



HAL
open science

Caractérisation des tumeurs et de leur évolution en TEP/TDM au ¹⁸F-FDG pour le suivi thérapeutique

Jacques-Antoine Maisonobe

► **To cite this version:**

Jacques-Antoine Maisonobe. Caractérisation des tumeurs et de leur évolution en TEP/TDM au ¹⁸F-FDG pour le suivi thérapeutique. Autre [cond-mat.other]. Université Paris Sud - Paris XI, 2012. Français. NNT : 2012PA112404 . tel-00844243

HAL Id: tel-00844243

<https://theses.hal.science/tel-00844243>

Submitted on 14 Jul 2013

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

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



Comprendre le monde,
construire l'avenir®

UNIVERSITE PARIS-SUD

ÉCOLE DOCTORALE : STITS

Laboratoire Imagerie et Modélisation en Neurobiologie et Cancérologie

Spécialité : Physique

THÈSE DE DOCTORAT

soutenue le 13/12/2012

par

Jacques-Antoine MAISONOBE

Caractérisation des tumeurs et de leur évolution
en TEP/TDM au ^{18}F -FDG pour le suivi
thérapeutique

Directeur de thèse : Irène BUVAT

Composition du jury :

Rapporteurs : Pr. Jacques DAR COURT
Pr. Damien HUG LO
Examineurs : Dr. Vincent FROUIN
Dr. Camilo GARCIA
Dr. Isabelle GARDIN

Résumé

La Tomographie par Emission de Positons (TEP) au Fluoro-déoxyglucose marqué au Fluor 18 (^{18}F -FDG), analogue du glucose, permet d'obtenir une image de la consommation de glucose dans l'organisme. La plupart des foyers tumoraux présentant une consommation excessive de glucose, son utilisation en oncologie permet d'améliorer la prise en charge des patients en diminuant le temps nécessaire pour évaluer l'efficacité des traitements tels que la chimiothérapie et la radiothérapie. Mon projet de recherche visait à proposer et améliorer des méthodes de quantification en TEP au ^{18}F -FDG afin de caractériser au mieux l'évolution métabolique des volumes tumoraux.

De nombreux facteurs biaisent la quantification en TEP. Parmi eux, l'Effet de Volume Partiel (EVP) reste difficile à corriger, notamment à cause de la faible résolution spatiale des images TEP. Afin de déterminer l'impact de la correction de l'EVP sur l'évaluation des réponses des tumeurs, une étude sur données simulées par Monte Carlo a tout d'abord été effectuée. Cette étude a été complétée par l'analyse de données TEP/TDM (Tomodensitométrie) acquises chez 40 patients atteints de cancers colorectaux métastatiques (CCM), traités par chimiothérapie à l'Institut Jules Bordet (Bruxelles). L'analyse de 101 tumeurs a montré que les critères tels que le SUV, n'incluant pas de correction de l'EVP, et qui reflètent alors le volume tumoral et son activité, prédisaient mieux l'évolution tumorale que les critères corrigés de l'EVP.

Compte tenus des résultats prometteurs récents de méthodes de caractérisation de l'hétérogénéité de la fixation du FDG dans les tumeurs, un second volet de notre travail a consisté à étudier l'intérêt de la prise en compte de la texture dans le cadre du suivi thérapeutique. L'application de l'analyse de texture aux cas de CCM étudiés précédemment n'a pas permis de démontrer une valeur ajoutée des indices de texture par rapport aux index quantitatifs couramment employés en clinique. Nous avons montré que cette conclusion s'expliquait en partie par la non-robustesse des indices de texture vis-à-vis des paramètres impliqués dans leur mesure.

Nous avons enfin cherché à évaluer une méthode d'Analyse Factorielle de Séquences d'Images Médicales (AFSIM), appliquée au contexte du suivi thérapeutique, pour caractériser l'évolution tumorale tout au long du traitement. Cette étude a porté sur 9 séries de 4 à 6 examens TEP/TDM de patients traités par radiothérapie au Centre Hospitalier Universitaire Henri Becquerel de Rouen. Outre l'information visuelle globale apportée par cette méthode, l'analyse quantitative des résultats obtenus a permis de caractériser l'hétérogénéité de la réponse vue par l'AFSIM. L'échec des index classiques, provenant entre autres de leur incapacité à distinguer les processus inflammatoires de l'activité métabolique tumorale, a permis de monter la valeur ajoutée de l'AFSIM par rapport aux index tels que le SUV maximal ou moyen.

Mots-clés : TEP/TDM, ^{18}F -FDG, suivi thérapeutique, quantification, SUV, effet de volume partiel, hétérogénéité, texture, analyse factorielle, AFSIM

Table des matières

RESUME.....	3
TABLE DES MATIERES	5
INTRODUCTION.....	9
CHAPITRE I: LA TOMOGRAPHIE PAR EMISSION DE POSITONS EN CANCEROLOGIE.....	13
I.1. Caractéristiques des cancers	14
I.1.1. Caractéristiques des cancers et particularités tissulaires.....	14
I.1.2. Métabolisme glucidique des cellules normales.....	15
I.1.3. Particularités du métabolisme des cellules cancéreuses : l'effet Warburg.....	16
I.2. L'imagerie TEP	17
I.2.1. Principe et fonctionnement.....	17
I.2.2. Métabolismes mis en évidence par la TEP/TDM en cancérologie	34
I.3. Etude du métabolisme glucidique à l'aide de la fixation du ¹⁸F-FDG	36
I.3.1. Les limites du traceur	36
I.3.2. Analyse qualitative	39
I.3.3. Mesure du métabolisme glucidique	39
I.3.4. Méthodes semi-quantitatives	46
I.3.5. En résumé	50
I.4. Conclusion	50
CHAPITRE II : LE SUIVI THERAPEUTIQUE EN CANCEROLOGIE : CRITERES ANATOMIQUES ET CRITERES TEP.....	53
II.1. Critères anatomiques.....	55
II.1.1. Principales recommandations.....	55
II.1.2. Autres critères : les critères de Choi.....	60
II.1.3. Conclusion sur les critères anatomiques.....	61
II.2. Critères métaboliques en TEP	61
II.2.1. Utilisation de la TEP au FDG en clinique	61
II.2.2. La quantification en TEP au FDG dans le contexte du suivi thérapeutique : difficultés et enjeux..	62
II.2.3. Facteurs affectant la mesure du SUV	62
II.2.4. Besoin de normalisation des pratiques	64
II.3. Recommandations et critères de classifications.....	65
II.3.1. Critères de l'EORTC.....	65
II.3.2. Critères PERCIST	69
II.3.3. Autres méthodes d'analyses des images TEP.....	73
II.4. Conclusion	91

CHAPITRE III : EFFET DE VOLUME PARTIEL : DEFINITION ET METHODES DE CORRECTION..... 93

III.1. Fonction de réponse	94
III.2. Définition de l'effet de volume partiel.....	95
III.3. Conséquences de l'effet de volume partiel	97
III.4. Facteurs influençant l'effet de volume partiel.....	99
III.4.1. Taille et forme de la tumeur	99
III.4.2. Contraste local.....	99
III.4.3. Résolution spatiale de l'image reconstruite.....	100
III.4.4. Echantillonnage spatial	101
III.4.5. Méthode de mesure	101
III.4.6. Conclusion.....	103
III.5. Méthodes de correction de l'EVP	104
III.5.1. Méthodes de corrections appliquées durant la reconstruction.....	104
III.5.2. Méthode de correction appliquée sur les images reconstruites.....	106
III.5.3. En résumé.....	122
III.5.4. Applications des méthodes de correction de volume partiel au suivi thérapeutique	124
III.6. Conclusion	127

CHAPITRE IV : COMPARAISON DES CAPACITES PREDICTIVES D'INDEX METABOLIQUES POUR L'EVALUATION DE LA REPONSE TUMORALE PRECOCE 129

IV.1. Introduction : faut-il corriger le SUV de l'effet de volume partiel pour évaluer la réponse tumorale?	130
IV.2. Etude sur données simulées.....	130
IV.2.1. Matériel et méthodes	130
IV.2.2. Résultats	140
IV.2.3. Discussion	142
IV.2.4. Conclusion.....	143
IV.3. Comparaison d'index métaboliques pour l'évaluation de la réponse précoce au traitement : application à des cas de cancers colorectaux métastatiques traités par chimiothérapie.....	143
IV.3.1. Matériels et méthodes.....	144
IV.3.1. Résultats	150
IV.3.1. Discussion	155
IV.3.2. Conclusion.....	162

CHAPITRE V : CARACTERISATION DE L'HETEROGENEITE MESUREE EN TEP-FDG ET INTERET DANS LE CONTEXTE DU SUIVI THERAPEUTIQUE 163

V.1. Mesure d'hétérogénéité tumorale en TEP : motivations	164
--	------------

V.2. Valeur prédictive et variabilité des index issus de l'analyse de texture : application à des cas de CCM	166
V.2.1. Matériels et méthodes.....	166
V.2.2. Résultats	180
V.2.3. Discussion	193
V.3. Conclusion	198
CHAPITRE VI : EVALUATION D'UNE METHODE D'IMAGERIE PARAMETRIQUE DE LA REPONSE TUMORALE	199
VI.1. Introduction : interprétation de la réponse tumorale sur des séries de plus de 2 examens TEP ..	200
VI.2. L'Analyse Factorielle appliquée au contexte du suivi thérapeutique.....	201
VI.3. Résolution du modèle.....	203
VI.3.1. Etapes préliminaires : recalage et sélection des voxels	204
VI.3.2. Analyse orthogonale.....	206
VI.3.3. Analyse oblique.....	207
VI.3.4. Projection	209
VI.4. Matériel et méthode	209
VI.4.1. Patients et protocole	209
VI.5. Résultats.....	215
VI.5.1. Lésions étudiées	215
VI.5.2. Index métaboliques	216
VI.5.3. Résultats de SULTAN.....	219
VI.5.4. Comparaison entre index métaboliques et résultats fournis par SULTAN.....	220
VI.6. Discussion	229
VI.7. Conclusion	232
CONCLUSION & PERSPECTIVES	235
LISTE DES FIGURES.....	241
LISTE DES TABLEAUX	251
LISTE DES ABREVIATIONS	257
BIBLIOGRAPHIE	263
PUBLICATIONS	285

Introduction

La tomographie par émission de positons (TEP) est une technique d'imagerie médicale non invasive, qui permet de rendre compte de la distribution de la fixation, en 3 dimensions, d'un radiotracteur injecté dans l'organisme. En mettant à profit les particularités du métabolisme des cellules cancéreuses, il est possible, grâce à la TEP, de les distinguer des autres tissus. Si de nombreux traceurs sont étudiés à l'heure actuelle, le plus utilisé en clinique est le 2-deoxy-2-¹⁸F-fluoro-D-glucose (ou FDG), un analogue du glucose, qui permet de visualiser et de quantifier, à l'échelle macroscopique, le métabolisme glucidique exacerbé des lésions cancéreuses. En couplant un imageur TEP à un tomodensitomètre (ou TDM), il est alors possible de localiser et de caractériser l'agressivité et le développement de la maladie chez les patients atteints de cancer. Outre l'intérêt diagnostique de cette technique d'imagerie, la TEP est de plus en plus utilisée pour suivre l'évolution de la maladie, ou évaluer sa réponse à un traitement. En effet, les modifications biologiques des volumes tumoraux, vues par la TEP, précèdent les modifications anatomiques (vues par exemple par la TDM seule), ce qui explique l'impact potentiel de l'imagerie métabolique dans la prise en charge thérapeutique des patients.

En dépit des progrès techniques, la résolution des images TEP reste en deçà de celles des images obtenues par les systèmes d'imageries anatomiques tels que la TDM ou l'Imagerie par Résonance Magnétique (IRM). Le problème du bruit dans les images reconstruites, et surtout leur résolution spatiale limitée, constituent des limites majeures à la quantification en TEP. L'Effet de Volume Partiel, ou EVP, qui est dû à la fois à la faible résolution spatiale du système d'imagerie et à l'échantillonnage des données, est une source de biais d'autant plus critique dans le contexte du suivi thérapeutique qu'il dépend de nombreux paramètres, ce qui rend sa correction complexe.

Les méthodes de quantification utilisées à l'heure actuelle en TEP clinique reposent en grande partie sur la détermination de l'évolution entre deux examens TEP non corrigés de l'EVP, d'un index semi-quantitatif, appelé valeur de fixation standardisée, ou SUV pour *Standardized Uptake Value*. Cet index caractérise l'activité métabolique de la tumeur étudiée. L'évaluation de la réponse précoce de la tumeur au traitement consiste alors à mesurer la valeur des SUV entre un examen de référence et un examen acquis durant la thérapie, et à conclure en comparant l'écart relatif entre ces 2 valeurs à des seuils empiriques. La simplicité d'obtention et d'interprétation de cet index justifie en grande partie son succès actuel. Cette approche, très simple, offre un large champ de réflexion à la recherche de méthodes plus évoluées.

En effet, si de nombreuses méthodes ont été proposées dans la littérature pour corriger l'activité mesurée de l'effet délétère de l'EVP, peu d'études ont étudié l'impact de cette correction sur la classification de la réponse thérapeutique précoce. Le rôle du volume métaboliquement actif des tumeurs (par opposition à l'activité métabolique) reste à mieux comprendre dans le cadre du pronostic et du suivi thérapeutique. De même, les résultats diagnostics récents très prometteurs de la mesure de l'hétérogénéité de la distribution du radiotracteur dans le volume métaboliquement actif des tumeurs restent à valider et à confirmer dans le contexte du suivi thérapeutique. Enfin, que ce soit en utilisant le SUV, le volume ou encore en analysant la texture de la tumeur, l'état de l'art actuel ne propose que très peu de

méthodes pour caractériser l'évolution du métabolisme glucidique à partir de l'information contenue à l'échelle du voxels. Il n'existe pas non plus de méthode qui permette de décrire synthétiquement l'évolution au long cours de la maladie, à partir de plus de deux examens TEP.

Dans ce contexte, **l'objectif de notre travail de thèse était d'étudier les valeurs ajoutées et les limites de différentes méthodes de quantification potentiellement utiles dans le contexte du suivi thérapeutique en TEP au FDG.**

Le premier chapitre du manuscrit rappelle les principes de base de la TEP et présente les différentes méthodes de quantification du métabolisme glucidique décrites à l'heure actuelle.

Le second chapitre expose les différentes méthodes de classification de la réponse thérapeutique basées sur l'imagerie anatomique d'une part, puis sur la TEP d'autre part.

Le chapitre 3 explique l'effet de volume partiel et décrit les différentes méthodes proposées dans la littérature pour le corriger, en mettant tout particulièrement l'accent sur les méthodes applicables dans le contexte de l'oncologie. En particulier, une synthèse des résultats de ces méthodes dans le contexte de l'oncologie est présentée.

Le chapitre 4 présente les résultats de deux études (une sur données simulées, une sur données cliniques) que nous avons réalisées pour mieux comprendre le rôle de l'effet de volume partiel et de sa correction dans le contexte du suivi thérapeutique en oncologie. Ce travail aboutira à des conclusions concernant l'impact de la correction de volume partiel sur l'évaluation de la réponse précoce de cancers colorectaux métastatiques à la chimiothérapie.

En travaillant la même cohorte de patients atteints de cancer colorectaux métastatiques, nous décrivons, dans le chapitre 5, nos résultats concernant la valeur prédictive d'index de texture. Nous discuterons alors leur valeur ajoutée par rapport aux index classiques utilisés en clinique à l'heure actuelle.

Enfin, dans le chapitre 6, nous évaluons une approche d'imagerie paramétrique permettant de décrire la réponse tumorale vue par un certain nombre de TEP successives au moyen d'une seule image paramétrique, tout en fournissant des indices utiles pour la classification de la réponse. Sur des données acquises chez neuf patients atteints de cancer du poumon non à petites cellules et traités par radiothérapie ou radio-chimiothérapie, nous montrerons comment cette méthode met en évidence l'hétérogénéité de la réponse tumorale.

En conclusion, nous dégagerons les perspectives ouvertes par notre travail dans le contexte global de la TEP et de ses évolutions prochaines.

Chapitre I

La Tomographie par Emission de Positons en cancérologie

Les cellules cancéreuses se différencient des cellules normales par plusieurs aspects. La connaissance des mécanismes biologiques liés à ces particularités permet d'améliorer les thérapies mais aussi les méthodes de diagnostic. C'est entre autre la capacité des cellules tumorales à reprogrammer leur métabolisme glucidique qui permet de les distinguer des autres tissus en Tomographie par Emission de Positons (TEP) au 2-deoxy-2-¹⁸F-fluoro-D-glucose (ou FDG). Après avoir présenté brièvement les caractéristiques des tissus cancéreux, nous exposerons donc le principe de la TEP et les différentes étapes qui permettent de passer de la détection du signal à l'obtention d'une image interprétable en routine clinique. Nous détaillerons enfin les différents types d'analyses d'images décrites dans la littérature qui permettent de caractériser le métabolisme glucidique des tumeurs à partir d'une image TEP au FDG.

I.1. Caractéristiques des cancers

I.1.1. Caractéristiques des cancers et particularités tissulaires

Une revue récente de la littérature par (Hanahan and Weinberg, 2011) a relevé que les tumeurs humaines développaient, au cours de leur évolution, 10 caractéristiques complémentaires (résumées dans la figure I.1):

- Une indépendance vis-à-vis des signaux de prolifération.
- Une insensibilité aux signaux antiprolifératifs.
- Le développement d'une résistance à la mort cellulaire.
- Une capacité à induire l'angiogenèse.
- Une capacité à développer des mécanismes d'invasion et de dissémination.
- L'acquisition d'un potentiel illimité de réplication.
- La capacité à ne pas activer le système immunitaire.
- La reprogrammation du métabolisme énergétique.
- L'instabilité génomique et les mutations.
- L'inflammation.

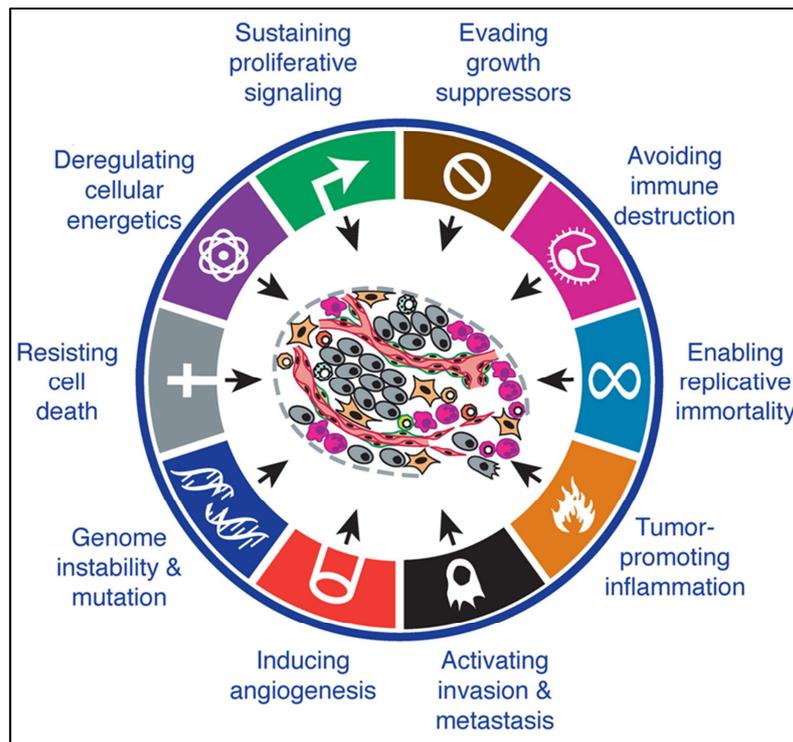


Figure I.1 : Résumé des 10 grandes caractéristiques des tissus tumoraux (Hanahan and Weinberg, 2011).

Ce sont les particularités du métabolisme glucidique, communes à la plupart des tumeurs humaines, qui ont justifié le succès clinique de la Tomographie par Emission de Positons, ou TEP, en oncologie. En effet pour maintenir un taux de prolifération élevé, les cellules tumorales doivent ajuster leur métabolisme énergétique. Pour ce faire, elles mettent en œuvre une manière de consommer le glucose très particulière, appelé *effet Warburg*, d'après le nom du physiologiste allemand, Otto Warburg, qui l'a mis en évidence dans les années 1920.

I.1.2. Métabolisme glucidique des cellules normales

Afin d'assurer leurs fonctions, les cellules non proliférantes des tissus différenciés des mammifères produisent l'énergie dont elles ont besoin sous forme d'Adénosine TriPhosphate (ATP), une coenzyme utilisée par la plupart des processus métaboliques. La production d'ATP se fait grâce à la glycolyse, littéralement la lyse du glucose, qui consiste en une transformation par réactions successives du glucose présent dans la cellule, en glucose-6-phosphate puis en pyruvate, un acide organique (figure I.2).

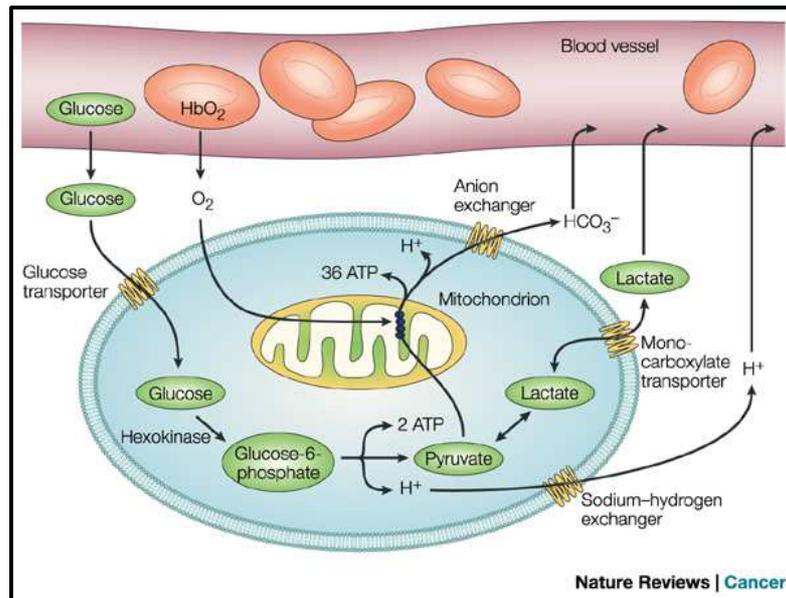


Figure I.2 : métabolisme glucidique des cellules de mammifères. Le sang apporte de l'oxygène (via l'hémoglobine) et du glucose aux cellules constituant les tissus. Le glucose entre dans le cytoplasme par l'intermédiaire de transporteurs spécifiques (famille des GLUT, protéines de la famille des transporteurs de glucose) où il est alors transformé en pyruvate. Ce pyruvate sera alors être soit réduit directement en lactate dans le cas d'absence d'oxygène (respiration anaérobie) pour ne former que 2 ATP / glucose soit être oxydé dans la mitochondrie et former jusqu'à 36 ATP (phosphorylation oxydative). (Schéma extrait de (Gatenby and Gillies, 2004)).

En l'absence d'oxygène, le pyruvate est transformé en acide lactique, ou lactate qui est ensuite excrété par la cellule. Le rendement en ATP de ce processus appelé *glycolyse anaérobie*, est très faible, de l'ordre de 2 ATP par unité de glucose.

En présence d'oxygène, une autre voie, la *phosphorylation oxydative*, beaucoup plus rentable, est favorisée au détriment de la glycolyse anaérobie. Elle correspond à l'oxydation d'une partie du pyruvate en H_2O et CO_2 au cours du cycle des Acides TriCarboxyliques (appelé cycle ATC, ou encore cycle de Krebs (Krebs and Johnson, 1937, 1980)), qui se déroule dans la mitochondrie (figure I.2). Parmi les produits mis en jeu lors de cette transformation, le NADH (ou *nicotinamide adenine dinucleotide* (NAD^+) réduit) produit durant l'ATC est utilisé pour alimenter la phosphorylation oxydative afin de maximiser la production d'ATP et de minimiser la production de lactate. L'oxydation d'une molécule de glucose permet, dans ces conditions, la formation d'un maximum de 36 ATP.

I.1.3. Particularités du métabolisme des cellules cancéreuses : l'effet Warburg

C'est en 1924 qu'Otto Warburg observe que le métabolisme glucidique des cellules tumorales diffère de celui des tissus normaux (Warburg, 1956a, b; Warburg and Posener, 1924). En effet, la plupart des cellules cancéreuses, et plus généralement les tissus en prolifération,

augmentent leur consommation de glucose et, indépendamment de la présence d'oxygène, diminuent leur phosphorylation oxydative tout en produisant une grande quantité de lactate (caractéristique d'une glycolyse anaérobie). Cette *glycolyse aérobie* accrue des cellules tumorale est aussi appelée *effet Warburg*.

L'une des questions qui se pose encore à l'heure actuelle dans le domaine de la biologie des cancers concerne l'intérêt évolutif de cette méthode de consommation du glucose par les cellules tumorales. En effet, la production d'ATP par cette voie est bien moindre que par phosphorylation oxydative, et elle est à peine plus rentable que la glycolyse anaérobie en terme de création d'énergie (~4 ATP produits par glucose consommé).

I.2. L'imagerie TEP

L'amélioration des connaissances relatives aux caractéristiques des cancers bénéficie des progrès faits dans le domaine de l'imagerie, et notamment l'imagerie métabolique. Après avoir présenté le principe de fonctionnement de la TEP, nous verrons en quoi son utilisation se justifie en oncologie.

I.2.1. Principe et fonctionnement

La TEP est une modalité d'imagerie fonctionnelle, par opposition aux systèmes d'imagerie anatomique tels que la tomodensitométrie (TDM) ou l'Imagerie par Résonance Magnétique (IRM) anatomique, par exemple. Elle permet la mesure *in-vivo* du lieu de désintégration d'un radiotracer administré dans l'organisme dans le but de mettre en évidence la fonction ciblée par le radiotracer. Nous allons résumer dans un premier temps les propriétés physiques des radiotraceurs utilisés en TEP et les moyens mis en œuvre pour les détecter. Nous expliciterons ensuite les étapes qui permettent de passer des signaux détectés à une image interprétable visuellement.

I.2.1.1. Bases de physique

L'étude du métabolisme en TEP se base sur l'utilisation de *marqueurs* radioactifs, qui correspondent à des atomes émetteurs de positons. L'objectif est de localiser dans l'organisme les zones d'accumulation d'une molécule vectrice propre à la fonction à étudier : le *vecteur*. C'est cette association *marqueur-vecteur*, encore appelée *traceur*, qui est administrée au sujet avant l'examen TEP.

I.2.1.1.1. Les émetteurs de positons

Lorsqu'un noyau possède un excès de nucléons, protons ou neutrons, il devient instable. Son retour à l'équilibre est permis par une succession d'une ou plusieurs désintégrations radioactives. Lorsque que ces atomes, ou *radionucléides*, présentent un excès de protons, ils se désexcitent de deux manières distinctes (on parle d'*embranchements*) : soit par émission β^+ soit par capture électronique.

- La première de ces désexcitations met en jeu la désintégration élémentaire suivante :



avec p correspondant à un proton, n un neutron, e^+ un positon, antiparticule de l'électron, de masse égale à celle d'un électron mais de charge opposée et ν_e un neutrino, particule élémentaire de charge électrique nulle et de masse négligeable.

Appliquée au *noyau mère* X de numéro atomique Z et de nombre de masse A, encore noté A_ZX , on obtient la formation du *noyau fille* ${}_{Z-1}^AY$ telle que :



Appliquée au fluor 18, noté ${}^{18}\text{F}$, élément instable de nombre de masse A=18 et de numéro atomique Z=9, par exemple, la désintégration s'écrit :



- La seconde, appelée capture électronique, s'écrit :



Soit dans le cas du ${}^{18}\text{F}$



Le *taux d'embranchement*, c'est-à-dire la probabilité pour que le noyau excité suive l'une ou l'autre de ces voies de désintégration, dépend des propriétés du noyau mère. Il est en faveur de l'émission de positon pour les noyaux de faible numéro atomique Z. Ce sont donc ces derniers qui sont considérés lors de l'élaboration de nouveaux traceurs en imagerie TEP. Ainsi, dans le cas du ${}^{18}\text{F}$, le taux d'embranchement est de 97% pour la désintégration β^+ et de 3% pour la capture électronique.

Dans le cadre de la désintégration β^+ , l'énergie libérée durant l'émission du positon est partagée entre le noyau fils, le neutrino et le positon. Les positons produits sont émis avec une

énergie cinétique qui varie de 0 à une énergie maximale E_{\max} de l'ordre du méga-électron volt (ou MeV, avec $1 \text{ eV} = 1,602 \times 10^{-19} \text{ J}$), où E_{\max} est principalement fonction la masse des noyaux mère et fille. L'énergie moyenne des positons ainsi émis par la source β^+ est environ égale au tiers de E_{\max} .

Une autre notion importante relative à l'utilisation des radiotraceurs en imagerie est le temps de demi-vie, propre à chaque élément radioactif et qui correspond au temps nécessaire pour que la moitié des éléments de l'échantillon se désintègre.

Les propriétés des principaux radionucléides utilisés en TEP sont résumées dans le tableau I.1.

Radionucléide	Demi-vie (min)	Energie maximale du positon après annihilation (MeV)	Distance effective moyenne (\bar{d}) parcourue par le positon dans l'eau (mm)
^{11}C	20	0,96	0,92
^{13}N	10	1,20	1,4
^{15}O	2	1,73	2,4
^{18}F	110	0,64	0,54
^{68}Ga	68	1,90	2,8
^{82}Rb	1	3,36	6,1

Tableau I.1 : Principales caractéristiques des radionucléides fréquemment utilisés en imagerie TEP.

I.2.1.1.2. Annihilation

Une fois émis, le positon perd rapidement son énergie cinétique par suite de collisions inélastiques avec les électrons du milieu. Une fois cette énergie dissipée, sur un parcours moyen de l'ordre de $10^{-1} \sim 10^{-2}$ cm, fonction de l'énergie d'émission du positon, il se recombine à un électron, de charge opposée. Très instable, la paire électron/positon ainsi formée, dont la durée de vie moyenne est de l'ordre de 10^{-10} secondes, va alors *s'annihiler* et produire simultanément 2 photons de 511 keV émis dans des directions quasiment opposées. Comme les positons ne sont pas complètement au repos avant cette annihilation, la distribution de l'angle entre les directions des 2 photons correspond à une Gaussienne centrée sur 180° et de largeur à mi-hauteur (LMH) d'environ $0,5^\circ$ (Colombino et al., 1965) suivant la description de (Derenzo et al., 1993).

La figure I.3 (droite) résume l'ensemble des évènements qui se produisent suite à l'émission du positon dans le milieu dans le cas du ^{18}F .

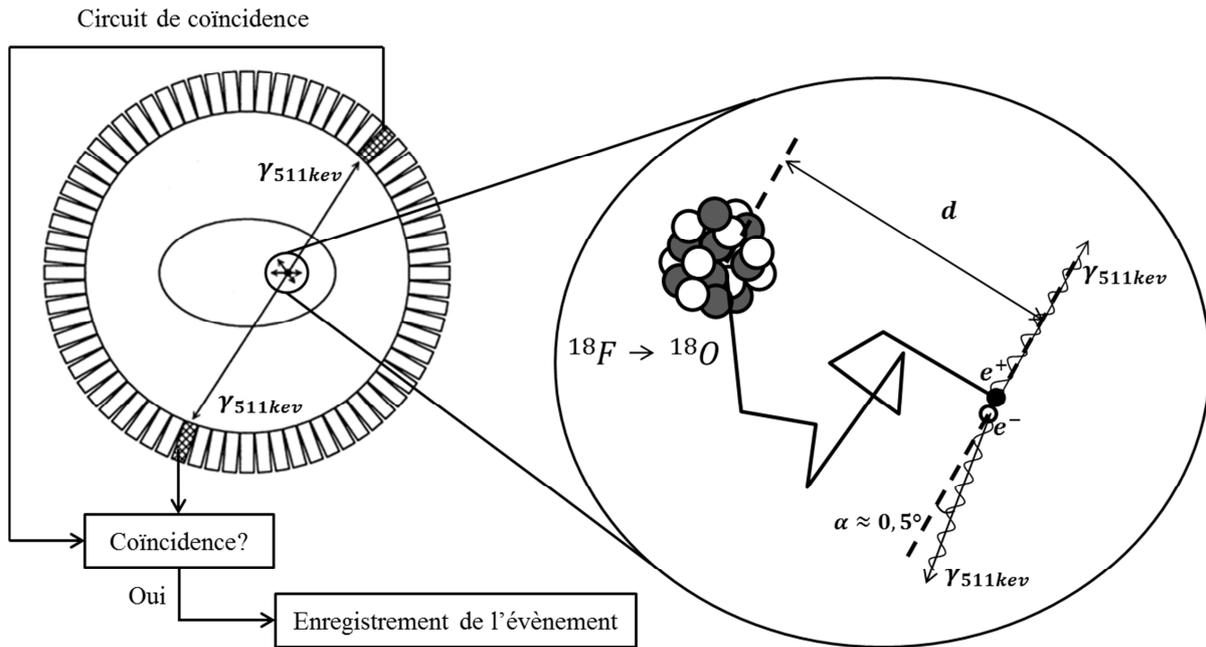


Figure I.3 : Principe de la TEP et de la mesure en coïncidence (gauche) et illustration des événements suivant l'émission du positon lors de la désexcitation d'un atome de ^{18}F (droite). d : distance effective moyenne parcourue par le positon. Elle correspond à la distance entre le noyau mère et le lieu de désintégration du positon, mesurée perpendiculairement à la ligne définie par les directions prises par les deux photons produits suite à l'annihilation.

I.2.1.1.3. Interaction rayonnement matière

Les photons issus de l'annihilation du positon interagissent avec la matière qu'ils traversent, c'est à dire la matière constituant le patient, et celle constituant les détecteurs qui vont être utilisés pour localiser la source émettrice de positon en TEP. Nous allons résumer les principales interactions subies par les photons (tableau I.2).

- Effet photoélectrique

Le photon incident cède toute son énergie à un électron orbital du milieu qu'il traverse. Si l'énergie ainsi transférée excède l'énergie de liaison de l'électron, celui-ci est alors éjecté de sa couche électronique : on parle alors de *photoélectron*. Cet effet implique plus généralement des électrons des couches internes (K ou L). Il est d'autant plus important que l'énergie du photon incident est faible et que le numéro atomique du milieu est élevé.

- Effet Compton

Le photon incident interagit avec un électron des couches les plus périphériques des atomes et seule une partie de son énergie est transmise à l'électron, sous forme d'énergie cinétique.

L'énergie restante est attribuée à un photon secondaire, ou *photon diffusé*, émis dans une direction différente de celle du photon primaire. Aux niveaux des énergies considérées en TEP, cet effet prédomine largement dans les tissus par rapport à l'effet photoélectrique.

- Effet Rayleigh

Contrairement à la diffusion Compton, la diffusion Rayleigh est une diffusion élastique, c'est-à-dire sans transfert d'énergie. Ainsi, le photon incident est dévié par un électron du milieu, mais sans perte d'énergie.

Interaction	Description
Effet photoélectrique	<p>Diagram illustrating the photoelectric effect: an incident photon ($\gamma_{511\text{keV}}$) is absorbed by an atom, and an electron (e^-) is ejected from the atom.</p>
Diffusion Compton	<p>Diagram illustrating Compton scattering: an incident photon ($\gamma_{511\text{keV}}$) interacts with an electron (e^-), resulting in a scattered photon ($\gamma' < 511\text{keV}$) and a recoiling electron (e^-).</p>
Effet Rayleigh	<p>Diagram illustrating Rayleigh scattering: an incident photon ($\gamma_{511\text{keV}}$) is deflected by an electron (e^-) without any energy loss, resulting in a scattered photon ($\gamma_{511\text{keV}}$).</p>

Tableau I.2 : 3 principales interactions entre rayonnement et matière pour des photons de 511keV.

I.2.1.2. Détection des photons

I.2.1.2.1. Types de scintillateurs

La détection des photons émis lors de l'annihilation passe par l'utilisation de détecteurs constitués de matériaux capables de les arrêter. Dans la mesure où le volume de ces détecteurs influe sur la résolution spatiale des images, il est important qu'ils soient les plus compacts

possibles. On favorise donc des matériaux dans lesquels l'effet photoélectrique prédomine aux énergies considérées, c'est-à-dire des matériaux de Z élevé. En TEP, on utilise principalement des *cristaux de scintillation*, dont les atomes produisent des photons de scintillation de quelques eV pour revenir à l'équilibre électronique suite à leur ionisation par les photons incidents. La lumière émise par ces cristaux est collectée par des *photodétecteurs* qui convertissent le signal optique en signal électrique. De nombreux paramètres rentrent en compte dans le choix du cristal utilisé tels que la constante de décroissance lumineuse des photons de scintillation, qui caractérise le taux de comptage du détecteur, ou la résolution en énergie, qui caractérise sa capacité à différencier les photons primaires des photons diffusés avant leur arrivée dans le détecteur. Le tableau I.3 résume certaines propriétés des cristaux les plus couramment utilisés en TEP.

Scintillateur	NaI(Tl)	BGO	GSO	LSO	LYSO
$1/\mu$ (mm)	25,9	11,2	15,0	12,3	12,6
Constante de décroissance (ns)	230	300	60	40	40
Résolution en énergie (en % de 511 keV)	7,8	20	8,9	<9	11
Probabilité d'absorption par effet photoélectrique à 511keV (%)	18	44	26	34	33
Masse volumique (g/cm^3)	3,67	7,13	6,71	7,35	7,19
Z effectif	50	73	58	65	64
Nombre de photons produits (10^3 photons /MeV)	41	9	8	30	30
Susceptibilité magnétique	Non	Non	Oui	Non	Non

Tableau I.3 : Caractéristiques physiques de certains des scintillateurs couramment employés en TEP. Données pour des photons de 511 keV. La définition du coefficient d'atténuation linéique μ est donnée au paragraphe I.2.1.6.

I.2.1.2.2. Bloc de détecteurs

Dans la plupart des tomographes, des *tubes photomultiplicateurs* sont couplés à des matrices de cristaux scintillants, pour amplifier et convertir le signal lumineux en un signal électronique. L'association des photomultiplicateurs aux matrices de cristaux forme un *bloc de détection*, plus compact et moins cher que les détecteurs de première génération (Casey and Nutt, 1986).

Il existe d'autres solutions technologiques que les tubes photomultiplicateurs, basées sur l'utilisation de semi-conducteurs, tels que les *photomultiplicateurs sur silicium* (SiPM) ou encore des photodiodes à avalanches (ou APD pour *Avalanche Photodiodes*), mises en œuvre dans les TEP les plus modernes. A caractéristiques comparables, ces photodétecteurs sont beaucoup plus compacts mais aussi insensibles aux champs magnétiques, d'où leur intérêt

pour la TEP/IRM. Plus de détails sur ces technologies sont fournis par (Lecomte, 2009), par exemple.

I.2.1.2.3. Détection en coïncidence

La localisation des zones de fixation du radiotracer dans le patient en TEP repose sur l'annihilation des positons dans la matière et l'émission en opposition et en simultanée des deux photons mono-énergétiques qui en résulte (511 keV). On utilise ainsi une *détection en coïncidence*, pour identifier le tube d'émission des 2 γ d'annihilations. Ainsi, des détecteurs tout autour de la source émettrice de positons étudiée et en ne prenant en compte que les cas où des paires de détecteurs mesurent quasi-simultanément un photon de 511 keV, on peut supposer que l'annihilation qui a donné naissance aux 2 γ mesurés s'est produite n'importe où dans le tube reliant les deux détecteurs. Cette tube, souvent décrit par une droite reliant les centres des détecteurs, est appelée Ligne De Réponse (LDR). La détection des coïncidences suivant un grand nombre de LDR permet, en utilisant la *reconstruction tomographique* (cf paragraphe I.2.1.5), d'estimer la distribution du radiotracer dans l'organisme.

D'un point de vue technique, la détection en coïncidence nécessite de relier à chaque détecteur un *analyseur d'amplitude multicanal*. Celui-ci permet de générer un signal analogique binaire pendant un temps τ , très bref, lorsque l'énergie d'un photon détecté, alors appelé *single*, est comprise dans un intervalle encadrant les 511 keV. Cet intervalle d'énergie est appelé *fenêtre en énergie*. Si un autre détecteur de l'anneau (figure I.3, gauche) génère lui aussi un signal binaire dans un laps de temps très court défini par une *fenêtre temporelle* de durée égale à 2τ , on considère qu'une coïncidence s'est produite sur la LDR joignant les deux détecteurs. L'évènement est alors enregistré. La longueur de la fenêtre temporelle dépend du type de cristal utilisé dans les détecteurs, mais aussi de l'électronique. Elle est de l'ordre de quelques nanosecondes pour les systèmes ne faisant pas usage de la mesure du temps de vol (cf paragraphe I.2.1.3).

Il est important de noter que le lieu d'émission du positon n'appartient pas à la LDR définie par les photons détectés du fait du parcours du positon avant annihilation. Ceci introduit, en plus de la non-colinéarité des directions d'émission des photons d'annihilation, une limite en résolution intrinsèque de la méthode de détection en coïncidence, que nous décrirons au paragraphe I.2.1.8.

I.2.1.2.4. Types d'évènements mesurés en TEP

Du fait des résolutions finies en énergie et en temps des différents constituants de la chaîne de mesure des coïncidences, certaines d'entre elles ne sont pas des *coïncidences vraies*, c'est-à-dire des coïncidences entre 2 photons issus d'une même annihilation (figure I.4A). Nous présentons les types de coïncidences parasites susceptibles d'être détectées.

- Coïncidences diffusées

Une coïncidence diffusée implique des photons qui ont interagi au moins une fois dans le patient avant d'être mesurés par les détecteurs. Si la perte d'énergie qui résulte de cette/ces interaction(s) n'est pas suffisante pour permettre de les discriminer en énergie par l'analyseur d'amplitude, ils sont enregistrés par le système de coïncidence. Il en résulte une LDR qui ne passe plus par le lieu d'annihilation (figure I.4B) et donc une dégradation de l'image.

- Coïncidences fortuites

Si les photons enregistrés par le système de coïncidence sont issus de deux annihilations distinctes (figure I.4C), on parle de coïncidence fortuite.

- Coïncidences multiples

Une coïncidence multiple est mesurée lorsqu'il y a détection de plus de deux photons dans une même fenêtre de coïncidence (figure I.4D). Il est possible d'ignorer l'ensemble de ces événements ou, si on fait l'hypothèse que l'information contenue dans ces mesures implique au moins une coïncidence vraie, d'annuler selon certains critères (énergétiques par exemple) une ou plusieurs des mesures pour ne garder qu'une LDR.

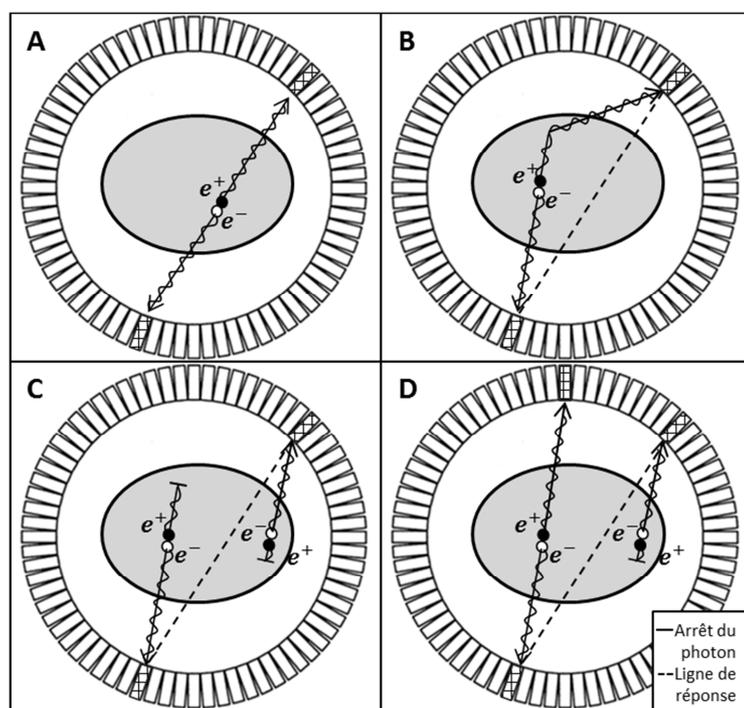


Figure I.4 : Principe d'une coïncidence vraie (A), d'une coïncidence diffusée (B), d'une coïncidence fortuite (C) et d'une coïncidence multiple (D) en TEP.

I.2.1.3. TEP temps de vol

Pour améliorer la localisation du lieu d'annihilation sur une LDR donnée et donc la résolution spatiale des images, les TEP cliniques les plus récentes permettent la mesure du décalage entre les temps d'arrivés des photons gammas sur les deux détecteurs, ou *temps de vol* (*Time of flight, TOF*, équation I.6). Connaissant la vitesse de la lumière (équation I.7), il est en effet possible de déterminer plus précisément le lieu d'annihilation des photons sur la LDR et ce d'autant plus précisément que la résolution temporelle de la TEP utilisée est grande (figure I.5).

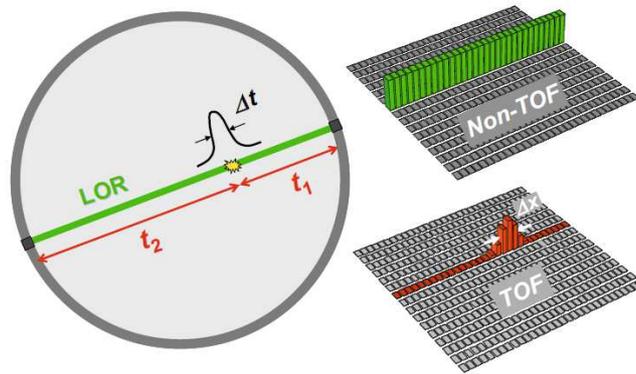


Figure I.5 : principe de la localisation du lieu d'annihilation avec la mesure de temps de vol (Time of Flight, TOF) (source : (Lecomte, 2009)).

avec

$$\Delta t = t_2 - t_1 \quad \text{Eq. I.6}$$

et

$$\Delta x = \frac{c \times \Delta t}{2} \quad \text{Eq. I.7}$$

A titre d'exemple, une TEP temps de vol avec une précision temporelle de 500 ps, correspondant à un ordre de grandeur sur les machines TEP temps de vol actuelles, permet de localiser le lieu d'annihilation du positon avec une précision de 7,5 cm sur la LDR. L'utilisation de la TOF permet de diminuer le bruit statistique et donc d'améliorer le rapport signal sur bruit dans l'image reconstruite (Budinger, 1983; Soussan et al., 2011).

I.2.1.4. Acquisition 2D/3D

I.2.1.4.1. Acquisition 2D

L'acquisition de données en TEP clinique est effectuée à l'aide de plusieurs anneaux de détecteurs coaxiaux, constitués de détecteurs similaires à ceux présentés précédemment. Les premières générations de TEP à anneaux multiples ne permettaient la prise en compte des coïncidences mesurées que dans les détecteurs d'un même anneau, appelés *plan droit*, ou en moyennant des coïncidences mesurées dans des anneaux directement voisins, pour former des *plans croisés* (figure I.6A). La prise en compte des coïncidences issues des plans croisés améliore l'échantillonnage axial, dans la mesure où les événements ainsi mesurés proviennent directement des espaces entre les plans droits. Pour limiter la quantité de photons diffusés et éviter de mesurer des photons d'annihilation issus de plans plus éloignés, la solution technologique la plus fréquemment utilisée consistait à placer des anneaux ou *septa* (en plomb ou tungstène) d'environ 1 mm d'épaisseur entre les anneaux.

Il est possible d'améliorer la sensibilité du système en augmentant la tolérance sur l'éloignement entre les plans croisés (figure I.6B), induisant cependant une diminution de la sensibilité au bord du champ de vue (ou FOV pour *Field Of View*, en anglais), où seuls les plans droits et directement voisins sont pris en compte (figure I.6A).

I.2.1.4.2. Acquisition 3D et stockage des données

Une augmentation de la sensibilité d'un facteur ~ 7 (Cherry et al., 1991) est possible en prenant en compte l'ensemble des combinaisons des coïncidences pouvant avoir lieu entre les différents plans (figure I.6C). Ceci nécessite la suppression des septa, afin de mesurer les lignes de réponses les plus obliques, induisant alors une augmentation du nombre de coïncidences diffusées d'un facteur ~ 3 (Cherry et al., 1991). Comme l'acquisition ne se fait plus seulement dans des plans parallèles ou quasi-parallèles, on parle d'acquisition 3D. L'augmentation de la sensibilité par rapport à des données acquises en 2D permet de diminuer la dose de traceur injectée, d'améliorer le rapport signal sur bruit ou encore de diminuer la durée des examens. L'augmentation de la sensibilité n'est pas uniforme dans la direction axiale et est d'autant plus faible que la mesure est faite au bord du FOV (Fahey, 2002).

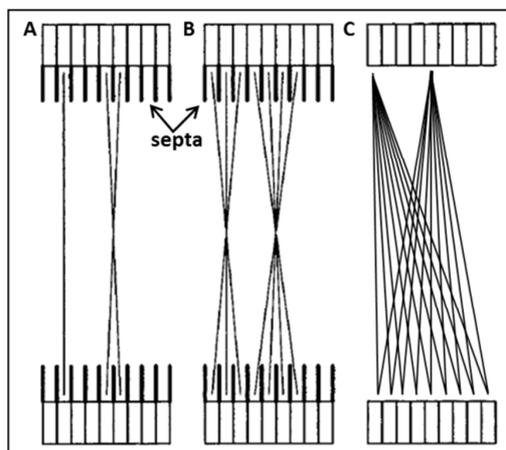


Figure I.6 : Définition des types d'acquisition en TEP : A) acquisition 2D avec plan droit et plan croisé B) 2D haute sensibilité avec plan droit et plan croisé et C) acquisition 3D (extrait et modifié de (Turkington, 2001)).

Les coïncidences détectées peuvent être enregistrées sous 2 formats :

Le format *mode-liste*, ou format séquentiel, consiste à stocker, par ordre chronologique de détection des coïncidences acquises, une liste d'informations telles que l'énergie des photons, la position des cristaux ayant détecté les différents événements et l'instant de cette détection ou encore le temps de vol, quand cette mesure est accessible.

Un autre format de stockage, appelé *sinogramme*, permet de stocker les données sous la forme de projections faites sur 180° autour du patient, avec des échantillonnages spatial, temporel et spectral prédéfinis.

I.2.1.5. Reconstruction des données

L'étape qui consiste à passer des données stockées à la distribution en 3D de l'activité mesurée est appelée *reconstruction*. Deux grands types d'algorithmes de reconstruction sont utilisés : les méthodes analytiques, telle que la rétroprojection filtrée, et les méthodes itératives. Les méthodes analytiques tendant à être de moins en moins utilisées en routine clinique, nous décrirons uniquement, dans la suite, la seconde série de méthodes. Nous renvoyons les lecteurs à (Cherry et al., 2006) pour de plus amples informations sur les méthodes analytiques.

Le principe de base des méthodes itératives consiste à estimer l'image par une suite d'itérations. Une première étape consiste à fournir une estimée grossière de l'image à reconstruire. Le processus itératif consiste alors à projeter les images estimées selon chaque LDR puis à les comparer aux projections mesurées. Les différences obtenues, mesurées par une *fonction de coût*, sont alors utilisées pour ajuster l'estimée initiale le long des LDR. Ce processus est répété jusqu'à ce que les écarts entre projections estimées et mesurées atteignent un certain niveau de tolérance.

Les méthodes itératives présentent des avantages : elles ont été conçues pour des images discrétisées, elles permettent de prendre en compte les propriétés statistiques des données mesurées ainsi que les phénomènes physiques interférant avec l'acquisition ou encore les particularités géométriques du tomographe. La contrepartie de ces algorithmes est un coût de calcul important, chaque itération nécessitant la mise en œuvre d'un grand nombre de projections/rétroprojections. Ces méthodes requièrent l'estimation de la *matrice système*, une matrice de grande dimension (composée de $\sim 10^{12}$ à 10^{16} valeurs) modélisant la probabilité qu'un positon émis soit détecté par coïncidence dans chacune des LDR possibles.

Les algorithmes les plus utilisés en routine clinique reposent sur une approche statistique. On notera en particulier ceux basés sur la maximisation de l'espérance (*Expectation Maximisation*, EM) :

- ML-EM : qui vise à maximiser la vraisemblance (*Maximum Likelihood*, ML) pour minimiser la fonction de coût (Shepp and Vardi, 1982).
- OS-EM (*Ordered Subsets*, OS, (Hudson and Larkin, 1994)) : une version accélérée de ML-EM qui minimise la fonction de coût par sous-ensembles (ou *subsets* en anglais) de projections réparties uniformément autour du volume.

En bref, dans ces 2 méthodes, la *vraisemblance* correspond à une mesure statistique qui est maximale quand les différences entre les projections estimées et mesurées sont minimales. L'algorithme de maximisation de l'*espérance* est un algorithme itératif qui permet de maximiser la vraisemblance lorsque l'on fait l'hypothèse que les données sont affectées d'un bruit de Poisson, ce qui est le cas des données issues des TEP. On notera que, dans le cas de l'algorithme OS-EM, la qualité de l'image finalement obtenue dépend en partie du nombre de sous-ensembles et d'itérations utilisées. En effet, le nombre d'itérations détermine la restitution des hautes fréquences mais aussi le niveau de bruit dans les images. Cette dépendance au nombre d'itérations existe aussi pour tout algorithme itératif.

I.2.1.6. Correction des données

Chacune des étapes permettant l'obtention des images TEP implique des phénomènes pouvant dégrader la qualité de l'image reconstruite. Nous allons présenter certaines méthodes de corrections utilisées pour réduire l'impact de ces phénomènes.

- Normalisation

La mesure en coïncidence repose sur l'utilisation d'un grand nombre de détecteurs. Cependant, la sensibilité de chaque LDR dépend de sa position dans l'espace et de la géométrie des blocs de détecteurs utilisés. De même, les particularités de chaque détecteur telles que leurs dimensions effectives, les propriétés physiques des cristaux de scintillations

ou encore le gain des photomultiplicateurs nécessitent qu'une correction soit appliquée afin de réduire les différences de sensibilité entre chaque LDR. Cette correction est appelée normalisation.

Une première série de méthodes (Defrise et al., 1991), dites directes, consistent à mesurer l'activité d'une source uniforme de géométrie connue. Les facteurs de normalisation appliqués à chaque LDR sont alors égaux au rapport entre le nombre de coïncidences attendues et le nombre de coïncidences effectivement mesurées. Les principaux inconvénients de cette technique sont d'une part la nécessité d'obtenir un taux de comptage élevé pour maximiser la précision de la correction appliquée à chaque LDR, d'autre part la prise en compte de sources de biais inhérentes à la mesure telle que la mesure des coïncidences diffusées.

Une autre méthode (Hoffman et al., 1989) consiste à diviser les facteurs de normalisation en deux composantes : l'efficacité de détection propre à chaque détecteur d'une part et la dépendance de la sensibilité de chaque LDR à l'angle d'incidence des photons sur le cristal, purement géométrique, d'autre part. L'avantage de cette méthode est, qu'une fois estimés, les facteurs de correction ne changent plus dans le temps. Toutefois, il est nécessaire de prendre en compte la dérive de l'électronique en effectuant régulièrement des mesures de sources connues placées au centre du FOV du scanner.

- Correction d'atténuation

Parmi les corrections nécessaires à l'obtention d'images interprétables en TEP, d'un point de vue visuel ou quantitatif, la correction d'atténuation est la plus importante. Comme nous l'avons vu précédemment, à 511 keV, il existe une probabilité non nulle pour que l'un ou les deux photons issus de l'annihilation d'un positon interagissent avec la matière (principalement par effet photoélectrique ou diffusion Compton, cf tableau I.2). Ceci réduit le nombre de coïncidences détectées et induit une sous-estimation de l'activité estimée. Le résultat de cette sous-estimation est illustré en figure I.7.

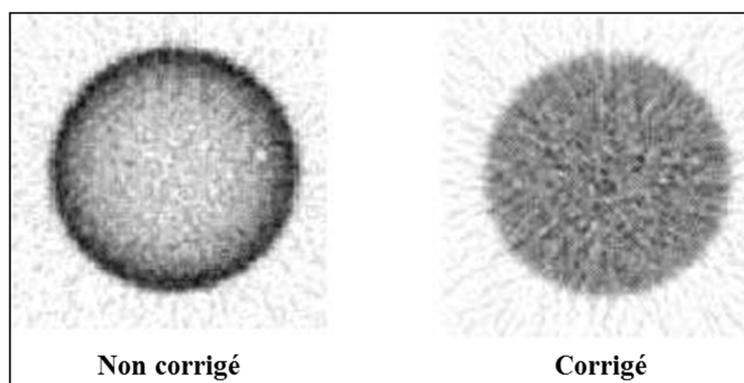


Figure I.7 : effet de la correction d'atténuation sur l'image d'un cylindre d'activité homogène. Extrait et traduit de (Cherry et al., 2006).

En considérant un flux de photons mono-énergétique I_0 émis en un point donné et $I(D)$ le flux de photons du même faisceau après sa traversée d'un milieu d'épaisseur D , on déduit une relation exponentielle entre ces deux flux telle que :

$$I(D) = I_0 e^{-\int_0^D \mu(x) dx} \quad \text{Eq. I.8}$$

où $\mu(x)$, exprimé en cm^{-1} , correspond au coefficient d'atténuation linéique propre au milieu se trouvant à la position x . Il dépend en particulier de l'énergie des photons.

L'atténuation des deux gammas issus d'une annihilation détectée le long d'une LDR ne dépend donc pas du lieu de l'annihilation sur cette même LDR. La probabilité pour que chacun des deux photons soit atténué ne dépend que de l'atténuation totale le long de la LDR (terme intégral de l'équation I.8).

La méthode de correction utilisée en routine clinique repose sur l'utilisation d'une source externe émettrice de rayons X issue d'un TDM, directement associé au système TEP (Beyer et al., 2000). L'image anatomique obtenue sert à déterminer la carte d'atténuation des photons de 511 keV en tout point du milieu. L'utilisation d'un tube à rayons X, dont l'énergie moyenne des photons est de l'ordre de 100 keV, nécessite cependant une conversion de la carte d'atténuation pour estimer l'atténuation des photons à 511 keV. Pour ce faire, la méthode la plus classique consiste à séparer les zones de fortes atténuations des zones de faibles atténuations sur l'image TDM. On applique ensuite, à chacune des deux composantes obtenues, un facteur d'échelle tabulé (Burger et al., 2002; Kinahan et al., 1998). L'application de facteurs d'échelle propres à différentes valeurs d'Unités Hounsfield (UH) permet de prendre en compte les différences de sections efficaces des effets Compton et photoélectrique entre les photons de 100 et 511 keV.

- Correction de la diffusion

La majeure partie des photons diffusés sont détectés avec une énergie inférieure à 511 keV. Il est donc possible de supprimer leur contribution en ne considérant que les photons compris dans une fenêtre d'énergie donnée (cf paragraphe I.2.1.2.3). Cette méthode est cependant insuffisante pour différencier les photons ayant une énergie légèrement inférieure à 511 keV et donc peu déviés durant leur parcours dans le patient.

Les méthodes de correction les plus précises utilisent des simulations numériques pour déterminer la part du rayonnement diffusé à partir de la distribution de l'atténuation (fournie par un TDM par exemple) et d'une estimation de la distribution de radioactivité. La première de ces estimations correspond à l'image reconstruite sans correction du diffusé. La part du diffusé peut alors être estimée en utilisant :

- un modèle analytique simplifié (Ollinger, 1996).
- une simulation Monte-Carlo (Levin et al., 1995).

Le processus est ensuite répété en utilisant comme nouvelle estimée de la distribution l'image ainsi corrigée. L'avantage de ces méthodes est qu'elles prennent en compte l'hétérogénéité de la distribution mesurée ainsi que les données relatives à l'atténuation de l'objet étudié.

- Correction des coïncidences fortuites

Il existe principalement deux méthodes principales pour estimer la distribution des coïncidences fortuites. La première méthode utilise la mesure du taux de comptage des singles pour chaque cristal. Si on note τ le temps d'ouverture d'une fenêtre temporelle émise par un analyseur d'amplitude multicanal suite à la mesure d'un single, alors le nombre de coïncidences fortuites R_{ij} mesurées par chaque LDR joignant les détecteurs i et j vaut :

$$R_{ij} = 2\tau N_i N_j \quad \text{Eq. I.9}$$

Avec N_i et N_j nombres de singles, mesurés par les détecteurs i et j respectivement.

On soustrait alors le nombre de coïncidences fortuites ainsi estimées au nombre de coïncidences totales détectées.

Une seconde méthode consiste à utiliser une seconde fenêtre de coïncidence décalée dans le temps : la *fenêtre de retard* (Williams et al., 1979). Le principe est le suivant : parallèlement à l'enregistrement des coïncidences par la méthode décrite au paragraphe I.2.1.2.3, chaque single détecté par un détecteur i ouvre une fenêtre temporelle de longueur τ , identique à celle utilisée pour le premier circuit de coïncidence, mais avec un retard r . Une fois la fenêtre ouverte, après le temps r , un autre single, issu d'une annihilation différente (si r a été pris grand devant τ), est détecté par un détecteur j et est alors enregistré. La probabilité de détecter des coïncidences fortuites est stationnaire dans le temps. L'ensemble des événements ainsi enregistrés, qui correspondent par définition à des coïncidences fortuites, permet d'estimer, pour chaque LDR, le nombre de coïncidences fortuites moyen enregistré par le système de coïncidences utilisant la fenêtre sans retard.

I.2.1.7. Etalonnage

Si les images obtenues suite aux processus de correction/reconstruction présentés précédemment sont visuellement interprétables, il reste encore une étape afin de pouvoir caractériser l'activité observée en une grandeur physique : *l'étalonnage*. L'étalonnage consiste à passer de la valeur numérique issue de l'image brute, généralement en nombre d'évènements par voxel, à l'activité des tissus par unité de volume, en Becquerel par millilitre (Bq/mL).

L'étalonnage consiste à acquérir, à l'aide du tomographe, la distribution d'une source d'activité homogène connue, généralement un cylindre placé au centre du FOV, et à

reconstruire l'image en appliquant l'ensemble des corrections décrites précédemment. La moyenne du nombre de coups mesurés dans le volume d'intérêt (VOI, *Volume of interest*) correspondant à la zone d'activité connue dans l'image reconstruite est alors utilisé afin de calculer le facteur de calibration par :

$$f_{calibration} = \frac{Activité\ mesurée/Volume}{Nombre\ de\ coups\ dans\ le\ VOI} \quad \text{Eq. I.10}$$

I.2.1.8. Limites de résolution

La résolution spatiale d'un système d'imagerie est la plus petite distance permettant de distinguer deux objets dans l'image. En TEP, elle dépend de plusieurs facteurs :

- Facteurs liés à l'origine du signal
 - Le parcours du positon avant son annihilation avec un électron du milieu. Il dépend de l'isotope utilisé mais aussi du milieu dans lequel se produit la désintégration.
 - La non colinéarité des photons émis suite à l'annihilation du positon. La résolution spatiale du système sera d'autant plus mauvaise que la distance entre deux détecteurs en coïncidence sera élevée, ce qui relie directement la résolution spatiale de la TEP au diamètre de l'anneau de détecteurs utilisé.

- Facteurs liés au tomographe
 - La géométrie des détecteurs joue un rôle important sur la résolution intrinsèque du tomographe. Plus la surface utile des détecteurs est grande, moins la résolution de l'image sera bonne. Ceci est lié à l'impossibilité d'échantillonner les fréquences spatiales f de l'objet étudiée si la fréquence d'échantillonnage utilisée n'est pas au moins supérieure à $2f$ (Shannon, 1949).
 - Les caractéristiques du détecteur, liées par exemple à la qualité du couplage entre cristaux et photomultiplicateurs ou à la nature du matériau constituant le scintillateur. Même si le cristal est choisi pour maximiser la proportion d'effet photoélectrique pour un rayonnement incident de 511 keV, l'effet Compton prédomine à cette énergie (cf paragraphe I.2.1.1.3). Il est donc impossible de connaître avec certitude la localisation exacte de la première interaction entre un photon incident et les électrons du cristal, dont les dimensions dans les 3 directions de l'espace sont de l'ordre de $4 \times 4 \times 30 \text{ mm}^3$.

- L'algorithme de reconstruction.
 - L'algorithme de reconstruction permet de compenser (reconstruction analytique) ou de modéliser directement (dans la matrice système des algorithmes itératifs) des effets physiques et/ou matériels propres à l'imageur

TEP utilisé. Les paramètres de l'algorithme utilisé permettent d'atteindre un compromis entre résolution spatiale et niveau de bruit dans l'image reconstruite pouvant impliquer que l'un soit volontairement dégradé au profit de l'autre (Qi and Leahy, 2006). Pour des raisons de confort d'interprétation visuelle, des *post-filtrages* consistant à convoluer l'image reconstruite par une gaussienne 3D de largeur à mi-hauteur donnée sont aussi souvent utilisés, ce qui dégrade la résolution spatiale dans l'image finale.

La qualité de l'image reconstruite issue d'un tomographe TEP est donc limitée par un grand nombre de facteurs allant du principe de détection des photons d'annihilation à l'algorithme de reconstruction utilisé. Au final, la résolution spatiale dans les images reconstruites interprétées par les médecins nucléaires en TEP à l'heure actuelle est de l'ordre de 5 à 8 mm.

Dans certaines conditions, la faible résolution des images TEP peut biaiser la mesure. Cet effet, appelé *effet de volume partiel* (EVP), ainsi que l'ensemble des méthodes proposées dans la littérature pour corriger de cet effet, seront présentés dans le chapitre III.

I.2.1.9. Imagerie multimodalité

L'imagerie multimodalité a pour but de fournir des images médicales issues de différentes modalités d'imagerie pour un même patient, puis mises en correspondance spatiale et temporelle.

Les modalités ayant le plus profité de l'approche multimodalité sont sans conteste l'imagerie TEP et l'imagerie TEMP (pour Tomographie d'Emission Monophotonique). En effet, la combinaison de la TEP ou de la TEMP à un imageur TDM permet la *fusion* des informations métaboliques et anatomiques en un seul examen. Cela a permis d'une part d'améliorer la localisation des zones fonctionnelles d'intérêt dans l'organisme, et d'autre part de fournir une carte d'atténuation précise, prérequis indispensable à l'obtention d'images TEP ou TEMP quantitatives.

Ce sont les travaux de (Hasegawa et al., 1990), qui ont initié le développement de l'imagerie multimodale en Médecine Nucléaire. L'objectif était de remplacer les sources radioactives, utilisées sur un système TEMP pour obtenir l'image de transmission, par un tube à rayons-X de basse énergie. La concrétisation en un premier prototype TEP/TDM pour le domaine médical s'est produit 10 ans plus tard (Beyer et al., 2000) et a remporté un succès rapide. Les TEP non couplé à un TDM ont maintenant quasiment disparu des hôpitaux à l'heure actuelle.

Si l'importance de la TEP/TDM en oncologie est à présent largement admise, le débat sur l'intérêt de la multimodalité renaît depuis peu avec le développement d'imageurs couplant TEP ou TEMP avec l'IRM. Bien que les preuves de faisabilité de tels systèmes soient contemporaines de celles des imageurs TEP/TDM (eg : (Shao et al., 1997)), leur apparition au

niveau hospitalier ne fait que débiter (Pichler et al., 2010). En bref, les difficultés de tels systèmes résident dans des problèmes technologiques :

- La sensibilité des détecteurs TEP aux champs magnétiques de l'IRM.
- La perturbation des champs magnétiques par les détecteurs TEP.
- La difficulté d'acquérir les données simultanément sans interférence mutuelle.

Ces nouveaux systèmes permettent une amélioration très nette du contraste dans les tissus mous (eg : cerveau), le recoupement entre les informations issues des images IRM fonctionnelles (eg : images de perfusion ou de diffusion) et les images TEP et enfin la réduction de la dose au patient (de l'ordre de ~14 mSv pour un TEP/TDM corps entier, soit 6 mSv de moins que la dose maximale annuelle autorisée pour les personnels exposés en France, la dose délivrée par la TDM représentant 80% de la dose TEP délivrée (Coulot et al., 2011)). Le besoin d'une image d'atténuation des photons de 511 keV pour corriger l'image TEP fait partie des points encore débattus à l'heure actuelle.

I.2.2. Métabolismes mis en évidence par la TEP/TDM en cancérologie

La TEP permet de mettre en évidence l'accumulation du traceur afin d'évaluer l'activité métabolique tumorale. Nous allons décrire les fonctions mises en évidence à l'aide de la TEP en routine clinique ainsi que les principales fonctions étudiées en recherche à l'heure actuelle. Le métabolisme glucidique, mis en évidence par le FDG et faisant l'objet des travaux de recherches présentés ici, sera détaillé plus amplement dans un second temps.

I.2.2.1. En routine clinique

L'utilisation de traceurs en routine clinique est encadrée en France par l'article L5121-8 du code de la santé publique. Il définit l'EMA (Agence Européenne d'Evaluation des Médicaments) à l'échelle européenne, et l'ANSM (l'Agence Nationale de Sécurité du Médicament), à l'échelle nationale, comme étant les seules entités autorisées à délivrer l'Autorisation de Mise sur le Marché (AMM) de nouvelles spécialités pharmaceutiques. L'obtention de cette AMM est un prérequis indispensable à la commercialisation d'un nouveau médicament par une entreprise. Le processus d'obtention de cette autorisation est long et coûteux et nécessite la démonstration rigoureuse de l'innocuité du médicament ainsi que celle de son efficacité, qu'il soit à visée curative ou diagnostique. Il se décompose généralement en 4 phases successives :

- Une phase préclinique, en laboratoire, généralement sur modèle animal, permet de définir les niveaux de doses de référence et la potentielle toxicité du produit.
- La phase I correspond à l'évaluation sur un faible nombre de volontaires (<100), sains ou non, de la tolérance et effets secondaires du produit.
- La phase II, subdivisée en 2 sous phases, a pour but :

- a) d'évaluer l'efficacité de la molécule sur une population limitée de malades (100~200 patients)
- b) de déterminer la dose de médicament adéquate (100~300 patients)
- Le but de la phase III est enfin de démontrer l'efficacité du médicament par rapport au traitement existant, ou à l'absence de traitement quand aucune stratégie thérapeutique n'est définie au moment des tests. Les tailles des cohortes peuvent ici atteindre plusieurs milliers de patients.

Chacune des phases mettant en œuvre des tests sur l'homme est rigoureusement encadrée par des comités d'éthiques (ou comité de protection des personnes) validant la pertinence, la légalité et l'éthique de l'étude.

Traceur	Mécanisme de captation ciblé	Distribution physiologique	Exemple d'indications
18F-Fluorodeoxyglucose (FDG)	Diffusion simplifiée par les transporteurs du glucose dans les cellules tumorales	-Cerveau -Cœur -Graisse brune -Élimination urinaire	Diagnostic: • Cancer primitif ou métastatique Stadification et suivi de la réponse thérapeutique pour des cancers: • Des voies aérodigestives supérieures • Pulmonaires • Colorectaux • Seins • Autres
18F-Fluorodopamine (FDOPA)	Synthèse dopaminergique	-Striatum -Pancréas -Voies biliaires -Duodénum -Élimination urinaire	Détection de récurrences/maladies résiduelles dans le cas de tumeurs: • Cérébrales primitives • Endocrines digestives
18F-Fluorocholeline (FCH)	Transport actif/passif de la choline vers les cellules	-Glandes salivaires -Foie -Pancréas -Rate -Intestin grêle -Élimination urinaire	Détection de récurrences/bilan d'extension: • Cancer de la prostate • Carcinome hépatocellulaire

Tableau I.4 : distribution physiologique et exemples d'indications des 3 traceurs ayant reçu l'AMM en France.

En France, les radiotraceurs TEP ayant obtenu l'AMM dans le cadre du diagnostic sont :

- La 6-¹⁸F-fluoro-L-dihydroxyphénylalanine (FDOPA), analogue d'un acide aminé aromatique naturel transformé par les tissus cibles en dopamine, un neurotransmetteur.
- La ¹⁸F-fluorocholeline (FCH), analogue de la choline, un nutriment essentiel rentrant en jeu dans de nombreux processus cellulaires (eg: synthèse de phospholipides membranaires, de neurotransmetteurs ou signalisation transmembranaire).
- Le 2-deoxy-2-¹⁸F-fluoro-D-glucose (FDG), analogue du glucose, source d'énergie des cellules.

Le tableau I.4 résume certaines des indications de ces radiotraceurs en oncologie ainsi que leurs propriétés principales. La FDOPA est cependant principalement utilisée en neurologie pour mettre en évidence la perte des neurones dopaminergiques au niveau du striatum (eg : maladie de Parkinson).

A l'heure actuelle, le FDG est le traceur le plus employé en clinique (>90% des examens TEP), en dépit, comme nous allons le voir, de son manque de spécificité et notamment de son incapacité à différencier les tissus en prolifération des zones d'inflammations chroniques. S'il permet de diagnostiquer, de déterminer le stade d'avancement de la maladie et d'évaluer la réponse à la thérapie de la plupart des cancers, il apporte cependant peu d'informations en cas de cancers de la prostate, de carcinomes hépatocellulaires et de cancers gastriques, d'où le développement de nouveaux traceurs. Le cas des cancers de la prostate est au cœur de ces recherches compte tenu de leur incidence, ce qui a justifié la mise sur le marché du FCH, cependant peu discriminant de la malignité de ce type de lésion (Schwaiger and Wester, 2011).

I.2.2.2. En phase de recherche

Si on ne considère que les traceurs TEP présentés dans la *Molecular Imaging Probes and Contrast Agent List* (MIP & CA List) fournie par la MICAD (*Molecular Imaging and Contrast Agent Database*), qui répertorie la grande majorité des sondes moléculaires utilisées dans des études sur l'homme ou l'animal à partir de bases de données fiables (eg : PubMed/MedLine), on comptabilise en 2012 plus de 1500 résultats, toutes indications et tous marqueurs confondus. Parmi eux, environ 550 ont été utilisés dans le contexte de l'oncologie. Certains, présentés dans le tableau I.5 par les fonctions qu'ils permettent de mettre en évidence, ont fourni des résultats très encourageants (Juweid and Cheson, 2006; Kumar et al., 2008; Zhu and Shim, 2011).

L'ensemble des travaux rapportés dans le présent document portant sur la quantification en TEP au FDG, nous allons à présent détailler comment il est possible de connaître le métabolisme glucidique des tumeurs à partir de la mesure de la fixation de ce traceur dans les tissus.

I.3. Etude du métabolisme glucidique à l'aide de la fixation du ^{18}F -FDG

I.3.1. Les limites du traceur

L'utilisation du TEP au ^{18}F -FDG fait consensus pour évaluer le métabolisme glucidique des tumeurs. Cependant, il existe, en plus des biais de mesure inhérents à l'obtention d'images TEP quantitatives vus précédemment, de nombreux mécanismes biologiques pouvant altérer l'interprétation de ces examens (Hoekstra et al., 2002). En effet :

- Une partie du FDG capté est utilisé par les macrophages du microenvironnement et non par les cellules tumorales elles-mêmes. Ce point est d'autant plus critique que la thérapie est avancée et que le système immunitaire réagit au traitement.

- La captation de FDG est liée au nombre de cellules cancéreuses, à l'hypoxie ou à d'autres formes de prolifération.
- Les mécanismes de réparation cellulaire peuvent masquer la mort cellulaire.
- Les changements de perfusion de la tumeur peuvent provenir de son évolution propre ou de la réaction au traitement.
- Le FDG, analogue chimique du glucose, diffère suffisamment de ce dernier pour être intégré plus rapidement que lui par les cellules et ne pas être phosphorylé aussi rapidement une fois dans le cytoplasme.

Si ce dernier point peut être pris en compte par certaines méthodes de quantification pour diminuer le biais introduit par l'utilisation de ce radiotracer (cf paragraphe I.3.3.1), les autres limites justifient l'utilisation et le développement d'autres radiotraceurs, plus spécifiques de la présence de cellules tumorales.

Malgré ces limites et la faible résolution des images TEP, il existe de nombreux moyens pour caractériser le métabolisme glucidique. Une vue d'ensemble de ces méthodes, allant de l'analyse qualitative aux différentes méthodes quantitatives proposées dans la littérature et présentées par ordre décroissant de complexité, est proposée et discutée.

Caractéristique étudiée	Mécanisme/voie cellulaire mise en jeu	Analoge biologique du traceur	Traceur en voie de développement correspondant	Exemples d'indications
Prolifération cellulaire	Synthèse de l'ADN	Thymidine	¹⁸ F-thymidine (FLT)	Diagnostic, stadification, suivi du traitement, prédiction de la réponse tumorale
Résistance à la mort programmée (apoptose)	Liaison à la phosphatidylserine membranaire	Annexine V	¹⁸ F-Annexine V	Suivi thérapeutique, localisation des foyers apoptotiques
Angiogenèse	Expression de l'intégrine $\alpha_v\beta_3$	Acide Arginine-glycine-aspartique	¹⁸ F-galacto-c (RGDFK)	Suivi de traitement, étude des mélanomes, glioblastome, cancers du sein, ovaires, prostate
Hypoxie	Diffusion et accumulation dans les cellules hypoxiques	-	¹⁸ F-misonidazole (FMISO)	Amélioration du plan de traitement en radiothérapie, choix de la chimiothérapie
Métabolisme	Synthèse protéique (acides aminés)	Méthionine	¹¹ C-métionine (MET)	Diagnostic, bilan d'extension, planification de traitement (RT), évaluation de traitement
		Acides aminés synthétiques	¹⁸ F-fluoroethyl (FET)	Suivi thérapeutique, gradation

Tableau I.5 : exemple de traceurs étudiés en recherche à l'heure actuelle et applications potentielles.

I.3.2. Analyse qualitative

L'analyse qualitative correspond à l'analyse visuelle d'un examen TEP statique par le praticien. Elle se base sur les différences de contraste entre la lésion et les tissus sains avoisinant, qui dépendent :

- Du temps entre injection et acquisition.
- Du glucose sanguin.
- Du taux d'insuline.
- De la nature des tissus voisins de la lésion.

Ces analyses, nécessaires pour le diagnostic ou les bilans d'extensions, sont cependant considérées comme peu reproductibles et sont déconseillées dans le contexte du suivi thérapeutique pour la plupart des localisations. Dans ce cas, la mesure de paramètres est recommandée. L'analyse qualitative reste cependant privilégiée pour l'étude des lymphomes par exemple (cf paragraphe II.2.1).

I.3.3. Mesure du métabolisme glucidique

Le modèle compartimental utilisé pour déterminer la consommation de glucose à partir du niveau de radioactivité mesuré par suite de l'émission de positons du ^{18}F -FDG chez l'Homme a été adapté par (Phelps et al., 1979; Reivich et al., 1979) du modèle de consommation du 2-deoxy-D-[14C]-glucose (2DG) dans le cerveau de rat (Sokoloff et al., 1977).

I.3.3.1. Analyse cinétique complète et hypothèses sous-jacentes

Cette description complète de la métabolisation du FDG se base généralement sur un modèle à 3 compartiments et 4 constantes d'échanges (figure I.8) permettant de décrire les passages entre :

- Le FDG circulant et le FDG intracellulaire, avec les constantes cinétiques K_1 et k_2 .
- La phosphorylation du FDG en FDG-6- PO_4 par les hexokinases une fois dans la membrane cellulaire (k_3) et la déphosphorylation par la glucose-6-phosphatase (G-6-P) du FDG-6- PO_4 en FDG (k_4 , introduite par (Phelps et al., 1979)).

Le FDG-6- PO_4 ne pouvant pas être métabolisé par la cellule, à la différence du glucose ayant suivi la même voie, il s'accumule dans la cellule, la constante de déphosphorylation étant négligeable devant celle de la phosphorylation du FDG. Les cellules tumorales ayant intégré une plus grande part de FDG sont donc les lieux d'une concentration de ^{18}F plus importante que les cellules saines moins avides en glucose, ce qui justifie l'utilisation de radiotraceurs en TEP comme marqueur de l'activité glucidique.

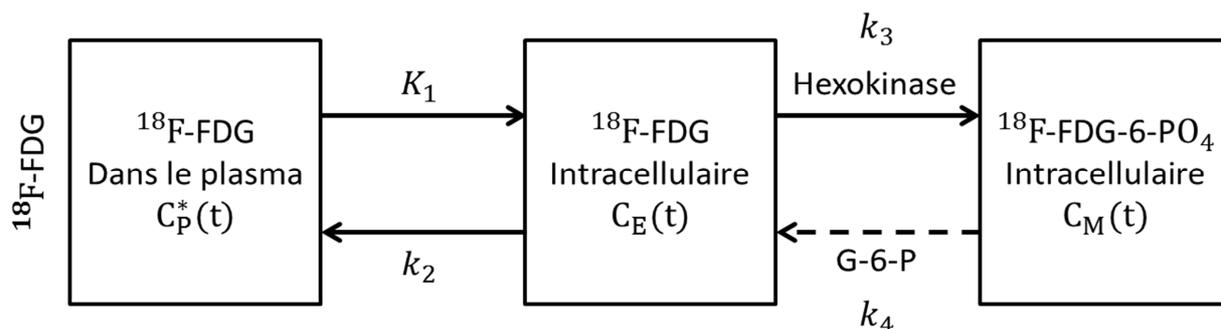


Figure I.8 : modèle compartimental utilisé pour décrire la métabolisation du FDG dans les tissus.

Ce modèle compartimental (Wienhard, 2002) implique plusieurs hypothèses, parmi lesquelles on retiendra que :

- Les tissus contenus dans la région d'intérêt étudiée sont homogènes et les cinétiques qui les caractérisent sont identiques. Autrement dit, les paramètres de flux sanguin, de transport entre plasma et tissu et les taux de phosphorylation du FDG et du glucose par les hexokinases sont identiques dans l'ensemble de l'échantillon.
- Ces taux, ainsi que la concentration de glucose dans le plasma et celles des intermédiaires de la voie glycolytique sont constants durant la mesure. Cela implique que la glycolyse soit à l'état d'équilibre.
- Les concentrations de glucose et de FDG dans le plasma artériel sont à peu près égales à leurs concentrations plasmatiques dans les capillaires.

L'hypothèse relative à l'homogénéité des tissus tumoraux étant rarement vérifiée en oncologie (eg : hétérogénéité de la vascularisation, variation locale du microenvironnement, proportion de tissus nécrotiques), l'estimation des paramètres du modèle est biaisée et induit une surestimation de la métabolisation du glucose dans le cas de la prise en compte de la déphosphorylation du FDG-6-PO₄ (Schmidt et al., 1992; Schmidt et al., 1991). Ces derniers travaux sur données simulées ont d'ailleurs conclu que la constante de déphosphorylation pourrait en fait traduire l'hétérogénéité de la tumeur. Des améliorations du modèle, visant à prendre en compte ces hétérogénéités, ont été proposées (Schmidt et al., 1992; Wu et al., 1995) mais les résultats obtenus par ces méthodes sont contradictoires (eg : (Hoekstra et al., 2002)). Des développements supplémentaires semblent donc requis, d'autant que l'augmentation du nombre de paramètres à déterminer (eg : 6 paramètres dans le cas de (Wu et al., 1995)) pourrait accroître la variabilité des valeurs estimées.

Le modèle à 3 constantes cinétiques permet d'exprimer la concentration de FDG, exprimée en kBq.mL^{-1} , dans le VOI, à chaque instant par :

$$FDG(t) = K_i^* \int_0^t C_p^*(\tau) \cdot d\tau + V_p C_p^*(t) \quad \text{Eq. I.11}$$

avec
$$K_i^* = \frac{K_1 \cdot k_3}{k_2 + k_3} \quad \text{Eq. I.12}$$

où :

$C_p^*(t)$: Concentration plasmatique de FDG à un instant t (en kBq.mL^{-1}).

V_p : Fraction de volume plasmatique dans le VOI (sans unité).

K_i^* : Débit net entrant de FDG dans le VOI, aussi appelée constante de liaison, exprimée en min^{-1} .

Connaissant K_i^* ainsi que la concentration plasmatique de glucose C_{glu} exprimée en mmol.L^{-1} , il est possible d'estimer le taux de métabolisation de glucose dans le VOI, ou constante de transport, exprimée en $\mu\text{mol.mL}^{-1}.\text{min}^{-1}$ et notée MR_{glu} , par la formule :

$$MR_{glu} = \frac{C_{glu} \cdot K_i^*}{LC} \quad \text{Eq. I.13}$$

où LC est une constante (en anglais, *Lumped Constant*), permettant de prendre en compte la différence de transport et de phosphorylation entre FDG et glucose.

Si l'analyse compartimentale est la méthode la plus précise pour analyser les données TEP, elle nécessite cependant des données complexes à acquérir en routine clinique :

- La mesure de la *fonction d'entrée artérielle*. Une alternative à la mesure invasive de la fonction d'entrée artérielle consiste à l'estimer à partir d'une zone artérielle tracée sur les images TEP (eg : courbe rouge en figure I.9B).
- Une acquisition d'images dynamiques TEP à partir du moment de l'injection afin de déterminer la concentration de FDG dans la région d'intérêt à chaque instant (tissue TAC).

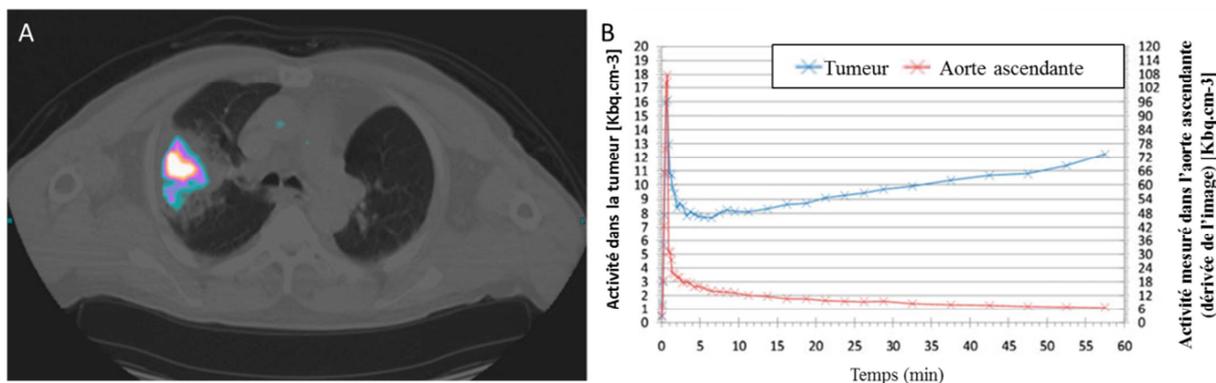


Figure I.9 : Données collectées pour l'analyse compartimentale d'un adénocarcinome pulmonaire (T2) du lobe supérieur droit (A). TAC de la tumeur (bleu) et de l'aorte ascendante (rouge) (B). (Extrait et traduit de (Vriens et al., 2010)).

Les caractères, invasif de la mesure de fonction d'entrée artérielle, et chronophage de l'acquisition dynamique, justifient le développement de méthodes plus adaptées à la clinique. Elles se décomposent en deux groupes : les méthodes cinétiques simplifiées et les méthodes semi-quantitatives.

I.3.3.2. Simplifications à l'analyse cinétique complète

De nombreuses méthodes d'analyse cinétiques simplifiées ont été proposées dans la littérature, dont certaines sont discutées dans les revues de (Buvat, 2007; Hoekstra et al., 2000; Tomasi et al., 2012). Nous présenterons ici en détail la plus utilisée en recherche, l'analyse graphique de Patlak (Patlak et al., 1983), et présenterons brièvement les apports et inconvénients des méthodes moins usitées.

I.3.3.3. Analyses graphiques de Patlak et ses dérivées

L'analyse graphique de Patlak repose sur la linéarisation des équations du modèle compartimental à 3 compartiments et 3 constantes cinétiques (i.e. où la déphosphorylation est négligée : $k_4 = 0 \text{ min}^{-1}$). (Patlak et al., 1983) mettent en effet en évidence qu'après un certain temps T , il existe une relation affine entre :

- Le rapport entre les variations dans le temps des activités du tissu (FDG(t)) et du plasma ($C_p^*(t)$) (terme de gauche de l'équation I.14).
- Le rapport entre l'intégrale de $C_p^*(t)$ entre l'injection et T et sa valeur instantanée (terme de droite de l'équation I.14).

tel qu'en traçant le graphique Y(X) (équation I.14'), la mesure de la pente et de l'ordonnée à l'origine de la droite obtenue permet de déterminer K_i^* et V_p , respectivement. Un exemple de courbe issue de l'analyse de Patlak de la tumeur montrée à la figure I.9A est présentée en figure I.10.

$$\frac{FDG(t)}{C_p^*(t)} = K_i^* \times \frac{\int_0^t C_p^*(\tau) \cdot d\tau}{C_p^*(t)} + V_p \quad \forall t > T \quad \text{Eq. I.14}$$

$$Y(t) = K_i^* \times X(t) + V_p \quad \text{Eq. I.14'}$$

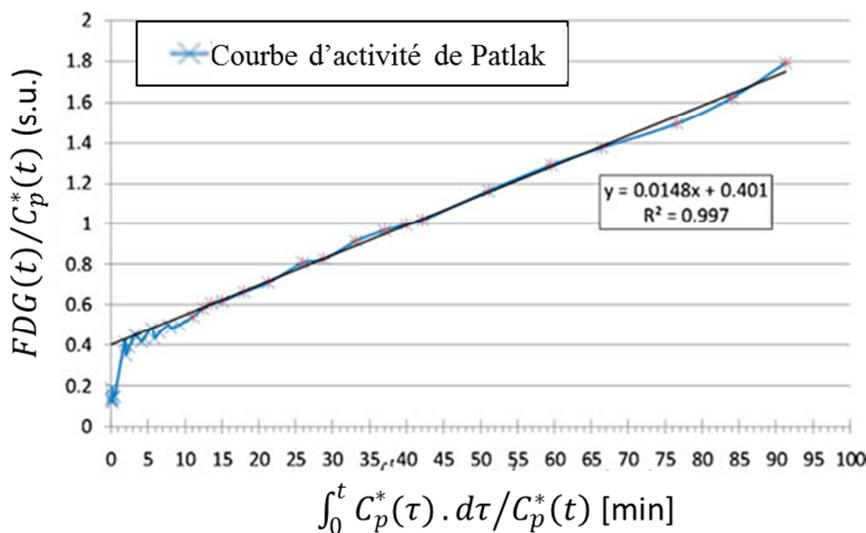


Figure I.10 : courbe obtenue suite à l'analyse de Patlak de la tumeur présentée à la figure I.9. Extrait et traduit de (Vriens et al., 2010).

Cette méthode graphique nécessite donc une acquisition dynamique simplifiée (la cadence d'acquisition des images peut être allégée par rapport à ce qui est nécessaire à l'analyse cinétique complète), débutant 15~30 minutes après injection, et toujours un prélèvement artériel.

L'un des avantages de l'analyse de Patlak provient de la linéarité de phénomène mesuré : elle est en effet beaucoup plus rapide mais aussi moins sensible au bruit que les méthodes d'analyses cinétiques complètes. Ce dernier point permet de l'appliquer à la mesure de la constante de débit entrant net de FDG (K_i^*) à l'échelle du voxel et donc de déterminer la valeur du taux de métabolisation du glucose en tout point de l'image reconstruite. On obtient ainsi une image paramétrique. Cette méthode ne permet cependant pas de distinguer les valeurs des paramètres cinétiques inclus dans K_i^* (cf équation I.12) ni de prendre en compte la déphosphorylation du FDG-6-PO₄.

Beaucoup de publications ont mis en œuvre l'analyse de Patlak, jugée aussi performante que l'analyse compartimentale (eg : (Hoekstra et al., 2002) comparant 34 méthodes d'analyses, incluant différents temps d'acquisition post-injection, sur 19 patients atteints de cancers du poumon ou (Krak et al., 2003) à partir des données de 20 patientes atteintes de cancers du sein). Elle est aussi à la base du développement de nombreuses autres méthodes, comme détaillé dans (Hoekstra et al., 2002). On notera par exemple une méthode simplifiée proposée par (Wong and Hicks, 1994) dont le but est de déterminer les paramètres déterminés par la méthode Patlak sans passer par l'acquisition en continue de la fonction d'entrée artérielle. Si l'écart trouvé sur la constante K_i^* par rapport à la méthode graphique de Patlak est de l'ordre de 3 à 4%, celui portant sur la valeur de V_p est compris entre 7 et 15%, dans une étude portant sur 7 sujets sains. Ce modèle a cependant été très peu utilisé dans la littérature.

Des alternatives au prélèvement artériel ou veineux ont aussi été proposées, se basant sur l'acquisition des dynamiques d'images thoraciques (de l'aorte, du ventricule ou de l'oreillette gauche) et donnant des résultats comparables aux méthodes de références (eg : (Chen et al., 1998)) mais nécessitant des durées d'acquisition longues et l'inclusion de la zone thoracique dans le FOV, non forcément compatibles avec la routine clinique. Ces approches sont de plus limitées par l'EVP et la captation de radiotraceurs dans les tissus proches de la région d'intérêt (Hapdey et al., 2011).

I.3.3.4. Analyse cinétique simplifiée

Pour éviter l'acquisition de données dynamiques, (Hunter et al., 1996) ont décrit une méthode cinétique simplifiée (SKA, *Simplified Kinetic Analysis*), basée sur une estimation (plutôt qu'une mesure) de la fonction d'entrée. Cette méthode requiert l'acquisition d'une image TEP statique de 15 minutes (effectuée 45 à 60 minutes après l'injection de FDG) ainsi qu'un échantillon veineux ~55 minutes post-injection. En faisant l'hypothèse que l'intégrale de la fonction d'entrée, entre le temps d'injection et la réalisation de l'examen TEP, est uniquement fonction:

- De la dose totale de radiotraceur injectée.
- De la concentration de FDG mesurée dans un échantillon sanguin tardif.
- Du poids du patient.

il est possible d'estimer le taux de métabolisation du glucose. La fonction d'entrée du patient est déduite en ajustant les paramètres d'un modèle correspondant à la somme de trois exponentielles (équation I.15).

$$C_p^*(t) = A_1 e^{-b_1 t} + A_2 e^{-b_2 t} + A_3 e^{-b_3 t} \quad \text{Eq. I.15}$$

avec

- $b_i (i \in [1,2,3])$ paramètres supposés communs à tous les patients et obtenus à partir d'une population témoin.
- A_1 et A_2 déterminés à partir de la masse maigre du patient et de la dose de FDG injectée.
- A_3 obtenu en ajustant le modèle mathématique défini pour $C_p^*(t)$ (équation I.15) en utilisant la mesure d'activité de l'échantillon veineux prélevé tardivement.

Si cette analyse est applicable en routine clinique, elle reste néanmoins perfectible de par la non distinction entre FDG métabolisé et non métabolisé dans la tumeur ($V_p = 0\%$). Cette valeur a pourtant été estimée comme pouvant varier de 6 à 67% d'une tumeur à l'autre au moment du TEP statique (Freedman et al., 2003). De plus, l'hypothèse selon laquelle la fonction d'entrée de tous les patients suivrait un modèle mathématique unique (équation I.15) est une hypothèse forte.

Afin d'améliorer cette méthode, (Sundaram et al., 2004) ont tenté de prendre en compte le FDG non métabolisé à partir de l'acquisition de plusieurs images mesurées tardivement, sans mesure de la fonction d'entrée. Cette méthode hybride, à mi-chemin entre l'analyse de Patlak et l'analyse cinétique simplifiée, repose aussi sur l'utilisation de la fonction d'entrée artérielle moyenne d'une population témoin, mise à l'échelle du patient auquel un prélèvement sanguin est effectué, 40 minutes après l'injection du radiotracer. Si les résultats obtenus sont proches de la méthode de référence de (Patlak et al., 1983), cette analyse permet de diminuer la variabilité et les biais observés avec la méthode de (Hunter et al., 1996), probablement du fait de la prise en compte du taux de FDG non métabolisé, paramètre important quand les tumeurs sont le siège d'une forte inflammation ou qu'elles présentent des zones de tissus non viables (Sundaram et al., 2004).

Récemment, une généralisation de SKA (*Extension of the Simplified Kinetic Analysis*, ESKA) a été proposée par (Hapdey et al., 2011) pour optimiser le modèle de fonction d'entrée. Le modèle proposé est dérivé d'un ensemble de fonctions d'entrées issues de patients témoins ayant subi le protocole d'intérêt. Une analyse en composantes principales de cet ensemble de fonctions est réalisée. Seules les 2 premières composantes (notées $C_{p-1}^*(t)$ et C_{p-2}^*) de cette analyse en composante principale, associées aux 2 plus grandes valeurs propres (i.e. les plus représentatives de la décomposition), sont retenues pour décrire la fonction d'entrée $C_p^*(t)$ de chaque patient. Le modèle à ajuster est donc de la forme :

$$C_p^*(t) = A_1 C_{p-1}^* + A_2 C_{p-2}^* \quad \text{Eq. I.16}$$

où A_1 et A_2 sont deux constantes permettant d'estimer la fonction d'entrée propre à chaque patient. Elles sont déterminées grâce à deux prélèvements sanguins effectués peu de temps après l'injection et à la fin de l'examen, respectivement.

