



HAL
open science

Évaluation de l'Hétérogénéité Tumorale en Imagerie Paramétrique

Charly Girot

► **To cite this version:**

Charly Girot. Évaluation de l'Hétérogénéité Tumorale en Imagerie Paramétrique. Cancer. Université Paris-Saclay, 2022. Français. NNT : 2022UPASL006 . tel-04010755

HAL Id: tel-04010755

<https://theses.hal.science/tel-04010755>

Submitted on 2 Mar 2023

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

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

Évaluation de l'Hétérogénéité Tumorale en Imagerie Paramétrique

*Evaluation of Tumor Heterogeneity with Parametric
Imaging*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 582 : cancérologie : biologie - médecine - santé
(CBMS)

Spécialité de doctorat : Recherche clinique, Innovation technologique,
Santé publique

Graduate School : Sciences de la vie et santé, Référent : Faculté de
médecine

Thèse préparée dans l'unité de recherche **BioMaps** (Université Paris-Saclay, CEA, CNRS, Inserm), sous la direction de **Stéphanie PITRE-CHAMPAGNAT**, Chargée de Recherche, CNRS, université Paris-Saclay, le co-encadrement de **Andreas VOLK**, Chargé de Recherche, Inserm, université Paris-Saclay

Thèse soutenue à Paris-Saclay, le 21 janvier 2022, par

Charly GIROT

Composition du jury

Charles-André CUENOD Hopital Européen Georges Pompidou Professeur des universités - Praticien Hospitalier , Université de Paris	Président
Benjamin LEMASSON Grenoble Institut des Neurosciences (GIN), Inserm Associate Professor, Université Grenoble Alpes	Rapporteur & Examineur
Arthur VAROQUAUX Centre de Résonance Magnétique Biologique et Médicale (CRMBM) Professeur, Aix Marseille Université	Rapporteur & Examineur
Isabelle BLOCH Laboratoire de Recherche en Informatique (LIP6), CNRS Professeure, Sorbonne Université	Examinatrice
Michèle KIND Praticienne Hospitalière, Institut Bergonié, Bordeaux	Examinatrice
Alexandre VIGNAUD NeuroSpin, CEA Directeur de Recherche, université Paris-Saclay	Examineur

Remerciements

Cette thèse a été effectuée au laboratoire BioMaps, anciennement IR4M, et je remercie les deux directeurs successifs, Luc Darasse et Vincent Lebon pour leur accueil. Par ailleurs, je remercie les cheffes d'équipes successives de l'équipe imagerie en cancérologie, Stéphanie Pitre-Champagnat, Nathalie Lassau et Corine Balleyguier, pour leur accueil, leurs conseils et leur accompagnement tout du long de mes travaux de thèse.

Je tiens ensuite à remercier du fond du cœur Stéphanie Pitre-Champagnat, directrice de thèse et Andréas Volk, co-directeur de thèse, pour leurs conseils précieux au cours de ces années difficiles, tant au niveau scientifique qu'au niveau humain, pour les moments de rigueur et les moments partagés. Je suis convaincu d'avoir tiré le meilleur de cette formation grâce à eux et à leur disponibilité, grâce à leurs encouragements. Leur engagement infailible m'a permis de toujours satisfaire ma curiosité tout en m'évitant l'égarement. Je les remercie enfin pour la patience dont ils ont su faire preuve à mon égard et pour l'équilibre parfait formé par ce duo d'encadrement. Pour cela et pour tout le reste, je les remercie infiniment.

Je remercie l'ensemble des membres du jury pour avoir manifesté leur intérêt pour mes travaux. Ma reconnaissance va ensuite à Benjamin Lemasson et Arthur Varoquaux qui ont accepté d'examiner mon mémoire en tant que rapporteur. Je remercie enfin Isabelle Bloch, Michèle Kind, Alexandre Vignaud et Charles-André Cuenod d'avoir participé à ce jury.

Ce travail se situant à l'interface de plusieurs disciplines, j'ai eu la chance de collaborer avec de nombreuses personnes qui ont toutes su apporter une brique différente et importante à ces travaux. Je tiens à leur exprimer toute ma gratitude.

Je remercie profondément l'équipe de radiologues de Gustave Roussy. François Bidault et Samy Ammari qui m'ont fait l'honneur de partager leurs données cliniques constituant tout un volet de cette thèse. Je les remercie également pour leur expertise clinique, leurs nombreuses interventions et leur bonne humeur quotidienne.

Merci à Claire Pellot-Baraka et Xavier Maitre qui ont également à plusieurs reprises partager des données ou du matériel et qui ont ouvert des discussions scientifiques de grand intérêt. Merci également à Claire pour la relecture du manuscrit et les critiques

précieuses.

Merci à Vincent Lebon, directeur du laboratoire BioMaps et chef de service du Service Hospitalier Frédéric Joliot, Tounes Saidi, Christine Baron et Fabien Rastello, pour leur support indéfectible, dans l'urgence, pour la mise en place du protocole volontaire sains. Cette étude n'aurait pas pu voir le jour sans eux.

Bien que les travaux principaux de ces années de recherche soient rendus dans ce manuscrit, d'autres projets se sont également intégrés à mon expérience.

Pour le projet calibration US, je tiens à remercier tout particulièrement Laurène Jourdain, en mémoire de nos nombreuses heures de manip, de nos allers retours avec le chariot, des fantômes d'agar et surtout des gouttes-gouttes, merci. Je remercie également Isaline Hoferer et Stéphanie, pour avoir pris ou repris la suite.

Merci à Nathalie Lassau pour les nombreuses discussions qui sur le plan scientifique m'ont permis de soulever de nombreuses questions sur mes travaux et de m'intéresser à de nouveaux domaines et qui sur le plan personnel et professionnel m'aura soutenu et orienté.

A la frontière de la recherche et de l'éducation, la vulgarisation de la radiomique n'aurait jamais été possible sans Salma Moalla, radiologue à Gustave Roussy, grâce à qui le poster a pu être primé à deux reprises, aux JFR et au RSNA. Merci d'avoir rendu cette tâche facile, pour le voyage à Chicago et pour nos nombreuses discussions autour d'un café.

Je remercie également Ingrid Leguerney, pour m'avoir supporté en bureau voisin, nos discussions et la découverte du petit-animal. J'exprime également toute mon amitié à Baya Benatsou que je remercie vivement pour les nombreuses discussions, pour son expertise US, sa disponibilité et sa gestion sans failles des projets.

Un très grand merci à Georges Willoquet, pour le support toujours impeccable et pour le sauvetage de ma thèse. J'exprime également toute mon amitié à Marylène Heurtaut pour tous ces instants de soutien, généralement un café à la main, partageant nos envies et nos plans de voyages.

Je remercie également mes voisins de bureau successifs qui ont su me supporter en commençant par Virginie Grand-Perret, puis les stagiaires Inas Sadoun, Roman Kozlov et leur aide pour la calibration US, Victor Le Chevalier pour ses travaux de développement et Benoit Charmettant.

Merci à Stéphanie Franchi-Abella et Catherine Adamsbaum, radiologues à l'hôpital du Kremlin-Bicêtre pour la collaboration en radiologie pédiatrique, qui malgré l'absence de résultats aura permis de consolider notre approche méthodologique.

Enfin, la réalisation de cette thèse n'aurait été ni envisageable ni possible sans le soutien précieux de mes proches.

Tout d'abord mes parents pour lesquels les adjectifs disponibles ne me permettent pas de les remercier suffisamment. Ils m'ont apporté non seulement tout au long de ma thèse, mais plus généralement de mes études, toute l'aide dont j'avais besoin, toute l'aide que l'on peut espérer. Leurs encouragements et leur croyance dans mes capacités, d'autant plus au cours de cette période au contexte difficile, sont inestimables.

Je remercie également ma famille, ma tante, mes grands-parents, les jumeaux, Olivia, Marie. Les nombreux appels de réconfort, le sport, les repas et les retrouvailles sont autant de moments m'ayant permis de surmonter ces années. Je tiens également à les féliciter pour les efforts de compréhension de mon sujet de thèse et les tentatives de restitutions nombreuses et variées. Merci à Jérémy, mon frère, pour avoir partagé différents appartements et différentes chambres avec moi qui nous ont permis de partager des moments privilégiés. Merci également pour avoir passé plusieurs semaines confinées avec moi.

Enfin, que cela soit à Lyon, à Paris, où les deux, il ne m'aurait pas été possible d'en arriver là aujourd'hui sans l'amitié inébranlable de mes amis. Je les remercie énormément pour tout ce qu'ils m'apportent au quotidien, pour tous ces moments de bonheur, pour les colocations, au quotidien ou à Lurange, pour tous ces événements tardifs et ces festivals frappés, pour les débriefs matinaux et les bières d'apéros. Merci.

Sommaire

Remerciements	3
Introduction	13
1 Les biomarqueurs pour la cancérologie	17
1 Les biomarqueurs pour la cancérologie : rôles et définitions	18
1.1 Biomarqueur, des définitions en perpétuelle évolution	20
1.2 Les différents rôles des biomarqueurs : descriptions simples et superposables	21
2 Diversité et méthodes d'acquisitions des biomarqueurs	23
2.1 Les biomarqueurs biochimiques en routine clinique	24
2.2 Les biomarqueurs d'imagerie en routine clinique	27
3 La nouvelle génération de biomarqueurs pour l'évaluation de la réponse thérapeutique	32
4 Processus de validation : la clef de voute des biomarqueurs d'imagerie .	40
5 L'hétérogénéité tumorale comme biomarqueur d'imagerie	46
5.1 Prise en compte de l'hétérogénéité tumorale dans les nouvelles stratégies thérapeutiques	47
5.2 État de l'art de l'évaluation de l'hétérogénéité tumorale en imagerie	49
6 Conclusion	53
2 Vers de Nouveaux Biomarqueurs de l'Hétérogénéité Tumorale en Imagerie Paramétrique : Concept, Développement et Caractérisation	67
1 Concept	68
2 Développement d'un nouveau biomarqueur	68
2.1 Le clustering et l'hétérogénéité des intensités	70
2.2 Les métriques de complexité de la distribution spatiale	77
2.3 D'un ensemble de métriques vers un unique biomarqueur	81
3 Caractérisation numérique du biomarqueur de l'hétérogénéité	85
3.1 Fantômes numériques hétérogènes	86

3.2	Influence de la connectivité sur les métriques d'hétérogénéité . . .	87
3.3	Influence de la résolution spatiale de l'image radiologique	91
3.4	Influence du contraste de l'image radiologique	95
4	Conclusion	100
3	Exploration préclinique des biomarqueurs d'imagerie de l'hétérogénéité	109
1	Angiogenèse : source d'hétérogénéité vasculaire tumorale	109
1.1	Le processus d'angiogenèse et l'hétérogénéité vasculaire tumorale .	110
1.2	Quantifier l'angiogenèse : l'imagerie de contraste, un outil incontournable	111
2	Étude comparative des indices d'hétérogénéité comme biomarqueurs d'imagerie de la réponse thérapeutique sur données préclinique	113
2.1	Description du protocole expérimental	113
2.2	Quantification de l'hétérogénéité tumorale	117
2.3	Résultats	119
2.4	Discussion	135
3	Conclusions et perspectives	140
4	Évaluation de l'intérêt clinique de biomarqueurs de l'hétérogénéité de l'image	147
1	Étude de la répétabilité du H_{index} en IRM de diffusion cérébrale	148
1.1	Enjeux et objectifs	148
1.2	Protocole d'acquisition clinique	148
1.3	Protocole de traitement et d'analyse des données	152
1.4	Résultats	157
1.5	Discussion	166
1.6	Conclusion	176
2	Application du H_{index} en DCE-IRM pour s'affranchir de la variabilité de l'AIF : Première étude rétrospective de faisabilité du H_{index} en cancérologie ORL	177
2.1	Enjeux et objectifs	177
2.2	Protocole d'acquisition clinique	178
2.3	Protocole de traitement et d'analyse des données DCE-IRM . . .	179
2.4	Résultats	183
2.5	Discussion	189
2.6	Conclusion	194
3	Application du H_{index} dans un contexte de radiothérapie de tumeurs cérébrales : Evaluation du potentiel diagnostique pour la différenciation de récurrence locale et de radionécrose en IRM de contraste et de perfusion . .	195
3.1	Enjeux et objectifs	195

3.2	Méthodologie du protocole clinique	196
3.3	Protocole de traitement et d'analyse des données	198
3.4	Résultats	200
3.5	Discussion	205
3.6	Conclusion	207
4	Conclusion	207
	Conclusion	217
	Contributions Scientifiques	221

Acronymes

ADC Apparent Diffusion Coefficient. 33

AIF Fonction d'entrée Artériel (Arterial Input Function). 39

AUC Aire sous la courbe. 34, 49

BI Biomarqueur d'Imagerie. 16, 25–28, 35, 38, 40, 41, 43–45, 47–49, 138

CT Computed Tomography. 26, 27, 43, 48

DCE-IRM IRM dynamique de contraste. 107, 110–113, 115, 134, 138

DSC-IRM IRM de contraste de susceptibilité dynamique. 194

DWI-IRM IRM de diffusion. 145

FDA U.S. Food and Drug Administration. 18, 19, 40

IA Intelligence Artificielle. 22, 50

IRM Imagerie par Résonance Magnétique. 26, 32–35, 43, 48

NIH U.S. National Institute of Health. 18, 19, 40

OMS Organisation Mondiale de la Santé. 18

OS Overall Survival. 19, 20, 41

PFS Progression Free Survival. 19, 20, 41

PPC Paramètre Pharmacocinétique. 177

ROI Région d'Intérêt. 48

SUV Standardized Uptake Value. 27, 30, 33, 35, 39

TEMP Tomographie par Emission Mono-Photonique. 26

TEP Tomographie par Emission de Positons. 16, 20, 30, 33, 39, 41, 48

US Echographie UltraSonore. 26, 32, 43, 48

WHO World Health Organization. 18

Introduction

En cancérologie, la prise en charge personnalisée des patients passe par une meilleure compréhension du développement tumoral mais également des mécanismes d'actions possibles. Ainsi, dans ce domaine peut-être plus qu'ailleurs, les praticiens peuvent s'appuyer sur les avancées d'une recherche pluridisciplinaire couvrant de nombreux champs d'investigations de la recherche fondamentale en biologie, pharmacologie, physique, numérique jusqu'aux investigations cliniques. Du point de vue fondamental, les avancées sont étroitement liées aux progrès technologiques et notamment instrumentaux permettant de caractériser et analyser toujours plus finement les tumeurs et leurs environnements tant à l'échelle moléculaire et génétique qu'à l'échelle tissulaire et macroscopique. A l'ère de la médecine de précision, l'objectif de ces recherches fondamentales est double avec, d'une part, le développement et l'évaluation de nouvelles stratégies thérapeutiques combinant ou non plusieurs actions biologiques, et, d'autre part, le développement d'outils pertinents permettant de guider finement le praticien vers un protocole thérapeutique personnalisé. Ces outils que sont les biomarqueurs visent à refléter différents aspects des tumeurs, de leur composition génétique à leur fonctionnement métabolique et à être employés du diagnostic jusqu'au suivi post-thérapeutique. Néanmoins, l'utilisation de ces biomarqueurs issus de la recherche fondamentale est confrontée à la réalité clinique du terrain avec une mise en pratique d'autant plus laborieuse que leur obtention est conditionnée à des techniques d'acquisitions de plus en plus complexes et des données de plus en plus massives. Ainsi, associé aux développements instrumentaux poussés aux limites, le domaine des sciences numériques est aujourd'hui particulièrement sollicité dans l'exploration des biomarqueurs. C'est notamment le cas de l'intelligence artificielle (IA) de plus en plus largement appliquée en radiologie.

En effet, en raison de la multiplication des examens radiologiques, l'intégration de l'Intelligence Artificielle (IA) dans le flux de travail du radiologue apparaît comme une évidence. Parmi les nombreuses applications à l'interface de ces deux domaines, celle qui s'intéresse à l'analyse et la quantification des images, plus particulièrement encore en cancérologie et qui se situe à l'interface entre ces deux domaines porte un nom : la radiomique. L'idée initiale de la radiomique visait à compléter le jeu de biomarqueurs d'imagerie existants et principalement basés sur des informations morphologiques, par

l'identification de caractéristiques dans les images, invisibles à l'œil expert du radiologue. Néanmoins, malgré de nombreux résultats encourageants, force est de constater qu'aucun biomarqueur issu de la radiomique n'a aujourd'hui trouvé sa place en routine clinique malgré près de 3500 études publiées depuis 2012. Ce bilan mitigé peut s'expliquer par un triple constat. i) La mesure d'un biomarqueur en radiologie dépend des performances des systèmes d'imagerie, en plus de leur nature. Cet aspect méthodologique est à l'origine des pistes de réflexion relatives à la standardisation des données ou bien des procédures d'acquisitions des études cliniques. ii) Le manque d'interprétabilité des résultats participe également à la limitation de la radiomique. iii) Enfin, cette approche est aujourd'hui réservée aux grands centres spécialisés qui sont en possession de bases de données conséquentes et nécessaires au développement d'algorithmes d'IA robustes.

Malgré ces limites, la radiomique rencontre un véritable engouement auprès des radiologues en vue d'exploiter l'hétérogénéité tumorale, attendue comme un biomarqueur pronostic ou prédictif en cancérologie. L'hétérogénéité des pixels composants la tumeur reflète l'hétérogénéité de la mesure physique du système d'imagerie concerné. Les récents développements instrumentaux permettent ainsi d'accéder à l'hétérogénéité des densités tissulaires en radiographie X, de l'activité métabolique en examen nucléaire ou bien encore de la vascularisation tumorale en imagerie ultrasonore et IRM. Il existe ainsi plusieurs centaines d'indices décrivant l'hétérogénéité des images en radiomique et comme mentionné plus haut, ils sont issus d'un formalisme mathématique difficilement interprétable dans la pratique radiologique. Le sujet de cette thèse se situe dans ce contexte, avec l'ambition de proposer de nouveaux indices descriptifs de l'hétérogénéité tumorale. Ces indices s'appuient sur une méthodologie favorisant d'une part leur robustesse vis à vis des procédures d'analyse d'image avec l'ambition de s'affranchir ou tout du moins d'alléger les procédures de standardisation évoquées et d'autre part leur l'interprétabilité en vue de guider facilement le praticien vers une stratégie thérapeutique personnalisée.

Pour situer le développement de nos biomarqueurs dans le cadre de l'évaluation au sens large de l'hétérogénéité tumorale, nous commencerons par présenter dans le premier chapitre, les fondements des biomarqueurs à travers leurs origines et les différents rôles qui peuvent leur être attribués. L'exposé sera ensuite centré sur la nouvelle génération de biomarqueurs pour l'évaluation de la réponse thérapeutique avec notamment l'apport de l'imagerie dite quantitative. Enfin, avant de nous intéresser à l'hétérogénéité tumorale en tant que biomarqueur d'imagerie, nous décrirons les recommandations actuelles concernant le développement de biomarqueurs d'imagerie en appuyant notamment sur les raisons des nombreux échecs de transfert vers la pratique clinique.

Le second chapitre détaillera le concept et les objectifs visés pour le développement de nouveaux biomarqueurs de l'hétérogénéité ainsi que les choix retenus pour leur construction. L'accent sera en particulier mis sur leur interprétation dans le contexte

de l'imagerie paramétrique. Une modélisation numérique complétera ce chapitre afin de proposer une caractérisation des biomarqueurs proposés notamment au regard de leurs sources de variabilités.

Une exploration de l'intérêt de nos biomarqueurs pour le suivi de l'efficacité thérapeutique fera l'objet du troisième chapitre. Il s'agit d'une analyse rétrospective de données pré-cliniques provenant d'une étude sur un modèle tumoral chez le petit animal soumis à un traitement anti-angiogénique. Nous y aborderons dans un premier temps la problématique de l'angiogenèse comme cible thérapeutique, et son lien avec l'hétérogénéité vasculaire tumorale. Nous présenterons ensuite les résultats d'une étude comparative des performances des biomarqueurs développés et des méthodes traditionnelles de quantification de l'hétérogénéité en IRM de contraste (DCE-IRM).

Le quatrième chapitre sera consacré au volet clinique de ces travaux de thèse. Dans un premier temps, nous présenterons une étude de répétabilité de la mesure des biomarqueurs évalués sur des cartes paramétriques en IRM de diffusion dans le cadre d'une étude clinique prospective sur volontaires sains. Par la suite, deux pistes d'applications cliniques des biomarqueurs développés seront présentées. Par des analyses rétrospectives de données IRM cliniques, il s'agira d'explorer le potentiel de notre méthodologie pour, d'une part, faciliter la standardisation de la DCE-IRM appliquée à l'examen des cancers ORL et d'autre part, établir le diagnostic différentiel précoce entre deux pathologies, les radionécroses et les récidives dans un contexte de traitement de tumeurs cérébrales.

Enfin, nous concluons en présentant d'abord un récapitulatif des résultats obtenus puis les perspectives qui se dégagent de l'ensemble de ce travail d'un point de vue clinique.

Chapitre 1

Les biomarqueurs pour la cancérologie

Le terme « Biomarqueur » provient de la contraction des termes marqueur et biologique. Ce terme qualifie toute entité en mesure de refléter un processus biologique sous-jacent (MA & GAMAGEDARA, 2015). Les biomarqueurs peuvent être d'origine biologique, chimique ou biophysique et sont des outils scientifiques largement utilisés de la médecine à l'écologie en passant par l'astronomie et la géologie. Dans un contexte médical, les biomarqueurs désignent un ensemble de signes cliniques, biologiques, représentatifs d'un état médical, mesurés objectivement et d'une manière reproductible (STRIMBU & TAVEL, 2010). Dans cette thèse, il est question des biomarqueurs pour la cancérologie, domaine dans lequel la recherche associée y est la plus développée, en raison notamment de l'incidence des cancers à travers le monde qui pourrait atteindre près de 30 millions de cas par an en 2040 (+40%)(BRAY et al., 2018). Appliquée à ce domaine, Hayes et al proposent en 1996 une définition du biomarqueur comme un indicateur « d'une altération moléculaire, cellulaire, tissulaire ou d'un processus qui informe sur l'état actuel, ou de manière plus intéressante, sur le comportement futur du cancer » (HAYES et al., 1996). Mais la particularité des biomarqueurs en cancérologie réside dans la très grande diversité des pathologies regroupées au sein de cette spécialité, ce qui leurs confèrent une définition à la fois très large avec un rôle différent à chaque instant de la prise en charge du patient, et évolutive en lien avec les progrès thérapeutiques, technologiques et méthodologiques. Dans ce contexte, la classification TNM (EDGE & COMPTON, 2010) par la Taille de la tumeur, du Nombre de ganglions lymphatiques localement envahis, et la présence de Métastases illustre pleinement ces deux aspects de la définition du biomarqueur pour la cancérologie. Plus précisément, ce regroupement de trois indicateurs permet d'extraire une classification du stade de la progression du cancer en attribuant un chiffre à chaque lettre de l'acronyme TNM. Ainsi on attribue un score T0, 1, 2, 3 ou 4 en fonction de la taille de la tumeur et de son étendue aux tissus voisins (figure 1.1), N0, 1, 2, etc. en fonction du nombre de ganglion atteint, et M0 ou 1 selon la présence de métastases dans l'organisme (figure 1.2). Adapté à un grand nombre

de cancer, cette classification TNM illustre les critères complexes d'un biomarqueur : 1) La quantification des indicateurs de taille, de nombre de ganglions et de présence de métastases est d'une très grande robustesse d'un centre clinique à l'autre simplifiant ainsi sa mise en œuvre par l'ensemble de la communauté médicale, point essentiel pour garantir un langage commun international pour une prise en charge optimale uniformisée ; 2) cette classification guide le clinicien avec un triple objectif : diagnostique, pronostic et prédictif de la survie et de l'efficacité thérapeutique, selon les types de cancer qu'elle décrit ; 3) enfin, la définition de la classification TNM, initialement établie en 1946 par Pierre Denoix, chirurgien à Gustave Roussy (Villejuif, France), est révisée périodiquement à l'échelle internationale selon les cancers et au regard des progrès, principalement ici technologiques avec les développements en imagerie médicale, le dernier en date étant l'avènement de l'imagerie tomographique par émission de positons (TEP) permettant d'accéder, entre autres, au bilan d'extension de la maladie, donc à N et M.

Aujourd'hui, il existe un regain d'intérêt pour le développement de nouveaux biomarqueurs en lien avec une meilleure compréhension des mécanismes du cancer, une analyse plus fine des données massives et complexes et enfin des performances technologiques croissantes. C'est en particulier le cas pour les Biomarqueur d'Imagerie (BI) où les enjeux actuels visent à identifier et valider des nouveaux indicateurs pour une caractérisation fonctionnelle, métabolique ou moléculaire de la tumeur. Dans ce premier chapitre, au regard de la spécificité des biomarqueurs en cancérologie, leurs différents rôles seront dans un premier temps décrits. Puis nous associerons la diversité des biomarqueurs à leur méthode d'acquisition en dissociant les indicateurs biochimiques de ceux d'imagerie. Enfin, je finirai ce chapitre en introduisant la problématique de la quantification de l'hétérogénéité tumorale comme nouvelle piste de biomarqueur pour les thérapies ciblées.

1 Les biomarqueurs pour la cancérologie : rôles et définitions

L'un des enjeux sociétaux largement abordé par la recherche académique et clinique ces vingt dernières années est la médecine personnalisée également nommée médecine de précision. Son concept s'appuie sur la capacité à établir un parcours de soin unique pour chaque patient, et les biomarqueurs en sont les chevilles ouvrières. Appliquée à la cancérologie, on comprend aisément la nécessité de la médecine personnalisée pour combiner le triptyque thérapeutique traditionnel que sont la chimiothérapie, la radiothérapie et la chirurgie avec les nouvelles thérapies ciblées et l'immunothérapie, visant les spécificités phénotypiques de chaque tumeur. Au regard de la diversité des

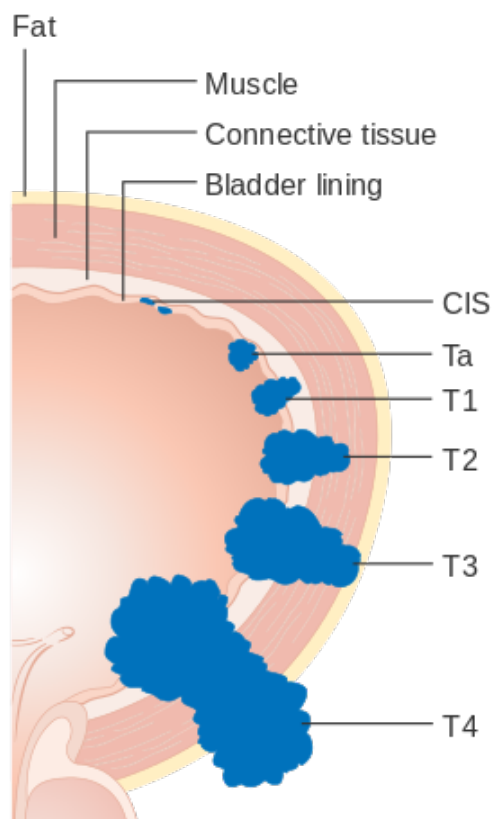


FIGURE 1.1 – Exemple de classification T du cancer de la vessie selon la taille et l'étendue de la tumeur aux tissus voisins. Tiré de Cancer Research UK.

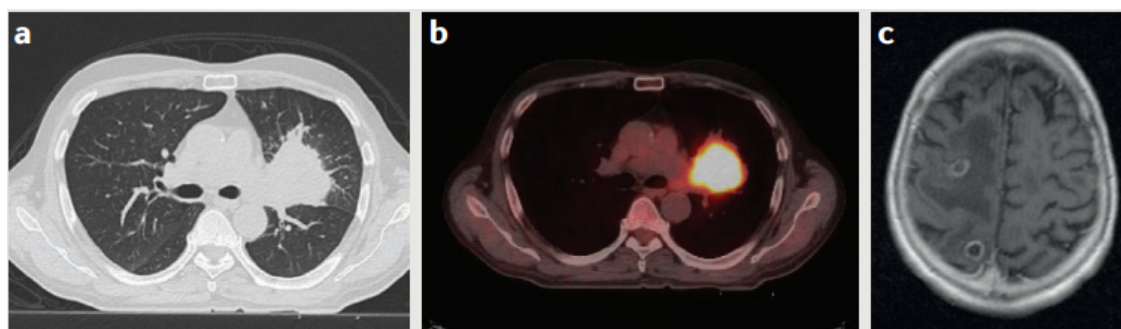


FIGURE 1.2 – Stratification TNM d'un patient avec un cancer du poumon non à petites cellules de stade IV T2 N0 M1. Le TNM est établi à partir de a) tumeur du poumon T2 en imagerie CT, b) aucune preuve de l'implication des ganglions N0 en imagerie TEP-CT et c) métastases cérébrales M1 en IRM. Tiré de O'CONNOR et al. (2017)

cancers et des traitements, il existe une multitude de biomarqueurs en oncologie avec des objectifs tout aussi variés. Dans cette première partie, nous proposons de préciser la définition du terme « biomarqueur », avant d'en décrire ses rôles aux différentes étapes clés de la prise en charge du cancer.

1.1 Biomarqueur, des définitions en perpétuelle évolution

La définition des biomarqueurs, depuis la première apparition du terme dans les années 1950 (PORTER, 1957), a eu besoin d'évoluer pour correspondre à la réalité clinique des techniques et des méthodes permettant d'accéder aux mesures pour décrire l'état du patient. Deux déclarations générales ont contribué à la définition du biomarqueur. L'Organisation Mondiale de la Santé (OMS) (angl. World Health Organization (WHO)), est la première en 1993 à définir ce que sont les biomarqueurs, à savoir « toutes les mesures qui reflètent une interaction entre un système biologique et un potentiel danger, qui peut être de nature chimique, physique ou biologique. Ces mesures peuvent être fonctionnelles, physiologiques, biochimiques au niveau cellulaire, ou au niveau des interactions moléculaires ». Cette définition se voulait volontairement très large afin de pouvoir être employée dans un large champ d'application mais ne permet pas de cadrer précisément le champ d'application de ces biomarqueurs ni de les différencier.

Pour mieux cadrer cette définition à l'échelle internationale, le U.S. National Institute of Health (NIH) et la U.S. Food and Drug Administration (FDA), deux organismes des États-Unis d'Amérique, offrent en 2001 une définition commune, plus précise, mieux adaptée aux applications cliniques modernes. Un biomarqueur est alors défini comme « une caractéristique mesurée objectivement et indicatrice d'un processus biologique normal ou pathogénique, ou d'une réponse pharmacologique à une exposition ou à une intervention thérapeutique ». Selon cette définition, les biomarqueurs peuvent donc être dérivés de tout relevé médical qu'ils soient physiologiques (pouls, pression sanguine, ...), moléculaires (sang, fluides corporels), génétiques, histologiques (tissus, biopsies) ou radiographiques.

En cancérologie, les biomarqueurs sont aujourd'hui omniprésents et apparaissent alors comme des outils indispensables pour évaluer objectivement les différents aspects du cancer à différents instants du suivi des patients. Dépistage, diagnostic, pronostic, stadification, prédiction, suivi de l'efficacité thérapeutique ou de la toxicité, sont autant de moments clés où les cliniciens ont recours à des biomarqueurs. Ces rôles bien souvent complémentaires vont être précisés dans la partie suivante.

1.2 Les différents rôles des biomarqueurs : descriptions simples et superposables

En cancérologie, on attribue différents rôles aux biomarqueurs. L'outil BEST « Biomarkers, EndpointS, and other Tools », est un outil, proposé une nouvelle fois par la FDA et le NIH, et mis à jour régulièrement depuis 2015 (FDA-NIH BIOMARKER WORKING GROUP, 2016). Cet outil prend la forme d'un glossaire et décrit au total 7 rôles correspondant à des utilisations cliniques différentes d'un biomarqueur. Bien que les descriptions de ces rôles soient généralistes et individuelles, il est néanmoins possible de les classer en sous-groupes. Les critères qui définissent ces rôles sont si riches et variés qu'il serait possible de former des groupes de différentes manières, en fonction du contexte d'utilisation ou encore au regard de l'objectif clinique. Pour permettre d'effectuer un tour d'horizon de ces différents rôles, nous proposons une analogie avec les mathématiques en considérant, d'un côté des biomarqueurs qui revêtent un rôle déterministe, c'est-à-dire, dont la mesure permet d'établir une certitude à l'instant t , et d'un autre côté les biomarqueurs dont l'évaluation indique une probabilité de l'évolution de la maladie. Ces différents rôles sont ainsi décrits dans la Table 1.1 en fonction de la nature de leur réponse, déterministe ou probabiliste.

Dans le cadre de cette thèse, nous nous intéressons principalement à quatre rôles décrits ci-dessous : deux déterministes, les biomarqueurs diagnostiques et de suivi, deux probabilistes, les biomarqueurs pronostiques et prédictifs.

Les biomarqueurs diagnostiques permettent de détecter ou de confirmer la présence de la maladie, généralement au moment du dépistage puis dans un second temps peuvent également permettre d'identifier les sous types cancéreux. Les biomarqueurs de suivi sont ceux pour lesquelles la mesure est répétée de manière longitudinale, ils sont utiles tout au long du parcours de soin et leur objectif est de classer la réponse thérapeutique. Cette réponse peut être partielle ou complète, ou au contraire indiquer une progression ou encore une stabilité de la maladie.

Un biomarqueur pronostique indique l'évolution probable de la maladie indépendamment du traitement, qu'il s'agisse d'une rémission, d'une récurrence ou d'une progression de la maladie. Les biomarqueurs prédictifs prévoient la probabilité de survenue d'un événement clinique favorable ou non, en réponse à un traitement. Il est à noter qu'en cancérologie, on entend par événement clinique, tout événement comme la récurrence de la maladie, le décès, la progression, ou tout signe clinique d'aggravation de la maladie. Deux indicateurs sont généralement utilisés pour juger de la durée entre le diagnostic ou le début du traitement et ces événements. Ces indicateurs sont la survie globale (angl. Overall Survival (OS)) qui est à ce jour le critère de jugement d'un traitement privilégié tandis que la survie sans progression (angl. Progression Free Survival (PFS)), autrement dit la durée de l'effet du traitement est considérée de manière

secondaire, en l'absence d'amélioration de la survie.

Faire la distinction entre les biomarqueurs pronostiques ou prédictifs n'est pas toujours simple. Nous proposons de l'illustrer à l'aide du critère TNM déjà introduit précédemment et pour lequel il a été mentionné qu'il regroupait différents objectifs cliniques. Cette classification est en effet un biomarqueur pronostique pour la majorité des types de cancer. C'est-à-dire qu'il permet d'établir un pronostic relatif à l'issue clinique de la maladie et distingue les cancers T1-2 des cancers T3-4 dont la taille est significativement plus importante dans ce deuxième cas et pour lesquels les chances de survies sont plus faibles. Il est également prédictif en permettant d'identifier les patients qui pourraient bénéficier de certaines thérapies, notamment aux traitements anti-angiogéniques indiqués en première ligne pour les cancers en phase métastatique, identifiés M1, du côlon, du sein, du rein ou du poumon.

De manière générale le recours aux biomarqueurs est primordial en cancérologie. En effet, ils permettent de réduire les temps d'exposition des patients à des traitements inefficaces en sélectionnant la meilleure voie thérapeutique possible. Cela est d'autant plus cruciale à un stade précoce des traitements puisque de multiples études, notamment en TEP (de GEUS-OEI et al., 2009 ; KOSTAKOGLU & CHESON, 2014 ; SALAVATI et al., 2009) ont aujourd'hui démontrées une relation entre la réponse précoce à un traitement et la réponse finale et donc à OS ou PFS.

En plus de cette diversité de rôles que peuvent posséder les biomarqueurs, ils se différencient également selon leur méthode d'acquisition. Nous allons donc préciser dans la section suivante, les différentes approches qui nous permettent de collecter des biomarqueurs pour la cancérologie.

Rôle	Utilisation Clinique	Exemple	Référence
Rôle déterministe			
Diagnostique	Détecter ou confirmer la présence de la maladie ; Identifier les sous types cancéreux	Expression génétique	Scott et al, 2014
Suivi	Évaluer le statut et l'étendue de la maladie via des mesures longitudinales ; Évaluer l'effet d'une intervention thérapeutique et classifier la réponse comme nulle, partielle, totale ou en progression	PSA	Freedland, Moul, 2007
			Suite à la page suivante

Suite de la page précédente

Rôle	Utilisation Clinique	Exemple	Référence
Pharmacodynamique	Démontrer l'apparition d'une réponse biologique après exposition à un agent ; Il s'agit d'un biomarqueur de suivi qui peut fournir des informations précoces sur l'efficacité d'un traitement au regard de l'issue clinique ou des informations pharmacologique au regard de problèmes de sureté	SUV	Keloff, 2005 ; Wahl, 2009
Sureté	Mesuré avant ou après une exposition à un agent pour évaluer la présence ou l'étendue des effets indésirables	Comptage de Neutrophile	Rizzo, 2010 ; Smith, 2015
Rôle probabiliste			
Pronostic	Identifier la probabilité de survenue d'un nouvel événement clinique, d'une récurrence ou d'une progression	TNM	Edge, 2010
Prédictif	Identifier les individus favorables ou non qui pourraient bénéficier des effets d'une thérapie	TNM	McLeod, 2005
Risque/Susceptibilité	Indiquer la probabilité de développer une maladie pour un individu ne présentant aucun signe clinique	BRCA1/2	Struewing, 1997 ; Thorlacijs, 1998
Sureté	Mesuré avant ou après une exposition à un agent pour évaluer la probabilité de subir des effets indésirables		

TABLE 1.1 – Les différents rôles attribués au biomarqueur accompagnés d'une définition de leur utilisation clinique ainsi que d'un exemple. Ces rôles sont reparties dans deux groupes, d'un point de vue mathématique : déterministe ou probabiliste

2 Diversité et méthodes d'acquisitions des biomarqueurs

Connaître la façon dont est déterminé un biomarqueur est un élément essentiel pour apprécier de manière triviale son universalité, c'est-à-dire son exploitabilité à

l'échelle inter-individu, inter-service médical, inter-machine ou toute autre échelle. En cancérologie, il existe deux grandes familles de biomarqueurs au regard de leur mode d'acquisition : ceux issus soit des analyses biochimiques soit de l'imagerie médicale, radiologique ou nucléaire. Dans ces deux familles, les découvertes de biomarqueurs sont guidées par des progrès technologiques tels que la multiplication des analyses génomiques pour les biochimiques et les évolutions des modalités d'imagerie, mais également par des avancées méthodologiques telles que l'émergence de l'Intelligence Artificielle (IA) avec le "deep-learning" et la radiomique en imagerie pour traiter les données massives. L'engouement portant sur les biomarqueurs associés à ces évolutions technologiques conduit actuellement à une amplification des recherches translationnelles dans le domaine avec le développement d'une quantité croissante d'indices potentiels. Ces deux familles de biomarqueurs, régulièrement combinées en routine clinique, vont être abordées dans cette deuxième partie de chapitre.

2.1 Les biomarqueurs biochimiques en routine clinique

Les biomarqueurs biochimiques résultent des dosages dérivés d'échantillons biologiques d'origines diverses : tissu prélevé en diagnostic initial ou bien en peropératoire, échantillons sanguins, urinaires, sérum ou tout autre fluide biologique. Les différentes méthodes d'analyses mises en jeu au regard de la diversité de la nature des échantillons sont nombreuses et il ne s'agit pas ici de les décrire de manière exhaustive. En revanche, il nous apparaît important de revenir sur l'analyse anatomopathologique, l'examen diagnostique phare en cancérologie permettant d'accéder à des biomarqueurs pronostiques, prédictifs et de suivi de référence, dit « gold standard ». Cette analyse regroupe l'histologie qui étudie le tissu tumoral et son architecture à l'échelle mésoscopique ainsi que la cytologie qui concerne l'échelle cellulaire. Elle permet ainsi de déterminer le caractère malin de la tumeur, précise le type cellulaire concerné (sarcome, carcinome, ...) et son agressivité avec le nombre de cellules en division. Cette analyse peut être complétée par différentes méthodes de marquage des cellules pour informer sur la présence de différentes protéines ou molécules et peut ainsi fournir des biomarqueurs prédictifs permettant d'orienter le patient vers une stratégie de traitement personnalisée. D'autres technologies plus récentes mettant en œuvres des automates permettent de déterminer le génotype et le phénotype des cellules présentes. En fonction de la molécule étudiée, on parle d'un champ différent de la biologie : la génomique avec le séquençage de l'ADN, la transcriptomique avec l'étude des ARN messagers produits à la transcription de l'ADN, la protéomique en identifiant les protéines et la métabolomique avec cette fois-ci la présence de métabolites (lipides, sucres, aminoacides) nécessaires entre autres à la prolifération et à la vascularisation tumorale. La figure 1.3 représente le lien biologique entre chacun de ces champs d'applications.

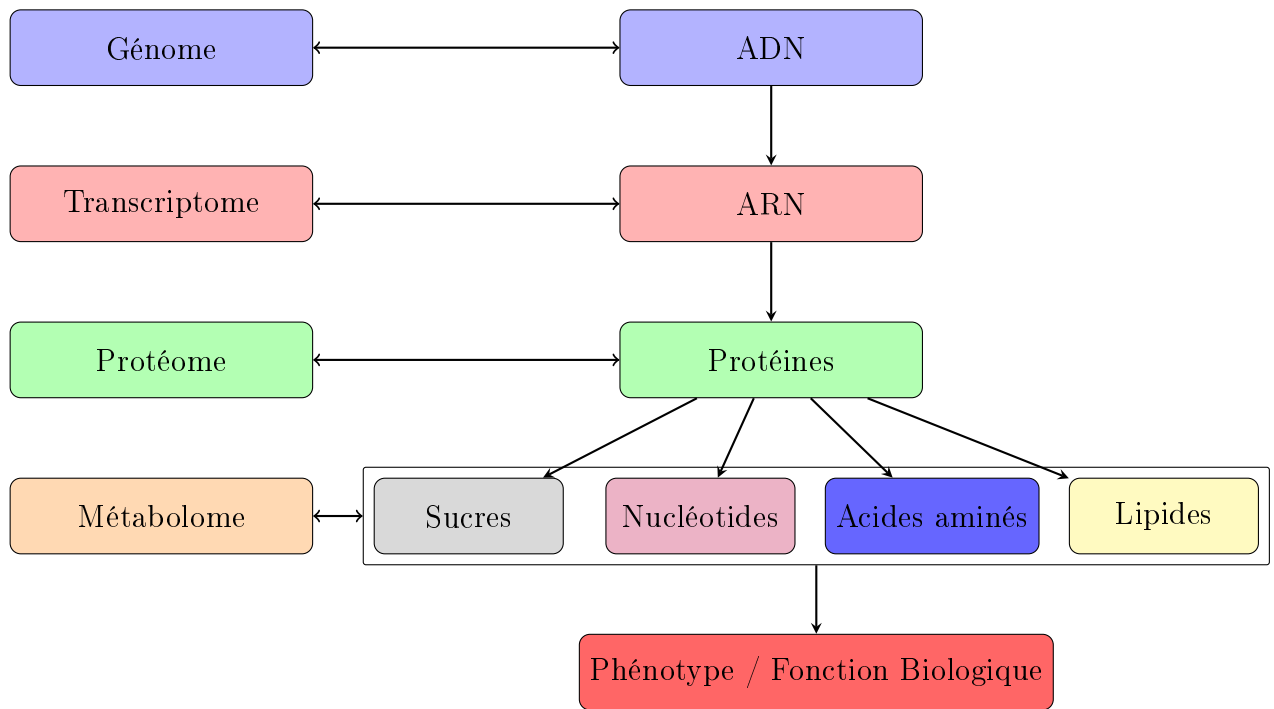


FIGURE 1.3 – Flux de l’information de l’ADN jusqu’au phénotype selon le dogme de la biologie. A chaque étape est associé l’outil d’analyse biologique correspondant, de l’analyse génomique (génomique) jusqu’à l’analyse métabolique (métabolome).

Au-delà du diagnostic, ces analyses permettent aujourd’hui d’établir des biomarqueurs pronostiques ou prédictif de l’efficacité thérapeutique. Par exemple, YAMAGUCHI et al. (2020) ont montré que pour les patients atteints de cancer gastriques et pour lesquelles la cytologie du lavage péritonéal (CY1) était positive, une chimiothérapie post-opératoire (gastrectomie) pourrait avoir un avantage en termes de survie pour ces patients avec un prolongement de 9.9 mois à >25 mois. De nombreux biomarqueurs biochimiques sont couramment employés en routine clinique et la table 1.2 en dresse une liste non-exhaustive.

Biomarqueur (fluide de mesure)	Technologie associée	Application (Organe)	Référence
Génomique			
Mutations BRCA-1 et BRCA-2 (sang)	Séquencage ADN	Pronostic et évaluation des risques (sein)	MARUVADA 38
Osteopontin (sang)	Microarray	Diagnostic, Pronostic (ovaire)	MARUVADA 60
Suite à la page suivante			

Suite de la page précédente

Biomarqueur (fluide de mesure)	Technologie associée	Application (Organe)	Référence
HER-2/neu (tissue,serum)	PCR,FISH	Pronostic et Traitement (sein)	MARUVADA 63
Protéomique			
PSA (Sérum)	Immunoassay	Diagnostic, Pronostic (prostate)	MARUVADA 52
CEA (Sérum)	Immunoassay	Diagnostic, Pronostic (colon, poumon, sein)	MARUVADA 68,70
EGFR (Sérum)	Immunoassay	Prognostic, Traitement (poumon)	MARUVADA 57
Métabolomique			
Profilage Métabolomique (sérum)	¹ H NMR Spectroscopy	Détection précoce et diagnostique (ovaire)	MARUVADA 75
Profilage métabolique dynamique basé sur les isotopes stables du glucose (SIDMAP) (cellules tumorales)	Spéctrometrie de masse	Découverte de traitement et traitement (pancréas)	MARUVADA 76

TABLE 1.2 – Liste non exhaustive de différents biomarqueurs issus d'échantillons biologiques pouvant être employés en cancérologie. Les biomarqueurs sont indiqués avec entre parenthèses l'échantillon biologique utilisé pour la mesure, la technologie employée pour effectuer la mesure, son rôle en routine clinique et la référence associée.

Dans ce schéma de diagnostic et de traitement du cancer, les biomarqueurs tiennent un rôle clé. Les découvertes ont permis de faire basculer certains cancers de fatal à curable, néanmoins pour aller encore plus loin et guérir toujours plus de patient, les stratégies futures cherchent à intercepter le cancer avant son apparition et donc

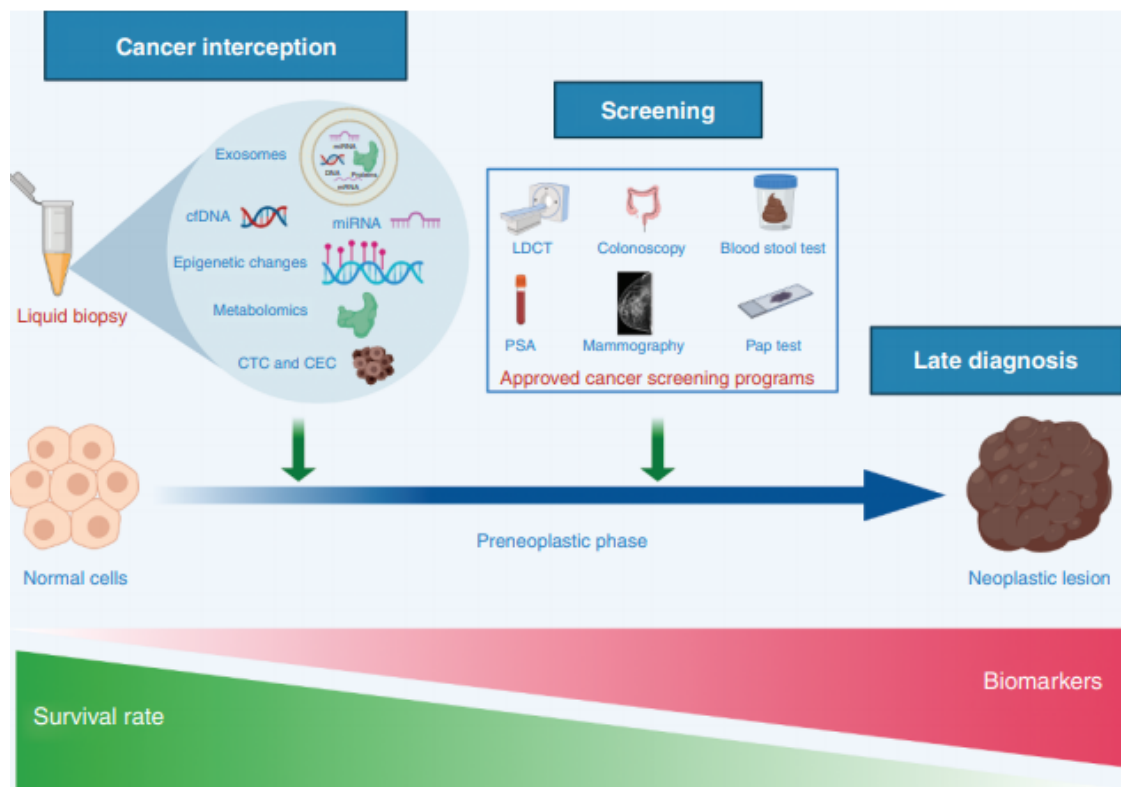


FIGURE 1.4 – Principe de l’interception du cancer et positionnement des biopsies dites liquide en amont des biomarqueurs de dépistage et de suivi du traitement. Tiré de SERRANO et al. (2020)

sa détection grâce aux biomarqueurs cités précédemment. Pour mener cette stratégie d’interception, les nouvelles technologies basées sur les biopsies liquides sont en plein essor. Ces technologies sont utilisées pour analyser les maladies lorsqu’elles sont dans un stade de pré-malignité et sont effectuées à partir des cellules circulantes dont des exemples de méthodes d’acquisitions sont représentés figure 1.4. Ce concept d’interception du cancer n’est pas nouveau mais représentait jusqu’à présent une utopie. Pour devenir une réalité, de nouvelles méthodes d’acquisition et d’analyse minimalement invasives basées sur les biopsies liquides et associées aux technologies les plus avancées de l’imagerie médicale sont nécessaires (SERRANO et al., 2020).

2.2 Les biomarqueurs d’imagerie en routine clinique

2.2.1 Définition d’un biomarqueur d’imagerie et modalités associées

Les avancées méthodologiques associées aux progrès technologiques permettent à l’imagerie médicale non seulement d’établir un diagnostic mais également de fournir des informations fonctionnelles, métaboliques, générant des BI pronostiques ou prédictifs. Ceux-ci peuvent dériver d’une ou plusieurs séries d’images acquises durant le parcours clinique d’un patient, du dépistage au suivi post-thérapeutique. En complément des

informations biochimiques, les BI ont pour avantage principal d'être accessible de manière non-invasive, ce qui leur confère la possibilité de multiplier les examens d'imagerie au cours du traitement. Ils sont par ailleurs accessibles en temps réel, pendant ou après l'acquisition. Enfin, l'imagerie permet d'examiner les tumeurs dans leur globalité à une échelle macroscopique en considérant leur hétérogénéité et permet d'évaluer de multiples lésions indépendamment et simultanément. Ces BI sont extraits à partir des différentes modalités d'imagerie dont les modes de détections définissent la nature de ces indicateurs. Dans ce contexte, les principales propriétés des systèmes d'imagerie vont être brièvement et schématiquement introduites.

Imagerie par Résonance Magnétique (IRM)

L'IRM est une modalité d'imagerie dont le principe repose sur la détection du signal émanant des protons (essentiellement de l'eau et des lipides) placés dans un champ magnétique intense après une perturbation de leur état d'aimantation par des impulsions radiofréquences (RF). Le signal IRM renseigne sur les propriétés physico-chimiques des tissus et dépend du champ magnétique, typiquement 1,5 ou 3 Teslas en routine clinique, et des séquences prédéfinies des impulsions RF. L'IRM possède les avantages de fournir des informations anatomiques de haute-résolution spatiale tout en possédant une grande profondeur de pénétration sans aucune problématique de radiation ionisante pour les patients.

Scanner (tomodensitométrie) angl. Computed Tomography (CT)(HSIEH, 2015)

Le scanner, également connue sous son abréviation anglaise « CT Scan » est une modalité d'imagerie qui repose sur la mesure des coefficients d'atténuation des tissus aux rayons X en fonction de celui de l'eau. L'unité de mesure de ce rapport de coefficient est l'unité Hounsfield (HU). Du fait de sa grande disponibilité, de sa facilité d'utilisation et de la rapidité d'acquisition de quelques secondes, le scanner est devenu la modalité d'imagerie incontournable en cancérologie, permettant de fournir des images de haute résolution anatomique.

Echographie UltraSonore (US)(Manual of Diagnostic Ultrasound. 1, 2011)

L'échographie est une imagerie d'interface qui repose sur le principe de la différence d'impédances acoustiques mise en jeu selon les densités tissulaires. Ainsi, l'imagerie acoustique permet de visualiser les interfaces entre les tissus, les organes et les lésions. Les échographies sont fréquemment faites en oncologie notamment grâce à leur disponibilité, leur faible coût et leur aspect non-ionisant.

Imagerie nucléaire(KHALIL, 2017)

La TEP et la Tomographie par Emission Mono-Photonique (TEMP) (angl. SPECT) sont deux modalités d'imagerie nucléaire. Elles reposent toutes deux sur la détection des rayonnements gamma d'un radiopharmaceutique injecté au patient et métabolisé ou accumulé par la lésion à examiner. La différence entre les deux modalités d'imagerie repose sur la nature du rayonnement radioactif du radiopharmaceutique émettant un

ou deux photons simultanément¹, ce qui d'un point de vue instrumental implique des principes de détection très différents (LIVIERATOS, 2012 ; OEHR, s. d.). Les protocoles standards permettent de réaliser des examens corps entier. Couplé au CT à la fois pour son information anatomique et ses mesures correctives de l'atténuation des tissus, cet examen est aujourd'hui incontournable dans le bilan d'extension de la maladie avec une sensibilité accrue au métabolisme de l'analogue du glucose, le désoxyglucose.

2.2.2 Utilisation des biomarqueurs d'imagerie, du diagnostic à l'évaluation de la réponse thérapeutique

De nombreux BI sont aujourd'hui utilisés en routine clinique (table 1.3). Ces indicateurs sont utilisés pour le dépistage, la détection des tumeurs et le suivi des patients dans le temps y compris pour l'évaluation de la réponse thérapeutique et pour l'identification des complications thérapeutiques.

En pratique, grâce à des mises à jour régulières et une recherche continue, les BI suivants, principalement basés sur la taille et l'aspect visuel de la tumeur servent aujourd'hui de référence. Nous l'avons déjà évoqué, mais le critère TNM est un des biomarqueurs les plus couramment utilisé pour stadifier les cancers. Le Breast Imaging Reporting and Data System (BI-RADS) est un autre biomarqueur basé sur un système de classification proposé par le American College of Radiology (ACR) depuis 1993 et dont la dernière mise à jour date de 2003 avec sa cinquième version. Son rôle est diagnostique pour le cancer du sein qui permet de les classer via les résultats d'imagerie d'IRM, d'US et de mammographie. Ce biomarqueur propose 6 niveaux de classification, (investigation incomplète, anomalie bénigne, probablement bénigne, anomalie suspecte, fortement suspecte, cancer prouvé par la biopsie) et permet d'adapter la prise en charge uniformément pour chaque patient. A noter que ce biomarqueur, initialement développé pour le sein est aujourd'hui dérivé pour différents organes comme le foie (LI-RADS), les poumons (Lung-RADS), la prostate (PI-RADS), et d'autres.

On notera également, l'existence de deux indicateurs issus de l'imagerie nucléaire qui ne s'appuient pas sur la taille des tumeurs. i) Un biomarqueur, à visée diagnostique uniquement, issu de l'imagerie TEP avec la quantification de la fixation d'un radiopharmaceutique analogue du glucose, le fluorodésoxyglucose (F18-FDG). La quantité de ce produit de contraste dans la lésion normalisée par la dose injectée et/ou le poids du patient est l'indicateur Standardized Uptake Value (SUV)². Ce biomarqueur reflète l'activité métabolique de la tumeur et permet ainsi, de dissocier les tissus malins versus bénins au-delà d'une valeur seuil autour de 3 (DIMITRAKOPOULOU-STRAUSS et

1. Les deux photons sont émis simultanément à la suite de l'annihilation d'un positron par le radiopharmaceutique en TEP

2. Il existe plusieurs indicateurs SUV qui varie selon la région d'intérêt considérée : le pixel d'intensité maximale, un volume d'1mL, une région tracée à la main ou bien définie par un seuil automatique.

al., 2002). ii) Un biomarqueur pronostic pour le cancer de la prostate, issu de l'imagerie TEMP, le BSI pour Bone Scan Index. Cet indicateur évalue l'étendue du cancer au niveau squelettique et représente la charge tumorale globale³ (DENNIS et al., 2012; SABBATINI et al., 1999).

Traditionnellement, pour évaluer la réponse thérapeutique chez les patients atteints de tumeurs solides, les cliniciens ont recours à des BI anatomiques qui se basent sur des mesures de la taille des tumeurs. On parle également de biomarqueur morphologique. Parmi les plus couramment utilisés, on retrouve le critère RECIST 1.1 développé en 2009 par l'équipe du Pr Therasse (EISENHAUER et al., 2009). Ce critère existe en réalité depuis 1970 et est le résultat de mises à jour successives des critères WHO (MILLER et al., 1981; WORLD HEALTH ORGANIZATION, 1979), RECIST (THERASSE et al., 2000) puis RECIST 1.1. L'objectif de ce biomarqueur est d'évaluer la charge tumorale globale en se basant sur l'évolution de la longueur du plus grand axe d'un nombre maximal (10) de lésions au cours du temps. Nous proposons table 1.4 un résumé des caractéristiques permettant de former le critère RECIST 1.1 afin d'évaluer la réponse thérapeutique ainsi qu'une comparaison aux caractéristiques de ses prédécesseurs.

L'avènement des nouvelles thérapies ciblées avec la nécessité de conduire une médecine dite personnalisée implique le développement de biomarqueurs de l'efficacité thérapeutique complémentaires aux critères morphologiques. C'est en particulier le cas des tumeurs du rein, stromales gastro-intestinales (GIST) métastatiques qui bénéficient aujourd'hui en première intention de thérapies ciblées mais pour lesquels l'évaluation via RECIST 1.1 n'est pas un bon indicateur (GRIMALDI et al., 2018; NISHINO, 2018). Nous présenterons dans la partie suivante cette nouvelle génération de BI complémentaires qui cherchent à évaluer autrement la réponse thérapeutique.

Biomarqueur d'Imagerie (BI)	Modalité	Cancer	Rôle	Décision potentielle	Utilité du Biomarqueur
RECIST (Volume morphologique)	TDM, IRM, US	Tumeurs solides	Réponse	Continuité, changement traitement	Recommandation actuelle pour l'évaluation de la réponse
PERCIST (volume métabolique)	TEP	Tumeurs solides	Réponse	Continuité, changement traitement	Recommandation actuelle pour l'évaluation de la réponse
					Suite à la page suivante

3. La charge tumorale globale correspond au nombre de cellules cancéreuses, à la taille de la tumeur, ou à la quantité de cancer dans le corps

Suite de la page précédente

Biomarqueur d'Imagerie (BI)	Modalité	Cancer	Rôle	Décision potentielle	Utilité du Biomarqueur
^{18}F FDG- SUV_{max}	TEP	Multiples	Réponse thérapeutique, Prédicatif	Continuité, arrêt, changement traitement	Prédicatif des contrôles loco-régionaux pour les cancers de la tête et du cou
ACR BI-RADS	Mammo- graphie	Sein	Diagnostic	Orienté la thérapie	Utilisé mondialement
Statut TNM	XR, TDM, IRM, TEP, TEMP, US	Presque tout les cancers	Pronostic	Orienté la thérapie	Utilisé mondialement, influence le management de presque tout les patients avec tumeur solide, Validé et qualifié exhaustivement
Indice de Scintigraphie Osseuse	TEMP	Prostate	Pronostic, Prédicatif	Orienté la thérapie	Les logiciels employés pour le calcul de l'indice nécessitent des approbations réglementaires
Fraction d'éjection ventriculaire gauche (FEVG)	US, Scintigraphie	Sein	Sureté	Orienté les thérapies	Guide un nombre important de patients (ex : trastuzumab), Confirmation de la décroissance de $\geq 10\%$ par répétition d'examen
T-score	DXA	Sein	Sureté	Guide le suivi thérapeutique	Les logiciels employés pour le calcul de l'indice nécessitent des approbations réglementaires
Absorption au delà du seuil	TEMP	Sein, Mélanome	Diagnostic	Détection peropératoire des ganglions sentinelles	Le seuil du BI correspond à un comptage de la radioactivité ≥ 3 écarts types du comptage moyen des tissus à au moins 200 mm du site d'injection
					Suite à la page suivante

Suite de la page précédente

Biomarqueur d'Imagerie (BI)	Modalité	Cancer	Rôle	Décision potentielle	Utilité du Biomarqueur
Catégorisation MARIBS	IRM	Sein	Risque	Pour les patients présentant des facteurs de risques génétiques comme les mutations BRCA1/2	Approuvé au Royaume-Uni

TABLE 1.3 – Liste non-exhaustive des Biomarqueurs d'imagerie validés à l'échelle internationale, adaptée de O'CONNOR et al. (2017). Les autres biomarqueurs d'imagerie fonctionnelle sont aujourd'hui encore au stade de l'essai clinique et ne sont pas utilisés en routine clinique. Abréviations : DXA - Absorptiométrie par Rayons X ; XR - Rayons X

3 La nouvelle génération de biomarqueurs pour l'évaluation de la réponse thérapeutique

Historiquement, la TEP représente la première modalité d'imagerie qui a bousculé les approches cliniques traditionnelles dans les années 2000 avec le premier biomarqueur à visée diagnostique issu de l'imagerie fonctionnelle quantitative. Par ailleurs, lorsqu'il s'agit d'évaluer l'efficacité thérapeutique, le SUV a également toute sa place avec la distinction entre une réponse complète ou partielle, et une maladie stable ou bien en progression (BENZ et al., 2008). En parallèle des études cliniques sur le potentiel du SUV comme biomarqueur pour la cancérologie, Choi et al ont montré dès 2004 sa corrélation avec l'imagerie conventionnelle CT pour évaluer la progression des GIST et l'efficacité thérapeutique en se basant cette fois-ci sur la quantification de la variation de la densité tumorale mesurée en unité Hounsfield UH associé au diamètre de la lésion (CHOI et al., 2004). Aujourd'hui, il en découle les critères CHOI de 2007 (CHOI et al., 2007) et les recommandations PERCIST introduites en 2009 pour l'évaluation par TEP (WAHL et al., 2009), deux approches complémentaires des indicateurs morphologiques décrits dans RECIST. De plus, ces nouveaux critères ont ouvert la voix aux développements de

	WHO	RECIST	RECIST 1.1
<i>Méthode pour la mesure</i>			
Lésions	Deux dimensions	Une dimension - Lésions cibles identifiées	Une dimension - Lésions cibles et ganglions identifiés
Mesure	Produit du plus long diamètre et du plus long diamètre perpendiculaire	Diamètre le plus long	Diamètre le plus long pour les lésions Diamètre le plus court pour les ganglion.
<i>Critère de réponse</i>			
Réponse complète (CR)	Disparition de toutes les lésions connues; Confirmation ≥ 4 semaines plus tard	Disparition de toutes les lésions connues; Confirmation ≥ 4 semaines plus tard	Disparition de toutes les lésions connues; Confirmation ≥ 4 semaines plus tard; L'axe court des nœuds lymphatiques pathologiques doit être $< 10\text{mm}$.
Réponse partielle (PR)	Réduction de taille $\geq 50\%$	Réduction de la somme des longs axes $\geq 30\%$	Réduction $\geq 30\%$ de la somme des longs axes depuis la baseline.
Maladie stable (SD)	Le patient ne répond pas aux critères PR ou PD	Le patient ne répond pas aux critères PR ou PD	La patient ne répond pas aux critères de PR ni de PD en utilisant la somme la plus faible comme référence.
Maladie en progression (PD)	Augmentation $\geq 25\%$ pour au moins une lésion mesurable; ou apparition de nouvelles lésions	Augmentation $\geq 20\%$ de la somme des longs axes depuis la référence; ou apparition de nouvelles lésions	Augmentation $\geq 20\%$ de la somme des longs axes en utilisant comme référence la somme la plus faible au cours de l'étude; la somme doit aussi démontrée une augmentation absolue $\geq 5\text{ mm}$; ou apparition de nouvelles lésions.

TABLE 1.4 – Caractéristiques principales des critères WHO, RECIST, RECIST 1.1 pour évaluer la réponse thérapeutique par le biais de mesure de différentes dimensions des lésions.

nouveaux biomarqueurs issus de données complexes d'images quantitatives fonctionnelles des tumeurs. En particulier, il y a eu un engouement pour extraire des critères basés sur le rehaussement du signal après injection d'agents de contraste, mais de manière plus générale à partir des modalités d'imagerie quantitative. C'est le cas avec l'IRM de contraste, de diffusion, de l'échographie dite également de contraste, et de l'élastographie par IRM ou par US et dont les principes de fonctionnement sont précisés ci-dessous.

IRM dynamique de contraste

L'acquisition dynamique d'images en IRM de contraste s'effectue après l'injection d'un agent de contraste à base de Gadolinium (p.ex. Gd-DOTA, Gd-DTPA). L'IRM dynamique de contraste implique l'acquisition d'une série d'image avant, pendant et après l'arrivée et le départ de l'agent de contraste dans la tumeur. A partir de la courbe de rehaussement du signal, en fonction de la pondération de la séquence et du modèle d'ajustement utilisé, il est possible d'extraire des informations relatives à la perfusion des tissus, au volume sanguin, à la perfusion microvasculaire et à la perméabilité capillaire. Deux types de techniques dynamiques permettent d'extraire ces informations. Les séquences dites de contraste DCE (Dynamic Contrast Enhanced), réalisées à l'aide de séquences pondérées T1, et les séquence dites de susceptibilité DSC (Dynamic Susceptibility- Contrast), réalisées à partir de séquences pondérées T2*. Pour la technique DCE, Il existe plusieurs modèles mathématiques compartimentaux permettant d'extraire des paramètres quantitatifs par une analyse pharmacocinétique de la courbe de rehaussement du signal, le plus connu étant le modèle développé par Tofts et Kermode (TOFTS & KERMODE, 1991) pour les tumeurs cérébrales. Parmi ces paramètres quantitatifs on retrouve, la constante du transfert de volume K^{trans} , aussi appelée le taux de remplissage, le volume fractionnel de l'espace extravasculaire extracellulaire (EES) v_e , le volume fractionnel de l'espace intravasculaire v_p et k_{ep} la constante du taux de transfert entre EES et le plasma, aussi appelé le taux de vidange et le temps de transit moyen MTT. On a la relation $k_{ep} = K^{trans}/v_e$. De faibles valeurs de K^{trans} indiquent une faible perfusion et/ou perméabilité capillaire. La technique DSC permet quant à elle d'extraire les paramètres quantitatifs de flux sanguin relatif (rBF) et de volume sanguin relatif (rBV). De manière générale, l'IRM de contraste est mise en œuvre pour suivre l'efficacité thérapeutique pour une variété de traitements, incluant la chimiothérapie pour les cancers du sein, de la vessie et les sarcomes osseux (PADHANI & KHAN, 2010). Il a été montré à plusieurs reprises pour le sein, qu'une diminution progressive du K^{trans} dans la tumeur était accompagnée d'une réponse à la chimiothérapie et qu'à l'inverse, une augmentation ou une stagnation du K^{trans} était prédictive d'une absence de réponse (AH-SEE et al., 2008 ; PADHANI et al., 2006). Toutefois, il apparait que l'IRM dynamique est une méthode largement sous utilisée dans sa pratique clinique, avec un intérêt qualitatif uniquement et peu de considérations pour les valeurs quantitatives extraites des cinétiques de rehaussement malgré les promesses

de ces paramètres quantitatifs. L'un des problèmes majeurs à la validation massive de ces indices réside dans la standardisation des protocoles d'acquisitions et d'analyses. Nous noterons tout de même l'existence d'un profil d'acquisition, c.à.d. d'un consensus relatif à l'acquisition en IRM de contraste qui a pour objectif d'harmoniser la technique sur les différents sites⁴ (SHUKLA-DAVE et al., 2019). Une approche rigoureuse à leur validation comme cela a été le cas pour le SUV en TEP, devrait permettre leur qualification pour une utilisation clinique. Nous y reviendrons dans la section suivante.

IRM de Diffusion – DWI (Diffusion Weighted IRM)

L'imagerie de diffusion est une technique particulière pour laquelle le contraste de l'image est lié directement à l'agitation thermique des molécules d'eau et non à un agent exogène. L'agitation des molécules d'eau dans les tissus n'est pas aléatoire et est restreinte, leur mouvement étant modifié et limité par les interactions entre les membranes des cellules, les macromolécules et les parois des vaisseaux sanguins. L'IRM de diffusion est la seule modalité à fournir des informations sur la cellularité⁵ des tissus, la tortuosité de l'espace extra-cellulaire et l'intégrité des membranes cellulaires. Il est possible de générer un paramètre quantitatif évaluant la diffusion dans les tissus. Ce paramètre est nommé ADC pour Apparent Diffusion Coefficient (mm^2/s). La pondération en diffusion est réalisée par l'application de gradients de champ magnétique supplémentaires dans la séquence IRM. La pondération est exprimée par la valeur de b (s/mm^2). Pour générer les cartes ADC, il faut typiquement réaliser au moins deux séquences avec des pondérations en diffusion différentes. L'ADC présente un intérêt certain pour l'évaluation de la réponse thérapeutique. En effet, tout processus qui conduit à la nécrose va nécessairement lever les contraintes cellulaires sur les mouvements des molécules d'eau et donc augmenter l'ADC. Puisque la mort cellulaire et les modifications vasculaires peuvent toutes deux précéder les modifications volumiques des tumeurs, l'ADC se présente comme un biomarqueur potentiel de l'efficacité précoce des thérapies qui induisent l'apoptose ou la nécrose. Bien que les cartes ADC soient largement employées pour leur aspect qualitatif en pratique clinique par une visualisation du contraste, de plus en plus d'études démontrent que l'augmentation de l'ADC tumoral reflète un succès thérapeutique pour les cancers du sein, du foie, du cerveau ou encore pour les sarcomes osseux (PADHANI & KHAN, 2010). Pour le sein, l'ADC peut indiquer la réponse au traitement de manière précoce en comparaison aux mesures de la taille des tumeurs (PICKLES et al., 2006), néanmoins, l'un des problèmes majeurs à son implantation clinique et à l'évolution du critère RECIST, réside une fois de plus dans la standardisation des approches de collection et d'analyse de ces données (PADHANI et al.,

4. http://qibawiki.rsna.org/images/1/1f/QIBA_DCE-MRI_Profile-Stage_1-Public_Comment.pdf

5. Terme issu des analyses histopathologiques ; Correspond à la quantité et à la densité des cellules dans un tissu.

2009). Comme pour l'IRM de contraste, un profil d'acquisition existe aujourd'hui pour faciliter l'harmonisation d'acquisition entre les différents sites, celui-ci est régulièrement mis-a-jour⁶. Sa diffusion et son utilisation pour permettre d'exploiter le plein potentiel de la technique se doit désormais d'être massive.

Echographie de contraste - DCE-US

Depuis les années 2000, il existe des agents de contraste dédiés à l'imagerie US sous forme de microbulles de gaz, généralement de l'hexafluorure de soufre. Ce produit de contraste, contrairement à celui employé en DCE-IRM, possède une taille micrométrique et est ainsi contraint à l'espace intravasculaire. C'est à partir d'un modèle mathématique de la courbe de rehaussement du signal qu'il est possible d'extraire des informations quantitatives comme l' Aire sous la courbe (AUC) (angl. Area Under the Curve) relative à la vascularisation de la zone d'intérêt en lien avec le débit et le volume vasculaire intratumoral ou encore le temps de transit (MTT, Mean Transit Time) relatif au temps moyen pris par les bulles pour traverser la tumeur et qui informe ainsi sur la tortuosité du système vasculaire intratumoral. Si les premières applications cliniques ont été publiées vers 2010, en particulier pour le diagnostic des tumeurs hépatiques, c'est en 2014 que l'équipe de N. Lassau et al ont montré que l'aire sous la courbe de rehaussement (AUC) était un biomarqueur prédictif de l'efficacité thérapeutique des anti-angiogéniques sur tout type de cancers métastasés avec une valeur seuil de 40% de diminution après un mois de traitement (LASSAU et al., 2014). Aujourd'hui, c'est un des rares nouveaux biomarqueurs à être validé à l'échelle multicentrique. Les prochaines évolutions de l'échographie de contraste sont attendues sur un volet moléculaire avec le développement d'agent de contraste couplées à des traceurs qui se fixent sur des cibles d'intérêt comme les facteurs de croissance vasculaire des protéines VEGF et qui permettrait ainsi de quantifier dans ce cas précis le processus d'angiogenèse. Bien que de nombreuses études précliniques ayant recours à des microbulles ciblées aient été couronnées de succès, la transition de la technique vers la clinique est lente, en raison des nombreuses réactions allergiques provoquées par l'avidine employée pour fonctionnaliser les microbulles. A ce jour, seul deux agents ciblés sont la cible d'essais cliniques pour les tumeurs hépatiques (SonazoidTM, approuvé pour un usage clinique au Japon) et pour les cancers du sein, des ovaires, de la prostate et du pancréas (BR55, Bracco Suisse SA, Geneva, Switzerland) (KÖSE et al., 2020).

Elastographie par IRM et US

L'élastographie est une technique utilisée pour quantifier la dureté des tissus, exprimée en kPa . Deux modalités d'imagerie permettent d'y accéder : l'IRM et l'échographie. Dans les deux cas, il est nécessaire de générer une déformation du tissu à l'échelle locale qu'elle soit manuelle, mécanique ou bien acoustique. En

6. https://qibawiki.rsna.org/images/7/7e/QIBADWIPProfile_as_of_2019-Feb-05.pdf

échographie, l'enjeu est d'autant plus important qu'il s'agit d'une modalité très accessible en comparaison avec l'IRM. La motivation principale pour l'utilisation de l'élastographie dans l'évaluation de la réponse thérapeutique est basée sur le lien connu entre la progression des tumeurs cancéreuses et le renforcement, le durcissement, de la structure des tissus tumoraux en comparaison des tissus sains environnants (PASZEK & WEAVER, 2004). Les applications cliniques de l'élastographie consistent principalement en l'évaluation de la fibrose hépatique et bien que l'échographie soit la plus avancée, l'IRM est également en train d'émerger avec succès comme une modalité non-invasive et une alternative aux biopsies (YIN et al., 2007). Par suite de ces applications, de nombreuses études, principalement en échographie, sont aujourd'hui publiées pour tenter de diagnostiquer les cancers du sein (CHEN et al., 2014) ou de la prostate (WOO et al., 2017) et de caractériser les lésions hépatiques en établissant des seuils de dureté permettant de différencier les lésions bénignes des lésions malignes (GERBER et al., 2017) ou pour identifier les différentes lésions malignes (CHOONG et al., 2017). Dans ce dernier contexte, j'ai eu l'occasion durant ma thèse de travailler avec une radiologue sur la faisabilité d'un biomarqueur de dureté pour l'évaluation de l'efficacité thérapeutique (MOALLA et al., 2020). Bien que l'élastographie montre des résultats prometteurs pour l'évaluation de la réponse dans de nombreuses études pré-cliniques, de nombreux défis techniques, comme la grande disparité des principes de mesure, l'amélioration de la résolution spatiale et de la qualité du signal en IRM ou encore des écarts importants dans la quantification pour les deux modalités, doivent encore être surpassés (ABRAMSON et al., 2016).

Au regard de la multitude des biomarqueurs potentiels qui sont en cours d'étude, et pour compléter celles déjà cités ci-dessus, le tableau suivant (table 1.5) propose un état de l'art non-exhaustif des BI de la réponse thérapeutique, validés pour une utilisation en essai clinique mais non qualifiés pour une utilisation de routine.

Cette liste de BI met en évidence la difficulté de ces nouveaux BI fonctionnelle à succéder aux critères morphologiques RECIST. Cette observation est d'autant plus surprenante pour l'évaluation de la réponse thérapeutique, notamment pour les thérapies ciblées, pour lesquelles RECIST n'est pas adapté, en particulier dans les temps précoces après l'initiation de la thérapie. Le biomarqueur SUV conserve ainsi son avance, probablement parce que très tôt les questions soulevées par la communauté scientifique ont concerné la standardisation des mesures au regard de la diversité des équipements, des protocoles d'injection et d'acquisition. Ainsi, il apparaît en filigrane la problématique de la validation d'un biomarqueur pour une utilisation à l'échelle internationale, qui va être à présent détaillée.

