THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Informatique

École doctorale Information Structures Systèmes I2S

Unité de recherche L2C UMR5221

IDENTIFICATION DES MARQUEURS IRM MORPHOLOGIQUES ET FONCTIONNELS PREDICTIFS DE LA CROISSANCE DES GLIOMES DE BAS GRADE : DEVELOPPEMENT D'UN ALGORITHME AUTOMATIQUE DE SEGMENTATION ET DE PREDICTION DE LA TRANSFORMATION

Présentée par Margaux VERDIER

Le 05 Décembre 2022

Sous la direction de Nicolas MENJOT DE CHAMPFLEUR

Devant le jury composé de

Luc TAILLANDIER, PU-PH, Centre hospitalier régional et universitaire de Nancy	Rapporteur
Damien GALANAUD, PU-PH, Hôpitaux Universitaires Pitié Salpêtrière - AP-HP	Rapporteur
Emmanuelle LE BARS, PhD, Ingénieure Physique Imagerie, CHU Gui de Chauliac	Co-encadrante
Matthieu LE PETIT COIFFE, PhD, Collaboration Manager France, Siemens Healthcare SAS	Co-encadrant
Amélie DARLIX, PH Neuro-oncologue, Institut du cancer de Montpellier (ICM)	Examinatrice
Christophe FIORIO, PU, LIRMM, équipe ICAR, CNRS, Université de Montpellier	Examinateur
Hugues DUFFAU, PU-PH Neurochirurgien, CHU Gui de Chauliac	Examinateur

UNIVERSITÉ DE MONTPELLIER

Table des matières

Introduction 1

I. Gliomes diffus de bas grade 3

a. Généralités 3

1. Origine cellulaire et moléculaire 3

2. Epidémiologie 3

3. Classification OMS 4

4. Histoire naturelle 4

b. Prise en charge clinique 5

1. Diagnostic clinique 5

2. Chirurgie éveillée 6

3. Identification histologique et de biologie moléculaire 8

4. Traitement de chimio/radiothérapie 9

5. Suivi radiologique 10

6. Identification de la transformation anaplasique 22

II. Segmentation automatique des gliomes de bas grade 26

a. Réseau de neurones convolutifs 26

1. Principe 27

2. L'algorithme U-Net 27

3. Implémentation nnU-Net 29

b. Données de référence : segmentation manuelle 32

c. Population 32

d. Article - Evaluation of an nnU-Net type automated tumor segmentation as a clinical volumetric tool for low-grade glioma follow-up 33

III. Caractérisation de la tumeur : analyse radiomique 49

a. Définition 50

b. Paramètres 51

1. L’analyse de premier ordre 51
2. Les paramètres de forme 51
3. L’analyse de second ordre : les paramètres de texture 51
c. Radiomique dans les gliomes diffus de bas-grade 55
d. Des images aux paramètres radiomiques 56
1. Population de l’étude radiomique 56
2. Chaine de traitement 60
IV. Prédiction de l’évolution du gliome 72
a. Réduction de variables 72
1. Principaux algorithmes 73
2. Méthode d’analyse 77
3. Résultats 78
b. Sélection du modèle 80
1. Le support vecteur machine 80
2. Les critères de sélection 84
3. Résultats 86
4. Discussion 87
c. Prédiction 88
1. SVM de validation 88
2. Pertinence du score de probabilité SVM 89
V. Conclusion 102
Références 105
Annexes 123
Annexe 1 - Tableau des 770 paramètres radiomiques 123
Annexe 2- Courbe d'évolution du diamètre moyen tumoral 150

Remerciements

Aux membres du jury :

Merci aux Professeurs *Luc Taillandier* et *Damien Galanaud* d'avoir accepté d'être rapporteurs de ce travail de thèse. Merci au Professeur *Hugues Duffau*, à Mme *Amélie Darlix* et M.*Christophe Fiorio* d'avoir accepté de venir examiner cette thèse lors de la soutenance. Merci à M. *Luc Bauchet* et Mme *Valérie Rigau* d'avoir accepté d'assister à ma soutenance de thèse.

J'ai hâte de vos retours et de nos discussions futures.

A mes encadrant·e·s :

Je remercie vivement le *Professeur Nicolas Menjot de Champfleur* d'avoir accepté d'être directeur de cette thèse et de m'avoir accompagnée pendant ces 3 ans (et 4mois).

Nicolas, merci pour ta confiance.

A *Emmanuelle Le Bars*, un énorme merci pour tout le soutien et les réflexions nouvelles apportées à chaque étape de ce travail. Emmanuelle, merci pour ta disponibilité et ton enthousiasme, merci pour cet environnement de travail stimulant que tu m'as apporté tout au long de cette thèse.

A *Matthieu Le Petit Coiffe*, merci pour ton encadrement et ta disponibilité, ce partenariat CIFRE m'a beaucoup appris et apporté.

A mes collègues :

A *Jeremy Deverdun*. Je me souviens encore mon premier jour de stage dans cette petite salle de recherche, les posters de « The IT Crowd » et les figurines Star Wars m'ont fait aussitôt réaliser que j'étais en présence d'une personne de très bon gout. Merci de m'avoir fait réaliser que l'informatique et le code étaient extraordinaires et surtout que j'étais capable de m'y affronter. Merci d'avoir fait de mon lieu de travail, un lieu agréable où je me rends de bon cœur, c'est très précieux. Enfin, un énorme merci pour ta présence pendant cette dernière semaine d'écriture intensive. J'espère que « Mobilette » et « Fifolette » continueront à collaborer dans le futur.

Merci *à Guillaume* et *Zakaria*, mes voisins de bureau, pour toutes ces aventures incroyables entre midi et deux.

A *Justine Belko*, merci du fond du cœur. Merci pour tout le travail que l'on a réalisé ensemble depuis le stage jusqu'à ma thèse, mais surtout merci pour ton amitié et ton soutien. J'espère qu'on continuera à faire des mots croisés à grande vitesse.

Merci aussi à *Amel Benali*, la 3ième super nana, qui a continué à me soutenir même après la fin de sa thèse.

A *Liesjet*, *Arthur* et *Céline* mais aussi à *Victor* merci pour votre soutien depuis le tout début de ma thèse et même avant !
A *Thierry Chaptal*, merci de t'être toujours montré disponible pour m'assister dans mon travail.
Je tiens à remercier *Marinette Moynier* et *Olivier Martin* ainsi que toute l'équipe du *service de neuroradiologie du CHU Gui de Chauliac* que je croise au quotidien. Aux nouveaux Docteurs en médecine *Thomas Dos Santos* et *Philippe Lam*, nos travaux de thèse se sont enrichis mutuellement par notre collaboration, merci pour votre gentillesse et votre enthousiasme.
Je tiens aussi à remercier toutes les personnes de *Siemens* avec lesquelles j'ai eu la chance de travailler. Merci à *Thomas Troalen* pour sa gentillesse et son intérêt constant pour mon travail.
Un grand merci à *Till Huelnhagen* et *Bénédicte Maréchal*, vous avez toujours été très disponibles et très intéressés par mon travail. Sans vous, cet article n'aurait pas pu voir le jour, je vous en suis très reconnaissante.

A ma famille :

Merci à mes merveilleux parents, *Annick* et *Frédou*. Merci de m'avoir soutenue pendant cette thèse et depuis toujours, et d'avoir constamment eu confiance en moi, même quand je ne savais pas trop où j'allais. Je crois que j'ai trouvé mon chemin ! Je vous aime.
A mon super frère *Joris*, merci pour tes encouragements depuis l'océan indien, je les ai bien reçus et ils m'ont toujours motivés.
Merci à ma tatie *Antoinette* qui a permis ma retraite d'écriture en Lozère et qui s'est tellement occupée de moi. Une pensée pleine d'affection pour mon tonton *Louis* qui aurait adoré que je lui raconte mon travail.
Aux *Mertens-Deckers* et à la petite *Veyrier*, ma deuxième famille, merci pour l'amour, les gentils petits mots et les bons desserts.
Merci à *Simon* qui est tout simplement le meilleur.

A mes ami·e·s :

A *Julie*, *Mathilde* et *Jessie*, mes plus vieilles amies, mais aussi à *Caroline*, *Raida* et *Sandra*, merci pour la force que vous m'avez donnée.
A toute ma troupe de « copaings », merci pour tous ces moments plein de joie

Merci à *toutes les personnes* que j'ai oublié de citer

Abréviations

GBG : Gliome de bas grade

GHG: Gliome de haut grade

OMS: Organisation mondiale de la Santé

IRM: Imagerie par résonnance magnétique

FLAIR: Fluid Attenuated Inversion Recovery

MTD ou DMT: Diamètre moyen tumoral

VDE ou VED: Vitesse d’expansion du diamètre moyen tumoral

CNN: Réseau de neurones convolutifs

DSC: Coefficient de similitude de Dice

Introduction

Les gliomes représentent 50% des tumeurs cérébrales primitives. Les gliomes de bas grade (GBG), classés grade 1 et 2 par l'Organisation mondiale de la Santé (OMS)[1], sont caractérisés par une croissance lente et continue[2]. En général, les GBG sont diagnostiqués chez les jeunes adultes, après une première crise d'épilepsie. Le principal traitement des patient·e·s atteint·e·s de GBG est la chirurgie en condition éveillée afin d'augmenter la survie tout en préservant les fonctions cérébrales importantes. Le prise en charge par la suite peut inclure des traitements de radiothérapie et/ou de chimiothérapie[3] La détermination du grade tumoral s'appuie sur l'analyse du tissu tumoral obtenu par biopsie ou exérèse de la lésion. Malheureusement, les GBG évoluent toujours vers un grade agressif, le gliome de haut grade (GHG) classé grade 3 et 4 par l'OMS. La transformation anaplasique est définie par l'évolution d'une tumeur de grade II vers un grade III et est un événement inéluctable dans l'histoire de la maladie. La médiane de survie, sans traitement, est de 6,7 ans pour les GBG et de 4,9 ans pour les gliomes en transformation anaplasique[4].

Le diagnostic, le suivi et le traitement des gliomes diffus de bas grade reposent sur la modalité d'Imagerie par Résonance Magnétique (IRM), considérée comme référence [5].

Les prédicteurs les plus précis de cette transformation, dite transformation anaplasique, sont actuellement la vitesse d'expansion du diamètre moyen (VED)[2] calculée à partir du volume tumoral[6] estimé sur la séquence T2 FLAIR[7] et le rapport du volume sanguin cérébral[8] extrait des cartographies de ce volume sanguin cérébral obtenues à partir de la séquence de perfusion en IRM. La prise de contraste renseignée en pratique courante sur l'imagerie morphologique 3D T1 suite à l'injection de produit de contraste (chélates de gadolinium), indique l'existence d'une rupture de la barrière hémato-encéphalique témoin d'une transformation maligne, mais n'est pas spécifique du moment de cette transformation anaplasique.

C'est donc le volume qui reste la mesure la plus précise pour mettre en évidence une évolution péjorative du gliome. Les GBG, de par leur statut infiltrant, diffus et en constante évolution, sont difficiles à segmenter sur la seule base des intensités de l'image. Bien que cela puisse être réalisé par des médecins avertis, la segmentation manuelle reste extrêmement chronophage. Des méthodes automatiques se développent de plus en plus[9–11]. L'accès difficile à un nombre important de données d'entrainement et l'aspect diffus des tumeurs de bas grade complexifie la mise en place d'une segmentation automatique fiable. De plus, ces algorithmes sont

très rarement entrainés sur des données hétérogènes semblables à celles de routine clinique (tumeurs opérées, traitées par chimiothérapie, machines d'acquisition différentes…).

Le caractère infiltrant, les contours flous, les évolutions lentes et hétérogènes du signal T2 FLAIR et du volume sanguin au sein de la tumeur sont des éléments typiques des GBG observés en IRM en routine clinique. Ces caractéristiques ne sont pourtant actuellement pas prises en compte ni exploitées comme prédicteurs de la transformation anaplasique en pratique courante. Cependant ces éléments semblent être en lien avec l'évolution anaplasique des gliomes et de l'activité mitotique d'une forme maligne[12]. La radiomique est une méthode consistant à extraire un grand nombre de paramètres quantitatifs (texture, forme, intensités...) au sein de la tumeur à partir d'images radiologiques. Cette technique est de plus en plus appliquée à l'étude des tumeurs[13–16] et parait intéressante pour l'étude de l'évolution longitudinale des GBG.

L'objectif final de ce travail de thèse est d'être capable de prédire la transformation anaplasique ou, a minima, l'évolution péjorative du gliome cérébral diffus de bas grade au cours de son évolution.

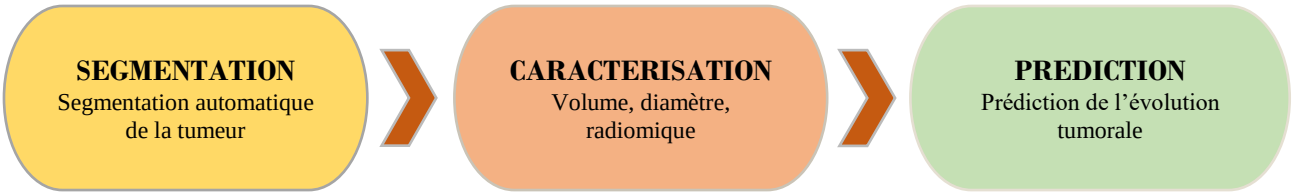


Figure 1 : Les trois objectifs du travail de thèse.

Pour parvenir à cet objectif, la mise en place d'une segmentation automatique du volume tumoral adapté à des imageries de GBG en routine clinique a été la première étape. Pour prédire l'évolution de la tumeur, il a fallu la caractériser au cours de son suivi longitudinal. Cette caractérisation a été possible par le calcul de paramètres quantitatifs morphologiques (volume, diamètre moyen de la tumeur…) mais aussi de paramètres plus fins, radiomiques, extraits au sein du volume tumoral sur des imageries anatomiques et de perfusion (texture, intensités…). Grace à des outils dérivés de l'intelligence artificielle, la combinaison de paramètres caractérisant l'évolution, voire la transformation anaplasique, a été mise en évidence. Pour finir, la capacité de ces paramètres quantitatifs à prédire une transformation maligne a été évaluée (*Figure* 1).

I. Gliomes diffus de bas grade

Les gliomes diffus de bas grade (GBG) sont des tumeurs primitives, c'est-à-dire développées à partir des tissus cérébraux, qui ont pour origine les cellules gliales du parenchyme cérébral. Ce sont des tumeurs infiltrantes, de croissance lente, qui touchent les jeunes adultes. Ces gliomes sont classés en astrocytomes ou oligodendrogliomes de grade 2 après leur étude histologique et moléculaire, conformément à la classification OMS en vigueur (la dernière étant celle de 2021)[1]. Ces tumeurs peuvent se développer dans l'ensemble du système nerveux central mais apparaissent le plus souvent dans la région fronto-temporo-insulaire ou le lobe frontal[17].

Les GBG évoluent préférentiellement le long des faisceaux de matière blanche, vers une transformation maligne, dite transformation anaplasique[18]. L'anaplasie désigne la perte de différenciation des cellules faisant en sorte qu'elles retournent à une forme immature et moins spécialisée ; elles se divisent rapidement et leur apparence et leur comportement est différent de cellules saines. Cette transformation laisse place à un gliome de haut-grade (GHG), grade 3, et peut aller jusqu'à sa forme la plus agressive, le glioblastome.

a. Généralités

1. Origine cellulaire et moléculaire

Les cellules souches neurales et les précurseurs des oligodendrocytes sont considérées comme les deux populations cellulaires à l'origine des gliomes diffus de bas grade. Leur naissance se fait par l'activation ou l'inactivation de gènes ou de cascades moléculaires au sein des cellules. Les gliomes possèdent des comportements biologiques différents selon leur type cellulaire ou leur région cérébrale d'origine, ce qui implique des vitesses de croissance et des réponses aux traitements différents[19].

2. Epidémiologie

Les gliomes, dans leur ensemble, représentent 50% des tumeurs cérébrales primitives chez l'adulte. Le registre central des tumeurs cérébrales des Etats-Unis a recensé entre 2012 et 2016 que les GBG représentent 6,5% des tumeurs primitives du système nerveux[17]. L'incidence pour les gliomes de bas grades, astrocytomes et oligodendrogliomes confondus, est de 0,76 cas pour 100 000 habitants par an, et touche majoritairement les jeunes adultes.

3. Classification OMS

Les gliomes sont classés de 1 à 4 selon leur degré de malignité :

- Grade 1 : Il correspond à des tumeurs à croissance lente et normalement bien circonscrites, bien qu'elles puissent envahir de grandes régions du cerveau. Ce sont des tumeurs bénignes qui, selon leur localisation, peuvent être traitées chirurgicalement.
- Grade 2 : ce sont des tumeurs à croissance lente, mais aux limites imprécises. Les tumeurs de grade I et II sont dite de "bas grade" (GBG)
- Grade 3 : Ce sont des tumeurs anaplasiques, les cellules se développent rapidement et de manière chaotique.
- Grade 4 : Ce grade désigne des tumeurs pouvant contenir divers types de cellules qui se multiplient rapidement et qui peuvent nécroser. Ce sont des tumeurs infiltrantes dont les limites ne sont pas facilement discernables en imagerie. Elles peuvent être des astrocytomes de grade 4 ou des glioblastomes.

Les tumeurs de grade 3 et 4 sont des tumeurs de "haut grade" (GHG).

4. Histoire naturelle

Les gliomes diffus de bas grade sont de croissance lente et leur origine temporelle est difficile à établir. Ils sont le plus souvent localisés dans les zones cérébrales temporales, frontales et insulaires qui sont des zones régissant un grand nombre d'activités fonctionnelles.

Dans un premier temps, la phase dite silencieuse est associée à une évolution lésionnelle très lente qui permet la mise en place d'une neuroplasticité et n'engendre donc pas de symptômes particuliers[20]. Au cours de cette période, la découverte de la lésion se fait de manière fortuite, lors d'un examen scanner ou IRM prescrit pour une raison différente[20].

La lésion gliale va évoluer en s'infiltrant et migrant préférentiellement le long des faisceaux de substance blanche[21]. Cette migration va engendrer l'apparition de crises d'épilepsie.

L'évolution du gliome diffus de bas-grade va se faire inexorablement vers la transformation anaplasique. Sans traitement, la durée de vie des patient·e·s est de 6,7ans[22,23].

L'objectif de la prise en charge clinique est de rallonger la durée de vie en repoussant le plus possible l'arrivée de cette transformation anaplasique. Des traitements sont appliqués pour réduire au maximum le volume et la croissance du gliome.

b. Prise en charge clinique

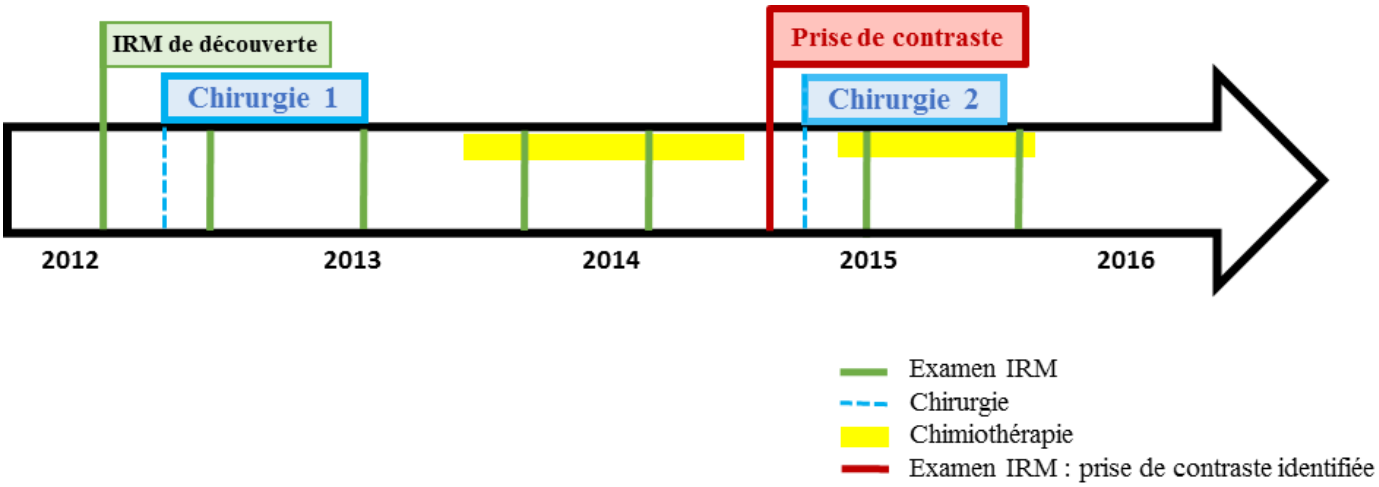


Figure 2: Exemple d'une prise en charge clinique dans le cas d'un gliome de bas grade.

La prise en charge clinique d'une personne diagnostiquée avec un gliome de bas grade se fait sur une longue période de temps. A partir de la découverte de la maladie, plusieurs traitements de type chirurgicaux ou encore de chimiothérapie ou radiothérapie peuvent avoir lieu. Les examens de neuroradiologie font office de surveillance de la croissance et d'évaluation de la réponse aux traitements de la lésion tumorale. Au cours du suivi, une prise de contraste sur l'imagerie T1 témoigne d'une transformation de la tumeur de bas grade en tumeur de haut grade (*Figure 2*).

1. Diagnostic clinique

La découverte d'un GBG se fait après une crise d'épilepsie pour 90% des cas. Ces crises sont partielles ou secondairement généralisées[24] et leur aspect dépend de la localisation du gliome responsable de la crise. Les 10% restant sont des découvertes fortuites suite à un examen IRM pour une pathologie différente[25]

C'est donc après un examen IRM que le diagnostic de gliomes diffus de bas grade est émis. Ces lésions possèdent une forme et un signal caractéristiques sur les séquences T1 et T2 FLAIR (*I.5.Suivi radiologique*).

Ces imageries permettent d'établir un diagnostic mais aussi de suspecter le stade d'évolution de la maladie. Le diagnostic final sera posé après une identification histologique de la tumeur, l'examen anatomopathologique.

Après la chirurgie, le résidu tumoral résultant de l'exérèse est analysé sur le plan anatomopathologique. Plusieurs lames sont réalisées pour une analyse microscopique des cellules tumorales, au cœur de la tumeur, en périphérie, ainsi que dans l'aspiration réalisée en fin d'opération permettant de collecter les résidus dans la cavité d'exérèse. Le but étant d'avoir des échantillons représentatifs de la tumeur et une vision globale de son état d'évolution. L'analyse se fait en fonction des directives établies par l'Organisation mondiale de la Santé (OMS), la dernière classification étant celle produite par l'OMS en 2021. En comparaison avec les versions antérieures, la distinction histologique d'astrocytome ou d'oligodendrogliome repose beaucoup plus sur les marqueurs moléculaires que sur l'aspect des cellules. Les types dits oligo-astrocytomes ou gliomes mixtes ne sont plus retenus dans cette nouvelle classification[1] car l'évolution des méthodes d'analyse permettent de mieux discriminer les types cellulaires.

2. Chirurgie éveillée

De nos jours, la technique qui a été démontrée comme permettant le gain de temps le plus important et une préservation de la qualité de vie des patient·e·s est la chirurgie en première intention. Avant tout autre traitement et dès la découverte de la lésion tumorale, la résection maximale du volume tumoral permet de ralentir l'avancée de la maladie de manière drastique. Il a été démontré qu'une résection totale ou au moins sub-totale avec un résidu de moins de 15mL impacte le cours naturel de croissance des gliomes de bas-grade[26]. Les pratiques cliniques tendent de plus en plus vers une prise en charge neuro-oncologique fonctionnelle et personnalisée[27]. Avec un traitement adapté, la médiane de survie des patient·e·s atteint·e·s de gliomes diffus de bas grade est passée à plus de 10 ans, permettant de juger ce suivi comparable à celui d'une maladie chronique. Il est donc nécessaire de mettre en place des techniques de traitement qui engendrent le moins possible de déficits neurologiques qui pourraient impacter la qualité de vie à long terme.

Le but d'un acte chirurgical dans le cadre de gliomes diffus de bas grade est de prélever le maximum de volume tumoral. Afin de respecter une balance « onco-fonctionnelle » et de s'adapter à chaque personne, la chirurgie éveillée a été jugée comme étant la méthode la plus efficace[28]. Au cours de la chirurgie éveillée, le·a neurochirurgien·enne utilise une stimulation électrique sur le cortex afin d'identifier les zones fonctionnelles essentielles, permettant d'éviter la résection de celles-ci. Après préparation de la zone cérébrale tumorale (incision du scalp, craniotomie puis incision de la dure mère), le·a patient·e est réveillé·e, et différentes tâches fonctionnelles guidées

par un·e neuropsychologue sont réalisées pour cartographier la zone fonctionnelle. Les tâches sont adaptées à la fonction. Si une zone concerne l'éloquence, une perturbation de la production de parole sera observée empêchant la dénomination d'objet ou la lecture par exemple. Au niveau du cortex moteur, la stimulation va déclencher des mouvements involontaires et au niveau du cortex sensitif des sensations de fourmillements.

Il a été montré que la réalisation d'une chirurgie en première intention était l'attitude la plus bénéfique à adopter. La survie globale est multipliée par deux comparée à une biopsie et un comportement de « wait and see », et ce indépendamment du type génétique[26,29]. Le volume du résidu tumoral post-chirurgie a son importance car il a été montré que le bénéfice pour la survie est garanti sur un résidu de moins de 15mL[26]. C'est pourquoi des résections plus larges que les limites observées en IRM ont été privilégiées afin d'augmenter le gain de survie global[30]. En effet, il a été démontré que l'imagerie ne représentait pas la totalité de la lésion et que l'infiltration de la tumeur se faisait au-delà des limites visibles sur l'imagerie FLAIR[31]. Les limites de la résection sont alors les limites fonctionnelles identifiées par la stimulation corticale au cours de la chirurgie éveillée. Une telle résection, dite « supra-totale », retarde la transformation anaplasique, contrairement à une résection se limitant à la zone identifiée par l'imagerie[28]. Une telle chirurgie est recommandée autant pour les gliomes symptomatiques qu'asymptomatiques. Dans le cas d'un gliome de bas grade découvert de manière fortuite, une résection de grande ampleur permet de retarder les symptômes, et donc la transformation, allongeant la durée de survie globale[23]. Dans le cadre d'une découverte suite à une crise d'épilepsie, une chirurgie supra-totale est bénéfique pour la préservation de la qualité de vie en sauvegardant les fonctions primaires et en réduisant l'épilepsie. L'épilepsie est très handicapante en ce qui concerne les gliomes de bas-grade car souvent résistante aux traitements médicamenteux, c'est pourquoi sa disparition permet une augmentation drastique de la qualité de vie.

La réalisation de plusieurs chirurgies éveillées au cours du suivi de la maladie est courante car, vu l'aspect continu de la croissance de ce type de gliomes, le retour à un stade plus agressif est inévitable après des chirurgies incomplètes. Des chirurgies en série permettent de stopper régulièrement la croissance tumorale et permettent aussi de tirer profit de la plasticité cérébrale[32,33]. La plasticité cérébrale est la capacité du système nerveux central à modifier son activité, à s'adapter, à se réorganiser en fonction de stimulus internes ou externes. C'est une activité qui se fait naturellement lors du vieillissement du cerveau mais qui a lieu aussi suite à un traumatisme. Dans le cas d'un gliome, la neuroplasticité qui entre en jeu est appelée « post-lésionnelle » et implique une

réorganisation et redistribution des zones fonctionnelles afin de pallier les déficiences de la zone lésée[20,34]. Plusieurs interventions chirurgicales permettront donc de supprimer au fur et à mesure les zones lésionnelles devenues non fonctionnelles afin d'atteindre la résection totale tout en préservant la qualité de vie.

3. Identification histologique et de biologie moléculaire

L'examen histologique consiste en une coloration des cellules grâce à un marquage hématoxyline/éosine (HE) qui permet de bien différencier les noyaux du cytoplasme. L'état de transformation est défini par le biais de plusieurs observations. La densité de cellules, leur morphologie et le nombre de noyaux en cours de mitose (division cellulaire) témoigne d'une transformation en cours, à plus ou moins grande échelle. Les zones de thrombose et de nécrose peuvent aussi être mises en évidence et sont le signe d'un stade de transformation plus agressif (glioblastome). D'autres marquages spécifiques pourront être réalisés dans certains cas, en complément, ainsi qu'une étude par technique immunohistochimique permettant de mettre en évidence des antigènes ayant une importance pronostique. Par exemple, le marquage de l'enzyme IDH qui, par son absence dans les cellules tumorales, mettra en évidence une mutation du gène du même nom[35]. L'utilisation de techniques cytogénétiques telle que la méthode d'hybridation fluorescente in situ (FISH) est aussi un outil indispensable afin de mettre en évidence certaines mutations génétiques.

Ces mutations donnent des indices sur le comportement évolutif du gliome et la réponse aux traitements. Il existe deux mutations clés qui sont recherchées en priorité pour effectuer le diagnostic.

Le gène IDH est responsable de la production de l'enzyme du même nom qui joue un rôle dans la protection des cellules contre le stress oxydatif. La mutation du gène IDH constitue un marqueur de bon pronostic, il augmente la survie globale et améliore la réponse à la chimiothérapie[36] comparativement aux gliomes IDH non-mutés, dit wild-type.

Parmi les gliomes IDH mutés, on distingue les astrocytomes diffus et les oligodendrogliomes par l'observation de l'intégrité des chromosomes 1 et 19, plus précisément la délétion ou non du bras court du chromosome 1 et du bras long du chromosome 19.

Une codélétion 1p/19q est le facteur déterminant pour le diagnostic d'oligodendrogliome synonyme d'une meilleure réponse à la chimiothérapie, et donc d'un meilleur pronostic[37].

Si la codélétion n'est pas présente, il s'agit d'un astrocytome IDH mutant et le grade de ce gliome est déterminé par une recherche plus pointue des paramètres moléculaires couplés aux signes d'anaplasie (nécrose, prise de contraste sur l'imagerie T1). Notamment la délétion homozygote des gènes CDKN2A et CDKN2B, impliqués dans le contrôle de la prolifération cellulaire. Cette délétion est d'un très mauvais pronostic et témoigne d'un astrocytome IDH mutant anaplasique grade 4 [38,39]. Dans le cas des gliomes IDH non mutés, une prolifération endothélio-capillaire (PEC), une nécrose, une augmentation de l'expression de facteurs de croissance épidermiques (EGFR), la mutation du promoteur de la protéine TERT ou bien le gain du chromosome 7 et la perte du 10 sont autant de facteurs en faveur d'une tumeur au stade avancé de gliobastome (grade 4)[40] (*Figure 3*).

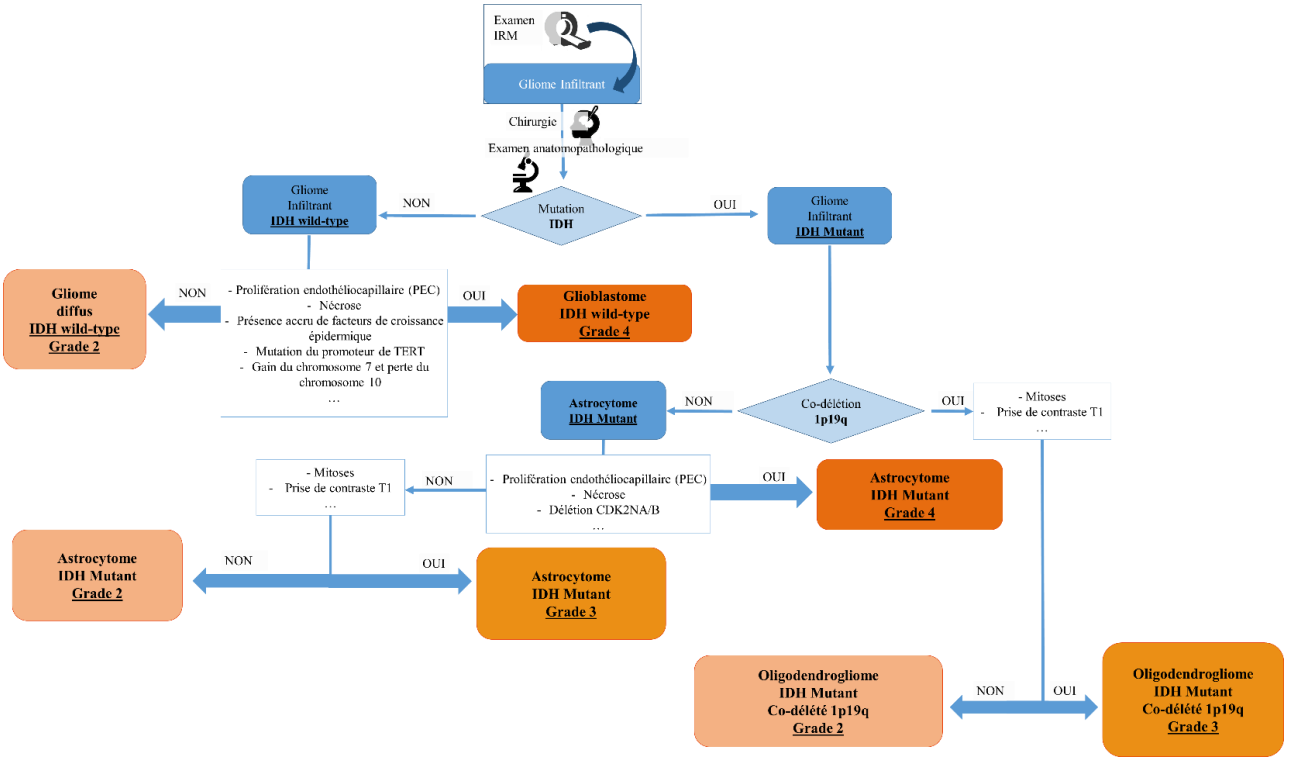


Figure 3 : Arbre décisionnel pour le diagnostic anatomopathologique des gliomes diffus de bas grade (OMS 2021)

4. Traitement de chimio/radiothérapie

Dans la prise en charge des gliomes de bas-grade, la chimiothérapie a un rôle important. Elle peut être utilisée à titre néoadjuvant, c'est-à-dire avant un traitement principal. Lorsqu'une lésion apparait très diffuse et qu'une résection totale est difficile à envisager, la chimiothérapie permet de réduire l'infiltration de la tumeur et donc de rendre, dans un second temps, la chirurgie possible et

plus efficace[41]. Lorsque cela est possible, c'est la chirurgie qui est réalisée en première intention et, dans le cas d'une évolution tumorale en période post opératoire, la chimiothérapie permet de contrôler l'augmentation de volume et de retarder une transformation anaplasique. La chimiothérapie permet aussi d'améliorer la qualité de vie en réduisant les crises d'épilepsie[42]. Il existe plusieurs types de chimiothérapie dans le cadre du traitement des gliomes de bas grade, correspondant à des molécules différentes. Le traitement par témozolomide a prouvé être mieux supporté par les patient·te·s et est préconisé en première intention. Il existe aussi le protocole PCV, coordination de l'effet de trois molécules : procarbazine, CCNU (ou lomustine) et vincristine ; ce protocole est plus rarement proposé car il possède une toxicité importante.

La radiothérapie n'est pas proposée en première ligne thérapeutique, mais plutôt pour des cas qui ne sont plus opérables et pour lesquels la maladie évolue malgré la chimiothérapie. En effet, la radiothérapie n'a pas d'impact sur la survie globale de la maladie[43,44]. De plus, quand elle est effectuée trop précocement elle a un effet délétère sur la qualité de vie à long terme avec un important déclin cognitif[45].

5. Suivi radiologique

Dans le cadre du suivi d'un gliome diffus de bas grade, l'exploration neuroradiologique par imagerie IRM est l'examen de première intention à partir de la découverte lésionnelle et tout au long du suivi de son évolution, en étroite collaboration avec les neurochirurgien·ne·s et oncologues.

L'imagerie par résonance magnétique (IRM) traditionnelle observe la résonance des protons (H+) dont la somme des spins, hors d'un champ, est une résultante magnétique nulle. La technique d'IRM s'appuie sur le fait que, soumis à un champ magnétique, les spins des particules du corps humain s'alignent et précessent autour de ce champ de manière parallèle ou antiparallèle. L'envoi d'un signal de radiofréquence (RF) correspondant à la fréquence de précession du spin des protons aboutit à un transfert d'énergie et à un phénomène de résonance. Cette phase d'excitation du système de spins produit une bascule de l'aimantation tissulaire. Les spins absorbent de l'énergie qu'ils restituent lorsque la fréquence RF est interrompue afin de retourner à leur état d'équilibre, cette étape correspond à la phase de relaxation. Ces phases d'excitation et de relaxation peuvent être décrites par une composante longitudinale et une composante transversale. Le temps de relaxation T1 correspond au temps pour lequel l'aimantation longitudinale retourne dans l'axe du champ magnétique, à 63% de sa valeur initiale. Le temps de relaxation T2, traduit le retour de

l'aimantation transversale à 37% de la fréquence de précession en place avant le signal RF. Ces deux temps T1 et T2 dépendent de la nature cellulaire mais aussi de l'environnement des tissus. La variation des pulses de RF (bascule de 90°, de 180°...), la variation des temps de répétition entre deux pulses de 90° (TR), l'ajout d'impulsion RF supplémentaire telle qu'une impulsion d'inversion récupération (180°) et le délai d'enregistrement de l'écho de signal (TE), sont autant de paramètres capables de pondérer une image en T1 ou T2 afin de mettre en évidence certains tissus[46].

Les gliomes diffus de bas grade sont des tumeurs à croissance lente qui évolueront toujours vers un stade de transformation maligne en tumeur de haut grade[47]. Les examens radiologiques réguliers reposent sur l'IRM cérébrale multimodale qui confirme le diagnostic de tumeur, et permet d'objectiver l'évolution lésionnelle, la vitesse de croissance et l'efficacité des traitements.

a) Acquisitions IRM

Certaines séquences sont particulièrement utilisées en routine clinique pour le suivi des GBG.

(1) L'imagerie 3D pondérée T1 avec injection de produit de contraste (3D T1 GADO)



Figure 4 : Imagerie pondérée T1 avec injection de produit de contraste

L'imagerie 3D pondérée T1 met en évidence en hypersignal les protons ayant des temps de relaxation T1 courts comme la graisse et ceux ayant T1 longs en hyposignal, c'est-à-dire l'eau et le liquide céphalo-rachidien (LCR). Les chélates de gadolinium, ou gadonlinium, sont le produit de contraste le plus couramment utilisé en IRM. Ce produit de contraste est injecté par voie intraveineuse. Disposant d'un effet sur les temps de relaxation T1 et T2 et persistant (quelques heures) sur le T1 des protons à proximité des chelates de gadolinium, un effet de rehaussement du signal est observé sur les imageries pondérées en T1. Ce produit est utilisé pour mettre en évidence le réseau sanguin et les vascularisations cérébrales, puisqu'il ne passe pas, en condition normale, la barrière hématoencéphalique. Le gliome de bas-grade apparaîtra en hyposignal s'il n'y a pas de rehaussement (*Figure 4*).

(2) L'imagerie T2 FLAIR

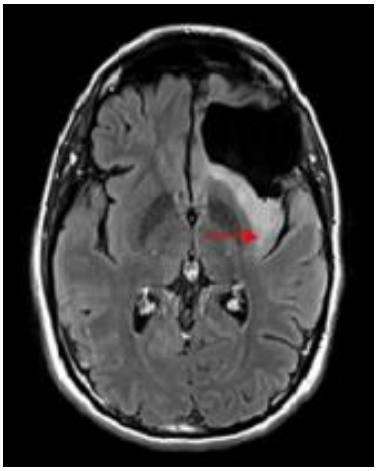


Figure 5 : Imagerie T2 FLAIR

Sur une image T2 FLAIR, un tissu avec un temps de relaxation T2 court, comme la graisse, apparaît en hyposignal alors que l'eau, qui possède un temps T2 long, apparaîtra en hypersignal. L'objectif de la séquence "FLuid Attenuated Inversion Recovery" (FLAIR) est de supprimer le signal de l'eau libre comme le liquide céphalo-rachidien (LCR). Les œdèmes (eau liée) et les tissus tumoraux (infiltration tumorale, inflammation) apparaissent en hypersignal (*Figure 5*).

(3) Imagerie de perfusion et cartographies de perfusion

La séquence utilisée est la séquence Echo de Gradient pondérée T2* couplée à un schéma de gradient ultra rapide dit Echo Planar Imaging (EPI) permettant la répétition d'un volume cérébral toutes les deux secondes sur une durée de l'ordre de 1min 30sec. Couplée à cette acquisition, une injection de produit de contraste (chélate de gadolinium) est réalisée de façon synchrone avec l'acquisition en bolus (0,2ml/kg pour un débit de 4 à 6cc/sec). Le raccourcissement du temps de relation T2* des protons à proximité des chélates de gadolinium induit une diminution de signal sur les imageries EG-EPI-T2*, ceci au cours du premier passage des chélates dans la circulation sanguine. La séquence est dénommée « Dynamic susceptibility contrast » : DSC. Cette technique de perfusion en IRM repose sur l'injection intraveineuse de produit de contraste et la mesure rapide des changements transitoires du signal pendant le passage du bolus des chélates de gadolinium dans le cerveau. La baisse du signal en T2*, due à la susceptibilité induite par le gadolinium, est enregistrée "voxel par voxel" afin de construire une courbe de l'intensité en fonction du temps. Le degré de diminution du signal dans un voxel du tissu cérébral dépend de la densité des capillaires sanguins au sein du voxel, mais aussi de la concentration du produit de contraste dans ces capillaires.

Dans l'hypothèse que la concentration en agent de contraste est proportionnelle au taux de relaxation transverse T2* et que l'intensité du signal pour une séquence pondérée en T2* a une relaxation mono-exponentielle, la courbe de concentration tissulaire en agent de contraste (C(t)) dans chaque voxel peut être déduite du signal IRM et utilisée pour le calcul de cartographies de

paramètres de perfusion. L'approche quantitative revient à calculer la réponse impulsionnelle caractéristique de chaque voxel en déconvoluant la concentration en agent de contraste, C(t), par la fonction d'entrée artérielle (AIF), puis à en extraire des paramètres associés à la perfusion. Le profil temporel de la réponse impulsionnelle pour chaque voxel dépend du flux cérébral sanguin, du temps de transit moyen, du volume sanguin cérébral et du temps au maximum (*Figure 6*).

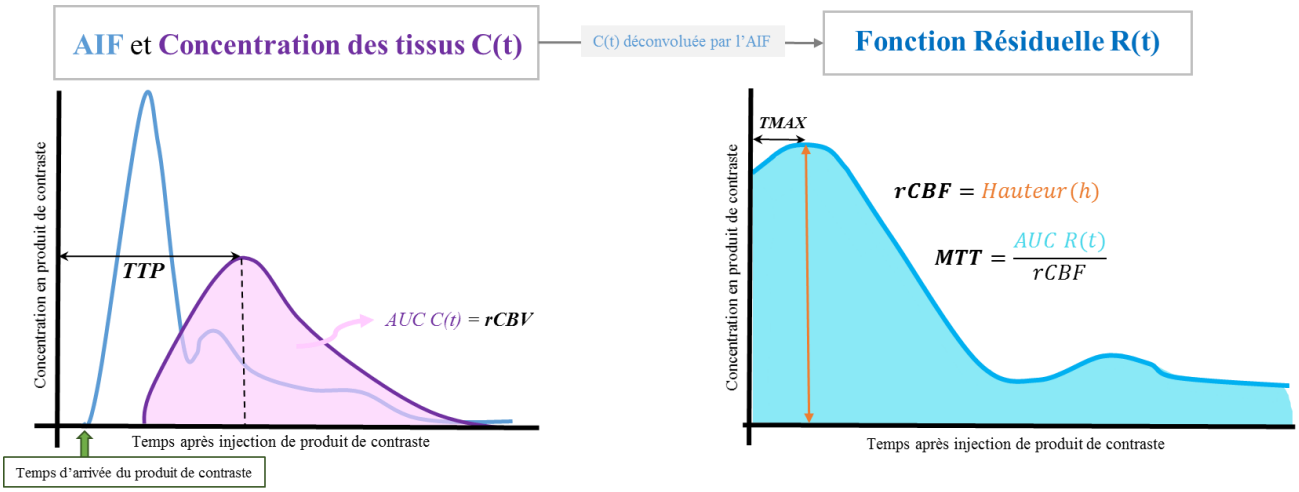


Figure 6: Paramètres d'imagerie de perfusion calculés à partir de : la courbe de concentration des tissus, C(t), (représentation inversée) produite par le passage du produit de contraste (chélates de gadolinium) dans la circulation sanguine cérébrale. Paramètres d'imagerie de perfusion calculés à partir de : la courbe de concentration des tissus, C(t), (représentation inversée) produite par le passage du produit de contraste (chélates de gadolinium) dans la circulation sanguine cérébrale. A partir de la C(t) sont calculés le temps d'arrivé au pic maximum de contraste (TTP) et l'aire sous la courbe correspond au volume sanguin cérébral (rCBV). La déconvolution de la C(t) avec la fonction d'entrée artérielle (AIF) donne la fonction résiduelle (R(t)) à partir de laquelle sont calculés le temps au contraste maximum (TMAX), le débit sanguin cérébral (rCBF) et le temps de transit moyen (MTT) qui est le rapport entre l'aire sous la courbe (AUC R(t)) et le rCBF.

L'analyse de la perfusion IRM suppose que l'agent de contraste reste confiné dans l'espace vasculaire durant toute la durée de l'acquisition. L'agent de contraste induit un raccourcissement des temps de relaxation T1 et T2/T2* des protons à sa proximité dû à la variation locale de champ magnétique induite par une concentration en agent de contraste localement importante. L'effet observé en imagerie EG-EPI-T2* au premier passage de l'agent de contraste est une chute de signal par effets de susceptibilité dans le compartiment vasculaire. Dès que l'agent de contraste atteint l'espace extracellulaire ou extravasculaire, son effet est majoritairement un raccourcissement du T1 caractérisé par un rehaussement de signal (effet de la pondération T1 résiduelle de la séquences EG-EPI-T2*). S'il existe une rupture de la barrière hémato-encéphalique, une accumulation a lieu dès le premier passage de l'agent de contraste dans l'espace extracellulaire, et la variation de signal

mesurée, au sein du voxel, ne représentera plus uniquement la variation de signal intravasculaire (*Figure 10*). L'extravasation de l'agent de contraste dans les lésions néovasculaires entraînera une sous-estimation de la mesure du rCBV[48,49].

A l'inverse, l'accumulation de l'agent de contraste peut engendrer un effet de susceptibilité magnétique important et induire une chute de signal T2* majorée, et dans ces régions, les courbes de concentration sont susceptibles de diminuer et de passer en dessous de la ligne de base, causant potentiellement une sur-estimation du volume sanguin relatif.

Tout modèle tentant de suivre avec précision les changements de perfusion doit tenir compte de la correction de cette extravasation afin de :

- Mieux distinguer les différents grades tumoraux[50]
- Différencier les effets du traitement (nécrose de radiation) de la récidive tumorale[51,52]
- Eviter la perte de corrélation entre le rCBV et le grade tumoral observé lorsque la correction de cette extravasation n'est pas appliquée.

La fonction d'entrée artérielle (AIF), représentant l'entrée de l'agent de contraste dans le tissu à l'échelle du voxel, est au cœur de la quantification du rCBF. La quantification de la perfusion implique la mesure de cette AIF et une analyse par déconvolution pour éliminer l'étalement temporel du signal mesuré (évolution de la concentration tissulaire dans le temps) associé à la forme de l'AIF. L'AIF correspond à la courbe de concentration moyennée sur un ensemble de pixels sélectionnés dans les segments proximaux des artères intracrâniennes.

En raison de son rôle fondamental, de nombreux progrès ont été réalisés concernant la manière de mesurer et l'endroit où mesurer cette AIF, son influence sur la quantification du rCBF en IRM (39).

L'ensemble des logiciels de post-traitement intègre maintenant la sélection automatique des voxels d'intérêt artériels et veineux afin de prendre en compte l'AIF dans le post traitement de la perfusion cérébrale en IRM.

Le logiciel OLEA Sphère 3.0[53] utilisé dans ce travail est capable de générer une fonction résiduelle, R(t), qui correspond à la déconvolution des courbes de concentration tissulaire par l'AIF et permet de générer un ensemble de cartes paramétriques à l'échelle du voxel (*Figure 6*) :

- Le volume sanguin régional relatif (rCBV) est obtenu en calculant l'aire sous la courbe de concentration en agent de contraste, C(t)

- Le temps d'arrivée à la variation maximum de contraste, correspond à la valeur crête de la courbe de C(t), c'est le TTP ("Time To Peak").
- Le débit sanguin cérébral (rCBF) correspond à la valeur maximale de la fonction résiduelle R(t)
- Le temps de transit moyen (MTT) est calculé en divisant l'aire sous la courbe de R(t) par le rCBF. Le MTT correspond à la largeur à mi-hauteur de la courbe de R(t).
- Le "Time to maximum", appelé TMAX, correspond au temps d'arrivée au pic de la courbe de R(t).
- La valeur K2 (ou Ktrans) est une estimation de la perméabilité magnétique. L'algorithme génère la carte des valeurs de K2 et celle d'un rCBV corrigé (rCBVc) de la possible extravasation de l'agent de contraste.
- La carte de la projection de l'intensité maximale dans le temps (tMIP) est construite à partir des données de perfusion brute, en affichant pour chaque pixel sa concentration maximale atteinte lors de l'acquisition. Cette carte représente donc la projection des intensités maximales au cours du temps de la concentration en produit de contraste.

Les données de rCBV et rCBF sont dites relatives, car les constantes de proportionnalité qui interviennent dans la conversion du signal en courbes de concentration sont inconnues. Ces valeurs dépendent de l'équipement d'imagerie, de l'agent de contraste, des paramètres d'injection et des paramètres d'acquisition.

Il existe d'autres séquences IRM qui ont montré de bons résultats dans la différenciation et le suivi de l'évolution des gliomes, mais qui sont moins généralement utilisées en pratique clinique courante.

(4) L'imagerie de diffusion

Cette imagerie est sensible aux micro-mouvements des molécules d'eau à l'intérieur des voxels dus à l'application de gradient de diffusion de part et d'autre de l'impulsion radio fréquences 180° d'une séquence dite de spin écho associée au mode EPI ultra rapide d'application des gradients. L'application des gradients de diffusion est paramétré par la valeur « b » dépendant de l'intensité, de la durée de l'application des gradients de diffusion et du temps séparant l'application des deux gradients de diffusion de part et d'autre de l'impulsion 180°. Dans les tissus cérébraux

cette diffusion est contrainte, contrairement à son aspect libre dans le LCR. Elle peut s'effectuer dans toutes les directions de l'espace (diffusion isotrope) ou préférentiellement dans une direction donnée (diffusion anisotrope) comme dans les fibres nerveuses. En situation pathologique, ces mouvements sont modifiés, dépendants de l'architecture, de l'état cellulaire, de la température et de la viscosité du milieu.

En réalisant des acquisitions pondérées en diffusion avec une valeur ou plusieurs valeurs de b et une nombre de directions de gradient pouvant aller de 6 à 256 directions, une analyse par tenseur de diffusion permet de générer un ensemble de paramètres caractérisant la diffusion tels que la diffusivité axiale, radiale, moyenne et la fraction d'anisotropie du tissu cérébral. De nombreux logiciels et modèles ont été développés pour l'analyse du tenseur afin de représenter les principaux faisceaux de la substance blanche représentés par quelques fibres myélinisées très longues qui permettent à des aires cérébrales distantes de communiquer entre elles.

Les études des gliomes à partir de l'imagerie de diffusion rapportent la capacité de cette modalité à mettre en évidence les changements structuraux de la tumeur :

- pour évaluer la réponse aux traitements[54], Castellano et a. montrent, à partir des analyses de l'histogramme des paramètres issus du tenseur DTI et de l'évolution de la diffusion au sein du gliome, les effets précoces de la chimiothérapie par temozolomide chez les patient·e·s atteint·e·s de gliome de bas grade.
- pour différencier gliomes de haut et bas grade[55,56], plusieurs nouveaux paramètres issus du tenseur permettent de les différencier, plus encore dans les régions tumorales solides non rehaussées avec injection de chélates de gadolinium, ce qui renforce la théorie selon laquelle les régions tumorales non rehaussées sont des régions de haut grade (*Figure 7*).

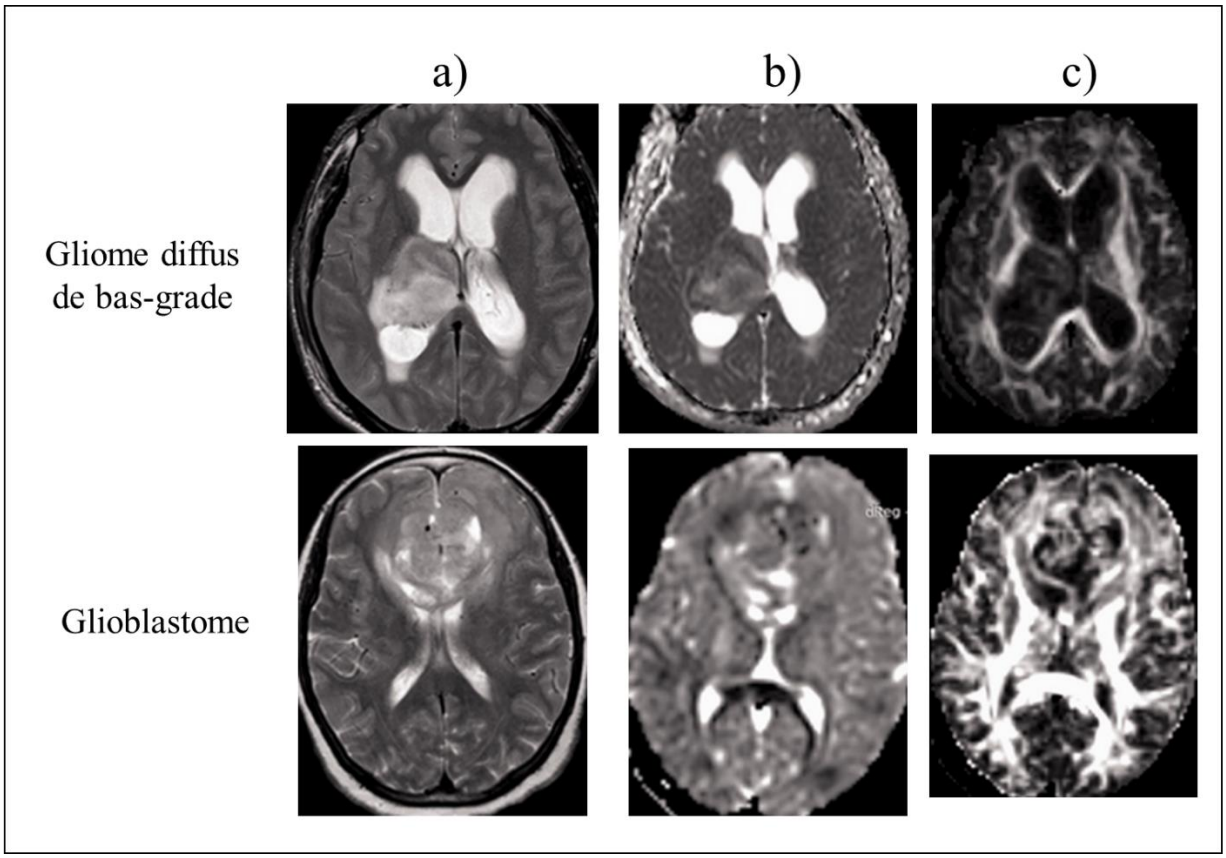


Figure 7:Comparaison de signal de tenseur de diffusion (DTI) entre gliomes de bas grades et glioblastomes (Extrait du travail d'El-Serougy et collaborateurs[56]). La colonne a) montre les images T2 FLAIR, la colonne b) les cartes de diffusion moyenne (MD) et la colonne c) la carte de fraction d'anisotropie (FA). Dans le cas des gliomes diffus de bas grade la carte de MD montre un signal homogène et la carte de FA montre un faible signal dans la partie solide de la tumeur. Pour le glioblastome, la carte de MD montre un iso-signal avec des zones linéaires à faible signal des parties solides de la tumeur, et la carte de FA montre un mélange de signaux hétérogènes, principalement faibles, de la partie solide de la tumeur.

(5) La spectroscopie par résonance magnétique nucléaire

C'est une technique de spectroscopie basée sur les propriétés magnétiques de certains noyaux atomiques. Le noyau le plus utilisé en spectroscopie chez l'homme est le proton. Le principe est de supprimer le signal des protons du pic de l'eau présents en grande quantité dans l'organisme et pouvant masquer les protons de molécules à étudier, afin de pouvoir observer et quantifier les protons des autres molécules du tissu cérébral.

Un *spectre RMN* comporte un ou des pics correspondant à la résonance des différents protons présents dans une molécule. La position du (ou des) pic(s) est déterminée par le

déplacement chimique (ppm) qui résulte de l'écran formé par le nuage électronique des noyaux d'hydrogène au sein de molécules.

Au niveau cérébral, plusieurs métabolites sont observables par spectroscopie du proton et sont à la base de son fonctionnement. Les plus importants d'entre eux sont observés et quantifiés généralement dans des études spectroscopiques à temps d'écho courts: le N- acétylaspartate (NAA) marqueur de la viabilité neuronale, la créatine/phosphocréatine (Cr/PCr) impliquée dans le métabolisme énergétique, la choline (Cho) principalement sous forme de phospholipides qui est un constituant important des membranes cellulaires/neuronales, le lactate (Lac) produit final de la glycolyse anaérobie, le myo- inositol (Ins) marqueur de la souffrance gliale, et d'autres métabolites plus difficilement observables et quantifiables, tels que le glucose (Glc), le glutamate (Glu), la glutamine (Gln), l'acide γ-amino- butyrique (GABA), l'aspartate (Asp), l'alanine, la taurine (Tau). Des lipides et des macromolécules sont aussi présents sur le spectre (*Figure 8*).

