



**HAL**  
open science

# Analyse des images de tomographie par émission de positons pour la prédiction de récurrence du cancer du col de l'utérus

Geoffrey Roman Jimenez

► **To cite this version:**

Geoffrey Roman Jimenez. Analyse des images de tomographie par émission de positons pour la prédiction de récurrence du cancer du col de l'utérus. Traitement du signal et de l'image [eess.SP]. Université de Rennes, 2016. Français. NNT : 2016REN1S037 . tel-01420492

**HAL Id: tel-01420492**

**<https://theses.hal.science/tel-01420492v1>**

Submitted on 20 Dec 2016

**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.



THÈSE / UNIVERSITÉ DE RENNES 1  
*sous le sceau de l'Université Bretagne Loire*

pour le grade de  
**DOCTEUR DE L'UNIVERSITÉ DE RENNES 1**  
*Mention : Traitement du Signal et télécommunications*  
**Ecole doctorale MATISSE**

présentée par

**Geoffrey Roman Jimenez**

Préparée à l'unité de recherche LTSI – INSERM 1099  
Laboratoire du Traitement du Signal et de l'Image  
ISTIC UFR Electronique et Informatique

---

**Analyse des images  
de tomographie par  
émission de positons  
pour la prédiction de  
récidive du cancer du  
col de l'utérus**

**Thèse soutenue à Rennes  
le 25/03/2016**

devant le jury composé de :

**Matthieu HATT**

CR, LATIM INSERM 1101 / *rapporteur*

**Bernard DUBRAY**

PU/PH, CLCC Rouen - LITIS / *rapporteur*

**Su RUAN**

PU, Université de Rouen - LITIS / *examineur*

**Henri DER SARKISSIAN**

DR, Keosys Medical Imaging / *examineur*

**Renaud DE CREVOISIER**

PU/PH, Université de Rennes 1 / *directeur de thèse*

**Oscar ACOSTA**

MCU, Université de Rennes 1 / *encadrant de thèse*



# TABLES DES MATIERES

<b>Tables des matières .....</b>	<b>3</b>
<b>Introduction.....</b>	<b>7</b>
<b>1 Contexte médico-scientifique et problématique .....</b>	<b>9</b>
1.1 Cancer du col de l'utérus et imagerie TEP .....	11
1.1.1 Le cancer du col de l'utérus .....	11
1.1.2 La tomographie par émission de positons : Principes et définitions .....	15
1.1.3 L'utilisation standard de l'imagerie TEP en oncologie .....	20
1.1.4 L'utilisation de la TEP au 18F-FDG dans la prédiction des cancers autres que le col utérin ....	24
1.1.5 L'utilisation de la TEP au 18F-FDG dans la prédiction des cancers du col utérin.....	27
1.2 L'analyse de l'imagerie TEP pour la prédiction de récurrence.....	31
1.2.1 Prétraitements d'image dans l'imagerie TEP.....	31
1.2.2 Segmentation de l'imagerie TEP.....	33
1.2.3 Quantification et extraction de paramètres de l'imagerie TEP .....	38
1.3 Enjeux méthodologiques dans l'analyse des images TEP pour la prédiction de récurrence .....	46
1.3.1 Faible spécificité du radiotracer FDG.....	46
1.3.2 Impact de la segmentation pour la prédiction .....	46
1.3.3 Limites de la quantification de l'imagerie TEP.....	47
<b>2 Objectifs de la Thèse .....</b>	<b>49</b>
<b>3 Etape de prétraitement : Suppression des artefacts d'hyperfixation vésicale dans les images TEP au 18F-FDG pour l'analyse du métabolisme des tumeurs du col utérin (article au journal CBM 2015).....</b>	<b>51</b>
3.1 Abstract.....	53
3.2 Introduction .....	53
3.3 Materials and Methods.....	55
3.3.1 PET/CT dataset.....	55
3.3.2 Detection of bladder metabolic artefact .....	57
3.3.3 Validation.....	60
3.4 Results.....	62
3.5 Discussion .....	65
3.6 Conclusion.....	67
<b>4 Segmentation et caractérisation du métabolisme tumoral pré- et per-thérapeutique des images TEP 18F-FDG pour la prédiction de récurrence (communication au workshop IGRT de MICCAI 2012) .....</b>	<b>69</b>
4.1 Abstract.....	70
4.2 Introduction .....	70

4.3	Materials and Methods .....	71
4.3.1	Data and Clinical Protocol .....	71
4.3.2	Segmentation & Characterization of the metabolic tumor region .....	71
4.3.3	Tumor recurrence prediction .....	75
4.4	Results .....	75
4.4.1	FGMC evaluation .....	75
4.4.2	Explanatory features identification .....	75
4.5	Discussion .....	78
4.6	Conclusion .....	79
<b>5</b>	<b>Impact de la méthode de segmentation sur la valeur pronostic du signal tumoral 18F-FDG TEP et de son évolution au cours du traitement (article publié au journal IRBM 2013) .....</b>	<b>81</b>
5.1	Abstract .....	83
5.2	Introduction .....	83
5.3	Materials and Methods .....	84
5.3.1	Data and clinical protocol .....	84
5.3.2	PET image analysis .....	84
5.4	Results .....	86
5.5	Discussion .....	87
5.6	Conclusion .....	88
<b>6</b>	<b>Valeur prédictive des paramètres extraits des images TEP au 18F-FDG pour la récurrence des tumeurs du col utérin (communication ESTRO 2014 et soumission article au journal IJROBP en février 2016) .....</b>	<b>89</b>
6.1	Abstract .....	91
6.2	Introduction .....	91
6.3	Material and Methods .....	92
6.3.1	Patients and tumor .....	92
6.3.2	Treatment and follow-up .....	92
6.3.3	18F-FDG PET/CT acquisitions .....	93
6.3.4	18F-FDG PET/CT image analysis .....	93
6.3.5	Statistical analysis .....	94
6.4	Results .....	94
6.4.1	Assessment of PET response to treatment (at 40 Gy) .....	95
6.4.2	Predictive value of 18F-FDG PET/CT parameters .....	95
6.5	Discussion .....	96
6.5.1	Feasibility and timing .....	96
6.5.2	Predictive value compared with literature .....	97
6.5.3	Optimal Cut-off .....	97
6.5.4	Limits .....	97

6.6	Conclusion.....	104
<b>7</b>	<b>Quantification par pondération d'appartenance des images TEP au 18F-FDG pour la prédiction de récurrence (Communication EMBC 2015 et ESTRO 2016 et article soumis au journal JNM février 2016)....</b>	<b>107</b>
7.1	Abstract.....	109
7.2	Introduction .....	110
7.3	Material and Methods .....	110
7.3.1	Patients and tumor .....	111
7.3.2	Treatment and follow-up.....	111
7.3.3	18F-FDG PET/CT acquisitions.....	112
7.3.3.1	18F-FDG PET image analyses.....	112
7.3.4	Statistical analysis .....	113
7.4	Results.....	114
7.4.1	Predictive value of hard-thresholding on recurrence.....	114
7.4.2	Predictive value of fuzzy-thresholding .....	114
7.5	Discussion .....	115
7.5.1	State of the Art .....	115
7.5.2	Fuzzy-thresholding and fuzzy-weighted quantification.....	115
7.5.3	Limitations .....	116
7.6	Conclusion.....	124
<b>8</b>	<b>Random Forest pour la prédiction de récurrence des cancers du col utérin à partir de la TEP au 18F-FDG</b>	<b>125</b>
8.1	Abstract.....	127
8.2	Introduction .....	127
8.3	Material and Methods .....	128
8.3.1	Patients and tumor .....	128
8.3.2	Treatment and follow-up.....	129
8.3.3	18F-FDG PET/CT acquisitions.....	129
8.3.4	18F-FDG PET analyses.....	129
8.3.5	statistical analyses .....	130
8.4	Results.....	131
8.5	Discussion .....	132
8.6	Appendix.....	140
8.7	Conclusion.....	142
<b>9</b>	<b>Conclusion et Perspectives .....</b>	<b>143</b>
9.1	Prétraitement : contributions et limites .....	143
9.1.1	Suppression de l'hyperfixation vésicale.....	143