Cette méthode a conduit à estimer des K_i^* d'autant plus proches des résultats fournis par l'analyse de Patlak et d'autant plus éloignés des résultats de la méthode SKA originale que le protocole d'injection étudié s'éloignait de celui proposé par (Hunter et al., 1996).

I.3.4. Méthodes semi-quantitatives

Les méthodes dérivées de l'analyse compartimentale permettent de déterminer des paramètres quantitatifs caractérisant les tumeurs. Cependant, les moyens à mettre en œuvre pour obtenir les données nécessaires à l'utilisation de ces méthodes sont souvent incompatibles avec la pratique clinique. Le compromis entre simplicité et estimation de grandeurs quantitatives a largement favorisé le développement et l'utilisation de méthodes semi-quantitatives. C'est la méthode de calcul du SUV (*Standardized Uptake Value*), que nous allons présenter, qui en a le plus bénéficié. Cette mesure utilise une unique TEP statique, généralement acquise une heure après l'injection du radiotracer, et ne nécessite pas plus d'informations que la dose injectée et le poids ou la taille du patient.

I.3.4.1. Valeur de fixation normalisée : le SUV

La valeur de fixation normalisée a été proposée dès 1941 par (Kenney et al., 1941) sous le nom de DAR (*Differential Absorption Ratio*) et a trouvé ses premières applications en TEP dans les années 80 (Kubota et al., 1985). Malgré de nombreux débats relatifs à son utilisation en oncologie (Keyes, 1995; Thie, 2004), il est à présent l'index quantitatif le plus utilisé en clinique (eg : (Beyer et al., 2011)). Pour cette raison, il fait régulièrement l'objet de recommandations dans des protocoles de normalisation, dont l'objet est de limiter les sources de variabilité entre les centres en uniformisant les pratiques (eg : (Wahl et al., 2009)). Nous présenterons ici les hypothèses simplificatrices qui permettent de passer de la mesure du métabolisme glucidique telle que décrite par l'analyse compartimentale à celle du SUV.

I.3.4.1.1. Définition du SUV

La fixation de traceur dans les tissus dépend en particulier de la dose injectée et du volume sanguin dans lequel l'activité se distribue. Une manière simple de normaliser la fixation mesurée consiste donc à prendre en compte ces deux facteurs, tels que :

$$SUV_{BW}(g \cdot ml^{-1}) = \frac{\text{Concentration d'activité } (\frac{kBq}{mL})}{\text{Dose injectée } (kBq) / \text{Poids du patient } (g)} \quad \text{Eq. I.17}$$

Cette définition du SUV normalisé par le poids (ou *Body Weight*, BW, en anglais) est généralement donnée sans dimension. Dans ce cas, on fait donc l'hypothèse implicite que la masse volumique du patient est de $1 \text{ g}\cdot\text{mL}^{-1}$ (masse volumique de l'eau).

Comme nous le verrons dans le chapitre II, il existe, en dehors du choix de la normalisation (i.e. du dénominateur), autant de moyen de définir le SUV d'une lésion que de manière de choisir le VOI représentative de la concentration d'activité de la tumeur (i.e. du numérateur). A titre d'exemple ici, on notera:

- L'utilisation croissante (Beyer et al., 2011; Wahl et al., 2009), en clinique ou en recherche, du SUV_{max} . Dans ce cas, le VOI utilisé pour calculer le SUV est le voxel de la tumeur ayant la concentration d'activité la plus élevée.
- L'utilisation du SUV_{peak} , dont le VOI (ou VOI_{peak}) utilisé pour calculer le numérateur est un volume fixe dans la tumeur. La concentration d'activité utilisée est alors la moyenne de la concentration d'activité des voxels contenus dans ce VOI. La définition exacte du VOI_{peak} doit encore faire consensus. Récemment, l'impact de cette définition sur la classification de la réponse tumorale a été étudiée (Vanderhoek et al., 2012), en considérant 24 définitions possibles du SUV_{peak} . Pour cela, les auteurs ont analysé des images TEP au ^{18}F -FLT de 17 patients, acquises avant et pendant le traitement. Les résultats montrent que la réponse intra tumorale vue par cet index dépend plus de la taille du VOI_{peak} considéré que de sa localisation dans la tumeur (centré sur le SUV_{max} ou placé dans la zone d'activité la plus élevée) ou de sa forme (circulaire ou sphérique). Nous reviendrons en détail sur les conclusions de cet article au paragraphe III.4.5.

I.3.4.1.2. Hypothèses simplificatrices et écart au modèle cinétique

Comme rappelé dans (Buvat, 2007), la mesure du métabolisme glucidique par le SUV_{BW} se base sur des simplifications du modèle compartimental, à savoir :

- L'évolution temporelle de la fixation de FDG dans la tumeur est négligée. On suppose de plus qu'au moment de réaliser l'examen TEP, la quantité de fixation de FDG dans la lésion est égale à la quantité de FDG piégé dans la tumeur. Or la fixation de FDG dans la tumeur continue généralement de croître au moment de l'acquisition des données. Par conséquent, la mesure dépend du temps qui s'est écoulé entre l'injection du radiotracteur et le début de l'acquisition.
- Comme pour la méthode SKA (Hunter et al., 1996), on considère que toute l'activité mesurée dans la tumeur correspond à du FDG métabolisé. Autrement dit, la fraction de FDG circulant est nulle. Cette hypothèse peut induire d'importants écarts entre SUV et métabolisme du FDG mesuré par des analyses de référence comme l'analyse de Patlak (Freedman et al., 2003; McDermott et al., 2007; Sundaram et al., 2004).

- La non distinction entre la métabolisation du glucose et celle du FDG. Dans le calcul du SUV, on fait l'hypothèse implicite que la constante localisée (LC, cf équation I.15) utilisée par le modèle compartimental pour rendre compte de
 - l'intégration plus rapide du FDG dans le cytoplasme d'une part,
 - de sa plus lente déphosphorylation une fois dans la cellule d'autre part,
 est identique dans toute la tumeur, invariante dans le temps, et égale à 1, ce qui n'est pas justifié d'après les études sur tissus cérébraux pathologiques (Wienhard, 2002).
- Qu'en utilisant la normalisation par le poids du patient (cf équation I.17), on considère que tous les tissus non tumoraux consomment une quantité totale de FDG proportionnelle au poids du patient. Comme relevé par (Keyes, 1995), cette hypothèse n'est pas vérifiée dans le cas des tissus adipeux, dont la fixation de FDG est inférieure à celle des autres tissus. Ce problème est majeur dans le contexte du suivi thérapeutique, où la modification de la masse grasseuse d'un patient durant le traitement peut engendrer des interprétations des variations du SUV_{BW} contradictoires.

Pour remédier à ce dernier problème et réduire la dépendance du SUV à la masse grasseuse du patient, d'autres définitions du volume de distribution ont été proposées. Les plus communément employées sont :

- La surface corporelle totale (BSA, *Body Surface Area*), exprimée en m^2 . Elle est couramment calculée à partir de formules empiriques faisant intervenir la taille et le poids du patient (eg : (Haycock et al., 1978), (Mosteller, 1987)). La formule préconisée par l'*European Organisation for Research and Treatment of Cancer* (EORTC), qui a décrit l'un des 2 principaux protocoles TEP pour le suivi du traitement en oncologie (cf paragraphe II.3.1) est la formule proposée par (Du Bois and Du Bois, 1916) :

$$BSA(m^2) = 0,00718 \times poids(kg)^{0,425} \times taille(cm)^{0,725} \quad \text{Eq. I.18}$$

- La masse maigre (LBM, *Lean Body Mass*, proposée par (Zasadny and Wahl, 1993)). Elle est également obtenue par des formules empiriques dépendant de la taille, du poids mais aussi du sexe et de l'âge du patient (cf équations I.19 et I.20).

$$\text{Femme :} \quad LBM(kg) = 1,07 \times poids(kg) - 148 \times \left(\frac{poids(kg)}{taille(cm)} \right)^2 \quad \text{Eq. I.19}$$

$$\text{Homme :} \quad LBM(kg) = 1,10 \times poids(kg) - 120 \times \left(\frac{poids(kg)}{taille(cm)} \right)^2 \quad \text{Eq. I.20}$$

Si ces deux normalisations permettent de réduire la corrélation entre le SUV et le poids du patient (Kim et al., 1994; Zasadny and Wahl, 1993), il a été montré, sur une population de 153 patientes, que les différentes formules proposées pour calculer la masse maigre induisait une variabilité importante des SUV ainsi normalisés (Erselcan et al., 2002). Pour remédier à ce problème, (Chan, 2012) a proposé une méthode automatique permettant de déterminer la masse maigre du patient à partir d'un TDM non-corps entier de faible dose. La comparaison de ces deux méthodes de normalisation a fourni des résultats contradictoires, soit en faveur du SUV_{BSA} ((Graham et al., 2000; Izquierdo-Garcia et al., 2009), soit en faveur du SUV_{LBM} (Kim and Gupta, 1996; Krak et al., 2003).

Enfin, comme le glucose est en compétition avec le FDG, il est généralement demandé au patient de se présenter à jeun au moment de l'examen. En effet, une glycémie élevée peut induire une plus faible fixation du radiotracer et donc une sous-estimation du SUV (Lindholm et al., 1993). Des études évaluant l'intérêt de la normalisation du SUV par la valeur du glucose sanguin n'ont pas fourni de conclusion univoque. Certains auteurs ont ainsi conclu sur son intérêt (eg : (Wong et al., 2005) pour la classification de lymphomes) et d'autres sur l'absence de bénéfice (eg : (Stahl et al., 2004) pour le suivi thérapeutique). Certains auteurs ont conclu que cette correction diminue la reproductibilité de la mesure du SUV (Hadi et al., 2008). (Cheebsumon et al., 2011b) ont enfin souligné le besoin d'utiliser des méthodes de mesure du glucose sanguin précises dans le contexte du suivi thérapeutique pour diminuer la variabilité test-retest¹ du SUV corrigé du niveau de glucose sanguin.

La simplicité d'obtention du SUV explique qu'il soit le seul paramètre largement adopté en routine clinique. Afin de valider sa fiabilité, de nombreuses études ont donc comparé cette estimée du métabolisme glucidique à celles fournies par les méthodes plus complexes. Les résultats obtenus mettent en évidence une bonne corrélation entre les valeurs prises par cet index et celles obtenues par des analyses plus complexes dans de nombreuses localisations étudiées (eg : tumeurs osseuses : (Brenner et al., 2004), cancers du poumon : (Cheebsumon et al., 2011b; de Geus-Oei et al., 2007), cancers du rein (Freedman et al., 2003), du sein (Krak et al., 2003), gastro-œsophagien (Kroep et al., 2003)). Certains de ces auteurs ont cependant relevé la faible capacité du SUV à reproduire les variations du métabolisme glucidique vues par les méthodes de référence, dans le contexte du suivi thérapeutique. La reproductibilité du SUV a elle aussi été largement étudiée à travers des études de type test-retest. Une méta-analyse récente de ces travaux (de Langen et al., 2012) a permis de montrer que la répétabilité du SUV était de l'ordre de ~30%. La reproductibilité s'est en outre avérée meilleure pour un SUV dont la concentration d'activité (numérateur de l'équation I.17) correspond à la moyenne de plusieurs voxels de la tumeur que pour le SUV_{max}. La répétabilité du SUV étant d'autant plus mauvaise que l'activité mesurée dans la lésion étudiée était faible, les auteurs ont enfin proposé de déterminer, en plus de la variation relative du SUV, sa variation absolue, et de comparer aussi la variation absolue à un seuil, de l'ordre de 1 g/mL. La combinaison de ces deux seuils pourrait permettre de s'assurer, dans le contexte du suivi thérapeutique, que toute

¹ Variabilité test-retest : test qui consiste à évaluer la capacité d'une méthode de mesure à reproduire un résultat à partir de deux mesures effectuées dans des conditions identiques. Plus cette variabilité est faible, plus la méthode étudiée est capable de reproduire un résultat quand aucun paramètre d'entrée ne change.

modification de la valeur du SUV entre deux examens est représentative d'une réponse physiologique de la tumeur et n'est pas uniquement due à des fluctuations statistiques. Ce point sera développé plus complètement dans le chapitre II, où l'influence de chacun des facteurs affectant la mesure du SUV sera détaillée.

I.3.4.2. Rapport tissus tumoraux / tissus sains

Le rapport entre l'activité moyenne dans la tumeur et dans les tissus sains, ou rapport T/N, est l'index semi-quantitatif le plus simple d'accès, mais souffre de plusieurs biais. Se basant sur les mêmes données que celles requises pour l'obtention du SUV (ie : d'un TEP statique acquis ~60 minutes après l'injection du radiotracteur), il possède les mêmes sources de biais que ce dernier. De plus, si la définition de la zone tumorale est un problème majeur en TEP, celui de la définition de tissu sain est elle aussi complexe. En effet, l'activité des tissus sains peut évoluer du fait du traitement, et ce même si la réponse tumorale est négative. De même, il n'est pas toujours aisé de définir une zone de référence, comme dans le cas de cancers métastatiques où la zone de tissu sain préconisée par les protocoles peut elle-même contenir des foyers tumoraux. Enfin, comme l'activité dans les tissus sains est généralement faible, une erreur sur la valeur de l'activité dans ces zones peut induire une erreur statistique importante dans celle du rapport. Pour ces raisons, le SUV est donc bien plus utilisé que le rapport T/N.

I.3.5. En résumé

La littérature propose de nombreuses méthodes pour caractériser, de manière plus ou moins précise, le métabolisme glucidique des tumeurs à partir d'examens TEP. Dans la pratique, les impératifs cliniques ont largement favorisé les index semi-quantitatifs, et surtout la valeur de fixation normalisée, ou SUV, pour sa simplicité d'obtention. Cependant, les nombreuses sources de biais affectant cette mesure et les hypothèses simplificatrices qui la lient aux grandeurs quantitatives d'intérêts (eg : K_i^* , MR_{glu}) impliquent de prendre de nombreuses précautions avant toute interprétation. Le tableau I.6 résume l'ensemble des avantages et inconvénients des différentes méthodes présentées précédemment.

I.4. Conclusion

La TEP est utile pour la caractérisation du métabolisme des maladies néoplasiques. Bien que peu de traceurs aient obtenues l'AMM à l'heure actuelle, l'effort fait en recherche sur les radiotraceurs innovants laisse penser qu'il deviendra possible de marquer séparément différentes caractéristiques des cancers. Actuellement, l'étude du métabolisme glucidique exacerbé des cellules tumorales, permise par l'emploi en routine clinique du FDG, revêt un intérêt particulier. En effet, si une partie des difficultés rencontrées pour quantifier les grandeurs biologiques relatives à l'emploi de ce radiotracteur sont propres aux mécanismes cellulaires mis en jeu (eg : distinction entre les captations concomitantes de FDG par les

cellules inflammatoire et tumorales), d'autres sont liées au principe même de la mesure en TEP (eg : problème de mouvement respiratoire, médiocre résolution spatiale des images reconstruites). La plupart des méthodes de quantification présentées dans ce chapitre sont transposables à d'autres types d'images TEP, i.e. que ce soit à des images TEP/IRM ou TEP/TDM temps de vol, au FDG ou non. En faisant le choix explicite de ne s'intéresser qu'à une classe d'index semi-quantitatif, dont les limites sont connues et largement décrites dans la littérature (Keyes, 1995), nous admettons donc *de facto* des biais, inhérents aux hypothèses sous-jacentes à leur définition. Nous expliciterons, dans la prochaine partie, les conséquences de ce choix sur la mesure et détaillerons les stratégies proposées dans la littérature pour en atténuer les principaux effets.

Méthode	Avantages	Inconvénients
Analyse cinétique complète	<ul style="list-style-type: none"> -Méthode de référence -Analyse quantitative -Connaissance des paramètres internes au modèle 	<ul style="list-style-type: none"> -Mesure nécessaire de la fonction d'entrée -Implémentation complexe -Durée d'examen longue (examen dynamique) -Besoin d'un échantillon sanguin
Méthode graphique de Patlak	<ul style="list-style-type: none"> -Analyse quantitative -Souvent considérée comme excellente alternative à l'analyse cinétique complète -Implémentation moins complexe que l'analyse compartimentale -Possibilité d'une analyse par pixel 	<ul style="list-style-type: none"> -Pas d'accès aux paramètres individuels du modèle compartimental -Non prise en compte de la déphosphorylation -Nécessité de mesurer la fonction d'entrée -Durée d'examen longue
Analyse cinétique simplifiée (SKA)	<ul style="list-style-type: none"> -Utilisation d'un examen statique -Echantillon de sang veineux 	<ul style="list-style-type: none"> -Hypothèses sous jacentes du modèle de fonction d'entrée à ajuster pas forcément adaptée à tous les patients (solution ESKA)
SUV	<ul style="list-style-type: none"> -Besoin d'un seul examen TEP statique -Pas de nécessité de prélèvement sanguin -Semi-quantitatif -Utilisation quasi-généralisée en routine clinique -Facile à mesurer 	<ul style="list-style-type: none"> -Non prise en compte de la différence entre FDG circulant et non circulant -Pas de distinction possible entre les cinétiques du glucose et du FDG -Mesure dépendante du temps entre acquisition et injection -Dépendant de la glycémie du patient -Dépendance de la valeur au poids du patient -Qualité de la mesure très dépendante de la qualité de l'image et des paramètres utilisés pour
Rapport T/N	<ul style="list-style-type: none"> -Besoin d'un seul examen TEP statique -Pas de nécessité de prélèvement sanguin -Semi-quantitatif -Facile à mesurer 	<ul style="list-style-type: none"> -Mesure dépendante du temps entre acquisition et injection -Définition subjective des tissus sains -Réponse des tissus sains au traitement -Possède tous les inconvénients du SUV

Tableau I.6 : résumé des principales méthodes de détermination du métabolisme glucidique à partir de la fixation du ¹⁸F-FDG mesurée dans le tissu d'intérêt, en TEP. Corrige les fautes.

Chapitre II

Le suivi thérapeutique en cancérologie :

Critères anatomiques et critères TEP

Si l'utilisation de la TEP/TDM au FDG est largement admise à l'heure actuelle en oncologie, c'est d'abord parce qu'elle permet d'obtenir des informations sur l'agressivité des lésions tumorales, leurs localisations vis-à-vis des structures anatomiques et leur extension dans l'organisme. Sachant que l'imagerie métabolique a le potentiel de mettre en évidence des modifications cellulaires des néoplasmes, induites par le traitement, plus précocement que les modifications anatomiques, cette modalité d'imagerie est aussi utilisée pour évaluer la réponse thérapeutique (Ben-Haim and Ell, 2009). La TEP/TDM est également utilisée pour l'évaluation de nouvelles stratégies de traitement et pour l'aide à la planification de traitement en radiothérapie (Bentzen, 2005; Grégoire et al., 2007). L'utilisation de la TEP pour l'évaluation de la réponse thérapeutique nécessite une méthodologie rigoureuse, que ce soit pour la réalisation des examens, mais aussi pour l'analyse des résultats obtenus.

Dans le contexte du suivi thérapeutique, l'utilisation des critères anatomiques reste cependant la norme. En 1979, l'Organisation Mondiale de la Santé (OMS, ou WHO pour *World Health Organization*) proposait les premières recommandations pour l'évaluation de la réponse des tumeurs solides (Miller et al., 1981; WHO, 1979), incluant des mesures sur données anatomiques. L'objectif de normalisation des pratiques, largement atteint, a cependant soulevé de nombreuses discussions dans les communautés médicales. La collaboration entre trois grands groupes de travail, à savoir : l'EORTC (*European Organization for Research and Treatment of Cancer*) et les NCI (*National Cancer Institutes*) américaines et canadiennes, ont permis de proposer en 2000 de nouveaux critères : les critères RECIST pour *Response Evaluation Criteria in Solid Tumors* (ou RECIST 1.0 (Therasse et al., 2000)). Ceux-ci prenaient en compte les limites des critères OMS et les avancées technologiques apparues dans l'intervalle. Une mise à jour de ces critères, appelée RECIST 1.1, a été proposée en 2009 (Eisenhauer et al., 2009). Cependant, ces critères ne sont pas adaptés à toutes les localisations ou traitements, d'où la proposition de métriques plus spécifiques.

Parallèlement à ces travaux sur données anatomiques, l'apparition et le potentiel prédictif précoce de la TEP clinique ont permis à l'EORTC, dès 1999, de proposer un protocole clinique permettant l'évaluation de la réponse au traitement. Il se base sur la mesure de SUV effectuée sur des examens au FDG acquis précocement après le début du traitement (Young et al., 1999). La généralisation des TEP en services hospitaliers a conduit à un accroissement rapide des connaissances dans ce domaine et à la proposition, dès 2009, d'un protocole plus complet : PERCIST, pour *PET Response Criteria in Solid Tumors* (Wahl et al., 2009). Si les objectifs de ces recommandations sont de permettre l'intégration et la comparaison des informations métaboliques quantitatives dans les protocoles de recherches cliniques (phases I à III), elles visent, à plus long terme, à être appliquées à la routine clinique.

Dans ce chapitre, nous exposerons tout d'abord les critères anatomiques les plus utilisés en routine clinique pour évaluer la réponse tumorale à la thérapie. Nous décrivons ensuite les critères de Choi (Choi et al., 2007), qui consistent à prendre en compte, en plus des variations des dimensions caractéristiques des lésions, celles de la densité tumorale vue par la TDM. Après avoir détaillé et discuté les critères propres à la TEP et préconisés à l'heure actuelle pour l'évaluation de la réponse à la thérapie, nous présenterons les méthodes semi-

quantitatives proposées dans la littérature pour extraire une plus grande information que la seule activité moyenne ou le volume métabolique. Les critères fonctionnels autres que ceux relatifs à la TEP débordant du cadre de ce travail, nous n'avons pas développé ces aspects.

II.1. Critères anatomiques

L'évaluation de la réponse tumorale en pratique clinique repose le plus souvent sur de nombreuses informations allant des taux de marqueurs sériques (eg : taux de PSA dans le cas de cancer de la prostate) en passant par la palpation, l'endoscopie ou encore l'imagerie. L'utilisation des critères anatomiques que nous allons présenter fait donc partie d'un tout. Le but de ces protocoles est de fournir une information quantitative permettant d'aider le praticien à évaluer le succès ou l'échec de la stratégie thérapeutique adoptée.

II.1.1. Principales recommandations

II.1.1.1. Critères OMS

Les critères OMS (WHO, 1979) visaient à normaliser les pratiques cliniques et à fournir un langage commun dans le cadre de l'évaluation de la réponse tumorale au traitement. Les recommandations portant sur les données radiologiques (TDM, IRM ou échographie) classent les lésions en deux catégories :

- Les tumeurs mesurables. Elles sont caractérisées par la mesure, sur les données anatomiques considérées, et par une approche quelconque, du produit de leurs deux plus grands axes orthogonaux. Quand plusieurs lésions sont observées dans un même organe, la réponse se base sur l'ensemble de ces lésions en considérant la somme de ces produits.
- Les lésions non mesurables (eg : lymphangites pulmonaires métastatiques) sont caractérisées en estimant leur évolution ou en relevant leur apparition/disparition.

Les réponses de chaque tumeur (tableau II.1), évaluées sur des examens distants d'au moins 4 semaines, sont classées en Réponse Complète (RC), Réponse Partielle (RP), Maladie Stable (MS) ou Maladie Progressive (MP) et permettent de déduire la réponse globale du patient au traitement. L'ensemble des critères, pour les deux catégories de lésions, mesurables ou non, sont indiqués dans le tableau II.1. Le cas des métastases osseuses, explicité dans le texte de (WHO, 1979), n'a pas été rapporté ici.

Type de lésions	Réponse Complète (RC)	Réponse Partielle (RP)	Maladie Stable (MS)	Maladie Progressive (MP)
Mesurables	• Disparition de toutes les lésions connues sur deux examens séparés ≥ 4 semaines	• $\Delta PdL \leq -50\%$ • Pas d'apparition de lésion • Pas de progression d'une seule lésion	• $-50\% < \Delta PdL < +25\%$	• $\Delta PdL > +25\%$ pour une ou plus d'une lésion • Apparition d'une nouvelle lésion
Non-mesurables	• Disparition de toutes les lésions connues pendant au moins 4 semaines	• ΔPdL estimé $\leq -50\%$	• Pas de changement significatif • $-50\% < \Delta PdL$ estimé $< +25\%$	• ΔPdL estimé $> +25\%$ pour une ou plus d'une lésion • Apparition d'une nouvelle lésion

Tableau II.1 : critères OMS de la réponse tumorale au traitement. L'évaluation se base sur de l'évolution du Produit des plus grandes Longueurs orthogonales (ΔPdL) de chaque tumeur mesurable ou de cette même grandeur estimée dans le cas des tumeurs non mesurables. Ces mesures se font sur des examens espacés d'au moins 4 semaines.

La réponse globale du patient se déduit de la classification par lésion de la manière suivante :

- La plus mauvaise réponse doit prévaloir dans le cas de patients n'ayant que des lésions mesurables.
- Si les lésions non mesurables sont classées MS, alors la classification en RC des lésions mesurables est inchangée mais la classification globale du patient est RP.
- S'il y a un nombre supérieur ou égale de RC ou RP par rapport aux MS dans l'évaluation par organe, alors la réponse globale est RP.
- Si une seule lésion apparaît ou si une lésion est classée MP, la réponse globale au traitement est MP.

Ces critères sont discutables sur plusieurs points. D'abord, la notion de lésion non mesurable nécessite une estimation subjective. De même, le seuil de classification en MP peut être atteint à cause de biais de mesure ; en effet, une augmentation de $\sim 11\%$ de la mesure de chacun des axes de l'une des lésions induit une augmentation du produit des longueurs de 25%. De plus, ce protocole n'a pas permis d'atteindre l'objectif de normalisation des pratiques visé par l'OMS, car il ne spécifiait pas, par exemple, le nombre maximal/minimal de lésions à prendre en compte dans l'évaluation.

II.1.1.2. Critères RECIST 1.0

Pour dépasser les limites de ces critères et prendre en compte l'apparition en routine clinique de nouvelles technologies telles que l'IRM, les critères RECIST 1.0 ont été proposés en 2000 (Therasse et al., 2000). Les différences principales viennent de :

- La définition du nombre maximal de lésions à prendre en compte (10 au total et 5 au maximum par organe, sans pour autant donner d'indications sur la méthode de sélection à appliquer).

- Seule la mesure du plus grand axe des lésions est préconisée. On considère donc la somme des plus grands axes pour l'évaluation de réponse.
- La définition précise de *lésion non mesurable*, qui implique la non mesure de toute lésion de moins de 20 mm (ou 10 mm quand un TDM hélicoïdal est utilisé pour l'évaluation).
- La définition de *lésions cibles*, mesurables, mesurées et choisies pour l'évaluation sur l'examen avant traitement en fonction de leurs plus grandes dimensions.
- La définition de *lésions non cibles*, dont certaines peuvent ne pas être mesurables mais contribuer malgré tout à la réponse globale de la maladie. Elles font l'objet de critères différents de ceux utilisés pour définir la réponse des lésions cibles.

Les critères RECIST 1.0 de la réponse tumorale à la thérapie des lésions cibles et non cibles sont donnés dans le tableau II.2.

Type de lésions	Réponse Complète (RC)	Réponse Partielle (RP)	Maladie Stable (MS)	Maladie Progressive (MP)
Cibles	<ul style="list-style-type: none"> • Disparition de toutes les lésions connues sur deux examens séparés ≥ 4 semaines 	<ul style="list-style-type: none"> • $\Delta SdL \leq -30\%$ en prenant comme référence la SdL obtenue sur l'examen initial 	<ul style="list-style-type: none"> • $-30\% < \Delta SdL < +20\%$ en prenant comme référence la plus petite SdL mesurée depuis l'examen de référence 	<ul style="list-style-type: none"> • $\Delta SdL \geq +20\%$ en prenant comme référence la plus petite SdL obtenue depuis l'examen de référence • Apparition d'une nouvelle lésion
Non-cibles	<ul style="list-style-type: none"> • Disparition de toutes les lésions connues • Normalisation des marqueurs tumoraux 	<ul style="list-style-type: none"> • Persistance d'une ou plusieurs lésions et/ou • Maintien du niveau des marqueurs tumoraux au dessus des limites acceptables 		<ul style="list-style-type: none"> • Progression sans équivoque d'une lésion existante et/ou • Apparition d'une ou plusieurs nouvelles lésions

Tableau II.2: critères RECIST 1.0 de la réponse tumorale au traitement. L'évaluation des lésions cibles se base sur l'évolution de la Somme de leurs plus grandes Longueurs (ΔSdL). Ces mesures se font sur des examens espacés d'au moins 4 semaines (et 6-8 semaines dans le cas d'étude de phase II).

L'évaluation de la réponse globale est similaire à celle proposée par la classification OMS une fois la réponse des lésions cibles et non cible déterminée :

- Réponse complète au traitement si les lésions cibles et non cibles sont classées RC.
- Réponse partielle si les lésions cibles répondent (partiellement ou complètement) et si les lésions non-cibles ne sont ni RC ni MP.
- Maladie stable si les lésions cibles sont classées MS et si les lésions non cibles sont non MP.
- Maladie progressive dès lors que de nouvelles lésions sont apparues ou si les lésions, cibles ou non, ont progressé.

Une comparaison rapide des critères RECIST 1.0 et OMS met en évidence qu'une augmentation de 20% des plus grandes dimensions orthogonales d'une tumeur correspond à une augmentation du produit de ces longueurs de 44%. Cette lésion, à la limite d'être classée MP selon les critères RECIST 1.0 (tableau II.2) est vue comme largement MP si on se réfère aux critères OMS (tableau II.1). Ce fort écart potentiel entre les deux critères montre que

RECIST 1.0 est moins sensible que les critères OMS. Pour illustrer cette différence, (Therasse et al., 2000) relèvent que parmi l'ensemble des 234 patients classés MP, les critères OMS et RECIST 1.0 classent des réponses MP à la même date dans 92% des cas. Dans 7,3% des cas, les critères OMS ont classés une maladie progressive effectivement progressive plus tôt que les critères basés sur une mesure unidimensionnelle.

Afin d'évaluer la validité des critères RECIST 1.0, (Therasse et al., 2006) ont effectué une revue de 60 publications les ayant mis en œuvre entre 2000 et Novembre 2005. Parmi elles, 7/8 études confirment que les critères OMS permettent de détecter plus précocement la récurrence que les critères RECIST 1.0. Cependant, la comparaison de la classification obtenue par des mesures en volume (3D) et non plus unidimensionnelles valident bien les critères RECIST 1.0 (Therasse et al., 2000). L'étude site par site des résultats fournis par RECIST 1.0 a enfin mis en évidence que ces critères n'étaient pas adaptés à l'évaluation de la réponse des mésothéliomes ou des tumeurs stromales gastro-intestinales (GIST, *Gastro-Intestinal Stromal Tumours*) traités par Imatinib, un inhibiteur de l'activité tyrosine kinase. Dans le premier cas, des modifications des critères RECIST 1.0 ont été proposées (eg : (Byrne and Nowak, 2004)) car la mesure unidimensionnelle n'est pas adaptée pour ce type de pathologie. Dans le second, plus complexe car la réponse à la thérapie des GIST peut impliquer une augmentation du volume de la lésion par exemple, (Stroszczyński et al., 2005) ont proposé d'associer un examen IRM à l'évaluation. (Choi, 2005) a proposé puis formalisé (Choi et al., 2007) la prise en compte des changements simultanés des dimensions et de la densité de la tumeur mesurés sur images TDM. Nous reviendrons en détails sur ces derniers critères au paragraphe II.1.2.

II.1.1.3. Critères RECIST 1.1

Pour améliorer les performances de classification de RECIST 1.0 et prendre en compte ses limites (Therasse et al., 2006), (Eisenhauer et al., 2009) ont publié une mise à jour de ce protocole, sous la forme des critères RECIST 1.1 (cf tableau II.3). Elle vise à permettre :

- La prise en compte d'un nombre plus faible de lésions (5 au total dont un maximum de 2 par organe au lieu des 10 lésions préconisées dans la première version) représentatives de tous les organes atteints par la maladie, sans modifier la réponse globale du patient.
- La prise en compte des ganglions lymphatiques comme pouvant faire partie des lésions mesurables (si leur petit axe est ≥ 15 mm) ou non mesurables (quand le petit axe est compris entre 10 et 15 mm). C'est la longueur du petit axe qui est intégrée dans la somme des longueurs utilisée pour définir la réponse des lésions cibles. En dehors de la mention relative aux ganglions, la notion de *lésion cible* et *lésion non cible* est inchangée par rapport à la définition proposée dans les critères RECIST 1.0.
- L'intégration « raisonnable » des données qualitatives issues du FDG-TEP pour la prise en compte de l'apparition de nouvelles lésions.
- La diminution du taux d'erreurs de classification quand seules des petites lésions sont considérées. La progression de la maladie n'est donc plus définie seulement par la

croissance de 20% de la somme des plus grandes dimensions des tumeurs mesurables, mais aussi par une croissance, en valeur absolue, de plus de 5 mm.

- La précision de la notion de « progression univoque » utilisée pour décrire les lésions progressives non mesurables (cf Tableau II.2).

Type de lésion	Réponse Complète (RC)	Réponse Partielle (RP)	Maladie Stable (MS)	Maladie Progressive (MP)
Cibles	Sur deux examens séparés ≥ 4 semaines : • Disparition de toutes les lésions mesurables • Tous les ganglions lymphatiques pathologiques doivent avoir un petit axe $< 10\text{mm}$	En prenant comme référence la SdL obtenue sur l'examen initial : • $\Delta\text{SdL}(\text{relatif}) \leq -30\%$	En prenant comme référence la plus petite SdL mesurée depuis l'examen de référence : • $\Delta\text{SdL}(\text{relatif}) > -30\%$ • $\Delta\text{SdL}(\text{relatif}) < +20\%$ • $\Delta\text{SdL}(\text{absolu}) < +5\text{mm}$	En prenant comme référence la plus petite SdL obtenue depuis l'examen de référence : • $\Delta\text{SdL}(\text{relatif}) \geq +20\%$ • $\Delta\text{SdL}(\text{absolu}) \geq +5\text{mm}$ et/ou • Apparition de nouvelles lésions
Non-cibles	• Disparition de toutes les lésions connues • Normalisation des marqueurs tumoraux • Tous les ganglions lymphatiques pathologiques doivent avoir un petit axe $< 10\text{mm}$	• Persistance d'une ou plusieurs lésions. et/ou • Maintien du niveau des marqueurs tumoraux au dessus des limites acceptables		• Progression sans équivoque d'une lésion existante et/ou • Apparition de nouvelles lésions

Tableau II.3 : critères RECIST 1.1 de la réponse tumorale au traitement. L'évaluation des lésions cibles repose sur l'évolution relative de la Somme de leurs plus grandes Longueurs ($\Delta\text{SdL}(\text{relatif})$) ou sur sa variation absolue ($\Delta\text{SdL}(\text{absolu})$). Ces mesures se font sur des examens espacés d'au moins 4 semaines.

Les règles relatives à l'établissement de la réponse globale du patient au traitement sont identiques à celles détaillées pour RECIST 1.0.

(Wahl et al., 2009) relèvent différents problèmes communs à l'ensemble des méthodes de classification anatomiques. Le premier vient de la perte d'information inhérente à la réduction des données continues mesurées (eg : produit des plus grandes longueurs orthogonales de chaque lésion mesurable) en une classification discrète de la réponse (ie : RC, RP, MS ou MP). De plus, dans le cas de nouvelles stratégies thérapeutiques, dont l'objectif est de cibler une fonction cellulaire précise, l'évaluation de la réponse tumorale par les critères RECIST peut ne pas être le meilleur moyen d'estimer l'efficacité réelle du traitement. C'est le cas particulier des GIST traités par imatinib ou des hépatomes traités par sorafenib², dont les résultats cliniques n'ont montré qu'une réduction très lente du volume tumoral malgré une augmentation notable de la survie des patients. Ce dernier point montre que la sensibilité des critères anatomiques tels que RECIST 1.0 ou 1.1 n'est donc pas suffisante pour ce type de stratégie thérapeutique (Benjamin et al., 2007; Ratain and Eckhardt, 2004).

Pour remédier à ce problème, des critères plus adaptés au type de pathologie ou à une population d'intérêt donnée ont été proposés. On notera par exemple, dans le cas des carcinomes hépatocellulaires, l'existence des critères EASL (*European Association for Study of the Liver*, (Bruix et al., 2001)) qui reposent sur la prise en compte du rehaussement du

² Le sorafenib et l'imatinib sont deux inhibiteurs de certaines kinases, des enzymes impliquées par exemple dans la prolifération cellulaire, l'invasion des tissus sains ou encore l'angiogénèse.

contraste vasculaire établi sur des données radiologiques avec injection de produit de contraste. Ces critères ont fourni de meilleurs résultats que les critères RECIST 1.0 (Fornier et al., 2009) ou 1.1 (Spira et al., 2011). Nous allons revenir à titre d'illustration sur les critères de Choi (Choi et al., 2007) proposés pour l'évaluation de la réponse des GIST.

II.1.2. Autres critères : les critères de Choi

Nous avons vu que les critères RECIST, et notamment des critères RECIST 1.0, sous estiment la réponse tumorale précoce de patients atteints de GIST métastatiques traités par imatinib (Choi et al., 2004). Choi et ses collègues ont proposé de les modifier afin de prendre en compte la variation de densité tumorale due au traitement. Pour cela, ils proposent d'exploiter des mesures du coefficient d'atténuation contenu dans les TDM d'évaluation (exprimé en UH). Les variations du volume anatomique des GIST traités par imatinib étant faibles, ils proposent, en sus, de diminuer les valeurs des seuils proposés dans les critères RECIST 1.0 (qui portent sur la variation des dimensions des tumeurs, cf tableau II.2) (Choi et al., 2007). Les seuils suggérés par Choi et al. sont comparés à ceux des critères RECIST 1.0, dans le tableau II.4. La sélection des lésions sur lesquelles s'appliquent ces critères sont les lésions cibles définies dans les critères RECIST 1.0 (cf paragraphe II.1.1.2).

Critères (par patient)	Réponse Complète (RC)	Réponse Partielle (RP)	Maladie Stable (MS)	Maladie Progressive (MP)
RECIST 1.0 (Therasse et al. 2000)	<ul style="list-style-type: none"> •Disparition de toutes les lésions •Pas de nouvelle lésion 	<ul style="list-style-type: none"> •Variation de la somme des plus grandes longueurs $\leq -30\%$ •Pas de nouvelle lésion •Pas de MP parmi les lésions non mesurables 	<ul style="list-style-type: none"> •Ni RC, ni RP, ni MP •Pas de détérioration symptomatique attribuée à la progression tumorale 	<ul style="list-style-type: none"> •Variation de la somme des plus grandes longueurs $\geq +20\%$ •Apparition d'une nouvelle lésion
Choi (Choi et al., 2007)	<ul style="list-style-type: none"> •Disparition de toutes les lésions •Pas de nouvelle lésion 	<ul style="list-style-type: none"> •Variation de la somme des plus grandes longueurs $\leq -10\%$ ou •Variation de la moyenne des coefficients d'atténuation TDM (UH) des lésions $\leq -15\%$ •Pas de nouvelle lésion •Pas de MP parmi les lésions non mesurables 	<ul style="list-style-type: none"> •Ni RC, ni RP, ni MP •Pas de détérioration symptomatique attribuée à la progression tumorale 	<ul style="list-style-type: none"> •Variation de la somme des plus grandes longueurs $\geq +10\%$ et •Variation de la moyenne des coefficients d'atténuation TDM (UH) des lésions $\geq +15\%$ •Apparition d'une nouvelle lésion

Tableau II.4 : comparaison des critères RECIST 1.0 et Choi. ΔSdL : variation de la somme des plus grandes longueurs des tumeurs mesurables, comme indiqué dans (Therasse et al., 2000). Les différences entre les critères sont indiquées en rouge.

Appliquée sur une cohorte de 40 patients ayant subi 2 examens TDM à 2 mois d'intervalle, (Choi et al., 2007) montrent une spécificité (respectivement sensibilité) de leur critère à identifier la réponse vue par la TEP de 97% (respectivement 100%) contre 52% (respectivement 100%) pour les critères RECIST 1.0. L'un des problèmes de cette évaluation repose sur la définition des critères utilisés pour définir la réponse tumorale de référence. En effet, pour valider leur approche, les auteurs se basent sur des mesures de SUV_{max} obtenues sur un examen TEP au FDG effectué au même moment (i.e. 2 mois après le début du

traitement) que les examens TDM d'évaluation. La définition de la réponse TEP utilisée ne se base cependant pas sur les critères EORTC proposés à l'époque (Young et al., 1999) (cf paragraphe II.3.1), ces derniers n'ayant pas inclus, pour leur définition, de cas de GIST. Les critères TEP utilisés par (Choi et al., 2007) correspondent en réalité à une combinaison de données non fournies ou portant sur les résultats de deux études pour lesquelles la définition des seuils de réponses optimum étaient établies sur des TEP acquises après 21-40 jours (Van den Abbeele et al., 2002) ou après 1 semaine de traitement. Ainsi par exemple, une étude récente (Dudeck et al., 2011), visant à comparer les capacités prédictives sur 3 mois des critères RECIST 1.0 et Choi sur 51 patients atteints de GIST, traités en seconde ligne avec du sunitinib, a montré que les critères Choi classaient des réponses partielles (31,4%, contre 2% pour les critères RECIST 1.0), les critères RECIST 1.0 prédisaient plus souvent des maladies stables (72,5% contre 41,1% pour les critères Choi). En étudiant la survie globale et la survie sans progression, les auteurs ont de plus montré que les patients classés MS par RECIST 1.0 n'avaient pas un plus mauvais devenir que les patients classés MS ou RP selon les critères Choi.

II.1.3. Conclusion sur les critères anatomiques

Bien que peu prédictifs en début de traitement et non adaptés à certaines localisations/traitements, les critères OMS et RECIST sont largement employés en recherche pour évaluer la réponse thérapeutique. La variabilité de la mesure et la latence importante entre réponse métabolique et anatomique mesurable limitent cependant leur utilisation en clinique. Nous allons exposer dans quelles conditions les données issues de la TEP, acquises dans les premières semaines après le début du traitement, offrent des perspectives pour remédier à ce problème et évaluer précocement l'efficacité de la stratégie thérapeutique employée.

II.2. Critères métaboliques en TEP

II.2.1. Utilisation de la TEP au FDG en clinique

Comme nous l'avons vu dans le chapitre I, il est possible, en prenant certaines précautions, d'obtenir des images TEP quantitatives. Une revue récente sur les méthodes d'évaluation de la réponse au traitement, (Abgral et al., 2011) relève que l'analyse qualitative a été peu à peu abandonnée au profit d'outils d'analyse quantitative dans le cas des tumeurs solides. Ainsi, (Beyer et al., 2011) ont par exemple constaté que dans la pratique clinique, 90% des institutions relevaient et incluaient une valeur du SUV dans les rapports d'analyse des examens TEP/TDM. Toujours selon cette étude, 91% des sites utilisent cet index pour le suivi du traitement, là où seulement 35% s'en servent pour différencier les lésions malignes des lésions bénignes. Cette étude est cependant biaisée car basée sur les données fournies par les centres ayant accepté de répondre au sondage qui leur a été soumis.

L'analyse visuelle est toujours mise en œuvre pour l'évaluation de la réponse thérapeutique des lymphomes (Abgral et al., 2011) :

- Pour l'évaluation de fin de traitement (comme c'est le cas pour les critères IHP (*International Harmonization Project in lymphoma*), décrits par (Cheson et al., 2007; Juweid et al., 2007), mais qui nécessitent cependant des mesures sur TDM).
- Pour l'évaluation précoce au cours du traitement (eg : échelle de Deauville, du nom de la ville où s'est produit le colloque international qui a permis à (Gallamini et al., 2009) de rédiger leurs critères).

Si la mesure du SUV est explicitement déconseillée dans les critères de Deauville, de nombreuses équipes ont cependant montré la supériorité des index semi-quantitatifs pour l'évaluation précoce (~2 cycles de chimiothérapie) de la réponse de ces maladies (Casasnovas et al., 2011; Lin et al., 2007).

II.2.2. La quantification en TEP au FDG dans le contexte du suivi thérapeutique : difficultés et enjeux

Si l'usage fréquent du SUV s'explique d'abord par sa simplicité d'obtention en comparaison aux méthodes plus quantitatives, il est nécessaire d'évaluer l'impact de ses limites. Après avoir montré comment la littérature propose de minimiser la variabilité de cette mesure, critique dans le cadre d'études multicentriques, nous présenterons les principaux protocoles permettant de classer la réponse tumorale à partir de la TEP au FDG. Nous présenterons enfin une vue d'ensemble des méthodes originales proposées dans la littérature et dont l'objectif est d'extraire plus d'information sur le métabolisme tumoral à partir des images TEP.

II.2.3. Facteurs affectant la mesure du SUV

Si la première partie de notre exposé a permis d'identifier les phénomènes qui pouvaient influencer la mesure du métabolisme glucidique en TEP, il reste à déterminer la part de chacun de ces paramètres sur la mesure du SUV. La littérature abondant sur ce sujet (eg : (Buvat, 2007; Huang, 2000; Keyes, 1995), nous renvoyons le lecteur à l'une des dernières revues effectuées sur ce sujet, proposée par (Adams et al., 2010), à laquelle cette section se réfère en partie.

II.2.3.1. Facteurs biologiques

Le SUV est le plus souvent mesuré sur des données statiques acquises environ 1 heure après l'injection du radiotracer. Les facteurs biologiques affectant la mesure sont donnés dans le tableau II.5.

Facteur	Description de l'effet sur la mesure du SUV	Ampleur de l'effet sur la mesure
Répartition de la masse corporelle	-Les tissus adipeux fixent moins le FDG que les tissus musculaires : •SUV d'autant plus élevé que la masse grasseuse du patient est élevée.	Chez les patients en surpoids, le SUV _{BW} dans les tissus sains (foie, rate ou sang) jusqu'à 2 fois plus élevé que chez les patients maigres (Zasadny and Wahl, 1993).
Calcul/mesure du facteur de pondération	-Le poids du patient peut changer pendant le traitement. -Utilisation de la surface corporelle totale/masse maigre plus pertinente?	-Diminution de la médiane du poids de patients atteints de cancers du poumons de 6,5% (hommes) et 11% (femmes) pendant le traitement (Harvie et al. 2003). -Diminution de la corrélation entre SUV et poids en utilisant d'autres facteurs de pondération.
Taux de glucose dans le sang	-Inhibition de la captation de FDG si taux de glucose sanguin trop élevé : • \nearrow glycémie \Rightarrow \searrow SUV	-Détérioration de la qualité de l'image. -Disparition de 30% des lésions connues quand glycémie trop élevée et sous-évaluation de la décroissance du SUV dans le cas de carcinomes colorectaux (Crippa et al. 1997).
Délai entre injection et acquisition (Δ Tacq)	-L'élimination du FDG est plus rapide dans les tissus normaux que dans les tissus tumoraux. - \nearrow Δ Tacq \Rightarrow \nearrow SUV	-SUV de nodules pulmonaires malins jusqu'à 30% plus grand pour un Δ Tacq passant de 60min à 120min (Lowe et al. 1995).
Mouvement respiratoire	-Mouvement respiratoire vu par la TEP et non pris en compte par le TDM utilisé pour la correction d'atténuation : •correction d'atténuation inexacte •SUV biaisé (mesure \searrow si amplitude du mouvement \nearrow) •Effet dépendant de la taille de la lésion, sa localisation, du cycle respiratoire du patient	-En fonction de la phase du cycle respiratoire à laquelle est acquis le TDM d'atténuation : variation du SUV _{max} de 30% et jusqu'à 21% de réduction de dimension du volume métabolique entre phases extrêmes du cycle respiratoire (Erdi et al. 2004). -Surestimation de 100% du volume des petites lésions à proximité du diaphragme (Kinahan and Fletcher, 2010).
Confort du patient	-Le stress ou mauvaises conditions de repos du patient après injection du traceur impliquent: • \nearrow SUV dans la graisse brune ou les muscles •Biais de la mesure	-La graisse brune est une source potentielle de faux-positifs. L'acquisition de 2 TEP/TDM rapprochées dans le temps de 32 patients a montré une \nearrow SUV _{max} de la graisse brune dans 76% des cas, variant de +12 à +192% (moyenne : 42%) (Alkhalil and Alavi, 2008).
Inflammation	Méconnaissance du processus inflammatoire dans la zone tumorale d'intérêt implique: • \nearrow SUV non liée à la consommation glucidique des cellules tumorales	-Inflammation \Rightarrow \nearrow du nombre de faux-positifs notamment lors de l'identification de lésions bénignes (Zhuang et al., 2001)

Tableau II.5 : facteurs biologiques affectant la mesure du SUV. (Extrait et complété de (Adams et al., 2010) et (Boellaard, 2009)).

II.2.3.2. Techniques

Les facteurs techniques ont un effet important sur la mesure. Si l'effet de chacun d'eux sur la valeur du SUV obtenue est faible en moyenne (<15%), leur accumulation peut entraîner une grande variabilité des mesures entre institutions (Boellaard, 2009). Ces facteurs incluent : les paramètres technologiques (types de détecteurs utilisés ou encore utilisation de la mesure temps de vol), les paramètres de reconstruction, l'erreur sur le facteur de calibration de la dose injectée, l'erreur de mesure sur le temps écoulé entre injection du radiotracer et début d'acquisition, l'utilisation d'agent de contraste pouvant modifier l'atténuation vue par le TDM utilisée pour la correction d'atténuation et le contourage du volume tumoral métabolique (manuel ou non). Un résumé de leur impact sur la mesure du SUV est présenté dans le tableau II.6.

Facteur	Détails des sources de biais	Impact de l'effet sur la mesure
Paramètres technologiques relatifs au tomographe	<ul style="list-style-type: none"> •Dimensions des cristaux •Méthode de correction des coïncidences fortuites •Mise en œuvre du temps de vol •Autre facteurs 	-Différences de $\pm 10\%$ sur la concentration d'activité mesurée, malgré le choix d'un protocole commun entre machines différentes (Geworski et al., 2002)
Paramètres de reconstruction	<ul style="list-style-type: none"> •Dimension des pixels de l'image reconstruite. •Algorithme utilisé et paramètres associés. •Utilisation d'un post-filtrage pendant/après la reconstruction. 	<p>-Le SUV moyen de voxels dans un VOI converge plus vite que le SUVmax en fonction du nombre d'itérations utilisées pour la reconstruction OSEM, du fait de l'augmentation du bruit dans l'image (Jaskowiak et al., 2005).</p> <p>-Différences non significatives entre SUV calculés sur image reconstruites par méthodes analytiques (rétroprojection filtrée) et itératives (OSEM) (Krak et al., 2005).</p> <p>-Description du choix des différents paramètres capitale : influence la mesure du SUV jusqu'à 30% (Boellaard, 2009).</p>
Erreur de calibration de la dose injectée	-La dose injectée au patient n'est pas celle prise en compte pour le calcul du SUV.	<p>-Variation du facteur de calibration de $\sim 4\%$ si aucune erreur de manipulation.</p> <p>-Jusqu'à plus de 10% dans le cas contraire (Kinahan and Fletcher, 2010).</p>
Erreur sur la base de temps utilisée par la TEP/TDM	-La dose injectée est mal corrigée de la décroissance radioactive qui se produit entre injection et acquisition.	-Un écart de 8 minutes induit une erreur de mesure sur le SUV de 5%.
Injection d'agent de contraste pour améliorer la visualisation de structures sur données TDM	-Augmentation artificielle du contraste vue par la TDM utilisée pour la correction d'atténuation.	<p>-Modifications significatives (faibles dans la tumeur $\sim 7\%$) des mesures SUV (max ou moyen) quelles que soient les structures, en phase veineuse ou artérielle (Behrendt et al. 2012).</p> <p>-Pas de modification de l'interprétation visuelle.</p>
Méthode de segmentation utilisée pour délimiter la zone tumorale	<p>-Le volume et la variabilité du volume d'intérêt dépendent de la méthode de segmentation utilisée, qu'elle soit manuelle, semi-automatique ou automatique.</p> <p>-Très grand nombre de méthodes décrites dans la littérature.</p> <p>-Méthode de référence non idéale (manuelle généralement) rend la validation de nouvelles méthodes complexes (Zaidi and El Naqa, 2010).</p>	<p>-Valeur absolue du SUV très dépendante de l'algorithme de segmentation utilisé (eg : voxel de plus grande intensité dans la tumeur, moyenne de tous ses voxels dont le SUV est $>2,5$).</p> <p>-Nombreuses définitions d'une même ROI pouvant induire de grande variation de la mesure du SUV d'une même tumeur (eg : variation de SUVpeak jusqu'à $\pm 50\%$ dans le cas du FLT, en fonction de la définition de la ROI considérée, (Vanderhoek et al., 2012)).</p>
Erreurs des programmes de lectures des données	-Après acquisition des données, les images sont reconstruites et transformées au format DICOM avant d'être transférées à la station de lecture. Cette station, intégrée ou non au tomographe, permet d'interpréter les images i.e. de lire les données au format SUV, de tracer des VOI et d'évaluer la localisation des foyers de fixations par rapport aux données anatomiques.	-A partir d'une image au format DICOM d'un TEP/TDM d'un fantôme NEMA de référence (ou DRO, <i>Digital Reference Object</i>) dont les paramètres sont connus du seul laboratoire de référence, (Kinahan et al., 2012) ont proposé à 13 centres volontaires de télécharger le DRO et de lire les fichiers DICOM sur leurs stations de lectures respectives. Les auteurs ont montré qu'une de ces stations avait obtenue des erreurs de plus de 60% lors de l'estimations du SUVmax des VOI du fantôme.

Tableau II.6 : vue d'ensemble des facteurs techniques pouvant influencer la mesure du SUV.

II.2.4. Besoin de normalisation des pratiques

La mesure du SUV dépend de la préparation du patient, de la machine utilisée, des protocoles d'acquisition, des paramètres de reconstruction ou encore des méthodes d'analyse d'image. Idéalement, pour diminuer les biais lors du suivi longitudinal d'un même patient, il est conseillé d'effectuer l'acquisition dans les conditions identiques quel que soit l'examen. Le trop grand nombre de différences observées dans la littérature dans le choix des facteurs pouvant influencer la mesure du SUV complique le partage des données entre les centres. Ce

partage permettrait pourtant la définition de critères communs utiles à la généralisation les seuils de classification de la réponse tumorale. Pour diminuer ces biais, la communauté a proposé dès 1998 de normaliser les pratiques (Schelbert et al., 1998). Les critères EORTC appliqués à la TEP (Young et al., 1999), proposés en 1999, sont longtemps restés les seuls critères d'évaluation de la réponse au traitement. En dehors de critères TEP spécifiques aux lymphomes évoqués précédemment (Cheson et al., 2007; Juweid et al., 2007) et sur lesquels nous ne reviendrons pas, seuls les critères PERCIST (Wahl et al., 2009), basés sur une bibliographie plus riche, ont été proposés pour répondre au besoin de normalisation des pratiques de la médecine nucléaire.

Une excellente revue de (Boellaard, 2009) a fait le bilan de l'ensemble des recommandations publiées dans l'intervalle entre la parution des critères EORTC et les critères PERCIST. Elle a entre autre permis de montrer que tant que l'objectif du praticien n'était pas de participer à une étude multicentrique et/ou ne visait pas à mesurer des grandeurs absolues, l'application de recommandations strictes (eg : (Boellaard et al., 2008) ou plus récemment encore (Boellaard et al., 2010)) n'était pas requise. Dans le cas contraire, l'utilisation de protocoles d'acquisition et d'analyse stricts rend la mesure du SUV très reproductible. En se basant sur les résultats d'une série d'études, (Weber, 2010) a ainsi relevé que la mise en œuvre de tels protocoles permettait de réduire l'écart-type des mesures répétées du SUV à environ 10%.

II.3. Recommandations et critères de classifications

II.3.1. Critères de l'EORTC

En publiant des critères de classification dès 1999, l'EORTC visait à proposer des méthodes de travail communes pour faciliter les études TEP multicentriques.

II.3.1.1. Recommandations pratiques

II.3.1.1.1. Préparation du patient

Les critères EORTC préconisent que le patient soit à jeun avant l'examen et que la glycémie soit mesurée. Un niveau de 4-7 mmol/L est attendu mais aucune recommandation n'est formulée quand une hyperglycémie est constatée. L'utilisation d'insuline est tolérée pour les patients diabétiques. L'utilisation de diazepam, un anxiolytique parfois utilisé pour ses propriétés anti-convulsivantes, est enfin suggérée pour permettre la diminution de la fixation musculaire physiologique, surtout quand la zone d'intérêt se situe dans la sphère ORL (Oto-Rhino-Laryngé).

Si l'utilisation d'un examen statique est choisie (pour la mesure du SUV), l'EORTC préconise de commencer l'acquisition 50-70 minutes après injection du radiotracer et de se servir de ce temps de référence pour tous les examens d'un même patient. Dans le cas où l'approche graphique de Patlak (recommandée) est retenue, il est conseillé d'effectuer une acquisition

dynamique d'au moins 60 minutes en parallèle d'une mesure en continue (ou permettant d'obtenir des résultats équivalents) de la fonction d'entrée.

II.3.1.1.2. Fréquence des examens TEP

Bien que reconnaissant un manque important de données pour optimiser leurs recommandations, l'EORTC suggère :

- L'acquisition d'une TEP, appelée TEP prétraitement, aussi proche que possible du début du traitement (avec un écart maximum de 2 semaines).
- Une latence de 1-2 semaine(s) est suggérée entre la fin de la chimiothérapie et l'examen TEP post-traitement pour éviter la mesure d'effets transitoires du traitement sur le volume tumoral, qui pourraient biaiser l'évaluation de la réponse.
- Le lien entre les changements métaboliques de la tumeur et le traitement par radiothérapie n'étant pas clair au moment de la rédaction du protocole, les auteurs ne donnent explicitement aucune recommandation pour ce type de traitement.

II.3.1.1.3. Mesure de la fixation

Considérant tous les points discutés dans le chapitre I concernant les différentes méthodes de quantification, l'EORTC, bien que soulignant son manque de sensibilité, préconise de mesurer le SUV normalisé par la surface corporelle totale (SUV_{BSA} , [m^2]) défini, dans le texte par les équations suivantes :

$$SUV_{BSA}(m^2) = \frac{\text{Activité dans les tissus } (Bq \cdot mL^{-1}) \times BSA}{\text{Dose injectée}(Bq)} \quad \text{Eq. II.1}$$

où la surface corporelle totale (BSA) est obtenue à partir de la formule empirique de (Du Bois and Du Bois, 1916) (cf équation I.18).

La correction du glucose circulant n'ayant pas permis de démontrer un avantage net sur la précision de la mesure au moment de la rédaction de ces critères, (Young et al., 1999) ne préconisent pas son utilisation.

II.3.1.1.4. Définition des volumes d'intérêts

Le choix du volume d'intérêt utilisé pour évaluer la réponse au traitement doit permettre de mesurer les évolutions de la fixation du FDG de la tumeur. Les recommandations sont donc les suivantes :

- Le VOI caractéristique de la tumeur défini sur la TEP prétraitement doit être tracé sur la région tumorale viable de forte fixation en FDG.
- Le même VOI doit être positionné sur les examens de suivi au plus près des zones tumorales initiales.
- Les mesures de fixation doivent inclure l'activité moyenne et maximale des voxels contenus dans le VOI.
- Les variations des dimensions de la fixation de FDG tumorale doivent être indiquées, i.e. en mesurant les dimensions des plus grands axes orthogonaux de la tumeur. Aucune information sur la méthode à adopter pour obtenir cette mesure n'est cependant indiquée.
- Pour rendre compte de l'effet de volume partiel (EVP), dont l'impact sur la mesure du SUV peut être important (cf chapitre III), les dimensions anatomiques des tumeurs étudiées doivent, si possible, être mesurées et comparées à la résolution de la TEP utilisée. Là encore, aucune méthode de mesure particulière n'est spécifiée.

Il est important de noter que ce protocole ne définit pas les notions de lésion cible ou de lésion mesurable. De plus, aucune information n'est donnée sur à la prise en compte de l'éventuelle activité métabolique des ganglions.

II.3.1.2. Critères de réponse tumorale

En considérant le niveau de reproductibilité de la mesure sur les TEP conventionnelles (de l'ordre de 10-20%) et le fait qu'une réduction précoce (de 15-25% après un cycle de chimiothérapie) ou plus tardive (de plus 25%) mais plus fiable (après 2 ou 3 cycles) du SUV soit prédictive de la réponse tumorale, l'EORTC a proposé les critères de classification donnés dans le tableau II.7.

La variation relative du SUV utilisée dans ce tableau est définie par :

$$\Delta SUV(\%) = \frac{SUV(T_2) - SUV(T_1)}{SUV(T_1)} \times 100 \quad \text{Eq. II.2}$$

où l'activité dans les tissus, utilisée pour calculer le SUV de la tumeur (équation II.1) peut correspondre à l'activité du voxel de plus forte intensité dans la tumeur ou à la moyenne des activités des voxels contenu dans le VOI qui caractérise la tumeur (cf paragraphe II.3.1.1.4). T_1 correspond au jour d'acquisition de l'examen prétraitement et T_2 à celui de l'examen de contrôle.

Réponse Complète (RC)	Réponse Partielle (RP)	Maladie Stable (MS)	Maladie Progressive (MP)
-Fixation de FDG dans le volume tumoral seulement imputable à l'activité des tissus sains avoisinants.	-Après un cycle de chimiothérapie: $\Delta\text{SUV} < -15 / -25\%$ -Après plus d'un cycle de chimiothérapie: $\Delta\text{SUV} < -25\%$	• $-15\% < \Delta\text{SUV} < 25\%$ et • Pas d'augmentation visible du volume métabolique (augmentation $< 20\%$ du plus grand axe).	• $\Delta\text{SUV} > 25\%$ ou • Croissance visible du volume métabolique tumoral (augmentation $> 20\%$ du plus grand axe). ou • Apparition de nouvelles zones de captation parmi les lésions métastatiques.

Tableau II.7 : détails des critères d'évaluation de la réponse tumorale selon l'EORTC (Young et al., 1999). ΔSUV : évolution relative du SUV entre l'examen TEP prétraitement et celui effectué pendant/après le traitement. La définition du SUV à utiliser, maximal ou moyen dans le VOI, n'est pas spécifiée dans le protocole.

II.3.1.3. Conclusions sur les critères EORTC en TEP

Si ces critères ont l'avantage d'avoir formalisé les bases de la normalisation en TEP, ils n'évitent pas les difficultés inhérentes à cet objectif de normalisation. Il subsiste en effet de nombreux paramètres mal définis et laissés à l'appréciation des utilisateurs. Il n'est ainsi pas fait mention de la manière d'évaluer la réponse globale du patient, de la nature ou du nombre de lésions à inclure dans l'analyse ou encore d'un critère permettant de distinguer les *lésions mesurables* des lésions *non mesurables*. Ainsi, si les auteurs insistent sur le fait que la littérature ne leur permet pas de préciser certains points (eg : choix du seuil de 15/25% de décroissance du SUV pour la réponse partielle), il en résulte des imprécisions délétères sur la classification de la réponse de la maladie.

En effet, si le choix de la normalisation à appliquer pour calculer le SUV est arrêté, la précision de la définition du VOI dans lequel il doit être calculé (voxel d'intensité maximal ou moyenne des SUV des voxels dans le VOI) n'est pas tranchée. Une étude récente de (Buvat et al., 2012) a par exemple montré que cette définition influençait la valeur du seuil permettant d'optimiser la discrimination entre les tumeurs répondant et ne répondant pas à la thérapie dans le cas de cancers colorectaux métastatiques. Pour ce type de lésions, une décroissance du SUV_{max} de 14% apparaît être le seuil optimum alors qu'il est de 22% pour un SUV moyen incluant l'activité de tous les voxels d'intensité supérieure à 40% du SUV_{max} .

De même, seule l'utilisation d'une méthode de segmentation manuelle étant suggérée, on peut imaginer que la variabilité de la mesure de longueur ou de SUV moyen, et donc la

classification de la réponse, soit très opérateur dépendant (Zaidi and El Naqa, 2010). Cette variabilité doit cependant prendre en compte le fait que seules des grandeurs relatives sont utilisées pour la classification, ce qui permet de compenser une partie des erreurs (Boellaard, 2011). Enfin, à la différence des critères RECIST par exemple, aucune indication n'est donnée sur la réponse globale du patient et sur la manière d'appliquer les critères précédents quand plusieurs lésions sont identifiées sur le TEP de référence (eg : mélange de réponses tumorales stables et progressives).

II.3.2. Critères PERCIST

Profitant d'une littérature plus fournie et de protocoles d'acquisitions optimisés afin de réduire la variabilité des mesures effectuées lors d'études multicentriques, (Wahl et al., 2009) proposent des critères plus complets.

II.3.2.1. Préparation du patient

L'un des éléments important du protocole PERCIST est la définition d'un protocole d'acquisition optimisé pour les études multicentriques. Il est basé sur le protocole néerlandais, très strict, publié par (Boellaard et al., 2008).

En bref, le patient doit être à jeun depuis 4 à 6 heures avant le début de l'examen, avoir une mesure de glycémie non corrigée $< 11,1$ mmol/L et ne pas avoir utilisé d'insuline. Comme pour les critères EORTC, l'acquisition des données doit se faire 50-70 minutes après injection. L'écart entre l'injection de FDG et le début de l'acquisition des examens de suivi ne doit pas être différent de plus de 15 minutes par rapport à celui de l'examen de référence (en respectant le délai de fixation minimum de 50 minutes). Enfin, il est recommandé d'effectuer l'acquisition de chaque TEP sur le même appareil et d'injecter la même dose de traceur, une tolérance de $\pm 20\%$ étant cependant admise.

II.3.2.2. Temps entre chaque examen

Le temps à laisser pour évaluer la réussite du traitement entre les examens TEP prétraitement et de celui de contrôle est encore débattu. Bien qu'il dépende probablement du type de tumeur ou encore de la nature du traitement, (Wahl et al., 2009) suggèrent que la mesure précoce, après 1 cycle de chimiothérapie, permettrait de déterminer le développement de résistance tumorale au traitement.

Par ailleurs, l'acquisition d'un TEP post-traitement permet de valider la réussite ou l'échec du traitement et/ou de proposer de nouvelles stratégies thérapeutiques. Si cette dernière option est envisagée, les auteurs préconisent, comme pour les critères EORTC, d'attendre au minimum 10 jours entre ce TEP d'évaluation et la fin de la chimiothérapie afin de ne pas mesurer

d'effets transitoires. Dans le cas d'une radiothérapie, il est recommandé d'attendre 8-12 semaines du fait des réactions inflammatoires (aigües) et de la formation de fibrose (chronique) liée au traitement.

II.3.2.3. Mesure de la fixation

Pour caractériser la réponse tumorale, seule la mesure du SUV est préconisée. Il est conseillé de le normaliser par la masse maigre (cf paragraphe I.3.4.1.2, ou SUV_{LBM} , que l'on notera simplement SUL pour la suite de la description des critères PERCIST) mais de ne pas la corriger du taux de glucose dans le sang. A l'inverse des critères EORTC, les critères PERCIST ne donnent pas explicitement de formule ou de méthode de référence pour calculer le facteur de normalisation par la masse maigre.

II.3.2.4. Définition du volume d'intérêt

L'une des particularités de ces critères est de comparer la fixation tumorale (SUL_{peak}) à une activité de référence (SUL_{ref}) afin de distinguer les lésions mesurables des lésions non mesurables. L'obtention de SUL_{ref} se base sur le calcul de la valeur moyenne (SUL_{fond}) et sur l'écart-type (noté σ_{fond}) de l'activité des voxels contenus dans une ROI placée dans un organe de référence. Les ROI à utiliser dépendent de l'organe considéré.

- Mesure de l'activité de référence (SUL_{ref})

Quand le foie est sain, on considère comme valeur de SUV_{fond} la moyenne des SUL des voxels contenus dans une ROI correspondant à une sphère de 3 cm de diamètre placé dans le lobe droit du foie. L'écart-type σ_{fond} (*foie*) de ces voxels est aussi calculé. La fixation du foie est considérée normale et peut être utilisée pour l'évaluation quand elle ne varie pas de plus de 20% (et de 0,3 unité SUL) par rapport à l'examen de référence. La valeur de référence considérée dans ce cas est :

$$SUV_{ref}(foie) = 1,5 \times SUL_{fond} + 2 \times \sigma_{fond}(foie) \quad \text{Eq. II.3}$$

Quand le foie est pathologique ou si les critères de stabilité dans le temps de cet organe ne sont pas satisfaits, SUL_{ref} et son écart-type sont déterminées à partir d'une ROI placée dans le compartiment sanguin. Celle-ci est obtenue en considérant l'activité des voxels d'un cylindre de diamètre 1 cm et de hauteur 2 cm dans la direction axiale, placée dans l'aorte thoracique descendante. La valeur de référence à employer dans ce cas s'écrit sous la forme :

$$SUV_{ref}(aorte) = 2,0 \times SUL_{fond} + 2 \times \sigma_{fond}(aorte) \quad \text{Eq. II.4}$$

Les critères de stabilité à utiliser pour valider l'utilisation du SUV_{ref} obtenu dans le compartiment sanguin sont identiques à ceux du foie.

On retiendra que le $SUV_{ref}(foie)$ est de l'ordre de 2,5, valeur souvent utilisée pour la segmentation de volume cible en radiothérapie (Nestle et al., 2005).

- Mesure de l'activité des lésions cibles

L'évaluation de la réponse des lésions cibles repose sur la mesure de la moyenne des SUL (noté SUL_{peak}) des voxels contenus dans une sphère de diamètre maximal égal à 1,2 cm (soit un volume ~ 1 mL). Cette sphère doit être placée dans la lésion de manière à maximiser la valeur du SUL_{peak} calculée (Wahl et al., 1993).

Si seule la mesure du SUL_{peak} peut permettre d'évaluer la réponse des tumeurs au traitement (cf critères de réponse présentés dans le tableau II.8), les auteurs proposent de relever systématiquement le SUL du voxel de plus grande intensité dans la lésion (SUL_{max}) et, à titre exploratoire, d'y ajouter la mesure :

- Des SUL moyens de chaque lésion mesurable dans des ROIs contenant tous les voxels de SUL supérieures à 50% et 70% du SUL_{peak} : $SUL_{50\%}$ et $SUL_{70\%}$
- Du TLG (*Total Lesion Glycolysis*), défini par (Larson et al., 1999) comme le produit, pour chaque lésion, du volume métabolique et de son activité moyenne. La ROI à considérer pour calculer le volume selon (Wahl et al., 2009) consiste à prendre tous les voxels de la lésion d'intérêt dont le SUL est supérieur à $SUL_{ref}(foie)$. Un facteur 3 sur l'écart-type (équation II.4, terme de droite) peut être utilisé au lieu du facteur 2 pour « les tumeurs très actives ».

II.3.2.5. Critères de réponse tumorale

Comme les critères anatomiques (RECIST 1.0 et RECIST 1.1), les critères PERCIST définissent tout d'abord la notion de lésion mesurable et de lésion cible.

- Lésion mesurable

Toute lésion est mesurable si son activité (SUL_{peak}) est supérieure à l'activité de référence (SUR_{ref}) établie précédemment.

- Lésion cible

Les lésions cibles correspondent aux **5 tumeurs** mesurables (au maximum 2 tumeurs par organe) présentant l'activité métabolique la plus intense à un instant donné (ie ces lésions peuvent ne pas être les mêmes d'un examen de suivi à l'autre).

L'évaluation de la réponse au traitement du patient se base sur la mesure du SUV_{peak} de la lésion la plus active mesurée sur chaque examen. Il est possible, en seconde intention, d'appliquer les critères de réponse à la somme des SUL_{peak} des lésions cibles définies sur chaque TEP de suivi (que ces lésions soient les mêmes entre les différents examens ou non) en seconde intention. La classification consiste à comparer la variation (relative et absolue) de cette grandeur entre la TEP prétraitement et les TEP post-traitement à différents seuils.

En bref, une réponse complète correspond à la disparition visuelle de toute tumeur métaboliquement active. Une réponse partielle est définie par la décroissance de plus de 30% et de 0,8 unité SUL de la tumeur la plus active. A l'inverse, une croissance de 30% et de 0,8 unité SUL, ou l'apparition de nouvelles lésions, décrit une maladie progressive.

Lésion	Réponse Complète (RC)	Réponse Partielle (RP)	Maladie Stable (MS)	Maladie Progressive (MP)
Mesurable	<ul style="list-style-type: none"> Disparition complète de la fixation des lésions mesurables ($SUL_{peak} < SUL_{ref}$ ET non distinguable de l'activité de fond du compartiment sanguin) Disparition de toute lésion (non distinguable de l'activité de fond du compartiment sanguin) Pas de nouvelle lésion avide en FDG et dont les caractéristiques de la fixation pourraient «évoquer celle d'un cancer» Si RECIST classe la maladie MP, vérifier le résultat à partir d'un autre examen 	<ul style="list-style-type: none"> Pour la tumeur d'activité la plus élevée <ul style="list-style-type: none"> $\Delta SUL_{peak} < -30\%$ et <ul style="list-style-type: none"> $\Delta SUL_{peak} < 0,8$ Quelle que soit la lésion, cible ou non <ul style="list-style-type: none"> $\Delta SUL_{peak} < +30\%$ et <ul style="list-style-type: none"> Pas d'$\uparrow > 30\%$ de la plus grande dimension <u>anatomique</u> (cf MP selon RECIST ou OMS; dans le cas contraire, vérifier le résultat avec un autre examen) 	<ul style="list-style-type: none"> Ni RC, ni RP, ni MP 	<ul style="list-style-type: none"> Pour la tumeur d'activité la plus élevée : Si l'\uparrow de fixation est non imputable à de l'inflammation/traitement: <ul style="list-style-type: none"> $\Delta SUL_{peak} > +30\%$ et <ul style="list-style-type: none"> $\Delta SUL_{peak} > 0,8$ ou <ul style="list-style-type: none"> \uparrow visible du volume métabolique ($\Delta TLG > +75\%$, sans $\searrow SUL$) Apparition de nouvelle lésion, dont la fixation serait «typique de celle d'un cancer»
Non mesurable	<ul style="list-style-type: none"> Disparition de toutes les lésions 	<ul style="list-style-type: none"> Quand non MP <ul style="list-style-type: none"> Persistance d'au moins une lésion ou <ul style="list-style-type: none"> Marqueurs tumoraux au dessus des limites acceptables 		<ul style="list-style-type: none"> Progression sans équivoque des lésions non cibles ou Apparition de nouvelle lésion, dont la fixation serait «typique de celle d'un cancer»

Tableau II.8 : critères PERCIST d'évaluation de la réponse tumorale au traitement. Dans le cas où 5 lésions cibles sont retenues sur la TEP prétraitement, la somme des SUL_{peak} peut être utilisée pour l'évaluation, afin de réduire les sources d'erreurs. ΔSUL : évolution, relative ou absolue (selon le seuil) entre l'examen TEP prétraitement et les examens de suivi, du SUV normalisé par la masse maigre.

Une autre métrique, basée sur l'utilisation du TLG, est proposée à des fins exploratoires. Dans ce cas, une croissance du TLG de plus de 75% peut servir de base à la description d'une maladie progressive. Les critères détaillés sont fournis dans le tableau II.8.

La réponse globale au traitement est évaluée de la manière suivante :

- Elle correspond à la réponse la plus optimiste obtenue à partir des lésions mesurables jusqu'à la progression ou à la reprise de la maladie.

- Si les lésions **non cibles**, mesurables ou non, sont classées RP/MS et que la réponse des lésions cibles est complète (CR), alors la réponse globale est une réponse métabolique partielle (RP).
- Si les lésions **non cibles** sont classées RP/MS, alors la classification en RP des lésions cibles n'est pas modifiée.

II.3.2.6. Conclusion sur les critères PERCIST en TEP

Si de nombreuses questions restent en suspens dans la version actuelle des critères PERCIST, ils permettent néanmoins de formaliser certains points déjà présents dans les critères anatomiques de références, comme par exemple, les notions de lésion mesurable et de lésion cible. A contrario donc, on note que ces critères, proposés pour fournir une base de travail solide (eg : choix de protocoles d'acquisitions stricts (Boellaard et al., 2008)) suggèrent aussi de nombreuses alternatives, afin de permettre d'accumuler des preuves pour affiner les choix des métriques et de seuils proposés dans une prochaine version du protocole. On notera par exemple qu'il est conseillé de rapporter différentes mesures du SUL (eg : maximal ou moyen dans un VOI segmenté selon différents critères) ou encore de relever la valeur du TLG à des fins exploratoires.

II.3.3. Autres méthodes d'analyses des images TEP

L'interprétation quantitative de l'information métabolique contenue dans ces examens repose, dans le meilleur des cas, sur la mesure du SUV maximal ou moyen d'une population de voxels définis comme caractéristiques de la lésion étudiée. De nombreuses méthodes ont été proposées dans la littérature pour prendre en compte d'autres informations contenues dans les images TEP. Les premières méthodes, décrites au paragraphe II.3.3.1, ont été développées pour caractériser l'hétérogénéité de la distribution du radiotracer dans la tumeur. Des études sur petits animaux ont en effet montré que la distribution de FDG dans les tissus tumoraux est corrélée aux caractéristiques du microenvironnement tumoral (eg : (Pugachev et al., 2005)) et donc à des informations d'intérêt pour le suivi thérapeutique.

Outre ces approches, qui consistent à évaluer l'évolution d'une information globale caractéristique de la distribution du FDG dans la tumeur à un instant donné, certains auteurs ont développé des méthodes permettant de décrire directement l'évolution tumorale par une analyse de la réponse thérapeutique à l'échelle du voxel. Ces techniques seront décrites au paragraphe II.3.3.2.

II.3.3.1. Méthodes de mesure d'hétérogénéité

L'état de l'art que nous présentons se base en partie sur la revue récente de (Asselin et al., 2012) relative à la mesure d'hétérogénéité dans les tumeurs humaines vue par la TEP et l'IRM. Nous limitons ici notre propos à l'imagerie TEP, et excluons les techniques relatives à

la TEP dynamique (eg : mesure de la dimension fractale des courbes activité-temps calculées à partir des examens TEP pour décrire l'hétérogénéité de la cinétique de fixation du radiotracer (Dimitrakopoulou-Strauss et al., 2001a; Dimitrakopoulou-Strauss et al., 2001b)).

II.3.3.1.1. Analyse visuelle

L'inclusion de paramètres de forme et de texture à des fins diagnostiques en TEP/TDM a été rapportée pour la première fois par (Miller et al., 2003) pour des cas de cancers du col de l'utérus. Elle consistait alors à cumuler un score selon différents critères tels que : la taille de la tumeur par rapport aux dimensions du pelvis, sa sphéricité (0 pour une tumeur sphérique, 1 pour une tumeur non sphérique, la plus grande valeur ayant un moins bon pronostic) et son hétérogénéité (de nulle à marquée selon que le SUV à l'intérieur du volume tumoral varie de 0 à 30%).

Plus tard, (Kalff et al., 2006) ont proposé des critères basés sur la reconnaissance de forme de la fixation tumorale de FDG pour caractériser la réponse thérapeutique de carcinomes rectaux. En se basant sur les travaux de (Singnurkar et al., 2010) étudiant notamment l'intérêt de la TEP/TDM pour l'évaluation de la réponse après l'ablation par radiofréquence de lésions pulmonaires en prenant en compte différentes formes de fixation de FDG, (Hofman and Hicks, 2010) ont discuté l'intérêt de ces approches par rapport à celles basées seulement sur le SUV, comme proposées par (Wahl et al., 2009) dans les critères PERCIST par exemple.

II.3.3.1.2. Utilisation de régions d'intérêt et analyse de forme

En partant du constat que la mesure du SUV_{max} mesurée à partir de la TEP au FDG de patients atteints de sarcomes était un facteur pronostic indépendant de la mort du patient et qu'il était plus fiable que les données habituellement utilisées en clinique telles que la biopsie prétraitement ou l'histologie post-opératoire (Eary et al., 2002), (O'Sullivan et al., 2005) ont proposé une méthode d'analyse de la distribution de la fixation intra tumorale afin de déterminer la progression de la maladie. En bref, elle consiste à déterminer l'écart entre la fixation de FDG effective de la tumeur et celle d'une tumeur de forme ellipsoïdale dont l'activité serait maximale au centre et d'autant plus faible qu'on s'en éloignerait. Les paramètres quantifiant l'hétérogénéité de la tumeur réelle sont déduits de ces écarts. L'application de cette analyse (Eary et al., 2008) aux données TEP de 234 patients a montré que cette méthode fournissait un facteur pronostic indépendant prédictif de la survie sans progression et de la survie globale du patient.

Sans aller jusqu'à cette généralisation de la mesure de l'excentricité de la tumeur, (El Naqa et al., 2009) ont illustré, sur 14 cas de cancers du col de l'utérus et 9 patients atteints de cancers de la sphère ORL, l'utilisation de grandeurs relatives à la forme de la tumeur (segmentées sur la TDM ou sur la TEP) pour prédire la réponse au traitement. On retiendra l'utilisation:

- De l'excentricité EX, définie par :

$$EX = 1 - \sqrt{\frac{a^2}{b^2}} \quad \text{Eq. II.5}$$

où a (respectivement b) est la valeur du demi-grand axe (respectivement petit axe) de l'ajustement du volume métabolique par un ellipsoïde.

- Du nombre d'Euler qui correspond à la différence entre le nombre d'amas de voxels (i.e. de voxel connexes) caractéristiques de la tumeur et le nombre de zones hypofixantes (ou trous) considérées comme des zones non actives et contenues dans le volume métabolique de la tumeur.
- De la solidité, qui mesure la convexité de la structure. Cette grandeur est définie par la proportion des voxels appartenant à l'enveloppe convexe du volume métabolique de la tumeur qui sont également présents dans le volume métabolique de cette tumeur.
- L'étendue : similaire à la solidité, cette grandeur représente la proportion de voxels appartenant au plus petit parallélépipède rectangle englobant le volume métabolique de la tumeur qui sont également présents dans le volume métabolique de cette tumeur.

Aucune application de ces métriques n'a cependant été rapportée par la suite dans la littérature.

II.3.3.1.3. Analyse d'histogrammes

Deux approches, se basant sur l'étude des niveaux d'intensités des SUV contenus dans les volumes tumoraux segmentés, sont décrites dans la littérature.

- Caractéristiques de premier ordre

La première approche, très utilisée en traitement d'image, consiste à extraire différentes statistiques directement issues de la répartition des valeurs de fixations normalisées de FDG dans la tumeur. On utilise, pour les calculer, un histogramme d'intensité représentatif de la distribution des valeurs des SUV des voxels contenus dans le volume segmenté étudié. Il est pour cela nécessaire de ré-échantillonner les niveaux de fixations de FDG dans la tumeur i.e. de les regrouper par valeurs de SUV proches, compris dans des intervalles de valeurs donnés. Ce ré-échantillonnage des valeurs permet de réaffecter un niveau de gris $v = V(i)$, avec $v \in [1, G]$ à chaque voxel i . Un exemple de fonction en escalier V utilisée pour ce ré-échantillonnage est donné à l'équation V.1. Une fois les valeurs de niveaux de gris attribuées à chaque voxel i , on définit une densité de probabilité $P(v)$ telle que :

$$P(v) = \frac{\text{Card}(\{i/V(i) = v\})}{n} \quad \text{Eq. II.6}$$

où Card correspond au nombre de voxels dont le niveau de gris attribué suite à l'application de la fonction de ré-échantillonnage V est égale à $V(i)$ et n est le nombre total de voxels représentatifs de la tumeur sur l'image TEP.

Cette méthode décrit les caractéristiques de premier ordre et englobe les grandeurs telles que :

- Les bornes supérieures et inférieures de la distribution.
- Ses quantiles caractéristiques (eg : 1^{er} et 3^{ème} quartiles ou encore la médiane).
- Son énergie, définie par :

$$Energie = \sum_{v=1}^G P(v)^2 \quad \text{Eq. II.7}$$

- Son entropie.

$$Entropie = \sum_{v=1}^G -P(v) \times \ln(P(v)) \quad \text{Eq. II.8}$$

- Différents moments centrés réduits d'ordre N , notés m_N (cf : équation II.9), qui peuvent correspondre, en fonction de la valeur de N considérée, à : la moyenne ($N=1$), la variance ($N=2$, caractéristique de la dispersion des valeurs autour de la moyenne), l'asymétrie (ou *skewness*, $N=3$), le coefficient d'aplatissement (*kurtosis*) (obtenu pour $N=4$), caractérisant la distribution.

$$m_N = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^N}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{N/2}} \quad \text{Eq. II.9}$$

où ici, x_i est une variable aléatoire continue telle que:

- x_i correspond à la valeur du SUV de chacun des n voxels i contenus dans le volume de la tumeur segmentée.
- \bar{x} est le SUV moyen dans ce même volume.

Cette approche a par exemple été mise en œuvre par (Tixier et al., 2011) dans le contexte de la prédiction de la réponse à partir des données TEP au FDG acquises avant le début du traitement de 41 patients atteints de cancer de l'œsophage. Aucun de ces index n'a cependant fourni de résultat encourageant dans ce contexte. De même, dans une étude visant à comparer

la variabilité de différents index caractérisant l'hétérogénéité, dont ceux présentés ci-dessus, (Galavis et al., 2010) ont conclu que seule l'énergie issue de ces caractéristiques au premier ordre avait un niveau de variabilité faible (<5%) en fonction des paramètres de reconstruction. Nous reviendrons en détail sur ces deux résultats dans la discussion de ce chapitre.

- Histogrammes Intensité Volume

En se basant sur l'Histogramme Dose Volume (ou DVH, (Drzymala et al., 1991; Mohan et al., 1987)) utilisé en radiothérapie pour caractériser les niveaux de dose reçus par les volumes cibles et les organes à risques, (El Naqa et al., 2009) ont décrit une méthode graphique qui permet de quantifier l'hétérogénéité tumorale : l'Histogramme Intensité Volume (IVH). Cet histogramme caractérise la distribution d'activité dans la tumeur en traçant le pourcentage du volume segmenté total (en ordonnée) dont le SUV est au moins égal à une valeur donnée en pourcentage du SUV_{max} (en abscisse). Un exemple de ce type de graphique est présenté en figure II.1.

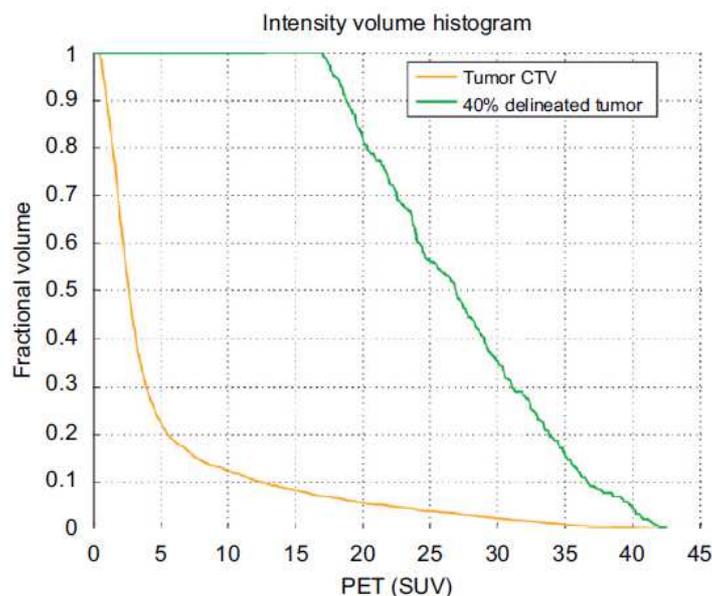


Figure II.1 : Histogrammes Intensité-Volume (IVH) d'une tumeur du col de l'utérus du Volume Tumoral Clinique (CTV, utilisé pour déterminer le plan de traitement en radiothérapie, courbe marron) et du volume segmenté sur la TEP grâce à un iso contour égal à 40% du SUV maximal de la tumeur (courbe verte). Extrait de (El Naqa et al., 2009).

A partir de ce graphique, (El Naqa et al., 2009) ont proposé d'extraire plusieurs métriques telles que :

- $I_{x\%}$: le SUV minimal correspondant à x% du volume.

- $V_{x\%}$: le pourcentage du volume étudié dont le SUV est au moins égale à X% du SUV_{max} de la tumeur.
- Des soustractions entre des valeurs particulières de $I_{x\%}$ (respectivement $V_{x\%}$), afin de rendre compte de l'hétérogénéité de la distribution de radiotracer dans le volume étudié. Les auteurs rapportent ainsi :
 - $I_{10-90} = I_{10\%} - I_{90\%}$
 - $V_{10-90} = V_{10\%} - V_{90\%}$

Parmi les différentes méthodes étudiées par (El Naqa et al., 2009), le calcul des IVH sur les TEP prétraitement de 9 patients atteints de cancers de la sphère ORL a permis de montrer que $V_{90\%}$ prédisait mieux la réponse à la radio-chimiothérapie à long terme (aire sous la courbe (ou AUC pour Area Under the Curve) de la ROC³ (*Receiver Operating Characteristic*) égale à 0,95) que les méthodes classiques tel que le SUV_{max} (AUC = 0,7). On notera cependant, sur cette même cohorte, les très bonnes performances du volume métabolique (AUC = 0,875).

Toujours dans cette étude, les métriques dérivées de l'IVH appliquées à une seconde cohorte constituée de 14 patientes atteintes de cancers du col de l'utérus ont conduit à des performances légèrement supérieures à celles, très faibles, du SUV ou du volume métabolique par exemple. Par exemple, V_{10-90} a conduit à une AUC de 0,65 contre 0,46 pour SUV_{max} et 0,57 pour le volume métabolique par exemple. Quelles que soient les cohortes, le nombre de patients considérés n'est pas suffisant pour dégager des différences significatives.

Pour évaluer l'apport de l'information d'hétérogénéité de la distribution de radiotracer caractérisée par l'AUC de l'IVH (AUC-IVH) par rapport au SUV_{max} , (van Velden et al., 2011) ont mis en œuvre cette méthode sur des données de patients atteints de cancers du poumon non à petites cellules (NSCLC, *Non Small Cell Lung Cancer*). Pour réintégrer les informations d'activité et/ou de volume, importantes dans le contexte du suivi thérapeutique, à la valeur d'AUC-IVH, (van Velden et al., 2011) proposent trois nouvelles définitions des IVH. Elles consistent à normaliser ou non les axes par le SUV_{max} et/ou le volume de la tumeur, respectivement. Dans ces cas, une renormalisation des valeurs obtenues est nécessaire. Ces métriques ont fourni des résultats encourageants dans le contexte du suivi thérapeutique, en particulier lorsque les images étaient préalablement corrigées de l'EVP par la méthode de Van Cittert (cf équation III.10) et débruitées. Dans leur revue sur les méthodes de quantification de l'hétérogénéité en TEP et IRM, (Asselin et al., 2012) soulignent cependant que même si ces métriques caractérisent bien des fixations en formes d'anneaux, elles apparaissent moins sensibles à l'hétérogénéité de la distribution du radiotracer que l'inverse du coefficient de variation (i.e. la moyenne divisée par l'écart-type).

En représentant la valeur absolue du volume en ordonnée en fonction du seuil, (en pourcentage du SUV_{max}) utilisé pour le segmenter (cf figure II.2), (Kidd and Grigsby, 2008)

³ La courbe ROC est un outil statistique utilisé pour évaluer les performances d'un classificateur binaire. Ces courbes sont utilisées en médecine pour évaluer les capacités prédictives d'une méthode de diagnostic par exemple. Une explication détaillée de cette méthode d'analyse est fournie au paragraphe IV.2.1.4.1.

ont estimé un paramètre d'hétérogénéité H caractéristique de la distribution de FDG dans la tumeur.

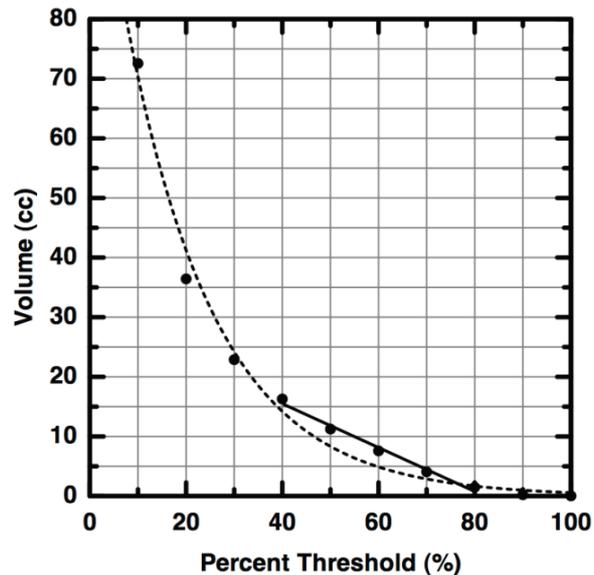


Figure II.2: courbe du volume en fonction du seuil exprimé en pourcentage du SUV_{max} , utilisé pour le segmenter. Les auteurs utilisent la pente de la droite d'ajustement (droite en trait plein), tracée à partir des points obtenus pour un seuil compris entre 40% et 80% du SUV_{max} de la tumeur, pour caractériser l'hétérogénéité de la distribution de FDG dans la lésion. La courbe en pointillés correspond à l'ajustement, par une fonction exponentielle décroissante, de l'ensemble des points (extrait de (Brooks and Grigsby, 2011)).

Le paramètre H est obtenu en effectuant une régression linéaire des points compris entre 40% et 80% (cf figure II.2) du SUV maximum, avec comme postulat qu'entre ces bornes, le volume décroît sous la forme :

$$V(mL) = V_0 - H \times S \quad \text{Eq. II.10}$$

avec V_0 constante, S le seuil (en % du SUV maximal dans la tumeur) et H (dV/dT) l'hétérogénéité de la tumeur. La pente de la droite de régression fournit la valeur d'intérêt. Les auteurs ont mesuré cet index sur les TEP prétraitement de 72 patientes atteintes de cancers du col de l'utérus. Ils ont montré que l'index était très corrélé au volume métabolique segmenté à partir d'une méthode semi-automatique basée sur un seuil égal à 40% du SUV_{max} ($R^2=0,88$). La valeur de H n'était cependant pas corrélée au type histologique de la lésion (valeur de $p = 0,5$) mais était associée à la probabilité de développer des métastases ganglionnaires (valeur de $p = 0,0009$).

La méthode proposée par (Kidd and Grigsby, 2008) a fait l'objet, trois ans après sa présentation, d'un rectificatif important. En effet, Grigsby et al. ont réexaminé les données et observé que dans leur cohorte de lésions, hormis pour des volumes supérieurs à 150 mL, le volume métabolique obtenu en utilisant un seuil égal à 80% du SUV_{max} était très petit (d'environ de 3 mL). En partant de ce constat, les auteurs ont montré que les derniers points considérés pour effectuer la régression linéaire nécessaire au calcul de H (figure II.2 et équation II.6) étaient indépendant de la tumeur i.e. que la pente était approximativement proportionnelle au volume de la tumeur et donc au volume obtenu en utilisant un seuil égal à 40% du SUV_{max} (Brooks and Grigsby, 2011). Seules les fluctuations statistiques expliquent alors que deux tumeurs de même volume aient une valeur de H différente. Malgré cette démonstration, (Huang et al., 2012) ont appliqué cette méthode sur les données de 40 patients atteints de carcinomes du nasopharynx. Leurs résultats leur ont permis de conclure que la mesure de l'hétérogénéité obtenue en calculant H était utile pour caractériser ce type de tumeur et prédire la survie des patients. On notera que Huang et al. avaient pourtant relevé une corrélation très élevée entre H et le volume métabolique ($R^2 = 0,97$), résultat qui les a amené à conclure que l'hétérogénéité tumorale était davantage liée aux dimensions de la tumeur qu'au SUV_{max} .

On retiendra que la caractérisation de l'hétérogénéité par les histogrammes présentés ci-dessus ignore toute information spatiale et notamment celle portant sur les relations entre les différents niveaux d'activité métabolique contenus à l'intérieur du volume tumoral (Brooks and Grigsby, 2011).

II.3.3.1.4. Analyse de texture

L'étude de la texture a souvent été utilisée pour l'analyse quantitative d'images médicales dans le but de distinguer des régions saines et pathologiques ou pour segmenter les volumes tumoraux, que ce soit en TDM (Tesar et al., 2008; Xu et al., 2006), en mammographie (Bellotti et al., 2006; Szekely et al., 2006) ou en IRM (Assefa et al., 2010; Sinha et al., 1997). Ces méthodes permettent entre autre de réintroduire l'information spatiale dans l'analyse des niveaux de gris contenus dans le volume d'intérêt.

Différentes techniques ont été décrites dans la littérature : celles basées sur des analyses statistiques d'une part (c'est le cas des grandeurs calculées à partir du contenu des matrices de cooccurrence (Haralick et al., 1973), qui permettent de déterminer les caractéristiques de second ordre) et celles basées sur les décompositions en ondelettes, initiées par (Unser, 1995).

