

**THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE
FRANCHE-COMTÉ**

PRÉPARÉE À L'UNIVERSITÉ DE BOURGOGNE

École doctorale n°37

Sciences Pour l'Ingénieur et Microtechniques

Doctorat d'Instrumentation et Informatique de l'Image

par

NÉRÉE PAYAN

Imagerie TEP au ^{18}F -FDG du cancer du sein : étude combinée de la perfusion tumorale et du métabolisme glucidique pour la caractérisation tumorale et la prédiction de la réponse à la chimiothérapie néoadjuvante

Thèse présentée et soutenue à Dijon, le 24 juin 2020

Composition du Jury :

DURAND EMMANUEL

HATT MATHIEU

SARRUT DAVID

COCHET ALEXANDRE

PRESLES BENOIT

VRIGNEAUD JEAN-MARC

Rapporteur

Rapporteur

Président du jury

Directeur de thèse

Co-encadrant de thèse

Co-encadrant de thèse

Résumé : La chimiothérapie néoadjuvante (CNA) est un traitement fréquemment proposé aux patientes présentant des tumeurs mammaires localement avancées ou volumineuses au diagnostic. L'obtention d'une réponse histologique complète (pCR) à l'issue de la CNA est un facteur de bon pronostic lié à une diminution du risque de récurrence, c'est pourquoi sa prédiction est devenue un objectif clé pour l'orientation de nouvelles stratégies thérapeutiques. Des études ont montré que les modifications précoces du métabolisme glucidique, évaluées par Tomographie par Émission de Positons (TEP) au 2-desoxy-2-18F-fluoro-D-glucose (¹⁸F-FDG) réalisée avant traitement et après une cure de CNA, permettaient au moins partiellement de prédire cette réponse. Toutefois, compte tenu de la diversité des cancers du sein et de la complexité des mécanismes qui sous-tendent à la réponse tumorale, il s'avère nécessaire d'avoir une approche multiparamétrique. Or, la TEP au ¹⁸F-FDG, qui est un examen de référence en imagerie médicale pour quantifier le métabolisme et la viabilité des tumeurs cancéreuses, permet également d'appréhender la perfusion tumorale sous couvert d'une imagerie dynamique acquise immédiatement après administration du radiotracer. Par ailleurs, l'étude de l'hétérogénéité tumorale, métabolique comme perfusionnelle, permet d'obtenir une caractérisation toujours plus poussée des lésions. L'objectif de cette thèse a donc été d'évaluer l'apport combiné du métabolisme et de la perfusion tumorale évaluée par la TEP au ¹⁸F-FDG réalisée avant traitement, pour la prédiction de la réponse à la CNA, en l'associant également à d'autres paramètres cliniques et biologiques ainsi qu'à des paramètres de texture caractérisant l'hétérogénéité tumorale. Les études ont été réalisées à l'aide d'une cohorte de 246 patients. Dans une première analyse nous avons évalué l'impact que pourraient avoir différentes méthodes de calcul des indices de texture sur leurs relations avec la réponse à la CNA. Les résultats ont ainsi permis d'identifier qu'une discrétisation relative (DR) de l'image semble plus appropriée pour identifier les patientes bonnes répondeuses. Dans une seconde étude, les relations entre les paramètres d'hétérogénéité tumorale avec les caractéristiques biologiques et moléculaires des tumeurs ont été évaluées. Les résultats ont montré une perfusion tumorale maximale significativement plus élevée pour les tumeurs de stade T3 et T4 et également chez les patients présentant un envahissement métastatique lymphatique régional (N+), sans qu'aucune différence métabolique significative ne soit observée. Par ailleurs, une différence significative d'hétérogénéité de perfusion a été notée entre les phénotypes tumoraux bien qu'aucune différence de perfusion globale n'ait été constatée. Enfin, une analyse de la prédiction de la réponse histologique à la CNA a été réalisée dans une dernière étude, à l'aide de modèles pronostiques utilisant la régression logistique et une sélection de variable univariée. Les meilleures performances de prédiction ont été notées pour les modèles combinant les paramètres cliniques et métaboliques (conventionnels et de texture), les résultats de perfusion ne permettant pas, a priori, d'améliorer la prédiction.

Abstract: Neoadjuvant chemotherapy (NAC) is a common treatment in patients with locally advanced or large breast cancer at diagnosis. A pathological complete response (pCR) at the end of NAC is recognized as a good surrogate marker of relapse-free survival. An early identification of the pathological response has then become a key parameter to monitor new therapeutic strategies. Studies, focusing on predictive biomarkers identification, have shown that early changes in tumour metabolism, assessed by Positron Emission Tomography (PET) using 2-desoxy-2-18F-fluoro-D-glucose (¹⁸F-FDG), allow the early assessment of the pathological response at the end of treatment. However, given the diversity of breast cancers and the complexity of the mechanisms underlying the tumour response, a multi-parametric approach is necessary. ¹⁸F-FDG PET, the gold standard for in vivo evaluation of tumour metabolism, also provides estimation of tumour blood flow using a dynamic acquisition performed immediately after injection. Furthermore, the evaluation of tumour heterogeneity, for both metabolism and perfusion, brings new insights into breast lesions characterization. The main objective of this thesis was then to evaluate the combined tumour metabolism and perfusion information, assessed by ¹⁸F-FDG PET performed before treatment, in order to predict pathological response to NAC. These biomarkers were also associated with clinical and biological parameters as well as texture parameters to quantify tumour heterogeneity. All the studies were conducted using a cohort of 246 patients. In a first analysis, we evaluated the impact of several texture features calculation methods on their relationships with the pathological response to NAC. The results showed that a relative rescaling for the PET images appears more appropriate to identify good responders. In a second study, conventional metabolism and perfusion parameters with their heterogeneity features were analysed according to tumour biological characteristics and molecular subtypes. Results showed that T3 and T4 stage tumours and tumours with lymph node involvement presented a higher perfusion, whereas no significant differences in SUV max or SUV mean were reported. In addition, tumour blood flow heterogeneity was different among breast cancer phenotypes and this was not previously observed with a global analysis of the perfusion. Finally, a treatment response prediction analysis was performed in a last study, with prognostic models using logistic regression and univariate feature selection. The best prediction performances were noted for models combining clinical and metabolic parameters (conventional and textural), whereas the tumour blood flow did not apparently improve the prediction.

REMERCIEMENTS

En tout premier lieu je tiens à remercier mon directeur de thèse, Pr. Alexandre Cochet, pour m'avoir accordé sa confiance et donné l'opportunité de réaliser cette thèse. Vous m'avez accompagnée depuis mon arrivée, en étant toujours disponible et à l'écoute. Pour votre précieux enseignement, pour le temps passé à me guider dans mes recherches, pour vos conseils tout au long de cette thèse et votre bienveillance, je vous témoigne ma sincère reconnaissance.

Je tiens également à exprimer mes remerciements à mes co-encadrants Benoît Presles et Jean-Marc Vrigneaud. Merci Benoît pour ta grande disponibilité, ton accompagnement constant, tes conseils et tes relectures toujours très complètes. Je te remercie également pour ta franchise et ton investissement dans cette thèse. Nos discussions constructives m'ont aidée dans mon travail, merci pour tout le temps que tu y as consacré.

Jean-Marc je te remercie pour ton enseignement, ta précieuse expertise et tes nombreux conseils. Merci également pour ta patience, ta bienveillance et tes encouragements permanents. Réaliser cette thèse sous ton encadrement fut extrêmement enrichissant, inspirant et motivant.

Je considère que ce fut une réelle chance d'avoir un directeur et des encadrants aussi présents, disponibles et à l'écoute. Vous m'avez tous permis de m'améliorer, de me former et de travailler à la frontière de trois domaines scientifiques différents et complémentaires. Merci pour la confiance que vous m'avez accordée et pour m'avoir permis de réaliser cette thèse au sein du laboratoire ImViA.

Je tiens aussi à exprimer ma sincère reconnaissance à David Sarrut qui m'a fait l'honneur de présider le jury de soutenance, ainsi qu'à Emmanuel Durand et Mathieu Hatt pour avoir accepté d'être rapporteurs des travaux de cette thèse. Merci à tous pour le temps passé à lire ce manuscrit et pour vos analyses, conseils et remarques.

Je souhaite également remercier chaleureusement le Pr. François Brunotte d'avoir offert de son temps en relectures et en précieux et sages conseils tout au long de cette thèse. Merci pour votre bienveillance et vos encouragements.

La thèse est une expérience professionnelle, mais également personnelle et synonyme de rencontres. Un grand merci donc à tous les Dijonnais, de souche ou de passage, que j'ai eu la chance de rencontrer et qui ont permis d'ajouter de la joie et de la bonne humeur à cette aventure. Merci à Roberto, Chloé, Cyril, Florian, Orval, Clément, Meriem, Arlsan, Roxane, Romain et tous les autres, pour votre bonne humeur, votre soutien et bien sûr pour toutes les parties de coinche et séances de bloc! Un merci tout particulier à Amira pour son expertise et tout

son temps passé à relire, écouter, conseiller... merci encore ! Merci également aux membres du laboratoire ImViA et du CGFL avec qui j'ai eu le plaisir de travailler.

Ma thèse n'aurait également pas été la même sans la présence d'un fameux Breton. Merci Alan d'avoir pris la peine d'écouter mes doutes et angoisses et d'avoir supporté mes plaintes à rallonge :). Avoir fait cette thèse en parallèle avec la tienne m'a beaucoup apporté, en soutien et conseils mais également en moment partagés.

Je tiens également à remercier mes amis de longue date, Anne, Antoine, Sofiane, Thomas et Julien, toujours présents malgré la distance. Merci pour votre soutien et vos encouragements sans faille. Nos moments passés ensemble sont toujours de réelles bouffées d'air et de joie. À Pauline, Mathou, Flo, Heidi, Sarah. Merci de faire partie de ma vie malgré les longs moments d'absence. Une grande pensée également à tous les autres copains Grenoblois ainsi qu'à toutes les riches rencontres de mon Master 2.

Un grand merci également à mes parents et ma famille, pour leur présence et leur accompagnement, avec une pensée toute particulière pour mes grands-parents. Je tiens également à remercier ma belle-famille pour leurs encouragements permanents.

Pour finir, je te remercie Aurélien, pour toutes les années passées à tes côtés, pour ton soutien indéfectible et ton amour, merci. Tu as été embarqué dans cette aventure qu'est la thèse et tu as su être présent en m'accompagnant tout au long de ces années, en m'encourageant quotidiennement, notamment dans les moments difficiles. Pour tout cela et pour tout le reste, je te remercie.

SOMMAIRE

Introduction	1
1 Cancer du sein et Tomographie par Émission de Positons : présentation et contexte	4
1.1 Le cancer du sein	5
1.1.1 Épidémiologie	5
1.1.2 Classification des cancers du sein	5
1.1.2.1 Classification histologique	6
1.1.2.2 Classification selon le grade histo-pathologique	6
1.1.2.3 Caractérisation et classification biologique des différents sous- groupes du cancer du sein	8
1.1.2.3.1 Récepteurs hormonaux (RH) et HER2	8
1.1.2.3.2 Classification génomique	9
1.1.2.3.3 Classification immunohistochimique	10
1.1.3 Le traitement par chimiothérapie néoadjuvante	10
1.1.3.1 Intérêts de la chimiothérapie néoadjuvante	11
1.1.3.2 La réponse à la chimiothérapie néoadjuvante	12
1.1.4 Évaluation précoce de la réponse au traitement	13
1.1.4.1 Évaluation de la réponse par l'imagerie	13
1.1.4.1.1 Volume tumoral	13
1.1.4.1.2 Métabolisme tumoral	14
1.1.4.1.3 Perméabilité capillaire et perfusion tumorale	14
1.1.4.1.4 La diffusion en IRM	15
1.1.4.1.5 Autres approches	16
1.1.4.2 Prédiction de la réponse	16
1.1.4.3 Hétérogénéité tumorale	17
1.2 La Tomographie par Émission de Positons	17

1.2.1	Principe et fonctionnement de la tomographie par émission de positons . .	17
1.2.1.1	Les radiotraceurs en TEP	18
1.2.1.2	Détection du signal en TEP	19
1.2.1.2.1	Interaction des photons avec la matière	19
1.2.1.2.2	Principe et système de détection	19
1.2.1.2.3	Modes d'acquisitions 2D/3D	22
1.2.1.2.4	Corrections du signal	24
1.2.1.3	La reconstruction de l'image en TEP	26
1.2.1.3.1	Reconstructions 2D/3D	26
1.2.1.3.2	Les méthodes analytiques	27
1.2.1.3.3	Les méthodes itératives	28
1.2.2	Limitations et propriétés de l'imagerie TEP	31
1.2.2.1	Le bruit	31
1.2.2.2	La résolution spatiale	32
1.2.2.3	L'effet de volume partiel	33
1.2.3	La TEP au ^{18}F -FDG en oncologie	34
1.2.3.1	Le 2-desoxy-2- ^{18}F -fluoro-D-glucose ou ^{18}F -FDG	35
1.2.3.2	Métabolisme glucidique et Standardized Uptake Value	36
1.2.3.3	L'étude de la perfusion tumorale via la TEP au ^{18}F -FDG	37
1.2.3.4	La place de la TEP au ^{18}F -FDG dans la prise en charge du cancer du sein	39
1.3	Rationnel de la thèse	41
2	De l'acquisition des données à l'analyse du métabolisme et de la perfusion tumorale	43
2.1	Protocole TREN	44
2.2	Description des paramètres de reconstruction	46
2.2.1	BLOB : modèle de l'image	46
2.2.2	Ordered Subsets (OS) : modèle du système	48
2.2.3	TF : Correction du temps de vol	48
2.3	Perfusion tumorale et imagerie dynamique	49
2.3.1	Création de l'image de perfusion	49

2.3.2	Acquisition des données et reconstruction des images dynamiques	50
2.3.3	Délinéation des régions d'intérêt	50
2.3.4	Identification du pic artériel et création de l'image paramétrique de perfusion	51
2.3.5	Reconstruction dynamique et statistique dans l'image	53
2.4	Métabolisme et volume tumoral	55
2.4.1	Segmentation des images en TEP	55
2.4.1.1	Les méthodes par seuillage	56
2.4.1.2	Les méthodes basées sur la détection de contours	57
2.4.1.3	Les méthodes de classification	58
2.4.2	Détermination des paramètres nécessaires à la méthode de seuillage adaptatif et segmentation du volume métabolique	61
2.4.3	Fantôme NEMA-IEC	61
2.4.4	Seuillage adaptatif de Schaefer <i>et al.</i> : étalonnage sur fantôme	62
2.4.5	Seuillage adaptatif de Schaefer <i>et al.</i> : application aux images cliniques . .	63
3	Analyse de l'hétérogénéité métabolique et perfusionnelle des tumeurs mammaires	67
3.1	Hétérogénéité tumorale : origine et motivations	68
3.1.1	Hétérogénéité tumorale et oncogenèse	68
3.1.2	Caractérisation de l'hétérogénéité tumorale	69
3.2	Analyse de texture	70
3.2.1	Principe et application de l'analyse de texture en imagerie médicale	70
3.2.1.1	Introduction à l'analyse de texture	70
3.2.1.2	Protocole d'application de l'analyse de texture aux images médicales	70
3.2.2	Matrices et paramètres de texture	73
3.2.2.1	Indices d'intensité	73
3.2.2.2	Indices morphologiques	73
3.2.2.3	Matrices de texture	75
3.2.3	Analyse de texture en TEP : applications et recommandations	78
3.2.3.1	Influence des paramètres d'acquisition, reconstruction et de traitement des images	79
3.2.3.1.1	Répétabilité des indices (études test-retest)	79

3.2.3.1.2	Reproductibilité des indices	80
3.2.3.2	Influence de la méthode de délinéation	81
3.2.3.3	Influence de la méthode de discrétisation	82
3.2.3.4	Recommandations vis à vis du volume tumoral	85
3.2.3.5	Choix méthodologiques pour les études avec les données de la base TREN	86
3.3	Etude comparative des méthodes de discrétisation	86
3.3.1	Rationnel et objectifs	86
3.3.2	Matériel et méthodes	87
3.3.2.1	Population étudiée et protocole d'acquisition des données	87
3.3.2.2	Discrétisation des images TEP	88
3.3.2.3	Extraction des paramètres des images	89
3.3.2.4	Analyse statistique	90
3.3.3	Résultats	90
3.3.3.1	Caractéristiques de la cohorte	90
3.3.3.2	Influence de la méthode de calcul des indices de texture sur leurs relations avec les paramètres TEP conventionnels	97
3.3.3.3	Différences des paramètres TEP entre les tumeurs des patientes pCR et non-pCR et influence de la méthode de discrétisation : analyse Mann-Whitney	98
3.3.3.4	Différences de texture entre les tumeurs des patientes pCR et non-pCR et influence de la méthode de discrétisation : analyse factorielle discriminante descriptive	100
3.3.4	Discussion et conclusion de l'étude	103
3.3.4.1	Influence de la méthode de discrétisation et du calcul des GLCM sur les indices de texture	103
3.3.4.2	Influence de la méthode de calcul des indices de texture dans la différentiation des patientes pCR et non-pCR	105
3.3.4.3	Méthode de discrétisation et objectif clinique	106
3.3.4.4	Limitations et conclusion de l'étude	106
3.4	Relations entre l'hétérogénéité tumorale et les paramètres biologiques et molé- culaires des tumeurs	107
3.4.1	Motivations et objectifs	107
3.4.2	Article	108

4	Évaluation de la prédiction de la réponse au traitement	123
4.1	Des données patients à la prédiction	124
4.1.1	Principe du Machine Learning	124
4.1.2	Préparation et sélection des données d'entrée	126
4.1.2.1	Préparation des données	126
4.1.2.2	Sélection des données	127
4.1.2.3	Gestion des données déséquilibrées	129
4.1.3	Algorithmes d'apprentissages	129
4.1.4	Stratégies d'apprentissages	130
4.1.5	Evaluation de l'erreur de prédiction	132
4.2	Prédiction de la réponse à la CNA pour les patientes atteintes d'un cancer du sein	135
4.2.1	Objectif principal	135
4.2.2	Matériel et méthodes	136
4.2.2.1	Population étudiée et protocole d'acquisition des données	136
4.2.2.2	Description des données	136
4.2.2.3	Sélection des variables	137
4.2.2.4	Stratégie d'apprentissage	138
4.2.2.5	Analyse statistique	139
4.2.3	Résultats	139
4.2.3.1	Caractéristiques de la cohorte	139
4.2.3.2	Courbes d'apprentissages et sélection de variables	141
4.2.3.3	Scores d'entraînement et de généralisation des différents modèles	142
4.2.3.4	Contribution des variables de texture	147
4.2.3.5	Récapitulatif des résultats d'apprentissage	148
4.2.4	Discussion et conclusion de l'étude	150
4.2.4.1	Performances de prédiction et apport des données de perfusion .	150
4.2.4.2	Performances de prédiction vis à vis de la littérature	151
4.2.4.3	Limites et conclusion de l'étude	153
5	Discussion et perspectives	155
6	Liste des publications	160

<i>SOMMAIRE</i>	viii
Bibliographie	205
Annexes	213
A Indices de texture de la GLCM	215
B Tableaux de l'étude sur la méthode de discrétisation	217
C Analyse sur l'amélioration de la qualité des images de faible statistique	220
C.1 Métriques	220
C.2 Segmentation des volumes d'intérêts	221
C.3 Caractérisation du biais induit par la statistique des images TEP	223
C.4 Amélioration de la qualité des images de faible statistique	226
C.4.1 Paramètres de reconstruction	227
C.4.2 Correction de l'effet de volume partiel	230
C.4.2.1 Détermination de la résolution spatiale du système	231
C.4.2.2 Analyse des résultats des corrections de l'EVP	233
D Protocole TREN	237

LISTE DES ABRÉVIATIONS

Acc	Accuracy
AT	Analyse de Texture
AUC	Area Under the Curve
BF	Blood Flow
Bq	Becquerel
CNA	Chimiothérapie Néo-Adjuvante
CV	Cross Validation
DA	Discrétisation Absolue
DCE-MRI	Dynamic Contrast Enhanced Magnetic Resonance Imaging
DR	Discrétisation Relative
ER	Estrogen receptor
EGFR	Epidermal Growth Factor Receptor
EVP	Effet de Volume Partiel
FOV	Field Of View
GLCM	Gray Level Co-occurrence Matrix
GLRLM	Gray Level Run-Length Matrix
GLSZM	Gray Level Size-Zone Matrix
HER2	Human Epidermal Growth Factor Receptor-2
IRM	Imagerie par Résonance Magnétique
LOOCV	Leave One-Out Cross Validation
LOR	Line Of Response
LC	Learning Curve
MATV	Metabolic Active Tumor Volume
MCC	Matthews correlation coefficient
ML	Machine Learning
NEMA	National Electrical Manufacturers Association
NGDLM	Neighborhood Gray-Level Difference Matrix
pCR	Pathological Complete Response
PR	Progesterone receptor
RH	Récepteurs hormonaux
ROI	Region Of Interest
SBR	Scarff-Bloom-Richardson
SMOTE	Synthetic Minority Oversampling Technique
SUV	Standardized Uptake Value
TEP	Tomographie par Émission de Positons
TEMP	Tomographie par Émission MonoPhotonique

TN True Negative

TOF Time Of Flight

TP True Positive

TSM Taux (d'incidence ou de mortalité) standardisé sur la population mondiale

TDM TomoDensitoMétrie

VOI Volume Of Interest

VM Volume Métabolique

VP Volume Perfusé

INTRODUCTION

Le cancer du sein est un problème de santé publique majeur dans la mesure où il constitue la première cause de décès par cancer chez la femme en France et dans le monde. Pour les tumeurs localement avancées ou volumineuses au diagnostic, il est fréquent de proposer aux patientes un traitement systémique de Chimiothérapie Néoadjuvante (CNA) en amont du traitement chirurgical. La CNA vise à réduire la taille tumorale afin de permettre une chirurgie mammaire conservatrice non envisageable initialement et permet de surcroît d'évaluer la chimiosensibilité de la tumeur. L'obtention d'une réponse histologique complète (pCR) à l'issue de la chimiothérapie étant un facteur de bon pronostic lié à une diminution du risque de récurrence, ce paramètre est devenu un bon marqueur de substitution permettant d'orienter les nouvelles stratégies thérapeutiques. De ce constat est née la nécessité d'identifier des marqueurs cliniques, radiologiques, biologiques ou moléculaires, prédictifs de la réponse à la CNA. Plusieurs approches utilisant des marqueurs cliniques ou radiologiques ont été utilisées. Toutefois, les modifications fonctionnelles précédant généralement les modifications morphologiques, l'intérêt pour l'imagerie fonctionnelle et moléculaire a rapidement émergé.

Dès les années 2000, il a été montré que les modifications précoces du métabolisme glucidique, évaluées par Tomographie par émission de Positons (TEP) au 2-desoxy-2-¹⁸F-fluoro-D-glucose (¹⁸F-FDG) en cours de CNA, pouvaient permettre d'évaluer précocement (après une ou deux cures) la réponse histologique en fin de traitement [1, 2]. Ces premiers résultats ont ensuite été pondérés selon le phénotype tumoral et le type de traitement [3]. Ces premières constatations ont été confirmées par de multiples études monocentriques et par quelques études multicentriques [4, 5]. Toutefois, cette évaluation précoce de la réponse métabolique s'est révélée insuffisante pour être pleinement intégrée dans les stratégies thérapeutiques. Compte tenu de la diversité des cancers du sein et de la complexité des mécanismes qui sous-tendent à la réponse tumorale, il s'avère nécessaire d'avoir une approche multiparamétrique, intégrant l'étude combinée de paramètres cliniques, biologiques, fonctionnels, moléculaires, voire génotypiques. Dans cette optique, notre équipe a développé une méthode d'évaluation combinée du métabolisme et de la perfusion tumorale à partir d'une imagerie TEP au ¹⁸F-FDG de premier passage. Néanmoins, l'utilité de cette combinaison comme marqueur précoce de la réponse à la CNA n'a, pour le moment, que très peu été étudiée.

Par ailleurs, un nouveau domaine dans l'analyse des images médicales s'est développé : la "radiomique". Cette dernière peut être définie comme le développement de méthodes d'extractions d'un grand nombre de caractéristiques descriptives tumorales, plus ou moins visibles à l'oeil nu. La radiomique englobe un grand nombre de paramètres quantitatifs, dont l'un

des sous-ensembles repose sur les outils mathématiques de l'analyse de texture qui permet l'évaluation de l'hétérogénéité de la totalité du volume tumoral de manière non-invasive. De nombreuses études ont employé la radiomique en oncologie, y compris dans le cas des cancers mammaires, où l'hétérogénéité du métabolisme tumoral a été évaluée et mise en relation avec d'autres paramètres cliniques, biologiques ou moléculaires. Sa valeur prédictive et pronostique a également été étudiée. Néanmoins, dans aucune de ces études l'analyse de l'hétérogénéité de la perfusion tumorale n'a été couplée avec l'information du métabolisme.

Enfin, un autre domaine, celui de l'intelligence artificielle, qui englobe les apprentissages dits automatiques (machine learning) et profonds (deep learning), permet une étude plus poussée des paramètres dans leur ensemble (biologie, génomique, imagerie...). Des modèles de prédiction appliqués aux données médicales ont vu le jour et leurs performances se sont accrues ces dernières années avec l'apport des informations de la radiomique.

La prédiction de la réponse à la CNA dans le cadre du cancer du sein n'a pas fait exception, l'hétérogénéité métabolique tumorale ayant été testée dans des modèles prédictifs. Toutefois, là encore, parmi toutes les études ayant été réalisées à partir des images TEP, aucune n'a évalué l'apport combiné du métabolisme glucidique et de la perfusion tumorale dans la prédiction de la réponse à la CNA.

Objectif principal de la thèse et organisation du manuscrit

L'objectif principal de cette thèse est d'étudier dans le cas du cancer du sein, l'apport de l'étude combinée du métabolisme et de la perfusion tumorale évaluée par la TEP au ^{18}F -FDG réalisée avant traitement, dans la prédiction de la réponse à la chimiothérapie néoadjuvante, en l'associant également à d'autres paramètres cliniques et biologiques.

Cette analyse a été effectuée par étapes, avec dans un premier temps deux analyses descriptives préliminaires suivies d'une analyse prédictive réalisée avec une méthode d'apprentissage automatique. L'hétérogénéité à la fois du métabolisme mais également de la perfusion ont été étudiées et utilisées comme informations complémentaires.

Ce manuscrit est organisé en quatre parties. Le premier chapitre expose le contexte clinique en définissant les caractéristiques générales des cancers du sein, les principaux éléments de leur prise en charge et les critères d'évaluation de la réponse précoce au traitement néoadjuvant. Il est également rappelé le principe de l'imagerie TEP en spécifiant lorsqu'il est nécessaire, les caractéristiques propres à l'imageur utilisé pour nos études. L'imagerie de perfusion par la TEP au ^{18}F -FDG est également présentée dans cette partie tout comme l'évaluation du métabolisme glucidique. Le rationnel de la thèse est développé en fin de chapitre.

Le second chapitre a pour objectif de décrire les grandes étapes qui ont permis la création de la base de données grâce à laquelle les études présentées dans ce manuscrit ont pu être réalisées. Pour cela, dans un premier temps sont présentés le protocole clinique et le

protocole d'acquisition et de reconstruction des images TEP puis dans un second temps sont détaillés les méthodes de reconstruction des images paramétriques de perfusion et de métabolisme.

Le troisième chapitre porte sur la caractérisation des tumeurs mammaires pour lesquelles les informations de la perfusion et du métabolisme tumoral sont étudiées. Dans cette partie sont exposés l'intérêt de la caractérisation de l'hétérogénéité tumorale en oncologie et le principe de l'analyse de texture des images TEP. Le chapitre comporte ensuite deux études descriptives. La première porte sur l'optimisation de la méthode de calcul des indices de textures dans le contexte clinique de la réponse à la CNA pour des cancers mammaires. Les résultats nous permettent de statuer sur une méthodologie à appliquer pour la seconde étude. Cette dernière cible plus spécifiquement les relations entre les indices d'hétérogénéité de perfusion et de métabolisme des tumeurs, avec les paramètres biologiques et les sous-types moléculaires. Cette étude est présentée sous forme d'article car déjà publiée dans *European Journal of Nuclear Medicine and Molecular Imaging*.

Enfin dans le quatrième et dernier chapitre, l'apport combiné du métabolisme et de la perfusion tumorale pour la prédiction de la réponse histologique complète à la CNA ont été évalués. Pour cela, dans une première partie sont exposés le principe de l'apprentissage automatique et les principales étapes nécessaires à son application. Dans une seconde partie, la méthodologie et les résultats de la prédiction de la réponse à la CNA sur notre base de données sont détaillés et discutés.

CANCER DU SEIN ET TOMOGRAPHIE PAR ÉMISSION DE POSITONS : PRÉSENTATION ET CONTEXTE

Sommaire

1.1 Le cancer du sein	5
1.1.1 Épidémiologie	5
1.1.2 Classification des cancers du sein	5
1.1.3 Le traitement par chimiothérapie néoadjuvante	10
1.1.4 Évaluation précoce de la réponse au traitement	13
1.2 La Tomographie par Émission de Positons	17
1.2.1 Principe et fonctionnement de la tomographie par émission de positons	17
1.2.2 Limitations et propriétés de l'imagerie TEP	31
1.2.3 La TEP au ^{18}F -FDG en oncologie	34
1.3 Rationnel de la thèse	41

Ce premier chapitre a pour objectif de décrire le contexte médical dans lequel s'inscrivent les différents travaux de cette thèse. Dans une première partie, une brève présentation du cancer du sein, de ses caractéristiques phénotypiques et de leur classification est faite afin de pouvoir identifier la population étudiée. La stratégie thérapeutique néoadjuvante est ensuite introduite afin d'amener aux enjeux cliniques que sont l'évaluation précoce et la prédiction de la réponse au traitement. Dans une seconde partie, la Tomographie par Émission de Positons (TEP) est présentée, du radiopharmaceutique à la reconstruction de l'image clinique finale. Les caractéristiques de la TEP seront exposées de manière générale en spécifiant à chaque fois en complément les caractéristiques de l'imageur avec lequel les acquisitions ont été réalisées. Deux paramètres biologiques extraits des images TEP sont ensuite présentés : le métabolisme glucidique et la perfusion tumorale. Ces deux paramètres vont être utilisés pour tous les travaux décrits dans ce manuscrit. Enfin dans une troisième partie sont résumés le rationnel de l'étude et les différents objectifs de cette thèse qui seront développés dans les chapitres suivants.

1.1/ LE CANCER DU SEIN

1.1.1/ ÉPIDÉMIOLOGIE

Le cancer du sein est le cancer le plus fréquent et la première cause de décès par cancer chez la femme en France et dans le monde [6]. Selon les estimations de l'Institut National du Cancer (INCa), 58 459 nouveaux cas ont été recensés en France en 2018 avec 12 146 décès pour cette même année. L'âge médian au diagnostic et au décès étaient respectivement de 63 et 74 ans en 2018. L'incidence du cancer du sein (standardisée à la population mondiale) a presque doublé entre 1990 et 2018, passant de près de 30 000 à plus de 58 400 cas annuels, hausse attribuable à l'augmentation du risque mais également à l'augmentation et au vieillissement de la population.

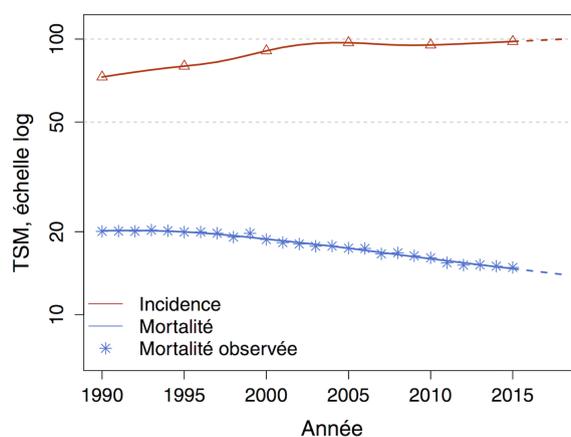


FIGURE 1.1 – Taux d'incidence et de mortalité en France selon l'année (taux standardisé monde TSM). Source : Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine de 1990 à 2018. Volume 1 - Tumeurs solides

Le taux de mortalité (standardisée à la population mondiale) suit quant à lui une évolution inverse, avec une diminution moyenne entre 1990 et 2018 de -1,3% par an et de -1,6% par an sur la période plus récente depuis 2010. Il reste néanmoins difficile, pour expliquer cette baisse de mortalité, de faire la part entre l'impact des diagnostics plus performants et plus précoces apportés par le dépistage et l'amélioration des techniques de traitements adjuvants et de leurs indications.

1.1.2/ CLASSIFICATION DES CANCERS DU SEIN

Le cancer du sein est une maladie hétérogène, non seulement au niveau morphologique, mais également au niveau biologique. En pratique, un examen anatomopathologique permet d'établir le diagnostic de malignité à partir d'une biopsie ou d'une pièce opératoire. Si la malignité est confirmée, la tumeur mammaire est alors classée selon son histologie, son grade et son sous-groupe moléculaire. Les types et grades histologiques ainsi que les sous-groupes tumoraux se caractérisent par des différences dans la prise en charge thérapeutique, la réponse aux traite-

ments et dans la survie des patients.

1.1.2.1/ CLASSIFICATION HISTOLOGIQUE

Le type histologique du cancer du sein va être un élément important pour la conduite thérapeutique et l'évaluation du pronostic. Les types les plus fréquents sont les adénocarcinomes, tumeurs qui se développent à partir des cellules épithéliales de la glande mammaire. La plupart des adénocarcinomes sont développés à partir des canaux galactophores (adénocarcinomes canaux) et plus rarement des lobules (adénocarcinomes lobulaires). Ils représentent à eux seuls la quasi-totalité des tumeurs malignes du sein (95%) [6]. Parmi ces cancers, on distingue les cancers *in situ* (confinés dans les canaux ou les lobules) et les cancers infiltrants (fig. 1.2) :

- Les carcinomes *in situ*

Les carcinomes *in situ* correspondent à une prolifération des cellules cancéreuses sans rupture de la membrane basale. Il n'y a donc théoriquement pas de risque d'envahissement ganglionnaire ou métastatique. Le carcinome canalaire est le plus fréquent. Il représente huit à neuf cancer *in situ* sur dix. Il se définit par la prolifération de cellules épithéliales malignes, confinées à l'intérieur des canaux galactophores et sans envahissement du tissu conjonctif voisin. Le carcinome ductal est quant à lui plus rare et ne représente seulement que 10 à 15% des cancers *in situ*. C'est un carcinome qui est localisé dans les canaux intra-lobulaires mammaires et qui ne présente pas d'envahissement du tissu conjonctif voisin. Les localisations bilatérales et multiples sont courantes.

- Les carcinomes infiltrants (ou invasifs)

Ils représentent la plupart des cancers du sein. Le cancer devient infiltrant à partir du moment où les cellules cancéreuses ont traversé la membrane basale et sont présentes dans les tissus avoisinants. Les cellules peuvent alors entrer en contact avec des vaisseaux sanguins et lymphatiques et être à l'origine d'une diffusion métastatique lymphogène et/ou hémotogène. La plupart des cancers infiltrants sont des adénocarcinomes canaux, mais on comptabilise au total 20 principaux types histologiques dont 18 sous-types mineurs [7].

La population de patientes sur laquelle ont été effectuées les différentes études durant cette thèse concernent uniquement des cancers infiltrants.

1.1.2.2/ CLASSIFICATION SELON LE GRADE HISTO-PATHOLOGIQUE

La biopsie permet à la fois d'identifier le type des tumeurs mais également leur grade. Ce dernier est défini par l'architecture des cellules tumorales, la forme et la taille de leur noyau et leur activité mitotique. Il existe plusieurs grades histopathologiques et la modalité la plus utilisée en France pour les définir dans le cas des cancers invasifs est celle de Scarff-Bloom-Richardson (SBR) modifiée par Elston et Ellis (tab. 1.1) [8]. Cette classification apporte une information pronostique indépendante de la taille tumorale et de l'extension ganglionnaire. A

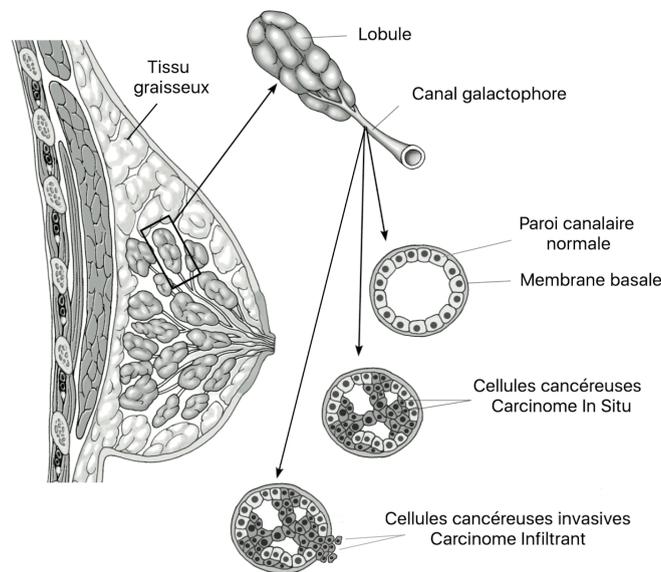


FIGURE 1.2 – Carcinome canalaire *in situ* et infiltrant du sein. Source (adaptée) : www.cancer.org

chaque caractéristique histologique est assigné un score de 1 à 3 dont l'addition détermine le grade histo-pathologique. Les trois grades SBR (bas, intermédiaire et haut) sont définis comme suit : Grade I (score = 3, 4 ou 5), Grade II (score = 6 ou 7) et Grade III (score = 8 ou 9). Les patientes présentant une tumeur de grade I et II ont des taux de survie globale et sans récurrences plus élevées que celles présentant une tumeur de grade III [9]. Le grade a également une valeur prédictive vis à vis de la réponse à la chimiothérapie, avec une meilleure réponse pour les tumeurs de grade III [10].

Grades SBR modifiés selon Eston et Ellis.	
	Score
<i>Formation de tubes :</i>	
la tumeur ne comprend que des tubes (>75%)	1
la tumeur est partiellement tubulaire (10-75%)	2
la tumeur ne comprend aucun tube (<10%)	3
<i>Atypies cytonucléaires :</i>	
noyaux régulier monomorphes	1
atypies modérées	2
noyaux pléiomorphes avec atypies marquées	3
<i>Index mitotique :</i>	
variable selon le diamètre du champ	1 à 3

TABLE 1.1 – Grades SBR modifiés selon Eston et Ellis.

1.1.2.3/ CARACTÉRISATION ET CLASSIFICATION BIOLOGIQUE DES DIFFÉRENTS SOUS-GROUPES DU CANCER DU SEIN

En complément du type histologique et de son grade, deux classifications supplémentaires, la classification génomique et la classification immunohistochimique (IHC), vont permettre de caractériser plus précisément les cancers du sein en mettant en évidence leur hétérogénéité cellulaire. Afin de faciliter la compréhension de ces classifications, une brève description des récepteurs hormonaux et du récepteur de croissance HER2 précède leur introduction.

1.1.2.3.1/ RÉCEPTEURS HORMONAUX (RH) ET HER2

- Les récepteurs hormonaux (RH)

L'oestrogène et la progestérone sont deux hormones femelles qui peuvent favoriser la croissance de certaines cellules, notamment celles de certains cancers du sein dits "hormono-sensibles". Les récepteurs de ces hormones se situent dans le noyau des cellules normales du sein ainsi que dans certains types de cellules cancéreuses. Une fois ces hormones fixées sur leurs récepteurs respectifs, elles peuvent affecter le comportement et la croissance des cellules. L'expression de ces récepteurs aux oestrogènes (*estrogen receptor* (ER)) et à la progestérone (*progesterone receptor* (PR)) est recherchée au sein du tissu tumoral car ils vont renseigner sur le comportement et la potentielle réponse de la tumeur à l'hormonothérapie. En fonction ou non de la présence des récepteurs hormonaux, les tumeurs sont :

- ER positives (ER+) ou ER négatives (ER-)
- PR positives (PR+) ou PR négatives (PR-)

Les tumeurs peuvent être uniquement ER+ ou PR+, mais aussi posséder les deux types de récepteurs.

- Les récepteurs HER2

Human Epidermal growth factor Receptor-2 ou HER2 est le nom plus connu du gène ErbB2 (qui est un proto-oncogène identifié au niveau du chromosome 17q21). Le récepteur associé est présent à la surface des cellules normales où il participe à la croissance et la différenciation cellulaire. Sur les cellules cancéreuses, la possible augmentation anormale de ce récepteur favorise leur prolifération. Les tumeurs surexprimant HER2 sont dites HER-positives (HER2+).

L'évaluation de la surexpression de HER2 est réalisée par deux principales méthodes :

- L'IHC : technique rapide qui marque la protéine surexprimée à la surface des cellules afin d'évaluer le niveau d'expression du récepteur HER2. Le résultat est noté sur une échelle de 0 à 3+. [11]
- L'hybridation *in situ* par fluorescence (FISH) : technique qui évalue l'amplification du gène HER2 présent dans les cellules cancéreuses. Cette méthode plus sensible est également plus onéreuse et est utilisée en cas de surexpression de niveau 2+ en IHC pour confirmer le statut HER2.

Une tumeur est dite HER2+ si son IHC est de 3+ ou s'il est de 2+ avec un FISH positif.

1.1.2.3.2/ CLASSIFICATION GÉNOMIQUE

Les études fondatrices de Perou *et al.* et Sorlie *et al.* [12–14] dans les années 2000 ont permis d'établir des profils d'expression génétique pour les tumeurs mammaires. Les auteurs ont classé ces dernières en cinq principaux sous-groupes :

- Les profils luminal A et B

Ce sont les formes les plus fréquentes des cancers du sein. Les gènes de ces tumeurs codent les protéines des cellules épithéliales de la lumière des canaux ou des lobules du sein, d'où leur nom. Ces tumeurs expriment les marqueurs associés aux récepteurs hormonaux (RH), certaines cytokératines (CK8, CK18) ainsi que des facteurs de transcription spécifiques de différenciation luminaire comme la GATA-binding protein 3 (GATA3) ou encore l'hépatocyte nuclear factor 3 alpha (FOXA1) [15]. Elles sont classées en deux sous-groupes :

- les tumeurs lumineuses A : elles correspondent généralement au génotype le plus fréquent. Elles sont caractérisées par une forte expression des RH contre une faible expression d'HER2 [12, 14].
- les tumeurs lumineuses B : moins fréquentes, elles expriment également fortement les RH et sont aussi caractérisées par l'expression d'EGFR (*epidermal growth factor receptor*) et d'HER2. Ces tumeurs, plus prolifératives que les lumineuses A, sont généralement de plus mauvais pronostic [16, 17].

- Le profil HER2-positif

Les tumeurs HER2 sont caractérisées par une forte expression du gène ErbB2 qui a pour conséquence une surexpression du récepteur HER2 associé. Ces tumeurs tendent à se développer et se propager plus rapidement que les cellules cancéreuses qui expriment HER2 en quantité normale [18]. Des traitements spécifiques existent : des anticorps qui inactivent les récepteurs HER2 en se liant à la région extracellulaire de la protéine HER2 (trastuzumab et pertuzumab) ou des inhibiteurs de la tyrosine kinase des récepteurs EGFR (ErbB1) et HER2 (ErbB2) dans le domaine intra-cellulaire (lapatinib) [19].

- Le profil Basal-like

Ces tumeurs expriment une partie des marqueurs "basaux" tels que les cytokératines CK 5 et 17, CK 14, EGFR, etc. Elles sont caractérisées par une forte expression des gènes prolifératifs, souvent avec une atteinte ganglionnaire au diagnostic et à haut risque métastatique [14]. Comme la plupart de ces cancers n'expriment ni les RH, ni les récepteurs HER2, ils sont généralement associés au phénotype "Triple Négatif" (TN). Le pronostic de ces tumeurs est moins bon que celui du sous-groupe luminaire malgré généralement une bonne réponse aux chimiothérapies [20–22].

- Le profil “Normal-like”

Beaucoup plus rare, ce sous-groupe est également plus difficile à caractériser. Ces tumeurs sont classées, comme les basal-like, par un profil triple négatif. Leur expression génomique est proche de celle observée dans le tissu mammaire normal et dans le tissu adipeux.

1.1.2.3.3/ CLASSIFICATION IMMUNOHISTOCHIMIQUE

Du fait du coût des analyses, l'identification génomique des sous-groupes tumoraux est actuellement peu utilisée en routine clinique. L'immunohistochimie a permis une classification substitutive de ces sous-types moléculaires [18,23,24] en se basant sur les récepteurs hormonaux (RH), le récepteur de facteur de croissance HER2 et sur le marqueur Ki-67 de la prolifération cancéreuse. La classification comporte les cinq sous-types clinico-pathologiques suivants [25,26] :

- Luminal A-like : tumeurs ER+ et PR+, absence de surexpression de HER2 et une expression de Ki-67 faible (<14%) ou intermédiaire (entre 14 et 19%)
- Luminal B-like : tumeurs ER+, absence de surexpression de HER2 et au moins une des deux conditions suivantes :
 - PR négatif ou faible (<20%) et un Ki-67 intermédiaire
 - Une expression de Ki-67 supérieure à 20% [27]
- Luminal B-like/HER2 positif : tumeurs ER+, surexpression de HER2
- HER2 positif (non-luminal) : tumeurs ER- et PR-, surexpression de HER2
- Triple Négatif : tumeurs ER- et PR-, absence de surexpression de HER2

Les patientes atteintes de cancers luminaux, bien que ces derniers soient connus pour être peu chimiosensibles, ont une survie à long terme plus favorable que les patientes atteintes de tumeurs HER2 ou triple négatives [14,28]. Le sous-type TN est celui de plus mauvais pronostic, avec généralement des rechutes précoces et touche plus fréquemment les femmes jeunes [29,30]. On lui associe communément le sous-type basal défini par la génomique. En pratique, même si la majorité des tumeurs TN sont de type basal, plus de 20% d'entre elles sont de profils différents (Claudin low, basal, interferon rich, moléculaire apocrine...) avec des pronostics et des comportements biologiques très différents, faisant des tumeurs triple négatives un sous-groupe très hétérogène [31,32].

Quelle que soit la classification biologique adoptée, la taxonomie du cancer du sein dans son ensemble a donc permis de progresser vers l'individualisation de la prise en charge thérapeutique, du pronostic et de la réponse au traitement.

1.1.3/ LE TRAITEMENT PAR CHIMIOTHÉRAPIE NÉOADJUVANTE

Le diagnostic du cancer du sein est basé à la fois sur l'examen clinique, sur l'anatomopathologie et sur l'imagerie. En effet, en complément de la classification tumorale, le cancer va également être classé selon son stade d'évolution par la classification TNM, qui prend en compte la taille des tumeurs (T), l'atteinte ganglionnaire (N), et l'envahissement métastatique à distance (M). Ce diagnostic va permettre de diriger le clinicien vers les options thérapeutiques adéquates,

l'option principale restant la chirurgie [33].

1.1.3.1/ INTÉRÊTS DE LA CHIMIOTHÉRAPIE NÉOADJUVANTE

La chimiothérapie néoadjuvante est un traitement cytotoxique systémique utilisé dans le cas des cancers infiltrants et effectué avant le traitement chirurgical. Son principal objectif est de réduire le volume tumoral et potentiellement le stade du cancer. Initialement, la CNA a été utilisée pour des cancers du sein localement avancés, non résécables et inflammatoires à des fins d'amélioration de la survie [34–39]. A partir de la fin des années 80, ce traitement a également été proposé afin d'augmenter les possibilités de conservation mammaire. L'objectif, pour les tumeurs résécables de taille importante, était de diminuer au maximum le volume tumoral afin de permettre une chirurgie la moins invasive possible. Cela a notamment permis des chirurgies conservatrices chez des femmes qui initialement nécessitaient une mastectomie. La CNA a également montré de très bons résultats en terme de réponse tumorale [40] et élimine dans certains cas le besoin de dissection des ganglions lymphatiques axillaires, épargnant ainsi aux patientes la morbidité à long terme du lymphœdème associé. En ce qui concerne l'amélioration de la survie pour les tumeurs initialement résécables, elle pourrait être théoriquement justifiée par trois mécanismes :

- le traitement plus précoce des cellules micro-métastatiques avant l'apparition de clones chimio-résistants [41]
- la limitation de la stimulation tumorale par la sécrétion de facteurs de croissance durant l'intervention chirurgicale [42–44]
- l'adaptation de la chimiothérapie à la réponse tumorale : évaluer *in vivo* l'effet de la CNA par l'étude de la réponse tumorale et ganglionnaire afin d'identifier les patientes mauvaises répondeuses.

En pratique, l'application d'une chimiothérapie avant chirurgie n'améliore pas la survie des patientes en comparaison au même traitement administré post chirurgie [45, 46], mais permet néanmoins d'augmenter le taux de chirurgies conservatrices et a été donc proposée comme norme de soin pour le traitement des stades précoces [47]. En effet, la CNA est à ce jour largement utilisée pour traiter les sous-types de cancer du sein Triple Négatif et HER2+, permettant des taux accrus de chirurgie conservatrice accompagnée d'une diminution de la dissection axillaire [48, 49]. Par ailleurs, la valeur pronostique de la réponse histologique à la CNA a, quant à elle, été clairement démontrée [50, 51]. De nombreuses études ont montré que la réponse à la CNA est étroitement corrélée à la survie sans récurrence et la survie globale, avec un plus faible risque de rechute pour les patientes en réponse histologique complète (*pathological complete response* (pCR)) [52–55]. La pCR a également été reconnue comme facteur pronostique pour les cancer localement avancés [56] ainsi que pour les formes opérables de plus faible grade [52, 57]. La réponse histologique à la CNA a donc été par la suite fréquemment utilisée comme marqueur substitutif de la survie dans les études évaluant, en situation néoadjuvante, l'efficacité de thérapies cytotoxiques.

1.1.3.2/ LA RÉPONSE À LA CHIMIOTHÉRAPIE NÉOADJUVANTE

La réponse histologique est évaluée sur la pièce opératoire (tumeur primitive et curage ganglionnaire). Il existe plusieurs classifications pour la caractériser. La classification de Chevallier [58] est actuellement celle de référence et a donc été utilisée pour les différents travaux présentés dans ce manuscrit. La classification de Chevallier évalue à la fois la réponse macroscopique et microscopique de la tumeur ainsi que celle des ganglions axillaires homolatéraux (tab. 1.2).

Classification de CHEVALLIER (1993)
Grade 1 : absence de toute cellule tumorale histologiquement et absence de métastase axillaire
Grade 2 : présence de carcinome <i>in situ</i> dans le sein, sans cellule tumorale invasive et absence de métastase axillaire
Grade 3 : carcinome invasif avec altération stromale
Grade 4 : absence d'altération des cellules tumorales

TABLE 1.2 – Classification de Chevallier pour l'évaluation histologique de la réponse à la chimiothérapie néoadjuvante des cancers du sein.

En pratique, la réponse histologique complète est définie par la disparition totale de toute lésion mesurable, i.e. l'absence de cellules résiduelles invasives dans le site tumoral et ganglionnaire (qu'il s'agisse de la présence de tissu cicatriciel, ou de restitution *ad integrum*). Cette définition correspond aux grades de Chevallier I et II.

- Réponse histologique et phénotype tumoral

La réponse histologique est également fortement influencée par les phénotypes tumoraux pour lesquels les probabilités et les significations pronostiques diffèrent. En effet, malgré leur agressivité, les cancers TN ont été identifiés comme prédictifs d'une bonne réponse histologique à la CNA. Le contraste entre le taux élevé de pCR et l'agressivité de ces tumeurs a été décrit comme "le paradoxe triple négatif" [54,59,60]. Le pronostic des patientes obtenant une pCR est excellent avec moins de 10% de récives à 5 ans, égal aux autres phénotypes en pCR. A l'inverse, les tumeurs TN qui ne sont pas identifiées comme répondeuses à l'issue de la CNA présentent un pronostic beaucoup plus sombre [54,59,60]. Pour le phénotype HER2, bien qu'il soit lui aussi de pronostic péjoratif, il a été montré que le taux de réponse augmentait significativement lors de l'ajout de Trastuzumab (spécifique des récepteurs HER2) à la CNA [61,62]. Enfin, pour le phénotype luminal, l'effet de la CNA est quant à lui controversé. En effet, ces tumeurs moins chimiosensibles présentent un taux de pCR inférieur à 5% à la fin des traitements. Cela n'est cependant pas synonyme de mauvais pronostic pour ce phénotype qui présente un taux de survie plus élevé que celui des tumeurs TN [25,63,64].

1.1.4/ ÉVALUATION PRÉCOCE DE LA RÉPONSE AU TRAITEMENT

L'un des avantages majeurs de la CNA est donc de permettre l'évaluation *in vivo* de la chimiosensibilité de la tumeur via l'analyse, lors du geste chirurgical, de la réponse histologique. Toutefois, la détermination de cette réponse reste tardive car elle est connue uniquement en fin du traitement néoadjuvant. L'analyse des réponses clinique, biologique et d'imagerie pourraient alors apporter des critères plus précoces pour l'évaluation de l'efficacité de la chimiothérapie, d'ordre pronostique. L'idée serait de permettre au clinicien de choisir et d'adapter au mieux son traitement pour pallier aux possibles résistances observées et éviter ainsi la sur-toxicité d'un traitement inefficace. L'évaluation précoce de la résistance au traitement se justifie donc par le besoin d'identifier le plus tôt possible les patientes mauvaises répondeuses, afin d'orienter en conséquence leur traitement de chimiothérapie néoadjuvante vers un traitement alternatif plus personnalisé.

Ainsi, de nombreuses études ont recherché si d'éventuels paramètres biologiques, comme le niveau d'expression des RH, p53, Bcl-2 ou encore le Ki-67, étaient corrélés à une réponse tumorale après CNA. Parmi les paramètres étudiés, les valeurs de Ki-67 sont apparues comme prédictives d'une bonne réponse à la CNA [65–67], sans toutefois aboutir à des modifications de prise en charge. L'examen clinique étant également limité, d'autres études se sont alors intéressées à l'évaluation précoce de la réponse à la CNA par l'imagerie.

1.1.4.1/ ÉVALUATION DE LA RÉPONSE PAR L'IMAGERIE

Les modalités d'imagerie utilisables sont multiples et comprennent notamment l'échographie, l'Imagerie par Résonance Magnétique (IRM) ou encore la tomographie par émission de positons (TEP). L'échographie est la technique couramment utilisée pour évaluer la réponse tumorale. L'IRM est plus généralement utilisée pour les cas difficiles à suivre, telles que les tumeurs présentes dans les seins denses ou les tumeurs multifocales. Enfin, la TEP quant à elle permet l'évaluation des variations précoces du métabolisme tumoral. Afin de prendre une décision thérapeutique la plus précoce possible, plusieurs paramètres ont été étudiés en imagerie : le volume tumoral d'une part et les caractéristiques physiologiques cellulaires évaluées par l'imagerie fonctionnelle et moléculaire d'autre part.

1.1.4.1.1/ VOLUME TUMORAL

L'un des principaux critères de réponse (ou non-réponse) au traitement est la modification du volume tumoral. L'appréciation de la réponse sera faite en fonction du pourcentage de réduction de ce volume (ou de certaines de ses dimensions) constaté au cours du traitement selon les critères RECIST (*Response Evaluation Criteria in Solid Tumors*) formulés en 2000 [68] et mis à jour depuis [69]. Cependant, cette réduction tumorale doit être importante pour statuer sur la réponse au traitement et s'avère moins pertinente pour les évaluations précoces. L'utilisation de l'imagerie moléculaire et fonctionnelle va permettre d'améliorer la précision de l'évaluation de la réponse en renseignant sur les caractéristiques physiologiques cellulaires et vasculaires.

1.1.4.1.2/ MÉTABOLISME TUMORAL

La TEP au ^{18}F -Fluorodésoxyglucose, noté ^{18}F -FDG, permet une mesure in vivo du métabolisme glucidique de la tumeur. Le principe et le fonctionnement de la TEP seront détaillés dans la seconde partie de ce chapitre. L'intérêt de cette imagerie en CNA repose sur le fait que les modifications du métabolisme tumoral sont généralement observées avant les changements anatomiques. De nombreuses études ont ainsi mis en évidence une corrélation entre les variations précoces du métabolisme glucidique des tumeurs et la réponse au traitement néoadjuvant, comme le résumant les méta-analyses sur le sujet [70–72]. D'autres études ont également évalué le rôle de la TEP au ^{18}F -FDG en tant que biomarqueur précoce de la réponse à la CNA en fonction des phénotypes tumoraux [3, 73, 74]. La TEP au ^{18}F -FDG est rapidement apparue comme un outil d'évaluation précoce de la réponse à la CNA en complément des examens conventionnels. La standardisation de l'évaluation de la réponse au traitement par l'imagerie TEP est actuellement établie par les critères PERCIST (*Positron Emission tomography Response Criteria in Solid Tumors*) qui ont été définis en 2009 [75]. L'évaluation de la réponse métabolique précoce par la TEP au ^{18}F -FDG a ensuite été utilisée pour l'étude d'escalade thérapeutique. L'étude AVATAXHER [76] par exemple a proposé un renforcement thérapeutique par ajout de bevacizumab, traitement antiangiogénique, aux patientes traitées par CNA dont le cancer du sein surexprimait HER2 et présentait une réponse métabolique précoce jugée insuffisante. Les résultats de l'étude ont montré que ce changement thérapeutique chez les patientes initialement mauvaises répondeuses augmentait la proportion de celles obtenant une pCR, soulignant ainsi l'intérêt de la TEP au ^{18}F -FDG dans ce type de stratégie thérapeutique.

L'évaluation et la prédiction de la réponse à la CNA par la TEP sont utilisées à l'heure actuelle uniquement dans le cadre de protocoles de recherche et nécessitent davantage de validation pour une pratique en routine clinique.

1.1.4.1.3/ PERMÉABILITÉ CAPILLAIRE ET PERFUSION TUMORALE

Il est possible d'évaluer l'apport sanguin d'une tumeur par différentes approches. Plusieurs données peuvent être quantifiées. La vascularisation tumorale est une mesure reliée au volume sanguin au sein de l'espace intra vasculaire tumoral et peut être quantifiée par une échographie de contraste. La perfusion tumorale quant à elle est reliée au débit sanguin au sein du tissu tumoral et peut être accompagnée de la mesure de la perméabilité capillaire. En IRM, ces données peuvent être mesurées par la délivrance d'un bolus de produit de contraste qui se distribue dans le réseau intra vasculaire, se diffuse dans le secteur interstitiel notamment en fonction de la perméabilité capillaire, puis en ressort vers le secteur plasmatique.

- IRM de perfusion

L'IRM de perfusion (DCE-MRI pour *Dynamic Contrast Enhanced Magnetic Resonance Imaging*) est une technique qui va donc permettre de quantifier la perméabilité capillaire.

Des chélates de gadolinium sont utilisés comme agent de contraste. Ce dernier est administré par voie veineuse et donc véhiculé tout d'abord par voie vasculaire puis qui se diffuse dans les tissus interstitiels. Via cet examen, des études ont pu montrer une corrélation entre les tumeurs très vascularisées et une altération de la survie globale et sans récurrence [77, 78]. L'IRM de perfusion est ainsi utilisée pour mesurer l'efficacité de thérapies anti-angiogéniques (qui limitent la néo-angiogénèse tumorale) ou de thérapies cytotoxiques responsables de l'altération du système vasculaire tumoral [78]. Cependant, l'application de l'IRM de perfusion en situation néoadjuvante a donné lieu à des résultats contradictoires [79–83], Yu *et al.* ayant notamment montré que les modifications de perfusion après un cycle de CNA s'avéraient être très modérés et sûrement trop précoces pour prédire à elles seules la réponse histologique tumorale [82].

- Échographie de contraste

L'échographie en contraste quant à elle permet d'évaluer la vascularisation tumorale via des agents de contraste ultrasonores qui restent confinés dans l'espace intravasculaire. L'évaluation de la perfusion ne dépend donc pas, avec cette technique, de la perméabilité capillaire. Dans le cas du cancer du sein, l'échographie de contraste a été montrée comme étant capable d'évaluer l'angiogénèse tumorale [84] avec des résultats corrélés à ceux de la perfusion tumorale évaluée en IRM dynamique [85]. L'intérêt de l'échographie de contraste dans l'évaluation précoce de la réponse à la CNA n'a fait que croître ces dernières années et présente des résultats encourageants [86, 87]. Cette technique a également pour avantage d'être moins onéreuse, d'avoir une meilleure disponibilité et une meilleure résolution temporelle. Toutefois, la variabilité des mesures, liée notamment à la variabilité inter observateur, reste l'une de ses principales limites [88].

- TEP

L'acquisition dynamique en TEP va permettre d'analyser la biodistribution du traceur au cours du temps. La TEP à l'eau marquée à l'oxygène 15 ($^{15}\text{O-H}_2\text{O}$) est l'outil de référence pour étudier la perfusion tumorale. De manière générale, la modification de la perfusion tumorale mesurée par la TEP à l' $^{15}\text{O-H}_2\text{O}$, avant et pendant la CNA, a été montrée comme étant un marqueur prédictif de la réponse histologique finale ainsi que de la survie [73, 89–93]. Par la suite, des études ont également montré que la TEP au $^{18}\text{F-FDG}$, avec acquisition au premier passage, présentait des résultats très corrélés à ceux obtenus avec l'eau marquée à l'oxygène 15 [94]. En se basant sur ce constat, d'autres travaux ont évalué la perfusion via la TEP au $^{18}\text{F-FDG}$ et ont présenté des résultats encourageants [95], [96]. Le principe de l'évaluation de la perfusion par la TEP au $^{18}\text{F-FDG}$ sera développé plus en détail dans la seconde partie de ce chapitre.

1.1.4.1.4/ LA DIFFUSION EN IRM

De manière générale, la diffusion en IRM permet d'obtenir des images dont le contraste va être influencé par les différences de mobilité des molécules d'eau présentes dans les tissus, au niveau intra et extra cellulaires. Le coefficient apparent de diffusion (ADC pour *Apparent Diffusion Coefficient*) est sensible à la densité cellulaire, à l'état des membranes cellulaires ainsi qu'à la

micro-structure tissulaire [97]. Les tumeurs ayant généralement une densité cellulaire élevée, cette dernière va restreindre la diffusion des molécules d'eau. La diminution de la densité cellulaire tumorale après un cycle de CNA va se traduire par une augmentation des valeurs d'ADC. L'IRM de diffusion a donc été étudiée comme biomarqueur de la réponse précoce à la CNA et a montré des résultats prometteurs [98, 99]. La diminution de la densité cellulaire des tumeurs à pour avantage d'être antérieure à la diminution du volume tumoral [100] et semble par ailleurs très variable d'un phénotype tumoral à un autre [101].

1.1.4.1.5/ AUTRES APPROCHES

D'autres approches moins répandues peuvent également être envisagées pour l'évaluation de la réponse précoce à la CNA. On peut en effet citer la spectroscopie RMN du proton [102–104] par mesure de la choline ou encore l'utilisation d'autre radiotraceurs émetteurs de positons pour la TEP tel que le 18F-fluoroestradiol (¹⁸F-FES). La TEP au ¹⁸F-FES va en effet permettre l'évaluation non invasive de l'expression des récepteurs aux oestrogènes, l'oestradiol pouvant alors être évalué comme potentiel biomarqueur de la réponse précoce aux traitements [105, 106].

1.1.4.2/ PRÉDICTION DE LA RÉPONSE

L'évaluation et la prédiction de la réponse à la CNA sont deux approches qui peuvent sembler similaires. Lorsque l'on parle d'évaluation, on considère l'évolution au cours du traitement d'un ou plusieurs paramètres et en fonction de leur évolution, on va, à l'aide de critères connus comme RECIST ou PERCIST, statuer sur la réponse (ou non réponse) au traitement thérapeutique. La prédiction de la réponse au traitement va également utiliser l'évolution de ces mêmes paramètres afin de prédire, avec une précision plus ou moins élevée, la réponse au traitement. Cette prédiction peut se faire grâce à des méthodes d'apprentissages supervisés ou non supervisés, communément regroupés dans les méthodes dites de "*machine learning*". Le principe de ces algorithmes sera exposé plus en détail dans le chapitre 4. Ce qui est toutefois intéressant de souligner à ce stade du manuscrit, c'est le fait qu'avec ces méthodes, une prédiction puisse être réalisée à l'aide uniquement des données acquises avant le début du traitement (données "*baseline*"). Cette propriété fait partie des aspects qui différencient ces méthodes par rapport à l'évaluation de la réponse aux traitements. Parmi les paramètres utilisés pour la prédiction de la réponse à la CNA, on peut citer :

- les paramètres cliniques, tels que les phénotypes tumoraux, l'âge, le grade tumoral, etc ;
- les paramètres biologiques, comme les marqueurs cytotoxiques, les récepteurs hormonaux, etc ;
- les paramètres d'imagerie, comme le volume, les valeurs de métabolisme, de perfusion tumorale ou encore des valeurs de paramètres de texture permettant l'évaluation de l'hétérogénéité tumorale.

L'intérêt pour la prédiction de la réponse à la CNA n'a fait que croître durant la dernière décennie avec, rien que dans ces trois dernières années, un grand nombre d'études sur le sujet. Si l'on se concentre sur les études incluant des données d'imagerie, plusieurs équipes se sont

intéressées à la prédiction de la réponse à la CNA aussi bien en utilisant l'imagerie d'ultrasons [107–112], l'IRM [113–128], la TEP au ^{18}F -FDG [129–132] ou encore en combinant plusieurs de ces modalités [125, 133–136]. En plus des paramètres d'imagerie conventionnels, la quasi totalité de ces études ont évalué le pouvoir prédictif de l'hétérogénéité tumorale, que ce soit en ultrasons, IRM ou TEP. Cette dernière est évaluée grâce à des méthodes d'analyse de texture. La caractérisation de l'hétérogénéité tumorale ainsi que le principe et l'application de l'analyse de texture seront exposés plus en détail dans le chapitre 3.

1.1.4.3/ HÉTÉROGÉNÉITÉ TUMORALE

L'étude de l'hétérogénéité tumorale est motivée par l'hypothèse que la présence d'hétérogénéité dans les images pourrait refléter certaines caractéristiques tumorales comme la présence de nécrose, de zones très perfusées, ou à forte prolifération cellulaire (chap. 2). Lorsque l'on parle d'hétérogénéité tumorale, il est important de souligner qu'il ne s'agit pas d'une hétérogénéité "absolue" mais de l'hétérogénéité d'un paramètre donné, telle que la captation métabolique de la tumeur. Dans le cas des cancers mammaires, l'hétérogénéité tumorale évaluée en imagerie fonctionnelle s'est avérée être corrélée à la réponse histologique à la CNA [137] et a donc été étudiée comme marqueur prédictif de cette même réponse. Des travaux ont montré que l'étude quantitative des images d'échographie permettait, avec l'analyse des paramètres de texture, d'améliorer la prédiction de la réponse à la CNA et de potentiellement guider la planification du traitement des patientes [108, 109, 138, 139]. De même, l'association de l'hétérogénéité tumorale en IRM de perfusion avec des variables cliniques s'est avérée prédictive de la réponse au traitement [119, 124, 128, 140]. Braman *et al.* (2017) ont par ailleurs montré que l'identification des variables de texture les plus prédictives pouvait varier en fonction des différents sous-types de récepteurs. Enfin, la TEP au ^{18}F -FDG a également permis d'identifier l'hétérogénéité métabolique comme potentiel marqueur prédictif de la réponse à la CNA dans le cancer du sein [129, 132, 134, 141, 142].

1.2/ LA TOMOGRAPHIE PAR ÉMISSION DE POSITONS

1.2.1/ PRINCIPE ET FONCTIONNEMENT DE LA TOMOGRAPHIE PAR ÉMISSION DE POSITONS

La Tomographie par Emission de Positons (TEP) est une technique d'imagerie moléculaire. Le principe de l'imagerie TEP repose sur l'administration au patient d'un radiotracer, généralement constitué d'une molécule vectrice et d'un isotope radioactif, qui va permettre de visualiser et/ou mesurer *in vivo* et de manière non invasive une fonction biologique ciblée. Les particules émises à partir de l'annihilation d'un électron et d'un positon issu de la désintégration de l'isotope radioactif, vont être détectées à leur sortie du patient et vont permettre, après une étape de reconstruction du signal, de générer l'image fonctionnelle. Au final, la construction d'une image TEP dépend de trois éléments clés : le radiotracer, la détection et la reconstruction.

1.2.1.1/ LES RADIOTRACEURS EN TEP

Les radiotraceurs sont des molécules dites “marquées” par un isotope radioactif, qui vont cibler des processus biologiques spécifiques. En TEP, ils doivent présenter certaines caractéristiques, telles que :

- avoir une grande spécificité pour l’organe, la fonction ou la pathologie ciblée afin de minimiser la fixation dans d’autres zones ;
- être la plupart du temps l’analogue d’une molécule présente dans l’organisme et ne pas perturber son comportement chimique ;
- être compatible avec une acquisition clinique : demi-vie de l’isotope radioactif, temps d’assimilation/fixation et d’acquisition.

D’un point de vue physique, les isotopes pour la TEP doivent également être émetteurs β^+ . Leur désintégration va se caractériser par l’émission d’un positon et d’un neutrino. Le positon émis va ensuite se propager dans la matière où il perd peu à peu son énergie cinétique suite aux interactions avec les autres particules du milieu. La distance moyenne parcourue est appelée le libre parcours moyen du positon et dépend de son énergie d’émission. Lorsque ce dernier a déposé la quasi-totalité de son énergie, il va ensuite entrer en collision avec un électron du milieu et s’annihiler. L’annihilation de cette particule avec son antiparticule conduit à la création de deux photons de même énergie (511 keV) émis en quasi opposition ($180^\circ \pm 0.2^\circ$) [143, 144] (fig. 1.3).

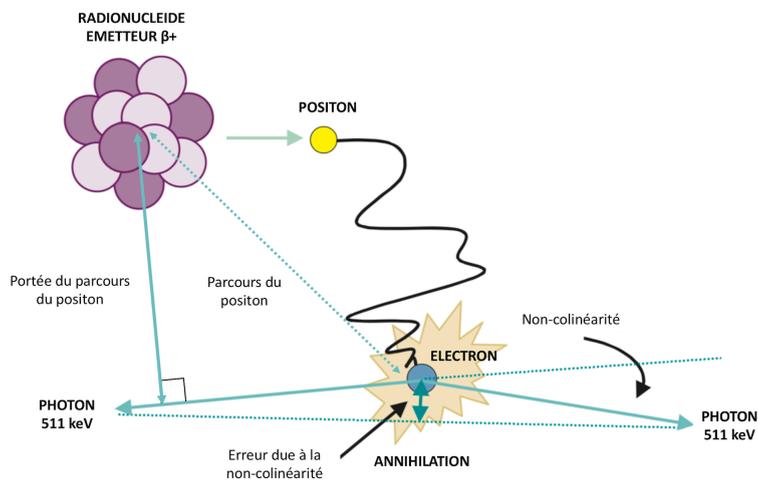


FIGURE 1.3 – Illustration de l’annihilation du positon et de l’électron. [145]

Les radionucléides les plus communs en TEP sont l’oxygène 15 (^{15}O), l’azote 13 (^{13}N), le carbone 11 (^{11}C) et le fluor 18 (^{18}F) dont les périodes, i.e. le temps au bout duquel la moitié des noyaux radioactifs se sont désintégrés, sont respectivement d’environ 2, 10, 20 et 110 minutes. Excepté pour le fluor 18, chacun de ces isotopes nécessite un cyclotron de production proche du lieu d’examen. En pratique, une fois le radiotraceur administré au patient, un certain temps d’attente est nécessaire afin d’obtenir l’accumulation du radiotraceur dans la zone ciblée. Le

temps d'attente va dépendre à la fois du radiotracer utilisé et du phénomène biologique que l'on souhaite observer. Les particules émises à partir de la désintégration de l'isotope radioactif peuvent ensuite être détectées à leur sortie du patient.

1.2.1.2/ DÉTECTION DU SIGNAL EN TEP

1.2.1.2.1/ INTERACTION DES PHOTONS AVEC LA MATIÈRE

Les photons d'annihilation vont interagir dans la matière avec les électrons périphériques du milieu. Aux énergies considérées en TEP, les photons peuvent interagir de trois façons dans la matière : soit par effet photoélectrique, soit par diffusion Compton, soit par diffusion Rayleigh (fig. 1.4). L'effet photoélectrique a lieu lorsque le photon transmet la totalité de son énergie sous forme d'énergie cinétique à un électron qui est alors expulsé du cortège électronique. La diffusion Compton quant à elle a lieu lorsque le photon ne transmet qu'une partie de son énergie à un électron qui est expulsé du cortège électronique, l'autre partie de l'énergie du photon initial étant diffusé sous la forme d'un photon d'énergie plus faible. Enfin, lorsque le photon se diffuse dans la matière sans perte d'énergie, on parle alors de diffusion Rayleigh. L'interaction majoritaire pour des photons de 511 keV dans les tissus est la diffusion Compton [146].

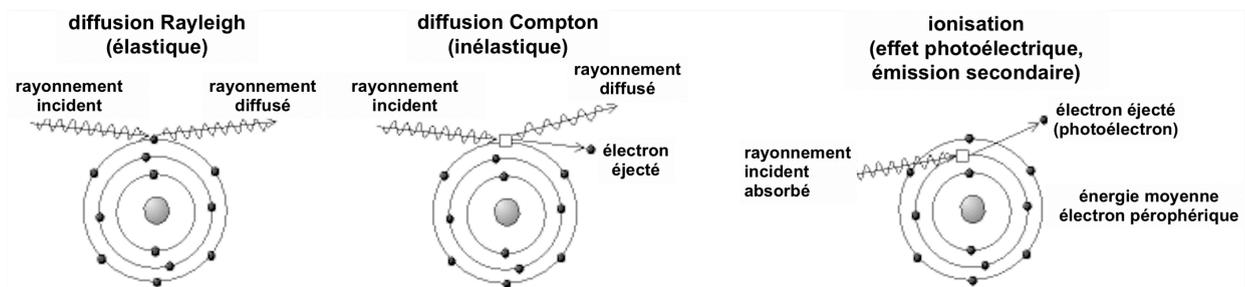


FIGURE 1.4 – Interactions possibles des photons dans la matière en TEP, *source : wikipédia*

1.2.1.2.2/ PRINCIPE ET SYSTÈME DE DÉTECTION

La détection en TEP repose sur les propriétés des photons issus du phénomène d'annihilation. Ces photons d'annihilation émis en quasi opposition vont pouvoir être détectés en coïncidence grâce à un système placé en couronne tout autour du patient (fig. 1.5). Les détecteurs utilisés en TEP sont composés de cristaux scintillateurs couplés à des photodétecteurs. L'énergie déposée lors de l'interaction du photon dans le cristal est transformée en photons lumineux puis convertie en signal électrique mesurable et amplifié par les photodétecteurs. Les cristaux utilisés doivent avoir une densité et un coefficient d'atténuation linéaire élevés afin de maximiser l'interaction des photons de 511 keV dans un petit volume de cristal (on parle de pouvoir d'arrêt du cristal). Ils doivent également avoir un bon rendement lumineux, i.e. un nombre élevé de photons lumineux produits par interaction, afin de faciliter l'analyse des photons incidents et d'améliorer

ainsi les caractéristiques de la résolution spatiale. La résolution en énergie du détecteur, i.e. sa capacité à détecter des rayonnements en énergie légèrement différents, va également impacter sur l'efficacité de détection des photons de 511 keV vis à vis des photons ayant perdu de l'énergie lors de leur diffusion dans la matière. Enfin, les cristaux doivent avoir une réponse rapide i.e. l'émission d'un grand nombre de photons lumineux par unité de temps (ce qui correspond à une constante de décroissance de luminescence faible), afin d'optimiser la résolution temporelle et les performances en taux de comptage du système. Meilleure est cette dernière, plus étroite pourra être la fenêtre temporelle de détection.

Les principaux cristaux utilisés en TEP et leurs caractéristiques sont illustrés dans le tableau 1.3.

cristaux scintillants					
Scintillateur	NaI(Tl)	BGO	GSO	LSO	LYSO
Densité (g/cm ³)	3,7	7,13	6,7	7,4	7,1
Épaisseur d'atténuation 1/ μ à 511 keV (mm)	25,9	11,2	15,0	12,3	12,6
Photofraction (%)	18	42	26	34	33
Décroissance de la scintillation (ns)	230	300	60/600	40	40
Rendement lumineux (photons/MeV)	38000	8200	8500	2600	26000-30000

TABLE 1.3 – Caractéristiques des principaux cristaux scintillants en TEP

Ces couplages cristaux-photodétecteurs sont généralement regroupés en blocs [147] puis juxtaposés pour former un anneau de détection (fig. 1.5). Cette couronne de détecteurs, d'environ 80 à 90 cm de diamètre interne selon les modèles d'imageurs, va permettre de détecter en coïncidence les photons et d'identifier ainsi une ligne de réponse (LOR pour *line of response*) sur laquelle l'annihilation a eu lieu. Plusieurs couronnes peuvent être alignées afin d'obtenir un cylindre de détection plus étendu.

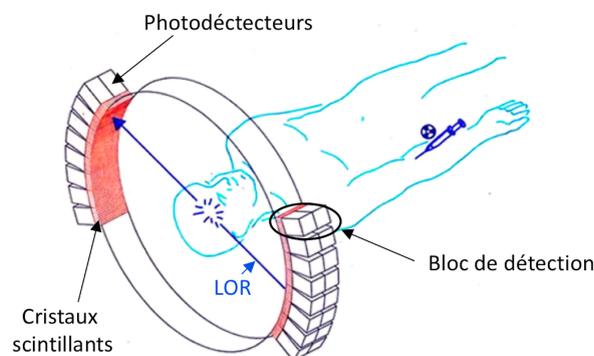


FIGURE 1.5 – Illustration schématique du système de détection en TEP [148]

L'imageur GEMINI TF : le GEMINI TrueFlight (Gemini TF) [149] utilisé pour tous les travaux qui seront décrits dans ce manuscrit ne possède toutefois pas une géométrie en bloc détecteurs. Ce tomographe dispose de modules dont chacun est composé d'un cristal de LYSO continu pixelisé, avec une grille de 23 x 44 pixels de 4 x 4 mm³ de section et 22 mm de profondeur

(fig. 1.6b). Chaque cristal est couplé à des photomultiplicateurs (PM) (fig. 1.6a) de 39 mm de diamètre. Au total, la lecture du signal est effectuée par un réseau de 420 PMs. Les modules sont placés de manière à ce que leur plus grande dimension soit parallèle à l'axe du tomographe. Le GEMINI TF possède une unique couronne de détection composée de 28 modules, avec 90,34 cm de distance séparant deux modules se faisant face.

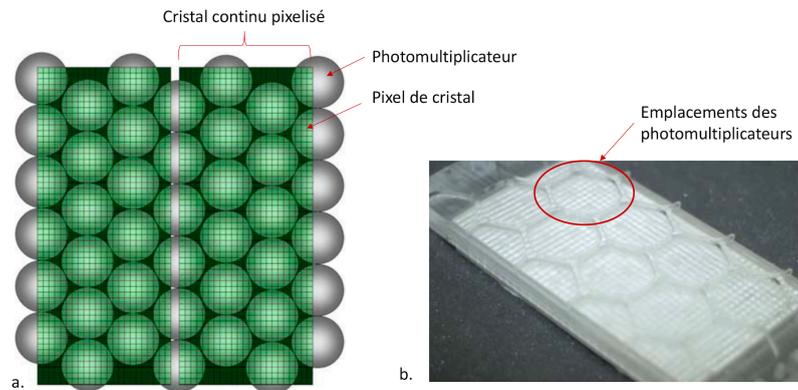


FIGURE 1.6 – (a) Schéma illustrant deux modules adjacents dont des PMs se chevauchent, (b) photo d'un cristal LYSO pixelisé d'un GEMINI TF. sources : (a) Daube-Witherspoon (2009), (b) Philips.

La détection en coïncidence est régie par deux critères principaux : la fenêtre temporelle de détection pendant laquelle deux photons sont considérés comme simultanés, généralement comprise entre 5 et 12 ns, ainsi que de la fenêtre en énergie. Si l'annihilation a eu lieu à égale distance des détecteurs, les photons seront collectés simultanément. Si l'annihilation est plus proche d'un des détecteurs, les photons seront alors détectés avec un léger décalage dans le temps. La mesure de ce décalage de temps d'arrivée sur les détecteurs, appelée également temps de vol (*time of flight* (TOF)), aide à cibler la position de l'annihilation sur la LOR (fig. 1.7). La prise en compte du temps de vol n'est pas systématique et dépend du modèle du tomographe.

L'imageur GEMINI TF : le GEMINI TF permet la prise en compte du temps de vol (dont il tire son nom). La résolution temporelle du système est de 585 ps et la fenêtre temporelle de coïncidence est fixée à 6 ns. La résolution énergétique moyenne sur tous les cristaux (11,5% de largeur à mi-hauteur, FWHM pour *full width at half maximum*, à 511 keV) permet un réglage par défaut de la fenêtre d'énergie de 440–665 keV.

Malgré l'optimisation des résolutions temporelles et en énergie des détecteurs, les coïncidences collectées ne sont pas toutes des coïncidences "vraies" (fig. 1.8a), i.e. issues de l'annihilation d'un même positon et dont les photons n'ont pas été diffusés dans le milieu. Elles peuvent également provenir de :

- **coïncidences diffusées :** elles sont observées si au moins un des deux photons issus d'une même annihilation a subi une ou plusieurs diffusions dans le patient ou dans le système de

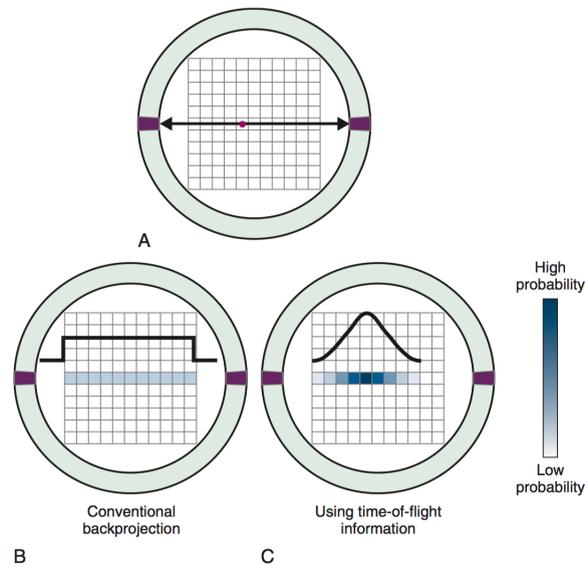


FIGURE 1.7 – (A) Détection en coïncidence de photons émis de la source (point rouge) (B) En absence de correction de temps de vol, la probabilité d’emplacement de la source est constante sur toute la LOR. (C) Avec l’information sur le temps de vol, la probabilité d’émission de la source suit une distribution gaussienne. [145]

détection avant d’être détectés (fig. 1.8b). La proportion de ces coïncidences va dépendre de la résolution en énergie du détecteur : plus la résolution est élevée, plus le nombre de coïncidences diffusées collectées sera faible.

- **coïncidences fortuites** : elles correspondent à la détection des deux photons issus de deux annihilations distinctes, ayant diffusé ou non (fig. 1.8c). Leur nombre dépend de la taille de la fenêtre temporelle.
- **coïncidences multiples** : elles correspondent à la détection de plus de deux photons dans la fenêtre temporelle de coïncidence (fig. 1.8d). Il existe plusieurs manières de traiter ces événements, dont la plus simple consiste à ne pas prendre en compte ces coïncidences détectées. À l’inverse, il peut être décidé de conserver l’intégralité des coïncidences détectées et les LOR correspondantes. Enfin, une dernière stratégie consiste à conserver uniquement la paire de la plus haute énergie.

En pratique, on définit ce que l’on appelle un champ de vue (FOV pour *Field Of View*) qui délimite l’espace “utile” dans lequel l’image va être reconstruite. Ce dernier est caractérisé par sa dimension axiale (qui va être définie par le nombre d’anneaux, le nombre de couronnes et la dimension des cristaux selon cet axe) et sa dimension transverse.

1.2.1.2.3/ MODES D’ACQUISITIONS 2D/3D

Les acquisitions en TEP peuvent se faire en mode bidimensionnel (2D) ou en mode tridimensionnel (3D). Dans le cas bidimensionnel, seules les coïncidences issues de cristaux provenant

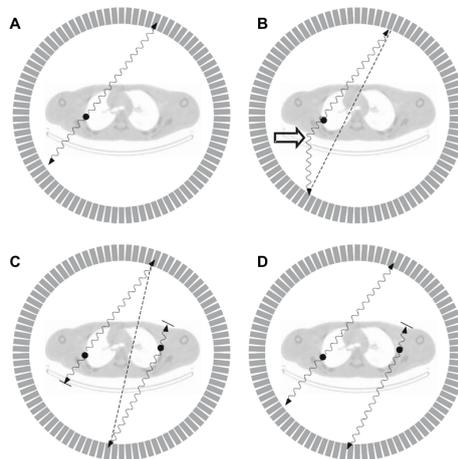


FIGURE 1.8 – Illustration des coïncidences détectées en TEP : les coïncidences vraies (A), diffusées (B), fortuites (C) et multiples (D). [143]

d'un même anneau (plan direct) ou d'un anneau adjacent (plan croisé) sont prises en compte. Des septas sont placés entre les anneaux et permettent ainsi une collimation axiale en stoppant les photons des LOR obliques (fig. 1.9). La fraction des coïncidences diffusées parmi les coïncidences recueillies en est réduite. En mode tridimensionnel, les coïncidences issues de la totalité des anneaux sont prises en compte. Le nombre de coïncidences détectées, augmenté d'un facteur allant de 5 à 10 [150, 151], permet d'améliorer la sensibilité de la détection et par conséquent les performances en signal sur bruit (SNR pour *Signal to Noise Ratio*) dans les images reconstruites. Néanmoins, cela introduit également un biais de sensibilité entre le centre et les bords du FOV axial, avec un nombre plus élevé de coïncidences détectées au centre. Ce biais peut être compensé lors de l'acquisition par un chevauchement du FOV. On observe également une proportion plus importante de coïncidences diffusées en mode 3D qu'en mode 2D.

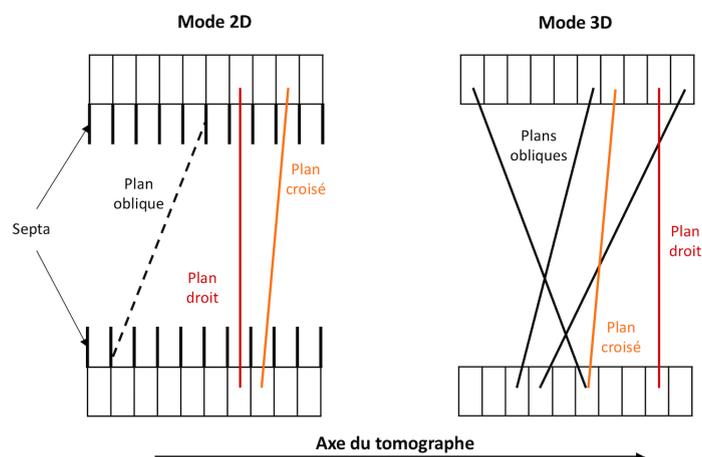


FIGURE 1.9 – Illustration des modes de détections 2D et 3D

L'imageur GEMINI TF : comme la plupart des appareils de TEP actuels, le GEMINI TF

est dépourvu de septa, l'acquisition se fait en 3D.

1.2.1.2.4/ CORRECTIONS DU SIGNAL

La quantification des images TEP est possible uniquement si des étapes de corrections du signal sont effectuées en amont, soit sur les données brutes mesurées, soit pendant le processus de reconstruction des images. Les corrections appliquées au signal permettent à la fois de “faire le tri” parmi toutes les coïncidences détectées (part. 1.2.1.2.2), de prendre en compte la dégradation du signal issue des interactions des photons et leur atténuation dans la matière, ou encore d'évaluer l'impact de la géométrie propre au tomographe et des caractéristiques et limites du système de détection sur le signal recueilli.

- La normalisation

Dans un système TEP, l'efficacité de détection de coïncidence peut varier entre différentes paires de photodétecteurs en raison de l'efficacité de la chaîne de détection, qui va dépendre de la manufacture de chaque cristal et de chaque photodétecteur, du système électronique et de la géométrie. La normalisation consiste à corriger le signal mesuré de ces inhomogénéités inhérentes au système de détection en appliquant des facteurs correctifs multiplicatifs. Ces facteurs peuvent être estimés via l'acquisition d'une source planaire (parallèle à l'axe du détecteur) décrivant un mouvement circulaire afin d'irradier de manière homogène tous les détecteurs [152]. La qualité de la correction va dépendre de la statistique de la mesure de référence. Les facteurs de correction mesurés englobant toutes les sources de perturbation du signal, certains travaux ont entrepris de les découpler de manière à évaluer chaque facteur séparément [152–154] puis de façon conjointe [155].

L'imageur GEMINI TF : le modèle de normalisation intégré dans le GEMINI TF est le modèle de Hoffman *et al.* [152] qui inclut la dépendance de la sensibilité en fonction de l'angle d'incidence des photons sur le cristal. Les facteurs de correction sont alors découplés en une composante géométrique et une composante efficacité.

- La correction d'atténuation

La possible interaction des photons d'annihilation dans le milieu réduit le nombre de coïncidences vraies détectées. Pour remédier à la sous-estimation de l'activité qui en résulte [156], une correction d'atténuation est effectuée afin d'obtenir des images qualitativement et quantitativement correctes. Cette correction est généralement faite à l'aide d'une acquisition tomomodensitométrique (TDM) qui permet l'évaluation des coefficients d'atténuation linéiques des photons dans les tissus tout en permettant également de faciliter l'interprétation des images TEP par l'apport d'une information anatomique. Ce sont les raisons pour lesquelles la plupart des imageurs TEP actuels sont couplés avec une machine à TDM.

L'imageur GEMINI TF : la correction d'atténuation est effectuée via une acquisition TDM.

- La correction des coïncidences fortuites et diffusées

Le signal recueilli en TEP est la somme des coïncidences vraies, des coïncidences diffusées et des coïncidences fortuites (part. 1.2.1.2.2). Afin de récupérer uniquement le signal provenant des coïncidences vraies, il est nécessaire de soustraire au signal recueilli les composantes du signal correspondant aux deux autres sources de coïncidences.

Il existe deux principales méthodes qui permettent d'estimer les coïncidences fortuites. La première consiste à évaluer pour chaque cristal le nombre d'événements simples. Ces derniers représentent tous les événements détectés à l'intérieur de la fenêtre en énergie, quelque soit leur instant d'arrivée par rapport à la fenêtre temporelle de coïncidence. La deuxième méthode consiste à utiliser une fenêtre temporelle retardée, qui contiendra des événements purement aléatoires et qui permet ainsi d'estimer la proportion d'événements aléatoires dans la fenêtre de coïncidences [157, 158].

En ce qui concerne les coïncidences diffusées, plusieurs méthodes ont été mises en place pour leur correction. De la plus intuitive à la plus complexe, on peut citer le fenêtrage en énergie qui va permettre de ne sélectionner que les photons dont l'énergie est proche de 511 keV (les photons possédant une énergie inférieure étant supposés avoir subi une diffusion), la méthode de "convolution-soustraction" [159], qui considère que la distribution des coïncidences diffusées est une convolution de la distribution des coïncidences vraies et peut donc être corrigée une fois le noyau de convolution identifié, l'estimation des coïncidences diffusées par la simulation de la diffusion simple [160–165], souvent nommée SSS pour *Single Scatter Simulation* et qui s'appuie notamment sur les données TDM, ou encore l'estimation des coïncidences diffusées par simulation Monte Carlo [166–170]. Cette dernière est la méthode la plus précise et la plus complète pour la correction des coïncidences diffusées mais nécessite des temps de calcul longs non compatibles avec la routine clinique.

L'imageur GEMINI TF : la correction des coïncidences fortuites est effectuée via l'application d'une fenêtre temporelle décalée. Pour ce qui est des coïncidences diffusées, elles sont corrigées via une méthode de simulation de la diffusion simple optimisée pour le TOF (TOF-SSS) [171].

- La correction du temps mort

Après avoir reçu un photon, le détecteur a une période, dite de "temps mort", pendant laquelle il ne peut plus comptabiliser de nouveaux événements. À un taux de comptage élevé, i.e. à forte activité, le temps mort de l'intégralité du système peut limiter l'efficacité de détection. Le taux de comptage réel du système n'augmente alors pas linéairement avec l'activité, biaisant ainsi la quantification. La méthode classique permettant de corriger l'effet de ces temps morts consiste à appliquer un facteur global multiplicatif sur les données de l'acquisition [172], ce facteur étant associé à différentes quantités d'activité.

L'imageur GEMINI TF : le facteur permettant de corriger le temps mort du détecteur est déterminé à partir du taux de comptage d'événements simples.

1.2.1.3/ LA RECONSTRUCTION DE L'IMAGE EN TEP

Une fois détectées, ces coïncidences sont ensuite enregistrées afin d'être utilisées pour la reconstruction de l'image TEP. Le stockage de ces données peut se faire de différentes manières, dont les deux principales sont :

- **le mode liste** : le stockage des coïncidences se fait de manière chronologique. L'énergie et l'index des cristaux sont conservés pour chaque coïncidence ainsi que l'information sur le TOF s'il est disponible.
- **le sinogramme** : l'enregistrement des données se fait via un échantillonnage de l'espace qui va permettre de représenter les LOR dans des plans perpendiculaires à l'axe du tomographe. Le nom «sinogramme» vient du fait qu'un seul point de l'image $f(x,y)$ trace une sinusoïde dans l'espace de projection. Le sinogramme d'un objet correspond à la superposition de toutes les sinusoïdes pondérées par la quantité d'activité à chaque point de l'objet imagé. Si l'on prend l'exemple d'une représentation en deux dimensions d'un plan transverse quelconque perpendiculaire à l'axe de l'imageur, chaque projection va être caractérisée par son angle d'inclinaison par rapport à l'axe du tomographe (ϕ) et sa distance x_r par rapport au centre du FOV. À chaque angle va correspondre une ligne du sinogramme sur laquelle sont comptabilisés tous les événements détectés sur l'ensembles des LOR concernées. L'ensemble de toutes les projections sur tous les angles ϕ et toutes les distances r forment le sinogramme (fig. 1.10). Pour un angle donné, l'application d'un échantillonnage régulier des distances est la méthode la plus communément utilisée. Toutefois, si l'on considère les LORs réelles (i.e. définie par deux cristaux), plus l'on se rapproche du bord du FOV transaxial, plus la distance entre les LORs est étroite. C'est ce que l'on nomme "l'effet d'arc" [173]. Un élément du sinogramme proche du bord du FOV contiendra donc plus de LORs réelles qu'au centre. La sensibilité des LORs se trouvant au bord du FOV est donc augmentée. Dans ce cas là, la correction de cet effet d'arc doit donc être prise en compte au moment de la reconstruction.

À partir de ces événements stockés, l'objectif de la reconstruction de l'image va être d'obtenir une estimation en volume de la distribution du radiotraceur dans l'organisme du patient.

L'imageur GEMINI TF : la géométrie du tomographe permet un FOV axial de 18 cm et transverse de 57,6 cm. Les données sont acquises en mode liste.

1.2.1.3.1/ RECONSTRUCTIONS 2D/3D

Dans le cas des acquisitions 2D, les coïncidences sont formées à partir de LOR dont les cristaux appartiennent à un même anneau (plan direct) ou à deux anneaux adjacents (plan croisé) (fig. 1.9). La reconstruction des données se fait alors en 2D pour chacun des plans (droits et croisés).

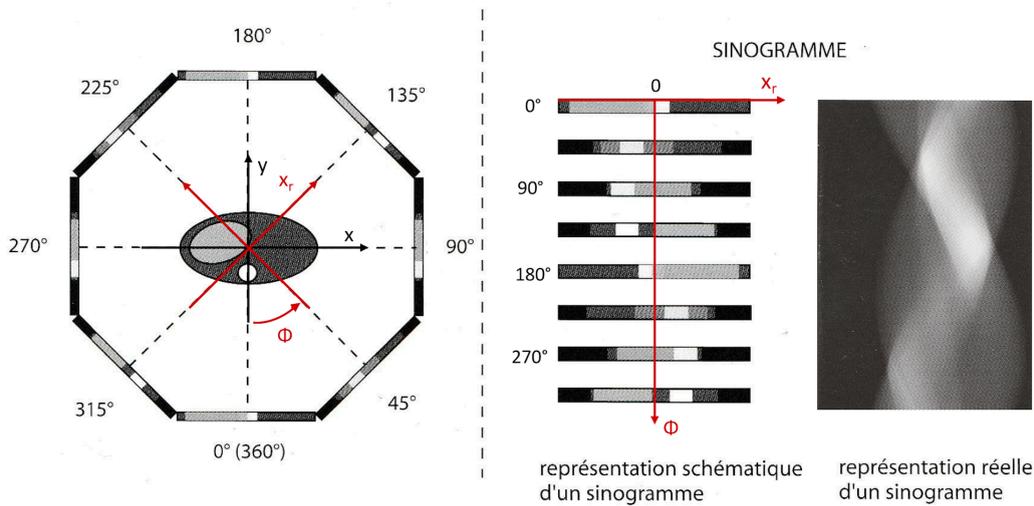


FIGURE 1.10 – Exemple schématique de la construction d'un sinogramme, adapté de [174]

Les images 2D sont ensuite concaténées pour former un volume 3D.

Dans le cas des acquisitions 3D, les images peuvent soit être reconstruites en 2D soit reconstruites en 3D. Lorsque les données sont reconstruites en 2D, ces dernières sont rééchantillonnées de manière à “recréer” des LORs uniquement dans les plans droits (ou croisés). La reconstruction se fait alors dans les mêmes conditions que précédemment. Plusieurs méthodes de rééchantillonnage existent. Parmi ces dernières, on peut citer le *Single Slice ReBinning* (SSRB) [175], le *Multi Slice ReBinning* (MSRB) [176] ou encore le *FOurier REbinning* (FORE) [177]. Dans le SSRB par exemple, les coïncidences issues de LORs obliques sont réaffectées aux plans droits ou croisés qui se trouvent à mi-chemin dans la direction axiale entre les deux cristaux impliqués dans la coïncidence. La reconstruction tridimensionnelle quant à elle exploite l'intégralité des données 3D sans rééchantillonnage. Les images 2D interprétées sont des coupes de ce volume 3D.

L'imageur GEMINI TF : La reconstruction se fait de manière tridimensionnelle.

Les algorithmes de reconstruction peuvent être classés en deux catégories : les méthodes analytiques et les méthodes itératives. Les méthodes analytiques étant moins utilisées pour les reconstructions TEP en routine clinique, elles seront décrites plus succinctement que les méthodes itératives.

1.2.1.3.2/ LES MÉTHODES ANALYTIQUES

Les approches analytiques ont été les premières méthodes utilisées pour les algorithmes de reconstruction. Elles sont basées sur l'inversion des relations mathématiques entre une fonction f décrivant un objet et ses projections linéaires p . Si l'on prend l'exemple d'un objet en deux dimensions, cela revient à retrouver à partir des projections $p(x_r, \phi)$ des LORs pour un angle donné, l'intégrale de l'objet $f(x, y)$ sur ces mêmes LORs (fig. 1.11). La transformation de $f(x, y)$

→ $p(x_r, \phi)$ est connu sous le nom de transformée de Radon. Le principe des méthodes analytiques consiste à inverser la transformée de Radon [178–180] afin de reconstruire l’objet f . Ces méthodes ont l’avantage d’être rapides d’exécution mais négligent cependant de nombreuses caractéristiques physiques propres aux données réelles, comme l’atténuation, la diffusion, la statistique de Poisson dans le comptage des photons ou encore le bruit [180]. Par ailleurs, les images reconstruites par rétroprojection simple sont difficilement directement exploitables en clinique car elles sont très artéfactées. Afin de remédier à ce problème, une étape de filtrage est nécessaire. Cette dernière se fait dans le domaine fréquentiel (domaine de Fourier) et est intégrée dans l’algorithme analytique le plus utilisé qui est la rétroprojection filtrée (*filtered back-projection* (FBP)). Le filtrage permet de limiter l’amplification du bruit dans l’image reconstruite, le compromis entre la résolution spatiale et le niveau de bruit sera déterminés par le choix du filtre et de sa fréquence de coupure.

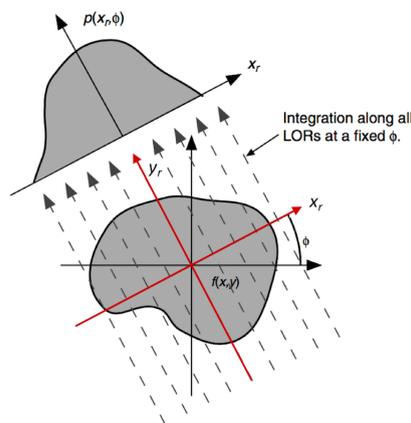


FIGURE 1.11 – Exemple de projection faite à partir de l’intégration sur toutes les LORs parallèles pour un angle ϕ donné [180]

1.2.1.3.3/ LES MÉTHODES ITÉRATIVES

Afin de remédier aux limitations des méthodes analytiques et se rapprocher au maximum de la réalité imagée, les méthodes itératives se basent sur un modèle linéaire plus général que le modèle de Radon qui va permettre l’inclusion de facteurs réalistes dans le processus de reconstruction, tels que le bruit, la diffusion ou encore l’atténuation des photons d’annihilation. Cependant, ces ajouts complexifient le modèle mathématique à résoudre, rendant le calcul de la solution plus difficile. C’est pourquoi en pratique ces algorithmes de reconstruction tentent d’estimer la distribution des traceurs de manière itérative : à partir d’une distribution arbitraire, les projections calculées sont comparées à celles mesurées. L’image estimée est affinée et corrigée progressivement jusqu’à ce qu’elle corresponde au mieux aux données mesurées. L’équation à résoudre peut s’écrire de façon simplifiée comme suit :

$$g = \mathbf{H}f \quad (1.1)$$

avec f qui représente la distribution du radiotracer, g qui correspond aux projections mesurées et H la matrice système dont chaque élément $h_{i,j}$ va contenir la probabilité que l'élément d'image f_j contribue à la mesure g_i . Cette matrice est dite "matrice système" car elle va permettre d'inclure de nombreuses caractéristiques physiques qui interviennent dans l'acquisition des données, comme la géométrie de l'imageur, les effets physiques comme l'atténuation ou le flou, ou encore la normalisation du système de détection. La modélisation du système par la matrice H peut se faire en factorisant chacune de ces contributions physiques [181] tel que :

$$H = H_{range}H_{norm}H_{flou}H_{geom}H_{attn} \quad (1.2)$$

où H_{range} va représenter le flou dû au libre parcours moyen du positon, H_{norm} va représenter les facteurs d'étalonnage et de normalisation des détecteurs, H_{flou} va correspondre au flou dans l'espace de projection dû aux contributions des coïncidences diffusées, à la non-colinéarité d'émissions des photons d'annihilation et à l'incertitude sur la profondeur d'interaction dans le cristal sur le flou dans l'espace de projection, H_{geom} va modéliser les probabilités de détection en fonction de la géométrie du système (angles solides) et H_{attn} va représenter les facteurs d'atténuations. En fonction des contributions physiques prises en compte, le modèle variera en complexité. Les mesures g étant connues et la matrice H estimée, il suffirait donc en théorie d'inverser l'équation afin de retrouver la distribution f recherchée. En pratique, la matrice H étant de très grande dimension et généralement pas carrée, son inversion est soit difficile soit impossible. Pour retrouver f on va donc essayer de minimiser de manière itérative les différences entre la distribution des projections estimées et la distribution réelle des projections enregistrées. Cette différence que l'on cherche à minimiser est définie, soit de manière fixe avec un nombre d'itérations pré-établies, soit de manière automatique en utilisant un critère d'arrêt (seuil) qui aura été prédéfini pour la fonction de coût utilisée.