Biomarqueur d'Imagerie (BI)	Modalité	Cancer	Rôle	Décision potentielle	Utilité du Biomarqueur
AUC	DCE- US	RCC, CHC, Colon, Sein, GIST, Mélanome	Réponse	Continuité, arrêt traitement	Une diminution $\geq 40\%$ de l'AUC est prédictif de la réponse thérapeutique à un agent antiangiogénique
^{18}F FDG-SUV	TEP	Tumeurs solides, Œsophage	Pronostic	Modalités de traitement : adjuvante, néo adjuvante, combinaisons	Très grande variabilité inter-individu et inter-tumeurs
^{18}F FLT-SUV	TEP	Tumeurs solides	Pronostic	Modalités de traitement : adjuvante, néo adjuvante, combinaisons	Quantification de la présence d'une haute activité de prolifération cellulaire dans les tissus
ADC, MRF	DWI- IRM	Tumeur solides	Risque, Prédictif	Utilité d'un autre examen diagnostic invasif (biopsie), thérapie adjuvante ou néo adjuvante, Traitement radical ou surveillance active	Risque de récurrence ou de métastase, Corrélation au grade tumoral (cervicale- MRF = ADC + T1 + T2)
					Continue page suivante

Suite de la page précédente

Biomarqueur	Modalité	Cancer	Rôle	Décision potentielle	Utilité du Biomarqueur
ADC	DWI-IRM	Sein, Rectal	Réponse	Continuité, changement traitement	Réponse aux thérapies néo adjuvantes à 12 semaines, meilleure détection (12 semaines) et prédiction de la réponse complète à la thérapie que les séquences classiques
Perfusion T2*	DSC-IRM	Cerveau	Diagnostic, Prédicatif	Type et instant de l'intervention/traitement	Semi-Quantitatif, distinction gliome de grade du II et III
Perfusion T2*	DSC-IRM	Cerveau	Réponse	Application traitement	Différencier les effets du traitement et la progression tumorale
APT	CEST-IRM	Gliome	Diagnostic, Prédicatif	Stratégie thérapeutique	Évaluation de la prolifération et corrélé au grade tumoral et à l'indice Ki67
Flux et volume sanguin	DCE-TDM	Rectal et Poumon	Diagnostic, Prédicatif	Dissection chirurgicale, Radiothérapie adjuvante, thérapie adjuvante	Détection des tumeurs rectales avec ganglions métastatiques avec le flux sanguin et prédiction survie avec la perméabilité dans les poumons
Perfusion et flux sanguin	DCE-TDM	Œsophage	Réponse	Futures traitements	Réponse à la chimiothérapie

Continue page suivante

Suite de la page précédente

Biomarqueur d'Imagerie (BI)	Modalité	Cancer	Rôle	Décision potentielle	Utilité du Biomarqueur
K^{trans} , k_{ep} , V_e	DCE- IRM	Cervical, Endométrial, Rectal, Sein	Risque, Pronostic, Prédicatif	Thérapies néoadjuvantes, adjuvantes ou stratégie multimodales	Réhaussement signal Pronostic PFS et OS pour cervicales ; faible k_{ep} associé à haut-risque de sous-types histologiques pour les cancers endométriaux ; K^{trans} et ADC prédictif des tumeurs du sein à faible risque
K^{trans} , k_{ep} , V_e , V_p	DCE- IRM	Multiple	Réponse	Changement de la stratégie thérapeutique	Evaluation de la réponse thérapeutique à un agent antiangiogénique
Densité UH	TDM	Gastro- intestinal	Réponse	Continuité, changement traitement	Diminution de la densité tumorale $\geq 15\%$ identifie les répondeurs mieux que RECIST en terme de sensitivité et spécificité

TABLE 1.5 – Liste non-exhaustive des Biomarqueurs d'imagerie permettant d'évaluer la réponse thérapeutique. Parmi ces biomarqueurs seule l'AUC en Échographie de contraste est aujourd'hui validée à l'échelle internationale. Les autres sont au stade de l'essai clinique et ne sont pas utilisés en routine clinique. Adapté de O'CONNOR et al. (2017)

4 Processus de validation : la clef de voute des biomarqueurs d'imagerie

Depuis la découverte des premiers BI à visée diagnostique, de nombreux biomarqueurs potentiels ont fait leur apparition dans la littérature alors que quelques-uns seulement sont utilisés en routine clinique. Parmi toutes ces publications, certaines ont proposé des biomarqueurs prometteurs, et quelques-unes même sont à l'origine d'indicateurs utilisés exhaustivement, mais finalement très peu permettent d'apporter un support d'aide à la décision aux praticiens cliniques. Lorsque l'on s'intéresse comme ici à l'évaluation de la réponse thérapeutique en cancérologie, le nombre de biomarqueurs utilisés en pratique clinique est encore plus faible en comparaison aux biomarqueurs diagnostiques par exemple. Les raisons de ces échecs de translation de la recherche vers la pratique

clinique sont multiples. On retiendra quatre raisons principales.

- La validation clinique d'un biomarqueur de nature probabiliste est difficile car corrélée à la survie elle nécessite d'explorer le devenir des patients sur de longues durées.
- De plus, ces biomarqueurs pronostiques ou prédictifs ne mesurent pas directement des processus biologiques comme l'angiogenèse par exemple ou des caractéristiques biologiques tumorales comme la surexpression d'une protéine. La corrélation avec les données histopathologiques sont alors difficiles et leur signification clinique imprécise. De plus, dans la pratique, la corrélation entre les données d'imagerie et les données histopathologiques est souvent difficile à mettre en œuvre en raison de la correspondance entre les données, de la taille des échantillons et du fait que l'acte d'exérèse implique une unique prise de données histologiques. Pour cette raison, le recours à des études précliniques est essentiel pour explorer les relations entre les modifications en imagerie et les modifications tissulaires pathologiques dû à un traitement (WATERTON & PYLKKANEN, 2012).
- Les biomarqueurs ont prouvé leur efficacité mais ne trouvent pas d'application utile car trop complexes à mettre en œuvre, souvent sous couvert d'une problématique de standardisation de procédure d'acquisition ou d'analyse, ou bien encore lorsque la différence avec l'indicateur de référence est faible. Ils restent généralement indéfiniment confinés à la littérature académique. Le K^{trans} est un exemple de biomarqueur complexe à mettre en œuvre. Bien qu'il ait été largement corrélé aux processus biologiques de l'angiogenèse, sa standardisation reste un sujet de recherche notamment pour la mesure de la Fonction d'entrée Artériel (Arterial Input Function) (AIF) qui constitue une source de variabilité importante des cartes paramétriques. Nous en reparlerons dans le chapitre 4 de ce rapport de thèse.
- Les biomarqueurs sont souvent étudiés et validés à l'échelle monocentrique. Le passage aux études multicentriques soulève la question de leurs sources de variabilité dans le cadre de l'imagerie quantitative avec un besoin de très grande robustesse. Ainsi en 2009, après plusieurs années de l'essor de la TEP, lors des mises à jour des critères de réponse basés sur des informations morphologiques, le groupe du Pr Therasse a estimé qu'il était encore trop tôt pour considérer les informations de ces biomarqueurs comme le SUV en remettant encore en cause leur stabilité, leur standardisation, ou leur faible disponibilité.

Pour permettre d'assurer et également d'accélérer le développement de ces biomarqueurs, des recommandations internationales existent. De nombreux consortiums académiques, sociétés savantes, industriels et cliniciens se sont interrogés sur la manière

dont ces BI, de leur acquisition jusqu'à leur analyse, pouvaient être standardisés et comment la terminologie pouvait être harmonisée. Parmi ces instances on retrouve d'une part la FDA et le NIH qui au travers de l'outil BEST par exemple définissent les rôles de ces biomarqueurs, fournissent un socle commun à toutes les recherches scientifiques (FDA-NIH BIOMARKER WORKING GROUP, 2016). D'autre part, les efforts d'autres acteurs ont permis de formuler des recommandations sur l'acquisition et l'analyse de nombreux BI. On retrouve entre autres, le US National Cancer Institute (NCI)⁷ par le biais du Quantitative Imaging Network (QIN), la Quantitative Imaging Biomarkers Alliance (QIBA)⁸, la European Society of Radiology (ESR)⁹, la International Society for Magnetic Resonance in Medicine (ISMRM)¹⁰, The Image Biomarker Standardization Initiative (IBSI)¹¹ ou encore la European Organisation for Research and Treatment of Cancer (EORTC)¹² et le Cancer Research UK (CRUK). Ces deux derniers acteurs, associés à d'autres experts du domaine, ont permis la création d'une feuille de route qui sert aujourd'hui de référence au processus de développement d'un BI. Cette feuille de route est largement réemployée par un consensus qui a été établi en 2017 par 78 auteurs réunis sous l'égide de James O'Connor. Intitulé Imaging biomarker roadmap for cancer studies (O'CONNOR et al., 2017). Nous revenons désormais sur les deux éléments qui constituent ce consensus : la feuille de route évoquée précédemment et les recommandations associées qui en découle.

Concernant la feuille de route, présentée figure 1.5, son objectif est de représenter le processus de développement des BI. Elle peut être résumée selon 3 niveaux de validations distincts :

1. la validation biologique qui valide les performances cliniques du biomarqueur,
2. la validation technique qui s'assure que le BI peut être mesurée précisément et avec robustesse,
3. la validation économique qui doit permettre de juger le compromis bénéfices santé face au coût économique vis-à-vis des biomarqueurs existants.

Le franchissement de ces différents niveaux permet une validation progressive du BI. Il s'agit tout d'abord de le valider pour une utilisation dans des essais cliniques, d'abord via des études rétrospectives, ou bien monocentriques ou encore à l'échelle de quelques centres. C'est dans un second temps qu'intervient l'objectif de qualifier le biomarqueur pour une utilisation clinique à large échelle et cela passe par la réalisation d'études prospectives multicentriques à l'échelle internationale.

7. <https://www.cancer.gov/>

8. <https://www.rsna.org/research/quantitative-imaging-biomarkers-alliance>

9. <https://www.myesr.org/>

10. <https://www.ismrm.org/>

11. <https://theibsi.github.io/>

12. <https://www.eortc.org/>

Il faut noter que cette feuille de route diffère pour les BI par rapport aux biomarqueurs biologiques. En effet, pour l'imagerie, les étapes de validations techniques et cliniques/biologiques se déroulent parallèlement et non séquentiellement. De plus, une validation technique avancée s'effectue dans la plupart des cas tardivement dans le processus, à savoir une validation de la reproductibilité multicentrique et multi-machines. En effet, lors de la réalisation d'études multicentriques, les difficultés techniques et administratives inhérentes à une collaboration à large échelle imposent généralement un délai supplémentaire à cette validation technique avancée. Enfin, les validations cliniques définitives, i.e. lorsque les BI sont comparés à la survie (OS, PFS) sont généralement repoussées dans l'attente d'une validation technique suffisamment avancée pour permettre de larges essais cliniques.

Concernant les recommandations, le second élément de ce consensus, il s'agit d'une liste de 14 recommandations qui sont proposées pour assurer une bonne stratégie de développement. Elles doivent permettre de surmonter les obstacles fréquemment rencontrés pour le passage d'un BI potentiel à un BI utilisé en routine clinique. Ces recommandations concernent le financement des études et leur publication, leur validation technique, biologique, et clinique, leur qualification définitive et enfin des recommandations sur le plan économique relatives à l'efficacité du biomarqueur. Ces recommandations permettent également d'identifier, de quantifier, de documenter et de rendre public les limitations rencontrées lors du développement. Parmi l'ensemble des recommandations, nous listons ci-dessous celles qui nous semblent essentielles et les plus critiques à mettre en place en recherche clinique.

Recommandations de validation technique

- *Recommandation N°3* : Tout comme pour la communauté des biomarqueurs biologiques, les centres et laboratoires d'imagerie doivent être identifiés comme compétent pour la mesure du BI. Cela peut passer par une accréditation. Cette approche, adoptée notamment par la communauté TEP, permet de garantir une calibration constante des équipements pour une quantification standardisée entre les différentes machines et les différents sites. Le report de résultat devrait de la même manière être standardisé comme c'est le cas pour les critères RADS.
- *Recommandation N°4* : La standardisation et l'harmonisation des pratiques doit passer par des lignes directrices mises à jour et revues régulièrement. En cause, l'évolution technologique permanente (p. ex. mise à jour logiciel ou composants) qui peut impliquer des déviations dans les BI. Des modèles statistiques adéquats, comme une régression linéaire multivariée par exemple, doivent être utilisés pour s'ajuster à ces déviations dans les BI au cours des études.
- *Recommandation N°5* : La précision des BI doit être démontrée précocement

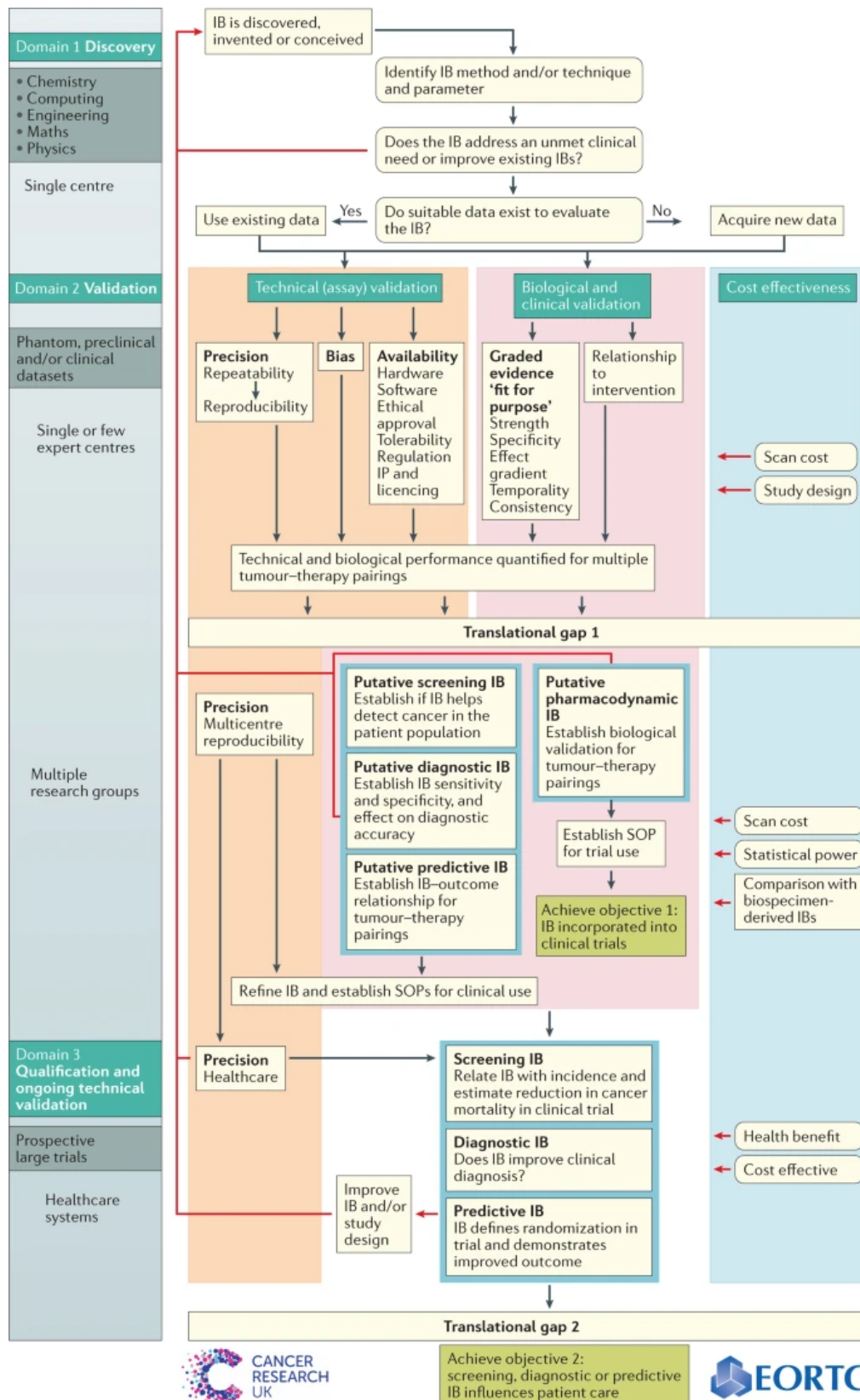


FIGURE 1.5 – Feuille de route détaillée du processus de validation d'un biomarqueur d'imagerie (BI). Le processus de développement d'un BI diffère de celui des biomarqueurs biologiques. Pour l'imagerie les étapes de validations techniques et biologiques/cliniques peuvent se dérouler parallèlement plutôt que séquentiellement et des validations techniques ont généralement lieu à la fin du processus comme la la reproductibilité multi-machines. Image reproduite de Cancer Research UK <http://www.cancerresearchuk.org/>

par des études de répétabilité monocentrique et des études de reproductibilité impliquant quelques centres. Cette recommandation est d'autant plus importante lorsqu'il s'agit de qualifier un BI de l'efficacité thérapeutique.

- *Recommandation N°6* : Lorsque le BI est suffisamment qualifié pour être utilisé en clinique, son utilisation en routine dépend principalement d'une étude multicentrique de grande ampleur (à l'échelle européenne par exemple) pour valider la reproductibilité pour différentes machines, différents agents de contrastes, différents traceurs, différents logiciels. En cas de déviation des résultats, au mieux une calibration commune ou à minima une prise en compte statistique de ces déviations doit être envisagée.

Recommandations de validation Biologique ou Clinique

- *Recommandation N°8* : La validation clinique est relativement tardive dans le processus de développement des BI en comparaison des biomarqueurs biologiques. Les études pré-cliniques ne sont donc pas à négliger et sont fortement encouragées. En effet, elles permettent de fournir des données très contrôlées pour permettre d'analyser les corrélations entre le BI et la pathologie, entre le BI et les effets d'une intervention thérapeutique.
- *Recommandation N°9* : Les méthodes de comparaison pour corréler les résultats d'anatomopathologie et les BI doivent être améliorées. Les méthodes actuelles échouent généralement car elles ne prennent pas en compte les différences d'hétérogénéité spatiale et d'échelles entre les deux mesures. La validation biologique ne peut être que limitée lorsque sont comparées des données d'imagerie (une coupe ou un volume) à des données d'anatomopathologie (simple section très fine (qq μm) en raison d'un alignement non parfait des données. Il est ainsi fortement recommandé d'avoir recours à des données 3D pour une meilleure colocalisation entre les mesures et réduire le biais d'échantillonnage.

Recommandation de rentabilité (balance cout-efficacité)

- *Recommandation N°12* : Le recours à l'imagerie est souvent perçu, parfois à tort, comme un examen coûteux. Dans ce contexte, les études à grande échelle devraient prendre en compte une analyse de rentabilité par rapport aux techniques concurrentes. De plus, les BI potentiels devraient dans la mesure du possible, être développé à partir d'exams radiologiques existants, validés et d'ores et déjà fortement répandus (US, IRM, CT par exemple).

L'imagerie clinique a considérablement transformée la médecine contemporaine et les BI ont un énorme potentiel dans l'amélioration de la pratique clinique en cancérologie.

Néanmoins, ils doivent faire l'objet d'une validation rigoureuse, véritable défi pour les équipes de recherche. Aujourd'hui ce défi est d'autant plus grand qu'un nombre en croissance exponentiel de nouveaux BI voit le jour chaque année visant à caractériser l'hétérogénéité tumorale. Avant de les présenter en détail, la notion d'hétérogénéité tumorale va être introduite dans la partie suivante.

5 L'hétérogénéité tumorale comme biomarqueur d'imagerie

Les connaissances actuelles des tumeurs, permettent de les définir comme des systèmes hétérogènes, c'est-à-dire, qui manquent d'unité car composés d'éléments de natures différentes. L'hétérogénéité tumorale possède de multiples aspects où il coexiste à la fois i) des variations inter-patients pour lesquels un même type de cancer ne présente pas la même réponse thérapeutique ; ii) une variation inter-tumorale au sein d'un même patient avec une diversité dans les caractéristiques génétiques, il peut s'agir de différentes lésions primaires, ou de lésions primaires et métastatiques ; iii) et enfin une variation intratumoral impliquant des différences entre les cellules d'une même tumeur. On s'intéresse par la suite à cette dernière hétérogénéité.

Au cours des dernières décennies, les connaissances des mécanismes de développement tumoral ont évolué. D'une vision simpliste avec un ensemble de cellules cancéreuses et leurs gènes associés s'accumulant, la compréhension de la genèse tumorale s'explique dorénavant comme un ensemble de processus complexes qui combinés favorisent l'apparition de tumeurs. Ainsi, l'évolution d'une cellule saine vers une cellule maligne est un processus multi-étapes. C'est en 2000 puis 2011 que Hanahan et Weinberg, conceptualisent le processus et les mécanismes de développement tumoral à l'aide de grandes caractéristiques (angl. hallmarks) (figures 1.6,1.7). Elles sont liées à la fois aux cellules cancéreuses mais également au micro-environnement tumoral et synthétisent l'ensemble des connaissances dont nous disposons aujourd'hui concernant le développement tumoral (HANAHAN & WEINBERG, 2000, 2011). Ces mécanismes de développement sont nombreux et possèdent un ordre d'apparition variable parmi le spectre de types et sous-types de cancer et sont donc à l'origine d'une grande hétérogénéité des cellules tumorales qui est présente évidemment d'une tumeur à l'autre mais également au sein d'une même tumeur. Ce dernier point est une limite considérable pour la fiabilité des biopsies tissulaires, ces prélèvements à l'échelle quasi-ponctuelle, sont effectués entre une et quatre fois dans la tumeur. Gerlinger et al. ont montré dans ce contexte que les biopsies ne représentent qu'une minorité des aberrations génétiques présentes dans la tumeur (GERLINGER et al., 2012)). L'imagerie, par sa capacité à explorer le tissu dans son ensemble, s'impose comme un outil indispensable pour prendre en compte cette hétérogénéité tumorale pour deux types d'application. La

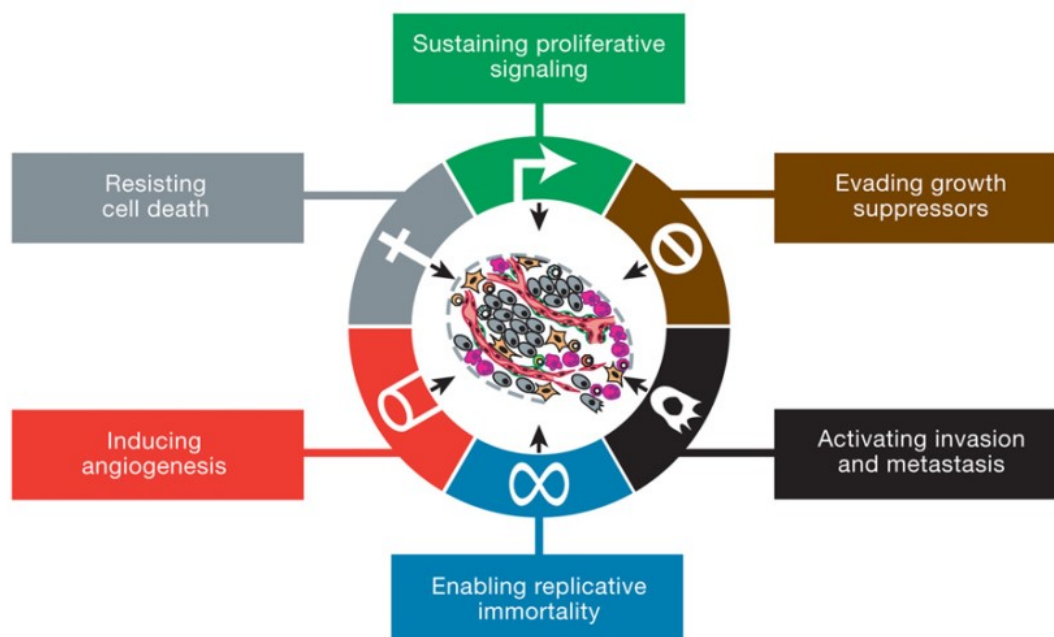


FIGURE 1.6 – Les "Hallmarks" du cancer sont les grandes caractéristiques identifiées par Hanahan et Weinberg en 2000. Ces caractéristiques sont distinctes et nécessaires à la croissance tumorale et à la dissémination métastatique. Ces caractéristiques acquises par les cellules tumorales sont les capacités à : proliférer, échapper aux suppresseurs de croissance, résister à la mort cellulaire, se répliquer à l'infini, induire l'angiogenèse, envahir les tissus et métastaser.

première visera à guider ou remplacer les biopsies invasives par l'utilisation de BI validés. La seconde serait d'intégrer au critère RECIST la caractérisation de l'hétérogénéité d'une image, selon la nature physique du système d'acquisition, pour suivre l'efficacité thérapeutique. C'est dans ce second contexte que se situent les travaux de cette thèse.

5.1 Prise en compte de l'hétérogénéité tumorale dans les nouvelles stratégies thérapeutiques

En raison des cellules de différents génotypes et phénotypes¹³ qui composent la tumeur, l'hétérogénéité intra-tumorale se situe à plusieurs niveaux avec une hétérogénéité génétique, une hétérogénéité cellulaire de l'immunité, une hétérogénéité métabolique et vasculaire avec une disparité dans la distribution d'oxygène et de nutriments et enfin une hétérogénéité dans la progression tumorale. Cette dernière participe à la fois la complexité des tumeurs et à la difficulté de les traiter à l'aide d'une thérapie unique. Malgré les excellents résultats au cours des traitements initiaux,

13. Ensemble des traits observables d'un organisme, le caractère d'une cellule

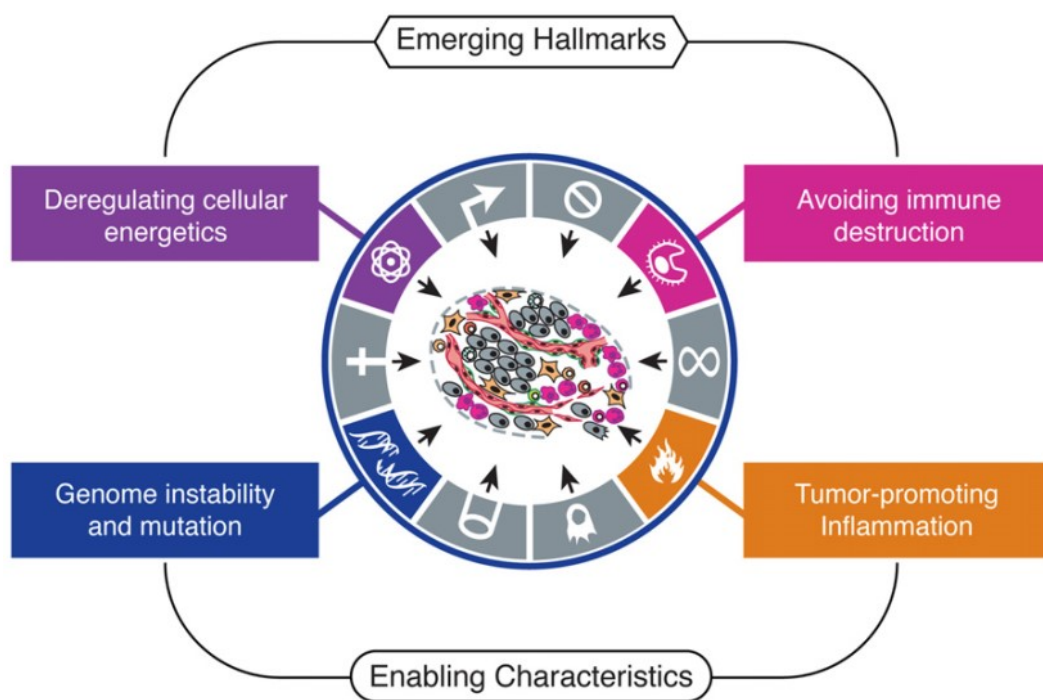


FIGURE 1.7 – Mise à jour en 2009 par Hanahan et Weinberg des connaissances des mécanismes du développement tumoral. La différence majeure par rapport aux Hallmarks de 2000 est la prise en compte des contributions du micro-environnement tumoral qui ajoute deux caractéristiques : l'instabilité génomique et l'état d'inflammation. Deux autres caractéristiques liées aux cellules cancéreuses sont également ajoutées : la dérégulation du métabolisme d'énergie cellulaire et la capacité à échapper au système immunitaire.

une partie non négligeable des cancers développent une résistance à la thérapie au cours du temps. Cette résistance pose un problème majeur dans la lutte contre le cancer et sa guérison complète.

Les différents mécanismes de résistance peuvent être repartis selon deux catégories : la résistance innée ou intrinsèque et la résistance acquise. La première n'est pas conditionnée par le traitement et est inhérente aux caractéristiques de la tumeur, alors que la seconde est induite par la thérapie. Ces résistances peuvent être le résultat de plusieurs causes et nous nous focalisons ici sur l'une d'entre elle, l'hétérogénéité intra-tumorale avec des cellules de sensibilités variables selon les agents thérapeutiques. Le traitement initial permet ainsi de détruire une portion de la tumeur alors que les cellules cancéreuses moins sensibles survivent. Ces cellules résistantes vont ainsi pouvoir proliférer, et éventuellement créer des métastases qui seront insensibles au traitement initial. C'est dans ce contexte que les nouvelles stratégies combinatoires ont été introduites en associant à la fois les chimiothérapies qui endommagent de manière non spécifique toutes les cellules à la croissance rapide et les thérapies ciblées qui s'attaquent aux voies de signalisation tumorale en relations avec les caractéristiques définies précédemment.

Les stratégies thérapeutiques combinatoires existent depuis 1965 lorsque James Holland et Emil Freireich démontrent pour la première fois la possibilité de combiner deux chimiothérapies pour lutter contre les leucémies aigues (FREI et al., 1965) en réduisant la charge tumorale et en prolongeant la survie de patients pédiatriques. Aujourd'hui, la combinaison de différentes stratégies thérapeutiques, deux ou plus, dans le but de cibler spécifiquement les voies de signalisation de développement ou de prolifération tumorale est un enjeu majeur de la lutte contre le cancer. Les thérapies combinatoires entravent simultanément le métabolisme des cellules cancéreuses en différents points pour améliorer l'efficacité thérapeutique, tout en réduisant les effets secondaires et les dosages et sont également utiles pour prévenir ou retarder l'apparition de mécanismes de résistances acquises (MOKHTARI et al., 2017). Ces thérapies combinées incitent au développement d'indicateurs parmi lesquels les BI permettraient à terme de classifier la réponse des patients selon les différents traitements (MASUI et al., 2013). Ces BI vont être à présent détaillés.

5.2 État de l'art de l'évaluation de l'hétérogénéité tumorale en imagerie

L'un des plus grands défis de la radiologie oncologique est désormais d'identifier et de valider ces nouveaux biomarqueurs potentiels de l'hétérogénéité intra-tumorale. Plus précisément, ils permettraient de renforcer la fiabilité des biomarqueurs existants, principalement morphologiques, et pourraient potentiellement fournir des informations

prédictives pertinentes de manière précoce après le début d'une thérapie (O'CONNOR et al., 2017). Les informations caractéristiques de ces biomarqueurs dépendent du système d'imagerie à l'origine des données quantitatives, et représentent donc différents aspects des hétérogénéités de la tumeur. Ainsi, il peut s'agir de l'hétérogénéité de densité tissulaire en CT, métabolique en TEP, de la vascularisation en IRM ou US, etc...

Pour quantifier cette hétérogénéité intra-tumorale en tant que BI, on distingue deux grandes approches.

La première approche méthodologique est de considérer le volume tumoral comme une seule et unique entité qui forme la Région d'Intérêt (ROI), et d'en extraire des indices. Parmi les méthodes qui permettent de quantifier l'hétérogénéité, on distingue quatre catégories d'indices¹⁴ (ALIC et al., 2014). (a) les méthodes non spatiales qui caractérisent l'hétérogénéité tumorale à l'aide de descripteurs non spatiaux. On retrouve par exemple des indices descriptifs de la fréquence de la distribution des niveaux de gris extraits des histogrammes d'intensités comme la déviation standard, la skewness ou le kurtosis. On parle généralement de statistiques du premier ordre ; (b) des méthodes spatiales qui s'intéressent à la répartition spatiale des niveaux de gris. On parle généralement de statistiques du second ordre ou d'analyse de texture décrite dès 1981 pour des applications en imagerie ultrasonore (LERSKI et al., 1981). Ces méthodes se réfèrent à un ensemble de méthodes mathématiques qui évaluent la localisation et l'intensité des niveaux de gris pour calculer des paramètres de texture qui fournissent une mesure de l'hétérogénéité spatiale. Ces indices sont basés sur la génération de matrices qui permettent de générer des paramètres comme le contraste, ou l'homogénéité ; (c) des analyses fractales qui consistent en des mesures statistiques qui ne tiennent pas compte des échelles et qui reflètent la complexité d'une forme et sa capacité à occuper l'espace ; (d) l'ensemble des algorithmes de traitement d'image qui permettent d'extraire des paramètres de textures mais qui ne sont pas basés sur des matrices. Ces algorithmes font généralement usage d'une transformation dans un autre domaine comme les filtres de Fourier ou les ondelettes. Grâce à la revue d'Alic et al datant de 2014 qui parcourt un total de 209 publications, toutes modalités confondues, la montée en puissance des études usant de ces différentes catégories pour quantifier l'hétérogénéité tumorale est impressionnante, que cela soit à visée diagnostique, de suivi de l'efficacité thérapeutique, pronostique ou prédictive. En effet, on dénombre avant 2007, une moyenne annuelle de publications inférieure à 10 puis une progression exponentielle année après année pour atteindre 66 publications en 2012 et dépasser les 1000 publications annuelles depuis 2019 (GUIOT et al., 2021). On notera également que parmi toutes ces publications, l'IRM en association avec les indices basés sur les histogrammes d'intensités et la texture des images, est la modalité la plus représentée pour l'évaluation de la réponse

14. Voir annexe pour une définition mathématique plus détaillée de ces indices

thérapeutique et a été très tôt appliquée au cerveau (SCHAD et al., 1993). De plus, seulement 13% de ces études étaient prospectives. Ce dernier point est important. Bien que les études rétrospectives soient nécessaires pour développer, tester et évaluer des BI en lien avec l'hétérogénéité tumorale, le réel objectif est d'évaluer leur performance dans des études cliniques prospectives (ALIC et al., 2014). Ce dernier point, associé au défi de la standardisation déjà évoqué précédemment, nous renvoi une nouvelle fois, 6 ans après la publication de cette revue, au fait qu'aucun BI évaluant l'hétérogénéité tumorale n'ait réussi à succéder aux critères RECIST.

La seconde approche méthodologique, considère à l'inverse le volume tumoral comme un ensemble d'entités aux propriétés distinctes et consiste à extraire des informations à partir de sous-régions tumorales d'intérêt. Le clustering est une technique qui consiste à découvrir des regroupements naturels dans un ensemble de données et qui permet de mettre en évidence ces sous-régions d'intérêt nommées « clusters ». Chaque cluster représente un ensemble de tissus aux propriétés similaires au regard de la modalité d'imagerie employée. Par exemple, à partir d'une carte paramétrique de l'AUC réalisé en échographie de contraste, les pixels regroupés dans un cluster représentent des tissus qui possèdent les mêmes caractéristiques vasculaires en termes de débit et de volume sanguin. Pour réaliser ce clustering, de nombreux algorithmes itératifs existent. Ils reposent tous sur le principe de regrouper des objets, ici des pixels ou des voxels, en se basant sur une définition de similarité entre ces objets et en utilisant l'ensemble des objets à sa disposition. Il n'y a donc pas de notion d'entraînement ou de validation avec le clustering, comme cela pourrait être le cas avec des approches de classification supervisées qui apprennent d'un jeu de données connu pour pouvoir classer les futures données entrantes. Les informations extraites à partir de ces sous-régions d'intérêt peuvent être classées en deux catégories (a) les paramètres statistiques conventionnels qui décrivent la distribution des histogrammes d'intensité de chaque cluster et (b) les paramètres qui décrivent l'organisation des objets dans ces clusters comme le nombre de voxels ou la fraction volumique de ces clusters par rapport au volume complet. En comparaison à l'approche précédente, l'utilisation de la technique de clustering pour évaluer l'hétérogénéité intra-tumorale en clinique est très peu présente dans la littérature. A l'exception de quelques publications sur des problématiques de diagnostic du cancer du sein (CHANG et al., 2012), de prédiction de la rechute (ANDERSEN et al., 2011) ou d'évaluation de la réponse thérapeutique (LONGO et al., 2015), la majorité des études sont précliniques. Les travaux qui seront présentés dans les chapitres suivants s'appuieront justement sur cette approche par clustering.

Quelle que soit l'approche utilisée, sachant que les différentes approches peuvent être combinées, la quantification de l'hétérogénéité des images des tumeurs conduit à générer un nombre d'indices particulièrement important. Convoqué aux différentes modalités d'imagerie des tumeurs, ce nombre de biomarqueurs potentiels multiparamétrique

devient considérable et leur analyse n'est plus possible par les approches traditionnelles. L'avènement des méthodes d'analyses des données massives comme l'IA ou le deep-learning prend ici tout son sens et est à l'origine d'une discipline émergente de la omique appliquée à la radiologie : la radiomique.

La radiomique est une approche quantitative de l'imagerie médicale qui est définie comme un processus illustré figure 1.8, permettant de générer une large quantité d'indices incluant entre autres ceux décrit précédemment et quantifiant l'hétérogénéité, mais également des indices relatifs à la forme de la tumeur et qui peuvent être combinés à des données cliniques ou encore génomique. On parle même de radio-génomique dans ce dernier cas (SALA et al., 2017). La caractéristique principale de la radiomique est qu'elle reconnaît les images non pas comme de simples photos mais comme un ensemble de données complexes qui peuvent être quantifiées objectivement, en opposition à une évaluation visuelle (van TIMMEREN et al., 2020a). La recherche de BI en oncologie au cours des 10 dernières années est majoritairement composée d'analyses radiomiques, quelle que soit la modalité, l'organe cible ou le rôle du biomarqueur (LAMBIN et al., 2017) et l'évolution du domaine est rapide avec aujourd'hui un accent fort sur la reproductibilité et l'interprétation des résultats (LIU et al., 2019) (figure 1.9). Cet essor de la radiomique a conduit également au développement de nombreux logiciels, où à l'intégration dans ceux déjà existants, comme LifEx¹⁵, 3D Slicer¹⁶, MITK¹⁷, pour ne citer qu'eux et qui permettent de faciliter la mise en place du processus de radiomique. Néanmoins, et comme précédemment, malgré l'émergence fulgurante de la méthode dans le domaine, à l'exception de quelques outils commerciaux d'aides au diagnostics (CAD : Computer Aided Diagnosis) marqués CE et/ou approuvés par la FDA pour un usage clinique (van LEEUWEN et al., 2021), aucun de ces outils, la radiomique et l'IA, n'ont permis, à l'heure actuelle de valider des biomarqueurs. Les raisons de cette absence de succès sont multiples et correspondent sensiblement aux recommandations évoquées précédemment : le faible taux d'études prospectives, le manque de standardisation des procédures d'acquisition et de calcul des indices, une recherche réservée aux grands centres de recherche détenteur de grandes bases de données (SHAIKH et al., 2017) et enfin une interprétabilité difficile des résultats. Ces points peuvent être considérés comme limitant l'intégration des BI vers la pratique clinique.

De manière générale, la mesure d'un indice dépend de la nature du système d'imagerie à l'origine des données exploitées, de ses performances instrumentales intrinsèques, des réglages pour la procédure d'acquisition et bien évidemment de la méthodologie pour extraire et calculer cet indice. On peut alors dissocier deux catégories de standardisation : l'une propre à l'acquisition (système, performances, réglages) et la seconde en lien avec

15. <https://www.lifexsoft.org/>

16. <https://www.slicer.org/>

17. [https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_\(MITK\)](https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK))

le traitement et l'analyse des images pour extraire l'indice. La standardisation des pratiques en radiomique n'en est aujourd'hui qu'à ses prémises. Bien que les études reportées s'intéressent principalement à l'intérêt clinique des potentiels BI, peu se sont jusque-là intéressées à la standardisation. SCHWIER et al. (2019) notaient par exemple que dans le cas particulier des applications de la radiomique en IRM du cancer de la prostate, il ne leur avait pas été possible d'identifier un ensemble de recommandations constantes concernant les différentes étapes de traitement ou d'analyse des images. Tout en ayant conscience qu'il s'agit d'un sujet relativement émergent, il faut tout de même noter une vague d'efforts au sein de la communauté (EUROPEAN SOCIETY OF RADIOLOGY et al., 2019). C'est notamment le cas de ZWANENBURG et al. (2020), au travers de l'initiative IBSI (Imaging Biomarker Standardisation Initiative) qui proposent depuis récemment la standardisation de 169 paramètres de radiomique pour l'imagerie scanner. Ces recommandations issues d'une large étude multicentrique doivent désormais être partagées et mises en pratique pour permettre des avancées. Concernant le volet de standardisation de l'acquisition, la communauté médicale est également consciente de sa nécessité. Ainsi, plusieurs études explorent l'impact de la variabilité de l'acquisition sur les mesures des BI (RAI et al., 2020). Au sein de l'équipe, nous sommes convaincus de son importance et dans ce contexte nous avons exploré cette variabilité en IRM (AMMARI et al., 2021) en scanner (CAMELLA et al., 2018) et en US (MOALLA et al., 2020) avec des mesures de répétabilité.

Une seconde limite importante à l'adhésion de la radiomique est son manque d'interprétation aisée en lien avec la pathologie. En effet, le développement rapide du domaine a conduit à l'exploration d'un nombre d'indices croissant, jusqu'à plusieurs centaines, et basés sur un formalisme mathématique peu intuitif. De plus, les analyses mathématiques pour mettre en évidence l'intérêt clinique des BI apparaissent complexes. Ainsi, bien qu'ils aient permis d'améliorer le pouvoir prédictif des résultats, ils ont en contrepartie complexifié l'interprétation clinique et en ont éloigné une possible signification biologique (van TIMMEREN et al., 2020b). À terme, cette limite dans l'interprétabilité des résultats risque de porter préjudice à l'établissement de la radiomique à large échelle clinique. Dans ce contexte TOMASZEWSKI et GILLIES (2021) dans leur revue, s'intéressent aux différentes approches émergentes permettant de réintroduire une signification biologique dans les résultats et qui tentent de combler le gap entre les résultats et leur interprétation.

6 Conclusion

Les progrès en cancérologie s'accompagnent de stratégies thérapeutiques de plus en plus personnalisées avec de nouvelles molécules activant des mécanismes anti-

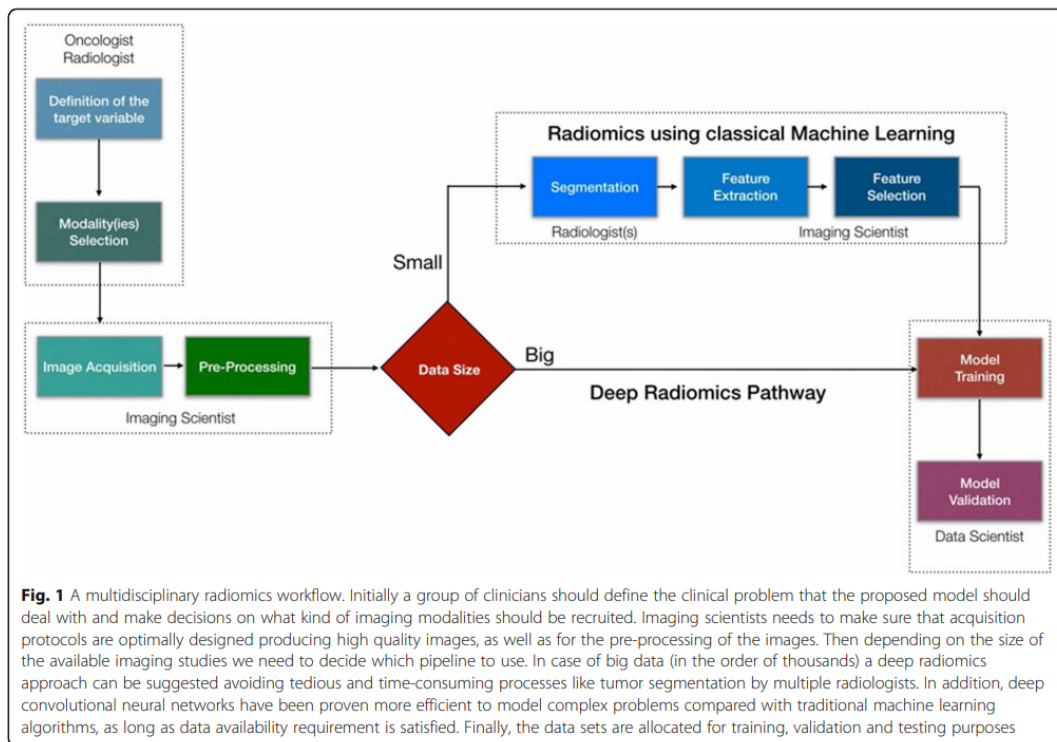


FIGURE 1.8 – Processus de génération d'un biomarqueur d'imagerie en radiomique. Extrait de Papanikolaou et al, 2020

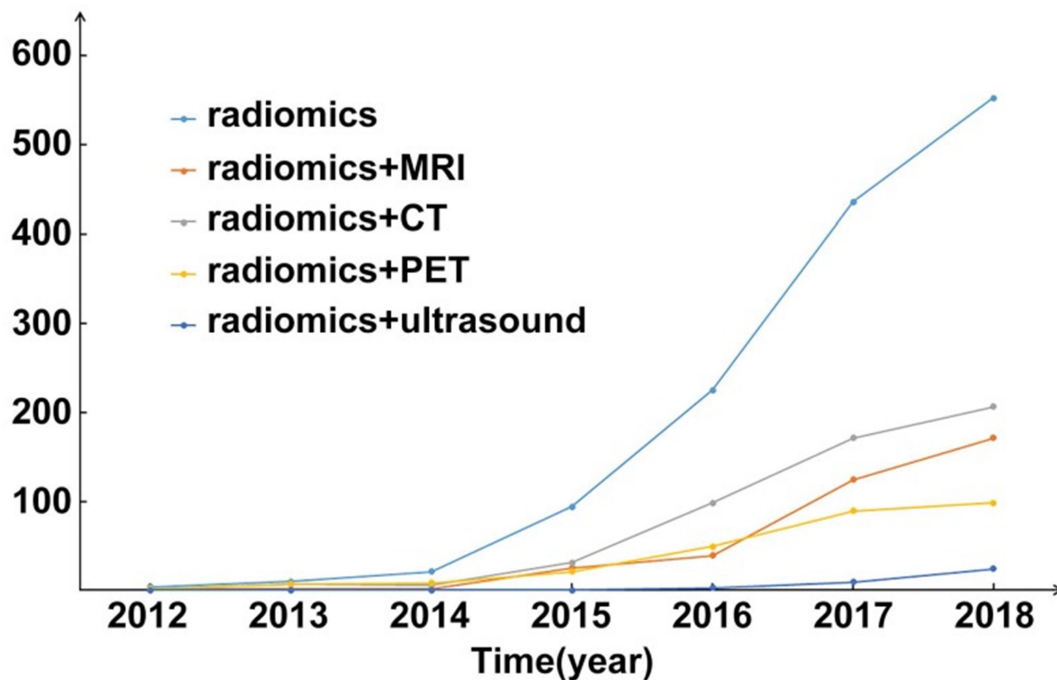


FIGURE 1.9 – Nombre de publications annuelles en radiomique relatives à la cancérologie et à l'hétérogénéité tumorale. Zhenyu, 2019.

tumoraux pas toujours corrélés à la diminution du volume tumoral. Pour guider le praticien parmi ces différents traitements, des nouveaux biomarqueurs sont développés et visent à réduire les temps d'exposition des patients à des traitements inefficaces en sélectionnant la meilleure voie thérapeutique possible. Aujourd'hui, au regard des nombreux développements de biomarqueurs d'imagerie, les attentes sont grandes en termes de pouvoir prédictif ou pronostic de la maladie. C'est en particulier le cas pour caractériser l'hétérogénéité des images des tumeurs qui pourrait être un marqueur de certaines thérapies ciblées prometteuses. Les résultats cliniques associés ne permettent pas encore d'en dégager de validation robuste mais sont révélateurs de certaines limites pour un transfert vers la routine clinique. En particulier la problématique de l'interprétabilité clinique des indices issus d'un formalisme mathématique parfois complexe est un point essentiel à soulever pour une adhésion complète de la méthode par la communauté médicale. Le second repose sur une trop grande variabilité des mesures des indices en raison des approches multicentriques de validation des biomarqueurs et qui mettent en évidence une hétérogénéité des systèmes de mesures, des procédures d'acquisitions, de traitement et d'analyse des images.

Bibliographie

- ABRAMSON, R. G., ARLINGHAUS, L. R., DULA, A. N., QUARLES, C. C., STOKES, A. M., WEIS, J. A., WHISENANT, J. G., CHEKMENEV, E. Y., ZHUKOV, I., WILLIAMS, J. M. & YANKEELOV, T. E. (2016). Mr imaging biomarkers in oncology clinical trials. *Magnetic Resonance Imaging Clinics of North America*, 24(1), 11-29. <https://doi.org/10.1016/j.mric.2015.08.002>
- AH-SEE, M.-L. W., MAKRIS, A., TAYLOR, N. J., HARRISON, M., RICHMAN, P. I., BURCOMBE, R. J., STIRLING, J. J., d'ARCY, J. A., COLLINS, D. J., PITTAM, M. R., RAVICHANDRAN, D. & PADHANI, A. R. (2008). Early Changes in Functional Dynamic Magnetic Resonance Imaging Predict for Pathologic Response to Neoadjuvant Chemotherapy in Primary Breast Cancer. *Clin Cancer Res*, 14(20), 6580-6589. <https://doi.org/10.1158/1078-0432.CCR-07-4310>
- ALIC, L., NIESSEN, W. J. & VEENLAND, J. F. (2014). Quantification of Heterogeneity as a Biomarker in Tumor Imaging : A Systematic Review (C. HATZIS, Éd.). *PLoS ONE*, 9(10), e110300. <https://doi.org/10.1371/journal.pone.0110300>
- AMMARI, S., PITRE-CHAMPAGNAT, S., DERCLE, L., CHOUZENOUX, E., MOALLA, S., REUZE, S., TALBOT, H., MOKOYOKO, T., HADCHITI, J., DIFFETOCQ, S., VOLK, A., EL HAIK, M., LAKISS, S., BALLEYGUIER, C., LASSAU, N. & BIDAULT, F. (2021). Influence of Magnetic Field Strength on Magnetic Resonance Imaging Radiomics Features in Brain Imaging, an In Vitro and In Vivo Study. *Frontiers in Oncology*, 10, 541663. <https://doi.org/10.3389/fonc.2020.541663>
- ANDERSEN, E. K. F., KRISTENSEN, G. B., LYNG, H. & MALINEN, E. (2011). Pharmacokinetic analysis and k-means clustering of DCEMR images for radiotherapy outcome prediction of advanced cervical cancers. *Acta Oncologica*, 50(6), 859-865. <https://doi.org/10.3109/0284186X.2011.578586>
- BENZ, M. R., ALLEN-AUERBACH, M. S., EILBER, F. C., CHEN, H. J., DRY, S., PHELPS, M. E., CZERNIN, J. & WEBER, W. A. (2008). Combined assessment of metabolic and volumetric changes for assessment of tumor response in patients with soft-tissue sarcomas. *J Nucl Med*, 49(10), 1579-1584. <https://doi.org/10.2967/jnumed.108.053694>
- BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R. L., TORRE, L. A. & JEMAL, A. (2018). Global cancer statistics 2018 : globocan estimates of incidence and

- mortality worldwide for 36 cancers in 185 countries. *CA : A Cancer Journal for Clinicians*, 68(6), 394-424. <https://doi.org/10.3322/caac.21492>
- CARAMELLA, C., ALLORANT, A., ORLHAC, F., BIDAULT, F., ASSELAIN, B., AMMARI, S., JARANOWSKI, P., MOUSSIER, A., BALLEYGUIER, C., LASSAU, N. & PITRE-CHAMPAGNAT, S. (2018). Can we trust the calculation of texture indices of CT images? a phantom study. *Medical Physics*, 45(4), 1529-1536. <https://doi.org/10.1002/mp.12809>
- CHANG, Y.-C., HUANG, Y.-H., HUANG, C.-S., CHANG, P.-K., CHEN, J.-H. & CHANG, R.-F. (2012). Classification of breast mass lesions using model-based analysis of the characteristic kinetic curve derived from fuzzy c-means clustering. *Magnetic Resonance Imaging*, 30(3), 312-322. <https://doi.org/10.1016/j.mri.2011.12.002>
- CHEN, L., HE, J., LIU, G., SHAO, K., ZHOU, M., LI, B. & CHEN, X. (2014). Diagnostic performances of shear-wave elastography for identification of malignant breast lesions : a meta-analysis. *Jpn J Radiol*, 32(10), 592-599. <https://doi.org/10.1007/s11604-014-0349-2>
- CHOI, H., CHARNSANGAVEJ, C., FARIA, S. C., MACAPINLAC, H. A., BURGESS, M. A., PATEL, S. R., CHEN, L. L., PODOLOFF, D. A. & BENJAMIN, R. S. (2007). Correlation of Computed Tomography and Positron Emission Tomography in Patients With Metastatic Gastrointestinal Stromal Tumor Treated at a Single Institution With Imatinib Mesylate : Proposal of New Computed Tomography Response Criteria. *JCO*, 25(13), 1753-1759. <https://doi.org/10.1200/JCO.2006.07.3049>
- CHOI, H., CHARNSANGAVEJ, C., FARIA, S. d. C., TAMM, E. P., BENJAMIN, R. S., JOHNSON, M. M., MACAPINLAC, H. A. & PODOLOFF, D. A. (2004). **CT Evaluation of the Response of Gastrointestinal Stromal Tumors After Imatinib Mesylate Treatment** : A Quantitative Analysis Correlated with FDG PET Findings. *American Journal of Roentgenology*, 183(6), 1619-1628. <https://doi.org/10.2214/ajr.183.6.01831619>
- CHOONG, K. L., WONG, Y. H., YEONG, C. H., GNANASUNTHARAM, G. K., GOH, K. L., YOONG, B. K., PONGNAPANG, N. & ABDULLAH, B. J. J. (2017). Elasticity Characterization of Liver Cancers Using Shear Wave Ultrasound Elastography : Comparison Between Hepatocellular Carcinoma and Liver Metastasis. *Journal of Diagnostic Medical Sonography*, 33(6), 481-488. <https://doi.org/10.1177/8756479317733713>
- de GEUS-OEI, L.-F., VRIENS, D., van LAARHOVEN, H. W., van der GRAAF, W. T. & OYEN, W. J. (2009). Monitoring and Predicting Response to Therapy with ¹⁸F-FDG PET in Colorectal Cancer : A Systematic Review. *J Nucl Med*, 50, 43S-54S. <https://doi.org/10.2967/jnumed.108.057224>

- DENNIS, E. R., JIA, X., MEZHERITSKIY, I. S., STEPHENSON, R. D., SCHODER, H., FOX, J. J., HELLER, G., SCHER, H. I., LARSON, S. M. & MORRIS, M. J. (2012). Bone Scan Index : A Quantitative Treatment Response Biomarker for Castration-Resistant Metastatic Prostate Cancer. *JCO*, 30(5), 519-524. <https://doi.org/10.1200/JCO.2011.36.5791>
- DIMITRAKOPOULOU-STRAUSS, A., STRAUSS, L. G., HEICHEL, T., WU, H., BURGER, C., BERND, L. & EWERBECK, V. (2002). The role of quantitative (18)F-FDG PET studies for the differentiation of malignant and benign bone lesions. *J Nucl Med*, 43(4), 510-518.
- EDGE, S. B. & COMPTON, C. C. (2010). The American Joint Committee on Cancer : the 7th Edition of the AJCC Cancer Staging Manual and the Future of TNM. *Ann Surg Oncol*, 17(6), 1471-1474. <https://doi.org/10.1245/s10434-010-0985-4>
- EISENHAUER, E., THERASSE, P., BOGAERTS, J., SCHWARTZ, L., SARGENT, D., FORD, R., DANCEY, J., ARBUCK, S., GWYTHYER, S., MOONEY, M., RUBINSTEIN, L., SHANKAR, L., DODD, L., KAPLAN, R., LACOMBE, D. & VERWEIJ, J. (2009). New response evaluation criteria in solid tumours : Revised RECIST guideline (version 1.1). *European Journal of Cancer*, 45(2), 228-247. <https://doi.org/10.1016/j.ejca.2008.10.026>
- EUROPEAN SOCIETY OF RADIOLOGY, DESOUSA, N. M., ACHTEN, E., ALBERICH-BAYARRI, A., BAMBERG, F., BOELLAARD, R., CLÉMENT, O., FOURNIER, L., GALLAGHER, F., GOLAY, X., HEUSSEL, C. P., JACKSON, E. F., MANNIESING, R., MAYERHOFER, M. E., NERI, E., O'CONNOR, J., OGUZ, K. K., PERSSON, A., SMITS, M., ... ZECH, C. J. (2019). Validated imaging biomarkers as decision-making tools in clinical trials and routine practice : current status and recommendations from the EIBALL* subcommittee of the european society of radiology (ESR). *Insights into Imaging*, 10(1), 87. <https://doi.org/10.1186/s13244-019-0764-0>
- FDA-NIH BIOMARKER WORKING GROUP. (2016). *BEST (Biomarkers, EndpointS, and other Tools) Resource*. Food and Drug Administration (US). Récupérée 17 août 2021, à partir de <http://www.ncbi.nlm.nih.gov/books/NBK326791/>
- FREI, E., KARON, M., LEVIN, R. H., FREIREICH, E. J., TAYLOR, R. J., HANANIAN, J., SELAWRY, O., HOLLAND, J. F., HOOGSTRATEN, B., WOLMAN, I. J., ABIR, E., SAWITSKY, A., LEE, S., MILLS, S. D., BURGERT, E. O., SPURR, C. L., PATTERSON, R. B., EBAUGH, F. G., JAMES, G. W. & MOON, J. H. (1965). The effectiveness of combinations of antileukemic agents in inducing and maintaining remission in children with acute leukemia. *Blood*, 26(5), 642-656.
- GERBER, L., FITTING, D., SRIKANTHARAJAH, K., WEILER, N., KYRIAKIDOU, G., BOJUNGA, J., SCHULZE, F., BON, D., ZEUZEM, S. & RUST, M. F. (2017).

- Evaluation of 2D- Shear Wave Elastography for Characterisation of Focal Liver Lesions. *JGLD*, 26(3), 283-290. <https://doi.org/10.15403/jgld.2014.1121.263.dsh>
- GERLINGER, M., ROWAN, A. J., HORSWELL, S., LARKIN, J., ENDESFELDER, D., GRONROOS, E., MARTINEZ, P., MATTHEWS, N., STEWART, A., TARPEY, P., VARELA, I., PHILLIMORE, B., BEGUM, S., McDONALD, N. Q., BUTLER, A., JONES, D., RAINE, K., LATIMER, C., SANTOS, C. R., ... SWANTON, C. (2012). Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *N Engl J Med*, 366(10), 883-892. <https://doi.org/10.1056/NEJMoa1113205>
- GRIMALDI, S., TERROIR, M. & CARAMELLA, C. (2018). Advances in Oncological Treatment : Limitations of RECIST 1.1 Criteria. *Q J Nucl Med Mol Imaging*, 62(2). <https://doi.org/10.23736/S1824-4785.17.03038-2>
- GUIOT, J., VAIDYANATHAN, A., DEPREZ, L., ZERKA, F., DANTHINE, D., FRIX, A.-N., LAMBIN, P., BOTTARI, F., TSOUTZIDIS, N., MIRAGLIO, B., WALSH, S., VOS, W., HUSTINX, R., FERREIRA, M., LOVINFOSSE, P. & LEIJENAAR, R. T. (2021). A review in radiomics : Making personalized medicine a reality via routine imaging. *Med Res Rev*, med.21846. <https://doi.org/10.1002/med.21846>
- HANAHAN, D. & WEINBERG, R. A. (2000). The Hallmarks of Cancer. *Cell*, 100(1), 57-70. [https://doi.org/10.1016/S0092-8674\(00\)81683-9](https://doi.org/10.1016/S0092-8674(00)81683-9)
- HANAHAN, D. & WEINBERG, R. A. (2011). Hallmarks of Cancer : The Next Generation. *Cell*, 144(5), 646-674. <https://doi.org/10.1016/j.cell.2011.02.013>
- HAYES, D. F., BAST, R. C., DESCH, C. E., FRITSCH, H., KEMENY, N. E., JESSUP, J. M., LOCKER, G. Y., MACDONALD, J. S., MENNEL, R. G., NORTON, L., RAVDIN, P., TAUBE, S. & WINN, R. J. (1996). Tumor Marker Utility Grading System : a Framework to Evaluate Clinical Utility of Tumor Markers. *JNCI Journal of the National Cancer Institute*, 88(20), 1456-1466. <https://doi.org/10.1093/jnci/88.20.1456>
- HSIEH, J. (2015). *Computed tomography : principles, design, artifacts, and recent advances* (Third edition). SPIE.
- KHALIL, M. M. (Éd.). (2017). *Basic Science of PET Imaging*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-40070-9>
- KÖSE, G., DARGUZYTE, M. & KIESSLING, F. (2020). Molecular ultrasound imaging. *Nanomaterials*, 10(10), 1935. <https://doi.org/10.3390/nano10101935>
- KOSTAKOGLU, L. & CHESON, B. D. (2014). Current role of FDG PET/CT in lymphoma. *European Journal of Nuclear Medicine and Molecular Imaging*, 41(5), 1004-1027. <https://doi.org/10.1007/s00259-013-2686-2>
- LAMBIN, P., LEIJENAAR, R. T., DEIST, T. M., PEERLINGS, J., de JONG, E. E., van TIMMEREN, J., SANDULEANU, S., LARUE, R. T., EVEN, A. J., JOCHEMS, A., van WIJK, Y., WOODRUFF, H., van SOEST, J., LUSTBERG, T., ROELOFS, E., van

- ELMPT, W., DEKKER, A., MOTTAGHY, F. M., WILDBERGER, J. E. & WALSH, S. (2017). Radiomics : the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*, 14(12), 749-762. <https://doi.org/10.1038/nrclinonc.2017.141>
- LASSAU, N., BONASTRE, J., KIND, M., VILGRAIN, V., LACROIX, J., CUINET, M., TAIEB, S., AZIZA, R., SARRAN, A., LABBE-DEVILLIERS, C., GALLIX, B., LUCIDARME, O., PTAK, Y., ROCHER, L., CAQUOT, L.-M., CHAGNON, S., MARION, D., LUCIANI, A., FEUTRAY, S., ... KOSCIELNY, S. (2014). Validation of Dynamic Contrast-Enhanced Ultrasound in Predicting Outcomes of Antiangiogenic Therapy for Solid Tumors : The French Multicenter Support for Innovative and Expensive Techniques Study. *Investigative Radiology*, 49(12), 794-800. <https://doi.org/10.1097/RLI.000000000000085>
- LERSKI, R. A., SMITH, M. J., MORLEY, P., BARNETT, E., MILLS, P. R., WATKINSON, G. & MACSWEEN, R. N. M. (1981). Discriminant analysis of ultrasonic texture data in diffuse alcoholic liver disease : 1. fatty liver and cirrhosis. *Ultrasonic Imaging*, 3(2), 164-172. <https://doi.org/10.1177/016173468100300203>
- LIU, Z., WANG, S., DONG, D., WEI, J., FANG, C., ZHOU, X., SUN, K., LI, L., LI, B., WANG, M. & TIAN, J. (2019). The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology : Opportunities and Challenges. *Theranostics*, 9(5), 1303-1322. <https://doi.org/10.7150/thno.30309>
- LIVIERATOS, L. (2012). Basic principles of SPECT and PET imaging. In I. FOGELMAN, G. GNANASEGARAN & H. van der WALL (Éd.), *Radionuclide and hybrid bone imaging* (p. 345-359). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02400-9_12
- LONGO, D. L., DASTRÙ, W., CONSOLINO, L., ESPAK, M., ARIGONI, M., CAVALLO, F. & AIME, S. (2015). Cluster analysis of quantitative parametric maps from DCE-MRI : application in evaluating heterogeneity of tumor response to antiangiogenic treatment. *Magnetic Resonance Imaging*, 33(6), 725-736. <https://doi.org/10.1016/j.mri.2015.03.005>
- MA, Y. & GAMAGEDARA, S. (2015). Biomarker analysis for oncology. *Biomarkers Med.*, 9(9), 845-850. <https://doi.org/10.2217/bmm.15.60>
- Manual of diagnostic ultrasound. 1* (2. ed). (2011). World Health Organization.
- MASUI, K., GINI, B., WYKOSKY, J., ZANCA, C., MISCHÉL, P. S., FURNARI, F. B. & CAVENEÉ, W. K. (2013). A tale of two approaches : complementary mechanisms of cytotoxic and targeted therapy resistance may inform next-generation cancer treatments. *Carcinogenesis*, 34(4), 725-738. <https://doi.org/10.1093/carcin/bgt086>

- MILLER, A. B., HOOGSTRATEN, B., STAQUET, M. & WINKLER, A. (1981). Reporting results of cancer treatment. *Cancer*, 47(1), 207-214. [https://doi.org/10.1002/1097-0142\(19810101\)47:1<207::aid-cncr2820470134>3.0.co;2-6](https://doi.org/10.1002/1097-0142(19810101)47:1<207::aid-cncr2820470134>3.0.co;2-6)
- MOALLA, S., GIROT, C., FRANCHI-ABELLA, S., AMMARI, S., BALLEYGUIER, C., LASSAU, N. & PITRE-CHAMPAGNAT, S. (2020). Methodological Study to Investigate the Potential of Ultrasound-Based Elastography and Texture as Biomarkers to Monitor Liver Tumors. *Diagnostics*, 10(10), 811. <https://doi.org/10.3390/diagnostics10100811>
- MOKHTARI, R. B., HOMAYOUNI, T. S., BALUCH, N., MORGATSKAYA, E., KUMAR, S., DAS, B. & YEGER, H. (2017). Combination therapy in combating cancer. *Oncotarget*, 8(23), 38022-38043. <https://doi.org/10.18632/oncotarget.16723>
- NISHINO, M. (2018). Tumor Response Assessment for Precision Cancer Therapy : Response Evaluation Criteria in Solid Tumors and Beyond. *American Society of Clinical Oncology Educational Book*, (38), 1019-1029. https://doi.org/10.1200/EDBK_201441
- O'CONNOR, J. P. B., ABOAGYE, E. O., ADAMS, J. E., AERTS, H. J. W. L., BARRINGTON, S. F., BEER, A. J., BOELLAARD, R., BOHNDIEK, S. E., BRADY, M., BROWN, G., BUCKLEY, D. L., CHENEVERT, T. L., CLARKE, L. P., COLLETTE, S., COOK, G. J., DESOUZA, N. M., DICKSON, J. C., DIVE, C., EVELHOCH, J. L., ... WATERTON, J. C. (2017). Imaging Biomarker Roadmap for Cancer Studies. *Nat Rev Clin Oncol*, 14(3), 169-186. <https://doi.org/10.1038/nrclinonc.2016.162>
- OEHR, P. (s. d.). PET and PET-CT in oncology, 1.
- PADHANI, A. R., HAYES, C., ASSERSOHN, L., POWLES, T., MAKRIS, A., SUCKLING, J., LEACH, M. O. & HUSBAND, J. E. (2006). Prediction of Clinicopathologic Response of Breast Cancer to Primary Chemotherapy at Contrast-enhanced MR Imaging : Initial Clinical Results. *Radiology*, 239(2), 361-374. <https://doi.org/10.1148/radiol.2392021099>
- PADHANI, A. R. & KHAN, A. A. (2010). Diffusion-weighted (DW) and dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) for monitoring anticancer therapy. *Targ Oncol*, 5(1), 39-52. <https://doi.org/10.1007/s11523-010-0135-8>
- PADHANI, A. R., LIU, G., MU-KOH, D., CHENEVERT, T. L., THOENY, H. C., TAKAHARA, T., DZIK-JURASZ, A., ROSS, B. D., VAN CAUTEREN, M., COLLINS, D., HAMMOUD, D. A., RUSTIN, G. J., TAOULI, B. & CHOYKE, P. L. (2009). Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker : Consensus and Recommendations. *Neoplasia*, 11(2), 102-125. <https://doi.org/10.1593/neo.81328>

- PASZEK, M. J. & WEAVER, V. M. (2004). The Tension Mounts : Mechanics Meets Morphogenesis and Malignancy. *J Mammary Gland Biol Neoplasia*, 9(4), 325-342. <https://doi.org/10.1007/s10911-004-1404-x>
- PICKLES, M. D., GIBBS, P., LOWRY, M. & TURNBULL, L. W. (2006). Diffusion changes precede size reduction in neoadjuvant treatment of breast cancer. *Magnetic Resonance Imaging*, 24(7), 843-847. <https://doi.org/10.1016/j.mri.2005.11.005>
- PORTER, K. A. (1957). Effect of homologous bone marrow injections in x-irradiated rabbits. *British Journal of Experimental Pathology*, 38(4), 401-412.
- RAI, R., HOLLOWAY, L. C., BRINK, C., FIELD, M., CHRISTIANSEN, R. L., SUN, Y., BARTON, M. B. & LINEY, G. P. (2020). Multicenter evaluation of MRI-based radiomic features : a phantom study. *Medical Physics*, 47(7), 3054-3063. <https://doi.org/10.1002/mp.14173>
- SABBATINI, P., LARSON, S. M., KREMER, A., ZHANG, Z.-F., SUN, M., YEUNG, H., IMBRIACO, M., HORAK, I., CONOLLY, M., DING, C., OUYANG, P., KELLY, W. K. & SCHER, H. I. (1999). Prognostic Significance of Extent of Disease in Bone in Patients With Androgen-Independent Prostate Cancer. *JCO*, 17(3), 948-948. <https://doi.org/10.1200/JCO.1999.17.3.948>
- SALA, E., MEMA, E., HIMOTO, Y., VEERARAGHAVAN, H., BRENTON, J., SNYDER, A., WEIGELT, B. & VARGAS, H. (2017). Unravelling tumour heterogeneity using next-generation imaging : radiomics, radiogenomics, and habitat imaging. *Clinical Radiology*, 72(1), 3-10. <https://doi.org/10.1016/j.crad.2016.09.013>
- SALAVATI, A., BASU, S., HEIDARI, P. & ALAVI, A. (2009). Impact of fluorodeoxyglucose PET on the management of esophageal cancer. *Nuclear Medicine Communications*, 30(2), 95-116. <https://doi.org/10.1097/MNM.0b013e32831af204>
- SCHAD, L. R., BLÜML, S. & ZUNA, I. (1993). IX. MR tissue characterization of intracranial tumors by means of texture analysis. *Magnetic Resonance Imaging*, 11(6), 889-896. [https://doi.org/10.1016/0730-725X\(93\)90206-S](https://doi.org/10.1016/0730-725X(93)90206-S)
- SCHWIER, M., van GRIETHUYSEN, J., VANGEL, M. G., PIEPER, S., PELED, S., TEMPANY, C., AERTS, H. J. W. L., KIKINIS, R., FENNESSY, F. M. & FEDOROV, A. (2019). Repeatability of multiparametric prostate MRI radiomics features. *Scientific Reports*, 9(1), 9441. <https://doi.org/10.1038/s41598-019-45766-z>
- SERRANO, M. J., GARRIDO-NAVAS, M. C., DIAZ MOCHON, J. J., CRISTOFANILLI, M., GIL-BAZO, I., PAUWELS, P., MALAPELLE, U., RUSSO, A., LORENTE, J. A., RUIZ-RODRIGUEZ, A. J., PAZ-ARES, L. G., VILAR, E., RAEZ, L. E., CARDONA, A. F. & ROLFO, C. (2020). Precision Prevention and Cancer Interception : The New Challenges of Liquid Biopsy. *Cancer Discov*, 10(11), 1635-1644. <https://doi.org/10.1158/2159-8290.CD-20-0466>

- SHAIKH, F. A., KOLOWITZ, B. J., AWAN, O., AERTS, H. J., von REDEN, A., HALABI, S., MOHIUDDIN, S. A., MALIK, S., SHRESTHA, R. B. & DEIBLE, C. (2017). Technical challenges in the clinical application of radiomics. *JCO Clinical Cancer Informatics*, (1), 1-8. <https://doi.org/10.1200/CCI.17.00004>
- SHUKLA-DAVE, A., OBUCHOWSKI, N. A., CHENEVERT, T. L., JAMBAWALIKAR, S., SCHWARTZ, L. H., MALYARENKO, D., HUANG, W., NOWOROLSKI, S. M., YOUNG, R. J., SHIROISHI, M. S., KIM, H., COOLENS, C., LAUE, H., CHUNG, C., ROSEN, M., BOSS, M. & JACKSON, E. F. (2019). Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J. Magn. Reson. Imaging*, 49(7), e101-e121. <https://doi.org/10.1002/jmri.26518>
- STRIMBU, K. & TAVEL, J. A. (2010). What are biomarkers? *Curr Opin HIV AIDS*, 5(6), 463-466. <https://doi.org/10.1097/COH.0b013e32833ed177>
- THERASSE, P., ARBUCK, S. G., EISENHAEUER, E. A., WANDERS, J., KAPLAN, R. S., RUBINSTEIN, L., VERWEIJ, J., GLABBEKE, M. V., van OOSTEROM, T., CHRISTIAN, M. C. & GWYTHYER, S. G. (2000). New Guidelines to Evaluate the Response to Treatment in Solid Tumors. *Journal of the National Cancer Institute*, 92(3), 12.
- TOFTS, P. S. & KERMODE, A. G. (1991). Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. *Magn. Reson. Med.*, 17(2), 357-367. <https://doi.org/10.1002/mrm.1910170208>
- TOMASZEWSKI, M. R. & GILLIES, R. J. (2021). The biological meaning of radiomic features. *Radiology*, 298(3), 505-516. <https://doi.org/10.1148/radiol.2021202553>
- van LEEUWEN, K. G., SCHALEKAMP, S., RUTTEN, M. J. C. M., van GINNEKEN, B. & de ROOIJ, M. (2021). Artificial intelligence in radiology : 100 commercially available products and their scientific evidence. *Eur Radiol*, 31(6), 3797-3804. <https://doi.org/10.1007/s00330-021-07892-z>
- van TIMMEREN, J. E., CESTER, D., TANADINI-LANG, S., ALKADHI, H. & BAESSLER, B. (2020a). Radiomics in medical imaging—“how-to” guide and critical reflection. *Insights Imaging*, 11(1), 91. <https://doi.org/10.1186/s13244-020-00887-2>
- van TIMMEREN, J. E., CESTER, D., TANADINI-LANG, S., ALKADHI, H. & BAESSLER, B. (2020b). Radiomics in medical imaging—“how-to” guide and critical reflection. *Insights into Imaging*, 11(1), 91. <https://doi.org/10.1186/s13244-020-00887-2>
- WAHL, R. L., JACENE, H., KASAMON, Y. & LODGE, M. A. (2009). From RECIST to PERCIST : Evolving Considerations for PET Response Criteria in Solid Tumors. *J Nucl Med*, 50, 122S-150S. <https://doi.org/10.2967/jnumed.108.057307>

- WATERTON, J. C. & PYLKKANEN, L. (2012). Qualification of imaging biomarkers for oncology drug development. *European Journal of Cancer*, 48(4), 409-415. <https://doi.org/10.1016/j.ejca.2011.11.037>
- WOO, S., SUH, C. H., KIM, S. Y., CHO, J. Y. & KIM, S. H. (2017). Shear-Wave Elastography for Detection of Prostate Cancer : A Systematic Review and Diagnostic Meta-Analysis. *American Journal of Roentgenology*, 209(4), 806-814. <https://doi.org/10.2214/AJR.17.18056>
- WORLD HEALTH ORGANIZATION. (1979). WHO handbook for reporting results of cancer treatment. <https://apps.who.int/iris/handle/10665/37200>
- YAMAGUCHI, T., TAKASHIMA, A., NAGASHIMA, K., MAKUUCHI, R., AIZAWA, M., OHASHI, M., TASHIRO, K., YAMADA, T., KINOSHITA, T., HATA, H., KAWACHI, Y., KAWABATA, R., TSUJI, T., HIHARA, J., SAKAMOTO, T., FUKAGAWA, T., KATAI, H., HIGUCHI, K. & BOKU, N. (2020). Efficacy of Postoperative Chemotherapy After Resection that Leaves No Macroscopically Visible Disease of Gastric Cancer with Positive Peritoneal Lavage Cytology (CY1) or Localized Peritoneum Metastasis (P1a) : A Multicenter Retrospective Study. *Ann Surg Oncol*, 27(1), 284-292. <https://doi.org/10.1245/s10434-019-07697-x>
- YIN, M., TALWALKAR, J. A., GLASER, K. J., MANDUCA, A., GRIMM, R. C., ROSSMAN, P. J., FIDLER, J. L. & EHMAN, R. L. (2007). Assessment of Hepatic Fibrosis With Magnetic Resonance Elastography. *Clinical Gastroenterology and Hepatology*, 5(10), 1207-1213.e2. <https://doi.org/10.1016/j.cgh.2007.06.012>
- ZWANENBURG, A., VALLIÈRES, M., ABDALAH, M. A., AERTS, H. J. W. L., ANDREARCZYK, V., APTE, A., ASHRAFINIA, S., BAKAS, S., BEUKINGA, R. J., BOELLAARD, R., BOGOWICZ, M., BOLDRINI, L., BUVAT, I., COOK, G. J. R., DAVATZIKOS, C., DEPEURSINGE, A., DESSEROIT, M.-C., DINAPOLI, N., DINH, C. V., ... LÖCK, S. (2020). The image biomarker standardization initiative : standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*, 295(2), 328-338. <https://doi.org/10.1148/radiol.2020191145>

Chapitre 2

Vers de Nouveaux Biomarqueurs de l'Hétérogénéité Tumorale en Imagerie Paramétrique : Concept, Développement et Caractérisation

Les connaissances évolutives des mécanismes de développement tumoral ont mis en lumière la nécessité de prendre en compte l'hétérogénéité intra-tumorale afin de renforcer le pouvoir diagnostique, pronostique, prédictif ou de suivi de l'efficacité thérapeutique des biomarqueurs d'imagerie. Malgré de nombreux développements dans le domaine, aucun biomarqueur d'imagerie fonctionnelle s'intéressant à l'hétérogénéité n'a aujourd'hui réussi sa translation vers la pratique clinique à large échelle avec pour causes principales un manque de standardisation des méthodes d'acquisition et d'analyse des images et une difficulté d'interprétation des biomarqueurs générés. Dans ce contexte, l'objet de ce travail de thèse était de développer des biomarqueurs d'imagerie dédiés à la quantification de l'hétérogénéité intra-tumorale en mettant l'accent sur leur interprétabilité.

Après avoir défini le concept et les objectifs visés pour le développement de ces nouveaux biomarqueurs potentiels, nous présenterons la méthodologie choisie pour quantifier les métriques correspondantes. Nous détaillerons ensuite chacune des étapes permettant la génération de ces indices. Notre présentation sera plus particulièrement focalisée sur leur interprétation en termes de modifications de l'hétérogénéité intra-tumorale. Enfin, la caractérisation de la méthode et de ses indices par le biais d'une simulation numérique fera l'objet d'une dernière partie.