La première application par (El Naqa et al., 2009) en TEP, dans le contexte du suivi thérapeutique, s'est limitée à la première série de méthodes (cf tableau II.9, colonnes 3 et 4). La matrice de cooccurrence, alors mise en œuvre et adaptée à des données à 3 dimensions, permet de mesurer la probabilité de trouver des voxels voisins d'intensité donnée dans toutes les directions de l'espace. Une fois calculée, cette matrice donne accès à différentes grandeurs locales telles que l'entropie, l'énergie, le contraste ou encore l'homogénéité. Pour illustrer l'intérêt des index dérivés de la matrice de cooccurrence, (El Naqa et al., 2009) ont comparé

des méthodes classiques telles que le SUV_{max} ou le SUV_{moyen} à celles d'*Homogeneity* et *Contrast*, dont les définitions mathématiques sont données au paragraphe. La comparaison de ces index portait sur les AUC des courbes ROC représentatives des capacités prédictives de ces méthodes à distinguer la persistance post-radiothérapie de la maladie chez un groupe de 14 patientes atteintes de cancers du col de l'utérus (colonne 3, tableau II.9) et la survie globale chez 9 patients atteints de cancers ORL (colonne 4, tableau II.9).

Les résultats exploratoires encourageants de (El Naqa et al., 2009) sur la capacité prédictive de ces index ont poussé d'autres équipes à évaluer l'intérêt de ces méthodes. Ainsi, (Galavis et al., 2010) ont étudié la variabilité, pour différents paramètres de reconstructions, des index dérivés des matrices de cooccurrence mais aussi de ceux relatifs aux caractéristiques d'ordres supérieurs tels que :

- La longueur de séries homogènes (Tang, 1998).
- La dépendance à l'intensité des voxels directement voisins (Sun and Wee, 1983).
- Les différences de niveaux de gris (Amadasun and King, 1989).

La robustesse des index de texture a été étudiée en fonction du type d'acquisition des données (2D ou 3D), de la dimension des voxels ou du type de reconstruction (nombre d'itérations OS-EM, ou valeur de la Largeur à Mi-Hauteur (LMH) du post-filtrage gaussien). Les bornes de l'écart à la moyenne des valeurs obtenues pour chacun des 50 index étudiés sur les données de 20 patients ont permis de mettre en évidence les limites de certains d'entre eux, soulignant ainsi la difficulté d'utiliser ces index pour des comparaisons multicentriques.

La colonne 5 du tableau II.9 donne la valeur absolue de l'écart à la moyenne le plus élevé présentée dans les travaux de Galavis et al. Les valeurs numériques données dans ces tableaux ont été estimées par numérisation des histogrammes présentés dans cet article. La méthode de numérisation employée implique une erreur d'estimation des valeurs de l'ordre de 3-5% environ.

Plus tard, mais sans référencer les résultats de Galavis et al., (Tixier et al., 2011) ont évalué les capacités de ces méthodes à prédire la réponse à la thérapie en utilisant les TEP statiques prétraitement de 41 patients atteints de cancer de l'œsophage. La comparaison portait sur une partie des index étudiés par (Galavis et al., 2010). Elle incluait aussi des statistiques issues de la matrice des longueurs de zones homogènes (proposées par (Thibault et al., 2011)) et excluait les index relatifs à la dépendance à l'intensité des voxels directement voisins (Sun and Wee, 1983) (cf tableau II.9, colonnes 6 et 7). Les résultats de l'analyse ROC visant à établir la capacité des index métaboliques à prédire la réponse tardive des tumeurs vue par RECIST 1.0 ont montré que les index classiques (i.e. dérivés du SUV) séparaient moins bien les tumeurs répondant des tumeurs ne répondant pas à la thérapie ($AUC < 0,6$) que les index dérivés de l'analyse de texture ($AUC > 0,82$). Le manque de détails sur les performances des différentes méthodes évaluées ne permet cependant pas de connaître au cas par cas leurs capacités prédictives.

Type de texture étudiée	Noms des index de texture	El Naqa et al. 2009		Galaxis et al. 2010	Tixier et al. 2011		Tixier et al. 2012
		Col	ORL		Sensibilité/Spécificité/AUC MS+MP vs RP+RC	Sensibilité/Spécificité/AUC MS+MP+RP vs RC	
Caractéristique d'ordre 2 (matrice de Coocurrence)	Homogeneity	0,62	0,83	36%	88%/73%/>0,82	92%/56%/?	1,8±11,5%
	Energy	0,73	0,50	2%	?/?>0,82	?/??	10,9±26,4%
	Correlation	-	-	50%	?/?>0,82	?/??	-0,6±27,7%
	Contrast (Inertia)	0,60	0,80	109%	?/?>0,82	?/??	5,4±24,0%
	Entropy	0,65	0,53	>10%	79%/82%/>0,82	92%/69%/?	-2,0±5,4%
	Dissimilarity	-	-	65%	?/?>0,82	?/??	-2,1±13,0%
	Sum of squares	-	-	167%	-	-	-
	Sum average	-	-	113%	-	-	-
	Sum variance	-	-	143%	-	-	-
	Sum entropy	-	-	14%	-	-	-
	Difference variance	-	-	90%	-	-	-
	Difference entropy	-	-	18%	-	-	-
	Information measure of correlation	-	-	27%	-	-	-
	Maximal correlation coefficient (1)	-	-	4%	-	-	-
	Maximal correlation coefficient (2)	-	-	146%	-	-	-
	Maximal probability	-	-	53%	-	-	-
	Diagonal moment	-	-	200%	-	-	-
	Difference energy	-	-	58%	-	-	-
	Inertia	-	-	109%	-	-	-
	Inverse difference moment	-	-	65%	-	-	-
	Sum energy	-	-	52%	-	-	-
	Cluster shade	-	-	200%	-	-	-
	Cluster prominence	-	-	200%	-	-	-
	GrayLevel Run Length (GLRL)	Short Run Emphasis (SRE)	-	-	161%	?/?>0,82	?/??
Long Run Emphasis (LRE)		-	-	181%	?/?>0,82	?/??	-
Low Gray-level Run Emphasis (LGRE)		-	-	4%	?/?>0,82	?/??	-
High Gray-level Run Emphasis (HGRE)		-	-	11%	?/?>0,82	?/??	-
Short Run Low Gray-level Emphasis (SRLGE)		-	-	156%	?/?>0,82	?/??	-
Short Run High Gray-level Emphasis (SRHGE)		-	-	173%	?/?>0,82	?/??	-
Long Run Low Gray-level Emphasis (LRLGE)		-	-	180%	?/?>0,82	?/??	-
Long Run High Gray-level Emphasis (LRHGE)		-	-	181%	?/?>0,82	?/??	-
Gray-Level Non-Uniformity run (GLNUr)		-	-	11%	76%/91%/0,89	85%/75%/?	-
Run Length Non-Uniformity (RLNU)		-	-	44%	?/?>0,82	?/??	-
Run Percentage (RP)		-	-	31%	?/?>0,82	?/??	-

Neighboring Gray Level (NGL)	Small number emphasis	-	-	14%	-	-	-
	Large number emphasis	-	-	119%	-	-	-
	Number nonuniformity	-	-	39%	-	-	-
	Second moment	-	-	29%	-	-	-
	Entropy-NGL	-	-	9%	-	-	-
Tone Difference (NCTD)	Coarseness	-	-	69%	?/?>0,82	?/?/?	-
	Contrast	-	-	169%	?/?>0,82	?/?/?	-
	Busyness	-	-	64%	?/?>0,82	?/?/?	-
Intensity-Size-Zone	Short Zone Emphasis (SZE)	-	-	-	?/?>0,82	?/?/?	-6,0±54,3%
	Long Zone Emphasis (LZE)	-	-	-	?/?>0,82	?/?/?	3,6±30,0%
	Low Gray-level Zone Emphasis (LGZE)	-	-	-	?/?>0,82	?/?/?	-4,0±55,3%
	High Gray-level Zone Emphasis (HGZE)	-	-	-	?/?>0,82	?/?/?	3,9±20,4%
	Short Zone Low Gray-level run Emphasis (SZLGE)	-	-	-	?/?>0,82	?/?/?	-7,0±67,6%
	Short Zone High Gray-level Emphasis (SZHGE)	-	-	-	?/?>0,82	?/?/?	1,0±31,2%
	Long Zone Low Gray-level Emphasis (LZLGE)	-	-	-	?/?>0,82	?/?/?	1,8±28,9%
	Long Zone High Gray-level Emphasis (LZHGE)	-	-	-	?/?>0,82	?/?/?	3,5±35,8%
	Gray-Level Non-Uniformity zone (GLNUz)	-	-	-	?/?>0,82	?/?/?	-9,7±24,0%
	Zone Length Non-Uniformity (ZLNU)	-	-	-	76%/91%/>0,82	92%/69%/0,85	11,2±23,1%
	Zone Percentage (ZP)	-	-	-	?/?>0,82	?/?/?	-2,7±16,9%

Tableau II.9 : résultats publiés concernant l'utilisation des méthodes d'analyse de texture sur des examens TEP. Les noms des index utilisés ici (colonne 2) correspondent à ceux que nous avons retenus lors de nos travaux présentés au chapitre V ou ceux choisis par les auteurs quand ceux-ci étaient présentés dans une seule publication. Les résultats donnés pour (El Naqa et al., 2009) (colonnes 3 et 4) sont des valeurs d'AUC des courbes ROC pour les deux cohortes de patients étudiées. Ceux de (Galavis et al., 2010) (colonne 5) correspondent à la valeur absolue la plus élevée de l'écart à la moyenne des valeurs des différents index obtenus sur des images reconstruites en utilisant différents paramètres de reconstruction. Ces valeurs sont déduites des histogrammes présentés dans l'article. Les colonnes 6 et 7 représentent respectivement les Spécificité, Sensibilité et AUC des courbes ROC des index de texture pour identifier un groupe de tumeurs de l'œsophage répondant à la thérapie d'un groupe ne répondant pas (colonne 6) et pour séparer un groupe répondant complètement par rapport à un autre groupe, constitué d'un tout autre type de lésions (colonne 7). La dernière colonne représente l'écart relatif entre deux examens TEP acquis à des intervalles de temps très courts, moyenné sur plusieurs lésions (± 1 écart-type). - : index non étudié dans la publication. ? : valeur non fournie dans l'étude.

Afin d'évaluer la reproductibilité de ces index, (Tixier et al., 2012) ont enfin comparé, pour 16 patients atteints de cancers de l'œsophage, les variations entre deux TEP acquises à moins de 4 jours d'intervalle, avant le début du traitement. Les index étudiés (cf tableau II.9, colonne 8) n'incluaient cependant plus ceux obtenus à partir des matrices des différences de niveaux de gris, ni ceux issus des longueurs de séries homogènes, du fait d'une corrélation élevée entre les index dérivés de ces matrices. La reproductibilité de certains de ces paramètres, ainsi que leur variabilité en fonction du paramètre de discrétisation ont permis de valider l'utilisation clinique de ces grandeurs et leur intérêt pour le suivi thérapeutique.

Nous reviendrons en détails sur les résultats des articles de (Galavis et al., 2010; Tixier et al., 2012; Tixier et al., 2011) dans le chapitre V. Les formules et définitions détaillées des différents index étudiés par (Tixier et al., 2011) seront précisées. Etant donné l'importance de ces résultats pour la suite de notre exposé, nous avons résumé les valeurs de performance des index classiques et des index issus des caractéristiques de premier ordre présentées dans le tableau II.10.

Noms de l'index issus de la caractéristique de premier ordre	EL Naqa et al. 2009		Galavis et al. 2010	Tixier et al. 2011		Tixier et al. 2012
	Col	ORL		Sensibilité/Spécificité/AUC MS+MP vs RP+RC	Sensibilité/Spécificité/AUC MS+MP+RP vs RC	
SUV maximum	0,46	0,7	-	53%/73%/0,59	46%/91%/0,7	4,7±19,5%
SUV minimum	0,46	0,6	-	?/?/<0,6	?/?/?	6,3±26,5%
SUV moyen	0,46	0,65	56%	71%/45%/<0,6	62%/81%/?	5,5±21,2%
SUV médian	-	-	86%	?/?/<0,6	?/?/?	-
Variance	-	-	95%	?/?/?	?/?/?	-
Ecart-type	0,53	0,675	-	?/?/?	?/?/?	1,2±23,2%
Coefficient de variation	-	-	38%	-	-	-
1/Coefficient de variation	-	-	-	-	-	4,1±24,1%
Skewness	-	-	200%	-	-	-0,3±27,5%
Kurtosis	-	-	200%	-	-	2,1±18,0%
Energie	-	-	2%	-	-	-
Entropie	-	-	0	-	-	-

Tableau II.10 : résultats des index classiques et issus des caractéristiques de premier ordre utilisées dans les articles de (El Naqa et al., 2009; Galavis et al., 2010; Tixier et al., 2012; Tixier et al., 2011). Les résultats donnés pour (El Naqa et al., 2009) (colonne 2 et 3) sont des valeurs d'AUC des courbes ROC pour les 2 cohortes de patients étudiées. Ceux de (Galavis et al., 2010) (colonne 4) correspondent à la valeur absolue la plus élevée de l'écart à la moyenne des valeurs des différents index obtenus sur des images reconstruites en utilisant différents paramètres de reconstruction. Ces valeurs sont déduites des histogrammes présentés dans l'article. Les colonnes 5 et 6 représentent les Spécificité, Sensibilité et AUC des courbes ROC des index, pour distinguer un groupe de tumeurs de l'œsophage répondant à la thérapie d'un groupe ne répondant pas (colonne 5) et pour séparer un groupe répondant complètement par rapport à un autre groupe, constitué d'un tout autre type de lésions (colonne 6). La dernière colonne rapporte l'écart relatif entre deux examens TEP acquis à des intervalles de temps très courts, moyenné sur plusieurs lésions (± 1 écart-type). - : index non étudié dans la publication.

II.3.3.2. Méthodes de détection des changements métaboliques à l'échelle du voxel

Pour décrire l'évolution du volume métabolique actif des tumeurs, certains auteurs ont proposé de décrire l'évolution tumorale non plus à partir d'une information moyenne mais en utilisant toute l'information de l'image via une analyse à l'échelle du voxel.

II.3.3.2.1. Imagerie paramétrique

Récemment, (Necib et al., 2011) ont proposé une méthode d'imagerie paramétrique (IP) pour décrire objectivement l'apparition de zones d'hétérogénéité ou de nécroses dans les tissus tumoraux entre 2 examens TEP successifs. Cette méthode consiste à analyser l'évolution des voxels de la tumeur entre des examens réalisés à plusieurs semaines d'intervalles et à identifier ainsi les zones dans lesquelles l'activité métabolique a varié significativement entre les examens. Elle permet d'aboutir à une image 3D paramétrique qui décrit la réponse de la maladie à la thérapie et à des paramètres décrivant les modifications métaboliques relatives aux tumeurs.

Les étapes et la théorie sous-jacentes sont les suivantes :

- i. *Recalages des volumes TEP* : pour suivre l'évolution de la tumeur entre 2 examens TEP successifs (notés respectivement $TEP(T_1)$ et $TEP(T_2)$, figure II.3A et II.3B) à l'échelle du voxel, il est nécessaire de recalcr les 2 volumes d'images, de sorte qu'un voxel représente la même portion de l'espace dans les 2 examens. Du fait de la médiocre résolution spatiale des images TEP et du grand nombre de points de repères disponibles dans les images TDM associées à ces TEP, les volumes TEP sont recalés à partir de recalages TDM/TDM. Les auteurs ont pour cela favorisé une transformation rigide par bloc (Ourselin et al., 2000) plutôt qu'un recalage élastique, qui pourrait déformer le volume étudié et donc biaiser l'analyse voxel à voxel.
- ii. *Calcul de l'image paramétrique* : une fois recalée, les images TEP sont soustraites afin d'obtenir une image de différence DIF, en unité SUV, telle que:

$$DIF = TEP(T_2) - TEP(T_1) \quad \text{Eq. II.11}$$

Après avoir défini un VOI_{IP} contenant la ou les tumeurs étudiées, un graphique biparamétrique est calculé afin d'identifier les voxels dont la fixation a évolué entre les deux examens. L'abscisse et l'ordonnée de ce graphique correspondent respectivement, pour chaque voxel du VOI_{IP} , au SUV dans l'examen $TEP(T_1)$ et à l'évolution du SUV dans l'image des différences DIF.

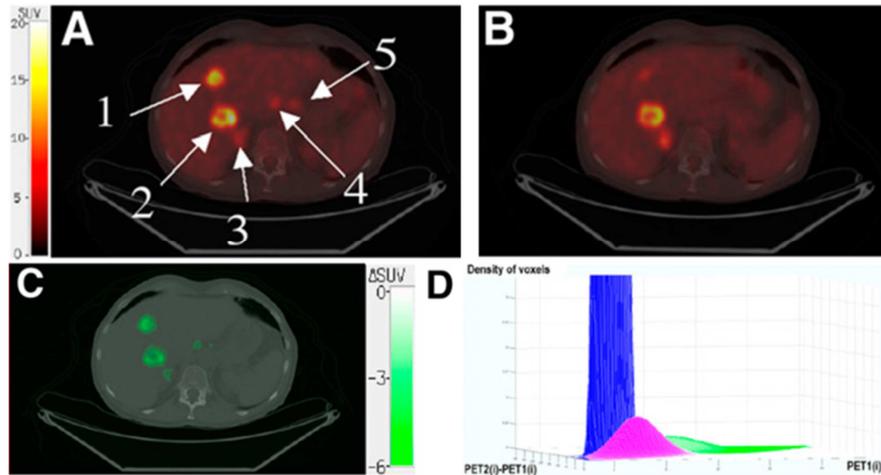


Figure II.3 : A : coupe transverse de la TEP/TDM (notée $TEP(T_1)$) acquise avant le début du traitement d'un patient atteint d'un cancer colorectal métastatique. Les flèches montrent 5 lésions métaboliquement actives. B : coupe transverse de la TEP/TDM acquise après 1 cycle de chimiothérapie ($TEP(2)$). C : image paramétrique obtenue suite à l'application de la méthode décrite par (Necib et al. 2011). D : graphique de soustraction avec identification de 3 composantes : le bruit correspond à l'amas bleu, la captation physiologique (relative au foie) est représentée en rose et l'évolution tumorale est représentée en vert. Extrait de (Necib et al. 2011).

A partir de ce graphique, les auteurs identifient les voxels ayant évolué en utilisant un modèle de mélanges gaussiens (ou GMM pour *Gaussian Mixture Model* en anglais). Cette méthode permet d'estimer, par une approche paramétrique, la distribution de variables aléatoires étudiées en les modélisant sous la forme d'une somme de K gaussiennes (figure II.3D), chacune étant caractérisée par une moyenne μ_k et un écart-type σ_k . L'identification des K gaussiennes se fait de façon itérative. Au final, on attribue à chaque voxel une probabilité $t_k(x_i)$ (où x_i correspondent aux coordonnées du voxel i dans l'espace $(TEP_1(i), DIF(i))$ d'appartenir à l'amas gaussien k . Un maximum de $K=4$ clusters est admis :

- 1 correspondant à un groupe de voxels dont l'activité n'évolue pas (assimilable au bruit).
- 1 agrégat décrivant les modifications physiologique de la fixation non liées à l'évolution de la lésion.
- 1 groupe de voxels dont l'activité décroît entre les deux examens.
- 1 groupe de voxels dont l'activité croît entre les deux examens.

iii. *Calcul d'index à partir des images paramétriques.* A partir des agrégats de voxels obtenus, (Necib et al., 2011) définissent deux index permettant de refléter l'information contenue dans les images paramétriques (figure II.3C) :

- Le premier correspond au volume tumoral dont l'activité a évolué entre les 2 examens. Il est défini par le plus grand volume de voxels connexes de chaque amas.

- Le second correspond à la moyenne des valeurs dans DIF de ces mêmes voxels.

Ces grandeurs permettent de définir la réponse de chaque tumeur : RC, RP ou MP. Il n'est en effet pas possible de définir une maladie stable par cette méthode puisque l'amas relatif aux voxels d'intensité fixe dans le temps est assimilé au bruit dans l'image.

Cette méthode a été évaluée sur une population de 28 patients atteints de cancers colorectaux métastatiques. Les résultats présentés par (Necib et al., 2011) montrent que l'IP donne des résultats au moins aussi bons que le protocole EORTC en terme de sensibilité et de spécificité de la détection de modification du métabolisme tumoral. De même, cette méthode apparaît comme étant plus sensible que les critères RECIST aux changements dus au traitement.

Si elle permet la prise en compte d'une plus grande partie de l'information disponible dans les examens TEP, cette méthode ne doit pas être utilisée sans une étude conjointe des deux acquisitions TEP initiales. L'IP ne permet pas de distinguer une réponse partielle de la tumeur à une réponse complète uniquement à partir des index estimé. De même, elle ne permet pas d'identifier les maladies stables sur les images paramétriques. Autrement dit, une tumeur qui n'évolue ni en volume ni en activité ne sera pas représentée sur des images paramétriques. De même, si une tumeur apparaît dans une zone située à l'extérieure du VOI_{IP} , la méthode ne permet pas de détecter la progression de la maladie. On notera enfin que les auteurs ont montré une bonne robustesse de leur méthode aux erreurs potentielles de recalage. L'étude de la variabilité des résultats en fonction des différentes opérations opérateur-dépendantes, tel le choix des voxels inclus dans le VOI_{PI} , serait un élément de discussion important.

II.3.3.2.2. Théorie du raisonnement paradoxal

Une méthode similaire à la précédente a été proposée très récemment pour décrire la réponse du volume métabolique tumoral à la thérapie (David et al., 2012). Elle utilise la *théorie du raisonnement plausible et paradoxal* de Dezert-Smarandache (ou DSMT, pour l'acronyme anglais), qui permet de combiner des informations incertaines et conflictuelles (Smarandache and Dezert, 2006). Comme la méthode d'IP, elle vise à décrire l'évolution d'un VOI donné. Les étapes relatives à son implémentation sur des données de 2 examens TEP de suivi sont les suivantes (cf figure II.4):

- i. Après avoir défini le VOI contenant la tumeur à étudier, une première étape de prétraitement est appliquée pour réduire l'effet de la résolution spatiale limitée du système d'imagerie TEP (i.e. de l'EVP). L'impact de cet effet sur la classification de la réponse au traitement ayant fait l'objet d'une étude à part entière dans notre travail de thèse, nous détaillerons dans le chapitre III les causes et conséquences de l'EVP ainsi que les méthodes décrites dans la littérature pour le compenser. L'algorithme mis en œuvre par (David et al., 2012) correspond à la méthode de déconvolution/débruitage utilisée par (Boussion et al., 2009) exposée au paragraphe III.5.2.1.1. Suite à cette opération, les 2 examens TEP étudiés sont recalés suivant une

approche similaire à celle décrite dans la méthode d'imagerie paramétrique vue précédemment, à partir des images TDM associées.

- ii. La méthode de détection des changements du métabolisme tumoral à l'échelle du voxel repose sur 3 étapes distinctes :
- *Cartographie de l'image.* La zone tumorale est d'abord segmentée par une approche floue, appelée FLAB (pour *Fuzzy Locally Adaptive Bayesian* (Hatt et al., 2009)), et appliquée sur les VOI des TEP(T₁) et TEP(T₂). Les zones tumorales résultantes sont donc différentes sur la TEP(T₁) et la TEP(T₂). En particulier, cette méthode fait l'hypothèse que les distributions d'activités segmentées suivent un GMM. FLAB permet donc l'estimation, pour chaque examen j des paramètres (μ_i, σ_i^2) des 2 distributions gaussiennes f_i^j recherchées, où i est soit le « volume tumoral métaboliquement actif » (noté VM), soit le « fond » (noté F).
 - *Changement de la signature.* Suite à la cartographie des images, chaque voxel est décrit par son activité et sa probabilité d'appartenance à l'une ou l'autre des 2 distributions gaussiennes, ou classes. L'appartenance de chaque voxel à une classe de l'instant T₁ à l'instant T₂ dépend donc de ces 2 paramètres. La seconde étape consiste à caractériser les changements entre les 2 cartographies obtenues à l'étape précédente. On considère pour cela 2 niveaux d'observation : un local et un autre global.
L'évolution globale de la tumeur est évaluée à l'aide d'une mesure de distance D entre les 2 distributions gaussiennes établies grâce à la segmentation floue. Quatre distances sont ainsi calculées : $D(f_{VM}^1, f_F^2)$, $D(f_{VM}^1, f_{VM}^2)$, $D(f_F^1, f_{VM}^2)$, $D(f_F^1, f_F^2)$.
L'évolution locale de la tumeur se base sur une mesure, pour chaque voxel, d'un contraste entre ses degrés d'appartenance aux différentes classes définies à l'étape de cartographie des images. La mesure de contraste proposée par (David et al., 2012) permet de caractériser l'évolution de la densité de probabilité et la variation de l'activité du voxel dans cette classe.
 - *Détection des changements par la théorie paradoxale.* Le but de la théorie paradoxale appliqué dans le contexte du suivi est de modéliser directement les changements en utilisant des fonctions de croyance spécifiques. Dans ce contexte, on définit un référentiel de base, composé de l'ensemble des hypothèses étudiées (ici, « changement » ou « non changement » entre les deux examens, notés θ_{chg} et θ_{nochg} respectivement), appelé *cadre de discernement* Θ tel que :

$$\Theta = \{\theta_{chang}, \theta_{no\ chang}\} \quad \text{Eq. II.12}$$

Ce cadre de discernement permet de définir un second référentiel, appelé *hyper power set*, noté D^\ominus , tel que :

$$D^\ominus = \{\emptyset, \theta_{chang}, \theta_{no\ chang}, \theta_{chang} \cup \theta_{no\ chang}, \theta_{chang} \cap \theta_{no\ chang}\} \quad \text{Eq. II.13}$$

qui sert de référentiel de définition pour évaluer la véracité d'une proposition. Il permet en outre de définir une grandeur appelée *masse de probabilité élémentaire*, qui permet de modéliser les différentes évolutions possibles des voxels de la tumeur. On utilise pour cela les changements de signatures obtenus à l'étape précédente. En plus des notions de « changement » et « non changement », la DSmT permet donc de modéliser aussi les notions d'« incertitude » ($\theta_{chang} \cup \theta_{no\ chang}$) et de « paradoxe » ($\theta_{chang} \cap \theta_{no\ chang}$). Au final, on obtient une *carte de changement* permettant de caractériser les voxels ayant évolué, ou pas, entre les 2 examens TEP.

- iii. *Classification*. La dernière étape consiste à combiner les données TEP et la carte de changement pour distinguer l'évolution de chacun des voxels étudiés. (David et al., 2012) distinguent 3 types d'évolution pour les voxels tumoraux (répondeurs, stables ou progressifs) et 1 type pour les voxels caractéristiques du fond (ou physiologiques) dans les deux examens. La distinction entre voxels de la tumeur ou du fond n'évoluant pas, la classification est effectuée en s'aidant des informations obtenues durant l'étape de cartographie de l'image. Les voxels en progression et répondant à la tumeur sont distingués en étudiant le signe du changement de la valeur du SUV de chaque voxel. (David et al., 2012) définissent enfin une carte de segmentation pour les TEP(T₁) et TEP(T₂) en combinant, dans le premier cas les voxels tumoraux ayant répondu et ceux étant restés stables, et dans le second cas ceux ayant progressés et ceux étant restés stables. La réponse de la tumeur est alors déduite en calculant l'évolution relative du volume et du SUV moyen ainsi calculés sur chaque examen.

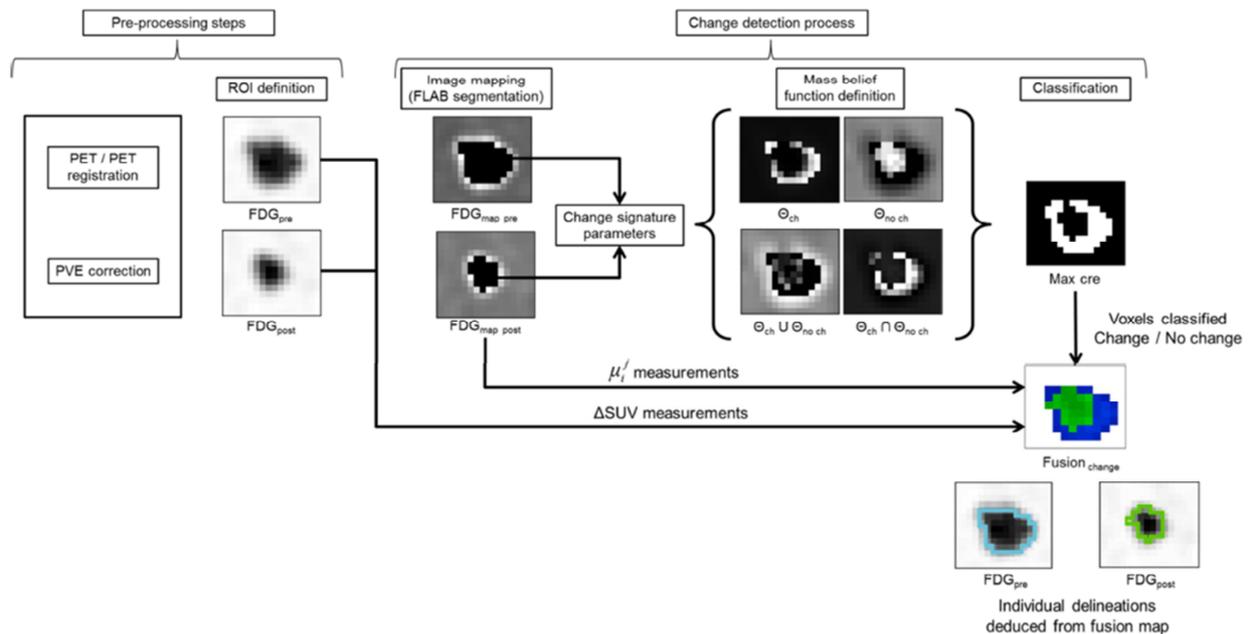


Figure II.4 : étapes mises en œuvre par (David et al., 2012). Après recalage, correction de l'effet de volume partiel et segmentation des données TEP, une première étape vise à cartographier la zone tumorale avant et après traitement au moyen d'une méthode de segmentation floue (les voxels noirs appartiennent à la tumeur, le reste est attribué au fond). A partir des fonctions de croyances dérivées de l'analyse paradoxale, on génère une image binaire représentant les changements dans l'image (voxels blancs) et les voxels n'ayant pas changé (voxels en noir). Pour distinguer parmi ces 2 familles de voxels ceux appartenant à la tumeur et ceux du fond d'une part, et ceux correspondant à une progression ou une diminution de l'activité de la tumeur d'autre part, on utilise soit les données segmentées, soit les signes de l'image de soustraction. On déduit de ce résultat une carte de changement de la tumeur à partir de laquelle on calcule l'évolution du SUV et du volume dans une zone contenant les voxels constants et décroissant sur la TEP initiale (voxels vert et bleu ici) et contenant les voxels stables et croissant sur le second examen TEP (vert et rouge, non présents dans l'exemple proposé).

Cette méthode a été validée sur des données simulées, pour lesquelles les distributions d'activité étudiées sont connues a priori, et sur des données de patients atteints de cancers de l'œsophage traités par radio-chimiothérapie. Les données simulées ont servi à évaluer la robustesse de la méthode en fonction du niveau de bruit dans l'image ainsi que l'erreur sur le volume (EV) ayant évolué ainsi que l'erreur de classification (EC) de la réponse.

Sur données simulées, la méthode de (David et al., 2012) apparaît plus précise et plus juste pour estimer le volume et la réponse que la méthode de segmentation adaptative (i.e. prenant en compte la valeur du fond local) semi-automatique classique proposée par (Nestle et al., 2005). Cette méthode fait l'objet d'une description détaillée au paragraphe IV.2.1.2. La méthodologie décrite ne donne cependant aucune information sur les données utilisées pour optimiser la méthode de segmentation adaptative à laquelle sont comparés les résultats de la méthode de (David et al., 2012). Il est donc difficile d'interpréter les forts écarts systématiques moyens obtenus par la méthode classique, que ce soit sur l'EV ou sur l'EC.

Dans le contexte de l'évaluation de la réponse thérapeutique, la méthode basée sur la théorie paradoxale a donné des résultats similaires aux index classiques tels que le SUV moyen ou maximal de la tumeur. Elle apparaît cependant plus à même de décrire l'évolution du volume métabolique que les méthodes de segmentation classiques, d'où son intérêt potentiel dans le contexte du suivi thérapeutique durant la radiothérapie, où la segmentation des volumes tumoraux peut devenir complexe à mesure que la dose augmente (Edet-Sanson et al., 2012).

La méthodologie mise en œuvre présente cependant quelques points à préciser. La méthode proposée par (David et al., 2012) prenant en compte les variations de fixations à l'échelle du voxel, il serait intéressant d'étudier la variabilité de leurs résultats en fonction du niveau de bruit dans les images, des paramètres de reconstruction employés ou encore de la méthode de déconvolution mise en œuvre. De plus, l'influence du choix du VOI tumoral étudié sur les résultats fournis par FLAB, et notamment sur les paramètres des classes obtenus, semble être un point important à évaluer.

Un dernier point de discussion porte sur la méthode employée pour distinguer les voxels de la tumeur échappant à la thérapie de ceux ayant répondu. En levant l'indétermination entre ces 2 types de réponses via le calcul de l'évolution relative du SUV dans chaque voxel classé comme ayant changé, le résultat semble sensible au bruit. Ceci est particulièrement critique vue l'étape de déconvolution mise en œuvre, et ce malgré l'utilisation d'une décomposition en ondelette des images afin de diminuer le bruit dans les données finalement corrigées de l'EVP. Ces traitements des données suggèrent une variabilité potentielle des résultats d'autant plus critique que la présence de zones échappant à la thérapie peut jouer un rôle important sur le choix d'une modification de la stratégie thérapeutique.

II.4. Conclusion

L'objet de ce chapitre était de décrire les méthodes proposées jusqu'à présent pour caractériser la réponse tumorale à partir des données anatomiques ou des données TEP. Dans ce dernier cas, il apparaît que les méthodes mises en œuvre en routine clinique reposent dans le meilleur des cas, sur la mesure du SUV, maximal ou moyen dans le volume d'intérêt. De nombreuses méthodes ont été proposées dans la littérature pour enrichir les données prises en compte, que ce soit par l'analyse de texture ou par des approches plus complexes telles que l'imagerie paramétrique ou plus récemment encore l'utilisation de la théorie du raisonnement paradoxal. Si l'application de ces méthodes à un plus grand type de lésions est nécessaire, leurs résultats préliminaires n'en sont pas moins prometteurs. Enfin, il apparaît qu'en dehors de la méthode d'analyse factorielle proposée par (Necib et al., 2008), sur laquelle nous reviendrons en détail au chapitre VI, aucune méthode proposée à l'heure actuelle ne permet de rendre en compte de l'ensemble de l'information contenue dans les images TEP pour statuer sur l'évolution de la maladie lorsque plus de 2 examens TEP sont disponibles.

Chapitre III

Effet de volume partiel : définition et méthodes de correction

L'accès généralisé à des images TEP quantitatives a été rendu possible grâce aux progrès dans le domaine de la reconstruction et au développement de la TEP/TDM. Cependant, l'Effet de Volume Partiel, ou EVP, décrit initialement par (Hoffman et al., 1979), entraîne des biais extrêmement importants (>100%) dans certaines conditions.

L'impact de l'EVP sur les mesures, et notamment sur celle du SUV, dépend des dimensions et de la forme de l'objet étudié. Il semble donc d'autant plus important de corriger de l'EVP dans le contexte du suivi thérapeutique que les formes et dimensions des tumeurs peuvent changer suite au traitement.

Après avoir détaillé les causes de l'EVP ainsi que ses conséquences sur la mesure, nous présenterons une revue des méthodes de correction proposées dans la littérature, en soulignant leurs applications en oncologie. Nous mettrons enfin en exergue les résultats de certaines de ces méthodes de correction dans le contexte du suivi thérapeutique.

III.1. Fonction de réponse

La résolution spatiale d'un système d'imagerie est définie comme la distance minimale entre deux sources ponctuelles telle que leurs images soient discernables. Elle est caractérisée par la Fonction De Réponse (FDR) du système d'imagerie qui correspond à l'image d'une source ponctuelle dans l'air. Malgré sa *non stationnarité* dans le champ de vue (figure III.1), due à la géométrie cylindrique des tomographes ainsi qu'à la réponse imparfaite des détecteurs, on la modélise souvent en :

- Négligeant ses variations dans le plan transverse.
- La décomposant sous la forme d'une gaussienne 3D dont chaque composante décrit son étalement dans les 3 directions de l'espace (radial, axial et tangential), auxquelles sont associées des largeurs à mi-hauteur (LMH) telles que :

$$FDR_{model}(x, y, z) = FDR_x(x) \times FDR_y(y) \times FDR_z(z) \quad \text{Eq. III.1}$$

avec :

$$FDR_i(i) = \frac{1}{\sigma_i(2\pi)^{1/2}} e^{-\frac{i^2}{2\sigma_i^2}} \quad \text{Eq. III.2}$$

où σ_i est tel que :

$$\sigma_i = \frac{LMH_i}{2\sqrt{2 \times \ln(2)}} \quad \text{Eq. III.3}$$

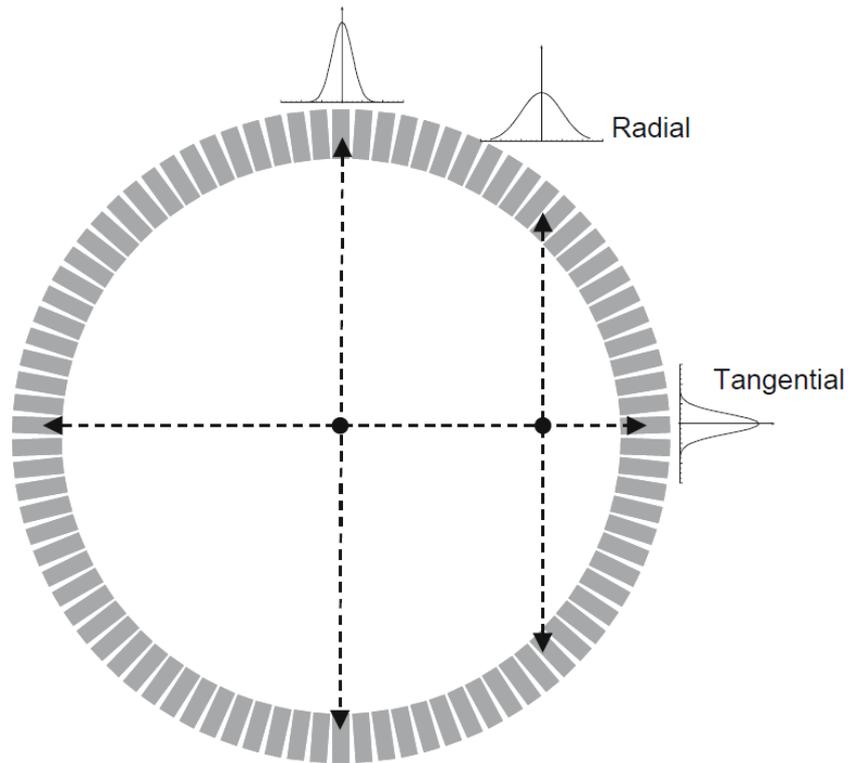


Figure III.1 : la résolution transaxiale inclut une composante tangentielle et une composante radiale. En éloignant la source radioactive de l'axe du tomographe, la probabilité pour que les photons incidents interagissent avec les scintillateurs de plusieurs détecteurs avant d'être absorbés augmente. Il est donc plus difficile de définir avec précision le lieu d'interaction des gammas de 511 keV quand la distance à l'axe augmente, d'où une résolution spatiale dégradée dans cette direction.

Des modèles plus sophistiqués ont été décrits afin de prendre en compte la non stationnarité de la FDR. Ils consistent à modéliser les phénomènes dégradant la résolution directement dans la matrice système utilisée pour la reconstruction itérative (cf paragraphe I.2.1.5). Nous reviendrons sur certaines de ces méthodes au paragraphe (cf paragraphe III.5.1.1).

III.2. Définition de l'effet de volume partiel

Le terme effet de volume partiel recouvre deux phénomènes distincts. L'EVP entraîne une différence entre l'intensité de l'objet mesurée et l'intensité réelle.

- Résolution spatiale limitée du système d'imagerie

La première cause de l'EVP est la résolution spatiale limitée du système d'imagerie. Dans le cas où l'objet mesuré n'est pas entouré d'activité, son signal se trouvera partagé avec celui des voxels qui lui sont voisins (*spill-out*, figure II.1, effet a). Par conséquent, l'activité mesurée en regard de l'objet sera plus faible que l'activité réelle (figure III.2, effet b). Inversement, quand une zone d'hypofixation est en contact avec une zone d'activité plus élevée, on mesure un signal plus intense sur les bords de la zone hypofixante (*spill in*, figure III.5C).

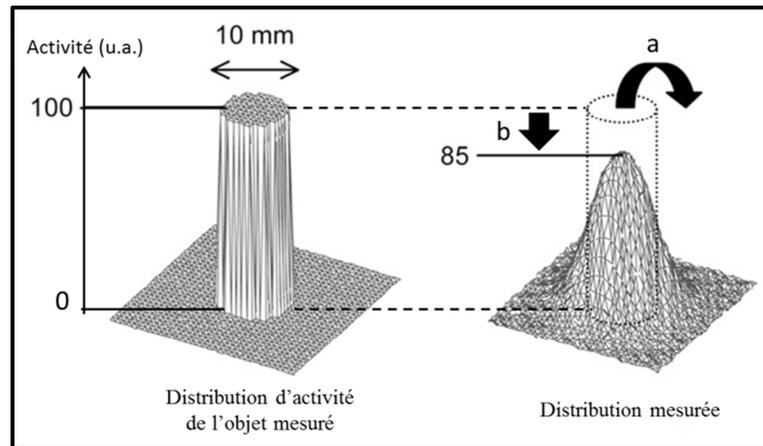


Figure III.2 : effet de la résolution spatiale du système d'imagerie sur la mesure d'une source circulaire de rayon 10mm et d'activité uniforme (en unités arbitraires) dans un fond d'activité nulle. Les conséquences sur la mesure sont un étalement du signal à l'extérieur des limites de l'objet mesuré (*spill-out*) et une diminution de l'activité maximale mesurée. (Extrait et adapté de (Soret et al., 2007)).

Connaissant la FDR du système d'imagerie, une description mathématique simplifiée de l'effet de la résolution finie du système d'imagerie sur l'image R mesurée s'écrit :

$$R = I \otimes FDR \quad \text{Eq. III.4}$$

avec I distribution d'activité réelle, FDR la fonction de réponse normalisée du système d'imagerie et \otimes l'opérateur de convolution à 3 dimensions.

- Echantillonnage des images reconstruites

Le deuxième phénomène désigné sous le terme d'EVP est dû à l'échantillonnage des images reconstruites. Il correspond au fait que les voxels de l'image reconstruite peuvent contenir des tissus d'activité métabolique différentes (figure III.3B). Le résultat de la discrétisation de l'image en voxels est alors un moyennage des activités des différentes sources d'émission contenues dans chaque voxel. Ce phénomène est généralement désigné par le terme anglais *tissue-fraction effect*. Cet effet est présent dans toutes les modalités d'imagerie, même celles dont la résolution spatiale est meilleure que celle de la TEP.

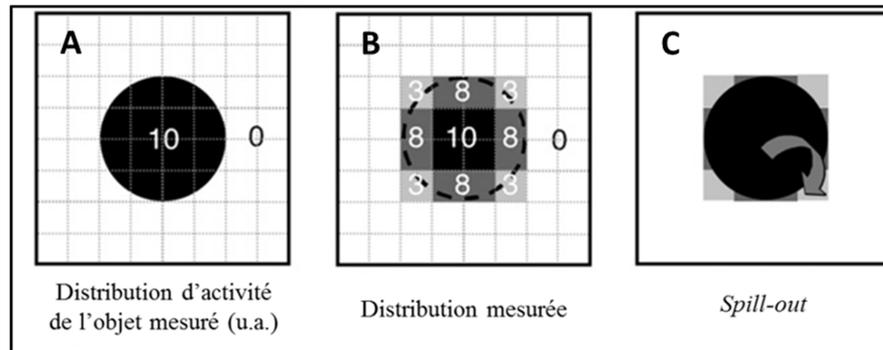


Figure III.3 : effet de l'échantillonnage sur la mesure d'un cercle d'activité homogène entouré d'une activité nulle (A). L'activité mesurée dans les pixels se trouvant à la périphérie du cercle inclut à la fois l'activité du signal provenant de l'objet et celle du fond. (B) Une partie du signal de l'objet étudié est mesurée au-delà de ses bords externes (*spill-out*, C). Extrait et traduit de (Soret et al., 2007).

On appelle donc EPV la combinaison du flou lié à la résolution spatiale finie du système d'imagerie et de l'échantillonnage des images reconstruites.

III.3. Conséquences de l'effet de volume partiel

L'EVP revêt une importance particulière en imagerie TEP compte tenu de la modeste résolution spatiale des images reconstruites. Le lien entre image idéale, c'est-à-dire non affectée de l'EVP, et image reconstruite, donné par l'équation III.3, correspond visuellement à un effet de lissage. Il est donc plus difficile de distinguer les frontières des objets étudiés du fait de la diminution du contraste apparent entre les zones de différentes intensités (effet conjoint du *spill-in* et *spill-out*, figure III.3). Cette diminution du contraste explique la difficulté de segmenter les données TEP et donc le grand nombre de méthodes proposées dans la littérature pour y parvenir (eg : (Zaidi and El Naqa, 2010)).

L'EVP a aussi un effet important sur la quantification. Du fait du *spill-out*, le signal mesuré pour une lésion métaboliquement active placée dans un milieu de faible activité sera d'autant plus étalé que les dimensions de l'objet seront faibles, comme nous le verrons dans le paragraphe III.4.1. Ceci est illustré en figure III.4. On constate que pour une sphère dont le diamètre est inférieur à 3 fois la LMH de la FDR, l'activité maximale mesurée dans l'objet est inférieure à la valeur attendue du fait de l'EVP.

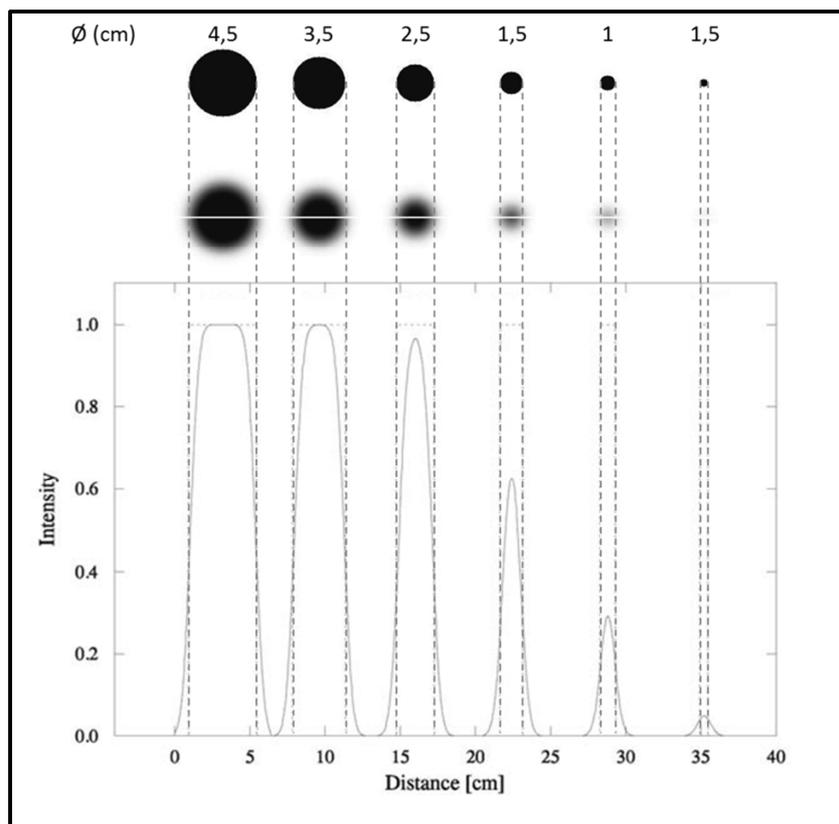


Figure III.4 : impact de l'effet de volume partiel sur la mesure de l'activité en fonction du diamètre de sources sphériques. Ligne 1) coupe transverse des sphères étudiées, d'activités homogènes unitaires identiques (u.a.) et de diamètres décroissants. Ligne 2) Effet de l'EVP (simulé ici par la convolution des images idéales par une FDR de LMH = 10 mm) sur l'image des objets étudiés et (ligne 3) profils comparés des distributions. Extrait et adapté de (Cherry et al., 2006).

Dans le cas où l'objet étudié se trouve dans un milieu d'activité non nulle, les effets conjugués du spill-out (figure III.5B) et du spill-in (figure III.5C) (appelés *spill-over*) ne se compensent que partiellement (figure III.5D), ce qui rend l'interprétation visuelle et la correction l'EVP particulièrement complexes.

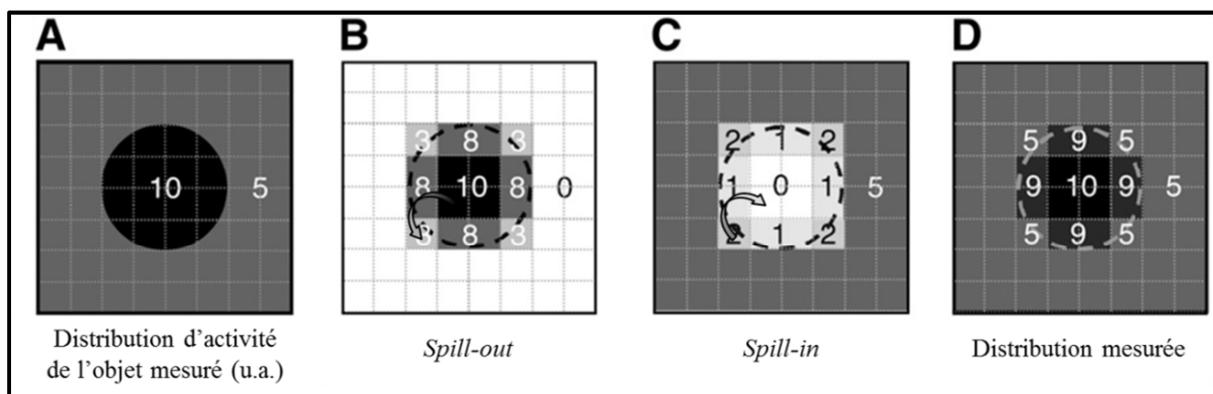


Figure III.5 : Objet mesuré (activité en u.a) (A), effet du spill-in (B) et du spill-out (C) sur cette même distribution et image mesurée résultant de ces deux effets (D). Extrait et adapté de (Soret et al., 2007).

III.4. Facteurs influençant l'effet de volume partiel

Les biais introduits par l'EVP dépendent de plusieurs paramètres parmi lesquels on distingue (Soret et al., 2007) :

- La taille et la forme de l'objet.
- Le contraste entre l'activité des tissus environnants et celui de l'objet étudié.
- La résolution spatiale des images reconstruites.
- Leur échantillonnage spatial.
- La métrique utilisée pour caractériser la distribution d'intérêt.

III.4.1. Taille et forme de la tumeur

L'EVP sera d'autant plus marqué que les dimensions caractéristiques de l'objet d'intérêt seront petites devant une longueur égale à 2 à 3 fois la LMH de la FDR de la TEP utilisée (Hoffman et al., 1979). Cet aspect est particulièrement critique dans le contexte du suivi thérapeutique. En effet, dans le cas où la thérapie entraîne une diminution importante du volume métabolique de la lésion sans pour autant modifier l'activité du métabolisme glucidique du volume résiduel, l'EVP peut se traduire par une sous-estimation de l'activité métabolique résiduelle.

Outre les dimensions de l'objet, sa forme peut jouer un rôle important sur l'impact de l'EVP (Kessler et al., 1984). En effet, si une des dimensions de la lésion étudiée est faible devant la LMH caractérisant la FDR dans cette même direction, alors l'activité mesurée sera sous-estimée, et ce indépendamment de ses dimensions dans les autres directions de l'espace. Pour évaluer l'impact de l'EVP sur un objet donné, il est donc important de retenir que sa sphéricité (cf équation V.22), traduisant l'écart entre sa forme et celle d'une sphère (de compacité égale à 1), doit être prise en compte (Mazziotta et al., 1981).

III.4.2. Contraste local

L'EVP a pour conséquence le partage d'une partie du signal propre à chacune des structures adjacentes.

L'effet de *spill-out* (figure III.5B) c'est-à-dire de contamination du signal provenant de l'objet vers les tissus environnants est complètement déterminé par l'activité, la forme et les dimensions de la distribution d'activité réelle de cet objet.

L'EVP affectant la mesure de l'activité d'un objet dépend aussi de l'activité des tissus environnant : c'est le *spill-in* (figure III.5C). Par exemple, une tumeur localisée dans le poumon, milieu d'activité faible voire nulle, sera beaucoup moins affectée par le *spill-in* qu'une tumeur localisée dans le foie. L'impact de l'EVP sera d'autant plus difficile à estimer que l'activité des tissus l'entourant diffèrera d'un bord à l'autre de la tumeur (figure III.6), comme c'est le cas pour des lésions localisées près du diaphragme par exemple. On notera aussi que les distributions d'activité mesurées dans des lésions pulmonaires sont particulièrement affectées par le mouvement respiratoire. Ceci a pour conséquence une

augmentation de l'effet de flou dans la direction du mouvement et une diminution de l'activité mesurée dans la structure étudiée.

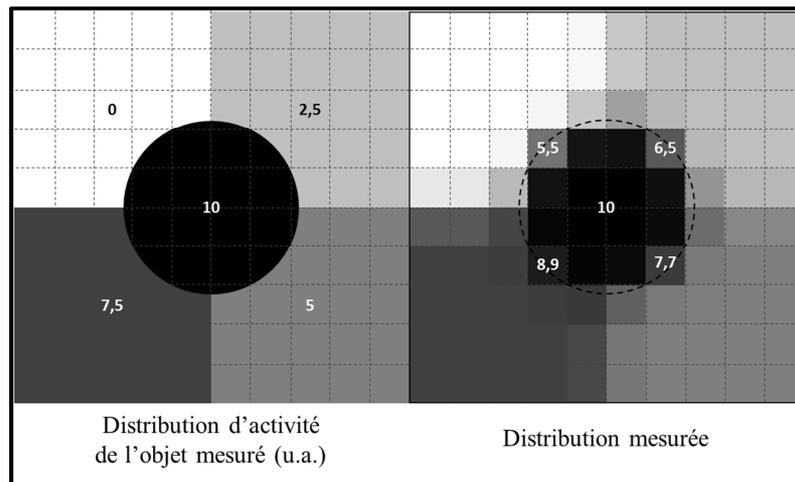


Figure III.6 : effet de l'activité des tissus environnant sur l'activité mesurée de l'objet étudié.

Comme nous l'avons vu précédemment en ce qui concerne la forme de l'objet, on comprend l'importance de la prise en compte de l'évolution de l'activité du fond situé à proximité des lésions étudiées, dans le contexte du suivi thérapeutique. L'activité métabolique des tissus environnants peut en effet changer suite au traitement (eg : réaction inflammatoire des tissus sains après radiothérapie) et ainsi modifier l'interprétation des variations du SUV des lésions. Toujours du fait du *spill-in*, la présence ou non d'une zone nécrotique à l'intérieur de la lésion réduira sa fixation apparente et donc la valeur du SUV mesurée (Soret et al., 2002).

III.4.3. Résolution spatiale de l'image reconstruite

La résolution spatiale du système d'imagerie caractérise l'étalement du signal dans l'image reconstruite (effet a, figure III.2). Pour cette raison, la valeur du SUV, pour un objet donné, diffèrera d'un tomographe à l'autre, i.e. en fonction de ses caractéristiques (eg : dimensions des cristaux, type d'acquisition), mais aussi des paramètres de reconstruction (méthode mise en œuvre, nombre d'itérations utilisées dans les reconstructions itératives (Jaskowiak et al., 2005), type de post-filtrage, etc cf paragraphe I.2.1.5). Pour ces raisons, dans le contexte du suivi thérapeutique, il est recommandé d'utiliser une même machine et de conserver ses réglages entre les acquisitions des données TEP d'un même patient (Boellaard et al., 2010; Wahl et al., 2009). La différence entre les résolutions spatiales de chaque tomographe explique en outre la difficulté de réaliser des études multicentriques et l'effort fait dans ce domaine pour minimiser les sources de variabilités (Westerterp et al., 2007).

III.4.4. Echantillonnage spatial

La résolution spatiale dans l'image reconstruite est aussi limitée par l'échantillonnage choisi. Ainsi, plus les dimensions des voxels utilisés pour représenter les données sont grandes, plus la probabilité pour qu'ils contiennent des informations métaboliques issues de tissus de natures différentes est grande (figure III.7). L'information ainsi moyennée peut alors masquer d'éventuelles zones d'hétérogénéité tumorale telles que des zones de nécrose et/ou accroître la difficulté de définir les bords des volumes tumoraux (figure III.4).

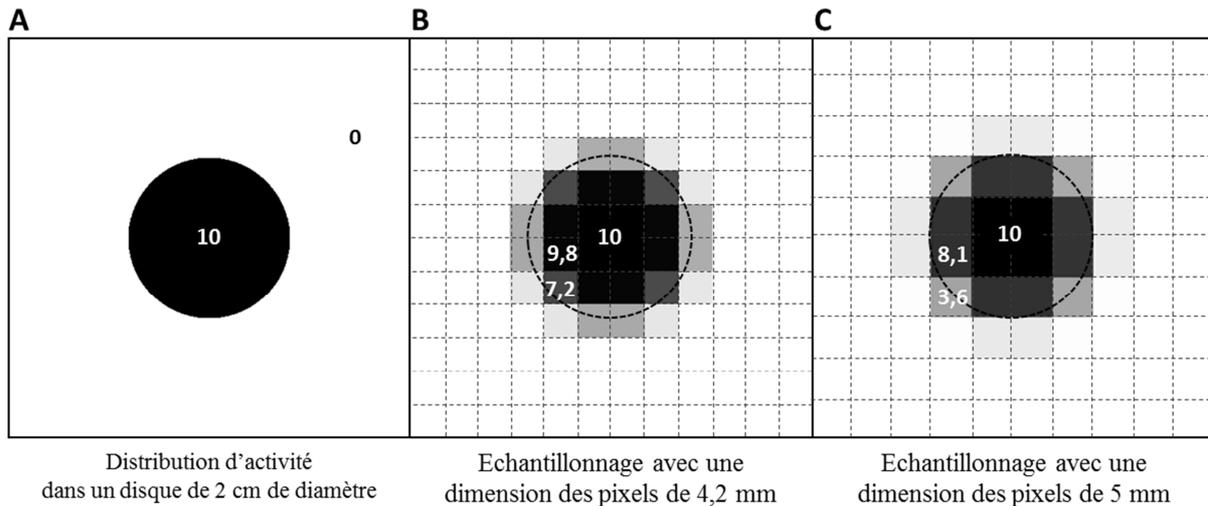


Figure III.7 : effet de l'échantillonnage sur la distribution d'activité (en u.a.) mesurée d'un disque (A) pour une taille de pixel de 4,2 et 5 cm (B et C, respectivement).

III.4.5. Méthode de mesure

L'impact de l'EVP sur la mesure du SUV dépend du VOI considéré pour caractériser la concentration d'activité de la tumeur (Soret et al., 2007). Comme nous l'avons vu au paragraphe I.3.4.1.1 ainsi que lors de la présentation des protocoles PERCIST (cf paragraphe II.3.2) et EORTC (cf paragraphe II.3.1), il n'y a pas de consensus sur la définition du VOI le mieux adapté, que ce soit pour le diagnostic, le suivi thérapeutique ou encore la définition des volumes cibles en radiothérapie (Zaidi and El Naqa, 2010).

- SUV_{max}

Le VOI le plus utilisé, que ce soit en routine clinique ou en recherche, est le voxel de plus forte intensité, utilisé pour calculer le SUV_{max} (Beyer et al., 2011; Wahl et al., 2009). La simplicité de son calcul et l'information physiologique qu'il contient (dans la mesure où il est sensé caractériser la zone la plus agressive de la lésion) justifient sa fréquente utilisation. De plus, en fonction de la forme et des dimensions de la tumeur, il est moins affecté par l'EVP ou le mouvement respiratoire que les autres index (Boellaard, 2012). Cependant, sa sensibilité au bruit et donc aux paramètres de reconstruction fait qu'il présente une répétabilité moins bonne

que les SUV moyens, mesurés à partir d'un plus grand nombre de voxels (de Langen et al., 2012).

- SUV_{peak}

Pour diminuer l'effet du bruit sur la mesure du SUV quand celui-ci est calculé à partir d'un seul voxel, (Wahl et al., 2009) (cf paragraphe II.3.2) ont récemment proposé de mesurer la concentration d'activité moyenne dans un VOI de volume fixe (VOI_{peak}) et dont la position dans la tumeur est choisie afin de maximiser la valeur du SUV finalement calculé (SUV_{peak}). (Boellaard, 2012) souligne que le SUV_{peak} serait moins sensible que le SUV_{max} aux caractéristiques de l'image telles que la résolution. L'un des problèmes relatifs à l'utilisation de cet index est l'absence de consensus sur la définition du VOI. Pour étudier l'influence du choix de la définition du VOI_{peak} sur les valeurs prises par le SUV_{peak} , et en se basant sur les données TEP ^{18}F -FLT acquises avant et pendant le traitement de 17 patients, (Vanderhoek et al., 2012) ont comparé les résultats fournis par 24 définitions du SUV_{peak} préalablement utilisées dans la littérature. Ces définitions différaient par la taille, la forme ou la localisation du VOI_{peak} dans la tumeur. Les auteurs ont ainsi démontré une forte dépendance des valeurs prises par cet index en fonction de la définition du VOI. En considérant les 24 valeurs de SUV_{peak} obtenues pour chaque tumeur, les auteurs ont calculé leurs variations relatives par rapport à leurs moyennes et ont constaté qu'elles étaient comprises entre -46% à +49%. En fonction du VOI_{peak} considéré, les auteurs ont de plus montré une variation des $\Delta SUV_{peak}(\%)$ (cf équation II.2), entre 2 examens, comprise entre +49% et -35% par rapport à la moyenne de ces valeurs obtenues pour toutes les tumeurs étudiées. Les comparaisons des évolutions relatives des SUV_{max} et des différents SUV_{peak} entre 2 examens TEP de chaque tumeur ont prouvé que $\Delta SUV_{max}(\%)$ était compris dans l'intervalle des 24 valeurs des $\Delta SUV_{peak}(\%)$ obtenus dans 70% des cas. Enfin, il apparaît que la réponse tumorale vue par les $\Delta SUV_{peak}(\%)$ est plus sensible à la taille du VOI_{peak} qu'à sa forme ou à sa localisation dans la tumeur.

- SUV moyen dans un volume segmenté

Le calcul du SUV moyen dans un volume segmenté permet de prendre en compte toute l'information métabolique de la tumeur. L'absence de consensus sur la méthode de segmentation la plus adaptée, le grand nombre d'algorithmes de segmentation proposés dans la littérature et l'absence de méthode de référence adéquate pour les valider n'ont cependant pas permis de justifier la généralisation d'index plus complexes que le SUV_{max} .

Des équipes ont comparé les résultats de certains algorithmes de segmentation suivant différents critères. (Cheebsumon et al., 2011c) ont par exemple comparé la justesse et la précision de différents algorithmes de segmentation semi-automatiques en se basant sur des données simulées et des données fantômes pour lesquels la dimension des tumeurs, le bruit dans l'image, les paramètres de reconstruction et le contraste tumeurs/tissus sains variaient. L'évaluation des performances de 6 méthodes couramment utilisées a permis de mettre en évidence que les caractéristiques des images reconstruites pouvaient avoir un effet important sur les performances des méthodes. Les auteurs soulignent en outre les performances médiocres des méthodes basées sur un algorithme de croissance de région sur des données

non lissées, en partie à cause du niveau de bruit dans les images. Cette observation justifie l'optimisation des algorithmes de segmentation utilisés au type d'images étudié.

Dans une seconde publication, (Cheebsumon et al., 2011a) ont évalué la capacité des mêmes algorithmes de segmentation à reproduire les valeurs de volume et d'activité moyenne sur les données TEP au ^{18}FDG et ^{18}FLT dynamiques de patients acquises dans un intervalle de temps court (variabilité test-retest). Si la variabilité test-retest des volumes métaboliques fournis par les différentes méthodes est très importante (jusqu'à 29%), il est apparu que celle des SUV moyens extraits des mêmes VOI est largement inférieure (<11%).

Enfin, dans leur méta-analyse sur la variabilité test-retest du SUV_{max} et du SUV moyen dans un isocontour égal à 50% du SUV_{max} , (de Langen et al., 2012) ont montré que ce dernier index avait une meilleure répétabilité que le SUV_{max} , c'est-à-dire une meilleure capacité à fournir une même valeur quand aucun paramètre (acquisition, métabolisme etc) ne change. Les différences de performances observées entre les index a justifié que les auteurs proposent d'adapter, pour chaque index, les seuils de réponse à la thérapie proposés initialement par (Wahl et al., 2009) dans PERCIST 1.0 en fonction de la VOI utilisée. Une telle adaptation des valeurs de seuils a par exemple été illustrée dans les travaux de (Buvat et al., 2012) (cf paragraphe II.3.1.3).

III.4.6. Conclusion

L'EVP dépend de plusieurs de facteurs, tels que le métabolisme de l'organe étudié et des tissus environnants, le type de tomographe utilisé, sa résolution spatiale, les dimensions et la forme de l'objet étudié, etc. Nous allons détailler les principales méthodes de correction l'EVP proposées dans la littérature. Leur objectif est d'inverser l'effet de la FDR du système d'imagerie sur l'image reconstruite pour remonter à la valeur de l'activité réelle et/ou à la distribution d'activité étudiée.

Cette opération peut être effectuée par déconvolution, c'est-à-dire en appliquant l'opération inverse de celle présentée à l'équation III.4, dans le domaine image, ou durant la reconstruction itérative par exemple. La déconvolution dans le domaine image entraîne cependant une augmentation du bruit. Dans les deux cas, l'image corrigée présente souvent des artefacts de Gibbs, ou artefact en anneaux, au voisinage des zones de fortes variations de contrastes. Ces artefacts sont notamment dus au fait que la méthode mise en œuvre tente de restaurer des informations de hautes fréquences spatiales perdues durant l'acquisition et donc absentes des données à corriger.

Un grand nombre de méthodes ont été décrites dans la littérature, que certains auteurs tels que (Soret et al., 2007), (Rousset et al., 2007), (Thomas et al., 2011) ou encore (Hutton et al., 2006) ont comparé en fonction, notamment, du domaine d'application.

III.5. Méthodes de correction de l'EVP

Nous allons détailler les principales techniques proposées dans la littérature pour corriger les données TEP de l'EVP. Ces méthodes sont généralement séparées en 2 groupes : celles appliquées durant l'étape de reconstruction et celles visant à corriger directement les images reconstruites. Nous distinguerons parmi ces deux groupes les techniques qui visent à corriger l'image et/ou la mesure dans une région d'intérêt et celles dont le but est de corriger l'image à l'échelle du voxel. Seules les méthodes visant à corriger de l'effet de la résolution finie des systèmes d'imagerie seront présentées.

Comme le soulignent (Rousset et al., 2007), la plupart des techniques de correction décrites dans la littérature ont été proposées initialement pour l'imagerie fonctionnelle du cerveau, les premières applications de la TEP sur l'Homme ayant porté sur cette localisation. Ceci explique que les méthodes de correction les plus sophistiquées nécessitent l'utilisation de données anatomiques provenant de l'IRM, recalées sur les données TEP. A la différence de la TDM, l'IRM est en effet capable de distinguer avec un très bon contraste les différents tissus mous tels que la Matière Blanche (MB) et la Matière Grise (MG). L'utilisation de la TDM étant à présent généralisée, notamment en oncologie, nous préciserons, dans notre présentation, dans quels cas les méthodes de correction ont été mises en œuvre en oncologie. Nous présenterons en outre certains des résultats liés à l'application de ces méthodes en cancérologie pour le diagnostic, la détermination du stade de la maladie ou encore la segmentation des volumes tumoraux. Enfin, nous discuterons de leur intérêt pour le suivi thérapeutique.

III.5.1. Méthodes de corrections appliquées durant la reconstruction

L'EVP provient en partie de la mauvaise résolution spatiale des images reconstruites. Un moyen de corriger de cet effet consiste donc à améliorer cette résolution durant l'étape de reconstruction des images. Les méthodes que nous allons détailler ici s'appliquent uniquement à l'échelle du voxel.

III.5.1.1. Amélioration de la résolution

La médiocre résolution des images TEP provient de nombreux paramètres physiques (cf paragraphe II.2.3.2). De même, la non stationnarité de la FDR provient en partie de la géométrie des tomographes et de la réponse imparfaite des détecteurs (cf paragraphe III.2.1.2.2 et tableau I.3). En modélisant certains de ces effets directement dans la matrice système utilisée pour reconstruire les données (cf paragraphe I.2.1.5), il est possible d'améliorer la résolution et de réduire la non stationnarité de la FDR dans le champ de vue.

La décomposition de la matrice système M proposée par (Reader et al., 2002) fait apparaître les différents effets affectant M dans *l'espace de mesure* et dans *l'espace image*, en plus de la

composante « réponse géométrique » G du système. Chaque élément G_{ij} de G représente la probabilité qu'une paire de photons émise dans un voxel j donné atteigne la paire de cristaux constituant la LDR i . Cette décomposition correspond au produit matriciel suivant :

$$M = WGB \quad \text{Eq. III.5}$$

où W (pour *weight* en anglais) prend en compte tous les effets décrits dans l'espace des mesures, G est la matrice de réponse géométrique et B (pour *blur*) modélise les effets dégradant la résolution de l'image (parcours du positon). Cette décomposition rend cependant difficile l'intégration de la non colinéarité des photons émis suite à l'annihilation des positons dans l'espace de mesure, dans la matrice G ou encore dans l'espace image.

Certaines des méthodes proposées dans la littérature ont pour but de compenser les phénomènes dégradant la résolution spatiale dans l'espace image (Reader et al., 2003), ce qui revient à définir la matrice B (équation III.5) comme la FDR du système d'imagerie.

Il existe 4 grands types de méthodes pour modéliser la matrice système. Elles se basent :

- Des simulations Monte Carlo (Engle and Kadrmas, 2007; Qi et al., 1998).
- Des modèles analytiques (Rahmim et al., 2008; Taschereau et al., 2011).
- Des mesures expérimentales ajustées (Alessio and Kinahan, 2008; Panin et al., 2006; Tohme and Qi, 2009).
- Des méthodes empiriques (Reader et al., 2003; Watson, 2011).

Pour plus de détails, nous renvoyons le lecteur à l'état de l'art exhaustif sur les méthodes de modélisation de la matrice système proposé par (Stute, 2010) dans son manuscrit de thèse.

La modélisation des effets dégradant la résolution dans la matrice système est proposée en option sur la plupart des tomographes cliniques et permet d'améliorer la résolution spatiale et le contraste dans les images reconstruites à niveau de bruit donné (Panin et al., 2006; Reader et al., 2003). Elle a par exemple été mise en œuvre en oncologie dans l'étude de (Hoetjes et al., 2010), dont l'objectif était de comparer différentes méthodes de correction de l'EVP selon des critères de précisions et de justesse de mesures du SUV. Nous décrirons précisément les conclusions de cette publication dans la discussion du présent chapitre (cf paragraphe III.5.4) du fait de son lien avec l'application des méthodes de correction d'EVP pour le suivi thérapeutique en oncologie.

III.5.1.2.A partir d'a priori anatomiques

Une alternative à la méthode précédente consiste à introduire des a priori anatomiques dans l'algorithme de reconstruction afin de contrôler le bruit et de renforcer les bords des structures fonctionnelles pré-segmentées. A partir de l'inclusion de données issues de la TDM ou de l'IRM dans le processus de reconstruction, on utilise l'approche du *Maximum A Posteriori*

(MAP), issue de la statistique bayésienne. Ce cadre de reconstruction permet par exemple d'ajouter des contraintes portant sur les variations entre les voxels voisins en fonction de la région anatomique à laquelle ils appartiennent (Bowsher et al., 1996).

Cette méthode, principalement utilisée pour la TEP en neurologie, du fait du besoin d'une mise en correspondance très précise des données anatomiques et métaboliques, a par exemple été mise en œuvre par (Baete et al., 2004; Bataille et al., 2007; Tang and Rahmim, 2009). On notera deux applications pour la TEP/TDM corps entier, proposées par (Comtat et al., 2002) et plus récemment de (Cheng-Liao and Qi, 2011).

Vue la difficulté de recaler précisément des données anatomiques sur les données TEP dans ce contexte (du fait du mouvement respiratoire par exemple), (Comtat et al., 2002) ont lissé les informations anatomiques binaires utilisées pour refléter l'incertitude sur la position exacte des structures anatomiques d'intérêt. Une étude sur données simulées et expérimentales a permis de démontrer la valeur ajoutée de la méthode qui apparaît comme un bon compromis entre recouvrement optimum du signal d'origine et compensation des erreurs potentielles dues aux écarts entre données anatomiques et métaboliques. Des performances similaires à la technique de correction de l'EVP GTM (*Geometrical Transfer Matrix*, sur données post-reconstruction, cf paragraphe III.5.2.2.3) ont été obtenues. Cette méthode n'est pas conseillée quand l'examen vise à détecter des tumeurs mais permet de diminuer les biais de mesure par rapport à une méthode ne faisant pas intervenir d'a priori anatomiques.

En utilisant la méthode de surface de niveaux (LSM, *Level Surface Method*) pour décrire les surfaces des structures métaboliquement actives contenues dans les images TEP et en les modifiant pour les adapter en parties aux informations anatomiques, (Cheng-Liao and Qi, 2011) ont proposé une nouvelle méthode MAP. Cette méthode a donné des résultats quantitatifs plus précis sur données simulées que les autres méthodes de reconstruction évaluées.

Outre ces articles, qui visaient à démontrer la faisabilité de ce type de méthodes dans le contexte de l'oncologie, aucune étude n'a utilisé ces méthodes sur des données cliniques.

III.5.2. Méthode de correction appliquée sur les images reconstruites

Le second type de méthodes de correction a pour objectif de corriger directement l'image reconstruite.

III.5.2.1. Correction à l'échelle du voxel

III.5.2.1.1. Correction par déconvolution itérative

- Problématique

L'objectif de la déconvolution est d'estimer la distribution non affectée de l'EVP (terme I de l'équation III.4) à partir de l'image reconstruite non corrigée R . Le principal problème est qu'il s'agit d'un problème mal posé, c'est-à-dire un problème pour lequel :

- Il n'y a pas toujours de solution
- Que cette solution, si elle existe, n'est pas unique
- Qu'elle ne dépend pas continûment des données, i.e. que de petites perturbations sur les données peuvent induire de « grandes » modifications de cette solution.

En effet dans notre cas, les conditions sur l'unicité de la solution et sur sa variation continue par rapport aux données ne sont pas vérifiées. Il n'y a en général pas de solution exacte pour estimer I du fait du bruit dans l'image mais il est cependant possible de s'en approcher.

- Méthodes de déconvolution classiques

On reformule l'équation III.4 telle qu'elle puisse s'écrire comme le résultat d'un opérateur, noté FDR , correspondant à la convolution de la distribution réelle par la FDR du système d'imagerie, mais en se plaçant dans le cas plus réaliste d'une image mesurée présentant une composante de bruit B telle que :

$$R = FDR(I) + B \quad \text{Eq. III.6}$$

Alors en cherchant à inverser l'opération de convolution FDR , comme cherchent par exemple à faire les méthodes de déconvolution classiques que nous allons présenter, on obtient :

$$FDR^{-1}(R) = FDR^{-1}(FDR(I)) + FDR^{-1}(B) \quad \text{Eq. III.7}$$

d'où :

$$FDR^{-1}(R) = I + FDR^{-1}(B) \quad \text{Eq. III.8}$$

Autrement dit, l'opération qui consiste à inverser l'effet de la résolution finie du système d'imagerie sur l'image mesurée permet d'obtenir la distribution réelle plus un terme d'amplification du bruit. Ceci explique qu'il soit nécessaire d'ajouter des étapes permettant de *régulariser la solution*, c'est-à-dire d'utiliser des contraintes pour obtenir une solution stable et unique.

De nombreux algorithmes ont été décrits dans la littérature. La plupart d'entre eux se basent sur la méthode itérative proposée par (van Cittert, 1931) en spectroscopie, dont la procédure originale repose sur l'algorithme suivant :

$$I^i = I^{i-1} + \alpha(R - FDR \otimes I^{i-1}) \quad \text{Eq. III.9}$$

Avec α paramètre de convergence (constant, généralement pris égal à 1) et I^i , i -ème estimée de I telle que pour $i=0$:

$$I^0 = R \quad \text{Eq. III.10}$$

Le terme de droite de l'équation III.9 est appelé résidu. La formulation proposée initialement par van Cittert (VC) n'est cependant qu'une méthode de correction de l'erreur non basée sur l'optimisation de critères statistiques connus (Carasso, 1999). Une autre formulation, plus couramment utilisée, et appelée *reblurred* van Cittert (rVC), consiste à convoluer le résidu de l'équation III.9 par la FDR tel que :

$$I^i = I^{i-1} + \alpha(FDR \otimes (R - FDR \otimes I^{i-1})) \quad \text{Eq. III.11}$$

Elle a été mise en œuvre dans des études comparatives sur données TEP par (Tohka and Reilhac, 2008), (Boussion et al., 2009), (Thomas et al., 2011) ou encore (Teo et al., 2007). Dans cette dernière implémentation de la méthode de VC (équation III.10), pour éviter le problème d'augmentation du bruit dans l'image déconvoluée, les auteurs cherchent seulement à faire converger la mesure corrigée du SUV moyen dans une VOI pré-segmentée sur la TEP non corrigée de l'EVP. Reconnaisant le trop fort niveau de bruit des images corrigées par VC, (Teo et al., 2007) insistent sur le fait que l'image déconvoluée ne doit pas être interprétée directement et que leur méthode ne doit être utilisée que d'un point de vue quantitatif. Les auteurs ont proposé de définir a priori le nombre d'itérations à utiliser pour atteindre la convergence. En outre, (Teo et al., 2007) imposent, tout comme (Tohka and Reilhac, 2006), d'ajouter une contrainte de non-négativité sur tous les pixels p de l'image déconvoluée telle qu'à chaque itération i :

$$I^i(p) \geq 0 \quad \text{Eq. III.12}$$

Il existe de nombreuses autres méthodes de déconvolution (cf (Crilly, 1991)) pour lesquelles l'expression du résidu diffère. Ainsi par exemple, la méthode de déconvolution de Lucy-Richardson (LR) (Lucy, 1974; Richardson, 1972), développée initialement pour le traitement d'image en astronomie, a aussi été mise en œuvre à de nombreuses reprises en TEP. Dans ce cas, l'étape de régularisation n'est plus additive (à la différence de la méthode de VC), mais multiplicative :

$$I^i = I^{i-1} \left(\frac{R}{FDR \otimes I^{i-1}} \otimes FDR \right) \quad \text{Eq. III.13}$$

L'avantage de cette formulation est qu'elle permet de limiter la propagation du bruit (Boussion et al., 2009).

- Amélioration des méthodes de déconvolution par débruitage du résidu

Pour limiter la propagation du bruit engendrée par la déconvolution, (Boussion et al., 2009) ont proposé d'ajouter des étapes de débruitage dans le processus itératif de déconvolution. La

méthode repose sur l'utilisation d'une décomposition en ondelettes du résidu. Cette méthode, couplée à l'algorithme de Lucy-Richardson (que nous appellerons dans la suite LRWBD, *Lucy-Richardson and Wavelet Based Denoising*), a fourni une plus grande précision et une meilleure justesse que la méthode ne mettant pas en œuvre les étapes de débruitage originale.

Cette méthode a été appliquée en oncologie par (Hatt et al., 2012a) pour évaluer la valeur prédictive et pronostique de différents index métaboliques mesurés sur une TEP au FDG acquises avant la radio-chimiothérapie de 50 patients atteints de cancers de l'œsophage. Ces auteurs ont montré que quelque soient les définitions des VOI utilisés pour calculer les concentrations d'activité nécessaire aux calculs des SUV (ici SUV_{max} , SUV_{peak} et le SUV moyen dans un volume segmenté par la méthode de (Hatt et al., 2009), cf équation I.17), LRWBD ne permettait pas d'améliorer la capacité des index à prédire :

- La réponse thérapeutique établie par les critères RECIST 1.0 (cf paragraphe II.1.1.2) établie 1 mois après la fin du traitement.
- La survie globale des patients.

- Déconvolution par un algorithme de type ML-EM

Une dernière classe de méthodes de déconvolution utilise l'algorithme de reconstruction ML-EM (Kirov et al., 2008). Elle ne nécessite cependant pas la détermination précise de la FDR, ce qui permet, en choisissant non plus de modéliser la FDR mais une fonction de déconvolution, plus large, de limiter l'augmentation de la variance dans l'image en calculant un *rapport d'imprécision*. Ce rapport permet de réduire localement la variance dans les régions d'activité uniforme, et inversement, d'augmenter le contraste dans les régions dont la variation d'activité est plus importante. Les paramètres de la méthode de régularisation changent localement afin de permettre la détection de lésion ou de bords de structures. L'avantage de cette méthode est qu'elle permet à l'utilisateur d'optimiser les résultats selon les objectifs qu'il se fixe. Dans le cas où l'objectif est d'améliorer la précision de la quantification, 8 paramètres doivent être optimisés, la plupart d'entre eux dépendant de la qualité de l'image ou des propriétés du tomographe. Enfin, (Kirov et al., 2008) soulignent que leur méthode nécessite d'être validée pour l'étude de grandes tumeurs hétérogènes, dans lesquelles le processus d'augmentation locale de la variance peut engendrer des artefacts.

Une méthode récente, basée elle aussi sur un algorithme ML-EM appliqué sur les données post-reconstruction a été proposée pour ce type précis de tumeur (Barbee et al., 2010). A la différence de la méthode précédente, elle nécessite la mesure précise de la FDR du système (modèle non stationnaire) et ne s'applique de plus qu'à un VOI défini préalablement par l'utilisateur via un masque binaire. Elle a fait l'objet d'une évaluation sur des acquisitions de fantômes présentant des volumes d'activité homogène mais une comparaison des résultats à ceux obtenus avec des méthodes plus classiques serait nécessaire pour évaluer la valeur ajoutée de cette méthode.

- Comparaisons des méthodes de déconvolution

L'avantage des méthodes de déconvolution est qu'elles ne nécessitent pas de segmentation anatomique des lésions mais seulement les données TEP reconstruites, ce qui permet de diminuer les biais liés à la mauvaise mise en correspondance de données. Ceci explique que

ces méthodes soient parmi les plus utilisées en recherche pour corriger la mesure d'activité dans le contexte de l'oncologie. A ce titre, la plupart des publications citées dans ce paragraphe ont eu pour objectif de comparer différentes méthodes de correction d'EVP, dont au moins 2 par déconvolution.

En comparant les effets de 4 méthodes de correction de volume partiel sur la mesure de l'évolution relative du SUV ainsi que sur sa variabilité test-retest, chez 15 patientes atteintes de cancer du sein et 10 patients atteints de NSCLC, (Hoetjes et al., 2010) ont constaté que les méthodes de déconvolution de :

- LR (suivant l'implémentation de (Tohka and Reilhac, 2008)).
- VC (selon la méthode de la convergence de la valeur moyenne du SUV de (Teo et al., 2007)).

conduisaient à des résultats presque identiques.

Dans leur travail visant à proposer une méthode de déconvolution permettant un meilleur contrôle du niveau bruit dans l'image corrigée, (Boussion et al., 2009) ont comparé les résultats de leur algorithme à ceux obtenus par les méthodes de déconvolution de VC et LR. Les auteurs ont mis en évidence une très nette augmentation du bruit dans les images déconvoluées par VC par rapport aux autres méthodes. L'augmentation du bruit dans les images corps entier de 13 patients, par rapport aux images non corrigées, est de $195 \pm 112\%$ en utilisant VC alors qu'elle n'est, pour les mêmes données, que de $29 \pm 14,1\%$ pour la méthode (non débruitée) de LR. En considérant les volumes métaboliques (VM) des lésions de ces 13 patients (moyenne $38 \pm 26 \text{cm}^3$ pour 22 VM) tracées manuellement par un médecin, les auteurs relèvent une augmentation de l'intensité moyenne corrigée par rapport à la valeur non corrigée de $59 \pm 26\%$ pour VC contre $39 \pm 17\%$ pour la méthode de LR non débruitée. Ce résultat n'est cependant pas cohérent avec les résultats obtenus par les auteurs sur les données fantômes, où le rapport entre le signal mesuré dans la sphère et le signal mesuré dans le fond (SBR, *Signal to Background Ratio*) a été trouvé comme étant très similaire pour les 4 algorithmes comparés. La méthode de VC surestimait cependant légèrement les valeurs mesurées par rapport aux autres méthodes de déconvolution étudiées. La différence de définition des VOI avec celle utilisée par (Teo et al., 2007) peut expliquer cet écart. Ces derniers proposent en effet de ne corriger que le SUV moyen de la zone de la tumeur la plus active, et non pas celui de zones pouvant inclure les bords de la tumeur, où la correction par déconvolution est rendue complexe du fait de la perte des informations de hautes fréquences durant l'acquisition des données. Les auteurs ont enfin montré que l'erreur sur les valeurs de SBR due à un écart sur la valeur de la LMH de la FDR utilisée pour la déconvolution augmentait plus vite pour la méthode de VC que pour celle de LR. On retiendra qu'un écart de 1 mm sur la LMH résulte sur une erreur de mesure de SBR de $\sim 5\%$ pour la méthode de VC, en accord avec les résultats de (Teo et al., 2007).

III.5.2.1.2. Transfert d'informations de hautes fréquences

Pour réintroduire les informations de hautes fréquences dans les images TEP, (Boussion et al., 2006) ont utilisé des données anatomiques de haute résolution (MMA, *Mutual Multiresolution Analysis*). Le principe consiste à extraire, transformer et intégrer les détails de données IRM ou TDM à des images de moins bonne résolution telles que des images TEP ou TEMP. Pour cela, une transformation en ondelettes discrète des deux images est effectuée afin d'obtenir les fréquences spatiales à un niveau de résolution qui leur est commun. A partir des détails de hautes fréquences contenus dans l'image anatomique et d'un modèle, des informations haute fréquence (détails ou texture par exemple) sont introduites dans l'image de plus basse résolution. Appliquée initialement par (Boussion et al., 2006) en 2D, cette méthode a été étendue à la 3D par (Le Pogam et al., 2008; Le Pogam et al., 2011). Dans cette dernière version, quand une corrélation entre données anatomiques et métaboliques est observée, la région concernée est corrigée de l'EVP. Dans le cas inverse, la nouvelle méthode ne corrige pas la zone de l'EVP, évitant ainsi la formation d'artefacts dans les données corrigées. L'image finalement obtenue peut donc présenter à la fois des régions corrigées et non corrigées de l'EVP. Cette méthode a été mise en œuvre récemment (Le Pogam et al., 2012) sur les données ^{18}F FLT-TEP/IRM(T1) de 12 patients présentant des tumeurs du cerveau (gliomes) et comparée à la méthode de déconvolution de (Boussion et al., 2009). Sans moyen de référence pour évaluer laquelle de ces méthodes fournissait la meilleure correction, les auteurs ont cependant relevé que MMA rehaussait de manière plus importante ($p=0,02$) la valeur de fixation de FLT moyenne dans un VOI tracé par un médecin et fournissait une image plus contrastée que LRWBD.

La méthode des ondelettes a été modifiée par (Shidahara et al., 2009) dans le contexte de la TEP en neurologie, pour lequel il est possible d'intégrer des données segmentées à l'aide d'un atlas. Les résultats de ces auteurs suggèrent que l'utilisation de l'atlas permet d'améliorer la quantification. Le problème de la mauvaise mise en correspondance des données anatomiques et métaboliques n'est cependant pas évoqué, ni celui des cas où les informations anatomiques ne se traduisent pas dans les données métaboliques, ce qui est susceptible d'introduire des artefacts (Le Pogam et al., 2008). Ces derniers points limitent l'application de ces méthodes au contexte général de l'oncologie.

III.5.2.1.3. Méthode basée sur le partitionnement

- Méthode de Meltzer (2 compartiments)

En utilisant les informations anatomiques issues des données TDM ou IRM, (Videen et al., 1988) ont proposé une méthode de correction de l'EVP à l'échelle du voxel, pour corriger la distribution d'activité d'une région d'intérêt donnée, appelée compartiment d'étude. Décrite initialement pour la correction de données 2D, cette méthode a ensuite été développée pour la 3D par (Meltzer et al., 1990). La motivation initiale était de mettre en évidence les zones d'atrophie dans le cerveau, en corrigeant la distribution d'activité des tissus cérébraux de celle

des tissus avoisinants. Cette approche nécessite la segmentation des données anatomiques en deux régions distinctes :

- Celle correspondant au cerveau et constituée à la fois de la MB et la MG : V_{MB+MG} .
- Le reste i.e. le liquide céphalo rachidien (LCR), fluide dans lequel baigne le cerveau et la moelle épinière : V_{LCR} .

En considérant l'image reconstruite comme la convolution de la somme des distributions d'activité des deux compartiments étudiés, on peut exprimer l'image R reconstruite par :

$$R(i) = (A_{MB+MG}(i) + A_{LCR}(i)) \otimes FDR(i) \quad \text{Eq. III.14}$$

où $A_{MB+MG}(i)$ est la valeur de l'activité réelle dans le compartiment constitué de la MB et de la MG, $A_{LCR}(i)$ celle du LCR, $FDR(i)$ la FDR du système d'imagerie et i le voxel de l'image considéré.

A partir de la segmentation des 2 compartiments, on déduit un masque binaire P_{MB+MG} , de dimension identique à celle de l'image TEP étudiée, tel que pour chaque voxel i de P , on ait :

$$P_{MB+MG}(i) = \begin{cases} 1 & \text{si } i \in V_{MB+MG} \\ 0 & \text{sinon} \end{cases} \quad \text{Eq. III.15}$$

En utilisant ce masque binaire et en faisant l'hypothèse supplémentaire que l'activité réelle dans le LCR est uniformément nulle, on obtient :

$$R(i) = [A_{MB+MG}(i) \times P_{MB+MG}(i)] \otimes FDR(i) \quad \text{Eq. III.16}$$

Pour résoudre cette équation et remonter à la valeur de la distribution d'activité réelle en tout point de l'image $A_{MB+MG}(i)$, une dernière hypothèse est que la concentration d'activité dans les tissus du compartiment étudié (V_{MG+MB}) fluctue peu par rapport à la moyenne d'activité dans ce compartiment (Meltzer et al., 1999), ce qui revient à écrire que :

$$R(i) \approx A_{MB+MG}(i) \times [P_{MB+MG}(i) \otimes FDR(i)] \quad \text{Eq. III.17}$$

d'où

$$A_{MB+MG}(i) \approx \frac{R(i)}{[P_{MB+MG}(i) \otimes FDR(i)]} \quad \text{Eq. III.18}$$

En d'autres termes, la valeur de l'activité corrigée de l'EVP en tout point du compartiment V_{MB+MG} , est égale à la valeur de l'activité mesurée en ce point divisée par la valeur obtenue après convolution du masque binaire de ce compartiment par la FDR du système d'imagerie. Il est important de retenir que, du fait des hypothèses utilisées, cette méthode implique que :

- L'autre compartiment (V_{LCR}) a une activité homogène.
- La correction d'EVP du compartiment dont on souhaite connaître la distribution d'activité réelle se fait à l'échelle du voxel. Elle ne porte donc pas sur la correction de

la valeur moyenne de son activité (à la différence de la méthode de coefficient de recouvrement, que nous détaillerons au paragraphe III.5.2.2.1).

- Méthode de Müller-Gärtner (3 compartiments)

Pour corriger non plus la distribution d'activité du cerveau dans son ensemble mais celle de la MG, (Müller-Gärtner et al., 1992) ont généralisé l'approche de (Meltzer et al., 1990) en distinguant les effets des activités de 2 compartiments (la MB et le LCR) sur celle d'un troisième (qui correspond alors au compartiment d'étude qu'est la MG). Pour ce faire, on segmente ces 3 régions anatomiques pour obtenir 3 masques binaires. Tout comme la méthode de (Meltzer et al., 1999), on suppose de plus que les activités dans la MB et le LCR sont homogènes. Elles sont estimées en considérant leurs valeurs moyennes respectives mesurées sur l'image TEP reconstruite. On obtient donc :

$$A_{MB}(i) = \overline{A_{MB}} \cdot P_{MB}(i) \quad \text{Eq. III.19}$$

et

$$A_{LCR}(i) = \overline{A_{LCR}} \cdot P_{LCR}(i) \quad \text{Eq. III.20}$$

où $A_{MB}(i)$ (respectivement $A_{LCR}(i)$) est la distribution d'activité de la MB (respectivement du LCR) en tout voxel i de l'image telle que sa valeur est prise égale à la valeur moyenne $\overline{A_{MB}}$ (respectivement $\overline{A_{LCR}}$) mesurée sur l'image TEP. En accord avec l'approche de (Müller-Gärtner et al., 1992), $\overline{A_{LCR}}$ est généralement fixé à 0. La valeur de $\overline{A_{MB}}$ peut-être calculée en considérant une grande Région d'intérêt (ROI) placée au centre de la MB sur l'image TEP reconstruite. P_{MB} et P_{LCR} correspondent aux masques binaires (cf équation III.15) de ces 2 compartiments obtenus par la segmentation des images anatomiques recalées sur les données TEP. En faisant la même approximation que pour le cas de 2 compartiments (équation III.18), on peut déterminer les valeurs d'activités corrigées du compartiment d'étude tel que :

$$A_{MG}(i) \approx \frac{R(i) - \overline{A_{MB}} \cdot [P_{MB}(i) \otimes FDR(i)] - \overline{A_{LCR}} \cdot [P_{LCR}(i) \otimes FDR(i)]}{[P_{MG}(i) \otimes FDR(i)]} \quad \text{Eq. III.21}$$

Comme dans le cas de 2 compartiments, cette méthode permet de corriger la distribution d'activité du seul compartiment d'étude (dans l'application de (Müller-Gärtner et al., 1992), c'est la MG), à l'échelle du voxel. (Meltzer et al., 1999) ont étendu cette méthode pour prendre en compte non plus 3 mais 4 compartiments.

- Modification de la méthode à 3 compartiments

Une modification de la méthode de (Müller-Gärtner et al., 1992) (appelée dans la suite *modified Müller-Gärtner*, mMGG) a été proposée par (Rousset et al., 1998b) afin d'estimer au mieux l'activité moyenne dans la MB ($\overline{A_{MB}}$, équation III.19). Cette méthode consiste à déterminer $\overline{A_{MB}}$, à partir d'une méthode de correction de volume partiel post-reconstruction

au niveau de la région : la méthode GTM (pour *Geometrical Transfer Matrix*, cf paragraphe III.5.2.2.3). Cette méthode fait l'objet d'une description détaillée par (Quarantelli et al., 2004).

- Applications

La méthode de partitionnement a été mise en œuvre à de nombreuses occasions en imagerie TEP cérébrale (eg : (Giovacchini et al., 2005; Madsen et al., 2011; Mikhno et al., 2008) mais n'a pas, à notre connaissance, été appliquée en oncologie. Ceci s'explique par le fait que les hypothèses sous-jacentes du modèle ne sont pas appropriées à ce contexte (Soret et al., 2007).

Différentes méthodes de correction d'EVP basées sur l'utilisation de segmentation des données anatomiques (dont celles de (Müller-Gärtner et al., 1992) et de (Quarantelli et al., 2004)) ont été comparées à des méthodes de déconvolution (rVC, cf paragraphe III.5.2.1.1), par (Thomas et al., 2011) dans le contexte de l'imagerie TEP cérébrale des plaques amyloïdes. Cette étude a montré que les performances de la correction de rVC étaient en deçà de celles des différentes méthodes de correction utilisant des informations anatomiques sur des données fantômes simulées. Les auteurs ont en particulier étudié le recouvrement de l'activité dans 10 zones cérébrales appartenant à la MG dans un fantôme simulant un patient atteint de la maladie d'Alzheimer. Dans ce cas, la différence absolue, moyennée sur les 10 structures, des rapports entre activités simulées et activités corrigées de l'EVP, est d'environ 20% quand rVC est utilisé, et de moins de 5% pour les autres méthodes de correction d'EVP évaluées. Les auteurs expliquent ces sous-performances de la méthode de rVC au fait que les critères utilisés pour arrêter l'algorithme ne permettaient pas une correction suffisante.

