scores des modèles tendent à s'atténuer.

Avec l'ajout d'un nombre plus important de patients, la définition des seuils des critères de sélection peut être ajustée. De même, des modifications dans la méthodologie, par exemple, l'estimation des performances de validation par validation croisée à 10 plis à la place de la [LOOCV](#) peuvent être introduites.

Finalement, l'application de la méthode a été illustrée sur la prédiction de la mutation de l'histone H3, mais elle peut évidemment être appliquée à d'autres études radiomiques sur les patients **GITC** comme la mutation des gènes ACVR1 et TP53. De plus, une analyse conjointe des trois types de mutations serait possible à condition d'augmenter suffisamment le nombre d'effectifs car plusieurs classes seraient définies.

Conclusion

Les résultats de ce chapitre montrent que les indices d'imagerie apportent de l'information quantitative sur les tumeurs **GITC** et peuvent améliorer la détermination des mutations associées à ces tumeurs en comparaison aux indices cliniques seuls. Pour les classifications des mutations associées aux patients **GITC**, les résultats obtenus par les différentes méthodes d'apprentissage automatique sont équivalents; aucune méthode ne se démarque. Les modèles radiomiques qui prédisent la mutation de l'histone H3.1 versus H3.3 présentent une meilleure performance que les modèles radiomiques pour l'identification des mutations des gènes ACVR1 et TP53. La méthodologie pour les analyses radiomiques multimodales, présentée dans la deuxième partie de ce chapitre, a l'avantage de pouvoir donner une prédiction pour tous les patients en utilisant l'ensemble des informations disponibles même quand elles sont incomplètes. Les modèles peuvent être régulièrement mis à jour pour tirer le meilleur profit possible des données disponibles. L'application de la méthodologie pour la prédiction de la mutation des histones H3.1 versus H3.3 suggère sa pertinence pour les études à données manquantes, car ses performances de test sont meilleures que les performances de test obtenues avec les modèles proposés dans la première étude du chapitre. En général, le faible nombre de patients disponibles pour la thèse est encore insuffisant pour caractériser finement les différents groupes. L'augmentation de la base de données est nécessaire pour que les modèles estimés soient valables et que les tendances montrées dans ce chapitre puissent être statistiquement validées. Les résultats sont toutefois encourageants et permettent d'avoir des perspectives intéressantes pour la suite du travail, avec l'intégration de l'ensemble des données du **PHRC BIOMEDE**.

5.5 Références

- BERTSIMAS, D., C. PAWLOWSKI et Y. D. ZHUO. 2018, «From predictive methods to missing data imputation : An optimization approach», *Journal of Machine Learning Research*, vol. 18, n° 196, p. 1–39. URL <http://jmlr.org/papers/v18/17-073.html>. 166
- CASTEL, D., C. PHILIPPE, R. CALMON, L. LE DRET, N. TRUFFAUX, N. BODDAERT, M. PAGÈS, K. R. TAYLOR, P. SAULNIER, L. LACROIX, A. MACKAY, C. JONES, C. SAINTE-ROSE, T. BLAUWBLOMME, F. ANDREIUOLO, S. PUGET, J. GRILL, P. VARLET et M.-A. DEBILY. 2015, «Histone h3f3a and hist1h3b k27m mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes», *Acta Neuropathologica*, vol. 130, n° 6, doi : 10.1007/s00401-015-1478-0, p. 815–827, ISSN 1432-0533. URL <https://doi.org/10.1007/s00401-015-1478-0>. 126
- CHAWLA, N., K. BOWYER, L. HALL et W. KEGELMEYER. 2002, «Smote : Synthetic minority over-sampling technique», *JAIR*, vol. 16, p. 321–357. 139
- COHEN, K. J., N. JABADO et J. GRILL. 2017, «Diffuse intrinsic pontine gliomas - Current management and new biologic insights. Is there a glimmer of hope?», *Neuro-Oncology*, vol. 19, n° 8, doi :10.1093/neuonc/nox021, p. 11 025 – 1034, ISSN 15235866. 126

- KOO, T. K. et M. Y. LI. 2016, «A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research», *Journal of Chiropractic Medicine*, vol. 15, n° 2, doi : 10.1016/j.jcm.2016.02.012, p. 155–163, ISSN 15563707. 136
- MACKAY, A., A. BURFORD, D. CARVALHO, E. IZQUIERDO, J. FAZAL-SALOM, K. R. TAYLOR, L. BJERKE, M. CLARKE, M. VINCI, M. NANDHABALAN, S. TEMELSO, S. POPOV, V. MOLINARI, P. RAMAN, A. J. WAANDERS, H. J. HAN, S. GUPTA, L. MARSHALL, S. ZACHAROULIS, S. VAIDYA, H. C. MANDEVILLE, L. R. BRIDGES, A. J. MARTIN, S. AL-SARRAJ, C. CHANDLER, H.-K. NG, X. LI, K. MU, S. TRABELSI, D. H.-B. BRAHIM, A. N. KISLJAKOV, D. M. KONOVALOV, A. S. MOORE, A. M. CARCABOSO, M. SUNOL, C. DE TORRES, O. CRUZ, J. MORA, L. I. SHATS, J. N. STAVALE, L. T. BIDINOTTO, R. M. REIS, N. ENTZ-WERLE, M. FARRELL, J. CRYAN, D. CRIMMINS, J. CAIRD, J. PEARS, M. MONJE, M.-A. DEBILY, D. CASTEL, J. GRILL, C. HAWKINS, H. NIKBAKHT, N. JABADO, S. J. BAKER, S. M. PFISTER, D. T. W. JONES, M. FOULADI, A. O. VON BUEREN, M. BAUDIS, A. RESNICK et C. JONES. 2017, «Integrated molecular meta-analysis of 1,000 pediatric high-grade and diffuse intrinsic pontine glioma», *Cancer cell*, vol. 32, n° 4, doi :10.1016/j.ccell.2017.08.017, p. 520—537.e5, ISSN 1535-6108. URL <http://europepmc.org/articles/PMC5637314>. 126
- PENG, H., F. LONG et C. DING. 2005, «Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, n° 8, doi :10.1109/TPAMI.2005.159, p. 1226–1238, ISSN 0162-8828. 163

Chapitre 6

Études radiomiques appliquées aux analyses de survie des patients atteints de GITC

« Ainsi mon dessein n'est pas d'enseigner la méthode que chacun doit suivre pour bien conduire sa raison, mais de faire voir en quelle sorte j'ai tâché de conduire la mienne. »

René Descartes



Ce chapitre considère les analyses radiomiques pour la prédiction de la survie globale des patients GITC. La radiomique a été évaluée pour la définition d'une signature de survie globale des patients GITC afin de définir deux groupes de patients ayant des survies globales différentes. Le chapitre commence par une analyse uni-variée réalisée pour chaque indice (indices cliniques, indices d'imagerie locaux et indices d'imagerie globaux) suivie d'une procédure de sélection d'indices proposée dans le chapitre précédent et finit avec la définition et évaluation des modèles radiomiques multi-variés proposés en mettant en œuvre différentes techniques d'apprentissage automatique.

Sommaire

6.1 Analyses uni-variées de la survie globale des patients atteints de GITC	170
6.1.1 Indices cliniques	170
6.1.2 Indices d'imagerie	175
6.2 Analyses multi-variées de la survie globale des patients GITC	180
6.2.1 Sélection d'indices	180
6.2.2 Modèles de régression	181
6.2.3 Résultats de prédiction de la survie globale	182
6.3 Discussion et conclusion	186
6.4 Références	189

6.1 Analyses uni-variées de la survie globale des patients atteints de GTC

En partant des images de diagnostic et des indices cliniques, nous avons étudié s’il était possible d’identifier les patients qui décèdent plus tôt que les autres. Contrairement aux études du chapitre 5, les patients des cohortes PREBIOMEDE et BIOMEDE ne peuvent pas être fusionnées, car pour tous les patients de la cohorte PREBIOMEDE, le traitement suivi était la radiothérapie alors que pour les patients de la cohorte BIOMEDE, les traitements choisis dépendaient également des mutations génétiques présentes dans la tumeur. L’effet des différents traitements est illustré sur la figure 6.1 où les courbes Kaplan-Meier sont présentées séparément par type de mutation pour les cohortes PREBIOMEDE et BIOMEDE. Tous les patients des deux cohortes qui présentaient les informations nécessaires (mutation correspondante et temps de survie) ont été utilisés pour la construction de chaque courbe. Les valeurs-p du test du Logrank sont affichées sur chaque graphe ainsi que les tableaux du nombre des patients à risque à différents temps d’analyse.

Il est à remarquer que pour les patients de la base PREBIOMEDE, les mutations sont de bons prédicteurs de la survie globale : la présence de la mutation TP53 est plus agressive que son absence, les tumeurs qui présentent une mutation de H3.3 suggèrent un risque plus élevé de décès et celles qui présentent une mutation H3.1 et l’absence de la mutation ACVR1 est plus agressive que sa présence. Cependant, avec les différents choix de thérapie qui ont été faits pour les patients de la cohorte BIOMEDE, les résultats trouvés avec PREBIOMEDE ne sont plus valables, ce qui tend à prouver une certaine efficacité de ces thérapies ciblées. C’est la raison pour laquelle, il n’est plus possible de fusionner l’ensemble des patients sans tenir compte de leur spécificité puisque les patients font partie d’essais thérapeutiques différents. Nous avons donc restreint l’analyse de la survie aux patients PREBIOMEDE.

Pour évaluer les relations entre les différents indices radiomiques et la survie globale des patients, j’ai commencé par l’analyse du comportement de chaque indice individuellement. Les mesures utilisées pour évaluer les indices ont été le RH, l’indice C et le test de ratio de vraisemblance pour les modèles de régression de Cox uni-variés à risque proportionnel. Tous les développements de ce chapitre ont été faits en langage R avec l’utilisation des bibliothèques *Survival*, *glmnet* et *randomForestSRC*.

6.1.1 Indices cliniques

Pour les 38 patients de la [cohorte radiomique 4Mod](#) les valeurs de l’indice C, du RH et la valeur-p du test de ratio de vraisemblance pour le volume, l’âge au diagnostic et le sexe du patient sont présentées dans le tableau 6.1. Tous ces indices respectent la condition des hasards proportionnels.

TABLEAU 6.1 – Cohorte radiomique 4Mod : évaluation des modèles uni-variés cliniques de survie globale.

	RH (95% Intervalle de confiance)	C-index (erreur-type)	Ratio de vraisemblance
Age	0,945 (0,856; 1,043)	0,546 (0,057)	p = 0,243
Volume	1 (0,9999; 1)	0,551 (0,057)	p = 0,595
Sexe	0,638 (0,323; 1,264)	0,551 (0,050)	p = 0,193

Pour tous les indices cliniques, l’intervalle de confiance du RH inclut la valeur 1, la valeur de l’indice C n’est pas élevée et les tests du ratio de vraisemblance ne rejettent pas l’hypothèse d’équivalence avec le modèle nul. Ces indices semblent être ainsi non-predictifs de la survie globale.

La figure 6.2 montre les valeurs des âges au diagnostic versus la survie globale et les courbes de Kaplan-Meier des patients séparés en deux groupes en fonction de l’âge médian des patients au moment du diagnostic.

La figure 6.3 montre les valeurs des volumes de VOI_{Tumeur} versus la survie globale et les courbes Kaplan-Meier des patients séparés en deux groupes en fonction du volume de VOI_{Tumeur} médian

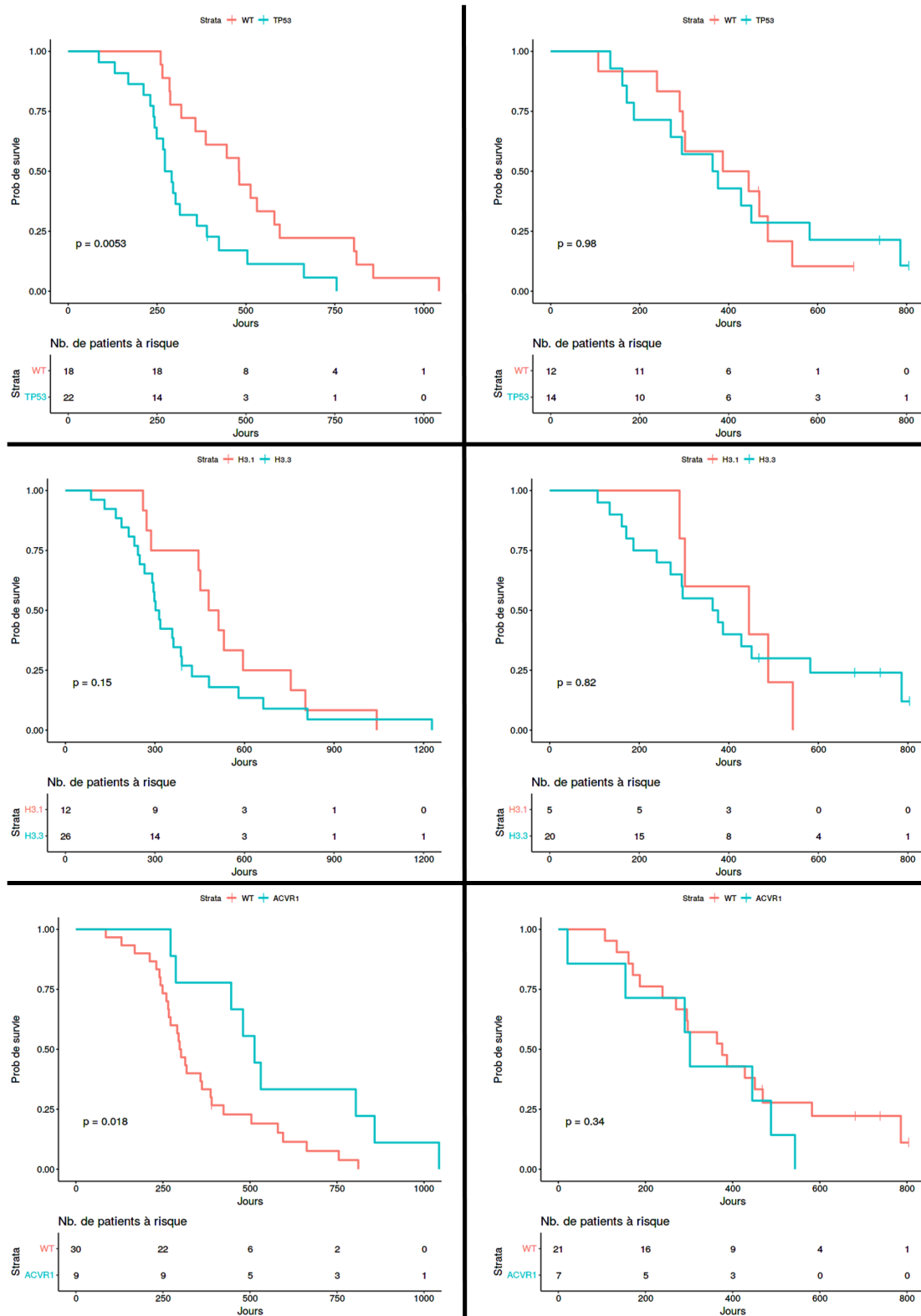


FIGURE 6.1 – Les courbes Kaplan-Meier de survie globale à partir du diagnostic pour les 38 patients de la base PREBIOMEDE (première colonne) et les patients de la base BIOMEDE (deuxième colonne). Les patients sont séparés en fonction des mutations TP53 (en haut), H3 (au milieu) et ACVR1 (en bas).

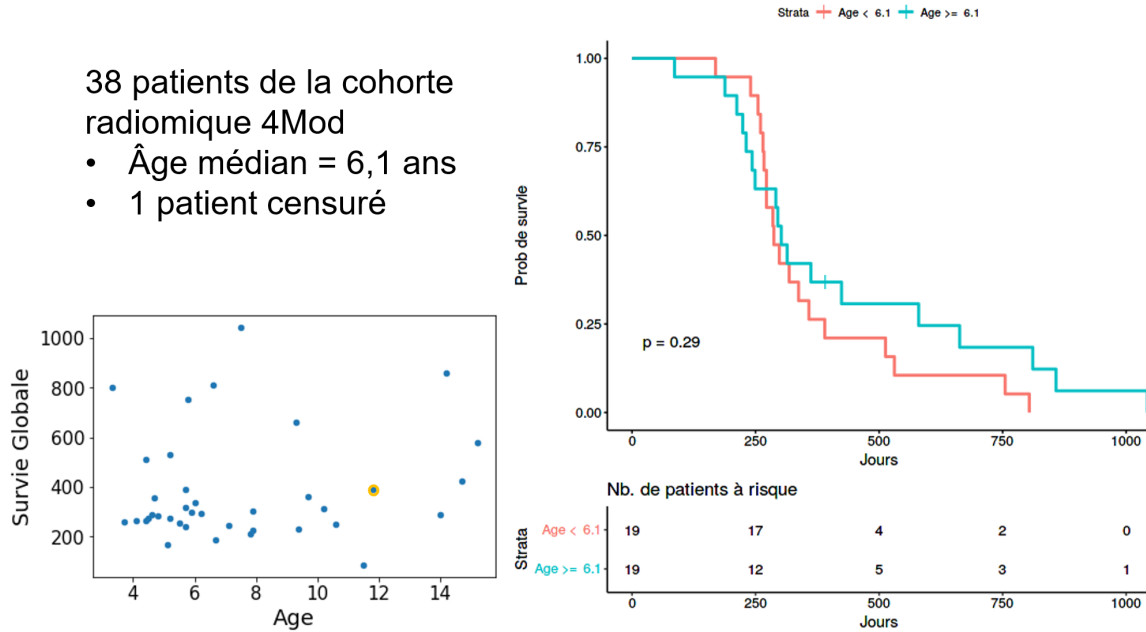


FIGURE 6.2 – A gauche, la survie globale en jours versus les âges au diagnostic pour les 38 patients de la cohorte radiomique 4Mod. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients qui ont l'âge au diagnostic plus grand (en turquoise) et plus petit (en corail) que l'âge médian. La valeur-p du test de log-rank est affichée. Sous les courbes Kaplan-Meier, le tableau avec le nombre de patients à risque associé est affiché.

des patients.

Finalement, la figure 6.4 montre l'histogramme des survies globales pour les filles et les garçons, ainsi que les courbes Kaplan-Meier correspondantes.

Les mêmes analyses ont été faites pour les 47 patients de la cohorte PREBIOMEDE pour lesquels les informations nécessaires étaient disponibles. Le tableau 6.2 montre les valeurs des RHs, indice C et les valeurs-p des tests de ratio de vraisemblance.

TABEAU 6.2 – Cohorte PREBIOMEDE : Évaluation des modèles uni-variés cliniques de survie globale.

	RH (95% Intervalle de confiance)	C-index (erreur-type)	Ratio de vraisemblance
Age	0,993 (0,917; 1,076)	0,500 (0,051)	p = 0,865
Volume	1 (0,9999; 1)	0,556 (0,051)	p = 0,373
Sexe	0,828 (0,457; 1,502)	0,536 (0,044)	p = 0,533

La figure 6.5 montre les valeurs des âges au diagnostic versus la survie globale et les courbes Kaplan-Meier des 47 patients de la cohorte PREBIOMEDE séparés en deux groupes : âge plus petit que l'âge médian de tous les patients et âge plus élevé que l'âge médian de tous les patients.

La figure 6.6 montre les valeurs des volumes de VOI_{Tumeur} versus la survie globale et les courbes Kaplan-Meier des 47 patients de la cohorte PREBIOMEDE séparés en deux groupes par le volume de la VOI_{Tumeur} plus grand ou plus petit que le volume de VOI_{Tumeur} médian de tous les patients.

Finalement, la figure 6.7 montre l'histogramme des survies globales pour les filles et les garçons, ainsi que les courbes Kaplan-Meier correspondantes aux 47 patients de la cohorte PREBIOMEDE.

Même en augmentant les effectifs, les indices cliniques ne s'avèrent pas des bons prédicteurs de la survie globale. Le volume de la tumeur VOI_{Tumeur} est le paramètre qui semble le plus prédictif, mais les différences restent non-significatives (valeur-p du test logRank > 0,05).

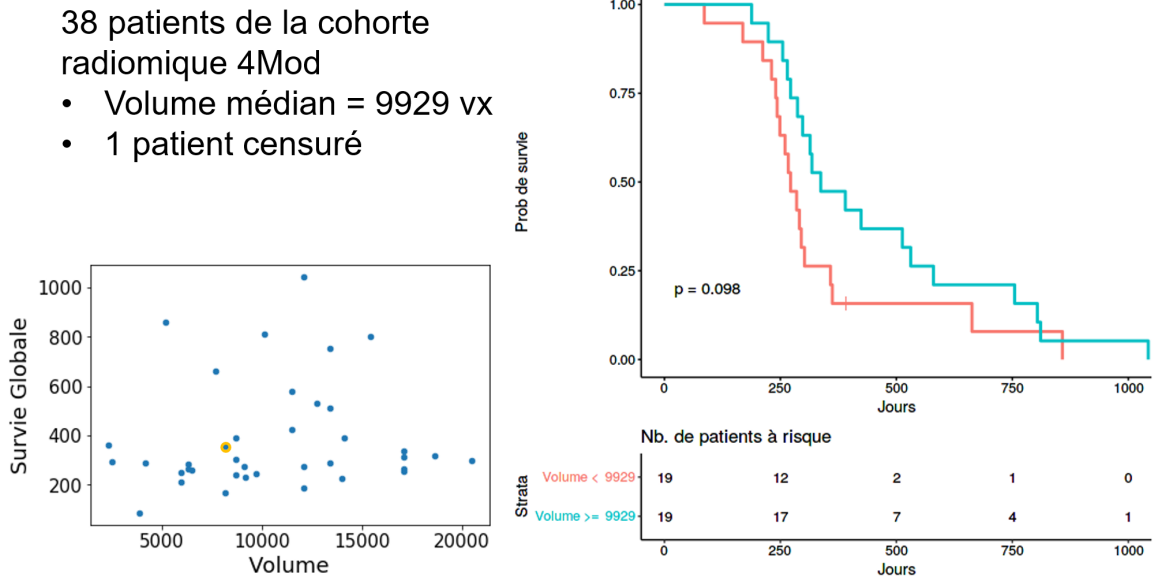


FIGURE 6.3 – A gauche, la survie globale en jours versus les volumes de VOI_{tumeur} pour les 38 patients de la cohorte radiomique 4Mod. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients qui ont le volume plus grand (en turquoise) et plus petit (en corail) que le volume médian. La valeur-p du test de log-rank est affichée. Sous les courbes Kaplan-Meier, le tableau avec le nombre de patients à risque associé est affiché.

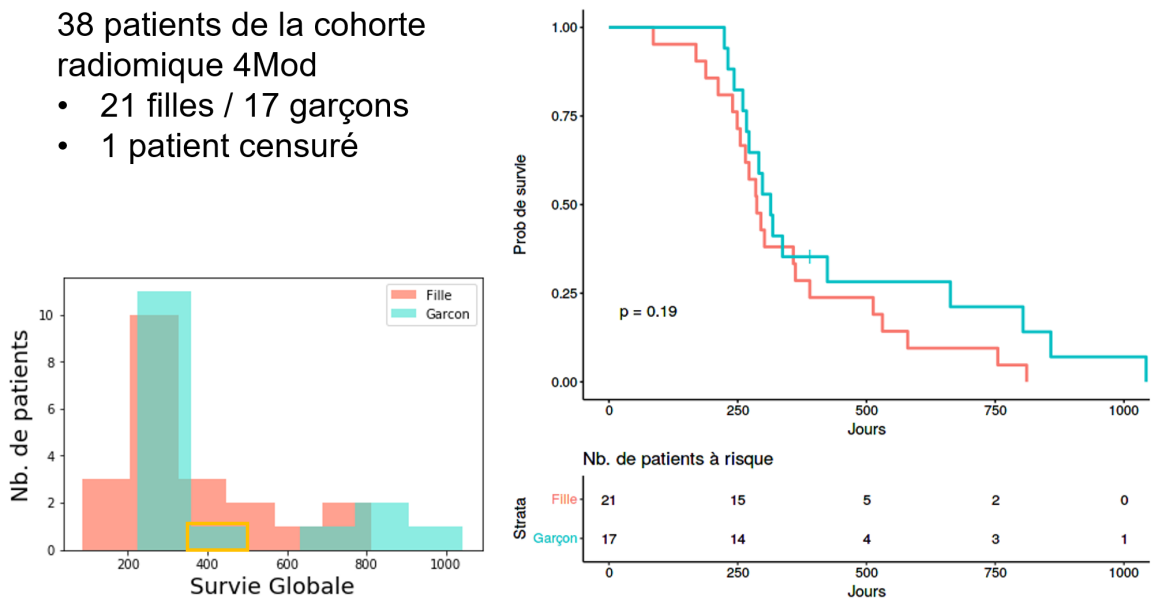


FIGURE 6.4 – A gauche, les histogrammes de la survie globale en jours pour les filles (en corail) et les garçons (en turquoise) de la cohorte radiomique 4Mod. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients filles (en corail) et garçons (en turquoise). La valeur-p du test de log-rank est affichée. Sous les courbes Kaplan-Meier, le tableau avec le nombre de patients à risque associé est affiché.

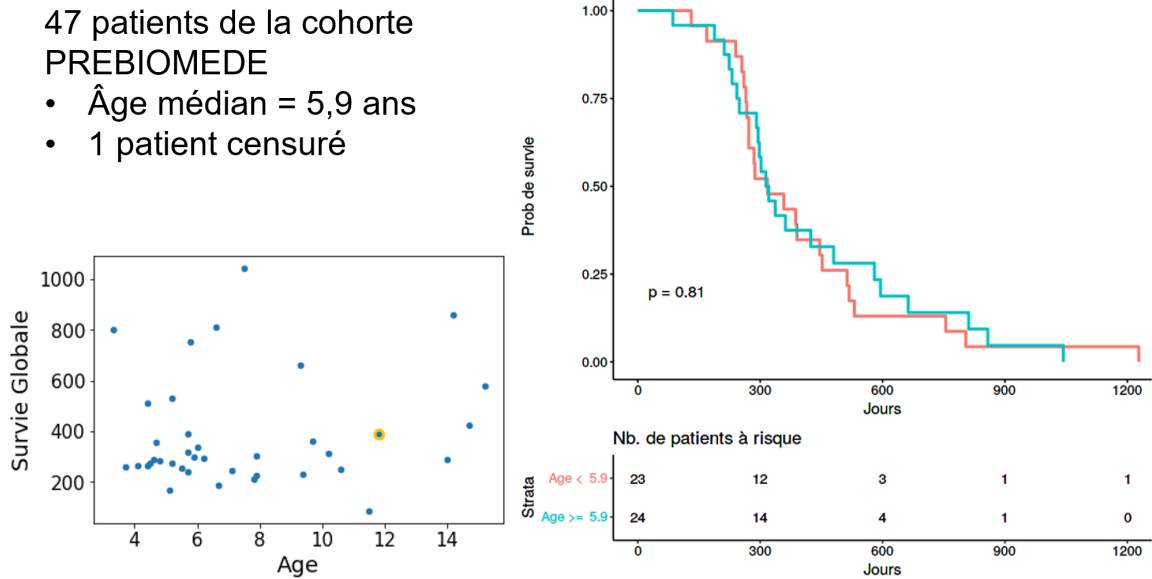


FIGURE 6.5 – A gauche, la survie globale en jours versus les âges au diagnostic pour les 47 patients de la cohorte PREBIOMEDE. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients qui ont l'âge au diagnostic plus grand (en turquoise) et plus petit (en corail) que l'âge médian. La valeur-p du test de log-rank est affichée. Sous les courbes Kaplan-Meier, le tableau avec le nombre de patients à risque associé est affiché.

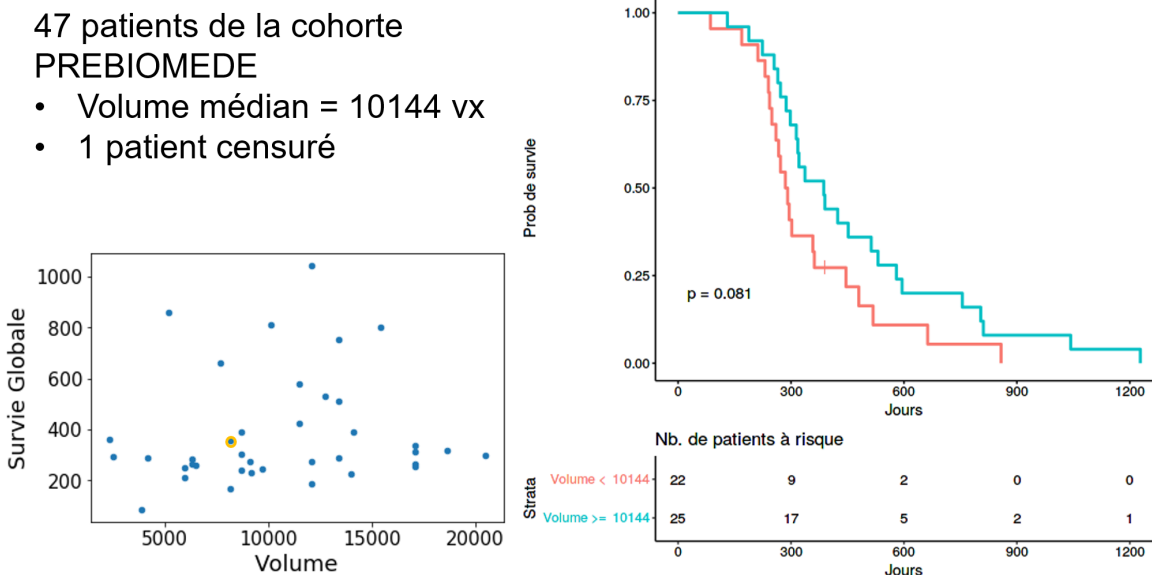


FIGURE 6.6 – A gauche, la survie globale en jours versus les volumes de VOI_{tumeur} pour les 47 patients de la cohorte PREBIOMEDE. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients qui ont le volume plus grand (en turquoise) et plus petit (en corail) que le volume médian. La valeur-p du test de log-rank est affichée. Sous les courbes Kaplan-Meier, le tableau avec le nombre de patients à risque associé est affiché.

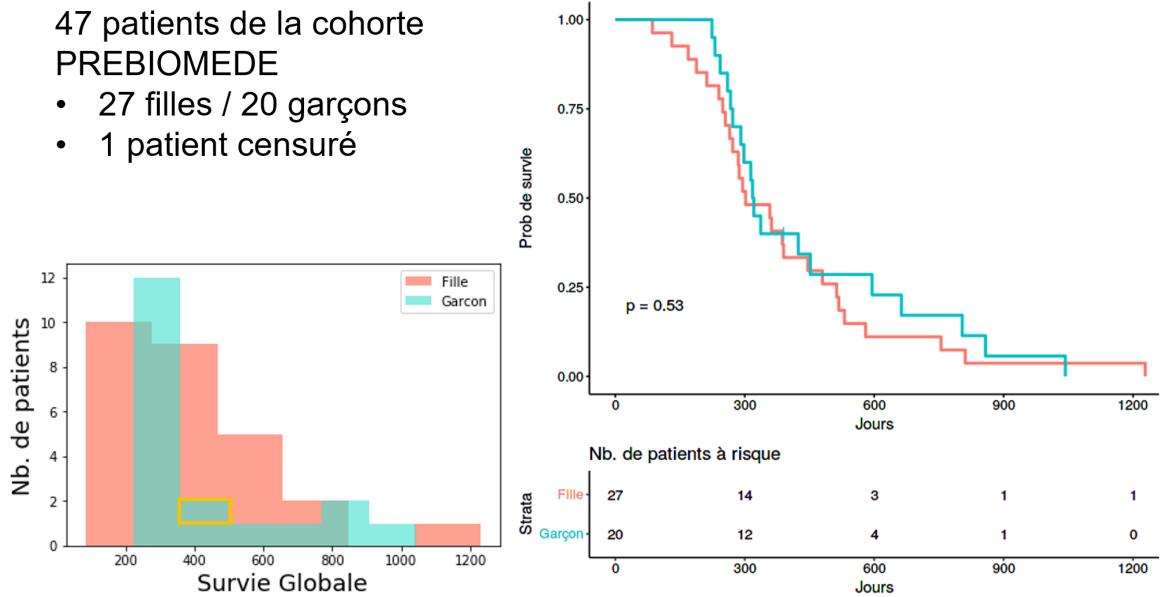


FIGURE 6.7 – A gauche, les histogrammes de la survie globale en jours pour les filles (en corail) et les garçons (en turquoise) de la cohorte PREBIOMEDE. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients filles (en corail) et garçons (en turquoise). La valeur-p du test de log-rank est affichée. Sous les courbes Kaplan-Meier, le tableau avec le nombre de patients à risque associé est affiché.

6.1.2 Indices d'imagerie

Les indices d'imagerie ont été normalisés par la soustraction de la moyenne et division par l'écart-type pour que les RH soient plus facilement interprétables et pas trop impactés par les indices qui présentent des valeurs très faibles ou très élevées. Dans le cas d'indices normalisés, un RH égal à 1,25 veut dire que l'augmentation d'un écart-type de la valeur de l'indice implique une probabilité 1,25 plus élevée de décès.

Pour les 38 patients de la [cohorte radiomique 4Mod](#) les valeurs de l'indice C, du RH et la valeur-p du test de ratio de vraisemblance ont aussi été utilisées pour évaluer les indices d'imagerie. Pour les indices globaux d'imagerie calculés dans les images T1w les résultats sont présentés dans le tableau en annexe A.7, pour les images T2w dans le tableau en annexe A.8, pour les images T1c dans le tableau en annexe A.9 et pour les images FLAIR dans le tableau en annexe A.10. Les équivalents pour les indices locaux sont présentés dans les tableaux en annexe A.11, A.12, A.13 et A.14. Tous ces indices respectent la condition de hasards proportionnels.

Parmi les 316 indices d'imagerie globaux (316 indices d'imagerie locaux), seulement sept (quatre) indices présentent un ratio de vraisemblance statistiquement significatif ($valeur - p < 0,05$) avec une valeur de RH qui n'inclut pas 1,00 dans son intervalle de confiance. Ces indices (globaux : T1w firstorder Kurtosis, T2w firstorder Skewness, T2w firstorder Maximum, T2w firstorder Kurtosis, FLAIR firstorder Skewness, FLAIR glcm ClusterShade, FLAIR glszm LowIntensityLargeAreaEmphasis; locaux : T2w firstorder Skewness, T2w glcm ClusterShade, FLAIR firstorder Skewness, FLAIR glcm ClusterShade) sont présentés dans le tableau 6.3 et sont en vert dans les tableaux annexes A.7-A.14. Comme pour l'étude de la classification de la mutation, il y a une prévalence d'indices du premier ordre. Ils représentent environ deux tiers des 11 indices. En plus de ces indices, d'autres indices présentent une valeur d'indice C assez élevée ($> 0,55$) mais la valeur de RH inclut 1,00 dans l'intervalle de confiance. Ces indices sont en bleu dans les tableaux annexes A.7-A.14 et ils sont majoritairement originaires des modalités T1c et FLAIR.