La quantification reste le plus souvent relative, même s'il est possible dans certaines conditions, et après calibration, de réaliser une quantification absolue de la concentration de métabolites.

Les domaines d'application de la spectroscopie par résonance magnétique sont essentiellement : les tumeurs, les pathologies inflammatoires et infectieuses et les pathologies métaboliques, principalement au niveau cérébral. Dans le cadre des tumeurs cérébrales et plus particulièrement des gliomes, la spectroscopie a plusieurs intérêts.

La valeur absolue du rapport Cho/NAA a été montrée plus élevée dans les lésions de haut grade que dans celles de bas grade. Le rapport Cho/NAA de 0,9 produit une sensibilité et une spécificité de 78 % dans la différenciation entre les lésions de bas grade et celles de haut grade. En combinant la spectroscopie et l'imagerie anatomique, une augmentation de la sensibilité a été observée, passant à 86 % et une spécificité à 80 %. Une forte corrélation a donc été démontrée enter le rapport Cho/Cr et Cho/NAA dans la différenciation des gliomes de bas grade et de haut grade[57].

Le statut histologique peut être discernable par spectroscopie comme par exemple la mutation IDH par détection du 2-Hydroxyglutarate[58], la codélétion 1p/19q[59] ou encore l'expression du Ki-67[60]. Le ratio choline/créatine permet de discerner parmi les gliomes ceux qui auront un statut plus agressif[61,62]. En général les GBG sont caractérisés par des niveaux de choline peu élevés, une

concentration de créatine qui augmente indique un gliome qui évolue rapidement, en cours de transformation[63] (*Figure 8*).

Les métabolites associés à la glutamine sont aussi étudiés car ils sont impliqués dans les processus inflammatoires qui, dans le processus d'évolution des GBG, peuvent être un moyen de quantifier de manière non-invasive la transformation[64–66].

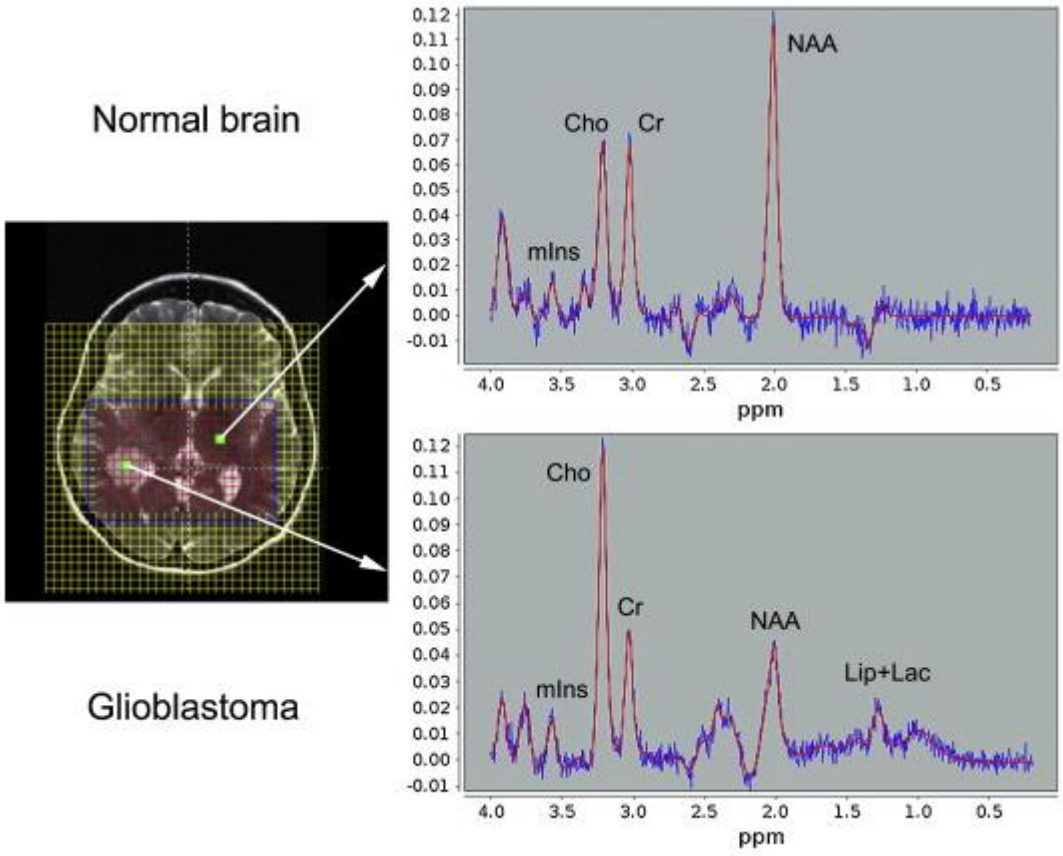


Figure 8: Spectroscopie protonique mono voxel du cerveau normal et d'un glioblastome (Extrait du travail Bulik et collaborateurs[63]) : graphiques de l'intensité du signal en fonction de sa fréquence (mIns = pic du myo-inositol, Cho = pic de la choline, Cr = pic de la créatine, NAA = pic du N-acétylaspartate)

b) Mesure d'évolution tumorale sur l'imagerie

Historiquement, plusieurs méthodes intégrant des données radiologiques ont été mises en place pour évaluer l'expansion de la lésion tumorale. En 2000, les critères RECIST unidimensionnels (Response Evaluation Criteria In Solid Tumours) ont été publiés dans le cadre de tumeurs solides[67]. Ces critères s'appuient sur la mesure du plus grand diamètre de la lésion (*Figure*

9). La mesure de ce diamètre se fait entre deux dates d'imagerie et vise à évaluer la réponse aux traitements en pourcentage d'évolution. Ces critères permettent de comptabiliser des lésions multiples en faisant la somme de chacun des plus grands diamètres, et incluent aussi la notion d'apparition et de disparition de lésions. Plusieurs études ont comparé les RECIST à d'autres types de mesures à plusieurs dimensions[68,69]. Ces études ont rapporté une concordance dans l'évaluation de la croissance tumorale, mais leur hétérogénéité et leur croissance irrégulière dans l'espace ont privilégié des mesures à plusieurs dimensions.

Les critères de RANO (Response Assessment in Neuro-Oncology) établis en 2010 furent longtemps utilisés pour évaluer le diamètre équivalent de la tumeur[70]. Les critères RANO ont pour origine les critères de MacDonald dont la première version a été éditée en 1990[71]. Ces critères permettent de définir l'indicateur de "survie globale" et aussi la "survie sans progression tumorale" (progression free survival = PFS) qui correspond à l'intervalle de temps dans lequel aucune évolution de la tumeur n'est observée après un traitement. Ces indicateurs s'appuient sur plusieurs facteurs cliniques et radiologiques y compris une mesure bi-dimensionnelle de la tumeur. Selon MacDonald, il s'agit de calculer le produit des deux diamètres maximaux du plan transverse, sur la lésion prenant le contraste en séquence pondérée T1. Mais, sachant que la localisation de la tumeur ne correspond pas exactement à la localisation de la prise de contraste, les critères de RANO prennent en compte les hypersignaux en T2 FLAIR en plus de la prise de contraste en T1[70]. Cette prise en compte de l'imagerie FLAIR a permis à ces critères d'être mieux adaptés pour l'évaluation des GBG.

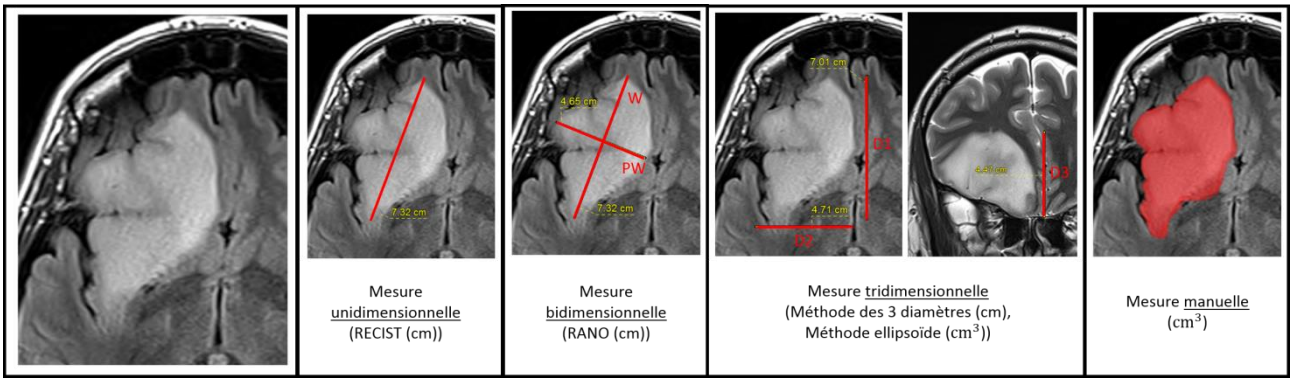


Figure 9: Différentes méthodes pour réaliser le suivi d'évolution des gliomes diffus de bas grade : La méthode unidimensionnelle utilisant le plus grand diamètre sur la coupe où la tumeur est la plus volumineuse. La mesure bidimensionnelle prenant en compte le rapport entre le plus grand diamètre et sa plus grande perpendiculaire au sein de la tumeur. La méthode tridimensionnelle consiste à calculer les deux plus grands diamètres dans la coupe axiale, et le plus grand diamètre dans la coupe coronale. Ces différentes mesures servent à calculer le diamètre moyen de la tumeur, ou son volume (méthode ellipsoïde). La mesure manuelle consiste à tracer coupe par coupe les limites de la tumeur afin d'en obtenir le volume.

Les critères RANO sont donc calculés dans le plan axial, sur les images pondérées T2 FLAIR. La mesure la plus longue sera appréciée par l'opérateur·ice et correspondra au diamètre W. La mesure la plus longue, perpendiculaire à la précédente, correspondra au diamètre PW (*Figure 9*). L'évolution de la tumeur entre deux examens IRM se fera par la comparaison du produit des deux diamètres perpendiculaires :

Équation 1

$$Diamètre\ équivalent = W \times PW$$

Il existe aussi une méthode tri-dimensionnelle appelée « méthode des 3 diamètres ». Elle est utilisée afin de déterminer le diamètre moyen de la tumeur (DMT). Il s'agit de la moyenne des 3 plus grands diamètres dans les coupes axiales, coronales et sagittales (*Figure 9*).

Équation 2

$$DMT = \frac{D1 + D2 + D3}{3}$$

Mais ce DMT est plus précisément calculé à partir du volume. La tumeur est considérée comme un objet de forme ellipsoïde dont la formule du volume est la suivante :

Équation 3

$$V = \frac{4}{3}\pi \times \mathrm{r1} \times \mathrm{r2} \times \mathrm{r3}$$

Pour l'évaluation du volume tumoral des gliomes en recherche clinique, plusieurs approximations ont été réalisées pour aboutir au calcul du volume simplifié suivant :

Équation 4

$$V = \frac{D1 \times D2 \times D3}{2}$$

La méthode de référence pour l'évaluation de l'évolution des gliomes de bas-grade est l'obtention d'un volume le plus proche possible de la réalité : la méthode ellipsoïde est acceptable mais n'égale pas le traçage manuel sur l'imagerie T2 FLAIR[6]. Cette segmentation manuelle est délicate à réaliser car le caractère diffus des gliomes de bas grade la rend fastidieuse, même pour un œil expert.

A partir du volume obtenu, soit par la mesure des 3 plus grands diamètres soit par traçage manuel (*Figure 9*), le DMT est donc calculé de la manière suivante :

Équation 5

$$DMT = (2 \times V)^{\frac{1}{3}}$$

Cette formule est couramment utilisée par des spécialistes en neurochirurgie et neuro-oncologie pour l'évaluation de la croissance tumorale[2,47,65,72].

A partir de ce DMT, la vitesse d'expansion du diamètre moyen (VED) est calculée en millimètres par an. Il s'agit de la pente de la courbe d'évolution du DMT au cours du temps.

La VED calculée à partir du volume tumoral tracé manuellement est la méthode la plus précise pour évaluer la croissance de la tumeur au cours du temps [2,72]. Elle est évaluée à 4mm/an en moyenne pour un gliome diffus de bas grade[47].

Le volume est donc la méthode actuelle la plus fiable pour évaluer la croissance tumorale, mais son estimation la plus précise se fait à partir du tracé manuel qui est une mesure chronophage, et son évolution, la VED, nécessite au moins 2 dates d'examen IRM pour être évaluée.

6. Identification de la transformation anaplasique

La transformation anaplasique désigne le passage d'un GBG (grade 2) en GHG (grade 3). L'identification de cette transformation se fait par un suivi IRM régulier, multiparamétrique.

La transformation anaplasique d'un gliome s'accompagne de la fabrication de nouveaux vaisseaux pour alimenter la tumeur. Cette production de nouveaux vaisseaux se fait de façon anarchique : les vaisseaux néoformés sont malformés et perméables. On assiste alors à une augmentation de la densité vasculaire et à une augmentation de la perméabilité, avec à terme une rupture de la barrière hématoencéphalique.

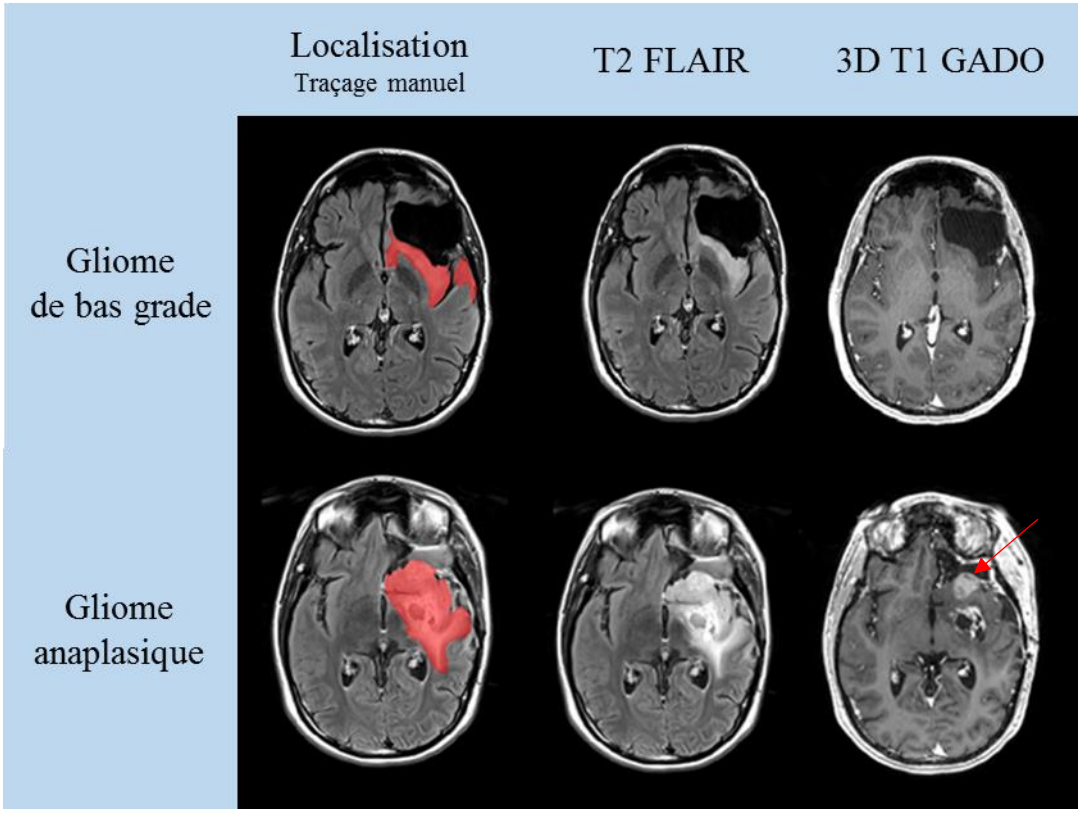


Figure 10 : Différences entre gliome de bas grade (GBG) et de gliome anaplasique (gliome de haut grade (GHG)) sous différentes modalités d'imagerie : La première colonne représente les limitations de la tumeur en rouge. En T2 FLAIR la lésion apparait en hyper signal, avec un signal plus intense pour le GHG. En imagerie 3D T1, avec produit de contraste, le GBG est en hypo signal et le GHG peut présenter une prise de contraste (flèche rouge) ou de la nécrose (taches noires).

Une prise de contraste peut être observée sur l'imagerie T1, dans le cadre du suivi de gliome (*Figure 10, flèche rouge*). La prise de contraste est très sensible à une transformation mais n'est pas spécifique de celle-ci car elle peut être absente chez certains gliomes transformés[40]. L'IRM de perfusion trouve alors tout son intérêt, en mettant en évidence l'absence d'hyperperfusion.

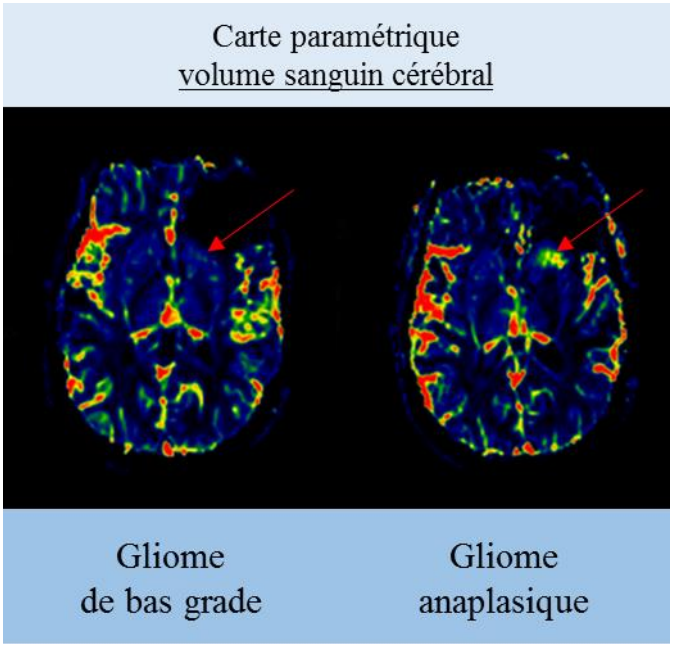


Figure 11 : Carte paramétriques de volume sanguin cérébral (rCBV). Comparaison des signaux de rCBV entre un gliome de bas grade et le même gliome à la date d'identification de la prise de contraste sur l'imagerie 3D T1 GADO. On observe que la lésion n'a pas le même aspect sur ces imageries (flèche rouge), on observe la rupture de la barrière hématoencéphalique caractéristique d'une transformation maligne par la fuite de sang, donc de produit de contraste, dans les tissus.

L'IRM de perfusion permet de mesurer les débits sanguins, le volume sanguin et les temps de transit, rendant compte de l'état de la microcirculation sanguine du cerveau. C'est pourquoi elle est très intéressante pour l'identification de la transformation anaplasique[73]. L'imagerie de perfusion pour l'étude des gliomes de bas grade a mis en évidence la corrélation entre l'augmentation du rCBV et la transformation anaplasique[74,75](*Figure 11*). Plus précisément, on estime que lorsque le rapport du rCBV entre une région tumorale anormale avec un rCBV augmenté divisé par le rCBV dans une zone controlatérale saine (substance blanche ou grise en fonction de la localisation) dépasse 1,75, le gliome se comportera de manière plus agressive et tendra vers une transformation rapide en gliome de haut-grade[8]. Il existe aussi une corrélation entre K2 et transformation, mais elle est moins forte et significative que celle avec le rapport du rCBV[50].

Les œdèmes et les tissus tumoraux apparaissent en hypersignal sur l'imagerie T2 FLAIR. C'est sur cette séquence que l'on identifie le mieux le gliome, et donc à partir de celle-ci que le suivi de la croissance volumique est réalisé (*Figure 10*). La dimension de la tumeur est étroitement liée au pronostic de survie des patient·te·s[76,77]. Une intervention chirurgicale est recommandée en première intention et il a été montré un impact bénéfique sur la durée de survie des patient·e·s si le

résidu tumoral était inférieur à 15mL[26], voire même 10mL[77]. Il est donc recommandé de réduire au maximum le volume tumoral afin d'améliorer le pronostic de survie.

La localisation du gliome et sa croissance entraînant un dépassement de la ligne médiane cérébrale constituent un critère pronostique qui permet d'évaluer à quel point la tumeur va être symptomatique et donc associée à une évolution péjorative. Il a été observé que l'évolution des GBG se traduit en imagerie IRM par une hétérogénéité de signal sur les séquences T2 FLAIR au sein de la lésion tumorale, également observable dans le cas des glioblastomes [78,79].

La vitesse de croissance naturelle d'un gliome de bas grade est estimée par une VED de 4mm/an[47]. L'augmentation de la VED à 8 mm/an (et plus) est corrélée avec l'augmentation du risque de transformation[2] et il existe une accélération de la vitesse de croissance d'un gliome diffus de bas grade dans les mois précédant la transformation anaplasique[80]. Ce facteur, calculé sur 6 mois minimum, est capable d'être plus juste pour prédire la transformation que le seul volume ou l'activité perfusionnelle au sein de la tumeur[81].

La VED est donc un facteur pronostique très pertinent et il est préconisé d'effectuer sa mesure dès le début de la découverte de la lésion[72].

Aujourd'hui, l'évaluation manuelle du volume tumoral comme moyen d'estimer le diamètre tumoral moyen, et ainsi la VED lors du suivi, reste la référence pour quantifier la croissance tumorale[2], mais elle est très chronophage en pratique clinique. L'évaluation de la croissance à partir de données quantitatives est importante, car une simple inspection visuelle ne permet pas toujours de détecter une progression sur deux examens. En effet, espacés de 3 ou 6 mois, deux examens peuvent sembler similaires et une simple analyse visuelle, même faite par un·e neuroradiologue expérimenté·e, peut laisser passer une augmentation de la taille de la tumeur[6]. Le développement d'outils automatiques est donc nécessaire afin de rendre ces mesures quantitatives plus accessibles en pratique clinique.

II. Segmentation automatique des gliomes de bas grade

Les gliomes, de par leur statut infiltrant, diffus et en constante évolution, sont difficiles à segmenter sur la seule base des intensités de l'image. Bien que cela puisse être réalisé par des médecins avertis, la segmentation manuelle reste extrêmement chronophage. Avec le développement des méthodes d'apprentissage dérivées de l'intelligence artificielle (IA), des applications ont naturellement été appliquées à la détection de tumeurs sur la base de l'imagerie.

Au cours des dernières années, des techniques de segmentation automatisée, basées sur l'entrainement à la reconnaissance d'images, ont été développées pour faciliter l'étude de la croissance des gliomes et de leur transformation maligne.

Différentes méthodes de segmentation d'image existent. Certaines se basent sur un seuillage simple des niveaux de gris (méthode d'Otsu)[82], d'autres au contraire utilisent des approches plus complexes nécessitant des connaissances a priori (segmentations probabilistes[83]). Dans notre cas, les GBG ont des localisations extrêmement variées avec un contraste faible et des limites diffuses en rapport avec leur caractère diffus et infiltrant, et ces méthodes trouvent donc leurs limites. A l'heure actuelle, c'est l'œil humain qui est le plus efficace pour segmenter sur la base des imageries FLAIR et T1. Pour mimer le comportement humain, des méthodes d'IA se développent. Souvent basées sur des principes d'application de filtres multi échelles, elles s'entraînent et apprennent à reconnaître les caractéristiques recherchées sur des données connues afin de pouvoir les identifier lorsque de nouvelles données leur sont soumises.

a. Réseau de neurones convolutifs

Une multitude de méthodes existent[84], connues sous le nom d'apprentissage profond (deep learning) avec des spécificités plus ou moins fortes vis-à-vis de l'objectif. On retrouvera des algorithmes de réseaux adversariaux (GAN) créant de nouvelles instances de données qui ressemblent aux données d'apprentissage, des réseaux de mémoire à long terme et à court terme (LSTM) qui sont capables de mémoriser des dépendances à long terme et sont utilisés notamment pour la reconnaissance vocale, les machines de Boltzmann restreintes (RBM) qui sont des réseaux neuronaux stochastiques qui peuvent apprendre à partir d'une distribution de probabilités sur un ensemble d'entrées et sont utilisés pour la classification.

Les réseaux de neurones convolutifs (CNN) quant à eux, se sont avérés particulièrement adaptés à la segmentation d'images.

1. Principe

Les réseaux de neurones convolutifs (CNN) se basent, comme la plupart des méthodes de classification issues de l'intelligence artificielle, sur des phases d'entraînement/validation. Brièvement une base de données dite d'entraînement est fournie à l'algorithme, elle contient les données (pour nous l'image) et les labels (la localisation binarisée de la tumeur). Chaque pixel faisant partie ou non de la tumeur. L'outil s'entraîne ainsi et, en apprenant les caractéristiques des données de niveaux de gris, apprend à classer chaque pixel de l'image tout en optimisant ses paramètres. Ses performances sont ensuite évaluées sur des données nouvelles qui lui sont inconnues (*Figure 12*). Il est ainsi indispensable de lui fournir un nombre de données important et surtout des données variées, sous peine de réduire drastiquement ses performances sur des données différentes de celles avec lesquelles il a appris[85].

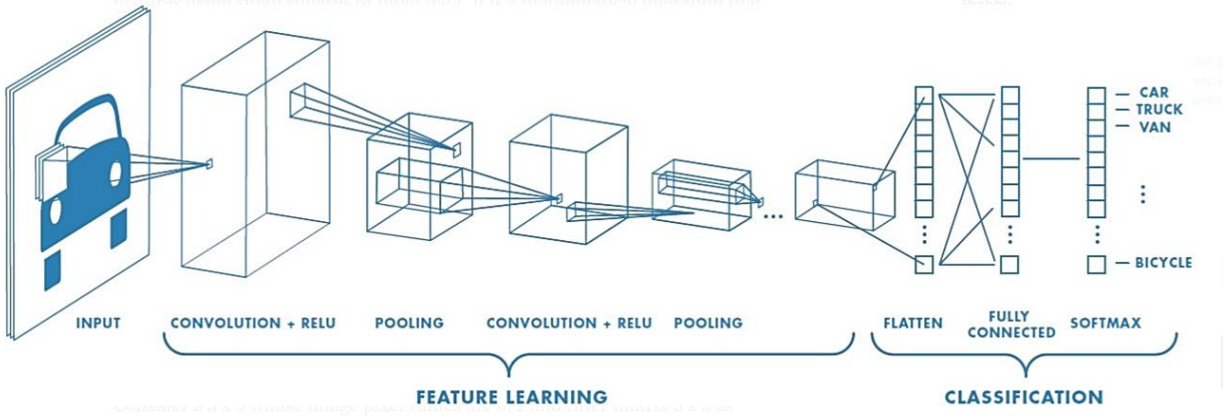


Figure 12: Exemple d'un réseau composé de nombreuses couches convolutives. Des filtres sont appliqués à chaque image d'apprentissage à différentes résolutions, et la sortie de chaque image convoluée est utilisée comme entrée de la couche suivante

Pour la segmentation de lésion tumorale, différentes méthodes existent, comme les modèles statistiques qui peuvent reconnaître les anomalies dans les images[86]. Les CNN sont petit à petit devenus la référence, et avec les développement des technologies et des puissances de calcul, les CNN sont devenus entièrement convolutifs (FCN)[87]. Certaines études ont même exploré des méthodes hybrides qui couplent CNN et modèles statistiques [88–90]

2. L'algorithme U-Net

Sur la base des FCN, l'architecture U-Net a été développée, organisée en une succession de phases de sous-échantillonnage et de sur-échantillonnage[91].

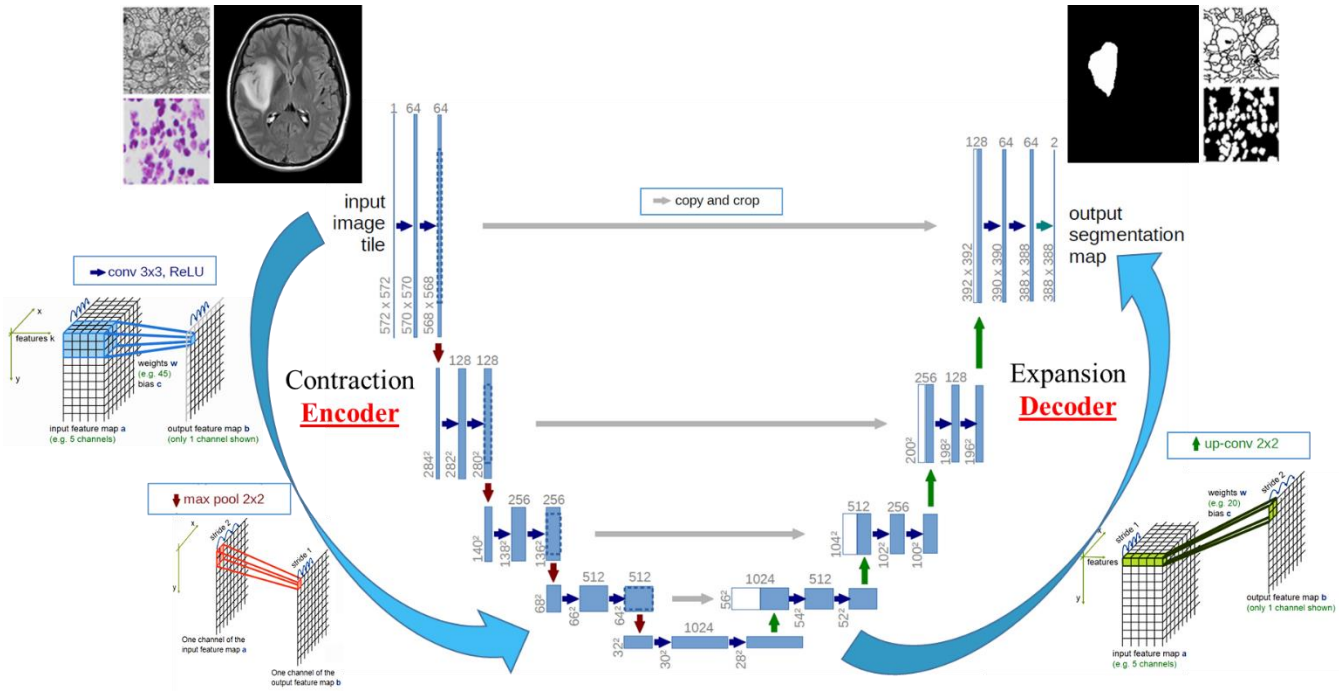


Figure 13 : Architecture du réseau U-Net : Première étape, la contraction, appelée « encoder », où sont réalisés une succession de convolutions (flèches bleues, conv 3x3) et de sous-échantillonnages par la valeur maximale (flèches rouges, max pool 2x2).La seconde étape est une étape d'expansion, appelée « decoder », où les images sont dilatées, sur-échantillonnées, par convolution (flèches vertes, up-conv 2x2), jusqu'à aboutir à une image à la taille originale avec la structure d'intérêt segmentée. Tout au long du processus, les concaténations entre « encoder » et « decoder » permettent de conserver l'information spatiale des structures de l'image.

Le réseau U-Net est composé d'une étape de contraction (« encoder ») suivi par une étape d'expansion (« decoder ») de l'image étudiée (Figure *13*). La contraction est réalisée par plusieurs étapes de convolutions. Chaque niveau est composé de deux convolutions par un filtre de 3x3 et avec la fonction d'activation ReLU (« conv 3x3 », flèche bleue). La fonction d'activation permet de filtrer les valeurs des cartes paramétriques en les rapportant à 0 ou 1. Chaque niveau est constitué de deux couches de convolutions, et le nombre de canaux change de 1 → 64, car le processus de convolution augmente la profondeur de l'image. Entre deux niveaux a lieu un processus de sous échantillonnage par la valeur maximale (« max pooling », flèche rouge) qui réduit de moitié la taille de l'image. Lors de la contraction, les informations structurales de l'image sont mises en évidence mais avec une perte d'information spatiale. La convolution transposée (flèche grise) est une technique de sur-échantillonnage par concaténation qui permet d'augmenter la taille des images et donc de transmettre les informations spatiales des images de l'encoder aux images du decoder.

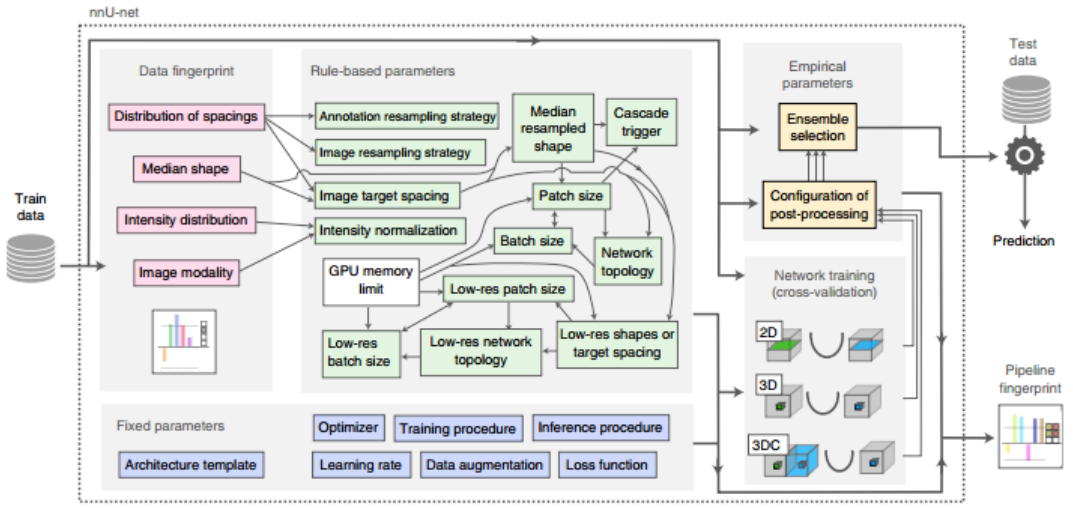
Après la contraction, l'image va être ramenée à sa taille originale lors des étapes d'expansion. Chaque étape de la voie d'expansion consiste en un sur-échantillonnage des cartes paramétriques suivi d'une convolution 2x2 ("up-convolution") qui divise par deux le nombre de

canaux de caractéristiques, d'une concaténation par convolution transposée, et de deux convolutions 3x3, chacune suivie d'une ReLU.

Les approches de segmentation pilotées par l'architecture U-Net ont montré de très bons résultats pour la segmentation des gliomes, indépendamment de leur organisation exacte en nombre de couches ou du type de fonction d'activation utilisée[10,92,93]. Faire varier les paramètres de l'algorithme de segmentation permet de s'adapter aux données d'entrée et de trouver la configuration permettant la meilleure segmentation possible de l'image. Alors que les algorithmes de segmentation permettent le traitement d'images 3D très bien résolues, certaines études nécessitent des algorithmes pouvant s'adapter à des données hétérogènes (différentes machines, champs d'acquisitions, paramètres de séquences…). La performance des algorithmes classiques dépend fortement des propriétés des données et des conditions matérielles. Il a été observé sur une review de 2017 par Litjens et al. que plusieurs études qui utilisent les mêmes algorithmes ont obtenu des résultats très différents[11].

3. Implémentation nnU-Net

C'est à partir de cette problématique que le CNN auto-adaptatif, l'algorithme nnU-Net ("no new net"), a été développé[94]. Le nnU-Net s'adapte automatiquement à des ensembles de données hétérogènes et permet une segmentation « prête à l'emploi ». Il exécute des règles systématiques qui permettent de générer des méthodes d'apprentissage profond, adapté au type d'images et au type de segmentation (tumeur cérébrale, sclérose en plaque, lésion pulmonaire…) sans qu'il soit nécessaire d'optimiser le set de données étudiées. En effet, les images sont reconnues par leur « signature » (Data finger printing) et le pipeline de traitement s'organise en fonction de ces informations (*Figure 14*).



Design choice	Required input	Automated (fixed, rule-based or empirical) configuration derived by distilling expert knowledge (more details in online methods)
Learning rate	–	Poly learning rate schedule (initial, 0.01)
Loss function	–	Dice and cross-entropy
Architecture template	–	Encoder–decoder with skip-connection ('U-Net-like') and instance normalization, leaky ReLU, deep supervision (topology-adapted in inferred parameters)
Optimizer	–	SGD with Nesterov momentum ($\mu = 0.99$)
Data augmentation	–	Rotations, scaling, Gaussian noise, Gaussian blur, brightness, contrast, simulation of low resolution, gamma correction and mirroring
Training procedure	–	1,000 epochs × 250 minibatches, foreground oversampling
Inference procedure	–	Sliding window with half-patch size overlap, Gaussian patch center weighting
Intensity normalization	Modality, intensity distribution	If CT, global dataset percentile clipping & z score with global foreground mean and s.d. Otherwise, z score with per image mean and s.d.
Image resampling strategy	Distribution of spacings	If anisotropic, in-plane with third-order spline, out-of-plane with nearest neighbor Otherwise, third-order spline
Annotation resampling strategy	Distribution of spacings	Convert to one-hot encoding → If anisotropic, in-plane with linear interpolation, out-of-plane with nearest neighbor Otherwise, linear interpolation

Image target spacing	Distribution of spacings	If anisotropic, lowest resolution axis tenth percentile, other axes median. Otherwise, median spacing for each axis. (computed based on spacings found in training cases)
Network topology, patch size, batch size	Median resampled shape, target spacing, GPU memory limit	Initialize the patch size to median image shape and iteratively reduce it while adapting the network topology accordingly until the network can be trained with a batch size of at least 2 given GPU memory constraints. for details see online methods.
Trigger of 3D U-Net cascade	Median resampled image size, patch size	Yes, if patch size of the 3D full resolution U-Net covers less than 12.5% of the median resampled image shape
Configuration of low-resolution 3D U-Net	Low-res target spacing or image shapes, GPU memory limit	Iteratively increase target spacing while reconfiguring patch size, network topology and batch size (as described above) until the configured patch size covers 25% of the median image shape. For details, see online methods.
Configuration of post-processing	Full set of training data and annotations	Treating all foreground classes as one; does all-but-largest-component-suppression increase cross-validation performance? Yes, apply; reiterate for individual classes No, do not apply; reiterate for individual foreground classes
Ensemble selection	Full set of training data and annotations	From 2D U-Net, 3D U-Net or 3D cascade, choose the best model (or combination of two) according to cross-validation performance

Figure 14 : Le nnU-Net (extrait de l'étude de F.Isensee et al.[94]) Pour une nouvelle tâche de segmentation, les propriétés de la base de données sont extraites sous la forme d'une signature du jeu de données ("data fingerprint" en rose). Un ensemble de règles heuristiques modélise les interdépendances des paramètres (flèches fines) et opère sur cette signature pour en déduire les "paramètres basés sur les règles" (vert) du pipeline qui dépendent des données. Ceux-ci sont complétés par des "paramètres fixes" (bleu), qui sont prédéfinis et ne nécessitent aucune adaptation. Jusqu'à trois configurations sont entraînées dans une validation croisée à cinq voies. Enfin, nnU-Net effectue automatiquement une sélection empirique de l'ensemble optimal de ces modèles et détermine si un post-traitement est nécessaire ("paramètres empiriques", jaune). Le tableau du bas montre les valeurs explicites ainsi que les formulations de règles résumées de tous les paramètres configurés (res= résolution).

Le nnU-Net s'est situé parmi les meilleures performances à plusieurs concours de segmentation automatique de données biomédicales, comme le *Medical Segmentation Decathlon*[95]. Il a obtenu la meilleure performance pour le « *Multimodal Brain Tumor Segmentation Challenge 2020* » (BraTS'20)[96] ; en effet cet algorithme a obtenu de très bons résultats pour la segmentation de lésions tumorales cérébrales avec un coefficient de similitude de Dice (DSC) de 0.89 (deux masques identiques auront un DSC de 1)[97].

En ce qui concerne la segmentation automatique de gliomes de bas-grade (GBG), de manière générale, les études présentant les meilleures performances sont entraînées et testées avec des cohortes de grades mixtes et incluent une majorité de gliomes de haut grade (GHG)[10,87,92,93,98,99] .

Très peu d'études se concentrent sur une segmentation automatisée spécifique aux GBG[100,101]. En ce qui concerne les GBG, la complexité de leur segmentation réside dans leurs contours flous, en raison de leur statut infiltrant, diffus et en constante évolution. De plus, la méthode automatique développée pour la segmentation des GBG doit être adaptée à la réalité clinique, c'est-à-dire à des acquisitions plus ou moins bien résolues (2D, 3D), avec des paramètres d'acquisitions IRM différents, et obtenues à partir de machines IRM différentes (constructeurs, champ magnétique).

La plupart des bases de données d'entraînement des algorithmes de la littérature comprennent une grande majorité d'images acquises sur les mêmes machines, à des champs magnétiques identiques, à haute résolution. De plus ce sont souvent des images obtenues à la phase pré-chirurgicale, en l'absence de cavité chirurgicale, ce qui rend la segmentation très délicate. Ces données sont donc malheureusement loin de la réalité des données issues du suivi clinique des GBG. Ces limites imposent d'avoir une population d'entraînement pour le modèle de segmentation automatique très proche de la réalité clinique et radiologique.

L'algorithme nnU-Net a été sélectionné pour la suite de ce travail afin de mettre en place l'outil de segmentation automatique des gliomes de bas grade. L'implémentation et l'optimisation de la méthode ont été réalisées via une collaboration entre l'Institut d'Imagerie Fonctionnelle Humaine (I2FH), situé au sein du service de neuroradiologie de l'Hôpital Gui de Chauliac à Montpellier, et l'équipe Siemens Healthineers Lausanne de Tobias Kober, Bénédicte Maréchal et Till Huelnhagen.

b. Données de référence : segmentation manuelle

Comme décrit précédemment, afin d'entraîner l'algorithme automatique, des données de référence doivent être définies. Les données de référence proposées dans ce travail sont les contours de lésions cérébrales de GBG tracées sur une image nifti ou DICOM, et ces contours sont exportés sous forme de masque binaire.

Pour cette étude, les segmentations manuelles de tumeurs ont été réalisées à partir des imageries T2 FLAIR de chaque examen IRM grâce au logiciel MRICron[102], et réparties entre 5 neuroradiologues ayant un minimum de 2 ans de pratique. Pour éviter les erreurs de segmentation manuelle, tous les masques ont été vérifiés par une tierce personne experte.

c. Population

La base SPECIFY (Suivi longitudinal rétrospectif de la Perfusion des gliomes cérébraux et confrontation aux données de Croissance tumorale en Imagerie par Résonnance Magnétique, ainsi qu'aux autres Facteurs pronostiques de la transformation anaplasique) a été mise en place par le service de neuroradiologie de l'hôpital Gui de Chauliac afin de collecter des données cliniques, anatomopathologiques, génétiques et radiologiques dans le suivi de gliomes de bas grade cérébraux, associant les données de perfusion et permettant ainsi d'évaluer la croissance tumorale de manière longitudinale (Clinicaltrials.gov : NCT04346472). C'est une base de données rétrospective qui permet une visualisation complète de l'historique médical des patient·e·s : les différentes IRM (date, machine, différentes séquences, diagnostic), les données cliniques, la ou les opérations subies (date, localisation, partielle ou totale, déficit pré et post-opératoire), les résultats anatomopathologiques de la tumeur excisée (histologie, génétique) et le descriptif des différents traitements de chimiothérapie et radiothérapie.

Pour l'entraînement de l'algorithme de segmentation automatique, 30 suivis de patient·e·s ont été sélectionnés au hasard parmi les 200 présents sur SPECIFY. Les examens sélectionnés dans le suivi pouvaient avoir été réalisés avant ou après la chirurgie. Ils devaient comporter au moins 1 examen avant tout traitement par chimiothérapie ou radiothérapie. L'examen IRM sélectionné devait aussi comporter les séquences T2 FLAIR (3D si disponible, sinon 2D), 3D T1 GADO et imagerie de perfusion.

La performance de l'algorithme a été évaluée à partir de deux populations sélectionnées selon les mêmes critères que les données d'entraînement. La première représente le suivi longitudinal de 6 sujets (69 examens IRM) et la seconde correspond à 49 sujets avec 2 dates d'examen IRM

chacun. Cette base de données a été utilisée pour la totalité des travaux de thèse et correspond à un total de 85 sujets, représentant 485 examens IRM (*Table 1, II.d.Article - Evaluation of an nnU-Net type automated tumor segmentation as a clinical volumetric tool for low-grade glioma follow-up*).

d. Article - Evaluation of an nnU-Net type automated tumor segmentation as a clinical volumetric tool for low-grade glioma follow-up

Evaluation of an nnU-Net type automated tumor segmentation as a clinical volumetric tool for low-grade glioma follow-up

Margaux Verdier[1,4], Jeremy Deverdun[1] PhD, Justine Belko[1], Thierry Chaptal[1], Nicolas Menjot de Champfleur[1,2,5] MD, PhD; Hugues Duffau[3,6] MD, PhD, Thomas Troalen[4] PhD, Bénédicte Maréchal[7,8] PhD, Till Huelnhagen[7,8] PhD, Emmanuelle Le Bars[1,2] PhD

[1]I2FH, Institut d'Imagerie Fonctionnelle Humaine, Department of Neuroradiology, Montpellier University Medical Center, Montpellier, France; [2]Department of Neuroradiology, Montpellier University Medical Center, Montpellier, France; [3]Department of Neurosurgery, Montpellier University Medical Center ; [4]Siemens Healthcare SAS, Saint-Denis, France; [5]Laboratoire Charles Coulomb, University of Montpellier, France ; [6]Institute for Neuroscience of Montpellier, INSERM U1051, Montpellier University Medical Center, Montpellier, France ; [7]Advanced Clinical Imaging Technology, Siemens Healthcare AG, Lausanne, Switzerland. [8]LTS5, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Introduction

Gliomas represent 50% of all primary brain tumors[1]. Low grade gliomas (LGG), classified grade 1 and 2 by the World Health Organization[2], are characterized by a slow and continuous growth preferentially along cerebral white matter tracts [3]. LGG always evolve toward an aggressive grade, but this malignant transformation can take a long time, resulting in long term monitoring over years. Several predictors of a malignant transformation exist such as the presence of contrast enhancement areas in T1-weighted MRI which is related to local blood-brain barrier disruption, but a non-enhancing lesion does not always show an absence of histological correlation with neo-angiogenesis. Another transformation predictor is tumor volume, as larger tumors are more likely to transform[4]. Nevertheless, the most precise predictor is the velocity of diameter expansion (VDE) [5] where patients with a VDE higher than 8mm/year are considered as already transformed. Lesion volume and VDE are commonly quantified using T2-weighted FLAIR acquisitions. However, the infiltrative nature and blurred contours of LGG make the lesion segmentation challenging compared to other tumor types[6]. VDE assessment can therefore require a large amount of time even for an experienced radiologist or neurosurgeon, limiting its systematic use in clinical practice.

In recent years, a rapid evolution of automated segmentation techniques has been observed to facilitate the study of glioma growth and their malignant transformation with increasing performance[7–13]. The U-Net architecture was developed, marked by a succession of down-sampling and up-sampling phases[14], showing good results independent of their organization in number of layers or activation functions[15–17]. In this context, the nnU-Net framework, a self-adapting CNN, was recently proposed and has shown good performance for image segmentation tasks[18]. This framework combines the U-net architecture with data preprocessing techniques to improve efficiency and simplify application[19]. While several automated solutions for glioma segmentation have been suggested, most studies evaluating their performance used mixed grades cohorts with a majority of high-grade glioma (HGG)[8,11,15–17,20]. Few studies focused specifically on automated segmentation of LGG and mainly based their training on pre-surgical images[12,13]. However, clinical reality shows frequent preventive surgical interventions to delay malignant transformation[21]. Therefore, this study evaluates the performance of an nnU-Net, trained on routine clinical LGG

images including pre- and post-surgical images to automatically segment LGG lesions for longitudinal follow-up.

Methods

485 MRI exams of 85 LGG patients were included (**Table 1**). Participants were selected randomly from the SPECIFY database (NCT04346472_UF9647). SPECIFY was developed in 2015 by the Montpellier University Hospital and includes imaging and clinical data of LGG patients between 2009 and 2021. All participants gave informed consent.

This study consists of three phases. First, the algorithm was trained with MRI data (2D/3D T2 FLAIR and contrast enhanced 3D T1-weighted) and *manual segmentation* (MS) of the LGG lesion by experienced neuroradiologists from a *training dataset*, in order to generate an *automated segmentation* (AS). Second, the performance of the trained algorithm was evaluated in an *evaluation cohort* for LGG follow-up, compared to MS. Third, a *clinical validation cohort* was created to assess the performance of the AS in clinical practice. In this phase, manually corrected segmentations (CS) were created by adding or removing areas from the AS when deemed necessary by the expert reader.

Segmentations:

Manual tumor segmentation was performed on T2 FLAIR images, distributed among 5 experienced neuroradiologists with a minimum of 2 years of practice. Both MS and CS were realized with MRIcron[22]. The MS tumor masks were used as ground truth for the *training* and the *evaluation* of the CNN, and the CS was used as ground truth for the *clinical validation.* To avoid manual segmentation errors, all MS were checked by another expert reader.

A self-adapting CNN with a 2D U-Net architecture was used for automated LGG segmentation. The nnU-Net was specifically designed for biomedical data with an optimized and standardized pipeline. It considers data properties, with a “dataset fingerprinting” step, and dynamically optimizes the CNN [19]. The 2D model is less resource intensive and allows faster training and inference compared to the 3D model. Prior to training, the 3D T1-weighted images were co-registered to the T2 FLAIR images using affine transformation. Both were then used as

input channels to the network, together with the corresponding manual segmentation (MS). The trained network was then used to generate automated segmentations (AS). During the training process, the *training dataset* (n=318) was split by the nnU-Net into an internal training (n=254) and a validation set (n=64).

Training dataset:

The training data set was composed of 318 MR exams (30 patients, ~11 exams per patient). Each MRI exam included post-contrast 3D T1-weighted and T2 FLAIR imaging with varying MRI systems and field strengths: 53% of the exams were performed at 3T (MAGNETOM Skyra, Siemens Healthcare, Germany; Ingenia, Philips Healthcare, Netherlands), and 47% at 1.5T (INTERA, Philips Healthcare, Netherlands; Optima, GE Healthcare, USA; MAGNETOM Avanto and Aera, Siemens Healthcare, Germany) (**Table 1**).

Evaluation cohort

69 MR exams corresponding to the complete follow-up of 6 randomly selected LGG patients (11.5±3.3 exams/patient) were included (**Table 1**). Lesion masks were automatically created by the nnU-Net trained as described above. For both AS and MS, tumor volumetric values (cm^3) were extracted and plotted over time. Volumetric values were used to estimate the mean tumor diameter (MTD (1)).

(1) $$\text{Mean tumor diameter (mm)} = (2 \times Volume)^{\frac{1}{3}} \times 10$$

The MTD was used to calculate the velocity of diameter expansion (VDE (2)), in millimeters per year, for AS and MS, between two time points (*time* in years).

(2) $$\text{Velocity of Diameter Expansion (mm/year)} = \frac{\Delta_{MTD}}{\Delta_{time}}$$

The VDE was plotted for every patient over time and was calculated for a minimum period of 6 months [23]. When participants had an initial surgery, VDE was determined from the first exam

post-surgery and in case of a chemotherapy, VDE was calculated between the first and the last MRI exam acquired during treatment period.

Clinical validation cohort

The clinical validation cohort was created to assess the AS performance in routine clinical practice. It was composed of 98 MRI exams of 49 patients (mean age 41 ± 8, 2 MRI exams per patient) that were submitted to the trained nnU-Net (**Table 1**). The 98 generated AS were then manually corrected to obtain the resulting CS.

	Training set (n = 30)	**Evaluation cohort (n = 6)**	**Clinical validation cohort (n = 49)**	
MRI exams (n)	318	69	98	
MRI exams per patient (mean (± sd))	11.2 (± 4.5)	11.5 (± 3.3)	2 (± 0)	
Demographics				
Sex (% (n))				
Male	57% (17)	50% (3)	53% (26)	
Female	43% (13)	50% (3)	47% (23)	
Age (mean (± sd))	35 ± 7	45 ± 11	41 ± 8	
MR field strength 3T (% (n))	56% (177)	38% (26)	59% (58)	*P-value*
MR vendors (% (n))				
Philips	1% (2)	-	-	0.62
Siemens	99% (175)	100% (26)	100% (58)	
MR field strength 1.5T (% (n))	44% (141)	62% (43)	41% (40)	
MR vendors (% (n))				
Philips	5% (7)	7% (3)	-	
GE	3% (4)	9% (4)	-	0.07
Siemens	92% (130)	84% (36)	100% (40)	
MRI Sequence				
T2 FLAIR (% (n))				
2D (resolution (mm^3))				
0.5×0.5×5	0.3% (1)	4.5% (3)	-	
0.9×0.9×5	29.5% (94)	42% (28)	92% (34)	
0.9×0.9×4	13.5% (43)	19% (13)	92% (6)	
0.75×0.75×3	50% (159)	34% (23)	92% (50)	0.38
3D (resolution (mm^3))				
1×1×1	6.7% (21)	2% (2)	8% (8)	
T1-weighted. gadolinium-enhanced (% (n))				
3D (1×1×1 mm^3)				
Turbo Gradient-Echo	89% (283)	81% (56)	91% (89)	
Gradient-Echo	10% (34)	17% (12)	5% (5)	0.05
Turbo Spin-Echo	1% (4)	2% (1)	4% (4)	

Table 1: Data description: *The* ***Training set*** *describes the training data of the automated segmentation (AS) algorithm composed of 318 MRI exams of 30 patients. To evaluate the trained algorithm for longitudinal follow-up, AS were generated for the entire follow up of 6 patients comprising a total of 69 MRI exams (****Evaluation cohort****). An evaluation of the AS for routine*

*clinical use was conducted in a wider population of 49 patients, with two MRI exams each (**Clinical validation cohort**). To assess the similarity between the various cohorts, the P-value was calculated with a Chi-squared test with regard to used magnetic field strength, machine vendors, T2 FLAIR or T1-weighted acquisition protocols. At a significance level of p=0.05, a significant difference in T1-weighted acquisition protocols was found between the data sets. The 3 other tests showed no significant differences between the cohorts.*

Evaluation metrics:

On a per-voxel basis, true positive (TP), true negative (TN), false positive (FP) and false negative (FN) values of the AS compared to the ground truth (MS or CS) were assessed. The Dice similarity coefficient (DSC) was calculated to measure the overlap between the manual and automated segmentation, as well as the AS sensitivity. The sensitivity reflects the probability for a voxel to be included in the AS when present in the ground truth.

The AS' performance to assess the volumetric follow-up equivalently to the ground truth was evaluated by Lin's concordance correlation coefficient[24] and a Bland-Altman test[25]. These tests were also calculated to evaluate the correlation between the AS and the MS-based VDE.

A Chi-squared test was performed to evaluate the similarity between cohorts regarding magnetic field strengths (1.5T/3T), vendors (Siemens/Philips/GE) and acquisition types (2D/3DFLAIR and 3D T1).

Results

The Chi-squared tests for the different 3T vendors (P-value = 0.62), 1.5T vendors (P-value = 0.07) and for the proportions of 2D/3D FLAIR (P-value = 0.38) showed no significant differences between cohorts. Cohorts differed significantly only for 3D T1 acquisition types (P-value = 0.05), yet with a majority of turbo gradient-echo sequences within each cohort (**Table 1**).

Evaluation cohort

The visual evaluation of the MS highlighted a segmentation error for patient 3 for which the neuroradiologist was misled by a large flow artifact present post-surgery and over the follow-up scans. Such artifacts can appear in 2D FLAIR images when the surgical cavity is linked to the

ventricles. The final manual ground truth segmentation in this patient was therefore based on the consensus between 3 other neuroradiologists.

The mean global sensitivity of the AS produced by the nnU-Net was 0.76 ± 0.14. The mean global DSC was 0.82 ± 0.13 for the AS results, across all patients and exams (range 0.68 ± 0.26 to 0.87 ± 0.03). (**Table 2**). **Figure 1** shows the tumor volume evolution for each patient, with indication of surgical interventions, if present. The MS-based volume estimations were systematically larger than those based on the AS.

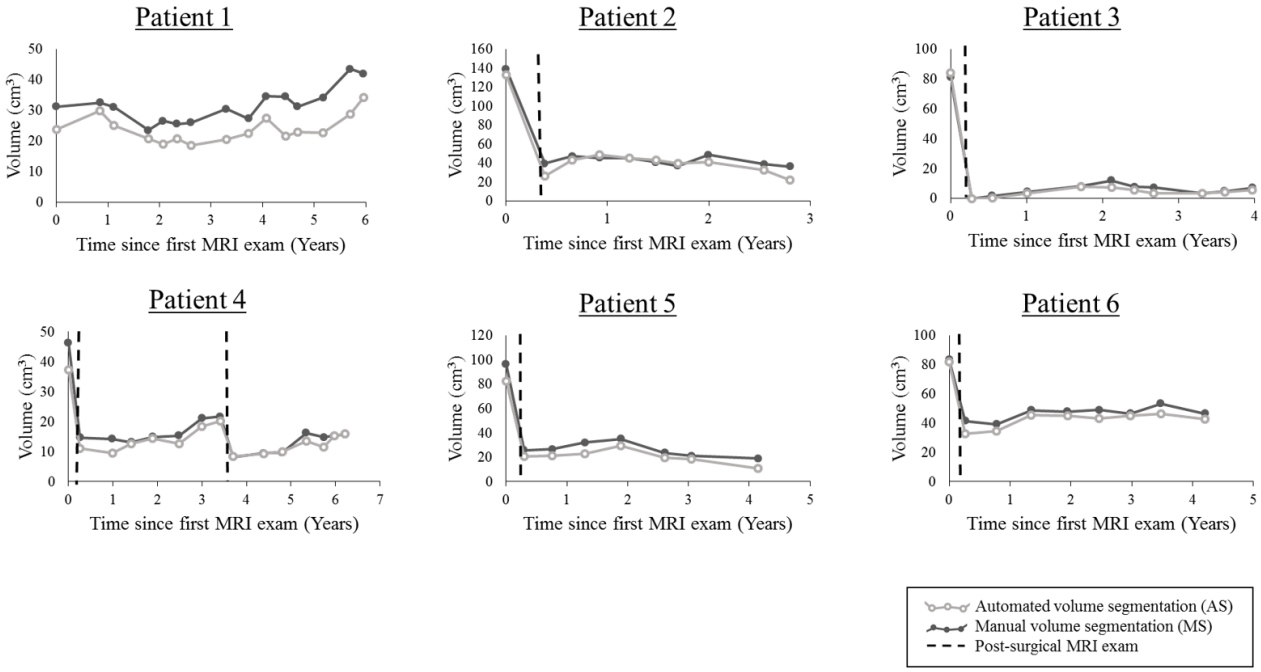


Figure 1: *Evaluation cohort: Evolution of the tumor volume over time. For each patient, the variation of the tumor volume over time was plotted for the automated (AS, empty dots) and manual segmentation (MS, full dots). In case the patient underwent a surgery (patient 2 to 6), the post-surgical exam is marked with a dashed line.*

The Bland-Altman analysis confirmed an average underestimation of the AS-derived tumor volume of 4.2 cm^3 (**Figure 2A**). Still, both methods showed an excellent concordance of 0.96 (Lin's correlation coefficient). The lowest mean DSC was found for patient 3 (0.68±0.26) (**Table 2**). At time point 2, after surgery, both manual and automated methods didn't find a tumor residue (**Supplementary material, Table1**). At time point 3, a small residue was manually identified that the AS failed to segment. Instead, it segmented a stronger hyperintensity that was found elsewhere in the brain distant from the tumor site. At time point 8, the AS failed to segment tumor areas with the lowest FLAIR hyperintensities **(Figure 3A)**. Still, the DSC ranged from 0.75 to 0.95 for 88% of

the data (**Supplementary material, Table 1**), highlighting a substantial overlap between MS and AS (**Figure 3B)**.