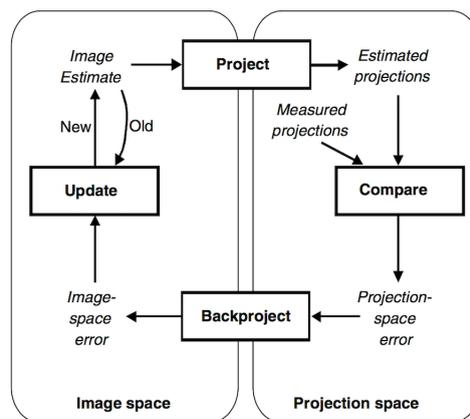


FIGURE 1.12 – Organigramme d'un algorithme générique de reconstruction itérative [180]

Les algorithmes itératifs sont basés soit sur des modèles algébriques soit sur des modèles probabilistes. Pour les modèles algébriques, l'estimation des projections est faite sans hypothèse statistique préalable, contrairement aux modèles probabilistes qui eux utilisent des hypothèses sur les lois statistiques associées aux données TEP mesurées. Ces derniers, qui sont les plus

utilisés en routine clinique, considèrent que le bruit qui affecte les données suit une distribution de Poisson [182]. Les principaux algorithmes utilisés en routine clinique sont les suivants :

- Maximum Likelihood Expectation Maximization (MLEM)

L'algorithme le plus largement utilisé est le *Maximum Likelihood Expectation Maximization* (MLEM) [183, 184] qui tend à maximiser l'espérance (*expectation maximisation* (EM)) en utilisant l'estimateur de vraisemblance (*maximum likelihood* (ML)) afin de trouver l'objet dont la probabilité d'observation est la plus élevée. Son comportement de convergence est connu et prévisible. Il a comme propriété d'avoir ses basses fréquences spatiales qui convergent en premier, suivies des fréquences plus élevées qui s'ajoutent au fur et à mesure des nouvelles itérations [185]. Plus le nombre d'itérations est élevé, plus l'algorithme se rapproche de la solution mais plus la variance, qui se manifeste comme du bruit, augmente dans l'image résultante. En pratique, afin de limiter de bruit dans l'image finale, la convergence doit être stoppée de manière prématurée [186, 187]. L'arrêt de la convergence est régi par le nombre d'itérations qui est prédéfini par l'utilisateur. La convergence de cet algorithme est lente et le temps de calcul requis à la reconstruction d'une image a longtemps été le principal frein à son utilisation en routine clinique.

- Ordered Subset Expectation Maximization (OSEM)

L'algorithme OSEM est une version accélérée de MLEM qui introduit des sous-étapes au sein d'une itération [188]. En effet, à chaque itération, au lieu d'appliquer l'algorithme MLEM directement sur l'ensemble des données mesurées, il va être appliqué sur un premier sous-ensemble de projections afin de faire une première correction de l'image estimée. Les sous-ensembles suivants vont ensuite permettre, les uns après les autres, d'affiner l'estimation de l'image jusqu'à l'itération suivante. Entre chaque sous-ensemble, les projections utilisées doivent être les plus espacées possibles afin de couvrir au maximum l'espace des projections de manière régulière/homogène par itération. Même si la convergence de cet algorithme n'a pas été démontrée (contrairement à MLEM), il a été reconnu de manière empirique que le résultat d'OSEM après n itérations est très proche de celui de MLEM après $n \times k$ itérations, k étant le nombre de sous-ensembles. Lorsqu'il est réduit à un seul sous-ensemble englobant toutes les données de projection, l'algorithme OSEM est équivalent à l'algorithme MLEM. Les propriétés d'OSEM restent très proches de celles de MLEM et le temps de traitement pour une itération d'OSEM est comparable à celui d'une itération de MLEM. La solution finale est donc atteinte beaucoup plus rapidement via OSEM. Néanmoins, le principal coût de l'utilisation de OSEM par rapport à MLEM vient de l'utilisation des sous-ensembles, qui permettent certes l'accélération de la convergence, mais augmentent également plus rapidement le bruit dans l'image [189]. La qualité des images va dépendre du choix du nombre d'itérations et du nombre de sous-ensembles qui vont déterminer l'arrêt de la convergence de l'algorithme. Une variante d'OSEM, nommée RAMLA (*Row-Action Maximum Likelihood Algorithm*) a également été proposé par Browne et De Pierro [190]. Cet algorithme, dont la convergence a été démontrée, utilise autant de sous-ensembles que de projections et applique à un terme

correctif de pondération qui permet de contrôler le taux de convergence.

La principale difficulté de ces algorithmes est de trouver un compromis entre la convergence vers la solution et la variance issue du processus itératif. C'est pourquoi pour chacun d'entre eux il est en pratique nécessaire de stopper de manière prématurée la convergence de l'algorithme. En parallèle, la gestion de l'amplification du bruit peut également être faite par un lissage des images durant le processus itératif, soit par l'application de filtres (de type gaussien généralement), soit en pondérant l'impact de chaque étape itérative via un paramètre de relaxation. La supériorité des méthodes itératives statistiques vis à vis des autres types de méthodes ayant été démontrée [191, 192], ces dernières sont utilisées dans la quasi-totalité des cas en clinique.

Cette présentation des algorithmes itératifs de reconstruction n'est pas exhaustive et ne concerne que les méthodes statistiques. Il existe de nombreux autres algorithmes [180, 193] qui n'ont pas été décrit dans ce manuscrit car ils sont moins fréquemment utilisés en routine clinique. L'algorithme et les paramètres de reconstruction qui ont été utilisés dans le cadre des travaux de cette thèse seront développés plus en détail dans le chapitre 2.

1.2.2/ LIMITATIONS ET PROPRIÉTÉS DE L'IMAGERIE TEP

Les données TEP sont affectées par de nombreux biais dont certaines corrections peuvent être appliquées afin d'optimiser la quantification en TEP (part. 1.2.1.2.4). Néanmoins, certaines caractéristiques intrinsèques au système ou à la reconstruction perdurent et font partie des caractéristiques et des limitations de l'imagerie TEP.

1.2.2.1/ LE BRUIT

Les données en TEP sont naturellement bruitées, de part la nature aléatoire de l'émission des positons et du compromis nécessaire entre le temps d'acquisition, l'activité injectée et la qualité d'image. Ce bruit est d'autant plus prononcé que la statistique des images reconstruites est faible. Afin d'améliorer la statistique de l'image, un compromis doit être fait entre les temps d'acquisitions et la dose injectée. La nature probabiliste des désintégrations radioactives des radiotraceurs rend Poissonienne la principale composante du bruit visible sur les sinogrammes. Cependant, la nature du bruit se complexifie [194] après la reconstruction. Le bruit dans les images TEP reconstruites est généralement caractérisé par la variance (ou l'écart type) des valeurs des voxels dans une zone de l'image identifiée ou connue comme homogène [195]. Ce bruit est souvent assimilé à une distribution Gaussienne et va dépendre des algorithmes de reconstruction ainsi que des filtres utilisés au cours des étapes itératives (le filtre le plus utilisé étant le filtre gaussien caractérisé par sa FWHM).

L'imageur GEMINI TF : dans le cas du GEMINI TF, le lissage de l'image est effectué durant

le processus itératif. Ce dernier ne se fait pas par un filtrage gaussien mais par l'application d'un paramètre de relaxation λ qui va permettre de pondérer la fonction erreur, c'est à dire le rapport des projections mesurées/estimées, à chaque mise à jour du processus itératif. Plus λ est faible, plus le poids de chaque itération est faible, ce qui a comme conséquence de limiter l'ajout de bruit dans l'image. Ce paramètre est compris entre 0.01 et 1. Plus ce dernier est proche de 0.01, plus l'image apparaît lissée. En routine clinique, ce paramètre est laissé par défaut égal à 1.

1.2.2.2/ LA RÉOLUTION SPATIALE

La résolution spatiale fait référence à la dimension minimale du plus petit détail observable. Elle se définit comme la plus petite distance nécessaire pour distinguer deux sources ponctuelles juxtaposées dans une image. Une des mesures simplifiées de la résolution spatiale du système consiste à imager une source ponctuelle ou linéaire dont on relève les profils d'intensité selon une direction donnée (radiale, tangentielle et axiale). Les valeurs de FWHM sont mesurées après interpolation linéaire des profils. La résolution spatiale du système selon la direction étudiée correspond alors à la valeur des FWHM mesurées.

Des limites intrinsèques à la physique de l'imagerie TEP vont limiter la résolution spatiale [196] : le libre parcours moyen du positon et la non-colinéarité des photons d'annihilation. En effet, comme vu précédemment (part. 1.2.1.2), le lieu d'annihilation ne correspondant pas au lieu d'émission du positon, le libre parcours moyen de ce dernier a un impact direct sur la résolution spatiale. Plus l'énergie d'émission est grande, plus le positon va parcourir de distance avant de s'annihiler. Pour le cas du fluor 18, son libre parcours moyen est de l'ordre de 0,5 mm. Par ailleurs, l'incertitude angulaire d'émission des photons qui sont émis à $180^\circ \pm 0,2^\circ$ ajoute également une part d'incertitude sur leur lieu d'émission et par conséquent sur le lieu de fixation du radiotraceur.

D'autres effets dû à la géométrie des tomographes ainsi qu'à l'électronique de détection vont également dégrader la résolution spatiale. Ces effets, qui vont dépendre des caractéristiques des imageurs, sont : la section de détection du cristal (ou pixel de cristal), la mesure de la profondeur d'interaction des photons dans les cristaux (DOI pour Depth Of Interaction), l'échantillonnage et les erreurs dues au multiplexage du système de détection. En effet, plus la section de détection du cristal est grande, moins l'échantillonnage spatial des projections acquises est fin. Par ailleurs, les photons d'annihilation pénètrent généralement une certaine distance dans l'anneau de détection avant d'interagir. Or, dans les systèmes où la DOI n'est pas mesurée, l'interaction est considérée comme ayant eu lieu soit à la surface du cristal soit à une profondeur fixe. De plus, il arrive également que le photon ait traversé plusieurs cristaux avant d'être détecté. L'absence d'information sur la DOI peut entraîner un décalage entre la LOR réelle et la LOR mesurée (dès lors que la LOR réelle n'est pas issues de cristaux "*back to back*"), dégradant ainsi la résolution spatiale. Cet effet est d'autant plus marqué que la LOR se situe loin du centre du FOV, accentuant ainsi les effets de parallaxe [197]. Plusieurs méthodes instrumentales permettent de calculer la DOI [198,199] limitant ainsi cette source d'erreur. La résolution spatiale va également dépendre de l'échantillonnage intrinsèque de la géométrie du système de détection qui s'avère être non uniforme : en fonction de l'endroit où l'on se place

dans le FOV, les voxels de l'image seront traversés par plus ou moins de LORs [200]. Enfin, pour de nombreux systèmes de détection, le nombre de cristaux est plus élevé que le nombre de photodétecteurs. Le traitement du signal par ces derniers peut alors se révéler imparfait, ce qui a pour conséquence de dégrader la résolution spatiale [196].

Pour finir, la résolution spatiale va également dépendre des conditions d'acquisition et des paramètres de reconstruction des images TEP. Plusieurs aspects de la reconstruction vont affecter la résolution spatiale, comme les étapes de filtrages, le type de fonction de base utilisée pour la représentation de l'image ou encore l'échantillonnage de l'image reconstruite [201]. De même, le bruit naturellement présent dans les images TEP va dégrader la résolution spatiale. L'estimation de son impact sur cette dernière est complexe, car il dépend de nombreux facteurs (nombre d'événements détectés, type d'algorithme de reconstruction utilisé, etc). Enfin, de par la durée des acquisitions TEP, certains mouvements de patients sont inévitables, comme les mouvements respiratoires ou cardiaques, et impactent en conséquence la résolution spatiale des images. Afin de limiter cet effet, plusieurs études se sont concentrées sur la corrections des mouvements via notamment des techniques de synchronisation (*gating*) [202].

L'imageur GEMINI TF : les cristaux de LYSO continus sont pixelisés avec une grille de 23 x 44 pixels de 4 x 4 mm³ de section et 22 mm de profondeur (fig. 1.6b). L'information de la DOI n'est pas disponible sur le GEMINI TF et aucune compensation de la dégradation de la résolution spatiale n'est prise en compte au niveau de la reconstruction. En conditions NEMA (*National Electrical Manufacturers Association*) [203], la résolution spatiale (FWHM) du système à 1 cm est de 4,8 mm en transverse et 4,6 mm en axiale. En condition d'acquisition et de reconstruction standards cliniques, la résolution spatiale du système a été évaluée dans le travail présenté en annexe C de ce manuscrit.

1.2.2.3/ L'EFFET DE VOLUME PARTIEL

L'effet de volume partiel (EVP) fait référence à deux phénomènes distincts qui vont impacter à la fois de manière quantitative et qualitative les images TEP [204]. Le premier phénomène concerne le flou présent dans l'image dû à la résolution spatiale limitée du système [Moses,1993]. Ce flou va provoquer un étalement du signal qui va provoquer une contamination mutuelle entre les régions de l'image présentant des activités différentes. L'activité d'une zone fixante va être susceptible d'être détectée à l'extérieur de la structure (*spill out*), mais l'activité environnante (autre structure fixante, fond, etc.) va également être susceptible de contaminer cette zone d'activité (*spill in*). Cet effet d'étalement va avoir pour conséquence de modifier la taille apparente des structures ainsi que leur activité, comme illustré sur la figure 1.13. La distinction entre deux structures proches est d'autant plus difficile que ce flou est important.

Le second phénomène est lié à l'échantillonnage relativement grossier des images TEP reconstruites (des voxels isotropes de 2 à 5 mm étant généralement utilisés en routine clinique), auquel est associé le problème de l'effet de fraction tissulaire (*tissue-fraction effect*). Ce dernier fait référence au fait qu'un même voxel puisse contenir plusieurs structures ne possédant pas la même concentration en activité. On aura donc par voxel une moyenne des activités apparentes pon-

dérées par la proportion des structures présentes dans le voxel. L'échantillonnage peut donc induire une sous ou surestimation d'activité en fonction des cas.

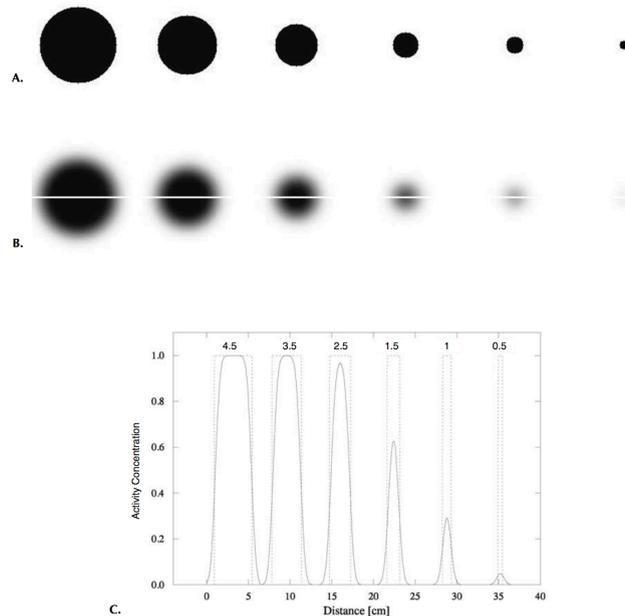


FIGURE 1.13 – Données simulées montrant l'effet de volume partiel sur un ensemble de sources sphériques (diamètre allant de 0,5 cm à 4,5 cm) de concentration d'activité égale (A). Ces données ont ensuite été filtrées (B) dans les trois dimensions avec un filtre gaussien (FWHM = 10 mm) afin de simuler un système de détection TEP. Comme le montrent les profils correspondants (C) plus la structure est petite, plus on observe une diminution de l'activité maximale ainsi qu'une augmentation de "l'étalement" spatial de l'activité. [205]

De manière générale, les petites structures sont les plus soumises à l'EVP, particulièrement celles dont la taille est inférieure à trois fois la résolution spatiale [206]. L'effet de volume partiel dépend donc à la fois de la résolution, de l'échantillonnage, du contraste et de la taille des structures. Plus la résolution du système est élevée, plus l'échantillonnage est fin, plus la structure est grande et moins le contraste est présent entre la structure et le fond, plus l'EVP sera faible.

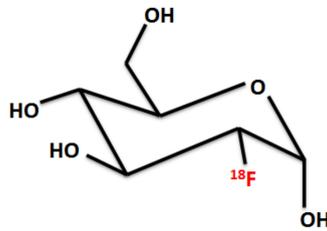
Plusieurs approches ont été étudiées afin de prendre en compte et corriger l'EVP dans les images TEP pendant ou après la reconstruction des images. Les méthodes qui peuvent être appliquées sont résumées dans les revues et méta-analyses sur le sujet [207, 208].

1.2.3/ LA TEP AU ^{18}F -FDG EN ONCOLOGIE

La TEP est donc une imagerie fonctionnelle et moléculaire dont l'objectif est de visualiser et quantifier *in vivo* un procédé biologique donné. En oncologie, la principale fonction étudiée est le métabolisme glucidique. En effet, les cellules tumorales présentent une consommation accrue de glucose par rapport aux cellules dites normales [209, 210] ce qui permet en partie de les identifier.

1.2.3.1/ LE 2-DESOXY-2-18F-FLUORO-D-GLUCOSE OU ^{18}F -FDG

L'étude du métabolisme glucidique en TEP est possible via l'utilisation d'un radiotracer, le 2-desoxy-2-18F-fluoro-D-glucose, noté ^{18}F -FDG, dont la molécule vectrice est un analogue du glucose avec, à la place d'un des groupements hydroxyyles, un isotope radioactif : le fluor 18 (^{18}F) (fig. 1.14). La molécule de ^{18}F -FDG va pénétrer dans les cellules de la même manière que

FIGURE 1.14 – 2-desoxy-2-18F-fluoro-D-glucose ou ^{18}F -FDG

le glucose, majoritairement par le biais des transporteurs GLUT-1. Le ^{18}F -FDG va ensuite être phosphorylé (^{18}F -FDG-6P) par l'hexokinase. Contrairement au glucose, le ^{18}F -FDG phosphorylé ne va pouvoir ni être métabolisé, ni déphosphorylé et se retrouve donc piégé dans la cellule et s'y accumule. Cette accumulation se fait de manière proportionnelle à la consommation de glucose par la cellule. Elle a lieu au sein des cellules normales mais est nettement plus marquée dans la plupart des cellules tumorales, les cellules médiatrices de l'inflammation activées, les cellules cérébrales ou musculaires sollicitées (ex : cardiomyocytes) ou de manière plus modérée dans le foie et l'intestin. On trouve également une concentration élevée de ^{18}F -FDG dans les reins et la vessie compte tenu de l'élimination urinaire du traceur, contrairement au glucose qui n'est pas présent dans les urines dans des conditions physiologiques. Le ^{18}F -FDG va rester majoritairement piégé dans la cellule jusqu'à la désintégration du ^{18}F en ^{18}O . L'atome d'oxygène pourra alors se lier à un ion H^+ et la molécule pourra être métabolisée par la cellule. Le ^{18}F -FDG est le traceur le plus communément utilisé en TEP.

En pratique, une imagerie TEP au ^{18}F -FDG nécessite que le patient soit à jeun au moins 6h avant l'examen afin principalement d'éviter l'hyperinsulinisme post-prandial, qui stimule l'entrée du glucose (et du ^{18}F -FDG) dans les cellules adipeuses et musculaires, aux dépens des autres cellules. Le radiotracer est injecté au patient par voie intraveineuse, avec une activité allant de 2 à 5 MBq par kg. S'ensuit une période de repos dont le délai dépend d'un compromis entre la période radioactive du fluor 18, le temps nécessaire à ce que le traceur se répartisse dans l'organisme et le temps qu'un bon rapport entre la fixation tumorale et la fixation vasculaire non spécifique du radiotracer, qui constitue du bruit de fond, soit atteint. Durant ce délai, qui dure généralement entre 60 et 90 min, le patient a comme consigne de se mouvoir le moins possible afin de limiter les fixations musculaires physiologiques du ^{18}F -FDG. Après élimination de l'accumulation du radiotracer dans la vessie, le patient est ensuite placé sur la table de l'imageur et doit rester immobile le temps de l'acquisition.

Les acquisitions TEP se font par plusieurs positions du détecteur que l'on nomme "pas de lit" et dont la grandeur varie en fonction du FOV axial propre à chaque imageur. Pour chaque pas

le signal est recueilli en moyenne pendant 1,5 à 4 minutes. En fonction du volume à étudier (corps entier ou zone anatomique limitée) et du FOV de l'imager, l'examen de routine clinique dure généralement entre 15 et 30 minutes. Une fois l'acquisition terminée, une première analyse visuelle des images est faite par le praticien afin de repérer la localisation des foyers tumoraux. Une analyse quantitative du métabolisme est ensuite réalisée afin de pouvoir caractériser plus en détail l'activité tumorale.

1.2.3.2/ MÉTABOLISME GLUCIDIQUE ET STANDARDIZED UPTAKE VALUE

L'évolution de l'accumulation du glucose et donc du ^{18}F -FDG est représentée à partir d'un modèle cinétique à deux compartiments [211] (fig. 1.15). La résolution complète de ce modèle cinétique peut être effectuée de manière plus ou moins simplifiée (l'analyse cinétique complète [212], la méthode de Patlak [213] ou encore l'analyse cinétique simplifiée [214]) et nécessite généralement des prélèvements sanguins couplés à une acquisition TEP pour le calcul de ce que l'on nomme la "fonction d'entrée" du modèle.

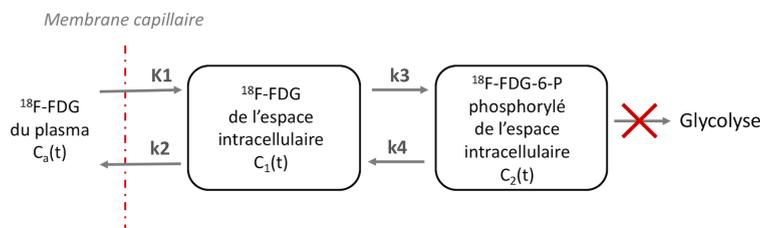


FIGURE 1.15 – Modèle cinétique permettant d'illustrer l'évolution de la fixation du ^{18}F -FDG

Dans ces cas précis, ces pratiques restent donc difficiles à appliquer en routine clinique malgré les diverses simplifications de calcul appliquées. Actuellement, la mesure la plus communément utilisée pour caractériser la fixation du ^{18}F -FDG en clinique [215] est une mesure semi-quantitative : la valeur de l'accumulation normalisée ou SUV (*Standardized Uptake Value*). Le SUV correspond à la fixation du radiotracer dans le tissu d'intérêt (kBq/mL) normalisée par l'activité injectée au patient et corrigée de la décroissance radioactive au temps de l'acquisition (kBq), divisé par le volume du patient (mL). En pratique, la masse volumique des patients étant considérée comme égale à celle de l'eau (1g/mL), le volume est remplacé par la masse du patient (g). Le SUV se définit donc selon l'équation suivante :

$$SUV = \frac{\text{Activité dans le tissu d'intérêt (kBq/mL)}}{(\text{l'activité injectée (kBq)} / \text{masse du patient (g)})} \quad (1.3)$$

Le SUV est une quantité sans dimension. Si la répartition du radiotracer se faisait de manière homogène dans l'organisme, le SUV serait en tout point égal à 1. Il existe différentes mesures de SUV dont les principales sont :

- le SUV_{\max} qui correspond à la valeur du voxel d'intensité maximale au sein du volume d'intérêt (volume of interest (VOI)). Le SUV_{\max} est la métrique la plus utilisée en routine clinique. Il a pour avantage de ne pas être dépendant directement de l'utilisateur. Il reste

néanmoins très dépendant du protocole d'acquisition (notamment le filtrage) et ne décrit le métabolisme des tumeur qu'à partir d'un unique voxel.

- le SUV_{mean} qui correspond à la moyenne des intensités des voxels compris dans le VOI. Il a pour avantage de décrire le métabolisme moyen du VOI dans sa globalité. Cependant, il va être beaucoup plus dépendant du volume de la région d'intérêt (et donc de l'utilisateur si la délimitation est manuelle) que le SUV_{max} . Il est également sensible, en particulier pour les petit volumes, à l'échantillonnage et la résolution des images TEP.
- le SUV_{peak} est une solution intermédiaire entre le SUV_{mean} et le SUV_{max} . Il correspond à la moyenne de l'intensité des voxels présents dans un volume fixe de la région d'intérêt autour du voxel d'intensité maximale ou de la zone de fixation la plus élevée. Le volume est fixe quelles que soient les tailles des VOIs et est généralement de 1 mL. Le SUV_{peak} a pour avantage d'être moins sensible à l'utilisateur que le SUV_{max} et moins sensible à la résolution des images que le SUV_{mean} . Néanmoins, le SUV_{peak} présente un biais qui va varier en fonction de la taille des VOIs et de la région considérée (régions 2D ou 3D) [216].

Il existe également des possibilités de normalisation du SUV en prenant en compte la glycémie qui influe sur la captation intracellulaire du ^{18}F -FDG. De même, les tissus adipeux ayant, dans des conditions de jeûne une captation du ^{18}F -FDG moindre que les autres tissus [217], il est possible de normaliser le SUV en prenant en considération la surface corporelle totale [218] ou encore la masse maigre (on parle alors de SUL (*Standard Uptake Lean*)) [219], afin de limiter voire corriger l'influence de la masse grasseuse des patients sur les valeurs de SUV tumoraux [75, 219, 220].

Toutes les analyses du métabolisme glucidique des tumeurs qui sont présentées dans ce manuscrit ont été réalisées en utilisant la définition standard du SUV (equ. 1.3).

1.2.3.3/ L'ÉTUDE DE LA PERFUSION TUMORALE VIA LA TEP AU ^{18}F -FDG

La perfusion tumorale fait également partie des paramètres biologiques qui peuvent être extraits des images TEP. En effet, comme présenté dans la première partie de ce chapitre (part. 1.1.4.1.3), en imagerie TEP, le marqueur de référence pour l'évaluation de la perfusion tumorale est l'eau marquée à l'oxygène 15 (^{15}O -H₂O) [221, 222]. Cependant, la période radioactive de l'oxygène 15 est de 2 minutes, ce qui nécessite la présence d'un cyclotron à proximité immédiate du centre hospitalier et réduit donc grandement sa disponibilité. Afin de contourner cette limitation, Mullani *et al.* ont développé et validé dans les années 80 un modèle de premier passage pour le calcul in vivo du débit sanguin pour l'évaluation de la perfusion myocardique en utilisant un autre marqueur : le Rubidium 82 [223]. Dans les années 2000, Zasadny *et al.* [93] ont de leur côté montré que la constante de transfert K1 (fig. 1.15) obtenue avec un autre radiotracteur, le ^{18}F -FDG, était très corrélée à celle obtenue avec l' ^{15}O -H₂O, suggérant que l'analyse cinétique compartimentale du ^{18}F -FDG pourrait apporter des informations sur la perfusion tumorale. Au vu de l'atout majeur que représenterait une unique injection pour l'évaluation de la perfusion et du métabolisme tumorale lors d'un examen d'imagerie fonctionnelle, Mullani *et al.* ont alors cherché à évaluer le débit sanguin tumoral par le modèle de premier passage en utilisant cette fois-ci du ^{18}F -FDG. Le modèle compartimental cinétique

du ^{18}F -FDG a alors été assimilé à un modèle mono-compartmental, considérant que sur les deux premières minutes post-injection où a lieu le premier passage du traceur au niveau de la tumeur, il est possible de s'affranchir du second compartiment correspondant au FDG phosphorylé, la glycolyse n'ayant pas encore eu lieu. En appliquant ce modèle, Mullani *et al.* ont alors montré que l'évaluation de la perfusion tumorale par la TEP au ^{18}F -FDG présentait une excellente corrélation avec le flux sanguin tumoral mesuré avec l' ^{15}O - H_2O [94] et pouvait donc être utilisée comme méthode alternative.

- Modèle mono-compartmental de Mullani *et al.* appliqué au ^{18}F -FDG

Si l'on considère un modèle mono-compartmental associé au FDG (fig. 1.16), on obtient un modèle analogue à celui proposé pour ^{15}O - H_2O :

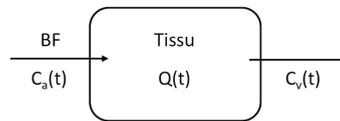


FIGURE 1.16 – Modèle monocompartmental de Kety

où le terme $Q(T)$ représente la quantité résiduelle du traceur dans le tissu au temps T , $Ca(t)$ la concentration artérielle et $Cv(t)$ la concentration veineuse du traceur au temps t et BF (*Blood Flow*) le flux sanguin du milieu vasculaire vers le tissu. La quantité de traceur dans le tissu ($Q(t)$) est égale à la quantité délivrée au tissu ($Ca(t)$) moins la quantité qui en sort ($Cv(t)$) (equ. 1.4).

$$Q(t) = \int_0^t BF.Ca(t)dt - \int_{t_d}^t BF.Cv(t)dt \quad (1.4)$$

Par ailleurs, le modèle de Mullani postule que lors du premier transit d'un bolus de ^{18}F -FDG à travers un tissu, il existe une période durant laquelle le traceur n'a pas encore quitté la région d'intérêt, de sorte que la concentration veineuse du traceur est momentanément nulle [224]. Pendant ce délai (time of delay : t_d) la plupart du traceur est retenu dans le tissu. En se basant sur ce principe et en considérant que le temps de mesure t est inférieur à t_d [223,225,226], le débit sanguin peut alors être estimé pendant cet intervalle de temps en utilisant l'équation suivante :

$$BF = \frac{Q(t)}{E. \int_0^t Ca(t)dt} \quad \text{avec} \quad Cv(t - t_d) = 0 \quad (1.5)$$

L'extraction et le piégeage métabolique d'un traceur dans les tissus variant en fonction de la nature du traceur et du tissu, un terme représentant la fraction d'extraction (E) du radiotracer doit être ajouté. Néanmoins, Mullani *et al.* ont constaté que la fraction d'extraction du ^{18}F -FDG dans le tissu tumoral était proche de celle de l' ^{15}O - H_2O , avec en moyenne seulement une baisse

de 14% et pouvait donc être considérée égale à 1 [94] ce qui simplifie l'équation comme suit :

$$BF = \frac{Q(t)}{\int_0^t Ca(t)dt} \quad (1.6)$$

L'erreur dans la mesure du débit sanguin due à la statistique de comptage peut théoriquement être minimisée lorsque le numérateur et le dénominateur de l'équation sont à leurs valeurs maximales [223]. Or, dans le cas spécifique du premier passage (avec $t < t_d$), il existe un temps t_m , appelé "temps du pic artériel" pour lequel le taux de comptage est considéré comme maximal. Ce temps correspond au moment où l'apport artériel au tissu cesse, l'entrée et le drainage veineux étant alors considérés comme nuls. À cet instant t_m , le taux de comptage est momentanément stable et l'on considère que la totalité du bolus injecté est dans le champ du détecteur. Le signal détecté atteint donc théoriquement son maximum [223]. L'équation du débit sanguin de premier passage peut donc s'écrire de la manière suivante :

$$BF = \frac{Q(t_m)}{\int_0^{t_m} Ca(t)dt} \quad (1.7)$$

- Perfusion tumorale et imagerie dynamique

Cette analyse de la perfusion tumorale est possible grâce à une imagerie TEP, dite de premier passage, réalisée directement après l'injection du radiotraceur. La perfusion tumorale n'est cependant pas directement accessible sur ces images TEP statiques. Cette dernière doit être modélisée par des modèles compartimentaux (comme le modèle de Mullani) dont les données biologiques ($Q(t)$, $Ca(t)$) sont accessibles via la reconstruction dynamique des images de premier passage. Les procédés d'acquisition et de reconstruction qui vont permettre la création de l'image paramétrique de perfusion seront décrits en détail dans le chapitre 2.

1.2.3.4/ LA PLACE DE LA TEP AU ^{18}F -FDG DANS LA PRISE EN CHARGE DU CANCER DU SEIN

De manière générale, la détection et la caractérisation des lésions mammaires primitives se fait via l'examen clinique, le bilan biologique et le bilan d'imagerie. Pour ce dernier, les techniques d'imageries conventionnelles (mammographie, échographie ou encore IRM mammaire) vont avoir une meilleure sensibilité que la TEP au ^{18}F -FDG pour la détection et la caractérisation des lésions mammaires primitives de stade précoce (< 2 cm) ou à faible captation de ^{18}F -FDG (carcinome lobulaire, phénotype luminal avec surexpression du Ki-67 et de SBR faible) [3, 227]. La TEP va en effet être utilisée en oncologie essentiellement pour les bilans d'extensions [228, 229]. Selon les recommandations du référentiel interrégional de prise en charge du cancer du sein 2019 [33], dans le cas spécifique des cancer du sein infiltrants non métastatiques, le bilan d'extension par l'imagerie est systématique :

- pour les tumeurs de plus de 5 cm ou à extension directe avec la paroi thoracique et/ou

- peau (tumeurs cT3-T4 de la classification TNM) ;
- en cas d'atteinte ganglionnaire (cN+) ;
- en cas d'envahissement ganglionnaire macro-métastatique confirmé après chirurgie.

Ce bilan d'extension de première intention peut se faire soit par un scanner thoraco-abdomino-pelvien et une scintigraphie osseuse, soit par une TEP-TDM au ^{18}F -FDG.

- Bilan loco-régional

Pour le bilan d'extension locorégional des envahissements axillaires, la TEP n'est pas indiquée pour les tumeurs de stade précoce (tumeurs < 2 cm) mais son utilité a toutefois été démontrée pour les cancers larges (> 3 cm), inflammatoires ou localement avancés [230–233], particulièrement à partir du stade IIB [234] et potentiellement IIA [231]. La TEP apparaît également intéressante pour l'évaluation de l'envahissement extra-axillaire [235–237].

- Bilan d'extension métastatique

Dans le cadre du bilan d'extension métastatique, la TEP montre dans la plupart des cas une meilleure sensibilité que les imageries conventionnelles, en particulier pour les stades localement avancés [231, 238] ou les tumeurs inflammatoires [233]. La valeur pronostique de la classification TNM par la TEP est alors supérieure à celle du bilan d'imagerie conventionnel.

- Valeur pronostique

Il a également été montré que le métabolisme glucidique évalué en TEP s'avère être corrélé à des facteurs histopathologiques pronostiques de la survie, comme les grades histopronostiques, l'index de prolifération Ki-67, la protéine p53 ou encore l'expression de récepteur hormonaux [239–242], ainsi qu'à la survie elle-même [243, 244].

- Recherche de récurrence

Une TEP au ^{18}F -FDG est indiquée en cas de forte suspicion de récurrence (symptômes cliniques, ré-ascension des biomarqueurs biologiques sanguins (CA15.3 et ACE), etc.). La TEP s'avère être un examen particulièrement performant pour l'authentification et la localisation des potentiels sites de récurrence, et supérieure aux bilans d'imagerie conventionnels [245, 246]. Pour ce qui est de la surveillance en général, la TEP ne fait pas partie des examens systématiques.

- Évaluation et prédiction de la réponse à la CNA

L'évaluation et la prédiction de la réponse à la CNA par la TEP est actuellement uniquement utilisée dans le cadre de protocoles de recherche clinique. Néanmoins, dans le cadre du cancer du sein, l'évaluation précoce de la réponse tumorale par l'étude du métabolisme glucidique a montré des résultats très encourageants (part. 1.1.4.1.2). Les données métaboliques avant et en cours de traitement néoadjuvant ont également été identifiées comme de potentiels marqueurs prédictif de la survie (part. 1.1.4.2).

1.3/ RATIONNEL DE LA THÈSE

L'un des enjeux majeurs en oncologie est donc d'arriver à proposer aux patients des thérapies personnalisées afin d'optimiser la réponse au traitement tout en minimisant les effets secondaires d'une thérapie inefficace. Dans le cas de la chimiothérapie néoadjuvante prescrite pour les cancers du sein localement avancés, la prédiction de la réponse au traitement fait partie des défis actuels. La volonté d'adapter au plus vite le traitement a ainsi vu naître la nécessité d'identifier des marqueurs (cliniques, radiologiques, biologiques, moléculaires) prédictifs précoces ou explicatifs de la qualité de la réponse histologique à la CNA.

Des marqueurs cliniques (modifications de la taille de la tumeur évaluée cliniquement) et radiologiques (critères RECIST) (part. 1.1.4.1.2 et 1.1.4.1.1) ont tout d'abord été décrits. Néanmoins, ces derniers ne permettent généralement qu'une évaluation tardive de la réponse tumorale. Des études se sont alors intéressées aux imageries fonctionnelles et moléculaires. En effet, les tumeurs malignes sont associées à une augmentation du métabolisme glucidique ainsi qu'à une augmentation du débit sanguin [247–250]. Dans la plupart des cas, ces deux phénomènes sont couplés mais il arrive dans certaines situations que ça ne soit pas le cas. Les variations de métabolisme et de la perfusion tumorale ont alors été considérées comme deux paramètres clés, ou en tout cas complémentaires, pour la caractérisation de la réponse des tumeurs au traitement [247, 251, 252]. Par ailleurs, les modifications métaboliques et perfusionnelles sont généralement observées avant les changements anatomiques, doublant ainsi l'intérêt pour les biomarqueurs issus de l'imagerie fonctionnelle et moléculaire. Les résultats des études ont alors identifié le métabolisme glucidique et la perfusion tumorale comme biomarqueurs prometteurs pour l'évaluation précoce de la réponse histologique à la CNA (part. 1.1.4.1.2 et 1.1.4.1.3).

Le métabolisme glucidique tumoral étant conventionnellement évalué par la TEP au ^{18}F -FDG en routine clinique, la perfusion quant à elle, peut être évaluée par de nombreuses techniques d'imagerie telles que la TEP à ^{15}O - H_2O , l'IRM de perfusion ou encore l'échographie de contraste (part. 1.1.4.1.3). Toutefois, l'analyse combinée du métabolisme et de la perfusion nécessiterait donc dans tous ces cas, soit l'utilisation de deux modalités d'imagerie, ce qui reste contraignant d'un point de vue organisationnel, soit uniquement d'une imagerie TEP mais avec deux radiotraceurs différents et donc deux injections au patient. Par ailleurs, en raison de la courte période radioactive de cet isotope, seule une minorité de centres peuvent exploiter la TEP à ^{15}O - H_2O . Les études de Mullani *et al.* sur l'évaluation du débit sanguin par le modèle de premier passage en utilisant cette fois-ci du ^{18}F -FDG (part. 1.2.3.3) présentent donc de multiples intérêts. En effet, l'excellente corrélation mesurée entre le flux sanguin tumoral évalué avec ^{15}O - H_2O et la perfusion tumorale évaluée par la TEP au ^{18}F -FDG [94] permettrait d'utiliser cette dernière comme outil pour la mesure de la perfusion tumorale en TEP, la rendant ainsi accessible à un plus grand nombre de centres hospitaliers. Par ailleurs, une unique administration de radiotraceur serait alors nécessaire pour l'évaluation conjointe du métabolisme et de la perfusion tumorale lors d'un même examen.

En se basant sur ce constat, Cochet *et al.* [95] ont alors montré les corrélations entre la perfusion tumorale en TEP au ^{18}F -FDG et les marqueurs immuno-histochimiques de la prolifération (Ki-67) et de la densité microvasculaire (CD34 et CD105) des tumeurs mammaires, identifiant ainsi la TEP au ^{18}F -FDG comme un outil fiable pour évaluer à la fois le métabolisme glucidique et la perfusion des tumeurs mammaires en routine clinique. Dans la continuité, deux autres études se sont intéressées aux corrélations entre la perfusion tumorale évaluée par la TEP au ^{18}F -FDG et la survie des patients chez les tumeurs Triples Négatives [96] et plus récemment vis à vis de la réponse à la CNA pour les tumeurs HER2-positives [253]. L'étude des patientes atteintes d'un carcinome mammaire Triple Négatif a montré que les tumeurs présentant une vascularisation tumorale persistante ou augmentée après le premier cycle de CNA présentaient un pronostic plus sombre, soulignant le potentiel intérêt de la perfusion dans le suivi de ces cancers. L'étude sur les tumeurs HER2-positives n'a, quant à elle, pas mis en évidence de corrélation entre les valeurs de perfusion et la réponse à la CNA. L'approche de l'évaluation métabolique combinée à l'évaluation de la perfusion tumorale pourrait notamment être particulièrement intéressante dans le contexte de la prescription personnalisée de thérapies anti-angiogéniques ciblant la perfusion tumorale [76].

Actuellement, l'évaluation de la perfusion tumorale mesurée en TEP au ^{18}F -FDG comme marqueur précoce de la réponse à la CNA n'a que très peu été étudiée. Or, la possibilité de mesurer le métabolisme et la perfusion tumorale via la TEP au ^{18}F -FDG fait de cette modalité une opportunité pour l'évaluation combinée de ces deux paramètres biologiques dans l'évaluation précoce de la réponse à la CNA.

DE L'ACQUISITION DES DONNÉES À L'ANALYSE DU MÉTABOLISME ET DE LA PERFUSION TUMORALE

Sommaire

2.1 Protocole TREN	44
2.2 Description des paramètres de reconstruction	46
2.2.1 BLOB : modèle de l'image	46
2.2.2 Ordered Subsets (OS) : modèle du système	48
2.2.3 TF : Correction du temps de vol	48
2.3 Perfusion tumorale et imagerie dynamique	49
2.3.1 Création de l'image de perfusion	49
2.3.2 Acquisition des données et reconstruction des images dynamiques . . .	50
2.3.3 Délimitation des régions d'intérêt	50
2.3.4 Identification du pic artériel et création de l'image paramétrique de perfusion	51
2.3.5 Reconstruction dynamique et statistique dans l'image	53
2.4 Métabolisme et volume tumoral	55
2.4.1 Segmentation des images en TEP	55
2.4.2 Détermination des paramètres nécessaires à la méthode de seuillage adaptatif et segmentation du volume métabolique	61
2.4.3 Fantôme NEMA-IEC	61
2.4.4 Seuillage adaptatif de Schaefer <i>et al.</i> : étalonnage sur fantôme	62
2.4.5 Seuillage adaptatif de Schaefer <i>et al.</i> : application aux images cliniques	63

Dans ce chapitre, nous allons décrire le processus de récolte des données, du recrutement des patientes à la récupération des informations de métabolisme et de perfusion tumorale qui ont permis de générer la base de données patients grâce à laquelle ont été effectuées les études décrites dans les chapitres suivants. Ainsi dans un premier temps est décrit le protocole clinique (TREN) à partir duquel ont été récoltées les données patients. Dans un second temps sont développées les caractéristiques de reconstruction des images cliniques. La création de l'image paramétrique de perfusion est ensuite expliquée pour enfin décrire dans une dernière partie la récolte de l'information du métabolisme glucidique et la segmentation du volume tumoral.

2.1/ PROTOCOLE TREN

Ce travail de thèse a été possible grâce à la mise en place il y a plusieurs années, d'un protocole de recherche clinique au sein du Centre George François Leclerc (CGFL) à Dijon. Ce protocole, nommé TREN (Tomographie par émission de positons au ^{18}F -FDG pour évaluer précocement la REponse au traitement Néoadjuvant des cancers du sein), a été initié en 2004 et est proposé aux femmes atteintes d'un cancer du sein invasif auxquelles une chimiothérapie néoadjuvante a été prescrite. Il a pour objectif principal d'évaluer conjointement les modifications métaboliques et perfusionnelles précoces observées dans la tumeur primitive (et les métastases ganglionnaires le cas échéant, en exploitant les données de la TEP au ^{18}F -FDG réalisée au diagnostic (TEP1) et après la première cure de chimiothérapie néoadjuvante (TEP2). L'objectif est de pouvoir identifier, d'une part des marqueurs métaboliques et/ou perfusionnels prédictifs précoces de la réponse histologique tumorale et d'autre part de déterminer des facteurs métaboliques et/ou perfusionnels pronostiques indiquant le risque de rechute locale ou métastatique. Une fois ces facteurs prédictifs et pronostiques identifiés, l'objectif serait de proposer des protocoles d'adaptation précoce des schémas de chimiothérapie selon l'évaluation de la réponse (ou de la non-réponse) tumorale évaluée avec ces marqueurs.

Les critères de jugement utilisés dans le protocole TREN sont le taux de réponse histologique complète (pCR) évaluée sur la pièce opératoire après traitement chirurgical selon la classification de Chevallier et la survie à 5 ans sans évolution ou récurrence de la maladie. Les critères d'inclusion concernent les femmes de 18 ans ou plus, atteintes d'un cancer du sein nouvellement diagnostiqué dont le diamètre tumoral, évalué en échographie, est supérieur à 2 cm, et ayant une indication de traitement néoadjuvant (stade II ou III de la classification UICC). Les patientes doivent également avoir été informées de l'étude, avoir signé l'accord de l'exploitation des données cliniques, biologiques et d'imagerie et être affiliées à la sécurité sociale. Les critères d'exclusion concernent les patientes atteintes de tumeurs inflammatoires ou avec extension métastatique, ayant eu des contre-indication à la CNA ou à la chirurgie, enceintes ou allaitantes, à forte glycémie (> 9 mmol/L), les patientes pour lesquelles une maladie psychiatrique compromettant la compréhension de l'information ou la réalisation de l'étude est reconnue, les patientes sous tutelle, curatelle ou sauvegarde de justice, ou encore les patientes réticentes à subir des études d'imagerie de manière répétée.

Le protocole TREN est une étude monocentrique de soins courants, prospective, observationnelle et non randomisée. En pratique, ce protocole est proposé à toute patiente traitée au CGFL et répondant aux critères d'inclusions. Les patientes sont informées oralement de la réalisation de cette dernière par les médecins investigateurs. Après un certain délai de réflexion, les patientes ne s'étant pas opposées à l'étude sont incluses dans le protocole. Leur non opposition est renseignée dans leur dossier médical par le médecin investigateur ayant recueilli le consentement.

La figure 2.1 représente le schéma simplifié du déroulement du protocole TREN (l'intégralité du protocole TREN est disponible en annexe D de ce manuscrit). Ce dernier consiste en une première acquisition TEP/TDM (TEP1) au ^{18}F -FDG effectuée avant le début du traitement dans

le cadre du bilan d'extension initial. Cet examen est effectué au moins 7 jours après la biopsie. Toutes les patientes ont eu pour consigne de jeûner pendant au moins 6 heures avant l'administration du ^{18}F -FDG. Un prélèvement intraveineux est également effectué avant l'administration du radiotracer pour le dosage de la glycémie plasmatique. La patiente est positionnée dans la caméra TEP en décubitus ventral, bras levés, sur un matelas dédié à la mammo-scintigraphie. Un bolus de 3 MBq/kg de ^{18}F -FDG est administré à 1mL/s et simultanément, une première acquisition TEP de 8 minutes, centrée sur le sein, suivie d'une acquisition tomодensitométrique de la même région pour le repérage anatomique (120 kVp, 120 mAs), sont effectuées. Cette première acquisition précoce va permettre la détermination de la perfusion tumorale. La patiente est ensuite installée en salle d'attente, en position couchée, au repos strict, pendant environ 50 minutes. Cette dernière est ensuite ré-installée en décubitus dorsal pour une acquisition TEP de la racine des cuisses jusqu'à la tête, suivie d'une acquisition TDM de la même région. Quarante-vingt-dix minutes après l'administration du radiotracer, une dernière acquisition TEP en décubitus ventral et centrée sur la région mammaire est réalisée, suivie d'une acquisition tomодensitométrique de la même région. Sur cette dernière acquisition, réalisée en deux pas de lit de 4 minutes chacun, sera évalué le métabolisme tumoral. La figure 2.2 résume le déroulement de la TEP1. De janvier 2009 à décembre 2010, les examens ont été effectués sur une TEP

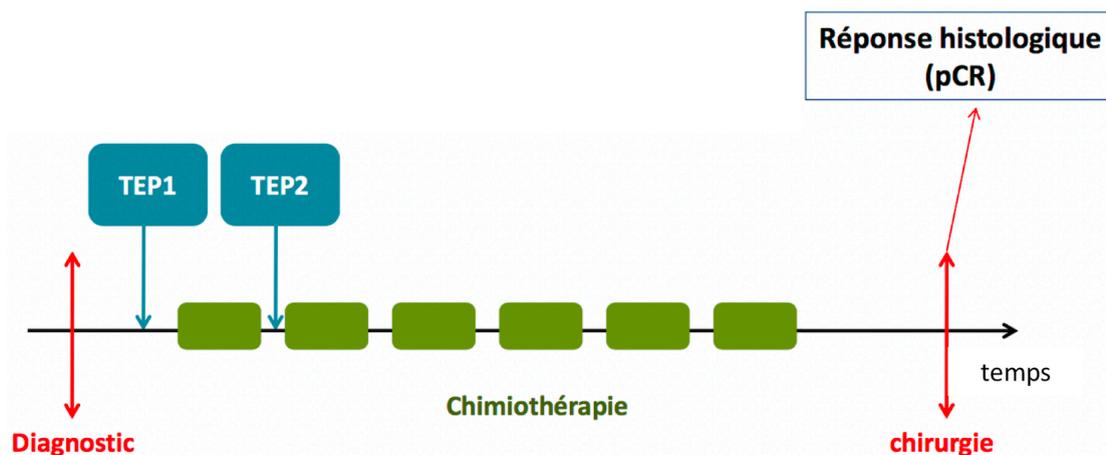


FIGURE 2.1 – Schéma simplifié du protocole TREN. Les données de la TEP au ^{18}F -FDG réalisée avant le début du traitement (TEP1) sont exploitées pour quantifier le métabolisme initial de la tumeur ainsi que la perfusion tumorale. Les données de la TEP au ^{18}F -FDG réalisée après une cure de chimiothérapie (TEP2) sont utilisées pour évaluer le métabolisme/la perfusion tumorale résiduel et la réponse métabolique/perfusionnelle en comparaison avec les résultats de la TEP1.

GEMINI GXL (Philips) équipée de cristaux de GSO (orthosilicate de gadolinium). À partir de décembre 2010, le service de médecine nucléaire du CGFL s'est doté d'une TEP GEMINI TF (Philips). Les données recueillies pour les travaux de cette thèse ont été extraits de la base TREN en prenant les patientes de janvier 2011 à mai 2017 de manière à avoir des examens TEP effectués sur un même imageur, le GEMINI TF. Les examens de 254 patientes ont été recueillis sur cette période. Après élimination des profils ayant des données manquantes ou partielles, une base de données de 246 patientes au total a pu être exploitée. Les études décrites dans les

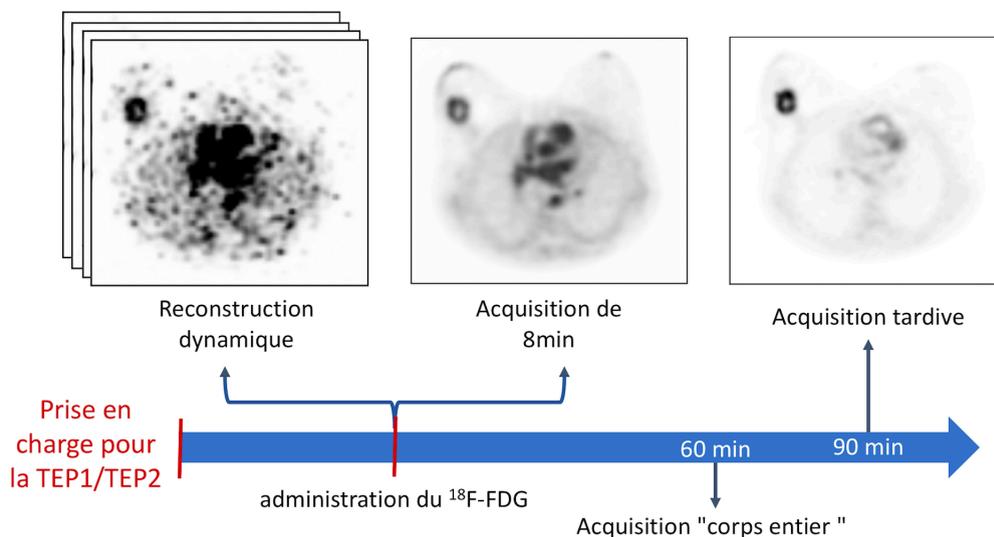


FIGURE 2.2 – Schéma simplifié du protocole de la TEP1/TEP2

prochains chapitres de ce manuscrit ont été effectuées à partir des données de la TEP 1. Les modalités d'acquisition, de reconstruction et de traitement des images seront exposées dans la suite de ce chapitre. Les modalités d'examens et de récolte des données étant les mêmes pour la TEP1 et la TEP2, le traitement des données ne sera exposé qu'une seule fois, avec des exemples issus de la TEP1.

2.2/ DESCRIPTION DES PARAMÈTRES DE RECONSTRUCTION

Les paramètres de reconstructions décrits dans cette partie s'appliquent à toutes les acquisitions TEP qui ont été effectuées dans le protocole TREN, i.e. aussi bien les images TEP précoces que tardives, de la TEP1 et de la TEP2. Pour tous ces examens, les images ont été reconstruites en utilisant l'algorithme BLOB-OS-TF en mode liste. Cette méthode de reconstruction est caractérisée par trois principaux paramètres :

2.2.1/ BLOB : MODÈLE DE L'IMAGE

Une image est à l'origine un signal réel et continu que l'on peut représenter par une fonction continue f . Afin d'être traité numériquement, cette dernière est généralement estimée par une combinaison d'un nombre fini de fonctions de base.

Soit r un point de coordonnées x , y et z le long des trois axes orthogonaux habituels et r_j , $j = (1, 2, 3, \dots, J)$ un ensemble de points disposés sur une grille régulière dans un espace de dimension 3 et dont les indices correspondent aux noeuds de cette grille (avec un pas d'échantillonnage d). La fonction continue $f(r)$ peut donc être estimée à partir d'une somme finie de fonctions de

base ϕ , comme suit :

$$f(r) \approx \sum f_j \phi_j(r - r_j) \quad (2.1)$$

avec f_j qui représente l'ensemble des coefficients de l'image. Le choix de la fonction de base ϕ va avoir un impact à la fois sur la modélisation de la fonction f mais également sur la cohérence des projections vis à vis des mesures, sur la complexité des algorithmes et sur la capacité de calcul nécessaire [254]. La fonction ϕ la plus courante en imagerie médicale est une fonction constante par morceaux, telle que :

$$\phi(r) = \begin{cases} 1 & \text{si } |r| < 0 \\ 0 & \text{sinon} \end{cases} \quad (2.2)$$

Cette fonction correspond à l'approche de type voxel (ou pixel en deux dimensions), où l'image est discrétisée en éléments identiques représentés par des parallélépipèdes droits, caractérisés par leurs trois dimensions selon les axes cartésiens x , y et z (respectivement x , y en deux dimensions). La représentation de l'image sera d'autant plus fine que le pas d'échantillonnage d de la grille est petit. La plupart des algorithmes de reconstruction en TEP utilisent cette discrétisation. Toutefois, certaines études ont souligné le fait que les voxels traditionnels sont des éléments de volume assez grossiers, que ce soit pour modéliser la distribution spatiale à reconstruire ou pour modéliser le processus de mesure. Les voxels ont une géométrie qui, de par le caractère discontinu à leurs bords, introduit inutilement des hautes fréquences dans l'image [255, 256].

Dans le cadre des reconstructions d'images tomographiques, il a été proposé des approches alternatives aux voxels [257]. L'une d'elle utilise des éléments de volume sphériques, dits "blobs" [258]. La fonction "blob" est définie par la généralisation de la fonction de Kaiser-Bessel. Cette dernière a la caractéristique d'être symétrique avec un profil radial semblable à une cloche qui décroît lentement vers zéro. La fonction ϕ associée est définie de la manière suivante :

$$\phi(r) = \begin{cases} \frac{I_m(\alpha \sqrt{1 - \frac{r^2}{a^2}})}{I_m(\alpha)} \times \left(\sqrt{1 - \frac{r^2}{a^2}}\right)^m & \text{si } 0 \leq r \leq \alpha \\ 0 & \text{sinon} \end{cases} \quad (2.3)$$

avec I_m la fonction de Kaiser-Bessel modifiée de première espèce d'ordre m , α le paramètre qui contrôle la forme de la fonction blob et a qui définit son rayon. Les limites des blobs, qui présentent une réduction progressive des valeurs près de leurs limites, ont tendance à adoucir les contours des objets présents dans l'image, rendant cette dernière plus lisse et homogène, en comparaison aux voxels. Les blobs sont placés de manière régulière dans l'espace, soit suivant une grille cartésienne, le plus commun, soit suivant une grille hexagonale. L'une des différences majeures par rapport à une discrétisation en voxel est le chevauchement entre blobs adjacents. Cette superposition permet une continuité dans les limites entre ces éléments élémentaires de l'image mais engendre une plus grande complexité dans le calcul des projections : une même LOR intersecte un nombre plus élevé de blobs que de voxels pour une résolution équivalente [259]. Néanmoins, la symétrie de rotation de ces éléments sphériques permet un calcul efficace des projections de l'image représentée [257]. Des études ont ainsi montré que l'utilisation de blobs lors des reconstructions itératives permettait une amélioration substantielle des performances

de reconstruction, qualitativement et quantitativement, par rapport à l'utilisation conventionnelle de voxels. Les images reconstruites à l'aide de blobs présentent moins de bruit pour une résolution équivalente [254].

L'algorithme avec lequel les images de la base de données ont été reconstruites a utilisé des blobs comme fonctions de base. Les paramètres m , a et α ont été fixés selon les recommandations d'utilisation de Philips à 2, 2.5 et 8.3689 respectivement. Ces critères sont ceux proposés pour les conditions d'acquisition standards corps entier et n'ont pas été spécifiquement optimisés pour une analyse radiomique et/ou dynamique.

2.2.2/ ORDERED SUBSETS (OS) : MODÈLE DU SYSTÈME

L'algorithme de reconstruction qui a été employé pour nos images est l'algorithme OSEM-3D, dont le principe a été décrit dans le chapitre précédent (part. 1.2.1.3.3). L'acquisition ainsi que la reconstruction des données se font de manière tridimensionnelle. Pour cet algorithme, l'arrêt de la convergence est régit par le nombre d'itérations et le nombre de sous-ensembles prédéfinis par l'utilisateur. Ces derniers ont été respectivement fixés à 3 et 33, ce qui correspond aux paramètres par défaut utilisés en routine clinique. Les données ont été stockées en mode liste. Les corrections des coïncidences fortuites et diffusées, du temps mort, de la décroissance et de l'atténuation ont été appliquées aux données avant et durant le processus itératif (part. 1.2.1.2.4) [260]. La gestion de l'amplification du bruit est effectuée grâce au paramètre de relaxation λ (part. 1.2.1.3.3) qui a été laissé à sa valeur par défaut égale à 1, ce qui correspond au lissage minimal proposé. Les images ont été reconstruites en matrices de 144 x 144 avec des voxels isotropes de 4mm. Tous ces paramètres de reconstruction correspondent à ceux utilisés en routine clinique.

2.2.3/ TF : CORRECTION DU TEMPS DE VOL

La prise en compte de l'information liée au temps de vol permet d'améliorer le rapport signal sur bruit des images TEP et donc d'avoir la possibilité éventuelle de réduire le temps d'acquisition [261–265]. Le principe théorique de cette correction a été présenté dans le chapitre précédent (part. 1.2.1.2.2). Au niveau de la reconstruction de l'image, la principale différence avec les systèmes sans correction de TOF est l'utilisation d'un noyau représentant la distribution de probabilité du lieu d'annihilation du positon sur la LOR. Cette distribution est supposée gaussienne et permet de limiter la zone dans laquelle l'événement est positionné. Ce noyau est caractérisé par sa FWHM qui est généralement égale à la résolution temporelle du système de détection de manière à obtenir une distribution d'activité reconstruite la plus fidèle possible vis à vis de la distribution réelle. Plusieurs études ont étudié l'impact des paramètres spécifiques au TOF sur la qualité des images et ont montré que des tailles de noyaux inadaptés conduisent à des reconstructions incorrectes [266–268]. De manière générale, une FWHM trop faible conduira à des reconstructions présentant des contrastes réduits avec une mauvaise restitution des bords des structures présentes dans l'image : dans une zone homogène, l'activité reconstruite aura tendance à se concentrer au centre de l'objet imagé [265]. Un noyau plus large quant à lui, permettra une meilleure restitution du contraste mais au détriment de l'uniformité :

l'activité reconstruite se concentrera vers les bords des structures [265]. Dans le cas spécifique du GEMINI TF, Vandenberghe *et al.* (2009) [264] ont également étudié la troncature du noyau ainsi que l'utilisation d'un algorithme SSS pour la simulation du diffusé dédié au système TOF (TOF-SSS) [269, 270]. Les auteurs concluent qu'une troncature de noyau limitée ($\pm 3\sigma$) peut être autorisée afin d'accélérer la vitesse de reconstruction (tout en conservant suffisamment de contraste) et que l'utilisation de l'algorithme TOF-SSS pour la correction du diffusé est nécessaire pour obtenir une quantification comparable à celle obtenue avec la reconstruction TEP classique.

Un noyau de reconstruction TOF de 14,1 cm, correspondant à la configuration classique, a été utilisé lors de la reconstruction des images.

2.3/ PERFUSION TUMORALE ET IMAGERIE DYNAMIQUE

2.3.1/ CRÉATION DE L'IMAGE DE PERFUSION

Le calcul de la perfusion tumorale nécessite la modélisation de la cinétique du traceur dans les tissus. Cette dernière peut être réalisée grâce à un modèle compartimental. Plusieurs modèles existent [271] et dans le cadre de cette thèse nous avons décidé d'utiliser le modèle mono-compartimental de Mullani [94] (part. 1.2.3.3) qui a montré une très bonne corrélation avec les résultats de perfusion mesurés avec l' ^{15}O - H_2O , méthode de référence pour l'évaluation de la perfusion tumorale en TEP. Pour rappel, l'équation générale du débit sanguin (BF) de premier passage selon le modèle de Mullani s'écrit de la manière suivante :

$$BF = \frac{Q(t_m)}{\int_0^{t_m} Ca(t)dt} \quad (2.4)$$

La perfusion tumorale (BF) peut ainsi être estimée à partir de la radioactivité totale mesurée au niveau de la région d'intérêt tumorale ($Q(t_m)$) au temps du pic artériel t_m , divisée par l'intégrale de la concentration artérielle ($Ca(t)$) du traceur du temps 0 au temps t_m ($\int_0^{t_m} Ca(t)dt$), qui représente l'apport de traceur au cours du temps à l'organe cible. Cet apport artériel correspond à ce que nous appellerons la fonction d'entrée artérielle du modèle.

Contrairement aux méthodes conventionnelles qui demandent des prélèvements sanguins pour l'évaluation de la fonction d'entrée artérielle, Mullani *et al.* ont proposé une technique permettant d'estimer de manière non invasive cette dernière, en mesurant la radioactivité au cours du temps dans une région d'intérêt centrée sur l'aorte [94]. Cette approche nécessite une acquisition TEP sur les deux premières minutes suivant l'administration d'un bolus de ^{18}F -FDG au patient. Cette acquisition permet la reconstruction dynamique des images dites de premier passage, à partir desquelles peut être identifié le temps de pic artériel.

Le calcul de la perfusion tumorale va donc nécessiter trois principales étapes :

- l'acquisition des données TEP et la reconstruction des images dynamiques ;
- la délimitation des régions d'intérêts permettant l'évaluation de la fonction d'entrée artérielle et le calcul de l'activité dans la région tumorale ;

- la détermination du temps de pic artériel et la construction de l'image paramétrique de perfusion.

2.3.2/ ACQUISITION DES DONNÉES ET RECONSTRUCTION DES IMAGES DYNAMIQUES

Suivant les instructions du protocole TREN, une acquisition TEP de 8 minutes avec un seul pas de lit est effectuée directement après l'administration du radiotracer au patient. Toutes les données brutes sont sauvegardées en mode liste. À partir de ces dernières sont reconstruites des images dynamiques sur les 2 premières minutes avec un échantillonnage temporel de 10s en utilisant les mêmes paramètres de reconstruction.

Au total, lors de ce premier examen TEP, pour chaque patiente sont conservées :

- l'image TEP précoce acquise sur et reconstruite sur 8 minutes (matrice image 144 x 144, voxels de 4 x 4 x 4 mm³);
- l'image TDM correspondante (120 kVp, 120 mAs, matrice image 512 x 512, voxels de 1,17 x 1,17 x 2,5 mm³);
- une première série dynamique correspondant aux deux premières minutes d'acquisition et comprenant 12 images TEP reconstruites avec un échantillonnage temporel de 10s.

2.3.3/ DÉLINÉATION DES RÉGIONS D'INTÉRÊT

L'étude de la perfusion tumorale nécessite la délimitation de plusieurs volumes d'intérêt : le volume tumoral perfusé (VP) pour le calcul de l'activité dans la région d'intérêt tumorale, un volume centré sur l'aorte ascendante pour l'évaluation de la fonction d'entrée artérielle et un volume placé dans le sein controlatéral pour l'évaluation du bruit de fond ("bkg" pour *background*). Ce dernier n'est pas nécessaire à l'évaluation de la perfusion tumorale via le modèle de Mullani, mais cette région a néanmoins été délimitée en prévision d'études sur l'amélioration de la qualité des images, avec notamment la correction de l'effet de volume partiel (part. 1.2.2.3). La dimension de ce volume a été choisie de manière à être proche du volume tumoral segmenté. La délimitation des volumes d'intérêt a été faite sur les images TEP précoces de 8 minutes, car ces dernières ont une bien meilleure statistique que les images dynamiques, ce qui permet une meilleure identification des VOIs. Contrairement à la délimitation du volume métabolique (VM) pour laquelle de nombreuses études ont évalué différentes méthodes de segmentation (chap. 4), à notre connaissance, aucune étude ni recommandation n'ont été présentées dans la littérature pour la délimitation du volume tumoral perfusé. La délimitation manuelle, qui reste par ailleurs la méthode de référence pour l'évaluation des techniques de segmentation automatiques et semi-automatiques du VM, a donc été choisie pour segmenter les régions d'intérêts de l'image TEP précoce. Cette délimitation bénéficie de l'expertise du praticien dans l'analyse des images TEP ainsi que des informations anatomiques qu'il possède via l'acquisition TDM et de sa fusion avec les images TEP.

Les VOIs tracées sur l'image TEP précoce ont ensuite été extraites au format DICOM (*Digital Image and COmmunication in Medicine*) RT-Struct puis superposées sur les acquisitions dynamiques (fig. 2.3).

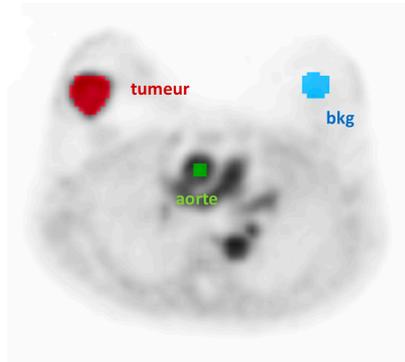


FIGURE 2.3 – Exemples des trois volumes d'intérêts : tumeur (rouge), aorte ascendante (vert) et tissus sain/bkg (bleu), délimitées sur la TEP précoce statique de 8 min

2.3.4/ IDENTIFICATION DU PIC ARTÉRIEL ET CRÉATION DE L'IMAGE PARAMÉTRIQUE DE PERFUSION

Les reconstructions dynamiques permettent l'évaluation de l'évolution de l'activité du radiotraceur lors des deux premières minutes après son administration au patient, soit avec un échantillonnage temporel de 10s. Une fois les volumes d'intérêts superposés sur les images dynamiques, la fonction d'entrée artérielle correspondant à l'évolution de l'activité moyenne du traceur en fonction du temps (la TAC pour *Time Activity Curve*) dans le VOI aortique est généré (fig. 2.4). Pour chaque série d'images dynamiques, le temps de premier passage t_m du traceur est visuellement identifié sur la TAC de l'aorte au niveau du point d'inflexion de la fonction, juste après le pic de premier passage, comme illustré sur la figure 2.4. L'intégrale de la concentration artérielle du traceur du temps 0 au temps t_m peut alors être calculée.

À chaque pas d'échantillonnage temporel correspond donc une image de la série dynamique. Une fois le temps de premier passage t_m identifié, l'image correspondante est sélectionnée. Sur cette image est mesurée la valeur d'activité tumorale $Q(t_m)$ pour chaque voxel de la VOI et une image paramétrique de perfusion est alors générée selon l'équation 1.7, avec dans chaque voxel, la valeur de la perfusion associée. Les TAC des tumeurs peuvent également être générées avec, pour un temps t donné, la valeur moyenne de l'activité totale mesurée $Q(t)$ au niveau de la région d'intérêt tumorale, comme illustré sur la figure 2.5.

En pratique, l'image de perfusion est générée pour la totalité de l'image TEP et non uniquement pour la région tumorale. Or, comme le modèle compartimental utilisé est dédié à l'évaluation de la perfusion tumorale, le calcul de la perfusion à l'extérieur de cette région d'intérêt n'a pas de sens. C'est pourquoi dans la suite des travaux de cette thèse, lorsque l'on parlera "d'image paramétrique de perfusion" on sous-entendra que l'on considère uniquement la région d'intérêt tumoral.

Le processus de la création de l'image paramétrique à partir de l'acquisition de l'image TEP précoce est résumé dans la figure 2.6.

Remarque : tous les calculs statistiques des VOIs (de premier et second ordre) des études présentées dans ce manuscrit, la création des images paramétriques et les étapes de segmentation

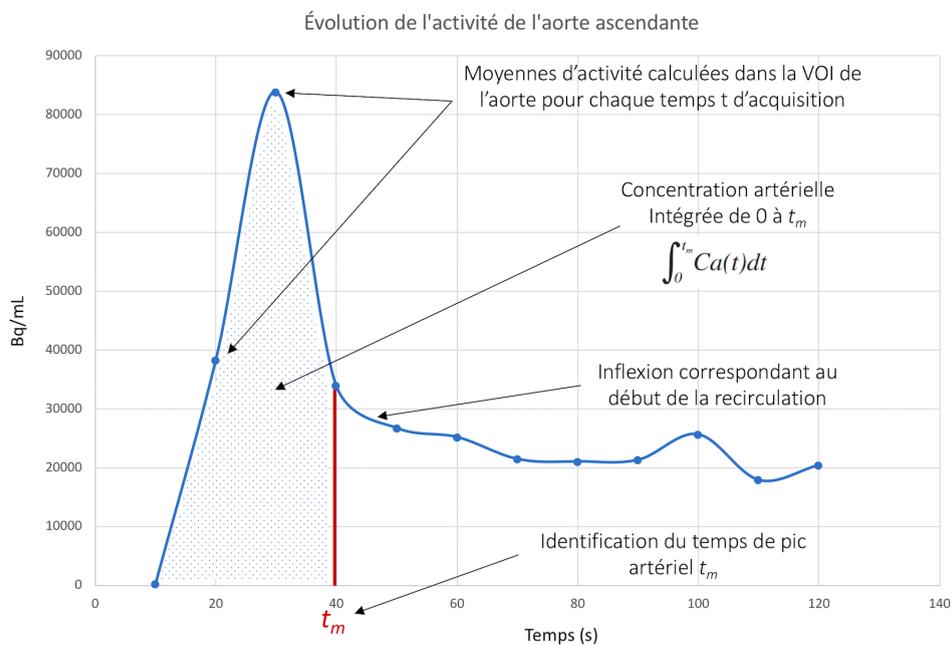


FIGURE 2.4 – Évolution de l'activité du traceur en fonction du temps, mesurée dans le VOI aortique lors d'une reconstruction dynamique avec un échantillonnage temporel de 5s (fonction d'entrée du modèle de Mullani). Détermination visuelle du temps t_m .

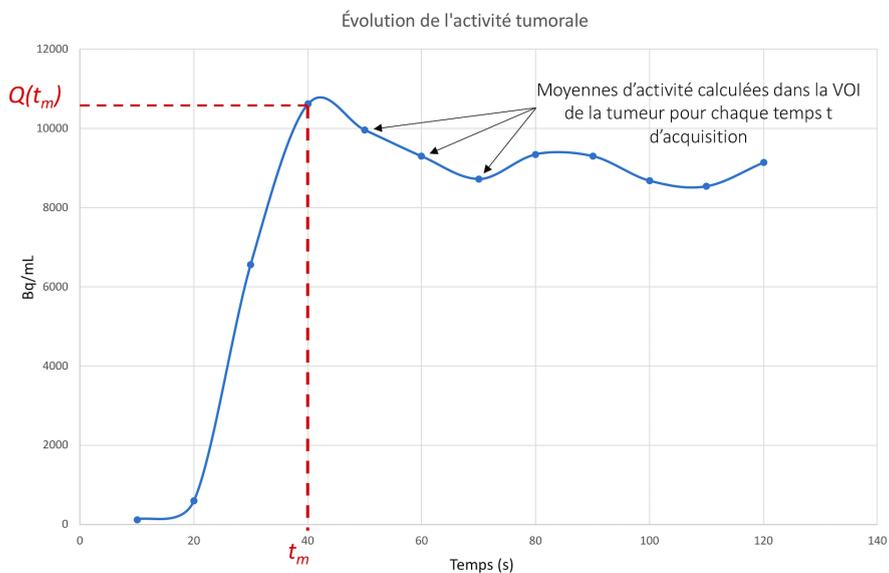


FIGURE 2.5 – Évolution de l'activité du traceur en fonction du temps, mesurée dans le VOI tumoral lors d'une reconstruction dynamique avec un échantillonnage temporel de 5s. La mesure de $Q(t_m)$ représente la quantité de traceur dans la tumeur au temps t_m .

(part. 2.4) ont été effectuées de manière automatique via l'utilisation d'outils click du logiciel *vv* [272]. Des outils permettant le calcul des images de SUV et de perfusion ont également été

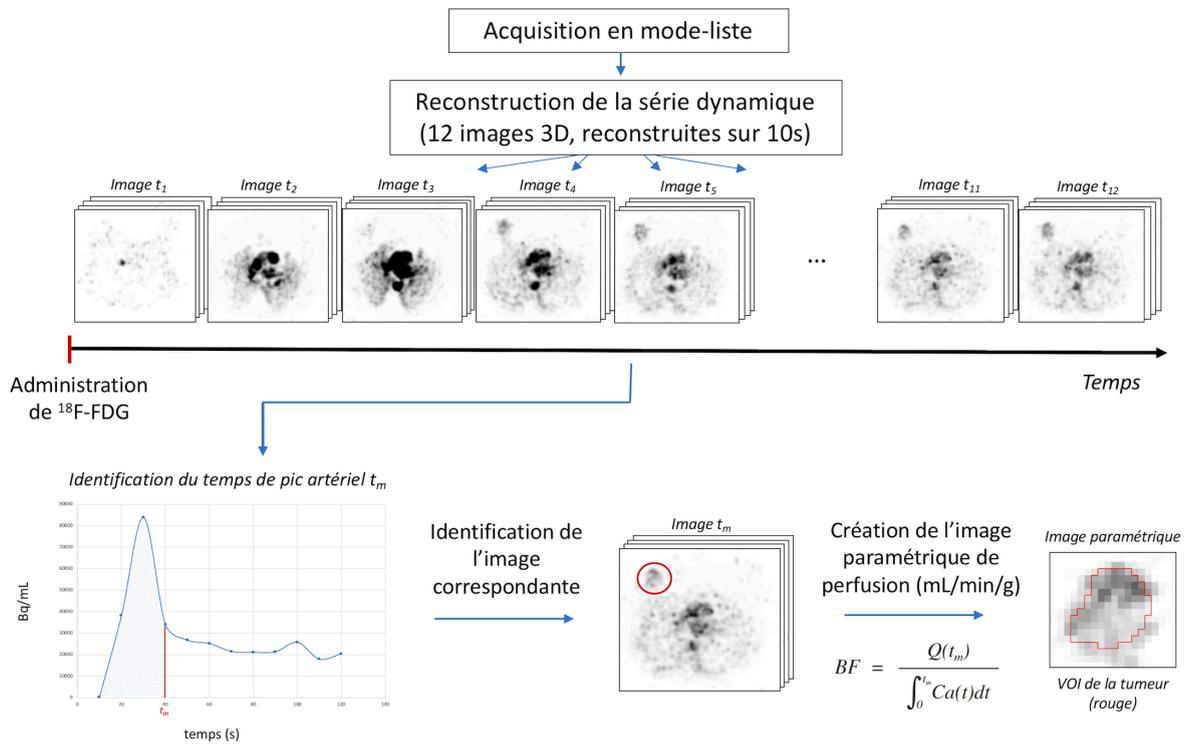


FIGURE 2.6 – Résumé du processus de création de l'image paramétrique de perfusion.

développés et intégrés au logiciel *vv*.

2.3.5/ RECONSTRUCTION DYNAMIQUE ET STATISTIQUE DANS L'IMAGE

L'image paramétrique va donc permettre la quantification de la perfusion tumorale. Toutefois, le signal étant reconstruit sur seulement 10s, ces images dynamiques ont une bien plus faible statistique que les images TEP statiques reconstruites durant 2 minutes et plus (fig. 2.7). Il existe plusieurs méthodes pour améliorer la qualité des images de faible statistique. Dans le contexte spécifique des acquisitions dynamiques, plusieurs approches ont été étudiées. On peut citer par exemple la correction du biais statistique dans les images [273, 274] ou encore l'optimisation des protocoles d'acquisition. On peut optimiser l'activité injectée de manière à maximiser la qualité des images dynamiques [275, 276] ou en étudiant l'influence de l'échantillonnage temporel sur la variabilité et le biais des mesures cinétiques [277–279]. D'autres équipes ont également évalué l'impact du choix du modèle cinétique, de manière à améliorer de manière générale la quantification des paramètres biologiques [279–282]. Enfin, de nombreuses études se sont concentrées sur des algorithmes de reconstruction dédiés aux images dynamiques. En effet, de manière conventionnelle, les acquisitions dynamiques en clinique sont reconstruites à partir de séquences d'images 3D indépendantes sur lesquelles sont appliqués des modèles cinétiques. Ces derniers permettent ensuite d'estimer des paramètres cinétiques et/ou fonctionnels dont les valeurs sont représentées sous la forme d'une image paramétrique de la région d'intérêt considérée. Or, le bruit présent dans ces séquences d'images 3D étant généralement bien plus élevé que

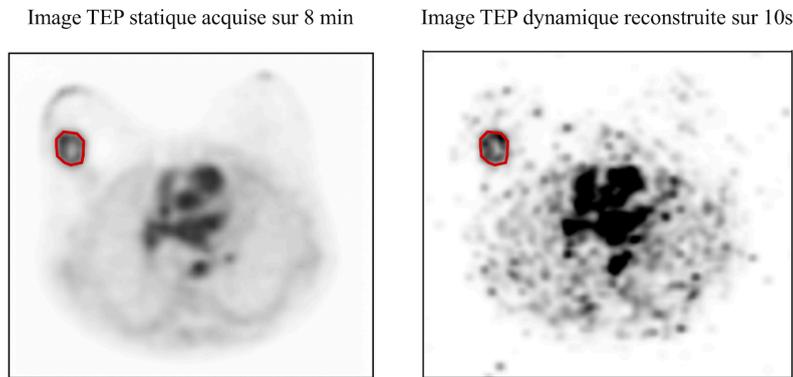


FIGURE 2.7 – Illustration visuelle de la qualité d’une image TEP statique reconstruite sur 8 minutes (gauche) et d’une image TEP dynamique reconstruite sur 10s (droite). En rouge le VOI de la tumeur, délinéé sur l’image statique puis superposée à l’image dynamique.

dans les images TEP classiques, il peut notamment induire un biais dans la quantification des paramètres mesurés. De manière optimale, la distribution du bruit, qui est spatialement corrélé dans les séquences d’images 3D reconstruites, devrait être prise en compte dans les analyses post-reconstruction des images. Cependant, pour les modèles d’algorithmes de reconstruction itératifs communément utilisés, la distribution du bruit reste rarement connue avec précision et varie en fonction des paramètres de reconstruction utilisés et du type d’algorithme appliqué. Des algorithmes dédiés aux reconstructions dynamiques ont alors émergés. Ces derniers tendent à pallier aux limites liées à la faible statistique en intégrant un modèle temporel de distribution du radiotracteur directement dans l’algorithme. Ces techniques de reconstruction 4D directes sont appliquées sur les données brutes dont le modèle de bruit est connu (Poisson). Elles permettent ainsi une modélisation de ce dernier plus juste et peuvent également implicitement prendre en compte les corrélations spatiales du bruit et imposer les contraintes de modélisation temporelle tout au long de la procédure de reconstruction itérative. Des études de reconstructions 4D ont été réalisées à l’aide de nombreux radiotraceurs, y compris le ^{18}F -FDG [283–288] et ont présenté des résultats très encourageant, avec une réduction significative du biais des mesures des paramètres. Les présentations des principaux algorithmes 4D et leurs caractéristiques peuvent être consultés dans la revue de Andrew J Reader et Jeroen Verhaeghe (2014) [289].

Dans le cadre d’un encadrement de stage de master 2, un travail a été entrepris afin de caractériser le biais quantitatif en fonction de la statistique des images pour ensuite essayer d’optimiser la qualité d’image de faible statistique, par l’ajustement des paramètres de reconstruction et la correction de l’effet de volume partiel. Cette étude, disponible en annexe de ce manuscrit (Annexe C), a montré des résultats prometteurs. Néanmoins, cette dernière ayant été réalisée tardivement au court de la thèse, il n’a pas été possible, faute de temps, d’intégrer ces améliorations aux images cliniques générées pour les études présentées dans ce manuscrit.

2.4/ MÉTABOLISME ET VOLUME TUMORAL

Dans le chapitre précédent, nous avons abordé l'étude du métabolisme glucidique via la TEP au ^{18}F -FDG et présenté différents indices permettant son évaluation, comme le SUV_{max} , SUV_{mean} , etc. La mesure du volume métabolique de la tumeur via la TEP au ^{18}F -FDG est également une approche intéressante pour le développement de nouveaux paramètres métaboliques d'ordre pronostique [290] ou prédictif [291] de la réponse au traitement, pour le suivi thérapeutique [292], ou encore pour la définition des contours tumoraux (GTV pour *Gross Tumor Volumes*) en radiothérapie [293]. Il existe plusieurs méthodes pour évaluer le VM sur les images TEP dont les principales sont décrites dans la suite de ce chapitre.

2.4.1/ SEGMENTATION DES IMAGES EN TEP

De manière générale, la segmentation des volumes d'intérêts est une étape importante dans l'analyse et la quantification des images TEP. Le principe de la segmentation vise à partitionner l'image en régions dont les voxels vont partager des caractéristiques communes, comme l'intensité, la texture, etc. En imagerie TEP, la segmentation vise principalement à distinguer les régions tumorales métaboliquement actives des régions avoisinantes généralement considérées comme le fond de l'image. La délimitation de ces hyperfixations a plusieurs applications :

- la détermination du VM qui va permettre de caractériser le stade tumoral initial et son évolution au cours du traitement ;
- la définition d'une région d'intérêt tumorale sur laquelle vont pouvoir être calculés de nombreux indices métaboliques ou perfusionnels dans le cas d'une imagerie de perfusion ;
- l'aide à la détermination du volume tumoral pour les applications de radiothérapie.

Cependant, la segmentation va se heurter en TEP à plusieurs obstacles. L'effet de volume partiel, par "l'étalement" des contours des structures fixantes et l'effet de fraction tissulaire, va impacter la qualité de la segmentation. Le bruit élevé dans les images TEP va également desservir la délimitation des contours, en particulier pour les images à faible statistique comme les images dynamiques ou encore pour les structures à faible contraste. Les mouvements du patient vont introduire un flou dans l'image qui va rendre encore plus difficile la délimitation exacte des contours des régions, et la possible hétérogénéité de l'accumulation du FDG dans les tumeurs complexifie également la segmentation de la région d'intérêt tumorale.

Afin d'évaluer le VM, la segmentation manuelle est une approche intuitive. Elle bénéficie de l'expertise du praticien dans l'analyse des images TEP, des informations médicales *a priori* qu'il possède, que ce soit en anatomie ou en oncologie en général mais également des informations présentes dans le dossier patient. Cependant, la segmentation manuelle reste très subjective et souffre donc de variations inter- et intra- opérateurs [294–298]. Par ailleurs, la délimitation manuelle se fait généralement coupe par coupe, elle est donc longue et fastidieuse à mettre en place.

Mues par le besoin d'outils de quantification précis et robustes pour l'analyse des images TEP, d'autres méthodes de segmentation, automatiques ou semi-automatiques [299] spécifiques aux images TEP ont alors émergé afin d'essayer de pallier aux limitations de la segmentation ma-

nuelle [300]. Cette section s'appuie sur notamment sur les revues de littératures de [299] et [301] ainsi que sur les dernières recommandations en date de l'*American Association of Physicists in Medicine (AAPM)* [302].

2.4.1.1/ LES MÉTHODES PAR SEUILLAGE

Les fixations tumorales en TEP sont généralement caractérisées par un hypersignal dans les images. Une des manières de discriminer la tumeur du fond de l'image est d'utiliser un seuil d'intensité au-delà duquel on va considérer le tissu comme tumoral. De nombreuses méthodes reposant sur ce principe ont été étudiées. La plupart d'entre elles reposent sur une simple discrimination des voxels en intensité, sans prise en compte d'informations spatiales. Toutefois, en raison de la variabilité inter- et intra- tumorale et des caractéristiques des images TEP (résolution spatiale, bruit), il n'y a à ce jour pas de consensus général sur la sélection d'un niveau de seuillage. Par conséquent, la détermination optimale du seuil demeure une tâche difficile dont l'étude reste un domaine de recherche actuel. Parmi les différentes techniques de segmentation, on distingue les méthodes à seuil fixe, les méthodes à seuil adaptatif et les méthodes adaptatives itératives qui reposent sur l'application successive de plusieurs seuils, jusqu'à convergence.

- Seuillages fixes

En seuillage fixe, tous les voxels dont l'intensité est supérieure à une valeur seuil fixée sont considérés comme appartenant au volume tumoral. Ce seuil correspond généralement soit à une valeur absolue de SUV, souvent supérieure à 2,5 SUV [303], soit à un seuil relatif à la valeur d'intensité maximale, généralement de 40 à 43% du SUV_{\max} en clinique [304]. Ces méthodes ont l'avantage d'être rapides et très simples à utiliser. Cependant, lorsque le seuil d'intensité est fixé comme un pourcentage de l'intensité maximale dans la tumeur, ce dernier va directement dépendre de la valeur du SUV_{\max} qui est particulièrement sensible au bruit dans l'image ainsi qu'à l'hétérogénéité tumorale. Une faible variation du SUV_{\max} peut engendrer une variation importante des contours tumoraux. Par ailleurs, pour les lésions à faible contraste avec le fond de l'image ou pour les petites lésions très impactées par l'EVP, un seuillage fixe va avoir tendance à inclure les voxels du fond dans la région d'intérêt, avec pour conséquence une surévaluation du volume métabolique. Enfin, les méthodes par seuillage fixe vont être généralement incapables de donner des résultats pertinents en présence d'hétérogénéité de la fixation, occasionnant généralement là aussi une sous-évaluation du volume d'intérêt. Les segmentations issues de seuillages fixes vont donc dépendre très fortement de la taille des lésions, du contraste d'activité entre la tumeur et le fond, du bruit présent dans l'image [305] ainsi que de l'hétérogénéité des lésions. Afin d'essayer d'intégrer ces paramètres aux méthodes de segmentation, des méthodes un peu plus complexes utilisant des seuillages adaptatifs ont donc émergé.

- Seuillages adaptatifs

La plupart des méthodes de seuillages adaptatifs prennent en compte certaines propriétés des sources (= lésions dans le cas clinique) pour le calcul du seuil optimal, comme le rapport d'activité entre la source et le fond [303, 306–308], la taille et la forme de la source [305], ou encore le flou induit par le mouvement [309, 310]. Plusieurs revues sur la segmentation en TEP en général ont répertorié les principales méthodes de seuillages adaptatifs et les ont comparé entre elles [299, 301, 311]. La plupart de ces méthodes déterminent le seuil optimal en s'appuyant sur des acquisitions de fantômes où la taille, la forme des sources et les rapports d'activité source/fond vont permettre de déterminer pour chaque situation le seuil à appliquer. Cette approche rend ces méthodes applicables à un plus large éventail de cas et tend à augmenter leur robustesse notamment contre l'EVP [312, 313]. Dans l'ensemble, la principale limitation de ces méthodes repose sur le fait que leurs paramètres doivent être optimisés pour chaque imageur et chaque configuration (protocole d'acquisition, de reconstruction etc.) [306]. Par ailleurs, le fait que ces méthodes soient optimisées via des fantômes composés généralement de formes géométriques simples (sphères) homogènes les rend pour la plupart assez peu performantes pour segmenter les volumes hétérogènes.

- Méthodes itératives

D'autres méthodes reposent sur le calcul d'une fonction seuil qui est préalablement calibrée en fonction des caractéristiques lésions/sources testées. La détermination du seuil optimal se fait itérativement jusqu'à convergence de cette fonction. Par exemple, dans les méthodes de Jentzen *et al.* (2007) ou Drever *et al.* (2007), la taille de la lésion ainsi que le rapport d'activité de la tumeur par rapport au fond sont pris en compte. La détermination du seuil optimal se fait alors de manière itérative jusqu'à convergence, i.e. lorsque le volume segmenté n'évolue plus entre deux seuils successifs. Ces méthodes nécessitent une étape initiale d'étalonnage afin de définir la fonction qui donne le seuil optimal en fonction des paramètres considérés, mais ne nécessitent pas obligatoirement une étape d'optimisation préalable [314], contrairement aux méthodes de seuillage adaptatifs non itératifs.

De manière générale, les méthodes par seuillages peuvent donner de bons résultats pour les distributions d'activité homogène mais restent néanmoins limitées pour les lésions de forme complexe et/ou de fixation hétérogène [315]. Les seuillages fixes restent néanmoins déconseillés selon les dernières recommandation de l'AAPM [302].

2.4.1.2/ LES MÉTHODES BASÉES SUR LA DÉTECTION DE CONTOURS

La segmentation de régions d'intérêt sur les images TEP peut également être réalisée à partir de critères morphologiques basés sur la détection des contours [316]. L'avantage de cette approche repose sur le fait que l'activité à l'intérieur et à l'extérieur de la région d'intérêt n'a pas besoin d'être uniforme ni d'être constante le long du contour pour obtenir une segmentation réussie. Avec cette approche, on s'intéresse donc aux variations des intensités dans l'image pour détecter les transitions entre différentes régions. Pour cela, des méthodes dérivatives sont généralement utilisées pour générer une image gradient. Les maxima de cette image

représentent alors les contours de l'objet à segmenter. On peut distinguer deux principales approches pour localiser ces maxima : la méthode des "contours actifs" (également connus sous l'appellation *snakes*) [316,317] et la méthode de "partage des eaux" (*watershed*) [318].

- Les contours actifs

La méthode des contours actifs se base sur un *a priori* initial géométrique de l'objet à segmenter qui se déforme itérativement pour s'adapter aux contours de l'objet. La déformation de ce contour est gérée par une fonction de coût appelée énergie. Cette dernière est définie de manière à prendre en compte à la fois la régularité du contour et les contraintes liées à l'image (intensité, forme, texture, etc). Cette technique a été adaptée aux images TEP [316, 319, 320] notamment en complément d'une méthode de segmentation basée sur le seuillage, en ajoutant une contrainte liée aux contours de la région à segmenter [321].

La méthode des "fronts de propagation", ou *Level Set*, a également été proposée [322] comme technique de modélisation des contours actifs et testée sur les images TEP [321, 322]. Cette méthode remplace l'information géométrique *a priori* des *snakes* par une interface bi ou tri-dimensionnelle qui va se propager en "suivant" les contours de l'objet. La vitesse de propagation va dépendre de nombreux paramètres pour lesquels le modèle doit être adapté [322, 323].

- La méthode de partage des eaux

La méthode du partage des eaux quant à elle considère l'image gradient comme une carte topographique avec plusieurs niveaux "d'altitude". Parmi les approches connues, on peut citer celle de la "montée des eaux" où l'on simule une montée des eaux à partir de minima régionaux. Au fur et à mesure de la montée des eaux, certaines régions, correspondant à deux minima distincts et donc initialement isolées, fusionnent. La ligne correspondant à l'altitude à partir de laquelle les bassins fusionnent permet d'en déterminer la frontière. D'autres approches, qui utilisent notamment des algorithmes par ruissellement ou topologiques, sont également utilisés pour construire cette ligne de partage des eaux [324, 325].

De manière générale, les méthodes de détection de contours vont être sensibles à la résolution spatiale, au flou et au bruit présent dans les images. Il est possible néanmoins de compenser ces effets par des étapes de pré-traitement telles que des méthodes de déconvolution ou de correction d'EVP pour le flou, ou encore par des méthodes de débruitage [326, 327]. Néanmoins, quelles que soient les méthodes de détection de contours, leurs performances vont dépendre à la fois des paramètres à ajuster (comme la vitesse pour les level sets), des filtres utilisés ainsi que des potentielles corrections pré-traitement effectuées.

2.4.1.3/ LES MÉTHODES DE CLASSIFICATION

Le principe des méthodes de classification repose sur la labellisation de chaque élément de l'image à une région donnée. Cette classification peut se faire via des méthodes dites d'agrégation qui vont chercher à diviser l'image en régions distinctes en groupant les voxels de mêmes

caractéristiques préalablement déterminées. La classification peut également être réalisée à l'aide d'approches probabilistes et statistiques, qui vont prendre en compte la répartition statistique des voxels dans l'image, en plus de leur information spatiale et d'intensité. Enfin, des méthodes d'apprentissage supervisé vont permettre d'estimer statistiquement les dépendances dans les données à des fins de classification. Ces méthodes sont largement étudiées à l'heure actuelle car elles montrent des résultats prometteurs [328].

- Les méthodes d'agrégation

Pour les méthodes d'agrégation, les voxels sont regroupés dans une même région s'ils ont des caractéristiques communes ou proches. L'une des méthodes d'agrégation la plus intuitive est la croissance de région (*region growing*), qui commence avec l'implantation d'une "graine", i.e. un point de départ au processus itératif, à partir duquel les voxels voisins vont progressivement être inclus dans la région initiale s'ils satisfont certains critères de similitude [329, 330]. Néanmoins, si les critères sont fixes, l'algorithme de croissance de région se rapproche très fortement d'un seuillage adaptatif auquel une contrainte de connectivité a été ajoutée.

L'une des méthodes classiques d'agrégation est la méthode des *k*-moyennes (*k-means*) qui vise à affecter un ensemble de point (ici les voxels) à des classes (ici des régions) dont le nombre est préalablement connu [331]. L'appartenance à une classe est déterminée par la moyenne des caractéristiques des éléments contenus dans cette classe. L'inclusion d'un élément est faite de manière à toujours minimiser la distance entre l'élément et la moyenne de la classe. Le processus d'inclusion d'un nouvel élément est généralement réalisé de manière itérative, une nouvelle moyenne est donc recalculée à chaque fois jusqu'à convergence de l'algorithme. Une autre méthode consiste à attribuer les éléments non pas directement à une classe, mais à une probabilité d'appartenance à la classe. Cette méthode est appelée *Fuzzy-C-means* (FCM) [332] car elle se base sur une approche de logique floue. La méthode de FCM a été utilisée pour la segmentation des images TEP [333, 334] et a montré de meilleurs résultats que certaines méthodes par seuillage (fixe et adaptatif) [335].

La principale limite de ces approches repose sur la connaissance *a priori* du nombre de classes. Elles sont également sensibles au bruit ainsi qu'à l'hétérogénéité des lésions. Toutefois, la prise en compte de ces deux derniers aspects par l'ajout d'un filtrage anisotropique et par la décomposition en ondelette à trous (*wavelet transform*) a permis d'améliorer la méthode FCM [336].

- Les approches statistiques

Les approches statistiques visent à classer les voxels en différentes régions en fonction de leurs propriétés statistiques de répartition dans l'image par une modélisation du bruit mais également en fonction de leurs caractéristiques spatiales. En effet, le bruit étant corrélé spatialement dans les images TEP reconstruites, la probabilité d'un voxel d'appartenir à la même classe que ses voisins est plus élevée que celle d'appartenir à une autre classe. Parmi les approches qui ont été proposées, la plupart sont basées sur l'inférence bayésienne. Les approches qui modélisent les interactions spatiales des voxels sont généralement basées sur les champs de

Markov et les modélisations du bruit sont le plus souvent définies à l'aide de distributions gaussiennes, même si d'autres modèles plus avancés ont également été proposés [337]. Parmi les études qui ont utilisé sur des images TEP des algorithmes de segmentation fondés sur des méthodes statistiques, on peut citer celle de Montgomery *et al.* [338] qui a couplé les champs de Markov à une approche multi-échelle qui repose sur une transformée en ondelettes à trous ou encore Aristophanous *et al.* [339] dont la méthode s'appuie sur une approche de mélange de modèle gaussiens (*gaussian mixture models*). Néanmoins, bien que ces algorithmes proposent des modèles robustes vis à vis de distributions bruitées, ils ne prennent pas en compte la notion de flou inhérente aux images TEP (part. 1.2.2). Afin d'essayer de palier à cette limitation, une autre approche, a été proposée [340–342]. Cette méthode de segmentation, appelée FLAB pour *Fuzzy Locally Adaptive Bayesian*, a été évaluée sur fantômes physiques ainsi que sur images TEP simulées et cliniques et a montré de très bons résultats en comparaison à d'autres méthodes de segmentation, y compris sur des lésions hétérogènes [343].

- Les méthodes par apprentissage supervisé

La particularité de ces méthodes réside dans l'apprentissage obligatoire au préalable du modèle de segmentation sur un ensemble de données dédiées, appelé généralement l'ensemble d'entraînement (*training set*). Dans le cas, par exemple, où les voxels doivent être classés en deux régions (représentant typiquement la tumeur et les tissus sains), l'entraînement du modèle se fait sur des images dans lesquelles l'appartenance (la labellisation) des voxels à l'une des deux régions est connue. Le principe de l'entraînement est ensuite de permettre à l'algorithme d'identifier sur un grand nombre d'images des paramètres discriminants qu'il juge représentatifs de chaque région. Ce sont ces paramètres qui seront ensuite utilisés sur de nouvelles images TEP à des fins de classification. L'application à la segmentation des images TEP est pour le moment expérimentale mais connaît un grand essor depuis quelques années. Un premier challenge MICCAI [328] a permis de comparer, sur des données de fantômes physiques ainsi que sur des images TEP simulées et cliniques, plusieurs techniques de segmentation d'images TEP dont plusieurs méthodes d'apprentissage supervisé. La méthode identifiée comme la plus performante s'est avérée être une technique d'apprentissage supervisé utilisant une approche par réseau de neurones convolutifs (CNN pour *Convolutional Neural Network*). Toutefois, de manière générale, pour les méthodes par apprentissage supervisé, la nécessité d'avoir un jeu de données d'entraînement peut être rapidement limitant.

De nombreuses méthodes de segmentation plus ou moins complexes ont donc été développées et/ou adaptées à l'imagerie TEP. Cependant, à l'heure actuelle, aucune de ces méthodes ne fait consensus. En ce basant sur l'analyse comparative des techniques de segmentation en TEP de Foster *et al.* (2014) ainsi que sur les études Schaefer *et al.* (2015) et Zaidi *et al.* (2012), nous avons décidé d'utiliser la méthode de seuillage adaptatif de Schaefer *et al.* (2008), en raison de son accessibilité et de ses performances, évaluées à la fois sur fantômes et sur images cliniques. L'étape préalable d'étalonnage sur fantôme et l'application de cette dernière à nos images cliniques sont développées dans la partie qui suit.

2.4.2/ DÉTERMINATION DES PARAMÈTRES NÉCESSAIRES À LA MÉTHODE DE SEUILLAGE ADAPTATIF ET SEGMENTATION DU VOLUME MÉTABOLIQUE

La méthode de Schaefer est une méthode de segmentation par seuillage adaptatif dont l'application aux images cliniques nécessite une calibration préalable sur fantôme physique.

2.4.3/ FANTÔME NEMA-IEC

Les fantômes physiques sont des objets utilisés en imagerie pour la caractérisation, via des tests standardisés, des performances machines. Ils ont généralement des parois en plexiglas avec des inserts et des volumes creux permettant l'ajout à l'intérieur de solutions radioactives. Leur géométrie et leur composition étant connus, ils servent ainsi de référence terrain pour l'évaluation des performances des méthodes d'estimation d'activité. Il existe plusieurs types de fantômes physiques. On peut citer les fantôme de type "Jaszczak", cylindres creux dont les dimensions sont choisies pour simuler la largeur d'un patient et qui peuvent contenir des sphères remplissables ou pleines pour simuler des lésions tumorales ; les fantômes développés spécifiquement pour l'imagerie cérébrale (RS 901T Head Phantom, Hoffman 3-D Brain Phantom) avec notamment des compartiments permettant de simuler la distribution de l'activité dans le cerveau ; ou encore les fantômes dit anthropomorphes qui simulent le thorax humain de manière plus ou moins réaliste (NEMA-IEC, Torso ECT/TOR/P, Head/Thorax phantom RS 800T). Ces derniers contiennent différents inserts simulant certaines parties anatomiques, comme la colonne vertébrale, les poumons, le foie ou le coeur. Le fantôme anthropomorphe NEMA-IEC a été uti-



FIGURE 2.8 – Fantôme anthropomorphe NEMA IEC [203]

lisé pour l'étape de calibration de la méthode de Schaefer (fig. 2.8). Ce fantôme est un cylindre de dimension $300 \times 230 \times 180$ mm. Il possède un insert creux cylindrique d'un diamètre de 51 mm qui s'étend verticalement en son centre et dont le remplissage est généralement fait avec un matériau de densité égale aux poumons. Il dispose également de 6 sphères creuses de 10, 13, 17, 22, 28, 37 mm de diamètre, centrées sur un rayon de 57 mm. Le fond ainsi que les sphères du fantôme peuvent être remplis avec des solutions avec ou sans radioactivité. Les volumes (le fond ou les sphères individuellement) peuvent être remplis avec des concentrations en activité différentes, ce qui permet d'étudier différents contrastes dans l'image.

2.4.4/ SEUILLAGE ADAPTATIF DE SCHAEFER *et al.* : ÉTALONNAGE SUR FANTÔME

La méthode de Schaefer *et al.* est une méthode de segmentation par seuillage adaptatif dont la valeur du seuil va être dépendante du contraste source/fond [308]. Dans leur étude, Schaefer *et al.* relie le seuil optimal (TS) en fonction de la valeur moyenne de SUV dans la région correspondant à 70% du SUV_{\max} et de l'activité du fond (bkg) de la manière suivante :

$$TS = a \times mSUV_{70} + b \times Bkg \quad (2.5)$$

Afin de pouvoir déterminer ce seuil TS, une procédure d'étalonnage décrite dans [308] est nécessaire pour la détermination des coefficients a et b qui varient en fonction des conditions d'acquisition des images TEP (ces derniers sont spécifiques à l'imageur, au protocole d'acquisition et de reconstruction des images).

- Protocole d'acquisition

Plusieurs acquisitions correspondant à des rapports de concentration d'activité sphère/fond de 2 :1, 5 :1, 10 :1, 20 :1 et 40 :1, ont été réalisées sur un fantôme NEMA-IEC, par ajouts progressifs d'activité dans le fond du fantôme. Les rapports des concentrations d'activité ont été choisis de manière à couvrir toute la gamme des rapports de contraste observés dans les images TEP tardives. Les reconstructions ont ensuite été faites selon le même protocole que les images TEP tardives du protocole TREN.