La figure 6.8 présente les valeurs des quatre indices locaux qui ont un ratio de vraisemblance statistiquement significatif avec la survie globale pour les 38 patients de la [cohorte radiomique](#)

TABLEAU 6.3 – Cohorte Radiomique 4Mod : évaluation des indices locaux calculés en FLAIR pour la prédiction de la survie globale.

Indices d'imagerie	RH (95% Intervalle de confiance)	C-index (erreur-type)	Ratio de vraisemblance
Globaux			
T1w.firstorder.Kurtosis	2,25 (1,24; 4,08)	0,60 (0,06)	p = 0,01
T2w.firstorder.Skewness	1,58 (1,04; 2,39)	0,57 (0,06)	p = 0,03
T2w.firstorder.Maximum	1,54 (1,07; 2,23)	0,64 (0,06)	p = 0,02
T2w.firstorder.Kurtosis	2,34 (1,45; 3,79)	0,59 (0,06)	p < 0,01
FLAIR.firstorder.Skewness	1,80 (1,21; 2,68)	0,63 (0,06)	p < 0,01
FLAIR.glc.ClusterShade	1,79 (1,26; 2,55)	0,65 (0,06)	p = 0,01
FLAIR.glszm.LowIntensityLargeAreaEmphasis	1,63 (1,11; 2,39)	0,57 (0,06)	p = 0,03
Locaux			
T2w.firstorder.Skewness	1,55 (1,06; 2,25)	0,61 (0,06)	p = 0,02
T2w.glc.ClusterShade	2,24 (1,38; 3,63)	0,60 (0,06)	p < 0,01
FLAIR.firstorder.Skewness	1,46 (1,03; 2,05)	0,61 (0,06)	p = 0,04
FLAIR.glc.ClusterShade	1,68 (1,05; 2,69)	0,61 (0,06)	p = 0,04

4Mod versus le temps de survie globale ainsi que les courbes Kaplan-Meier de deux groupes stratifiés par la valeur médiane de l'indice. On peut remarquer que l'indice *ClusterShade*, une mesure de l'asymétrie sur la moyenne des intensités de la matrice GLCM, présente des écarts de valeurs très élevés pour certains patients. Il n'est pas robuste et sera éliminé par la sélection d'indices dans la suite.

La figure 6.9 présente les valeurs des quatre indices globaux calculés dans les images T1w et T2w qui présentent un ratio de vraisemblance statistiquement significatif avec la survie globale pour les 38 patients de la cohorte radiomique 4Mod versus le temps de survie globale ainsi que les courbes Kaplan-Meier de deux groupes stratifiés par la valeur médiane de l'indice. La figure 6.10 présente le même type de résultat pour les indices globaux calculés dans les images FLAIR et T1c qui présentent le ratio de vraisemblance statistiquement significatif. Pour T1c firstorder Kurtosis les courbes Kaplan-Meier des deux groupes séparés par la valeur médiane sont similaires; cependant, si un autre seuil avait été choisi (par exemple, la valeur 4,00), les deux groupes présenteraient des courbes assez différentes.

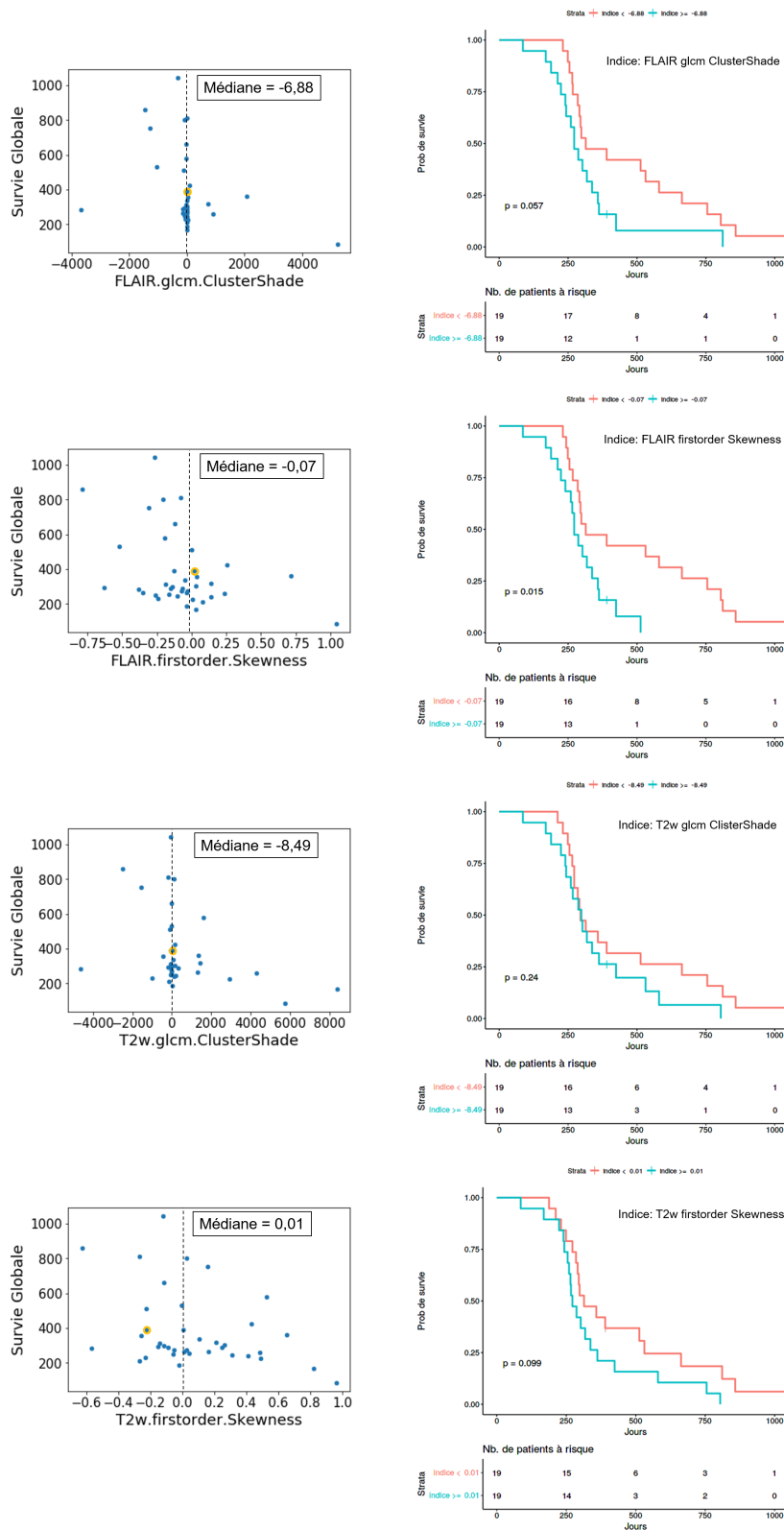


FIGURE 6.8 – A gauche, la survie globale en jours versus les valeurs des indices locaux qui présentent un ratio de vraisemblance statistiquement significatif pour la cohorte radiomique 4Mod. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients stratifiés par la valeur médiane de l'indice. Les valeurs-p des tests de log-rank sont affichées. Sous les courbes Kaplan-Meier, les tableaux avec le nombre de patients à risque associé sont affichés.

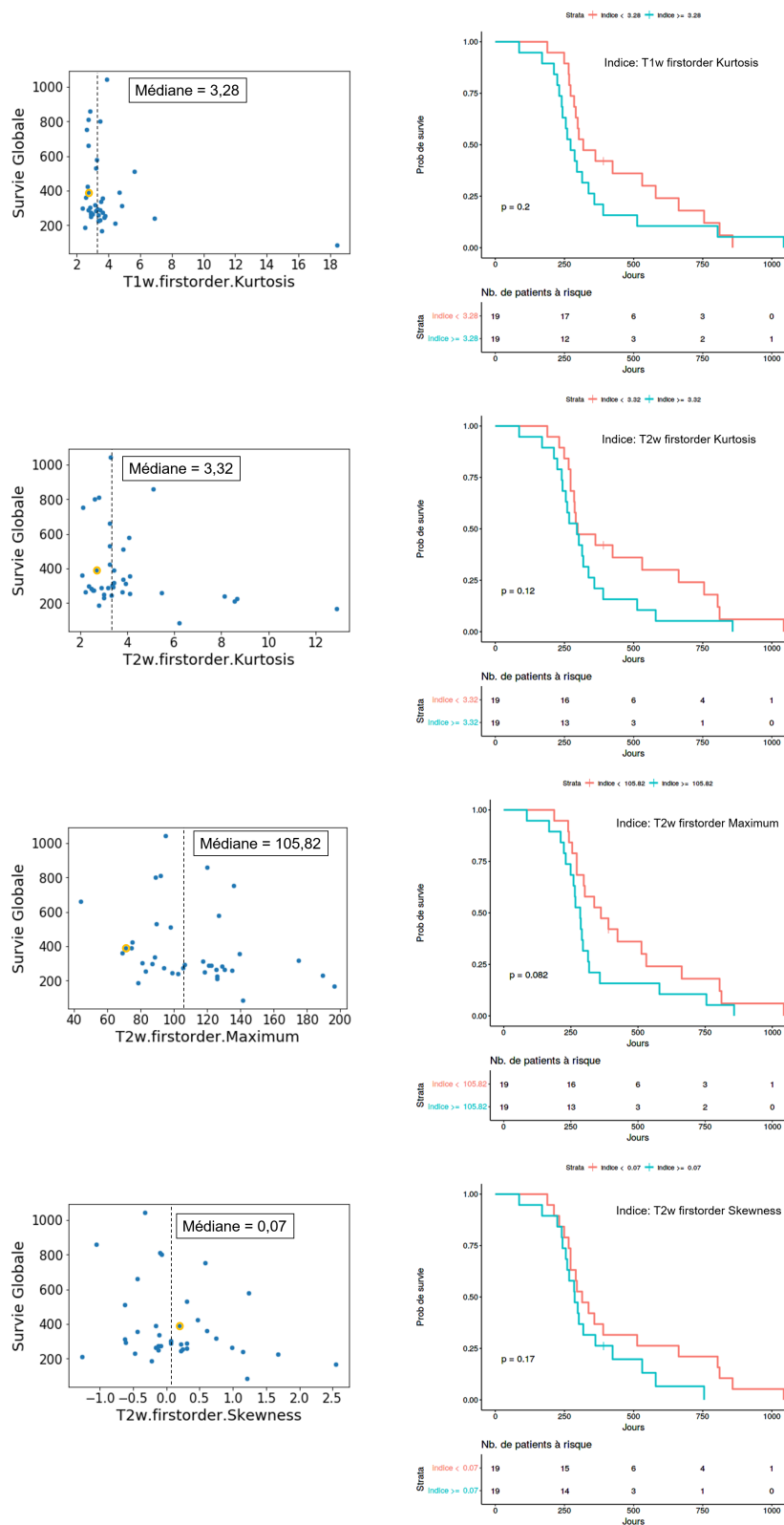


FIGURE 6.9 – A gauche, la survie globale en jours versus les valeurs des indices globaux calculés en T1w et T2w qui présentent un ratio de vraisemblance statistiquement significatif pour la cohorte radiomique 4Mod. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients stratifiés par la valeur médiane de l'indice. Les valeurs-p des tests de log-rank sont affichées. Sous les courbes Kaplan-Meier, les tableaux avec le nombre de patients à risque associé sont affichés.

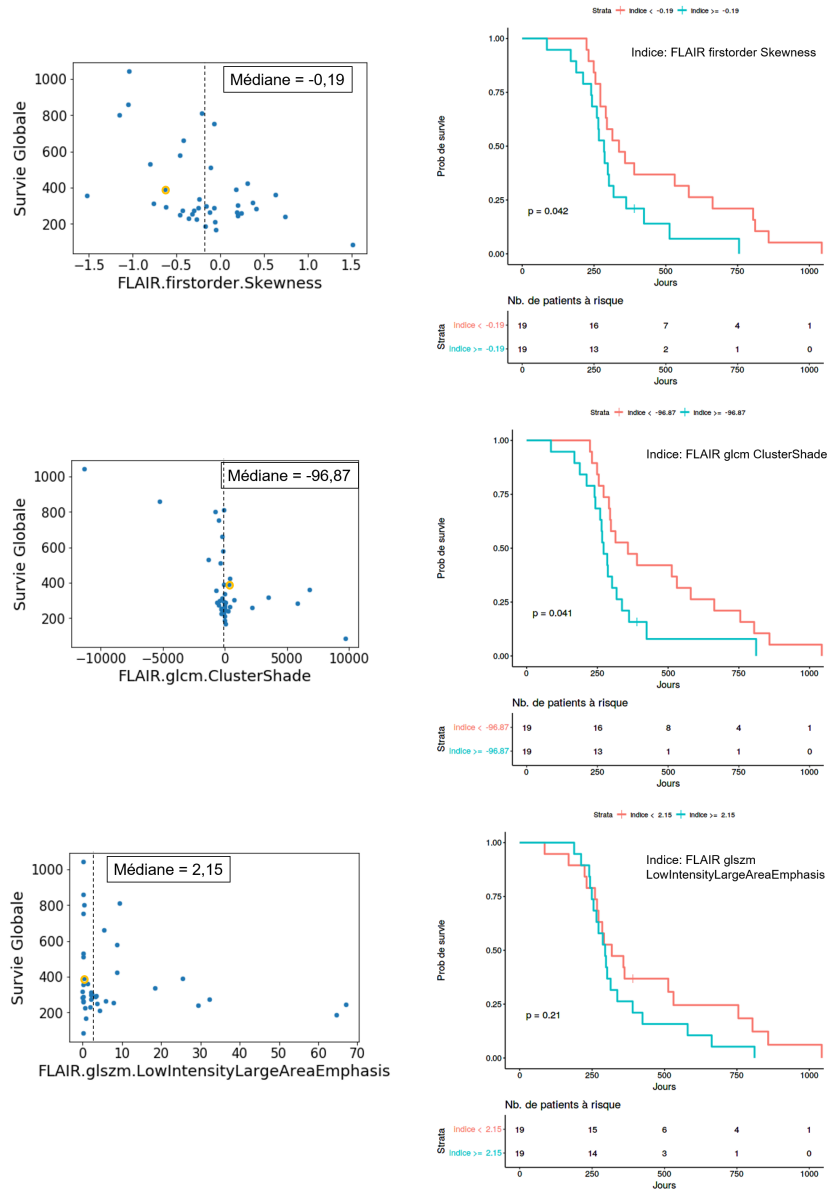


FIGURE 6.10 – A gauche, la survie globale en jours versus les valeurs des trois indices globaux calculés en FLAIR qui présentent un ratio de vraisemblance statistiquement significatif pour la cohorte radiomique 4Mod. Le patient censuré est affiché avec le contour jaune. A droite, les courbes Kaplan-Meier des patients stratifiés par la valeur médiane de l'indice. Les valeurs-p des tests de log-rank sont affichées. Sous les courbes Kaplan-Meier, les tableaux avec le nombre de patients à risque associé sont affichés.

6.2 Analyses multi-variées de la survie globale des patients GITC

L'analyse multi-variée de la survie globale a été conduite comme pour les études de classification de différentes mutations : en incluant seulement les indices cliniques; en incluant seulement les indices d'imagerie; en combinant les indices cliniques et d'imagerie. Cependant, comme montré sur la figure 6.1, pour l'étude de la survie globale, il est impossible d'analyser ensemble les cohortes BIOMEDE et PREBIOMEDE parce que les patients ont reçu des traitements différents qui ont un impact sur leur durée de survie. Ainsi, dans cette étude d'analyse de la survie globale, seulement les résultats de validation (en LOOCV) des 38 patients de la [cohorte radiomique 4Mod](#) sont présentés.

6.2.1 Sélection d'indices

Pour l'analyse multi-variée de la survie globale, les mêmes étapes de sélection d'indices utilisées pour l'analyse des mutations ont été appliquées. La seule adaptation faite est le changement du critère de $AUC > 0,75$ par le critère $\text{indice } C > 0,55$.

La figure 6.11 montre le nombre d'indices moyen (\pm l'écart-type) qui ont été sélectionnés à chaque étape pour chaque pli de LOOCV. Après les trois étapes de sélection, 13 ± 2 indices locaux et 7 ± 1 indices globaux ont été sélectionnés dans chaque pli. Les indices sélectionnés dans chaque pli sont souvent les mêmes, sept indices locaux (*FLAIR glszm LowIntensitySmallAreaEmphasis*, *T1c firstorder Maximum*, *T1c glrlm ShortRunLowGrayLevelEmphasis*, *T2w firstorder Maximum*, *T2w glszm ZoneVariance*, *T1w firstorder Energy*, *T1w glszm IntensityVariability*) et six globaux (*FLAIR firstorder 90Percentile*, *T1c glcm DifferenceEntropy*, *T1c glcm DifferenceVariance*, *T2w firstorder Median*, *T2w glrlm RunVariance*, *T1w glszm ZoneVariance*) ont été sélectionnés dans plus de 26 plis (sur les 38). Considérant les indices cliniques, le sexe et le volume ont été sélectionnés dans 22 (sur 38) plis de LOOCV, alors que l'âge au diagnostic a été sélectionné dans 11 plis. Il y a eu cinq plis de LOOCV où aucun paramètre clinique n'a été sélectionné, conduisant à des modèles combinés équivalents aux modèles d'imagerie. Parmi les indices d'imagerie fréquemment sélectionnés, il n'y a aucun des 11 indices les plus performants dans les analyses uni-variées, ils sont tous éliminés dans la première étape de sélection (la robustesse aux dilatations et érosions des VOI_{tumeur}), cependant, les indices sélectionnés sont en majorité près du seuil considéré pour mettre en évidence les indices les plus performants en analyse uni-variée (cf. tableaux A.7-A.14).

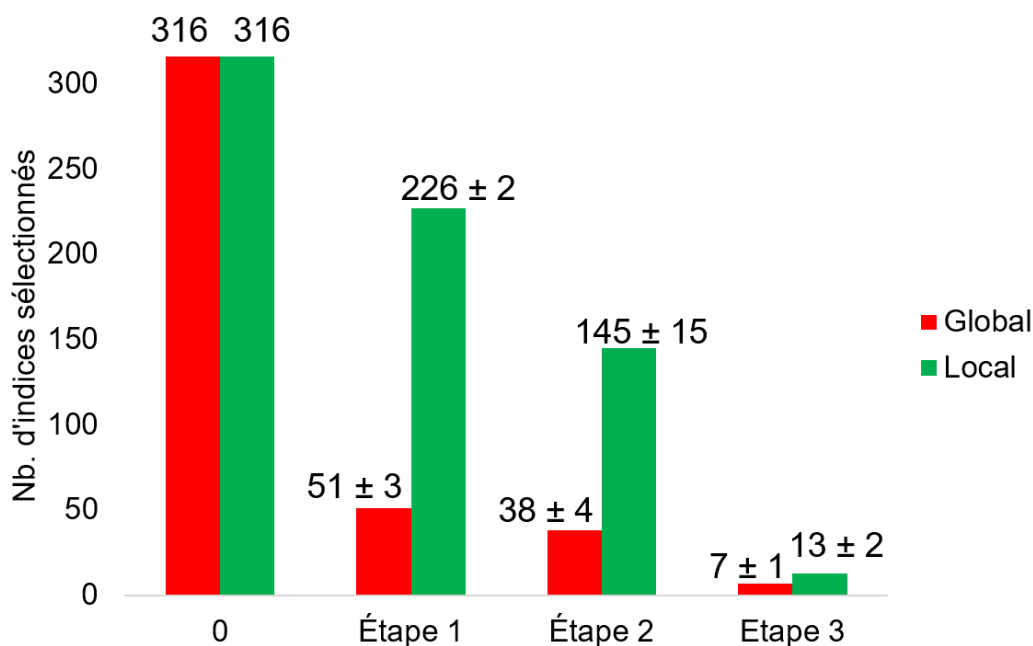


FIGURE 6.11 – Nombre moyen d'indices sélectionnés à chaque étape de la procédure de sélection d'indice pour les 38 patients de la cohorte 4Mod.

6.2.2 Modèles de régression

Pour l'ensemble des 38 plis de LOOCV les indices ont été normalisés avec la soustraction de la moyenne et la division par l'écart-type en considérant les 37 patients d'entraînement. Les mêmes valeurs des paramètres de standardisation ont été utilisées pour standardiser les indices du patient de validation.

Comme pour les études de classification, il n'existe pas de méthode de régression reconnue comme la plus performante dans tous les cas. Par conséquent, trois méthodes fréquemment appliquées dans la littérature ont été utilisées. La première est le modèle de Cox à risques proportionnels avec régression Ridge, la seconde, le modèle de Cox à risques proportionnels combiné avec une méthode de contraction des coefficients de la régression LASSO, et la troisième, la méthode des forêts aléatoires de survie avec un algorithme de chasse de variables (variable hunting, présenté par CHEN et ISHWARAN [2013]).

Pour chaque pli de LOOCV la valeur médiane des risques estimés pour les patients de l'entraînement a été utilisée comme seuil pour classer le patient de test comme un patient à risque élevé ou un patient à risque faible. Tous les patients classés comme à risque faible ou élevé par la régression de la LOOCV estimée sans eux-mêmes ont été regroupés ensemble. Une courbe de Kaplan-Meier pour les patients classés comme à faible risque, ainsi qu'une courbe de Kaplan-Meier pour les patients classés comme à risque élevé ont été tracées.

Pour classer chaque patient dans les deux courbes de Kaplan-Meier, tous les autres patients ont été utilisés en les regroupant; ainsi les observations ne sont pas indépendantes entre elles. Une stratégie de permutation a été utilisée pour évaluer la significativité statistique car il n'est pas possible d'utiliser les courbes produites avec le LOOCV directement à cause des observations dépendantes. Le processus complet d'analyse de survie globale (sélection, régression, calcul du test de logRank des courbes de validation) a été répété 500 fois avec des permutations aléatoires des temps de survie globale des patients (SIMON et collab. [2011]). La significativité statistique du test de logRank a été estimée par la distribution des 500 valeurs du test obtenues par les permutations.

Modèle de Cox à risques proportionnels avec régression Ridge : Pour l'étude rapportée dans ce paragraphe, la valeur de λ_R a été choisie dans chaque pli de LOOCV par la valeur qui minimise l'erreur d'une validation croisée (interne à la LOOCV) à 10 plis.

Modèle de Cox à risques proportionnels avec LASSO : Comme pour la régression Ridge, la valeur de λ_L a été choisie dans chaque pli de LOOCV par la valeur qui minimise l'erreur d'une validation croisée (interne à la LOOCV) à 10 plis.

Modèle de forêt aléatoire de survie avec sélection d'indices par profondeur minimale : En utilisant le paquet R *randomForestSRC*, les paramètres par défaut ont été utilisés : 500 arbres et taille des nœuds terminaux égale à 2.

6.2.3 Résultats de prédiction de la survie globale

Les courbes de Kaplan-Meier de validation sont présentées sur la figure 6.12. Pour la comparaison, les courbes de Kaplan-Meier d'entraînement sont présentées sur la figure 6.13, elles ont été estimées en utilisant les 38 patients de la *cohorte radiomique 4Mod* pour estimer le modèle de survie et le seuil pour séparer les deux groupes. Malgré les bons résultats d'entraînement, la méthode de Cox avec régression Ridge mélange voire inverse les courbes de validation des patients à risque élevé et à faible risque pour tous les types d'indices. La méthode de Cox avec LASSO présente ce comportement pour les modèles qui ont comme entrée les indices globaux alors que les forêts aléatoires de survie avec sélection d'indices par profondeur minimale ont tendance à mélanger les deux groupes (risque faible et risque élevé) pour tous les types d'entrée sauf les indices d'imagerie locaux, où la séparation va dans le sens attendu.

Les méthodes de Cox avec LASSO et forêts aléatoires avec sélection d'indices sont plus performantes en termes de validation que la méthode de Cox avec régression Ridge, pour les configurations d'entrée utilisant les indices d'imagerie locaux. Cependant, les valeurs-p du test du logRank ne sont pas statistiquement significatives (valeur-p > 0.16) pour aucun des 15 cas présentés sur la figure 6.12, sauf pour le cas du modèle Cox avec LASSO et indices d'imagerie globaux (valeur-p = 0,01), qui inverse les courbes. C'est le seul cas où les deux courbes (haut risque et bas risque) n'ont pas d'intersection.

Pour les modèles d'indices d'imagerie globaux seuls, en moyenne, $1,4 \pm 0,6$ indices sont sélectionnés par LASSO, et $2,0 \pm 0,0$ par la réduction de profondeurs minimale des forêts aléatoires. Pour les modèles combinés d'indices d'imagerie globaux combinés avec les indices cliniques, en moyenne, $1,9 \pm 0,9$ indices sont sélectionnés par LASSO, et $2,0 \pm 0,2$ par la réduction de profondeurs minimale des forêts aléatoires. Pour les modèles d'indices d'imagerie locaux seuls, en moyenne, $3,1 \pm 0,9$ indices sont sélectionnés par LASSO, et $2,9 \pm 0,3$ par la réduction des profondeurs minimales des forêts aléatoires. Pour les modèles combinés d'indices d'imagerie locaux combinés avec les indices cliniques, en moyenne, $1,9 \pm 0,9$ indices sont sélectionnés par LASSO, et $2,9 \pm 0,5$ par la réduction de profondeurs minimale des forêts aléatoires.

Pour les modèles de Cox avec LASSO et forêts aléatoires avec les indices d'imagerie locaux qui sont le plus performants en LOOCV, les indices utilisés dans chaque pli sont montrés dans le tableau 6.4.

Pour le modèle de Cox avec LASSO, la combinaison d'indices la plus courante est *FLAIR glszm LowIntensitySmallAreaEmphasis* et *T2w firstorder Maximum*, combinés 32 fois, quelques fois avec d'autres indices supplémentaires. Ces deux indices figurent parmi les indices le plus fréquemment sélectionnés dans les plis de LOOCV. Même si ces deux indices ne sont pas parmi les 11 indices les plus performants en analyse uni-variée, ils présentent une bonne performance : *FLAIR glszm LowIntensitySmallAreaEmphasis* (RH = 1,43 (1,00; 2,06); indice C = 0,60 (0,06); valeur-p (ratio de vraisemblance) = 0,06, cf. tableau A.14) et *T2w firstorder Maximum* (RH = 1,29 (0,90; 1,85); indice C = 0,59 (0,06); valeur-p (ratio de vraisemblance) = 0,18, cf. tableau A.12). Pour les forêts aléatoires, la combinaison d'indices inclus dans le modèle est moins stable mais l'indice *FLAIR glszm LowIntensitySmallAreaEmphasis* est également fréquemment sélectionné (19 fois). Les valeurs de

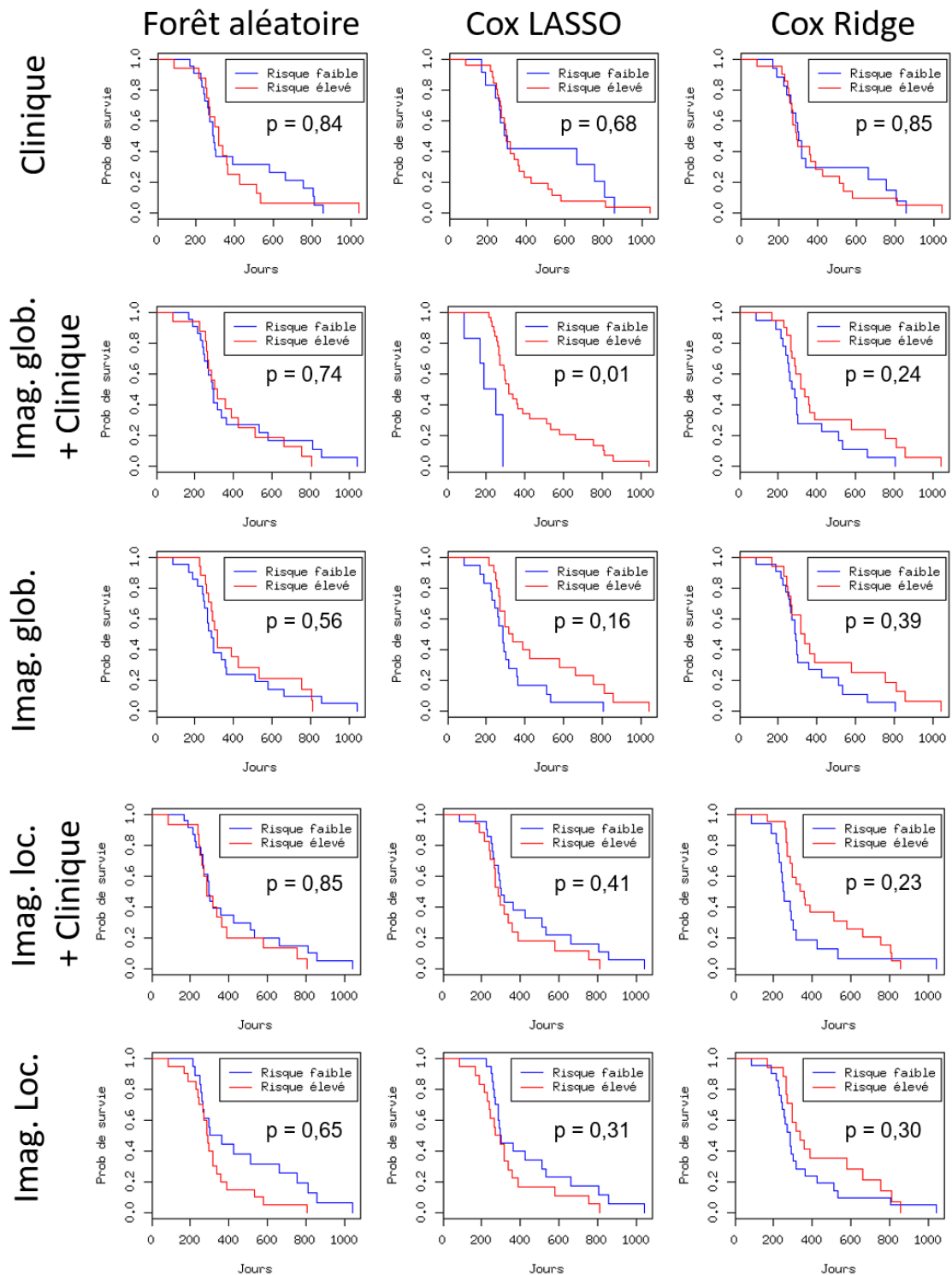


FIGURE 6.12 – Courbes de Kaplan-Meier de validation résultant de la classification du patient de validation de chaque pli de la LOOCV. Les valeurs-p des test logRank sont affichées sur chaque graphe.