The VDE plots over time (**Figure 4**) to illustrate the tumor evolution along the follow-up, during or after a treatment (surgery, chemotherapy) showed a substantial concordance between MS and AS with a Lin's correlation coefficient of 0.7. The Bland-Altman analysis confirmed a slight VDE underestimation of 0.1mm/year for the AS (**Figure 2B**) and highlighted two extreme points. These points are from patient 2 before and after chemotherapy. Between points A and B, the MS provided a VDE of 3.81 mm/years while due to errors in the AS at time point A (DSC of 0.78), the VDE calculation from the automated method is higher with 15.93 mm/years. Same errors are found between time points C and D with a VDE of 3.42 for the ground truth and 19.67 for the automated method, also due to AS errors (DSC of 0.74) (**Figure 4**). Apart from these two points the AS provided VDE calculations comparable to the MS.

Clinical validation cohort

The mean volume segmented by AS and CS were respectively 32.83 cm^3 (± 27.09) and 33.72 cm^3 (± 28.05) including small and large lesions. The mean DSC was 0.94 ± 0.13 and the median DSC was 0.98 ± 0.13. The mean sensitivity was 0.91 ± 0.17 (**Table 2**). The Bland-Altmann analysis showed a slightly higher volume for the CS method with an average difference of 1 cm^3 and limits of agreement between +9.1 to -7.1 cm^3 (**Figure 2C**). For 81% of the dataset, the DSC was >0.9, for 15% between 0.9 and 0.7, for 1% between 0.7 and 0.5, and 3% showed a DSC <0.5 (**Figure 5**). Manual correction of the AS was necessary in 19% of the cases, with a DSC<0.9 and a major correction for only 4% (3 exams) with a DSC<0.7.

		Segmentation results (± sd)				Lin's concordance correlation coefficient		
		Mean global DSC	Mean DSC/patient	Mean global sensitivity	Mean sensitivity/patient	**Volume**	**MTD**	**VDE**
Evaluation cohort (MS/AS)	Patient 1	0.82 ± 0.13	0.83 ± 0.04	0,76 ± 0.14	0.73 ± 0.07	0.96 Excellent	0.96 Excellent	0.7 Substantial
	Patient 2		0.87 ± 0.06		0.83 ± 0.12			
	Patient 3		0.68 ± 0.26		0.63 ± 0.27			
	Patient 4		0.84 ± 0.05		0.81 ± 0.07			
	Patient 5		0.84 ± 0.06		0.75 ± 0.09			
	Patient 6		0.87 ± 0.03		0.83 ± 0.05			
Clinical validation cohort (CS/AS)	Whole population : 49 patients, 98 MRI exams	0,93 ± 0,13		0,91 ± 0,17		0,99 Excellent		

Table 2: *Evaluation cohort results: The Dice Similarity Coefficient (DSC) and sensitivity between MS and AS were calculated. For all the AS across all patients, the mean global DSC was 0.81 ± 0.14. The Lin's concordance correlation coefficient between MS and AS was 0.96 for both volumes and Mean Tumor Diameter (MTD) calculation, which is an*

excellent result. For the VDE the concordance correlation coefficient was 0.7 between the two methods which is a substantial concordance.

Clinical validation cohort results: The mean Dice Similarity Coefficient (DSC) of the whole clinical validation cohort was 0.93 ± 0.13 and the mean sensitivity was 0.91± 0.17. The Lin's concordance correlation coefficient between the corrected (CS) and the automated segmentation (AS) was excellent with 0.99.

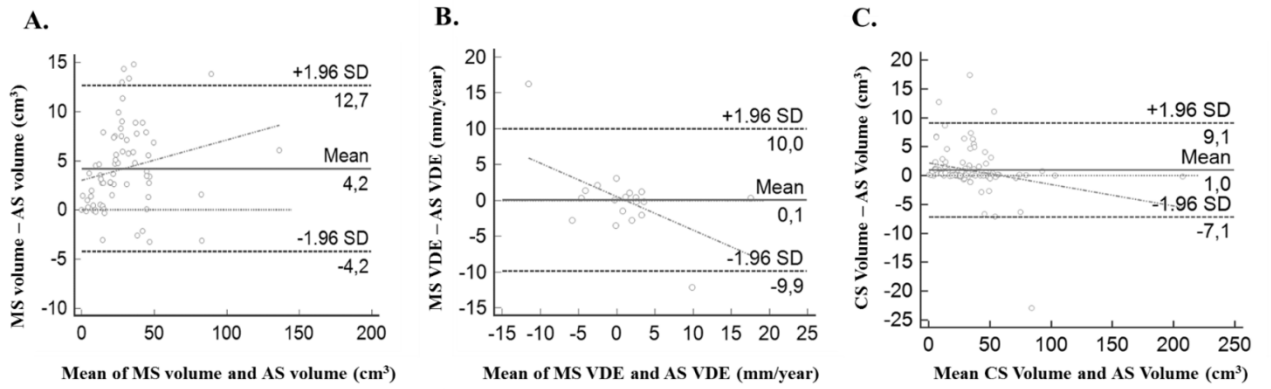


Figure 2: *Bland-Altman plots*

A*. The segmented volumes for the evaluation cohort: Manual segmentation (MS) and automated segmentation (AS) volumes showing a systematic underestimation of the AS of 4.2 cm³.*

B*. The VDE calculations based on the segmented volumes for the evaluation cohort: Manual segmentation (MS) and automated segmentation (AS) velocity of diameter expansion (VDE) showing a slight underestimation of the AS of 0.1 mm/year.*

C*. The segmented volumes for the clinical validation cohort: Comparison of the automated segmentation (AS) and corrected segmentation (CS) showing a slight underestimation of the AS volumes of 1 cm³.*

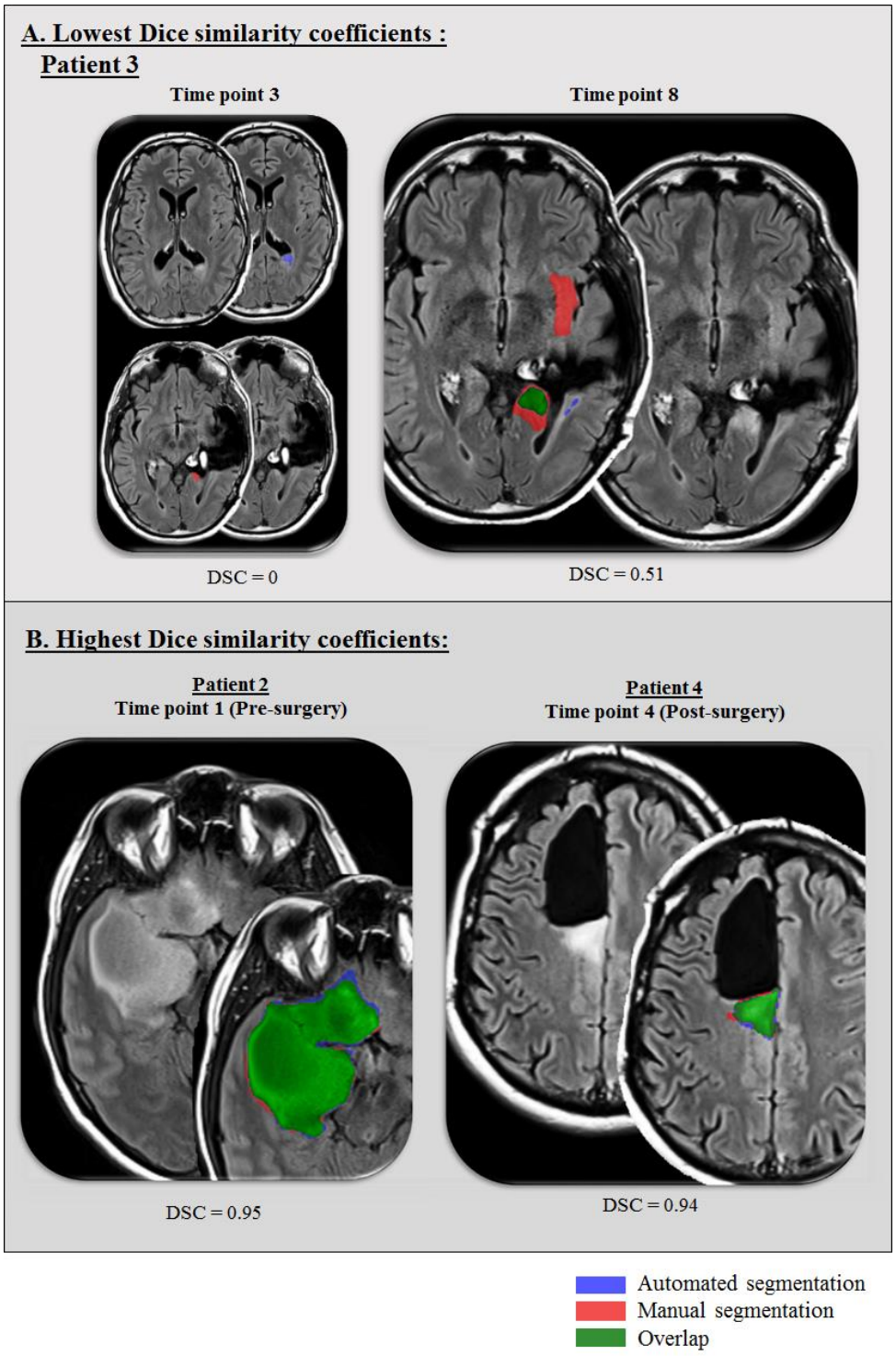


Figure 3: *Example segmentation results for the evaluation cohort*

A. *Lowest Dice similarity coefficients (DSC), the case of patient 3 (mean DSC = 0.68±0.26).*

Time point 3: The automated segmentation (AS) failed to find the tumor residue close to the surgical cavity identified by the manual segmentation (MS, in red) and it falsely segmented a FLAIR hyperintense area (blue), identified as a flow artifact due to the connection between the ventricles and the surgery cavity.

Time point 8: Diffusely hyperintense tumor regions were missed by the AS.

B. *Cases showing the highest Dice similarity coefficients (DSC): Left: Pre-surgery case, patient 2 time point 1 (DSC = 0.95). Right: Post-surgery case, patient 4 time point 4 (DSC= 0.94).*

The true positive values (green) represent the overlap between the automated segmentation (AS) and the manually segmented tumor volume (MS). The MS is considered as ground truth. False negative voxels (red) are present in the MS but not selected by the AS while false positive voxels (blue) are only selected by the AS. The AS performed well in both cases irrespective of whether the MRI exam being pre or post-surgery.

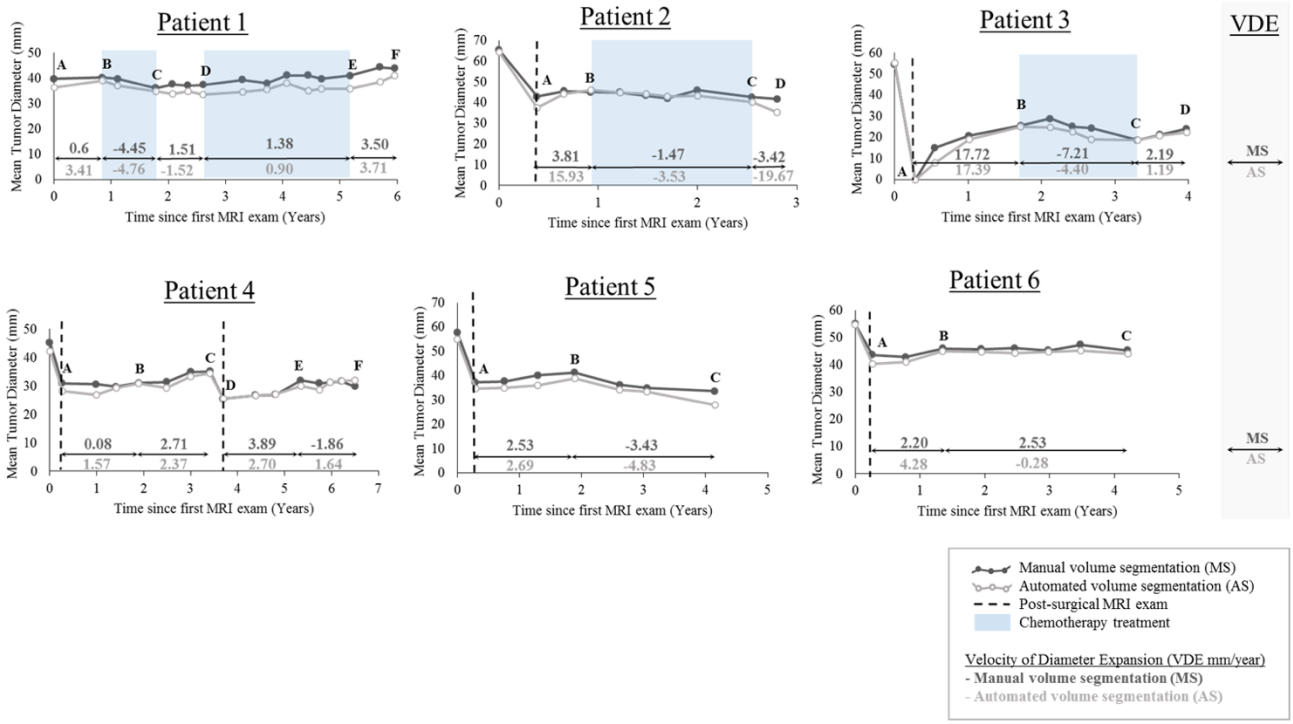


Figure 4: *Evaluation cohort: Mean tumor diameter (mm) for automated (empty dots) and manual segmentation (full dots) over the years, the Velocity of diameter expansion (VDE) is calculated between specific time points. The VDE describes the tumor diameter evolution between two MRI exams and evaluate the tumor evolution along the follow-up, to evaluate the need of a treatment (surgery, chemo or radiotherapy) or its efficiency. Periods of chemotherapy treatment are highlighted in blue. In case there was a surgery, the post-surgical exam is marked with a dashed line. In case of an initial surgery (patient 2 to 6), the VDE is calculated starting from the post-surgery MRI exam (A). In case there was a surgery during the follow-up, the VDE is calculated until the pre-surgery exam (patient 4, C) and the next VDE calculation starts from the post-surgery exam (patient 4, D).*

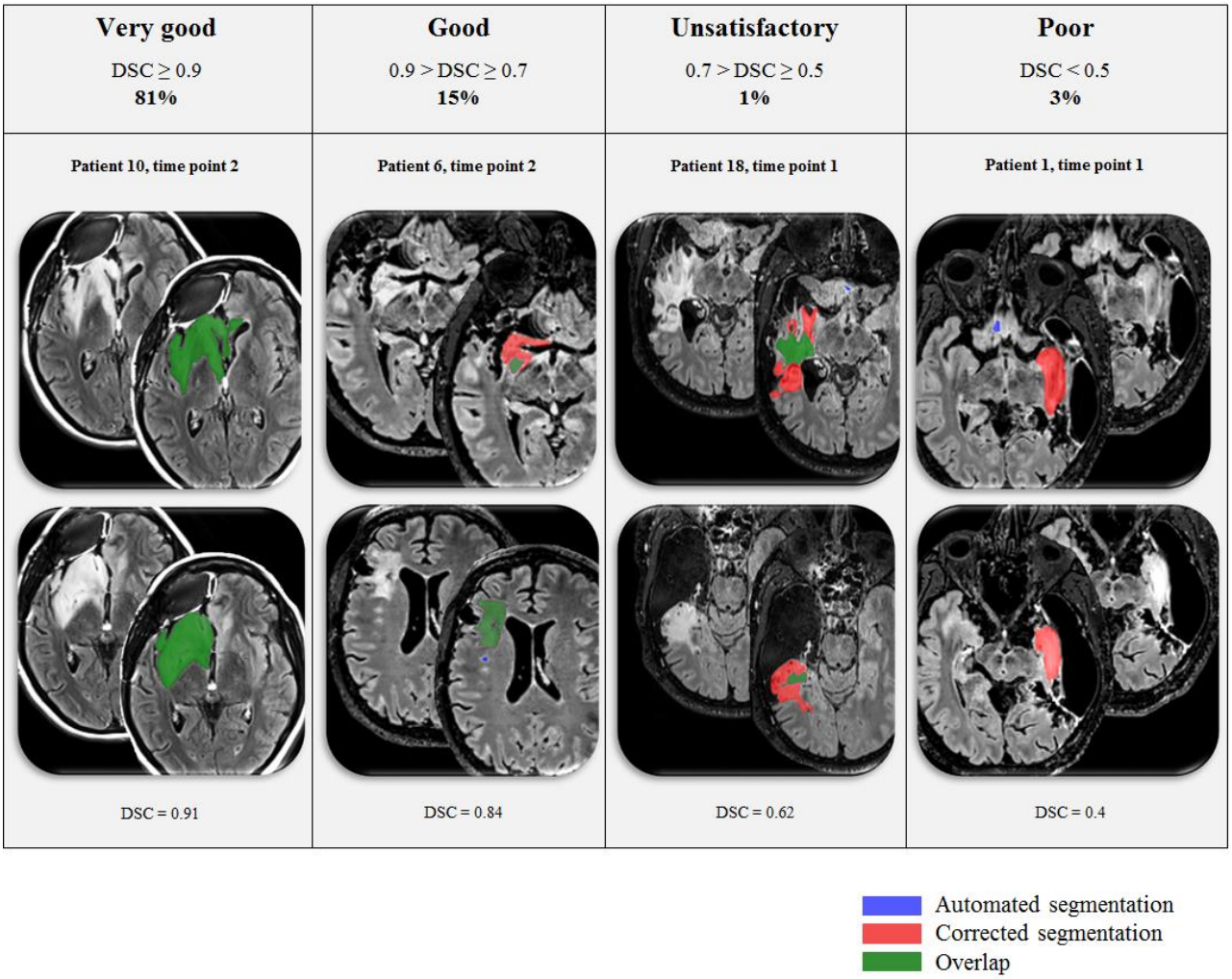


Figure 5: *Results of the clinical validation. For the validation, the automated segmentation (AS) of 98 MRI exams (49 patients, 2 time points per patient) were corrected, if necessary, by an experienced neuroradiologist. The AS and the corrected segmentation (CS) were compared by Dice Similarity Coefficient (DSC). For each patient shown in this figure, different slices of the FLAIR images of one time point are displayed. The overlap between AS and CS, i.e. the uncorrected area, is shown in green. In blue, the segmentation removed and in red the one added by the corrector (AS errors) are indicated. In this clinical evaluation cohort, 81% of the exams had a DSC higher than 0.9 which is a very good performance. 15% had a good performance with DSC between 0.7 and 0.9. For 1 among the 98 CS (1%), the segmentation showed unsatisfactory performance with a DSC between 0.5 and 0.7. For 2 of the cases (3%) the performance was poor with a DSC lower than 0.5.*

Discussion

We evaluated the feasibility of automated tumor segmentation with an nnU-Net-based algorithm for longitudinal LGG follow-up. Until now, only few studies focused specifically on LGG and their progression over time [12,13]. In the evaluation cohort, the performance of the produced

AS was good compared to the ground truth (mean DSC 0.82). Mean DSCs described in the literature are slightly higher, ranging from 0.85-0.91 but these studies included mainly high-grade glioma (HGG)[8,11,15–17,20] which can be more easily delineated. The AS was strongly correlated with MS tumor volume, yet the AS had a tendency to underestimate the volume. Surgery did not seem to influence the AS performance. The proportion of false negatives in the AS was low, except for one subject in which the AS failed (mean DSC of 0.68). For this patient, a flow artifact appeared in the FLAIR images. These cases are common in 2D FLAIR acquisition and can be tricky for the manual evaluation. Nevertheless, the AS avoided errors in this artifact due to the use of the 3D T1 acquisitions where the artefact is not present. The AS tended to fail segmenting the lowest intensities of the tumor. Nevertheless, the majority of subjects showed DSC values ≥ 0.83.

By our knowledge, nnU-Net's ability to estimate VDE, being the best indicator of LGG evolution[5], compared to ground truth clinical data has not been evaluated before. We demonstrated that VDE follow-up with AS was concordant with MS, with a Lin's correlation coefficient of 0.7, confirming the capacity to estimate the VDE based on AS extracted volume. VDE calculations showed little differences between MS and AS apart from two points. This variability may be related to LGG's infiltrative character, known to cause volume delineation variability even between experienced neuroradiologists[26]. Small differences in volume measures can strongly impact the VDE over time. Nevertheless, we confirmed nnU-Net's capacity to segment LGG, despite tumor heterogeneity. The AS, when available in clinical routine, could be a time-saving tool, although segmentation maps should always be visually checked by a radiologist before diagnosis.

The clinical validation of the AS compared to its correction (CS) confirmed the tendency of the AS to underestimate the tumor volume. The mean DSC was high with 0.93. More importantly, using the AS as a guidance for manual tumor contouring reduced the mean difference and amplitude of volume variations. In 81% of the cases the AS required little to no corrections, and in an additional 15% only minor corrections were required, as indicated by a DSC between 0.7-0.9. Only 4% of the images required major corrections. Thus, the LGG trained nnU-Net, could be a reliable tool enabling an easier and faster LGG segmentation in clinical practice.

The dataset used to train the nnU-Net was on the one hand homogeneous, due to the longitudinal character with ~11 exams per subject, and on the other hand heterogeneous with variable MR field-strength and machine vendors. Both reflect clinical reality. Larger multi-centric

validation studies are required taking into account both homo- and heterogeneity aspects before introducing the nnU-Net as a common automated tool for LGG assessment. In addition, we suggest performing an AS correction experiment on follow-up data to improve and facilitate the VDE calculation. Ideally this would be done with a multi-reader consensus. This may allow more accurate manual ground-truth establishment and improve comparability between methods. Nevertheless, whether one or multiple readers establish the manual segmentation, it remains an estimation of the actual tumor volume[27]. Finally, 3D FLAIR acquisitions with higher spatial resolution compared to their 2D counterparts are becoming more common in clinical practice also thanks to advanced acceleration techniques. Our training set included only a small percentage of 3D FLAIR images. In the clinical evaluation, 4 among 8 cases of 3D FLAIR exams showed the lowest DSC values suggesting that the algorithm should be trained with more 3D images. As 3D manual segmentation is more complicated and time consuming than 2D, the development of automated tools is even more crucial. An optimal automated 3D FLAIR segmentation should limit the presence of flow artifacts and volume variations resulting in better VDE estimations.

In summary, we validated the use of a nnU-Net trained to automatically segment and follow-up LGG. The network was able to provide a volumetric and VDE evaluation over time. It showed a very good performance when compared with the corrections of an experienced neuroradiologist. The data involved in this study reflected clinical reality, i.e., surgical interventions, chemotherapy treatments, long follow-ups, FLAIR hyper signal variations, including both 3D and 2D FLAIR images from different MR vendors and field strengths. Nevertheless, the 2D nnU-Net demonstrates promising performance for the evaluation of a LGG follow-up. Future investigations could evaluate the benefits of training with more 3D data and using a 2.5D or 3D architecture to develop a reliable support for radiologists in LGG growth assessment.

References:

1. Goodenberger, M. L. & Jenkins, R. B. Genetics of adult glioma. *Cancer Genet.* **205**, 613–621 (2012).

2. Figarella-Branger, D. *et al.* [The 2021 WHO classification of tumours of the central nervous system]. *Ann. Pathol.* S0242-6498(21)00257–1 (2021) doi:10.1016/j.annpat.2021.11.005.

3. Duffau, H. & Mandonnet, E. The 'onco-functional balance' in surgery for diffuse low-grade glioma: integrating the extent of resection with quality of life. *Acta Neurochir. (Wien)* **155**, 951–957 (2013).

4. Mariani, L. *et al.* The impact of tumour volume and surgery on the outcome of adults with supratentorial WHO grade II astrocytomas and oligoastrocytomas. *Acta Neurochir. (Wien)* **146**, 441–448 (2004).

5. Pallud, J. *et al.* Velocity of tumor spontaneous expansion predicts long-term outcomes for diffuse low-grade gliomas. *Neuro-Oncol.* **15**, 595–606 (2013).

6. Mandonnet, E. *et al.* Continuous growth of mean tumor diameter in a subset of grade II gliomas. *Ann. Neurol.* **53**, 524–528 (2003).

7. Badrigilan, S. *et al.* Deep learning approaches for automated classification and segmentation of head and neck cancers and brain tumors in magnetic resonance images: a meta-analysis study. *Int. J. Comput. Assist. Radiol. Surg.* **16**, 529–542 (2021).

8. Kamnitsas, K. *et al.* Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **36**, 61–78 (2017).

9. Long, J., Shelhamer, E. & Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *ArXiv14114038 Cs* (2015).

10. Hussain, S., Anwar, S. M. & Majid, M. Segmentation of glioma tumors in brain using deep convolutional neural network. *Neurocomputing* **282**, 248–261 (2018).

11. Cui, S., Mao, L., Jiang, J., Liu, C. & Xiong, S. Automatic Semantic Segmentation of Brain Gliomas from MRI Images Using a Deep Cascaded Neural Network. *J. Healthc. Eng.* **2018**, 4940593 (2018).

12. Li, Z. *et al.* Low-Grade Glioma Segmentation Based on CNN with Fully Connected CRF. *J. Healthc. Eng.* **2017**, 9283480 (2017).

13. Yang, Q., Zhang, H., Xia, J. & Zhang, X. Evaluation of magnetic resonance image segmentation in brain low-grade gliomas using support vector machine and convolutional neural network. *Quant. Imaging Med. Surg.* **11**, 300–316 (2021).

14. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *ArXiv150504597 Cs* (2015).

15. Yan, B., Cao, M., Gong, W. & Wei, B. Multi-scale brain tumor segmentation combined with deep supervision. *Int. J. Comput. Assist. Radiol. Surg.* (2021) doi:10.1007/s11548-021-02515-w.

16. McKinley, R., Meier, R. & Wiest, R. Ensembles of Densely-Connected CNNs with Label-Uncertainty for Brain Tumor Segmentation. in *BrainLes@MICCAI* (2018). doi:10.1007/978-3-030-11726-9_40.

17. Naser, M. A. & Deen, M. J. Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. *Comput. Biol. Med.* **121**, 103758 (2020).

18. Isensee, F., Jaeger, P. F., Full, P. M., Vollmuth, P. & Maier-Hein, K. H. nnU-Net for Brain Tumor Segmentation. *ArXiv201100848 Cs Eess* (2020).

19. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J. & Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).

20. Wang, P. & Chung, A. C. S. Relax and focus on brain tumor segmentation. *Med. Image Anal.* **75**, 102259 (2022).

21. Duffau, H. Diffuse low-grade glioma, oncological outcome and quality of life: a surgical perspective. *Curr. Opin. Oncol.* **30**, 383–389 (2018).

22. NITRC: MRIcron: Tool/Resource Info. https://www.nitrc.org/projects/mricron.

23. Brasil Caseiras, G. *et al.* Low-Grade Gliomas: Six-month Tumor Growth Predicts Patient Outcome Better than Admission Tumor Volume, Relative Cerebral Blood Volume, and Apparent Diffusion Coefficient. *Radiology* **253**, 505–512 (2009).

24. Lin, L. I.-K. A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics* **45**, 255–268 (1989).

25. Bland, J. M. & Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet Lond. Engl.* **1**, 307–310 (1986).

26. Bø, H. K. *et al.* Intra-rater variability in low-grade glioma segmentation. *J. Neurooncol.* **131**, 393–402 (2017).

27. Visser, M. *et al.* Inter-rater agreement in glioma segmentations on longitudinal MRI. *NeuroImage Clin.* **22**, 101727 (2019).

III. Caractérisation de la tumeur : analyse radiomique

Afin de prédire l'évolution de la tumeur, il est nécessaire de la comprendre et de la caractériser en extrayant un maximum de données caractéristiques de son évolution par imageries IRM.

La plupart des approches précédemment décrites prédisant la transformation anaplasique, se basent sur des variations morphologiques de la tumeur[103,104] ou des marqueurs simples et indépendants comme la perfusion calculée à partir du volume [8]. La prise de contraste renseignée en pratique courante sur l'imagerie morphologique 3D T1 suite à l'injection de produit de contraste, indique l'existence d'une rupture de la barrière hémato-encéphalique, mais n'est pas spécifique du moment de la transformation anaplasique. Ces métriques simples ne reflètent au final qu'une partie des éléments visibles, et l'approche de la transformation reste difficile à évaluer malgré ces paramètres.

Il a été observé que l'évolution des GBG se traduit en imagerie IRM par une majoration de l'hétérogénéité du signal sur les séquences T2 FLAIR au sein de la lésion tumorale. Le caractère infiltrant, des contours flous, les évolutions lentes et hétérogènes du signal T2 FLAIR et du volume sanguin au sein de la tumeur sont des éléments typiques des GBG observés en IRM en routine clinique. Ces caractéristiques ne sont pourtant actuellement pas prises en compte ni exploitées comme prédictives de la transformation anaplasique en pratique courante. Les hétérogénéités de signal des séquences T2 FLAIR sont des paramètres davantage étudiés pour évaluer la croissance des glioblastomes [7]. Nous savons également que les glioblastomes (GHG) présentent d'importantes hétérogénéités de signal en T2 FLAIR en IRM[12,105]. Ces hétérogénéités pourraient donc être en lien avec l'activité mitotique et traduire la transformation maligne du gliome. Néanmoins, ces études restent basées sur des observations, et depuis quelques années, avec l'évolution des algorithmes d'intelligence artificielle, des méthodes permettant de quantifier les variations de signal au sein de l'image se sont développées.

Appréhender les variations et caractéristiques du signal issues de l'imagerie multimodale au sein de la tumeur constitue la base de ce que l'on appelle la radiomique, que nous nous proposons d'utiliser afin de décrire finement les caractéristiques des GBG au cours de leur évolution, dans le cadre du suivi radiologique.

a. Définition

Le terme « radiomique » est apparu en 2012. La radiomique se base sur un ensemble de données quantitatives extraites à partir d'imagerie radiologique (CT, PET, IRM…) afin de caractériser des lésions tumorales solides[13].

Elle permet une conversion numérique de la richesse texturale de l'image et permet une analyse plus détaillée et quantitative que celle réalisée à l'œil nu. Le nombre de paramètres extraits est important et leur exploitation s'est faite grâce au développement des algorithmes d'intelligence artificielle. Ces algorithmes sont capables de réaliser des tests statistiques de grande ampleur afin de mettre en évidence la combinaison de paramètres permettant de répondre au mieux à la problématique étudiée. En premier lieu, toutes les données possibles sont extraites, puis l'algorithme essaye d'en tirer une redondance pouvant amener à une hypothèse. C'est une méthode « data driven », c'est-à-dire s'appuyant sur un recueil très important de paramètres. Cette désignation s'applique aussi à des disciplines caractérisées par la terminaison « -omiques ». Comme la génomique, par exemple, étudie les informations des gènes, la radiomique extrait des informations des microstructures de l'image, et toutes deux visent à décrire et/ou comprendre le phénotype des phénomènes ou anomalies biologiques étudiées[106].

L'un des avantages d'une approche radiomique, dans le cadre de l'oncologie et plus particulièrement du suivi des GBG, est le suivi radiologique fréquent (3-6mois). L'ensemble des imageries produites constitue autant de données potentiellement exploitables par une étude radiomique.

En utilisant l'évolution de centaines de caractéristiques dérivées des images, elle permet de parvenir à une identification plus sensible et plus fiable des types de tumeurs. La radiomique peut servir aussi à caractériser les différents types cellulaires[14,107,108], surveiller le développement et la progression de la maladie ou sa réponse au traitement[109–111].

De manière plus globale, la radiomique permettra de caractériser l'hétérogénéité tumorale, c'est-à-dire la grande variabilité des phénotypes des cancers, et elle s'insère bien dans une démarche de médecine personnalisée orientant vers une thérapie ciblée pour chaque patient·e·s. Elle pourra s'intégrer à des données cliniques, biologiques ou encore génétiques afin de mieux appréhender chaque tumeur.

b. Paramètres

Pour étudier les paramètres radiomiques d'une lésion tumorale, plusieurs métriques quantitatives peuvent être extraites de la région de l'image désignée comme représentant la zone lésionnelle (masque de segmentation). Ces paramètres quantitatifs sont de plusieurs types :

1. L'analyse de premier ordre

Ces informations directement liées aux intensités de l'image sont aussi appelées « analyses de premier ordre » et sont directement issues des histogrammes de niveaux de gris de l'image. On pourra alors extraire l'intensité maximale, minimale, moyenne, médiane… Ce sont aussi des informations liées à la distribution de ces intensités sur l'histogramme, décrite par les mesures de variance, de « skewness » analysant la symétrie de la distribution, ou encore de « kurtosis » décrivant une distribution plus ou moins « pointue » des intensités. L'intensité de l'image peut aussi être étudiée par l'énergie qui décrit l'ampleur des valeurs des voxels dans une image, ou encore par l'entropie qui est une mesure de l'aléatoire de la distribution des voxels dans l'image (***Erreur ! Source du renvoi introuvable.***.*A*).

2. Les paramètres de forme

Ces paramètres peuvent être calculés sur la base du volume ou sur une coupe de la lésion, et sont indépendants de la distribution de l'intensité du niveau de gris de l'image. Les caractéristiques sont dérivées de la forme approximative définie par un maillage en 3 ou 2 dimensions. Dans un espace en 2 dimensions on pourra extraire par exemple la surface, le périmètre… Dans un espace en 3 dimensions le volume, la sphéricité, la « compactness » qui mesure la compacité de la forme de la tumeur par rapport à une sphère, représentant le maximum de compacité, et un certain nombre d'autres paramètres. Que ce soit en 2D ou en 3D, on pourra calculer la sphéricité, le diamètre maximum et minimum du maillage ou encore l'élongation qui montre la relation entre les deux plus grands axes composant la forme du maillage (***Erreur ! Source du renvoi introuvable.***.*B*)

3. L'analyse de second ordre : les paramètres de texture

La mesure quantitative de paramètres de texture, extraits d'images radiologiques caractérisant l'hétérogénéité tumorale à un niveau non visible à l'œil nu et corrélés à la nature biologique des tissus, semble émerger comme un nouveau biomarqueur tumoral au même titre que l'hétérogénéité histologique ou moléculaire. Son utilisation pourrait permettre de prédire la nature biologique d'un tissu, la réponse thérapeutique ou le pronostic d'une lésion tumorale. L'analyse de la texture intervient souvent dans la classification des images, mais il n'existe pas de définition

universelle. Elle peut être considérée comme un motif global résultant de la répétition de sous-motifs locaux[112] ou comme une région où un ensemble de propriétés sont variables, constantes ou encore périodiques[113]. De nombreuses méthodes de gestion de la texture ont été développées, basées sur les différentes façons de caractériser la texture.

Les images peuvent être filtrées pour mettre en évidence les différents motifs qui la composent afin de caractériser sa macro-structure, c'est la **méthode structurelle**[114]. La représentation résultante est une synthèse de l'image.

Il existe des **méthodes basées sur des modèles mathématiques** complexes comme les modèles fractals[115], ou stochastiques[116], mais elles sont moins couramment utilisées compte tenu de la puissance de calcul nécessaire.

Les propriétés de texture peuvent aussi être analysées dans un espace différent de celui de l'image, comme à l'échelle des fréquences ou à l'échelle des dimensions. Parmi ces **méthodes de transformation**, on peut citer la transformée de Fourier[117], la transformée de Gabor[118] ou encore la transformée en ondelettes[119]. Ces 3 méthodes peuvent aussi être étudiées de manière concomitante[120].

La troisième méthode, la plus couramment utilisée, est **la méthode statistique**. Elle met en évidence les propriétés régissant la distribution des pixels et les relations des valeurs de niveau de gris dans l'image (***Erreur ! Source du renvoi introuvable.***.*C*). Cette analyse des niveaux de gris de l'image se fait avec des matrices visant à mettre en évidence diverses relations entre les valeurs de niveaux de gris.

La « gray-level cooccurrence matrix », la matrice de dépendance spatiale des niveaux de gris (GLCM)[121], calcule la fréquence à laquelle des paires de pixels ayant des valeurs et une relation spatiale spécifique apparaissent. La relation spatiale est spécifiée par une dimension et une direction, en général 4 directions (0°, 45°, 90° et 135°) pour une analyse 2D, et 13 directions en 3D. A partir de la GLCM, plusieurs paramètres sont calculés. Les plus courants sont le contraste qui représente la variation locale d'intensités, la corrélation qui est la probabilité d'apparition jointe de paires de pixels spécifiques, l'homogénéité qui mesure la proximité de la distribution des éléments de la GLCM par rapport à celle de sa diagonale, et enfin l'énergie qu'on appelle aussi uniformité, calculée par la somme des éléments de la GLCM au carré.

Il existe plusieurs autres méthodes basées sur l'analyse des niveaux de gris de l'image par le biais d'une matrice, notamment la « run length matrix » (GLRLM)[122] qui consiste à parcourir

l'image, toujours dans une direction donnée, à la recherche de séries de pixels ayant la même valeur de niveau de gris. Ou encore la « size zone matrix » (GLSZM) qui quantifie les zones de niveau de gris dans une image. Une zone de niveau de gris est définie comme le nombre de voxels connectés qui partagent la même intensité de niveau de gris (***Erreur ! Source du renvoi introuvable.***.*C*).

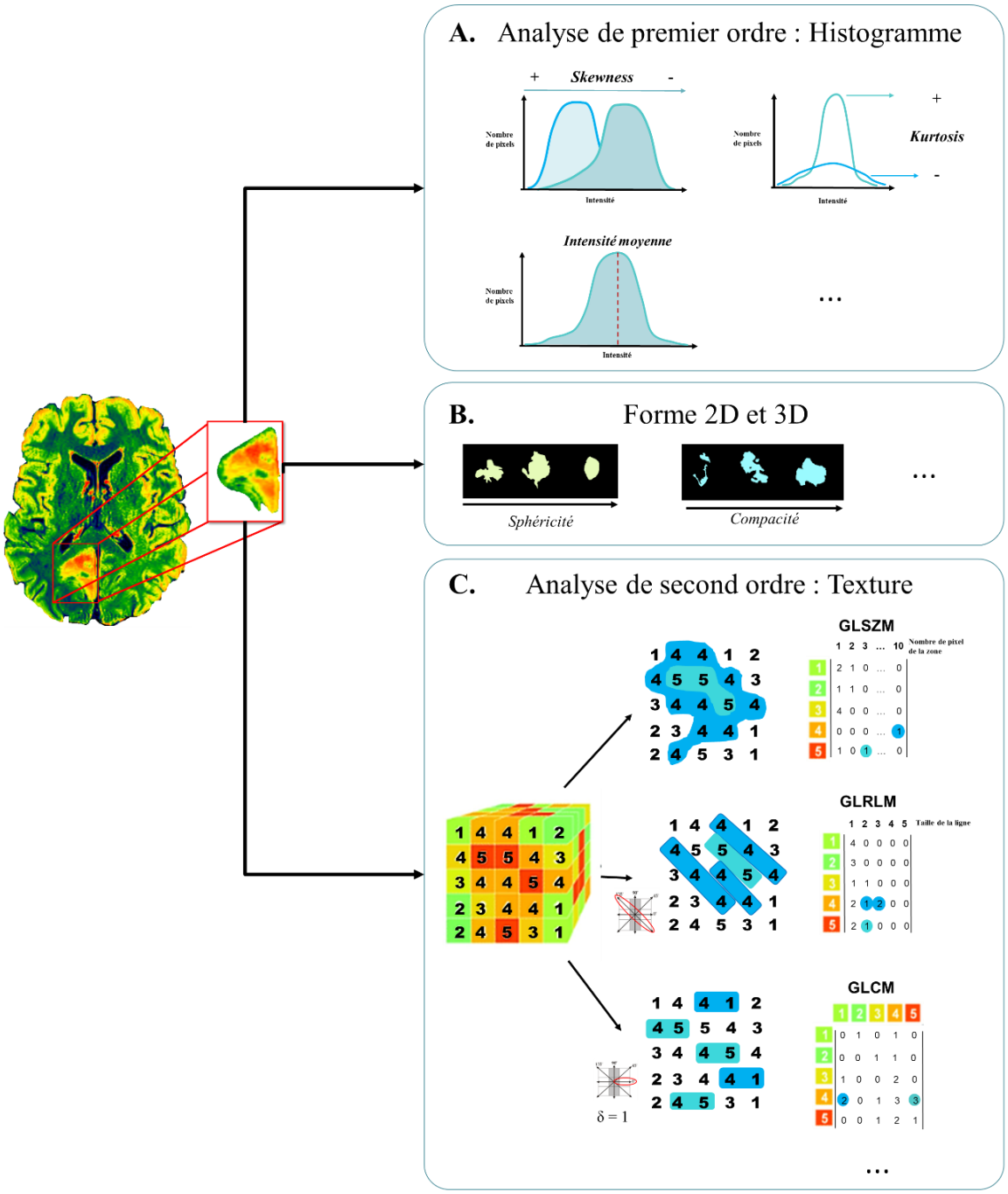


*Figure 15: Etude radiomique d'une lésion tumorale. **A.** Analyse de premier ordre se basant sur l'histogramme du nombre de pixels en fonction des intensités de l'image. **B.** Analyse de la forme du masque lésionnel (2D et 3D). **C.** Analyse de second ordre mettant en évidence la texture de l'image. Cette analyse se base sur l'étude des niveaux de gris de l'image en réalisant des matrices dépendantes d'une certaine organisation de ceux-ci. Par exemple, la matrice de co-occurrence des niveaux de gris, GLCM, va parcourir l'image pixel par pixel, et compter les répétitions (co-occurrences) des pixels entre eux à une distance de 1 (δ=1) et dans une orientation de 0°. On obtient donc, avec ces paramètres, 2 fois le motif «4 1» et 3 fois le motif «4 5».*

Avec le temps, de plus en plus de paramètres ont pu être calculés à partir des différentes méthodes d'analyse radiomique. C'est pourquoi, avec la multiplication des études radiomiques, il a fallu mettre en place des directives pour l'utilisation de ces données. L'initiative de standardisation des biomarqueurs d'image, l'IBSI (« image biomarker standardisation initiative »), est une collaboration internationale indépendante qui vise à standardiser l'extraction de biomarqueurs d'image à haut débit, à partir d'imagerie radiologique, afin de réaliser une analyse quantitative (radiomique). Le manque de reproductibilité et de validation des études d'analyse quantitative d'images est un défi majeur, car il existe un manque de directives et de définitions consensuelles pour le processus de traduction de l'imagerie acquise en biomarqueurs d'image. L'IBSI vise à fournir une nomenclature, des définitions des biomarqueurs et des valeurs de référence pour vérifier le traitement des images et les calculs des biomarqueurs ainsi que des directives pour la mise en forme des résultats d'études radiomiques[123].

c. Radiomique dans les gliomes diffus de bas-grade

La radiomique basée sur l'IRM est une modalité non invasive qui est devenue un domaine d'intérêt particulier pour le classement des tumeurs, mais n'a pour l'instant que peu d'application en pratique clinique. C'est néanmoins une technique prometteuse qui peut répondre à certains enjeux en oncologie[15].

La radiomique, utilisée conjointement avec une analyse histopathologique des gliomes, permet de mettre en évidence les différents types cellulaires qui les caractérisent. Cela peut être utilisé pour distinguer deux groupes avec une caractéristique précise comme la présence ou non de la co-délétion 1p/19q qui est facteur de bon pronostic[124], ou bien pour discriminer plus largement les différents sous- types cellulaires[125,126]

La méthode radiomique associée à l'intelligence artificielle s'est déjà montrée efficace pour discriminer GBG et GHG : que ce soit en étudiant seulement les métriques d'une seule matrice de niveaux de gris (GLSZM) et dans 2 modalités d'imageries anatomiques[16], ou bien en extrayant un très grand nombre de variables à partir de différents types d'imagerie[127]. Ces méthodes radiomiques utilisant un très grand nombre de paramètres, ont été appliquées à la discrimination des différents types de gliomes[128], ou encore entre gliomes et autres types de lésion cérébrale[129].

Néanmoins, les études permettant de distinguer GBG de GHG n'identifient pas de métriques spécifiques à une transformation anaplasique car elles prennent en compte des groupes hétérogènes en termes d'évolution. Les groupes identifiés comme GHG comportent des grades 3 et 4 sans

prendre en compte l'aspect évolutif des gliomes, et ne permettent pas de savoir si ce sont des GHG secondaires ou primaires (issus ou non de la transformation anaplasique d'un GBG). De plus, ces études ne permettent pas d'identifier des biomarqueurs radiomiques pouvant être utilisés comme biomarqueurs de la transformation anaplasique en routine clinique, car les grades inclus dans ces études sont parfois des populations hétérogènes de grade 1 et 2 pour les GBG et 3 et 4 pour les GHG[16,127,130].

d. Des images aux paramètres radiomiques

1. Population de l'étude radiomique

L'extraction des paramètres radiomiques ainsi que l'optimisation de la chaîne de traitement associée a été réalisée sur la base d'une cohorte incluant à la fois des données de bas grade, GBG non transformés, ainsi que des gliomes de haut grade (GHG), qui représente la finalité de l'état d'évolution du bas grade (*Figure 16*). Précisons que cette cohorte sera utilisée pour l'intégralité des analyses de caractérisation mais également de prédiction qui suivra. Le choix d'inclure des GHG a ainsi été réalisé en prévision de la partie suivante (*IV - Prédiction de l'évolution du gliome*).

Les patient·e·s GBG sont issu.e.s de la base de données SPECIFY correspondant aux suivis utilisés dans l'entraînement et le test de l'algorithme de segmentation automatique. Ce sont des suivis longs au cours desquels plusieurs interventions chirurgicales sont parfois effectuées. Les imageries IRM réalisées directement après les actes chirurgicaux ont été éliminées de la cohorte, ce sont en effet des imageries de contrôle difficilement lisibles en raison des stigmates de la chirurgie. Au cours de certains suivis, une transformation anaplasique pouvant avoir lieu (*voir I.4 - Histoire naturelle*), le GBG est donc considéré comme un GHG à partir de ce point et pour les examens de suivi postérieurs. Enfin, un suivi ayant présenté des artefacts pour plusieurs acquisitions, il a été décidé de le retirer de la cohorte. La cohorte de GBG étudiée est donc composée du suivi longitudinal de 35 sujets GBG, ce qui représente 282 examens IRM. Au sein de cette cohorte, on distinguera 2 groupes. Le premier est composé des examens GBG du début du suivi. Ce sont des examens réalisés en pré ou post-chirurgie, mais hors de tout traitement par chimiothérapie, et ils sont considérés comme les plus éloignés de la transformation anaplasique. Cela représente 70 examens IRM qui seront labélisés « GBG ». Le deuxième groupe est composé des 212 examens IRM composant le suivi des GBG, ces examens seront donc labellisés « suivi GBG ». C'est au long de ces examens que le GBG va évoluer, et que la pratique clinique va tenter de freiner son évolution par des traitements de chimiothérapie ou par chirurgie (*Figure 16*).

Une cohorte de GHG issus de GBG sera indispensable pour identifier par la suite de potentiels biomarqueurs de la transformation anaplasique. En effet, dans notre suivi des GBG, le nombre de transformations anaplasiques identifiées par prise de contraste est trop faible pour permettre une analyse statistique comparative fiable avec les GBG non transformés. Il a donc été nécessaire de compléter avec des patient·e·s externes à nos données de suivi. Ces GHG sélectionnés sont des tumeurs ayant subi une transformation anaplasique. Ils sont dits GHG secondaires et ont été identifiés dans le cadre de l'activité clinique du service de neurochirurgie du CHU de Montpellier. L'examen retenu est au choix celui réalisé à la date de prise en charge (pour les GHG secondaires à l'admission, validés par une analyse anatomopathologique) ou celui réalisé à la date d'identification de la transformation par prise de contraste à l'imagerie (pour les GHG secondaires transformés lors d'un suivi de GBG). La base de données est composée de 86 examens IRM (*Figure 16*).

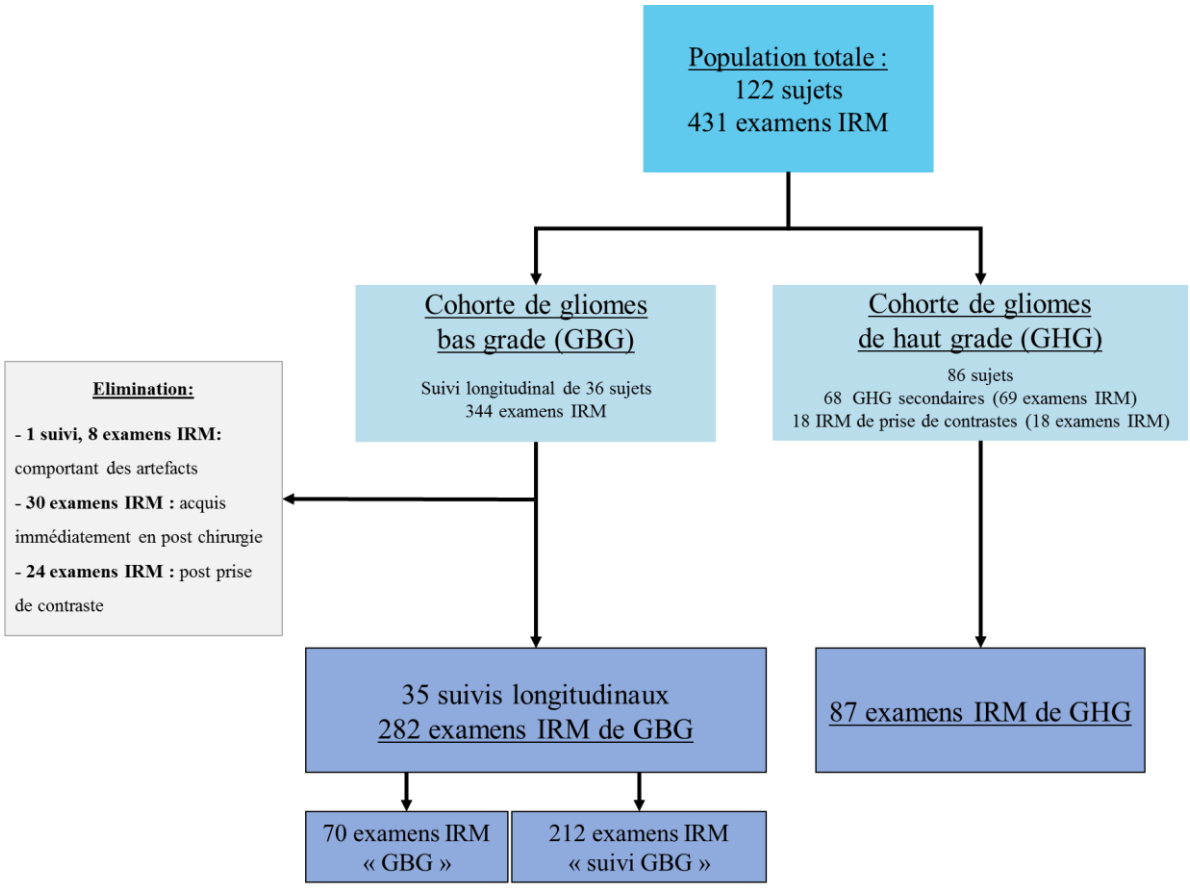


Figure 16: Organigramme de la composition des deux cohortes. La cohorte gliomes de bas grades (GBG) est composée de 282 examens IRM qui représentent 35 suivis longitudinaux. Ces 282 examens composent 2 groupes, 70 examens labélisés « GBG » correspondent aux imageries du début du suivi où on considère que le gliome n'a pas encore beaucoup évolué, les 212 restant sont des examens de « suivi GBG ».La cohorte de gliomes de haut grade (GHG) est composée de 68 GHG secondaires et de 18 examens IRM où apparait une prise de contraste qui témoigne de la transformation anaplasique du gliome. Cette cohorte contient donc 87 examens IRM de GHG.

Tous ces examens sont issus de la routine clinique et ont été colligés rétrospectivement, ils sont donc très hétérogènes (*Tableau 1*). Tout d'abord en matière de machine d'acquisition IRM, il y a eu plusieurs modèles de constructeurs différents au fil du temps (Siemens, Général Electrics, Philips) et aussi différents champs magnétique de 3 Tesla ou 1,5 Tesla. On observe aussi une variation de résolution au niveau des différentes acquisitions (T2 FLAIR, 3D T1 GADO et perfusion). Tous ces aspects sont autant de facteurs pouvant influer sur l'analyse de l'image. Pour réaliser une étude radiomique, il est primordial que les images soient comparables.

C'est pourquoi un prétraitement commun à toutes les imageries de la base de données a été mis en place afin d'homogénéiser cette base avant l'analyse radiomique.

	Cohorte de bas grade **Grade 2**	**Cohorte de haut grade** **Grade 3 secondaire**
Examens IRM (n)	282	87
Examens IRM par sujet (moyenne (± sd))	8 (± 4)	1
Nombre de sujets (n)	35	86
Sexe (% (n))		
Homme	54% (19)	57% (49)
Femme	46% (16)	43% (37)
Age moyen (± sd)	35 ± 8	43 ± 13
IRM à champs 3T (% (n))	52% (147)	59% (52)
Constructeur (% (n))		
Siemens	100% (147)	98% (51)
Philips		2% (1)
IRM à champs 1.5T (% (n))	48% (135)	41% (35)
Constructeur (% (n))		
Siemens	95% (128)	97% (34)
Philips	4% (5)	
GE	1% (2)	3% (1)
Séquence IRM		
T2 FLAIR (% (n))		
2D (résolution (mm^3))		
0.5×0.5×5	0,5% (1)	
0.9×0.9×6	32% (89)	13% (11)
0.9×0.9×5	3% (8)	9% (8)
0.9×0.9×4	14% (41)	25% (22)
0.75×0.75×3	46% (132)	47% (41)
3D (résolution (mm^3))		
1×1×1	4% (10)	5% (4)
1.5×1×1	0,5% (1)	1% (1)
T1 avec injection de Gadolinium (% (n))		
3D (1×1×1 mm^3)		
Turbo Gradient-Echo	88,3% (249)	83% (71)
Gradient-Echo	11% (31)	17% (15)
Turbo Spin-Echo	0,7% (2)	
Perfusion		
T2* 2D (résolution (mm^3))		
1.8×1.8×4	53% (150)	61% (53)
1.8×1.8×5	47% (132)	39% (34)

Tableau 1 : *Informations démographiques sur la population utilisée dans l'étude radiomique.*

2. Chaîne de traitement

Les acquisitions réalisées au sein de notre protocole tel que défini précédemment aboutissent à des images de résolutions et parfois d'orientations différentes. Il a donc été indispensable de trouver les transformations permettant de superposer une image à une autre, c'est-à-dire de les ramener dans un espace commun. Les procédures en traitement d'image IRM qui permettent ces transformations sont la réorientation, la corégistration et la normalisation. Dans un premier temps, ces outils vont être décrits, puis leur application à la base de données sera expliquée (*Figure 17*).

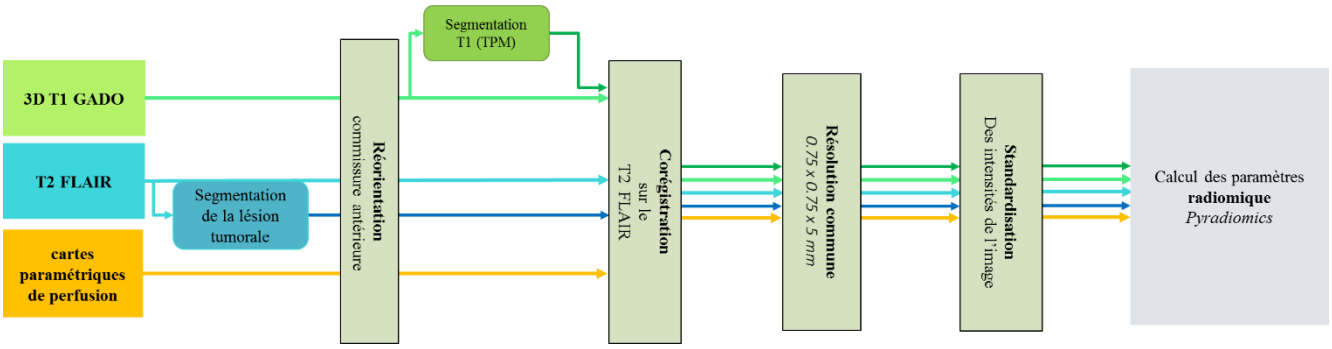


Figure 17 : Les étapes de prétraitement des données avant l'extraction des paramètres radiomiques : réorientation, corégistration, report à une résolution commune et normalisation des intensités.