1 Concept

Comme évoqué précédemment, actuellement les recherches académiques et industrielles considèrent deux approches complémentaires dans la quantification d'un biomarqueur d'hétérogénéité : soit il intègre l'ensemble du volume tumoral, soit il rend compte de l'hétérogénéité de sous-régions intra-tumorales. Dans le premier cas, on retrouve les approches de radiomique et d'intelligence artificielle associées, qui, néanmoins, à ce jour n'ont pas encore produit de biomarqueurs qualifiés pour un usage clinique. Toutefois, les résultats déjà obtenus (BI et al., 2019 ; ZHANG & ZHU, 2018) soulignent la nécessité de développer des indices relatifs à l'hétérogénéité mais qui permettent également une interprétation physiologique. Au-delà des apprentissages de ces domaines en termes de gestion, de traitement et d'analyse des données d'imagerie, beaucoup d'efforts sont encore nécessaires, notamment dans la standardisation des protocoles, des analyses et dans l'interprétation des marqueurs. Dans le second cas des biomarqueurs quantifiant l'hétérogénéité des sous-régions intra-tumorales, les travaux précliniques menés au sein de notre équipe avaient mis en évidence l'intérêt des indices intra-tumoraux dans l'évaluation des traitements anti-angiogéniques (PITRE-CHAMPAGNAT et al., 2015) au travers de la génération de cartes paramétriques. Les travaux de cette thèse se situent dans cette continuité avec l'ambition de développer de nouveaux biomarqueurs facilement interprétables par les radiologues et considérant les deux approches décrites ci-dessus. Plus précisément, le développement de nouveaux biomarqueurs quantitatifs de l'hétérogénéité intra-tumorale a été entrepris selon les critères méthodologiques suivants :

1. Être extraits à partir de cartes paramétriques obtenues en imagerie quantitative fonctionnelle, ceci facilitant l'interprétabilité du nouvel indice qui en découle en lien direct avec l'hétérogénéité de l'image et de la modalité employée.
2. Présenter une grande robustesse pour simplifier le processus de standardisation dans le cas des études multicentriques, multi-machines.
3. Fournir une représentation spatiale plus intuitive que les indices habituellement employés dans ce cadre comme les indices de texture par exemple.

2 Développement d'un nouveau biomarqueur

L'hétérogénéité d'une image, que ce soit celle d'une tumeur ou d'une autre structure, est une notion subjective qui fait appel en partie à notre perception visuelle de celle-ci. Ainsi pour définir un biomarqueur associé et donc quantifier l'hétérogénéité, il nous faut la définir (Li and Reynolds, 1994). Le dictionnaire "Petit Robert" indique



FIGURE 2.1 – Comparer la complexité des images (a) et (b) est subjectivement difficile. da Silva et al. proposent d'étudier leur cartes d'attentions respectives (c) et (d) qui semblent indiquer que (a) est perceptuellement plus simple que (b). Extrait de da SILVA et COURBOULAY (2011).

que l'hétérogénéité représente ce qui est "composé d'éléments de natures différentes". Nous proposons donc une définition opérationnelle de l'hétérogénéité d'une image en distinguant deux niveaux : l'intensité des pixels ou voxels dont elle est constituée, et leur répartition spatiale les uns par rapport aux autres. L'hétérogénéité des intensités sera caractérisée par la distribution de l'histogramme et notre capacité à y identifier un nombre de gammes d'intensités distinctes. L'hétérogénéité spatiale d'une image, sous entendue l'hétérogénéité de la distribution spatiale des pixels, sera caractérisée par la complexité de l'image. Bien que la notion même de complexité ne soit pas moins floue que celle d'hétérogénéité, son mérite est de pouvoir être rattachée aux travaux effectués sur la perception visuelle. La complexité d'une image y est décrite comme la difficulté pour un individu à focaliser son attention sur un point ou une zone resserrée d'une image (da SILVA & COURBOULAY, 2011). Par exemple, la figure 2.1 proposée par da Silva et al, illustre qu'il est possible de classer les images par ordre de complexité et donc en fonction de leur hétérogénéité spatiale. Nous proposerons ainsi un ensemble de métriques permettant de quantifier la complexité des images tout en ayant conscience que la définition opérationnelle que nous proposons pourrait ne pas être en accord avec l'ensemble des points de vue qualitatifs et visuels.

Ainsi, les approches de radiomique proposent notamment de quantifier l'hétérogénéité d'une image en considérant ces deux aspects, l'intensité et la distribution spatiale, imbriqués. C'est en particulier le cas avec les indices de textures définis en fonction du niveau d'intensité et de l'arrangement spatiale. A l'inverse, le principe des biomarqueurs développés au cours de cette thèse repose sur une approche traitant séparément ces deux niveaux d'hétérogénéités. Ainsi, la méthode proposée est constituée de deux étapes :

1. Une étape de traitement d'image pour considérer l'hétérogénéité des intensités des pixels permettant de subdiviser l'image en clusters qui correspondent à différentes régions d'intérêt.
2. Une étape d'analyse pour quantifier l'hétérogénéité spatiale à travers trois métriques

caractéristiques de la complexité d'une image basées sur les propriétés des patches¹ composant un cluster, en particulier leur nombre, leur fraction volumique, et la distance moyenne entre ces patches.

L'ensemble de ces étapes appliquées à une imagerie paramétrique permet ainsi de décrire pleinement l'hétérogénéité d'une image tout en fournissant des indicateurs intuitifs tout au long du processus d'analyse de ces mêmes images.

Nous décrirons dans un premier temps le principe de fonctionnement des algorithmes de clustering employés dans la littérature pour évaluer l'hétérogénéité des intensités (2.1). Dans un second temps, les indices que nous avons retenus pour quantifier la complexité de l'organisation spatiale seront décrits et l'accent sera mis sur leur interprétabilité (2.2). L'ensemble de ces résultats permettra de définir un indice de l'hétérogénéité unique à deux échelles et dont la formulation sera explicitée dans la dernière partie de cette section (2.3).

2.1 Le clustering et l'hétérogénéité des intensités

Le clustering ou « partitionnement de données » est un terme qui représente un ensemble de techniques de classification non-supervisées qui consistent à regrouper des objets ou segmenter des données en plusieurs catégories dans un contexte où ils ne possèdent pas d'étiquettes, c.à.d. que leur véritable appartenance à une catégorie n'est pas connue à l'avance. Ces catégories sont appelées clusters. Un cluster est ainsi formé de plusieurs objets semblables se distinguant des objets présents dans les autres clusters. Plus généralement, le clustering assigne à un cluster des objets qui partagent des caractéristiques communes (CEBECI & YILDIZ, 2015), donc dans notre cas des pixels de niveaux d'intensité similaires.

Les différents algorithmes de clustering reposent sur deux éléments, à savoir la similarité et la représentation des données.

La similarité correspond à la métrique qui est employée pour établir la distance entre deux objets. Cette métrique est choisie en fonction du domaine d'application de l'algorithme de clustering et de l'objectif visé. Deux métriques de similarité parmi les plus répandues sont la distance Euclidienne et la distance Manhattan. La première informe sur le chemin rectiligne le plus court entre deux points. La seconde, également nommée distance « entres-blocs urbains » ou la distance « taxi », est obtenue en mesurant la distance entre deux points en parcourant le chemin sur une grille. Alors que la distance Euclidienne est employée dans la quasi-totalité des domaines scientifiques, la Manhattan trouve généralement des applications pour la navigation ou dans les algorithmes d'apprentissage très volumineux (AGGARWAL et al., 2001).

1. Un patch correspond à un ensemble de pixels connectés spatialement les uns aux-autres.

Concernant les modèles de représentation des données, il en existe quatre types décrits ci-dessous et qui dépendent de plusieurs critères. :

1. Les modèles hiérarchiques : chaque objet est assimilé à un cluster distinct. Il y a donc au départ autant de clusters que d'objets. Les deux clusters les plus semblables sont regroupés et l'opération est répétée jusqu'à que tous les objets soient regroupés dans un seul cluster. Les résultats sont généralement proposés sous la forme d'un dendrogramme permettant de sélectionner le niveau de hiérarchie des clusters à partir duquel les séparer. Ces modèles sont facilement interprétables mais s'adaptent difficilement à de grands jeux de données.
2. Les modèles distributifs : les objets sont répartis en fonction de leur probabilité d'appartenance à une même distribution. Ces modèles souffrent généralement d'un excès d'apprentissage, c'est-à-dire qu'ils ont tendance à considérer le bruit ou les valeurs aberrantes au même niveau que les données utiles.
3. Les modèles de densité : les objets sont attribués à des clusters au regard de leur répartition et de leur regroupement dans l'espace des données.
4. Les modèles centroïdes : chaque cluster est défini par un centroïde, l'équivalent d'un centre de gravité dans l'espace des données et les objets sont attribués au cluster pour lequel la similarité avec le centroïde est la plus importante.

En raison du caractère non-supervisé du processus de clustering, la fiabilité du résultat ne peut pas toujours être vérifiée. Un algorithme ne peut donc pas nécessairement être jugé meilleur qu'un autre. Le choix de la méthode doit ainsi être fait au regard des données et de l'utilisation prévue. Dans le cadre de cette thèse, les « objets » seront représentés par les pixels ou voxels² caractérisés par leur niveau d'intensité et les distances entre objets correspondent aux différences d'intensité. Le choix du modèle est fait au regard des considérations suivantes. Les données d'entrée des algorithmes de clustering sont des vecteurs d'informations sur une échelle, une dimension (1D) de la taille du nombre de pixels. Ce nombre de pixels, dépend d'une part de la taille de la région d'intérêt et de la taille des pixels de l'image. Dans le cas de l'imagerie en cancérologie, la région d'intérêt définie par la tumeur peut s'étendre de quelques pixels à plusieurs milliers. Pour cette raison, les modèles hiérarchiques (1) ne seront pas employés en raison de leur difficile adaptation à de grands jeux de données. De plus, le bruit est un paramètre inhérent de l'imagerie et bien que celui-ci ne se retrouve pas dans les cartes paramétriques générées, ces dernières contiennent fréquemment des valeurs aberrantes ; les modèles distributifs (2) sont donc généralement peu adaptés sans des traitements

2. Nous confondrons au cours de ces explications les termes pixels et voxels qui représentent respectivement les plus petits éléments graphiques distinguables dans une image en 2 dimensions (2D) ou en 3 dimensions (3D).

préalables. Les modèles de densités (3), bien qu'adaptables aux jeux de données 1D, ont été développés dans un premier temps pour des jeux de données à plusieurs dimensions. Enfin, les modèles centroïdes (4) s'adaptent très bien à de larges volumes de données. Néanmoins, la nécessité de connaître à priori le nombre de clusters à former constitue leur principale limitation. Cette contrainte peut toutefois être contournée. En effet, le nombre de clusters à former peut être soit déduit d'hypothèses physiologiques soit estimé par le calcul. Dans le premier cas, l'objectif du clustering conditionne le nombre de clusters à former. Par exemple, lorsque l'objectif est de discriminer les zones vascularisées des non-vascularisées en échographie de contraste par exemple, il fait sens de générer deux clusters. Dans le deuxième cas, lorsqu'il n'y a pas d'à priori physiologiques, différentes méthodes qui seront introduites par la suite existent pour permettre d'estimer le nombre de clusters optimal à former.

Le choix du modèle repose donc finalement sur le type et la taille des données. Généralement, plusieurs modèles peuvent être employés pour un même jeu de données. Dans ce cas, le choix du modèle doit se faire en toute conscience de ses avantages et de ses limitations. Nous faisons ici le choix des modèles centroïdes qui semblent le mieux adaptés à nos données. Ainsi, nous décrivons par la suite deux algorithmes basés sur ce modèle et qui seront employés pour réaliser le processus de clustering. Nous décrivons également la méthodologie employée pour déterminer le nombre de clusters à former.

2.1.1 L'algorithme des K-moyennes (KM)

L'algorithme des K-moyennes (Macqueen J, 1967, Proc Berkely ..), plus largement connu sous son nom anglais « K-means », ou encore la méthode des centres mobiles, est un des algorithmes de clustering les plus répandus. Il s'agit d'un processus itératif qui calcule les K centroïdes $C = \{c_1, c_2, \dots, c_K\}$ à partir des N données d'entrée $x = \{x_1, x_2, \dots, x_N\}$ et le nombre de clusters K à former.

L'objectif de KM est de minimiser la fonction objective de l'erreur quadratique pour les K centroïdes :

$$J(x) = \sum_{j=1}^K \sum_{i=1}^N \|x_i^{(j)} - c_j\|^2 \quad (2.1)$$

avec $x_i^{(j)}$ l'ensemble des objets dans le cluster j .

Les étapes de KM, illustrés figure 2.2, sont les suivantes :

1. Les centroïdes C sont choisis aléatoirement parmi X.
2. Les distances entre chaque objet x_i et les centroïdes c_j sont calculés
3. Assignation : les objets sont assignés au cluster pour lequel la distance avec le centroïde est la plus faible

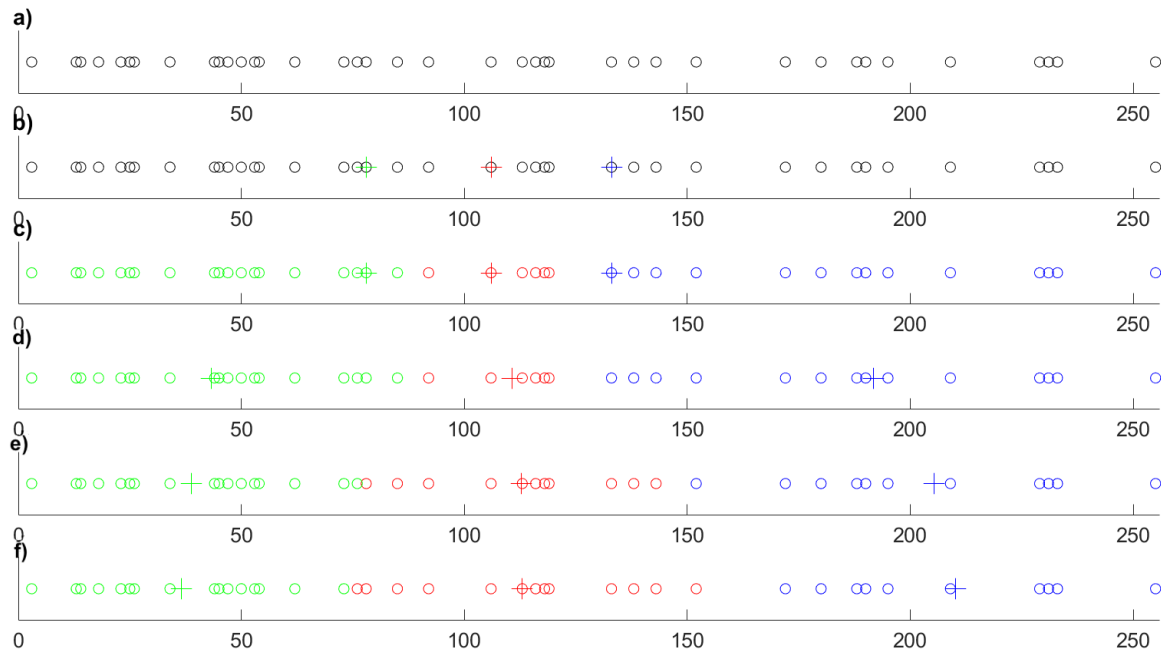


FIGURE 2.2 – Illustration du processus itératif des K-moyennes. a) données d'entrée représentant la valeur de l'intensité de 40 pixels compris entre 0 et 255 (u.a). Pour former $K=3$ clusters. b) 3 pixels sont sélectionnés aléatoirement comme centroïdes initiaux : (craie) de valeurs d'intensités 80(vert), 105 (rouge) et 135 (bleu) . c) Représentation de l'itération 1 ; Chaque pixel est assigné au cluster dont le centroïde est le plus proche. d) Les centroïdes sont mis-à-jour après le calcul de l'intensité moyenne des pixels appartenant aux différents clusters correspondants ; Fin de l'itération 1. e) Résultat final obtenu après 6 itérations correspondants à la répétition des étapes c) et d) jusqu'à ce qu'aucun pixel ne change de cluster d'attribution d'une itération à la suivante.

4. Mise à jour : les centroïdes sont recalculés avec la nouvelle assignation de chaque objet : $c_j = \sum_{i=1}^m \frac{x_i}{m}, 1 \leq k \leq K$
5. Si aucun point n'a été assigné à un nouveau cluster ou que le nombre maximal d'itération est atteint, l'algorithme est arrêté, sinon les étapes 3) et 4) sont répétées.

2.1.2 L'algorithme Fuzzy-C-Means (FCM)

Originellement développé par Bezdek (1981) pour contrer certaines limitations de l'algorithme des K-moyennes, l'algorithme FCM est une extension de KM. La principale différence réside dans le fait que FCM est un algorithme de clustering dont le seuillage est souple au contraire des seuillages durs pour les K-moyennes. En d'autres termes, les objets ne sont pas membres d'un unique cluster mais membre de plusieurs clusters avec un certain degré d'appartenance à chacun. En imagerie, lorsque FCM est employé sans la volonté d'un recouvrement des clusters, les pixels sont attribués au cluster pour lequel leur degré d'appartenance est le plus fort. En pratique FCM diffère très peu de KM dans son processus itératif, avec l'ajout d'un terme de floutage U qui incorpore

les degré d'appartenance des observations à chaque cluster. De ce fait, la fonction à minimiser devient :

$$J(x) = \sum_{j=1}^K \sum_{i=1}^N w_{ij}^m \|x_i^{(j)} - c_j\|^2, m \in [1, \text{inf}] \quad (2.2)$$

ou m est un hyperparamètre de floutage, avec habituellement $m = 2$ et

$$w_{ij}^m = \frac{1}{\sum_{k=1}^K \left(\frac{\|x_i - c_j\|}{\|c_i - c_k\|} \right)^{\frac{2}{m-1}}}, c_j = \frac{\sum_{i=1}^N w_{ij}^m \cdot x_i}{\sum_{i=1}^N w_{ij}^m} \quad (2.3)$$

Les différentes étapes de l'algorithme sont alors exactement les mêmes que pour KM avec l'application de ces équations modifiées.

2.1.3 Détermination du nombre de clusters

Les algorithmes de clustering basés sur des modèles de représentation des données par centroïdes possèdent comme contrainte principale de devoir connaître au préalable le nombre de clusters K à former. De manière générale, dès lors que le partitionnement réel des données n'est pas connu, le nombre de clusters à former peut être déterminé de deux manières. La première consiste à demander à l'utilisateur d'imposer un nombre de clusters jugé optimal avec pour conséquence d'introduire un biais dans les résultats basé sur l'interprétation de l'utilisateur et un manque de reproductibilité des résultats. La seconde consiste à employer des méthodes numériques qualitatives ou quantitatives permettant de déterminer ce nombre optimal. Cette problématique a attiré une attention particulière au cours des vingt dernières années et de nombreuses approches mathématiques ont été proposées (KODINARIYA & MAKWANA, 2013). Nous nous sommes intéressés à deux d'entre elles, la « méthode des coudes » et la « méthode de la silhouette ».

La méthode du coude (angl. Elbow Method) est la méthode la plus ancienne permettant de déterminer K pour un jeu de données (THORNDIKE, 1953). Le principe est basé sur la définition même du clustering qui est de former des clusters de sorte que la somme des variations au sein de chaque cluster soit la plus faible possible. L'idée est de visualiser au travers de la courbe de l'inertie, à partir de quel nombre, l'ajout d'un nouveau cluster n'ajoute pas de nouvelles informations, pouvant même créer une structure contre-intuitive comme illustré figure 2.3. Cette figure, purement illustrative, représente un ensemble de points définis par leurs coordonnées et où le résultat pour un nombre de clusters mal défini a priori provoque la fusion de plusieurs clusters ou la division contre-intuitive d'un cluster. Le K optimal peut alors être défini par la procédure suivante :

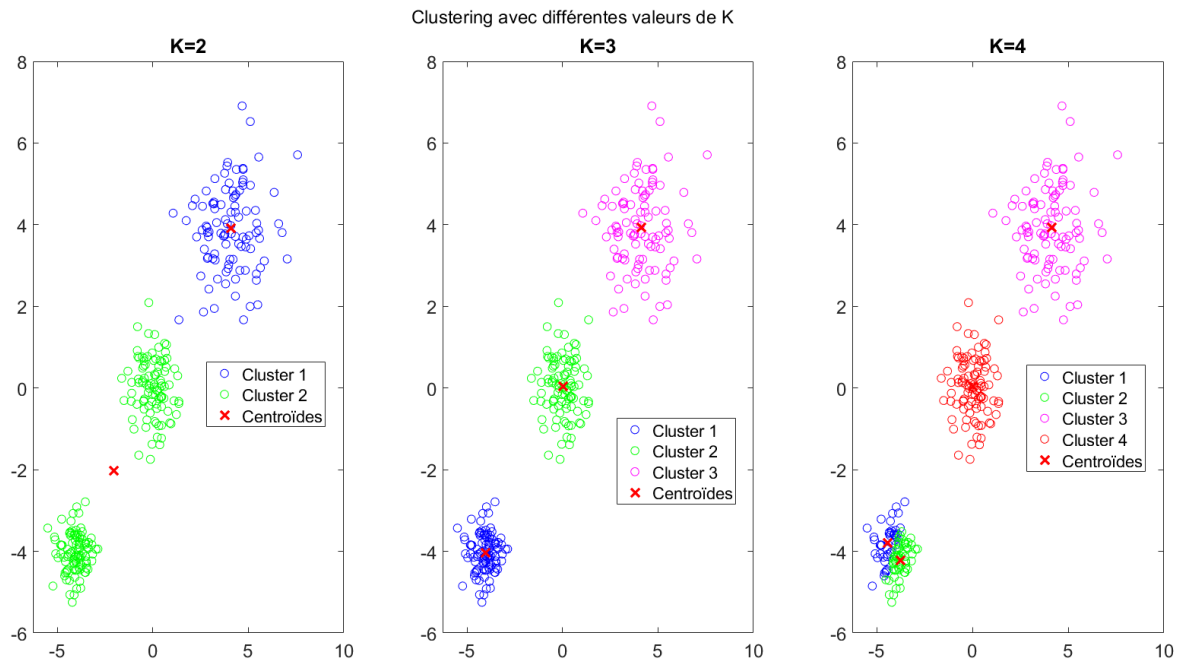


FIGURE 2.3 – Représentation de trois ensembles de points définis par leurs coordonnées géométriques (x,y) (u.a.) et illustration du résultat du clustering de gauche à droite pour $K=2, 3$ et 4 clusters. Pour $K=2$, deux ensembles de points visuellement dissociés sont fusionnés dans le même cluster. Pour $K=4$, un ensemble de points est divisé en deux clusters. $K=3$ permet de représenter le jeu de données originales correctement.

1. Appliquer le processus de clustering aux données pour différentes valeurs de K (de 1 à 10 par exemple).
2. Pour chaque K , calculer la somme des erreurs quadratiques de chaque cluster.
3. Afficher la courbe de la somme des erreurs quadratiques en fonction de K .
4. Identifier le K optimal à partir de la localisation du pli dans la courbe, considérée comme un indicateur.

Toutefois, bien que cette méthode soit facile à illustrer sur des cas d'exemple comme nous venons de le faire, le « coude » ne peut pas toujours être identifié sans ambiguïté (KETCHEN JR. & SHOOK, 1996) rendant cet outil subjectif et peu fiable.

La méthode du coude étant plus une aide à la décision qu'une véritable métrique permettant d'estimer le K optimal de manière objective, nous nous sommes intéressés à la méthode proposée par ROKACH et MAIMON (2005), et employée par ANDERSEN et al. (2011), sur des données d'imagerie. Il s'agit d'une variante de la méthode dite de la « silhouette » KAUFMAN L. et ROUSSEUW P. (1987). Le principe de cette optimisation est d'avoir recours non seulement à la somme des erreurs quadratiques comme pour la méthode précédente mais elle calcule également la similarité entre les différents clusters. Ces deux notions sont respectivement représentées par les métriques

Variance intra-cluster moyenne (VICM), c'est à dire la somme des erreurs quadratiques de chaque cluster que l'on cherche à minimiser et la Distance minimal inter-clusters (DMIC)³ que l'on cherche à maximiser. Les équations pour VICM (2.4) et DMIC(2.5) sont :

$$VICM_K = \sum_{k=1}^K \sum_{x \in C_k} \frac{\|x_i - \mu_k\|}{N_k} \quad (2.4)$$

$$DMIC_K = \frac{K}{\min_{i \neq j} \|\mu_i - \mu_j\|}, \text{ pour } K \geq 2 \quad (2.5)$$

μ_k le K^{eme} centroïde, x_i le i^{eme} élément du cluster C_k , N_k le nombre d'éléments dans le cluster k , K le nombre de clusters

Lorsque K augmente, le nombre d'éléments par clusters diminue ce qui entraine logiquement une diminution de la variance et donc de VICM tandis que DMIC augmente (K augmente tandis que la distance entre les clusters diminue). Le nombre de clusters K optimal est alors défini comme le minimum de la somme de ces métriques normalisées.

Ces deux méthodes, dont l'une constitue une aide visuelle à la décision tandis que l'autre est une métrique de validation, ne sont pas exclusives et sont donc deux outils qui peuvent être utilisés conjointement pour déterminer le nombre optimal de clusters. La figure 2.4 illustre pour le jeu de données employé, la mise en application de ces deux méthodes. Le résultat est ici sans équivoque. En effet, pour ces données illustratives, la séparation entre les clusters est quasi-totale et contrôlée par la génération des points. Néanmoins, pour une application en condition réelle, la séparation des clusters n'est pas toujours aussi franche. De plus, ces méthodes sont employées lorsque le nombre de clusters n'est pas connu a priori et ne sont pas basées sur des hypothèses physiologiques. Ainsi, nous ne possédons aucun critère permettant de justifier l'existence d'un cluster et le recours à ces méthodes constitue un recours objectif à la détermination du nombre de clusters à former.

En conclusion, le processus de clustering, associé à une méthode de détermination du nombre de clusters optimal, permet d'évaluer l'hétérogénéité des intensités de l'image par la formation d'un nombre de clusters représentant au mieux la répartition des intensités. Dans le contexte particulier de l'imagerie en cancérologie, l'application de ce processus à des images paramétriques permet une identification non-supervisée de plusieurs sous-régions tumorales caractérisées par des valeurs similaires du paramètre mesuré. La section suivante est consacrée à l'évaluation de la complexité spatiale de ces sous-régions.

3. En réalité, DMIC est le ratio entre le nombre de clusters et la distance minimal entre les clusters. Nous choisissons ici de conserver le terme distance minimal pour cette métrique pour s'accorder avec la littérature.

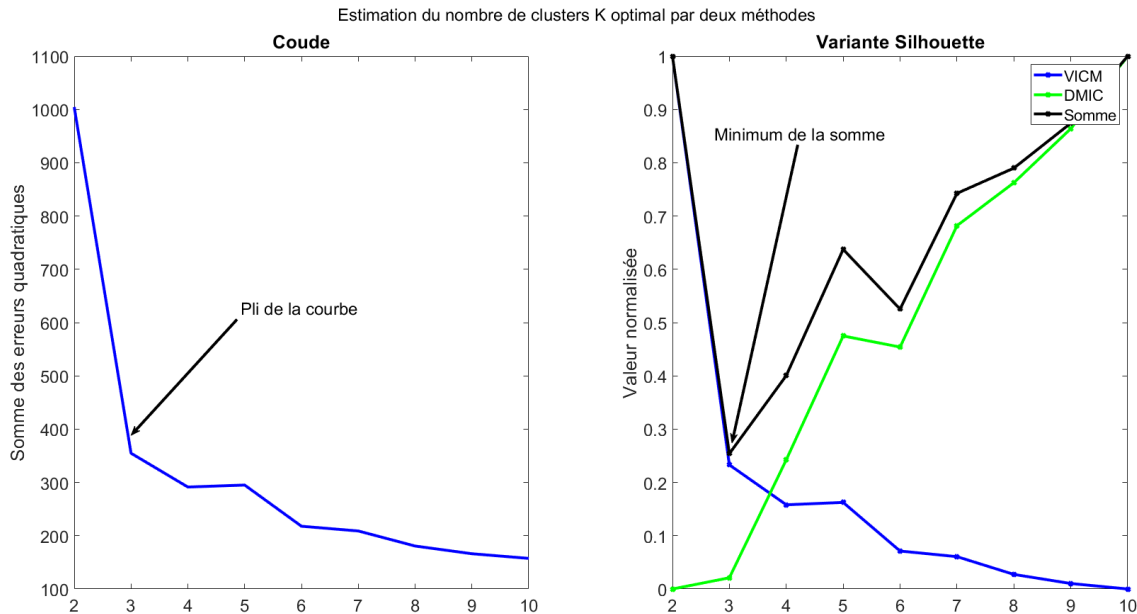


FIGURE 2.4 – Courbes permettant de déterminer le nombre de clusters optimal avec, à gauche la méthode du coude et à droite, la variante de la méthode de la silhouette. Les deux méthodes estiment un nombre de clusters optimal $K=3$. Pour la méthode de la silhouette, les métriques sont normalisées. Les unités sont arbitraires, et les résultats obtenues sont indiqués par des flèches noires.

2.2 Les métriques de complexité de la distribution spatiale

Pour quantifier l'hétérogénéité de la distribution spatiale, nous nous sommes donc appuyés sur une évaluation de la complexité des clusters comme évoqué dans la partie concept de ce chapitre. Pour rappel, la complexité est relative à la difficulté pour un observateur de trouver un point de fixation dans une image. Les métriques qui ont permis de quantifier cela et qui vont être définies ci-dessous offrent selon les critères décrits en première partie de ce chapitre, une représentation intuitive de la complexité.

Les métriques ont ainsi été choisies sur la base de la définition donnée par LI et REYNOLDS (1995), pour décrire l'hétérogénéité spatiale d'une carte catégorique⁴ qui stipule qu'elle « peut être mesurée par la complexité de la composition et de la configuration de ses patchs ». Dans cette définition, un patch correspond à un ensemble de pixels, connectés spatialement les uns aux autres. Un cluster peut donc être composé de multiples patchs. **De cette façon nous pouvons quantifier l'hétérogénéité spatiale définie par la complexité de l'image en se basant sur la composition des patchs caractérisée par leur nombre et leur proportion dans l'image, ainsi que sur la configuration des patchs qui correspond à leur arrangement spatial, leur forme, le contraste entre les patchs voisins, la connectivité entre**

4. Carte paramétrique ou image dans laquelle les pixels appartiennent à une catégorie. Typiquement, une image binaire avec les pixels de valeurs 0 identifiant l'arrière-plan et les pixels de valeur 1 identifiant les pixels d'intérêt.

les patchs similaires et leur anisotropie.

Dans le cadre de la méthodologie développée dans cette thèse, après le processus de clustering, les cartes paramétriques analysées consistent en un ensemble de K cartes catégoriques, K étant le nombre de clusters. Pour décrire la complexité de la distribution spatiale de chacune de ces cartes, trois métriques ont été retenues et nous allons les définir avant d'aborder leur interprétation.

2.2.1 Définitions des métriques de la complexité spatiale

Au regard de la définition précédente, trois métriques ont été sélectionnées pour décrire la complexité spatiale des clusters. Deux sont relatives à la complexité de leur composition : le nombre de patchs et la fraction volumique. La troisième est relative à la complexité de la configuration des patchs du cluster : la distance moyenne entre les patchs. Il s'agit du seul indice retenu parmi ceux relatifs à la configuration, proposés par Li & Reynolds. En effet, les indices qui dépendent du contraste ou du niveau d'intensité ne trouvent pas d'application dans nos clusters construits sur des valeurs d'intensités similaires, la forme de ces patchs ne semble pas non plus une information pertinente au regard du nombre de patchs et de la résolution spatiale offerte par les modalités d'imagerie quantitatives associée à une très faible taille de certains patchs ne recouvrant que quelques pixels. Nous décrivons désormais chacune de ces trois métriques.

Le nombre de patchs

Le nombre de patchs qui ne sont pas connectés spatialement est obtenu à partir d'une analyse en composantes connectées. Il s'agit d'une application algorithmique basée sur la théorie des graphes⁵ au cours de laquelle une étiquette unique est attribuée à chaque patch isolé. Un patch ne peut être connecté à aucun autre patch par le biais d'aucun pixel. Pour effectuer cette opération de comptage, l'algorithme requiert de connaître le degré de liaison à employer pour considérer deux pixels ou voxels voisins, ce degré de liaison est nommé connectivité. En 2D, les pixels peuvent être connectés par leurs faces et par leurs angles, on parle donc d'analyse 4- ou 8-connexions. En 3D, les voxels peuvent être connectés par leurs faces, leurs arêtes et leurs angles, on parle alors d'analyse 6-, 18- ou 26-connexions. Ce principe de connectivité est illustré sur la figure 2.5.

La fraction volumique

La fraction de volume représente le ratio entre le volume occupé par le cluster et le volume total de la ROI. Il est exprimé en pourcentage. La somme de la fraction volumique de tous les clusters d'une ROI est de 100%. Cette métrique représente la contribution de la complexité de l'organisation spatiale du cluster à l'hétérogénéité globale de la ROI.

5. Structures mathématiques utilisées pour modéliser les relations deux-a-deux entre des objets

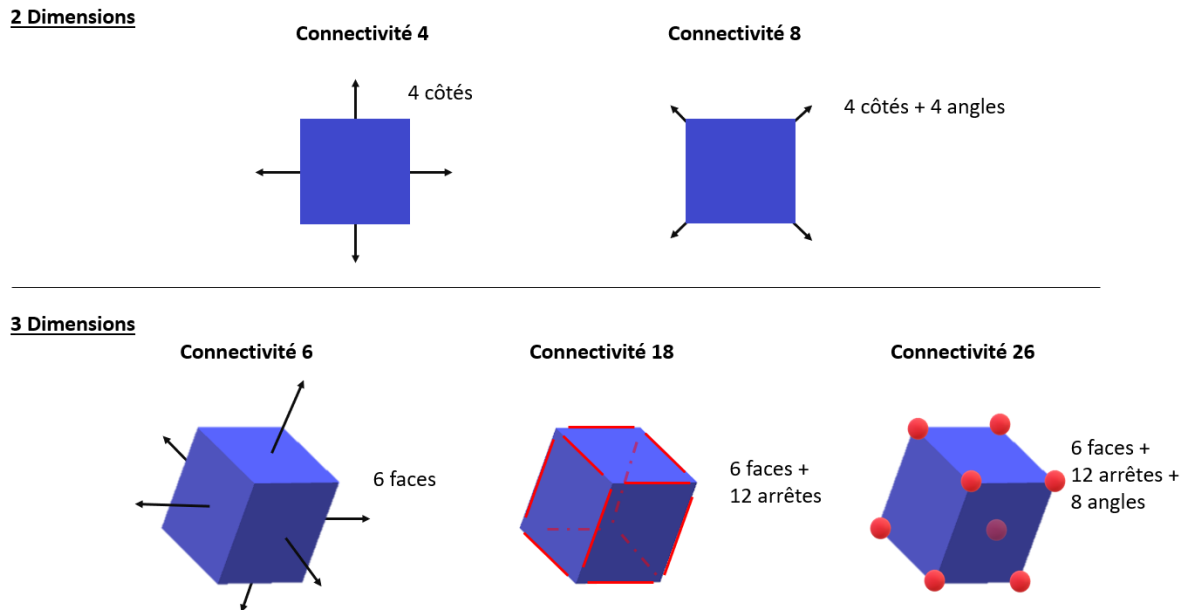


FIGURE 2.5 – Schéma de principe représentant la connectivité entre deux pixels (2D) avec 4 ou 8 connexions et entre deux voxels (3D) avec 6, 18 ou 26 connexions.

La distance moyenne

Cette métrique mesure la séparation spatiale moyenne entre tous les patches d'un cluster identifiés précédemment. La séparation est représentée par la distance euclidienne. Les patches sont représentés par leur barycentre et la distance est calculée pour chaque paire de patches.

2.2.2 Interprétation des métriques

Ces trois métriques ont été sélectionnées pour leur simplicité de mise en œuvre et de compréhension et donc d'interprétation. En effet, si l'on en revient à la relation entre la complexité des images et la difficulté à focaliser la vue sur une zone restreinte de l'image, décrire l'évolution de l'hétérogénéité des images devient relativement intuitif à l'aide de ces métriques. Pour illustrer ces notions, nous représentons à l'aide d'illustrations schématiques, différents cas de figures d'évolution de la complexité dans des images. Ces images générées représentent des cartes paramétriques composées d'un seul cluster, de compositions et de configurations variées.

La figure 2.6 représente une succession d'images générées avec une complexité croissante. En effet, alors que les patches occupent la même fraction volumique de l'image, leur nombre croît. Au fur et à mesure de l'évolution du nombre de patches, la granularité de l'image augmente imposant une focalisation visuelle plus complexe. La complexité de l'image évolue alors au même titre que le nombre de patches.

A l'inverse, pour un nombre de patches constant, la complexité du cluster ne varie

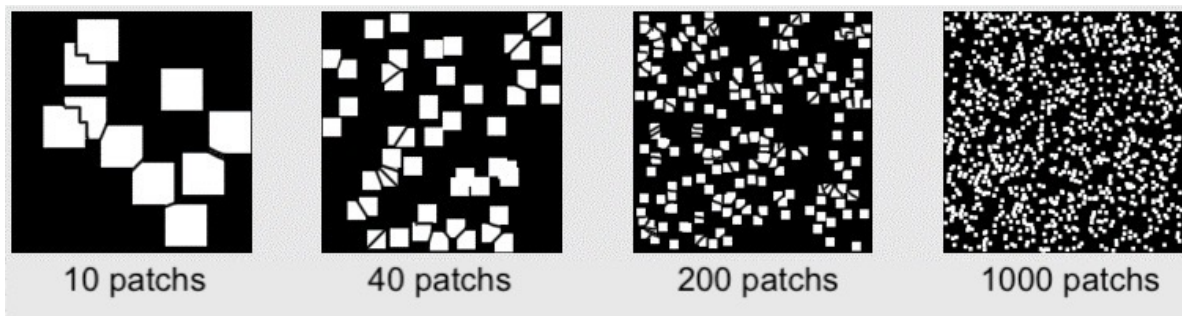


FIGURE 2.6 – Illustration schématique d'un cluster dans une ROI composée d'un nombre de patches croissant de gauche à droite. L'ensemble des pixels du cluster occupe pour chaque image la même fraction volumique. Les pixels appartenant au cluster valent 1 (blanc). La ROI correspond à l'image complète. La complexité de la carte augmente de gauche à droite avec un nombre de patches croissants de 10, 40, 200 et 1000.

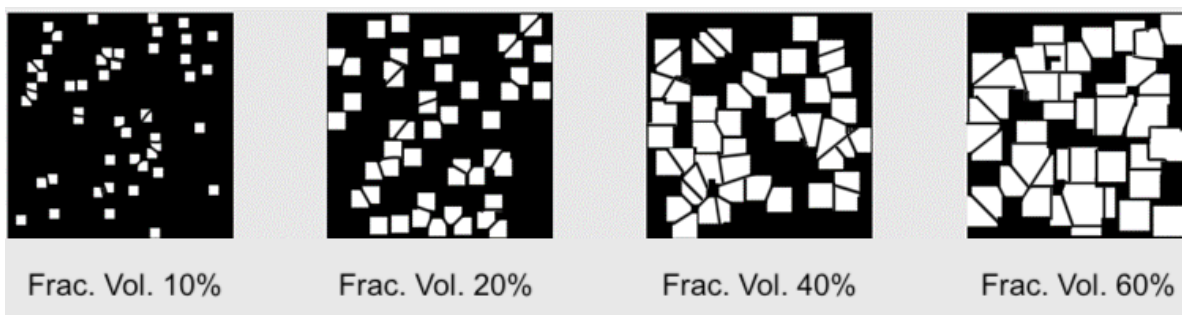


FIGURE 2.7 – Illustration schématique d'un cluster dans une ROI. De gauche à droite, un nombre constant de patches et une fraction volumique croissante avec respectivement une contribution au volume total de l'image et donc à l'hétérogénéité de l'image de 10, 20, 40 et 60%.

pas, comme illustré figure 2.7. Néanmoins, la contribution de la complexité du cluster à l'hétérogénéité globale de l'image croît lorsque la fraction volumique augmente. La combinaison de la complexité d'un cluster et de sa contribution à l'hétérogénéité globale de l'image est illustrée figure 2.8 qui propose une première interprétation de l'utilisation simultanée de ces deux métriques.

L'interprétation de ces métriques est relativement aisée et elle permet d'identifier des différences de composition mais il nous paraissait nécessaire d'introduire un indice relatif à la configuration spatiale des patches. Pour illustrer l'intérêt de cet indice, la figure 2.9 contient trois exemples présentant visiblement des configurations (ou organisations spatiales) différentes mais pour lesquelles les deux premières métriques présentées ci-dessus ne permettent pas de quantifier les différences. C'est pourquoi la configuration est prise en compte par un indice qui mesure des distances moyennes entre les patches. La figure 2.10 illustre le principe du calcul de cette métrique. L'image A implique une configuration plus complexe que l'image B. En effet, en se basant de nouveau sur la notion de perception visuelle de la complexité, une augmentation de la distance implique

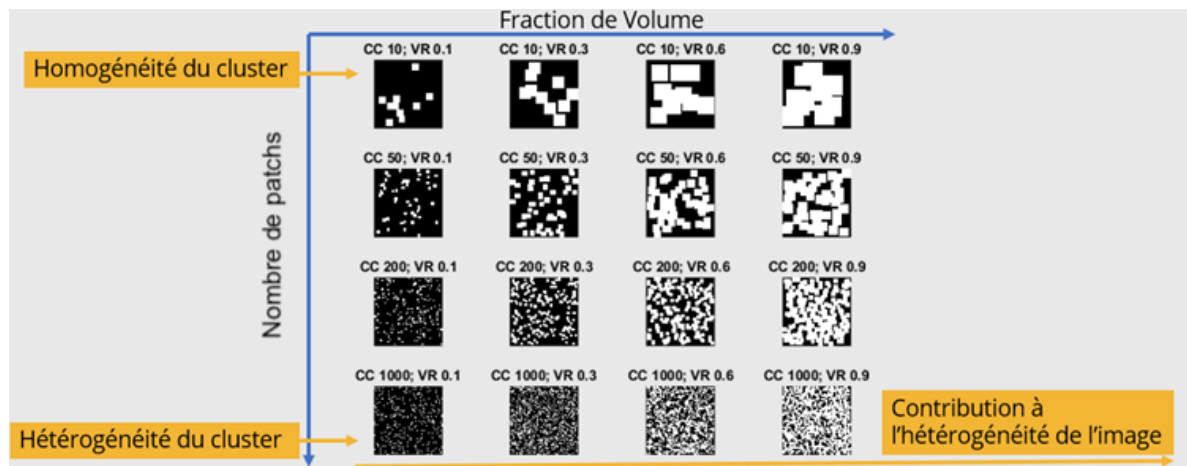


FIGURE 2.8 – Mise en correspondance de la relation du nombre de patches (CC) et de la fraction volumique (VR) au regard de l'interprétation de l'hétérogénéité tumorale. Le nombre de patches influe sur la complexité de l'organisation spatiale et la fraction volumique sur la contribution du cluster à l'hétérogénéité de l'image.

une moins bonne focalisation et donc une plus grande complexité de l'image.

Ces illustrations numériques prenant en compte de manière simpliste un unique cluster ont permis de clarifier l'interprétation des trois métriques. En plus, et dans une démarche de recherche exploratoire, nous avons développé un indice global regroupant ces trois métriques et décliné à deux échelles, celle de la tumeur et celle des clusters. Il a pour ambition de faciliter l'interprétation et par là même, l'exploitation des indices d'hétérogénéité par la communauté médicale. Il est décrit dans la section suivante.

2.3 D'un ensemble de métriques vers un unique biomarqueur

Si les trois métriques semblent être adaptées à la description de la complexité spatiale des cartes paramétriques des tumeurs, cela devient rapidement plus compliqué lorsqu'il s'agit d'étendre ce raisonnement à une carte paramétrique comprenant de multiples niveaux de gris. En effet, après le processus de clustering ces cartes paramétriques comprennent plusieurs clusters, à partir desquels les métriques peuvent être extraites. Il résulte alors pour chaque carte paramétrique un ensemble de nombreuses métriques. Bien que ces métriques trouvent leur intérêt dans leur facilité d'interprétation individuelle, une interprétation croisée de la totalité des métriques conduit inévitablement à une représentation mentale laborieuse avec un nombre de clusters élevé.

Dans ce contexte et avec l'objectif permanent de proposer un biomarqueur facilement interprétable voire intuitif et généralisable aux différentes modalités d'imagerie, nous proposons un nouvel indice prenant en compte les clusters et les métriques présentées. Nommé index de l'hétérogénéité ou H_{index} , nous allons désormais décrire sa formulation mathématique que nous déclinons à deux échelles, celle de la tumeur et celle des clusters.

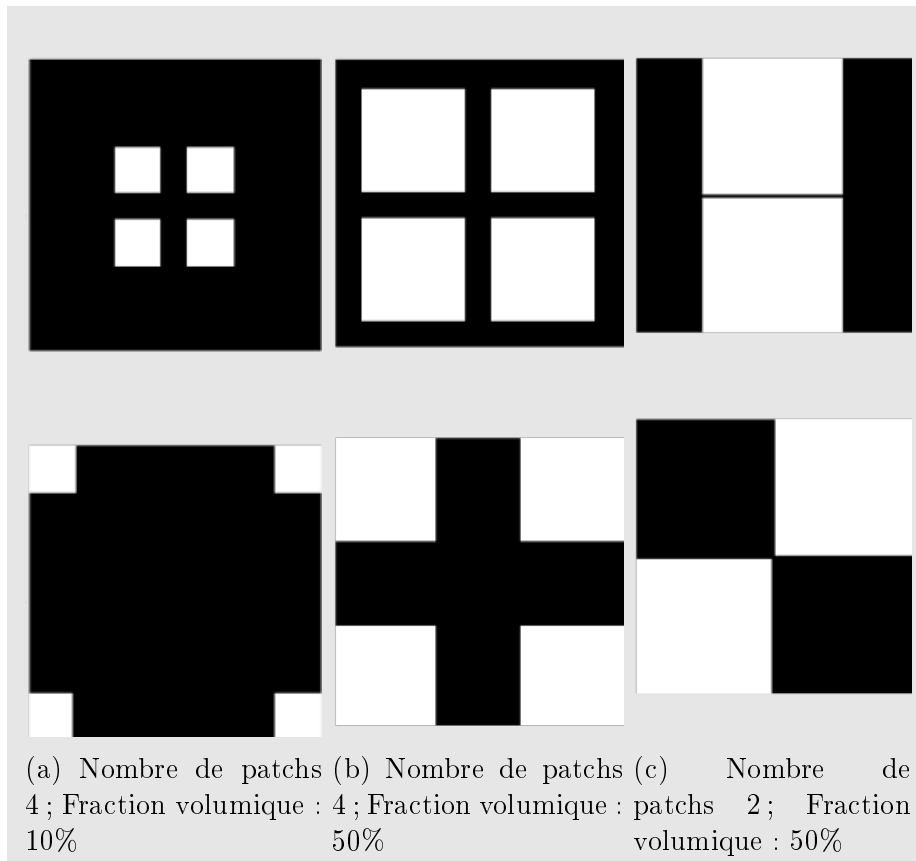


FIGURE 2.9 – Illustration de différentes images contenant un cluster. Deux à deux (verticalement), les images possèdent la même composition avec un nombre de patches et une fraction volumique identiques. À l'inverse leur configuration est distincte. Ces exemples illustrent l'insuffisance des métriques de la composition pour différencier l'organisation spatiale de ces paires d'images. Le cluster possède les caractéristiques suivantes : (a) 4 patches, 10% du volume total, (b) 4 patches, 50% du volume total, (c) 2 patches, 50% du volume total.

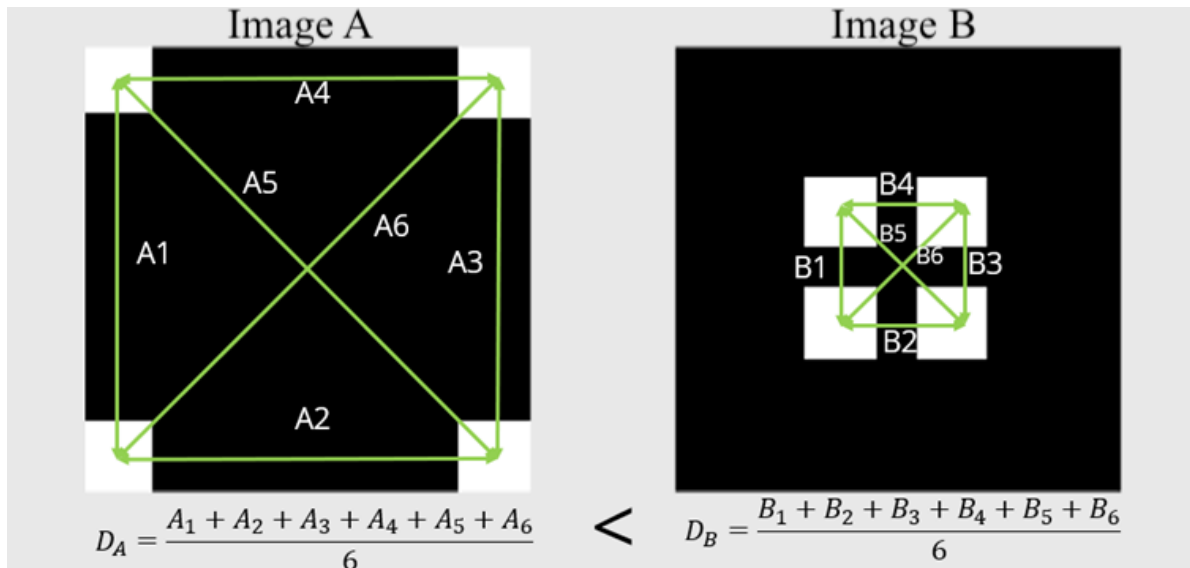


FIGURE 2.10 – Illustration de la métrique distance moyenne entre les patches permettant de différencier les deux organisations par leur configuration spatiale. L'image A possède une distance moyenne entre ses patches supérieure à celle de l'image B et est ainsi plus complexe d'après la définition opérationnelle.

2.3.1 Définition du H_{index}

Le H_{index} vise à considérer l'hétérogénéité d'une image comme la somme de la complexité des clusters qui la composent. La complexité au sein d'un unique cluster, le $H_{cluster}$, peut être définie par la formule suivante (2.6) :

$$H_{cluster} = N \times D \quad (2.6)$$

avec, N le nombre de patches, D la distance moyenne inter-patches.

L'hétérogénéité de l'image définie par la somme des complexités de tous ses clusters permet d'en déduire la formule du H_{index} suivante (2.7) :

$$H_{index} = \sum_{k=1}^K H_{cluster}(k) \times fV(k) \quad (2.7)$$

avec, K le nombre de clusters, fV la fraction volumique.

Avec N le nombre de patches composant un cluster, D la distance moyenne entre ces patches et fV la fraction volumique du cluster par rapport au volume tumoral.

Le H_{index} constitue donc un unique indice numérique quantifiant l'hétérogénéité d'une image à partir d'un processus de clustering et de métriques définissant la complexité de la tumeur et de ses sous-régions. La figure 2.11 propose un schéma récapitulatif de l'ensemble de la méthodologie permettant d'obtenir le H_{index} par l'extraction des différentes métriques et l'application du processus de clustering.

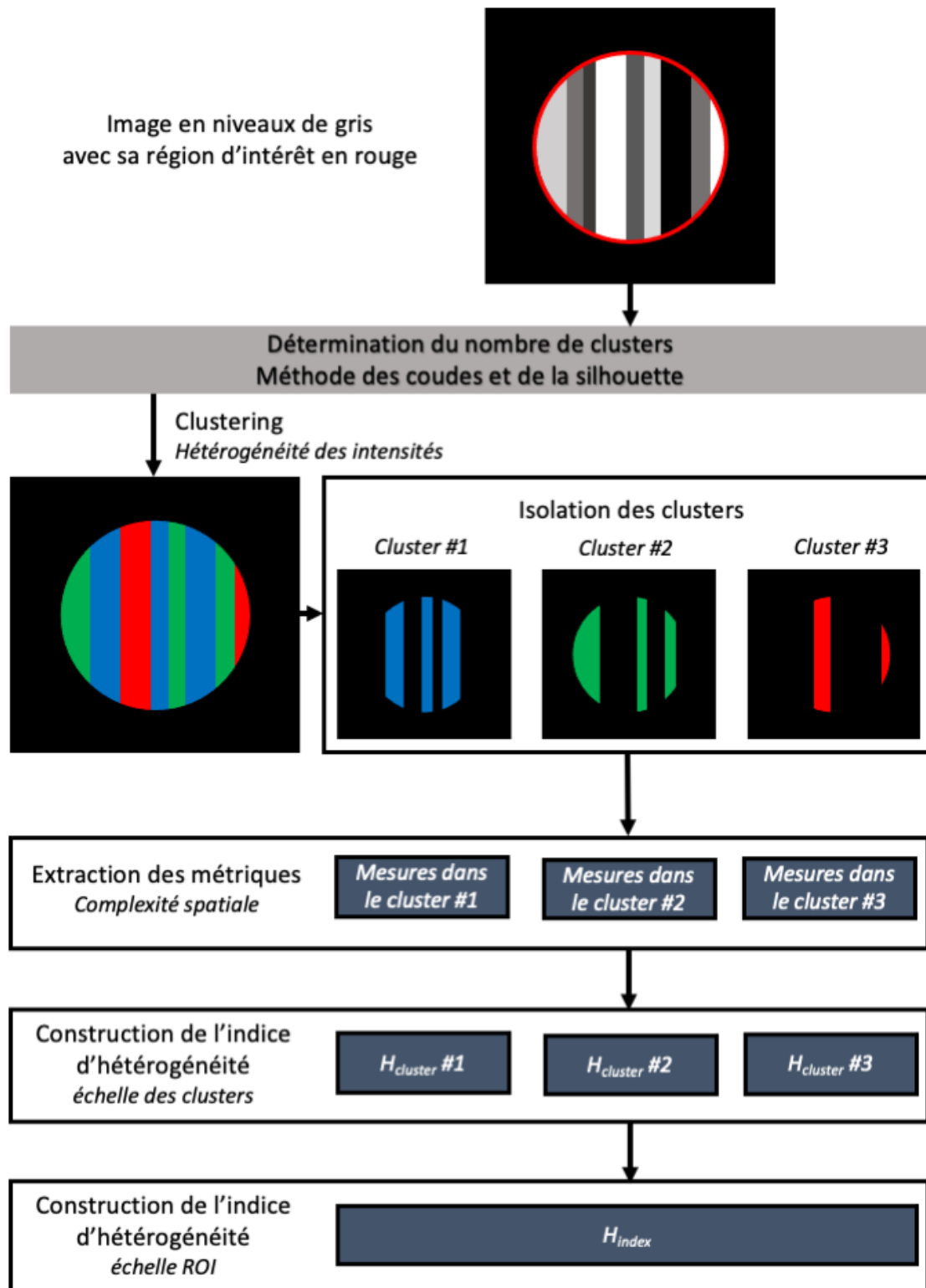


FIGURE 2.11 – Schéma récapitulatif de la méthodologie permettant à partir d'une image d'obtenir le H_{index} . L'image proposée schématise une carte paramétrique en niveaux de gris ainsi qu'une région d'intérêt symbolisant les contours d'une ROI (rouge). A partir de la détermination objective du nombre de clusters, le clustering est appliqué afin d'isoler les sous-régions de l'image. Les métriques sont ensuite extraites afin de déterminer le H_{index} et le $H_{cluster}$.

2.3.2 Interprétation

L'interprétation de la variation du H_{index} est évidemment corrélée à celles des métriques qui le constituent. Ainsi, l'évaluation de la complexité de l'image repose sur 1) le nombre de patches composant les clusters, 2) la distance en lien avec la configuration des patches et enfin 3) la fraction volumique qui pondère l'hétérogénéité de la région d'intérêt étudiée par la contribution du cluster correspondant. Par la suite, nous considérons à la fois le H_{index} mais également l'indice à l'échelle du cluster, le $H_{cluster}$. L'intérêt ici sera de considérer des effets de variation d'hétérogénéité au sein ou entre les clusters qui pourraient être masqués par l'extraction d'une moyenne globale. Ainsi il devient possible de quantifier l'hétérogénéité par le biais de la complexité de l'organisation spatiale, tout en convenant que la notion d'hétérogénéité reste un élément subjectif puisqu'il est généralement consécutif à une interprétation visuelle et individuelle de l'image médicale. L'utilisation de ce biomarqueur sous la forme d'un unique indice impose néanmoins certaines conditions méthodologiques conventionnelles. 1) La comparaison du H_{index} entre deux images peut se faire uniquement si les images sont issues de la même modalité et avec le même nombre de clusters. 2) Les métriques définies et employées pour caractériser la complexité d'une image reposent sur une succession de processus de traitement et d'analyse d'images. Par conséquent, les sources de variabilité qui pourraient avoir une incidence sur leur valeur doivent être connues et maîtrisées. Le nombre de patches par exemple est nécessairement corrélé à la taille des pixels, tandis que la distance entre les patches mérite d'être comparée au volume global de la ROI.

Ainsi la suite de ce chapitre sera consacrée à la caractérisation des processus et des métriques décrits jusqu'à présent au regard des différentes sources de variabilités et qui constitue tout naturellement l'étape suivante de validation. Dans ce chapitre dédié à la définition et à la caractérisation théorique du biomarqueur, nous nous sommes ainsi focalisés sur une approche de validation numérique qui est décrite ci-après.

3 Caractérisation numérique du biomarqueur de l'hétérogénéité

La modélisation numérique est une méthode conventionnelle lorsqu'il s'agit d'étudier ou de prédire des comportements physiques complexes. Appliquée à l'imagerie médicale, le recours à des fantômes numériques vise à mimer une configuration anatomique ou pathologique particulière. La modélisation numérique a pour objectif soit d'évaluer les performances théoriques d'une méthode de détection comme les systèmes d'imagerie TEP avec la plateforme GATE (SARRUT et al., 2021) ou bien de personnaliser le

traitement comme l'optimisation de la pose de stent au sein d'une artère (BERTI et al., 2021; ZHANG et al., 2016). Les progrès des modélisations numériques permettent aujourd'hui d'envisager de générer à terme un « jumeau numérique ». Ce terme a initialement été introduit par le Dr. Michael Grieves (GRIEVES, 2019), pour signifier qu'il était aujourd'hui possible de considérer une parfaite représentation numérique d'un objet et donc du corps humain et ainsi soumettre ce miroir virtuel à un ensemble de tests. Néanmoins, en cancérologie, dû à la part méconnue des mécanismes de formation et la grande hétérogénéité de la pathologie, le « jumeau numérique » n'est pas d'actualité.

Dans le cadre de ce travail, le recours à des modélisations numériques simulant l'hétérogénéité présente au sein des images, vise à identifier les paramètres influençant la variabilité de toutes les métriques précédemment définies afin de pouvoir guider l'interprétation dans un contexte clinique. Les différentes sources de variabilité sont : i) les méthodes de traitement d'image pour lesquelles peuvent être envisagés différents algorithmes pour le clustering ainsi que différentes connectivités pour déterminer le nombre de patches ; ii) l'origine des images médicales qui soulève la question des performances instrumentales du système d'acquisition et en particulier de sa résolution spatiale donc de la taille des pixels de l'image, et du contraste, donc les différences d'intensité entre les pixels.

Cette section présentera donc premièrement le principe de simulation adopté et en fournira des exemples avant de s'intéresser à la caractérisation de la méthodologie développée en fonction des sources de variabilités identifiées.

3.1 Fantômes numériques hétérogènes

L'objectif visé par la simulation des fantômes numériques est de proposer un premier niveau de validation de la méthodologie établie pour quantifier l'hétérogénéité tumorale à partir d'une carte paramétrique. Pour caractériser chacun des processus que sont le clustering et l'extraction de métriques quantitatives, la simulation doit être en mesure de générer plusieurs clusters dont la composition sera parfaitement maîtrisée avec un nombre de patches et une fraction volumique connus. En revanche, la configuration des patches au sein des clusters i.e. leur emplacement, leur disposition et leur forme, sera déterminée par un processus aléatoire. Ainsi la métrique distance moyenne entre les patches sera déterminée par la simulation aléatoire. Le développement des simulations a été réalisé sous Matlab (The Mathworks, Natick, United States) à l'aide de l'algorithme suivant :

Algorithme 1 : Algorithme de simulation de cartes catégoriques

```

Nrow : nombre de lignes de l'image ;
Ncol : nombre de colonnes ;
K : nombre de clusters à générer;
Npatches : nombre de patch par cluster;
FVolumique : fraction volumique de chaque cluster;
Vals : Valeur d'intensités des clusters;
for  $k \leftarrow 1$  to  $K$  do
    Générer le cluster k for  $i \leftarrow 1$  to  $Npatch$  do
        Génération aléatoire du patch  $i$ ;
        while Fraction Volumique du cluster  $\leq$   $FVolumique_k$  do
            Ajouter un pixel de valeur  $Vals_k$  connecté à au moins un pixel du
            cluster en fonction de la connectivité (4 ou 8 en 2D) et qui n'est pas
            connecté à un autre patch du cluster;
        end
    end
end

```

Pour faciliter la compréhension de cette partie, nous avons d'abord considéré des exemples de fantômes numériques hétérogènes représentant des images simulées de taille 50x50 pixels et dotés en premier lieu d'un unique cluster. Au fil des simulations, l'hétérogénéité spatiale des images est progressive et associée à un ensemble de compositions et de configurations des patches représentant la diversité des situations cliniques possibles. Ainsi la figure 2.12 montre des fantômes numériques présentant une évolution de la composition spatiale avec un nombre de patches croissant de 8 à 200 et une fraction volumique constante et donc une augmentation de la complexité. La figure 2.13 représente pour un nombre de patches constant à 40, une évolution de la contribution du cluster à la complexité de l'image avec une évolution progressive de la fraction volumique de 5 à 50 %. Enfin, il est évidemment possible de tendre vers des configurations encore plus réalistes en augmentant cette fois-ci le nombre de clusters. La figure 2.14 présente des simulations pour 1, 2 ou 3 clusters. La somme des nombres de patches pour chaque image y est de 30 et la somme des fractions volumiques des clusters pour chaque image est de 30%.

3.2 Influence de la connectivité sur les métriques d'hétérogénéité

Dans le cadre de cette étude numérique, l'étude de l'influence de la connectivité a été faite à partir de la génération d'images contenant un unique cluster, 10 patches, une fraction volumique de 20%. Ces cartes contiennent ainsi des pixels dont la valeur est «

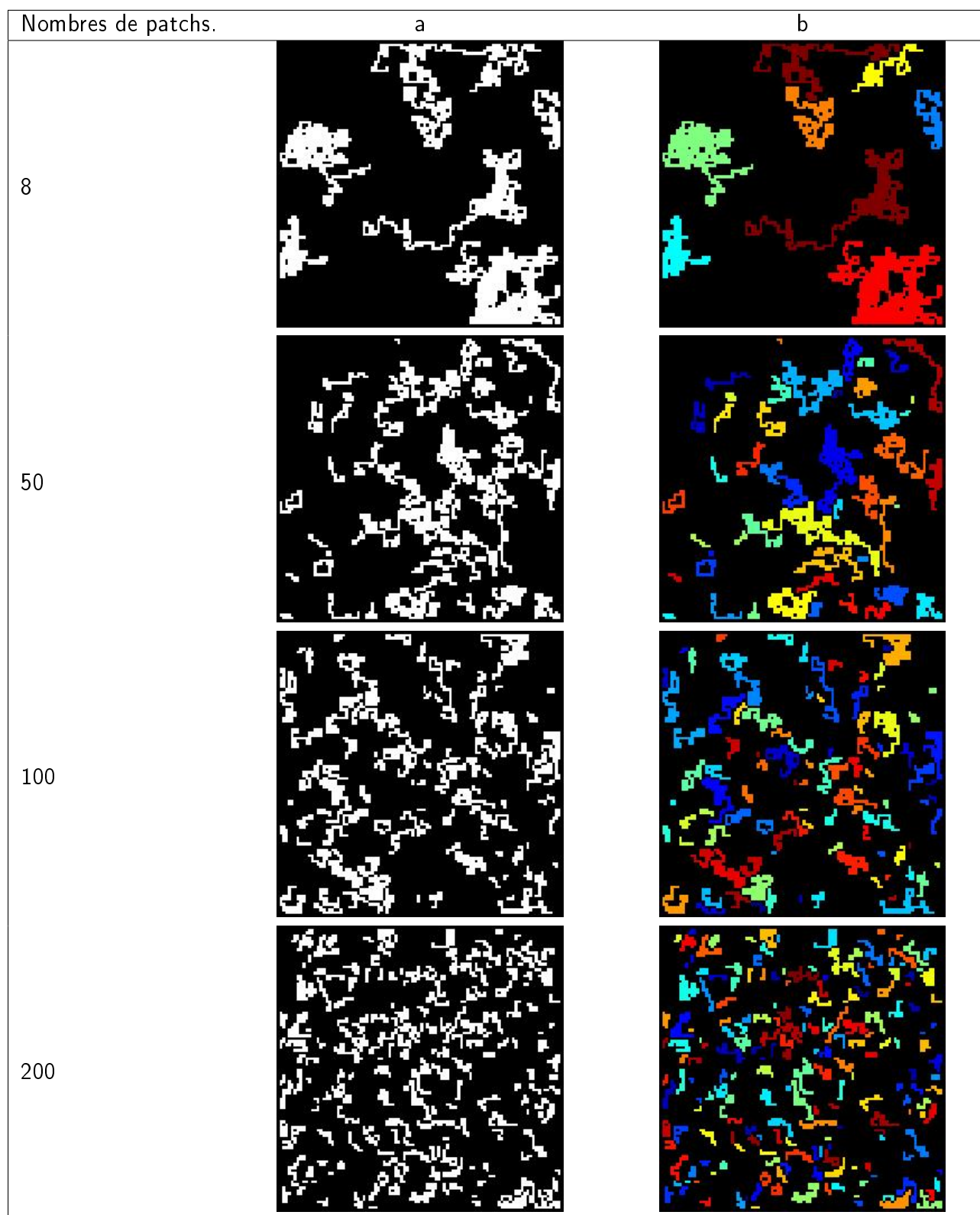


FIGURE 2.12 – Exemples de cartes paramétriques simulées comportant 1 cluster de fraction volumique constante (20%) et de différents nombre de patchs : 8, 50, 100, 200.

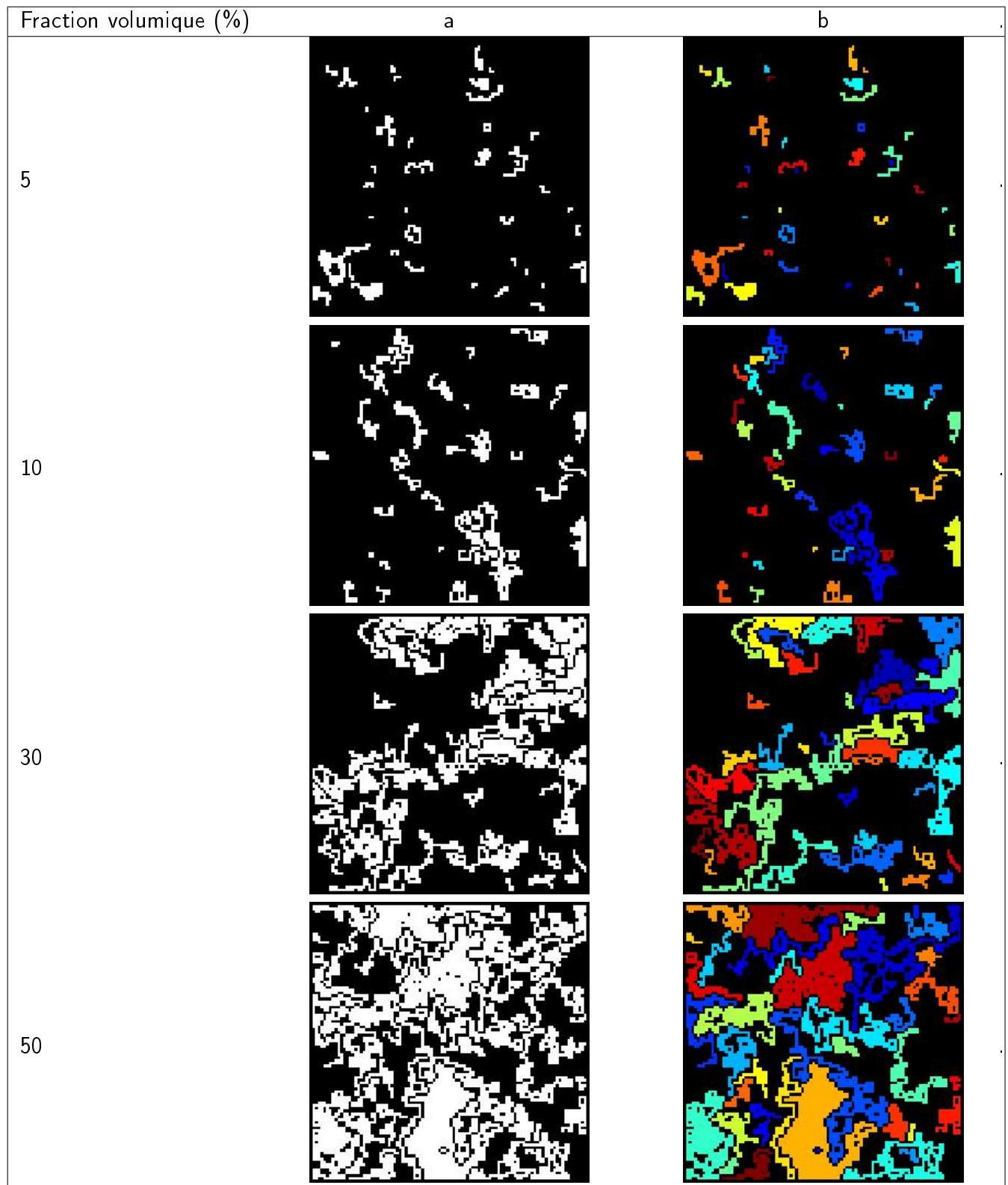


FIGURE 2.13 – Exemples de cartes paramétriques simulées comportant 1 cluster. Le nombre de patches est constant (40) pour chaque simulation et les fractions volumiques croissantes : 5, 10, 30, 50%

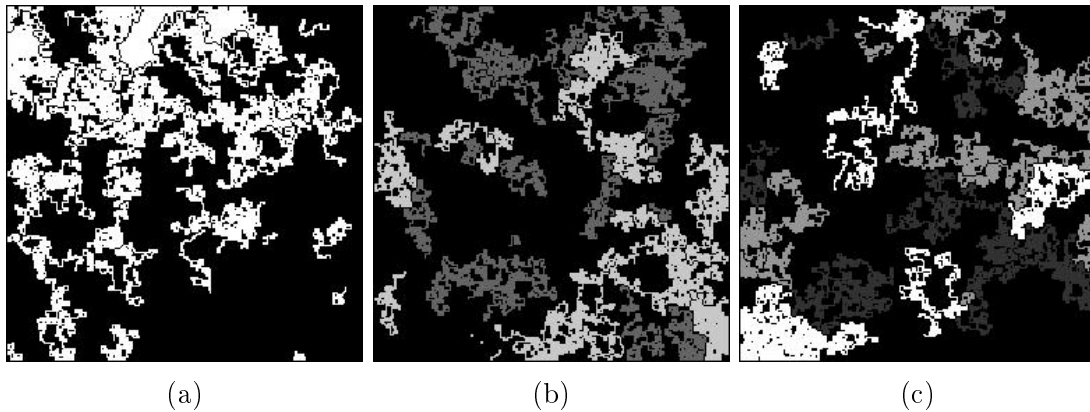


FIGURE 2.14 – Simulations de cartes paramétriques illustrant la possibilité de générer des clusters différents en terme de nombre, de compositions et de configurations. (a) Carte composée d'un cluster de 30 patches et de fraction volumique 30%. (b) Carte composée de 2 clusters de 15 patches chacun représentant respectivement 15 % du volume. (c) Carte composée de 3 clusters de 10 patches chacun représentant respectivement 10% du volume.

1 » lorsque le pixel appartient au cluster, « 0 » sinon. Parmi l'ensemble des métriques proposées pour l'exploration d'un biomarqueur de l'hétérogénéité, nous nous intéressons ici tout particulièrement au nombre de patches. En effet, contrairement à la distance et à la fraction volumique, le nombre de patches est le seul indice qui résulte d'une mise en œuvre algorithmique complexe et qui dépend notamment de la connectivité (cf 2.2.1). Ainsi, cette section aborde l'importance du choix de la connectivité lors du calcul du nombre de patches et l'influence d'une modification de la résolution d'image au cours du processus.

La connectivité utilisée lors de l'analyse est un paramètre majeur dans l'évaluation de la complexité. Pour caractériser son importance, deux images 2D ont été simulées avec les deux types de connectivité possibles pour la construction de leurs configurations spatiales. La première (figure 2.15.a) est générée avec une connectivité 4, les pixels de chaque patch ne peuvent donc pas avoir de voisins appartenant à un autre patch dans les quatre directions : haut, bas, gauche droite. La seconde (figure 2.15.b) est générée avec une connectivité 8, les pixels ne peuvent avoir aucun voisin appartenant à un autre patch dans les 4 directions précédentes et les 4 diagonales. Chaque figure représente la carte générée, avec des couleurs artificielles permettant d'identifier visuellement chaque patch généré. Une fois ces images simulées, nous les avons analysées pour en extraire les métriques et l'indice $H_{cluster}$ en considérant les deux types de connectivités, mais cette fois-ci, pour déterminer le nombre de patches. Les valeurs des métriques déterminées lorsque les images sont analysées avec la même connectivité que celle choisie pour leur construction, représentent les vraies valeurs. Celles-ci sont ensuite à comparer avec les cas de figures où les deux connectivités (analyse versus construction) sont différentes. Les résultats sont présentés table 2.1.

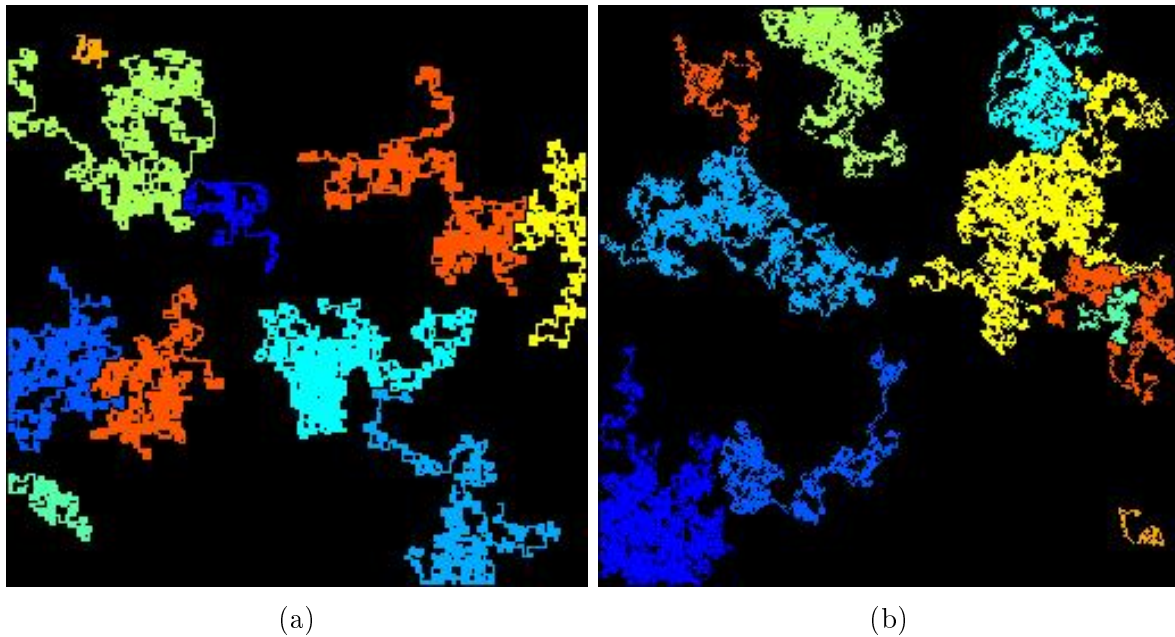


FIGURE 2.15 – Cartes paramétriques simulées composées d'un cluster de 10 patches occupants 20% de la surface de l'image selon deux connectivités : (a) 4 avec ainsi aucun voisins entre deux patches dans les directions haut, bas, gauche, droite ;(b) 8 avec aucun voisins dans les quatres directions précédentes et les quatres diagonales.

Il résulte que lorsque la connectivité de l'analyse est supérieure à celle mise en œuvre pour la construction de l'image simulée alors deux métriques sont sous-estimées : le nombre de patch à 7 au lieu de 10 attendu, ce qui implique une diminution du $H_{cluster}$ de 30% avec une valeur de 354,8 au lieu de 513. Lorsque la connectivité utilisée pour l'analyse est inférieure à celle de la construction, trois métriques sont impactées. La distance moyenne entre les patches diminue de 10% en passant de 48,9 u.a. à 43,7. A l'inverse le nombre de patches et le $H_{cluster}$ sont grandement surestimés, d'un facteur 13 pour le premier et d'un facteur 12 pour le second.

Par conséquent, la connectivité doit toujours être choisie à sa valeur la plus haute possible. Néanmoins dans la pratique clinique, lorsque la véritable configuration 3D n'est pas connue, ce choix est fait de manière arbitraire mais doit impérativement être constant sur l'ensemble des patients une fois choisie. Enfin, il est à noter qu'une connectivité trop élevée pourrait engendrer le regroupement inapproprié de patches initialement distincts.

3.3 Influence de la résolution spatiale de l'image radiologique

En radiologie, la diversité des constructeurs, des machines, des protocoles et des modalités est à l'origine d'une grande variabilité de la résolution spatiale des images radiologiques. Dans ce contexte, l'harmonisation des résolutions spatiales pour l'ensemble des images d'un jeu de données est devenue pratique courante et plus particulièrement encore avec l'émergence de l'intelligence artificielle pour laquelle

Connectivité générée	4		8	
Connectivité de l'analyse	4	8	8	4
Résultats	Vérité	Analyse	Vérité	Analyse
Nombre de patches	10	7	10	131
Fraction volumique (%)	20	20	20	20
Distance moyenne (u.a)	51.3	50.7	48.9	43.7
$H_{cluster}$	513.0	354.8	489.3	5721.3

TABLE 2.1 – Détermination des métriques et du $H_{cluster}$ à partir d'analyses en composantes connectées selon les différentes connectivités possibles. La colonne "Connectivité générée" 4 correspond à la figure 2.15.a ; 8 correspond à la figure 2.15.b. La distance est une unité arbitraire représentant le nombre de pixels d'écarts.

les algorithmes d'apprentissages profonds requièrent une parfaite correspondance des résolutions entre les images. Dans ce dernier cas, les modifications sont généralement effectuées en considérant une dégradation de la résolution d'image afin de correspondre au format du réseau de neurones employé et pour limiter l'espace mémoire occupé par chaque image. Dans le cadre de nos travaux sur un biomarqueur de l'hétérogénéité, il nous apparaît important de caractériser l'impact d'une telle modification sur les trois métriques et le $H_{cluster}$.