---

9.2	Segmentation : contributions et limites.....	144
9.2.1	L'impact de la segmentation sur la prédiction de récidence .....	144
9.3	Quantification : contributions et limites .....	145
9.3.1	Les paramètres identifiés comme prédictifs de la récidence post-traitement des cancers du col de l'utérus.....	145
9.3.2	L'évaluation de l'évolution du signal tumoral au cours du traitement pour la prédiction ....	146
9.4	Perspectives .....	147
	<b>Valorisation scientifique du travail de thèse .....</b>	<b>149</b>
	Articles de journaux publiés .....	149
	Articles de journaux soumis.....	149
	Communications en conférence avec comité de lecture .....	149
	<b>Références.....</b>	<b>151</b>

## INTRODUCTION

Le cancer du col de l'utérus constitue aujourd'hui le quatrième type de cancer le plus diagnostiqué et la quatrième cause de décès lié au cancer chez la femme dans le monde. En 2012, un total de 527 600 nouveaux cas de cancer du col utérin, ainsi que 265 700 décès liés à celui-ci, ont été recensés dans le monde.

De nombreuses options thérapeutiques existent pour traiter cette maladie. Parmi elles, la radiothérapie est aujourd'hui largement préconisée, notamment pour les cancers du col utérin localement avancés dont le traitement, standardisé, se base sur la combinaison d'une radiothérapie externe, d'une chimiothérapie concomitante et suivi d'une curiethérapie. Pour ce type de cancer, ce traitement offre de très bonnes réponses et la majorité des patientes sont en rémission à la fin de leur traitement.

Cependant, au cours des mois qui suivent le traitement, certaines des patientes vont présenter une récurrence de leur cancer (environ 30% pour les cancers du col de l'utérus localement avancés). Ces récurrences post-traitement étant généralement très virulentes, les traitements de rattrapage sont souvent peu efficaces, réduisant ainsi considérablement les chances de survies de ces patientes. La possibilité de détecter précocement le risque de récurrence pour les patientes traitées est alors cruciale car celle-ci permettrait d'adapter la stratégie thérapeutique des patientes à risques en vue d'augmenter leur chance de survie.

La tomographie par émission de positons (TEP) au fluorodésoxyglucose marqué au fluor 18 (18F-FDG) est une technique d'imagerie fonctionnelle fréquemment utilisée en oncologie car elle reflète l'activité métabolique glucidique des cellules. Au-delà de son utilisation bien établie dans le bilan diagnostique du cancer et dans la détection des rechutes post-traitements, celle-ci est aujourd'hui utilisée en cours de thérapie pour évaluer la réponse métabolique au traitement ou pour l'aide au ciblage radiothérapeutique de la tumeur. Liée à l'agressivité tumorale, le métabolisme glucidique des tumeurs – au diagnostic du cancer ou en cours de traitement – est également de plus en plus étudié pour sa valeur pronostique sur les résultats carcinologiques. Dans la problématique de la prédiction, la TEP est alors évaluée dans sa capacité à fournir des paramètres permettant de discriminer, dans un groupe de patients traités, les patients à risques de récurrence.

L'exploration et l'identification de facteurs prédictifs nécessite toutefois une analyse fine et non biaisée du signal tumoral afin d'extraire des paramètres qui puissent être reproductibles et standardisés. De nombreux facteurs limitent néanmoins l'analyse de l'imagerie TEP ; la faible résolution spatiale, l'effet de volume partiel, la présence de bruit et d'artefact ou encore le faible contraste rencontré, sont autant de facteurs à prendre en compte dans l'analyse.

Une première limitation dans l'analyse de la TEP 18F-FDG est sa non spécificité aux cellules cancéreuses, car l'augmentation du métabolisme glucidique peut aussi être marqué lors de phénomènes inflammatoires ou infectieux. L'inflammation des tissus sains, en cours d'irradiation, est donc une source potentielle de perte de spécificité du signal. Ainsi, l'analyse des images en cours de traitement nécessite de prendre en compte la présence d'artefacts inflammatoires et de la fonte tumorale dans l'analyse. Par ailleurs, l'hyperfixation physiologique des organes environnants peut également être la source d'une perte de spécificité dans l'analyse du signal. Dans la localité du col utérin, la vessie, adjacente à l'organe atteint, peut fixer le 18F-FDG de manière comparable à la tumeur créant un continuum d'hyperfixation pouvant gêner l'analyse.

Une deuxième difficulté pour l'analyse du signal TEP se situe au niveau de l'isolation du signal tumoral du reste de l'image. En effet, l'imagerie TEP souffre d'un faible contraste, d'une faible résolution spatiale et de la présence de bruit d'acquisition. Les délimitations de la tumeur peuvent donc varier d'une image à l'autre,



en fonction de la méthode de segmentation choisie et des variations statistiques entre les examens et les patients. Une des techniques les plus utilisées en routine clinique consiste à réaliser un seuillage d'intensité de l'image TEP. Cette méthodologie est cependant très sensible au seuil choisi et a montré des limites à s'adapter à la taille des tumeurs et aux variations de bruit et de contraste entre les examens et les patients. De nombreuses études se sont penchées sur cette problématique afin de proposer des méthodologies s'adaptant aux spécificités de chaque image et d'augmenter la précision et la reproductibilité de la délinéation. Cependant, malgré les avancées considérables dans ce domaine, aucune méthodologie ne fait aujourd'hui consensus et très peu d'études ont comparé leur impact sur la valeur prédictive des paramètres ainsi extraits. L'importance de la segmentation TEP dans la prédiction de récurrence est une question encore très peu étudiée et son impact sur les paramètres extraits reste encore à établir.

L'évaluation de la captation glucidique des images TEP est généralement focalisée sur des paramètres simples mesurant l'hyperfixation maximale, le volume métabolique, la glycolyse totale de la lésion ou encore la valeur moyenne des hyperfixations au sein de la lésion. Certaines études proposent de prendre en compte de nouvelles mesures, notamment liées à l'hétérogénéité de l'hyperfixation au sein de la tumeur, afin de caractériser plus finement le métabolisme tumoral. La pertinence de ces paramètres reste encore à démontrer dans l'objectif de prédiction.

L'analyse de l'évolution du signal TEP est aujourd'hui couramment utilisée pour l'évaluation de la réponse thérapeutique, notamment au travers des critères RECIST, EORTC et/ou PERCIST qui s'attachent à mesurer l'évolution de paramètres extraits des images acquises au cours du traitement. Dans ce même contexte, certaines études ont également proposé l'utilisation d'images paramétriques, obtenues par recalage d'espace anatomique entre les différents examens (généralement pré- et per-thérapeutique). A ce jour, ces mesures d'évolution du signal ont très peu été étudiées dans le contexte de la prédiction.

Ainsi, d'importantes questions se posent encore aujourd'hui sur les méthodologies à adopter au sein du processus de quantification du métabolisme tumoral pour permettre une extraction de paramètres pertinents à la prédiction de récurrence.

Dans le cadre du traitement radiothérapeutique des cancers du col de l'utérus, ces travaux de thèse visent à explorer la valeur prédictive des images TEP au 18F-FDG, pré- et per-traitement, au regard de différentes méthodes d'analyse d'images et d'évaluer la pertinence de ces traitements dans l'extraction de paramètres prédictifs de la récurrence et/ou de la survie des patientes.