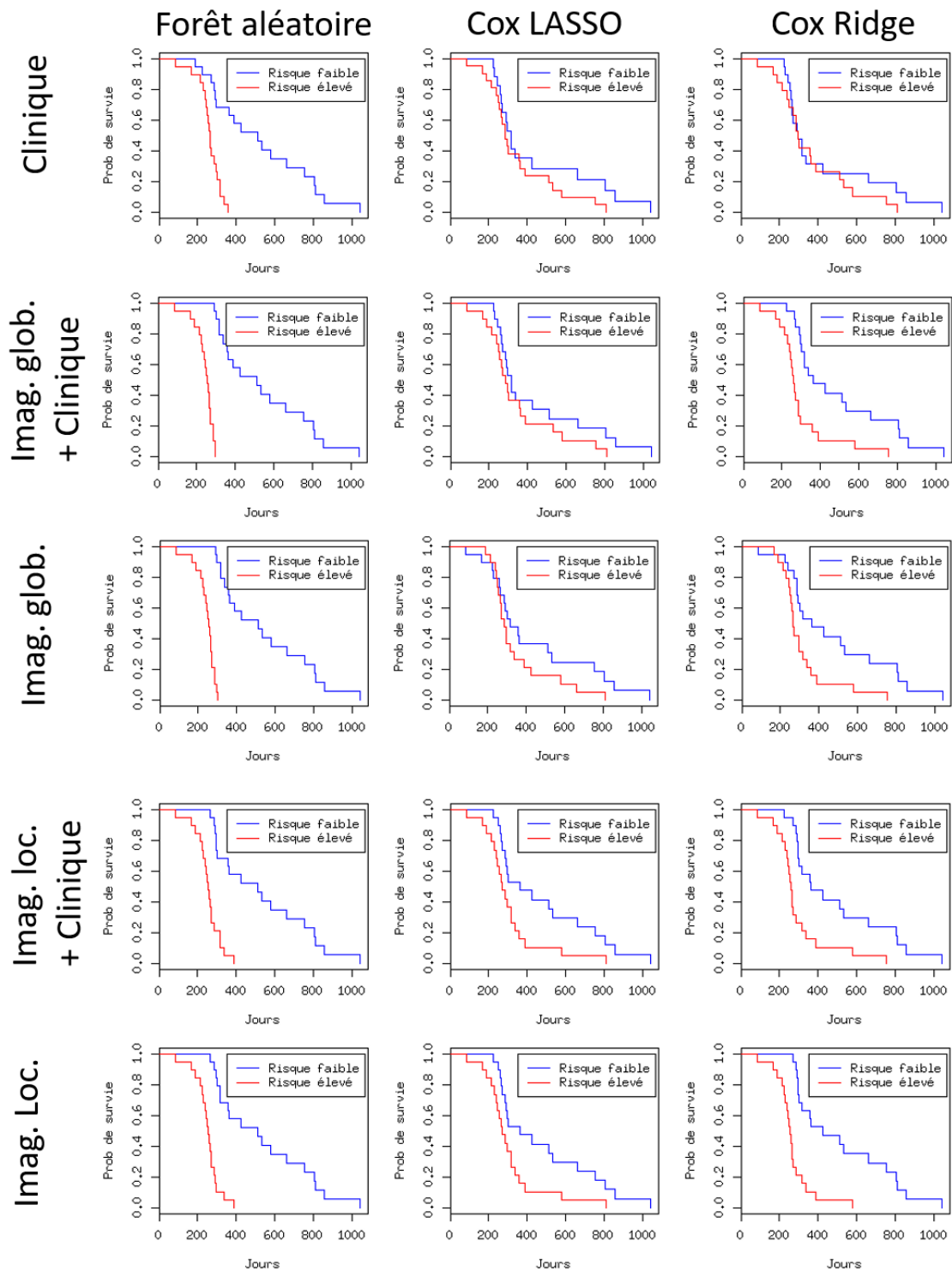


FIGURE 6.13 – Courbes de Kaplan-Meier d’entraînement, calculées avec les patients utilisés pour estimer la signature de survie et séparés par la valeur médiane comme étant à risque faible ou élevé de décès.

pli	Forêt aléatoire	Cox LASSO
1	T1c firstorder Range, T1c glrlm HighGrayLevelRunEmphasis	FLAIR glrlm LongRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy, T1w glszm HighIntensityLargeAreaEmphasis
2	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glcm Dissimilarity, FLAIR glcm lmc1	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
3	FLAIR glcm lmc1, T2w glszm ZoneVariance, T1w glrlm ShortRunLowGrayLevelEmphasis	T1c firstorder Maximum, T2w firstorder Maximum
4	FLAIR glcm lmc1, FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glcm ClusterShade	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum
5	FLAIR firstorder RobustMeanAbsoluteDeviation, FLAIR glszm IntensityVariability	FLAIR glrlm LongRunLowGrayLevelEmphasis, T1c firstorder Maximum, T2w firstorder Maximum
6	T1c glrlm LongRunLowGrayLevelEmphasis, FLAIR glcm lmc1, T1c firstorder Maximum	FLAIR glrlm LongRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy, T1w glszm HighIntensityLargeAreaEmphasis
7	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm HighGrayLevelRunEmphasis, T1c glszm SmallAreaEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm LongRunLowGrayLevelEmphasis, T1c glszm IntensityVariability, T2w firstorder Maximum
8	T1c glrlm ShortRunLowGrayLevelEmphasis, T2w glszm SmallAreaEmphasis, T1w glrlm ShortRunLowGrayLevelEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
9	FLAIR glcm lmc1, FLAIR glszm LowIntensitySmallAreaEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
10	FLAIR firstorder RobustMeanAbsoluteDeviation, T1c glszm LowIntensitySmallAreaEmphasis, FLAIR glszm LowIntensityEmphasis	FLAIR glszm LowIntensityEmphasis, T2w firstorder Maximum, T1w firstorder Energy
11	FLAIR glcm lmc1, T1c firstorder Maximum, T1c glszm LowIntensitySmallAreaEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
12	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glszm LowIntensitySmallAreaEmphasis, FLAIR glcm lmc1, T1c firstorder Maximum	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm LowGrayLevelRunEmphasis, T2w firstorder Maximum
13	FLAIR glcm lmc1, FLAIR glszm LowIntensitySmallAreaEmphasis , T1c firstorder Maximum	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
14	T1w glszm IntensityVariability, T1c glcm DifferenceVariance, T1w glrlm ShortRunLowGrayLevelEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
15	FLAIR firstorder RobustMeanAbsoluteDeviation, FLAIR firstorder Maximum, FLAIR glszm LowIntensitySmallAreaEmphasis , T1c firstorder Maximum	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c firstorder Median, T1c glszm SmallAreaEmphasis, T1c glrlm LongRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
16	FLAIR glszm LowIntensitySmallAreaEmphasis , FLAIR firstorder RobustMeanAbsoluteDeviation, T2w glszm SmallAreaEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c firstorder Median, T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
17	FLAIR firstorder RobustMeanAbsoluteDeviation, T1c firstorder Median	FLAIR glrlm LongRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy, T1w glszm HighIntensityLargeAreaEmphasis
18	T1c firstorder Range, FLAIR glcm lmc1, FLAIR glszm LowIntensityEmphasis	FLAIR glszm LowIntensityEmphasis, T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
19	FLAIR firstorder RobustMeanAbsoluteDeviation, FLAIR firstorder Maximum, T1c firstorder Median	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum, T1w firstorder Energy
20	FLAIR glcm lmc1, T2w firstorder Maximum, FLAIR glrlm LongRunLowGrayLevelEmphasis	FLAIR glrlm LongRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
21	FLAIR glszm LowIntensitySmallAreaEmphasis , FLAIR firstorder Median, T1c glrlm HighGrayLevelRunEmphasis	FLAIR firstorder Median, FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
22	FLAIR firstorder RobustMeanAbsoluteDeviation, FLAIR glszm LowIntensitySmallAreaEmphasis , T1w glszm HighIntensityLargeAreaEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
23	T1c firstorder Maximum, T1w glrlm ShortRunLowGrayLevelEmphasis, T1w glszm IntensityVariability	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
24	FLAIR firstorder RobustMeanAbsoluteDeviation, FLAIR glszm IntensityVariability, FLAIR glszm LowIntensitySmallAreaEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum
25	FLAIR glszm LowIntensitySmallAreaEmphasis , T1w firstorder Energy, FLAIR glcm lmc1	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
26	FLAIR firstorder RobustMeanAbsoluteDeviation, T1c firstorder Maximum, FLAIR firstorder Maximum	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
27	FLAIR glszm LowIntensitySmallAreaEmphasis , FLAIR glcm lmc1, T1w glszm IntensityVariability	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w firstorder Energy
28	T1c firstorder Maximum T1c glrlm RunVariance	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
29	FLAIR firstorder Maximum, T1c firstorder Maximum, T1w glszm IntensityVariability	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glrlm ShortRunLowGrayLevelEmphasis, T2w firstorder Maximum, T1w glszm HighIntensityLargeAreaEmphasis
30	T1c firstorder Maximum, FLAIR firstorder RobustMeanAbsoluteDeviation	FLAIR glrlm LongRunLowGrayLevelEmphasis, T2w firstorder Maximum
31	T1c firstorder Maximum, FLAIR firstorder RobustMeanAbsoluteDeviation, FLAIR glszm LowIntensitySmallAreaEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c firstorder Median, T2w firstorder Maximum
32	FLAIR glszm LowIntensitySmallAreaEmphasis , T1c firstorder Maximum, T1w glszm IntensityVariability	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
33	T1c glrlm ShortRunLowGrayLevelEmphasis, FLAIR glcm lmc1, FLAIR glszm LowIntensitySmallAreaEmphasis	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
34	FLAIR glszm LowIntensitySmallAreaEmphasis , FLAIR firstorder RobustMeanAbsoluteDeviation, T1c firstorder Maximum	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
35	T1c glcm DifferenceVariance, T1w glszm IntensityVariability, FLAIR firstorder RobustMeanAbsoluteDeviation	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
36	FLAIR firstorder RobustMeanAbsoluteDeviation, FLAIR glszm LowIntensitySmallAreaEmphasis , T1w firstorder Energy	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
37	FLAIR glcm lmc1, FLAIR glszm LowIntensitySmallAreaEmphasis , T1c glcm DifferenceVariance	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum
38	T1w glszm HighIntensityLargeAreaEmphasis, T1c firstorder Maximum, T1w glszm IntensityVariability	FLAIR glszm LowIntensitySmallAreaEmphasis , T2w firstorder Maximum, T1w firstorder Energy

TABLEAU 6.4 – Indices d'imagerie locaux sélectionnés à chaque pli de la LOOCV pour les modèles de forêt aléatoire et Cox avec LASSO. Les indices les plus fréquemment sélectionnés sont en gras.

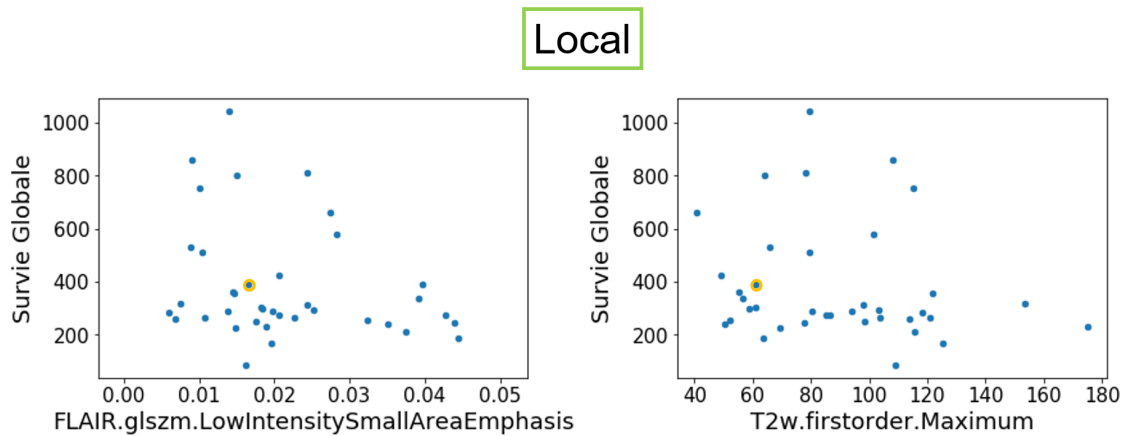


FIGURE 6.14 – La survie globale en jours versus les valeurs des indices locaux *FLAIR glszm LowIntensitySmallAreaEmphasis* et *T2w firstorder Maximum* pour la cohorte Radiomique 4Mod. Le patient censuré est affiché avec le contour jaune.

ces indices versus sa survie globale sont illustrées sur la figure 6.14. Les tumeurs présentant les valeurs de *FLAIR glszm LowIntensitySmallAreaEmphasis* et *T2w firstorder Maximum* locaux élevées tendent à avoir une durée de survie courte.

Les résultats suggèrent aussi que pour une base de données de faible effectif comme PREBIO-MEDE les modèles qui utilisent très peu d'indices (environ 3) sont les plus performants. Cependant, la base de données semble être trop petite pour trouver des modèles avec des résultats statistiquement significatifs, même si quelques indices semblent porter de l'information sur la survie globale.

6.3 Discussion et conclusion

Dans ce chapitre, des analyses radiomiques ont été conduites avec les images de diagnostic et les indices cliniques pour répondre à une tâche de prédiction de la survie globale des patients en les classant comme étant à risque faible ou élevé de décès. Les expérimentations ont été réalisées avec un sous-ensemble de la cohorte PREBIO-MEDE, la [cohorte radiomique 4Mod](#), pour laquelle les patients présentaient les quatre types d'imagerie structurale (*T1w*, *T2w*, *T1c* et *FLAIR*) au moment du diagnostic.

Il a été montré (cf. figure 6.1) que pour les patients PREBIO-MEDE qui ont reçu le même traitement, les mutations étaient des bons prédicteurs de la survie globale. Ainsi, la figure 6.15 montre les courbes de survie des patients stratifiés par leur classification en [LOOCV](#) de la mutation de l'histone H3 pour le modèle de régression logistique combinant les indices d'imagerie globaux et les indices cliniques. On peut noter qu'en fonction de la prédiction de la mutation de l'histone H3, les courbes Kaplan-Meier sont écartées. Les informations sur la survie globale sont donc présentes de manière indirecte. Cependant, ce modèle ne sera plus valable sur les nouveaux patients de la cohorte BIOMEDE puisque les mutations sont utilisées à l'heure actuelle pour guider les nouveaux traitements qui modifient les temps de survie des patients.

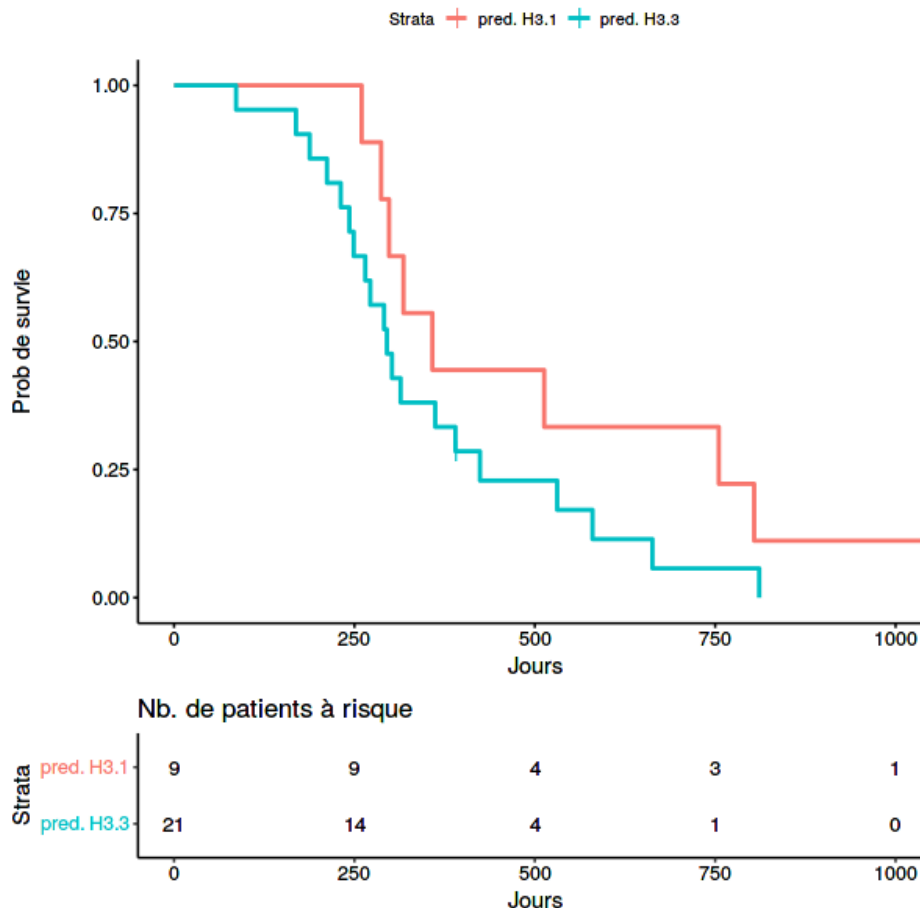


FIGURE 6.15 – Courbes de Kaplan-Meier des patients prédits en LOOCV comme mutés en H3.1 et H3.3 par une régression logistique faite sur un modèle combinant les indices radiomiques d’imagerie globaux et les indices cliniques.

Comme mentionné dans les chapitre 2, les études à petits effectifs de [HARWARD et collab. \[2018\]](#) (17 patients) et [HARGRAVE et collab. \[2008\]](#) (39 patients) n’ont trouvé aucun indice radiologique prédicteur de la survie. En supposant que les indices d’hétérogénéité en imagerie contiendraient des informations pertinentes, les analyses radiomiques pour la prédiction de la survie ont été réalisées. Même si les performances des modèles finaux ne sont pas statistiquement significatives, nous avons trouvé des indices locaux en [T2w](#) (*T2w firstorder Maximum*) et en [FLAIR](#) (*FLAIR glszm LowIntensitySmallAreaEmphasis*) assez prometteurs, car ils performent bien individuellement et sont souvent sélectionnés dans les modèles multi-variés les plus performants.

Les modèles multi-variés les plus performants sont la forêt aléatoire avec sélection d’indices par profondeur minimale et le modèle de Cox avec LASSO qui utilisent en moyenne deux indices d’imagerie globaux pour la [cohorte radiomique 4Mod](#). Parmi ces indices, il n’y avait aucun indice originaire des images [T1c](#), alors que la présence d’un anneau rehaussé sur les images [T1c](#) était un indice présent dans la signature proposée par [JANSEN et collab. \[2014\]](#) et validée par [VELDHUIJZEN VAN ZANTEN et collab. \[2017\]](#). Des tests supplémentaires devront être réalisés pour vérifier nos résultats.

Dans cette étude, la séparation des patients à risque faible et à risque élevé a été déterminée par la valeur médiane de la signature de régression de survie. Les résultats ont été aussi évalués pour une séparation des patients à chaque pli de la [LOOCV](#) par le seuil qui minimise la valeur-p du test de logrank des patients d’entraînement. Les résultats correspondant à cette procédure sont affichés sur la figure 6.16. Ils ne changent pas de façon notable les résultats mais les performances des modèles de forêt aléatoires et Cox LASSO sont moins bonnes que celles obtenues avec une séparation par la médiane.

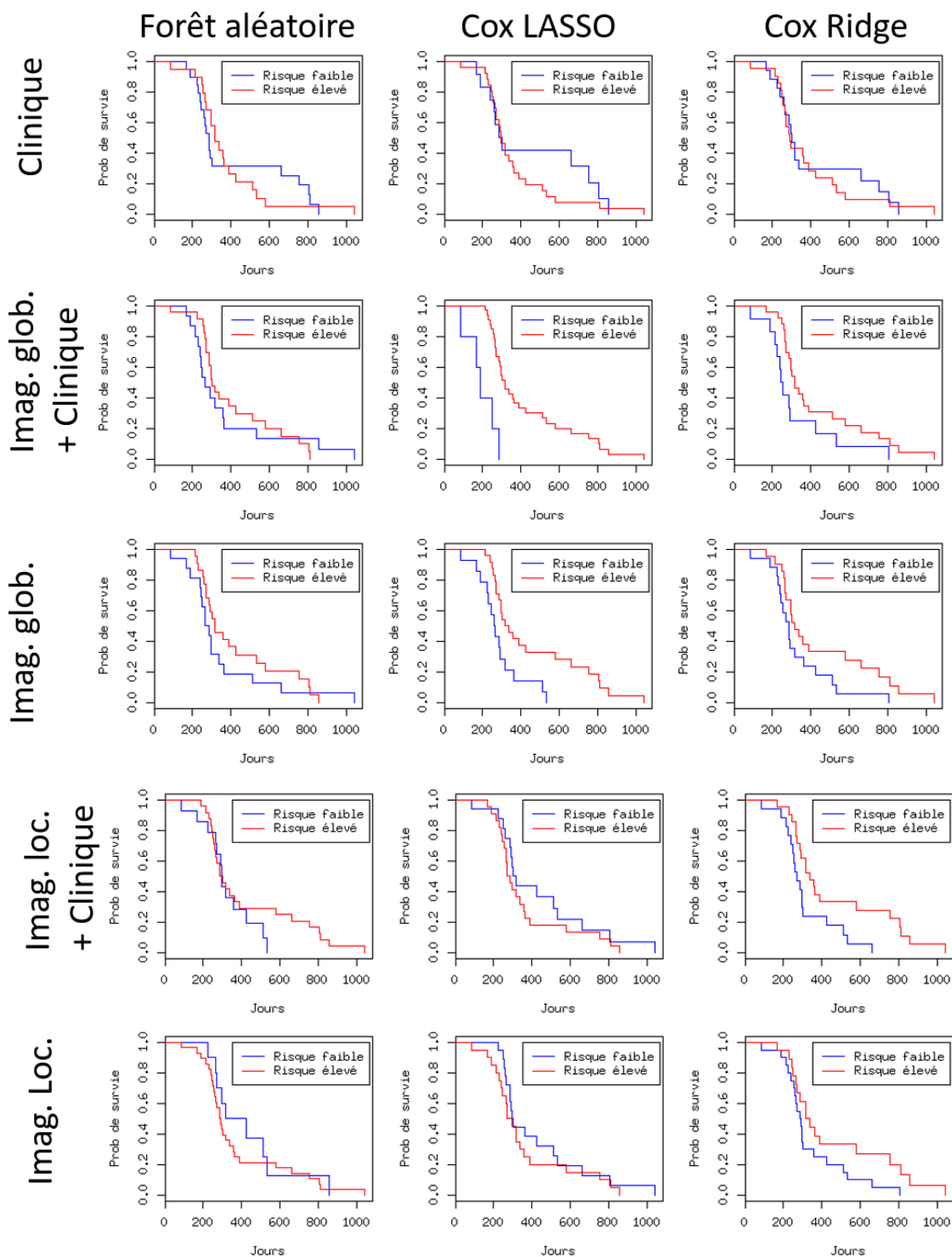


FIGURE 6.16 – Courbes Kaplan-Meier de validation suite à une classification du patient de validation de chaque pli de la LOOCV, le seuil choisi à chaque pli pour séparer les patients à risque faible et à risque élevé étant la valeur qui minimise la valeur-p du test du logrank sur les patients d’entraînement.

En réalité, les patients atteints du GTC sont considérés comme survivants à long terme s'ils sont vivants deux ans après le diagnostic (HOFFMAN et collab. [2018]). Dans la cohorte radiomique 4Mod, seulement 5 patients ont vécu plus de deux ans après le diagnostic, ce nombre a été considéré comme insuffisant pour caractériser le groupe.

Dans le futur, l'idéal serait d'analyser les survies des patients en fonction de leur mutations et des traitements reçus pour prédire l'efficacité de ces différents traitements. Pour cela un nombre plus important de sujets est nécessaire et la base de données BIOMEDE prise dans son intégralité devrait aider à la construction de tels modèles.

Conclusion

Les résultats de ce chapitre montrent que les modèles qui séparent le mieux les patients à risque faible et à risque élevé de décès sont les modèles avec deux ou trois indices. Même si les résultats restent très modérés en prédiction, ils semblent meilleurs que les résultats des études précédentes focalisées sur de faibles effectifs, n'utilisant que des indices radiologiques simples ou catégoriels. Cela suggère que certains indices radiomiques peuvent porter des informations pertinentes. Toutefois, le faible nombre de patients impacte les performances des prédicteurs de la survie globale. L'intérêt de modèles radiomiques utilisant plusieurs indices nécessite l'augmentation de la base de données pour que les modèles estimés soient robustes et que les tendances montrées dans ce chapitre puissent être validées.

6.4 Références

- CHEN, X. et H. ISHWARAN. 2013, «Pathway hunting by random survival forests», *Bioinformatics*, vol. 29, n° 1, p. 99–105. 181
- HARGRAVE, D., N. CHUANG et E. BOUFFET. 2008, «Conventional mri cannot predict survival in childhood diffuse intrinsic pontine glioma», *Journal of Neuro-Oncology*, vol. 86, n° 3, doi : 10.1007/s11060-007-9473-5, p. 313–319, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-007-9473-5>. 187
- HARWARD, S., S. HARRISON FARBER, M. MALINZAK, O. BECHER et E. M. THOMPSON. 2018, «T2-weighted images are superior to other mr image types for the determination of diffuse intrinsic pontine glioma intratumoral heterogeneity», *Child's Nervous System*, vol. 34, n° 3, doi : 10.1007/s00381-017-3659-8, p. 449–455, ISSN 1433-0350. URL <https://doi.org/10.1007/s00381-017-3659-8>. 187
- HOFFMAN, L., S. VELDHIJZEN VAN ZANTEN et collab.. 2018, «Clinical, radiologic, pathologic, and molecular characteristics of long-term survivors of diffuse intrinsic pontine glioma (DIPG) : A collaborative report from the international and european society for pediatric oncology dipg registries», *J Clin Oncol*, vol. 36, n° 19, p. 1963–1972. URL <https://www.ncbi.nlm.nih.gov/pubmed/29746225>. 189
- JANSEN, M. H., S. E. VELDHIJZEN VAN ZANTEN, E. SANCHEZ ALIAGA, M. W. HEYMANS, M. WARMUTH-METZ, D. HARGRAVE, E. J. VAN DER HOEVEN, C. E. GIDDING, E. S. DE BONT, O. S. ESHGHI, R. REDDINGIUS, C. M. PEETERS, A. Y. SCHOUTEN-VAN MEETEREN, R. H. GOOSKENS, B. GRANZEN, G. M. PAARDEKOOPER, G. O. JANSSENS, D. P. NOSKE, F. BARKHOF, C. M. KRAMM, W. P. VANDERTOP, G. J. KASPERS et D. G. VAN VUURDEN. 2014, «Survival prediction model of children with diffuse intrinsic pontine glioma based on clinical and radiological criteria», *Neuro-Oncology*, vol. 17, n° 1, doi :10.1093/neuonc/nou104, p. 160–166, ISSN 1522-8517. URL <https://doi.org/10.1093/neuonc/nou104>. 187

SIMON, R. M., J. SUBRAMANIAN, M. C. LI et S. MENEZES. 2011, «Using cross-validation to evaluate predictive accuracy of survival risk classifiers based on high-dimensional data», *Briefings in Bioinformatics*, vol. 12, n° 3, doi :10.1093/bib/bbr001, p. 203–214. URL <http://dx.doi.org/10.1093/bib/bbr001>. 181

VELDHUIJZEN VAN ZANTEN, S. E. M., A. LANE, M. W. HEYMANS, J. BAUGH, B. CHANEY, L. M. HOFFMAN, R. DOUGHMAN, M. H. A. JANSEN, E. SANCHEZ, W. P. VANDERTOP, G. J. L. KASPERS, D. G. VAN VUURDEN, M. FOULADI, B. V. JONES et J. LEACH. 2017, «External validation of the diffuse intrinsic pontine glioma survival prediction model : a collaborative report from the international dipg registry and the siopg registry», *Journal of Neuro-Oncology*, vol. 134, n° 1, doi : 10.1007/s11060-017-2514-9, p. 231–240, ISSN 1573-7373. URL <https://doi.org/10.1007/s11060-017-2514-9>. 187

Conclusion et perspectives

Pour traduire le fait que les images contiennent beaucoup plus d'informations que celles qui sont actuellement extraites et interprétées, le concept de radiomique a été introduit en 2012. La possibilité de prédire le phénotype et le pronostic de patients à partir des informations quantitatives de l'imagerie a attiré l'attention de chercheurs qui travaillent sur différentes pathologies. Ainsi, depuis cette date, le nombre de publications citant le terme radiomique a cru de façon exponentielle.

Le **GITC** est une tumeur récemment classifiée grâce aux découvertes des sous-types moléculaires associés à ces tumeurs pédiatriques. Par conséquent, des études sont actuellement conduites pour mieux décrire le fonctionnement de cette la maladie. Les patients sont désormais diagnostiqués par une combinaison de trois facteurs : les symptômes cliniques, l'**IRM** et l'analyse moléculaire. Dans ce contexte, les études radiomiques se sont révélées être très pertinentes pour aider à la compréhension de la maladie en explorant les informations présentes dans l'**IRM** de façon plus approfondie.

Ce travail de thèse représente, à notre connaissance, les premières études radiomiques réalisées sur des patients atteints de **GITC**. Les premières étapes de traitement de données d'imagerie et d'extraction des indices ont été réalisées pour rendre les données plus robustes avant la définition des modèles prédictifs. Malgré les promesses apportées par des approches de l'apprentissage profond, nous avons opté pour créer des modèles classiques en utilisant des caractéristiques extraites de façon classique pour faciliter l'interprétation des résultats. En outre, l'effectif est trop faible pour les approches d'apprentissage profond.

Le chapitre 3 a introduit ainsi une méthode de standardisation d'intensités des images adaptée à l'**IRM** multi-modale du cerveau pédiatrique. En effet, si la radiomique s'appuie principalement sur les indices calculés à partir de l'imagerie, il est indispensable d'avoir des images comparables. Nos résultats montrent l'efficacité de la méthode **hWS** sur plus de 1500 images de **GITC** ainsi que sur d'autres lésions cérébrales. Cette standardisation d'intensités a été appliquée systématiquement sur toutes les images multi-modales utilisées pour les analyses radiomiques. Le pipeline proposé a été automatisé afin de faciliter l'introduction de nouvelles sessions d'imagerie et de nouveaux patients dans les analyses radiomiques. En perspective, l'ajout d'une méthode comme **Combat** sur le pipeline de standardisation d'intensités des images des **GITC** acquises avec des champs magnétiques d'intensités différentes seraient intéressante au vue des résultats positifs obtenus sur la cohorte de lésions cérébrales (LC).

A l'issue des analyses du chapitre 4, nos recommandations sur les choix méthodologiques de calcul d'indices de texture sont les suivantes :

1. Discrétiser les niveaux de gris sur les images normalisées en regroupant les valeurs d'intensité dans un intervalle de largeur constante pour tous les patients
2. Utiliser un volume d'intérêt de taille et formes constantes et à défaut être vigilants quant aux volumes différents

Ces recommandations ont été appliquées pour le calcul des indices d'imagerie sur lesquels nous avons travaillé. Pour l'utilisation des volumes d'intérêt constants, une méthode d'extraction d'indices d'imagerie locaux a été introduite. Cependant, les résultats du chapitre 5 ne nous permettent pas de confirmer l'hypothèse selon laquelle parmi ces indices locaux certains refléteraient des informations importantes pour la prédiction des mutations génétiques associées au **GITC**, une

fois libérés de leur corrélation avec le volume.

La méthode de sélection d'indices a éliminé les indices fortement corrélés au volume par le critère de robustesse à la segmentation. C'est la raison pour laquelle, beaucoup plus d'indices globaux ont été éliminés en comparaison avec les indices locaux. Cependant, suite aux étapes ultérieures de la sélection, qui recherchent les indices pertinents pour la classification, tout en éliminant ceux qui sont trop corrélés, le nombre final d'indices globaux et locaux sélectionnés sont du même ordre de grandeur. Même si le nombre d'indices locaux pertinents est plus élevé que le nombre d'indices globaux pertinents, les performances des modèles de prédiction utilisant les indices globaux sont légèrement supérieures à celles des modèles de prédiction utilisant les indices locaux.

Les analyses de survie du chapitre 6 ont suggéré une meilleure correspondance entre les indices d'imagerie locaux et la survie globale. Ce point doit cependant être validé avec une base de données de taille plus conséquente, car les résultats ne sont pas statistiquement significatifs à présent. Cependant, avec la nouvelle prise en charge thérapeutique des patients atteints de **GITC** inclus dans les **PHRC BIOMEDE** et **BIOMEDE 2.0**, les analyses de survie devront intégrer les différentes thérapies comme co-facteurs dans les modèles.

Comme les bases de données utilisées pour les prédictions des mutations et de la survie globale sont à petits effectifs, les seuils pour sélectionner les indices pertinents sont stricts, pour réduire le nombre d'indices à inclure dans le modèle. Avec une base de données élargie, ces seuils peuvent être revus pour inclure d'avantage d'indices. Ainsi l'intérêt des indices d'imagerie locaux n'a pas été établi; des études supplémentaires avec une base de données plus large doivent être réalisées pour conclure sur ce sujet.

Certains indices accessibles avec PyRadiomics n'ont pas été exploités, comme les indices calculés sur les images transformées en ondelettes, ou sur les images de gradient. Ils n'ont pas été calculés car l'introduction d'une source de complexité et variabilité supplémentaire n'a pas été considérée comme prioritaire. En effet, les indices qui se montrent finalement les plus pertinents pour les études menées dans cette thèse sont majoritairement les indices simples de premier ordre.

Les indices de forme sont aussi fréquemment utilisés dans les études radiomiques et reflètent la géométrie des tumeurs. Ce sont aussi les indices de forme qui sont utilisés dans les critères d'analyse de progression des tumeurs comme le *Response Assessment in Neuro-Oncology (RANO)*. Ils n'ont pas été utilisés parce que les régions tumorales ont été approchées par des sphères pour que les régions étudiées soient indépendantes d'un opérateur expert délimitant les tumeurs. L'étude doit être poursuivie en incluant ces paramètres. Les analyses longitudinales doivent aussi être réalisées pour étudier la progression de la tumeur au cours du temps et la réponse au traitement des patients en fonction des différentes thérapies proposées.

Les résultats des modèles multi-variés pour la prédiction des variantes de la mutation H3 montrent que la combinaison des indices cliniques avec les indices d'imagerie sont plus performantes que les indices cliniques ou d'imagerie seuls. Néanmoins, les résultats obtenus pour la prédiction des mutations TP53 et ACVR1 sont moins encourageants, principalement pour la mutation du gène ACVR1. Les résultats de prédiction des variantes de la mutation H3 obtenus sur les données de test fournies pour des données extraites de BIOMEDE sont moins bien que ceux obtenus en validation. Ce résultat est principalement dû au faible nombre de sujets inclus. Ce petit nombre de sujets génère des difficultés pour établir le modèle entre les indices et la variable à prédire, qui soit efficace mais non-spécifique de la base de données utilisée pour l'entraînement. Dans les petites cohortes, la présence de valeurs aberrantes dans les données d'entraînement est également difficile à repérer et très préjudiciable. L'élargissement de la base de données est ainsi la priorité majeure pour que les études radiomiques puissent être poursuivies. Dans le cas d'une maladie rare comme le **GITC**, un effort doit être fait pour que chaque nouveau patient diagnostiqué puisse être inclus dans l'étude radiomique.

Les résultats du cas d'usage de la méthodologie multi-modèles proposée dans la deuxième partie du chapitre 5 sont très encourageants. Tous les patients de la cohorte PREBIOMEDE pré-

sentant la mutation des variantes H3.1 et H3.3 de l'histone H3 ont été inclus dans la conception du modèle même si toutes les modalités n'étaient pas présentes et une prédiction a été faite pour tous les patients de la cohorte BIOMEDE présentant la mutation des variantes H3.1 et H3.3 de l'histone H3, même ceux qui présentaient des données incomplètes. Les performances en test (indice de Youden égale à 0,45, score F1-pondéré valant 0,81) sont bonnes malgré les différents facteurs qui ne facilitent pas les conditions de test : la base de test BIOMEDE que nous avons utilisée présente des classes plus déséquilibrées (20% patients H3.1 et 80% patients H3.3) que la base d'entraînement PREBIOMEDE (32% patients H3.1 et 68% patients H3.3) et les deux bases ont été acquises de façon séparée dans le temps (PREBIOMEDE entre 2009 et 2014; BIOMEDE entre 2014 et 2019). Dans cette thèse, seul le cas d'usage sur la mutation H3 a été présenté mais la méthodologie multi-modèles peut être appliquée sur les autres études de prédiction. De plus, avec une augmentation de la base de données, l'ajout des patients H3-WT qui sont plus rares pourrait être considéré.

Enfin, comme le PHRC BIOMEDE est toujours en cours d'inclusion, le modèle d'analyse proposé dans cette thèse (de la standardisation d'intensités à la création des modèles de prédiction) permet l'inclusion rapide de nouveaux patients. Ainsi il sera possible d'améliorer les modèles de prédiction et d'accroître la robustesse des résultats progressivement.

Le tableau qui suit résume les messages principaux de conclusion et de perspectives.

Conclusions

1. La méthode **hWS** avec les modifications proposées a normalisé avec succès les intensités des images d'**IRM** des patients atteints de **GITC**.
2. Les indices de texture doivent être calculés sur les images normalisées avec une discrétisation des niveaux de gris présentant une largeur constante.
3. Pour la prédiction des variantes de la mutation de l'histone H3, les modèles combinant les indices cliniques avec les indices d'imagerie sont plus performants que les modèles utilisant les indices cliniques ou les indices d'imagerie seuls.
4. Les analyses de survie multi-variées nécessitent l'augmentation de la base de données pour espérer obtenir des résultats statistiquement significatifs.
5. La méthodologie multi-modèles a permis l'inclusion de patients à données d'imagerie manquantes et présente des résultats très encourageants.
6. Le pipeline de standardisation d'intensités et la création des modèles de prédiction permettent l'inclusion rapide de patients.