- Détermination des coefficients

Afin d'évaluer l'activité du fond, une série de six VOIs a été tracée à une distance d'au moins 5 mm des sphères. Le SUV moyen de toutes ces VOIs a été utilisé comme activité moyenne du fond. Pour déterminer le seuillage TS optimal pour chaque sphère, une segmentation par croissance de région a été effectuée, en faisant varier les seuils de manière itérative, par pas de 0,01 (SUV). La graine à partir de laquelle débute le processus de croissance de région a été placée manuellement au centre de chaque sphère. Une contrainte de volume a ensuite été appliquée, de manière à choisir le seuil d'intensité TS qui minimise l'erreur entre le volume de la région segmentée et le volume réel des sphères (à $\pm 2\%$). Ces seuils ont été recueillis pour chacune des sphères et pour chaque contraste étudié. Une autre série de VOIs ont été générés avec un seuillage fixé à 70% du SUV_{\max} . La valeur moyenne de SUV dans ces VOIs ($mSUV_{70}$) a été récoltée. Pour chaque acquisition et pour chaque sphère, le contraste C sphère/fond de l'image a ensuite été déterminé de la manière suivante :

$$C = \frac{mSUV_{70} - Bkg}{Bkg} \quad (2.6)$$

Les seuils optimaux normalisés TS_{norm} ($TS/(mSUV_{70} - Bkg)$) en fonction des valeurs de contraste ont été tracés pour chaque taille de sphère, comme illustré sur la figure 2.9. Une courbe d'ajus-

tement a ensuite été effectuée sur les données selon la relation suivante :

$$\frac{TS}{mSUV_{70} - Bkg} = (k/C) + a \quad (2.7)$$

Deux fonctions de régression ont été utilisées, une première pour les volumes inférieurs à 1,15 mL et une seconde pour les volumes supérieurs à 1,15 mL. À partir de cette relation, la valeur du seuil optimal TS peut être retrouvée via l'équation 2.5 en prenant $b = k - a$.

À l'issue de cette cet étalonnage, les valeurs des paramètres a et b ont pu être déterminés pour le système GEMINI TF dans les conditions d'acquisition et de reconstruction étudiées. Leurs valeurs sont les suivantes :

- $a = 0,44261$ et $b = 0,48281$ ($R = 0.96$) pour des sphères de volume égal ou supérieur à 1,15 mL
- $a = 0,61352$ et $b = 0,62179$ ($R = 0.99$) pour des sphère de volume inférieur à 1,15 mL

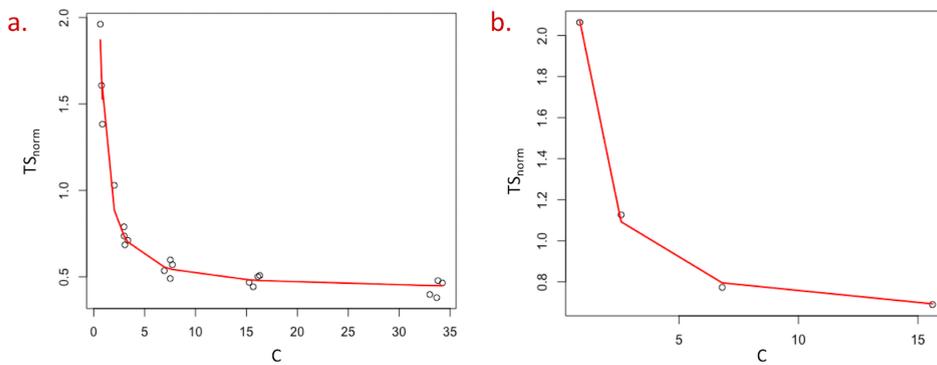


FIGURE 2.9 – Seuils optimaux normalisés TS_{norm} ($TS/(mSUV_{70} - bkg)$) en fonction du contraste image pour les sphère de volume supérieur à 1,15 mL (a) et inférieur à 1,15 mL (b)

Une fois ces coefficients déterminés, la méthode de seuillage adaptatif a pu être appliquée aux images TEP tardives préalablement converties en valeurs de SUV selon l'équation 2.5.

2.4.5/ SEUILLAGE ADAPTATIF DE SCHAEFER *et al.* : APPLICATION AUX IMAGES CLINIQUES

L'application de l'algorithme de Schaefer *et al.* sur les données cliniques nécessite le tracé d'une région d'intérêt centrée sur la tumeur afin d'identifier le SUV_{max} pour pouvoir ensuite générer le VOI correspondant à 70% du SUV_{max} sur lequel sera calculé le SUV moyen ($mSUV_{70}$). Il est également nécessaire de tracer un volume d'intérêt sur le sein controlatéral sur lequel sera calculé l'activité du fond. De manière à automatiser cette étape pour l'intégralité de la base de données, nous avons décidé d'utiliser les VOIs qui ont été tracés sur la TEP de premier passage pour localiser la région tumorale et la région de bruit de fond du sein controlatéral de la TEP tardive. Les patientes ayant été re-positionnées entre ces deux examens, une étape de recalage des images a été nécessaire afin de pouvoir transposer les VOIs d'un examen à l'autre.

- Recalage des régions d'intérêt

Le recalage est un procédé qui implique deux images distinctes, une image dite mobile $I_M(x)$ et une image fixe $I_F(x)$, toutes deux de dimension d et définies dans leur propre domaine spatial : $\Omega_F \subset \mathbb{R}^d$ et $\Omega_M \subset \mathbb{R}^d$ respectivement. Lors d'un recalage, l'image mobile $I_M(x)$ est déformée de manière à correspondre à l'image fixe $I_F(x)$. Le principe du recalage est donc de trouver un déplacement $u(x)$ qui permet à $I_M(x + u(x))$ d'être aligné spatialement sur $I_F(x)$, ce qui revient à trouver une transformation $T(x) = x + u(x)$ qui aligne spatialement $I_M(T(x))$ sur $I_F(x)$. La qualité de l'alignement est définie par une mesure de distance ou de similarité. De manière générale, le problème de recalage est formulé comme un problème d'optimisation où il faut trouver la transformation qui minimise une certaine fonction de coût C :

$$\hat{T} = \underset{T}{\operatorname{argmin}} C(T, I_F, I_M) \quad (2.8)$$

Cette dernière correspond à une mesure de similarité avec éventuellement l'ajout d'un terme de régularisation. Dans le cas de notre étude, le recalage souhaité n'a pas pour objectif d'être extrêmement précis, il doit seulement permettre une identification "grossière" de la zone tumorale dans laquelle sera recherchée la valeur d'intensité maximale de la tumeur (SUV_{\max}). Pour cette raison, il a été décidé de se limiter aux transformations rigides (translations, rotations). L'utilisation de ce type de transformation ne nécessite pas de terme de régularisation dans la fonction de coût. Pour notre étude, l'Information Mutuelle (IM) a été utilisée comme mesure de similarité [344, 345]. Cette mesure suppose qu'il existe une relation statistique entre les distributions de probabilités des intensités dans l'image fixe et l'image mobile :

$$IM(T, I_F, I_M) = \sum_{f \in \Omega_F} \sum_{m \in \Omega_M} p(f, m) \log\left(\frac{p(f, m)}{p_F(f)p_M(m)}\right) \quad (2.9)$$

avec $p(f, m)$ la probabilité conjointe de trouver la valeur f dans l'image fixe là où l'on trouve la valeur m dans l'image mobile. $p_F(f)$ et $p_M(m)$ sont respectivement les probabilités de trouver les valeurs f et m dans l'image fixe et l'image mobile. Cette mesure de similarité, bien qu'habituellement utilisée pour des recalages entre images de modalités différentes [346], a l'avantage de fonctionner aussi pour les recalages monomodales. Pour son application, le nombre de bins de l'histogramme a été fixé à 32. L'optimisation a été effectuée par une descente de gradient stochastique et une approche pyramidale avec 4 niveaux a été employée.

Toutes les opérations de recalage ont été effectuées grâce aux outils elastix [347, 348] sur les images TDM associées aux images TEP. Le recalage n'a pas été effectué directement sur les images TEP car ces dernières ne présentent pas les mêmes caractéristiques entre l'image précoce et l'image tardive (la fixation du traceur ayant évolué).

- Délimitation du volume métabolique

Une fois le VOI tumoral recalé sur l'image TEP tardive, une dilatation de ce dernier est effectuée afin de garantir l'inclusion de la totalité de la tumeur et remédier aux éventuels erreurs de recalage. La valeur du SUV_{\max} est alors récupérée dans cette région, l'isocontour correspondant à 70% du SUV_{\max} est tracé et la valeur d'intensité moyenne dans cette région est sauvegardée ($m\text{SUV}_{70}$). Sur le VOI recalé correspondant aux tissus sains, la moyenne des intensités est

calculée (bkg). Une fois les valeurs de $mSUV_{70}$ et d'activité dans le fond récupérées, le seuillage adaptatif de Schaefer *et al.* permettant la segmentation du volume métabolique a été appliqué.

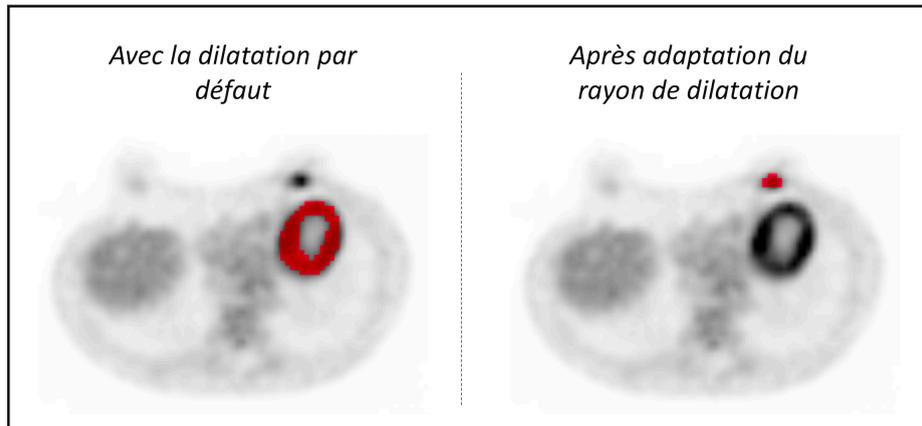


FIGURE 2.10 – Illustration d’une tumeur proche du coeur chez une patiente de 67 ans atteinte d’un carcinome canalaire luminal du sein gauche. À gauche, un exemple où une partie du coeur de la patiente a été incluse durant la phase de dilatation, occasionnant un VOI finale (rouge) non représentative de la tumeur. À droite, le VOI corrigé (rouge) de la tumeur, segmenté après diminution du rayon de dilatation.

Toutefois, durant l’étape de dilatation, pour certaines tumeurs situées proches de la paroi thoracique, la région d’intérêt dilatée incluait une partie du coeur. Dans ces cas-là, la valeur maximale, correspondant normalement au SUV_{max} de la tumeur, s’avérait être fausse et surestimée, comme illustré sur la figure 2.10. Ces cas ont pu être isolés et corrigés en diminuant le facteur de dilatation jusqu’à ce que la région d’intérêt n’inclue plus aucune partie du coeur, tout en veillant néanmoins à ce que l’ensemble du volume tumoral soit compris. En ce qui concerne les VOIs correspondant aux tissus sains, dans la quasi-totalité des cas, le recalage rigide a permis d’obtenir un VOI dans le sein controlatéral sur la TEP tardive sans transformation additionnelle. Cependant, dans certains cas, le VOI recalé s’est retrouvé soit à la frontière, soit en dehors du sein controlatéral. Pour ces cas-là, une nouvelle région d’intérêt a été tracée manuellement via le logiciel ITKSnap [349]

Enfin, l’une des limites de l’algorithme de segmentation automatique de Schaefer *et al.*, qui est une limite majeure des méthodes de seuillages fixes et adaptatifs, concerne son application aux tumeurs hétérogènes. Typiquement, dans le cas de tumeurs nécrosées, le volume tumoral segmenté ne va pas inclure ces régions de nécrose. Or, ces dernières ont été intégrées dans les délimitations manuelles de la perfusion tumorale. Dans le cadre de cette thèse, nous avons donc fait le choix d’incorporer ces régions nécrosées ou peu fixantes dans le volume métabolique afin de pouvoir étudier en parallèle la perfusion tumorale et le métabolisme glucidique de la tumeur, sur un volume tumoral équivalent. Par ailleurs, l’étude de l’hétérogénéité tumorale, qui fait également partie des travaux de cette thèse, a été un argument supplémentaire à l’incorporation de ces régions dans le volume tumoral métabolique. Ces régions non incluses par la segmentation automatique ont donc été identifiées et corrigée manuellement via ITKSnap (fig. 2.11). Seules

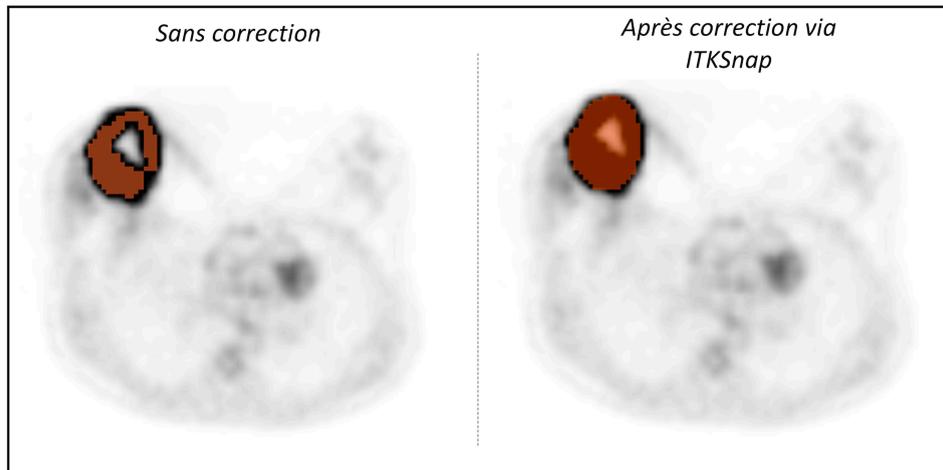


FIGURE 2.11 – Exemple VOIs corrigée chez une patiente de 45 ans atteinte d’un carcinome canalaire Triple Négatif au sein droit.

les tumeurs montrant des zones non incluses complètement englobées dans le VOI issu de la segmentation automatique ont été corrigées (31 patientes au total).

Une étape de vérification finale a été effectuée par la visualisation de l’intégralité des images TEP tardives et des régions d’intérêts tumorales et de bruit de fond associées.

EN RÉSUMÉ Ce travail a permis de construire l’image paramétrique nécessaire à l’évaluation de la perfusion tumorale. En parallèle, une étude sur la quantification du biais dans les images de faible statistique et sur l’amélioration de leur qualité a montré des résultats prometteurs pour une future application sur les images dynamiques de la base de données (Annexe C). Enfin, la segmentation du volume tumoral sur les images TEP tardives a permis de générer la région d’intérêt pour l’évaluation du métabolisme glucidique. À l’issu de toutes ces étapes, une base de données de 246 patientes contenant les images cliniques, les volumes d’intérêts et les images paramétriques de perfusion, a été créée.

ANALYSE DE L'HÉTÉROGÉNÉITÉ MÉTABOLIQUE ET PERFUSIONNELLE DES TUMEURS MAMMAIRES

Sommaire

3.1 Hétérogénéité tumorale : origine et motivations	68
3.1.1 Hétérogénéité tumorale et oncogenèse	68
3.1.2 Caractérisation de l'hétérogénéité tumorale	69
3.2 Analyse de texture	70
3.2.1 Principe et application de l'analyse de texture en imagerie médicale .	70
3.2.2 Matrices et paramètres de texture	73
3.2.3 Analyse de texture en TEP : applications et recommandations	78
3.3 Etude comparative des méthodes de discrétisation	86
3.3.1 Rationnel et objectifs	86
3.3.2 Matériel et méthodes	87
3.3.3 Résultats	90
3.3.4 Discussion et conclusion de l'étude	103
3.4 Relations entre l'hétérogénéité tumorale et les paramètres biologiques et moléculaires des tumeurs	107
3.4.1 Motivations et objectifs	107
3.4.2 Article	108

En se basant sur la base de données générée, l'objectif de ce troisième chapitre est de caractériser les tumeurs mammaires via leur analyse perfusionnelle et métabolique. Pour ce faire, une étude sur l'hétérogénéité des tumeurs, évaluée à la fois sur les images métaboliques et sur les images de perfusion a été effectuée via une analyse de texture. Deux analyses ont été réalisées, une première dont l'objectif était d'évaluer l'impact d'une méthode de calcul des paramètres de texture sur la classification des patientes bonnes et mauvaises répondeuses, et une seconde centrée sur la caractérisation des tumeurs mammaires en fonction des relations observées entre les indices d'hétérogénéité tumorale et les paramètres biologiques et moléculaires des tumeurs. Ce troisième chapitre se découpe en quatre parties : une première introduisant l'hétérogénéité tumorale et l'intérêt clinique potentiel qui en découlent, une seconde partie décrivant le principe

de l'analyse de texture, une troisième partie portant sur l'étude de la méthodologie de calcul des paramètres de texture et enfin une dernière partie traitant sur l'hétérogénéité tumorale perfusionnelle et métabolique et ses relations avec les paramètres biologiques et moléculaires des tumeurs.

3.1/ HÉTÉROGÉNÉITÉ TUMORALE : ORIGINE ET MOTIVATIONS

3.1.1/ HÉTÉROGÉNÉITÉ TUMORALE ET ONCOGÈNESE

Dans leur revue de 2011, Hanahan et Weinberg [350] rappellent les 10 principaux mécanismes biologiques développées par les cellules tumorales au cours de leur évolution :

- L'indépendance vis à vis des signaux de prolifération
- L'insensibilité aux signaux supresseurs de croissance
- Le développement d'une résistance à la mort cellulaire (apoptose)
- La capacité à induire l'angiogénèse
- La capacité d'invasion et de dissémination métastatique
- Un potentiel de réplification illimité
- La capacité d'échapper au système immunitaire
- La reprogrammation du métabolisme énergétique
- Une instabilité génomique et l'apparition de mutations
- Une inflammation favorisant le processus tumoral

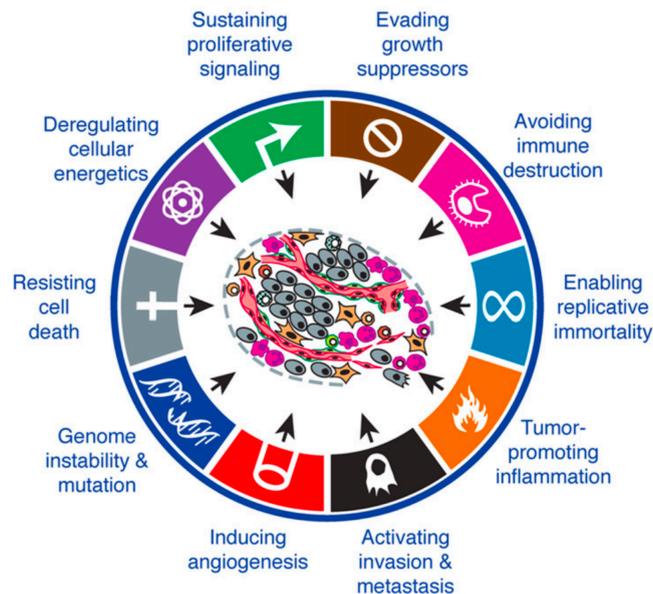


FIGURE 3.1 – Résumé des 10 grandes caractéristiques des tissus tumoraux [350]

Durant l'oncogénèse, les processus de mutations et de sélections donnent naissance à des populations de cellules cancéreuses distinctes au sein d'une même tumeur. En effet, les cellules

peuvent présenter des caryotypes différents [351], des expressions génétiques différentes [352] ou encore des différences au sein du génome [353] occasionnant ainsi une hétérogénéité dans la tumeur. Différents types d'hétérogénéité tumorale peuvent être identifiés [354] :

- l'hétérogénéité intra-tumorale issue des différences cellulaires au sein d'une même tumeur ;
- l'hétérogénéité intra-métastases issue des différences cellulaires au sein d'une même métastase ;
- l'hétérogénéité inter-métastase, qui correspond à des différences observées entre plusieurs métastases d'un même patient [355] ;
- l'hétérogénéité inter-patient, qui représente les différences tumorales entre patients.

L'hétérogénéité intra-tumorale regroupe rien qu'à elle seule plusieurs types d'hétérogénéités qui peuvent être d'ordre cellulaire, mais également tissulaire ou macroscopique. En effet, de par l'évolution des cellules cancéreuses et l'environnement dans lequel elles se trouvent, des hétérogénéités tissulaires régionales peuvent apparaître dans l'architecture stromale, au niveau de l'angiogenèse, dans la consommation en oxygène ou encore dans le métabolisme glucidique [356]. En effet, la néovascularisation des tumeurs faiblement régulée et la prolifération intensive de leurs cellules conduit à la formation de zones au sein des tumeurs plus ou moins éloignées du réseau vasculaire. L'apport d'oxygène et de nutriments se fait alors de manière hétérogène. La baisse de disponibilité en oxygène, i.e. l'hypoxie, ainsi qu'en nutriments, comme le glucose, entraînent également une augmentation de l'activité glycolytique des cellules dans ces zones [357]. Ces régions d'hypoxie faiblement perfusées peuvent conduire à des hétérogénéités macroscopiques comme des nécroses si l'apport vasculaire devient absent au fur et à mesure de la croissance tumorale.

Ces hétérogénéités présentent donc un intérêt clinique car elles sont le reflet des aspects du développement des tumeurs.

3.1.2/ CARACTÉRISATION DE L'HÉTÉROGÉNÉITÉ TUMORALE

L'évaluation de l'hétérogénéité intra-tumorale, considérée comme un des reflets de la nature et de l'évolution du cancer, représente un défi en oncologie [358] dans la mesure où elle pourrait être prise en compte dans la stratégie thérapeutique à des fins de médecine personnalisée. La caractérisation de cette hétérogénéité peut être réalisée à l'aide de biopsies. Ces dernières permettent l'analyse des tissus à l'échelle cellulaire et offrent donc une excellente résolution spatiale. Toutefois, cette approche reste un acte invasif et limité aux points de prélèvement. De nombreuses équipes se sont alors intéressées à l'analyse de texture des images médicales, notamment des images TEP, pour la quantification de l'hétérogénéité tumorale. Cette dernière, visible sur l'imagerie, pourrait être en partie le reflet des hétérogénéités sur des échelles inférieures. L'examen d'imagerie offre, certes, une moins bonne résolution spatiale que la biopsie, mais permet l'étude de l'ensemble de la lésion et ce, de manière non-invasive.

3.2/ ANALYSE DE TEXTURE

3.2.1/ PRINCIPE ET APPLICATION DE L'ANALYSE DE TEXTURE EN IMAGERIE MÉDICALE

3.2.1.1/ INTRODUCTION À L'ANALYSE DE TEXTURE

L'analyse de texture (AT) est un procédé qui est utilisé pour extraire des informations quantitatives, représentatives de ce que l'on nomme la texture d'une image, en renseignant sur l'arrangement spatial des intensités de l'image étudiée. La texture d'une image va ainsi être caractérisée par des indices de texture reflétant par exemple l'homogénéité (ou l'entropie) des intensités, la régularité (ou l'irrégularité) de certains modèles structurels des intensités de l'image, etc. L'avantage de l'analyse de texture est qu'elle est à même d'identifier aussi bien des motifs de texture macroscopiques, reconnaissables et visibles à l'oeil nu, comme le sont par exemple les zones de larges nécroses des tumeurs, mais également des caractéristiques de textures plus fines, moins (ou non) accessibles à une expertise visuelle.

L'application de l'analyse de texture aux images médicales, toutes modalités confondues, fait partie de la radiomique (*radiomic* en anglais), dont le terme a été introduit en 2010 par Gillies *et al.* puis repris en 2012 par Lambin *et al.* [359,360]. La radiomique est définie comme l'extraction d'un grand nombre de paramètres quantitatifs en imagerie (intensité, forme, texture...), dont on va étudier la complémentarité et la corrélation avec d'autres caractéristiques biologiques, comme l'activité génomique des tumeurs [361,362], leur classification par stade, histologie ou encore leur profil moléculaire [363–366]. La radiomique a également donné des résultats prometteurs dans le monitoring des thérapies en oncologie [367,368] notamment dans la prédiction de la réponse à la CNA dans le cas du cancer du sein (part. 1.1.4.3).

L'analyse de texture des images médicales, utilisée en parallèle ou complément de biomarqueurs conventionnels cliniques et/ou biologiques, est actuellement un sujet de recherche majeur pour le développement de la médecine personnalisée.

3.2.1.2/ PROTOCOLE D'APPLICATION DE L'ANALYSE DE TEXTURE AUX IMAGES MÉDICALES

De l'intérêt croissant dans le milieu médical pour la radiomique est né en 2016 et régulièrement mis à jour depuis (la dernière version, v11, datant de 2019), une initiative de normalisation des biomarqueurs en imagerie nommée IBSI (Image Biomarker Standardisation Initiative) [369]. L'IBSI est une collaboration internationale indépendante qui vise à normaliser l'extraction des indices radiomiques des images médicales. Cette initiative est née du manque de reproductibilité et de validation des études radiomiques de plus en plus nombreuses. L'objectif de l'IBSI est donc de fournir une nomenclature et des définitions pour la radiomique en général, de donner des directives d'application ou encore de référencer des outils de vérification des calculs des indices de texture. Ainsi, avant de pouvoir extraire les indices de texture des images médicales, tout un processus de préparation des données doit être effectué. Les principales étapes de ce processus

tel qu'il est présenté dans IBSI sont les suivantes :

A. La conversion des intensités des données brutes

Certaines modalités d'imagerie nécessitent une conversion des données brutes en valeurs plus représentatives cliniquement. C'est le cas en TEP lors de la conversion des images en valeurs de SUV, ou encore en valeur de perfusion comme expliqué dans le chapitre précédent. Cette conversion des intensités doit être effectuée avant le calcul des paramètres de texture sur l'image.

B. Les traitements post-acquisition

Le traitement des images post-acquisition a pour objectif d'améliorer la qualité de l'image reconstruite. Cette étape n'est pas systématique et va dépendre de la modalité d'imagerie utilisée. On peut citer par exemple la correction d'artefacts pour les acquisitions TDM [370], les étapes de débruitage et de correction de non-uniformité des intensités en résonance magnétique [371–374] ou encore la correction de l'EVP en TEP [375,376].

C. La segmentation du volume d'intérêt

L'étape suivante consiste à définir la zone sur laquelle les indices de texture vont être calculés. Les méthodes de segmentation dépendent des propriétés des images, elles vont donc varier selon les modalités d'imagerie utilisées. Pour l'imagerie TEP, une description générale des méthodes de segmentation a été faite dans le chapitre précédent, partie 2.4. L'influence de la méthode de segmentation sur les valeurs des indices de texture sera discutée plus en détail dans la partie 3.2.3.2 de ce même chapitre.

D. L'interpolation

Le calcul des indices de texture doit être invariant par rotation, afin de permettre la comparaison entre différents échantillons. Pour cela, la discrétisation de l'image doit être faite avec des voxels isotropes. Si ces derniers sont fréquents en imagerie TEP, ce n'est néanmoins pas le cas en imagerie TDM ou IRM pour lesquelles les voxels ont généralement une dimension axiale plus grande que les dimensions radiales et tangentielles. Dans le cas de voxels non-isotropes, un rééchantillonnage de l'image et donc une interpolation est alors nécessaire. L'étape d'interpolation est importante pour la cohérence des mesures, les indices de texture étant sensibles aux changements de taille des voxels [377–381]. Dans le cas des données TEP à notre disposition, les voxels étant isotropes, cette étape n'a pas été obligatoire. Plusieurs aspects, qui découlent de ce ré-échantillonnage spatial, comme le choix de l'algorithme d'interpolation ou encore la re-segmentation des VOIs ne seront donc pas abordés ici, mais sont décrits en détail dans [369].

E. La discrétisation des intensités

La discrétisation des intensités présentes dans le VOI est généralement nécessaire car, en regroupant les voxels de même intensité, elle va permettre de restreindre la taille de la matrice de texture à calculer et également limiter l'impact du bruit [369,382]. Cette discrétisation doit trouver un compromis entre la minimisation des valeurs possibles pour un voxel et la conservation des informations de l'image. Plusieurs méthodes ont été décrites

dans la littérature. Parmi celles-ci, deux méthodes ont été identifiées comme les plus communément employées : la méthode dite “absolue” ou à “largeur de bin fixe” (FBS pour *Fixed Bin Size*) et la méthode dite “relative” ou à “nombre de bin fixe” (FBN pour *Fixed Bin Number*).

- Discrétisation Relative (DR)

La DR consiste à rééchantillonner l’intensité des voxels entre les valeurs minimales et maximales présentes dans le VOI. Cette approche peut être décomposée en deux étapes : en premier on redimensionne les intensités minimales et maximales des voxels entre 0 et 1, puis on les rééchantillonne avec un nombre fixe (D) de valeurs discrètes [383,384]. La DR est associée dans la littérature à une discrétisation dite en “nombre de bin fixe” (FBN) car on discrétise l’image en un nombre fixe de valeurs d’intensités (16, 32, 64, 128, etc) [384,385]. L’équation permettant cette discrétisation est la suivante (où les valeurs d’intensité discrétisées sont arrondies à l’inférieur) :

$$DR(x) = \left\lfloor D \times \frac{I(x) - I_{min}}{I_{max} - I_{min}} \right\rfloor \quad (3.1)$$

avec $I(x)$ et $DR(x)$ les valeurs d’intensité dans un voxel x avant et après la discrétisation de l’image, I_{min} et I_{max} les intensités minimales et maximales, et D le nombre de bins du rééchantillonnage. Une légère variante, avec l’ajout d’un +1 au dénominateur de cette équation a également été utilisée dans la littérature [384,386]. Dans la suite de ce manuscrit, les deux approches seront notées DR.

- Discrétisation Absolue (DA)

La DA consiste à rééchantillonner l’intensité des voxels entre des bornes d’intensité fixées. La démarche est relativement similaire à celle de la DR : les intensités sont redimensionnées entre des valeurs fixées puis rééchantillonnées avec un nombre fixe (D) de valeurs discrètes [387]. L’équation correspondant à une discrétisation absolue est la suivante (où les valeurs d’intensité discrétisées sont arrondies à l’inférieur) :

$$DA(x) = \left\lfloor D \times \frac{I(x) - I_{boundary\ min}}{I_{boundary\ max} - I_{boundary\ min}} \right\rfloor \quad (3.2)$$

avec $I(x)$ et $AR(x)$ les valeurs d’intensité dans un voxel x avant et après la discrétisation de l’image, $I_{boundary\ min}$ et $I_{boundary\ max}$ les valeurs d’intensités minimales et maximales fixées et D le nombre de bin du rééchantillonnage. Les bornes d’intensités étant fixes d’une image à l’autre, la composante $(I_{boundary\ max} - I_{boundary\ min})/D$ est donc constante. Cette dernière correspond à la largeur de bin utilisée pour la discrétisation. Cette valeur ne variant pas d’une image à l’autre, la DA est associée dans la littérature à une méthode de “largeur de bin fixe” (FBS). Une légère variante de discrétisation absolue a également été proposé par Leijennar *et al.* [388]. Les deux méthodes seront notées DA dans la suite du manuscrit. La DR a comme inconvénient de ne pas conserver d’une tumeur à l’autre l’échelle des intensités respectives, dans la mesure où toutes les valeurs des voxels vont être redimensionnées entre les valeurs minimales et maximales de chaque image, comme illustré figure 3.2. Toutefois, cette méthode de discrétisation a l’avantage

de permettre une comparaison directe des valeurs des caractéristiques de texture à travers différentes VOI [369]. La DA, quant à elle, a l'avantage de conserver l'échelle des intensités entre les images (fig. 3.2). Le choix optimal de la méthode de discrétisation va dépendre de plusieurs facteurs, dont le type d'imagerie, à savoir s'il s'agit d'une imagerie dont l'unité est clairement définie (comme la TEP (SUV) ou TDM (Unité Hounsfield)), ou arbitraire (comme l'IRM) [369]. L'impact du choix de la méthode de discrétisation en imagerie TEP sur les valeurs des indices de texture sera discuté plus en détail dans la partie 3.2.3.3 de ce même chapitre et feront l'objet de la partie 3. Les discrétisations absolues et relatives sont les deux méthodes plus couramment employées dans la littérature, néanmoins d'autres approches moins courantes ont également été utilisées, comme l'égalisation d'histogramme [389] ou l'algorithme de clustering de Max-Lloyd [390].

Une fois les intensités discrétisées, les calculs des matrices de textures peuvent être effectués. Ces matrices ainsi que les différents indices de texture qui en sont extraits sont décrits dans la partie suivante.

3.2.2/ MATRICES ET PARAMÈTRES DE TEXTURE

La radiomique englobe un grand nombre de paramètres. Avant de définir les différentes matrices de texture, une brève présentation des indices d'intensité et morphologiques, est réalisée. Par la suite est détaillée la présentation des quatre principales matrices de texture. Pour faciliter la comparaison avec la littérature, les noms des matrices et des indices présentés seront écrits en anglais.

3.2.2.1/ INDICES D'INTENSITÉ

Les indices d'intensité vont décrire la distribution globale des intensités dans le VOI. Ils ne nécessitent pas l'étape de discrétisation de l'image pour être mesurés. Parmi les indices conventionnels utilisés on peut citer la valeur maximale, la valeur minimale, la moyenne ou encore l'écart type des intensités des images. D'autres paramètres vont quant à eux décrire l'allure de l'histogramme de l'image, comme son asymétrie (*skewness*) ou son aplatissement (*kurtosis*) comme illustré figure 3.3. Une liste exhaustive de tous ces indices peut être consultée dans [369]. Les indices d'intensité et de premier ordre reflètent une information globale qui ne renseignent pas sur l'arrangement spatial des intensités dans l'image. Différents arrangements d'intensité peuvent avoir des valeurs d'indices identiques si l'histogramme de l'image reste inchangé [391].

3.2.2.2/ INDICES MORPHOLOGIQUES

La géométrie des lésions peut s'avérer être représentative de leur statut bénin/malin, comme dans le cancer du sein où le caractère spiculé des lésions est généralement corrélée au caractère invasif et donc à un statut malin de stade avancé contrairement aux lésions aux bords bien définis correspondant plus régulièrement à des tumeurs bénignes ou moins agressives [392–394]. Des indices morphologiques renseignant sur la forme des tumeurs peuvent donc

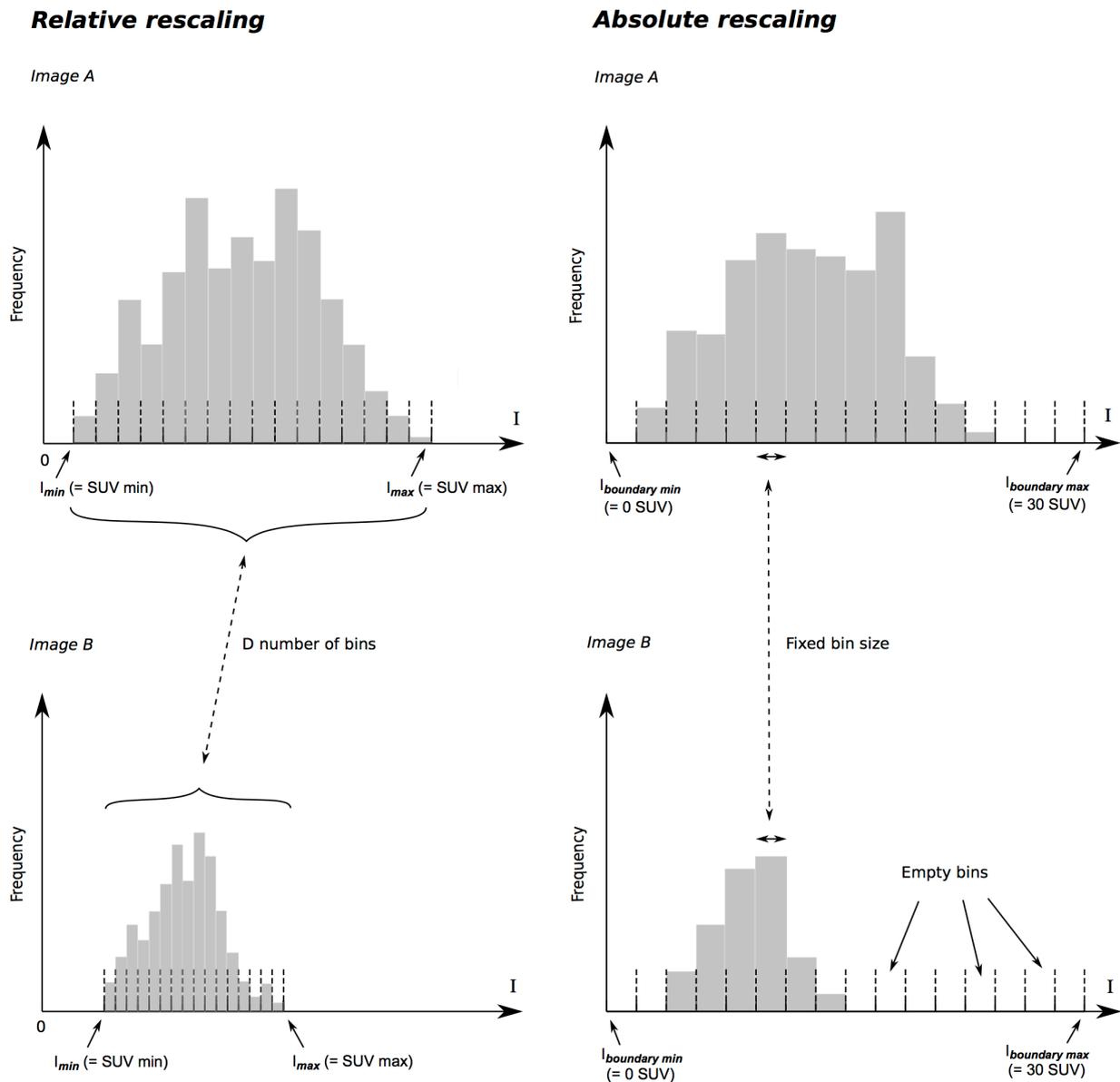


FIGURE 3.2 – Histogrammes simulés de deux images (A et B) sur lesquelles sont appliquées une DR (à gauche) et une DA (à droite) avec D égale à 16 bins. Avec une DR, les valeurs d'intensité sont normalisées pour chaque VOI avec un nombre fixe de bins, ce qui induit des largeurs de bins différentes entre les images. Avec une DA, le choix de bornes fixes permet de conserver la même échelle de SUV dans les images avec une largeur de bin fixe et un nombre de niveau d'intensité pour le rééchantillonnage fixe qui conduit à des bins «vides» comme illustré dans l'image B.

être calculés et utilisés en complément des indices d'intensité, de premier ordre et de texture. Parmi ces paramètres, on peut citer la sphéricité, la compacité, l'irrégularité, le plus large diamètre ou encore le ratio volume/surface

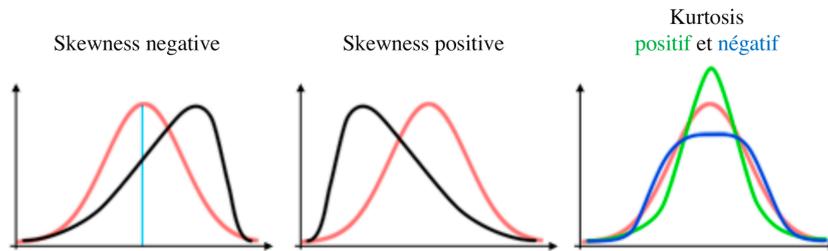


FIGURE 3.3 – Illustrations des mesures de skewness et kurtosis, qui permettent d'évaluer l'écart de l'histogramme de l'image par rapport à une distribution Gaussienne.

3.2.2.3/ MATRICES DE TEXTURE

Les indices d'ordre 2 et supérieurs vont permettre l'analyse de l'hétérogénéité intratumorale. Ces derniers sont extraits de matrices de textures qui caractérisent les relations existantes entre les voxels d'un même volume d'intérêt. Les indices de texture vont ainsi permettre de renseigner sur l'arrangement spatial des voxels dans l'image. On distingue deux types d'approches : l'analyse locale, qui va renseigner sur la relation d'un voxel avec son/ses voisins, et l'analyse régionale, qui va renseigner sur les relations entre des groupes de voxels de même intensité. Les quatre matrices principalement utilisées sont :

- la matrice de co-occurrence (GLCM : *Gray Level Co-occurrence Matrix*) et la matrice de différence des niveaux de gris (NGDLM : *Neighborhood Gray-Level Difference Matrix*) pour une étude locale de la texture de l'image ;
- la matrice de longueur des séries homogènes (GLRLM : *Gray Level Run-Length Matrix*) et la matrice des tailles de zones homogènes (GLSZM : *Gray Level Size-Zone Matrix*) pour une étude régionale de la texture de l'image.
- la matrice de co-occurrence (GLCM)

La GLCM, proposée par Haralick *et al.* dans les années 70 [395] va renseigner sur la probabilité d'observer un couple d'intensités situés à une distance d donnée l'une de l'autre et dans une direction de l'espace $\theta = (u, v)$ fixée. Les indices extraits de la GLCM sont des paramètres de second ordre car ils vont renseigner sur la relation entre une paire de voxels. La GLCM est une matrice de taille $D \times D$ avec D le nombre de niveaux d'intensités présentes dans l'image après discrétisation. Chaque élément (i, j) de la GLCM correspond à la probabilité d'avoir deux voxels d'intensité i et j , séparés par un vecteur (d, θ) .

En pratique, le remplissage de la matrice va donc dépendre de deux paramètres : la distance d correspond à l'espacement entre les deux voxels étudiés et la direction θ d'observation. Généralement la distance d est choisie égale à 1, de manière à étudier les relations entre deux voxels voisins. Pour une direction θ donnée, les paires d'intensités voxels sont répertoriées. La matrice est ensuite généralement rendue symétrique afin de prendre en compte la direction $-\theta$ puis normalisée afin de retrouver des probabilités d'observation de couples d'intensités [395]. Un exemple de calcul de GLCM est illustré figure 3.4.

Pour les directions d'observation, chaque voxel ayant 26 connexités avec ses voisins pour

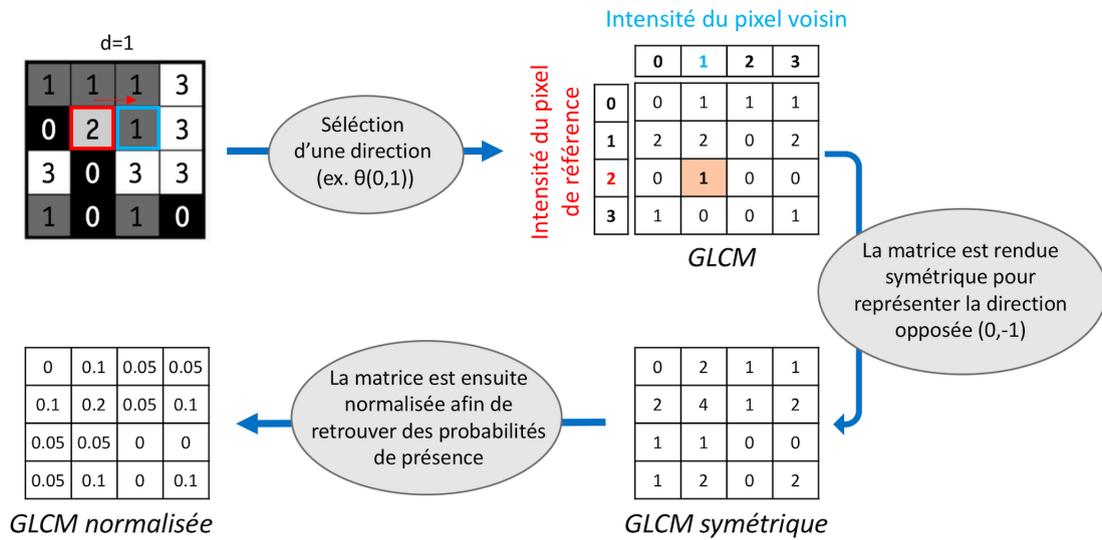


FIGURE 3.4 – Illustration du calcul de la matrice GLCM

une image 3D (resp. 8 connexités en 2D), 13 directions (resp. 4 en 2D) sont nécessaires pour couvrir la totalité de l'espace, comme l'illustre la figure 3.5.

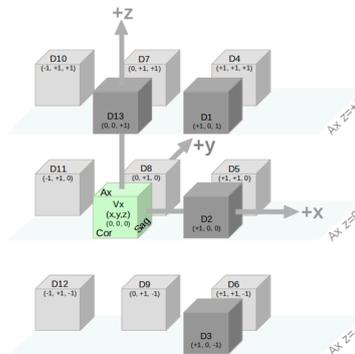


FIGURE 3.5 – Illustration des 13 directions possibles pour le calcul des matrices, [396]

Deux approches ont été présentées dans la littérature pour le calcul des indices de texture [397] pour des images 3D : soit 13 GLCM sont calculées pour chacune des 13 directions, soit une unique GLCM est calculée en sommant puis en normalisant les probabilités calculées sur les 13 directions. La première option est la plus répandue. L'indice de texture choisi est calculé pour chacune des 13 directions. Les 13 valeurs sont ensuite moyennées pour obtenir la valeur finale du paramètre de texture sur le VOI étudié. Dans le cas de la seconde option, une unique GLCM est calculée et donc un unique indice est extrait. Cette méthode est actuellement celle recommandée par l'IBSI. Il existe également des approches en 2 dimensions [369] qui ne seront toutefois pas développées ici car elles sont peu utilisées en pratique. À partir de la matrice de co-occurrence, de nombreux indices de texture peuvent être calculés [395]. Les indices utilisés dans ce manuscrit sont présentés en annexe A, une liste exhaustive étant

disponible dans [369].

- la matrice de différence des niveaux de gris (NGDLM)

La NGDLM [398] va permettre d'extraire des indices qui vont renseigner sur la relation entre un voxel et ses voisins. La NGDLM renseigne sur la différence d'intensité d'un voxel et l'intensité moyenne de ses voisins situés à une distance d . Comme pour la GLCM, cette distance est généralement prise égale à 1. La NGDLM est donc une matrice de taille $D \times 1$, avec D le nombre de niveaux d'intensités présentes dans l'image après discrétisation. Les principaux indices calculés à partir de cette matrice sont présentés en détail dans [369]. Un exemple de calcul de NGDLM est illustré figure 3.6.

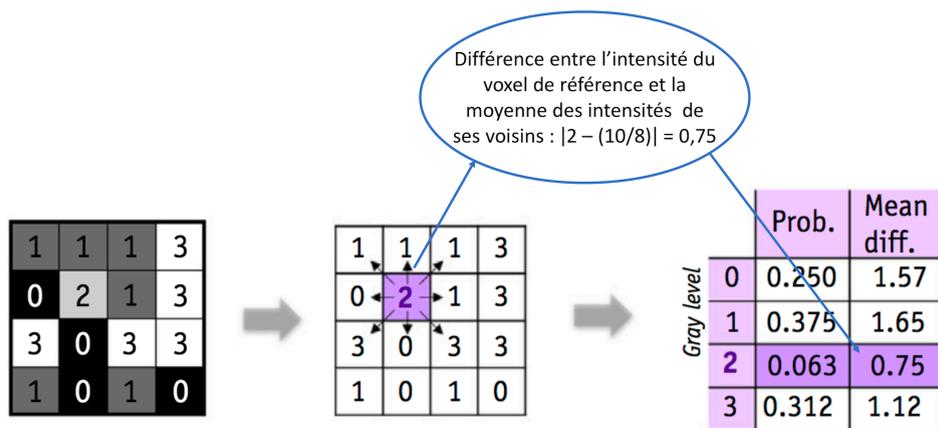


FIGURE 3.6 – Exemple de matrice NGLDM, adaptée de [399]. Sur cette illustration, les probabilités de présence des valeurs d'intensités sont ajoutées à la matrice.

- la matrice de longueur des séries homogènes (GLRLM)

La GLRLM [400] va renseigner sur la longueur des séries de voxels de même intensité dans l'image, suivant une direction θ donnée. Tout comme pour la GLCM, 13 directions sont possibles en 3D. La GLRLM va être de taille $D \times L$, avec D le nombre de niveaux d'intensités présentes dans l'image après discrétisation et L la taille maximale de longueurs de séries homogènes observées dans l'image. De la même manière que la GLCM, soit 13 matrices sont calculées pour chacune

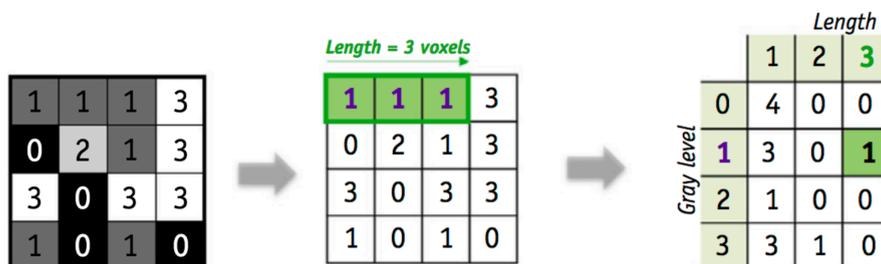


FIGURE 3.7 – Exemple de matrice GLRLM, adaptée de [399]

des directions θ , soit une unique matrice est calculée en normalisant la somme des probabilités calculées sur les 13 directions. Un exemple de calcul de GLRLM est illustré figure 3.7. Les principaux indices extraits des matrices GLRLM peuvent être consultés dans [369].

- la matrice des tailles de zones homogènes (GLSZM)

La GLSZM [401] va renseigner sur la dimension de zones (connexité) de même intensité dans l'image. Tout comme la GLRM, la GLSZM va être de taille $D \times S$, avec D le nombre de niveaux d'intensités présentes dans l'image après discrétisation et S la dimension de la zone homogène observée dans l'image. Un exemple de calcul de GLSZM est illustré figure 3.8. Les principaux indices extraits des matrices GLSZM peuvent également être consultés dans [369].

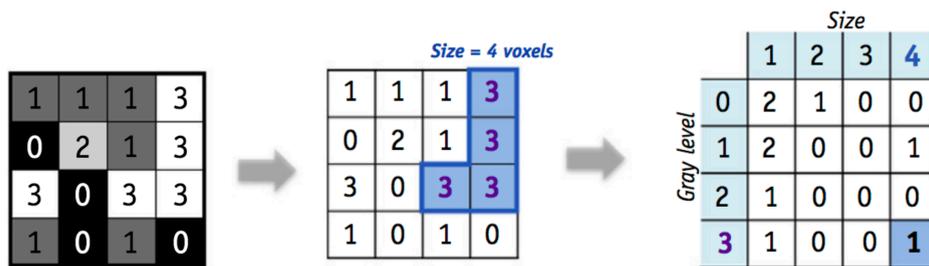


FIGURE 3.8 – Exemple de matrice GLSZM adaptée de [399]

3.2.3/ ANALYSE DE TEXTURE EN TEP : APPLICATIONS ET RECOMMANDATIONS

Depuis les dix dernières années, un nombre croissant d'études portant sur l'analyse de texture ont été publiées. La figure 3.9 rapporte le nombre de publications répertoriées entre 2010 et 2019, via une recherche pubmed effectuée le 01 mars 2020 avec les mots clés "texture analysis" ou "radiomic" ou "radiomics". Au total, 1076 publications sont renseignées pour l'année

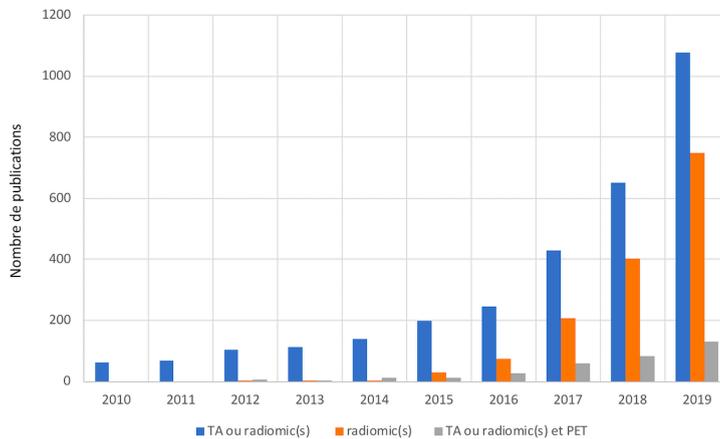


FIGURE 3.9 – Évolution du nombre d'articles répertoriés dans pubmed sur l'analyse de texture, de 2010 à 2019

2019, plus de deux fois le nombre d'études recensées deux ans auparavant (429 publications en 2017). Le terme "radiomic" étant plus récent, on observe les premières études utilisant ce terme à partir de 2012. Néanmoins, on peut noter qu'en ajoutant le mot clé "PET", les études de texture sur les images TEP représentent moins de 15% des études répertoriées en 2019 (152 publications).

Parmi ces articles sur la radiomique en TEP, les principales thématiques abordées sont :

- des études méthodologiques évaluant l'impact d'un ou plusieurs paramètres sur les valeurs des indices de texture
- des études méthodologiques proposant des méthodes d'analyse et d'application de l'analyse de texture
- des études cliniques évaluant les relations entre les indices de texture et des paramètres cliniques
- des études cliniques évaluant l'apport de l'analyse de texture pour répondre à des questions cliniques comme la classification de tumeurs, l'évaluation de la réponse à des traitements ou encore la prédiction de la survie
- des revues de la littérature

Les différents paramètres méthodologiques impactant les valeurs des indices de texture des images TEP sont présentés ci-après. Les choix qui en découlent pour le traitement des images de notre base de données y sont présentés.

3.2.3.1/ INFLUENCE DES PARAMÈTRES D'ACQUISITION, RECONSTRUCTION ET DE TRAITEMENT DES IMAGES

L'évaluation de la variabilité des indices de texture est essentielle pour une comparaison des résultats entre différentes études. Cette évaluation peut être effectuée par des études de répétabilité qui vont évaluer la variabilité d'un indice mesuré plusieurs fois dans les mêmes conditions, ainsi que par des études de reproductibilité qui vont quant à elles évaluer les variations de valeurs d'un indice lorsque ce dernier est mesuré dans des conditions d'observations différentes. De nombreux travaux se sont ainsi intéressés à l'impact d'un ou plusieurs paramètres sur les valeurs d'indices de texture en imagerie TEP. Parmi les paramètres testés, on peut citer l'échantillonnage, les paramètres de reconstruction, l'impact du type d'imageurs et des différents traitements d'images (filtrage, corrections d'EVP, de mouvement ou du bruit). Ces études ont été réalisées via des acquisitions sur fantôme et/ou sur des données patients.

3.2.3.1.1/ RÉPÉTABILITÉ DES INDICES (ÉTUDES TEST-RETEST)

Les études de répétabilité des indices ont pour objectif d'évaluer la variabilité de leurs valeurs entre deux mesures effectuées dans les mêmes conditions. La variabilité des mesures peut être évaluée via différentes métriques, dont l'une des principales est le calcul du coefficient de corrélation intra-classe (ICC *pour Intraclass Correlation Coefficient*). L'ICC est compris entre 0 et 1. Un ICC de 1 correspond à une répétabilité parfaite. Généralement, un indice est dit

répétable lorsque son ICC est supérieur à 0.8. Une autre approche utilisée dans la littérature est l'exploitation des graphiques de Bland-Altman où la différence moyenne entre les mesures est évaluée ainsi que l'intervalle comprenant 95% des mesures. Plusieurs publications ont effectué des études de répétabilité sur les indices de texture [402–405] sur des images cliniques de cancers du poumon (NSCLC pour Non-Small Cell Lung Cancer) ou de l'oesophage. Dans les études de Leijenaar *et al.* [402] et Tixier *et al.* [403], tous les indices issus de la matrice de co-occurrence apparaissent très reproductibles avec un $ICC \geq 80\%$. L'étude de Van Velden *et al.* [405] explore à la fois la répétabilité des indices et leur reproductibilité en fonction des paramètres de reconstruction et de la méthode de segmentation. Ils constatent que la majorité des indices (98%) basés sur l'intensité, la forme et la texture ont une répétabilité comparable à celle observée pour les mesures conventionnelles de SUV (SUV_{mean} , SUV_{max} , SUV_{peak} , etc.), avec des ICC pour la GLCM supérieurs à 80%, excepté pour l'homogeneity (EANM-reconstruction et DR) et le contrast (PSF-reconstruction et DR). Enfin, Desseroit *et al.* [404] évaluent la répétabilité à la fois sur les images TEP et sur les images TDM associées. Les indices considérés ont été évalués à l'aide de graphiques de Bland-Altman et la répétabilité des paramètres a été définie en fonction des moyennes et écart-types (SD pour *Standard Deviation*) des différences entre les mesures. Les résultats de cette étude montrent que les indices issus de la matrice de co-occurrence sont majoritairement fiables, avec seulement cinq métriques moyennement répétables ($16,7\% < SD < 22,2\%$) et trois peu répétables ($SD > 22,2\%$) sur les 23 étudiées au total.

3.2.3.1.2/ REPRODUCTIBILITÉ DES INDICES

La reproductibilité des paramètres est évaluée en mesurant les variations de valeurs des indices entre deux mesures dans des conditions d'observations différentes. Plusieurs études ont donc essayé de quantifier la reproductibilité des indices de texture en fonction de différents protocoles d'acquisition et reconstruction sur des images cliniques [378, 379, 381, 406, 407], sur fantômes [378, 408] ou sur image simulées [409]. Parmi ces travaux, certains testent différents algorithmes de reconstruction (OSEM, RAMLA, BLOB, TOF, PSF, etc) et/ou différents nombres d'itérations [378, 379, 381, 406–409], font varier la taille des matrices de reconstruction [378, 379, 381, 406], le mode (2D,3D) et le temps d'acquisition [378, 406, 408] ou appliquent différents filtrages [378, 379, 381, 408, 409]. Les résultats de ces études sont variables, notamment à cause des différents types d'imageurs et des méthodes de calcul des indices. Toutefois, des tendances se dégagent. On peut en effet noter que la taille de la matrice de reconstruction semble être l'une des plus grandes sources de variabilité et que l'indice Entropy calculé via la GLCM est l'un des paramètres le plus robuste.

L'impact de la correction (ou non-correction) des mouvements en TEP sur les valeurs des indices de texture a également été étudié sur des images cliniques de cancers pulmonaires [410–412]. Les résultats de l'étude de Olivier *et al.* [411] montrent que les mouvements respiratoires impactent les valeurs des indices. Seul 26% d'entre eux (56 indices au total) présentant des variations inférieures à 5% entre les images 3D et 4D. Néanmoins, l'étude à différentes phases d'une acquisition 4D a montré une bonne reproductibilité des indices au

cours du cycle respiratoire [410, 411]. Par la suite, Olivier *et al.* [413] ont étudié l'impact du bruit sur les valeurs des indices de texture. Dans leur étude, plusieurs bruits Gaussiens (de SD variables) ont été appliqués sur des images 3D et 4D de fantômes et de patients atteints de cancers du poumon. Les résultats soulignent l'impact majeur du bruit dans la plupart des indices étudiés. Les paramètres les plus sensibles au bruit sont ceux issus des GLSZM, mais il est également mentionné que le bruit dégrade les indices issus de la GLCM.

La diversité des imageurs et l'utilisation de protocoles d'acquisition et de reconstruction variant d'une étude à l'autre, les comparaisons entre travaux de différents Centres restent difficiles malgré les analyses de robustesse des indices. Pour remédier à cela, une harmonisation des valeurs des paramètres de TEP a été proposée par Orhac *et al.* [414]. Dans cette étude, les auteurs ont appliqué l'harmonisation ComBat [415], initialement décrite pour l'analyse génomique, sur des images TEP de tissus hépatiques sains et de lésions mammaires acquises dans deux départements différents et avec des résolutions spatiales significativement différentes. Après harmonisation, plus aucune des distributions des indices de texture testés n'était significativement différente entre les deux Centres. Cette méthode d'harmonisation est basée uniquement sur les données des patients acquises dans les différents instituts, sans nécessité de calibration/test sur fantômes, ce qui la rend appropriée pour l'analyse rétrospective des données. Par ailleurs, les auteurs soulignent le fait que cette harmonisation est directement applicable aux indices de texture calculés et non aux images TEP, il n'est donc pas nécessaire de réduire la qualité d'un ensemble d'images pour correspondre à la qualité inférieure d'un autre. Cette harmonisation s'applique aussi bien aux indices de texture qu'aux paramètres de SUV [414] et semble être un moyen prometteur de construire des modèles radiomiques à partir de données regroupées de différents Centres [416, 417]. L'utilisation de nouveaux patients après une harmonisation des paramètres reste néanmoins non triviale (la transformation pouvant mener à des valeurs de paramètres aberrantes) et la pluralité des appareils et paramètres d'acquisition/reconstruction utilisés pourrait rendre le nombre de labels à utiliser pour l'harmonisation limitant pour une estimation robuste.

3.2.3.2/ INFLUENCE DE LA MÉTHODE DE DÉLINÉATION

Comme il a été exposé au chapitre précédent (part. 2.4), il n'y a pour le moment pas de consensus sur une méthode optimale de délinéation des volumes tumoraux en TEP. Ainsi, dans le cas des études utilisant l'analyse de texture à des fins de classification ou pour la prédiction de la réponse au traitement, les modèles utilisés doivent être robustes vis à vis de la méthode de segmentation des volumes tumoraux afin qu'ils soient le plus reproductible possible.

Leijenaar *et al.* [402] ont dans un premier temps étudié la répétabilité des indices de texture extraits d'images TEP de cancers pulmonaires segmentés manuellement par cinq experts différents. Les auteurs ont observé une stabilité avec un ICC > 0.8 pour les indices issus des matrices GLCM et GLRLM. Orhac *et al.* [418] ont quant à eux comparé les valeurs des indices de texture calculés à partir de VOI segmentés soit avec un seuillage adaptatif, soit

avec un seuillage fixe (40% SUV_{max}) pour trois types de cancers différents (sein, pulmonaire et cancer colorectal métastatique). Seuls cinq paramètres de texture, dont l'entropie extraite de la GLCM, ont présenté une faible variabilité (avec une $SD_{indice} < SD_{SUV_{mean}}$) et 14 ont montré une variabilité modérée ($SD_{SUV_{mean}} < SD_{indice} < SD_{VM}$) dont l'*homogeneity* et la *dissimilarity* de la GLCM.

Hatt *et al.* [419] ont quant à eux étudié en parallèle l'impact de la segmentation et celui de la correction (ou non-correction) de l'effet de volume partiel sur les valeurs des indices de texture. Dans cette étude, les auteurs ont évalué les performances prédictives des indices en fonction des méthodes de segmentation dans l'évaluation de la réponse au traitement pour des cancers de l'œsophage. Les résultats de cette étude montrent que quatre indices de texture (dont l'*homogeneity* et la *dissimilarity* issus de la GLCM) sont statistiquement plus prédictifs lorsqu'une segmentation FLAB a été appliquée par rapport à une segmentation par seuillage fixe. Par ailleurs, excepté pour trois indices de texture régionaux (intensity variability, size-zone variability et high intensity emphasis), les paramètres de textures apparaissent peu sensibles vis à vis de la correction de l'EVP, avec uniquement de faibles différences ($<10\pm 15\%$) de notées entre les valeurs moyennes des indices extraits des images corrigées et non corrigées. Toutefois, les fortes corrélations entre la plupart des indices calculés et le volume tumoral sont à prendre en considération vis à vis d'un éventuel biais.

Cheng *et al.* [420] se sont ensuite intéressés aux capacités pronostiques d'indices de texture issues d'une analyse régionale (GLSZM) dans le cas de cancers ORL (OPSCC oropharyngeal squamous cell carcinoma) en utilisant deux méthodes de segmentation. Parmi tous les indices testés, le ZLNU a été identifié comme ayant une valeur pronostique quelle que soit la méthode de segmentation. Enfin, plus récemment, Altazi *et al.* [380] ont comparé l'impact d'une segmentation manuelle et d'une segmentation semi-automatique sur les valeurs des indices de texture mesurés pour des cancers du col de l'utérus. Sur 79 paramètres testés, 10 présentent des résultats de reproductibilité élevée.

Toutefois, afin de s'affranchir du choix de la méthode de segmentation, une méthode alternative a été proposée par Foley *et al.* [421]. Dans cette étude, les auteurs utilisent un algorithme basé sur les arbres décisionnels pour sélectionner la méthode de segmentation la plus appropriée [422] à l'étude radiomique de cancers de l'œsophage. L'avantage de l'utilisation d'un tel algorithme réside dans le choix automatique de la méthode la plus appropriée pour chaque individu parmi les différentes méthodes testées (seuillage adaptatif, agrégation, clustering ...).

3.2.3.3/ INFLUENCE DE LA MÉTHODE DE DISCRÉTISATION

L'influence de la discrétisation des images sur la robustesse des indices de texture a été étudiée dans plusieurs études et présente cependant des résultats contradictoires. En effet, dans le cas de lésions pulmonaires, une meilleure répétabilité des indices calculés avec une DA a été observée dans [405, 423] alors que pour cette même pathologie, les indices de texture mesurés après une DR ont été identifiés comme ceux ayant une plus faible variabilité dans une autre étude [404]. De la même manière, une étude sur fantôme a identifié les indices

de texture calculés avec un DR comme ayant une plus faible dépendance aux facteurs de reconstruction ainsi qu’au bruit présent dans les images [424] alors que d’autres études ont montré que majoritairement, les indices de texture calculés avec une DA s’avéraient être plus reproductibles que les indices calculés avec un DR vis à vis de la statistique des images TEP [425] et des paramètres d’acquisition et de reconstruction [408].

Outre l’influence sur la reproductibilité des indices, le choix de cette discrétisation ainsi que le choix des paramètres utilisés (nombre de bin, largeur de bin) ont également un impact direct sur les valeurs des indices de texture mesurés. Si l’on compare les méthodes absolue et relative précédemment présentées, plusieurs études ont montré une forte corrélation entre les indices de texture et le volume tumoral lorsque l’image est discrétisée de manière relative, et une corrélation entre les indices de texture et les valeurs SUV_{\max} lorsque celle-ci est discrétisée de manière absolue. L’information apportée par la texture est donc potentiellement redondante avec le volume tumoral ou avec les valeurs de SUV_{\max} en fonction de l’approche de discrétisation employée. Par ailleurs, la comparaison des indices de texture calculés avec une DA ou une DR a également souligné des différences majeures dans les distributions de ces derniers. En effet, les indices *entropy* et *homogeneity* issus de la GLCM varient généralement selon la même tendance lorsqu’ils sont calculés sur des images ayant été discrétisées de manière relative alors qu’ils présentent des variations de distribution opposées lorsque l’image a été discrétisée de manière absolue, ce qui paraît plus intuitif. De cette constatation s’en est suivie la question de l’interprétation des paramètres de texture. Une première étude a été réalisée par Tixier *et al.* [383] sur l’évaluation de la corrélation entre les valeurs des paramètres de texture mesurés après une DR et l’appréciation visuelle de l’hétérogénéité de tumeurs pulmonaires déterminée par des experts. Dans cette étude, l’hétérogénéité tumorale a été évaluée visuellement selon trois niveaux (1 : distribution homogène, 2 : hétérogénéité modérée, 3 : hétérogénéité élevée) et quantitativement par trois paramètres de texture (*homogeneity*, *entropy* et *dissimilarity*). Les résultats montrent des corrélations modérées ($0,4 < \rho < 0,6$) entre les paramètres de texture et les scores visuels. Toutefois, cette corrélation peut sembler contre-intuitive dans la mesure où le paramètre de texture *homogeneity* étudié augmente de concert avec l’hétérogénéité visuelle. Dans une autre étude, Orhac *et al.* [426] évaluent les corrélations entre les indices de texture sur des images TEP en pré-clinique et les mêmes indices calculés sur les images d’autoradiographie des coupes histologiques de ces mêmes tumeurs. Dans cette étude, les indices de texture sont calculés après une DA des images. Les résultats montrent des corrélations modérées à élevées ($0,57 < \rho < 0,85$) entre les indices mesurés sur les deux modalités. Dans une autre publication, afin d’approfondir l’interprétation des paramètres de texture, Orhac *et al.* [427] simulent des images de sphères comprenant des structures macroscopiques homogènes sur lesquelles plusieurs niveaux de bruits sont appliqués. Une comparaison directe entre les mêmes indices mesurés après deux discrétisations différentes d’une même image initiale plus ou moins bruitée est réalisée. Les résultats de cette étude, complétés par ceux de [428] semblent suggérer que les indices calculés après une DR ou une DA reflètent potentiellement différents types d’hétérogénéités. En effet, dans [428] l’auteur suggère que les sphères où aucune structure macroscopique homogène n’est présente mais où un bruit important occasionne une grande va-

riabilité de fixation, peuvent être vues comme ayant une “micro-hétérogénéité” en comparaison à la “macro-hétérogénéité” observée lorsque des structures, dont la taille est comparable à la taille de la sphère simulée, sont incluses. Les indices de texture mesurés après une DA montrent dans ces travaux une corrélation avec l’hétérogénéité macroscopique observée sur les images simulées [427] contrairement aux indices mesurés après une DR qui semblent potentiellement refléter davantage une “micro-hétérogénéité” [428]. Toutefois, dans [428] l’auteur met également en garde contre le potentiel biais induit par une DR qui pourrait “créer” artificiellement de la texture sur une image bruitée considérée cliniquement comme homogène. Enfin, plus récemment, une autre étude [423] a également comparé, comme Tixier *et al.* [383], les valeurs de plusieurs indices de texture par rapport à des scores visuels d’hétérogénéité tumorale évaluée par trois experts sur des lésions pulmonaires. Pour cette étude, les indices de texture ont été calculés après une DR, après une DA et après une variante de la DA qui utilise une taille de bin fixe comme présenté dans [388]. Les résultats de l’étude montrent une faible à moyenne corrélation entre les indices mesurés après une DA et l’appréciation visuelle ($0,14 < \rho < 0,37$) et une absence de corrélation entre cette dernière et les indices mesurés après une DR ($\rho < 0,12$).

Néanmoins, la naissance de la radiomique vient en grande partie du constat que les images cliniques sont sous-exploitées par l’évaluation visuelle. L’argument qui tend à privilégier la DA par rapport à la DR (qui permet une corrélation intuitive des indices avec l’hétérogénéité macroscopique) peut donc être discuté [429]. En effet, par définition, l’analyse de texture permet de quantifier aussi bien une texture visible à l’œil nu, ce qui est *a priori* le cas lorsqu’une DA est employée, qu’une texture peu ou non visible à l’œil nu. On peut donc dans ce cas concevoir l’utilisation d’une DR comme un complément d’information moins accessible de prime abord au clinicien.

L’information clinique apportée par l’analyse de texture et son interprétation vont donc être possiblement dépendantes de la méthode de discrétisation. Le choix de la méthode de calcul des indices de texture se pose alors. Actuellement, les recommandations de l’IBSI laissent une marge de manoeuvre dans le choix de la méthode de discrétisation des images TEP. En effet, bien qu’il soit notifié que l’utilisation d’une DA permette de conserver les échelles des valeurs des intensités discrétisées et soit donc pertinente dans l’analyse des images TEP, la DR reste parmi les choix applicables pour cette modalité (cf. [369] tableau 3.1, version de mars 2020). Par ailleurs, il est également mentionné que la méthode de calcul des indices de texture, peut être optimisée en fonction des objectifs cliniques visés. C’est d’ailleurs en partie le principe qui a été employé par les premières études qui ont évalué l’impact de la discrétisation sur la classification tumorale à partir des images cliniques en TEP. En effet, Leijenaar *et al.* [388] ont évalué, sur des images cliniques de cancers pulmonaires (NSCLC), la capacité de différents indices de texture discrétisés de manière absolue et relative à classer les patients. En conclusion de leurs études, les auteurs identifient les indices issus des images discrétisées de manière absolue comme plus appropriés pour la classification des patients dans ce contexte clinique. De la même manière, Orhac *et al.* [387] ont montré que l’utilisation de la méthode de DA conduit à une meilleure différenciation entre des lésions pulmonaires et des régions de tissus sains et

fournit également une meilleure classification entre différents types histologiques de cancer du poumon (NSCLC). Une troisième étude, menée par Shen *et al.* [430] dans le cas du cancer du col de l'utérus, a montré que les indices de texture calculés après une DR des images permettait d'obtenir une meilleure identification des cancers avec une atteinte métastatique ganglionnaire. Enfin, une dernière étude de Lucia *et al.* [431] a montré le caractère prédictif de paramètres de texture dans la prédiction de la survie sans récurrence des cancers du col de l'utérus, uniquement après une discrétisation des images faite par égalisation d'histogramme [432]. Ces études, dont les choix méthodologiques diffèrent selon les types de cancers étudiés, sont à notre connaissance les seules actuellement à avoir étudié l'impact de différentes méthodes de discrétisation sur la classification de cas cliniques.

Par ailleurs, au-delà du choix de l'approche de discrétisation, l'impact des valeurs choisies pour les paramètres utilisés dans chacune des méthodes a également été étudié [384, 385, 387]. Pour ce qui est de la DR, de nombreuses études ont testé l'influence du nombre de bin allant de 8 à 128. Au vu des résultats, afin de maximiser la reproductibilité des paramètres et afin de limiter la corrélation des indices de texture avec l'information du volume métabolique, un consensus a été trouvé pour un nombre de bin de 32 ou 64. On peut néanmoins noter que la majorité des études publiées à ce jour emploie 64 niveaux d'intensité pour la discrétisation des images TEP. Pour ce qui est de la DA, les limites utilisées vont dépendre de la plage des intensités présentes dans les images, et donc de la nature de ces dernières (SUV, perfusion). En règle générale, il est conseillé de prendre des valeurs qui englobent la totalité des intensités observées dans les images étudiées.

Enfin, la manière de calculer la GLCM [397] s'est avérée avoir un impact sur la corrélation entre les indices extraits de cette dernière et les volumes tumoraux. En effet, Hatt *et al.* [397] ont ainsi montré dans leur étude que le calcul d'une unique GLCM donnait des valeurs d'entropie et de dissimilarité moins corrélées au volume tumoral que lors de l'utilisation de 13 GLCM pour le calcul de ces indices. Cette analyse a été testée pour une DR, l'impact du calcul de la GLCM n'ayant pas été étudié à notre connaissance sur des indices de second ordre mesurés après une DA.

3.2.3.4/ RECOMMANDATIONS VIS À VIS DU VOLUME TUMORAL

En parallèle des études évaluant la corrélation entre les indices de texture et le volume tumoral, s'est posé la question d'un volume minimal requis pour l'application de l'analyse de texture en TEP. L'un des premiers papiers à avoir fait référence à un volume minimal pour une analyse fiable de l'hétérogénéité est celui de Brooks and Grigsby [433]. Dans leur étude, les auteurs suggèrent que l'indice *entropy* (GLCM), mesuré sur un volume tumoral inférieur à 45 cm³, refléterait davantage le volume de la lésion que son hétérogénéité. Ce résultat a été par la suite nuancé, les auteurs n'ayant testé qu'un seul indice et ayant procédé à une discrétisation des intensités de l'image en utilisant plus de 150 bins. Par la suite, Hatt *et al.* [397] ont montré que les corrélations des indices avec les volumes tumoraux peuvent être minimisées en adaptant

la méthode de calcul de la GLCM, abaissant la limite minimale de volume à 10 cm^3 . Plutôt que d'exclure les tumeurs de petites tailles, les auteurs suggèrent alors d'étudier les corrélations entre les indices d'hétérogénéité et les volumes tumoraux afin d'analyser le caractère complémentaire ou redondant de ces indices en fonction de la taille des lésions.

3.2.3.5/ CHOIX MÉTHODOLOGIQUES POUR LES ÉTUDES AVEC LES DONNÉES DE LA BASE TREN

Si l'on reprend les différentes étapes du protocole d'application de l'analyse de texture aux images médicales, les choix qui ont été pris pour notre base de données sont les suivants :

- La conversion des intensités des images a été effectuée de manière à d'obtenir une image paramétrique de perfusion tumorale ainsi qu'une image de SUV. Cette étape de conversion a été présentée et décrite dans le chapitre précédent.
- Aucun post-traitement n'a été appliqué aux images reconstruites. Toutefois, cette étape est décrite dans l'IBSI comme non obligatoire lorsqu'un processus de réduction de bruit est déjà implémenté pendant la reconstruction de l'image, ce qui peut être considéré comme le cas pour nos données de par le lissage effectué lors de la reconstruction des images (part. 1.2.2.1). Aucune correction de l'EVP n'a été appliquée en première intention, ce qui n'aura *a priori* que peu d'influence sur les valeurs des indices de texture calculés [419].
- Les segmentations des volumes d'intérêt ont été réalisées de manière manuelle sur les images de perfusion tumorale et à l'aide d'un seuillage adaptatif pour les images métaboliques. Ces choix reposent sur les arguments exposés au chapitre précédent (part. 2.4.2).
- Les indices de texture étant calculés uniquement sur les images TEP dont les voxels sont par défaut isotropes pour nos images, aucune interpolation a été nécessaire.
- Au vu des résultats exposés dans la littérature, il a été décidé d'optimiser le choix de la méthode de discrétisation en fonction des objectifs cliniques. Ce point fait l'objet de l'étude de la partie suivante de ce chapitre.
- Enfin, les lésions mammaires étudiées étant majoritairement de petite taille, il a été décidé d'analyser uniquement l'analyse locale de l'hétérogénéité tumorale. De ce fait, les indices de texture issus des GLCM, qui apparaissent dans l'ensemble relativement robustes vis à vis des divers paramètres méthodologiques étudiés dans la littérature, ont été utilisés dans les travaux de cette thèse.

3.3/ ETUDE COMPARATIVE DES MÉTHODES DE DISCRÉTISATION

3.3.1/ RATIONNEL ET OBJECTIFS

L'objectif principal de la thèse est d'étudier l'apport combiné du métabolisme et de la perfusion tumorale évaluée par la TEP au ^{18}F -FDG dans la prédiction de la réponse à la chimiothérapie néoadjuvante. Pour ce faire, nous avons dans un premier temps souhaité quantifier l'hétérogénéité tumorale dont le caractère prédictif, pour l'évaluation de la réponse à la CNA des lésions mammaires, a été mis en lumière par de nombreuses études (part. 1.1.4.3).

Toutefois, l'efficacité des indices de texture en tant que biomarqueurs des réponses thérapeutiques pourrait dépendre en partie du choix de la stratégie de discrétisation, celle-ci pouvant possiblement refléter différentes caractéristiques tumorales (part. 3.2.3.3). Le choix n'étant pas arrêté sur la méthode de discrétisation optimale des images TEP (part. 3.2.3.3), nous avons décidé, d'effectuer un travail préliminaire afin d'observer si une méthode de discrétisation apparaissait plus appropriée au contexte clinique étudié. En effet, à notre connaissance, seules les trois études précédemment citées [387, 388, 430] ont évalué l'impact de la méthode de discrétisation des images TEP pour la classification des tumeurs, et aucune n'a été menée dans le cas du cancer du sein. Par ailleurs, les images dynamiques présentent un rapport signal sur bruit bien plus faible que les images métaboliques ce qui peut potentiellement impacter les valeurs des indices de texture. Or, à notre connaissance, aucune étude n'a été réalisée sur les images de perfusion tumorale.

L'objectif de cette étude est donc de faire une analyse préliminaire permettant d'identifier la méthode de calcul des indices de texture qui se présenterait comme la plus appropriée pour identifier les patientes ayant eu une réponse histologique complète. Ce travail se découpe/présente en deux parties :

1. l'exploration de l'impact de la méthode de discrétisation sur la relation entre les indices de texture et les paramètres plus conventionnels (SUV, paramètres de perfusion, volumes tumoraux)
2. l'évaluation des différences de valeurs des indices entre les patientes pCR et non-pCR afin d'observer si ces différences sont plus ou moins marquées en fonction de la méthode employée pour les calculs des indices de texture, c'est à dire en fonction de la discrétisation (DA ou DR) et du mode de calcul des GLCM (1 unique matrice ou 13). Cette étude a fait l'objet d'un article qui a été soumis dans Plos One.

3.3.2/ MATÉRIEL ET MÉTHODES

3.3.2.1/ POPULATION ÉTUDIÉE ET PROTOCOLE D'ACQUISITION DES DONNÉES

De janvier 2011 à mai 2017, les patientes incluses dans le protocole TREN (cancers du sein nouvellement diagnostiqués, de stade II ou III et avec une indication de CNA) ont été recrutées de manière prospective sur les recommandations d'oncologues pour un examen TEP/TDM de ^{18}F -FDG avant le début de leur traitement néoadjuvant. La taille tumorale et l'atteinte ganglionnaire ont été évaluées par échographie avant le début du traitement. La classification TNM des tumeurs a été établie conformément à la 8e édition du American Joint Committee on Cancer (AJCC) [434]. Les biopsies des tumeurs primitives ont permis d'évaluer les marqueurs moléculaires (RP, ER, HER2, etc.) ainsi que les caractéristiques histologiques des tumeurs, et ont permis d'établir leurs grades selon le système de Scarff-Bloom-Richardson (SBR) [8]. Les tumeurs ont été considérées comme ER+ et/ou PR+ si au moins 10% des cellules étaient marquées positivement. La surexpression de HER2 a été estimée par immunohistochimie et classée sur une échelle allant de 0 à 3+. Les tumeurs présentant une surexpression de niveau 3+

ont été considérées comme HER2-positives. L'hybridation in situ par fluorescence (FISH) a été utilisée pour confirmer le statut HER2+ pour les cas ambigus (surexpression de niveau 2+). Les patientes ont été classées en trois groupes selon leurs sous-types moléculaires : Luminal (RH positifs, HER2 négatif), HER2 (HER2 positif) et triple négatifs (TN) (RH négatif, HER2 négatif).

Pour ce qui est du traitement néoadjuvant, ce dernier a débuté après la réalisation de la TEP1. Les patientes ont subi divers régimes de CNA :

- les femmes atteintes de tumeurs HER2+ ont reçu des schémas thérapeutiques à base de trastuzumab et de docétaxel,
- les femmes atteintes de tumeurs triple négatives ou luminales ont reçu, soit six cycles de chimiothérapie séquentielle avec des anthracyclines et des taxanes, soit six cycles de FEC100 (épirubicine 100 mg/m², 5-fluorouracile 500 mg/m² et cyclophosphamide 500 mg/m²).

La prise en charge médicale n'a pas été adaptée à la réponse métabolique précocement observée, excepté en cas de progression tumorale. Dans ce cas précis, la décision de poursuivre, de modifier ou d'interrompre le traitement néoadjuvant a été prise par l'oncologue responsable de la prise en charge la patiente. L'exérèse de la tumeur a par la suite été effectuée dans le mois suivant la fin de la CNA. En fonction de l'appréciation du chirurgien et de la taille du résidu tumoral, les patientes ont pu bénéficier soit d'une chirurgie conservatrice, soit d'une mastectomie. La réponse histologique a été évaluée sur l'examen anatomopathologique de la pièce opératoire, selon la classification de Chevallier. En ce qui concerne le protocole d'acquisition et de reconstruction des données, il s'agit de celui du protocole TREN qui a été décrit précédemment dans le chapitre 2. Il ne sera donc pas rappelé ici.

3.3.2.2/ DISCRÉTISATION DES IMAGES TEP

Les intensités des VOIs tumoraux des images paramétriques de perfusion et des images de métabolisme glucidique ont été discrétisés de manière relative (DR) et de manière absolue (DA) (part. 1.2.1.3.3) selon les équations 3.1 et 3.2 respectivement. Pour la DR, I_{min} et I_{max} correspondent aux intensités minimales et maximales observées dans le VOI. Pour ce qui est de la DA, les intensités minimales et maximales $I_{boundary\ min}$ et $I_{boundary\ max}$ doivent être déterminées en amont et rester fixes pour toutes les lésions étudiées. Au vu des valeurs de métabolisme et de perfusion tumorale observées dans notre cohorte de patients, les limites minimales et maximales ont été fixées à 0 et 30 SUV pour les images de métabolisme et à 0 et 1 mL/min/g pour les images de perfusion. Ces choix ont été faits de manière à englober, pour chaque paramètre biologique étudié, la totalité des valeurs observées dans la cohorte.

Pour l'analyse du métabolisme, les VOI des tumeurs ont été rééchantillonnées avec 64 niveaux d'intensités ($D=64$) comme suggéré par la littérature [384, 435, 436]. Pour ce qui est de l'analyse de la perfusion, aucune référence dans la littérature n'a été trouvée. Les tumeurs ont alors été rééchantillonnées en testant 16, 32 et 64 niveaux d'intensités ($D=16, 32, 64$) en

gardant à l'esprit que les images de perfusion présentent une plus faible statistique dans l'image que celles de métabolisme. Avec les limites choisies, ces valeurs D correspondent pour la DA à des largeurs de bin comprises entre 0,06 et 0,02 mL/min/g et entre 0.47 et 1.87 SUV environ.

3.3.2.3/ EXTRACTION DES PARAMÈTRES DES IMAGES

- Paramètres conventionnels

Des paramètres de premier ordre reflétant le métabolisme et la perfusion tumorale de manière globale ont été extraits des images paramétriques de SUV et de perfusion respectivement. Les paramètres qui ont été calculés sont les moyennes (SUV_{mean} et BF_{mean}), les écarts types (SUV_{sd} et BF_{sd}), les valeurs d'intensité maximales (SUV_{max} et BF_{max}), les volumes tumoraux (VM et VP) ainsi que le *Total Lesion Glycolysis* ($TLG = VM \times SUV_{\text{mean}}$) a été extrait des images métaboliques. Les paramètres skewness et kurtosis ont également été calculés à partir des images paramétriques de perfusion et de métabolisme.

- Paramètres de texture

Des paramètres de second ordre ont été extraits des images afin de décrire et quantifier localement l'arrangement spatial des voxels. Après discrétisation, neuf indices de texture ont été calculés à partir des matrices GLCM. Ces paramètres ont été calculés avec une distance d égale à 1 décrivant ainsi la relation entre deux voxels consécutifs. Les indices que nous avons sélectionnés pour cette étude sont :

- *entropy* et *energy* pour la quantification de l'uniformité et la régularité de la texture ;
- *homogeneity*, *Inverse Different Moment (IDM)*, *Inertia (Contrast)* pour évaluer l'homogénéité et le contraste des distributions d'intensité dans l'image ;
- *dissimilarity* et *correlation* pour quantifier les dépendances linéaires et les variations d'intensité entre paires de voxels ;
- *cluster shade* et *cluster prominence* pour renseigner la symétrie des matrices GLCM.

Les formules associées à tous ces indices sont disponibles en annexe A. Deux méthodes ont été utilisées pour le calcul des indices de texture : la première consiste à utiliser 13 GLCM, une pour chaque direction spatiale, tandis que la seconde consiste à ne calculer qu'une unique GLCM, qui est la somme des 13 matrices précédentes. Ces deux méthodes font parties des stratégies listées par l'IBSI (figures 13.e et 13.f de la version de l'IBSI de mars 2020). Après normalisation des GLCM, les indices de texture sont calculés pour chaque matrice et moyennés dans le cas de 13 GLCM uniquement, pour obtenir les valeurs finales des indices. Chaque paramètre de texture du métabolisme sera donc calculé au total quatre fois :

- avec une DR et 13 GLCM
- avec une DA et 13 GLCM
- avec une DR et 1 GLCM
- avec une DA et 1 GLCM

Les indices de texture de la perfusion tumorale seront quant à eux calculés pour chacune des

quatre configurations ci-dessus de trois manières différentes : une première fois en utilisant un rééchantillonnage avec 16 niveaux d'intensités, une seconde fois avec 32 niveaux d'intensités et une dernière avec 64 niveaux d'intensités, soit au total 12 valeurs pour un même indice.

3.3.2.4/ ANALYSE STATISTIQUE

L'analyse statistique a été réalisée avec RStudio (version 1.0.153) [437] et tous les indices TEP ont été extraits des images paramétriques à l'aide des outils déjà préexistants et d'autres qui ont été implémentés dans le logiciel vv [272]. Les indices de textures qui ont été intégrés ont été soumis à une étape de vérification en comparant les valeurs obtenues avec celle du logiciel LIFE_x [438], conforme aux recommandations IBSI. Une analyse descriptive a été réalisée afin d'évaluer s'il existe des différences de valeurs des indices de texture entre les tumeurs des patientes pCR et celles des patientes non-pCR. Cette analyse a également pour objectif d'observer si ces différences (si différences il y a) sont plus ou moins marquées en fonction de la méthode employée pour les calculs des indices (méthode de discrétisation (DA ou DR), calcul de la GLCM (1 unique matrice ou 13)). Ce travail a d'abord été réalisé sur l'ensemble de la population puis de manière plus ciblée sur les sous-groupes HER2 et TN. Cette analyse n'a toutefois pas été réalisée sur le sous-groupe Luminal car la réponse histologique complète comme marqueur substitutif de la survie est connue pour être peu pertinent pour ce phénotype tumoral (part. 1.1.3.2).

L'évaluation des différences de valeurs entre les indices des patientes pCR et non-pCR ont dans un premier temps été effectuées à l'aide de tests de Mann-Whitney bilatéraux dont le seuil de significativité a été fixé à 0,05. Une correction de Benjamini-Hochberg a ensuite été appliquée aux p-values. Le test de Mann-Whitney évaluant individuellement chaque variable, nous avons par la suite réalisé une analyse descriptive multivariée sur nos données. Cette analyse a été effectuée par une analyse factorielle discriminante (AFD) descriptive à l'aide du package Discriminer disponible sous R. Cette méthode vise à séparer deux classes (ou plus) en recherchant la combinaison linéaire des variables qui maximise la séparation entre les groupes. Les combinaisons linéaires de ces variables sont communément appelées variables discriminantes. Le rapport de corrélation η^2 ($0 \leq \eta^2 \leq 1$) a été utilisé comme indicateur du niveau de séparation des groupes. Un rapport de corrélation égal à 1 correspond à une discrimination parfaite entre les deux classes. Enfin, l'étude de la corrélation entre les indices de texture et les paramètres globaux a été effectuée à l'aide des coefficients de corrélation de Spearman (ρ).

3.3.3/ RÉSULTATS

3.3.3.1/ CARACTÉRISTIQUES DE LA COHORTE

Les paramètres de texture ont été calculés sur les 246 patientes disponibles dans la base de données. Toutefois, les indices de texture n'ont pas pu être calculés pour 29 des patientes pour lesquelles les tumeurs présentaient des formes à la fois très anisotropes et de très petit volume, réduisant ainsi la cohorte à 217 patientes. Parmi ces dernières, la connaissance de la réponse

histologique à la CNA n'a été renseignée que pour 207 d'entre elles. Par ailleurs, certaines patientes ayant des tumeurs dont les valeurs maximales de perfusion étaient très faibles, ont eu, après une DA, l'intégralité des valeurs comprises dans le VOI regroupées dans un unique bin. Pour ces tumeurs, le calcul de l'indice de corrélation nécessitait une division par zéro, rendant cet indice inaccessible. Afin d'avoir pour toutes les patientes les mêmes indices évalués, il a été décidé de ne pas sélectionner ces patientes pour ce travail. Même si avec une DR ce souci ne s'est pas présenté, le même nombre de patientes a été étudié, quelle que soit la méthode de discrétisation des images. Au final, l'étude a donc été menée sur 196 patientes, dont l'âge moyen est de 50 ans. Le VM médian est de 6,5 ml avec un intervalle interquartile de 8,3 ml et le volume médian perfusé est de 4,4 ml avec un intervalle interquartile de 7,3 ml. Bien qu'une grande proportion des volumes tumoraux soient de petite taille (<10 cc), toutes les lésions ont été analysées sans aucune limitation de taille afin de conserver un échantillon le plus représentatif possible des lésions du cancer du sein. Les carcinomes canauxiaux représentent 95,4% des patientes, 3,6% sont lobulaires et 1% ne sont pas classées. Soixante-treize (37,2%) patientes sont de phénotype HER2, 71 (36,2%) de phénotype luminal et 52 (26,6%) de phénotype TN. Parmi la population de patients, 47 (24%) ont atteint une réponse histologique complète contre 149 (76%) non-pCR. Les caractéristiques des patients et des tumeurs sont résumées dans le tableau 3.1.

Caractéristiques des patientes et des tumeurs	
	Nombre de patients (%) N = 196
Type histologique	
Canalaire	187 (95,4)
Lobulaire	7 (3,6)
Autre	2 (1)
Status ER	
Positive	84 (42,9)
Négative	112 (57,1)
Status PR	
Positive	111 (56,6)
Négative	85 (43,4)
Status HER2	
Positive	73 (37,2)
Négative	123 (62,8)
Phénotype tumoral	
ER+/HER2-	71 (36,2)
HER2+	73 (37,2)
TN	52 (26,6)
Réponse histologique	
pCR	47 (24)
non-pCR	149 (76)
pCR par phénotype	
ER+/HER2-	2 (1)
HER2+	23 (11,7)
TN	22 (11,2)
	Moyenne (\pm écart interquartile)
MATV (mL)	
Totalité de la population	6.5 (\pm 8.3)
ER+/HER2-	7.2 (\pm 8.5)
HER2+	4.5 (\pm 6.2)
TN	6.9 (\pm 19.4)
SUV_{max}	
Totalité de la population	5.6 (\pm 4.5)
ER+/HER2-	4.4 (\pm 3.2)
HER2+	6.0 (\pm 5.1)
TN	7.0 (\pm 5.5)
SUV_{mean} (mL/min/g)	
Totalité de la population	0.16 (\pm 0.13)
ER+/HER2-	0.16 (\pm 0.16)
HER2+	0.16 (\pm 0.12)
TN	0.17 (\pm 0.10)
Volume perfusé (mL)	
Totalité de la population	9.0 (\pm 7.8)
ER+/HER2-	7.1 (\pm 5.3)
HER2+	9.1 (\pm 8.2)
TN	12.2 (\pm 10.3)
BF_{max} (mL/min/g)	
Totalité de la population	0.33 (\pm 0.27)
ER+/HER2-	0.29 (\pm 0.24)
HER2+	0.30 (\pm 0.28)
TN	0.35 (\pm 0.24)
BF_{mean} (mL/min/g)	
Totalité de la population	0.18 (\pm 0.1)
ER+/HER2-	0.18 (\pm 0.09)
HER2+	0.19 (\pm 0.12)
TN	0.18 (\pm 0.08)

TABLE 3.1 – Caractéristiques des patientes et des tumeurs

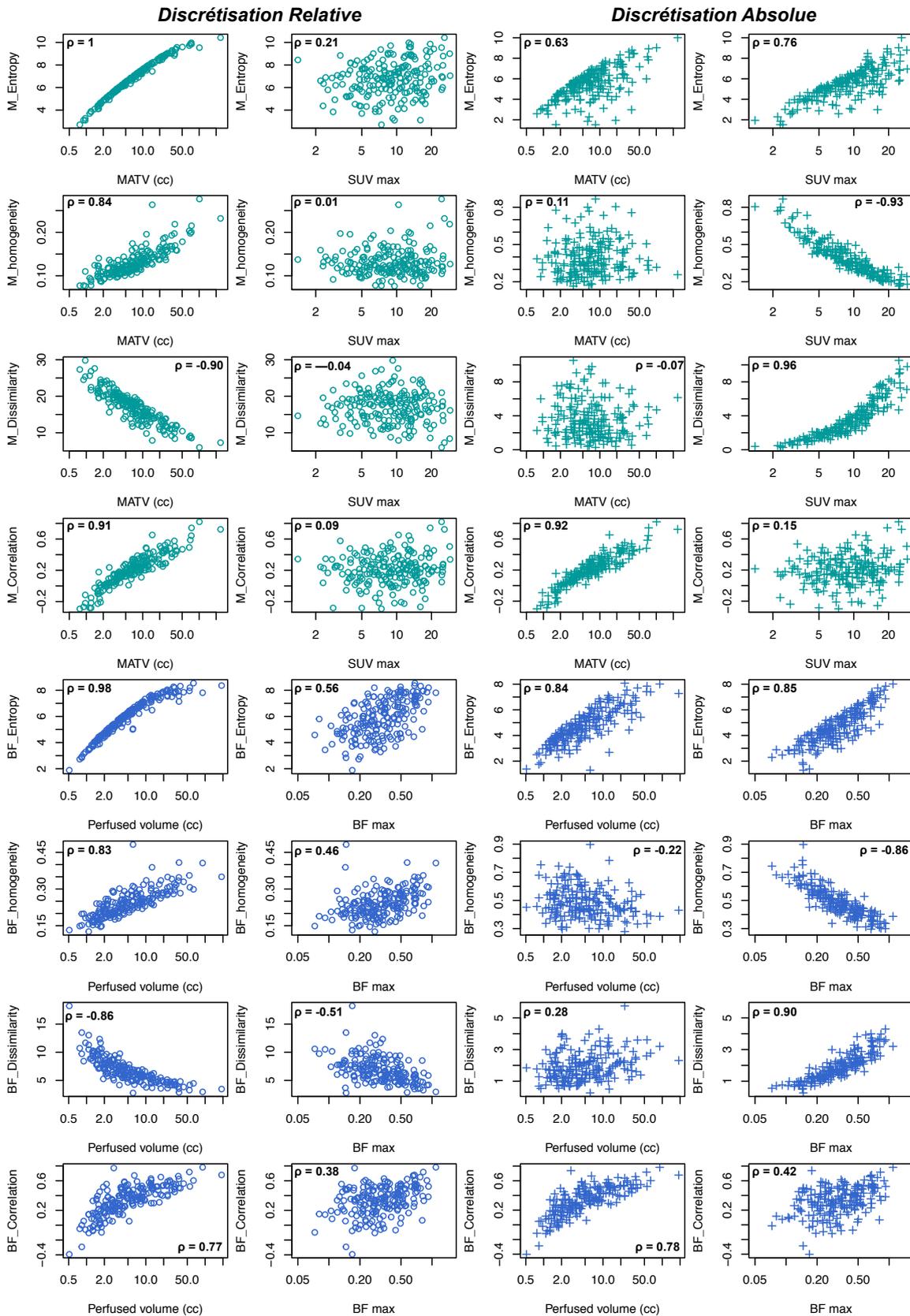


FIGURE 3.10 – Distributions des indices de texture de perfusion (bleu) et métabolisme (vert) en fonction des paramètres TEP, après une discrétisation relative (ronds) et absolue (croix) de l'image, les indices ayant été calculés avec 13 GLCM et avec un rééchantillonnage de 32 bins pour les images de perfusion

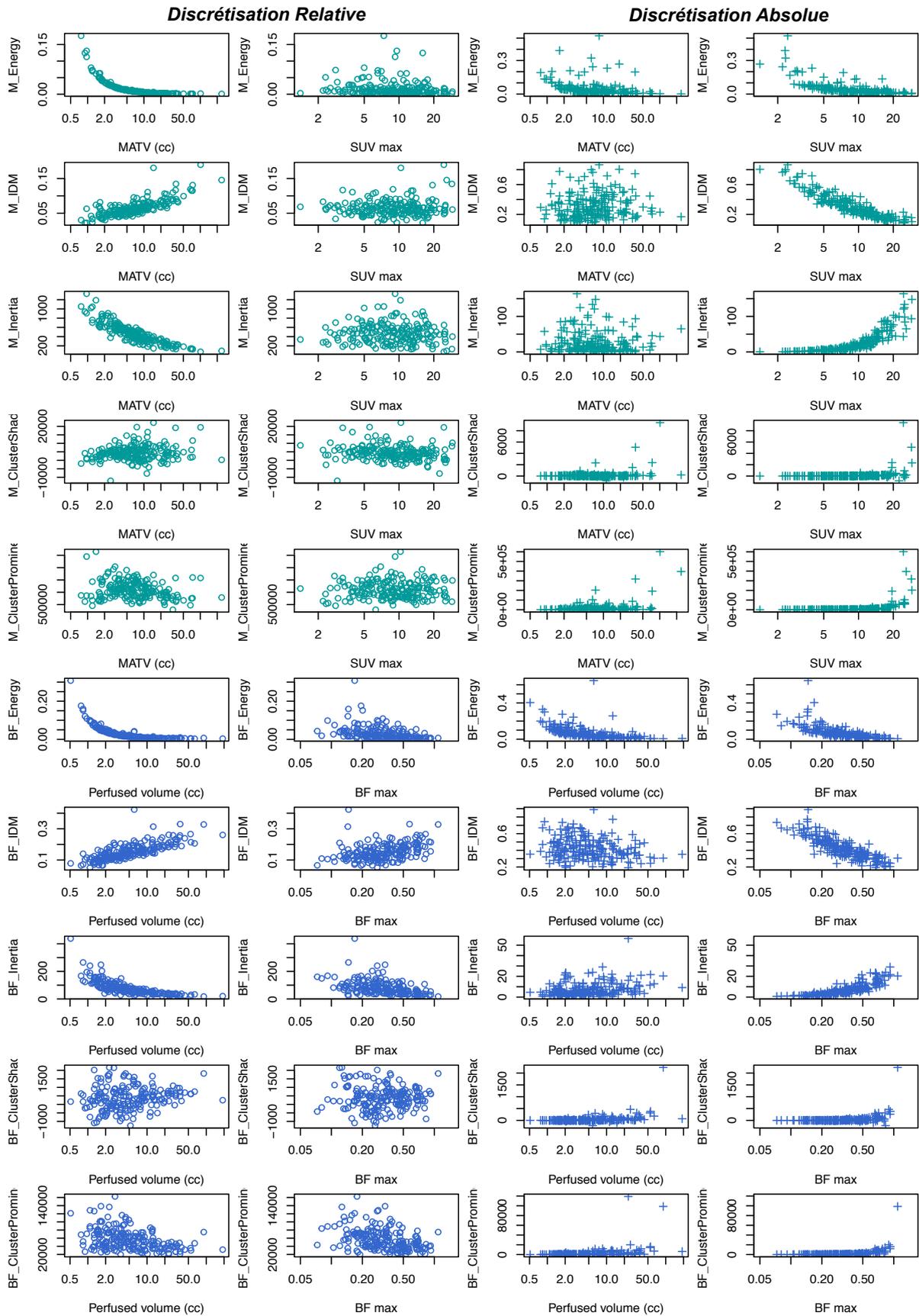


FIGURE 3.11 – Distributions des indices de texture de perfusion (bleu) et métabolisme (vert) en fonction des paramètres TEP, après une discrétisation relative (ronds) et absolue (croix) de l'image, les indices ayant été calculés avec 13 GLCM et avec un rééchantillonnage de 32 bins pour les images de perfusion

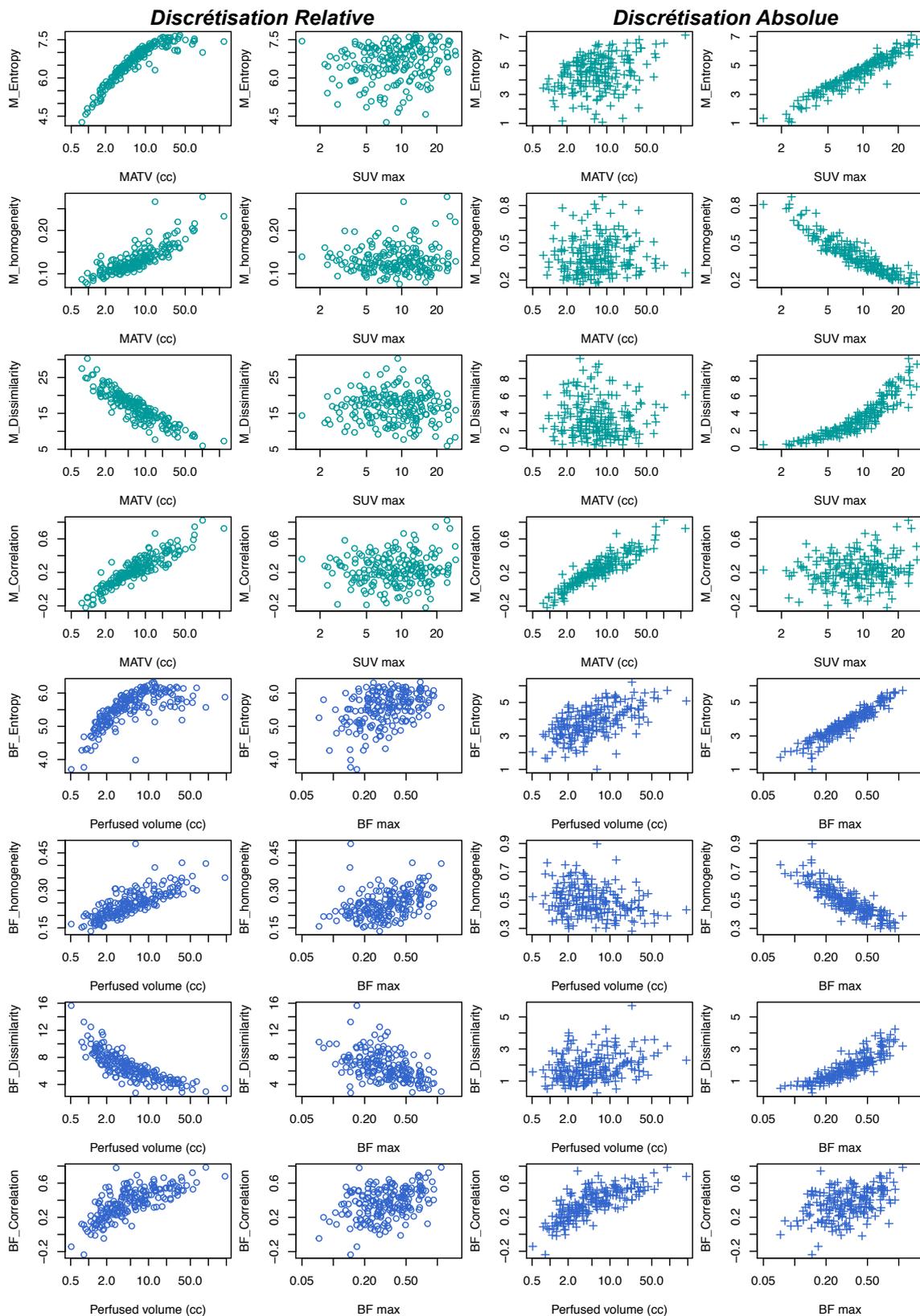


FIGURE 3.12 – Distributions des indices de texture de perfusion (bleu) et métabolisme (vert) en fonction des paramètres TEP, après une discrétisation relative (ronds) et absolue (croix) de l'image, les indices ayant été calculés avec 1 GLCM et avec un rééchantillonnage de 32 bins pour les images de perfusion

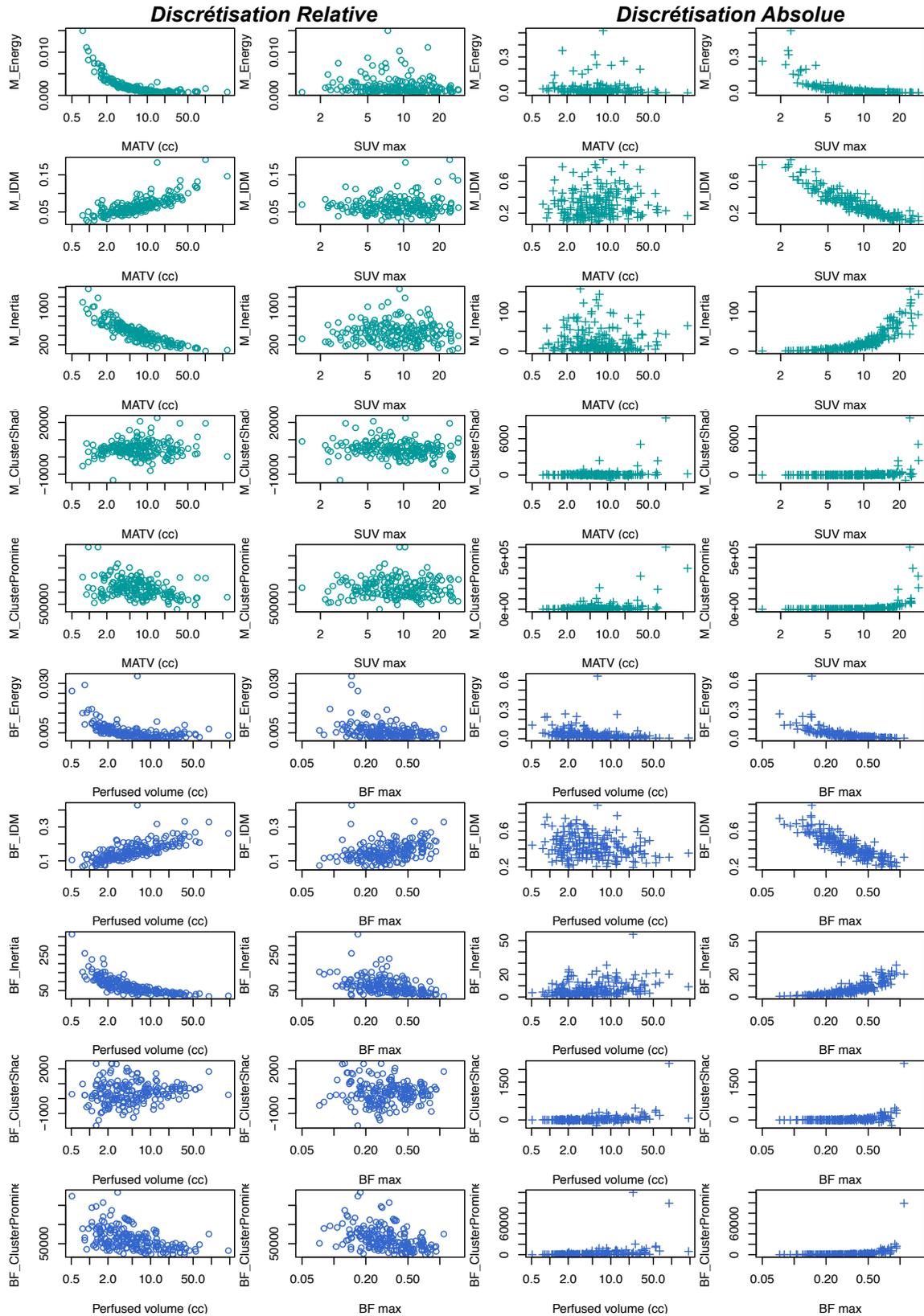


FIGURE 3.13 – Distributions des indices de texture de perfusion (bleu) et métabolisme (vert) en fonction des paramètres TEP, après une discrétisation relative (ronds) et absolue (croix) de l'image, les indices ayant été calculés avec 1 GLCM et avec un rééchantillonnage de 32 bins pour les images de perfusion

3.3.3.2/ INFLUENCE DE LA MÉTHODE DE CALCUL DES INDICES DE TEXTURE SUR LEURS RELATIONS AVEC LES PARAMÈTRES TEP CONVENTIONNELS

Les résultats (fig. 3.10, 3.11 et tableau B.2 page 219) des corrélations entre les indices de texture métaboliques, le VM et le SUV_{\max} montrent que la plupart des indices de texture calculés après une DR sont fortement corrélés avec le VM tandis que la plupart des indices de texture calculés après une DA sont fortement corrélés avec le SUV_{\max} . On note également que les indices *entropy* et *energy* calculés après une DA restent néanmoins corrélées au VM ($\rho = 0,89$ et $-0,88$ pour $Entropy_{13GLCM}$ et $Energy_{13GLCM}$ respectivement et $\rho = 0,68$ et $-0,66$ pour $Entropy_{1GLCM}$ et $Energy_{1GLCM}$). Pour ce qui est des indices de texture issues de l'image paramétrique de perfusion, des relations similaires à celles des indices métaboliques ont été observées entre les indices de texture du BF, le VP et le BF_{\max} . Les indices *entropy* et *energy* calculés après une DA restent, en plus d'être corrélées au BF_{\max} , moyennement à fortement corrélés au volume perfusé ($0,67 \geq \rho \geq 0,94$ et $-0,92 \geq \rho \geq -0,63$ pour $Entropy_{13GLCM}$ et $Energy_{13GLCM}$ et $0,5 \geq \rho \geq 0,74$ et $-0,46 \geq \rho \geq -0,7$ pour $Entropy_{1GLCM}$ et $Energy_{1GLCM}$ respectivement). Par ailleurs, il est intéressant de noter que l'indice de corrélation est quant à lui toujours corrélé au volume tumoral et peu corrélé à l'intensité maximale de la tumeur quelle que soit la méthode de discrétisation. Les distributions des indices étudiés en fonction des volumes d'intérêts (VM ou VP) ou des valeurs d'intensité maximales (SUV_{\max} ou BF_{\max}) sont illustrés sur les figures 3.10, 3.11, 3.12 et 3.13 et tous les coefficients de Spearman sont répertoriés dans les tableaux B.1 page 218 et B.2 page 219.