Après une description du contexte médical et de la problématique scientifique puis de la problématique méthodologique au chapitre 1, nous exposerons, au chapitre 2, les objectifs de cette thèse. Les chapitres 3, 4, 5, 6, 7 et 8 présentent des études menées dans l'optique de répondre aux diverses problématiques soulevées. Ce manuscrit s'achève par une conclusion générale des résultats obtenus et la présentation des perspectives de ces travaux de thèse.

# 1 CONTEXTE MEDICO-SCIENTIFIQUE ET PROBLEMATIQUE

*Ce premier chapitre présente le contexte clinique et scientifique de ces travaux. Nous aborderons tout d'abord certaines généralités concernant le cancer du col de l'utérus, de ses origines aux résultats thérapeutiques, en passant par son dépistage clinique et les options thérapeutiques. Nous dresserons ensuite un état des lieux de l'utilisation de la tomographie par émission de positons au 18F-fluorodésoxyglucose en oncologie pour établir un état de l'art sur la prédiction de récurrence. Enfin, nous établirons les limitations méthodologiques actuelles dans l'analyse des images pré- et per-thérapeutiques et les pistes et objectifs pour lever ces verrous.*

---

## CONTENU

<b>1</b>	<b>CONTEXTE MEDICO-SCIENTIFIQUE ET PROBLEMATIQUE .....</b>	<b>9</b>
1.1	Cancer du col de l'utérus et imagerie TEP .....	11
1.1.1	Le cancer du col de l'utérus .....	11
1.1.2	La tomographie par émission de positons : Principes et définitions .....	15
1.1.3	L'utilisation standard de l'imagerie TEP en oncologie .....	20
1.1.4	L'utilisation de la TEP au 18F-FDG dans la prédiction des cancers autres que le col utérin ....	24
1.1.5	L'utilisation de la TEP au 18F-FDG dans la prédiction des cancers du col utérin.....	27
1.2	L'analyse de l'imagerie TEP pour la prédiction de récurrence.....	31
1.2.1	Prétraitements d'image dans l'imagerie TEP.....	31
1.2.2	Segmentation de l'imagerie TEP .....	33
1.2.3	Quantification et extraction de paramètres de l'imagerie TEP .....	38
1.3	Enjeux méthodologiques dans l'analyse des images TEP pour la prédiction de récurrence .....	46
1.3.1	Faible spécificité du radiotracer FDG.....	46
1.3.2	Impact de la segmentation pour la prédiction .....	46
1.3.3	Limites de la quantification de l'imagerie TEP.....	47

---



## 1.1 CANCER DU COL DE L'UTERUS ET IMAGERIE TEP

### 1.1.1 LE CANCER DU COL DE L'UTERUS

#### Généralités et épidémiologies

Le cancer est une pathologie qui désigne la prolifération anarchique de cellules au sein de l'organisme. A la différence des cellules saines, les cellules cancéreuses ont la particularité de pouvoir se multiplier sans limite et d'échapper au phénomène d'apoptose, c'est-à-dire de la mort cellulaire programmée. Ces cellules prolifèrent jusqu'à former une masse appelée tumeur maligne. Cette tumeur envahit alors peu à peu l'organe hôte, altère son fonctionnement et peut provoquer sa mort biologique. Certaines des cellules cancéreuses possèdent également la capacité de se détacher de leur tissu d'origine pour migrer vers une autre partie de l'organisme et former une nouvelle tumeur appelée métastase.

Dans cette thèse nous nous concentrerons sur le cancer développé au sein du col de l'utérus.

Le col de l'utérus (appelé aussi col utérin ou cervix) est le point de communication entre l'utérus et le sommet du vagin. Il a pour fonction de sécréter en permanence un mucus appelé glaire cervicale. Cette dernière joue un rôle important dans la reproduction et constitue une barrière de protection contre les infections. Le col utérin comprend deux parties : le canal endocervical (ou endocol) situé dans la partie haute du col et l'exocol, qui lui est situé près du vagin (Figure 1.1). Ces deux canaux sont entièrement recouverts d'une muqueuse, composée d'un tissu de surface appelé épithélium et d'un tissu conjonctif en profondeur. C'est de l'épithélium que se développe la quasi-totalité des cancers ; que l'on nomme carcinomes. Ces derniers peuvent se développer à partir de l'épithélium malpighien de l'exocol (carcinomes épidermoïde) ou à partir de l'épithélium glandulaire de l'endocol (adénocarcinomes).

Les statistiques globales sur le cancer de l'agence internationale pour la recherche sur le cancer (IARC) estime qu'en 2012, un total de 527 600 nouveaux cas de cancer du col utérin, ainsi que 265 700 décès liés à celui-ci, ont été recensés dans le monde [1]. Ce cancer représente le quatrième cancer le plus diagnostiqué et la quatrième cause de décès lié au cancer chez la femme.

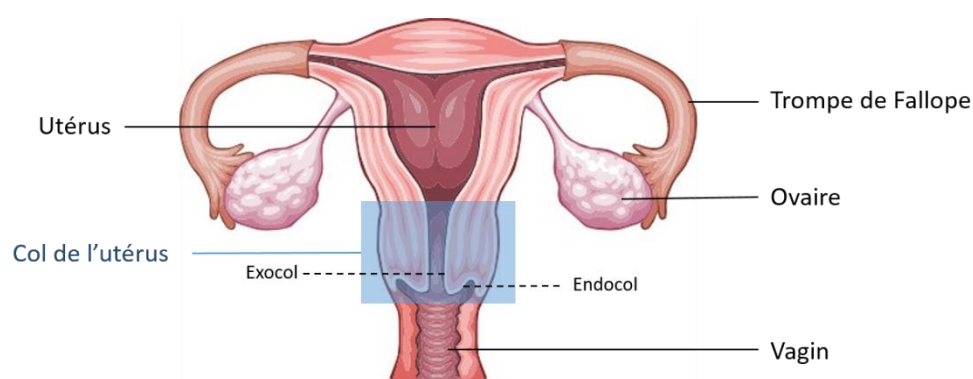


Figure 1.1 : Représentation de l'appareil reproducteur féminin

#### Dépistage, diagnostic et classification

Un cancer du col de l'utérus est suspecté si une anomalie est décelée lors d'un examen de dépistage (frottis cervico-utérin) ou si des symptômes sont apparus.

Pour établir le diagnostic, des prélèvements biopsiques sont réalisés au niveau des lésions et l'examen anatomopathologique de ces prélèvements confirme le diagnostic de cancer du col de l'utérus. L'étendue de la maladie est ensuite déterminée grâce à des examens d'imagerie par résonance électromagnétique (IRM) et par tomographie par émission de positons (TEP).

L'ensemble des examens du diagnostic permet de caractériser précisément chaque cancer et de définir notamment le type de cellules impliquées (type histologique), la profondeur de la tumeur dans la muqueuse, son extension éventuelle aux organes voisins ou aux ganglions lymphatiques proches et son extension éventuelle à des organes éloignés (métastases). Généralement, les carcinomes du col utérin sont décrits au travers de la classification de la Fédération International de Gynécologique et d'Obstétrique (FIGO) [2].