Un autre résultat intéressant de Thomas et al. porte sur les coefficients de variations (i.e. l'écart-type sur la moyenne), pour chaque structure, des valeurs d'activités calculées grâce à des mesures effectuées sur plusieurs jeux de données simulées. Les résultats montrent que la correction d'EVP augmente la variabilité inter-sujet des activités mesurées par rapport à la mesure non corrigée, et ce quelle que soit la zone cérébrale étudiée. Pour la majorité des zones cérébrales, la variabilité des valeurs corrigées était plus proche de celle des valeurs d'activité simulées (ou valeurs vraies). En d'autres termes, selon ces auteurs, la correction d'EVP augmenterait la variabilité inter-patient des valeurs d'activités mesurées, mais les rapprocheraient alors des variations réelles des activités des structures étudiées.

III.5.2.2. Corrections à l'échelle d'une région d'intérêt

III.5.2.2.1. Méthode du coefficient de recouvrement

- Principes de la méthode

L'une des premières méthodes de correction de l'EVP, et aussi l'une des plus simples à mettre en œuvre, a été proposée par (Hoffman et al., 1979). Elle consiste à multiplier la valeur de fixation de l'objet mesurée sur l'image TEP reconstruite dans une ROI ($A_{\text{mesurée}}$) et à diviser cette valeur par un Coefficient de Recouvrement, ou CR (ou RC *Recovery Coefficient* en

anglais), tabulé, pour obtenir la valeur corrigée de l'EVP : $A_{\text{corrigée}}$ (équation III.22). L'une des hypothèses implicites de ce modèle est donc que l'activité de l'objet d'intérêt est homogène dans la ROI étudiée.

$$RC = \frac{A_{\text{mesurée}}}{A_{\text{corrigée}}} \quad \text{Eq. III.22}$$

La détermination de la valeur de CR à utiliser pour l'objet étudié nécessite la connaissance de ses dimensions ainsi que de la FDR du système d'imagerie. Il est donc nécessaire d'utiliser des CR adaptés au type d'objet étudié (notamment à sa forme, cf paragraphe III.4.1) et que les valeurs utilisées aient été déterminées sur un tomographe de FDR similaire à celle du tomographe utilisé pour acquérir les mesures à corriger. (Hoffman et al., 1979) concluent de plus que cette méthode de correction est bien adaptée pour corriger la mesure d'activité d'objets dont les dimensions sont supérieures ou égales à plus de 1 fois la LMH de la FDR du système d'imagerie. Comme le propose (Hoffman et al., 1979), l'estimation des CR à appliquer pour corriger l'activité mesurée de l'objet étudié peut se faire par deux moyens. Le premier consiste à estimer les dimensions de l'objet à partir des images anatomiques. On fait alors l'hypothèse implicite que le volume anatomique est identique au volume fonctionnel. La seconde approche consiste à déterminer directement les dimensions du volume fonctionnel à partir des images fonctionnelles par segmentation par exemple. Cette dernière méthode, sur lesquels nous reviendrons plus loin du fait de son application dans le contexte du suivi thérapeutique (cf paragraphe III.5.4), a été mise en œuvre par (Avril et al., 1997; Gallivanone et al., 2011). Une étude de faisabilité de la méthode de (Hoffman et al., 1979) qui portait sur des données fantômes, réalisée par (Geworski et al., 2000), a montré que cette technique de correction était applicable pour des lésions dont les dimensions étaient supérieures à 1,5 fois la LMH. Les CR sont généralement calculés pour des objets de formes simples via des acquisitions de données fantômes tels que des sphères ou des cylindres d'activité homogène. On notera qu'il est possible d'appliquer le CR à une valeur de $A_{\text{mesurée}}$ calculée comme l'activité moyenne ou maximale dans une ROI, si le calcul du CR a été basé sur des mesures d'activités moyennes ou maximales, respectivement.

Afin d'adapter cette méthode à des cas où l'activité du fond local est non nulle, et donc de corriger la mesure du spill-over, (Kessler et al., 1984) ont proposé des valeurs de CR prenant en compte non seulement la forme mais aussi la valeur du contraste entre l'objet et le fond. Il est possible d'adapter la méthode de (Kessler et al., 1984) à des volumes de formes quelconques sans utiliser de table pour déterminer la valeur du CR (Soret et al., 2007). Pour cela, il est nécessaire de se donner un masque binaire P_{objet} représentatif de l'objet étudié.

$$P_{\text{objet}}(i) = \begin{cases} 1 & \text{si } i \in V_{\text{objet}} \\ 0 & \text{sinon} \end{cases} \quad \text{Eq. III.23}$$

où V_{objet} correspond au volume segmenté de l'objet étudié et i désigne un voxel dans le repère de l'image TEP considérée. Ce masque peut être déterminé soit à partir d'une image anatomique, soit en segmentant directement les données TEP par exemple. On notera A_{fond}

l'activité mesurée sur l'image TEP du fond local dans lequel est plongé l'objet étudié, $A_{\text{mesuré}}$ l'activité moyenne de l'objet dans ce masque projeté sur l'image TEP. En faisant l'hypothèse que :

- Le fond n'est pas affecté par l'EVP.
- L'activité de l'objet mesuré est homogène.

on peut alors appliquer l'équation III.14 proposée dans le cadre de la méthode de (Meltzer et al., 1990). En faisant l'hypothèse proposée par (Meltzer et al., 1999) (cf équation III.18), on en déduit que :

$$A_{\text{mesurée}} = (A_{\text{corrigé}} - A_{\text{fond}}) \times RC + A_{\text{fond}} \quad \text{Eq. III.24}$$

ou encore que :

$$A_{\text{corrigé}} = \frac{(A_{\text{mesuré}} - A_{\text{fond}})}{RC} + A_{\text{fond}} \quad \text{Eq. III.24'}$$

où CR correspond à la moyenne, dans V_{objet} , des valeurs obtenues suite à la convolution de P_{objet} par la FDR du système d'imagerie (cf équation III.18).

L'utilisation de la méthode de (Kessler et al., 1984) nécessite de vérifier des hypothèses fortes, telles que :

- Le fait que l'activité dans l'objet d'intérêt soit homogène.
- L'activité du fond ne soit pas affectée de l'EVP (ce qui n'est pas vérifié quand l'objet est à l'interface de milieu d'activités différentes par exemple, ou encore quand le fond considéré est lui-même constitué de milieux d'activités différentes (Links et al., 1996; Soret et al., 2007)).

- Application de la méthode du CR en oncologie

Malgré ces hypothèses, et surtout du fait de la simplicité de sa mise en œuvre, la méthode du CR proposée par (Kessler et al., 1984) a été utilisée pour corriger les mesures d'activité en TEP dans le domaine de l'oncologie (Avril et al., 1997; Avril et al., 2001; Bural et al., 2009; Gallivanone et al., 2011; Geworski et al., 2000; Hickeson et al., 2002; Jentzen et al., 2008; Menda et al., 2001; Sakaguchi et al., 2010; Srinivas et al., 2009; van Heijl et al., 2010; Vesselle et al., 2004). Certains de ces travaux ont démontré la faisabilité de cette méthode pour des études cliniques ou ont proposé différentes méthodes d'implémentation (Gallivanone et al., 2011; Geworski et al., 2000; Sakaguchi et al., 2010; Srinivas et al., 2009). Elle a aussi été mise en œuvre pour dans le cadre de nombreux travaux de recherche clinique.

Dans leurs résultats, (Avril et al., 1997) montrent qu'en corrigeant à la fois de l'EVP par la méthode de CR et du glucose sanguin ($SUV_{\text{CR_glu}}$), la mesure du SUV permettait de mieux distinguer les tumeurs malignes des tumeurs bénignes dans le cas de cancers du sein. Des résultats équivalents, basés sur la mesure du SUV corrigé de l'EVP par la méthode du CR

(SUV_{CR}), ont été obtenus par (Hickeson et al., 2002) sur des cancers du poumon. Pour cette dernière localisation, la comparaison des capacités de différentes définitions du SUV_{max} à séparer lésions pulmonaires bénignes et malignes a montré que la méthode de correction par CR ne permettait pas d'améliorer la classification. Au contraire, la comparaison des aires sous les courbes ROC (cf paragraphe IV.2.1.4.1), a montré que le SUV_{CR} était le moins à même de séparer les deux groupes de lésions ($AUC=0,864$), le meilleur index étant le SUV_{max} normalisé par le poids du patient (0,915).

Toujours pour les cancers du sein, et en s'appuyant sur leurs résultats précédents, (Avril et al., 2001) ont utilisé le SUV_{CR_glu} pour voir à quelles données histologiques cette mesure de l'activité métabolique était la plus corrélée. Les résultats ont montré des corrélations faibles entre cette définition du SUV et les données histologiques, suggérant que la mesure de la fixation de FDG utilisée ne caractérisait pas le comportement biologique des cancers du sein. Les auteurs n'ont trouvé aucune corrélation entre la fixation de FDG et les dimensions de la tumeur (coefficient de corrélation de Pearson égale à 0,06). Ils n'ont cependant pas cherché à évaluer l'effet de la correction d'EVP sur les corrélations étudiées.

Dans une série de travaux mettant en œuvre cette correction, (Vesselle et al., 2000) ont montré l'intérêt de la correction d'EVP par le CR chez des patients atteints de cancers du poumons non à petites cellules (NSCLC, *Non Small Cell Lung Cancer*) du fait d'un lien fort entre le SUV_{max} corrigé de l'EVP et le degré de différenciation cellulaire mesuré histologiquement ($p < 10^{-4}$). Les auteurs ont aussi relevé que le fait d'appliquer le CR supprimait la corrélation entre le diamètre anatomique de la lésion et le SUV_{max} (le coefficient de corrélation de Spearman passant alors de 0,56 ($p < 10^{-3}$) à -0,04 ($p = 0,83$)). Toujours chez des patients NSCLC ($N=178$), mais dans le contexte de la stadification, (Vesselle et al., 2004) ont ensuite prouvé que la correction de la valeur du SUV_{max} par la méthode de CR faisait disparaître la corrélation observée entre la mesure d'activité non corrigé et le grade de la tumeur. Dans une dernière étude, (Vesselle et al., 2007) ont enfin montré, en se basant sur une étude portant sur la survie de patients atteints de NSCLC (Higashi et al., 2002), que la correction d'EVP par CR du SUV ne permettait plus de prédire la survie ($p = 0,08$) pourtant établie avec la mesure de fixation non corrigée ($p < 10^{-3}$). Ce dernier résultat doit cependant être soumis à caution puisque les valeurs de CR utilisées n'étaient pas adaptées au tomographe de (Higashi et al., 2002), ni à sa mesure du SUV (les valeurs de SUV rapportées étant des SUV moyens dans une ROI et non le SUV_{max}).

III.5.2.2.2. Méthode d'ajustement

En se basant sur les travaux de (Kessler et al., 1984), (Chen et al., 1999) ont proposé une méthode applicable en oncologie permettant d'estimer à la fois le volume métabolique de l'objet étudié et son activité corrigée de l'EVP. En supposant que les valeurs d'activité dans l'objet et dans le fond sont homogènes et que l'objet étudié est sphérique, les auteurs ont décrit un algorithme itératif dont l'objectif consiste à minimiser la différence entre la distribution d'activité mesurée et un modèle dans lequel varient :

- La localisation de la sphère,
- Son diamètre,
- Son activité,
- L'activité du fond.

Cette méthode nécessite l'optimisation de 3 paramètres pour des résultats optimaux. Elle a été comparée à celle du CR par ses auteurs sur des données fantômes contenant des implants d'activités homogènes de formes elliptiques ou circulaires. Elle a permis d'améliorer significativement l'exactitude des valeurs d'activité estimées.

Dans leurs travaux, (Tylski et al., 2010) ont proposé de modifier la méthode de (Chen et al., 1999) afin de prendre en compte n'importe quelle forme de tumeur. Cette modification repose sur une estimation initiale grossière du volume tumoral. On utilise pour cela une méthode de segmentation semi-automatique prenant en compte l'activité de fond (Nestle et al., 2005) (cf paragraphe IV.2.1.2) que l'on dilate pour obtenir un volume englobant le volume métabolique réel de la tumeur. On utilise alors un algorithme itératif minimisant une fonction de coût afin d'estimer la distribution corrigée de l'EVP qui se rapproche le plus, une fois convoluée par la FDR du système, de la distribution mesurée. Sur données fantôme et données simulées, la méthode d'ajustement a atteint une précision, en terme de segmentation, similaire à la méthode de (Nestle et al., 2005), elle-même optimisée pour les données traitées. Toujours dans les travaux de (Tylski et al., 2010), la méthode d'ajustement a fait l'objet d'une comparaison exhaustive à la technique de CR. Connaissant les valeurs d'activités à estimer sur différents jeux de données simulées, les auteurs ont étudié l'erreur moyenne sur la mesure du SUV, maximum ou moyen. Les VOI utilisés pour calculer ces SUV moyens étaient obtenus à partir de différentes méthodes de segmentation semi-automatique, puis corrigées ou non de l'EVP par la méthode du CR. Les résultats montrent que le SUV moyen segmenté par la méthode de (Nestle et al., 2005) et corrigé par la méthode du CR d'une part, et la méthode d'ajustement d'autre part, se sont avérées être les plus robustes des 10 étudiées. Elles sont en outre plus précises que le SUV_{max} , pourtant peu affecté par l'EVP.

III.5.2.2.3. Geometric Transfer Matrix (GTM)

La méthode de CR de (Kessler et al., 1984) ne permet de prendre en compte que le spill-over entre 2 compartiments : le fond et l'objet (cf équation III.24'). La méthode GTM permet de corriger de l'EVP les activités moyennes de n'importe quel nombre de compartiments.

Proposée initialement par (Rousset et al., 1998a) dans le contexte de la neurologie, cette méthode nécessite la segmentation sur les données IRM des N régions étudiées, mises en correspondance avec les données TEP.

Le modèle suppose que les distributions d'activité à l'intérieur de chacune des N VOI_k segmentées (avec k numéro attribué au compartiment k tel que $k \in [1;N]$) sont uniformes et

égales à leurs moyennes respectives, notées $\overline{A_{mesurée\ k}}$. Ces activités sont calculées directement sur les images TEP. Elles permettent de former un vecteur $\overrightarrow{A_{mesuré}}$ tel que :

$$\overrightarrow{A_{mesuré}} = [\overline{A_{mesuré\ k}}] / \overrightarrow{A_{mesuré}} \in \mathcal{M}_{N,1}(\mathbb{R}) \quad \text{Eq. III.25}$$

où $\mathcal{M}_{N,1}(\mathbb{R})$ correspond à l'espace des matrices de dimensions N lignes et 1 colonne et à valeurs dans \mathbb{R} .

La seconde étape consiste à déterminer :

- La contribution du signal de chaque compartiment sur le signal mesuré dans les N-1 autres compartiments.
- La part de l'activité émise par chaque compartiment et détectée dans ce même compartiment.

Pour définir les NxN contributions possibles entre les activités de chaque couple de compartiments (k,l) (avec $k \in [1;N]$ et $l \in [1;N]$), on calcule les coefficients g_{kl} . Ces coefficients permettent de définir la *matrice de contamination croisée* G telle que :

$$G = \begin{pmatrix} g_{1,1} & \cdots & g_{1,N} \\ \vdots & \ddots & \vdots \\ g_{N,1} & \cdots & g_{N,N} \end{pmatrix} \quad \text{Eq. III.26}$$

Il existe plusieurs méthodes pour déterminer les valeurs des g_{kl} (Frouin et al., 2002) :

- *L'approche image* consiste à utiliser les N masques binaires P_k (cf équation III.23) déterminés à l'aide des segmentations des compartiments sur les données anatomiques. On convolue le masque P_k de chaque structure par la FDR du système d'imagerie TEP considéré puis on calcule la moyenne des valeurs ainsi obtenues dans chaque VOI_l. Les valeurs diagonales de G représentent alors la part de l'activité émise dans une structure et détectée dans cette même structure. Les termes g_{kl} (respectivement g_{lk}), tels que $k \neq l$, représentent la contamination en activité du compartiment l (respectivement k) dans le compartiment k (respectivement l).
- *L'approche sinogramme* (Rousset et al., 1998a) consiste à utiliser un simulateur analytique afin d'obtenir les sinogrammes correspondant à ces masques binaires puis à les convoluer par la fonction gaussienne représentant la FDR du détecteur. On reconstruit alors les sinogrammes ainsi dégradés par la résolution du système d'imagerie en utilisant le même algorithme que celui utilisé pour reconstruire les images TEP étudiées.

En se basant sur des données simulées et des données fantômes, les résultats de la méthode GTM obtenus en mettant en œuvre ces deux approches ont été comparés par (Frouin et al., 2002) dans le contexte de l'imagerie TEP des striata. Ces deux approches conduisent à des résultats tout à fait comparables, l'approche image présentant cependant un gain en temps de calcul important (de l'ordre de 1000 sur des données TEP 3D).

En écrivant les activités moyennes $\overrightarrow{A_{corrigée k}}$ non affectée de l'EVP de chaque compartiment k sous la forme d'un vecteur $\overrightarrow{A_{corrigée}}$, et en faisant l'hypothèse que les activités mesurées de chaque compartiment peuvent se décomposer sous la forme d'une combinaison linéaire des activités réelles des différentes structures, on obtient :

$$\overrightarrow{A_{mesurée}} = \mathbf{G} \times \overrightarrow{A_{corrigée}} \quad \text{Eq. III.27}$$

d'où, en inversant \mathbf{G} :

$$\overrightarrow{A_{corrigée}} = \mathbf{G}^{-1} \times \overrightarrow{A_{mesurée}} \quad \text{Eq. III.28}$$

Comme la méthode du CR de (Kessler et al., 1984), la méthode GTM ne vise pas à fournir une image corrigée de l'EVP (comme le font par exemple les méthodes de déconvolution) mais estime seulement des valeurs moyennes d'activités corrigées de l'EVP dans les N compartiments segmentés sur les données anatomiques. Dans ce contexte, (Rousset et al., 1998a) ont montré que leur méthode permettait de réduire les biais sur la mesure d'activité dans les petites structures cérébrales à hauteur de 0-6%, contre 25-50% en l'absence de correction de l'EVP. En outre, seul $\overrightarrow{A_{corrigée}}$ est affecté par les erreurs de mise en correspondance des données anatomiques et métaboliques (Frouin et al., 2002) alors que $\overrightarrow{A_{corrigée}}$ et \mathbf{G}^{-1} sont tous deux affectés par des erreurs de segmentation (Tohka and Reilhac, 2008).

- Modifications de la méthode GTM

De nombreuses modifications de la méthode GTM ont été proposées :

- (Du et al., 2005) ont démontré les limites de la méthode GTM en TEMP quand l'algorithme de reconstruction itératif mis en œuvre n'est pas linéaire i.e. quand la somme des reconstructions de 2 jeux de projections n'est pas égale à la reconstruction de leur somme. Pour améliorer la correction d'EVP dans ce cas, ces auteurs ont appliqué une petite perturbation aux données reconstruites en utilisant l'approche sinogramme pour calculer les coefficients g_{kl} . Sur des données fantôme et des données simulées dans le contexte de la TEMP du cerveau, la méthode pGTM (*perturbation based GTM*) a réduit significativement les biais par rapport à la méthode GTM non modifiée (Du et al., 2006).
- En considérant un voxel comme un compartiment pour la méthode GTM, il serait théoriquement possible de corriger l'image de l'EVP à l'échelle du voxel. Ceci nécessiterait cependant l'inversion d'une matrice de perturbation de taille très grande et mal conditionnée i.e. très sensible au moindre changement dans l'image et donc au bruit. Pour résoudre ce problème, (Segobin et al., 2010) ont développé la méthode PARSLR (*PARtially-Segmented Lucy-Richardson*) qui consiste à coupler l'approche

GTM à la méthode de déconvolution de LR. Appliquée sur des données fantôme 2D du cerveau, PARSLR n'a sous-estimé la valeur d'activité mesurée que de 0,9% contre 5,2% pour la méthode de LR et 6,5% pour GTM. Appliquée à l'échelle du voxel, la méthode GTM a conduit à une sous-estimation de 157% de la valeur corrigée du fait du mauvais conditionnement de la matrice G.

- Une dernière mise en œuvre de la méthode GTM, appelée RBV (*Region Based Voxel-wise*) a été proposée par (Thomas et al., 2011) pour estimer une image corrigée de l'EVP. Elle consiste à appliquer d'abord la méthode GTM pour estimer les activités moyennes $\overline{A_{corrigée k}}$ dans les N régions segmentées. A partir de ces valeurs, (Thomas et al., 2011) utilisent les N masques binaires et leur attribuent leurs valeurs $\overline{A_{corrigée k}}$ respectives pour créer une distribution homogène par partie, notée S. La dernière étape de cette méthode consiste à appliquer une correction d'EVP à l'échelle du voxel, en introduisant les informations issues de S dans l'image TEP reconstruite R telle que l'image finalement corrigée s'exprime :

$$I(i) = R(i) \times \frac{S(i)}{S(i) \otimes FDR(i)} \quad \text{Eq. III.29}$$

Cette méthode permet de corriger de l'EVP au sein d'un même type de tissu, tels que les différents noyaux contenus dans la MG : le thalamus, l'amygdale ou encore l'hippocampe. Utilisée pour la quantification des plaques amyloïdes, (Thomas et al., 2011) ont montré une correction quasi-parfaite de l'activité mesurée sur données simulées, et ce quel que soit la zone du cerveau étudiée. La variabilité des résultats obtenus par RBV en fonction de la qualité de la segmentation des données anatomiques utilisée doit encore être établie.

- Applications de la méthode GTM

Cette méthode a été très utilisée en TEP ou en TEMP cérébrale ou cardiaque. Elle a par exemple été comparée aux méthodes de déconvolution de LR et rVC (cf paragraphe III.5.2.1.1) par (Tohka and Reilhac, 2008), sur des données TEP dynamiques simulées mettant en œuvre du ^{11}C -Raclopride. Dans ce contexte, les paramètres physiologiques d'intérêt sont le potentiel de liaison du radio ligand (PL) et les TAC déduites de ces images dans un VOI. Les résultats montrent la supériorité de la méthode GTM pour la quantification de ces grandeurs quand une segmentation parfaite des données anatomique est utilisée. Les méthodes de déconvolution ont cependant atteint une précision comparable à celle de la méthode GTM dans des conditions optimales quand les régions utilisées pour déterminer les TAC et PL étaient adaptées aux méthodes de déconvolution (i.e. érodées, afin de prendre en compte la non correction de l'effet de l'échantillonnage des images, cf paragraphe III.2). La méthode de (Rousset et al., 1998a) s'est cependant avérée moins robuste que les méthodes de déconvolution en cas d'erreurs de recalage ou de mauvaise segmentation des données anatomiques. Elle est aussi difficilement transposable pour l'imagerie des cancers du fait de la

segmentation nécessaire des données anatomiques et de l'hypothèse d'homogénéité de la distribution d'activité dans les différents compartiments.

Comme la méthode de (Müller-Gärtner et al., 1992), l'application de la méthode GTM est rendue complexe pour la TEP en oncologie du fait des besoins en recalages et segmentation des données. Elle n'a pas été appliquée dans ce domaine.

III.5.3. En résumé

Le choix de la stratégie de correction d'EVP la plus adaptée à un problème dépend de nombreux paramètres tels que :

- La nature des données d'entrée (TEP/IRM ou TEP/TDM, avec ou sans injection de produit de contraste etc).
- La localisation étudiée.
- Le type de correction souhaitée (i.e. d'une correction de l'image ou directement de celle de la mesure).
- De la validité des hypothèses sous-jacentes dans le contexte (eg : distribution homogène dans le compartiment).

Un grand nombre des méthodes de correction de l'EVP appliquées aux images reconstruites et présentées ici ont été développées pour l'imagerie TEP et TEMP du cerveau. Une particularité de cette localisation est la quasi-absence de mouvement des tissus cérébraux et donc la relative simplicité de la mise en correspondance entre les données anatomiques et métaboliques. De plus, ces méthodes de correction nécessitent une segmentation très soignée de structures anatomiques (MB, MG et LCF par exemple), visibles en IRM, mais souvent non distinguables par la TDM autrement que par l'injection de produit de contraste. Pour ces raisons, la plupart des méthodes utilisées en imagerie cérébrale sont difficilement transposables en oncologie. Ceci explique que seules les méthodes, très simples, du CR, proposées par (Hoffman et al., 1979; Kessler et al., 1984) et de déconvolution, qui ne nécessitent pas forcément d'a priori anatomiques pour être utilisées, aient été mise en œuvre dans ce contexte.

En nous basant sur la classification des méthodes de correction proposée par (Soret et al., 2007), le tableau 1 résume les méthodes de correction de l'EVP présentées dans ce chapitre et appliquées au moins une fois dans le contexte de la correction d'EVP pour l'imagerie TEP des cancers.

Propriétés	Nom de la méthode	Modélisation plus fine de la FDR	Déconvolution	Transfert d'informations de hautes fréquences	Coefficient de recouvrement	Méthode d'ajustement
Hypothèses						
Volume tumoral connu?	Non	Non	Non	Non	Oui, au moins la/les dimension(s) caractéristique(s)	Oui
Activité tumorale homogène?	Non	Non	Non	Non	Oui	Oui
Besoin de données anatomiques?	Non	Non	Non	Oui	Non pas nécessairement	Non
Informations générales sur la méthode						
Échelle de la correction	Voxel	Voxel	Voxel	Voxel (ou locale eg : Le Pogam (2012))	Régionale	Régionale
Moment d'application de la correction	Pendant la reconstruction	Après la reconstruction	Après la reconstruction	Après la reconstruction	Après la reconstruction	Après la reconstruction
Résultat fourni par la méthode	<ul style="list-style-type: none"> Image moins affectée de l'EVP FDR plus stationnaire dans le FOV 	<ul style="list-style-type: none"> Généralement, image corrigée de l'EVP Parfois, valeur d'activité corrigée (eg : Teo (2007)) 	<ul style="list-style-type: none"> Image corrigée de l'EVP 	<ul style="list-style-type: none"> Valeur d'activité (maximale ou moyenne) corrigée de l'EVP 	<ul style="list-style-type: none"> Valeur d'activité moyenne corrigée de l'EVP Estimation du volume métabolique non affecté de l'EVP 	
Références liées à l'implémentation ou à la mise en œuvre de la méthode	<p>Qi (1998)</p> <p>Reader (2003)</p> <p>Panin (2006)</p> <p>Rahmim (2008)</p> <p>Alessio (2008)</p> <p>Tohme (2009)</p> <p>Hoetjies (2010)</p> <p>Taschereau (2011)</p> <p>Watson (2011)</p>	<p>van Cittert (1931)</p> <p>Richardson (1972)</p> <p>Lucy (1974)</p> <p>Teo (2007)</p> <p>Kirov (2008)</p> <p>Tohka (2008)</p> <p>Boussion (2009)</p> <p>Hatt (2009)</p> <p>Hoeijes (2010)</p> <p>Barbee (2010)</p>	<p>Boussion 2006,2009)</p> <p>Le Pogam (2008)</p> <p>Shidahara (2009)</p> <p>Le Pogam (2008, 2011, 2012)</p>	<p>Hoffman (1979)</p> <p>Kessler (1984)</p> <p>Links (1996)</p> <p>Avril (1997,2001)</p> <p>Geworski (2000)</p> <p>Vesselle (2000, 2004, 2007)</p> <p>Menda (2001)</p> <p>Hickeson (2002)</p> <p>Soret (2007)</p> <p>Jentzen (2008)</p> <p>Bural (2009)</p> <p>Srinivas (2009)</p> <p>Sakaguchi (2010)</p> <p>van Heijl (2010)</p> <p>Gallivanone (2011)</p>	<p>Chen (1999)</p> <p>Tylski (2007,2010)</p>	

Tableau III.1 : résumé des méthodes de correction de l'effet de volume partiel présentées dans le chapitre III et ayant été appliquées au moins une fois en oncologie.

III.5.4. Applications des méthodes de correction de volume partiel au suivi thérapeutique

La quantification absolue en TEP, comme celle proposée dans les protocoles les plus récents (Wahl et al., 2009), requiert de limiter au maximum les biais sur la mesure de l'activité métabolique. En ce sens, la correction de l'EVP, dont l'impact sur la mesure peut-être importante, paraît nécessaire pour évaluer la réponse à la thérapie. Cependant, en cancérologie à ce jour, et en dehors des travaux de (Maisonobe et al., 2011; Maisonobe et al., 2010), que nous présenterons au chapitre suivant, seules trois études ont cherché à établir l'influence de la correction d'EVP sur l'évolution relative du SUV entre deux examens et son potentiel impact sur la réponse thérapeutique.

- Mise en œuvre de différentes stratégies de correction en oncologie et impact sur l'évolution relative du SUV entre deux examens

En considérant différentes stratégies de correction d'EVP applicables en oncologie, (Hoetjes et al., 2010) ont cherché à évaluer leurs performances et leur impact sur les valeurs des SUV dans le contexte du suivi précoce. La comparaison portait sur 5 méthodes de correction :

- Deux méthodes de correction appliquées durant l'étape de reconstruction (Reader et al., 2003) (cf paragraphe III.5.1.1).
- Une méthode de CR prenant en compte la forme de l'objet par convolution d'un masque binaire (cf paragraphe III.5.2.2.1).
- Deux méthodes de déconvolution (VC (Teo et al., 2007) et LR (Tohka and Reilhac, 2008), cf paragraphe III.5.2.1.1).

Les trois méthodes de correction post-reconstruction ont été appliquées sur des données reconstruites par un algorithme OS-EM classiques i.e. ne corrigeant pas de l'EVP. La correction d'EVP portait sur une valeur du SUV normalisée par le poids du patient et corrigée du glucose plasmatique.

Comme nous l'avons vu au paragraphe III.5.2.1.1, les auteurs de cette étude ont montré que les deux méthodes de déconvolution conduisaient quasiment aux mêmes résultats, et ont donc limité la suite de leur étude aux résultats de la méthode de LR.

L'étude sur données simulées a permis de mettre en évidence que pour des volumes de sphères supérieures à 4 mL, toutes les méthodes comparées corrigeaient les valeurs d'activité mesurées avec une bonne précision. Quand les volumes utilisés pour la mesure de l'activité n'étaient plus mesurés à partir du VOI segmenté par une méthode semi-automatique appliquée sur les données TEP mais à partir des VOI exacts, le recouvrement de l'activité réelle était alors de plus de 80%, même pour une sphère de moins de 1 mL. Les données fantôme, qui contenaient un plus grand nombre de sphères de volume compris entre 1 et 10 mL, ont confirmé ce résultat pour toute sphère de plus de 1 mL.

L'application sur données patients visait à évaluer l'impact potentiel des corrections sur la mesure du SUV chez 15 patientes atteints de cancer du sein avant et après 1 cycle de chimiothérapie (pour un total de 18 lésions) ainsi que la variabilité test-retest des valeurs corrigées chez 10 patients atteints de NSCLC (30 lésions). En moyennant les évolutions relatives des SUV après un cycle de chimiothérapie, i.e. les Δ SUV (cf équation II.2) de toutes les lésions obtenues avant et après correction dans la première étude, les auteurs ont montré que seule la moyenne des Δ SUV corrigés par la méthode de CR n'était pas significativement différente de celle issue des valeurs non corrigées. Les auteurs soulignent cependant le faible écart entre la moyenne des valeurs non corrigées (environ -30,5%) et les valeurs corrigées par déconvolution ou durant le processus de reconstruction (comprises entre -26% et -27%). L'étude de la variabilité test-retest a montré une augmentation moyenne du SUV entre 2 TEP effectuées à 2 jours d'intervalle de 9,8% sans correction d'EVP contre 10,8% avec correction d'EVP, la différence entre ces valeurs étant non significative ($p > 0,17$).

L'analyse de ces résultats a permis aux auteurs de conclure que la correction d'EVP pouvait permettre de supprimer l'effet de la modification du volume sur la mesure de l'évolution de l'activité métabolique sans pour autant diminuer la précision de cette mesure. En fonction de la méthode, la correction pouvait changer significativement la valeur du Δ SUV et donc potentiellement la classification de la réponse. L'absence de différence entre la valeur moyenne des Δ SUV corrigées par CR et celle des valeurs non corrigées a été attribuée au fait que cette méthode repose sur l'utilisation d'un VOI, à la différence des 3 autres méthodes de correction étudiées. La difficulté de segmenter les données en TEP (Zaidi and El Naqa, 2010) peut expliquer l'impact potentiel de ce paramètre sur les résultats d'une méthode qui repose en grande partie sur cette étape.

(Gallivanone et al., 2011) ont proposé et évalué une méthode de correction d'EVP utilisable en clinique et basée sur la méthode du CR. Les valeurs de ces CR sont alors déterminées à partir de courbes obtenues via des données fantômes. De manière assez similaire à l'application sur données patient proposée par (Hoetjes et al., 2010), Gallivanone et al. ont ensuite cherché à évaluer la faisabilité de cette méthode sur les données FDG-TEP de 16 patients atteints de différents types de cancers (pancréas, sein, œsophage, ORL, etc), métastatiques ou non.

Pour cette deuxième étape, 24 lésions (9 primitives et 15 métastases) ont été considérées. Ces patients ont tous subi deux examens TEP : un avant le début de leurs traitements respectifs (chimiothérapie et/ou radiothérapie) et un examen de suivi (en moyenne 6 mois plus tard, intervalle de 2 à 10 mois). L'analyse des valeurs mesurée sur les TEP initiales a montré une modification significative entre les valeurs de SUV corrigées de l'EVP et les valeurs non corrigées (soit une augmentation de plus de 25% pour 52% des lésions segmentées par une méthode automatique sur la TEP).

En considérant l'évolution relative du SUV entre les deux examens (cf équation II.2), les auteurs ont observé une réduction de Δ SUV (maximal ou moyen dans la lésion) par rapport à

la valeur non corrigée dans 63% des cas et une augmentation dans le cas contraire. Une analyse statistique des résultats a montré une différence significative entre les différentes distributions de Δ SUV avant et après correction d'EVP par cette méthode ($p < 0,01$) suggérant l'importance potentielle de cette correction pour le suivi thérapeutique. Malgré le manque d'homogénéité de la cohorte et les différences entre les temps écoulés entre les 2 examens TEP de suivi, (Gallivanone et al., 2011) ont cherché à illustrer l'importance de ces changements de Δ SUV sur la classification de la réponse tumorale. En appliquant les seuils de réponse proposés par les critères EORTC utilisés pour le suivi en TEP (cf paragraphe II.3.1.2), les auteurs ont illustré la modification potentielle de la réponse thérapeutique sur une lésion (passant de maladie stable à maladie répondant à la thérapie après application de la correction d'EVP).

- Impact de la correction d'EVP sur la réponse thérapeutique précoce

Récemment, (Hatt et al., 2012b) ont évalué la valeur prédictive de différents index extraits de TEP acquises avant puis après une et deux semaines de traitement par chimiothérapie pré-chirurgicale chez 28 patients atteints de cancers rectaux localement avancés. L'objectif des auteurs était d'évaluer les capacités de différents index tels que le SUV_{max} , le volume métabolique (VM), le SUV moyen dans ce volume (SUV_{moyen}), le SUV_{peak} , la glycolyse totale de la lésion (i.e. le produit de VM par SUV_{moyen}) et celles des mêmes index corrigés de l'EVP (méthode mise en œuvre non détaillée dans le résumé du poster) à distinguer entre différents groupes de réponse établis à partir d'une analyse histopathologique.

Leurs résultats ont mis en évidence que seul le SUV moyen mesuré sur la TEP prétraitement était corrélé à la réponse anatomopathologique. En particulier, les auteurs ont remarqué qu'en utilisant la TEP acquise après une semaine, aucun des index étudié ne prédisait la réponse. En considérant les données TEP acquises après deux semaines de traitement, seul le VM ne prédisait plus la réponse. Dans ces conditions, l'AUC de la courbe ROC du TLG (0,79) était alors la plus élevée. La significativité de la supériorité du TLG par rapport aux autres index étudiés n'a cependant pas été spécifiée. (Hatt et al., 2012b) ont enfin relevé que la classification de la réponse thérapeutique, à partir d'une TEP prétraitement seule ou en considérant l'évolution des index métaboliques après deux semaines de traitement ne semblait pas dépendre de la correction d'EVP.

A la suite de ces résultats, il apparaît donc que si la correction d'EVP semble pouvoir modifier significativement l'estimation de l'évolution relative du SUV entre deux examens (Gallivanone et al., 2011; Hoetjes et al., 2010), son impact, quand il est significatif (Hoetjes et al., 2010), pourrait cependant ne pas améliorer la classification de la réponse par rapport aux index plus classiques (Hatt et al., 2012b). En outre, même si selon (Hoetjes et al., 2010), la correction d'EVP permet de supprimer le lien potentiel entre activité et volume métabolique, rien n'indique que les seuils proposés dans la littérature pour la classification de la réponse doivent être appliqués sans être ajusté en conséquence (Buvat et al., 2012). Cette question est

d'autant plus importante que les corrections appliquées et présentées dans ce chapitre reposent sur des hypothèses résolument différentes (cf tableau III.1).

III.6. Conclusion

Ce chapitre avait pour objectif de donner une vue d'ensemble des méthodes de corrections d'EVP en TEP, en mettant l'accent sur les méthodes applicables en oncologie. Nous avons vu que dans ce contexte, la difficulté de segmenter les volumes tumoraux sur l'imagerie anatomiques, du fait du faible contraste entre tissu sain et pathologique, était un frein à l'application de méthodes pourtant largement éprouvées en imagerie TEMP et TEP du cerveau.

Le chapitre suivant étudie l'impact de deux méthodes de correction de l'EVP sur la réponse thérapeutique précoce de patients atteints de cancer colorectaux métastatiques. Une étude préliminaire sur données simulées aura pour objectif de mieux comprendre le rôle respectif du volume métabolique et de l'activité dans le contexte du suivi thérapeutique.

Chapitre IV

Comparaison des capacités prédictives d'index métaboliques pour l'évaluation de la réponse tumorale précoce

IV.1. Introduction : faut-il corriger le SUV de l'effet de volume partiel pour évaluer la réponse tumorale?

Si les travaux proposés dans la littérature suggèrent un impact potentiel de la correction de l'effet de volume partiel (EVP) sur l'estimation de la réponse tumorale au traitement (cf paragraphe III.5.4), aucun ne statue sur son bénéfice pour la classification des patients en répondeurs ou non répondeurs. L'hypothèse que nous souhaitons valider dans le travail qui va être présenté ici était qu'une mesure corrigée de l'EVP, et donc à priori moins biaisée, permettait de mieux distinguer les tumeurs répondant des tumeurs ne répondant pas à la thérapie.

A partir de données simulées numériquement, nous avons d'abord tenté de mettre en évidence l'influence de la correction de l'EVP en fonction de la tâche de classification. Nous avons ensuite cherché à comparer nos résultats à ceux obtenus sur des données cliniques de patients atteints de cancer colorectaux métastatiques (CCM) traités par chimiothérapie et pour lequel nous possédions une évaluation de la réponse à long terme. Les 2 ensembles de résultats présentés dans ce chapitre ont fait l'objet de présentations orales aux congrès de la *Society of Nuclear Medicine* (Maisonobe et al., 2011; Maisonobe et al., 2010).

IV.2. Etude sur données simulées

IV.2.1. Matériel et méthodes

IV.2.1.1. Données simulées

L'implémentation et la validation de méthodes d'analyse d'images en TEP s'appuient sur différents types de données : les images simulées et celles de fantômes physiques, pour lesquelles on connaît *a priori* les valeurs à estimer, et les images de patients.

Les données simulées ont pour but d'imiter les images provenant d'un tomographe. Elles peuvent être obtenues à partir de simulations analytiques, dont nous ne détaillerons pas le principe ici, ou par des simulations Monte Carlo. Dans les deux cas, il est nécessaire de définir des données d'entrée, qui décrivent les propriétés physiques des objets pour lesquels on souhaite connaître la distribution d'activité vue par la TEP. Ces données contiennent entre autres la distribution d'activité à simuler, la distribution de densité et enfin un modèle du tomographe considéré ainsi que les paramètres d'acquisition associés. Une fois ces données définies, les simulations Monte Carlo permettent de générer des données brutes telles que détectées par les tomographes. Ces données brutes sont ensuite reconstruites pour obtenir les images simulées.

Le principe de base des simulations Monte Carlo est d'échantillonner des distributions à l'aide d'un grand nombre de tirages aléatoires. Ces nombres sont obtenus à partir de générateurs de nombres pseudo-aléatoires, i.e. des programmes générant des séquences de nombres ayant des propriétés se rapprochant le plus possible de celles de nombres aléatoires. Une fois déterminés, ils sont utilisés pour simuler les interactions de chaque particule dans la matière,

et en particulier, les processus intervenant dans la réalisation d'un examen de médecine nucléaire.

Parmi les simulateurs Monte Carlo utilisés pour simuler des examens TEP, on peut citer SimSET (Harrison et al., 1994) ou GATE (Jan et al., 2004) par exemple. C'est ce dernier qui a été mis en œuvre pour obtenir les images TEP utilisées pour cette étude. *GATE* (pour *Geant4 Application for Emission Tomography*) est développé depuis 2001 dans la collaboration *OpenGATE*, impliquant 17 laboratoires internationaux. Il est basé sur Geant4 (Agostinelli et al., 2003), un des principaux codes de simulation utilisé pour simuler le transport de particules à travers la matière. Geant4 a été développé initialement par le Conseil Européen pour la Recherche Nucléaire (CERN). Dans le contexte de l'imagerie TEP, GATE permet de simuler le transport de particules dans le patient mais aussi dans les détecteurs. De nombreuses caractéristiques des tomographes peuvent alors être prise en compte telles que la géométrie des cristaux qui les constituent ou encore leur temps mort par exemple.

Pour réaliser notre étude, nous avons utilisé des données simulées disponibles dans l'équipe. Elles modélisent des acquisitions réalisées avec un modèle validé du tomographe Philips Gemini GXL, dérivé de celui proposé par (Lamare et al., 2007). La prise en compte de toutes les caractéristiques du système de détection rend la modélisation très précise. Les distributions d'activité et d'atténuation ont été générées à partir d'un fantôme voxélisé NCAT (pour *NURBS CARDiac Torso phantom*, figure IV.1), qui permet de simuler les mouvements respiratoires et cardiaques du patient durant l'examen TEP.

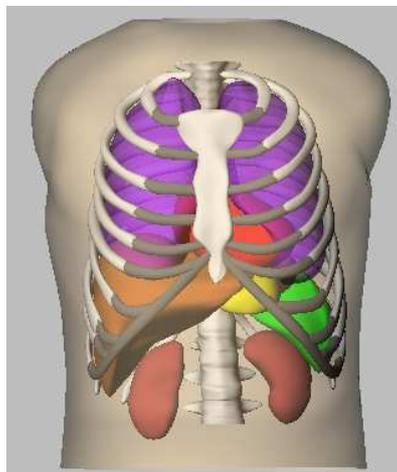


Figure IV.1 : rendu 3D du fantôme NCAT (*Nurb-based CARDiac-Torso phantom*) utilisé pour simuler des données réalistes (extrait de www.bme.unc.edu/~wsegars/)

Les tumeurs simulées ont été modélisées par des sphères d'activités homogènes placées dans différentes localisations : 4 dans le poumon droit, 2 dans le poumon gauche et une dans le foie (Figure IV.2). Le contraste entre l'activité de la sphère du foie (Figure IV.2a, flèche du bas) et celle du fond diffère de celui des autres sphères, cette dernière a été exclue de l'étude afin d'avoir un jeu de données homogène. Les sphères avaient un diamètre de 1, 2 et 3 cm.

L'activité dans les sphères a été fixée de façon à obtenir 7 valeurs de SUV différentes : {1; 1,5; 2; 3; 4; 8; 12}. Les données utilisées ici correspondaient à des acquisitions non synchronisées au signal respiratoire. Elles incluaient donc un flou cinétique important au niveau des tumeurs simulées (cf figure IV.2).

Au total, nous disposons de 21 examens TEP simulés (i.e. un examen contenant des tumeurs ayant toutes une des sept valeurs de SUV et une des trois valeurs de volumes possibles), avec, pour chacun d'entre eux, 7 tumeurs de caractéristiques identiques, mais variant en taille et/ou en activité d'un examen à l'autre. Finalement, un réplicat de ces images a été effectué : ce jeu de données correspond à la répétition des simulations précédentes mais avec une initialisation différente du générateur de nombres pseudo-aléatoires.

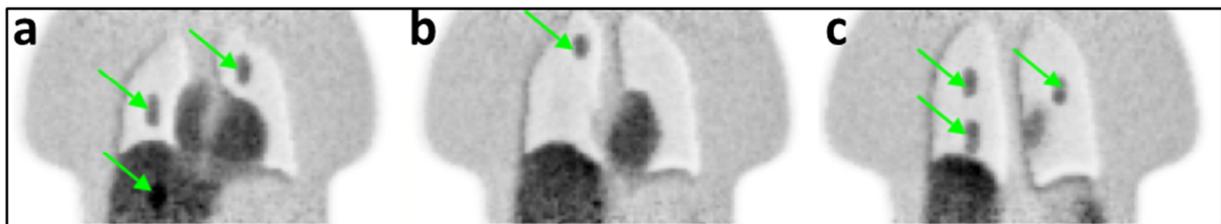


Figure IV.2 : Fantôme NCAT simulé avec GATE. Les sphères, représentant des tumeurs, sont indiquées avec les flèches vertes. Les images étant non synchronisées à la respiration. On remarque que le flou cinétique est important et modifie la forme des sphères. Extrait du rapport de thèse de (Tylski, 2009).

Les distributions étudiées, dont l'activité totale était comprise entre 43 et 45 MBq, ont été simulées afin d'obtenir des images équivalentes à celles obtenues lors d'une acquisition de deux minutes par pas de lit. Les images ont été reconstruites avec l'algorithme OS-EM 3D (5 itérations et 5 sous-ensembles), et corrigées de l'atténuation. S'agissant de données simulées, les coïncidences détectées sont parfaitement caractérisées. Il est donc possible de déterminer, sans erreur, si la coïncidence détectée est vraie ou diffusée. Dans notre étude, seules les coïncidences vraies ont été traitées, ce qui revient à supposer que des corrections parfaites de diffusion et de coïncidences fortuites ont été utilisées. La taille des voxels dans les images reconstruites est de $4 \times 4 \times 4 \text{ mm}^3$ et les images reconstruites ont été post-filtrées avec une gaussienne 3D de LMH égale à 4 mm.

IV.2.1.2. Caractérisation du métabolisme tumoral

Pour comparer la capacité prédictive des index à distinguer entre différentes évolutions possibles des volumes tumoraux simulés, nous avons calculé, pour chaque tumeur simulée, différentes estimées du SUV. Une étape préalable a consisté à optimiser les différents algorithmes de segmentation nécessaires à ces calculs. Cette optimisation a été réalisée dans le cadre du travail de thèse de (Tylski, 2009), en déterminant la valeur des paramètres minimisant, en moyenne, la différence entre le volume segmenté sur l'image TEP et le volume réel de la tumeur. Les données utilisées pour cette optimisation correspondaient à des examens TEP simulés réalistes, incluant des tumeurs non nécessairement sphériques, mais d'activités homogènes, localisées dans des examens TEP de sujets sains selon la méthode décrite par (Stute et al., 2011). Ces données avaient été reconstruites dans des conditions

identiques à celles des données NCAT présentées au paragraphe IV.2.1.1. Elles ont fait l'objet d'une publication par (Tylski et al., 2010).

Les SUV comparés dans cette étude sont les suivants :

- **SUV_{max}** : le SUV d'intensité maximale dans le volume tumoral
- **SUV_{peak}** : le SUV moyen dans un cube de 3 voxels de côté, centré sur le voxel de plus forte intensité. Cette définition, différente de celle proposée dans le protocole PERCIST par (Wahl et al., 2009), est considérée comme l'une des définitions envisageable de la région de volume fixe à utiliser (Boellaard, 2012).
- **SUV_{53%}** : il correspond au SUV moyen dans un volume segmenté à partir d'un algorithme de croissance de région et dont l'ensemble des voxels agrégés ont une intensité supérieure ou égale à 53% du SUV_{max}. La valeur de 53% utilisée ici est issue de l'optimisation décrite ci-dessus.
- **SUV_{fond}** : il correspond au SUV moyen dans une région obtenue par croissance de région et dont le seuil est défini en prenant en compte l'activité moyenne de la zone la plus active de la tumeur ainsi que l'activité du fond locale (Nestle et al., 2005). La définition précise de ce seuil S_{fond} est :

$$S_{fond} = A_{fond} + e_{opt} \times A_{70\%} \quad \text{Eq. IV.1}$$

où A_{fond} correspond à la moyenne de l'activité dans une région définie par une couronne de 2 voxels d'épaisseur placée à 8 mm de la région obtenue en utilisant un seuil égale à 40% du SUV_{max}. $A_{70\%}$ est égale à l'activité moyenne dans une région connexe où tous les voxels ont une activité supérieure à 70% du SUV_{max} ($VOI_{70\%}$) et e_{opt} est le paramètre optimisé à partir des images de tumeurs simulées réalistes décrites précédemment. La valeur utilisée ici pour e_{opt} est de 0,32.

- **SUV_{cr}** : qui correspond au SUV_{fond} corrigé de l'EVP par la méthode de coefficient de recouvrement (CR) proposée par (Kessler et al., 1984) (cf paragraphe III.5.2.2.1). Ce coefficient a été calculé en convoluant le masque binaire de chaque tumeur obtenu en utilisant la segmentation mettant en œuvre le seuil optimisé S_{fond} par la Fonction de Réponse (FDR) dans les images reconstruites. Une fois cette méthode de segmentation optimisée sur données simulées, la Largeur à mi-hauteur (LMH) utilisée a été optimisée en minimisant l'écart entre le SUV corrigé de l'EVP par la méthode de CR et la valeur réellement simulée. La valeur de LMH retenue suite à cette optimisation est de 8,9 mm.

La figure IV.3 résume les définitions des VOI utilisées pour calculer les différents index SUV comparés dans cette étude.

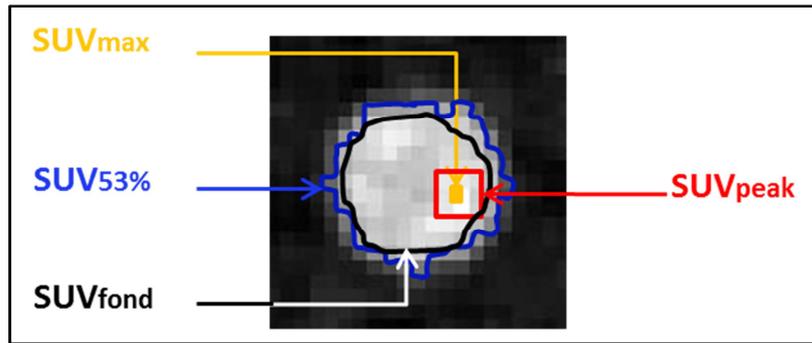


Figure IV.3 : illustration des contours utilisés pour calculer les différentes définitions de SUV comparés dans l'étude sur données simulées.

IV.2.1.3. Appariement des images

A partir des données NCAT, il est possible, pour chacune des 6 tumeurs, d'apparier les images pour simuler des évolutions de volume et/ou d'activité. La disponibilité de 2 réplicats pour chaque jeu de données permet aussi de simuler des cas où le volume ET l'activité réelle de la tumeur simulée ne changent pas. Pour évaluer la capacité des index SUV à distinguer différents types d'évolutions tumorales, nous avons apparié les données NCAT de deux manières distinctes :

- 1^{ère} tâche de discrimination : évaluation de la capacité des index SUV à distinguer les deux groupes de tumeurs suivants :
 - Tumeurs dont l'activité simulée est inchangée entre deux mesures, quel que soit le changement de volume (figure IV.4, ligne A). Il est possible de former 378 combinaisons (7 activités x 9 combinaisons de volumes possibles par sphère x 6 sphères) pour constituer ce groupe (G1)
 - Tumeurs pour lesquelles l'activité simulée change, quelle que soit l'évolution du volume. Ce groupe sera noté G2. Il est possible de former 2268 combinaisons pour former ce groupe (i.e. 42 évolutions de SUV x 9 combinaisons de volumes x 6 sphères).

- 2^{ème} tâche de discrimination : il s'agit d'évaluer la capacité des index à distinguer les 2 groupes suivants :
 - Un groupe G3 de tumeurs pour lesquelles l'activité simulée ET le volume sont inchangés (figure IV.4, ligne B). 126 cas (7 activités simulées x 3 diamètres x 6 sphères) constituent ce groupe.
 - Un groupe de tumeurs G4 pour lequel au moins un paramètre change. Ce groupe peut donc contenir des tumeurs dont le volume change mais pas l'activité, dont l'activité change mais pas le volume et enfin des tumeurs pour lesquelles activité et volume varient. Ce groupe est constitué de toutes les combinaisons constituant G2 (soit 2268 combinaisons) et de tous les cas pour

lesquels l'activité ne change pas mais où le diamètre change (soit 7 activités x 6 évolutions de volumes x 6 sphères, donc 252 cas). Au total, il est possible d'inclure 2520 combinaisons dans G4.

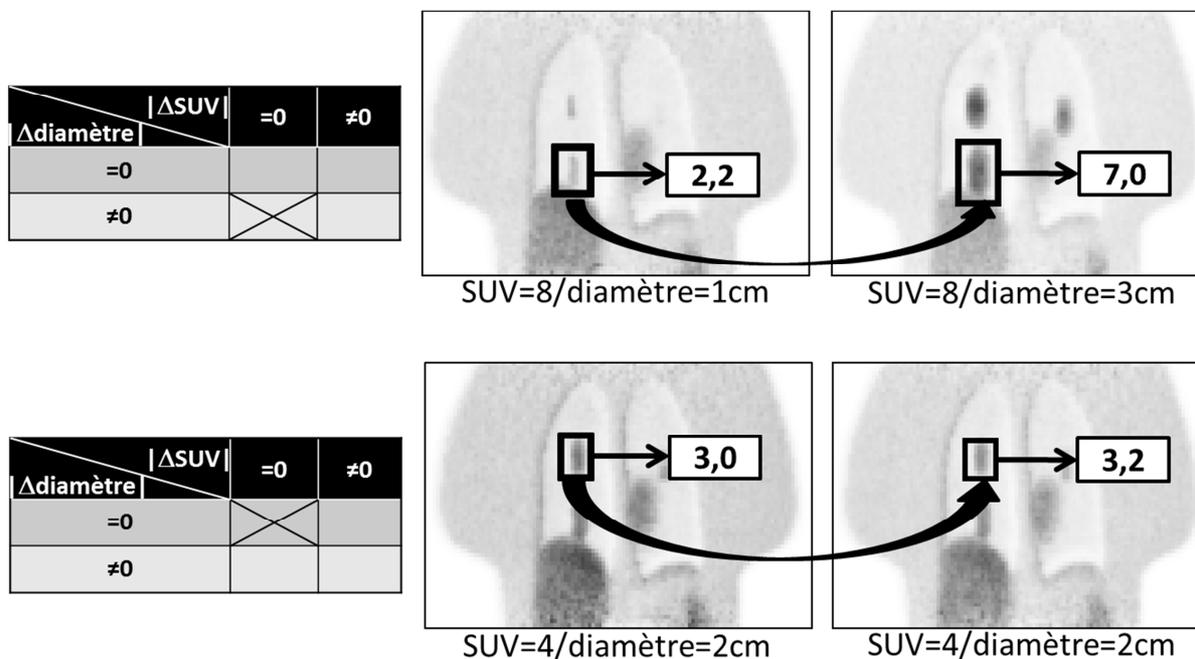


Figure IV.4 : exemple d'appariement de données utilisé pour simuler différents types d'évolution des tumeurs. Première ligne : cas d'une tumeur dont l'activité réelle reste égale à 8 (unités SUV) mais dont le diamètre augmente de 1 cm à 3 cm. Le SUV_{fond} mesuré passe alors de 2,2 à 7,0 g/mL. Ligne 2 : cas d'une tumeur n'évoluant pas entre les deux examens mais pour lesquelles deux simulations utilisant des séries de nombres aléatoires différentes ont été utilisées. Dans ce cas, le SUV_{fond} mesuré change légèrement (de 3 à 3,2.)

IV.2.1.4. Analyse statistique

IV.2.1.4.1. Caractérisation d'un test diagnostique : définitions

On caractérise souvent les performances d'un test diagnostique par ses *sensibilité* et *spécificité*. Cette approche repose sur une vision binaire du problème de diagnostic : les sujets sont sains ou malades et la maladie est détectée ou ne l'est pas.

Pour les sujets malades, on définit comme:

- Vrais Positifs (VP) les sujets dont la maladie est détectée par le test.
- Faux Négatifs (FN) les sujets dont la maladie n'est pas détectée.

Pour les sujets sains, on appelle :

- Vrais Négatifs (VN) ceux pour lesquels aucune maladie n'est détectée.
- Faux Positifs (FP) ceux pour lesquels le test détecte une maladie.

La sensibilité, aussi appelée Fraction de Vrais Positifs (FVP), est la capacité d'un test à donner un résultat positif si la maladie est présente. Elle est égale au rapport du nombre de vrais positifs sur le nombre de sujets malades de l'échantillon (équation IV.2).

$$\text{Sensibilité} \stackrel{\text{def}}{=} \frac{VP}{VP + FN} \quad \text{Eq. IV.2}$$

La spécificité est la capacité d'un test à donner un résultat négatif en l'absence de maladie. Elle est égale au rapport du nombre de vrais négatifs sur le nombre de sujets sains (équation IV.3).

$$\text{Spécificité} \stackrel{\text{def}}{=} \frac{VN}{VN + FP} \quad \text{Eq. IV.3}$$

La Fraction de Faux Positifs (FFP) est définie par :

$$\text{Fraction de faux positif} \stackrel{\text{def}}{=} \frac{FP}{VN + FP} = 1 - \text{Spécificité} \quad \text{Eq. IV.4}$$

Une autre façon de mesurer les performances d'un système est de déterminer la qualité des prédictions issues du système, souvent en calculant deux figures de mérite :

- la Valeur Prédictive Positive (VPP), qui représente la fraction de cas réellement positifs parmi tous les cas identifiés comme positifs par le test :

$$\text{Valeur prédictive positive} \stackrel{\text{def}}{=} \frac{VP}{VP + FP} \quad \text{Eq. IV.5}$$

- la Valeur Positive Négative (VPN), qui représente la fraction de cas réellement négatifs parmi les cas identifiés comme négatifs par le test :

$$\text{Valeur prédictive positive} \stackrel{\text{def}}{=} \frac{VN}{VN + FN} \quad \text{Eq. IV.6}$$

Contrairement aux sensibilité et spécificité, les valeurs prédictives dépendent de la prévalence de la pathologie, c'est-à-dire la proportion d'individus atteints par la maladie dans la population considérée. Elles mesurent donc essentiellement l'utilité diagnostique d'un système, mais pas sa justesse intrinsèque.

Considérons un échantillon représentatif d'une population pour laquelle nous savons que les individus qui la composent sont effectivement malades et un autre échantillon constitué de sujets sains. Pour une méthode diagnostique D fournissant, par exemple, une valeur d'autant plus grande que la maladie est exprimée, on peut tracer, pour chaque échantillon, un diagramme représentant la fréquence des valeurs fournies par la méthode D. Ce diagramme représente les distributions des mesures pour les deux échantillons, que nous avons représenté dans la figure IV.5 par 2 gaussiennes.

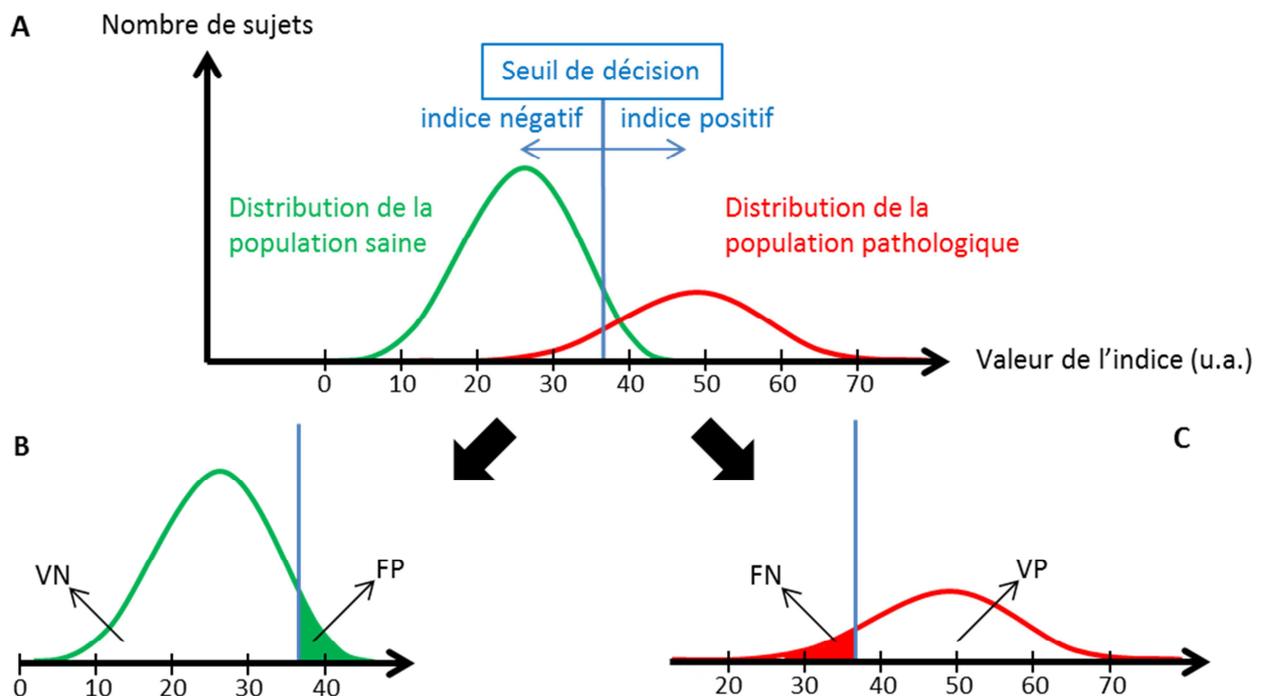


Figure IV.5 : Illustration des distributions des valeurs fournies par une méthode de diagnostic pour deux populations distinctes, dont les réponses sont connues a priori par une méthode de référence. En choisissant un seuil de décision (A), correspondant à une valeur donnée de l'indice évalué, on définit la proportion de Vrais Négatifs (VN) et Faux Positifs (FP) dans la population saine (B) et celle de Faux Négatifs (FN) et Vrais Positifs (VP) dans la population malade (C).

En définissant une valeur seuil au-delà de laquelle la valeur de l'indice mesurée chez un patient le classe comme malade, on peut en déduire une sensibilité et une spécificité. Ces deux grandeurs étant liées et évoluant dans des sens inverses, le choix du seuil correspond à un compromis. Si l'objectif de la méthode de diagnostic est, par exemple, de classer comme malade 100% des patients réellement malades (forte sensibilité), il est nécessaire d'utiliser un seuil très faible, ce qui induit une augmentation importante du nombre de faux positifs et donc une chute de la spécificité. Il est donc nécessaire de se doter d'un outil permettant de caractériser les performances de la méthode de classification binaire indépendamment du choix du seuil. Nous avons utilisé l'analyse ROC.

IV.2.1.4.1. Courbes ROC

Développée durant la seconde guerre mondiale par l'armée américaine pour améliorer la capacité des signaux radars à prédire la présence d'avions japonais, la courbe ROC (pour *Receiver Operating Characteristic*) a par la suite été utilisée en théorie de la détection du signal et en psychophysique (Green and Swets, 1966), puis en médecine pour l'évaluation de tests diagnostiques (Swets, 1988). Cette courbe représente la fraction de vrais positifs (ou FVP i.e. la sensibilité, équation IV.2) en fonction de la fraction de faux positifs (ou FFP i.e. 1-Spécificité, équation IV.4) pour l'ensemble des seuils possibles. Des exemples de courbes ROC sont données en figure IV.6. L'indice 1, dont la courbe ROC passe par le point de

coordonnées (0 ;1), correspond à une méthode de diagnostic parfaite car capable de séparer sans erreur les deux populations. L'indice 2 est caractéristique d'une méthode non informative i.e. purement aléatoire : il est incapable de distinguer les deux populations. L'un des intérêts de la courbe ROC est de permettre de comparer les capacités prédictives d'indices tels que les indices 3 et 4, pour lequel on constate ici que l'indice 4 présente une meilleure sensibilité et spécificité que l'indice 3.

Les exemples présentés ici reposent sur l'hypothèse que les distributions des valeurs des indices suivent des lois normales chez les deux populations: on parle alors d'approche paramétrique. Dans le cas où cette hypothèse n'est pas justifiée, ce qui peut être mis en évidence par différents tests statistiques dont le plus usité est celui de Kolmogorov-Smirnov, les courbes obtenues sont affines par morceaux : la courbe sera d'autant plus lisse que le nombre d'individus constituant les échantillons sera élevé.

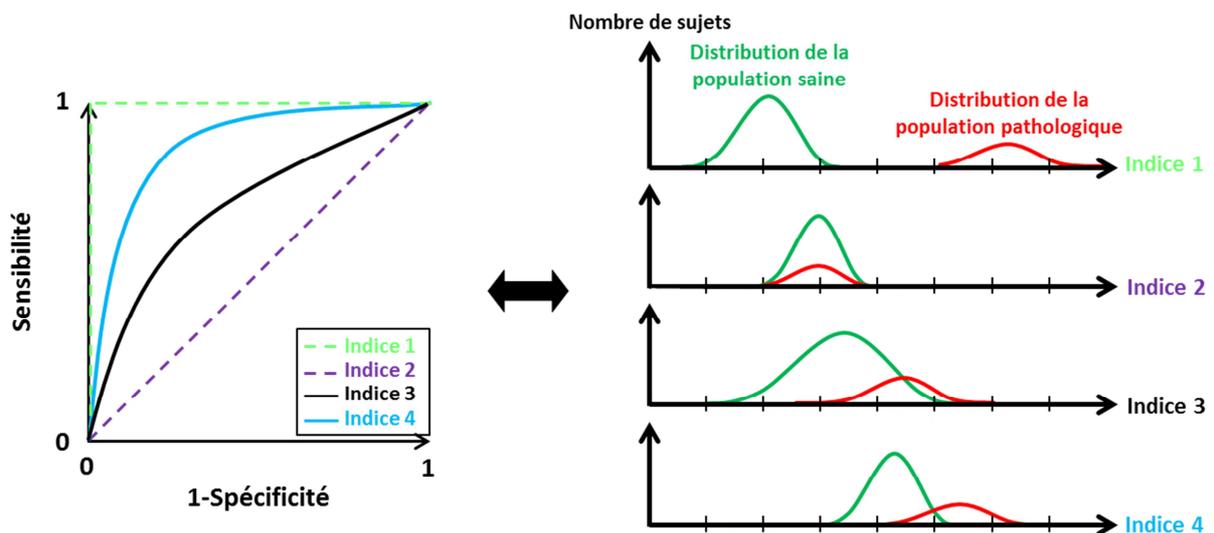


Figure IV.6 : Exemple de courbes ROC (gauche) et des distributions de valeurs correspondantes pour deux populations et quatre méthodes de diagnostic. Les indices 1 et 2 correspondent à des cas extrêmes où la méthode étudiée permet de distinguer totalement (indice 1) ou pas du tout (index 2) les différents groupes de patients.

Il est possible de comparer les performances de détection de chaque indice à l'aide de leurs Aires Sous la Courbe ROC, ou AUC (pour *Area Under the Curve*, en anglais). Cette AUC varie de 0 à 1. La valeur de 1 est atteinte dans le cas d'un test parfait (eg : indice 1, figure IV.6). A l'inverse, si les distributions des résultats obtenus pour les sujets malades et pour les sujets sains se superposent, la sensibilité est égale à la FFP pour toutes les valeurs du test : la courbe ROC d'un tel test est donc superposée à la première diagonale et son AUC est égale à 0,5. Un tel test n'a donc aucun apport diagnostique. Enfin, une AUC inférieure à 0,5 est obtenue pour un test dont la réponse est inversée, i.e. pour lequel un sujet effectivement malade présente un résultat inférieur à celui d'un sujet sain. Mathématiquement, l'AUC correspond à la probabilité d'identifier correctement l'image avec anomalie quand une image

avec et une image sans anomalie sont présentées simultanément à un observateur (Hanley and McNeil, 1982).

La fiabilité des courbes ROC dépend de la position du point considéré sur la courbe. Dans le cas d'une approche paramétrique, l'écart-type en tout point est :

$$\sigma_{FFP} = \sqrt{\frac{FFP(1 - FFP)}{Np^- - 1}} \quad \text{Eq. IV.7}$$

où σ_{FFP} est l'écart-type de la courbe ROC suivant l'axe des abscisses et Np^- correspond au nombre de cas sains considérés.

De même, l'écart-type de la courbe ROC suivant l'axe des ordonnées, σ_{FVP} , vaut :

$$\sigma_{FVP} = \sqrt{\frac{FVP(1 - FVP)}{Np^+ - 1}} \quad \text{Eq. IV.8}$$

en notant Np^+ le nombre de cas constituant l'échantillon pathologique considérée.

On déduit des équations 7 et 8 que les écart-types des points en abscisses et ordonnées des courbes ROC varient inversement avec le nombre de sujets sains et pathologiques considérés. Pour ne pas avoir un biais plus grand sur la fraction de faux (respectivement de vrais) positifs que sur celle de vrais (respectivement de faux) positifs, il est donc conseillé d'étudier des groupes de même taille. Pour cette raison, nous avons considéré dans cette étude l'ensemble des cas du groupe offrant le moins de combinaisons possibles (128 pour G1 et 378 pour G2) et tiré au sort un même nombre de combinaisons dans le groupe qui lui était comparé. L'étude a été répétée sur 100 échantillons tirés au hasard afin de s'assurer que les résultats présentés étaient représentatifs des différentes combinaisons possibles.

IV.2.1.4.1. Comparaison des AUC

Afin d'évaluer la capacité des différents index à séparer les groupes G1/G2 d'une part et G3/G4 d'autre part, nous avons effectué une analyse ROC non-paramétrique. Nous avons systématiquement calculé les AUC et écart-types associés, et avons comparé la significativité des différences entre toutes les paires d'AUC possibles. Les échantillons n'étant pas indépendants, nous avons pour cela utilisé la méthode proposée par (Hanley and McNeil, 1983). En considérant une erreur de type I de 5%, nous avons donc testé, pour chaque paire d'indices A et B, l'hypothèse H_0 : « il n'y a pas de différence entre l'AUC de l'indice A et celle de l'indice B ».

Afin de limiter l'inflation du risque de type I (c'est-à-dire de rejeter l'hypothèse H_0 alors qu'elle est vraie), liée à la réalisation de plusieurs tests statistiques, nous avons utilisé la correction de Bonferroni (Bland and Altman, 1995). Cette méthode consiste à diviser le seuil

de significativité utilisé par le nombre de comparaisons effectuées. Dans notre cas, ce nombre est égale au nombre de combinaisons de 2 indices parmi les 5 étudiés, i.e. 10. On considèrera donc que la différence entre les AUC étudiées est significative quand la valeur de p obtenue pour chaque comparaison de paires d'AUC sera inférieure ou égale à 0,005.

IV.2.2. Résultats

IV.2.2.1. Capacités des index à distinguer des variations de SUV

La première tâche de discrimination vise à établir quel index est le plus à même de séparer un groupe de tumeurs dont le SUV ne varie pas (G1) d'un groupe de tumeurs dont le SUV varie (G2), et ce quelles que soient les variations de volumes. Les courbes ROC des 5 index comparés sont montrées sur la figure IV.7.

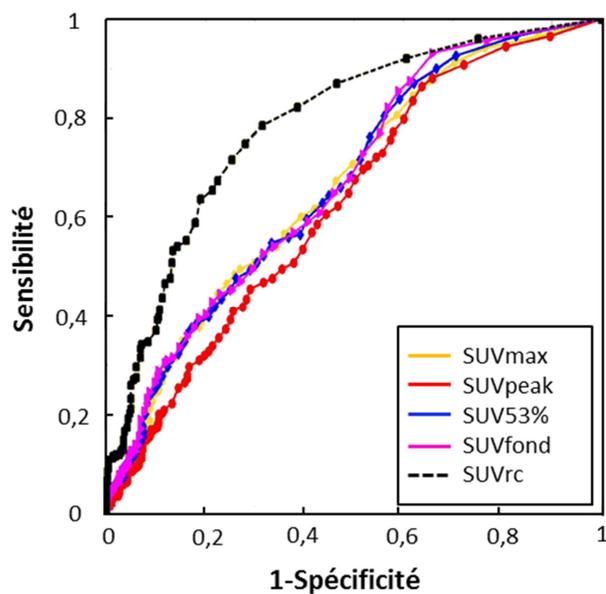


Figure IV.7 : courbes ROC des 5 indices pour la séparation des groupes de tumeurs G1 et G2 (N=378 lésions par groupe).

Les AUC et écart-types correspondant sont donnés dans le tableau IV.1.

Définition du SUV	AUC (σ)
SUVmax	0,66 (0,02)
SUVpeak	0,63 (0,02)
SUV53%	0,67 (0,02)
SUVfond	0,67 (0,02)
SUVrc	0,79 (0,02)

Tableau IV.1 : valeurs des AUC (1 écart-type) des méthodes comparées pour la séparation des groupes de tumeurs G1 et G2.

La comparaison deux à deux des AUC, en prenant en compte la correction de Bonferroni, met en évidence que la seule méthode significativement différente des autres est celle corrigée de l'EVP (SUV_{rc}, $p < 0,0001$). Même sans la correction de Bonferroni, aucune méthode non corrigée de l'EVP n'apparaît significativement différente des autres ($p > 0,1$).

IV.2.2.2. Capacité des index à identifier un groupe pour lequel aucun paramètre ne change

Dans un second temps, nous avons étudié la capacité des index à séparer un groupe de tumeurs pour lesquelles activité et volume sont inchangés (G3) d'un groupe pour lequel au moins un paramètre change (G4). Les capacités des 5 index à distinguer ces types d'évolution sont données sur la figure IV.8. Leurs AUC et écart-types associés sont donnés dans le tableau IV.2.

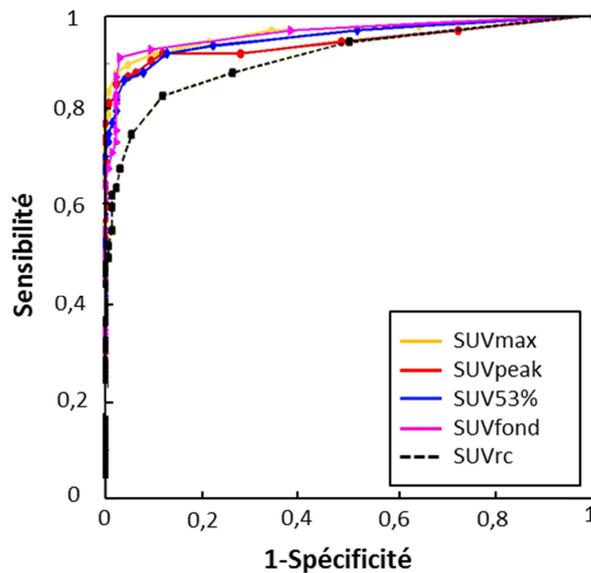


Figure IV.8 : courbes ROC des 5 indices comparés pour la séparation des groupes G3 et G4 de tumeurs (N=126 lésions par groupe).

Définition du SUV	AUC (σ)
SUVmax	0,96 (0,01)
SUVpeak	0,95 (0,02)
SUV53%	0,96 (0,01)
SUVfond	0,97 (0,01)
SUVrc	0,92 (0,02)

Tableau IV.2 : valeurs des AUC (1 écart-type) des méthodes comparées pour la séparation des groupes de tumeurs G3 et G4.

En ne prenant en compte que les paires d'AUC deux à deux, sans considérer le problème de comparaison multiple, seule une différence significative a été observée entre SUV_{rc} et SUV_{fond} avec une valeur de p de 0,03. La modification du seuil de significativité par la méthode de Bonferroni ne permet plus de distinguer une seule différence entre les aires sous les courbes ROC présentées dans le tableau IV.2.

IV.2.3. Discussion

L'objectif de cette étude sur données simulées était de déterminer dans quel cas la correction de l'EVP permettait d'améliorer la classification de l'évolution tumorale.

Les résultats obtenus en figure IV.8 montrent clairement que la correction de volume partiel mise en œuvre permet d'améliorer la classification quand l'objectif est de distinguer des lésions dont l'activité change de celles dont l'activité ne change pas, indépendamment de l'évolution du volume métabolique. Ce premier résultat laisse donc penser que si l'objectif est de déterminer un changement d'activité indépendamment de tout autre changement, alors la mesure du SUV doit être corrigée de l'EVP.

A contrario, quand l'objectif est de séparer un groupe de lésions dont au moins un paramètre change contre un autre où ni l'activité, ni le volume ne changent, alors la correction de volume partiel n'améliore pas la classification.

Ces résultats peuvent être expliqués par le fait que l'EVP rend le SUV dépendant du volume métabolique. C'est ce qui peut être observé sur la figure III.4, sur laquelle on constate que plus le volume de la sphère étudiée est faible, plus son activité mesurée est faible. Ainsi, si l'objectif de la mesure du SUV est de détecter un changement de volume et/ou de SUV, alors il semble préférable de ne pas corriger de l'EVP, car l'évolution du SUV non corrigé reflète à la fois des changements de volume et d'activité.

A l'opposé, si l'objectif clinique est de détecter uniquement un changement d'activité métabolique, alors cette étude montre que la correction d'EVP, qui permet de supprimer la dépendance de l'activité mesurée au volume métabolique, est nécessaire, et permet d'augmenter très significativement les performances de classification.

IV.2.3.1. Limites de l'étude

D'un point de vue méthodologique, le travail réalisé sur simulations présente des limites. Le premier est lié à la nature des données. Tout d'abord, comme nous pouvons le voir sur les figures IV.2 et IV.4, le fantôme NCAT utilisé conduit à des images très fortement affectées par le mouvement respiratoire. De plus, les distributions d'activité des tumeurs étudiées sont celles de sphères homogènes, loin des distributions d'activité réelles parfois observables sur données cliniques, ce qui rend complexe toute transposition directe des résultats à des données cliniques. Ce dernier point est d'autant plus critique que les données utilisées ne présentaient pas un niveau de bruit similaire à ceux des images cliniques, obtenus dans des conditions d'acquisition et reconstructions similaires. En effet, la suppression a priori des coïncidences fortuites et diffusées, plutôt que leur correction par des méthodes implémentées en clinique, rend les données un peu trop idéales. Ceci affecte probablement le SUV_{max} dont la sensibilité au bruit a été rapportée (Boellaard, 2012; Soret et al., 2007). La correction de l'EVP par coefficient de recouvrement (Kessler et al., 1984) mise en œuvre est en revanche bien adaptée au modèle de lésion considéré, ces dernières étant d'activité homogène.

IV.2.4. Conclusion

Si la correction de l'EVP permet de rendre les mesures d'activité indépendantes du volume des tumeurs, il reste à déterminer, en clinique, quel est l'objectif de la mesure d'activité. En mesurant le SUV, l'objectif est-il de détecter uniquement de modifications d'activité ou, au contraire, a-t-il pour but de mesurer des changements globaux de la tumeur ?

Nous avons cherché à répondre à cette nouvelle question à partir d'une base de données fournie par l'Institut Jules Bordet de Bruxelles.

IV.3. Comparaison d'index métaboliques pour l'évaluation de la réponse précoce au traitement : application à des cas de cancers colorectaux métastatiques traités par chimiothérapie

Le manque de consensus sur la définition de l'index métabolique le mieux adapté et le plus prédictif dans le contexte du suivi thérapeutique, ainsi que le grand nombre de métriques proposées dans la littérature, nécessitent que soient menées des études supplémentaires pour statuer sur les performances des différents index. Ce sont pour ces raisons que (Wahl et al., 2009) ont proposé, à titre exploratoire, de rapporter, en plus du SUL_{peak} , différentes valeurs telles que le SUL_{max} ou encore la charge métabolique totale (ou TLG pour *Total Lesion Glycolysis*, (Larson et al., 1999)). L'intérêt pronostique de ce dernier index, qui combine explicitement informations de volume et d'activité, a en effet été mis en évidence pour certaines localisations (e.g, (Cazaentre et al., 2010; Hatt et al., 2011) mais sa valeur ajoutée dans le contexte du suivi thérapeutique ne fait pas l'unanimité (eg, (Benz et al., 2008).

L'objectif de l'étude que nous allons présenter est de clarifier l'impact de l'EVP et de sa correction par différentes méthodes sur l'évaluation précoce de la réponse au traitement chez des patients atteints de CCM. A cette fin, nous avons comparé les performances de 8 indices :

- 5 SUV non corrigés de l'EVP.
- 2 index corrigés de l'EVP.
- Le TLG.

Nous avons utilisé les TEP/TDM acquises avant puis 15 jours après le début de la chimiothérapie de patients atteints de CCM et comparé leurs capacités à prédire la réponse définie par les critères RECIST 1.0 utilisant une mesure de l'évolution de la longueur des lésions cibles après 6 à 8 semaines de traitement.

IV.3.1. Matériels et méthodes

IV.3.1.1. Patients

Les données de 40 patients atteints de CCM traités à l'hôpital Jules Bordet de Bruxelles ont été utilisées pour cette étude. Ces patients ont été recrutés initialement pour une étude prospective dont le but était d'évaluer le rôle clinique de la TEP précoce. Cette étude a été approuvée par le comité d'éthique de l'Institut Jules Bordet et référencée sur clinicaltrials.gov sous le numéro NCT00741481. Les traitements suivis sont indiqués dans le tableau IV.3. Aucune thérapie ciblée (eg : chimiothérapie anti-VEGF ou anti-EGFR et pour lesquels les critères RECIST ne sont pas nécessairement adaptés) n'a été utilisée pour traiter ces patients.

Caractéristiques	N (%)
Nombre total de patients	
Sites des différentes lésions étudiées	
• Primaires	3
• Foie	70
• Poumon	12
• Péritoine	9
• Autres	7
Total	101
Classification RECIST 1.0	
• Réponse Partielle	27
• Maladie Stable	55
• Maladie Progressive	19
Age (en années)	
• Médiane	65
• Intervalle	17-83
Genre	
• Homme	23 (58%)
• Femme	17 (42%)
Lignes de chimiothérapie	
• Première ligne	29 (72%)
• Seconde ligne	11 (28%)
Schémas de traitements	
• FOLFOX	20 (50%)
• FOLFIRI	13 (33%)
• FOLFOX + bevacizumab	1 (3%)
• FOLFIRI + bevacizumab	4 (10%)
• FOLFIRI + panitumumab	1 (3%)
• Capecitabine	1 (3%)

Tableau IV.3 : caractéristiques des patients et lésions étudiés.

IV.3.1.2. Données TDM

La réponse thérapeutique a été déterminée en comparant deux jeux de TDM : un première TDM a été acquise sur un scanner hélicoïdal, avec ou sans injection de produit de contraste par voie intraveineuse (en fonction de la lésion) en moyenne 9 jours (de 0 à 26 jours) avant le premier TEP. Un second scanner d'évaluation a été acquis environ 6 à 8 semaines après le début du traitement, et plus tôt (3 patients, soit 5% des lésions considérées) en cas de suspicion de récurrence. L'épaisseur de coupe variait de 3 à 5 mm suivant le scanner utilisé. Les dimensions des lésions ont été mesurées par un radiologue expérimenté, après sélection des lésions en accord avec un médecin nucléaire. Les lésions cibles prises en compte dans cette étude ont été choisies suivant les critères RECIST 1.0 ((Therasse et al., 2000), cf paragraphe II.1.1.2), avec la restriction suivante : seules les lésions clairement identifiées à la fois sur la TEP/TDM de prétraitement et la TDM diagnostique ainsi que celles pour lesquelles le plus long diamètre anatomique était supérieure à 15 mm, ont été analysées. Chaque tumeur a été assignée à un des quatre types de réponse thérapeutique tardive : Réponse Complète (RC), Réponse Partielle (RP), Maladie Stable (MS) et Maladie Progressive (MP).

IV.3.1.3. Données TEP/TDM

Tous les patients ont subi un premier examen FDG-TEP/TDM, ou examen de référence, en moyenne 9 jours après le scanner de diagnostic (intervalle [0-26] jours). Quatorze jours après le début du traitement, un second examen TEP, ou *examen précoce*, a été effectué. Toutes ces données ont été acquises sur un même tomographe (TEP/TDM *GE Discovery LS*), 60 minutes (moyenne : 67 minutes, intervalle [52-88] minutes) après l'injection de 4 MBq/kg de FDG. Elles ont ensuite été reconstruites à partir d'un algorithme OS-EM intégré, avec 2 itérations et 28 sous-ensembles, puis post-filtrées à partir d'une gaussienne 3D de largeur à mi-hauteur de 5,45 mm. Ces images sont corrigées de l'atténuation à partir des données TDM et de la diffusion par convolution-soustraction (Bailey and Meikle, 1994). Les images TDM ont été acquises à partir d'un scanner hélicoïdal multi-détecteur (4 coupes) *LightSpeed (GE Medical Systems)*. La tension était de 120 kV et la gamme de courant variait de 30 mA à 200 mA, fixée par l'algorithme *Auto-mA (GE)*. Les autres paramètres d'acquisition du scanner étaient : un pitch de 1,5 et une vitesse de déplacement de la table de 15 mm par rotation. Les dimensions transverses de la matrice des images scanner étaient de 512x512 voxels (taille d'un pixel : 0,98 mm x 0,98 mm) avec une épaisseur de coupe de 5 mm. Les images TEP se composent quant à elles de 128x128 voxels en transverse (taille d'un pixel : 3,91 mm x 3,91 mm) avec une épaisseur de coupe de 4,25 mm.

Finalement, les images TEP ont été exprimées en unités SUV, en utilisant la normalisation par le poids du patient (cf équation I.17).

Afin de mieux visualiser les données mise en œuvre pour ce travail, nous résumons en figure IV.9 le protocole de déroulement des examens dans le temps.

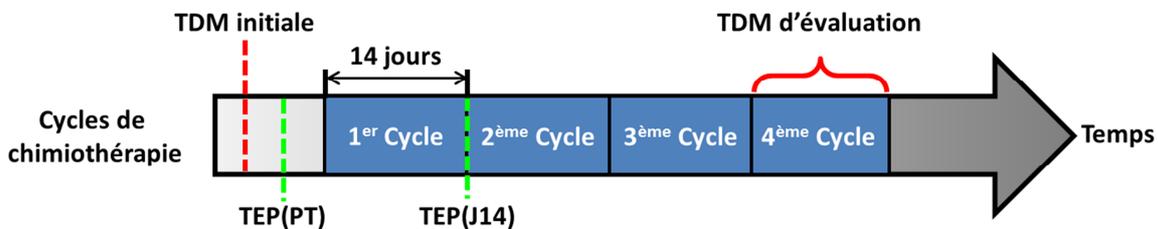


Figure IV.9 : frise chronologique indiquant les dates d'acquisition des données par rapport au traitement.

IV.3.1.4. Caractérisation du métabolisme tumoral

Huit index métaboliques ont été calculés pour caractériser l'activité tumorale vue par l'imagerie TEP au ^{18}F -FDG. Tous ces index ont été estimés à l'intérieur d'une large VOI dessinée manuellement, centrée sur la lésion, et incluant au moins 50% d'activité de fond. Si nécessaire, le volume a été ajusté afin que chaque VOI ne contienne qu'une seule région de fixation élevée. L'ensemble de ces segmentations manuelles a été effectué par un même opérateur, à partir d'une version recherche d'OWS (*Oncology WorkStation*) fournie par Dosisoft (version 1.0.0.2.8).

- Volume métaboliquement actif

Le volume métaboliquement actif des lésions (noté VM dans la suite quand il s'agira du volume calculé sur les lésions étudiées ici) a été obtenu à partir de la méthode de segmentation proposée par (Nestle et al., 2005) et décrite dans le paragraphe IV.2.1.2.

Pour assurer la connexité entre les voxels définissant le VM, nous avons considéré la plus grande région connexe obtenue à partir de l'application du seuil défini à l'équation IV.1, et avons systématiquement validé visuellement les régions résultantes.

L'activité de fond, A_{fond} , propre à chaque lésion, a été calculée à partir de la méthode décrite en IV.2.1.2. Pour éviter l'inclusion de voxels non caractéristiques de l'activité du fond, seuls des voxels appartenant à l'intersection de la VOI définie manuellement et de la région utilisée pour calculer A_{fond} ont été pris en compte.

Le paramètre e_{opt} (équation. IV.1) a été optimisé à partir de 3 acquisitions d'un même fantôme de Jaszczak composé de 6 sphères de volumes {0,5 ; 1 ; 2 ; 4 ; 8 ; 16} mL. Ces 3 acquisitions ont été effectuées sur le même tomographe et avec les mêmes paramètres de reconstruction que ceux utilisés pour obtenir les données patients. Seul le rapport entre activité de fond et activité dans les sphères a été modifié entre chaque acquisition. Ces rapports étaient de 2,96 :1 ; 5,88 :1 et 10 :1. Une image obtenue avec un contraste de 10 :1 est présentée sur la figure IV.10. Une valeur de e_{opt} égale à 0,30 a été obtenue en minimisant l'erreur moyenne entre le volume réel et le volume mesuré en utilisant S_{fond} , où l'erreur moyenne a été calculée en considérant toutes les sphères et tous les contrastes.

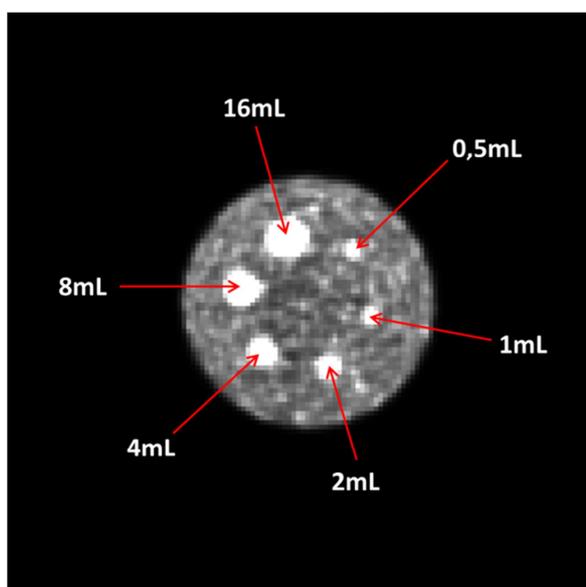


Figure IV.10 : coupe transverse de l'acquisition sur données fantômes de contraste le plus élevé (10 :1) utilisée pour optimiser la méthode de segmentation prenant en compte l'activité de fond (Nestle et al., 2005).

- SUV

Les capacités prédictives de 6 index SUV ont été évaluées, en incluant 4 index sans correction de l'EVP et 2 corrigés de l'EVP. Les définitions de 4 de ces 6 index ont été présentées dans le paragraphe IV.2.1.2. Il s'agit de SUV_{max} , SUV_{peak} , SUV_{fond} et SUV_{cr} . Pour ce dernier index, qui correspond à la correction de volume partiel de la mesure d'activité obtenue via SUV_{fond} , nous avons calculé le coefficient de recouvrement en supposant la fonction de réponse du système d'imagerie gaussienne, isotrope et stationnaire, de LMH égale à 7 mm. Cette valeur de la résolution spatiale a été estimée en minimisant l'erreur quadratique moyenne sur le volume des 18 sphères extraites des images TEP du fantôme de Jaszczak décrites ci-dessus (6 sphères x 3 contrastes). La correction de spill-in a été prise en compte en utilisant la valeur d'activité moyenne du fond local A_{fond} décrite au paragraphe IV.2.1.2.

En plus de ces 4 index, nous avons ajouté une autre mesure d'activité corrigée de l'EVP, obtenue en effectuant une déconvolution 3D basée sur l'algorithme de Van Cittert (VC) proposé par (Teo et al., 2007) (cf III.5.2.1.1), en utilisant 12 itérations et un paramètre de convergence égale à 1. La valeur moyenne d'activité corrigée SUV_{decon} a ensuite été calculée en utilisant la région utilisée pour calculer $A_{70\%}$ (cf équation IV.1). La valeur moyenne du SUV calculé dans cette dernière région, $SUV_{70\%}$, sans correction de volume partiel, a aussi été systématiquement relevée.

- Charge métabolique totale (TLG)

Le TLG de chaque lésion a été calculé comme le produit de VM par SUV_{fond} (Larson et al., 1999).

Pour chaque index et chaque lésion de chaque patient, nous avons ensuite calculé l'évolution relative entre les 2 examens TEP. Par exemple pour le SUV_{fond} , cette évolution relative est définie par :

$$\Delta SUV_{fond}(\%) = \frac{SUV_{fond}(J14) - SUV_{fond}(PT)}{SUV_{fond}(PT)} \times 100 \quad \text{Eq. IV.9}$$

où PT et J14 correspondent au moment d'acquisition des données TEP par rapport au début du traitement, c'est-à-dire avant le début du traitement (prétraitement) et après 14 jours de chimiothérapie (cf figure IV.9).

IV.3.1.5. Analyse de données

La moyenne, l'écart-type et l'intervalle de valeurs prises par l'ensemble des lésions ont été calculés pour chacun des 8 index au début et après 2 semaines de traitement. Pour évaluer l'accord entre les index, nous avons tracé les graphiques de Bland-Altman (Martin Bland and Altman, 1986) entre les valeurs des différents index mesurés sur la TEP prétraitement.

Le graphique de Bland-Altman, du nom de ses inventeurs, évalue la cohérence des résultats produits par deux méthodes d'estimation d'une même grandeur. Cette méthode est

particulièrement adaptée en l'absence de méthode de référence. Elle correspond à la représentation, pour chaque cas étudié, de la différence entre les estimées issues des deux méthodes en fonction de la moyenne des deux estimées. Sur le graphique ainsi obtenu, la moyenne des différences correspond au biais moyen entre les deux méthodes. Si on fait l'hypothèse que les différences suivent une loi normale, alors 95% des différences doivent être comprises entre cette valeur moyenne et $\pm 1,96$ fois son écart-type. C'est la raison pour laquelle on représente également ces deux lignes sur le graphique. Cette analyse ne permet pas de déterminer laquelle des deux méthodes évaluées est la plus juste mais plutôt à déterminer si les deux méthodes sont interchangeables. La figure IV.11 donne 4 exemples de graphiques de Bland-Altman, ainsi que leurs interprétations.

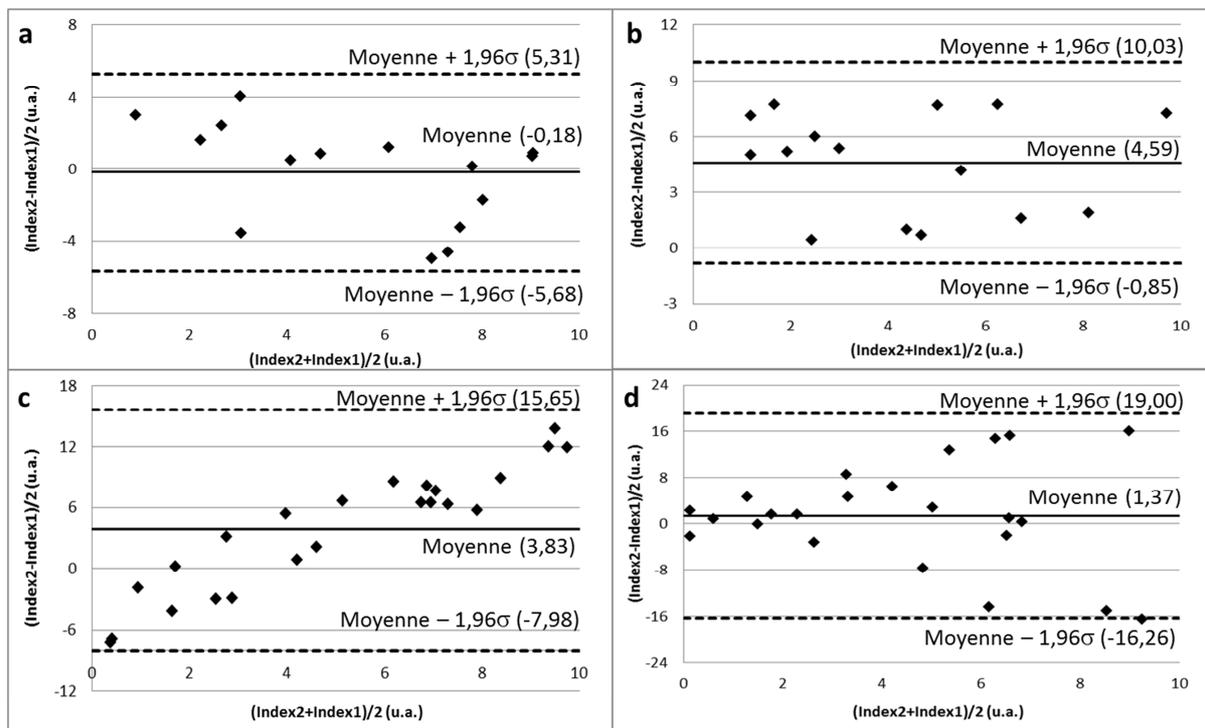


Figure IV.11 : exemples de graphiques de Bland-Altman. (a) : cas pour lequel un très bon accord moyen entre les méthodes évaluées est observé. (b) : exemple où il existe une erreur systématique absolue entre les 2 index comparés. (c) : cas de deux index dont l'écart varie linéairement avec l'amplitude de la mesure et (d) cas de deux index dont les valeurs diffèrent d'autant plus que la valeur mesurée est grande.