Pour ce qui est de l'influence de la méthode de calcul de la matrice de co-occurrence, les valeurs des coefficients de Spearman restent majoritairement inchangées, excepté pour les indices *entropy* et *energy* pour lesquels on peut noter quelques différences.

Nombre de bins	Indices de texture	Indices de premier ordre	AR		RR	
			13 GLCM	1 GLCM	13 GLCM	1 GLCM
16	Entropy	BF_{\max}	0.93	0.95	0.51	0.06
	Energy		-0.92	-0.92	-0.48	-0.01
32	Entropy		0.8	0.95	0.56	0.39
	Energy		-0.86	-0.94	-0.55	-0.34
64	Entropy		0.72	0.91	0.56	0.53
	Energy		-0.74	-0.9	-0.56	-0.5
16	Entropy	0.67	0.5	0.93	0.28	
	Energy	-0.63	-0.46	-0.88	-0.16	
32	Entropy	VP	0.84	0.59	0.98	0.78
	Energy		-0.8	-0.55	-0.97	-0.67
64	Entropy		0.94	0.74	0.99	0.95
	Energy		-0.92	-0.7	-0.99	-0.91

TABLE 3.2 – Coefficient de corrélation de Spearman (ρ) entre les indices métaboliques *entropy* et *energy*, et le VP et le BF_{\max}

En effet, lorsque l'image est discrétisée avec la méthode relative, ces indices vont être moins corrélés aux volumes tumoraux ainsi qu'aux valeurs de BF_{\max} que lorsqu'ils sont calculés à partir

d'une unique GLCM. Cette diminution de corrélation est plus ou moins marquée selon le nombre de bin utilisés pour l'échantillonnage de l'image paramétrique de perfusion. De même, lorsque l'image est discrétisée avec la méthode absolue, ces deux indices de texture sont également moins corrélés aux volumes tumoraux lorsqu'ils sont calculés à partir d'une unique GLCM. En revanche, les corrélations avec les distributions des intensités maximales (SUV_{\max} et BF_{\max}) augmentent. Les coefficients de Spearman associés aux indices *entropy* et *energy* sont regroupés dans les tableaux 3.3 pour le métabolisme et le tableau 3.2 pour la perfusion.

Indices de texture	Indices de premier ordre	AR		RR	
		13 GLCM	1 GLCM	13 GLCM	1 GLCM
Entropy	SUV_{\max}	0.76	0.95	0.21	0.22
Energy		-0.79	-0.94	-0.22	-0.21
Entropy	VM	0.63	0.29	1	0.96
Energy		-0.58	-0.27	-1	-0.92

TABLE 3.3 – Coefficient de corrélation de Spearman (ρ) entre les indices métaboliques *entropy* et *energy*, et le VM et le SUV_{\max}

Enfin, si l'on s'intéresse à l'influence du nombre de niveaux d'intensités D utilisés pour le rééchantillonnage des images de perfusion, on peut noter que les distributions des indices en fonction du volume perfusé ou du BF_{\max} restent globalement semblables. En revanche, les valeurs des indices vont varier en fonction du rééchantillonnage, quelle que soit la méthode de discrétisation. On note également que pour certains indices les différences observées vont parfois être plus ou moins marquées en fonction de la taille de la lésion. Par exemple, après une DR, de plus grands écarts de valeurs d'*entropy* sont observés entre les différents rééchantillonnages (D=16, 32 et 64) pour les lésions de grande taille et inversement pour l'indice *dissimilarity*. Ce phénomène est également observé avec le BF_{\max} où, après une DA, l'*energy* va en général avoir de plus grands écarts de valeurs entre les différents rééchantillonnages pour les lésions avec un BF_{\max} faible et inversement pour la *dissimilarity*. Pour finir, on peut noter que l'indice de corrélation n'est que très peu impacté par le choix de la valeur de D. Les distributions des indices de texture calculées avec les différentes valeurs de rééchantillonnage sont illustrées figures 3.14 et 3.15.

3.3.3.3/ DIFFÉRENCES DES PARAMÈTRES TEP ENTRE LES TUMEURS DES PATIENTES PCR ET NON-PCR ET INFLUENCE DE LA MÉTHODE DE DISCRÉTISATION : ANALYSE MANN-WHITNEY

En ce qui concerne l'étude de l'ensemble de la population, aucune différence significative n'a été observée pour les paramètres conventionnels de perfusion ni pour les indices de texture associés entre les patientes pCR et non-pCR, quelle que soit la méthode de discrétisation, le nombre de niveaux d'intensité du rééchantillonnage ou encore la méthode de calcul des GLCM. Pour ce qui est des indices de texture du métabolisme, six indices calculés après une DR ont montré des différences significatives entre les deux groupes de patientes, pour les deux méthodes de calcul des GLCM. En revanche, aucun indice calculé après une DA n'est apparu

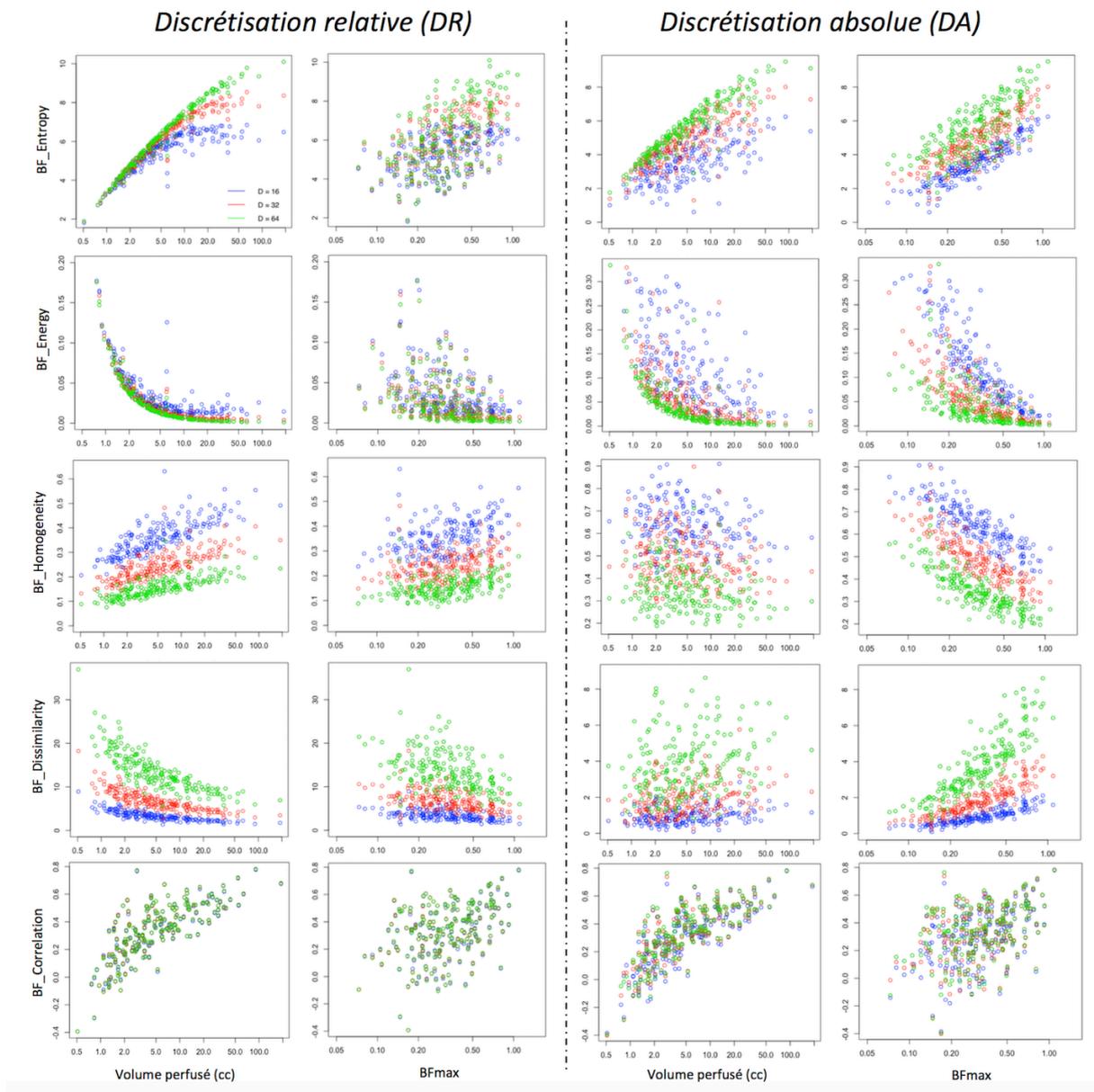


FIGURE 3.14 – Distributions des indices de texture de perfusion en fonction du BF_{max} et du volume perfusé, après une discrétisation relative (gauche) et absolue (droite) de l'image en utilisant 16 (bleu), 32 (rouge) et 64 (vert) niveau d'intensité de rééchantillonnage.

significatif. Parmi les paramètres de premier ordre, seul le volume métabolique a montré une p-value significative entre les deux groupes. Les résultats des indices ayant des p-values significatives après correction (< 0.05) sont regroupés dans le tableau 3.4.

En ce qui concerne les analyses par phénotype, après la correction de Benjamini-Hochberg, aucun indice, que ce soit pour le sous-groupe HER2 ou TN, n'est apparu significativement différent entre les deux groupes de patientes.

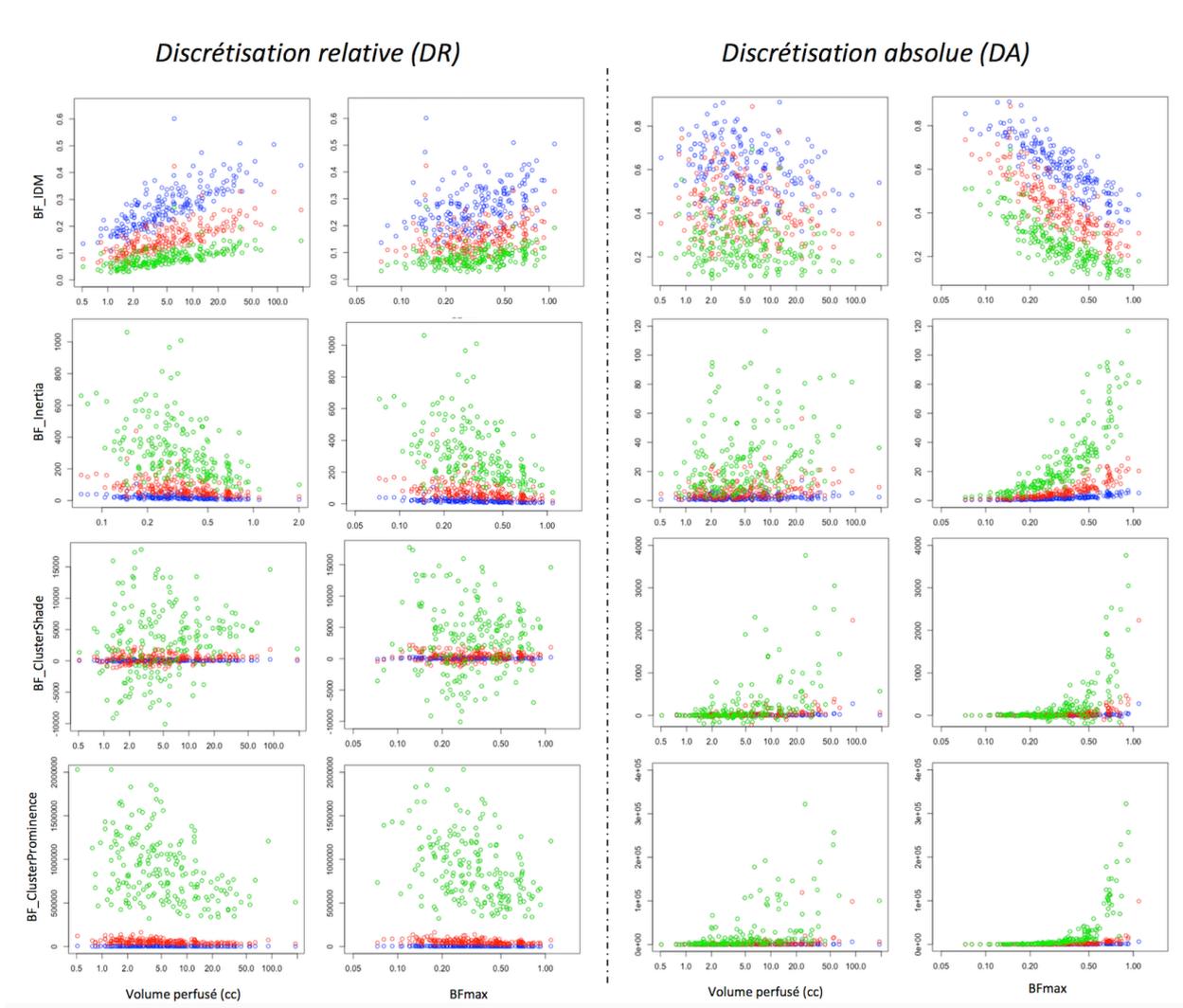


FIGURE 3.15 – Distributions des indices de texture de perfusion en fonction du BF_{max} et du volume perfusé, après une discrétisation relative (gauche) et absolue (droite) de l'image en utilisant 16 (bleu), 32 (rouge) et 64 (vert) niveau d'intensité de rééchantillonnage.

3.3.3.4/ DIFFÉRENCES DE TEXTURE ENTRE LES TUMEURS DES PATIENTES PCR ET NON-PCR ET INFLUENCE DE LA MÉTHODE DE DISCRÉTISATION : ANALYSE FACTORIELLE DISCRIMINANTE DESCRIPTIVE

L'AFD a été appliquée sur la totalité de la population ainsi qu'aux sous-groupes HER2 et TN. Pour chaque analyse, la totalité des indices a été utilisée de manière conjointe, ce qui implique au total 31 variables par modèle. Chaque indice de texture de la perfusion tumorale pouvant être calculé de 12 manières différentes (part. 3.3.2.3), nous avons au total comparé 12 modèles d'AFD (6 par méthode de discrétisation).

Les résultats sur l'étude de la totalité de la population, présentés dans le tableau 3.5, montrent des rapports de corrélation η^2 légèrement plus élevés pour les modèles où les indices

Différences des indices de textures entre les groupes pCR et non-pCR				
Textures features	13 GLCM		1 GLCM	
	AR (pCR vs. non-pCR)	RR (pCR vs. non-pCR)	AR (pCR vs. non-pCR)	RR (pCR vs. non-pCR)
M_energy	ns	0.023 (0.014 vs. 0.027)	ns	0.04 (0.002 vs. 0.003)
M_entropy	ns	0.023 (6.98 vs. 6.17)	ns	0.035 (6.7 vs. 6.3)
M_homogeneity	ns	0.024 (0.13 vs. 0.12)	ns	0.033 (0.14 vs. 0.12)
M_inertia	ns	0.016 (439 vs. 577)	ns	0.016 (426 vs. 561)
M_correlation	ns	0.016 (0.24 vs. 0.13)	ns	0.016 (0.26 vs. 0.16)
M_dissimilarity	ns	0.016 (16.4 vs. 18.9)	ns	0.016 (16.1 vs. 18.6)

TABLE 3.4 – Résultats des tests de Mann-Whitney entre les groupes pCR et non-pCR pour l'ensemble de la population. Les valeurs des p-values significatives ($< 0,05$, après correction de tests multiples) sont indiquées avec la valeur moyenne des indices de texture pour chaque groupe entre parenthèses lorsqu'ils sont significatifs. DA : discrétisation absolue, GLCM : matrice de co-occurrence de niveau de gris, IDM : inverse different moment, M : métabolisme, ns : non significatif, DR : discrétisation relative.

de texture utilisés ont été calculés après une DR des images métaboliques et de perfusion. Le meilleur rapport de corrélation observé est de 0,247, ce qui signifie que le modèle est capable de distinguer les patientes pCR des non-pCR dans 24,7% des cas. Ce résultat a été noté pour des indices de texture calculés après une DR en utilisant 13-GLCM et avec un rééchantillonnage des images de perfusion de 32 bins (modèle 1).

Pour les analyses par phénotypes, le meilleur rapport de corrélation ($\eta^2 = 0,509$) pour le sous-groupe HER2 a été trouvé pour le modèle dont les indices de texture ont été calculés après une DR en utilisant 13-GLCM et avec un rééchantillonnage des images de perfusion de 64 bins (modèle 2). En ce qui concerne le sous-groupe TN, le meilleur rapport de corrélation, égal à 0,678, a été observé pour le modèle dont les indices de texture ont été calculés après une DR en utilisant une unique GLCM et avec un rééchantillonnage des images de perfusion de 16 bins (modèle 3).

Que ce soit pour la population totale ou pour le sous-groupe HER2, les rapports de corrélation les plus élevés sont toujours observés pour les modèles où les indices de texture sont calculés après une DR des images, quelle que soit la méthode de calcul de la GLCM ou de rééchantillonnage des images de perfusion. En revanche, ce n'est pas le cas pour le phénotype TN. En effet, si l'on prend les modèles où les indices de texture ont été calculés à partir de 13 GLCM, le meilleur rapport de corrélation est alors observé pour un modèle où les indices de texture ont été calculé après une DA.

Si l'on se réfère aux observations sur nos données ainsi qu'à la littérature [397], le principal impact de la méthode de calcul de la GLCM concerne les relations entre les indices de texture et les volumes tumoraux. Or, en observant les caractéristiques volumiques des tumeurs par phéno-

Résultats de l'AFD entre les groupes pCR et non-pCR						
Population (pCR vs. non-pCR)	BF bins	M bins	13 GLCM		1 GLCM	
			AR	RR	AR	RR
Totalité de la population (47 vs. 149)	16	64	0.243	0.247*	0.224	0.242
	32	64	0.235	0.247	0.226	0.229
	64	64	0.224	0.244	0.214	0.245
Phénotype HER2 (23 vs. 50)	16	64	0.414	0.457	0.417	0.421
	32	64	0.441	0.472	0.427	0.423
	64	64	0.445	0.509	0.414	0.454
Phénotype TN (22 vs. 30)	16	64	0.555	0.574	0.53	0.678
	32	64	0.607	0.56	0.566	0.614
	64	64	0.581	0.561	0.57	0.605

TABLE 3.5 – Rapports de corrélation (η^2) de l'AFD descriptive pour l'ensemble de la population, pour les sous-groupes HER2 et TN. Les rapports de corrélation maximaux sont écrits en gras pour chaque groupe de patients. La valeur marquée d'une * a été arrondie au supérieur. DA : discrétisation absolue, BF : *blood flow*, GLCM : matrice de co-occurrence du niveau de gris, M : métabolisme, DR : discrétisation relative.

Volumes tumoraux du phénotype TN				
Volumes	VP (mL)		VM (mL)	
	médiane	écart interquartile	médiane	écart interquartile
Totalité	6.592	20.032	6.88	19.36
≤ 70 mL	5.952	9.984	6.400	9.216
≤ 50 mL	5.888	9.68	6.336	9.152
≤ 40 mL	5.408	8.864	6.208	8.464
≤ 30 mL	4.928	5.824	5.632	6.784
≤ 25 mL	4.288	5.152	5.280	6.144

TABLE 3.6 – Médiane et interquartiles des volumes tumoraux du phénotype TN en fonction de différents seuils volumiques.

type, on constate que le sous-groupe TN est celui dont les volumes tumoraux et leurs dispersions sont en moyenne les plus importants (cf. tableau 3.1). En prenant en compte cette information, nous avons décidé d'explorer l'impact, pour ce phénotype, du volume sur la classification de l'AFD des patientes pCR et non-pCR. Pour cela, nous avons sélectionné des sous-ensembles des patientes TN ayant des volumes tumoraux inférieurs ou égaux à 25, 30, 40, 50 et 70 mL (tab. 3.6). Une AFD a ensuite été appliquée à chaque sous-ensembles de patientes. Les résultats, regroupés dans le tableau 3.7, montrent que lorsque les indices de texture sont calculés avec une unique GLCM, les modèles donnant les meilleurs rapports de corrélation sont toujours ceux où les indices de texture ont été calculés après une DR, quelle que soit la taille des tumeurs étudiées. En revanche, lorsque les indices de texture sont calculés à partir de 13 GLCM, on

Impact du volume des tumeurs TN sur les résultats de l'AFD						
Volume tumoraux	Totalité des tumeurs	≤ 70 mL	≤ 50 mL	≤ 40 mL	≤ 30 mL	≤ 25 mL
Nombre de patients	52	49	48	46	42	38
Meilleure méthode en utilisant 13 GLCM (nb de bins)	AR(32)	AR(32)	AR(32)	AR(64)	AR(64)/RR(16)	RR(16)
Meilleure méthode en utilisant 1 GLCM (nb de bins)	RR(16)	RR(64)	RR(16)	RR(16)	RR(16)	RR(16)

TABLE 3.7 – Rapports de corrélation (η^2) de l'AFD descriptive pour les différents sous-groupes TN

remarque que les modèles séparant au mieux les groupes de patientes pCR et non-pCR vont être variables : on observe les meilleurs séparations pCR/non-pCR en utilisant des modèles où les indices de texture ont été calculé après une DR lorsque l'on sélectionne des (petits) volumes tumoraux inférieurs à 30 mL. En revanche, on obtient une meilleure séparation pCR/non-pCR en utilisant des modèles où les indices de texture ont été calculés après une DA si l'on inclut les tumeurs de plus de 30 mL. Au final, pour les trois groupes étudiés, i.e. la population totale, le phénotype HER2 et le phénotype TN, les poids canoniques des variables discriminantes des modèles 1, 2 et 3 ont été collectés et regroupés dans le tableau 3.8.

3.3.4/ DISCUSSION ET CONCLUSION DE L'ÉTUDE

L'objectif principal de cette étude préliminaire était d'identifier la méthode de calcul des indices de texture qui se présenterait comme la plus appropriée pour identifier les patientes ayant eu une réponse histologique complète.

3.3.4.1/ INFLUENCE DE LA MÉTHODE DE DISCRÉTISATION ET DU CALCUL DES GLCM SUR LES INDICES DE TEXTURE

Cette première étude d'analyse de texture sur des images TEP montre que les indices de texture de perfusion présentent globalement les mêmes relations que ceux du métabolisme avec les paramètres TEP conventionnels. En ce qui concerne l'influence de la méthode de calcul GLCM pour chaque méthode de discrétisation, les résultats ont montré que seulement deux des indices étudiés, l'*entropy* et l'*energy*, semblaient être impactés par le choix de la méthode de calcul GLCM. Ces deux indices calculés avec une unique GLCM se sont avérés moins corrélés aux volumes tumoraux (métabolique et perfusé) et cela pour les deux méthodes de discrétisation des images. Cette observation est en accord avec une précédente étude dont les résultats montrent une baisse de la corrélation entre l'*entropy* (calculée à partir d'une DR) et le volume métabolique [439]. Cette étude rapporte également une baisse de la corrélation entre l'indice *dissimilarity* et le VM, phénomène qui n'est cependant pas observé dans notre étude. Toutefois, la baisse de corrélation pour cet indice est particulièrement marquée pour les gros volumes tumoraux et a été constatée pour des volumes ≥ 25 mL, ce qui pourrait expliquer que cela ne soit pas observée sur notre cohorte de patients.

Indices de texture	Population totale (n = 196)	Phénotype HER2 (n = 73)	Phénotype TN (n = 52)
SUV _{mean}	1.33	1.33	1.69
SUV _{sd}	1.80	1.84	2.79
M _{skewness}	0.89	1.00	1.12
M _{kurtosis}	0.70	0.29	1.54
VM	1.15	0.38	0.34
SUV _{max}	2.91	0.80	4.37
TLG	0.73	0.15	1.10
M _{energy}	0.60	0.22	0.79
M _{entropy}	0.32	0.08	0.00
M _{IDM}	6.16	1.74	18.98
M _{homogeneity}	10.47	5.27	26.48
M _{inertia}	5.32	4.64	10.89
M _{correlation}	0.17	2.06	0.50
M _{dissimilarity}	10.27	6.79	18.50
M _{clustershade}	0.64	0.97	0.30
M _{clusterprominence}	0.31	0.27	0.04
VP	0.05	1.57	1.34
BF _{max}	0.92	1.11	1.11
BF _{mean}	0.27	0.05	0.94
BF _{sd}	0.60	0.85	0.61
BF _{skewness}	0.38	0.23	0.09
BF _{kurtosis}	0.02	0.07	0.33
BF _{energy}	0.36	1.94	2.19
BF _{entropy}	0.38	2.44	1.47
BF _{IDM}	7.25	8.14	18.37
BF _{homogeneity}	9.89	10.61	14.45
BF _{inertia}	4.27	2.72	6.47
BF _{correlation}	0.50	0.07	3.26
BF _{dissimilarity}	7.44	5.58	13.03
BF _{clustershade}	0.15	0.04	0.46
BF _{clusterprominence}	0.24	0.13	1.39

TABLE 3.8 – Valeurs absolues des poids canoniques de l’AFD descriptive pour les différents paramètres TEP. Les 10 poids canoniques les plus élevés pour chaque population sont indiqués en gras. BF : *blood flow*, IDM : *inverse different moment*, M : métabolisme, VM, volume métabolique, SUV : *standardized uptake value*, TLG : *total lesion glycolysis*.

En ce qui concerne l’échantillonnage des images de perfusion, les distributions des indices en fonction du volume perfusé ou du BF_{max} restent globalement semblables, contrairement à leurs valeurs qui vont varier en fonction du nombre de niveaux d’intensités D utilisés. Les différences sont toutefois plus marquées pour certains indices en fonction de la taille du volume perfusé ou de l’intensité maximale dans la tumeur : de plus grands écarts de valeurs d’*entropy* sont par exemple observés entre les différents rééchantillonnages pour des larges volumes perfusés que pour des petits volumes.

Enfin, il est intéressant de noter que l'indice de corrélation semble robuste à la fois à la méthode de discrétisation mais également à l'échantillonnage des images. Ce constat est en accord avec des précédents résultats d'analyses effectuées sur des images de métabolisme tumoral [388, 418, 423].

3.3.4.2/ INFLUENCE DE LA MÉTHODE DE CALCUL DES INDICES DE TEXTURE DANS LA DIFFÉRENTIATION DES PATIENTES pCR ET NON-pCR

Pour l'ensemble de la population, les résultats de Mann-Whitney ont montré que certains indices de texture calculés après une DR étaient significativement différents entre les deux groupes, contrairement aux indices calculés après une DA, suggérant ainsi une meilleure identification des patientes pCR et non-pCR en utilisant une DR. Toutefois, le VM présentant également une valeur de p-value significative et étant corrélés aux indices de texture lorsqu'ils sont calculés après une DR, il n'est pas possible, sans analyse multivariée, d'évaluer l'apport réel des indices de texture par rapport au volume métabolique dans la classification des patientes.

L'AFD a également montré une meilleure classification des patientes pour les modèles utilisant des indices de texture calculés après une DR. Néanmoins, dans l'étude de la population totale, la différence de performance avec les modèles utilisant des indices calculés après une DA reste faible, ce qui encourage à la prudence vis à vis de l'interprétation de ces résultats. En ce qui concerne les sous-groupes HER2 et TN, les résultats de l'AFD montrent des différences plus marquées entre les modèles, avec toujours une meilleure identification des groupes pCR/non-pCR pour ceux utilisant les indices calculés après une DR. Ces modèles permettent de distinguer les patientes pCR des non-pCR dans 50,9% et 67,8% des cas pour HER2 et TN respectivement. Ces scores de classification sont bien meilleurs que ceux observés pour la population totale (24,7%), ce qui peut notamment être expliqué par la proportion pCR/non-pCR plus équilibrée dans ces sous-groupes : 23,8%, 31,5% et 41,2% de pCR pour l'ensemble de la population, les sous-types HER2 et TN respectivement.

Pour ce qui est du phénotype TN, la variabilité des résultats entre les modèles utilisant 13 GLCM et ceux utilisant une unique GLCM semble avoir un lien avec la dispersion et la taille des volumes tumoraux présents dans ce sous-groupe. Une analyse utilisant une unique GLCM apparaît alors plus robuste.

Enfin, l'étude des poids canoniques des variables discriminantes a montré que parmi tous les paramètres métaboliques standards, seul le SUV_{max} présentait un poids canonique relativement élevé pour l'ensemble de la population et le sous-groupe TN. Bien que le SUV_{max} ne puisse individuellement pas permettre une classification des patientes pCR et non-pCR comme l'illustrent les résultats du Mann-Whitney ainsi que d'autres études [3, 142], il pourrait néanmoins être utile pour cette classification lorsqu'il est considéré conjointement avec d'autres caractéristiques TEP. L'examen des poids canoniques montre également que seuls des faibles poids sont attribués aux paramètres globaux de perfusion, quelle que soit la population étudiée. Cette observation peut être mise en parallèle avec les résultats du Mann-Whitney où aucune des caractéristiques globales de perfusion n'a présenté de différence significative. Une étude récente a

également rapporté que le débit sanguin des lésions mammaires avant le début du traitement néoadjuvant ne semble pas associé à une réponse histologique complète [253]. Cependant, il convient de noter qu'en fonction de la population étudiée, quatre à six indices de texture de la perfusion sont parmi les dix poids canoniques des variables les plus élevés. Cela suggère que l'ajout de l'information de l'hétérogénéité de la perfusion combinée aux paramètres métaboliques améliore la classification des patientes. Enfin, la plupart des indices métaboliques qui sont apparus significatifs avec l'analyse Mann-Whitney ont des poids canoniques parmi les 10 plus élevés et sont également supérieurs au poids du VM. Ces résultats semblent donc montrer l'apport additionnel des indices de texture par rapport à celui du volume tumoral.

3.3.4.3/ MÉTHODE DE DISCRÉTISATION ET OBJECTIF CLINIQUE

Parmi les trois études qui ont évalué l'impact de la méthode de discrétisation des images TEP pour la classification des tumeurs [387,388,430], Orhac *et al.* et Leijenaar *et al.* ont conclu qu'une discrétisation absolue permettait une meilleure classification des tumeurs pulmonaires. De leur côté, Shen *et al.* optent pour l'utilisation d'une DR afin d'obtenir de meilleures performances dans la prédiction des métastases ganglionnaires pelviennes des cancers du col de l'utérus. Dans le contexte du cancer du sein, les résultats actuels suggèrent également que les indices de texture du métabolisme et de perfusion calculés après une DR permettent une meilleure classification des patientes pCR et non-pCR. Tous ces résultats semblent illustrer que l'efficacité des indices de texture à pouvoir répondre à un objectif clinique spécifique peut être influencé par la méthodologie de calcul de ces derniers. Les résultats divergents de ces quatre études peuvent éventuellement s'expliquer en partie par la diversité des caractéristiques des cancers étudiés et des objectifs cliniques visés. En effet, chacune des quatre études ayant des objectifs variés (différenciation des types de cancers, prédiction de métastases, réponse au traitement, etc.), la comparaison directe des résultats peut sembler délicate. Dans la présente étude, nous avons décidé d'utiliser le statut de réponse au traitement comme variable définissant nos deux groupes de patientes à identifier. Ce choix a été fait avec l'objectif d'optimiser la méthode de calcul des indices de texture pour une future étude sur la prédiction de la réponse à la CNA. Toutefois, il est important de garder à l'esprit que si le choix d'un autre aspect clinique avait été sélectionné, comme la présence/quantité de marqueurs de prolifération ou de vascularisation (Ki-67, CD34, etc.), la conclusion aurait potentiellement pu être différente.

3.3.4.4/ LIMITATIONS ET CONCLUSION DE L'ÉTUDE

L'une des limitations de cette étude reste les faibles volumes tumoraux sur lesquels l'analyse de texture est effectuée. Néanmoins, la petite taille des tumeurs est l'une des caractéristiques du cancer du sein et le choix de l'utilisation d'indices de texture locaux (et non régionaux) a été décidée vis à vis de cette dernière. Aucune correction d'effet de volume partiel n'a par ailleurs été effectuée. Par ailleurs, le biais de quantification inhérent à la faible statistique des images dynamiques pourrait également affecter les résultats [273]. Enfin, il a été démontré que les acquisitions TEP de haute résolution améliorent la caractérisation des lésions mammaires [440]. Toutefois, la modélisation de la PSF n'est malheureusement pas disponible

sur l'imagerie utilisé dans cette étude et les matrices de voxels ont, quant à elles, été générées selon les mêmes conditions qu'en routine clinique, c'est à dire dans notre cas avec des voxels de taille assez importante ($4 \times 4 \times 4 \text{ mm}^3$).

Pour conclure, cette étude a confirmé que les relations des indices de texture du métabolisme avec les paramètres TEP conventionnels varient en fonction de la méthode de discrétisation. Ce travail a également permis d'observer que les indices de texture de la perfusion présentent globalement les mêmes comportements que ceux du métabolisme, malgré les différences de caractéristiques entre ces images. Par ailleurs, en utilisant les informations de la perfusion et du métabolisme tumoral, les résultats suggèrent qu'une DR permettrait de sensiblement mieux identifier les patients pCR des non-pCR dans la présente cohorte. En outre, l'analyse factorielle discriminante descriptive suggère que l'information de l'hétérogénéité de la perfusion combinée aux paramètres métaboliques de la TEP améliore la classification des patientes. Enfin, malgré les corrélations observées, les indices de texture métaboliques semblent apporter une information additionnelle complémentaire à celle du volume tumoral.

3.4/ RELATIONS ENTRE L'HÉTÉROGÉNÉITÉ TUMORALE ET LES PARAMÈTRES BIOLOGIQUES ET MOLÉCULAIRES DES TUMEURS

3.4.1/ MOTIVATIONS ET OBJECTIFS

Les études présentées dans ce chapitre ont pour objectif d'essayer de caractériser les tumeurs mammaires via leur information perfusionnelle et métabolique. La précédente analyse avait pour objectifs d'identifier si une méthode de calcul des indices de texture apparaissait plus appropriée au contexte clinique étudié. Au vu des résultats, les paramètres qui ont été sélectionnés pour cette nouvelle étude sont les suivants :

- une discrétisation relative des images TEP en utilisant 13 GLCM. En effet, même si une unique GLCM est apparue plus robuste sur l'étude du phénotype TN, nous avons toutefois décidé de privilégier le choix de 13 GLCM qui a donné de meilleurs résultats sur l'analyse de la totalité de la cohorte et sur le sous-groupe HER2 ;
- un rééchantillonnage des images de perfusion effectué avec 32 bins. Ce choix a été fait de manière à privilégier le nombre de bins donnant le meilleur résultat sur l'étude de la totalité de la cohorte dans un esprit de généralisation de l'application des indices de texture. Par ailleurs, au vu du plus faible rapport signal sur bruit dans les images de perfusion par rapport aux images de métabolisme, ce choix semble être un compromis afin de limiter partiellement l'impact du bruit ;
- en ce qui concerne les paramètres de texture, en recoupant les résultats du Mann-Whitney et des poids canoniques des variables discriminantes de l'AFD, sept indices de texture se sont distingués : l'entropy, l'energy, l'inertia, l'homogeneity, la dissimilarity, la correlation et IDM. Parmi ces paramètres, quatre ont été sélectionnés pour l'étude de l'hétérogénéité tumorale en se basant sur leur robustesse et répétabilité : l'*entropy*, l'*homogeneity*, la

dissimilarity et la *correlation*.

Le travail présenté par la suite s'est focalisé sur la caractérisation des tumeurs mammaires en analysant leur hétérogénéité de perfusion et de métabolisme. En effet, l'étude conjointe des paramètres TEP conventionnels et des indices de texture métaboliques est une opportunité pour étudier plus en détail les relations entre le métabolisme glucidique des tumeurs et d'autres paramètres cliniques, biologiques ou même représentatifs de l'agressivité tumorale [441–445]. Nous avons donc souhaité étudier en parallèle l'hétérogénéité de la perfusion tumorale qui pourrait potentiellement elle aussi, permettre une meilleure caractérisation des lésions mammaires et approfondir par la même occasion, les relations entre la perfusion de la tumeur et son métabolisme. L'objectif de la présente étude est donc d'évaluer les relations entre les paramètres d'hétérogénéité tumorale, de métabolisme et de perfusion, avec les caractéristiques biologiques des tumeurs et les sous-types moléculaires.

Concernant les patientes incluses dans cette étude, la connaissance de la réponse histologique à la CNA n'étant pas étudiée, la cohorte de patientes a pu être utilisée dans son ensemble avec 217 patientes.

3.4.2/ ARTICLE



Biological correlates of tumor perfusion and its heterogeneity in newly diagnosed breast cancer using dynamic first-pass ^{18}F -FDG PET/CT

Neree Payan¹ · Benoit Presles¹ · François Brunotte¹ · Charles Coutant² · Isabelle Desmoulins³ · Jean-Marc Vrigneaud^{1,4} · Alexandre Cochet^{1,4}

Received: 3 April 2019 / Accepted: 1 July 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Purpose The aim of this prospective study is to analyze the global tumor blood flow (BF) and its heterogeneity in newly diagnosed breast cancer (BC) according to tumor biological characteristics and molecular subtypes. These perfusion parameters were compared to those classically derived from metabolic studies to investigate links between perfusion and metabolism.

Methods Two hundred seventeen newly diagnosed BC patients underwent a ^{18}F -FDG PET/CT exam before any treatment. A 2-min dynamic acquisition, centered on the chest, was performed immediately after intravenous injection of 3 MBq/kg of ^{18}F -FDG, followed by a two-step static acquisition 90 min later. Tumor BF was calculated (in ml/min/g) using a single compartment kinetic model. In addition to standard PET parameters, texture features (TF) describing the heterogeneity of tumor perfusion and metabolism were extracted. Patients were divided into three groups: Luminal (HR+/HER2-), HER2 (HER2+), and TN (HR-/HER2-). Global and TF parameters of BF and metabolism were compared in different groups of patients according to tumor biological characteristics.

Results Tumors with lymph node involvement showed a higher perfusion, whereas no significant differences in SUV_{max} or SUV_{mean} were reported. TN tumors had a higher metabolic activity than HER2 and luminal tumors but no significant differences in global BF values were noted. HER2 tumors exhibited a larger tumor heterogeneity of both perfusion and metabolism compared to luminal and TN tumors. Heterogeneity of perfusion appeared well correlated to that of metabolism.

Conclusions The study of breast cancer perfusion shows a higher BF in large tumors and in tumors with lymph node involvement, not paralleled by similar modifications in tumor global metabolism. In addition, the observed correlation between the perfusion heterogeneity and the metabolism heterogeneity suggests that tumor perfusion and consequently the process of tumor angiogenesis might be involved in the metabolism heterogeneity previously shown in BC.

Keywords Breast cancer · ^{18}F -FDG PET/CT · Blood flow · Heterogeneity · Textural features

This article is part of the Topical Collection on Oncology – Chest

✉ Neree Payan
nerea.payan@u-bourgogne.fr

- ¹ ImViA EA 7535, University of Burgundy, Dijon, France
- ² Department of Surgical Oncology, Georges-François Leclerc Cancer Center, Dijon, France
- ³ Department of Medical Oncology, Georges-François Leclerc Cancer Center, Dijon, France
- ⁴ Department of Nuclear Medicine, Georges-François Leclerc Cancer Center, Dijon, France

Introduction

Breast cancer (BC) is the most common cancer and the second leading cause of death by cancer among women. This cancer has various molecular subtypes that differ in aggressiveness, treatment, and outcome [39, 52]. As a great heterogeneity regarding genetic profile, aggressiveness, and therapeutic options [43] persist, it appears essential to develop additional diagnostic tools to better characterize tumors and to guide therapeutic strategies, using biology, next-generation sequencing, and imaging. Tumors are often associated with an increase in glucose metabolism [31]

and in the context of BC, baseline metabolism is positively associated with tumor aggressiveness [15]. Beyond glucose metabolism, tumors have an increased blood flow (BF), due to the angiogenesis [35]. Since a mismatch between tumor metabolism and tumor blood flow has been reported [30], evaluation of these two different parameters appears essential to fully characterize breast tumors. Different imaging modalities, such as dynamic contrast-enhanced magnetic resonance imaging (MRI), ^{15}O -water positron emission tomography/computed tomography (PET/CT) and 1-h dynamic ^{18}F -fluorodeoxyglucose (^{18}F -FDG) PET/CT, have yielded kinetic parameters reflecting vascular permeability and perfusion. Methods based on the 2-min half-life ^{15}O remain the gold standard for perfusion evaluation [25]. However, because an on-site cyclotron is needed, only a restricted number of research centers are able to use ^{15}O -water. Furthermore, ^{18}F -FDG PET has become a standard for primary staging of large and locally advanced breast cancer [9] and a full 1-h dynamic ^{18}F -FDG PET acquisition allows tumor blood flow and metabolism evaluation after a single injection of ^{18}F -FDG. As a 1-h acquisition is hardly suitable in clinical routine, an alternative method allowing the measurement of the tumor blood flow from a first-pass ^{18}F -FDG PET has been developed by Mullani et al. [33]. The calculated blood flow [10, 20, 21] showed an excellent correlation with the tumor perfusion measured with ^{15}O -water [34]. In the context of BC, few studies have suggested the interest of the perfusion evaluation to monitor BC therapies and its ability to provide predictive and prognostic information [10, 13, 20, 21, 29, 46].

Recently, investigation of tumor heterogeneity using texture analysis (TA) has emerged [44, 45]. Standard PET/CT parameters combined with TA [2, 3, 19] appear promising to bring valuable information on the relationship between metabolism heterogeneity and clinical, biological parameters, tumor subtypes, or tumor aggressiveness [16, 28, 32, 44, 45]. In a similar way, the present paper hypothesizes that perfusion heterogeneity analysis might bring valuable additional information to metabolic imaging with ^{18}F -FDG PET/CT. Indeed, the reasons why tumors are heterogeneous are still largely unknown and one of the explanations could be the vascular disorders related to the anarchic nature of the tumor angiogenesis [1]. Nonetheless, although several previous studies have quantified ^{18}F -FDG uptake intratumoral heterogeneity, to the best of our knowledge, none was carried out on tumor BF heterogeneity. The first-pass ^{18}F -FDG PET/CT represents an opportunity to assess tumor perfusion in addition to tumor metabolism and to evaluate in a single exam, perfusion heterogeneity and its relation to metabolism heterogeneity.

The objective of this prospective study is to analyze the global tumor blood flow and its heterogeneity in a large cohort of 217 newly diagnosed breast cancer cases

according to tumor biological characteristics and molecular subtypes. These perfusion parameters were compared to those classically derived from metabolic studies to further investigate links between perfusion and metabolism.

Materials and methods

Patients and study design

From January 2011 to May 2017, 217 patients with newly diagnosed stage II or III BC were prospectively recruited, based on oncologists' recommendations for a ^{18}F -FDG PET/CT scan before any treatment. Patients with high glycemia (> 9 mmol/l) and unable or unwilling to undergo PET scan were excluded. The institutional review board approved this prospective study as a current-care study. The medical team documented patient non-opposition in source documents and in the information notice provided to the patient, and the need for a signature indicating informed consent was not required. Core biopsies of primary tumors were used to evaluate histological characteristics, such as histological type and tumor grade, using the Scarff–Bloom–Richardson (SBR) system [6]. Tumor, lymph nodes, metastasis, and classification of malignant tumors (TNM) were established in accordance with the 8th edition of the American Joint Committee on Cancer (AJCC) staging system [4]. Molecular markers examined included estrogen receptor (ER), progesterone receptor (PR) and HER2 expression. ER and PR status were considered positive if tumors showed at least 10% of positive cells. HER2 status was assessed by immunohistochemistry or fluorescent in situ hybridization (FISH) and graded from 0 to 3+. Scores of 3+ were considered positive. HER2 amplification confirmed by FISH was used to assess positivity in ambiguous cases (2+). Patients were classified into three groups according to their molecular subtypes: luminal (hormone receptors (HR)-positive, HER2-negative), HER2 (HER2-positive) and triple negative (TN) (HR-negative, HER2-negative).

PET-CT protocol

All images were acquired with a Gemini TruFlight PET/CT scanner (Philips Medical Systems, Eindhoven, The Netherlands) providing an axial field of view of 18 cm and a transaxial slice thickness of 4 mm. An automatic PET infusion system (Bayer Medical Care, Inc., Indianola, PA, USA) was used to inject a bolus of 3 MBq/kg of ^{18}F -FDG. Simultaneously with injection, an early first 8-min ^{18}F -FDG PET scan was performed, in a prone position and centered on the breast. A low-dose CT scan (120 kVp, automatic tube current modulation, 0.5-s rotation time, 16×1.5 mm

collimation, a pitch of 0.7) of the same region was carried out for anatomical registration and attenuation correction. The dynamic first-pass was reconstructed based on twelve 10-s frames extracted from the first 2 min of the PET acquisition. A whole-body emission and transmission scans were acquired 60 min later, in a supine position. Finally, 90 min after injection, the delayed static image was acquired in a prone position with a two-step PET/CT scan restricted to the chest. The emission acquisition time was 4 min per bed position. All the PET images were reconstructed using a first 3D-ordered subset iterative (OSEM) time of flight reconstruction technique with three iterations and 33 subsets. Images matrix sizes of 144×144 with an isotropic voxel size of 4 mm were used. The emission data were corrected for random coincidences, decay, dead time, scattering and attenuation.

Parametric images

Measurement of blood flow

The tumor blood flow measurement using dynamic PET of first-pass ¹⁸F-FDG uptake is based on the one-compartment first-pass model of Mullani et al [33, 34] and has already been explained in detail in the literature [10, 33, 34].

Several volumes of interest (VOIs) were manually delineated on the early static PET image and the corresponding CT image using a dedicated workstation (Extended Brilliance Workspace 3.5; Philips): a first VOI representing the contours of the primary tumor, a second VOI into the ascending aorta, and a third VOI in the contralateral normal breast. The latter was used as a background reference. All the VOIs were superimposed on the dynamic PET image

and the time–activity curve (TAC) of the ascending aorta was then calculated. Counting rate reached a maximum when the arterial input to the tumor ceased and when the tracer had not left the region of interest (both arterial input and venous drainage were zero). This time T_m corresponds to the end of the first pass of the tracer and was visually determined on the arterial TAC. A parametric BF image (Fig. 1), containing a BF value in each tumor voxel, was then computed with in-house scripts using the vv tools [42], which are based on the Insight Segmentation and Registration Toolkit (ITK) library [51], by using the following equation:

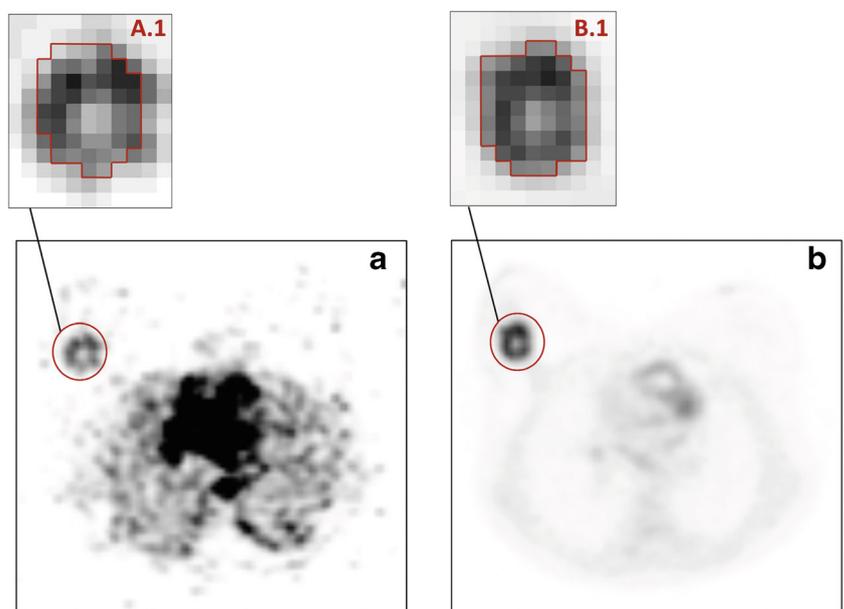
$$BF = \frac{Q(T_m)}{E \int_0^{T_m} Ca(t)dt} \quad (1)$$

where $Q(T_m)$ is the amount of tracer measured in each tumor voxel at time T_m and $Ca(t)$ is the arterial concentration of the tracer measured at time t . The extraction fraction E of ¹⁸F-FDG was assumed to be equal to one [20].

Measurement of glucose metabolism

Semi-automatic segmentation using a contrast-dependent method of metabolic active tumor volumes (MATVs) were performed on the delayed static PET images. This segmentation is based on the determination of an optimal threshold value as a function of the measured contrast in the image. It requires a calibration procedure described in detail by [41]. For this purpose, phantom acquisitions were performed using the NEMA/IEC body phantom with sphere-to-background activity concentration ratios ranging from 40:1 to 2:1. These ratios were chosen to span the whole range of contrast ratios observed in the delayed PET images.

Fig. 1 Parametric blood flow (ml/min/g) image (a) from ¹⁸F-FDG PET first-pass dynamic (10-s frames) acquisition and SUV image (b) from delayed (4-min per bed position) acquisition, for a 63-year-old woman with TN invasive ductal carcinoma of right breast. A.1: tumor VOI from parametric blood flow image after a rescaling using 32 grey levels. B.1: tumor VOI from SUV image after a rescaling using 64 grey levels. Primary tumor BF_{max}: 0.5, BF_{mean}: 0.2, SUV_{max}: 13.5 and SUV_{mean}: 8.1



A series of six VOIs were drawn with a distance of at least 5 mm from the spheres. The mean standardized uptake values (SUV) of all these VOIs was used as the background value. All the spheres were automatically segmented by varying a threshold in steps of 0.01. From the optimal threshold values yielding the best matching with the true spheres' volume ($\pm 2\%$), a regression function was calculated to determine the relation with the normalized thresholds and the contrast ratios, as defined in [41]. Two regression functions were used, a first one for the small volumes (less than 2 ml) and a second one for the larger volumes (more than 2 ml). The adaptive threshold algorithm was computed by using in-house tools based on the vv software. All the necrotic tumor regions were included in the tumors VOIs. Late PET images were then converted to SUV (Fig. 1) by using the classical expression:

$$SUV = \frac{\text{Tumor activity concentration (MBq/ml)}}{\text{Injected dose (MBq) / body weight (g)}} \quad (2)$$

PET-derived features

Standard PET parameters reflecting tumor metabolism and perfusion were extracted from the SUV PET and blood flow parametric images, respectively. The computed features were the average, (SUV_mean and BF_mean for the metabolism and the perfusion, respectively), the standard deviation, (SUV_sd and BF_sd) and the maximum (SUV_max and BF_max) tumor values. In addition, the metabolic active tumor volume and the total lesion glycolysis (TLG = MATV \times SUV_mean) were calculated for the metabolism. All these first-order statistics provide global information on tumor perfusion and metabolism, without any spatial information on the voxel arrangement. On the contrary, TA using second-order statistics can locally describe and quantify the spatial arrangement of the voxels. Several studies have been conducted in order to determine robust and reproducible texture features (TF) from tumor metabolic PET images [18, 27, 49] but, to the best of our knowledge, none have been conducted on the tumor perfusion images from first-pass ^{18}F -FDG PET in BC. Therefore, the same TFs were selected to study both metabolism and perfusion in order to make comparisons easier. Four TFs were used for the analysis of tumor heterogeneity: entropy, inverse different moment (IDM), correlation, and dissimilarity. The entropy is related to the orderliness in the image (how regular the pixel value differences are). This feature quantifies the randomness of the voxel intensity distribution.

$$Entropy = - \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} P(i, j) \ln (P(i, j)) \quad (3)$$

The IDM, also called local homogeneity, quantifies the homogeneity within an image and will be high if the intensities of each pair of voxels are similar.

$$Inverse\ different\ moment = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} \frac{P(i, j)}{1 + (i - j)^2} \quad (4)$$

The correlation measures the linear dependency of grey levels and the dissimilarity quantifies the variation of grey-level voxel pairs.

$$Correlation = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} \frac{P(i, j)(i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} \quad (5)$$

$$Dissimilarity = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} P(i, j) |i - j| \quad (6)$$

These features were computed in 3D using grey-level co-occurrence matrices (GLCM), which take into account pairs of voxels arrangement. For each of the 13 spatial directions, a GLCM matrix was computed between consecutive voxels (i.e., a 1-distance relationship). $P(i, j)$ is the probability value recorded for a pair of voxels in the co-occurrence matrix, μ_i and μ_j correspond to the averages on row i and column j , σ_i and σ_j are the variances on row i and column j and $Nbins$ is the number of grey levels used in the image quantization. The TF values were calculated for each matrix and then averaged to obtain the final TF values.

All the texture parameters were calculated with a fixed number of 64 bins for the metabolism analysis as recommended in the literature [17, 18, 47, 49], and with a number of 32 bins for the perfusion analysis, to minimize the impact of the noise which is higher in parametric blood flow than in metabolic images. Original PET images were rescaled between the minimum and the maximum intensity within the tumor VOI (a relative rescaling) (Fig. 1).

Statistical analysis

Statistical analysis was performed with R [40]. For all the features extracted from the parametric images, average, median, and standard deviation values were calculated. Shapiros' test was used to test the normality of the variable distribution. Since most distributions were not normal, relations between features and phenotypes were assessed using the Kruskal–Wallis test, considering a significance level of 0.05. If statistical significance was found, pairwise comparisons using a Mann–Whitney test with a Bonferroni adjustment was computed to find out differences between phenotypes. Associations between PET features and clinical and physiological parameters were investigated using a

two-sided Mann–Whitney test. Correlations between PET features were evaluated with the Spearman rank test.

Results

Patient characteristics

The study was conducted on 217 patients with a median age of 49 years and an average primary tumor diameter of 34.5 mm. The median MATV was 6.1 ml with an interquartile range of 8.2 ml and the median perfused volume was 4.0 ml with an interquartile range of 7.1 ml. Ductal carcinoma represented 93.6% of the patients, 5.5% were lobular, and 0.9% were not classified. Seventy-nine patients were HER2-positive, 81 luminal/HER2-negative, and 57 had TN tumors. Patient and tumor characteristics are summarized in Table 1.

Correlation between PET features

Figure 2 shows the Spearman correlation heat map of PET features (metabolism and perfusion). Regarding perfusion features, BF.correlation, BF.entropy and BF.IDM were strongly correlated with each other ($r > 0.74$), whereas BF.dissimilarity was inversely correlated with the other perfusion textural features ($r < -0.71$). Except for a weak correlation between BF.max and BF.entropy ($r = 0.61$), BF.max and BF.mean were not correlated with perfusion textural features. Concerning metabolic features, M.correlation, M.IDM and M.entropy were strongly correlated with each other ($r > 0.70$) and with MATV ($r > 0.73$) and TLG, whereas M.dissimilarity was strongly inversely correlated with the other metabolic textural features, MATV, and TLG. SUV.max and SUV.mean were not correlated with the other metabolic features. Regarding the correlation between perfusion and metabolic features, except for a weak correlation between BF.max and TLG ($r = 0.64$), there was no significant correlation between standard metabolism features (SUV.max, SUV.mean, MATV, TLG) and perfusion standard features (BF.max and BF.mean). In contrast, BF.correlation, BF.entropy and BF.IDM were correlated with M.correlation, M.entropy, MATV, and TLG, and inversely correlated with M.dissimilarity. Finally, BF.dissimilarity was correlated with M.dissimilarity and inversely correlated with the other metabolic TFs.

Relationship between PET-derived features and tumor characteristics

Table 2 summarizes the PET-derived features in association with the tumor characteristics. Concerning perfusion parameters, there was no difference according to ER, PR, and HER2 status. In contrast, BF.max, BF.entropy and

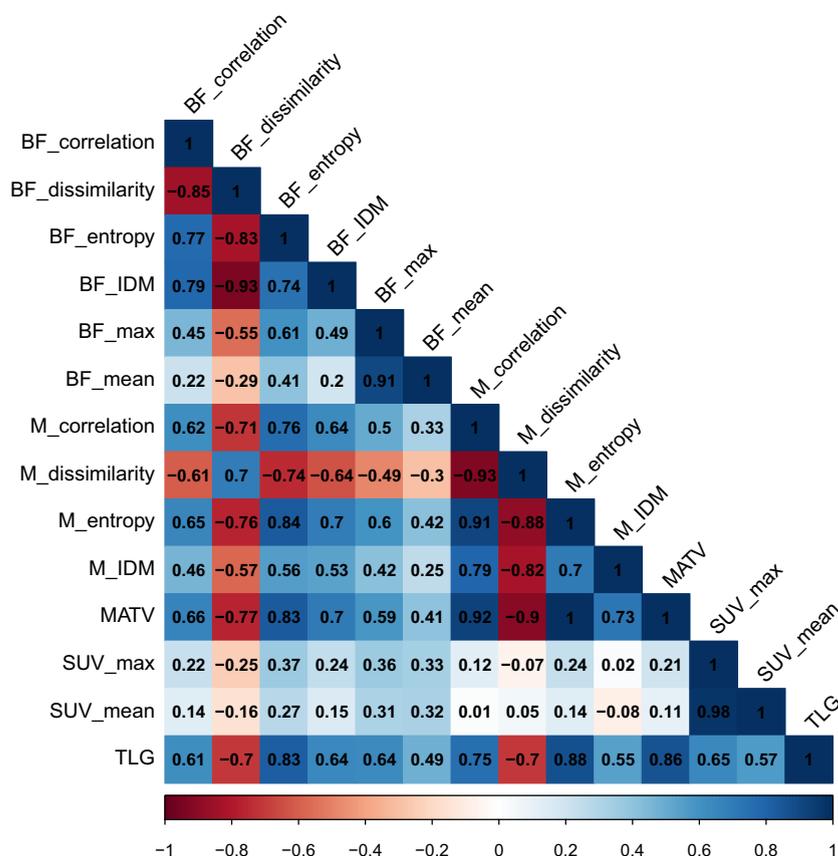
Table 1 Patient and tumor characteristics

	Number of patients (%) <i>N</i> = 217
Menopausal	
Yes	93 (42.9)
No	124 (57.1)
T stage	
T1	23 (10.6)
T2	150 (69.1)
T3	12 (5.5)
T4	32 (14.8)
N stage	
N0	74 (34.1)
N1	104 (47.9)
N2	20 (9.2)
N3	19 (8.8)
SBR	
Grade 1 and 2	115 (53)
Grade 3	102 (47)
Histological type	
Ductal	203 (93.6)
Lobular	12 (5.5)
Others	2 (0.9)
ER status	
Positive	124 (57.1)
Negative	93 (42.9)
PR status	
Positive	91 (41.9)
Negative	126 (58.1)
HER2 status	
Positive	79 (36.4)
Negative	138 (63.6)
Breast cancer sub-group	
ER+/HER2-	81 (37.3)
HER2+	79 (36.4)
TN	57 (26.3)

ER: estrogen receptor; PR: progesterone receptor; TN: triple negative

BF.correlation were significantly higher in the event of T stage 3–4, lymph node involvement or ductal carcinoma, whereas BF.dissimilarity was lower. T stage 3–4 tumors also exhibited higher BF.IDM. Finally, BF.max was higher for SBR grade III. Regarding metabolic features, both SUV.max and SUV.mean were significantly higher in the event of SBR grade III, ductal carcinoma, ER negative, and PR negative. In contrast, no differences were found for SUV.max and SUV.mean concerning HER2 status, T-stage and N-stage. Higher MATV and TLG were found in patients with T-stage 3–4, N⁺ status, and with HER2-status. TLG was also higher in the event of SBR grade III and ductal carcinoma. M.entropy, M.IDM and M.correlation

Fig. 2 Spearman rank correlation heat map between blood flow texture features, metabolism texture features, and standard PET features. The dark blue colors represent strong positive correlations (close to 1), the dark red represent strong inverse correlations (close to -1). The light/white represent an absence of correlation (close to 0). BF: blood flow; IDM: inverse different moment; M: metabolism; MATV: metabolic active tumor volume; SUV: standardized uptake value; TLG: total lesion glycolysis



were higher in patients with T-stage 3–4 and HER2 negative, whereas M_dissimilarity was lower. M_entropy and M_correlation were also higher in patients with N+ status.

Comparison between phenotypes

Table 3, Figs. 3, and 4 summarize the comparison of blood flow and metabolism characteristics between BC phenotypes. Regarding global BF analysis, no significant differences among phenotypes were observed. TF analysis showed no significant differences between HER2 and Luminal tumors. However, TN tumors had a significantly higher entropy than luminal tumors (entropy: TN = 6.18±1.65, luminal = 5.49±1.45, p=0.011), and a significantly higher homogeneity than the HER2 ones (IDM: HER2 = 0.14±0.05, TN = 0.17±0.06, p=0.015; correlation: HER2 = 0.29±0.22, TN = 0.37±0.23, p=0.034). Concerning metabolism analysis, TN tumors showed higher values of SUV_max and SUV_mean than HER2 and luminal tumors. In addition, TF analysis showed that HER2 tumors had a significantly lower homogeneity and entropy than luminal tumors (IDM: HER2 = 0.06±0.02, luminal = 0.07±0.02, p<0.001; entropy: HER2 = 6.22±1.53, luminal = 6.80±1.59, p=0.034). HER2 tumors had also a lower homogeneity and entropy than TN ones (IDM: TN

= 0.07±0.03, entropy: TN = 6.86±1.93, p=0.023 and p=0.047). No significant differences in heterogeneity were reported between TN and luminal tumors.

Discussion

The present paper shows that adding blood flow evaluation, from the first pass of 18F-FDG, to the more conventional delayed 18F-FDG uptake, brings new insights into breast cancer biology. Global blood flow increases in the context of large tumors and axillary lymph node invasion. Texture analysis shows that perfusion heterogeneity parallels that of metabolism. These results suggest a complementary role of advanced textural features and standard PET imaging parameters for the in vivo biological characterization of BC lesions, supporting the hypothesis that advanced PET imaging analysis can provide additional information.

As previously reported, global tumor perfusion was found to be almost identical in the three BC subtypes [20, 46], while glucose uptake significantly increased from luminal to HER2 and TN tumors [9, 10, 15, 16, 20, 32]. The increased metabolism of some tumors, especially the TN subtype, has led to the mismatch concept between perfusion and tumor metabolism [10]. Regarding the

Table 2 Relationship between PET-derived features, histological factors, and tumor characteristics (Mann–Whitney *p* values). The average of the feature values for each group is presented in parentheses when significant

PET features	T stage	N stage	SBR	Histology	Estrogen receptor	Progesterone receptor	HER2 status
	1-2 vs. 3-4	0 vs. 1-3	1-2 vs. 3	ductal vs. lobular	- vs. +	- vs. +	- vs. +
Blood flow	BF_max (ml/min/g)	0.03 (0.33 vs. 0.46)	0.036 (0.33 vs. 0.39)	<0.001 (0.37 vs. 0.18)	0.10	0.18	0.69
	BF_mean (ml/min/g)	0.063	<0.001 (0.13 vs. 0.19)	0.006 (0.18 vs. 0.11)	0.24	0.23	0.79
Entropy	<0.001 (5.55 vs. 6.41)	0.002 (5.27 vs. 5.96)	0.06	0.009 (5.77 vs. 4.62)	0.14	0.18	0.54
	IDM	0.002 (0.15 vs. 0.17)	0.41	0.07	0.85	0.80	0.054
Dissimilarity	<0.001 (7.24 vs. 5.96)	0.020 (7.5 vs. 6.7)	0.21	0.029 (6.91 vs. 8.48)	0.81	0.87	0.22
	Correlation	0.023 (0.29 vs. 0.38)	0.49	0.01 (0.32 vs. 0.16)	0.94	0.96	0.16
Metabolism	SUV_max	0.30	<0.001 (7.7 vs. 13.2)	<0.001 (10.6 vs. 4.6)	<0.001 (12.4 vs. 8.6)	<0.001 (11.7 vs. 8.2)	0.84
	SUV_mean	0.16	0.08	<0.001 (4.7 vs. 7.8)	<0.001 (7.3 vs. 5.3)	<0.001 (7.0 vs. 5.0)	0.71
MATV	<0.001 (9.45 vs. 19.21)	<0.001 (6.8 vs. 13.8)	0.44	0.22	0.22	0.43	0.003 (13.9 vs. 7.2)
	TLG	0.012 (66.5 vs. 127.6)	<0.001 (51.8 vs. 110.0)	0.004 (82.4 vs. 19.4)	0.27	0.17	0.025 (97.8 vs. 45.9)
Entropy	<0.001 (6.38 vs. 7.49)	<0.001 (6.1 vs. 6.9)	0.31	0.18	0.25	0.45	0.004 (6.83 vs. 6.22)
	IDM	<0.001 (0.06 vs. 0.08)	0.37	0.97	0.13	0.39	<0.001 (0.070 vs. 0.056)
Dissimilarity	<0.001 (17.9 vs. 14.6)	0.06	0.89	0.84	0.11	0.25	0.002 (16.5 vs. 18.4)
	Correlation	<0.001 (0.16 vs. 0.32)	0.001 (0.13 vs. 0.22)	0.70	0.082 (0.17 vs. 0.21)	0.15	0.004 (0.22 vs. 0.14)

BF: blood flow; IDM: inverse different moment; MATV: metabolic active tumor volume; SUV: standardized uptake value; TLG: total lesion glycolysis

Table 3 Comparison between blood flow and metabolism characteristics for each breast cancer phenotype

PET features		BC subgroups values			BC subgroups comparison		
		Luminal (mean ± sd)	HER2 (mean ± sd)	TN (mean ± sd)	Luminal-HER2 (p value)	Luminal-TN (p value)	HER2-TN (p value)
Blood flow	BF.max (ml/min/g)	0.32±0.19	0.37±0.29	0.40±0.21	–	–	–
	BF.mean (ml/min/g)	0.17±0.09	0.18±0.12	0.18±0.08	–	–	–
	Entropy	5.49±1.45	5.62±1.47	6.18±1.65	–	0.011	–
	IDM	0.15±0.05	0.14±0.05	0.17±0.06	–	–	0.015
	Dissimilarity	7.07±2.21	7.20±2.41	6.53±2.61	–	–	–
	Correlation	0.29±0.23	0.29±0.22	0.37±0.23	–	–	0.034
Metabolism	SUV_max	8.0±5.0	10.0±5.1	13.8±6.8	0.020	<0.001	0.006
	SUV_mean	4.9±3.0	6.1±3.1	8.0±3.8	0.021	<0.001	0.018
	TLG	56.1±87.0	45.9±61.7	157.2±362.7	–	0.033	0.003
	Entropy	6.80±1.59	6.22±1.53	6.86±1.93	0.034	–	0.047
	IDM	0.07±0.02	0.06±0.02	0.07±0.03	<0.001	–	0.023
	Dissimilarity	16.5±4.0	18.4±3.8	16.6±5.3	0.008	–	–
	Correlation	0.22±0.19	0.14±0.19	0.21±0.25	0.014	–	–

BF: blood flow; IDM: inverse different moment; TLG: total lesion glycolysis

relationship between global perfusion parameters and tumor characteristics, no association between blood flow average (BF.mean) and hormonal receptors or SBR was observed, as previously reported [10, 20, 32]. Although no perfusion difference was highlighted in a previous paper from our institution [20], a significantly higher perfusion was however noted, in the present study, for ductal carcinomas. The larger number of patients here (217 vs. 150) may potentially explain this difference in the results. As there is an important numeric disproportion between the lobular and the ductal tumor population, results must be carefully interpreted. However, the observation of a higher perfusion for ductal carcinomas is in line with some previous studies where the expression of vascular endothelial growth factor (VEGF) protein was significantly higher in invasive ductal carcinomas than in invasive lobular carcinomas [11, 26].

Moreover, in the present study, T3-T4 tumors were found to have a higher BF_max than T1-T2 tumors and this observation has not been reported before. It could be argued that tumor blood flow increases with size but the blood flow “max” is calculated within a small part of a tumor, i.e., a cube of 64 mm³ representing less than 100 mg of tumor. Yet, this observation is in line with a recent study [23] where a significantly lower microvessel density was associated with a smaller tumor size. Fifty percent of tumors with a diameter between 2 and 5 cm had a microvessel density > 72.9/mm² by comparison to only 38% in tumors with a diameter smaller than 2 cm. Moreover, we noted that, in patients with lymph nodes involvement (N⁺), primary tumors also exhibited a

higher BF. Lymph node involvement reflects tumor local invasion and is an important prognostic factor. Similarly, a borderline correlation has been reported [5] between vascular proliferation index and positive nodal status in a series of 110 patients with locally advanced breast cancer. Supporting these observations, a color Doppler ultrasound study showed in a large cohort of breast carcinomas (368 patients) that the peak systolic flow velocity was significantly associated with tumor size and the presence of axillary lymph node metastases [7]. In the same way, a direct association has also been reported between expression of VEGF and the tumor size and the number of metastatic lymph nodes [11].

The differences observed in perfusion depending on tumor size and lymph node involvement might reflect an increased angiogenesis. Indeed, it has been suggested that angiogenesis in breast cancer is regulated by vascular factors secreted by tumor cells and has a close relationship with axillary nodule metastasis [8]. Angiogenesis plays an important role in both tumor growth and metastasis of many types of cancers and is recognized as an important hallmark of oncogenesis [14]. This might provide an explanation to the higher perfusion observed in primary tumors with clinical lymph node involvement, as extensive angiogenesis could facilitate cancer metastasis. These differences in perfusion according to T-stage and N-stage are not paralleled by similar differences in tumor metabolism, as no significant differences were noted for both SUV_max or SUV_mean. This observation might suggest that studying perfusion might provide further insights in breast cancer invasiveness.

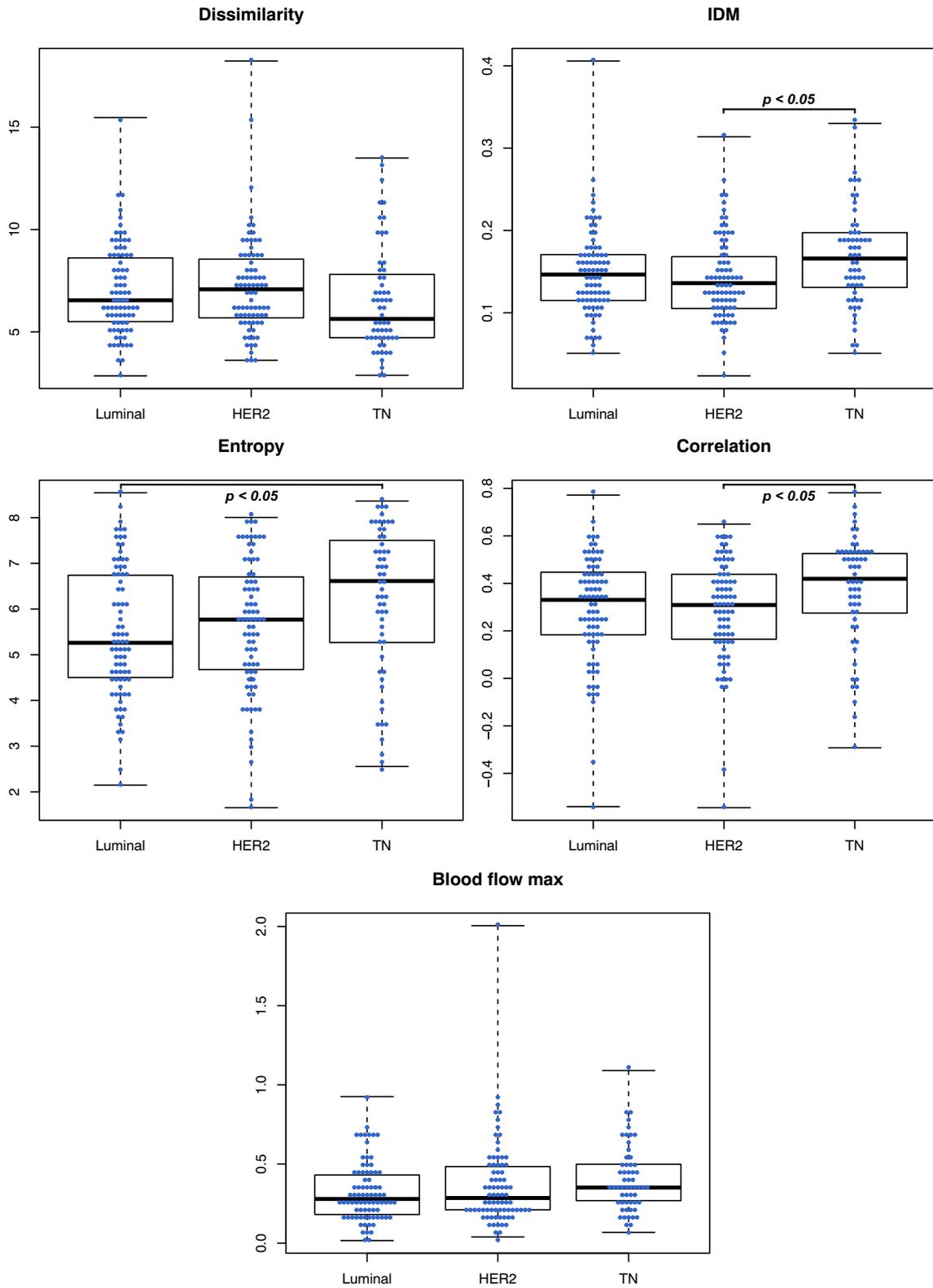


Fig. 3 Distribution of PET features among phenotypes for blood flow. IDM: inverse different moment

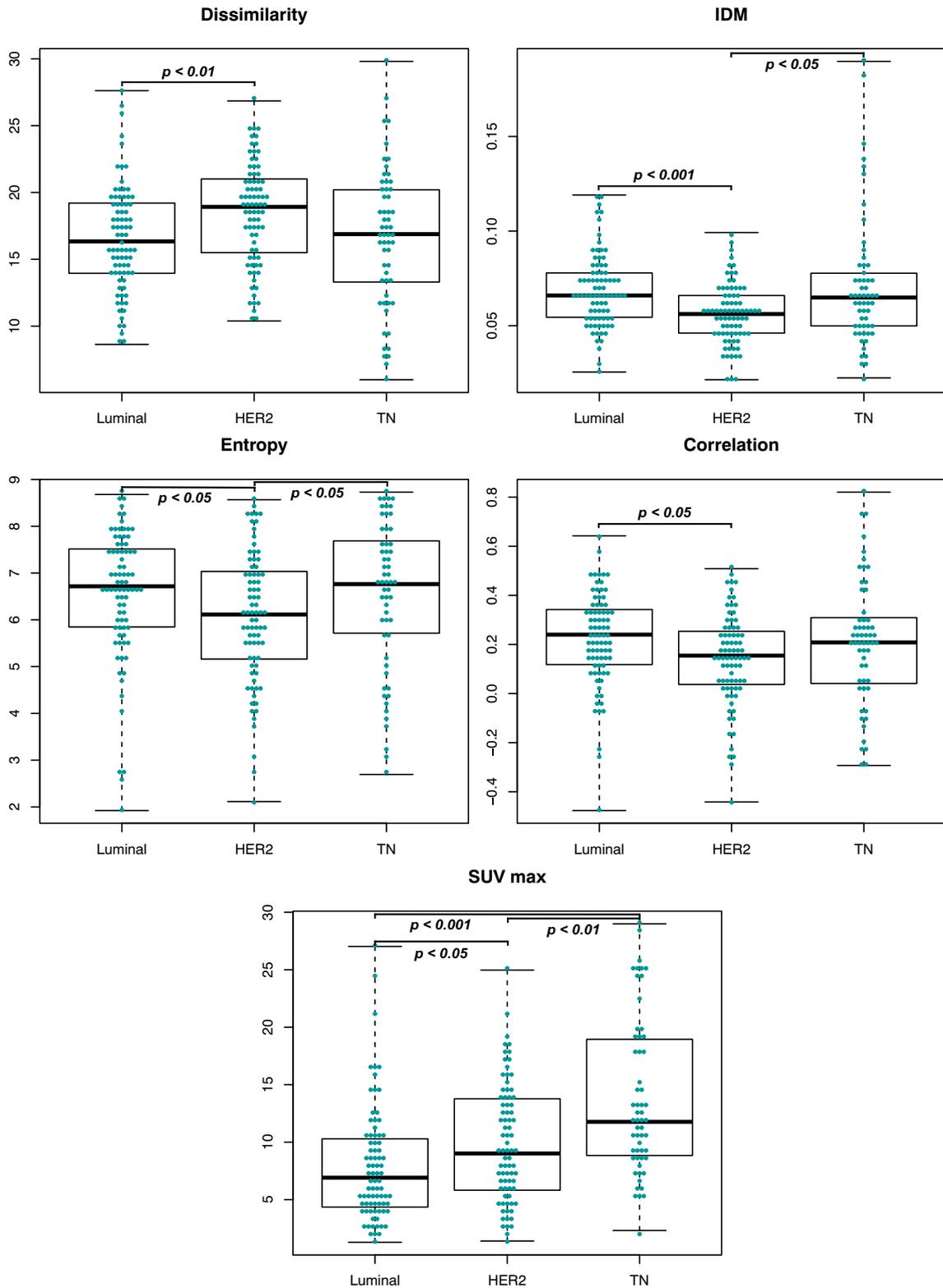


Fig. 4 Distribution of PET features among phenotypes, for metabolism. IDM: inverse different moment; SUV: standardized uptake value

In recent years, several studies have investigated the heterogeneity of ^{18}F -FDG uptake in different tumor types. Some research has focused on breast cancer and various image-derived features, reflecting heterogeneity, have been proposed. Most of these studies have been mainly descriptive, attempting to find correlations with tumor subtypes and treatment responses. In the present study, using a single ^{18}F -FDG PET/CT exam, we were able to characterize both tumor perfusion and metabolism heterogeneity and look for possible links between these two tumor hallmarks' heterogeneity. Texture feature interpretation can be influenced by the method used to calculate it. Depending on the PET image rescaling method [27, 37], it was shown that the texture feature will not have the same behavior, and potentially not the same meaning [27, 28, 36–38]. Bearing this question in mind, a limited number of texture features known to be robust and reproducible [17, 19, 28, 37, 38, 49] were selected in the present study. Moreover, similar texture features were used to study both perfusion and metabolism, thus making comparisons easier.

To the best of our knowledge, no prior study has addressed perfusion heterogeneity using PET/CT in BC. Among a large population covering different BC phenotypes, the differences found between HER2 and TN tumors were significant, with a higher heterogeneity in the HER2 subtype. This difference was also found between HER2 tumors and luminal tumors, but was not significant. Metabolism heterogeneity analysis highlighted similar differences between phenotypes, with HER2 tumors significantly more heterogeneous than luminal and TN tumors. Consequently, HER2 subtype appeared here, as the most heterogeneous phenotype, in both perfusion and metabolism. It is worth noting that perfusion heterogeneity matched metabolism heterogeneity.

One possible interpretation of these facts might be related to the heterogeneous distribution of vascular density within the tumor, due to angiogenesis [1]. Indeed, previous published works underlined the relationship between the HER2-positive tumors and angiogenesis [22, 24, 50]. Moreover, for this BC subtype, clinical trials have suggested that addition of an antiangiogenic treatment to a conventional neoadjuvant chemotherapy, can increase the proportion of patients who achieve a pathological complete response [12], supporting the link between HER2 and angiogenesis.

Finally, unlike global perfusion parameters that show only a limited correlation with standard metabolism features [10, 46], perfusion TFs and metabolism TFs were highly correlated (for example, $r=0.84$ for the correlation between M_entropy and BF_entropy). As the heterogeneity is assessed by comparison of neighbor pixels, this could reflect some anarchic distribution of perfusion and metabolism. One study has investigated the heterogeneity of microvessels in

BC [1] at the microscopic level. The authors concluded that heterogeneity is even more important inside a given tumor than among various types of tumors. Results obtained here may suggest that understanding perfusion heterogeneity at the pixel level could possibly be a step forward in the comprehension of metabolism heterogeneity.

Study limitations

One of the main limitations of this study is the small volume of the tumors. However, small tumor size is one of breast cancer's characteristics and the choice of TFs was done taking into account this aspect, using local TFs only. Considering the wide range of techniques available and the contradictory results, a careful interpretation of heterogeneity should be done. Regarding perfusion and metabolism tumor heterogeneities comparison, we performed a global rather than a voxel-to-voxel comparison given that perfusion and metabolism were evaluated from two different acquisitions.

No correction for partial volume effect was applied and the low count statistics in the dynamic images might add a bias that could affect the results [48]. By improving the signal-to-noise ratio and the spatial resolution of the dynamic short time frames, better image quality should be obtained, leading to new or more significant results. Improvement of spatial resolution and sensitivity provided by new digital PET or dedicated breast PET scanners should improve the analysis in this range of tumor sizes. Moscoso et al. [32] recently investigated the dedicated breast PET resolution improvement. They showed that this technique enables a more accurate heterogeneity characterization, but they focused only on metabolism analysis.

Conclusions

In this study, a significantly higher perfusion in tumors with lymph node involvement has been highlighted, whereas no significant difference was noted in tumor metabolism. In addition, tumor blood flow heterogeneity was different in different breast cancer phenotypes and this was not previously observed with a global analysis of the perfusion. A combination of standard PET parameters and textural features of both tumor perfusion and metabolism could represent a step forward towards a better understanding of tumor behavior and therapeutic strategies adaptation, in the era of personalized medicine. The next step would be to investigate if all these new texture features and their early changes during therapy might improve clinical outcome prediction or treatment response, as potential biomarkers of therapeutic response.

Acknowledgements The authors thank the technologists from the Department of Nuclear Medicine of the Georges-François Leclerc Center for their help in the development of this study, Genevieve Laporte for her help with the clinical database, and Isabel Gregoire for the English revision of the manuscript.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval This study was conducted in accordance with the Declaration of Helsinki and approved by the institutional ethics committee of the Georges-François Leclerc Center.

Informed consent The patient non-opposition was recorded in source documents by the medical team and used as patient informed consent.

References

- Ahlgren J, Risberg B, Villman K, Bergh J. Angiogenesis in invasive breast carcinoma—a prospective study of tumour heterogeneity. *Eur J Cancer*. 2002;38(1):64–69.
- Alic L, Niessen WJ, Veenland JF. Quantification of heterogeneity as a biomarker in tumor imaging: a systematic review. *PLoS One*. 2014;9(10):e110,300.
- Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern*. 1989;19(5):1264–1274.
- Amin MB, Edge SB, Greene FL, Byrd DR, Brookland RK, Washington MK, Gershenwald JE, Compton CC, Hess KR, Sullivan DC, Milburn Jessup J, Brierley JD, Gaspar LE, Schilsky RL, Balch CM, Winchester DP, Asare EA, Madera M, Gress DM, Meyer LR. *AJCC Cancer Staging Manual*. Springer. 2018.
- Arnes JB, Stefansson IM, Straume O, Baak JP, Lønning PE, Foulkes WD, Akslen LA. Vascular proliferation is a prognostic factor in breast cancer. *Breast Cancer Res Treat*. 2012;133(2):501–510.
- Bhargava R, Striebel J, Beriwal S, Flickinger JC, Onisko A, Ahrendt G, Dabbs DJ. Prevalence, morphologic features and proliferation indices of breast carcinoma molecular classes using immunohistochemical surrogate markers. *Int J Clin Exp Pathol*. 2009;2(5):444.
- Chao TC, Luo YF, Chen SC, Chen MF. Color Doppler ultrasound in breast carcinomas: relationship with hormone receptors, DNA ploidy, s-phase fraction, and histopathology. *Ultrasound Med Biol*. 2001;27(3):351–355.
- Choi WWL, Lewis MM, Lawson D, Yin-Goen Q, Birdsong GG, Cotsonis GA, Cohen C, Young AN. Angiogenic and lymphangiogenic microvessel density in breast carcinoma: correlation with clinicopathologic parameters and VEGF-family gene expression. *Mod Pathol*. 2005;18(1):143–152.
- Cochet A, Dygai-Cochet I, Riedinger JM, Humbert O, Berriolo-Riedinger A, Toubeau M, Guiu S, Coutant C, Coudert B, Fumoleau P, Brunotte F. 18F-FDG PET/CT provides powerful prognostic stratification in the primary staging of large breast cancer when compared with conventional explorations. *Eur J Nucl Med Mol Imaging*. 2013;41(3):428–437.
- Cochet A, Pigeonnat S, Khoury B, Vrigneaud JM, Touzery C, Berriolo-Riedinger A, Dygai-Cochet I, Toubeau M, Humbert O, Coudert B, Fumoleau P, Arnould L, Brunotte F. Evaluation of breast tumor blood flow with dynamic first-pass 18F-FDG PET/CT: comparison with angiogenesis markers and prognostic factors. *J Nucl Med*. 2012;53(4):512–520.
- Coradini D, Pellizzaro C, Veneroni S, Ventura L, Daidone MG. Infiltrating ductal and lobular breast carcinomas are characterised by different interrelationships among markers related to angiogenesis and hormone dependence. *Br J Cancer*. 2002;87(10):1105–1111.
- Coudert B, Pierga JY, Mouret-Reynier MA, Kerrou K, Ferrero JM, Petit T, Kerbrat P, Dupré PF, Bachelot T, Gabelle P, Giard S, Coeffic D, Bougnoux P, Prevost JB, Paintaud G, Thibault G, Hernandez J, Coudert M, Arnould L, Berriolo-Riedinger A. Use of [18F]-FDG PET to predict response to neoadjuvant trastuzumab and docetaxel in patients with HER2-positive breast cancer, and addition of bevacizumab to neoadjuvant trastuzumab and docetaxel in [18F]-FDG PET-predicted non-responders (AVATAXHER): an open-label, randomised phase 2 trial. *Lancet Oncol*. 2014;15(13):1493–1502.
- Dunnwald LK, Gralow JR, Ellis GK, Livingston RB, Linden HM, Specht JM, Doot RK, Lawton TJ, Barlow WE, Kurland BF, Schubert EK, Mankoff DA. Tumor metabolism and blood flow changes by positron emission tomography: relation to survival in patients treated with neoadjuvant chemotherapy for locally advanced breast cancer. *J Clin Oncol*. 2008;26(27):4449–4457.
- Folkman J. Tumor angiogenesis: therapeutic implications. *N Engl J Med*. 1971;285(21):1182–1186.
- Groheux D, Giacchetti S, Moretti JL, Porcher R, Espié M, Lehmann-Che J, de Roquancourt A, Hamy AS, Cuvier C, Vercellino L, Hindié E. Correlation of high 18F-FDG uptake to clinical, pathological and biological prognostic factors in breast cancer. *Eur J Nucl Med Mol Imaging*. 2011;38(3):426–435.
- Groheux D, Majdoub M, Tixier F, Le Rest CC, Martineau A, Merlet P, Espié M, de Roquancourt A, Hindié E, Hatt M, Visvikis D. Do clinical, histological or immunohistochemical primary tumour characteristics translate into different (18)F-FDG PET/CT volumetric and heterogeneity features in stage II/III breast cancer. *Eur J Nucl Med Mol Imaging*. 2015;42(11):1682–1691.
- Hatt M, Majdoub M, Vallières M, Tixier F, Le Rest CC, Groheux D, Hindié E, Martineau A, Pradier O, Hustinx R, Perdrisot R, Guillevin R, El Naqa I, Visvikis D. 18F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. *J Nucl Med*. 2015;56(1):38–44.
- Hatt M, Tixier F, Le Rest CC, Pradier O, Visvikis D. Robustness of intratumour 18F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging*. 2013;40(11):1662–1671.
- Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present any future. *Eur J Nucl Med Mol Imaging*. 2017;44(1):151–165.
- Humbert O, Lasserre M, Bertaut A, Fumoleau P, Coutant C, Brunotte F, Cochet A. Breast cancer blood flow and metabolism on dual-acquisition 18F-FDG PET: correlation with tumor phenotype and neoadjuvant chemotherapy response. *J Nucl Med*. 2018;59(7):1035–1041.
- Humbert O, Riedinger JM, Vrigneaud JM, Kanoun S, Dygai-Cochet I, Berriolo-Riedinger A, Toubeau M, Depardon E, Lasserre M, Tisserand S, Fumoleau P, Brunotte F, Cochet A. 18F-FDG PET-derived tumor blood flow changes after 1 cycle of neoadjuvant chemotherapy predicts outcome in triple-negative breast cancer. *J Nucl Med*. 2016;57(11):1707–1712.
- Konecny GE, Meng YG, Untch M, Wang HJ, Bauerfeind I, Epstein M, Stieber P, Vernes JM, Gutierrez J, Hong K, Beryt M, Hepp H, Slamon DJ, Pegram MD. Association between

- HER-2/neu and vascular endothelial growth factor expression predicts clinical outcome in primary breast cancer patients. *Clin Cancer Res.* 2004;10(5):1706–1716.
23. Kraby MR, Opdahl S, Russnes HG, Bofin AM. Microvessel density in breast cancer: the impact of field area on prognostic informativeness. *J Clin Pathol.* 2019.
 24. Kumar R, Yarmand-Bagheri R. The role of HER2 in angiogenesis. *Semin Oncol.* 2001;28(5 Suppl 16):27–32.
 25. Laking G, Price P. Radionuclide imaging of perfusion and hypoxia. *Eur J Nucl Med Mol Imaging.* 2010;37(S1):20–29.
 26. Lee AH, Dublin EA, Bobrow LG, Poulos R. Invasive lobular and invasive ductal carcinoma of the breast show distinct patterns of vascular endothelial growth factor expression and angiogenesis. *J Pathol.* 1998;185(4):394–401.
 27. Leijenaar RTH, Nalbantov G, Carvalho S, van Elmpt WJC, Troost EGC, Boellaard R, Aerts HJWL, Gillies RJ, Lambin P. The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep.* 2015;5(11):075.
 28. Lemarignier C, Martineau A, Teixeira L, Vercellino L, Espié M, Merlet P, Groheux D. Correlation between tumour characteristics, SUV measurements, metabolic tumour volume, TLG and textural features assessed with F-FDG PET in a large cohort of oestrogen receptor-positive breast cancer patients. *Eur J Nucl Med Mol Imaging.* 2017;44(7):1145–1154.
 29. Mankoff DA, Dunnwald LK, Gralow JR, Ellis GK, Charlop A, Lawton TJ, Schubert EK, Tseng J, Livingston RB. Blood flow and metabolism in locally advanced breast cancer: relationship to response to therapy. *J Nucl Med.* 2002;43(4):500–509.
 30. Mankoff DA, Dunnwald LK, Partridge SC, Specht JM. Blood flow-metabolism mismatch: good for the tumor, bad for the patient. *Clin Cancer Res.* 2009;15(17):5294–5296.
 31. Mérida I, Ávila-Flores A. Tumor metabolism: new opportunities for cancer therapy. *Clin Transl Oncol.* 2006;8(10):711–716.
 32. Moscoso A, Ruibal Á, Domínguez-Prado I, Fernández-Ferreiro A, Herranz M, Albaina L, Argibay S, Silva-Rodríguez J, Pardo-Montero J, Aguiar P. Texture analysis of high-resolution dedicated breast F-FDG PET images correlates with immunohistochemical factors and subtype of breast cancer. *Eur J Nucl Med Mol Imaging.* 2018;45(2):196–206.
 33. Mullani NA, Gould KL. First-pass measurements of regional blood flow with external detectors. *J Nucl Med.* 1983;24(7):577–581.
 34. Mullani NA, Herbst RS, O'Neil RG, Gould KL, Barron BJ, Abbruzzese JL. Tumor blood flow measured by PET dynamic imaging of first-pass 18F-FDG uptake: a comparison with 150-labeled water-measured blood flow. *J Nucl Med.* 2008;49(4):517–523.
 35. Neufeld G, Kessler O. Pro-angiogenic cytokines and their role in tumor angiogenesis. *Cancer Metastasis Rev.* 2006;25(3):373–385.
 36. Orhac F, Nioche C, Soussan M, Buvat I. Understanding changes in tumor texture indices in PET: a comparison between visual assessment and index values in simulated and patient data. *J Nucl Med.* 2017;58(3):387–392.
 37. Orhac F, Soussan M, Chouahnia K, Martinod E, Buvat I. 18F-FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer. *PLoS One.* 2015;10(12):e0145063.
 38. Orhac F, Soussan M, Maisonneuve JA, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med.* 2014;55(3):414–422.
 39. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge Ø, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D. Molecular portraits of human breast tumours. *Nature.* 2000;406(6797):747–752.
 40. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2017. <http://www.R-project.org/>.
 41. Schaefer A, Kremp S, Hellwig D, Rube C, Kirsch CM, Nestle U. A contrast-oriented algorithm for FDG-PET-based delineation of tumour volumes for the radiotherapy of lung cancer: derivation from phantom measurements and validation in patient data. *Eur J Nucl Med Mol Imaging.* 2008;35(11):1989–1999.
 42. Seroul P, Sarrut D. VV: A viewer for the evaluation of 4D image registration. MIDAS Journal (Medical Image Computing and Computer-Assisted Intervention MICCAI'2008, Workshop - Systems and Architectures for Computer Assisted Interventions), 1–8. 2008. <http://vv.creatis.insa-lyon.fr/>.
 43. Shi Y, Jin J, Ji W, Guan X. Therapeutic landscape in mutational triple-negative breast cancer. *Mol Cancer.* 2018;17(1):99.
 44. Son SH, Kim DH, Hong CM, Kim CY, Jeong SY, Lee SW, Lee J, Ahn BC. Prognostic implication of intratumoral metabolic heterogeneity in invasive ductal carcinoma of the breast. *BMC Cancer.* 2014;14:585.
 45. Soussan M, Orhac F, Boubaya M, Zelek L, Ziolo M, Eder V, Buvat I. Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer. *PLoS One.* 2014;9(4):e94017.
 46. Specht JM, Kurland BF, Montgomery SK, Dunnwald LK, Doot RK, Gralow JR, Ellis GK, Linden HM, Livingston RB, Allison KH, Schubert EK, Mankoff DA. Tumor metabolism and blood flow as assessed by positron emission tomography varies by tumor subtype in locally advanced breast cancer. *Clin Cancer Res.* 2010;16(10):2803–2810.
 47. Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. *J Nucl Med.* 2012;53(5):693–700.
 48. Walker MD, Asselin MC, Julyan PJ, Feldmann M, Talbot PS, Jones T, Matthews JC. Bias in iterative reconstruction of low-statistics PET data: benefits of a resolution model. *Phys Med Biol.* 2011;56(4):931–949.
 49. Yan J, Chu-Shern JL, Loi HY, Khor LK, Sinha AK, Quek ST, Tham IWK, Townsend D. Impact of image reconstruction settings on texture features in 18F-FDG PET. *J Nucl Med.* 2015;56(11):1667–1673.
 50. Yang W, Klos K, Yang Y, Smith TL, Shi D, Yu D. ErbB2 overexpression correlates with increased expression of vascular endothelial growth factors A, C, and D in human breast carcinoma. *Cancer.* 2002;94(11):2855–61.
 51. Yoo TS, Ackerman MJ, Lorensen WE, Schroeder W, Chalana V, Aylward S, Metaxas D, Whitaker R. Engineering and algorithm design for an image processing API: a technical report on ITK - the insight toolkit. In: Westwood J, editor. *Proceedings of medicine meets virtual reality.* Amsterdam: IOS Press; 2002. P. 586–592.
 52. Zardavas D, Irrthum A, Swanton C, Piccart M. Clinical management of breast cancer heterogeneity. *Nat Rev Clin Oncol.* 2015;12(7):381–394.

EN RÉSUMÉ Dans ce troisième chapitre, plusieurs méthodes pour le calcul des indices de texture ont été étudiées en faisant varier le mode de discrétisation des images, le calcul des GLCM et le nombre de niveaux d'intensités utilisés pour le rééchantillonnage des images de perfusion. L'objectif de cette étude était d'identifier la méthodologie qui serait la plus appropriée pour identifier, sur les données d'imagerie acquises avant le début du traitement néoadjuvant, les patientes ayant eu une réponse histologique complète. Suite à ce travail, il a été décidé d'appliquer pour les calculs des indices de texture, une discrétisation relative (DR) en utilisant 13 GLCM, avec un rééchantillonnage des images de perfusion et de métabolisme de 32 et 64 bins respectivement. Une fois cette méthodologie définie, une seconde étude évaluant les relations entre les paramètres d'hétérogénéité tumorale, aussi bien métaboliques que perfusionnels, avec les caractéristiques biologiques des tumeurs et les sous-types moléculaires a été réalisée. Les résultats ont montré une perfusion significativement plus élevée pour les tumeurs avec atteinte ganglionnaire sans que cela soit observé sur les données de métabolisme. Par ailleurs, une différence significative d'hétérogénéité de perfusion a été notée entre les phénotypes tumoraux sans toutefois qu'une différence globale soit observée. Cette étude a ainsi souligné l'intérêt de l'analyse combinée de l'hétérogénéité de métabolisme et de perfusion pour une meilleure compréhension du comportement tumoral, paramètre clé pour l'adaptation des stratégies thérapeutiques en vue de médecines personnalisées. Pour faire suite à cette analyse, une troisième étude a été réalisée sur l'apport de l'information combinée du métabolisme, de la perfusion tumorale et de leurs hétérogénéités dans la prédiction de la réponse à la chimiothérapie néoadjuvante. Ce travail est présenté dans le chapitre qui suit.