On distingue 5 stades FIGO (0, I, II, III, et IV) eux-mêmes pouvant être subdivisés en 2 sous-stades (A et B) avec différents degrés (1 et 2). Ceux-ci permettent d'évaluer de manière précise l'étendue du cancer du col :

- **Stade 0 :**  
Le carcinome de stade 0 est une forme tumorale précancéreuse (i.e. cancer in situ) qui se caractérise par l'absence d'invasion des tissus environnants.
- **Stade I :**  
Le carcinome de stade I est strictement limité au col utérin. On ne doit pas prendre en compte l'extension au corps utérin.
  - **Stade IA :** L'invasion est limitée à l'invasion stromale mesurée ne dépassant pas 5 mm en profondeur et 7 mm en largeur.
    - **Stade IA1 :** L'invasion stromale ne dépassant pas 3 mm de profondeur et 7 mm de largeur.
    - **Stade IA2 :** L'invasion stromale comprise entre 3 et 5 mm de profondeur et ne dépassant pas 7 mm de largeur.
  - **Stade IB :** Les lésions cliniques sont limitées au col ou aux lésions infracliniques plus importantes que dans le Stade IA.
    - **Stade IB1 :** Lésions cliniques de taille ne dépassant pas 4 cm.
    - **Stade IB2 :** Lésions cliniques de taille supérieure à 4 cm.
- **Stade II :**  
Le carcinome de stade II s'étend au-delà du col, mais sans atteindre les parois pelviennes. Il affecte le vagin, mais pas au-delà de ses deux tiers supérieurs.
  - **Stade IIA :** Pas d'atteinte des paramètres. L'invasion touche les deux tiers supérieurs du vagin.
  - **Stade IAB :** Atteinte des paramètres. La paroi pelvienne n'est pas touchée.
- **Stade III :**  
Le carcinome de stade III est étendu à la paroi pelvienne. Il n'existe pas de zone non envahie par le cancer entre la tumeur et la paroi pelvienne. La tumeur touche le tiers inférieur du vagin.
  - **Stade IIIA :** Pas d'extension à la paroi pelvienne, mais atteinte du tiers inférieur du vagin.
  - **Stade IIIB :** Extension à la paroi pelvienne, hydronéphrose ou rein muet.
- **Stade IV :**  
Le carcinome de stade IV s'est étendu au-delà du petit bassin ou a envahi la muqueuse de la vessie et/ou du rectum.
  - **Stade IVA :** Extension de la tumeur aux organes pelviens adjacents.
  - **Stade IVB :** Extension aux organes distants.

Au-delà de cette classification, l'usage clinique tend à utiliser les termes de stade « précoce » pour les stades IA, IB et IIA, stade « localement avancé » pour les stades IIB, III et IVA et stade « avancé » pour le stade IVB.

## Options thérapeutiques des cancers du col de l'utérus

Le traitement du cancer du col de l'utérus fait appel, selon le stade de la maladie, à la chirurgie, la radiothérapie externe, la curiethérapie et la chimiothérapie, utilisées seules ou associées.

Les différentes options thérapeutiques les plus généralement préconisées, en fonction du stade de la maladie, sont succinctement décrites ci-après.

### *La chirurgie*

La chirurgie est principalement utilisée pour traiter les tumeurs précoces de moins de 4 cm (stade IA1 à IB1). Elle a pour but de retirer la totalité de la tumeur et d'éliminer le risque de récurrence. L'intervention la plus fréquemment réalisée est la colpohystérectomie élargie qui consiste en l'ablation de l'utérus, des paramètres et de la partie supérieure du vagin. Généralement, les ovaires sont également retirés et il est souvent nécessaire de retirer les ganglions lymphatiques du pelvis (curage ganglionnaire). La colpohystérectomie peut également être précédée d'une curiethérapie afin de réduire préalablement la taille de la tumeur et faciliter son extraction.

### *La radiochimiothérapie concomitante*

La radiochimiothérapie concomitante, associant une radiothérapie externe, une curiethérapie et une chimiothérapie, est le traitement de référence des cancers du col de plus de 4 cm et/ou localement avancés (stades IB2, II, III et IVA).

La chimiothérapie repose sur l'administration par intraveineuse hebdomadaire de cisplatine (sels de platine), un médicament anticancéreux agissant sur la division incontrôlée des cellules cancéreuses (arrêt de la mitose). La chimiothérapie est un traitement systémique qui ne cible pas seulement les cellules cancéreuses. Son rôle est de potentialiser l'effet de l'irradiation, elle améliore le contrôle local et métastatique [3,4].

La curiethérapie est une technique de radiothérapie qui consiste à placer des éléments radioactifs (iridium ou césium) au contact direct de la tumeur. Du fait de sa radioactivité, les sources libèrent spontanément des rayons ionisants qui vont détruire les cellules cancéreuses de manière très localisée. La dose de rayons ionisants délivrée en radiothérapie se mesure en Gray (Gy). Dans le cas du cancer du col, les sources radioactives sont introduites directement dans le vagin et placées contre la tumeur (curiethérapie endocavitaire).

La radiothérapie externe est une technique de radiothérapie reposant sur l'utilisation d'un accélérateur linéaire de particules (photons), focalisé sur la tumeur et visant à détruire les cellules cancéreuses de l'extérieur du corps de la patiente. La source des rayons étant externe, ces derniers traversent le corps, et donc certaines cellules saines, avant d'atteindre la tumeur. La radiothérapie conformationnelle, très utilisée de nos jours, vise à délivrer précisément la dose de rayonnement vers le volume de la tumeur, tout en épargnant le plus possible les tissus sains.

L'association de ces traitements a pour avantage de renforcer l'action respective de chacun des traitements. En effet, la chimiothérapie rend les cellules plus sensibles aux rayonnements et l'action combinée de la curiethérapie et de la radiothérapie externe permet de cibler de manière précise le volume tumoral et de préserver au mieux les tissus sains.

### *La chimiothérapie exclusive*

Dans le cas des tumeurs de stade IVB, ayant atteint des organes éloignés (métastases), le traitement repose principalement sur une chimiothérapie en vue de ralentir, voire d'arrêter la progression fortement avancée de la maladie. Elle est parfois associée à une radiothérapie permettant d'attaquer fortement le foyer principal du cancer.

### Résultats carcinologiques et enjeux des traitements du cancer du col de l'utérus

A la suite de son traitement, un patient est dit en rémission si aucune cellule cancéreuse n'est décelée lors des examens médicaux. La guérison du cancer est définie par le temps ; on parle d'une rémission du cancer sans rechute dans le temps. Les résultats carcinologiques sont alors définis comme la probabilité de survie (et/ou de récurrence), à un temps donné, dans les mois ou années après le traitement. Ces résultats sont très dépendants de la situation anatomique du cancer, du grade tumoral et du volume initial, etc. De plus, ceux-ci se basent sur une analyse de groupe et ne permettent pas de prévoir les chances de survie d'une patiente en particulier.

Les résultats carcinologiques du cancer du col utérin, à cinq ans après le traitement, en fonction du stade du cancer et du traitement associé, sont présentés dans le Tableau 1.1 .

**Tableau 1.1 Résultats thérapeutiques à 5 ans en fonction du stade FIGO et du traitement associé (source : Amit et al. [5])**

FIGO Stage Criteria	Treatment Options	5-Year Survival
O. Carcinoma in situ	Conization, TAH	96%–100%
IA. Microscopic lesions	Conization, modified hysterectomy <sup>a</sup>	92%–94%
IB1. Macroscopic lesion ≤4 cm	RH, RTL, XRT/Chemo	80.7%
IB2. Macroscopic lesion ≥4 cm	RH, XRT/Chemo	79%
IIA. Vaginal involvement	RH, XRT/Chemo	76%
IIB. Parametrial involvement	XRT/Chemo	73.3%
IIIA. Lower one-third of vagina	XRT/Chemo	50.5%
IIIB. Side wall extension and/or hydronephrosis	XRT/Chemo	46%
IVA. Adjacent organ involvement	XRT/Chemo	29%
IVB. Distant metastasis	Chemo, palliative XRT	22%

*Abbreviations:* Chemo, chemotherapy; RH, radical hysterectomy; RTL, radical trachelectomy; TAH, total abdominal hysterectomy; XRT, radiotherapy.

<sup>a</sup> Individualization by patient's age and fertility desire.

La rechute de la maladie peut se manifester de différentes manières ; elle peut être locale ou locorégionale si la récurrence est située dans l'organe de la tumeur initiale et/ou dans son voisinage, et métastatiques si celle-ci a atteint un organe différent.