De manière générale, la résolution spatiale d'une image peut être dégradée ou améliorée. Il y a dégradation lorsque plusieurs pixels sont regroupés pour n'en former qu'un seul de plus grande taille, et amélioration lorsqu'un pixel est divisé en plusieurs de plus petite taille. Dans les deux cas, une interpolation est nécessaire pour déterminer la valeur des nouveaux pixels. Dans le cas particulier des cartes catégoriques, la méthode d'interpolation ne peut être que celle des voisins les plus proches (angl. Nearest Neighbor) car elle ne crée pas de nouvelles structures et la valeur attribuée aux pixels résultants ne peut être qu'une valeur, et donc une catégorie, qui existait dans la carte d'origine.

La figure 2.16 illustre une dégradation de la résolution spatiale de l'image initiale de 300x300 pixels à 50x50 pixels et qui s'accompagne donc d'une augmentation d'un facteur 6 de la taille élémentaire du pixel⁶. L'image initiale simulée contient un unique cluster, 10 patches et d'une fraction volumique de 20%. La figure 2.17 illustre à l'inverse l'amélioration de la résolution spatiale de l'image simulée de 50x50 pixels à 300x300 pixels avec donc une diminution d'un facteur 6 de la taille élémentaire du pixel. Les résultats de l'extraction des métriques et de l' $H_{cluster}$ obtenus dans les deux cas de figures sont reportés dans la table 2.2.

Lorsque la résolution spatiale de l'image est dégradée avec une augmentation de la taille élémentaire du pixel, la composition de l'image s'en retrouve modifiée avec l'apparition ou la disparition de certaines connexions entre les patches et donc une

6. Ce raisonnement est établi en considérant la dimension de l'image simulée (en unité de longueur) fixe.

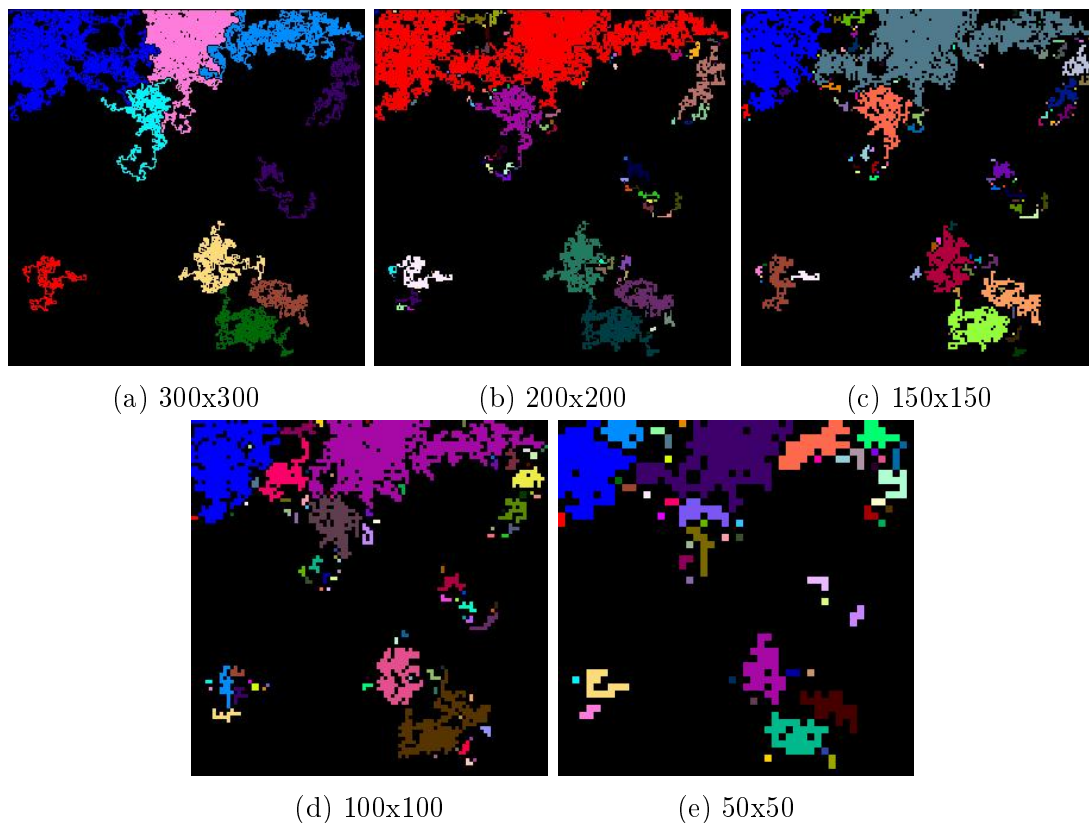


FIGURE 2.16 – (a) Carte paramétrique simulée de taille 300x300 pixels composée d'un cluster de 10 patches occupant 20% de l'image. Dégradations de la résolution de 200x200 jusqu'à 50x50 (b-e) avec une augmentation de la taille des pixels. Les couleurs sont arbitraires et ont pour unique objectif de distinguer les patches voisins.

modification du nombre de patches résultant de l'analyse en composantes connectées. Ainsi, il apparaît une augmentation du nombre de patches avec respectivement 107 et 122 patches pour les images de taille 200x200 pixels et 150x150 pixels. En effet, la modification de la résolution implique qu'en raison de l'interpolation effectuée de nombreux patches se retrouvent séparés par un pixel appartenant à l'arrière-plan. Le nombre de patches pour les résolutions 100x100 pixels et 50x50 pixels est toujours supérieur à celui de l'image originale mais diminue progressivement par rapport aux résolutions modifiées précédentes. La dégradation de la résolution spatiale y est tellement importante que les patches fusionnent et leur nombre se retrouve logiquement corrélé au nombre de pixels décroissants. A l'inverse, une amélioration de la résolution spatiale ne modifie pas la configuration de l'image et les valeurs des métriques et du $H_{cluster}$ restent inchangés.

Par conséquent, si une modification de la résolution d'image est nécessaire dans le processus de traitement et d'analyse pour harmoniser le jeu de données, celle-ci doit systématiquement être améliorée pour envisager une analyse de son hétérogénéité par les métriques étudiées.

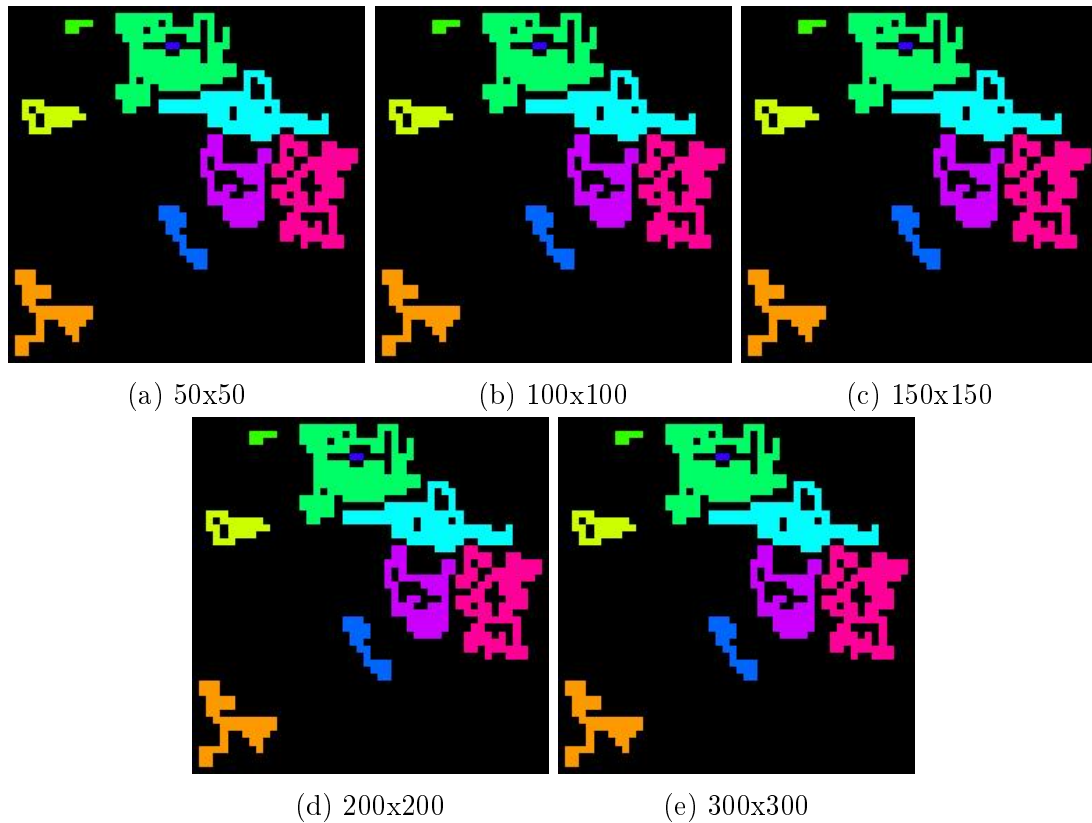


FIGURE 2.17 – (a) a) Carte paramétrique simulée de taille 50x50 pixels composée d'un cluster de 10 patches occupant 20% de l'image. Amélioration de la résolution de 100x100 jusqu'à 300x300 (b-e) avec une diminution de la taille des pixels. Les couleurs sont arbitraires et ont pour unique objectif de distinguer les patches voisins.

	Dégradation de la résolution				
	300x300	200x200	150x150	100x100	50x50
Nombre de patches	10	107	122	100	66
Fraction volumique (%)	20.0	19.9	20.0	20.3	20.3
Distance moyenne (u.a)	154.7	125.3	134.6	130.7	139.5
$H_{cluster}$	1547	13405	16418	13074	9209
	Amélioration de la résolution				
	50x50	100x100	150x150	200x200	300x300
Nombre de patches	10	10	10	10	10
Fraction volumique (%)	20	20	20	20	20
Distance moyenne (u.a)	23.9	23.9	23.9	23.9	23.9
$H_{cluster}$	238.8	238.8	238.8	238.8	238.8

TABLE 2.2 – Détermination des métriques et du $H_{cluster}$ pour une carte paramétrique simulée composé d'un cluster de 10 patches et de fraction volumique 20%. La première colonne correspond à la résolution originale, les suivantes aux images de résolution modifiée : (haut) dégradation de la résolution spatiale ; (bas) amélioration de la résolution spatiale.

3.4 Influence du contraste de l'image radiologique

L'image radiologique est définie par sa modalité, sa résolution spatiale mais également par son contraste. Ce paramètre est également extrêmement variable selon les systèmes d'imagerie, les constructeurs, et les protocoles d'acquisition. La normalisation des intensités ou égalisation d'histogramme est une étape généralement effectuée en radiomique pour que les différentes données soient rapportées à une échelle commune sans introduire de différences dans les gammes d'intensités. Dans le contexte d'essais cliniques multicentriques, ces étapes de normalisation apparaissent comme des facteurs critiques à l'amélioration des modèles de prédiction (PAPANIKOLAOU et al., 2020 ; UM et al., 2019). Dans le cadre de l'étude des cartes paramétriques, les sources de variabilités peuvent reposer sur les différentes modélisations mathématiques possibles et sur les différents protocoles d'acquisitions pour extraire les paramètres quantitatifs.

Dans le cadre de notre approche, l'étape de clustering nous permet de nous affranchir de la variabilité des contrastes et laisse place à l'évaluation de l'organisation spatiale de l'image. Dans ce contexte, outre le choix du nombre de clusters déjà abordé précédemment (cf section 2.1.3), nous avons étudié la fiabilité du clustering en présence d'une variation de contraste. Pour ce faire, une première étude permet de démontrer l'indépendance de nos métriques et du $H_{cluster}$ au contraste des images. Une seconde étude visait quant à elle à évaluer l'influence du type d'algorithme de clustering sur les métriques.

3.4.1 Indépendance par rapport aux niveaux de gris/ indépendance par rapport au contraste

Avec l'objectif fixé de génération d'un biomarqueur qui soit utilisable au sein de protocoles multicentriques, la simulation suivante vise à démontrer l'indépendance des indices d'hétérogénéité aux variations de contraste et ce, grâce au processus de clustering.

A partir d'une image simulée de taille 200x200 pixels et dotée d'une configuration spatiale fixe, les valeurs d'intensités des pixels sont déterminées selon une distribution gaussienne des intensités caractérisée par sa valeur moyenne et son écart-type notés [moyenne, écart-type]. Ainsi, nous avons simulé trois images avec les distributions gaussiennes aux paramètres respectifs : [150,25], [150, 10], [200,10]. La composition des trois images résultantes est exactement similaire puisque les valeurs d'intensités des pixels sont attribuées selon chaque gaussienne par le même tirage aléatoire. Les histogrammes de ces trois images sont présentés figure 2.18. Il faut noter qu'ici, la composition spatiale initiale des images et en particulier le nombre de patches est aléatoire. Il est le seul résultat du processus d'attribution de la valeur des pixels. Une fois les trois images générées, elles ont été soumises au processus de clustering avec $K=3$ clusters à former. Ce nombre ayant été choisi par la méthode développée dans la section 2.1.3 et appliquée ici à la première image. Ce nombre K n'est pas déterminé pour

Image Cluster #	1			2			3		
	1	2	3	1	2	3	1	2	3
Nombres de patches	659	710	669	659	707	660	675	688	681
Fraction Volumique	2.6	4.5	2.8	2.7	4.4	2.8	2.6	4.6	2.7
Distance Moyenne	110	111	109	111	110	111	112	110	111
$H_{cluster}$	19042	35856	21009	19753	34905	20812	19654	35051	21139
H_{index}		75907			75470			75844	

TABLE 2.3 – Métriques, $H_{cluster}$ et H_{index} extraits à partir des trois images présentant des contrastes différents et après le processus de clustering.

les autres images puisque dans un contexte d'étude multicentrique, une fois déterminée dans un centre, la méthodologie doit être appliquée de la même manière dans les autres centres et donc en ayant recours à un même nombre de clusters. Les différentes métriques, le $H_{cluster}$ et le H_{index} ont été extraits du résultat de chaque clustering.

Il résulte du processus de clustering, les histogrammes présentés figure 2.19. L'identification des 3 clusters est similaire avec le nombre de pixels par cluster similaire pour chaque image. L'indice ajusté de Rand (IAR) qui est une mesure de la similarité entre deux partitions est calculé deux à deux entre les résultats du clustering. Il en résulte que l'IAR entre l'image 1 et 2 est de 0.9849, entre l'image 1 et 3 de 0.9847 et de 0.9847 entre l'image 2 et 3. Cet indice confirme la très forte similarité entre les résultats du clustering.

Enfin, les métriques et les indices extraits des différentes images sont donnés à titre indicatif table 2.3. On retrouve alors des métriques avec de très faibles variations pour chaque cluster entre les images et il en résulte des H_{index} pour chaque image respectivement de 75907, 75470, 75844. En effet, il est attendu qu'une infime variation dans le résultat du processus de clustering n'implique qu'une très faible variation sur les métriques et indices qui en sont extraits.

Ces résultats illustrent un avantage concurrentiel du biomarqueur et des métriques que nous proposons pour l'évaluation de l'hétérogénéité. Bien que ces résultats doivent être confirmés dans un contexte clinique, le recours à ces métriques pourrait ainsi faciliter les processus d'harmonisation des données (van TIMMEREN et al., 2020) et ainsi participer à la standardisation des protocoles pour des études multicentriques. De plus, la question de l'importance et de la fiabilité de ces processus d'harmonisation des intensités appliqués en imagerie paramétrique commence à émerger au sein de la communauté (CROMBÉ et al., 2021).

3.4.2 Choix de l'algorithme de clustering

Les deux algorithmes proposés, K-moyenne et FCM, possèdent une différence majeure comme indiqué dans la partie 2.1 de ce chapitre. Pour rappel, le premier effectue un

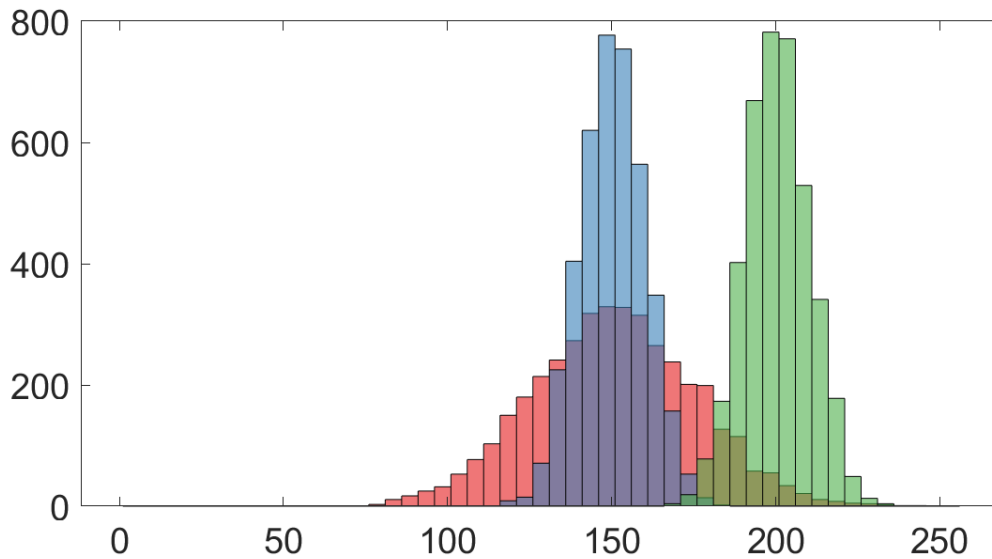


FIGURE 2.18 – Histogrammes des trois images simulées avec des distributions gaussiennes d'intensités différentes. En rouge, un histogramme de l'image aux intensités attribuées aléatoirement par une gaussienne de moyenne 150 et d'écart type 25, en bleu de moyenne 150 et d'écart type 10, en vert de moyenne 200 et d'écart type 10. (u.a.)

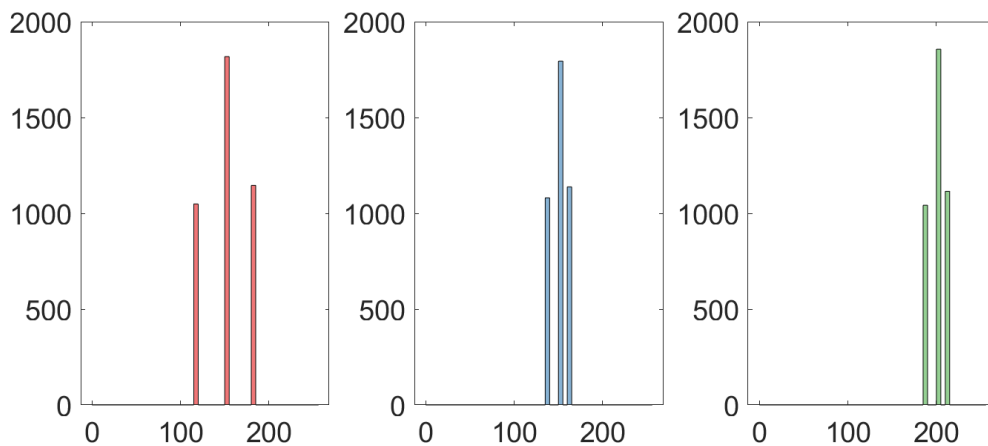


FIGURE 2.19 – Histogrammes des trois images après le processus de clustering. Les centroïdes des clusters sont respectivement pour chaque image : 120, 150 et 180 pour la première image (rouge) ; 138, 150, 162 pour la deuxième (bleu) ; 188, 200, 212 pour la troisième (vert) (u.a.)

clustering dit dure, c'est-à-dire qu'un objet appartient à un seul et unique cluster, tandis que l'autre effectue un clustering dit doux où les objets peuvent appartenir avec plus ou moins de probabilités à différents clusters. Dans ce dernier cas, lors de la reconstruction des images, les pixels sont attribués au cluster pour lequel la probabilité d'appartenance est la plus grande. L'avantage principal de l'algorithme des K-moyenne réside dans sa rapidité, sa robustesse et sa simplicité d'implémentation. Nous ne redémontrons pas ici ces avantages, de nombreuses études sur différents jeux de données ayant d'ores et déjà reportées des résultats similaires (CEBECI & YILDIZ, 2015; PANDA et al., 2012; R, 2012). Son inconvénient est qu'il peut ne pas réussir à séparer correctement des clusters pour lesquels les valeurs se superposent. FCM a été développé pour contrer cette problématique. Néanmoins, dans le contexte de l'imagerie radiologique et paramétrique, la constitution optimale des clusters n'est pas connue a priori. Dans le cadre de notre étude, nous nous sommes intéressés à l'influence de l'algorithme de clustering sur nos indices de l'hétérogénéité, et ce lorsque le contraste de l'image est variable.

La simulation consiste en une image simulée de taille 200x200 et comme lors de la simulation précédente, les valeurs d'intensités de pixels sont attribuées à partir de trois gaussiennes centrées sur la même valeur, 150 u.a mais d'écart-types différents et respectivement de 10, 30 et 60 u.a. Les histogrammes des trois images simulées ci-après nommées respectivement Image 1, 2 et 3 dans l'ordre décroissant de leur écart-type, sont présentés figure 2.20. Chacune de ces images est soumise à un processus de clustering selon les deux algorithmes et les histogrammes résultants sont proposés figure 2.21.

Les résultats du clustering sont comparés deux à deux par l'indice de Rand et la valeur du H_{index} résultant. Ainsi, pour l'image 1, l'IAR qui compare les résultats par les deux méthodes de clustering est de 0.9987 indiquant un résultat très similaire. Pour l'image 2 et 3, l'indice de rand est respectivement 0.9984 et 1, ce dernier indiquant un résultat identique. L'ensemble des données calculées sont présentés table 2.4 qui indique une très grande similitude des résultats entre les deux algorithmes de clustering. Les H_{index} extraits sont respectivement pour les K-moyennes et FCM : 63620 et 62648 pour l'image 1; 62315 et 62556 pour l'image 2; 61520 et 61520 pour l'image 3.

En conséquence, les deux algorithmes n'influencent pas les valeurs des indices d'hétérogénéité, et ce quel que soit le contraste des images. Si l'algorithme des K-moyennes pourrait être privilégié en raison de sa rapidité d'exécution, dans un contexte clinique la durée d'exécution des deux algorithmes reste faible, de l'ordre de quelques secondes. Ainsi, le choix de l'algorithme de clustering reste subjectif mais au regard d'une utilisation multicentrique du biomarqueur, il doit faire partie intégrante du processus de standardisation de la méthodologie et en ayant conscience des avantages et des limitations de chacun.

Métrique	Image 1		Image 2		Image 3	
	KM	FCM	KM	FCM	KM	FCM
Nombres Patches	657	663	636	655	624	624
	696	678	658	663	653	653
	697	677	687	667	673	673
Fraction Volumique (%)	2.53	2.59	2.56	2.77	2.52	2.52
	4.46	4.57	4.53	4.44	4.53	4.53
	2.96	2.76	2.94	2.81	2.97	2.97
Distances (u.a)	92.6	92.9	95.1	95.8	94.2	94.2
	92.6	93.5	93.7	93.8	96.2	96.2
	92.8	92.3	94.5	94.3	91.9	91.9
$H_{cluster}$	15489	16035	15433	17354	14779	14779
	28890	29082	27849	27545	28408	28408
	19241	17531	19032	17657	18333	18333
H_{index}	63620	62648	62315	62556	61520	61520

TABLE 2.4 – Valeurs des métriques et des indices de l'hétérogénéité pour la comparaison des algorithmes des K-moyennes et de FCM pour des images de différents contrastes.

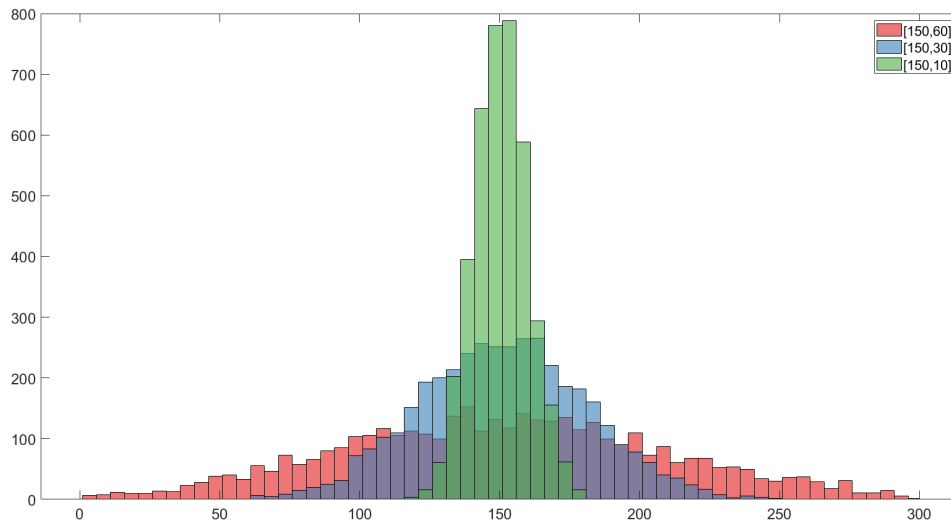


FIGURE 2.20 – Histogrammes de trois images aux contrastes différents caractérisés respectivement par les distributions gaussiennes caractérisées par une moyenne commune de 150 (u.a.) et d'écart type respectifs 60 (rouge, Image 1), 30 (bleu, image 2), 10 (vert, image 3).

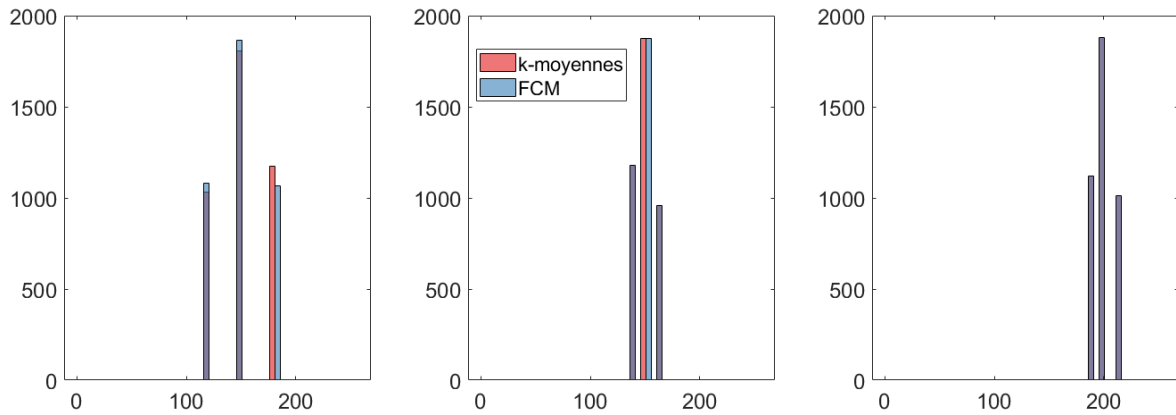


FIGURE 2.21 – Histogrammes résultant du processus de clustering pour les deux algorithmes K-moyennes et FCM pour chacune des trois images simulées. En rouge, les histogrammes après traitement par les K-moyennes, en bleu par FCM, en gris le recouvrement des deux.

4 Conclusion

Les travaux réalisés et présentés dans ce chapitre ont permis de définir un jeu de biomarqueurs potentiels évaluant l'hétérogénéité tumorale. Leur élaboration a été réalisée à partir de deux processus distincts et successifs. D'une part le clustering, qui permet de mettre en évidence l'hétérogénéité des intensités par l'identification de sous-régions tumorales d'intérêts. Appliquées à des cartes paramétriques, les sous-régions identifiées peuvent refléter l'hétérogénéité des tissus néoplasiques ou bien des processus physiologiques au sein de la tumeur. D'autre part, la complexité de la distribution spatiale de l'image d'une tumeur est basée sur des métriques intuitives que sont le nombre de patches, leur distance moyenne, leur fraction volumique et le H_{index} à l'échelle tumorale et à l'échelle des clusters. Le fil conducteur de l'élaboration de ces indices était leur interprétabilité plus intuitive pour décrire l'évolution de l'hétérogénéité d'une image.

Une étude par simulation numérique a complété cette partie théorique. L'objectif était d'identifier les avantages et les limitations des processus de génération des indices et d'en identifier les sources de variabilités en regard des règles usuelles de traitement et d'analyse d'image. Plus précisément, le recours à des simulations de fantômes numériques, dont l'objectif était de représenter des systèmes tumoraux hétérogènes, nous a permis de valider qu'un modèle de clustering adéquat permettait de s'affranchir des effets de contraste et des niveaux de gris. Indépendamment des niveaux de gris, le second processus d'extraction de métriques et d'indices descripteurs de la complexité permettait ainsi de se focaliser sur l'hétérogénéité spatiale. Leur caractérisation a permis de mettre en évidence que d'une part leurs performances n'étaient pas impactées par une amélioration de la résolution spatiale et que d'autre part la connectivité devait être

préféablement surestimée et standardisée.

Ces différentes caractérisations s'inscrivent dans le cadre d'une validation technique préliminaire et recommandée par le consensus concernant le développement de biomarqueurs d'imagerie quantitative (O'CONNOR et al., 2017)(cf chapitre 1.4). Bien que ce consensus évoque exclusivement le recours à des fantomes physiques, au regard de notre problématique d'évaluation de l'hétérogénéité en imagerie, il n'existe à l'heure actuelle aucun fantome standardisé bien que quelques approches soient décrites dans la littérature (VALLADARES et al., 2020). En effet, ces fantômes physiques devraient contenir un niveau de complexité permettant dans des dispositions réalistes de caractériser chaque indice (FOURNIER et al., 2021) et pour chaque modalité. Dans ce contexte le recours aux fantômes numériques a permis de contourner cette limitation matérielle. De plus, ces simulations permettent de s'affranchir de la variabilité inhérente aux systèmes d'acquisitions due à l'hétérogénéité des réglages et aux biais de mesures. Enfin, d'un point de vue pratique, le recours à des fantômes numériques permet de réduire largement les coûts et de multiplier les études à partir d'un nombre potentiellement infini de simulations. Le principal obstacle à notre étude et plus généralement à l'ensemble des études portant sur l'hétérogénéité tumorale, qu'elles soient physiques ou numérique, demeure dans la nécessité d'établir une définition préalable de l'hétérogénéité afin de garantir une interprétation sans-équivoque de ce qui est mesuré.

Les caractéristiques de notre méthodologie étant établies, nous allons à présent détailler les différentes applications qui lui ont été consacrées en considérant les $H_{cluster}$, H_{index} et les métriques qui les constituent. Le chapitre 3 proposera au travers d'une étude préclinique, un premier niveau de validation biologique des indices d'hétérogénéité développés dans cette thèse. Le chapitre 4 abordera un premier niveau de validation clinique avec une étude de répétabilité sur volontaires sains, puis les pistes d'applications cliniques pouvant justifier l'intérêt clinique de ces nouveaux indices.

Bibliographie

- AGGARWAL, C. C., HINNEBURG, A. & KEIM, D. A. (2001). On the Surprising Behavior of Distance Metrics in High Dimensional Space [Series Title : Lecture Notes in Computer Science]. In J. VAN DEN BUSSCHE & V. VIANU (Éd.). G. GOOS, J. HARTMANIS & J. van LEEUWEN (**typeredactors**), *Database Theory — ICDT 2001* (p. 420-434). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-44503-X_27
- ANDERSEN, E. K. F., KRISTENSEN, G. B., LYNG, H. & MALINEN, E. (2011). Pharmacokinetic analysis and k-means clustering of DCEMR images for radiotherapy outcome prediction of advanced cervical cancers. *Acta Oncologica*, 50(6), 859-865. <https://doi.org/10.3109/0284186X.2011.578586>
- BERTI, F., ANTONINI, L., POLETTI, G., FIUZA, C., VAUGHAN, T. J., MIGLIAVACCA, F., PETRINI, L. & PENNATI, G. (2021). How to Validate in silico Deployment of Coronary Stents : Strategies and Limitations in the Choice of Comparator. *Frontiers in Medical Technology*, 3, 702656. <https://doi.org/10.3389/fmedt.2021.702656>
- BI, W. L., HOSNY, A., SCHABATH, M. B., GIGER, M. L., BIRKBAK, N. J., MEHRTASH, A., ALLISON, T., ARNAOUT, O., ABBOSH, C., DUNN, I. F., MAK, R. H., TAMIMI, R. M., TEMPANY, C. M., SWANTON, C., HOFFMANN, U., SCHWARTZ, L. H., GILLIES, R. J., HUANG, R. Y. & AERTS, H. J. W. L. (2019). Artificial intelligence in cancer imaging : Clinical challenges and applications. *CA : A Cancer Journal for Clinicians*, caac.21552. <https://doi.org/10.3322/caac.21552>
- CEBECI, Z. & YILDIZ, F. (2015). Comparison of K-Means and Fuzzy C-Means Algorithms on Different Cluster Structures. *Journal of Agricultural Informatics*, 6(3). <https://doi.org/10.17700/jai.2015.6.3.196>
- CROMBÉ, A., BUY, X., HAN, F., TOUPIN, S. & KIND, M. (2021). Assessment of repeatability, reproducibility, and performances of t2 mapping-based radiomics features : a comparative study. *Journal of Magnetic Resonance Imaging*, 54(2), 537-548. <https://doi.org/10.1002/jmri.27558>

- da SILVA, M. P. & COURBOULAY, V. (2011). Une nouvelle mesure de complexité pour les images basée sur l'attention visuelle. *XXIIIème colloque GRETSI - Bordeaux*, 5.
- FOURNIER, L., COSTARIDOU, L., BIDAUT, L., MICHOUX, N., LECOUVET, F. E., de GEUS-OEI, L.-F., BOELLAARD, R., OPREA-LAGER, D. E., OBUCHOWSKI, N. A., CAROLI, A., KUNZ, W. G., OEI, E. H., O'CONNOR, J. P. B., MAYERHOEFER, M. E., FRANCA, M., ALBERICH-BAYARRI, A., DEROOSE, C. M., LOEWE, C., MANNIESING, R., ... DESOUSA, N. M. (2021). Incorporating radiomics into clinical trials : expert consensus on considerations for data-driven compared to biologically driven quantitative biomarkers. *European Radiology*. <https://doi.org/10.1007/s00330-020-07598-8>
- GRIEVES, M. W. (2019, janvier). Virtually intelligent product systems : digital and physical twins. In S. FLUMERFELT, K. G. SCHWARTZ, D. MAVRIS & S. BRICENO (Éd.), *Complex systems engineering : theory and practice* (p. 175-200). American Institute of Aeronautics ; Astronautics, Inc. <https://doi.org/10.2514/5.9781624105654.0175.0200>
- KAUFMAN L. & ROUSSEEUW P. (1987). Clustering by means of Medoids. *Statistical Data Analysis Based on the L1-Norm and Related Methods*. Springer US.
- KETCHEN JR., D. J. & SHOOK, C. L. (1996). THE APPLICATION OF CLUSTER ANALYSIS IN STRATEGIC MANAGEMENT RESEARCH : AN ANALYSIS AND CRITIQUE. *Strategic Management Journal*, 17(6), 441-458. [https://doi.org/10.1002/\(SICI\)1097-0266\(199606\)17:6<441::AID-SMJ819>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G)
- KODINARIYA, T. & MAKWANA, P. (2013). International Journal of Advance Research in Computer Science and Management Studies. 1(6), 7.
- LI, H. & REYNOLDS, J. F. (1995). On Definition and Quantification of Heterogeneity. *Oikos*, 73(2), 280. <https://doi.org/10.2307/3545921>
- O'CONNOR, J. P. B., ABOAGYE, E. O., ADAMS, J. E., AERTS, H. J. W. L., BARRINGTON, S. F., BEER, A. J., BOELLAARD, R., BOHNDIEK, S. E., BRADY, M., BROWN, G., BUCKLEY, D. L., CHENEVERT, T. L., CLARKE, L. P., COLLETTE, S., COOK, G. J., DESOUSA, N. M., DICKSON, J. C., DIVE, C., EVELHOCH, J. L., ... WATERTON, J. C. (2017). Imaging Biomarker Roadmap for Cancer Studies. *Nat Rev Clin Oncol*, 14(3), 169-186. <https://doi.org/10.1038/nrclinonc.2016.162>
- PANDA, S., SAHU, S., JENA, P. & CHATTOPADHYAY, S. (2012). Comparing Fuzzy-C Means and K-Means Clustering Techniques : A Comprehensive Study [Series Title : Advances in Intelligent and Soft Computing]. In D. C. WYLD, J. ZIZKA & D. NAGAMALAI (Éd.), *Advances in Computer Science, Engineering & Applications* (p. 451-460). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30157-5_45

- PAPANIKOLAOU, N., MATOS, C. & KOH, D. M. (2020). How to develop a meaningful radiomic signature for clinical use in oncologic patients. *Cancer Imaging*, 20(1), 33. <https://doi.org/10.1186/s40644-020-00311-4>
- PITRE-CHAMPAGNAT, S., LEGUERNEY, I., BOSQ, J., PERONNEAU, P., KIESSLING, F., CALMELS, L., COULOT, J. & LASSAU, N. (2015). Dynamic Contrast-Enhanced Ultrasound Parametric Maps to Evaluate Intratumoral Vascularization : *Investigative Radiology*, 50(4), 212-217. <https://doi.org/10.1097/RLI.000000000000101>
- R, J. B. (2012). A comparative analysis of fuzzy c-means clustering and k means clustering algorithms. *International Journal Of Computational Engineering Research*, 737-739.
- ROKACH, L. & MAIMON, O. (2005). Clustering Methods. In O. MAIMON & L. ROKACH (Éd.), *Data Mining and Knowledge Discovery Handbook* (p. 321-352). Springer-Verlag. https://doi.org/10.1007/0-387-25465-X_15
- SARRUT, D., BAŁA, M., BARDIÈS, M., BERT, J., CHAUVIN, M., CHATZIPAPAS, K., DUPONT, M., ETXEBESTE, A., FANCHON, L. M., JAN, S., KAYAL, G., KIROV, A. S., KOWALSKI, P., KRZEMIEN, W., LABOUR, J., LENZ, M., LOUDOS, G., MEHADJI, B., MÉNARD, L., ... RONCALI, E. (2021). Advanced Monte Carlo simulations of emission tomography imaging systems with GATE. *Physics in Medicine & Biology*, 66(10), 10TR03. <https://doi.org/10.1088/1361-6560/abf276>
- THORNDIKE, R. L. (1953). Who belongs in the family? *Psychometrika*, 18(4), 267-276. <https://doi.org/10.1007/BF02289263>
- UM, H., TIXIER, F., BERMUDEZ, D., DEASY, J. O., YOUNG, R. J. & VEERARAGHAVAN, H. (2019). Impact of image preprocessing on the scanner dependence of multi-parametric MRI radiomic features and covariate shift in multi-institutional glioblastoma datasets. *Physics in Medicine & Biology*, 64(16), 165011. <https://doi.org/10.1088/1361-6560/ab2f44>
- VALLADARES, A., BEYER, T. & RAUSCH, I. (2020). Physical imaging phantoms for simulation of tumor heterogeneity in PET, CT, and MRI : an overview of existing designs. *Medical Physics*, 47(4), 2023-2037. <https://doi.org/10.1002/mp.14045>
- van TIMMEREN, J. E., CESTER, D., TANADINI-LANG, S., ALKADHI, H. & BAESSLER, B. (2020). Radiomics in medical imaging—“how-to” guide and critical reflection. *Insights into Imaging*, 11(1), 91. <https://doi.org/10.1186/s13244-020-00887-2>
- ZHANG, Q., MENG, Z., ZHANG, Y., YAO, K., LIU, J., ZHANG, Y., JING, L., YANG, X., PALIWAL, N., MENG, H. & WANG, S. (2016). Phantom-based experimental validation of fast virtual deployment of self-expandable stents for cerebral aneurysms. *BioMedical Engineering OnLine*, 15(S2), 125. <https://doi.org/10.1186/s12938-016-0250-6>

ZHANG, Q. & ZHU, S.-C. (2018). Visual Interpretability for Deep Learning : a Survey [arXiv : 1802.00614]. *arXiv :1802.00614 [cs]*. Récupérée 22 octobre 2021, à partir de <http://arxiv.org/abs/1802.00614>

Annexe

Indices de texture de la GLCM

La GLCM, matrice de co-occurrence des niveaux de gris, initialement décrite par Haralick [1] s'intéresse à l'arrangement spatial des paires de voxels pour en extraire des indices de texture. La GLCM se calcule dans 13 directions en 3D (4 en 2D) à une distance habituellement fixée à 1 voxel. Cette distance définit la relation entre les voxels qui est analysée. La GLCM est calculé à partir de la moyenne de chaque matrice évaluée dans les 13 directions. Il est possible de former 6 indices de texture à partir de la GLCM. Avant de donner leur définition, l'équation 2.8 définit la formation d'une matrice de la GLCM dans une direction pour la case (i,j) de la matrice. i et j correspondent aux niveaux de gris mis en correspondance.

$$GLCM_{\Delta x \Delta y}(i, j) = \frac{1}{\text{Paires}_{ROI}} \sum_{p=1}^{N-\Delta x} \sum_{q=1}^{M-\Delta y} \begin{cases} 1 & \text{si } (I(p, q) = i, I(p + \Delta x, q + \Delta y) = j) \\ & \text{et } I(p, q), I(p + \Delta x, q + \Delta y) \in ROI \\ 0 & \text{sinon} \end{cases} \quad (2.8)$$

$I(p, q)$ correspond au voxel (p, q) dans l'image I de taille $N \times M$. Paires_{ROI} correspond au nombre de voxels appartenant à la ROI.

Nous donnons maintenant l'équation des indices de texture calculés à partir de la GLCM à savoir l'homogénéité (2.9), l'énergie (2.10), le contraste (2.11), la corrélation (2.12), l'entropie 2.13 et la dissimilarité (2.14).

$$\text{Homogeneite} = \sum_i \sum_j \frac{GLCM(i, j)}{1 + |i - j|} \quad (2.9)$$

$$\text{Energie} = \sum_i \sum_j GLCM(i, j)^2 \quad (2.10)$$

$$\text{Contraste} = \sum_i \sum_j (i - j)^2 \cdot GLCM(i, j) \quad (2.11)$$

$$\text{Correlation} = \sum_i \sum_j \frac{(i - \mu_i) \cdot (j - \mu_j) \cdot GLCM(i, j)}{\sigma_i \cdot \sigma_j} \quad (2.12)$$

avec μ_i, μ_j la moyenne de la ligne i ou de la colonne j, et σ_i, σ_j la variance de la ligne ou de la colonne j.

$$\text{Entropie} = - \sum_i \sum_j GLCM(i, j) \cdot \log_{10}(GLCM(i, j) + \epsilon) \quad (2.13)$$

avec $\epsilon = 2e-16$.

$$Dissimilarite = \sum_i \sum_j |i - j|.GLCM(i, j) \quad (2.14)$$

[1] Haralick, Robert M., K. Shanmugam, et Its'Hak Dinstein. « Textural Features for Image Classification ». IEEE Transactions on Systems, Man, and Cybernetics SMC-3, n 6 (novembre 1973) : 610-21. <https://doi.org/10.1109/TSMC.1973.4309314>.

Chapitre 3

Exploration préclinique des biomarqueurs d'imagerie de l'hétérogénéité

Les indices quantitatifs de l'hétérogénéité tumorale développés durant cette thèse ont été appliqués en premier lieu à des données précliniques issues de l'évaluation thérapeutique d'une molécule anti-angiogénique par imagerie en IRM dynamique de contraste DCE-IRM. L'objectif de notre étude était d'évaluer l'apport potentiel de ces indices en tant que biomarqueurs précoces prédictifs de l'efficacité de ce type de traitement ciblé. Il s'agissait d'une première étude rétrospective de type "preuve de concept" réalisée sur des cartes paramétriques quantitatives obtenues dans des conditions d'expérimentation préclinique bien définies.

Après une brève introduction du processus d'angiogenèse, qui est à l'origine de l'hétérogénéité vasculaire tumorale, nous présenterons en détail une étude des performances de nos indices appliqués aux images paramétriques issues de DCE-IRM. Enfin nous discuterons du rôle potentiel de ces indices en tant que biomarqueurs.

1 Angiogenèse : source d'hétérogénéité vasculaire tumorale

Parmi les caractéristiques du développement tumoral (cf Chap 1.5), la capacité à induire et maintenir l'angiogenèse – la création de nouveaux vaisseaux sanguins pour alimenter des tissus – est une caractéristique de la genèse tumorale qui a connu beaucoup d'attention au cours des dernières décennies. En effet, dans le contexte de la médecine de précision, de nombreuses thérapies ciblant les voies de signalisation de l'angiogenèse ont fait leur apparition. Ainsi, ces thérapies sont cytostatiques et induisent généralement des modifications fonctionnelles et métaboliques de la tumeur avant toute modification morphologique. D'un point de vue des biomarqueurs, ceux extraits de

l'imagerie morphologique, permettent d'évaluer la réponse aux thérapies cytotoxiques conventionnelles, mais apparaissent peu adaptés à l'évaluation de la réponse, notamment de manière précoce dès les premières modifications induites par les thérapies ciblées. Le risque pour les patients est alors de subir une exposition prolongée à des traitements inefficaces tout en retardant l'établissement d'un protocole de soin adapté. Considérant cette limitation, il est ainsi urgent de développer de nouveaux biomarqueurs quantifiant ces modifications de manière précoce. Pour cela, une des stratégies consiste à s'appuyer sur l'imagerie médicale pour évaluer les modifications de l'hétérogénéité fonctionnelle et/ou métabolique des tumeurs, et plus particulièrement dans le cas de l'angiogenèse à l'évaluation des modifications de l'hétérogénéité vasculaire.

Ainsi, dans cette section, nous rappellerons tout d'abord brièvement le concept de l'angiogenèse et nous détaillerons son impact sur l'hétérogénéité tumorale. Enfin, nous nous intéresserons aux modalités d'imageries qui permettent de quantifier ce processus.

1.1 Le processus d'angiogenèse et l'hétérogénéité vasculaire tumorale

Les tumeurs, comme les tissus sains, requièrent un apport en nutriment, en oxygène et la possibilité d'éliminer les déchets métaboliques ainsi que le dioxyde de carbone. Ces exigences sont assurées par la néovascularisation de la tumeur suite au processus d'angiogenèse initié à partir du réseau vasculaire existant. En effet, lorsque des tissus ne sont pas suffisamment alimentés, ils génèrent des signaux biochimiques comme notamment des facteurs de croissance de l'endothélium vasculaire (VEGF : Vascular Endothelial Growth Factor) qui stimulent l'angiogenèse. Le développement tumoral repose largement sur ce processus. En effet, en l'absence de circulation sanguine, le volume tumoral ne peut dépasser 2 mm^3 , et cette absence de support vasculaire peut rendre les tumeurs nécrotiques ou apoptotiques (CARMELIET & JAIN, 2000). Dans un tissu sain, la prolifération cellulaire est naturellement régulée par des facteurs pro- et anti-angiogéniques. Néanmoins, lorsque les tumeurs acquièrent la capacité d'induction de l'angiogenèse, les facteurs pro-angiogéniques comme le VEGF sont sécrétés de manière permanente. Le réseau vasculaire tumoral nouvellement créé par cette stimulation dérégulée est totalement désorganisé et anarchique avec notamment des vaisseaux hétérogènes en dimensions, tortueux (WEIDNER et al., 1992), mal formés, dilatés et avec une quantité de bifurcations anormales (LESS et al., 1991). Les flux sanguins sont aberrants et les micro-vaisseaux sont particulièrement perméables. Une illustration schématique de ce réseau vasculaire tumoral comparé à un réseau sain a été proposée par l'équipe de TREPS et GAVARD (2015)(cf. figure 1). La perméabilité capillaire exacerbée conduit à des fuites importantes de macromolécules, ce qui engendre la formation d'œdèmes et une pression interstitielle importante conduisant à une délivrance hétérogène des traitements au sein de la tumeur (AZZI et al., 2013).

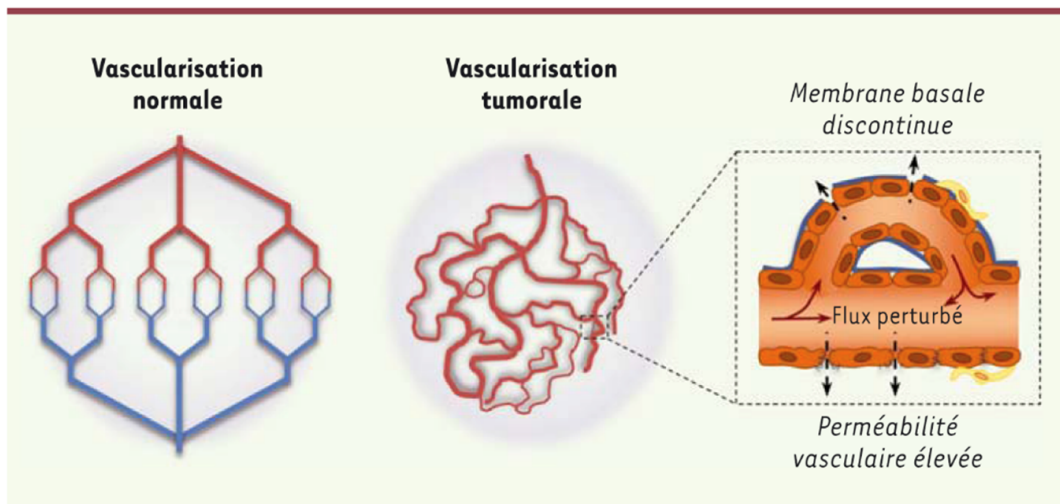


FIGURE 1 – Illustration comparative du réseau vasculaire dans les tissus sains et dans les tissus cancéreux. Le réseau vasculaire tumoral est désordonné, tortueux et perméable en raison du processus d'angiogenèse permanent. Illustration extraite de TREPS et GAVARD (2015)

Au cours des dernières décennies, les recherches ont conduit à l'apparition de traitements ciblant spécifiquement l'angiogenèse tumorale. L'objectif de ces thérapies ciblées était initialement d'inhiber le processus d'angiogenèse afin de stopper la croissance tumorale par privation de ses ressources nutritives. JAIN (2005) ont montré que l'utilisation de ces thérapies pouvait également permettre une normalisation du réseau vasculaire tumoral. Ces molécules, comme le bevacizumab, un anticorps bloquant l'expression de VEGF et qui a démontré son efficacité contre les cancers métastatiques colorectaux et rénaux en combinaison avec une chimiothérapie (CARMELIET & JAIN, 2011a), permettent en effet de normaliser la vascularisation tumorale en établissant une structure et un fonctionnement physiologique du réseau sanguin tumoral (CARMELIET & JAIN, 2011b). Cette normalisation du réseau pourrait permettre de faciliter la diffusion des traitements cytotoxiques au sein de la tumeur, mais également de la sensibiliser aux radiothérapies ou d'assister les immunothérapies. On s'intéresse alors dans la partie suivante aux possibilités offertes par l'imagerie pour étudier ces modifications vasculaires induites par des thérapies anti-angiogéniques.

1.2 Quantifier l'angiogénèse : l'imagerie de contraste, un outil incontournable

A l'inverse des thérapies cytotoxiques, les traitements anti-angiogéniques n'affectent pas nécessairement les caractéristiques morphologiques, plus particulièrement dans les temps précoces après l'initiation du traitement. De nombreuses techniques d'imagerie moléculaires ou fonctionnelles ont été développées pour suivre les modifications

physiopathologiques de la fonctionnalité et de la structure de la vascularisation tumorale. Parmi ces techniques d'imagerie on retrouve l'imagerie TEP et l'imagerie de contraste en IRM, en échographie et en scanner déjà introduites dans le chapitre 1.3.

Parmi ces modalités, l'IRM et l'échographie de contraste, possèdent l'avantage de ne pas avoir recours à des radiations ionisantes, critère utile en cancérologie en raison de leur usage répété pour le suivi de la réponse thérapeutique¹. Dans les essais cliniques, le délai entre deux acquisitions cherche à être raccourci afin de permettre une détection précoce des modifications de la vascularisation tumorale suite à l'administration d'un agent anti-angiogénique (JAIN, 2005). D'un point de vue validation des modalités d'imagerie et des biomarqueurs associés, l'IRM et l'échographie de contraste présentent toutes deux bien des différences. L'échographie de contraste, étudiée à l'échelle mono et multicentrique est largement détaillée dans différents guidelines européennes et internationales (DIETRICH et al., 2020). Le biomarqueur d'imagerie associé est l'aire sous la courbe de rehaussement du signal qui a été validé sur 539 patients répartis sur 19 centres cliniques en France. Ce paramètre permet de cibler les thérapies anti-angiogéniques efficaces pour les patients et ce après un seul mois d'administration. Cette modalité reste pour l'instant peu développée dans les services de radiologie malgré ces avantages qui sont d'être non ionisante, peu couteuse et facile d'accès. L'obtention d'un remboursement spécifique de l'acte en 2020 devrait néanmoins permettre de développer la technique. De plus, malgré l'amélioration continue de la modalité, les études quantitatives en 3D qui cherchent notamment à réduire la variabilité inter et intra-opérateur de l'échographie 2D, demeurent pour l'instant des sujets d'études mais qui pourraient apporter une forte valeur ajoutée à la technique d'après le constat que les images 2D ne peuvent pas considérer l'hétérogénéité tumorale au même titre que l'imagerie volumique 3D (GRAND-PERRET, 2019). En considérant principalement ce dernier point, la DCE-IRM apparait alors comme la modalité de choix pour l'évaluation de la vascularisation tumorale.

La modalité a été appliquée dans de nombreuses études précliniques ou cliniques pour l'évaluation de l'efficacité d'agents anti-angiogéniques et les paramètres quantitatifs extraits étaient corrélés à diverses caractéristiques physiopathologiques chez l'animal mais également chez l'humain (SALEM & O'CONNOR, 2016). Aujourd'hui, cette modalité est validée uniquement à l'échelle mono-centrique car elle souffre d'une variabilité dans les protocoles et procédures. Cette variabilité repose sur les différences intrinsèques des machines, des séquences et des procédures d'analyses qui limitent une validation multicentrique large. Dans le cas de la validation clinique de l'échographie de contraste les sources potentielles de variabilité avaient été écartées en fixant une même méthode et une même référence d'équipement dans les 19 centres (LASSAU et al., 2014). Malgré son

1. Dans le cas particulier de l'IRM de contraste, il faut néanmoins tenir compte de l'usage d'agents de contraste basés sur le gadolinium, récemment critiqué pour ses éventuels effets secondaires (SCHIEDA et al., 2020)

absence de validation multicentrique, la DCE-IRM reste une technique de choix pour la détection des variations subtiles de la micro-vascularisation tumorale et ainsi pour pouvoir évaluer à terme l'hétérogénéité globale de la tumeur (YANG & KNOPP, 2011). Et le respect scrupuleux des guidelines existantes (SHUKLA-DAVE et al., 2019) pourrait permettre de faciliter une validation à grande échelle.

La suite de ce travail sera ainsi consacrée à la présentation d'une étude préclinique en DCE-IRM 3D qui constitue une première application de la méthode développée pour quantifier l'hétérogénéité vasculaire tumorale et plus particulièrement les modifications de cette hétérogénéité au cours d'un traitement anti-angiogénique.

2 Étude comparative des indices d'hétérogénéité comme biomarqueurs d'imagerie de la réponse thérapeutique sur données préclinique

L'objectif de cette étude était de comparer les performances de nos indices à ceux issus de l'analyse de texture typiquement utilisés pour évaluer l'hétérogénéité des images. En effet, si les indices de textures ont contribué à l'importance de la prise en compte de l'hétérogénéité tumorale dans la stratégie thérapeutique, leur validité n'est pas encore établie et leur interprétation reste difficile pour décrire une évolution physiologique. Nos indices pourraient être complémentaires, voire plus pertinents, pour évaluer ces modifications au cours du temps.

Après avoir décrit le protocole expérimental ayant servi à générer les données, incluant le modèle animal choisi, la molécule thérapeutique administrée et la procédure d'acquisition des images, nous détaillerons dans une seconde partie la méthodologie appliquée à ce jeu de données précliniques. Les résultats obtenus seront également présentés avec une comparaison aux indices de textures en tant que biomarqueurs potentiels de l'hétérogénéité vasculaire pour l'évaluation de la réponse thérapeutique.

2.1 Description du protocole expérimental

Notre étude préclinique s'appuie sur des données acquises en 2009 par l'équipe d'Andreas Volk, co-responsable de ces travaux de thèse. L'objectif initial de l'étude était d'évaluer l'efficacité de la DCE-IRM 3D pour suivre l'efficacité du Sunitinib comme agent anti-angiogénique sur les tumeurs. Aujourd'hui cette molécule est utilisée en routine clinique pour les cancers du rein et les tumeurs gastro-intestinales. Dans cette partie nous décrivons le protocole expérimental ayant servi à collecter les données et qui nous permet rétrospectivement de les analyser. Nous présenterons donc brièvement, le modèle souris, le protocole de traitement et enfin le protocole d'acquisition d'images employé

(EL TANNIR EL TAYARA et al., 2011 ; VAUTIER et al., 2013).

Modèle préclinique et tumoral

L'étude a porté sur une cohorte de 15 souris immunodéficientes (nudes SWISS). Les souris ont subi une xéngreffe² de tumeur humaine colorectale (TC302, Institut Curie, France), avec implantation en sous-cutanée au niveau abdominal.

Protocole thérapeutique Les souris étaient réparties en deux groupes : groupe traité (n=9) et groupe contrôle (n=6). Le traitement anti-angiogénique consistait en une administration orale quotidienne de Sunitinib® de 60 mg/kg à partir du jour 1, les souris contrôles recevaient une solution véhicule- seule. Le Sunitinib® est un inhibiteur de multiples récepteurs de tyrosine kinase impliqués dans le processus d'angiogenèse au même titre que les VEGF décrits précédemment.

Le volume tumoral au jour 1 était en moyenne de 419 ± 203 mm³ et les souris étaient suivies sur une durée de 10 jours post-traitement par IRM anatomique et 3D DCE-IRM pour la mesure du volume tumoral et la génération de cartes paramétriques quantitatives. Le protocole thérapeutique et d'imagerie est schématisé sur la figure 2.

Les examens IRM d'intérêt pour notre étude étaient ceux acquis avec la DCE-IRM à Jour 1 (juste avant le début du traitement) et à Jour 4. Plus précisément, nous avons analysé les cartes paramétriques K^{trans} de la perméabilité et de la perfusion tumorale générées à partir de ces acquisitions. Les résultats des examens anatomiques de l'étude initiale ont servi de référence pour confirmer la réponse thérapeutique sur la base du volume tumoral, en croissance pour le groupe contrôle, notamment aux examens du jour 11, et en stagnation pour le groupe traité (EL TANNIR EL TAYARA et al., 2011).

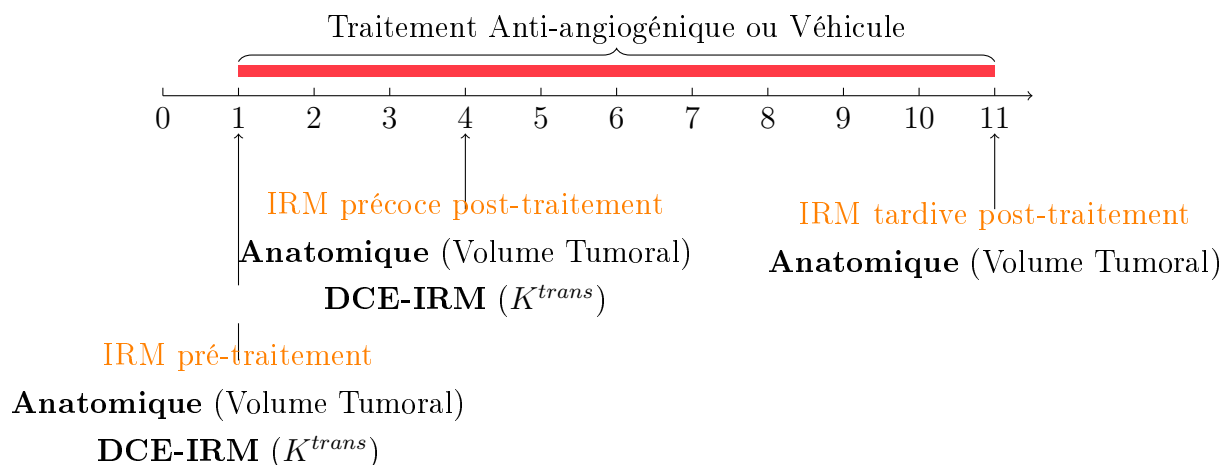


FIGURE 2 – Protocole de traitement et d'acquisition des images au cours de l'étude pré-clinique sur souris soumises à un traitement anti-angiogénique et suivies sur une période de 11 jours.

Protocole d'acquisition d'image en DCE-IRM (VAUTIER et al., 2013)

2. Désigne la transplantation d'un greffon ou le donneur et le receveur sont d'origine biologique différentes

Les acquisitions IRM ont été réalisées avec un appareil IRM 4,7 T dédié au petit animal (Bruker Biospin, Allemagne). La séquence employée était la séquence RETIA (VAUTIER et al., 2013), une technique d'acquisition radiale 3D multi écho de gradient. Son schéma d'acquisition radial la rend peu sensible aux mouvements physiologiques, la 3D permet d'accéder à l'hétérogénéité de la tumeur dans sa globalité par la génération de cartes paramétriques en 3D et finalement le multi-écho de gradient permet de corriger les effets $R2^*$. A cela s'ajoute l'acquisition simultanée d'une fonction d'entrée artérielle, (AIF : Arterial Input Function) sur une coupe du cœur avec une résolution temporelle élevée (2 sec). L'antenne non-commerciale utilisée était une sonde quadrature de type "cage d'oiseau" fabriquée par Joël Mispelter (MISPELTER et al., 2015) de diamètre $D=35$ mm et de longueur $l=50$ mm.

Avant l'injection du produit de contraste, des images anatomiques à haute-résolution spatiale étaient acquises avec le même champ de vue 3D que lors de l'acquisition radiale DCE-IRM. Ces images couvrant la totalité de la tumeur étaient acquises avec une séquence écho de spin multi-coupes déclenchée par la respiration. Les paramètres de la séquence étaient : $TR \approx 1$ sec, $TE = 10,7$ msec, matrice 128×128 , épaisseur de coupe $0,47$ mm, nombre de coupes 64. Ces images servaient à la segmentation des tumeurs.

Durant l'acquisition, les animaux étaient anesthésiés avec de l'isoflurane (AErrane®), Baxter, France) et la température corporelle stabilisée. L'agent de contraste (AC) utilisé pour l'acquisition dynamique était un AC expérimental P846 (Guerbet, France) de taille moyenne avec une masse moléculaire de $3,5$ kDa. Il était injecté à une dose de $0,089$ mmol/kg avec un débit d'injection de $600 \mu\text{L}/\text{min}$ pour un volume total de $200 \mu\text{L}$.

Au cours d'une acquisition en DCE-IRM, étaient continuellement acquis 16 jeux de données 3D correspondant à la tumeur incluant une acquisition 3D pré-contraste, et simultanément, 1024 (64×16) images 2D au niveau du cœur permettant de mesurer l'AIF à haute résolution temporelle. Les paramètres des séquences 2D et 3D sont résumés dans le tableau 3.1 et la figure 3 illustre l'acquisition simultanée d'un volume 3D et de 64 fois la même coupe 2D du cœur.

Les images ont été reconstruites par un algorithme de « regridding » standard en C++ (JACKSON et al., 1991) avec une résolution spatiale isotrope de $500 \mu\text{m}$. Pour les données 3D l'application d'une procédure « sliding window » a conduit à une résolution temporelle de 30 s. Les concentrations d'AC dans les voxels (tumeur) ou dans la ROI de l'AIF (cœur) ont été estimées à partir des valeurs $R1(t)$ corrigées par $R2^*$ en utilisant une relaxivité $r1 = 15 \text{ s}^{-1} \text{ mM}^{-1}$. Les constantes de temps de l'AIF ont été mesurées par un ajustement bi-exponentiel des courbes temps-concentration.

L'étude préclinique initiale a porté entre autres sur l'extraction des cartes paramétriques de la perméabilité vasculaire via la quantification des valeurs K^{trans} dans

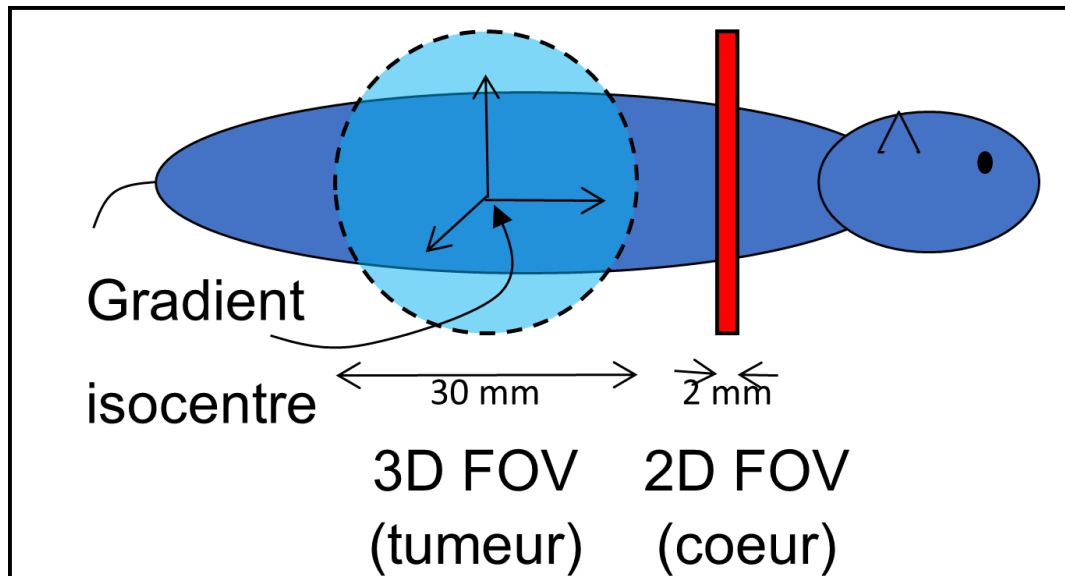


FIGURE 3 – Principe de l'acquisition simultanée du volume 3D tumoral et de 64 coupes 2D centrées sur le cœur pour mesurer l'AIF. Reproduite de EL TANNIR EL TAYARA et al. (2011).

	IRM de Contraste	
	Volume 3D	Coupe 2D
TE1 (ms)	0,9	1,9
δ TE (ms)	1,11	1,11
TR (ms)	31	31
Largeur de bande (kHz)	100	100
Echos	10	10
Matrice	64×64×64	64×64
Champ de vue (FOV)	30×30×30 mm ³	30×30 mm ²
Points de lecture	64	64
Projections	4096	64
Temps d'acquisition	2 min	2 s
Epaisseur de coupe (mm)	0,47	2

TABLE 3.1 – Paramètres d'acquisitions de la séquence RETIA pour l'acquisition dynamique de contraste.

chaque voxel à l'aide du modèle pharmacocinétique de Tofts-Kermode (TOFTS & KERMODE, 1991). Le paramètre K^{trans} correspond à la constante de transfert à travers la paroi capillaire vers le compartiment interstitiel et représente une combinaison complexe entre la perfusion et la perméabilité capillaire. Le volume tumoral était segmenté manuellement à partir des images haute-résolution écho de spin puis le contour était transféré sur les cartes K^{trans} . C'est donc l'hétérogénéité de ces cartes paramétriques que nous avons cherché à évaluer à partir des indices ad hoc. Notons que le processus mathématique de reconstruction des cartes K^{trans} produit des voxels de valeur non-définis, annotés NaN (« not-a-number »). Ces voxels résultaient soit d'un ajustement qui ne correspondaient pas à des valeurs physiques acceptables, soit d'un faible rapport signal sur bruit, soit d'une augmentation R1 insuffisante (VAUTIER et al., 2013). La proportion de voxels NaN était de $23,8\% \pm 13,2\%$. Ils ont été exclus des analyses suivantes.

2.2 Quantification de l'hétérogénéité tumorale

La méthodologie de quantification de l'hétérogénéité développée précédemment a donc été appliquée aux cartes paramétriques K^{trans} dérivées de l'imagerie DCE-IRM préclinique. Dans le contexte de l'évaluation de la réponse thérapeutique, nous nous sommes intéressés à l'évolution de nos indices d'hétérogénéité entre les cartes acquises avant l'initiation de la thérapie (Jour 1) et celles acquises Jour 4. Pour rappel, notre méthode consiste en l'application successive d'un processus de clustering puis d'extraction de métriques de la complexité spatiale pour former un biomarqueur d'imagerie. Nous détaillons désormais l'application de chacune de ces étapes.

2.2.1 Clustering

L'algorithme des K-moyennes a été employé pour le clustering. Il a fallu tout d'abord définir le nombre de clusters optimal pour partitionner les cartes. Pour rappel, chaque cluster représente un ensemble de voxels aux valeurs paramétriques, ici celle du K^{trans} , similaires, c'est à dire un ensemble de voxels où le tissu a des propriétés de perfusion et de perméabilité similaires.

La recherche du nombre de clusters optimal a été réalisée dans la plage $K = [2, 8]$. Pour chaque K, le clustering était répété 30 fois avec des initialisations aléatoires afin de réduire la probabilité de résultat convergeant vers des minima locaux. Les centroïdes résultant sont ceux qui sont apparus majoritairement parmi les 30 répétitions.

Cette optimisation a été appliquée à chacune des cartes paramétriques. Toutefois, pour rendre possible la comparaison des cartes entre elles après le clustering, elles ont toutes été soumises au processus de clustering avec le même nombre de cluster. Le nombre de cluster optimal K choisi était celui avec le plus grand taux d'apparition

parmi toutes les cartes. Une fois le choix établi, les cartes ont été soumises au processus de clustering avec K clusters à former. Le clustering est de nouveau effectué 30 fois pour éviter la convergence vers des minima locaux et le résultat correspond aux centroïdes apparu le plus grand nombre de fois au cours des 30 répétitions.

Ce processus d'optimisation et de clustering est effectué sous MATLAB (The MathWorks Inc., Natick, MA, USA).

2.2.2 Extraction et sélection des paramètres

Pour décrire la complexité de l'organisation spatiale des cartes, les indices d'hétérogénéité décrits dans le chapitre 2.2.2 à savoir le nombre de patches, la fraction volumique des clusters et la distance moyenne inter-patches ont été extraits. Ces indices permettent ainsi de former les H-cluster et le H-index. Pour l'extraction de l'indice nombre de patches, la connectivité employée pour l'analyse en composantes connectées était de 6.

Ces indices ont été comparés à des indices du second-ordre statistique évalués sur le volume d'intérêt tumoral entier. Parmi l'ensemble des indices du second-ordre disponibles, notre attention s'est focalisée sur ceux spécifiquement liés à l'hétérogénéité spatiale. Ainsi les 6 indices de texture retenues étaient issus de la matrice GLCM : Homogénéité, Energie, Contraste, Corrélation, Entropie, Dissimilarité. Les définitions mathématiques des paramètres de texture employés au cours de cette étude sont fournis en annexe du chapitre 2. Ces paramètres étaient extraits à l'aide du logiciel en libre accès LIFEx en utilisant les paramètres de calcul suivant : nombre de niveaux de gris 256 – Mise à l'échelle des intensités absolues avec valeur maximum de $K^{trans} = 0,8 \text{ min}^{-1}$ correspondant à la valeur maximum de K^{trans} parmi toutes les cartes à notre disposition. Pour la définition mathématique de l'ensemble des indices disponibles dans LIFEx, le lecteur peut se diriger vers la documentation officielle disponible à l'adresse <http://www.lifex.soft.org> (Nioche et al, 2018, Orsay, France).

Ont été également comparés des indices statistiques conventionnels du premier ordre qui décrivent la distribution des intensités du volume tumoral entier et de chacun des clusters. La valeur moyenne, la médiane et le volume étaient calculés pour la ROI tumorale 3D mais également au sein de chaque cluster. La moyenne et la médiane visaient à montrer des différences globales de la distribution de la perfusion et de la perméabilité. Le volume tumoral était quant à lui employé en tant que biomarqueur de référence de l'efficacité thérapeutique.

Enfin, la corrélation entre tous les indices a été analysée et établie avec un coefficient de Pearson supérieur ou égal à 0,95. En cas de corrélation dans un groupe d'indices, un seul indice était conservé. Les groupes correspondent aux indices formant le H_{index} , aux indices de textures et aux indices conventionnels.

2.2.3 Analyses Statistiques

L'objectif de l'analyse était de démontrer si un ou plusieurs indices d'hétérogénéité d'image étaient modifiés significativement par l'effet du traitement. Pour ce faire, les indices ont été analysés individuellement selon trois types de comparaison :

- (i) L'évolution entre le Jour 1 et le Jour 4 au sein du groupe de souris contrôles
- (ii) L'évolution entre le Jour 1 et le Jour 4 au sein du groupe de souris traitées
- (iii) Les différences d'évolution entre Jour 1 et Jour 4 entre les deux groupes

Pour les comparaisons (i) et (ii), les valeurs des indices pré-traitement (Jour 1) et post-traitement (Jour 4) ont été soumises à des tests de Wilcoxon appariés. Pour la comparaison (iii), la différence de valeurs des indices entre avant et après traitement a été comparée entre les groupes à l'aide de tests de Wilcoxon non-appariés. Les résultats statistiques présentant une valeur de test p inférieure ou égale à 0,05 ont été considérés comme démontrant une différence statistique significative. Une correction pour comparaison multiple a été appliquée à tous les indices sélectionnés à l'aide d'une procédure de Benjamini-Hochberg (BH). Au cours de cette procédure, le taux de fausse-découverte était appliqué à 0,05. Toutes les analyses statistiques ont été conduites sous R (R Core Team, 2019).

2.3 Résultats

2.3.1 Optimisation du processus de clustering

Le processus d'optimisation a été appliqué à chacune des 30 cartes paramétriques disponibles. Il résulte de ce processus que le nombre de clusters optimal pour ce jeu de données est $K=4$. En effet, bien que la méthode du coude illustrée figure 4 ne permette pas d'établir ce nombre de manière franche, la méthode établissant un compromis entre la variance intra-cluster moyenne (VICM) et la distance moyenne inter-clusters, (DMIC) dont les résultats sont reportés figure 5, indique qu'une majorité des cartes ($n=21$) trouvent leur nombre optimal de clusters à $K=4$, tandis qu'une unique carte possède un $K=3$ et huit cartes trouvent un nombre de clusters optimal à $K=5$ (figure 6). De plus la différence moyenne de la somme de VICM et DMIC était très faible entre $K=4$ et $K=5$ ($0,033 \pm 0,030$). Ainsi, la valeur majoritaire a été retenue et le nombre de clusters optimal alors employée par la suite a été $K=4$ correspondant au meilleur compromis pour l'ensemble des données.

Les clusters sont par la suite numérotés #1 à #4 dans l'ordre croissant de leur valeur moyenne de K^{trans} correspondant à la valeur du centroïde. Ces valeurs moyennes pour l'ensemble des cartes étaient respectivement de 0,8 ; 2,2 ; 4,4 et 9,9 (10^{-2} min^{-1}). La figure 7 illustre le résultat du clustering sur une unique coupe et propose également

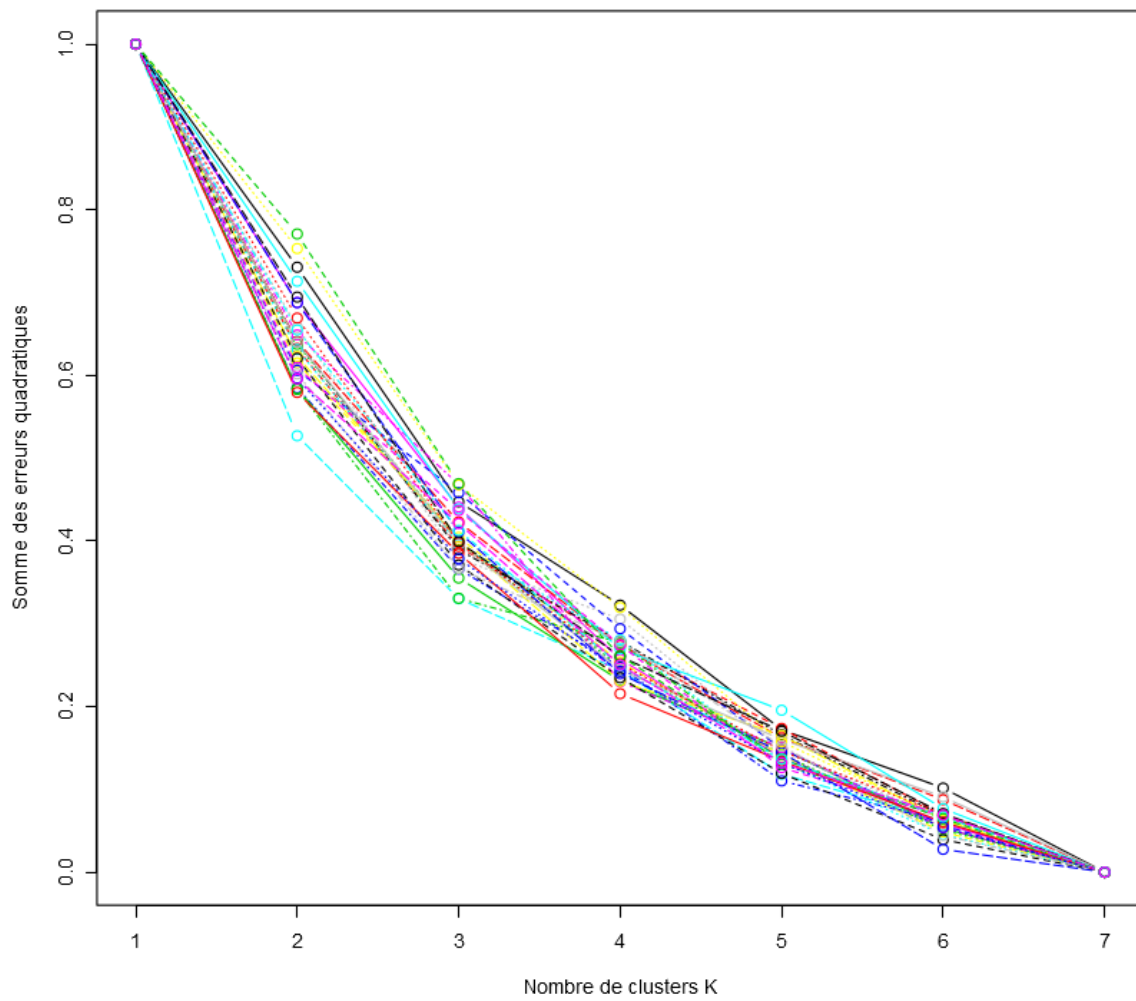


FIGURE 4 – Méthode du coude appliquée à chacune des 30 cartes paramétriques pour déterminer le nombre de clusters K optimal.