Enfin, ces premières études radiomiques réalisées au moment du diagnostic suggèrent que les indices radiomiques apportent des informations pour la prise en charge des patients atteints de **GITC mais l'augmentation de la base de données est nécessaire pour la confirmation des résultats. Pour cela, le pipeline d'analyse proposée dans cette thèse est automatisé et conçu de façon à pouvoir inclure facilement de nouveaux patients.**

Perspectives

1. Pour les images multi-centriques, l'harmonisation des différents indices radiomiques peut nécessiter une méthode d'harmonisation spécifique du type Combat, ce qui va se produire avec la prise en compte de toutes les données BIOMEDE.
2. L'hypothèse que les indices locaux reflètent des informations supplémentaires importantes reste à tester avec une base de données de taille plus importante.
3. Les patients H3-**WT** peuvent être inclus dans l'analyse si leur nombre est suffisamment important.
4. La méthodologie multi-modèle peut être appliquée pour les prédictions des mutations des gènes ACVR1 et TP53.
5. D'autres indices d'imagerie, comme les indices calculés sur les images transformées ondelettes, peuvent être inclus.
6. Les analyses longitudinales doivent être réalisées en fonction des différentes thérapies proposées et des examens d'imagerie disponibles pour définir l'évolution pour chaque patients.
7. Les tumeurs peuvent être segmentées pour le calcul et l'utilisation des indices de forme.
8. Les analyses radiomiques peuvent inclure des informations dérivées des images de perfusion **DSC** ou **ASL** et des images de diffusion (en cours dans le laboratoire IMIV).
9. La méthode d'harmonisation d'images d'**IRM** proposée peut être appliquée à d'autres études (en cours au laboratoire IVIM pour les cancers du sein et du poumon).

Liste de publications

Articles de revue

- **J. Goya-Outi**, F. Orlhac, R. Calmon, A. Alentorn, C. Nioche, C. Philippe, S. Puget, N. Boddaert, I. Buvat, J. Grill, V. Frouin, F. Frouin, Computation of reliable textural indices from multimodal brain MRI: suggestions based on a study of patients with diffuse intrinsic pontine glioma, *Physics in Medicine and Biology*, 63(10):105003 (2018)
- C. Nioche, F. Orlhac, S. Boughdad, S. Reuzé, **J. Goya-Outi**, C. Robert, C. Pellot-Barakat, M. Soussan, F. Frouin, I. Buvat, LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity, *Cancer Research*, 78(16):4786-4789 (2018)

Article de conférence

- **J. Goya Outi**, R. Calmon, F. Orlhac, C. Philippe, N. Boddaert, S. Puget, V. Frouin, I. Buvat, J. Grill, F. Frouin, Can Structural MRI Radiomics Predict DIPG Histone H3 Mutation and Patient Overall Survival at Diagnosis Time?, *IEEE-EMBS International Conference on Biomedical and Health Informatics*, 19-22 mai 2019, Chicago, USA

Résumés de conférence

- **J. Goya Outi**, F. Orlhac, R. Calmon, C. Nioche, C. Philippe, A. Alentorn, J. Grill, V. Frouin, F. Frouin, Impact of volume size on radiomic features computed from MR images: an illustration in pediatric neuroimaging, *European Society for Magnetic Resonance in Medicine and Biology annual meeting*, 19-21 octobre 2017, Barcelone, Spain
- **J. Goya Outi**, R. Calmon, A. Alentorn, C. Philippe, F. Orlhac, C. Nioche, J. Grill, V. Frouin, F. Frouin, A pipeline to standardize intensities of MR images for reliable longitudinal studies in neuro-oncology, *European Society for Magnetic Resonance in Medicine and Biology annual meeting*, 19-21 octobre 2017, Barcelone, Spain
- **J. Goya-Outi**, F. Orlhac, R. Calmon, C. Philippe, S. Puget, N. Boddaert, I. Buvat, J. Grill, V. Frouin, F. Frouin, Volume-independent radiomic features from T2w-FLAIR MRI could reveal mutation of histones in diffuse intrinsic pontine glioma, *International Society for Magnetic Resonance in Medicine annual meeting*, 16-21 juin 2018, Paris, France
- **J. Goya-Outi**, F. Orlhac, R. Calmon, V. Frouin, C. Philippe, N. Boddaert, S. Puget, J. Grill, I. Buvat, F. Frouin, Can MRI radiomics predict the type of histone H3 mutations in diffuse intrinsic pontine glioma?, *International Cancer Imaging Society meeting*, 8-9 octobre 2018, Menton, France
- **J. Goya Outi**, V. Frouin, N. Boddaert, J. Grill, F. Frouin, Modèles radiomiques adaptés aux données manquantes en imagerie multi-modale, *Recherche en Imagerie et Technologies pour la Santé*, 20-22 mai 2019, Tours, France

- **J. Goya Outi**, A. Pister, H. Urien, C. Philippe, A. Grigis, R. Calmon, N. Boddaert, J. Grill, V. Frouin, E. Frouin, A pipeline and its associated quality control to standardize structural MR images in neuro-oncology: a pre-requisite for radiomic studies, Congrès National d'Imagerie du Vivant, 5-6 février 2019, Paris, France

Annexe A

Annexes

A.1 Tableaux annexes

TABLEAU A.1 – Les patients avec GITC de la cohorte PREBIOMEDE. Les patients de la cohorte étoile, dont les sessions immédiatement avant et après biopsie ont été utilisées comme référence dans la méthode de standardisation d'intensités (dans le chapitre 3), sont en gras. Les patients de la cohorte radiomique 4Mod sont soulignés en rouge.

Sujet	Sexe ¹	Âge	ACVR1 ²	TP53 ²	H3 ²	Survie Globale (jours)	Status ³	Nombre de sessions d'imagerie
sub – 1102	0	6,9	1	0	1	480	1	6
sub – 1139	0	9,7	0	1	3	362	1	4
sub – 1454	0	9,3	0	1	0	>52	0	2
sub – 1535	0	2,2	–	–	3	1228	1	16
sub – 1644	0	5,8	0	1	1	755	1	18
sub – 1756	1	3,3	1	0	1	804	1	12
sub – 1829	0	5,5	–	–	–	255	1	6
sub – 1872	1	10,2	0	1	3	314	1	5
sub – 1952	1	9,3	0	1	3	663	1	4
sub – 2019	0	11,5	0	1	3	86	1	3
sub – 2053	1	14,2	1	0	0	858	1	18
sub – 2452	0	5,2	1	1	1	272	1	1
sub – 3037	1	7,5	1	0	1	1043	1	30
sub – 3151	0	6,6	0	0	3	811	1	2
sub – 3182	0	7,8	0	1	3	212	1	5
sub – 3593	1	5,7	0	0	3	318	1	6
sub – 3712	0	10,6	0	1	3	249	1	1
sub – 4023	0	6,2	0	1	3	295	1	5
sub – 4385	0	7,9	0	1	3	302	1	7
sub – 4437	0	4,8	–	0	–	285	1	5
sub – 4596	1	11,8	0	1	3	>390	0	4
sub – 4722	0	5,7	0	1	0	240	1	5
sub – 5039	1	7,1	0	1	3	243	1	5
sub – 5196	1	6	–	–	–	321	1	5
sub – 5272	1	5,9	0	–	3	298	1	7
sub – 5328	0	5,7	0	1	3	131	1	4
sub – 5460	0	4,4	1	0	1	513	1	7
sub – 5503	0	4,1	–	–	–	518	1	2
sub – 5506	0	5,7	1	0	1	446	1	6
sub – 5951	1	9,4	0	1	3	231	1	5
sub – 6046	1	3,7	0	0	1	260	1	2
sub – 6121	0	4,6	1	0	1	287	1	8
sub – 6194	1	14,7	0	1	3	424	1	9
sub – 6450	1	7,9	–	–	–	224	1	2
sub – 7450	1	14	0	1	3	291	1	6
sub – 7494	1	6,0	–	–	–	337	1	5
sub – 7905	1	1,4	–	–	1	452	1	2
sub – 7931	0	15,2	0	0	3	580	1	9
sub – 8044	0	5,7	0	1	3	390	1	6
sub – 8818	0	3,9	0	0	3	387	1	9
sub – 9165	0	4,1	0	0	3	265	1	6
sub – 9331	0	6,7	–	–	3	188	1	5
sub – 9564	0	5,2	1	0	1	531	1	3
sub – 9857	1	4,4	0	1	0	267	1	6
sub – 9944	1	4,5	0	1	2	272	1	4
sub – 9987	0	5,1	0	1	3	169	1	5
sub – 9999	0	4,7	0	0	3	358	1	2
sub – 8888	1	8,6	0	0	1	595	1	1

¹ Fille = 0; Garçon = 1

² Non-muté = 0; muté = 1; – = pas de biopsie ou résultat pas concluant

³ censuré = 0; décès = 1

TABLEAU A.2 – Les patients avec GITC de la cohorte BIOMEDE.

Sujet	Sexe ¹	Âge	ACVR1 ²	TP53 ²	H3 ²	Survie Globale (jours)	Status ³	Nombre de sessions d'imagerie
sub – 006	1	5,8	0	1	3	270	1	7
sub – 009	0	5,5	0	0	3	107	1	4
sub – 016	0	3,5	1	0	1	290	1	7
sub – 028	0	4,1	1	0	1	543	1	3
sub – 037	1	6,3	–	–	–	>1183	0	1
sub – 038	1	9,1	0	1	3	295	1	6
sub – 040	0	6,5	0	0	3	387	1	6
sub – 047	1	9,3	0	1	3	364	1	5
sub – 051	1	6,9	0	1	3	134	1	3
sub – 058	0	12,8	0	1	3	786	1	6
sub – 071	0	7,0	0	1	3	161	1	6
sub – 073	1	10,1	1	–	–	21	1	1
sub – 080	1	7,1	1	–	–	154	1	1
sub – 082	1	10,4	0	1	3	>804	0	5
sub – 084	1	6,7	0	1	3	376	1	6
sub – 086	0	11,4	0	1	3	451	1	6
sub – 092	1	10,2	0	1	3	738	0	8
sub – 095	0	5,7	1	0	1	302	1	6
sub – 100	1	7,2	0	0	3	>681	0	6
sub – 102	1	9,6	0	1	3	582	1	4
sub – 104	1	16,2	0	0	3	239	1	6
sub – 108	0	4,7	–	–	–	547	1	6
sub – 110	1	7,1	0	1	3	171	1	5
sub – 114	1	15,4	0	1	3	187	1	7
sub – 115	1	3,4	1	0	1	445	1	6
sub – 124	1	19,0	0	1	3	428	1	3
sub – 135	1	4,7	1	0	1	488	1	6
sub – 139	0	3,3	0	0	0	469	1	6
sub – 143	1	6,8	0	0	3	>467	0	3
sub – 149	1	6,0	0	0	3	297	1	6
sub – 160	1	16,5	–	–	–	>341	0	4
sub – 162	1	14,2	–	–	–	>348	0	4
sub – 163	1	30,2	–	–	–	200	1	5
sub – 166	1	9,7	–	–	–	213	1	4
sub – 178	1	4,9	–	–	–	139	1	3
sub – 190	1	7,4	–	–	–	>214	0	2
sub – 218	0	5,6	–	–	–	>49	0	2
sub – 219	0	11	–	–	–	>49	0	1

¹ Fille = 0; Garçon = 1² Non-muté = 0; muté = 1; – = pas de biopsie ou résultat pas concluant ou résultat non encore connu³ censuré = 0; décès = 1

TABLEAU A.3 – Les patients de lésions cérébrales de la cohorte LC. Ces patients ont été imaged dans une IRM à 1,5T puis dans une IRM à 3T dans un intervalle de temps de 2 à 80 jours sans avoir eu de traitement, de chirurgie et sans évolution visible des lésions.

Sujet	T1c ¹	FLAIR ¹	Nombre de lésions
Patient 01	1	1	2
Patient 02	1	1	1
Patient 03	1	1	1
Patient 05	1	1	3
Patient 06	1	1	1
Patient 09	0	1	2
Patient 10	1	1	1
Patient 11	1	1	1
Patient 12	1	1	1
Patient 14	1	1	2
Patient 15	1	0	1
Patient 16	1	1	1
Patient 18	1	1	1
Patient 19	1	1	1
Patient 20	0	1	1
Patient 21	0	1	1
Patient 22	0	1	1
Patient 23	1	1	1
Patient 81	1	1	1
Patient 82	0	1	1

¹ Non = 0; Oui = 1

TABLEAU A.4 – AUC des indices d'imagerie pour prédire la mutation H3.1 ou H3.3. Les valeurs d'AUC entre 0,75 et 0,869 sont affichées en bleu. Les valeurs d'AUC supérieures ou égales à 0,869 sont en vert.

	T1w		T2w		T1c		FLAIR	
	global	local	global	local	global	local	global	local
firstorder.InterquartileRange	0,574	0,551	0,676	0,597	0,716	0,722	0,676	0,705
firstorder.Skewness	0,727	0,744	0,523	0,557	0,670	0,733	0,625	0,659
firstorder.Uniformity	0,614	0,511	0,699	0,602	0,705	0,693	0,688	0,727
firstorder.MeanAbsoluteDeviation	0,619	0,562	0,659	0,591	0,716	0,722	0,670	0,716
firstorder.Energy	0,676	0,852	0,682	0,545	0,676	0,602	0,682	0,722
firstorder.RobustMeanAbsoluteDeviation	0,580	0,557	0,665	0,591	0,733	0,727	0,670	0,705
firstorder.Median	0,881	0,852	0,580	0,551	0,733	0,670	0,710	0,733
firstorder.TotalEnergy	0,676	0,852	0,682	0,545	0,676	0,602	0,682	0,722
firstorder.Maximum	0,619	0,869	0,517	0,523	0,727	0,568	0,716	0,716
firstorder.RootMeanSquared	0,875	0,852	0,568	0,545	0,642	0,602	0,710	0,722
firstorder.90Percentile	0,818	0,869	0,517	0,506	0,534	0,506	0,739	0,739
firstorder.Minimum	0,824	0,830	0,716	0,636	0,818	0,886	0,648	0,511
firstorder.Entropy	0,614	0,517	0,699	0,614	0,710	0,710	0,699	0,722
firstorder.StandardDeviation	0,642	0,568	0,659	0,608	0,716	0,716	0,693	0,722
firstorder.Range	0,659	0,574	0,659	0,608	0,784	0,722	0,710	0,716
firstorder.Variance	0,642	0,580	0,659	0,614	0,716	0,716	0,693	0,722
firstorder.10Percentile	0,847	0,835	0,659	0,619	0,864	0,875	0,608	0,614
firstorder.Kurtosis	0,636	0,693	0,585	0,642	0,625	0,540	0,619	0,619
firstorder.Mean	0,875	0,852	0,568	0,545	0,642	0,602	0,710	0,722
glcm.SumVariance	0,534	0,500	0,699	0,625	0,653	0,688	0,716	0,761
glcm.Homogeneity1	0,523	0,534	0,562	0,580	0,727	0,727	0,778	0,767
glcm.Homogeneity2	0,523	0,545	0,557	0,574	0,722	0,710	0,778	0,761
glcm.ClusterShade	0,756	0,790	0,545	0,523	0,795	0,807	0,722	0,739
glcm.MaximumProbability	0,545	0,545	0,665	0,557	0,693	0,693	0,727	0,739
glcm.Idmn	0,636	0,614	0,625	0,540	0,511	0,574	0,534	0,506
glcm.SumVariance2	0,619	0,551	0,665	0,602	0,710	0,710	0,688	0,699
glcm.Contrast	0,585	0,517	0,619	0,619	0,739	0,739	0,756	0,756
glcm.DifferenceEntropy	0,574	0,511	0,614	0,608	0,727	0,739	0,767	0,761
glcm.InverseVariance	0,528	0,528	0,557	0,591	0,722	0,716	0,767	0,761
glcm.Entropy	0,591	0,506	0,676	0,602	0,722	0,710	0,739	0,744
glcm.Dissimilarity	0,562	0,506	0,580	0,608	0,733	0,739	0,767	0,767
glcm.DifferenceVariance	0,614	0,523	0,659	0,636	0,733	0,739	0,750	0,756
glcm.Idn	0,653	0,625	0,636	0,648	0,591	0,545	0,557	0,528
glcm.Idm	0,523	0,545	0,557	0,574	0,722	0,710	0,778	0,761
glcm.Correlation	0,631	0,585	0,756	0,659	0,528	0,545	0,523	0,585
glcm.Autocorrelation	0,528	0,506	0,699	0,625	0,642	0,688	0,716	0,767
glcm.SumEntropy	0,619	0,528	0,710	0,619	0,705	0,705	0,710	0,716
glcm.AverageIntensity	0,517	0,523	0,693	0,642	0,642	0,688	0,722	0,773
glcm.Energy	0,557	0,517	0,665	0,580	0,705	0,699	0,727	0,744
glcm.SumSquares	0,642	0,591	0,659	0,608	0,716	0,716	0,688	0,716
glcm.ClusterProminence	0,693	0,597	0,653	0,591	0,727	0,727	0,699	0,699
glcm.SumAverage	0,523	0,523	0,688	0,636	0,642	0,682	0,722	0,773
glcm.Imc2	0,631	0,545	0,750	0,636	0,665	0,659	0,540	0,682
glcm.Imc1	0,597	0,580	0,710	0,619	0,557	0,670	0,540	0,659
glcm.DifferenceAverage	0,562	0,506	0,580	0,608	0,733	0,739	0,767	0,767
glcm.Id	0,523	0,534	0,562	0,580	0,727	0,727	0,778	0,767
glcm.ClusterTendency	0,619	0,551	0,665	0,602	0,710	0,710	0,688	0,699
grlm.ShortRunLowGrayLevelEmphasis	0,506	0,568	0,670	0,716	0,568	0,602	0,710	0,795
grlm.GrayLevelVariance	0,653	0,580	0,665	0,614	0,716	0,716	0,693	0,716
grlm.LowGrayLevelRunEmphasis	0,517	0,574	0,665	0,716	0,568	0,602	0,705	0,801
grlm.GrayLevelNonUniformityNormalized	0,619	0,511	0,693	0,602	0,710	0,688	0,693	0,727
grlm.RunVariance	0,523	0,568	0,534	0,545	0,705	0,705	0,767	0,767
grlm.GrayLevelNonUniformity	0,574	0,511	0,545	0,608	0,540	0,688	0,500	0,716
grlm.LongRunEmphasis	0,506	0,568	0,523	0,545	0,710	0,710	0,761	0,761
grlm.ShortRunHighGrayLevelEmphasis	0,562	0,511	0,693	0,625	0,653	0,699	0,722	0,756
grlm.RunLengthNonUniformity	0,682	0,568	0,705	0,551	0,722	0,705	0,688	0,767
grlm.ShortRunEmphasis	0,500	0,557	0,523	0,551	0,710	0,705	0,756	0,767
grlm.LongRunHighGrayLevelEmphasis	0,534	0,511	0,699	0,642	0,636	0,665	0,710	0,773
grlm.RunPercentage	0,500	0,568	0,523	0,551	0,710	0,705	0,756	0,761
grlm.LongRunLowGrayLevelEmphasis	0,534	0,597	0,659	0,699	0,574	0,614	0,716	0,812
grlm.RunEntropy	0,648	0,528	0,722	0,625	0,705	0,710	0,676	0,705
grlm.HighGrayLevelRunEmphasis	0,545	0,506	0,699	0,625	0,653	0,693	0,716	0,756
grlm.RunLengthNonUniformityNormalized	0,500	0,562	0,528	0,551	0,710	0,705	0,756	0,767
glszm.GrayLevelVariance	0,727	0,602	0,653	0,614	0,739	0,722	0,699	0,722
glszm.LowIntensityLargeAreaEmphasis	0,574	0,602	0,568	0,591	0,636	0,648	0,733	0,778
glszm.HighIntensitySmallAreaEmphasis	0,653	0,545	0,710	0,642	0,705	0,710	0,705	0,756
glszm.SmallAreaEmphasis	0,602	0,500	0,585	0,551	0,727	0,767	0,517	0,733
glszm.LargeAreaEmphasis	0,523	0,545	0,506	0,551	0,670	0,705	0,693	0,756
glszm.ZoneVariance	0,523	0,545	0,506	0,545	0,670	0,699	0,693	0,744
glszm.SizeZoneVariabilityNormalized	0,602	0,500	0,580	0,568	0,733	0,773	0,517	0,739
glszm.LowIntensitySmallAreaEmphasis	0,614	0,500	0,699	0,659	0,642	0,653	0,716	0,773
glszm.HighIntensityVariabilityNormalized	0,648	0,551	0,693	0,636	0,688	0,716	0,722	0,739
glszm.IntensityVariabilityNormalized	0,693	0,517	0,665	0,614	0,722	0,699	0,693	0,722
glszm.ZonePercentage	0,534	0,528	0,545	0,580	0,716	0,727	0,761	0,767
glszm.LowIntensityEmphasis	0,619	0,523	0,670	0,699	0,648	0,676	0,716	0,784
glszm.SizeZoneVariability	0,665	0,506	0,699	0,591	0,773	0,744	0,807	0,761
glszm.IntensityVariability	0,574	0,597	0,614	0,619	0,665	0,631	0,727	0,665
glszm.ZoneEntropy	0,693	0,517	0,750	0,597	0,761	0,710	0,733	0,705
glszm.HighIntensityLargeAreaEmphasis	0,506	0,545	0,602	0,659	0,739	0,682	0,534	0,591

TABLEAU A.5 – AUC des indices d'imagerie prédire la présence de mutation du gène TP53. Les valeurs d'AUC entre 0,75 et 0,85 sont affichées en bleu. Les valeurs d'AUC supérieures ou égales à 0,85 sont en vert.

	T1w		T2w		T1c		FLAIR	
	global	local	global	local	global	local	global	local
firstorder.InterquartileRange	0,615	0,600	0,727	0,654	0,642	0,631	0,635	0,708
firstorder.Skewness	0,515	0,554	0,535	0,627	0,562	0,654	0,650	0,585
firstorder.Uniformity	0,619	0,631	0,742	0,673	0,635	0,619	0,642	0,715
firstorder.MeanAbsoluteDeviation	0,608	0,612	0,719	0,658	0,623	0,631	0,646	0,715
firstorder.Energy	0,692	0,719	0,715	0,619	0,696	0,692	0,712	0,785
firstorder.RobustMeanAbsoluteDeviation	0,608	0,612	0,727	0,650	0,646	0,635	0,638	0,708
firstorder.Median	0,704	0,719	0,612	0,623	0,765	0,727	0,800	0,788
firstorder.TotalEnergy	0,692	0,719	0,715	0,619	0,696	0,692	0,712	0,785
firstorder.Maximum	0,538	0,700	0,588	0,612	0,596	0,523	0,727	0,758
firstorder.RootMeanSquared	0,708	0,719	0,631	0,619	0,692	0,692	0,796	0,785
firstorder.90Percentile	0,754	0,723	0,631	0,627	0,527	0,577	0,769	0,777
firstorder.Minimum	0,692	0,700	0,835	0,546	0,692	0,735	0,719	0,588
firstorder.Entropy	0,627	0,608	0,738	0,685	0,627	0,635	0,658	0,712
firstorder.StandardDeviation	0,600	0,615	0,712	0,665	0,612	0,631	0,654	0,715
firstorder.Range	0,631	0,615	0,750	0,654	0,635	0,631	0,750	0,708
firstorder.Variance	0,600	0,608	0,712	0,665	0,612	0,623	0,654	0,719
firstorder.10Percentile	0,700	0,688	0,550	0,608	0,777	0,750	0,735	0,769
firstorder.Kurtosis	0,542	0,600	0,512	0,592	0,527	0,577	0,677	0,581
firstorder.Mean	0,708	0,719	0,631	0,619	0,692	0,688	0,796	0,785
glcm.SumVariance	0,504	0,554	0,835	0,700	0,558	0,600	0,808	0,746
glcm.Homogeneity1	0,627	0,631	0,631	0,646	0,665	0,673	0,735	0,754
glcm.Homogeneity2	0,627	0,631	0,619	0,638	0,669	0,662	0,738	0,754
glcm.ClusterShade	0,585	0,627	0,538	0,538	0,654	0,665	0,642	0,604
glcm.MaximumProbability	0,658	0,650	0,735	0,654	0,669	0,642	0,669	0,723
glcm.Idmn	0,523	0,554	0,665	0,558	0,562	0,638	0,592	0,554
glcm.SumVariance2	0,573	0,588	0,719	0,650	0,604	0,612	0,658	0,708
glcm.Contrast	0,619	0,638	0,681	0,669	0,646	0,650	0,746	0,758
glcm.DifferenceEntropy	0,627	0,627	0,673	0,658	0,658	0,677	0,742	0,758
glcm.InverseVariance	0,623	0,635	0,612	0,654	0,665	0,669	0,731	0,754
glcm.Entropy	0,642	0,638	0,738	0,669	0,650	0,635	0,712	0,723
glcm.Dissimilarity	0,631	0,627	0,646	0,662	0,662	0,669	0,735	0,758
glcm.DifferenceVariance	0,615	0,631	0,696	0,677	0,627	0,650	0,765	0,765
glcm.Idn	0,504	0,558	0,700	0,550	0,565	0,608	0,604	0,515
glcm.Idm	0,627	0,631	0,619	0,638	0,669	0,662	0,738	0,754
glcm.Correlation	0,504	0,519	0,681	0,608	0,650	0,623	0,523	0,550
glcm.Autocorrelation	0,504	0,550	0,838	0,704	0,542	0,592	0,804	0,742
glcm.SumEntropy	0,615	0,612	0,750	0,673	0,612	0,623	0,665	0,715
glcm.AverageIntensity	0,504	0,558	0,842	0,708	0,538	0,588	0,812	0,750
glcm.Energy	0,623	0,654	0,727	0,650	0,635	0,635	0,696	0,723
glcm.SumSquares	0,592	0,612	0,708	0,665	0,608	0,619	0,650	0,715
glcm.ClusterProminence	0,562	0,569	0,688	0,654	0,588	0,608	0,685	0,704
glcm.SumAverage	0,508	0,554	0,835	0,704	0,538	0,581	0,812	0,750
glcm.lmc2	0,512	0,573	0,715	0,685	0,550	0,600	0,504	0,704
glcm.lmc1	0,531	0,519	0,654	0,688	0,712	0,577	0,573	0,685
glcm.DifferenceAverage	0,631	0,627	0,646	0,662	0,662	0,669	0,735	0,758
glcm.Id	0,627	0,631	0,631	0,646	0,665	0,673	0,735	0,754
glcm.ClusterTendency	0,573	0,588	0,719	0,650	0,604	0,612	0,658	0,708
gtrlm.ShortRunLowGrayLevelEmphasis	0,538	0,535	0,865	0,754	0,512	0,535	0,827	0,758
gtrlm.GrayLevelVariance	0,604	0,608	0,708	0,665	0,608	0,623	0,658	0,719
gtrlm.LowGrayLevelRunEmphasis	0,538	0,538	0,862	0,750	0,500	0,538	0,819	0,762
gtrlm.GrayLevelNonUniformityNormalized	0,627	0,623	0,738	0,673	0,650	0,619	0,646	0,715
gtrlm.RunVariance	0,642	0,646	0,585	0,627	0,654	0,658	0,723	0,738
gtrlm.GrayLevelNonUniformity	0,669	0,615	0,512	0,677	0,535	0,615	0,577	0,712
gtrlm.LongRunEmphasis	0,635	0,642	0,596	0,623	0,650	0,654	0,719	0,738
gtrlm.ShortRunHighGrayLevelEmphasis	0,546	0,562	0,835	0,688	0,562	0,608	0,808	0,738
gtrlm.RunLengthNonUniformity	0,692	0,635	0,719	0,631	0,715	0,658	0,738	0,762
gtrlm.ShortRunEmphasis	0,627	0,635	0,600	0,631	0,654	0,658	0,727	0,762
gtrlm.LongRunHighGrayLevelEmphasis	0,519	0,515	0,846	0,712	0,531	0,581	0,808	0,750
gtrlm.RunPercentage	0,623	0,638	0,596	0,627	0,650	0,654	0,719	0,750
gtrlm.LongRunLowGrayLevelEmphasis	0,515	0,550	0,854	0,738	0,515	0,565	0,827	0,762
gtrlm.RunEntropy	0,604	0,600	0,769	0,681	0,623	0,627	0,642	0,712
gtrlm.HighGrayLevelRunEmphasis	0,527	0,558	0,835	0,696	0,562	0,604	0,808	0,738
gtrlm.RunLengthNonUniformityNormalized	0,627	0,635	0,604	0,631	0,654	0,658	0,727	0,762
glszm.GrayLevelVariance	0,612	0,615	0,704	0,665	0,604	0,623	0,688	0,712
glszm.LowIntensityLargeAreaEmphasis	0,508	0,596	0,685	0,658	0,581	0,608	0,735	0,754
glszm.HighIntensitySmallAreaEmphasis	0,550	0,581	0,827	0,692	0,592	0,623	0,808	0,735
glszm.SmallAreaEmphasis	0,542	0,581	0,585	0,619	0,592	0,665	0,558	0,708
glszm.LargeAreaEmphasis	0,581	0,658	0,569	0,631	0,615	0,654	0,654	0,746
glszm.ZoneVariance	0,577	0,654	0,565	0,646	0,615	0,658	0,654	0,742
glszm.SizeZoneVariabilityNormalized	0,538	0,581	0,585	0,619	0,596	0,669	0,554	0,708
glszm.LowIntensitySmallAreaEmphasis	0,558	0,581	0,850	0,712	0,585	0,608	0,858	0,762
glszm.HighIntensityEmphasis	0,550	0,573	0,835	0,696	0,569	0,627	0,808	0,727
glszm.IntensityVariabilityNormalized	0,635	0,615	0,735	0,662	0,635	0,627	0,665	0,712
glszm.ZonePercentage	0,615	0,650	0,608	0,642	0,646	0,658	0,731	0,758
glszm.LowIntensityEmphasis	0,542	0,577	0,854	0,727	0,577	0,596	0,838	0,762
glszm.SizeZoneVariability	0,715	0,650	0,758	0,650	0,727	0,673	0,869	0,765
glszm.IntensityVariability	0,712	0,604	0,669	0,623	0,742	0,627	0,777	0,735
glszm.ZoneEntropy	0,669	0,615	0,796	0,638	0,627	0,627	0,712	0,700
glszm.HighIntensityLargeAreaEmphasis	0,612	0,673	0,662	0,565	0,681	0,792	0,515	0,665

TABLEAU A.6 – AUC des indices d'imagerie pour prédire la présence de mutation du gène ACVR1. Les valeurs d'AUC entre 0,75 et 0,85 sont affichées en bleu. Les valeurs d'AUC supérieures ou égales à 0,85 sont en vert.