Dans le cas du cancer du col localement avancé, environ 30% des patientes présenteront une récurrence locale dans les deux premières années après la fin du traitement. Les facteurs prédictifs de rechute locale connus et le plus souvent utilisés sont : le stade clinique, la présence d'un envahissement ganglionnaire au diagnostic et la qualité de la réponse au traitement [6–8]. La récurrence du cancer est par nature très virulente et les tentatives de rattrapages peuvent entraîner de fortes complications [9]. C'est pourquoi les chances de survies des patientes présentant une récurrence sont généralement très faibles [10].

### 1.1.2 LA TOMOGRAPHIE PAR EMISSION DE POSITONS : PRINCIPES ET DEFINITIONS

La TEP est une technique d'imagerie fonctionnelle qui reflète l'activité métabolique des cellules tumorales. Le principe de la TEP repose sur la détection de photons de deux photons  $\gamma$ , émis après injection intraveineuse d'un traceur radioactif, émetteur de positons  $\beta^+$ .

#### Le 18F-FDG et l'émission de positons

Le radio-traceur le plus couramment utilisé en TEP est le 18F-FDG, analogue du glucose marqué au fluor. L'augmentation de la glycolyse est l'une des caractéristiques des cellules tumorales, associée à une surexpression des transporteurs transmembranaires du glucose, en particulier GLUT-1. Injecté par voie intraveineuse, le 18F-FDG est transporté dans la cellule, où il est rapidement transformé en 18F-FDG-6-phosphate, lieu où il reste piégé. Le 18F-FDG-6-phosphate s'accumule ainsi dans les cellules cancéreuses. Il s'accumule aussi dans les cellules inflammatoires qui sur-expriment aussi les transporteurs GLUT-1 et GLUT-3 [11]. Il est donc à noter que les foyers d'hyperfixations en cas de pathologie infectieuse ou inflammatoire peuvent donc être à l'origine d'erreurs d'interprétation.

A la désintégration  $\beta^+$  du fluor 18, le positon émis parcourt quelques millimètres (parcours moyens du fluor 18 : 2,6 mm) dans le tissu jusqu'à rencontrer un électron. Il se produit alors une réaction d'annihilation durant laquelle deux photons de 511 keV sont émis dans des direction opposées à 180°.

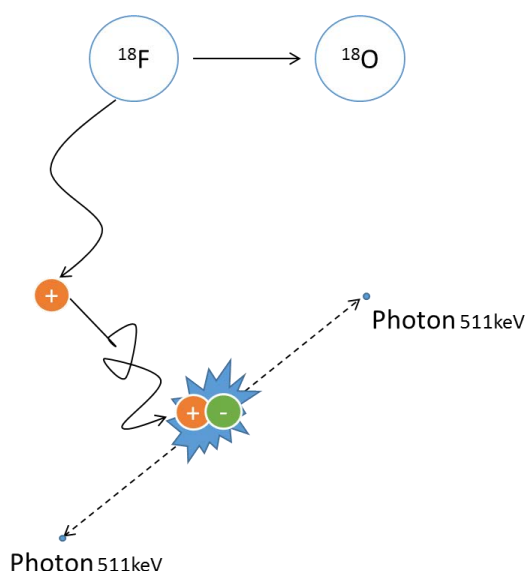


Figure 1.2 Principe de la tomographie par émission de positons au 18F-FDG

#### Détection des photons et reconstructions

C'est de la détection en coïncidence des photons émis lors de l'annihilation que la TEP tire son principe. Les photons émis à 180° l'un de l'autre sont alors détectés par les cristaux du détecteur formant un anneau autour du patient. La captation de tous les coïncidences de photons par la caméra TEP permet, après l'application d'un algorithme de reconstruction, de retrouver la localisation de chacune des interactions et de former une image volumique de celles-ci.

La caméra TEP, aussi appelé tomographe à émission de positons, est aujourd'hui généralement accompagnée d'un scanner de tomodensitométrie (TDM). L'acquisition de TDM, en plus d'apporter une information anatomique à l'information métabolique, permet également de corriger les erreurs de détection de coïncidence inhérentes à l'acquisition TEP.



La Figure 1.3 présente un exemple d'examen TEP au 18F-FDG pré-thérapeutique, fusionnée avec son image TDM, acquise pour une patiente atteinte d'un cancer du col utérin localement avancé. La zone en surbrillance correspond à l'hyperfixation de la tumeur au sein du col.

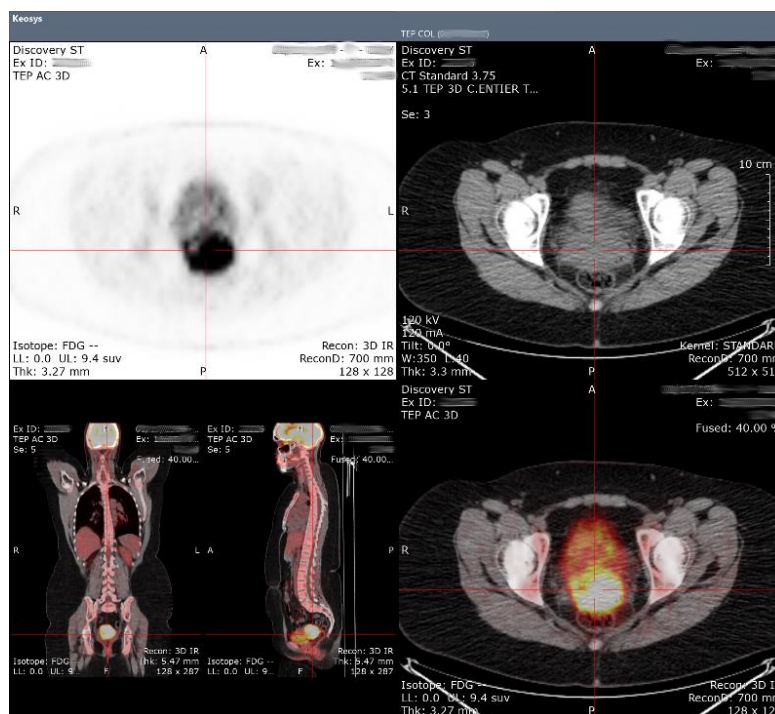


Figure 1.3 Examen TEP au 18F-FDG pré-thérapeutique, fusionné avec l'examen TDM, d'une patiente atteinte d'un cancer du col utérin localement avancé.

### Normalisation du signal de la TEP

La valeur de chaque pixel volumique de l'image (voxel) correspond à une concentration radioactive du 18F-FDG dans le volume du voxel, mesurée en Bq/mL. La fixation tissulaire du radiotracteur 18F-FDG dépend, entre autres, de la quantité de l'activité administrée, de la corpulence et de la masse grasseuse du patient, des autres organes, du taux d'insuline et du nombre de transporteurs de glucose du tissu.

Afin de normaliser la fixation du FDG, l'index le plus communément utilisé est le SUV (Standardized Uptake Value). Il s'agit d'un indice de fixation normalisée, sans dimension, qui se calcule dans un tissu d'intérêt, à un instant donné, comme le rapport de la fixation (en kBq/mL) à l'activité totale injectée (en kBq) rapporté par le volume total du patient (en mL).

$$SUV_{bw} = \frac{\text{fixation dans le tissu d'intérêt} \left( \frac{kBq}{mL} \right)}{\text{activité totale (kBq)} / \text{volume total (mL)}}$$

En pratique, le SUV est souvent calculé en rapportant la dose injectée au poids du patient (en g), on parle alors de  $SUV_{bw}$  (bw pour *Body Weight* en anglais) :

$$SUV_{bw} = \frac{\text{fixation dans le tissu d'intérêt} \left( \frac{kBq}{mL} \right)}{\text{Activité Totale (kBq)}} \times \text{poids (g)}$$

En supposant que le patient a une masse volumique de 1 g/mL, le  $SUV_{bw}$  reste donc une quantité sans dimension. Le  $SUV_{bw}$  est fréquemment utilisé car il est simple à calculer et permet de comparer de façon quantitative des examens réalisés pour un même patient, à différents temps de traitement. On notera qu'en pratique, la *Activité Totale* correspond à la dose totale injectée corrigée de sa décroissance physique entre le moment de l'injection et le moment de l'acquisition TEP en pondérant l'activité totale par un facteur  $\lambda$  calculé comme suit :

$$\lambda = 2^{\left( \frac{T}{T_{1/2}} \right)}$$

Avec  $T$  le temps passé entre l'injection et le début de l'acquisition (généralement aux alentours de 1h) et  $T_{1/2}$  le temps de demi-vie du 18F-FDG (6588 s).