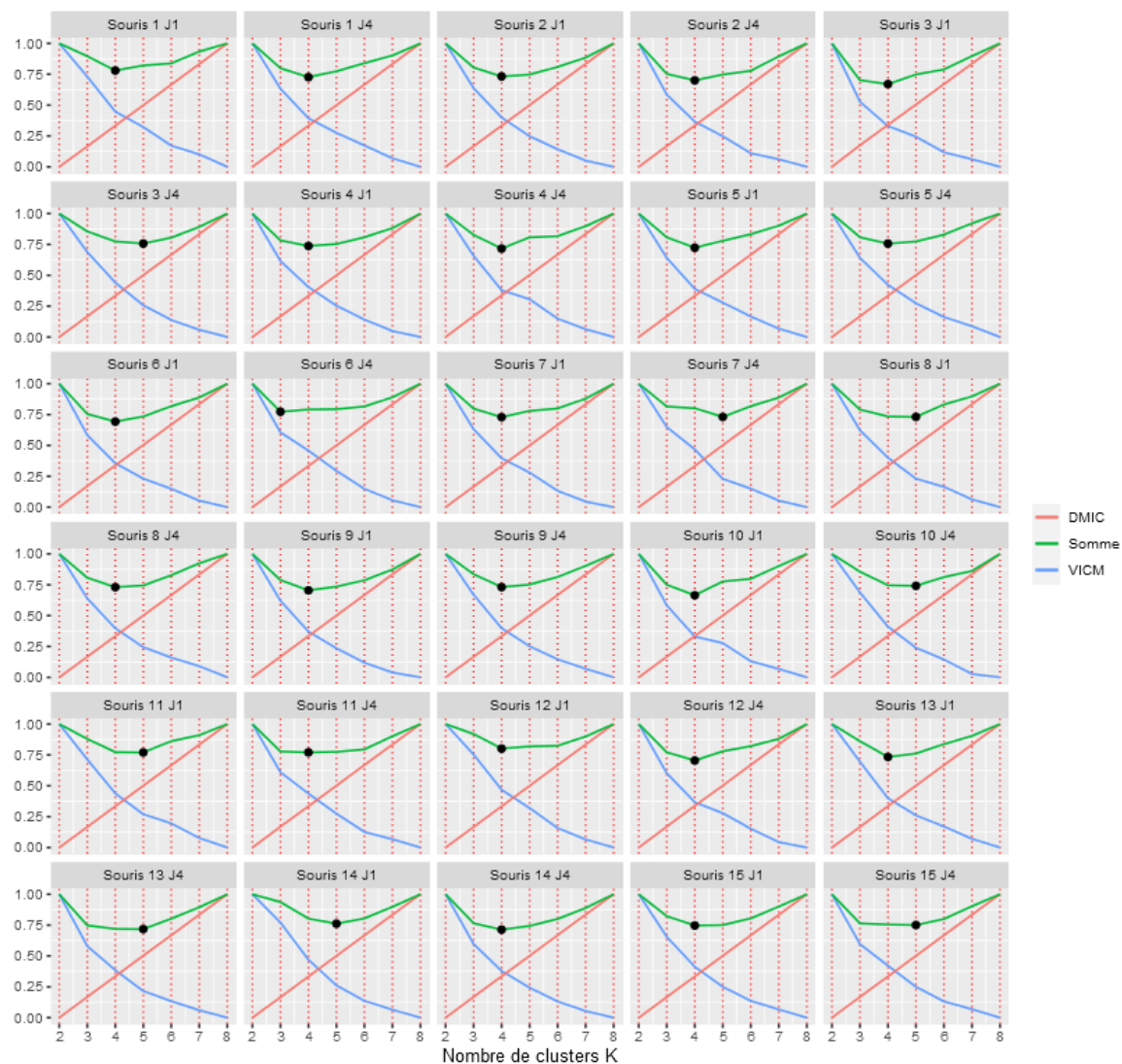


FIGURE 5 – Méthode de minimisation de la variance intra-cluster et de maximisation de la distance inter-clusters appliquée à chacune des 30 cartes paramétriques pour déterminer le nombre de clusters K optimal.

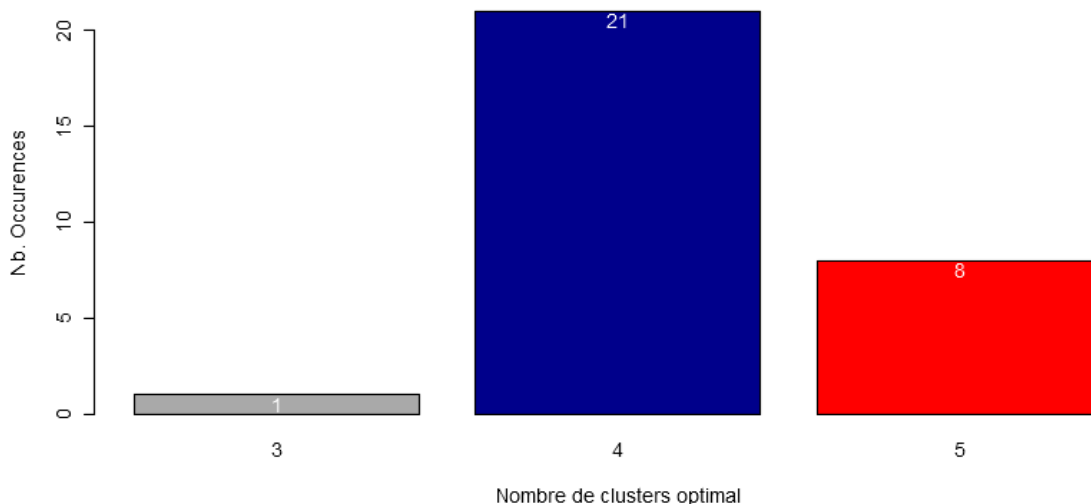


FIGURE 6 – Bilan des résultats d’optimisation du nombre de clusters optimal pour chaque carte. La majorité des cartes paramétriques sont au mieux représentées par un nombre $K=4$ clusters.

une superposition de chacun des clusters à l’image originale. Ces clusters ont ensuite été utilisés de manière individuelle en tant que carte catégorique où chaque pixel est défini comme appartenant au cluster ou non, et peuvent également être utilisés comme masque pour extraire les sous-populations de pixels et leurs valeurs d’intensités pour chaque cluster étudié et notamment pour calculer les indices statistiques conventionnels.

2.3.2 Appréciation Visuelle de l’hétérogénéité

Une appréciation visuelle des coupes 2D extraites des cartes paramétriques 3D permet d’effectuer une évaluation qualitative de l’évolution de l’hétérogénéité dans les images. Comme illustré sur la figure 8 à l’aide de coupes représentatives, l’hétérogénéité tumorale du K^{trans} évolue sur la période des trois premiers jours de suivi. En effet, dans le cadre du groupe traité, cette hétérogénéité semble diminuer au cours du temps avec le transfert des patches au sein des clusters intermédiaires (#2 et #3) vers le cluster de faible intensité (#1). A l’inverse, pour le groupe contrôle, la complexité dans l’organisation spatiale des clusters tend à augmenter au cours du temps avec des disparités plus importantes dans les niveaux d’intensités et leur organisation. La méthode que nous avons développée cherche ainsi à quantifier cette hétérogénéité spatiale afin d’en comparer les évolutions au sein et entre chacun des groupes.

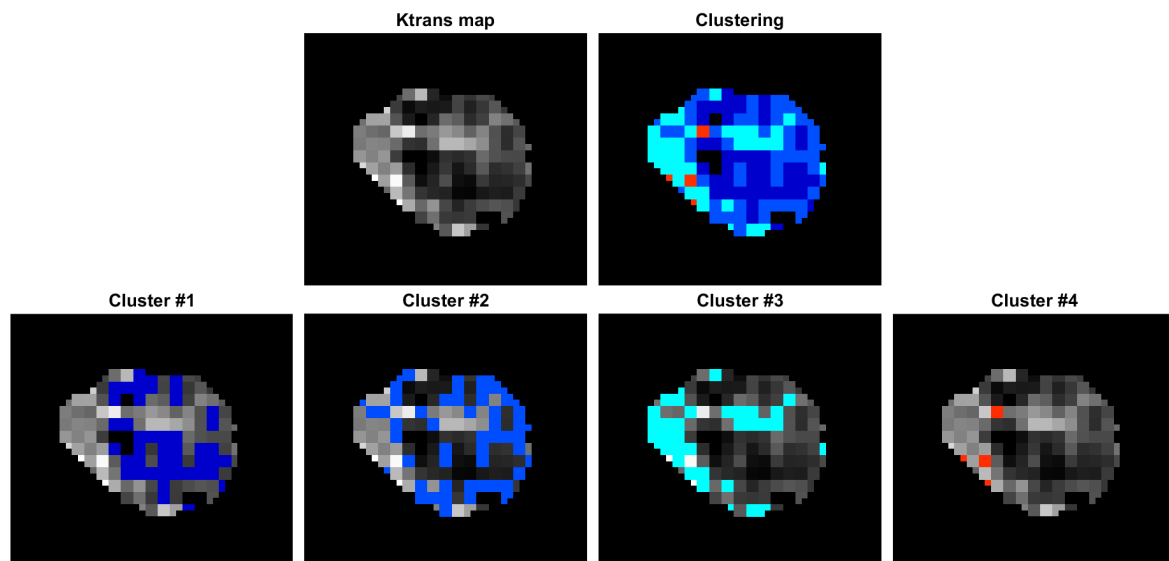


FIGURE 7 – Unique coupe 2D extraite d'une carte paramétrique 3D du K^{trans} située en haut à gauche. En haut à droite, est présenté le résultat du clustering. En bas, chacun des quatre clusters sont illustrés de manière isolée et superposés à leur carte paramétrique, ordonnés dans l'ordre croissant de la valeur moyenne de K^{trans} , celle des centroïdes.

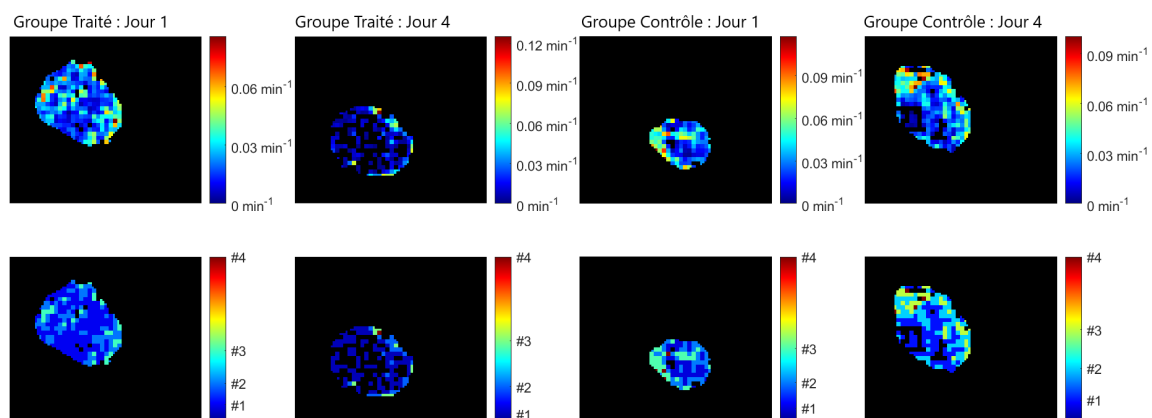


FIGURE 8 – Unique coupe 2D extraite d'une carte paramétrique 3D pour une souris traitée (gauche) et une souris contrôle (droite). On observe une évolution de l'hétérogénéité entre les groupes contrôles et traités à J1 et à J4. La ligne du haut représente les cartes paramétriques originales et la ligne de bas le résultat du clustering.

2.3.3 Extraction et sélection d'indices

Au total, ce sont 38 indices qui ont été extraits de chaque carte paramétrique : 17 de nos indices dont 4 $H_{cluster}$ et le H_{index} ; 15 indices statistiques conventionnels dont 3 indices au sein des 4 clusters et 3 indices au sein de la ROI tumorale complète ; enfin, 6 indices de textures ont été retenus et établis sur la ROI tumorale complète.

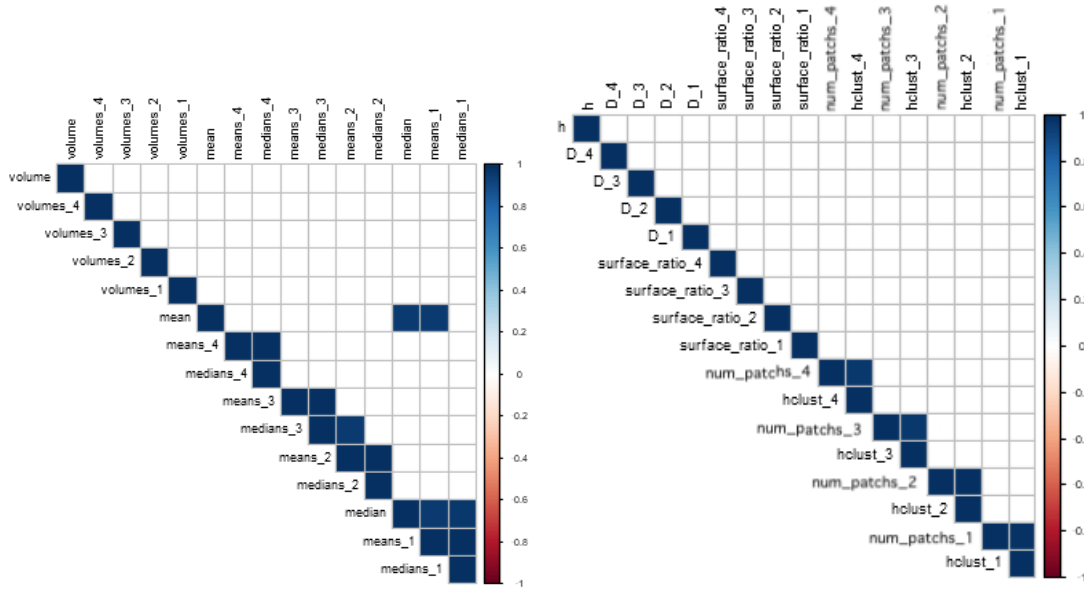
A partir de ces 38 indices, les corrélations ont été étudiées pour sélectionner un nombre limité d'indices indépendants (figure 9). Concernant les paramètres conventionnels, la moyenne et la médiane étaient corrélées pour chaque cluster et pour la ROI. Seule la médiane a ainsi été conservée, à l'exception de la médiane du cluster #1 corrélée à la médiane de la ROI. Parmi les indices de textures extraits de la GLCM, le contraste et l'entropie étaient respectivement corrélés à la dissimilarité et à l'homogénéité et n'ont donc pas été retenus. Parmi nos indices d'hétérogénéité, les $H_{cluster}$ et le nombre de patches de chaque cluster étaient également corrélés. Seuls les $H_{cluster}$ ont ainsi été conservés. Une ultime vérification a été effectuée pour s'assurer de la non-corrélation de l'ensemble de ces indices. Aucun indice supplémentaire n'a ainsi été rejeté (figure 10).

Finalement, un total de 26 indices a été retenu :

- 13 indices de complexité
 - 4 fractions volumiques
 - 4 distances moyennes inter patches
 - 4 $H_{cluster}$
 - H_{index}
- 9 indices conventionnels
 - 4 médianes (clusters #2 à #4 et ROI)
 - 5 volumes (clusters #1 à #4 et ROI)
- 4 indices de texture
 - Énergie, Corrélation, Homogénéité, Dissimilarité

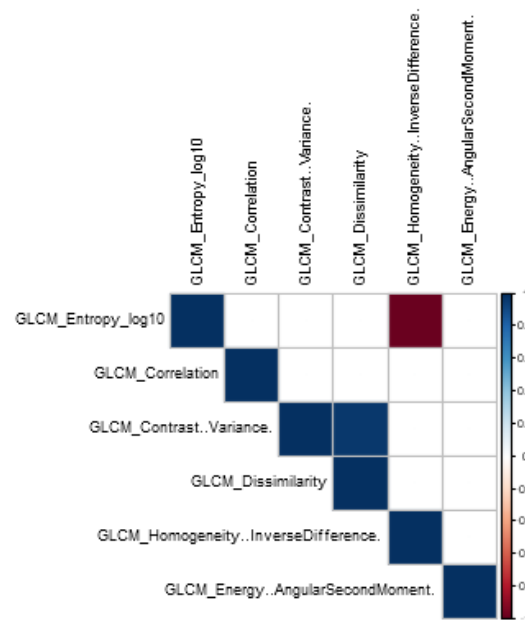
2.3.4 Analyses statistiques

Avant administration du traitement au jour 1, nous observons une absence de différence significative entre les valeurs des 26 indices des deux groupes. La table 3.2 permet ainsi de vérifier une homogénéité des deux groupes avant initiation du traitement.



(a)

(b)



(c)

FIGURE 9 – Corrélations entre les indices : (a) conventionnelles ; (b) de complexité intratumorale dont le H_{index} et (c) de texture issus de la GLCM. Seules les corrélations dont le coefficient de Pearson est ≥ 0.95 sont affichées.

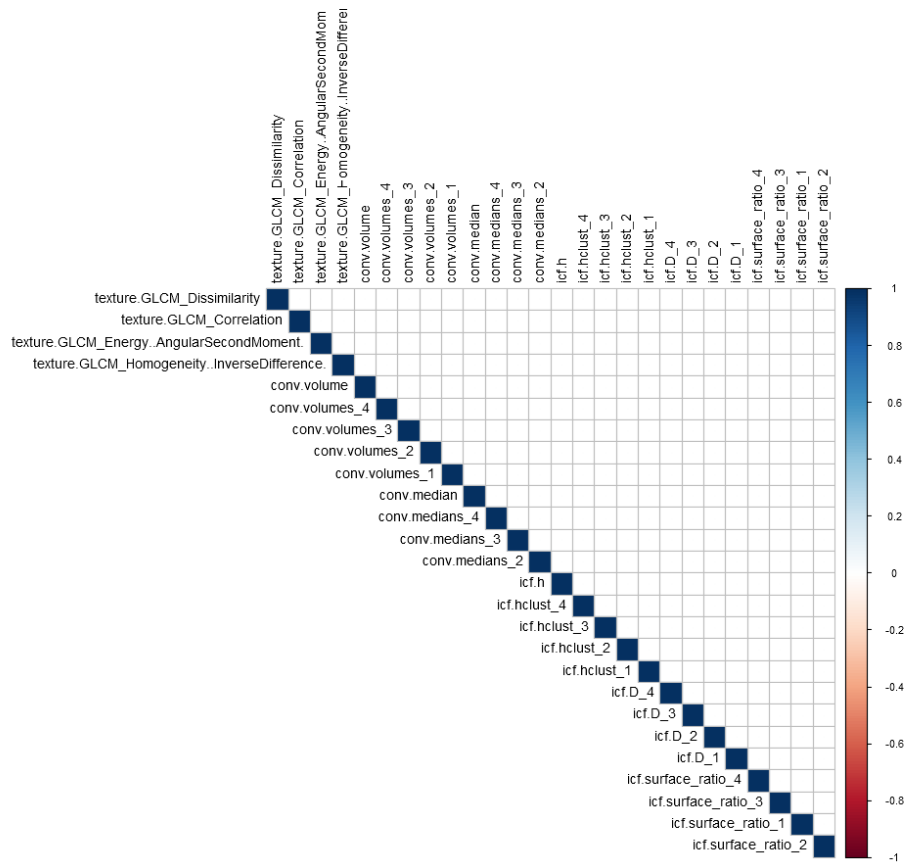


FIGURE 10 – Corrélations entre tous les indices après élimination de ceux corrélés dans les différents groupes.

	Indices	p-value	Contrôle Moyenne \pm ec.type	Traité Moyenne \pm ec.type	
Sous-régions / clusters	Complexité				
	Fraction volumique #1 (%)	0.863	50.2 \pm 2.6	50.4 \pm 1.4	
	Fraction volumique #2 (%)	0.863	35.9 \pm 1.3	35.9 \pm 0.8	
	Fraction volumique #3 (%)	0.776	12.1 \pm 1.8	12.0 \pm 0.7	
	Fraction volumique #4 (%)	0.529	1.8 \pm 0.3	1.6 \pm 0.3	
	Distance moyenne #1 (mm)	0.529	6.062 \pm 0.145	6.253 \pm 0.295	
	Distance moyenne #2 (mm)	1	5.774 \pm 0.141	5.736 \pm 0.306	
	Distance moyenne #3 (mm)	0.776	6.063 \pm 0.168	5.964 \pm 0.340	
	Distance moyenne #4 (mm)	0.954	6.024 \pm 0.253	6.04 \pm 0.393	
	$H_{cluster}$ #1	0.954	390 \pm 35	497 \pm 139	
	$H_{cluster}$ #2	0.689	769 \pm 111	963 \pm 174	
	$H_{cluster}$ #3	0.864	1025 \pm 137	1074 \pm 145	
	$H_{cluster}$ #4	0.864	334 \pm 58	303 \pm 45	
	H_{index}	0.607	498 \pm 51	578 \pm 89	
		Statistiques conventionnelles			
		Médiane #2 (min ⁻¹)	0.327	0.024 \pm 0.003	0.020 \pm 0.003
		Médiane #3 (min ⁻¹)	0.388	0.045 \pm 0.007	0.036 \pm 0.006
		Médiane #4 (min ⁻¹)	0.689	0.095 \pm 0.016	0.082 \pm 0.014
		Volume #1 (mm ³)	1	199.133 \pm 15.144	206.165 \pm 27.255
		Volume #2 (mm ³)	0.776	142.993 \pm 9.995	145.081 \pm 17.009
	Volume #3 (mm ³)	0.775	48.053 \pm 7.765	48.809 \pm 6.681	
	Volume #4 (mm ³)	0.607	7.330 \pm 1.431	6.412 \pm 1.223	
Région Tumorale ROI	Statistiques conventionnelles				
	Médiane (min ⁻¹)	0.388	0.017 \pm 0.003	0.014 \pm 0.002	
	Volume (mm ³)	0.776	474.496 \pm 24.915	503.656 \pm 76.868	
	Texture				
	Homogénéité	0.456	0.466 \pm 0.032	0.495 \pm 0.028	
	Énergie	0.327	0.017 \pm 0.005	0.025 \pm 0.006	
	Corrélation	0.049	0.486 \pm 0.015	0.43 \pm 0.023	
	Dissimilarité	0.606	3.14 \pm 0.537	2.709 \pm 0.436	

TABLE 3.2 – Valeurs moyennes, écarts types, et valeurs de p résultantes des tests de Wilcoxon non-appariés entre le groupe contrôle et le groupe traité au jour 1 avant l'initiation du traitement.

Dans le cadre de la première comparaison, à savoir l'évolution des indices entre les jours 1 et 4 chez les souris contrôles, les valeurs de p des analyses statistiques sont reportées dans la table 3.3 ainsi que les valeurs moyennes et les déviations standards de chacun des indices aux examens IRM des jours 1 et 4. Les tests de Wilcoxon appariés ont été effectués sur les valeurs des indices pour les deux examens au sein de chaque groupe. La valeur q correspond au résultat de la correction pour de multiples comparaison par la méthode de Benjamini-Hochberg (BH) et est reportée pour tous les indices dont la significativité statistique persistait après correction. Au sein du groupe contrôle, aucun indice ne présente de différence statistique significative au cours du temps après correction de BH.

La seconde comparaison visant à étudier l'évolution des indices au cours des jours 1 et 4 chez la souris traitée est résumée dans la table 3.4. Dans ce groupe, on note plusieurs indices variant de manière significative. Ainsi, parmi les 13 indices évaluant la complexité à l'échelle des clusters, incluant le H_{index} , après correction de BH, quatre indices font preuve d'une différence statistique significative avant et après traitement (table 3.5) : les fractions volumiques des clusters #1 ($p=0,004$), #2 ($p=0,004$), #3 ($p=0,012$), et le $H_{cluster}$ #3 ($p=0,004$). Parmi les indices de texture extraits de la GLCM, et après correction de BH, aucun indice ne permet de démontrer une différence significative entre le début du traitement anti-angiogénique et J4. Sur les 9 indices de statistiques conventionnels, 3 sont significativement différents au cours du temps : les volumes des clusters #2 ($p=0,004$) et #3 ($p=0,004$) et la médiane de la ROI tumorale ($p=0,008$).

Enfin, la dernière comparaison cette fois-ci entre les deux groupes est présentée dans la table 3.5 qui reporte de la même manière les valeurs moyennes et les écarts types des deltas-indices, à savoir la différence pour chaque groupe entre la valeur de l'indice à l'examen du jour 4 et celle à l'examen du jour 1. Les résultats des analyses statistiques de Wilcoxon non-appariés sont également reportés ainsi que le résultat de la correction de BH pour les indices significatifs après l'application de la correction. Parmi nos 13 indices, 4 permettaient de différencier statistiquement leur évolution entre les groupes : les fractions volumiques des clusters #1 ($p=0.005$) et #2 ($p=0.0008$), le $H_{cluster}$ #3 ($p=0.003$) et le H_{index} ($p=0.003$). Les indices de texture ne permettent pas non plus dans ce cas de démontrer une différence significative entre les groupes. Enfin, 3 indices de statistiques conventionnelles permettent également de différencier significativement l'évolution entre les deux groupes : les volumes du cluster #2 ($p=0,0004$), #3 ($p=0,012$) et celui de la ROI tumorale ($p=0,003$).

Indices	p-value	Pré- traitement Moyenne \pm ec.type	Post- Traitement Moyenne \pm ec.type	q-value	
Sous-régions / clusters	Complexité				
	Fraction volumique #1 (%)	0.687	50.2 \pm 2.6	51.8 \pm 3.7	-
	Fraction volumique #2 (%)	0.562	35.9 \pm 1.3	34.4 \pm 1.7	-
	Fraction volumique #3 (%)	1	12.1 \pm 1.8	12.0 \pm 1.9	-
	Fraction volumique #4 (%)	1	1.8 \pm 0.3	1.8 \pm 0.5	-
	Distance moyenne #1 (mm)	0.094	6.062 \pm 0.145	6.549 \pm 0.123	-
	Distance moyenne #2 (mm)	0.094	5.774 \pm 0.141	6.186 \pm 0.109	-
	Distance moyenne #3 (mm)	0.062	6.063 \pm 0.168	6.486 \pm 0.117	-
	Distance moyenne #4 (mm)	0.031	6.024 \pm 0.253	6.677 \pm 0.17	-
	H-cluster #1	0.062	390 \pm 35	544 \pm 83	-
	H-cluster #2	0.062	769 \pm 111	1090 \pm 132	-
	H-cluster #3	0.219	1025 \pm 137	1181 \pm 110	-
	H-cluster #4	0.844	334 \pm 58	406 \pm 89	-
	H-index	0.031	498 \pm 51	645 \pm 69	-
	Statistiques conventionnelles				
	Médiane #2 (min ⁻¹)	0.687	0.024 \pm 0.003	0.021 \pm 0.003	-
	Médiane #3 (min ⁻¹)	0.312	0.045 \pm 0.007	0.039 \pm 0.004	-
	Médiane #4 (min ⁻¹)	0.437	0.095 \pm 0.016	0.08 \pm 0.008	-
	Volume #1 (mm ³)	0.312	199.1 \pm 15.1	245.7 \pm 24.4	-
	Volume #2 (mm ³)	0.219	143.0 \pm 10.0	162.5 \pm 10.9	-
	Volume #3 (mm ³)	0.562	48.0 \pm 7.8	56.1 \pm 8.8	-
	Volume #4 (mm ³)	1	7.3 \pm 1.4	8.4 \pm 2.3	-
Région Tumorale ROI	Statistiques conventionnelles				
	Médiane (min ⁻¹)	0.437	0.017 \pm 0.003	0.015 \pm 0.003	-
	Volume (mm ³)	0.062	474.496 \pm 24.915	585.218 \pm 24.188	-
	Texture]				
	Homogénéité	0.687	0.466 \pm 0.032	0.493 \pm 0.027	-
	Énergie	0.687	0.017 \pm 0.005	0.023 \pm 0.005	-
	Corrélation	1	0.486 \pm 0.015	0.481 \pm 0.031	-
	Dissimilarité	0.687	3.14 \pm 0.537	2.668 \pm 0.396	-

TABLE 3.3 – Valeurs moyennes et écarts types des métriques pour le groupe contrôle au jour 1 et au jour 4. Les valeurs de p, résultantes des tests de Wilcoxon appariés indiquent la différence statistique entre J1 et J4.

Indices	p-value	Pré- traitement Moyenne ± ec.type	Post- Traitement Moyenne ± ec.type	q-value
Sous-régions / clusters	Complexité			
Fraction volumique #1 (%)	0.004	50.4 ± 1.4	68.8 ± 2.4	0.101
Fraction volumique #2 (%)	0.004	35.9 ± 0.8	22.4 ± 1.6	0.050
Fraction volumique #3 (%)	0.012	12.0 ± 0.7	7.3 ± 1.0	0.043
Fraction volumique #4 (%)	0.734	1.6 ± 0.3	1.5 ± 0.3	-
Distance moyenne #1 (mm)	0.300	6.253 ± 0.295	5.963 ± 0.354	-
Distance moyenne #2 (mm)	0.652	5.736 ± 0.306	5.837 ± 0.288	-
Distance moyenne #3 (mm)	0.734	5.964 ± 0.34	5.98 ± 0.362	-
Distance moyenne #4 (mm)	0.910	6.04 ± 0.393	5.96 ± 0.458	-
H-cluster #1	0.570	497 ± 139	492 ± 124	-
H-cluster #2	0.496	963 ± 174	1161 ± 317	-
H-cluster #3	0.004	1074 ± 145	631 ± 112	0.034
H-cluster #4	0.027	304 ± 46	196 ± 33	-
H-index	0.020	578 ± 90	403 ± 100	-
	Statistiques conventionnelles			
Médiane #2 (min ⁻¹)	0.496	0.02 ± 0.003	0.02 ± 0.002	-
Médiane #3 (min ⁻¹)	0.129	0.036 ± 0.006	0.047 ± 0.005	-
Médiane #4 (min ⁻¹)	0.359	0.082 ± 0.014	0.098 ± 0.011	-
Volume #1 (mm ³)	0.496	206.165 ± 27.255	197.749 ± 32.747	-
Volume #2 (mm ³)	0.004	145.081 ± 17.009	64.773 ± 10.799	0.025
Volume #3 (mm ³)	0.004	48.809 ± 6.681	19.874 ± 2.797	0.020
Volume #4 (mm ³)	0.055	6.312 ± 1.223	3.86 ± 0.506	-
Région Tumorale ROI	Statistiques conventionnelles			
Médiane (min ⁻¹)	0.008	0.014 ± 0.002	0.008 ± 0.001	0.033
Volume (mm ³)	0.055	503.656 ± 76.868	461.99 ± 73.877	-
	Texture			
Homogénéité	0.055	0.495 ± 0.028	0.54 ± 0.021	-
Énergie	0.027	0.025 ± 0.006	0.044 ± 0.008	-
Corrélation	0.055	0.43 ± 0.023	0.507 ± 0.019	-
Dissimilarité	0.910	2.709 ± 0.436	2.643 ± 0.336	-

TABLE 3.4 – Valeurs moyennes et écarts types des métriques pour le groupe traité au jour 1 et au jour 4. Les valeurs de p, résultantes des tests de Wilcoxon appariés indiquent la différence statistique entre J1 et J4. Les valeurs de q, résultantes de la correction BH sont reportées pour les indices qui conservaient leur significativité. Ces indices significatifs après correction sont surlignés en vert.

Indices	p-value	Gp Contrôle Moyenne \pm ec.type	Groupe Traité Moyenne \pm ec.type	q-value
Sous-régions / clusters	Complexité			
Fraction volumique #1 (%)	0.005	0.016 \pm 0.04	0.183 \pm 0.028	0.021
Fraction volumique #2 (%)	0.0008	-0.015 \pm 0.015	-0.135 \pm 0.017	0.010
Fraction volumique #3 (%)	0.088	-0.001 \pm 0.022	-0.047 \pm 0.012	-
Fraction volumique #4 (%)	0.864	0 \pm 0.006	-0.001 \pm 0.003	-
Distance moyenne #1 (mm)	0.018	0.487 \pm 0.198	-0.29 \pm 0.194	-
Distance moyenne #2 (mm)	0.224	0.102 \pm 0.146	5.736 \pm 0.306	-
Distance moyenne #3 (mm)	0.049	0.412 \pm 0.177	0.102 \pm 0.146	-
Distance moyenne #4 (mm)	0.036	0.653 \pm 0.203	-0.08 \pm 0.321	-
H-cluster #1	0.181	154 \pm 62	-5 \pm 78	-
H-cluster #2	0.388	321 \pm 118	198 \pm 235	-
H-cluster #3	0.003	155 \pm 129	-444 \pm 76	0.024
H-cluster #4	0.066	72 \pm 74	-108 \pm 34	-
H-index	0.003	146 \pm 55	-175 \pm 59	0.018
	Statistiques conventionnelles			
Médiane #2 (min^{-1})	0.272	-0.004 \pm 0.004	0 \pm 0.002	-
Médiane #3 (min^{-1})	0.145	-0.006 \pm 0.009	0.01 \pm 0.005	-
Médiane #4 (min^{-1})	0.181	-0.015 \pm 0.018	0.017 \pm 0.012	-
Volume #1 (mm^3)	0.145	46.534 \pm 28.648	-8.416 \pm 13.455	-
Volume #2 (mm^3)	0.0004	19.505 \pm 10.857	-80.308 \pm 14.609	0.010
Volume #3 (mm^3)	0.012	8.087 \pm 9.468	-28.935 \pm 6.256	0.045
Volume #4 (mm^3)	0.328	1.11 \pm 2.768	-2.453 \pm 1.18	-
Région Tumorale ROI	Statistiques conventionnelles			
Médiane (min^{-1})	0.328	-0.003 \pm 0.003	-0.006 \pm 0.002	-
Volume (mm^3)	0.003	110.722 \pm 33.866	-41.666 \pm 18.244	0.014
	Texture			
Homogénéité	0.689	0.026 \pm 0.04	0.046 \pm 0.025	-
Énergie	0.272	0.005 \pm 0.008	0.02 \pm 0.008	-
Corrélation	0.049	-0.006 \pm 0.032	0.077 \pm 0.033	-
Dissimilarité	0.689	-0.473 \pm 0.659	-0.066 \pm 0.395	-

TABLE 3.5 – Deltas des valeurs moyennes et leurs écarts types des métriques pour chaque groupe entre J4 et J1. Les valeurs de p, résultantes des tests de Wilcoxon non-appariés indiquent la différence statistique entre les deux groupes. Les valeurs de q, résultantes de la correction BH sont reportées pour les indices qui conservaient leur significativité. Ces indices significatifs après correction sont surlignés en vert.

Sous-régions / clusters	Indices	p-value en fonction de la connectivité		
		6	18	26
	Complexité			
	Distance moyenne #1 (mm)	-	0.002	-
	Distance moyenne #2 (mm)	-	-	-
	Distance moyenne #3 (mm)	-	-	-
	Distance moyenne #4 (mm)	-	-	-
	H-cluster #1	-	-	-
	H-cluster #2	-	-	-
	H-cluster #3	0.003	0.008	-
	H-cluster #4	-	-	-
	H-index	0.003	0.003	0.008

TABLE 3.6 – Influence du paramètre connectivité employé pour l'analyse en composantes connectées sur les métriques qui en dépendent. La valeur de p issue du test de Wilcoxon est indiquée uniquement pour les indices qui étaient significatifs après correction de BH, "-" est indiqué dans le cas inverse.

L'ensemble des indices permettant de mettre en évidence une différence significative dans l'évolution entre les deux groupes sont représentés sur la figure 11 pour les indices de la complexité de l'organisation spatiale et sur la figure 12 pour les indices conventionnels. Pour chacun de ces indices et pour chaque groupe est indiquée la variation entre sa valeur au jour 1 et au jour 4 ainsi que la différence d'évolution entre les groupes. On notera que l'indice H_{index} du groupe contrôle augmente (Delta $H_{index} > 0$) à l'image des hétérogénéités grandissantes appréciées visuellement en figure 8. A l'inverse pour le groupe traité, la valeur de l'indice diminue (Delta $H_{index} < 0$) comme observée sur la figure 8 avec une tendance à l'homogénéisation des valeurs de K^{trans} .

Enfin, l'influence de la connectivité sur nos deltas-indices a été évalué en comparant les résultats significatifs au regard de sa valeur 6, 18 ou 26. Peu de différences émanaient de la comparaison entre les résultats pour 6 et 18 avec une variation de la valeur de p du $H_{cluster}$ #3 respectivement de 0,003 à 0,008. Pour la connectivité 26, le $H_{cluster}$ #3 n'était plus significatif. Ces résultats sont reportés dans la table 3.6 pour la distance moyenne, les $H_{cluster}$ et le H_{index} . La fraction volumique n'est pas impactée par la modification de la connectivité. La figure 13 illustre par exemple l'influence de la connectivité sur le nombres de patches du cluster #3 et on observe par exemple dans le groupe contrôle, une diminution du nombre de patches à J1 et à J4 avec l'augmentation de la connectivité ce qui implique donc une diminution du H_{index} .

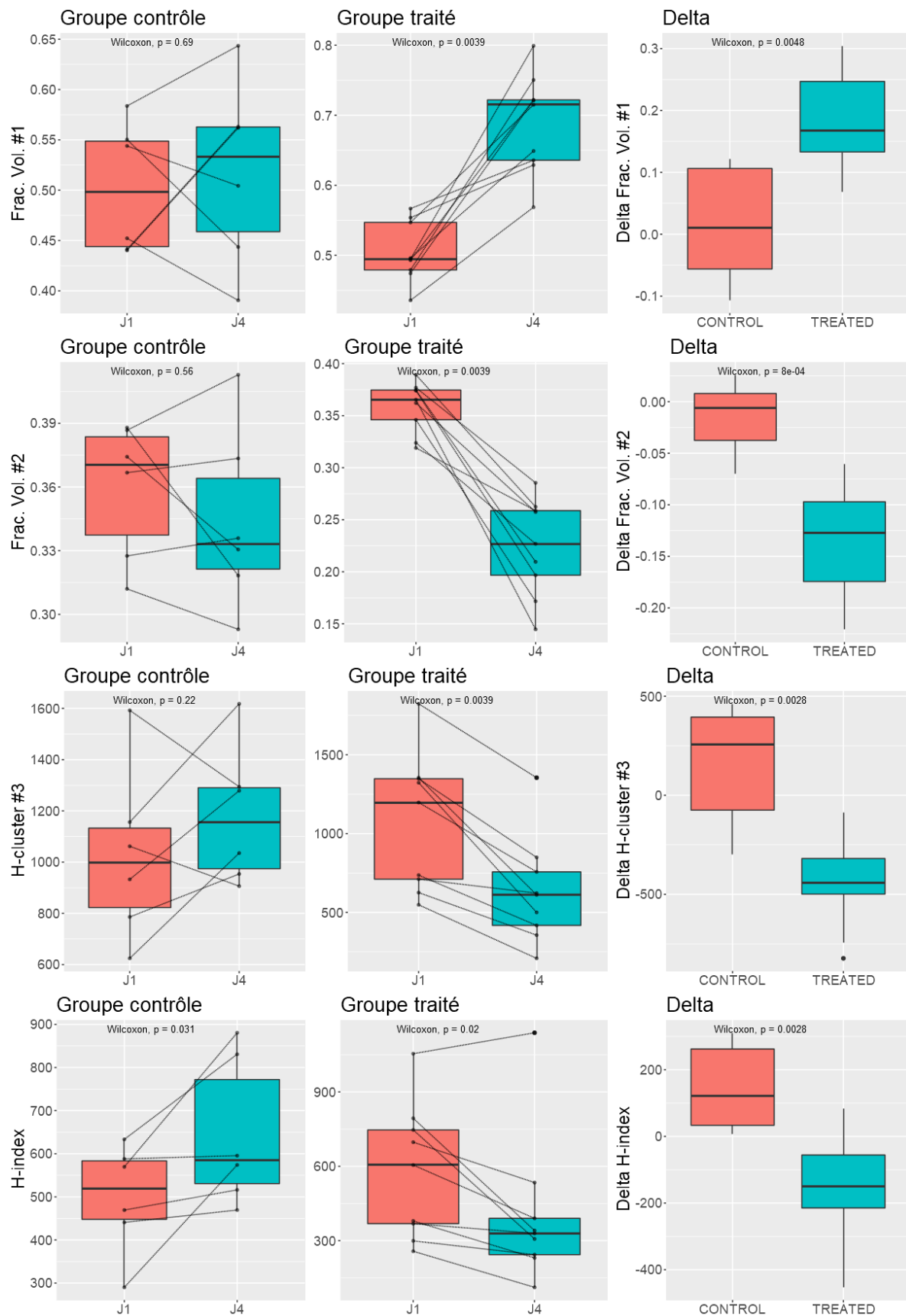


FIGURE 11 – Boxplots illustrant les différences entre J1 et J4 pour les métriques, le H_{index} et le $H_{cluster}$ #3 pour les groupes contrôlés (colonne 1), traités (colonne 2) et pour les valeurs deltas entre les groupes (colonne 3). Ces métriques sont celles qui présentaient une différence significative après correction de BH dans la comparaison entre les deux groupes.

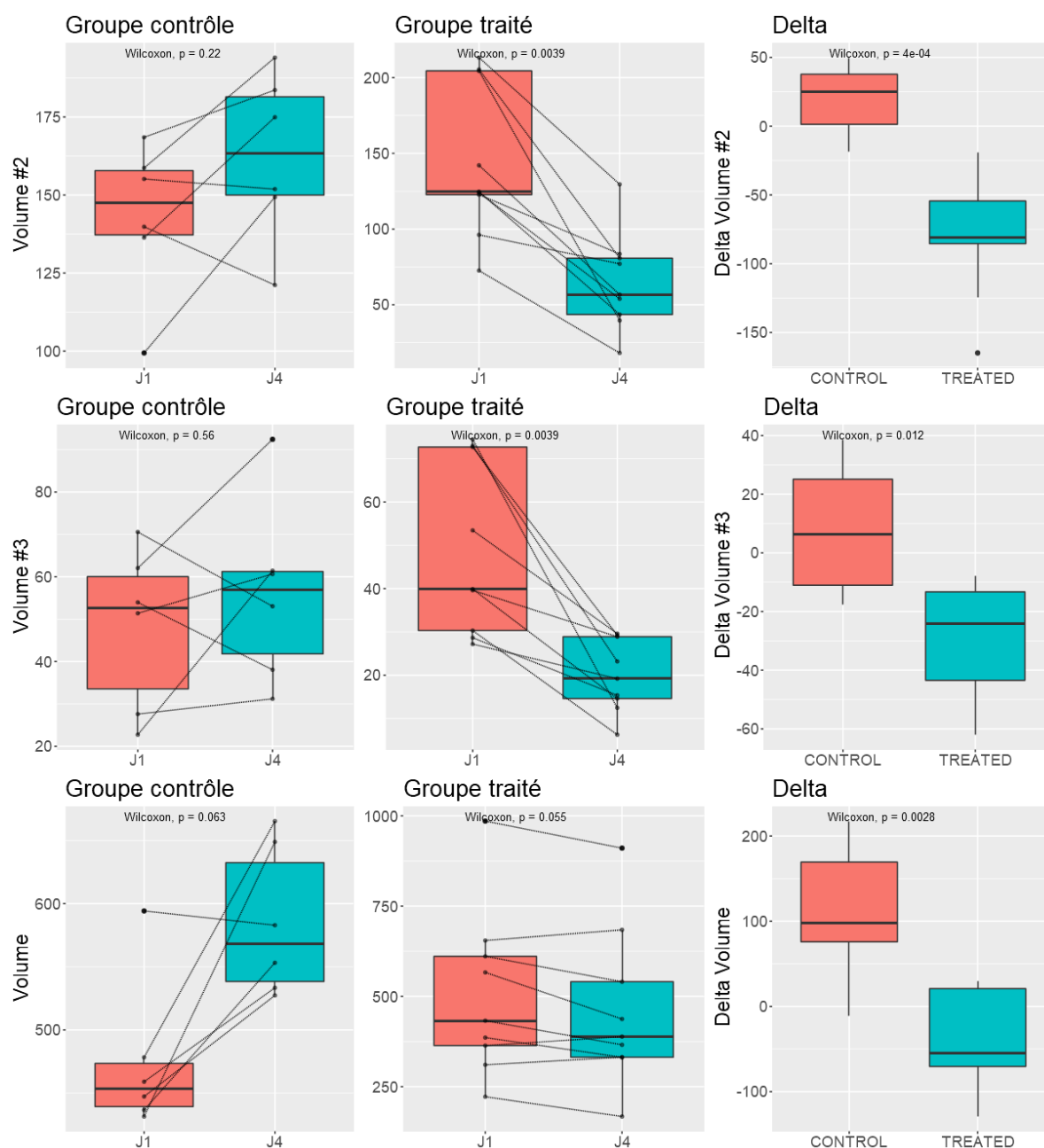


FIGURE 12 – Boxplots illustrant les différences entre J1 et J4 pour les volumes #2, #3 et la ROI pour les groupes contrôles (colonne 1), traités (colonne 2) et pour les valeurs deltas entre les groupes (colonne 3). Ces métriques sont celles qui présentaient une différence significative après correction de BH dans la comparaison entre les deux groupes.

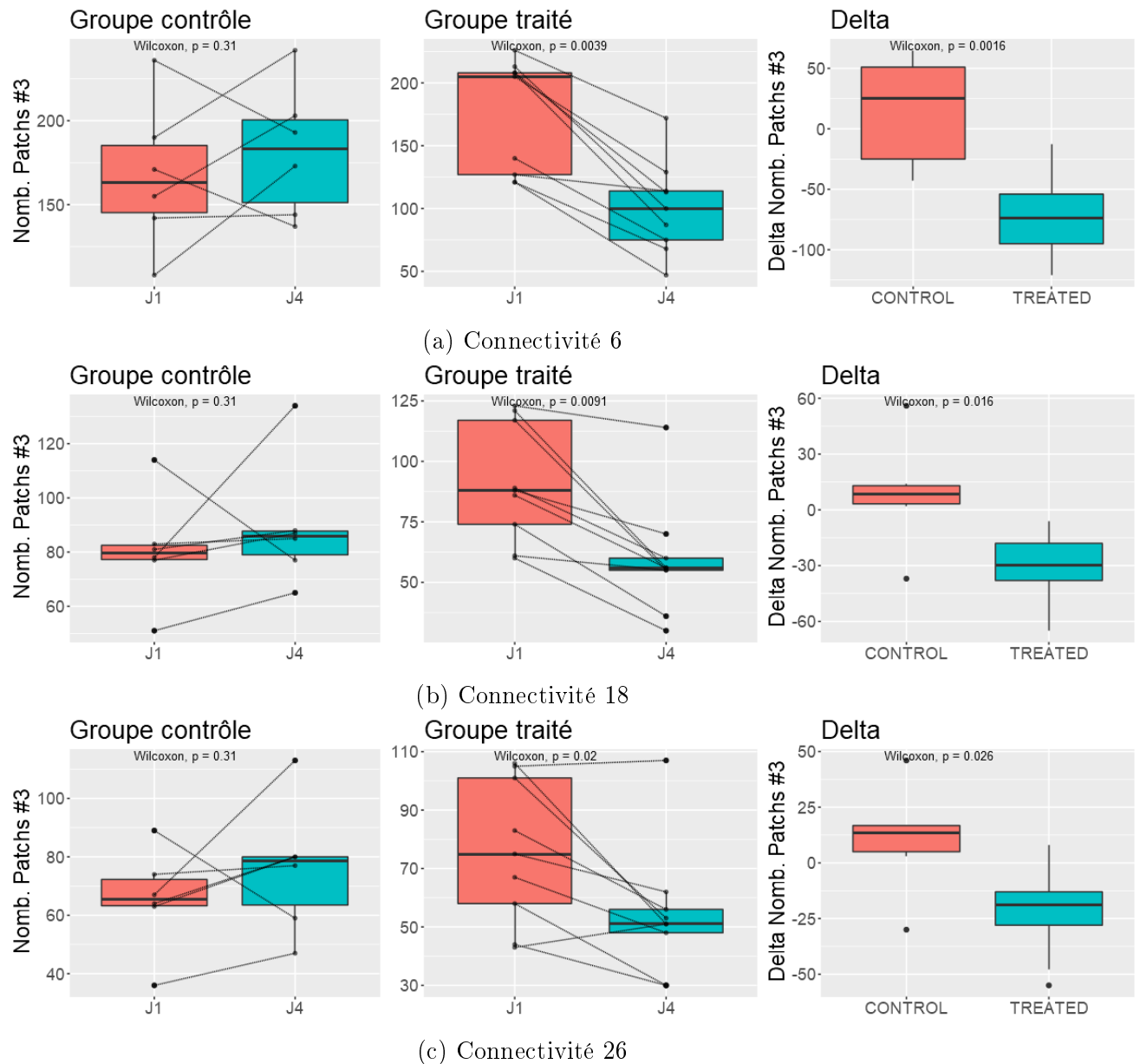


FIGURE 13 – Influence de la connectivité sur le nombre de patches du cluster #3 résultant de l'analyse en composantes connectées.

2.4 Discussion

En oncologie, l'intérêt croissant pour les thérapies ciblées et en particulier pour les molécules anti-angiogéniques ont conduit à un nombre croissant d'études s'intéressant à la micro-vascularisation et à l'hétérogénéité vasculaire tumorale en raison de la capacité de ces thérapies à normaliser et à réorganiser la structure micro-vasculaire (JAIN, 2005). Certaines études ont tenté d'exploiter l'hétérogénéité des images radiologiques des tumeurs via des analyses de texture pour évaluer la réponse thérapeutique (FAVE et al., 2017 ; LIU et al., 2016 ; PANZERI et al., 2018), mais aucune de ces approches n'a à ce jour réussi à atteindre un usage clinique, probablement à cause d'un manque de

standardisation et de guidelines (LAMBIN et al., 2017), des méthodes, mais également suite à un défaut d'interprétabilité des résultats, critère nécessaire à la validation d'un biomarqueur.

Dans cette étude préclinique rétrospective, nous décrivons une première application de notre méthode de quantification de l'hétérogénéité par la description de la complexité de l'organisation spatiale de sous-régions intra-tumorales. Cette étude a déjà fait l'objet d'une publication qui néanmoins ne s'intéressait pas au H_{index} mais uniquement aux métriques qui le composent (GIROT et al., 2021). Cette application consistait en un suivi longitudinal de l'hétérogénéité micro-vasculaire en imagerie DCE-IRM sur un modèle murin expérimental. La méthode associe un processus de clustering et l'extraction d'indices permettant d'identifier et de quantifier différents niveaux de complexité d'une carte paramétrique. Cette approche est originale au regard des autres études qui majoritairement évaluent soit l'hétérogénéité spatiale moyennée sur la globalité de la tumeur (ROSE et al., 2009), soit l'hétérogénéité des intensités en analysant uniquement les histogrammes des sous-régions tumorales et qui ne considèrent donc pas leur distribution spatiale (LONGO et al., 2015). A notre connaissance, seules quelques études se sont intéressées au processus de clustering pour la quantification de l'hétérogénéité intra-tumorale. LONGO et al. (2015) ont eu recours à un processus de clustering pour extraire de sous régions tumorales des indices statistiques conventionnels pour évaluer la réponse thérapeutique. ANDERSEN et al. (2011) ont eu recours à la fraction volumique des clusters formés par les cartes K^{trans} et v_e pour évaluer la réponse thérapeutique. TORHEIM et al. (2016) ont analysé le nombre et la tailles des patchs formés par des clusters multiparamétriques à partir des paramètres K^{trans} et v_e mais également à partir de paramètres issus d'autres modèles pharmacocinétiques afin d'identifier les régions tumorales liées à la rechute dans le cas des cancers cervicaux. COQUERY et al. (2018) ont comparé l'apport du clustering en IRM multiparamétrique via la taille des clusters à l'analyse conventionnelle des ROI pour détecter les variations d'hétérogénéité induite par différentes combinaisons de traitements.

A partir des résultats obtenus dans cette étude préclinique, la méthode proposée apparaît plus sensible que les analyses de texture pour permettre un suivi de l'efficacité thérapeutique sous nos conditions expérimentales. En effet, les analyses statistiques démontrent qu'aucun indice de la GLCM ne permettait de différencier significativement les groupes traité et contrôle entre eux, ou bien d'établir une différence entre avant et après traitement au sein de chaque groupe. A l'inverse, les autres catégories d'indices ont révélé des différences significatives. Ainsi, 5 indices évalués au sein des clusters (4 complexités, 1 conventionnel) ont varié significativement au cours du traitement pour le groupe traité et 6 dont 4 concernés précédemment (4 complexités, 2 conventionnels) variaient significativement différemment pour chaque groupe au terme de quatre jours de suivi.

Le constat fait ici sur les indices de texture ne prend pas en considération deux indices significatif au test de Wilcoxon mais qui étaient éliminés après la correction de BH, à savoir l'énergie et la corrélation. Cette correction est appliquée pour ne pas retenir les faux positifs dans un contexte de comparaisons statistiques multiples. Ce type de correction est évidemment essentiel en particulier lorsque le nombre d'indices testés est supérieur à la taille de l'échantillon étudié. Toutefois, le recours à ce type de correction est un sujet de débat puisque certains auteurs conseillent de ne pas y avoir recours lorsqu'une étude est exploratoire et faite de comparaisons non-prévues entre différentes catégories d'indices dont le potentiel pourrait être considérée dans de futures études (ARMSTRONG, 2014). Ainsi, on peut relever des études conséquentes sur 100 patients explorant au minimum 128 indices de texture (AHMED et al., 2013), et les acteurs précisent qu'aucune correction n'était appliquée puisqu'il s'agissait d'une étude exploratoire cherchant à émettre des hypothèses. Ainsi, et faute de recul sur ce type d'analyse, il apparait prudent d'envisager les prochaines études avec et sans les corrections de tests multiples et d'y associer tous les indices, ceux issus de l'analyse de texture compris.

En ce qui concerne nos métriques et indices de l'hétérogénéité, nous nous sommes également intéressés à l'influence de la connectivité comme discuté dans le chapitre 2. Il a notamment été vérifié qu'une variation de la connectivité de 6 comme dans cette étude, vers 18 puis 26, produisait un jeu d'indices légèrement différent. En effet, le choix d'une connectivité trop élevée peut engendrer un regroupement important des patchs. Nous avons ainsi observé par exemple qu'avec une connectivité de 6, les 200 patchs observés dans le cluster #3 chez une souris contrôle à J4, étaient réduits au nombre de 74 avec l'emploi d'une connectivité 26. Ce regroupement grossier des patchs ne permettait plus de décrire l'hétérogénéité spatiale des images, entraînant ainsi une perte de significativité de l'indice associé. Par conséquence, nous recommandons d'avoir recours à la connectivité comprise entre 6 et 18 avec ce jeu de données. La standardisation de cet unique hyperparamètre de notre étude est donc essentiel et nous recommandons de la fixer pour l'ensemble des données d'une étude, plus particulièrement dans l'éventualité d'une étude multicentrique. En revanche, les indices de texture dépendent de multiples paramètres qui doivent être défini par l'utilisateur : le nombre de niveau de gris à employer pour la quantification, l'intervalle des niveaux de gris à retenir et différents processus de pré-traitement des images qui peuvent également s'additionner.

Outre ces paramètres, les indices de texture sont également dépendant de sources de variabilité externes tels que les réglages et les performances des systèmes d'imagerie ou l'utilisation de définitions mathématiques variantes entre les différents logiciels couramment employés. Dans le cas des performances machines, des efforts sont fourni par la communauté, notamment en scanner, pour permettre de standardiser les protocoles de collecte de ces indices et de les rendre exploitable sur différents sites (ZWANENBURG

et al., 2020) mais de nombreux efforts doivent encore être faits particulièrement en ce qui concerne l'IRM. En comparaison et comme souligné dans le chapitre 2, notre méthode permettrait de s'affranchir des variations de contrastes inter-machines grâce au processus de clustering. Par extension aux cartes paramétriques, comme celles de K^{trans} , avec la variabilité établie de l'AIF, notre méthode représente un avantage indéniable dans la perspective de protocole d'études multicentriques validant un biomarqueur uniquement si sa mesure est indépendante du système d'imagerie.

Au regard des résultats obtenus, le nouveau jeu d'indices caractérisant l'hétérogénéité des cartes paramétriques constituait une première preuve de concept pour le développement d'un biomarqueur d'imagerie de l'efficacité thérapeutique d'un traitement anti-angiogénique. Néanmoins, pour considérer cette étude comme une première étape de validation de potentiels biomarqueurs d'imagerie, ces indices devraient d'après les recommandations internationales être confrontés à la "biologie sous-jacente" d'après les recommandations internationales. Or les analyses histopathologiques permettant de confirmer la variation de l'hétérogénéité vasculaire n'ont pas été prévues dans l'étude initiale et c'est pour cela que ce chapitre aborde l'exploration préclinique du biomarqueur et non sa validation. De plus, même si ces analyses avaient été conduites, au regard de la qualité variable des échantillons, qui ne peuvent pas représenter la tumeur dans sa globalité, et qui ne peuvent pas être comparés avant et après traitement, établir une corrélation entre la physiopathologie et un biomarqueur d'imagerie est parfois impossible (WATERTON & PYLKKANEN, 2012). Toutefois, cela n'empêche pas trois autres résultats qui soulignent le lien avec la "biologie sous-jacente".

Premièrement, un déplacement de la distribution des histogrammes vers des valeurs d'intensités plus faible dans le groupe traité a été reporté à la fois dans ce travail et dans l'étude originale (EL TANNIR EL TAYARA et al., 2011) pour la valeur médiane du K^{trans} mesurée sur la ROI tumorale complète. Ce résultat valide l'évolution des fractions volumiques des clusters intermédiaires #2 et #3 qui après l'administration de la thérapie anti-angiogénique diminuent significativement et s'accompagne d'une augmentation concomitante de la fraction volumique du cluster #1 de valeur plus faible de K^{trans} . Cette observation est donc en cohérence avec une diminution de la perfusion et de la perméabilité capillaire après traitement. En associant la fraction volumique et la distance moyenne, au nombre de patches, le $H_{cluster}$ nous renseigne sur l'hétérogénéité de la vascularisation. Ainsi la diminution du H_{index} au cours du temps traduit une homogénéisation des cartes K^{trans} pour le groupe traité avec un ΔH_{index} négatif. A l'inverse du groupe contrôle, bien qu'aucun indice ne soit significatif statistiquement, le ΔH_{index} est dans ce cas positif ce qui indique quantitativement une augmentation de l'hétérogénéité vasculaire, phénomène physiologique attendu lors de la croissance tumorale associée avec un processus de néo-angiogenèse. Néanmoins, même si le statut thérapeutique des souris avait été confirmé préalablement par les analyses de volumes,

ces observations devront par la suite être associées à des analyses anatomopathologiques qui soient corrélables à une évaluation de la variation de l'hétérogénéité en imagerie.

Deuxièmement, bien que le volume tumoral tende à décroître pour le groupe traité lors de l'examen au jour 4, cette variation n'était pas significative dans notre étude. Le volume tumoral est généralement utilisé comme biomarqueur de référence en oncologie. Néanmoins, en pratique clinique, au cours d'une thérapie anti-angiogénique, il a été montré que les évaluations morphologiques des tumeurs basées sur le critère RECIST sous-estimaient largement la réponse tumorale à ces stratégies thérapeutiques (SMITH et al., 2010). Par exemple, aucune modification volumique tumorale n'a été observée au cours du traitement pour respectivement 50 et 75% des patients atteints de carcinomes hépatocellulaires (MOTZER et al., 2018) ou de tumeurs neuroendocriniennes gastroentéropancréatique (LUO et al., 2017). Cette absence de réponse des critères RECIST s'explique principalement par le fait que les modifications vasculaires apparaissent généralement avant toute modification volumique de la tumeur. Nos résultats précliniques sont ainsi cohérents avec ces résultats cliniques.

Finalement, aucun des indices n'a permis de démontrer une différence significative dans le cluster #4 de plus forte intensité, généralement considéré comme le "point chaud" (angl. hot spot) des images. Ce résultat, inattendu en comparaison aux précédentes études qui portaient toutes leur attention sur ces points chauds, est contre-intuitif puisque ces zones sont généralement considérées comme celles offrant un maximum d'informations. Par exemple, un intervalle de variation plus important des valeurs de K^{trans} a été identifié par HAYES et al. (2002) dans ces "points chauds" par rapport à la ROI tumorale. La différence de résultat avec notre étude peut s'expliquer par la définition de ces "points chauds". Alors qu'ils sont définis comme l'ensemble des pixels au-delà du 75ème percentile de l'histogramme tumoral par Hayes et al, notre méthode de clustering non-supervisée résulte en des clusters avec finalement très peu de voxels et les fractions volumiques impliquées représentant 2% du volume tumoral.

Cette étude comportait certaines limitations, partiellement mentionnées ci-dessus. Premièrement, les données sont originaires d'une précédente étude et ont été analysées rétrospectivement. Néanmoins, ces données 3D étaient acquises selon une méthodologie robuste avec le même protocole, la même machine, et les mêmes paramètres IRM. L'hétérogénéité des jeux de données constitue souvent une limitation des études cliniques rétrospectives, et de fait, ces données précliniques analysées ont été considérées toutes comme utiles et adaptées à une preuve de concept de la méthodologie développée au cours de cette thèse. Par ailleurs, bien que ces résultats soient significatifs, même après une correction de BH, ils doivent être considérés comme préliminaires en raison de la faible taille du jeu de données qui n'était pas initialement conçu pour des considérations d'hétérogénéité tumorale. Deuxièmement, en l'absence de données histologiques, l'interprétation de nos résultats en termes de propriétés physiologiques

micro-vasculaires semble cohérente aux attentes, d'autant plus que la correspondance du K^{trans} avec la biologie est généralement admise.

3 Conclusions et perspectives

Nous avons présenté une première application de la méthode développée au cours de ce travail de thèse pour analyser l'hétérogénéité des cartes paramétriques. Nous avons montré au travers de cette application à des cartes K^{trans} dérivées de l'imagerie DCE-IRM de contraste que cette méthode permettait de fournir de potentiels biomarqueurs quantitatifs de l'hétérogénéité pour l'évaluation de la réponse thérapeutique d'un traitement anti-angiogénique. De plus, il apparaît que notre méthode présente des avantages en comparaison aux analyses de texture avec une interprétabilité plus aisée des résultats au regard de la physiologie tumorale explorée et une plus faible variabilité lorsqu'il s'agira d'aborder des approches multicentriques. L'enjeu du chapitre suivant sera d'explorer des applications cliniques potentiels des biomarqueurs identifiés dans ce contexte préclinique afin d'étendre notamment la validation technique de ces BI, de démontrer leur capacité à simplifier la standardisation des procédures d'analyses et enfin de tester leur potentiel dans un contexte clinique différent du suivi de l'efficacité anti-angiogénique.

Bibliographie

- AHMED, A., GIBBS, P., PICKLES, M. & TURNBULL, L. (2013). Texture analysis in assessment and prediction of chemotherapy response in breast cancer : texture analysis in breast cancer MRI. *Journal of Magnetic Resonance Imaging*, 38(1), 89-101. <https://doi.org/10.1002/jmri.23971>
- ANDERSEN, E. K. F., KRISTENSEN, G. B., LYNG, H. & MALINEN, E. (2011). Pharmacokinetic analysis and k-means clustering of DCEMR images for radiotherapy outcome prediction of advanced cervical cancers. *Acta Oncologica*, 50(6), 859-865. <https://doi.org/10.3109/0284186X.2011.578586>
- ARMSTRONG, R. A. (2014). When to use the bonferroni correction. *Ophthalmic and Physiological Optics*, 34(5), 502-508. <https://doi.org/10.1111/opo.12131>
- AZZI, S., HEBDA, J. K. & GAVARD, J. (2013). Vascular Permeability and Drug Delivery in Cancers. *Frontiers in Oncology*, 3. <https://doi.org/10.3389/fonc.2013.00211>
- CARMELIET, P. & JAIN, R. K. (2000). Angiogenesis in cancer and other diseases. *Nature*, 407(6801), 249-257. <https://doi.org/10.1038/35025220>
- CARMELIET, P. & JAIN, R. K. (2011a). Molecular mechanisms and clinical applications of angiogenesis. *Nature*, 473(7347), 298-307. <https://doi.org/10.1038/nature10144>
- CARMELIET, P. & JAIN, R. K. (2011b). Principles and mechanisms of vessel normalization for cancer and other angiogenic diseases. *Nature Reviews Drug Discovery*, 10(6), 417-427. <https://doi.org/10.1038/nrd3455>
- COQUERY, N., SERDUC, R., RÉMY, C., BARBIER, E. L. & LEMASSON, B. (2018). Cluster versus ROI analysis to assess combined antiangiogenic therapy and radiotherapy in the f98 rat-glioma model. *NMR in Biomedicine*, 31(8), e3933. <https://doi.org/10.1002/nbm.3933>
- DIETRICH, C. F., NOLSØE, C. P., BARR, R. G., BERZIGOTTI, A., BURNS, P. N., CANTISANI, V., CHAMMAS, M. C., CHAUBAL, N., CHOI, B. I., CLEVERT, D.-A., CUI, X., DONG, Y., D'ONOFRIO, M., FOWLKES, J. B., GILJA, O. H., HUANG, P., IGNEE, A., JENSSEN, C., KONO, Y., ... ZHENG, R. (2020). Guidelines and Good Clinical Practice Recommendations for Contrast Enhanced Ultrasound (CEUS) in the Liver – Update 2020 – WFUMB in Cooperation with EFSUMB,

- AFSUMB, AIUM, and FLAUS. *Ultraschall in der Medizin - European Journal of Ultrasound*, 41(5), 562-585. <https://doi.org/10.1055/a-1177-0530>
- EL TANNIR EL TAYARA, N., BEN ACHOUR, N., WALCZAK, C., NEMATI, F., MISPELTER, J., DECAUDIN, D., VAUTIER, J. & VOLK, A. (2011). Anti-angiogenic therapy follow-up in a mouse tumor model by a novel 3D radial multi-gradient echo DCE MRI technique with individual AIF measurement. *Proceedings of the International Society for Magnetic Resonance in Medicine*, 19, 1056.
- FAVE, X., ZHANG, L., YANG, J., MACKIN, D., BALTER, P., GOMEZ, D., FOLLOWILL, D., JONES, A. K., STINGO, F., LIAO, Z., MOHAN, R. & COURT, L. (2017). Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Scientific Reports*, 7(1), 588. <https://doi.org/10.1038/s41598-017-00665-z>
- GIROT, C., VOLK, A., WALCZAK, C., LASSAU, N. & PITRE-CHAMPAGNAT, S. (2021). New method for quantification of intratumoral heterogeneity : a feasibility study on ktrans maps from preclinical DCE-MRI. *Magnetic Resonance Materials in Physics, Biology and Medicine*. <https://doi.org/10.1007/s10334-021-00930-3>
- GRAND-PERRET, V. (2019). *Echographie de contraste 3D pour l'évaluation de la microvascularisation tumorale* (thèse de doct.) [2019SACLS575]. <http://www.theses.fr/2019SACLS575/document>
- HAYES, C., PADHANI, A. R. & LEACH, M. O. (2002). Assessing changes in tumour vascular function using dynamic contrast-enhanced magnetic resonance imaging. *NMR in Biomedicine*, 15(2), 154-163. <https://doi.org/10.1002/nbm.756>
- JACKSON, J., MEYER, C., NISHIMURA, D. & MACOVSKI, A. (1991). Selection of a convolution function for Fourier inversion using gridding (computerised tomography application). *IEEE Transactions on Medical Imaging*, 10(3), 473-478. <https://doi.org/10.1109/42.97598>
- JAIN, R. K. (2005). Normalization of tumor vasculature : an emerging concept in antiangiogenic therapy. *Science*, 307(5706), 58-62. <https://doi.org/10.1126/science.1104819>
- LAMBIN, P., LEIJENAAR, R. T., DEIST, T. M., PEERLINGS, J., de JONG, E. E., van TIMMEREN, J., SANDULEANU, S., LARUE, R. T., EVEN, A. J., JOCHEMS, A., van WIJK, Y., WOODRUFF, H., van SOEST, J., LUSTBERG, T., ROELOFS, E., van ELMPT, W., DEKKER, A., MOTTAGHY, F. M., WILDBERGER, J. E. & WALSH, S. (2017). Radiomics : the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*, 14(12), 749-762. <https://doi.org/10.1038/nrclinonc.2017.141>
- LASSAU, N., BONASTRE, J., KIND, M., VILGRAIN, V., LACROIX, J., CUINET, M., TAIEB, S., AZIZA, R., SARRAN, A., LABBE-DEVILLIERS, C., GALLIX, B., LUCIDARME, O., PTAK, Y., ROCHER, L., CAQUOT, L.-M., CHAGNON,

- S., MARION, D., LUCIANI, A., FEUTRAY, S., ... KOSCIELNY, S. (2014). Validation of Dynamic Contrast-Enhanced Ultrasound in Predicting Outcomes of Antiangiogenic Therapy for Solid Tumors : The French Multicenter Support for Innovative and Expensive Techniques Study. *Investigative Radiology*, 49(12), 794-800. <https://doi.org/10.1097/RLI.000000000000085>
- LESS, J. R., SKALAK, T. C., SEVICK, E. M. & JAIN, R. K. (1991). Microvascular architecture in a mammary carcinoma : branching patterns and vessel dimensions. *Cancer Research*, 51(1), 265-273.
- LIU, J., MAO, Y., LI, Z., ZHANG, D., ZHANG, Z., HAO, S. & LI, B. (2016). Use of texture analysis based on contrast-enhanced MRI to predict treatment response to chemoradiotherapy in nasopharyngeal carcinoma : MRI in treatment response prediction. *Journal of Magnetic Resonance Imaging*, 44(2), 445-455. <https://doi.org/10.1002/jmri.25156>
- LONGO, D. L., DASTRÙ, W., CONSOLINO, L., ESPAK, M., ARIGONI, M., CAVALLO, F. & AIME, S. (2015). Cluster analysis of quantitative parametric maps from DCE-MRI : application in evaluating heterogeneity of tumor response to antiangiogenic treatment. *Magnetic Resonance Imaging*, 33(6), 725-736. <https://doi.org/10.1016/j.mri.2015.03.005>
- LUO, Y., CHEN, J., HUANG, K., LIN, Y., CHEN, M., XU, L., LI, Z.-P. & FENG, S.-T. (2017). Early evaluation of sunitinib for the treatment of advanced gastroenteropancreatic neuroendocrine neoplasms via CT imaging : RECIST 1.1 or choi criteria? *BMC Cancer*, 17(1), 154. <https://doi.org/10.1186/s12885-017-3150-7>
- MISPELTER, J., LUPU, M. & BRIGUET, A. (2015). *NMR probeheads for biophysical and biomedical experiments : theoretical principles & practical guidelines* (2. ed). Imperial College Press.
- MOTZER, R. J., TANNIR, N. M., MCDERMOTT, D. F., ARÉN FRONTERA, O., MELICHAR, B., CHOUEIRI, T. K., PLIMACK, E. R., BARTHÉLÉMY, P., PORTA, C., GEORGE, S., POWLES, T., DONSKOV, F., NEIMAN, V., KOLLMANNSSBERGER, C. K., SALMAN, P., GURNEY, H., HAWKINS, R., RAVAUD, A., GRIMM, M.-O., ... ESCUDIER, B. (2018). Nivolumab plus ipilimumab versus sunitinib in advanced renal-cell carcinoma. *New England Journal of Medicine*, 378(14), 1277-1290. <https://doi.org/10.1056/NEJMoa1712126>
- PANZERI, M. M., LOSIO, C., DELLA CORTE, A., VENTURINI, E., AMBROSI, A., PANIZZA, P. & DE COBELLI, F. (2018). Prediction of chemoresistance in women undergoing neo-adjuvant chemotherapy for locally advanced breast cancer : volumetric analysis of first-order textural features extracted from multiparametric MRI. *Contrast Media & Molecular Imaging*, 2018, 1-7. <https://doi.org/10.1155/2018/8329041>

- ROSE, C. J., MILLS, S. J., O'CONNOR, J. P. B., BUONACCORSI, G. A., ROBERTS, C., WATSON, Y., CHEUNG, S., ZHAO, S., WHITCHER, B., JACKSON, A. & PARKER, G. J. M. (2009). Quantifying spatial heterogeneity in dynamic contrast-enhanced MRI parameter maps. *Magnetic Resonance in Medicine*, 62(2), 488-499. <https://doi.org/10.1002/mrm.22003>
- SALEM, A. & O'CONNOR, J. P. (2016). Assessment of tumor angiogenesis. *Magnetic Resonance Imaging Clinics of North America*, 24(1), 45-56. <https://doi.org/10.1016/j.mric.2015.08.010>
- SCHIEDA, N., van der POL, C. B., WALKER, D., TSAMPALIEROS, A. K., MARALANI, P. J., WOO, S. & DAVENPORT, M. S. (2020). Adverse events to the gadolinium-based contrast agent gadoteric acid : systematic review and meta-analysis. *Radiology*, 297(3), 565-572. <https://doi.org/10.1148/radiol.2020200073>
- SHUKLA-DAVE, A., OBUCHOWSKI, N. A., CHENEVERT, T. L., JAMBAWALIKAR, S., SCHWARTZ, L. H., MALYARENKO, D., HUANG, W., NOWOROLSKI, S. M., YOUNG, R. J., SHIROISHI, M. S., KIM, H., COOLENS, C., LAUE, H., CHUNG, C., ROSEN, M., BOSS, M. & JACKSON, E. F. (2019). Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *Journal of Magnetic Resonance Imaging*, 49(7), e101-e121. <https://doi.org/10.1002/jmri.26518>
- SMITH, A. D., LIEBER, M. L. & SHAH, S. N. (2010). Assessing tumor response and detecting recurrence in metastatic renal cell carcinoma on targeted therapy : importance of size and attenuation on contrast-enhanced CT. *American Journal of Roentgenology*, 194(1), 157-165. <https://doi.org/10.2214/AJR.09.2941>
- TOFTS, P. S. & KERMODE, A. G. (1991). Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. fundamental concepts. *Magnetic Resonance in Medicine*, 17(2), 357-367. <https://doi.org/10.1002/mrm.1910170208>
- TORHEIM, T., GROENDAHL, A. R., ANDERSEN, E. K. F., LYNG, H., MALINEN, E., KVAAL, K. & FUTSAETHER, C. M. (2016). Cluster analysis of dynamic contrast enhanced MRI reveals tumor subregions related to locoregional relapse for cervical cancer patients. *Acta Oncologica*, 55(11), 1294-1298. <https://doi.org/10.1080/0284186X.2016.1189091>
- TREPS, L. & GAVARD, J. (2015). L'angiogenèse tumorale : Quand l'arbre de vie tourne mal. *médecine/sciences*, 31(11), 989-995. <https://doi.org/10.1051/medsci/20153111013>
- VAUTIER, J., EL TANNIR EL TAYARA, N., WALCZAK, C., MISPELTER, J. & VOLK, A. (2013). Radial multigradient-echo DCE-MRI for 3d K^{trans} mapping with individual arterial input function measurement in mouse tumor models : 3d radial

- DCE-MRI in mice with individual AIF. *Magnetic Resonance in Medicine*, 70(3), 823-828. <https://doi.org/10.1002/mrm.24504>
- WATERTON, J. C. & PYLKKANEN, L. (2012). Qualification of imaging biomarkers for oncology drug development. *European Journal of Cancer*, 48(4), 409-415. <https://doi.org/10.1016/j.ejca.2011.11.037>
- WEIDNER, N., FOLKMAN, J., POZZA, F., BEVILACQUA, P., ALLRED, E. N., MOORE, D. H., MELI, S. & GASPARINI, G. (1992). Tumor angiogenesis : a new significant and independent prognostic indicator in early-stage breast carcinoma. *JNCI : Journal of the National Cancer Institute*, 84(24), 1875-1887. <https://doi.org/10.1093/jnci/84.24.1875>
- YANG, X. & KNOPP, M. V. (2011). Quantifying tumor vascular heterogeneity with dynamic contrast-enhanced magnetic resonance imaging : a review. *Journal of Biomedicine and Biotechnology*, 2011, 1-12. <https://doi.org/10.1155/2011/732848>
- ZWANENBURG, A., VALLIÈRES, M., ABDALAH, M. A., AERTS, H. J. W. L., ANDREARCZYK, V., APTE, A., ASHRAFINIA, S., BAKAS, S., BEUKINGA, R. J., BOELLAARD, R., BOGOWICZ, M., BOLDRINI, L., BUVAT, I., COOK, G. J. R., DAVATZIKOS, C., DEPEURSINGE, A., DESSEROIT, M.-C., DINAPOLI, N., DINH, C. V., ... LÖCK, S. (2020). The image biomarker standardization initiative : standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*, 295(2), 328-338. <https://doi.org/10.1148/radiol.2020191145>

Chapitre 4

Évaluation de l'intérêt clinique de biomarqueurs de l'hétérogénéité de l'image

Un biomarqueur d'imagerie est initialement développé pour répondre à un objectif bien défini qui peut évoluer au fil des études cliniques. Appliqué à la cancérologie, ce biomarqueur peut s'appliquer à un type de cancer particulier, ou bien à une thérapie spécifique. L'intérêt clinique d'un biomarqueur ne se mesure pas uniquement par sa capacité à identifier un événement clinique mais également par ses capacités à faire évoluer les pratiques. L'ambition initiale des biomarqueurs de l'hétérogénéité se situe dans l'évaluation de l'efficacité des thérapies ciblées, comme les molécules anti-angiogéniques abordées dans le chapitre précédent, mais cette ambition s'étend aujourd'hui surtout vers les nouvelles immunothérapies pour contribuer à une médecine de précision. Dans le cadre de cette thèse, nous avons choisi d'explorer quelques pistes d'applications potentielles de ce nouveau biomarqueur en condition clinique préalablement à des études cliniques prospectives de suivi thérapeutique. Nous avons orienté les premières analyses cliniques au regard des études en cours au sein du service de radiologie de Gustave Roussy. Ainsi nous avons retenu deux pistes d'application en IRM. La première vise à étudier la capacité du H_{index} à s'affranchir de la difficulté récurrente à standardiser la fonction d'entrée artérielle (AIF) pour extraire des paramètres quantitatifs en DCE-IRM. La seconde piste d'application vise à explorer la capacité du H_{index} à établir un diagnostic précoce pour différencier les récurrences cérébrales de la radionécrose.