	T1w		T2w		T1c		FLAIR	
	global	local	global	local	global	local	global	local
firstorder.InterquartileRange	0,599	0,522	0,549	0,538	0,571	0,560	0,599	0,610
firstorder.Skewness	0,764	0,720	0,670	0,626	0,604	0,643	0,797	0,725
firstorder.Uniformity	0,615	0,527	0,571	0,511	0,571	0,560	0,593	0,626
firstorder.MeanAbsoluteDeviation	0,626	0,516	0,533	0,538	0,555	0,566	0,610	0,610
firstorder.Energy	0,654	0,824	0,665	0,516	0,654	0,725	0,670	0,709
firstorder.RobustMeanAbsoluteDeviation	0,604	0,511	0,533	0,544	0,588	0,571	0,599	0,610
firstorder.Median	0,852	0,819	0,533	0,516	0,753	0,720	0,742	0,720
firstorder.TotalEnergy	0,654	0,824	0,665	0,516	0,654	0,725	0,670	0,709
firstorder.Maximum	0,626	0,846	0,610	0,577	0,549	0,593	0,637	0,665
firstorder.RootMeanSquared	0,846	0,824	0,505	0,516	0,742	0,725	0,720	0,709
firstorder.90Percentile	0,868	0,857	0,560	0,588	0,654	0,681	0,692	0,692
firstorder.Minimum	0,769	0,791	0,588	0,516	0,698	0,753	0,566	0,566
firstorder.Entropy	0,615	0,527	0,566	0,505	0,571	0,566	0,599	0,621
firstorder.StandardDeviation	0,637	0,527	0,522	0,522	0,560	0,566	0,610	0,615
firstorder.Range	0,615	0,549	0,538	0,511	0,610	0,571	0,621	0,604
firstorder.Variance	0,637	0,538	0,522	0,511	0,560	0,555	0,610	0,615
firstorder.10Percentile	0,819	0,802	0,522	0,516	0,753	0,753	0,599	0,659
firstorder.Kurtosis	0,665	0,626	0,544	0,538	0,593	0,560	0,654	0,538
firstorder.Mean	0,846	0,824	0,505	0,516	0,742	0,720	0,720	0,709
glcm.SumVariance	0,566	0,560	0,588	0,522	0,511	0,538	0,648	0,654
glcm.Homogeneity1	0,560	0,577	0,533	0,522	0,588	0,588	0,698	0,687
glcm.Homogeneity2	0,560	0,577	0,533	0,527	0,582	0,571	0,709	0,687
glcm.ClusterShade	0,780	0,824	0,681	0,599	0,648	0,648	0,819	0,808
glcm.MaximumProbability	0,527	0,538	0,588	0,522	0,588	0,577	0,615	0,654
glcm.Idmn	0,687	0,648	0,610	0,615	0,604	0,533	0,549	0,599
glcm.SumVariance2	0,637	0,538	0,522	0,533	0,555	0,549	0,599	0,599
glcm.Contrast	0,522	0,544	0,505	0,511	0,588	0,577	0,654	0,659
glcm.DifferenceEntropy	0,533	0,588	0,516	0,505	0,588	0,599	0,670	0,670
glcm.InverseVariance	0,560	0,571	0,538	0,511	0,582	0,577	0,692	0,687
glcm.Entropy	0,560	0,544	0,571	0,511	0,582	0,577	0,643	0,637
glcm.Dissimilarity	0,538	0,577	0,511	0,500	0,593	0,588	0,670	0,676
glcm.DifferenceVariance	0,505	0,549	0,538	0,527	0,566	0,577	0,654	0,659
glcm.Idn	0,692	0,637	0,582	0,582	0,626	0,538	0,566	0,604
glcm.Idm	0,560	0,577	0,533	0,527	0,582	0,571	0,709	0,687
glcm.Correlation	0,670	0,593	0,687	0,500	0,505	0,544	0,544	0,511
glcm.Autocorrelation	0,571	0,560	0,588	0,522	0,522	0,538	0,648	0,659
glcm.SumEntropy	0,643	0,511	0,577	0,500	0,560	0,560	0,588	0,615
glcm.AverageIntensity	0,582	0,582	0,582	0,538	0,527	0,538	0,654	0,670
glcm.Energy	0,511	0,527	0,549	0,522	0,566	0,577	0,637	0,654
glcm.SumSquares	0,637	0,544	0,522	0,516	0,560	0,555	0,604	0,615
glcm.ClusterProminence	0,670	0,582	0,538	0,522	0,560	0,560	0,626	0,604
glcm.SumAverage	0,582	0,582	0,582	0,533	0,527	0,533	0,654	0,670
glcm.Imc2	0,670	0,527	0,621	0,516	0,533	0,522	0,560	0,571
glcm.Imc1	0,637	0,566	0,610	0,505	0,544	0,522	0,626	0,549
glcm.DifferenceAverage	0,538	0,577	0,511	0,500	0,593	0,588	0,670	0,676
glcm.Id	0,560	0,577	0,533	0,522	0,588	0,588	0,698	0,687
glcm.ClusterTendency	0,637	0,538	0,522	0,533	0,555	0,549	0,599	0,599
grlm.ShortRunLowGrayLevelEmphasis	0,582	0,604	0,615	0,621	0,555	0,522	0,643	0,692
grlm.GrayLevelVariance	0,648	0,538	0,527	0,511	0,555	0,555	0,610	0,615
grlm.LowGrayLevelRunEmphasis	0,593	0,604	0,610	0,615	0,544	0,516	0,637	0,698
grlm.GrayLevelNonUniformityNormalized	0,626	0,533	0,566	0,511	0,571	0,555	0,599	0,626
grlm.RunVariance	0,566	0,571	0,544	0,544	0,566	0,582	0,698	0,692
grlm.GrayLevelNonUniformity	0,632	0,544	0,588	0,505	0,549	0,549	0,588	0,615
grlm.LongRunEmphasis	0,566	0,571	0,555	0,544	0,571	0,588	0,687	0,687
grlm.ShortRunHighGrayLevelEmphasis	0,522	0,549	0,582	0,522	0,505	0,549	0,654	0,648
grlm.RunLengthNonUniformity	0,654	0,577	0,665	0,533	0,681	0,577	0,670	0,692
grlm.ShortRunEmphasis	0,571	0,566	0,555	0,533	0,571	0,577	0,681	0,692
grlm.LongRunHighGrayLevelEmphasis	0,560	0,577	0,588	0,538	0,538	0,516	0,643	0,654
grlm.RunPercentage	0,566	0,571	0,555	0,538	0,571	0,577	0,681	0,687
grlm.LongRunLowGrayLevelEmphasis	0,599	0,599	0,610	0,599	0,533	0,505	0,654	0,709
grlm.RunEntropy	0,665	0,511	0,588	0,505	0,560	0,566	0,577	0,604
grlm.HighGrayLevelRunEmphasis	0,544	0,555	0,588	0,522	0,505	0,544	0,648	0,648
grlm.RunLengthNonUniformityNormalized	0,577	0,577	0,555	0,533	0,571	0,577	0,681	0,692
glszm.GrayLevelVariance	0,709	0,577	0,544	0,511	0,577	0,560	0,626	0,621
glszm.LowIntensityLargeAreaEmphasis	0,621	0,593	0,505	0,511	0,505	0,533	0,654	0,703
glszm.HighIntensitySmallAreaEmphasis	0,588	0,500	0,582	0,538	0,544	0,555	0,637	0,643
glszm.SmallAreaEmphasis	0,538	0,593	0,500	0,571	0,560	0,610	0,533	0,637
glszm.LargeAreaEmphasis	0,571	0,544	0,566	0,533	0,549	0,582	0,637	0,687
glszm.ZoneVariance	0,577	0,544	0,566	0,527	0,549	0,582	0,637	0,681
glszm.SizeZoneVariabilityNormalized	0,533	0,599	0,505	0,555	0,566	0,615	0,533	0,643
glszm.LowIntensitySmallAreaEmphasis	0,500	0,533	0,577	0,560	0,533	0,527	0,681	0,681
glszm.HighIntensityEmphasis	0,577	0,511	0,582	0,533	0,511	0,555	0,648	0,632
glszm.IntensityVariabilityNormalized	0,692	0,522	0,560	0,505	0,571	0,555	0,599	0,621
glszm.ZonePercentage	0,549	0,560	0,549	0,527	0,577	0,588	0,676	0,692
glszm.LowIntensityEmphasis	0,505	0,538	0,582	0,599	0,527	0,533	0,654	0,687
glszm.SizeZoneVariability	0,582	0,555	0,621	0,522	0,626	0,599	0,764	0,676
glszm.IntensityVariability	0,538	0,588	0,632	0,538	0,670	0,566	0,720	0,731
glszm.ZoneEntropy	0,698	0,522	0,621	0,511	0,610	0,566	0,632	0,615
glszm.HighIntensityLargeAreaEmphasis	0,522	0,533	0,621	0,588	0,615	0,643	0,527	0,621

TABLEAU A.7 – Cohorte Radiomique 4Mod : Évaluation des indices globaux calculés en T1w pour la prédiction de la survie globale. Les indices présentant l'indice-C plus grand que 0,55 sont en bleu. Les indices présentant en plus le RH sans la valeur 1 dans son intervalle de confiance et un ratio de vraisemblance statistiquement significatif sont en vert.

	RH (95% Intervalle de confiance)	indice-C (erreur-type)	Ratio de vraisemblance
T1w.firstorder.InterquartileRange	0,85 (0,60; 1,21)	0,54 (0,06)	p = 0,34
T1w.firstorder.Skewness	1,07 (0,71; 1,62)	0,52 (0,06)	p = 0,74
T1w.firstorder.Uniformity	1,23 (0,84; 1,80)	0,51 (0,06)	p = 0,29
T1w.firstorder.MeanAbsoluteDeviation	0,88 (0,63; 1,22)	0,52 (0,06)	p = 0,42
T1w.firstorder.Energy	0,95 (0,66; 1,37)	0,53 (0,06)	p = 0,80
T1w.firstorder.RobustMeanAbsoluteDeviation	0,86 (0,61; 1,22)	0,53 (0,06)	p = 0,37
T1w.firstorder.Median	1,10 (0,79; 1,52)	0,54 (0,06)	p = 0,58
T1w.firstorder.TotalEnergy	0,95 (0,66; 1,37)	0,53 (0,06)	p = 0,80
T1w.firstorder.Maximum	1,13 (0,69; 1,84)	0,47 (0,06)	p = 0,63
T1w.firstorder.RootMeanSquared	1,11 (0,80; 1,54)	0,55 (0,06)	p = 0,54
T1w.firstorder.90Percentile	1,07 (0,77; 1,49)	0,53 (0,06)	p = 0,68
T1w.firstorder.Minimum	1,07 (0,78; 1,48)	0,51 (0,06)	p = 0,68
T1w.firstorder.Entropy	0,87 (0,62; 1,24)	0,51 (0,06)	p = 0,44
T1w.firstorder.StandardDeviation	0,90 (0,65; 1,23)	0,51 (0,06)	p = 0,48
T1w.firstorder.Range	0,99 (0,70; 1,39)	0,47 (0,06)	p = 0,96
T1w.firstorder.Variance	0,89 (0,65; 1,22)	0,51 (0,06)	p = 0,45
T1w.firstorder.10Percentile	1,14 (0,82; 1,58)	0,53 (0,06)	p = 0,43
T1w.firstorder.Kurtosis	2,25 (1,24; 4,08)	0,60 (0,06)	p = 0,01
T1w.firstorder.Mean	1,11 (0,80; 1,54)	0,55 (0,06)	p = 0,54
T1w.glc.M.SumVariance	1,04 (0,72; 1,51)	0,51 (0,06)	p = 0,84
T1w.glc.M.Homogeneity1	0,94 (0,64; 1,39)	0,55 (0,06)	p = 0,77
T1w.glc.M.Homogeneity2	0,95 (0,65; 1,40)	0,55 (0,06)	p = 0,80
T1w.glc.M.ClusterShade	0,88 (0,64; 1,21)	0,51 (0,06)	p = 0,41
T1w.glc.M.MaximumProbability	1,18 (0,80; 1,73)	0,51 (0,06)	p = 0,42
T1w.glc.M.Idmn	0,94 (0,68; 1,31)	0,50 (0,06)	p = 0,71
T1w.glc.M.SumVariance2	0,88 (0,64; 1,22)	0,52 (0,06)	p = 0,41
T1w.glc.M.Contrast	1,14 (0,78; 1,66)	0,53 (0,06)	p = 0,50
T1w.glc.M.DifferenceEntropy	1,05 (0,72; 1,54)	0,53 (0,06)	p = 0,80
T1w.glc.M.InverseVariance	0,94 (0,63; 1,38)	0,54 (0,06)	p = 0,74
T1w.glc.M.Entropy	0,93 (0,65; 1,33)	0,49 (0,06)	p = 0,69
T1w.glc.M.Dissimilarity	1,11 (0,75; 1,64)	0,53 (0,06)	p = 0,61
T1w.glc.M.DifferenceVariance	1,11 (0,78; 1,58)	0,53 (0,06)	p = 0,57
T1w.glc.M.Idn	0,93 (0,67; 1,29)	0,52 (0,06)	p = 0,67
T1w.glc.M.Idm	0,95 (0,65; 1,40)	0,55 (0,06)	p = 0,80
T1w.glc.M.Correlation	0,71 (0,49; 1,04)	0,59 (0,06)	p = 0,08
T1w.glc.M.Autocorrelation	1,03 (0,71; 1,49)	0,52 (0,06)	p = 0,88
T1w.glc.M.SumEntropy	0,85 (0,59; 1,21)	0,52 (0,06)	p = 0,35
T1w.glc.M.AverageIntensity	1,02 (0,70; 1,49)	0,51 (0,06)	p = 0,90
T1w.glc.M.Energy	1,16 (0,79; 1,70)	0,49 (0,06)	p = 0,44
T1w.glc.M.SumSquares	0,89 (0,65; 1,22)	0,50 (0,06)	p = 0,45
T1w.glc.M.ClusterProminence	0,90 (0,66; 1,22)	0,48 (0,06)	p = 0,44
T1w.glc.M.SumAverage	1,03 (0,71; 1,49)	0,51 (0,06)	p = 0,89
T1w.glc.M.Imc2	0,73 (0,51; 1,06)	0,59 (0,06)	p = 0,10
T1w.glc.M.Imc1	1,29 (0,88; 1,88)	0,58 (0,06)	p = 0,19
T1w.glc.M.DifferenceAverage	1,11 (0,75; 1,64)	0,53 (0,06)	p = 0,61
T1w.glc.M.Id	0,94 (0,64; 1,39)	0,55 (0,06)	p = 0,77
T1w.glc.M.ClusterTendency	0,88 (0,64; 1,22)	0,52 (0,06)	p = 0,41
T1w.glr.M.ShortRunLowGrayLevelEmphasis	1,06 (0,75; 1,52)	0,48 (0,06)	p = 0,73
T1w.glr.M.GrayLevelVariance	0,89 (0,65; 1,22)	0,50 (0,06)	p = 0,46
T1w.glr.M.LowGrayLevelRunEmphasis	1,06 (0,75; 1,52)	0,48 (0,06)	p = 0,73
T1w.glr.M.GrayLevelNonUniformityNormalized	1,21 (0,83; 1,77)	0,52 (0,06)	p = 0,31
T1w.glr.M.RunVariance	0,96 (0,66; 1,39)	0,54 (0,06)	p = 0,82
T1w.glr.M.GrayLevelNonUniformity	1,02 (0,72; 1,45)	0,46 (0,06)	p = 0,91
T1w.glr.M.LongRunEmphasis	0,95 (0,66; 1,39)	0,54 (0,06)	p = 0,81
T1w.glr.M.ShortRunHighGrayLevelEmphasis	1,02 (0,71; 1,46)	0,51 (0,06)	p = 0,91
T1w.glr.M.RunLengthNonUniformity	0,97 (0,67; 1,41)	0,52 (0,06)	p = 0,88
T1w.glr.M.ShortRunEmphasis	1,07 (0,73; 1,56)	0,55 (0,06)	p = 0,74
T1w.glr.M.LongRunHighGrayLevelEmphasis	1,02 (0,70; 1,48)	0,50 (0,06)	p = 0,94
T1w.glr.M.RunPercentage	1,07 (0,73; 1,57)	0,55 (0,06)	p = 0,74
T1w.glr.M.LongRunLowGrayLevelEmphasis	1,07 (0,75; 1,52)	0,48 (0,06)	p = 0,74
T1w.glr.M.RunEntropy	0,84 (0,59; 1,19)	0,52 (0,06)	p = 0,31
T1w.glr.M.HighGrayLevelRunEmphasis	1,02 (0,71; 1,46)	0,52 (0,06)	p = 0,91
T1w.glr.M.RunLengthNonUniformityNormalized	1,08 (0,73; 1,59)	0,55 (0,06)	p = 0,70
T1w.gls.M.GrayLevelVariance	0,90 (0,66; 1,22)	0,48 (0,06)	p = 0,46
T1w.gls.M.LowIntensityLargeAreaEmphasis	1,13 (0,77; 1,65)	0,44 (0,06)	p = 0,57
T1w.gls.M.HighIntensitySmallAreaEmphasis	0,95 (0,69; 1,32)	0,48 (0,06)	p = 0,77
T1w.gls.M.SmallAreaEmphasis	1,08 (0,70; 1,68)	0,55 (0,06)	p = 0,73
T1w.gls.M.LargeAreaEmphasis	1,05 (0,71; 1,54)	0,44 (0,06)	p = 0,81
T1w.gls.M.ZoneVariance	1,05 (0,71; 1,54)	0,44 (0,06)	p = 0,81
T1w.gls.M.SizeZoneVariabilityNormalized	1,11 (0,71; 1,72)	0,55 (0,06)	p = 0,65
T1w.gls.M.LowIntensitySmallAreaEmphasis	1,11 (0,80; 1,53)	0,48 (0,06)	p = 0,55
T1w.gls.M.HighIntensityEmphasis	0,95 (0,69; 1,32)	0,47 (0,06)	p = 0,78
T1w.gls.M.IntensityVariabilityNormalized	1,16 (0,82; 1,64)	0,51 (0,06)	p = 0,40
T1w.gls.M.ZonePercentage	1,15 (0,78; 1,70)	0,54 (0,06)	p = 0,49
T1w.gls.M.LowIntensityEmphasis	1,10 (0,80; 1,53)	0,49 (0,06)	p = 0,58
T1w.gls.M.SizeZoneVariability	1,07 (0,74; 1,55)	0,49 (0,06)	p = 0,72
T1w.gls.M.IntensityVariability	1,13 (0,78; 1,65)	0,50 (0,06)	p = 0,53
T1w.gls.M.ZoneEntropy	0,85 (0,60; 1,20)	0,52 (0,06)	p = 0,34
T1w.gls.M.HighIntensityLargeAreaEmphasis	0,99 (0,71; 1,39)	0,56 (0,06)	p = 0,96

TABLEAU A.8 – Cohorte Radiomique 4Mod : Évaluation des indices globaux calculés en T2w pour la prédiction de la survie globale. Les indices présentant l'indice-C plus grand que 0,55 sont en bleu. Les indices présentant en plus le RH sans la valeur 1 dans son intervalle de confiance et un ratio de vraisemblance statistiquement significatif sont en vert.

	RH (95% Intervalle de confiance)	indice-C (erreur-type)	Ratio de vraisemblance
T2w.firstorder.InterquartileRange	0,96 (0,71; 1,29)	0,51 (0,06)	p = 0,79
T2w.firstorder.Skewness	1,58 (1,04; 2,39)	0,57 (0,06)	p = 0,03
T2w.firstorder.Uniformity	1,00 (0,72; 1,38)	0,51 (0,06)	p = 0,98
T2w.firstorder.MeanAbsoluteDeviation	0,98 (0,73; 1,31)	0,49 (0,06)	p = 0,88
T2w.firstorder.Energy	0,96 (0,67; 1,37)	0,53 (0,06)	p = 0,81
T2w.firstorder.RobustMeanAbsoluteDeviation	0,96 (0,71; 1,29)	0,51 (0,06)	p = 0,79
T2w.firstorder.Median	1,30 (0,88; 1,92)	0,56 (0,06)	p = 0,21
T2w.firstorder.TotalEnergy	0,96 (0,67; 1,37)	0,53 (0,06)	p = 0,81
T2w.firstorder.Maximum	1,54 (1,07; 2,23)	0,64 (0,06)	p = 0,02
T2w.firstorder.RootMeanSquared	1,32 (0,89; 1,96)	0,54 (0,06)	p = 0,18
T2w.firstorder.90Percentile	1,17 (0,81; 1,69)	0,54 (0,06)	p = 0,41
T2w.firstorder.Minimum	1,29 (0,84; 1,97)	0,59 (0,06)	p = 0,24
T2w.firstorder.Entropy	1,02 (0,74; 1,41)	0,54 (0,06)	p = 0,92
T2w.firstorder.StandardDeviation	1,01 (0,75; 1,35)	0,53 (0,06)	p = 0,95
T2w.firstorder.Range	1,20 (0,85; 1,69)	0,57 (0,06)	p = 0,30
T2w.firstorder.Variance	0,98 (0,74; 1,30)	0,47 (0,06)	p = 0,89
T2w.firstorder.10Percentile	1,36 (0,91; 2,01)	0,57 (0,06)	p = 0,14
T2w.firstorder.Kurtosis	2,34 (1,45; 3,79)	0,59 (0,06)	p < 0,01
T2w.firstorder.Mean	1,32 (0,89; 1,96)	0,54 (0,06)	p = 0,18
T2w.glm.SumVariance	1,04 (0,73; 1,50)	0,50 (0,06)	p = 0,82
T2w.glm.Homogeneity1	0,92 (0,67; 1,28)	0,56 (0,06)	p = 0,64
T2w.glm.Homogeneity2	0,92 (0,66; 1,27)	0,56 (0,06)	p = 0,60
T2w.glm.ClusterShade	1,08 (0,81; 1,43)	0,55 (0,06)	p = 0,62
T2w.glm.MaximumProbability	0,93 (0,66; 1,31)	0,56 (0,06)	p = 0,66
T2w.glm.Idmn	1,21 (0,89; 1,64)	0,57 (0,06)	p = 0,21
T2w.glm.SumVariance2	0,98 (0,75; 1,30)	0,47 (0,06)	p = 0,91
T2w.glm.Contrast	1,01 (0,75; 1,37)	0,55 (0,06)	p = 0,94
T2w.glm.DifferenceEntropy	1,04 (0,76; 1,42)	0,55 (0,06)	p = 0,81
T2w.glm.InverseVariance	0,91 (0,66; 1,26)	0,56 (0,06)	p = 0,58
T2w.glm.Entropy	1,02 (0,74; 1,41)	0,52 (0,06)	p = 0,90
T2w.glm.Dissimilarity	1,03 (0,76; 1,41)	0,56 (0,06)	p = 0,84
T2w.glm.DifferenceVariance	1,01 (0,75; 1,36)	0,56 (0,06)	p = 0,96
T2w.glm.Idn	1,21 (0,88; 1,68)	0,57 (0,06)	p = 0,24
T2w.glm.Idm	0,92 (0,66; 1,27)	0,56 (0,06)	p = 0,60
T2w.glm.Correlation	0,98 (0,71; 1,35)	0,56 (0,06)	p = 0,89
T2w.glm.Autocorrelation	1,04 (0,73; 1,49)	0,50 (0,06)	p = 0,83
T2w.glm.SumEntropy	1,01 (0,73; 1,40)	0,53 (0,06)	p = 0,95
T2w.glm.AverageIntensity	1,00 (0,69; 1,44)	0,51 (0,06)	p = 0,98
T2w.glm.Energy	0,97 (0,71; 1,34)	0,53 (0,06)	p = 0,86
T2w.glm.SumSquares	0,99 (0,75; 1,31)	0,46 (0,06)	p = 0,94
T2w.glm.ClusterProminence	1,01 (0,77; 1,33)	0,56 (0,06)	p = 0,94
T2w.glm.SumAverage	1,00 (0,69; 1,44)	0,50 (0,06)	p = 0,99
T2w.glm.Imc2	0,97 (0,68; 1,40)	0,50 (0,06)	p = 0,89
T2w.glm.Imc1	1,01 (0,71; 1,44)	0,52 (0,06)	p = 0,94
T2w.glm.DifferenceAverage	1,03 (0,76; 1,41)	0,56 (0,06)	p = 0,84
T2w.glm.Id	0,92 (0,67; 1,28)	0,56 (0,06)	p = 0,64
T2w.glm.ClusterTendency	0,98 (0,75; 1,30)	0,47 (0,06)	p = 0,91
T2w.grlm.ShortRunLowGrayLevelEmphasis	1,05 (0,77; 1,42)	0,52 (0,06)	p = 0,77
T2w.grlm.GrayLevelVariance	0,98 (0,74; 1,30)	0,46 (0,06)	p = 0,90
T2w.grlm.LowGrayLevelRunEmphasis	1,05 (0,77; 1,42)	0,52 (0,06)	p = 0,78
T2w.grlm.GrayLevelNonUniformityNormalized	1,00 (0,72; 1,38)	0,51 (0,06)	p = 0,99
T2w.grlm.RunVariance	0,88 (0,63; 1,22)	0,57 (0,06)	p = 0,43
T2w.grlm.GrayLevelNonUniformity	0,98 (0,68; 1,41)	0,51 (0,06)	p = 0,90
T2w.grlm.LongRunEmphasis	0,88 (0,63; 1,22)	0,56 (0,06)	p = 0,44
T2w.grlm.ShortRunHighGrayLevelEmphasis	1,05 (0,73; 1,50)	0,50 (0,06)	p = 0,81
T2w.grlm.RunLengthNonUniformity	0,97 (0,67; 1,39)	0,54 (0,06)	p = 0,86
T2w.grlm.ShortRunEmphasis	1,13 (0,81; 1,58)	0,57 (0,06)	p = 0,45
T2w.grlm.LongRunHighGrayLevelEmphasis	1,02 (0,71; 1,47)	0,50 (0,06)	p = 0,91
T2w.grlm.RunPercentage	1,14 (0,82; 1,58)	0,57 (0,06)	p = 0,45
T2w.grlm.LongRunLowGrayLevelEmphasis	1,04 (0,77; 1,41)	0,51 (0,06)	p = 0,80
T2w.grlm.RunEntropy	0,99 (0,72; 1,37)	0,48 (0,06)	p = 0,96
T2w.grlm.HighGrayLevelRunEmphasis	1,04 (0,73; 1,49)	0,50 (0,06)	p = 0,83
T2w.grlm.RunLengthNonUniformityNormalized	1,13 (0,82; 1,58)	0,57 (0,06)	p = 0,45
T2w.glszm.GrayLevelVariance	1,01 (0,76; 1,34)	0,54 (0,06)	p = 0,94
T2w.glszm.LowIntensityLargeAreaEmphasis	0,98 (0,73; 1,31)	0,55 (0,06)	p = 0,88
T2w.glszm.HighIntensitySmallAreaEmphasis	1,07 (0,75; 1,52)	0,51 (0,06)	p = 0,73
T2w.glszm.SmallAreaEmphasis	0,95 (0,68; 1,31)	0,51 (0,06)	p = 0,74
T2w.glszm.LargeAreaEmphasis	0,96 (0,69; 1,32)	0,56 (0,06)	p = 0,79
T2w.glszm.ZoneVariance	0,96 (0,69; 1,33)	0,56 (0,06)	p = 0,80
T2w.glszm.SizeZoneVariabilityNormalized	0,95 (0,69; 1,31)	0,51 (0,06)	p = 0,74
T2w.glszm.LowIntensitySmallAreaEmphasis	1,01 (0,74; 1,39)	0,50 (0,06)	p = 0,95
T2w.glszm.HighIntensityEmphasis	1,07 (0,75; 1,52)	0,51 (0,06)	p = 0,73
T2w.glszm.IntensityVariabilityNormalized	1,00 (0,73; 1,36)	0,52 (0,06)	p = 0,98
T2w.glszm.ZonePercentage	1,11 (0,80; 1,54)	0,56 (0,06)	p = 0,54
T2w.glszm.LowIntensityEmphasis	1,02 (0,75; 1,40)	0,51 (0,06)	p = 0,89
T2w.glszm.SizeZoneVariability	1,02 (0,73; 1,42)	0,50 (0,06)	p = 0,93
T2w.glszm.IntensityVariability	1,10 (0,74; 1,63)	0,52 (0,06)	p = 0,63
T2w.glszm.ZoneEntropy	1,03 (0,75; 1,40)	0,53 (0,06)	p = 0,86
T2w.glszm.HighIntensityLargeAreaEmphasis	0,92 (0,64; 1,34)	0,57 (0,06)	p = 0,67

TABLEAU A.9 – Cohorte Radiomique 4Mod : Évaluation des indices globaux calculés en T1c pour la prédiction de la survie globale. Les indices présentant l'indice-C plus grand que 0,55 sont en bleu.

	RH (95% Intervalle de confiance)	indice-C (erreur-type)	Ratio de vraisemblance
T1c.firstorder.InterquartileRange	1,07 (0,78; 1,46)	0,53 (0,06)	p = 0,70
T1c.firstorder.Skewness	1,35 (0,93; 1,94)	0,60 (0,06)	p = 0,12
T1c.firstorder.Uniformity	0,90 (0,65; 1,25)	0,55 (0,06)	p = 0,53
T1c.firstorder.MeanAbsoluteDeviation	1,09 (0,80; 1,49)	0,56 (0,06)	p = 0,60
T1c.firstorder.Energy	0,96 (0,67; 1,37)	0,53 (0,06)	p = 0,81
T1c.firstorder.RobustMeanAbsoluteDeviation	1,08 (0,78; 1,48)	0,53 (0,06)	p = 0,66
T1c.firstorder.Median	1,13 (0,83; 1,56)	0,51 (0,06)	p = 0,44
T1c.firstorder.TotalEnergy	0,96 (0,67; 1,37)	0,53 (0,06)	p = 0,81
T1c.firstorder.Maximum	1,12 (0,84; 1,50)	0,57 (0,06)	p = 0,45
T1c.firstorder.RootMeanSquared	1,20 (0,87; 1,65)	0,53 (0,06)	p = 0,28
T1c.firstorder.90Percentile	1,17 (0,85; 1,61)	0,55 (0,06)	p = 0,34
T1c.firstorder.Minimum	1,04 (0,77; 1,39)	0,45 (0,06)	p = 0,81
T1c.firstorder.Entropy	1,12 (0,82; 1,53)	0,55 (0,06)	p = 0,49
T1c.firstorder.StandardDeviation	1,10 (0,81; 1,49)	0,56 (0,06)	p = 0,56
T1c.firstorder.Range	1,09 (0,81; 1,46)	0,58 (0,06)	p = 0,57
T1c.firstorder.Variance	1,06 (0,77; 1,45)	0,56 (0,06)	p = 0,73
T1c.firstorder.10Percentile	1,13 (0,84; 1,51)	0,54 (0,06)	p = 0,43
T1c.firstorder.Kurtosis	1,47 (1,01; 2,15)	0,55 (0,06)	p = 0,06
T1c.firstorder.Mean	1,20 (0,87; 1,65)	0,53 (0,06)	p = 0,28
T1c.glcM.SumVariance	1,08 (0,79; 1,48)	0,57 (0,06)	p = 0,64
T1c.glcM.Homogeneity1	0,88 (0,63; 1,22)	0,57 (0,06)	p = 0,44
T1c.glcM.Homogeneity2	0,88 (0,63; 1,23)	0,57 (0,06)	p = 0,45
T1c.glcM.ClusterShade	1,04 (0,76; 1,43)	0,57 (0,06)	p = 0,82
T1c.glcM.MaximumProbability	0,92 (0,66; 1,28)	0,54 (0,06)	p = 0,63
T1c.glcM.Idmn	0,84 (0,58; 1,22)	0,50 (0,06)	p = 0,38
T1c.glcM.SumVariance2	1,05 (0,77; 1,43)	0,56 (0,06)	p = 0,77
T1c.glcM.Contrast	1,12 (0,81; 1,56)	0,58 (0,06)	p = 0,52
T1c.glcM.DifferenceEntropy	1,15 (0,83; 1,59)	0,58 (0,06)	p = 0,40
T1c.glcM.InverseVariance	0,88 (0,63; 1,23)	0,57 (0,06)	p = 0,46
T1c.glcM.Entropy	1,12 (0,82; 1,54)	0,56 (0,06)	p = 0,49
T1c.glcM.Dissimilarity	1,14 (0,82; 1,58)	0,57 (0,06)	p = 0,45
T1c.glcM.DifferenceVariance	1,11 (0,80; 1,53)	0,58 (0,06)	p = 0,55
T1c.glcM.Idn	0,95 (0,64; 1,40)	0,48 (0,06)	p = 0,78
T1c.glcM.Idm	0,88 (0,63; 1,23)	0,57 (0,06)	p = 0,45
T1c.glcM.Correlation	0,94 (0,65; 1,36)	0,53 (0,06)	p = 0,74
T1c.glcM.Autocorrelation	1,08 (0,79; 1,48)	0,57 (0,06)	p = 0,63
T1c.glcM.SumEntropy	1,12 (0,81; 1,54)	0,56 (0,06)	p = 0,49
T1c.glcM.AverageIntensity	1,11 (0,81; 1,51)	0,57 (0,06)	p = 0,53
T1c.glcM.Energy	0,90 (0,64; 1,25)	0,56 (0,06)	p = 0,52
T1c.glcM.SumSquares	1,05 (0,77; 1,44)	0,56 (0,06)	p = 0,75
T1c.glcM.ClusterProminence	0,99 (0,72; 1,35)	0,43 (0,06)	p = 0,95
T1c.glcM.SumAverage	1,11 (0,82; 1,52)	0,57 (0,06)	p = 0,51
T1c.glcM.Imc2	1,10 (0,76; 1,58)	0,52 (0,06)	p = 0,61
T1c.glcM.Imc1	0,95 (0,67; 1,34)	0,51 (0,06)	p = 0,75
T1c.glcM.DifferenceAverage	1,14 (0,82; 1,58)	0,57 (0,06)	p = 0,45
T1c.glcM.Id	0,88 (0,63; 1,22)	0,57 (0,06)	p = 0,44
T1c.glcM.ClusterTendency	1,05 (0,77; 1,43)	0,56 (0,06)	p = 0,77
T1c.glrM.ShortRunLowGrayLevelEmphasis	0,93 (0,65; 1,34)	0,57 (0,06)	p = 0,70
T1c.glrM.GrayLevelVariance	1,06 (0,77; 1,45)	0,56 (0,06)	p = 0,73
T1c.glrM.LowGrayLevelRunEmphasis	0,93 (0,65; 1,33)	0,57 (0,06)	p = 0,68
T1c.glrM.GrayLevelNonUniformityNormalized	0,90 (0,64; 1,25)	0,55 (0,06)	p = 0,51
T1c.glrM.RunVariance	0,86 (0,61; 1,20)	0,57 (0,06)	p = 0,36
T1c.glrM.GrayLevelNonUniformity	0,88 (0,63; 1,22)	0,56 (0,06)	p = 0,43
T1c.glrM.LongRunEmphasis	0,86 (0,62; 1,20)	0,57 (0,06)	p = 0,37
T1c.glrM.ShortRunHighGrayLevelEmphasis	1,08 (0,79; 1,47)	0,57 (0,06)	p = 0,64
T1c.glrM.RunLengthNonUniformity	0,99 (0,69; 1,41)	0,52 (0,06)	p = 0,95
T1c.glrM.ShortRunEmphasis	1,16 (0,83; 1,62)	0,57 (0,06)	p = 0,37
T1c.glrM.LongRunHighGrayLevelEmphasis	1,08 (0,79; 1,47)	0,57 (0,06)	p = 0,64
T1c.glrM.RunPercentage	1,16 (0,83; 1,63)	0,56 (0,06)	p = 0,37
T1c.glrM.LongRunLowGrayLevelEmphasis	0,91 (0,64; 1,29)	0,58 (0,06)	p = 0,60
T1c.glrM.RunEntropy	1,11 (0,81; 1,52)	0,56 (0,06)	p = 0,52
T1c.glrM.HighGrayLevelRunEmphasis	1,08 (0,79; 1,47)	0,57 (0,06)	p = 0,64
T1c.glrM.RunLengthNonUniformityNormalized	1,16 (0,83; 1,62)	0,57 (0,06)	p = 0,37
T1c.glszm.GrayLevelVariance	1,07 (0,79; 1,46)	0,57 (0,06)	p = 0,65
T1c.glszm.LowIntensityLargeAreaEmphasis	0,86 (0,62; 1,20)	0,58 (0,06)	p = 0,35
T1c.glszm.HighIntensitySmallAreaEmphasis	1,08 (0,79; 1,47)	0,56 (0,06)	p = 0,63
T1c.glszm.SmallAreaEmphasis	1,12 (0,80; 1,57)	0,55 (0,06)	p = 0,51
T1c.glszm.LargeAreaEmphasis	0,85 (0,61; 1,18)	0,57 (0,06)	p = 0,31
T1c.glszm.ZoneVariance	0,85 (0,61; 1,19)	0,57 (0,06)	p = 0,31
T1c.glszm.SizeZoneVariabilityNormalized	1,12 (0,80; 1,57)	0,55 (0,06)	p = 0,50
T1c.glszm.LowIntensitySmallAreaEmphasis	0,93 (0,66; 1,32)	0,58 (0,06)	p = 0,69
T1c.glszm.HighIntensityEmphasis	1,08 (0,80; 1,47)	0,57 (0,06)	p = 0,62
T1c.glszm.IntensityVariabilityNormalized	0,86 (0,62; 1,19)	0,54 (0,06)	p = 0,36
T1c.glszm.ZonePercentage	1,18 (0,85; 1,63)	0,58 (0,06)	p = 0,35
T1c.glszm.LowIntensityEmphasis	0,93 (0,66; 1,32)	0,57 (0,06)	p = 0,69
T1c.glszm.SizeZoneVariability	1,06 (0,78; 1,44)	0,53 (0,06)	p = 0,73
T1c.glszm.IntensityVariability	1,07 (0,73; 1,58)	0,50 (0,06)	p = 0,72
T1c.glszm.ZoneEntropy	1,12 (0,82; 1,52)	0,55 (0,06)	p = 0,49
T1c.glszm.HighIntensityLargeAreaEmphasis	0,88 (0,63; 1,25)	0,57 (0,06)	p = 0,46

TABLEAU A.10 – Cohorte Radiomique 4Mod : Évaluation des indices globaux calculés en FLAIR pour la prédiction de la survie globale. Les indices présentant l'indice-C plus grand que 0,55 sont en bleu. Les indices présentant en plus le RH sans la valeur 1 dans son intervalle de confiance et un ratio de vraisemblance statistiquement significatif sont en vert.