En pratique le  $SUV_{bw}$  est donc calculé comme suit :

$$SUV_{bw} = \frac{\text{fixation dans le tissu d'intérêt} \left( \frac{kBq}{mL} \right)}{\text{Activité Totale (kBq)} \times \lambda} \times \text{poids (g)}$$

Nous noterons toutefois que cette mesure est très discutée quant à sa fiabilité et sa reproductibilité [12–14], notamment car elle approxime de nombreux facteurs comme l'évolution temporelle de la captation du FDG dans les cellules durant l'acquisition, la répartition non-uniforme du traceur dans la corps ou la non prise en compte de la masse grasseuse du patient dans la métabolisation du glucose.

Afin d'aborder certaines de ces points, la fixation de chaque voxel est parfois corrigée par la mesure de la masse grasseuse du patient (Lean Body Mass ou LBM en anglais) ou la surface du corps (Boy Surface Area ou BSA en anglais) plutôt que par le poids du patient. Ces mesures sont nommées respectivement  $SUV_{lbm}$  (parfois raccourcis en SUL) et  $SUV_{bsa}$  et se calcule par :

$$SUV_{lbm} \text{ ou } SUL = \frac{\text{fixation dans le tissu d'intérêt} \left( \frac{kBq}{mL} \right)}{\text{activité totale (kBq)} \times \lambda} \times LBM(g)$$

$$\text{Avec } LBM = \begin{cases} 1.10 \times \text{poids} - 120 \times \left( \frac{\text{poids}}{\text{taille}} \right)^2 & \text{pour les hommes} \\ 1.07 \times \text{poids} - 148 \times \left( \frac{\text{poids}}{\text{taille}} \right)^2 & \text{pour les femmes} \end{cases}$$

et

$$SUV_{bsa} = \frac{\text{fixation dans le tissu d'intérêt} \left( \text{en } \frac{kBq}{mL} \right)}{\text{Activité totale (en kBq)} \times \lambda} \times BSA \text{ (en } cm^2)$$

$$\text{Avec } BSA = 71.84 \times \text{taille}^{0.725} \times \text{poids}^{0.425}$$

Ces mesures sont utilisées pour les critères de l'EORTC [15] pour le  $SUV_{bsa}$  et par les critères PERCIST [16].

Aujourd'hui, le  $SUV_{bw}$  reste l'index de référence en routine clinique pour l'évaluation du métabolisme glucidique et est généralement rapporté sous l'appellation « SUV ».

### Mesures du métabolisme tumoral couramment utilisé en clinique

L'analyse du métabolisme tumoral au travers de l'image TEP-FDG s'effectuent généralement par des mesures de SUV caractérisant la lésion dans l'image.

Le premier paramètre à avoir été extraits des images TEP est basé sur la mesure de l'intensité maximale au sein de la tumeur, le  $SUV_{max}$ . Cette mesure présente l'avantage de ne nécessiter aucune segmentation au préalable. Sa simplicité d'extraction en a fait une première analyse de choix pour une mesure instantanée de l'agressivité tumorale. Toutefois, cette mesure est connue pour être très sensible au bruit, à la durée et au paramètres de l'acquisition, la rendant très peu reproductible [17–19].

Conséquemment, la mesure du pic de SUV (peak SUV ou  $SUV_{peak}$  en anglais), définie comme la moyenne des SUV dans une boule de 1,2 cm de diamètre centrée autour du  $SUV_{max}$ , pour les tumeurs de diamètre supérieur à 2 cm [16,20], a été proposée comme alternative plus robuste au bruit de l'image [21].

D'autres paramètres caractérisant la lésion sont également couramment utilisés comme le volume métabolique de la tumeur (Metabolic Tumor Volume ou MTV en anglais), la valeur moyenne de SUV au sein de la lésion (mean SUV ou  $SUV_{mean}$  en anglais) ou encore la mesure de la glycolyse totale de la lésion (*Total Lesion Glycolysis* ou TLG en anglais) [22].

La mesure de la glycolyse totale de la lésion (*Total Lesion Glycolysis* ou TLG en anglais) a été proposée pour la première fois en 1999 par Larson *et al.* [23]. Ce paramètre est défini par la multiplication du volume métabolique tumoral et la valeur moyenne d'hyperfixation au sein de la lésion.

Le calcul des paramètres TEP les plus utilisés pour mesurer l'hyperfixation tumorale est décrit dans le Tableau 1.2.

Tableau 1.2 : Paramètres les plus couramment extraits des images TEP

Nom du paramètres	Equation	Définition
<b>SUVmax</b>	$Max(I(X))$	Valeur de l'hyperfixation maximale à l'intérieur de la lésion
<b>SUVpeak</b>	$SUV_{peak} = \frac{1}{N_{max}} \cdot \sum_{i \in S_{max}} I(x_i)$	Le pic de SUV. $SUV_{max}$ moyenné par ses voisins (se trouvant dans une sphère de 1.2cm)
<b>SUVmean</b>	$\frac{1}{N} \cdot \sum_i I(x_i)$	Moyenne d'hyperfixation dans la région tumorale
<b>MTV</b>	$Vol_{voxel} \cdot N$	Volume de la lésion
<b>TLG</b>	$TLG = SUV_{mean} \times MTV = Vol_{voxel} \cdot \sum_i I(x_i)$	Glycolyse totale de la lésion. Correspond à la somme des valeurs de voxels au sein de la région
Avec : $X = \{x_1, x_2, \dots, x_i, x_N\}$ , les N voxels dans le volume d'intérêt, $I(x_i)$ l'intensité SUV du voxel $x_i$ $Vol_{voxel}$ le volume d'un voxel et $S_{max}$ les $N_{max}$ voxels se trouvant dans le voisinage de $SUV_{max}$ et dans une sphère de 1,2cm		

### Les limitations de l'imagerie TEP

Malgré son utilisation en constante évolution dans le domaine de l'oncologie, l'analyse des images TEP reste sujette à de nombreuses contraintes et limitations autant technologiques et méthodologiques que scientifiques.

En effet, l'évaluation de la captation glucidique est affectée par certains facteurs inhérentes à la nature de l'acquisition TEP comme le bruit, la faible résolution spatiale, l'effet de volume partiel et le faible contraste rencontré.

#### *Le bruit dans l'imagerie TEP*

Le bruit dans l'imagerie TEP provient tout d'abord de la nature stochastique de la mesure de concentration d'activité radioactive du radiotracteur, couramment modélisée par une loi de poisson. La deuxième source de bruit provient (et dépend) de la reconstruction des images TEP. Ce bruit est généralement modélisé par un bruit gaussien.

#### *La faible résolution de l'imagerie TEP*

La faible résolution spatiale est également une limitation forte de l'imagerie TEP. Les techniques d'acquisitions TEP actuels ne permettent pas d'obtenir une résolution en dessous de 5 mm. En comparaison aux autres modalités telles que la TDM ou l'IRM, dont les résolutions spatiales peuvent être submillimétriques. La première raison tient au fait du parcours aléatoire du positon à la désintégration  $\beta^+$  du fluor 18, dont déplacement moyen est 2,6 mm. De plus, les deux photons émis suite à l'annihilation de ce positon avec un électron ne sont systématiquement entièrement colinéaires, la variabilité peut aller jusqu'à  $0.5^\circ$ . Une autre raison réside dans les caractéristiques physiques de la caméra TEP, dont le positionnement, le matériau et la taille des cristaux peuvent influencer la résolution spatiale.