Avant de présenter ces deux applications et conformément aux recommandations relatives au développement d'un biomarqueur d'imagerie, nous avons mené une étude de répétabilité du H_{index} dans un contexte clinique sur volontaires sains examinés par IRM de diffusion (DWI-IRM) quantitative.

1 Étude de la répétabilité du H_{index} en IRM de diffusion cérébrale

1.1 Enjeux et objectifs

Le processus de développement d'un biomarqueur d'imagerie implique plusieurs étapes de validation dites techniques dans les recommandations d'O'Connor déjà discutées en chapitre 1 (O'CONNOR et al., 2017). L'un des critères de cette validation technique concerne la précision de mesure du biomarqueur et notamment l'évaluation de sa répétabilité en premier lieu par le biais d'une étude monocentrique sur un objet physique, un fantôme, ou à partir de données précliniques ou cliniques.

Dans le contexte de ce travail de thèse, le choix a été fait de réaliser cette évaluation sur volontaires sains via une étude clinique prospective. En effet, le développement de fantômes simulant l'hétérogénéité tumorale présente un réel défi, en particulier en IRM. Quelques approches sont décrites dans la littérature (VALLADARES et al., 2020) mais à l'heure actuelle un fantôme standardisé n'est pas disponible. Une étude clinique a été préférée à une étude préclinique car l'environnement de recherche associé à l'unité BioMaps nous le permettait facilement tout en respectant scrupuleusement la réglementation en vigueur. L'objectif du H_{index} étant de mesurer une hétérogénéité, nous nous sommes orientés vers l'IRM du cerveau puisque ce dernier possède une hétérogénéité naturelle complexe et que son repositionnement lors d'une étude test-retest est relativement aisé et reproductible en l'absence de mouvements physiologiques. Enfin, l'IRM de diffusion a été la modalité choisie pour accéder à une carte paramétrique quantitative, l'ADC, sans besoin d'injecter aux volontaires un produit de contraste.

Cette étude clinique sur volontaires sains consistait donc à déterminer la précision du biomarqueur en termes de variabilité et de répétabilité du H_{index} mesuré dans différentes régions d'intérêt présentant visuellement différents degrés d'hétérogénéité. Après avoir décrit la mise en place du protocole et la méthodologie d'analyse choisie, les résultats de l'étude seront présentés et discutés.

1.2 Protocole d'acquisition clinique

1.2.1 Déploiement du protocole volontaire sain

Notre étude s'inscrit dans le cadre du protocole de recherche générique « Méthodologie de Recherche en IRM » (MRIRM) qui regroupe les études sur volontaires sains réalisées sur le site du Service Hospitalier Frédéric Joliot (SHFJ) à Orsay. Cette recherche, dont le promoteur est le Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA), a pour objectif général l'« évaluation de nouvelles méthodes d'imagerie par

résonance magnétique pour l'étude de la structure et du fonctionnement du corps humain». Elle a fait l'objet d'une approbation par un comité de protection des personnes (CPP) et permet notamment la réalisation de deux examens successifs dans les mêmes conditions pour un même volontaire, nécessaires à notre étude de répétabilité. La sous-étude correspondante réalisée dans le cadre de cette thèse s'intitulait "Variabilité Inter et Intra Individu d'un paramètre de l'hétérogénéité en IRM de diffusion » (MRIRM-VIIPHD - Numéro d'amendement : 21.01587.000020).

L'étude prévoyait l'inclusion de 30 volontaires au maximum sur une durée de 36 mois.

Les critères d'inclusions des volontaires sains dans le cadre de l'étude MRIRM-VIIPHD étaient les suivants :

- Hommes et Femmes
- Entre 20 et 30 ans inclus.
- Ayant la capacité de coopérer et de rester immobile 1h30 au maximum
- Ayant la capacité à comprendre les instructions données
- Ayant signé le consentement éclairé
- Bénéficiant d'une couverture sociale
- N'étant pas en période d'exclusion pour une autre étude de recherche

Les critères de non-inclusions étaient :

- Personne privée de liberté
- Majeur protégé
- Grossesse
- Contre-indications à l'IRM
 - Implants (mécaniques ou électroniques : implants cochléaires, stimulateurs cardiaques ou neuronaux, pompes à infusion, clips anévrismaux magnétiques, ...)
 - Corps étrangers métalliques intraoculaires ou dans le système nerveux
 - Objets métalliques susceptibles de concentrer le champ de radiofréquence (tatouages, piercings, dispositif intra-utérin, ...)
- Claustrophobie ou anxiété importante
- Antécédents cardiovasculaire récents (<1 an) et sévères
- Toute affection grave rendant l'acquisition difficile ou inexploitable

Les données des volontaires sont protégées par le Règlement Général de la Protection des Données (RGPD) et les volontaires étaient informés du transfert temporaire de leur données pour la durée d'analyse des images vers le site de Gustave Roussy. Les conditions de transfert de données entre le site d'acquisitions du SHFJ et le site d'analyse de Gustave Roussy, ont fait l'objet d'un accord entre les deux parties notamment au regard de la RGPD et de la durée de stockage des données.

Les volontaires bénéficient d'une indemnisation de 80€ couvrant leur frais de déplacement et leur temps de disponibilité, intégralement prise en charge par Gustave Roussy. La mise au point du protocole d'acquisition sur la plateforme TEP/IRM 3T du SHFJ et la prise en charge des volontaires ont été réalisées en collaboration avec des manipulateurs radios du SHFJ.

1.2.2 Descriptif technique du protocole d'acquisition

Les acquisitions ont été menées sur un imageur 3 Tesla Signa TM PET/MR (GE Healthcare). Il s'agit d'un imageur hybride IRM-TEP, incorporant dans une seule machine les systèmes TEP et IRM. Seule la modalité IRM a été utilisée dans notre étude. Les acquisitions ont été réalisées à l'aide de l'antenne tête 3T à 32 canaux (Nova MedicalTM).

Le protocole d'acquisition des séquences était composé d'une séquence de localisation, d'une séquence anatomique axiale pondérée T2 et d'une séquence de diffusion IVIM. Les séquences dites IVIM évaluent les mouvements microscopiques des molécules d'eau causés par les effets de diffusion et la perfusion capillaire. Elles font l'objet d'un regain de popularité, notamment en oncologie car elles ne nécessitent pas d'agents de contraste pour accéder à une information de perfusion. Les effets de perfusion sont accentués par l'acquisition d'images à de multiples valeurs de faible pondération en diffusion ($b < 300 \text{ mm}^2/\text{s}$) permettant d'accéder à la pseudo diffusion D^* et à la fraction de perfusion f . Le modèle d'ajustement des courbes bi-exponentiel proposé par le LE BIHAN et al. (1986) prend en compte les effets de perfusion et diffusion et est fourni dans l'eq 4.1.

$$S/S_0 = fe^{-b(D+D^*)} + (1 - f)e^{-bD} \quad (4.1)$$

Avec S l'intensité du signal, S_0 l'intensité de base, sans sensibilisation à la diffusion, D le coefficient de diffusion mesuré, D^* le coefficient de pseudodiffusion reflétant la perfusion capillaire, et f la fraction perfusée autrement dit le pourcentage de voxels occupés par des capillaires.

Pour plus de détails concernant les séquences IVIM qui sont hors de notre étude, les lecteurs intéressés peuvent consulter la revue de LEBIHAN et al. (2019). Notons, qu'à partir des séquences IVIM, il est possible d'accéder aux paramètres énoncés mais également, par la sélection des valeurs de b , 0 et $1000 \text{ mm}^2/\text{s}$, de générer la carte paramétrique d'ADC actuellement plus largement utilisée en routine clinique et plus

Séquence	Axiale DWI IVIM	Positionnement QIBA DWI
Champ magnétique (Teslas)	3	1.5 ou 3
Nombre de canaux d'antenne tête	32	Idéal
Nombre de b	12/2 ¹	Idéal ≥ 3 , Acceptable 2
Valeur max. de b (s/mm^2)	1000	Idéal
Nombre de coupes	32	NC
Épaisseur de coupe (mm)	4 (jointive)	Idéal
Champ de vue (mm)	256x256	220-240 sur chaque axe
Matrice d'acquisition	160x160	Idéal
Orientation	Oblique	Transversal-Axial
Direction Fréquence	R/L	R/L
Nombre de moyennes	2	Idéal
Facteur d'accélération	2	Idéal
TR (ms)	3400	Acceptable
TE (ms)	5200	Idéal

TABLE 4.1 – Résumé des paramètres d'acquisition du protocole volontaire sain pour la séquence IVIM en comparaison des recommandations QIBA (SHUKLA-DAVE et al., 2019). Ces recommandations stipulent soit un paramètre fixe soit, lorsque cela est mentionné, une recommandation *Idéale* ou *Cible/Acceptable*. NC : non communiqué

largement explorée au cours de ce travail.

La séquence anatomique pondérée T2 était une séquence écho de spin aux paramètres suivants : nombre de coupes 64 ; coupes jointives ; épaisseur de coupe 4 mm ; matrice d'acquisition 160x160 ; champ de vue 256x256 mm² ; TR = 3400 ms ; TE = 100 ms ; Durée d'acquisition 2 min 30 sec. Les paramètres de la séquence IVIM sont fournis table 4.1 et comparés aux paramètres recommandés pour la réalisation d'une séquence de diffusion sur le cerveau par QIBA (SHUKLA-DAVE et al., 2019). Cette séquence a été réalisée avec 12 valeurs de pondérations de diffusion b (s/mm^2) : 0, 10, 20, 30, 40, 50, 80, 100, 200, 400, 800, 1000.

Contrairement aux recommandations QIBA préconisant une orientation de coupe en axial strict, et après consultation de nos collègues radiologues sur les pratiques usuelles en IRM cérébrale, nous avons choisi une orientation légèrement oblique pour chaque acquisition. En effet, pour garantir un repérage anatomique similaire pour l'ensemble des acquisitions, l'orientation des coupes axiales a été définie sur une coupe sagittale centrale selon la règle CACP (Corps-calleux antérieur, Corps-calleux Postérieur). La figure 4.1 illustre le positionnement du champ de vue.

Dans le contexte d'une étude sur la répétabilité des mesures, les volontaires étaient soumis à deux séances d'acquisitions successives. Chaque séance durait approximativement 25 minutes entre l'installation et la fin de l'acquisition. Pour considérer la variabilité de repositionnement du patient et des régions d'intérêt, il était demandé aux volontaires de sortir entièrement de la machine entre les deux acquisitions pour une durée oscillant entre 3 et 5 minutes.

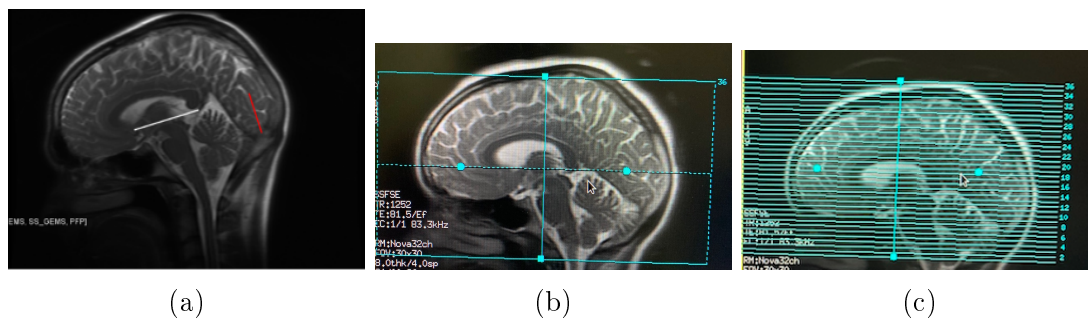


FIGURE 4.1 – Illustration de (a) l'alignement du plan oblique sur le corps calleux en imagerie anatomique ; (b) Alignement en pratique de la boîte du champ de vue sur le corps calleux dans le plan sagittal pour l'acquisition multi-coupes. (c) Lorsque l'alignement est effectuée, la boîte est remplacée pour prendre en compte l'ensemble du cerveau.

1.3 Protocole de traitement et d'analyse des données

Dans cette section nous décrivons l'ensemble des procédures post-acquisition qui regroupent la génération des cartes paramétriques, la définition des régions d'intérêt et les analyses statistiques. La figure 4.2 propose de récapituler les principales étapes du traitement des images et de l'analyse.

1.3.1 Génération des cartes paramétriques

A partir des images de diffusion acquises, les cartes paramétriques ont été générées à l'aide du logiciel Olea Sphère (v3.0.18, Olea Medical Solutions, France). Pour chaque acquisition d'IRM de diffusion, quatre cartes paramétriques ont été générées : une carte ADC issue d'un ajustement mono-exponentiel des courbes d'intensités en fonction de b , et les cartes D , D^* et f issues d'un ajustement bi-exponentiel. Pour l'ajustement mono-exponentiel, seules les valeurs de b minimum et maximum, à savoir 0 et 1000 étaient utilisées, afin de reproduire les conditions d'acquisition typiquement utilisées en IRM de diffusion clinique. A l'inverse, les 12 valeurs de b étaient employées pour l'ajustement bi-exponentielle. Avant l'ajustement des courbes voxel par voxel, une correction du mouvement et un filtrage spatial étaient effectuées par le logiciel. Les images étaient ensuite toutes converties au format NIFTI à l'aide de la librairie *dicom2nifti* en python.

1.3.2 Définition des régions d'intérêt et recalage

Différentes régions d'intérêt (ROI), cinq au total, ont été définies. L'emplacement de ces ROI a été défini par un radiologue expert selon deux critères. Premièrement, leur repérage anatomique devait être relativement simple afin de pouvoir les reproduire au mieux pour chaque acquisition et chaque volontaire. Ce repérage était facilité par le positionnement CACP des coupes lors de l'acquisition. Deuxièmement, elles devaient

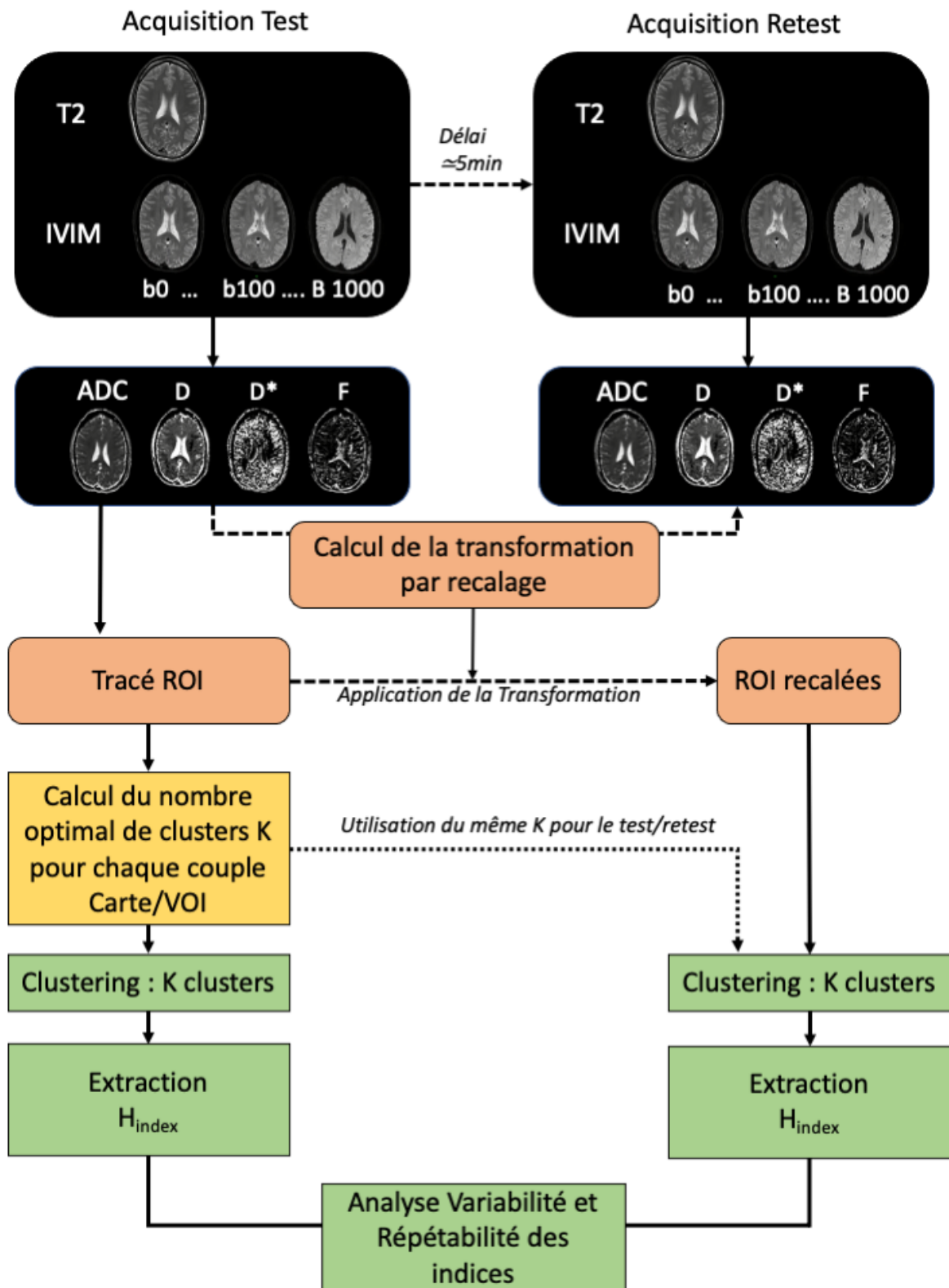


FIGURE 4.2 – Schéma récapitulatif du protocole volontaire sain MRIRM-VIIPHD. En noir, les étapes d'acquisition et de génération des cartes paramétriques. En orange, les étapes de tracé et de recalage des ROI. En jaune, l'étape d'optimisation du nombre de clusters. En vert, les processus d'extraction et d'analyse des indices.

être placées pour disposer de différents cas de figures représentant différents niveaux d'hétérogénéité d'image et des volumes variés. La description de chacune de ces ROI, qui correspondent à des volumes multi-coupes, est donnée ci-dessous, accompagnée d'illustrations montrant pour deux coupes représentatives des ROI tracées sur l'image pondérée $T2^2$ et la carte ADC pour comparaison (cf. Figure 4.3). La ROI #1, située dans l'hémisphère gauche, comprenait la tête du noyau caudé. La ROI #2, dans l'hémisphère gauche toujours, comprenait le striatum, la capsule extrême, la capsule externe, la capsule interne et le thalamus. La ROI #3 correspondait au splénium du corps calleux. La ROI #4 est l'extension de la ROI #3 dans la direction postérieure jusqu'au sillon pariéto-occipital. Enfin, la ROI #5 comprenait l'hémisphère droit, ventricules exclus. Les ROI choisies pour leur homogénéité tissulaires étaient les ROI #1 et #3 comprenant respectivement de la matière grise et de la matière blanche, toutes deux de volumes différents. Les ROI #2, #4 et #5 étaient choisies pour représenter des niveaux d'hétérogénéités supérieures aux régions précédentes, et de volumes également différents. La ROI #2 était composée de matière grise et blanche. Les ROI #4 et #5 comprenaient de la matière blanche, de la matière grise et du liquide céphalo-rachidien (LCR).

Les régions d'intérêt étaient tracées sur les cartes ADC en s'aidant des images pondérées T2 pour le repérage anatomique. Cela permettait de garantir un emplacement optimal notamment en raison de la déformation anatomique présente sur les images de diffusion en comparaison aux images T2. Cette déformation est notamment dû à l'utilisation d'une séquence écho planar en IRM de diffusion.

Les régions d'intérêt étaient définies sur les images de l'acquisition test puis transposées aux images retest. Pour garantir une reproductibilité optimale des positions des ROI entre les deux acquisitions, la transformation permettant de passer de l'image ADC retest à l'image ADC test a été calculée à partir d'un processus de recalage des images. L'inverse de cette transformation était ensuite appliqué à la ROI afin de la positionner sur l'image retest. Notons que l'application d'une transformation de recalage aux images aurait impliquée une interpolation des pixels avec une possible modification de la structure fine des images. Son application à la ROI permettait ainsi de conserver la structure fine des images originales. L'interpolation était réalisée par la méthode des plus proches voisins, seule méthode permettant d'éviter la modification des bordures des ROI.

La délimitation des ROI a été réalisée sous ITK-SNAP (YUSHKEVICH et al., 2006). Les ROI étaient enregistrées au format NIFTI et le recalage effectué à l'aide de la librairie *SimpleITK* en python avec les paramètres suivants : transformation : rigide étant donné qu'aucune déformation anatomique ou qu'aucun facteur d'échelle n'intervient entre les deux acquisitions ; optimiseur : Descente de gradient à pas réguliers afin de faciliter la

2. A titre illustratif uniquement puisque les ROI sont tracés sur la carte ADC pour les analyses comme précisé plus loin

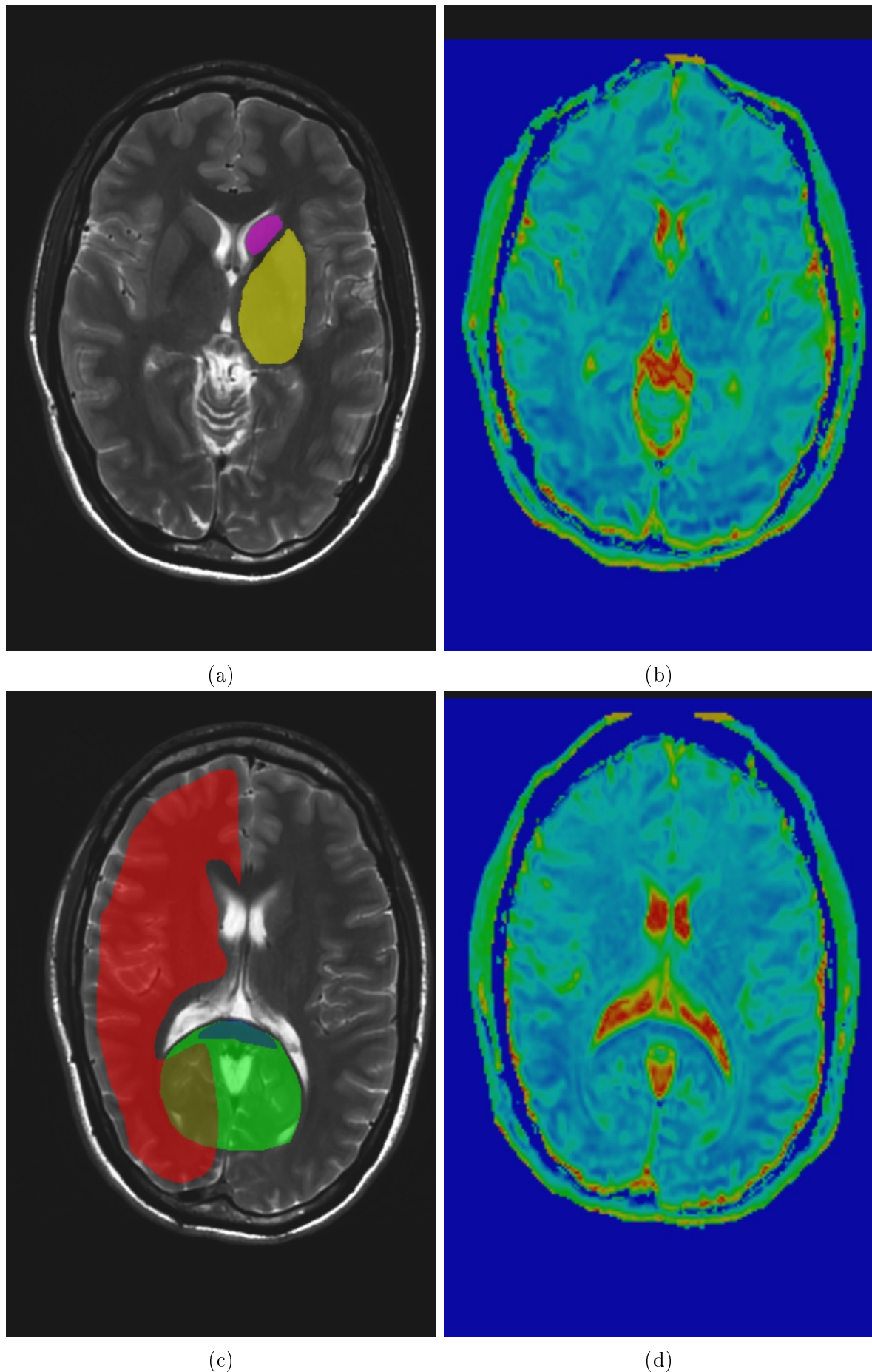


FIGURE 4.3 – Illustration des ROI sur deux coupes sélectionnées. (a) Coupe illustrant les ROI #1 (rose) et #2 (jaune). (b) Carte ADC correspondante de la coupe pour les ROI #1 et #2. (c) Coupe illustrant les ROI #3 (bleu), #4 (vert) et #5 (rouge). (d) Carte ADC correspondante de la coupe pour les ROI #3, #4 et #5.

convergence du recalage avec les paramètres ; taux d'apprentissage 2.0, pas minimum $1e^{-6}$, tolérance d'amplitude $1e^{-8}$; métrique de comparaison des images : information mutuelle de Mattes adaptée à l'imagerie paramétrique ADC ; Approche multi-résolution de facteurs [8,4] et de sigmas [2,1] permettant de faciliter la convergence. Enfin, le recalage était parcimonieusement corrigé manuellement lorsque la procédure *SimpleITK* n'offrait pas de solution satisfaisante visuellement. Ces corrections manuelles ont été effectuées sous *ITKSnap*.

1.3.3 Extraction des indices

La valeur moyenne et le H_{index} ont été extraits à partir de chaque ROI. Pour le clustering, l'algorithme Fuzzy C-means était employé. Le nombre de clusters était optimisé pour chaque individu au regard des deux acquisitions. En cas de différence entre l'acquisition test et l'acquisition retest, le nombre de clusters maximum était utilisé. Cette optimisation était réalisée pour chacune des quatre cartes paramétriques et chacune des cinq ROI. Pour le calcul du H_{index} , l'analyse en composantes connectées était réalisée avec une connectivité de valeur 18.

Pour les cartes ADC uniquement, les indices de texture issus de la GLCM, initialement décrits dans le chapitre 2, ont été extraits pour chacune des 5 ROI en vue d'établir leur robustesse et de les comparer au H_{index} . Le calcul de ces indices a été effectué à partir de LIFEx avec les paramètres suivants : les intensités étaient absolues et dans l'intervalle 0 et $3,4.10^{-3}$ mm²/s ; le nombre de barres d'histogrammes était fixé à 64 ; les calculs étaient effectués en 3D.

1.3.4 Analyses des données

L'analyse des données a été effectuée sous R (version 4.1.1, R Core Team, 2021). Pour chaque évaluation qu'il s'agisse de la variabilité ou de la répétabilité, l'analyse a tout d'abord portée sur les cartes ADC majoritairement utilisées en routine clinique et, pour le H_{index} et la valeur moyenne, les résultats ont ensuite été comparés à ceux obtenus sur les cartes D, D* et f issues de la méthode IVIM.

Pour aborder la répétabilité et la variabilité, trois métriques et un graphique, couramment rencontrés dans la littérature ont été utilisés : l'erreur de biais absolue (EBA) moyenne accompagnée des graphes Bland-Altman, le coefficient de corrélation intra classes (ICC) et une statistique plus traditionnelle avec le test de Wilcoxon.

L'EBA est un critère permettant de rendre compte de la précision des mesures intra-individus. D'un point de vue mathématique, il correspond à la moyenne des différences absolue entre deux acquisitions (eq 4.2) et a été calculé en complément du biais non-absolu moyen fourni par les graphes de Bland-Altman puisqu'il nous apparaissait plus judicieux dans le cadre d'une étude de test-retest de prendre en compte la différence

absolues entre les deux mesures. L'EBA est accompagné par les graphes de Bland-Altman qui permettent de représenter l'erreur relative moyenne (le biais non absolu) et les limites de concordances (LOA – limits of agreement) contenant 95 % des différences entre les deux mesures. Ensuite, les distributions ne pouvant pas être qualifiées de normales suite aux tests de normalité effectués, des tests de Wilcoxon appariés ont été utilisés pour évaluer la différence statistique entre deux jeux de mesures. Enfin, l'ICC, est une mesure de répétabilité largement répandue dans la littérature. Elle prend en considération à la fois la variabilité des mesures intra-individus et la variabilité entre les individus (eq 4.3).

Ces métriques ont été calculées pour analyser la variabilité et la répétabilité du H_{index} et de la valeur moyenne pour chaque ROI. Les valeurs d'ICC ont été, d'une part, comparées entre le H-index et la valeur moyenne du paramètre et, d'autre part, confrontées aux guidelines de KOO et LI (2016), répandues dans la littérature et qui mentionnent qu'une valeur d'ICC inférieure à 0,5 indique une faible répétabilité, qu'entre 0,5 et 0,75 la répétabilité est considérée comme modérée, et enfin bonne pour des valeurs comprises entre 0,75 et 0,9. Au-delà de 0,9 la répétabilité de la mesure est dite excellente. Pour le calcul de l'ICC, le package R « irr » était utilisé. L'ICC était basé sur le modèle mixte bidirectionnel (two-way mixed model), à concordance absolue et avec une unique métrique (single rater). Ce modèle correspond à la notation ICC(3,1) d'après la convention de SHROUT et FLEISS (1979).

$$EBA = \frac{1}{N} \sum_{i=1}^N |Test(i) - Retest(i)| \quad (4.2)$$

avec N , le nombre d'individus.

$$ICC(3, 1) = \frac{MS_{BS} - MS_{WS}}{MS_{BS} + (k - 1)MS_{WS}} \approx \frac{\sigma_{BS}}{\sigma_{BS} + \sigma_{WS}} \quad (4.3)$$

avec MS_{BS} et σ_{BS} respectivement l'erreur quadratique et la variance inter-sujets ; MS_{WS} et σ_{WS} l'erreur quadratique et la variance intra-sujet.

1.4 Résultats

1.4.1 Patients et Régions d'intérêts

Au total, 25 volontaires sains ont été inclus dans l'étude entre le 1er et 29 octobre 2021. Sur ces 25, deux volontaires ont été exclus de la suite de l'étude. Le premier en raison d'un arrêt de la procédure d'acquisition. Le second en raison d'une qualité d'image insuffisante sur le retest contenant des artefacts d'origines inconnues. Finalement, les données de 23 volontaires ont été utilisées pour l'analyse statistique. La table 4.2, précise les caractéristiques des volontaires et le volume des ROI. De plus, alors que les ROI #1

Caractéristiques	Valeurs
Volontaires	
Sexe	
Femmes	12/23 (52.2%)
Hommes	11/23 (47.8 %)
Indice de Masse Corporel (IMC) ($kg.m^{-2}$)	21.8 ± 2.3 (17.5 - 25.4)
Age (années)	25.2 ± 2.0 (22-29)
Régions d'intérêt - Volume (mm^3)	
ROI #1	1054
ROI #2	9815
ROI #3	1723
ROI #4	17554
ROI #5	121957

TABLE 4.2 – Caractéristiques de la population de volontaires sains et des cinq ROI.

à #4 se répartissent sur généralement 3 ou 4 coupes, la ROI #5 est à la marge, très volumineuse, ne représentant pas une zone anatomique bien définie et s'étalant sur 7 à 8 coupes en fonction des individus. Le nombre de clusters employé pour chaque mesure est indiqué en annexe. Les valeurs descriptives extraites à partir des cartes ADC pour l'ensemble des acquisitions test et retest sont fournis dans la table 4.3 et illustrées figure 4.4 et 4.5 respectivement pour la valeur moyenne et le H_{index} . Ces dernières indiquent des différences statistiques selon les ROI et en particulier les résultats concernant le H_{index} valident des régions d'hétérogénéité de diffusion et de perfusion très différentes.

1.4.2 Comparaison statistique entre les deux acquisitions

Pour chaque comparaison statistique entre l'ensemble des mesures test et des mesures retest, pour chaque ROI et pour chaque carte paramétrique, aucune différence statistique significative n'a été observée entre les données. Pour l'ADC, les valeurs de p évaluées pour chaque ROI étaient comprises entre 0,21 et 0,75 pour le H_{index} et 0,13 et 0,71 pour l'ADC moyen. Pour les cartes D, D* et F, l'ordre de grandeur était le même à l'exception du paramètre D* moyen pour la ROI #2 où p=0,056.

1.4.3 Évaluation du biais sur les cartes ADC

L'EBA pour les mesures du H_{index} et de la valeur moyenne d'ADC entre les acquisitions test et retest est reportée dans la table 4.3. Cette dernière est accompagnée des graphes Bland-Altman (figures 4.6 et 4.7) illustrant les différences entre les mesures test et retest en fonction de leurs moyennes pour chaque sujet, ainsi que le biais (non absolu) moyen et les LOA à 95%. Il apparaît alors que la valeur moyenne de l'ADC dans chaque ROI possède un faible biais absolu (EBA) [1,2; 3,4] %. Les valeurs de H_{index} possèdent un biais absolu moyen plus élevé de [11,3; 28,5] % pour les ROI 1-3 et 5 et de 52,9 % pour

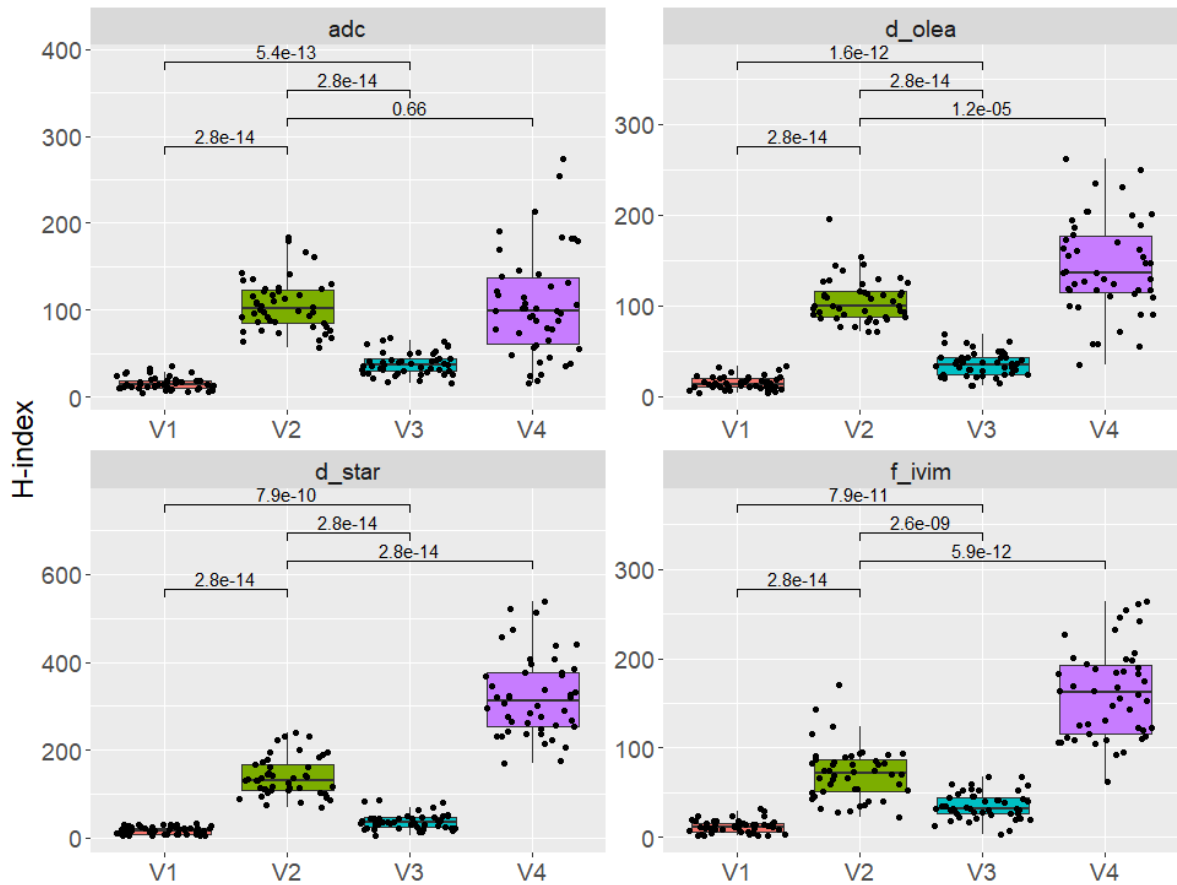


FIGURE 4.4 – Boxplots des H_{index} (u.a) pour les ROI #1 à #4 pour l'ensemble des acquisitions test et retest. V1 à V4 correspondent aux ROIS #1 à #4. La ROI #5 possédant des valeurs de H_{index} à la marge et très supérieures par rapport aux autres ROI n'est pas représentée ci-contre mais possédait une différence significative évidente avec les autres ROI.

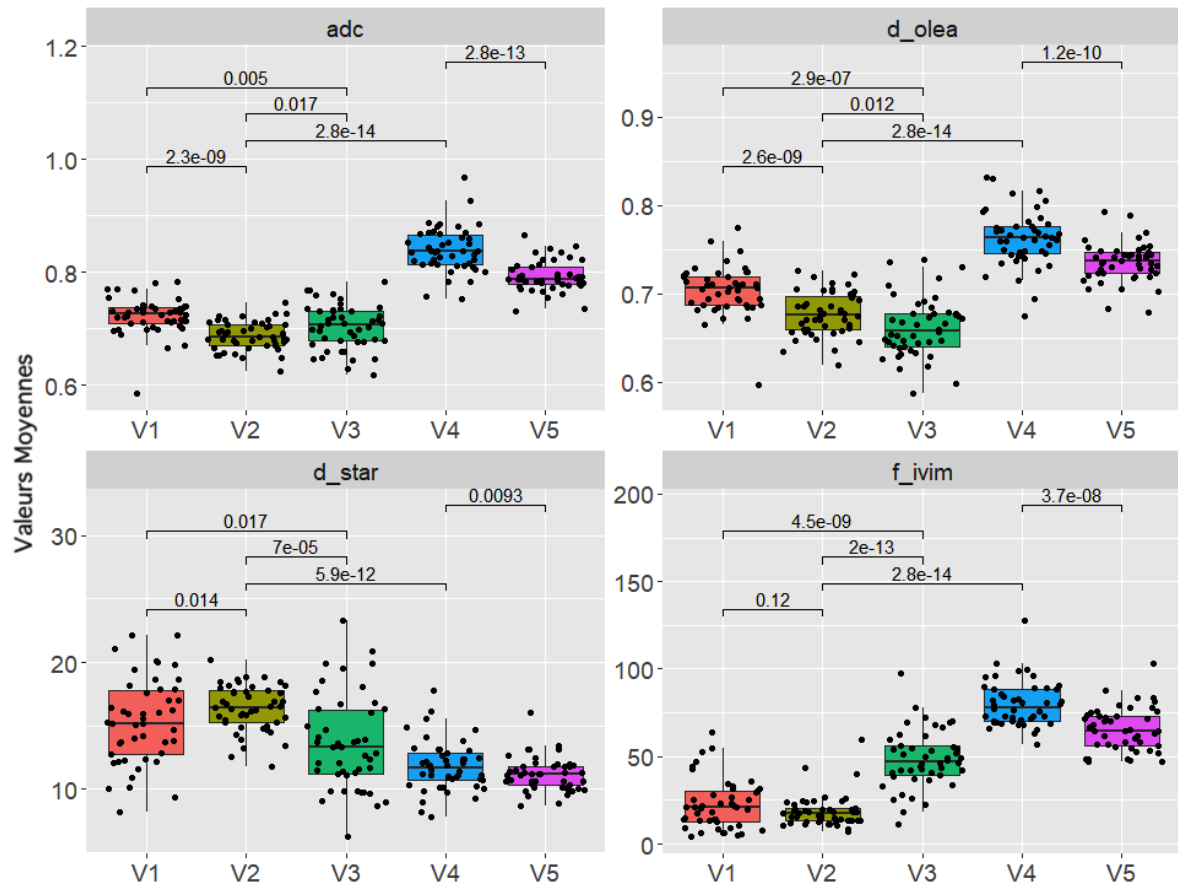


FIGURE 4.5 – Boxplots des valeurs moyennes des paramètres ADC ($10^{-3} \text{ mm}^2/\text{s}$), D ($10^{-3} \text{ mm}^2/\text{s}$), D^* ($10^{-3} \text{ mm}^2/\text{s}$), f (u.a) pour les ROI #1 à #5 pour l'ensemble des acquisitions test et retest. V1 à V5 correspondent aux ROIS #1 à #5.

ROI	Moyenne Test	Moyenne Retest	EBA	EBA (%)	p-value
H_{index} (u.a)					
ROI #1	15.3±6.6	16.0±8.3	2.4	11.3	0.5
ROI #2	109.1±30.6	105.9±30.5	28.3	28.5	0.56
ROI #3	39.5±13.2	37.0±13.1	7.6	21	0.21
ROI #4	107.6±66.5	107.6±55.3	32.9	52.9	0.75
ROI #5	1662.2±329.0	1657.5±384.0	208.7	13.2	0.69
ADC moyen ($10^{-3} \text{ mm}^2/\text{s}$)					
ROI #1	0.724 ±0.025	0.720±0.039	0.024	3.4	0.56
ROI #2	0.684±0.024	0.690±0.026	0.016	2.4	0.13
ROI #3	0.698±0.038	0.708±0.040	0.021	3.1	0.31
ROI #4	0.839±0.083	0.837±0.042	0.014	1.6	0.45
ROI #5	0.792±0.026	0.793±0.028	0.009	1.2	0.71

TABLE 4.3 – Valeurs moyennes du H_{index} et de l'ADC moyen pour chaque ROI pour l'examen test et l'examen retest de l'erreur de biais absolue moyenne (EBA), également en pourcentage, entre les deux acquisitions. La valeur de p indique la différence statistique de Wilcoxon apparié entre le test et le retest.

la ROI #4.

1.4.4 Évaluation de l'ICC sur les cartes ADC

Pour chaque région d'intérêt, l'ICC est reporté dans la table 4.4 pour le H_{index} et pour la valeur moyenne. En termes d'ICC, d'après les conventions, le paramètre ADC moyen est modérément répétable pour les ROI #1 à #3 et bien répétable pour les ROI #4 et #5. Concernant le H_{index} , l'ICC indique une faible répétabilité pour la ROI #2, une répétabilité modérée pour la ROI#5, une bonne répétabilité pour les ROI #3 et #4, et une excellente répétabilité pour la ROI #1.

Toutefois, en considérant les intervalles de confiances à 95%, et d'après les recommandations pour la présentation des résultats de l'ICC par KOO et LI (2016), il apparait que l'ADC moyen devrait être considéré comme faiblement à modérément répétable pour les ROI #1, #2 et #3. Les ROI #4 et #5 conservent leur statut de bien répétable. Le H_{index} devrait lui être considéré comme non répétable pour la ROI#2, faiblement à modérément pour la ROI #5, de modérément à bien répétable pour les ROI #3 et #4, et enfin, de bien à excellent pour la ROI #1.

1.4.5 Analyse de texture

Les valeurs descriptives des indices de texture extraits à partir des cartes ADC sont indiqués dans la table 4.5 ainsi que les résultats liés à l'analyse de la variabilité avec l'EBA moyen et la valeur de p issue des test de Wilcoxon. Pour les comparaisons

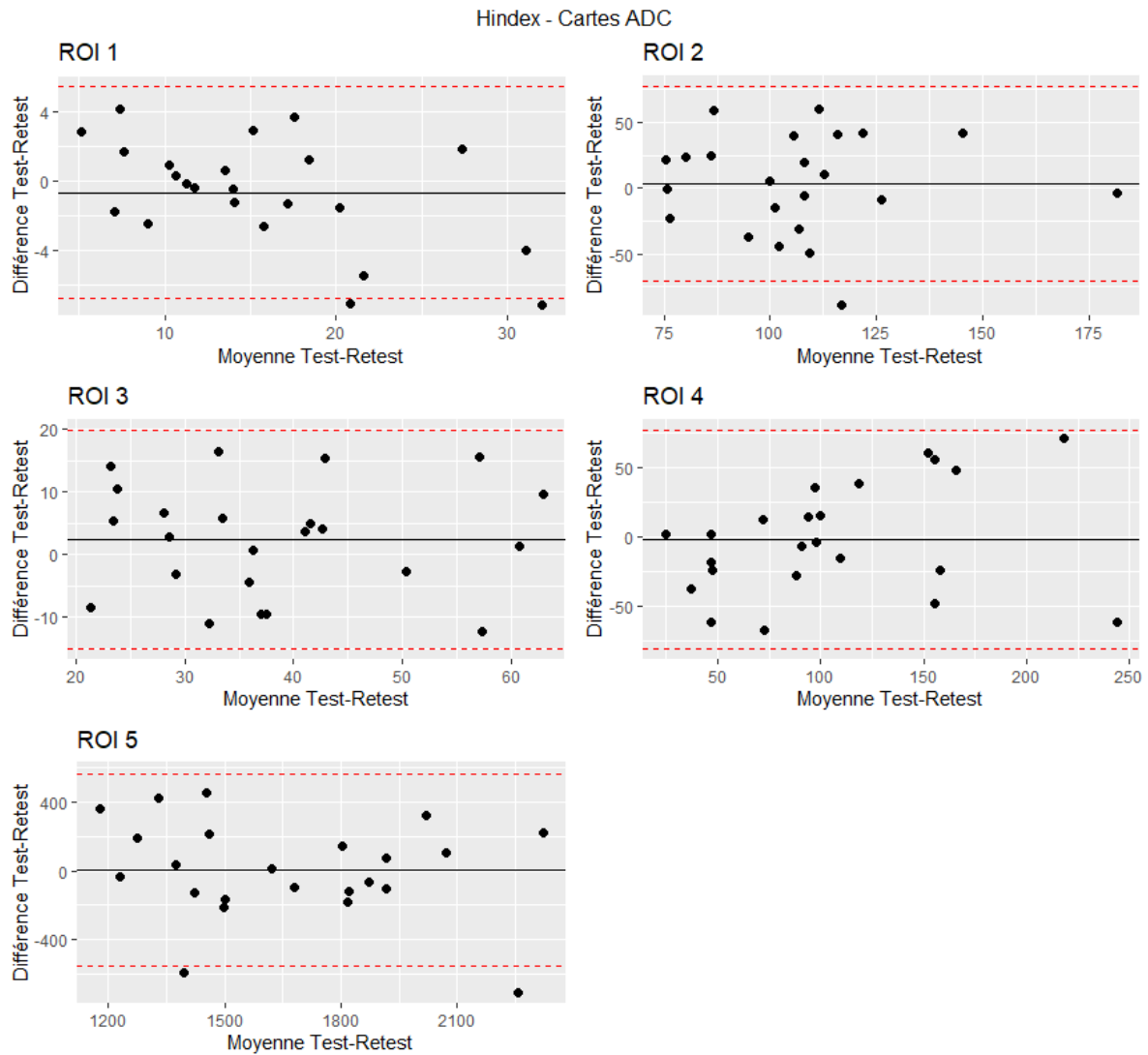


FIGURE 4.6 – Graphes de Bland-Altman pour le H_{index} (u.a) mesuré à partir des cartes ADC et représenté pour chacune des ROI. La ligne pleine noire correspond au biais non-absolu moyen. Les lignes pointillées rouges correspondent aux LOA à 95%.

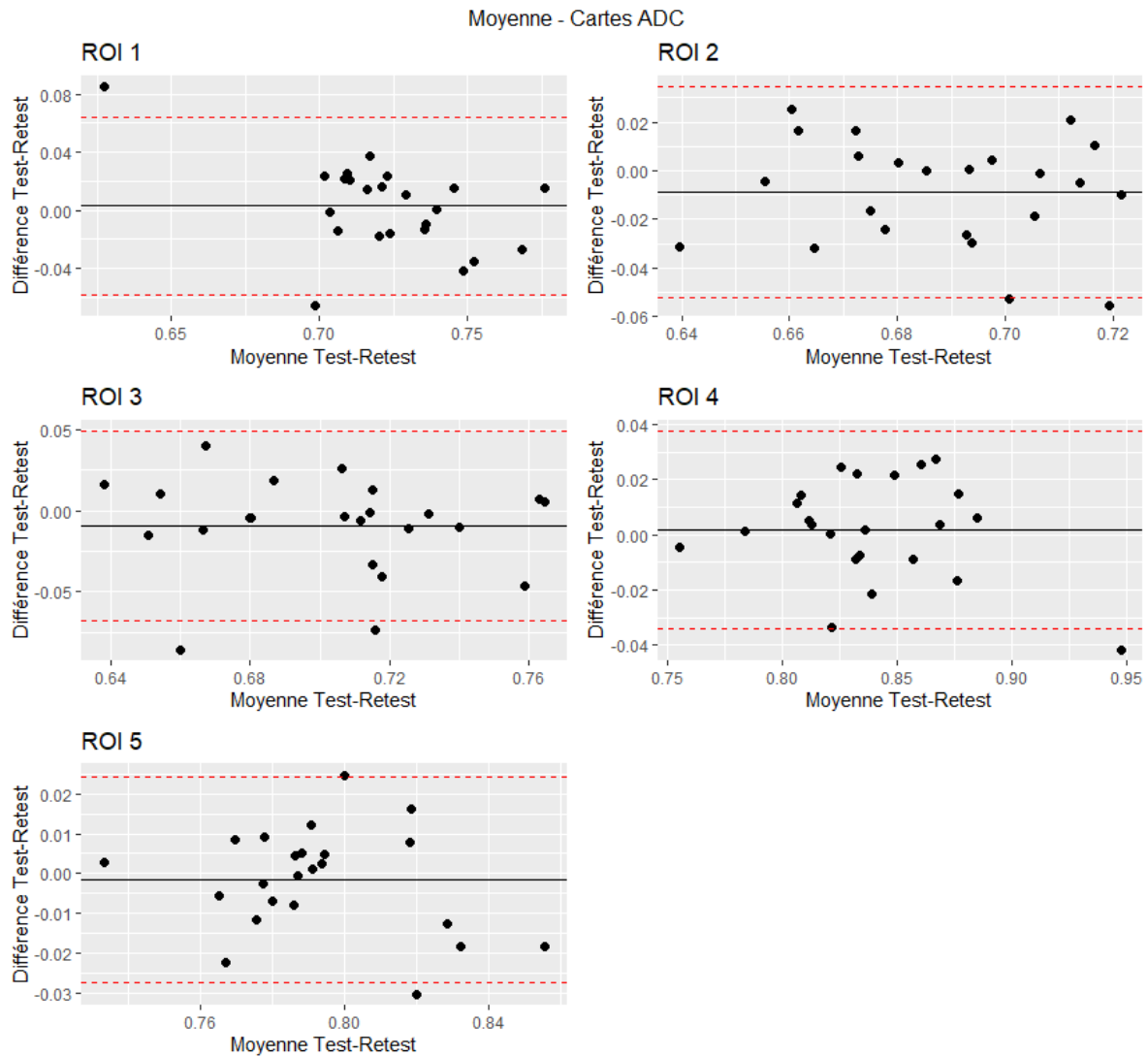


FIGURE 4.7 – Graphes Bland-Altman pour le paramètre ADC ($10^{-3}mm^2/s$) moyen et représenté pour chacune des ROI. La ligne pleine noire correspond au biais non-absolu moyen. Les lignes pointillés rouges correspondent aux LOA à 95%.

ROI	ICC	CI 95%	F-value	p-value
H_{index}				
ROI#1	0.91	[0.81 0.96]	22.3	9e-11
ROI#2	0.31	[-0.14 0.65]	1.9	8.5e-2
ROI#3	0.76	[0.52 0.89]	7.7	2.8e-6
ROI#4	0.78	[0.56 0.90]	7.9	2.5e-6
ROI#5	0.69	[0.40 0.85]	5.3	8.8e-5
ADC Moyen				
ROI#1	0.57	[0.21 0.79]	3.6	1.9e-3
ROI#2	0.66	[0.34 0.84]	5.1	2.6e-4
ROI#3	0.70	[0.44 0.86]	6.0	2.2e-5
ROI#4	0.89	[0.78 0.95]	18.4	4.3e-10
ROI#5	0.88	[0.75 0.9]	16.3	1.6e-9

TABLE 4.4 – Répétabilité du paramètre moyen d'ADC et du H_{index} au regard de chacune des ROI. La répétabilité est évaluée par l'ICC, son intervalle de confiance à 95%. La valeurs de F et de p indiquent la fiabilité de l'ICC.

statistiques entre l'ensemble des mesures test et retest, aucune différence statistique n'a été observée pour chacun des indices.

Les intervalles d'EBA moyens en pourcentage pour chacun des indices selon les ROI étaient pour l'homogénéité [4,2 ; 12,0], pour l'énergie [17,6 ; 39,4], pour le contraste [31,4 ; 169,4], pour la corrélation [7,4 ; 121,6], pour l'entropie [3,3 ; 12,0] et pour la dissimilarité [12,4 ; 40,9]. La ROI #1, de plus petit volume, possédait l'EBA le plus élevé pour tous les indices à l'exception du contraste pour lequel l'EBA moyen maximal était observé pour la ROI #2. A l'inverse la ROI #5 correspondant au volume le plus important, disposait de l'EBA le plus faible sauf pour l'entropie où l'EBA de la ROI #4 atteignant 3,3 contre 4,12 pour la ROI #5 et pour l'énergie pour laquelle la ROI #2 démontrait un EBA de 17,6 % contre 18,5 pour la ROI #5. En comparaison avec le H_{index} (EBA moyen = [13,2 ; 52,9] %) comme illustré sur la figure 4.8, le contraste et la corrélation apparaissent plus variable, l'entropie et l'homogénéité semblent moins variables et l'énergie et la dissimilarité se situent dans le même ordre de grandeur.

Concernant la répétabilité, pour chaque région d'intérêt, l'ICC est reporté dans la table 4.6 pour chaque indice de texture accompagné des intervalles de confiance à 95% et des résultats du test F associé. En termes d'ICC, d'après les conventions, le contraste était non répétable (ICC<0,5) quelle que soit la ROI. La dissimilarité n'était pas répétable pour les ROI #1, #2 et #4 et modérément pour les #3 et #5. L'homogénéité n'était pas répétable pour la ROI #1 et elle l'était modérément pour les #2 à #5. L'énergie n'était pas répétable pour la ROI #1, modérément pour les #2, #3 et #5 et bien répétable pour la ROI #4. L'entropie n'était pas répétable pour la ROI #1, modérément pour les #2 et #3 et bien répétable pour les #4 et #5. Enfin la corrélation n'était pas répétable pour la ROI #2, modérément pour les #3, #4 et #5

Indice & ROI	Moyenne Test	Moyenne Retest	EBA	EBA (%)	p-value
Homogénéité					
#1	0.542	0.562	0.06	12.0	0.24
#2	0.642	0.637	0.03	4.9	0.43
#3	0.592	0.580	0.04	7.1	0.33
#4	0.475	0.474	0.03	6.7	1.00
#5	0.558	0.550	0.02	4.2	0.46
Énergie					
#1	0.064	0.072	0.022	39.4	0.13
#2	0.074	0.069	0.013	17.6	0.14
#3	0.062	0.061	0.014	23.8	0.73
#4	0.023	0.023	0.004	20.4	0.99
#5	0.035	0.035	0.006	18.5	0.71
Contraste					
#1	10.19	8.02	8.99	123.4	0.28
#2	2.18	3.83	2.24	169.4	0.27
#3	3.41	5.00	2.61	77.1	0.06
#4	19.14	18.57	6.11	31.4	0.52
#5	9.12	10.63	2.23	31.5	0.13
Corrélation					
#1	0.15	0.14	0.07	121.6	0.72
#2	0.49	0.50	0.11	27.3	0.38
#3	0.33	0.32	0.15	43.7	0.92
#4	0.56	0.56	0.06	11.2	0.89
#5	0.55	0.53	0.04	7.4	0.16
Entropie					
#1	1.40	1.36	0.16	12.0	0.33
#2	1.37	1.41	0.09	7.0	0.16
#3	1.39	1.42	0.09	6.7	0.38
#4	2.04	2.04	0.07	3.3	0.92
#5	1.82	1.82	0.07	4.12	0.96
Dissimilarité					
#1	1.77	1.61	0.74	40.9	0.36
#2	0.98	1.06	0.22	24.1	0.27
#3	1.22	1.36	0.29	24.4	0.12
#4	2.62	2.62	0.41	14.9	0.89
#5	1.76	1.87	0.19	12.5	0.25

TABLE 4.5 – Valeurs moyennes de chacun des indices de texture issus de la GLCM à partir des cartes ADC pour chaque ROI pour l'examen test et l'examen retest et l'EBA moyen, également en pourcentage, entre les deux acquisitions. La valeur de p indique la différence statistique de Wilcoxon entre le test et le retest.

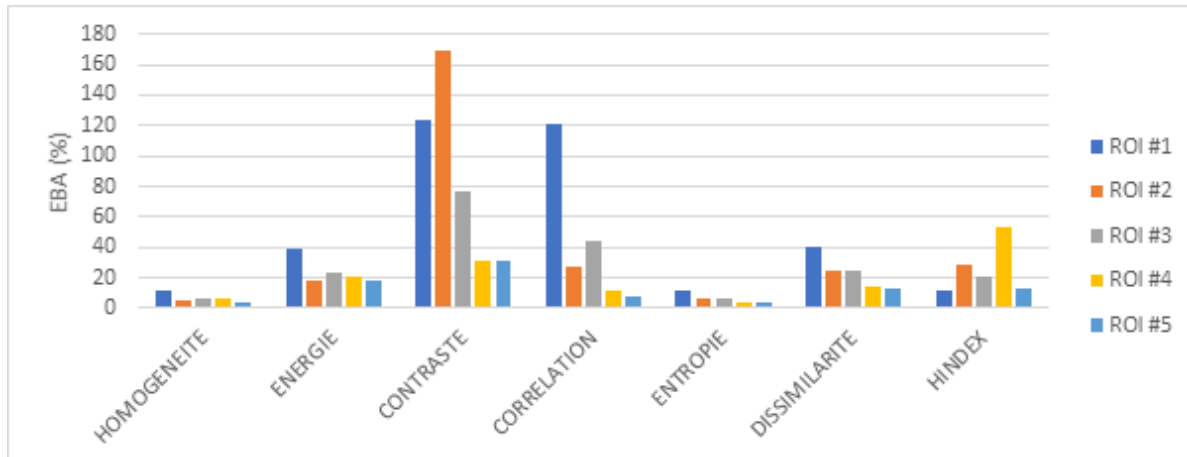


FIGURE 4.8 – Comparaison des EBA moyens en pourcentage pour chacun des indices de texture issus de la GLCM avec ceux du H_{index} pour les cinq ROI à partir des cartes ADC.

et excellentement pour la ROI #1. Finalement nous notons que pour la ROI #3, tous les indices étaient modérément répétables.

La répétabilité des indices de texture issus de la GLCM est comparée en figure 4.9 à la répétabilité du H_{index} au travers des valeurs d'ICC. Le H_{index} était ainsi le plus répétable pour les ROI #1 et #3, les indices de texture obtenaient de meilleures performances pour la ROI #2 avec l'homogénéité, l'entropie et l'énergie ($ICC > 0,5$), pour la ROI #4 avec l'entropie $ICC=0,84$ vs $0,78$ pour le H_{index} , et enfin pour la ROI #5 avec l'entropie et la corrélation ($ICC > 0,69$ du H_{index}) et notamment pour l'entropie étant le seul indice bien répétable ($ICC > 0,75$).

1.4.6 Évaluation du biais et de l'ICC sur les cartes D, D* et f

Les tables 4.7, 4.8, 4.9 reportent respectivement les valeurs moyennes pour chaque carte, le biais absolu moyen, l'ICC et son intervalle de confiance à 95% ainsi que la valeur de p associée à l'ICC. Les ICC de ces tableaux ainsi que ceux de l'ADC, sont résumées dans la figure 4.10 représentant les ICC et leurs intervalles de confiances ainsi que les limites des conventions d'interprétation du ICC. Il ressort de ces résultats que pour les cartes D* et f, la valeur du paramètre moyen ne possédait pas une meilleure répétabilité que le H_{index} . Pour la carte D, le paramètre moyen surpasse systématiquement les performances du H_{index} mais sa bonne répétabilité était observée dans les ROI #4 et #5 uniquement, donc dans les ROI de volumes importants ($>15\text{cm}^3$).

1.5 Discussion

Nous nous sommes intéressés à la répétabilité et à la variabilité d'un biomarqueur potentiel de l'hétérogénéité, le H_{index} , au cours d'une étude prospective sur volontaires

ROI & indice	ICC	CI 95%	F-value	p-value
Homogénéité				
#1	0.21	[-0.19 0.56]	1.6	0.15
#2	0.69	[0.39 0.85]	5.3	1.2e-4
#3	0.62	[0.31 0.82]	4.5	3.4e-4
#4	0.56	[0.19 0.78]	3.4	2.8e-3
#5	0.64	[0.320.82]	4.5	3.2e-4
Energie				
#1	0.28	[-0.20 0.61]	1.8	8.24 e-2
#2	0.73	[0.46 0.87]	6.7	1.72 e-5
#3	0.64	[0.32 0.83]	4.5	3.81 e-4
#4	0.78	[0.54 0.90]	7.7	5.18 e-6
#5	0.61	[0.27 0.81]	4.0	8.39 e-4
Contraste				
#1	-0.1	[-0.43 0.40]	0.9	0.52
#2	-0.06	[-0.42 0.33]	0.8	0.62
#3	0.32	[-0.06 0.63]	2.0	0.05
#4	0.35	[-0.07 0.67]	2.0	0.05
#5	0.38	[0.01 0.68]	2.4	0.02
Corrélation				
#1	0.91	[0.79 0.95]	20.0	4.7 e-10
#2	0.28	[-0.15 0.62]	1.8	9.58 e-2
#3	0.51	[0.12 0.75]	3.0	6.39 e-3
#4	0.62	[0.28 0.82]	4.18	7.11 e-4
#5	0.70	[0.42 0.85]	5.9	4.40 e-5
Entropie				
#1	0.46	[0.07 0.72]	2.7	1.12 e-2
#2	0.68	[0.38 0.84]	5.5	8.10 e-5
#3	0.74	[0.50 0.88]	7.0	8.84 e-6
#4	0.84	[0.66 0.93]	11.4	1.53 e-7
#5	0.75	[0.51 0.89]	7.0	1.17 e-5
Dissimilarité				
#1	0.01	[-0.40 0.41]	1.0	0.48
#2	0.35	[-0.04 0.65]	2.1	0.04
#3	0.55	[0.21 0.78]	3.7	1.61 e-3
#4	0.43	[0.024 0.71]	2.4	0.020
#5	0.56	[0.22 0.78]	3.7	1.5 e-3

TABLE 4.6 – Répétabilité des indices de texture issus de la GLCM à partir des cartes ADC pour chaque ROI. La répétabilité est évaluée par l'ICC et son intervalle de confiance à 95%. Les valeurs de F et de p indiquent la fiabilité de l'ICC.

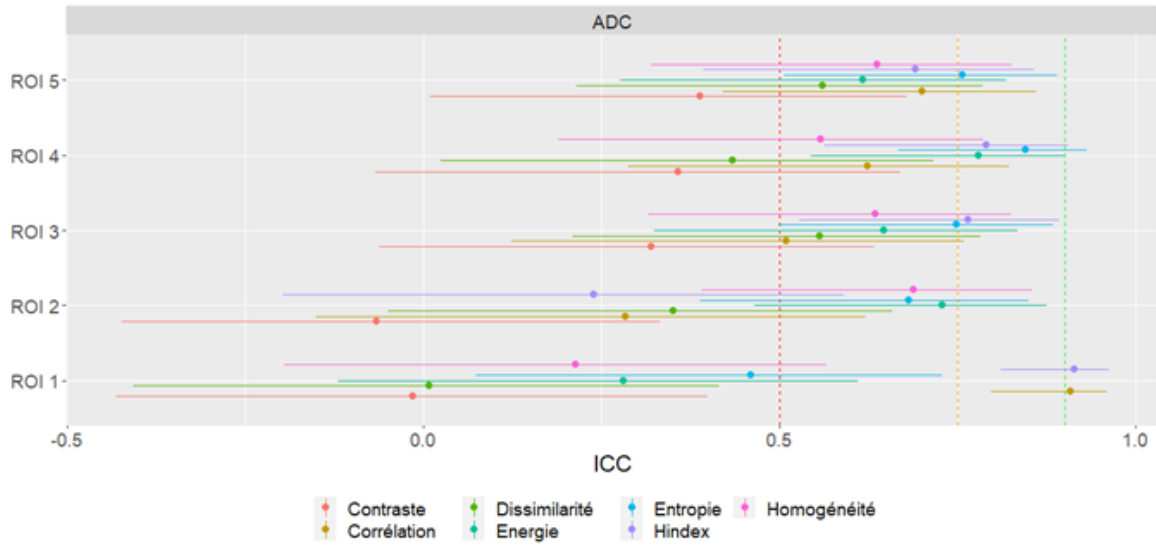


FIGURE 4.9 – Comparaison de l'ICC des indices de texture avec l'ICC du H_{index} . Les indices sont évalués à partir des cartes ADC pour chaque ROI.

ROI	Moyenne Test	Moyenne Retest	EBA	EBA (%)	ICC	CI 95%	p-value
H_{index} (u.a)							
ROI#1	15.0±8.8	14.9 ±5.3	7.0	80.3	0.28	[-0.14 0.62]	9.4e-2
ROI#2	103.1±27.5	109.0±21.4	17.3	18.1	0.65	[0.34 0.83]	1.9e-4
ROI#3	35.9±13.5	34.1 ±13.3	10.1	28.9	0.57	[0.22 0.79]	1.6e-3
ROI#4	145.2±55.7	140.9±49.4	42.3	40.2	0.51	[0.13 0.76]	5.5e-3
ROI#5	1839.2±339.8	1768.8±370.6	230.0	12.5	0.57	[0.22 0.79]	1.47e-3
D Moyen ($10^{-3} mm^2/s$)							
ROI#1	0.707 ±0.023	0.701±0.031	0.022	3.1	0.34	[-0.06 0.65]	4.9e-2
ROI#2	0.678 ±0.027	0.677±0.0	0.015	2.2	0.71	[0.42 0.86]	5.9e-5
ROI#3	0.660±0.032	0.663±0.035	0.022	3.4	0.62	[0.29 0.82]	6.0e-4
ROI#4	0.764±0.029	0.760±0.034	0.016	2.1	0.79	[0.54 0.90]	1.3e-6
ROI#5	0.736±0.019	0.734 ±0.025	0.013	1.7	0.76	[0.51 0.89]	8.8e-6

TABLE 4.7 – Répétabilité du H_{index} et de la valeur moyenne de D au regard de chacune des ROI. La variabilité est évaluée par l'erreur de biais absolue (EBA) et la répétabilité l'ICC et son intervalle de confiance ainsi que la valeur de p associée.

ROI	Moyenne Test	Moyenne Retest	EBA	EBA(%)	ICC	CI 95%	p-value
H_{index} (u.a)							
ROI#1	16.6±8.3	16.2±8.8	7.6	56.4	0.31	[-0.11 0.64]	7.0e-2
ROI#2	128.5±44.7	147.9±42.4	36.7	34.5	0.45	[0.08 0.71]	8.9e-3
ROI#3	40.5±40.5	35.0±17.1	15.3	45.4	0.46	[0.09 0.73]	8.9e-3
ROI#4	327.4±94.0	317.9±91.2	56.0	17.5	0.73	[0.48 0.88]	1.7e-5
ROI#5	3688.9±586.3	3660.4±633.3	394.4	11.0	0.66	[0.34 0.83]	2.9e-4
D* Moyen ($10^{-3} \text{ mm}^2/\text{s}$)							
ROI#1	15.06±2.9	15.58±3.8	3.75	25.3	0.17	[-0.26 0.54]	2.1e-1
ROI#2	16.59±2.2	16.18±1.4	2.13	13.5	-0.08	[-0.49 0.33]	6.5e-1
ROI#3	13.74±3.2	13.47±4.2	2.95	21.0	0.45	[0.04 0.72]	1.5e-2
ROI#4	11.87±1.8	11.87±2.2	1.97	17.7	0.45	[0.04 0.72]	1.6e-1
ROI#5	11.19±1.2	11.18±1.5	1.11	10.2	0.33	[-0.09 0.65]	5.8e-2

TABLE 4.8 – Répétabilité du H_{index} et de la valeur moyenne de D* au regard de chacune des ROI. La variabilité est évaluée par l'erreur de biais absolue (EBA) et la répétabilité l'ICC et son intervalle de confiance ainsi que la valeur de p associée.

ROI	Moyenne Test	Moyenne Retest	EBA	EBA (%)	ICC	CI 95%	p-value
H_{index} (u.a)							
ROI#1	11.9±7.2	12.0 ±7.4	9.0	114.2	-0.06	[-0.48 0.36]	0.6
ROI#2	67.6±32.7	75.6±26.5	31.7	67.2	0.11	[-0.3 0.49]	0.29
ROI#3	37.8±15.7	32.4±14.3	17.3	48.4	-0.05	[-0.44 0.34]	0.61
ROI#4	158.10±56.3	164.34±45.6	39.1	28.1	0.58	[0.22 0.79]	1.5e-2
ROI#5	1988.9±336.2	1983.6±278.0	254.7	12.8	0.49	[0.09 0.74]	8e-3
f Moyen (u.a.)							
ROI#1	23.1±15.1	22.1±13.6	14.4	97.0	0.13	[-0.30 0.51]	0.27
ROI#2	15.8±4.9	21.8±11.7	9.6	70.3	-0.001	[-0.32 0.36]	0.50
ROI#3	46.5 ±17.1	50.5±16.0	13.0	40.8	0.54	[0.18 0.77]	0.50
ROI#4	79.6±12.0	80.6±14.1	10.6	13.5	0.40	[-0.01 0.69]	0.03
ROI#5	64.3±11.8	67.24±13.1	8.4	14.0	0.53	[0.17 0.77]	2.9e-3

TABLE 4.9 – Répétabilité du H_{index} et de la valeur moyenne de f au regard de chacune des ROI. La variabilité est évaluée par l'erreur de biais absolue (EBA) et la répétabilité l'ICC et son intervalle de confiance ainsi que la valeur de p associée.

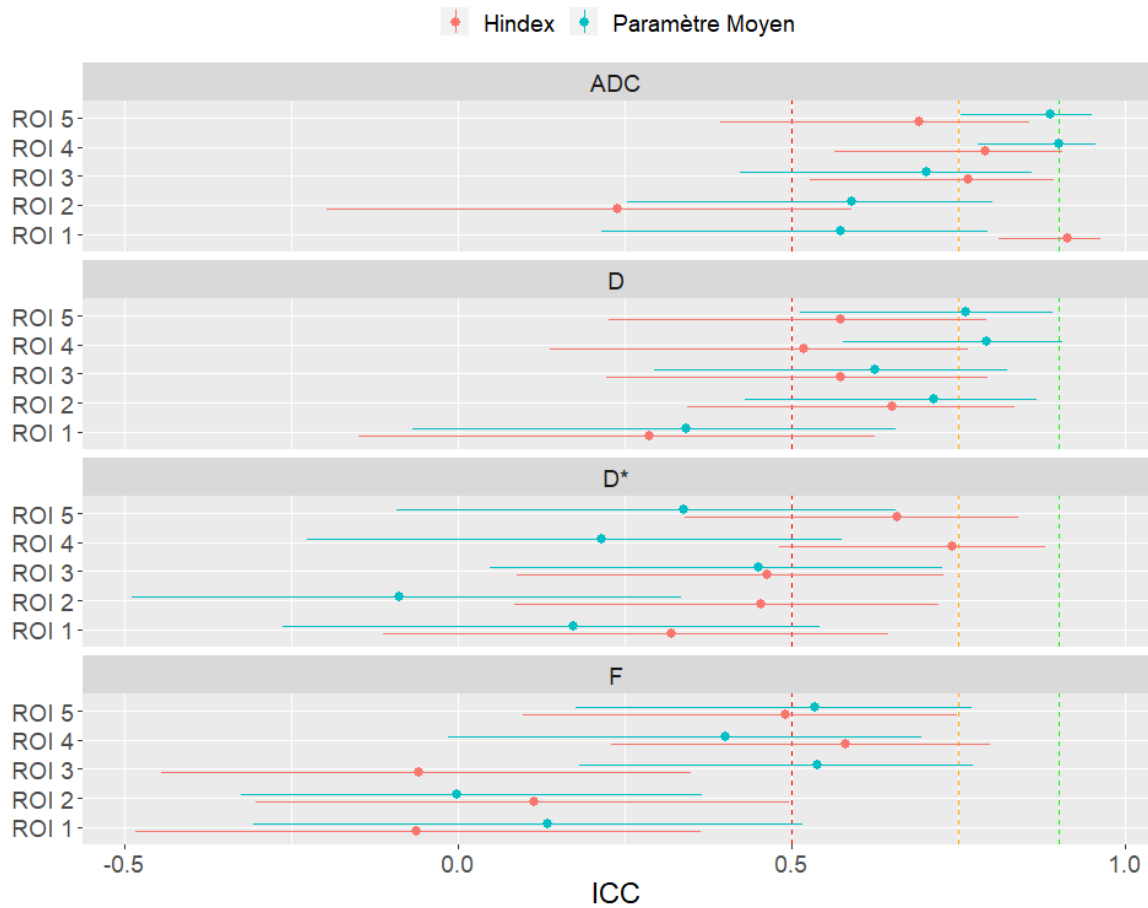


FIGURE 4.10 – Représentation graphique des ICC du H_{index} (rouge) et de la valeur moyenne (bleu) et leurs intervalles de confiance à 95% pour chaque ROI et chaque carte paramétrique. Les lignes en pointillés représentent les limites des conventions avec en rouge la limite entre faiblement et modérément répétable, en orange la limite entre modérément et bien répétable, en vert la limite entre une bonne répétabilité et une excellente.

sains dans le cadre de l'imagerie paramétrique IRM de diffusion. L'EBA représentait une métrique d'évaluation de la variabilité tandis que l'ICC représentait une métrique de la répétabilité entre les deux acquisitions réalisées dans des conditions identiques. Ces conditions remplissaient les recommandations méthodologiques pour les études de répétabilité du groupe QIBA à savoir la même procédure de mesure, le même système, le même opérateur, les mêmes conditions opérationnelles, la même localisation et surtout une réalisation au cours d'une courte période de temps permettant de minimiser l'effet de composantes extérieures sur la variance (RAUNIG et al., 2015). Dans le cadre de cette étude, nous avons présenté et discuté les données du H_{index} . Toutefois, les autres métriques qui le composent ont été évidemment calculés mais leur répétabilité n'est pas présentée ici au regard du manque d'intérêt pour une étude de test-retest.

Le critère de répétabilité était influencé par les sources de variabilité suivantes : l'orientation des coupes, le positionnement des ROI, la variabilité intrinsèque du système d'acquisition et enfin la variabilité de la mesure du H_{index} reflétant la variabilité de la mesure de l'hétérogénéité des images acquises. Lors des différentes acquisitions, le volontaire était entièrement remplacé dans la machine conduisant à un positionnement des coupes axiales légèrement variable. Un recalage était effectué pour obtenir une correspondance entre les coupes des deux acquisitions mais au regard de la résolution de 4mm dans le plan sagittal, coupes jointives, et des effets de volumes partiels, le repositionnement des coupes pouvait être imparfait et donc engendrer une variabilité entre les deux acquisitions. Pour obtenir une information de répétabilité qui n'aurait été relative qu'à la mesure, il aurait fallu acquérir deux images consécutives au cours de la même séquence sans remplacer le volontaire. Néanmoins, en vue d'une utilisation clinique, la prise en compte de cette variabilité d'acquisition au sein de la variabilité totale nous paraissait nécessaire.

A première vue, les valeurs moyennes des ADC et du H_{index} varient peu entre l'acquisition test et retest et ce pour toutes les régions. C'est également ce qui ressort avec les tests statistiques de Wilcoxon appariés qui n'ont pas permis de relever de différences significatives pour chacune des ROI entre les acquisitions test et retest. Toutefois si ce type d'analyse statistique est assez conventionnel, il n'est probablement pas des plus pertinents dans les études de test-retest (SHESKIN, 2011).

Globalement, nos résultats ont montré que pour les cartes ADC, les valeurs moyennes d'ADC étaient peu variables au regard du biais et bien répétables selon l'ICC uniquement pour les deux ROI #4 et #5 qui étaient 2 à 12 fois plus grandes que les trois autres. Ce premier résultat souligne la variabilité des mesures d'ADC qui est atténuée avec un volume d'intérêt croissant comme observé par CHEN et al. (2021) lors d'une étude sur le pancréas. De plus, la répétabilité des valeurs moyennes du paramètre ADC était comparable aux résultats de HAKULINEN et al. (2012) également dans différentes régions anatomiques cérébrales sur volontaires sains.