L'évolution relative moyenne des index métaboliques entre l'examen TEP PT et l'examen TEP acquis après 2 semaines a été calculée pour 2 groupes de lésions :

- i. GR : groupe de lésions correspondant à l'ensemble des lésions classées par les critères RECIST 1.0 comme Répondant Partiellement (RP) ou Complètement (RC).
- ii. GNR : groupe de lésions constitué des Maladies Stables (MS) ou Progressives (MP), toujours selon la classification RECIST 1.0.

Les distributions de valeurs des index étudiées diffèrent significativement de la distribution normale selon des tests de Kolmogorov-Smirnov. Nous avons donc évalué la significativité de la différence entre les médianes des distributions des évolutions relatives (équation IV.9) observées entre les groupes GR et GNR en utilisant le test de Wilcoxon, en considérant un niveau de significativité $\alpha=0,05$.

Pour comparer les capacités des 8 index à prédire la réponse à la chimiothérapie vue plus tard par RECIST 1.0, une analyse ROC non paramétrique a été effectuée (Hanley and McNeil, 1982) à partir des groupes GR et GNR. Ces courbes ROC ont été caractérisées par leurs AUC. La significativité des différences entre chacune des 28 paires d'AUC a été évaluée via une analyse bilatérale des variances par rangs de Friedman (Siegel and Castellan, 1988).

Enfin, nous avons étudié la variabilité de chaque index en calculant les coefficients de variations (noté CV, équation IV.10) de la différence absolue des index entre les 2 examens TEP (équation IV.11)

$$CV_{index} = \frac{\sqrt{\frac{1}{N_{MS}} \sum_{lésion=1}^{N_{MS}} (\Delta index_{lésion} - \overline{\Delta index})^2}}{\overline{\Delta index}} \quad \text{Eq. IV.10}$$

avec $\Delta index_{lésion}$ et $\overline{\Delta index}$ définis par :

$$\Delta index_{lésion} = |index_{lésion}(J14) - index_{lésion}(PT)| \quad \text{Eq. IV.11}$$

où

$$\overline{\Delta index} = \frac{1}{N_{MS}} \sum_{lésion=1}^{N_{MS}} |index_{lésion}(J14) - index_{lésion}(PT)| \quad \text{Eq. IV.12}$$

Le CV a été calculé seulement pour les N_{MS} tumeurs classées MS par les critères RECIST 1.0. Ces lésions sont en effet les seules pour lesquelles les variations des index entre les TEP PT et J14 sont supposées être négligeables, d'où le fait que les CV calculés représentent principalement la variabilité des index.

IV.3.1. Résultats

IV.3.1.1. Tumeurs

Chacun des 40 patients étudiés présentait un nombre moyen de 3 lésions (nombre qui variait entre 1 et 8 lésions). 101 lésions satisfaisaient les critères de sélection présentés au paragraphe IV.3.1.2. Parmi elles, 3 correspondaient à des lésions primaires, 70 étaient localisées dans le

foie, 12 dans les poumons, 9 dans le péritoine et 7 dans d'autres zones (cf tableau IV.3). Pour ces lésions, les critères RECIST ont conclu à 27 RP, 55 MS, 19 MP et aucune RC.

IV.3.1.2. Volume métabolique tumoral

Le VM tumoral moyen des 101 lésions étudiées variait, avant traitement, de 1,0 à 382 mL (médiane : 8,5 mL, moyen \pm écart-type : $34,4 \pm 66,4$ mL). La figure IV.12 montre la proportion de VM mesurés sur la TEP avant traitement plus petite ou plus grande que 5mL, pour l'ensemble des tumeurs ainsi que pour les lésions répondant et ne répondant pas à la thérapie séparément. Un test du χ^2 a montré que les distributions des deux groupes de volumes (plus petit et plus grand que 5 mL avant traitement) des tumeurs répondant et ne répondant pas à la thérapie n'étaient pas statistiquement différentes.

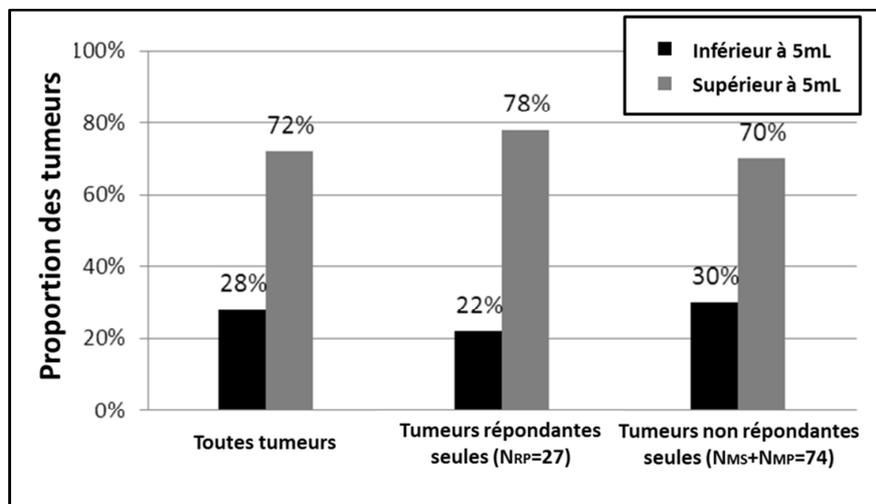


Figure IV.12 : Proportion de tumeurs dont le volume est inférieur ou supérieur à 5 ml pour l'ensemble des tumeurs étudiées (histogramme de gauche), seulement les lésions classées comme répondant à la thérapie (milieu) et seulement les lésions ne répondant pas (droite).

IV.3.1.3. Index métaboliques

Un résumé des statistiques des valeurs des différents index obtenues sur les données TEP est donné dans le tableau IV.4. Tous les index ont décliné significativement après 2 semaines de traitement (la comparaison entre les colonnes 2 et 3 a été effectuée à partir d'un test de Wilcoxon par rang bivarié avec un seuil de significativité de 0,05). Le test de Mann-Whitney, utilisé pour comparer les évolutions relatives de chaque index pour les deux groupes de tumeurs (colonnes 4 et 5), montre que les médianes de ces deux groupes sont significativement différentes pour tous les index sauf pour VM.

Index	TEP(PT)	TEP(J14)	Variation relative (%)	
			Répondantes	Non-répondantes
SUV _{fond} (g/mL)	6,1 ± 2,2 (1,7 – 12,4)	5,1 ± 1,9* (1,2 – 10,4)	-30,5 ± 15,2 (-72,4 – -8,4)	-6,7 ± 29,9* (-56,5 – 171,1)
SUV _{peak} (g/mL)	7,4 ± 3,4 (1,4 – 21,3)	6,0 ± 2,9* (1,1 – 18,0)	-33,5 ± 17,0 (-74,0 – -3,4)	-8,8 ± 37,1* (-65,2 – 250,4)
SUV _{70%} (g/mL)	8,3 ± 3,4 (2,1 – 18,5)	6,9 ± 3,1* (1,3 – 17,1)	-31,0 ± 18,3 (-79,4 – -1,5)	-7,3 ± 41,4* (-66,7 – 226,7)
SUV _{max} (g/mL)	10,5 ± 4,3 (2,5 – 22,3)	8,7 ± 3,9* (1,5 – 21,1)	-32,3 +/- 18,3 (-78,1 – 3,5)	-6,9 +/- 37,3* (-56,3 – 151,2)
SUV _{cr} (g/mL)	7,9 +/- 2,7 (2,5 – 14,6)	6,7 +/- 2,5* (1,7 – 12,9)	-29,3 +/- 18,2 (-76,6 – 0,2)	-6,3 +/- 30,0* (-56,3 – 151,2)
TLG (g)	252,4 +/- 559,3 (3,6 – 3644,9)	161,1 +/- 293,6* (2,2 – 1574,3)	-37,2 +/- 40,7 (-81,2 – 3,5)	3,3 +/- 109,9* (-76,6 – 834,1)
SUV _{decon} (g/mL)	11,7 +/- 4,7 (3,1 – 27,2)	9,8 +/- 4,5* (1,8 – 21,9)	-28,0 +/- 22,4 (-83,8 – 6,7)	-6,6 +/- 45,0* (-73,6 – 261,6)
VM (mL)	34,4 +/- 66,4 (1,0 – 381,5)	27,4 +/- 45,7* (1,0 – 262,4)	-4,9 +/- 61,4 (-77,1 – 218,2)	6,0 +/- 59,2 (-74,9 – 254,9)

Tableau IV.4 : valeurs moyenne \pm 1 écart-type (et intervalle de valeurs) pour les 8 index métaboliques étudiés pour l'ensemble des lésions sur les TEP (colonne 2) avant traitement et à J14 (colonne 3) (* : valeur de p inférieure à 0,05, caractéristique d'une différence significative entre les colonnes 2 et 3). Moyenne des variations relatives (\pm 1 écart-type et intervalle de valeurs) des index après 2 semaines de chimiothérapie pour le groupe de tumeurs répondant (colonne 4) et ne répondant pas à la thérapie (colonne 5, * : valeur de p < 0,05 entre les colonnes 4 et 5).

IV.3.1.4. Graphiques de Bland-Altman

La figure IV.13 montre les graphiques de Bland-Altman comparant les index métaboliques non corrigés de l'EVP (figures IV.13 a à f) ainsi que celui portant sur les 2 index corrigés étudiés (figure IV.13 g). La forte relation linéaire obtenue entre la plupart index (en dehors du graphique f) suggère que les valeurs comparées sont très corrélées. On note ainsi par exemple que SUV_{cr} est en moyenne égale à 70% de SUV_{decon} (figure IV.13h).

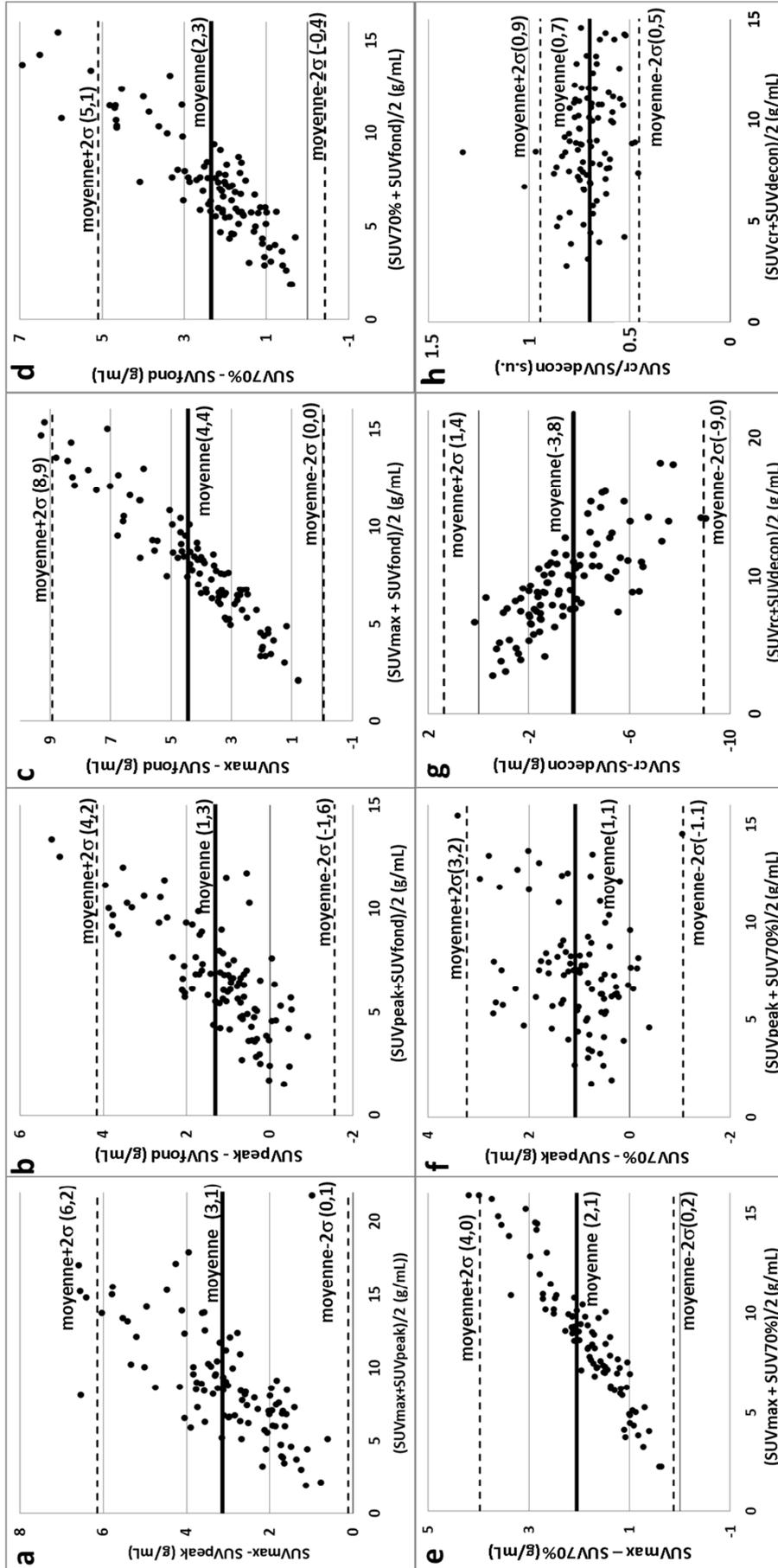


Figure IV.13: graphiques de Bland-Altman comparant les différents index non corrigés de l'EVP (a à f) et des 2 index corrigés (g) pour les 101 lésions étudiées avant le début du premier cycle de chimiothérapie. La différence moyenne entre chaque paire d'index ainsi que cette moyenne ± 2 écart-type (σ) sont présentés en lignes pleines et en pointillés, respectivement. Les valeurs correspondantes sont données entre parenthèses. Le graphique h représente les valeurs du rapport entre SUVcr et SUVdecon en fonction de la moyenne des différentes valeurs. Il illustre la corrélation entre les valeurs prises par ces deux index, également mise en évidence par la forte relation linéaire présente sur le graphique g.

IV.3.1.5. Analyse ROC

La figure IV.14 présente les courbes ROC décrivant la capacité des 8 index métaboliques à détecter la réponse à la thérapie. Les AUC (± 1 écart-type) de ces courbes ROC sont respectivement de : $0,81 \pm 0,04$ pour $\Delta\text{SUV}_{\text{fond}}$, $0,79 \pm 0,05$ pour $\Delta\text{SUV}_{\text{peak}}$, $0,77 \pm 0,05$ pour $\Delta\text{SUV}_{\text{max}}$, $0,77 \pm 0,06$ pour $\Delta\text{SUV}_{\text{cr}}$, $0,75 \pm 0,05$ pour $\Delta\text{SUV}_{70\%}$, $0,74 \pm 0,06$ pour ΔTLG , $0,69 \pm 0,06$ pour $\Delta\text{SUV}_{\text{decon}}$, et $0,58 \pm 0,07$ pour ΔVM .

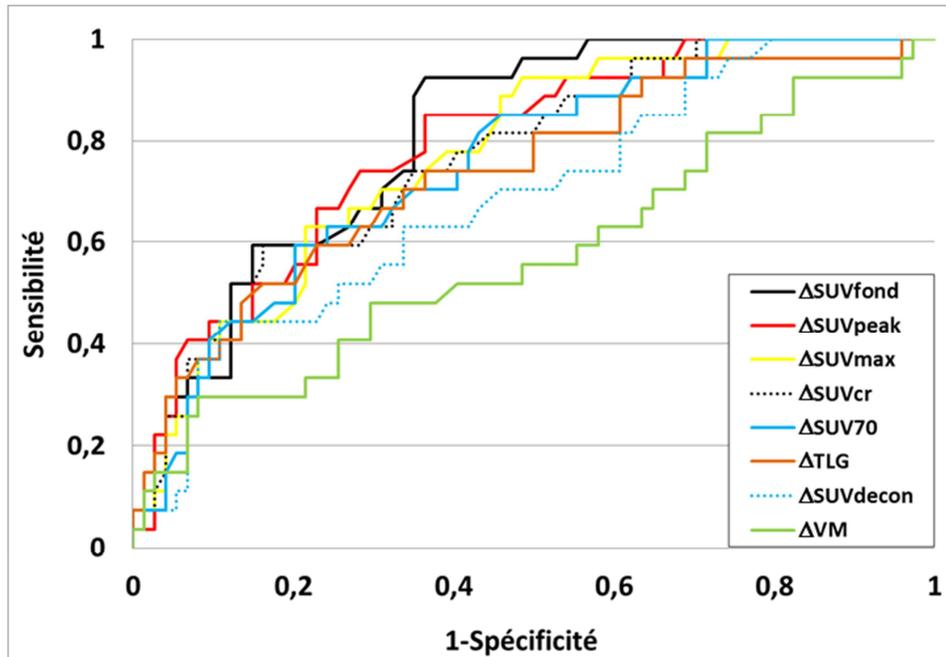


Figure IV.14 : Courbes ROC de $\Delta\text{SUV}_{\text{fond}}$, $\Delta\text{SUV}_{\text{peak}}$, $\Delta\text{SUV}_{\text{max}}$, $\Delta\text{SUV}_{\text{cr}}$, $\Delta\text{SUV}_{\text{decon}}$, $\Delta\text{SUV}_{70\%}$, ΔTLG et ΔVM calculés à partir d'une TEP/TDM acquise précocement après le début du traitement (J14) pour l'identification des lésions répondant ($N_{\text{RP}} = 27$) et ne répondant pas à la thérapie ($N_{\text{MS}} + N_{\text{MP}} = 74$) suivant la réponse RECIST 1.0 établie sur données anatomiques après 6 à 8 semaines de traitement.

Le test bilatéral d'analyse des variances par rangs de Friedman, non-paramétrique, a montré que ces AUC n'étaient pas toutes identiques. La comparaison des 28 paires d'AUC par une procédure prenant en compte la comparaison multiple et donc l'inflation du risque de type I a montré que seules les $\text{SUV}_{70\%}$, SUV_{fond} et SUV_{cr} avaient une AUC significativement supérieure à celle de VM ($p < 0,05$), alors que SUV_{max} , $\text{SUV}_{\text{decon}}$, SUV_{peak} et TLG n'étaient pas significativement différentes. $\text{SUV}_{\text{decon}}$ présente des performances assez médiocres, avec une AUC légèrement plus faible que celles des AUC associées aux autres indices SUV. Aucune comparaison entre paire d'index n'a montré d'AUC significativement différente des autres.

IV.3.1.6. Coefficients de variation

Les CV des différents index entre les deux examens TEP des 55 lésions stables est de 0,72 pour SUV_{fond} , 0,82 pour SUV_{peak} , 0,76 pour SUV_{max} et SUV_{cr} , 0,75 pour $\text{SUV}_{70\%}$, 1,00 pour $\text{SUV}_{\text{decon}}$, 1,90 pour TLG et 1,70 pour VM.

IV.3.1. Discussion

L'objectif de cette étude était de clarifier l'impact de la définition de l'index semi-quantitatif utilisé, sur des données cliniques, pour caractériser l'évolution du métabolisme glucidique de lésions cancéreuses à partir d'une TEP acquise avant le traitement et d'une TEP acquise 2 semaines après le début de la chimiothérapie. Nous avons en particulier cherché à évaluer la pertinence de la correction de l'EVP dans ce contexte.

IV.3.1.1. Interchangeabilité des index SUV

Toutes les définitions du SUV utilisées sont censées être représentatives de l'activité métabolique de la lésion étudiée. Il est donc intéressant d'évaluer si elles sont interchangeables.

Les graphiques de Bland-Altman (figure IV.13) démontrent des relations de proportionnalité entre les différents index. Par définition, le SUV_{max} est plus élevé que le SUV_{fond} ou le SUV_{peak} mais les figures IV.13a et IV.13c montrent aussi que plus le SUV est élevé, plus la différence entre les 2 indices est élevée. On constate de plus que le SUV_{peak} est en moyenne plus élevé que SUV_{fond} bien qu'il puisse être plus petit pour certaines tumeurs (cf figure IV.13b). En effet, pour de petites lésions, le SUV_{max} se situe parfois dans un voxel situé à proximité des bords du volume segmenté. Dans ces cas, le voxel de plus forte intensité est potentiellement entouré par des voxels de faible activité, non caractéristiques de la lésion mais de son voisinage direct. Un tel cas est illustré figure IV.15. La définition du SUV_{peak} utilisé ici peut alors conduire à inclure des voxels de faible activité dans la VOI_{peak} , mais pas nécessairement dans celle utilisée pour calculer le SUV_{fond} , ce qui explique alors que la valeur de SUV_{peak} puisse être plus petite que celle de SUV_{fond} . Bien que différente de la définition de VOI_{peak} proposée par (Wahl et al., 2009) dans les critères PERCIST 1.0 (cf. paragraphe II.3.2), notre définition fait partie des définitions utilisées dans la littérature (Boellaard, 2012). Dans leur article portant sur la comparaison des performances du SUV_{peak} en fonction du choix du VOI_{peak} , (Vanderhoek et al., 2012) ont montré que le facteur qui influençait la variation intra tumorale de ΔSUV n'était pas tant la position du VOI dans la tumeur que ses dimensions. Le choix de la méthode la plus adaptée pour former une sphère de 1 mL, comme proposé par (Wahl et al., 2009), dans une matrice de voxel de 4 mm de côté, reste cependant à préciser. On remarque enfin que plus l'activité de la tumeur est élevée et plus la différence entre les valeurs fournies par SUV_{peak} et SUV_{fond} est importante.

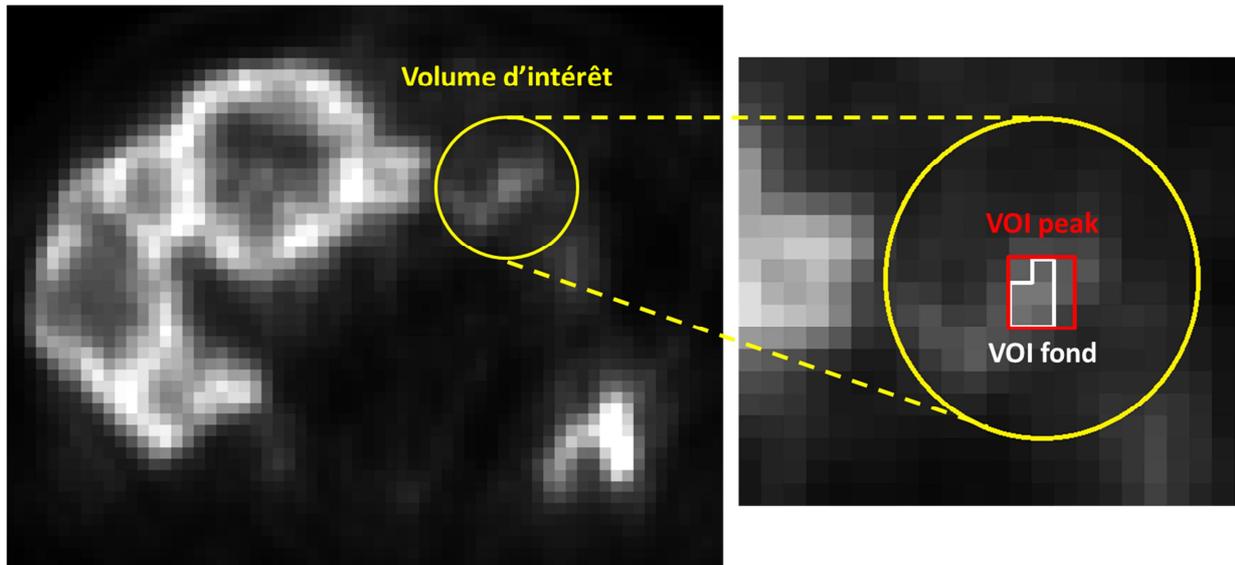


Figure IV.15 : cas d'une lésion pour laquelle la VOI utilisée pour calculer le SUV_{peak} inclus des voxels situés à l'extérieur du volume segmenté par la méthode de (Nestle et al., 2005). Il en découle une valeur de SUV_{peak} plus petite que celle du SUV_{fond},

La figure IV.13 montre aussi que pour l'ensemble des SUV non corrigés de l'EVP, la valeur obtenue dépend davantage de la définition de l'index utilisée quand l'activité de la lésion à estimer est élevée. Les fortes relations linéaires mises en évidence sur les graphiques de Bland-Altman (figure IV.13) suggèrent qu'en moyenne, un index peut être remplacé par un autre en utilisant un facteur d'échelle. C'est le cas par exemple pour les SUV_{CR} et SUV_{decon} (figure IV.13 h) pour lesquels il apparaît qu'en moyenne, le premier vaut 70% du second. De la même manière, SUV_{max} est en moyenne plus grand que SUV_{peak} de 46%, que SUV_{fond} de 71% et que SUV_{70%} de 24%.

Le coefficient de variation suggère que tous les SUV non corrigés de l'EVP ont des variabilités similaires, c'est-à-dire de 2,1 à 2,4 fois inférieures à celle du VM. Ces résultats vont dans le sens de ceux présentés récemment par (Cheebsumon et al., 2011a) portant sur la capacités de différents algorithmes de segmentation à reproduire les mêmes VOI sur des données TEP de patients acquises à très peu de temps d'intervalle (ou variabilité test-retest), avant le début du traitement. Ces auteurs ont en effet démontré que les variabilités des SUV moyen étaient jusqu'à 10 fois inférieures à celles des VM. Enfin, la variabilité du TLG est la plus élevée parmi tous les index étudiés.

IV.3.1.2. Corrections de volume partiel

Comme nous l'avons vu dans le chapitre III, l'EVP est la source de biais la plus importante quand il s'agit de caractériser l'activité tumorale des petites tumeurs (Soret et al., 2007). Dans notre cas, la résolution spatiale des images TEP reconstruites étant de ~ 7 mm, l'activité mesurée dans toute tumeur de moins de ~ 5 mL (i.e. volume équivalent d'une sphère de diamètre égale à 3 fois cette LMH) est fortement affectée de l'EVP. Dans cette étude, 28%

des lésions présentaient un VM plus petit que cette limite (figure IV.12). Une restriction de notre analyse, visant à ne prendre en compte que ces lésions, n'a cependant pas permis de démontrer de différence significative entre les capacités des différents index à distinguer entre la réponse et l'absence de réponse à la thérapie. Nous avons donc comparé les différents index en prenant en compte l'ensemble des lésions, après avoir pris soin de vérifier que les proportions de petits VM (<5 mL) et de VM plus grands (>5 mL) n'étaient pas statistiquement différentes entre les groupe de tumeurs GR et GNR.

Deux méthodes de correction de volume partiel post-reconstruction ont été évaluées. La première était basée sur la déconvolution d'images reconstruites par la méthode proposée par (Teo et al., 2007), utilisant une déconvolution de VC (cf paragraphe III.5.2.1.1). Comme nous l'avons vu, le nombre d'itérations à utiliser pour cet algorithme doit permettre d'atteindre la convergence de la valeur d'activité d'intérêt sans augmenter le bruit de l'image corrigée de manière trop forte importante. Nous avons vérifié, en termes de différences statistiques significatives entre AUC, que les résultats obtenus restaient inchangés quand quatre itérations, au lieu de douze, étaient utilisées pour l'algorithme de déconvolution de Van Cittert (VC). Un autre paramètre important de cette méthode de correction est la valeur de seuil, en pourcentage du SUV_{max} , utilisée pour définir le volume dans lequel est mesurée l'activité moyenne corrigée de l'EVP. (Teo et al., 2007) propose d'utiliser un seuil égal à 80% du SUV_{max} . Ce seuil s'est avéré mal adapté dans notre cas. Nous avons donc utilisé un seuil de 70%, identique à celui utilisé pour calculer $SUV_{70\%}$. Nous avons aussi implémenté la méthode de correction de Lucy-Richardson (LR) (Lucy, 1974) (cf paragraphe III.5.2.1.1) afin de corriger l'activité métabolique moyenne dans le $VOI_{70\%}$ (SUV_{LR}). La corrélation entre les valeurs de SUV corrigées obtenues par ces deux méthodes sur les données TEP prétraitement des 101 tumeurs s'est avérée très élevée (coefficient de détermination R^2 égale à 0,99), avec un écart moyen très faible (égale à 0,47 g/mL, cf graphique de Bland-Altman figure IV.16). Ceci suggère que les 2 méthodes de déconvolution sont interchangeable. En terme de performances de classification, les 2 méthodes de déconvolution se sont avérées comparables ($AUC \pm 1$ écart-type de $0,68 \pm 0,06$ pour la méthode de LR contre $0,69 \pm 0,06$ pour VC).

D'un point de vue quantitatif, nos résultats sont donc en accord avec les observations de (Hoetjes et al., 2010). Ces auteurs ont en effet rapporté avoir obtenu des valeurs de SUV corrigées quasi identiques en utilisant ces deux méthodes de déconvolution. Au contraire, (Boussion et al., 2009) rapportent de mauvaises performances pour la méthode de VC. Cependant, la comparaison entre les méthodes de VC et LR faite par ces auteurs porte sur des valeurs corrigées obtenues dans des volumes segmentés manuellement par des médecins, correspondant à des seuils sur le SUV_{max} compris entre 40 et 50%. Or les bords des tumeurs peuvent contenir des variations d'activité rapides caractéristiques du passage d'un tissu à un autre. Après déconvolution, ces zones peuvent inclure des artefacts de Gibbs, qu'il est important de ne pas intégrer dans l'analyse, qu'elle soit quantitative ou non. Pour cette raison, (Teo et al., 2007) ont proposé de corriger la valeur d'activité dans un VOI très petite autour du SUV_{max} . Ces auteurs ont montré que le choix d'un VOI trop grand (dans leur étude, plus grand que $VOI_{70\%}$) revient à prendre en compte des régions pour lesquelles la convergence n'est pas atteinte.

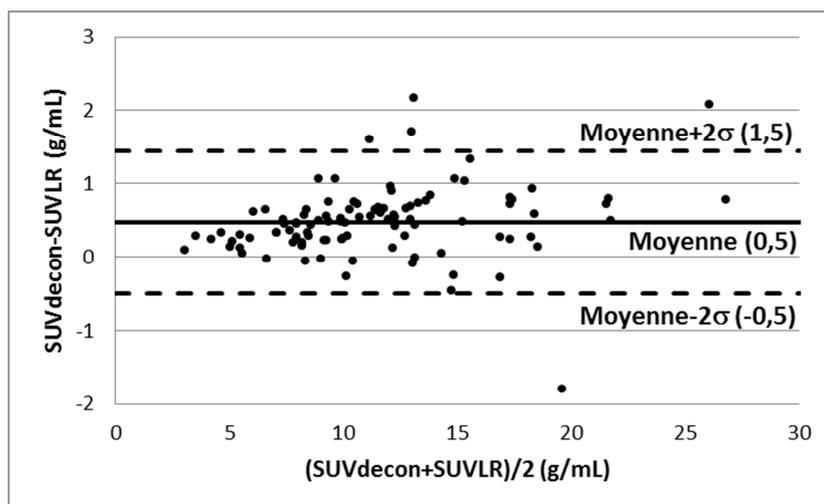


Figure IV.16 : graphique de Bland-Altman entre les valeurs de SUV corrigées par la méthode de déconvolution de VC (SUV_{decon}) et LR (SUV_{LR}) pour les 101 lésions sur la TEP avant traitement.

L'autre méthode de correction de l'EVP que nous avons utilisée est basée sur le CR (cf paragraphe III.5.2.2.1). Elle requiert la segmentation préalable de la lésion ainsi qu'une estimée de la résolution spatiale dans les images reconstruites. Pour la segmentation, nous nous sommes basés sur les performances de la méthode de segmentation de (Nestle et al., 2005), qui a été jugée équivalente à une méthode prenant en compte l'EVP pour déterminer le volume métabolique (Tylski et al., 2010). Concernant le choix de la LMH, les résultats se sont avérés robustes quand une valeur de 6 mm et une autre de 8 mm ont été utilisées au lieu de celle de 7 mm, utilisée pour les résultats présentés ici.

Comme observé dans les paragraphes III.3 et III.5.2.2, la mesure de la fixation de FDG est souvent corrélée à celle du volume à cause de l'EVP. Pour évaluer l'efficacité des corrections de l'EVP, nous avons donc calculé les coefficients de Pearson entre le VM et chacun des index SUV, à partir des résultats obtenus sur les données prétraitement et après 1 cycle de chimiothérapie. Des coefficients de corrélation plus faibles ont été trouvés pour les deux SUV corrigés de l'EVP (0,14 pour SUV_{cr} et 0,09 pour SUV_{decon}) comparés à ceux obtenus pour les index non corrigés de l'EVP (0,29 pour SUV_{fond} , 0,40 pour SUV_{peak} , 0,36 pour SUV_{max} et 0,23 pour $SUV_{70\%}$), suggérant que les corrections de l'EVP ont été efficaces. SUV_{decon} n'est pas apparu corrélé linéairement au VM ($p=0,21$) à la différence de SUV_{cr} . En évaluant plus précisément la différence entre SUV_{cr} et SUV_{decon} , on constate que SUV_{decon} était en moyenne plus élevé que SUV_{cr} (cf tableau IV.4 et figure IV.13h), ce qui, complété par les valeurs de corrélations au VM, suggère que SUV_{decon} est plus efficace que SUV_{cr} . Enfin, la valeur moyenne, pour toutes les tumeurs, du rapport entre SUV_{decon} , qui correspond à la valeur corrigé de l'EVP de $SUV_{70\%}$, et $SUV_{70\%}$ est de 1,43 (écart-type = 0,22). Celle du rapport entre SUV_{cr} et SUV_{fond} est de 1,33 (écart-type = 0,14). La proximité de ces deux valeurs confirme que les deux méthodes de corrections ont bien été efficaces et que les différences entre les

résultats obtenus étaient principalement dues à la nature de la région considérée pour la mesure.

IV.3.1.3. Variation relative des index entre les deux examens

Le seul index métabolique pour lequel la variation relative entre les deux examens n'était pas statistiquement différente entre les lésions répondant et celles ne répondant pas à la thérapie est le VM (cf tableau IV.4). Cette observation est cohérente avec les observations de (Cheebsumon et al., 2011a) montrant la forte variabilité test-retest des valeurs de volumes métaboliques, contrairement aux mesures d'activité qui en sont dérivées. En outre, la chimiothérapie induit une diminution lente du volume métabolique, avec une diminution probablement nulle ou faible du volume tumoral après 2 semaines (un cycle) de chimiothérapie chez les lésions répondant à la thérapie. Ceci explique au moins partiellement pourquoi, dans cette étude, la correction de l'EVP n'améliore la valeur prédictive des index extraits des données TEP au FDG acquises durant le traitement.

Les courbes ROC (figure IV.14) suggèrent que tous les index SUV ont des performances similaires pour distinguer la réponse de la non réponse tumorale, telle qu'obtenue ultérieurement par les critères RECIST 1.0, en dehors de celles du SUV_{decon} qui conduit à une courbe plus proche de la diagonale. Les courbes ROC montrent aussi que les changements de VM apportent une information permettant de discriminer les deux types de réponse, puisque l'AUC est significativement supérieure à 0,5 ($p < 10^{-5}$). Même si VM est bien moins informatif que les index dérivés du SUV (en dehors de SUV_{decon}), la suppression de l'information relative au volume métabolique, contenue dans les index non corrigés de l'EVP, pourrait être préjudiciable. C'est ce qui est observé lorsque l'on compare les AUC obtenus avec SUV_{cr} et SUV_{decon} à celles obtenues avec les index non corrigés de l'EVP (figure IV.14). En particulier, SUV_{decon} , qui correspond au SUV le mieux corrigé de l'EVP, a des performances de classification plus mauvaises que celles des index non corrigés. Ces mauvaises performances pourraient aussi être expliquées par la forte variabilité de SUV_{decon} , comparées à celles des autres index SUV non corrigés de l'EVP, mise en évidence par l'analyse des CV. Cependant, le CV du TLG était bien supérieure à celui du SUV_{decon} , avec une AUC pourtant supérieure (0,74). Ceci suggère que les faibles performances de classification de SUV_{decon} ne peuvent pas être expliquées par la seule variabilité de l'index.

La comparaison de SUV_{cr} et SUV_{fond} seuls (en ne prenant pas en compte les comparaisons multiples entre les 8 index), qui correspond à 2 index calculés à partir du même VOI mais avec et sans correction d'EVP, montre que la correction d'EVP réduit de manière significative l'AUC décrivant les performances de discrimination des index ($p=0,02$). Ce dernier résultat suggère qu'en supprimant l'information de volume implicitement contenue dans le SUV_{fond} du fait de l'EVP, le SUV_{cr} apporte moins d'information, au sens de la réponse tumorale globale, que quand l'information de volume est implicitement incluse dans la mesure d'activité non corrigée de l'EVP. Ce résultat est confirmé par la différence significative observée en ne comparant que les AUC des $SUV_{70\%}$ et SUV_{decon} entre elles ($p=0,03$).

Au regard des résultats obtenus sur données simulées et présentés au IV.2.2, il apparaît donc que la mesure d'un unique paramètre métabolique tel que le SUV non corrigé de l'EVP soit plus informative en clinique que la mesure d'un SUV corrigé de l'EVP, car elle reflète n'importe quel type de changement, qu'il soit liée à l'activité métabolique ou au changement de volume seul.

Nous avons aussi cherché à évaluer si un autre index, correspondant au produit du volume métabolique seul et de l'activité corrigée de l'EVP, permettait de mieux distinguer la réponse tumorale au sens de RECIST 1.0 que le TLG combinant le VM et un SUV non corrigé de l'EVP. L'index défini par le produit du VM et de SUV_{decon} a conduit à une AUC de $0,70 \pm 0,07$, alors que l'AUC pour le produit du SUV_{CR} et du VM était de $0,75 \pm 0,06$. Aucune de ces deux valeurs n'est apparue significativement supérieure à celle du TLG défini par (Larson et al., 1999), ce qui suggère que ces différentes définitions du TLG ne permettent pas d'améliorer la classification entre tumeurs répondant et tumeurs ne répondant pas à la thérapie.

Comme le changement de VM contient une information utile pour la classification de la réponse tumorale (AUC significativement supérieure à 0,5), nous avons examiné la classification basée sur un graphique en 2D présentant sur un premier axe la variation relative de l'activité corrigée de l'EVP (ΔSUV_{decon}) et sur un second celle de l'évolution de volume métabolique (ΔVM) (figure IV.17). Nous n'avons cependant pas réussi à améliorer la classification de la réponse tumorale à partir de ce graphique : l'inclusion de l'information de volume à partir d'un seul index non corrigé de l'EVP apparaît plus robuste qu'en considérant conjointement le changement de volume et celui du SUV corrigé de l'EVP.

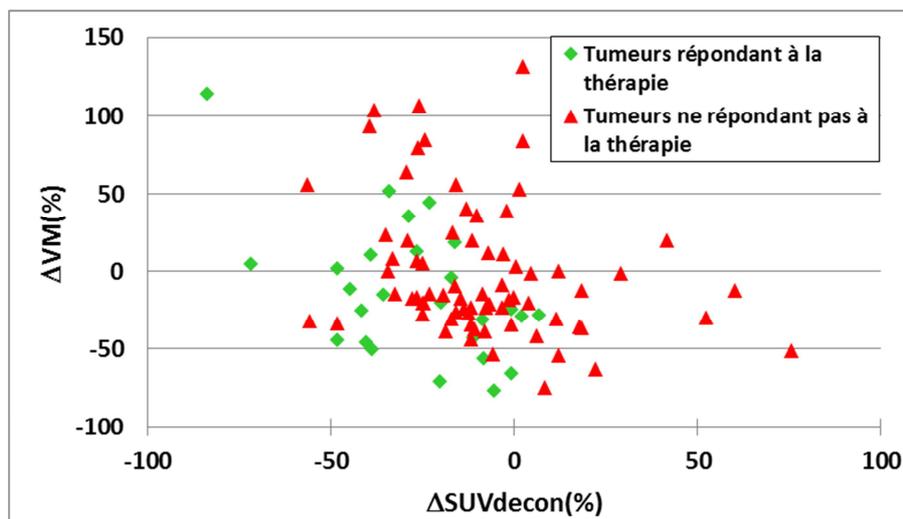


Figure IV.17 : graphique 2D représentant la variation relative de SUV_{decon} (axe des abscisses) et de volume métabolique (axe des ordonnées) entre les deux examens TEP.

IV.3.1.4. Limites de l'étude

Dans cette étude, nous avons cherché à évaluer la possibilité d'évaluer la réponse au traitement à partir d'une TEP acquise après 2 semaines de traitement en utilisant comme référence la réponse tumorale vue par les critères RECIST 1.0 sur des données TDM acquises après 6-8 semaines de traitement. L'utilisation des critères RECIST comme référence pour établir la réponse tumorale métabolique précoce n'est cependant pas idéale. Dans ce contexte, une TEP tardive aurait été un meilleur moyen de valider la capacité d'une TEP précoce à prédire la réponse à long terme à la thérapie. Des études supplémentaires, similaires à celles menées par (Vesselle et al., 2007; Vesselle et al., 2004), ou (Soussan et al., 2012) par exemple, et portant sur l'intérêt de la correction d'EVP pour évaluer la survie globale et la survie sans progression, sont donc nécessaires. Les index calculés sur les TEP précoces et présentant les AUC les plus élevées correspondent donc aux index prédisant mieux la réponse vue 4 à 6 semaines plus tard par la TDM. Nous n'avons pas cherché ici à prédire la réponse à partir de la TEP précoce seule, comme cela a pu être proposé par exemple par (Tixier et al., 2011) sur les cancers de l'œsophage, mais seulement à évaluer précocement (après 14 jours de traitement) la réponse vue à plus long terme par la TDM. Cette approche sera étudiée dans le chapitre suivant, dans lequel nous avons mis en œuvre la même cohorte de patients.

Parmi les 101 lésions étudiées, notre échantillon contenait 75% de lésions ne répondant pas à la thérapie au sens de RECIST. 1.0. Ceci implique que nos résultats surestiment probablement la spécificité (cf paragraphe IV.2.1.4.1) des méthodes et donc que les courbes ROC comparées ici tendent à présenter un biais en faveur d'un déplacement des courbes vers la diagonale. Cependant, le mauvais équilibre entre le nombre de tumeurs répondant et ne répondant pas à la thérapie est pris en compte dans l'analyse statique effectuée et ne biaise donc pas la comparaison entre les différents index.

Cette étude portait principalement sur l'évaluation précoce de la réponse tumorale. Le rôle de la correction de l'EVP à un stade plus avancé de la thérapie, c'est-à-dire quand la décroissance du volume métaboliquement actif est plus important pour les tumeurs qui répondent, nécessite aussi d'être évalué. De même, l'évolution de la fixation du FDG au cours du traitement peut dépendre de la thérapie utilisée. Par exemple, l'imatinib (une thérapie ciblée utilisée notamment pour le traitement des tumeurs stromales gastro-intestinales), induit une décroissance de la fixation tumorale de FDG dans une période de quelques heures à quelques jours après le début du traitement alors que la thérapie endocrine, tel que le tamoxifène (un modulateur sélectif des récepteurs des œstrogènes utilisé pour le traitement des cancers du sein) implique une augmentation de la fixation de FDG dans le même intervalle de temps (Contractor and Aboagye, 2009).

Enfin, nos résultats ont été obtenus pour un type de lésions particulier, chez des patients atteints de cancers colorectaux métastatiques. Leur généralisation à d'autres localisations reste donc à confirmer mais sont cependant en accord avec les résultats récents de (Hatt et al., 2012c) sur des cancers rectaux avancés (cf paragraphe III.5.4). De plus, il serait intéressant d'étudier, dans le même contexte, les capacités prédictives d'autres informations caractéristiques du métabolisme des lésions telles que celles fournies par l'analyse de texture proposée par exemple par (El Naqa et al., 2009). En effet, des résultats récents, portant sur des

cancers de l'œsophage (Tixier et al., 2011), ont démontré que ces index distinguaient mieux les lésions répondant des lésions ne répondant pas à la thérapie que les index classiques quand ils étaient mesurés sur la TEP avant le début du traitement. Cette question sera abordée dans le chapitre suivant.

IV.3.2. Conclusion

Chez 40 patients atteints de cancers colorectaux métastatiques, pour lesquels 101 lésions ont été étudiées (dont 28% faisaient moins de 5mL), nous avons mis en évidence que le SUV_{max} , le SUV_{fond} , le SUV_{peak} le $SUV_{70\%}$ et le TLG calculés sur une TEP précoce (c'est-à-dire 2 semaines après le début du traitement) et comparés à leur valeur de référence, mesurées sur une TEP réalisée avant le traitement, prédisaient la réponse tardive vue par les critères RECIST 1.0 sur données TDM après 6 à 8 semaines de thérapie. La mesure de l'activité métabolique par des SUV corrigés de l'EVP n'a pas amélioré la distinction entre lésions répondant et lésions ne répondant pas à la thérapie, probablement car la correction de l'EVP supprime l'information relative à l'information de volume métaboliquement actif. Le changement de volume métabolique seul, entre le début du traitement et après 2 semaines de chimiothérapie, prédit mal la réponse tumorale vue, à plus long terme, par la TDM, probablement du fait de la forte variabilité des estimées de volume métabolique.

Chapitre V

Caractérisation de l'hétérogénéité mesurée en TEP-FDG et intérêt dans le contexte du suivi thérapeutique

V.1. Mesure d'hétérogénéité tumorale en TEP : motivations

Pour caractériser le métabolisme tumoral, l'état de l'art actuel propose, dans le meilleur des cas, de déterminer le volume et/ou le SUV moyen ou maximal d'une population de voxels définis comme caractéristiques de la lésion étudiée. Cependant, la simple analyse visuelle de la fixation de FDG visualisée par la TEP (exemple figure V.1) suggère que la distribution intra-tumorale de radiotracer peut être extrêmement hétérogène.

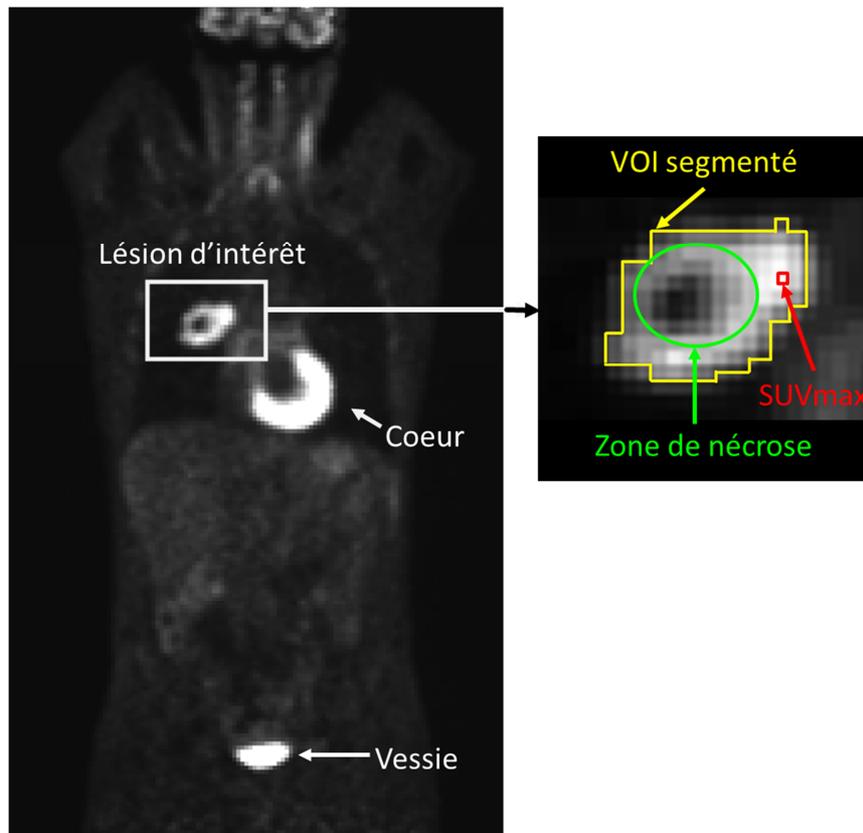


Figure V.1 : coupe coronale de la TEP/TDM au FDG d'un patient atteint d'un cancer du poumon non à petites cellules présentant une importante nécrose centrale. Extrait de (Orlhac, 2012)

Comme nous l'avons vu au paragraphe II.3.3, certains auteurs ont donc tenté d'établir le lien entre les mécanismes biologiques tumoraux et les zones d'hétérogénéité observées dans les examens TEP au FDG.

Dans une étude sur un modèle animal de tumeur de la prostate par exemple, (Pugachev et al., 2005) ont étudié, par une approche histologique, la corrélation entre la fixation de FDG et l'hypoxie (i.e. les zones de faible concentration en oxygène), le flux sanguin et la prolifération cellulaire. Leurs résultats ont permis de mettre en évidence que, pour le type de tumeur étudié, une forte concentration en FDG était indicatrice de l'hypoxie tumorale, dont l'expression est un facteur de mauvais pronostic, mais pas des deux autres fonctions étudiées. Une étude antérieure, portant sur les données de 49 patients ayant bénéficié d'une TEP au FDG et une

autre au ^{18}F MISO (un marqueur de l'hypoxie, cf tableau I.5), a montré que la corrélation entre les lieux d'accumulation de ces traceurs semblait tumeur spécifique (Rajendran et al., 2004).

Toujours à partir de modèles animaux, (Henriksson et al., 2007) ont étudié le lien entre la fixation de FDG et les données histologiques de tumeurs issues de xéno greffes de carcinomes épidermoïdes implantés chez 15 souris. La méthode mise en œuvre consistait alors à mesurer l'activité en FDG à l'aide d'un compteur puits de différentes sections de tumeurs (4 sections/tumeurs, 1 tumeur par animal) et à comparer les valeurs obtenues au nombre de cellules tumorales viables mesurées par histopathologie. Les auteurs ont constaté que l'activité en FDG était plus importante dans les sections de tumeurs contenant plus de 50% de cellules tumorales viables. Les résultats montrent cependant un écart très faible entre les valeurs d'activité des deux groupes d'échantillons de tumeurs, et une valeur de significativité limite : $p=0,03$. De plus, les quartiers de tumeurs contenant une plus grande portion de volume nécrosé avaient une activité en FDG plus faible que ceux contenant de plus petits volumes de nécrose. Ces résultats leur ont permis de conclure que l'hétérogénéité intra-tumorale observable par l'histopathologie était corrélée au niveau d'activité du FDG dans les tissus.

Il est donc possible d'imaginer qu'une mesure de l'hétérogénéité de la fixation du radiotracer permette d'extraire des caractéristiques utiles dans le cadre de l'évaluation de la réponse tumorale ou de sa prédiction à plus long terme. Cependant, le niveau de résolution des TEP à l'heure actuelle limite *de facto* la nature des hétérogénéités tissulaires que cette modalité permet d'observer. Ainsi par exemple, si différents clones d'une même tumeur se trouvent mélangés au sein de zones de faibles dimensions (un voxel cubique de 4 mm de côté contenant par exemple $\sim 10^7$ cellules tumorales), alors l'hétérogénéité qui en résulte ne sera pas mesurable.

De nombreuses approches largement documentées en traitement d'image et déjà appliquées en imagerie anatomique sont depuis peu utilisées en TEP pour extraire des informations relatives aux caractéristiques de forme et de texture des volumes métaboliques (cf paragraphe II.3.3). Ainsi, la mesure de l'hétérogénéité, par les méthodes d'analyse de texture, a conduit récemment à de bons résultats pronostiques (Tixier et al., 2011). Cependant, aucune étude n'a jusqu'à présent permis de déterminer la capacité de ces index à prédire la réponse précoce au traitement. Une première étude a donc eu pour objectif d'évaluer la valeur pronostique de ces index sur les TEP prétraitement de patients atteints de cancers colorectaux métastatiques ainsi que leur apport pour la prédiction précoce de la réponse à partir d'une TEP acquise peu de temps après le début du traitement.

En outre, l'évaluation de ces méthodes n'a porté que sur leur reproductibilité, l'influence du choix du paramètre de discrétisation (Tixier et al., 2012) et sur leur variabilité en fonction des paramètres de reconstruction (Galavis et al., 2010). Nous présenterons, dans la partie suivante, des résultats visant à déterminer la variabilité des index de texture en fonction de la façon dont la tumeur est segmentée. Ce problème est d'autant plus important qu'il n'y a pas de consensus sur l'algorithme de segmentation semi-automatique optimum pour définir le

volume tumoral en TEP (Zaidi and El Naqa, 2010). Nous étudierons, dans un second temps, la normalisation la plus adaptée à l'analyse de texture en TEP.

Le travail décrit dans ce chapitre a été essentiellement réalisé dans le cadre du stage de Master 2 de Fanny Orlhac qui s'est déroulé du 01/02/2012 au 31/07/2012 au laboratoire IMNC et que j'ai co-encadré.

V.2. Valeur prédictive et variabilité des index issus de l'analyse de texture : application à des cas de CCM

V.2.1. Matériels et méthodes

Pour cette étude, nous avons considéré une sous-population des données TEP/TDM de la cohorte de patients présentée dans le chapitre IV.

En dehors des calculs d'aires sous les courbes (AUC) ROC, réalisées sous ROCKIT (Version 2011 disponible via <http://metz-roc.uchicago.edu>), tous les calculs et traitements d'images ont été effectués sous *Matlab* (*MathWorks*, Version 7.13).

V.2.1.1. Patients

La population de tumeurs que nous avons considérée a été obtenue en appliquant plusieurs critères d'exclusions. Ces critères portent :

- Sur les volumes des lésions étudiées.
- Sur la préparation des patients avant l'acquisition des données.

Caractéristiques des patients	N
Nombre de patients étudiés	28
Age (années)	
• Moyenne	63
• Intervalle	28-83
Sexe	
• Femme	10
• Homme	18
Ligne de traitement	
• Première	22
• Deuxième	6
Traitement	
• FOLFIRI	13
• FOLFOX	14
• Capecitabine	1
Nombre de lésions par patient	
• Moyenne/patient	3
• Intervalle	1-7
• Total	72
Localisation des lésions	
• Primitive	2
• Foie	59
• Péritoine	4
• Autres	7

Tableau V.1 : caractéristiques des patients atteints d'un cancer colorectal métastatique étudiés.

Ces critères seront explicités aux paragraphes V.2.1.3 et V.2.1.4.3. Au final, les données FDG-TEP de 28 patients ont été étudiées. Les caractéristiques des patients sont fournies dans le tableau V.1.

V.2.1.2. Données TDM

La réponse thérapeutique a été déterminée à partir des données TDM présentées au paragraphe IV.3.1.2. Ainsi, l'application des critères RECIST 1.0 ((Therasse et al., 2000), cf paragraphe II.1.1.2) a permis de déterminer la réponse à 6-8 semaines après le début du traitement. La classification de la réponse de chaque tumeur s'est faite suivant 4 groupes : Réponse Complète (RC), Réponse Partielle (RP), Maladie Stable (MS) et Maladie Progressive (MP).

V.2.1.3. Données TEP/TDM

Afin d'évaluer la capacité des index issus de la TEP à prédire la réponse précoce, un premier critère d'exclusion des données a porté sur l'écart maximal admissible des temps post-injection des deux examens de chaque patient. En accord avec les recommandations PERCIST (Wahl et al., 2009), 11 minutes d'intervalle au maximum ont été tolérées (moyenne : 4 minutes, intervalle [0-11] minutes). Cette valeur correspond au meilleur compromis statistique possible : le choix d'un seuil moins strict (mais inférieur à 15 minutes) ne permettait en effet pas d'inclure plus de lésions répondantes.

Les données utilisées ici ont été présentées au paragraphe IV.3.1.3. Pour les examens TEP de ces 28 patients, l'acquisition des données prétraitement a débuté, en moyenne, 69 minutes (intervalle [58-88] minutes) après injection de 4 MBq/kg. Celle des données TEP/TDM acquises après 14 jours a été effectuée, en moyenne, 68 minutes après injection du radiotracteur (intervalle [52-92] minutes).

V.2.1.4. Analyse quantitative des données TEP

Une fois les images converties en unités SUV en utilisant une normalisation par le poids (cf équation I.17), une première étape de segmentation des volumes tumoraux a été effectuée.

V.2.1.4.1. Segmentation des volumes tumoraux

L'ensemble des index ont été calculés, pour chaque tumeur, à partir de volumes métaboliques, noté VM_S , où S correspond à la définition du seuil utilisé. Ils ont été obtenus à partir d'algorithmes de croissance de région, pour lesquels les seuils utilisés correspondent à :

- $S_{40\%}$, égal à 40% du SUV maximum (SUV_{max}) de la tumeur
- S_{fond} (cf équation IV.1), qui correspond à une valeur de seuil prenant en compte l'activité de fond local (A_{fond}), suivant la définition de (Nestle et al., 2005). Les

valeurs des différents paramètres utilisées pour obtenir le seuil propre à chaque tumeur sont définies par :

- $A_{70\%}$ est le SUV moyen dans le volume obtenu en utilisant $S_{70\%}$.
- A_{fond} est le SUV moyen du fond local. Cette valeur correspond à la moyenne des SUV des voxels contenus dans une coque de 2 voxels d'épaisseur, dont le bord intérieur est situé à 6 voxels de $VM_{70\%}$ et dont on ne considère que les voxels dont l'activité est strictement inférieure à 2,5 unités SUV. Ce dernier point permet de s'assurer qu'aucun voxel inclus dans la région utilisée pour calculer A_{fond} n'appartient à une autre lésion. Cette définition de A_{fond} , qui diffère de celle proposée au IV.3.1.4, a été décrite par (Cheebsumon et al., 2011c) dans son étude sur la comparaison de méthodes de segmentation.
- e_{opt} est une constante optimisée à partir de 3 acquisitions de fantôme Jaszczak. Nous renvoyons le lecteur au paragraphe IV.3.1.4 pour les détails relatifs à l'obtention de la valeur 0,30 utilisée ici.

Quelle que soit la méthode utilisée, une fermeture morphologique a été appliquée à l'ensemble des volumes métaboliques afin d'assurer l'inclusion des zones tumorales nécrotiques dans l'étude de texture. Pour assurer la convergence de la segmentation, un ajustement manuel a été effectué quand les algorithmes de croissance de région divergeaient visiblement ou incluaient d'autres lésions que celles pour lesquelles les critères RECIST 1.0 avaient été établis.

Le seuil $S_{40\%}$ n'a été utilisé que dans le but de comparer les performances de classification des différents index de texture en fonction de la méthode de segmentation utilisée.

V.2.1.4.2. Mesures de l'activité métabolique

A partir des volumes métaboliques précédemment décrits, différents index caractéristiques de l'activité métabolique ont été calculés:

- SUV_{max} : le SUV du voxel d'activité la plus élevée dans la tumeur.
- SUV_{peak} : moyenne des SUV des voxels contenus dans un VOI parallélépipédique de 3 voxels de côté et dont la position est choisie de telle sorte que :
 - Le SUV moyen des voxels contenu dans le VOI soit le plus élevé possible
 - Son centre appartienne à VM_{fond}

On notera que pour les dimensions des voxels des images utilisées ici ($3,91 \times 3,91 \times 4,25 \text{mm}^3$), la sphère de 1,2 cm de diamètre (i.e. ~ 3 voxels) proposée par (Wahl et al., 2009) pour définir le VOI_{peak} dans PERCIST implique l'inclusion, l'exclusion ou l'interpolation des valeurs des SUV des voxels définissant les sommets du parallélépipède circonscrit à cette sphère et aligné sur la maille de l'image. Nous avons ici choisi d'inclure les SUV des voxels des sommets dans le calcul du SUV_{peak} .

- SUV_{fond} correspondant au SUV moyen dans le volume segmenté à partir du seuil S_{fond} .
- VM_{fond} : le volume métabolique de la tumeur segmentée par la méthode de (Nestle et al., 2005)

- La glycolyse totale de la lésion (*Total Lesion Glycolysis*, (Larson et al., 1999)), qui correspond au produit de SUV_{fond} par VM_{fond} .

V.2.1.4.3. Analyse de texture

Pour que le calcul d'un paramètre de texture ait un sens, il doit être effectué sur un volume relativement grand, incluant suffisamment de voxels. Pour ne pas trop réduire la cohorte de tumeurs étudiées, nous avons choisi de nous limiter aux lésions de volumes supérieurs à 4,8 mL (~80 voxels). En effet, d'après (Soret et al., 2007), l'EVP introduit des biais importants pour des structures de longueur caractéristique inférieures à 2~3 LMH. En considérant que les images TEP avec lesquelles nous travaillons ont une résolution spatiale égale à 7 mm (cf paragraphe IV.3.1.4), et en considérant la sphère de diamètre équivalent à 3 fois cette dimension, on en déduit un volume limite de 4,8 mL, en deçà duquel nous estimons que le calcul d'un paramètre de texture ne décrirait pas la texture réelle de la tumeur.

- Ré-échantillonnage
 - Méthode de ré-échantillonnage relatif (R_{relatif})

Afin de comparer nos résultats à ceux de la littérature, nous avons ré-échantillonné les niveaux de fixations de FDG en accord avec la méthode de discrétisation proposée par (Tixier et al., 2011). Cette étape vise à réduire l'impact du bruit sur la mesure de texture en regroupant les valeurs de SUV proches d'un même niveau d'intensité. Ce ré-échantillonnage des valeurs permet de réaffecter un niveau de gris $V(i)$ à chaque voxel i tel que :

$$V(i) = 2^d \times \frac{SUV(i) - SUV_{\min}}{SUV_{\max} - SUV_{\min} + 1} \quad \text{Eq. V.1}$$

Avec:

- d : paramètre de discrétisation entier naturel tel que $d \in [3;7]$. Le choix de sa valeur permet de définir le nombre de niveaux de gris (égale à $n=2^d$) considérés pour l'analyse de texture.
- SUV_{\min} (respectivement SUV_{\max}) la valeur du SUV d'intensité la plus faible (respectivement la plus forte) dans la tumeur
- $SUV(i)$: valeur du SUV dans le voxel i de la tumeur

Par défaut, nous avons utilisé une valeur de d égale à 6, soit une discrétisation des SUV des voxels sur 64 niveaux de gris, en accord avec (Tixier et al., 2011). Les gammes de valeurs prises par chaque niveau de gris étant relatives aux bornes des SUV des voxels propres à chaque tumeur, nous noterons cette méthode de discrétisation R_{relatif} dans la suite de l'exposé.

- Méthode de ré-échantillonnage absolu (R_{absolu})

La normalisation R_{relatif} , appliquée en amont de la mesure de texture dans les volumes métaboliques vus par la TEP, fait disparaître de ces mesures la valeur d'activité réelle, et notamment la valeur d'activité maximum. Autrement dit, deux lésions dont la distribution d'activité est identique mais dont les activités maximales diffèrent conduiraient à des valeurs d'index d'hétérogénéité identiques.

De plus, en utilisant les SUV minimum et maximum propres à chaque tumeur mais en imposant un nombre de niveaux de gris identique, le calcul de R_{relatif} utilise un nombre de valeurs qui n'est pas forcément représentatif des variations d'activités physiologiques intratumorales de toutes les lésions étudiées. En effet, en utilisant un ré-échantillonnage sur 64 niveaux de gris, R_{relatif} donnerait, pour une tumeur de $SUV_{\text{max}} = 20 \text{ g/mL}$ et $SUV_{\text{min}} = 8 \text{ g/mL}$ (i.e. 40% du SUV_{max}), 1 niveau de gris tous les 0,19 g/mL. Par contre, pour une tumeur dont les SUV des voxels qui la constituent seraient compris dans l'intervalle [1,6; 4], R_{relatif} donnerait 1 niveau de gris tous les 0,04 g/mL.

Pour intégrer l'information d'intensité dans la mesure de texture, il est possible d'utiliser un intervalle de SUV commun à toutes les tumeurs (R_{absolu}). Par cette méthode, un même niveau de gris est attribué à tous les SUV des voxels de toutes les tumeurs compris dans un intervalle de valeurs défini à l'avance. Pour englober tous les cas de figures rencontrés dans notre cohorte, les bornes maximales et minimales utilisées pour décrire l'échelle de niveaux de gris commune sont prises en remplaçant, dans l'équation V.1, SUV_{min} (respectivement SUV_{max}) par le plus petit (respectivement le plus grand) SUV minimal (respectivement maximal) de toutes les tumeurs de la cohorte.

- Calculs des matrices de texture

Une fois le ré-échantillonnage effectué, quatre matrices de texture ont été calculées, pour chaque examen et chaque tumeur. Quand nécessaire, ces algorithmes, issus des méthodes décrites au paragraphe 3.3.1.4, et donc définies pour des applications 2D, ont été adaptés à l'analyse de données 3D en moyennant les paramètres obtenus pour chacune des 13 directions qui permettent de passer d'un voxel à ses plus proches voisins. Les algorithmes mis en œuvre sont illustrés en figure V.2, pour la coupe coronale de la tumeur présentée à la figure V.1. Pour simplifier l'explication, seules 4 valeurs de ré-échantillonnage ont été utilisées dans cet exemple.

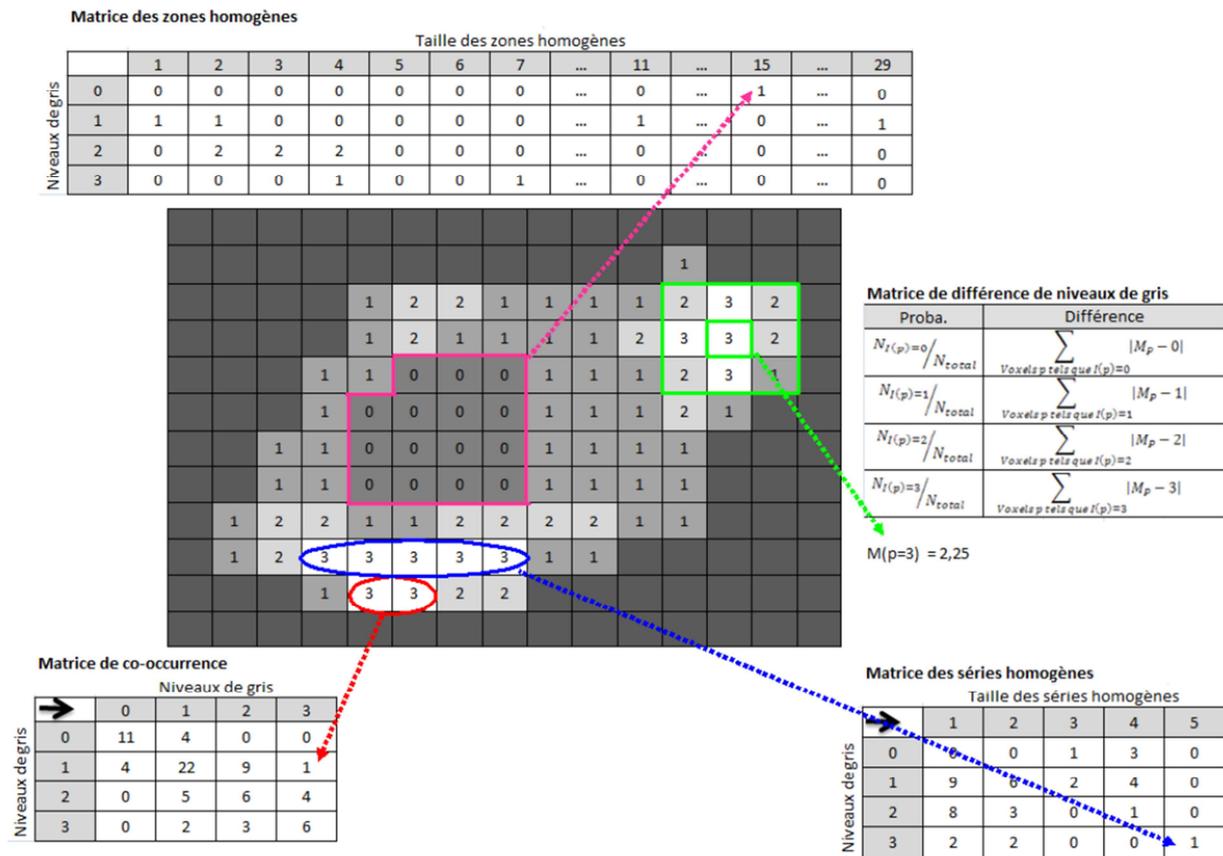


Figure V.2 : Principe, pour le cas d'une analyse à 2 dimensions, de la construction des 4 matrices utilisées pour l'analyse de la texture de la fixation de FDG vue par la TEP. Le ré-échantillonnage utilisé ici à des fins d'illustration ne se base que sur 4 valeurs de niveau de gris ($d=2$). Les pixels les plus sombres n'appartiennent pas au volume segmenté et ne sont donc pas pris en compte pour l'analyse. Les flèches situées dans les cases supérieures gauches des matrices de séries homogènes et de cooccurrence correspondent au sens de parcours choisi a priori pour définir, respectivement, les tailles des séries homogènes et paires de niveaux de gris rencontrés dans l'image ré-échantillonnée. Extrait de (Orlhac, 2012).

Les matrices calculées à partir des données contenues dans le volume segmenté correspondent à :

- Matrice de cooccurrence (*Cooccurrence matrix*, C) : elle caractérise les relations d'intensité entre les couples de voxels voisins. Une fois normalisée par l'ensemble des occurrences possibles, chaque élément de C, noté $C(i,j,k)$, correspond à la probabilité de trouver une paire de voxels d'intensités i et j dans l'une des k directions de l'espace (avec $k \in [1; 13]$).
- Matrice de longueur des séries homogènes (*gray level Run length matrix*, R) : elle renseigne sur la longueur des séries (ou *run*) de même niveaux de gris dans l'image. Chaque élément $R(i,j,k)$ de R donne le nombre de séries homogènes de longueur j et d'intensité i dans la direction k .
- Matrice de différence des niveaux de gris (*Neighborhood gray level difference matrix*, N avec $N \in \mathcal{M}_{n,2}(\mathbb{N}^+) / n=2^d$, où $\mathcal{M}_{n,2}(\mathbb{N}^+)$ correspond à l'espace des matrices de

dimension n lignes et 2 colonnes à valeurs dans \mathbb{N}^+) : elle décrit les différences entre chaque voxel et ses voisins. $\forall i \in [1; n]$:

- $N(i,1)$ est égal à la probabilité d'occurrence d'un voxel d'intensité i
- $N(i,2) = \sum_p |i - \bar{A}_p|$ avec $p \in \{\text{voxels d'intensité } i\}$ et \bar{A}_p valeur moyenne des niveaux de gris des 26 voxels plus proches voisins du voxel p.
- Matrice de longueur de zones homogènes (*gray level Zone length matrix*, Z) : elle donne la longueur des zones de niveaux de gris identiques. $\forall i \in [1; n]$ et $\forall j \in [1; m]$ avec m longueur de la plus grande zone, $Z(i,j)$ correspond au nombre de zones de longueur j et d'intensité i dans la tumeur. Ces matrices, définies initialement pour décrire les textures d'images 2D, ont été développées pour permettre le dénombrement des tailles de toutes les régions de pixels de mêmes intensités. Elles se construisent sur le même principe que la matrice de différences de longueurs des séries homogènes. Elles ne visent pas à dénombrer des segments d'une longueur donnée dans une direction choisie mais directement les tailles des régions de mêmes niveaux d'intensité.

En fonction de la manière dont ces matrices sont analysées, il est possible d'extraire différents paramètres caractéristiques de la distribution des niveaux de gris dans l'image. Ces paramètres sont résumés dans les tableaux V.2, V.3 et V.4 pour les matrices de cooccurrence (C), de longueur des séries homogènes (R) et de différence des niveaux de gris (N), respectivement. Les caractéristiques extraites des matrices de longueur des zones homogènes (Z) étant identiques à celles issues des longueurs des séries homogènes, nous n'avons pas reproduit l'ensemble des formules. En effet ici, seule la notion de zone (*Zone*) remplace celle de longueur (*Run*). Au total, ce sont 31 paramètres dérivés des matrices de texture qui ont été extraits de chacun des volumes segmentés, puis utilisés pour caractériser l'hétérogénéité tumorale vue par la TEP. Une définition et une interprétation sommaire de leur sens physique sont fournies dans le tableau V.5.

Nom de la caractéristique	Définition mathématique
<i>Homogeneity</i>	$\sum_{i,j} \frac{C(i,j)}{1 + i - j }$ Eq. V.2
<i>Energy</i>	$\sum_{i,j} C(i,j)^2$ Eq. V.3
<i>Correlation</i>	$\sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)C(i,j)}{\sigma_i \sigma_j}$ Eq. V.4
<i>Contrast (Inertia)</i>	$\sum_{i,j} i - j ^2 C(i,j)$ Eq. V.5
<i>Entropy</i>	$-\sum_{i,j} C(i,j) \times \ln(C(i,j))$ Eq. V.6
<i>Dissimilarity</i>	$\sum_{i,j} C(i,j) \times i - j $ Eq. V.7

Tableau V.2 : caractéristiques extraites des matrices de cooccurrences C. μ_i et σ_i correspondent respectivement à la moyenne et à l'écart-type des $C(i,j)$ pour tous les niveaux de gris j, i étant fixé.

Nom de la caractéristique	Définition mathématique
<i>Short Run Emphasis (SRE)</i>	$\frac{1}{\theta} \times \sum_{i,j} \frac{R(i,j)}{j^2}$ Eq. V.8
<i>Long Run Emphasis (LRE)</i>	$\frac{1}{\theta} \times \sum_{i,j} R(i,j) \times j^2$ Eq. V.9
<i>Low Gray-level Run Emphasis (LGRE)</i>	$\frac{1}{\theta} \times \sum_{i,j} \frac{R(i,j)}{i^2}$ Eq. V.10
<i>High Gray-level Run Emphasis (HGRE)</i>	$\frac{1}{\theta} \times \sum_{i,j} R(i,j) \times i^2$ Eq. V.11
<i>Short Run Low Gray-level Emphasis (SRLGE)</i>	$\frac{1}{\theta} \times \sum_{i,j} \frac{R(i,j)}{i^2 \times j^2}$ Eq. V.12
<i>Short Run High Gray-level Emphasis (SRHGE)</i>	$\frac{1}{\theta} \times \sum_{i,j} \frac{R(i,j) \times i^2}{j^2}$ Eq. V.13
<i>Long Run Low Gray-level Emphasis (LRLGE)</i>	$\frac{1}{\theta} \times \sum_{i,j} \frac{R(i,j) \times j^2}{i^2}$ Eq. V.14
<i>Long Run High Gray-level Emphasis (LRHGE)</i>	$\frac{1}{\theta} \times \sum_{i,j} R(i,j) \times i^2 \times j^2$ Eq. V.15
<i>Gray-Level Non-Uniformity run (GLNUR)</i>	$\frac{1}{\theta} \times \sum_i \left(\sum_j R(i,j) \right)^2$ Eq. V.16
<i>Run Length Non-Uniformity (RLNU)</i>	$\frac{1}{\theta} \times \sum_j \left(\sum_i R(i,j) \right)^2$ Eq. V.17
<i>Run Percentage (RP)</i>	$\frac{\theta}{n \times m}$ Eq. V.18

Tableau V.3 : caractéristiques issues des matrices de longueurs des séries homogènes. θ correspond au nombre de séries homogènes dans la tumeur. n est le nombre de niveaux de gris utilisés pour la discrétisation et m la longueur de la plus grande série. Une fois calculées pour chacune des 13 directions de parcours, les valeurs obtenues pour chacun des index sont moyennées.

Nom de la caractéristique	Définition mathématique
<i>Coarseness</i>	$\sum_{i,j} \frac{1}{N(i,1) \times N(i,2)}$ Eq. V.19
<i>Contrast</i>	$\left(\frac{1}{G \times (G-1)} \times \sum_{i,j} N(i,1) \times N(j,1) \times (i-j)^2 \right) \times \left(\frac{1}{E} \times \sum_i N(j,2) \right)$ Eq. V.20
<i>Busyness</i>	$\frac{\sum_i N(i,1) \times N(i,2)}{\sum_{i,j} i \times N(i,1) - j \times N(j,1)}$ Eq. V.21

Tableau V.4 : caractéristiques issues des matrices de différences de niveaux de gris (N). G correspond au nombre de différents niveaux de gris et E au nombre de voxels contenus dans le volume segmenté.

Nom de la matrice de texture	Définition mathématique	Interprétation de l'index
Matrice de co-occurrence (C)	<i>Homogeneity</i> <i>Energy (ou moment angulaire d'ordre 2)</i> <i>Correlation</i> <i>Contrast / Inertia</i> <i>Entropy</i> <i>Dissimilarity</i>	Homogénéité locale des paires de voxels Uniformité de la texture Dépendance linéaire des niveaux de gris dans l'image Mesure des variations locales dans les paires de voxels Complexité et désordre de la texture Variation des niveaux de gris dans les paires de voxels
Matrice de longueur des séries homogènes (R)	<i>Short Run Emphasis (SRE)</i> <i>Long Run Emphasis (LRE)</i> <i>Low Gray-level Run Emphasis (LGRE)</i> <i>High Gray-level Run Emphasis (HGRE)</i> <i>Short Run Low Gray-level Emphasis (SRLGE)</i> <i>Short Run High Gray-level Emphasis (SRHGE)</i> <i>Long Run Low Gray-level Emphasis (LRLGE)</i> <i>Long Run High Gray-level Emphasis (LRHGE)</i> <i>Gray-Level Non-Uniformity run (GLNUr)</i> <i>Run Length Non-Uniformity (RLNU)</i> <i>Run Percentage (RP)</i>	Distribution des segments courts Distribution des segments longs Distribution des segments de faible intensité Distribution des segments de haute intensité Distribution des segments courts avec une intensité faible (combinaison de <i>SRE</i> et <i>LRE</i>) Distribution des segments courts et d'intensité élevée (combinaison des <i>SRE</i> et <i>HGRE</i>) Distribution des segments longs et de faibles intensité (combinaison <i>LRE</i> et <i>LGRE</i>) Distribution des segments longs et d'intensité élevée (combinaison <i>LRE</i> et <i>HGRE</i>) Non uniformité des niveaux de gris Non uniformité des longueurs de segments Homogénéité et distribution des séries dans l'image
Matrice de différence des niveaux de gris (N)	<i>Coarseness</i> <i>Contrast</i> <i>Busyness</i>	Granulosité de l'image - Fréquence spatiale des changements d'intensité
Matrice de longueur des zones homogènes (Z)	<i>Short Zone Emphasis (SZE)</i> <i>Long Zone Emphasis (LZE)</i> <i>Low Gray-level Zone Emphasis (LGZE)</i> <i>High Gray-level Zone Emphasis (HGZE)</i> <i>Short Zone Low Gray-level Emphasis (SZLGE)</i> <i>Short Zone High Gray-level Emphasis (SZHGE)</i> <i>Long Zone Low Gray-level Emphasis (LZLGE)</i> <i>Long Zone High Gray-level Emphasis (LZHGE)</i> <i>Gray-Level Non-Uniformity zone (GLNUz)</i> <i>Zone Length Non-Uniformity (ZLNU)</i> <i>Zone Percentage (ZP)</i>	Distribution des petites zones Distribution des grandes zones Distribution des zones de faible intensité Distribution des zones de haute intensité Distribution des petites zones avec une intensité faible (combinaison de <i>SZE</i> et <i>LZE</i>) Distribution des petites zones et d'intensité élevée (combinaison des <i>SZE</i> et <i>HGZE</i>) Distribution des grandes zones et de faibles intensité (combinaison <i>LZE</i> et <i>LGZE</i>) Distribution des grandes zones et d'intensité élevée (combinaison <i>LZE</i> et <i>HGZE</i>) Non uniformité des niveaux de gris Non uniformité des tailles des zones Homogénéité et distribution des zones dans l'image

Tableau V.5 : liste des index extraits des différentes matrices de textures étudiées. Nous renvoyons le lecteur au manuscrit de thèse de doctorat de (Thibault, 2009) pour de plus amples informations.

Vue le faible nombre de publications ayant mis en œuvre l'analyse de texture en imagerie TEP, nous avons résumé, dans les tableaux V.6A et V.6B, les correspondances entre les noms utilisés par (Galavis et al., 2010), (Tixier et al., 2011) et (Tixier et al., 2012) et ceux utilisés pour définir les différents index présentés.

Les matrices de longueurs de zones homogènes, issues des travaux de (Thibault et al., 2011), ont été mises en œuvre pour la première fois par (Tixier et al., 2012; Tixier et al., 2011), d'où le fait que (Galavis et al., 2010) ne les aient pas présentées (tableau V.6B, colonne 3). De même, constatant les fortes corrélations entre les valeurs obtenues par les matrices R et Z, (Tixier et al., 2012) n'ont pas jugé nécessaire de rapporter les résultats obtenus par la matrice de longueurs des séries homogènes dans leurs second article (cf tableau V.6A, colonne 5).

Matrice correspondante	Noms des index de texture employés dans l'étude des CCM	Correspondance avec les notations employées dans les autres articles			Remarques
		Galavis et al. 2010	Tixier et al. 2011	Tixier et al. 2012	
Matrice de Cooccurrence	Homogeneity	Homogeneity (9)	Homogeneity	Homogeneity	
	Energy	Energy (7)	Second angular moment	Second angular moment	La définition de Galavis n'est pas précisée
	Correlation	Correlation (11)	Correlation	Correlation	
	Contrast (Inertia)	Contrast – GLMC (10)	Contrast (Inertia)	Correlation (Inertia)	
	Entropy	Entropy – GLMC (17)	Entropy	Entropy	
	Dissimilarity	Dissimilarity (15)	Dissimilarity	Dissimilarity	
	SRE	SRE (32)	SRE	-	
	LRE	LRE (33)	LRE	-	
	LGRE	LGRE (37)	Low-intensity run emphasis	-	
	HGRE	HGRE (38)	High-intensity run emphasis	-	
Matrice de longueur de séries homogène	SRLGE	SRLGE (39)	Low-intensity short-run emphasis	-	
	SRHGE	SRHGE (40)	High-intensity short-run emphasis	-	
	LRLGE	LRLGE (41)	Low-intensity long-run emphasis	-	
	LRHGE	LRHGE (42)	High-intensity long-run emphasis	-	
	GLNUr	Gray-level nonuniformity (34)	Intensity variability	-	
	RLNU	Run length nonuniformity (35)	Run-length variability	-	
	RP	Run Percent (36)	Run percentage	-	

Tableau V.6A : correspondances entre les définitions des index issus des matrices C et L présentées et celles utilisées dans les 3 principales publications ((Galavis et al., 2010), (Tixier et al., 2011) et (Tixier et al., 2012)) traitant de l'intérêt clinique de l'analyse de texture en TEP.

Matrice correspondante	Noms des index de texture employés dans l'étude des CCM		Correspondance avec les notations employées dans les autres articles		Remarques
	Galavis et al. 2010	Tixier et al. 2011	Tixier et al. 2012		
Matrice de différence de niveaux de gris	Coarseness	Coarseness (48)	Coarseness	-	-
	Contrast	Contrast-NGL (49)	Contrast	-	-
	Busyness	Busyness (50)	Busyness	-	-
	SZE	-	Small-zone emphasis	Small-area emphasis	-
	LZE	-	Large-zone emphasis	Large-area emphasis	-
	LGZE	-	Low-intensity zone emphasis	Low-intensity emphasis	Formule non donnée par Tixier 2012.
	HGZE	-	High-intensity zone emphasis	High-intensity emphasis	-
	SZLGE	-	Low-intensity short-zone emphasis	Low-intensity small-area emphasis	-
	SZHGE	-	High-intensity short-zone emphasis	High-intensity small-area emphasis	La définition considérée pour l'étude CCM et celle de Tixier et al. 2012 diffèrent. Celles que nous avons utilisées, en accord avec celle proposée par Thibault 2009, correspondent à celle que Tixier et al. 2012 ont attribuée à <i>High-Intensity Large-Area Emphasis</i> .
	LZLGE	-	Low-intensity large-zone emphasis	Low-intensity large-area emphasis	La formule utilisée dans l'étude CCM diffère de celle proposée dans Tixier 2012. Celle que nous avons utilisée pour définir cet index correspond à celle utilisée par Tixier et al. 2012 pour définir <i>High-Intensity Small Area Emphasis</i> .
Matrice de longueur de zones homogènes	LZHGE	-	High-intensity large-zone emphasis	High-intensity large-area emphasis	La définition utilisée dans l'étude CCM est identique à la définition proposée par Tixier et al. 2011. Elle diffère cependant entre l'étude CCM et celle donnée dans Tixier 2012, où elle fait apparaître un terme i^2 en dénominateur de $R(i,j)$, non présent chez Thibault 2009. Ce même terme i^2 est présent dans la formule de <i>Size Zone variability</i> , que l'on retrouve bien chez Tixier et al. 2011 mais plus dans les travaux de Tixier et al. 2012.
	GLNUz	-	Intensity variability	Intensity variability	La définition considérée est identique à l'étude de Tixier et al. 2011. Elle diffère de celle prise par Tixier et al. 2012, où elle est alors différente de la définition donnée par Thibault 2009.
	ZLNU	-	Size-zone variability	Size-zone variability	La définition diffère entre l'étude CCM et celle donnée par Tixier et al. 2012. Elle est alors exactement égale à l'inverse de la définition de <i>Large Area Emphasis</i> donnée par ces mêmes auteurs.
	ZP	-	Zone percentage	Zone percentage	

Tableau V.6B : correspondances entre les définitions des index issus des matrices N et Z présentées et celles utilisées dans les 3 principales publications ((Galavis et al., 2010), (Tixier et al., 2011) et (Tixier et al., 2012)) traitant de l'intérêt clinique de l'analyse de texture en TEP.

L'une des principales difficultés dans le travail de comparaison que nous avons mené vient des écarts entre les définitions des index de textures extraits de Z (cf colonne 6 tableau V.6B):

- Certaines définitions d'index diffèrent entre (Tixier et al., 2011) et (Tixier et al., 2012), comme c'est le cas que pour *Intensity Variability (GLNUz)* et *Size-Zone Variability (ZLNU)*. Ces écarts sont cependant difficiles à recenser dans la mesure où très peu de formules sont fournies dans (Tixier et al., 2011).
- Certains index définis dans (Tixier et al., 2012) (cf tableau V.7) et extraits de Z ne suivent pas à la même définition que les index de même nom définis dans (Thibault, 2009). Ces dernières formules sont elles-mêmes dérivées des formules issues des travaux de (Tang, 1998; Xu et al., 2004), relatives aux matrices de longueur des séries homogènes (cf tableau V.3). Pour pouvoir comparer nos résultats aux travaux initiaux de (Tixier et al., 2011), qui portent sur l'intérêt clinique des index de texture, et pour lequel il est légitime de supposer que les formules sont en accord avec la littérature (cf écarts entre (Tixier et al., 2011) et (Tixier et al., 2012), tableau V6.B), nous avons donc calculé les index selon les formules usuelles, proposées par (Thibault, 2009). Les formules proposées dans (Tixier et al., 2012) et rapportées à l'identique dans le tableau V.7, laissent en effet penser que certaines erreurs se sont probablement glissées dans le formulaire fourni, notamment du fait:
 - De notations utilisant des indices non référencés pour les expressions mathématiques de *Small-Area Emphasis* et *Large-Area Emphasis*.
 - Du remplacement systématique du signe multiplié par une virgule dans les termes de sommation.

Le fait de considérer les formules usuelles permet en outre de conserver les interprétations des index fournis dans le tableau V.5 (colonne 3).

Formulas for Regional Heterogeneity Features			
Feature	Formula	Feature	Formula
Small-area emphasis	$\frac{1}{\Omega} \sum_{i=1}^M \sum_{j=1}^N \frac{z(i,j)}{j^2}$	High-intensity emphasis	$\frac{1}{\Omega} \sum_{i=1}^M \sum_{j=1}^N i^2, z(i,j)$
Large-area emphasis	$\frac{1}{\Omega} \sum_{i=1}^M \sum_{j=1}^N j^2, z(i,j)$	Low-intensity small-area emphasis	$\frac{1}{\Omega} \sum_{i=1}^M \sum_{j=1}^N \frac{z(i,j)}{i^2, j^2}$
Intensity variability	$\frac{1}{\Omega} \sum_{i=1}^M \left[\sum_{j=1}^N \frac{z(i,j)}{j^2} \right]^2$	High-intensity small-area emphasis	$\frac{1}{\Omega} \sum_{i=1}^M \sum_{j=1}^N i^2, j^2, z(i,j)$
Size-zone variability	$\frac{1}{\Omega} \sum_{i=1}^M \left[\sum_{j=1}^N \frac{z(i,j)}{j^2} \right]^2$	Low-intensity large-area emphasis	$\frac{1}{\Omega} \sum_{i=1}^M \sum_{j=1}^N \frac{j^2, z(i,j)}{i^2}$
Zone percentage	$\Omega / \sum_{i=1}^M \sum_{j=1}^N j^2, z(i,j)$	High-intensity large-area emphasis	$\frac{1}{\Omega} \sum_{i=1}^M \sum_{j=1}^N \frac{i^2, z(i,j)}{j^2}$

Ω = number of homogeneous areas within tumor; z = intensity size-zone matrix; M = discretization value; N = size of largest homogeneous area within tumor; z(i,j) = number of areas with intensity i and size j.

Tableau V.7 : Reproduction du formulaire présenté par (Tixier et al., 2012) pour définir les index de texture déduits de la matrice de longueur des zones homogènes.

V.2.1.4.4. Analyse de forme

Afin d'évaluer le potentiel prédictif de la forme des lésions, nous avons enfin étudié la sphéricité de la tumeur, définie par :

$$Sphéricité = \pi^{1/3} \times \frac{(6 \times Volume)^{2/3}}{Aire} \quad \text{Eq. V.22}$$

où *Volume* et *Aire* sont obtenus, pour chaque tumeur et chaque examen, par une triangulation de Delaunay (Wadell, 1932). Cette méthode correspond à une méthode de maillage par des triangles qui vise à maximiser le plus petit angle de l'ensemble des triangles utilisés. La valeur obtenue varie entre 0 (cas d'un plan infiniment fin) et 1 (pour une sphère parfaite).

V.2.1.5. Analyse statistique

- Valeurs prédictives des index

Dans un premier temps, la capacité des 37 index décrits précédemment à prédire la réponse à la thérapie à partir de l'examen prétraitement (PT) seul a été évaluée.

Par la suite, nous avons cherché à déterminer la capacité prédictive de la variation de ces index (notée Δ_{index}) entre la TEP prétraitement (TEP(PT)) et la TEP précoce (i.e. acquise après 14 jours de traitement, que l'on notera dans la suite TEP(J14)). La formule employée a été présentée à l'équation IV.9.

Les valeurs des index utilisées pour ces deux premières analyses ont été obtenues à partir de la méthode de segmentation définie par le seuil $S_{\text{fond}} : VM_{\text{fond}}$.

Pour ces deux études, une analyse ROC non paramétrique (cf paragraphe IV.2.1.4.1) a été utilisée pour déterminer la capacité des index à séparer les 2 groupes de lésions identifiés par RECIST 1.0 après 6-8 semaines de traitement :

- Lésions répondant à la thérapie (RC et RP) d'une part
- Lésions ne répondant pas à la thérapie (MS et MP) d'autre part.

Les aires sous la courbe ROC et leurs écart-types ont été systématiquement rapportés et comparés à la valeur de référence 0,5 en utilisant un test non paramétrique de Wilcoxon ($\alpha=0,05$).

- Variabilité des index métaboliques
 - En fonction du paramètre et de la méthode de discrétisation

Nous avons évalué l'effet du choix du paramètre de discrétisation sur les capacités prédictives des index de texture. A partir de mesures effectuées sur la TEP prétraitement seule, nous

avons rapporté les valeurs des AUC des courbes ROC de chaque index de texture pour 5 niveaux de discrétisation.

- En fonction du volume métabolique considéré

Afin d'évaluer l'impact du choix de la définition du volume métabolique sur les valeurs prises par les index de textures, nous avons tout d'abord étudié l'accord entre les valeurs de VM_{fond} et $VM_{40\%}$ établies sur la TEP(PT) à l'aide d'un graphique de Bland-Altman (cf IV.3.1.5).

Nous avons ensuite étudié l'impact de la méthode de segmentation sur les valeurs prises par les différents index. Seules les valeurs des index calculés sur les examens TEP prétraitement ont été étudiées ici. Pour chaque index étudié, l'écart relatif moyen (et son écart-type) des valeurs obtenues pour toutes les tumeurs de la cohorte en utilisant $S_{40\%}$ et S_{fond} a été rapporté. L'écart relatif moyen, noté $\bar{\Delta}$, correspond à la moyenne, sur toutes les tumeurs, des écart-relatifs Δ , est défini par :

$$\Delta(\%) = \frac{index(S_{40\%}) - index(S_{fond})}{\left(\frac{index(S_{40\%}) + index(S_{fond})}{2}\right)} \times 100 \quad \text{Eq. V.23}$$

Enfin, pour évaluer l'effet de l'algorithme de segmentation sur la capacité des différents index à prédire la réponse RECIST tardive à partir de la TEP prétraitement, nous avons calculé l'AUC des courbes ROC de chacun d'entre eux et comparé chacune des valeurs obtenues à la valeur de référence 0,5 par un test de Wilcoxon. Pour déterminer si les AUC des différents index calculés étaient, dans l'ensemble, plus prédictifs en utilisant l'une ou l'autre des méthodes de segmentation, nous avons effectué un test non paramétrique apparié bilatéral de Wilcoxon (ou *Wilcoxon signed rank test*) en considérant toutes les paires d'AUC obtenues pour chaque index à partir de VM_{fond} et $VM_{40\%}$ ($\alpha=0,05$).

- Corrélations entre les index métaboliques étudiés

Afin de mettre en évidence les liens entre les différents index extraits des données TEP, nous avons calculé les coefficients de corrélation de Pearson obtenus entre les valeurs de tous les index, pour les valeurs calculées à partir des TEP prétraitements. Nous avons cherché à mettre en évidence les liens entre les index :

- calculés sur une même matrice de texture
- issus des matrices de séries et de zones homogènes
- de texture et classiques ou de forme : SUV, VM_{fond} , TLG ou sphéricité.

- Evaluation de la méthode de discrétisation absolue

Enfin, nous avons comparé les capacités prédictives des index (issus de la TEP(PT) seule ou en étudiant leurs évolutions relatives entre la TEP(PT) et la TEP(J14)) en fonction de la méthode de ré-échantillonnage, R_{absolu} ou $R_{relatif}$, en utilisant 64 valeurs de discrétisation, à partir d'une analyse ROC. La supériorité des différentes AUC par rapport à 0,5 a été

systématiquement testée à l'aide d'un test de Wilcoxon. De même, pour les deux études, un test de Wilcoxon apparié a été effectué pour évaluer la différence entre les médianes des deux groupes d'AUC formés de tous les index de texture par l'une et l'autre méthode de ré-échantillonnage. Pour évaluer plus finement si la différence obtenue pour chaque index entre chaque paire d'AUC par l'une et l'autre des 2 méthodes de ré-échantillonnage était significative, nous avons mis en œuvre la méthode de (Hanley and McNeil, 1983). Cette analyse statistique permet de prendre en compte les corrélations entre les valeurs des index calculés par l'une et l'autre des méthodes de discrétisation lorsqu'elles ont été obtenues à partir d'une même population de tumeurs. Enfin, les corrélations entre les valeurs des différents index de texture obtenues en utilisant R_{absolu} d'une part et les index classiques d'autre part, ont été systématiquement rapportées.

V.2.2.Résultats

V.2.2.1. Lésions étudiées

- En considérant seulement la TEP prétraitement

La cohorte initiale, présentée au paragraphe IV.3.1.1, regroupait 40 patients CCM (Hendlisz et al., 2011). Une fois les critères d'exclusion appliqués ($VM_{\text{fond}} > 4,8$ mL), les données TEP(PT) de 28 patients, soit 72 lésions (cf tableau V.1), ont été considérées pour obtenir les résultats relatifs à la capacité de la TEP prétraitement à prédire la réponse anatomique vue par RECIST 6-8 semaines après le début de la chimiothérapie.

Le volume métabolique moyen (respectivement médian) des 72 lésions avant traitement, obtenu en utilisant S_{fond} , était de 73,9 mL (respectivement 20,8 mL) (intervalle [5,0-519,0] mL). Les critères RECIST 1.0 établis après 6-8 semaines de traitement ont donné :

- Aucune lésion RC
- 21 RP
- 38 MS
- 13 MP

- En considérant la TEP prétraitement et la TEP précoce (J14)

En appliquant des critères d'exclusion portant sur l'écart entre les délais post-injection des TEP(PT) et TEP(J14) (≤ 11 minutes), seules les données de 18 des 27 patients, soit un total de 45 tumeurs, ont été étudiées. Pour ces lésions, les valeurs caractéristiques des VM_{fond} sur la TEP(J14) étaient de 71,1 mL en moyenne (médiane : 48,9 mL) (intervalle : [3,2-238,6] mL).