	RH (95% Intervalle de confiance)	indice-C (erreur-type)	Ratio de vraisemblance
FLAIR.firstorder.InterquartileRange	0,71 (0,49; 1,04)	0,61 (0,06)	p = 0,06
FLAIR.firstorder.Skewness	1,80 (1,21; 2,68)	0,63 (0,06)	p < 0,01
FLAIR.firstorder.Uniformity	1,41 (0,98; 2,01)	0,60 (0,06)	p = 0,06
FLAIR.firstorder.MeanAbsoluteDeviation	0,71 (0,50; 1,03)	0,61 (0,06)	p = 0,05
FLAIR.firstorder.Energy	0,94 (0,66; 1,35)	0,53 (0,06)	p = 0,75
FLAIR.firstorder.RobustMeanAbsoluteDeviation	0,70 (0,48; 1,02)	0,61 (0,06)	p = 0,05
FLAIR.firstorder.Median	0,86 (0,61; 1,19)	0,53 (0,06)	p = 0,36
FLAIR.firstorder.TotalEnergy	0,94 (0,66; 1,35)	0,53 (0,06)	p = 0,75
FLAIR.firstorder.Maximum	1,01 (0,68; 1,48)	0,48 (0,06)	p = 0,98
FLAIR.firstorder.RootMeanSquared	0,91 (0,64; 1,28)	0,52 (0,06)	p = 0,59
FLAIR.firstorder.90Percentile	0,83 (0,58; 1,19)	0,55 (0,06)	p = 0,31
FLAIR.firstorder.Minimum	1,30 (0,93; 1,82)	0,60 (0,06)	p = 0,12
FLAIR.firstorder.Entropy	0,73 (0,51; 1,05)	0,59 (0,06)	p = 0,08
FLAIR.firstorder.StandardDeviation	0,73 (0,51; 1,04)	0,59 (0,06)	p = 0,07
FLAIR.firstorder.Range	0,84 (0,57; 1,23)	0,56 (0,06)	p = 0,35
FLAIR.firstorder.Variance	0,73 (0,50; 1,06)	0,59 (0,06)	p = 0,08
FLAIR.firstorder.10Percentile	1,13 (0,81; 1,56)	0,54 (0,06)	p = 0,47
FLAIR.firstorder.Kurtosis	1,03 (0,77; 1,39)	0,54 (0,06)	p = 0,84
FLAIR.firstorder.Mean	0,91 (0,64; 1,29)	0,52 (0,06)	p = 0,59
FLAIR.g lcm.SumVariance	0,75 (0,50; 1,12)	0,58 (0,06)	p = 0,13
FLAIR.g lcm.Homogeneity1	1,28 (0,90; 1,82)	0,57 (0,06)	p = 0,17
FLAIR.g lcm.Homogeneity2	1,28 (0,91; 1,82)	0,57 (0,06)	p = 0,16
FLAIR.g lcm.ClusterShade	1,79 (1,26; 2,55)	0,65 (0,06)	p = 0,01
FLAIR.g lcm.MaximumProbability	1,31 (0,92; 1,87)	0,58 (0,06)	p = 0,14
FLAIR.g lcm.Idmn	1,00 (0,74; 1,34)	0,52 (0,06)	p = 0,98
FLAIR.g lcm.SumVariance2	0,73 (0,50; 1,06)	0,60 (0,06)	p = 0,08
FLAIR.g lcm.Contrast	0,87 (0,59; 1,29)	0,56 (0,06)	p = 0,48
FLAIR.g lcm.DifferenceEntropy	0,79 (0,55; 1,14)	0,56 (0,06)	p = 0,21
FLAIR.g lcm.InverseVariance	1,28 (0,90; 1,82)	0,57 (0,06)	p = 0,17
FLAIR.g lcm.Entropy	0,74 (0,51; 1,06)	0,59 (0,06)	p = 0,09
FLAIR.g lcm.Dissimilarity	0,82 (0,56; 1,19)	0,56 (0,06)	p = 0,27
FLAIR.g lcm.DifferenceVariance	0,91 (0,61; 1,34)	0,56 (0,06)	p = 0,62
FLAIR.g lcm.Idn	0,98 (0,71; 1,34)	0,51 (0,06)	p = 0,89
FLAIR.g lcm.Idm	1,28 (0,91; 1,82)	0,57 (0,06)	p = 0,16
FLAIR.g lcm.Correlation	0,71 (0,50; 1,02)	0,58 (0,06)	p = 0,06
FLAIR.g lcm.Autocorrelation	0,74 (0,50; 1,11)	0,57 (0,06)	p = 0,12
FLAIR.g lcm.SumEntropy	0,73 (0,51; 1,04)	0,59 (0,06)	p = 0,08
FLAIR.g lcm.AverageIntensity	0,74 (0,51; 1,07)	0,57 (0,06)	p = 0,09
FLAIR.g lcm.Energy	1,41 (0,99; 2,01)	0,60 (0,06)	p = 0,07
FLAIR.g lcm.SumSquares	0,73 (0,50; 1,07)	0,59 (0,06)	p = 0,09
FLAIR.g lcm.ClusterProminence	0,76 (0,50; 1,14)	0,59 (0,06)	p = 0,15
FLAIR.g lcm.SumAverage	0,74 (0,51; 1,07)	0,57 (0,06)	p = 0,10
FLAIR.g lcm.Imc2	0,72 (0,50; 1,05)	0,58 (0,06)	p = 0,08
FLAIR.g lcm.Imc1	1,28 (0,88; 1,85)	0,55 (0,06)	p = 0,18
FLAIR.g lcm.DifferenceAverage	0,82 (0,56; 1,19)	0,56 (0,06)	p = 0,27
FLAIR.g lcm.Id	1,28 (0,90; 1,82)	0,57 (0,06)	p = 0,17
FLAIR.g lcm.ClusterTendency	0,73 (0,50; 1,06)	0,60 (0,06)	p = 0,08
FLAIR.g lrlm.ShortRunLowGrayLevelEmphasis	1,13 (0,85; 1,49)	0,56 (0,06)	p = 0,42
FLAIR.g lrlm.GrayLevelVariance	0,73 (0,50; 1,06)	0,59 (0,06)	p = 0,08
FLAIR.g lrlm.LowGrayLevelRunEmphasis	1,14 (0,86; 1,51)	0,57 (0,06)	p = 0,39
FLAIR.g lrlm.GrayLevelNonUniformityNormalized	1,40 (0,98; 2,01)	0,60 (0,06)	p = 0,07
FLAIR.g lrlm.RunVariance	1,25 (0,89; 1,75)	0,56 (0,06)	p = 0,21
FLAIR.g lrlm.GrayLevelNonUniformity	1,22 (0,85; 1,75)	0,54 (0,06)	p = 0,28
FLAIR.g lrlm.LongRunEmphasis	1,24 (0,89; 1,74)	0,56 (0,06)	p = 0,21
FLAIR.g lrlm.ShortRunHighGrayLevelEmphasis	0,75 (0,50; 1,11)	0,57 (0,06)	p = 0,12
FLAIR.g lrlm.RunLengthNonUniformity	0,90 (0,63; 1,29)	0,55 (0,06)	p = 0,57
FLAIR.g lrlm.ShortRunEmphasis	0,81 (0,57; 1,13)	0,56 (0,06)	p = 0,22
FLAIR.g lrlm.LongRunHighGrayLevelEmphasis	0,74 (0,50; 1,10)	0,59 (0,06)	p = 0,11
FLAIR.g lrlm.RunPercentage	0,81 (0,57; 1,13)	0,56 (0,06)	p = 0,22
FLAIR.g lrlm.LongRunLowGrayLevelEmphasis	1,19 (0,88; 1,60)	0,55 (0,06)	p = 0,29
FLAIR.g lrlm.RunEntropy	0,73 (0,51; 1,04)	0,59 (0,06)	p = 0,08
FLAIR.g lrlm.HighGrayLevelRunEmphasis	0,74 (0,50; 1,11)	0,57 (0,06)	p = 0,12
FLAIR.g lrlm.RunLengthNonUniformityNormalized	0,81 (0,57; 1,14)	0,56 (0,06)	p = 0,22
FLAIR.g lszm.GrayLevelVariance	0,75 (0,51; 1,11)	0,57 (0,06)	p = 0,13
FLAIR.g lszm.LowIntensityLargeAreaEmphasis	1,63 (1,11; 2,39)	0,57 (0,06)	p = 0,03
FLAIR.g lszm.HighIntensitySmallAreaEmphasis	0,79 (0,53; 1,17)	0,55 (0,06)	p = 0,21
FLAIR.g lszm.SmallAreaEmphasis	1,15 (0,85; 1,55)	0,57 (0,06)	p = 0,36
FLAIR.g lszm.LargeAreaEmphasis	1,25 (0,91; 1,71)	0,56 (0,06)	p = 0,20
FLAIR.g lszm.ZoneVariance	1,25 (0,91; 1,71)	0,56 (0,06)	p = 0,20
FLAIR.g lszm.SizeZoneVariabilityNormalized	1,14 (0,85; 1,53)	0,57 (0,06)	p = 0,39
FLAIR.g lszm.LowIntensitySmallAreaEmphasis	1,15 (0,86; 1,53)	0,56 (0,06)	p = 0,37
FLAIR.g lszm.HighIntensityEmphasis	0,77 (0,52; 1,15)	0,56 (0,06)	p = 0,18
FLAIR.g lszm.IntensityVariabilityNormalized	1,35 (0,95; 1,92)	0,57 (0,06)	p = 0,10
FLAIR.g lszm.ZonePercentage	0,83 (0,58; 1,18)	0,56 (0,06)	p = 0,28
FLAIR.g lszm.LowIntensityEmphasis	1,16 (0,86; 1,54)	0,55 (0,06)	p = 0,35
FLAIR.g lszm.SizeZoneVariability	0,85 (0,58; 1,25)	0,57 (0,06)	p = 0,40
FLAIR.g lszm.IntensityVariability	0,95 (0,66; 1,35)	0,54 (0,06)	p = 0,75
FLAIR.g lszm.ZoneEntropy	0,71 (0,50; 1,01)	0,58 (0,06)	p = 0,05
FLAIR.g lszm.HighIntensityLargeAreaEmphasis	1,05 (0,77; 1,45)	0,52 (0,06)	p = 0,75

TABLEAU A.11 – Cohorte Radiomique 4Mod : Évaluation des indices locaux calculés en T1w pour la prédiction de la survie globale. Les indices présentant l'indice-C plus grand que 0,55 sont en bleu.

	RH (95% Intervalle de confiance)	indice-C (erreur-type)	Ratio de vraisemblance
T1w.firstorder.InterquartileRange	0,96 (0,66; 1,39)	0,50 (0,06)	p = 0,83
T1w.firstorder.Skewness	0,91 (0,69; 1,20)	0,54 (0,06)	p = 0,52
T1w.firstorder.Uniformity	1,11 (0,73; 1,69)	0,49 (0,06)	p = 0,64
T1w.firstorder.MeanAbsoluteDeviation	0,96 (0,67; 1,40)	0,51 (0,06)	p = 0,85
T1w.firstorder.Energy	1,12 (0,81; 1,55)	0,56 (0,06)	p = 0,49
T1w.firstorder.RobustMeanAbsoluteDeviation	0,96 (0,66; 1,39)	0,51 (0,06)	p = 0,82
T1w.firstorder.Median	1,12 (0,81; 1,55)	0,56 (0,06)	p = 0,49
T1w.firstorder.TotalEnergy	1,12 (0,81; 1,55)	0,56 (0,06)	p = 0,49
T1w.firstorder.Maximum	1,16 (0,81; 1,65)	0,55 (0,06)	p = 0,42
T1w.firstorder.RootMeanSquared	1,12 (0,81; 1,55)	0,56 (0,06)	p = 0,49
T1w.firstorder.90Percentile	1,12 (0,81; 1,56)	0,55 (0,06)	p = 0,49
T1w.firstorder.Minimum	1,10 (0,79; 1,53)	0,53 (0,06)	p = 0,57
T1w.firstorder.Entropy	0,96 (0,63; 1,45)	0,49 (0,06)	p = 0,83
T1w.firstorder.StandardDeviation	0,97 (0,67; 1,41)	0,50 (0,06)	p = 0,87
T1w.firstorder.Range	1,02 (0,69; 1,50)	0,51 (0,06)	p = 0,93
T1w.firstorder.Variance	0,99 (0,70; 1,41)	0,50 (0,06)	p = 0,96
T1w.firstorder.10Percentile	1,12 (0,81; 1,54)	0,53 (0,06)	p = 0,51
T1w.firstorder.Kurtosis	1,05 (0,75; 1,48)	0,53 (0,06)	p = 0,77
T1w.firstorder.Mean	1,12 (0,81; 1,55)	0,56 (0,06)	p = 0,49
T1w.g lcm.SumVariance	1,09 (0,75; 1,58)	0,54 (0,06)	p = 0,66
T1w.g lcm.Homogeneity1	0,91 (0,60; 1,39)	0,54 (0,06)	p = 0,67
T1w.g lcm.Homogeneity2	0,93 (0,61; 1,40)	0,54 (0,06)	p = 0,71
T1w.g lcm.ClusterShade	0,89 (0,66; 1,21)	0,60 (0,06)	p = 0,46
T1w.g lcm.MaximumProbability	1,07 (0,71; 1,60)	0,48 (0,06)	p = 0,76
T1w.g lcm.Idmn	0,85 (0,61; 1,20)	0,56 (0,06)	p = 0,37
T1w.g lcm.SumVariance2	0,96 (0,68; 1,36)	0,50 (0,06)	p = 0,82
T1w.g lcm.Contrast	1,27 (0,83; 1,96)	0,52 (0,06)	p = 0,29
T1w.g lcm.DifferenceEntropy	1,09 (0,72; 1,66)	0,53 (0,06)	p = 0,68
T1w.g lcm.InverseVariance	0,92 (0,60; 1,39)	0,54 (0,06)	p = 0,69
T1w.g lcm.Entropy	1,00 (0,66; 1,53)	0,52 (0,06)	p = 1,00
T1w.g lcm.Dissimilarity	1,19 (0,77; 1,84)	0,53 (0,06)	p = 0,43
T1w.g lcm.DifferenceVariance	1,22 (0,81; 1,85)	0,51 (0,06)	p = 0,36
T1w.g lcm.Idn	0,85 (0,61; 1,19)	0,56 (0,06)	p = 0,34
T1w.g lcm.Idm	0,93 (0,61; 1,40)	0,54 (0,06)	p = 0,71
T1w.g lcm.Correlation	0,78 (0,54; 1,12)	0,58 (0,06)	p = 0,18
T1w.g lcm.Autocorrelation	1,08 (0,74; 1,57)	0,54 (0,06)	p = 0,69
T1w.g lcm.SumEntropy	0,92 (0,61; 1,39)	0,50 (0,06)	p = 0,69
T1w.g lcm.AverageIntensity	1,07 (0,73; 1,57)	0,53 (0,06)	p = 0,73
T1w.g lcm.Energy	1,09 (0,73; 1,62)	0,48 (0,06)	p = 0,68
T1w.g lcm.SumSquares	1,00 (0,70; 1,42)	0,50 (0,06)	p = 0,99
T1w.g lcm.ClusterProminence	0,97 (0,72; 1,31)	0,49 (0,06)	p = 0,86
T1w.g lcm.SumAverage	1,07 (0,73; 1,57)	0,54 (0,06)	p = 0,73
T1w.g lcm.Imc2	0,84 (0,56; 1,27)	0,51 (0,06)	p = 0,41
T1w.g lcm.Imc1	1,15 (0,80; 1,65)	0,52 (0,06)	p = 0,43
T1w.g lcm.DifferenceAverage	1,19 (0,77; 1,84)	0,53 (0,06)	p = 0,43
T1w.g lcm.Id	0,91 (0,60; 1,39)	0,54 (0,06)	p = 0,67
T1w.g lcm.ClusterTendency	0,96 (0,68; 1,36)	0,50 (0,06)	p = 0,82
T1w.g lrlm.ShortRunLowGrayLevelEmphasis	0,95 (0,64; 1,41)	0,55 (0,06)	p = 0,80
T1w.g lrlm.GrayLevelVariance	0,99 (0,70; 1,40)	0,50 (0,06)	p = 0,95
T1w.g lrlm.LowGrayLevelRunEmphasis	0,96 (0,65; 1,42)	0,55 (0,06)	p = 0,82
T1w.g lrlm.GrayLevelNonUniformityNormalized	1,11 (0,73; 1,68)	0,49 (0,06)	p = 0,64
T1w.g lrlm.RunVariance	0,96 (0,65; 1,43)	0,56 (0,06)	p = 0,85
T1w.g lrlm.GrayLevelNonUniformity	1,12 (0,73; 1,71)	0,50 (0,06)	p = 0,62
T1w.g lrlm.LongRunEmphasis	0,96 (0,64; 1,42)	0,56 (0,06)	p = 0,83
T1w.g lrlm.ShortRunHighGrayLevelEmphasis	1,09 (0,75; 1,60)	0,53 (0,06)	p = 0,65
T1w.g lrlm.RunLengthNonUniformity	1,09 (0,72; 1,64)	0,56 (0,06)	p = 0,69
T1w.g lrlm.ShortRunEmphasis	1,07 (0,71; 1,60)	0,56 (0,06)	p = 0,75
T1w.g lrlm.LongRunHighGrayLevelEmphasis	1,05 (0,73; 1,51)	0,53 (0,06)	p = 0,78
T1w.g lrlm.RunPercentage	1,07 (0,71; 1,60)	0,56 (0,06)	p = 0,75
T1w.g lrlm.LongRunLowGrayLevelEmphasis	0,97 (0,66; 1,43)	0,55 (0,06)	p = 0,89
T1w.g lrlm.RunEntropy	0,91 (0,61; 1,38)	0,50 (0,06)	p = 0,67
T1w.g lrlm.HighGrayLevelRunEmphasis	1,08 (0,74; 1,58)	0,54 (0,06)	p = 0,68
T1w.g lrlm.RunLengthNonUniformityNormalized	1,08 (0,72; 1,63)	0,56 (0,06)	p = 0,70
T1w.g lszm.GrayLevelVariance	0,98 (0,69; 1,39)	0,50 (0,06)	p = 0,91
T1w.g lszm.LowIntensityLargeAreaEmphasis	1,09 (0,77; 1,54)	0,46 (0,06)	p = 0,63
T1w.g lszm.HighIntensitySmallAreaEmphasis	1,09 (0,74; 1,59)	0,52 (0,06)	p = 0,67
T1w.g lszm.SmallAreaEmphasis	1,17 (0,78; 1,76)	0,56 (0,06)	p = 0,43
T1w.g lszm.LargeAreaEmphasis	1,11 (0,77; 1,59)	0,47 (0,06)	p = 0,59
T1w.g lszm.ZoneVariance	1,12 (0,78; 1,61)	0,48 (0,06)	p = 0,56
T1w.g lszm.SizeZoneVariabilityNormalized	1,21 (0,80; 1,82)	0,56 (0,06)	p = 0,36
T1w.g lszm.LowIntensitySmallAreaEmphasis	1,09 (0,73; 1,64)	0,48 (0,06)	p = 0,67
T1w.g lszm.HighIntensityEmphasis	1,05 (0,72; 1,54)	0,52 (0,06)	p = 0,79
T1w.g lszm.IntensityVariabilityNormalized	1,12 (0,75; 1,67)	0,51 (0,06)	p = 0,58
T1w.g lszm.ZonePercentage	1,16 (0,76; 1,78)	0,54 (0,06)	p = 0,49
T1w.g lszm.LowIntensityEmphasis	1,04 (0,70; 1,53)	0,47 (0,06)	p = 0,85
T1w.g lszm.SizeZoneVariability	1,24 (0,81; 1,89)	0,55 (0,06)	p = 0,33
T1w.g lszm.IntensityVariability	1,17 (0,81; 1,69)	0,58 (0,06)	p = 0,40
T1w.g lszm.ZoneEntropy	0,90 (0,60; 1,33)	0,49 (0,06)	p = 0,59
T1w.g lszm.HighIntensityLargeAreaEmphasis	1,03 (0,72; 1,46)	0,44 (0,06)	p = 0,87

TABLEAU A.12 – Cohorte Radiomique 4Mod : Évaluation des indices locaux calculés en T2w pour la prédiction de la survie globale. Les indices présentant l'indice-C plus grand que 0,55 sont en bleu. Les indices présentant en plus le RH sans la valeur 1 dans son intervalle de confiance et un ratio de vraisemblance statistiquement significatif sont en vert.

	RH (95% Intervalle de confiance)	indice-C (erreur-type)	Ratio de vraisemblance
T2w.firstorder.InterquartileRange	1,02 (0,76; 1,38)	0,53 (0,06)	p = 0,89
T2w.firstorder.Skewness	1,55 (1,06; 2,25)	0,61 (0,06)	p = 0,02
T2w.firstorder.Uniformity	0,95 (0,67; 1,35)	0,53 (0,06)	p = 0,79
T2w.firstorder.MeanAbsoluteDeviation	1,04 (0,77; 1,41)	0,54 (0,06)	p = 0,81
T2w.firstorder.Energy	1,25 (0,84; 1,87)	0,53 (0,06)	p = 0,28
T2w.firstorder.RobustMeanAbsoluteDeviation	1,02 (0,76; 1,38)	0,53 (0,06)	p = 0,88
T2w.firstorder.Median	1,22 (0,81; 1,82)	0,53 (0,06)	p = 0,35
T2w.firstorder.TotalEnergy	1,25 (0,84; 1,87)	0,53 (0,06)	p = 0,28
T2w.firstorder.Maximum	1,29 (0,90; 1,85)	0,59 (0,06)	p = 0,18
T2w.firstorder.RootMeanSquared	1,25 (0,84; 1,86)	0,53 (0,06)	p = 0,28
T2w.firstorder.90Percentile	1,23 (0,84; 1,78)	0,56 (0,06)	p = 0,29
T2w.firstorder.Minimum	1,28 (0,85; 1,95)	0,57 (0,06)	p = 0,24
T2w.firstorder.Entropy	1,06 (0,76; 1,49)	0,53 (0,06)	p = 0,73
T2w.firstorder.StandardDeviation	1,05 (0,77; 1,43)	0,53 (0,06)	p = 0,77
T2w.firstorder.Range	1,09 (0,79; 1,52)	0,55 (0,06)	p = 0,59
T2w.firstorder.Variance	1,04 (0,77; 1,39)	0,53 (0,06)	p = 0,81
T2w.firstorder.10Percentile	1,27 (0,85; 1,91)	0,54 (0,06)	p = 0,25
T2w.firstorder.Kurtosis	1,29 (0,87; 1,90)	0,57 (0,06)	p = 0,22
T2w.firstorder.Mean	1,25 (0,84; 1,86)	0,53 (0,06)	p = 0,28
T2w.glcmm.SumVariance	1,00 (0,72; 1,41)	0,53 (0,06)	p = 0,98
T2w.glcmm.Homogeneity1	0,91 (0,65; 1,27)	0,56 (0,06)	p = 0,58
T2w.glcmm.Homogeneity2	0,90 (0,65; 1,26)	0,56 (0,06)	p = 0,55
T2w.glcmm.ClusterShade	2,24 (1,38; 3,63)	0,60 (0,06)	p < 0,01
T2w.glcmm.MaximumProbability	0,92 (0,64; 1,30)	0,56 (0,06)	p = 0,62
T2w.glcmm.Idmn	1,05 (0,76; 1,43)	0,53 (0,06)	p = 0,78
T2w.glcmm.SumVariance2	1,04 (0,78; 1,39)	0,54 (0,06)	p = 0,80
T2w.glcmm.Contrast	1,04 (0,76; 1,42)	0,56 (0,06)	p = 0,81
T2w.glcmm.DifferenceEntropy	1,07 (0,78; 1,48)	0,55 (0,06)	p = 0,67
T2w.glcmm.InverseVariance	0,90 (0,65; 1,25)	0,55 (0,06)	p = 0,54
T2w.glcmm.Entropy	1,08 (0,77; 1,52)	0,54 (0,06)	p = 0,66
T2w.glcmm.Dissimilarity	1,06 (0,77; 1,46)	0,55 (0,06)	p = 0,73
T2w.glcmm.DifferenceVariance	1,04 (0,76; 1,42)	0,56 (0,06)	p = 0,83
T2w.glcmm.Idn	1,03 (0,73; 1,45)	0,52 (0,06)	p = 0,87
T2w.glcmm.Idm	0,90 (0,65; 1,26)	0,56 (0,06)	p = 0,55
T2w.glcmm.Correlation	0,93 (0,66; 1,31)	0,51 (0,06)	p = 0,67
T2w.glcmm.Autocorrelation	1,00 (0,72; 1,41)	0,53 (0,06)	p = 0,98
T2w.glcmm.SumEntropy	1,05 (0,75; 1,48)	0,52 (0,06)	p = 0,77
T2w.glcmm.AverageIntensity	0,99 (0,71; 1,39)	0,48 (0,06)	p = 0,97
T2w.glcmm.Energy	0,94 (0,67; 1,33)	0,55 (0,06)	p = 0,73
T2w.glcmm.SumSquares	1,04 (0,78; 1,40)	0,53 (0,06)	p = 0,78
T2w.glcmm.ClusterProminence	1,05 (0,79; 1,39)	0,55 (0,06)	p = 0,75
T2w.glcmm.SumAverage	1,00 (0,71; 1,40)	0,47 (0,06)	p = 0,98
T2w.glcmm.lmc2	0,99 (0,7; 1,40)	0,49 (0,06)	p = 0,96
T2w.glcmm.lmc1	0,98 (0,71; 1,37)	0,51 (0,06)	p = 0,92
T2w.glcmm.DifferenceAverage	1,06 (0,77; 1,46)	0,55 (0,06)	p = 0,73
T2w.glcmm.Id	0,91 (0,65; 1,27)	0,56 (0,06)	p = 0,58
T2w.glcmm.ClusterTendency	1,04 (0,78; 1,39)	0,54 (0,06)	p = 0,80
T2w.glrmm.ShortRunLowGrayLevelEmphasis	1,04 (0,74; 1,46)	0,50 (0,06)	p = 0,83
T2w.glrmm.GrayLevelVariance	1,04 (0,77; 1,39)	0,53 (0,06)	p = 0,80
T2w.glrmm.LowGrayLevelRunEmphasis	1,04 (0,74; 1,46)	0,50 (0,06)	p = 0,84
T2w.glrmm.GrayLevelNonUniformityNormalized	0,96 (0,68; 1,35)	0,53 (0,06)	p = 0,80
T2w.glrmm.RunVariance	0,89 (0,64; 1,25)	0,57 (0,06)	p = 0,51
T2w.glrmm.GrayLevelNonUniformity	0,96 (0,68; 1,36)	0,53 (0,06)	p = 0,82
T2w.glrmm.LongRunEmphasis	0,89 (0,64; 1,25)	0,57 (0,06)	p = 0,50
T2w.glrmm.ShortRunHighGrayLevelEmphasis	1,01 (0,72; 1,41)	0,54 (0,06)	p = 0,96
T2w.glrmm.RunLengthNonUniformity	1,13 (0,81; 1,58)	0,56 (0,06)	p = 0,49
T2w.glrmm.ShortRunEmphasis	1,13 (0,81; 1,58)	0,56 (0,06)	p = 0,48
T2w.glrmm.LongRunHighGrayLevelEmphasis	0,99 (0,71; 1,39)	0,48 (0,06)	p = 0,97
T2w.glrmm.RunPercentage	1,12 (0,80; 1,57)	0,56 (0,06)	p = 0,49
T2w.glrmm.LongRunLowGrayLevelEmphasis	1,02 (0,73; 1,43)	0,49 (0,06)	p = 0,89
T2w.glrmm.RunEntropy	1,05 (0,75; 1,47)	0,52 (0,06)	p = 0,79
T2w.glrmm.HighGrayLevelRunEmphasis	1,01 (0,72; 1,41)	0,53 (0,06)	p = 0,97
T2w.glrmm.RunLengthNonUniformityNormalized	1,13 (0,81; 1,58)	0,56 (0,06)	p = 0,48
T2w.glszm.GrayLevelVariance	1,05 (0,78; 1,41)	0,54 (0,06)	p = 0,77
T2w.glszm.LowIntensityLargeAreaEmphasis	0,98 (0,70; 1,36)	0,55 (0,06)	p = 0,89
T2w.glszm.HighIntensitySmallAreaEmphasis	1,03 (0,74; 1,44)	0,54 (0,06)	p = 0,86
T2w.glszm.SmallAreaEmphasis	1,08 (0,79; 1,48)	0,55 (0,06)	p = 0,64
T2w.glszm.LargeAreaEmphasis	0,92 (0,66; 1,29)	0,56 (0,06)	p = 0,62
T2w.glszm.ZoneVariance	0,93 (0,66; 1,29)	0,57 (0,06)	p = 0,64
T2w.glszm.SizeZoneVariabilityNormalized	1,07 (0,78; 1,47)	0,54 (0,06)	p = 0,68
T2w.glszm.LowIntensitySmallAreaEmphasis	1,02 (0,72; 1,44)	0,47 (0,06)	p = 0,92
T2w.glszm.HighIntensityEmphasis	1,02 (0,73; 1,43)	0,53 (0,06)	p = 0,90
T2w.glszm.IntensityVariabilityNormalized	0,97 (0,69; 1,35)	0,52 (0,06)	p = 0,84
T2w.glszm.ZonePercentage	1,11 (0,80; 1,54)	0,56 (0,06)	p = 0,55
T2w.glszm.LowIntensityEmphasis	1,00 (0,71; 1,40)	0,53 (0,06)	p = 0,99
T2w.glszm.SizeZoneVariability	1,08 (0,78; 1,50)	0,55 (0,06)	p = 0,63
T2w.glszm.IntensityVariability	1,21 (0,86; 1,69)	0,60 (0,06)	p = 0,26
T2w.glszm.ZoneEntropy	1,06 (0,76; 1,48)	0,54 (0,06)	p = 0,74
T2w.glszm.HighIntensityLargeAreaEmphasis	0,73 (0,53; 1,02)	0,62 (0,06)	p = 0,06

TABLEAU A.13 – Cohorte Radiomique 4Mod : Évaluation des indices locaux calculés en T1c pour la prédiction de la survie globale. Les indices présentant l'indice-C plus grand que 0,55 sont en bleu.