Dans le cas du H_{index} il est difficile d'identifier la source de sa bonne ou mauvaise répétabilité. La mesure de l'hétérogénéité via le H_{index} présentait des biais non absolus moyens variant tous autour de zéro et les EBA moyens (en %) moins élevés étaient obtenus pour les ROI #1, #3 et #5. En d'autres termes, les différences individuelles étaient distribuées plus ou moins symétriquement autour de 0. De plus, au regard des valeurs des EBA pour ces trois ROI, le H_{index} présentait une faible variabilité intra-individu attribuée aux mesures et au positionnement des ROI. Notons qu'une variabilité d'un biomarqueur d'imagerie entre 15 et 30% est typiquement admise (BANE et al., 2018; LASSAU et al., 2017). Concernant les deux ROI avec un EBA élevé, #2 et #4, cette variabilité était importante. Dans le cas de la ROI #4 dont la localisation est proche des ventricules, son positionnement systématique s'est avéré dans la pratique nettement plus délicat que pour les autres ROI. Notons également que l'EBA élevé pour la ROI #4 résultait également d'une forte contribution d'une valeur aberrante. Le choix de ne pas l'exclure de l'étude a été maintenu car l'inspection visuelle des images ne permettait pas de les disqualifier par manque de qualité d'image. Toutefois, étant donné que ce volontaire aurait été exclu d'après les recommandations relatives aux valeurs aberrantes (RAUNIG et al., 2015), nous notons que son exclusion permettrait de diminuer le biais de 52,9 à 40,1 %, ce qui reste un EBA élevé. Dans le cas de la ROI #2, la variabilité du H_{index} pourrait s'expliquer par la distribution des valeurs d'intensité concernée. Plus précisément, il s'agissait de la région anatomique contenant une grande diversité de structures tissulaires associée à de faibles variations des valeurs d'ADC comme le montre la figure 4.3. D'un point de vue quantitatif, cette ROI était caractérisée par une hétérogénéité élevée (H_{index} entre 106 et 109 u.a.) mais avec la plus faible valeur moyenne d'ADC parmi les cinq régions (entre 0,684 et 0,690 $10^{-3}mm^2/s$) et un écart type faible. Ce résultat souligne la difficulté en général d'étudier des paramètres quantifiant des hétérogénéités sur des zones de faible contraste d'intensité pouvant limiter la stabilité du processus de clustering.

Concernant la répétabilité du H_{index} avec l'analyse de l'ICC, c'est à dire en considérant la variabilité supplémentaire inter-individus, nos résultats ont montré sur les cartes ADC une bonne répétabilité pour les ROI #1, #3 et #4. Pour la ROI #4, nous avons relevé une grande variabilité intra-individu avec un EBA élevé mais qui était associée à une encore plus forte variabilité inter-individus avec des écarts types importants au sein de la population. La part de variance pouvant être attribuée à la mesure se retrouvait ainsi plus faible expliquant une bonne répétabilité au vu de la valeur de l'ICC malgré les variances importantes. A l'inverse pour la ROI #5, où il est apparu une répétabilité modérée, les deux variances étaient relativement faibles et l'ICC se retrouvait ainsi sous-estimée par l'évaluation au sein d'une population relativement homogène. Ce dernier point était attendu si l'on admet une variabilité de l'hétérogénéité de l'ADC entre les volontaires sains sur une aussi grande ROI. Enfin, la faible répétabilité observée dans la

ROI #2 s'explique probablement de la même manière que précédemment au regard de la distribution des valeurs d'intensités et qui conduisent à une variabilité des fractions volumiques.

La comparaison de la variabilité du H_{index} aux indices de textures est également intéressante. Tout d'abord, d'après les récentes recommandations (BRYNOLFSSON et al., 2017), l'intervalle de valeurs pour l'analyse de texture sur les cartes ADC était fixé à des seuils physiologiques entre 0 et 3,4 mm²/s. Cette valeur maximale a été observée dans nos données et était également reportés par PIERPAOLI et al. (1996) dans les régions des cartes ADC correspondantes au liquide céphalorachidien. Ensuite, il apparaît que le volume des ROI impacte d'avantage les indices de texture qui se trouvent globalement plus variables pour les ROI #1, #2 et #3 de plus faibles volumes. Cette tendance n'était pas observée pour le H_{index} . Le contraste et la corrélation présentaient globalement une variabilité très élevée et une répétabilité faible, à l'exception de la corrélation pour la ROI #1. Pour les autres indices de texture issus de la GLCM, à savoir l'homogénéité, l'énergie, l'entropie et la dissimilarité, la variabilité était de même ordre de grandeur que pour le H_{index} et il est difficile de dégager une tendance dans la comparaison avec le H_{index} pour la répétabilité au travers des ICC. L'entropie et l'énergie démontraient tout de même une variabilité et une répétabilité similaire à celle du H_{index} et constituent ainsi de véritables challengers du H_{index} . Ce dernier tire alors son bénéfice au regard de son interprétabilité, en particulier en comparaison avec celle de l'énergie qui reste un concept difficile à transposer d'un point de vue médicale. Dans le cadre de l'entropie, celui-ci rend compte du désordre dont on comprend facilement l'interprétation en radiologie. Cet indice est d'ailleurs souvent proposé par différentes études comme potentiel biomarqueur d'imagerie sur les cartes ADC en le corrélant au score de Gleason pour la prostate (BAEK et al., 2020) ou bien encore comme biomarqueur pronostic de la rechute pour les cancers cervicaux (LUCIA et al., 2018). Ainsi, l'entropie apparaît comme un indice complémentaire au H_{index} dans son approche de quantification de l'hétérogénéité. Par ailleurs, nos résultats s'accordent globalement avec ceux de RODRIGUES et al. (2021) qui montrent sur des cartes ADC tumorales de la sphère ORL qu'aucun indice de la GLCM n'était fortement répétable, avec un ICC > 0,9. Dans le cas de notre étude, on peut noter que cette forte répétabilité est observée pour un unique indice, la corrélation (ICC=0,91) mais qui en même temps dispose d'un EBA moyen 121,6%. Enfin, rappelons la forte dépendance des analyses de texture aux paramètres de calcul employés et notamment aux nombres d'intensités de gris qu'il conviendrait d'étudier plus en détail lors de futurs travaux. Dans ce sens MERISAARI et al. ; SCHWIER et al. ; TRAVERSO et al. (2020, 2019, 2019) s'accordent sur la forte dépendance des indices aux processus de prétraitement des images et concluent que leur répétabilité varie grandement mais reste faible à modérée quel que soit le paramétrage du nombre de valeurs d'intensités.

Concernant les autres cartes paramétriques, D, D* et f, les EBA du H_{index} étaient

relativement élevés à l'exception de la ROI #5. Les ICC démontraient pour toutes les ROI des indices faiblement ou modérément répétables et ne surpassait pas les performances des valeurs moyennes des cartes paramétriques, également relativement faibles, à l'exception des ROI #4 et #5 de la carte D*. Ces résultats étaient attendus au regard de la variabilité supérieure des paramètres de perfusion et de pseudo diffusion issus des séquences IVIM en comparaison aux valeurs de l'ADC (LEBIHAN et al., 2019). Dans le cas de la carte D, les résultats différaient de ceux relatifs aux carte ADC. Il est probable que la détermination de ces coefficients de diffusion à partir de 2 ou 12 valeurs de b soient à l'origine de cette différence. Notons par ailleurs que l'ADC présentait globalement des ICC supérieurs à D. Ces résultats dégradés pour D en comparaison de l'ADC pourraient également s'expliquer par la présence d'une quantité importantes d'artéfacts dans les cartes paramétriques, également présents dans les cartes D* et f, qui se traduisaient par la présence de zones de pixels d'intensités très faibles, quasi-nuls. Ces artéfacts diminuaient la valeur moyenne de diffusion et impactait potentiellement fortement le H_{index} . L'origine de ces artéfacts reste à déterminer.

De manière générale, nous pouvons relever que l'interprétation de la répétabilité au regard des valeurs de l'ICC n'est pas aisée en raison des différents modèles qui peuvent être employées et des valeurs seuils de comparaisons qui peuvent être adoptées.

Nous avons choisi pour notre étude le modèle ICC(3,1) d'après les recommandations de KOO et LI (2016) qui suggèrent qu'un modèle à effet mixte est le plus approprié pour des études test-retest. En effet, des mesures répétées ne peuvent pas être considérées comme des échantillons aléatoires (PORTNEY & WATKINS, 2009). Néanmoins, et malgré l'existence depuis 2016 des recommandations de Koo et Li, qui suggèrent de reporter systématiquement le modèle (« one-way », « two-way ») le type (« simple », « moyenne ») et l'accord (« absolue », « cohérent), il n'y a à notre connaissance, pour l'évaluation de la répétabilité de l'ADC, que deux études l'ayant reportées (MERISAARI et al., 2020 ; SCHWIER et al., 2019) et l'une d'entre elle propose un modèle ICC(1,1) (« oneway », aléatoire, simple) qui ne semble pas appropriée pour la comparaison test-retest.

Concernant les valeurs seuils, il n'existe à ce jour aucun consensus avec d'une part les recommandations sur lesquels nous nous sommes appuyées et d'autre part des seuils variables. Ces seuils peuvent être soit comparés à un ICC dit de référence, c'est le cas de SCHWIER et al. (2019) qui comparent l'ICC des indices de radiomique à l'ICC du volume qu'ils ont considéré comme la référence suite à une précédente étude, mais sont également parfois choisi de manière plus aléatoire. Par exemple, TRAVERSO et al. (2019) ont eu recours à un seuil binaire de 0,85 pour indiquer la répétabilité ou non des indices de radiomique dans le contexte des cartes ADC de cancers colorectaux. RODRIGUES et al. (2021) dans le cadre des modifications précoces induites par un traitement pour les carcinomes ORL estiment que la répétabilité est excellente pour un $ICC \geq 0,8$ pour les indices de texture et d'histogramme. MERISAARI et al. (2020) utilisent 0,75 comme

le seuil de haute répétabilité des indices de radiomique dans le cadre des cancers de la prostate. HAKULINEN et al. (2012), pour l'ADC cérébrale sur volontaires sains, indique qu'un $ICC \geq 0,8$ possède une excellente répétabilité et que dans l'intervalle $[0,6 \ 0,79]$ sa répétabilité est modérée, et enfin CHEN et al. (2021) indiquent seulement que dans l'intervalle $[0 \ 0,2]$ la répétabilité de l'ADC évaluée sur le pancréas sain est faible. Dans ce contexte, il nous a semblé approprié de se comparer d'une part aux recommandations et à l'ICC de la valeur moyenne de l'ADC qui reste aujourd'hui l'indice le plus étudié dans la littérature.

Notre étude présentait certaines limites. Tout d'abord, le nombre de clusters utilisé n'était pas le même pour tous les individus et contrairement à ce qui aurait pu être attendu intuitivement, les plus gros volumes n'étaient pas représentés par un plus grand nombre de clusters. Cela soulève la question de la pertinence de notre méthode objective de détermination du nombre de clusters dans le cas précis de l'application à des données d'IRM de diffusion cérébrale. De futures études devront s'intéresser plus largement au choix du nombre de clusters. Ensuite, l'étude a été effectuée sur une population de volontaires sains. Les ICC calculés sur une population saine qui tend à être homogène peut impliquer une sous-estimation de la répétabilité en raison d'une faible variance biologique (HAGIWARA et al., 2020). De plus, sur une population saine, les ROI ont été choisies afin d'obtenir des zones plus ou moins hétérogènes mais la démarche est limitée par une variation de l'hétérogénéité qui était finalement logiquement associée à un volume croissant. Dans cette optique, nous souhaiterions par la suite appliquer cette démarche d'une à un jeux de données représentant une pathologie comme celui proposé par FEDOROV et al. (2018) sur 15 patients et disponible publiquement sur la plateforme The Cancer Imaging Archive (TCIA). Nous souhaiterions également employé notre méthode dans le cadre de l'évaluation de la réponse thérapeutique où une modification de l'hétérogénéité pourrait être observée au cours du temps et ainsi valider la fiabilité de nos indices. Pour cette dernière, les limites de temps imposé par ce travail de thèse ne permettait pas la réalisation d'une étude clinique prospective comme nous avons pu le faire ici qui vise à déterminer l'existence d'une corrélation avec la survie du patient. En effet, les études de survie nécessitent de pouvoir suivre les patients sur un minimum de cinq années.

Enfin, il aurait également été possible d'étudier la variabilité et la reproductibilité sous d'autres angles comme par exemple le calcul du coefficient de répétabilité qui peut informer sur la différence minimale mesurable pour indiquer une différence mais qui n'aurait pas de sens dans le contexte de volontaires sains. Enfin, nous soutenons qu'une standardisation, là encore, des méthodes d'analyses statistiques de la répétabilité est nécessaire et que l'application des recommandations est urgente.

1.6 Conclusion

En conclusion, cette étude a souligné la difficulté méthodologique pour évaluer la variabilité et la répétabilité d'un biomarqueur de l'hétérogénéité d'imagerie tant par le choix de l'objet d'étude que par l'analyse mathématique à mettre en œuvre pour les quantifier. Notre étude s'est focalisée sur la détermination de la variabilité et de la répétabilité du H_{index} pour évaluer l'hétérogénéité de la diffusion de l'eau sur des cartes paramétriques d'ADC. Dans le cas de trois ROI sur cinq, ce paramètre était bien répétable au regard des recommandations pour les seuils des ICC. Pour les deux autres ROI, le processus de recalage et la stabilité du processus de clustering sembleraient être à l'origine d'une robustesse plus modérée. Ce dernier point devra être investigué par la suite dans un contexte clinique d'imagerie de tissus pathologiques impliquant des cas de figures bien différents de ceux rencontrés.

2 Application du H_{index} en DCE-IRM pour s'affranchir de la variabilité de l'AIF : Première étude rétrospective de faisabilité du H_{index} en cancérologie ORL

2.1 Enjeux et objectifs

Le cancer de la sphère oto-rhino-laryngée (ORL), représente le cinquième cancer le plus fréquent en France. Le carcinome épidermoïde (CE) (angl. Nasopharyngeal Carcinoma) représente l'histologie majoritaire de ces cancers. A l'heure actuelle, un des problèmes majeurs dans la gestion des patients atteints de CE réside dans la difficulté à évaluer le statut et la distribution spatiale de l'oxygénation ou de l'hypoxie des tumeurs primaires ou secondaires et donc dans la difficulté à individualiser les thérapies.

Dans la gestion des CE, l'évaluation précoce de la réponse au traitement par la détection des modifications fonctionnelles avec l'objectif de prédire les issues cliniques (MUI et al., 2021) fait partie des axes majoritaires de recherche. Parmi les modalités d'imageries non-invasives disponibles, l'IRM semble la plus puissante et versatile pour la gestion des CE (ZHENG et al., 2014; ZHENG et al., 2015). Comme cela a déjà été mentionné, au-delà du suivi des modifications morphologiques, l'IRM fonctionnelle quantitative permet de suivre l'efficacité d'un traitement par la mesure des modifications physiologiques du micro-environnement tumoral (HARRY et al., 2010; JACKSON et al., 2005).

Pour évaluer l'une des caractéristiques d'intérêt des CE, l'hypoxie et la microperfusion tumorale, il est possible de s'intéresser aux paramètres fonctionnels de la microcirculation à partir de la DCE-IRM (GAUSTAD et al., 2020) et au mouvement Brownien des molécules d'eau en IRM de diffusion à l'aide des séquences IVIM (DAPPA et al., 2017). Néanmoins, la DCE-IRM est aujourd'hui privilégiée à l'IRM de diffusion qui souffre d'une moins bonne résolution spatiale, d'un rapport signal sur bruit plus faible et d'une déformation concomitante de l'image autour du crâne et du nasopharynx.

Comme indiqué dans le chapitre précédent, la DCE-IRM permet de fournir des informations relatives à l'angiogenèse tumorale par la quantification de paramètres relatifs à la microvascularisation. Lorsque les paramètres sont évalués avant l'initiation d'un traitement, il a été démontré qu'ils permettaient : 1) de différencier les CE d'autres pathologies comme les lymphomes ou les hyperplasies lymphoïdes par l'analyse des courbes d'intensités (SONG et al., 2020); 2) d'établir le grade T de la classification TNM des CE avec des valeurs d'aires sous la courbe et de fraction volumique extracellulaire/extravasculaire plus élevées pour les CE aux stades T3-4 associés à un flux sanguin tumoral supérieur que les CE classés T1-2 (CAO et al., 2008). Au regard de l'intérêt de ces paramètres dans l'évaluation prétraitement des CE, de nombreuses études

se sont attelées à démontrer leur pouvoir pronostic et/ou prédictif par la mesure de leur changement au cours des traitements (QIN et al., 2019 ; ZHENG et al., 2015 ; ZHENG et al., 2017) et plus particulièrement en évaluant les modifications d'hétérogénéité intra-tumorales visant à quantifier l'aspect tortueux de leurs structures vasculaires (GONZALES, 2014).

Actuellement et malgré les promesses de la DCE-IRM et de ses paramètres associés, la limitation technique principale à son utilisation clinique pour le suivi de l'efficacité thérapeutique réside comme cela a déjà été évoqué dans le manque de standardisation.

Dans ce contexte, nous nous sommes intéressés à un des critères souffrant du manque de standardisation, la cinétique de la concentration plasmatique, l'AIF, qui conditionne les valeurs des paramètres extraits en DCE-IRM en intervenant dans le modèle mathématique compartimental. La mesure de cette cinétique est en effet effectuée dans des espaces intravasculaires artériels et sa mesure n'est pas toujours triviale pour l'utilisateur avec comme source de variabilité la sélection d'une ROI souvent de très petite dimension. Pour participer à la standardisation de la procédure, des logiciels de traitement et d'analyse d'image permettent en plus du mode manuel habituel, de sélectionner l'AIF de manière automatique. Dans ce contexte, nous avons cherché à étudier de manière rétrospective, la variabilité des biomarqueurs d'hétérogénéité mesurés à partir des cartes paramétriques déterminées selon des AIF dites automatiques versus manuels. Ainsi, l'étude suivante, menée à l'institut Gustave Roussy et au SHFJ dans l'unité BioMaps vise à démontrer que l'utilisation du H_{index} comme biomarqueur de l'hétérogénéité permettrait de s'affranchir des difficultés de mesure de l'AIF et ainsi de simplifier les processus de standardisations. La méthode de cette étude va être à présent précisée.

2.2 Protocole d'acquisition clinique

2.2.1 Description du protocole clinique

Les données collectées pour la réalisation de cette étude s'inscrivent dans le cadre de deux protocoles de recherche clinique loi Jardé 3 : HNNS-PETMR, ID-RCB No :2018-A00517-48 et HNR-PETMR, ID-RCB No :2018-A00182-53. Ces protocoles visaient une étude commune sur l'intérêt potentiel de la TEP-IRM pour le cancer ORL avec d'une part pour aider au bilan ganglionnaire initial et d'autre part, la recherche de récurrence. Le système d'imagerie impliqué dans cette étude est la même que celle introduite dans le protocole MRIRM présenté précédemment, sur le site du SHFJ. Nous ne nous intéressons au cours de ce travail qu'à la composante IRM des données et plus particulièrement à celles obtenues par DCE-IRM.

Au cours de ce protocole, 34 patients ont été inclus entre septembre 2018 et juin 2020.

Les 34 patients disposaient d'une preuve histologique d'un CE. Aucun ne présentait de contre-indication à la réalisation d'images TEP ou IRM. Tous les patients avaient fourni leur consentement éclairé à la réalisation de l'examen selon la réglementation en vigueur. Les critères d'inclusions comprenaient notamment : homme ou femme adulte, indication de chirurgie ganglionnaire. Les critères de non-inclusion comprenaient notamment une contre-indication à l'IRM. Aucun traitement n'avait été initié pour l'ensemble des patients.

2.2.2 Descriptif technique du protocole d'acquisition

La DCE-IRM de la sphère ORL était une des composantes du protocole complet d'acquisition multimodale incluant plusieurs protocoles d'imagerie avec l'IRM, la TEP ORL, l'IRM Poumon, l'IRM et la TEP corps entier. L'examen DCE-IRM était réalisé avec une injection de 0,2 ml/kg d'agent de contraste au Gadolinium (DOTAREM, Guerbet, France) avec un retard de 25 sec par rapport au début de l'acquisition et un débit de 2 mL/sec.

La séquence DCE-IRM (Axiale T1 LAVA) était une séquence écho de gradient en 3 dimensions de type SPGR (SPoiled Gradient-Recalled, pondérée T1). Les caractéristiques de la séquence étaient les suivantes : plan de coupe axial, TE = 1,4 ms ; TR = 5,2 ms ; Angle de bascule = 30° ; Champ de vue = 256x256x40 mm³ ; Matrice = 160x160x8 voxels ; résolution temporelle = 3,35 sec ; durée d'acquisition 6 min. Une bande de saturation d'épaisseur 6 cm était positionnée parallèlement au champ de vue, immédiatement en dessous de ce dernier, afin de compenser le phénomène d'entrée de coupe.

2.3 Protocole de traitement et d'analyse des données DCE-IRM

2.3.1 Génération des cartes paramétriques et extraction des indices

A partir des images de DCE-IRM, quatre cartes étaient générées. Chacune de ces cartes correspondait à un Paramètre Pharmacocinétique (PPC) : la carte de la constante de transfert K^{trans} (min⁻¹), relative à la perfusion et à la perméabilité capillaire ; la carte K_{ep} (min⁻¹), correspondant au taux de reflux ; et enfin les cartes v_e (%) et v_p (%) représentant respectivement le volume extravasculaire/extracellulaire ou le volume plasmatique par unité volumique de tissus. Pour générer ces cartes, il est nécessaire de mesurer l'AIF pour chaque série d'images. A l'aide du logiciel Olea Sphère (v3.0.18, Olea Medical Solutions, France) utilisé dans cette étude et pour le calcul des cartes paramétriques, et du modèle mathématique de Tofts étendu (TOFTS et al., 1999), différentes options sont offertes concernant la sélection de l'AIF et le prétraitement des images. Elles vont être précisées ici.

Comme déjà dit, la sélection de voxels intra-artériels contribuant à l'AIF peut

se faire de deux manières : manuelle ou automatique. La méthode manuelle requiert l'intervention d'un radiologue et consistait à sélectionner quatre voxels à main levée dans l'artère carotide interne. A l'inverse, la méthode automatique d'Olea est basée sur la méthode décrite par MOURIDSEN et al. (2006) et consiste à sélectionner sans intervention de l'utilisateur 50 voxels dans la totalité du volume imagé pour obtenir une AIF moyenne. Le choix de l'AIF, partie intégrante des modélisations pharmacocinétiques, influence nécessairement la valeur des PPC et a été précédemment étudiée (ZIAYEE et al., 2018).

Outre le choix de l'AIF, différentes étapes de prétraitement des images peuvent être effectués avant de générer les cartes paramétriques comme le filtrage spatial et la correction de mouvement. Nous nous intéressons particulièrement à cette dernière. La DCE-IRM étant une acquisition dynamique, des mouvements involontaires ou physiologiques du patient au cours d'une acquisition continue sur plusieurs minutes sont inévitables. Dans la sphère ORL, les mouvements physiologiques comprennent des mouvements musculaires notamment de la langue. Les mouvements involontaires de la tête peuvent être compensés par une opération de recalage rigide. Lors d'analyses préliminaires, un effet de cette correction de mouvement sur l'AIF a en effet été observé.

Dans cette étude nous avons cherché à déterminer la variabilité du H_{index} , et de manière secondaire celles des PPC, selon la façon d'obtenir l'AIF. Le logiciel Olea permet l'obtention de l'AIF en considérant quatre modes possibles :

- Sélection automatique de l'AIF – Avec correction de mouvement : mode Auto-Avec
- Sélection automatique de l'AIF – Sans correction de mouvement : mode Auto-Sans
- Sélection manuelle de l'AIF – Avec correction de mouvement : mode Manuel-Avec
- Sélection manuelle de l'AIF – Sans correction de mouvement : mode Manuel-Sans

Ces quatre AIF ont été appliquées pour extraire le H_{index} et les PPC (K^{trans} , K_{ep} , v_e , v_p) dans deux ROI différentes : la première région correspondait à la tumeur et la seconde à une région musculaire. Elles ont été définies manuellement sur les images DCE-IRM assistées d'une visualisation anatomique avec les images pondérées T2. Cette segmentation, a été réalisée par F. Bidaut, un radiologue expert avec vingt ans d'expérience. La figure 4.11 illustre sur une image pondérée T2 un exemple de localisation de ces ROI et est accompagnée de la figure 4.12 des cartes paramétriques correspondantes sur la même coupe. Dans le cadre de la détermination du H_{index} , le nombre de clusters optimal était calculé pour chaque ROI pour l'ensemble de la population. L'algorithme des K-moyennes était employé pour le clustering et une connectivité de 18 pour l'analyse en composantes connectés permettant d'obtenir le nombre de patches, une des métriques composant le H_{index} . Enfin, les indices de textures issus de la GLCM, déjà discutés, ont également été extraits. Aucune recommandation

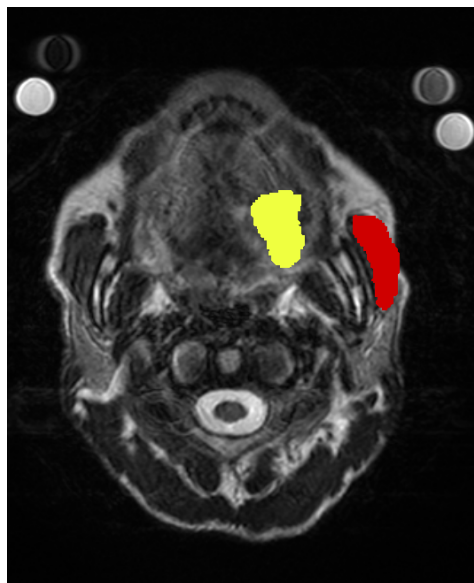


FIGURE 4.11 – Coupe Axiale T2 représentant les ROI muscle (rouge) et tumeur (jaune).

n'a été établie à notre connaissance concernant la radiomique dans le contexte de la DCE-IRM en ORL. Ainsi, les recommandations de BRYNOLFSSON et al. (2017) ont été suivis, à savoir : utiliser des images de mêmes résolutions, l'intervalle des intensités était fixé manuellement pour refléter une information physiologique pertinente des cartes, le nombre de niveaux de gris était de 256 afin d'obtenir la meilleure résolution possible. Les intervalles de valeurs observés ainsi retenues pour les cartes K^{trans} , K_{ep} , v_e , v_p étaient respectivement $[0-0,9] \text{ min}^{-1}$ $[0-7.5] \text{ min}^{-1}$ $[0-0,6] \text{ u.a.}$ $[0-0,08] \text{ u.a.}$

2.3.2 Analyse des données

Toutes les analyses des données ont été réalisées avec R sous RStudio (version 4.1.1, R Core Team, 2021). L'étude a porté sur deux volets distincts : le premier pour étudier l'influence des quatre modes de mesure de l'AIF sur l'AIF elle-même, et le second pour déterminer l'influence de ces modes sur les valeurs extraites de H_{index} et des quatre PPC que sont K^{trans} , K_{ep} , v_e et v_p .

Dans le cadre du premier volet, le critère retenu pour étudier la variabilité de l'AIF selon les modes de mesure a été choisi comme étant son amplitude maximale, exprimé en mM. Ainsi, nous avons cherché à établir si ces amplitudes variaient significativement en fonction des modes, comparés deux à deux :

1. Auto-Avec versus Manuel-Avec
2. Auto-Sans versus Manuel - Sans
3. Auto-Avec versus Auto-Sans
4. Manuel-Avec versus Manuel-Sans

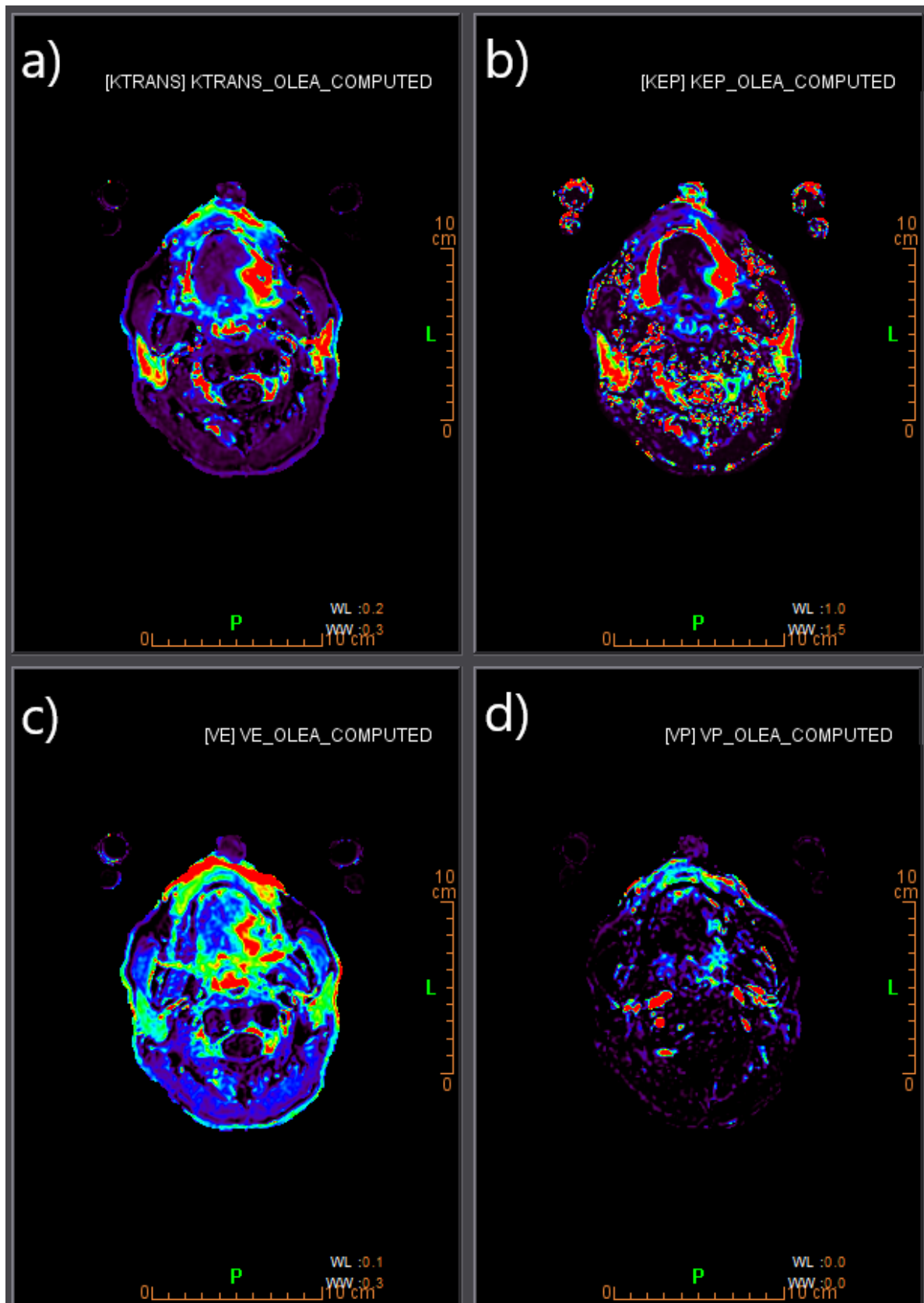


FIGURE 4.12 – Coupes axiales des cartes paramétriques (a) K^{trans} (b) K_{ep} (c) v_e (d) v_p . La coupe axiale représentée est la même que pour l'image T2 figure 4.11

Pour chacun de ces modes, des tests de Wilcoxon appariés ont été réalisés pour déterminer si la différence observée était significative entre les amplitudes des AIF sur les 34 patients. Pour illustrer les différences possibles des amplitudes des AIF selon les quatre modes, des graphiques de Bland-Altman ont été réalisés. Ces graphiques ont été associés à une représentation des distributions cumulées sur les 34 patients, ce qui a pour avantage d'illustrer cette fois-ci sur un même graphique les distributions des amplitudes maximales de l'AIF pour les quatre modes. La construction cumulative de cette représentation permet de mettre en évidence la valeur médiane correspondant à la valeur maximale scindant la population en deux parties égales. Ainsi cette représentation graphique des distributions cumulées accède visuellement à la valeur médiane des AIF et permet d'observer le biais présent entre chacun des modes de génération de l'AIF.

Le second volet de l'analyse des données a porté sur l'étude de l'influence des quatre modes de génération de l'AIF sur la détermination du H_{index} et des PPC moyens extraits. Pour cela un test statistique de Kruskal-Wallis et un ensemble de tests de Wilcoxon appariés ont été réalisés sur les indices selon les comparaisons deux à deux précédemment définies. Les tests de Wilcoxon et de Kruskal-Wallis ont été employées en raison de la distribution non-normale des données après que cette hypothèse de non-normalité ait été testée. Une valeur de $p \leq 0,05$ était considérée comme démontrant une différence statistique significative entre les valeurs. De plus, et comme pour le premier volet de l'analyse des données, des graphiques de Bland-Altman et des distributions cumulées ont également été réalisés pour illustrer les différences de distribution du H_{index} et des valeurs moyennes des PPC en fonction des modes de mesure de l'AIF.

Pour compléter l'approche statistique de cette étude, nous avons également déterminé l'ICC comme paramètre de stabilité sur les 34 patients. Plus précisément, il mesure la part de variance qui peut être attribuée à la méthode de mesure de l'AIF par rapport à la variance inter-individus. Le modèle d'ICC employée était le modèle ICC(3,1) (cf section 1.3.4). La répétabilité du H_{index} quantifiant l'hétérogénéité des PPC a également été comparée à celle des indices de textures extraits dans les ROI. Il s'agit des indices extraits de la GLCM déjà discutés dans le chapitre 3. Les ICC déterminés ont été analysés et interprétés selon les conventions déjà mentionnées dans la première partie de ce chapitre.

2.4 Résultats

2.4.1 Patients et régions d'intérêts

Les caractéristiques générales des données patients et les volumes des ROI sont fournies dans la table 4.10. Sur l'ensemble des 34 patients, les tumeurs n'ont pas pu être identifiées pour 3 patients en DCE-IRM car elles étaient situées en dehors du champ

Caractéristiques	Valeurs
Patients (n=34)	
<i>Sexe</i>	
Femmes	11 (32%)
Hommes	23 (68%)
Age	64.6 ± 12.8 [24-91]
Régions d'intérêt	
Volume (cm^3)	
Tumeur (n=31)	10.5 ± 19.1 [0.7 - 105]
Muscle (n=34)	5.8 ± 3.0 [1.4 - 13.4]

TABLE 4.10 – Caractéristiques générales des données patients et des ROI tumeurs et muscles.

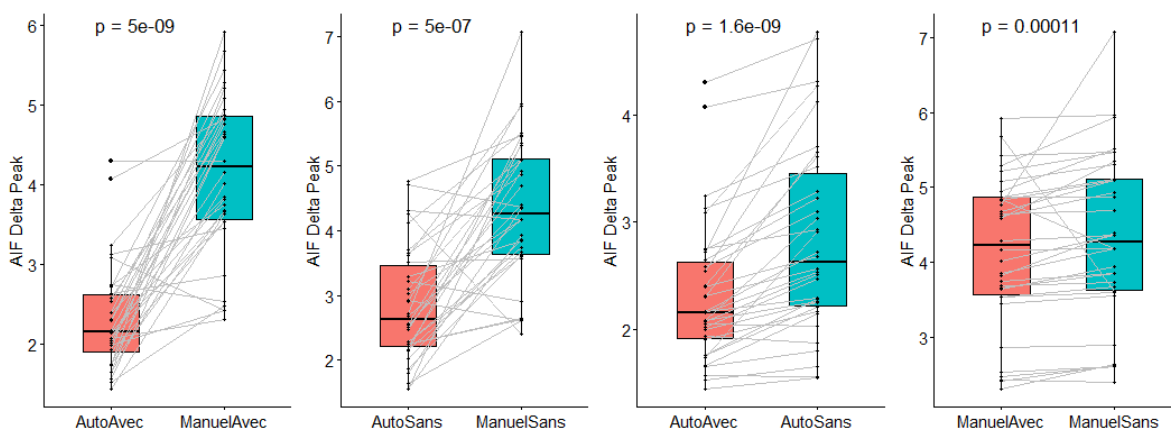


FIGURE 4.13 – Évaluation des différences statistiques sur l'amplitude maximale de l'AIF (mM) entre les modes de mesure deux à deux (AutoAvec ; ManuelAvec), (AutoSans ; ManuelSans), (AutoAvec ; AutoSans), (ManuelAvec ; ManuelSans). Les différences statistiques évaluées par des tests de Wilcoxon appariés sont indiquées par leur valeur de p.

de vue. Par ailleurs, il apparaît une grande disparité dans les volumes tumoraux avec un facteur 150 entre les volumes minimum et maximum. Les volumes des ROI définies dans du tissu musculaire occupent une plage de valeur plus faible avec un maximum de 15 cm^3 , le facteur entre les valeurs minimum et maximum pour le muscle étaient alors de 10.

2.4.2 Étude de l'influence des modes de mesure sur l'AIF

Les amplitudes maximales pour les courbes d'AIF étaient significativement différentes entre elles selon les modes de mesures automatiques ou manuels, avec ou sans corrections de mouvements. Ces différences sont illustrées sur la figure 4.13, où les quatre modes sont comparés, les tests de Wilcoxon appariés délivraient tous une valeur de $p \leq 1,10e-4$.

En complément de ces tests, et à l'aide de la figure 4.14, il apparaît que les valeurs

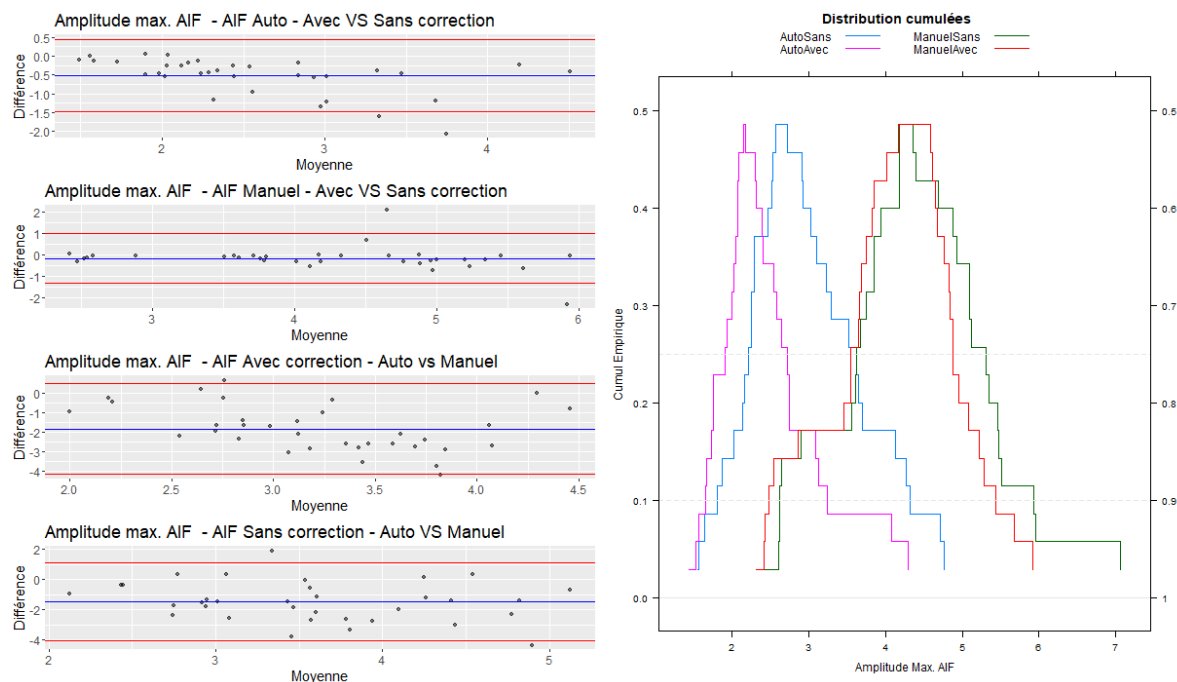


FIGURE 4.14 – (Gauche) Graphes de Bland-Altman comparant la différence et la moyenne des amplitudes maximales de l'AIF pour les quatre couples de modes de mesures de l'AIF. Les lignes bleus représentent le biais moyen entre les modes et les lignes rouges les LOA à 95%. (Droite) Distribution cumulées pour les quatre modes. L'amplitude de l'AIF est exprimée en mM.

et les intervalles de confiances des amplitudes maximums d'AIF étaient plus élevés pour les AIF extraites manuellement par rapport aux automatiques d'un facteur presque 2. Ces valeurs étaient pour le mode Auto-Avec=2,31 mM, Auto-Sans=2,84 mM, Manuel-Avec=4,13 mM, Manuel-Sans=4,32 mM. Pour les AIF dites manuelles, les médianes correspondantes au pic de la distribution cumulée, apparaissent quasi superposables avec, toutefois, une différence significative dans la distribution des valeurs sur les 34 patients selon les modes avec et sans correction de mouvement ($p=1,1e-4$). Dans le cas des AIF obtenues automatiquement, on observe à la fois une diminution significative de la médiane avec la correction de mouvement ($p=1,6e-9$), et un rétrécissement de la largeur de la distribution cumulée. En d'autres termes, le mode de génération de l'AIF automatique versus manuel entraîne une diminution globale de l'amplitude de l'AIF et la correction de mouvement à tendance à réduire sa dispersion inter-patients. Cette observation était également perçue sur les graphs de Bland-Altman, mais de manière moins évidente.

2.4.3 Optimisation du nombre de clusters

Préalable à la détermination du H_{index} , l'optimisation du nombre de clusters a été réalisée indépendamment sur l'ensemble des 34 patients pour chacune des ROI. Il

Méthode Correction de mouvement		Auto		Manuelle	
		Sans	Avec	Sans	Avec
H_{index} Tumeur	K^{trans}	53.2±91.7	52.3±87.6	57.6±104.3	55.5±92.5
	v_e	55.0±97.1	55.5±94.5	57.5±102.9	58.0±103.3
	v_p	46.3±47.4	45.1±44.7	49.5±59.7	42.6±48.9
	K_{ep}	48.9±76.6	48.9±80.4	48.3±79.5	47.3±69.0
Muscle	K^{trans}	28.3 ±33.7	29.3±34.8	25.5±33.4	27.3±33.8
	v_e	26.0±33.3	28.3±32.4	27.2±34.6	27.2±33.5
	v_p	15.1±32.8	16.0±33.1	12.6±30.8	13.9±31.2
	K_{ep}	28.0±31.2	27.0±30.4	28.1±32.2	28.7±34.0
Valeur Moyenne					
Tumeur	K^{trans}	0.49±0.131	0.300±0.168	0.163±0.077	0.168±0.077
	v_e	0.223±0.088	0.243±0.085	0.185±0.069	0.184±0.064
	v_p	0.018±0.010	0.023±0.012	0.012±0.008	0.013±0.008
	K_{ep}	1.303±0.78	1.43±0.97	0.99±0.39	1.01±0.34
Muscle	K^{trans}	0.076±0.059	0.088±0.072	0.049±0.029	0.050±0.03
	v_e	0.105±0.040	0.113±0.035	0.084±0.026	0.082±0.022
	v_p	0.0022±0.0030	0.0026±0.0037	0.0011±0.0017	0.0011±0.0018
	K_{ep}	0.759±0.411	0.805±0.497	0.625±0.275	0.623±0.286

TABLE 4.11 – Valeurs du H_{index} et valeurs moyennes des PPC et écarts types pour les deux ROI tumeur et muscle et pour chacun des 4 modes d'obtention de l'AIF. Les valeurs moyennes de K^{trans} et K_{ep} sont exprimés en min^{-1} , toutes les autres en u.a.

en ressort que dans plus de 90% des cas la valeur de K=3 clusters ressortait. Seules quelques cartes évaluées sur la tumeur ou les muscles, notamment les plus gros volumes, proposaient une valeur de K=4. Le choix d'utiliser le même nombre de clusters, K=3, pour l'ensemble des cartes paramétriques et des ROI a alors été fait.

2.4.4 Etude de l'influence des modes de mesure de l'AIF sur le H_{index} et les PPC

Les valeurs des H_{index} déterminés pour les 4 cartes paramétriques et les valeurs moyenne des PPC associées sont résumées dans la table 4.11 pour chacune des ROI selon les 4 modes de mesure de l'AIF. Les analyses statistiques de Kruskal-Wallis conduites pour le H_{index} (table 4.12) et la texture (table 4.13) sur l'ensemble des 4 modes ont montré qu'il n'y avait aucune différence significative observée ($p>0,05$). A l'inverse, pour les valeurs moyennes des PPC, la totalité des mesures étaient significativement différentes, sauf pour le K_{ep} du muscle ($p=0,16$).

Pour une lecture plus fine de l'influence des modes de mesure de l'AIF, la différence statistique a été évaluée en les comparant deux à deux. Ces résultats sont également reportés sur la table 4.12 pour le H_{index} et les valeurs moyennes des PPC. Il en ressort que toutes les moyennes des PPC étaient significativement différentes pour les trois

ROI et Image	4 modes Kruskal-Wallis (p)	Modes d'obtention de l'AIF Wilcoxon entre les modes deux à deux			
		AA-MA	AS-MS	AA-AS	MA-MS
Tumeur					
K^{trans}					
H_{index}	0.99	0.34	0.75	0.87	0.29
Moyenne	1.9e-5 **	1.9e-6 **	2.2e-5 **	8.4e-7 **	0.08
K_{ep}					
H_{index}	0.93	0.96	0.76	0.84	0.59
Moyenne	0.046 *	5.3e-6 **	2.2e-5 **	0.046 *	0.16
v_e					
H_{index}	0.99	0.3	0.74	0.24	0.84
Moyenne	0.0029 **	1.5e-4 **	0.005 **	0.006 *	0.32
v_p					
H_{index}	0.99	0.44	0.92	0.78	0.26
Moyenne	4.6e-5 **	1.0e-7 **	2.2e-6 **	9.6e-5 **	0.03 *
Muscle					
K^{trans}					
H_{index}	0.82	0.017 *	0.016 *	0.95	0.21
Moyenne	0.005 **	2.9e-8 **	2.2e-6 **	1e-6 **	0.13
K_{ep}					
H_{index}	0.99	0.10	0.45	0.41	0.9
Moyenne	0.16	5.1 e-6 **	6.5e-6 **	0.16	0.5
v_e					
H_{index}	0.9	0.95	0.16	0.1	0.47
Moyenne	1.9e-4 **	3.2e-6 **	2.5e-4 **	0.015 *	0.91
v_p					
H_{index}	0.39	0.26	9e-4 **	0.96	0.059
Moyenne	2.5e-4 **	1.3e-8 **	5.8e-7 **	0.12	0.68

TABLE 4.12 – Valeurs de p correspondantes aux évaluations statistiques des H_{index} et des paramètres moyens extraits à partir des cartes paramétriques avec : les tests de Kruskal-Wallis comparant les 4 modes de mesure de l'AIF et les tests de Wilcoxon appariés comparant les modes deux à deux. Les comparaisons sont effectués pour les deux ROI : tumeur et muscle. AA : Auto-Avec ; MA : Manuel-Avec ; AS : Auto-Sans ; MS : Manuel-Sans. * indique des valeurs de $p \leq 0,05$ et ** indique $p \leq 0,005$.

Tumeur	K^{trans}	K_{ep}	v_e	v_p
Homogénéité	0.79	0.82	0.99	0.86
Énergie	0.96	0.92	0.99	0.90
Contraste	0.99	0.83	0.97	0.95
Corrélation	0.96	0.98	0.96	0.71
Entropie	0.99	0.98	0.99	0.95
Dissimilarité	0.99	0.87	0.98	0.86
Muscle				
Homogénéité	0.75	0.86	0.40	0.12
Énergie	0.75	0.92	0.43	0.12
Contraste	0.96	0.57	0.74	0.63
Corrélation	0.56	0.98	0.59	0.61
Entropie	0.85	0.88	0.73	0.11
Dissimilarité	0.89	0.66	0.60	0.22

TABLE 4.13 – Valeurs de p pour les tests de Kruskal-Wallis entre les 4 modes de mesure de l'AIF pour les indices de texture issus de la GLCM évalués sur la tumeur et le muscle à partir des différentes cartes paramétriques.

comparaisons impliquant un mode de génération automatique de l'AIF, avec ou sans correction de mouvement. En revanche, la comparaison des modes manuels d'AIF, avec versus sans correction de mouvement, ne présentait pas de différence significative exceptée pour la carte v_p de la tumeur ($p=0,03$). Concernant, le H_{index} , lui ne montrait pas de différence significative dans la tumeur, quels que soit les modes comparés. En revanche dans le muscle, il y avait des différences significatives pour le H_{index} extrait des cartes K^{trans} pour les 2 comparaisons automatiques vs manuelles, avec ou sans correction de mouvement et également pour le H_{index} extrait de la carte v_p en comparant les modes automatique vs manuel, sans correction de mouvement ($p=9e-4$).

Enfin, les graphes de Bland-Altman et de distributions cumulées présentés sur la figure 4.15 illustrent les distributions pour le H_{index} qui tendent à être superposables quel que soit le mode de mesure de l'AIF. A l'inverse, les distributions cumulées des valeurs moyennes des K^{trans} présentaient des comportements différents, tout comme pour les autres paramètres (résultats non présentés). Ces comportements sont similaires à ceux de l'amplitude de l'AIF décrits dans la partie précédente. Dans le cas des AIF obtenues automatiquement, on observe une différence de la médiane qui a tendance à augmenter avec la correction de mouvement. Lorsque l'AIF est déterminée manuellement, les deux distributions cumulées avec et sans correction de mouvement semblent superposables et de médianes inférieures à celles des modes automatiques. En d'autres termes, le mode de mesure de l'AIF automatique entraîne cette fois-ci une évaluation supérieure des valeurs des PPC par rapport au mode manuel et la correction de mouvement ne semble pas modifier la dispersion des valeurs PPC inter-patients bien qu'elle les surestime en

mode automatique.

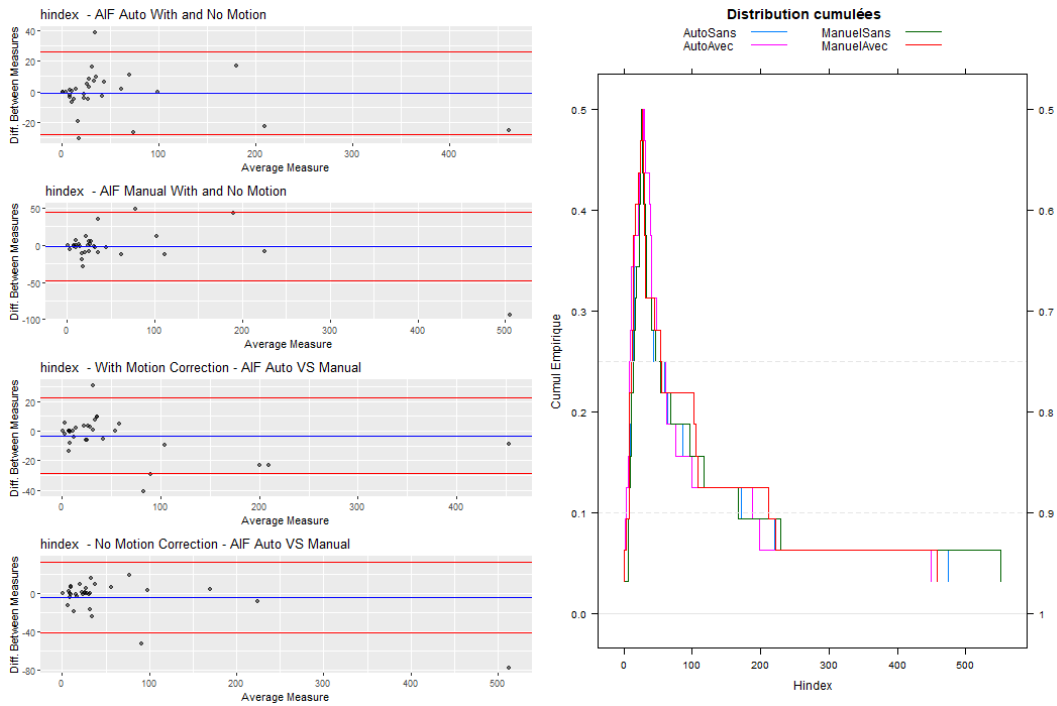
2.4.5 Etude de la stabilité du H_{index} , des indices de texture et des PPC moyens selon les modes d'AIF

Pour étudier la stabilité des différents paramètres décrivant l'hétérogénéité des cartes incluant le H_{index} et les indices de texture, les valeurs d'ICC ont été déterminées. Les figures 4.16 et 4.17 illustrent les ICC ainsi calculées pour les régions tumorales et musculaires, respectivement pour les cartes v_e , v_p , et K^{trans} , K_{ep} . Le critère d'ICC montre que le H_{index} est globalement excellentement stable quel que soit le mode de mesure de l'AIF. En effet, son ICC est toujours supérieur à 0,9 à l'exception de la ROI tumorale de la carte v_p où ICC=0,88. Concernant les valeurs moyennes PPC, les ICC indiquent des mesures globalement très peu ou modérément stables à l'exception des tissus musculaires associés aux cartes v_p et K_{ep} pour lesquelles il y avait une bonne stabilité avec un ICC compris entre 0,75 et 0,9. En ce qui concerne les indices de texture, les résultats étaient plus dispersés. L'entropie, la dissimilarité et la corrélation étaient stables ($0,75 < ICC < 0,90$) ou très stables ($ICC > 0,9$) quelle que soit la carte ou la ROI. Le contraste était stable ou très stable à l'exception de la carte K_{ep} du muscle pour laquelle la stabilité était modérée ($0,5 < ICC < 0,75$). L'énergie était très stable pour les cartes v_e toutes ROI confondues et pour les cartes K_{ep} , K^{trans} évaluées sur le muscle, stable pour le v_p muscle, modérément stable pour le K^{trans} tumeur et non stable pour K_{ep} et v_p tumeur ($ICC < 0,5$).

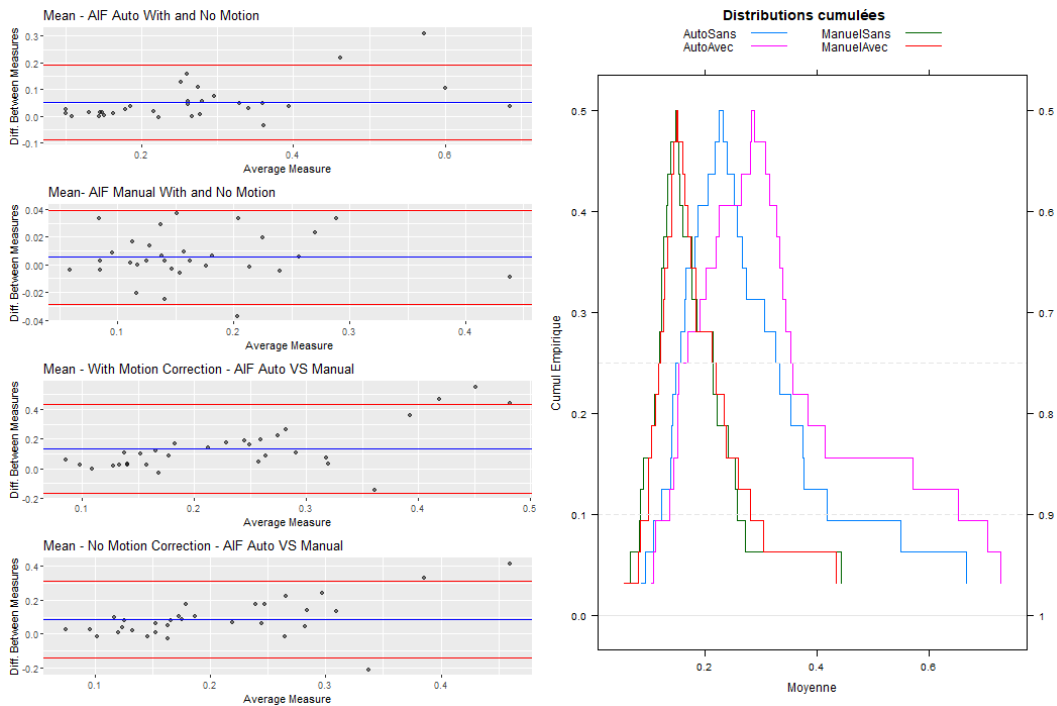
2.5 Discussion

L'objectif de cette étude était d'analyser l'influence des différents modes de mesure de l'AIF en DCE-IRM sur les valeurs de notre indice d'hétérogénéité dans le cadre clinique de l'imagerie du cancer ORL chez 34 patients. Le H_{index} a été extrait à partir des cartes paramétriques K^{trans} , v_e , v_p et K_{ep} dans deux régions d'intérêt, la tumeur et le muscle. L'AIF pouvait être obtenu de deux manières, automatiquement ou manuellement et une correction de mouvement pouvait être appliquée ou non. Les combinaisons de ces options représentaient les quatre modes de mesure de l'AIF.

Les analyses menées sur l'amplitude maximale de l'AIF, qui correspondait à la métrique retenue pour caractériser les variations de l'AIF, ont montré que ces modes de mesure influent de manière significative. L'amplitude des AIF déterminées manuellement étaient en moyenne 78% plus élevées que lorsqu'elles étaient mesurées automatiquement. La correction de mouvement influait également significativement sur l'amplitude de l'AIF. Dans le cas des AIF mesurées manuellement, la correction de mouvement diminuait légèrement la dispersion des amplitudes maximales relevées chez les 34 patients. En revanche, pour les AIF mesurées automatiquement, la correction

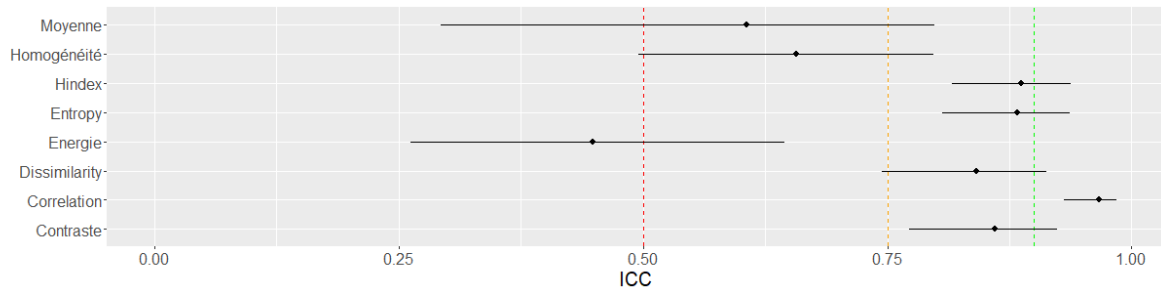


(a)

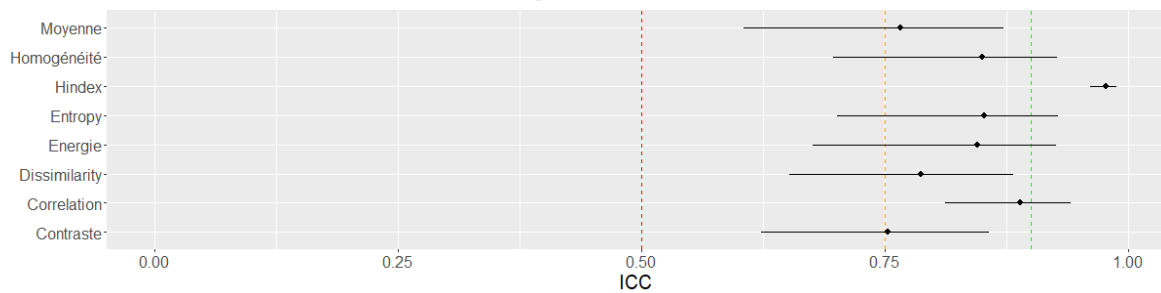


(b)

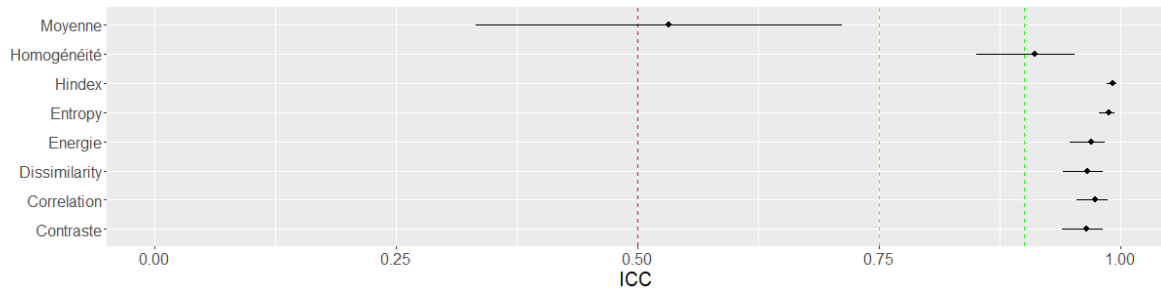
FIGURE 4.15 – (Gauche) Graphes Bland-Altman et (Droite) distribution des différences cumulées pour les cartes K^{trans} évaluées sur la tumeur d'après les quatre modes de mesure de l'AIF avec (a) le H_{index} et (b) la valeur moyenne (min^{-1}). Pour les graphes de Bland-Altman, les lignes bleus représente le biais moyen entre les modes et les lignes rouges les LOA à 95% (a) pour le H_{index} , (b) pour la valeur moyenne.



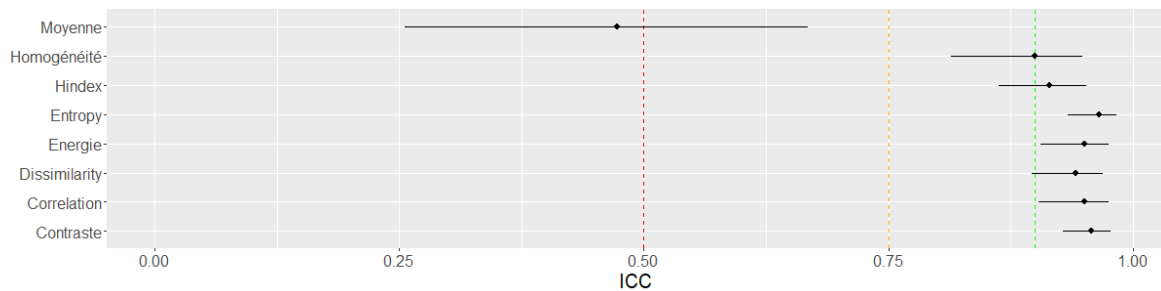
(a) v_p - Tumeur



(b) v_p - Muscle

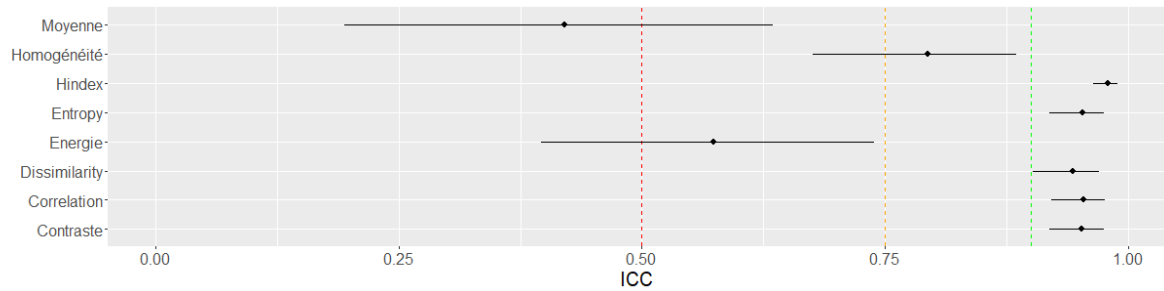


(c) v_e - Tumeur

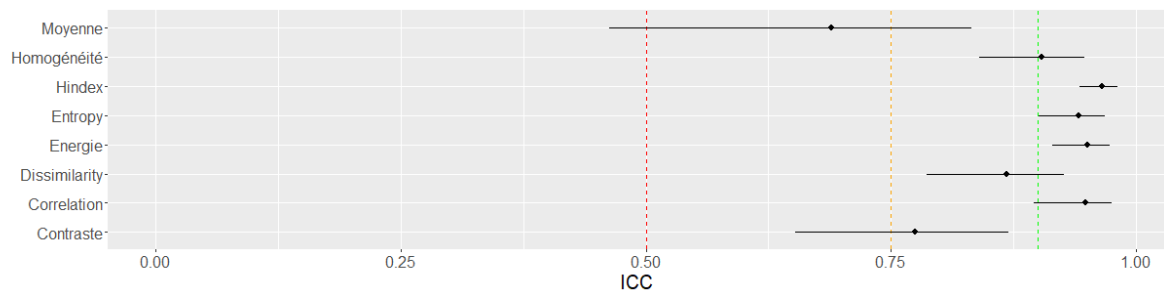


(d) v_e - Muscle

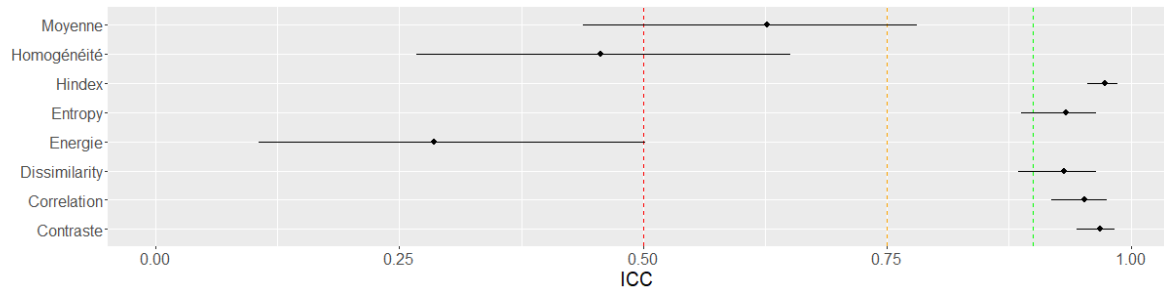
FIGURE 4.16 – Représentation graphique des ICC évalués pour le H_{index} , les indices de texture issus de la GLCM et la moyenne pour les cartes v_p et v_e et les ROI (a,c) tumeur et (b,d) muscle.



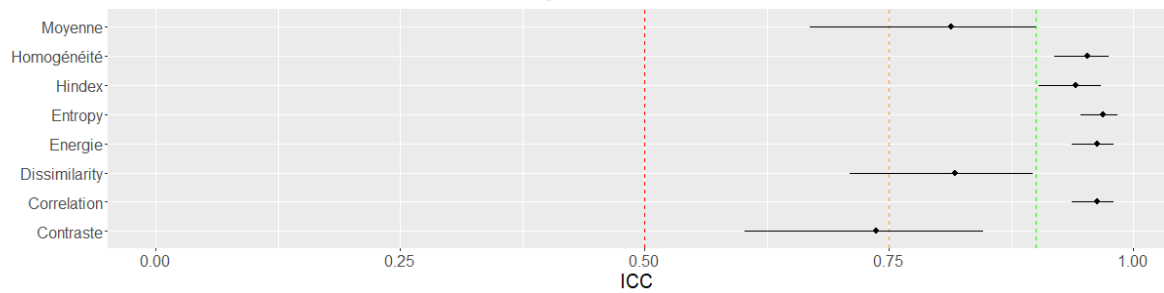
(a) K^{trans} - Tumeur



(b) K^{trans} - Muscle



(c) K_{ep} - Tumeur



(d) K_{ep} - Muscle

FIGURE 4.17 – Représentation graphique des ICC pour le H_{index} , les indices de texture issus de la GLCM et la moyenne pour les cartes K^{trans} et K_{ep} et les ROI (a,c) tumeur et (b,d) muscle.

de mouvement impliquait à la fois une diminution significative de 19% de la valeur de l'amplitude et une diminution de la dispersion des amplitudes ainsi relevées. La différence notable entre les mesures automatiques et manuelles s'explique par la méthode de sélection des voxels contribuant à l'AIF. La méthode automatique, décrite par MOURIDSEN et al. (2006), basée sur une approche de clustering, sélectionne 50 voxels dans la totalité du volume pour former l'AIF. L'inspection visuelle de ces voxels a montré qu'une portion d'entre eux ne se situaient pas au centre d'artères au calibre important et pouvaient ainsi être affectés par des effets de volume partiel induisant une sous-estimation des valeurs. Par contre, dans le cas de la sélection manuelle de l'AIF, un radiologue expert sélectionnait seulement 4 voxels au centre d'une artère carotide interne de calibre important. Ces 4 voxels étant situés sur la même coupe, il est ainsi possible de considérer que ces voxels représentent l'espace intravasculaire artériel avec un minimum de volume partiel. La correction de mouvement influait moins sur l'amplitude de l'AIF que pour la méthode automatique et la méthode manuelle semblait ainsi plus robuste à cette opération. Néanmoins, la méthode manuelle présente l'inconvénient d'être consommatrice de temps et opérateur dépendante avec ce que cela implique sur la reproductibilité et les besoins d'opérateurs entraînés spécifiquement. Une méthode semi-automatique, qui permettrait à l'utilisateur de restreindre la zone de recherche automatique de voxels pour l'AIF n'a pas été étudiée dans ce travail et devra faire l'objet d'une étude complémentaire. En effet, l'étude menée par ZIAYEE et al. (2018) employait une méthode semi-automatique basée sur la même méthode qu'au cours de notre étude pour la détermination automatique de l'AIF mais dans une zone restreinte. Leurs résultats démontraient ainsi des AIF semi-automatiques et manuels similaires. De plus, d'autres approches de détermination automatique de l'AIF pourraient être envisagées en considérant les résultats de CARROLL et al. (2003) et de SANZ-REQUENA et al. (2015) qui employaient des méthodes automatiques sensiblement différentes de celle implémentée dans Olea Sphere et obtenaient des résultats ne démontrant pas de différence significative entre les deux modes.

La valeur des PPC étant directement influencée par la mesure de l'AIF, avoir établi la variabilité de l'AIF selon leur mode de mesure nous permettait ainsi d'évaluer selon ces mêmes modes de mesures la variabilité des PPC, et de la comparer à celle du H_{index} et des indices de textures, obtenues à partir des cartes paramétriques correspondantes. Ainsi, les valeurs des PPC étaient significativement impactées par les modes de mesure de l'AIF. Quel que soit le tissu cible, les méthodes de mesure automatiques et manuelles de l'AIF produisaient des résultats significativement différents avec les méthodes automatiques surévaluant les PPC. Enfin, la correction de mouvement impliquait que les PPC différaient significativement mais de manière moins marquée qu'en fonction de la méthode de sélection de l'AIF. Aucune étude à notre connaissance ne s'était jusque-là intéressée à l'influence de la correction de mouvement. En ce qui concerne

le H_{index} , notons tout d'abord que dans le cadre de cette étude de la variabilité, le H_{index} uniquement était présenté. Les métriques qui le composent et les $H_{cluster}$ étaient évidemment calculés mais leur stabilité n'a pas été présentée ici au regard du manque d'intérêt dans un contexte autre que celui de l'évaluation du suivi thérapeutique. En effet, dans notre étude, la variabilité du H_{index} englobe la variabilité des $H_{cluster}$. A l'inverse, des PPC, le H_{index} n'était pas impacté par les modes d'obtention de l'AIF quel que soit le mode à l'exception des comparaisons automatiques vs manuels lorsque le H_{index} était évalué sur le muscle pour les cartes K^{trans} et v_p . Enfin, les indices de texture n'étaient en partie pas non plus impactés par le mode de mesure de l'AIF. Ces résultats sont cohérents avec les résultats obtenus par un autre membre de notre équipe, Wanxin Dong, dans le cadre de son stage de Master avec le radiologue F.Bidault. Lors de ces travaux, ils se sont intéressés au rapport des PPC tumeur sur muscle et ont démontré que ce rapport était stable au regard du mode de mesure de l'AIF. Dans ce sens, il semblerait que la mesure de l'AIF n'affecte que les valeurs des PPC de manière relative et ne modifie donc pas l'hétérogénéité des images.

Enfin, la stabilité du H_{index} a été comparée à celle des PPC moyens et à celles des indices de texture typiquement utilisés pour évaluer l'hétérogénéité d'images. Les calculs de l'ICC, ont tout d'abord permis de confirmer les résultats de la précédente analyse et ainsi de démontrer que les PPC moyens n'étaient globalement pas stable et que le H_{index} était stable. Plus précisément, l'ICC des PPC moyens était inférieure à 0,75 dans 6 cas sur 8. En ce qui concerne le H_{index} , il était extrêmement stable ($ICC > 0,9$) pour toutes les ROI et cartes paramétriques, et stable pour la carte v_e évaluée sur la tumeur ($ICC = 0,88$). Pour les indices de texture, ils étaient globalement tous stable ou très stables, à l'exception de l'énergie évaluée sur K^{trans} , K_{ep} et v_p sur la tumeur et du contraste pour la carte K_{ep} sur le muscle. Enfin, aucune autre étude à notre connaissance ne s'est à l'heure actuelle intéressée à la variabilité des indices de texture pour des cartes de DCE-IRM au regard des modes de mesure de l'AIF. Néanmoins, comme démontré par SCHWIER et al. (2019), la variabilité des indices de texture dépend de plusieurs paramètres et notamment des différents processus de préparation des images, ce qui n'est pas le cas du H_{index} .

2.6 Conclusion

En conclusion, nous avons démontré que notre indice de quantification de l'hétérogénéité était stable au regard des différents modes de mesure de l'AIF possibles. Ce résultat est particulièrement prometteur dans un contexte de standardisation des procédures d'acquisition et de traitement des données. Le H_{index} pourrait alors être un biomarqueur indépendant de la méthode de mesure de l'AIF, source de variabilité des mesures de PPC limitant en partie la validation multicentrique de la DCE-IRM

en cancérologie pour le suivi de l'efficacité thérapeutique (O'CONNOR et al., 2017). Les indices de texture sont également apparus comme stables mais dans une moindre mesure. Cette piste de l'application du H_{index} pour s'affranchir de la variabilité de l'AIF en DCE-IRM apparait attractive et nécessite d'être approfondie en particulier en abordant notamment la reproductibilité inter-machines sur différents sites.

3 Application du H_{index} dans un contexte de radiothérapie de tumeurs cérébrales : Evaluation du potentiel diagnostic pour la différenciation de récurrence locale et de radionécrose en IRM de contraste et de perfusion

3.1 Enjeux et objectifs

Les métastases cérébrales représentent les tumeurs cérébrales les plus fréquentes chez l'adulte soit 25% des lésions (ANDREWS et al., 2004 ; AOYAMA et al., 2006). Dans ce cas l'approche thérapeutique bien établie repose sur la radiochirurgie stéréotaxique (RCS) qui implique la délivrance de fortes doses de radiation focalisées sur la tumeur (MEHRABIAN et al., 2017). En comparaison à la radiothérapie complète du cerveau, la RCS limite les récurrences locales tumorales (RLT), risque majeure après le traitement initial, et permet une amélioration du temps de survie (ANDREWS et al., 2004 ; HOEFNAGELS et al., 2009). En revanche dans ce cas il existe un plus grand risque de radionécrose (RN) apparaissant typiquement plusieurs mois après la fin du traitement (GREENE-SCHLOESSER et al., 2012). Diagnostiquer et dissocier les RN des RLT est primordial puisque les options thérapeutiques diffèrent sensiblement. Les RN sont principalement traitées par un immuno-suppresseur, le dexaméthasone, ou par un anti-angiogénique, le bevacizumab dans les cas persistants (SHAW & BATES, 1984), tandis que les RLT sont traitées à l'aide du triptyque chirurgie, radiation, chimiothérapie (MEHRABIAN et al., 2017).

Dans l'optique de réduire le nombre d'interventions diagnostiques invasives pour les biopsies cérébrales, différentes modalités d'imagerie sont aujourd'hui explorées et visent à distinguer les RLT et les RN. Dans le paysage clinique actuel, l'imagerie TEP combinée à des acides aminés radiomarqués (^{18}F F-DOPA TEP, ^{18}F F-FET TEP) est aujourd'hui considérée la plus prometteuse avec une spécificité située entre 80 et 90 % (CECCON et al., 2016 ; CICONE et al., 2021 ; CICONE et al., 2015 ; ROMAGNA et al., 2016). Néanmoins, sa faible disponibilité et son aspect ionisant imposent d'explorer des pistes additionnelles comme l'imagerie IRM de contraste qui reste la modalité de choix pour l'évaluation post-traitement des patients atteints de tumeurs cérébrales. Notons

qu'à ce jour il n'existe pas de modalité gold-standard pour la différenciation des RLT et des RN.

Dans ce contexte, des radiologues de l'équipe se sont intéressés à la question (AMMARI S et al., 2021). A partir d'IRM multiparamétrique comprenant des séquences pondérées T1, de diffusion (DWI) et de contraste de susceptibilité dynamique (DSC), leurs résultats ont montré l'intérêt des cartes TRAM (Treatment Response Assessment Map) qui établissent la différence entre les images pondérées T1 3 min post-injection de chélate de Gd et la même acquisition 90 minutes plus tard. D'une part ces cartes TRAM se sont avérées efficaces pour guider le radiologue dans la différenciation entre les RN des RLT mais sont apparues encore meilleures que l'analyse qualitative du radiologue (radiologue seul) en terme de sensibilité et spécificité. De plus, en considérant les différences physiologiques entre les deux pathologies qui sont susceptibles de se traduire également à l'image, une approche radiomique avait complétée cette étude et présentée une efficacité intermédiaire. De nombreuses autres études se sont intéressées à la radiomique en IRM multiparamétrique pour l'évaluation des métastases cérébrales et notamment pour l'évaluation de la réponse thérapeutique et des récives métastatiques avec des résultats encourageants (cf review LOHMANN et al. (2020)) mais qui n'atteignaient pas ceux de l'imagerie TEP. Bien qu'au sein de notre équipe, les performances de la signature radiomique étaient inférieures à la TRAM, ces résultats étaient toutefois satisfaisants et prometteurs.

Notre étude menée dans ce contexte a été rétrospective et utilisait un ensemble d'images issues de différentes modalités IRM. Il s'agissait d'évaluer l'intérêt de notre indice d'hétérogénéité, le H_{index} , pour le diagnostic différentiel des RN et des RLT. L'avantage du H_{index} par rapport à l'étude radiomique précédente résidait en particulier dans son interprétabilité, largement abordée au cours des précédents chapitres, par rapport aux signatures radiomiques. De plus, si les cartes TRAM donnent de bons résultats, elles imposent une acquisition tardive après 90 minutes ce qui n'est pas toujours envisageable en routine clinique. Dans ce contexte, l'ambition de notre étude a été d'extraire le H_{index} uniquement à partir des images issues de la première séance d'acquisition post-injection d'agent de contraste, donc à partir d'images pondérées T1 et à partir de cartes paramétriques issues de la DSC-IRM.