ÉVALUATION DE LA PRÉDICTION DE LA RÉPONSE AU TRAITEMENT

Sommaire

4.1 Des données patients à la prédiction	124
4.1.1 Principe du Machine Learning	124
4.1.2 Préparation et sélection des données d'entrée	126
4.1.3 Algorithmes d'apprentissages	129
4.1.4 Stratégies d'apprentissages	130
4.1.5 Evaluation de l'erreur de prédiction	132
4.2 Prédiction de la réponse à la CNA pour les patientes atteintes d'un cancer du sein	135
4.2.1 Objectif principal	135
4.2.2 Matériel et méthodes	136
4.2.3 Résultats	139
4.2.4 Discussion et conclusion de l'étude	150

Dans le chapitre précédent, nous avons étudié les relations entre les différentes variables caractérisant les tumeurs mammaires et nous avons statué par la même occasion sur une méthode de calcul des paramètres de textures, en accord avec le contexte clinique étudié. Dans ce dernier chapitre nous avons voulu aller au delà de l'analyse descriptive réalisée précédemment en étudiant cette fois ci l'apport combiné du métabolisme et de la perfusion tumorale dans la prédiction de la réponse histologique complète à la CNA. Afin de pouvoir réaliser cette analyse, une préparation et une sélection des données sont nécessaires avant l'utilisation de techniques d'apprentissages automatiques. Dans ce chapitre, nous allons donc dans un premier temps définir les outils et les étapes nécessaires à l'application d'un modèle d'apprentissage, en allant du traitement des données à l'analyse finale de la performance de la prédiction. Dans un second temps, les résultats de l'étude effectuée sur les données cliniques pour la prédiction de la réponse à la CNA sera présentée. Comme pour l'analyse de texture, de nombreuses appellations seront écrites en anglais afin de simplifier le parallèle avec la littérature.

4.1/ DES DONNÉES PATIENTS À LA PRÉDICTION

Cette première partie a pour objectif de présenter le principe de l'apprentissage automatique, le *machine learning* (ML), et de décrire plus spécifiquement les méthodes et les stratégies qui ont été employées pour la prédiction de la réponse à la CNA avec nos données patients.

4.1.1/ PRINCIPE DU MACHINE LEARNING

- Présentation des apprentissages automatiques

Le machine learning peut être défini comme une catégorie de modèles mathématiques et statistiques qui donnent à des ordinateurs la capacité d'apprendre à partir de données d'entrées et d'améliorer leurs performances sans être explicitement programmés pour [446]. Dans la pratique, cela signifie que l'on va donner des données d'entrée (*input*) à un algorithme à partir desquelles ce dernier va modéliser une règle de décision propre à la problématique et aux données étudiées, pour générer une prédiction en sortie (*output*) qui pourra être mise à jour si de nouvelles données sont ajoutées au modèle. Parmi les méthodes de ML on distingue les apprentissages supervisés, semi-supervisés et non supervisés. En apprentissage non supervisé, le nombre de classes, leur nature ou encore les labels des données d'entrée sont inconnus. Dans ce cas là, l'algorithme cherche à déterminer la structure sous-jacente des données sans entraînement initial ou information additionnelle. Dans le cas des apprentissages supervisés, les données d'entrées sont labellisées, c'est à dire que chaque observation est associée à ses caractéristiques de sorties : l'*output* de ces données est donc connu. L'apprentissage passe alors par une phase d'entraînement et une phase de test. Durant la phase d'entraînement, l'objectif est de trouver un modèle qui relie le plus fidèlement possible les données d'input et d'output du jeu de données d'entraînement. La phase de test quant à elle va permettre d'évaluer les performances du modèle entraîné en effectuant une prédiction sur une nouvelle donnée d'entrée. Enfin, l'apprentissage semi-supervisé se trouve à mi-chemin entre l'apprentissage supervisé et l'apprentissage non supervisé, avec un ensemble de données d'entrées dont une partie seulement est labellisée. De manière générale, on distingue deux types de problèmes à résoudre qui dépendent de la nature des variables à prédire :

- Lorsque les variables à prédire sont continues, comme par exemple le temps écoulé avant la rechute d'un patient, on parle alors de problèmes de régression [447].
- Lorsque les variables à prédire sont discrètes, comme par exemple le grade d'une tumeur, on parle alors de classification [448].

Dans la présente étude, le problème à résoudre est un problème supervisé de classification binaire où nous allons prédire la réponse histologique complète au traitement néoadjuvant des patientes. La variable à prédire est dichotomique et sépare les patientes bonnes répondeuses (pCR) des non-répondeuses (non-pCR). Les différents points abordés dans la suite de ce chapitre concernent les apprentissages supervisés de classification.

- Performances et biais

La généralisation d'un modèle correspond à sa capacité, une fois entraîné, à effectuer des pré-

dictions sur des données de test que l'algorithme n'aura jamais vu auparavant. Un modèle de ML est considéré comme bon lorsque ses performances de généralisation sont élevées. Toutefois, un modèle de prédiction n'est jamais parfait et présente toujours des différences plus ou moins grandes entre les valeurs prédites et les valeurs réelles. Un modèle trop simple (de variance faible) ne va pas être capable de s'ajuster aux données occasionnant ainsi un biais sur les performances, y compris sur celles des données d'entraînement. Le risque est alors d'avoir un sous-apprentissage (*underfitting*), où l'on considère que les variables choisies pour décrire les données d'entrées ne sont pas suffisantes pour modéliser leurs relations. En revanche, un modèle trop complexe (de variance élevée) risque le sur-apprentissage (*overfitting*). Le modèle est alors extrêmement performant sur les données d'entraînement mais s'applique mal aux données de test et a donc des performances de généralisation faibles. La difficulté en ML est donc de trouver un compromis entre la façon précise et pertinente de modéliser les données d'entraînement afin qu'elles soient représentatives des caractéristiques des variables à prédire, sans tomber toutefois dans un sur-apprentissage qui peut refléter le bruit présent dans les données, non représentatif du modèle sous-jacent (fluctuations et erreurs de mesures, de labellisation, valeurs extrêmes, etc). On parle alors de compromis biais-variance. L'une des méthodes qui permet de visualiser ce dernier est de générer ce que l'on nomme des courbes d'apprentissage (*learning curve*). Ces dernières vont représenter, en fonction du nombre d'observations, les performances du modèle obtenues sur les données d'entraînement et sur les données de test (ou de validation part.4.1.3) comme illustré figure 4.1. Sur ces courbes, l'écart entre les scores des données d'entraînement et de test doivent être les plus proches et les plus élevés possibles, afin de minimiser la variance. Généralement, l'écart de scores diminue lorsque le nombre d'observations augmente. En cas de sous-apprentissage, il est conseillé de complexifier le modèle en augmentant par exemple le nombre de variables afin de diminuer le biais. Si l'écart entre les scores d'entraînement et de tests sont importants, on est alors en cas de sur-apprentissage. Si le phénomène d'*overfitting* peut partiellement être réduit en augmentant le nombre d'observations, cela reste néanmoins une approche limitée car en pratique les données ne sont généralement pas illimitées, comme c'est le cas lors de cohortes de patients en médecine. Afin de diminuer la variance, d'autres approches sont alors conseillées comme augmenter la régularisation ou réduire la complexité du modèle.

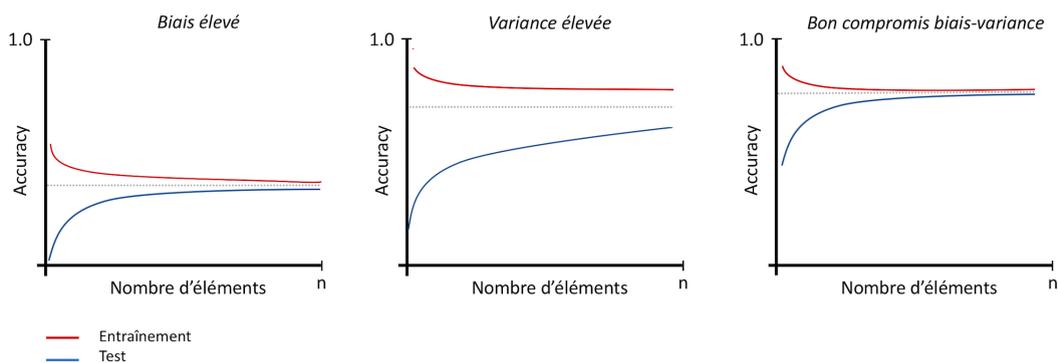


FIGURE 4.1 – Exemple de courbes d'apprentissages qui illustrent le compromis biais-variance

4.1.2/ PRÉPARATION ET SÉLECTION DES DONNÉES D'ENTRÉE

4.1.2.1/ PRÉPARATION DES DONNÉES

La préparation des données d'entrée est une étape importante du processus de ML et contribue à l'efficacité d'un modèle. Les données d'entrée sont souvent des variables de nature différente, elles peuvent être continues, ordinales, dichotomiques, catégorielles ... [449]. Ces variables peuvent être indépendantes ou non, corrélées ou non, parfois avec des données manquantes et sont généralement d'échelles différentes. Deux types de transformations ont été appliquées à nos données en fonction de la nature des variables :

- Variables catégorielles

La plupart des algorithmes (excepté pour les arbres décisionnels par exemple) ne prennent pas en compte directement les variables catégorielles, ces dernières doivent être binarisées avant d'être introduites dans le modèle. Plusieurs cas de figure se présentent. Soit la variable est dichotomique et dans ce cas là, la conversion en entier de valeur 0 ou 1 est directe. C'est le cas par exemple pour la variable qui représente la surexpression des récepteurs HER2 qui est alors encodée avec 1 pour le statut HER2+ et 0 pour le statut HER2- (ou vice versa). Soit la variable catégorielle a plus de deux valeurs possibles mais qui peuvent être regroupées dans deux ensembles distincts, comme par exemple le statut tumoral pour lequel les stades T1 et T2 peuvent être regroupés dans un premier ensemble et les stades T3 et T4 dans un second, séparant ainsi les tumeurs par degré de sévérité. La variable peut alors être encodée comme précédemment. Enfin, si la variable possède plus de deux possibilités qui ne peuvent pas être regroupées en deux sous-ensembles de manière pertinente, comme c'est le cas du phénotype tumoral par exemple, chacune des possibilités doit alors être encodée comme une variable booléenne [450]. Ainsi, si l'on prend l'exemple des trois phénotypes tumoraux, si une tumeur est Luminale, la variable correspondant au statut luminal sera égale à 1, celle du statut HER2 sera égale à 0 et celle de TN à 0 également. La variable "phénotype tumoral" est donc remplacée par trois variables phénotypes (HER2, Luminal et TN). En pratique, puisque l'information d'une des variables peut alors être déduite des deux autres, il est fortement recommandé de ne conserver que l'information de deux variables sur trois afin d'éviter la colinéarité entre ces variables encodées.

- Variables continues

Parmi les variables continues, de grandes disparités de valeurs peuvent être présentes. Par exemple, dans les données patients étudiées, la perfusion tumorale va être comprise entre 0 et 1 mL/min/g, les valeurs de contrast (*inertia*) sont de l'ordre de 10^1 à 10^2 et celles du *cluster shade* oscillent entre -10^4 et 10^4 . Or, de nombreux algorithmes supposent que les distributions des variables sont centrées sur 0 avec une variance de même ordre. Si une variable possède une variance de plusieurs ordres de grandeurs supérieure à celle d'autres variables du modèle, l'algorithme pourrait alors donner un poids bien plus élevé à cette variable faussant ainsi l'apprentissage. Afin d'éviter ce type de situation, les variables doivent donc être normalisées ou

standardisées. Dans le cas de nos données, nous avons appliqué une standardisation (*standard scaling*) de manière à avoir les données centrées autour de 0 avec un écart type de 1 :

$$x = \frac{(x - \mu)}{\sigma} \quad (4.1)$$

avec μ la moyenne de la variable x et σ son écart type.

On peut également souligner que la plupart des algorithmes de ML fonctionnent rarement avec des données manquantes. Or, il est très courant d'en avoir en pratique, comme par exemple en médecine où l'information de certains paramètres biologiques peuvent manquer pour certaines patientes. Dans ce cas là, plusieurs options sont possibles : soit la patiente associée à la donnée manquante est exclue du modèle, soit la donnée manquante est complétée de manière artificielle [451]. Son remplacement peut se faire de plusieurs manières en fonction du type et de la signification de la variable. Pour les variables continues on peut utiliser par exemple la moyenne ou la médiane de la variable sur la population étudiée ou le mode dans le cas de variables catégorielles. Dans le cas de nos données, les variables manquantes ne représentant qu'une très faible proportion de la population totale (7 patientes sur 254, soit moins de 3%), nous avons préféré les exclure. Enfin, il est également possible de corriger les données des valeurs extrêmes considérées comme aberrantes [452]. Toutefois ces dernières peuvent être le reflet de bruit ou peuvent également, dans certains cas, contenir de l'information. Afin d'être sûr de ne pas enlever une potentielle valeur informative, nous avons préféré conserver la totalité des valeurs même si ces dernières auraient pu être considérées comme aberrantes.

4.1.2.2/ SÉLECTION DES DONNÉES

Avant de construire un modèle, il arrive qu'un très grand nombre de paramètres soient disponibles, dépassant parfois même le nombre d'observations. Généralement il est alors conseillé de faire une sélection de variables (feature selection) qui a pour objectifs :

- de simplifier le modèle afin d'améliorer sa généralisation et limiter l'*overfitting* ;
- de faciliter l'interprétation des résultats avec un nombre de variables explicatives réduit ;
- de réduire les potentielles corrélations issues du hasard entre la variable explicative et la variable à prédire ;
- de permettre de réduire le temps ou les ressources nécessaires à l'exécution du modèle.

La sélection de variables a pour finalité de sélectionner les paramètres qui sont les plus représentatifs de la variable à prédire. Pour cela, on recense trois méthodes principales :

- Les méthodes filtres (filters methods) [453]

Ces méthodes sont appliquées avant la phase d'entraînement et sont généralement indépendantes de l'algorithme utilisé [454]. Elles sont reconnues comme robustes vis à vis du sur-apprentissage et simples d'application [455]. Néanmoins, la relation entre la variable à prédire et la variable explicative testée est évaluée de manière individuelle et ne prend donc pas en compte les relations possibles entre les variables explicatives [453]. Le risque est alors de sélectionner des variables apportant une information redondante et non complémentaire.

- Les méthodes enveloppantes (*wrappers methods*) [456]

Contrairement aux méthodes filtres, les *wrappers methods* intègrent au sein du processus de sélection un algorithme d'apprentissage qui va pouvoir évaluer les performances de classification d'un groupe de variables au cours de la sélection. La phase d'apprentissage est réalisée en deux parties : une phase d'entraînement et une phase de validation pour évaluer la performance du groupe de variables testées. Au final, les sous-ensembles de variables explicatives sélectionnées par cette méthode s'avèrent être bien adaptés à l'algorithme de classification utilisé, mais ne seront pas forcément valides si l'on change le classifieur. Parmi ces méthodes, on retrouve la méthode de *Sequential Forward Selection* (SFS) ou *Sequential Floating Forward Selection* (SFFS) [457] où les sous-ensembles de variables sont construits en ajoutant une à une les variables au modèle et en évaluant à chaque fois les performances du nouvel ensemble de paramètres.

- Les méthodes intégrées (*embedded methods*)

Ces méthodes vont incorporer une étape de sélection de variables pendant le processus d'apprentissage. Contrairement aux méthodes *wrappers* l'utilisation d'un jeu de validation n'est pas nécessaire pour tester le sous-ensemble de variables sélectionnées, l'apprentissage se fait sur la totalité de jeu de données disponibles (hormis les observations du test). Les méthodes intégrées étudient l'impact de chaque variable sur le score de prédiction du modèle durant sa construction. Généralement ces méthodes introduisent des pénalisations qui vont conditionner les nombres de variables explicatives sélectionnées. Parmi les méthodes communes, on peut citer Lasso (*Least Absolute Shrinkage and Selection Operator*) [458], Ridge [459] ou encore Elastic Net [460]. L'un des avantages des méthodes intégrées par rapport aux *wrappers* est leur rapidité d'exécution.

La sélection de variables peut être mise en parallèle avec les méthodes de réduction de dimension qui visent également à réduire le nombre de paramètres impliqués dans le modèle prédictif. Toutefois, ces approches diffèrent des méthodes de sélection dans la mesure où les variables initiales sont transformées en nouvelles variables qui sont généralement une combinaison linéaire de ces dernières. Parmi les approches les plus connues, on peut citer l'analyse en composante principale (ACP) ou l'analyse linéaire discriminante (ALD). Toutefois, avec ces approches, les variables finales compliquent généralement la compréhension et l'interprétation du modèle. Dans nos travaux sur la prédiction de la réponse à la CNA nous avons préféré utiliser des méthodes de sélection de variables en testant une méthode filtre basée sur des tests statistiques univariés. Le choix d'appliquer une méthode filtre a été effectué de manière à privilégier dans un premier temps une méthode simple et robuste de sélection de variable, pour pouvoir par la suite augmenter en complexité. Toutefois, seuls les résultats obtenus après une sélection de variable univariée ont pu être présentés dans ce manuscrit. L'application de méthodes de sélection de variables intégrées sont en cours de développement et d'analyse.

4.1.2.3/ GESTION DES DONNÉES DÉSÉQUILIBRÉES

Lors des problèmes de classification, il arrive très souvent que les deux classes de la variable à prédire soient déséquilibrées au sein de la population étudiée. Dans le cas de notre étude par exemple, environ 25% des patientes présentent une réponse histologique complète à la CNA (pCR) contre 75% non-pCR. Ce déséquilibre pouvant être limitant pour certains algorithmes, plusieurs méthodes ont été développées afin de remédier à cette possible limitation. L'une des options consiste à prendre en compte l'asymétrie des classes dans les modèles en appliquant une matrice de coût : le déséquilibre est pondéré en donnant un poids plus élevé aux observations de la classe minoritaire. La seconde option est de rééchantillonner les classes de manière à les équilibrer. La première approche est de sous-échantillonner la classe majoritaire, soit de manière aléatoire, soit en utilisant des sélections un peu plus avancées, par des méthodes de clustering ou par des méthodes creusant les écarts entre les deux classes [143]. Dans tous les cas, le sous-échantillonnage, en fonction du déséquilibre, peut induire une perte considérable de données. La seconde approche consiste donc à sur-échantillonner la classe minoritaire. Deux principales méthodes sont alors possibles : soit les observations de la classe minoritaire sont aléatoirement répliquées, soit les données sont augmentées de manière synthétique. Cette dernière approche est généralement effectuée avec l'algorithme SMOTE (*Synthetic Minority Oversampling Technique*) [461].

4.1.3/ ALGORITHMES D'APPRENTISSAGES

Dans cette partie est décrit l'algorithme de régression logistique qui a été utilisé dans notre étude pour la prédiction de la réponse à la CNA. Les autres algorithmes couramment utilisés en ML seront cités par la suite mais n'ayant pas été utilisés dans cette thèse, il ne seront donc pas présentés en détail.

LA RÉGRESSION LOGISTIQUE La régression logistique est une méthode qui va permettre de prédire une variable dichotomique y à partir de variables descriptives/explicatives $x = (x_1, x_2, \dots, x_n)$. Ces dernières peuvent aussi bien être quantitatives que catégorielles (encodées). En pratique, la probabilité d'un élément d'appartenir à l'une des deux classes de la variable y est modélisée par une fonction logistique, qui présente une forme sigmoïde. Afin de rendre le modèle linéaire, cette fonction sigmoïde est généralement transformée avec la fonction Logit, de manière à attribuer un poids $w = (\beta_1, \beta_2, \dots, \beta_n)$ aux n variables explicatives. L'objectif de la fonction *Logit* est alors de trouver des paramètres β_i qui maximisent la probabilité $p(y|x, w)$ d'un élément d'appartenir à l'une des deux classes. Cette optimisation est réalisée en utilisant le logarithme du maximum de vraisemblance (*LL* pour *Log-Likelihood*) :

$$LL = - \sum_{i=1}^l \log(1 + e^{-y_i \mathbf{w}^T x_i}) \quad (4.2)$$

avec l le nombre d'observations (patients). Cela revient à minimiser la fonction de coût de

la régression logistique qui correspond à la forme négative du LL , soit :

$$\underset{\mathbf{w}}{\operatorname{argmin}} \sum_{i=1}^l \log(1 + e^{-y_i \mathbf{w}^T x_i}) \quad (4.3)$$

Afin de limiter le sur-apprentissage, il est également possible de pénaliser les modèles complexes via une régularisation. Cette dernière permet de minimiser à la fois la fonction de coût et la complexité en ajoutant un facteur de régularisation :

$$\underset{\mathbf{w}}{\operatorname{argmin}} \sum_{i=1}^l \log(1 + e^{-y_i \mathbf{w}^T x_i}) + \lambda R(\mathbf{w}) \quad (4.4)$$

avec $R(w)$ une fonction de régularisation. Cette fonction impose une pénalité qui sera d'autant plus importante que λ est élevé. En pratique, la régression logistique que nous avons appliqué à nos données a été calculée à partir des outils de la bibliothèque Python Scikit-learn [462]. Dans cette dernière, il est possible d'ajouter un terme de régularisation L2 qui va favoriser une variance faible. Ce terme correspond au carré scalaire du vecteur poids w . Afin de faciliter les calculs de gradients, ce terme est également pondéré par un facteur $\frac{1}{2}$. Dans Scikit-learn, la fonction de coût à minimiser est la suivante :

$$\min \frac{1}{2} w^T w + C \sum_{i=1}^l \log(1 + e^{-y_i w^T x_i}) \quad (4.5)$$

On note que le coefficient λ a été remplacé par le paramètre C qui va permettre ici de contrôler le degré de la régularisation. Plus ce dernier est grand, plus le terme de régularisation $\min \frac{1}{2} w^T w$ est négligeable par rapport au terme correspondant à la fonction LL et donc moins le modèle est pénalisé. Ce paramètre est appelé "hyperparamètre" et peut être optimisé durant le processus d'apprentissage.

Le modèle de régression logistique a comme avantage d'être facile à paramétrer et à interpréter et représente une première étape avant de tester des modèles plus complexes. De nombreux autres algorithmes d'apprentissage existent, on peut noter parmi les plus connus les k -plus proches voisins (k -NN pour *K-nearest-neighbour*) [463], les machines à support de vecteur (SVM pour *Support vector machine*) [464], les forêts aléatoires (*Random Forest*) [465, 466] ou encore les réseaux de neurones (ANN pour *Artificial Neural Network*) [467].

4.1.4/ STRATÉGIES D'APPRENTISSAGES

En ML, le modèle doit être évalué en fonction de sa performance de prédiction sur un jeu de données inconnues. C'est pour cela qu'en pratique, si l'on souhaite évaluer le modèle sur des données à disposition, il faut alors "simuler" des données inconnues en scindant en deux l'ensemble total en un jeu d'entraînement, sur lequel le modèle va être construit et entraîné, et un jeu de test, sur lequel la performance dite de généralisation de ce dernier sera évaluée. Il est impératif que les données d'entraînement soit distinctes des données de test sous peine

de biaiser les performances de généralisation du modèle [468]. Ce découpage est effectué de manière à avoir la plupart du temps 10 à 25% des données totales utilisées pour le jeu de test. Par ailleurs, certaines sélection de variables (les *wrappers*) vont nécessiter un jeu de données de validation pour pouvoir être appliquées. De la même manière, certains algorithmes vont avoir des hyper-paramètres à optimiser durant la phase d'entraînement, comme c'est le cas pour la régression logistique. Pour ces cas là, le jeu d'entraînement doit donc être lui aussi scindé en deux parties afin d'isoler un jeu de validation distinct du jeu de test.

Afin d'effectuer ces découpages, il est donc nécessaire d'appliquer ce que l'on nomme une *holdout method*. Cette dernière consiste à séparer de manière aléatoire l'ensemble des données en deux ensembles, un d'entraînement et un de test. L'erreur de prédiction du modèle est alors mesurée sur le jeu de test. Toutefois, afin de s'affranchir d'un découpage aléatoire possiblement avantageux (ou désavantageux), il existe plusieurs approches :

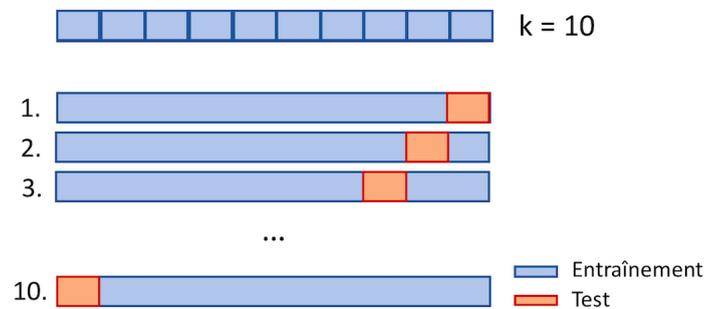
- la validation croisée (CV pour *Cross Validation*)

La validation croisée est une méthode qui permet de découper un ensemble de données afin d'optimiser et d'estimer la fiabilité d'un modèle en évaluant ses performances de généralisation sur des données inconnues. On parle alors de *k-fold cross validation* qui va permettre de découper un ensemble de données N en k échantillons. Un des k échantillons est alors sélectionné comme jeu de données de validation sur lequel est calculé un score de performance et les $k - 1$ autres échantillons constituent le jeu d'entraînement sur lequel le modèle est optimisé et entraîné. On répète ensuite cette opération en sélectionnant un autre échantillon k pour le jeu de validation, et ainsi de suite. Il y a alors autant de cycles d'apprentissages qu'il y a d'échantillons k et les erreurs de prédiction sont moyennées sur la totalité des cycles. Chaque donnée est comprise exactement une fois en tant qu'élément du jeu de validation. Lors du découpage des données en deux ensembles d'entraînement et de test, si une *k-fold CV* est employée, il est commun de choisir un k compris entre 10 et 4 (pour avoir entre 10 et 25% des données dans l'ensemble test). Un cas particulier de la CV est la *leave one-out cross validation* (LOOCV), qui fonctionne comme le *k-fold* avec comme spécificité de n'avoir à chaque fois qu'une seule donnée dans le jeu de validation. k est alors égal à la taille du jeu de données total. Pour les sélections *wrappers* de variables et l'optimisation des modèles, il est fortement conseillé de réaliser une *k-fold CV* (ou une LOOCV).

- les *holdouts* répétés

Il est également possible d'effectuer un *holdout* de manière répétée (avec par exemple 50, 100 ou 200 itérations) toujours avec des découpages aléatoires du jeu de données pour évaluer d'une manière moins biaisée les performances de généralisation des modèles (on parle de *Shuffle splits* ou de *bootstrap* lorsque les observations peuvent être répétées dans un même ensemble). Par ailleurs, la validation croisée peut elle aussi être répétée avec des découpages aléatoires (on parle de *repeated k-fold*). Les résultats de chaque itération sont alors moyennés ou combinés par vote afin d'évaluer les performances des modèles et leurs erreurs de prédictions sont alors moyennées sur les n répétitions, comme c'est le cas en validation croisée.

Enfin, il est important de noter que la découpe de la base de données en un ensemble d'entraînement et un ensemble de test implique que les sélections de variables précédemment décrites

FIGURE 4.2 – Exemple d’une validation croisée avec $k = 10$

soient effectuées uniquement sur le jeu de données d’entraînement tout comme l’optimisation des modèles d’apprentissages. Par ailleurs, lorsque le jeu de données est déséquilibré, il est également conseillé de veiller à ce que les données soient stratifiées durant les découpages en sous-ensembles (entraînement, validation et test) afin de garantir la même proportion représentative des deux classes de la variable à prédire (*stratified k-folds*, *stratified shuffle splits*, etc). Le modèle d’apprentissage général utilisé dans notre étude sera présenté dans la partie méthode 4.2.2.4 .

Remarque : pour ce qui est de l’optimisation des hyperparamètres, ces derniers pouvant être nombreux, il existe un outil qui teste de manière automatique et exhaustive toutes les combinaisons possibles : le grid-search. Le grid search est alors appliqué généralement à l’ensemble des données d’entraînement et il procède en interne à une CV, il n’est ainsi pas nécessaire d’appliquer "à la main" une k-fold CV pour la partie validation.

4.1.5/ EVALUATION DE L’ERREUR DE PRÉDICTION

De nombreuses métriques permettent de mesurer l’erreur de prédiction et d’évaluer les performances des modèles de ML. Dans cette partie vont être décrites les métriques que nous avons utilisées.

- Matrice de confusion

Soit y le vecteur correspondant aux classes de chaque patiente, avec $y = 1$ pour les patientes pCR et $y = 0$ pour les non-pCR. Une méthode relativement simple pour évaluer les performances d’un modèle est de comparer les valeurs réelles de cette variable à prédire y avec les valeurs estimées. La matrice de confusion va répertorier sous forme de matrice les vrais positifs (TP pour True Positive) qui correspondent à des prédictions justes de la classe 1 de y , les vrais négatifs (TN pour Triple Negative) qui correspondent à des prédictions justes de la classe 0 de y , les faux positifs (FP pour False Positive) qui correspondent à des prédictions de classe 1 qui sont réellement de la classe 0 et enfin les faux négatifs (FN pour False Negative) qui correspondent à des prédictions de classe 0 qui sont en réalité de la classe 1 (fig. 4.3). Cette matrice permet à la fois de quantifier l’erreur mais permet également de la situer en indiquant sur quel aspect le modèle se trompe. À partir de cette matrice sont calculés de nombreux

		Valeur réelle	
		+	-
Valeur prédite	+	Vrais Positifs (TP)	Faux Positifs (FP)
	-	Faux Négatifs (FN)	Vrais Négatifs (TN)

FIGURE 4.3 – Matrice de confusion

indices qui sont décrits ci-après.

- L'exactitude (*Accuracy*)

L'exactitude de la mesure décrit la proportion des prédictions correctes du modèle sur le total des prédictions à effectuer. Elle se définit comme suit :

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4.6)$$

Cette métrique très couramment utilisée dans la littérature n'est cependant pas réellement représentative de la performance du modèle lorsque les jeux de données sont déséquilibrés. Si l'on considère un modèle dont la prédiction correspond à la classe majoritaire représentée à 75%, le score minimal d'*accuracy* serait alors de 0,75, ce qui pourrait sembler être un score performant. Or, le score d'*accuracy* est alors plus représentatif de la proportion des classes que de la justesse, on parle du "paradoxe de l'*accuracy*" [469]. L'*accuracy* est une métrique comprise entre 0 et 1 (avec 1 pour une justesse maximale).

- Précision (*precision*), Rappel (*recall*) et *f1 score*

La précision correspond à la proportion des éléments réellement positifs parmi l'ensemble des cas qui ont été estimés positifs par le modèle. La précision est également connue sous le terme de valeur prédictive positive (PPV pour *Predictive Positive Value*) et se calcule de la manière suivante :

$$Precision = \frac{TP}{(TP + FP)} \quad (4.7)$$

Le rappel quant à lui renseigne sur la capacité du modèle à retrouver les vrais positifs. Il est également appelé sensibilité ou taux de vrais positifs (TPR pour *True Positive Rate*) et se calcule de la manière suivante :

$$Rappel = \frac{TP}{(TP + FN)} \quad (4.8)$$

L'évaluation conjointe de la précision et du rappel renseigne de manière assez complète les performances du modèle. Toutefois, ces deux métriques sont généralement sous tension : l'amélioration de l'une se fait au détriment de l'autre. Une des manière de les évaluer conjointement

est l'utilisation du *f1 score*, qui est la moyenne harmonique de la précision et du rappel. Sa formule est la suivante :

$$f_{\beta} = \frac{(1 + \beta^2) \times \text{rappel} \times \text{precision}}{\beta^2 \times \text{precision} + \text{rappel}} \quad (4.9)$$

Le paramètre β peut être adapté en fonction du poids que l'on souhaite attribuer à chacune des métriques. Un $\beta < 1$ accordera plus de valeur à la précision qu'au rappel et un $\beta > 1$ allouera plus de poids au rappel vis à vis de la précision. Prendre un β égal à 1 donnera autant de poids à chacune des métriques et c'est ce qui a été choisi dans notre étude.

Le *f1 score* est conseillé, contrairement à l'*accuracy*, pour l'évaluation des performances des modèles pour lesquels les distributions des données sont déséquilibrées. Le *f1 score* est une métrique comprise entre 0 et 1, tout comme la précision et le rappel. Néanmoins, la principale limite de ces métriques est qu'elles ne considèrent qu'une unique classe, la classe positive. En effet l'information des vrais négatifs n'est jamais exploitée. Si l'on souhaite s'intéresser à l'information des deux classes de la variable à prédire, il est alors également conseillé de calculer le coefficient de corrélation de Matthews (MCC pour *Matthews correlation coefficient*).

- Coefficient de corrélation de Matthews (MCC)

Afin de prendre en compte les quatre informations de la matrice de confusion, il est possible d'utiliser le *Matthews correlation coefficient*. La classe à prédire et la classe réelle sont alors considérées comme deux variables dont on calcule la corrélation : plus cette corrélation est élevée, meilleure est la prédiction. Les valeurs du *Matthews correlation coefficient* sont comprises entre -1 et 1. Lorsque la prédiction est parfaite (FP=FN=0) le MCC aura une valeur de 1. Si aucune prédiction n'est conforme à la réalité, sa valeur sera alors égale à -1. Un MCC de 0 signifie que l'algorithme n'est pas meilleur qu'une prédiction aléatoire. La formule du MCC est la suivante :

$$MCC = \frac{(TP \times TN - FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4.10)$$

Une valeur élevée de MCC signifiera donc que les deux classes sont bien prédites, même si ces dernières sont déséquilibrées.

- La courbe ROC (*Receiver Operating Characteristic curve*)

La courbe ROC va permettre de représenter les performances d'un modèle de classification pour différentes valeurs des seuils associant les probabilités aux prédictions. La courbe ROC trace le taux de vrais positifs en fonction du taux de faux positifs (fig. 4.4) pour différents seuils testés. Une manière de synthétiser l'information de la courbe ROC est d'évaluer son aire sous la courbe (AUC pour *Area Under the Curve*). Un classifieur est considéré comme idéal lorsque le taux de vrais positif est de 100% et donc le taux de faux positifs nul. Ce cas idéal est représenté par la courbe verte qui correspond alors à une AUC de 1. Plus la courbe ROC se rapproche du coin supérieur gauche, plus l'AUC est proche de 1 et meilleure est la performance du classifieur. Enfin, une AUC de 0.5 issue de la ROC représentée par la diagonale

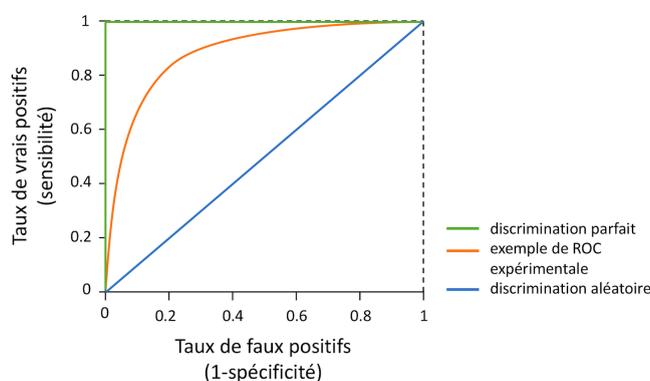


FIGURE 4.4 – Illustration d’une courbe ROC

bleu représente le score d’une classification purement aléatoire et donc non informative. L’*accuracy*, le *f1 score*, le MCC et l’AUC de la courbe ROC représentent les quatre métriques que nous avons utilisées pour estimer la performance de notre modèle de régression logistique.

4.2/ PRÉDICTION DE LA RÉPONSE À LA CNA POUR LES PATIENTES ATTEINTES D’UN CANCER DU SEIN

4.2.1/ OBJECTIF PRINCIPAL

A notre connaissance, aucune étude n’a évalué l’apport conjoint du métabolisme et de la perfusion tumorale évalués par la TEP au ^{18}F -FDG dans la prédiction de la réponse à la CNA. Afin de réaliser cette analyse, nous avons décidé d’effectuer les processus de prédiction en trois étapes :

1. évaluer les performances de prédiction de modèles qui ne disposent uniquement que de données cliniques (C) ;
2. évaluer les performances de prédiction de modèles pour lesquels les données cliniques et les données de métabolisme tumoral sont disponibles (C+M) ;
3. enfin, évaluer les performances de prédiction de modèles pour lesquels la totalité des paramètres à notre disposition sont combinés et utilisés, c’est à dire les données cliniques, les données de métabolisme glucidique et les données de perfusion tumorale (C+M+BF).

Pour chaque modèle concerné, l’étude des paramètres de métabolisme et de perfusion sera réalisée en prenant en compte les indices globaux et les paramètres de texture. Toutes les données sont celles disponibles avant le début de la CNA.

4.2.2/ MATÉRIEL ET MÉTHODES

4.2.2.1/ POPULATION ÉTUDIÉE ET PROTOCOLE D'ACQUISITION DES DONNÉES

La population a déjà été décrite dans les chapitres précédents. Sur les 217 patientes, les données de la réponse histologique sont connues pour 207 d'entre elles. Parmi ces dernières, quatre ont été exclues de l'étude car elles ont reçu des traitements de chimiothérapie adjuvante non conventionnels (Avastin, Taxol Carboplatine). Au total, 204 patientes ont été incluses, dont 50 (24.6%) présentent une réponse histologique complète et 153 (74.4%) sont non-pCR. La variable à prédire présente donc des classes déséquilibrées.

En ce qui concerne le protocole d'acquisition des données, il a été décrit en détail dans les deux chapitres précédents et ne sera donc pas rappelé ici.

4.2.2.2/ DESCRIPTION DES DONNÉES

Les données qui ont été utilisées dans cette études ont été regroupées en 3 catégories :

- Les données cliniques

Les données cliniques regroupent les informations biologiques et histopathologiques ainsi que l'âge des patientes. Mise à part cette dernière variable, tous les autres paramètres sont des variables catégorielles et ont dû être encodés de manière à pouvoir être correctement traitées par l'algorithme de ML. Les variables cliniques étudiées sont les suivantes :

- les récepteurs hormonaux, i.e. les statuts PR, ER et HER2. Ces trois paramètres ont été encodés de manière binaire avec 1 pour leur statut positif, 0 pour leur statut négatif.
- les types histologiques, i.e. lobulaire ou canalaire. Comme pour les récepteurs, cette variable a été encodée directement de manière binaire.
- le statut ménopause (oui/non) : a été directement transposé en réponse binaire.
- la plurifocalité (oui/non) : a été directement transposé en réponse binaire.
- le stade T tumoral : les stades T1 et T2 ont été groupés dans un même ensemble et les stades T3 et T4 dans un second ensemble pour être encodés de manière binaire par la suite.
- le statut ganglionnaire N : tout comme pour le stade tumoral, les données ont été regroupées en deux ensembles afin d'être encodées. Un premier groupe avec uniquement le stade N0 et un second regroupant les stades N1, N2 et N3 ont été réalisés.
- le grade SBR et le grade SBR-mitotique : là aussi deux ensembles ont été créés pour chacune des variables. Un premier groupe comprenant les grades 1 et 2 et un second ensemble pour le grade 3.
- le phénotype tumoral : luminal, HER2 ou TN. Cette variable a été remplacée par trois variables binaires représentant chacune un phénotype.
- le traitement : comprend trois modalités différentes qui ont été encodées en trois variables binaires. Les trois modalités de traitement qui ont été identifiées sont les suivantes :
 - 3 cycles de FEC100 suivis de 3 cycles de taxotère
 - 3 cycles de FEC100 suivis de 3 cycles de taxotère-herceptine

— 6 cycles de taxotère-herceptine

Pour les données ternaires (le type tumorale et le type de traitement), si ces variables sont sélectionnées dans le modèle d'apprentissage, seules deux des trois modalités sont conservées par variable.

- Les données métaboliques

Les données métaboliques comprennent les paramètres globaux : VM, SUV_{\max} , SUV_{mean} , SUV_{sd} , TLG, *skewness* et *kurtosis* ainsi que les 9 indices de textures issus de la GLCM et utilisés dans le chapitre précédent : *entropy*, *energy*, *inertia*, *homogeneity*, *IDM*, *dissimilarity*, *correlation*, *cluster shade* et *cluster prominence*. Ces indices ont été calculés après une discrétisation relative (DR) de l'image de SUV, en utilisant 13 GLCM et avec un rééchantillonnage de 64 bins. Comme il a été mentionné dans le chapitre précédent, le choix des indices de texture a été fait de manière à réaliser une analyse locale de l'hétérogénéité. La totalité des indices étudiés précédemment ont été utilisés dans cette étude afin de laisser la sélection des variables écarter les paramètres non informatifs.

- Les données de perfusion

Tout comme pour le métabolisme, les données de perfusion comprennent les paramètres globaux : VP, BF_{\max} , BF_{mean} , BF_{sd} , *skewness* et *kurtosis* ainsi que les 9 indices de texture : *entropy*, *energy*, *inertia*, *homogeneity*, *IDM*, *dissimilarity*, *correlation*, *cluster shade* et *cluster prominence*. Ces indices ont été calculés après une discrétisation relative (DR) de l'image de perfusion, en utilisant 13 GLCM et avec un rééchantillonnage de 32 bins.

Pour chaque modèle, toutes les données ont été standardisées de manière à avoir les données centrées autour de 0 avec un écart type de 1.

4.2.2.3/ SÉLECTION DES VARIABLES

La méthode de sélection de variable que nous avons utilisé est une méthode filtre qui consiste à appliquer des tests statistiques univariés de Mann-Whitney et de Chi^2 entre les variables explicatives et la variable à prédire y qui, dans notre cas, représente la réponse histologique selon la classification de Chevallier (pCR ou non-pCR). Les tests de Mann-Whitney ont été appliqués aux variables continues et les tests de Chi^2 ont été utilisés pour les variables catégorielles. Pour chacun d'entre eux, plusieurs seuils de significativité ont été testés : 0.05, 0.1, 0.15 et 0.2 comme cela a été fait dans [129]. Une fois les variables significatives sélectionnées, les paramètres corrélés entre eux ont ensuite été exclus afin de limiter les informations redondantes qui pourraient potentiellement pénaliser le modèle. Pour cela, plusieurs seuils de corrélation ont été testés : 0.7, 0.75 et 0.8. Ainsi, lorsque deux variables présentaient une corrélation supérieure au seuil fixé, seule l'une d'entre elles était conservée. Le choix de la variable à garder peut être fait de manière arbitraire. Pour cette étude, nous avons fait le choix de conserver la variable ayant eu la plus petite p-value en sortie des tests univariés. Pour ce qui est des variables trai-

tement et phénotype tumoral, dont l'encodage génère trois variables distinctes, si ces dernières sont sélectionnées, seulement deux des trois modalités sont conservées (part. 4.1.2.1). Le choix a également été fait en relation avec les valeurs des p-values, la variable avec la plus haute p-value parmi les trois étant alors exclue.

4.2.2.4/ STRATÉGIE D'APPRENTISSAGE

La stratégie d'apprentissage et d'évaluation des performances de généralisation a été faite en plusieurs étapes. Un premier *holdout* a été effectué de manière à scinder l'ensemble des données en deux parties : une d'entraînement (80%) et une de test (20%). Ce *holdout* a été réalisé de manière stratifiée. Sur les 80% d'entraînement, les sélections de variables univariées ont ensuite été effectuées en testant à chaque fois tous les paramètres associés (seuil de significativité et coefficients de corrélation). Pour chaque ensemble de variables sélectionnées, une validation croisée sur l'ensemble d'entraînement a par la suite été effectuée afin d'optimiser l'hyperparamètre C du modèle de régression logistique. Cette optimisation a été effectuée de manière exhaustive. Pour chaque étape de validation, la valeur de C permettant d'obtenir le meilleur score de validation est conservée. La métrique de *f1 score* a été utilisée pour cette optimisation car elle est reconnue comme pertinente pour l'évaluation des performances de classification des données déséquilibrées. Une valeur optimale de C a ainsi été associée à chaque ensemble de variables testé. Une fois le modèle optimal identifié (i.e. avec les variables explicatives sélectionnées et la valeur de C déterminée), il est ensuite entraîné sur l'ensemble d'entraînement (80%) puis évalué sur les données du test (20%). Pour chacun des modèles étudiés, les scores d'entraînement, de validation et de test sont conservés. Enfin, le *holdout* avec lequel la base de données a été scindée en deux va être répété 50 fois de manière aléatoire afin de minimiser les possibles biais. Les résultats de tests et d'entraînement sont alors moyennés sur ces 50 itérations. La stratégie d'apprentissage utilisée a été effectuée en se basant sur [470] et est illustrée figure 4.5.

Nous avons également souhaité comparer les scores moyens obtenus aux scores de *dummy classifiers*. Les classes (pCR/non-pCR) de notre variable à prédire étant déséquilibrées, ces derniers ont été utilisés afin de prédire sur l'ensemble de test uniquement la classe majoritaire. La proportion de non-pCR dans l'ensemble de test (et d'entraînement) étant en moyenne de 75%, on s'attend alors à des scores d'*accuracy* pour les *dummy classifiers* de l'ordre de 0,75. En ce qui concerne les autres métriques, les scores devraient correspondre à ceux d'une prédiction aléatoire, c'est à dire à une AUC de ROC de 0,5 et un *f1 score* et un MCC de 0.

Enfin, nous avons également souhaité corriger le déséquilibre des classes en pondérant dans un premier temps les données afin d'appliquer un poids plus élevé aux observations de la classe minoritaire, et dans un second temps en rééchantillonnant les données via SMOTE [461]. Ces corrections ont été effectuées uniquement sur les données d'entraînement et n'ont en aucun cas modifié les données de test. Au final, pour chaque approche C, C+M et C+M+BF, les scores moyens sur les 50 itérations ont été calculés pour tous les modèles, avec et sans correction du déséquilibre des données, puis ont été comparés aux résultats des *dummy classifiers*.

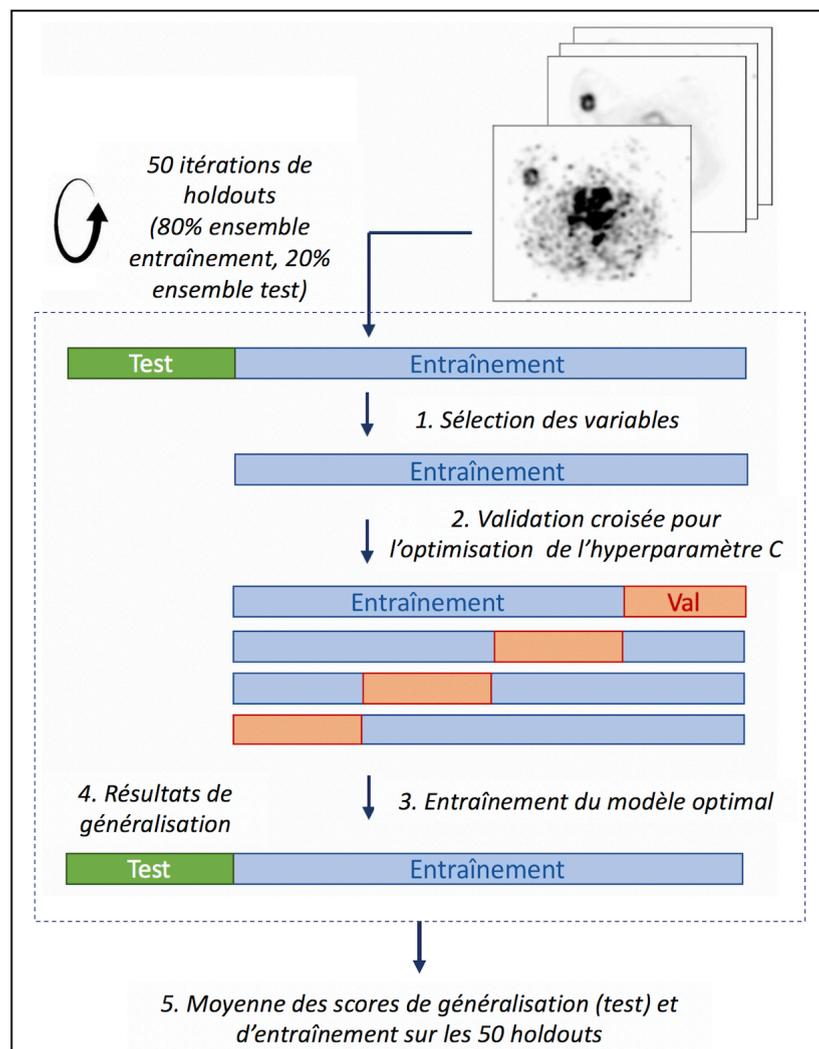


FIGURE 4.5 – Stratégie d'apprentissage

4.2.2.5/ ANALYSE STATISTIQUE

Pour chaque configuration (C, C+M et C+M+BF avec ou sans correction du déséquilibre des classes), les différences entre les meilleurs résultats de généralisation récoltés sur les 50 *holdouts* ont été évaluées à l'aide de tests de Mann-Whitney bilatéraux dont le seuil de significativité a été fixé à 0,05.

4.2.3/ RÉSULTATS

4.2.3.1/ CARACTÉRISTIQUES DE LA COHORTE

L'étude a été menée sur 203 patientes dont l'âge moyen est de 50 ans. Le VM médian est de 6,14 mL avec un intervalle interquartile de 8,2 mL et le volume médian perfusé est de 4,0 mL avec un intervalle interquartile de 7,0 mL. Les carcinomes canauxaires représentent 95,6% des patientes,

3,4% sont lobulaires et 1% ne sont pas classées. Soixante seize (37,4%) des patientes sont de phénotype HER2, 76 (37,5%) de phénotype luminal et 51 (25,1%) de phénotype TN. Parmi la population, 50 (24,6%) ont atteint une réponse histologique complète contre 153 (74,4%) non-pCR. Le jeu des données a été scindé en deux cohortes : une d'entraînement (80%, soit environ 160 patientes) et une de test (20%, soit environ 40 patientes) où la proportion de pCR/non-pCR (approximativement 25%/75%) a été conservée. Les caractéristiques des patients sont résumées dans le tableau 4.1.

Caractéristiques des patientes et des tumeurs		
	pCR (%)	npCR (%)
	50 (24.6)	153 (75.4)
Ménopausée		
Oui	33 (16.3)	83 (40.9)
Non	17 (8.4)	70 (34.5)
T-Stage		
1-2	44 (21.7)	117 (57.6)
3-4	6 (3)	36 (17.7)
N-Stage		
0	13 (6.4)	49 (24.1)
1-3	37 (18.2)	104 (51.2)
SBR		
1-2	17 (8.4)	93 (45.8)
3	33 (16.3)	60 (29.6)
Type histologique		
Canalaire	48 (23.6)	146 (71.9)
Lobulaire	2 (1)	5 (2.5)
Autre	0	2 (1)
Status ER		
Positive	12 (5.9)	106 (52.2)
Négative	38 (18.7)	47 (23.2)
Status PR		
Positive	8 (3.9)	78 (38.4)
Négative	42 (20.7)	75 (36.9)
Status HER2		
Positive	27 (13.3)	49 (24.1)
Négative	23 (11.3)	104 (51.2)
Phénotype tumoral		
ER+/HER2-	2 (1)	74 (36.5)
HER2+	27 (13.3)	49 (24.1)
TN	21 (10.3)	30 (14.8)

TABLE 4.1 – Caractéristiques des patientes et des tumeurs

4.2.3.2/ COURBES D'APPRENTISSAGES ET SÉLECTION DE VARIABLES

Pour chaque approche C, C+M et C+M+BF, les courbes d'apprentissages ont été tracées sur les données d'entraînement, une première fois avec la totalité des variables disponibles, puis une seconde fois uniquement avec les variables sélectionnées par les tests univariés. Ces courbes ont été générées en faisant augmenter progressivement le nombre de patients disponibles (*training samples*) et en effectuant des validations croisées avec, à chaque fois, 75% des données allouées à l'entraînement et 25% à la validation. Cette validation croisée a été faite de manière stratifiée afin de conserver dans chaque ensemble la proportion de chaque classe.

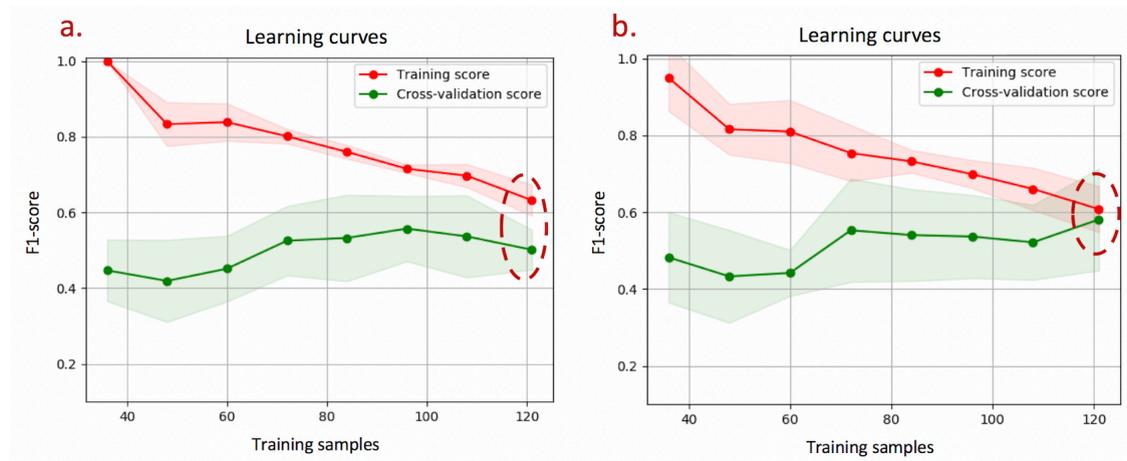


FIGURE 4.6 – Exemple de courbes d'apprentissages réalisées par validations croisées pour le modèle C avant (a) et après (b) sélection de variables, qui représentent les scores d'*accuracy* d'entraînement (rouge) et de validation (vert), pour le quatrième *holdout* effectué parmi les 50.

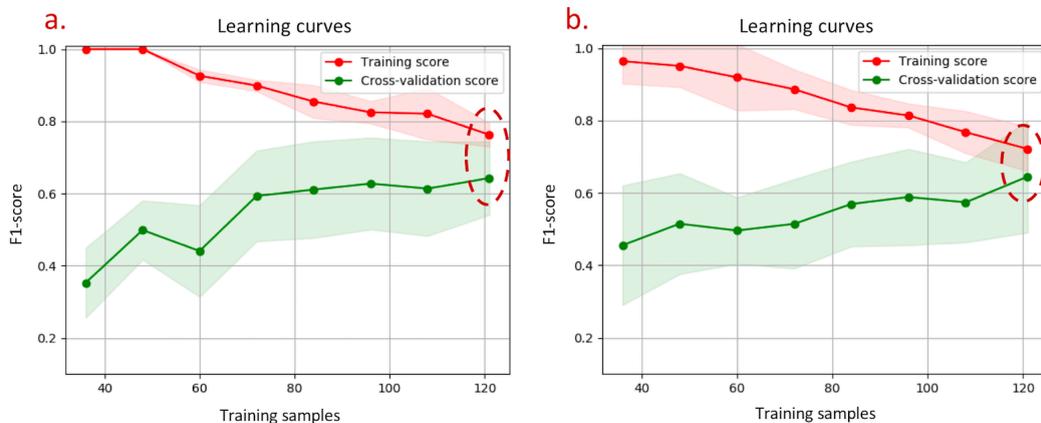


FIGURE 4.7 – Exemple de courbes d'apprentissages réalisées par validations croisées pour le modèle C+M avant (a) et après (b) sélection de variables, qui représentent les scores d'*accuracy* d'entraînement (rouge) et de validation (vert), pour le quatrième *holdout* effectué parmi les 50.

Sans sélection de variables, le modèle C comporte au total 15 variables (après encodage), le modèle C+M est composé lui de 31 variables et le modèle C+M+BF possède 46 variables. Les

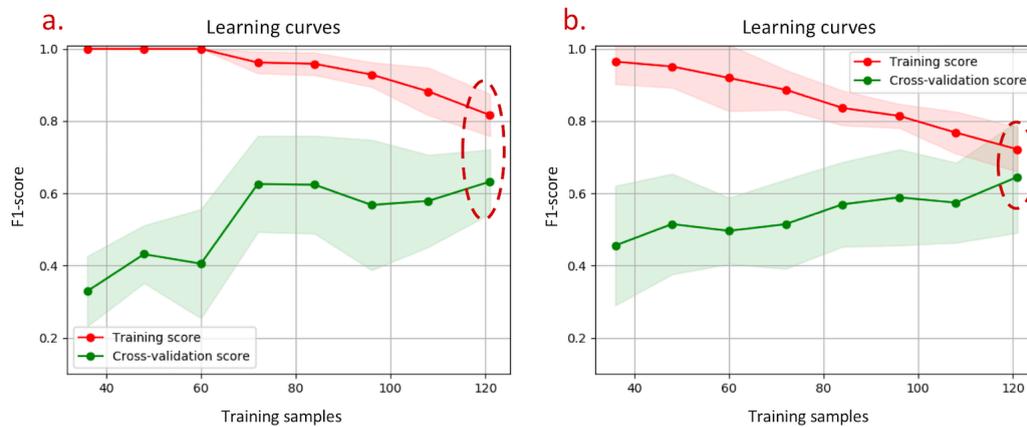


FIGURE 4.8 – Exemple de courbes d’apprentissages réalisées par validations croisées pour le modèle C+M+BF avant (a) et après (b) sélection de variables, qui représentent les scores d’*accuracy* d’entraînement (rouge) et de validation (vert), pour le quatrième *holdout* effectué parmi les 50.

courbes d’apprentissages générées avec la totalité des paramètres à disposition montrent que plus le modèle possède de variables, plus sa variance augmente. En effet, comme il est illustré sur les figures 4.6a, 4.7a et 4.8a, l’écart entre les *f1-scores* d’entraînement et de validation sont plus importants pour le modèle C+M+BF que pour C+M qui a lui même un écart plus grand que le modèle C. On note également que plus le nombre de patients augmente, plus cette variance diminue. Toutefois le nombre de patients étant limité, l’une des options pour améliorer les scores de validation et réduire les variances observées consiste à diminuer le nombre de degré de liberté du modèle. Une des approches alors possible est de réaliser une sélection de variables. En effet, après sélection, nous avons pu observer pour chacun des modèles une amélioration des scores de validation ainsi qu’une réduction de la variance, comme illustré sur les figures 4.6b, 4.7b et 4.8b. Ces courbes d’apprentissages ont ainsi souligné la plus-value d’effectuer une sélection de variable en amont de la prédiction.

4.2.3.3/ SCORES D’ENTRAÎNEMENT ET DE GÉNÉRALISATION DES DIFFÉRENTS MODÈLES

Pour chacune des configurations (C, C+M et C+M+BF) et pour chaque modèle testé (ensemble des variables sélectionnées et hyperparamètre C optimisé), les scores d’entraînements et de tests ont été moyennés sur les 50 *holdouts*. Les résultats finaux de ces scores ont été représentés sous forme de heatmaps pour chaque configuration de sélection de variables, c’est à dire avec en abscisse les seuils alpha des p-values significatives des tests univariés et en ordonnées les valeurs de corrélations de Spearman maximales tolérées entre variables. Plus les seuils alpha et les valeurs de corrélations maximales sont élevés, plus les modèles acceptent un grand nombre de variables. Pour chaque modèle, les paramètres sélectionnés ont été sauvegardés et un *majority voting* (avec un seuil fixé à 50%) a été effectué afin d’identifier celles présentes dans la majorité des *holdouts*. Ces résultats ont été également comparé avec ceux des *dummy classifiers* ainsi qu’avec ceux obtenus après un entraînement réalisé avec une pondération ou

un suréchantillonnage de la classe minoritaire. L'optimisation de l'hyperparamètre C ayant été effectuée en utilisant comme métrique le f1 score, les résultats de généralisation pour tous les modèles sont présentés avec cette même métrique. Pour les meilleurs modèles de chacune des configurations (C, C+M et C+M+BF), les performances d'*accuracy*, d'AUC de ROC et de MCC sont également renseignées.

Sans effectuer de correction sur le déséquilibre des classes durant le processus d'entraînement, on observe qu'en moyenne **les meilleurs résultats sont obtenus pour la configuration C+M, avec un score f1 de généralisation maximal de 0,5** obtenue avec sélection de variable réalisée avec un seuil alpha de 0,1 et une corrélation maximale entre variables de 0,7. Pour cette configuration, les variables sélectionnées dans plus de 50% des *holdouts* sont **le phénotype tumoral, le type de traitement, les récepteurs hormonaux, les grades SBR et SBR mitose, le volume métabolique et deux indices de textures : l'*inertia* et l'*energy*.**

Le meilleur *f1 score* (0,46) observé pour le modèle composé uniquement de variables cliniques est obtenu pour un alpha de 0,2 et une corrélation maximale entre variables de 0,7. Les variables sélectionnées dans plus de 50% des cas sont **le phénotype tumoral, le type de traitement, les récepteurs hormonaux, les grades SBR et SBR mitose et le stade T tumoral.**

Enfin, le meilleur score de généralisation pour le modèle C+M+BF correspondant à un score f1 de 0,49, est noté pour un alpha de 0,05 et une corrélation maximale entre variables de 0,7. Les variables sélectionnées dans la majorité des *holdouts* sont alors les mêmes que pour la meilleure configuration C+M. Le *majority voting* ne rapporte donc aucune variable de perfusion. Les premiers modèles pour lesquels un des paramètres de perfusion est sélectionné dans plus de la moitié des *holdouts* sont ceux obtenus pour des p-values $\geq 0,15$. Les scores f1 de généralisation des différents modèles sont présentés figures 4.9a et les scores d'*accuracy*, d'AUC des ROC et de MCC associés aux meilleurs *f1 scores* sont répertoriés dans le tableau 4.2.

En ce qui concerne les *f1 scores* d'entraînement, on observe que de manière générale, ces derniers augmentent entre les modèles C, C+M et C+M+BF, et sont d'autant plus élevés en moyenne que les seuils des p-values et des coefficients de corrélation augmentent, comme le montre la figure 4.9b. Cela semble cohérent dans la mesure où l'augmentation du nombre de variables favorise l'apprentissage sur l'ensemble d'entraînement. Cela peut néanmoins en contrepartie occasionner un surapprentissage, creusant les écarts entre les scores d'entraînement et de généralisation (comme la tendance pourrait le suggérer pour la configuration C+M+BF).

Si les classes sont équilibrées durant l'entraînement, que ce soit en pondérant ou en suréchantillonnant la classe minoritaire, les résultats de généralisation sont tout le temps supérieurs à ceux observés précédemment (fig. 4.10 et 4.11). On note également que les écart-types des *f1 scores* sont réduits, aussi bien sur les données d'entraînement que de test.

Lorsque la classe minoritaire est pondérée, les meilleurs scores de générali-

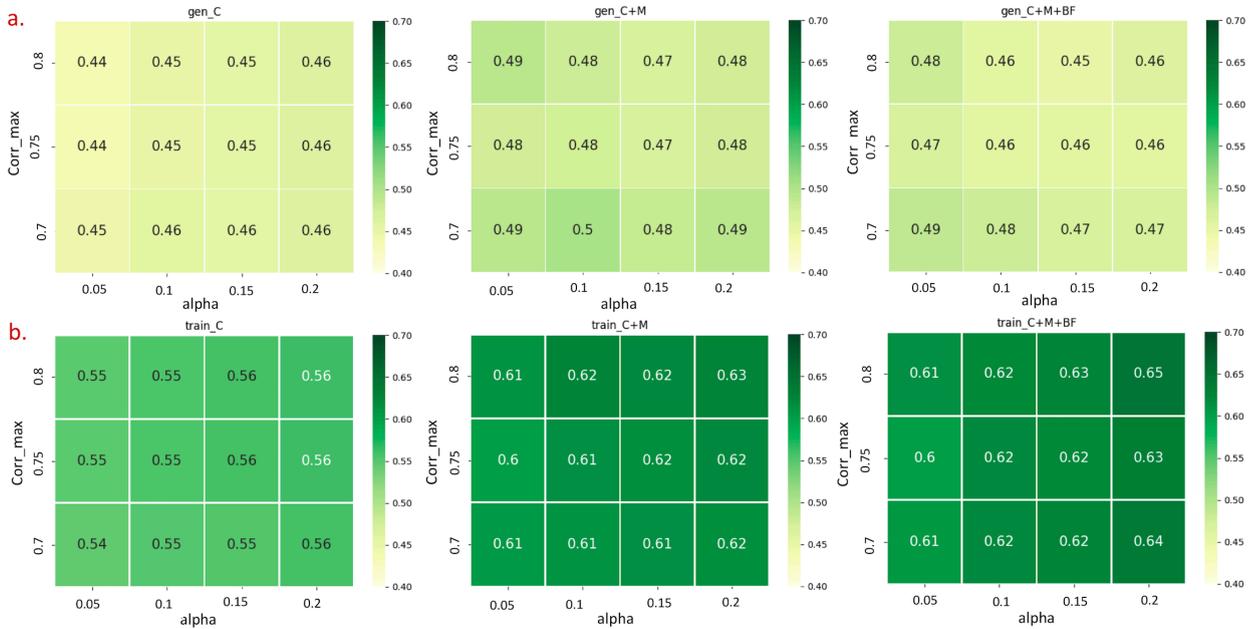


FIGURE 4.9 – Résultats des $f1$ scores de généralisation (a) et d’entraînement (b) moyennés sur les 50 itérations pour les trois modèles, C, C+M et C+M+BF

	C		C + M		C + M + BF		Dummy classifier	
	Entraînement	Test	Entraînement	Test	Entraînement	Test	Entraînement	Test
f1 score	0,56 ($\pm 0,06$)	0,46 ($\pm 0,14$)	0,61 ($\pm 0,05$)	0,50 ($\pm 0,14$)	0,61 ($\pm 0,05$)	0,49 ($\pm 0,14$)	0	0
AUC	0,70 ($\pm 0,04$)	0,65 ($\pm 0,08$)	0,73 ($\pm 0,03$)	0,67 ($\pm 0,08$)	0,73 ($\pm 0,03$)	0,67 ($\pm 0,08$)	0,5	0,5
MCC	0,45 ($\pm 0,06$)	0,34 ($\pm 0,16$)	0,51 ($\pm 0,06$)	0,39 ($\pm 0,16$)	0,50 ($\pm 0,06$)	0,37 ($\pm 0,17$)	0	0
Accuracy	0,81 ($\pm 0,02$)	0,77 ($\pm 0,05$)	0,83 ($\pm 0,02$)	0,79 ($\pm 0,06$)	0,83 ($\pm 0,02$)	0,78 ($\pm 0,06$)	0,75	0,76

TABLE 4.2 – Résultats d’entraînement et de test moyennés sur les 50 *holdouts* pour chacun des trois modèles les plus performants des configurations C, C+M et C+M+BF et résultats du dummy classifier.

sation sont observés pour les configurations C+M et C+M+BF (que l’on notera $C+M_{\text{balanced}}$ et $C+M+BF_{\text{balanced}}$), avec un $f1$ maximal de 0.6 (légèrement supérieur pour $C+M_{\text{balanced}}$ à la troisième décimale), obtenus pour un seuil alpha de 0,05 et une corrélation maximale entre variables de 0,7 (fig. 4.10a). Les variables sélectionnées dans plus de 50% des *holdouts* sont les mêmes dans les deux cas que celles de la configuration C+M étudiée précédemment.

Lorsque la classe minoritaire est sur-échantillonnée, le meilleur score de généralisation est noté pour la configuration C+M (que l’on notera $C+M_{\text{smote}}$), avec un $f1$ score maximal de 0,61 obtenu pour un alpha de 0,05 et une corrélation maximale entre variables de 0,8 (fig. 4.11a). Les variables sélectionnées dans la majorité des *holdouts* sont également les mêmes que pour la configuration C+M avec cependant une variable supplémentaire : le paramètre de texture *homogeneity*.

De manière générale, les résultats de généralisation sont toutefois légèrement meilleurs avec la pondération de la classe minoritaire qu’avec son suréchantillonnage, même si le meilleur

score ($f1\ score = 0.61$) est noté pour $C+M_{smote}$ (fig. 4.10a et 4.11a). Les écart-types des métriques sont quant à eux globalement équivalents entre les deux méthodes de corrections du déséquilibre des données.

Pour toutes les configurations (C, C+M et C+M+BF), les meilleurs modèles pour lesquels le déséquilibre des classes a été corrigé ont à chaque fois des performances de généralisation significativement plus élevées ($p\text{-value} < 0,001$) que les meilleurs modèles où aucune correction n'a été appliquée. Par ailleurs, si l'on compare les modèles optimaux, les différences entre les modèles $C+M_{balanced}$ et $C_{balanced}$ sont proches du seuil de significativité ($p\text{-value} = 0,053$) et le modèle $C+M_{smote}$ est quant à lui, significativement plus performant que le modèle C_{smote} ($p\text{-value} = 0,008$).

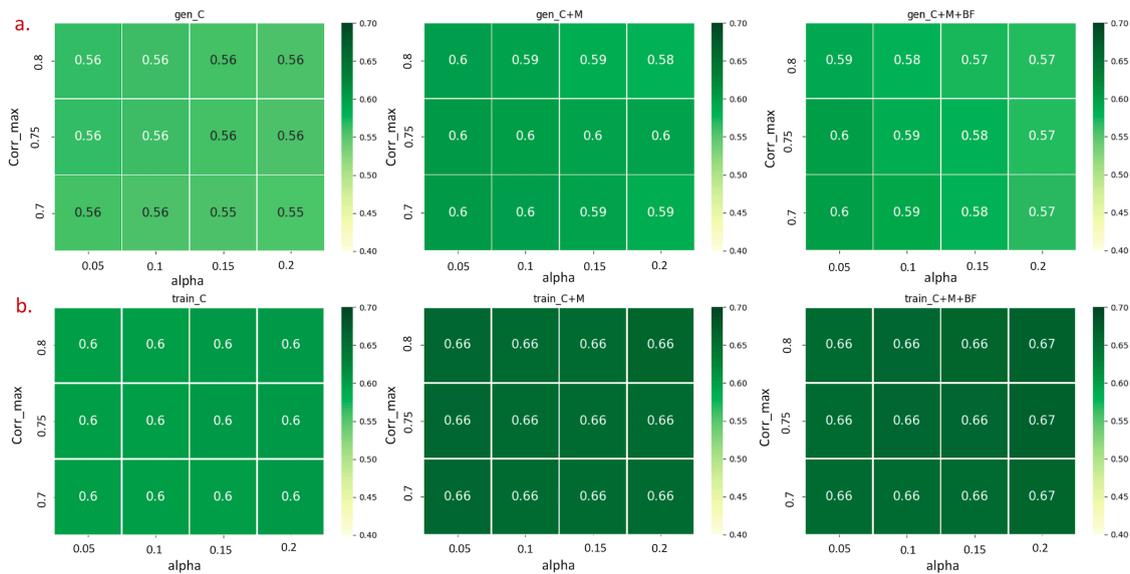


FIGURE 4.10 – Résultats des $f1\ scores$ de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations, effectuées avec une pondération de la classe minoritaire pour l'optimisation des trois modèles, C, C+M et C+M+BF

	C		C + M		C + M + BF	
	Entraînement	Test	Entraînement	Test	Entraînement	Test
f1 score	0,60 ($\pm 0,02$)	0,56 ($\pm 0,09$)	0,66 ($\pm 0,03$)	0,60 ($\pm 0,08$)	0,66 ($\pm 0,03$)	0,60 ($\pm 0,08$)
AUC	0,76 ($\pm 0,02$)	0,73 ($\pm 0,08$)	0,81 ($\pm 0,03$)	0,77 ($\pm 0,07$)	0,81 ($\pm 0,02$)	0,76 ($\pm 0,07$)
MCC	0,46 ($\pm 0,03$)	0,41 ($\pm 0,14$)	0,55 ($\pm 0,05$)	0,47 ($\pm 0,12$)	0,55 ($\pm 0,05$)	0,46 ($\pm 0,13$)
Accuracy	0,72 ($\pm 0,03$)	0,70 ($\pm 0,06$)	0,78 ($\pm 0,03$)	0,73 ($\pm 0,06$)	0,78 ($\pm 0,03$)	0,73 ($\pm 0,07$)

TABLE 4.3 – Résultats d'entraînement et de test moyennés sur les 50 *holdouts* pour chacun des trois modèles les plus performants des configurations C, C+M et C+M+BF, effectués avec une pondération de la classe minoritaire pour l'optimisation des modèles.

Pour ce qui est des résultats moyens d'entraînement, on observe sur les figures 4.10b et

4.11b que les *f1 scores* sont également supérieurs aux scores d'entraînement obtenus sans la correction du déséquilibre des classes (fig. 4.9b). Là encore, de manière générale, les *f1 scores* augmentent entre les configurations C, C+M et C+M+BF (*balanced* et *smote*), et sont d'autant plus élevés que les seuil des p-values et les coefficients de corrélation augmentent. Si l'on compare les deux méthodes de correction, les résultats montrent en moyenne des *f1 scores* supérieurs pour les modèles entraînés avec un suréchantillonnage de la classe minoritaire (figures 66b et 67b). De manière générale, les écarts de scores entre les performances d'entraînement et de généralisation sont réduits lorsque les données sont équilibrées durant la phase d'apprentissage.

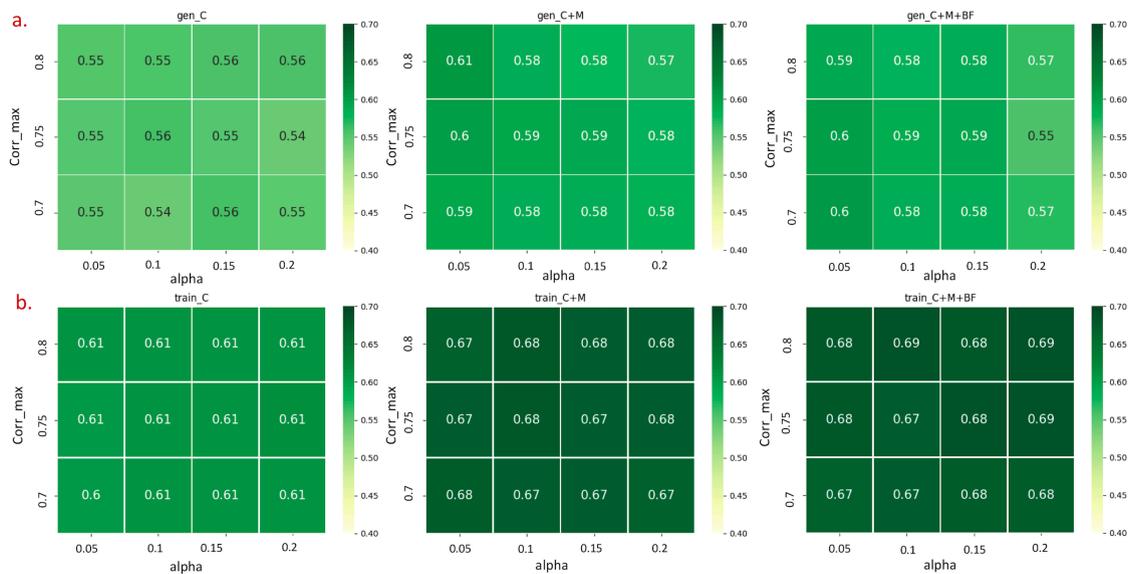


FIGURE 4.11 – Résultats des *f1 scores* de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations, effectués avec un suréchantillonnage par SMOTE de la classe minoritaire pour l'optimisation des trois modèles, C, C+M et C+M+BF

	<i>C</i>		<i>C + M</i>		<i>C + M + BF</i>	
	Entraînement	Test	Entraînement	Test	Entraînement	Test
f1 score	0,61 ($\pm 0,02$)	0,56 ($\pm 0,08$)	0,67 ($\pm 0,03$)	0,61 ($\pm 0,09$)	0,67 ($\pm 0,03$)	0,60 ($\pm 0,09$)
AUC	0,77 ($\pm 0,02$)	0,73 ($\pm 0,07$)	0,81 ($\pm 0,02$)	0,76 ($\pm 0,08$)	0,82 ($\pm 0,02$)	0,76 ($\pm 0,08$)
MCC	0,47 ($\pm 0,03$)	0,4 ($\pm 0,12$)	0,56 ($\pm 0,05$)	0,47 ($\pm 0,13$)	0,56 ($\pm 0,05$)	0,46 ($\pm 0,14$)
Accuracy	0,73 ($\pm 0,03$)	0,70 ($\pm 0,06$)	0,79 ($\pm 0,03$)	0,75 ($\pm 0,07$)	0,79 ($\pm 0,03$)	0,74 ($\pm 0,07$)

TABLE 4.4 – Résultats d'entraînement et de test moyennés sur les 50 *holdouts* pour chacun des trois modèles les plus performants des configurations C, C+M et C+M+BF, effectués avec un suréchantillonnage par SMOTE de la classe minoritaire pour l'optimisation des modèles.

En ce qui concerne les performances des *dummy classifiers*, les scores f1 moyennés et de MCC sont tous nuls et les moyennes des AUC des ROC sont de 0,5. Ces résultats correspondent donc aux scores d'une classification purement aléatoire. Les valeurs d'*accuracy* sont quant à elles de 0,76, ce qui est représentatif de la proportion de la classe majoritaire. Les scores des *dummy classifiers* sont renseignés dans le tableau 4.2.

4.2.3.4/ CONTRIBUTION DES VARIABLES DE TEXTURE

La valeur ajoutée des paramètres de texture a ensuite été évaluée pour les configurations C+M et C+M+BF pour lesquelles les 50 itérations d'apprentissages et de tests ont été réitérées sans permettre aux modèles de sélectionner les indices de textures. **Les résultats montrent que les performances ont globalement diminué lorsque ces derniers sont indisponibles, mais pas de manière significative** (p-value = 0.28 et 0.48 pour C+M et C+M+BF respectivement). **Les meilleurs $f1$ scores passent de 0,5 à 0,47 et de 0,49 à 0,47 pour les configurations C+M et C+M+BF respectivement.** Le meilleur modèle est noté pour un alpha de 0,15 et une corrélation maximale de 0,7 pour la configuration C+M. Les variables sélectionnées dans plus de 50% des *holdouts* sont le volume métabolique, le SUV_{max} , le phénotype tumoral, le type de traitement, les récepteurs hormonaux, les grades SBR et SBR mitose et le stade T tumoral. En ce qui concerne la configuration C+M+BF, le meilleur score moyen de $f1$ est observé pour un alpha de 0,1 et une corrélation maximale de 0,7. Les variables sélectionnées dans la majorité des *holdouts* sont le volume métabolique, le phénotype tumoral, le type de traitement, les récepteurs hormonaux ainsi que les grades SBR et SBR mitose. La différence entre les variables sélectionnées pour les configurations C+M et C+M+BF vient essentiellement de la valeur du seuil de significativité des tests univariés. En effet, pour la configuration C+M+BF, avec un alpha de 0,15 et une corrélation maximale de 0,7, on retrouve exactement les mêmes variables présentes dans plus de 50% des *holdouts* que celles de la configuration C+M.

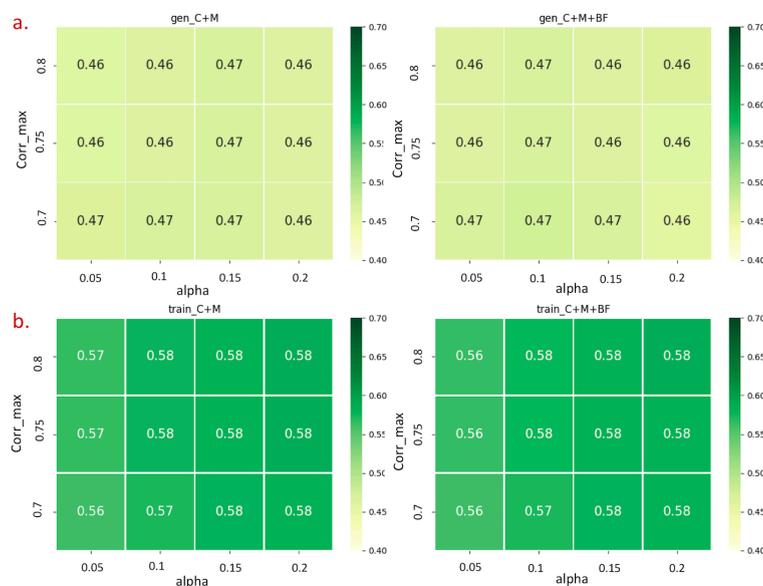


FIGURE 4.12 – Résultats des $f1$ scores de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations, pour les modèles C+M et C+M+BF sans variable de texture

Lorsque les données sont équilibrées durant la phase d'entraînement, les scores de généralisation sont alors égaux ou inférieurs à ceux obtenus avec les indices de texture inclus (fig. 4.10 et 4.11 et fig. 4.13 et 4.14). Les meilleurs scores de généralisation sont obtenus pour des seuils alpha de 0,05 et pour des coefficients de corrélation allant de 0,7 à 0,8. Les variables sélection-

	<i>C + M</i>		<i>C + M + BF</i>	
	Entraînement	Test	Entraînement	Test
f1 score	0,58 ($\pm 0,05$)	0,47 ($\pm 0,14$)	0,57 ($\pm 0,05$)	0,47 ($\pm 0,14$)
AUC	0,71 ($\pm 0,03$)	0,66 ($\pm 0,07$)	0,71 ($\pm 0,03$)	0,66 ($\pm 0,07$)
MCC	0,48 ($\pm 0,06$)	0,35 ($\pm 0,16$)	0,47 ($\pm 0,06$)	0,36 ($\pm 0,16$)
Accuracy	0,82 ($\pm 0,02$)	0,77 ($\pm 0,05$)	0,82 ($\pm 0,02$)	0,78 ($\pm 0,05$)

TABLE 4.5 – Résultats d’entraînement et de test moyennés sur les 50 *holdouts* pour les deux modèles les plus performants des configurations C+M et C+M+BF effectués sans les variables de textures.

nées dans la majorité des *holdouts* sont alors les mêmes pour tous ces modèles et comprennent le volume métabolique, le phénotype tumoral, le type de traitement, les récepteurs hormonaux et les grades SBR et SBR mitose.

Les scores d’entraînement sont quant à eux toujours inférieurs à ceux obtenus pour les modèles impliquant des indices de texture.

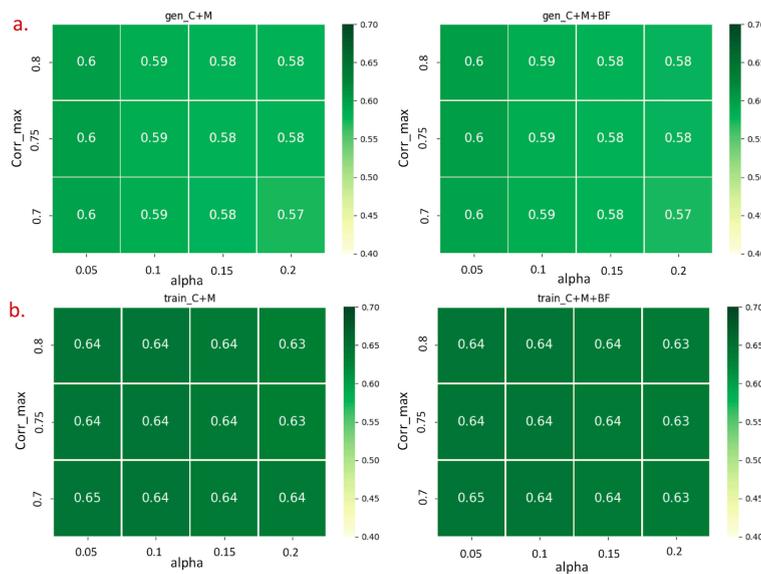


FIGURE 4.13 – Résultats des $f1$ scores de généralisation (a) et d’entraînement (b) moyennés sur les 50 itérations, effectuées avec une pondération de la classe minoritaire pour l’entraînement des deux modèles C+M et C+M+BF sans variable de texture

4.2.3.5/ RÉCAPITULATIF DES RÉSULTATS D’APPRENTISSAGE

Les modèles C+M présentent donc de manière générale de meilleurs scores de généralisation que les modèles C et C+M+BF. Les différences de résultats sont non significatives lorsque le déséquilibre des classes n’est pas corrigé, mais le deviennent lorsqu’une correction via SMOTE est appliquée (p -value = 0,008). On note également que les corrections du déséquilibre des classes durant la phase d’entraînement permettent d’améliorer significativement les scores de

	<i>C + M</i>		<i>C + M + BF</i>	
	Entraînement	Test	Entraînement	Test
f1 score	0,64 ($\pm 0,03$)	0,60 ($\pm 0,08$)	0,64 ($\pm 0,03$)	0,60 ($\pm 0,08$)
AUC	0,80 ($\pm 0,03$)	0,77 ($\pm 0,07$)	0,80 ($\pm 0,03$)	0,77 ($\pm 0,07$)
MCC	0,52 ($\pm 0,05$)	0,46 ($\pm 0,13$)	0,52 ($\pm 0,05$)	0,47 ($\pm 0,12$)
Accuracy	0,76 ($\pm 0,03$)	0,73 ($\pm 0,06$)	0,76 ($\pm 0,03$)	0,73 ($\pm 0,06$)

TABLE 4.6 – Résultats d’entraînement et de tests moyennés sur les 50 *holdouts* pour les deux modèles les plus performants des configurations C+M et C+M+BF, effectués sans variables de textures et avec une pondération de la classe minoritaire pour l’optimisation des modèles.

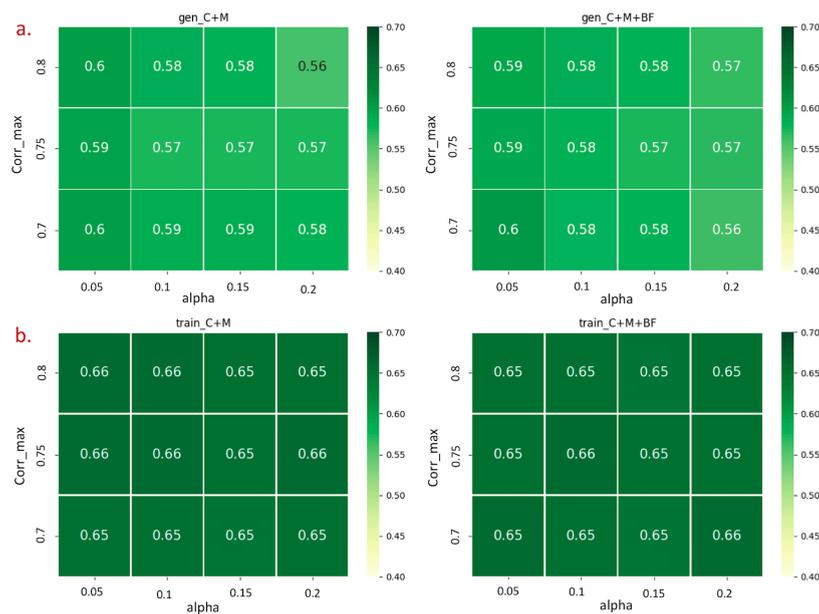


FIGURE 4.14 – Résultats des *f1 scores* de généralisation (a) et d’entraînement (b) moyennés sur les 50 itérations, effectuées avec un suréchantillonnage par SMOTE de la classe minoritaire pour l’entraînement des deux modèles C+M et C+M+BF sans variable de texture.

	<i>C + M</i>		<i>C + M + BF</i>	
	Entraînement	Test	Entraînement	Test
f1 score	0,65 ($\pm 0,02$)	0,60 ($\pm 0,1$)	0,65 ($\pm 0,03$)	0,60 ($\pm 0,1$)
AUC	0,80 ($\pm 0,02$)	0,76 ($\pm 0,08$)	0,80 ($\pm 0,02$)	0,76 ($\pm 0,08$)
MCC	0,53 ($\pm 0,03$)	0,46 ($\pm 0,15$)	0,53 ($\pm 0,04$)	0,47 ($\pm 0,15$)
Accuracy	0,77 ($\pm 0,02$)	0,75 ($\pm 0,07$)	0,78 ($\pm 0,02$)	0,75 ($\pm 0,07$)

TABLE 4.7 – Résultats d’entraînement et de tests moyennés sur les 50 *holdouts* pour les deux modèles les plus performants des configurations C+M et C+M+BF, effectués sans variables de textures et avec un suréchantillonnage par SMOTE de la classe minoritaire pour l’optimisation des modèles.

généralisation des modèles optimaux pour chacune des configurations (p-values $< 0,001$).

L'utilisation des paramètres de texture permet quant à elle d'améliorer les performances de généralisation des modèles lorsque ces derniers ont été évalués sans équilibrer les données d'apprentissages, mais pas de manière significative. Lorsqu'une correction du déséquilibre des classes est effectué, les performances de généralisation des modèles sans indices de texture obtiennent alors des scores très proches de ceux des modèles qui les incluent.

Enfin, les résultats de chaque configuration présentent des scores de généralisation meilleurs que ceux obtenus avec les *dummy classifiers*.

4.2.4/ DISCUSSION ET CONCLUSION DE L'ÉTUDE

L'objectif principal de cette étude était d'évaluer l'apport combiné de la perfusion et du métabolisme dans la prédiction de la réponse à la CNA. Pour se faire, nous avons réalisé plusieurs prédictions en ne disposant dans un premier temps que des variables cliniques, puis des variables cliniques et métaboliques combinées pour enfin ajouter aux données les variables de perfusion et observer si les résultats de prédiction en étaient améliorés. **Les données qui ont été utilisées sont uniquement celle disponibles avant le début de la CNA.**

4.2.4.1/ PERFORMANCES DE PRÉDICTION ET APPORT DES DONNÉES DE PERFUSION

De manière générale, les résultats montrent que les modèles utilisant les données cliniques et métaboliques, avec les paramètres de texture associés, présentent les meilleurs scores de généralisation. Lorsqu'aucune correction du déséquilibre des classes n'est effectuée, on observe un *f1 score* maximal de 0,5 ($\pm 0,14$), une AUC de la ROC de 0,67 ($\pm 0,08$), un MCC de 0,39 ($\pm 0,16$) et une *accuracy* de 0,79 ($\pm 0,06$). Lorsque les classes ont été équilibrées durant la phase d'entraînement, on note un *f1 score* maximal de 0,61 ($\pm 0,09$), une AUC de la ROC de 0,76 ($\pm 0,08$), un MCC de 0,47 ($\pm 0,13$) et une *accuracy* de 0,75 ($\pm 0,07$). Ces derniers résultats sont alors significativement meilleurs que ceux des modèles utilisant les données cliniques seules ou combinées aux données de perfusion et de métabolisme. Pour toutes les configurations, les variables qui sont sélectionnées dans plus de 50% des cas sont le phénotype tumoral, le type de traitement, les récepteurs hormonaux, les grades SBR et SBR mitose, le volume métabolique et des indices de texture (*inertia*, *energy* et *homogeneity*). Par ailleurs, les modèles étudiés permettent d'obtenir des résultats de généralisation de *f1 score*, d'AUC, d'*accuracy* et de MCC toujours supérieurs à ceux des *dummy classifiers*, et ce, malgré la proportion très élevée dans la population générale des non-pCR (75%).

Néanmoins, on remarque que l'ajout des données de perfusion ne semble pas améliorer les scores moyens de prédiction, et semble même pénaliser les modèles au vu des scores de généralisation plus faibles que ceux observés pour la configuration C+M. Toutefois, pour des seuils alpha et des corrélations maximales faibles, les modèles C+M+BF présentent des résultats qui sont parfois égaux à ceux des modèles C+M. Or dans ces cas là, aucune variable de perfusion n'apparaît dans plus de 50% des *holdouts*. La sélection de variable étant plus restrictive pour ces valeurs de corrélations et de seuils, elle ne permet pas, *a priori*, la sélection de paramètres de perfusion. La contribution de cette dernière serait donc majoritairement absente des modèles, ce qui

expliquerait les égalités dans les résultats. De même, lorsque ces derniers sont légèrement plus faibles pour les modèles C+M+BF, on peut alors supposer que les itérations dans lesquelles des paramètres de perfusion ont quand même été sélectionnés occasionnent sur ces mêmes *holdouts*, des résultats de généralisation plus faibles que les modèles utilisant uniquement les données cliniques et métaboliques. Par ailleurs, plus les seuils alpha et les corrélation maximales augmentent, plus les écarts entre les scores de généralisation et ceux d'entraînement sont élevés, ce qui traduit une augmentation de la variance des modèles. De même, les meilleurs scores d'entraînement sont notés pour les configurations C+M+BF avec des valeurs de seuils alpha et de corrélation maximales élevés. Or, ces configurations correspondent aux modèles qui possèdent le plus grand nombre de variables. Ces scores pourraient alors être le résultat d'un *overfitting* responsable de l'augmentation de la variance.

4.2.4.2/ PERFORMANCES DE PRÉDICTION VIS À VIS DE LA LITTÉRATURE

Dans la littérature, à notre connaissance, aucune étude n'a étudié l'apport conjoint du métabolisme, de la perfusion tumorale et de son hétérogénéité évalués par la TEP au ^{18}F -FDG dans la prédiction de la réponse à la CNA. En revanche, des études ont examiné le caractère prédictif des paramètres de texture dans ce contexte. Parmi celles-ci, il est nécessaire de distinguer les travaux qui ont évalué la valeur prédictive de certains paramètres en réalisant des études descriptives et ceux qui ont évalué les performances des paramètres sur la prédiction de la réponse à la CNA par des méthodes d'apprentissages supervisés.

En effet, les premiers réalisent des tests univariés ou multivariés pour identifier si certains paramètres ont des valeurs significativement différentes entre les classes pCR et non-pCR, et évaluent également en analyse multivariée la contribution de chaque variable dans la classification. Dans ces études, l'intégralité des données sont utilisées pour évaluer les performances des variables. C'est typiquement la démarche qui a été décrite dans la partie 3.3 du chapitre précédent. Il en va de même avec les analyses ROC qui permettent d'évaluer la valeur prédictive des variables. Toutefois, dans ces travaux, aucune étape d'apprentissage, ni de validation n'est effectuée. Bien que l'on parle alors parfois de prédiction dans ces études, il s'agit plus exactement d'une évaluation de la valeur prédictive des paramètres. Les résultats sont par conséquent complémentaires mais non comparables directement avec les résultats de prédiction. Si l'on se concentre sur les études incluant des paramètres de textures, on peut citer ainsi les travaux de Molina Garcia *et al.* [132] (N=68, pCR : 26%), Yoon *et al.* [134] (N=83, pCR : 55%) et Ha *et al.* [141] (N=73, pCR : 23%) où des associations significatives entre les indices de textures et la réponse histologique complète ont été observées. En revanche, Lemarignier *et al.* [471] (N=168, pCR : 12%) et Chen *et al.* [142] (N=61, pCR : 41%) n'ont quant à eux observé aucune association significative entre les indices de texture issue de la TEP réalisée avant le début de traitement néoadjuvant et la réponse à la CNA. Toutefois, Chen *et al.* ont montré que les variations des indices de texture avant et après deux cures de CNA avaient une valeur prédictive vis à vis de la réponse au traitement. Néanmoins, aucune performance d'apprentissage ou de généralisation ne sont disponibles dans tous ces travaux.

A notre connaissance, trois études ont évalué par apprentissage supervisé les performances des indices de texture issus des images TEP acquises avant le début du traitement néoadjuvant, pour la prédiction de la réponse à la CNA. Antunovic *et al.* [130] (N=79, pCR :24%) ont ainsi testé différents modèles en incluant ou non des indices de textures et ont observé de meilleures performances avec les modèles couplant les paramètres de texture avec des données cliniques. Les auteurs ont alors rapporté des scores de prédiction d'AUC de ROC allant de 0,70 à 0,73. L'ajout des indices de texture permet alors d'améliorer les performances des modèles, mais pas de manière significative. Toutefois, dans cette étude, les mêmes données patients ont été utilisées pour construire et évaluer le modèle, ce qui occasionne une surestimation des scores de performance, comme le soulignent les auteurs dans l'article. Ces derniers ont alors également évalué leurs modèles en réalisant 100 itérations de validations croisées. Les résultats moyens d'AUC sont alors compris entre 0,56 et 0,62. Dans une seconde étude, Lee *et al.* [129] ont comparé sur une large cohorte de patients (N=435, pCR :23%) plusieurs modèles prédictifs, en utilisant dans un premier temps uniquement des données cliniques, puis uniquement des données de texture et enfin en combinant les variables de texture et les paramètres cliniques. Dans cette étude, les auteurs scindent leur jeu de données en deux ensembles, un de test et un d'entraînement. Les scores d'entraînement montrent des AUC de ROC de 0,62 à 0,79 et des performances de généralisation de 0,64 à 0,80. Les meilleurs résultats sont obtenus pour des modèles combinant les données cliniques et les indices de texture. Les différences de scores vis à vis des autres modèles testés sont alors significatives sur les jeux d'entraînement, mais pas sur les jeux de tests. Enfin, Li *et al.* [472] (N=100, pCR :50%) ont très récemment évalué l'apport des indices de textures issus de la TEP mais également de l'image TDM pour la prédiction de la réponse à la CNA. Les auteurs montrent alors que les modèles combinant les paramètres cliniques et les indices radiomiques issus de la TEP et des images TDM donnent de meilleurs résultats que sans l'analyse radiomique, avec respectivement une AUC de ROC et une *accuracy* de 0,95 et 0,86 pour les scores d'entraînement et de 0,73 et 0,8 pour les scores de généralisation. Toutefois, les auteurs ne mentionnent pas si ces différences sont significatives entre les modèles.

Sans correction du déséquilibre des classes, les résultats d'AUC de généralisation obtenus dans la présente étude ($0,67 \pm 0,08$) restent donc inférieurs à ceux recensés dans la littérature. Néanmoins, les auteurs de [130] n'ayant pas testé leur modèle sur un jeu de données indépendant, les scores de généralisation présentés sont très certainement surestimés. Peut être serait-il plus juste alors de comparer les résultats avec ceux de leurs validations croisées moyennées sur 100 itérations. Lee *et al.* [129] ont quant à eux évalué les performances de leurs modèles sur un unique jeu de test, sans effectuer plusieurs itérations permettant de s'affranchir d'un découpage de la population possiblement avantageux ou désavantageux. Il convient donc de rester prudent sur la comparaison d'une valeur de score unique par rapport à une moyenne calculée sur 50 itérations. Enfin, Li *et al.* [472] ont utilisé les données de texture des images TDM dans leur modèle optimal ce qui *a priori* améliore les performances de généralisation. Par ailleurs, là encore il s'agit de résultats évalués sur un unique ensemble de test et qui plus est, sur une population dont les proportions des classes pCR/non-pCR sont équilibrées. De plus, il est intéressant de souligner que lorsqu'une correction du déséquilibre des classes est effectué

sur nos données d'entraînement (AUC de ROC de généralisation maximale de 0,77 ($\pm 0,07$)), les écarts avec les scores de généralisation des études précédemment décrites se réduisent, et les performances sont alors meilleures que celles d'Antunovic *et al.* [130]. Par ailleurs, dans deux de ces trois études, une amélioration des résultats est observée lorsque les données de texture sont ajoutées aux modèles prédictifs, mais pas de manière significative, ce qui est également le cas dans notre étude. Enfin, pour chaque modèle, il est intéressant de noter que les variables cliniques, comme le phénotype tumoral ou les récepteurs hormonaux, apparaissent de manière récurrente parmi les variables ayant un fort impact prédictif.

Il est important de souligner que, tout comme la démarche réalisée dans [130], l'objectif de la présente étude était d'identifier si une configuration permettait d'obtenir en moyenne des scores de généralisation plus élevés qu'une autre. Cette approche ne permet pas d'identifier un modèle unique mais une catégorie/un sous-ensemble de variables informatives pour la classification étudiée. Les variables sélectionnées vont varier entre les différents *holdouts* au sein d'une même configuration donnant des scores finaux de généralisation et d'entraînement issus de modèles différents. Cela concerne également la régularisation des modèles où l'optimisation de l'hyperparamètre C est réalisée à chaque *holdouts* et en fonction de chaque jeu de variables sélectionnées et va donc varier régulièrement entre les modèles d'une même configuration.

4.2.4.3/ LIMITES ET CONCLUSION DE L'ÉTUDE

Cette analyse a permis, en moyennant les résultats sur les 50 itérations, de montrer que les paramètres issus du métabolisme tumoral et les paramètres de texture associés permettent une amélioration de la prédiction à la CNA, à l'inverse des paramètres de perfusion. Toutefois, de nombreux aspects doivent encore être étudiés avant de statuer de manière définitive sur la non-information additionnelle des paramètres de perfusion pour la prédiction de la réponse à la CNA. En effet, cette étude a été effectuée en utilisant une unique sélection de variable univariée. Les deux principales limites de ce type de sélection sont la non prise en compte de la colinéarité et corrélation des variables sélectionnées, et la non considération des interactions combinées des variables entre elles. Le premier aspect a été limité dans notre étude par l'ajout d'une sélection basée sur les corrélations entre les variables sélectionnées de manière à ne conserver que celles présentant des corrélations en deçà d'un certain seuil. Néanmoins, le second point constitue une limite majeure dans l'analyse. En effet, si une variable de manière isolée n'apporte *a priori* pas d'information au modèle, elle pourrait néanmoins s'avérer informative si elle était couplée à une/des autre(s) variable(s). Or, la sélection actuelle ne permet pas de sélectionner ce type de variable. Il convient donc, dans de futurs travaux, d'inclure des sélections de variables enveloppantes ou intégrées afin de prendre un compte cet aspect et d'observer l'impact que cela aurait sur les variables sélectionnées et sur les résultats de généralisation. Par ailleurs, cette classification supervisée a été effectuée en utilisant la régression logistique, qui est une approche qui a comme avantage d'être facile à paramétrer et à interpréter et reste une première étape pertinente avant de tester des modèles plus complexes. Il serait donc intéressant dans un second temps, de comparer plusieurs types d'algorithmes, ayant des

approches différentes, non-linéaires comme les forêts aléatoires, puis observer l'impact sur les résultats de prédiction de réponse à la CNA pour notre cohorte de patientes.

EN RÉSUMÉ Dans ce dernier chapitre, nous avons évalué l'apport combiné du métabolisme et de la perfusion tumorale évalués par la TEP au ^{18}F -FDG dans la prédiction de la réponse à la CNA. L'analyse par apprentissage supervisé a été réalisée en comparant plusieurs modèles et a permis de mettre en évidence l'amélioration des performances de prédiction par l'utilisation des paramètres de texture du métabolisme. En ce qui concerne les données de perfusion, ces dernières n'ont par contre, pas permis d'améliorer les scores de généralisation. Toutefois, il serait intéressant de tester d'autres approches de sélection de variables ainsi que d'autres algorithmes de classification afin de ne pas écarter de manière prématurée les données de perfusion dans la prédiction de la réponse à la CNA.

DISCUSSION ET PERSPECTIVES

L'évaluation de la perfusion tumorale à partir d'une imagerie TEP au ^{18}F -FDG de premier passage [94, 95, 253, 473] possède deux avantages majeurs : l'utilisation du ^{18}F -FDG permet l'analyse combinée du métabolisme et de la perfusion tumorale à partir d'une unique injection de radiotracer et la rend accessible à la majorité des centres hospitaliers. Toutefois, en dépit de l'intérêt reconnu de l'imagerie fonctionnelle et moléculaire dans l'évaluation précoce et la prédiction de la réponse à la CNA des cancers du sein, l'utilisation combinée de ces deux biomarqueurs n'a, à l'heure actuelle, que très peu été étudiée.

Grâce aux données du protocole TREN, l'objectif des travaux de recherche présentés dans ce manuscrit a été d'évaluer conjointement ces deux biomarqueurs dans la prédiction de la réponse à la CNA des patientes atteintes de cancers du sein.