Parmi elles, RECIST 1.0 a distingué :

- Aucune lésion RC
- 9 RP
- 27 MS
- 9 MP

V.2.2.2. Pouvoir prédictif des index métaboliques à partir de la TEP prétraitement seule

En ne se basant que sur la TEP(PT), nous avons considéré un ré-échantillonnage sur 64 valeurs de niveaux de gris et seulement les volumes métaboliques segmentés en utilisant S_{fond} , nous avons construit les courbes ROC des 37 index métaboliques. Celles de Homogeneity et de SUV_{peak} sont représentées en figure V.3.

Les AUC (± 1 écart-types) des courbes de tous les index sont représentées en figure V.4. Aucun d'entre eux n'a obtenu une AUC supérieure à celle de SUV_{peak} (0,66). Sans considérer le problème de comparaison multiple, l'AUC de ce dernier index est d'ailleurs la seule à être significativement supérieure à 0,5 ($p=0,032$). On constate en particulier qu'aucun des index de texture étudié n'obtient une AUC supérieure à 0,63.

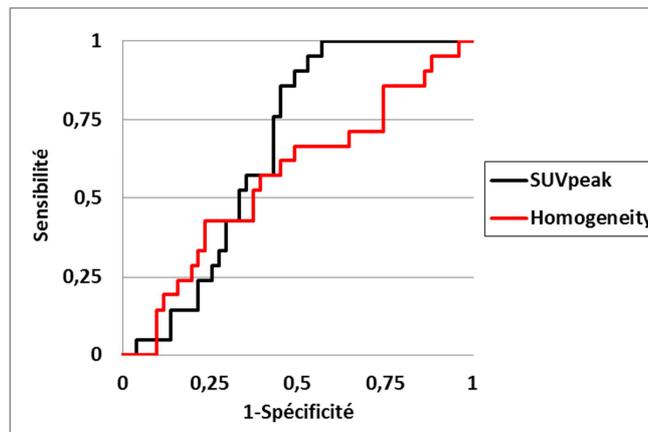


Figure V.3 : courbes ROC de SUV_{peak} et Homogeneity pour la distinction entre tumeurs CR et PR d'une part et MS et MP d'autre part, à partir d'un examen TEP prétraitement seul. Extrait et modifié de (Orlhac, 2012).

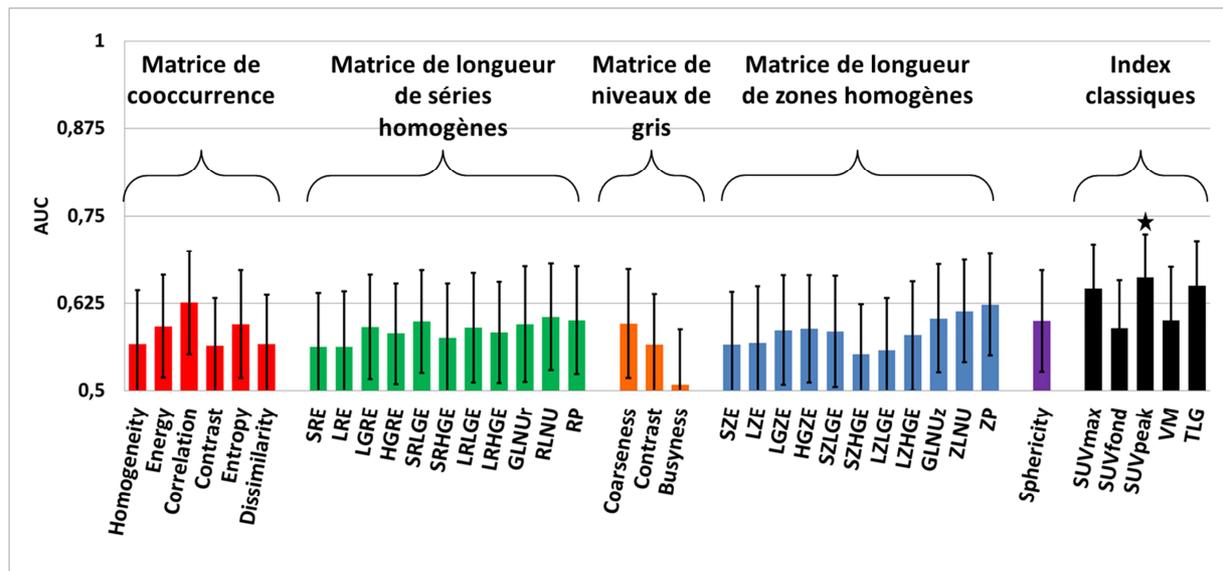


Figure V.4 : Aires sous les courbes ROC (AUC) et écarts-types associés des 37 index calculés à partir de l'examen TEP(PT). Extrait et modifié de (Orlhac, 2012). * : valeur significativement supérieure à 0,5 ($p<0,05$).

V.2.2.3. Capacités prédictives des index métaboliques avec prise en compte de la TEP précoce

Nous avons ensuite comparé la capacité des différents index à prédire précocement la réponse anatomique tardive vue par RECIST, à partir de l'évolution de ces index entre TEP(PT) et TEP(J14). Les AUC issues de l'analyse ROC, basées sur les données de 45 lésions colorectales, sont présentées en figure V.5.

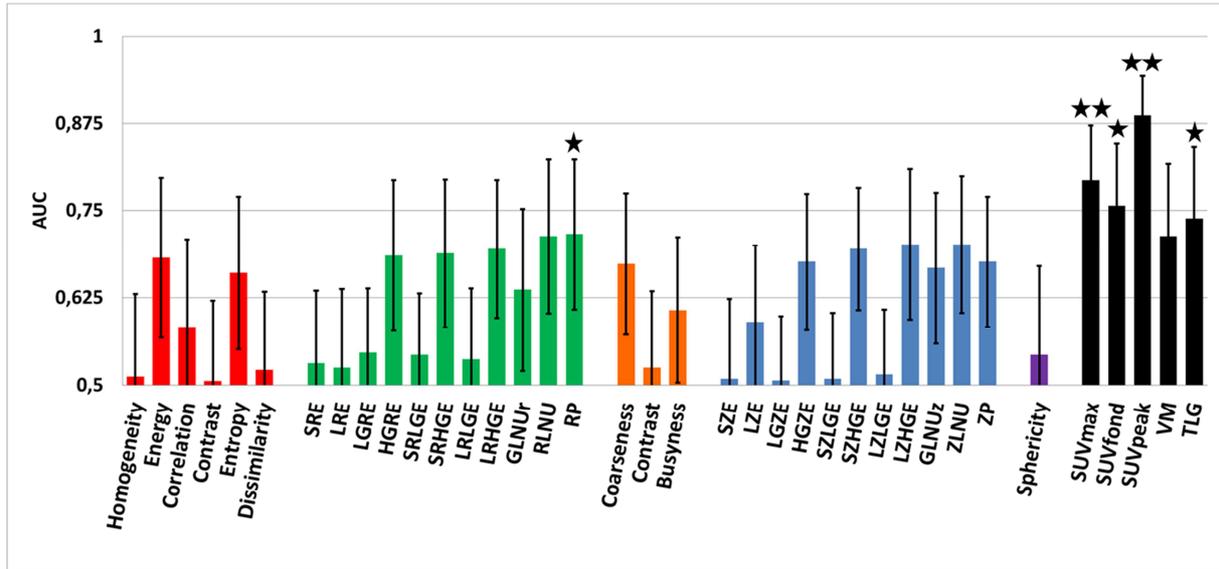


Figure V.5 : AUC (et écart-types) des 37 index mesurant l'évolution relative entre les deux examens TEP. AUC significativement supérieure à 0,5 (* : $p < 0,05$; ** : $p < 0,01$).

En dehors de VM_{fond} , tous les index classiques ont une AUC significativement supérieure à 0,5. Parmi les 31 index de textures considérés, seul *Run Percentage* obtient une p-value inférieure au seuil de significativité ($p=0,047$) avec une AUC de 0,716.

V.2.2.4. Variabilité des paramètres de texture

V.2.2.4.1. En fonction du paramètre de discrétisation

L'objectif suivant était d'évaluer l'impact du choix du nombre de niveaux de gris utilisés pour le ré-échantillonnage sur la capacité prédictive des index de texture calculés sur la TEP(PT) seule. Nous avons pour cela systématiquement calculé, pour 8, 16, 32, 64 et 128 niveaux de discrétisation, les AUC des courbes ROC obtenues par chaque index. Pour 74% de ces index (23/31), les performances dépendent peu du nombre de niveaux de gris considéré d'après les faibles différences observées entre les différentes valeurs obtenues (cf figure V.6 : *Homogeneity*, *Correlation*, *GLNUz*, *ZLNU*). A contrario, *Energy*, *Entropy*, *LRLGE* et *SZLGE* montrent des performances quasi-nulles quand seulement 8 valeurs sont utilisées pour ré-échantillonner les valeurs de fixations intra tumorales ; leurs capacités à distinguer les lésions répondant à la thérapie des lésions ne répondant pas sont relativement stables pour les autres niveaux de discrétisation. Enfin, on observe des variations importantes des AUC, bien que non significatives, pour des degrés de ré-échantillonnage intermédiaires (16, 32 ou 64), pour *Busyness*, *SZE*, *SZHGE* et *ZP*.

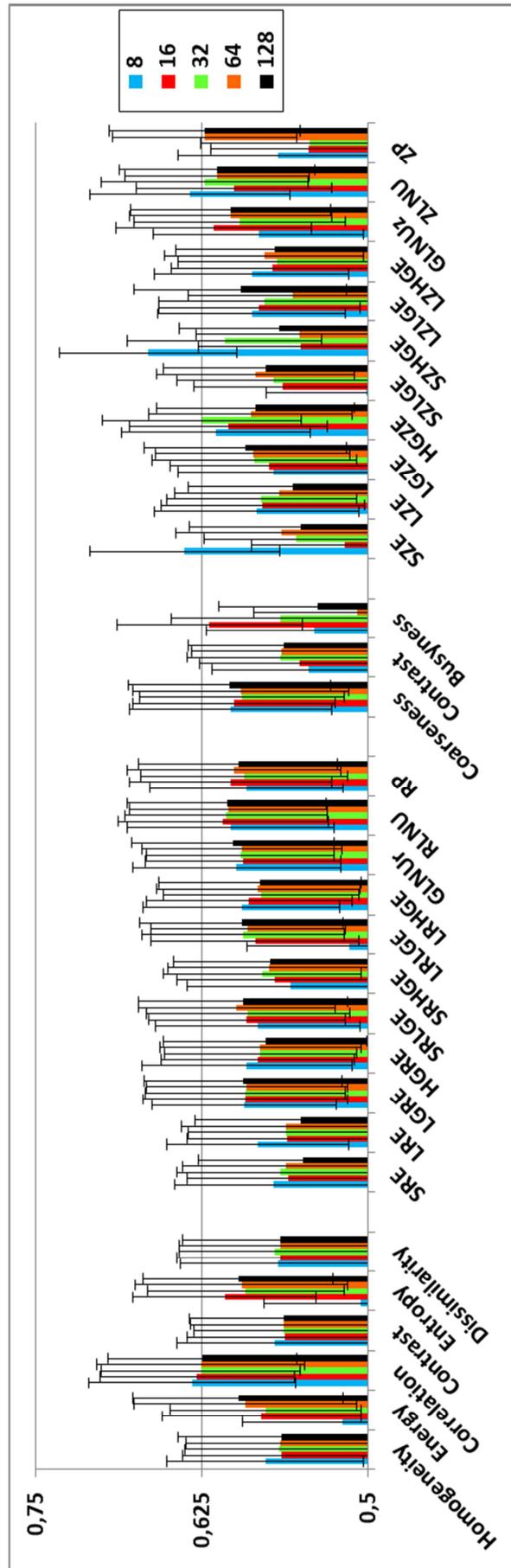


Figure V.6 : AUC (et écart-types correspondants) de chaque paramètre de texture en fonction du niveau de discrétisation utilisé. Afin de mieux mettre en évidence les différences potentielles, l'intervalle de valeur de ces AUC a été limité à [0,5-0,75] et non pas [0,5-1] comme c'était le cas dans la figure V.4 par exemple.

V.2.2.4.2. En fonction de la méthode utilisée pour segmenter le volume métabolique

Nous avons enfin comparé la variabilité des mesures de texture en fonction du volume métabolique utilisé : basé sur un seuil prenant en compte le fond local (S_{fond}) ou sur un seuil proportionnel au SUV maximal dans la tumeur ($S_{40\%}$).

Ne considérant que les données TEP prétraitement, l'utilisation de $S_{40\%}$ a fourni un $VM_{40\%}$ moyen, médian, maximal et minimal de 63,0 mL, 18,7 mL, 3,8 mL et 429,5 mL respectivement. Pour mémoire (cf paragraphe V.2.2.1), ces valeurs étaient respectivement de 73,9 mL, 20,8 mL, 5,0 mL et 519,0 mL pour VM_{fond} .

Le graphique de Bland Altman permet de caractériser l'accord entre les volumes métaboliques fournis par ces deux méthodes de segmentation (figure V.7) à partir des données TEP(PT) seules. L'écart moyen (et écart-type) entre $VM_{40\%}$ et VM_{fond} est de 10,8 ($\pm 30,3$) mL, avec des valeurs en moyenne plus élevées pour VM_{fond} . Seules deux tumeurs ont des $VM_{40\%}$ plus petits que la limite de 4,8 mL fixée précédemment pour l'analyse de texture (3,8 mL et 4,5 mL). Elles ont été conservées pour l'analyse comparative des index.

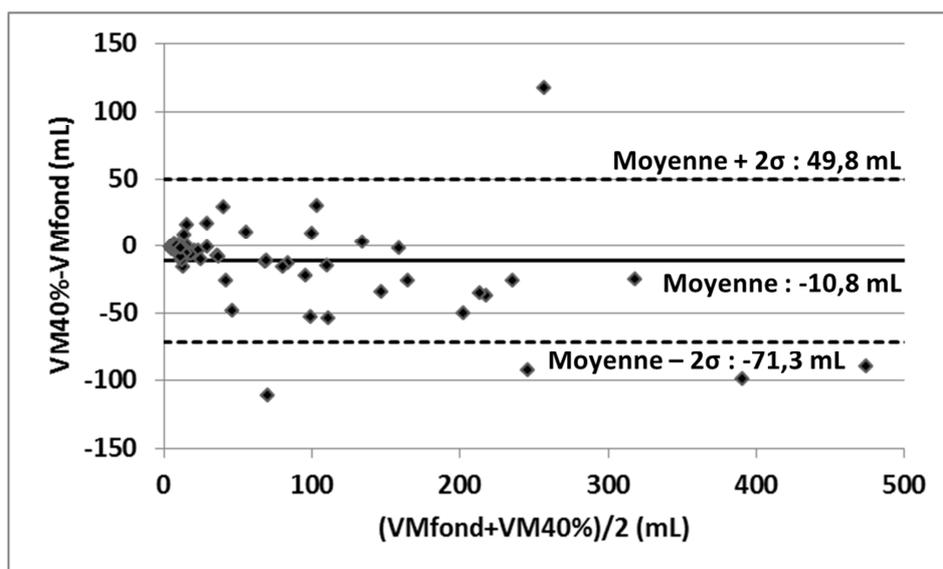


Figure V.7 : Graphique de Bland-Altman comparant les valeurs des volumes métaboliques obtenus en utilisant deux méthodes de segmentation.

Les écarts relatifs moyens obtenus par chaque index, sur les données TEP prétraitement et à partir des 2 méthodes de segmentation sont fournis en figure V.8. Seul *Busyness* obtient un $\bar{\Delta}$ supérieur à 5% en valeur absolue ($-8,3 \pm 480,3\%$). En dehors de cet index, on constate que les dispersions des valeurs autour de $\bar{\Delta}$ sont parfois importantes, comme c'est le cas pour *LGRE*, *GLNur*, *LGZE* par exemple. On note enfin la grande robustesse de certains paramètres de texture tels que *Homogeneity*, *Entropy*, *SRE*, *SZE* dont les écarts relatifs moyens sont comparables à ceux obtenus par les différents SUV rapportés dans cette étude.

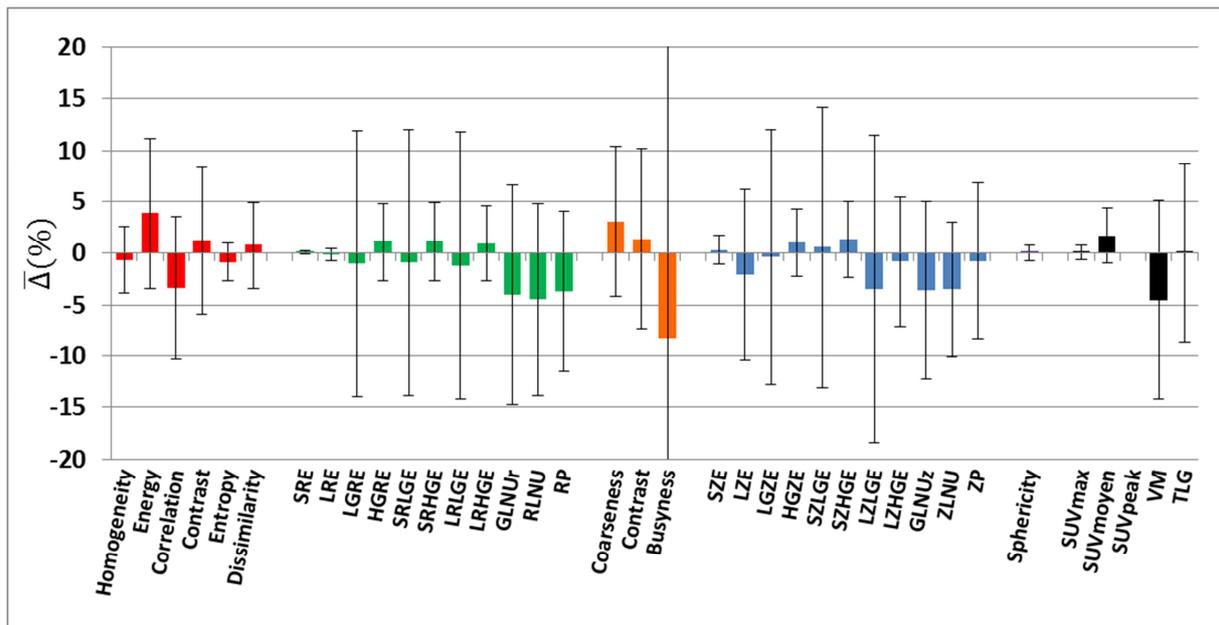


Figure V.8 : Ecarts relatifs moyens $\bar{\Delta}(\%)$ (± 1 écart-type) calculés en considérant toutes les tumeurs pour chaque index et déterminé à partir des méthodes de segmentation S_{fond} et $S_{40\%}$. La valeur présentée pour SUV_{moyen} correspond à l'écart relatif moyen calculé en comparant les SUV_{fond} et $SUV_{40\%}$. L'échelle choisie ne permet pas de représenter l'écart-type obtenu pour Busyness, qui vaut 480,3%.

Afin de déterminer l'impact de la segmentation non plus seulement sur les valeurs prises par les index mais aussi sur la capacité prédictive des index calculés sur la TEP prétraitement, nous avons rapporté les AUC (± 1 écart-type) des courbes ROC obtenues à partir de S_{fond} et $S_{40\%}$. Nous avons reproduit les résultats déjà présentés en figure V.4 pour faciliter la comparaison entre les méthodes de segmentation pour chacun des index étudiés (figure V.9). La comparaison par paire des AUC obtenues pour les 2 types de segmentation n'a pas permis de dégager de différences significatives.

Nous avons enfin cherché à évaluer si les AUC des différents index calculés étaient, dans l'ensemble, plus prédictifs, en utilisant l'une ou l'autre des méthodes de segmentation. L'hypothèse H_0 : « la distribution de la différence deux à deux des AUC obtenues par les 37 index étudiés en utilisant les deux méthodes de segmentation est symétrique, continue et de médiane nulle » a été rejeté avec une p-value égale à 0,002 (test de Wilcoxon), en faveur des résultats obtenus en utilisant S_{fond} .

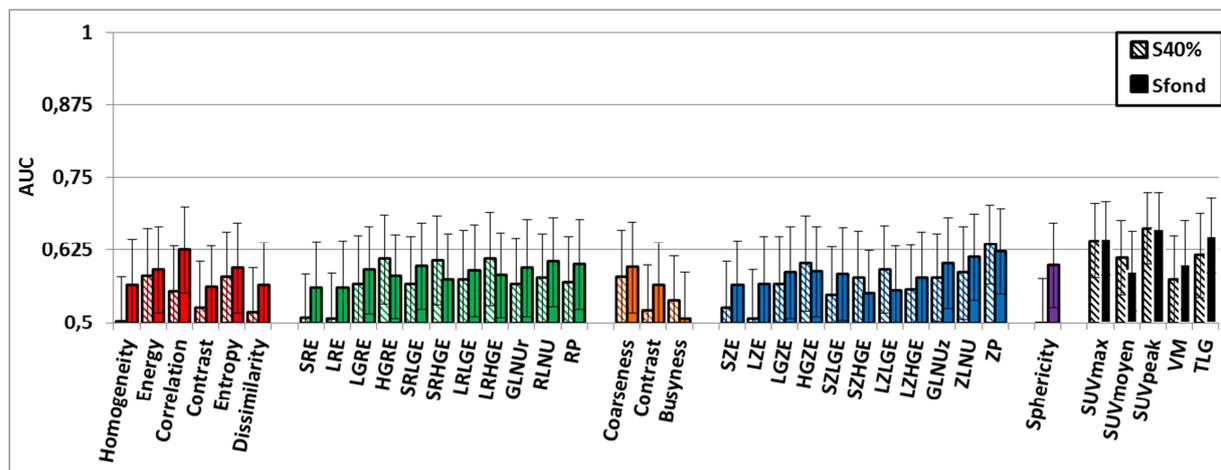


Figure V.9 : comparaison des AUC des courbes ROC des différents index obtenues en segmentant les volumes des lésions à l'aide d'un seuil égal à 40% du SUV_{max} (barres hachurées) et sur un seuil prenant en compte le fond local (barres pleines). Seul l'examen TEP(PT) est pris en compte ici. Les barres d'erreurs rapportées représentent 1 écart-type.

V.2.2.5. Corrélation entre les index métaboliques

Afin de mettre en évidence les relations entre les différents index métaboliques, nous avons rapporté les coefficients de corrélations entre leurs valeurs mesurées sur la TEP prétraitement seule. Le tableau V.8 donne les coefficients de Pearson entre les index issus de la matrice de cooccurrence. Toutes les valeurs sont supérieures à 0,78 (en valeur absolue).

	Homogeneity	Energy	Correlation	Contrast (inertia)	Entropy	Dissimilarity
Homogeneity	1,00	-0,78	0,90	-0,92	0,81	-0,97
Energy		1,00	-0,83	0,85	-0,92	0,84
Correlation			1,00	-0,86	0,86	-0,89
Contrast (inertia)				1,00	-0,82	0,99
Entropy					1,00	-0,84
Dissimilarity						1,00

Tableau V.8 : coefficients de Pearson obtenus entre les index extraits des matrices de cooccurrence et calculés sur la TEP(PT). En vert les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8.

De même, nous avons observé des corrélations très fortes entre certains index issus des matrices de longueur des séries homogènes d'une part et ceux provenant de la matrice de longueurs de zones homogènes d'autre part (tableaux V.9 et V.10). Des corrélations/anti-corrélations parfaites (coefficient de Pearson égal à 1 ou -1) ont été obtenues pour 5 couples d'index issus de la matrice de longueur des séries homogènes.

	SRE	LRE	LGRE	HGRE	SRLGE	SRHGE	LRLGE	LRHGE	GLNU _r	RLNU	RP
SRE	1,00	-1,00	0,64	0,05	0,65	0,08	0,62	-0,07	-0,79	-0,71	-0,75
LRE		1,00	-0,64	-0,04	-0,65	-0,07	-0,62	0,08	0,80	0,72	0,76
LGRE			1,00	-0,20	1,00	-0,18	1,00	-0,27	-0,52	-0,49	-0,53
HGRE				1,00	-0,19	1,00	-0,22	0,99	0,25	0,33	0,33
SRLGE					1,00	-0,17	1,00	-0,26	-0,52	-0,49	-0,53
SRHGE						1,00	-0,20	0,99	0,22	0,31	0,31
LRLGE							1,00	-0,29	-0,51	-0,49	-0,52
LRHGE								1,00	0,35	0,42	0,43
GLNU _r									1,00	0,98	0,98
RLNU										1,00	0,99
RP											1,00

Tableau V.9 : coefficients de Pearson obtenus entre les index extraits des matrices de longueur de séries homogènes. En vert les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu inférieure à 0,1.

	SZE	LZE	LGZE	HGZE	SZLGE	SZHGE	LZLGE	LZHGE	GLNU _z	ZLNU	ZP
SZE	1,00	-0,88	0,64	-0,10	0,66	0,24	-0,01	-0,74	-0,73	-0,63	-0,52
LZE		1,00	-0,58	0,14	-0,58	-0,15	0,15	0,82	0,82	0,71	0,48
LGZE			1,00	-0,29	0,99	-0,07	0,53	-0,58	-0,53	-0,50	-0,44
HGZE				1,00	-0,24	0,93	-0,44	0,58	0,42	0,50	0,55
SZLGE					1,00	-0,03	0,46	-0,56	-0,51	-0,48	-0,41
SZHGE						1,00	-0,39	0,29	0,14	0,25	0,32
LZLGE							1,00	-0,18	-0,06	-0,12	-0,33
LZHGE								1,00	0,89	0,85	0,74
GLNU _z									1,00	0,97	0,79
ZLNU										1,00	0,80
ZP											1,00

Tableau V.10 : coefficients de Pearson obtenus entre les index extraits des matrices de longueur de zones homogènes. En vert les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu inférieure à 0,1.

Compte tenu des liens entre les définitions mathématiques des index extraits de ces deux matrices, nous avons rapporté les corrélations entre les index extraits par ces deux méthodes (tableau V.11). En dehors des paires (*LRLGE*, *LZLGE*) et (*LRHGE*, *LZHGE*), nous avons obtenus des corrélations très élevées ($R > 0,81$).

	SRE	LRE	LGRE	HGRE	SRLGE	SRHGE	LRLGE	LRHGE	GLNUr	RLNU	RP
SZE	0,96	-0,96	0,64	0,05	0,64	0,08	0,62	-0,06	-0,75	-0,68	-0,72
LZE	-0,96	0,96	-0,57	-0,01	-0,58	-0,04	-0,55	0,10	0,86	0,77	0,80
LGZE	0,65	-0,65	0,99	-0,18	0,99	-0,16	0,98	-0,25	-0,52	-0,49	-0,53
HGZE	-0,11	0,12	-0,30	0,98	-0,30	0,97	-0,32	0,98	0,37	0,44	0,45
SZLGE	0,66	-0,66	0,96	-0,13	0,97	-0,11	0,95	-0,20	-0,51	-0,47	-0,51
SZHGE	0,21	-0,19	-0,08	0,94	-0,08	0,95	-0,10	0,90	0,09	0,18	0,17
LZLGE	-0,08	0,08	0,59	-0,46	0,58	-0,46	0,62	-0,45	-0,01	-0,05	-0,07
LZHGE	-0,78	0,79	-0,58	0,48	-0,58	0,46	-0,58	0,58	0,89	0,87	0,89
GLNUz	-0,77	0,78	-0,53	0,30	-0,53	0,27	-0,53	0,39	0,99	0,99	0,99
ZLNU	-0,67	0,68	-0,51	0,40	-0,51	0,37	-0,50	0,48	0,94	0,99	0,98
ZP	-0,49	0,49	-0,46	0,48	-0,46	0,46	-0,47	0,55	0,76	0,78	0,81

Tableau V.11 : coefficients de Pearson obtenus entre les index extraits des matrices de longueur de zones homogènes et ceux extraits des matrices de longueur de séries homogènes. En vert, les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu inférieure à 0,1.

De moindres corrélations ont été observées entre les grandeurs dérivées des matrices de différences de niveaux de gris (tableau V.12), où seules *Contrast* et *Coarseness* présentent une valeur de leur coefficient de Pearson intermédiaire (R=0,79).

	Coarseness	Contrast	Busyness
Coarseness	1,00	0,79	0,04
Contrast		1,00	0,02
Busyness			1,00

Tableau V.12 : coefficient de Pearson obtenus entre les index extraits des matrices de différences de niveaux de gris. En bleu, les coefficients dont la valeur absolue est inférieure à 0,1.

Enfin, pour évaluer l'apport d'information entre index de texture et index classiques, nous avons calculé systématiquement l'ensemble des corrélations entre les valeurs des index de texture et les index classiques ou de forme (tableau V.13). Les corrélations entre SUV et index de texture sont toutes faibles, voire nulles. En revanche, les corrélations calculées entre VM et ces mêmes index de texture sont parfois très élevées. Ainsi, la corrélation entre VM et *Entropy* d'une part, et VM et *ZP* d'autre part, est élevée (R=0,78). Des valeurs de corrélation encore plus importantes (>0,87) sont enfin obtenues entre VM et *GLNUr*, *RLNU*, *RP*, *LZHGE*, *GLNUz* ou encore *ZLNU*. Ces derniers index de texture caractérisent les dimensions ou l'intensité des zones homogènes des tumeurs.

	Sphericity	SUVmax	SUVfond	SUVpeak	VM	TLG
Homogeneity	-0,33	0,04	-0,25	0,01	0,62	0,45
Energy	0,09	-0,23	-0,03	-0,25	-0,58	-0,48
Correlation	-0,28	0,22	-0,03	0,24	0,67	0,55
Contrast(inertia)	0,26	-0,02	0,25	0,01	-0,56	-0,43
Entropy	-0,08	0,28	0,08	0,31	0,78	0,67
Dissimilarity	0,29	-0,03	0,25	0,00	-0,60	-0,45
SRE	0,23	-0,13	0,14	-0,11	-0,71	-0,54
LRE	-0,23	0,13	-0,14	0,12	0,73	0,56
LGRE	0,15	-0,11	0,06	-0,12	-0,49	-0,40
HGRE	0,26	0,37	0,51	0,49	0,32	0,39
SRLGE	0,15	-0,11	0,06	-0,12	-0,49	-0,40
SRHGE	0,26	0,36	0,51	0,48	0,30	0,37
LRLGE	0,14	-0,11	0,06	-0,12	-0,48	-0,40
LRHGE	0,24	0,39	0,50	0,51	0,41	0,46
GLNUr	-0,09	0,27	0,07	0,27	0,98	0,85
RLNU	-0,04	0,31	0,14	0,33	1,00	0,92
RP	-0,06	0,28	0,11	0,30	0,99	0,89
Coarseness	0,14	-0,25	-0,05	-0,27	-0,73	-0,60
Contrast	0,31	-0,04	0,23	-0,01	-0,56	-0,43
Busyness	0,10	-0,07	-0,04	-0,05	-0,05	-0,09
SZE	0,18	-0,11	0,15	-0,09	-0,68	-0,51
LZE	-0,20	0,15	-0,11	0,13	0,77	0,61
LGZE	0,16	-0,13	0,05	-0,13	-0,49	-0,40
HGZE	0,17	0,38	0,47	0,49	0,44	0,46
SZLGE	0,16	-0,14	0,05	-0,13	-0,47	-0,38
SZHGE	0,19	0,35	0,51	0,45	0,17	0,24
LZLGE	-0,11	-0,09	-0,15	-0,14	-0,05	-0,13
LZHGE	0,01	0,35	0,21	0,40	0,87	0,81
GLNUz	-0,08	0,28	0,10	0,29	0,99	0,88
ZLNU	-0,03	0,33	0,18	0,35	0,99	0,92
ZP	0,01	0,24	0,17	0,29	0,78	0,73

Tableau V.13 : coefficients de corrélations de Pearson entre index classiques et index de texture. En vert, les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu, inférieure à 0,1.

Pour illustrer le type de corrélations obtenues, nous avons ajouté le graphique $GLNU_z$ en fonction de VM_{fond} (figure V.10A). Pour évaluer l'influence du choix de la méthode de segmentation sur ces résultats, et notamment celles relatives au volume (colonne 6 du tableau V.13), nous avons aussi étudié les corrélations entre les mêmes index mais obtenus à partir de la segmentation utilisant $S_{40\%}$. A titre d'illustration, nous avons représenté le graphique de $GLNU_z$ calculé dans les volumes segmentés en utilisant $S_{40\%}$, en fonction de $VM_{40\%}$ (figure V.10B). Dans l'ensemble, des résultats similaires à ceux présentés dans le tableau V.11 ont été obtenus pour tous les index en utilisant $S_{40\%}$ (résultats non présentés).

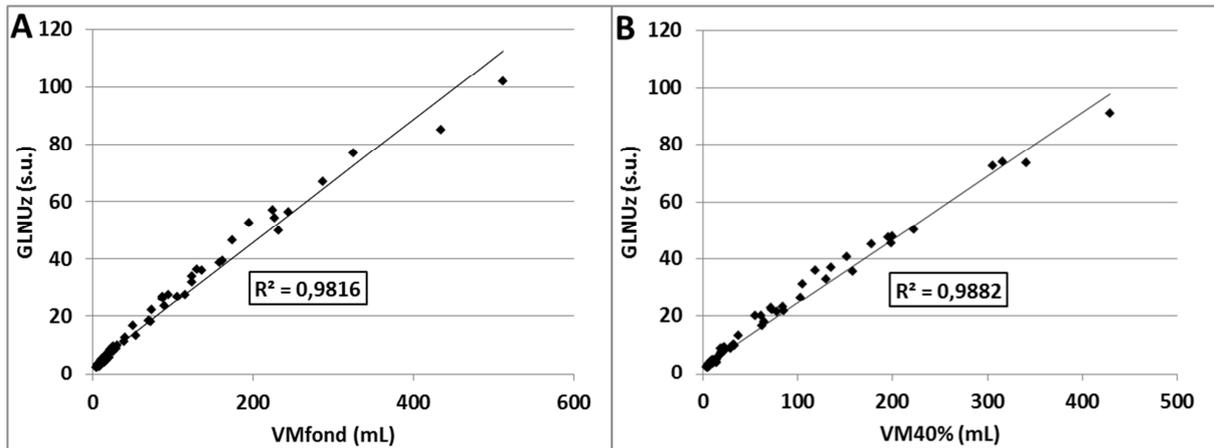


Figure V.10 : corrélation entre GLNUz et le volume métabolique calculé à partir du seuil S_{fond} (A) et $S_{40\%}$ (B)

V.2.2.6. Comparaison de la capacité prédictive des méthodes d'analyse de texture pour deux méthodes de ré-échantillonnage différentes

En considérant les volumes segmentés via le seuil S_{fond} et avec un paramètre de discrétisation de 64 niveaux de gris, nous avons comparé les capacités prédictives des index de texture à partir de la TEP prétraitement seule (figure V.11) ainsi que celles obtenues en calculant l'évolution relative (cf équation IV.9) des grandeurs après 14 jours de traitement (figure V.13) pour 2 méthodes de ré-échantillonnage différentes (R_{absolu} et $R_{relatif}$).

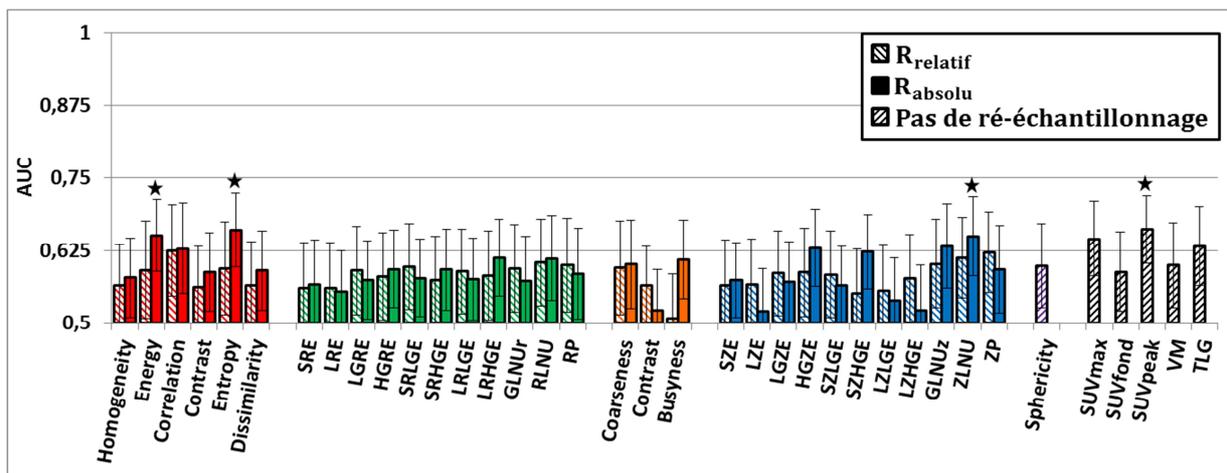


Figure V.11 : AUC (+/- 1 ET) des courbes ROC des index métaboliques pour deux méthodes de ré-échantillonnage à partir de la TEP prétraitement seule. * : AUC significativement supérieure à 0,5 ($p < 0,05$).

En se basant sur la TEP prétraitement seule, R_{absolu} permet d'améliorer la capacité prédictive de *Entropy*, *Energy* et *ZLNU*, rendant leurs AUC significativement supérieures à 0,5 ($p < 0,05$, figure V.11). Dans l'ensemble, le test des rangs signés de Wilcoxon montre qu'il n'y a pas de différence significative entre les différentes paires d'AUC ($p = 0,34$). La méthode de ré-échantillonnage proposée permet de rendre prédictifs *Entropy*, *Energy* et *ZLNU*. Cependant, la comparaison par paire d'AUC obtenue pour chacun des index de texture montre qu'aucune des AUC obtenue par la méthode de ré-échantillonnage mettant en œuvre une échelle de niveau de gris commune à toutes les tumeurs n'est différent de celle obtenue en utilisant la méthode de discrétisation décrite dans la littérature.

En appliquant la nouvelle méthode de ré-échantillonnage au calcul de l'évolution relative des différents index (figure V.12), 18/31 index de texture deviennent prédictifs ($AUC > 0,5$) sans pour autant dépasser les performances de SUV_{peak} . Seul un index (*RP*) voit son AUC décroître de manière significative, en utilisant R_{relatif} ($p = 0,004$). Le test de Wilcoxon visant à comparer les deux groupes de paires d'AUC obtenues par l'une et l'autre des méthodes montre une différence très significative entre les médianes ($p = 1,8 \times 10^{-4}$). Outre la différence significative observée entre les deux AUC obtenues pour RP en faveur de R_{relatif} , la comparaison des AUC de chaque index montre que R_{absolu} permet d'améliorer de manière significative le caractère prédictif d'*Homogeneity*, *Contrast (Inertia)*, *Dissimilarity*, *SRE*, *Contrast*, *LGZE* et *ZLNU* (étoiles rouges sur la figure V.12).

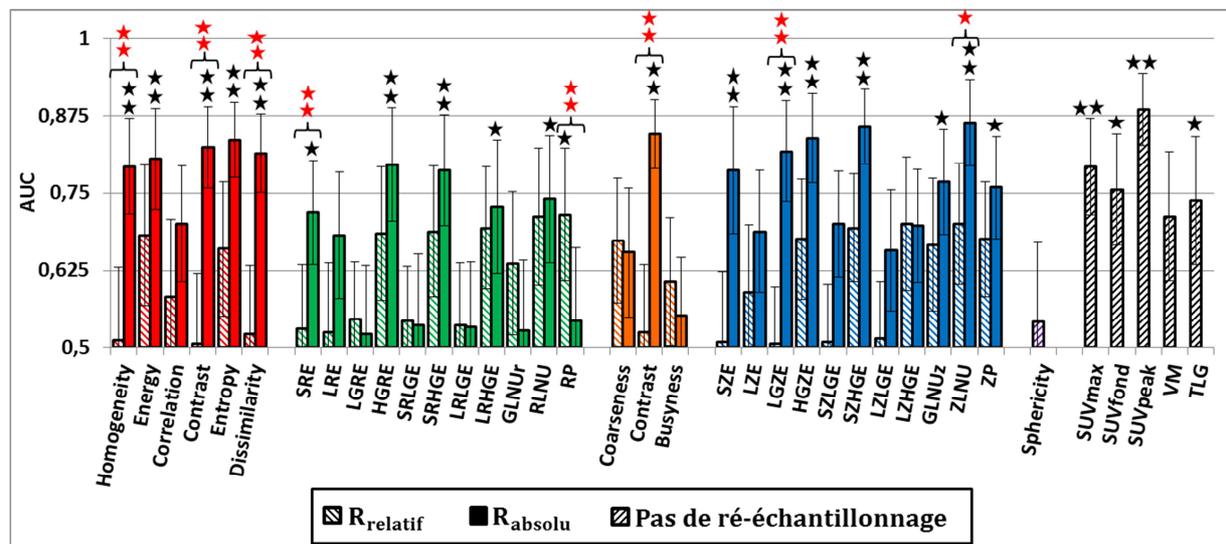


Figure V.12 : AUC (+/- 1 ET) des courbes ROC des index métaboliques pour deux méthodes de ré-échantillonnage, obtenues en calculant l'évolution relative des index entre l'examen prétraitement et après 14 jours de chimiothérapie. AUC significativement supérieure à 0,5 (étoiles noires telles que * : $p < 0,05$; ** : $p < 0,01$). Les étoiles rouges marquent les cas où les paires d'AUC obtenues pour chacun des index sont significativement différentes en utilisant l'une ou l'autre méthode de ré-échantillonnage.

Les corrélations entre index classiques et index de texture obtenus à partir de cette méthode de discrétisation sont donnés dans le tableau V.14.

	Sphericity	SUVmax	SUVmean	SUVpeak	Volume	TLG
Homogeneity	-0,40	-0,78	-0,89	-0,79	0,05	-0,12
Energy	-0,22	-0,67	-0,64	-0,69	-0,33	-0,40
Correlation	-0,28	0,24	-0,01	0,26	0,67	0,60
Contrast (inertia)	0,40	0,81	0,90	0,81	-0,09	0,05
Entropy	0,16	0,78	0,65	0,80	0,63	0,71
Dissimilarity	0,41	0,82	0,93	0,83	-0,08	0,07
SRE	0,33	0,61	0,74	0,62	-0,19	0,00
LRE	-0,29	-0,54	-0,66	-0,55	0,18	0,00
LGRE	-0,25	-0,50	-0,58	-0,52	-0,06	-0,14
HGRE	0,44	0,94	0,97	0,95	0,14	0,31
SRLGE	-0,26	-0,53	-0,61	-0,55	-0,08	-0,16
SRHGE	0,44	0,93	0,97	0,94	0,12	0,29
LRLGE	-0,21	-0,40	-0,48	-0,42	-0,02	-0,10
LRHGE	0,44	0,94	0,97	0,96	0,20	0,37
GLNUr	-0,17	0,02	-0,15	0,03	0,85	0,69
RLNU	-0,02	0,37	0,21	0,39	0,99	0,98
RP	-0,13	0,13	-0,03	0,15	0,96	0,84
Coarseness	0,13	-0,29	-0,08	-0,31	-0,68	-0,60
Contrast	0,40	0,46	0,67	0,48	-0,33	-0,22
Busyness	-0,04	0,11	0,12	0,12	-0,02	0,02
SZE	0,26	0,70	0,77	0,70	0,03	0,10
LZE	-0,12	-0,24	-0,26	-0,23	0,17	0,01
LGZE	-0,19	-0,42	-0,48	-0,43	-0,08	-0,13
HGZE	0,40	0,96	0,96	0,96	0,19	0,35
SZLGE	-0,24	-0,52	-0,59	-0,54	-0,08	-0,17
SZHGE	0,38	0,95	0,94	0,93	0,13	0,29
LZLGE	-0,13	-0,25	-0,27	-0,25	0,13	-0,01
LZHGE	-0,08	-0,23	-0,27	-0,22	0,27	0,09
GLNUz	-0,02	0,41	0,26	0,43	0,97	0,97
ZLNU	0,09	0,59	0,44	0,61	0,87	0,96
ZP	0,27	0,80	0,79	0,81	0,33	0,49

Tableau V.14 : corrélation entre index de texture et de forme et index de texture calculées sur les TEP prétraitement en utilisant 64 niveaux de gris et un ré-échantillonnage commun pour toutes les tumeurs de la cohorte. En vert, les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu, inférieure à 0,1.

La méthode de ré-échantillonnage utilisée introduit de fortes corrélations entre les SUV (maximum, moyen ou peak) et *Homogeneity*, *Contrast*, *Entropy*, *Dissimilarity*, *HGRE*, *SRHGE*, *LRHGE*, *HGZE*, *SZHGE* et *ZP* (colonnes 3, 4 et 5).

V.2.3. Discussion

Il a récemment été montré que l'analyse de texture du métabolisme glucidique intra-tumoral vu par la FDG-TEP avait, dans des cas de cancers de l'œsophage, une valeur pronostique prédictive supérieure à celle des index SUV classiquement utilisés en clinique (Tixier et al., 2011). L'intérêt de cette approche serait de mesurer des informations issues de la reconnaissance de forme visuelle, dont l'intérêt potentiel a été discuté par (Hofman and Hicks, 2010), ou encore de caractériser d'autres informations physiologiques d'intérêt mises indirectement en évidence par la fixation de FDG dans les tissus tumoraux (eg : (Pugachev et al., 2005)).

A partir des résultats constatés dans la littérature, nous avons d'abord cherché à valider l'intérêt de cette approche dans le cas de cancers colorectaux métastatiques, quand elle était appliquée à des données TEP acquises avant le début du traitement. Nous nous sommes donc concentrés sur les index décrits dans la littérature comme étant prédictifs de la réponse anatomique de référence en clinique.

Les résultats obtenus dans notre cas (figure V.4) montrent clairement qu'aucun index de texture ou de forme n'est plus prédictif de la réponse à la thérapie que le SUV_{peak} , dont l'AUC, égale à 0,63, est la seule à être significativement supérieure à 0,5. Cette tendance est inverse à celle décrite par la littérature dans le cas de cancers de l'œsophage (où l'ensemble des AUC des index classiques étaient inférieures à celles obtenues par les méthodes d'analyse de texture pour (Tixier et al., 2011)).

Nous avons aussi cherché à évaluer la capacité prédictive de l'évolution de ces index de texture entre une TEP prétraitement et une TEP acquise 14 jours après le début de la chimiothérapie (figure V.5). Si la faible taille de la cohorte de tumeurs étudiée (45 lésions dont seulement 9 répondantes) rend la comparaison entre les capacités prédictives des méthodes difficile, on constate cependant que la tendance observée précédemment se confirme. Les méthodes mesurant la fixation moyenne ou maximale de FDG dans le volume tumoral apparaissent prédictives de la réponse tumorale. A contrario, dans ce contexte, seul un index de texture (*Run Percentage*) a obtenu une AUC significativement supérieure à 0,5, avec une valeur de p à peine inférieure à la limite fixée ($p=0,047$). Ce résultat, qui ne permet donc pas d'invalider pour autant les performances très comparables de *LRHGE*, *RLNU*, *SZHGE*, *LZHGE* et *ZLNU*, dont les AUC sont toutes supérieures à 0,69, est cependant représentatif de l'absence de valeur ajoutée des index de texture dans le contexte de l'évaluation de la réponse tumorale précoce des CCM.

Ces différences par rapport à la littérature constatées, nous avons cherché à caractériser l'impact du choix de différents paramètres sur les performances et la variabilité des index de texture.

Le premier paramètre étudié a été le nombre de niveaux de gris utilisés pour le ré-échantillonnage. Pour évaluer son influence, nous avons rapporté les AUC des courbes ROC pour 5 valeurs du paramètre (figure V.6). Cette comparaison a mis en évidence l'invariance du résultat précédent (figure V.4) obtenu pour 64 valeurs de niveaux de gris, pour la plupart

des paramètres de texture. On constate cependant que pour 4 index de texture (*Energy*, *Entropy*, *LRLGE* et *SZLGE*), le choix d'une discrétisation sur 8 niveaux de gris fait chuter leurs performances de classification. De même, on constate que les performances de *Busyness*, *SZE*, *SZHGE* et *ZP* semblent assez sensibles au nombre de niveaux de gris considérés pour l'analyse. Ces résultats, à confirmer sur une plus grosse cohorte de tumeurs, peuvent cependant servir de base de travail pour favoriser l'étude plus approfondie de certains index de texture par rapport à d'autres.

En nous appuyant sur les données de (Tixier et al., 2011), nous avons analysé les différences méthodologiques pouvant expliquer les tendances observées. Pour discuter nos résultats par rapport à l'existant, nous avons rapporté, dans les tableaux V.6A et V.6B, les correspondances entre les noms donnés aux différents index utilisés ici et ceux utilisés dans les travaux de (Tixier et al., 2012; Tixier et al., 2011). Une difficulté est que (Tixier et al., 2012) ne déclarent pas de changements de définition par rapport à l'article de 2011 et effectuent malgré tout une corrélation directe entre les résultats présentés dans ces deux articles. Les choix faits pour établir les correspondances entre les noms des index utilisés dans notre étude et ceux proposés dans les travaux successifs de Tixier et al. se sont appuyés:

- Sur les définitions et descriptions données aux différents index extraits de R dans le domaine du traitement d'image (Tang, 1998; Thibault, 2009; Xu et al., 2004).
- Sur le fait que ces références soient elles-mêmes utilisées par (Tixier et al., 2011) et qu'elles correspondent alors aux définitions les plus courantes.
- Sur l'interchangeabilité entre les termes employés pour nommer les index extraits de R par (Tixier et al., 2011), et les termes utilisés dans la littérature. A titre d'exemple, *intensity* devient *gray-level* (cf tableau V.6A colonnes 3 et 4). Ces inversions se retrouvent logiquement dans les noms des index extraits de Z dans cette publication (cf tableau V.6B, col 3 et 4). Ne considérant plus R dans leur seconde publication, et au regard des formules données pour *SZE* ou *LZE* par exemple, il paraît légitime de remplacer le terme *zone* par *area* et entre *small* par *short*, pour comparer (Tixier et al., 2011) et (Tixier et al., 2012).

Le premier article de cette équipe ne précise pas les dimensions des voxels ni la résolution spatiale dans les images reconstruites. Il est donc impossible d'émettre d'hypothèses sur le fait que le tomographe utilisé dans leurs études permette de distinguer des variations de fixation de FDG plus fines que celles visibles avec le tomographe que nous avons utilisé. Nous avons relevé les différences suivantes entre notre étude et la leur :

- Le volume minimal de lésion considéré : 10 mL contre 5 mL dans la présente étude.
- Le type de lésions : œsophages primaires contre colorectales métastatiques.
- Méthode de segmentation : *Fuzzy Locally Adaptive Bayesian* (Hatt et al., 2009) contre une méthode par croissance de région prenant en compte le fond local.

Pour ne pas réduire trop la cohorte de lésions, nous avons décidé de n'exclure que les volumes les plus affectés par l'EVP, compte tenu de la résolution de nos images reconstruites (7 mm).

Dans leurs travaux, (Tixier et al., 2011) n'indiquent pas la distribution des volumes des lésions analysées, soulignant uniquement que seules les tumeurs primaires avaient été étudiées du fait du trop petit nombre de voxels dans les lésions métastatiques et ganglionnaires. Nous avons considéré ici qu'un volume médian de tumeurs égal à 20,8 mL était suffisant pour appliquer une analyse de texture. En accord avec le volume minimal des lésions considérées par (Tixier et al., 2011) dans le cas des cancers de l'œsophage (10 mL), nous avons cependant validé le fait que les tendances observées sur la figure V.4 étaient conservées lorsque seules les tumeurs de plus de 10 mL étaient considérées (résultats non présentés du fait du nombre trop faible de lésions répondantes considérées pour l'analyse).

Une seconde différence pouvant expliquer les faibles performances des index de texture dans notre cas provient de la nature des lésions étudiées. En effet, l'étude présentée ici se base en majeure partie sur des métastases et seules 3% des lésions considérées sur la TEP(PT) sont des tumeurs primitives (cf tableau V.1). L'hétérogénéité de ces lésions dépend tout d'abord de leurs origines clonales, c'est-à-dire des particularités de la cellule qui leur a donné naissance (Langley and Fidler, 2007). Autrement dit, deux métastases issues de la même tumeur primaire peuvent présenter des caractéristiques différentes. De plus, contrairement à Tixier et al., notre cohorte regroupait 82% de lésions localisées dans le foie (tableau V.1). Or l'hétérogénéité d'une métastase donnée dépend aussi de son lieu d'implantation (Langley and Fidler, 2007). Enfin, les différences observées pourraient être dues à la différence de vascularisation entre le foie et l'œsophage. En effet, la prolifération tumorale est liée au degré d'hypoxie (Chan and Giaccia, 2007) qui détermine en partie la fixation de FDG de manière tumeur dépendante (Pugachev et al., 2005; Rajendran et al., 2004).

Une autre différence méthodologique importante entre l'étude de Tixier et al. et la nôtre vient de la méthode de segmentation utilisée. Pour étudier l'impact de ce paramètre sur les capacités prédictives et sur la variabilité des index de texture, nous avons considéré une autre méthode de segmentation souvent rapportée dans la littérature et basée sur un seuil de 40% du SUV maximum de la tumeur. L'étude du graphique de Bland-Altman comparant les VM fournis par les 2 méthodes a démontré un écart moyen de 10,8 mL (figure V.7). Bien qu'importante absolue, cette valeur correspond à un écart relatif de ~15% par rapport au volume métabolique moyen des lésions obtenu en utilisant S_{fond} (cf paragraphe V.3.1). La variabilité des valeurs prises par certains index de texture en fonction de la méthode de segmentation (figure V.8) est comparable à celle, très faible, du SUV maximum (*SRE*, *LRE*, *SZE* et *Sphericity*). Elle est similaire à celle obtenue par le SUV moyen calculé à partir de ces deux régions pour *Homogeneity*, *Entropy* et *Dissimilarity*, issus de la matrice de cooccurrence, pour lesquels une répétabilité voisine de celle du SUV_{max} a été rapportée par (Tixier et al., 2012). La variation moyenne est cependant aussi importante que celle du VM pour *Energy* (dont la valeur a été considérée comme peu dépendante des paramètres de reconstruction ou d'acquisition par (Galavis et al., 2010)) *GLNU_z* et *ZLNU*, dont les forts potentiels prédictifs et les faibles variabilités ont été soulignés à deux reprises par (Tixier et al., 2012; Tixier et al., 2011). Même si les performances de classification des index de textures apparaissent dans l'ensemble légèrement supérieures en utilisant S_{fond} , cette comparaison par index des paires d'AUC permet de conclure que les faibles performances des

index de texture ne viennent pas du choix de la méthode de segmentation utilisée. En effet, SUV_{peak} reste le seul index à être prédictif de la réponse RECIST établie après 6-8 semaines de traitement.

Pour mieux comprendre les liens entre les grandeurs étudiées et voir en quoi elles apportaient une information différente de celle fournie par les index classiques, nous avons étudié les corrélations entre ces différentes grandeurs. Nous avons pour cela considéré une discrétisation de 64 valeurs et le volume métabolique segmenté par la méthode de (Nestle et al., 2005). Une première constatation porte sur les nombreuses corrélations élevées obtenues entre les différents paramètres déduits de la matrice de cooccurrence (tableau V.8). Les résultats présentés à la figure V.8 montrent la bonne reproductibilité des valeurs d'*Entropy* en fonction de la segmentation utilisée. De plus, les résultats de (Galavis et al., 2010) mettent particulièrement en valeur la robustesse de cet index (cf tableau II.9 colonne 5). Enfin, ceux de (Tixier et al., 2011) prouvant la supériorité de sa capacité prédictive ainsi que sa bonne répétabilité (Tixier et al., 2012) nous permettent de souligner encore l'intérêt potentiel de cet index. Nous souhaitons cependant mettre l'accent sur le fait que nos résultats (figures V.5 et V.4) ainsi que ceux de (El Naqa et al., 2009), portant sur deux cohortes de patients, certes très réduites, ne semblent pas valider son intérêt pour toutes les localisations (cf tableau II.9 colonnes 3 et 4). Enfin, comme indiqué par (Tixier et al., 2011), des corrélations très élevées (dans notre cas, $R > 0,94$ cf tableau V.11) ont été trouvées entre la plupart des définitions communes utilisées pour définir les index issus des matrices de séries de longueurs et de zones homogènes (cf tableau V.5), en dehors des paires *LZLGE/LRLGE*, *LZHGE/LRLGE* et *ZP/LP* pour lesquelles des valeurs de $R < 0,81$ ont été obtenues.

N'ayant pas étudié les paramètres extraits des matrices de longueurs de séries homogènes, (Galavis et al., 2010) ont cependant relevé une faible variabilité de *Gray-level Non Uniformity (GLNU_r)*, *Low Gray level Run Emphasis (LGRE)* et *High Gray-level Gray Run Emphasis (HGRE)* en fonction des paramètres de reconstruction, ce qui incite à considérer sérieusement ces index pour d'autres études. De la même manière, bien que relatifs à la matrice *Z*, et outre des différences notables entre les définitions utilisées dans leurs deux publications successives, Tixier et al. ont aussi insisté sur l'intérêt de ces index. Selon ces auteurs, les répétabilités de *GLNU_z* et de *ZLNU* sont comparables à celle de SUV_{max} ((Tixier et al., 2012), cf tableau II.9). Les corrélations relevées dans notre étude entre ces deux index et le volume métabolique ($R = 0,99$, cf tableau V.13 et figure V.10A, et ce quel que soit la segmentation considérée (figure V.10B)) expliquent probablement l'absence de valeur ajoutée de ces paramètres par rapport au volume métabolique dans notre étude. L'intérêt prédictif (Tixier et al., 2011) ainsi que la variabilité test-retest (Tixier et al., 2012) du volume métabolique n'ayant pas été étudiés pour les cas des cancers de l'œsophage dans ces deux études, aucune information sur d'éventuelles corrélations entre VM et les index de texture n'a été rapportée dans ces travaux. Le volume métabolique étant particulièrement prédictif de la réponse au traitement pour cette localisation (Hatt et al., 2012a; Hyun et al., 2010), à la différence des cas de CCM étudiés ici (figure V.4, V.5), il serait intéressant de comparer les capacités prédictives des volumes métaboliques à celles de *ZLNU* et *GLNU_z* par exemple, dans le cas de cancers de l'œsophage. Ce travail permettrait de valider l'ajout d'information par l'analyse

de texture par rapport au simple calcul du volume métabolique. En particulier, nous pensons que la comparaison d'un grand nombre de paramètres, qui justifie l'approche systématique mise en œuvre par Tixier *et al.*, Galavis *et al.* et nous même, doit non seulement permettre de :

- Démontrer la valeur ajoutée d'approches complexes et nécessitant une interprétation particulière, pas aussi simple et directe que celles du volume ou de l'activité métabolique d'une lésion.
- Permettre de s'assurer que les nouvelles informations extraites des données TEP par de nouvelles méthodes d'analyses apportent réellement une nouvelle connaissance sur la complexité du métabolisme glucidique des tumeurs. La validation de l'apport de nouvelles informations est d'autant plus importante qu'un très petit nombre de paramètres est extrait des données TEP à l'heure actuelle. On gardera par exemple en tête le problème de méthodologie rencontré par (Brooks and Grigsby, 2011) (cf II.3.3.1.3) dont la définition de l'hétérogénéité s'est avérée être, par construction, comparable à la valeur du volume métabolique.

Enfin, nous avons évalué la capacité prédictive des index de textures quand la notion de fixation de FDG était réintroduite dans la mesure de texture par le biais de la méthode de ré-échantillonnage utilisée. Nous avons pour cela utilisé une échelle commune de niveaux de gris pour toutes les tumeurs. La motivation de l'évaluation de cette méthode repose sur le fait que la caractérisation de texture par les index ne doit pas être liée à l'agencement d'un nombre artificiellement élevé de niveaux de gris, mal adapté à des tumeurs dont les variations d'intensité des SUV des voxels qui la constituent seraient trop faibles. En d'autres termes, nous avons cherché à éviter de mesurer des textures représentatives de cas où les différences entre les activités correspondant à un niveau de gris et son suivant ne seraient pas significatives de différences physiologiques ou biologiques entre les zones de la tumeur mais au contraire liées aux seules limites de la TEP en tant qu'instrument de mesure (eg : fort bruit dans l'image, faible résolution etc).

R_{absolu} ne permet pas d'augmenter significativement les performances des index (seuls 3 d'entre eux deviennent faiblement prédictifs et ne dépassent pas les performances de SUV_{peak} , figure V.11). On constate cependant que dans le cas où l'on mesure l'évolution relative entre le début de la chimiothérapie et 14 jours de traitement, 18 index voient leurs capacités prédictives augmenter suffisamment pour conduire à une AUC supérieure à 0,5, comparable à celle obtenue par SUV_{peak} , contre 1 seul (RP) en faveur de la méthode de ré-échantillonnage décrite dans la littérature (figure V.4). Pour 7 d'entre eux, l'augmentation de la valeur prédictive, évaluée en comparant chaque pair d'AUC obtenue par R_{absolu} d'une part et R_{relatif} d'autre part, s'est avérée significative (figure V.12). La réintroduction de la notion d'activité dans le ré-échantillonnage augmente logiquement la corrélation entre les différentes définitions du SUV étudiées et les index de texture (tableau V.14). L'intérêt de cette normalisation reste à évaluer sur une cohorte étendue pour valider le couplage intrinsèque des informations de SUV et de texture en un seul index.

V.3. Conclusion

L'application de l'analyse de texture à des cas de cancers colorectaux métastatiques n'a pas permis de valider les résultats observés dans le cas de cancers de l'œsophage à partir d'une TEP prétraitement seule. La mise en œuvre de ces méthodes dans le contexte du suivi thérapeutique précoce n'a elle non plus pas amélioré la classification vue par le SUV, couramment employé en routine clinique. D'après l'ensemble de nos résultats, il apparaît donc que les paramètres de texture, leur applicabilité, et leur supériorité par rapport aux paramètres utilisés jusqu'à maintenant restent, selon nous, à démontrer. L'étude des différences méthodologiques entre les études précédemment publiées et notre étude n'a pas permis d'expliquer totalement les différences de conclusions. Cependant, l'étude de la corrélation entre les index de texture les plus mis en avant dans la littérature a permis de montrer qu'ils étaient particulièrement liés au volume métabolique, qui, dans le cas de cancers colorectaux, est peu prédictif de la réponse au traitement, à l'inverse des cancers de l'œsophage par exemple. L'étude d'une nouvelle méthode de discrétisation, basée sur une mise à l'échelle des valeurs de SUV commune à toutes les tumeurs a permis d'augmenter la capacité prédictive des index de texture de manière significative pour 7/31 index, les rendant prédictif dans 18 cas. L'intérêt de cette approche reste à valider sur une cohorte plus importante.

Chapitre VI

Evaluation d'une méthode d'imagerie paramétrique de la réponse tumorale

VI.1. Introduction : interprétation de la réponse tumorale sur des séries de plus de 2 examens TEP

En oncologie, le suivi du patient à l'aide de la TEP/TDM au FDG repose souvent sur une évaluation visuelle de l'évolution des lésions entre les examens, comme c'est le cas pour les patients atteints de cancers du poumon non à petites cellules (ou NSCLC, *Non Small Cell Lung Cancers*) par exemple (Hicks, 2009). Pourtant, dans le but de normaliser les pratiques entre les institutions, la quantification est de plus en plus fréquemment utilisée (Boellaard, 2009; Wahl et al., 2009; Young et al., 1999). Malgré les problèmes de définition relatifs, notamment, au choix du VOI le mieux adapté à la mesure (cf paragraphe II.3.1.3), le SUV semble s'imposer comme l'index quantitatif de référence. Quelle que soit la définition retenue, le SUV décrit une information moyenne du métabolisme de la tumeur.

Pour dépasser cette approche très simplificatrice de la description du métabolisme tumoral, quelques méthodes plus sophistiquées ont été proposées dans la littérature, telles que celles que nous avons discutées au paragraphe II.3.3 et celles visant à rendre compte de l'hétérogénéité de la distribution de FDG au sein du volume tumoral (cf paragraphe II.3.3.1). Cependant, dans ces approches aussi, une caractéristique du tissu tumoral est résumée par une valeur unique. Quelques méthodes ont aussi été proposées pour caractériser l'évolution tumorale à l'échelle du voxel, entre deux examens TEP, et ainsi aboutir à une quantification plus précise des informations extraites de ces données (cf paragraphe II.3.3.2).

Le suivi des patients consiste à comparer entre eux au moins deux examens (en général un acquis avant, l'autre pendant/après le traitement), mais parfois plus. Si des recommandations ont été proposées pour évaluer la réponse tumorale dans le cas où deux examens ont été effectués (Wahl et al., 2009; Young et al., 1999), il n'y a, à l'heure actuelle, pas de recommandations relatives à la caractérisation de la réponse tumorale lorsque plus de deux TEP doivent être comparés.

Outre l'usage de la TEP pour le suivi thérapeutique, les progrès récents de la radiothérapie (RT) vont aussi vers une intégration des informations métaboliques lors de la planification de traitement. En effet, le développement de méthodes de traitement de haute précision telles que la RT par modulation d'intensité (appelée IMRT pour *Intensity Modulated RT*, en anglais) permettent de délivrer des doses d'irradiation homogènes très élevées aux volumes cibles tout en épargnant les tissus sains voisins. Ces approches requièrent une définition très précise du volume tumoral macroscopique (ou GTV pour *Gross Tumor Volume*) à traiter. Dans ce contexte, l'utilisation des images anatomiques pour définir les volumes cibles peut entraîner la définition d'un GTV significativement différent du volume tumoral réel. C'est le cas par exemple quand les caractéristiques, vues par la modalité d'imagerie utilisée, entre les tissus tumoraux et les tissus sains voisins, sont similaires. De la même manière, ces modalités ne permettent généralement pas de distinguer l'envahissement des ganglions les plus proches de la lésion, alors qu'il est important de les inclure dans les volumes à traiter durant la RT.

La TEP/TDM, qui est capable de distinguer les lésions bénignes des lésions malignes tout en fournissant une information de localisation précise pour la RT, a donc un rôle important à

jouer pour l'amélioration des plans de traitement. Mieux encore, le développement de radiotraceurs tel que le FMISO (cf tableau I.5), un marqueur de l'hypoxie et donc des tissus tumoraux les plus résistants aux rayons ionisants, pourrait permettre de cibler encore plus spécifiquement certaines zones tumorales. L'intégration des informations issues de la TEP dans les plans de traitements en RT pourrait donc permettre non seulement d'améliorer la définition des GTV mais aussi d'augmenter spécifiquement les doses délivrées aux zones des volumes tumoraux les plus radio résistantes (via le FMISO) ou les plus agressives (avec le FDG), pour tenter d'améliorer le contrôle tumoral. Cette approche appelée, *dose painting* (Bentzen, 2005; Ling et al., 2000), est, à l'heure actuelle, l'une des méthodes de traitement en RT les plus étudiées et les plus prometteuses (Grégoire et al., 2012). Toujours dans l'idée d'améliorer le contrôle local de la maladie et aussi de diminuer la dose aux tissus sains, certains auteurs (Yan et al., 1997) ont enfin proposé de modifier le traitement par RT, planifié avant le début de la thérapie, au cours du traitement. Cette approche permet alors d'adapter la distribution de doses aux transformations anatomiques du patient engendrées par l'irradiation. On parle alors de *radiothérapie adaptative*.

Ces méthodes de traitement par RT offrent un rôle potentiel à la TEP, qui pourrait permettre d'améliorer les plans de traitement pour prendre en compte les particularités biologiques (*dose painting*) et les changements (*radiothérapie adaptative*) locaux de la tumeur.

Compte tenu de la complexité de l'évolution tissulaire de la tumeur et de sa réponse au traitement, et en conformité avec les travaux de (Necib et al., 2011), nous émettons l'hypothèse que la réponse de la tumeur serait mieux caractérisée par l'analyse des changements au niveau du voxel. L'objectif du travail que nous allons présenter dans ce chapitre est d'étudier la valeur ajoutée d'une méthode d'imagerie paramétrique qui décrit les modifications de la tumeur au fil du temps de façon plus complète que la comparaison de deux valeurs pour chaque examen. À cette fin, le laboratoire a développé une approche d'imagerie paramétrique qui permet de résumer la réponse tumorale vue par un certain nombre de TEP/TDM successives sur une seule image paramétrique, tout en fournissant des index utiles pour la classification de la réponse.

Notre travail a consisté à valider cette approche et à la comparer aux résultats fournis par des méthodes d'analyse plus classiques sur une série de données de 9 patients atteints d'un NSCLC traités par RT ou radio-chimiothérapie (RT-CT).

VI.2. L'Analyse Factorielle appliquée au contexte du suivi thérapeutique

Pour étudier l'évolution temporelle de lésions cancéreuses à l'échelle du voxel, nous avons utilisé l'analyse factorielle, qui permet d'analyser des séries d'images indexées par le temps. La méthode SULTAN appliquée ici (pour SUIvi Longitudinal en Tomographie par ANalyse factorielle) a été développée par (Necib, 2010) et correspond à une application de l'Analyse Factorielle de Séquences d'Images Médicales (AFSIM) utilisée initialement sur des séries

d'images scintigraphiques (Benali et al., 1994; Buvat et al., 1998; Di Paola et al., 1982). Par la suite, l'AFSIM a été mise en œuvre sur des séries d'images dynamiques que ce soit en TEP ou en IRM par exemple (Bonnerot et al., 1992; Frouin et al., 1999; Frouin et al., 2001; Janier et al., 2002; Klein et al., 2010; Millet et al., 2012). Cette méthode a aussi été utilisée sur des séries d'images TEMP ou scintigraphiques en indexant les séries d'images sur l'énergie et non plus sur le temps afin de séparer la contribution des photons diffusés de celle des photons primaires (eg : (Mas et al., 1990)). Quelle que soit l'application, l'AFSIM estime un petit nombre de courbes, appelées *facteurs* (représentant des Courbes Activités Temps – TAC ou des spectres selon que les images sont indexées en temps ou en énergie, respectivement), caractéristiques des tendances majeures dans la série d'images initiale. On associe à ces facteurs des images, appelées *images factorielles*, qui représentent les compartiments. Les facteurs et images factorielles associées correspondent à une décomposition linéaire du signal contenu dans la séquence d'images initiale.

Dans l'application que nous en avons faite, l'AFSIM analyse des séries de données acquises à plusieurs jours d'intervalles, pendant le suivi thérapeutique (cf figure VI.1). Dans ce contexte, SULTAN est utilisé pour décomposer une séquence d'images TEP statiques S_{it} (où i indexe un voxel du volume étudié tel que $i \in [1; N]$ avec N égal au nombre de voxels dans l'image et $t \in [1; T]$ où T correspond au nombre d'exams) en un petit nombre d'images factorielles $\tilde{A}_k(i)$ associées aux facteurs $\tilde{F}_k(t)$ avec $k \in [1; K]$ où K est le nombre de facteurs (généralement compris entre 2 et 4). L'objectif de la méthode est de décrire l'évolution temporelle du signal dans chaque voxel à partir de la combinaison linéaire des K facteurs tel que :

$$S_{it} = \sum_k \tilde{A}_k(i) \times \tilde{F}_k(t) + \varepsilon_{it} \quad \text{Eq. VI.1}$$

où ε_{it} représente le bruit dans l'image et les erreurs du modèle.

Dans l'équation VI.1, seul S_{it} est connu, alors que les images, les facteurs et le terme d'erreur ε sont estimés. Nous détaillons maintenant la méthode de résolution du modèle.

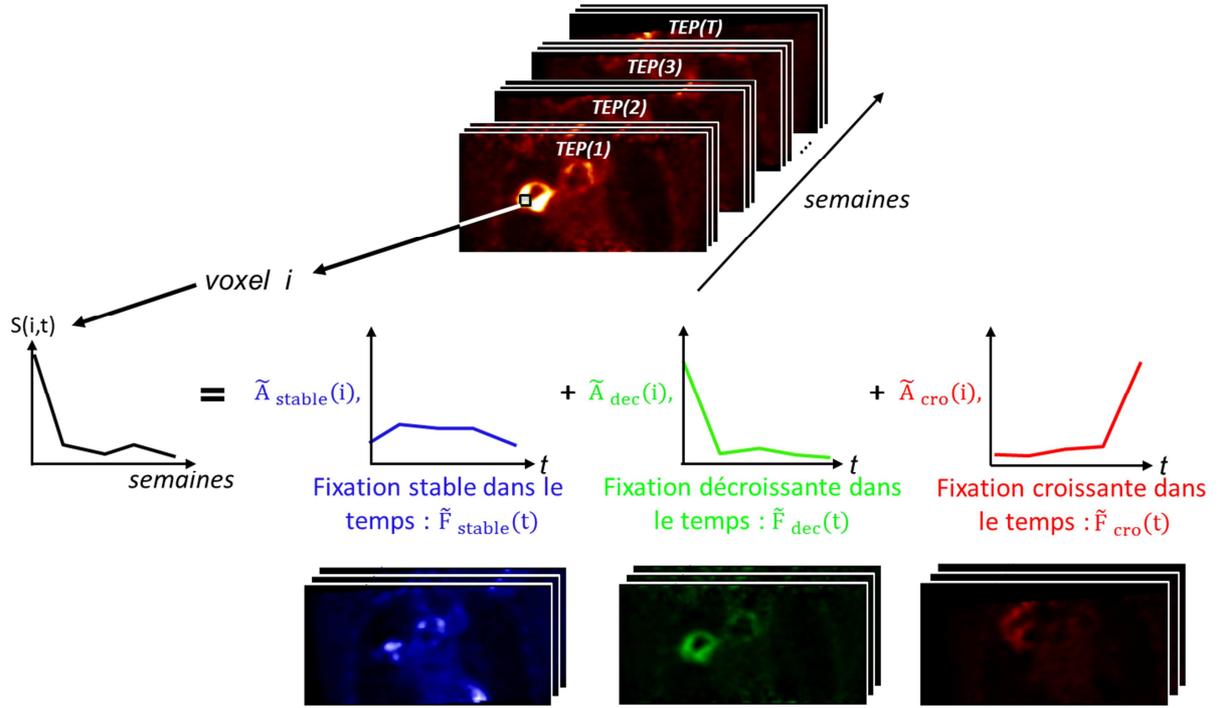


Figure VI.1 : illustration du principe de l'analyse factorielle appliquée à une série de T examens TEP acquis durant la radiothérapie d'un patient atteint de NSCLC. On montre ici la décomposition, sous la forme d'une somme linéaire de 3 facteurs (3 courbes de couleurs) de l'activité d'un des N voxels inclus dans le volume étudié. Les images factorielles associées à ces facteurs (en dessous des courbes respectives) donnent pour chaque voxel i de l'image le poids associé au facteur correspondant.

VI.3. Résolution du modèle

Nous nous baserons sur les travaux de (Buvat et al., 1998) pour présenter le modèle et sa méthode de résolution, en nous plaçant dans le contexte du suivi thérapeutique.

Considérons une séquence de T images, chacune constituée de N voxels. La série d'images correspond alors à un ensemble de N vecteurs S_i (avec $i \in [1; N]$) possédant chacun T composantes, notées S_{it} (avec $t \in [1; T]$), en accord avec les notations précédentes. Chaque S_i est appelé *trixel* et représente la variation temporelle du signal contenu dans le voxel i.

En utilisant le formalisme unifié de (Benali et al., 1994), le modèle à résoudre par l'AFSIM est :

$$S_{it} = \frac{S_{..}}{K} \sum_{k=1}^K \tilde{A}_k(i) \times \tilde{F}_k(t) + \varepsilon_{it} \quad \text{Eq. VI.2}$$

Avec $S_{..}$, somme totale du signal tel que :

$$S_{..} = \sum_{i=1}^N \sum_{t=1}^T S_{it} \quad \text{Eq. VI.3}$$

Le modèle est résolu en 5 étapes :

- i. Le recalage des examens TEP statiques, c'est-à-dire la mise en correspondance spatiale des voxels pour suivre leur évolution dans le temps.
- ii. La sélection des N voxels à étudier.
- iii. Le filtrage du bruit (i.e. l'estimation de ε_{it}). Cette étape correspond à l'*analyse orthogonale* et conduit à un sous-espace d'étude.
- iv. L'identification des facteurs dans ce sous-espace d'étude, par un changement de base : c'est l'*analyse oblique*.
- v. Le calcul des facteurs de pondération $\tilde{A}_k(i)$ associés à chaque voxel, qui correspondent aux images factorielles.

VI.3.1. Etapes préliminaires : recalage et sélection des voxels

Les deux étapes préalables à la résolution du modèle présenté à l'équation VI.2 correspondent à la mise en correspondance des examens TEP statiques et à la sélection des voxels dont nous souhaitons décrire l'évolution au cours du traitement.

- Mise en correspondance des examens TEP

Le modèle d'analyse factorielle que nous avons présenté suppose qu'il est possible de suivre l'évolution du signal de chaque voxel dans le temps. Autrement dit, les voxels dont on étudie les variations doivent correspondre à la même petite région anatomique d'un examen TEP au suivant. Il est donc nécessaire de mettre en correspondance les régions étudiées dans les séries d'examens TEP, c'est-à-dire de les recalcr. Le principe de la démarche que nous avons adoptée ici est présenté sur la figure VI.2.

En considérant chacune des TDM($t>1$), fusionnées aux examens TEP(t) et utilisées pour la localisation anatomique, nous avons déterminé, par une méthode semi-automatique, les transformations qui permettaient de recalcr un volume d'intérêt, noté $\text{VOI}_{\text{recalé}}(t>1)$, sur la même zone anatomique, $\text{VOI}_{\text{recalé}}$, définie sur la TDM initiale (TDM(1)). La transformation utilisée, qui est obtenue grâce à une méthode de recalage rigide, vise à maximiser un critère de similitude entre les informations anatomiques contenues dans les volumes à recalcr. Elle nécessite de déterminer 6 paramètres, ou degrés de liberté, constitués de 3 rotations et de 3 translations. Nous avons utilisé la méthode de *block matching* (Ourselin et al., 2000) implémentée dans le logiciel *Isogray* (Dosisoft) qui effectue une optimisation multi-échelle et une interpolation trilineaire. Ce choix a été validé dans les travaux de (Necib et al., 2011). Le $\text{VOI}_{\text{recalé}}$ était toujours centré sur la lésion étudiée ou le groupe de lésions étudiées quand celles-ci étaient situées dans la même zone anatomique et à proximité (à ~ 5 cm au maximum) les unes des autres. Si les lésions étudiées étaient trop éloignées les unes des autres, plusieurs recalages indépendants ont été effectués, un pour chaque lésion. Le $\text{VOI}_{\text{recalé}}$ incluait de plus

les repères osseux les plus proches (e.g : côtes et/ou vertèbres) et se réduisait, quand cela était possible, à une partie du poumon pour que l'hypothèse d'un recalage rigide reste réaliste.

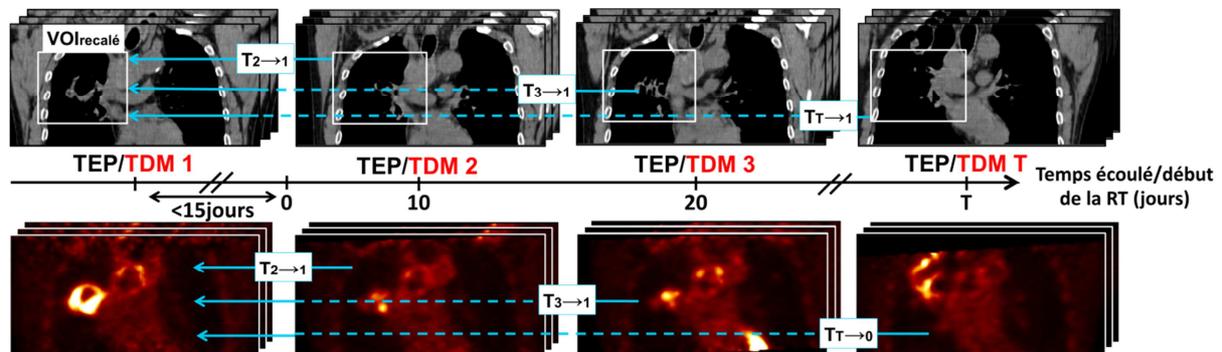


Figure VI.2 : méthode utilisée pour recalibrer les données TEP entre elles. On délimite, sur chaque examen TDM de la série, la zone à recalibrer, contenant la lésion étudiée, des tissus sains avoisinants et les repères osseux caractéristiques les plus proches (ici, les côtes par exemple). A partir d'un recalibrage rigide, on obtient la transformation $T_{i \rightarrow 1}$ à appliquer à chaque TDM(i) pour mettre en correspondance la zone étudiée sur la TDM(1) (première ligne de la figure). Une fois ces transformations $T_{i \rightarrow 1}$ déterminées, on les applique à chacune des TEP(i) associées afin de recalibrer la même zone mais entre les TEP(i) et la TEP(1) (ligne 2).

Une fois déterminées les transformations T , elles ont été automatiquement appliquées aux examens TEP correspondants (figure VI.2, ligne 2). Dans certains cas, un ajustement manuel d'une ou plusieurs TEP de la série a été effectué afin de corriger des mouvements anatomiques locaux des lésions. Ces derniers peuvent en effet ne pas être précisément retranscrits par le recalibrage anatomique, notamment quand les repères osseux utilisés pour le recalibrage des données TDM sont trop éloignés de la lésion.

- Sélection des voxels étudiés par SULTAN

Après l'étape de recalibrage, il est nécessaire de sélectionner les voxels à considérer pour l'analyse factorielle. Nous avons pour cela défini un VOI_{SULTAN} manuellement pour chaque lésion étudiée. Ce dernier est automatiquement reproduit sur l'ensemble des examens TEP recalibrés. Chaque VOI_{SULTAN} contenait la lésion d'intérêt ainsi que des tissus sains voisins, dont la fixation en FDG était supposée ne pas changer au cours du temps (ce qui exclut par exemple l'œsophage). Au final, les VOI obtenus contenaient 2 à 3 fois plus de tissus sains que de volume tumoral.

En dehors de l'étape de recalibrage sur données anatomiques, pour lesquelles nous avons utilisé le logiciel *Isogray* (*Dosisoft*), l'ensemble des résultats de SULTAN que nous présenterons a été obtenu à l'aide du logiciel *Pixies* (*Apteryx*, version 1.11).

VI.3.2. Analyse orthogonale

La résolution du modèle linéaire décrit par l'équation VI.2 passe par celle d'un modèle d'analyse de la covariance, dit *modèle à effet fixe* (Causinus, 1986), afin d'estimer la composante non bruitée \tilde{S}_{it} tel que :

$$S_{it} = \tilde{S}_{it} + \varepsilon_{it} \quad \text{Eq. VI.4}$$

Le modèle à effet fixe suppose que le signal observé S_{it} est la somme d'une composante fixe (i.e. sans bruit) et d'une erreur aléatoire ε_{it} . En résolvant ce modèle, on obtient un sous-espace \mathbb{S} de \mathbb{R}^T , appelé *sous-espace d'étude*, de dimension $Q < T$, tel que la projection de tout S_{it} sur ce sous-espace donne sa composante non bruitée \tilde{S}_{it} .

L'espace d'étude \mathbb{S} est obtenu en effectuant une décomposition orthogonale utilisant une métrique appropriée. Cette métrique peut être estimée ou déduite des données initiales. Dans le cas d'images TEP reconstruites par la méthode OS-EM, le niveau de bruit dans les images varie approximativement de façon monotone avec le signal (Buvat, 2002). Il est donc préférable d'utiliser la métrique du χ^2 , optimale dans le cas où le bruit suit une distribution de Poisson (Benali et al., 1993). La décomposition orthogonale correspondante est alors celle de l'Analyse Factorielle des Correspondances (AFC). Cette AFC décrit, dans un espace de dimension minimale, la partie de l'information initiale la plus caractéristique. Les trixels non bruités peuvent alors être écrits sous la forme :

$$\tilde{S}_{it} = \frac{S_{i.}S_{.t}}{S_{..}} + \sum_{q=1}^{K-1} \frac{S_{i.}S_{.t}}{S_{..}} \frac{1}{\sqrt{\lambda_p}} \varphi_q(t) \psi_q(i) \quad \text{Eq. VI.5}$$

En notant

$$S_{i.} = \sum_{t=1}^T S_{it} \quad \text{Eq. VI.6}$$

Et

$$S_{.t} = \sum_{i=1}^N S_{it} \quad \text{Eq. VI.7}$$

Les valeurs propres λ_p , vecteurs propres $\varphi_q(j)$ et images propres $\psi_q(i)$ résultent de la décomposition orthogonale. A l'issue de l'analyse orthogonale, on obtient donc une base de $K-1$ vecteurs propres $\{\varphi_q(t)\}_{t=1..T}$ qui engendrent l'espace \mathbb{S} contenant les trixels non bruités.

VI.3.3. Analyse oblique

L'étape suivante, appelée analyse oblique, a pour but d'estimer les K facteurs \tilde{F}_k et les images factorielles associées (i.e. les K \tilde{A}_k) L'hypothèse est que ces facteurs et images appartiennent à l'espace d'étude \mathbb{S} issu de l'analyse orthogonale. Pour cette estimation, nous avons résolu le modèle dit *modèle courbes* (Benali et al., 1994), qui fait l'hypothèse que les facteurs peuvent être écrits sous la forme :

$$\tilde{F}_k(t) = \frac{S_{.t}}{S_{..}} + \sum_{q=1}^{K-1} \beta_{kq} \varphi_q(t) \quad \text{Eq. VI.8}$$

Où les β_{kq} correspondent aux coordonnées de $\tilde{F}_k(t)$ dans \mathbb{S} . En utilisant ce modèle, les images factorielles sont données par :

$$\tilde{A}_k(i) = \frac{S_{i.}}{S_{..}} + \sum_{q=1}^{K-1} \frac{1}{S_{..}} \gamma_{kq} \psi_q(i) \quad \text{Eq. VI.9}$$

Où les γ_{kq} sont les coordonnées des $\tilde{A}_k(i)$ dans \mathbb{S} . Dans ce modèle, les coordonnées des facteurs et des images dans \mathbb{S} sont liés par la relation matricielle suivante :

$$\beta \gamma^t = K \mathbf{1}_{K \times K} \quad \text{Eq. VI.10}$$

Avec β (resp. γ) matrice de dimensions $K \times K$ de terme β_{kq} (resp. γ_{kq}) et dont la première colonne est constituée de 1. $\mathbf{1}_{K \times K}$ correspond à la matrice identité de dimensions $K \times K$ et γ^t transposée de γ .

La résolution de ce modèle se fait de manière itérative en utilisant des contraintes de non négativité, de normalisation et des connaissances a priori. En bref, la procédure itérative peut être décrite de la manière suivante : à partir d'un premier jeu de facteurs $\tilde{F}_k(t)$, on déduit les images associées $\tilde{A}_k(i)$. Ces images sont alors modifiées pour satisfaire les différentes contraintes. En utilisant ces nouvelles images, on déduit alors les $\tilde{F}_k(t)$ correspondants. Ces nouveaux facteurs sont à leur tour modifiés pour satisfaire les différentes contraintes. Ces modifications permettent de recalculer les images associées aux facteurs. Le processus se répète jusqu'à ce qu'un critère d'arrêt soit vérifié. Ce critère peut porter sur le nombre d'itérations, sur le nombre maximal de valeurs négatives acceptables ou sur la différence maximale acceptable entre deux estimations consécutives des facteurs ou des images associées.

La procédure itérative est plus précisément la suivante :

- i. On initialise les coordonnées β_{kq} des facteurs $\tilde{F}_k(t)$. On impose $\beta_{k0}=1$ pour tous les facteurs.
- ii. On calcule les γ_{kq} correspondant grâce à l'équation VI.10.

- iii. On normalise les γ_{kq} de manière à vérifier les équations relatives au modèle dual du modèle courbes (i.e. *le modèle images* cf (Benali et al., 1994)).
- iv. On calcule les images factorielles $\tilde{A}_k(i)$ correspondant aux γ_{kq} à partir de l'équation VI.9 puis on leur applique des contraintes de non négativité. Cette étape revient à annuler tout $\tilde{A}_k(i)$ dont la valeur est négative. Toujours en cherchant à vérifier les équations du modèle images (cf étape iii.), on déduit puis normalise les nouveaux γ_{kq} .
- v. En utilisant à nouveau l'équation VI.10, on calcule les coordonnées β_{kq} des facteurs correspondant aux nouvelles coordonnées γ_{kq} des images factorielles obtenues à l'étape iv.
- vi. Les facteurs $\tilde{F}_k(t)$ sont déduits des γ_{kq} via l'équation VI.8 puis les contraintes de non-négativité sont appliquées sur les facteurs. Si $\tilde{F}_k(t) < 0$, alors il est remplacé par 0. En plus de ces contraintes, notre application de l'AFSIM au suivi thérapeutique ajoute une contrainte supplémentaire portant sur le fait que l'un des facteurs soit constant. Cette contrainte est imposée de la même manière que les contraintes de positivité. On rend alors constant le facteur le plus constant, en remplaçant chaque valeur de la courbe par la valeur moyenne de la courbe avant contrainte, puis on normalise. Les β_{kq} sont ensuite recalculés à partir des nouveaux $\tilde{F}_k(t)$ et normalisés tels que $\beta_{k0} = 1$.
- vii. Retour à l'étape 2 sauf si le nombre d'itérations a permis d'atteindre le critère d'arrêt imposé.

Les contraintes imposées a priori peuvent ne pas être entièrement satisfaites du fait :

- De leur incompatibilité,
- De l'absence de courbes dans l'espace d'étude \mathbb{S} qui permettent de les satisfaire.

L'algorithme présenté ci-dessus peut être sujet à des instabilités : la modification des fonctions annule l'effet de la modification des compartiments, et inversement. Pour résoudre ce problème, il est possible de relaxer les contraintes (Buvat et al., 1998). Cette opération revient à ce que les fonctions et les compartiments ne soient modifiés que s'ils s'éloignent fortement des contraintes, en prenant en compte les écart-types associés aux $A_k(i)$ et $F_k(j)$. Elle permet de stabiliser la convergence en rendant les contraintes compatibles. Cette opération requiert le choix d'un facteur de relaxation qui est une mesure, en nombre d'écart-type, de la quantité dont un point d'une fonction ou d'un voxel d'un compartiment est autorisé à fluctuer sans être modifié. Quand nécessaire, le facteur de relaxation utilisé a été fixé de manière globale, c'est-à-dire identique pour chaque contrainte. Sa valeur était alors gardée la plus faible possible afin de permettre la convergence de l'algorithme vers une solution stable et satisfaisant au mieux les contraintes.

Pour réaliser l'analyse, il est nécessaire de choisir le nombre K de facteurs. Nous avons considéré dans notre étude que 2 ou 3 facteurs étaient suffisants lorsque nous avons $T \geq 3$ (soit au moins 3 examens). Ce choix permet de représenter les trixels n'évoluant pas (facteur

constant), ceux diminuant et/ou ceux augmentant (cf figure VI.1). Le choix entre deux ou trois facteurs dépendait de la convergence de l'algorithme ainsi que du respect des contraintes (et notamment de la positivité des $\tilde{A}_k(i)$) dans le résultat obtenu après convergence. Si trois facteurs étaient choisis initialement mais que l'algorithme ne convergait pas ou conduisait à deux facteurs ayant la même tendance, alors l'analyse était recommencée avec deux facteurs.

VI.3.4. Projection

Une fois les facteurs $\tilde{F}_k(t)$ obtenus, les images factorielles associées sont calculées sur l'ensemble de l'image, et non pas seulement sur la région étudiée (i.e. sur la région contenant les N trixels d'intérêt). Pour ce faire, on projette chaque trixel bruité sur la base des facteurs obtenue. Pour cette projection, il est possible de renormaliser les $\tilde{F}_k(t)$ ou les $\tilde{A}_k(i)$ selon la norme L1 (la somme des valeurs absolues), L2 (la racine carré de la somme des carrés des valeurs), la somme des valeurs positives ou la moyenne.

Nous avons ici sélectionné une normalisation des facteurs sur la valeur moyenne :

$$\frac{\sum_{t=1}^T \tilde{F}_k(t)}{T} = 1 \quad \text{Eq. VI.11}$$

Cette normalisation permet de représenter approximativement le facteur constant par une courbe constante de valeur $\tilde{F}_{constant}(t) = 1 \forall t$.

VI.4. Matériel et méthode

VI.4.1. Patients et protocole

Pour évaluer l'intérêt de l'analyse factorielle dans le contexte du suivi thérapeutique, nous avons étudié les données TEP de 10 patients atteints de NSCLC de grades II à IV. Ces patients ont tous été traités par RT) ou RT-CT au centre hospitalo-universitaire Henri Becquerel de Rouen. Les caractéristiques de cette cohorte sont données dans le tableau VI.1. Le protocole de recherche pour lequel ces patients, dûment informés, étaient enrôlés, a été approuvé par un comité d'éthique. L'étude, close en Janvier 2009, dont le but initial était *d'étudier la modification de l'activité métabolique tumorale mesurée par TEP-FDG, tout au long d'une radiothérapie, chez des patients ayant un cancer du poumon*, a été enregistrée sur e-cancer.fr sous l'identifiant RECF0645 (RTEP1).

Caractéristiques	N
Nombre total de patients	10
Age (en années)	
• Moyenne (écart-type)	61 (8)
• Intervalle	40-82
Genre	
• Homme	9
• Femme	1
Traitements	
• RT	4
• RT+CT	6
Types histologiques	
• Carcinomes épidermoïdes	6
• Adénocarcinomes	2
• Cellules indifférenciées	2

Tableau VI.1 : caractéristiques des patients

L'inclusion des patients impliquait la vérification de certains critères :

- i. Que l'histologie ait permis de démontrer la présence d'un NSCLC invasif.
- ii. Qu'une tumeur montre une fixation en FDG significative (par rapport au fond) lors de l'inclusion du patient. Si l'acquisition d'un examen TEP n'était pas automatique durant le diagnostic, la mesure d'une activité métabolique significative (i.e. au moins deux fois supérieure à celle du fond) sur la FDG-TEP devait être effectuée au plus tôt 21 jours avant la première séance de RT.
- iii. D'avoir au moins une tumeur évaluable par les critères RECIST.
- iv. Que le patient soit éligible pour un traitement par RT à visée curative, et donc que la tumeur soit localisée et non métastatique.
- v. Etre âgé de plus de 18 ans.
- vi. Que le patient ait donné son consentement éclairé par écrit.

Enfin, les critères d'exclusions portaient sur le fait que le/la patient(e):

- i. Souffre d'un diabète mal régulée (glycémie >10 mmol).
- ii. Soit en enceinte ou allaite.
- iii. Soit déjà inclus dans un autre protocole de recherche mettant en œuvre une molécule expérimentale.

Les détails précis des différents critères présentés ici sont fournis dans l'article de (Edet-Sanson et al., 2012), qui portait sur la même cohorte.

VI.4.1.1. Protocole de traitement

Les doses prescrites à l'isocentre du volume cible prévisionnel (ou *Planning Target Volume*) étaient de 66 Gy pour les RT-CT et de 70 Gy pour les RT seules.

Le fractionnement appliqué correspondait à une dose de 2 Gy/fraction et de 5 fractions/semaine, délivrées par un accélérateur de 18MV.

Les contraintes de dose appliquées étaient :

- Une dose au PTV comprise entre 95 et 107% de la dose prescrite,
- Un maximum de 30% du volume total de poumon recevant au maximum 20%,
- Une dose maximale à la moelle épinière de 46 Gy.

Les plans de traitements ont été réalisés à l'aide d'une station Eclipse (*Varian Medical System*).

VI.4.1.2. Données TEP/TDM

Outre les critères précisés au paragraphe VI.4.1, l'inclusion définitive des patients dans le protocole reposait sur la présence d'une activité significative dans une FDG-TEP effectuée dans les 15 jours précédents la première fraction d'irradiation (PET(1)). Après le début de la RT/RT-CT, 4 à 5 examens FDG-TEP ont été acquis tous les 14 à 16 Gy (i.e. 1 examen tous les 10 jours), c'est-à-dire toutes les 7 fractions et ce jusqu'à 70 Gy. Tous les examens TEP ont été acquis sur un tomographe *Biograph LSO Sensation 16* (*Siemens Medical Solutions*, Hoffman Estates, Knoxville, TN, USA).

Pour diminuer l'irradiation du patient due à l'imagerie TEP, une dose réduite de FDG (3,5MBq/kg) a été injectée après 30 minutes de repos. Le début des examens a commencé en moyenne 60 minutes (± 5 minutes) par une TDM dans la direction cranio-caudale. Les paramètres de cette TDM étaient : une tension d'alimentation de 120 kVp (pour *Peak voltage*, ou tension maximale appliquée au tube à rayons-X), une charge de 100-150 mAs fonction du poids du patient et optimisée par un logiciel de réduction de dose (*Caredose*, *Siemens Medical Solution*).

L'examen TEP réalisé avant le traitement (appelé TEP prétraitement ci-après) incluait deux acquisitions distinctes.

- La première, faisant suite à l'acquisition de la TDM, correspondait à un examen avec synchronisation (ou *gating*) respiratoire, utilisant un système *Anzai AZ-733V* (*Anzai Medical*, Tokyo, Japon), de 15 minutes, centré sur le thorax. Cette technique a pour objectif de réduire l'effet du mouvement sur les images TEP et notamment de réduire la sous-estimation de l'activité métabolique mesurée due au flou cinétique.
- Suite à cet examen, une autre acquisition sans *gating* a été effectuée de la mi-cuisse à la base du crâne.