3.2 Méthodologie du protocole clinique

3.2.1 Description du protocole clinique

Notre étude s'appuie sur les données provenant de l'étude monocentrique rétrospective menée par les radiologues de l'équipe à Gustave Roussy (Villejuif) entre Juin 2013 et Mai 2018. 77 patients symptomatiques dont la fin de la radiothérapie pour

lésion métastatique cérébrale solitaire se situait entre 7 et 15 mois ont été enrôlés dans l'étude. Tous les examens d'imagerie avaient été réalisés à Gustave Roussy. L'étude était approuvée par les comités locaux institutionnelles en accord avec la déclaration d'Helsinki. Une lettre d'information à l'attention des patients les informait de leur participation à l'étude, en accord avec les guidelines nationales et institutionnelles.

Les critères d'inclusion à l'étude étaient :

- Les patients dont la RCS avait bien fonctionnée. Le critère était en lien avec la réduction du volume tumoral.
- Mais qui ont présenté plusieurs mois par la suite une preuve de RN ou RLT.

Les critères d'exclusions étaient :

- Les patients pour lesquelles l'acquisition IRM présentait d'importants artefacts de mouvements.
- Les patients n'ayant pas reçu d'injection de gadolinium.

Pour le diagnostic de la radionécrose composé d'un diagnostic initial et d'un diagnostic final, le standard de référence correspondait au consensus d'experts munis des données d'imagerie et des analyses histopathologiques (biopsie). Le diagnostic initial RN ou RLT était effectué par un comité d'experts composé de deux neuro-oncologues, trois radiothérapeutes et deux neurochirurgiens. Le diagnostic final était réalisé soit après un suivi de trois ans par une analyse histopathologique après résection chirurgicale, ou par un examen clinique et l'imagerie.

3.2.2 Descriptif technique du protocole d'acquisition IRM

Les acquisitions étaient réalisées sur un IRM 3 Teslas (MR 750, Discovery, General Electric Healthcare). Les examens d'imagerie réalisés comprenaient les séquences suivantes : pondérée T1 pré-contraste, T2-Flair, DWI, DSC, pondérée T1 post-contraste (3 min post-injection, T1w3), pondérée T1 post-contraste (70 à 90 minutes après injection, T1w90). Nous nous sommes intéressé au cours de notre étude aux séquences T1w3 et à la DSC. Les paramètres de la séquence T1w étaient les suivants : séquence 3D écho de gradient rapide ; TR=6,1 ms, TE=2,1ms, épaisseur de coupe = 1 mm, Champ de vue 260x260x146 mm³ ; taille des voxels 0,8x0,8x1,2 mm³ ; matrice d'acquisition : 300x300x146. La séquence DSC était acquise en écho-planar, pondérée T2*, effectuant 50 mesures à intervalles de 2 à 3 s en fonction du nombre de coupes. Les paramètres étaient : TE=62,4 ms, angle de bascule=90° ; Matrice d'acquisition : 128x128 ; Champ de vue : 260x256 mm². Une injection en bolus d'agent de contraste, l'acide gadotérique (Dotarem®), était effectuée à une dose de 0,1 mmol/kg avec un débit de 5ml/s suivi par 20 ml de solution saline au débit de 3 ml/s. Les 8 premières acquisitions étaient

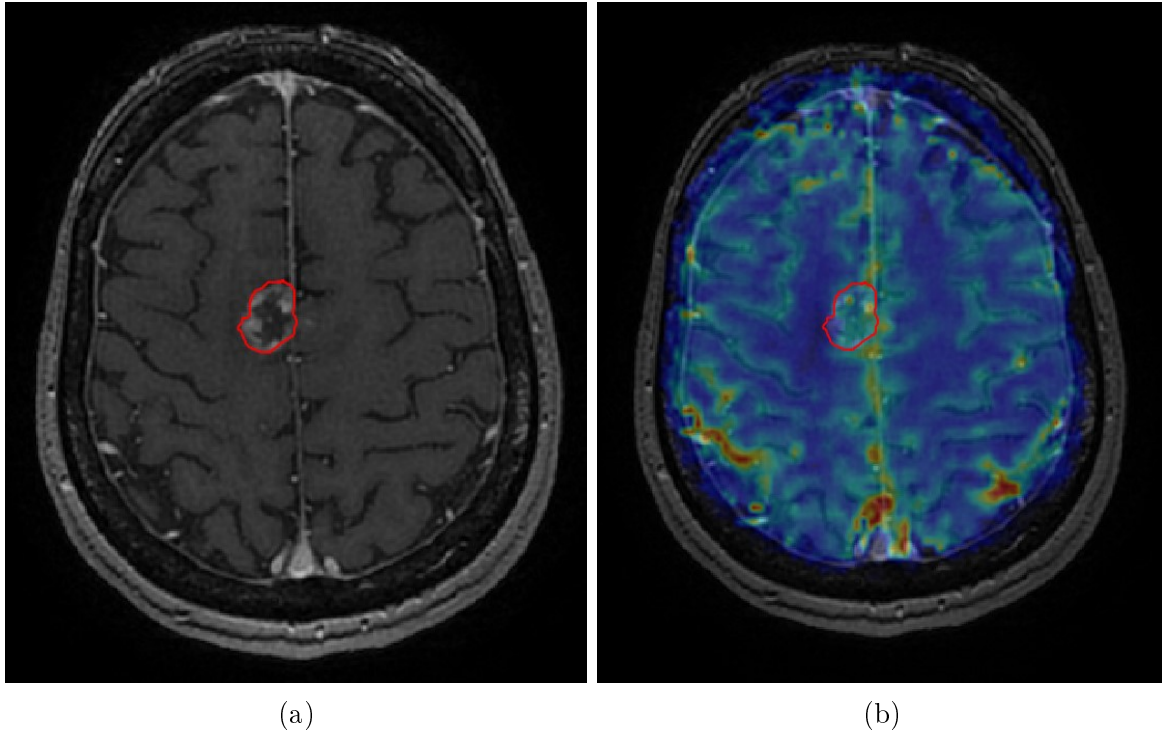


FIGURE 4.18 – Coupe Axiale IRM autour d'une lésion. (a) Image T1w3 : T1, 3 min post-injection de contraste. (b) Superposition de la carte rBV sur la séquence morphologique T1w3.

réalisées avant l'arrivée de l'agent de contraste pour établir une référence pré-contraste. La figure 4.18 présente une coupe représentative d'une lésion sur une séquence pondérée T1 3 minutes après l'injection de l'agent de contraste et sur une carte rBV.

3.3 Protocole de traitement et d'analyse des données

3.3.1 Génération des cartes paramétriques, régions d'intérêts et extraction du H_{index}

Les cartes paramétriques rBF, rBV, MTT, K2 et rBV corrigés (rBVc) étaient générées et extraites à l'aide d'Olea Sphère (v3.0.18, Olea Medical Solutions, France) à partir des images de la séquence DSC-IRM. Le paramètre rBF (relative Blood Flow) correspond au volume de sang passant à travers une quantité de tissu par unité de temps. Le paramètre rBV (relative Blood Volume) correspond au volume sanguin dans une quantité de tissu. MTT (Mean Transit Time) équivaut au temps moyen, en secondes, passé en moyenne par les globules rouges dans un volume de tissu. K2 est une estimation de la perméabilité capillaire et rBVc correspond donc à rBV corrigé des fuites d'agent de contraste d'après la carte K2.

Les volumes d'intérêts étaient réalisés semi-automatiquement sous Olea sphère par un radiologue doté de 5 ans d'expérience. La procédure semi-automatique consistait en

un contour automatique basé sur un seuillage des niveaux de gris et était ensuite corrigée manuellement si nécessaire. Le contour était ainsi réalisé le long de la bordure tumorale réhaussée en contraste.

Le H_{index} était extrait pour chaque type d'image c'est à dire des images pondérées T1 et des 5 cartes paramétriques rBF, rBV, MTT, K2 et rBV corrigés (rBVc). Le nombre de clusters était optimisé par type d'image et constant pour l'ensemble des individus. Le processus de clustering était fuzzy-c-means. L'analyse en composantes connectées était réalisée avec une connectivité 18.

3.3.2 Analyses des performances diagnostiques du H_{index}

Deux analyses complémentaires ont été réalisées pour analyser les performances du H_{index} dans le diagnostic différentiel de RLT et RN : l'une traditionnelle associant les statistiques et les études de courbes ROC, et la seconde à partir d'un apprentissage machine.

Dans le cadre de la première analyse, une étude préalable de corrélation entre les H_{index} extraits à partir des 5 types d'images (T1,rBF, rBV, MTT, K2 et rBVc) a été faite en vue d'éliminer des redondances. Parmi les corrélations, la plus significative était conservée. Le seuil du coefficient de Pearson à 0,95 était utilisé comme seuil de corrélation. Une fois cette étape faite, l'analyse visait à déterminer l'existence de différences statistiques pour le H_{index} entre les groupes RN et RLT à l'aide d'un test de Wilcoxon non apparié. Une valeur de $p \leq 0.05$ était considérée significative. Cette approche a été mise en œuvre en considérant l'ensemble des volumes des tumeurs cérébrales et elle a été déclinée jusqu'à déterminer un volume minimal en-dessous de laquelle l'extraction du H_{index} n'est plus pertinente pour éventuellement les exclure du jeu de données si nécessaire. Les performances du H_{index} ont ensuite été analysées indépendamment pour chacune des images, par les courbes ROC (fonction d'efficacité du récepteur) afin d'obtenir l'AUC, et d'en extraire le seuil optimal défini par l'index de Youden et la sensibilité, la spécificité, l'accuracy, la précision correspondants. Plus précisément, la courbe ROC est une mesure de la performance d'un indice dans la distinction entre deux groupes. Dans cette étude, RLT est considéré comme la condition positive et RN comme la condition négative. La courbe est obtenue en calculant pour chaque valeur possible de l'indice étudié, sa sensibilité et sa spécificité. La sensibilité correspond au taux de vrais positifs, dans notre cas le taux de RLT prédites comme RLT, et constitue les ordonnées de la courbe ROC. La spécificité correspond au taux de vrais négatifs ici le taux de RN prédite RN, ou 1-taux de faux positifs, et constitue l'abscisse de la courbe ROC. L'AUC correspond à l'aire sous la courbe ROC, comprise entre 0,5 et 1. Plus elle est proche de 1, meilleures sont les performances diagnostiques pour différencier RN versus RLT. L'index de Youden est défini comme la somme de

la spécificité et de la sensibilité moins une unité. A sa valeur maximale, il est possible de déterminer la valeur de l'indice étudié permettant d'obtenir le meilleur compromis entre la spécificité et la sensibilité. Enfin, la précision correspond à la valeur prédictive positive, autrement dit à 1 moins le taux de faux positifs, et l'accuracy correspond au taux de réussite de prédiction réussies pour les deux conditions.

Une seconde analyse s'est intéressée à optimiser la spécificité de la méthode en cherchant une combinaison mathématique du H_{index} des différentes images. Pour cela nous nous sommes appuyé sur une méthode d'apprentissage machine. Cette analyse a été réalisée par l'utilisation d'un arbre de décision pour classification binaire avec 80% des données utilisées pour l'entraînement du modèle et 20% pour le test. Ce modèle d'apprentissage permettait notamment d'obtenir un résultat interprétable grâce à une succession d'opérations binaires pour chaque paramètre employé.

Toutes les analyses des données précédentes ont été conduites sous R (version 4.1.1, R Core Team, 2021). En particulier pour l'approche d'apprentissage machine, le package « caret » a été employé.

3.4 Résultats

3.4.1 Caractéristiques des données patients

Les caractéristiques de la population et des régions d'intérêts sont fournies dans la table 4.14. L'étude comportait 118 examens réalisés sur un total de 77 patients. Parmi ces 77 patients, 25 ont effectué de multiples examens entre 7 et 15 mois après la RCS. Plus précisément 24 patients ont effectué deux examens avec un écart minimum de 4 mois entre les deux examens, et un patient a effectué jusqu'à 5 examens espacés de 2 à 5 mois. Ces examens ont été considérés comme indépendants.

Pour l'extraction du H_{index} , l'optimisation du nombre de clusters résultait pour chaque type d'image en $K_{T1}=6$, $K_{rBF}=4$, $K_{rBV}=4$, $K_{MTT}=3$, $K_{K2}=3$, et $K_{rBVc}=3$ clusters.

3.4.2 Évaluation des performances du H_{index} pour le diagnostic RN vs RLT

Les corrélations entre les H_{index} des différentes images post-injection ont été évaluées et observées entre les images K2 et rBVc ($R^2=0,98$), et entre rBV, rBF et MTT ($R^2>0,95$). Ainsi les images retenues pour la suite des analyses ROC étaient au nombre de 3 : T1w3, rBF et rBVc.

A l'aide des tests de Wilcoxon non-appariés, la significativité du H_{index} pour différencier les RN des RLT a été évaluée. Il apparaît sur la figure 4.19 que la mesure du H_{index} démontrait une différence statistique significative pour des lésions supérieures à $0,6 \text{ cm}^3$ et une cassure de la courbe avec une baisse importante de p au delà de

Caractéristiques	Valeurs
Age (années)	56.9 ± 10.6 [24 - 78]
Sexe	
Nombre patients	77
Nombre d'examens	118
Femmes	76 (66.3%)
Hommes	42 (33.7%)
Tumeur primaire	
Bronchique	57 (45.8%)
Mélanome	22 (18.6%)
Carcinome du sein	19 (16.1%)
Carcinome papillaire de la thyroïde	8 (6.7%)
Ovaire	3 (2.5%)
Sarcome	4 (3.4%)
Cellules claires du rein	3 (2.5%)
Autres	2 (1.7%)
Traitement Systémique	
Immunothérapie (ICI)	44
Chimiothérapie ou Thérapie ciblée	74
Diagnostic Final	
Radionécrose (RN)	74
Récidive locale tumorale (RLT)	44
Volume lésions ROI (cm ³)	8.98 ± 12.63 [46.50 - 80.57]

TABLE 4.14 – Résumé des caractéristiques de la population étudiée pour le diagnostic des RN ou des RLT.

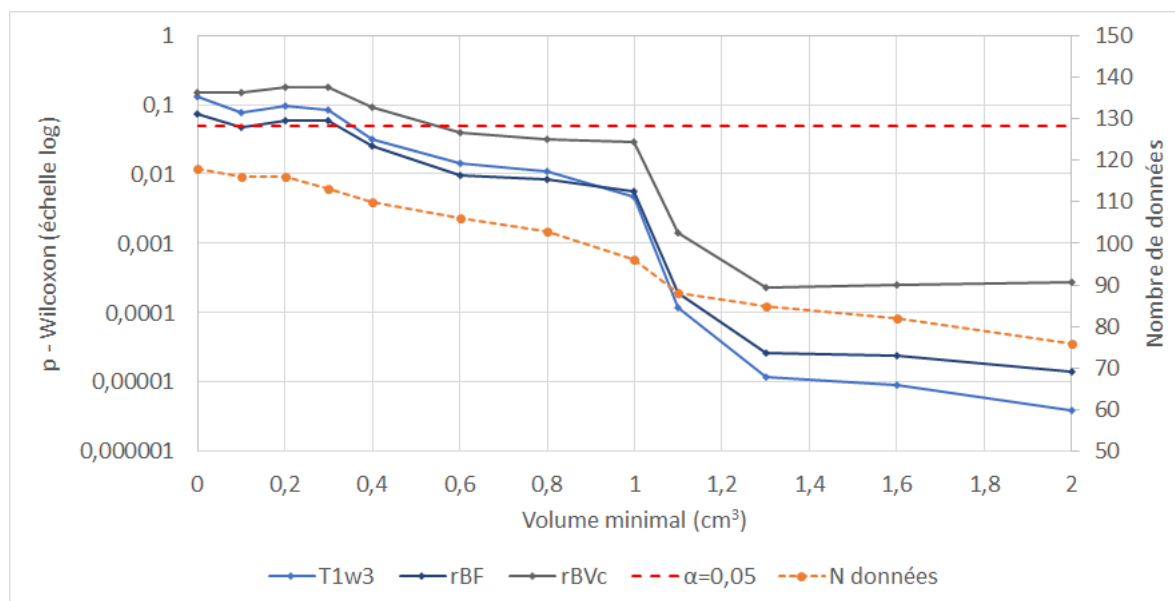


FIGURE 4.19 – Différences statistiques du H_{index} entre les groupes RN et RLT pour chaque carte paramétrique en fonction du volume minimal (cm^3) des lésions admises dans les données. Axe principal (échelle log) : p-value, Axe secondaire (droite) : Nombre de données. La courbe en pointillés rouge correspond à la limite 0,05 en dessous de laquelle la différence est considérée comme significative. La courbe orange en pointillés, correspond au nombre de données restantes après exclusion des volumes inférieurs à la limite de l'axe des abscisses. Les courbes pleines correspondent à la valeur de p pour chaque type d'image.

1 cm^3 . En effet, à partir de cette limite basse de volume, tous les H_{index} extraits de chaque type d'image post-injection démontraient un $p < 0,05$ dans les tests de Wilcoxon. Le tableau 4.15 relève l'ensemble des valeurs de ce graphique ainsi que le nombre de données après exclusion des volumes inférieurs aux différents seuils, au total et par groupe. Plus précisément, les H_{index} évalués sur les images pondérées T1, et les cartes paramétriques rBF étaient statistiquement significatifs dans la différenciation des RN et des RLT à partir de 0,4 cm^3 et à partir de 0,6 cm^3 pour les cartes rBvc. Les performances du H_{index} ont donc ensuite été évaluées par l'analyse des courbes ROC pour des volumes supérieures ou égaux à 1,1 cm^3 uniquement.

Les courbes ROC ont été produites et analysées pour évaluer les performances du H_{index} à partir des données retenues pour trois volumes minimum différents : 1,1 ; 1,5 ; et 2 cm^3 . Les résultats AUC, sensibilité, spécificité, accuracy, précision et meilleur seuil pour chaque image sont fournis dans les tables 4.16, 4.17, 4.18 respectivement pour les trois limites de volumes minimums. Il en ressort que les H_{index} mesurés sur T1w3 et rBF obtenaient systématiquement de meilleures performances ($AUC \geq 74,2\%$) que la signature radiomique lors de l'étude précédente (AMMARI S et al., 2021) ($AUC = 72,1\%$) et rBvc obtenaient de meilleures performances qu'à partir des volumes supérieurs à 1,5 cm^3 ($AUC \geq 73,6\%$). En comparaison à la TRAM seule ($AUC = 76,8\%$) obtenue

Vol. Min (cm ³)	Wilcoxon - p			Données		
	T1w3	rBF	rBVc	N	N RN	N RLT
0	0,13	0,073	0,15	118	74	44
0,1	0,077	0,048	0,15	116	73	43
0,2	0,094	0,058	0,18	116	72	43
0,3	0,084	0,058	0,18	113	70	42
0,4	0,032	0,025	0,091	110	69	40
0,6	0,014	0,0094	0,04	106	65	38
0,8	0,011	0,0083	0,032	103	64	37
1	0,0047	0,0055	0,029	96	60	35
1,1	0,00012	0,00019	0,0014	88	57	31
1,3	0,000012	0,000026	0,00023	85	55	29
1,6	0,0000093	0,000024	0,00025	82	53	28
2	0,0000039	0,000014	0,00028	76	48	27

TABLE 4.15 – Valeurs de p pour les tests de Wilcoxon (RN vs RLT) pour chaque type d'image au regard du volume minimal retenu pour les lésions. Sont également reportés le nombre de données restantes après exclusion au total et pour chacun des groupes.

après 90 minutes post-injection, les H_{index} issus des images T1w3 et rBF obtenaient de meilleures performances ($AUC \geq 77,3\%$) à partir des volumes supérieurs à $1,5 \text{ cm}^3$. En comparaison aux radiologues associés à la TRAM ($AUC=78,7\%$), le H_{index} obtenait des meilleures performances ($AUC=81,0\%$) avec les images T1w3 uniquement pour les volumes supérieures à 2 cm^3 . Dans ce cas de figure la sensibilité était accrue avec les images T1w3 (85.2%) et rBF (74.0%) par rapport au radiologue seul (31.3%), au radiologue aidé de la TRAM (65.6%) et la TRAM seule (60.9%). En revanche, la signature radiomique obtenait une sensibilité de 78% au détriment de sa spécificité de 60%. Au regard de la spécificité, la TRAM et TRAM+radiologue était au-delà de 91% tandis que le H_{index} obtenu immédiatement en post-injection se situait au maximum à 68,7% pour la carte T1, 81,2% pour la carte rBF et 75% pour la carte rBVc. Par ailleurs, les seuils de H_{index} correspondants étaient respectivement pour les images T1w3, rBF et rBVc de 469,8, 411,6 et 388,2. Ces seuils ne variaient pas en fonction des volumes tumoraux sélectionnés au-dessus de 1 cm^3 , $1,5 \text{ cm}^3$, 2 cm^3 .

Enfin, les H_{index} issus des trois types d'imageries identifiées, étaient utilisées pour créer un arbre de décision binaire en considérant des volumes supérieurs à 2 cm^3 . Le jeu de données d'entraînement comportait 60 examens (37 RN, 23 RLT) et le jeu de données test 15 (11 RN, 4 RLT). Il en résulte un arbre de décision qui contenait une unique branche séparant le jeux de données de manière optimale à l'aide du H_{index} extrait de la carte rBF à la valeur seuil du $H_{index} = 606$. Les performances n'étaient pas améliorées en comparaison aux indices seuls : sensibilité : 0,5 (RLT : 2 succès, 2 échecs). ; spécificité : 91,7% (RN : 11 succès, 1 échec) ; Précision : 50% ; Accuracy : 73,7%.

Image	T1w3	rBF	rBVc
AUC (%)	74.9	74.2	70.7
AUC CI 95%	63.3-86.6	62.3-86.1	58.7-82.3
Sensibilité(%)	74.1	64.5	61.3
Spécificité(%)	73.6	84.2	78.6
Accuracy(%)	73.8	77.2	72.4
Précision(%)	60.5	68.9	61.3
Meilleur Seuil	469.8	411.5	388.2

TABLE 4.16 – Résumé de l'analyse des performances du H_{index} à partir de la courbe ROC et du meilleur seuil pour la différentiation des RN et des RLT. Les données dont les volumes des lésions étaient **inférieures à 1,1 cm^3** étaient éliminées.

Image	T1w3	rBF	rBVc
AUC	78.5	77.3	73.6
AUC CI 95%	67.8-89.2	66.0-88.7	62.0-85.2
Sensitivité	79.3	68.9	65.5
Spécificité	71.7	83.0	76.9
Accuracy	74.3	78.0	72.8
Précision	60.5	68.9	61.2
Meilleur Seuil	469.8	411.5	388.2

TABLE 4.17 – Résumé de l'analyse des performances du H_{index} à partir de la courbe ROC et du meilleur seuil pour la différentiation des RN et des RLT. Les données dont les volumes des lésions étaient **inférieures à 1,5 cm^3** étaient éliminées.

Image	T1w3	rBF	rBVc
AUC	81.0	79.2	74.8
AUC CI 95%	70.7-91.1	67.7-90.7	62.8-86.9
Sensitivité	85.2	74.0	70.3
Spécificité	68.7	81.2	75.0
Accuracy	74.7	78.7	73.3
Précision	60.5	68.9	61.3
Meilleur Seuil	469.8	411.6	388.2

TABLE 4.18 – Résumé de l'analyse des performances du H_{index} à partir de la courbe ROC et du meilleur seuil pour la différentiation des RN et des RLT. Les données dont les volumes des lésions étaient **inférieures à 2 cm^3** étaient éliminées.

3.5 Discussion

Au cours de cette étude, nous nous sommes intéressés au diagnostic différentiel de deux pathologies, la radionécrose et la récurrence tumorale, dans le contexte d'un suivi post-thérapeutique de métastases cérébrales par l'imagerie multiparamétrique en IRM. Dans ce cas, l'examen IRM était composé d'une imagerie T1w post-injection de contraste et d'une séquence DSC-IRM permettant d'obtenir cinq cartes paramétriques rBF, rBD, MTT, K2, rBV corrigée (rBVc). Cette étude rétrospective s'inscrivait dans le prolongement de celle menée précédemment au sein de l'équipe et qui s'était appuyée d'une part sur la carte TRAM et d'autre part sur une analyse radiomique à partir de l'imagerie T1w tardive après 90 minutes post-injection. L'ambition de notre étude était d'explorer l'intérêt du H_{index} développé dans cette thèse pour aider au diagnostic différentiel et ce, dès le début de l'examen d'IRM, soit immédiatement post-injection.

En premier lieu, l'étude des corrélations entre les H_{index} extraits des différentes cartes paramétriques issues de l'imagerie DSC-IRM et de l'imagerie T1w post-injection de contraste a révélée des corrélations importantes toutefois attendues entre les différentes images. En effet, la carte rBVc s'appuie sur la carte K2 pour corriger la carte rBV originale, et les cartes rBF, rBV et MTT sont mathématiquement liées par la relation entre le débit, le volume et le temps. Ces analyses ont ainsi permis de focaliser la suite de notre étude sur les trois types d'imageries non corrélées qui étaient T1w 3 min post-injection de contraste, rBF et rBVc.

Le H_{index} était extrait à partir de chacun de ces types d'images dans les ROI identifiées par un radiologue expert et correspondantes aux lésions. Une première analyse a permis de démontrer à l'aide des traditionnels tests de Wilcoxon le pouvoir diagnostique du H_{index} dans la différentiation des pathologies. Toutefois, il a aussi été observé que l'analyse de l'hétérogénéité faisait sens avec des différences significatives pour des lésions dont le volume était supérieur à $1,1 \text{ cm}^3$. En effet, en dessous de cette limite, les H_{index} selon les types d'images ne démontraient pas tous une différence significative. Ces résultats peuvent s'expliquer par la capacité d'extraire des informations sur l'hétérogénéité spatiale d'une image au regard des performances spatiales des systèmes d'acquisition. En effet, en DSC-IRM, la résolution spatiale est de 5mm^3 ($1 \times 1 \times 5 \text{ mm}^3$), des ROI de 1cm^3 de volume comprenaient environ 200 voxels et ce répartis sur plusieurs coupes. La pertinence d'évaluer l'hétérogénéité d'une structure si petite mérite d'être posée. Il est difficile de tirer de plus amples conclusions sur l'impact du volume des ROI pour les analyses d'hétérogénéité car peu d'études s'intéressent à son influence. Parmi elles, une étude évoque un lien dans le cadre de l'IRM de diffusion des cancers cervicaux, où les indices de texture extraits dans des petits volumes diffèrent des gros volumes (WORMALD et al., 2020) et une autre étude suggère que les indices de texture doivent être interprétés avec précautions lors de l'étude de différents volumes d'intérêts en raison

de la forte corrélation entre certains indices et le volume (GOYA-OUTI et al., 2018). Ainsi, ces volumes de petites tailles étaient exclus des analyses suivantes.

L'analyse des courbes ROC appliquée au H_{index} extrait à partir de ces trois types d'images a permis de démontrer des performances globales satisfaisantes de notre méthodologie dans ce contexte exploratoire. Dans le cadre de l'étude menée précédemment au sein de l'unité (AMMARI S et al., 2021) la carte TRAM avait aidé le radiologue à différencier les RN des RLT avec une AUC=78,4%, dans le cas d'une utilisation TRAM seule sans radiologue l'AUC était de 76,8%, bien meilleure que le radiologue seul en terme de sensibilité et spécificité avec une AUC de seulement 57,4%. Une approche radiomique supplémentaire sur la carte TRAM permettait d'obtenir une AUC de valeur intermédiaire égale à 72,1%. De nombreuses autres études se sont intéressées à la radiomique en IRM multiparamétrique pour l'évaluation des métastases cérébrales et notamment pour l'évaluation de la réponse thérapeutique et des récives métastatiques avec des résultats encourageants avec des AUC comprises entre 0,73 et 0,98 (cf Review (LOHMANN et al., 2020)).

En comparaison, le H_{index} a obtenu de meilleures performances sous condition de considérer des lésions supérieures à un volume seuil. En effet, pour les volumes ≥ 2 cm³, le H_{index} obtenait une AUC=81% pour les images T1w, 79,2% pour les cartes rBF et 74,8% pour les cartes rBVc. Ces résultats sont par ailleurs facilement interprétables. Ils suggèrent que les RLT possèdent un H_{index} supérieur au seuil relevé dans les courbes ROC et présentent ainsi une hétérogénéité plus importante qu'elle soit tissulaire, déterminée à partir des cartes T1w3 ou bien vasculaire à partir des cartes rBF et rBVc. Enfin, et dans un contexte de routine clinique possédant des contraintes temporelles importantes, contrairement à l'étude précédente qui mettait en œuvre une séquence d'images longues nécessitant un intervalle de 90 minutes entre les deux acquisitions T1w, le H_{index} extrait des premières acquisition IRM de contraste uniquement démontrait de bonnes performances.

Enfin, la combinaison des H_{index} issus des différents types d'images dans l'implémentation d'un arbre de décision binaire n'a pas permis d'obtenir de meilleurs résultats. On remarque d'ailleurs que la valeur seuil du H_{index} était élevée, de valeur 606 (u.a), en comparaison avec 411 (u.a) pour la carte rBF lors de l'approche ROC. Néanmoins, d'une part, la faible taille de la cohorte (<100) n'était peut être pas suffisante et d'autre part, les variables cliniques (âge, type de tumeur primaire, immunothérapie O/N, symptômes, chirurgie O/N, diagnostic initial) n'ont pas été incluses dans la construction du modèle puisqu'il s'agissait de s'intéresser au pouvoir diagnostique de l'hétérogénéité uniquement et plus particulièrement de notre H_{index} . De futures analyses pourront s'intéresser à l'utilisation des variables cliniques mais également aux valeurs moyennes des paramètres explorés. Cette méthode de classification binaire avait été préférée à d'autres algorithmes tout aussi commun

d'apprentissage machine en raison de ses capacités à fournir un résultat totalement interprétable en terme d'hétérogénéité puisqu'il s'agit d'une succession de décisions binaires pour chacun des paramètres sélectionnés par le modèle d'apprentissage selon leur aptitude à séparer les groupes RN et RLT considérés.

3.6 Conclusion

En conclusion, nous avons montré que le H_{index} était capable de procurer d'excellentes performances diagnostiques pour dissocier les radionécroses des récidives tout en permettant une interprétation sur l'hétérogénéité de l'image de la tumeur. Ces résultats ont par ailleurs été obtenus dans des conditions particulièrement favorables pour notre méthode puisqu'elle a pu s'appliquer sur des images acquises très tôt durant l'examen contrairement à la méthode TRAM qui nécessitait une séquence supplémentaire à 90 minutes après l'injection du produit de contraste. Cette application pour le H_{index} apparaît donc comme une première piste d'application comme biomarqueur en cancérologie avec toutefois la nécessité d'explorer plus finement la limitation du volume tumoral pour y apporter éventuellement une correction.

4 Conclusion

Les travaux présentés dans ce chapitre ont permis d'explorer différentes pistes d'applications possibles d'un biomarqueur de l'hétérogénéité, le H_{index} , et d'évaluer son apport potentiel, dans un contexte clinique d'imagerie en cancérologie. Plus précisément, trois études distinctes ont permis d'explorer tout d'abord la robustesse de cet indice au travers de sa répétabilité dans le contexte de l'IRM de diffusion cérébrale, puis son intérêt clinique potentiel pour faciliter la standardisation de l'AIF dans le contexte de la DCE-IRM de la sphère ORL et enfin d'évaluer ses performances diagnostiques pour la différenciation de la radionécrose et des récidives locales tumorales dans le cadre de l'IRM de contraste neurologique. Nous revenons sur ces différents résultats.

L'évaluation de la répétabilité d'un biomarqueur d'imagerie est une étape obligatoire, conformément aux recommandations internationales, afin de pouvoir l'employer tout d'abord dans le cadre d'une étude clinique prospective à grande échelle avant même de pouvoir songer à son utilisation en routine. L'étude prospective menée ici sur volontaires sains dans le cadre de l'IRM de diffusion cérébrale constituait ainsi une première étape de la validation technique du H_{index} . Les résultats obtenus démontraient que l'évaluation de l'hétérogénéité à partir du H_{index} en IRM de diffusion était faisable et les résultats encourageants permettent désormais d'envisager une évaluation plus approfondie de ce biomarqueur. Plus précisément, la bonne répétabilité intra-sujets qui a été mesurée durant cette thèse devra être confirmée à l'échelle de plusieurs appareils IRM

et de plusieurs départements de radiologie afin de caractériser pleinement la robustesse en intégrant cette-fois ci, la reproductibilité multicentrique. Enfin, avec une ambition initiale pour suivre l'efficacité thérapeutique, la robustesse du H_{index} devrait par la suite être évaluée dans un contexte clinique pathologique pour lequel l'hétérogénéité est bien différente de celle rencontrée sur le cerveau de volontaires sains.

La standardisation de la mesure d'AIF est une des raisons limitantes à la qualification des biomarqueurs issus de la DCE-IRM pour une utilisation de routine. En raison de la difficulté de la mesure et de l'aspect chronophage de la méthode, une approche de détermination automatique de l'AIF est souvent proposée par les logiciels de quantification. Les résultats présentés dans ce chapitre ont montré que ce mode de génération influait tout d'abord sur l'AIF elle-même ainsi que sur les paramètres pharmacocinétiques associés. A l'inverse, le H_{index} n'était pas sensible aux modes de détermination de l'AIF et sa stabilité a été comparée à celles des indices de texture issus de la GLCM. A l'exception de deux indices qu'étaient l'énergie et le contraste, la stabilité de ces indices était également bonne. Nous avons ainsi suggéré que l'hétérogénéité spatiale des images était peu sensible aux variations de l'AIF. Ces résultats étaient d'ailleurs concordant avec la caractérisation du H_{index} sur fantôme numérique effectuée plus tôt au cours de ce travail de thèse. Ces résultats nous encouragent à une utilisation plus large du H_{index} dans le cadre d'une étude multicentrique afin de valider ces résultats en vue de s'affranchir de la standardisation de l'AIF.

Enfin, une dernière application s'est intéressée aux performances diagnostiques du H_{index} . Les méthodes actuelles permettant d'effectuer le diagnostic différentiel entre la radionécrose et la récurrence locale des tumeurs métastatiques cérébrales impliquent soit de l'imagerie TEP ionisante, soit un examen IRM long associant deux acquisitions identiques espacées de 90 minutes. Dans ce contexte, nous avons démontré que l'évaluation de l'hétérogénéité des lésions par le H_{index} à partir d'un seul examen T1w post-contraste et/ou DSC-IRM détenait un fort potentiel avec une interprétation intuitive des résultats. Une limitation a toutefois été relevée sur les volumes tumoraux exploitables par la méthode de détermination du H_{index} . Il s'agissait ainsi d'une première piste d'application clinique du H_{index} en tant que biomarqueur d'imagerie diagnostique.

En conclusion, ce chapitre a exploré des pistes d'applications à partir de diverses modalités d'IRM (DWI, DCE, DSC) et toutes ont démontré des résultats qui nous encouragent à poursuivre les efforts pour le développement d'un biomarqueur de l'hétérogénéité intuitif et facile à mettre en œuvre. Ces applications ont également permis de souligner les pistes d'améliorations futures du H_{index} avec notamment le sujet concernant le volume des lésions étudiées soulevé lors de l'étude diagnostic et le critère relatif au nombre de clusters à employer soulevé lors de l'étude de répétabilité.

Bibliographie

- AMMARI S, RABIEE B, JUVINA S, ANTONIOS L, SALLÉ DE CHOU R, BALLEYGUIER C, ASSI T, BIDAULT F, EL HAİK M, GARCIA GCTE, LAURENCE M, LIMKIN E, BOCKEL S, QUILLEN T A, KHETTAB M, ROBERT C, REUZE S, MOALLA S, GIROT C, ... DERCLE L. (2021). (soumis) Performance of conventional and machine-learning approaches for the diagnosis of tumor recurrence on MRI after radiation therapy of brain metastases. *Diagnostics*.
- ANDREWS, D. W., SCOTT, C. B., SPERDUTO, P. W., FLANDERS, A. E., GASPAR, L. E., SCHELL, M. C., WERNER-WASIK, M., DEMAS, W., RYU, J., BAHARY, J.-P., SOUHAMI, L., ROTMAN, M., MEHTA, M. P. & CURRAN, W. J. (2004). Whole brain radiation therapy with or without stereotactic radiosurgery boost for patients with one to three brain metastases : phase III results of the RTOG 9508 randomised trial. *The Lancet*, 363(9422), 1665-1672. [https://doi.org/10.1016/S0140-6736\(04\)16250-8](https://doi.org/10.1016/S0140-6736(04)16250-8)
- AOYAMA, H., SHIRATO, H., TAGO, M., NAKAGAWA, K., TOYODA, T., HATANO, K., KENJYO, M., OYA, N., HIROTA, S., SHIOURA, H., KUNIEDA, E., INOMATA, T., HAYAKAWA, K., KATOH, N. & KOBASHI, G. (2006). Stereotactic radiosurgery plus whole-brain radiation therapy vs stereotactic radiosurgery alone for treatment of brain metastases : a randomized controlled trial. *JAMA*, 295(21), 2483. <https://doi.org/10.1001/jama.295.21.2483>
- BAEK, T. W., KIM, S. H., PARK, S. J. & PARK, E. J. (2020). Texture analysis on bi-parametric MRI for evaluation of aggressiveness in patients with prostate cancer. *Abdominal Radiology*, 45(12), 4214-4222. <https://doi.org/10.1007/s00261-020-02683-4>
- BANE, O., HECTORS, S. J., WAGNER, M., ARLINGHAUS, L. L., ARYAL, M. P., CAO, Y., CHENEVERT, T. L., FENNESSY, F., HUANG, W., HYLTON, N. M., KALPATHY-CRAMER, J., KEENAN, K. E., MALYARENKO, D. I., MULKERN, R. V., NEWITT, D. C., RUSSEK, S. E., STUPIC, K. F., TUDORICA, A., WILMES, L. J., ... TAOUILI, B. (2018). Accuracy, repeatability, and interplatform reproducibility of T1 quantification methods used for DCE-MRI : Results from a multicenter phantom study. *Magnetic Resonance in Medicine*, 79(5), 2564-2575. <https://doi.org/10.1002/mrm.26903>

- BRYNOLFSSON, P., NILSSON, D., TORHEIM, T., ASKLUND, T., KARLSSON, C. T., TRYGG, J., NYHOLM, T. & GARPEBRING, A. (2017). Haralick texture features from apparent diffusion coefficient (ADC) MRI images depend on imaging and pre-processing parameters. *Scientific Reports*, 7(1), 4041. <https://doi.org/10.1038/s41598-017-04151-4>
- CAO, Y., POPOVTZER, A., LI, D., CHEPEHA, D. B., MOYER, J. S., PRINCE, M. E., WORDEN, F., TEKNOS, T., BRADFORD, C., MUKHERJI, S. K. & EISBRUCH, A. (2008). Early prediction of outcome in advanced head-and-neck cancer based on tumor blood volume alterations during therapy : a prospective study. *International Journal of Radiation Oncology*Biophysics*Physics*, 72(5), 1287-1290. <https://doi.org/10.1016/j.ijrobp.2008.08.024>
- CARROLL, T. J., ROWLEY, H. A. & HAUGHTON, V. M. (2003). Automatic calculation of the arterial input function for cerebral perfusion imaging with MR imaging. *Radiology*, 227(2), 593-600. <https://doi.org/10.1148/radiol.2272020092>
- CECCON, G., LOHMANN, P., STOFFELS, G., JUDOV, N., FILSS, C. P., RAPP, M., BAUER, E., HAMISCH, C., RUGE, M. I., KOCHER, M., KUCHELMEISTER, K., SELHAUS, B., SABEL, M., FINK, G. R., SHAH, N. J., LANGEN, K.-J. & GALLDIKS, N. (2016). Dynamic O 18 f-fluoroethyl)-l-tyrosine positron emission tomography differentiates brain metastasis recurrence from radiation injury after radiotherapy. *Neuro-Oncology*, now149. <https://doi.org/10.1093/neuonc/now149>
- CHEN, S., LIU, R., MA, C., BIAN, Y., LI, J., YANG, P., WANG, M. & LU, J. (2021). Repeatability of apparent diffusion coefficient at 3.0 tesla in normal pancreas. *Cureus*. <https://doi.org/10.7759/cureus.15734>
- CICONE, F., CARIDEO, L., SCARINGI, C., ROMANO, A., MAMEDE, M., PAPA, A., TOFANI, A., CASCINI, G. L., BOZZAO, A., SCOPINARO, F. & MINNITI, G. (2021). Long-term metabolic evolution of brain metastases with suspected radiation necrosis following stereotactic radiosurgery : longitudinal assessment by f-DOPA PET. *Neuro-Oncology*, 23(6), 1024-1034. <https://doi.org/10.1093/neuonc/noaa239>
- CICONE, F., MINNITI, G., ROMANO, A., PAPA, A., SCARINGI, C., TAVANTI, F., BOZZAO, A., MAURIZI ENRICI, R. & SCOPINARO, F. (2015). Accuracy of f-DOPA PET and perfusion-MRI for differentiating radionecrotic from progressive brain metastases after radiosurgery. *European Journal of Nuclear Medicine and Molecular Imaging*, 42(1), 103-111. <https://doi.org/10.1007/s00259-014-2886-4>
- DAPPA, E., ELGER, T., HASENBURG, A., DÜBER, C., BATTISTA, M. J. & HÖTKER, A. M. (2017). The value of advanced MRI techniques in the assessment of cervical cancer : a review. *Insights into Imaging*, 8(5), 471-481. <https://doi.org/10.1007/s13244-017-0567-0>

- FEDOROV, A., SCHWIER, M., CLUNIE, D., HERZ, C., PIEPER, S., KIKINIS, R., TEMPANY, C. & FENNESSY, F. (2018). An annotated test-retest collection of prostate multiparametric MRI. *Scientific Data*, 5(1), 180281. <https://doi.org/10.1038/sdata.2018.281>
- GAUSTAD, J.-V., HAUGE, A., WEGNER, C. S., SIMONSEN, T. G., LUND, K. V., HANSEM, L. M. K. & ROFSTAD, E. K. (2020). DCE-MRI of Tumor Hypoxia and Hypoxia-Associated Aggressiveness. *Cancers*, 12(7), E1979. <https://doi.org/10.3390/cancers12071979>
- GOYA-OUTI, J., ORLHAC, F., CALMON, R., ALENTORN, A., NIOCHE, C., PHILIPPE, C., PUGET, S., BODDAERT, N., BUVAT, I., GRILL, J., FROUIN, V. & FROUIN, F. (2018). Computation of reliable textural indices from multimodal brain MRI : suggestions based on a study of patients with diffuse intrinsic pontine glioma. *Physics in Medicine & Biology*, 63(10), 105003. <https://doi.org/10.1088/1361-6560/aabd21>
- GREENE-SCHLOESSER, D., ROBBINS, M. E., PEIFFER, A. M., SHAW, E. G., WHEELER, K. T. & CHAN, M. D. (2012). Radiation-induced brain injury : A review. *Frontiers in Oncology*, 2. <https://doi.org/10.3389/fonc.2012.00073>
- HAGIWARA, A., FUJITA, S., OHNO, Y. & AOKI, S. (2020). Variability and Standardization of Quantitative Imaging : Monoparametric to Multiparametric Quantification, Radiomics, and Artificial Intelligence. *Investigative Radiology*, 55(9), 601-616. <https://doi.org/10.1097/RLI.0000000000000666>
- HAKULINEN, U., BRANDER, A., RYYMIN, P., ÖHMAN, J., SOIMAKALLIO, S., HELMINEN, M., DASTIDAR, P. & ESKOLA, H. (2012). Repeatability and variation of region-of-interest methods using quantitative diffusion tensor MR imaging of the brain. *BMC Medical Imaging*, 12(1), 30. <https://doi.org/10.1186/1471-2342-12-30>
- HARRY, V. N., SEMPLE, S. I., PARKIN, D. E. & GILBERT, F. J. (2010). Use of new imaging techniques to predict tumour response to therapy. *The Lancet Oncology*, 11(1), 92-102. [https://doi.org/10.1016/S1470-2045\(09\)70190-1](https://doi.org/10.1016/S1470-2045(09)70190-1)
- HOEFNAGELS, F. W. A., LAGERWAARD, F. J., SANCHEZ, E., HAASBEEK, C. J. A., KNOL, D. L., SLOTMAN, B. J. & PETER VANDERTOP, W. (2009). Radiological progression of cerebral metastases after radiosurgery : assessment of perfusion MRI for differentiating between necrosis and recurrence. *Journal of Neurology*, 256(6), 878-887. <https://doi.org/10.1007/s00415-009-5034-5>
- JACKSON, A., BUCKLEY, D. L., PARKER, G. J. et al. (2005). *Dynamic contrast-enhanced magnetic resonance imaging in oncology* (T. 12). Springer.
- KOO, T. K. & LI, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>

- LASSAU, N., COIFFIER, B., FAIVRE, L., BENATSOU, B., BIDAULT, S., GIRARD, E., ASSELAIN, B., PITRE-CHAMPAGNAT, S. & KOSCIELNY, S. (2017). Study of inpatient variability and reproducibility of quantitative tumor perfusion parameters evaluated with dynamic contrast-enhanced ultrasonography. *Investigative Radiology*, 52(3), 148-154. <https://doi.org/10.1097/RLI.000000000000324>
- LE BIHAN, D., BRETON, E., LALLEMAND, D., GRENIER, P., CABANIS, E. & LAVAL-JEANTET, M. (1986). MR imaging of intravoxel incoherent motions : application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2), 401-407. <https://doi.org/10.1148/radiology.161.2.3763909>
- LEBIHAN, D., IIMA, M., FEDERAU, C. & SIGMUND, E. E. (Éd.). (2019). *Intravoxel incoherent motion (IVIM) MRI : principles and applications*. Pan Stanford Publishing.
- LOHMANN, P., KOCHER, M., RUGE, M. I., VISSER-VANDEWALLE, V., SHAH, N. J., FINK, G. R., LANGEN, K.-J. & GALLDIKS, N. (2020). PET/MRI Radiomics in Patients With Brain Metastases. *Frontiers in Neurology*, 11, 1. <https://doi.org/10.3389/fneur.2020.00001>
- LUCIA, F., VISVIKIS, D., DESSEROIT, M.-C., MIRANDA, O., MALHAIRE, J.-P., ROBIN, P., PRADIER, O., HATT, M. & SCHICK, U. (2018). Prediction of outcome using pretreatment 18f-FDG PET/CT and MRI radiomics in locally advanced cervical cancer treated with chemoradiotherapy. *European Journal of Nuclear Medicine and Molecular Imaging*, 45(5), 768-786. <https://doi.org/10.1007/s00259-017-3898-7>
- MEHRABIAN, H., DESMOND, K. L., SOLIMAN, H., SAHGAL, A. & STANISZ, G. J. (2017). Differentiation between radiation necrosis and tumor progression using chemical exchange saturation transfer. *Clinical Cancer Research*, 23(14), 3667-3675. <https://doi.org/10.1158/1078-0432.CCR-16-2265>
- MERISAARI, H., TAIMEN, P., SHIRADKAR, R., ETTALA, O., PESOLA, M., SAUNAVAARA, J., BOSTRÖM, P. J., MADABHUSHI, A., ARONEN, H. J. & JAMBOR, I. (2020). Repeatability of radiomics and machine learning for DWI : short-term repeatability study of 112 patients with prostate cancer. *Magnetic Resonance in Medicine*, 83(6), 2293-2309. <https://doi.org/10.1002/mrm.28058>
- MOURIDSEN, K., CHRISTENSEN, S., GYLDENSTED, L. & ØSTERGAARD, L. (2006). Automatic selection of arterial input function using cluster analysis. *Magnetic Resonance in Medicine*, 55(3), 524-531. <https://doi.org/10.1002/mrm.20759>
- MUI, A. W., LEE, A. W., LEE, V. H., NG, W., VARDHANABHUTI, V., MAN, S. S., CHUA, D. T., LAW, S. C. & GUAN, X. (2021). Prognostic and therapeutic evaluation of nasopharyngeal carcinoma by dynamic contrast-enhanced (DCE), diffusion-weighted (DW) magnetic resonance imaging (MRI) and magnetic resonance

- spectroscopy (MRS). *Magnetic Resonance Imaging*, 83, 50-56. <https://doi.org/10.1016/j.mri.2021.07.003>
- O'CONNOR, J. P. B., ABOAGYE, E. O., ADAMS, J. E., AERTS, H. J. W. L., BARRINGTON, S. F., BEER, A. J., BOELLAARD, R., BOHNDIEK, S. E., BRADY, M., BROWN, G., BUCKLEY, D. L., CHENEVERT, T. L., CLARKE, L. P., COLLETTE, S., COOK, G. J., DESOUSA, N. M., DICKSON, J. C., DIVE, C., EVELHOCH, J. L., ... WATERTON, J. C. (2017). Imaging biomarker roadmap for cancer studies. *Nature Reviews Clinical Oncology*, 14(3), 169-186. <https://doi.org/10.1038/nrclinonc.2016.162>
- PIERPAOLI, C., JEZZARD, P., BASSER, P. J., BARNETT, A. & DI CHIRO, G. (1996). Diffusion tensor MR imaging of the human brain. *Radiology*, 201(3), 637-648. <https://doi.org/10.1148/radiology.201.3.8939209>
- PORTNEY, L. G. & WATKINS, M. P. (2009). *Foundations of clinical research : applications to practice* (3rd ed) [OCLC : ocn166454308]. Pearson/Prentice Hall.
- QIN, Y., YU, X., HOU, J., HU, Y., LI, F., WEN, L., LU, Q. & LIU, S. (2019). Prognostic value of the pretreatment primary lesion quantitative dynamic contrast-enhanced magnetic resonance imaging for nasopharyngeal carcinoma. *Academic Radiology*, 26(11), 1473-1482. <https://doi.org/10.1016/j.acra.2019.01.021>
- RAUNIG, D. L., MCSHANE, L. M., PENNELLO, G., GATSONIS, C., CARSON, P. L., VOYVODIC, J. T., WAHL, R. L., KURLAND, B. F., SCHWARZ, A. J., GÖNEN, M., ZAHLMANN, G., KONDRATOVICH, M. V., O'DONNELL, K., PETRICK, N., COLE, P. E., GARRA, B., SULLIVAN, D. C. & QIBA TECHNICAL PERFORMANCE WORKING GROUP. (2015). Quantitative imaging biomarkers : a review of statistical methods for technical performance assessment. *Statistical Methods in Medical Research*, 24(1), 27-67. <https://doi.org/10.1177/0962280214537344>
- RODRIGUES, A., LOMAN, K., NAWROCKI, J., HOANG, J. K., CHANG, Z., MOWERY, Y. M., OYEKUNLE, T., NIEDZWIECKI, D., BRIZEL, D. M. & CRACIUNESCU, O. (2021). Establishing ADC-Based Histogram and Texture Features for Early Treatment-Induced Changes in Head and Neck Squamous Cell Carcinoma. *Frontiers in Oncology*, 11, 708398. <https://doi.org/10.3389/fonc.2021.708398>
- ROMAGNA, A., UNTERRAINER, M., SCHMID-TANNWALD, C., BRENDL, M., TONN, J.-C., NACHBICHLER, S. B., MUACEVIC, A., BARTENSTEIN, P., KRETH, F.-W. & ALBERT, N. L. (2016). Suspected recurrence of brain metastases after focused high dose radiotherapy : can [18f]FET- PET overcome diagnostic uncertainties ? *Radiation Oncology*, 11(1), 139. <https://doi.org/10.1186/s13014-016-0713-8>
- SANZ-REQUENA, R., PRATS-MONTALBÁN, J. M., MARTÍ-BONMATÍ, L., ALBERICH-BAYARRI, Á., GARCÍA-MARTÍ, G., PÉREZ, R. & FERRER, A. (2015). Automatic individual arterial input functions calculated from PCA outperform manual and population-averaged approaches for the pharmacokinetic modeling of DCE-MR

- images : PCA-based automatic individual AIF. *Journal of Magnetic Resonance Imaging*, 42(2), 477-487. <https://doi.org/10.1002/jmri.24805>
- SCHWIER, M., van GRIETHUYSEN, J., VANGEL, M. G., PIEPER, S., PELED, S., TEMPANY, C., AERTS, H. J. W. L., KIKINIS, R., FENNESSY, F. M. & FEDOROV, A. (2019). Repeatability of multiparametric prostate MRI radiomics features. *Scientific Reports*, 9(1), 9441. <https://doi.org/10.1038/s41598-019-45766-z>
- SHAW, P. J. & BATES, D. (1984). Conservative treatment of delayed cerebral radiation necrosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 47(12), 1338-1341. <https://doi.org/10.1136/jnnp.47.12.1338>
- SHESKIN, D. (2011). *Handbook of parametric and nonparametric statistical procedures* (5th ed). CRC Press.
- SHROUT, P. E. & FLEISS, J. L. (1979). Intraclass correlations : uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428. <https://doi.org/10.1037/0033-2909.86.2.420>
- SHUKLA-DAVE, A., OBUCHOWSKI, N. A., CHENEVERT, T. L., JAMBAWALIKAR, S., SCHWARTZ, L. H., MALYARENKO, D., HUANG, W., NOWOROLSKI, S. M., YOUNG, R. J., SHIROISHI, M. S., KIM, H., COOLENS, C., LAUE, H., CHUNG, C., ROSEN, M., BOSS, M. & JACKSON, E. F. (2019). Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *Journal of Magnetic Resonance Imaging*, 49(7), e101-e121. <https://doi.org/10.1002/jmri.26518>
- SONG, C., CHENG, P., CHENG, J., ZHANG, Y., SUN, M., XIE, S. & ZHANG, X. (2020). Differential diagnosis of nasopharyngeal carcinoma and nasopharyngeal lymphoma based on DCE-MRI and RESOLVE-DWI. *European Radiology*, 30(1), 110-118. <https://doi.org/10.1007/s00330-019-06343-0>
- TOFTS, P. S., BRIX, G., BUCKLEY, D. L., EVELHOCH, J. L., HENDERSON, E., KNOPP, M. V., LARSSON, H. B., LEE, T.-Y., MAYR, N. A., PARKER, G. J., PORT, R. E., TAYLOR, J. & WEISSKOFF, R. M. (1999). Estimating kinetic parameters from dynamic contrast-enhanced t1-weighted MRI of a diffusable tracer : standardized quantities and symbols. *Journal of Magnetic Resonance Imaging*, 10(3), 223-232. [https://doi.org/10.1002/\(SICI\)1522-2586\(199909\)10:3<223::AID-JMRI2>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1522-2586(199909)10:3<223::AID-JMRI2>3.0.CO;2-S)
- TRAVERSO, A., KAZMIERSKI, M., SHI, Z., KALENDRALIS, P., WELCH, M., NISSEN, H. D., JAFFRAY, D., DEKKER, A. & WEE, L. (2019). Stability of radiomic features of apparent diffusion coefficient (ADC) maps for locally advanced rectal cancer in response to image pre-processing. *Physica Medica*, 61, 44-51. <https://doi.org/10.1016/j.ejmp.2019.04.009>

- VALLADARES, A., BEYER, T. & RAUSCH, I. (2020). Physical imaging phantoms for simulation of tumor heterogeneity in PET, CT, and MRI : an overview of existing designs. *Medical Physics*, 47(4), 2023-2037. <https://doi.org/10.1002/mp.14045>
- WORMALD, B. W., DORAN, S. J., IND, T. E., D'ARCY, J., PETTS, J. & DESOUZA, N. M. (2020). Radiomic features of cervical cancer on t2-and diffusion-weighted MRI : prognostic value in low-volume tumors suitable for trachelectomy. *Gynecologic Oncology*, 156(1), 107-114. <https://doi.org/10.1016/j.ygyno.2019.10.010>
- YUSHKEVICH, P. A., PIVEN, J., HAZLETT, H. C., SMITH, R. G., HO, S., GEE, J. C. & GERIG, G. (2006). User-guided 3d active contour segmentation of anatomical structures : significantly improved efficiency and reliability. *NeuroImage*, 31(3), 1116-1128. <https://doi.org/10.1016/j.neuroimage.2006.01.015>
- ZHENG, D., CHEN, Y., CHEN, Y., XU, L., CHEN, W., YAO, Y., DU, Z., DENG, X. & CHAN, Q. (2014). Dynamic contrast-enhanced MRI of nasopharyngeal carcinoma : a preliminary study of the correlations between quantitative parameters and clinical stage : DCE-MRI of nasopharyngeal carcinoma. *Journal of Magnetic Resonance Imaging*, 39(4), 940-948. <https://doi.org/10.1002/jmri.24249>
- ZHENG, D., CHEN, Y., LIU, X., CHEN, Y., XU, L., REN, W., CHEN, W. & CHAN, Q. (2015). Early response to chemoradiotherapy for nasopharyngeal carcinoma treatment : value of dynamic contrast-enhanced 3.0 t MRI : assessing NPC chemoradiotherapy by DCE-MRI. *Journal of Magnetic Resonance Imaging*, 41(6), 1528-1540. <https://doi.org/10.1002/jmri.24723>
- ZHENG, D., YUE, Q., REN, W., LIU, M., ZHANG, X., LIN, H., LAI, G., CHEN, W., CHAN, Q. & CHEN, Y. (2017). Early responses assessment of neoadjuvant chemotherapy in nasopharyngeal carcinoma by serial dynamic contrast-enhanced MR imaging. *Magnetic Resonance Imaging*, 35, 125-131. <https://doi.org/10.1016/j.mri.2016.08.011>
- ZIAYEE, F., MÜLLER-LUTZ, A., GROSS, J., QUENTIN, M., ULLRICH, T., HEUSCH, P., ARSOV, C., RABENALT, R., ALBERS, P., ANTOCH, G., WITTSACK, H. & SCHIMMÖLLER, L. (2018). Influence of arterial input function (AIF) on quantitative prostate dynamic contrast-enhanced (DCE) MRI and zonal prostate anatomy. *Magnetic Resonance Imaging*, 53, 28-33. <https://doi.org/10.1016/j.mri.2018.06.004>

Annexe

Etude de la répétabilité du H_{index} en IRM de diffusion cérébrale

Résultats détaillés du nombre de clusters optimal employé pour chaque individu au regard de la région d'intérêt pour chaque carte paramétrique. Le nombre de clusters était défini comme le maximum obtenu pour l'optimisation entre l'acquisition test et retest.

Individu # étude	ROI					ROI					ROI					ROI				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	5	5	5	4	5	4	5	5	5	5	6	4	5	5	5	4	5	5	6	5
2	5	5	5	4	5	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5
3	5	5	5	5	5	5	5	5	5	5	4	4	5	5	5	5	5	4	5	5
4	4	5	4	5	5	5	4	5	5	4	4	4	6	5	5	5	5	5	5	5
5	5	5	4	5	5	5	5	5	5	5	6	4	5	4	5	5	5	5	5	5
6	5	5	5	5	5	4	5	5	5	5	4	4	5	5	6	5	5	5	5	5
7	4	4	5	4	5	6	4	6	5	5	4	4	5	6	5	5	5	5	5	5
8	5	5	5	5	5	6	5	5	5	5	4	4	6	5	5	5	5	5	5	5
10	5	4	5	5	5	5	5	5	5	5	5	4	6	5	5	4	5	5	5	5
11	5	5	5	5	5	4	5	6	5	6	4	4	6	5	5	4	5	6	5	5
12	5	4	5	4	5	4	5	5	5	5	5	5	5	5	5	6	5	5	5	5
13	5	5	4	4	5	5	5	5	4	5	4	4	5	5	5	5	5	5	5	5
14	5	5	5	4	5	5	5	4	5	5	5	5	4	5	5	5	4	4	6	5
15	5	5	5	5	5	5	5	5	6	6	4	4	5	5	6	5	5	5	5	6
16	4	5	5	5	5	5	5	5	5	4	5	4	5	5	5	4	5	5	5	5
17	4	5	4	4	5	5	5	5	5	5	5	4	6	5	5	5	5	5	5	5
18	4	4	5	5	5	4	5	5	5	5	3	4	4	4	4	4	5	5	5	5
19	5	5	5	5	5	5	5	5	5	5	4	5	5	4	4	5	5	5	5	5
20	5	5	5	5	5	5	5	4	5	6	6	4	6	5	5	5	5	5	5	5
21	5	4	5	5	5	4	5	5	5	5	5	4	5	5	5	5	5	5	5	5
22	5	5	5	4	5	5	5	4	4	4	4	4	4	5	5	4	5	5	5	5
23	5	5	5	5	5	5	5	5	5	5	4	4	6	5	5	5	5	5	5	5
24	5	5	5	4	5	4	4	5	5	4	5	4	6	5	5	5	5	5	5	5
Carte	ADC					D					D*					F				

TABLE 4.19 – Nombre de clusters optimal utilisé pour chaque individu et chaque ROI pour l'étude MRIRM de répétabilité du H_{index}

Conclusion

Présente à tous les stades du développement tumoral et à toutes les échelles, cellulaires à tissulaires, l'hétérogénéité tumorale est un sujet de recherche en pleine expansion. L'un des enjeux clés pour l'avenir de la lutte contre le cancer et plus particulièrement dans le cadre de la médecine personnalisée, réside dans la quantification de cette hétérogénéité notamment grâce à l'imagerie médicale qui par le développement de nouveaux biomarqueurs d'imagerie quantitative constitue un outil incontestable. Toutefois, la translation des recherches scientifiques vers la pratique clinique n'en est qu'à ses prémises. Trois axes majeurs peuvent expliquer ce retard : le manque de standardisation des procédures, le manque d'interprétabilité des résultats toujours plus complexes et enfin le manque de définition établie de ce qu'est l'hétérogénéité qui demeure encore une notion subjective. Dans le domaine de la radiomique, de nombreux efforts ont récemment été accomplis d'une part pour faciliter la standardisation des processus d'acquisitions des images et de leur analyse et d'autre part, pour rendre les résultats interprétables notamment par l'ajout d'une composante en lien avec une caractéristique phénotypique du cancer. Malgré ces efforts, aucun biomarqueur d'imagerie issu de la radiomique et plus généralement lié à l'hétérogénéité en imagerie n'a aujourd'hui fait son apparition dans la pratique clinique. Ces observations ont conduit, dès 2015, notre équipe à s'intéresser à la notion d'hétérogénéité tumorale par le biais de méthodes alternatives à la radiomique.

C'est dans le prolongement de ces ambitions que s'est inscrit mon travail de recherche présenté dans ce document et qui comporte trois volets. Le premier volet méthodologique visait à proposer une méthode de quantification de l'hétérogénéité d'une image puis à la caractérisation des métriques comme biomarqueurs potentiels avec comme motivation une interprétabilité aisée des résultats. Le second volet constituait une application directe de la méthode. Il s'agissait d'une preuve de concept à partir d'une étude préclinique démontrant le potentiel de notre méthode pour la caractérisation de l'hétérogénéité vasculaire tumorale en DCE-IRM dans un contexte de suivi de l'efficacité thérapeutique. Enfin, le dernier volet visait à élargir le champ d'applications en cancérologie d'un des biomarqueurs développés, le H_{index} regroupant trois études cliniques dont une prospective. La première était un préalable méthodologique et visait à valider la répétabilité du H_{index} en IRM de diffusion. La seconde s'est intéressée à

explorer son potentiel d'application pour faciliter la standardisation de la DCE-IRM dans le contexte de la cancérologie ORL. Et enfin, la dernière application explorait son potentiel pour le diagnostic différentiel de la radionécrose et des récives locales en IRM neurologique multimodale. Les principaux résultats de la thèse sont donc établis selon ces trois volets : méthodologique, préclinique, et clinique.

D'un point de vue méthodologique, l'enjeu était de quantifier l'hétérogénéité tout en maîtrisant l'interprétation des résultats. Pour ce faire, étant donné la subjectivité de la notion d'hétérogénéité, une définition basée sur la perception visuelle et donc d'approche intuitive a été établie. A partir de cette définition, une succession de processus a été appliquée pour extraire des indices quantifiables. Le premier processus, le clustering, associé à une méthode d'optimisation du nombre de clusters permettait d'identifier différentes régions d'intérêts à partir des cartes paramétriques étudiées et donc de mettre en avant l'hétérogénéité des intensités. Un second processus visait à extraire de chacune de ces sous régions, des métriques indépendantes de l'intensité des pixels afin de caractériser la complexité de leur organisation spatiale. La combinaison de ces différentes métriques caractérisant d'une part l'hétérogénéité des intensités et de la distribution spatiale des pixels permettait enfin de former un unique indice, le H_{index} . Dans ce travail méthodologique, une étude de caractérisation de l'ensemble de ces métriques et indices a été menée à partir de fantômes numériques conçus pour refléter différentes hétérogénéités d'images d'après nos définitions. Ainsi, l'influence de la connectivité, le seul hyperparamètre de notre méthode, et l'influence de la résolution d'image ont permis d'établir des recommandations pour la suite de ces travaux. La connectivité devait être maximisée lorsque possible et la résolution des images ne devait pas être dégradée spatialement afin de conserver les mêmes degrés d'hétérogénéité. Enfin, la robustesse à des variations de contraste a été testée et deux algorithmes de clustering, les K-moyennes et Fuzzy-C-Means, ont été comparés. Il apparaissait que le clustering permettait de s'affranchir des variations de contraste et que les deux algorithmes fournissaient des résultats similaires.

D'un point de vue préclinique, il s'agissait de la première application de la méthode développée pour explorer son potentiel dans le contexte du suivi de l'efficacité thérapeutique. Cette étude rétrospective en DCE-IRM était ainsi la première à conclure que nos indices apparaissent comme de bons potentiels biomarqueurs de l'efficacité d'une thérapie anti-angiogénique. En effet, alors que les mesures d'hétérogénéité démontraient des résultats similaires pour l'ensemble des souris avant le début du traitement, l'hétérogénéité des tumeurs diminuaient significativement de l'ordre de 40% au bout de quatre jours de traitement tandis que l'hétérogénéité des tumeurs des souris contrôles avait tendance à augmenter. De plus, les résultats d'une précédente étude menée sur cette cohorte, démontraient que le volume augmentait significativement au terme de onze jours de traitement pour les souris contrôles et stagnait pour les souris traitées. Ainsi,

notre méthode permettait de détecter de manière précoce, au terme de quatre jours, des différences significatives liées à l'application d'un traitement anti-angiogénique, contrairement au critère volumique habituellement considéré comme une référence pour les traitements cytotoxiques, et contrairement également aux indices de texture habituellement considérés pour la quantification de l'hétérogénéité des images.

D'un point de vue clinique, trois études ont été menées dont deux dédiées à explorer le champ d'application du H_{index} et son potentiel clinique. Une première étude prospective sur des volontaires sains a été finalisée avant la fin de cette thèse. Il s'agissait d'une étape clef de la validation d'un biomarqueur en évaluant la répétabilité du H_{index} . Cette étude menée en IRM de diffusion a particulièrement souligné la difficulté méthodologique pour évaluer rigoureusement la variabilité et la répétabilité d'un biomarqueur de l'hétérogénéité d'imagerie tant par le choix de l'objet d'étude que par l'analyse mathématique à mettre en œuvre pour les quantifier. Le H_{index} était bien répétable, au regard des recommandations sur les seuils des ICC, pour évaluer l'hétérogénéité de la diffusion de l'eau sur des cartes paramétriques d'ADC dans trois ROI sur cinq. Si dans un des deux cas, la raison d'une répétabilité modérée était attendue, l'autre cas de figure défavorable a mis en évidence une faiblesse dans le processus de clustering sur des valeurs d'intensité trop proches. Une seconde étude clinique, rétrospective, s'est intéressée à la problématique de la standardisation de l'AIF en imagerie DCE-IRM dans le contexte des cancers ORL. Les résultats montraient que le H_{index} pourrait permettre de faciliter la standardisation de cette modalité par l'utilisation d'un indice qui était stable quel que soit les quatre modes utilisés pour mesurer l'AIF. Ces résultats permettent ainsi d'envisager des études multicentriques pour lesquels la standardisation de l'AIF impacte fortement la mise en pratique clinique des biomarqueurs d'imagerie associées à la modalité. Enfin, une dernière étude clinique rétrospective visait à établir en IRM neurologique multimodale, le diagnostic différentiel entre la radionécrose et les récurrences locales tumorales dans le cadre du suivi post-thérapeutique des métastases cérébrales. Ce diagnostic, aujourd'hui effectué soit à l'aide d'un examen IRM compliqué associant sur 90 minutes plusieurs acquisitions en IRM de contraste, soit à partir de l'imagerie TEP ionisante, requiert le développement de nouveaux outils adaptés à des examens possiblement répétés tous les 4 à 6 mois en moyenne. Les résultats obtenus avec la méthode du H_{index} étaient prometteurs puisqu'ils permettaient à partir d'une série d'images IRM de contraste d'obtenir des résultats similaires en un temps d'acquisition court. Toutefois, ne surpassant pas l'imagerie TEP qui a notamment fait preuve d'une très forte spécificité dans ce diagnostic, ces résultats ne prétendent pas offrir une alternative à la pratique clinique et ils ont permis de soulever la problématique de la taille suffisante des lésions pour la quantification de l'hétérogénéité, identifiée comme au-delà d' 1cm^3 .

A la vue de la diversité des résultats obtenus au cours de cette thèse, les perspectives

scientifiques cliniques sont nombreuses. A court-terme, il s'agit tout d'abord de poursuivre l'étude de la répétabilité du H_{index} en exploitant des images issues de fantômes en vue de maîtriser ses caractéristiques. De plus les bons résultats obtenus en relation avec la variabilité de l'AIF incitent à prolonger l'étude du H_{index} comme méthode simplifiant la standardisation de la DCE-IRM. En effet, les résultats obtenus avec cette modalité étaient les plus prometteurs et de nombreuses étapes sont encore nécessaires pour la validation du H_{index} en tant que biomarqueur d'imagerie de l'efficacité thérapeutique. Dans un premier temps, il s'agirait de continuer l'étude menée sur la variabilité de l'AIF, à une échelle multicentrique afin de vérifier nos hypothèses de facilitation de la standardisation et en parallèle de démontrer une relation biologique entre notre biomarqueur potentiel et l'hétérogénéité vasculaire tumorale. A plus long-terme, il s'agirait d'appliquer la méthode à une étude prospective de suivi de l'efficacité thérapeutique pour laquelle des temps plus longs sont nécessaires car elles nécessitent un suivi patient sur plusieurs années. D'un point de vue méthodologique, des améliorations sont possibles et différentes pistes sont à explorer. Premièrement, la définition du nombre de clusters, bien qu'objective, doit être plus longuement étudiée afin de définir sa pertinence au regard de chacune des applications. En particulier l'influence de l'homogénéité des intensités et de la taille des lésions sur le processus de clustering devra être plus largement étudiée et pourrait également accompagner l'évolution de la méthode définissant le nombre de clusters. Deuxièmement, le processus de clustering était effectué à l'aide des algorithmes classiques sans aucun à priori physiologique. Le recours à des algorithmes plus avancés qui ferait notamment usage des dernières avancées en intelligence artificielle serait à explorer, tout en maintenant un fort degré d'interprétabilité des résultats. Dans le cas de l'imagerie multimodale, un clustering multimodale pourrait également être envisagé en combinaison avec notre méthode, tout en conservant les mêmes garanties. Enfin, malgré son manque crucial d'interprétabilité, la radiomique constitue un outil majeur de la nouvelle génération de biomarqueurs d'imagerie et les deux approches ne doivent pas mutuellement s'exclure. Avec le regain d'intérêt de la communauté pour comprendre et interpréter biologiquement les résultats de la radiomique, il est envisageable que le H_{index} puisse tout à fait s'intégrer dans un processus plus large que celui que nous avons décrit tout au long de ce travail.

Contributions Scientifiques

Publications

C. Girot, A. Volk, C. Walczak, N. Lassau, S. Pitre-Champagnat. New method for quantification of intratumoral heterogeneity : a feasibility study on K^{trans} maps from preclinical DCE-MRI. Magn Reson Mater Phy 34, 845–857 (2021). <https://doi.org/10.1007/s10334-021-00930-3>

S. Moalla, C. Girot, S. Franchi-Abella, S. Ammari, C. Balleyguier, N. Lassau, S. Pitre-Champagnat. Methodological Study to Investigate the Potential of Ultrasound-Based Elastography and Texture as Biomarkers to Monitor Liver Tumors. Diagnostics. 2020 ; 10(10) :811. <https://doi.org/10.3390/diagnostics10100811>

S. Ammari, B. Rabiee, S. Juvina, L. Antonios, R. Sallé de Chou, C. Balleyguier, T. Assi, F. Bidault, M. El Haik, GCTE. Garcia, M. Laurence, E. Limkin, S. Bockel, A. Quillent, M. Khettab, C. Robert, S. Reuze, S. Moalla, C. Girot, N. Lassau, E. Chouzenoux, L. Derclé. Performance of conventional and machine-learning approaches for the diagnosis of tumor recurrence on MRI after radiation therapy of brain metastases, Diagnostics, Soumis Janvier 2022.

Posters & Abstracts - Conférences

C. Girot, S. Moalla, S. Pitre-Champagnat, F. Bidault, S. Ammari, C. Balleyguier, A. Volk, N. Lassau. Radiomique pour les nuls : L'analyse de l'image par la texture mais pas que ..., Journées Françaises de Radiologie 2021 - **Poster Primé**

C. Girot, S. Moalla, S. Pitre-Champagnat, F. Bidault, S. Ammari, C. Balleyguier, A. Volk, N. Lassau. Radiomics for Dummies : Image Texture Analysis and much more ..., Radiological Society of North America, Annual Meeting 2021 - **Poster Primé, Certificate of Merit**

C. Girot, A. Volk, C. Walczak, S. Pitre-Champagnat. Anti-angiogenic therapy response in murine subcutaneous tumors by cluster analysis of parametric maps from DCE-MRI, IEEE MIC, 2019

Title : Evaluation of tumor heterogeneity with parametric imaging

Keywords : Cancer, Radiology, Biomarker, Heterogeneity

Abstract : In oncology and the era of precision medicine, the use of imaging biomarkers became essential for personalized diagnosis and therapeutic monitoring. Except for PET imaging, imaging biomarkers used in routine practice are mainly based on anatomical information without considering tumor physiology and the associated heterogeneity. However, this tissular or vascular intra-tumoral heterogeneity is now recognized as a prognostic or predictive factor and should be considered in the development of new biomarkers, which is made possible by technological and methodological advances. From a practical point of view, MRI imaging, whether contrast or diffusion, provides quantitative information on tumor micro-vascularization and associated heterogeneity resulting from the neo-angiogenic process. To complement conventional intra-tumor heterogeneity quantification methods, such as texture analysis, which generally suffers from a lack of interpretability, we propose an image processing and analysis method aiming to provide an intuitive representation of heterogeneity. Thus, based on parametric maps, we have developed indices evaluating heterogeneity as the cumulative complexity of pixel intensities and their spatial organization in different tumor subregions. These regions, identified by a clustering process, characterize the heterogeneity of the measured imaging parameter intensities. The spatial organization of these subregions and their constituent elements are then analyzed and quantified to provide a single index of heterogeneity at the tumor and clusters scales. Through different numerical, preclinical, and clinical applications, the performances of this new index have been characterized following international recommendations regarding the development of imaging biomarkers. These applications consisted first of a proof of concept on MRI data from a small animal tumor model. It illustrated the potential clinical interest of our index for monitoring therapeutic efficacy in the context of targeted therapies. Then, different applications to clinical MRI allowed us to identify potential clinical contributions, especially about the standardization of image acquisition and analysis protocols, and thus to progress in the validation process of this index as an imaging biomarker. All the results obtained during this thesis are promising and allow us to envisage the use of this method in a wider framework of clinical investigations.

Titre : Évaluation de l'hétérogénéité tumorale en imagerie paramétrique

Mots clés : Cancer, Radiologie, Biomarqueur, Hétérogénéité

Résumé : En cancérologie et à l'ère de la médecine de précision, l'utilisation de biomarqueurs d'imagerie est devenue indispensable pour le diagnostic et le suivi thérapeutique personnalisé. A l'exception de l'imagerie TEP, les biomarqueurs d'imagerie utilisés en routine sont principalement basés sur des informations anatomiques sans tenir compte de la physiologie tumorale et de son hétérogénéité associée. Or, cette hétérogénéité qu'elle soit tissulaire ou vasculaire intra-tumorale est aujourd'hui reconnue comme un facteur pronostic ou prédictif et devrait être considérée dans l'élaboration des nouveaux biomarqueurs dont les développements sont rendus possibles grâce aux avancées technologiques et méthodologiques. D'un point de vue pratique, l'imagerie IRM qu'elle soit de contraste ou de diffusion permet aujourd'hui d'accéder à des informations quantitatives relatives à la micro-vascularisation tumorale et à son hétérogénéité, résultat du processus de néo-angiogenèse. Pour compléter les méthodes conventionnelles de quantification de l'hétérogénéité intra-tumorale, comme l'analyse de texture, qui souffrent généralement d'un manque d'interprétabilité, nous proposons une méthode de traitement et d'analyse d'image visant à fournir une représentation intuitive de l'hétérogénéité. Ainsi, à partir de cartes paramétriques, nous avons développé des indices évaluant l'hétérogénéité comme la complexité cumulée des intensités des pixels et de leur organisation spatiale dans différentes sous-régions tumorales. Ces régions, identifiées par un processus de clustering, caractérisent l'hétérogénéité des intensités du paramètre d'imagerie mesuré. L'organisation spatiale de ces sous-régions et des éléments qui les constituent est alors analysée et quantifiée afin de fournir un unique indice de l'hétérogénéité à l'échelle tumorale et à l'échelle des clusters. Au travers de différentes applications numériques, précliniques et cliniques, les performances de ce nouvel indice ont été caractérisées en accord avec les recommandations internationales relatives au développement de biomarqueurs d'imageries. Ces applications consistaient tout d'abord en une preuve de concept sur des données d'IRM provenant d'un modèle tumoral chez le petit animal. Elle a illustré l'intérêt clinique éventuel de notre indice pour le suivi de l'efficacité thérapeutique dans le contexte des thérapies ciblées. Puis, différentes applications à des données d'IRM cliniques ont permis d'identifier des apports cliniques potentiels notamment au regard de la standardisation des protocoles d'acquisitions et d'analyses des images et ainsi de progresser dans le processus de validation de cet indice comme biomarqueurs d'imagerie. L'ensemble des résultats obtenus durant cette thèse sont prometteurs et permettent d'envisager une utilisation de cette méthode dans un plus large cadre d'investigations cliniques.