Notons que dans le cadre de cette thèse, l'ensemble des traitements d'image ont été réalisés en utilisant MATLAB[131] et la toolbox SPM12[132]. SPM12 constitue la bibliothèque de programmation sous MATLAB qui fait référence en la matière et propose un grand nombre de fonctions assurant le traitement des données IRM.

Toutes les images ont étés converties depuis le format DICOM au format nifti. Pour les imageries 3D T1 GADO et T2 FLAIR, la conversion est faite par le logiciel XellSior développé au sein de la plateforme I2FH. Les cartographies de perfusion ont été extraites par Oléa[133] et converties en nifti par Matlab[131] avec la toolbox SPM12[132].

a) Réorientation

Afin de faciliter les étapes ultérieures de prétraitement, toutes les images doivent être, dans un premier temps, réorientées. Il s'agit d'orienter les images en fonction d'un point commun qui deviendra l'origine de chacune d'entre elles.

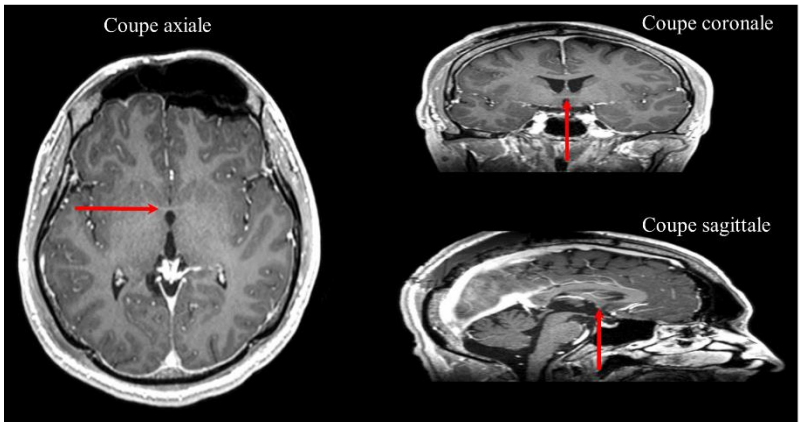


Figure 18: Localisation de la commissure antérieure sur l'imagerie 3D T1 GADO en coupe axiale, coronale et sagittale.

Pour chaque acquisition, à chaque modalités d'imagerie (T2 FLAIR, 3D T1 GADO et aux 8 cartes paramétriques de perfusion) une structure anatomique a été choisie comme origine.

La commissure antérieure est une structure de fibres nerveuses de substance blanche reliant les deux hémisphères cérébraux. Cette structure est facilement identifiable sur l'imagerie 3D T1 GADO dans les plans axial, coronal et sagittal (*Figure 18*).

Les coordonnées de ce point ont été colligées sur l'imagerie 3D T1 GADO à chaque date d'examen. Pour chaque examen, les différentes imageries sont réorientées afin que les coordonnées de la commissure soient l'origine de l'espace d'acquisition, c'est-à-dire égale à 0 dans les 3 directions.

b) Segmentation tissulaire

Au niveau cérébral, on distingue principalement trois tissus : la matière grise, la matière blanche et le liquide cérébro-spinal. La segmentation vise le plus souvent à identifier ces tissus dans l'image IRM. En effet, en pondération T1, ces tissus disposent d'un contraste différent.

L'identification de ces tissus sur les imageries 3D T1 GADO est indispensable afin de définir un masque « parenchyme » incluant la matière grise et la matière blanche nécessaire à la réalisation de la standardisation des intensités de la base de données (*e) - Standardisation des intensités*).

SPM utilise une approche de segmentation unifiée probabiliste[83] basée sur l'intensité des voxels et l'utilisation de cartes de probabilités des tissus (TPM) obtenues à partir de nombreux sujets sains. Cette méthode permet entre autres la création d'un « template » (modèle) adapté à une cohorte donnée, et la normalisation des images sur ce modèle. Les cartes de TPM sont déformées de manière non rigide afin de correspondre au mieux aux données du sujet : elles produisent ainsi une

segmentation spécifique au sujet. Les informations de déformations sont stockées dans un seul fichier « flow field » (champ de vecteurs) considéré constant dans le temps[134]. Cela permet une normalisation plus fine ; cependant les données se retrouvent dans un espace qui est propre à leur cohorte (du fait de la création du « template »).

Dans une première étape, on cherche à décrire l'histogramme des voxels comme une superposition de distributions gaussiennes. En effet si l'on regarde l'histogramme d'une image pondérée en T1, on observe la présence de plusieurs pics correspondant aux tissus de contraste différents. Une fois les gaussiennes obtenues, elles sont normalisées par l'histogramme total afin d'obtenir les probabilités d'appartenance. On obtient ainsi une segmentation « brute ». Cette étape est cependant très sensible aux inhomogénéités de signal dans l'image, qui résultent de la non-uniformité de B0, de la bobine RF, d'artéfacts, etc.[135]. SPM intègre donc une étape préalable d'estimation et de correction de ces biais permettant de limiter leur impact. Une fois cette première segmentation obtenue, elle est comparée aux cartes de probabilités disponibles (TPM), en utilisant une approche bayésienne. L'image en niveau de gris et la segmentation seront considérées comme les réalisations de deux champs aléatoires. Les TPM étant des cartes de probabilités d'appartenance de chaque voxel à un tissu, elles doivent être adaptées aux populations de sujets. En pédiatrie par exemple, la répartition des différents tissus est très différente de celle de l'adulte. Des TPM spécifiques créées sur la base de sujets jeunes existent pour réaliser la segmentation. SPM intègre les TPM pour 6 tissus : la matière grise, la matière blanche, le liquide cérébro-spinal et même l'os, la graisse, et l'air. Cette deuxième étape requiert une normalisation. Les cartographies sont situées dans l'espace MNI. Une normalisation « Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra » (DARTEL) est réalisée pour réaligner les TPM et les segmentations brutes issues de la première étape (*Figure 19*). En résumé, le biais de l'image est corrigé, une première segmentation brute est réalisée, puis les TPM et les segmentations sont recalées. Le processus est réalisé plusieurs fois jusqu'à obtenir une segmentation correcte.

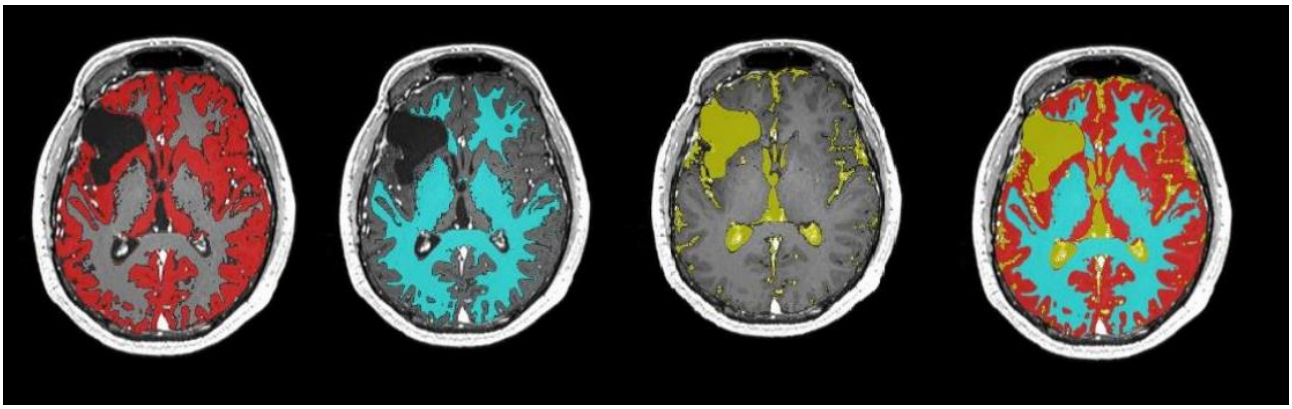


Figure 19: Segmentation des tissus sur le 3D T1 GADO. De gauche à droite : Matière grise, Matière blanche, LCR et les trois segmentations ensemble

c) La corégistration

Lors d'acquisitions comme la nôtre, contenant de multiples modalités, il est important de les positionner dans le même espace de façon à être en mesure de naviguer facilement d'une image à l'autre, et de les superposer, par exemple pour pouvoir superposer un même masque sur toutes les imageries. On commence ainsi par sélectionner une image dite de référence correspondant à un volume représentatif de l'acquisition. Tous les autres volumes seront alors transformés pour correspondre à celui-ci.

Cependant, du fait des différences d'intensités pour une même région, la minimisation des scores de différences entre les images, telle qu'elle est produite par un simple réalignement, ne permet pas d'aboutir à un résultat satisfaisant. Il faut donc trouver une transformation de corps rigide (6 paramètres) qui relie les deux images sans passer par la minimisation simple de la différence des intensités.

Le processus de co-registration est ramené au problème de l'optimisation d'une fonction de coût. Pour calculer cette fonction de coût, une première approche consiste à extraire indépendamment des deux images une même information générique, par exemple la segmentation des différents tissus. On minimise ensuite la moyenne des différences au carré de ces informations génériques[83]. Une autre approche consiste à calculer l'information mutuelle. Il s'agit d'un concept issu de la théorie de l'information. Pour deux images données, l'information mutuelle permet de quantifier leur ressemblance. Son calcul est basé sur la densité conjointe de probabilité des niveaux de gris des deux images. Pour cela, on calcule un histogramme conjoint des niveaux de gris. Les axes de l'histogramme représentent l'intensité des voxels sur chaque image, chaque point (a,b) du plan correspond à une intensité a sur la première image et à une intensité b sur la deuxième. On attribue à

chaque point un niveau de gris, en fonction du nombre total de voxels ayant une intensité a sur la première image et une intensité b sur la deuxième (*Figure 20,A.*). Lorsque les images sont alignées, certains points de l'histogramme ont un niveau de gris élevé, et les autres un niveau proche de 0 puisque tous les voxels coïncident (*Figure 20,B.*). En revanche, lorsque les images ne sont pas bien alignées, le nombre de coïncidences diminue et le niveau de gris se disperse dans l'histogramme. L'information mutuelle permet de quantifier cette information[136].

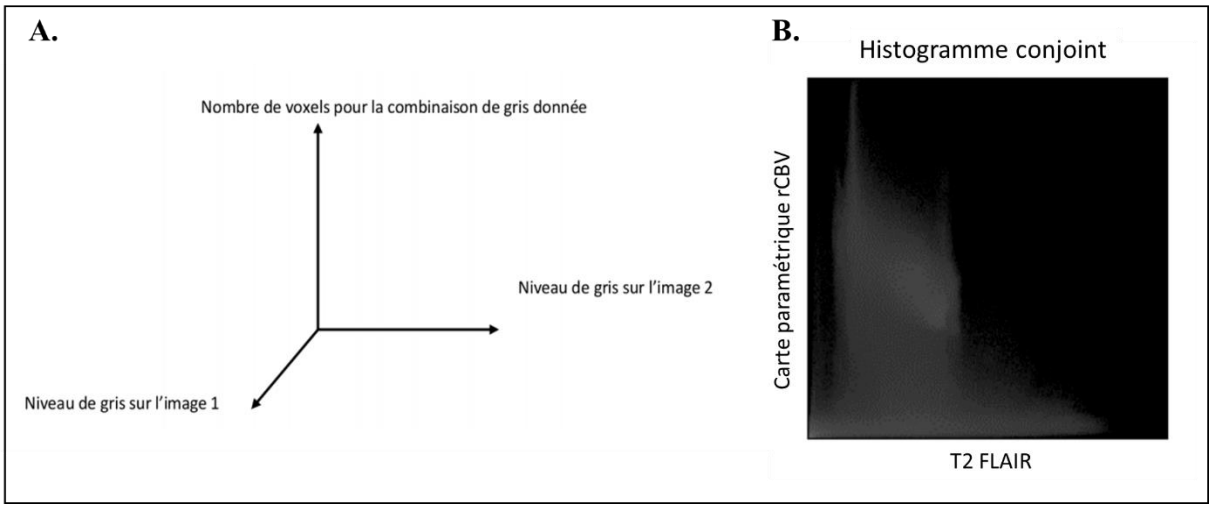


Figure 20: La corégistration. A. : Axes d'un histogramme conjoint. Le nombre de voxels pour une combinaison de niveau de gris donnée peut être représenté par une couleur. B. : Exemple d'histogramme conjoint obtenu par SPM après corégistration d'une carte de rCBV

Le calcul des paramètres radiomiques a été réalisé au sein du masque de segmentation tumoral. Ce masque a été tracé manuellement sur l'imagerie T2 FLAIR. C'est donc cette imagerie qui a été utilisée comme référence pour la corégistration de la base de données. Pour chaque examen IRM, les imageries 3D T1 GADO, les masques de segmentations des différents tissus ainsi que les 8 cartes paramétriques de perfusion ont été corégistrées dans l'espace de leurs T2 FLAIR (et donc masque de segmentation tumorale) respectifs.

d) Interpolation à une résolution commune

La base de données utilisée est une base de données de pratique clinique. C'est-à-dire que les examens sont très hétérogènes, parfois réalisés dans différents centres avec des pratiques différentes. Certains examens datent de 2010 et vont jusqu'à 2020. Les machines sont différentes, possèdent des constructeurs différents et des champs d'acquisitions qui peuvent être 3T ou 1,5T.

Les acquisitions T2 FLAIR sur lesquelles le masque manuel de la lésion tumorale a été réalisé ont des résolutions très différentes. Afin de pouvoir comparer les différents volumes

tumoraux, il faut ramener les images étudiées à une résolution commune. Ces différentes résolutions ont aussi un impact sur les paramètres radiomiques, car un grand nombre d'entre eux sont calculés à partir des informations au sein des voxels, informations qui ne seront comparables que si les résolutions des images sont les mêmes. Comparer des données avec des résolutions différentes peut induire des biais avec des sur ou sous-estimations de volume à certaines dates où les résolutions fluctuent.

Nous nous proposons ainsi d'identifier, dans notre base de données, la résolution commune à laquelle rapporter l'intégralité des données afin de minimiser les erreurs inhérentes à l'interpolation notamment vers des résolutions plus élevées.

Pour mettre en évidence cette résolution commune, il faut évaluer l'impact de l'interpolation sur les images de la base de données, particulièrement les imageries T2 FLAIR sur lesquelles sont tracés les volumes car ce sont les moins bien résolues. Il y a une plus grande perte d'information quand une image très peu résolue est reportée à une haute résolution : il y a interpolation donc création d'information dans l'image. Il est important que les masques de segmentation tracés restent dans les tissus de la lésion tumorale et n'intègrent pas de tissu sain.

La résolution commune choisie devra générer le moins d'erreur possible dans les segmentations manuelles et les paramètres radiomiques calculés au sein du masque.

(1) Protocole

A partir des 315 imageries T2 FLAIR, 29 résolutions différentes, caractéristiques de la base de données, ont été identifiées. Une image T2 FLAIR de résolution 1x1x1mm^3 a été corégistrée pour constituer un groupe d'images T2 FLAIR et leur masque de segmentation manuelle respectif dans les 29 résolutions identifiées. Afin de simuler l'effet de la transformation de la base de données, les possibles interpolations et extrapolations induites par un changement de résolution, ces 29 images ont été corégistrées aux résolutions d'une large gamme de résolutions. Cette gamme comprend 42 résolutions allant de résolutions très fines (0.5x0.5x1) à de très basses résolutions (3x3x4). Les 29 T2 FLAIR caractéristiques sont chacune reportée dans les 42 résolutions de la gamme et aboutissent à un groupe de 1218 T2 FLAIR accompagnées de leur masque respectif constituant le groupe de simulation de la base de données.

Une image T2 FLAIR de résolution 1x1x1mm^3 a été corégistrée dans les 42 résolutions de la gamme afin de constituer le groupe contrôle. Les 1218 T2 FLAIR représentatifs de la base de données seront comparés aux 42 T2 FLAIR du groupe contrôle (*Figure 21*).

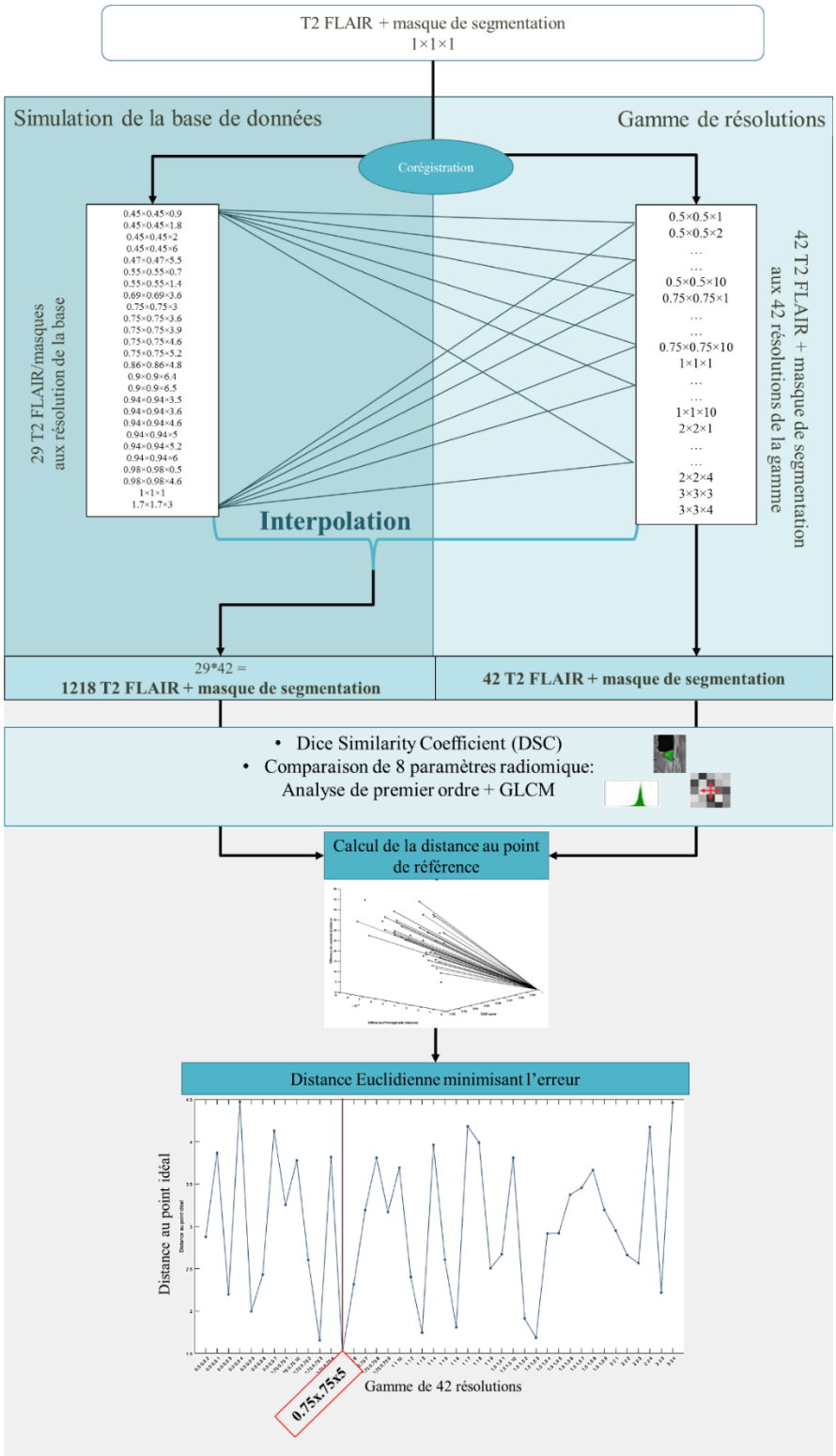


Figure 21 : Méthode pour la mise en évidence de la résolution commune : A partir d'une image T1 FLAIR de référence, deux cohortes sont réalisées. La base de données est simulée par la corégistration de l'image référence aux 29 dimensions retrouvées dans la base de données. La deuxième cohorte est réalisée par la corégistration de l'image de référence dans une gamme de 42 résolutions. La comparaison entre données interpolées ou non interpolées se fera par : le calcul du coefficient de Dice (DSC) et l'extraction de 8 paramètres radiomiques au sein de la tumeur. Afin d'identifier la résolution à laquelle l'interpolation génère le moins d'erreur, la distance euclidienne au point idéal est calculée. Le point idéal aura un DSC=1 (masques de segmentation identiques) et aucune différence entre les paramètres radiomiques (=0).

(2) Métriques évaluées

Les paramètres qui doivent varier le moins possible après report de la base de données à une résolution commune, sont le masque de segmentation de la tumeur et les paramètres radiomiques.
Les masques tumoraux des deux groupes ont été comparés par le coefficient de Dice (Dice similarity coefficient, DSC) qui évalue la proportion de recouvrement commun entre deux masques volumiques de segmentation.

Pour évaluer la texture, une matrice de cooccurrence de type GLCM a été calculée pour chaque masque tumoral. A partir de cette matrice ont été extraites les métriques de contraste, homogénéité, énergie et corrélation. Une analyse statistique des niveaux de gris de l'image a aussi été réalisée afin de calculer les paramètres de premier ordre de kurtosis, entropie, variance et intensité médiane.
Ces paramètres ont été choisis car ils apparaissent significatifs dans les études radiomiques cherchant à classer les gliomes en fonction de leur état de transformation[137–139].

(3) Sélection de la résolution commune

Chacune des images à une résolution donnée, parmi les 42 de la cohorte « gamme de résolutions », seront comparées aux 29 interpolations des images de la cohorte « simulation de la base de données » dans cette résolution.
Le point parfait possède, pour ces 29 comparaisons, un DSC médian égal à 1 traduisant deux masques recouvrant la lésion de manière identique, et la médiane des différences égale à zéro pour chacun des 8 paramètres de texture.

Parmi les résolutions du groupe contrôle, la résolution, se rapprochant le plus de ce point parfait a été identifiée par une comparaison à facteurs multiples. La distance euclidienne à ce point parfait est calculée pour chaque comparaison.

Équation 6

$$D = \sqrt{(1 - DSCmédian)^2 + (0 - \Delta_{contraste})^2 + (0 - \Delta_{homogénéité})^2 + \cdots + (0 - \Delta_{intensité_médiane})^2}$$

La résolution pour laquelle la médiane des DSC est la plus proche de 1 et la médiane des différences (Δ) de paramètres radiomiques la plus proche de zéro, est 0,75 x 0,75 x 5mm (*Figure 21*).

Les imageries T2 FLAIR, 3D T1 GADO et les cartes paramétriques de perfusion ont été rapportées à cette résolution commune par interpolation (reslicing) avec SPM12.

e) Standardisation des intensités

L'homogénéisation de la base de données passe aussi par la standardisation des intensités de l'image. Parmi chaque modalité d'imagerie, il y a des paramètres d'acquisition différents qui donnent des images avec une gamme de niveaux de gris qui varient. Les différentes modalités d'imagerie, T2 FLAIR, 3DT1 GADO et chacune des 8 cartes paramétriques de perfusion sont normalisées indépendamment.

L'approche de normalisation décrite par Nyùl et al. a été utilisée[140].

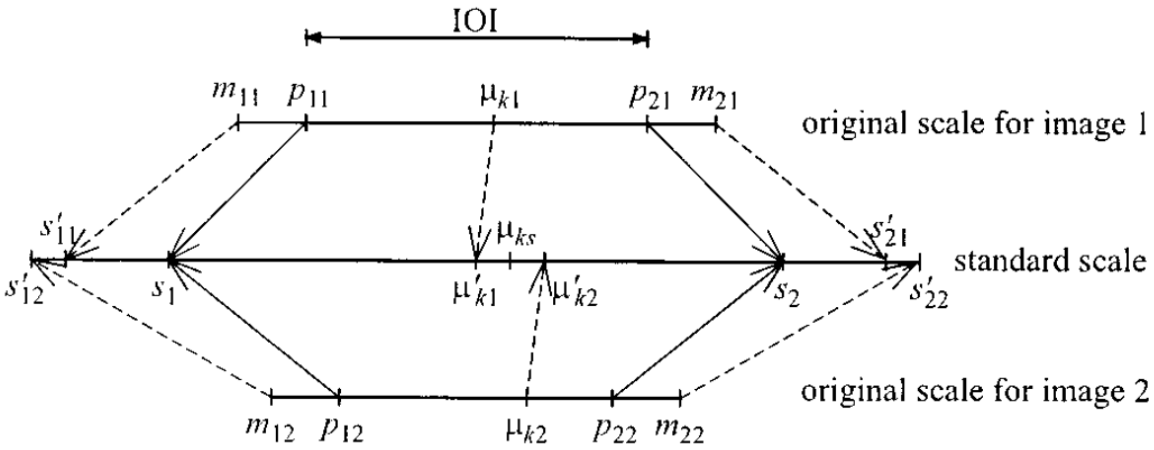


Figure 22 : *Méthode d'identification des paramètres de l'histogramme standardisé. Pour illustration, seulement deux images sont comparées ici. Pour le nombre d'images* $j = 1, 2$; m_{1j} *et* m_{2j} *sont le minimum et le maximum des intensités dans l'image ;* p_{1j} *et* p_{2j} *sont les percentiles minimum et maximum ;* μ_{kj} *est l'un des marqueurs de l'histogramme et* μ'_{kj} *est la valeur projetée de* μ_{kj} *sur l'échelle standardisée ; Les moyennes des valeurs projetées* μ'_{kj} *sont notées* μ_{ks}, *avec* s *le paramètre recherché dans l'histogramme à l'échelle standardisée.*

Par modalité d'imagerie, un histogramme standardisé des valeurs de l'imagerie est construit à partir de chaque IRM, quel que soit le sujet (*Figure 22*). Les paramètres utilisés sont similaires à ceux décrits par Bedekar et al.[141] car ils ont été testés et validés dans le cadre de l'analyse tumorale en incluant des paramètres de perfusion (rCBV). C'est-à-dire 11 points de repères, des marqueurs, liés à différents percentiles (10, 20, 30, 40, 50, 60, 70, 80, 90, 99.8) avec une plage d'intensité normalisée en fonction de la modalité étudiée. L'histogramme de chaque image est finalement projeté sur l'histogramme normalisé des valeurs obtenues pour la modalité d'imagerie étudiée (*Figure 23*).

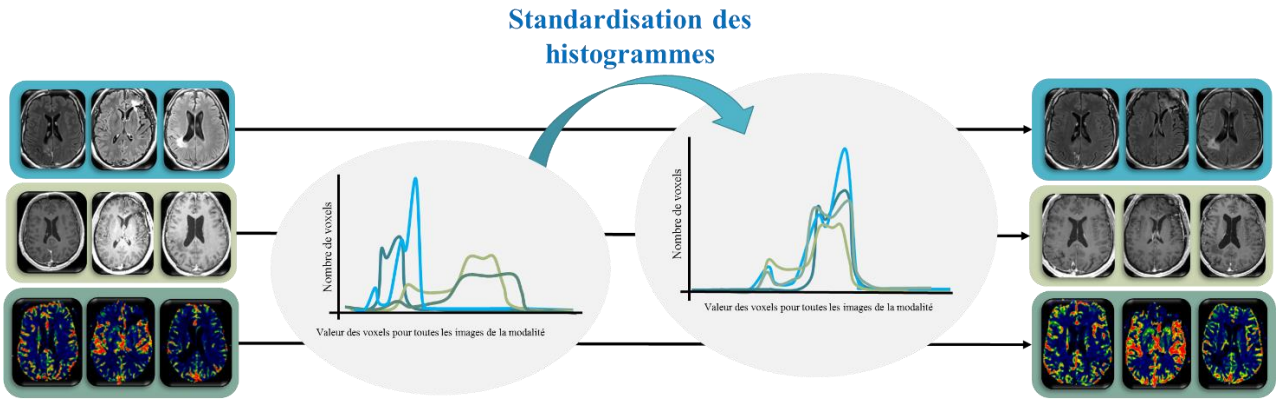


Figure 23 : Standardisation. Impact de l'alignement des histogrammes sur les intensités des images

f) Extraction des paramètres radiomique

Il existe plusieurs logiciels proposés pour effectuer le calcul des paramètres radiomiques dans une image. Tous ont leurs caractéristiques propres et sont plus ou moins accessibles. Parmi ceux mentionnés dans la littérature, certains proposent leur code source en accès libre (« open-source »[142–144], d'autres sont sous forme de logiciels gratuits[145], ou encore sont des produits commercialisés[146]. Dans certains de ces cas, le code source des fonctions de calcul n'est pas toujours accessible au public, ce qui empêche la comparaison des résultats entre les études de la littérature. En effet, sélectionner la « meilleure » méthode d'extraction des résultats nécessite de pouvoir les comparer dans les mêmes conditions. Même parmi les solutions « open-source », le nombre de paramètres est différent et les manières de calculer peuvent varier. Parmi les méthodes open-sources, la méthode Pyradiomics est plus largement utilisée. Les paramètres radiomiques calculés par ce logiciel suivent les recommandations de l'« Image biomarker standardisation initiative », l'IBSI[123]. Les logiciels radiomiques conformes à l'IBSI semblent augmenter la fiabilité statistique des paramètres[147]. C'est pourquoi le logiciel Pyradiomics a été choisi pour cette étude.

(1) Pyradiomics

Pour l'analyse radiomique, le package Pyradiomics a donc été utilisé[148]. Il s'agit d'un package python open-source qui extrait des caractéristiques radiomiques, suivant les recommandations de l'IBSI, destinées au traitement d'imagerie médicale.

Les données d'entrée sont les différentes modalités d'imagerie avec leur masque de segmentation respectif au format nifti ; en sortie, les paramètres radiomiques sont au format JSON («JavaScript Object Notation »), qui est un format de données textuelles.

Les 107 paramètres Pyradiomics ont été extraits par Python, collectés et enregistrés sous Matlab. Parmi ces 107 paramètres, on compte 18 paramètres de premier ordre représentant la distribution des intensités des voxels dans le masque, et sont calculés à partir de l'histogramme des intensités des voxels. Les 24 paramètres calculés par GLCM mettent en évidence la distribution des valeurs de pixels cooccurrents. 14 paramètres sont liés à la forme du masque de segmentation. A partir d'une GLRLM, 16 paramètres mettent en évidence les séries de niveaux de gris, c'est-à-dire une ligne de pixels dans une certaine direction avec la même valeur d'intensité. 16 paramètres ont été calculés par une Gray Level Size Zone Matrix (GLSZM) qui quantifie les zones de l'image à niveaux de gris identiques. Une Neighbouring Gray Tone Difference Matrix (NGTDM) a calculé 5 paramètres, elle quantifie la différence entre une valeur de gris et la valeur de gris moyenne de ses voisins à une distance donnée. Enfin, 14 paramètres ont été calculés à partir d'une Grey Level Dependance Matrix (GLDM) mettant en évidence la dépendance des niveaux de gris entre eux dans l'image.

(2) Sélection des paramètres radiomiques

Afin d'être le plus spécifique possible de la transformation anaplasique, il faudra identifier les paramètres radiomiques qui ne sont pas corrélés au volume. En effet, le volume, et ses dérivés, sont les données de référence pour définir la progression. Il est donc indispensable de ne pas avoir de lien direct entre la métrique définissant la progression et les paramètres pour lesquels la capacité à détecter la progression va être testée. Par exemple, un paramètre dont la valeur est un nombre de voxels sera directement associé au volume tumoral. Nous avons donc recherché des paramètres radiomiques indépendants du volume.

Pour chacune des modalités d'imageries étudiées (T2 FLAIR, 3D T1 et 8 cartes paramétriques de perfusion), 107 paramètres radiomiques sont calculés, ce qui représente 1070 paramètres radiomiques par acquisition IRM. Une partie d'entre eux sont calculés à partir du volume, du nombre de voxels, ou d'une autre mesure corrélée à la taille de la région d'intérêt (masque de segmentation de la région tumorale). Afin d'identifier un groupe de paramètres caractéristiques de la transformation anaplasique corrélés aux informations de structure et de texture de l'image, il est nécessaire de s'affranchir le plus possible des paramètres corrélés à des valeurs de dimension de la tumeur. En effet, l'évolution du volume est un paramètre qui a beaucoup de poids dans la prédiction de la transformation, Ici, ce sont les variations d'information au sein de l'image, témoins d'une transformation anaplasique des tissus, que l'on cherche à mettre en évidence, un groupe de paramètres indépendants du volume tumoral.

On teste donc pour chacun des 930 paramètres radiomiques différents si la valeur du paramètre est corrélée au volume tumoral pour chaque modalité d'imagerie à chaque acquisition IRM.

On sélectionne les paramètres pour lesquels il y a au moins une modalité d'imagerie non corrélée au volume.

Shape	Firstorder	GLCM	GLDM	GLRLM	GLSZM	NGTDM
Elongation	**10Percentile**	**Autocorrelation**	DependenceEntropy	GrayLevelNonUniformity	GrayLevelNonUniformity	**Busyness**
Flatness	**90Percentile**	**ClusterProminence**	DependenceNonUniformity	**GrayLevelNonUniformityNormalized**	**GrayLevelNonUniformityNormalized**	Coarseness
LeastAxisLength	Energy	**ClusterShade**	**DependenceNonUniformityNormalized**	**GrayLevelVariance**	**GrayLevelVariance**	**Complexity**
MajorAxisLength	**Entropy**	**ClusterTendency**	**DependenceVariance**	**HighGrayLevelRunEmphasis**	**HighGrayLevelZoneEmphasis**	**Contrast**
Maximum2DDiameterColumn	**InterquartileRange**	**Contrast**	GrayLevelNonUniformity	**LongRunEmphasis**	**LargeAreaEmphasis**	**Strength**
Maximum2DDiameterRow	**Kurtosis**	**Correlation**	**GrayLevelVariance**	**LongRunHighGrayLevelEmphasis**	**LargeAreaHighGrayLevelEmphasis**	
Maximum2DDiameterSlice	**Maximum**	**DifferenceAverage**	**HighGrayLevelEmphasis**	**LongRunLowGrayLevelEmphasis**	**LargeAreaLowGrayLevelEmphasis**	
Maximum3DDiameter	**MeanAbsoluteDeviation**	**DifferenceEntropy**	**LargeDependenceEmphasis**	**LowGrayLevelRunEmphasis**	**LowGrayLevelZoneEmphasis**	
MeshVolume	**Mean**	**DifferenceVariance**	**LargeDependenceHighGrayLevelEmphasis**	**RunEntropy**	SizeZoneNonUniformity	
MinorAxisLength	**Median**	**InverseDifference**	**LargeDependenceLowGrayLevelEmphasis**	RunLengthNonUniformity	**SizeZoneNonUniformityNormalized**	
Sphericity	**Minimum**	**InverseDifferenceMoment**	**LowGrayLevelEmphasis**	**RunLengthNonUniformityNormalized**	**SmallAreaEmphasis**	
SurfaceArea	**Range**	**InverseDifferenceMomentNormalized**	**SmallDependenceEmphasis**	**RunPercentage**	**SmallAreaHighGrayLevelEmphasis**	
SurfaceVolumeRatio	**RobustMeanAbsoluteDeviation**	**InverseDifferenceNormalized**	**SmallDependenceHighGrayLevelEmphasis**	**RunVariance**	SmallAreaLowGrayLevelEmphasis	
VoxelVolume	**RootMeanSquared**	InformationalMeasureOfCorrelation1	SmallDependenceLowGrayLevelEmphasis	**ShortRunEmphasis**	ZoneEntropy	
	Skewness	InformationalMeasureOfCorrelation2		**ShortRunHighGrayLevelEmphasis**	**ZonePercentage**	
	TotalEnergy	**InverseVariance**		**ShortRunLowGrayLevelEmphasis**	ZoneVariance	
	Uniformity	**JointAverage**				
	Variance	**JointEnergy**				
		JointEntropy				
		MaximalCorrelationCoefficient				
		MaximumProbability				
		SumAverage				
		SumEntropy				
		SumSquares				

Tableau 2: Paramètres radiomiques issus de Pyradiomics (n=107). Les paramètres sectionnés après l'élimination des paramètres liés au volume sont surlignés en gris (n=77).

IV. Prédiction de l'évolution du gliome

Comme évoqué dans notre objectif principal, la 3ième partie de notre travail vise à utiliser les informations radiomiques recueillies dans l'étape précédente et caractérisant la tumeur, afin d'identifier précocement des marqueurs de la transformation ou a minima d'une évolution péjorative du GBG.

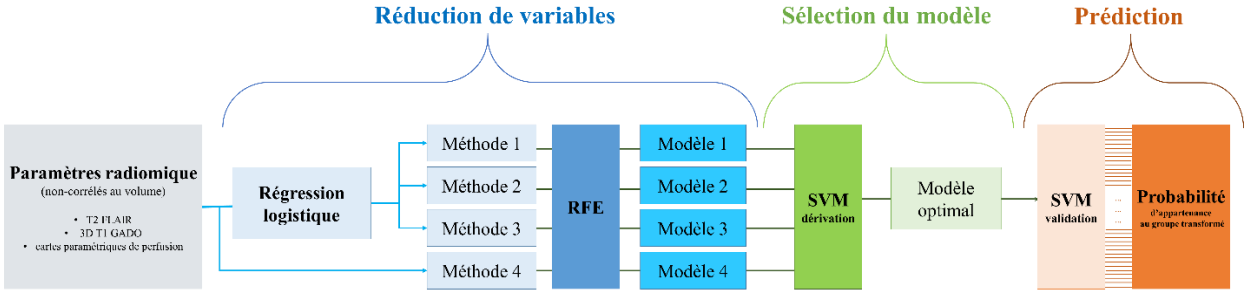


Figure 24: Workflow permettant d'aboutir à la prédiction d'une croissance proche de celle d'un gliome de haut grade au cours du suivi des gliomes de bas grade.

Notre objectif est ici clairement, d'aboutir à un outil et/ou une métrique permettant de renseigner sur le statut évolutif de la tumeur. Nous proposons ainsi un schéma d'analyse en 3 étapes (Figure 24) intégrant, à partir des données radiomiques, (1) une phase de réduction de variables, souvent indispensable pour améliorer la qualité des modèles prédictifs, (2) une phase de sélection du modèle optimal, et (3) une phase de prédiction consistant, à partir du modèle choisi, à obtenir un score de probabilité reflétant la proximité du gliome avec un profil de haut grade. Notre hypothèse étant que cette probabilité permettra d'identifier, de façon transversale, des profils d'évolution péjoratifs à partir des données obtenues au cours du suivi.

Nous utiliserons la même population d'étude qu'à l'étape précédente (III.1 - Population de l'étude radiomique) en définissant les GHG comme étant le groupe qu'il faudra prédire à partir des algorithmes de classification.

a. Réduction de variables

Grâce à la radiomique, des groupes de données aux caractéristiques très proches peuvent être différenciés. Le fait de générer autant de variables permet de mettre en évidence des dissimilarités fines. Mais plus le nombre de variables est élevé, plus il est probable de trouver une combinaison qui sépare les deux groupes par simple hasard. On perd ainsi la capacité de mettre en évidence les vraies variables discriminantes. Le plus souvent, il est nécessaire de réaliser une réduction du nombre de paramètres radiomiques afin d'éviter cet « overfitting », un sur-ajustement lié au

déséquilibre entre le très grand nombre de paramètres et le nombre de cas étudiés. Ce déséquilibre peut mener à de fausses conclusions liées au hasard statistique.

1. Principaux algorithmes

Il existe plusieurs méthodes de réduction du nombre de variables, mais aucune recommandation actuelle ne permet de trancher entre elles. Toutes ces approches ont pour but de n'identifier que les variables pertinentes en éliminant les variables considérées comme du « bruit ».

Parmi les approches les plus connues, 4 vont être présentées ici.

La sélection de variables peut être réalisée à partir d'un test statistique qui mettra en évidence les relations entre les différents paramètres (t-test, ANOVA, Wilcoxon…). L'analyse des composantes principales (PCA)[149] est aussi souvent utilisée et se sert de l'algèbre linéaire pour transformer les données étudiées en données compressées dans un nombre de dimensions et de composante principale donné[150]. Un algorithme issu de l'apprentissage automatique d'élimination récursive des paramètres (ou RFE)[151] peut aussi être utilisé : il fonctionne en supprimant de manière récursive des paramètres et en construisant un modèle sur les paramètres restants. Il utilise la précision du modèle (accuracy) pour identifier les paramètres (et la combinaison de paramètres) qui contribuent au mieux à la prédiction recherchée. Enfin, il est courant d'utiliser l'algorithme de "Least Absolute Shrinkage and Selection Operator" (LASSO)[152,153] qui effectue une régression logistique tout en pénalisant les paramètres peu informatifs afin de les éliminer du modèle final.

Certaines études incluent la comparaison de plusieurs méthodes de réduction de variables afin de sélectionner celle qui présente la meilleure performance[154]. Les méthodes le plus souvent utilisées restent : (1) la réduction sur la base d'analyse statistique standard, où l'on recherche les variables étant statistiquement différentes entre les groupes afin de les inclure dans le modèle (Wilcoxon[16], régression logistique) ; et (2) l'élimination récursive de variables (ou RFE) combinant un classifieur et un algorithme de calcul de rang des paramètres afin de définir, sur la base d'un critère de classification (par exemple la précision), les variables à retenir[127,128].

Les deux approches utilisées vont être décrites, puis leur efficacité sur les données de l'étude sera testée.

a) Réduction par analyse statistique : Régression logistique

La régression logistique[155] est un modèle statistique linéaire qui mesure l'association entre une variable quantitative et la survenue d'un événement qualitatif, souvent représenté de manière

dichotomique (0 ou 1)[156]. Dans le cas d'un problème dichotomique comme ici, c'est une fonction sigmoïde qui a été utilisée comme fonction de régression logistique afin de caractériser la distribution des paramètres radiomiques en fonction du statut GBG (0) ou GHG (1). Le seuil de classification est généralement fixé à 50% de la fonction de régression logistique. Lorsque la valeur observée est supérieure à ce seuil, l'événement (0 ou 1) est susceptible de se produire, et si elle est inférieure, c'est l'inverse (*Figure 25*).

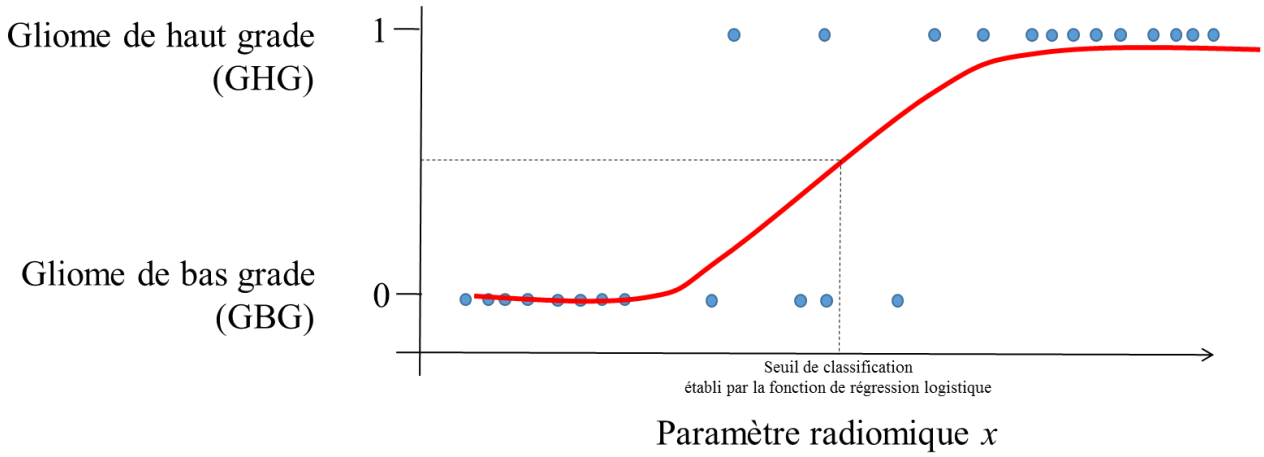


Figure 25: Régression logistique de forme sigmoïde (rouge). Elle est modélisée sur la base de l'observation de la population, classée de manière dichotomique haut (GHG) ou bas grade (GBG). Le seuil de classification est fixé à 50%.

La significativité de la régression logistique est évaluée par le calcul de la p-value. La p-value teste si le classement par la régression logistique du paramètre radiomique étudié est fiable et non dû au hasard. Plus un paramètre a des valeurs spécifiques et très différentes pour chacun des deux groupes, plus le seuil de classification est fiable, et donc plus la valeur de la p-value sera faible.

b) Apprentissage automatique : Élimination récursive de paramètres (RFE)

A partir des données de sortie d'un classifieur (SVM, Random forest…), l'élimination récursive de paramètres (RFE)[151] tente de sélectionner les variables en classant avec un test itératif des ensembles de paramètres de plus en plus petits. En sortie du classifieur, chaque paramètre a un poids qui lui est associé : à chaque itération du RFE, les paramètres les moins bien classés sont éliminés jusqu'à ce que toutes les caractéristiques de l'ensemble de données soient utilisées, ce qui signifie arriver alors à un modèle à une seule variable (ou jusqu'à une limite définie) (***Erreur ! Source du renvoi introuvable.***).

Dans notre étude, nous faisons le choix d'utiliser un classifieur de type « random forest » (RF) ou forêt d'arbres décisionnels. Comme nous le verrons dans les parties suivantes, un classifieur de type support vecteur machine (SVM) sera utilisé pour la prédiction GHG, GBG. Il nous paraissait important de ne pas réutiliser le même algorithme dans l'étape de réduction de variable, ceci afin de limiter les biais et maximiser la fiabilité des résultats.

Le RFE est une méthode d'apprentissage automatique qui permet de gérer des études à grand nombre de paramètres, et qui met en évidence les prédicteurs les plus représentatifs de l'hypothèse étudiée[157]. L'algorithme RFE permet de réduire le nombre de paramètres, car un trop grand nombre de paramètres corrélés entre eux a un impact sur la capacité du RF à identifier les prédicteurs les plus forts[158].

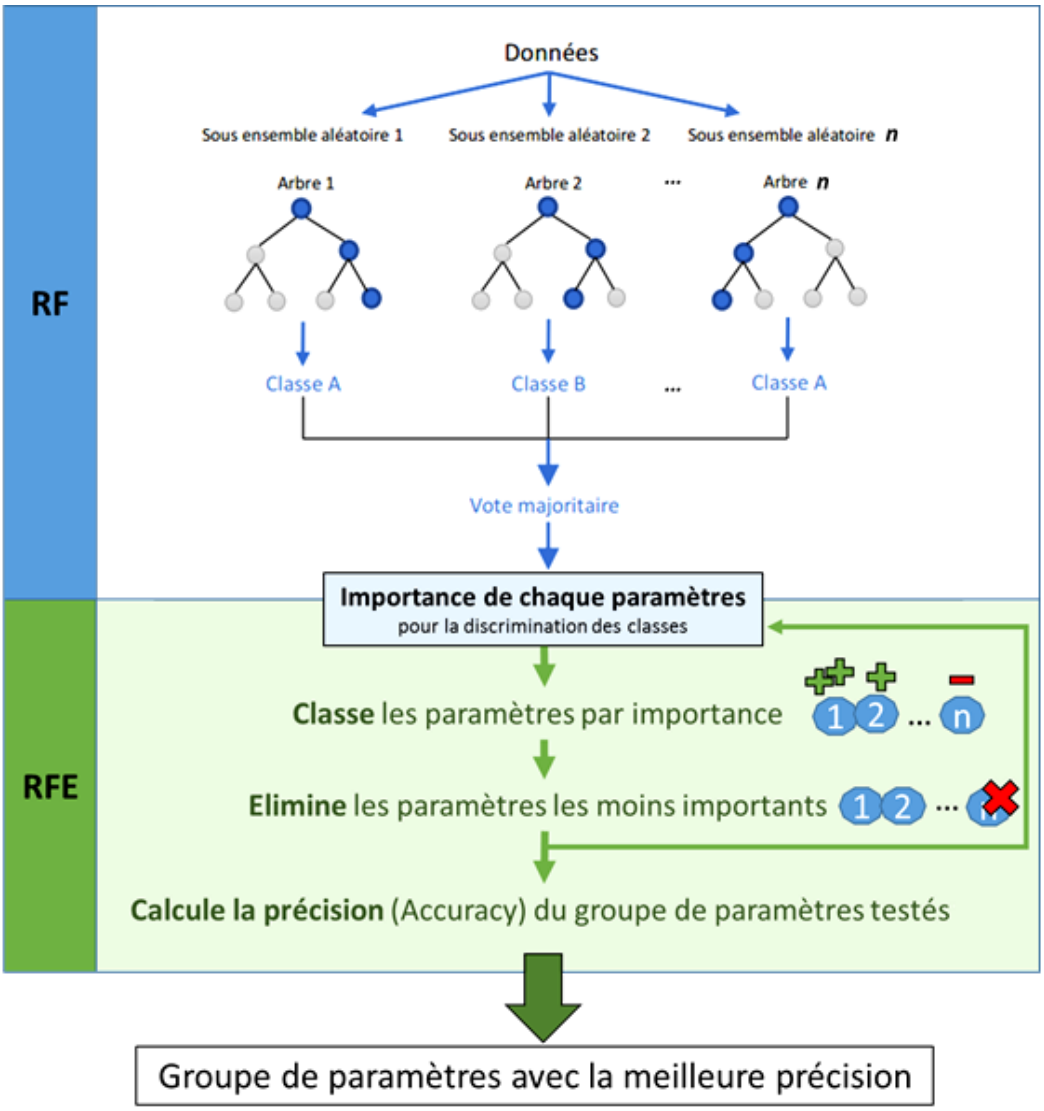


Figure 26: Régression de variable par RFE : Le classifieur en entrée de l'élimination récursive de paramètres (RFE) est de type « forêt d'arbres décisionnels » (« random forest », RF). Le classifieur permet d'attribuer à chaque paramètre, testé dans plusieurs conformations, un critère d'importance. La RFE va, au cours de plusieurs itérations, classer les paramètres par importance et éliminer les moins importants jusqu'à utilisation de tous les paramètres. A chaque itération, la précision (Accuracy) est calculée. Le groupe de paramètres en sortie de RFE est celui qui présente la meilleure précision.

Le résultat obtenu après l'application du RF-RFE est un groupe de paramètres présentant la meilleure précision (accuracy) quant à la différenciation entre GBG et GHG. La précision se calcule à partir des paramètres suivants : le nombre de prédictions exactes représenté par les vrais positifs (VP) et les vrais négatifs (VN), et le nombre de fausses prédictions avec les faux positifs (FP) et les faux négatifs (FN).

Équation 7

$$Précision = \frac{\mathrm{VP} + \mathrm{VN}}{\mathrm{VP} + \mathrm{FP} + \mathrm{VN} + \mathrm{FN}}$$

2. Méthode d'analyse

Afin d'identifier le groupe de paramètres radiomiques les plus pertinents pour notre phase de prédiction, nous avons défini 4 méthodes (*Figure 24*, ***Erreur ! Source du renvoi introuvable.***) que nous comparerons :

- **Méthode 1** : tous les paramètres significativement différents entre les GHG et GBG (régression logistique, p-value < **0.05**) ont été sélectionnés (N=202, *Annexe 1*), puis inclus dans une analyse en RFE
- **Méthode 2** : sélection des 5 paramètres (maximum) les plus significativement différents entre les groupes GHG et GBG (régression logistique, p-value < **0.05**) pour chaque modalité (N=50, *Annexe 1*), puis inclus dans une analyse RFE
- **Méthode 3** : tous les paramètres significativement différents entre les GHG et GBG (régression logistique, p-value < **0.01**) ont été sélectionnés (N=87, *Annexe 1*), puis inclus dans une analyse en RFE
- **Méthode 4** : sélection de variables par analyse RFE uniquement sur l'intégralité des variables disponibles (N = 770, *Annexe 1*)

Le critère de sélection du RFE a été défini comme étant la précision.

3. Résultats

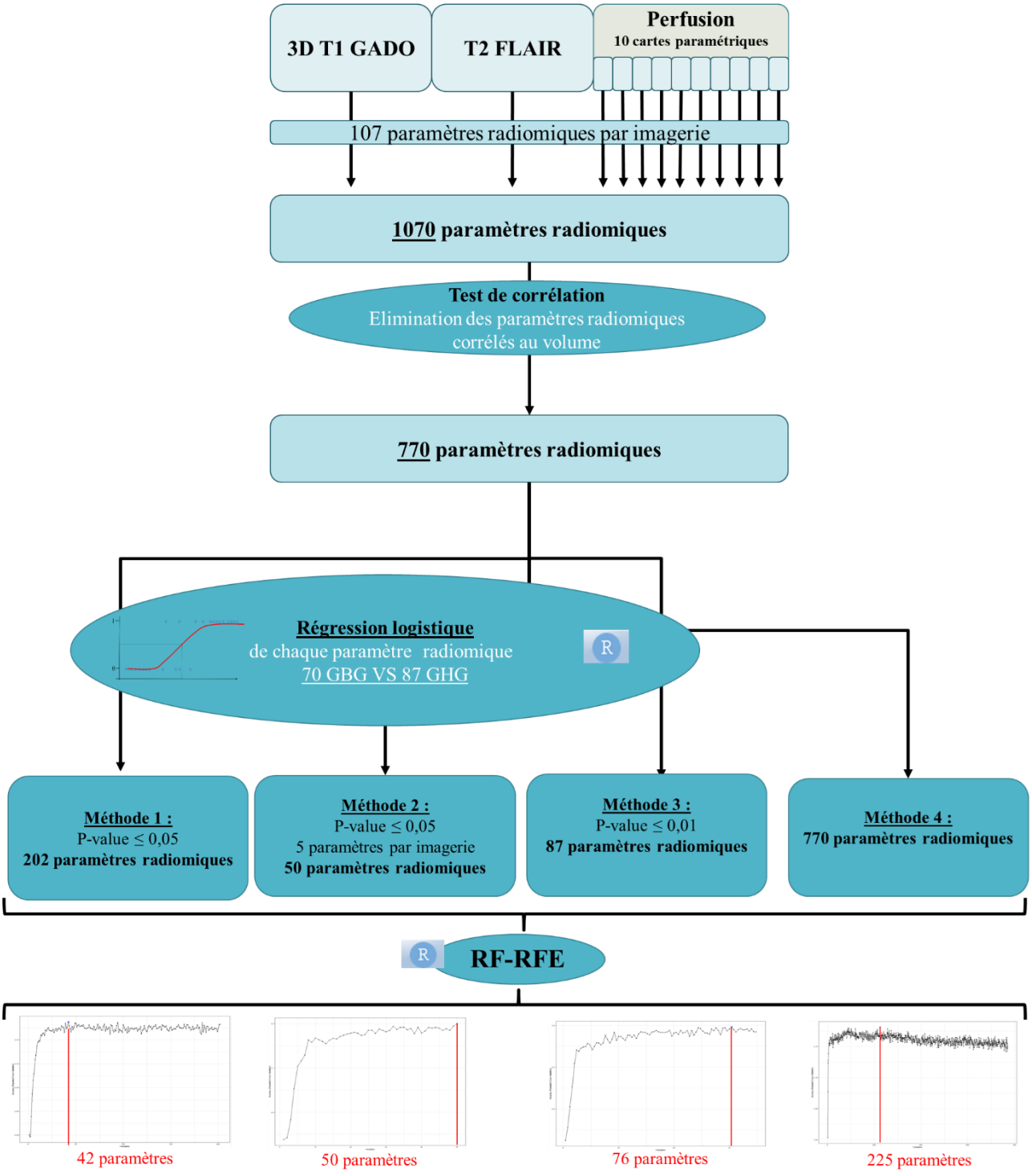


Figure 27: Méthodes de sélection et différents résultats de la réduction des paramètres radiomiques

Après le RFE, un total de 42 paramètres sur 202 ont été retenus pour la méthode 1 (Précision = 0,71), 50 sur 50 pour la méthode 2 (Précision = 0,81), 76 sur 87 pour la méthode 3 et 225 sur 770 pour la méthode 4 (Précision= 0.69) (***Erreur ! Source du renvoi introuvable.***).

Notons que pour la méthode 2, aucune variable n'a pu être exclue.

Pour la méthode 1, 14.3% (N=6) des paramètres radiomiques sont issus d'une analyse de premier ordre de l'histogramme du masque tumoral, 14.3% (N=6) sont des paramètres calculés à partir de la matrice identifiant les organisations linéaires de niveaux de gris dans l'image GLRLM, 9.5% (N=6) sont extraits de la matrice de dépendance des niveaux de gris de l'image GLDM, 42.9% (N=18) sont calculés à partir de la matrice de cooccurrence GLCM, 9.5% (N=6) à partir de la matrice GLSZM qui identifie les zones de l'image à niveaux de gris identiques, et enfin 9.5% (N=6) à partir de la matrice NGTDM mettant en évidence la différence entre une valeur de gris et les valeurs voisines.

Pour la méthode 2, 12% (N=6) des paramètres radiomiques proviennent d'une analyse de premier ordre, 24% (N=12) sont calculés à partir d'une GLCM, et 24% (N=12) à partir d'une GLDM. 20% (N=10) sont des paramètres calculés à partir de la GLRLM, et les derniers 20% (N=10) à partir de la matrice GLSZM.

Pour la méthode 3, 27.6% (N=21) des paramètres radiomiques proviennent d'une analyse de premier ordre. 31.6% (N=24) sont calculés à partir d'une GLCM, 27.6% (N=7) d'une GLDM, 13.2% (N=10) sont des paramètres calculés à partir de la GLRLM, 14.5% (N=11) à partir de la matrice GLSZM, et les derniers 3.9% (N=3) à partir de la NGTDM

Pour la méthode 4, 20.9% (N=47) des paramètres radiomiques proviennent d'une analyse de premier ordre. 26.2% (N=59) sont calculés à partir d'une GLCM, 11.1% (N=25) à partir d'une GLDM, 15.1% (N=34) sont calculés à partir de la GLRLM, 15.6% (N=35) à partir de la matrice GLSZM, et les derniers 11.1% (N=25) à partir de la NGTDM (*Tableau 3*).