Pour tous les examens qui ont suivi (i.e. pour toutes les TEP acquises durant la RT), seule la région du thorax (des clavicules à la partie supérieure du diaphragme) a été scannée (avec et sans *gating* respiratoire). Il est important de noter ici que seuls les examens non synchronisés

à la respiration ont été utilisés pour cette étude. Ces acquisitions statiques ont débuté en moyenne 86 minutes après l'injection du radiotracer (intervalle : [73-116 minutes]).

Les dimensions des images TDM utilisées pour la correction d'atténuation étaient de 512x512x33 voxels (dimensions d'un voxel : 0,98x0,98x5,0 mm³) alors que les examens TDM utilisés pour la localisation anatomique et la mise en correspondance entre les examens successifs présentaient une épaisseur de coupe de 3 mm, chacune espacée de 2 mm pour une dimensions des voxels en transaxial de 0,98x0,98 mm². Les examens TEP étaient reconstruits dans une matrice de dimensions 168x168x81 (dimensions d'un voxel : 4,06x4,06x2,0 mm³), au moyen d'un algorithme itératif *Fourier-Rebinning* (FORE) et *Attenuation-Weighted Ordered Subset Expectation Maximization* avec 4 itérations et 8 sous-ensembles (station *Leonardo*, *Siemens Medical Solution*). Ces examens étaient corrigés des coïncidences fortuites, du diffusé et de l'atténuation à partir des données TDM. Enfin, un filtre Gaussien de LMH de 5 mm a été appliqué après la reconstruction.

Pour pouvoir étudier l'évolution tumorale à partir des examens TEP, nous avons imposé que tous les examens TEP de suivi aient un délai entre injection et acquisition inférieur ou égal à 10 minutes par rapport au même temps pour la TEP prétraitement. Après l'exclusion de tous les examens ne respectant pas cette contrainte, neuf patients avec au moins quatre TEP/TDM ont été inclus dans notre étude. Les images TEP ont enfin toutes été converties en unités SUV normalisées par le poids (cf équation I.17).

VI.4.1.3. Analyse quantitative des examens TEP

Une première partie du travail d'analyse des données a consisté à mettre en œuvre la méthode d'analyse factorielle décrite au paragraphe VI.3 sur les données TEP des patients NSCLC.

VI.4.1.3.1. Méthodes de suivi classique

Par la suite, trois index SUV ont été calculés, pour chaque examen TEP et pour chaque lésion, à partir d'une station de recherche OWS (*Dosisoft*, version 1.0.0.43).

- SUV_{max} : le SUV du voxel d'activité la plus élevée dans la lésion étudiée. Quand l'activité de la lésion n'était pas discernable de celle du fond local, le SUV le plus élevé de la zone décrite comme tumorale sur la TEP(1) a été relevé.
- SUV_{adapt} : SUV moyen dans une région segmentée à partir d'un seuil adaptatif T_{adapt} prenant en compte le contraste entre l'activité dans la tumeur et celle du fond. Cette méthode, proposée par (Vauclin et al., 2009), vise à déterminer la valeur optimale de T_{adapt} par une approche itérative. Elle nécessite une fonction de calibration qui permet d'attribuer, à une valeur de contraste mesurée, celle du seuil à appliquer. Le principe de l'algorithme itératif est le suivant :

- i. une étape préliminaire permet de former un volume VOI_0 englobant la lésion et de déterminer une première valeur de seuil T_{h1} à appliquer aux voxels contenus dans VOI_0 .
- ii. L'algorithme itératif débute alors en considérant, dans VOI_0 , l'amas de voxels connexes dont l'activité est supérieur à T_{h1} et qui permet de former VOI_{TH1} .
- iii. A partir de VOI_{TH1} , on détermine une mesure du contraste locale $Cont_{meas}$.
- iv. A l'aide de la fonction de calibration, dérivée de mesure sur données fantômes et qui permet de définir les paramètres A et B de la formule VI.12, on déduit de $Cont_{meas}$ une nouvelle valeur de seuil, notée T_{h2} . Si T_{h1} est égale à T_{h2} , on arrête l'algorithme et T_{adapt} vaut T_{h2} , sinon on recommence à l'étape 2 en remplaçant T_{h1} par T_{h2} .

$$T_{h2} = A + B/C_{mes} \quad \text{Eq. VI.12}$$

Les valeurs des paramètres A et B utilisées dans cette étude, identiques à celles présentés par (Vauclin et al., 2009), ont été optimisées par l'auteur de cette technique pour le système d'imagerie utilisé pour acquérir les données patients décrites au paragraphe VI.4.1.2. La calibration de cet algorithme de segmentation s'est basée sur des acquisitions de données fantôme sphérique (lequel contenait 6 sphères de volume variant de 0,55 à 17,2 mL), selon 5 niveaux de contrastes sphères/fond.

- $SUV_{40\%}$: SUV moyen dans une zone obtenue par croissance de région et dont tous les voxels connexes ont une activité supérieure ou égale à 40% du SUV_{max} de la lésion.

A partir des deux algorithmes de segmentation utilisés pour calculer $SUV_{40\%}$ et SUV_{adapt} , nous avons aussi relevé systématiquement, quand les algorithmes utilisés avaient convergé, les volumes métaboliques correspondants (notés $VM_{40\%}$ et VM_{adapt} , respectivement).

En nous basant sur les valeurs des VM_{adapt} mesurés sur la TEP de référence et fournis par la méthode de segmentation la plus adaptée à notre cohorte, à savoir la méthode mettant en œuvre un seuil adaptatif, nous avons cherché à évaluer si le seul volume métabolique, mesuré sur la TEP initiale, permettait de séparer, avant toute analyse d'hétérogénéité, les ganglions des tumeurs. Nous avons pour cela testé si les volumes médians de ces deux groupes de lésions étaient significativement différents par un test non paramétrique de Wilcoxon.

VI.4.1.3.2. Comparaison des différentes approches

- Faisabilité des approches en fonction de l'examen TEP

Nous avons tout d'abord cherché à évaluer la capacité des cinq index considérés (SUV_{max} , SUV_{adapt} et VM_{adapt} d'une part et $SUV_{40\%}$ et $VM_{40\%}$ d'autre part) ainsi que de SULTAN à fournir un résultat non aberrant sur l'activité métabolique de la lésion pour chacun des TEP. Pour cela, nous avons calculé, pour chacun des 6 examens TEP (TEP prétraitement et jusqu'à 5 TEP de suivi), le nombre de fois où la méthode fournissait un résultat aberrant. Nous avons ainsi déduit le taux d'échec de chaque méthode en fonction de moment où les données sont acquises durant le traitement.

- Réponse des lésions à la thérapie vue par les index métaboliques

Les critères d'exclusion appliqués aux données n'ont pas permis d'inclure tous les examens acquis à un moment précis du traitement. Il n'a donc pas été possible d'appliquer un seuil commun pour caractériser l'évolution du SUV de la lésion entre la TEP initiale (ou TEP de référence) et la dernière TEP prise en compte (ou TEP finale, mesurée au temps t_{final}) dans la mesure où la dernière TEP considérée n'était pas nécessairement acquise à un moment identique du traitement pour chacun des patient. Nous avons cependant calculé, pour chaque patient et chacun des cinq index étudiés, l'évolution relative entre le TEP prétraitement et le dernier TEP, par :

$$\Delta Index(\%) = \frac{Index(t_{final}) - Index(t_{référence})}{Index(t_{référence})} \times 100 \quad \text{Eq. VI.13}$$

- Résultats de SULTAN

Les facteurs obtenus par SULTAN décrivent les principales évolutions suivies par les voxels du volume étudié. Les images associées permettent de connaître la proportion de chaque facteur dans le volume. Pour mesurer l'hétérogénéité de la réponse à la radiothérapie dans chaque lésion, nous avons calculé la moyenne, sur i , des $\tilde{A}_k(i)$ (notés $\overline{\tilde{A}(i)}$) ainsi que la proportion du facteur décroissant dans le volume métabolique segmenté sur la TEP prétraitement. Pour cela, nous avons considéré le VM obtenu par la méthode adaptative décrite au paragraphe VI.4.1.3.1. L'index ainsi calculé, que nous noterons $R_{décroissant}$, est donné par :

$$R_{\text{décroissant}} = \frac{\overline{\tilde{A}_{\text{décroissant}}}}{\sum_{k=1}^K \overline{\tilde{A}_k}} \quad \text{Eq. VI.14}$$

avec $\overline{\tilde{A}_k}$ tel que :

$$\overline{\tilde{A}_k} = \sum_j \frac{\tilde{A}_k(j)}{M} \quad \forall j \in VM_{\text{adapt}}(t_{\text{référence}}) \quad \text{Eq. VI.15}$$

où $t_{\text{référence}}$ correspond au moment de l'acquisition de la première TEP utilisée pour l'analyse factorielle et M au nombre de voxel contenus dans VM_{adapt} . $R_{\text{décroissant}}$ rend compte de la proportion du signal de la tumeur principalement décrit par le facteur décroissant. Autrement dit, une tumeur n'évoluant pas doit avoir un $R_{\text{décroissant}}$ proche de 0. De même, une tumeur principalement décrite par le facteur décroissant aura une valeur de $R_{\text{décroissant}}$ proche de 1. Dans le cas où deux facteurs sont obtenus et où $R_{\text{décroissant}}$ est proche de 0,5, SULTAN permet de mettre en évidence une réponse hétérogène : l'évolution de la tumeur est alors en moyenne bien décrite par un mélange équilibré des différents facteurs.

- Comparaison des résultats fournis par SUV_{max} et SUV_{adapt} à ceux fournis par SULTAN

Nous avons enfin étudié les liens entre les évolutions du SUV_{max} tout au long du traitement pour chacune des lésions, du SUV_{adapt} , quand cela était possible, ainsi que ceux du facteur décroissant obtenu en utilisant SULTAN. Pour pouvoir comparer les évolutions de $SUV_{\text{max}}(t)$, $\tilde{F}_{\text{décroissant}}(t)$ et SUV_{adapt} , nous avons choisi de normaliser les valeurs des points de la courbe du facteur décroissant de manière à ce que :

$$\tilde{F}_{\text{décroissant}}(t_{\text{référence}}) = SUV_{\text{max}}(t_{\text{référence}}) \quad \text{Eq. VI.16}$$

VI.5. Résultats

VI.5.1. Lésions étudiées

La liste des lésions étudiées chez les patients ainsi que les détails des examens analysés pour chaque patient (i.e. pour lesquels une différence maximum de 10 minutes entre le temps post-injection des différents examens était respectée) est présentée dans le tableau VI.2. Pour le patient 1, l'examen de référence considéré correspond au premier examen acquis durant le traitement et non pas à l'examen prétraitement. Au total, nous avons étudié l'évolution de 16 lésions, incluant 5 tumeurs et 11 ganglions. Pour les lésions 13 et 15 du patient 9, l'examen 3 a été exclu de l'étude à cause d'un décalage du champ de vue dans la direction cranio-caudale ne permettant pas de suivre l'évolution de tout le volume de ces 2 lésions.

Patient N°	Sexe	Age au premier examen	Lésion N°	Type de lésion	Examens TEP considérés
1	M	52	1	G	2-3-4-5--
			2	T	
2	F	38	3	G	1-2-4-5-6
			4	G	
3	M	74	5	T	1-2-3-4-5-6
4	M	68	6	G	1-2-3-4-5-6
5	M	55	7	T	1-2-3-4-5
6	M	59	8	G	1-2-3-4
7	M	82	9	T	1-2-5-6
			10	G	
			11	G	
8	M	76	12	G	1-2-3-4-5
9	M	51	13	T	1-2-3-5
			14	G	1-2-3-4-5
			15	G	1-2-3-5
			16	G	1-2-3-4-5

Tableau VI.2 : informations relatives aux données considérées pour l'analyse de données. G = ganglion. T = tumeur.

A partir de la segmentation des volumes métaboliques effectués sur les TEP initiales par la méthode adaptative, nous avons observé que les lésions avaient un volume moyen (± 1 écart-type) de $11,3 \text{ mL} \pm 13,9$ (médiane = $4,2 \text{ mL}$, intervalle compris entre $1,0$ et $44,5 \text{ mL}$). Les 5 tumeurs étudiées, de volume médian égal à $29,0 \text{ mL}$, étaient significativement plus volumineuses que les 11 ganglions (volume médian égal à $1,9 \text{ mL}$) (test de Wilcoxon, $p=0,0014$).

VI.5.2. Index métaboliques

Le taux d'échec des différentes méthodes, c'est-à-dire leur incapacité à fournir un résultat en fonction du moment auquel le TEP est acquis durant le traitement, est présenté en figure VI.3. Les méthodes de segmentation basées sur l'utilisation d'un seuil n'ont pas convergé dans 88% des cas (14/16 examens) pour $T_{40\%}$ et dans 56% des cas pour T_{adapt} (9/16 examens) quand elles ont été appliquées sur la dernière image de la série. A contrario, SUV_{max} a fourni une mesure quel que soit la fixation de FDG dans la lésion (i.e. dans les 76 volumes TEP considérés). En considérant l'ensemble des données traitées, $T_{40\%}$ et T_{adapt} n'ont pas permis de caractériser les lésions dans 31/76 (41%) et 15/76 (20%) des cas, respectivement. Le suivi complet n'a pu être obtenu que pour 1 lésion sur 16 en utilisant $T_{40\%}$ contre 7 lésions sur 16 en utilisant T_{adapt} . Ces deux méthodes de segmentation échouent d'autant plus que l'examen est acquis tardivement après le début de traitement. Pour une dose de plus de $\sim 42 \text{ Gy}$, ces algorithmes ne permettent pas de segmenter les lésions dans plus de 40% des cas.

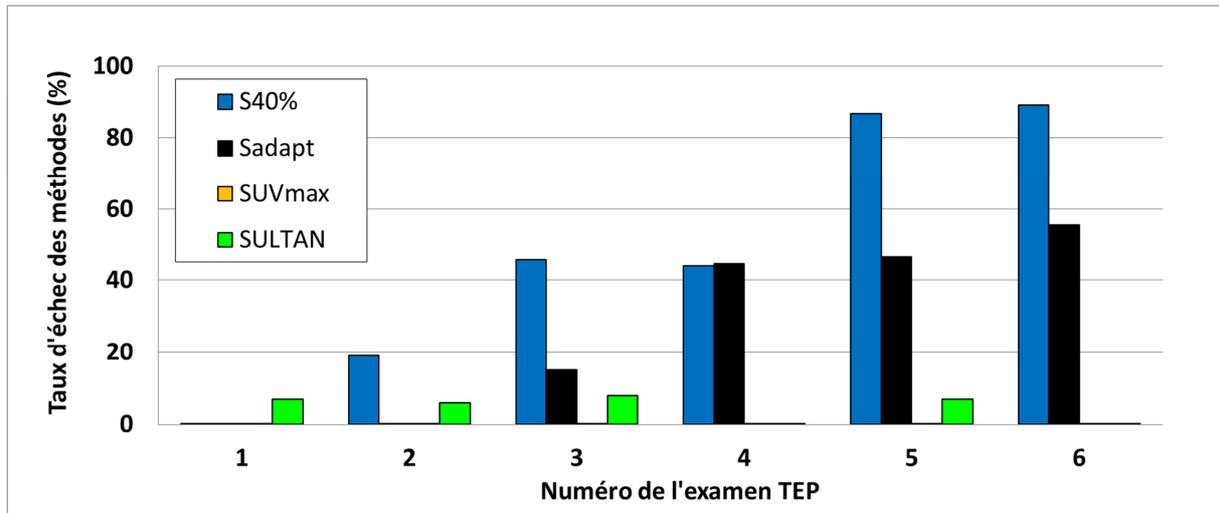


Figure VI.3 : taux d'échec des différentes méthodes en fonction du numéro de l'examen TEP/TDM de la série (pour un total de 76 volumes d'images TEP considérées).

Les valeurs sur la $TEP(t_{référence})$ ainsi que l'évolution relative entre les $TEP(t_{référence})$ et $TEP(t_{finale})$ considérées pour chaque lésion (tableau VI.2, colonne 6) obtenues pour les 5 index métaboliques sont données dans le tableau VI.3. Les échecs des méthodes de segmentation, quand elles étaient appliquées sur le dernier examen, ont aussi été rapportés. Ces échecs s'expliquent par la faible activité résiduelle dans les lésions à la fin du traitement. Le tableau VI.3 montre une décroissance du SUV_{max} de toutes les lésions entre les examens TEP initial et final supérieure à 25%.

Compte tenu des meilleures performances de l'algorithme de segmentation utilisant le seuil adaptatif par rapport à celui utilisant un seuil de 40% (tableau VI.3), nous avons utilisé les volumes VM_{adapt} pour caractériser les lésions sur les TEP de référence (colonne 5).

N° Lésion	VM40%			VMadapt			SUVmax			SUV40%			SUVadapt		
	tref (mL)	tfin (mL)	$\Delta VM(\%)$	tref (mL)	tfin (mL)	$\Delta VM(\%)$	tref (g/mL)	tfin (g/mL)	$\Delta SUV(\%)$	tref (g/mL)	tfin (g/mL)	$\Delta SUV(\%)$	tref (g/mL)	tfin (g/mL)	$\Delta SUV(\%)$
1	14,9	-	-	15,3	6,4	-58	10,5	6,2	-41	6,3	-	-	6,3	3,6	-42
2	8,6	-	-	8,7	-	-	9,7	3,6	-63	5,6	-	-	5,6	-	-
3	3,1	-	-	3,1	-	-	10,7	2,3	-79	6,0	-	-	6	-	-
4	4,0	-	-	1,0	-	-	3,4	1,8	-46	1,7	-	-	2,3	-	-
5	41,0	18,1	-56	44,5	16,7	-62	13,6	6,5	-52	9,0	3,6	-60	8,7	3,6	-59
6	5,7	-	-	6,0	-	-	7,9	2,4	-70	5,0	-	-	4,9	-	-
7	35,5	-	-	37,0	-	-	7,7	5,6	-27	4,5	-	-	4,4	-	-
8	7,7	8,2	7	5,3	4,1	-23	7,9	6,0	-25	4,7	3,3	-30	5,3	4	-24
9	28,0	-	-	29,0	21,6	-26	14,9	3,5	-77	9,0	-	-	8,9	2,2	-75
10	1,9	-	-	1,6	2,3	44	18,2	3,5	-81	11,0	-	-	11,6	2,6	-78
11	2,4	-	-	1,9	-	-	13,9	2,4	-83	8,6	-	-	7,9	-	-
12	1,6	-	-	1,6	-	-	7,8	2,6	-67	4,5	-	-	4,5	-	-
13	6,5	-	-	2,9	2,3	-21	4,5	2,5	-44	2,5	-	-	3,1	2	-35
14	4,7	-	-	1,5	-	-	5,0	1,6	-68	2,7	-	-	3,7	-	-
15	16,0	-	-	15,3	4,4	-71	8,2	3,8	-54	4,7	-	-	4,7	2,6	-45
16	2,8	-	-	1,4	-	-	5,7	1,6	-72	3,2	-	-	3,9	-	-

Tableau VI.3 : valeurs des différents index mesurés sur la TEP de référence (t_{ref}) et sur la dernière TEP considérée (t_{fin}) pour les 16 lésions étudiées. La troisième colonne de chaque index correspond à l'évolution relative entre ces deux examens (cf équation VI.13). Un tiret signifie un échec de la méthode de segmentation à fournir une segmentation cohérente du volume métabolique de la lésion.

VI.5.3. Résultats de SULTAN

SULTAN a fourni des résultats pour 15/16 des lésions étudiées (figure VI.3) mais n'a pas permis de suivre l'évolution de la lésion 13. Dans ce cas particulier, on remarque (figure VI.4) que la lésion est placée à proximité d'une zone de scoliose (figure VI.4A) rendant difficile le repositionnement à l'identique du patient au fil des examens TEP. Plusieurs tentatives de recalages, visant :

- à n'inclure dans $VOI_{\text{recalé}}$ que la lésion 13,
- à utiliser différents repères osseux situés à proximité,
- à utiliser des recalages manuels ou semi-automatiques élastiques à partir des images TEP seules

n'ont pas permis d'aboutir à une analyse SULTAN pertinente. De plus, l'instabilité dans le temps de l'activité du fond local, et notamment de celle du poumon entourant la lésion (cf figures VI.4 D et E) peuvent expliquer la difficulté de satisfaire la contrainte portant sur l'existence d'un facteur constant.

Les résultats obtenus sur toutes les autres lésions ont systématiquement fourni au moins 2 facteurs : l'un étant décroissant, l'autre constant. L'analyse factorielle a montré que l'une des lésions (lésion 7) nécessitait l'utilisation d'un troisième facteur, croissant (cf figure VI.1), pour satisfaire les différentes contraintes imposées dans le modèle.

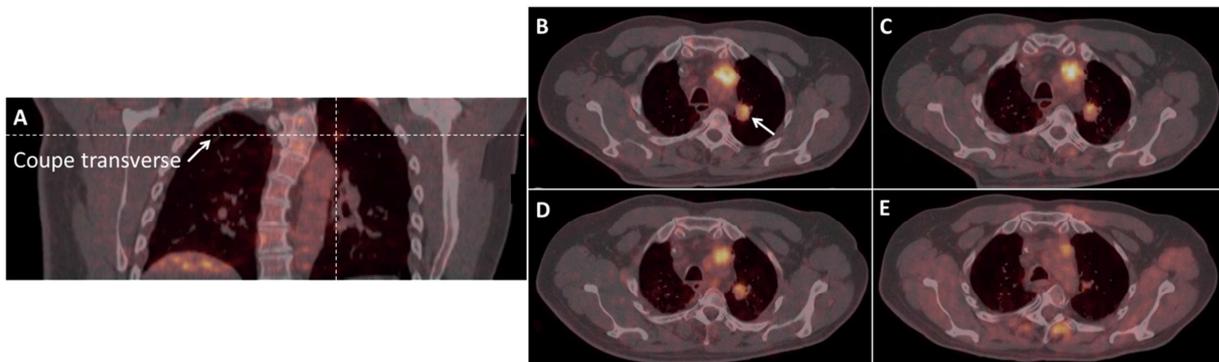


Figure VI.4 : A) coupe coronale de la TEP prétraitement du patient 9. En pointillés, les plans axiaux et triaxiaux se coupent au niveau de la lésion 13. B) coupe transverse de la TEP/TDM correspondant au plan pointé par une flèche blanche dans l'image A. Les images C-E) correspondent aux fusions des TEP/TDM des examens 2, 3 et 5 respectivement, dans le même plan de coupe (sans recalage ici). Images extraites de *OWS (Dosisoft)*.

Un ajustement manuel des volumes recalés a été nécessaire pour 8/60 (13%) des recalages effectués, et concernaient 5 des 16 lésions étudiées (10, 11, 12, 14 et 16). Ces ajustements correspondaient à des translations des données TEP déjà recalées à partir des données TDM, par un maximum de 2 voxels dans les 3 directions de l'espace.

Pour chaque lésion, nous avons calculé la valeur moyenne des $\tilde{A}_k(i)$ dans le volume VM_{adapt} obtenu sur la TEP de référence (tableau IV.4, colonnes 2 à 4) ainsi que les valeurs de

$R_{\text{décroissant}}$ correspondante pour chaque lésion. En dehors des lésions 7 et 8, toutes les lésions expriment plus fortement le facteur décroissant que le facteur constant. En particulier, la lésion 8 correspond à la lésion dont le $\Delta\text{SUV}_{\text{max}}$ était aussi le plus faible de la population de lésion considérée (tableau VI.3).

N° lésion	Valeurs des \bar{A}_k pour les différents facteurs obtenus (normalisées en SUV)			$R_{\text{décroissant}}$ (%)
	Constant	Décroissant	Croissant	
1	1,2	3,0	-	71
2	1,0	3,2	-	76
3	0,7	2,6	-	79
4	0,5	0,9	-	64
5	2,3	3,7	-	62
6	0,7	1,8	-	72
7	0,9	0,9	0,2	45
8	2,6	1,7	-	40
9	1,4	3,1	-	69
10	0,8	5,5	-	87
11	0,6	4,3	-	88
12	0,6	2,2	-	79
13	-	-	-	-
14	0,3	2,0	-	87
15	1,3	1,9	-	59
16	0,5	1,5	-	75

Tableau VI.4 : valeurs moyennes des $\bar{A}_k(j)$ dans le $\text{VM}_{\text{adapt}}(t_{\text{référence}})$ de chaque lésion (\bar{A}_k) pour chacun des 2 ou 3 facteurs obtenus (colonnes 2 à 4).

VI.5.4. Comparaison entre index métaboliques et résultats fournis par SULTAN

Les facteurs décroissants obtenus pour chaque lésion et normalisés par $\text{SUV}_{\text{max}}(t_{\text{référence}})$, ainsi que l'évolution de SUV_{max} durant le traitement, pour chacune des 15 lésions étudiées, sont donnés en figure VI.5. Le coefficient de détermination R^2 entre les variations de SUV_{max} et le facteur décroissant étaient supérieures à 0,95 pour 11/15 (73%) lésions et comprises entre 0,63 et 0,89 pour les 4 autres lésions. En dehors de ces 4 lésions, on constate donc un bon accord entre les évolutions de SUV_{max} et du facteur décroissant.

Pour les lésions 1, 3, 5, 7 et 8, pour lesquelles le facteur décroissant semble moins fortement lié à l'évolution du SUV_{max} dans le temps, nous avons cherché à expliquer les discordances en examinant l'ensemble des informations fournies par SULTAN. Nous présentons cette analyse détaillée.

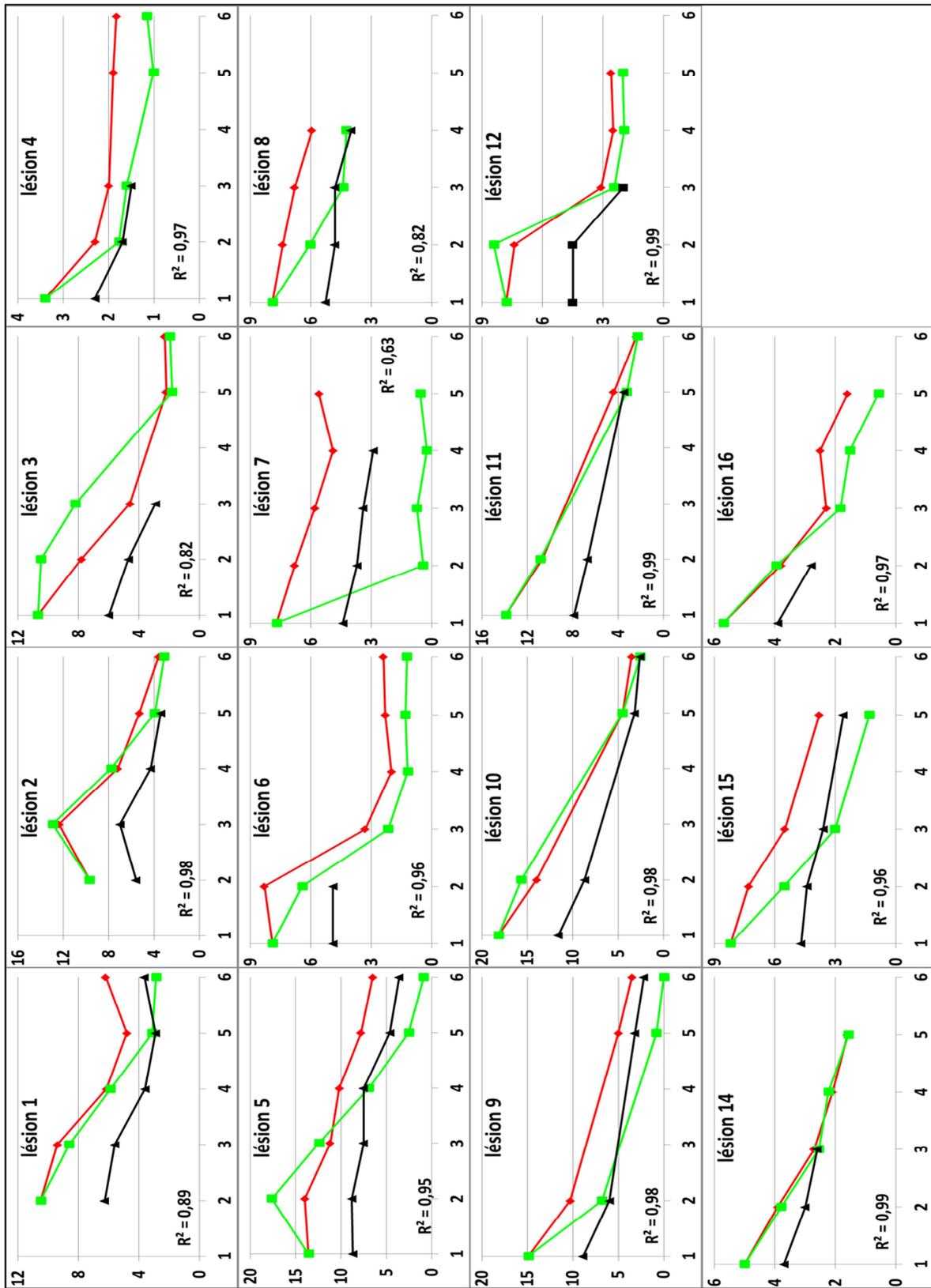


Figure VI.5 : Evolution du SUVmax (en rouge), du facteur décroissant issu de l'analyse factorielle et normalisé par la valeur de SUVmax(1) (en vert) et des valeurs successivement prises par le SUVadapt pour toutes les lésions étudiées. Le coefficient de détermination présenté pour chaque graphe est la valeur de R^2 entre les valeurs successives de SUVmax et celles du facteur décroissant.

- Lésion 1

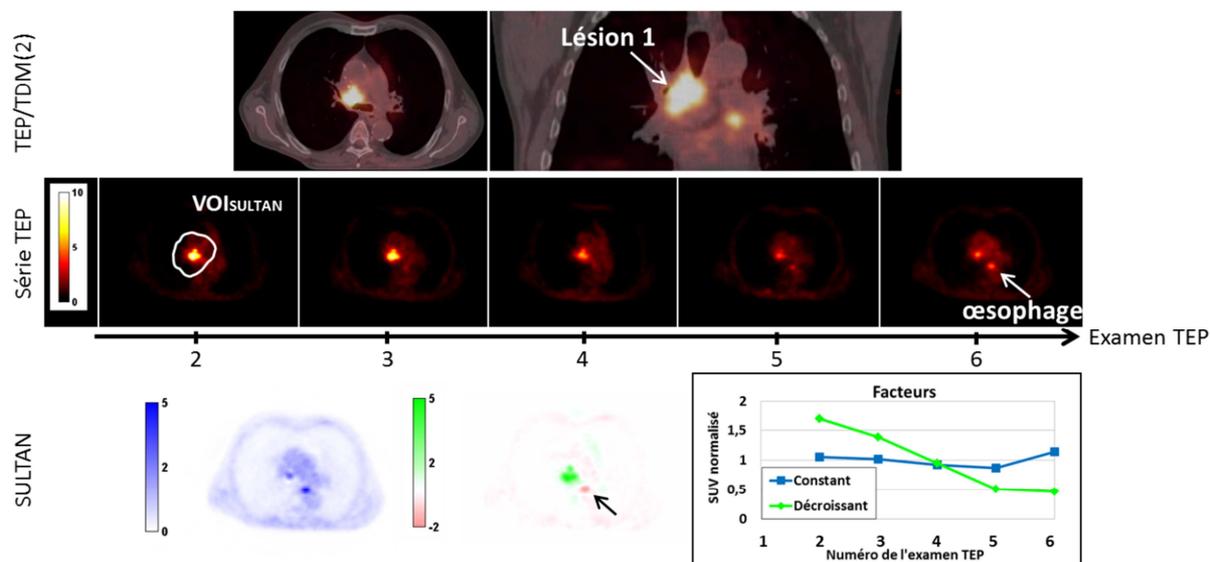


Figure VI.6 : 1^{ère} ligne : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (2 ici) montrant la lésion 1. 2^{ème} ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 1 et images factorielles et facteurs résultants (3^{ème} ligne). L'œsophage (flèche blanche sur la TEP(5)) ne fait pas partie de VOI_{SULTAN} analysée. Les valeurs négatives obtenues dans l'image factorielle associée au facteur décroissant (image 2, ligne 2, flèche noire) ne font donc pas partie des résultats interprétés de cette analyse.

La figure VI.5 met en évidence une décroissance plus rapide du facteur décroissant que celle du SUV_{max} , ce dernier augmentant même entre les deux derniers examens pris en compte pour l'analyse. En considérant le facteur constant (figure VI.6, courbes bleue en bas à droite), décroissant légèrement durant les 4 premiers examens et qui augmente lors du dernier examen, il est possible d'attribuer la croissance du SUV_{max} et du SUV_{adapt} observée sur la figure VI.5 à l'augmentation de l'activité de l'ensemble des tissus, sains ou non. Cette explication se justifie visuellement par l'augmentation de l'activité sur les deux coupes transverses des examens TEP finaux présentés à la figure VI.6 et sur lesquels une augmentation de l'activité médiastinale est observable entre les quatrième et cinquième examens. Ces résultats suggèrent que SULTAN permet de distinguer la contribution du signal commune à l'ensemble des tissus, et qui peut, par exemple, être liée à une réaction inflammatoire de tous les tissus présents dans le champ irradié, et celle plus propre à l'évolution de l'activité de la lésion.

Il est important de retenir que la valeur négative obtenue dans l'image factorielle associée au facteur décroissant (flèche noire, figure VI.6) n'est pas représentative de l'évolution de la fixation dans cette zone, correspondant anatomiquement à celle de l'œsophage (flèche blanche sur la TEP(5)). En effet, le VOI_{SULTAN} utilisé ici pour caractériser l'évolution de la tumeur (figure VI.6, TEP(1)) ne contenait pas de voxels de cet organe. Pour étudier l'évolution de cette zone indépendamment ou non de celle de la lésion, il est nécessaire de

l'inclure à la VOI_{SULTAN} , sans quoi la projection finalement obtenue ne saurait décrire convenablement cette région. Une étude complémentaire a montré qu'en incluant cette zone, un troisième facteur, croissant lors du dernier examen, permettait de décrire l'œsophagite visible sur 2 derniers examens de la série montrée à la figure VI.6.

- Lésion 3

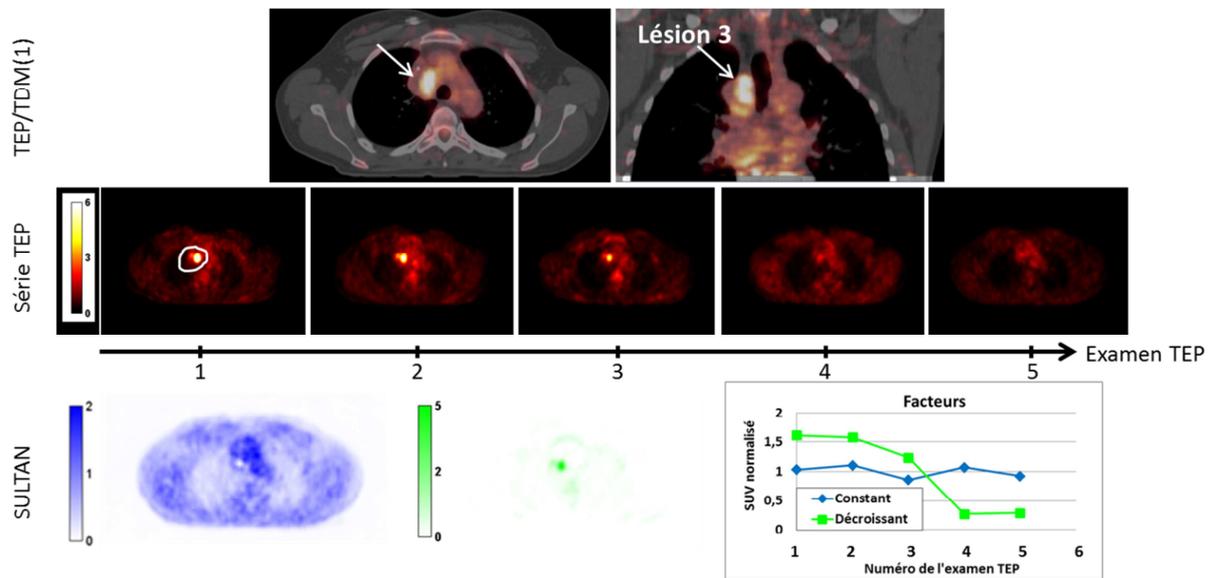


Figure VI.7 : 1^{ère} ligne : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (1 ici) montrant la lésion 3. 2^{ème} ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 3 et images factorielles et facteurs résultants (3^{ème} ligne).

La figure VI.5 met en évidence une décroissance du SUV_{max} de la lésion 3 durant le traitement et une stabilisation de cet index entre les 2 derniers examens. Les examens TEP consécutifs (figure VI.7) montrent quant à eux que l'activité de la lésion diminue pour atteindre celle du fond dans la TEP(4). Au contraire de l'évolution du SUV_{max} , celle du facteur décroissant montre une activité constante entre les deux premiers examens puis une décroissance rapide de l'activité. Le plateau final obtenu entre les 2 examens finaux est, par contre, en accord avec l'évolution suivie par le SUV_{max} .

Les différences observées s'expliquent, comme dans le cas de la lésion 1, par le fait que le SUV_{max} ne représente pas l'évolution de l'ensemble de la lésion, à la différence de l'analyse factorielle. Ceci est confirmé par la décroissance de l'activité moyenne de la lésion vue par SUV_{adapt} , qui est moins rapide que celle du SUV_{max} (figure VI.5) pour les 3 premiers examens.

L'écart entre l'évolution du SUV_{max} et celle du facteur décroissant s'explique aussi par la croissance concomitante du volume métabolique durant les 2 premières séances de radiothérapie. La segmentation du volume métabolique par la méthode adaptative sur ces 3

premiers examens (figure VI.8, et examens TEP de la figure VI.7) montre en effet une augmentation du VM_{adapt} , inverse à celle de SUV_{adapt} de cette lésion.

Dans ce cas, il apparaît donc que le SUV_{max} , qui est la seule méthode classique qui permet de suivre le métabolisme de la lésion tout au long du traitement, n'est pas suffisant pour interpréter la complexité de l'évolution de cette lésion. En ne se référant qu'à l'évolution de cet index, la diminution de -57% du SUV_{max} entre TEP prétraitement et le second examen après le début de la RT laisse en effet penser à une réponse tumorale très nette. Cependant, couplée à l'évolution du volume métabolique de la lésion, qu'il n'est possible de décrire que sur 3 examens car les méthodes de segmentation échouent ensuite, cette décroissance du SUV_{max} apparaît moins univoque et met en évidence le besoin de décrire la lésion dans son ensemble et non pas à partir de l'activité d'un unique voxel. Le facteur décroissant obtenu par SULTAN, qui ne décroît que de 23% entre la TEP(1) et la TEP(3), laisse penser à un effet moindre du traitement par rapport à ce qui est suggéré par le SUV_{max} seul, en accord avec l'évolution de VM_{adapt} . SULTAN montre donc qu'il est possible de prendre en compte l'évolution de tous les voxels de la lésion mais surtout de décrire cette évolution quel que soit le moment du traitement, ce que ne permettent ni SUV_{max} , ni VM_{adapt} .

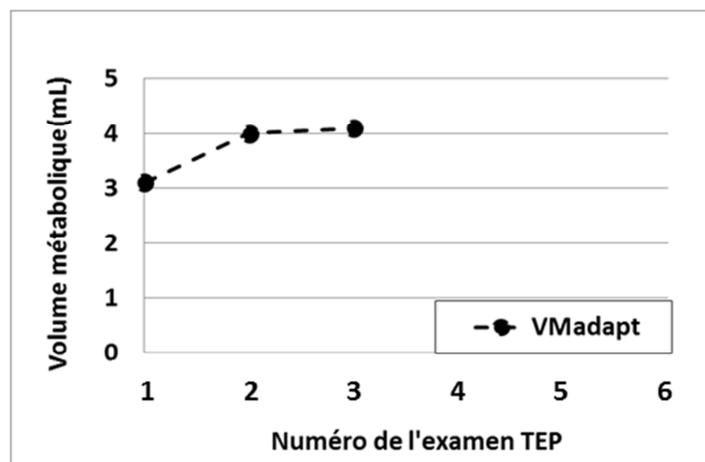


Figure VI.8 : évolution du volume métabolique de la lésion 3 vue par VM_{adapt} .

- Lésion 7

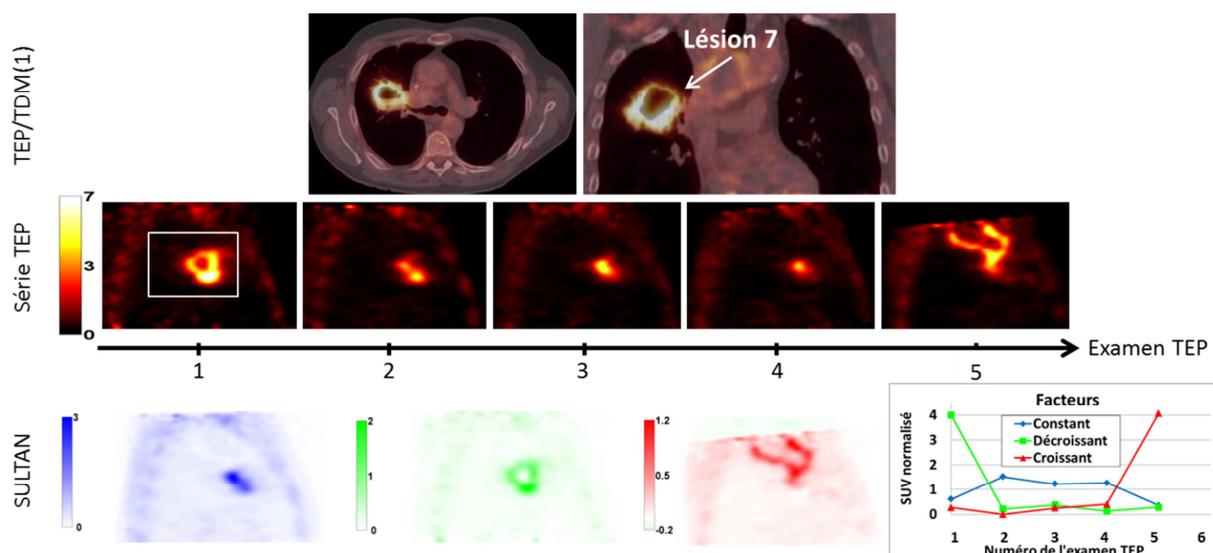


Figure VI.9 : 1^{ère} ligne : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (1 ici) montrant la lésion 7. 2^{ème} ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 7 et images factorielles et facteurs résultants (3^{ème} ligne).

L'analyse factorielle de la lésion 7 nécessite l'utilisation de 3 facteurs (figure VI.9, 2^{ème} ligne). La différence importante entre l'évolution du facteur décroissant et celle du SUV_{max} observée figure VI.5 s'explique par la nature du phénomène décrit par le facteur décroissant. En effet, le SUV_{max} montre une décroissance lente de l'activité du voxel de plus forte intensité jusqu'au TEP(4) puis une augmentation sur la dernière TEP ; a contrario, le facteur décroissant indique une disparition quasi complète de toute activité dans une zone de la tumeur entre la TEP(1) et la TEP(2), correspondant en partie au volume tumoral observé sur la TEP(1), d'après l'image factorielle associée. L'analyse des examens TEP représentés en figure VI.9 montre en effet que le volume tumoral initial, nécrosé, diminue de manière très importante entre les 2 premiers examens. La forme de la tumeur sur la TEP(2)) inclut une zone active à l'endroit où se trouvait la nécrose initialement (cf TEP(1)) et cette zone reste active jusqu'à la TEP(4). Ce phénomène d'affaissement de la tumeur puis de disparition est décrit par le facteur « constant » (figure VI.9) qui représente principalement l'évolution de l'activité dans la zone où ce phénomène se produit (cf zone de forte expression de ce facteur, au centre de la zone tumorale, dans l'image factorielle associée, figure VI.9). Sur la dernière TEP, on constate la formation d'une nouvelle zone métaboliquement active juste au-dessus du foyer résiduel de la TEP(4). Cette zone correspond à la zone rouge présente dans l'image factorielle associée au facteur croissant (figure VI.9) dont les valeurs prises jusqu'à la TEP(4) sont proches de 0 puis croissent rapidement entre le TEP(4) et la TEP(5). D'un point de vue médical, cette augmentation importante de l'activité a été associée à la formation d'une zone d'infection, donc sans rapport direct avec la tumeur pour laquelle le patient était initialement traité.

Pour comprendre l'augmentation du SUV_{max} sur le dernier examen (figure VI.5), nous avons évalué la proportion des valeurs prises par les $\tilde{A}_k(i)$ de chacune de 3 images factorielles obtenues, en fonction des positions successivement prises par le SUV_{max} dans la série d'images TEP (figure VI.10B). Nous avons de plus déterminé, dans la figure VI.11A, l'évolution de la distance euclidienne du voxel contenant le SUV_{max} dans le temps par rapport à sa position initiale, i.e. dans la TEP(1).

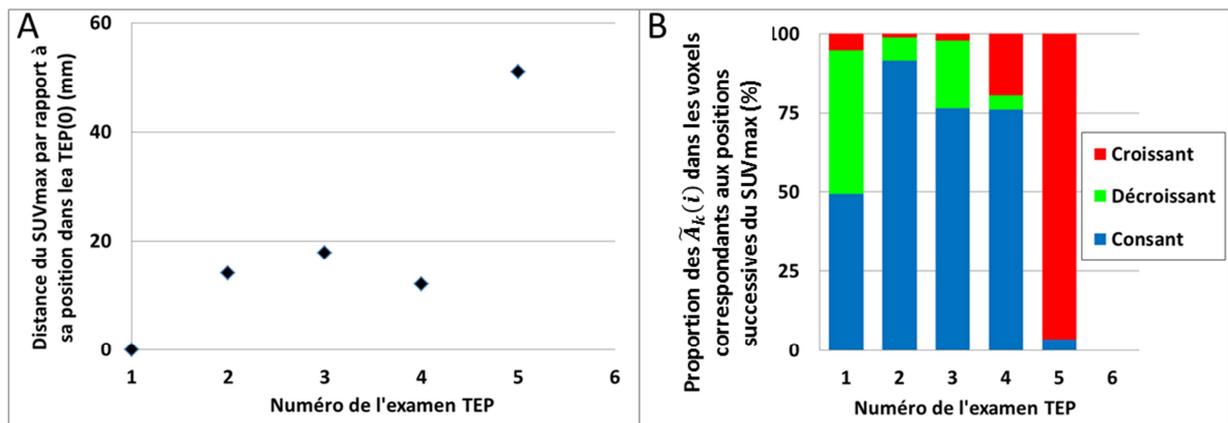


Figure VI.10 : A : distance du voxel contenant le SUV_{max} dans les examens TEP($t > 1$), par rapport à la localisation du voxel contenant le SUV_{max} dans la TEP(1). B : Proportion des $\tilde{A}_k(i)$ dans chacun des voxels occupés par le SUV_{max} de la lésion 7.

Il apparaît que si le SUV_{max} appartient bien à la zone décrite par le facteur constant dans les examens 2 à 4, il décrit quasi intégralement, dans le dernier examen, l'activité de la zone inflammatoire (figure VI.11B), c'est-à-dire celle d'un voxel dont l'activité a brutalement augmenté entre la TEP(4) et la TEP(5). Ce changement est à mettre en rapport avec le déplacement du voxel contenant le SUV_{max} dans l'examen 5 (figure VI.10A), qui semble en accord avec l'idée que l'activité qu'il caractérise ne correspond plus à celle du tissu initialement étudié.

Les résultats obtenus pour cette lésion suggèrent donc que le SUV_{max} , seul index permettant de suivre l'évolution de cette lésion du début à la fin du traitement, est mal adapté pour décrire les changements métaboliques de la lésion. A contrario, SULTAN peut permettre de distinguer entre les différentes évolutions des tissus irradiés et à séparer, à la manière des méthodes de segmentation floue (eg : (Hatt et al., 2009)), les évolutions des zones d'une même tumeur par exemple.

- Lésion 8

La figure VI.11 fournit les résultats de l'analyse factorielle relatifs à la lésion 8. L'analyse visuelle des examens TEP successifs montrent la faible évolution de l'activité métabolique de ce ganglion durant la radiothérapie, en accord avec la décroissance du SUV_{max} (qui passe alors de 7,9 à 6,0 $g \cdot mL^{-1}$ entre l'examen de référence et la TEP(3), soit une diminution de 25%) relevée dans le tableau VI.3 (colonne 10) et dans la figure VI.6. Cette diminution est à mettre en parallèle avec la variation plus importante du facteur décroissant dans le même intervalle, dont les valeurs varient de 1,4 (TEP(1)) à 0,75 (TEP(4)), soit une décroissance de 46% (figure VI.11). Cependant, on constate que pour cette lésion, comme pour la lésion 7, l'évolution du métabolisme glucidique est principalement décrite par le facteur constant (tableau VI.4). Autrement dit, l'écart entre le facteur décroissant et l'évolution du SUV_{max} s'explique par le fait que cette lésion est bien décrite par la combinaison d'un facteur décroissant et d'un facteur constant, comme on peut encore le voir sur les images factorielles associées aux lésions.

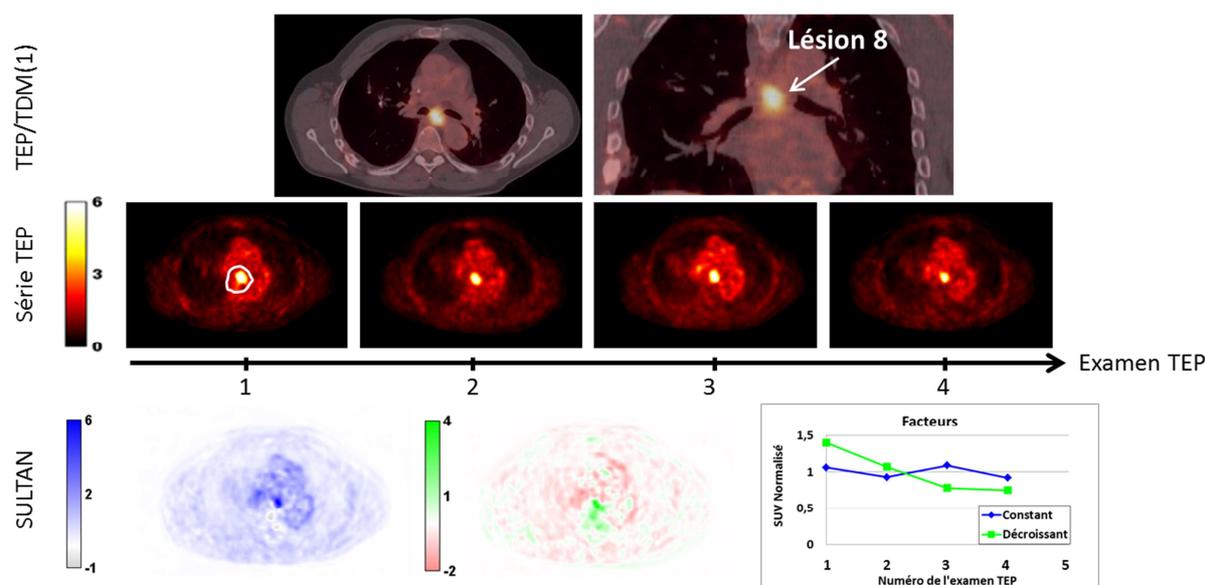


Figure VI.11 : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (1 ici) montrant la lésion 8. 2^{ème} ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 8 et images factorielles et facteurs résultants (3^{ème} ligne).

L'avantage de l'analyse factorielle dans ce cas est qu'elle permet de rendre compte du fait que l'activité de fond n'est pas stable dans le temps (via le facteur constant), en accord avec l'impression visuelle (cf activité de la zone médiastinale dans les examens TEP 1 à 4 de la figure VI.11). Selon nous, l'expression des deux facteurs au niveau de la tumeur permet de séparer la composante du signal spécifique à la tumeur de celle se comportant comme les tissus non tumoraux voisins.

Malgré un coefficient de détermination élevé entre les valeurs prises successivement par le SUV_{max} et celles du facteur décroissant de la lésion 5 (figure 5, $R^2=0,95$), nous avons enfin cherché à expliquer les différences visibles entre ces courbes.

- Lésion 5

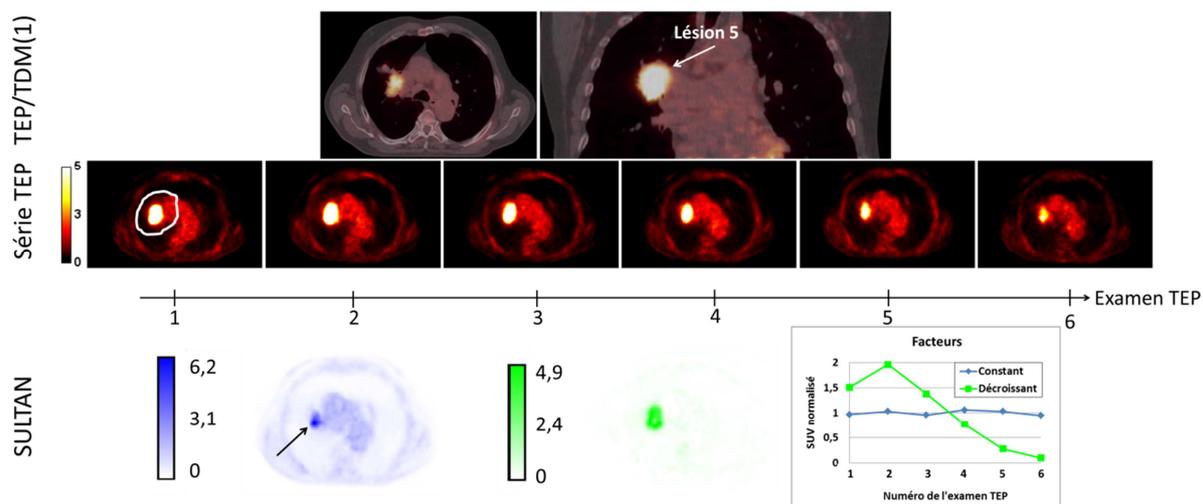


Figure VI.12 : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (1 ici) montrant la lésion 5. 2^{ème} ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 5 et images factorielles et facteurs résultants (3^{ème} ligne).

Comme dans le cas de la lésion 7, le facteur décroissant (figure VI.12, troisième ligne) donne une image de l'évolution des voxels localisés sur les bords de la tumeur. En effet, le centre de la couronne verte a une intensité plus faible que sur les bords (figure VI.12, image factorielle associée au facteur décroissant). A contrario, la zone de la lésion exprimant le plus fortement le facteur constant (cf flèche noire sur l'image factorielle associée au facteur constant, figure VI.12, troisième ligne) coïncide avec la partie centrale du volume métabolique initial, qui reste active du début à la fin du traitement comme le montre les examens TEP successifs.

En comparant l'évolution du volume métabolique de cette lésion ($VM_{adapt}(t)$, cf figure VI.13) à celle du facteur décroissant, on constate que ce dernier décrit en partie l'évolution de l'activité dans les voxels du bord de la tumeur et dont la fixation de FDG augmente entre la TEP(1) et la TEP(2). Le facteur décroissant traduit donc, comme dans le cas de la lésion 3, à la fois une modification d'activité et une modification du volume métaboliquement actif dans l'espace. Ce résultat montre l'importance de prendre en compte à la fois les informations images, fournies par les images factorielles et issues des examens TEP qui ont permis de d'obtenir ces résultats, et les informations courbes, représentées par les facteurs pour interpréter les résultats obtenus par SULTAN.

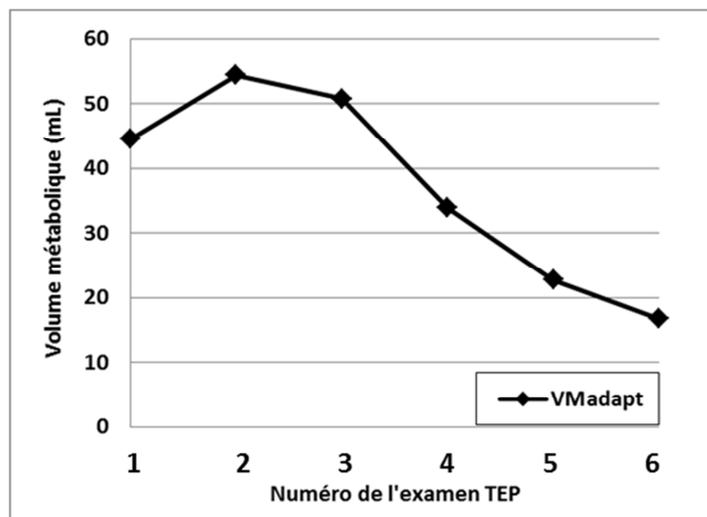


Figure VI.13 : évolution du VM_{adapt} de la lésion 5 durant le traitement.

VI.6. Discussion

L'objectif de ce travail était d'étudier la valeur ajoutée d'une approche d'analyse factorielle appliquée à des séries de plus de 3 examens TEP/TDM statiques acquis tout au long d'une RT ou d'une RT-CT de patients atteints de NSCLC. L'intérêt de cette approche dans le contexte du suivi thérapeutique est potentiellement double. Elle permet en effet :

- De rendre compte de l'hétérogénéité de la réponse tumorale à l'échelle du voxel, à la différence des méthodes régionales utilisées en routine clinique à l'heure actuelle.
- De prendre en compte les particularités des évolutions d'une population de voxels à partir d'un très petit nombre d'a priori sur le résultat attendu.

Le premier objectif de notre étude était d'évaluer la capacité de cinq méthodes classiques d'une part, et de SULTAN d'autre part, à suivre l'évolution de 5 tumeurs et de 11 ganglions durant le traitement. En accord avec les résultats déjà publiés sur cette cohorte (Edet-Sanson et al., 2012), l'accroissement de la dose cumulée diminue la capacité des algorithmes de segmentation semi-automatiques considérés à distinguer les zones tumorales des tissus sains avoisinants (figure VI.3). La méthode adaptative proposée par (Vauclin et al., 2009) obtient de meilleures performances que celle basée sur un seuil égal à 40% du SUV_{max} . Le taux d'échec de l'algorithme prenant en compte le contraste local reste en effet relativement faible jusqu'à ~20 Gy (15%) et se stabilise à ~50% au-delà de cette dose. La méthode utilisant $S_{40\%}$ ne permet, quant à elle, de suivre l'évolution que d'environ 50% des lésions une fois la dose de 10 Gy dépassée, pour atteindre les 90% d'échec de segmentation au-delà de 30 Gy. Ceci a justifié que nous ne considérions dans la suite que les informations fournies par la méthode adaptative. Parmi les méthodes classiques évaluées, SUV_{max} apparaît comme la seule méthode permettant de suivre l'évolution de l'activité de la lésion indépendamment de la dose délivrée au volume cible.

SULTAN a permis de suivre l'évolution de 15 des 16 lésions étudiées. Le seul échec correspond à l'une des 5 tumeurs de la cohorte (lésion 13, figure VI.4) située dans le lobe

pulmonaire supérieur gauche. Cet échec a été attribué en partie à la difficulté de recaler les données anatomiques entre elles dans cette zone. De plus, il s'est avéré ici très complexe de définir une VOI_{SULTAN} incluant un fond local dont l'activité soit suffisamment stable pour valider l'hypothèse d'un facteur constant dans le temps. Si SULTAN n'a pas permis de suivre l'évolution de cette lésion, SUV_{max} et S_{adapt} ont, eux, fournis des résultats cohérents (tableau VI.3).

Pour quantifier le degré d'hétérogénéité de la réponse tumorale au traitement vu par SULTAN, nous avons calculé la valeur moyenne, dans le $VM_{adapt}(t_{référence})$, des $\tilde{A}_k(i)$ associés aux différents facteurs (tableau VI.4). Si la cohorte de tumeurs étudiées contient peu de lésions dont la réponse semble hétérogène, il est intéressant de noter que celles qui obtiennent les $R_{décroissant}$ les plus faibles (0,45 et 0,40 pour les lésions 7 et 8 respectivement, cf tableau VI.4) apparaissent effectivement évoluer de manière très hétérogène durant le traitement, comme l'ont montré les discussions relatives à ces cas. Ce point semble donc valider l'utilisation du facteur $R_{décroissant}$ pour résumer l'information image. Cependant, si le calcul de $R_{décroissant}$ semble bien adaptée à notre cohorte de lésions (car le volume tumoral segmenté sur la TEP de référence finit par décroître entre le début et la fin du traitement), il doit être modifié pour des cas d'échappement à la thérapie, que nous n'avons pas observé dans cette cohorte. En considérant un cas pour lequel un facteur croissant décrirait une augmentation du signal à l'extérieur du volume tumoral défini initialement, le volume à considérer pour prendre en compte l'hétérogénéité de la réponse pourrait, par exemple, être le plus grand volume segmenté de la lésion durant les examens de suivi. D'autre part, une étude complémentaire très intéressante consisterait à étudier le lien entre le degré d'hétérogénéité de la réponse (vue par $R_{décroissant}$ par exemple) à la présence, avant le début du traitement, ou à l'apparition de populations, au cours de la thérapie, de clonogènes métaboliquement actifs résistant à l'irradiation.

Enfin, pour 6 des 16 tumeurs étudiées, (4, 5, 7, 8, 9 et 15, i.e. 3/5 tumeurs et 3/11 ganglions) le facteur décroissant représentait moins de 70% de l'évolution de la tumeur (tableau VI.4, colonne VI.5). En dehors de la lésion 4, toutes ces lésions ont un VM_{adapt} important ($>8,3$ mL), supérieur à la médiane de l'échantillon de volume métabolique VM_{adapt} considéré (4,2 mL). Ceci suggère que l'hétérogénéité de la réponse est davantage liée au volume métaboliquement actif qu'à la nature de la lésion (i.e. ganglion ou tumeur). Cette observation reste cependant à confirmer sur une cohorte de tumeurs faisant intervenir de plus petits volumes tumoraux, puis qu'il existe une différence significative entre la médiane du VM_{adapt} des ganglions et de lésions considérés dans cette étude ($p=0,0014$).

Nous avons ensuite cherché à déterminer l'apport de SULTAN lorsque le facteur décroissant était peu corrélé à l'évolution du SUV_{max} (figure VI.5). Nous avons constaté que dans les cas où la lésion était très petite (<3 mL, cas des lésions 4, 10, 11, 12, 14 et 16), une grande corrélation ($R^2 \geq 0,97$) était observée entre les 2 courbes. Ceci s'explique simplement par le fait que de telles lésions sont bien décrites par un très petit nombre de voxels et en particulier par le SUV_{max} , qui contient, du fait de l'EVP, une information importante sur l'activité métabolique de la lésion. Cette observation suggère un intérêt moindre de l'analyse factorielle dans les cas de très petites lésions. Ceci est d'autant plus vrai que SULTAN requiert un

prétraitement des données et notamment leur recalage successif pouvant nécessiter un ajustement manuel (dans 13% des cas rapporté ici). Ces étapes sont particulièrement critiques pour les petits volumes tumoraux. Dans ces cas, il apparaît donc que le SUV_{max} , seule méthode classique fournissant un résultat quel que soit la réduction du volume de la lésion (tableau VI.3), correspond à la meilleure stratégie de suivi parmi les différentes méthodes comparées dans cette étude. Ce point reste à nuancer pour les cas de lésions de volume un peu supérieurs, comme nous l'avons vu lors de l'analyse de la lésion 3, dont le volume initial était de 3,1 mL. Dans ce cas, et en accord avec l'interprétation des résultats relatifs à la lésion 5, nous avons en effet montré que SULTAN donnait une information sur une population de voxels dont l'évolution de l'activité pouvait différer largement de celle décrite par le SUV_{max} (figures VI.7 et VI.8). Dans ces cas, SULTAN permet d'intégrer une information sur l'évolution de tout le volume tumoral mais nécessite parfois (eg : lésion 7) une analyse visuelle de chaque examen pour pouvoir expliquer précisément les résultats.

L'identification des limites de l'analyse factorielle appliquée au contexte du suivi thérapeutique est un point d'autant plus important que pour les petits volumes, les méthodes de segmentation sont très vite limitées pour la description de l'évolution du VM (figure VI.8). A l'inverse, si le SUV_{max} permet toujours d'obtenir une valeur (figure VI.3) et donc de quantifier l'évolution du métabolisme glucidique de la lésion, sa localisation exacte n'est pas toujours simple à déterminer, comme dans le cas de l'évolution complexe de la lésion 7 (figure VI.10B).

Dans les cas où l'écart entre l'évolution du SUV_{max} et celle du facteur décroissant était plus importante ($R^2 < 0,9$), nous avons examiné au cas par cas les différences constatées, dans le but d'identifier les configurations où SULTAN présente une valeur ajoutée par rapport à l'analyse de l'évolution du SUV_{max} . Pour les 2 plus grosses lésions de la cohorte i.e. de 44,5 mL (lésion 5, figure VI.12) et 37 mL (lésion 7, figure VI.9) et dont la réponse est hétérogène ($R_{décroissant} < 0,62$, tableau VI.4), l'analyse factorielle a permis de visualiser dans l'espace différents types de comportement de la réponse qui ne sont pas distinguables par les méthodes classiques (e.g : zone d'inflammation ou affaissement du volume tumoral), si ce n'est par une analyse visuelle. De la même manière, l'analyse factorielle permet de séparer la contribution du signal propre à l'évolution de la tumeur de celui lié à la fluctuation, tout au long du traitement, de l'activité de tissus environnant (pouvant être dû à l'effet du traitement par exemple, lésions 1 et 8, figure VI.6 et VI.11 respectivement). Il apparaîtrait donc élégant de démontrer la capacité de SULTAN à distinguer entre la fixation de FDG liée à l'inflammation des tissus irradiés et celle relative à la prolifération tumorale en étudiant la corrélation entre les grandeurs anatomopathologiques associées à ces informations et les index quantitatifs extraits de l'analyse factorielle, tel que $R_{décroissant}$ par exemple.

Les limites du travail résident principalement dans le fait que, dans la cohorte de lésion étudiée, la majeure partie des lésions répondent très positivement et de manière homogène au traitement (13/15, soit 87% des lésions étudiées par SULTAN ont en effet un $R_{décroissant} > 0,6$ ainsi qu'un $\Delta SUV_{max} < -50\%$, cf tableau IV.4 Et IV.3 respectivement). Par conséquent, nous n'avons pas pu étudier de nombreux cas de réponse hétérogène pour lesquelles l'information apportée par SULTAN pourrait permettre, contrairement aux méthodes classiques, de

distinguer différents types de réponses à long terme. En outre, le trop faible nombre de gros volumes tumoraux considérés ne permet pas de généraliser l'interprétation proposée pour expliquer les différences entre facteur décroissant et évolution du SUV_{max} .

Comme nous l'avons vu, SULTAN inclut quelques étapes dépendantes de l'opérateur, comme le choix de la VOI à recalibrer et celui de la VOI utilisée pour définir les voxels d'intérêt (VOI_{SULTAN}). L'une des limites de la méthode, qui a été rencontrée par exemple lors de l'étude de la lésion 13, est la difficulté, dans certaines configurations, de recalibrer les données anatomiques entre elles. Le problème de mise en correspondance entre données anatomiques et données métaboliques s'est avéré critique lors de l'application de SULTAN sur une cohorte de patients atteints de cancers ORL métastatiques (Maisonobe et al., 2012), et notamment pour les lésions placées au niveau du diaphragme. Nous reviendrons sur cette étude en conclusion de ce chapitre. Ce problème, commun à toutes les méthodes d'analyse des séries d'images TEP à l'échelle du voxel (David et al., 2012; Necib et al., 2011) représente selon nous la plus grande limite à l'application de la méthode SULTAN.

VI.7. Conclusion

Le travail présenté dans ce chapitre a permis d'évaluer l'intérêt de l'analyse factorielle dans le contexte du suivi thérapeutique sur 9 patients atteints de NSCLC. Si la méthodologie mise en œuvre ne permet pas de justifier son utilisation par rapport au SUV_{max} pour des très petites lésions (<3 mL), SULTAN a cependant permis de suivre l'évolution de 15/16 (94%) lésions jusqu'au dernier examen TEP, à la différence des méthodes de segmentation classiques qui échouent dans plus de 40% des cas quand un examen est acquis à mi-traitement, i.e. quand la moitié de la dose prescrite a été délivrée. En analysant les cas où une grande différence était constatée entre l'évolution du SUV_{max} et le facteur décroissant obtenu pour chaque lésion, nous avons pu mettre en évidence l'apport de SULTAN, et notamment sa capacité à décrire dans l'espace l'hétérogénéité de la réponse tumorale. Selon nous, deux études complémentaires devraient être effectuées pour mieux étudier l'intérêt de l'analyse de l'hétérogénéité de la réponse tumorale à la thérapie.

- La première consisterait à étudier la capacité de la SULTAN à distinguer des cas d'échappement à la thérapie sur une cohorte homogène suffisamment importante. Un travail préliminaire a été effectué dans ce sens en dernière année de thèse. Il portait sur une petite cohorte de 5 patients présentant des cancers ORL et traités à l'institut Jules Bordet par Sorafenib (Maisonobe et al., 2012). A l'inverse de la cohorte NSCLC présentée ici, cette cohorte de lésions (N=21) contenait cependant peu de réponses positives au traitement. Une étude importante à mener, et pour laquelle SULTAN pourrait apporter une information plus riche que la seule mesure de ΔSUV_{max} , consisterait à évaluer l'intérêt de la mesure de l'hétérogénéité de la réponse, vue par exemple sur $R_{décroissant}$, et son lien avec la survie sans progression.

- De la même manière, pour valider les informations contenues dans les index quantitatifs qu'il est possible d'extraire de SULTAN (eg : $R_{\text{décroissant}}$), nous pensons que cette méthode devrait être appliquée sur une cohorte de tumeurs pour lesquelles une analyse anatomopathologique serait disponible. Cette étude pourrait permettre de valider la mesure de l'hétérogénéité de la réponse que nous avons discutée ici et ainsi comparer les résultats de cette approche macroscopique à un niveau microscopique.

Nous pensons donc que la méthode décrite ici revêt un intérêt particulier pour la radiothérapie et le suivi thérapeutique dans la mesure où elle permet de distinguer les principales évolutions tissulaires tumorales, de les localiser de manière précise dans l'espace ou encore de spécifier le degré d'hétérogénéité de la réponse à la thérapie.

Conclusion
&
Perspectives

Commercialisée dès 2001, la TEP/TDM présente des avantages largement admis de nos jours. En oncologie, que ce soit pour la détection, le bilan d'extension, le diagnostic ou le suivi thérapeutique, son intérêt a été démontré à maintes reprises dans des études cliniques. Ces travaux se basent dans la majorité des cas sur la mise en évidence de la fixation d'un analogue du glucose, le FDG, dont la consommation par les cellules cancéreuses est, le plus souvent, exacerbée.

Si les corrections mises en œuvre, l'amélioration des pratiques et les avancées technologiques permettent d'obtenir des images quantitatives en routine clinique, les protocoles d'analyse des données utilisés à l'heure actuelle peuvent encore être optimisés. En effet, en dépit de l'existence de recommandations pour l'interprétation des images dans le contexte de l'évaluation de la réponse à la thérapie, comme celles données dans les protocoles EORTC et PERCIST, les meilleures manières d'extraire une information quantitative à partir des examens TEP et de l'interpréter restent à déterminer. En imagerie des lésions tumorales, potentiellement petites, le rôle de l'effet de volume partiel (EVP) mérite aussi d'être exploré et mieux compris. Enfin, on peut imaginer que des méthodes plus sophistiquées que la simple estimation d'une valeur caractérisant une lésion puissent permettre d'extraire des informations plus riches à partir des examens TEP. Notre travail de thèse a essentiellement consisté à explorer différentes voies permettant d'améliorer l'interprétation quantitative des examens TEP, dans le cadre du suivi thérapeutique.

Tout d'abord, pour déterminer quel index était le plus à même de discriminer entre des évolutions tumorales distinctes, nous avons évalué, sur des données simulées, la capacité de différents index semi-quantitatifs à séparer des groupes de tumeurs dont les évolutions étaient connues a priori. Nous avons montré que si l'objectif est de distinguer une évolution d'activité métabolique seule, alors le SUV corrigé de l'EVP est le mieux adapté. A l'inverse, dans le cas où la problématique est la discrimination entre un groupe de tumeurs n'évoluant pas et un groupe de tumeurs pour lequel au moins un paramètre change (ici, son volume métabolique ou son activité), alors le fait de corriger le SUV de l'EVP, et donc de supprimer l'information de volume métabolique que cette grandeur contient, s'avère délétère : il est préférable de considérer le SUV non corrigé de l'EVP pour la classification.

Nous avons ensuite cherché à savoir laquelle des deux situations étudiées sur données simulées était pertinente dans le contexte clinique. Pour répondre à cette question, nous avons étudié les données TEP/TDM de 40 patients, soit 101 lésions (dont 28% faisaient moins de 5 mL), atteints de cancers colorectaux métastatiques traités par chimiothérapie. Nous avons alors montré que les évolutions relatives de SUV_{max} , SUV_{fond} , SUV_{peak} , $SUV_{70\%}$ et TLG, mesurées entre un examen TEP réalisé avant le début du traitement et un autre acquis deux semaines après le début de la chimiothérapie prédisaient la réponse tardive vue par les critères RECIST 1.0 sur données TDM après 6 à 8 semaines de thérapie. La mesure de l'activité métabolique par des SUV corrigés de l'EVP n'améliore pas la distinction entre lésions répondant et lésions ne répondant pas à la thérapie. Ceci est probablement dû au fait que la correction de l'EVP supprime l'information sur le volume métaboliquement actif contenue dans la mesure d'activité. Le changement, entre le début du traitement et après deux semaines de chimiothérapie, du volume métabolique seul prédit mal la réponse tumorale vue à plus long

terme par la TDM, probablement du fait de la forte variabilité des estimées de volume métabolique.

Pour augmenter la quantité d'informations extraites des données TEP et dépasser la seule prise en compte du volume et/ou du SUV, quelques équipes ont proposé de mesurer l'hétérogénéité de la fixation de radiotracer dans les lésions. Des résultats très prometteurs ont été obtenus chez des patients atteints de cancer de l'œsophage. Nous avons étudié l'intérêt des mesures de texture sur notre cohorte de patients atteints de cancers colorectaux métastatiques, dans le but de 1/ Valider la valeur pronostique de ces index sur une autre pathologie ; 2/ Etudier leurs capacités à évaluer la réponse précoce ; 3/ Déterminer les sources de variabilités potentielles de ces index en fonction de paramètres tels que la méthode de ré-échantillonnage ou de segmentation, par exemple.

Pour le pronostic de la réponse thérapeutique à partir de la TEP initiale, les résultats de cette étude n'ont pas confirmé les résultats encourageants rapportés précédemment dans le cas du cancer de l'œsophage. La comparaison des valeurs d'index de texture avant thérapie et pendant la thérapie n'a, elle non plus, pas permis d'améliorer la classification réalisée au moyen du SUV et employée en routine clinique. D'après l'ensemble de nos résultats, les paramètres de texture, leur applicabilité et leur supériorité par rapport aux paramètres utilisés jusqu'à présent restent à démontrer. L'étude des différences méthodologiques entre les travaux précédents et les nôtres n'a pas permis d'expliquer totalement les écarts observés. Cependant, l'analyse de la corrélation entre les index de texture les plus mis en avant dans la littérature a permis de montrer une corrélation très forte avec le volume métabolique, qui, dans le cas de cancers colorectaux, à l'inverse des cancers de l'œsophage, est peu prédictif de la réponse au traitement. Le développement d'une nouvelle méthode de discrétisation pour le calcul de ces index de texture a cependant permis d'augmenter la capacité prédictive de ces index de manière significative pour 7/31 index, les rendant prédictif dans 18 cas. Malgré ce gain, les index de texture ne sont pas apparus plus prédictifs que les métriques classiques telles que le SUV_{peak} sur la cohorte étudiée. L'intérêt de cette approche reste à valider sur une cohorte plus importante.

L'information prise en compte à l'heure actuelle sur les données TEP est très partielle, car elle ne reflète qu'une information moyenne globale de chaque lésion. De plus, aucune méthode n'a été décrite pour prendre en compte l'évolution du métabolisme glucidique entre plus de 2 examens. Pour pallier à ces deux problèmes, nous avons évalué une méthode d'analyse factorielle, appelée SULTAN, chez neuf patients atteints de cancer du poumon non à petites cellules et traités par radiothérapie ou radio-chimiothérapie. Les résultats suggèrent que la complexité de la méthodologie mise en œuvre ne se justifie pas par rapport à la mesure, très simple, du SUV_{max} , pour des très petites lésions (<3 mL). Cependant, SULTAN a permis de suivre l'évolution de 15/16 (94%) lésions jusqu'au dernier examen TEP considéré, à la différence des méthodes de segmentation classiques qui échouent dans plus de 40% des cas quand un examen est acquis au milieu du traitement. En analysant les cas où une grande différence était constatée entre l'évolution du SUV_{max} et les résultats produits par SULTAN, nous avons mis en évidence l'apport de SULTAN, et notamment sa capacité à décrire l'hétérogénéité spatiale de la réponse tumorale.

Perspectives

Les trois études réalisées dans ce travail de thèse visaient à explorer la valeur ajoutée, les limites et l'intérêt potentiel en routine clinique de techniques d'analyse appliquées au domaine de la TEP en oncologie. Pour ce faire, nous avons considéré exclusivement des données TEP/TDM au FDG, qui est de loin le radiotracer le plus utilisé à l'heure actuelle en oncologie.

L'utilisation d'autres traceurs, tels que le FLT ou le FMISO, devrait apporter des informations utiles à la détermination des paramètres cellulaires les plus prédictifs de l'évolution de la maladie. La mise en œuvre de ces traceurs permettra de déduire les informations macroscopiques qui rendent compte des paramètres cellulaires sous-jacents. Les méthodes de traitement d'images que nous avons étudiées sont applicables sur n'importe quel type de données TEP, et ce quel que soit le radiotracer utilisé. Il n'est donc pas exclu que sur des images acquises avec des traceurs plus spécifiques, les paramètres de texture s'avèrent apporter des informations plus riches que sur des examens basés sur le FDG. L'idée serait alors d'identifier les grandeurs les plus à même de décrire, par exemple, la présence de nécrose, de zones d'hypoxie ou encore l'hétérogénéité du microenvironnement cellulaire de la tumeur.

Les résultats que nous avons obtenus laissent penser que l'intérêt des informations de texture n'est pas univoque et pourrait dépendre de la nature des lésions étudiées ou de la qualité des images TEP utilisées. En dehors de ces causes, qui pourraient expliquer les écarts importants entre ce qui a été décrit dans la littérature et nos résultats, il apparaît nécessaire de chercher à quelle information biologique les grandeurs que nous calculons se rapportent. Dans cette optique, la corrélation d'informations anatomopathologiques à celles des index de texture, maintenant initiée dans le cadre d'une collaboration avec le Dr. Soussan, nous semble aussi primordiale que prometteuse.

Quelle que soit la méthode de correction de l'EVP mise en œuvre, nous avons constaté que la mesure de l'activité métabolique seule réduisait la capacité de la TEP, acquise précocement au début du traitement, à prédire l'évolution du volume anatomique à long terme. La valeur ajoutée des stratégies de correction de l'EVP utilisées apparaît donc décevante, de même que les alternatives que nous avons proposées pour coupler à la fois les informations de volume et d'activité. Il en ressort que le SUV non corrigé de l'EVP reste un index actuellement très robuste pour rendre compte de l'activité tumorale. Dans un futur proche, l'amélioration de la résolution spatiale des tomographes, qui permettrait d'accroître la précision des méthodes de segmentation et de diminuer l'effet de l'EVP sur la mesure d'activité, devrait conduire à une mesure plus fiable des paramètres tels que le TLG, ou l'analyse biparamétrique de l'évolution du SUV corrigé de l'EVP et de l'évolution du volume métabolique.

La méthode SULTAN que nous avons évaluée dans le dernier chapitre de ce manuscrit a été proposée initialement pour l'amélioration du suivi dans le cas où plus de deux examens sont acquis au cours du traitement. Dans ce contexte, le suivi de la maladie est en effet rendu complexe notamment du fait de la difficulté de discriminer entre l'activité du fond et celle de

la lésion résiduelle à mesure que le traitement progresse. Ceci rend la segmentation des lésions complexe et leur caractérisation, par des index tels que le volume ou le SUV, délicate. Un point à approfondir dans le futur serait de combiner les données issues de SULTAN aux analyses de texture que nous avons étudié afin de décrire plus précisément l'hétérogénéité de la réponse à la thérapie.

Si SULTAN est plus complexe à mettre en œuvre que les méthodes classiques, sa précision, de l'ordre du voxel, la rend cependant intéressante pour l'amélioration des plans de traitement en radiothérapie. Le potentiel de cette méthode est d'autant plus évident avec l'avènement des méthodes de *dose painting* et de radiothérapie adaptative. La visualisation des différentes zones de réponse et de progression de la maladie dans l'espace pourrait en effet permettre d'augmenter les doses d'irradiation prescrites aux volumes cibles durant le traitement et donc d'améliorer à la fois le contrôle de la maladie tout en diminuant la dose aux tissus sains.

Les progrès à venir en TEP permettront de repousser les limites de la quantification, que nous avons évoquées tout au long de ce manuscrit. Ces évolutions passeront probablement par 1/ la combinaison des informations d'examen TEP issues de l'injection de différents traceurs, 2/ l'amélioration des performances des méthodes de segmentation grâce à des images moins bruitées, 3/ l'usage de méthodes de quantification plus évoluées que les méthodes actuelles, 4/ la mise en commun et l'évaluation, sur de plus grandes cohortes, de méthodes de quantifications plus performantes ou encore 5/ le fait d'imposer des pratiques communes à tous les centres, à l'échelle européenne notamment. En parallèle selon nous, la compréhension de la complexité des mécanismes biologiques mis en œuvre dans les cancers (Hanahan and Weinberg, 2011) nécessitera la combinaison d'informations provenant de différentes modalités d'imagerie, comme cela a été suggéré récemment lors du lancement du projet Radiomics (Lambin et al., 2012) par exemple. Dans cette optique, la compréhension des limites des informations qu'il est possible d'extraire des examens TEP s'inscrit dans une démarche nécessaire visant à mieux déterminer et à mieux comprendre les paramètres clés de ces maladies.

Liste des figures

Chapitre I

Figure I.1 : Résumé des 10 grandes caractéristiques des tissus tumoraux (Hanahan and Weinberg, 2011).	15
Figure I.2 : métabolisme glucidique des cellules de mammifères. Le sang apporte de l'oxygène (via l'hémoglobine) et du glucose aux cellules constituant les tissus. Le glucose entre dans le cytoplasme par l'intermédiaire de transporteurs spécifiques (famille des GLUT, protéines de la famille des transporteurs de glucose) où il est alors transformé en pyruvate. Ce pyruvate sera alors être soit réduit directement en lactate dans le cas d'absence d'oxygène (respiration anaérobie) pour ne former que 2 ATP / glucose soit être oxydé dans la mitochondrie et former jusqu'à 36 ATP (phosphorylation oxydative). (Schéma extrait de (Gatenby and Gillies, 2004))......	16
Figure I.3 : Principe de la TEP et de la mesure en coïncidence (gauche) et illustration des évènements suivant l'émission du positon lors de la désexcitation d'un atome de ^{18}F (droite). d : distance effective moyenne parcourue par le positon. Elle correspond à la distance entre le noyau mère et le lieu de désintégration du positon, mesurée perpendiculairement à la ligne définie par les directions prises par les deux photons produits suite à l'annihilation.....	20
Figure I.4 : Principe d'une coïncidence vraie (A), d'une coïncidence diffusée (B), d'une coïncidence fortuite (C) et d'une coïncidence multiple (D) en TEP.	24
Figure I.5 : principe de la localisation du lieu d'annihilation avec la mesure de temps de vol (Time of Flight, TOF) (source : (Lecomte, 2009)).	25
Figure I.6 : Définition des types d'acquisition en TEP : A) acquisition 2D avec plan droit et plan croisé B) 2D haute sensibilité avec plan droit et plan croisé et C) acquisition 3D (extrait et modifié de (Turkington, 2001)).	27
Figure I.7 : effet de la correction d'atténuation sur l'image d'un cylindre d'activité homogène. Extrait et traduit de (Cherry et al., 2006).	29
Figure I.8 : modèle compartimental utilisé pour décrire la métabolisation du FDG dans les tissus.	40
Figure I.9 : Données collectées pour l'analyse compartimentale d'un adénocarcinome pulmonaire (T2) du lobe supérieur droit (A). TAC de la tumeur (bleu) et de l'aorte ascendante (rouge) (B). (Extrait et traduit de (Vriens et al., 2010)).	42
Figure I.10 : courbe obtenue suite à l'analyse de Patlak de la tumeur présentée à la figure I.9. Extrait et traduit de (Vriens et al., 2010).	43

Chapitre II

Figure II.1 : Histogrammes Intensité-Volume (IVH) d'une tumeur du col de l'utérus du Volume Tumoral Clinique (CTV, utilisé pour déterminer le plan de traitement en radiothérapie, courbe marron) et du volume segmenté sur la TEP grâce à un iso contour égal à 40% du SUV maximal de la tumeur (courbe verte). Extrait de (El Naqa et al., 2009). **77**

Figure II.2: courbe du volume en fonction du seuil exprimé en pourcentage du SUVmax, utilisé pour le segmenter. Les auteurs utilisent la pente de la droite d'ajustement (droite en trait plein), tracée à partir des points obtenus pour un seuil compris entre 40% et 80% du SUVmax de la tumeur, pour caractériser l'hétérogénéité de la distribution de FDG dans la lésion. La courbe en pointillés correspond à l'ajustement, par une fonction exponentielle décroissante, de l'ensemble des points (extrait de (Brooks and Grigsby, 2011)). **79**

Figure II.3 : A : coupe transverse de la TEP/TDM (notée TEP(T1)) acquise avant le début du traitement d'un patient atteint d'un cancer colorectal métastatique. Les flèches montrent 5 lésions métaboliquement actives. B : coupe transverse de la TEP/TDM acquise après 1 cycle de chimiothérapie (TEP(2)). C : image paramétrique obtenue suite à l'application de la méthode décrite par (Necib et al . 2011). D : graphique de soustraction avec identification de 3 composantes : le bruit correspond à l'amas bleu, la captation physiologique (relative au foie) est représentée en rose et l'évolution tumorale est représentée en vert. Extrait de (Necib et al. 2011). **86**

Figure II.4 : étapes mises en œuvre par (David et al., 2012). Après recalage, correction de l'effet de volume partiel et segmentation des données TEP, une première étape vise à cartographier la zone tumorale avant et après traitement au moyen d'une méthode de segmentation floue (les voxels noirs appartiennent à la tumeur, le reste est attribué au fond). A partir des fonctions de croyances dérivées de l'analyse paradoxale, on génère une image binaire représentant les changements dans l'image (voxels blancs) et les voxels n'ayant pas changé (voxels en noir). Pour distinguer parmi ces 2 familles de voxels ceux appartenant à la tumeur et ceux du fond d'une part, et ceux correspondant à une progression ou une diminution de l'activité de la tumeur d'autre part, on utilise soit les données segmentées, soit les signes de l'image de soustraction. On déduit de ce résultat une carte de changement de la tumeur à partir de laquelle on calcule l'évolution du SUV et du volume dans une zone contenant les voxels constants et décroissant sur la TEP initiale (voxels vert et bleu ici) et contenant les voxels stables et croissant sur le second examen TEP (vert et rouge, non présents dans l'exemple proposé). **90**

Chapitre III

Figure III.1 : la résolution transaxiale inclut une composante tangentielle et une composante radiale. En éloignant la source radioactive de l'axe du tomographe, la probabilité pour que les photons incidents interagissent avec les scintillateurs de plusieurs détecteurs avant d'être absorbés augmente. Il est donc plus difficile de définir avec précision le lieu d'interaction des gammas de 511 keV quand la distance à l'axe augmente, d'où une résolution spatiale dégradée dans cette direction. **95**

Figure III.2: effet de la résolution spatiale du système d'imagerie sur la mesure d'une source circulaire de rayon 10mm et d'activité uniforme (en unités arbitraires) dans un fond d'activité nulle. Les conséquences sur la mesure sont un étalement du signal à l'extérieur des limites de l'objet mesuré (*spill-out*) et une diminution de l'activité maximale mesurée. (Extrait et adapté de (Soret et al., 2007)). **96**

Figure III.3: effet de l'échantillonnage sur la mesure d'un cercle d'activité homogène entouré d'une activité nulle (A). L'activité mesurée dans les pixels se trouvant à la périphérie du cercle inclut à la fois l'activité du signal provenant de l'objet et celle du fond. (B) Une partie du signal de l'objet étudié est mesurée au-delà de ses bords externes (*spill-out*, C). Extrait et traduit de (Soret et al., 2007). **97**

Figure III.4: impact de l'effet de volume partiel sur la mesure de l'activité en fonction du diamètre de sources sphériques. Ligne 1) coupe transverse des sphères étudiées, d'activités homogènes unitaires identiques (u.a.) et de diamètres décroissants. Ligne 2) Effet de l'EVP (simulé ici par la convolution des images idéales par une FDR de LMH = 10 mm) sur l'image des objets étudiés et (ligne 3) profils comparés des distributions. Extrait et adapté de (Cherry et al., 2006). **98**

Figure III.5 : Objet mesuré (activité en u.a) (A), effet du spill-in (B) et du spill-out (C) sur cette même distribution et image mesurée résultant de ces deux effets (D). Extrait et adapté de (Soret et al., 2007). **98**

Figure III.6 : effet de l'activité des tissus environnant sur l'activité mesurée de l'objet étudié. **100**

Figure III.7: effet de l'échantillonnage sur la distribution d'activité (en u.a.) mesurée d'un disque (A) pour une taille de pixel de 4,2 et 5 cm (B et C, respectivement). **101**

Chapitre IV

Figure IV.1 : rendu 3D du fantôme NCAT (*Nurb-based Cardiac-Torso phantom*) utilisé pour simuler des données réalistes (extrait de www.bme.unc.edu/~wsegars/) **131**

Figure IV.2 : Fantôme NCAT simulé avec GATE. Les sphères, représentant des tumeurs, sont indiquées avec les flèches vertes. Les images étant non synchronisées à la respiration. On remarque que le flou cinétique est important et modifie la forme des sphères. Extrait du rapport de thèse de (Tylski, 2009). **132**

Figure IV.3 : illustration des contours utilisés pour calculer les différentes définitions de SUV comparés dans l'étude sur données simulées. **134**

Figure IV.4 : exemple d'appariement de données utilisé pour simuler différents types d'évolution des tumeurs. Première ligne : cas d'une tumeur dont l'activité réelle reste égale à 8 (unités SUV) mais dont le diamètre augmente de 1 cm à 3 cm. Le SUVfond mesuré passe alors de 2,2 à 7,0 g/mL. Ligne 2 : cas d'une tumeur n'évoluant pas entre les deux examens mais pour lesquelles deux simulations utilisant des séries de nombres aléatoires différentes ont été utilisées. Dans ce cas, le SUVfond mesuré change légèrement (de 3 à 3,2.) **135**

Figure IV.5 : Illustration des distributions des valeurs fournies par une méthode de diagnostic pour deux populations distinctes, dont les réponses sont connues a priori par une méthode de référence. En choisissant un seuil de décision (A), correspondant à une valeur donnée de l'indice évalué, on définit la proportion de Vrais Négatifs (VN) et Faux Positifs (FP) dans la population saine (B) et celle de Faux Négatifs (FN) et Vrais Positif (VP) dans la population malade (C). **137**

Figure IV.6 : Exemple de courbes ROC (gauche) et des distributions de valeurs correspondantes pour deux populations et quatre méthodes de diagnostic. Les indices 1 et 2 correspondent à des cas extrêmes où la méthode étudiée permet de distinguer totalement (indice 1) ou pas du tout (index 2) les différents groupes de patients. **138**

Figure IV.7 : courbes ROC des 5 indices pour la séparation des groupes de tumeurs G1 et G2 (N=378 lésions par groupe). **140**

Figure IV.8 : courbes ROC des 5 indices comparés pour la séparation des groupes G3 et G4 de tumeurs (N=126 lésions par groupe). **141**

Figure IV.9 : frise chronologique indiquant les dates d'acquisition des données par rapport au traitement. **146**

Figure IV.10: coupe transverse de l'acquisition sur données fantômes de contraste le plus élevé (10 :1) utilisée pour optimiser la méthode de segmentation prenant en compte l'activité de fond (Nestle et al., 2005). **147**

Figure IV.11 : exemples de graphiques de Bland-Altman. (a) : cas pour lequel un très bon accord moyen entre les méthodes évaluées est observé. (b) : exemple où il existe une erreur systématique absolue entre les 2 index comparés. (c) : cas de deux index dont l'écart varie linéairement avec l'amplitude de la mesure et (d) cas de deux index dont les valeurs diffèrent d'autant plus que la valeur mesurée est grande. **149**

Figure IV.12 : Proportion de tumeurs dont le volume est inférieur ou supérieur à 5 ml pour l'ensemble des tumeurs étudiées (histogramme de gauche), seulement les lésions classées comme répondant à la thérapie (milieu) et seulement les lésions ne répondant pas (droite). **151**

Figure IV.13 : graphiques de Bland-Altman comparant les différents index non corrigés de l'EVP (a à f) et des 2 index corrigés (g) pour les 101 lésions étudiées avant le début du premier cycle de chimiothérapie. La différence moyenne entre chaque paire d'index ainsi que cette moyenne ± 2 écart-type (σ) sont présentés en lignes pleines et en pointillés, respectivement. Les valeurs correspondantes sont données entre parenthèses. Le graphique h représente les valeurs du rapport entre SUVcr et SUVdecon en fonction de la moyenne des différentes valeurs. Il illustre la corrélation entre les valeurs prises par ces deux index, également mise en évidence par la forte relation linéaire présente sur le graphique g. **153**

Figure IV.14 : Courbes ROC de Δ SUVfond, Δ SUVpeak, Δ SUVmax, Δ SUVcr, Δ SUVdecon, Δ SUV70%, Δ TLG et Δ VM calculés à partir d'une TEP/TDM acquise précocément après le début du traitement (J14) pour l'identification des lésions répondant (NRP = 27) et ne répondant pas à la thérapie (NMS + NMP = 74) suivant la réponse RECIST 1.0 établie sur données anatomiques après 6 à 8 semaines de traitement. **154**

Figure IV.15 : cas d'une lésion pour laquelle la VOI utilisée pour calculer le SUVpeak inclus des voxels situés à l'extérieur du volume segmenté par la méthode de (Nestle et al., 2005). Il en découle une valeur du SUVpeak plus petite que celle du SUVfond, **156**

Figure IV.16 : graphique de Bland-Altman entre les valeurs de SUV corrigées par la méthode de déconvolution de VC (SUVdecon) et LR (SUVLR) pour les 101 lésions sur la TEP avant traitement. **158**

Figure IV.17 : graphique 2D représentant la variation relative de SUVdecon (axe des abscisses) et de volume métabolique (axe des ordonnées) entre les deux examens TEP. **160**

Chapitre V

Figure V.1 : coupe coronale de la TEP/TDM au FDG d'un patient atteint d'un cancer du poumon non à petites cellules présentant une importante nécrose centrale. Extrait de (Orlhac, 2012)..... **164**

Figure V.2 : Principe, pour le cas d'une analyse à 2 dimensions, de la construction des 4 matrices utilisées pour l'analyse de la texture de la fixation de FDG vue par la TEP. Le ré-échantillonnage utilisé ici à des fins d'illustration ne se base que sur 4 valeurs de niveau de gris ($d=2$). Les pixels les plus sombres n'appartiennent pas au volume segmenté et ne sont donc pas pris en compte pour l'analyse. Les flèches situées dans les cases supérieures gauches des matrices de séries homogènes et de cooccurrence correspondent au sens de parcours choisi a priori pour définir, respectivement, les tailles des séries homogènes et paires de niveaux de gris rencontrés dans l'image ré-échantillonnée. Extrait de (Orlhac, 2012). **171**

Figure V.3 : courbes ROC de SUV_{peak} et Homogeneity pour la distinction entre tumeurs CR et PR d'une part et MS et MP d'autre part, à partir d'un examen TEP prétraitement seul. Extrait et modifié de (Orlhac, 2012). **181**

Figure V.4 : Aires sous les courbes ROC (AUC) et écarts-types associés des 37 index calculés à partir de l'examen TEP(PT). Extrait et modifié de (Orlhac, 2012). * : valeur significativement supérieure à 0,5 ($p<0,05$)..... **181**

Figure V.5 : AUC (et écart-types) des 37 index mesurant l'évolution relative entre les deux examens TEP. AUC significativement supérieure à 0,5 (* : $p<0,05$; ** : $p<0,01$). **182**