	RH (95% Intervalle de confiance)	indice-C (erreur-type)	Ratio de vraisemblance
T1c.firstorder.InterquartileRange	1,13 (0,81 ; 1,59)	0,56 (0,06)	p = 0,49
T1c.firstorder.Skewness	1,12 (0,79 ; 1,59)	0,53 (0,06)	p = 0,53
T1c.firstorder.Uniformity	0,87 (0,63 ; 1,22)	0,56 (0,06)	p = 0,42
T1c.firstorder.MeanAbsoluteDeviation	1,12 (0,81 ; 1,56)	0,56 (0,06)	p = 0,50
T1c.firstorder.Energy	1,25 (0,90 ; 1,75)	0,53 (0,06)	p = 0,19
T1c.firstorder.RobustMeanAbsoluteDeviation	1,13 (0,80 ; 1,58)	0,56 (0,06)	p = 0,50
T1c.firstorder.Median	1,23 (0,89 ; 1,71)	0,54 (0,06)	p = 0,22
T1c.firstorder.TotalEnergy	1,25 (0,90 ; 1,75)	0,53 (0,06)	p = 0,19
T1c.firstorder.Maximum	1,20 (0,87 ; 1,65)	0,57 (0,06)	p = 0,28
T1c.firstorder.RootMeanSquared	1,25 (0,90 ; 1,75)	0,53 (0,06)	p = 0,18
T1c.firstorder.90Percentile	1,23 (0,89 ; 1,70)	0,54 (0,06)	p = 0,23
T1c.firstorder.Minimum	1,06 (0,78 ; 1,44)	0,49 (0,06)	p = 0,70
T1c.firstorder.Entropy	1,16 (0,83 ; 1,61)	0,56 (0,06)	p = 0,39
T1c.firstorder.StandardDeviation	1,12 (0,81 ; 1,55)	0,56 (0,06)	p = 0,49
T1c.firstorder.Range	1,14 (0,83 ; 1,56)	0,58 (0,06)	p = 0,44
T1c.firstorder.Variance	1,10 (0,79 ; 1,53)	0,56 (0,06)	p = 0,60
T1c.firstorder.10Percentile	1,13 (0,83 ; 1,53)	0,52 (0,06)	p = 0,45
T1c.firstorder.Kurtosis	1,18 (0,83 ; 1,67)	0,54 (0,06)	p = 0,38
T1c.firstorder.Mean	1,25 (0,90 ; 1,75)	0,53 (0,06)	p = 0,19
T1c.glcM.SumVariance	1,12 (0,80 ; 1,55)	0,57 (0,06)	p = 0,53
T1c.glcM.Homogeneity1	0,87 (0,62 ; 1,21)	0,57 (0,06)	p = 0,42
T1c.glcM.Homogeneity2	0,87 (0,62 ; 1,22)	0,56 (0,06)	p = 0,43
T1c.glcM.ClusterShade	1,03 (0,75 ; 1,42)	0,54 (0,06)	p = 0,85
T1c.glcM.MaximumProbability	0,89 (0,64 ; 1,24)	0,56 (0,06)	p = 0,50
T1c.glcM.Idmn	1,01 (0,68 ; 1,51)	0,51 (0,06)	p = 0,96
T1c.glcM.SumVariance2	1,09 (0,78 ; 1,52)	0,56 (0,06)	p = 0,62
T1c.glcM.Contrast	1,15 (0,81 ; 1,62)	0,57 (0,06)	p = 0,46
T1c.glcM.DifferenceEntropy	1,16 (0,83 ; 1,61)	0,57 (0,06)	p = 0,40
T1c.glcM.InverseVariance	0,87 (0,62 ; 1,22)	0,57 (0,06)	p = 0,43
T1c.glcM.Entropy	1,16 (0,83 ; 1,62)	0,57 (0,06)	p = 0,40
T1c.glcM.Dissimilarity	1,16 (0,83 ; 1,63)	0,57 (0,06)	p = 0,41
T1c.glcM.DifferenceVariance	1,12 (0,80 ; 1,57)	0,58 (0,06)	p = 0,51
T1c.glcM.Idn	1,10 (0,74 ; 1,62)	0,52 (0,06)	p = 0,63
T1c.glcM.Idm	0,87 (0,62 ; 1,22)	0,56 (0,06)	p = 0,43
T1c.glcM.Correlation	1,07 (0,75 ; 1,52)	0,51 (0,06)	p = 0,72
T1c.glcM.Autocorrelation	1,12 (0,81 ; 1,55)	0,57 (0,06)	p = 0,52
T1c.glcM.SumEntropy	1,15 (0,83 ; 1,61)	0,56 (0,06)	p = 0,40
T1c.glcM.AverageIntensity	1,15 (0,83 ; 1,60)	0,57 (0,06)	p = 0,41
T1c.glcM.Energy	0,89 (0,64 ; 1,23)	0,55 (0,06)	p = 0,47
T1c.glcM.SumSquares	1,10 (0,78 ; 1,53)	0,56 (0,06)	p = 0,60
T1c.glcM.ClusterProminence	1,03 (0,74 ; 1,44)	0,57 (0,06)	p = 0,85
T1c.glcM.SumAverage	1,16 (0,84 ; 1,61)	0,57 (0,06)	p = 0,39
T1c.glcM.Imc2	1,16 (0,83 ; 1,63)	0,56 (0,06)	p = 0,38
T1c.glcM.Imc1	0,87 (0,63 ; 1,20)	0,57 (0,06)	p = 0,41
T1c.glcM.DifferenceAverage	1,16 (0,83 ; 1,63)	0,57 (0,06)	p = 0,41
T1c.glcM.Id	0,87 (0,62 ; 1,21)	0,57 (0,06)	p = 0,42
T1c.glcM.ClusterTendency	1,09 (0,78 ; 1,52)	0,56 (0,06)	p = 0,62
T1c.glrM.ShortRunLowGrayLevelEmphasis	0,84 (0,59 ; 1,19)	0,58 (0,06)	p = 0,32
T1c.glrM.GrayLevelVariance	1,10 (0,79 ; 1,53)	0,57 (0,06)	p = 0,60
T1c.glrM.LowGrayLevelRunEmphasis	0,84 (0,59 ; 1,19)	0,58 (0,06)	p = 0,32
T1c.glrM.GrayLevelNonUniformityNormalized	0,87 (0,63 ; 1,22)	0,56 (0,06)	p = 0,42
T1c.glrM.RunVariance	0,87 (0,62 ; 1,21)	0,57 (0,06)	p = 0,40
T1c.glrM.GrayLevelNonUniformity	0,87 (0,62 ; 1,22)	0,56 (0,06)	p = 0,42
T1c.glrM.LongRunEmphasis	0,87 (0,62 ; 1,21)	0,56 (0,06)	p = 0,40
T1c.glrM.ShortRunHighGrayLevelEmphasis	1,12 (0,81 ; 1,56)	0,57 (0,06)	p = 0,51
T1c.glrM.RunLengthNonUniformity	1,16 (0,83 ; 1,63)	0,57 (0,06)	p = 0,39
T1c.glrM.ShortRunEmphasis	1,16 (0,83 ; 1,62)	0,57 (0,06)	p = 0,39
T1c.glrM.LongRunHighGrayLevelEmphasis	1,12 (0,81 ; 1,56)	0,57 (0,06)	p = 0,51
T1c.glrM.RunPercentage	1,16 (0,83 ; 1,62)	0,57 (0,06)	p = 0,39
T1c.glrM.LongRunLowGrayLevelEmphasis	0,85 (0,60 ; 1,20)	0,58 (0,06)	p = 0,34
T1c.glrM.RunEntropy	1,16 (0,83 ; 1,61)	0,57 (0,06)	p = 0,40
T1c.glrM.HighGrayLevelRunEmphasis	1,12 (0,81 ; 1,56)	0,58 (0,06)	p = 0,51
T1c.glrM.RunLengthNonUniformityNormalized	1,16 (0,83 ; 1,63)	0,57 (0,06)	p = 0,39
T1c.glszm.GrayLevelVariance	1,09 (0,79 ; 1,51)	0,56 (0,06)	p = 0,61
T1c.glszm.LowIntensityLargeAreaEmphasis	0,89 (0,65 ; 1,23)	0,57 (0,06)	p = 0,46
T1c.glszm.HighIntensitySmallAreaEmphasis	1,11 (0,80 ; 1,54)	0,57 (0,06)	p = 0,55
T1c.glszm.SmallAreaEmphasis	1,17 (0,83 ; 1,63)	0,55 (0,06)	p = 0,38
T1c.glszm.LargeAreaEmphasis	0,90 (0,65 ; 1,23)	0,56 (0,06)	p = 0,49
T1c.glszm.ZoneVariance	0,90 (0,65 ; 1,24)	0,56 (0,06)	p = 0,50
T1c.glszm.SizeZoneVariabilityNormalized	1,17 (0,84 ; 1,63)	0,55 (0,06)	p = 0,37
T1c.glszm.LowIntensitySmallAreaEmphasis	0,86 (0,61 ; 1,22)	0,57 (0,06)	p = 0,38
T1c.glszm.HighIntensityEmphasis	1,11 (0,80 ; 1,54)	0,58 (0,06)	p = 0,54
T1c.glszm.IntensityVariabilityNormalized	0,86 (0,62 ; 1,20)	0,57 (0,06)	p = 0,38
T1c.glszm.ZonePercentage	1,18 (0,84 ; 1,65)	0,57 (0,06)	p = 0,35
T1c.glszm.LowIntensityEmphasis	0,86 (0,61 ; 1,21)	0,57 (0,06)	p = 0,37
T1c.glszm.SizeZoneVariability	1,18 (0,84 ; 1,65)	0,56 (0,06)	p = 0,36
T1c.glszm.IntensityVariability	0,99 (0,68 ; 1,45)	0,45 (0,06)	p = 0,97
T1c.glszm.ZoneEntropy	1,15 (0,83 ; 1,59)	0,56 (0,06)	p = 0,42
T1c.glszm.HighIntensityLargeAreaEmphasis	0,93 (0,67 ; 1,30)	0,52 (0,06)	p = 0,68

TABLEAU A.14 – Cohorte Radiomique 4Mod : Évaluation des indices locaux calculés en FLAIR pour la prédiction de la survie globale. Les indices présentant l'indice-C plus grand que 0,55 sont en bleu. Les indices présentant en plus le RH sans la valeur 1 dans son intervalle de confiance et un ratio de vraisemblance statistiquement significatif sont en vert.

	RH (95% Intervalle de confiance)	indice-C (erreur-type)	Ratio de vraisemblance
FLAIR.firstorder.InterquartileRange	0,80 (0,54; 1,19)	0,60 (0,06)	p = 0,25
FLAIR.firstorder.Skewness	1,46 (1,03; 2,05)	0,61 (0,06)	p = 0,04
FLAIR.firstorder.Uniformity	1,36 (0,94; 1,97)	0,58 (0,06)	p = 0,10
FLAIR.firstorder.MeanAbsoluteDeviation	0,81 (0,55; 1,21)	0,59 (0,06)	p = 0,30
FLAIR.firstorder.Energy	0,87 (0,61; 1,25)	0,53 (0,06)	p = 0,46
FLAIR.firstorder.RobustMeanAbsoluteDeviation	0,80 (0,54; 1,20)	0,60 (0,06)	p = 0,26
FLAIR.firstorder.Median	0,86 (0,60; 1,23)	0,54 (0,06)	p = 0,40
FLAIR.firstorder.TotalEnergy	0,87 (0,61; 1,25)	0,53 (0,06)	p = 0,46
FLAIR.firstorder.Maximum	0,89 (0,60; 1,32)	0,54 (0,06)	p = 0,55
FLAIR.firstorder.RootMeanSquared	0,87 (0,61; 1,25)	0,53 (0,06)	p = 0,46
FLAIR.firstorder.90Percentile	0,85 (0,58; 1,24)	0,54 (0,06)	p = 0,39
FLAIR.firstorder.Minimum	1,02 (0,74; 1,40)	0,53 (0,06)	p = 0,92
FLAIR.firstorder.Entropy	0,76 (0,52; 1,13)	0,59 (0,06)	p = 0,17
FLAIR.firstorder.StandardDeviation	0,82 (0,55; 1,23)	0,59 (0,06)	p = 0,33
FLAIR.firstorder.Range	0,86 (0,57; 1,30)	0,58 (0,06)	p = 0,46
FLAIR.firstorder.Variance	0,90 (0,60; 1,34)	0,58 (0,06)	p = 0,58
FLAIR.firstorder.10Percentile	0,94 (0,67; 1,32)	0,51 (0,06)	p = 0,72
FLAIR.firstorder.Kurtosis	1,16 (0,78; 1,73)	0,54 (0,06)	p = 0,47
FLAIR.firstorder.Mean	0,87 (0,61; 1,25)	0,53 (0,06)	p = 0,46
FLAIR.g lcm.SumVariance	0,85 (0,56; 1,30)	0,59 (0,06)	p = 0,44
FLAIR.g lcm.Homogeneity1	1,27 (0,87; 1,84)	0,56 (0,06)	p = 0,21
FLAIR.g lcm.Homogeneity2	1,28 (0,88; 1,85)	0,57 (0,06)	p = 0,19
FLAIR.g lcm.ClusterShade	1,68 (1,05; 2,69)	0,61 (0,06)	p = 0,04
FLAIR.g lcm.MaximumProbability	1,35 (0,95; 1,92)	0,56 (0,06)	p = 0,11
FLAIR.g lcm.Idmn	0,83 (0,55; 1,24)	0,56 (0,06)	p = 0,36
FLAIR.g lcm.SumVariance2	0,90 (0,60; 1,35)	0,59 (0,06)	p = 0,60
FLAIR.g lcm.Contrast	0,93 (0,63; 1,39)	0,57 (0,06)	p = 0,72
FLAIR.g lcm.DifferenceEntropy	0,82 (0,56; 1,20)	0,56 (0,06)	p = 0,29
FLAIR.g lcm.InverseVariance	1,26 (0,87; 1,82)	0,56 (0,06)	p = 0,23
FLAIR.g lcm.Entropy	0,76 (0,52; 1,11)	0,57 (0,06)	p = 0,15
FLAIR.g lcm.Dissimilarity	0,86 (0,58; 1,27)	0,57 (0,06)	p = 0,42
FLAIR.g lcm.DifferenceVariance	0,96 (0,65; 1,43)	0,56 (0,06)	p = 0,84
FLAIR.g lcm.Idn	0,86 (0,57; 1,30)	0,55 (0,06)	p = 0,47
FLAIR.g lcm.Idm	1,28 (0,88; 1,85)	0,57 (0,06)	p = 0,19
FLAIR.g lcm.Correlation	0,72 (0,48; 1,09)	0,61 (0,06)	p = 0,11
FLAIR.g lcm.Autocorrelation	0,84 (0,55; 1,29)	0,59 (0,06)	p = 0,41
FLAIR.g lcm.SumEntropy	0,76 (0,51; 1,12)	0,59 (0,06)	p = 0,16
FLAIR.g lcm.AverageIntensity	0,78 (0,52; 1,17)	0,60 (0,06)	p = 0,22
FLAIR.g lcm.Energy	1,35 (0,95; 1,92)	0,57 (0,06)	p = 0,10
FLAIR.g lcm.SumSquares	0,91 (0,60; 1,36)	0,58 (0,06)	p = 0,63
FLAIR.g lcm.ClusterProminence	1,10 (0,75; 1,64)	0,42 (0,06)	p = 0,63
FLAIR.g lcm.SumAverage	0,78 (0,52; 1,17)	0,60 (0,06)	p = 0,21
FLAIR.g lcm.Imc2	0,71 (0,48; 1,05)	0,60 (0,06)	p = 0,09
FLAIR.g lcm.Imc1	1,29 (0,86; 1,95)	0,60 (0,06)	p = 0,21
FLAIR.g lcm.DifferenceAverage	0,86 (0,58; 1,27)	0,57 (0,06)	p = 0,42
FLAIR.g lcm.Id	1,27 (0,87; 1,84)	0,56 (0,06)	p = 0,21
FLAIR.g lcm.ClusterTendency	0,90 (0,60; 1,35)	0,59 (0,06)	p = 0,60
FLAIR.g lrlm.ShortRunLowGrayLevelEmphasis	1,45 (0,99; 2,11)	0,58 (0,06)	p = 0,06
FLAIR.g lrlm.GrayLevelVariance	0,90 (0,60; 1,35)	0,58 (0,06)	p = 0,59
FLAIR.g lrlm.LowGrayLevelRunEmphasis	1,45 (0,99; 2,11)	0,59 (0,06)	p = 0,06
FLAIR.g lrlm.GrayLevelNonUniformityNormalized	1,36 (0,94; 1,97)	0,58 (0,06)	p = 0,10
FLAIR.g lrlm.RunVariance	1,30 (0,91; 1,85)	0,57 (0,06)	p = 0,16
FLAIR.g lrlm.GrayLevelNonUniformity	1,36 (0,94; 1,98)	0,58 (0,06)	p = 0,11
FLAIR.g lrlm.LongRunEmphasis	1,30 (0,91; 1,85)	0,56 (0,06)	p = 0,16
FLAIR.g lrlm.ShortRunHighGrayLevelEmphasis	0,85 (0,56; 1,29)	0,59 (0,06)	p = 0,43
FLAIR.g lrlm.RunLengthNonUniformity	0,78 (0,54; 1,12)	0,57 (0,06)	p = 0,17
FLAIR.g lrlm.ShortRunEmphasis	0,77 (0,54; 1,11)	0,57 (0,06)	p = 0,16
FLAIR.g lrlm.LongRunHighGrayLevelEmphasis	0,83 (0,55; 1,27)	0,60 (0,06)	p = 0,37
FLAIR.g lrlm.RunPercentage	0,77 (0,54; 1,11)	0,57 (0,06)	p = 0,17
FLAIR.g lrlm.LongRunLowGrayLevelEmphasis	1,44 (1,00; 2,09)	0,59 (0,06)	p = 0,06
FLAIR.g lrlm.RunEntropy	0,77 (0,52; 1,14)	0,59 (0,06)	p = 0,18
FLAIR.g lrlm.HighGrayLevelRunEmphasis	0,85 (0,56; 1,29)	0,59 (0,06)	p = 0,41
FLAIR.g lrlm.RunLengthNonUniformityNormalized	0,77 (0,54; 1,11)	0,57 (0,06)	p = 0,17
FLAIR.g lszm.GrayLevelVariance	0,91 (0,60; 1,37)	0,58 (0,06)	p = 0,64
FLAIR.g lszm.LowIntensityLargeAreaEmphasis	1,37 (0,99; 1,91)	0,57 (0,06)	p = 0,08
FLAIR.g lszm.HighIntensitySmallAreaEmphasis	0,89 (0,59; 1,35)	0,57 (0,06)	p = 0,57
FLAIR.g lszm.SmallAreaEmphasis	0,93 (0,66; 1,32)	0,51 (0,06)	p = 0,69
FLAIR.g lszm.LargeAreaEmphasis	1,33 (0,96; 1,85)	0,57 (0,06)	p = 0,11
FLAIR.g lszm.ZoneVariance	1,34 (0,96; 1,86)	0,56 (0,06)	p = 0,10
FLAIR.g lszm.SizeZoneVariabilityNormalized	0,94 (0,67; 1,33)	0,50 (0,06)	p = 0,73
FLAIR.g lszm.LowIntensitySmallAreaEmphasis	1,43 (1,00; 2,06)	0,60 (0,06)	p = 0,06
FLAIR.g lszm.HighIntensityEmphasis	0,86 (0,57; 1,31)	0,59 (0,06)	p = 0,47
FLAIR.g lszm.IntensityVariabilityNormalized	1,33 (0,92; 1,91)	0,58 (0,06)	p = 0,13
FLAIR.g lszm.ZonePercentage	0,80 (0,55; 1,17)	0,56 (0,06)	p = 0,24
FLAIR.g lszm.LowIntensityEmphasis	1,42 (0,99; 2,04)	0,58 (0,06)	p = 0,06
FLAIR.g lszm.SizeZoneVariability	0,86 (0,59; 1,25)	0,56 (0,06)	p = 0,42
FLAIR.g lszm.IntensityVariability	0,86 (0,60; 1,24)	0,53 (0,06)	p = 0,43
FLAIR.g lszm.ZoneEntropy	0,75 (0,52; 1,09)	0,58 (0,06)	p = 0,13
FLAIR.g lszm.HighIntensityLargeAreaEmphasis	1,09 (0,76; 1,55)	0,50 (0,06)	p = 0,66

A.2 Documents annexes

A.2.1 Extrait de la documentation Pyradiomics v.1.1.1

Radiomic Features

This section contains the definitions of the various features that can be extracted using PyRadiomics. They are subdivided into the following classes:

- *First Order Features* (19 features)
- *Shape Features* (13 features)
- *Gray Level Co-occurrence Matrix (GLCM) Features* (28 features)
- *Gray Level Size Zone Matrix (GLSZM) Features* (16 features)
- *Gray Level Run Length Matrix (GLRLM) Features* (16 features)

All feature classes, with the exception of shape can be calculated on either the original image and/or a derived image, obtained by applying one of several filters. The shape descriptors are independent of gray value, and are extracted from the label mask. If enabled, they are calculated separately of enabled input image types, and listed in the result as if calculated on the original image.

First Order Features

```
class radiomics.firstorder.RadiomicsFirstOrder (inputImage, inputMask, **kwargs)
    Bases: radiomics.base.RadiomicsFeaturesBase
```

First-order statistics describe the distribution of voxel intensities within the image region defined by the mask through commonly used and basic metrics.

Let:

\mathbf{X} denote the three dimensional image matrix with N voxels

$\mathbf{P}(i)$ the first order histogram with N_I discrete intensity levels, where N_I is defined by the number of levels, calculated based on the `binWidth` parameter.

$p(i)$ be the normalized first order histogram and equal to $\frac{\mathbf{P}(i)}{\sum \mathbf{P}(i)}$

Following additional settings are possible:

- `voxelArrayShift [0]`: Integer, This amount is added to the gray level intensity in features Energy, Total Energy and RMS, this is to prevent negative values. ___If using CT data, or data normalized with mean 0, consider setting this parameter to a fixed value (e.g. 2000) that ensures non-negative numbers in the image. Bear in mind however, that the larger the value, the larger the volume confounding effect will be. ___

getEnergyFeatureValue ()

Calculate the Energy of the image array.

$$energy = \sum_{i=1}^N (\mathbf{X}(i) + c)^2$$

Here, c is optional value, defined by `voxelArrayShift`, which shifts the intensities to prevent negative values in \mathbf{X} . This ensures that voxels with the lowest gray values contribute the least to Energy, instead of voxels with gray level intensity closest to 0.

Energy is a measure of the magnitude of voxel values in an image. A larger values implies a greater sum of the squares of these values. This feature is volume-confounded, a larger value of c increases the effect of volume-confounding.

getTotalEnergyFeatureValue ()

Calculate the Total Energy of the image array.

$$total\ energy = V_{voxel} \sum_{i=1}^N (\mathbf{X}(i) + c)^2$$

Here, c is optional value, defined by `voxelArrayShift`, which shifts the intensities to prevent negative values in \mathbf{X} . This ensures that voxels with the lowest gray values contribute the least to Energy, instead of voxels with gray level intensity closest to 0.

Total Energy is the value of Energy feature scaled by the volume of the voxel in cubic mm. This feature is volume-confounded, a larger value of c increases the effect of volume-confounding.

getEntropyFeatureValue ()

Calculate the Entropy of the image array.

$$entropy = - \sum_{i=1}^{N_I} p(i) \log_2 (p(i) + \epsilon)$$

Here, ϵ is an arbitrarily small positive number ($\approx 2.2 \times 10^{-16}$).

Entropy specifies the uncertainty/randomness in the image values. It measures the average amount of information required to encode the image values.

getMinimumFeatureValue ()

Calculate the Minimum Value in the image array.

$$minimum = \min(\mathbf{X})$$

get10PercentileFeatureValue ()

Calculate the 10th percentile in the image array.

get90PercentileFeatureValue ()

Calculate the 90th percentile in the image array.

pyradiomics Documentation, Release 1.1.1

getMaximumFeatureValue ()

Calculate the Maximum Value in the image array.

$$maximum = \max(\mathbf{X})$$

getMeanFeatureValue ()

Calculate the Mean Value for the image array.

$$mean = \frac{1}{N} \sum_{i=1}^N \mathbf{X}(i)$$

getMedianFeatureValue ()

Calculate the Median Value for the image array.

getInterquartileRangeFeatureValue ()

Calculate the interquartile range of the image array.

interquartile range = $\mathbf{P}_{75} - \mathbf{P}_{25}$, where \mathbf{P}_{25} and \mathbf{P}_{75} are the 25th and 75th percentile of the image array, respectively.

getRangeFeatureValue ()

Calculate the Range of Values in the image array.

$$range = \max(\mathbf{X}) - \min(\mathbf{X})$$

getMeanAbsoluteDeviationFeatureValue ()

Calculate the Mean Absolute Deviation for the image array.

$$mean\ absolute\ deviation = \frac{1}{N} \sum_{i=1}^N |\mathbf{X}(i) - \bar{X}|$$

Mean Absolute Deviation is the mean distance of all intensity values from the Mean Value of the image array.

getRobustMeanAbsoluteDeviationFeatureValue ()

Calculate the Robust Mean Absolute Deviation for the image array.

$$robust\ mean\ absolute\ deviation = \frac{1}{N_{10-90}} \sum_{i=1}^{N_{10-90}} |\mathbf{X}_{10-90}(i) - \bar{X}_{10-90}|$$

Robust Mean Absolute Deviation is the mean distance of all intensity values from the Mean Value calculated on the subset of image array with gray levels in between, or equal to the 10th and 90th percentile.

getRootMeanSquaredFeatureValue ()

Calculate the Root Mean Squared of the image array.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (\mathbf{X}(i) + c)^2}$$

Here, c is optional value, defined by `voxelArrayShift`, which shifts the intensities to prevent negative values in \mathbf{X} . This ensures that voxels with the lowest gray values contribute the least to RMS, instead of voxels with gray level intensity closest to 0.

RMS is the square-root of the mean of all the squared intensity values. It is another measure of the magnitude of the image values. This feature is volume-confounded, a larger value of c increases the effect of volume-confounding.

getStandardDeviationFeatureValue ()

Calculate the Standard Deviation of the image array.

$$standard\ deviation = \sqrt{\frac{1}{N} \sum_{i=1}^N (\mathbf{X}(i) - \bar{X})^2}$$

Standard Deviation measures the amount of variation or dispersion from the Mean Value.

getSkewnessFeatureValue (*axis=0*)

Calculate the Skewness of the image array.

$$skewness = \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N} \sum_{i=1}^N (\mathbf{X}(i) - \bar{X})^3}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (\mathbf{X}(i) - \bar{X})^2} \right)^3}$$

Where μ_3 is the 3rd central moment.

Skewness measures the asymmetry of the distribution of values about the Mean value. Depending on where the tail is elongated and the mass of the distribution is concentrated, this value can be positive or negative.

Related links:

<https://en.wikipedia.org/wiki/Skewness>

getKurtosisFeatureValue (*axis=0*)

Calculate the Kurtosis of the image array.

$$kurtosis = \frac{\mu_4}{\sigma^4} = \frac{\frac{1}{N} \sum_{i=1}^N (\mathbf{X}(i) - \bar{X})^4}{\left(\frac{1}{N} \sum_{i=1}^N (\mathbf{X}(i) - \bar{X})^2 \right)^2}$$

Where μ_4 is the 4th central moment.

Kurtosis is a measure of the ‘peakedness’ of the distribution of values in the image ROI. A higher kurtosis implies that the mass of the distribution is concentrated towards the tail(s) rather than towards the mean. A lower kurtosis implies the reverse: that the mass of the distribution is concentrated towards a spike near the Mean value.

Related links:

<https://en.wikipedia.org/wiki/Kurtosis>

getVarianceFeatureValue ()

Calculate the Variance in the image array.

$$variance = \sigma^2 = \frac{1}{N} \sum_{i=1}^N (\mathbf{X}(i) - \bar{X})^2$$

Variance is the the mean of the squared distances of each intensity value from the Mean value. This is a measure of the spread of the distribution about the mean.

getUniformityFeatureValue ()

Calculate the Uniformity of the image array.

$$uniformity = \sum_{i=1}^{N_i} p(i)^2$$

Uniformity is a measure of the sum of the squares of each intensity value. This is a measure of the heterogeneity of the image array, where a greater uniformity implies a greater heterogeneity or a greater range of discrete intensity values.

Shape Features

class radiomics.shape.**RadiomicsShape** (*inputImage, inputMask, **kwargs*)

Bases: *radiomics.base.RadiomicsFeaturesBase*

pyradiomics Documentation, Release 1.1.1

In this group of features we included descriptors of the three-dimensional size and shape of the tumor region. Let in the following definitions denote V the volume (mm^3) and A the surface area of the volume (mm^2) of interest.

getVolumeFeatureValue ()

Calculate the volume of the tumor region in cubic millimeters.

getSurfaceAreaFeatureValue ()

Calculate the surface area of the tumor region in square millimeters using a marching cubes algorithm.

$$A = \sum_{i=1}^N \frac{1}{2} |a_i b_i \times a_i c_i|$$

Where:

N is the number of triangles forming the surface mesh of the volume (ROI)

$a_i b_i$ and $a_i c_i$ are the edges of the i^{th} triangle formed by points a_i , b_i and c_i

getSurfaceVolumeRatioFeatureValue ()

Calculate the surface area to volume ratio of the tumor region

$$\text{surface to volume ratio} = \frac{A}{V}$$

Here, a lower value indicates a more compact (sphere-like) shape.

getSphericityFeatureValue ()

Calculate the Sphericity of the tumor region.

$$\text{sphericity} = \frac{\sqrt[3]{36\pi V^2}}{A}$$

Sphericity is a measure of the roundness of the shape of the tumor region relative to a sphere. It is a dimensionless measure, independent of scale and orientation. The value range is $0 < \text{sphericity} \leq 1$, where a value of 1 indicates a perfect sphere (a sphere has the smallest possible surface area for a given volume, compared to other solids).

Note: This feature is correlated to Compactness 1, Compactness 2 and Spherical Disproportion. In the default parameter file provided in the `pyradiomics\bin` folder, Compactness 1 and Compactness 2 are therefore disabled.

getCompactness1FeatureValue ()

Calculate the compactness (1) of the tumor region.

$$\text{compactness 1} = \frac{V}{\sqrt{\pi A^3}}$$

Similar to Sphericity, Compactness 1 is a measure of how compact the shape of the tumor is relative to a sphere (most compact). It is therefore correlated to Sphericity and redundant. It is provided here for completeness. The value range is $0 < \text{compactness 1} \leq \frac{1}{6\pi}$, where a value of $\frac{1}{6\pi}$ indicates a perfect sphere.

By definition, $\text{compactness 1} = \frac{1}{6\pi} \sqrt{\text{compactness 2}} = \frac{1}{6\pi} \sqrt{\text{sphericity}^3}$.

Note: This feature is correlated to Compactness 2, Sphericity and Spherical Disproportion. In the default parameter file provided in the `pyradiomics\bin` folder, Compactness 1 and Compactness 2 are therefore disabled.

getCompactness2FeatureValue ()

Calculate the Compactness (2) of the tumor region.

$$\text{compactness 2} = 36\pi \frac{V^2}{A^3}$$

Similar to Sphericity and Compactness 1, Compactness 2 is a measure of how compact the shape of the tumor is relative to a sphere (most compact). It is a dimensionless measure, independent of scale and orientation. The value range is $0 < \text{compactness 2} \leq 1$, where a value of 1 indicates a perfect sphere.

By definition, $\text{compactness 2} = (\text{sphericity})^3$

Note: This feature is correlated to Compactness 1, Sphericity and Spherical Disproportion. In the default parameter file provided in the `pyradiomics\bin` folder, Compactness 1 and Compactness 2 are therefore disabled.

getSphericalDisproportionFeatureValue ()

Calculate the Spherical Disproportion of the tumor region.

$$\text{spherical disproportion} = \frac{A}{4\pi R^2} = \frac{A}{\sqrt[3]{36\pi V^2}}$$

Where R is the radius of a sphere with the same volume as the tumor, and equal to $\sqrt[3]{\frac{3V}{4\pi}}$.

Spherical Disproportion is the ratio of the surface area of the tumor region to the surface area of a sphere with the same volume as the tumor region, and by definition, the inverse of Sphericity. Therefore, the value range is $\text{spherical disproportion} \geq 1$, with a value of 1 indicating a perfect sphere.

Note: This feature is correlated to Compactness 1, Sphericity and Spherical Disproportion. In the default parameter file provided in the `pyradiomics\bin` folder, Compactness 1 and Compactness 2 are therefore disabled.

getMaximum3DDiameterFeatureValue ()

Calculate the largest pairwise euclidean distance between tumor surface voxels. Also known as Feret Diameter.

getMaximum2DDiameterSliceFeatureValue ()

Calculate the largest pairwise euclidean distance between tumor surface voxels in the row-column plane.

getMaximum2DDiameterColumnFeatureValue ()

Calculate the largest pairwise euclidean distance between tumor surface voxels in the row-slice plane.

getMaximum2DDiameterRowFeatureValue ()

Calculate the largest pairwise euclidean distance between tumor surface voxels in the column-slice plane.

getElongationFeatureValue ()

Calculate the elongation of the shape, which is defined as:

$$\text{Elongation} = \frac{\lambda_{\text{longest}}}{\lambda_{\text{intermediate}}}$$

Here, λ_{longest} and $\lambda_{\text{intermediate}}$ are the lengths of the largest and second largest principal component axes.

References:

- Andrey P, Kieu K, Kress C, Lehmann G, Tirichine L, Liu Z, et al. Statistical analysis of 3D images detects regular spatial distributions of centromeres and chromocenters in animal and plant nuclei. PLoS Comput Biol. 2010;6:27.

getFlatnessFeatureValue ()

Calculate the flatness of the shape, which is defined as:

$$\text{Flatness} = \frac{\lambda_{\text{intermediate}}}{\lambda_{\text{shortest}}}$$

pyradiomics Documentation, Release 1.1.1

Here, $\lambda_{\text{intermediate}}$ and $\lambda_{\text{shortest}}$ are the lengths of the second largest and smallest principal component axes.

References:

- Andrey P, Kieu K, Kress C, Lehmann G, Tirichine L, Liu Z, et al. Statistical analysis of 3D images detects regular spatial distributions of centromeres and chromocenters in animal and plant nuclei. PLoS Comput Biol. 2010;6:27.

Gray Level Co-occurrence Matrix (GLCM) Features

class `radiomics.glm.RadiomicsGLCM` (*inputImage, inputMask, **kwargs*)

Bases: `radiomics.base.RadiomicsFeaturesBase`

A Gray Level Co-occurrence Matrix (GLCM) of size $N_g \times N_g$ describes the second-order joint probability function of an image region constrained by the mask and is defined as $\mathbf{P}(i, j|\delta, \alpha)$. The $(i, j)^{\text{th}}$ element of this matrix represents the number of times the combination of levels i and j occur in two pixels in the image, that are separated by a distance of δ pixels in direction α , and N_g is the number of discrete gray level intensities. The distance δ from the center voxel is defined as the distance according to the infinity norm. For $\delta = 1$, this assumes 26-connectivity in 3D and for $\delta = 2$ a 98-connectivity.

Note that pyradiomics by default computes symmetrical GLCM!

As a two dimensional example, let the following matrix \mathbf{I} represent a 5x5 image, having 5 discrete grey levels:

$$\mathbf{I} = \begin{bmatrix} 1 & 2 & 5 & 2 & 3 \\ 3 & 2 & 1 & 3 & 1 \\ 1 & 3 & 5 & 5 & 2 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 2 & 4 & 3 & 5 \end{bmatrix}$$

For distance $\delta = 1$ (considering pixels with a distance of 1 pixel from each other) in directions $\alpha = 0^\circ$ and opposite $\alpha = 180^\circ$ (i.e., to the left and right from the pixel with the given value), the following symmetrical GLCM is obtained:

$$\mathbf{P} = \begin{bmatrix} 6 & 4 & 3 & 0 & 0 \\ 4 & 0 & 2 & 1 & 3 \\ 3 & 2 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 2 \end{bmatrix}$$

Let:

ϵ be an arbitrarily small positive number ($\approx 2.2 \times 10^{-16}$)

$\mathbf{P}(i, j)$ be the co-occurrence matrix for an arbitrary δ and α

$p(i, j)$ be the normalized co-occurrence matrix and equal to $\frac{\mathbf{P}(i, j)}{\sum \mathbf{P}(i, j)}$

N_g be the number of discrete intensity levels in the image

$p_x(i) = \sum_{j=1}^{N_g} P(i, j)$ be the marginal row probabilities

$p_y(j) = \sum_{i=1}^{N_g} P(i, j)$ be the marginal column probabilities

μ_x be the mean gray level intensity of p_x and defined as $\mu_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) i$

μ_y be the mean gray level intensity of p_y and defined as $\mu_y = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) j$

σ_x be the standard deviation of p_x

σ_y be the standard deviation of p_y

$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)$, where $i + j = k$, and $k = 2, 3, \dots, 2N_g$

$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)$, where $|i - j| = k$, and $k = 0, 1, \dots, N_g - 1$

$HX = -\sum_{i=1}^{N_g} p_x(i) \log_2(p_x(i) + \epsilon)$ be the entropy of p_x

$HY = -\sum_{j=1}^{N_g} p_y(j) \log_2(p_y(j) + \epsilon)$ be the entropy of p_y

$HXY = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2(p(i, j) + \epsilon)$ be the entropy of $p(i, j)$

$HXY1 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2(p_x(i)p_y(j) + \epsilon)$

$HXY2 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(j) \log_2(p_x(i)p_y(j) + \epsilon)$

By default, the value of a feature is calculated on the GLCM for each angle separately, after which the mean of these values is returned. If distance weighting is enabled, GLCM matrices are weighted by weighting factor W and then summed and normalised. Features are then calculated on the resultant matrix. Weighting factor W is calculated for the distance between neighbouring voxels by:

$W = e^{-\|d\|^2}$, where d is the distance for the associated angle according to the norm specified in setting ‘weightingNorm’.

The following class specific settings are possible:

- symmetricalGLCM** [True]: boolean, indicates whether co-occurrences should be assessed in two directions per angle, which results in a symmetrical matrix, with equal distributions for i and j . A symmetrical matrix corresponds to the GLCM as defined by Haralick et al.
- weightingNorm** [None]: string, indicates which norm should be used when applying distance weighting. Enumerated setting, possible values:
 - ‘manhattan’: first order norm
 - ‘euclidean’: second order norm
 - ‘infinity’: infinity norm.
 - ‘no_weighting’: GLCMs are weighted by factor 1 and summed
 - None: Applies no weighting, mean of values calculated on separate matrices is returned.

In case of other values, an warning is logged and option ‘no_weighting’ is used.