	Méthode 1	Méthode 2	Méthode 3	Méthode 4
Premier ordre	14.3% (N=6)	12% (N=6)	27.6% (N=21)	20.9% (N=47)
GLCM	42.9% (N=18)	24% (N=12)	31.6% (N=24)	26.2% (N=59)
GLDM	9.5% (N=6)	24% (N=12)	27.6% (N=7)	11.1% (N=25)
GLRLM	14.3 (N=6)	20% (N=10)	13.2% (N=10)	15.1% (N=34)
GLSZM	9.5% (N=6)	20% (N=10)	14.5% (N=11)	15.6% (N=35)
NGTDM	9.5% (N=6)		3.9% (N=3)	11.1% (N=25)

Tableau 3:Répartition des différents paramètres radiomiques issus de la sélection par RFE pour chaque méthode de sélection. Les paramètres sont issus soit d'une analyse de premier ordre, d'une matrice de cooccurrence des niveaux de gris (GLCM), d'une matrice de dépendance des niveaux de gris (GLDM), d'une matrice identifiant les lignes de niveaux de gris identiques (GLRLM) , d'une matrice identifiant des zones de niveaux de gris (GLSZM) ou encore d'une matrice mettant en évidence les différences entre une valeur de gris et les valeurs voisines (NGTDM).

b. Sélection du modèle

Les 4 méthodes de réduction de variables aboutissent à des combinaisons différentes de variables. Pour choisir la combinaison la plus performante et fiable, nous ne pouvons pas seulement nous baser sur la précision fournie par le RFE. Il est indispensable d'utiliser des approches permettant d'évaluer les réelles performances des modèles, et surtout leur capacité de généralisation via des approches dites de « validation croisée ». Afin de limiter les biais, nous faisons le choix d'utiliser le support vecteur machine ou « SVM » pour la sélection du modèle ainsi que pour la prédiction d'un algorithme de classification reconnu, dérivé du champ de l'intelligence artificielle. Cela permettra de combiner les avantages du Random Forest (utilisé dans le RFE) et du SVM.

1. Le support vecteur machine

Le SVM est une méthode d'apprentissage automatique qui permet de résoudre une série de problèmes de classification, tout comme la méthode Random Forest (RF). La méthode RF ayant été utilisée pour la réduction de variables et afin de diminuer les biais induit par l'utilisation d'un seul type de méthode, le SVM a été utilisé pour choisir le modèle le plus performant.

Le SVM est un algorithme qui apprend par l'exemple à attribuer des labels aux objets (classifieur). Il répond à deux principes généraux, (1) optimiser la précision du classificateur (maximiser le pourcentage de labels corrects attribués aux nouveaux échantillons) et (2) optimiser sa reproductibilité (garantir que le classificateur peut s'appliquer à de nouvelles données).

Dans un premier temps, les paramètres sélectionnés à l'étape précédente par le RFE sont utilisés comme entrée pour le SVM. Ensuite, le classifieur est entraîné et testé. L'étape finale est

l'évaluation des performances du classificateur : sensibilité, spécificité et précision (*Équation 7*). Ces tests de performance sont calculés à partir des scores vrai positif (VP), vrai négatif (VN), faux positif (FP) et faux négatif (FN).

Logiquement, le test ultime des performances du modèle se fait sur de nouvelles données, inconnues du SVM. La meilleure méthode consiste donc à diviser l'ensemble de données en groupes d'entraînement et de tests dès le début.

Une fois qu'un classifieur tel qu'un SVM a été généré, il peut ensuite être appliqué à de nouveaux échantillons pour prédire à quel groupe ils appartiennent.

Le SVM utilise un apprentissage supervisé. Un ensemble de données est analysé afin de conduire à un modèle qui sépare les données en un nombre souhaité et distinct de classes.

Pendant la phase d'apprentissage, le modèle fait correspondre les données d'entrée aux données de sortie sur la base de nombreux exemples de paires entrée-sortie. Dans cette phase, le SVM est entraîné à l'aide d'observations d'exemples pour lesquels le groupe d'appartenance est identifié et connu (par exemple « gliome de bas grade » et « gliomes de haut grade »). Les caractéristiques correspondant à un ensemble d'observations peuvent être utilisées pour entraîner une fonction de décision qui génère des affectations de classe (ou labels).

La fonction de décision est représentée par un hyperplan optimal qui sépare (ou classifie) les paramètres appartenant aux différentes classes, sur la base de modèles de données. Cet hyperplan sera ensuite utilisé pour définir le label le plus probable pour de nouvelles données (non analysées par le SVM). Les paramètres utilisés pour définir l'hyperplan sont référencés par des coordonnées en fonction de leurs relations et forment ce qu'on appelle les vecteurs de support (*Figure 23*).

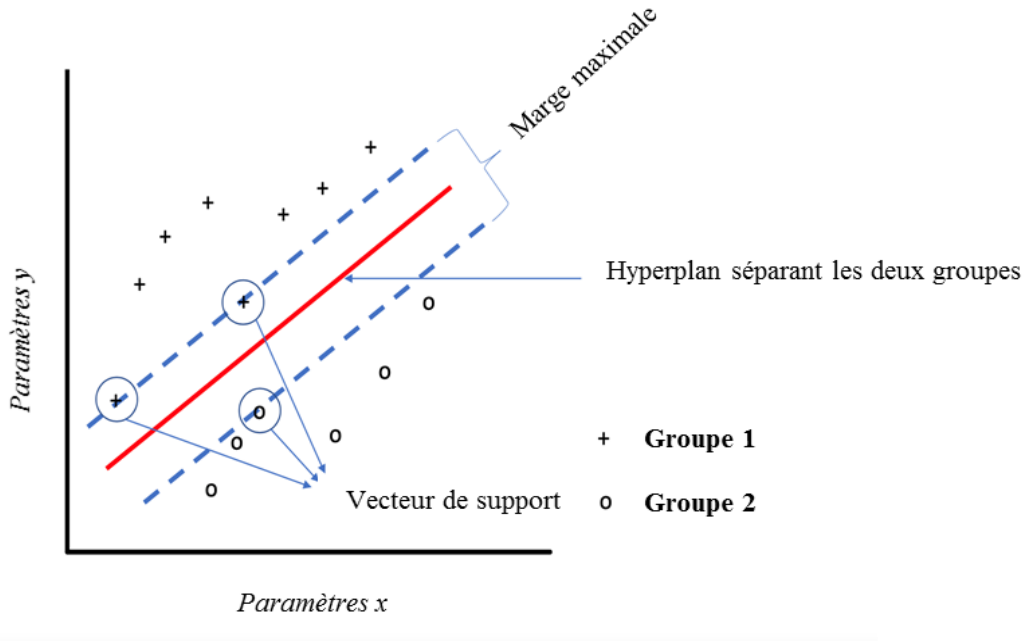


Figure 28: Illustration de l'hyperplan qui sépare au maximum les vecteurs de support de chacun des 2 groupes.

L'hyperplan de marge maximale est celui qui maximise la distance (la "marge") entre les vecteurs de support des deux labels de classe, afin de prédire la classification correcte pour de nouvelles données. L'hyperplan peut être une ligne lorsque deux dimensions sont étudiées, ou un plan bidimensionnel lorsque 3 caractéristiques sont utilisées.

Une des limites des SVM est le sur-ajustement (« overfitting »), qui correspond à la perte de la capacité de généralisation de l'apprentissage. Les caractéristiques particulières de la radiomique (principalement le petit nombre d'échantillons et le très grand nombre de variables) peuvent théoriquement influencer cet « overfitting », c'est-à-dire aboutir à un ajustement excessif du classifieur SVM. L'application de marge douce peut aider à éviter ce problème. C'est un hyperplan qui paraît moins adapté qu'une marge « rigide » à discerner les deux groupes d'entraînement, mais qui aura une meilleure performance face à de nouvelles données (*Figure 29*).

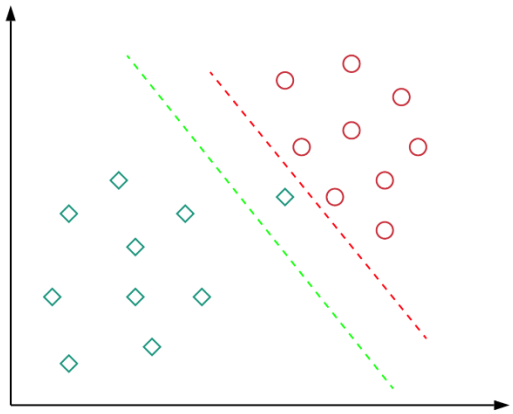


Figure 29: "Marge douce". L'hyperplan rouge sépare parfaitement tous les points de formation. Cependant, l'hyperplan vert a une marge plus faible et peut être plus apte à généraliser avec des données non vues.

Le SVM a recours à des noyaux ou kernels. Ces fonctions noyaux sont des fonctions mathématiques qui permettent la projection des données à classer dans une espace à plus grande dimension, un espace vectoriel[159]. Dans cette espace, le SVM est capable d'identifier l'hyperplan qui sépare le mieux les deux classes, non séparables dans un espace de plus faible dimension.

Lorsqu'on entraîne un modèle sur des données labélisées, on émet l'hypothèse qu'il doit également fonctionner avec de nouvelles données. Afin de s'assurer de l'exactitude ou non de ses prédictions, une confirmation supplémentaire est nécessaire.

La validation croisée est une méthode statistique qui permet d'évaluer la capacité de généralisation d'un modèle à partir d'une hypothèse. Dans notre étude, parmi les 4 modèles découlant des 4 méthodes de sélection, il s'agit d'identifier celui qui a la meilleure performance et fiabilité pour discriminer nos deux groupes (GBG et GHG).

Nous avons reproduit la méthodologie de Nishi et al.[160] en utilisant une validation croisée imbriquée en 10 niveaux (« 10 fold cross-validation ») pour la dérivation du modèle et l'évaluation interne.

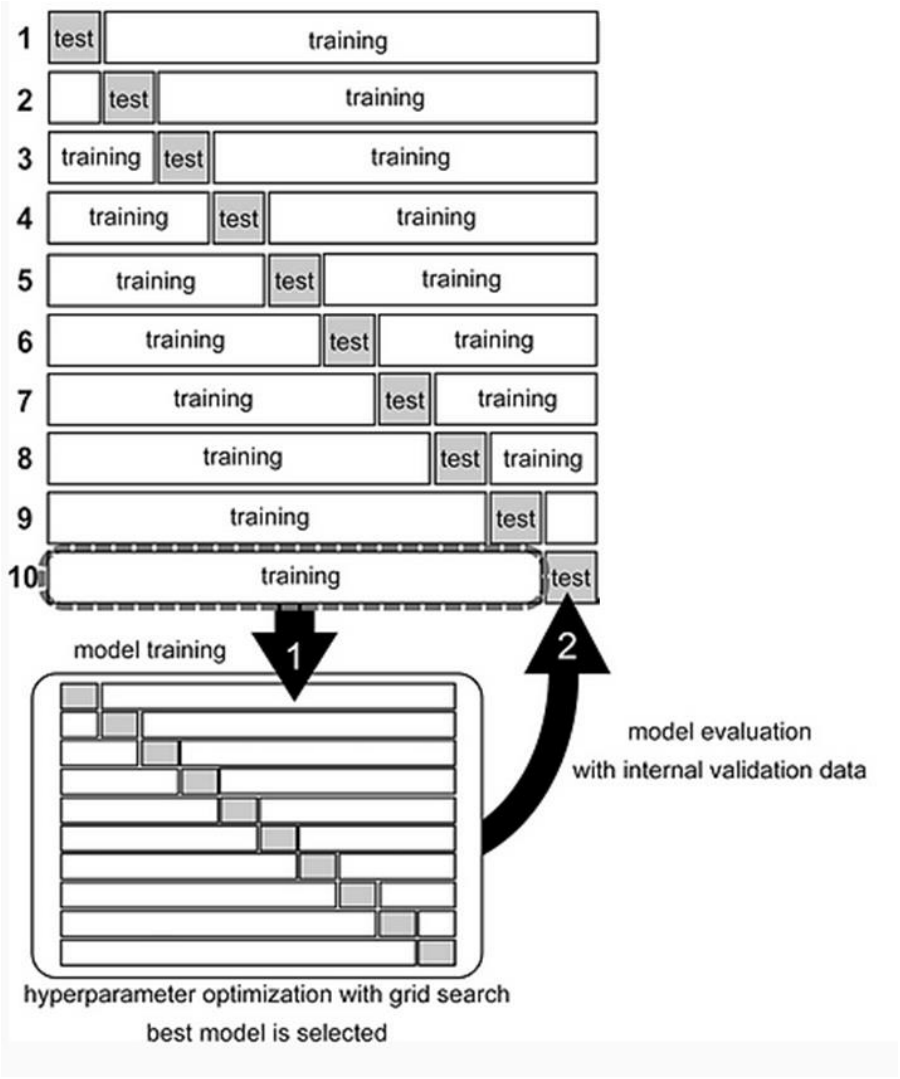


Figure 30: Validation croisée imbriquée en 10 niveaux (Figure extraite de Nishi et al.(158)

En bref, cette validation croisée divise l'ensemble des données en données d'entraînement et données de test menant à 10 modèles avec 10 divisions différentes (*Figure 30*). À chaque étape, un modèle est formé en optimisant, en interne, les hyper paramètres des données d'entraînement avec un algorithme de grille de recherche. Une fois le modèle créé, la sensibilité, la spécificité, la précision et l'aire sous la courbe ROC (AUC) ont été calculées pour chaque ensemble de données de validation interne. Enfin, pour chacun des 10 modèles, on obtient sensibilité, spécificité, précision et AUC (*Figure 30*).

2. Les critères de sélection

Chacun des 4 groupes de paramètres radiomiques a été utilisé comme base pour un SVM avec validation croisée imbriquée, afin que celui-ci calcule leurs performances pour la discrimination entre GBG et GHG. Pour chaque groupe, la précision et l'aire sous la courbe ROC

(AUC) ont été calculées, ainsi que les écarts types de ces valeurs. La courbe ROC est une courbe reflétant la qualité du modèle en rapportant la sensibilité en fonction de la spécificité.

Nous pouvons ainsi définir un modèle considéré parfait comme étant un modèle présentant une précision de 1, une AUC de 1 avec des déviations standard nulles. La distance euclidienne des performances des modèles à ce point parfait permettra d'identifier le meilleur modèle :

Équation 8

$$D = \sqrt{(AUC_{ref} - AUC)^2 + (Prec_{ref} - Prec)^2 + (SD(AUC)_{ref} - SD(AUC))^2 + (SD(Prec)_{ref} - SD(Prec))^2}$$

Avec AUC_{ref} l'AUC du point parfait (1) et $Prec_{ref}$ la précision de ce même point (1), leur déviation standard associées est notée $SD(AUC)_{ref}$ (0) et $SD(Prec)_{ref}$ (0). AUC, $Prec$, $SD(AUC)$ et $SD(Prec)$ sont les paramètres du modèle testé.

3. Résultats

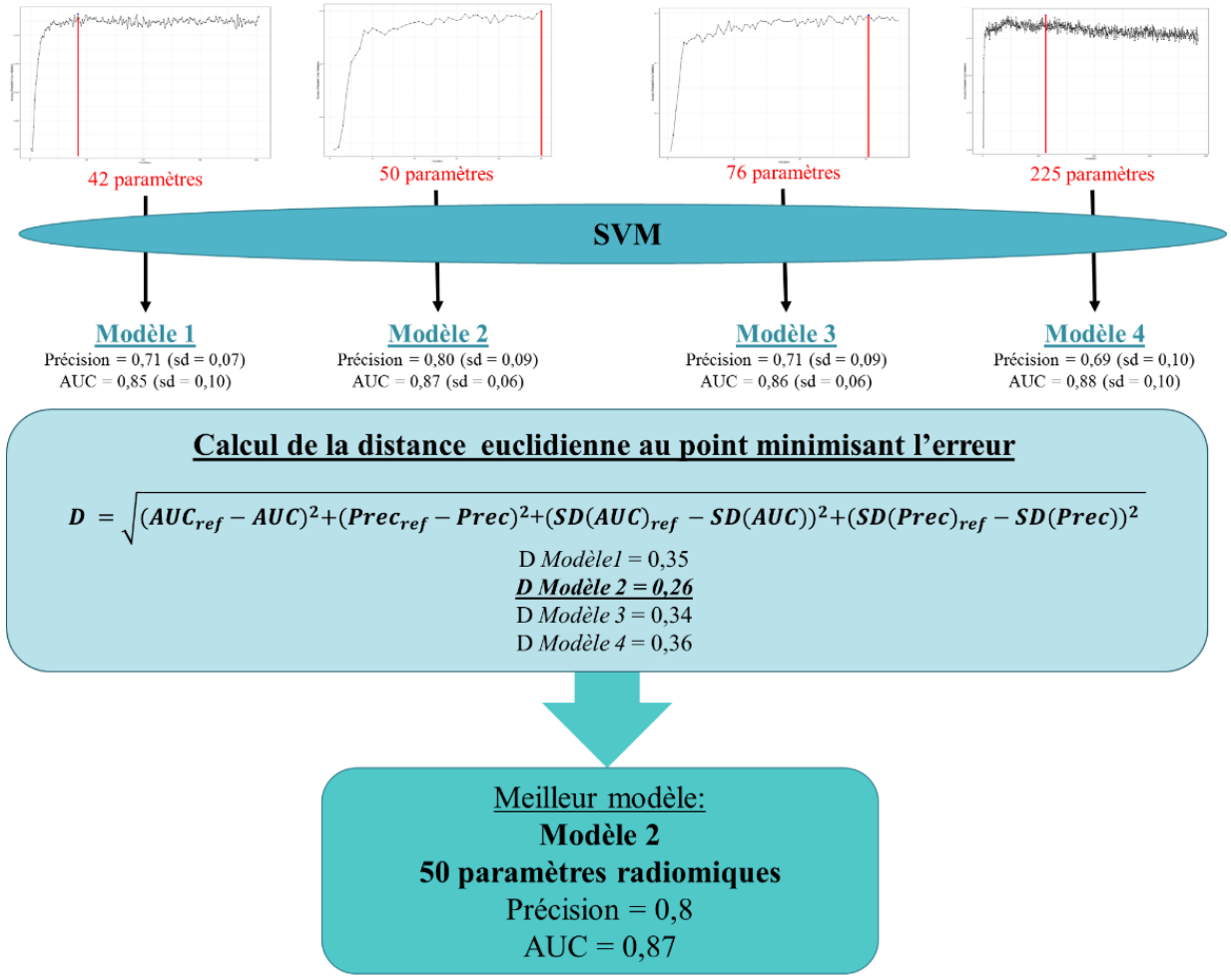


Figure 31 : Résultats de la sélection de modèles : Évaluation de la performance de chaque modèle par SVM et calcul de la distance euclidienne au point parfait.

Le modèle 1, correspondant à la méthode de sélection 1 ($p < 0.05$), présentait une AUC de 0.85 ± 0.1 et une précision de 0.71 ± 0.07.

Le modèle 2, correspondant à la méthode de sélection 2 (5 meilleurs variables par modalité), a obtenu une AUC de 0.87 ± 0.06 et une précision de 0.8 ± 0.09 (Tableau 4).

Le modèle 3, correspondant à la méthode de sélection 3 ($p < 0.01$), présentait une AUC de 0.86 ± 0.06 et une précision de 0.71 ± 0.09.

Le modèle 4, correspondant à la méthode de sélection 4 (toutes variables), présentait une AUC de 0.88 ± 0.1 et une précision de 0.69 ± 0.1.

La méthode 2 a obtenu la plus faible distance au point optimal (0.26) contre 0.35 pour la méthode 1, 0.34 pour la méthode 3 et 0.36 pour la méthode 4 (*Figure 31*).

Méthode de sélection 2 (5 paramètres par modalité d'imagerie à $p < 0.05$)	
T2 FLAIR	**3D T1 GADO**
FLAIR_original_firstorder_Maximum	T1_original_glcm_SumEntropy
FLAIR_original_ngtdm_Complexity	T1_original_glcm_MaximumProbability
FLAIR_original_firstorder_Range	T1_original_glrlm_RunEntropy
FLAIR_original_glcm_Idn	T1_original_glcm_ClusterShade
FLAIR_original_glrlm_RunEntropy	T1_original_firstorder_Entropy
Cartes paramétrique des perfusion	
rBF	**TMAX**
rBF_original_glrlm_RunEntropy	TMAX_original_gldm_LargeDependenceHighGrayLevelEmphasis
rBF_original_glszm_GrayLevelNonUniformityNormalized	TMAX_original_gldm_DependenceNonUniformityNormalized
rBF_original_glcm_SumEntropy	TMAX_original_gldm_SmallDependenceEmphasis
rBF_original_glszm_HighGrayLevelZoneEmphasis	TMAX_original_gldm_LargeDependenceEmphasis
rBF_original_glszm_SmallAreaHighGrayLevelEmphasis	TMAX_original_glszm_ZonePercentage
rBV	**K2**
rBV_original_glszm_GrayLevelNonUniformityNormalized	K2_original_glrlm_LongRunHighGrayLevelEmphasis
rBV_original_glrlm_RunEntropy	K2_original_glszm_SmallAreaHighGrayLevelEmphasis
rBV_original_glcm_SumEntropy	K2_original_gldm_HighGrayLevelEmphasis
rBV_original_gldm_LargeDependenceHighGrayLevelEmphasis	K2_original_gldm_SmallDependenceHighGrayLevelEmphasis
rBV_original_glrlm_LongRunHighGrayLevelEmphasis	K2_original_glcm_Autocorrelation
rBVc	**tMIP**
rBV_corrected_original_glrlm_RunEntropy	tMIP_original_firstorder_Skewness
rBV_corrected_original_glszm_GrayLevelNonUniformityNormalized	tMIP_original_glcm_SumEntropy
rBV_corrected_original_firstorder_InterquartileRange	tMIP_original_glszm_GrayLevelNonUniformityNormalized
rBV_corrected_original_glcm_JointAverage	tMIP_original_glrlm_RunEntropy
rBV_corrected_original_glcm_SumAverage	tMIP_original_firstorder_Kurtosis
MTT	**TTP**
MTT_original_glcm_Idn	TTP_original_gldm_LargeDependenceEmphasis
MTT_original_gldm_DependenceNonUniformityNormalized	TTP_original_gldm_LargeDependenceHighGrayLevelEmphasis
MTT_original_gldm_SmallDependenceEmphasis	TTP_original_glrlm_RunPercentage
MTT_original_glszm_ZonePercentage	TTP_original_gldm_DependenceVariance
MTT_original_glcm_Id	TTP_original_glrlm_LongRunEmphasis

Tableau 4: Paramètres radiomiques de la méthode 2 calculés sur les différentes modalités d'imagerie : 3D T1 GADO, T2 FLAIR et les cartes paramétriques de perfusion rBF, TMAX, rBV, K2,rBVc, tMIP, MTT et TTP.

4. Discussion

C'est la méthode 2 qui a obtenu la meilleure performance. Ce modèle est celui dont la méthode sélectionnait les meilleurs paramètres radiomiques de chaque modalité d'imagerie. Ce résultat va dans le sens des études actuelles qui ont identifié la transformation anaplasique comme un phénomène multimodal pouvant être décrit par plusieurs types d'imagerie (*Tableau 4*).

On peut tout d'abord citer les imageries anatomiques. L'imagerie 3D T1 GADO permet d'identifier la prise de contraste qui est synonyme de la transformation, et le suivi volumique est réalisé sur les imageries T2 FLAIR. A partir de l'imagerie T2, une hétérogénéité de signal au cours du suivi peut permettre d'identifier un état de croissance accéléré[12] ou encore une zone en cours de transformation[161].

La transformation anaplasique implique une néoangiogenèse au cours de laquelle de multiples vaisseaux vont se former afin d'assister la multiplication cellulaire de la tumeur. Cette prolifération vasculaire rapide et chaotique[162] explique l'intérêt des études pour les imageries de perfusion. Il a été montré que les cartes paramétriques de volume sanguin cérébral (rCBV) permettent l'identification d'une aggravation du comportement évolutif des GBG[8,50,75]. Si les variations de rCBV sont représentatives d'une transformation, il est logique que les signaux de débit rCBF[163,164], de modification de vascularisation (tMIP)[165] et de perméabilité vasculaire K2[50] montrent aussi une modification entre les GBG et les GHG étudiés dans cette étude. On peut aussi s'attendre à ce que les paramètres de temps de transit moyen (MTT), de temps au pic de contraste dans les tissus (TTP) et de temps pour atteindre le maximum de la fonction résidu (Tmax) soient modifiés (*I.5.a)(3).Imagerie de perfusion et cartographies de perfusion*).

Ce modèle comportant des paramètres extraits de diverses modalités d'imagerie s'ancre dans une logique, de plus en plus étudiée, d'étude multimodale des gliomes afin d'identifier une transformation[166,167,127].

c. Prédiction

La dernière étape de ce travail consiste à prédire un statut pour tous nouveaux examens IRM, en l'occurrence les dates IRM de suivi des 35 patient·e·s de cette étude: GHG ou GBG. Les approches de SVM permettent d'estimer une probabilité allant de 0 à 1 d'appartenir à l'un ou l'autre des groupes. Plus on est proche de 0, plus l'algorithme considère que la tumeur étudiée est un GBG, et au contraire, plus on se rapproche de 1, plus l'algorithme considère être en présence d'une lésion qui n'est pas de bas grade, plus proche d'un GHG. Nous pouvons considérer cette métrique de probabilité comme un score composite issue de l'analyse SVM qui résume les paramètres radiomiques permettant de distinguer les deux groupes.

Ainsi, nous faisons l'hypothèse que ce score, bien que méthodologiquement conçu pour être indépendant du volume, y soit corrélé. On s'attend à ce que la radiomique soit sensible aux phénomènes biologiques qui induisent une croissance de la lésion tumorale. Ce score de probabilité évoluerait ainsi au cours du suivi, traduisant de façon indépendante une évolution des GBG.

1. SVM de validation

A ce stade, le modèle SVM choisi, comme sa généralisation, a déjà été mis à l'épreuve par l'approche précédente (*IV.1.Le support vecteur machine*). Il n'est donc plus nécessaire d'appliquer des entraînements avec validation croisée. Au contraire, nous souhaitons maintenant que

l'algorithme apprenne avec un maximum de données pour obtenir la meilleure classification c'est-à-dire l'appartenance à l'un des deux groupes (GBG ou GHG). Ainsi l'intégralité des données d'entraînement a été utilisée, à savoir les GBG confirmés et les GHG confirmés. Les dates d'examen labellisées GBG (N=70) représentaient des GBG en début de suivi, considérées plus éloignées d'une prise de contraste, et les examens labellisés GHG des gliomes de haut grade secondaires (N=86). Le modèle ainsi construit a ensuite été utilisé pour estimer, pour chaque examen de suivi, la probabilité d'appartenance à l'un ou l'autre des 2 groupes.

2. Pertinence du score de probabilité SVM

Ce score de probabilité peut avoir de multiples applications. Si on l'analyse brutalement, il peut permettre de classer les examens en GHG ou GBG sur la base d'un seuil rigide de 0,5. Mais il est également possible d'évaluer l'évolution de ce score au cours du suivi : Évolue-t-il ? Dans le même sens que le volume ? Varie-t-il plus précocement ? Nous pouvons également nous poser la question de sa capacité à prédire non pas les groupes GHG / GBG mais les marqueurs connus comme étant signe de dégradation (VED etc.).

a) Prédiction du statut GHG – GBG

L'algorithme a appris à différencier les GHG des GBG, le score de probabilité généré a ainsi pour fonction première théorique de classer les nouvelles données.

Les examens (nouvelles données) ont été les examens IRM du suivi des GBG (n=212, voir *Figure 16*) et les examens IRM considérés comme des GHG après l'identification de la transformation (non inclus dans l'apprentissage n=38).

(1) Méthodologie

Afin de définir le statut de chaque IRM, un seuil est appliqué sur le score de probabilité. Les examens ayant un score inférieur à 0,5 sera labellisé GBG, et ceux supérieurs à 0,5 GHG. Pour ces données, le statut exact du sujet au moment de l'IRM n'est pas clairement connu, nous considèrerons alors que la prise de contraste représente la définition du statut GHG, les examens suivant une prise de contraste étant eux même définis GHG. Tous les autres seront définis GBG.

Les taux de vrai positif (VP), vrai négatif (VN), faux négatif (FN), faux positif (FP), ainsi que la sensibilité, et la spécificité seront calculés (*Tableau 5*).

Équation 9

$$Sen = \frac{VP}{VP + FN}$$

Équation 10

$$Spe = \frac{VN}{FP + VN}$$

(2) Résultats

VN	VP	FP	FN	Sen	Spe
58.4% (n=146)	7.6% (n=19)	26.4% (n=66)	7.6% (n=19)	0.5	0.69

Tableau 5: Résultats de la classification SVM au seuil de 0.5

Les taux de VP représentent une probabilité calculée par le SVM supérieure au seuil de 0,5 pour un examen post prise de contraste (GHG), et les VN représentent des probabilités inférieures à 0,5 pour des examens GHG de suivi (*Tableau 5*). Les erreurs représentées par les FP sont des examens appartenant au groupe GHG de suivi avec une probabilité calculée par le SVM supérieure à 0,5, donc faussement classés GHG. Enfin, les erreurs comptabilisées par les FN sont des examens GHG avec un seuil inférieur à 0,5.

Il a été observé que le SVM, avec un seuil de 0,5, permettait de classer correctement 50% des examens GHG (19/38) et 67% des GBG (146/212). La sensibilité était de 0,5 et la spécificité de 0,69.

(3) Discussion

En considérant un seuil rigoureux et admis de 0,5, la classification affiche des résultats modérés. Cela peut s'expliquer par le fait que nous avions à valider très peu de gliomes transformés au cours du suivi. De plus, si la prise de contraste représente un des marqueurs spécifiques de la transformation en GHG, il arrive cependant qu'un GBG devienne GHG sans prise de contraste apparente[168]. Le SVM apprend à reconnaître les groupes par une multitude de métriques radiomiques, en l'occurrence 5 par modalité. La prise de contraste visible en imagerie pondérée T1 avec injection de produit de contraste ne représente au final que 10% de l'information avec laquelle il a appris : il est donc possible que les FP observés traduisent une évolution péjorative de la

tumeur. Il est cependant impossible de le confirmer en l'absence d'analyse anatomopathologique. Ici, on cherche à prédire les GHG, et les résultats des mesures de sensibilité et spécificité traduisent la difficulté de cette prédiction fine ; si on avait visé à prédire les GBG, ces valeurs auraient été bien meilleures pour cette population mieux connue, générant donc moins d'erreurs.

Néanmoins, ces résultats ne sont pas décourageants car ce qui est recherché ici, c'est un score composite de toutes les données d'entraînement reflétant une évolution des GBG.

b) Évolution du score de probabilité dans le suivi

Indépendamment d'une notion de seuil, il semble pertinent d'analyser la variation du score de probabilité au cours de son suivi. Bien que nous ayons mis en place une série de prétraitements visant à s'assurer de l'indépendance des métriques sources vis-à-vis du volume, il n'en reste pas moins que les deux paramètres peuvent être corrélés car reflétant en théorie tous deux des aspects d'évolution de la tumeur.

(1) Méthodologie

Les évolutions du score de probabilité seront tracées au cours du suivi pour les 35 patient·e·s de la base de données, présentant des profils différents. Les variations des dimensions de la tumeur seront indiquées comme référence par le diamètre moyen tumoral (DMT (mm)) qui sera calculé à partir du volume (*Équation 5*). Enfin, les interventions chirurgicales, les traitements de chimiothérapie et radiothérapie seront également indiqués sur les courbes.

La corrélation entre le score de probabilité et le volume sera évaluée par un coefficient de corrélation de Pearson, le seuil de significativité étant définie à $p < 0.05$.

(2) Résultats

Pour les 35 patient·e·s de la base de données de cette étude, les courbes d'évolution de DMT (mm) ont été représentées conjointement à la valeur du score de probabilité, à chaque date d'examen IRM et au cours du temps (en mois). La totalité de ces courbes sera disponible en *Annexe 2*.

De par le nombre de sujets et le fait que chaque suivi est unique, chacun possède une courbe d'évolution du suivi qui lui est propre. On peut néanmoins mettre en évidence des enchaînements de décision de traitements (chirurgie, chimiothérapie…) caractéristiques du suivi des GBG. En effet la majorité des suivis présente une chute drastique de volume, survenant juste après l'examen IRM de diagnostic, indiquant la survenue d'une chirurgie en première intention. Notons que certains suivis

ne commencent pas lors de l'examen de découverte du gliome, ceci étant dû à la sélection des données, car l'imagerie de perfusion n'était souvent pas disponible.

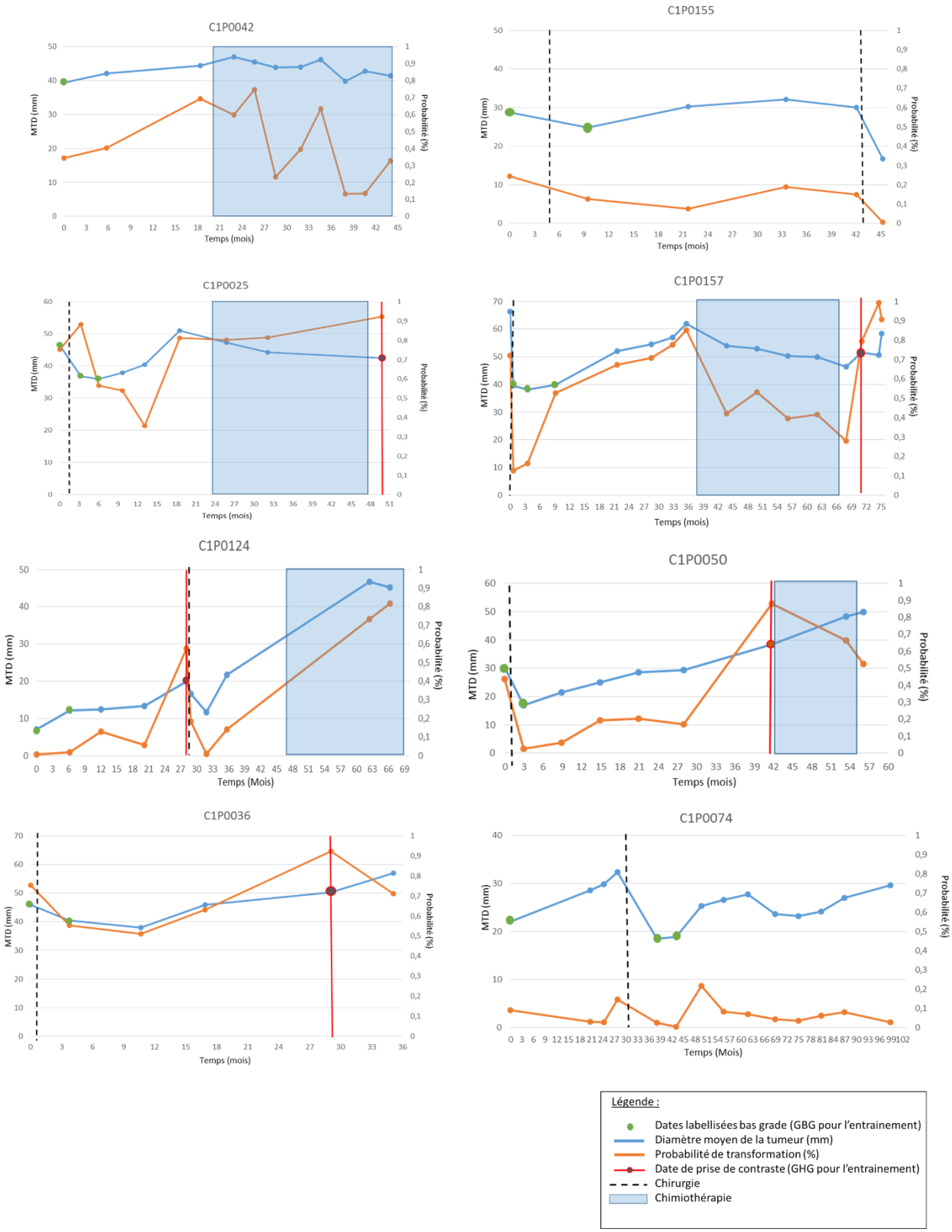


Figure 32: Exemple de 8 différents profils d'évolution volumique (diamètre moyen tumoral -MTD ou DMT- en bleu) et des valeurs de probabilité calculées par le SVM pour chaque date d'examen (en orange) en fonction du temps (mois). Les dates d'examen sélectionnées pour l'entraînement du SVM sont notées en vert et les dates d'examen de l'apparition de la prise de contraste, labélisées GHG pour l'entraînement, en rouge. La chirurgie entre deux points est notée en pointillés noirs, la période de chimiothérapie est en bleu.

Quelques exemples décrivant différents cas de figure peuvent néanmoins être présentés (Figure 32).

Au cours du suivi des deux premiers sujets présentés ici, aucune prise de contraste n'est rapportée. Pour le premier (C1P0042), la courbe de diamètre moyen (en bleu) traduit un volume évoluant légèrement, de 40 mm à 45 mm en 18 mois, puis une évolution stable sous traitement de chimiothérapie. Le score de probabilité quant à lui augmente avant le traitement, de 0,35 à 0,7, puis diminue globalement au cours du temps de traitement, cependant certaines dates montrent une probabilité plus importante, créant une courbe très fluctuante. A contrario, le profil du deuxième (C1P0155) montre une courbe de DMT et de score de probabilité qui évolue très peu, avec les mêmes tendances.

Le troisième (C1P0025) et le quatrième (C1P0157) suivi montrent tous les deux l'apparition d'une prise de contraste après la fin du traitement de chimiothérapie. Pour C1P0025, la courbe de DMT évolue au cours du temps et se stabilise sous chimiothérapie, jusqu'à la date de prise de contraste en sortie de traitement. La courbe de probabilité diminue d'abord drastiquement puis un pic apparaît juste avant la chimiothérapie, à 0,8 ; le traitement n'influe pas sur son ascension qui continue le long du traitement jusqu'à la dernière date du suivi, celle de la transformation, avec une valeur de plus de 0,9. Le suivi de C1P0157 est similaire : avant le traitement, les deux courbes augmentent, mais pendant la chimiothérapie la courbe de probabilité de transformation baisse drastiquement, de 0,8 à 0,3, pour revenir à une valeur similaire à celle constatée lors de l'examen pré-chimiothérapie ; elle continue d'augmenter pour les deux dates suivantes, jusqu'à plus de 0,9.

Pour les deux suivis suivants, la chimiothérapie intervient après la date de transformation anaplasique. Le suivi du sujet C1P0124 montre un pic d'évolution de volume, de 13mm à 20mm en 6 mois, et un important pic de score de probabilité (de 0,05 à 0,59). A l'examen de la prise de contraste, à partir de l'examen post opératoire et jusqu'à la fin du suivi, les deux paramètres augmentent continuellement, sans effet du traitement. Pour C1P0050, une évolution lente mais continuelle du volume a lieu au cours du temps malgré une transformation et une période de traitement (de 28mm à 50mm en 53mois). Une augmentation importante du score de probabilité est observée à la date de transformation anaplasique, le score passe de 0.15 à 0.85 entre deux dates espacées de 16 mois, puis ce score diminue au cours de la chimiothérapie.

Les deux derniers suivis, C1P0036 et C1P0074, ont des profils très différents. Le premier montre une évolution volumique plutôt stable et de hautes valeurs de probabilité (constamment au-

dessus de 0,5), même aux premières dates du suivi en post opératoire. À l'inverse, l'évolution volumique du second fluctue beaucoup, d'autant plus avec une intervention chirurgicale au milieu du suivi. Néanmoins, la courbe du score de probabilité reste très basse (constamment au-dessous de 0,25).

Globalement, et ce malgré certaines variations, on observe une tendance commune pour les courbes d'évolution de diamètre et pour celles des scores de probabilité au cours du temps. La corrélation entre le volume (cm^3) et le score de probabilité a été confirmée par le calcul du coefficient de corrélation de Pearson. Le test a montré une corrélation très significative, p-value inférieure à 0.0001 avec un coefficient de corrélation de 0,6 (*Figure 33*).

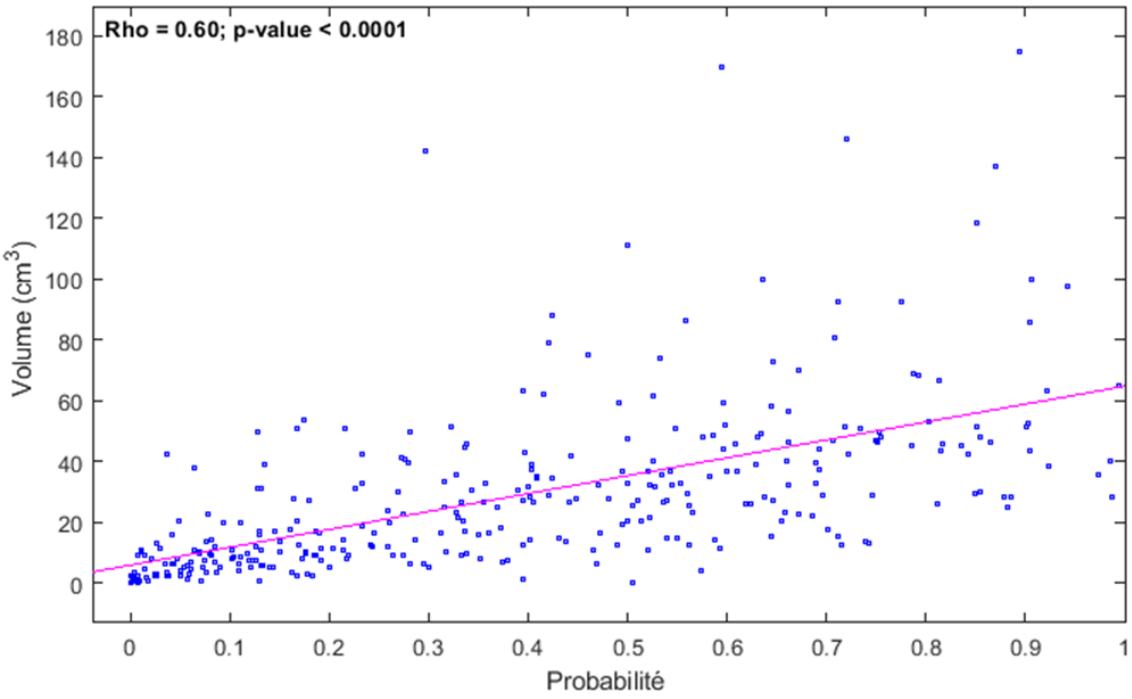


Figure 33 : Résultats du calcul du coefficient de Pearson : le score de probabilité est corrélé au volume (p-value < 0.0001, coefficient de Pearson (rho) = 0.6).

(3) Discussion

Du fait de la grande variabilité des temporalités et des interventions, l'analyse des profils ne peut être réalisée via des outils statistiques. L'inspection visuelle de ces courbes montre de façon marquante des évolutions similaires du score de probabilité et du volume, quelquefois de manière plus précoce pour le score de probabilité que pour celui du volume. Le score de probabilité apparaît

corrélé au volume. Comme nous l'avons évoqué en introduction de cette partie, les métriques ayant permis de générer cette probabilité ont été choisies pour être indépendantes du volume, ce qui n'empêche pas de trouver des associations, bien logiques puisque toutes deux décrivent la même entité : la tumeur. Les GHG ont tendance à être de plus gros volume [47,169,170], or plus la probabilité est proche de 1, plus l'algorithme classe le sujet en haut grade. De plus, il paraîtrait cohérent que les évolutions elles-mêmes soient corrélées puisqu'un volume qui augmente fortement (VED élevée) est souvent rapporté comme ayant une évolution péjorative [2], ce que nous espérons déceler également avec la radiomique. En pratique clinique courante, c'est le volume et son évolution (VED) qui guident la grande majorité des décisions thérapeutiques ; mais parfois, comme on peut le voir sur les profils d'évolutions (*Figure 32*), le volume augmente drastiquement entre deux dates d'examen sans signes avant-coureurs.

Nous avons réalisé une analyse complémentaire afin d'évaluer la corrélation entre la VED (entre deux examens en mm/an) et la variation de probabilité (rapportée par an) qui confirme une association (coefficient de corrélation de Pearson, rho = 0.38, p-value < 0.0001).

Nous remarquons également que la valeur seuil que nous avions utilisée précédemment (0.5) pour la classification GHG / GBG (*IV.2.a) Prédiction du statut GHG – GBG*) ne suffit effectivement pas pour décrire la variation de volume observée. En effet, les variations de volume ne sont pas toujours représentées par le score de probabilité, ou encore, à l'inverse, le seuil est dépassé pour des dates à la variation de volume faible et où le statut GBG est validé par l'analyse anatomopathologique post-chirurgie (*Figure 32 C1P0074, C1P0036, Annexe 2*).

c) Prédiction de la VED

Nous l'évoquions précédemment, la VED est un critère important d'évolution de la tumeur. En moyenne un GBG évolue à 4mm/an[47] de VED, et une évolution de 8mm/an ou plus est de mauvais pronostic[2].

Dans une logique de prédiction d'évolution du GBG, il semble pertinent d'évaluer la capacité de prédiction de la VED à partir du score de probabilité calculé sur un unique examen.

Est-il possible, pour un nouveau sujet, sans antériorité IRM, de prédire sans se tromper un seuil VED du GBG ? Par exemple, une probabilité supérieure à 0.7 est-elle systématiquement associée à des VED supérieures à 6 mm ?

(1) Méthodologie

Dans une visée davantage clinique, nous faisons le choix d'essayer de prédire avec certitude la valeur minimale de VED du GBG. Pour ce faire, nous utiliserons une approche itérative avec estimation des seuils optimaux. On définit ainsi un algorithme d'optimisation :

$$
\begin{aligned}
&pour\ seuil_{VED}\ allant\ de\ 0\ à\ VED_{max}\\
&\quad groupe_{VED} = VED > seuil_{VED}\\
&\quad pour\ seuil_{proba}\ allant\ de\ 0\ à\ 1\\
&\quad\quad groupe_{proba} = probabilite > seuil_{proba}\\
&\quad\quad [VP, VN, FP, FN] = estimationClassification\left(groupe_{VED}, groupe_{proba}\right)\\
&\quad fin\ pour\\
&fin\ pour\\
&combinaisonOptimale = \left(VP == \max\left(VP(FP == 0)\right)\right)\ \&\ (FP == 0)\\
&combinaisonOptimale = combinaisonOptimale\ \&\ (FN =\\
&\quad\quad = \min\left(FN(combinaisonOptimale)\right)
\end{aligned}
$$

Figure 34: Algorithme d'identification du seuil de probabilité permettant de prédire sans erreur une VED. VED_{max} représente la valeur maximale de VED observée dans la base de données. $estimationClassification$ est une fonction qui compare deux groupes pour estimer les vrais positifs (VP), faux positifs (FP), vrais négatifs (VN), faux négatifs (FN) pour un seuil de probabilité et un seuil de VED donné. $combinaisonOptimale$ représente la matrice binaire optimale présentant le maximum de vrais positifs (VP), aucun faux positif (FP), et le moins de faux négatifs (FN).

Afin de tester le seuil trouvé, nous avons appliqué notre méthode de prédiction sur des données issues d'une thèse de médecine du service de neuroradiologie (Philippe LAM, *Étude de l'effet de témozolomide sur la connectivité fonctionnelle chez des patients opérés pour lésion gliale de bas grade*). Cette étude porte sur des patient·e·s GBG avant chimiothérapie et 3 mois après le début de la chimiothérapie. Sur un échantillon de 35 sujets, soit 70 IRM, les données de T1, T2 FLAIR, et de perfusion ont été récupérées, la tumeur a été segmentée automatiquement via notre méthode nnU-Net, puis corrigée manuellement par un neuroradiologue (P.L). Les données ont ensuite été prétraitées automatiquement via les chaînes de traitement décrites dans ce travail de thèse, afin d'obtenir le score de probabilité de chaque examen. Pour les sujets présentant des probabilités supérieures au seuil défini, la VED sera estimée en récupérant les examens antérieurs. L'objectif ici est de confirmer que les probabilités supérieures au seuil identifié correspondent bien à la VED cible.

(2) Résultats

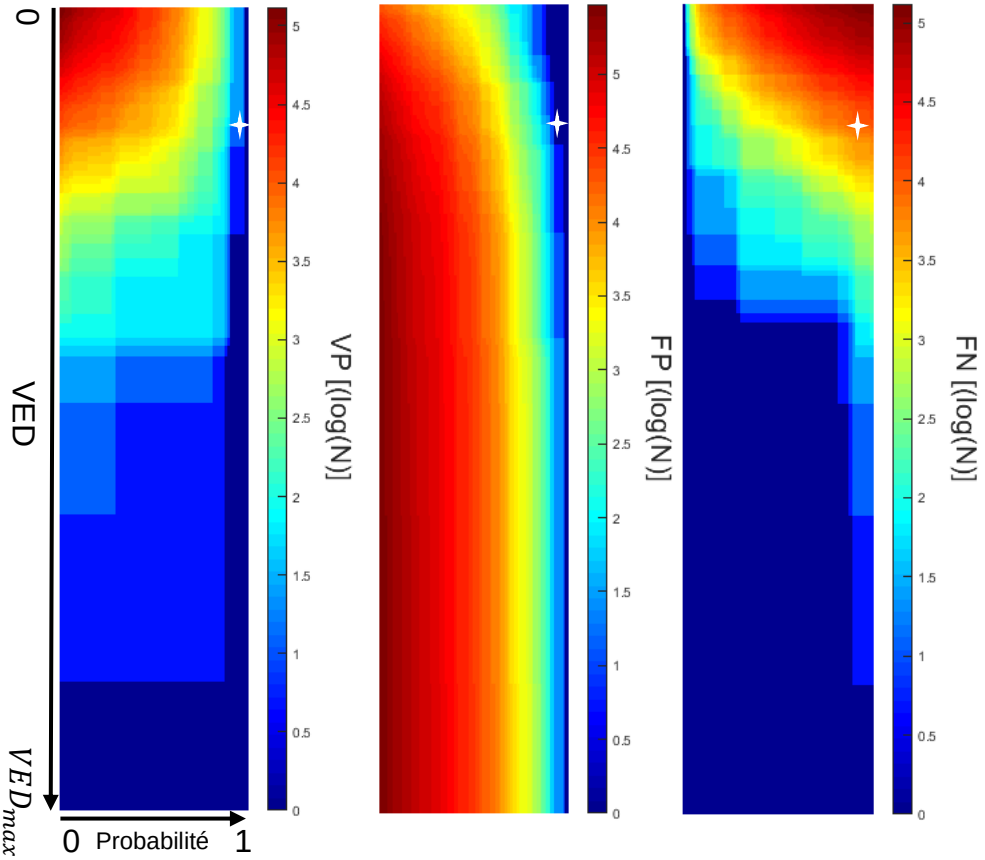


Figure 35: Répartition des vrais positifs (VP), faux positifs (FP) et faux négatifs (FN) en fonction des seuils de VED et de probabilités choisis. Par souci de clarté, le nombre à chaque position a été log-transformé. La combinaison optimale est renseignée par une étoile blanche dans chaque illustration.

Un seuil optimal respectant les critères d'absence de faux positifs, le maximum de vrais positifs et le minimum de faux négatifs (*Figure 35*, étoile blanche) a pu être identifié pour un score de probabilité supérieur à 0.93. Au-dessus de cette probabilité, tous les examens présentent une VED supérieure ou égale à 6.2 mm/an. En revanche, seuls 6.8% des sujets présentant une VED supérieure ou égale à ce seuil sont identifiés comme étant des GHG.

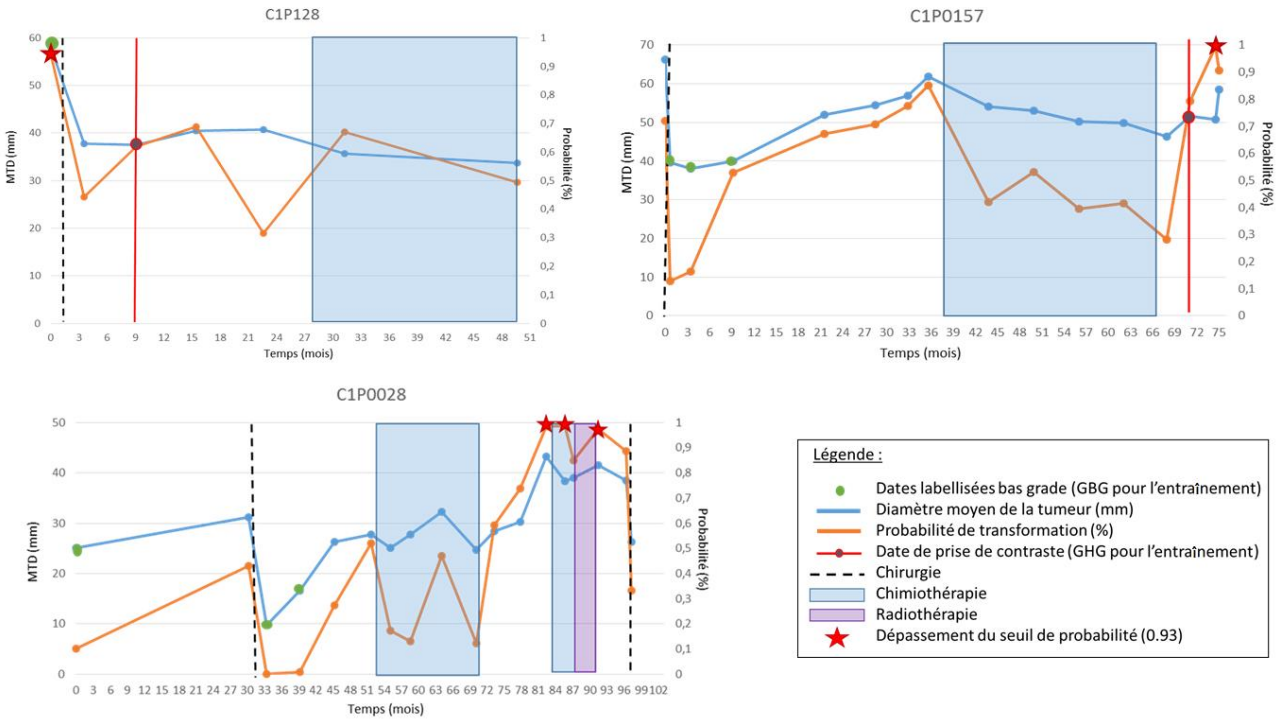


Figure 36: Profils au cours desquels le seuil de prédiction de la VED est dépassé (étoile rouge).

Pour 5 examens appartenant au profil de 3 sujets différents, le seuil prédictif de la VED a été dépassé *(Figure 36, étoile rouge*). Pour le sujet C1P128, le seuil de 0,93 est dépassé dès le premier examen du suivi, au moment de la prise en charge. L'examen IRM à 3 mois post-chirurgie n'identifie pas de transformation anaplasique, mais celui à 9 mois comporte une prise de contraste. Pour C1P0157 décrit précedemment (*Figure 32*), le seuil est dépassé après l'identification de la prise de contraste. Enfin, pour C1P0028, le dépassement du seuil a lieu après une période de 10 mois d'évolution volumique consécutive et après presque 7 ans de suivi (80 mois). Les deux premières dates dépassant le seuil interviennent juste avant le début et à la fin d'un cycle de chimiothérapie. Le troisième point qui dépasse le seuil sur ce suivi arrive après le traitement par radiothérapie effectué à la suite de la chimiothérapie.

Les probabilités obtenues sur les sujets avant et pendant la chimiothérapie sont toutes inférieures au seuil optimal de 0.93 (*Tableau 6*).

Probabilité pré-chimiothérapie	Probabilité 3 mois	Variation de volume (%)
0,24	0,06	1,38
0,73	0,11	-19,59
0,24	0,09	-41,06
0,74	0,59	43,95
0,23	0,27	-19,65
0,81	0,83	12,83
0,72	0,7	24,6
0,44	0,18	-0,83
0,66	0,25	-20,23
0,34	0,21	-30,77
0,12	0,05	-58,84
0,73	0,42	-63,68
0,14	0,15	-17,13
0,49	0,03	-86,99
0,04	0,04	12,01
0,23	0,14	-15,91
0,35	0,1	14,81
0,18	0,17	-75,61
0,23	0,08	-41,88
0,71	0,42	-39,67
0,56	**0,89**	108,95
0,73	0,77	10,97
0,33	0,36	14,79
0,37	0,39	7,45
0,26	0,2	-28,45
0,13	0,61	44,61
0,17	0,3	49,45
0,1	0,13	17,42
0,06	0,09	19,42
0,45	0,2	-10,13
0,52	0,15	9,68
0,2	0,24	4,88
0,08	0,06	-14,93
0,38	0,26	-4,3
0,07	0,08	23,38

Tableau 6 : Probabilité estimée pré-chimiothérapie et 3 mois après le début de la chimiothérapie. La variation de volume mesurée entre ces dates est rapportée dans la colonne de droite en %. Aucun examen ne présente une probabilité supérieure au seuil de 0,93. Le sujet surligné en jaune présente une probabilité proche (0,89) du seuil identifié.

(3) Discussion

Un seuil de probabilité a pu être trouvé (0,93) garantissant une parfaite prédiction d'une VED ≥ 6.2 mm/an. De façon intéressante, la VED identifiée a été rapportée comme étant un indicateur d'évolution péjorative. Cette valeur est en effet au-dessus du seuil de croissance d'un GBG[47]. En revanche, bien que tous les examens au-dessus de ce seuil correspondent à une VED ≥ 6.2 mm/an, ces examens ne représentent que 6.8% des VED ≥ 6.2 mm/an. Ce résultat reflète les contraintes que nous avons imposées à notre algorithme (aucun faux positif). Nous aurions pu faire le choix d'autoriser quelques erreurs afin d'améliorer la proportion de VED ≥ 6.2 mm/an identifiée, mais nous désirions avoir une pertinence clinique claire.