#### *L'effet de volume partiel*

L'effet de volume partiel est une conséquence de la résolution finie de l'imagerie TEP et correspond à la diffusion floue des frontières de structures ainsi qu'au mélange de leurs signaux au sein de l'image. Le système d'imagerie TEP est généralement modélisé par la convolution d'une image avec une fonction d'étalement de point (*Point Spread Function* ou PSF en anglais). Cette convolution entraîne donc un étalement du signal dans l'image induisant des structures aux frontières floues et une baisse de contraste dans l'image.

### 1.1.3 L'UTILISATION STANDARD DE L'IMAGERIE TEP EN ONCOLOGIE

Développée initialement pour l'étude du cerveau et du cœur, l'imagerie TEP est aujourd'hui devenue un outil de référence en oncologie. L'analyse des images TEP en générale, fait l'objet de très nombreuses études dans la littérature. Dans [24], Foster *et al.* ont dénombré plus DE 1 700 000 études cliniques portant sur cette modalité de par le monde en 2011. L'utilisation de la TEP au 18F-FDG en cancérologie a fait l'objet de Standards, Options et Recommandations (SOR) en 2003 sous l'égide de la Fédération nationale des centres de lutte contre le cancer (Bourguet *et al.*) publié 2002 puis actualisée en 2003 [25].

Parallèlement, la Société Française de Radiologie (SFR) et la Société Française de Biophysique et de Médecine Nucléaire (SFBMN) ont rédigé en 2005 un premier guide du bon usage des examens d'imagerie (GBU), mis à jour en 2013 à partir de différentes revues de la littérature et méta-analyses concernant la TEP-FDG ( <http://gbu.radiologie.fr> ).

Ces recommandations sont basées sur une étude systématique de la littérature et permettent d'établir un état de l'art de l'utilisation de l'imagerie TEP en oncologie et cela en fonction de la localité du cancer.

Globalement, l'imagerie TEP en oncologie est, ou peut-être, utilisée tout le long de la prise en charge du patient, du diagnostic de la maladie primitive au suivi clinique post-thérapeutique :

- pour le diagnostic du cancer :
  - diagnostic de la maladie primitive,
  - bilan d'extension (recherche précise de l'étendue du cancer et d'éventuelles métastases) ;
- lors du traitement radiothérapeutique :
  - Pour le ciblage radio-thérapeutique,
  - Pour l'évaluation de la réponse tumorale au traitement ;
- lors du suivi clinique post-traitement :
  - Pour la détection de maladie résiduelle et de récurrence,
  - Pronostic du devenir du patient (récidive et/ou survie globale).

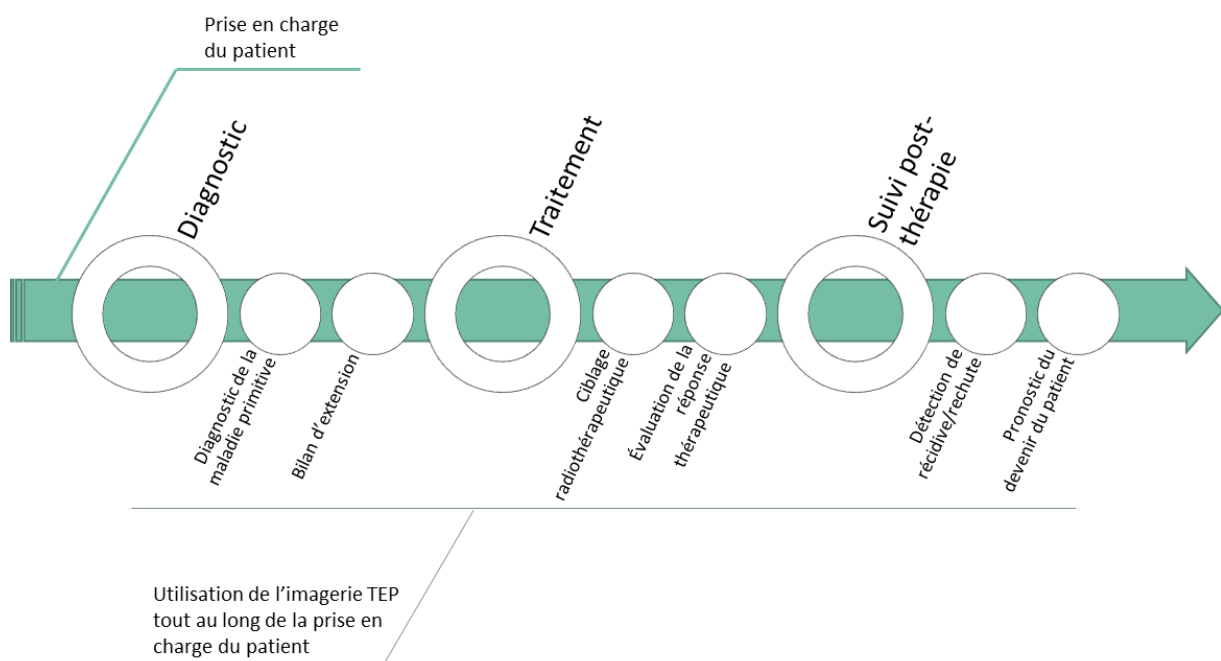


Figure 1.4 La TEP au 18F-FDG et son utilisation en cancérologie

### L'imagerie TEP au 18F-FDG pour le diagnostic, le bilan d'extension et la détection de récurrence

Profitant d'une information directement reliée à l'activité glucidique des cellules, l'imagerie TEP permet de discerner rapidement la présence d'une activité métabolique tumorale. La TEP-FDG a donc rapidement été adoptée pour le diagnostic, la stadification et l'évaluation des rechutes post-traitement dans diverses localités de cancers. Il est cependant important de remarquer que le métabolisme du FDG n'est pas spécifique aux cellules cancéreuses. Dans leur méta-analyse de la littérature sur le sujet [26], Czernin *et al.* ont montré que l'utilisation associée de la TEP et de la TDM permettait d'obtenir de meilleures performances de diagnostic qu'avec les images prises séparément.

Dans la localité du col, la 18F-FDG-TEP a montré une bonne sensibilité de détection de la tumeur cervicale, quel que soit leur type histologique, mais elle reste insuffisante pour le bilan d'extension local, où l'IRM reste l'examen de référence pour préciser la topographie, les dimensions de la tumeur ainsi que son extension locorégionale au sein du pelvis. Sa valeur dans le bilan d'extension ganglionnaire a toutefois été confirmée dans une méta-analyse, effectuée par Choi *et al.* en 2010, qui comparait les performances diagnostiques de la TEP avec la TDM et l'IRM pour la stadification ganglionnaire [27].

Pour les Stades < IB2 la sensibilité de la TEP-TDM est de 10% pour la détection des métastases ganglionnaires pelviennes (90% de faux négatif) [28]. La TEP TDM s'avère incapable de détecter de micro métastases. Si l'on considère les métastases > 5mm la sensibilité et la spécificité sont respectivement de 100% et 99.6% [29,30]. La lymphadénectomie pelvienne reste indispensable.

Pour les stades > IB2 la présence d'adénopathies, en particulier au niveau lomboaortique, est l'un des facteurs pronostiques majeurs, les taux de survie à 5 ans rapportés pour des patientes présentant des adénopathies lomboaortiques envahies varient entre 12 et 30 % [7]. Le risque d'envahissement ganglionnaire lomboaortique est corrélé au stade FIGO (5% pour les stades IB, 16% pour les stades II et 25% dans les stades III) et à la présence de ganglions pelviens (26% en cas d'adénopathies pelviennes vs 0,9% en l'absence de ganglions pelviens) [31,32]. L'enjeu est donc d'évaluer de manière fiable le statut ganglionnaire et donc d'éviter la réalisation d'un curage lomboaortique qui peut se révéler morbide. En 2005, Havrilevsky *et al.* ont rapporté dans une méta-analyse une sensibilité et une spécificité poolées de la TEP au 18F-FDG pour la détection des ganglions lomboaortiques (n=136) de 84 % et 95% respectivement. La TEP TDM a montré sa supériorité sur l'IRM pour l'évaluation de l'atteinte lomboaortique [33,34].