References

- Haralick, R., Shanmugan, K., Dinstein, I; Textural features for image classification; IEEE Transactions on Systems, Man and Cybernetics; 1973(3), p610-621
- https://en.wikipedia.org/wiki/Co-occurrence_matrix
- http://www.fp.ualgary.ca/mhallbey/the_glcm.htm

getAutocorrelationFeatureValue ()

Calculate and return the mean Autocorrelation.

$$autocorrelation = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)ij$$

Autocorrelation is a measure of the magnitude of the fineness and coarseness of texture.

pyradiomics Documentation, Release 1.1.1

getAverageIntensityFeatureValue ()

Return the mean gray level intensity of the i distribution.

$$\mu_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) i$$

Warning: As this formula represents the average of the distribution of i , it is independent from the distribution of j . Therefore, only use this formula if the GLCM is symmetrical, where $p_x(i) = p_y(j)$, where $i = j$.

getClusterProminenceFeatureValue ()

Using coefficients μ_x and μ_y , calculate and return the mean Cluster Prominence.

$$cluster\ prominenc = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x(i) - \mu_y(j))^4 p(i, j)$$

Cluster Prominence is a measure of the skewness and asymmetry of the GLCM. A higher values implies more asymmetry about the mean while a lower value indicates a peak near the mean value and less variation about the mean.

getClusterShadeFeatureValue ()

Using coefficients μ_x and μ_y , calculate and return the mean Cluster Shade.

$$cluster\ shade = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x(i) - \mu_y(j))^3 p(i, j)$$

Cluster Shade is a measure of the skewness and uniformity of the GLCM. A higher cluster shade implies greater asymmetry about the mean.

getClusterTendencyFeatureValue ()

Using coefficients μ_x and μ_y , calculate and return the mean Cluster Tendency.

$$cluster\ tendency = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x(i) - \mu_y(j))^2 p(i, j)$$

Cluster Tendency is a measure of groupings of voxels with similar gray-level values.

getContrastFeatureValue ()

Using the squared difference between gray values of neighbouring paris, calculate and return the mean Contrast.

$$contrast = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 p(i, j)$$

Contrast is a measure of the local intensity variation, favoring $P(i, j)$ values away from the diagonal ($i = j$). A larger value correlates with a greater disparity in intensity values among neighboring voxels.

getCorrelationFeatureValue ()

Using coefficients μ_x , μ_y , σ_x and σ_y , calculate and return the mean Correlation.

$$correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) i j - \mu_x(i) \mu_y(j)}{\sigma_x(i) \sigma_y(j)}$$

Correlation is a value between 0 (uncorrelated) and 1 (perfectly correlated) showing the linear dependency of gray level values to their respective voxels in the GLCM.

getDifferenceAverageFeatureValue ()

Using coefficient p_{x-y} , calculate and return the mean Difference Average.

$$difference\ average = \sum_{k=0}^{N_g-1} k p_{x-y}(k)$$

Difference Average measures the relationship between occurrences of pairs with similar intensity values and occurrences of pairs with differing intensity values.

getDifferenceEntropyFeatureValue ()

Using coefficient p_{x-y} , calculate and return the mean Difference Entropy.

$$difference\ entropy = \sum_{k=0}^{N_g-1} p_{x-y}(k) \log_2 (p_{x-y}(k) + \epsilon)$$

Difference Entropy is a measure of the randomness/variability in neighborhood intensity value differences.

getDifferenceVarianceFeatureValue ()

Using coefficients p_{x-y} and DifferenceAverage (DA) calculate and return the mean Difference Variance.

$$difference\ variance = \sum_{k=0}^{N_g-1} (1 - DA)^2 p_{x-y}(k)$$

Difference Variance is a measure of heterogeneity that places higher weights on differing intensity level pairs that deviate more from the mean.

getDissimilarityFeatureValue ()

Calculate and return the mean Dissimilarity.

$$dissimilarity = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i - j| p(i, j)$$

Dissimilarity is a measure of local intensity variation defined as the mean absolute difference between the neighbouring pairs. A larger value correlates with a greater disparity in intensity values among neighboring voxels.

getEnergyFeatureValue ()

Calculate and return the mean Energy.

$$energy = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i, j))^2$$

Energy (or Angular Second Moment) is a measure of homogeneous patterns in the image. A greater Energy implies that there are more instances of intensity value pairs in the image that neighbor each other at higher frequencies.

getEntropyFeatureValue ()

Calculate and return the mean Entropy.

$$entropy = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2 (p(i, j) + \epsilon)$$

Entropy is a measure of the randomness/variability in neighborhood intensity values.

getHomogeneity1FeatureValue ()

Calculate and return the mean Homogeneity 1.

$$homogeneity\ 1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + |i - j|}$$

pyradiomics Documentation, Release 1.1.1

Homogeneity 1 is a measure of the similarity in intensity values for neighboring voxels. It is a measure of local homogeneity that increases with less contrast in the window.

getHomogeneity2FeatureValue ()

Calculate and return the mean Homogeneity 2.

$$homogeneity\ 2 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + |i - j|^2}$$

Homogeneity 2 is a measure of the similarity in intensity values for neighboring voxels.

getImc1FeatureValue ()

Using coefficients HX , HY , HXY and $HXY1$, calculate and return the mean Informal Measure of Correlation 1.

$$IMC\ 1 = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

getImc2FeatureValue ()

Using coefficients HXY and $HXY2$, calculate and return the mean Informal Measure of Correlation 2.

$$IMC\ 2 = \sqrt{1 - e^{-2(HXY2 - HXY)}}$$

getIdmFeatureValue ()

Calculate and return the mean Inverse Difference Moment.

$$IDM = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + |i - j|^2}$$

IDM (inverse difference moment) is a measure of the local homogeneity of an image. IDM weights are the inverse of the Contrast weights (decreasing exponentially from the diagonal $i=j$ in the GLCM).

getIdmnFeatureValue ()

Calculate and return the mean Inverse Difference Moment Normalized.

$$IDMN = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + \left(\frac{|i-j|^2}{N_g^2}\right)}$$

IDMN (inverse difference moment normalized) is a measure of the local homogeneity of an image. IDMN weights are the inverse of the Contrast weights (decreasing exponentially from the diagonal $i = j$ in the GLCM). Unlike Homogeneity2, IDMN normalizes the square of the difference between neighboring intensity values by dividing over the square of the total number of discrete intensity values.

getIdFeatureValue ()

Calculate and return the mean Inverse Difference.

$$ID = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + |i - j|}$$

ID (inverse difference) is another measure of the local homogeneity of an image. With more uniform gray levels, the denominator will remain low, resulting in a higher overall value.

getIdnFeatureValue ()

Calculate and return the mean Inverse Difference Normalized.

$$IDN = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + \left(\frac{|i-j|}{N_g}\right)}$$

IDN (inverse difference normalized) is another measure of the local homogeneity of an image. Unlike Homogeneity1, IDN normalizes the difference between the neighboring intensity values by dividing over the total number of discrete intensity values.

getInverseVarianceFeatureValue ()

Calculate and return the mean Inverse Variance.

$$inverse\ variance = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{|i-j|^2}, i \neq j$$

getMaximumProbabilityFeatureValue ()

Calculate and return the mean Maximum Probability.

$$maximum\ probability = \max(p(i,j))$$

Maximum Probability is occurrences of the most predominant pair of neighboring intensity values.

getSumAverageFeatureValue ()

Coefficient p_{x+y} , calculate and return the mean Sum Average.

$$sum\ average = \sum_{k=2}^{2N_g} p_{x+y}(k)k$$

Sum Average measures the relationship between occurrences of pairs with lower intensity values and occurrences of pairs with higher intensity values.

getSumEntropyFeatureValue ()

Using coefficient p_{x+y} , calculate and return the mean Sum Entropy.

$$sum\ entropy = \sum_{k=2}^{2N_g} p_{x+y}(k) \log_2(p_{x+y}(k) + \epsilon)$$

Sum Entropy is a sum of neighborhood intensity value differences.

getSumVarianceFeatureValue ()

Using coefficients p_{x+y} and SumEntropy (SE) calculate and return the mean Sum Variance.

$$sum\ variance = \sum_{k=2}^{2N_g} (k - SE)^2 p_{x+y}(k)$$

Sum Variance is a measure of heterogeneity that places higher weights on neighboring intensity level pairs that deviate more from the mean.

getSumVariance2FeatureValue ()

Using coefficients p_{x+y} and SumAverage (SA) calculate and return the mean Sum Variance 2.

$$sum\ variance\ 2 = \sum_{k=2}^{2N_g} (k - SA)^2 p_{x+y}(k)$$

Sum Variance 2 is a measure of heterogeneity that places higher weights on neighboring intensity level pairs that deviate more from the mean.

This formula differs from SumVariance in that instead of subtracting the SumEntropy from the intensity, it subtracts the SumAverage, which is the mean of intensities and not its entropy

getSumSquaresFeatureValue ()

Using coefficients i and μ_x , calculate and return the mean Sum of Squares (also known as Variance) of the i distribution.

$$sum\ squares = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 p(i,j)$$

Sum of Squares or Variance is a measure in the distribution of neighboring intensity level pairs about the mean intensity level in the GLCM.

pyradiomics Documentation, Release 1.1.1

Warning: This formula represents the variance of the distribution of i and is independent from the distribution of j . Therefore, only use this formula if the GLCM is symmetrical, where $p_x(i) = p_y(j)$, where $i = j$

Gray Level Size Zone Matrix (GLSZM) Features

class radiomics.glszm.**RadiomicsGLSZM**(*inputImage*, *inputMask*, ***kwargs*)

Bases: *radiomics.base.RadiomicsFeaturesBase*

A Gray Level Size Zone (GLSZM) quantifies gray level zones in an image. A gray level zone is defined as a the number of connected voxels that share the same gray level intensity. A voxel is considered connected if the distance is 1 according to the infinity norm. This yields a 26-connected region in a 3D image, and an 8-connected region in a 2D image. In a gray level size zone matrix $P(i, j)$ the $(i, j)^{\text{th}}$ element describes the number of times a gray level zone with gray level i and size j appears in image.

As a two dimensional example, consider the following 5x5 image, with 5 discrete gray levels:

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix}$$

The GLSZM then becomes:

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Let:

$\mathbf{P}(i, j)$ be the size zone matrix

$p(i, j)$ be the normalized size zone matrix, defined as $p(i, j) = \frac{\mathbf{P}(i, j)}{\sum \mathbf{P}(i, j)}$

N_g be the number of discrete intensity values in the image

N_s be the number of discrete zone sizes in the image

N_p be the number of voxels in the image

References

- Guillaume Thibault; Bernard Fertil; Claire Navarro; Sandrine Pereira; Pierre Cau; Nicolas Levy; Jean Sequeira; Jean-Luc Mari (2009). "Texture Indexes and Gray Level Size Zone Matrix. Application to Cell Nuclei Classification". Pattern Recognition and Information Processing (PRIP): 140-145.

- https://en.wikipedia.org/wiki/Gray_level_size_zone_matrix

getSmallAreaEmphasisFeatureValue ()

Calculate and return the Small Area Emphasis (SAE) value.

$$SAE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathbf{P}(i, j)}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i, j)}$$

A measure of the distribution of small size zones, with a greater value indicative of more smaller size zones and more fine textures.

getLargeAreaEmphasisFeatureValue ()

Calculate and return the Large Area Emphasis (LAE) value.

$$LAE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j) j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j)}$$

A measure of the distribution of large area size zones, with a greater value indicative of more larger size zones and more coarse textures.

getGrayLevelNonUniformityFeatureValue ()

Calculate and return the Gray Level Non-Uniformity (GLN) value.

$$GLN = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_s} \mathbf{P}(i,j))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j)}$$

Measures the variability of gray-level intensity values in the image, with a lower value indicating more homogeneity in intensity values.

getGrayLevelNonUniformityNormalizedFeatureValue ()

Calculate and return the Gray Level Non-Uniformity Normalized (GLNN) value.

$$GLNN = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_s} \mathbf{P}(i,j))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \mathbf{P}(i,j)^2}$$

Measures the variability of gray-level intensity values in the image, with a lower value indicating a greater similarity in intensity values. This is the normalized version of the GLN formula.

getSizeZoneNonUniformityFeatureValue ()

Calculate and return the Size-Zone Non-Uniformity (SZN) value.

$$SZN = \frac{\sum_{j=1}^{N_s} (\sum_{i=1}^{N_g} \mathbf{P}(i,j))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j)}$$

Measures the variability of size zone volumes in the image, with a lower value indicating more homogeneity in size zone volumes.

getSizeZoneNonUniformityNormalizedFeatureValue ()

Calculate and return the Size-Zone Non-Uniformity Normalized (SZNN) value.

$$SZNN = \frac{\sum_{j=1}^{N_s} (\sum_{i=1}^{N_g} \mathbf{P}(i,j))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \mathbf{P}(i,j)^2}$$

Measures the variability of size zone volumes throughout the image, with a lower value indicating more homogeneity among zone size volumes in the image. This is the normalized version of the SZN formula.

getZonePercentageFeatureValue ()

Calculate and return the Zone Percentage (ZP) value.

$$ZP = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j)}{N_p}$$

Measures the coarseness of the texture by taking the ratio of number of zones and number of voxels in the ROI. Values are in range $\frac{1}{N_p} \leq ZP \leq 1$, with higher values indicating a larger portion of the ROI consists of small zones (indicates a more fine texture).

getGrayLevelVarianceFeatureValue ()

Calculate and return the Gray Level Variance (GLV) value.

$$GLV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j) (i - \mu)^2, \text{ where}$$

$$\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j) i$$

pyradiomics Documentation, Release 1.1.1

Measures the variance in gray level intensities for the zones.

getZoneVarianceFeatureValue ()

Calculate and return the Zone Variance (ZV) value.

$$ZV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i, j)(j - \mu)^2, \text{ where}$$

$$\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i, j)j$$

Measures the variance in zone size volumes for the zones.

getZoneEntropyFeatureValue ()

Calculate and return the Zone Entropy (ZE) value.

$$ZE = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i, j) \log_2(p(i, j) + \epsilon)$$

Here, ϵ is an arbitrarily small positive number ($\approx 2.2 \times 10^{-16}$).

Measures the uncertainty/randomness in the distribution of zone sizes and gray levels. A higher value indicates more heterogeneity in the texture patterns.

getLowGrayLevelZoneEmphasisFeatureValue ()

Calculate and return the Low Gray Level Zone Emphasis (LGLZE) value.

$$LGLZE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i, j)}{i^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i, j)}$$

Measures the distribution of lower gray-level size zones, with a higher value indicating a greater proportion of lower gray-level values and size zones in the image.

getHighGrayLevelZoneEmphasisFeatureValue ()

Calculate and return the High Gray Level Zone Emphasis (HGLZE) value.

$$HGLZE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i, j)i^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i, j)}$$

Measures the distribution of the higher gray-level values, with a higher value indicating a greater proportion of higher gray-level values and size zones in the image.

getSmallAreaLowGrayLevelEmphasisFeatureValue ()

Calculate and return the Small Area Low Gray Level Emphasis (SALGLE) value.

$$SALGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i, j)}{i^2 j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i, j)}$$

Measures the proportion in the image of the joint distribution of smaller size zones with lower gray-level values.

getSmallAreaHighGrayLevelEmphasisFeatureValue ()

Calculate and return the Small Area High Gray Level Emphasis (SAHGLE) value.

$$SAHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i, j)i^2}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i, j)}$$

Measures the proportion in the image of the joint distribution of smaller size zones with higher gray-level values.

getLargeAreaLowGrayLevelEmphasisFeatureValue ()

Calculate and return the Large Area Low Gray Level Emphasis (LALGLE) value.

$$LALGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathbf{P}(i,j)^2}{i^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j)}$$

Measures the proportion in the image of the joint distribution of larger size zones with lower gray-level values.

getLargeAreaHighGrayLevelEmphasisFeatureValue ()

Calculate and return the Large Area High Gray Level Emphasis (LAHGLE) value.

$$LAHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j) i^2 j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j)}$$

Measures the proportion in the image of the joint distribution of larger size zones with higher gray-level values.

Gray Level Run Length Matrix (GLRLM) Features

class radiomics.glrml.RadiomicsGLRLM(*inputImage, inputMask, **kwargs*)

Bases: *radiomics.base.RadiomicsFeaturesBase*

A Gray Level Run Length Matrix (GLRLM) quantifies gray level runs in an image. A gray level run is defined as the length in number of pixels, of consecutive pixels that have the same gray level value. In a gray level run length matrix $\mathbf{P}(i, j|\theta)$, the (i, j) th element describes the number of times a gray level i appears consecutively j times in the direction specified by θ .

As a two dimensional example, consider the following 5x5 image, with 5 discrete gray levels:

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix}$$

The GLRLM for $\theta = 0$, where 0 degrees is the horizontal direction, then becomes:

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Let:

$\mathbf{P}(i, j|\theta)$ be the run length matrix for an arbitrary direction θ

$p(i, j|\theta)$ be the normalized run length matrix, defined as $p(i, j|\theta) = \frac{\mathbf{P}(i,j|\theta)}{\sum \mathbf{P}(i,j|\theta)}$

N_g be the number of discrete intensity values in the image

N_r be the number of discrete run lengths in the image

N_p be the number of voxels in the image

By default, the value of a feature is calculated on the GLRLM for each angle separately, after which the mean of these values is returned. If distance weighting is enabled, GLRLMs are weighted by the distance between neighbouring voxels and then summed and normalised. Features are then calculated on the resultant matrix. The distance between neighbouring voxels is calculated for each angle using the norm specified in 'weightingNorm'.

The following class specific settings are possible:

- `weightingNorm [None]`: string, indicates which norm should be used when applying distance weighting. Enumerated setting, possible values:

pyradiomics Documentation, Release 1.1.1

- ‘manhattan’: first order norm
- ‘euclidean’: second order norm
- ‘infinity’: infinity norm.
- ‘no_weighting’: GLCMs are weighted by factor 1 and summed
- None: Applies no weighting, mean of values calculated on separate matrices is returned.

In case of other values, an warning is logged and option ‘no_weighting’ is used.

References

- Galloway MM. 1975. Texture analysis using gray level run lengths. Computer Graphics and Image Processing, 4(2):172-179.
- Chu A., Sehgal C.M., Greenleaf J. F. 1990. Use of gray value distribution of run length for texture analysis. Pattern Recognition Letters, 11(6):415-419
- Xu D., Kurani A., Furst J., Raicu D. 2004. Run-Length Encoding For Volumetric Texture. International Conference on Visualization, Imaging and Image Processing (VIIP), p. 452-458
- Tang X. 1998. Texture information in run-length matrices. IEEE Transactions on Image Processing 7(11):1602-1609.
- Tustison N., Gee J. Run-Length Matrices For Texture Analysis. Insight Journal 2008 January - June.

getShortRunEmphasisFeatureValue ()

Calculate and return the mean Short Run Emphasis (SRE) value for all GLRLMs.

$$SRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)}{i^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)}$$

A measure of the distribution of short run lengths, with a greater value indicative of shorter run lengths and more fine textural textures.

getLongRunEmphasisFeatureValue ()

Calculate and return the mean Long Run Emphasis (LRE) value for all GLRLMs.

$$LRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)}$$

A measure of the distribution of long run lengths, with a greater value indicative of longer run lengths and more coarse structural textures.

getGrayLevelNonUniformityFeatureValue ()

Calculate and return the mean Gray Level Non-Uniformity (GLN) value for all GLRLMs.

$$GLN = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} P(i,j|\theta))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)}$$

Measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values.

getGrayLevelNonUniformityNormalizedFeatureValue ()

Calculate and return the Gray Level Non-Uniformity Normalized (GLNN) value.

$$GLNN = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} P(i,j|\theta))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)^2}$$

Measures the similarity of gray-level intensity values in the image, where a lower GLNN value correlates with a greater similarity in intensity values. This is the normalized version of the GLN formula.

getRunLengthNonUniformityFeatureValue ()

Calculate and return the mean Run Length Non-Uniformity (RLN) value for all GLRLMs.

$$RLN = \frac{\sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_g} \mathbf{P}(i,j|\theta) \right)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta)}$$

Measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image.

getRunLengthNonUniformityNormalizedFeatureValue ()

Calculate and return the mean Run Length Non-Uniformity Normalized (RLNN) value for all GLRLMs.

$$RLNN = \frac{\sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_g} \mathbf{P}(i,j|\theta) \right)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta)}$$

Measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image. This is the normalized version of the RLN formula.

getRunPercentageFeatureValue ()

Calculate and return the mean Run Percentage (RP) value for all GLRLMs.

$$RP = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta)}{N_p}$$

Measures the coarseness of the texture by taking the ratio of number of runs and number of voxels in the ROI. Values are in range $\frac{1}{N_p} \leq RP \leq 1$, with higher values indicating a larger portion of the ROI consists of short runs (indicates a more fine texture).

getGrayLevelVarianceFeatureValue ()

Calculate and return the Gray Level Variance (GLV) value.

$$GLV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)(i - \mu)^2, \text{ where}$$

$$\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)i$$

Measures the variance in gray level intensity for the runs.

getRunVarianceFeatureValue ()

Calculate and return the Run Variance (RV) value.

$$RV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)(j - \mu)^2, \text{ where}$$

$$\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta)j$$

Measures the variance in runs for the run lengths.

getRunEntropyFeatureValue ()

Calculate and return the Run Entropy (RE) value.

$$RE = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j|\theta) \log_2(p(i,j|\theta) + \epsilon)$$

Here, ϵ is an arbitrarily small positive number ($\approx 2.2 \times 10^{-16}$).

getLowGrayLevelRunEmphasisFeatureValue ()

Calculate and return the mean Low Gray Level Run Emphasis (LGLRE) value for all GLRLMs.

pyradiomics Documentation, Release 1.1.1

$$LGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta)}{i^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta)}$$

Measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.

getHighGrayLevelRunEmphasisFeatureValue ()

Calculate and return the mean High Gray Level Run Emphasis (HGLRE) value for all GLRLMs.

$$HGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta) i^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta)}$$

Measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.

getShortRunLowGrayLevelEmphasisFeatureValue ()

Calculate and return the mean Short Run Low Gray Level Emphasis (SRLGLE) value for all GLRLMs.

$$SRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta)}{i^2 j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta)}$$

Measures the joint distribution of shorter run lengths with lower gray-level values.

getShortRunHighGrayLevelEmphasisFeatureValue ()

Calculate and return the mean Short Run High Gray Level Emphasis (SRHGLE) value for all GLRLMs.

$$SRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta) i^2}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta)}$$

Measures the joint distribution of shorter run lengths with higher gray-level values.

getLongRunLowGrayLevelEmphasisFeatureValue ()

Calculate and return the mean Long Run Low Gray Level Emphasis (LRLGLE) value for all GLRLMs.

$$LRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta) j^2}{i^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta)}$$

Measures the joint distribution of long run lengths with lower gray-level values.

getLongRunHighGrayLevelEmphasisFeatureValue ()

Calculate and return the mean Long Run High Gray Level Emphasis (LRHGLE) value for all GLRLMs.

$$LRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta) i^2 j^2}{i^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta)}$$

Measures the joint distribution of long run lengths with higher gray-level values.

Developers

This section contains information on how to add or customize the feature classes and filters available in PyRadiomics. PyRadiomics enumerates the available feature classes and input image types at initialization of the toolbox. These are available from the global `radiomics` namespace by use of the functions `getFeatureClasses ()` and `getInputImageTypes ()`, respectively. Individual features in a feature class are enumerated at initialization of the class. See also the [contributing guidelines](#)

Signature of a feature class

Each feature class is defined in a separate module, the module name is used as the feature class name (e.g. if module `tex.py` matches the feature class signature, it is available in the PyRadiomics toolbox as the ‘`tex`’ feature class). In the module a class should be defined that fits the following signature:

Annexe B

Liste des acronymes

- ADC** *apparent diffusion coefficient*. 5, 50, 51
- ASL** *arterial spin labelling*. 5, 14, 50, 194
- AUC** aire sous la courbe ROC. 31, 32, 98, 123, 125–127, 129, 130, 133, 137, 147, 162, 163, 165, 180
- BIDS** *brain imaging data structure*. 66
- CBF** *cerebral blood flow*. 5
- CBV** *cerebral blood volume*. 5
- CCI** coefficient de corrélation intra-classe. 136, 137
- Cho** Choline. 6, 50
- courbe ROC** fonction d'efficacité du récepteur. 31
- courbes ROC** fonctions d'efficacité du récepteur. 98
- CQ** contrôle automatique de qualité. 66, 67, 69, 71, 77, 80, 84
- Cr** Créatine. 6
- DCE** *dynamic contrast enhancement*. 5
- DH** distance de Hellinger. 67, 69, 73, 74, 77, 80, 82–85
- DICOM** *Digital imaging and communications in medicine*. 52, 66
- DSC** *dynamic susceptibility contrast*. 5, 194
- DTI** *diffusion tensor images*. 6, 50
- DWI** *diffusion weighted images*. 6, 14
- EGFR** récepteur du facteur de croissance épidermique. 8
- FLAIR** *fluid-attenuated inversion recovery*. 4, 5, 8, 46, 49, 51–54, 64, 67, 69–80, 85, 87, 88, 93, 113, 123, 126, 130, 133, 140, 144, 148, 159, 162, 165, 175, 176, 186, 187
- GBM** glioblastome multiforme. 6, 8, 14, 52
- GITC** gliome infiltrant du tronc cérébral. 2, 6, 8, 46, 48–52, 54, 64, 87, 88, 112, 116, 117, 123, 134, 139, 151, 158, 162, 167, 189, 191, 192, 194
- GLCM** matrice de cooccurrence des niveaux de gris. 17–19, 93, 176
- GLRLM** *gray level run-length matrix*. 17–19, 93, 117
- GLSZM** *gray level size zone matrix*. 17, 19, 93, 96
- hWS** *hybrid white stripe*. 62–65, 67, 69, 79–81, 87, 88, 191, 194

- IBSI** *image biomarker standardisation initiative*. 16, 18, 19
- IDH** *isocitrate déhydrogénase*. 6, 8, 14
- indice C** indice de concordance. 14, 32, 170, 172, 175, 180, 182
- IRB** *Institutional Review Board*. 52
- IRM** imagerie par résonance magnétique. 2, 4, 5, 8, 14, 15, 27, 28, 33, 46, 49, 51, 52, 54, 62–64, 69, 77, 80, 84, 87, 88, 92, 93, 95, 96, 98, 112, 117, 123, 154, 191, 194
- LASSO** *Least Absolute Shrinkage and Selection Operator*. 21, 181
- LCR** liquide céphalorachidien. 4, 62
- LOOCV** validation croisée avec exclusion d'un élément. 134, 135, 137, 139–141, 144, 146–149, 151, 152, 156, 162, 163, 166, 180–182, 186, 187
- MB** matière blanche. 4, 64, 65, 67–69, 71–74, 77, 78, 80–85, 87, 88, 106, 117
- MG** matière grise. 4, 62, 64, 67–70, 72, 75–78, 80–85, 88
- MGMT** méthylguanine méthyltransférase. 8
- MNI** *Montreal Neurological Institute*. 64, 65, 67, 82, 106
- mRMR** *minimum redundancy maximum relevance*. 22, 163, 164
- MTT** *mean transit time*. 5
- NAA** N-acétylaspartate. 6, 50
- NIFTI** *Neuroimaging Informatics Technology Initiative*. 52
- OMS** organisation mondiale de la santé. 6, 48
- PHRC** programme hospitalier de recherche clinique. 48, 54, 151, 167, 192, 193
- RANO** *Response Assessment in Neuro-Oncology*. 49
- RAVEL** *Removal of Artificial Voxel Effect by Linear regression*. 62, 80, 87
- RH** rapport de hasards. 32, 33, 51, 170, 172, 175, 182
- SMOTE** *Synthetic minority Over-samplig Technique*. 139, 144, 154
- T1c** *T1-weighted* après injection de produit de contraste. 4, 5, 8, 14, 46, 51–54, 64, 67, 69–80, 85, 87, 88, 93, 106, 123, 126, 130, 133, 140, 144, 145, 148, 151, 162, 175, 176, 186, 187
- T1w** *T1-weighted*. 4, 8, 46, 51–54, 64, 66–77, 92, 93, 123, 126, 130, 133, 140, 144, 145, 151, 154, 155, 159, 162, 175, 176, 186
- T2w** *T2-weighted*. 4, 8, 14, 33, 46, 49, 51–54, 63–67, 69–76, 84, 92, 93, 123, 126, 130, 133, 144, 148, 155, 162, 165, 175, 176, 186, 187
- TDM** tomodensitométrie. 2, 28, 33
- TE** temps d'écho. 2, 4, 53, 54
- TEP** tomographie par émission de positons. 2, 62, 87, 92, 93, 116
- TR** temps de répétition. 2, 4, 53, 54
- WS** *white stripe*. 64, 66, 67, 69, 80, 81
- WT** *wild-type*. 46, 53, 54, 127–132, 160, 162, 165, 193, 194

Annexe C

Glossaire

cohorte radiomique 4Mod La cohorte de 38 patients PREBIOMEDE qui présentent les quatre modalités d'imagerie structurelle dans la session de diagnostique. Utilisée pour les premières analyses radiomiques.. [123](#), [126](#), [127](#), [130](#), [131](#), [133–137](#), [140](#), [141](#), [144](#), [145](#), [147](#), [148](#), [150](#), [162](#), [165](#), [166](#), [170](#), [175](#), [176](#), [180](#), [182](#), [186](#), [187](#), [189](#)

cohorte étoile La première cohorte (en gras dans le tableau [A.1](#)) de patients PREBIOMEDE utilisée comme référence de qualité, car pour ces données, toutes les étapes du pipeline de standardisation ont été validées visuellement et individuellement.. [52](#), [53](#), [66](#), [67](#), [69–71](#), [77](#), [80](#), [92](#), [93](#), [123](#)

Annexe D

Liste des symboles

$\mu_{i,j}^{m*}$ mode de l'histogramme d'intensités des voxels correspondant aux couches 80 à 120 du repère MNI pour le patient i , visite j , modalité m . 64

τ Une valeur de tolérance autour de la mode de la fonction de distribution cumulative de la matière blanche.. 64, 66

$VOI_{biopsie}$ Volume sphérique de rayon 4,5mm autour de la région où la biopsie a été faite.. 92, 93, 95, 98, 106, 117

VOI_{tumeur} Volume sphérique tracé à l'intérieur de la tumeur.. 112, 113, 117, 123, 136, 137, 170, 172

Titre : Développements en radiomique pour une meilleure caractérisation du gliome infiltrant du tronc cérébral à partir d'imagerie par résonance magnétique

Mots clés : Radiomique, IRM, Neuro-oncologie, Apprentissage Automatique

Résumé : La radiomique suppose que des informations pertinentes non repérables visuellement peuvent être trouvées en calculant une grande quantité d'indices quantitatifs à partir des images médicales. En cancérologie, ces informations pourraient caractériser le phénotype de la tumeur et définir le pronostic du patient. Le GITC est une tumeur pédiatrique rare diagnostiquée d'après des signes cliniques et son apparence en IRM. Cette thèse présente les premières études radiomiques pour des patients atteints de GITC. Comme les intensités en IRM clinique sont exprimées en unités arbitraires, la première étape de l'étude a été la standardisation des images. Une méthode de normalisation basée sur l'estimation de l'intensité dans la matière blanche d'apparence normale s'est avérée efficace sur plus de 1500 volumes d'images. Des études méthodologiques sur le calcul des indices de texture ont abouti aux recommandations suivantes : (a) discrétiser les niveaux de gris avec une largeur constante pour tous les patients, (b) utiliser un volume d'intérêt constant ou faire attention au biais introduit par des volumes de taille et forme différentes. En s'appuyant sur ces recommandations, les indices radiomiques issus de 4 modalités d'IRM ont été systématiquement analysés en vue de prédire les principales mutations génétiques associées aux GITC et la survie globale des patients au moment du diag-

nostic. Un pipeline de sélection d'indices a été proposé et différentes méthodes d'apprentissage automatique avec validation croisée ont été mises en œuvre pour les deux tâches de prédiction. La combinaison des indices cliniques avec les indices d'imagerie est plus efficace que les indices cliniques ou d'imagerie seuls pour la prédiction des deux principales mutations de l'histone H3 (H3.1 versus H3.3) associées au GITC. Comme certaines modalités d'imagerie étaient manquantes, une méthodologie adaptée à l'analyse des bases de données d'imagerie multi-modales avec données manquantes a été proposée pour pallier les limites de recueil des données d'imagerie. Cette approche permet d'intégrer de nouveaux patients. Les résultats du test externe de prédiction des deux principales mutations de l'histone H3 sont encourageants. Concernant la survie, certains indices radiomiques semblent informatifs. Toutefois, le faible nombre de patients n'a pas permis d'établir les performances des prédicteurs proposés. Enfin, ces premières études radiomiques suggèrent la pertinence des indices radiomiques pour la prise en charge des patients atteints de GITC en absence de biopsie mais l'augmentation de la base de données est nécessaire pour confirmer ces résultats. La méthodologie proposée dans cette thèse peut être appliquée à d'autres études cliniques.

Title : Developments in radiomics for improving diffuse intrinsic pontine glioma characterization using magnetic resonance imaging

Keywords : Radiomics, MRI, Neuro-oncology, Machine Learning

Abstract : Radiomics is based on the assumption that relevant, non-visually identifiable information can be found by calculating a large amount of quantitative indices from medical images. In oncology, this information could characterize the phenotype of the tumor and define the prognosis of the patient. DIPG is a rare pediatric tumor diagnosed by clinical signs and MRI appearance. This work presents the first radiomic studies for patients with DIPG. Since clinical MRI intensities are expressed in arbitrary units, the first step in the study was image standardization. A normalization method based on intensity estimation of the normal-appearing white matter has been shown to be effective on more than 1500 image volumes. Methodological studies on the calculation of texture indices have then defined the following recommendations: (a) discretize gray levels with a constant width for all patients, (b) use a constant volume of interest or pay attention to the bias introduced by volumes of different size and shape. Based on these recommendations, radiomic indices from four MRI modalities were systematically analyzed to predict the main genetic mutations associated with DIPG and the overall survival of patients at the time of diagnosis. An index selec-

tion pipeline was proposed and different cross-validated machine learning methods were implemented for both prediction tasks. The combination of clinical indices with imaging indices is more effective than the clinical or imaging indices alone for the prediction of the two main mutations in histone H3 (H3.1 versus H3.3) associated with DIPG. As some imaging modalities were missing, a methodology adapted to the analysis of multi-modal imaging databases with missing data was proposed to overcome the limitations of the collection of imaging data. This approach made it possible to integrate new patients. The results of the external prediction test for the two main mutations of H3 histone are encouraging. Regarding survival, some radiomic indices seem to be informative. However, the small number of patients did not make it possible to establish the performance of the proposed predictors. Finally, these first radiomic studies suggest the relevance of the radiomic indices for the management of patients with DIPG in the absence of biopsy but the database need to be increased in order to confirm these results. The proposed methodology can be applied to other studies.