Aucun·e des patient·e·s issu·e·s des nouvelles données de GBG, inclus·es dans le projet de thèse de médecine de Philippe Lam, ne présentait une probabilité supérieure à 0,93, ce qui ne nous a pas permis d'identifier de potentielles VED élevées compte tenu du critère choisi. Comme nous venons de l'évoquer il est quasi-certain, notamment au vu des évolutions de volumes entre les deux IRM, que plusieurs sujets présentent des VED ≥ 6,2 mm/an. Il est intéressant de noter l'observation d'un sujet (*Tableau 6, jaune*) qui présente une probabilité proche du seuil (0.89) avec un volume qui double en trois mois (+108,95%). Cela n'invalide pas pour autant notre seuil, puisque nous avons vu que seuls 6.8% de ces VED sont correctement identifiées vis-à-vis du groupe d'appartenance GBG-GHG. Pour invalider notre méthode, il aurait fallu qu'un sujet présente une probabilité supérieure à 0,93 avec une VED inférieure à 6,2 mm/an, ce qui n'a pas été le cas.

L'intérêt de cette approche est de pouvoir, pour un·e patient·e sans aucune imagerie antérieure disponible, et donc sans aucune VED accessible, statuer quant à son profil d'évolution sur la base d'une seule acquisition. Au sein de notre base de données, de multiples examens, notamment lors des premières dates préopératoires, n'ont pas de VED associée. Parmi ces examens on peut remarquer un sujet (C1P0128) qui, pour sa première date préopératoire, présente un score de 0,94. Ce qui signifierait une forte probabilité de transformation en GHG. Son antériorité n'étant pas accessible il est impossible de confirmer sa VED. Nous avons cependant accès à son profil d'évolution (*Figure 36*, C1P0128). C'est un sujet qui a été opéré dans les jours qui ont suivi son diagnostic, qui n'a pas été identifié comme étant un GHG en post-opératoire et qui a malgré tout présenté une prise de contraste 9 mois après. Cette probabilité mise en évidence au premier examen, concordante avec une VED ≥ 6,2mm/an, est donc confirmée par l'évolution péjorative qui a suivi.

L'outil de prédiction a appris à identifier GHG et GBG, il n'a en revanche pas explicitement appris à identifier les patient·e·s sur la base d'un seuil de VED, c'est-à-dire à prédire les VED péjoratives. Pour les examens dont nous disposons dans notre base de données, les statuts de transformation ont été définis sur la base soit de la prise de contraste, soit d'une confirmation anatomo-pathologique du grade si elle est disponible. Pour finir, nous n'avons, sur les examens de suivi, que des informations parcellaires et aucune confirmation claire du statut de la tumeur. Voilà pourquoi il semblerait pertinent, dans de futurs travaux, de faire apprendre explicitement à l'algorithme, non pas de classer par rapport au statut GHG – GBG, mais par rapport à un seuil de VED, par exemple le seuil de 8mm/an qui est représentatif d'une évolution péjorative[2]. La VED a pour avantage d'être une information accessible pour la plupart des suivis (hors chirurgie) avec une réelle pertinence clinique.

V. Conclusion

Ce travail de thèse a permis de mettre en place une chaîne de traitement complète incluant une segmentation efficace et automatique de la tumeur, sa caractérisation à partir des données d'imageries multimodales, et enfin la prédiction de son statut GHG ou GBG sur la base d'un unique examen IRM.

L'algorithme de segmentation (nnU-Net) a été entraîné exclusivement sur des examens IRM de GBG caractéristiques de leur suivi clinique (chirurgies, chimiothérapies, différentes machines d'acquisition…). Cette base de données d'entraînement a permis à la segmentation automatique d'avoir une bonne performance sur de nouvelles données avec un coefficient de similarité de Dice moyen de 0.82 sur 6 suivis longitudinaux, et de 0.93 sur 49 sujets indépendants.

Le point crucial reste la collecte de ce type de données : que ce soit pour la segmentation tumorale ou pour le modèle SVM, toutes les méthodes dérivées de l'intelligence artificielle ont besoin de se ré-entraîner sur un nombre de données important. Pour améliorer la qualité de la segmentation automatique, il va par exemple être important d'apprendre avec davantage de données IRM différentes (notamment pour un mode d'acquisition 3D) et de façon récurrente. Les séquences IRM évoluant, l'outil risque de perdre en efficacité s'il n'est pas entraîné à les segmenter.

La justification souvent évoquée par les médecins pour ne pas segmenter le volume tumoral, et ce malgré sa pertinence reconnue face aux autres métriques[2,6,72], est l'aspect chronophage de ce travail lorsqu'il est réalisé manuellement. La mise en place de notre outil de segmentation devrait,

nous l'espérons, les faire changer d'avis. Comme nous l'avons montré, il est efficace et ne nécessite que quelques minimes corrections le plus souvent rapides à effectuer.

La chaîne de traitement développée dans ce travail de thèse a permis d'extraire les paramètres radiomiques des masques tumoraux, analysables sur un ensemble d'examen IRM obtenu dans plusieurs centres de radiologie, machines IRM et paramètres d'acquisition, afin de les caractériser et de différencier les paramètres témoins d'une transformation. A partir de ces paramètres radiomiques, un score de prédiction de la transformation en GHG a été généré pour chaque date d'examen IRM de la base de données. Ce score a montré une réelle pertinence en évoluant de façon similaire au volume, bien que basé sur des métriques radiomiques complètement indépendantes. Nous avons d'ailleurs pu établir un seuil pour ce score de probabilité de transformation correspondant à une VED $\geq$ 6.2 mm/an, sans avoir besoin d'accéder à l'antériorité de l'évolution du volume tumoral du sujet.

Nos analyses avaient pour but initial de prédire de façon précoce la transformation anaplasique. Force est de constater que cet objectif est encore trop ambitieux et méritera d'être davantage étudié. En effet, les résultats que nous avons obtenus pour la classification des suivis d'IRM de GBG montrent un certain nombre d'erreurs. Bien qu'une partie de ces erreurs puissent être imputées à l'algorithme, ce dernier a appris sur des données que nous lui avons fournies. Les examens de référence que nous avons utilisés pour les GBG, avec pour seuls critères le fait d'être en début de suivi et aussi le statut de bas grade validé par l'examen anatomopathologique post chirurgie, peuvent être source de biais et avoir faussement augmenté les erreurs de l'algorithme. Nous le savons, nous n'avons pas systématiquement accès à la réalité du statut des gliomes. Les données d'histopathologie, réelle méthode de référence, ne sont pas disponibles pour tous les suivis. L'illustration en est l'examen préopératoire identifié par notre algorithme comme ayant une VED élevée (*Figure 36*) mais qui a malgré tout été considéré comme bas grade et utilisé dans l'entraînement de notre algorithme. De plus, les histoires évolutives des patient·e·s peuvent être différentes (nombre de chirurgies, traitements…), mais surtout les GBG peuvent avoir des types cellulaires différents. Il a été décrit que certains paramètres radiomiques peuvent être impactés par les différents types cellulaires[124–126].

Ce travail met en évidence le besoin de simplifier au maximum les données de référence. Il sera donc intéressant de se focaliser en premier lieu sur une métrique accessible et validée qui représente un indicateur d'évolution péjorative, la VED. Nous pourrions tout à fait ré-entraîner

notre algorithme afin de, non plus prédire la transformation, mais les VED supérieures à 8 mm/an ou 6 mm/an comme mis en évidence par notre seuil. Le SVM pourrait alors apprendre uniquement sur une base de données de gliomes de bas grade pour lequel un suivi IRM est disponible, limitant ainsi le besoin d'avoir recours à des données externes de GHG.

Malgré tout, notre score de prédiction basé sur les paramètres de radiomique évolue globalement dans le même sens que le volume, en apportant une information supplémentaire qui ne nécessite pas d'antériorité. Les événements biologiques qui font évoluer le volume influent donc aussi sur des informations intrinsèques aux signaux des imageries IRM anatomiques et perfusionnelles, captés par les paramètres radiomiques. On peut alors suspecter que la variation de VED et la variation de probabilité sont corrélées. Sont-elles plus fines pour prédire une évolution péjorative ? Des études supplémentaires devront être menées pour répondre à ces interrogations.

Avec l'automatisation de tous les outils et les tests réalisés pour ce travail de thèse (de la segmentation à la prédiction), le temps de segmentation et de traitement de nouvelles données va être réduit drastiquement et facilitera la réalisation de nouvelles études.

A notre connaissance, aucune étude n'a exploré la réalisation d'une prédiction de la vitesse d'expansion du GBG sur la base d'une étude radiomique d'imagerie multimodale longitudinale. C'est un sujet innovant avec un objectif concret proche des objectifs de l'évaluation clinique en radiologie et un sujet d'intérêt pour le neurochirurgien et l'anatomopathologiste. Ce premier travail permet d'envisager une évolution des développements algorithmiques incluant de nombreux paramètres issus de l'exploitation des données IRM pour la discrimination fine de l'évolution des gliomes de bas grade.

Références

1. Louis, D. N. *et al.* The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-Oncol.* **23**, 1231–1251 (2021).
2. Pallud, J. *et al.* Velocity of tumor spontaneous expansion predicts long-term outcomes for diffuse low-grade gliomas. *Neuro-Oncol.* **15**, 595–606 (2013).
3. Duffau, H. & Mandonnet, E. The 'onco-functional balance' in surgery for diffuse low-grade glioma: integrating the extent of resection with quality of life. *Acta Neurochir. (Wien)* **155**, 951–957 (2013).
4. Pallud, J. *et al.* Prognostic value of initial magnetic resonance imaging growth rates for World Health Organization grade II gliomas. *Ann. Neurol.* **60**, 380–383 (2006).
5. Taillandier, L. & Bauchet, L. Référentiel de l'ANOCEF pour les gliomes de l'adulte. (2012).
6. Gui, C., Lau, J. C., Kosteniuk, S. E., Lee, D. H. & Megyesi, J. F. Radiology reporting of low-grade glioma growth underestimates tumor expansion. *Acta Neurochir. (Wien)* **161**, 569–576 (2019).
7. Radbruch, A. *et al.* Relevance of T2 signal changes in the assessment of progression of glioblastoma according to the Response Assessment in Neurooncology criteria. *Neuro-Oncol.* **14**, 222–229 (2012).
8. Law, M. *et al.* Gliomas: predicting time to progression or survival with cerebral blood volume measurements at dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. *Radiology* **247**, 490–498 (2008).
9. Hussain, S., Anwar, S. M. & Majid, M. Segmentation of glioma tumors in brain using deep convolutional neural network. *Neurocomputing* **282**, 248–261 (2018).
10. Naser, M. A. & Deen, M. J. Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. *Comput. Biol. Med.* **121**, 103758 (2020).

11. Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017).
12. Li, M. *et al.* T2/FLAIR Abnormity Could be the Sign of Glioblastoma Dissemination. *Front. Neurol.* **13**, (2022).
13. Lambin, P. *et al.* Radiomics: extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer Oxf. Engl. 1990* **48**, 441–446 (2012).
14. Kim, J. Y. *et al.* Incorporating diffusion- and perfusion-weighted MRI into a radiomics model improves diagnostic performance for pseudoprogression in glioblastoma patients. *Neuro-Oncol.* **21**, 404–414 (2019).
15. Lambin, P. *et al.* Radiomics: the bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **14**, 749–762 (2017).
16. Suárez-García, J. G., Hernández-López, J. M., Moreno-Barbosa, E. & Celis-Alonso, B. de. A simple model for glioma grading based on texture analysis applied to conventional brain MRI. *PLOS ONE* **15**, e0228972 (2020).
17. Ostrom, Q. T. *et al.* CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015. *Neuro-Oncol.* **20**, iv1–iv86 (2018).
18. Soffietti, R. *et al.* Guidelines on management of low-grade gliomas: report of an EFNS-EANO Task Force. *Eur. J. Neurol.* **17**, 1124–1133 (2010).
19. Darlix, A. *et al.* Epidemiology for primary brain tumors: a nationwide population-based study. *J. Neurooncol.* **131**, 525–546 (2017).
20. Duffau, H. Diffuse low-grade gliomas and neuroplasticity. *Diagn. Interv. Imaging* **95**, 945–955 (2014).

21. Mandonnet, E., Capelle, L. & Duffau, H. Extension of paralimbic low grade gliomas: toward an anatomical classification based on white matter invasion patterns. *J. Neurooncol.* **78**, 179–185 (2006).

22. Pignatti, F. *et al.* Prognostic factors for survival in adult patients with cerebral low-grade glioma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **20**, 2076–2084 (2002).

23. Duffau, H. Higher-Order Surgical Questions for Diffuse Low-Grade Gliomas: Supramaximal Resection, Neuroplasticity, and Screening. *Neurosurg. Clin. N. Am.* **30**, 119–128 (2019).

24. Smits, A. & Jakola, A. S. Clinical Presentation, Natural History, and Prognosis of Diffuse Low-Grade Gliomas. *Neurosurg. Clin. N. Am.* **30**, 35–42 (2019).

25. Mandonnet, E., Taillandier, L. & Duffau, H. Prise en charge des gliomes diffus de bas grade cérébraux de découverte fortuite: vers une politique de dépistage dans la population générale? *Bull. Académie Natl. Médecine* **201**, 311–324 (2017).

26. Roelz, R. *et al.* Residual Tumor Volume as Best Outcome Predictor in Low Grade Glioma – A Nine-Years Near-Randomized Survey of Surgery vs. Biopsy. *Sci. Rep.* **6**, 32286 (2016).

27. Mandonnet, E., Taillandier, L. & Duffau, H. Prise en charge des gliomes diffus de bas grade cérébraux de découverte fortuite: vers une politique de dépistage dans la population générale? *Bull. Académie Natl. Médecine* **201**, 311–324 (2017).

28. Duffau, H. Long-term outcomes after supratotal resection of diffuse low-grade gliomas: a consecutive series with 11-year follow-up. *Acta Neurochir. (Wien)* **158**, 51–58 (2016).

29. Jakola, A. S. *et al.* Surgical resection versus watchful waiting in low-grade gliomas. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **28**, 1942–1948 (2017).

30. Capelle, L. *et al.* Spontaneous and therapeutic prognostic factors in adult hemispheric World Health Organization Grade II gliomas: a series of 1097 cases: clinical article. *J. Neurosurg.* **118**, 1157–1168 (2013).

31. Pallud, J. *et al.* Diffuse low-grade oligodendrogliomas extend beyond MRI-defined abnormalities. *Neurology* **74**, 1724–1731 (2010).

32. Jungk, C. *et al.* Prognostic value of the extent of resection in supratentorial WHO grade II astrocytomas stratified for IDH1 mutation status: a single-center volumetric analysis. *J. Neurooncol.* **129**, 319–328 (2016).

33. Ahmadi, R. *et al.* Long-term outcome and survival of surgically treated supratentorial low-grade glioma in adult patients. *Acta Neurochir. (Wien)* **151**, 1359–1365 (2009).

34. Duffau, H. Hodotopy, neuroplasticity and diffuse gliomas. *Neurochirurgie.* **63**, 259–265 (2017).

35. Métais, A. & Rousseau, A. Le diagnostic histo-moléculaire des tumeurs gliales et glioneuronales. *Ann. Pathol.* **41**, 137–153 (2021).

36. Barthel, F. P., Wesseling, P. & Verhaak, R. G. W. Reconstructing the molecular life history of gliomas. *Acta Neuropathol. (Berl.)* **135**, 649–670 (2018).

37. Wesseling, P., van den Bent, M. & Perry, A. Oligodendroglioma: pathology, molecular mechanisms and markers. *Acta Neuropathol. (Berl.)* **129**, 809–827 (2015).

38. Novel, improved grading system(s) for IDH-mutant astrocytic gliomas - PubMed. https://pubmed.ncbi.nlm.nih.gov/29687258/.

39. Appay, R. *et al.* CDKN2A homozygous deletion is a strong adverse prognosis factor in diffuse malignant IDH-mutant gliomas. *Neuro-Oncol.* **21**, 1519–1528 (2019).

40. Osborn, A. G., Louis, D. N., Poussaint, T. Y., Linscott, L. L. & Salzman, K. L. The 2021 World Health Organization Classification of Tumors of the Central Nervous System: What Neuroradiologists Need to Know. *AJNR Am. J. Neuroradiol.* **43**, 928–937 (2022).

41. Mandonnet, E. & Duffau, H. An attempt to conceptualize the individual onco-functional balance: Why a standardized treatment is an illusion for diffuse low-grade glioma patients. *Crit. Rev. Oncol. Hematol.* **122**, 83–91 (2018).

42. Duffau, H. & Taillandier, L. New concepts in the management of diffuse low-grade glioma: Proposal of a multistage and individualized therapeutic approach. *Neuro-Oncol.* **17**, 332–342 (2015).

43. Karim, A. B. M. F. *et al.* Randomized trial on the efficacy of radiotherapy for cerebral low-grade glioma in the adult: European Organization for Research and Treatment of Cancer Study 22845 with the Medical Research Council study BRO4: an interim analysis. *Int. J. Radiat. Oncol. Biol. Phys.* **52**, 316–324 (2002).

44. van den Bent, M. J. *et al.* Long-term efficacy of early versus delayed radiotherapy for low-grade astrocytoma and oligodendroglioma in adults: the EORTC 22845 randomised trial. *Lancet Lond. Engl.* **366**, 985–990 (2005).

45. Douw, L. *et al.* Cognitive and radiological effects of radiotherapy in patients with low-grade glioma: long-term follow-up. *Lancet Neurol.* **8**, 810–818 (2009).

46. Currie, S., Hoggard, N., Craven, I. J., Hadjivassiliou, M. & Wilkinson, I. D. Understanding MRI: basic MR physics for physicians. *Postgrad. Med. J.* **89**, 209–223 (2013).

47. Mandonnet, E. *et al.* Continuous growth of mean tumor diameter in a subset of grade II gliomas. *Ann. Neurol.* **53**, 524–528 (2003).

48. Quarles, C. C., Ward, B. D. & Schmainda, K. M. Improving the reliability of obtaining tumor hemodynamic parameters in the presence of contrast agent extravasation. *Magn. Reson. Med.* **53**, 1307–1316 (2005).

49. Schmainda, K. M. *et al.* Characterization of a First-Pass Gradient-Echo Spin-Echo Method to Predict Brain Tumor Grade and Angiogenesis. *AJNR Am. J. Neuroradiol.* **25**, 1524–1532 (2004).

50. Law, M. *et al.* Comparison of cerebral blood volume and vascular permeability from dynamic susceptibility contrast-enhanced perfusion MR imaging with glioma grade. *AJNR Am. J. Neuroradiol.* **25**, 746–755 (2004).

51. Covarrubias, D. J., Rosen, B. R. & Lev, M. H. Dynamic magnetic resonance perfusion imaging of brain tumors. *The Oncologist* **9**, 528–537 (2004).

52. Paulson, E. S. & Schmainda, K. M. Comparison of dynamic susceptibility-weighted contrast-enhanced MR methods: recommendations for measuring relative cerebral blood volume in brain tumors. *Radiology* **249**, 601–613 (2008).

53. Olea Medical SA. Olea Sphere 3.0.

54. Castellano, A. *et al.* Evaluation of low-grade glioma structural changes after chemotherapy using DTI-based histogram analysis and functional diffusion maps. *Eur. Radiol.* **26**, 1263–1273 (2016).

55. Seow, P. *et al.* Neural Fiber Integrity in High- Versus Low-Grade Glioma using Probabilistic Fiber Tracking. *Acad. Radiol.* **28**, 1721–1732 (2021).

56. El-Serougy, L., Abdel Razek, A. A. K., Ezzat, A., Eldawoody, H. & El-Morsy, A. Assessment of diffusion tensor imaging metrics in differentiating low-grade from high-grade gliomas. *Neuroradiol. J.* **29**, 400–407 (2016).

57. Bradac, O. *et al.* Recognition of anaplastic foci within low-grade gliomas using MR spectroscopy. *Br. J. Neurosurg.* **28**, 631–636 (2014).

58. Andronesi, O. C. *et al.* Detection of 2-Hydroxyglutarate in IDH-Mutated Glioma Patients by In Vivo Spectral-Editing and 2D Correlation Magnetic Resonance Spectroscopy. *Sci. Transl. Med.* **4**, 116ra4-116ra4 (2012).

59. Bourdillon, P. *et al.* Prediction of anaplastic transformation in low-grade oligodendrogliomas based on magnetic resonance spectroscopy and 1p/19q codeletion status. *J. Neurooncol.* **122**, 529–537 (2015).

60. Su, C. *et al.* Amide Proton Transfer Imaging Allows Detection of Glioma Grades and Tumor Proliferation: Comparison with Ki-67 Expression and Proton MR Spectroscopy Imaging. *AJNR Am. J. Neuroradiol.* **38**, 1702–1709 (2017).

61. Gao, W. *et al.* Cho/Cr ratio at MR spectroscopy as a biomarker for cellular proliferation activity and prognosis in glioma: correlation with the expression of minichromosome maintenance protein 2. *Acta Radiol. Stockh. Swed. 1987* **60**, 106–112 (2019).

62. Ando, K. *et al.* [Usefulness of Cho/Cr ratio in proton MR spectroscopy for differentiating residual/recurrent glioma from non-neoplastic lesions]. *Nihon Igaku Hoshasen Gakkai Zasshi Nippon Acta Radiol.* **64**, 121–126 (2004).

63. Bulik, M., Jancalek, R., Vanicek, J., Skoch, A. & Mechl, M. Potential of MR spectroscopy for assessment of glioma grading. *Clin. Neurol. Neurosurg.* **115**, 146–153 (2013).

64. Ekici, S., Risk, B. B., Neill, S. G., Shu, H.-K. & Fleischer, C. C. Characterization of dysregulated glutamine metabolism in human glioma tissue with 1H NMR. *Sci. Rep.* **10**, 20435 (2020).

65. Sahin, N. *et al.* Advanced MR Imaging Techniques in the Evaluation of Nonenhancing Gliomas: Perfusion-Weighted Imaging Compared with Proton Magnetic Resonance Spectroscopy and Tumor Grade. *Neuroradiol. J.* **26**, 531–541.

66. Anselmi, M. *et al.* Diagnostic accuracy of proton magnetic resonance spectroscopy and perfusion-weighted imaging in brain gliomas follow-up: a single institutional experience. *Neuroradiol. J.* **30**, 240–252 (2017).

67. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer Oxf. Engl. 1990* **45**, 228–247 (2009).

68. Galanis, E. *et al.* Validation of neuroradiologic response assessment in gliomas: measurement by RECIST, two-dimensional, computer-assisted tumor area, and computer-assisted tumor volume methods. *Neuro-Oncol.* **8**, 156–165 (2006).

69. Shah, G. D. *et al.* Comparison of linear and volumetric criteria in assessing tumor response in adult high-grade gliomas. *Neuro-Oncol.* **8**, 38–46 (2006).

70. van den Bent, M. J. *et al.* Response assessment in neuro-oncology (a report of the RANO group): assessment of outcome in trials of diffuse low-grade gliomas. *Lancet Oncol.* **12**, 583–593 (2011).

71. Macdonald, D. R., Cascino, T. L., Schold, S. C. & Cairncross, J. G. Response criteria for phase II studies of supratentorial malignant glioma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **8**, 1277–1280 (1990).

72. Pallud, J. *et al.* Quantitative morphological magnetic resonance imaging follow-up of low-grade glioma: a plea for systematic measurement of growth rates. *Neurosurgery* **71**, 729–739; discussion 739-740 (2012).

73. Cha, S. Perfusion MR imaging: basic principles and clinical applications. *Magn. Reson. Imaging Clin. N. Am.* **11**, 403–413 (2003).

74. Sugahara, T. *et al.* Value of dynamic susceptibility contrast magnetic resonance imaging in the evaluation of intracranial tumors. *Top. Magn. Reson. Imaging TMRI* **10**, 114–124 (1999).

75. Law, M. *et al.* Glioma Grading: Sensitivity, Specificity, and Predictive Values of Perfusion MR Imaging and Proton MR Spectroscopic Imaging Compared with Conventional MR Imaging. *AJNR Am. J. Neuroradiol.* **24**, 1989–1998 (2003).

76. Pignatti, F. *et al.* Prognostic factors for survival in adult patients with cerebral low-grade glioma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **20**, 2076–2084 (2002).

77. Capelle, L. *et al.* Spontaneous and therapeutic prognostic factors in adult hemispheric World Health Organization Grade II gliomas: a series of 1097 cases: clinical article. *J. Neurosurg.* **118**, 1157–1168 (2013).

78. Pallud, J. *et al.* Prognostic significance of imaging contrast enhancement for WHO grade II gliomas. *Neuro-Oncol.* **11**, 176–182 (2009).

79. Aum, D. J. *et al.* Molecular and cellular heterogeneity: the hallmark of glioblastoma. *Neurosurg. Focus* **37**, E11 (2014).

80. Rees, J. *et al.* Volumes and growth rates of untreated adult low-grade gliomas indicate risk of early malignant transformation. *Eur. J. Radiol.* **72**, 54–64 (2009).

81. Brasil Caseiras, G. *et al.* Low-grade gliomas: six-month tumor growth predicts patient outcome better than admission tumor volume, relative cerebral blood volume, and apparent diffusion coefficient. *Radiology* **253**, 505–512 (2009).

82. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979).

83. Ashburner, J. & Friston, K. J. Unified segmentation. *NeuroImage* **26**, 839–851 (2005).

84. Badrigilan, S. *et al.* Deep learning approaches for automated classification and segmentation of head and neck cancers and brain tumors in magnetic resonance images: a meta-analysis study. *Int. J. Comput. Assist. Radiol. Surg.* **16**, 529–542 (2021).

85. Waldrop, M. M. News Feature: What are the limits of deep learning? *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1074–1077 (2019).

86. Zhao, Z., Yang, G., Lin, Y., Pang, H. & Wang, M. Automated glioma detection and segmentation using graphical models. *PLOS ONE* **13**, e0200745 (2018).

87. Kamnitsas, K. *et al.* Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **36**, 61–78 (2017).

88. Zhao, Z., Yang, G., Lin, Y., Pang, H. & Wang, M. Automated glioma detection and segmentation using graphical models. *PLoS ONE* **13**, e0200745 (2018).

89. Anwar, S. M., Yousaf, S. & Majid, M. Brain tumor segmentation on Multimodal MRI scans using EMAP Algorithm. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.* **2018**, 550–553 (2018).

90. Li, Z. *et al.* Low-Grade Glioma Segmentation Based on CNN with Fully Connected CRF. *J. Healthc. Eng.* **2017**, 9283480 (2017).

91. Yang, Q., Zhang, H., Xia, J. & Zhang, X. Evaluation of magnetic resonance image segmentation in brain low-grade gliomas using support vector machine and convolutional neural network. *Quant. Imaging Med. Surg.* **11**, 300–316 (2021).

92. Yan, B., Cao, M., Gong, W. & Wei, B. Multi-scale brain tumor segmentation combined with deep supervision. *Int. J. Comput. Assist. Radiol. Surg.* **17**, 561–568 (2022).

93. McKinley, R., Meier, R. & Wiest, R. Ensembles of Densely-Connected CNNs with Label-Uncertainty for Brain Tumor Segmentation. in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (eds. Crimi, A. et al.) 456–465 (Springer International Publishing, 2019). doi:10.1007/978-3-030-11726-9_40.

94. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation | Nature Methods. https://www.nature.com/articles/s41592-020-01008-z.

95. Antonelli, M. *et al.* The Medical Segmentation Decathlon. Preprint at https://doi.org/10.48550/arXiv.2106.05735 (2021).

96. Mehta, R. *et al.* QU-BraTS: MICCAI BraTS 2020 Challenge on Quantifying Uncertainty in Brain Tumor Segmentation - Analysis of Ranking Scores and Benchmarking Results. Preprint at https://doi.org/10.48550/arXiv.2112.10074 (2022).

97. Isensee, F., Jaeger, P. F., Full, P. M., Vollmuth, P. & Maier-Hein, K. H. nnU-Net for Brain Tumor Segmentation. Preprint at https://doi.org/10.48550/arXiv.2011.00848 (2020).

98. Cui, S., Mao, L., Jiang, J., Liu, C. & Xiong, S. Automatic Semantic Segmentation of Brain Gliomas from MRI Images Using a Deep Cascaded Neural Network. *J. Healthc. Eng.* **2018**, 4940593 (2018).

99. Wang, P. & Chung, A. C. S. Relax and focus on brain tumor segmentation. *Med. Image Anal.* **75**, 102259 (2022).

100. Li, Z. *et al.* Low-Grade Glioma Segmentation Based on CNN with Fully Connected CRF. *J. Healthc. Eng.* **2017**, e9283480 (2017).

101. Evaluation of magnetic resonance image segmentation in brain low-grade gliomas using support vector machine and convolutional neural network - Yang - Quantitative Imaging in Medicine and Surgery. https://qims.amegroups.com/article/view/51817/html.

102. MRIcron V1.0.20190902.

103. Pallud, J. *et al.* Velocity of tumor spontaneous expansion predicts long-term outcomes for diffuse low-grade gliomas. *Neuro-Oncol.* **15**, 595–606 (2013).

104. Inano, R. *et al.* Voxel-based clustered imaging by multiparameter diffusion tensor images for glioma grading. *NeuroImage Clin.* **5**, 396–407 (2014).

105. Pope, W. B. *et al.* Relationship between Gene Expression and Enhancement in Glioblastoma Multiforme: Exploratory DNA Microarray Analysis1. *Radiology* **249**, 268–277 (2008).

106. Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563–577 (2016).

107. Bae, S. *et al.* Radiomic MRI Phenotyping of Glioblastoma: Improving Survival Prediction. *Radiology* **289**, 797–806 (2018).

108. Rathore, S. *et al.* Radiomic MRI signature reveals three distinct subtypes of glioblastoma with different clinical and molecular characteristics, offering prognostic value beyond IDH1. *Sci. Rep.* **8**, 5087 (2018).

109. Chato, L. & Latifi, S. Machine Learning and Radiomic Features to Predict Overall Survival Time for Glioblastoma Patients. *J. Pers. Med.* **11**, 1336 (2021).

110. Kim, J. Y. *et al.* Incorporating diffusion- and perfusion-weighted MRI into a radiomics model improves diagnostic performance for pseudoprogression in glioblastoma patients. *Neuro-Oncol.* **21**, 404–414 (2019).

111. Bae, S. *et al.* Radiomic MRI Phenotyping of Glioblastoma: Improving Survival Prediction. *Radiology* **289**, 797–806 (2018).

112. Zucker, S. W. Toward a model of texture. *Comput. Graph. Image Process.* **5**, 190–202 (1976).

113. Sklansky, J. Image Segmentation and Feature Extraction. *IEEE Trans. Syst. Man Cybern.* **8**, 237–247 (1978).

114. Haralick, R. M. Statistical and structural approaches to texture. *Proc. IEEE* **67**, 786–804 (1979).

115. Shboul, Z. A., Reza, S. M. S. & Iftekharuddin, K. M. Quantitative MR Image Analysis for Brian Tumor. *VipIMAGE 2017 Proc. VI ECCOMAS Themat. Conf. Comput. Vis. Med. Image Process. Porto Port. Oct. 18-20 2017 VipIMAGE Conf. 2017 Porto Port.* **27**, 10–18 (2018).

116. Yu, Y. *et al.* Characterizing Brain Tumor Regions Using Texture Analysis in Magnetic Resonance Imaging. *Front. Neurosci.* **15**, 634926 (2021).

117. Ronald Bracewell. *The Fourier Transform And Its Applications Bracewell*. (2000).

118. Qiu, S. & Feichtinger, H. Discrete Gabor structures and optimal representations. *IEEE Trans Signal Process* (1995) doi:10.1109/78.469862.

119. Livens, S., Scheunders, P., Wouwer, G. & Dyck, D. Wavelets for texture analysis, an overview. in 581–585 vol.2 (1997). doi:10.1049/cp:19970958.

120. Zhu, H. *et al.* A new local multiscale Fourier analysis for medical imaging. *Med. Phys.* **30**, 1134–1141 (2003).

121. Nanni, L., Brahnam, S., Ghidoni, S., Menegatti, E. & Barrier, T. Different Approaches for Extracting Information from the Co-Occurrence Matrix. *PLOS ONE* **8**, e83554 (2013).

122. Tang, X. Texture information in run-length matrices. *IEEE Trans. Image Process.* **7**, 1602–1609 (1998).

123. Zwanenburg, A., Leger, S., Vallières, M. & Löck, S. Image biomarker standardisation initiative. *Radiology* **295**, 328–338 (2020).

124. Casale, R., Lavrova, E., Sanduleanu, S., Woodruff, H. C. & Lambin, P. Development and external validation of a non-invasive molecular status predictor of chromosome 1p/19q co-deletion based on MRI radiomics analysis of Low Grade Glioma patients. *Eur. J. Radiol.* **139**, 109678 (2021).

125. Rathore, S. *et al.* Radiomic MRI signature reveals three distinct subtypes of glioblastoma with different clinical and molecular characteristics, offering prognostic value beyond IDH1. *Sci. Rep.* **8**, 5087 (2018).

126. Lam, L. H. T. *et al.* Molecular subtype classification of low-grade gliomas using magnetic resonance imaging-based radiomics and machine learning. *NMR Biomed.* e4792 (2022) doi:10.1002/nbm.4792.

127. Malik, N. *et al.* MRI radiomics to differentiate between low grade glioma and glioblastoma peritumoral region. *J. Neurooncol.* **155**, 181–191 (2021).

128. Zacharaki, E. I. *et al.* Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn. Reson. Med.* **62**, 1609–1618 (2009).

129. Ortiz-Ramón, R., Ruiz-España, S., Mollá-Olmos, E. & Moratal, D. Glioblastomas and brain metastases differentiation following an MRI texture analysis-based radiomics approach. *Phys. Medica PM Int. J. Devoted Appl. Phys. Med. Biol. Off. J. Ital. Assoc. Biomed. Phys. AIFB* **76**, 44–54 (2020).

130. Vamvakas, A. *et al.* Imaging biomarker analysis of advanced multiparametric MRI for glioma grading. *Phys. Medica PM Int. J. Devoted Appl. Phys. Med. Biol. Off. J. Ital. Assoc. Biomed. Phys. AIFB* **60**, 188–198 (2019).

131. The MathWorks Inc. MATLAB 2018a.

132. Statistical Parametric Mapping: The Analysis of Functional Brain Images - 1st Edition. https://www.elsevier.com/books/statistical-parametric-mapping-the-analysis-of-functional-brain-images/penny/978-0-12-372560-8.

133. Olea Medical SA. Olea Sphere 3.0.

134. Ashburner, J. A fast diffeomorphic image registration algorithm. *NeuroImage* **38**, 95–113 (2007).

135. Kwan, R. K., Evans, A. C. & Pike, G. B. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans. Med. Imaging* **18**, 1085–1097 (1999).

136. Collignon, A. *et al.* Automated multi-modality image registration based on information theory. *Bizais* (1995).

137. Tang, T. T., Zawaski, J. A., Francis, K. N., Qutub, A. A. & Gaber, M. W. Image-based Classification of Tumor Type and Growth Rate using Machine Learning: a preclinical study. *Sci. Rep.* **9**, 12529 (2019).

138. Ditmer, A. *et al.* Diagnostic accuracy of MRI texture analysis for grading gliomas. *J. Neurooncol.* **140**, 583–589 (2018).

139. Petrujkić, K. *et al.* Computational quantitative MR image features - a potential useful tool in differentiating glioblastoma from solitary brain metastasis. *Eur. J. Radiol.* **119**, 108634 (2019).

140. Nyúl, L. G., Udupa, J. K. & Zhang, X. New variants of a method of MRI scale standardization. *IEEE Trans. Med. Imaging* **19**, 143–150 (2000).

141. Bedekar, D., Jensen, T. & Schmainda, K. M. Standardization of relative cerebral blood volume (rCBV) image maps for ease of both inter- and intrapatient comparisons. *Magn. Reson. Med.* **64**, 907–913 (2010).

142. Zhang, L. *et al.* IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. *Med. Phys.* **42**, 1341–1353 (2015).

143. Szczypiński, P. M., Strzelecki, M., Materka, A. & Klepaczko, A. MaZda--a software package for image texture analysis. *Comput. Methods Programs Biomed.* **94**, 66–76 (2009).

144. van Griethuysen, J. J. M. *et al.* Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* **77**, e104–e107 (2017).

145. Nioche, C. *et al.* LIFEx: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res.* **78**, 4786–4789 (2018).

146. Court, L. *et al.* Computational resources for radiomics. *Transl. Cancer Res.* **5**, 340–348 (2016).

147. Fornacon-Wood, I. *et al.* Reliability and prognostic value of radiomic features are highly dependent on choice of feature extraction platform. *Eur. Radiol.* **30**, 6241–6250 (2020).

148. van Rossum, G. Python 3.10.4.

149. Lever, J., Krzywinski, M. & Altman, N. Principal component analysis. *Nat. Methods* **14**, 641–642 (2017).

150. Hu, Z. *et al. A Deep Learning Model with Radiomics Analysis Integration for Glioblastoma Post-Resection Survival Prediction*. (2022).

151. Guyon, I., Weston, J., Barnhill, S. & Vapnik, V. Gene Selection for Cancer Classification using Support Vector Machines. *Mach. Learn.* **46**, 389–422 (2002).

152. Osman, A. F. I. A Multi-parametric MRI-Based Radiomics Signature and a Practical ML Model for Stratifying Glioblastoma Patients Based on Survival Toward Precision Oncology. *Front. Comput. Neurosci.* **13**, 58 (2019).

153. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).

154. Cao, X. *et al.* Differentiating solitary brain metastases from glioblastoma by radiomics features derived from MRI and 18F-FDG-PET and the combined application of multiple models. *Sci. Rep.* **12**, 5722 (2022).

155. Berkson, J. Application to the Logistic Function to Bio-Assay. *J. Am. Stat. Assoc.* (1944) doi:10.2307/2280041.

156. Masson, E. Comprendre la régression logistique. *EM-Consulte* https://www.em-consulte.com/article/842576/comprendre-la-regression-logistique.

157. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).

158. Gregorutti, B., Michel, B. & Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **27**, 659–678 (2017).

159. Muller, K.-R., Mika, S., Ratsch, G., Tsuda, K. & Scholkopf, B. An introduction to kernel-based learning algorithms. *IEEE Trans. Neural Netw.* **12**, 181–201 (2001).

160. Nishi, H. *et al.* Predicting Clinical Outcomes of Large Vessel Occlusion Before Mechanical Thrombectomy Using Machine Learning. *Stroke* **50**, 2379–2388 (2019).

161. Jakola, A. S. *et al.* Spatial distribution of malignant transformation in patients with low-grade glioma. *J. Neurooncol.* **146**, 373–380 (2020).

162. Cahill, D. & Turcan, S. Origin of Gliomas. *Semin. Neurol.* **38**, 5–10 (2018).

163. Aydin, S., Fatihoğlu, E., Koşar, P. N. & Ergün, E. Perfusion and permeability MRI in glioma grading. *Egypt. J. Radiol. Nucl. Med.* **51**, 2 (2020).

164. Satar, Z., Hotton, G. & Samandouras, G. Systematic review—Time to malignant transformation in low-grade gliomas: Predicting a catastrophic event with clinical, neuroimaging, and molecular markers. *Neuro-Oncol. Adv.* **3**, vdab101 (2021).

165. Hou, B. L. *et al.* Magnetic Resonance Imaging Parameters and Their Impact on Survival of Patients with Glioblastoma: Tumor Perfusion Predicts Survival. *World Neurosurg.* S1878-8750(18)32908–5 (2018) doi:10.1016/j.wneu.2018.12.085.

166. Zhuge, Y. *et al.* Automated glioma grading on conventional MRI images using deep convolutional neural networks. *Med. Phys.* **47**, 3044–3053 (2020).

167. Caulo, M. *et al.* Data-driven grading of brain gliomas: a multiparametric MR imaging study. *Radiology* **272**, 494–503 (2014).

168. Pallud, J. *et al.* Prognostic significance of imaging contrast enhancement for WHO grade II gliomas. *Neuro-Oncol.* **11**, 176–182 (2009).

169. Altieri, R. *et al.* Natural history of de novo high grade glioma: first description of growth parabola. *J. Neurosurg. Sci.* **64**, 399–403 (2020).

170. Altieri, R. *et al.* Radiological evaluation of ex novo high grade glioma: velocity of diametric expansion and acceleration time study. *Radiol. Oncol.* **55**, 26–34 (2020).

Annexes

Annexe 1 - Tableau des 770 paramètres radiomiques

Tableau des 770 paramètres radiomiques. Pour chaque paramètres, N= nombre de cas par classe (gliome de bas grade (GBG) ou haut grade (GHG)), médiane de la valeur de la variable, (x-y)=intervalle de confiance. Les deux dernières colonnes sont les p-value en uni varié (Wilcoxon) et régression logistique (RL). La méthode de sélection 1 est composée de tous les paramètres à p-value<0,05 (vert et jaune), la méthode 2 sélectionne 5 paramètres par modalité d'imagerie avec une p-value<0,005, la méthode 3 sélectionne les paramètres avec une p-value<0,01 (vert). La méthode 4 prend en compte tous les paramètres de la liste.

Variable	GBG	GHG	p-value (wilcoxon)	p-value (RL)
rBF_original_glrlm_RunEntropy	N=70 6.29 (5.86-6.56)	N=86 6.69 (6.48-7)	0	0
rBF_original_glszm_GrayLevelNonUniformityNormalized	N=70 0.02 (0.01-0.02)	N=86 0.01 (0.01-0.01)	0	1,00E-04
rBV_corrected_original_glrlm_RunEntropy	N=70 6.19 (5.77-6.55)	N=86 6.64 (6.41-6.91)	0	1,00E-04
rBV_corrected_original_glszm_GrayLevelNonUniformityNormalized	N=70 0.02 (0.01-0.02)	N=86 0.01 (0.01-0.01)	0	1,00E-04
FLAIR_original_firstorder_Maximum	N=70 6280.22 (5925.48-6616.41)	N=86 7205.19 (6235.64-9105.39)	0	2,00E-04
FLAIR_original_ngtdm_Complexity	N=70 206956.57 (137638.16-286096.08)	N=86 325302.16 (222081.91-618739.56)	0	2,00E-04
FLAIR_original_firstorder_Range	N=70 5807.76 (5023.44-6550.1)	N=86 6890.14 (5774.76-9031.63)	0	4,00E-04
rBF_original_glcm_SumEntropy	N=70 6.71 (6.09-7.11)	N=86 7.26 (6.97-7.64)	0	4,00E-04
rBF_original_glszm_HighGrayLevelZoneEmphasis	N=70 2490.21 (1466.36-3279.78)	N=86 3727.98 (2798.36-4636.44)	0	4,00E-04
rBF_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 2318.32 (1340.94-2988.78)	N=86 3355.41 (2494.22-4139.98)	0	4,00E-04
rBF_original_firstorder_InterquartileRange	N=70 510.02 (415.41-706.87)	N=86 709.25 (583.98-1022.79)	0	5,00E-04
rBF_original_glcm_JointAverage	N=70 30.94 (25.12-38.23)	N=86 40.43 (32.7-46.86)	0	5,00E-04
rBF_original_glcm_SumAverage	N=70 61.89 (50.24-76.45)	N=86 80.86 (65.4-93.72)	0	5,00E-04
rBV_corrected_original_firstorder_InterquartileRange	N=70 462.82 (384.32-667.51)	N=86 660.06 (527.84-937.63)	0	5,00E-04
rBV_corrected_original_glcm_JointAverage	N=70 30.74 (22.9-36.25)	N=86 38.5 (33.34-46.85)	0	5,00E-04

rBV_corrected_original_glcm_SumAverage	N=70 61.48 (45.79-72.49)	N=86 77.01 (66.67-93.71)	0	5,00E-04
rBF_original_firstorder_MeanAbsoluteDeviation	N=70 472.6 (315.55-542.36)	N=86 561.74 (485.78-705.75)	0	6,00E-04
rBF_original_firstorder_RobustMeanAbsoluteDeviation	N=70 237.78 (182.3-306.68)	N=86 319.9 (262.06-450.25)	0	6,00E-04
rBF_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 1441.73 (738.89-1978.87)	N=86 2087.92 (1434.44-2952.37)	0	6,00E-04
rBV_corrected_original_glszm_HighGrayLevelZoneEmphasis	N=70 2490.7 (1203.82-3280.57)	N=86 3603.91 (2702.36-4442.01)	0	6,00E-04
rBF_original_firstorder_RootMeanSquared	N=70 1055.92 (793.27-1240.24)	N=86 1294.59 (1062.81-1476.23)	0	7,00E-04
rBF_original_glrlm_HighGrayLevelRunEmphasis	N=70 1902.39 (1105.6-2607.03)	N=86 2842.27 (1990.24-3731.85)	0	7,00E-04
rBF_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 1884.98 (1086.51-2579.62)	N=86 2805.07 (1951.25-3697.43)	0	7,00E-04
rBV_corrected_original_firstorder_RobustMeanAbsoluteDeviation	N=70 205.24 (167.02-296.25)	N=86 300.92 (243.04-410.6)	0	8,00E-04
rBV_corrected_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 2233.22 (1061.68-2952.43)	N=86 3268.58 (2443.86-3941.91)	0	8,00E-04
rBF_original_glcm_DifferenceEntropy	N=70 5.03 (4.55-5.31)	N=86 5.38 (5.07-5.64)	0	9,00E-04
rBF_original_gldm_HighGrayLevelEmphasis	N=70 1846.39 (1042.41-2548.79)	N=86 2735.53 (1909.64-3636.15)	0	9,00E-04
rBV_corrected_original_glrlm_HighGrayLevelRunEmphasis	N=70 1832.74 (901.99-2387.43)	N=86 2673.69 (1949.81-3568.55)	0	9,00E-04
rBV_corrected_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 1802.46 (891.83-2358.66)	N=86 2640.08 (1923.92-3530.38)	0	9,00E-04
rBV_corrected_original_glcm_SumEntropy	N=70 6.62 (5.98-7.04)	N=86 7.2 (6.87-7.51)	0	0,001
rBF_original_firstorder_90Percentile	N=70 1552.69 (1114.45-1869.31)	N=86 1993.67 (1599.75-2432.43)	0	0,0011
rBF_original_glcm_Autocorrelation	N=70 1262.59 (782.49-1832.65)	N=86 2092.54 (1493.05-2966.87)	0	0,0011
rBV_corrected_original_firstorder_RootMeanSquared	N=70 966.91 (632.28-1115.27)	N=86 1182.88 (969.55-1369.68)	0	0,0011
rBV_corrected_original_gldm_HighGrayLevelEmphasis	N=70 1780.11 (880.26-2320.54)	N=86 2579.14 (1886.29-3464.15)	0	0,0011
rBF_original_firstorder_Entropy	N=70 6 (5.57-6.36)	N=86 6.41 (6.16-6.8)	0	0,0012
rBV_corrected_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 1432.5 (666.27-1804.47)	N=86 2007.14 (1345.12-2762)	0	0,0012
rBF_original_firstorder_Mean	N=70 785.78 (627.22-946.22)	N=86 1010.91 (813.42-1148.19)	0	0,0013
rBV_corrected_original_firstorder_MeanAbsoluteDeviation	N=70 414.55 (279.87-532.85)	N=86 545.19 (452.22-671.41)	0	0,0013
rBV_corrected_original_firstorder_90Percentile	N=70 1326.24 (910.01-1686.54)	N=86 1800.22 (1420.01-2208.9)	0	0,0014
rBV_corrected_original_firstorder_Mean	N=70 649.51 (451.32-803.57)	N=86 853.22 (696.67-1008.18)	0	0,0015
rBV_corrected_original_glcm_Autocorrelation	N=70 1200.52 (644.87-1768.49)	N=86 1985.53 (1462.67-2830.95)	0	0,0015
rBV_corrected_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 1935.21 (950.36-2589.89)	N=86 2894.01 (2100.27-3808.63)	0	0,0016
rBF_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 1985.82 (1210.1-2747.2)	N=86 3049.6 (2204.41-4058.52)	0	0,0017
rBF_original_ngtdm_Complexity	N=70 165860.39 (45547.67-256739.23)	N=86 258538.8 (219171.12-289620.02)	0	0,0018
rBF_original_glcm_InverseVariance	N=70 0.14 (0.12-0.16)	N=86 0.12 (0.1-0.14)	0,0011	0,0019

rBV_corrected_original_glcm_DifferenceEntropy	N=70 4.86 (4.42-5.28)	N=86 5.31 (4.95-5.63)	0	0,0019
rBF_original_glrlm_GrayLevelVariance	N=70 714.98 (413.53-1018.87)	N=86 1086.52 (828.28-1438.47)	0	0,0022
rBV_corrected_original_firstorder_Median	N=70 455.6 (341.13-556.47)	N=86 603.79 (464.72-772.2)	0	0,0024
MTT_original_glcm_Idn	N=70 0.92 (0.9-0.93)	N=86 0.93 (0.93-0.95)	0	0,0025
rBF_original_firstorder_Variance	N=70 438278.28 (251701.68-623218.95)	N=86 668733.19 (513068.84-894745.39)	0	0,0026
rBF_original_gldm_GrayLevelVariance	N=70 701.37 (403.27-997.65)	N=86 1070.06 (821.63-1433.2)	0	0,0026
MTT_original_gldm_DependenceNonUniformityNormalized	N=70 0.53 (SD =0.11)	N=86 0.44 (SD =0.1)	0	0,0028
rBF_original_glszm_GrayLevelVariance	N=70 896.8 (499.72-1248.67)	N=86 1284.99 (1018.6-1618.65)	0	0,0028
rBV_corrected_original_firstorder_Entropy	N=70 5.92 (5.51-6.28)	N=86 6.39 (6.07-6.67)	0	0,0028
MTT_original_gldm_SmallDependenceEmphasis	N=70 0.75 (0.69-0.81)	N=86 0.69 (0.62-0.73)	0	0,0029
rBF_original_firstorder_Median	N=70 571.28 (475.68-717.8)	N=86 733.65 (578.61-869.99)	0,0001	0,003
rBF_original_glcm_SumSquares	N=70 603.75 (263.58-943.52)	N=86 988.75 (726.16-1388.87)	0	0,003
rBF_original_glcm_ClusterTendency	N=70 1765.51 (776.49-2763.79)	N=86 2910.59 (2315.02-4306.68)	0	0,0032
MTT_original_glszm_ZonePercentage	N=70 0.81 (0.75-0.85)	N=86 0.75 (0.69-0.79)	0	0,0034
rBF_original_glcm_DifferenceAverage	N=70 14.18 (10.3-18.37)	N=86 18.34 (14.47-22.09)	0,0001	0,0036
tMIP_original_firstorder_Skewness	N=70 2.54 (1.79-3.26)	N=86 2.17 (1.59-2.92)	0,0177	0,004
rBV_corrected_original_glcm_DifferenceAverage	N=70 13.92 (SD =6.2)	N=86 17.53 (SD =5.53)	0,0002	0,0041
rBV_corrected_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.56 (0.51-0.61)	N=86 0.59 (0.56-0.63)	0,0035	0,0044
rBV_original_glszm_GrayLevelNonUniformityNormalized	N=70 0.02 (0.02-0.03)	N=86 0.02 (0.01-0.02)	0	0,0046
rBV_corrected_original_glszm_SmallAreaEmphasis	N=70 0.78 (0.74-0.81)	N=86 0.8 (0.77-0.82)	0,0035	0,0047
rBF_original_glszm_SmallAreaEmphasis	N=70 0.78 (0.76-0.81)	N=86 0.8 (0.78-0.82)	0,0096	0,0051
MTT_original_glcm_Id	N=70 0.12 (0.11-0.14)	N=86 0.14 (0.12-0.16)	0,0001	0,0059
rBF_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.57 (0.54-0.62)	N=86 0.6 (0.56-0.63)	0,0102	0,006
rBV_corrected_original_glrlm_GrayLevelVariance	N=70 652.15 (299.74-1030.42)	N=86 1080.21 (788.5-1370.73)	0	0,0061
rBV_original_glrlm_RunEntropy	N=70 5.86 (5.41-6.17)	N=86 6.19 (5.93-6.39)	0	0,0063
rBV_corrected_original_glcm_InverseVariance	N=70 0.15 (SD =0.04)	N=86 0.13 (SD =0.04)	0,0035	0,0064
MTT_original_ngtdm_Contrast	N=70 0.99 (0.67-1.48)	N=86 0.6 (0.43-0.77)	0	0,0069
rBV_corrected_original_firstorder_Variance	N=70 402082.76 (186242.56-636263.32)	N=86 663535.11 (481914.67-850037.14)	0	0,0071
rBV_corrected_original_gldm_GrayLevelVariance	N=70 642.85 (296.52-1018.17)	N=86 1062.93 (771.05-1360.46)	0	0,0071
MTT_original_glszm_SmallAreaEmphasis	N=70 0.88 (SD =0.04)	N=86 0.85 (SD =0.03)	0	0,0073

MTT_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.73 (SD =0.07)	N=86 0.68 (SD =0.06)	0	0,0075
FLAIR_original_glcm_Idn	N=70 0.92 (0.9-0.93)	N=86 0.94 (0.93-0.95)	0	0,0077
rBV_corrected_original_glcm_SumSquares	N=70 550.98 (224.25-966.47)	N=86 952.65 (671.74-1337.54)	0	0,0079
FLAIR_original_glrlm_RunEntropy	N=70 6.94 (6.62-7.07)	N=86 6.96 (6.83-7.19)	0,1259	0,0084
MTT_original_glcm_InverseVariance	N=70 0.06 (0.05-0.07)	N=86 0.07 (0.06-0.09)	0,0001	0,0086
tMIP_original_glcm_SumEntropy	N=70 8.57 (7.8-9.2)	N=86 9.36 (8.6-10.18)	0	0,0086
tMIP_original_glszm_GrayLevelNonUniformityNormalized	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0	0,0087
MTT_original_glcm_Idmn	N=70 0.98 (0.97-0.99)	N=86 0.99 (0.98-0.99)	0	0,0093
MTT_original_glcm_Idm	N=70 0.06 (0.05-0.08)	N=86 0.08 (0.06-0.09)	0,0001	0,0094
rBV_corrected_original_glcm_ClusterTendency	N=70 1532.98 (611.23-2879.59)	N=86 2661.63 (2030.53-4020.51)	0	0,0094
TTP_original_gldm_LargeDependenceEmphasis	N=70 11.22 (5.76-17.58)	N=86 20.06 (11.85-28.91)	0	0,0095
FLAIR_original_glcm_SumEntropy	N=70 7.46 (7.22-7.58)	N=86 7.57 (7.38-7.77)	0,0016	0,0096
TMAX_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 5845.46 (4283.36-9105.96)	N=86 8187.76 (5317.24-13505.54)	0,005	0,0102
MTT_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.97 (0.96-0.98)	N=86 0.96 (0.94-0.96)	0	0,0103
TTP_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 50727.92 (28153.88-91825.83)	N=86 113599.76 (67769.45-175533.78)	0	0,0103
TTP_original_glrlm_RunPercentage	N=70 0.93 (0.91-0.96)	N=86 0.9 (0.87-0.93)	0	0,0104
tMIP_original_glrlm_RunEntropy	N=70 8.04 (7.54-8.56)	N=86 8.59 (8-9.41)	0	0,0106
rBV_corrected_original_ngtdm_Complexity	N=70 159868.3 (27598.7-282189.63)	N=86 265713.65 (215472.89-323259.54)	0	0,0109
FLAIR_original_firstorder_Entropy	N=70 6.82 (6.49-6.97)	N=86 6.78 (6.64-7.07)	0,486	0,011
MTT_original_glcm_DifferenceAverage	N=70 24.91 (21.44-33.17)	N=86 22.5 (18.37-26.34)	0,0018	0,0111
TTP_original_gldm_DependenceVariance	N=70 3.95 (1.86-6.42)	N=86 6.8 (4.61-8.9)	0	0,0118
FLAIR_original_glcm_DifferenceEntropy	N=70 5.69 (5.33-5.89)	N=86 5.61 (5.39-5.91)	0,7989	0,0123
FLAIR_original_firstorder_Skewness	N=70 -0.24 (-0.47-0.06)	N=86 -0.02 (-0.55-0.34)	0,0943	0,0124
rBV_corrected_original_glszm_GrayLevelVariance	N=70 920.04 (368.88-1301.12)	N=86 1313.56 (964.09-1584.79)	0	0,0128
T1_original_glcm_SumEntropy	N=70 7.61 (7.18-7.92)	N=86 7.94 (7.75-8.18)	0	0,0131
TTP_original_glrlm_LongRunEmphasis	N=70 1.29 (1.15-1.47)	N=86 1.54 (1.28-1.75)	0	0,0132
TTP_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.9 (0.85-0.93)	N=86 0.84 (0.8-0.89)	0	0,0134
TTP_original_glrlm_ShortRunEmphasis	N=70 0.96 (0.94-0.97)	N=86 0.93 (0.91-0.96)	0	0,0136
MTT_original_glrlm_ShortRunEmphasis	N=70 0.99 (0.98-0.99)	N=86 0.98 (0.98-0.99)	0	0,014
rBF_original_firstorder_Skewness	N=70 2.28 (1.85-2.71)	N=86 2.04 (1.58-2.46)	0,0293	0,0144

rBV_original_glcm_SumEntropy	N=70 6.25 (5.58-6.58)	N=86 6.6 (6.26-6.92)	0	0,0151
K2_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 15749930.56 (1745138.78-28933507.12)	N=86 32163206.18 (14415281.67-45417300.02)	0,0005	0,0155
rBV_corrected_original_firstorder_Range	N=70 5178.43 (3488.78-5440.95)	N=86 5247.25 (5122.99-5555.39)	0,0084	0,0165
TTP_original_glszm_ZonePercentage	N=70 0.51 (0.36-0.64)	N=86 0.34 (0.24-0.51)	0	0,0165
FLAIR_original_glcm_JointEnergy	N=70 0 (0-0)	N=86 0 (0-0)	0	0,0168
T1_original_glcm_MaximumProbability	N=70 0 (0-0)	N=86 0 (0-0)	0	0,0169
rBF_original_glcm_Contrast	N=70 635.8 (263.56-996.52)	N=86 928.82 (648.73-1214.99)	0,0001	0,0196
T1_original_glrlm_RunEntropy	N=70 6.98 (6.68-7.19)	N=86 7.18 (7.03-7.38)	0	0,0204
T1_original_glcm_ClusterShade	N=70 40042.72 (-10442.65-168238.53)	N=86 158530.12 (39542.32-665181.45)	0,0001	0,0205
TTP_original_gldm_SmallDependenceEmphasis	N=70 0.44 (0.33-0.58)	N=86 0.31 (0.23-0.46)	0,0001	0,0207
FLAIR_original_firstorder_Kurtosis	N=70 3.24 (2.64-3.74)	N=86 3.9 (3.34-5.11)	0	0,021
FLAIR_original_glcm_Idmn	N=70 0.98 (0.97-0.99)	N=86 0.99 (0.99-0.99)	0	0,0215
rBV_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 3029.31 (2084.32-3926.01)	N=86 4510.91 (3408.75-6414.55)	0	0,0217
TTP_original_glrlm_RunVariance	N=70 0.12 (0.06-0.21)	N=86 0.24 (0.12-0.37)	0	0,0217
rBV_corrected_original_firstorder_Maximum	N=70 5076.38 (3282.5-5329.79)	N=86 5134.73 (5000-5438.23)	0,0202	0,022
MTT_original_glcm_Contrast	N=70 1380.63 (936.84-2263.72)	N=86 1142.87 (762.12-1615.31)	0,0254	0,0226
rBV_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 1140.48 (745.7-1495.04)	N=86 1496.21 (1169.28-2096.11)	0,0001	0,0227
rBV_original_firstorder_90Percentile	N=70 905.28 (638.97-1180.98)	N=86 1088.76 (858.68-1381.38)	0,0003	0,0228
TTP_original_glcm_Idm	N=70 0.19 (0.13-0.25)	N=86 0.26 (0.18-0.31)	0,0004	0,0232
TTP_original_gldm_DependenceNonUniformityNormalized	N=70 0.23 (0.16-0.33)	N=86 0.15 (0.12-0.23)	0	0,0232
rBV_original_firstorder_MeanAbsoluteDeviation	N=70 339.35 (SD =139.63)	N=86 428.01 (SD =135.36)	0,0001	0,0234
FLAIR_original_glcm_MaximumProbability	N=70 0 (0-0)	N=86 0 (0-0)	0	0,0236
rBF_original_firstorder_Range	N=70 5000 (4280.11-5001.91)	N=86 5010.05 (5000-5088.26)	0	0,024
rBV_corrected_original_firstorder_Minimum	N=70 -109 (-125.35--84.32)	N=86 -111.3 (-142.57--92.98)	0,0936	0,0245
rBV_original_firstorder_RootMeanSquared	N=70 685.1 (532.62-796.73)	N=86 794.41 (639.43-975.72)	0,0003	0,0246
K2_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 14517357.26 (1531320.91-26561140.64)	N=86 29575860.26 (12835452.23-41438347.66)	0,0012	0,025
K2_original_gldm_HighGrayLevelEmphasis	N=70 15292654.72 (1688052.11-28249673.56)	N=86 31051241.14 (13700685.41-43990502.77)	0,0012	0,0252
K2_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 13255322.24 (1386703.04-25399342.88)	N=86 28138957.29 (11890043.33-38379045.93)	0,0018	0,0253
K2_original_glcm_Autocorrelation	N=70 15196803.66 (1695547.11-	N=86 31120496.84 (13675601.96-	0,001	0,0254