Nous avons donc, dans un premier temps, généré la base de données sur laquelle tous nos travaux ont pu être effectués. Cette étape, présentée dans le chapitre 2, a permis d'obtenir, à partir des images de TEP dynamique et statique, acquises avant le début du traitement néoadjuvant, les données de perfusion et de métabolisme tumoral pour un total de 246 patientes. Pour ce faire, des outils informatiques permettant la création et l'exploitation des images paramétriques de perfusion et de métabolisme, ainsi que des scripts permettant le traitement automatique de la base de données ont été développés.

Afin de caractériser le plus précisément possible les lésions mammaires, l'étude de l'hétérogénéité tumorale a également été réalisée. Cette dernière a pu être quantifiée par l'extraction d'indices de texture des images paramétriques de perfusion et de métabolisme. Ainsi, dans une première étude présentée dans le chapitre 3 (part. 3.3) nous avons tout d'abord évalué l'impact que pourraient avoir différentes méthodes de calcul des indices de texture sur leurs relations avec la réponse à la CNA. Des analyses statistiques univariées et multivariées ont ainsi permis d'identifier qu'une **discrétisation relative (DR) de l'image semble plus appropriée pour identifier les patientes bonnes répondeuses**. à la suite de cette étude, nous avons également décidé d'utiliser 32 bins pour le rééchantillonnage des images paramétriques de perfusion et d'utiliser 13 matrices de cooccurrence pour le calcul des indices de textures. L'analyse multivariée a également montré que **les indices de texture**

métaboliques semblent apporter une information additionnelle à celle du volume tumoral. De même, les résultats suggèrent que **l'information de l'hétérogénéité de la perfusion combinée aux paramètres métaboliques améliore l'identification des réponses histologiques des patientes.**

Une fois la méthode de calcul des indices de texture définie, nous avons évalué dans une seconde étude, présentée dans ce même chapitre 3 (part. 3.4), les relations entre les paramètres d'hétérogénéité tumorale avec les caractéristiques biologiques et moléculaires des tumeurs. Les résultats ont montré **une perfusion maximale (BF_{\max}) significativement plus élevée pour les tumeurs de stade T3 et T4 et également chez les patients présentant un envahissement métastatique lymphatique régional (N+) (BF_{\max} et BF_{mean}).** Or, ces différences de perfusion selon les stades T et N ne sont pas observées sur les valeurs globales de métabolisme tumoral. L'étude de la perfusion pourrait donc fournir des informations supplémentaires sur le caractère invasif du cancer du sein. Par ailleurs, **une différence significative d'hétérogénéité de perfusion a été notée entre les phénotypes tumoraux bien qu'aucune différence de perfusion globale n'ait été observée.** Ces résultats soulignent donc l'intérêt de l'analyse combinée de l'hétérogénéité du métabolisme et de la perfusion pour une meilleure compréhension du comportement tumoral.

Enfin, l'analyse de la prédiction de la réponse histologique à la CNA a été réalisée dans une dernière étude présentée dans le chapitre 4. Plusieurs configurations ont été testées et **les meilleures performances ont été notées pour les modèles combinant les paramètres cliniques et métaboliques (globaux et de texture).** Les variables de perfusion ne permettent donc pas une amélioration de la prédiction, malgré les résultats encourageant de notre première étude.

Bien que ces premiers résultats soient de prime abord décevants vis à vis de l'apport de la perfusion dans la prédiction de la réponse à la CNA, il convient d'énumérer les limites de nos travaux. Ces dernières sont de deux types : de nature méthodologique, ou liées au modèle étudié.

1. **La stratégie d'apprentissage supervisé** présente actuellement des limites qui pourraient impacter les résultats de prédiction. La principale limite vient du type de sélection de variable utilisé. Cette sélection univariée ne permet pas d'étudier la valeur prédictive de variables combinées entre elles, ce qui peut potentiellement nuire à l'inclusion de variables informatives de perfusion. Cette hypothèse est encouragée par les résultats de l'étude de Mankoff *et al.* [89] où les auteurs ont montré que le rapport du métabolisme et de la perfusion tumorale évaluée par la TEP à l' $^{15}\text{O}-\text{H}_2\text{O}$ était significativement corrélé ($p\text{-value} = 0,01$) à la réponse à la CNA, contrairement au paramètre de perfusion pris indépendamment. Cette observation a également été confirmée par Spetch *et al.* [73] quelques années plus tard.

Par ailleurs, via les résultats de notre seconde étude, nous avons pu noter des corrélations moyennement à fortement élevées entre les indices de texture du métabolisme et ceux de

perfusion. Malgré l'élimination des paramètres fortement corrélés durant le processus de sélection de variables, il se peut néanmoins que certains indices de texture de perfusion soient toujours trop redondants avec les paramètres de texture métaboliques, faisant baisser les performances de généralisation des modèles.

Il apparaît donc nécessaire d'approfondir ce travail de prédiction en utilisant des sélections de variables enveloppantes ou intégrées, qui vont permettre de prendre en compte les relations des indices entre eux. Par ailleurs, de récentes études montrent une variabilité des résultats en fonction des choix de méthodes de sélection de variables ainsi que des algorithmes de classification utilisés [474–476]. Il serait donc intéressant de tester d'autres algorithmes de classification non linéaires, tels que les arbres décisionnels, qui pourraient éventuellement mieux correspondre à la classification des données étudiées.

- 2. Le faible rapport signal sur bruit des images TEP dynamiques** limite l'analyse quantitative de la perfusion, comme nous l'avons souligné dans le chapitre 2. L'amélioration de ce rapport signal sur bruit pourrait donc influencer l'apport de la perfusion dans la prédiction de la réponse à la CNA et améliorerait également l'analyse de texture (chap. 3). Trois pistes de développement sont à l'oeuvre :

- L'étude présentée en annexe de ce manuscrit (travail de stage de Master 2 réalisé par M. Benjamin Lalande), portant sur l'optimisation des paramètres de reconstruction sur la TEP Gemini TF. Cette étude a montré des résultats prometteurs quant à l'amélioration de la qualité des images de faible statistique. L'une des perspectives à ce travail serait donc de re-générer l'intégralité de la base de données avec les paramètres de reconstruction optimisés afin d'effectuer de nouveau l'analyse de la prédiction de la réponse à la CNA. Les résultats obtenus avec et sans amélioration de la qualité des images de perfusion pourraient alors être comparés entre eux.
- Le débruitage des données dynamiques via des méthodes d'apprentissage profond. Ce travail est en cours de réalisation en partenariat scientifique avec *General Electric Health Care* (GEHC) dans le cadre d'un contrat doctoral.
- L'acquisition des données avec un imageur TEP "numérique", dans lequel les photomultiplicateurs analogiques classiques sont remplacés par des photodétecteurs utilisant des matrices de microcellules de photodiodes à avalanche (SiPMs). Cette nouvelle génération de machines, plus sensible (meilleure résolution temporelle, champ de vue axial étendu), permettrait d'améliorer le rapport signal sur bruit des images dynamiques. Depuis 2017, tous les examens réalisés dans le cadre du protocole TREN sont acquis sur ce type de machine (Discovery MI 4, GEHC).

Il pourrait donc également être intéressant d'ajouter ces examens à la base de données en procédant à une harmonisation des paramètres extraits des images [414].

- 3. Le choix des variables étudiées** a été fait de manière à ne sélectionner que les données disponibles avant le début de la CNA, de manière à obtenir une prédiction au sens strict de la chimiosensibilité. Or, de nombreuses études, dont certaines réalisées par notre équipe, ont montré la valeur prédictive de la réponse tumorale métabolique précoce, évaluée après une ou deux cures de traitement [1–3, 5]. De la même manière, une association entre les variations de perfusion après un cycle ou à moyen terme de la CNA et le pronostic des

patientes a été observé [90, 92, 473]. Les données d'imagerie après un cycle de CNA étant également disponibles dans la base de données TREN, il serait donc intéressant d'ajouter les informations des variations des paramètres conventionnels TEP, comme le SUV_{max} , le SUV_{mean} mais également le BF_{max} et le BF_{mean} , dans la prédiction de la réponse au traitement. Les variations des indices de texture seront probablement plus difficiles à étudier, la réduction du volume métabolique tumoral étant généralement importante même après une seule cure de chimiothérapie.

Par ailleurs, des indices morphologiques, extraits de la tomodensitométrie, pourraient également être ajoutés, ces derniers permettant de compléter la caractérisation des lésions [392–394].

Enfin, au-delà des données phénotypiques, des données génotypiques pourraient également être intégrées dans les modèles. Une étude récente a montré l'intérêt de combiner les informations métaboliques et génotypiques pour prédire la réponse histologique [477]. Notre équipe a initié une étude pilote portant sur l'identification de profils transcriptomiques permettant de prédire la réponse histologique, combinés à la réponse métabolique, dans les cancers du sein triple négatifs (étude TRANSTEP). Une acquisition TEP au ^{18}F -FDG de premier passage a été réalisée pour tous les patients inclus dans cette étude, les données de perfusion tumorale y seront donc intégrées.

4. **Le choix du modèle** (cancer du sein et traitement néoadjuvant) étudié dans ce travail de thèse nous semble pertinent du fait de l'enjeu (cancer le plus fréquent chez la femme), de l'objectif (prédiction de la réponse pour adapter le traitement) et de la grande hétérogénéité des tumeurs mammaires. Cependant, le faible volume des tumeurs mammaires, même larges ou localement avancées, par rapport à d'autres types de tumeurs, rend assez difficile l'utilisation des paramètres de texture. L'étude combinée du métabolisme et de la perfusion pourrait être particulièrement intéressante pour évaluer la réponse à des stratégies thérapeutiques associant chimiothérapie cytotoxique et antiangiogéniques, ce qui n'est pas le cas dans la prise en charge néoadjuvante des cancers du sein, malgré les résultats encourageants de l'étude AVATAXHER [5].

Nous allons appliquer cette approche à d'autres modèles, notamment dans le cadre de l'étude PREMETHÉP. Cette étude multicentrique (PHRC-K 2014) a pour objectif d'évaluer la réponse des carcinomes hépatocellulaires (CHC) avancés à un traitement par sorafenib, en associant la TEP au ^{18}F -FDG et la TEP à la ^{18}F -fluorocholine, un traceur de la synthèse lipidique. Les CHC avancés sont généralement des tumeurs très volumineuses, et le sorafenib est une molécule anti tyrosine kinase qui a à la fois un effet antimitotique et antiangiogénique. Une acquisition dynamique au premier passage est réalisée immédiatement après l'injection de ^{18}F -FDG dans le cadre de cette étude. Nous serons confrontés toutefois à d'autres difficultés, liés aux mouvements de la tumeur durant l'acquisition, et au caractère multicentrique de cette étude.

En conclusion, malgré les limites soulevées, **l'analyse radiomique du métabolisme et de la perfusion tumorale, en utilisant la TEP au ^{18}F -FDG avec une acquisition au premier passage, a apporté des résultats encourageants pour la compréhension**

des comportements tumoraux dans le modèle utilisé, bien que l'ajout de la perfusion tumorale ne permette pas, *a priori* d'améliorer la prédiction de la réponse au traitement.

LISTE DES PUBLICATIONS

ARTICLE DANS UNE REVUE AVEC COMITÉ DE LECTURE

Biological correlates of tumor perfusion and its heterogeneity in newly diagnosed breast cancer using dynamic first-pass ^{18}F -FDG PET/CT, N. **Payan**, B.Presles, F.Brunotte, C.Coutant, I.Desmoulins, JM.Vrigneaud, A.Cochet. Eur J Nucl Med Mol Imaging. 2020 ;47(5) :1103-1115.

COMMUNICATIONS ORALES À DES CONFÉRENCES INTERNATIONALES

Texture analysis of tumour metabolism and perfusion for newly diagnosed breast cancer using dynamic first-pass ^{18}F -FDG PET/CT, N. **Payan**, B.Presles, JM.Vrigneaud, A.Cochet. 31st Annual Congress of the European-Association-of-Nuclear-Medicine (EANM), European Association for Nuclear Medicine, Oct 2018, Dusseldorf, Germany.

Contribution of tumour blood-flow for prediction of response to neoadjuvant chemotherapy in breast cancer using ^{18}F -FDG PET/CT. N. **Payan**, B.Presles, JM.Vrigneaud, A.Cochet. 32nd Annual Congress of the European-Association-of-Nuclear-Medicine (EANM), European Association for Nuclear Medicine, Oct 2019, Barcelona, Spain.

BIBLIOGRAPHIE

- [1] BERRIOLO-RIEDINGER, A., TOUZERY, C., RIEDINGER, J.-M., TOUBEAU, M., COUDERT, B., ARNOULD, L., BOICHOT, C., COCHET, A., FUMOLEAU, P., AND BRUNOTTE, F. **[18F]FDG-PET predicts complete pathological response of breast cancer to neoadjuvant chemotherapy.** *Eur. J. Nucl. Med. Mol. Imaging* 34, 12 (Dec. 2007), 1915–1924.
- [2] ROUSSEAU, C., DEVILLERS, A., SAGAN, C., FERRER, L., BRIDJI, B., CAMPION, L., RICAUD, M., BOURBOULOUX, E., DOUTRIAUX, I., CLOUET, M., BERTON-RIGAUD, D., BOURIEL, C., DELECROIX, V., GARIN, E., ROUQUETTE, S., RESCHE, I., KERBRAT, P., CHATAL, J. F., AND CAMPONE, M. **Monitoring of early response to neoadjuvant chemotherapy in stage II and III breast cancer by [18F]Fluorodeoxyglucose positron emission tomography.** *Journal of Clinical Oncology* 24, 34 (2006), 5366–5372.
- [3] HUMBERT, O., BERRIOLO-RIEDINGER, A., RIEDINGER, J. M., COUDERT, B., ARNOULD, L., COCHET, A., LOUSTALOT, C., FUMOLEAU, P., AND BRUNOTTE, F. **Changes in 18F-FDG tumor metabolism after a first course of neoadjuvant chemotherapy in breast cancer : influence of tumor subtypes.** *Ann. Oncol.* 23, 10 (Oct. 2012), 2572–2577.
- [4] GEBHART, G., GAMEZ, C., HOLMES, E., ROBLES, J., GARCIA, C., CORTES, M., DE AZAMBUJA, E., FAURIA, K., VAN DOOREN, V., AKTAN, G., COCCIA-PORTUGAL, M. A., B. KIM, S., VUYLSTEKE, P., CURE, H., EIDTMANN, H., BASELGA, J., PICCART, M., FLAMEN, P., AND DI COSIMO, S. **18F-FDG PET/CT for early prediction of response to neoadjuvant lapatinib, trastuzumab, and their combination in HER2-Positive breast cancer : Results from Neo-ALTTO.** *Journal of Nuclear Medicine* 54, 11 (2013), 1862–1868.
- [5] COUDERT, B., PIERGA, J.-Y., MOURET-REYNIER, M.-A., KERROU, K., FERRERO, J.-M., PETIT, T., KERBRAT, P., DUPRÉ, P.-F., BACHELOT, T., GABELLE, P., GIARD, S., COEFFIC, D., BOUGNOUX, P., PREVOST, J.-B., PAINTAUD, G., THIBAUT, G., HERNANDEZ, J., COUDERT, M., ARNOULD, L., AND BERRIOLO-RIEDINGER, A. **Use of [18F]-FDG PET to predict response to neoadjuvant trastuzumab and docetaxel in patients with HER2-positive breast cancer, and addition of bevacizumab to neoadjuvant trastuzumab and docetaxel in [18F]-FDG PET-predicted non-responders (AVATAXHER) : an open-label, randomised phase 2 trial.** *The Lancet Oncology* 15, 13 (2014), 1493–1502.
- [6] **Institut national du cancer (inca).**

- [7] HANBY, A. M., AND WALKER, C. **Tavassoli FA, devilee p : Pathology and genetics : Tumours of the breast and female genital organs. WHO classification of tumours series - volume IV. lyon, france : IARC press.** *Breast Cancer Research* 6, 3 (2004).
- [8] ELSTON, C. W., AND ELLIS, I. O. **Pathological prognostic factors in breast cancer. i. the value of histological grade in breast cancer : experience from a large study with long-term follow-up. c. w. elston & i. o. ellis. histopathology 1991 ; 19 ; 403-410.** *Histopathology* 41, 3A (Sept. 2002), 151–2, discussion 152–3.
- [9] NIXON, A. J., SCHNITT, S. J., GELMAN, R., GAGE, I., BORNSTEIN, B., HETEKIDIS, S., RECHT, A., SILVER, B., HARRIS, J. R., AND CONNOLLY, J. L. **Relationship of tumor grade to other pathologic features and to treatment outcome of patients with early stage breast carcinoma treated with breast-conserving therapy.** *Cancer* 78, 7 (Oct. 1996), 1426–1431.
- [10] PINDER, S. E., MURRAY, S., ELLIS, I. O., TRIHIA, H., ELSTON, C. W., GELBER, R. D., GOLDBIRSCH, A., LINDTNER, J., CORTÉS-FUNES, H., SIMONCINI, E., BYRNE, M. J., GOLOUH, R., RUDENSTAM, C. M., CASTIGLIONE-GERTSCH, M., AND GUSTERSON, B. A. **The importance of the histologic grade of invasive breast carcinoma and response to chemotherapy.** *Cancer* 83, 8 (Oct. 1998), 1529–1539.
- [11] PRESS, M. F., MICHAEL F. PRESS, SLAMON, D. J., FLOM, K. J., PARK, J., ZHOU, J.-Y., AND BERNSTEIN, L. **Evaluation of HER-2/neu gene amplification and overexpression : Comparison of frequently used assay methods in a molecularly characterized cohort of breast cancer specimens.** *Journal of Clinical Oncology* 20, 14 (2002), 3095–3105.
- [12] PEROU, C. M., SØRLIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., REES, C. A., POLLACK, J. R., ROSS, D. T., JOHNSEN, H., AKSLEN, L. A., FLUGE, Ø., PERGAMENSCHIKOV, A., WILLIAMS, C., ZHU, S. X., LØNNING, P. E., BØRRESEN-DALE, A.-L., BROWN, P. O., AND BOTSTEIN, D. **Molecular portraits of human breast tumours.** *Nature* 406, 6797 (2000), 747–752.
- [13] SØRLIE, T., PEROU, C. M., TIBSHIRANI, R., AAS, T., GEISLER, S., JOHNSEN, H., HASTIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., THORSEN, T., QUIST, H., MATESE, J. C., BROWN, P. O., BOTSTEIN, D., LØNNING, P. E., AND BØRRESEN-DALE, A. L. **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc. Natl. Acad. Sci. U. S. A.* 98, 19 (Sept. 2001), 10869–10874.
- [14] SORLIE, T., TIBSHIRANI, R., PARKER, J., HASTIE, T., MARRON, J. S., NOBEL, A., DENG, S., JOHNSEN, H., PESICH, R., GEISLER, S., DEMETER, J., PEROU, C. M., LØNNING, P. E., BROWN, P. O., BØRRESEN-DALE, A.-L., AND BOTSTEIN, D. **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc. Natl. Acad. Sci. U. S. A.* 100, 14 (July 2003), 8418–8423.
- [15] KRIJGSMAN, O., ROEPMAN, P., ZWART, W., CARROLL, J. S., TIAN, S., DE SNOO, F. A., BENDER, R. A., BERNARDS, R., AND GLAS, A. M. **A diagnostic gene profile**

- for molecular subtyping of breast cancer associated with treatment response.** *Breast Cancer Research and Treatment* 133, 1 (2012), 37–47.
- [16] PARKER, J. S., MULLINS, M., CHEANG, M. C. U., LEUNG, S., VODUC, D., VICKERY, T., DAVIES, S., FAURON, C., HE, X., HU, Z., QUACKENBUSH, J. F., STIJLEMAN, I. J., PALAZZO, J., MARRON, J. S., NOBEL, A. B., MARDIS, E., NIELSEN, T. O., ELLIS, M. J., PEROU, C. M., AND BERNARD, P. S. **Supervised risk predictor of breast cancer based on intrinsic subtypes.** *J. Clin. Oncol.* 27, 8 (Mar. 2009), 1160–1167.
- [17] REIS-FILHO, J. S., WESTBURY, C., AND PIERGA, J.-Y. **The impact of expression profiling on prognostic and predictive testing in breast cancer.** *J. Clin. Pathol.* 59, 3 (Mar. 2006), 225–231.
- [18] CAREY, L. A., PEROU, C. M., LIVASY, C. A., DRESSLER, L. G., COWAN, D., CONWAY, K., KARACA, G., TROESTER, M. A., TSE, C. K., EDMISTON, S., DEMING, S. L., GERADTS, J., CHEANG, M. C. U., NIELSEN, T. O., MOORMAN, P. G., EARP, H. S., AND MILLIKAN, R. C. **Race, breast cancer subtypes, and survival in the carolina breast cancer study.** *JAMA* 295, 21 (June 2006), 2492–2502.
- [19] KONECNY, G. E., PEGRAM, M. D., VENKATESAN, N., FINN, R., YANG, G., RAHMEH, M., UNTCH, M., RUSNAK, D. W., SPEHAR, G., MULLIN, R. J., KEITH, B. R., GILMER, T. M., BERGER, M., PODRATZ, K. C., AND SLAMON, D. J. **Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells.** *Cancer Res.* 66, 3 (Feb. 2006), 1630–1639.
- [20] DENT, R., HANNA, W. M., TRUDEAU, M., RAWLINSON, E., SUN, P., AND NAROD, S. A. **Pattern of metastatic spread in triple-negative breast cancer.** *Breast Cancer Res. Treat.* 115, 2 (May 2009), 423–428.
- [21] SMID, M., WANG, Y., ZHANG, Y., SIEUWERTS, A. M., YU, J., KLIJN, J. G. M., FOEKENS, J. A., AND MARTENS, J. W. M. **Subtypes of breast cancer show preferential site of relapse.** *Cancer Res.* 68, 9 (May 2008), 3108–3114.
- [22] KENNECKE, H., YERUSHALMI, R., WOODS, R., CHEANG, M. C. U., VODUC, D., SPEERS, C. H., NIELSEN, T. O., AND GELMON, K. **Metastatic behavior of breast cancer subtypes.** *J. Clin. Oncol.* 28, 20 (July 2010), 3271–3277.
- [23] DE RONDE, J. J., HANNEMANN, J., HALFWERK, H., MULDER, L., STRAVER, M. E., VRANCKEN PEETERS, M.-J. T. F. D., WESSELING, J., VAN DE VIJVER, M., WESSELS, L. F. A., AND RODENHUIS, S. **Concordance of clinical and molecular breast cancer subtyping in the context of preoperative chemotherapy response.** *Breast Cancer Res. Treat.* 119, 1 (Jan. 2010), 119–126.
- [24] BHARGAVA, R., STRIEBEL, J., BERIWAL, S., FLICKINGER, J. C., ONISKO, A., AHRENDT, G., AND DABBS, D. J. **Prevalence, morphologic features and proliferation indices of breast carcinoma molecular classes using immunohistochemical surrogate markers.** *Int. J. Clin. Exp. Pathol.* 2, 5 (Feb. 2009), 444–455.
- [25] GNANT, M., HARBECK, N., AND THOMSEN, C. **St. gallen 2011 : Summary of the consensus discussion.** *Breast Care* 6, 2 (Apr. 2011), 136–141.

- [26] GOLDHIRSCH, A., WINER, E. P., COATES, A. S., GELBER, R. D., PICCART-GE BHART, M., THÜRLIMANN, B., SENN, H.-J., AND PANEL MEMBERS. **Personalizing the treatment of women with early breast cancer : highlights of the st gallen international expert consensus on the primary therapy of early breast cancer 2013.** *Ann. Oncol.* 24, 9 (Sept. 2013), 2206–2223.
- [27] MAISONNEUVE, P., DISALVATORE, D., ROTMENSZ, N., CURIGLIANO, G., COLLEONI, M., DELLAPASQUA, S., PRUNERI, G., MASTROPASQUA, M. G., LUINI, A., BASSI, F., PAGANI, G., VIALE, G., AND GOLDHIRSCH, A. **Proposed new clinicopathological surrogate definitions of luminal a and luminal B (HER2-negative) intrinsic breast cancer subtypes.** *Breast Cancer Research* 16, 3 (2014).
- [28] ROUZIER, R., PEROU, C. M., SYMMANS, W. F., IBRAHIM, N., CRISTOFANILLI, M., ANDERSON, K., HESS, K. R., STEC, J., AYERS, M., WAGNER, P., MORANDI, P., FAN, C., RABIUL, I., ROSS, J. S., HORTOBAGYI, G. N., AND PUSZTAI, L. **Breast cancer molecular subtypes respond differently to preoperative chemotherapy.** *Clin. Cancer Res.* 11, 16 (Aug. 2005), 5678–5685.
- [29] FOULKES, W. D., SMITH, I. E., AND REIS-FILHO, J. S. **Triple-negative breast cancer.** *N. Engl. J. Med.* 363, 20 (Nov. 2010), 1938–1948.
- [30] MORRIS, G. J., NAIDU, S., TOPHAM, A. K., GUILLES, F., XU, Y., MCCUE, P., SCHWARTZ, G. F., PARK, P. K., ROSENBERG, A. L., BRILL, K., AND MITCHELL, E. P. **Differences in breast carcinoma characteristics in newly diagnosed African–American and caucasian patients.** *Cancer* 110, 4 (2007), 876–884.
- [31] REIS-FILHO, J. S., AND TUTT, A. N. J. **Triple negative tumours : a critical review.** *Histopathology* 52, 1 (2007), 108–118.
- [32] LEHMANN, B. D., AND PIETENPOL, J. A. **Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes.** *J. Pathol.* 232, 2 (Jan. 2014), 142–150.
- [33] **Référentiel interrégional de prise en charge du cancer du sein 2019.**
- [34] HORTOBAGYI, G. N., AMES, F. C., BUZDAR, A. U., KAU, S. W., MCNEESE, M. D., PAULUS, D., HUG, V., HOLMES, F. A., ROMSDAHL, M. M., FRASCHINI, G., MCBRIDE, C. M., MARTIN, R. G., AND MONTAGUE, E. **Management of stage III primary breast cancer with primary chemotherapy, surgery, and radiation therapy.** *Cancer* 62, 12 (1988), 2507–2516.
- [35] JAIYESIMI, I. A., BUZDAR, A. U., AND HORTOBAGYI, G. **Inflammatory breast cancer : a review.** *J. Clin. Oncol.* 10, 6 (June 1992), 1014–1024.
- [36] CHAUVERGNE, J., DURAND, M., HOERNI, B., COHEN, P., AND LAGARDE, C. **[induction chemotherapy in high risk breast cancer. results of a prospective therapeutic study (author’s transl)].** *Bull. Cancer* 66, 1 (1979), 9–16.
- [37] JACQUILLAT, C., WEIL, M., BAILLET, F., BOREL, C., AUCLERC, G., DE MAUBLANC, M. A., HOUSSET, M., FORGET, G., THILL, L., AND SOUBRANE, C. **Results of neoadjuvant chemotherapy and radiation therapy in the breast-conserving treat-**

- ment of 250 patients with all stages of infiltrative breast cancer. *Cancer* 66, 1 (July 1990), 119–129.
- [38] BONADONNA, G., VERONESI, U., BRAMBILLA, C., FERRARI, L., LUINI, A., GRECO, M., BARTOLI, C., D. YOLDI, G. C., ZUCALI, R., RILKE, F., ANDREOLA, S., SILVESTRINI, R., FRONZO, G. D., AND VALAGUSSA, P. **Primary chemotherapy to avoid mastectomy in tumors with diameters of three centimeters or more.** *JNCI Journal of the National Cancer Institute* 82, 19 (1990), 1539–1545.
- [39] KAUFMANN, M., VON MINCKWITZ, G., SMITH, R., VALERO, V., GIANNI, L., EIERMANN, W., HOWELL, A., COSTA, S. D., BEUZEBOC, P., UNTCH, M., BLOHMER, J.-U., SINN, H.-P., SITTEK, R., SOUCHON, R., TULUSAN, A. H., VOLM, T., AND SENN, H.-J. **International expert panel on the use of primary (preoperative) systemic treatment of operable breast cancer : Review and recommendations.** *Journal of Clinical Oncology* 21, 13 (2003), 2600–2608.
- [40] PERLOFF, M., LESNICK, G. J., KORZUN, A., CHU, F., HOLLAND, J. F., THIRLWELL, M. P., ELLISON, R. R., CAREY, R. W., LEONE, L., AND WEINBERG, V. **Combination chemotherapy with mastectomy or radiotherapy for stage III breast carcinoma : a cancer and leukemia group B study.** *Journal of Clinical Oncology* 6, 2 (1988), 261–269.
- [41] COLDMAN, A. J., AND GOLDIE, J. H. **Factors affecting the development of permanent drug resistance and its impact upon neoadjuvant chemotherapy.** *Recent Results Cancer Res.* 103 (1986), 69–78.
- [42] FISHER, B., GUNDUZ, N., COYLE, J., RUDOCK, C., AND SAFFER, E. **Presence of a growth-stimulating factor in serum following primary tumor removal in mice.** *Cancer Res.* 49, 8 (Apr. 1989), 1996–2001.
- [43] FISHER, B., GUNDUZ, N., AND SAFFER, E. A. **Interrelation between tumor cell proliferation and 17-fluoresceinated estrone binding following primary tumor removal, radiation, cyclophosphamide, or tamoxifen.** *Cancer Res.* 43, 11 (Nov. 1983), 5244–5247.
- [44] FISHER, B., SAFFER, E., RUDOCK, C., COYLE, J., AND GUNDUZ, N. **Effect of local or systemic treatment prior to primary tumor removal on the production and response to a serum growth-stimulating factor in mice.** *Cancer Res.* 49, 8 (Apr. 1989), 2002–2004.
- [45] MAURI, D., PAVLIDIS, N., AND IOANNIDIS, J. P. A. **Neoadjuvant versus adjuvant systemic treatment in breast cancer : a meta-analysis.** *J. Natl. Cancer Inst.* 97, 3 (Feb. 2005), 188–194.
- [46] MIEOG, J. S. D., VAN DER HAGE, J. A., AND VAN DE VELDE, C. J. H. **Neoadjuvant chemotherapy for operable breast cancer.** *Br. J. Surg.* 94, 10 (Oct. 2007), 1189–1200.
- [47] KIROVA, Y. M., CARROLL, S., FOURQUET, A., OFFERSEN, B., ARISTEI, C., AND Y. CHEN, J. **The st gallen international expert consensus on the primary therapy of early breast cancer 2017 : the point of view of an international panel of experts in radiation oncology.** *Annals of Oncology* 29, 1 (2018), 280–281.

- [48] LEE, S. C., GRANT, E., SHETH, P., GARCIA, A. A., DESAI, B., JI, L., GROSHEN, S., HWANG, D., YAMASHITA, M., AND HOVANESSIAN-LARSEN, L. **Accuracy of Contrast-Enhanced ultrasound compared with magnetic resonance imaging in assessing the tumor response after neoadjuvant chemotherapy for breast cancer.** *Journal of Ultrasound in Medicine* 36, 5 (2017), 901–911.
- [49] MAN, V. C. M., DEPARTMENT OF SURGERY, QUEEN MARY HOSPITAL, POKFULAM, KONG, H., AND CHEUNG, P. S. Y. **Neoadjuvant chemotherapy increases rates of breast-conserving surgery in early operable breast cancer.** *Hong Kong Medical Journal* (2017).
- [50] MATHEW, J., ASGEIRSSON, K. S., CHEUNG, K. L., CHAN, S., DAHDA, A., AND ROBERTSON, J. F. R. **Neoadjuvant chemotherapy for locally advanced breast cancer : a review of the literature and future directions.** *Eur. J. Surg. Oncol.* 35, 2 (Feb. 2009), 113–122.
- [51] RASTOGI, P., ANDERSON, S. J., BEAR, H. D., GEYER, C. E., KAHLENBERG, M. S., ROBIDOUX, A., MARGOLESE, R. G., HOEHN, J. L., VOGEL, V. G., DAKHIL, S. R., TAMKUS, D., KING, K. M., PAJON, E. R., WRIGHT, M. J., ROBERT, J., PAIK, S., MAMOUNAS, E. P., AND WOLMARK, N. **Preoperative chemotherapy : Updates of national surgical adjuvant breast and bowel project protocols B-18 and B-27.** *Journal of Clinical Oncology* 26, 5 (2008), 778–785.
- [52] FISHER, B., BRYANT, J., WOLMARK, N., MAMOUNAS, E., BROWN, A., FISHER, E. R., WICKERHAM, D. L., BEGOVIC, M., DECILLIS, A., ROBIDOUX, A., MARGOLESE, R. G., CRUZ, JR, A. B., HOEHN, J. L., LEES, A. W., DIMITROV, N. V., AND BEAR, H. D. **Effect of preoperative chemotherapy on the outcome of women with operable breast cancer.** *J. Clin. Oncol.* 16, 8 (Aug. 1998), 2672–2685.
- [53] GUARNERI, V., BROGLIO, K., KAU, S.-W., CRISTOFANILLI, M., BUZDAR, A. U., VALERO, V., BUCHHOLZ, T., MERIC, F., MIDDLETON, L., HORTOBAGYI, G. N., AND GONZALEZ-ANGULO, A. M. **Prognostic value of pathologic complete response after primary chemotherapy in relation to hormone receptor status and other factors.** *Journal of Clinical Oncology* 24, 7 (2006), 1037–1044.
- [54] LIEDTKE, C., MAZOUNI, C., HESS, K. R., ANDRÉ, F., TORDAI, A., MEJIA, J. A., FRASER SYMMANS, W., GONZALEZ-ANGULO, A. M., HENNESSY, B., GREEN, M., CRISTOFANILLI, M., HORTOBAGYI, G. N., AND PUSZTAI, L. **Response to neoadjuvant therapy and Long-Term survival in patients with Triple-Negative breast cancer.** *Journal of Clinical Oncology* 26, 8 (2008), 1275–1281.
- [55] WOLMARK, N., WANG, J., MAMOUNAS, E., BRYANT, J., AND FISHER, B. **Preoperative chemotherapy in patients with operable breast cancer : Nine-Year results from national surgical adjuvant breast and bowel project B-18.** *JNCI Monographs* 2001, 30 (2001), 96–102.
- [56] KUERER, H. M., NEWMAN, L. A., SMITH, T. L., AMES, F. C., HUNT, K. K., DHINGRA, K., THERIAULT, R. L., SINGH, G., BINKLEY, S. M., SNEIGE, N., BUCHHOLZ, T. A., ROSS, M. I., MCNEESE, M. D., BUZDAR, A. U., HORTOBAGYI, G. N., AND

- SINGLETARY, S. E. **Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy.** *J. Clin. Oncol.* 17, 2 (Feb. 1999), 460–469.
- [57] BONADONNA, G., VALAGUSSA, P., BRAMBILLA, C., FERRARI, L., MOLITERNI, A., TEREZIANI, M., AND ZAMBETTI, M. **Primary chemotherapy in operable breast cancer : eight-year experience at the milan cancer institute.** *J. Clin. Oncol.* 16, 1 (Jan. 1998), 93–100.
- [58] CHEVALLIER, B., ROCHE, H., OLIVIER, J. P., CHOLLET, P., AND HURTELOUP, P. **Inflammatory breast cancer. pilot study of intensive induction chemotherapy (fec-hd) results in a high histologic response rate.** *American journal of clinical oncology* 16, 3 (1993), 223–228.
- [59] MINCKWITZ, G. V., VON MINCKWITZ, G., UNTCH, M., BLOHMER, J.-U., COSTA, S. D., EIDTMANN, H., FASCHING, P. A., GERBER, B., EIERMANN, W., HILFRICH, J., HUOBER, J., JACKISCH, C., KAUFMANN, M., KONECNY, G. E., DENKERT, C., NEKLJUDOVA, V., MEHTA, K., AND LOIBL, S. **Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes.** *Journal of Clinical Oncology* 30, 15 (2012), 1796–1804.
- [60] CAREY, L. A., DEES, E. C., SAWYER, L., GATTI, L., MOORE, D. T., COLLICCHIO, F., OLLILA, D. W., SARTOR, C. I., GRAHAM, M. L., AND PEROU, C. M. **The triple negative paradox : Primary tumor chemosensitivity of breast cancer subtypes.** *Clinical Cancer Research* 13, 8 (2007), 2329–2334.
- [61] COUDERT, B. P., ARNOULD, L., MOREAU, L., CHOLLET, P., WEBER, B., VANLEMENS, L., MOLUÇON, C., TUBIANA, N., CAUSERET, S., L. MISSET, J., FEUTRAY, S., MERY-MIGNARD, D., GARNIER, J., AND FUMOLEAU, P. **Pre-operative systemic (neo-adjuvant) therapy with trastuzumab and docetaxel for HER2-overexpressing stage II or III breast cancer : results of a multicenter phase II trial.** *Annals of Oncology* 17, 3 (2006), 409–414.
- [62] GIANNI, L., EIERMANN, W., SEMIGLAZOV, V., MANIKHAS, A., LLUCH, A., TJULANDIN, S., ZAMBETTI, M., VAZQUEZ, F., BYAKHOW, M., LICHINITSER, M., CLIMENT, M. A., CIRUELOS, E., OJEDA, B., MANSUTTI, M., BOZHOK, A., BARONIO, R., FEYEREISLOVA, A., BARTON, C., VALAGUSSA, P., AND BASELGA, J. **Neoadjuvant chemotherapy with trastuzumab followed by adjuvant trastuzumab versus neoadjuvant chemotherapy alone, in patients with HER2-positive locally advanced breast cancer (the NOAH trial) : a randomised controlled superiority trial with a parallel HER2-negative cohort.** *The Lancet* 375, 9712 (2010), 377–384.
- [63] BHARGAVA, R., BERIWAL, S., DABBS, D. J., OZBEK, U., SORAN, A., JOHNSON, R. R., BRUFISKY, A. M., LEMBERSKY, B. C., AND AHRENDT, G. M. **Immunohistochemical surrogate markers of breast cancer molecular classes predicts response to neoadjuvant chemotherapy : a single institutional experience with 359 cases.** *Cancer* 116, 6 (Mar. 2010), 1431–1439.

- [64] GNANT, M., HARBECK, N., AND THOMSEN, C. **St. Gallen/Vienna 2017 : A brief summary of the consensus discussion about escalation and De-Escalation of primary breast cancer treatment.** *Breast Care* 12, 2 (May 2017), 102–107.
- [65] MAKRIS, A., POWLES, T. J., ALLRED, D. C., ASHLEY, S. E., TROTT, P. A., ORMEROD, M. G., TITLEY, J. C., AND DOWSETT, M. **Quantitative changes in cytological molecular markers during primary medical treatment of breast cancer : a pilot study.** *Breast Cancer Res. Treat.* 53, 1 (Jan. 1999), 51–59.
- [66] CHANG, J., POWLES, T. J., ALLRED, D. C., ASHLEY, S. E., CLARK, G. M., MAKRIS, A., ASSERSOHN, L., GREGORY, R. K., OSBORNE, C. K., AND DOWSETT, M. **Biologic markers as predictors of clinical outcome from systemic therapy for primary operable breast cancer.** *J. Clin. Oncol.* 17, 10 (Oct. 1999), 3058–3063.
- [67] MATSUBARA, N., MUKAI, H., FUJII, S., AND WADA, N. **Different prognostic significance of ki-67 change between pre- and post-neoadjuvant chemotherapy in various subtypes of breast cancer.** *Breast Cancer Res. Treat.* 137, 1 (Jan. 2013), 203–212.
- [68] THERASSE, P., ARBUCK, S. G., EISENHAEUER, E. A., WANDERS, J., KAPLAN, R. S., RUBINSTEIN, L., VERWEIJ, J., VAN GLABBEKE, M., VAN OOSTEROM, A. T., CHRISTIAN, M. C., AND GWYTHYER, S. G. **New guidelines to evaluate the response to treatment in solid tumors. european organization for research and treatment of cancer, national cancer institute of the united states, national cancer institute of canada.** *J. Natl. Cancer Inst.* 92, 3 (Feb. 2000), 205–216.
- [69] EISENHAEUER, E. A., THERASSE, P., BOGAERTS, J., SCHWARTZ, L. H., SARGENT, D., FORD, R., DANCEY, J., ARBUCK, S., GWYTHYER, S., MOONEY, M., RUBINSTEIN, L., SHANKAR, L., DODD, L., KAPLAN, R., LACOMBE, D., AND VERWEIJ, J. **New response evaluation criteria in solid tumours : revised RECIST guideline (version 1.1).** *Eur. J. Cancer* 45, 2 (Jan. 2009), 228–247.
- [70] MGHANGA, F. P., LAN, X., BAKARI, K. H., LI, C., AND ZHANG, Y. **Fluorine-18 fluorodeoxyglucose positron emission tomography-computed tomography in monitoring the response of breast cancer to neoadjuvant chemotherapy : a meta-analysis.** *Clin. Breast Cancer* 13, 4 (Aug. 2013), 271–279.
- [71] CHENG, X., LI, Y., LIU, B., XU, Z., BAO, L., AND WANG, J. **18F-FDG PET/CT and PET for evaluation of pathological response to neoadjuvant chemotherapy in breast cancer : a meta-analysis.** *Acta radiol.* 53, 6 (July 2012), 615–627.
- [72] WANG, Y., ZHANG, C., LIU, J., AND HUANG, G. **Is 18F-FDG PET accurate to predict neoadjuvant therapy response in breast cancer ? a meta-analysis.** *Breast Cancer Res. Treat.* 131, 2 (Jan. 2012), 357–369.
- [73] SPECHT, J. M., KURLAND, B. F., MONTGOMERY, S. K., DUNNWALD, L. K., DOOT, R. K., GRALOW, J. R., ELLIS, G. K., LINDEN, H. M., LIVINGSTON, R. B., ALLISON, K. H., SCHUBERT, E. K., AND MANKOFF, D. A. **Tumor metabolism and blood flow as assessed by positron emission tomography varies by tumor subtype in locally advanced breast cancer.** *Clin. Cancer Res.* 16, 10 (May 2010), 2803–2810.

- [74] KEAM, B., IM, S.-A., KOH, Y., HAN, S.-W., OH, D.-Y., CHO, N., KIM, J. H., HAN, W., KANG, K. W., MOON, W. K., KIM, T.-Y., PARK, I. A., NOH, D.-Y., CHUNG, J.-K., AND BANG, Y.-J. **Early metabolic response using FDG PET/CT and molecular phenotypes of breast cancer treated with neoadjuvant chemotherapy.** *BMC Cancer* 11 (Oct. 2011), 452.
- [75] WAHL, R. L., JACENE, H., KASAMON, Y., AND LODGE, M. A. **From RECIST to PERCIST : Evolving considerations for PET response criteria in solid tumors.** *J. Nucl. Med.* 50 Suppl 1 (May 2009), 122S–50S.
- [76] COUDERT, B., PIERGA, J.-Y., MOURET-REYNIER, M.-A., KERROU, K., FERRERO, J.-M., PETIT, T., KERBRAT, P., DUPRÉ, P.-F., BACHELOT, T., GABELLE, P., GIARD, S., COEFFIC, D., BOUGNOUX, P., PREVOST, J.-B., PAINAUD, G., THIBAUT, G., HERNANDEZ, J., COUDERT, M., ARNOULD, L., AND BERRIOLO-RIEDINGER, A. **Use of [18F]-FDG PET to predict response to neoadjuvant trastuzumab and docetaxel in patients with HER2-positive breast cancer, and addition of bevacizumab to neoadjuvant trastuzumab and docetaxel in [18F]-FDG PET-predicted non-responders (AVATAXHER) : an open-label, randomised phase 2 trial.** *The Lancet Oncology* 15, 13 (2014), 1493–1502.
- [77] PICKLES, M. D., MANTON, D. J., LOWRY, M., AND TURNBULL, L. W. **Prognostic value of pre-treatment DCE-MRI parameters in predicting disease free and overall survival for breast cancer patients undergoing neoadjuvant chemotherapy.** *Eur. J. Radiol.* 71, 3 (Sept. 2009), 498–505.
- [78] O'CONNOR, J. P. B., JACKSON, A., PARKER, G. J. M., AND JAYSON, G. C. **DCE-MRI biomarkers in the clinical evaluation of antiangiogenic and vascular disrupting agents.** *Br. J. Cancer* 96, 2 (Jan. 2007), 189–195.
- [79] MANTON, D. J., CHATURVEDI, A., HUBBARD, A., LIND, M. J., LOWRY, M., MARAVEYAS, A., PICKLES, M. D., TOZER, D. J., AND TURNBULL, L. W. **Neoadjuvant chemotherapy in breast cancer : early response prediction with quantitative MR imaging and spectroscopy.** *Br. J. Cancer* 94, 3 (Feb. 2006), 427–435.
- [80] PADHANI, A. R., HAYES, C., ASSERSOHN, L., POWLES, T., MAKRIS, A., SUCKLING, J., LEACH, M. O., AND HUSBAND, J. E. **Prediction of clinicopathologic response of breast cancer to primary chemotherapy at contrast-enhanced MR imaging : initial clinical results.** *Radiology* 239, 2 (May 2006), 361–374.
- [81] PICKLES, M. D., LOWRY, M., MANTON, D. J., GIBBS, P., AND TURNBULL, L. W. **Role of dynamic contrast enhanced MRI in monitoring early response of locally advanced breast cancer to neoadjuvant chemotherapy.** *Breast Cancer Res. Treat.* 91, 1 (May 2005), 1–10.
- [82] YU, H. J., CHEN, J.-H., MEHTA, R. S., NALCIOGLU, O., AND SU, M.-Y. **MRI measurements of tumor size and pharmacokinetic parameters as early predictors of response in breast cancer patients undergoing neoadjuvant anthracycline chemotherapy.** *J. Magn. Reson. Imaging* 26, 3 (Sept. 2007), 615–623.

- [83] CHEN, J.-H., AND SU, M.-Y. **Clinical application of magnetic resonance imaging in management of breast cancer patients receiving neoadjuvant chemotherapy.** *Biomed Res. Int.* 2013 (June 2013), 348167.
- [84] CHEN, M., WANG, W.-P., JIA, W.-R., TANG, L., WANG, Y., ZHAN, W.-W., AND FEI, X.-C. **Three-dimensional contrast-enhanced sonography in the assessment of breast tumor angiogenesis : correlation with microvessel density and vascular endothelial growth factor expression.** *J. Ultrasound Med.* 33, 5 (May 2014), 835–846.
- [85] WANG, J.-W., ZHENG, W., LIU, J.-B., CHEN, Y., CAO, L.-H., LUO, R.-Z., LI, A.-H., AND ZHOU, J.-H. **Assessment of early tumor response to cytotoxic chemotherapy with dynamic contrast-enhanced ultrasound in human breast cancer xenografts.** *PLoS One* 8, 3 (Mar. 2013), e58274.
- [86] JIA, K., LI, L., WU, X. J., HAO, M. J., AND XUE, H. Y. **Contrast-enhanced ultrasound for evaluating the pathologic response of breast cancer to neoadjuvant chemotherapy : A meta-analysis.** *Medicine* 98, 4 (Jan. 2019), e14258.
- [87] SARACCO, A., SZABÓ, B. K., TÁNCZOS, E., BERGH, J., AND HATSCHKEK, T. **Contrast-enhanced ultrasound (CEUS) in assessing early response among patients with invasive breast cancer undergoing neoadjuvant chemotherapy.** *Acta radiol.* 58, 4 (Apr. 2017), 394–402.
- [88] TANG, M.-X., MULVANA, H., GAUTHIER, T., LIM, A. K. P., COSGROVE, D. O., ECKERSLEY, R. J., AND STRIDE, E. **Quantitative contrast-enhanced ultrasound imaging : a review of sources of variability.** *Interface Focus* 1, 4 (Aug. 2011), 520–539.
- [89] MANKOFF, D. A., DUNNWALD, L. K., GRALOW, J. R., ELLIS, G. K., CHARLOP, A., LAWTON, T. J., SCHUBERT, E. K., TSENG, J., AND LIVINGSTON, R. B. **Blood flow and metabolism in locally advanced breast cancer : relationship to response to therapy.** *J. Nucl. Med.* 43, 4 (Apr. 2002), 500–509.
- [90] MANKOFF, D. A., DUNNWALD, L. K., GRALOW, J. R., ELLIS, G. K., SCHUBERT, E. K., TSENG, J., LAWTON, T. J., LINDEN, H. M., AND LIVINGSTON, R. B. **Changes in blood flow and metabolism in locally advanced breast cancer treated with neoadjuvant chemotherapy.** *J. Nucl. Med.* 44, 11 (Nov. 2003), 1806–1814.
- [91] TSENG, J., DUNNWALD, L. K., SCHUBERT, E. K., LINK, J. M., MINOSHIMA, S., MUZI, M., AND MANKOFF, D. A. **^{18}F -FDG kinetics in locally advanced breast cancer : correlation with tumor blood flow and changes in response to neoadjuvant chemotherapy.** *J. Nucl. Med.* 45, 11 (Nov. 2004), 1829–1837.
- [92] DUNNWALD, L. K., GRALOW, J. R., ELLIS, G. K., LIVINGSTON, R. B., LINDEN, H. M., SPECHT, J. M., DOOT, R. K., LAWTON, T. J., BARLOW, W. E., KURLAND, B. F., SCHUBERT, E. K., AND MANKOFF, D. A. **Tumor metabolism and blood flow changes by positron emission tomography : relation to survival in patients treated with neoadjuvant chemotherapy for locally advanced breast cancer.** *J. Clin. Oncol.* 26, 27 (Sept. 2008), 4449–4457.

- [93] ZASADNY, K. R., TATSUMI, M., AND WAHL, R. L. **FDG metabolism and uptake versus blood flow in women with untreated primary breast cancers.** *Eur. J. Nucl. Med. Mol. Imaging* 30, 2 (Feb. 2003), 274–280.
- [94] MULLANI, N. A., HERBST, R. S., O’NEIL, R. G., GOULD, K. L., BARRON, B. J., AND ABBRUZZESE, J. L. **Tumor blood flow measured by PET dynamic imaging of first-pass 18F-FDG uptake : a comparison with 15o-labeled water-measured blood flow.** *J. Nucl. Med.* 49, 4 (Apr. 2008), 517–523.
- [95] COCHET, A., PIGEONNAT, S., KHOURY, B., VRIGNEAUD, J.-M., TOUZERY, C., BERRIOLO-RIEDINGER, A., DYGAI-COCHET, I., TOUBEAU, M., HUMBERT, O., COUDERT, B., FUMOLEAU, P., ARNOULD, L., AND BRUNOTTE, F. **Evaluation of breast tumor blood flow with dynamic first-pass 18F-FDG PET/CT : comparison with angiogenesis markers and prognostic factors.** *J. Nucl. Med.* 53, 4 (Apr. 2012), 512–520.
- [96] HUMBERT, O., RIEDINGER, J.-M., VRIGNEAUD, J.-M., KANOUN, S., DYGAI-COCHET, I., BERRIOLO-RIEDINGER, A., TOUBEAU, M., DEPARDON, E., LASSERE, M., TISSERAND, S., FUMOLEAU, P., BRUNOTTE, F., AND COCHET, A. **18F-FDG PET-Derived tumor blood flow changes after 1 cycle of neoadjuvant chemotherapy predicts outcome in Triple-Negative breast cancer.** *J. Nucl. Med.* 57, 11 (Nov. 2016), 1707–1712.
- [97] O’FLYNN, E. A. M., AND DESOUSA, N. M. **Functional magnetic resonance : biomarkers of response in breast cancer.** *Breast Cancer Research* 13, 1 (2011).
- [98] PARK, S. H., MOON, W. K., CHO, N., CHANG, J. M., IM, S.-A., PARK, I. A., KANG, K. W., HAN, W., AND NOH, D.-Y. **Comparison of diffusion-weighted MR imaging and FDG PET/CT to predict pathological complete response to neoadjuvant chemotherapy in patients with breast cancer.** *Eur. Radiol.* 22, 1 (Jan. 2012), 18–25.
- [99] WU, L.-M., HU, J.-N., GU, H.-Y., HUA, J., CHEN, J., AND XU, J.-R. **Can diffusion-weighted MR imaging and contrast-enhanced MR imaging precisely evaluate and predict pathological response to neoadjuvant chemotherapy in patients with breast cancer ?** *Breast Cancer Res. Treat.* 135, 1 (Aug. 2012), 17–28.
- [100] PICKLES, M. D., GIBBS, P., LOWRY, M., AND TURNBULL, L. W. **Diffusion changes precede size reduction in neoadjuvant treatment of breast cancer.** *Magnetic Resonance Imaging* 24, 7 (2006), 843–847.
- [101] RICHARD, R., THOMASSIN, I., CHAPPELLIER, M., SCEMAMA, A., DE CREMOUX, P., VARNA, M., GIACCHETTI, S., ESPIÉ, M., DE KERVILER, E., AND DE BAZELAIRE, C. **Diffusion-weighted MRI in pretreatment prediction of response to neoadjuvant chemotherapy in patients with breast cancer.** *European Radiology* 23, 9 (2013), 2420–2431.
- [102] BAEK, H.-M., CHEN, J.-H., NALCIOGLU, O., AND SU, M.-Y. **Proton MR spectroscopy for monitoring early treatment response of breast cancer to neo-adjuvant chemotherapy.** *Ann. Oncol.* 19, 5 (May 2008), 1022–1024.

- [103] BAEK, H.-M., CHEN, J.-H., NIE, K., YU, H. J., BAHRI, S., MEHTA, R. S., NALCIOGLU, O., AND SU, M.-Y. **Predicting pathologic response to neoadjuvant chemotherapy in breast cancer by using MR imaging and quantitative 1H MR spectroscopy.** *Radiology* 251, 3 (June 2009), 653–662.
- [104] TOZAKI, M., OYAMA, Y., AND FUKUMA, E. **Preliminary study of early response to neoadjuvant chemotherapy after the first cycle in breast cancer : comparison of 1H magnetic resonance spectroscopy with diffusion magnetic resonance imaging.** *Jpn. J. Radiol.* 28, 2 (Feb. 2010), 101–109.
- [105] DEHDASHTI, F., MORTIMER, J. E., TRINKAUS, K., NAUGHTON, M. J., ELLIS, M., KATZENELLENBOGEN, J. A., WELCH, M. J., AND SIEGEL, B. A. **PET-based estradiol challenge as a predictive biomarker of response to endocrine therapy in women with estrogen-receptor-positive breast cancer.** *Breast Cancer Research and Treatment* 113, 3 (2009), 509–517.
- [106] CHAE, S. Y., KIM, S.-B., AHN, S. H., KIM, H. O., YOON, D. H., AHN, J.-H., JUNG, K. H., HAN, S., OH, S. J., LEE, S. J., KIM, H. J., SON, B. H., GONG, G., LEE, H. S., AND MOON, D. H. **A randomized feasibility study of F-Fluoroestradiol PET to predict pathologic response to neoadjuvant therapy in estrogen Receptor-Rich postmenopausal breast cancer.** *J. Nucl. Med.* 58, 4 (Apr. 2017), 563–568.
- [107] TADAYYON, H., GANGEH, M., SANNACHI, L., TRUDEAU, M., PRITCHARD, K., GHANDI, S., EISEN, A., LOOK-HONG, N., HOLLOWAY, C., WRIGHT, F., RAKOVITCH, E., VESPRINI, D., TRAN, W. T., CURPEN, B., AND CZARNOTA, G. **prediction of breast tumour response to chemotherapy using quantitative ultrasound imaging and artificial neural networks.** *Oncotarget* 10, 39 (June 2019), 3910–3923.
- [108] SANNACHI, L., GANGEH, M., TADAYYON, H., GANDHI, S., WRIGHT, F. C., SLODKOWSKA, E., CURPEN, B., SADEGHI-NAINI, A., TRAN, W., AND CZARNOTA, G. J. **Breast cancer treatment response monitoring using quantitative ultrasound and texture analysis : Comparative analysis of analytical models.** *Transl. Oncol.* 12, 10 (Oct. 2019), 1271–1281.
- [109] SANNACHI, L., GANGEH, M., TADAYYON, H., SADEGHI-NAINI, A., GANDHI, S., WRIGHT, F. C., SLODKOWSKA, E., CURPEN, B., TRAN, W., AND CZARNOTA, G. J. **Response monitoring of breast cancer patients receiving neoadjuvant chemotherapy using quantitative ultrasound, texture, and molecular features.** *PLoS One* 13, 1 (Jan. 2018), e0189634.
- [110] BAUMGARTNER, A., TAUSCH, C., HOSCH, S., PAPASSOTIROPOULOS, B., VARGA, Z., RAGETH, C., AND BAEGE, A. **Ultrasound-based prediction of pathologic response to neoadjuvant chemotherapy in breast cancer patients.** *Breast* 39 (June 2018), 19–23.
- [111] SHIA, W.-C., HUANG, Y.-L., WU, H.-K., AND CHEN, D.-R. **Using flow characteristics in Three-Dimensional power doppler ultrasound imaging to predict complete responses in patients undergoing neoadjuvant chemotherapy.** *J. Ultrasound Med.* 36, 5 (May 2017), 887–900.

- [112] TADAYYON, H., SANNACHI, L., GANGEH, M. J., KIM, C., GHANDI, S., TRUDEAU, M., PRITCHARD, K., TRAN, W. T., SLODKOWSKA, E., SADEGHI-NAINI, A., AND CZARNOTA, G. J. **A priori prediction of neoadjuvant chemotherapy response and survival in breast cancer patients using quantitative ultrasound.** *Sci. Rep.* 7 (Apr. 2017), 45733.
- [113] EUN, N. L., KANG, D., SON, E. J., PARK, J. S., YOUK, J. H., KIM, J.-A., AND GWEON, H. M. **Texture analysis with 3.0-t MRI for association of response to neoadjuvant chemotherapy in breast cancer.** *Radiology* 294, 1 (Jan. 2020), 31–41.
- [114] LIU, Z., LI, Z., QU, J., ZHANG, R., ZHOU, X., LI, L., SUN, K., TANG, Z., JIANG, H., LI, H., XIONG, Q., DING, Y., ZHAO, X., WANG, K., LIU, Z., AND TIAN, J. **Radiomics of multiparametric MRI for pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer : A multicenter study.** *Clin. Cancer Res.* 25, 12 (June 2019), 3538–3547.
- [115] PEREIRA, N. P., CURI, C., OSÓRIO, C. A. B. T., MARQUES, E. F., MAKDISSI, F. B., PINKER, K., AND BITENCOURT, A. G. V. **Diffusion-Weighted magnetic resonance imaging of patients with breast cancer following neoadjuvant chemotherapy provides early prediction of pathological response - a prospective study.** *Sci. Rep.* 9, 1 (Nov. 2019), 16372.
- [116] LO GULLO, R., ESKREIS-WINKLER, S., MORRIS, E. A., AND PINKER, K. **Machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy.** *Breast* 49 (Nov. 2019), 115–122.
- [117] KONTOPODIS, E., VENIANAKI, M., MANIKIS, G. C., NIKIFORAKI, K., SALVETTI, O., PAPADAKI, E., PAPADAKIS, G. Z., KARANTANAS, A. H., AND MARIAS, K. **Investigating the role of Model-Based and Model-Free imaging biomarkers as early predictors of neoadjuvant breast cancer therapy outcome.** *IEEE J Biomed Health Inform* 23, 5 (Sept. 2019), 1834–1843.
- [118] MACHIREDDY, A., THIBAUT, G., TUDORICA, A., AFZAL, A., MISHAL, M., KEMMER, K., NAIK, A., TROXELL, M., GORANSON, E., OH, K., ROY, N., JAFARIAN, N., HOLTORF, M., HUANG, W., AND SONG, X. **Early prediction of breast cancer therapy response using multiresolution fractal analysis of DCE-MRI parametric maps.** *Tomography* 5, 1 (Mar. 2019), 90–98.
- [119] XIONG, Q., ZHOU, X., LIU, Z., LEI, C., YANG, C., YANG, M., ZHANG, L., ZHU, T., ZHUANG, X., LIANG, C., LIU, Z., TIAN, J., AND WANG, K. **Multiparametric MRI-based radiomics analysis for prediction of breast cancers insensitive to neoadjuvant chemotherapy.** *Clin. Transl. Oncol.* (Apr. 2019).
- [120] HA, R., CHIN, C., KARCICH, J., LIU, M. Z., CHANG, P., MUTASA, S., PASCUAL VAN SANT, E., WYNN, R. T., CONNOLLY, E., AND JAMBAWALIKAR, S. **Prior to initiation of chemotherapy, can we predict breast tumor response ? deep learning convolutional neural networks approach using a breast MRI tumor dataset.** *J. Digit. Imaging* 32, 5 (Oct. 2019), 693–701.

- [121] DRUKKER, K., EDWARDS, A., DOYLE, C., PAPAIOANNOU, J., KULKARNI, K., AND GIGER, M. L. **Breast MRI radiomics for the pretreatment prediction of response to neoadjuvant chemotherapy in node-positive breast cancer patients.** *J Med Imaging (Bellingham)* 6, 3 (July 2019), 034502.
- [122] TAHMASSEBI, A., WENGERT, G. J., HELBICH, T. H., BAGO-HORVATH, Z., ALAEI, S., BARTSCH, R., DUBSKY, P., BALTZER, P., CLAUSER, P., KAPETAS, P., MORRIS, E. A., MEYER-BAESE, A., AND PINKER, K. **Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients.** *Invest. Radiol.* 54, 2 (Feb. 2019), 110–117.
- [123] CAIN, E. H., SAHA, A., HAROWICZ, M. R., MARKS, J. R., MARCOM, P. K., AND MAZUROWSKI, M. A. **Multivariate machine learning models for prediction of pathologic response to neoadjuvant therapy in breast cancer using MRI features : a study using an independent validation set.** *Breast Cancer Res. Treat.* 173, 2 (Jan. 2019), 455–463.
- [124] JAHANI, N., COHEN, E., HSIEH, M.-K., WEINSTEIN, S. P., PANTALONE, L., HYLTON, N., NEWITT, D., DAVATZIKOS, C., AND KONTOS, D. **Prediction of treatment response to neoadjuvant chemotherapy for breast cancer via early changes in tumor heterogeneity captured by DCE-MRI registration.** *Sci. Rep.* 9, 1 (Aug. 2019), 12114.
- [125] KIM, Y., KIM, S. H., SONG, B. J., KANG, B. J., YIM, K.-I., LEE, A., AND NAM, Y. **Early prediction of response to neoadjuvant chemotherapy using dynamic Contrast-Enhanced MRI and ultrasound in breast cancer.** *Korean J. Radiol.* 19, 4 (July 2018), 682–691.
- [126] LEE, J., KIM, S. H., AND KANG, B. J. **Pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer : Perfusion metrics of dynamic contrast enhanced MRI.** *Sci. Rep.* 8, 1 (June 2018), 9490.
- [127] MINARIKOVA, L., BOGNER, W., PINKER, K., VALKOVIČ, L., ZARIC, O., BAGO-HORVATH, Z., BARTSCH, R., HELBICH, T. H., TRATTNIG, S., AND GRUBER, S. **Investigating the prediction value of multiparametric magnetic resonance imaging at 3 T in response to neoadjuvant chemotherapy in breast cancer.** *European Radiology* 27, 5 (2017), 1901–1911.
- [128] BRAMAN, N. M., ETESAMI, M., PRASANNA, P., DUBCHUK, C., GILMORE, H., TIWARI, P., PLECHA, D., AND MADABHUSHI, A. **Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI.** *Breast Cancer Res.* 19, 1 (May 2017), 57.
- [129] LEE, H., LEE, D.-E., PARK, S., KIM, T. S., JUNG, S.-Y., LEE, S., KANG, H. S., LEE, E. S., SIM, S. H., PARK, I. H., LEE, K. S., KWON, Y. M., KONG, S. Y., JOO, J., JEONG, H. J., AND KIM, S.-K. **Predicting response to neoadjuvant chemotherapy in patients with breast cancer : Combined statistical modeling**

- using clinicopathological factors and FDG PET/CT texture parameters. *Clin. Nucl. Med.* 44, 1 (Jan. 2019), 21–29.
- [130] ANTUNOVIC, L., DE SANCTIS, R., COZZI, L., KIRIENKO, M., SAGONA, A., TORRISI, R., TINTERRI, C., SANTORO, A., CHITI, A., ZELIC, R., AND SOLLINI, M. **PET/CT radiomics in breast cancer : promising tool for prediction of pathological response to neoadjuvant chemotherapy.** *Eur. J. Nucl. Med. Mol. Imaging* 46, 7 (July 2019), 1468–1477.
- [131] GARCIA-VICENTE, A. M., PÉREZ-BETETA, J., AMO-SALAS, M., MOLINA, D., JIMENEZ-LONDOÑO, G. A., SORIANO-CASTREJÓN, A. M., PENA PARDO, F. J., AND MARTÍNEZ-GONZÁLEZ, A. **Predictive and prognostic potential of volume-based metabolic variables obtained by a baseline F-FDG PET/CT in breast cancer with neoadjuvant chemotherapy indication.** *Rev. Esp. Med. Nucl. Imagen Mol.* 37, 2 (Mar. 2018), 73–79.
- [132] MOLINA-GARCÍA, D., GARCÍA-VICENTE, A. M., PÉREZ-BETETA, J., AMO-SALAS, M., MARTÍNEZ-GONZÁLEZ, A., TELLO-GALÁN, M. J., SORIANO-CASTREJÓN, Á., AND PÉREZ-GARCÍA, V. M. **Intratumoral heterogeneity in 18F-FDG PET/CT by textural analysis in breast cancer as a predictive and prognostic subrogate.** *Annals of Nuclear Medicine* 32, 6 (2018), 379–388.
- [133] HAYASHI, N., TSUNODA, H., NAMURA, M., OCHI, T., SUZUKI, K., YAMAUCHI, H., AND NAKAMURA, S. **Magnetic resonance imaging combined with second-look ultrasonography in predicting pathologic complete response after neoadjuvant chemotherapy in primary breast cancer patients.** *Clin. Breast Cancer* 19, 1 (Feb. 2019), 71–77.
- [134] YOON, H.-J., KIM, Y., CHUNG, J., AND KIM, B. S. **Predicting neo-adjuvant chemotherapy response and progression-free survival of locally advanced breast cancer using textural features of intratumoral heterogeneity on F-18 FDG PET/CT and diffusion-weighted MR imaging.** *Breast J.* 25, 3 (May 2019), 373–380.
- [135] CHO, N., IM, S.-A., CHEON, G. J., PARK, I.-A., LEE, K.-H., KIM, T.-Y., KIM, Y. S., KWON, B. R., LEE, J. M., SUH, H. Y., AND SUH, K. J. **Integrated 18F-FDG PET/MRI in breast cancer : early prediction of response to neoadjuvant chemotherapy.** *European Journal of Nuclear Medicine and Molecular Imaging* 45, 3 (2018), 328–339.
- [136] GALLIVANONE, F., PANZERI, M. M., CANEVARI, C., LOSIO, C., GIANOLLI, L., DE COBELLI, F., AND CASTIGLIONI, I. **Biomarkers from in vivo molecular imaging of breast cancer : pretreatment 18F-FDG PET predicts patient prognosis, and pretreatment DWI-MR predicts response to neoadjuvant chemotherapy.** *Magnetic Resonance Materials in Physics, Biology and Medicine* 30, 4 (2017), 359–373.
- [137] ASHRAF, A., GAONKAR, B., MIES, C., DEMICHELE, A., ROSEN, M., DAVATZIKOS, C., AND KONTOS, D. **Breast DCE-MRI kinetic heterogeneity tumor markers : Preliminary associations with neoadjuvant chemotherapy response.** *Transl. Oncol.* 8, 3 (June 2015), 154–162.

- [138] ZHANG, Q., YUAN, C., DAI, W., TANG, L., SHI, J., LI, Z., AND CHEN, M. **Evaluating pathologic response of breast cancer to neoadjuvant chemotherapy with computer-extracted features from contrast-enhanced ultrasound videos.** *Phys. Med.* 39 (July 2017), 156–163.
- [139] SADEGHI-NAINI, A., SANNACHI, L., TADAYYON, H., TRAN, W. T., SLODKOWSKA, E., TRUDEAU, M., GANDHI, S., PRITCHARD, K., KOLIOS, M. C., AND CZARNOTA, G. J. **Chemotherapy-Response monitoring of breast cancer patients using quantitative Ultrasound-Based Intra-Tumour heterogeneities.** *Sci. Rep.* 7, 1 (Sept. 2017), 10352.
- [140] THIBAUT, G., TUDORICA, A., AFZAL, A., CHUI, S. Y.-C., NAIK, A., TROXELL, M. L., KEMMER, K. A., OH, K. Y., ROY, N., JAFARIAN, N., HOLTORF, M. L., HUANG, W., AND SONG, X. **DCE-MRI texture features for early prediction of breast cancer therapy response.** *Tomography* 3, 1 (Mar. 2017), 23–32.
- [141] HA, S., PARK, S., BANG, J.-I., KIM, E.-K., AND LEE, H.-Y. **Metabolic radiomics for pretreatment 18F-FDG PET/CT to characterize locally advanced breast cancer : Histopathologic characteristics, response to neoadjuvant chemotherapy, and prognosis.** *Scientific Reports* 7, 1 (2017).
- [142] CHENG, L., ZHANG, J., WANG, Y., XU, X., ZHANG, Y., ZHANG, Y., LIU, G., AND CHENG, J. **Textural features of 18F-FDG PET after two cycles of neoadjuvant chemotherapy can predict pCR in patients with locally advanced breast cancer.** *Annals of Nuclear Medicine* 31, 7 (2017), 544–552.
- [143] DALE L BAILEY, DAVID W TOWNSEND, P. E. V., AND MAISEY, M. N. **Positron Emission Tomography.** Springer, 2005.
- [144] DEBENEDETTI, S., COWAN, C. E., KONNEKER, W. R., AND PRIMAKOFF, H. **On the angular distribution of Two-Photon annihilation radiation.** *Physical Review* 77, 2 (1950), 205–212.
- [145] CHERRY, S. R., SORENSON, J. A., AND PHELPS, M. E. **Physics in Nuclear Medicine E-Book.** Elsevier Health Sciences, Feb. 2012.
- [146] ZAIDI, H., AND KORAL, K. F. **Scatter modelling and compensation in emission tomography.** *Eur. J. Nucl. Med. Mol. Imaging* 31, 5 (May 2004), 761–782.
- [147] CASEY, M. E., AND NUTT, R. **A multicrystal two dimensional BGO detector system for positron emission tomography.** *IEEE Transactions on Nuclear Science* 33, 1 (1986), 460–463.
- [148] JONES, T., AND TOWNSEND, D. **History and future technical innovation in positron emission tomography.** *Journal of Medical Imaging* 4, 1 (2017), 011013.
- [149] SURTI, S., KUHN, A., WERNER, M. E., PERKINS, A. E., KOLTHAMMER, J., AND KARP, J. S. **Performance of philips gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities.** *J. Nucl. Med.* 48, 3 (Mar. 2007), 471–480.

- [150] CHERRY, S. R., DAHLBOM, M., AND HOFFMAN, E. J. **3D PET using a conventional multislice tomograph without septa**. *J. Comput. Assist. Tomogr.* 15, 4 (July 1991), 655–668.
- [151] TOWNSEND, D. W., GEISSBUHLER, A., DEFRISE, M., HOFFMAN, E. J., SPINKS, T. J., BAILEY, D. L., GILARDI, M. C., AND JONES, T. **Fully three-dimensional reconstruction for a PET camera with retractable septa**. *IEEE Trans. Med. Imaging* 10, 4 (1991), 505–512.
- [152] HOFFMAN, E. J., GUERRERO, T. M., GERMANO, G., DIGBY, W. M., AND DAHLBOM, M. **PET system calibrations and corrections for quantitative and spatially accurate images**. *IEEE Transactions on Nuclear Science* 36, 1 (1989), 1108–1112.
- [153] DEFRISE, M., TOWNSEND, D. W., BAILEY, D., GEISSBUHLER, A. M. C., AND JONES, T. **A normalization technique for 3D PET data**. *Physics in Medicine and Biology* 36, 7 (1991), 939–952.
- [154] BADAWI, R. D., AND MARSDEN, P. K. **Developments in component-based normalization for 3D PET**. *Physics in Medicine and Biology* 44, 2 (1999), 571–594.
- [155] BAI, B., LI, Q., HOLDSWORTH, C. H., ASMA, E., TAI, Y. C., CHATZIOANNOU, A., AND LEAHY, R. M. **Model-based normalization for iterative 3D PET image reconstruction**. *Physics in Medicine and Biology* 47, 15 (2002), 2773–2784.
- [156] BUVAT, I. **Quantification in emission tomography : Challenges, solutions, and performance**. *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 571, 1-2 (2007), 10–13.
- [157] BRASSE, D., KINAHAN, P. E., LARTIZIEN, C., COMTAT, C., CASEY, M., AND MICHEL, C. **Correction methods for random coincidences in fully 3D whole-body PET : impact on data and image quality**. *J. Nucl. Med.* 46, 5 (May 2005), 859–867.
- [158] WILLIAMS, C. W., CRABTREE, M. C., AND BURGESS, S. G. **Design and performance characteristics of a positron emission computed axial Tomograph–ECAT®-II**. *IEEE Transactions on Nuclear Science* 26, 1 (1979), 619–627.
- [159] BAILEY, D. L., AND MEIKLE, S. R. **A convolution-subtraction scatter correction method for 3D PET**. *Physics in Medicine and Biology* 39, 3 (1994), 411–424.
- [160] BARNEY, J. S., ROGERS, J. G., HARROP, R., AND HOVERATH, H. **Object shape dependent scatter simulations for PET**. *IEEE Transactions on Nuclear Science* 38, 2 (1991), 719–725.
- [161] OLLINGER, J. M. **Model-based scatter correction for fully 3D PET**. *Phys. Med. Biol.* 41, 1 (Jan. 1996), 153–176.
- [162] WERLING, A., BUBLITZ, O., DOLL, J., ADAM, L.-E., AND BRIX, G. **Fast implementation of the single scatter simulation algorithm and its use in iterative image reconstruction of PET data**. *Physics in Medicine and Biology* 47, 16 (2002), 2947–2960.
- [163] WATSON, C. C., CASEY, M. E., MICHEL, C., AND BENDRIEM, B. **Advances in scatter correction for 3D PET/CT**. *IEEE Symposium Conference Record Nuclear Science*. (2004).

- [164] FEI, G., RYOKO, Y., MITSUO, W., AND HUA-FENG, L. **An effective scatter correction method based on single scatter simulation for a 3D whole-body PET scanner.** *Chinese Physics B* 18, 7 (2009), 3066–3072.
- [165] ACCORSI, R., ADAM, L.-E., WERNER, M. E., AND KARP, J. S. **Optimization of a fully 3D single scatter simulation algorithm for 3D PET.** *Physics in Medicine and Biology* 49, 12 (2004), 2577–2598.
- [166] HOLDSWORTH, C. H., LEVIN, C. S., FARQUHAR, T. H., DAHLBOM, M., AND HOFFMAN, E. J. **Investigation of accelerated monte carlo techniques for PET simulation and 3D PET scatter correction.** *IEEE Transactions on Nuclear Science* 48, 1 (2001), 74–81.
- [167] HOLDSWORTH, C. H., LEVIN, C. S., JANECEK, M., DAHLBOM, M., AND HOFFMAN, E. J. **Performance analysis of an improved 3-D PET monte carlo simulation and scatter correction.** *IEEE Transactions on Nuclear Science* 49, 1 (2002), 83–89.
- [168] BARRET, O., ADRIAN CARPENTER, T., CLARK, J. C., ANSORGE, R. E., AND FRYER, T.D. **Monte carlo simulation and scatter correction of the GE advance PET scanner with SimSET and geant4.** *Physics in Medicine and Biology* 50, 20 (2005), 4823–4840.
- [169] LEVIN, C. S., DAHLBOM, M., AND HOFFMAN, E. J. **A monte carlo correction for the effect of compton scattering in 3-D PET brain imaging.** *IEEE Transactions on Nuclear Science* 42, 4 (1995), 1181–1185.
- [170] CASTIGLIONI, I., CREMONESI, O., GILARDI, M. C., BETTINARDI, V., RIZZO, G., SAVI, A., BELLOTTI, E., AND FAZIO, F. **Scatter correction techniques in 3D PET : a monte carlo evaluation.** *IEEE Transactions on Nuclear Science* 46, 6 (1999), 2053–2058.
- [171] VANDENBERGHE, S., VAN ELMBT, L., GUERCHAFT, M., CLEMENTEL, E., VERHAEGHE, J., BOL, A., LEMAHIEU, I., AND LONNEUX, M. **Optimization of time-of-flight reconstruction on philips GEMINI TF.** *Eur. J. Nucl. Med. Mol. Imaging* 36, 12 (Dec. 2009), 1994–2001.
- [172] GERMANO, G., AND HOFFMAN, E. J. **Investigation of count rate and deadtime characteristics of a high resolution PET system.** *Journal of Computer Assisted Tomography* 12, 5 (1988), 836–846.
- [173] FAHEY, F. H. **Data acquisition in PET imaging.** *J. Nucl. Med. Technol.* 30, 2 (June 2002), 39–49.
- [174] DILLENSEGER, J.-P., AND MOERSCHEL, E. **Guide des technologies de l'imagerie médicale et de la radiothérapie : quand la théorie éclaire la pratique.** Elsevier Masson, May 2009.
- [175] DAUBE-WITHERSPOON, M. E., AND MUEHLLEHNER, G. **Treatment of axial data in three-dimensional PET.** *J. Nucl. Med.* 28, 11 (Nov. 1987), 1717–1724.
- [176] LEWITT, R. M., MUEHLLEHNER, G., AND KARP, J. S. **Three-dimensional image reconstruction for PET by multi-slice rebinning and axial image filtering.** *Physics in Medicine and Biology* 39, 3 (1994), 321–339.

- [177] DEFRISE, M., KINAHAN, P. E., TOWNSEND, D. W., MICHEL, C., SIBOMANA, M., AND NEWPORT, D. F. **Exact and approximate rebinning algorithms for 3-D PET data.** *IEEE Transactions on Medical Imaging* 16, 2 (1997), 145–158.
- [178] FIDDY, M. A. **The radon transform and some of its applications.** *Optica Acta : International Journal of Optics* 32, 1 (1985), 3–4.
- [179] KINAHAN, P. E., AND ROGERS, J. G. **Analytic 3D image reconstruction using all detected events.** *IEEE Transactions on Nuclear Science* 36, 1 (1989), 964–968.
- [180] WERNICK, M. N., AND AARSVOLD, J. N. **Emission Tomography : The Fundamentals of PET and SPECT.** Elsevier, Dec. 2004.
- [181] DAHLBOM, M. **Physics of PET and SPECT imaging.**
- [182] BARRETT, H. H., AND SWINDELL, W. **Radiological Imaging : The Theory of Image Formation, Detection, and Processing.** Elsevier, Dec. 2012.
- [183] SHEPP, L. A., AND VARDI, Y. **Maximum likelihood reconstruction for emission tomography.** *IEEE Transactions on Medical Imaging* 1, 2 (1982), 113–122.
- [184] LANGE, K., AND CARSON, R. **EM reconstruction algorithms for emission and transmission tomography.** *J. Comput. Assist. Tomogr.* 8, 2 (Apr. 1984), 306–316.
- [185] LALUSH, D. S., AND TSUI, B. M. W. **Mean-variance analysis of block-iterative reconstruction algorithms modeling 3D detector response in SPECT.** *IEEE Nuclear Science Symposium Conference Record* (1007).
- [186] VEKLEROV, E., AND LLACER, J. **Stopping rule for the MLE algorithm based on statistical hypothesis testing.** *IEEE Trans. Med. Imaging* 6, 4 (1987), 313–319.
- [187] JOHNSON, V. E. **A note on stopping rules in EM-ML reconstructions of ECT images.** *IEEE Trans. Med. Imaging* 13, 3 (1994), 569–571.
- [188] HUDSON, H. M., AND LARKIN, R. S. **Accelerated image reconstruction using ordered subsets of projection data.** *IEEE Trans. Med. Imaging* 13, 4 (1994), 601–609.
- [189] LALUSH, D. S., AND TSUI, B. M. **Performance of ordered-subset reconstruction algorithms under conditions of extreme attenuation and truncation in myocardial SPECT.** *J. Nucl. Med.* 41, 4 (Apr. 2000), 737–744.
- [190] BROWNE, J., AND DE PIERRO, A. B. **A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography.** *IEEE Transactions on Medical Imaging* 15, 5 (1996), 687–699.
- [191] MESINA, C. T., BOELLAARD, R., JONGBLOED, G., VAN DER VAART, A. W., AND LAMMERTSMA, A. A. **Experimental evaluation of iterative reconstruction versus filtered back projection for 3D [15o]water PET activation studies using statistical parametric mapping analysis.** *Neuroimage* 19, 3 (July 2003), 1170–1179.
- [192] RAZIFAR, P., LUBBERINK, M., SCHNEIDER, H., LÅNGSTRÖM, B., BENGTSSON, E., AND BERGSTRÖM, M. **Non-isotropic noise correlation in PET data reconstructed by FBP but not by OSEM demonstrated using auto-correlation function.** *BMC Medical Imaging* 5, 1 (2005).

- [193] QI, J., AND LEAHY, R. M. **Iterative reconstruction techniques in emission computed tomography**. *Physics in Medicine and Biology* 51, 15 (2006), R541–R578.
- [194] BARRETT, H. H., WILSON, D. W., AND TSUI, B. M. W. **Noise properties of the EM algorithm. i. theory**. *Physics in Medicine and Biology* 39, 5 (1994), 833–846.
- [195] LODGE, M. A., RAHMIM, A., AND WAHL, R. L. **Simultaneous measurement of noise and spatial resolution in PET phantom images**. *Phys. Med. Biol.* 55, 4 (Feb. 2010), 1069–1081.
- [196] DERENZO, S. E., MOSES, W. W., HUESMAN, R. H., AND BUDINGER, T. F. **Critical instrumentation issues for < 2mm resolution, high sensitivity brain pet**.
- [197] DERENZO, S. E., BUDINGER, T. F., HUESMAN, R. H., CAHOON, J. L., AND VULETICH, T. **Imaging properties of a positron tomograph with 280 bgo crystals**. *IEEE Transactions on Nuclear Science* 28, 1 (1981), 81–89.
- [198] TER-POGOSSIAN, M. M., MULLANI, N. A., HOOD, J., HIGGLINS, C. S., AND MICHAEL CURRIE, C. **A multislice positron emission computed tomograph (PETT IV) yielding transverse and longitudinal images**. *Radiology* 128, 2 (1978), 477–484.
- [199] WIENHARD, K., SCHMAND, M., CASEY, M. E., BAKER, K., BAO, J., ERIKSSON, L., JONES, W. F., KNOESS, C., LENOX, M., LERCHER, M., LUK, P., MICHEL, C., REED, J. H., RICHERZHAGEN, N., TREFFERT, J., VOLLMAR, S., YOUNG, J. W., HEISS, W. D., AND NUTT, R. **The ECAT HRRT : performance and first clinical application of the new high resolution research tomograph**. *IEEE Nuclear Science Symposium. Conference Record (Cat. No.00CH37149)* (2000).
- [200] MOSES, W. W. **Fundamental limits of spatial resolution in PET**. *Nucl. Instrum. Methods Phys. Res. A* 648 Supplement 1 (Aug. 2011), S236–S240.
- [201] KOTASIDIS, F. A., AND ZAIDI, H. **Experimental evaluation and basis function optimization of the spatially variant image-space PSF on the ingenuity PET/MR scanner**. *Med. Phys.* 41, 6 (June 2014), 062501.
- [202] VISVIKIS, D., CHEZE LE REST, C., AND LAMARE, F. **Le mouvement respiratoire en imagerie fonctionnelle du cancer : une revue des effets et des méthodes de correction**. *Traitement du Signal* (2006).
- [203] NATIONAL ELECTRICAL MANUFACTURERS ASSOCIATION. **Performance measurements of positron emission tomographs**. Tech. Rep. NU2-2012, NEMA, 2012.
- [204] SORET, M., BACHARACH, S. L., AND BUVAT, I. **Partial-volume effect in PET tumor imaging**. *J. Nucl. Med.* 48, 6 (June 2007), 932–945.
- [205] PHELPS, M. E. **PET : Physics, Instrumentation, and Scanners**. Springer Science & Business Media, Dec. 2006.
- [206] KESSLER, R. M., ELLIS, J. R., AND EDEN, M. **Analysis of emission tomographic scan data : Limitations imposed by resolution and background**. *Journal of Computer Assisted Tomography* 8, 3 (1984), 514–522.
- [207] ROUSSET, O., RAHMIM, A., ALAVI, A., AND ZAIDI, H. **Partial volume correction strategies in PET**. *PET Clin.* 2, 2 (Apr. 2007), 235–249.

- [208] GARGOURI, S., MOUELI, A., SAYADI, M., LABIDI, S., MAHERSI, M., AND ZAYED, S. **Brief review of partial volume corrections in PET/CT for lung cancer assessment.** *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)* (2018).
- [209] SMITH, T. A. **Facilitative glucose transporter expression in human cancer tissue.** *Br. J. Biomed. Sci.* 56, 4 (1999), 285–292.
- [210] WARBURG, O. **On respiratory impairment in cancer cells.** *Science* 124, 3215 (Aug. 1956), 269–270.
- [211] CARSON, R. E. **Tracer Kinetic Modeling in PET.** Springer, 2005.
- [212] MANKOFF, D. A., MUZI, M., AND ZAIDIY, H. **Quantitative Analysis in Nuclear Oncologic Imaging.** Springer, 2006.
- [213] PATLAK, C. S., AND BLASBERG, R. G. **Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. generalizations.** *J. Cereb. Blood Flow Metab.* 5, 4 (Dec. 1985), 584–590.
- [214] HUNTER, G. J., HAMBERG, L. M., ALPERT, N. M., CHOI, N. C., AND FISCHMAN, A. J. **Simplified measurement of deoxyglucose utilization rate.** *J. Nucl. Med.* 37, 6 (June 1996), 950–955.
- [215] BUVAT, I. **Les limites du SUV.** *Médecine Nucléaire* 31, 4 (2007), 165–172.
- [216] VANDERHOEK, M., PERLMAN, S. B., AND JERAJ, R. **Impact of the definition of peak standardized uptake value on quantification of treatment response.** *J. Nucl. Med.* 53, 1 (Jan. 2012), 4–11.
- [217] KEYES, JR, J. W. **SUV : standard uptake or silly useless value?** *J. Nucl. Med.* 36, 10 (Oct. 1995), 1836–1839.
- [218] DU BOIS, D., AND DU BOIS, E. F. **A formula to estimate the approximate surface area if height and weight be known. 1916.** *Nutrition* 5, 5 (Sept. 1989), 303–11; discussion 312–3.
- [219] ZASADNY, K. R., AND WAHL, R. L. **Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-d-glucose : variations with body weight and a method for correction.** *Radiology* 189, 3 (1993), 847–850.
- [220] SUGAWARA, Y., ZASADNY, K. R., NEUHOFF, A. W., AND WAHL, R. L. **Reevaluation of the standardized uptake value for FDG : variations with body weight and methods for correction.** *Radiology* 213, 2 (Nov. 1999), 521–525.
- [221] ITO, H., KANNO, I., KATO, C., SASAKI, T., ISHII, K., OUCHI, Y., IIDA, A., OKAZAWA, H., HAYASHIDA, K., TSUYUGUCHI, N., ISHII, K., KUWABARA, Y., AND SENDA, M. **Database of normal human cerebral blood flow, cerebral blood volume, cerebral oxygen extraction fraction and cerebral metabolic rate of oxygen measured by positron emission tomography with 15o-labelled carbon dioxide or water, carbon monoxide and oxygen : a multicentre study in japan.** *Eur. J. Nucl. Med. Mol. Imaging* 31, 5 (May 2004), 635–643.

- [222] KUDOMI, N., WATABE, H., HAYASHI, T., AND IIDA, H. **Separation of input function for rapid measurement of quantitative CMRO₂ and CBF in a single PET scan with a dual tracer administration method.** *Physics in Medicine and Biology* 52, 7 (2007), 1893–1908.
- [223] MULLANI, N. A., AND GOULD, K. L. **First-pass measurements of regional blood flow with external detectors.** *J. Nucl. Med.* 24, 7 (July 1983), 577–581.
- [224] SAPIRSTEIN, L. A. **Fractionation of the cardiac output of rats with isotopic potassium.** *Circ. Res.* 4, 6 (Nov. 1956), 689–692.
- [225] MULLANI, N. A., GOLDSTEIN, R. A., GOULD, K. L., MARANI, S. K., FISHER, D. J., O'BRIEN, JR, H. A., AND LOBERG, M. D. **Myocardial perfusion with rubidium-82. i. measurement of extraction fraction and flow with external detectors.** *J. Nucl. Med.* 24, 10 (Oct. 1983), 898–906.
- [226] YOSHIDA, K., MULLANI, N., AND GOULD, K. L. **Coronary flow and flow reserve by PET simplified for clinical applications using rubidium-82 or nitrogen-13-ammonia.** *J. Nucl. Med.* 37, 10 (Oct. 1996), 1701–1712.
- [227] HUMBERT, O., BERRIOLO-RIEDINGER, A., COCHET, A., GAUTHIER, M., CHARON-BARRA, C., GUIU, S., DESMOULINS, I., TOUBEAU, M., DYGAÏ-COCHET, I., COUTANT, C., FUMOLEAU, P., AND BRUNOTTE, F. **Prognostic relevance at 5 years of the early monitoring of neoadjuvant chemotherapy using (18)F-FDG PET in luminal HER2-negative breast cancer.** *Eur. J. Nucl. Med. Mol. Imaging* 41, 3 (Mar. 2014), 416–427.
- [228] HUSTINX, R., BÉNARD, F., AND ALAVI, A. **Whole-body FDG-PET imaging in the management of patients with cancer.** *Semin. Nucl. Med.* 32, 1 (Jan. 2002), 35–46.
- [229] WOOD, K. A., HOSKIN, P. J., AND SAUNDERS, M. I. **Positron emission tomography in oncology : A review.** *Clinical Oncology* 19, 4 (2007), 237–255.
- [230] KOOLEN, B. B., VALDÉS OLMO, R. A., VOGEL, W. V., VRANCKEN PEETERS, M. J. T. F. D., RODENHUIS, S., RUTGERS, E. J. T., AND ELKHUIZEN, P. H. M. **Pre-chemotherapy 18F-FDG PET/CT upstages nodal stage in stage II-III breast cancer patients treated with neoadjuvant chemotherapy.** *Breast Cancer Res. Treat.* 141, 2 (Sept. 2013), 249–254.
- [231] COCHET, A., DYGAÏ-COCHET, I., RIEDINGER, J.-M., HUMBERT, O., BERRIOLO-RIEDINGER, A., TOUBEAU, M., GUIU, S., COUTANT, C., COUDERT, B., FUMOLEAU, P., AND BRUNOTTE, F. **18F-FDG PET/CT provides powerful prognostic stratification in the primary staging of large breast cancer when compared with conventional explorations.** *Eur. J. Nucl. Med. Mol. Imaging* 41, 3 (Mar. 2014), 428–437.
- [232] FUSTER, D., DUCH, J., PAREDES, P., VELASCO, M., MUÑOZ, M., SANTAMARÍA, G., FONTANILLAS, M., AND PONS, F. **Preoperative staging of large primary breast cancer with [18f]fluorodeoxyglucose positron emission tomography/computed tomography compared with conventional imaging procedures.** *J. Clin. Oncol.* 26, 29 (Oct. 2008), 4746–4751.

- [233] ALBERINI, J.-L., LEREBOURS, F., WARTSKI, M., FOURME, E., LE STANC, E., GONTIER, E., MADAR, O., CHEREL, P., AND PECKING, A. P. **18f-fluorodeoxyglucose positron emission tomography/computed tomography (FDG-PET/CT) imaging in the staging and prognosis of inflammatory breast cancer.** *Cancer* 115, 21 (Nov. 2009), 5038–5047.
- [234] GROHEUX, D., MORETTI, J.-L., BAILLET, G., ESPIE, M., GIACCHETTI, S., HINDIE, E., HENNEQUIN, C., VILCOQ, J.-R., CUVIER, C., TOUBERT, M.-E., FILMONT, J.-E., SARANDI, F., AND MISSET, J.-L. **Effect of (18)F-FDG PET/CT imaging in patients with clinical stage II and III breast cancer.** *Int. J. Radiat. Oncol. Biol. Phys.* 71, 3 (July 2008), 695–704.
- [235] CERMIK, T. F., MAVI, A., BASU, S., AND ALAVI, A. **Impact of FDG PET on the preoperative staging of newly diagnosed breast cancer.** *Eur. J. Nucl. Med. Mol. Imaging* 35, 3 (Mar. 2008), 475–483.
- [236] EUBANK, W. B., MANKOFF, D. A., TAKASUGI, J., VESSELLE, H., EARLY, J. F., SHANLEY, T. J., GRALOW, J. R., CHARLOP, A., ELLIS, G. K., LINDSLEY, K. L., AUSTIN-SEYMOUR, M. M., FUNKHOUSER, C. P., AND LIVINGSTON, R. B. **18fluorodeoxyglucose positron emission tomography to detect mediastinal or internal mammary metastases in breast cancer.** *Journal of Clinical Oncology* 19, 15 (2001), 3516–3523.
- [237] AUKEMA, T. S., STRAVER, M. E., PEETERS, M.-J. T. F. D. V., RUSSELL, N. S., GILHUIJS, K. G. A., VOGEL, W. V., RUTGERS, E. J. T., AND OLMOS, R. A. V. **Detection of extra-axillary lymph node involvement with FDG PET/CT in patients with stage II-III breast cancer.** *Eur. J. Cancer* 46, 18 (Dec. 2010), 3205–3210.
- [238] WOLFF, A. C., AND DAVIDSON, N. E. **Primary systemic therapy in operable breast cancer.** *J. Clin. Oncol.* 18, 7 (Apr. 2000), 1558–1569.
- [239] CRIPPA, F., SEREGNI, E., AGRESTI, R., CHIESA, C., PASCALI, C., BOGNI, A., DECISE, D., DE SANCTIS, V., GRECO, M., DAIDONE, M. G., AND BOMBARDIERI, E. **Association between [18f]fluorodeoxyglucose uptake and postoperative histopathology, hormone receptor status, thymidine labelling index and p53 in primary breast cancer : a preliminary observation.** *Eur. J. Nucl. Med.* 25, 10 (Oct. 1998), 1429–1434.
- [240] BOS, R., VAN DER HOEVEN, J. J. M., VAN DER WALL, E., VAN DER GROEP, P., VAN DIEST, P. J., COMANS, E. F. I., JOSHI, U., SEMENZA, G. L., HOEKSTRA, O. S., LAMMERTSMA, A. A., AND MOLTHOFF, C. F. M. **Biologic correlates of (18)fluorodeoxyglucose uptake in human breast cancer measured by positron emission tomography.** *J. Clin. Oncol.* 20, 2 (Jan. 2002), 379–387.
- [241] SHIMODA, W., HAYASHI, M., MURAKAMI, K., OYAMA, T., AND SUNAGAWA, M. **The relationship between FDG uptake in PET scans and biological behavior in breast cancer.** *Breast Cancer* 14, 3 (2007), 260–268.
- [242] GARCÍA GARCÍA-ESQUINAS, M., GARCÍA-SÁENZ, J. A., ARRAZOLA GARCÍA, J., ENRIQUE FUENTES FERRER, M., FURIÓ, V., RODRIGUEZ REY, C., ROMÁN, J. M., AND

- CARRERAS DELGADO, J. L. **18F-FDG PET-CT imaging in the neoadjuvant setting for stages II-III breast cancer : association of loco-regional SUVmax with classical prognostic factors.** *Q. J. Nucl. Med. Mol. Imaging* 58, 1 (Mar. 2014), 66–73.
- [243] INOUE, T., YUTANI, K., TAGUCHI, T., TAMAKI, Y., SHIBA, E., AND NOGUCHI, S. **Preoperative evaluation of prognosis in breast cancer patients by [(18)F]2-Deoxy-2-fluoro-D-glucose-positron emission tomography.** *J. Cancer Res. Clin. Oncol.* 130, 5 (May 2004), 273–278.
- [244] BABA, S., ISODA, T., MARUOKA, Y., KITAMURA, Y., SASAKI, M., YOSHIDA, T., AND HONDA, H. **Diagnostic and prognostic value of pretreatment SUV in 18F-FDG/PET in breast cancer : comparison with apparent diffusion coefficient from diffusion-weighted MR imaging.** *J. Nucl. Med.* 55, 5 (May 2014), 736–742.
- [245] CHAMPION, L., BRAIN, E., GIRAUDET, A.-L., LE STANC, E., WARTSKI, M., EDELINE, V., MADAR, O., BELLET, D., PECKING, A., AND ALBERINI, J.-L. **Breast cancer recurrence diagnosis suspected on tumor marker rising : value of whole-body 18FDG-PET/CT imaging and impact on patient management.** *Cancer* 117, 8 (Apr. 2011), 1621–1629.
- [246] AUKEMA, T. S., RUTGERS, E. J. T., VOGEL, W. V., TEERTSTRA, H. J., OLDENBURG, H. S., VRANCKEN PEETERS, M. T. F. D., WESSELING, J., RUSSELL, N. S., AND VALDÉS OLMOS, R. A. **The role of FDG PET/CT in patients with locoregional breast cancer recurrence : a comparison to conventional imaging techniques.** *Eur. J. Surg. Oncol.* 36, 4 (Apr. 2010), 387–392.
- [247] BÖGLER, O., AND MIKKELSEN, T. **Angiogenesis in glioma : molecular mechanisms and roadblocks to translation.** *Cancer J.* 9, 3 (May 2003), 205–213.
- [248] NEUFELD, G., AND KESSLER, O. **Pro-angiogenic cytokines and their role in tumor angiogenesis.** *Cancer Metastasis Rev.* 25, 3 (Sept. 2006), 373–385.
- [249] MÉRIDA, I., AND AVILA-FLORES, A. **Tumor metabolism : new opportunities for cancer therapy.** *Clin. Transl. Oncol.* 8, 10 (Oct. 2006), 711–716.
- [250] WEBER, W. A., SCHWAIGER, M., AND AVRIL, N. **Quantitative assessment of tumor metabolism using FDG-PET imaging.** *Nucl. Med. Biol.* 27, 7 (Oct. 2000), 683–687.
- [251] JOHNSON, B. F., CLAY, T. M., HOBEIKA, A. C., KIM LYERLY, H., AND MORSE, M. A. **Vascular endothelial growth factor and immunosuppression in cancer : current knowledge and potential for new therapy.** *Expert Opinion on Biological Therapy* 7, 4 (2007), 449–460.
- [252] HERBST, R. S. **Therapeutic options to target angiogenesis in human malignancies.** *Expert Opinion on Emerging Drugs* 11, 4 (2006), 635–650.
- [253] HUMBERT, O., LASSERRE, M., BERTAUT, A., FUMOLEAU, P., COUTANT, C., BRUNOTTE, F., AND COCHET, A. **Breast cancer blood flow and metabolism on Dual-Acquisition F-FDG PET : Correlation with tumor phenotype and neoadjuvant chemotherapy response.** *J. Nucl. Med.* 59, 7 (July 2018), 1035–1041.

- [254] MATEJ, S., AND LEWITT, R. M. **Practical considerations for 3-D image reconstruction using spherically symmetric volume elements.** *IEEE Transactions on Medical Imaging* 15, 1 (1996), 68–78.
- [255] HANSON, K. M., AND WECKSUNG, G. W. **Local basis-function approach to computed tomography.** *Applied Optics* 24, 23 (1985), 4028.
- [256] LEWITT, R. M. **Multidimensional digital image representations using generalized Kaiser–Bessel window functions.** *Journal of the Optical Society of America A* 7, 10 (1990), 1834.
- [257] DEFRISE, M., AND GULLBERG, G. T. **Image reconstruction.** *Physics in Medicine and Biology* 51, 13 (2006), R139–R154.
- [258] LEWITT, R. M. **Alternatives to voxels for image representation in iterative reconstruction algorithms.** *Phys. Med. Biol.* 37, 3 (Mar. 1992), 705–716.
- [259] HU, Z., WANG, W., GUALTIERI, E. E., HSIEH, Y. L., KARP, J. S., MATEJ, S., PARMA, M. J., TUNG, C. H., WALSH, E. S., WERNER, M., AND GAGNON, D. **An LOR-based fully-3d PET image reconstruction using a blob-basis function.** *2007 IEEE Nuclear Science Symposium Conference Record* (2007).
- [260] WANG, W., HU, Z., GUALTIERI, E. E., PARMA, M. J., WALSH, E. S., SEBOK, D., L. HSIEH, Y., H. TUNG, C., SONG, X., GRIESMER, J. J., KOLTHAMMER, J. A., POPESCU, L. M., WERNER, M., KARP, J. S., AND GAGNON, D. **Systematic and distributed Time-of-Flight list mode PET reconstruction.** *2006 IEEE Nuclear Science Symposium Conference Record* (2006).
- [261] LEWELLEN, T. K. **Time-of-flight PET.** *Seminars in Nuclear Medicine* 28, 3 (1998), 268–275.
- [262] MUEHLLEHNER, G., AND KARP, J. S. **Positron emission tomography.** *Phys. Med. Biol.* 51, 13 (July 2006), R117–37.
- [263] CONTI, M. **Focus on time-of-flight PET : the benefits of improved time resolution.** *European Journal of Nuclear Medicine and Molecular Imaging* 38, 6 (2011), 1147–1157.
- [264] VANDENBERGHE, S., VAN ELMBT, L., GUERCHAFT, M., CLEMENTEL, E., VERHAEGHE, J., BOL, A., LEMAHIEU, I., AND LONNEUX, M. **Optimization of time-of-flight reconstruction on philips GEMINI TF.** *Eur. J. Nucl. Med. Mol. Imaging* 36, 12 (Dec. 2009), 1994–2001.
- [265] VANDENBERGHE, S., MIKHAYLOVA, E., D’HOE, E., MOLLET, P., AND KARP, J. S. **Recent developments in time-of-flight PET.** *EJNMMI Phys* 3, 1 (Dec. 2016), 3.
- [266] VUNCKX, K., ZHOU, L., MATEJ, S., DEFRISE, M., AND NUYS, J. **Fisher information-based evaluation of image quality for time-of-flight PET.** *IEEE Trans. Med. Imaging* 29, 2 (Feb. 2010), 311–321.
- [267] VANDENBERGHE, S., MATEJ, S., DAUBE-WITHERSPOON, M. E., GUERCHAFT, M., VERHAEGHE, J., BOL, A., VAN ELMBT, L., LEMAHIEU, I., AND KARP, J. S. **Determining timing resolution from TOF-PET emission data.** *2007 IEEE Nuclear Science Symposium Conference Record* (2007).