Figure V.6 : AUC (et écart-types correspondants) de chaque paramètre de texture en fonction du niveau de discrétisation utilisé. Afin de mieux mettre en évidence les différences potentielles, l'intervalle de valeur de ces AUC a été limité à [0,5-0,75] et non pas [0,5-1] comme c'était le cas dans la figure V.4 par exemple..... **183**

Figure V.7 : Graphique de Bland-Altman comparant les valeurs des volumes métaboliques obtenus en utilisant deux méthodes de segmentation. **184**

Figure V.8 : Ecarts relatifs moyens $\Delta(\%)$ (± 1 EC) calculés en considérant toutes les tumeurs pour chaque index et déterminé à partir des méthodes de segmentation Sfond et S40%. La valeur présentée pour SUV_{moyen} correspond à l'écart relatif moyen calculé en comparant les SUV_{fond} et SUV_{40%}. L'échelle choisie ne permet pas de représenter l'écart-type obtenu pour Busyness, qui vaut 480,3% **185**

Figure V.9 : comparaison des AUC des courbes ROC des différents index obtenues en segmentant les volumes des lésions à l'aide d'un seuil égal à 40% du SUV_{max} (barres

hachurées) et sur un seuil prenant en compte le fond local (barres pleines). Seul l'examen TEP(PT) est pris en compte ici. Les barres d'erreurs rapportées représentent 1 EC.....186

Figure V.10 : corrélation entre GLNUz et le volume métabolique calculé à partir du seuil Sfond (A) et S40% (B).....190

Figure V.11 : AUC (+/- 1 ET) des courbes ROC des index métaboliques pour deux méthodes de ré-échantillonnage à partir de la TEP prétraitement seule. * : AUC significativement supérieure à 0,5 ($p < 0,05$).....190

Figure V.12 : AUC (+/- 1 ET) des courbes ROC des index métaboliques pour deux méthodes de ré-échantillonnage, obtenues en calculant l'évolution relative des index entre l'examen prétraitement et après 14 jours de chimiothérapie. AUC significativement supérieure à 0,5 (étoiles noires telles que * : $p < 0,05$; ** : $p < 0,01$). Les étoiles rouges marquent les cas où les paires d'AUC obtenues pour chacun des index sont significativement différentes en utilisant l'une ou l'autre méthode de ré-échantillonnage.191

Chapitre VI

Figure VI.1 : illustration du principe de l'analyse factorielle appliquée à une série de T examens TEP acquis durant la radiothérapie d'un patient atteint de NSCLC. On montre ici la décomposition, sous la forme d'une somme linéaire de 3 facteurs (3 courbes de couleurs) de l'activité d'un des N voxels inclus dans le volume étudié. Les images factorielles associées à ces facteurs (en dessous des courbes respectives) donnent pour chaque voxel i de l'image le poids associé au facteur correspondant.....203

Figure VI.2 : méthode utilisée pour recalcr les données TEP entre elles. On délimite, sur chaque examen TDM de la série, la zone à recalcr, contenant la lésion étudiée, des tissus sains avoisinants et les repères osseux caractéristiques les plus proches (ici, les côtes par exemple). A partir d'un recalcr rigide, on obtient la transformation $T_i \rightarrow 1$ à appliquer à chaque TDM(i) pour mettre en correspondance la zone étudiée sur la TDM(1) (première ligne de la figure). Une fois ces transformations $T_i \rightarrow 1$ déterminées, on les applique à chacune des TEP(i) associées afin de recalcr la même zone mais entre les TEP(i) et la TEP(1) (ligne 2).205

Figure VI.3 : taux d'échec des différentes méthodes en fonction du numéro de l'examen TEP/TDM de la série (pour un total de 76 volumes d'images TEP considérées).....217

Figure VI.4 : A) coupe coronale de la TEP prétraitement du patient 9. En pointillés, les plans axiaux et triaxiaux se coupent au niveau de la lésion 13. B) coupe transverse de la TEP/TDM correspondant au plan pointé par une flèche blanche dans l'image A. Les images C-E) correspondent aux fusions des TEP/TDM des examens 2, 3 et 5 respectivement, dans le même plan de coupe (sans recalcr ici). Images extraites de OWS (*Dosisoft*).219

Figure VI.5 : Evolution du SUVmax (en rouge), du facteur décroissant issu de l'analyse factorielle et normalisé par la valeur de SUVmax(1) (en vert) et des valeurs successivement prises par le SUVadapt pour toutes les lésions étudiées. Le coefficient de détermination présenté pour chaque graphe est la valeur de R2 entre les valeurs successives de SUVmax et celles du facteur décroissant.....**221**

Figure VI.6 : 1ère ligne : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (2 ici) montrant la lésion 1. 2ème ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 1 et images factorielles et facteurs résultants (3ème ligne). L'œsophage (flèche blanche sur la TEP(5)) ne fait pas partie de VOISULTAN analysée. Les valeurs négatives obtenues dans l'image factorielle associée au facteur décroissant (image 2, ligne 2, flèche noire) ne font donc pas partie des résultats interprétés de cette analyse.....**222**

Figure VI.7 : 1ère ligne : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (1 ici) montrant la lésion 3. 2ème ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 3 et images factorielles et facteurs résultants (3ème ligne).....**223**

Figure VI.8 : évolution du volume métabolique de la lésion 3 vue par VMadapt.....**224**

Figure VI.9 : 1ère ligne : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (1 ici) montrant la lésion 7. 2ème ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 7 et images factorielles et facteurs résultants (3ème ligne).....**225**

Figure VI.10 : A : distance du voxel contenant le SUVmax dans les examens TEP(t) $t > 1$, par rapport à la localisation du voxel contenant le SUVmax dans la TEP(1). B : Proportion des **Aki** dans chacun des voxels occupés par le SUVmax de la lésion 7.....**226**

Figure VI.11 : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (1 ici) montrant la lésion 8. 2ème ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 8 et images factorielles et facteurs résultants (3ème ligne).....**227**

Figure VI.12 : coupes transaxiale et coronale de la fusion des données TEP/TDM de l'examen de référence (1 ici) montrant la lésion 5. 2ème ligne : coupes transaxiales de la série d'images utilisées pour l'analyse factorielle de la lésion 5 et images factorielles et facteurs résultants (3ème ligne).....**228**

Figure VI.13 : évolution du VMadapt de la lésion 5 durant le traitement.**229**

Liste des tableaux

Chapitre I

Tableau I.1 : Principales caractéristiques des radionucléides fréquemment utilisés en imagerie TEP.....	19
Tableau I.2 : 3 principales interactions entre rayonnement et matière pour des photons de 511keV.....	21
Tableau I.3 : Caractéristiques physiques de certains des scintillateurs couramment employés en TEP. Données pour des photons de 511 keV. La définition du coefficient d'atténuation linéique μ est donnée au paragraphe I.2.1.6.....	22
Tableau I.4 : distribution physiologique et exemples d'indications des 3 traceurs ayant reçu l'AMM en France.	35
Tableau I.5 : exemple de traceurs étudiés en recherche à l'heure actuelle et applications potentielles.....	38

Chapitre II

Tableau II.1 : critères OMS de la réponse tumorale au traitement. L'évaluation se base sur de l'évolution du Produit des plus grandes Longueurs orthogonales (ΔPdL) de chaque tumeur mesurable ou de cette même grandeur estimée dans le cas des tumeurs non mesurables. Ces mesures se font sur des examens espacés d'au moins 4 semaines.	56
Tableau II.2 : critères RECIST 1.0 de la réponse tumorale au traitement. L'évaluation des lésions cibles se base sur l'évolution de la Somme de leurs plus grandes Longueurs (ΔSdL). Ces mesures se font sur des examens espacés d'au moins 4 semaines (et 6~8 semaines dans le cas d'étude de phase II).	57
Tableau II.3 : critères RECIST 1.1 de la réponse tumorale au traitement. L'évaluation des lésions cibles repose sur l'évolution relative de la Somme de leurs plus grandes Longueurs ($\Delta SdL(\text{relatif})$) ou sur sa variation absolue ($\Delta SdL(\text{absolu})$). Ces mesures se font sur des examens espacés d'au moins 4 semaines.	59
Tableau II.4 : comparaison des critères RECIST 1.0 et Choi. ΔSdL : variation de la somme des plus grandes longueurs des tumeurs mesurables, comme indiqué dans (Therasse et al., 2000). Les différences entre les critères sont indiquées en rouge.	60
Tableau II.5 : facteurs biologiques affectant la mesure du SUV. (Extrait et complété de (Adams et al., 2010) et (Boellaard, 2009))......	63

Tableau II.6 : vue d'ensemble des facteurs techniques pouvant influencer la mesure du SUV. 64

Tableau II.7 : détails des critères d'évaluation de la réponse tumorale selon l'EORTC (Young et al., 1999). Δ SUV : évolution relative du SUV entre l'examen TEP prétraitement et celui effectué pendant/après le traitement. La définition du SUV à utiliser, maximal ou moyen dans le VOI, n'est pas spécifiée dans le protocole. 68

Tableau II.8 : critères PERCIST d'évaluation de la réponse tumorale au traitement. Dans le cas où 5 lésions cibles sont retenues sur la TEP prétraitement, la somme des SULpeak peut être utilisée pour l'évaluation, afin de réduire les sources d'erreurs. Δ SUL : évolution, relative ou absolue (selon le seuil) entre l'examen TEP prétraitement et les examens de suivi, du SUV normalisé par la masse maigre. 72

Tableau II.9 : résultats publiés concernant l'utilisation des méthodes d'analyse de texture sur des examens TEP. Les noms des index utilisés ici (colonne 2) correspondent à ceux que nous avons retenus lors de nos travaux présentés au chapitre V ou ceux choisis par les auteurs quand ceux-ci étaient présentés dans une seule publication. Les résultats donnés pour (El Naqa et al., 2009) (colonnes 3 et 4) sont des valeurs d'AUC des courbes ROC pour les deux cohortes de patients étudiées. Ceux de (Galavis et al., 2010) (colonne 5) correspondent à la valeur absolue la plus élevée de l'écart à la moyenne des valeurs des différents index obtenus sur des images reconstruites en utilisant différents paramètres de reconstruction. Ces valeurs sont déduites des histogrammes présentés dans l'article. Les colonnes 6 et 7 représentent respectivement les Spécificité, Sensibilité et AUC des courbes ROC des index de texture pour identifier un groupe de tumeurs de l'œsophage répondant à la thérapie d'un groupe ne répondant pas (colonne 6) et pour séparer un groupe répondant complètement par rapport à un autre groupe, constitué d'un tout autre type de lésions (colonne 7). La dernière colonne représente l'écart relatif entre deux examens TEP acquis à des intervalles de temps très courts, moyenné sur plusieurs lésions (± 1 écart-type). - : index non étudié dans la publication. ? : valeur non fournie dans l'étude. 83

Tableau II.10 : résultats des index classiques et issus des caractéristiques de premier ordre utilisées dans les articles de (El Naqa et al., 2009; Galavis et al., 2010; Tixier et al., 2012; Tixier et al., 2011). Les résultats donnés pour (El Naqa et al., 2009) (colonne 2 et 3) sont des valeurs d'AUC des courbes ROC pour les 2 cohortes de patients étudiées. Ceux de (Galavis et al., 2010) (colonne 4) correspondent à la valeur absolue la plus élevée de l'écart à la moyenne des valeurs des différents index obtenus sur des images reconstruites en utilisant différents paramètres de reconstruction. Ces valeurs sont déduites des histogrammes présentés dans l'article. Les colonnes 5 et 6 représentent les Spécificité, Sensibilité et AUC des courbes ROC des index, pour distinguer un groupe de tumeurs de l'œsophage répondant à la thérapie d'un groupe ne répondant pas (colonne 5) et pour séparer un groupe répondant complètement par rapport à un autre groupe, constitué d'un tout autre type de lésions (colonne 6). La dernière colonne rapporte l'écart relatif entre deux examens TEP acquis à des intervalles de temps très

courts, moyenné sur plusieurs lésions (± 1 écart-type). - : index non étudié dans la publication.
..... 84

Chapitre III

Tableau III.1 : résumé des méthodes de correction de l'effet de volume partiel présentées dans le chapitre III et ayant été appliquées au moins une fois en oncologie..... 123

Chapitre IV

Tableau IV.1 : valeurs des AUC (1 écart-type) des méthodes comparées pour la séparation des groupes de tumeurs G1 et G2..... 141

Tableau IV.2 : valeurs des AUC (1 écart-type) des méthodes comparées pour la séparation des groupes de tumeurs G3 et G4..... 142

Tableau IV.3: caractéristiques des patients et lésions étudiés..... 145

Tableau IV.4 : valeurs moyenne ± 1 écart-type (et intervalle de valeurs) pour les 8 index métaboliques étudiés pour l'ensemble des lésions sur les TEP (colonne 2) avant traitement et à J14 (colonne 3) (* : valeur de p inférieure à 0,05, caractéristique d'une différence significative entre les colonnes 2 et 3). Moyenne des variations relatives (± 1 écart-type et intervalle de valeurs) des index après 2 semaines de chimiothérapie pour le groupe de tumeurs répondant (colonne 4) et ne répondant pas à la thérapie (colonne 5, * : valeur de p < 0,05 entre les colonnes 4 et 5). 152

Chapitre V

Tableau V.1 : caractéristiques des patients atteints d'un cancer colorectal métastatique étudiés. 166

Tableau V.2 : caractéristiques extraites des matrices de cooccurrences C . μ_i et σ_i correspondent respectivement à la moyenne et à l'écart-type des $C(i,j)$ pour tous les niveaux de gris j , i étant fixé..... 172

Tableau V.3 : caractéristiques issues des matrices de longueurs des séries homogènes. θ correspond au nombre de séries homogènes dans la tumeur. n est le nombre de niveaux de gris utilisés pour la discrétisation et m la longueur de la plus grande série. Une fois calculées pour chacune des 13 directions de parcours, les valeurs obtenues pour chacun des index sont moyennées. 173

Tableau V.4 : caractéristiques issues des matrices de différences de niveaux de gris (N). G correspond au nombre de différents niveaux de gris et E au nombre de voxels contenus dans le volume segmenté.	173
Tableau V.5 : liste des index extraits des différentes matrices de textures étudiées. Nous renvoyons le lecteur au manuscrit de thèse de doctorat de (Thibault, 2009) pour de plus amples informations.	174
Tableau V.6A : correspondances entre les définitions des index issus des matrices C et L présentées et celles utilisées dans les 3 principales publications ((Galavis et al., 2010), (Tixier et al., 2011) et (Tixier et al., 2012)) traitant de l'intérêt clinique de l'analyse de texture en TEP.	175
Tableau V.6B : correspondances entre les définitions des index issus des matrices N et Z présentées et celles utilisées dans les 3 principales publications ((Galavis et al., 2010), (Tixier et al., 2011) et (Tixier et al., 2012)) traitant de l'intérêt clinique de l'analyse de texture en TEP.	176
Tableau V.7 : Reproduction du formulaire présenté par (Tixier et al., 2012) pour définir les index de texture déduits de la matrice de longueur des zones homogènes.	177
Tableau V.8 : coefficients de Pearson obtenus entre les index extraits des matrices de cooccurrence et calculés sur la TEP(PT). En vert les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8.	186
Tableau V.9 : coefficients de Pearson obtenus entre les index extraits des matrices de longueur de séries homogènes. En vert les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu inférieure à 0,1.	187
Tableau V.10 : coefficients de Pearson obtenus entre les index extraits des matrices de longueur de zones homogènes. En vert les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu inférieure à 0,1.	187
Tableau V.11 : coefficients de Pearson obtenus entre les index extraits des matrices de longueur de zones homogènes et ceux extraits des matrices de longueur de séries homogènes. En vert, les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu inférieure à 0,1.	188
Tableau V.12 : coefficient de Pearson obtenus entre les index extraits des matrices de différences de niveaux de gris. En bleu, les coefficients dont la valeur absolue est inférieure à 0,1.	188

Tableau V.13 : coefficients de corrélations de Pearson entre index classiques et index de texture. En vert, les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu, inférieure à 0,1..... **189**

Tableau V.14 : corrélation entre index de texture et de forme et index de texture calculées sur les TEP prétraitement en utilisant 64 niveaux de gris et un ré-échantillonnage commun pour toutes les tumeurs de la cohorte. En vert, les coefficients dont la valeur absolue est supérieure à 0,9, en jaune supérieure à 0,8 et en bleu, inférieure à 0,1. **192**

Chapitre VI

Tableau VI.1 : caractéristiques des patients..... **210**

Tableau VI.2 : informations relatives aux données considérées pour l'analyse de données. G = ganglion. T = tumeur. **216**

Tableau VI.3 : valeurs des différents index mesurés sur la TEP de référence (t_{ref}) et sur la dernière TEP considérée (t_{fin}) pour les 16 lésions étudiées. La troisième colonne de chaque index correspond à l'évolution relative entre ces deux examens (cf équations VI.13). Un tiret signifie un échec de la méthode de segmentation à fournir une segmentation cohérente du volume métabolique de la lésion **218**

Tableau VI.4 : valeurs moyennes des ***Akj*** dans le $VM_{adapt}(t_{référence})$ de chaque lésion (***Ak***) pour chacun des 2 ou 3 facteurs obtenus (colonnes 2 à 4). **220**

Liste des abréviations

2DG : 2-deoxy-D-[14C]-glucose

AFC : Analyse Factorielle des Correspondances

AFSIM : Factorielle de Séquences d'Images Médicales

AMM : Autorisation de Mise sur le Marché

ANSM : Agence Nationale de Sécurité du Médicament

APD : Avalanche PhotoDiode

ATC : Acide TriCarboxilique

ATP : Adénosine TriPhosphate

AUC : Aire sous la courbe

Bq : Becquerel

BSA : Body Surface Area

BW : body weight

CCM : cancer colorectaux métastatiques

CERN : Conseil Européen pour la Recherche Nucléaire

CR : Coefficient de recouvrement

CTV : Clinical Tumor Volume

DAR : Differential Absorption Ratio

DICOM : Digital Imaging and Communications in Medicine

DRO : Digital Reference Object

DSmT : Dezert-Smarandache theory

DVH : Histogrammes Dose Volume

EASL : European Association for Study of the Liver

EC : erreur de classification

EMA : Agence Européenne d'Evaluation des Médicaments

EORTC : European Organisation for Research and Treatment of Cancer

ESKA : Extension of the Simplified Kinetic Analysis

eV : électronvolt

EV : erreur sur le volume

EVP : Effet de Volume Partiel

FCH : 18F-fluorocholine
FDG : 2-deoxy-2-18F-fluoro-D-glucose
FDOPA : 6-18F-fluoro-L-dihydroxyphénylalanine
FDR : fonction de réponse
FET : 18F-fluoroethyl
FFP : Fraction de Faux Positifs
FLAB : Fuzzy Locally Adaptive Bayesian
FLT : 18F-thymidine
FMISO : 18F-Misonidazole
FN : Faux Négatifs
FOV : champs de vue Field of view
FP : Faux Positifs
FVP : Fraction de Vrais Positifs
GATE : Geant4 Application for Emission Tomography
GIST : Gastro-Intestinal Stromal Tumours
GLNur : Gray-Level Non-Uniformity run
GLNUz : Gray-Level Non-Uniformity zone
GLUT : GLUcose Transporter
GMM : Gaussian Mixture Model
GTM Geometric Transfer Matrix
GTV : pour Gross Tumor Volume
HGRE : High Gray-level Run Emphasis
HGZE : High Gray-level Zone Emphasis
HIV : Histogramme Intensité Volume
IHP : International Harmonization Project in lymphoma
IMRT : Intensity Modulated RadioTherapy
IP : Imagerie Paramétrique
IRM : Imagerie par Résonance Magnétique
LBM : Lean Body Mass

LC : Lumped Constant
LCR : liquide céphalo rachidien
LDR : Ligne De Réponse
LGRE : Low Gray-level Run Emphasis
LGZE : Low Gray-level Zone Emphasis
LMH : Largeur à Mi-Hauteur
LR : Lucy-Richardson
LRE : Long Run Emphasis
LRHGE : Long Run High Gray-level Emphasis
LRLGE : Long Run Low Gray-level Emphasis
LRWBD : Lucy-Richardson and Wavelet Based Denoising
LSM : Level Surface Method
LZE : Long Zone Emphasis
LZHGE : Long Zone High Gray-level Emphasis
LZLGE : Long Zone Low Gray-level Emphasis
MAP : Maximum A Posteriori
MET : 18F-métionine
MICAD : Molecular Imaging and Contrast Agent Database
MIP & CA List : Molecular Imaging Probes and Contrast Agent List
ML-EM : Maximum Likelihood Expectation Maximization
MMA : Mutual Multiresolution Analysis
MP : Maladie Progressive
MS : Maladie Stable
NADH : Nicotinamide Adenine Dinucleotide
NCAT : NURBS CARDiac Torso phantom
NCI : National Cancer Institutes :
NEMA : National Electrical manufacturers Association
NSCLC : Non Small Cell Lung Cancer
OMS (WHO) : Organisation Mondiale de la Santé (World Health Organization)

ORL : Oto-Rhino-Laryngé

OS-EM : Ordered Subsets Expectation Maximization

OWS : Oncology WorkStation

PARSLR : PARTially-Segmented Lucy-Richardson

PERCIST : PET Response Criteria in Solid Tumors

pGTM : perturbation based GTM

PL : Potentiel de liaison

RBV : Region Based Voxel-wise)

RC : Réponse Complète

RECIST : Response Evaluation Criteria in Solid Tumors

RGDfK : 18F-galacto-c

RLNU : Run Length Non-Uniformity

ROC : Receiver Operating Characteristic

ROI : région d'intérêt

RP : Run Percentage

RP : Réponse Partielle

RT : radiothérapie

RT-CT : radio-chimiothérapie

rVC : reblurred van Cittert

SBR : Signal to Background Ratio

SiPM : PhotoMultiplificateurs sur Silicium

SKA : Simplified Kinetic Analysis,

SKM : Simplified Kinetic Analysis

SRE : Short Run Emphasis

SRHGE : Short Run High Gray-level Emphasis

SRLGE : Short Run Low Gray-level Emphasis

SULTAN : SUivi Longitudinal en Tomographie par ANalyse factorielle

SUV : Standardized Uptake Value

SZE : Short Zone Emphasis

SZHGE : Short Zone High Gray-level Emphasis
SZLGE : Short Zone Low Gray-level Emphasis
TAC : Courbe Activité Temps
TDM : TomoDensitoMétrie
TEMP : Tomographie d'Emission MonoPhotonique
TEP : Tomographie par Emission de Positons
TLG : total legion glycolysis
UH : Unité Hounsfield
VC : van Cittert
VN : Vrais Négatifs
VOI : Volume d'Intérêt
VP : Vrais Positifs
VPN : Valeur Positive Négative
VPP : Valeur Prédictive Positive
ZLNU : Zone Length Non-Uniformity
ZP : Zone Percentage

Bibliographie

Abgral, R., Le Roux, P.Y., Blanc-Béguin, F., Eugène, T., Querellou, S., Bodet-Milin, C., Kraeber-Bodéré, F., and Salaün, P.Y. (2011). Évaluation de la réponse thérapeutique par tomographie par émission de positons (TEP) au 18fluoro-désoxyglucose (FDG) en oncologie-hématologie. *Médecine Nucléaire* 35, 600-607.

Adams, M.C., Turkington, T.G., Wilson, J.M., and Wong, T.Z. (2010). A systematic review of the factors affecting accuracy of SUV measurements. *AJR Am J Roentgenol* 195, 310-320.

Agostinelli, S., Allison, J., Amako, K., Apostolakis, J., Araujo, H., Arce, P., Asai, M., Axen, D., Banerjee, S., Barrand, G., *et al.* (2003). Geant4—a simulation toolkit. *Nucl Instrum Methods Phys Res A* 506, 250-303.

Alessio, A.M., and Kinahan, P.E. (2008). Application of a spatially variant system model for 3-D whole-body PET image reconstruction. Paper presented at: 5th IEEE Int Symp on Biomed Imaging.

Alkhalaf, K., and Alavi, A. (2008). Quantitative assessment of FDG uptake in brown fat using standardized uptake value and dual-time-point scanning. *Clin Nucl Med* 33, 663-667.

Amadasun, M., and King, R. (1989). Textural features corresponding to textural properties. *IEEE Trans Sys, Man and Cybernetics* 19, 1264-1274.

Assefa, D., Keller, H., Ménard, C., Laperriere, N., Ferrari, R.J., and Yeung, I. (2010). Robust texture features for response monitoring of glioblastoma multiforme on T1-weighted and T2-FLAIR MR images: a preliminary investigation in terms of identification and segmentation. *Med Phys* 37, 1722-1736.

Asselin, M.C., O'Connor, J.P., Boellaard, R., Thacker, N.A., and Jackson, A. (2012). Quantifying heterogeneity in human tumours using MRI and PET. *Eur J Cancer* 48, 447-455.

Avril, N., Bense, S., Ziegler, S.I., Dose, J., Weber, W., Laubenbacher, C., Römer, W., Jänicke, F., and Schwaiger, M. (1997). Breast imaging with fluorine-18-FDG PET: quantitative image analysis. *J Nucl Med* 38, 1186-1191.

Avril, N., Menzel, M., Dose, J., Schelling, M., Weber, W., Jänicke, F., Nathrath, W., and Schwaiger, M. (2001). Glucose metabolism of breast cancer assessed by 18F-FDG PET: histologic and immunohistochemical tissue analysis. *J Nucl Med* 42, 9-16.

Baete, K., Nuyts, J., Van Laere, K., Van Paesschen, W., Ceysens, S., De Ceuninck, L., Gheysens, O., Kelles, A., Van den Eynden, J., Suetens, P., *et al.* (2004). Evaluation of anatomy based reconstruction for partial volume correction in brain FDG-PET. *Neuroimage* 23, 305-317.

Bailey, D.L., and Meikle, S.R. (1994). A convolution-subtraction scatter correction method for 3D PET. *Phys Med Biol* 39, 411-424.

Barbee, D.L., Flynn, R.T., Holden, J.E., Nickles, R.J., and Jeraj, R. (2010). A method for partial volume correction of PET-imaged tumor heterogeneity using expectation maximization with a spatially varying point spread function. *Phys Med Biol* 55, 221-236.

Bataille, F., Comtat, C., Jan, S., Sureau, F.C., and Trebossen, R. (2007). Brain PET partial-volume compensation using blurred anatomical labels. Paper presented at: IEEE Trans Nucl Sci.

Behrendt, F.F., Temur, Y., Verburg, F.A., Palmowski, M., Krohn, T., Pietsch, H., Kuhl, C.K., and Mottaghy, F.M. (2012). PET/CT in lung cancer: Influence of contrast medium on quantitative and clinical assessment. *Eur Radiol* 22, 2458-2464.

Bellotti, R., De Carlo, F., Tangaro, S., Gargano, G., Maggipinto, G., Castellano, M., Massafra, R., Cascio, D., Fauci, F., Magro, R., *et al.* (2006). A completely automated CAD system for mass detection in a large mammographic database. *Med Phys* 33, 3066-3075.

Ben-Haim, S., and Ell, P. (2009). 18F-FDG PET and PET/CT in the evaluation of cancer treatment response. *J Nucl Med* 50, 88-99.

Benali, H., Buvat, I., Frouin, F., Bazin, J.P., and Di Paola, R. (1993). A statistical model for the determination of the optimal metric in factor analysis of medical image sequences (FAMIS). *Phys Med Biol* 38, 1065-1080.

Benali, H., Buvat, I., Frouin, F., Bazin, J.P., and Di Paola, R. (1994). Foundations of factor-analysis of medical image sequences - a unified approach and some practical implications. *Image and Vision Computing* 12, 375-385.

Benjamin, R.S., Choi, H., Macapinlac, H.A., Burgess, M.A., Patel, S.R., Chen, L.L., Podoloff, D.A., and Charnsangavej, C. (2007). We should desist using RECIST, at least in GIST. *J Clin Oncol* 25, 1760-1764.

Bentzen, S.M. (2005). Theragnostic imaging for radiation oncology: dose-painting by numbers. *Lancet Oncol* 6, 112-117.

Benz, M.R., Allen-Auerbach, M.S., Eilber, F.C., Chen, H.J., Dry, S., Phelps, M.E., Czernin, J., and Weber, W.A. (2008). Combined assessment of metabolic and volumetric changes for assessment of tumor response in patients with soft-tissue sarcomas. *J Nucl Med* 49, 1579-1584.

Beyer, T., Czernin, J., and Freudenberg, L.S. (2011). Variations in clinical PET/CT operations: results of an international survey of active PET/CT users. *J Nucl Med* 52, 303-310.

Beyer, T., Townsend, D.W., Brun, T., Kinahan, P.E., Charron, M., Roddy, R., Jerin, J., Young, J., Byars, L., and Nutt, R. (2000). A combined PET/CT scanner for clinical oncology. *J Nucl Med* 41, 1369-1379.

Bland, J.M., and Altman, D.G. (1995). Multiple significance tests: the Bonferroni method. *Brit Med J* 310, 170-171.

Boellaard, R. (2009). Standards for PET image acquisition and quantitative data analysis. *J Nucl Med* 50 *Suppl* 1, 11S-20S.

Boellaard, R. (2011). Need for standardization of 18F-FDG PET/CT for treatment response assessments. *J Nucl Med* 52 *Suppl* 2, 93S-100S.

Boellaard, R. (2012). Mutatis mutandis: harmonize the standard! *J Nucl Med* 53, 1-3.

Boellaard, R., O'Doherty, M.J., Weber, W.A., Mottaghy, F.M., Lonsdale, M.N., Stroobants, S.G., Oyen, W.J., Kotzerke, J., Hoekstra, O.S., Pruim, J., *et al.* (2010). FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging* 37, 181-200.

Boellaard, R., Oyen, W.J., Hoekstra, C.J., Hoekstra, O.S., Visser, E.P., Willemsen, A.T., Arends, B., Verzijlbergen, F.J., Zijlstra, J., Paans, A.M., *et al.* (2008). The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. *Eur J Nucl Med Mol Imaging* 35, 2320-2333.

- Bonnerot, V., Charpentier, A., Frouin, F., Kalifa, C., Vanel, D., and Di Paola, R. (1992). Factor analysis of dynamic magnetic resonance imaging in predicting the response of osteosarcoma to chemotherapy. *Invest Radiol* 27, 847-855.
- Boussion, N., Cheze Le Rest, C., Hatt, M., and Visvikis, D. (2009). Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging. *Eur J Nucl Med Mol Imaging* 36, 1064-1075.
- Boussion, N., Hatt, M., Lamare, F., Bizais, Y., Turzo, A., Cheze-Le Rest, C., and Visvikis, D. (2006). A multiresolution image based approach for correction of partial volume effects in emission tomography. *Phys Med Biol* 51, 1857-1876.
- Bowsher, J.E., Johnson, V.E., Turkington, T.G., Jaszczak, R.J., Floyd, C.R., and Coleman, R.E. (1996). Bayesian reconstruction and use of anatomical a priori information for emission tomography. *IEEE Trans Med Imaging* 15, 673-686.
- Brenner, W., Vernon, C., Muzi, M., Mankoff, D.A., Link, J.M., Conrad, E.U., and Eary, J.F. (2004). Comparison of different quantitative approaches to 18F-fluoride PET scans. *J Nucl Med* 45, 1493-1500.
- Brooks, F.J., and Grigsby, P.W. (2011). Current measures of metabolic heterogeneity within cervical cancer do not predict disease outcome. *Radiat Oncol* 6, 69.
- Bruix, J., Sherman, M., Llovet, J.M., Beaugrand, M., Lencioni, R., Burroughs, A.K., Christensen, E., Pagliaro, L., Colombo, M., Rodés, J., *et al.* (2001). Clinical management of hepatocellular carcinoma. Conclusions of the Barcelona-2000 EASL conference. European Association for the Study of the Liver. *J Hepatol* 35, 421-430.
- Budinger, T.F. (1983). Time-of-flight positron emission tomography: status relative to conventional PET. *J Nucl Med* 24, 73-78.
- Bural, G., Torigian, D.A., Houseni, M., Basu, S., Srinivas, S., and Alavi, A. (2009). Tumor metabolism measured by partial volume corrected standardized uptake value varies considerably in primary and metastatic sites in patients with lung cancer. A new observation. *Hell J Nucl Med* 12, 218-222.
- Burger, C., Goerres, G., Schoenes, S., Buck, A., Lonn, A.H., and Von Schulthess, G.K. (2002). PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients. *Eur J Nucl Med Mol Imaging* 29, 922-927.
- Buvat, I. (2002). A non-parametric bootstrap approach for analysing the statistical properties of SPECT and PET images. *Phys Med Biol* 47, 1761-1775.
- Buvat, I. (2007). Les limites du SUV. *Médecine Nucléaire* 31, 165-172.
- Buvat, I., Benali, H., and Di Paola, R. (1998). Statistical distribution of factors and factor images in factor analysis of medical image sequences. *Phys Med Biol* 43, 1695-1711.
- Buvat, I., Necib, H., Garcia, C., Wagner, A., Vanderlinden, B., Emonts, P., Hendlisz, A., and Flamen, P. (2012). Lesion-based detection of early chemosensitivity using serial static FDG PET/CT in metastatic colorectal cancer. *Eur J Nucl Med Mol Imaging* 39, 1628-1634.
- Byrne, M.J., and Nowak, A.K. (2004). Modified RECIST criteria for assessment of response in malignant pleural mesothelioma. *Ann Oncol* 15, 257-260.
- Carasso, A. (1999). Linear and nonlinear image deblurring: a documented study. *SIAM J Numer Anal* 36, 1659-1689.

Casasnovas, R.O., Meignan, M., Berriolo-Riedinger, A., Bardet, S., Julian, A., Thieblemont, C., Vera, P., Bologna, S., Brière, J., Jais, J.P., *et al.* (2011). SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma. *Blood* 118, 37-43.

Casey, M., and Nutt, R. (1986). A multycrystal two dimensional BGO detector system for positron emission tomography. *IEEE Trans Nucl Sci* 33, 460-463.

Caussinus, H. (1986). *Models and uses of principal component analysis* (Leiden: DSWO press).

Cazaentre, T., Morschhauser, F., Vermandel, M., Betrouni, N., Prangère, T., Steinling, M., and Huglo, D. (2010). Pre-therapy 18F-FDG PET quantitative parameters help in predicting the response to radioimmunotherapy in non-Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging* 37, 494-504.

Chan, D.A., and Giaccia, A.J. (2007). Hypoxia, gene expression, and metastasis. *Cancer Metastasis Rev* 26, 333-339.

Chan, T. (2012). Computerized method for automatic evaluation of lean body mass from PET/CT: comparison with predictive equations. *J Nucl Med* 53, 130-137.

Cheebsumon, P., van Velden, F.H., Yaqub, M., Frings, V., de Langen, A.J., Hoekstra, O.S., Lammertsma, A.A., and Boellaard, R. (2011a). Effects of image characteristics on performance of tumor delineation methods: a test-retest assessment. *J Nucl Med* 52, 1550-1558.

Cheebsumon, P., Velasquez, L.M., Hoekstra, C.J., Hayes, W., Kloet, R.W., Hoetjes, N.J., Smit, E.F., Hoekstra, O.S., Lammertsma, A.A., and Boellaard, R. (2011b). Measuring response to therapy using FDG PET: semi-quantitative and full kinetic analysis. *Eur J Nucl Med Mol Imaging* 38, 832-842.

Cheebsumon, P., Yaqub, M., van Velden, F.H., Hoekstra, O.S., Lammertsma, A.A., and Boellaard, R. (2011c). Impact of [¹⁸F]FDG PET imaging parameters on automatic tumour delineation: need for improved tumour delineation methodology. *Eur J Nucl Med Mol Imaging* 38, 2136-2144.

Chen, C.H., Muzic, R.F., Nelson, A.D., and Adler, L.P. (1999). Simultaneous recovery of size and radioactivity concentration of small spheroids with PET data. *J Nucl Med* 40, 118-130.

Chen, K., Bandy, D., Reiman, E., Huang, S.C., Lawson, M., Feng, D., Yun, L.S., and Palant, A. (1998). Noninvasive quantification of the cerebral metabolic rate for glucose using positron emission tomography, 18F-fluoro-2-deoxyglucose, the Patlak method, and an image-derived input function. *J Cereb Blood Flow Metab* 18, 716-723.

Cheng-Liao, J., and Qi, J. (2011). PET image reconstruction with anatomical edge guided level set prior. *Phys Med Biol* 56, 6899-6918.

Cherry, S.R., Dahlbom, M., and Cherry, S. (2006). *PET: Physics, Instrumentation, and Scanners* (Springer).

Cherry, S.R., Dahlbom, M., and Hoffman, E.J. (1991). 3D PET using a conventional multislice tomograph without septa. *J Comput Assist Tomogr* 15, 655-668.

Cheson, B.D., Pfistner, B., Juweid, M.E., Gascoyne, R.D., Specht, L., Horning, S.J., Coiffier, B., Fisher, R.I., Hagenbeek, A., Zucca, E., *et al.* (2007). Revised response criteria for malignant lymphoma. *J Clin Oncol* 25, 579-586.

Choi, H. (2005). Critical issues in response evaluation on computed tomography: lessons from the gastrointestinal stromal tumor model. *Curr Oncol Rep* 7, 307-311.

Choi, H., Charnsangavej, C., de Castro Faria, S., Tamm, E.P., Benjamin, R.S., Johnson, M.M., Macapinlac, H.A., and Podoloff, D.A. (2004). CT evaluation of the response of gastrointestinal stromal tumors after imatinib mesylate treatment: a quantitative analysis correlated with FDG PET findings. *AJR Am J Roentgenol* 183, 1619-1628.

Choi, H., Charnsangavej, C., Faria, S.C., Macapinlac, H.A., Burgess, M.A., Patel, S.R., Chen, L.L., Podoloff, D.A., and Benjamin, R.S. (2007). Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J Clin Oncol* 25, 1753-1759.

Colombino, P., Fiscella, B., and Trossi, L. (1965). Study of positronium in water and ice from 22 to -144 °C by annihilation quanta measurements. *Il Nuovo Cimento xxxviii(2)*, 707-723.

Comtat, C., Kinahan, P.E., Fessler, J.A., Beyer, T., Townsend, D.W., Defrise, M., and Michel, C. (2002). Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels. *Phys Med Biol* 47, 1-20.

Contractor, K.B., and Aboagye, E.O. (2009). Monitoring predominantly cytostatic treatment response with 18F-FDG PET. *J Nucl Med* 50 *Suppl 1*, 97S-105S.

Coulot, J., Magnier, F., Chauchat, P., Guilabert, N., Ricard, M., Dreuil, S., and Lumbroso, J. (2011). Comparaison de la dose efficace en TEP/TDM et en TDM. *Médecine Nucléaire* 35, 156-162.

Crilly, P.B. (1991). A quantitative evaluation of various iterative deconvolution algorithms. *IEEE Trans Instrum Meas* 40, 558-562.

Crippa, F., Gavazzi, C., Bozzetti, F., Chiesa, C., Pascali, C., Bogni, A., De Sanctis, V., Decise, D., Schiavini, M., Cucchetti, G., *et al.* (1997). The influence of blood glucose levels on [18F]fluorodeoxyglucose (FDG) uptake in cancer: a PET study in liver metastases from colorectal carcinomas. *Tumori* 83, 748-752.

David, S., Visvikis, D., Quellec, G., Cheze le Rest, C., Fernandez, P., Allard, M., Roux, C., and Hatt, M. (2012). Image change detection using paradoxical theory for patient follow-up quantitation and therapy assessment. *IEEE Trans Med Imaging* 31, 1743-1753.

de Geus-Oei, L.F., van der Heijden, H.F., Visser, E.P., Hermsen, R., van Hoorn, B.A., Timmer-Bonte, J.N., Willemsen, A.T., Pruijm, J., Corstens, F.H., Krabbe, P.F., *et al.* (2007). Chemotherapy response evaluation with 18F-FDG PET in patients with non-small cell lung cancer. *J Nucl Med* 48, 1592-1598.

de Langen, A.J., Vincent, A., Velasquez, L.M., van Tinteren, H., Boellaard, R., Shankar, L.K., Boers, M., Smit, E.F., Stroobants, S., Weber, W.A., *et al.* (2012). Repeatability of 18F-FDG uptake measurements in tumors: a metaanalysis. *J Nucl Med* 53, 701-708.

Defrise, M., Townsend, D.W., Bailey, D., Geissbuhler, A., Michel, C., and Jones, T. (1991). A normalization technique for 3D PET data. *Phys Med Biol* 36, 939-952.

Derenzo, S., Moses, W., Huesman, R., and Budinger, T. (1993). Critical instrumentation issues for <2 mm resolution, high sensitivity brain PET. In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in Brain PET* (Elsevier).

Di Paola, R., Bazin, J.P., Aubry, F., Aurengo, A., Cavailloles, F., Herry, J.Y., and Kahn, E. (1982). Handling of dynamic sequences in nuclear medicine. *IEEE Trans Nucl Sci* 29, 1310-1321.

Dimitrakopoulou-Strauss, A., Strauss, L.G., and Burger, C. (2001a). Quantitative PET studies in pretreated melanoma patients: a comparison of 6-[18F]fluoro-L-dopa with 18F-FDG and (15)O-water using compartment and noncompartment analysis. *J Nucl Med* 42, 248-256.

Dimitrakopoulou-Strauss, A., Strauss, L.G., Schwarzbach, M., Burger, C., Heichel, T., Willeke, F., Mechtersheimer, G., and Lehnert, T. (2001b). Dynamic PET 18F-FDG studies in patients with primary and recurrent soft-tissue sarcomas: impact on diagnosis and correlation with grading. *J Nucl Med* 42, 713-720.

Drzymala, R.E., Mohan, R., Brewster, L., Chu, J., Goitein, M., Harms, W., and Urie, M. (1991). Dose-volume histograms. *Int J Radiat Oncol Biol Phys* 21, 71-78.

Du Bois, D., and Du Bois, E.F. (1916). A formula to estimate the approximate surface area if height and weight be known. *Arch Int Med* 17, 863-871.

Du, Y., Tsui, B.M., and Frey, E.C. (2005). Partial volume effect compensation for quantitative brain SPECT imaging. *IEEE Trans Med Imaging* 24, 969-976.

Du, Y., Tsui, B.M., and Frey, E.C. (2006). Model-based compensation for quantitative 123I brain SPECT imaging. *Phys Med Biol* 51, 1269-1282.

Dudeck, O., Zeile, M., Reichardt, P., and Pink, D. (2011). Comparison of RECIST and Choi criteria for computed tomographic response evaluation in patients with advanced gastrointestinal stromal tumor treated with sunitinib. *Ann Oncol* 22, 1828-1833.

Eary, J.F., O'Sullivan, F., O'Sullivan, J., and Conrad, E.U. (2008). Spatial heterogeneity in sarcoma 18F-FDG uptake as a predictor of patient outcome. *J Nucl Med* 49, 1973-1979.

Eary, J.F., O'Sullivan, F., Powitan, Y., Chandhury, K.R., Vernon, C., Bruckner, J.D., and Conrad, E.U. (2002). Sarcoma tumor FDG uptake measured by PET and patient outcome: a retrospective analysis. *Eur J Nucl Med Mol Imaging* 29, 1149-1154.

Edet-Sanson, A., Dubray, B., Doyeux, K., Back, A., Hapdey, S., Modzelewski, R., Bohn, P., Gardin, I., and Vera, P. (2012). Serial assessment of FDG-PET FDG uptake and functional volume during radiotherapy (RT) in patients with non-small cell lung cancer (NSCLC). *Radiother Oncol* 102, 251-257.

Eisenhauer, E.A., Therasse, P., Bogaerts, J., Schwartz, L.H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., *et al.* (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 45, 228-247.

El Naqa, I., Grigsby, P., Apte, A., Kidd, E., Donnelly, E., Khullar, D., Chaudhari, S., Yang, D., Schmitt, M., Laforest, R., *et al.* (2009). Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit* 42, 1162-1171.

Engle, J., and Kadrmas, D. (2007). Modeling the spatially-variant point spread function in a fast projector for improved fully-3D PET reconstruction. In *J Nucl Med*, meeting abstract, pp. 417P.

Erdi, Y.E., Nehmeh, S.A., Pan, T., Pevsner, A., Rosenzweig, K.E., Mageras, G., Yorke, E.D., Schoder, H., Hsiao, W., Squire, O.D., *et al.* (2004). The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. *J Nucl Med* 45, 1287-1292.

Erselcan, T., Turgut, B., Dogan, D., and Ozdemir, S. (2002). Lean body mass-based standardized uptake value, derived from a predictive equation, might be misleading in PET studies. *Eur J Nucl Med Mol Imaging* 29, 1630-1638.

Fahey, F.H. (2002). Data acquisition in PET imaging. *J Nucl Med Technol* 30, 39-49.

Forner, A., Ayuso, C., Varela, M., Rimola, J., Hessheimer, A.J., de Lope, C.R., Reig, M., Bianchi, L., Llovet, J.M., and Bruix, J. (2009). Evaluation of tumor response after locoregional therapies in hepatocellular carcinoma: are response evaluation criteria in solid tumors reliable? *Cancer* 115, 616-623.

Freedman, N.M., Sundaram, S.K., Kurdziel, K., Carrasquillo, J.A., Whatley, M., Carson, J.M., Sellers, D., Libutti, S.K., Yang, J.C., and Bacharach, S.L. (2003). Comparison of SUV and Patlak slope for monitoring of cancer therapy using serial PET scans. *Eur J Nucl Med Mol Imaging* 30, 46-53.

Frouin, F., De Cesare, A., Bouchareb, Y., Todd-Pokropek, A., and Herment, A. (1999). Spatial regularization applied to factor analysis of medical image sequences (FAMIS). *Phys Med Biol* 44, 2289-2306.

Frouin, F., Merlet, P., Bouchareb, Y., Frouin, V., Dubois-Randé, J.L., De Cesare, A., Herment, A., Syrota, A., and Todd-Pokropek, A. (2001). Validation of myocardial perfusion reserve measurements using regularized factor images of H(2)(15)O dynamic PET scans. *J Nucl Med* 42, 1737-1746.

Frouin, V., Comtat, C., Reilhac, A., and Grégoire, M.C. (2002). Correction of partial-volume effect for PET striatal imaging: fast implementation and study of robustness. *J Nucl Med* 43, 1715-1726.

Galavis, P.E., Hollensen, C., Jallow, N., Paliwal, B., and Jeraj, R. (2010). Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol* 49, 1012-1016.

Gallamini, A., Fiore, F., Sorasio, R., and Meignan, M. (2009). Interim positron emission tomography scan in Hodgkin lymphoma: definitions, interpretation rules, and clinical validation. *Leuk Lymphoma* 50, 1761-1764.

Gallivanone, F., Stefano, A., Grosso, E., Canevari, C., Gianolli, L., Messa, C., Gilardi, M., and Castiglioni, I. (2011). PVE correction in PET-CT whole-body oncological studies from PVE-affected images. *IEEE Trans Nucl Sci* 58, 736-747.

Gatenby, R.A., and Gillies, R.J. (2004). Why do cancers have high aerobic glycolysis? *Nat Rev Cancer* 4, 891-899.

Geworski, L., Knoop, B.O., de Cabrejas, M.L., Knapp, W.H., and Munz, D.L. (2000). Recovery correction for quantitation in emission tomography: a feasibility study. *Eur J Nucl Med* 27, 161-169.

Geworski, L., Knoop, B.O., de Wit, M., Ivancević, V., Bares, R., and Munz, D.L. (2002). Multicenter comparison of calibration and cross calibration of PET scanners. *J Nucl Med* 43, 635-639.

Giovacchini, G., Toczek, M.T., Bonwetsch, R., Bagic, A., Lang, L., Fraser, C., Reeves-Tyer, P., Herscovitch, P., Eckelman, W.C., Carson, R.E., *et al.* (2005). 5-HT 1A receptors are reduced in temporal lobe epilepsy after partial-volume correction. *J Nucl Med* 46, 1128-1135.

Graham, M.M., Peterson, L.M., and Hayward, R.M. (2000). Comparison of simplified quantitative analyses of FDG uptake. *Nucl Med Biol* 27, 647-655.

Green, D., and Swets, J. (1966). Signal detection theory and psychophysics (John Wiley and Sons Inc.).

Grégoire, V., Haustermans, K., Geets, X., Roels, S., and Lonneux, M. (2007). PET-based treatment planning in radiotherapy: a new standard? *J Nucl Med* 48 *Suppl 1*, 68S-77S.

Grégoire, V., Jeraj, R., Lee, J.A., and O'Sullivan, B. (2012). Radiotherapy for head and neck tumours in 2012 and beyond: conformal, tailored, and adaptive? *Lancet Oncol* 13, e292-300.

Hadi, M., Bacharach, S.L., Whatley, M., Libutti, S.K., Straus, S.E., Rao, V.K., Wesley, R., and Carrasquillo, J.A. (2008). Glucose and insulin variations in patients during the time course of a FDG-PET study and implications for the "glucose-corrected" SUV. *Nucl Med Biol* 35, 441-445.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646-674.

Hanley, J.A., and McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiol* 143, 29-36.

Hanley, J.A., and McNeil, B.J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 148, 839-843.

Hapdey, S., Buvat, I., Carson, J.M., Carrasquillo, J.A., Whatley, M., and Bacharach, S.L. (2011). Searching for alternatives to full kinetic analysis in 18F-FDG PET: an extension of the simplified kinetic analysis method. *J Nucl Med* 52, 634-641.

Haralick, R.M., Shanmugam, K., and Dinstein, I.H. (1973). Textural features for image classification. *IEEE Trans Sys, Man and Cybernetics* SMC-3, 610-621.

Harrison, R.L., Vannoy, S.D., Kohlmyer, S., Sossi, V., and Lewellen, T.K. (1994). The effect of scatter on quantitation in positron volume imaging of the thorax. Paper presented at: IEEE Nucl Sci Symp Med Imaging Conf.

Harvie, M.N., Campbell, I.T., Thatcher, N., and Baidam, A. (2003). Changes in body composition in men and women with advanced nonsmall cell lung cancer (NSCLC) undergoing chemotherapy. *J Hum Nutr Diet* 16, 323-326.

Hasegawa, B.H., Gingold, E.L., Reilly, S.M., Liew, S.C., and Cann, C.E. (1990). Description of a simultaneous emission-transmission CT system. *Proceedings of SPIE* 1231, 50-60.

Hatt, M., Cheze le Rest, C., Turzo, A., Roux, C., and Visvikis, D. (2009). A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. *IEEE Trans Med Imaging* 28, 881-893.

Hatt, M., Le Pogam, A., Visvikis, D., Pradier, O., and Cheze Le Rest, C. (2012a). Impact of partial-volume effect correction on the predictive and prognostic value of baseline 18F-FDG PET images in esophageal cancer. *J Nucl Med* 53, 12-20.

Hatt, M., van Stiphout, R., Le Pogam, A., Lammering, G., Visvikis, D., and Lambin, P. (2012b). 18F-FDG PET derived indices and impact of partial volume effects correction for prediction of pathological response of locally advanced rectal cancer tumor. *J Nucl Med*, meeting abstract 53, 1228.

Hatt, M., van Stiphout, R., le Pogam, A., Lammering, G., Visvikis, D., and Lambin, P. (2012c). Early prediction of pathological response in locally advanced rectal cancer based on sequential (18)F-FDG PET. *Acta Oncol Ahead of print*.

Hatt, M., Visvikis, D., Albarghach, N.M., Tixier, F., Pradier, O., and Cheze-le Rest, C. (2011). Prognostic value of 18F-FDG PET image-based parameters in oesophageal cancer and impact of tumour delineation methodology. *Eur J Nucl Med Mol Imaging* 38, 1191-1202.

Haycock, G.B., Schwartz, G.J., and Wisotsky, D.H. (1978). Geometric method for measuring body surface area: a height-weight formula validated in infants, children, and adults. *J Pediatr* 93, 62-66.

Hendlisz, A., Golfinopoulos, V., Garcia, C., Covas, A., Emonts, P., Ameye, L., Paesmans, M., Deleporte, A., Machiels, G., Toussaint, E., *et al.* (2011). Serial FDG-PET/CT for early outcome prediction in patients with metastatic colorectal cancer undergoing chemotherapy. *Ann Oncol* 23, 1687-1693.

Henriksson, E., Kjellen, E., Wahlberg, P., Ohlsson, T., Wennerberg, J., and Brun, E. (2007). 2-Deoxy-2-[18F] fluoro-D-glucose uptake and correlation to intratumoral heterogeneity. *Anticancer Res* 27, 2155-2159.

Hickeson, M., Yun, M., Matthies, A., Zhuang, H., Adam, L.E., Lacorte, L., and Alavi, A. (2002). Use of a corrected standardized uptake value based on the lesion size on CT permits accurate characterization of lung nodules on FDG-PET. *Eur J Nucl Med Mol Imaging* 29, 1639-1647.

Hicks, R.J. (2009). Role of 18F-FDG PET in assessment of response in non-small cell lung cancer. *J Nucl Med* 50 *Suppl 1*, 31S-42S.

Higashi, K., Ueda, Y., Arisaka, Y., Sakuma, T., Nambu, Y., Oguchi, M., Seki, H., Taki, S., Tonami, H., and Yamamoto, I. (2002). 18F-FDG uptake as a biologic prognostic factor for recurrence in patients with surgically resected non-small cell lung cancer. *J Nucl Med* 43, 39-45.

Hoekstra, C.J., Hoekstra, O.S., Stroobants, S.G., Vansteenkiste, J., Nuyts, J., Smit, E.F., Boers, M., Twisk, J.W., and Lammertsma, A.A. (2002). Methods to monitor response to chemotherapy in non-small cell lung cancer with 18F-FDG PET. *J Nucl Med* 43, 1304-1309.

Hoekstra, C.J., Paglianiti, I., Hoekstra, O.S., Smit, E.F., Postmus, P.E., Teule, G.J., and Lammertsma, A.A. (2000). Monitoring response to therapy in cancer using [18F]-2-fluoro-2-deoxy-D-glucose and positron emission tomography: an overview of different analytical methods. *Eur J Nucl Med* 27, 731-743.

Hoetjes, N.J., van Velden, F.H., Hoekstra, O.S., Hoekstra, C.J., Krak, N.C., Lammertsma, A.A., and Boellaard, R. (2010). Partial volume correction strategies for quantitative FDG PET in oncology. *Eur J Nucl Med Mol Imaging* 37, 1679-1687.

Hoffman, E.J., Guerrero, T.M., Germano, G., Digby, W.M., and Dahlbom, M. (1989). PET system calibrations and corrections for quantitative and spatially accurate images. *IEEE Trans Nucl Sci* 36, 1108-1112.

Hoffman, E.J., Huang, S.C., and Phelps, M.E. (1979). Quantitation in positron emission computed tomography: 1. Effect of object size. *J Comput Assist Tomogr* 3, 299-308.

Hofman, M.S., and Hicks, R.J. (2010). Restaging: should we persist without pattern recognition? *J Nucl Med* 51, 1830-1832.

Huang, B., Chan, T., Kwong, D.L., Chan, W.K., and Khong, P.L. (2012). Nasopharyngeal Carcinoma: Investigation of Intratumoral Heterogeneity With FDG PET/CT. *AJR Am J Roentgenol* 199, 169-174.

Huang, S.C. (2000). Anatomy of SUV. Standardized uptake value. *Nucl Med Biol* 27, 643-646.

Hudson, H.M., and Larkin, R.S. (1994). Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* 13, 601-609.

Hunter, G.J., Hamberg, L.M., Alpert, N.M., Choi, N.C., and Fischman, A.J. (1996). Simplified measurement of deoxyglucose utilization rate. *J Nucl Med* 37, 950-955.

Hutton, B.F., Olsson, A., Som, S., Erlandsson, K., and Braun, M. (2006). Reducing the influence of spatial resolution to improve quantitative accuracy in emission tomography: A comparison of potential strategies. *Nucl Instrum Methods Physics Res A* 569, 462-466.

Hyun, S.H., Choi, J.Y., Shim, Y.M., Kim, K., Lee, S.J., Cho, Y.S., Lee, J.Y., Lee, K.H., and Kim, B.T. (2010). Prognostic value of metabolic tumor volume measured by 18F-fluorodeoxyglucose positron emission tomography in patients with esophageal carcinoma. *Ann Surg Oncol* 17, 115-122.

Izquierdo-Garcia, D., Davies, J.R., Graves, M.J., Rudd, J.H., Gillard, J.H., Weissberg, P.L., Fryer, T.D., and Warburton, E.A. (2009). Comparison of methods for magnetic resonance-guided [18-F]fluorodeoxyglucose positron emission tomography in human carotid arteries: reproducibility, partial volume correction, and correlation between methods. *Stroke* 40, 86-93.

Jan, S., Santin, G., Strul, D., Staelens, S., Assié, K., Autret, D., Avner, S., Barbier, R., Bardiès, M., Bloomfield, P.M., *et al.* (2004). GATE: a simulation toolkit for PET and SPECT. *Phys Med Biol* 49, 4543-4561.

Janier, M.F., Mazzadi, A.N., Lionnet, M., Frouin, F., André-Fouët, X., Cinotti, L., Revel, D., and Croisille, P. (2002). Factor analysis of medical image sequences improves evaluation of first-pass MR imaging acquisitions for myocardial perfusion. *Acad Radiol* 9, 26-39.

Jaskowiak, C.J., Bianco, J.A., Perlman, S.B., and Fine, J.P. (2005). Influence of reconstruction iterations on 18F-FDG PET/CT standardized uptake values. *J Nucl Med* 46, 424-428.

Jentzen, W., Weise, R., Kupferschläger, J., Freudenberg, L., Brandau, W., Bares, R., Burchert, W., and Bockisch, A. (2008). Iodine-124 PET dosimetry in differentiated thyroid cancer: recovery coefficient in 2D and 3D modes for PET(/CT) systems. *Eur J Nucl Med Mol Imaging* 35, 611-623.

Juweid, M.E., and Cheson, B.D. (2006). Positron-emission tomography and assessment of cancer therapy. *N Engl J Med* 354, 496-507.

Juweid, M.E., Stroobants, S., Hoekstra, O.S., Mottaghy, F.M., Dietlein, M., Guermazi, A., Wiseman, G.A., Kostakoglu, L., Scheidhauer, K., Buck, A., *et al.* (2007). Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. *J Clin Oncol* 25, 571-578.

Kalff, V., Duong, C., Drummond, E.G., Matthews, J.P., and Hicks, R.J. (2006). Findings on 18F-FDG PET scans after neoadjuvant chemoradiation provides prognostic stratification in patients with locally advanced rectal carcinoma subsequently treated by radical surgery. *J Nucl Med* 47, 14-22.

Kenney, J.M., Marinelli, L.D., and Woodard, H.Q. (1941). Tracer studies with radioactive phosphorus in malignant neoplastic disease. *Radiology* 37, 683-690.

- Kessler, R.M., Ellis, J.R., and Eden, M. (1984). Analysis of emission tomographic scan data: limitations imposed by resolution and background. *J Comput Assist Tomogr* 8, 514-522.
- Keyes, J.W. (1995). SUV: standard uptake or silly useless value? *J Nucl Med* 36, 1836-1839.
- Kidd, E.A., and Grigsby, P.W. (2008). Intratumoral metabolic heterogeneity of cervical cancer. *Clin Cancer Res* 14, 5236-5241.
- Kim, C.K., and Gupta, N.C. (1996). Dependency of standardized uptake values of fluorine-18 fluorodeoxyglucose on body size: comparison of body surface area correction and lean body mass correction. *Nucl Med Commun* 17, 890-894.
- Kim, C.K., Gupta, N.C., Chandramouli, B., and Alavi, A. (1994). Standardized uptake values of FDG: body surface area correction is preferable to body weight correction. *J Nucl Med* 35, 164-167.
- Kinahan, P., Pierce, L., Elston, B., Clunie, D., and Nelson, D. (2012). An FDG-PET/CT digital reference object for testing ROI and SUV calculations. *J Nucl Med*, meeting abstract 53, 607.
- Kinahan, P.E., and Fletcher, J.W. (2010). Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. *Semin Ultrasound CT MR* 31, 496-505.
- Kinahan, P.E., Townsend, D.W., Beyer, T., and Sashin, D. (1998). Attenuation correction for a combined 3D PET/CT scanner. *Med Phys* 25, 2046-2053.
- Kirov, A.S., Piao, J.Z., and Schmidlein, C.R. (2008). Partial volume effect correction in PET using regularized iterative deconvolution with variance control based on local topology. *Phys Med Biol* 53, 2577-2591.
- Klein, R., Beanlands, R.S., Wassenaar, R.W., Thorn, S.L., Lamoureux, M., DaSilva, J.N., Adler, A., and deKemp, R.A. (2010). Kinetic model-based factor analysis of dynamic sequences for 82-rubidium cardiac positron emission tomography. *Med Phys* 37, 3995-4010.
- Krak, N.C., Boellaard, R., Hoekstra, O.S., Twisk, J.W., Hoekstra, C.J., and Lammertsma, A.A. (2005). Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *Eur J Nucl Med Mol Imaging* 32, 294-301.
- Krak, N.C., van der Hoeven, J.J., Hoekstra, O.S., Twisk, J.W., van der Wall, E., and Lammertsma, A.A. (2003). Measuring [(18)F]FDG uptake in breast cancer during chemotherapy: comparison of analytical methods. *Eur J Nucl Med Mol Imaging* 30, 674-681.
- Krebs, H.A., and Johnson, W.A. (1937). Metabolism of ketonic acids in animal tissues. *Biochem J* 31, 645-660.
- Krebs, H.A., and Johnson, W.A. (1980). The role of citric acid in intermediate metabolism in animal tissues. *FEBS Lett* 117 Suppl, K1-10.
- Kroep, J.R., Van Groeningen, C.J., Cuesta, M.A., Craanen, M.E., Hoekstra, O.S., Comans, E.F., Bloemena, E., Hoekstra, C.J., Golding, R.P., Twisk, J.W., *et al.* (2003). Positron emission tomography using 2-deoxy-2-[18F]-fluoro-D-glucose for response monitoring in locally advanced gastroesophageal cancer; a comparison of different analytical methods. *Mol Imaging Biol* 5, 337-346.
- Kubota, K., Matsuzawa, T., Ito, M., Ito, K., Fujiwara, T., Abe, Y., Yoshioka, S., Fukuda, H., Hatazawa, J., and Iwata, R. (1985). Lung tumor imaging by positron emission tomography using C-11 L-methionine. *J Nucl Med* 26, 37-42.

Kumar, R., Dhanpathi, H., Basu, S., Rubello, D., Fanti, S., and Alavi, A. (2008). Oncologic PET tracers beyond [(18)F]FDG and the novel quantitative approaches in PET imaging. *Q J Nucl Med Mol Imaging* 52, 50-65.

Lamare, F., Cresson, T., Savean, J., Cheze Le Rest, C., Reader, A.J., and Visvikis, D. (2007). Respiratory motion correction for PET oncology applications using affine transformation of list mode data. *Phys Med Biol* 52, 121-140.

Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R.G., Granton, P., Zegers, C.M., Gillies, R., Boellard, R., Dekker, A., *et al.* (2012). Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 48, 441-446.

Langle, R.R., and Fidler, I.J. (2007). Tumor cell-organ microenvironment interactions in the pathogenesis of cancer metastasis. *Endocr Rev* 28, 297-321.

Larson, S.M., Erdi, Y., Akhurst, T., Mazumdar, M., Macapinlac, H.A., Finn, R.D., Casilla, C., Fazzari, M., Srivastava, N., Yeung, H.W., *et al.* (1999). Tumor treatment response based on visual and quantitative changes in global tumor glycolysis using PET-FDG imaging. The visual response score and the change in total lesion glycolysis. *Clin Positron Imaging* 2, 159-171.

Le Pogam, A., Bousson, N., Hatt, M., Turkheimer, F.E., Prunier-Aesch, C., Guilloteau, D., Baulieu, J.L., and Visvikis, D. (2008). A 3D multi resolution local analysis approach for correction of partial volume effects in emission tomography. Paper presented at: IEEE Nucl Sci Symp Med Imaging Conf.

Le Pogam, A., Hatt, M., Descourt, P., Bousson, N., Tsoumpas, C., Turkheimer, F.E., Prunier-Aesch, C., Baulieu, J.L., Guilloteau, D., and Visvikis, D. (2011). Evaluation of a 3D local multiresolution algorithm for the correction of partial volume effects in positron emission tomography. *Med Phys* 38, 4920-4923.

Le Pogam, A., Lamare, F., Hatt, M., Fernandez, P., Cheze Le Rest, C., and Visvikis, D. (2012). MRI data driven partial volume effects correction in PET imaging using 3D local multi-resolution analysis. *Nucl Instrum Methods Phys Res A in press - uncorrected proof*.

Lecomte, R. (2009). Novel detector technology for clinical PET. *Eur J Nucl Med Mol Imaging* 36 Suppl 1, S69-85.

Levin, C.S., Dahlbom, M., and Hoffman, E.J. (1995). A Monte Carlo correction for the effect of Compton scattering in 3-D PET brain imaging. *IEEE Trans Nucl Sci* 42, 1181-1185.

Lin, C., Itti, E., Haioun, C., Petegnief, Y., Luciani, A., Dupuis, J., Paone, G., Talbot, J.N., Rahmouni, A., and Meignan, M. (2007). Early 18F-FDG PET for prediction of prognosis in patients with diffuse large B-cell lymphoma: SUV-based assessment versus visual analysis. *J Nucl Med* 48, 1626-1632.

Lindholm, P., Minn, H., Leskinen-Kallio, S., Bergman, J., Ruotsalainen, U., and Joensuu, H. (1993). Influence of the blood glucose concentration on FDG uptake in cancer--a PET study. *J Nucl Med* 34, 1-6.

Ling, C.C., Humm, J., Larson, S., Amols, H., Fuks, Z., Leibel, S., and Koutcher, J.A. (2000). Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. *Int J Radiat Oncol Biol Phys* 47, 551-560.

Links, J.M., Zubietta, J.K., Meltzer, C.C., Stumpf, M.J., and Frost, J.J. (1996). Influence of spatially heterogeneous background activity on "hot object" quantitation in brain emission computed tomography. *J Comput Assist Tomogr* 20, 680-687.

Lowe, V.J., DeLong, D.M., Hoffman, J.M., and Coleman, R.E. (1995). Optimum scanning protocol for FDG-PET evaluation of pulmonary malignancy. *J Nucl Med* 36, 883-887.

Lucy, L.B. (1974). An iterative technique for the rectification of observed distributions. *Astron J* 79, 745-745.

Madsen, K., Haahr, M.T., Marnier, L., Keller, S.H., Baaré, W.F., Svarer, C., Hasselbalch, S.G., and Knudsen, G.M. (2011). Age and sex effects on 5-HT(4) receptors in the human brain: a [(11)C]SB207145 PET study. *J Cereb Blood Flow Metab* 31, 1475-1481.

Maisonobe, J.-A., Garcia, C., Lalami, Y., Lefebvre, Y., Awada, A., Flamen, P., and Buvat, I. (2012). Parametric imaging in PET/CT identifies non-responding lesions in head and neck cancer patients undergoing antiangiogenic treatment. *J Nucl Med*, meeting abstract 53, 2295.

Maisonobe, J.-A., Garcia, C., Vanderlinden, B., Hendlisz, A., Flamen, P., and Buvat, I. (2011). What is the best index to characterize early tumor response in FDG PET? *J Nucl Med*, meeting abstract 52, 49.

Maisonobe, J.-A., Stute, S., Necib, H., and Buvat, I. (2010). Is partial volume correction useful for detecting tumor uptake changes in patient monitoring? *J Nucl Med*, meeting abstract 51, 27.

Martin Bland, J., and Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327, 307-310.

Mas, J., Hannequin, P., Ben Younes, R., Bellaton, B., and Bidet, R. (1990). Scatter correction in planar imaging and SPECT by constrained factor analysis of dynamic structures (FADS). *Phys Med Biol* 35, 1451-1465.

Mazziotta, J.C., Phelps, M.E., Plummer, D., and Kuhl, D.E. (1981). Quantitation in positron emission computed tomography: 5. Physical--anatomical effects. *J Comput Assist Tomogr* 5, 734-743.

McDermott, G.M., Welch, A., Staff, R.T., Gilbert, F.J., Schweiger, L., Semple, S.I., Smith, T.A., Hutcheon, A.W., Miller, I.D., Smith, I.C., *et al.* (2007). Monitoring primary breast cancer throughout chemotherapy using FDG-PET. *Breast Cancer Res Treat* 102, 75-84.

Meltzer, C.C., Kinahan, P.E., Greer, P.J., Nichols, T.E., Comtat, C., Cantwell, M.N., Lin, M.P., and Price, J.C. (1999). Comparative evaluation of MR-based partial-volume correction schemes for PET. *J Nucl Med* 40, 2053-2065.

Meltzer, C.C., Leal, J.P., Mayberg, H.S., Wagner, H.N., and Frost, J.J. (1990). Correction of PET data for partial volume effects in human cerebral cortex by MR imaging. *J Comput Assist Tomogr* 14, 561-570.

Menda, Y., Bushnell, D.L., Madsen, M.T., McLaughlin, K., Kahn, D., and Kernstine, K.H. (2001). Evaluation of various corrections to the standardized uptake value for diagnosis of pulmonary malignancy. *Nucl Med Commun* 22, 1077-1081.

Mikhno, A., Devanand, D., Pelton, G., Cuasay, K., Gunn, R., Upton, N., Lai, R.Y., Libri, V., Mann, J.J., and Parsey, R.V. (2008). Voxel-based analysis of 11C-PIB scans for diagnosing Alzheimer's disease. *J Nucl Med* 49, 1262-1269.

Miller, A.B., Hoogstraten, B., Staquet, M., and Winkler, A. (1981). Reporting results of cancer treatment. *Cancer* 47, 207-214.

Miller, T.R., Pinkus, E., Dehdashti, F., and Grigsby, P.W. (2003). Improved prognostic value of 18F-FDG PET using a simple visual analysis of tumor characteristics in patients with cervical cancer. *J Nucl Med* 44, 192-197.

Millet, P., Moulin-Sallanon, M., Tournier, B.B., Dumas, N., Charnay, Y., Ibáñez, V., and Ginovart, N. (2012). Quantification of dopamine D(2/3) receptors in rat brain using factor analysis corrected [(18)F]Fallypride images. *Neuroimage* 62, 1455-1468.

Mohan, R., Brewster, L.J., and Barest, G.D. (1987). A technique for computing dose volume histograms for structure combinations. *Med Phys* 14, 1048-1052.

Mosteller, R.D. (1987). Simplified calculation of body-surface area. *N Engl J Med* 317, 1098.

Müller-Gärtner, H.W., Links, J.M., Prince, J.L., Bryan, R.N., McVeigh, E., Leal, J.P., Davatzikos, C., and Frost, J.J. (1992). Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects. *J Cereb Blood Flow Metab* 12, 571-583.

Necib, H. (2010). Caractérisation de l'évolution tumorale au cours d'une thérapie au moyen d'examen TEP/TDM (Thèse de l'Université Paris-Sud XI).

Necib, H., Dusart, M., Tylski, P., Vanderlinden, B., and Buvat, I. (2008). Parametric imaging for longitudinal monitoring of tumors in FDG PET. *J Nucl Med*, meeting abstract 49, 120P.

Necib, H., Garcia, C., Wagner, A., Vanderlinden, B., Emonts, P., Hendlisz, A., Flamen, P., and Buvat, I. (2011). Detection and characterization of tumor changes in 18F-FDG PET patient monitoring using parametric imaging. *J Nucl Med* 52, 354-361.

Nestle, U., Kremp, S., Schaefer-Schuler, A., Sebastian-Welsch, C., Hellwig, D., Rube, C., and Kirsch, C.M. (2005). 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, 1342-1348.

O'Sullivan, F., Roy, S., O'Sullivan, J., Vernon, C., and Eary, J. (2005). Incorporation of tumor shape into an assessment of spatial heterogeneity for human sarcomas imaged with FDG-PET. *Biostatistics* 6, 293-301.

Ollinger, J.M. (1996). Model-based scatter correction for fully 3D PET. *Phys Med Biol* 41, 153-176.

Orlhac, F. (2012). Caractérisation de l'hétérogénéité tumorale en Tomographie par Emission de Positons (TEP) au 18F-FDG : méthode, valeur pronostique, intérêt pour le suivi thérapeutique (rapport de stage de Master 2, soutenu le 21/06/2012).

Ourselin, S., Roche, A., Prima, S., and Ayache, N. (2000). Block matching: a general framework to improve robustness of rigid registration of medical images. Paper presented at: 3rd Int Conf Med Rob Med Robot Image Comput Assist Surg (MICCAI 2000) (Springer).

Panin, V.Y., Kehren, F., Michel, C., and Casey, M. (2006). Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Med Imaging* 25, 907-921.

Patlak, C.S., Blasberg, R.G., and Fenstermacher, J.D. (1983). Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab* 3, 1-7.

- Phelps, M.E., Huang, S.C., Hoffman, E.J., Selin, C., Sokoloff, L., and Kuhl, D.E. (1979). Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18)2-fluoro-2-deoxy-D-glucose: validation of method. *Ann Neurol* 6, 371-388.
- Pichler, B.J., Kolb, A., Nägele, T., and Schlemmer, H.P. (2010). PET/MRI: paving the way for the next generation of clinical multimodality imaging applications. *J Nucl Med* 51, 333-336.
- Pugachev, A., Ruan, S., Carlin, S., Larson, S.M., Campa, J., Ling, C.C., and Humm, J.L. (2005). Dependence of FDG uptake on tumor microenvironment. *Int J Radiat Oncol Biol Phys* 62, 545-553.
- Qi, J., and Leahy, R.M. (2006). Iterative reconstruction techniques in emission computed tomography. *Phys Med Biol* 51, R541-578.
- Qi, J., Leahy, R.M., Cherry, S.R., Chatziioannou, A., and Farquhar, T.H. (1998). High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. *Phys Med Biol* 43, 1001-1013.
- Quarantelli, M., Berkouk, K., Prinster, A., Landeau, B., Svarer, C., Balkay, L., Alfano, B., Brunetti, A., Baron, J.C., and Salvatore, M. (2004). Integrated software for the analysis of brain PET/SPECT studies with partial-volume-effect correction. *J Nucl Med* 45, 192-201.
- Rahmim, A., Tang, J., Lodge, M.A., Lashkari, S., Ay, M.R., Lautamäki, R., Tsui, B.M., and Bengel, F.M. (2008). Analytic system matrix resolution modeling in PET: an application to Rb-82 cardiac imaging. *Phys Med Biol* 53, 5947-5965.
- Rajendran, J.G., Mankoff, D.A., O'Sullivan, F., Peterson, L.M., Schwartz, D.L., Conrad, E.U., Spence, A.M., Muzi, M., Farwell, D.G., and Krohn, K.A. (2004). Hypoxia and glucose metabolism in malignant tumors: evaluation by [18F]fluoromisonidazole and [18F]fluorodeoxyglucose positron emission tomography imaging. *Clin Cancer Res* 10, 2245-2252.
- Ratain, M.J., and Eckhardt, S.G. (2004). Phase II studies of modern drugs directed against new targets: if you are fazed, too, then resist RECIST. *J Clin Oncol* 22, 4442-4445.
- Reader, A.J., Ally, S., Bakatselos, F., Manavaki, R., Walledge, R.J., Jeavons, A.P., Julyan, P.J., Zhao, S., Hastings, D.L., and Zweit, J. (2002). One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays. *IEEE Trans Nucl Sci* 49, 693-699.
- Reader, A.J., Julyan, P.J., Williams, H., Hastings, D.L., and Zweit, J. (2003). EM algorithm system modeling by image-space techniques for PET reconstruction. *IEEE Trans Nucl Sci* 50, 1392-1397.
- Reivich, M., Kuhl, D., Wolf, A., Greenberg, J., Phelps, M., Ido, T., Casella, V., Fowler, J., Hoffman, E., Alavi, A., *et al.* (1979). The [18F]fluorodeoxyglucose method for the measurement of local cerebral glucose utilization in man. *Circ Res* 44, 127-137.
- Richardson, W.H. (1972). Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America* 62, 55-59.
- Rousset, O., Rahmim, A., Alavi, A., and Zaidi, H. (2007). Partial volume correction strategies in PET. *PET Clinics* 2, 235-249.
- Rousset, O.G., Ma, Y., and Evans, A.C. (1998a). Correction for partial volume effects in PET: principle and validation. *J Nucl Med* 39, 904-911.

Rousset, O.G., Ma, Y., Wong, D.F., and Evans, A.C. (1998b). Pixel-versus region based partial volume correction in PET. In *Quantitative functional brain imaging with positron emission tomography* (Academic Press), pp. 67-75.

Sakaguchi, Y., Mizoguchi, N., Mitsumoto, T., Mitsumoto, K., Himuro, K., Ohya, N., Kaneko, K., Baba, S., Abe, K., Onizuka, Y., *et al.* (2010). A simple table lookup method for PET/CT partial volume correction using a point-spread function in diagnosing lymph node metastasis. *Ann Nucl Med* 24, 585-591.

Schelbert, H.R., Hoh, C.K., Royal, H.D., Brown, M., Dahlbom, M.N., Dehdashti, F., and Wahl, R.L. (1998). Procedure guideline for tumor imaging using fluorine-18-FDG. *J Nucl Med* 39, 1302-1305.

Schmidt, K., Lucignani, G., Moresco, R.M., Rizzo, G., Gilardi, M.C., Messa, C., Colombo, F., Fazio, F., and Sokoloff, L. (1992). Errors introduced by tissue heterogeneity in estimation of local cerebral glucose utilization with current kinetic models of the [18F]fluorodeoxyglucose method. *J Cereb Blood Flow Metab* 12, 823-834.

Schmidt, K., Mies, G., and Sokoloff, L. (1991). Model of kinetic behavior of deoxyglucose in heterogeneous tissues in brain: a reinterpretation of the significance of parameters fitted to homogeneous tissue models. *J Cereb Blood Flow Metab* 11, 10-24.

Schwaiger, M., and Wester, H.J. (2011). How many PET tracers do we need? *J Nucl Med* 52 *Suppl* 2, 36S-41S.

Segobin, S.H., Matthews, J.C., Markiewicz, P.J., and Herholz, K. (2010). A hybrid between region-based and voxel-based methods for Partial Volume correction in PET. Paper presented at: IEEE Nucl Sci Symp Med Imaging Conf.

Shannon, C.E. (1949). Communication in the presence of noise. *Proc IRE* 37, 10-21.

Shao, Y., Cherry, S.R., Farahani, K., Meadors, K., Siegel, S., Silverman, R.W., and Marsden, P.K. (1997). Simultaneous PET and MR imaging. *Phys Med Biol* 42, 1965-1970.

Shepp, L.A., and Vardi, Y. (1982). Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging* 1, 113-122.

Shidahara, M., Tsoumpas, C., Hammers, A., Boussion, N., Visvikis, D., Suhara, T., Kanno, I., and Turkheimer, F.E. (2009). Functional and structural synergy for resolution recovery and partial volume correction in brain PET. *Neuroimage* 44, 340-348.

Siegel, S., and Castellan, N. (1988). *Nonparametric Statistics for the Behavioral Sciences*, 2nd edition edn (New York, NY: McGraw-Hill).

Singnurkar, A., Solomon, S.B., Gönen, M., Larson, S.M., and Schöder, H. (2010). 18F-FDG PET/CT for the prediction and detection of local recurrence after radiofrequency ablation of malignant lung lesions. *J Nucl Med* 51, 1833-1840.

Sinha, S., Lucas-Quesada, F.A., DeBruhl, N.D., Sayre, J., Farria, D., Gorczyca, D.P., and Bassett, L.W. (1997). Multifeature analysis of Gd-enhanced MR images of breast lesions. *J Magn Reson Imaging* 7, 1016-1026.

Smarandache, F., and Dezert, J. (2006). *Advances and applications of DSMT for information fusion: collected works* (American Research Press).

Sokoloff, L., Reivich, M., Kennedy, C., Des Rosiers, M.H., Patlak, C.S., Pettigrew, K.D., Sakurada, O., and Shinohara, M. (1977). The [14C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J Neurochem* 28, 897-916.

- Soret, M., Bacharach, S.L., and Buvat, I. (2007). Partial-volume effect in PET tumor imaging. *J Nucl Med* 48, 932-945.
- Soret, M., Riddell, C., Hapdey, S., and Buvat, I. (2002). Biases affecting the measurements of tumor-to-background activity ratio in PET. *IEEE Trans Nucl Sci* 49, 2112-2118.
- Soussan, M., Cyrta, J., Maisonobe, J.-A., Chouahnia, K., Morere, J.-F., Eder, V., and Buvat, I. (2012). Performance of SUV- and volume-based parameters in FDG PET/CT to predict histopathological response in non small cell lung cancer (NSCLC). *J Nucl Med*, meeting abstract 53, 1393-.
- Soussan, M., Galas, J.L., Ouvrier, M.J., Pop, G., Neuman, A., and Weinmann, P. (2011). Impact clinique de la technique du temps-de-vol en TEP. *Médecine Nucléaire* 35, 401-405.
- Spira, D., Fenchel, M., Lauer, U.M., Claussen, C.D., Gregor, M., Bitzer, M., and Horgler, M. (2011). Comparison of different tumor response criteria in patients with hepatocellular carcinoma after systemic therapy with the multikinase inhibitor sorafenib. *Acad Radiol* 18, 89-96.
- Srinivas, S.M., Dhurairaj, T., Basu, S., Bural, G., Surti, S., and Alavi, A. (2009). A recovery coefficient method for partial volume correction of PET images. *Ann Nucl Med* 23, 341-348.
- Stahl, A., Ott, K., Schwaiger, M., and Weber, W.A. (2004). Comparison of different SUV-based methods for monitoring cytotoxic therapy with FDG PET. *Eur J Nucl Med Mol Imaging* 31, 1471-1478.
- Stroszczyński, C., Jost, D., Reichardt, P., Chmelik, P., Gaffke, G., Kretzschmar, A., Schneider, U., Felix, R., and Hohenberger, P. (2005). Follow-up of gastro-intestinal stromal tumours (GIST) during treatment with imatinib mesylate by abdominal MRI. *Eur Radiol* 15, 2448-2456.
- Stute, S. (2010). Modélisation avancée en simulation Monte Carlo de tomographie par émission de positons pour l'amélioration de la reconstruction et de la quantification. (Thèse de l'Université Paris-Sud XI).
- Stute, S., Carlier, T., Cristina, K., Noblet, C., Martineau, A., Hutton, B., Barnden, L., and Buvat, I. (2011). Monte Carlo simulations of clinical PET and SPECT scans: impact of the input data on the simulated images. *Phys Med Biol* 56, 6441-6457.
- Sun, C., and Wee, W.G. (1983). Neighboring gray level dependence matrix for texture classification. *Comput Vis Graph Image Proces* 23, 341-352.
- Sundaram, S.K., Freedman, N.M., Carrasquillo, J.A., Carson, J.M., Whatley, M., Libutti, S.K., Sellers, D., and Bacharach, S.L. (2004). Simplified kinetic analysis of tumor 18F-FDG uptake: a dynamic approach. *J Nucl Med* 45, 1328-1333.
- Swets, J.A. (1988). Measuring the accuracy of diagnostic systems. *Science* 240, 1285-1293.
- Szekely, N., Toth, N., and Pataki, B. (2006). A hybrid system for detecting masses in mammographic images. *IEEE Trans Instru Meas* 55, 944-952.
- Tang, J., and Rahmim, A. (2009). Bayesian PET image reconstruction incorporating anato-functional joint entropy. *Phys Med Biol* 54, 7063-7075.
- Tang, X. (1998). Texture information in run-length matrices. *IEEE Trans Image Proces* 7, 1602-1609.

Taschereau, R., Rannou, F.R., and Chatziioannou, A.F. (2011). A modeled point spread function for a noise-free system matrix. Paper presented at: IEEE Nucl Sci Symp Med Imaging Conf.

Teo, B.K., Seo, Y., Bacharach, S.L., Carrasquillo, J.A., Libutti, S.K., Shukla, H., Hasegawa, B.H., Hawkins, R.A., and Franc, B.L. (2007). Partial-volume correction in PET: validation of an iterative postreconstruction method with phantom and patient data. *J Nucl Med* 48, 802-810.

Tesar, L., Shimizu, A., Smutek, D., Kobatake, H., and Nawano, S. (2008). Medical image analysis of 3D CT images based on extension of Haralick texture features. *Comput Med Imaging Graph* 32, 513-520.

Therasse, P., Arbuck, S.G., Eisenhauer, E.A., Wanders, J., Kaplan, R.S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A.T., Christian, M.C., *et al.* (2000). 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, 205-216.

Therasse, P., Eisenhauer, E.A., and Verweij, J. (2006). RECIST revisited: a review of validation studies on tumour assessment. *Eur J Cancer* 42, 1031-1039.

Thibault, G. (2009). Indices de formes et de textures : de la 2D vers la 3D. Application au classement de noyaux de cellules. (Thèse de l'Université Aix-Marseille).

Thibault, G., Angulo, J., and Meyer, F. (2011). Advanced statistical matrices for texture characterization: Application to DNA chromatin and microtubule network classification. Paper presented at: IEEE Int Conf Image Proces.

Thie, J.A. (2004). Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med* 45, 1431-1434.

Thomas, B.A., Erlandsson, K., Modat, M., Thurfjell, L., Vandenberghe, R., Ourselin, S., and Hutton, B.F. (2011). The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. *Eur J Nucl Med Mol Imaging* 38, 1104-1119.

Tixier, F., Hatt, M., Le Rest, C.C., Le Pogam, A., Corcos, L., and Visvikis, D. (2012). Reproducibility of Tumor Uptake Heterogeneity Characterization Through Textural Feature Analysis in 18F-FDG PET. *J Nucl Med* 53, 693-700.

Tixier, F., Le Rest, C.C., Hatt, M., Albarghach, N., Pradier, O., Metges, J.P., Corcos, L., and Visvikis, D. (2011). Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med* 52, 369-378.

Tohka, J., and Reilhac, A. (2006). A Monte Carlo study of deconvolution algorithms for partial volume correction in quantitative PET. Paper presented at: IEEE Nucl Sci Symp Conf Rec.

Tohka, J., and Reilhac, A. (2008). Deconvolution-based partial volume correction in Raclopride-PET and Monte Carlo comparison to MR-based method. *Neuroimage* 39, 1570-1584.

Tohme, M.S., and Qi, J. (2009). Iterative image reconstruction for positron emission tomography based on a detector response function estimated from point source measurements. *Phys Med Biol* 54, 3709-3725.

Tomasi, G., Turkheimer, F., and Aboagye, E. (2012). Importance of quantification for the analysis of PET data in oncology: review of current methods and trends for the future. *Mol Imaging Biol* 14, 131-146.

Turkington, T.G. (2001). Introduction to PET instrumentation. *J Nucl Med Technol* 29, 4-11.

Tylski, P. (2009). Caractérisation des tumeurs et de leur évolution en TEP/TDM au 18F-FDG (Paris-Sud (XI)).

Tylski, P., Stute, S., Grotus, N., Doyeux, K., Hapdey, S., Gardin, I., Vanderlinden, B., and Buvat, I. (2010). Comparative assessment of methods for estimating tumor volume and standardized uptake value in (18)F-FDG PET. *J Nucl Med* 51, 268-276.

Unser, M. (1995). Texture classification and segmentation using wavelet frames. *IEEE Trans Image Proces* 4, 1549-1560.

van Cittert, P.H. (1931). Zum Einfluß der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. II. *Zeitschrift für Physik A Hadrons and Nuclei* 69, 298-308.

Van den Abbeele, A., Badawi, R., Cliche, J., Janicek, M., Tetrault, R., Spangler, T., Potter, A., Merriam, R., Silberman, S., Dimitrijevic, S., *et al.* (2002). 18F-FDG-PET predicts response to imatinib mesylate (Gleevec) in patients with advanced gastrointestinal stromal tumors (GIST). Paper presented at: *Proc Am Soc Clin Oncol* 21 (abstract 1610).

van Heijl, M., Omloo, J.M., van Berge Henegouwen, M.I., van Lanschot, J.J., Sloof, G.W., and Boellaard, R. (2010). Influence of ROI definition, partial volume correction and SUV normalization on SUV-survival correlation in oesophageal cancer. *Nucl Med Commun* 31, 652-658.

van Velden, F.H., Cheebsumon, P., Yaqub, M., Smit, E.F., Hoekstra, O.S., Lammertsma, A.A., and Boellaard, R. (2011). Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies. *Eur J Nucl Med Mol Imaging* 38, 1636-1647.

Vanderhoek, M., Perlman, S.B., and Jeraj, R. (2012). Impact of the definition of peak standardized uptake value on quantification of treatment response. *J Nucl Med* 53, 4-11.

Vauclin, S., Doyeux, K., Hapdey, S., Edet-Sanson, A., Vera, P., and Gardin, I. (2009). Development of a generic thresholding algorithm for the delineation of 18FDG-PET-positive tissue: application to the comparison of three thresholding models. *Phys Med Biol* 54, 6901-6916.

Vesselle, H., Freeman, J.D., Wiens, L., Stern, J., Nguyen, H.Q., Hawes, S.E., Bastian, P., Salskov, A., Vallières, E., and Wood, D.E. (2007). Fluorodeoxyglucose uptake of primary non-small cell lung cancer at positron emission tomography: new contrary data on prognostic role. *Clin Cancer Res* 13, 3255-3263.

Vesselle, H., Schmidt, R.A., Pugsley, J.M., Li, M., Kohlmyer, S.G., Vallires, E., and Wood, D.E. (2000). Lung cancer proliferation correlates with [F-18]fluorodeoxyglucose uptake by positron emission tomography. *Clin Cancer Res* 6, 3837-3844.

Vesselle, H., Turcotte, E., Wiens, L., Schmidt, R., Takasugi, J.E., Lalani, T., Vallières, E., and Wood, D.E. (2004). Relationship between non-small cell lung cancer fluorodeoxyglucose uptake at positron emission tomography and surgical stage with relevance to patient prognosis. *Clin Cancer Res* 10, 4709-4716.

Videen, T.O., Perlmutter, J.S., Mintun, M.A., and Raichle, M.E. (1988). Regional correction of positron emission tomography data for the effects of cerebral atrophy. *J Cereb Blood Flow Metab* 8, 662-670.

Vriens, D., Visser, E.P., de Geus-Oei, L.F., and Oyen, W.J. (2010). Methodological considerations in quantification of oncological FDG PET studies. *Eur J Nucl Med Mol Imaging* 37, 1408-1425.

Wadell, H. (1932). Volume, shape, and roundness of rock particles. *J Geol* 40, 443-451.

Wahl, R.L., Jacene, H., Kasamon, Y., and Lodge, M.A. (2009). From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. *J Nucl Med* 50 *Suppl 1*, 122S-150S.

Wahl, R.L., Zasadny, K., Helvie, M., Hutchins, G.D., Weber, B., and Cody, R. (1993). Metabolic monitoring of breast cancer chemohormonotherapy using positron emission tomography: initial evaluation. *J Clin Oncol* 11, 2101-2111.

Warburg, O. (1956a). On respiratory impairment in cancer cells. *Science* 124, 269-270.

Warburg, O. (1956b). On the origin of cancer cells. *Science* 123, 309-314.

Warburg, O., and Posener, K. (1924). Ueber den stoffwechsel der tumoren. *Biochem Z* 152, 319.

Watson, C.C. (2011). Estimating effective model kernel widths for PSF reconstruction in PET. Paper presented at: IEEE Nucl Sci Symp Med Imaging Conf.

Weber, W.A. (2010). Quantitative analysis of PET studies. *Radiother Oncol* 96, 308-310.

Westerterp, M., Pruijm, J., Oyen, W., Hoekstra, O., Paans, A., Visser, E., van Lanschot, J., Sloof, G., and Boellaard, R. (2007). Quantification of FDG PET studies using standardised uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters. *Eur J Nucl Med Mol Imaging* 34, 392-404.

WHO (1979). WHO handbook for reporting results of cancer treatment, Vol 48, Offset Publication (Geneva, Switzerland: World Health Organization).

Wienhard, K. (2002). Measurement of glucose consumption using [(18)F]fluorodeoxyglucose. *Methods* 27, 218-225.

Williams, C.W., Crabtree, M.C., and Burgiss, S.G. (1979). Design and performance characteristics of a positron emission computed axial tomograph-ECAT-II. *IEEE Trans Nucl Sci* 26, 619-627.

Wong, C.Y., Thie, J., Parling-Lynch, K.J., Zakalik, D., Margolis, J.H., Gaskill, M., Hill, J., Qing, F., Fink-Bennett, D., and Nagle, C. (2005). Glucose-normalized standardized uptake value from (18)F-FDG PET in classifying lymphomas. *J Nucl Med* 46, 1659-1663.

Wong, W.H., and Hicks, K. (1994). A clinically practical method to acquire parametric images of unidirectional metabolic rates and blood spaces. *J Nucl Med* 35, 1206-1212.

Wu, H.M., Huang, S.C., Choi, Y., Hoh, C.K., and Hawkins, R.A. (1995). A modeling method to improve quantitation of fluorodeoxyglucose uptake in heterogeneous tumor tissue. *J Nucl Med* 36, 297-306.

Xu, D., Kurani, A., Furst, J., and Raicu, D. (2004). Run-length encoding for volumetric texture. Paper presented at: The 4th IASTED International Conference on Visualization, Imaging, and Image Processing (Marbella, Spain).

Xu, Y., Sonka, M., McLennan, G., Guo, J., and Hoffman, E.A. (2006). MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. *IEEE Trans Med Imaging* 25, 464-475.

Yan, D., Vicini, F., Wong, J., and Martinez, A. (1997). Adaptive radiation therapy. *Phys Med Biol* 42, 123-132.

Young, H., Baum, R., Cremerius, U., Herholz, K., Hoekstra, O., Lammertsma, A.A., Pruim, J., and Price, P. (1999). Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. *Eur J Cancer* 35, 1773-1782.

Zaidi, H., and El Naqa, I. (2010). PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. *Eur J Nucl Med Mol Imaging* 37, 2165-2187.

Zasadny, K.R., and Wahl, R.L. (1993). 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, 847-850.

Zhu, A., and Shim, H. (2011). Current molecular imaging positron emitting radiotracers in oncology. *Nucl Med Mol Imaging* 45, 1-14.

Zhuang, H., Pourdehnad, M., Lambright, E.S., Yamamoto, A.J., Lanuti, M., Li, P., Mozley, P.D., Rossman, M.D., Albelda, S.M., and Alavi, A. (2001). Dual time point 18F-FDG PET imaging for differentiating malignant from inflammatory processes. *J Nucl Med* 42, 1412-1417.

Publications

1. Articles dans des revues avec comité de lecture

- Maisonobe J-A, Garcia CA, Necib H, Vanderlinden B, Hendlisz A, Flamen P, Buvat I. *Comparison of PET metabolic indices for the early assessment of tumor response in metastatic colorectal cancer patients treated by polychemotherapy.* Ahead of print in *European Journal of Nuclear Medicine and Molecular Imaging*.
- Soussan M, Chouahnia K, Maisonobe J-A, Moubaya M, Eder V, Morère J-F, Buvat I. Prognosis implications of volume-based measurements on FDG-PET/CT in stage III non small cell lung cancer after induction chemotherapy. En révision au *European Journal of Nuclear Medicine and Molecular Imaging*.

2. Communications sans actes

- Maisonobe J-A, Stute S, Necib H, and Buvat I. (2010). *Is partial volume correction useful for detecting tumor uptake changes in patient monitoring?* J Nucl Med, meeting abstract 51, 27.
- Maisonobe J-A, Garcia C, Vanderlinden B, Hendlisz A, Flamen P, and Buvat I. (2011). *What is the best index to characterize early tumor response in FDG PET?* J Nucl Med, meeting abstract 52, 49.
- Maisonobe J-A, Hapdey S, Edet-Sanson A, Dubray B, Gardin I, Vera P, and Buvat I. (2012b). *Parametric imaging for assessing tumor response in patients with non-small cell lung cancer.* J Nucl Med, meeting abstract 53, 267.

3. Posters

- Maisonobe J-A, Garcia C, Lalami Y, Lefebvre Y, Awada A, Flamen P, and Buvat I. (2012). *Parametric imaging in PET/CT identifies non-responding lesions in head and neck cancer patients undergoing antiangiogenic treatment.* J Nucl Med, meeting abstract 53, 2295.
- Maisonobe J-A, Hapdey S, Edet-Sanson A, Dubray B, Véra P, Buvat I. (2012). *Parametric Imaging for assessing tumor response follow-up with PET/CT during radiotherapy in patients with non-small cell lung cancer.* Molecular Imaging in Radiation Oncology (MIRO) Conference Records.
- Soussan M, Cyrtta, Maisonobe J-A, Chouahnia K, Morère J.-F, Eder V, and Buvat I. (2012). *Performance of SUV- and volume-based parameters in FDG PET/CT to predict histopathological response in non small cell lung cancer (NSCLC).* J Nucl Med, meeting abstract 53, 1393.
- Soussan M, Maisonobe J-A, Boubaya M, Chouahnia K, Morère J.-F, Eder V, and Buvat I. (2012b). *Prognosis value of FDG PET image-based parameters in stage III non-small cell lung cancer.* J Nucl Med meeting abstract 53, 1227.

4. Article en cours de rédaction

- *Parametric imaging to assess tumor response in serial FDG-PET/CT images during radiotherapy in patients with non-small cell lung cancers.*