	28028201.36)	43998628.41)		
K2_original_glrlm_HighGrayLevelRunEmphasis	N=70 15348186.62 (1687663.2-28249369.63)	N=86 31067376.68 (13703187.8-44073790.19)	0,0012	0,0256
rBV_corrected_original_glcm_Contrast	N=70 598.62 (235.6-935.3)	N=86 899.78 (601.62-1237.53)	0,0001	0,0257
K2_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 15257507.96 (1674207.3-28087301.53)	N=86 30945178.29 (13634528.49-43847911.66)	0,0013	0,0258
K2_original_glszm_HighGrayLevelZoneEmphasis	N=70 15517965.1 (1682928.62-28246017.74)	N=86 31132850.82 (13720026.16-44294403.16)	0,0011	0,0259
rBF_original_firstorder_Maximum	N=70 5000 (4266.21-5000)	N=86 5000 (5000-5074.75)	0	0,0261
TTP_original_glcm_Id	N=70 0.27 (SD =0.09)	N=86 0.32 (SD =0.09)	0,0008	0,0263
TMAX_original_gldm_DependenceNonUniformityNormalized	N=70 0.29 (0.2-0.39)	N=86 0.2 (0.15-0.29)	0	0,0268
MTT_original_firstorder_RootMeanSquared	N=70 1279.57 (1103.59-1505.35)	N=86 1244 (1135.91-1364.1)	0,1977	0,027
tMIP_original_firstorder_Kurtosis	N=70 12.24 (7.88-17.06)	N=86 9.08 (6.1-15.2)	0,0214	0,0271
rBV_original_glrlm_HighGrayLevelRunEmphasis	N=70 1068.95 (699.51-1406.93)	N=86 1398.3 (1073.54-1981.69)	0,0001	0,0275
TTP_original_glcm_Idn	N=70 0.95 (0.93-0.97)	N=86 0.97 (0.96-0.98)	0	0,0275
rBV_original_firstorder_RobustMeanAbsoluteDeviation	N=70 168.16 (126.93-224.63)	N=86 203.73 (167.01-275.46)	0,0003	0,0279
tMIP_original_firstorder_Entropy	N=70 7.95 (7.46-8.5)	N=86 8.52 (7.89-9.36)	0	0,0279
rBF_original_glcm_DifferenceVariance	N=70 366.91 (140.35-613.69)	N=86 564.54 (405.74-677.64)	0,0002	0,028
rBV_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 1050.93 (689.97-1395.15)	N=86 1381.86 (1049.35-1953.68)	0,0001	0,028
T1_original_firstorder_Entropy	N=70 6.85 (6.58-7.06)	N=86 7.04 (6.9-7.27)	0,0001	0,0284
MTT_original_firstorder_90Percentile	N=70 2007.97 (1720.09-2415.18)	N=86 1870.93 (1726.46-2115.15)	0,1359	0,0292
rBV_original_gldm_HighGrayLevelEmphasis	N=70 1021.8 (665.37-1350.6)	N=86 1330.84 (1037.41-1908.48)	0,0002	0,0292
MTT_original_glcm_SumSquares	N=70 1031.5 (712.52-1704.07)	N=86 1012.5 (698.78-1269.86)	0,1536	0,0293
rBV_original_glszm_HighGrayLevelZoneEmphasis	N=70 1569.34 (1015.24-2056.89)	N=86 2100.08 (1583.53-2777.26)	0	0,0293
rBV_original_glcm_DifferenceEntropy	N=70 4.73 (4.29-5.07)	N=86 5.04 (4.73-5.34)	0,0001	0,0295
rBF_original_gldm_SmallDependenceEmphasis	N=70 0.53 (SD =0.1)	N=86 0.55 (SD =0.1)	0,4181	0,0297
rBV_original_glcm_JointAverage	N=70 21.74 (18.63-25.43)	N=86 25.35 (21.92-29.98)	0,0005	0,0301
rBV_original_glcm_SumAverage	N=70 43.48 (37.26-50.86)	N=86 50.69 (43.85-59.96)	0,0005	0,0301
FLAIR_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 13300.72 (9830.33-18361.52)	N=86 14928.08 (11732.64-17689.53)	0,2143	0,032
rBV_corrected_original_gldm_SmallDependenceEmphasis	N=70 0.51 (0.46-0.57)	N=86 0.55 (0.47-0.6)	0,1685	0,032
rBF_original_glcm_ClusterProminence	N=70 33970956.58 (8849703.62-70183064.4)	N=86 76663001.54 (49017287.27-109612043.65)	0	0,0328
rBV_original_firstorder_InterquartileRange	N=70 371.87 (292.68-498.31)	N=86 454.91 (366.62-602.72)	0,0005	0,034

T1_original_glszm_GrayLevelNonUniformityNormalized	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,0005	0,0344
rBV_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 1454.34 (900.89-1855.41)	N=86 1876.59 (1449.15-2562.25)	0	0,035
TMAX_original_gldm_SmallDependenceEmphasis	N=70 0.52 (SD =0.17)	N=86 0.42 (SD =0.14)	0,0001	0,0355
MTT_original_ngtdm_Strength	N=70 33.27 (14.22-70.29)	N=86 14.98 (10.19-22.74)	0	0,0356
K2_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 24364730.8 (3797116.64-47520342.08)	N=86 52318416.69 (24279639.66-77362653.85)	0,0003	0,0358
rBV_original_firstorder_Mean	N=70 392.8 (278.15-476.62)	N=86 463.53 (366.63-561.99)	0,0011	0,0366
T1_original_glszm_LowGrayLevelZoneEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,0368
T1_original_glcm_JointEnergy	N=70 0 (0-0)	N=86 0 (0-0)	0	0,0369
TMAX_original_gldm_LargeDependenceEmphasis	N=70 9.75 (5.98-14.12)	N=86 14.18 (9.47-21.58)	0	0,037
MTT_original_firstorder_RobustMeanAbsoluteDeviation	N=70 430.02 (340.03-486.16)	N=86 397.07 (342.09-453.54)	0,0657	0,0371
rBV_original_glcm_ClusterTendency	N=70 1157.33 (547.87-1635.27)	N=86 1697.44 (1045.97-2545.47)	0,0001	0,0373
tMIP_original_firstorder_Minimum	N=70 -271.36 (-346.91--210.93)	N=86 -309.68 (-393.52--254.57)	0,0055	0,0373
TTP_original_firstorder_Uniformity	N=70 0.04 (0.02-0.05)	N=86 0.05 (0.03-0.06)	0,0075	0,0373
TMAX_original_glszm_ZonePercentage	N=70 0.57 (SD =0.18)	N=86 0.46 (SD =0.16)	0,0001	0,0377
rBV_original_glcm_SumSquares	N=70 418.16 (189.68-614.02)	N=86 588.13 (359.31-907.64)	0,0001	0,0378
MTT_original_firstorder_MeanAbsoluteDeviation	N=70 633.56 (519.02-763.37)	N=86 597.11 (516.68-696.16)	0,1426	0,038
MTT_original_firstorder_InterquartileRange	N=70 1023.56 (819.63-1158.15)	N=86 939.06 (792.18-1099.31)	0,0684	0,0391
rBV_corrected_original_firstorder_10Percentile	N=70 121.39 (77.21-200.67)	N=86 180.63 (76.01-264.21)	0,0541	0,0391
MTT_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 3763.55 (2812.12-6147.32)	N=86 3438.62 (2592.4-4116.99)	0,1173	0,0397
MTT_original_firstorder_Variance	N=70 711255.73 (465154.44-1088883.5)	N=86 673369.92 (470895-834435.2)	0,2319	0,0402
MTT_original_gldm_GrayLevelVariance	N=70 1138.13 (743.79-1741.85)	N=86 1077.44 (754.38-1335.72)	0,2333	0,0404
rBF_original_gldm_DependenceNonUniformityNormalized	N=70 0.29 (0.25-0.34)	N=86 0.31 (0.26-0.36)	0,4279	0,0404
rBF_original_glszm_ZonePercentage	N=70 0.6 (SD =0.1)	N=86 0.61 (SD =0.1)	0,5735	0,0411
rBV_original_glrlm_GrayLevelVariance	N=70 501.03 (289.74-687.03)	N=86 690.9 (465.9-978.83)	0,0001	0,0411
rBV_original_gldm_GrayLevelVariance	N=70 473.06 (281.2-670.88)	N=86 674.36 (446.27-963.65)	0,0001	0,042
rBV_original_firstorder_Variance	N=70 295497.51 (175674.03-419922.6)	N=86 421216.25 (278969.66-602533.68)	0,0001	0,0423
rBV_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 801.09 (497.09-1062.15)	N=86 1034.44 (705.94-1512.35)	0,0007	0,0428
rBV_corrected_original_glszm_ZonePercentage	N=70 0.58 (0.53-0.64)	N=86 0.61 (0.52-0.66)	0,2347	0,0431
TMAX_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.92 (0.88-0.94)	N=86 0.88 (0.84-0.91)	0	0,0437

FLAIR_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 17137.39 (12258.18-20726.57)	N=86 18646.29 (16084.98-21334.71)	0,0439	0,0438
rBV_corrected_original_firstorder_Skewness	N=70 2.32 (1.9-2.9)	N=86 2.09 (1.72-2.62)	0,1149	0,0439
tMIP_original_glcm_DifferenceEntropy	N=70 6.91 (SD =0.99)	N=86 7.61 (SD =1.1)	0	0,0446
FLAIR_original_glszm_GrayLevelNonUniformityNormalized	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,6066	0,0449
TTP_original_firstorder_Kurtosis	N=70 15.1 (7.15-32.85)	N=86 26.68 (13.34-44.47)	0,0021	0,0449
MTT_original_glcm_ClusterTendency	N=70 2809.39 (1916.86-4253.98)	N=86 2855.36 (1963.41-3601.06)	0,4054	0,045
rBF_original_firstorder_10Percentile	N=70 246.84 (187.87-317.1)	N=86 283.99 (187.27-363.92)	0,0986	0,045
rBV_original_glcm_Autocorrelation	N=70 644.96 (448.79-876.3)	N=86 856.18 (645.86-1237.26)	0,0001	0,0463
T1_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,001	0,048
TTP_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.03 (0.02-0.05)	N=86 0.04 (0.03-0.06)	0,0094	0,0482
T1_original_firstorder_Uniformity	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,001	0,0493
K2_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 23479455.89 (3724286.67-39465590.62)	N=86 43000717.58 (19909756.53-62220153.21)	0,0004	0,0505
TMAX_original_glrlm_ShortRunEmphasis	N=70 0.97 (0.95-0.97)	N=86 0.95 (0.93-0.96)	0	0,0517
rBV_original_firstorder_Entropy	N=70 5.6 (5.13-5.84)	N=86 5.85 (5.51-6.11)	0,0001	0,052
rBV_original_glcm_DifferenceAverage	N=70 11.79 (8.71-14.79)	N=86 14.28 (11-17.71)	0,0025	0,052
MTT_original_gldm_HighGrayLevelEmphasis	N=70 4586.49 (3490.32-7068.21)	N=86 4475.47 (3290.91-5428.54)	0,3006	0,0535
TTP_original_firstorder_Entropy	N=70 5.38 (SD =0.88)	N=86 5.13 (SD =0.74)	0,0601	0,0541
MTT_original_glrlm_GrayLevelVariance	N=70 1141.29 (745.06-1731.11)	N=86 1078.35 (751.2-1334.55)	0,2389	0,0546
MTT_original_glcm_DifferenceVariance	N=70 649.87 (389.93-1012.24)	N=86 581.82 (363.28-824.87)	0,1928	0,0547
TMAX_original_glcm_Idm	N=70 0.14 (0.11-0.2)	N=86 0.2 (0.14-0.25)	0,0005	0,055
rBF_original_firstorder_Kurtosis	N=70 9.29 (6.8-13.89)	N=86 8.07 (5.63-10.56)	0,0214	0,0554
TMAX_original_glcm_Id	N=70 0.21 (0.16-0.26)	N=86 0.26 (0.2-0.32)	0,0005	0,0554
MTT_original_glszm_GrayLevelVariance	N=70 1205.12 (758.27-1776.37)	N=86 1146.1 (798.31-1430.93)	0,3856	0,0555
K2_original_glcm_JointAverage	N=70 3870.2 (1299.49-5288.5)	N=86 5569.77 (3693.77-6600.24)	0,0012	0,0557
K2_original_glcm_SumAverage	N=70 7740.41 (2598.98-10577)	N=86 11139.55 (7387.55-13200.48)	0,0012	0,0557
TTP_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 9486.71 (6879.59-10301.19)	N=86 10606.97 (9416.88-11830.66)	0	0,057
K2_original_firstorder_Minimum	N=70 -98938.83 (-133223.76--32343.43)	N=86 -139914.76 (-168310.81--100337.72)	0,0013	0,0577
rBF_original_glcm_Correlation	N=70 0.48 (0.36-0.54)	N=86 0.51 (0.46-0.59)	0,0021	0,0578
TMAX_original_glrlm_RunPercentage	N=70 0.95 (0.93-0.96)	N=86 0.92 (0.89-0.94)	0,0001	0,0589

MTT_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 4614.76 (3474.85-7078.41)	N=86 4488.8 (3415.43-5522.21)	0,374	0,061
MTT_original_glrlm_RunPercentage	N=70 0.98 (0.98-0.99)	N=86 0.98 (0.97-0.98)	0	0,0615
MTT_original_glrlm_HighGrayLevelRunEmphasis	N=70 4644.36 (3506.96-7115.67)	N=86 4550.56 (3448.89-5588.06)	0,3817	0,0621
rBV_original_glcm_Contrast	N=70 466.93 (194.4-706.62)	N=86 653.17 (401.81-1021.7)	0,0013	0,0622
rBV_corrected_original_glcm_DifferenceVariance	N=70 352.14 (127.61-594.11)	N=86 535.97 (371.54-718.38)	0,0003	0,0638
MTT_original_firstorder_Kurtosis	N=70 5.23 (4.31-6.17)	N=86 5.99 (4.97-8.47)	0,0024	0,0642
TMAX_original_gldm_DependenceVariance	N=70 3.65 (2.05-5.46)	N=86 5.26 (3.67-7.41)	0	0,0643
FLAIR_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,9135	0,0649
MTT_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 4812.23 (3656.42-7351.46)	N=86 4873.21 (3578.85-5856.51)	0,4995	0,0662
rBF_original_glcm_Id	N=70 0.21 (0.19-0.24)	N=86 0.2 (0.17-0.23)	0,003	0,0666
rBV_corrected_original_gldm_DependenceNonUniformityNormalized	N=70 0.29 (0.25-0.33)	N=86 0.3 (0.24-0.35)	0,3895	0,0668
FLAIR_original_ngtdm_Contrast	N=70 0.99 (0.68-1.57)	N=86 0.49 (0.31-0.69)	0	0,0672
FLAIR_original_firstorder_Uniformity	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,8796	0,0682
K2_original_glcm_SumEntropy	N=70 8.81 (7.89-9.39)	N=86 9.45 (9.07-10.03)	0	0,0692
MTT_original_glcm_JointAverage	N=70 60.83 (SD =15.08)	N=86 58.19 (SD =14.98)	0,2771	0,0716
MTT_original_glcm_SumAverage	N=70 121.65 (SD =30.17)	N=86 116.37 (SD =29.95)	0,2771	0,0716
rBV_original_glcm_DifferenceVariance	N=70 312.26 (121.91-457.41)	N=86 395.18 (281.54-628.48)	0,0011	0,0716
MTT_original_firstorder_Mean	N=70 1002.36 (858.1-1126.26)	N=86 925.89 (851.15-1058.64)	0,1207	0,0719
FLAIR_original_glcm_ClusterTendency	N=70 2585.26 (2078.42-3456.45)	N=86 2569.49 (2005.1-3759.12)	0,7256	0,0723
MTT_original_glcm_Autocorrelation	N=70 3779.32 (2712.65-6066.91)	N=86 3723.87 (2691.94-4504.66)	0,4618	0,0726
FLAIR_original_glszm_HighGrayLevelZoneEmphasis	N=70 21138.63 (13828.94-24481.79)	N=86 22916.7 (19445.57-25378.34)	0,0132	0,0729
rBV_original_glszm_SmallAreaEmphasis	N=70 0.76 (0.74-0.8)	N=86 0.78 (0.76-0.79)	0,0458	0,0747
FLAIR_original_glcm_ClusterShade	N=70 -37654.08 (-65880.26-625.28)	N=86 798.12 (-45829.95-81289.5)	0,0006	0,0763
T1_original_glcm_DifferenceEntropy	N=70 5.66 (5.46-5.98)	N=86 5.83 (5.59-6.12)	0,0059	0,0773
rBV_original_firstorder_Range	N=70 5261.66 (4162.43-5521.79)	N=86 5281.82 (5145.01-5566.25)	0,0769	0,0773
T1_original_firstorder_Skewness	N=70 0.34 (-0.17-0.85)	N=86 0.77 (0.37-1.66)	0,0001	0,0774
rBV_original_firstorder_Maximum	N=70 5107.31 (3938.37-5337.46)	N=86 5156.65 (5000-5349.64)	0,0727	0,0783
TTP_original_glcm_InverseVariance	N=70 0.19 (SD =0.08)	N=86 0.22 (SD =0.07)	0,007	0,0786
tMIP_original_glcm_Correlation	N=70 0.44 (0.3-0.54)	N=86 0.49 (0.43-0.55)	0,0032	0,0804
MTT_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 4691.09 (3445.12-7132.18)	N=86 4396.69 (3631.58-5352.78)	0,4706	0,0819

MTT_original_firstorder_Median	N=70 889.03 (744.3-1000.41)	N=86 841.49 (711.57-964.72)	0,134	0,0856
rBV_original_glszm_GrayLevelVariance	N=70 701.03 (382.45-941.01)	N=86 919.54 (664.11-1221.28)	0,0001	0,0867
FLAIR_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 21391.99 (14077.55-24595.98)	N=86 22789.29 (19538.51-25209.07)	0,0212	0,087
rBF_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.02 (0.02-0.03)	N=86 0.02 (0.01-0.02)	0	0,0872
rBV_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.54 (0.51-0.6)	N=86 0.56 (0.53-0.59)	0,047	0,0896
FLAIR_original_glcm_ClusterProminence	N=70 20868747.63 (14660864.83-36782739.36)	N=86 26804772.79 (14156751.65-63948872.06)	0,1001	0,0902
TMAX_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.61 (SD =0.14)	N=86 0.55 (SD =0.1)	0,0051	0,0921
TMAX_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.03 (0.02-0.04)	N=86 0.03 (0.03-0.05)	0,0537	0,0926
rBV_corrected_original_glcm_Id	N=70 0.22 (0.2-0.25)	N=86 0.2 (0.17-0.24)	0,0033	0,093
rBV_corrected_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.02 (0.02-0.03)	N=86 0.02 (0.01-0.02)	0	0,0938
FLAIR_original_glrlm_HighGrayLevelRunEmphasis	N=70 21663.03 (14391.46-24920.61)	N=86 23219.64 (19852.51-25692.48)	0,0201	0,0948
MTT_original_glszm_HighGrayLevelZoneEmphasis	N=70 5092.67 (3698.01-7625.98)	N=86 4930.99 (4039.39-5981.69)	0,6625	0,0955
TTP_original_firstorder_RootMeanSquared	N=70 2144.76 (2061.62-2312.93)	N=86 2094.21 (2028.55-2180.2)	0,0082	0,0956
T1_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,096
FLAIR_original_gldm_HighGrayLevelEmphasis	N=70 21706.15 (14424.62-24952.81)	N=86 23265.35 (19869.53-25757.64)	0,0195	0,0968
TMAX_original_ngtdm_Strength	N=70 63.45 (36.74-118.68)	N=86 34.96 (19.66-65.11)	0	0,0981
T1_original_glrlm_LowGrayLevelRunEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,1002
rBV_original_glcm_ClusterProminence	N=70 20986907.16 (4434292.93-44453685.1)	N=86 40568801.45 (23414629.16-82552975.8)	0,0001	0,1004
TMAX_original_firstorder_Entropy	N=70 5.74 (5.42-6.25)	N=86 5.52 (4.94-6.1)	0,0435	0,1007
TTP_original_firstorder_90Percentile	N=70 2427.01 (2259.74-2706.44)	N=86 2314.57 (2191.08-2464.99)	0,0061	0,1053
T1_original_gldm_LowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,1054
TMAX_original_glcm_InverseVariance	N=70 0.12 (0.08-0.17)	N=86 0.17 (0.12-0.23)	0,0027	0,1054
rBV_original_glcm_ClusterShade	N=70 113778.12 (36898.22-193299.91)	N=86 187136.77 (118132.52-330049.27)	0,0001	0,106
FLAIR_original_glszm_GrayLevelVariance	N=70 1059.81 (826.6-1355.65)	N=86 1015.6 (812.92-1307.32)	0,7824	0,1074
TMAX_original_firstorder_Uniformity	N=70 0.03 (0.02-0.04)	N=86 0.03 (0.03-0.05)	0,0636	0,1076
rBV_original_firstorder_Median	N=70 220.91 (191.54-284.68)	N=86 258.73 (196-331.02)	0,0348	0,1083
FLAIR_original_glszm_LowGrayLevelZoneEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,1085
MTT_original_firstorder_Skewness	N=70 1.12 (0.82-1.28)	N=86 1.21 (0.88-1.67)	0,0684	0,1092
MTT_original_firstorder_Entropy	N=70 6.76 (6.5-7.08)	N=86 6.76 (6.57-6.96)	0,6217	0,1097

MTT_original_glcm_Correlation	N=70 0.33 (0.24-0.4)	N=86 0.39 (0.33-0.43)	0,0009	0,1099
TMAX_original_glrlm_LongRunEmphasis	N=70 1.23 (1.15-1.39)	N=86 1.35 (1.23-1.53)	0,0003	0,1113
FLAIR_original_glcm_Autocorrelation	N=70 23069.63 (14854.35-25994.47)	N=86 24119.77 (19884-26670.97)	0,0443	0,1133
rBV_corrected_original_glcm_ClusterProminence	N=70 31309226.96 (4435660.98-80098710.99)	N=86 78600666.19 (45652293.98-112673888.97)	0	0,114
rBV_original_glcm_Correlation	N=70 0.35 (0.26-0.46)	N=86 0.4 (0.36-0.46)	0,0019	0,1144
TMAX_original_firstorder_Median	N=70 266.29 (211.96-390.72)	N=86 220.32 (173.91-297.95)	0,0016	0,1149
T1_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,1155
TTP_original_ngtdm_Busyness	N=70 0.05 (0.02-0.1)	N=86 0.1 (0.06-0.19)	0	0,1163
TMAX_original_glcm_Idn	N=70 0.94 (0.91-0.96)	N=86 0.96 (0.93-0.97)	0,0009	0,1212
K2_original_glszm_LowGrayLevelZoneEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0,0005	0,1232
TMAX_original_glszm_SmallAreaEmphasis	N=70 0.82 (0.77-0.86)	N=86 0.77 (0.72-0.82)	0,0015	0,1255
TMAX_original_firstorder_Mean	N=70 511.11 (368.94-756.02)	N=86 430.52 (278.7-609.92)	0,0038	0,1307
rBV_corrected_original_glcm_Correlation	N=70 0.46 (0.34-0.55)	N=86 0.51 (0.45-0.58)	0,0029	0,132
TTP_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.55 (SD =0.13)	N=86 0.5 (SD =0.1)	0,0213	0,1331
TTP_original_glcm_DifferenceEntropy	N=70 4.56 (SD =0.95)	N=86 4.29 (SD =0.86)	0,0664	0,1335
FLAIR_original_glrlm_GrayLevelVariance	N=70 1012.82 (754.63-1299.14)	N=86 924.18 (762.1-1232.19)	0,4554	0,1353
K2_original_glrlm_RunEntropy	N=70 8.61 (7.97-8.98)	N=86 9 (8.65-9.48)	0,0002	0,1366
FLAIR_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,1368
FLAIR_original_firstorder_Variance	N=70 631010.57 (466913.22-811137.29)	N=86 574026.67 (472725.37-764161.41)	0,4468	0,1372
FLAIR_original_gldm_GrayLevelVariance	N=70 1009.12 (745.64-1297.51)	N=86 918.76 (756.33-1222.09)	0,4426	0,1373
FLAIR_original_glrlm_LowGrayLevelRunEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,1384
TTP_original_firstorder_Mean	N=70 2093.43 (2017.2-2203.66)	N=86 2040.88 (1993.45-2094.42)	0,0043	0,1388
rBF_original_glcm_Idm	N=70 0.14 (0.12-0.17)	N=86 0.13 (0.1-0.15)	0,0095	0,1392
K2_original_glcm_Correlation	N=70 0.26 (SD =0.21)	N=86 0.36 (SD =0.18)	0,0011	0,1393
TMAX_original_firstorder_InterquartileRange	N=70 497.52 (282.36-744.84)	N=86 393.8 (203.76-557.73)	0,0144	0,1397
FLAIR_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.97 (0.95-0.98)	N=86 0.96 (0.95-0.97)	0,0059	0,141
FLAIR_original_gldm_LowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,1415
rBV_original_glcm_InverseVariance	N=70 0.16 (0.13-0.18)	N=86 0.15 (0.12-0.17)	0,1077	0,1423
rBV_original_ngtdm_Complexity	N=70 103132.35 (30392.55-226049.16)	N=86 253016.36 (166797.37-325606.65)	0	0,1423

FLAIR_original_glcm_SumSquares	N=70 894.61 (707.58-1240.27)	N=86 863.53 (647.06-1167.27)	0,6141	0,1424
tMIP_original_firstorder_InterquartileRange	N=70 2235.88 (1566.67-3558.6)	N=86 3341.36 (1968.38-6647.86)	0,0012	0,1428
TTP_original_glszm_LargeAreaEmphasis	N=70 35.22 (6.16-128.97)	N=86 310.26 (54.14-1196.39)	0	0,1431
FLAIR_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 23218.72 (15437.8-26711.94)	N=86 24996.05 (21259.66-28103.11)	0,0104	0,1434
TTP_original_firstorder_Skewness	N=70 1.86 (0.82-2.8)	N=86 1.96 (0.68-3.85)	0,5063	0,1442
tMIP_original_ngtdm_Complexity	N=70 7010656.82 (1836761.66-53079282.16)	N=86 39524935.87 (3937530.45-147041903.97)	0,001	0,1444
TTP_original_glszm_SmallAreaEmphasis	N=70 0.76 (SD =0.09)	N=86 0.73 (SD =0.08)	0,0414	0,1453
FLAIR_original_glrlm_ShortRunEmphasis	N=70 0.99 (0.98-0.99)	N=86 0.98 (0.98-0.99)	0,0062	0,1474
TTP_original_glcm_SumEntropy	N=70 5.97 (SD =0.94)	N=86 5.86 (SD =0.81)	0,4602	0,1477
TTP_original_glcm_Idmn	N=70 0.99 (0.98-1)	N=86 1 (0.99-1)	0,0001	0,1479
MTT_original_glcm_ClusterProminence	N=70 41285941.41 (17651217.03-99888926.16)	N=86 50284746.5 (22237749.68-93049750.92)	0,9815	0,1493
FLAIR_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,1515
FLAIR_original_glrlm_RunPercentage	N=70 0.98 (0.97-0.99)	N=86 0.98 (0.97-0.98)	0,0069	0,1517
TMAX_original_glszm_LowGrayLevelZoneEmphasis	N=70 0.02 (0.02-0.03)	N=86 0.02 (0.01-0.02)	0,0001	0,1526
TTP_original_glrlm_RunEntropy	N=70 5.76 (SD =0.75)	N=86 5.69 (SD =0.57)	0,5677	0,1531
rBF_original_firstorder_Minimum	N=70 0 (-12.05-0)	N=86 0 (-17.16-0)	0,0805	0,1573
tMIP_original_glcm_JointAverage	N=70 127.72 (99.94-195.15)	N=86 186.75 (128-298.23)	0,0017	0,1612
tMIP_original_glcm_SumAverage	N=70 255.44 (199.87-390.3)	N=86 373.49 (256-596.46)	0,0017	0,1612
FLAIR_original_glszm_ZonePercentage	N=70 0.8 (0.71-0.87)	N=86 0.75 (0.7-0.81)	0,0034	0,1615
TMAX_original_glrlm_RunVariance	N=70 0.11 (0.06-0.17)	N=86 0.15 (0.1-0.26)	0,0005	0,1627
FLAIR_original_firstorder_MeanAbsoluteDeviation	N=70 636.28 (542.37-747.56)	N=86 582.72 (538.2-699.84)	0,1295	0,1636
rBV_corrected_original_glcm_Idm	N=70 0.15 (0.13-0.17)	N=86 0.13 (0.11-0.16)	0,0065	0,1648
FLAIR_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 57609.13 (38875.75-89015.47)	N=86 75874.5 (60644.76-94982.49)	0,0013	0,1736
FLAIR_original_glcm_DifferenceVariance	N=70 410.5 (270.61-532.26)	N=86 362.93 (258.51-510.89)	0,5317	0,174
rBF_original_ngtdm_Strength	N=70 29.41 (16.9-48.41)	N=86 13.95 (7.9-23.01)	0	0,1751
rBV_corrected_original_firstorder_Kurtosis	N=70 10.45 (7.1-14.42)	N=86 8.71 (6.36-11.8)	0,0647	0,1772
TMAX_original_firstorder_RobustMeanAbsoluteDeviation	N=70 229.46 (140.16-378.43)	N=86 171 (90.58-278)	0,0197	0,1786
rBV_original_firstorder_10Percentile	N=70 -5.87 (-24.76-19.31)	N=86 -0.04 (-42.47-34.44)	0,9333	0,179
rBF_original_glszm_LowGrayLevelZoneEmphasis	N=70 0.01 (0-0.01)	N=86 0 (0-0.01)	0,0002	0,1801

FLAIR_original_gldm_SmallDependenceEmphasis	N=70 0.74 (0.63-0.83)	N=86 0.68 (0.62-0.75)	0,004	0,183
K2_original_firstorder_90Percentile	N=70 2873.33 (2394.25-4651.05)	N=86 3997.1 (2737.36-5767.15)	0,0129	0,1833
TMAX_original_firstorder_Minimum	N=70 -176.04 (-232.87--134.25)	N=86 -205.94 (-326.31--142.12)	0,1077	0,184
tMIP_original_firstorder_RobustMeanAbsoluteDeviation	N=70 977.78 (696.86-1758.97)	N=86 1570.64 (886.76-3105.31)	0,0021	0,1862
FLAIR_original_gldm_DependenceVariance	N=70 0.53 (0.33-0.88)	N=86 0.71 (0.51-0.99)	0,0028	0,1865
K2_original_firstorder_Range	N=70 144471.3 (68225.08-203882.38)	N=86 198731.56 (140026.21-226818.39)	0,0013	0,1879
tMIP_original_firstorder_Median	N=70 2228.24 (1830.29-2817.03)	N=86 2984.76 (2099.53-3924.27)	0,0011	0,1879
FLAIR_original_glrlm_LongRunEmphasis	N=70 1.06 (1.03-1.09)	N=86 1.07 (1.05-1.09)	0,0089	0,1888
TMAX_original_firstorder_90Percentile	N=70 1230.25 (764.67-2147.18)	N=86 782.57 (530.84-1617.86)	0,0049	0,1943
K2_original_glcm_DifferenceEntropy	N=70 7.42 (6.76-7.97)	N=86 7.9 (7.46-8.43)	0,0009	0,1965
FLAIR_original_glszm_SmallAreaEmphasis	N=70 0.87 (0.81-0.91)	N=86 0.84 (0.81-0.87)	0,0077	0,1966
TTP_original_glcm_JointEnergy	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,1199	0,2027
TMAX_original_glrlm_RunEntropy	N=70 6.11 (5.69-6.56)	N=86 5.97 (5.57-6.46)	0,2811	0,2069
rBV_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 4468.51 (2345.61-6897.14)	N=86 9005.58 (5066.02-14846.42)	0	0,2072
FLAIR_original_gldm_LargeDependenceEmphasis	N=70 2.59 (1.97-3.74)	N=86 3.19 (2.55-3.94)	0,0032	0,2077
TMAX_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 9056.22 (6050.89-20399.09)	N=86 29633.44 (12197.39-111481.06)	0	0,2083
TTP_original_glcm_MaximumProbability	N=70 0.01 (0.01-0.02)	N=86 0.02 (0.01-0.02)	0,1157	0,211
FLAIR_original_glrlm_RunVariance	N=70 0.02 (0.01-0.03)	N=86 0.02 (0.02-0.03)	0,0088	0,2154
TTP_original_firstorder_Range	N=70 5929.9 (5166.48-7066.14)	N=86 6488.83 (5592.02-7651.48)	0,0053	0,2167
FLAIR_original_glcm_JointAverage	N=70 150.43 (118.88-159.76)	N=86 153.22 (139.62-161.8)	0,0582	0,2171
FLAIR_original_glcm_SumAverage	N=70 300.85 (237.77-319.51)	N=86 306.43 (279.23-323.6)	0,0582	0,2171
MTT_original_glcm_DifferenceEntropy	N=70 5.84 (5.42-6.13)	N=86 5.72 (5.5-6.01)	0,2065	0,2216
K2_original_ngtdm_Busyness	N=70 0 (0-0)	N=86 0 (0-0)	0,7123	0,2251
T1_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0	0,2281
tMIP_original_firstorder_Mean	N=70 3266.37 (2276.56-4841.25)	N=86 4338.74 (2953.3-7159.19)	0,0064	0,2293
MTT_original_glrlm_RunEntropy	N=70 6.9 (6.63-7.19)	N=86 6.93 (6.74-7.11)	0,8824	0,2296
TMAX_original_ngtdm_Contrast	N=70 0.72 (0.37-1.96)	N=86 0.27 (0.13-0.84)	0,0001	0,2318
TTP_original_glcm_DifferenceAverage	N=70 12.33 (6.61-21.54)	N=86 9.45 (5.37-13.71)	0,045	0,237
rBF_original_glcm_Idn	N=70 0.93 (0.92-0.94)	N=86 0.93 (0.91-0.94)	0,1663	0,2383
MTT_original_gldm_LargeDependenceEmphasis	N=70 2.79 (2.21-3.68)	N=86 3.55 (2.97-5.34)	0	0,2411

FLAIR_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.7 (SD =0.12)	N=86 0.65 (SD =0.09)	0,009	0,2443
TMAX_original_firstorder_MeanAbsoluteDeviation	N=70 514.47 (325.05-790.75)	N=86 387.87 (250.39-629.59)	0,0428	0,2467
rBF_original_glcm_ClusterShade	N=70 171115.44 (47092.64-319016.01)	N=86 316033.62 (203771.02-418057.07)	0	0,2572
TTP_original_firstorder_Maximum	N=70 6014.98 (5180.23-7043.94)	N=86 6488.83 (5584.41-7683.17)	0,0128	0,264
rBV_corrected_original_glszm_LowGrayLevelZoneEmphasis	N=70 0.01 (0-0.01)	N=86 0 (0-0.01)	0	0,2647
T1_original_glcm_Idn	N=70 0.94 (0.92-0.96)	N=86 0.95 (0.94-0.96)	0,0002	0,2718
tMIP_original_firstorder_90Percentile	N=70 6221.69 (3852.43-10979.22)	N=86 8722.65 (5369.53-17077.53)	0,003	0,2731
rBV_corrected_original_ngtdm_Contrast	N=70 0.48 (0.29-0.63)	N=86 0.55 (0.36-0.88)	0,027	0,2733
MTT_original_glcm_ClusterShade	N=70 158293.26 (67331.88-309744.73)	N=86 146924.97 (84988.81-291193.88)	0,9446	0,2736
TMAX_original_firstorder_RootMeanSquared	N=70 1043.13 (659.08-1406.59)	N=86 844.95 (601.99-1228.4)	0,0515	0,274
TTP_original_ngtdm_Strength	N=70 31 (14.06-59.55)	N=86 19.82 (11.05-33.22)	0,0049	0,2741
TMAX_original_glcm_DifferenceEntropy	N=70 5.28 (4.64-5.92)	N=86 4.91 (4.29-5.7)	0,0824	0,2743
TMAX_original_glszm_GrayLevelNonUniformityNormalized	N=70 0.02 (0.01-0.02)	N=86 0.02 (0.01-0.03)	0,321	0,2762
FLAIR_original_firstorder_Minimum	N=70 344.17 (-8.66-1234.92)	N=86 112.68 (-11.45-414.45)	0,0198	0,2769
tMIP_original_glcm_ClusterTendency	N=70 54498.58 (14790.31-131696.1)	N=86 88373.48 (22312.72-303899.97)	0,0265	0,2794
TTP_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 114094.81 (28373.56-688023.21)	N=86 1250989.05 (258781.6-5870142.03)	0	0,2799
tMIP_original_firstorder_RootMeanSquared	N=70 4866.2 (2943.78-7834.11)	N=86 6090.17 (3861.39-10706.46)	0,014	0,2801
TTP_original_firstorder_Minimum	N=70 0 (-13.64-470.78)	N=86 0 (-13-0)	0,0482	0,2835
rBV_corrected_original_glcm_Idn	N=70 0.94 (0.93-0.95)	N=86 0.93 (0.92-0.95)	0,4532	0,2862
FLAIR_original_glcm_Idm	N=70 0.06 (0.05-0.08)	N=86 0.07 (0.06-0.09)	0,0074	0,2875
TMAX_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0.05 (0.04-0.08)	N=86 0.04 (0.02-0.07)	0,07	0,2882
tMIP_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 39733.78 (17097.54-104378.86)	N=86 65685.8 (30535.4-195877.14)	0,0081	0,2884
tMIP_original_glcm_Autocorrelation	N=70 25059.99 (12856.67-60864.71)	N=86 46151.7 (24073.86-137464.88)	0,0023	0,2919
tMIP_original_gldm_HighGrayLevelEmphasis	N=70 39074.01 (16594.8-102929.51)	N=86 63939.94 (29399.03-193599.3)	0,0102	0,2919
TTP_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 3351.1 (2118.52-5221.21)	N=86 2553.59 (1720.43-4067.36)	0,0568	0,2922
tMIP_original_glrlm_HighGrayLevelRunEmphasis	N=70 39474.99 (16795.13-103648.23)	N=86 64358.85 (30015.47-194528.81)	0,0093	0,2925
rBV_corrected_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 4791.01 (3189.32-7411.78)	N=86 8761.98 (6864.13-11823.06)	0	0,2932
tMIP_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 39413.09 (16724.96-103530.04)	N=86 64098.08 (29904.41-194212.46)	0,0095	0,2935
tMIP_original_firstorder_MeanAbsoluteDeviation	N=70 2141.75 (1240.34-3613.89)	N=86 2942.21 (1521.26-5646.85)	0,0162	0,2947
rBV_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.03 (0.02-0.04)	N=86 0.02 (0.02-0.03)	0,0007	0,2952

tMIP_original_glszm_HighGrayLevelZoneEmphasis	N=70 43544.1 (18724.81-110018.28)	N=86 69118.81 (32396.84-204739.83)	0,0084	0,2967
TMAX_original_glcm_DifferenceAverage	N=70 19.51 (14.22-29.7)	N=86 15.44 (9.36-24.32)	0,0354	0,2987
K2_original_firstorder_Entropy	N=70 8.57 (7.85-8.9)	N=86 8.92 (8.57-9.4)	0,0007	0,2992
tMIP_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 42779.08 (17808.05-108673.82)	N=86 67055.01 (30270.98-200876.46)	0,01	0,2996
TMAX_original_glszm_LargeAreaEmphasis	N=70 23.44 (5.3-74.62)	N=86 106.3 (24.52-446.9)	0	0,3026
FLAIR_original_firstorder_RobustMeanAbsoluteDeviation	N=70 453.85 (381.46-540.98)	N=86 411.42 (367.08-483.23)	0,0403	0,3043
rBV_original_gldm_DependenceVariance	N=70 3.08 (2.04-3.98)	N=86 3.71 (2.46-5.75)	0,0075	0,3047
tMIP_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 37630.67 (14923.48-99046.39)	N=86 59256.93 (25183.89-186167.01)	0,0132	0,3059
MTT_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,5132	0,3062
FLAIR_original_firstorder_InterquartileRange	N=70 1091.31 (877.53-1315.04)	N=86 998.56 (875.04-1197.66)	0,0494	0,312
TMAX_original_glcm_Idmn	N=70 0.98 (0.97-0.99)	N=86 0.99 (0.98-0.99)	0,0128	0,313
rBF_original_ngtdm_Contrast	N=70 0.53 (0.33-0.69)	N=86 0.66 (0.44-1.04)	0,0077	0,3133
rBV_corrected_original_ngtdm_Strength	N=70 32.79 (20.85-46.34)	N=86 16.57 (10.07-26.1)	0	0,3157
TMAX_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 1939.27 (1073.58-3439.17)	N=86 1296.01 (691.05-2733.66)	0,0424	0,3184
T1_original_ngtdm_Busyness	N=70 0.03 (0.01-0.05)	N=86 0.04 (0.02-0.07)	0,0002	0,3243
K2_original_ngtdm_Contrast	N=70 26.92 (5.98-98.37)	N=86 18.33 (5.02-71.67)	0,3107	0,3263
FLAIR_original_glcm_Correlation	N=70 0.47 (0.33-0.57)	N=86 0.53 (0.45-0.61)	0,0006	0,3312
FLAIR_original_glcm_Id	N=70 0.12 (0.1-0.15)	N=86 0.14 (0.12-0.16)	0,0058	0,3317
TTP_original_firstorder_MeanAbsoluteDeviation	N=70 308.73 (159.71-441.95)	N=86 230.28 (134.64-342.55)	0,1124	0,3352
TTP_original_firstorder_Median	N=70 2045.07 (2009.49-2093.19)	N=86 2026.34 (1990.3-2060.35)	0,0376	0,3371
FLAIR_original_glcm_Contrast	N=70 971.27 (614.61-1336.68)	N=86 773.25 (568.04-1191.58)	0,0817	0,3378
rBF_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 4854.97 (3364.49-7047.26)	N=86 9045.28 (7111.38-14141.26)	0	0,3429
FLAIR_original_firstorder_RootMeanSquared	N=70 4034.41 (3845-4178.13)	N=86 3975.54 (3827.45-4127.84)	0,3895	0,3443
TMAX_original_glcm_ClusterProminence	N=70 165297483.15 (36593196.55-282936766.26)	N=86 143425815.2 (39889734.77-296006734.15)	0,8155	0,3479
FLAIR_original_ngtdm_Busyness	N=70 0.03 (0.01-0.06)	N=86 0.05 (0.03-0.09)	0,0003	0,3495
rBV_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0.76 (0.27-1.64)	N=86 1.22 (0.55-2.12)	0,0221	0,3496
rBV_original_glszm_LowGrayLevelZoneEmphasis	N=70 0.01 (0.01-0.02)	N=86 0.01 (0.01-0.02)	0,11	0,3545
FLAIR_original_firstorder_90Percentile	N=70 4903.97 (4751.65-5176.25)	N=86 4785.19 (4668.89-4976.34)	0,0202	0,3555
FLAIR_original_gldm_DependenceNonUniformityNormalized	N=70 0.51 (SD =0.15)	N=86 0.45 (SD =0.11)	0,0022	0,3565

MTT_original_glrlm_LongRunEmphasis	N=70 1.06 (1.04-1.09)	N=86 1.08 (1.06-1.13)	0,0001	0,3588
rBV_original_firstorder_Kurtosis	N=70 14.93 (10.6-22.03)	N=86 14.49 (11-19.8)	0,9844	0,3594
MTT_original_gldm_DependenceVariance	N=70 0.65 (0.45-1.07)	N=86 1.01 (0.71-1.77)	0	0,3616
rBF_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 6411.49 (3768.62-9225.91)	N=86 11478.51 (9034.86-17257.02)	0	0,3693
K2_original_firstorder_RootMeanSquared	N=70 7916.07 (3470.33-15031.52)	N=86 12426.3 (6154.59-21101.68)	0,0045	0,3698
K2_original_glcm_Idmn	N=70 1 (0.99-1)	N=86 1 (0.99-1)	0,9702	0,3758
tMIP_original_glcm_SumSquares	N=70 17866.52 (5508.25-45381.48)	N=86 31292.69 (7450.36-97656.48)	0,0357	0,3759
MTT_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 9791.4 (7644.64-12955.04)	N=86 9957.19 (8254.89-13410.27)	0,264	0,3822
tMIP_original_firstorder_Variance	N=70 12543461.01 (3524026.1-35479288.07)	N=86 21926224.02 (4559352.64-71508771.03)	0,0502	0,3917
tMIP_original_gldm_GrayLevelVariance	N=70 20069.99 (5640.99-56768.18)	N=86 35083.15 (7303.1-114415.97)	0,0502	0,3917
tMIP_original_glrlm_GrayLevelVariance	N=70 20190.18 (5658.54-57239.38)	N=86 34938.58 (8044.19-114016.08)	0,0494	0,3973
MTT_original_gldm_LowGrayLevelEmphasis	N=70 0.01 (0-0.02)	N=86 0.01 (0-0.02)	0,9645	0,3979
MTT_original_firstorder_Uniformity	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,4662	0,398
tMIP_original_glcm_InverseVariance	N=70 0.04 (SD =0.02)	N=86 0.04 (SD =0.02)	0,1934	0,3985
MTT_original_firstorder_10Percentile	N=70 63.81 (-7.36-189.67)	N=86 78.87 (-40.64-208.05)	0,905	0,4005
MTT_original_glrlm_LowGrayLevelRunEmphasis	N=70 0.01 (0-0.02)	N=86 0.01 (0-0.02)	0,8796	0,4016
rBV_corrected_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.92 (0.91-0.93)	N=86 0.93 (0.9-0.94)	0,7797	0,4023
MTT_original_ngtdm_Busyness	N=70 0.09 (0.03-0.16)	N=86 0.18 (0.12-0.29)	0	0,4036
TMAX_original_glrlm_LowGrayLevelRunEmphasis	N=70 0.06 (0.04-0.09)	N=86 0.05 (0.03-0.08)	0,1233	0,4037
MTT_original_ngtdm_Complexity	N=70 309512.28 (165615.71-575526.62)	N=86 481652.95 (286032.29-766449.42)	0,0019	0,4045
tMIP_original_glcm_Idn	N=70 0.94 (0.93-0.95)	N=86 0.94 (0.92-0.95)	0,4883	0,4047
K2_original_glszm_GrayLevelNonUniformityNormalized	N=70 0 (0-0.01)	N=86 0 (0-0)	0,002	0,4072
TMAX_original_glcm_SumEntropy	N=70 6.65 (5.92-7.06)	N=86 6.42 (5.73-7.15)	0,6091	0,4078
rBV_original_ngtdm_Contrast	N=70 0.32 (0.22-0.49)	N=86 0.34 (0.23-0.58)	0,7661	0,4183
rBF_original_glcm_Idmn	N=70 0.98 (0.98-0.99)	N=86 0.98 (0.98-0.99)	0,0636	0,4229
tMIP_original_glcm_ClusterProminence	N=70 33605275824.13 (2455583695.08-375448909986.29)	N=86 131165536766.93 (4047218730.28-1116464610086.68)	0,0843	0,4233
TMAX_original_glcm_ClusterShade	N=70 575585.98 (145668.3-963348.62)	N=86 471292.77 (152745.57-874130.13)	0,6166	0,4312
K2_original_firstorder_InterquartileRange	N=70 2430.04 (1740.98-3494.53)	N=86 2804.71 (1982.98-4046.88)	0,0793	0,4328
tMIP_original_glszm_GrayLevelVariance	N=70 21449.34 (5951.23-61880.87)	N=86 35711.41 (8140.07-117218.97)	0,0568	0,4346

rBV_original_gldm_LargeDependenceEmphasis	N=70 8.91 (6.55-11.26)	N=86 10.14 (7.51-13.06)	0,0695	0,4353
rBV_original_glszm_LargeAreaEmphasis	N=70 16.74 (6.95-47.05)	N=86 29.59 (12.59-73.51)	0,006	0,4367
TTP_original_ngtdm_Contrast	N=70 0.25 (0.07-0.58)	N=86 0.1 (0.04-0.25)	0,0002	0,4387
K2_original_glcm_InverseVariance	N=70 0.03 (0.02-0.03)	N=86 0.03 (0.02-0.04)	0,4404	0,4414
rBV_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0.05 (0.02-0.1)	N=86 0.07 (0.04-0.1)	0,135	0,4429
tMIP_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.82 (SD =0.07)	N=86 0.83 (SD =0.07)	0,7029	0,4525
K2_original_glszm_GrayLevelVariance	N=70 95698.14 (19445.09-313597.61)	N=86 237700.77 (70522.13-589437.8)	0,0158	0,4531
T1_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0,0001	0,4538
rBV_original_gldm_SmallDependenceEmphasis	N=70 0.49 (SD =0.11)	N=86 0.48 (SD =0.09)	0,687	0,4544
T1_original_gldm_DependenceNonUniformityNormalized	N=70 0.51 (0.45-0.57)	N=86 0.47 (0.42-0.52)	0,002	0,4545
FLAIR_original_glcm_DifferenceAverage	N=70 22.92 (17.54-26.58)	N=86 19.25 (16.58-24.63)	0,0212	0,4581
rBV_corrected_original_glrlm_ShortRunEmphasis	N=70 0.97 (0.96-0.97)	N=86 0.97 (0.96-0.97)	0,6859	0,4585
TMAX_original_gldm_LowGrayLevelEmphasis	N=70 0.07 (0.04-0.1)	N=86 0.06 (0.03-0.1)	0,1763	0,462
FLAIR_original_firstorder_Mean	N=70 3953 (3788.4-4085.09)	N=86 3889.1 (3748.77-4050.34)	0,4706	0,4646
rBF_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.93 (0.91-0.94)	N=86 0.93 (0.91-0.94)	0,8909	0,4655
MTT_original_glszm_GrayLevelNonUniformityNormalized	N=70 0.01 (0.01-0.01)	N=86 0.01 (0.01-0.01)	0,846	0,4664
T1_original_gldm_DependenceVariance	N=70 0.5 (0.38-0.65)	N=86 0.61 (0.51-0.77)	0,0012	0,4672
tMIP_original_glszm_SmallAreaEmphasis	N=70 0.92 (SD =0.03)	N=86 0.92 (SD =0.03)	0,7094	0,4672
rBV_corrected_original_firstorder_Uniformity	N=70 0.02 (0.02-0.03)	N=86 0.02 (0.01-0.02)	0	0,4697
TMAX_original_glcm_Correlation	N=70 0.32 (SD =0.12)	N=86 0.35 (SD =0.09)	0,0585	0,4718
TTP_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0 (0-0.02)	N=86 0.01 (0-0.03)	0,5063	0,4733
TTP_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0.04 (0.01-0.34)	N=86 0.1 (0.02-0.45)	0,0403	0,4742
T1_original_glszm_ZonePercentage	N=70 0.8 (SD =0.07)	N=86 0.77 (SD =0.06)	0,0015	0,4767
MTT_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0.01 (0-0.02)	N=86 0.01 (0-0.02)	0,7552	0,4767
rBV_corrected_original_glcm_Idmn	N=70 0.99 (0.98-0.99)	N=86 0.98 (0.98-0.99)	0,2404	0,4781
TMAX_original_ngtdm_Complexity	N=70 471095.27 (141113.51-653866.29)	N=86 428757.6 (216219.08-758037.76)	0,1965	0,4819
TTP_original_glcm_ClusterShade	N=70 56985.81 (4507.5-173236.81)	N=86 39447.54 (2433.73-107091.9)	0,1844	0,4821
rBV_original_gldm_LowGrayLevelEmphasis	N=70 0.03 (0.02-0.05)	N=86 0.03 (0.02-0.05)	0,4237	0,483
T1_original_gldm_SmallDependenceEmphasis	N=70 0.74 (SD =0.08)	N=86 0.7 (SD =0.07)	0,0018	0,4859
MTT_original_glrlm_RunVariance	N=70 0.02 (0.02-0.03)	N=86 0.03 (0.02-0.05)	0,0001	0,4859

T1_original_gldm_LargeDependenceEmphasis	N=70 2.63 (SD =0.7)	N=86 2.99 (SD =0.7)	0,0017	0,4882
TTP_original_firstorder_Variance	N=70 343642.34 (86465.45-601209.73)	N=86 208547.08 (96038.73-437024.07)	0,2795	0,4882
TTP_original_gldm_GrayLevelVariance	N=70 550.06 (138.22-960.7)	N=86 333.6 (154.04-698.16)	0,2763	0,4888
K2_original_gldm_DependenceNonUniformityNormalized	N=70 0.74 (0.68-0.81)	N=86 0.71 (0.66-0.76)	0,0288	0,4896
FLAIR_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 45357.58 (30073.79-74189.1)	N=86 62869.66 (48042.84-83842.93)	0,0006	0,4959
tMIP_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0.01 (0.01-0.03)	N=86 0.02 (0.01-0.03)	0,2223	0,4961
TMAX_original_firstorder_Skewness	N=70 3.06 (2.06-4.49)	N=86 4.17 (2.76-5.67)	0,0046	0,4978
tMIP_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0.2 (0.03-0.51)	N=86 0.29 (0.14-0.75)	0,0278	0,4999
tMIP_original_firstorder_10Percentile	N=70 825.23 (593.89-1076.31)	N=86 941.72 (535.57-1349.18)	0,1475	0,5006
tMIP_original_ngtdm_Busyness	N=70 0.01 (0.01-0.02)	N=86 0.02 (0.01-0.04)	0,0528	0,5009
rBF_original_gldm_DependenceVariance	N=70 2.03 (1.52-3.08)	N=86 2.47 (1.81-3.44)	0,0824	0,5034
TTP_original_glrlm_GrayLevelVariance	N=70 574.43 (152.48-977.1)	N=86 349.23 (178.34-740.4)	0,2973	0,5056
FLAIR_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0,0005	0,5075
TTP_original_glszm_GrayLevelNonUniformityNormalized	N=70 0.02 (0.02-0.03)	N=86 0.02 (0.02-0.03)	0,5224	0,5087
FLAIR_original_ngtdm_Strength	N=70 17.8 (9.97-50.28)	N=86 12.16 (6.88-19.91)	0,0017	0,5116
TTP_original_glcm_SumSquares	N=70 447.4 (112.8-889.95)	N=86 302.99 (119.91-628.1)	0,309	0,5129
TTP_original_glcm_ClusterTendency	N=70 1179.2 (303.79-2275.54)	N=86 804.89 (309.58-1699.92)	0,3875	0,5145
rBV_original_firstorder_Skewness	N=70 2.91 (2.23-3.58)	N=86 2.99 (2.45-3.38)	0,7688	0,5162
tMIP_original_ngtdm_Contrast	N=70 6.01 (1.97-14.39)	N=86 4.62 (1.8-11.04)	0,4094	0,5178
tMIP_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.98 (0.97-0.98)	N=86 0.98 (0.97-0.98)	0,8572	0,5188
TMAX_original_glcm_JointEnergy	N=70 0 (0-0.01)	N=86 0 (0-0)	0,9702	0,5196
tMIP_original_glcm_DifferenceAverage	N=70 67.6 (40.23-117.22)	N=86 83.31 (49.09-166.4)	0,0494	0,5199
rBV_original_glcm_Id	N=70 0.24 (0.21-0.28)	N=86 0.23 (0.2-0.26)	0,1465	0,5231
rBV_original_glrlm_LowGrayLevelRunEmphasis	N=70 0.03 (0.02-0.04)	N=86 0.03 (0.02-0.04)	0,5248	0,5232
T1_original_glcm_Correlation	N=70 0.48 (0.4-0.57)	N=86 0.56 (0.46-0.64)	0,0058	0,5234
tMIP_original_glrlm_ShortRunEmphasis	N=70 0.99 (0.99-0.99)	N=86 0.99 (0.99-0.99)	0,8684	0,5236
tMIP_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 60783.88 (24063.59-121581.86)	N=86 114096.82 (49677.4-272513.69)	0,0012	0,5249
tMIP_original_gldm_DependenceVariance	N=70 0.48 (0.33-0.7)	N=86 0.62 (0.4-0.95)	0,0524	0,5268
TTP_original_glcm_Contrast	N=70 582.25 (168.2-1264.67)	N=86 417.43 (166.68-820.95)	0,1892	0,5269
K2_original_gldm_DependenceVariance	N=70 0.26 (0.16-0.46)	N=86 0.32 (0.24-0.73)	0,0105	0,5273