- [268] DAUBE-WITHERSPOON, M. E., SURTI, S., MATEJ, S., WERNER, M., JAYANTHI, S., AND KARP, J. S. **Influence of Time-of-Flight kernel accuracy in TOF-PET reconstruction**. *2006 IEEE Nuclear Science Symposium Conference Record* (2006).
- [269] WERNER, M. E., SURTI, S., AND KARP, J. S. **Implementation and evaluation of a 3D PET single scatter simulation with TOF modeling**. *2006 IEEE Nuclear Science Symposium Conference Record* (2006).
- [270] WATSON, C. C. **Extension of single scatter simulation to scatter correction of Time-of-Flight PET**. *IEEE Nuclear Science Symposium Conference Record* (2005).
- [271] WATABE, H., IKOMA, Y., KIMURA, Y., NAGANAWA, M., AND SHIDAHARA, M. **PET kinetic analysis—compartmental model**. *Annals of Nuclear Medicine* 20, 9 (2006), 583–588.
- [272] SEROUL, P., AND SARRUT, D. **Vv : a viewer for the evaluation of 4d image registration**. In *MIDAS Journal (Medical Image Computing and Computer-Assisted Intervention MICCAI'2008, Workshop-Systems and Architectures for Computer Assisted Interventions)* (2008), Springer, New York, pp. 1–8.
- [273] WALKER, M. D., ASSELIN, M.-C., JULYAN, P. J., FELDMANN, M., TALBOT, P. S., JONES, T., AND MATTHEWS, J. C. **Bias in iterative reconstruction of low-statistics PET data : benefits of a resolution model**. *Phys. Med. Biol.* 56, 4 (Feb. 2011), 931–949.
- [274] JIAN, Y., PLANETA, B., AND CARSON, R. E. **Evaluation of bias and variance in low-count OSEM list mode reconstruction**. *Phys. Med. Biol.* 60, 1 (Jan. 2015), 15–29.
- [275] WALKER, M. D., MATTHEWS, J. C., ASSELIN, M.-C., WATSON, C. C., SALEEM, A., DICKINSON, C., CHARNLEY, N., JULYAN, P. J., PRICE, P. M., AND JONES, T. **Development and validation of a variance model for dynamic PET : uses in fitting kinetic data and optimizing the injected activity**. *Phys. Med. Biol.* 55, 22 (Nov. 2010), 6655–6672.
- [276] WALKER, M. D., MATTHEWS, J. C., ASSELIN, M.-C., SALEEM, A., DICKINSON, C., CHARNLEY, N., JULYAN, P. J., PRICE, P. M., AND JONES, T. **Optimization of the injected activity in dynamic 3D PET : a generalized approach using patient-specific NECs as demonstrated by a series of 15O-H2O scans**. *J. Nucl. Med.* 50, 9 (Sept. 2009), 1409–1417.
- [277] WIYAPORN, K., TOCHAROENCHAI, C., PUSUWAN, P., HIGUCHI, T., FUNG, G. S. K., FENG, T., PARK, M. J., AND TSUI, B. M. W. **Optimization of imaging protocols for myocardial blood flow (MBF) quantification with f-flurpiridaz PET**. *Phys. Med.* 42 (Oct. 2017), 127–134.
- [278] LEE, B. C., MOODY, J. B., WEINBERG, R. L., CORBETT, J. R., FICARO, E. P., AND MURTHY, V. L. **Optimization of temporal sampling for 82rubidium PET myocardial blood flow quantification**. *Journal of Nuclear Cardiology* 24, 5 (2017), 1517–1529.

- [279] PALARD-NOVELLO, X., BLIN, A.-L., LE JEUNE, F., GARIN, E., SALAÜN, P.-Y., DEVILLERS, A., GAMBAROTA, G., QUERELLOU, S., BOURGUET, P., AND SAINT-JALMES, H. **Optimization of temporal sampling for f-choline uptake quantification in prostate cancer assessment.** *EJNMMI Res.* 8, 1 (June 2018), 49.
- [280] ZHONG, M., AND KUNDU, B. K. **Optimization of a model corrected blood input function from dynamic FDG-PET images of small animal heart in vivo.** *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)* (2012).
- [281] LI, Y., AND KUNDU, B. K. **An improved optimization algorithm of the three-compartment model with spillover and partial volume corrections for dynamic FDG PET images of small animal hearts in vivo.** *Phys. Med. Biol.* 63, 5 (Feb. 2018), 055003.
- [282] WANG, G., CORWIN, M. T., OLSON, K. A., BADAWI, R. D., AND SARKAR, S. **Dynamic PET of human liver inflammation : impact of kinetic modeling with optimization-derived dual-blood input function.** *Phys. Med. Biol.* 63, 15 (July 2018), 155004.
- [283] TSOUMPAS, C., TURKHEIMER, F. E., AND THIELEMANS, K. **Study of direct and indirect parametric estimation methods of linear models in dynamic positron emission tomography.** *Medical Physics* 35, 4 (2008), 1299–1309.
- [284] TANG, J., KUWABARA, H., WONG, D. F., AND RAHMIM, A. **Direct 4D reconstruction of parametric images incorporating anato-functional joint entropy.** *Phys. Med. Biol.* 55, 15 (Aug. 2010), 4261–4272.
- [285] WANG, G., FU, L., AND QI, J. **Maximum a posteriori reconstruction of the patlak parametric image from sinograms in dynamic PET.** *Phys. Med. Biol.* 53, 3 (Feb. 2008), 593–604.
- [286] PETIBON, Y., RAKVONGTHAI, Y., EL FAKHRI, G., AND OUYANG, J. **Direct parametric reconstruction in dynamic PET myocardial perfusion imaging : in vivo studies.** *Phys. Med. Biol.* 62, 9 (May 2017), 3539–3565.
- [287] VERHAEGHE, J., GRAVEL, P., MIO, R., FUKASAWA, R., ROSA-NETO, P., P. SOUCY, J., THOMPSON, C. J., AND READER, A. J. **Motion-compensated fully 4D PET reconstruction using PET data supersets.** *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)* (2009).
- [288] ANGELIS, G. I., THIELEMANS, K., TZIORTZI, A. C., TURKHEIMER, F. E., AND TSOUMPAS, C. **Convergence optimization of parametric MLEM reconstruction for estimation of patlak plot parameters.** *Computerized Medical Imaging and Graphics* 35, 5 (2011), 407–416.
- [289] READER, A. J., AND VERHAEGHE, J. **4D image reconstruction for emission tomography.** *Physics in Medicine and Biology* 59, 22 (2014), R371–R418.
- [290] LUCIGNANI, G., AND LARSON, S. M. **Doctor, what does my future hold? the prognostic value of FDG-PET in solid tumours.** *Eur. J. Nucl. Med. Mol. Imaging* 37, 5 (May 2010), 1032–1038.

- [291] HATT, M., VISVIKIS, D., PRADIER, O., AND REST, C. C.-L. **Baseline 18F-FDG PET image-derived parameters for therapy response prediction in oesophageal cancer.** *European Journal of Nuclear Medicine and Molecular Imaging* 38, 9 (2011), 1595–1606.
- [292] WAHL, R. L., JACENE, H., KASAMON, Y., AND LODGE, M. A. **From RECIST to PERCIST : Evolving considerations for PET response criteria in solid tumors.** *J. Nucl. Med.* 50 Suppl 1 (May 2009), 122S–50S.
- [293] PAN, T., AND MAWLAWI, O. **PET/CT in radiation oncology.** *Med. Phys.* 35, 11 (Nov. 2008), 4955–4966.
- [294] FOX, J. L., RENGAN, R., O’MEARA, W., YORKE, E., ERDI, Y., NEHMEH, S., LEIBEL, S. A., AND ROSENZWEIG, K. E. **Does registration of PET and planning CT images decrease interobserver and intraobserver variation in delineating tumor volumes for non–small-cell lung cancer ?** *International Journal of Radiation Oncology*Biography*Physics* 62, 1 (2005), 70–75.
- [295] CALDWELL, C. B., MAH, K., UNG, Y. C., DANJOUX, C. E., BALOGH, J. M., NIMU GANGULI, S., AND EHRLICH, L. E. **Observer variation in contouring gross tumor volume in patients with poorly defined non-small-cell lung tumors on CT : the impact of 18 FDG-hybrid PET fusion.** *International Journal of Radiation Oncology*Biography*Physics* 51, 4 (2001), 923–931.
- [296] STEENBAKKERS, R. J. H. M., DUPPEN, J. C., FITTON, I., DEURLOO, K. E. I., ZIJP, L. J., COMANS, E. F. I., UITTERHOEVE, A. L. J., RODRIGUS, P. T. R., KRAMER, G. W. P., BUSSINK, J., DE JAEGER, K., BELDERBOS, J. S. A., NOWAK, P. J. C. M., VAN HERK, M., AND RASCH, C. R. N. **Reduction of observer variation using matched CT-PET for lung cancer delineation : a three-dimensional analysis.** *Int. J. Radiat. Oncol. Biol. Phys.* 64, 2 (Feb. 2006), 435–448.
- [297] FIORINO, C., RENI, M., BOLOGNESI, A., CATTANEO, G. M., AND CALANDRINO, R. **Intra- and inter-observer variability in contouring prostate and seminal vesicles : implications for conformal treatment planning.** *Radiother. Oncol.* 47, 3 (June 1998), 285–292.
- [298] GIRAUD, P., ELLES, S., HELFRE, S., DE RYCKE, Y., SERVOIS, V., CARETTE, M. F., ALZIEU, C., BONDIAU, P. Y., DUBRAY, B., TOUBOUL, E., HOUSSET, M., ROSENWALD, J. C., AND COSSET, J. M. **Conformal radiotherapy for lung cancer : different delineation of the gross tumor volume (GTV) by radiologists and radiation oncologists.** *Radiother. Oncol.* 62, 1 (Jan. 2002), 27–36.
- [299] FOSTER, B., BAGCI, U., MANSOOR, A., XU, Z., AND MOLLURA, D. J. **A review on segmentation of positron emission tomography images.** *Comput. Biol. Med.* 50 (July 2014), 76–96.
- [300] VAUCLIN, S., GARDIN, I., DOYEUX, K., HAPDEY, S., EDET-SANSON, A., AND VERA, P. **Segmentation des images TEP au 18F-FDG. principe et revue de la littérature.** *Médecine Nucléaire* 34, 6 (2010), 358–369.

- [301] LEE, J. A. **Segmentation of positron emission tomography images : some recommendations for target delineation in radiation oncology.** *Radiother. Oncol.* 96, 3 (Sept. 2010), 302–307.
- [302] HATT, M., LEE, J. A., SCHMIDTLEIN, C. R., NAQA, I. E., CALDWELL, C., DE BERNARDI, E., LU, W., DAS, S., GEETS, X., GREGOIRE, V., AND OTHERS. **Classification and evaluation strategies of auto-segmentation approaches for pet : Report of aapm task group no. 211.** *Medical physics* 44, 6 (2017), e1–e42.
- [303] NESTLE, U., KREMP, S., SCHAEFER-SCHULER, A., SEBASTIAN-WELSCH, C., HELLWIG, D., RÜBE, C., AND KIRSCH, C.-M. **Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer.** *J. Nucl. Med.* 46, 8 (Aug. 2005), 1342–1348.
- [304] MEIGNAN, M., SASANELLI, M., CASASNOVAS, R. O., LUMINARI, S., FIORONI, F., CORIANI, C., MASSET, H., ITTI, E., GOBBI, P. G., MERLI, F., AND VERSARI, A. **Metabolic tumour volumes measured at staging in lymphoma : methodological evaluation on phantom experiments and patients.** *Eur. J. Nucl. Med. Mol. Imaging* 41, 6 (June 2014), 1113–1122.
- [305] BRAMBILLA, M., MATHEOUD, R., SECCO, C., LOI, G., KRENGLI, M., AND INGLESE, E. **Threshold segmentation for PET target volume delineation in radiation treatment planning : the role of target-to-background ratio and target size.** *Med. Phys.* 35, 4 (Apr. 2008), 1207–1213.
- [306] ÖLLERS, M., BOSMANS, G., VAN BAARDWIJK, A., DEKKER, A., LAMBIN, P., TEULE, J., THIMISTER, W., RHAMY, A., AND DE RUYSSCHER, D. **The integration of PET-CT scans from different hospitals into radiotherapy treatment planning.** *Radiotherapy and Oncology* 87, 1 (2008), 142–146.
- [307] ERDI, Y. E., MAWLAWI, O., LARSON, S. M., IMBRIACO, M., YEUNG, H., FINN, R., AND HUMM, J. L. **Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding.** *Cancer* 80, S12 (1997), 2505–2509.
- [308] SCHAEFER, A., KREMP, S., HELLWIG, D., RÜBE, C., KIRSCH, C.-M., AND NESTLE, U. **A contrast-oriented algorithm for FDG-PET-based delineation of tumour volumes for the radiotherapy of lung cancer : derivation from phantom measurements and validation in patient data.** *Eur. J. Nucl. Med. Mol. Imaging* 35, 11 (Nov. 2008), 1989–1999.
- [309] RIEGEL, A. C., BUCCI, M. K., MAWLAWI, O. R., JOHNSON, V., AHMAD, M., SUN, X., LUO, D., CHANDLER, A. G., AND PAN, T. **Target definition of moving lung tumors in positron emission tomography : correlation of optimal activity concentration thresholds with object size, motion extent, and source-to-background ratio.** *Med. Phys.* 37, 4 (Apr. 2010), 1742–1752.
- [310] YAREMKO, B., RIAUKA, T., ROBINSON, D., MURRAY, B., ALEXANDER, A., MCEWAN, A., AND ROA, W. **Thresholding in PET images of static and moving targets.** *Phys. Med. Biol.* 50, 24 (Dec. 2005), 5969–5982.

- [311] BOUDRAA, A. O., AND ZAIDI, H. **Image segmentation techniques in nuclear medicine imaging**. *Quantitative Analysis in Nuclear Medicine Imaging* (2006), 308–357.
- [312] SORET, M., BACHARACH, S. L., AND BUVAT, I. **Partial-volume effect in PET tumor imaging**. *J. Nucl. Med.* 48, 6 (June 2007), 932–945.
- [313] DAISNE, J.-F., SIBOMANA, M., BOL, A., DOUMONT, T., LONNEUX, M., AND GRÉGOIRE, V. **Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios : influence of reconstruction algorithms**. *Radiother. Oncol.* 69, 3 (Dec. 2003), 247–250.
- [314] VAN DALEN, J. A., HOFFMANN, A. L., DICKEN, V., VOGEL, W. V., WIERING, B., RUERS, T. J., KARSSEMEIJER, N., AND OYEN, W. J. G. **A novel iterative method for lesion delineation and volumetric quantification with FDG PET**. *Nucl. Med. Commun.* 28, 6 (June 2007), 485–493.
- [315] VENEL, Y., GARHI, H., DE MURET, A., L. BAULIEU, J., BARILLOT, I., AND PRUNIER-AESCH, C. **Comparaison de six méthodes de segmentation du volume tumoral sur la 18F-FDG TEP-TDM avec le volume de référence anatomopathologique dans les cancers bronchopulmonaires non à petites cellules**. *Médecine Nucléaire* 32, 6 (2008), 339–353.
- [316] HSU, C.-Y., LIU, C.-Y., AND CHEN, C.-M. **Automatic segmentation of liver PET images**. *Comput. Med. Imaging Graph.* 32, 7 (Oct. 2008), 601–610.
- [317] KASS, M., WITKIN, A., AND TERZOPOULOS, D. **Snakes : Active contour models**. *International Journal of Computer Vision* 1, 4 (1988), 321–331.
- [318] BEUCHER, S., AND MEYER, F. **The morphological approach to segmentation : The watershed transformation**. *Mathematical Morphology in Image Processing* (2018), 433–481.
- [319] ERASMUS, J. J., GLADISH, G. W., BROEMELING, L., SABLOFF, B. S., TRUONG, M. T., HERBST, R. S., AND MUNDEN, R. F. **Interobserver and intraobserver variability in measurement of Non-Small-Cell carcinoma lung lesions : Implications for assessment of tumor response**. *Journal of Clinical Oncology* 21, 13 (2003), 2574–2582.
- [320] ABDOLI, M., DIERCKX, R. A. J. O., AND ZAIDI, H. **Contourlet-based active contour model for PET image segmentation**. *Med. Phys.* 40, 8 (Aug. 2013), 082507.
- [321] LI, H., THORSTAD, W. L., BIEHL, K. J., LAFOREST, R., SU, Y., SHOGHI, K. I., DONNELLY, E. D., LOW, D. A., AND LU, W. **A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours**. *Medical Physics* 35, 8 (2008), 3711–3721.
- [322] SETHIAN, J. A. **Level Set Methods and Fast Marching Methods : Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science**. Cambridge University Press, June 1999.
- [323] OSHER, S., AND SETHIAN, J. A. **Fronts propagating with curvature-dependent speed : Algorithms based on Hamilton-Jacobi formulations**. *Journal of Computational Physics* 79, 1 (1988), 12–49.

- [324] VINCENT, L., AND SOILLE, P. **Watersheds in digital spaces : an efficient algorithm based on immersion simulations.** *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 6 (1991), 583–598.
- [325] ROERDINK, J. B., AND MEIJSTER, A. **The watershed transform : Definitions, algorithms and parallelization strategies.** *Fundamenta informaticae* 41, 1, 2 (2000), 187–228.
- [326] GEETS, X., LEE, J. A., BOL, A., LONNEUX, M., AND GRÉGOIRE, V. **A gradient-based method for segmenting FDG-PET images : methodology and validation.** *Eur. J. Nucl. Med. Mol. Imaging* 34, 9 (Sept. 2007), 1427–1438.
- [327] WANET, M., LEE, J. A., WEYNAND, B., DE BAST, M., PONCELET, A., LACROIX, V., COCHE, E., GRÉGOIRE, V., AND GEETS, X. **Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer : A comparison with threshold-based approaches, CT and surgical specimens.** *Radiotherapy and Oncology* 98, 1 (2011), 117–125.
- [328] HATT, M., LAURENT, B., OUAHABI, A., FAYAD, H., TAN, S., LI, L., LU, W., JAOUEN, V., TAUBER, C., CZAKON, J., DRAPEJKOWSKI, F., DYRKA, W., CAMARASU-POP, S., CERVENANSKY, F., GIRARD, P., GLATARD, T., KAIN, M., YAO, Y., BARILLOT, C., KIROV, A., AND VISVIKIS, D. **The first MICCAI challenge on PET tumor segmentation.** *Med. Image Anal.* 44 (Feb. 2018), 177–195.
- [329] ADAMS, R., AND BISCHOF, L. **Seeded region growing.** *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16, 6 (1994), 641–647.
- [330] PAVLIDIS, T., AND T. LIOW, Y. **Integrating region growing and edge detection.** *Proceedings CVPR '88 : The Computer Society Conference on Computer Vision and Pattern Recognition* (1990).
- [331] MACQUEEN, J. **Some methods for classification and analysis of multivariate observations : UCLA, western manag. sci.** In *Inst. Working Paper 96* (1966).
- [332] DUNN, J. C. **A fuzzy relative of the ISODATA process and its use in detecting compact Well-Separated clusters.** *Journal of Cybernetics* 3, 3 (1973), 32–57.
- [333] HATT, M., LAMARE, F., BOUSSION, N., TURZO, A., COLLET, C., SALZENSTEIN, F., ROUX, C., JARRITT, P., CARSON, K., CHEZE-LE REST, C., AND VISVIKIS, D. **Fuzzy hidden markov chains segmentation for volume determination and quantitation in PET.** *Phys. Med. Biol.* 52, 12 (June 2007), 3467–3491.
- [334] ZHU, W., AND JIANG, T. **Automation segmentation of PET image for brain tumors.** *IEEE Nuclear Science Symposium. Conference Record (IEEE Cat. No.03CH37515)* (2003).
- [335] ZAIDI, H., ABDOLI, M., FUENTES, C. L., AND EL NAQA, I. M. **Comparative methods for PET image segmentation in pharyngolaryngeal squamous cell carcinoma.** *Eur. J. Nucl. Med. Mol. Imaging* 39, 5 (May 2012), 881–891.
- [336] BELHASSEN, S., AND ZAIDI, H. **A novel fuzzy c-means algorithm for unsupervised heterogeneous tumor quantification in PET.** *Med. Phys.* 37, 3 (Mar. 2010), 1309–1324.

- [337] DELIGNON, Y., MARZOUKI, A., AND PIECZYNSKI, W. **Estimation of generalized mixtures and its application in image segmentation.** *IEEE Trans. Image Process.* 6, 10 (1997), 1364–1375.
- [338] MONTGOMERY, D. W. G., AMIRA, A., AND ZAIDI, H. **Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model.** *Med. Phys.* 34, 2 (Feb. 2007), 722–736.
- [339] ARISTOPHANOUS, M., PENNEY, B. C., MARTEL, M. K., AND PELIZZARI, C. A. **A gaussian mixture model for definition of lung tumor volumes in positron emission tomography.** *Medical Physics* 34, 11 (2007), 4223–4235.
- [340] AYACHE, N., DELINGETTE, H., AND SERMESANT, M. **Functional Imaging and Modeling of the Heart : 5th International Conference, FIMH 2009 Nice, France, June 3-5, 2009 Proceedings.** Springer, May 2009.
- [341] HATT, M., CHEZE LE REST, C., ALBARGHACH, N., PRADIER, O., AND VISVIKIS, D. **PET functional volume delineation : a robustness and repeatability study.** *Eur. J. Nucl. Med. Mol. Imaging* 38, 4 (Apr. 2011), 663–672.
- [342] HATT, M., LE REST, C. C., TURZO, A., ROUX, C., AND VISVIKIS, D. **A fuzzy locally adaptive bayesian segmentation approach for volume determination in PET.** *IEEE Transactions on Medical Imaging* 28, 6 (2009), 881–893.
- [343] HATT, M., CHEZE LE REST, C., DESCOURT, P., DEKKER, A., DE RUYSSCHER, D., OELLERS, M., LAMBIN, P., PRADIER, O., AND VISVIKIS, D. **Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications.** *Int. J. Radiat. Oncol. Biol. Phys.* 77, 1 (May 2010), 301–308.
- [344] MATTES, D., HAYNOR, D. R., VESSELLE, H., LEWELLEN, T. K., AND EUBANK, W. **PET-CT image registration in the chest using free-form deformations.** *IEEE Trans. Med. Imaging* 22, 1 (Jan. 2003), 120–128.
- [345] MAES, F., COLLIGNON, A., VANDERMEULEN, D., MARCHAL, G., AND SUETENS, P. **Multimodality image registration by maximization of mutual information.** *IEEE Trans. Med. Imaging* 16, 2 (Apr. 1997), 187–198.
- [346] THÉVENAZ, P., AND UNSER, M. **Optimization of mutual information for multi-resolution image registration.** *IEEE Trans. Image Process.* 9, 12 (2000), 2083–2099.
- [347] KLEIN, S., STARING, M., MURPHY, K., VIERGEVER, M. A., AND PLUIM, J. **elastix : A toolbox for Intensity-Based medical image registration.** *IEEE Transactions on Medical Imaging* 29, 1 (2010), 196–205.
- [348] SHAMONIN, D. **Fast parallel image registration on CPU and GPU for diagnostic classification of alzheimer’s disease.** *Frontiers in Neuroinformatics* 7 (2013).
- [349] YUSHKEVICH, P. A., PIVEN, J., HAZLETT, H. C., SMITH, R. G., HO, S., GEE, J. C., AND GERIG, G. **User-guided 3D active contour segmentation of anatomical structures : significantly improved efficiency and reliability.** *Neuroimage* 31, 3 (July 2006), 1116–1128.

- [350] HANAHAN, D., AND WEINBERG, R. A. **Hallmarks of cancer : The next generation.** *Cell* 144, 5 (2011), 646–674.
- [351] HÖGLUND, M., GISSELSSON, D., SÄLL, T., AND MITELMAN, F. **Coping with complexity.** *Cancer Genetics and Cytogenetics* 135, 2 (2002), 103–109.
- [352] SHIBATA, D., SCHAEFFER, J., H. LI, Z., CAPELLA, G., AND PERUCHO, M. **Genetic heterogeneity of the c-k-ras locus in colorectal adenomas but not in adenocarcinomas.** *JNCI Journal of the National Cancer Institute* 85, 13 (1993), 1058–1063.
- [353] GERLINGER, M., ROWAN, A. J., HORSWELL, S., MATH, M., 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., NOHADANI, M., EKLUND, A. C., SPENCER-DENE, B., CLARK, G., PICKERING, L., STAMP, G., GORE, M., SZALLASI, Z., DOWNWARD, J., FUTREAL, P. A., AND SWANTON, C. **Intratumor heterogeneity and branched evolution revealed by multiregion sequencing.** *N. Engl. J. Med.* 366, 10 (Mar. 2012), 883–892.
- [354] VOGELSTEIN, B., PAPADOPOULOS, N., VELCULESCU, V. E., ZHOU, S., DIAZ, JR, L. A., AND KINZLER, K. W. **Cancer genome landscapes.** *Science* 339, 6127 (Mar. 2013), 1546–1558.
- [355] CAMPBELL, P. J., YACHIDA, S., MUDIE, L. J., STEPHENS, P. J., PLEASANCE, E. D., STEBBINGS, L. A., MORSBERGER, L. A., LATIMER, C., McLAREN, S., LIN, M.-L., MCBRIDE, D. J., VARELA, I., NIK-ZAINAL, S. A., LEROY, C., JIA, M., MENZIES, A., BUTLER, A. P., TEAGUE, J. W., GRIFFIN, C. A., BURTON, J., SWERDLOW, H., QUAIL, M. A., STRATTON, M. R., IACOBUZIO-DONAHUE, C., AND FUTREAL, P. A. **The patterns and dynamics of genomic instability in metastatic pancreatic cancer.** *Nature* 467, 7319 (Oct. 2010), 1109–1113.
- [356] O’CONNOR, J. P. B., ROSE, C. J., WATERTON, J. C., CARANO, R. A. D., PARKER, G. J. M., AND JACKSON, A. **Imaging intratumor heterogeneity : Role in therapy response, resistance, and clinical outcome.** *Clinical Cancer Research* 21, 2 (2015), 249–257.
- [357] WARBURG, O. **On the origin of cancer cells.** *Science* 123, 3191 (1956), 309–314.
- [358] MARUSYK, A., AND POLYAK, K. **Tumor heterogeneity : causes and consequences.** *Biochim. Biophys. Acta* 1805, 1 (Jan. 2010), 105–117.
- [359] GILLIES, R. J., ANDERSON, A. R., GATENBY, R. A., AND MORSE, D. L. **The biology underlying molecular imaging in oncology : from genome to anatome and back again.** *Clin. Radiol.* 65, 7 (July 2010), 517–521.
- [360] LAMBIN, P., RIOS-VELAZQUEZ, E., LEIJENAAR, R., CARVALHO, S., VAN STIPHOUT, R. G. P. M., GRANTON, P., ZEGERS, C. M. L., GILLIES, R., BOELLARD, R., DEKKER, A., AND AERTS, H. J. W. L. **Radiomics : extracting more information from medical images using advanced feature analysis.** *Eur. J. Cancer* 48, 4 (Mar. 2012), 441–446.

- [361] SEGAL, E., SIRLIN, C. B., OOI, C., ADLER, A. S., GOLLUB, J., CHEN, X., CHAN, B. K., MATCUK, G. R., BARRY, C. T., CHANG, H. Y., AND KUO, M. D. **Decoding global gene expression programs in liver cancer by noninvasive imaging.** *Nat. Biotechnol.* 25, 6 (June 2007), 675–680.
- [362] DIEHN, M., NARDINI, C., WANG, D. S., MCGOVERN, S., JAYARAMAN, M., LIANG, Y., ALDAPE, K., CHA, S., AND KUO, M. D. **Identification of noninvasive imaging surrogates for brain tumor gene-expression modules.** *Proc. Natl. Acad. Sci. U. S. A.* 105, 13 (Apr. 2008), 5213–5218.
- [363] LIANG, C., HUANG, Y., HE, L., CHEN, X., MA, Z., DONG, D., TIAN, J., LIANG, C., AND LIU, Z. **The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage I-II and stage III-IV colorectal cancer.** *Oncotarget* 7, 21 (May 2016), 31401–31412.
- [364] WU, W., PARMAR, C., GROSSMANN, P., QUACKENBUSH, J., LAMBIN, P., BUSSINK, J., MAK, R., AND AERTS, H. J. W. L. **Exploratory study to identify radiomics classifiers for lung cancer histology.** *Front. Oncol.* 6 (Mar. 2016), 71.
- [365] MU, W., CHEN, Z., LIANG, Y., SHEN, W., YANG, F., DAI, R., WU, N., AND TIAN, J. **Staging of cervical cancer based on tumor heterogeneity characterized by texture features on (18)F-FDG PET images.** *Phys. Med. Biol.* 60, 13 (July 2015), 5123–5139.
- [366] LI, H., ZHU, Y., BURNSIDE, E. S., HUANG, E., DRUKKER, K., HOADLEY, K. A., FAN, C., CONZEN, S. D., ZULEY, M., NET, J. M., SUTTON, E., WHITMAN, G. J., MORRIS, E., PEROU, C. M., JI, Y., AND GIGER, M. L. **Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set.** *NPJ Breast Cancer* 2 (May 2016).
- [367] LIMKIN, E. J., SUN, R., DERCLE, L., ZACHARAKI, E. I., ROBERT, C., REUZÉ, S., SCHERNBERG, A., PARAGIOS, N., DEUTSCH, E., AND FERTÉ, C. **Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology.** *Ann. Oncol.* 28, 6 (June 2017), 1191–1206.
- [368] GILLIES, R. J., KINAHAN, P. E., AND HRICAK, H. **Radiomics : Images are more than pictures, they are data.** *Radiology* 278, 2 (Feb. 2016), 563–577.
- [369] ZWANENBURG, A., LEGER, S., VALLIÈRES, M., AND LÖCK, S. **Image biomarker standardisation initiative.** *arXiv preprint arXiv :1612.07003* (2016).
- [370] GJESTEBY, L., DE MAN, B., JIN, Y., PAGANETTI, H., VERBURG, J., GIANTSOURI, D., AND WANG, G. **Metal artifact reduction in CT : Where are we after four decades ?** *IEEE Access* 4 (2016), 5826–5849.
- [371] GUDBJARTSSON, H., AND PATZ, S. **The rician distribution of noisy mri data.** *Magnetic Resonance in Medicine* 34, 6 (1995), 910–914.
- [372] SLED, J. G., ZIJDENBOS, A. P., AND EVANS, A. C. **A nonparametric method for automatic correction of intensity nonuniformity in MRI data.** *IEEE Transactions on Medical Imaging* 17, 1 (1998), 87–97.

- [373] VOVK, U., PERNUS, F., AND LIKAR, B. **A review of methods for correction of intensity inhomogeneity in MRI.** *IEEE Transactions on Medical Imaging* 26, 3 (2007), 405–421.
- [374] BALAFAR, M. A., RAMLI, A. R., SARIPAN, M. I., AND MASHOHOR, S. **Review of brain MRI image segmentation methods.** *Artificial Intelligence Review* 33, 3 (2010), 261–274.
- [375] BOUSSION, N., LE REST, C. C., HATT, M., AND VISVIKIS, D. **Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging.** *European Journal of Nuclear Medicine and Molecular Imaging* 36, 7 (2009), 1064–1075.
- [376] SORET, M., BACHARACH, S. L., AND BUVAT, I. **Partial-Volume effect in PET tumor imaging.** *Journal of Nuclear Medicine* 48, 6 (2007), 932–945.
- [377] SHAFIQ-UL-HASSAN, M., ZHANG, G. G., LATIFI, K., ULLAH, G., HUNT, D. C., BALAGURUNATHAN, Y., ABDALAH, M. A., SCHABATH, M. B., GOLDFOF, D. G., MACKIN, D., COURT, L. E., GILLIES, R. J., AND MOROS, E. G. **Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels.** *Med. Phys.* 44, 3 (Mar. 2017), 1050–1062.
- [378] SHIRI, I., RAHMIM, A., GHAFFARIAN, P., GERAMIFAR, P., ABDOLLAHI, H., AND BITARAFAN-RAJABI, A. **The impact of image reconstruction settings on 18F-FDG PET radiomic features : multi-scanner phantom and patient studies.** *Eur. Radiol.* 27, 11 (Nov. 2017), 4498–4509.
- [379] YAN, J., CHU-SHERN, J. L., LOI, H. Y., KHOR, L. K., SINHA, A. K., QUEK, S. T., THAM, I. W. K., AND TOWNSEND, D. **Impact of image reconstruction settings on texture features in 18F-FDG PET.** *J. Nucl. Med.* 56, 11 (Nov. 2015), 1667–1673.
- [380] ALTAZI, B. A., ZHANG, G. G., FERNANDEZ, D. C., MONTEJO, M. E., HUNT, D., WERNER, J., BIAGIOLI, M. C., AND MOROS, E. G. **Reproducibility of F18-FDG PET radiomic features for different cervical tumor segmentation methods, gray-level discretization, and reconstruction algorithms.** *J. Appl. Clin. Med. Phys.* 18, 6 (Nov. 2017), 32–48.
- [381] BAILLY, C., BODET-MILIN, C., COUESPEL, S., NECIB, H., KRAEBER-BODÉRE, F., ANSQUER, C., AND CARLIER, T. **Revisiting the robustness of PET-Based textural features in the context of Multi-Centric trials.** *PLoS One* 11, 7 (July 2016), e0159984.
- [382] YIP, S. S. F., AND HUGO J W. **Applications and limitations of radiomics.** *Physics in Medicine and Biology* 61, 13 (2016), R150–R166.
- [383] TIXIER, F., HATT, M., VALLA, C., FLEURY, V., LAMOUR, C., EZZOUHRI, S., INGRAND, P., PERDRISOT, R., VISVIKIS, D., AND LE REST, C. C. **Visual versus quantitative assessment of intratumor 18F-FDG PET uptake heterogeneity : prognostic value in non-small cell lung cancer.** *J. Nucl. Med.* 55, 8 (Aug. 2014), 1235–1241.
- [384] TIXIER, F., LE REST, C. C., HATT, M., ALBARGHACH, N., PRADIER, O., METGES, J.-P., CORCOS, L., AND VISVIKIS, D. **Intratumor heterogeneity characterized**

- by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J. Nucl. Med.* 52, 3 (Mar. 2011), 369–378.
- [385] EL NAQA, I., GRIGSBY, P., APTE, A., KIDD, E., DONNELLY, E., KHULLAR, D., CHAUDHARI, S., YANG, D., SCHMITT, M., LAFOREST, R., THORSTAD, W., AND DEASY, J. O. **Exploring feature-based approaches in PET images for predicting cancer treatment outcomes.** *Pattern Recognit.* 42, 6 (June 2009), 1162–1171.
- [386] YANG, F., THOMAS, M. A., DEHDASHTI, F., AND GRIGSBY, P. W. **Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer.** *Eur. J. Nucl. Med. Mol. Imaging* 40, 5 (May 2013), 716–727.
- [387] ORLHAC, F., SOUSSAN, M., CHOUAHNIA, K., MARTINOD, E., AND BUVAT, I. **18F-FDG PET-Derived textural indices reflect Tissue-Specific uptake pattern in Non-Small cell lung cancer.** *PLoS One* 10, 12 (Dec. 2015), e0145063.
- [388] LEIJENAAR, R. T. H., NALBANTOV, G., CARVALHO, S., VAN ELMPT, W. J. C., TROOST, E. G. C., BOELLAARD, R., AERTS, H. J. W. L., GILLIES, R. J., AND LAMBIN, P. **The effect of SUV discretization in quantitative FDG-PET radiomics : the need for standardized methodology in tumor texture analysis.** *Sci. Rep.* 5 (Aug. 2015), 11075.
- [389] CARLES, M., BACH, T., TORRES-ESPALLARDO, I., BALTAS, D., NESTLE, U., AND MARTÍ-BONMATÍ, L. **Significance of the impact of motion compensation on the variability of PET image features.** *Phys. Med. Biol.* 63, 6 (Mar. 2018), 065013.
- [390] VALLIÈRES, M., FREEMAN, C. R., SKAMENE, S. R., AND EL NAQA, I. **A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities.** *Phys. Med. Biol.* 60, 14 (July 2015), 5471–5496.
- [391] DERCLE, L., AMMARI, S., BATESON, M., DURAND, P. B., HASPINGER, E., MASSARD, C., JAUDET, C., VARGA, A., DEUTSCH, E., SORIA, J.-C., AND FERTÉ, C. **Limits of radiomic-based entropy as a surrogate of tumor heterogeneity : ROI-area, acquisition protocol and tissue site exert substantial influence.** *Sci. Rep.* 7, 1 (Aug. 2017), 7952.
- [392] SAHINER, B., CHAN, H. P., PETRICK, N., HELVIE, M. A., AND HADJIISKI, L. M. **Improvement of mammographic mass characterization using spiculation measures and morphological features.** *Med. Phys.* 28, 7 (July 2001), 1455–1465.
- [393] RANGAYYAN, R. M., MUDIGONDA, N. R., AND DESAUTELS, J. E. **Boundary modeling and shape analysis methods for classification of mammographic masses.** *Med. Biol. Eng. Comput.* 38, 5 (Sept. 2000), 487–496.
- [394] RANGAYYAN, R. M., EL-FARAMAWY, N. M., DESAUTELS, J. E., AND ALIM, O. A. **Measures of acutance and shape for classification of breast tumors.** *IEEE Trans. Med. Imaging* 16, 6 (Dec. 1997), 799–810.

- [395] HARALICK, R. M., SHANMUGAM, K., AND DINSTEIN, I. **Textural features for image classification**. *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*, 6 (1973), 610–621.
- [396] NIOCHE, C., ORLHAC, F., BOUGHDAD, S., REUZÉ, S., GOYA-OUTI, J., ROBERT, C., PELLOT-BARAKAT, C., SOUSSAN, M., FROUIN, F., AND BUVAT, I. **LIFEx : A free-ware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity**. *Cancer Research* 78, 16 (2018), 4786–4789.
- [397] HATT, M., MAJDOUB, M., VALLIERES, M., TIXIER, F., LE REST, C. C., GROHEUX, D., HINDIE, E., MARTINEAU, A., PRADIER, O., HUSTINX, R., PERDRISOT, R., GUILLEVIN, R., EL NAQA, I., AND VISVIKIS, D. **18F-FDG PET uptake characterization through texture analysis : Investigating the complementary nature of heterogeneity and functional tumor volume in a Multi-Cancer site patient cohort**. *Journal of Nuclear Medicine* 56, 1 (2015), 38–44.
- [398] AMADASUN, M., AND KING, R. **Textural features corresponding to textural properties**. *IEEE Transactions on systems, man, and Cybernetics* 19, 5 (1989), 1264–1274.
- [399] REUZÉ, S., SCHERNBERG, A., ORLHAC, F., SUN, R., CHARGARI, C., DERCLE, L., DEUTSCH, E., BUVAT, I., AND ROBERT, C. **Radiomics in nuclear medicine applied to radiation therapy : Methods, pitfalls, and challenges**. *International Journal of Radiation Oncology*Biophysics* 102, 4 (2018), 1117–1142.
- [400] GALLOWAY, M. M. **Texture analysis using gray level run lengths**. *Computer Graphics and Image Processing* 4, 2 (1975), 172–179.
- [401] THIBAUT, G., FERTIL, B., NAVARRO, C., PEREIRA, S., CAU, P., AND LEVY, N. **Texture indexes and gray level size zone matrix**. *Application to Cell Nuclei Classification*. *PRIP* (2009), 140–145.
- [402] LEIJENAAR, R. T. H., CARVALHO, S., VELAZQUEZ, E. R., VAN ELMPT, W. J. C., PARMAR, C., HOEKSTRA, O. S., HOEKSTRA, C. J., BOELLAARD, R., DEKKER, A. L. A. J., GILLIES, R. J., AERTS, H. J. W. L., AND LAMBIN, P. **Stability of FDG-PET radiomics features : an integrated analysis of test-retest and inter-observer variability**. *Acta Oncol.* 52, 7 (Oct. 2013), 1391–1397.
- [403] TIXIER, F., HATT, M., LE REST, C. C., LE POGAM, A., CORCOS, L., AND VISVIKIS, D. **Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET**. *J. Nucl. Med.* 53, 5 (May 2012), 693–700.
- [404] DESSEROIT, M.-C., TIXIER, F., WEBER, W. A., SIEGEL, B. A., LE REST, C. C., VISVIKIS, D., AND HATT, M. **Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of Non-Small cell lung cancer tumors : A repeatability analysis in a prospective multicenter cohort**. *Journal of Nuclear Medicine* 58, 3 (2017), 406–411.
- [405] VAN VELDEN, F. H. P., KRAMER, G. M., FRINGS, V., NISSEN, I. A., MULDER, E. R., DE LANGEN, A. J., HOEKSTRA, O. S., SMIT, E. F., AND BOELLAARD, R. **Repeatability of radiomic features in Non-Small-Cell lung cancer [(18)F]FDG-PET/CT**

- studies : Impact of reconstruction and delineation.** *Mol. Imaging Biol.* 18, 5 (Oct. 2016), 788–795.
- [406] GALAVIS, P. E., HOLLENSSEN, C., JALLOW, N., PALIWAL, B., AND JERAJ, R. **Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters.** *Acta Oncol.* 49, 7 (Oct. 2010), 1012–1016.
- [407] CORTES-RODICIO, J., SANCHEZ-MERINO, G., GARCIA-FIDALGO, M. A., AND TOBALINA-LARREA, I. **Identification of low variability textural features for heterogeneity quantification of 18F-FDG PET/CT imaging.** *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)* 35, 6 (2016), 379–384.
- [408] PFAEHLER, E., BEUKINGA, R. J., DE JONG, J. R., RIEMER H J, SLUMP, C. H., RUDI A J, AND BOELLAARD, R. **Repeatability of 18 F-FDG PET radiomic features : A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method.** *Medical Physics* 46, 2 (2019), 665–678.
- [409] NYFLOT, M. J., YANG, F., BYRD, D., BOWEN, S. R., SANDISON, G. A., AND KINAHAN, P. E. **Quantitative radiomics : impact of stochastic effects on textural feature analysis implies the need for standards.** *J Med Imaging (Bellingham)* 2, 4 (Oct. 2015), 041002.
- [410] YIP, S., MCCALL, K., ARISTOPHANOUS, M., CHEN, A. B., AERTS, H. J. W. L., AND BERBECO, R. **Comparison of texture features derived from static and respiratory-gated PET images in non-small cell lung cancer.** *PLoS One* 9, 12 (Dec. 2014), e115510.
- [411] OLIVER, J. A., BUDZEVICH, M., ZHANG, G. G., DILLING, T. J., LATIFI, K., AND MOROS, E. G. **Variability of image features computed from conventional and Respiratory-Gated PET/CT images of lung cancer.** *Transl. Oncol.* 8, 6 (Dec. 2015), 524–534.
- [412] GROOTJANS, W., TIXIER, F., VAN DER VOS, C. S., VRIENS, D., LE REST, C. C., BUSSINK, J., OYEN, W. J., DE GEUS-OEI, L.-F., VISVIKIS, D., AND VISSER, E. P. **The impact of optimal respiratory gating and image noise on evaluation of intratumor heterogeneity on 18f-fdg pet imaging of lung cancer.** *Journal of nuclear medicine* 57, 11 (2016), 1692–1698.
- [413] OLIVER, J. A., BUDZEVICH, M., HUNT, D., MOROS, E. G., LATIFI, K., DILLING, T. J., FEYGELMAN, V., AND ZHANG, G. **Sensitivity of image features to noise in conventional and Respiratory-Gated PET/CT images of lung cancer : Uncorrelated noise effects.** *Technol. Cancer Res. Treat.* 16, 5 (Oct. 2017), 595–608.
- [414] ORLHAC, F., BOUGHDAD, S., PHILIPPE, C., STALLA-BOURDILLON, H., NIOCHE, C., CHAMPION, L., SOUSSAN, M., FROUIN, F., FROUIN, V., AND BUVAT, I. **A postreconstruction harmonization method for multicenter radiomic studies in PET.** *J. Nucl. Med.* 59, 8 (Aug. 2018), 1321–1328.
- [415] JOHNSON, W. E., EVAN JOHNSON, W., LI, C., AND RABINOVIC, A. **Adjusting batch effects in microarray expression data using empirical bayes methods.** *Biostatistics* 8, 1 (2007), 118–127.

- [416] LUCIA, F., VISVIKIS, D., VALLIÈRES, M., DESSEROIT, M.-C., MIRANDA, O., ROBIN, P., BONAFFINI, P. A., ALFIERI, J., MASSON, I., MERVOYER, A., AND OTHERS. **External validation of a combined pet and mri radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy.** *European journal of nuclear medicine and molecular imaging* 46, 4 (2019), 864–877.
- [417] DISSAUX, G., VISVIKIS, D., DA-ANO, R., PRADIER, O., CHAJON, E., BARILLOT, I., DUVERGE, L., MASSON, I., ABGRAL, R., RIBEIRO, M.-J. S., AND OTHERS. **Pretreatment 18f-fdg pet/ct radiomics predict local recurrence in patients treated with stereotactic body radiotherapy for early-stage non-small cell lung cancer : A multicentric study.** *Journal of Nuclear Medicine* 61, 6 (2020), 814–820.
- [418] ORLHAC, F., SOUSSAN, M., MAISONOBE, J.-A., GARCIA, C. A., VANDERLINDEN, B., AND BUVAT, I. **Tumor texture analysis in 18F-FDG PET : relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis.** *J. Nucl. Med.* 55, 3 (Mar. 2014), 414–422.
- [419] HATT, M., TIXIER, F., CHEZE LE REST, C., PRADIER, O., AND VISVIKIS, D. **Robustness of intratumour 18F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma.** *Eur. J. Nucl. Med. Mol. Imaging* 40, 11 (Oct. 2013), 1662–1671.
- [420] CHENG, N.-M., FANG, Y.-H. D., LEE, L.-Y., CHANG, J. T.-C., TSAN, D.-L., NG, S.-H., WANG, H.-M., LIAO, C.-T., YANG, L.-Y., HSU, C.-H., AND YEN, T.-C. **Zone-size nonuniformity of 18F-FDG PET regional textural features predicts survival in patients with oropharyngeal cancer.** *European Journal of Nuclear Medicine and Molecular Imaging* 42, 3 (2015), 419–428.
- [421] FOLEY, K. G., HILLS, R. K., BERTHON, B., MARSHALL, C., PARKINSON, C., LEWIS, W. G., CROSBY, T. D. L., SPEZI, E., AND ROBERTS, S. A. **Development and validation of a prognostic model incorporating texture analysis derived from standardised segmentation of PET in patients with oesophageal cancer.** *Eur. Radiol.* 28, 1 (Jan. 2018), 428–436.
- [422] BERTHON, B., MARSHALL, C., EVANS, M., AND SPEZI, E. **ATLAAS : an automatic decision tree-based learning algorithm for advanced image segmentation in positron emission tomography.** *Phys. Med. Biol.* 61, 13 (July 2016), 4855–4869.
- [423] FORGÁCS, A., BÉRESOVÁ, M., GARAI, I., LASSEN, M. L., BEYER, T., DIFRANCO, M. D., BERÉNYI, E., AND BALKAY, L. **Impact of intensity discretization on textural indices of [F]FDG-PET tumour heterogeneity in lung cancer patients.** *Phys. Med. Biol.* 64, 12 (June 2019), 125016.
- [424] PRESOTTO, L., BETTINARDI, V., DE BERNARDI, E., BELLI, M. L., CATTANEO, G. M., BROGGI, S., AND FIORINO, C. **PET textural features stability and pattern discrimination power for radiomics analysis : An “ad-hoc” phantoms study.** *Fisica Medica* 50 (2018), 66–74.
- [425] BRANCHINI, M., ZORZ, A., ZUCCHETTA, P., BETTINELLI, A., DE MONTE, F., CECCHIN, D., AND PAIUSCO, M. **Impact of acquisition count statistics reduction**

- and SUV discretization on PET radiomic features in pediatric 18F-FDG-PET/MRI examinations. *Phys. Med.* 59 (Mar. 2019), 117–126.
- [426] ORLHAC, F., THÉZÉ, B., SOUSSAN, M., BOISGARD, R., AND BUVAT, I. **Multiscale texture analysis : From 18F-FDG PET images to histologic images.** *J. Nucl. Med.* 57, 11 (Nov. 2016), 1823–1828.
- [427] ORLHAC, F., NIOCHE, C., SOUSSAN, M., AND BUVAT, I. **Understanding changes in tumor texture indices in PET : A comparison between visual assessment and index values in simulated and patient data.** *J. Nucl. Med.* 58, 3 (Mar. 2017), 387–392.
- [428] ORLHAC, F. **Beyond the measurement of SUV in PET imaging : Properties and potential of the parameters of texture to characterize tumors.** Theses, Université Paris Sud - Paris XI, Sept. 2015.
- [429] HATT, M., TIXIER, F., VISVIKIS, D., AND CHEZE LE REST, C. **Radiomics in PET/CT : More than meets the eye ?** *J. Nucl. Med.* 58, 3 (Mar. 2017), 365–366.
- [430] SHEN, W.-C., CHEN, S.-W., LIANG, J.-A., HSIEH, T.-C., YEN, K.-Y., AND KAO, C.-H. **[18]fluorodeoxyglucose positron emission tomography for the textural features of cervical cancer associated with lymph node metastasis and histological type.** *Eur. J. Nucl. Med. Mol. Imaging* 44, 10 (Sept. 2017), 1721–1731.
- [431] LUCIA, F., VISVIKIS, D., DESSEROIT, M.-C., MIRANDA, O., MALHAIRE, J.-P., ROBIN, P., PRADIER, O., HATT, M., AND SCHICK, U. **Prediction of outcome using pretreatment 18 f-fdg pet/ct and mri radiomics in locally advanced cervical cancer treated with chemoradiotherapy.** *European journal of nuclear medicine and molecular imaging* 45, 5 (2018), 768–786.
- [432] VALLIÈRES, M., FREEMAN, C. R., SKAMENE, S. R., AND EL NAQA, I. **A radiomics model from joint fdg-pet and mri texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities.** *Physics in Medicine & Biology* 60, 14 (2015), 5471.
- [433] BROOKS, F. J., AND GRIGSBY, P. W. **The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake.** *J. Nucl. Med.* 55, 1 (Jan. 2014), 37–42.
- [434] AMIN, M. B., EDGE, S. B., GREENE, F. L., BYRD, D. R., BROOKLAND, R. K., WASHINGTON, M. K., GERSHENWALD, J. E., COMPTON, C. C., HESS, K. R., SULLIVAN, D. C., MILBURN JESSUP, J., BRIERLEY, J. D., GASPAR, L. E., SCHILSKY, R. L., BALCH, C. M., WINCHESTER, D. P., ASARE, E. A., MADERA, M., GRESS, D. M., AND MEYER, L. R. **AJCC Cancer Staging Manual.** Springer, Mar. 2018.
- [435] DESSEROIT, M.-C., TIXIER, F., WEBER, W. A., SIEGEL, B. A., LE REST, C. C., VISVIKIS, D., AND HATT, M. **Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of Non-Small cell lung cancer tumors : A repeatability analysis in a prospective multicenter cohort.** *J. Nucl. Med.* 58, 3 (2016), 406–411.

- [436] HATT, M., TIXIER, F., CHEZE LE REST, C., PRADIER, O., AND VISVIKIS, D. **Robustness of intratumour 18F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma.** *Eur. J. Nucl. Med. Mol. Imaging* 40, 11 (Oct. 2013), 1662–1671.
- [437] **Website.** R : A language and environment for statistical computing. Accessed : 2019-7-30.
- [438] NIOCHE, C., ORLHAC, F., BOUGHDAD, S., REUZÉ, S., GOYA-OUTI, J., ROBERT, C., PELLOT-BARAKAT, C., SOUSSAN, M., FROUIN, F., AND BUVAT, I. **Lifex : a free-ware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity.** *Cancer research* 78, 16 (2018), 4786–4789.
- [439] HATT, M., MAJDOUB, M., VALLIÈRES, M., TIXIER, F., LE REST, C. C., GROHEUX, D., HINDIÉ, E., MARTINEAU, A., PRADIER, O., HUSTINX, R., PERDRISOT, R., GUILLEVIN, R., EL NAQA, I., AND VISVIKIS, D. **18F-FDG PET uptake characterization through texture analysis : investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort.** *J. Nucl. Med.* 56, 1 (Jan. 2015), 38–44.
- [440] AIDE, N., SALOMON, T., BLANC-FOURNIER, C., GRELLARD, J.-M., LEVY, C., AND LASNON, C. **Implications of reconstruction protocol for histo-biological characterisation of breast cancers using FDG-PET radiomics.** *EJNMMI Res.* 8, 1 (Dec. 2018), 114.
- [441] GROHEUX, D., MAJDOUB, M., TIXIER, F., LE REST, C. C., MARTINEAU, A., MERLET, P., ESPIÉ, M., DE ROQUANCOURT, A., HINDIÉ, E., HATT, M., AND VISVIKIS, D. **Do clinical, histological or immunohistochemical primary tumour characteristics translate into different (18)F-FDG PET/CT volumetric and heterogeneity features in stage II/III breast cancer ?** *Eur. J. Nucl. Med. Mol. Imaging* 42, 11 (Oct. 2015), 1682–1691.
- [442] LEMARIGNIER, C., MARTINEAU, A., TEIXEIRA, L., VERCELLINO, L., ESPIÉ, M., MERLET, P., AND GROHEUX, D. **Correlation between tumour characteristics, SUV measurements, metabolic tumour volume, TLG and textural features assessed with 18F-FDG PET in a large cohort of oestrogen receptor-positive breast cancer patients.** *European Journal of Nuclear Medicine and Molecular Imaging* 44, 7 (2017), 1145–1154.
- [443] MOSCOSO, A., RUIBAL, Á., DOMÍNGUEZ-PRADO, I., FERNÁNDEZ-FERREIRO, A., HERRANZ, M., ALBAINA, L., ARGIBAY, S., SILVA-RODRÍGUEZ, J., PARDO-MONTERO, J., AND AGUIAR, P. **Texture analysis of high-resolution dedicated breast F-FDG PET images correlates with immunohistochemical factors and subtype of breast cancer.** *Eur. J. Nucl. Med. Mol. Imaging* 45, 2 (Feb. 2018), 196–206.
- [444] SON, S. H., KIM, D.-H., HONG, C. M., KIM, C.-Y., JEONG, S. Y., LEE, S.-W., LEE, J., AND AHN, B.-C. **Prognostic implication of intratumoral metabolic heterogeneity in invasive ductal carcinoma of the breast.** *BMC Cancer* 14 (Aug. 2014), 585.

- [445] SOUSSAN, M., ORLHAC, F., BOUBAYA, M., ZELEK, L., ZIOL, M., EDER, V., AND BUVAT, I. **Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer**. *PLoS One* 9, 4 (Apr. 2014), e94017.
- [446] SAMUEL, A. L. **Some studies in machine learning using the game of checkers. I**. *Computer Games I* (1988), 335–365.
- [447] SPECHT, D. F. **A general regression neural network**. *IEEE Transactions on Neural Networks* 2, 6 (1991), 568–576.
- [448] YPSILANTIS, P.-P., SIDDIQUE, M., SOHN, H.-M., DAVIES, A., COOK, G., GOH, V., AND MONTANA, G. **Predicting response to neoadjuvant chemotherapy with PET imaging using convolutional neural networks**. *PLOS ONE* 10, 9 (2015), e0137036.
- [449] WANG, L. **Heterogeneous data and big data analytics**. *Automatic Control and Information Sciences* 3, 1 (2017), 8–15.
- [450] POTDAR, K., S., T., AND D., C. **A comparative study of categorical variable encoding techniques for neural network classifiers**. *International Journal of Computer Applications* 175, 4 (2017), 7–9.
- [451] RICHMAN, M. B., TRAFALIS, T. B., AND ADRIANTO, I. **Missing Data Imputation Through Machine Learning Algorithms**. Springer, 2009.
- [452] HODGE, V., AND AUSTIN, J. **A survey of outlier detection methodologies**. *Artificial Intelligence Review* 22, 2 (2004), 85–126.
- [453] GUYON, I., AND ELISSEEFF, A. **An Introduction to Feature Extraction**. Springer, 2006.
- [454] JOHN, G. H., KOHAVI, R., AND PFLEGER, K. **Irrelevant features and the subset selection problem**. *Machine Learning Proceedings 1994* (1994), 121–129.
- [455] CHANDRASHEKAR, G., AND SAHIN, F. **A survey on feature selection methods**. *Computers & Electrical Engineering* 40, 1 (2014), 16–28.
- [456] KOHAVI, R., AND JOHN, G. H. **Wrappers for feature subset selection**. *Artificial Intelligence* 97, 1-2 (1997), 273–324.
- [457] WHITNEY, A. W. **A direct method of nonparametric measurement selection**. *IEEE Transactions on Computers C-20*, 9 (1971), 1100–1103.
- [458] TIBSHIRANI, R. **Regression shrinkage and selection via the lasso**. *Journal of the Royal Statistical Society : Series B (Methodological)* 58, 1 (1996), 267–288.
- [459] FRANK, L. E., AND FRIEDMAN, J. H. **A statistical view of some chemometrics regression tools**. *Technometrics* 35, 2 (1993), 109–135.
- [460] ZOU, H., AND HASTIE, T. **Regularization and variable selection via the elastic net**. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 67, 2 (2005), 301–320.
- [461] CHAWLA, N. V., BOWYER, K. W., HALL, L. O., AND KEGELMEYER, W. P. **SMOTE : Synthetic minority over-sampling technique**. *Journal of Artificial Intelligence Research* 16 (2002), 321–357.

- [462] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISSEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., AND OTHERS. **Scikit-learn : Machine learning in python**. *the Journal of machine Learning research* 12 (2011), 2825–2830.
- [463] COVER, T., AND HART, P. **Nearest neighbor pattern classification**. *IEEE Transactions on Information Theory* 13, 1 (1967), 21–27.
- [464] VAPNIK, V. N. **Estimation of dependences based on empirical data**. Springer Verlag, 1982.
- [465] SAFAVIAN, S. R., AND LANDGREBE, D. **A survey of decision tree classifier methodology**. *IEEE Transactions on Systems, Man, and Cybernetics* 21, 3 (1991), 660–674.
- [466] QUINLAN, J. R. **Induction of decision trees**. *Machine Learning* 1, 1 (1986), 81–106.
- [467] RUMELHART, D. E., MCCLELLAND, J. L., AND AU. **Parallel distributed processing**.
- [468] GIANCRISTOFARO, R. A., AND SALMASO, L. **Model performance analysis and model validation in logistic regression**. *Statistica* 63, 2 (2007), 375–396.
- [469] VALVERDE-ALBACETE, F. J., AND PELÁEZ-MORENO, C. **100% classification accuracy considered harmful : The normalized information transfer factor explains the accuracy paradox**. *PloS one* 9, 1 (2014).
- [470] WEN, J., SAMPER-GONZÁLEZ, J., ROUTIER, A., BOTTANI, S., DURRLEMAN, S., BURGOS, N., AND COLLIOT, O. **Beware of feature selection bias ! example on alzheimer’s disease classification from diffusion mri**.
- [471] LEMARIGNIER, C., MARTINEAU, A., TEIXEIRA, L., VERCELLINO, L., ESPIÉ, M., MERLET, P., AND GROHEUX, D. **Correlation between tumour characteristics, SUV measurements, metabolic tumour volume, TLG and textural features assessed with 18F-FDG PET in a large cohort of oestrogen receptor-positive breast cancer patients**. *European Journal of Nuclear Medicine and Molecular Imaging* 44, 7 (2017), 1145–1154.
- [472] LI, P., WANG, X., XU, C., LIU, C., ZHENG, C., FULHAM, M. J., FENG, D., WANG, L., SONG, S., AND HUANG, G. **18F-FDG PET/CT radiomic predictors of pathologic complete response (pCR) to neoadjuvant chemotherapy in breast cancer patients**. *European Journal of Nuclear Medicine and Molecular Imaging* 47, 5 (2020), 1116–1126.
- [473] HUMBERT, O., RIEDINGER, J.-M., VRIGNEAUD, J.-M., KANOUN, S., DYGAÏ-COCHET, I., BERRIOLO-RIEDINGER, A., TOUBEAU, M., DEPARDON, E., LASSERE, M., TISSERAND, S., FUMOLEAU, P., BRUNOTTE, F., AND COCHET, A. **18F-FDG PET-Derived tumor blood flow changes after 1 cycle of neoadjuvant chemotherapy predicts outcome in Triple-Negative breast cancer**. *J. Nucl. Med.* 57, 11 (Nov. 2016), 1707–1712.
- [474] DEIST, T. M., DANKERS, F. J., VALDES, G., WIJSMAN, R., HSU, I.-C., OBERIJE, C., LUSTBERG, T., VAN SOEST, J., HOEBERS, F., JOCHEMS, A., AND OTHERS. **Machine learning algorithms for outcome prediction in (chemo) radiotherapy : An empirical comparison of classifiers**. *Medical physics* 45, 7 (2018), 3449–3459.

- [475] LEGER, S., ZWANENBURG, A., PILZ, K., LOHAUS, F., LINGE, A., ZÖPHEL, K., KOTZERKE, J., SCHREIBER, A., TINHOFFER, I., BUDACH, V., AND OTHERS. **A comparative study of machine learning methods for time-to-event survival data for radio-mics risk modelling.** *Scientific reports* 7, 1 (2017), 1–11.
- [476] PARMAR, C., GROSSMANN, P., BUSSINK, J., LAMBIN, P., AND AERTS, H. J. **Machine learning methods for quantitative radiomic biomarkers.** *Scientific reports* 5 (2015), 13087.
- [477] GROHEUX, D., BIARD, L., LEHMANN-CHE, J., TEIXEIRA, L., BOUHIDEL, F. A., POIROT, B., BERTHEAU, P., MERLET, P., ESPIE, M., RESCHE-RIGON, M., AND OTHERS. **Tumor metabolism assessed by fdg-pet/ct and tumor proliferation assessed by genomic grade index to predict response to neoadjuvant chemotherapy in triple negative breast cancer.** *European journal of nuclear medicine and molecular imaging* 45, 8 (2018), 1279–1288.
- [478] BOELLAARD, R., DELGADO-BOLTON, R., OYEN, W. J., GIAMMARILE, F., TATSCH, K., ESCHNER, W., VERZIJLBERGEN, F. J., BARRINGTON, S. F., PIKE, L. C., WEBER, W. A., AND OTHERS. **Fdg pet/ct : Eanm procedure guidelines for tumour imaging : version 2.0.** *European journal of nuclear medicine and molecular imaging* 42, 2 (2015), 328–354.
- [479] BOELLAARD, R., DELGADO-BOLTON, R., OYEN, W. J. G., GIAMMARILE, F., TATSCH, K., ESCHNER, W., VERZIJLBERGEN, F. J., BARRINGTON, S. F., PIKE, L. C., WEBER, W. A., STROOBANTS, S., DELBEKE, D., DONOHUE, K. J., HOLBROOK, S., GRAHAM, M. M., TESTANERA, G., HOEKSTRA, O. S., ZIJLSTRA, J., VISSER, E., HOEKSTRA, C. J., PRUIM, J., WILLEMSSEN, A., ARENDS, B., KOTZERKE, J., BOCKISCH, A., BEYER, T., CHITI, A., KRAUSE, B. J., AND EUROPEAN ASSOCIATION OF NUCLEAR MEDICINE (EANM). **FDG PET/CT : EANM procedure guidelines for tumour imaging : version 2.0.** *Eur. J. Nucl. Med. Mol. Imaging* 42, 2 (Feb. 2015), 328–354.
- [480] STROTHER, S. C., CASEY, M. E., AND HOFFMAN, E. J. **Measuring PET scanner sensitivity : relating countrates to image signal-to-noise ratios using noise equivalent counts.** *IEEE Transactions on Nuclear Science* 37, 2 (1990), 783–788.
- [481] REILHAC, A., TOMEÏ, S., BUVAT, I., MICHEL, C., KEHEREN, F., AND COSTES, N. **Simulation-based evaluation of OSEM iterative reconstruction methods in dynamic brain PET studies.** *Neuroimage* 39, 1 (Jan. 2008), 359–368.
- [482] VAN VELDEN, F. H. P., KLOET, R. W., VAN BERCKEL, B. N. M., LAMMERTSMA, A. A., AND BOELLAARD, R. **Accuracy of 3-dimensional reconstruction algorithms for the High-Resolution research tomograph.** *Journal of Nuclear Medicine* 50, 1 (2008), 72–80.
- [483] KESSLER, R. M., ELLIS, J. R., AND EDEN, M. **Analysis of emission tomographic scan data : Limitations imposed by resolution and background.** *Journal of Computer Assisted Tomography* 8, 3 (1984), 514–522.
- [484] GOEDICKE, A., BERKER, Y., VERBURG, F. A., BEHRENDT, F. F., WINZ, O., AND MOTTAGHY, F. M. **Study-parameter impact in quantitative 90-yttrium PET**

- imaging for radioembolization treatment monitoring and dosimetry.** *IEEE Trans. Med. Imaging* 32, 3 (Mar. 2013), 485–492.
- [485] WILLOWSON, K. P., THE QUEST INVESTIGATOR TEAM, TAPNER, M., AND BAILEY, D. L. **A multicentre comparison of quantitative 90Y PET/CT for dosimetric purposes after radioembolization with resin microspheres.** *European Journal of Nuclear Medicine and Molecular Imaging* 42, 8 (2015), 1202–1222.
- [486] THOMAS, B. A., CUPLOV, V., BOUSSE, A., MENDES, A., THIELEMANS, K., HUTTON, B. F., AND ERLANDSSON, K. **PETPVC : a toolbox for performing partial volume correction techniques in positron emission tomography.** *Phys. Med. Biol.* 61, 22 (Nov. 2016), 7975–7993.
- [487] PRIETO, E., MARTÍ-CLIMENT, J. M., ARBIZU, J., GARRASTACHU, P., DOMÍNGUEZ, I., QUINCOES, G., GARCÍA-VELLOSO, M. J., LECUMBERRI, P., GÓMEZ-FERNÁNDEZ, M., AND RICHTER, J. A. **Evaluation of spatial resolution of a PET scanner through the simulation and experimental measurement of the recovery coefficient.** *Comput. Biol. Med.* 40, 1 (Jan. 2010), 75–80.

TABLE DES FIGURES

1.1	Taux d'incidence et de mortalité en France selon l'année (taux standardisé monde TSM). Source : Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine de 1990 à 2018. Volume 1 - Tumeurs solides	5
1.2	Carcinome canalaire <i>in situ</i> et infiltrant du sein. Source (adaptée) : www.cancer.org	7
1.3	Illustration de l'annihilation du positon et de l'électron. [145]	18
1.4	Interactions possibles des photons dans la matière en TEP, <i>source : wikipédia</i>	19
1.5	Illustration schématique du système de détection en TEP [148]	20
1.6	(a) Schéma illustrant deux modules adjacents dont des PMs se chevauchent, (b) photo d'un cristal LYSO pixelisé d'un GEMINI TF. sources : (a) Daube-Witherspoon (2009), (b) Philips.	21
1.7	(A) Détection en coïncidence de photons émis de la source (point rouge) (B) En absence de correction de temps de vol, la probabilité d'emplacement de la source est constante sur toute la LOR. (C) Avec l'information sur le temps de vol, la probabilité d'émission de la source suit une distribution gaussienne. [145]	22
1.8	Illustration des coïncidences détectées en TEP : les coïncidences vraies (A), diffusées (B), fortuites (C) et multiples (D). [143]	23
1.9	Illustration des modes de détections 2D et 3D	23
1.10	Exemple schématique de la construction d'un sinogramme, adapté de [174]	27
1.11	Exemple de projection faite à partir de l'intégration sur toutes les LORs parallèles pour un angle ϕ donné [180]	28
1.12	Organigramme d'un algorithme générique de reconstruction itérative [180]	29
1.13	Données simulées montrant l'effet de volume partiel sur un ensemble de sources sphériques (diamètre allant de 0,5 cm à 4,5 cm) de concentration d'activité égale (A). Ces données ont ensuite été filtrées (B) dans les trois dimensions avec un filtre gaussien (FWHM = 10 mm) afin de simuler un système de détection TEP. Comme le montrent les profils correspondants (C) plus la structure est petite, plus on observe une diminution de l'activité maximale ainsi qu'une augmentation de "l'étalement" spatial de l'activité. [205]	34
1.14	2-desoxy-2-18F-fluoro-D-glucose ou ^{18}F -FDG	35
1.15	Modèle cinétique permettant d'illustrer l'évolution de la fixation du ^{18}F -FDG	36

1.16	Modèle monocompartimental de Kety	38
2.1	Schéma simplifié du protocole TREN. Les données de la TEP au ^{18}F -FDG réalisée avant le début du traitement (TEP1) sont exploitées pour quantifier le métabolisme initial de la tumeur ainsi que la perfusion tumorale. Les données de la TEP au ^{18}F -FDG réalisée après une cure de chimiothérapie (TEP2) sont utilisées pour évaluer le métabolisme/la perfusion tumorale résiduel et la réponse métabolique/perfusionnelle en comparaison avec les résultats de la TEP1.	45
2.2	Schéma simplifié du protocole de la TEP1/TEP2	46
2.3	Exemples des trois volumes d'intérêts : tumeur (rouge), aorte ascendante (vert) et tissus sain/bkg (bleu), délimitées sur la TEP précoce statique de 8 min	51
2.4	Évolution de l'activité du traceur en fonction du temps, mesurée dans le VOI aortique lors d'une reconstruction dynamique avec un échantillonnage temporel de 5s (fonction d'entrée du modèle de Mullani). Détermination visuelle du temps t_m	52
2.5	Évolution de l'activité du traceur en fonction du temps, mesurée dans le VOI tumoral lors d'une reconstruction dynamique avec un échantillonnage temporel de 5s. La mesure de $Q(t_m)$ représente la quantité de traceur dans la tumeur au temps t_m	52
2.6	Résumé du processus de création de l'image paramétrique de perfusion.	53
2.7	Illustration visuelle de la qualité d'une image TEP statique reconstruite sur 8 minutes (gauche) et d'une image TEP dynamique reconstruite sur 10s (droite). En rouge le VOI de la tumeur, délimité sur l'image statique puis superposée à l'image dynamique.	54
2.8	Fantôme anthropomorphe NEMA IEC [203]	61
2.9	Seuils optimaux normalisés TS_{norm} ($TS/(mSUV_{70} - bkg)$) en fonction du contraste image pour les sphère de volume supérieur à 1,15 mL (a) et inférieur à 1,15 mL (b)	63
2.10	Illustration d'une tumeur proche du coeur chez une patiente de 67 ans atteinte d'un carcinome canalaire luminal du sein gauche. À gauche, un exemple où une partie du coeur de la patiente à été incluse durant la phase de dilatation, occasionnant un VOI finale (rouge) non représentative de la tumeur. À droite, le VOI corrigé (rouge) de la tumeur, segmenté après diminution du rayon de dilatation.	65
2.11	Exemple VOIs corrigée chez une patiente de 45 ans atteinte d'un carcinome canalaire Triple Négatif au sein droit.	66
3.1	Résumé des 10 grandes caractéristiques des tissus tumoraux [350]	68

3.2	Histogrammes simulés de deux images (A et B) sur lesquelles sont appliquées une DR (à gauche) et une DA (à droite) avec D égale à 16 bins. Avec une DR, les valeurs d'intensité sont normalisées pour chaque VOI avec un nombre fixe de bins, ce qui induit des largeurs de bins différentes entre les images. Avec une DA, le choix de bornes fixes permet de conserver la même échelle de SUV dans les images avec une largeur de bin fixe et un nombre de niveau d'intensité pour le rééchantillonnage fixe qui conduit à des bins «vides» comme illustré dans l'image B.	74
3.3	Illustrations des mesures de skewness et kurtosis, qui permettent d'évaluer l'écart de l'histogramme de l'image par rapport à une distribution Gaussienne.	75
3.4	Illustration du calcul de la matrice GLCM	76
3.5	Illustration des 13 directions possibles pour le calcul des matrices, [396]	76
3.6	Exemple de matrice NGLDM, adaptée de [399]. Sur cette illustration, les probabilités de présence des valeurs d'intensités sont ajoutées à la matrice.	77
3.7	Exemple de matrice GLRLM, adaptée de [399]	77
3.8	Exemple de matrice GLSZM adaptée de [399]	78
3.9	Évolution du nombre d'articles répertoriés dans pubmed sur l'analyse de texture, de 2010 à 2019	78
3.10	Distributions des indices de texture de perfusion (bleu) et métabolisme (vert) en fonction des paramètres TEP, après une discrétisation relative (ronds) et absolue (croix) de l'image, les indices ayant été calculés avec 13 GLCM et avec un rééchantillonnage de 32 bins pour les images de perfusion	93
3.11	Distributions des indices de texture de perfusion (bleu) et métabolisme (vert) en fonction des paramètres TEP, après une discrétisation relative (ronds) et absolue (croix) de l'image, les indices ayant été calculés avec 13 GLCM et avec un rééchantillonnage de 32 bins pour les images de perfusion	94
3.12	Distributions des indices de texture de perfusion (bleu) et métabolisme (vert) en fonction des paramètres TEP, après une discrétisation relative (ronds) et absolue (croix) de l'image, les indices ayant été calculés avec 1 GLCM et avec un rééchantillonnage de 32 bins pour les images de perfusion	95
3.13	Distributions des indices de texture de perfusion (bleu) et métabolisme (vert) en fonction des paramètres TEP, après une discrétisation relative (ronds) et absolue (croix) de l'image, les indices ayant été calculés avec 1 GLCM et avec un rééchantillonnage de 32 bins pour les images de perfusion	96
3.14	Distributions des indices de texture de perfusion en fonction du BF_{max} et du volume perfusé, après une discrétisation relative (gauche) et absolue (droite) de l'image en utilisant 16 (bleu), 32 (rouge) et 64 (vert) niveau d'intensité de rééchantillonnage.	99

3.15	Distributions des indices de texture de perfusion en fonction du BF_{\max} et du volume perfusé, après une discrétisation relative (gauche) et absolue (droite) de l'image en utilisant 16 (bleu), 32 (rouge) et 64 (vert) niveau d'intensité de ré-échantillonnage.	100
4.1	Exemple de courbes d'apprentissages qui illustrent le compromis biais-variance	125
4.2	Exemple d'une validation croisée avec $k = 10$	132
4.3	Matrice de confusion	133
4.4	Illustration d'une courbe ROC	135
4.5	Stratégie d'apprentissage	139
4.6	Exemple de courbes d'apprentissages réalisées par validations croisées pour le modèle C avant (a) et après (b) sélection de variables, qui représentent les scores d' <i>accuracy</i> d'entraînement (rouge) et de validation (vert), pour le quatrième <i>holdout</i> effectué parmi les 50.	141
4.7	Exemple de courbes d'apprentissages réalisées par validations croisées pour le modèle C+M avant (a) et après (b) sélection de variables, qui représentent les scores d' <i>accuracy</i> d'entraînement (rouge) et de validation (vert), pour le quatrième <i>holdout</i> effectué parmi les 50.	141
4.8	Exemple de courbes d'apprentissages réalisées par validations croisées pour le modèle C+M+BF avant (a) et après (b) sélection de variables, qui représentent les scores d' <i>accuracy</i> d'entraînement (rouge) et de validation (vert), pour le quatrième <i>holdout</i> effectué parmi les 50.	142
4.9	Résultats des <i>f1 scores</i> de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations pour les trois modèles, C, C+M et C+M+BF	144
4.10	Résultats des <i>f1 scores</i> de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations, effectuées avec une pondération de la classe minoritaire pour l'optimisation des trois modèles, C, C+M et C+M+BF	145
4.11	Résultats des <i>f1 scores</i> de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations, effectuées avec un suréchantillonnage par SMOTE de la classe minoritaire pour l'optimisation des trois modèles, C, C+M et C+M+BF	146
4.12	Résultats des <i>f1 scores</i> de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations, pour les modèles C+M et C+M+BF sans variable de texture	147
4.13	Résultats des <i>f1 scores</i> de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations, effectuées avec une pondération de la classe minoritaire pour l'entraînement des deux modèles C+M et C+M+BF sans variable de texture	148

4.14	Résultats des <i>f1 scores</i> de généralisation (a) et d'entraînement (b) moyennés sur les 50 itérations, effectuées avec un suréchantillonnage par SMOTE de la classe minoritaire pour l'entraînement des deux modèles C+M et C+M+BF sans variable de texture.	149
C.1	Image TEP du fantôme NEMA IEC avec contraste sphère-fond de 10 :1. En couleur les 12 régions d'intérêts tracées pour l'évaluation du bruit de fond (bkg). Image reconstruite avec l'algorithme BLOB-OS-TF, 3 itérations, 33 sous-ensembles.	221
C.2	Coefficients de recouvrement et coefficient de variation en fonction du diamètre des sphères (reconstruction avec 3 itérations, 33 sous-ensemble et un noyau de TOF de 14,1 cm). T-0.50 : seuillage fixe à 50% de l'intensité maximale, T-0.41 : seuillage fixe à 41% de l'intensité maximale, Initiale : croissance de région avec le volume réel des sphères comme critère d'arrêt, AT-0.41 : seuillage adaptatif (Imalytics) avec un seuil fixe de 41% de l'intensité maximale pour toutes les sphères, AT-var : seuillage adaptatif (Imalytics) avec un seuil variable pour chaque sphères, DILAT-1 : dilatation avec contrainte sur l'aspect sphérique, DILAT-1 : dilatation avec contrainte sur la différence du volume segmenté, NEMA : segmentation utilisée dans les recommandations des tests EANM [478].	223
C.3	(a) Illustration des différents contrastes et des répliques (b) aux différents temps d'acquisitions.	224
C.4	Comparaison des coefficients de variation en fonction du temps d'acquisition entre les mesures réalisées sur les images TEP et le modèle idéal de Poisson. Illustration à droite de l'image de la différence entre un bruit poissonien présent sur l'image TEP de 3h30 et le bruit d'une image avec un temps d'acquisition de 5s.	225
C.5	Coefficients de recouvrement moyennés sur les dix répliques en fonction du temps d'acquisition pour les sphères de 10 (a) et de 37 (b) mm de diamètre et en fonction de la taille des sphères (c) pour tous les temps d'acquisition (reconstructions avec 3 itérations et 33 sous-ensembles).	225
C.6	Biais du coefficient de recouvrement pour les dix répliques pour la plus petite sphère ($\varnothing = 10\text{mm}$) pour des temps d'acquisition de 5s (a) et 10s (b) (reconstructions avec 3 itérations et 33 sous-ensembles).	226
C.7	Influence du nombre d'itérations sur le coefficient de recouvrement en fonction de la taille des sphères (a) et sur le coefficient de variation (b) (reconstruction avec 33 sous-ensembles, un noyau de TOF de 14,1 cm et un paramètre de relaxation λ de 1).	227
C.8	Influence du nombre de sous-ensembles sur le coefficient de recouvrement en fonction du diamètre des sphères (a) et sur le coefficient de variation (b) (reconstruction avec 2 itérations, un noyau de TOF de 14,1 cm et un paramètre de relaxation λ de 1).	227

C.9	Influence du nombre d'itérations et de sous-ensembles ($n \times s$) sur le coefficient de recouvrement (a) et sur le coefficient de variation (b), pour des images reconstruites avec un noyau de TOF de 14,1 cm et un paramètre de relaxation λ de 1. Les modèles indiqués par les flèches correspondent au modèle de convergence $n \times s = 180$ avec 2 itérations et 90 sous-ensembles et au modèle de convergence $n \times s = 1 \times 33$ comprenant 1 itération et 33 sous-ensembles.	228
C.10	(a) Coefficients de recouvrement en fonction du diamètre des sphères pour différentes valeurs du paramètre de relaxation λ , (b) coefficients de variation en fonction du paramètre de relaxation λ (reconstruction avec 3 itérations, 33 sous-ensemble et un noyau de TOF de 14,1 cm).	229
C.11	(a) Coefficients de recouvrement en fonction du diamètre des sphères pour différentes valeurs de noyau de TOF, (b) coefficients de variation en fonction de la largeur du noyau de TOF (reconstruction avec 7 itérations, 8 sous-ensembles et un paramètre de relaxation de 1).	230
C.12	Graphique récapitulatif des coefficients de recouvrement en fonction des coefficients de variation pour tous les ensembles de paramètres de reconstruction testés. Illustration à droite de l'image reconstruite avec les paramètres standards et les paramètres optimisés (HQ : noyau de TOF de 18,7 cm, SB : paramètre de relaxation de 0,5).	231
C.13	Illustration du fantôme NEMA IEC avec des contrastes infinis	232
C.14	Profils de la source ligne, selon la direction radiale et tangentielle pour différentes coupes selon la direction axiale	233
C.15	écarts des Chi^2 en fonction des FWHM	234
C.16	Valeurs des RC en fonction du diamètre des sphères après différentes méthodes après différentes méthodes de correction de l'EVP	234
C.17	Graphique récapitulatif des coefficients de recouvrement en fonction des coefficients de variation pour les différentes corrections d'EVP testées.	235
C.18	évaluation de la qualité visuelle et des profils d'intensité pour la reconstruction initiale (bleu), la reconstruction optimisée (rouge) et la reconstruction optimisée corrigée de l'EVP (vert)	236

LISTE DES TABLES

1.1	Grades SBR modifiés selon Eston et Ellis.	7
1.2	Classification de Chevallier pour l'évaluation histologique de la réponse à la chimiothérapie néoadjuvante des cancers du sein.	12
1.3	Caractéristiques des principaux cristaux scintillants en TEP	20
3.1	Caractéristiques des patientes et des tumeurs	92
3.2	Coefficient de corrélation de Spearman (ρ) entre les indices métaboliques <i>entropy</i> et <i>energy</i> , et le VP et le BF_{\max}	97
3.3	Coefficient de corrélation de Spearman (ρ) entre les indices métaboliques <i>entropy</i> et <i>energy</i> , et le VM et le SUV_{\max}	98
3.4	Résultats des tests de Mann-Whitney entre les groupes pCR et non-pCR pour l'ensemble de la population. Les valeurs des p-values significatives ($< 0,05$, après correction de tests multiples) sont indiquées avec la valeur moyenne des indices de texture pour chaque groupe entre parenthèses lorsqu'ils sont significatifs. DA : discrétisation absolue, GLCM : matrice de co-occurrence de niveau de gris, IDM : inverse different moment, M : métabolisme, ns : non significatif, DR : discrétisation relative.	101
3.5	Rapports de corrélation (η^2) de l'AFD descriptive pour l'ensemble de la population, pour les sous-groupes HER2 et TN. Les rapports de corrélation maximaux sont écrits en gras pour chaque groupe de patients. La valeur marquée d'une * a été arrondie au supérieur. DA : discrétisation absolue, BF : <i>blood flow</i> , GLCM : matrice de co-occurrence du niveau de gris, M : métabolisme, DR : discrétisation relative.	102
3.6	Médiane et interquartiles des volumes tumoraux du phénotype TN en fonction de différents seuils volumiques.	102
3.7	Rapports de corrélation (η^2) de l'AFD descriptive pour les différents sous-groupes TN	103
3.8	Valeurs absolues des poids canoniques de l'AFD descriptive pour les différents paramètres TEP. Les 10 poids canoniques les plus élevés pour chaque population sont indiqués en gras. BF : <i>blood flow</i> , IDM : <i>inverse different moment</i> , M : métabolisme, VM, volume métabolique, SUV : <i>standardized uptake value</i> , TLG : <i>total lesion glycolysis</i>	104

4.1	Caractéristiques des patientes et des tumeurs	140
4.2	Résultats d'entraînement et de test moyennés sur les 50 <i>holdouts</i> pour chacun des trois modèles les plus performants des configurations C, C+M et C+M+BF et résultats du dummy classifier.	144
4.3	Résultats d'entraînement et de test moyennés sur les 50 <i>holdouts</i> pour chacun des trois modèles les plus performants des configurations C, C+M et C+M+BF, effectués avec une pondération de la classe minoritaire pour l'optimisation des modèles.	145
4.4	Résultats d'entraînement et de test moyennés sur les 50 <i>holdouts</i> pour chacun des trois modèles les plus performants des configurations C, C+M et C+M+BF, effectués avec un suréchantillonnage par SMOTE de la classe minoritaire pour l'optimisation des modèles.	146
4.5	Résultats d'entraînement et de test moyennés sur les 50 <i>holdouts</i> pour les deux modèles les plus performants des configurations C+M et C+M+BF effectués sans les variables de textures.	148
4.6	Résultats d'entraînement et de tests moyennés sur les 50 <i>holdouts</i> pour les deux modèles les plus performants des configurations C+M et C+M+BF, effectués sans variables de textures et avec une pondération de la classe minoritaire pour l'optimisation des modèles.	149
4.7	Résultats d'entraînement et de tests moyennés sur les 50 <i>holdouts</i> pour les deux modèles les plus performants des configurations C+M et C+M+BF, effectués sans variables de textures et avec un suréchantillonnage par SMOTE de la classe minoritaire pour l'optimisation des modèles.	149
B.1	Coefficients de corrélation de Spearman (ρ) entre le BF_{\max} , BF_{mean} , le VP et les indices de texture de perfusion.	218
B.2	Coefficients de corrélation de Spearman (ρ) entre le SUV_{\max} , le SUV_{mean} et le VM et les indices de texture métaboliques	219

ANNEXES

INDICES DE TEXTURE DE LA GLCM

Entropy : indice relié au désordre présent dans l'image (quantifie la régularité des différences). L'*Entropy* augmente lorsque la distribution des intensités présentes dans l'image est aléatoire.

$$Entropy = - \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} P(i, j) \ln (P(i, j)) \quad (A.1)$$

Energy ou *Uniformity* ou *Angular Second Moment* : quantifie l'uniformité de la texture. L'énergie augmente si le nombre de paires de voxel de même intensités augmente.

$$Energy = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} (P(i, j))^2 \quad (A.2)$$

Homogeneity : indice qui quantifie l'homogénéité de la distribution des intensités dans l'image, il vaut 1 pour une image constante.

$$Homogeneity = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} \frac{P(i, j)}{1 + |i - j|^2} \quad (A.3)$$

Inverse Different Moment (IDM) ou *local homogeneity* : quantifie l'homogénéité dans une image et sera élevé si les intensités de chaque paire de voxels sont similaires. Cet indice est similaire à l'*homogeneity* mais donne moins de poids aux éléments qui s'éloignent de la diagonale de la matrice.

$$Inverse\ Different\ Moment = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} \frac{P(i, j)}{1 + (i - j)^2} \quad (A.4)$$

Correlation : mesure la dépendance linéaire des niveaux d'intensité dans l'image.

$$Correlation = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} \frac{P(i, j)(i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} \quad (A.5)$$

Dissimilarity : quantifie les variations d'intensités des paires de voxels.

$$Dissimilarity = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} P(i, j) |i - j| \quad (A.6)$$

Inertia ou *Contrast* : quantifie le contraste dans une image. Cet indice sera élevé si la différence d'intensité entre deux voxels d'une même paire est grande.

$$Inertia = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} (i - j)^2 P(i, j) \quad (A.7)$$

Cluster Prominence : mesure l'asymétrie de la matrice de cooccurrence et sera élevé si la GLCM est asymétrique.

$$Cluster Prominence = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} P(i, j) [i + j + \mu_i + \mu_j]^4 \quad (A.8)$$

Cluster Shade : mesure le manque de symétrie de la matrice GLCM.

$$Cluster Shade = \sum_{i=1}^{Nbins} \sum_{j=1}^{Nbins} P(i, j) [i + j + \mu_i + \mu_j]^3 \quad (A.9)$$

avec $P(i, j)$ la probabilité de l'élément (i, j) de la matrice de co-occurrence normalisée, μ_i et μ_j les moyennes sur la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne de la matrice, σ_i et σ_j les variances sur la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne de la matrice et $Nbins$ le nombre de bins utilisés lors de la discrétisation de l'image.

B

TABLEAUX DE L'ÉTUDE SUR LA MÉTHODE DE DISCRÉTISATION

Relations entre les indices de texture de perfusion et le BF_{\max} , BF_{mean} et le volume perfusé									
Nombre de bins	64			32			16		
	BF_{\max}	BF_{mean}	VP	BF_{\max}	BF_{mean}	VP	BF_{\max}	BF_{mean}	VP
13 GLCM-RR-based									
BF_energy	-0.56	-0.35	-0.99	-0.55	-0.36	-0.97	-0.48	-0.36	-0.88
BF_entropy	0.56	0.34	0.99	0.56	0.35	0.98	0.51	0.36	0.93
BF_IDM	0.42	0.14	0.75	0.45	0.16	0.8	0.46	0.17	0.82
BF_homogeneity	0.46	0.17	0.81	0.46	0.18	0.83	0.47	0.18	0.83
BF_inertia	-0.53	-0.27	-0.87	-0.53	-0.27	-0.87	-0.52	-0.27	-0.86
BF_correlation	0.38	0.15	0.77	0.38	0.15	0.77	0.38	0.15	0.76
BF_dissimilarity	-0.51	-0.24	-0.86	-0.51	-0.24	-0.86	-0.51	-0.24	-0.86
BF_clustershade	0.05	-0.19	0.19	0.05	-0.19	0.19	0.05	-0.19	0.2
BF_clusterprominence	-0.42	-0.36	-0.4	-0.41	-0.35	-0.39	-0.41	-0.35	-0.37
13 GLCM-AR-based									
BF_energy	-0.74	-0.57	-0.92	-0.86	-0.73	-0.8	-0.92	-0.84	-0.63
BF_entropy	0.72	0.53	0.94	0.85	0.7	0.84	0.93	0.83	0.67
BF_IDM	-0.85	-0.89	-0.2	-0.86	-0.89	-0.21	-0.86	-0.88	-0.24
BF_homogeneity	-0.86	-0.89	-0.21	-0.86	-0.89	-0.22	-0.86	-0.88	-0.22
BF_inertia	0.91	0.87	0.33	0.91	0.87	0.33	0.91	0.87	0.32
BF_correlation	0.39	0.16	0.77	0.42	0.19	0.78	0.51	0.28	0.79
BF_dissimilarity	0.89	0.88	0.28	0.9	0.88	0.28	0.89	0.88	0.28
BF_clustershade	0.51	0.25	0.5	0.5	0.24	0.5	0.47	0.23	0.48
BF_clusterprominence	0.95	0.8	0.62	0.95	0.8	0.62	0.95	0.8	0.61
1 GLCM-RR-based									
BF_energy	-0.5	-0.36	-0.91	-0.34	-0.31	-0.67	-0.01	-0.13	-0.16
BF_entropy	0.53	0.36	0.95	0.39	0.32	0.78	0.06	0.15	0.28
BF_IDM	0.41	0.13	0.74	0.44	0.15	0.79	0.46	0.17	0.82
BF_homogeneity	0.46	0.17	0.81	0.46	0.17	0.82	0.47	0.18	0.83
BF_inertia	-0.53	-0.27	-0.86	-0.53	-0.27	-0.86	-0.53	-0.27	-0.86
BF_correlation	0.38	0.14	0.75	0.38	0.14	0.75	0.37	0.13	0.75
BF_dissimilarity	-0.51	-0.24	-0.86	-0.51	-0.24	-0.86	-0.51	-0.24	-0.85
BF_clustershade	0.02	-0.21	0.16	0.02	-0.21	0.17	0.03	-0.21	0.18
BF_clusterprominence	-0.49	-0.4	-0.53	-0.48	-0.39	-0.52	-0.47	-0.38	-0.5
1 GLCM-AR-based									
BF_energy	-0.9	-0.8	-0.7	-0.94	-0.88	-0.55	-0.92	-0.87	-0.46
BF_entropy	0.91	0.79	0.74	0.95	0.87	0.59	0.95	0.88	0.5
BF_IDM	-0.86	-0.9	-0.23	-0.87	-0.9	-0.23	-0.87	-0.88	-0.25
BF_homogeneity	-0.87	-0.9	-0.24	-0.88	-0.9	-0.24	-0.87	-0.88	-0.24
BF_inertia	0.92	0.88	0.34	0.92	0.88	0.35	0.91	0.88	0.34
BF_correlatio	0.39	0.15	0.76	0.42	0.18	0.77	0.52	0.29	0.78
BF_dissimilarity	0.9	0.89	0.3	0.9	0.89	0.3	0.89	0.88	0.3
BF_clustershade	0.51	0.25	0.5	0.5	0.24	0.49	0.48	0.23	0.47
BF_clusterprominence	0.95	0.8	0.6	0.95	0.8	0.6	0.95	0.81	0.6

TABLE B.1 – Coefficients de corrélation de Spearman (ρ) entre le BF_{\max} , BF_{mean} , le VP et les indices de texture de perfusion.

Relations entre les indices de texture de métabolisme et le SUV_{\max}, SUV_{mean} et le VM			
Indices de texture	SUV_{\max}	SUV_{mean}	VM
<i>13 GLCM-RR-based</i>			
M_energy	-0.21	-0.13	-1
M_entropy	0.21	0.13	1
M_IDM	0	-0.09	0.76
M_homogeneity	0	0.08	0.84
M_inertia	-0.06	0.04	-0.91
M_correlation	0.09	0	0.91
M_dissimilarity	-0.04	0.06	-0.90
M_clustershade	-0.22	-0.30	0.11
M_clusterprominence	-0.01	0.02	-0.21
<i>13 GLCM-AR-based</i>			
M_energy	-0.79	-0.74	-0.58
M_entropy	0.76	0.71	0.63
M_IDM	-0.93	-0.93	0.11
M_homogeneity	-0.93	-0.9	0.11
M_inertia	0.96	0.95	-0.05
M_correlation	0.15	0.05	0.92
M_dissimilarity	0.95	0.95	-0.07
M_clustershade	0.62	0.54	0.16
M_clusterprominence	0.98	0.94	0.23
<i>1 GLCM-RR-based</i>			
M_energy	-0.21	-0.17	-0.92
M_entropy	0.22	0.16	0.96
M_IDM	0	-0.1	0.78
M_homogeneity	0	-0.1	0.85
M_inertia	-0.07	0.04	-0.91
M_correlation	0.07	-0.02	0.91
M_dissimilarity	-0.4	-0.07	-0.90
M_clustershade	-0.22	-0.30	0.05
M_clusterprominence	-0.04	0	-0.3
<i>1 GLCM-AR-based</i>			
M_energy	-0.94	-0.92	-0.27
M_entropy	0.95	0.92	0.29
M_IDM	-0.93	-0.93	0.11
M_homogeneity	-0.93	-0.94	0.11
M_inertia	0.96	0.95	-0.05
M_correlation	0.13	0.03	0.91
M_dissimilarity	0.95	0.95	-0.07
M_clustershade	0.64	0.56	0.12
M_clusterprominence	0.98	0.94	0.22

TABLE B.2 – Coefficients de corrélation de Spearman (ρ) entre le SUV_{\max} , le SUV_{mean} et le VM et les indices de texture métaboliques

ANALYSE SUR L'AMÉLIORATION DE LA QUALITÉ DES IMAGES DE FAIBLE STATISTIQUE

L'objectif du travail de stage était d'essayer de caractériser la variabilité de la mesure et le biais quantitatif induit par une faible statistique dans l'image. Pour cela, l'analyse de la restitution du signal et du bruit en fonction de différents temps d'acquisition a été entreprise. Une fois cette étude effectuée, un travail sur l'amélioration de la qualité de ces images par l'optimisation des paramètres de reconstruction et l'utilisation d'une correction de l'effet de volume partiel a ensuite été réalisé.

De manière générale, les tests de quantification et les études de qualité des images peuvent être réalisés à partir de trois types de données : les données issues de fantômes physiques, les données provenant d'images simulées et les données d'images cliniques de patients. L'analyse sur l'évaluation et l'amélioration de la qualité des images décrite dans ce chapitre a été réalisée sur fantôme physique NEMA IEC présenté dans le chapitre 2.

C.1/ MÉTRIQUES

La qualité d'une image en TEP est généralement évaluée en prenant en compte la justesse de la restitution du signal dans l'image reconstruite et le bruit présent dans cette image. De nombreuses métriques permettent d'évaluer ces paramètres et nous avons choisi d'utiliser les deux suivantes : le coefficient de recouvrement ("RC" pour *Recovery Coefficient*) et le coefficient de variation ("CoV" pour *Coefficient of Variation*).

- Le coefficient de recouvrement (RC)

Le coefficient de recouvrement est une métrique qui permet de quantifier la restitution du signal dans l'image. Cette métrique qui est fréquemment utilisée dans les protocoles NEMA, a donc été choisie pour caractériser la qualité de nos images TEP. Par définition, le RC correspond à la concentration mesurée sur l'image dans la région d'intérêt (ici les sphères du fantôme) divisée

par la concentration théorique réelle imagée (l'activité réelle introduite dans les sphères). La valeur de RC est comprise entre 0 et 1, un RC égal à 1 représentant une restitution totale du signal. En pratique, le RC peut être exprimé individuellement pour chaque sphère ou en moyennant sur l'ensemble des six sphères.

$$RC = \frac{[kBq/mL]_{ROI \ image}}{[kBq/mL]_{sphere \ réelle}} \quad (C.1)$$

- Le coefficient de variation (CoV)

Le coefficient de variation (CoV) permet de décrire la variation des valeurs d'intensités dans une région. Cette méthode a été choisie ici car elle permet une mesure simple pour mesurer le bruit présent dans les images. Le CoV est défini comme le rapport entre l'écart type (SD) et la moyenne des valeurs des régions étudiées :

$$CoV(\%) = \frac{SD_{ROI \ image \ Bkg}}{[kBq/mL]_{ROI \ image \ Bkg}} \quad (C.2)$$

Plus la valeur de CoV est élevée, plus les images sont bruitées. En pratique, une moyenne est effectuée sur les 12 ROIs qui ont été placées dans le fond du fantôme, comme le montre la figure C.1, afin de pouvoir calculer pour chaque image les valeurs de CoV.

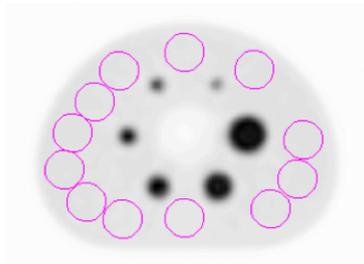


FIGURE C.1 – Image TEP du fantôme NEMA IEC avec contraste sphère-fond de 10 :1. En couleur les 12 régions d'intérêt tracées pour l'évaluation du bruit de fond (bkg). Image reconstruite avec l'algorithme BLOB-OS-TF, 3 itérations, 33 sous-ensembles.

C.2/ SEGMENTATION DES VOLUMES D'INTÉRÊTS

La segmentation des volumes d'intérêts est une étape clé dans l'analyse et la quantification des images TEP. Le choix de la méthode de segmentation peut avoir un impact significatif sur la restitution du signal. Comme il a été décrit dans la partie 2.4 du chapitre 2, de nombreuses méthodes de segmentation existent bien qu'aucune ne fasse consensus à l'heure actuelle. Pour les analyses sur fantôme, nous avons donc choisi d'étudier plusieurs approches pour la délimitation des VOIs :

- Des seuillages fixes, avec comme seuils 41% et 50% de l'intensité maximale dans les zones

fixantes. Ces derniers correspondent aux recommandations EANM pour l'imagerie TEP au ^{18}F -FDG [479].

- Une méthode de délimitation “manuelle”, qui consiste à effectuer une croissance de région avec le volume réel des sphères comme critère d'arrêt.
- Une méthode de seuillage adaptatif implémentée dans la console de recherche Philips Imalytics a également été utilisée. Cette dernière, basée sur une méthode de croissance de région à partir d'une graine placée manuellement, a été appliquée dans un premier temps avec un seuil fixe pour toutes les sphères, puis dans un second temps avec un seuil adapté à chaque sphère. Elle prend en compte également l'information du bruit de fond.
- Deux algorithmes de segmentation basés sur des croissances de région, qui ont été implémentés et exécutés avec les outils clitk du logiciel vv. Ces algorithmes, effectuent une opération morphologique de dilatation à partir de graines disposées aux centres théoriques des sphères, jusqu'à se rapprocher au maximum du volume réel des sphères. La première méthode a comme contrainte géométrique de conserver l'aspect sphérique des volumes tandis que la seconde méthode se doit de minimiser la différence entre les volumes segmentés et les volumes réels.

La segmentation manuelle par croissance de région, le seuillage adaptatif et les algorithmes de dilatation, de par les seuils et critères d'arrêt choisis, sont des approches qui vont être spécifiques à l'analyse sur fantôme sans pouvoir être directement transposables sur des images cliniques. Nous avons néanmoins souhaité les utiliser pour cette analyse sur fantôme afin d'évaluer l'impact de la méthode de segmentation sur la restitution du signal dans des images de faible statistique. Pour cela, l'influence de la méthode de segmentation sur les valeurs de RC a été étudiée pour une reconstruction faite avec les paramètres standards (3 itérations, 33 sous-ensembles, un noyau TOF de 14,1cm et un paramètre de relaxation λ de 1) pour une acquisition de 5s avec un contraste sphères/fond de 10 :1. La valeur de contraste a été choisie en fonction des valeurs des rapports signal sur bruit (SNR pour *Signal to Noise Ratio*) mesurés sur les images de perfusion de la population TREN (la valeur moyenne de SNR étant proche de 8). Le choix de la durée d'acquisition a quant à lui été fait en fonction des données du protocole TREN. En effet, en plus de la série dynamique de 10s, une série dynamique de 5s est également reconstruite et sauvegardée. Afin de se placer dans le cas le plus défavorable, celui où la statistique de l'image est la plus faible, nous avons donc décidé d'évaluer la restitution du signal sur des acquisitions fantôme de 5s.

Les résultats C.2 montrent que l'impact du choix de la méthode de segmentation est d'autant plus prononcé que les tailles de sphères diminuent. Parmi les méthodes donnant les meilleurs RC sur les petites sphères, on peut retenir la délimitation manuelle par croissance de région, le seuillage adaptatif implémenté dans Imalytics adapté pour chaque sphère ou encore les deux seuillages effectués par dilatation. La figure C.2, qui représente (à droite) la différence entre les volumes segmentés et les volumes théoriques, montre quant à elle que la méthode manuelle et la méthode par seuillage adaptatif adapté pour chaque sphère sont celles pour lesquelles les volumes segmentés sont les plus proches des volumes réels. La segmentation manuelle par

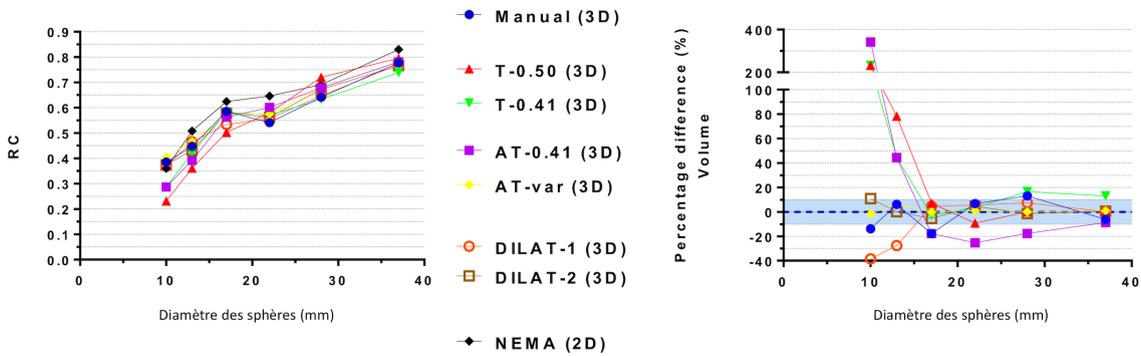


FIGURE C.2 – Coefficients de recouvrement et coefficient de variation en fonction du diamètre des sphères (reconstruction avec 3 itérations, 33 sous-ensemble et un noyau de TOF de 14,1 cm). T-0.50 : seuillage fixe à 50% de l'intensité maximale, T-0.41 : seuillage fixe à 41% de l'intensité maximale, Initiale : croissance de région avec le volume réel des sphères comme critère d'arrêt, AT-0.41 : seuillage adaptatif (Imalytics) avec un seuil fixe de 41% de l'intensité maximale pour toutes les sphères, AT-var : seuillage adaptatif (Imalytics) avec un seuil variable pour chaque sphères, DILAT-1 : dilatation avec contrainte sur l'aspect sphérique, DILAT-2 : dilatation avec contrainte sur la différence du volume segmenté, NEMA : segmentation utilisée dans les recommandations des tests EANM [478]

croissance de région qui utilise le volume réel des sphères comme critère d'arrêt a donc été utilisée pour la suite des analyses sur fantôme.

C.3/ CARACTÉRISATION DU BIAIS INDUIT PAR LA STATISTIQUE DES IMAGES TEP

Le biais quantitatif sur les images TEP de faible statistique a été mis en évidence par de nombreuses études [480–482] et a notamment été quantifié pour des reconstructions itératives comme OSEM [273,274]. Ce biais doit pouvoir être évalué et corrigé si nécessaire pour pouvoir maintenir une cohérence de quantification dans les études d'images dynamiques. Le biais étant dépendant de la statistique de l'image, l'objectif a donc été de caractériser la variabilité de la mesure sur des images reconstruites avec différents temps d'acquisitions. Ces derniers ont été choisis de manière à être à la fois représentatifs des temps d'acquisition cliniques mais également représentatifs des temps d'échantillonnage des acquisitions dynamiques utilisées pour l'évaluation de la perfusion tumorale.

- Acquisitions et reconstructions des images

Cinq acquisitions correspondant à cinq rapports de contraste sphères/fond ont été effectuées avec une concentration en activité de 3 kBq/mL dans le fond du fantôme au moment de l'acquisition. Les concentrations d'activité dans les sphères ont été déterminées de manière à obtenir des rapports de contrastes de 2:1, 5:1, 10:1, 20:1 et 30:1 afin de couvrir la plage de

contraste observée sur les images cliniques entre les tumeurs et le bruit de fond. Différents temps d'acquisitions correspondant à différentes statistiques dans l'image ont été réalisés avec au total sept temps d'acquisitions différents : 5, 10, 15, 30, 60, 120 et 480s. Dix réplicats ont été réalisés pour chacun des temps. Une acquisition de 3h30 a été effectuée, ce qui a permis d'avoir une acquisition de référence non bruitée (fig. C.3). Toutes les données brutes ont été sauvegardées en mode liste.