T1_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.7 (0.66-0.75)	N=86 0.67 (0.63-0.71)	0,0033	0,529
FLAIR_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0,0007	0,5301
T1_original_glszm_SmallAreaEmphasis	N=70 0.87 (SD =0.04)	N=86 0.85 (SD =0.04)	0,0027	0,5305
tMIP_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 64127.46 (25154.98-124108.67)	N=86 112712.69 (50059.44-267860.56)	0,0008	0,5309
K2_original_firstorder_RobustMeanAbsoluteDeviation	N=70 1124.84 (763.57-1680.84)	N=86 1353.3 (959.13-1985.78)	0,0793	0,5357
TMAX_original_glcm_MaximumProbability	N=70 0.02 (0.01-0.03)	N=86 0.02 (0.01-0.03)	0,4794	0,5384
rBF_original_glrlm_ShortRunEmphasis	N=70 0.97 (0.96-0.97)	N=86 0.97 (0.96-0.98)	0,8017	0,5388
rBV_corrected_original_glcm_ClusterShade	N=70 143419.19 (28816.52-342055.75)	N=86 309668.39 (179143.88-413701.37)	0,0001	0,5424
rBV_original_glszm_ZonePercentage	N=70 0.55 (SD =0.11)	N=86 0.54 (SD =0.1)	0,4683	0,5449
rBV_corrected_original_glcm_JointEnergy	N=70 0 (0-0)	N=86 0 (0-0)	0	0,5479
tMIP_original_glrlm_RunPercentage	N=70 0.99 (0.99-0.99)	N=86 0.99 (0.98-0.99)	0,6911	0,5524
rBV_original_ngtdm_Busyness	N=70 0.31 (0.16-0.47)	N=86 0.61 (0.33-1.16)	0	0,5528
rBV_original_ngtdm_Strength	N=70 29.96 (18.94-54.99)	N=86 19.78 (12.15-26.81)	0	0,5528
K2_original_ngtdm_Complexity	N=70 61316072.55 (9759582.58-275606686.61)	N=86 343566537.65 (75350631.18-1186191226.3)	0,0001	0,555
MTT_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0.1 (0.03-0.53)	N=86 0.24 (0.06-0.8)	0,0909	0,555
TMAX_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 3264.15 (1691.79-4939.3)	N=86 2995.12 (1669.7-4741.83)	0,4094	0,5556
MTT_original_glcm_MaximumProbability	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,1763	0,5592
TTP_original_glszm_GrayLevelVariance	N=70 786.61 (265.11-1213.32)	N=86 678.7 (471.57-1124.55)	0,9418	0,5621
TMAX_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 2429.73 (1315.81-4031.21)	N=86 1909.63 (1012.87-3478.44)	0,1397	0,5623
MTT_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0.01 (0-0.03)	N=86 0.02 (0-0.04)	0,5506	0,5624
rBV_corrected_original_glrlm_LongRunEmphasis	N=70 1.19 (1.15-1.24)	N=86 1.19 (1.15-1.26)	0,8017	0,5642
T1_original_glrlm_RunVariance	N=70 0.02 (SD =0.01)	N=86 0.02 (SD =0.01)	0,0048	0,566
rBV_corrected_original_glrlm_RunPercentage	N=70 0.96 (0.94-0.96)	N=86 0.95 (0.94-0.96)	0,9531	0,5664
K2_original_glrlm_LongRunEmphasis	N=70 1.03 (1.02-1.04)	N=86 1.03 (1.02-1.05)	0,0291	0,567
T1_original_ngtdm_Complexity	N=70 319641.64 (131474.39-595338)	N=86 741758.76 (362975.31-1385961.81)	0	0,5673
TTP_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 0.07 (0.01-0.85)	N=86 0.43 (0.08-2.37)	0,0004	0,5673
K2_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0 (0-0.01)	N=86 0 (0-0)	0,0096	0,5682
K2_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,6728	0,5682
rBF_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0.22 (0.07-0.64)	N=86 0.33 (0.17-0.91)	0,0734	0,5705

T1_original_ngtdm_Strength	N=70 36.48 (21.88-73)	N=86 30.93 (17.75-48.94)	0,0621	0,5708
rBF_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0.02 (0.01-0.04)	N=86 0.02 (0.01-0.05)	0,5771	0,5708
T1_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.97 (SD =0.01)	N=86 0.96 (SD =0.01)	0,0028	0,5718
T1_original_glszm_GrayLevelVariance	N=70 1392.07 (922.67-1783.45)	N=86 1745.53 (1262.17-2663.7)	0,0008	0,5728
T1_original_glrlm_GrayLevelVariance	N=70 1320.62 (851.27-1705.08)	N=86 1634.76 (1162.19-2411.85)	0,0018	0,5729
T1_original_firstorder_Variance	N=70 822157.4 (528261.6-1060679.83)	N=86 1016805.18 (722436.63-1499651.66)	0,0019	0,573
T1_original_gldm_GrayLevelVariance	N=70 1314.38 (845.45-1694.38)	N=86 1623.8 (1156.52-2398.75)	0,0019	0,573
T1_original_glrlm_RunPercentage	N=70 0.98 (SD =0.01)	N=86 0.98 (SD =0.01)	0,0033	0,5732
T1_original_glcm_Contrast	N=70 953.73 (656.25-1576.4)	N=86 1334.9 (817.78-2002.51)	0,0066	0,5737
T1_original_glcm_SumSquares	N=70 1101.5 (749.37-1483.57)	N=86 1455.99 (1043.96-2215.98)	0,0003	0,5738
T1_original_glcm_ClusterTendency	N=70 3492 (2340.84-4581.38)	N=86 4369.53 (3335.62-7105.14)	0,0001	0,5741
T1_original_glrlm_LongRunEmphasis	N=70 1.06 (SD =0.02)	N=86 1.06 (SD =0.02)	0,0046	0,5767
T1_original_gldm_LargeDependenceHighGrayLevelEmphasis	N=70 41310.16 (29852.13-51461.94)	N=86 58427.96 (38924.15-86119.55)	0,0001	0,5784
T1_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 33732.4 (24334.28-40753.6)	N=86 48398.11 (32760.82-69651.58)	0	0,5798
rBV_corrected_original_glrlm_RunVariance	N=70 0.08 (0.06-0.11)	N=86 0.08 (0.06-0.12)	0,475	0,5798
rBV_original_glrlm_RunPercentage	N=70 0.94 (0.93-0.96)	N=86 0.94 (0.93-0.95)	0,2333	0,5804
T1_original_glrlm_ShortRunEmphasis	N=70 0.99 (SD =0)	N=86 0.98 (SD =0)	0,0035	0,5808
tMIP_original_gldm_LargeDependenceEmphasis	N=70 2.17 (1.91-2.63)	N=86 2.38 (1.93-2.84)	0,1797	0,5829
rBV_corrected_original_ngtdm_Busyness	N=70 0.17 (0.08-0.26)	N=86 0.35 (0.2-0.6)	0	0,5852
TMAX_original_gldm_HighGrayLevelEmphasis	N=70 2288.49 (1291.73-3924.81)	N=86 1840.94 (939.42-3324.95)	0,1173	0,5856
rBF_original_gldm_LargeDependenceEmphasis	N=70 6.68 (5.31-8.93)	N=86 7.21 (5.65-9.02)	0,4662	0,587
rBV_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.91 (0.89-0.93)	N=86 0.91 (0.89-0.92)	0,2156	0,5875
T1_original_glszm_HighGrayLevelZoneEmphasis	N=70 16810.2 (12429.47-21860.27)	N=86 22762.1 (14125.37-32240.92)	0,0008	0,5879
rBF_original_ngtdm_Busyness	N=70 0.18 (0.08-0.29)	N=86 0.38 (0.2-0.64)	0	0,5888
T1_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 17798.22 (12373.16-22843.17)	N=86 23256.35 (14892.96-33195.48)	0,001	0,589
T1_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 14675.43 (10578.36-18747.77)	N=86 19475.83 (11963.07-28441.23)	0,0012	0,5893
rBV_corrected_original_gldm_DependenceVariance	N=70 2.29 (1.79-3.23)	N=86 2.83 (1.91-4.1)	0,0769	0,5893
T1_original_glcm_DifferenceVariance	N=70 448.44 (267.67-837.45)	N=86 718.71 (443.2-1090.23)	0,0004	0,5896
T1_original_glrlm_HighGrayLevelRunEmphasis	N=70 16942.51 (11802.28-21688.87)	N=86 22074.7 (13910.55-31772.95)	0,0012	0,5896
T1_original_gldm_HighGrayLevelEmphasis	N=70 16906.81 (11767.2-21709.43)	N=86 22004.37 (13844.71-31735.45)	0,0013	0,5898

T1_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 16734.46 (11712.83-21395.02)	N=86 21789.39 (13650.5-31462.08)	0,0013	0,5898
tMIP_original_glrlm_LowGrayLevelRunEmphasis	N=70 0.01 (0-0.02)	N=86 0.01 (0-0.01)	0,7961	0,5907
MTT_original_firstorder_Minimum	N=70 -443.52 (-633.53--295.77)	N=86 -441.49 (-687.01--288.48)	0,9815	0,591
tMIP_original_gldm_LowGrayLevelEmphasis	N=70 0.01 (0-0.02)	N=86 0.01 (0-0.02)	0,5722	0,5914
TMAX_original_glrlm_HighGrayLevelRunEmphasis	N=70 2444.22 (1328.47-4071.17)	N=86 1976.93 (1044.13-3502.39)	0,1536	0,5919
rBF_original_glrlm_RunVariance	N=70 0.07 (0.06-0.1)	N=86 0.07 (0.06-0.11)	0,5178	0,5921
T1_original_gldm_SmallDependenceHighGrayLevelEmphasis	N=70 12621.57 (9137.39-16084.3)	N=86 16842.84 (9814.86-24197.77)	0,0037	0,5931
T1_original_glcm_Autocorrelation	N=70 15179.88 (11323.07-20183.8)	N=86 20684.05 (12874.4-30399.39)	0,0012	0,5935
TTP_original_glcm_DifferenceVariance	N=70 424.27 (102.66-781.41)	N=86 308.36 (130.38-603.81)	0,4383	0,5936
TMAX_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 2.3 (0.89-14.34)	N=86 8.68 (2.08-30.79)	0,0014	0,5943
rBF_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 0.23 (0.07-1.01)	N=86 0.54 (0.16-2.6)	0,037	0,5953
K2_original_firstorder_MeanAbsoluteDeviation	N=70 2999.53 (1652.15-6326.11)	N=86 4658.45 (2612.26-8613.81)	0,0443	0,5956
TTP_original_glszm_LowGrayLevelZoneEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0,1216	0,5965
rBV_corrected_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0.02 (0.01-0.04)	N=86 0.02 (0.01-0.05)	0,99	0,5981
T1_original_firstorder_Median	N=70 2378.12 (1962.54-2678.15)	N=86 2495.8 (2153.79-2770.41)	0,1128	0,5982
tMIP_original_glszm_LargeAreaEmphasis	N=70 1.71 (1.51-2.17)	N=86 2.05 (1.57-3.1)	0,0275	0,5987
rBF_original_gldm_LowGrayLevelEmphasis	N=70 0.01 (0.01-0.02)	N=86 0.01 (0.01-0.03)	0,8293	0,599
K2_original_glrlm_RunVariance	N=70 0.01 (0.01-0.02)	N=86 0.01 (0.01-0.02)	0,033	0,6031
rBV_corrected_original_glrlm_LowGrayLevelRunEmphasis	N=70 0.01 (0.01-0.02)	N=86 0.01 (0.01-0.02)	0,2002	0,6036
rBF_original_glrlm_LongRunEmphasis	N=70 1.17 (1.14-1.23)	N=86 1.18 (1.14-1.23)	0,8544	0,6043
rBV_corrected_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 7125.33 (4005.6-10107.39)	N=86 11884.36 (8754.35-21238.4)	0	0,6065
rBV_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 1.36 (0.3-4.43)	N=86 3.43 (0.97-16.79)	0,0022	0,6078
rBV_original_glcm_MaximumProbability	N=70 0.01 (0.01-0.02)	N=86 0.01 (0.01-0.02)	0,0344	0,6097
rBV_original_glrlm_LongRunEmphasis	N=70 1.23 (1.17-1.31)	N=86 1.25 (1.19-1.3)	0,2389	0,6102
K2_original_glcm_JointEnergy	N=70 0 (0-0)	N=86 0 (0-0)	0	0,6123
K2_original_glcm_Idn	N=70 0.97 (0.96-0.99)	N=86 0.98 (0.96-0.98)	0,4468	0,6132
T1_original_firstorder_Kurtosis	N=70 4.6 (3.42-7.1)	N=86 6.02 (3.97-10.37)	0,0081	0,6137
TMAX_original_gldm_GrayLevelVariance	N=70 1342.6 (491.78-2104.29)	N=86 920.29 (414.06-1876.68)	0,1016	0,6143
T1_original_glcm_ClusterProminence	N=70 48389294.25 (21136818.92-100744333.19)	N=86 114751922.03 (42579678.23-332989624.28)	0	0,6146

TMAX_original_firstorder_Variance	N=70 838348.91 (307548.54-1314833.33)	N=86 574899.77 (258960.93-1173223.95)	0,1008	0,6148
T1_original_firstorder_10Percentile	N=70 1337.76 (937.04-1734.08)	N=86 1414.77 (895.05-1712.04)	0,7296	0,6171
rBV_original_glrlm_ShortRunEmphasis	N=70 0.96 (0.95-0.97)	N=86 0.96 (0.95-0.97)	0,3056	0,6178
TTP_original_gldm_LowGrayLevelEmphasis	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,9078	0,6185
tMIP_original_glcm_MaximumProbability	N=70 0 (0-0.01)	N=86 0 (0-0)	0,0559	0,6194
TMAX_original_ngtdm_Busyness	N=70 0.29 (0.11-0.57)	N=86 0.72 (0.33-1.12)	0	0,6244
T1_original_glszm_LargeAreaEmphasis	N=70 1.97 (1.73-2.41)	N=86 2.29 (1.95-2.82)	0,0014	0,6252
rBV_original_glrlm_RunVariance	N=70 0.1 (0.07-0.15)	N=86 0.11 (0.08-0.16)	0,1619	0,6288
tMIP_original_glrlm_LongRunEmphasis	N=70 1.04 (1.03-1.05)	N=86 1.04 (1.03-1.06)	0,3023	0,6292
TMAX_original_glrlm_GrayLevelVariance	N=70 1385.41 (503.24-2153.62)	N=86 973.84 (454.23-1922.05)	0,1199	0,6298
tMIP_original_glrlm_RunVariance	N=70 0.02 (0.01-0.02)	N=86 0.02 (0.01-0.03)	0,1435	0,6306
K2_original_firstorder_Maximum	N=70 37049.43 (20365.59-51231.52)	N=86 42198.41 (26605.29-73214.36)	0,0602	0,6317
FLAIR_original_firstorder_10Percentile	N=70 2924.51 (2712.83-3178.32)	N=86 2916.28 (2811.96-3107.31)	0,8796	0,6326
TMAX_original_glszm_HighGrayLevelZoneEmphasis	N=70 3513.65 (1831.74-5308.49)	N=86 3205.81 (1835.48-4956.65)	0,5018	0,6351
K2_original_glrlm_GrayLevelVariance	N=70 97075.85 (18358.98-331652.3)	N=86 238637.84 (59456.69-687377.86)	0,0251	0,6353
rBV_corrected_original_glcm_MaximumProbability	N=70 0.01 (0.01-0.01)	N=86 0.01 (0-0.01)	0,0003	0,641
rBV_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0.03 (0.01-0.04)	N=86 0.03 (0.02-0.04)	0,6292	0,6412
T1_original_glcm_Idmn	N=70 0.99 (0.98-1)	N=86 0.99 (0.99-1)	0,0016	0,6473
T1_original_firstorder_Mean	N=70 2428.46 (1991.85-2712.62)	N=86 2557.58 (2255.65-2890.62)	0,0169	0,6478
tMIP_original_firstorder_Range	N=70 29804.03 (15090.01-50100.64)	N=86 33061.12 (16195.04-60611.9)	0,1718	0,6497
MTT_original_glcm_JointEnergy	N=70 0 (0-0)	N=86 0 (0-0)	0,006	0,6502
tMIP_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0.01 (0-0.01)	N=86 0.01 (0-0.01)	0,8965	0,6504
T1_original_ngtdm_Contrast	N=70 0.88 (0.61-1.4)	N=86 0.54 (0.41-0.87)	0,0001	0,6505
K2_original_glcm_DifferenceAverage	N=70 124.84 (67.62-191)	N=86 167.93 (97.09-265.57)	0,0592	0,6514
K2_original_glcm_DifferenceVariance	N=70 67759.13 (13455.07-276597.07)	N=86 157287.87 (45988.88-482336.69)	0,0177	0,6545
rBV_original_firstorder_Minimum	N=70 -158.33 (-208.54--117.82)	N=86 -156.45 (-216.92--104.76)	0,9702	0,6548
TMAX_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 1.61 (0.6-3.57)	N=86 2 (0.82-3.9)	0,495	0,6559
tMIP_original_glcm_Idmn	N=70 0.99 (0.98-0.99)	N=86 0.99 (0.98-0.99)	0,5201	0,6609
rBV_corrected_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0.01 (0.01-0.02)	N=86 0.01 (0.01-0.02)	0,1525	0,6657
rBV_corrected_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0.29 (0.11-0.71)	N=86 0.37 (0.17-0.96)	0,1359	0,6669

K2_original_glszm_SmallAreaEmphasis	N=70 0.94 (0.93-0.96)	N=86 0.94 (0.92-0.95)	0,1173	0,6687
FLAIR_original_firstorder_Median	N=70 3959.44 (3775.07-4120.34)	N=86 3936.86 (3713.36-4113.28)	0,5306	0,6691
TTP_original_glszm_HighGrayLevelZoneEmphasis	N=70 8078.33 (5791.32-9449.16)	N=86 8020.59 (7461.1-9063.11)	0,3856	0,6691
tMIP_original_ngtdm_Strength	N=70 482.21 (193.93-964.9)	N=86 415.92 (103.8-1031.22)	0,2375	0,6702
tMIP_original_firstorder_Maximum	N=70 29584.41 (14810.19-49733.67)	N=86 32729.18 (15992.37-60294.96)	0,1832	0,6707
rBF_original_firstorder_Uniformity	N=70 0.02 (0.02-0.03)	N=86 0.02 (0.01-0.02)	0	0,6715
rBF_original_glcm_MaximumProbability	N=70 0.01 (0-0.01)	N=86 0 (0-0.01)	0,0004	0,6734
TTP_original_glcm_ClusterProminence	N=70 25173173.83 (2596322.45-53318778.27)	N=86 16252761.3 (4611649.98-53164313.68)	0,9333	0,6736
tMIP_original_glszm_LowGrayLevelZoneEmphasis	N=70 0 (0-0)	N=86 0 (0-0)	0,005	0,6764
K2_original_firstorder_Skewness	N=70 -4.15 (SD =4.08)	N=86 -4.75 (SD =3.25)	0,3195	0,6779
MTT_original_glszm_LargeAreaEmphasis	N=70 2.21 (1.72-3.22)	N=86 3.45 (2.42-7.54)	0	0,6781
TTP_original_glrlm_LowGrayLevelRunEmphasis	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,8628	0,6801
K2_original_gldm_SmallDependenceEmphasis	N=70 0.88 (0.85-0.92)	N=86 0.87 (0.84-0.89)	0,0278	0,6837
K2_original_glrlm_LowGrayLevelRunEmphasis	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,3551	0,6906
rBV_original_glcm_Idm	N=70 0.17 (0.14-0.2)	N=86 0.16 (0.13-0.19)	0,2717	0,6982
K2_original_firstorder_Variance	N=70 62135806.58 (11386011.86-207893845.43)	N=86 148644107.67 (36575031.38-439916728.29)	0,0278	0,6999
K2_original_gldm_GrayLevelVariance	N=70 99416.28 (18219.24-332634.93)	N=86 237831.43 (58521.76-703863.51)	0,0278	0,6999
TTP_original_glcm_JointAverage	N=70 83.5 (70.55-88.31)	N=86 83.11 (81.05-85.82)	0,6573	0,7003
TTP_original_glcm_SumAverage	N=70 167 (141.09-176.62)	N=86 166.22 (162.09-171.63)	0,6573	0,7003
K2_original_glcm_ClusterTendency	N=70 168212.96 (30220.62-608401.87)	N=86 459609.97 (116051.63-1252013.12)	0,0182	0,7018
T1_original_firstorder_90Percentile	N=70 3562.19 (3186.27-3786.52)	N=86 3725.05 (3392.1-4053.98)	0,0023	0,7066
TTP_original_firstorder_InterquartileRange	N=70 256.83 (173.02-400.29)	N=86 208.81 (144.22-276.76)	0,0214	0,7071
tMIP_original_gldm_DependenceNonUniformityNormalized	N=70 0.65 (0.59-0.72)	N=86 0.67 (0.59-0.73)	0,777	0,7079
TTP_original_firstorder_RobustMeanAbsoluteDeviation	N=70 117.2 (72.38-191.85)	N=86 93.47 (62.04-132.98)	0,024	0,7083
tMIP_original_glrlm_GrayLevelNonUniformityNormalized	N=70 0.01 (0-0.01)	N=86 0 (0-0.01)	0	0,7099
T1_original_firstorder_RootMeanSquared	N=70 2607.39 (2212.52-2869.03)	N=86 2747.96 (2438.85-3058.15)	0,0047	0,7109
rBV_corrected_original_gldm_LowGrayLevelEmphasis	N=70 0.01 (0.01-0.03)	N=86 0.01 (0.01-0.03)	0,3423	0,711
TMAX_original_glcm_JointAverage	N=70 30.2 (24.01-40.91)	N=86 27.31 (22.26-38.42)	0,1965	0,7187
TMAX_original_glcm_SumAverage	N=70 60.39 (48.03-81.82)	N=86 54.62 (44.52-76.84)	0,1965	0,7187

MTT_original_glszm_LargeAreaHighGrayLevelEmphasis	N=70 8504.2 (6603.1-11900.94)	N=86 9120.9 (7821.14-11987.05)	0,0775	0,72
rBV_original_gldm_DependenceNonUniformityNormalized	N=70 0.25 (0.21-0.32)	N=86 0.26 (0.22-0.29)	0,582	0,7202
T1_original_firstorder_InterquartileRange	N=70 1112.53 (920.78-1312.24)	N=86 1166.28 (1054.07-1381.32)	0,055	0,7204
K2_original_glcm_SumSquares	N=70 60151.25 (13478.82-233010.75)	N=86 172872.09 (43332.16-480151.43)	0,0244	0,7204
rBF_original_glrlm_LowGrayLevelRunEmphasis	N=70 0.01 (0.01-0.02)	N=86 0.01 (0.01-0.02)	0,6343	0,7207
rBF_original_glcm_JointEnergy	N=70 0 (0-0)	N=86 0 (0-0)	0	0,7214
T1_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
FLAIR_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
K2_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
MTT_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
rBF_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
rBV_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
rBV_corrected_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
TMAX_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
tMIP_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
TTP_original_shape_MajorAxisLength	N=70 59.27 (SD =24.52)	N=86 69.22 (SD =18.69)	0,006	0,722
MTT_original_glcm_SumEntropy	N=70 7.39 (7.03-7.73)	N=86 7.48 (7.26-7.69)	0,2686	0,7232
K2_original_glrlm_ShortRunEmphasis	N=70 0.99 (0.99-1)	N=86 0.99 (0.99-1)	0,0166	0,7252
K2_original_firstorder_Uniformity	N=70 0 (0-0.01)	N=86 0 (0-0)	0,0136	0,7266
T1_original_firstorder_RobustMeanAbsoluteDeviation	N=70 469.51 (388.83-558.54)	N=86 499.97 (449.83-578.95)	0,0559	0,7272
TTP_original_glcm_Autocorrelation	N=70 7024.22 (5196.89-7995.14)	N=86 6996.75 (6665.47-7565.31)	0,5992	0,7317
K2_original_gldm_LargeDependenceLowGrayLevelEmphasis	N=70 0 (0-0.16)	N=86 0.02 (0-0.3)	0,4928	0,7373
T1_original_glcm_JointAverage	N=70 119.97 (102.51-139.41)	N=86 140.55 (110.54-171.42)	0,0016	0,7385
T1_original_glcm_SumAverage	N=70 239.95 (205.01-278.83)	N=86 281.1 (221.08-342.84)	0,0016	0,7385
K2_original_firstorder_Median	N=70 303.47 (34.45-605.46)	N=86 437.75 (-71.74-906.2)	0,304	0,7398
K2_original_glszm_ZonePercentage	N=70 0.91 (0.89-0.94)	N=86 0.9 (0.88-0.92)	0,024	0,7419
MTT_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 0.06 (0.02-0.49)	N=86 0.25 (0.04-1.85)	0,0169	0,7444
tMIP_original_glcm_JointEnergy	N=70 0 (0-0)	N=86 0 (0-0)	0	0,745
tMIP_original_glcm_Idm	N=70 0.04 (0.03-0.05)	N=86 0.04 (0.03-0.05)	0,1685	0,7538
K2_original_ngtdm_Strength	N=70 2193.22 (669.82-4000.1)	N=86 2814.57 (1214.24-4885.96)	0,3141	0,7545

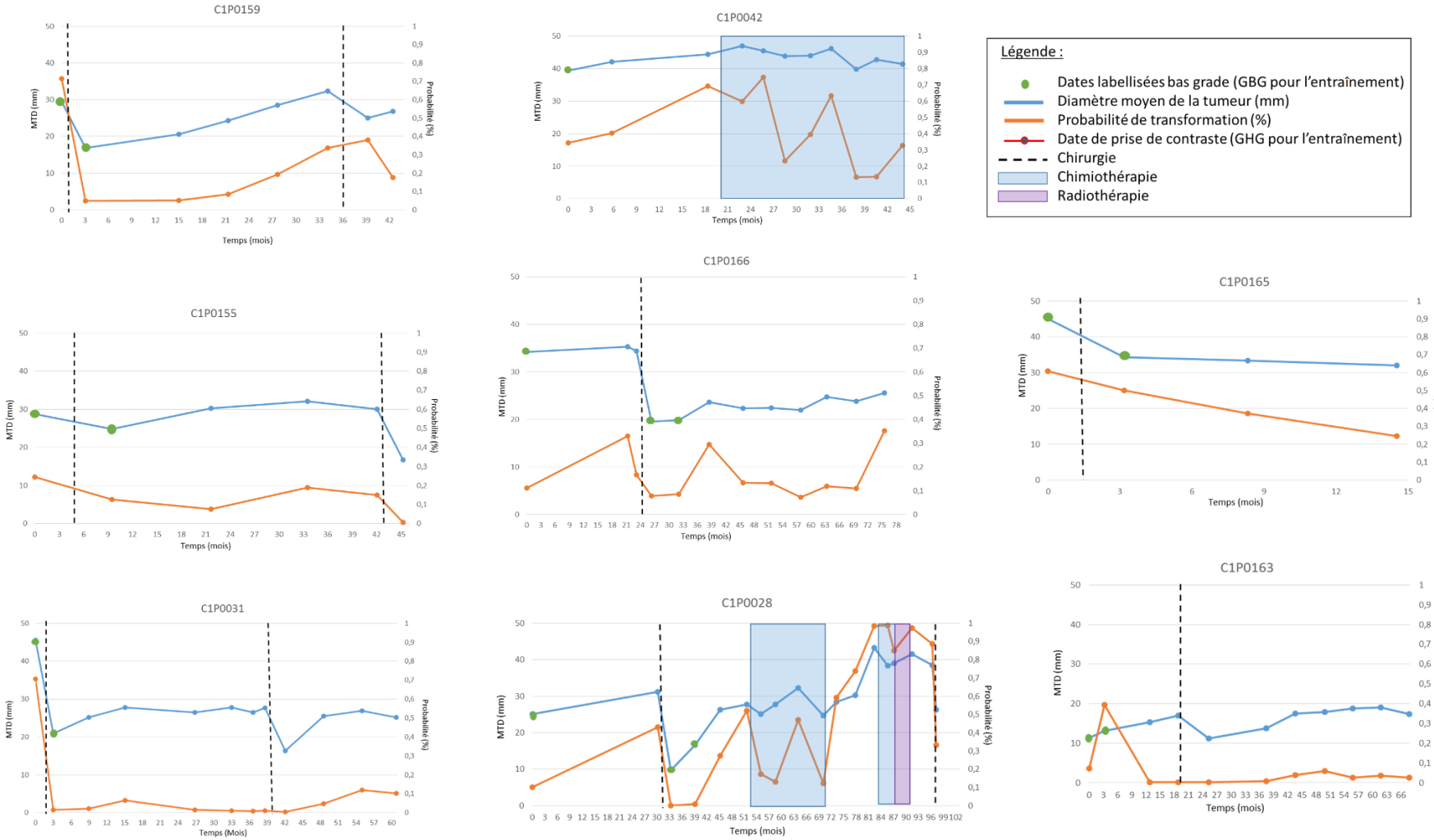
rBV_original_glcm_JointEnergy	N=70 0 (0-0)	N=86 0 (0-0)	0,0004	0,756
tMIP_original_firstorder_Uniformity	N=70 0.01 (0-0.01)	N=86 0 (0-0.01)	0,0001	0,7578
TMAX_original_glszm_GrayLevelVariance	N=70 1897.99 (862.49-2676.83)	N=86 1705.86 (884.53-2357.83)	0,3954	0,7618
TTP_original_glrlm_HighGrayLevelRunEmphasis	N=70 7404.53 (5778.99-8723.03)	N=86 7252.42 (6824.5-8066.63)	0,9929	0,764
rBV_original_glcm_Idn	N=70 0.95 (0.93-0.95)	N=86 0.94 (0.93-0.95)	0,7552	0,7648
TMAX_original_glcm_DifferenceVariance	N=70 1030.01 (428.16-1529.61)	N=86 890.16 (407.18-1564.81)	0,5893	0,7677
tMIP_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 0.12 (0.02-0.45)	N=86 0.29 (0.08-1.24)	0,005	0,7694
rBV_corrected_original_gldm_LargeDependenceEmphasis	N=70 7.17 (5.84-9.32)	N=86 7.48 (5.98-10.79)	0,4362	0,7764
rBF_original_glszm_LargeAreaEmphasis	N=70 8.37 (5.14-19.15)	N=86 12.16 (6.19-27)	0,0532	0,777
T1_original_glcm_DifferenceAverage	N=70 22.3 (18.33-26.33)	N=86 23.32 (19.12-28.58)	0,1856	0,7815
K2_original_firstorder_Kurtosis	N=70 30.32 (13.34-88.3)	N=86 34.66 (18.22-76.99)	0,6017	0,7836
K2_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0 (0-0)	N=86 0 (0-0.01)	0,3107	0,7953
K2_original_glcm_Contrast	N=70 86539.16 (23389.06-353658.01)	N=86 214575.49 (57166.98-617862.64)	0,024	0,7999
T1_original_firstorder_MeanAbsoluteDeviation	N=70 693.73 (567.61-775.71)	N=86 729.61 (661.65-877.53)	0,0076	0,8059
TTP_original_gldm_HighGrayLevelEmphasis	N=70 7364.71 (5724.46-8606.42)	N=86 7170 (6776.11-7960.99)	0,9815	0,8084
TMAX_original_firstorder_Kurtosis	N=70 14.88 (7.03-28.74)	N=86 25.2 (12.01-46.35)	0,0032	0,8091
K2_original_glcm_MaximumProbability	N=70 0 (0-0)	N=86 0 (0-0)	0,011	0,8145
rBV_original_firstorder_Uniformity	N=70 0.03 (0.03-0.04)	N=86 0.03 (0.02-0.03)	0,0022	0,8218
K2_original_glszm_LargeAreaEmphasis	N=70 1.41 (1.25-1.78)	N=86 1.5 (1.36-2.67)	0,0082	0,8242
K2_original_gldm_LowGrayLevelEmphasis	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,4426	0,8244
TMAX_original_firstorder_10Percentile	N=70 -56.04 (-129.69-27.01)	N=86 -23.26 (-138.94-35.25)	0,821	0,8267
K2_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 0 (0-0.07)	N=86 0.01 (0-0.54)	0,5674	0,8306
T1_original_firstorder_Minimum	N=70 -710.65 (-1384.34--211.93)	N=86 -1021.7 (-1670.84--378.59)	0,1077	0,8343
TTP_original_ngtdm_Complexity	N=70 133824.44 (33743.22-272991.5)	N=86 159931.27 (76014.2-335909.96)	0,0836	0,8353
TTP_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,7633	0,8377
TMAX_original_firstorder_Range	N=70 6843.24 (6154.91-7639.3)	N=86 7157.94 (5847.45-7946.72)	0,2375	0,8379
FLAIR_original_glszm_LargeAreaEmphasis	N=70 2.03 (1.56-3.12)	N=86 2.6 (1.98-3.66)	0,0023	0,8405
TMAX_original_glrlm_LongRunLowGrayLevelEmphasis	N=70 0.11 (0.06-0.16)	N=86 0.1 (0.05-0.18)	0,7444	0,848
rBV_corrected_original_glszm_LargeAreaEmphasis	N=70 9.75 (5.98-21.89)	N=86 14.86 (7.16-47.13)	0,0519	0,8497
FLAIR_original_glcm_InverseVariance	N=70 0.06 (0.04-0.07)	N=86 0.07 (0.06-0.09)	0	0,8529

K2_original_glszm_SizeZoneNonUniformityNormalized	N=70 0.86 (0.83-0.91)	N=86 0.85 (0.82-0.89)	0,125	0,8582
K2_original_firstorder_Mean	N=70 -484.8 (-1672.22-432.07)	N=86 -416.6 (-3470.1-586.9)	0,731	0,8695
T1_original_firstorder_Range	N=70 9418.29 (5387.64-12001.7)	N=86 11605.69 (8863.17-14027.93)	0,0001	0,8717
rBV_corrected_original_glszm_LargeAreaLowGrayLevelEmphasis	N=70 0.35 (0.11-1.14)	N=86 0.69 (0.19-3.76)	0,04	0,8744
rBF_original_glrlm_RunPercentage	N=70 0.96 (0.95-0.96)	N=86 0.96 (0.95-0.96)	0,9815	0,877
TMAX_original_glcm_Autocorrelation	N=70 1312.67 (792.37-2588.15)	N=86 1125.82 (637.49-2110.04)	0,294	0,8774
T1_original_firstorder_Maximum	N=70 8227.03 (4983.84-10905.22)	N=86 10447.01 (8443.02-12889.3)	0,0001	0,8789
MTT_original_glszm_LowGrayLevelZoneEmphasis	N=70 0 (0-0.01)	N=86 0 (0-0.01)	0,1053	0,8832
K2_original_firstorder_10Percentile	N=70 -3487.46 (-6533.35--1771.04)	N=86 -4142.89 (-6845.44--2245.33)	0,2724	0,8841
rBV_original_glcm_Idmn	N=70 0.99 (0.98-0.99)	N=86 0.99 (0.98-0.99)	0,9787	0,887
TMAX_original_glrlm_LongRunHighGrayLevelEmphasis	N=70 2577.56 (1494.77-4291.05)	N=86 2165.44 (1205.34-3743.53)	0,2907	0,891
tMIP_original_gldm_SmallDependenceEmphasis	N=70 0.83 (0.79-0.87)	N=86 0.84 (0.79-0.88)	0,7907	0,8923
TTP_original_firstorder_10Percentile	N=70 1783.52 (1664.77-1858.47)	N=86 1813.1 (1719.71-1871.27)	0,3817	0,8974
T1_original_glcm_InverseVariance	N=70 0.06 (SD =0.02)	N=86 0.06 (SD =0.02)	0,3357	0,9075
MTT_original_firstorder_Maximum	N=70 6134.67 (SD =1887.93)	N=86 7072.85 (SD =1877.42)	0,0024	0,9085
tMIP_original_glcm_Contrast	N=70 14351.46 (5223.33-51051.81)	N=86 29276.07 (6919.48-91005.95)	0,0728	0,9089
T1_original_glcm_Id	N=70 0.12 (0.11-0.14)	N=86 0.13 (0.11-0.14)	0,3934	0,9114
TTP_original_glrlm_ShortRunHighGrayLevelEmphasis	N=70 7085.27 (5297.88-8296.47)	N=86 6721.93 (6247.14-7637.21)	0,6859	0,9169
K2_original_gldm_LargeDependenceEmphasis	N=70 1.71 (1.43-2.17)	N=86 1.85 (1.64-2.42)	0,0115	0,9206
K2_original_glrlm_RunLengthNonUniformityNormalized	N=70 0.99 (0.98-0.99)	N=86 0.98 (0.98-0.99)	0,0133	0,9268
tMIP_original_glcm_ClusterShade	N=70 28345901.97 (4203029.76-168082116.82)	N=86 75594482.26 (3796972.7-374459670.87)	0,1416	0,9327
T1_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
FLAIR_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
K2_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
MTT_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
rBF_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
rBV_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
rBV_corrected_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
TMAX_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387

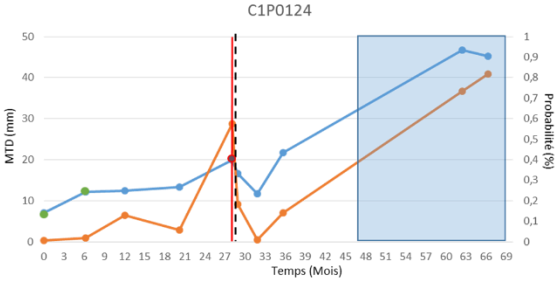
tMIP_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
TTP_original_shape_Sphericity	N=70 0.42 (0.35-0.55)	N=86 0.46 (0.38-0.53)	0,2375	0,9387
TMAX_original_firstorder_Maximum	N=70 6641.33 (5778.07-7402.69)	N=86 6911.36 (5702.18-7705.77)	0,3405	0,94
TTP_original_glcm_Correlation	N=70 0.28 (SD =0.09)	N=86 0.31 (SD =0.1)	0,0856	0,941
T1_original_glcm_Idm	N=70 0.06 (SD =0.02)	N=86 0.07 (SD =0.02)	0,2613	0,9447
K2_original_glcm_ClusterShade	N=70 -422473616.04 (-3707635223.01--6445776.28)	N=86 -1003838387.11 (-8642503564.65--152175199.07)	0,0195	0,9448
TMAX_original_glcm_Contrast	N=70 1510.75 (597.13-2448.24)	N=86 1214.47 (504.14-2312.47)	0,221	0,945
TMAX_original_glcm_SumSquares	N=70 1196.12 (396.36-2002.99)	N=86 889.29 (356.64-1858.1)	0,2361	0,9501
TMAX_original_glcm_ClusterTendency	N=70 3294.47 (968.53-5763.92)	N=86 2418.6 (943.08-5075.12)	0,2549	0,9537
tMIP_original_glcm_Id	N=70 0.08 (0.06-0.1)	N=86 0.07 (0.05-0.09)	0,119	0,9647
tMIP_original_glszm_ZonePercentage	N=70 0.88 (0.85-0.9)	N=86 0.88 (0.85-0.91)	0,874	0,9657
rBF_original_glrlm_ShortRunLowGrayLevelEmphasis	N=70 0.01 (0.01-0.02)	N=86 0.01 (0.01-0.02)	0,5341	0,9671
K2_original_glcm_Idm	N=70 0.03 (0.02-0.03)	N=86 0.03 (0.02-0.04)	0,2156	0,9747
K2_original_glrlm_RunPercentage	N=70 0.99 (0.99-0.99)	N=86 0.99 (0.99-0.99)	0,0223	0,9783
MTT_original_firstorder_Range	N=70 6636.46 (SD =1962.63)	N=86 7588.35 (SD =1911.16)	0,0027	0,9809
K2_original_glcm_Id	N=70 0.06 (0.05-0.08)	N=86 0.06 (0.05-0.08)	0,357	0,9838
K2_original_glcm_ClusterProminence	N=70 3482425567757.58 (34839050698.25-32056312681804.5)	N=86 10221841001391.7 (2147360241741.41-100521815574969)	0,0167	0,9875
TTP_original_glszm_SmallAreaHighGrayLevelEmphasis	N=70 6422.52 (4563.67-8064.2)	N=86 6078.51 (5360.17-7626.64)	0,8182	0,9881
tMIP_original_glcm_DifferenceVariance	N=70 8883.55 (2971.08-34599.99)	N=86 22418.57 (3638.15-64595.74)	0,0793	0,9942

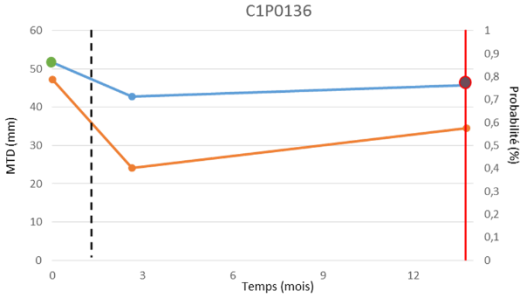
Annexe 2- Courbe d'évolution du diamètre moyen tumoral

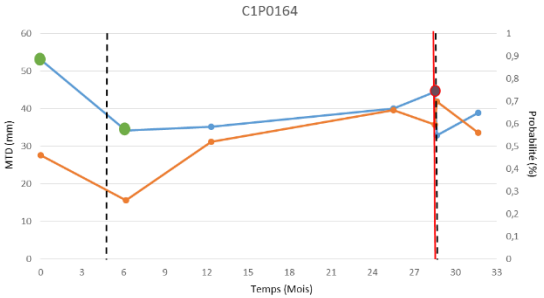
Courbe d'évolution du diamètre moyen tumoral, en bleu (MTD(mm)) et de la probabilité de transformation (entre 0 et 1)) issue du SVM en orange au cours du temps en mois

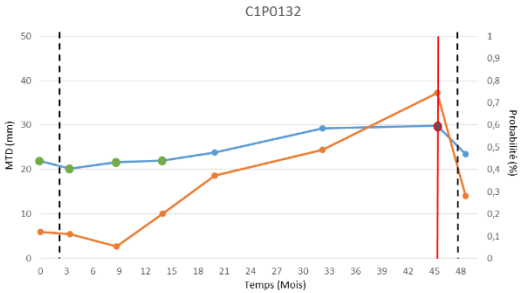


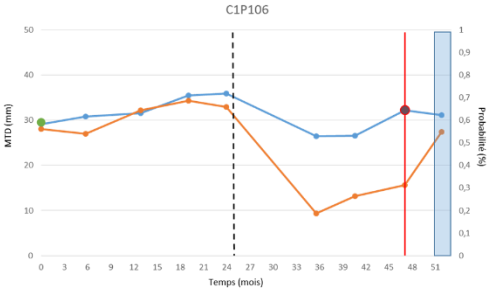
Annexes

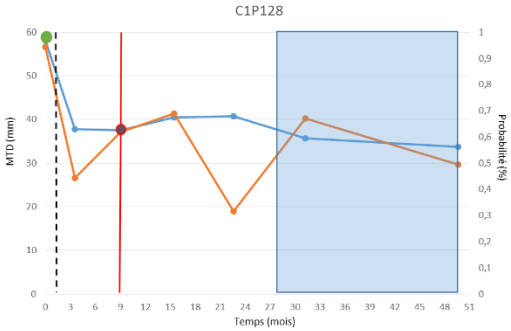

C1P0124
MTD (mm)
Probabilité (%)
Temps (Mois)

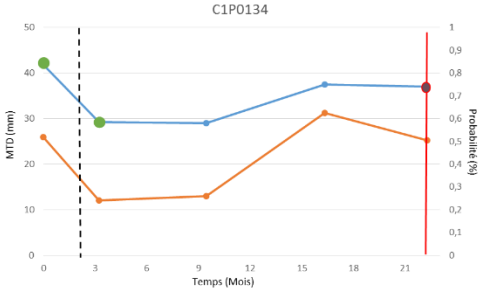

C1P0136
MTD (mm)
Probabilité (%)
Temps (mois)

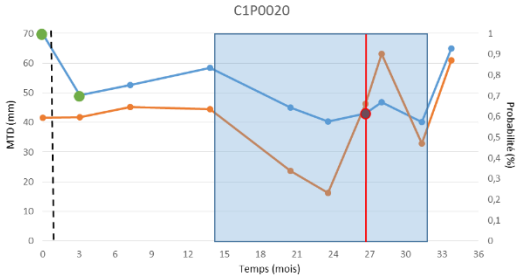

C1P0164
MTD (mm)
Probabilité (%)
Temps (Mois)

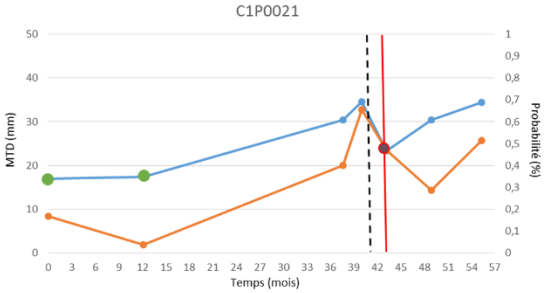

C1P0132
MTD (mm)
Probabilité (%)
Temps (Mois)


C1P106
MTD (mm)
Probabilité (%)
Temps (mois)

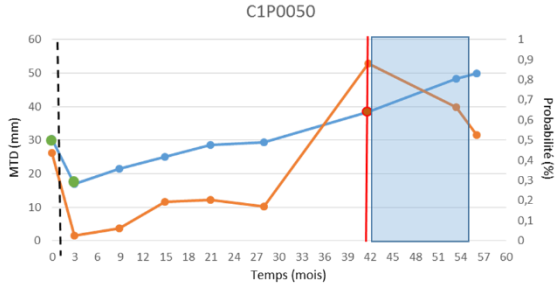

C1P128
MTD (mm)
Probabilité (%)
Temps (mois)

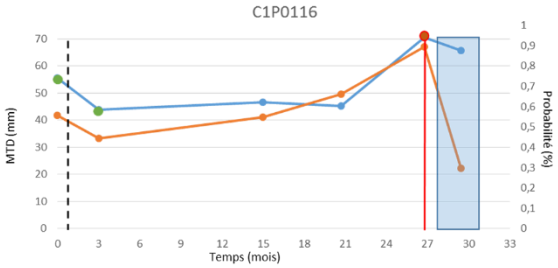

C1P0134
MTD (mm)
Probabilité (%)
Temps (Mois)

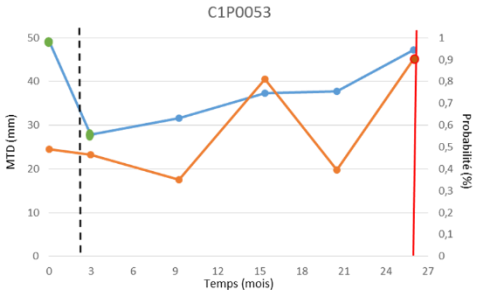

C1P0020
MTD (mm)
Probabilité (%)
Temps (mois)

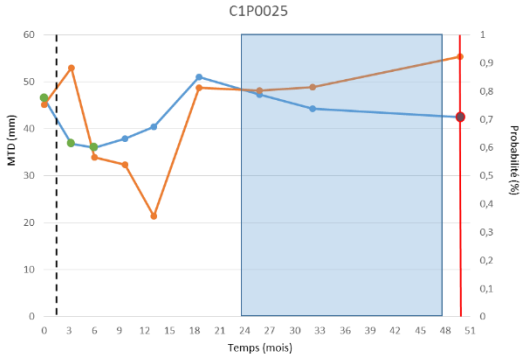

C1P0021
MTD (mm)
Probabilité (%)
Temps (mois)

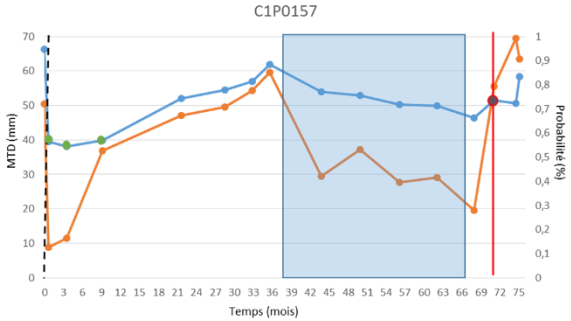
Annexes

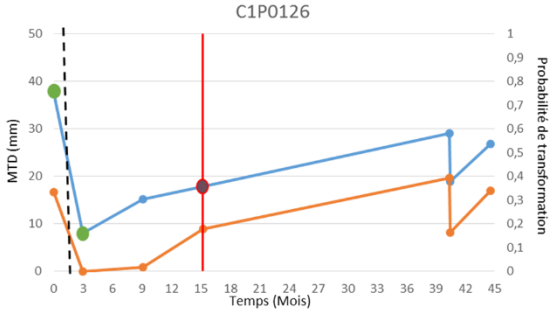

C1P0050
MTD (mm)
Probabilité (%)
Temps (mois)

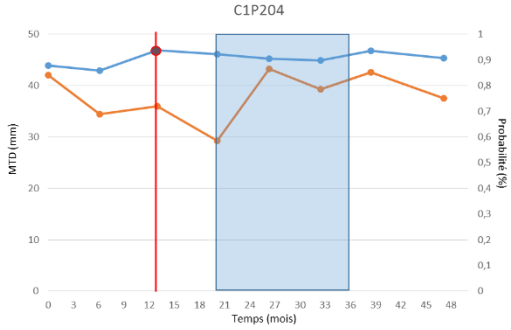

C1P0116
MTD (mm)
Probabilité (%)
Temps (mois)

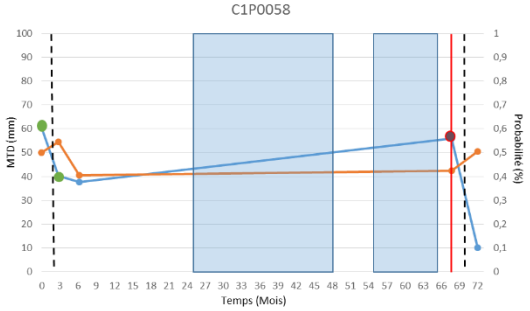

C1P0053
MTD (mm)
Probabilité (%)
Temps (mois)

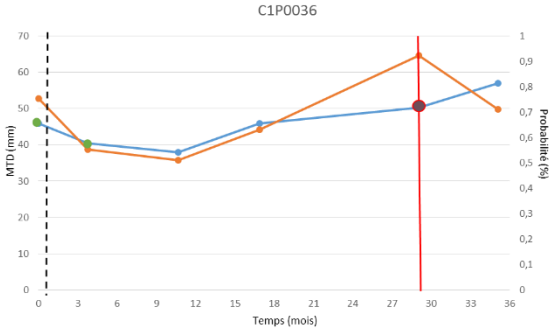

C1P0025
MTD (mm)
Probabilité (%)
Temps (mois)


C1P0157
MTD (mm)
Probabilité (%)
Temps (mois)

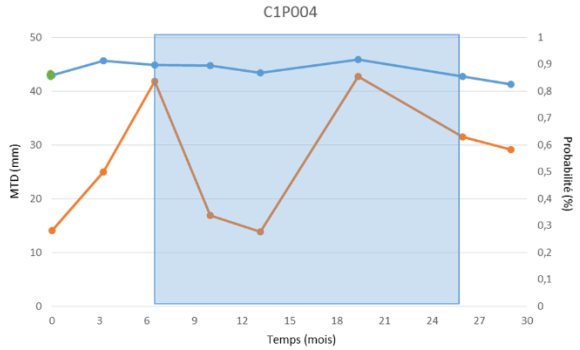

C1P0126
MTD (mm)
Probabilité de transformation
Temps (Mois)

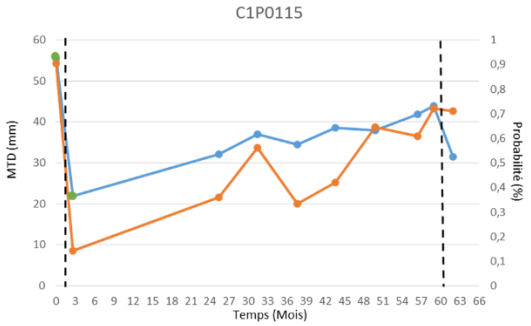

C1P204
MTD (mm)
Probabilité (%)
Temps (mois)

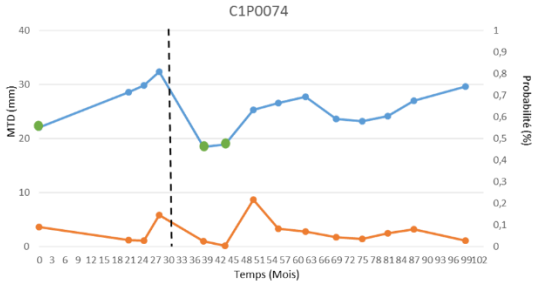

C1P0058
MTD (mm)
Probabilité (%)
Temps (Mois)

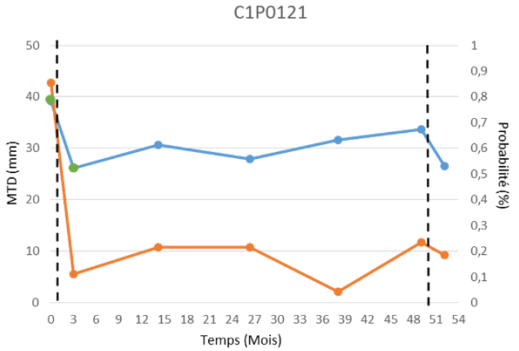

C1P0036
MTD (mm)
Probabilité (%)
Temps (mois)

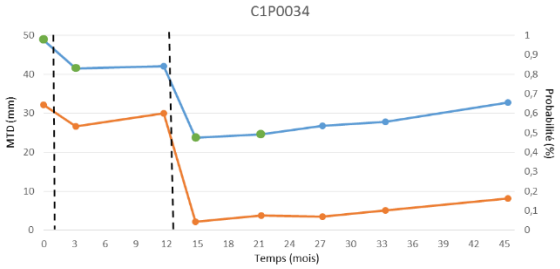
Annexes

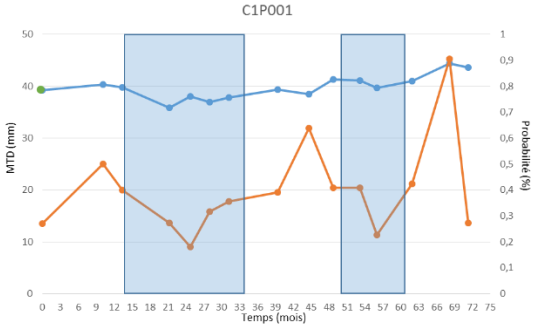

C1P004
MTD (mm)
Probabilité (%)
Temps (mois)

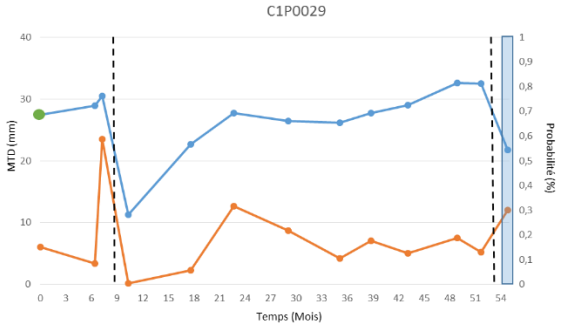

C1P0115
MTD (mm)
Probabilité (%)
Temps (Mois)

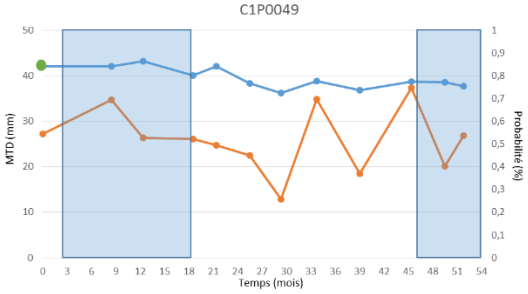

C1P0074
MTD (mm)
Probabilité (%)
Temps (Mois)

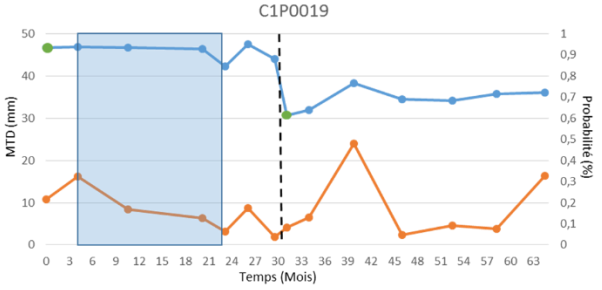

C1P0121
MTD (mm)
Probabilité (%)
Temps (Mois)


C1P0034
MTD (mm)
Probabilité (%)
Temps (mois)


C1P001
MTD (mm)
Probabilité (%)
Temps (mois)


C1P0029
MTD (mm)
Probabilité (%)
Temps (Mois)


C1P0049
MTD (mm)
Probabilité (%)
Temps (mois)


C1P0019
MTD (mm)
Probabilité (%)
Temps (Mois)

Annexes