L'étude sur l'évaluation et l'amélioration de la qualité des images a été effectuée dans un

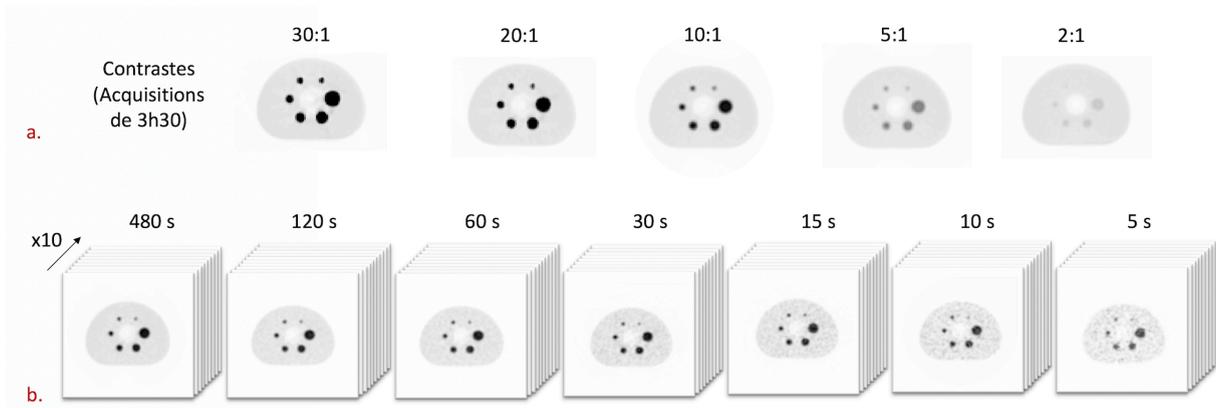


FIGURE C.3 – (a) Illustration des différents contrastes et des réplicats (b) aux différents temps d'acquisitions.

premier temps pour un contraste de 10 :1 avec la volonté d'étudier les autres contrastes par la suite. Les images correspondantes ont été reconstruites avec les mêmes paramètres que les images cliniques, c'est à dire en utilisant l'algorithme BLOB-OS-TF avec 3 itérations, 33 sous-ensembles, un paramètre de relaxation λ de 1 et un noyau TOF de 14.1 cm. Aucun traitement post-reconstruction n'a été effectué. Les images ont été reconstruites en matrices de 144 x 144 avec des voxels isotropes de 4 mm.

- Présentation des résultats

La variabilité des mesures et le biais quantitatif en fonction de la statistique des images ont été étudiés pour 10 acquisitions statistiquement équivalentes, répétées pour chaque temps de reconstruction. La segmentation des sphères a été faite sur les images de l'acquisition de 3h30 en utilisant une méthode de croissance de région, dont la graine initiale a été placée manuellement au centre des sphères et dont le critère d'arrêt est le volume réel de ces dernières.

Afin de pouvoir observer le comportement du bruit en fonction de la statistique des images, nous avons tracé les valeurs de CoV mesurées sur les images reconstruites pour chaque temps d'acquisition. La figure C.4 compare ainsi les valeurs de CoV sur nos images avec le comportement classique d'un bruit Poissonien. Il est intéressant de noter que les données mesurées sur l'image s'écartent du modèle Poissonien, soulignant l'intérêt de bien caractériser le bruit propre au modèle d'acquisition et de reconstruction retenus.

Sur les graphiques a et b de la figure C.6 sont représentés les coefficients de recouvrement

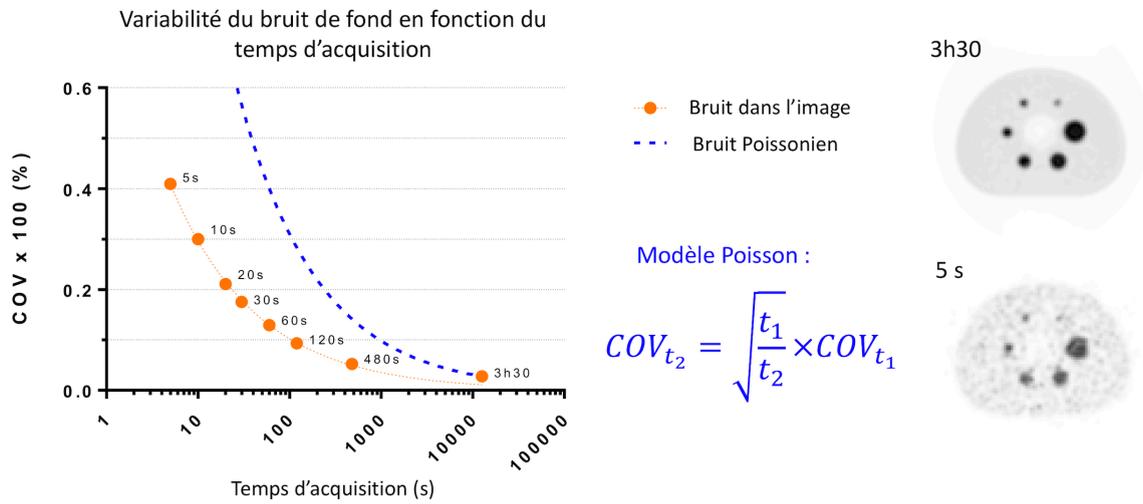


FIGURE C.4 – Comparaison des coefficients de variation en fonction du temps d’acquisition entre les mesures réalisées sur les images TEP et le modèle idéal de Poisson. Illustration à droite de l’image de la différence entre un bruit poissonien présent sur l’image TEP de 3h30 et le bruit d’une image avec un temps d’acquisition de 5s.

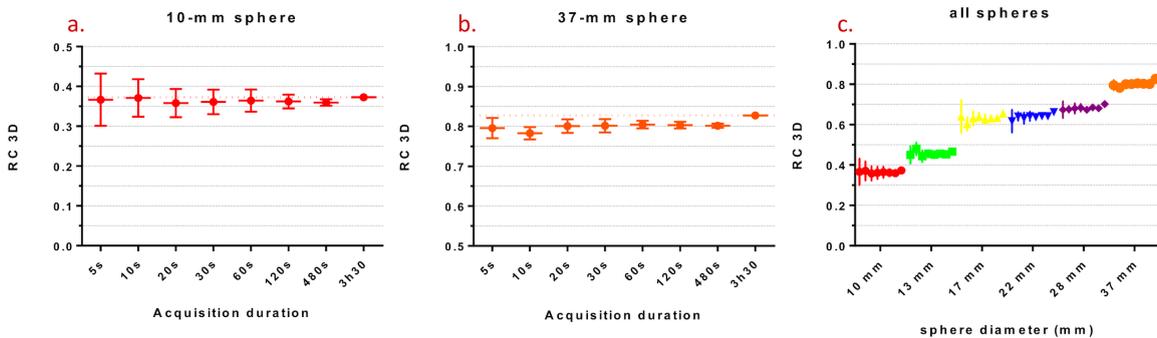


FIGURE C.5 – Coefficients de recouvrement moyennés sur les dix répliques en fonction du temps d’acquisition pour les sphères de 10 (a) et de 37 (b) mm de diamètre et en fonction de la taille des sphères (c) pour tous les temps d’acquisition (reconstructions avec 3 itérations et 33 sous-ensembles).

(± écart type) pour la plus petite et la plus grande sphère respectivement, en fonction du temps d’acquisition. Pour les deux sphères, plus le temps d’acquisition augmente, plus la variation du RC entre les 10 répliques est faible, soulignant moins de variabilité dans les mesures à forte statistique. Par ailleurs, on peut également noter que plus le temps d’acquisition est long, meilleur est le RC quelle que soit la taille de sphère (fig. C.6a, b et c). De même, comme l’illustre le graphique c de cette même figure, plus la sphère est grande, meilleur est le RC. Les faibles valeurs de RC pour les petites sphères sont en partie dû à l’effet de volume partiel dont l’impact est particulièrement marqué pour les petites structures, en particulier celles dont la dimension est inférieure à trois fois la résolution spatiale du système [483] (part. 1.2.2.2).

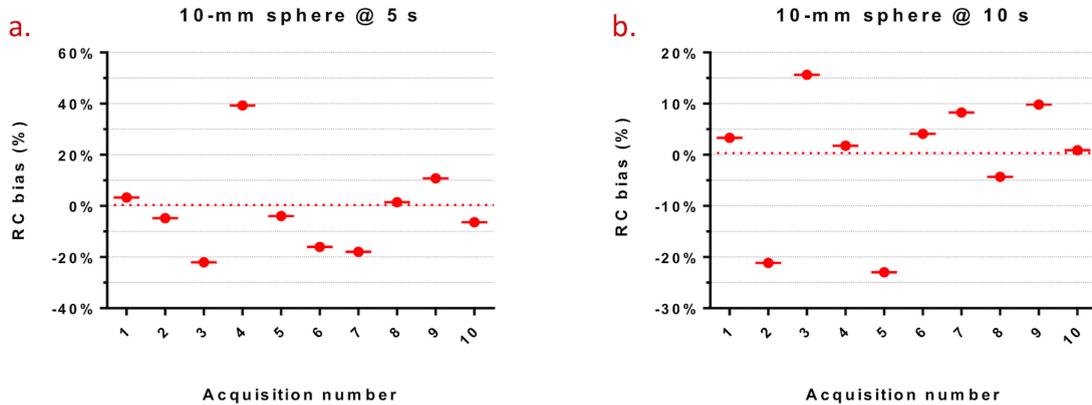


FIGURE C.6 – Biais du coefficient de recouvrement pour les dix réplicats pour la plus petite sphère ($\varnothing = 10\text{mm}$) pour des temps d'acquisition de 5s (a) et 10s (b) (reconstructions avec 3 itérations et 33 sous-ensembles).

La figure C.6 illustre quant à elle, pour des acquisitions de 10 (a) et 5 (b) secondes, le biais du coefficient de recouvrement sur les 10 réplicats de la plus petite des sphères. Ce dernier a été mesuré par rapport aux valeurs de RC de l'acquisition de 3h30. Nous observons ainsi que ce biais varie d'un réplica à l'autre avec une plage de valeurs de 2% à 39% pour les acquisitions de 5s et de 1% à 23% pour les acquisitions de 10s. Compte tenu de la variabilité des mesures, le biais maximal observé peut être deux fois plus important pour les acquisitions de 5s que pour celles de 10s.

C.4/ AMÉLIORATION DE LA QUALITÉ DES IMAGES DE FAIBLE STATISTIQUE

Comme énoncé au chapitre 2 (part. 2.3.5), il existe plusieurs approches pour améliorer la qualité des images de faible statistique. Dans le cas de notre étude, l'idée est de trouver une manière d'améliorer la qualité des images de faible statistique, afin de pouvoir par la suite réaliser les corrections correspondantes sur les images cliniques de perfusion de notre base de donnée. L'amélioration doit donc pouvoir être effectuée sans modifier le protocole d'acquisition. Par ailleurs, seuls les algorithmes de reconstruction implémentés dans les consoles Philips peuvent être utilisés et aucun d'entre eux n'est malheureusement dédié aux reconstructions dynamiques. Enfin, le choix du modèle cinétique est lié à la fois au choix du traceur utilisé et au tissu étudié. Le modèle mono-compartmental de Mullani a été choisi en conséquence, sa modification n'a donc, dans un premier temps, pas été envisagée. Afin de pouvoir améliorer la qualité des images, nous nous sommes donc concentrés sur la modification des paramètres de reconstruction de l'algorithme utilisé en clinique, sur l'influence des méthodes de segmentation des VOIs et sur la correction de l'effet de volume partiel. **Tous les calculs ont été effectués sur les reconstructions de 5s.**

C.4.1/ PARAMÈTRES DE RECONSTRUCTION

L'algorithme BLOB-OS-TF permet à l'utilisateur de modifier plusieurs paramètres impactant la reconstruction de l'image :

- le nombre d'itérations et de sous-ensembles, qui vont définir la convergence de l'algorithme
 - le paramètre de relaxation λ , qui va impacter sur le filtrage de l'image reconstruite
 - la largeur du noyau du temps de vol, qui va avoir un impact sur le contraste et l'uniformité de l'image
- Influence des nombres d'itérations et de sous-ensembles

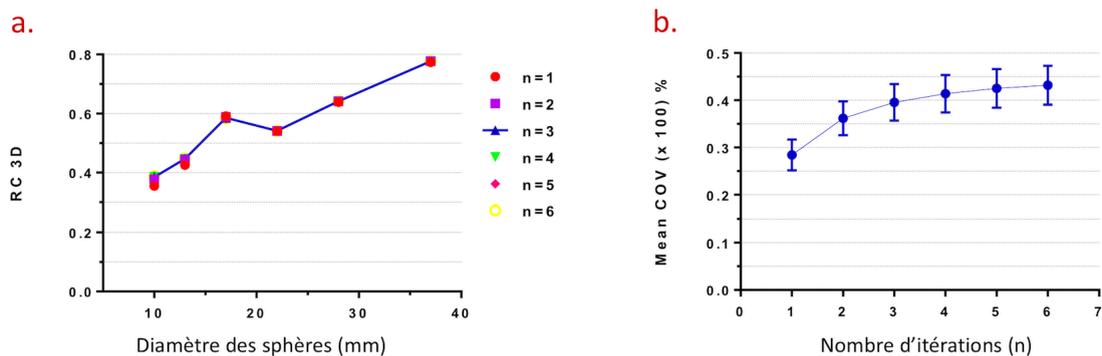


FIGURE C.7 – Influence du nombre d'itérations sur le coefficient de recouvrement en fonction de la taille des sphères (a) et sur le coefficient de variation (b) (reconstruction avec 33 sous-ensembles, un noyau de TOF de 14,1 cm et un paramètre de relaxation λ de 1).

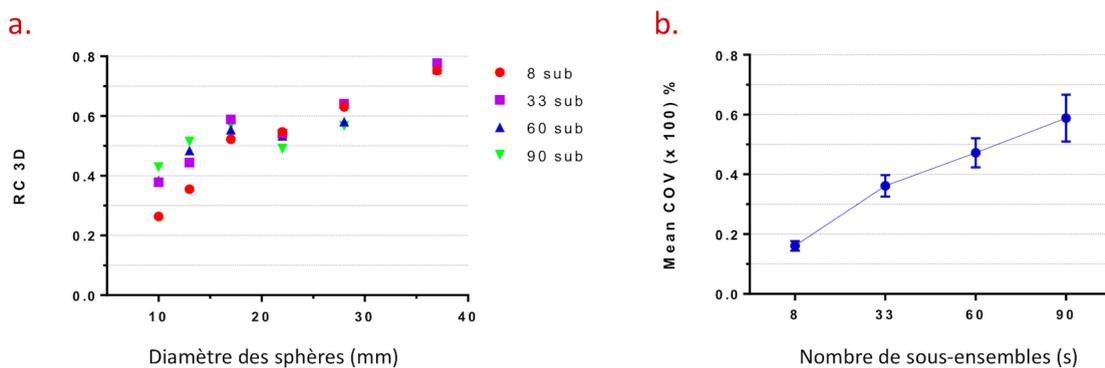


FIGURE C.8 – Influence du nombre de sous-ensembles sur le coefficient de recouvrement en fonction du diamètre des sphères (a) et sur le coefficient de variation (b) (reconstruction avec 2 itérations, un noyau de TOF de 14,1 cm et un paramètre de relaxation λ de 1).

Le comportement de l'algorithme BLOB-OS-TF est celui d'OSEM (part. 1.2.1.3.3) : les basses fréquences spatiales vont converger en premier, suivies des fréquences plus élevées qui vont s'ajouter au fur et à mesure des nouvelles étapes du processus itératif. Ainsi, plus le nombre d'itérations est élevé, plus l'algorithme va ajouter du signal et se rapprocher de la solution mais

plus la proportion de hautes fréquences, qui composent le bruit dans l'image résultante, va augmenter. Il en est de même avec les sous-ensembles, qui, en permettant l'accélération de la convergence, ajoutent également plus rapidement le bruit dans l'image. Les valeurs du nombre d'itérations et de sous-ensembles ne sont pas modifiables directement dans les protocoles de reconstruction, ils nécessitent la modification des fichiers systèmes. Le choix de leurs valeurs étant libre, nous avons donc décidé de tester quatre valeurs de sous-ensembles : 8, 33, 60, 90 et fait varier le nombre d'itérations de 1 à 7 afin d'analyser l'impact de ces paramètres sur la qualité des images reconstruites. Ces valeurs ont été choisies en se basant sur la configuration standard de la GEMINI TF (3 itérations, 33 sous-ensembles) ainsi que sur des études évaluant l'amélioration de la qualité des images TEP à l'yttrium 90 [484, 485] (ces dernières présentant une plus faible statistique que les images TEP au ^{18}F -FDG).

De manière générale, les résultats montrent que le nombre d'itérations impacte peu les valeurs de RC comme illustré sur la figure C.7a. En revanche, le bruit dans l'image augmente lorsque le nombre d'itérations augmente (fig. C.7b).

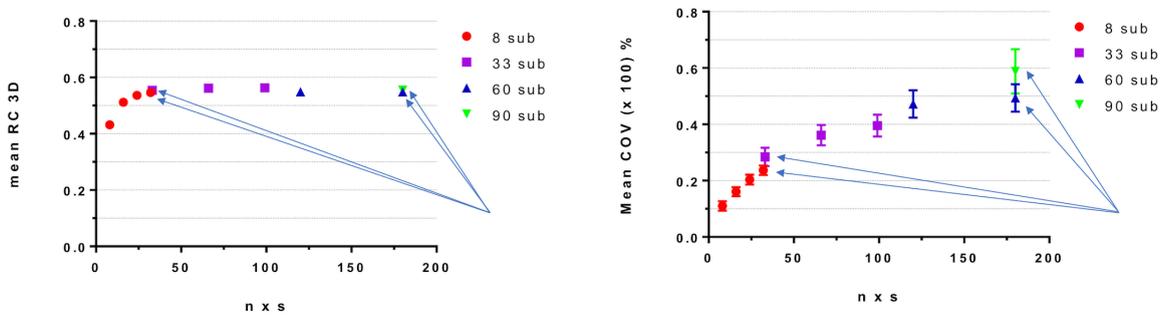


FIGURE C.9 – Influence du nombre d'itérations et de sous-ensembles ($n \times s$) sur le coefficient de recouvrement (a) et sur le coefficient de variation (b), pour des images reconstruites avec un noyau de TOF de 14,1 cm et un paramètre de relaxation λ de 1. Les modèles indiqués par les flèches correspondent au modèle de convergence $n \times s = 180$ avec 2 itérations et 90 sous-ensembles et au modèle de convergence $n \times s = 1 \times 33$ comprenant 1 itération et 33 sous-ensembles.

Il en va de même pour l'impact du nombre de sous-ensembles sur le bruit dans l'image reconstruite : les valeurs de CoV sont d'autant plus élevées que le nombre de sous-ensembles augmente (fig. C.8b). Pour ce qui est de la restitution du signal, l'augmentation du nombre de sous-ensembles affecte principalement les petites sphères où un RC plus élevé est noté pour les sous-ensembles les plus grands, comme illustré dans la figure C.8a.

Afin d'observer conjointement l'influence du nombre d'itérations et le nombre de sous-ensembles, les valeurs de RC et de CoV ont été évaluées pour différentes valeurs de convergence de l'algorithme. Cette dernière est définie comme le produit du nombre d'itérations multiplié par le nombre de sous-ensembles. On remarque ainsi sur la figure C.9 que pour deux situations au même taux de convergence ($n \times s$) (indiquées par les flèches bleues sur les graphiques), les modèles présentent la même valeur de RC mais des valeurs de CoV (et donc de bruit) très

différentes. Pour une même valeur du taux de convergence, la partition iterations/subsets a donc une importance : 5 itérations et 8 sous-ensembles ne sont pas équivalents en termes de bruit à 8 itérations et 5 sous-ensembles.

- Influence du paramètre de relaxation

Comme énoncé au chapitre précédent, le paramètre de relaxation λ permet de pondérer la contribution du bruit à chaque itération. Plus λ est faible, plus l'ajout de bruit dans l'image est limité. Plusieurs valeurs pour ce paramètre sont accessibles en modifiant le protocole de reconstruction. Par défaut, il lui est attribué une valeur de 1 correspondant au lissage minimal. Le paramètre de relaxation peut également être fixé à 0,7 ; 0,6 et 0,5 ; avec un lissage de plus en plus prononcé à mesure que sa valeur diminue, comme illustré sur la figure C.10b où les valeurs de CoV sont d'autant plus grandes que λ augmente. à contrario, les valeurs du paramètre de relaxation n'ont *a priori* pas ou peu d'influence sur la restitution du signal, excepté éventuellement pour les très petites structures, où le lissage peut engendrer une perte de signal plus importante, comme illustré sur la figure C.10a.

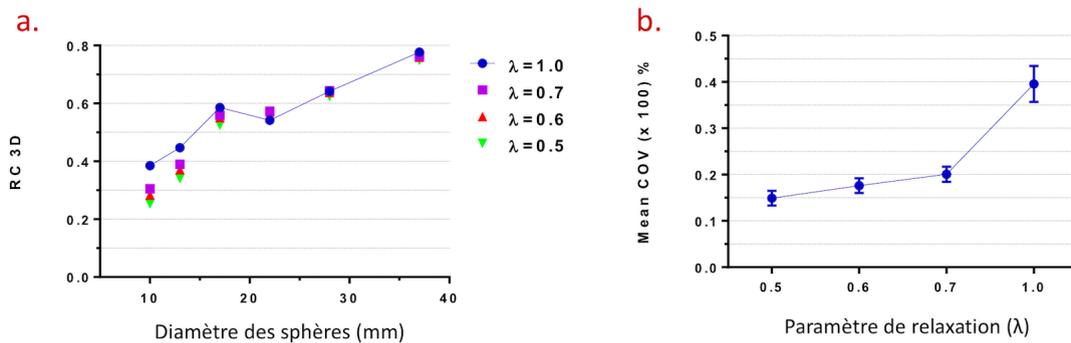


FIGURE C.10 – (a) Coefficients de recouvrement en fonction du diamètre des sphères pour différentes valeurs du paramètre de relaxation λ , (b) coefficients de variation en fonction du paramètre de relaxation λ (reconstruction avec 3 itérations, 33 sous-ensemble et un noyau de TOF de 14,1 cm).

- Influence du noyau de reconstruction TOF

Deux largeurs de noyaux (FWHM) sont accessibles directement via le protocole de reconstruction : une première par défaut de 14,1 cm et une seconde de 18,7 cm qui correspond au mode “qualité supérieure” (HQ pour *High Quality*). Avec un noyau plus large, le RC va sensiblement augmenter (fig. C.11a) mais le bruit sera plus important. La figure C.11b illustre ce phénomène où l'on peut toutefois noter une diminution significative de l'écart type des valeurs de CoV pour le mode HQ, traduisant une meilleure homogénéité dans l'image. Compte tenu de ces observations, plusieurs combinaisons de paramètres de reconstruction ont été testées, faisant ainsi varier à la fois le nombre d'itérations, le nombre de sous-ensembles, les valeurs du paramètre de relaxation et la taille du noyau de reconstruction TOF. Les résultats de RC et de CoV sont représentés sur la figure C.12 où chaque point représente un jeu différent de paramètres de

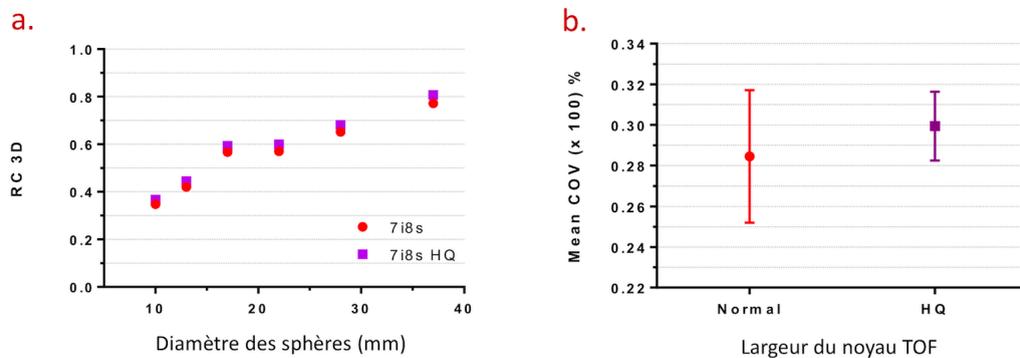


FIGURE C.11 – (a) Coefficients de recouvrement en fonction du diamètre des sphères pour différentes valeurs de noyau de TOF, (b) coefficients de variation en fonction de la largeur du noyau de TOF (reconstruction avec 7 itérations, 8 sous-ensembles et un paramètre de relaxation de 1).

reconstruction. Cette représentation des données à l’avantage de permettre de sélectionner la combinaison de paramètres optimaux donnant le meilleur rapport entre la restitution du signal et le bruit présent dans l’image. L’analyse des résultats nous a permis d’identifier les paramètres optimaux suivants : 7 itérations, 8 sous-ensembles, un lissage “B” correspondant à un λ de 0,5 et un noyau TOF de 18,7 cm correspondant au mode “HQ”. Par rapport aux paramètres de reconstruction standard appliqués initialement (3 itérations, 33 sous-ensembles, un λ de 1 et noyau TOF de 14,3 cm), le taux de convergence a donc diminué, un lissage plus important a été appliqué à l’image et l’option HQ a été sélectionnée. Comme illustrée sur la figure C.12, cette combinaison de paramètres (point rouge sur la figure C.12) permet d’avoir un RC équivalent à celui d’une reconstruction standard (point bleu sur la figure C.12) pour une valeur de CoV près de deux fois inférieur.

C.4.2/ CORRECTION DE L’EFFET DE VOLUME PARTIEL

L’effet de volume partiel est la combinaison à la fois du flou lié à la résolution spatiale et de l’échantillonnage plus ou moins grossier des images TEP (part. 1.2.2.3). La correction de cet effet fait partie des méthodes permettant d’améliorer la qualité des images TEP reconstruites. Or, la majorité des lésions tumorales mammaires de la base de données TREN sont sujettes à l’effet de volume partiel à cause de leur petite taille. Nous avons donc souhaité évaluer l’impact d’une correction de l’EVP sur les images de faible statistique afin d’observer et quantifier les possibles améliorations de la qualité de ces images.

Plusieurs approches ont été proposées afin de corriger l’EVP sur les images de TEP. Ces méthodes sont généralement soit accessibles sur des consoles de recherches cliniques, soit via des logiciels (libres ou payants) ou bibliothèques informatiques. Afin de corriger l’EVP sur nos images, nous avons décidé d’utiliser l’outil PET PVC [486] disponible en accès libre. L’idée était d’évaluer une méthode de correction qui puisse être appliquée par la suite à la base de données. Il était donc nécessaire d’avoir un outil qui fournisse en sortie une image corrigée sur

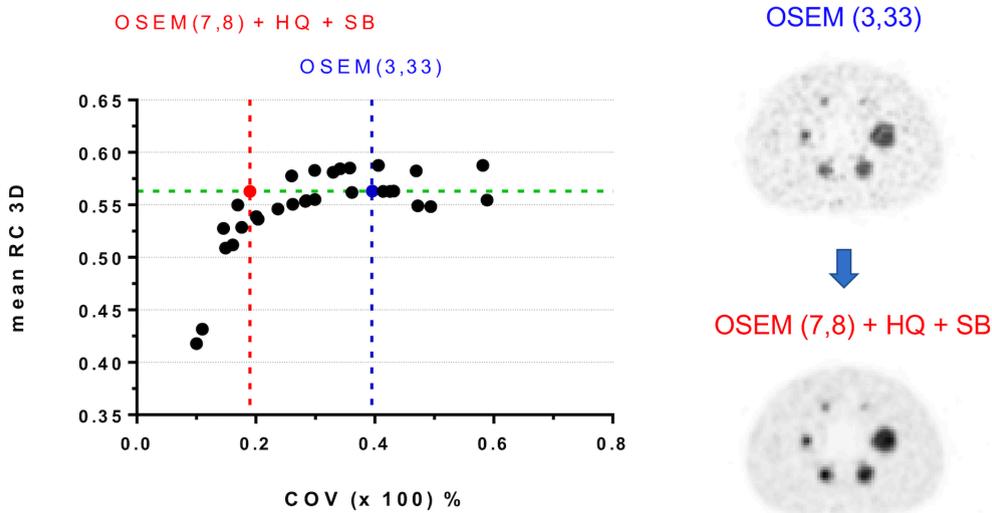


FIGURE C.12 – Graphique récapitulatif des coefficients de recouvrement en fonction des coefficients de variation pour tous les ensembles de paramètres de reconstruction testés. Illustration à droite de l’image reconstruite avec les paramètres standards et les paramètres optimisés (HQ : noyau de TOF de 18,7 cm, SB : paramètre de relaxation de 0,5).

laquelle divers paramètres (valeurs de perfusion, analyse de texture) pourraient ensuite être calculés. L’outil PET PVC permet d’obtenir cette image corrigée, contrairement, par exemple, aux outils de correction de la console de recherche Imalytics. Pour pouvoir appliquer les corrections proposées par l’outil PET PVC sur les données, deux paramètres sont néanmoins nécessaires : un masque du fantôme comprenant toutes les régions d’intérêts des sphères doit être généré et la valeur de la résolution spatiale du système doit être renseignée.

Comme il est conseillé d’avoir, pour le masque, une résolution la plus fine possible, ce dernier a été généré à partir de la segmentation de l’acquisition TDM du fantôme (voxels de dimension 1,17 x 1,17 x 2,5 mm³). Toutes les mesures de RC ont été faites en appliquant ce masque sur les images TEP.

C.4.2.1/ DÉTERMINATION DE LA RÉOLUTION SPATIALE DU SYSTÈME

En ce qui concerne la détermination de la résolution spatiale, les méthodes classiques réalisées à l’aide de sources ponctuelles ou linéaires donnent des résultats très satisfaisants dans des conditions d’échantillonnages élevés, comme c’est le cas pour l’évaluation des performances des systèmes d’imagerie [203]. Toutefois, ces conditions d’échantillonnage ne sont cependant pas représentatives des protocoles de reconstructions cliniques, où un échantillonnage plus “grossier” est généralement effectué. Prieto *et al.* [487] ont présenté une méthode alternative aux approches conventionnelles qui permet de mesurer la résolution spatiale du système dans des conditions cliniques. Cette méthode se base sur des valeurs de RC mesurés et simulés, qui permettent d’estimer la résolution spatiale en évaluant son impact sur la sous-estimation de l’activité pour des

objets sphériques. Selon leur étude, lorsque la résolution doit être mesurée dans des conditions cliniques, l'approche en utilisant les RC produit des valeurs de résolution plus représentatives que les méthodes classiques. Comme il s'agit ici de corriger l'EVP dans des conditions de reconstructions cliniques, nous avons donc décidé d'évaluer la résolution spatiale de notre système via la méthode de Prieto *et al.* directement sur des images de fantôme, puis de la comparer à une estimation faite avec une source linéaire (fig. C.13).

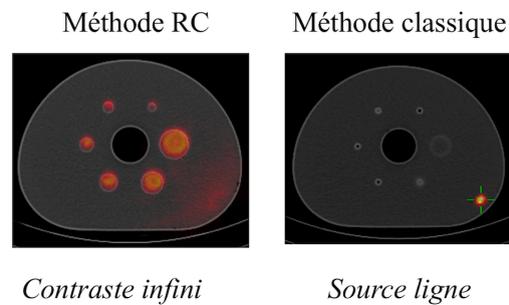


FIGURE C.13 – Illustration du fantôme NEMA IEC avec des contrastes infinis

- Méthode basée sur les profils d'activité d'une source ligne

Une source ligne, de 1 mm de diamètre externe a été placée au sein du fantôme NEMA IEC, parallèle à l'axe du tomographe. Le fond du fantôme a été rempli d'eau sans ajout d'activité (fond froid), permettant un contraste infini. Une acquisition de 30 minutes a été effectuée et les images ont été reconstruites en utilisant l'algorithme BLOB-OS-TF, avec les paramètres standards (3 itérations, 33 sous-ensembles). La résolution spatiale a ensuite été estimée dans le plan transverse, perpendiculairement à la position de la source de ligne, en mesurant la FWHM des profils d'activité tracés sur l'image reconstruite selon l'axe radial et l'axe tangentiel. Pour trois positions selon la direction axiale, les valeurs de FWHM ont été mesurées après interpolation linéaire des données des voxels, comme illustré sur la figure C.14. Les valeurs de FWHM ont ensuite été moyennées afin d'obtenir la résolution moyenne de l'image sur le plan transverse. En utilisant cette méthode, la résolution spatiale a été estimée à $9,1 \pm 0,5$ cm avec les paramètres cliniques de reconstruction.

- Méthode basée sur les valeurs de RC

Cette méthode consiste à comparer les valeurs de RC mesurées expérimentalement sur les images TEP d'un fantôme physique, à des valeurs de RC théoriques obtenues en dégradant progressivement une image simulée idéale du fantôme physique par une fonction Gaussienne. La valeur de résolution spatiale retenue est celle pour laquelle le noyau Gaussien donne le meilleur ajustement entre les valeurs théoriques et expérimentales.

Pour la mesure expérimentale des RC, les six sphères du fantôme ont été remplies avec une solution homogène de ^{18}F -FDG de 7 kBq/mL. Aucune activité n'a été placée dans le fond du fantôme de manière à avoir, comme pour la méthode classique, un contraste infini. Une acquisition de 30 minutes a ensuite été effectuée et les images ont été reconstruites selon le protocole de reconstruction standard précédemment décrit. Le RC a ensuite été calculé sur les images

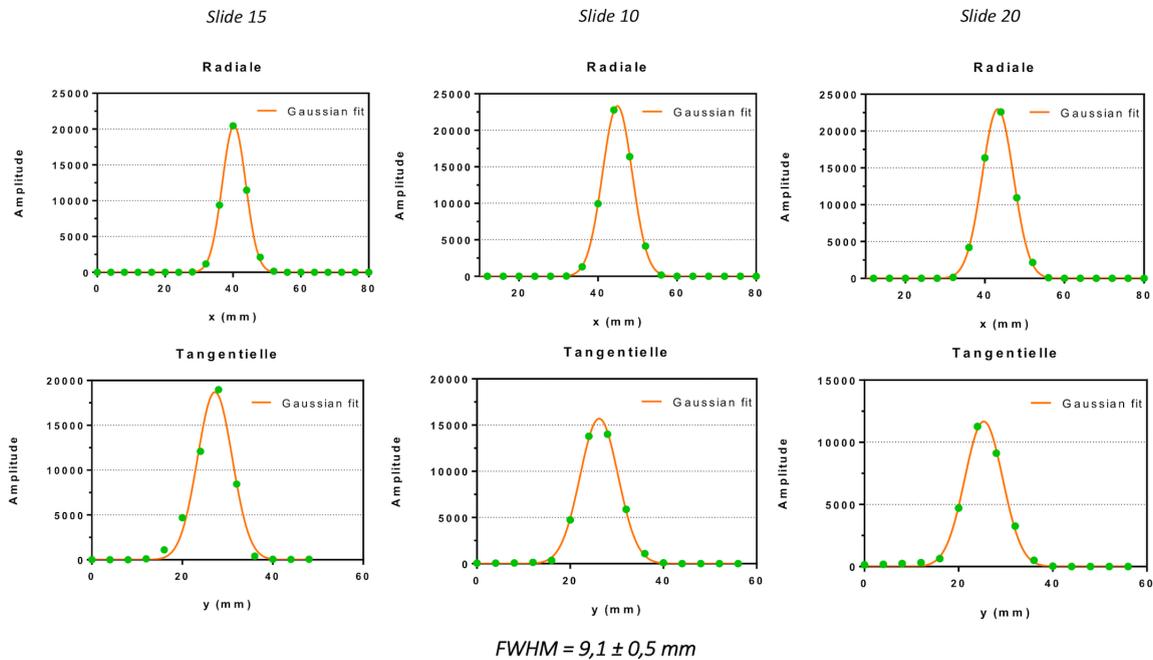


FIGURE C.14 – Profils de la source ligne, selon la direction radiale et tangentielle pour différentes coupes selon la direction axiale

reconstruites pour chacune des sphères. En ce qui concerne le calcul des RC théoriques, un modèle numérique a été généré à partir de la segmentation des images TDM du fantôme NEMA IEC. Une image “parfaite” a ensuite été créée dans les mêmes conditions que l’acquisition expérimentale, c’est à dire avec une concentration d’activité de 7 kBq/mL dans les sphères et aucune activité dans le fond. Cette image a ensuite été progressivement dégradée avec l’application de différents filtres Gaussiens de valeurs de FWHM allant de 5 à 15 mm avec un pas de 0,1. Pour chaque filtre, les valeurs de RC ont été calculés pour toutes les sphères. Les valeurs de chaque RC en fonction du diamètre des sphères ont été tracées. Les différences entre les courbes théoriques et la courbe expérimentale ont ensuite été évaluées par la méthode du χ^2 de Pearson, dont les écarts ont été tracés sur la figure C.15. La résolution du système a alors été estimée par la valeur de FWHM du modèle théorique minimisant la différence avec les valeurs expérimentales, ce qui correspond pour nos images à une FWHM de 8,6 mm. La valeur de la résolution spatiale estimée via les RC est donc comprise dans l’écart type de la valeur déterminée par la méthode classique, ce qui valide cette approche dans les conditions de l’étude. La valeur de 8,6 mm a donc été choisie comme résolution spatiale du système pour l’application de l’outil PET PVC.

C.4.2.2/ ANALYSE DES RÉSULTATS DES CORRECTIONS DE L’EVP

L’outil PET PVC propose 8 méthodes classiques de corrections de l’EVP. Parmi ces dernières, deux grandes catégories sont identifiées : les méthodes basées sur les masques anatomiques et les méthodes basées sur la déconvolution [208]. Par ailleurs, 14 techniques hybrides, basées sur la combinaison de plusieurs méthodes classiques, sont également disponibles. Les

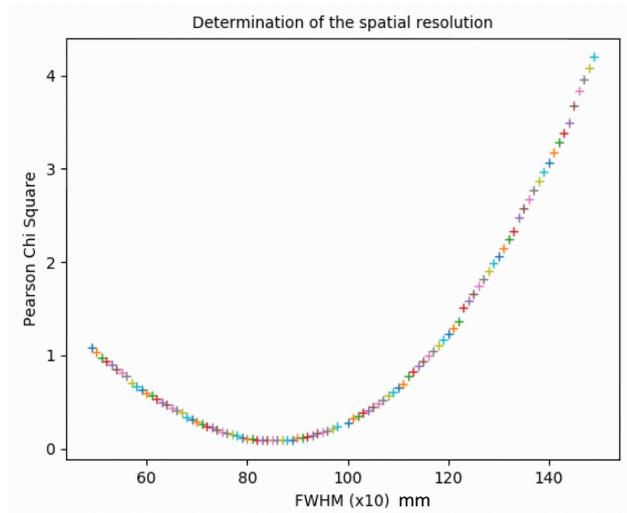


FIGURE C.15 – écarts des χ^2 en fonction des FWHM

descriptions détaillées de toutes ces approches peuvent être trouvées dans la publication de Thomas *et al.* [486] présentant l’outil PET PVC. Les méthodes permettant d’obtenir une image corrigée en sortie ont été appliquées aux images TEP reconstruites avec les paramètres standards et optimisés. Les images corrigées correspondantes ont été sauvegardées et pour chacune d’entre elles, les RC ont été mesurés. Excepté la méthode RVB VC qui présente globalement

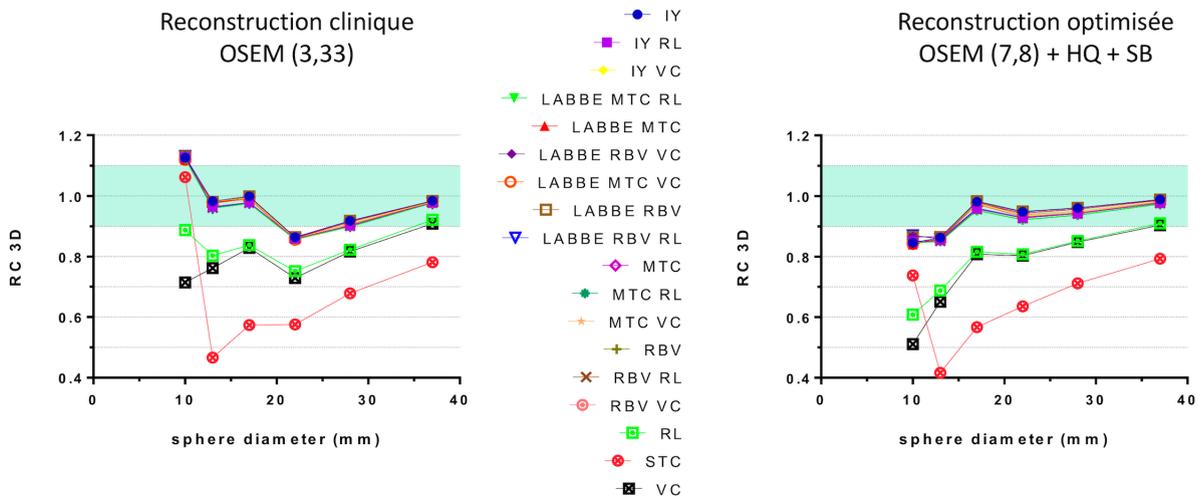


FIGURE C.16 – Valeurs des RC en fonction du diamètre des sphères après différentes méthodes après différentes méthodes de correction de l’EVP

des résultats similaires aux images non corrigées de l’EVP, toutes les autres corrections permettent d’améliorer la restitution du signal dans l’image. Cette amélioration est d’autant plus marquée pour les petites sphères, pour lesquelles les valeurs de RC présentent une augmentation allant jusqu’à doubler pour certaines corrections. L’amélioration de la restitution du signal

est toutefois moins élevée pour les images reconstruites avec les paramètres optimisés, mais présente néanmoins des RC très proches de 1 pour les corrections les plus performantes. Par

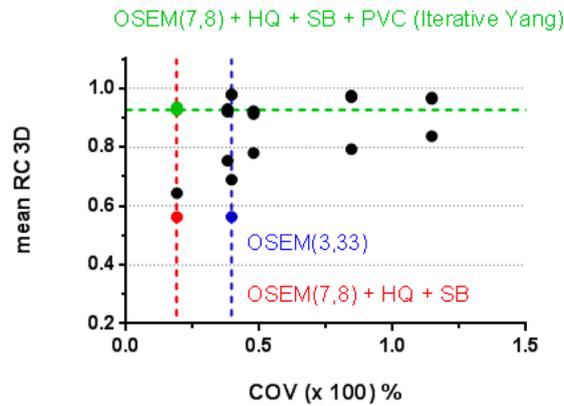


FIGURE C.17 – Graphique récapitulatif des coefficients de recouvrement en fonction des coefficients de variation pour les différentes corrections d’EVP testées.

ailleurs, aucune surestimation de l’activité n’est observée, contrairement aux RC des images reconstruites de manière standard, qui présentent pour la plus petite sphère des valeurs de RC la plupart du temps supérieures à 1. L’étude combinée de la restitution du signal et du bruit dans l’image (fig. C.17) a permis d’identifier la méthode itérative de Yang (IY) comme la correction permettant le meilleur compromis. L’impact sur l’évaluation de la qualité visuelle des images est illustré sur la figure C.18. Sur cette dernière sont représentées l’image reconstruite avec les paramètres initiaux, l’image reconstruite avec les paramètres optimisés et l’image reconstruite avec les paramètres optimisés avec correction de l’effet de volume partiel, accompagnées des profils d’activité de la sphère de plus grand diamètre. On observe ainsi que la reconstruction optimisée sans correction d’EVP présente, pour la plus grande sphère, un profil d’activité plus homogène que celui issu de la reconstruction standard et que la reconstruction optimisée avec correction d’EVP présente un profil également plus homogène et dont les transitions sont plus marquées. Cette analyse sur la qualité des images de faible statistique montre donc des résultats prometteurs quant à la possible amélioration des images de faible statistique comme le sont les images de perfusion.

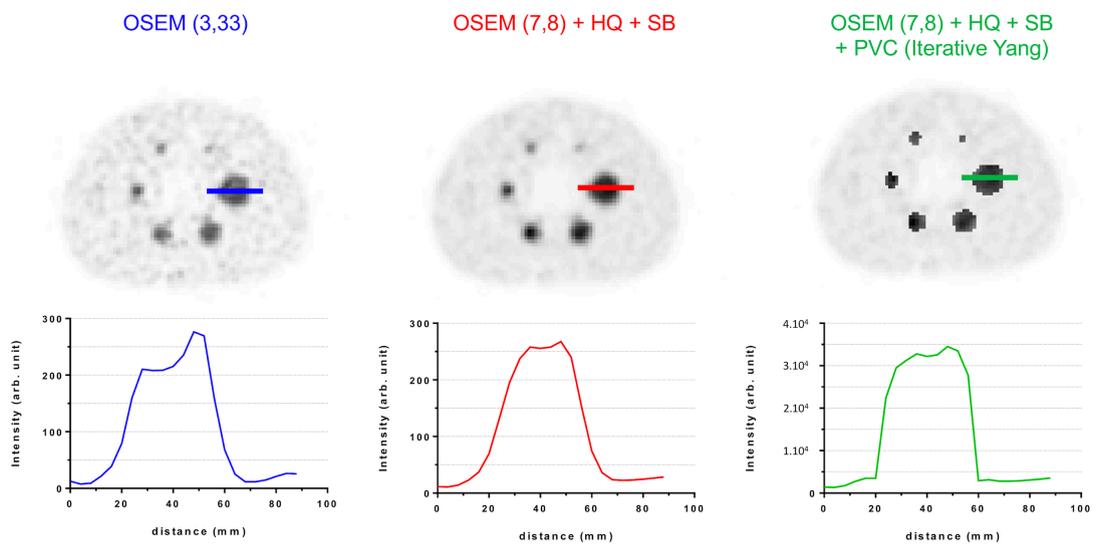


FIGURE C.18 – évaluation de la qualité visuelle et des profils d'intensité pour la reconstruction initiale (bleu), la reconstruction optimisée (rouge) et la reconstruction optimisée corrigée de l'EVP (vert)

D

PROCOLE TREN

N° ID RCB :

Etude de soins courants évaluant l'utilité de la Tomographie par Emission de Positons (TEP) au 18F-Fluorodésoxyglucose (FDG) pour évaluer précocement la REponse au traitement Néoadjuvant des cancers du sein (TREN)

GESTIONNAIRE

Centre Georges-François Leclerc
1, rue professeur Marion
BP 77980
21079 Dijon

Tél : 03 80 73 75 00
Fax : 03 80 73 77 53

INVESTIGATEUR COORDONNATEUR

Pr Alexandre COCHET
MEDECINE NUCLEAIRE
Centre GF Leclerc
1, rue professeur Marion
BP 77980
21079 Dijon

Tél : 03 80 73 75 22
Fax : 03 80 73 77 32
e-mail : acochet@cgfl.fr

COMITE DE REDACTION DE L'ESSAI (CRE)

Pr A. COCHET
Dr O. HUMBERT
Pr F.BRUNOTTE
M. J. SKRZYPSKI

Sommaire

1. INTRODUCTION ET RATIONNEL DE L'ESSAI	4
2. OBJECTIFS	8
2.1 Objectif principal :	8
2.2 Objectifs secondaires	8
3. CRITERES DE JUGEMENTS.....	8
3.1 Critère de jugement principal	8
3.2 Critères de jugements secondaires	8
4. METHODOLOGIE	8
4.1 Schéma de l'étude	8
4.2 Nombre de sujets inclus	9
4.3 Analyses statistiques	9
5. CRITERES DE SELECTION	10
5.1 Critères d'inclusion.....	10
5.2 Critères de non-inclusion.....	10
6. DEROULEMENT DE L'ETUDE	10
6.1 Inclusion.....	10
6.2 TEP 1	11
6.3 Début du traitement néoadjuvant	11
6.4 TEP2	12
6.5 Poursuite du traitement néoadjuvant et chirurgie	12
6.6 Modalités de suivi	12
6.7 Description des règles d'arrêt définitif ou temporaire.....	13
6.8. Inscription au fichier national pour les recherches :	13
6.9 Durée de l'étude.....	13
7. évaluation de la sécurité	13
7.1. Paramètres d'évaluation de la sécurité et notification/ déclaration	13
7.2 Événements / Effets indésirables, définitions.....	13
7.2.1 Définition d'un événement indésirable grave attendu (EIG-A).....	14
7.2.2 Définition d'un événement indésirable grave inattendu (EIG-I)	14
7.2.3 Critère d'intensité.....	14
8. SURVEILLANCE DE L'ETUDE	15
8.1 Comité Indépendant de Surveillance de l'Essai	15
8.2 Comité de pilotage	15

9. ASSURANCE QUALITE	15
10. PROPRIETES DES DONNEES ET CONFIDENTIALITE	16
11. REGLES DE PUBLICATIONS	16
12. ASPECTS ETHIQUES ET REGLEMENTAIRES.....	17
12.1 Comité de Protection des Personnes	17
12.2 Autorité compétente	17
12.3 Information et non-opposition des participants	17
12.4 Responsabilités du gestionnaire	17
12.5 Commission nationale de l’informatique et des libertés.....	18
12.6 Fédération des Comités de Patients pour la Recherche Clinique en Cancérologie	18
12.7 Responsabilités des investigateurs.....	18
13. ANNEXE : Déclaration d’HELSINKI	20

1. INTRODUCTION ET RATIONNEL DE L'ESSAI

Le cancer du sein est un problème majeur de santé publique puisqu'il constitue la première cause de décès par cancer chez la femme en Europe **(1)**. Pour les patients présentant au moment du diagnostic une tumeur volumineuse ou localement avancée, il est fréquemment proposé un traitement systémique par chimiothérapie et/ou thérapie ciblée précédant le traitement chirurgical (traitement néoadjuvant) dont le principal objectif est de réduire le volume tumoral et donc de réaliser un geste chirurgical conservateur **(2)**. De surcroît, le traitement néoadjuvant fournit des informations pronostiques puisque l'obtention d'une réponse histologique complète (pCR) est un facteur de bon pronostic, lié à un risque moindre de récurrence **(3)**. Actuellement le taux de pCR est de l'ordre de 20%, bien qu'il existe d'importantes variations selon le sous-type tumoral et le type de traitement administré **(3)**. L'objectif actuel des nouvelles stratégies thérapeutiques est d'augmenter ce taux de réponse. Ceci nécessite cependant une évaluation très précoce de la chimiosensibilité tumorale, afin d'adapter au plus vite le traitement. D'où la nécessité d'identifier, par des recherches translationnelles, des marqueurs (cliniques, radiologiques, biologiques, moléculaires) prédictifs précoces ou explicatifs de la qualité de la réponse histologique obtenue.

Aucun marqueur biologique n'a fait la preuve de son efficacité dans cette situation **(4)**. Des marqueurs cliniques (modifications de la taille de la tumeur évaluée cliniquement) et radiologiques (critères RECIST) ont été décrits mais ne permettent généralement qu'une évaluation tardive de la réponse tumorale **(5, 6)**. Ainsi, il existe une place pour l'imagerie fonctionnelle et moléculaire dans cette indication.

La Tomographie par Emission de Positons (TEP) au ¹⁸F-Fluorodesoxyglucose (FDG) constitue un outil unique d'évaluation *in vivo* du métabolisme glucidique des tumeurs. La TEP au FDG s'est révélée utile pour prédire précocement la réponse tumorale à la chimiothérapie néoadjuvante. De nombreuses études ont démontré qu'il existe une corrélation entre les modifications de métabolisme tumoral évalué par TEP au FDG après une ou deux cures de chimiothérapie, et la réponse histologique évaluée sur la pièce opératoire **(7-9) (Figure 1)**. Cependant, ces études préliminaires ont été réalisées sur des effectifs limités, sans tenir compte de la grande diversité des cancers du sein. Il est effectivement possible d'identifier 3 sous-types de cancer du sein, selon l'expression des récepteurs hormonaux (RH) et du récepteur pour les facteurs de croissance épidermiques humains (Human Epidermal Growth Factor Receptor 2, HER2): patients HER2+ (tumeurs surexprimant HER2), patients RH+ HER2- (tumeurs surexprimant RH mais pas HER2), et enfin patients triple négatifs (pas de surexpression ni de RH ni de HER2). Ces différents sous-types tumoraux présentent une agressivité variable, et peuvent faire l'objet d'approches thérapeutiques différentes notamment concernant le développement de thérapies ciblées **(3)**. Des travaux récents ont montré que la réponse métabolique précoce (après une ou deux cures de chimiothérapie) était très variable selon le sous-type de cancer du sein, soulignant donc l'intérêt d'une analyse par sous-groupes des données de réponse métabolique **(10-13)**.

Malgré les questions encore en suspens, l'évaluation de la réponse métabolique précoce par TEP au FDG est d'ores et déjà utilisée en routine clinique dans des études multicentriques d'escalade thérapeutique. L'étude AVATAXHER (EUDRACT 2009-013410-26, médecin coordonnateur Dr Bruno Coudert), a consisté à proposer un renforcement thérapeutique par ajout de bevacizumab, un

traitement antiangiogénique, aux patientes traitées par chimiothérapie néoadjuvante pour un cancer du sein surexprimant HER2 et présentant une réponse métabolique précoce jugée insuffisante, faisant craindre l'absence de réponse histologique complète **(14)**. Les résultats sont en cours d'analyse, mais sont très prometteurs quant au rôle potentiel de la TEP au FDG dans ce type de stratégie d'escalade thérapeutique.

Nous avons également développé une technique d'analyse des images TEP, permettant d'évaluer la perfusion tumorale conjointement au métabolisme glucidique, lors d'un même examen **(15)**. Cette approche est particulièrement intéressante dans le contexte de la prescription personnalisée de thérapies antiangiogéniques ciblant la perfusion tumorale.

L'évaluation de la réponse métabolique précoce repose donc sur la comparaison de la TEP initiale (TEP1), avant de débiter le traitement néoadjuvant, et de la TEP de contrôle (TEP2), après une ou deux cures. La TEP au FDG initiale est réalisée en routine pour le bilan d'extension (recherche de métastases à distance), suivant les recommandations de l'Institut National du Cancer (INCa) **(16)**. Le « guide du bon usage des examens d'imagerie médicale » recommandé par l'HAS, indique que la TEP au FDG peut être utile en phase néoadjuvante pour évaluer la réponse métabolique précoce **(17)**. Le Collège National des Gynécologues et Obstétriciens Français (CNGOF) recommande une surveillance précoce des patientes sous chimiothérapie néoadjuvante afin de ne pas méconnaître une progression sous traitement, rare mais toujours possible **(18, 19)**. Ainsi, les 2 examens (TEP1 et TEP2) se sont imposés dans la prise en charge standard des patientes traitées par chimiothérapie néoadjuvante pour un cancer du sein large ou localement avancé.

Ainsi, notre objectif est d'évaluer conjointement les modifications métaboliques et perfusionnelles précoces observées dans la tumeur primitive en exploitant les données de la TEP au FDG réalisée au diagnostic et après la première cure de chimiothérapie néoadjuvante, afin d'identifier d'une part des marqueurs métaboliques prédictifs précoces de la réponse histologique tumorale et d'autre part des facteurs métaboliques pronostiques indiquant le risque de rechute locale ou métastatique. Une fois ces facteurs prédictifs et pronostiques identifiés, cela permettra de développer des protocoles d'adaptation précoce des schémas de chimiothérapies selon l'évaluation de la réponse (ou de la non-réponse) tumorale évaluée par ces marqueurs.

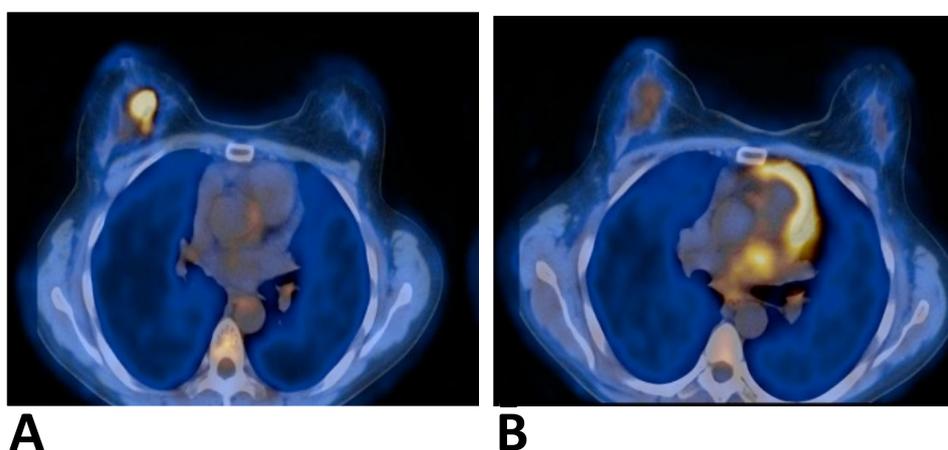


Figure 1 : Exemple de modification du métabolisme tumoral avant traitement (A) et après une cure de chimiothérapie néoadjuvante (B) chez une patiente présentant un cancer du sein droit.

● références bibliographiques

1. Malvezzi M, Bertuccio P, Levi F, La Vecchia C, Negri E. European cancer mortality predictions for the year 2013. *Ann Oncol* 2013;24:792-800.
2. Bonadonna G, Valagussa P, Brambilla C, et al. Primary chemotherapy in operable breast cancer: eight-year experience at the Milan Cancer Institute. *J Clin Oncol*. 1998;16(1):93-100.
3. von Minckwitz G, Untch M, Blohmer JU, et al. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J Clin Oncol* 2012;30:1796-1804.
4. Bottini A, Berruti A, Tampellini M, et al. Influence of neoadjuvant chemotherapy on serum tumor markers CA 15-3, MCA, CEA, TPS and TPA in breast cancer patients with operable disease. *Tumour biology* 1997;18(5):301-10.
5. Chevallier B, Roche H, Olivier JP, et al. Inflammatory breast cancer. Pilot study of intensive induction chemotherapy (FEC-HD) results in a high histologic response rate. *Am J Clin Oncol* 1993 Jun;16(3):223-8.
6. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000;92(3):205-16.
7. Cochet A, Generali D, Fox SB, et al. Positron emission tomography and neoadjuvant therapy of breast cancer. *J Natl Cancer Inst Monograph* 2011;43:111-115.
8. Wang Y, Zhang C, Liu J, et al. Is 18F-FDG PET accurate to predict neoadjuvant therapy response in breast cancer? A meta-analysis. *Breast Cancer Res Treat* 2012;131:357-369.
9. Kostakoglu L. Radionuclide response assessment of breast cancer. *Semin Nucl Med* 2013;43:299-316.
10. Humbert O, Berriolo-Riedinger A, Riedinger JM, et al. Changes in 18F-FDG tumor metabolism after a first course of neoadjuvant chemotherapy in breast cancer: influence of tumor subtypes. *Ann Oncol* 2012;23(10):2572-2577.
11. Groheux D, Hindié E, Giacchetti S, et al. Triple-negative breast cancer: early assessment with 18F-FDG PET/CT during neoadjuvant chemotherapy identifies patients who are unlikely to achieve a pathological complete response and are at high risk of early relapse. *J Nucl Med* 2012;53(2):249-254.
12. Groheux D, Giacchetti S, Hatt M, et al. HER2-overexpressing breast cancer: FDG uptake after two cycles of chemotherapy predicts the outcome of neoadjuvant treatment. *Br J Cancer* 2013;109(5):1157-1164.

13. Koolen BB, Pengel KE, Wesseling J, et al. FDG PET/CT during neoadjuvant chemotherapy may predict response in ER-positive/HER2-negative and triple negative, but not in HER2-positive breast cancer. *Breast* 2013;22(5):691-697.
14. Cochet A, Kerrou K, Nabholz JM, et al. An open-label randomized, multicenter, phase II study on neoadjuvant treatment with trastuzumab plus docetaxel versus trastuzumab plus docetaxel plus bevacizumab according to positron emission tomography (PET) value modification in patients with early stage HER2-positive breast cancer (AVATAXHER): Design description. *J Clin Oncol (meeting abstract)* 2012;30(15):TP S646.
15. Cochet A, Pigeonnat S, Khoury B, et al. Evaluation of breast tumor blood flow with dynamic first-pass ¹⁸F-FDG PET/CT: comparison with angiogenesis markers and prognostic factors. *J Nucl Med* 2012;53(4):512-520.
16. Cancer du sein infiltrant non métastatique - Questions d'actualités - Synthèse. Collection Avis & Recommandations, INCa, Boulogne-Billancourt, juillet 2012.
17. Guide du bon usage des examens d'imagerie médicale. Société Française de Radiologie, Société Française de Médecine Nucléaire. <http://gbu.radiologie.fr>. Accessed February 14 2014.
18. Chimiothérapie néoadjuvante. Mises à jour en Gynécologie Médicale 2008. Collège National des Gynécologues et Obstétriciens Français. http://www.cngof.asso.fr/d_livres/2008_GM_317_mauriac.pdf. Accessed February 14 2014.
19. Fisher B, Brown A, Mamounas E, et al. Effect of preoperative chemotherapy on local-regional disease in women with operable breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-18. *J Clin Oncol* 1997;15(7):2483-2493.

2. OBJECTIFS

2.1 Objectif principal :

Evaluer l'intérêt de la réponse métabolique précoce pour prédire la réponse histologique complète (pCR) au décours de la chimiothérapie néoadjuvante chez des patientes pris en charge au Centre Georges-François Leclerc pour un cancer du sein opérable, en exploitant les données de la TEP au FDG réalisée avant traitement (TEP1) et après une cure de chimiothérapie (TEP2), ces deux examens étant réalisés dans le cadre de la prise en charge standard de ces patients.

2.2 Objectifs secondaires

1°) Evaluer l'intérêt de la réponse métabolique précoce pour prédire la survie sans évolution de la maladie.

2°) Evaluer les modifications de la perfusion tumorale avant et après une cure de chimiothérapie néoadjuvante en exploitant les données des TEP au FDG (TEP1 et TEP2) réalisées en mode dynamique.

3°) Evaluer les relations entre les caractéristiques métaboliques et les caractéristiques immuno-histo-chimiques des tumeurs mammaires.

4°) Evaluer l'impact des marqueurs métaboliques extra-tumoraux tels que le métabolisme de la glande mammaire saine

3. CRITERES DE JUGEMENTS

3.1 Critère de jugement principal

Taux de réponse histologique complète (pCR) évaluée sur la pièce opératoire après traitement chirurgical selon la classification de Chevallier, au décours de la chimiothérapie néoadjuvante.

3.2 Critères de jugements secondaires

1°) Taux de survie à 5 ans sans évolution ou récurrence de la maladie.

2°) Taux de réponse histologique complète (pCR) évaluée sur la pièce opératoire après traitement chirurgical, au décours de la chimiothérapie néoadjuvante.

4. METHODOLOGIE

4.1 Schéma de l'étude

Il s'agit d'une étude monocentrique de soins courants, prospective, observationnelle, non randomisée. La **Figure 2** représente le schéma simplifié de l'étude.

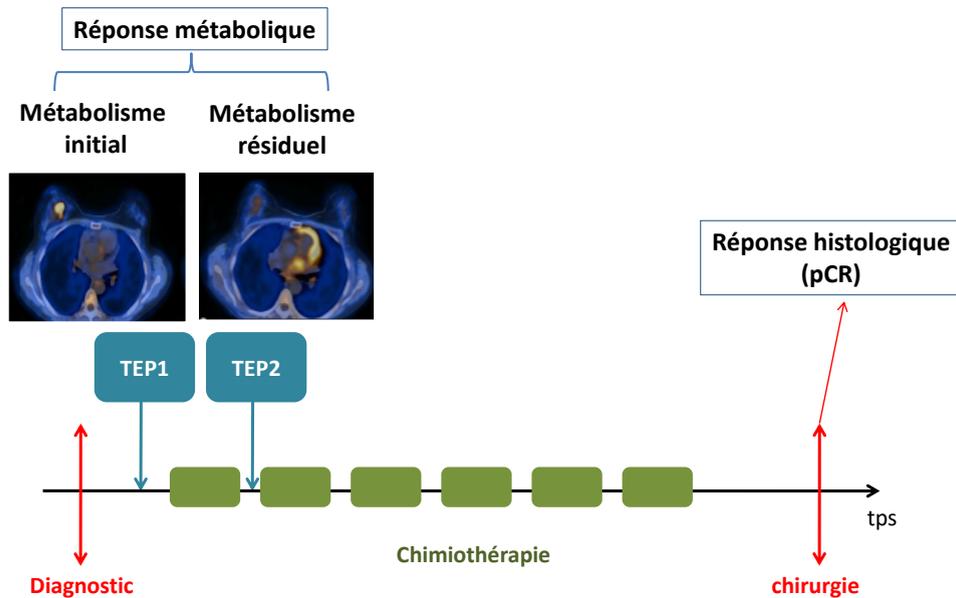


Figure 2: schéma simplifié de l'étude TREN. Les données de la TEP au FDG réalisée avant le début du traitement (TEP1) seront exploitées pour quantifier le métabolisme initial de la tumeur. Les données de la TEP au FDG réalisée après une cure de chimiothérapie (TEP2) seront utilisées pour évaluer le métabolisme tumoral résiduel et la réponse métabolique en comparaison avec les résultats de la TEP1.

4.2 Nombre de sujets inclus

Trois groupes de patients seront étudiés séparément, en fonction du phénotype immuno-histo-chimique de la tumeur: patients HER2+ (tumeurs surexprimant HER2), patients RH+ HER2- (tumeurs surexprimant RH mais pas HER2), et enfin patients triple négatifs (pas de surexpression ni de RH ni de HER2). D'après les données de la littérature et notre expérience personnelle au CGFL, la répartition attendue des patients présentant un cancer du sein large ou localement avancé dans chacun des 3 groupes est la suivante: HER2+: 25% ; RH+ HER2-: 60% ; Triple négatifs: 15%. Le taux de pCR espéré dans chacun de ces groupes est respectivement: 40%, 10% et 50%.

On prévoit donc le recrutement de 200 patients de manière à obtenir un minimum de 30 patients dans le groupe avec l'effectif le plus faible (triples négatifs).

On prévoit une durée de recrutement de 4 ans soit 50 patients par an, ce qui correspond approximativement à la moitié du nombre de patients pris en charge chaque année au CGFL pour un traitement néoadjuvant du cancer du sein.

4.3 Analyses statistiques

Pour chaque examen TEP (TEP1 et TEP2) le métabolisme tumoral (SUV), le volume et le débit de perfusion tumoral (BF) seront calculés. Pour chaque patient, les variations des différents paramètres métaboliques et perfusionnels entre la TEP1 et la TEP2 (par exemple, Δ SUV et Δ BF) seront calculées. Ces résultats seront, dans un premier temps, comparés à la réponse histologique

évaluée sur la pièce opératoire. Ils seront ensuite comparés au pronostic des patients (un suivi minimal de 5 ans sera nécessaire)

L'analyse des données sera effectuée selon les sous-types de tumeur du sein. Dans chaque groupe, une régression logistique sera effectuée afin d'identifier les facteurs prédictifs indépendants de la réponse histologique. Une analyse ROC sera également réalisée pour déterminer une valeur seuil optimale des modifications métaboliques et perfusionnelles permettant de prédire la réponse histologique. Enfin une analyse de Cox sera réalisée pour évaluer la valeur pronostique des paramètres TEP.

5. CRITERES DE SELECTION

5.1 Critères d'inclusion

- ✓ Patiente ayant été informé de l'étude
- ✓ Patiente affilié à un régime de sécurité sociale
- ✓ Femme de 18 ans ou plus (seuls les cancers du sein féminins seront inclus)
- ✓ Patient présentant un cancer du sein prouvé histologiquement
- ✓ Patient adressée au CGFL avec une Indication de traitement néoadjuvant (stade II ou III classification UICC)
- ✓ Patient ayant signé au moment de son admission au CGFL l'autorisation d'exploitation des données cliniques, biologiques et d'imagerie

5.2 Critères de non-inclusion

Liés à la maladie :

- ✓ Présence de métastase(s) à distance

Liés au traitement :

- ✓ Contre-indication au traitement néoadjuvant
- ✓ Contre-indication à la chirurgie

Liés aux conditions des patients :

- ✓ Refus de la patiente
- ✓ Maladie grave non équilibrée, infection sous-jacente susceptible d'empêcher la patiente de recevoir le traitement
- ✓ Grossesse ou allaitement
- ✓ Diabète non contrôlé lors des examens TEP (glycémie > 8 mmol/L)
- ✓ Maladie psychiatrique compromettant la compréhension de l'information ou la réalisation de l'étude
- ✓ Patient sous tutelle, curatelle ou sauvegarde de justice

6. DEROULEMENT DE L'ETUDE

6.1 Inclusion

Les recruteurs seront les médecins titulaires (oncologues médicaux, chirurgiens, médecins nucléaires) du CGFL.

Les patientes répondant aux critères de sélection seront sélectionnées pour participer à l'étude TREN parmi l'ensemble des patients pris en charge au CGFL pour traitement néoadjuvant d'un cancer du sein.

Les patientes seront informées oralement de la réalisation de l'étude par les médecins investigateurs et après un délai de réflexion, il leur sera demandé s'ils s'opposent à leur participation.

Indépendamment de cette étude, il est demandé aux patientes prises en charge au CGFL de signer une autorisation d'exploitation des données cliniques, biologiques et d'imagerie réalisées au CGFL.

La non opposition des patientes sera renseignée dans leurs dossiers médicaux par le médecin investigateur ayant recueilli le consentement.

6.2 TEP 1

La TEP1 sera réalisée avant de débuter le traitement néoadjuvant, dans le cadre du bilan d'extension initial. L'examen sera réalisé selon les modalités suivantes:

- ✓ Délai d'au moins 5 jours après la biopsie
- ✓ Patient à jeun depuis au moins 6 heures avant l'injection de FDG.
- ✓ Pesée et mesure de la taille du patient
- ✓ Prélèvement intraveineux pour un dosage de la glycémie plasmatique.
- ✓ Patient positionné dans la caméra TEP en décubitus ventral, bras levés, sur un matelas dédié à la mammo-scintigraphie
- ✓ Injection intraveineuse en bolus (1 mL/s) de 3 MBq/Kg de FDG
- ✓ Immédiatement après injection, réalisation d'une acquisition TEP dynamique centrée sur les seins pendant 8 minutes suivie d'une acquisition tomодensitométrique de la même région pour le repérage anatomique (120 kVp, 120 mAs)
- ✓ Patient installé en salle d'attente, en position couchée, au repos strict, pendant environ 50 minutes
- ✓ Ré-installation en décubitus dorsal pour une acquisition TEP de la racine des cuisses jusqu'à la tête, suivie d'une acquisition tomодensitométrique de la même région pour le repérage anatomique et la correction d'atténuation (120 kVp, 120 mAs).
- ✓ Réalisation d'une dernière acquisition en décubitus ventral sur le matelas dédié, bras levés, pour une acquisition TEP centrée sur la région mammaire (2 pas de 3 minutes), suivie d'une acquisition tomодensitométrique de la même région pour le repérage anatomique et la correction d'atténuation (120 kVp, 120 mAs).

Les données (Standardized Uptake Value [SUV] pour l'étude du métabolisme tumoral, débit de perfusion [BF] pour l'étude de la perfusion tumorale) seront recueillies.

6.3 Début du traitement néoadjuvant

Le traitement néoadjuvant sera débuté après la réalisation de la TEP1, selon les modalités habituelles. La prise en charge médicale ne sera pas adaptée à la réponse métabolique précocement observée, sauf en cas de progression tumorale. Dans ce cas, la décision de poursuivre, de modifier ou d'interrompre le traitement néoadjuvant sera laissée à l'oncologue prenant en charge la patiente.

Les patients étant pris en charge dans le cadre du soin courant, aucun suivi de l'utilisation des traitements ne sera réalisé.

6.4 TEP2

La TEP2 sera réalisée au moins deux semaines après le début de la première cure de chimiothérapie, idéalement la veille ou le matin de la seconde cure. Cet examen sera donc réalisé dans le cadre de l'évaluation précoce de la réponse à la chimiothérapie et du dépistage précoce d'une évolution sous traitement. L'examen sera réalisé selon les modalités suivantes:

- ✓ Patient à jeun depuis au moins 6 heures avant l'injection de FDG.
- ✓ Pesée et mesure de la taille du patient
- ✓ Prélèvement intraveineux pour un dosage de la glycémie plasmatique.
- ✓ Patient positionné dans la caméra TEP en décubitus ventral, bras levés, sur un matelas dédié à la mammo-scintigraphie
- ✓ Injection intraveineuse en bolus (1 mL/s) de 3 MBq/Kg de FDG
- ✓ Immédiatement après injection, réalisation d'une acquisition TEP dynamique centrée sur les seins pendant 8 minutes suivie d'une acquisition tomодensitométrique de la même région pour le repérage anatomique (120 kVp, 120 mAs)
- ✓ Patient installé en salle d'attente, en position couchée, au repos strict, pendant environ 70 minutes
- ✓ Ré-installation en décubitus ventral sur le matelas dédié, bras levés, pour une acquisition TEP centrée sur la région mammaire (2 pas de 3 minutes), en respectant le même délai d'acquisition que pour la TEP1 (environ 90 +/- 10 minutes après l'injection de FDG), suivie d'une acquisition tomодensitométrique de la même région pour le repérage anatomique et la correction d'atténuation (120 kVp, 120 mAs).

Les données (SUV pour l'étude du métabolisme tumoral, BF pour l'étude de la perfusion tumorale) seront recueillies.

6.5 Poursuite du traitement néoadjuvant et chirurgie

Le traitement néoadjuvant sera poursuivi selon les modalités habituelles (généralement 6 cures de chimiothérapie au total). Le traitement chirurgical sera également réalisé selon les modalités habituelles.

L'étude histologique de la pièce opératoire inclut une analyse de la réponse histologique au traitement néoadjuvant selon la classification de Chevallier.

6.6 Modalités de suivi

Le suivi clinique, biologique et radiologique des patients se fera selon les schémas habituels de surveillance du cancer du sein.

6.7 Description des règles d'arrêt définitif ou temporaire

L'essai peut être suspendu ou arrêté par le gestionnaire en concertation avec le coordonnateur à la demande de l'Autorité Compétente et /ou du Comité de Protection des Personnes (CPP) pour les raisons suivantes :

- ✓ Un recrutement insuffisant de patients,
- ✓ Une qualité de recueil de données insuffisante.

68. Inscription au fichier national pour les recherches :

Les patients participant à cette étude ne seront pas inscrits au fichier national pour les recherches.

6.9 Durée de l'étude

- ✓ Durée d'inclusion : 2 ans
- ✓ Durée de suivi pour une personne : 5 ans.
- ✓ Durée totale de l'étude : 7 ans

7. évaluation de la sécurité

7.1. Paramètres d'évaluation de la sécurité et notification/ déclaration

Conformément à la réglementation en vigueur, il n'y a pas d'obligation concernant la notification et le suivi des effets indésirables pour l'investigateur. Cependant, l'investigateur dans le cadre de cet essai relèvera les effets indésirables liés à la recherche en soins courants. Ce dernier notifiera ces événements aux services de vigilance de l'établissement conformément aux procédures en vigueur dans l'institution.

De même, il n'existe pas d'obligation pour le gestionnaire à déclarer les EIG. Ce dernier enregistrera pour information les EIG qui lui seront notifiés par l'investigateur.

7.2 Événements / Effets indésirables, définitions

Est considéré comme un événement indésirable grave (EIG) tout événement :

- ✓ Entraînant le décès,
- ✓ Mettant en jeu le pronostic vital,
- ✓ Entraînant une hospitalisation ou une prolongation d'hospitalisation,
- ✓ Provoquant une invalidité permanente ou une incapacité temporaire grave,
- ✓ Provoquant une anomalie congénitale, une malformation fœtale ou un avortement,

- ✓ Médicalement significatif.

Les termes invalidité et incapacité correspondent à tout handicap physique ou psychique temporaire ou permanent, cliniquement significatif et retentissant sur l'activité physique et/ou la qualité de vie du patient.

Est considéré comme médicalement significatif tout événement clinique ou résultat de laboratoire considéré comme grave par l'investigateur et ne correspondant pas aux critères de gravité définis ci-dessus. Ils peuvent faire courir un risque au patient et nécessitent une intervention médicale pour prévenir une issue correspondant à l'un des critères de gravité mentionnés précédemment (exemples : surdosages, seconds cancers, grossesses et faits nouveaux peuvent être considérés comme médicalement significatifs).

N'est pas considéré comme un EIG :

Une hospitalisation < à 24 heures,

Une hospitalisation programmée préalablement au début de l'essai et/ou prévue par le protocole (biopsie, chimiothérapie..).

Toute hospitalisation et tout décès en relation avec la progression du cancer et survenant au-delà de la période de notification définie ci-dessous.

Les événements survenant dans le cadre de la progression de la maladie incluant le décès ne seront également pas déclarés en tant qu'EIG.

7.2.1 Définition d'un événement indésirable grave attendu (EIG-A)

Un EIG-A est un événement déjà mentionné dans la version la plus récente de la brochure investigateur ou dans le résumé des caractéristiques du produit (RCP) pour les médicaments ayant déjà une autorisation de mise sur le marché (AMM). Cette définition s'applique également au médicament de l'essai lorsqu'il est administré pour une même population hors indication de l'AMM.

Considérant la réalisation de TEP au FDG en routine clinique et l'exploitation des données qui en résulte, aucun EIG n'est attendu.

7.2.2 Définition d'un événement indésirable grave inattendu (EIG-I)

Un EIG-I est un événement non mentionné ou différent par sa nature, son intensité, son évolution par rapport à la brochure investigateur ou au résumé des caractéristiques du produit (RCP) pour les médicaments ayant une autorisation de mise sur le marché (AMM).

7.2.3 Critère d'intensité

Le critère d'intensité ne doit pas être confondu avec le critère de gravité qui sert de guide pour définir les obligations de déclaration.

L'intensité des événements sera estimée selon l'extrait de la classification CTC-AE version 4.0. L'intensité des événements indésirables non listés dans cette classification sera appréciée selon les qualificatifs suivants :

Légère (grade 1) : n'affecte pas l'activité quotidienne habituelle du patient

Modérée (grade 2) : perturbe l'activité quotidienne habituelle du patient

Sévère (grade 3) : empêche l'activité quotidienne habituelle du patient

Très Sévère (grade 4) : impose des mesures de réanimation/ menace le pronostic vital

Décès (grade 5)

7.3 Evénements/effets indésirables attendus liés à la participation à cette étude

Considérant la longue expérience de la réalisation de TEP au FDG en routine clinique, aucun événement indésirable, grave ou non grave, n'est attendu.

Les 2 examens TEP-TDM au FDG seront réalisés dans le cadre de la prise en charge standard des patients traités par chimiothérapie néoadjuvante.

L'injection de FDG est réalisée en routine clinique depuis 2003 au CGFL (plus de 3000 examens en 2013). Aucun effet indésirable lié à cette injection n'a été observé.

L'irradiation liée à une TEP-TDM au FDG est évaluée en moyenne à 10 millisievert (mSv) (en incluant l'irradiation liée à l'injection du FDG et l'irradiation liée aux scanners de repérage).

La participation à cette étude aura pour seule conséquence la réalisation de 2 scanners de repérages supplémentaires réalisés immédiatement après l'acquisition dynamique, correspondant à une sur-irradiation d'environ 4 mSv au total (à titre de comparaison, un examen TDM thoraco-abdomino-pelvien à visée diagnostique entraîne une irradiation d'environ 15 à 20 mSv).

8. SURVEILLANCE DE L'ETUDE

8.1 Comité Indépendant de Surveillance de l'Essai

Compte tenu de la nature de l'essai aucun comité de surveillance de l'essai ne sera mis en place.

8.2 Comité de pilotage

L'étude sera encadrée et surveillée par un comité de pilotage composé de membres participant à l'étude.

9. ASSURANCE QUALITE

Afin de garantir l'authenticité et la crédibilité des données conformément aux BPC, le gestionnaire met en place un système d'assurance qualité qui comprend :

- la gestion de l'essai selon les procédures du CGFL,
- le contrôle qualité des données du site investigateur par le moniteur dont le rôle est :
- de vérifier la concordance et la cohérence des données du cahier d'observation par rapport aux documents-sources,
- l'identification de toutes les données à recueillir directement dans les cahiers d'observation, qui seront considérés comme des données sources,

- de s'assurer, le cas échéant, que les personnes susceptibles de se prêter à la recherche ne participent pas déjà à une recherche qui pourrait rendre impossible leur inclusion dans la recherche proposée. Il s'assure également qu'elles n'ont pas participé à une recherche pour laquelle actuellement une période d'exclusion est requise.

10. PROPRIETES DES DONNEES ET CONFIDENTIALITE

L'investigateur s'engage, pour lui-même et pour toutes les personnes amenées à suivre le déroulement de l'essai, à garantir la confidentialité de toutes les informations fournies par le Centre Georges François Leclerc jusqu'à la publication des résultats de l'essai.

Cette obligation de confidentialité ne s'appliquera pas aux renseignements que l'investigateur sera amené à communiquer aux patients dans le cadre de leur participation à l'essai ni aux informations déjà publiées.

L'investigateur s'engage à ne pas publier, divulguer ou utiliser, de quelque façon que ce soit, directement ou indirectement, les informations scientifiques ou techniques en relation avec l'essai.

Néanmoins, conformément à l'article R 5121-13 du Code de la Santé Publique, le centre et l'investigateur pourront donner des informations relatives à l'essai :

- au ministre chargé de la santé,
- aux médecins inspecteurs de santé publique,
- aux pharmaciens inspecteurs de santé publique,
- au Directeur Général et aux inspecteurs de l'ANSM.

L'essai ne pourra faire l'objet d'aucun commentaire écrit ou oral sans l'accord du gestionnaire ; l'ensemble des informations communiquées ou obtenues pendant la réalisation de l'essai appartenant de plein droit au centre Georges François Leclerc qui pourra librement en disposer.

11. REGLES DE PUBLICATIONS

Toutes les informations résultant de cet essai sont considérées comme confidentielles, au moins jusqu'à ce que l'analyse appropriée et le contrôle par le gestionnaire, l'investigateur coordonnateur et le statisticien de l'essai soient achevés.

Toutes publications, abstracts ou présentations comprenant les résultats de l'essai doivent être soumis pour approbation au gestionnaire (Centre Georges François Leclerc).

Par ailleurs, toutes communications, manuscrits ou présentations doivent comporter une rubrique qui mentionne impérativement au Centre Georges François Leclerc, toutes les institutions, les investigateurs, les groupes coopérateurs, les sociétés savantes qui ont contribué à la réalisation de l'essai, de même que les organismes qui ont supporté financièrement la recherche.

Pour la publication principale, française ou anglophone, les auteurs seront:

- les coordonnateurs de l'étude
- les investigateurs au prorata du nombre de patients recrutés
- le méthodologiste/statisticien de l'étude
- un représentant du gestionnaire

De même, les publications des résultats annexes comporteront le nom de la personne ayant réalisé le travail annexe ainsi que les noms de toutes les autres personnes concernées par ce travail annexe.

En cas de litige, l'ordre des auteurs sera arbitré par le Comité de Rédaction de l'Essai (CRE).

12. ASPECTS ETHIQUES ET REGLEMENTAIRES

L'essai clinique doit être conduit conformément :

- aux principes éthiques de la dernière version en vigueur de la déclaration d'Helsinki,
- aux Bonnes Pratiques Cliniques de la Conférence Internationale d'Harmonisation (ICH–E6, 17/07/96),
- à la Directive Européenne (2001/20/CE) sur la conduite des essais cliniques,
- à la loi Huriet (n° 88-1138) du 20 décembre 1988 relative à la Protection des Personnes se prêtant à la Recherche Biomédicale et modifiée par la loi de Santé Publique (n° 2004-806) du 9 août 2004,
- à la loi Informatique et Libertés n° 78-17 du 6 janvier 1978 modifiée par la loi n° 2004-801 du 6 août 2004 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel,
- à la loi bioéthique n° 2004-800 du 6 août 2004.

12.1 Comité de Protection des Personnes

Avant de réaliser une recherche biomédicale sur l'être humain, le gestionnaire est tenu d'en soumettre le projet à l'avis de l'un des comités de protection des personnes compétents pour le lieu où l'investigateur coordonnateur exerce son activité.

La demande d'avis du projet de recherche biomédicale est adressée au comité par le gestionnaire.

Les demandes de modifications substantielles des projets initiaux sont également adressées par le gestionnaire du comité pour avis.

12.2 Autorité compétente

Avant de réaliser ou de faire réaliser une recherche biomédicale le gestionnaire de cette recherche enregistre l'essai auprès de l'autorité compétente.

12.3 Information et non-opposition des participants

Préalablement à la réalisation d'une recherche biomédicale sur une personne, cette dernière doit être informée par l'investigateur lors d'une consultation et après un délai de réflexion suffisant, signifier sa non-opposition à sa participation. La trace de cette non-opposition doit être enregistrée dans le dossier source du patient.

L'information destinée aux participants de l'essai doit comprendre l'ensemble des éléments définis dans la loi de santé publique du 9 août 2004 et doit être écrite de façon simple, dans un langage compréhensible par le patient

12.4 Responsabilités du gestionnaire

Le gestionnaire de l'essai clinique est la personne physique ou morale qui prend l'initiative de la recherche biomédicale sur l'être humain, en assure la gestion et vérifie que son financement est prévu.

Le gestionnaire doit être établi dans la communauté européenne ou à défaut disposer d'un représentant légal dans un état membre.

Les principales responsabilités du gestionnaire sont :

- l'enregistrement de l'essai dans une base de données publique et l'obtention du n° ID RCB,
- la demande d'avis du projet initial et des amendements substantiels auprès du comité de protection des personnes (CPP),
- l'information de l'essai aux investigateurs,
- la déclaration à l'autorité compétente du début et de la fin de l'essai,
- la rédaction du rapport final de l'essai
- l'information des résultats de l'essai à l'autorité compétente, au CPP et aux participants à la recherche,
- l'archivage des documents essentiels de l'essai dans le classeur gestionnaire pour une durée minimale de 15 ans après la fin de la recherche

12.5 Commission nationale de l'informatique et des libertés

Le fichier informatique utilisé pour réaliser la présente recherche fera l'objet d'une autorisation de la CNIL en application des articles 40-1 est suivants de la loi « informatique et liberté ». Les données médicales concernant les patients font l'objet d'un traitement informatique et ne seront transmises qu'au gestionnaire, ainsi le cas échéant aux autorités sanitaires habilités, dans les conditions garantissant leur confidentialité. Les patients pourront exercer leurs droits d'accès et de rectification auprès du médecin investigateur qui les suit. Ce dernier transmettra cette demande au gestionnaire.

12.6 Fédération des Comités de Patients pour la Recherche Clinique en Cancérologie

La Fédération des Comités de Patients pour la Recherche Clinique en Cancérologie (FCPRCC) a été créée à l'initiative de la Fédération des Centres de Lutte Contre le Cancer (FNCLCC) et de la Ligue Nationale Contre le Cancer afin de relire les protocoles d'essais cliniques en cancérologie. Cette Fédération des Comités de Patients, coordonnée par le Bureau d'Etudes Cliniques et Thérapeutiques (BECT) de la FNCLCC, regroupe à la fois les comités de patients de la Ligue ainsi que d'autres établissements de santé. Elle s'engage à relire le protocole et à proposer des améliorations portant notamment sur la qualité de la lettre d'information, la mise à disposition d'un plan de traitement et de surveillance, la suggestion de mesures visant à améliorer le confort des patients.

12.7 Responsabilités des investigateurs

L'investigateur principal s'engage à conduire l'essai clinique conformément au protocole qui a été approuvé par le CPP.

L'investigateur ne doit apporter aucune modification au protocole sans l'autorisation écrite du gestionnaire et sans que le CPP aient donné leur autorisation sur les modifications proposées.

Il est de la responsabilité de l'investigateur principal :

- de fournir au gestionnaire son curriculum vitae ainsi que ceux des collaborateurs,
- d'identifier les membres de son équipe qui participent à l'essai et de définir leurs responsabilités,
- d'assurer le recrutement des patients après autorisation du gestionnaire,

Il est de la responsabilité de chaque investigateur :

- d'informer et de recueillir la non-opposition de chaque participant à la recherche avant toute procédure de sélection spécifique à l'essai,
- de compléter régulièrement les cahiers d'observation (CRF) pour chacun des patients inclus dans l'essai
- de dater, corriger et de signer les corrections des CRF pour chacun des patients inclus dans l'essai,
- d'accepter éventuellement les visites des inspecteurs des autorités de tutelle.

Toute la documentation relative à l'essai (protocole, consentements, cahiers d'observation, dossier investigateur, etc...), ainsi que les documents originaux (résultats de laboratoire, radiologies, comptes rendus de consultations, rapports d'exams cliniques, etc.) sont considérés comme confidentielle et doivent être détenus dans un lieu sûr. L'investigateur Principal devra conserver les données ainsi qu'une liste d'identification des patientes pendant une durée minimale de 15 ans après la fin de l'étude.

13. ANNEXE : Déclaration d'HELSINKI

Déclaration d'Helsinki de L'AMM - Principes éthiques applicables à la recherche médicale impliquant des êtres humains

Adoptée par la 18e Assemblée générale de l'AMM, Helsinki, Finlande, Juin 1964 et amendée par les

29e Assemblée générale de l'AMM, Tokyo, Octobre 1975 35e Assemblée générale de l'AMM, Venise, Octobre 1983

41e Assemblée générale de l'AMM, Hong Kong, Septembre 1989

48e Assemblée générale de l'AMM, Somerset West (Afrique du Sud), Octobre 1996

52e Assemblée générale de l'AMM, Edimbourg, Ecosse, Octobre 2000

53e Assemblée générale de l'AMM, Washington, Etats Unis, 2002 (ajout d'une note de clarification pour le paragraphe 29)

55e Assemblée générale de l'AMM, Tokyo, Japon 2004 (ajout d'une note de clarification concernant le paragraphe 30)

59e Assemblée générale de l'AMM, Séoul, Corée, Octobre 2008

A-INTRODUCTION

1. L'Association Médicale Mondiale (AMM) a élaboré la Déclaration d'Helsinki comme un énoncé de principes éthiques applicables à la recherche médicale impliquant des êtres humains, y compris la recherche sur du matériel biologique humain et sur des données identifiables. La Déclaration est conçue comme un tout indissociable. Aucun paragraphe ne peut être appliqué sans tenir compte de tous les autres paragraphes pertinents.

2. Cette Déclaration s'adresse principalement aux médecins. L'AMM invite cependant les autres participants à la recherche médicale impliquant des êtres humains à adopter ces principes.

3. Le devoir du médecin est de promouvoir et de sauvegarder la santé des patients, y compris celles des personnes impliquées dans la recherche médicale. Le médecin comprend celles des personnes impliquées dans la recherche médicale. Le médecin consacre son savoir et sa conscience à l'accomplissement de ce devoir.

4. La Déclaration de Genève de l'AMM engage les médecins en ces termes: «La santé de mon patient prévaut sur toutes les autres considérations » et le Code International d'Ethique Médicale déclare qu'un «médecin doit agir dans le meilleur intérêt du patient lorsqu'il le soigne».

5. Le progrès médical est basé sur la recherche qui, en définitive, doit comprendre des études impliquant des êtres humains. Des possibilités appropriées de participer à la recherche médicale devraient être offertes aux populations qui y sont sous- représentées.

6. Dans la recherche médicale impliquant des êtres humains, le bien-être de chaque personne impliquée dans la recherche doit prévaloir sur tous les autres intérêts.

7. L'objectif premier de la recherche médicale impliquant des êtres humains est de comprendre les causes, le développement et les effets des maladies et d'améliorer les interventions préventives, diagnostiques et thérapeutiques (méthodes, procédures et traitements). Même les meilleures interventions courantes doivent être évaluées en permanence par des recherches portant sur leur sécurité, leur efficacité, leur pertinence, leur accessibilité et leur qualité.

8. Dans la pratique médicale et la recherche médicale, la plupart des interventions comprennent des risques et des inconvénients.

9. La recherche médicale est soumise à des normes éthiques qui promeuvent le respect de tous les êtres humains et qui protègent leur santé et leurs droits. Certaines populations faisant l'objet de recherches sont particulièrement vulnérables et ont besoin d'une protection spéciale. Celles-ci incluent les personnes qui, d'elles-mêmes, ne sont pas en mesure de donner ou de refuser leur consentement et celles qui peuvent être vulnérables à la coercition ou à des influences indues.

10. Dans la recherche médicale impliquant des êtres humains, les médecins devraient tenir compte des normes et standards éthiques, légaux et réglementaires applicables dans leur propre pays ainsi que des normes et standards internationaux. Les protections garanties par la présente Déclaration aux personnes impliquées dans la recherche ne peuvent être restreintes ou exclues par aucune disposition éthique, légale ou réglementaire, nationale ou internationale.

B-PRINCIPES APPLICABLES A TOUS LES TYPES DE RECHERCHE MEDICALE

11. Il est du devoir des médecins participant à la recherche médicale de protéger la vie, la santé, la dignité, l'intégrité, le droit à l'auto-détermination, la vie privée et la confidentialité des informations des personnes impliquées dans la recherche.

12. La recherche médicale impliquant des êtres humains doit se conformer aux principes scientifiques généralement acceptés, se baser sur une connaissance approfondie de la littérature scientifique, sur d'autres sources pertinentes d'informations et sur des expériences appropriées en laboratoire et, le cas échéant, sur les animaux. Le bien être des animaux utilisés dans la recherche doit être respecté

13. Une prudence particulière s'impose dans la conduite de recherches susceptibles de nuire à l'environnement.

14. La conception et la conduite de toutes les études impliquant des êtres humains doivent être clairement décrites dans un protocole de recherche. Ce protocole devrait contenir une déclaration sur les enjeux éthiques en question et indiquer comment les principes de la présente Déclaration ont été pris en considération. Le protocole devrait inclure des informations concernant le financement, les promoteurs, les affiliations institutionnelles, d'autres conflits d'intérêts potentiels, les incitations pour les personnes impliquées dans la recherche et les mesures prévues pour soigner et/ou dédommager celles ayant subi un préjudice en raison de leur participation à l'étude.

Le protocole devrait mentionner les dispositions prévues après l'étude afin d'offrir aux personnes impliquées un accès aux interventions identifiées comme bénéfiques dans le cadre de l'étude ou à d'autres soins ou bénéfices appropriés.

15. Le protocole de recherche doit être soumis à un comité d'éthique de la recherche pour évaluation, commentaires, conseils et approbation avant que l'étude ne commence. Ce comité doit être indépendant du chercheur, du promoteur et de toute autre influence indue. Il doit prendre en considération les lois et réglementations du ou des pays où se déroule la recherche, ainsi que les normes et standards internationaux, mais ceux-ci ne doivent pas permettre de restreindre ou exclure l'une des protections garanties par la présente Déclaration aux personnes impliquées dans la recherche. Le comité doit avoir un droit de suivi sur les études en cours. Le chercheur doit fournir au comité des informations sur le suivi, notamment concernant tout

évènement indésirable grave. Aucune modification ne peut être apportée au protocole sans évaluation et approbation par le comité.

16. La recherche médicale impliquant des êtres humains doit être conduite uniquement par des personnes scientifiquement qualifiées et expérimentées. La recherche impliquant des patients ou des volontaires en bonne santé nécessite la supervision d'un médecin ou d'un autre professionnel de santé qualifié et compétent. La responsabilité de protéger les personnes impliquées dans la recherche doit toujours incomber à un médecin ou à un autre professionnel de santé et jamais aux personnes impliquées dans la recherche même si celles-ci ont donné leur consentement.

17. La recherche médicale impliquant une population ou une communauté défavorisée ou vulnérable se justifie uniquement si la recherche répond aux besoins et priorités sanitaires de cette population ou communauté et si, selon toute vraisemblance, les résultats de la recherche seront bénéfiques à cette population ou communauté.

18. Toute recherche médicale impliquant des êtres humains doit préalablement faire l'objet d'une évaluation soigneuse des risques et des inconvénients prévisibles pour les personnes et les communautés impliquées dans la recherche, par rapport aux bénéfices prévisibles pour elles et les autres personnes ou communautés affectées par la pathologie étudiée.

19. Tout essai clinique doit être enregistré dans une banque de données accessible au public avant que ne soit recruté la première personne impliquée dans la recherche.

20. Les médecins ne sont pas autorisés à participer à une recherche impliquant des êtres humains sans avoir la certitude que les risques inhérents ont été correctement évalués et pourront être gérés de manière satisfaisante. Les médecins doivent cesser immédiatement une étude dès que les risques s'avèrent dépasser les bénéfices potentiels ou dès l'instant où des résultats positifs et bénéfiques ont été démontrés.

21. Une recherche médicale impliquant des êtres humains ne peut être conduite que si l'importance de l'objectif dépasse les risques et inconvénients inhérents pour les personnes impliquées dans la recherche.

22. La participation de personnes capables à une recherche médicale doit être un acte volontaire. Bien qu'il puisse être opportun de consulter les membres de la famille ou les responsables de la communauté, aucune personne capable ne peut être impliquée dans une étude sans qu'elle ait donné librement son consentement.

23. Toutes les précautions doivent être prises pour protéger la vie privée et la confidentialité des informations personnelles concernant les personnes impliquées dans la recherche, et pour minimiser l'impact de l'étude sur leur intégrité physique, mentale et sociale.

24. Dans la recherche médicale impliquant des personnes capables, toute personne pouvant potentiellement être impliquée dans la recherche doit être correctement informé des objectifs, des méthodes, des sources de financement, de tout éventuel conflit d'intérêts, des affiliations institutionnelles du chercheur, des bénéfices escomptés et des risques potentiels de l'étude, des désagréments qu'elle peut engendrer et de tout autre aspect pertinent de l'étude. La personne pouvant potentiellement être impliquée dans la recherche doit être informée de son droit de refuser de participer à l'étude ou de s'en retirer à tout moment sans mesure de rétorsion. Une attention particulière devrait être accordée aux besoins d'informations spécifiques de chaque personne pouvant potentiellement être impliquée dans la recherche ainsi qu'aux méthodes adoptées pour fournir les informations. Lorsque le médecin ou une autre personne qualifiée en

la matière a la certitude que la personne concernée a compris les informations, il doit alors solliciter son consentement libre et éclairé, de préférence par écrit. Si le consentement ne peut pas être donné par écrit, le consentement non écrit doit être formellement documenté en présence d'un témoin.

25. Pour la recherche médicale utilisant des tissus ou des données d'origine humaine, les médecins doivent normalement solliciter le consentement pour le prélèvement, l'analyse, le stockage et/ou la réutilisation. Il peut se présenter des situations où il est impraticable, voire impossible d'obtenir le consentement ou que cela mettrait en péril la validité de la recherche. Dans de telles situations, la recherche peut être entreprise uniquement après évaluation et approbation d'un comité d'éthique de la recherche.

26. Lorsqu'il sollicite le consentement éclairé d'une personne pour sa participation à une recherche, le médecin devrait être particulièrement attentif lorsque cette dernière est dans une relation de dépendance avec lui ou pourrait donner son consentement sous la contrainte. Dans ce cas, le consentement éclairé devrait être sollicité par une personne qualifiée en la matière et complètement indépendante de cette relation.

27. Lorsque la recherche implique des personnes incapables, le médecin doit solliciter le consentement éclairé de leur représentant légal. Les personnes incapables ne doivent pas être incluses dans une étude qui n'a aucune chance de leur être bénéfique sauf si cette étude vise à améliorer la santé de la population qu'elles représentent, qu'elle ne peut pas être réalisée avec des personnes capables et qu'elle ne comporte que des risques et des inconvénients minimes.

28. Lorsqu'une personne considérée comme incapable est en mesure de donner son assentiment concernant sa participation à la recherche, le médecin doit solliciter cet assentiment en complément du consentement de son représentant légal. Le refus de la personne pouvant potentiellement être impliquée dans la recherche devrait être respecté.

29. La recherche impliquant des personnes physiquement ou mentalement incapables de donner leur consentement, par exemple des patients inconscients, peut être menée uniquement si l'état physique ou mental empêchant de donner un consentement éclairé est une caractéristique nécessaire de la population sur laquelle porte cette recherche. Dans de telles circonstances, le médecin devrait solliciter le consentement éclairé du représentant légal. En l'absence d'un représentant légal et si la recherche ne peut pas être retardée, l'étude peut être lancée sans le consentement éclairé. Dans ce cas, le protocole de recherche doit mentionner les raisons spécifiques d'impliquer des personnes dont l'état les rend incapables de donner leur consentement éclairé et l'étude doit être approuvée par un comité d'éthique de la recherche. Le consentement pour maintenir la personne concernée dans la recherche devrait, dès que possible, être obtenu de la personne elle-même ou de son représentant légal.

30. Les auteurs, rédacteurs et éditeurs ont tous des obligations éthiques concernant la publication des résultats de recherche. Les auteurs ont le devoir de mettre à la disposition du public les résultats de leurs recherches sur les êtres humains. Ils ont la responsabilité de fournir des rapports complets et précis. Ils devraient se conformer aux directives acceptées en matière d'éthique pour la rédaction de rapports. Les résultats aussi bien négatifs et non concluants que positifs devraient être publiés ou rendus publics par un autre moyen. La publication devrait mentionner les sources de financement, les affiliations institutionnelles et les conflits d'intérêts. Les rapports de recherche non-conformes aux principes de la présente Déclaration ne devraient pas être acceptés pour publication.

C. PRINCIPES ADDITIONNELS POUR LA RECHERCHE MEDICALE ASSOCIEE A DES SOINS MEDICAUX

31. Le médecin peut associer la recherche médicale à des soins médicaux uniquement dans la mesure où la recherche se justifie par sa valeur potentielle en matière de prévention, de diagnostic ou de traitement et si le médecin a de bonnes raisons de penser que la participation à l'étude ne portera pas atteinte à la santé des patients concernés.

32. Les bénéfices, les risques, les inconvénients, ainsi que l'efficacité d'une nouvelle intervention doivent être testés et comparés à ceux de la meilleure intervention courante avérée, sauf dans les circonstances suivantes :

L'utilisation de placebo, ou le fait de ne pas administrer de traitement, est acceptable lorsqu'il n'existe pas d'intervention courante avérée; ou l'utilisation d'un placebo afin de déterminer l'efficacité ou la sécurité d'une intervention est nécessaire pour des raisons de méthodologie incontournables et est nécessaire pour des raisons de méthodologie incontournables et scientifiquement fondées, et les patients recevant le placebo ou aucun traitement ne courent aucun risque de préjudices graves ou irréversibles. Le plus grand soin doit être apporté afin d'éviter tout abus de cette option.

33. A la fin de l'étude, les patients impliqués ont le droit d'être informés des conclusions de l'étude et de profiter de tout bénéfice en résultant, par exemple, d'un accès aux interventions identifiées comme bénéfiques dans le cadre de l'étude ou à d'autres soins ou bénéfices appropriés.

34. Le médecin doit fournir des informations complètes au patient sur la nature des soins liés à la recherche. Le refus d'un patient de participer à une étude ou sa décision de s'en retirer ne doit jamais interférer avec la relation patient-médecin.

35. Dans le cadre du traitement d'un patient, faute d'interventions avérées ou faute d'efficacité de ces interventions, le médecin, après avoir sollicité les conseils d'experts et avec le consentement éclairé du patient ou de son représentant légal, peut recourir à une intervention non avérée si, selon son appréciation professionnelle, elle offre une chance de sauver la vie, rétablir la santé ou alléger les souffrances du patient. Dans toute la mesure du possible, cette intervention devrait faire l'objet d'une recherche pour en évaluer la sécurité et l'efficacité. Dans tous les cas, les nouvelles informations devraient être enregistrées et, le cas échéant, rendues publiques.

© World Medical Association, Inc. - All Rights reserved.

© Asociación médica mundial - Todos los derechos reservados.

© L'Association Médicale Mondiale - Tous droits réservés.