Les SOR de 2003 ainsi que le GBU de 2005 recommandent par ailleurs d'utiliser la TEP-FDG en option en cas de récurrence avérée du cancer du col utérin, en particulier pour décider de la stratégie thérapeutique.

### Le ciblage de la radiothérapie

La TEP au 18F-FDG fournit des informations biologiques sur les voies métaboliques impliquées dans les phénomènes de radiorésistance, dont fait partie le métabolisme glucidique. Celle-ci apparaît donc comme particulièrement intéressante pour la délimitation du volume tumoral macroscopique dans la problématique de ciblage radiothérapeutique. La projection des images TEP-FDG sur des images TDM de dosimétrie a été étudiée pour déterminer si la perception de la tumeur par la TEP permettait d'optimiser la définition des champs d'irradiation compte tenu de la meilleure spécificité de la modalité.

Dans leur revue sur le sujet [35], Rossier *et al.* ont montré que l'association des informations de la TEP et de la TDM pourrait permettre soit d'améliorer la définition des volumes irradiés, soit d'optimiser l'irradiation au sein de la tumeur (voir Tableau 1.3). Ainsi, la TEP à visée de ciblage thérapeutique a un rôle déjà bien établi dans certaines maladies comme le cancer du poumon, et un rôle prometteur pour d'autres (cancers ORL, digestifs, de la prostate, du système nerveux central, et gynécologiques).

Dans la localité du col utérin, l'utilisation du volume métabolique tumoral analysé sur imagerie TEP a permis d'améliorer la couverture tumorale lors du ciblage ainsi que la détection des adénopathies sans augmenter l'irradiation des organes à risques.

On notera que l'utilisation de la TEP dans la problématique de ciblage nécessite une délimitation très fine du volume tumoral visible sur les images TEP, nécessitant la définition d'une méthode de segmentation validée, fiable et reproductible. De nombreuses méthodes de segmentation ont été proposées dans la littérature à ce sujet [36–39]. Ces méthodes feront l'objet d'une analyse plus détaillée dans la suite de ce manuscrit.

**Tableau 1.3 Apport de la TEP au 18F-FDG couplé à la TDM sur la détermination des volumes à irradier et sur la protection des organes à risques (Rossier *et al.* 2012 [35])**

Indication (tumeur)	Avantage volume tumoral biologique (TEP) sur tumeur	Effets sur les organes à risque
<i>Poumon</i>		
Cancer du poumon non à petites cellules	Haute sensibilité Diminution du volume tumoral biologique [8,9]	Diminution de l'irradiation [8,9]
Cancer du poumon à petites cellules	À déterminer	À déterminer
<i>Digestif</i>		
Œsophage	Données contradictoires [19–21]	Données contradictoires [19–21]
Rectum	Meilleure corrélation avec la pathologie [27]	À déterminer
Canal anal	Meilleure identification des adénopathies inguinales [31]	À déterminer
Pancréas	Meilleure détection de la tumeur et des adénopathies [34]	Augmentation de l'irradiation mais sans augmentation significative des toxicités [34]
<i>Prostate</i>	<i>Gold-standard</i> pour la détection de récurrence après chirurgie radicale ou radiothérapie [46]	À déterminer
<i>ORL</i>	Diminution du volume tumoral macroscopique avec bonne corrélation avec la pathologie [57]	À déterminer
<i>Système nerveux</i>	Augmentation du volume tumoral biologique avec bonne reproductibilité [64]	Augmentation de l'irradiation mais sans augmentation significative des toxicités [64]
<i>Gynécologie</i>		
Col utérin	Meilleures couverture tumorale [74,75] et détection des adénopathies	Pas d'augmentation d'irradiation des organes à risque [75]
Endomètre	À déterminer	À déterminer
<i>Lymphomes</i>		
Hodgkin	Modifications du volume tumoral biologique [82,83]	À déterminer
Lymphome non hodgkinien	À déterminer	À déterminer

### L'évaluation de la réponse thérapeutique et suivi du patient

L'imagerie TEP au 18F-FDG est également très étudiée pour l'évaluation précoce de la réponse thérapeutique, en complément de la TDM. En effet, les effets du traitement radiothérapeutique, qui passe par une altération du métabolisme des cellules cancéreuses, sont très précocement observables en imagerie TEP.

C'est pourquoi de nombreuses études se sont penchées sur l'extraction de paramètres caractérisant le métabolisme tumoral – et son évolution en cours de traitement, – permettant de quantifier la réponse de la tumeur à la radiothérapie [16,40–43]. Les critères de réponse les plus connus sont les critères de l'EORTC [15] et les critères PERCIST [16] parus respectivement en 1999 et 2009.

Les premiers critères d'évaluation de l'évolution des tumeurs utilisant l'imagerie TEP ont été proposés par l'EORTC en 1999 [15]. Ceux-ci étaient basés sur des critères de classification de la variation du SUV entre deux examens (généralement pré- et post-traitement) pour l'évaluation de la réponse tumorale au traitement. La variation, noté  $\Delta\text{SUV}$ , est définie par :

$$\Delta\text{SUV} = \frac{\text{SUV}_{T_1} - \text{SUV}_{T_2}}{\text{SUV}_{T_1}} \times 100$$

Où  $\text{SUV}_{T_1}$  et  $\text{SUV}_{T_2}$  correspondent respectivement aux mesures de SUV au premier et au deuxième examen (généralement pré- ou post-thérapeutique). La mesure du SUV peut être prise sur l'intensité maximale ( $\text{SUV}_{\text{max}}$ ) ou moyenne ( $\text{SUV}_{\text{mean}}$ ) au sein de la tumeur. Les critères de classification, mesurés en fonction du  $\Delta\text{SUV}$ , sont : la réponse complète, la réponse partielle, maladie stable et maladie en progression. Ces critères ont cependant été établis sur un nombre limité d'études disponibles à cette époque ne concernant que quelques localités de cancers.

Profitant d'une littérature plus élargie, les critères PERCIST [16] ont été définis en 2009 afin d'améliorer les critères de classification de la réponse tumorale. Aujourd'hui, ces critères sont majoritairement utilisés dans



la recherche clinique mais très peu en routine car l'acquisition d'exams supplémentaires (notamment per-thérapeutique) n'est pas une application standard et nécessite davantage de validations pour être généralisée.

#### 1.1.4 L'UTILISATION DE LA TEP AU 18F-FDG DANS LA PREDICTION DES CANCERS AUTRES QUE LE COL UTERIN

Récemment, l'imagerie TEP a également été étudiée pour sa valeur pronostique sur la réponse thérapeutique de la tumeur et, plus globalement, sur le devenir du patient.

En effet l'activité métabolique des tumeurs ainsi que son évolution au cours du traitement reflète son agressivité ainsi que sa résistance au traitement. La mesure quantifiée de l'information présente dans l'imagerie TEP pourrait donc permettre d'obtenir des informations déterminantes sur la nature même de la tumeur, de prédire la réussite ou l'échec du traitement et de l'adapter en vue, *in fine*, d'améliorer la prise en charge thérapeutique du patient.

De nombreuses études ont utilisées la TEP au 18F-FDG pour quantifier la réponse tumorale à un traitement par chimiothérapie [42,44-46]. Dans ces études, les modifications de captation du FDG, mesurées par une analyse visuelle du SUV, étaient soit prédictives du devenir du patient en terme de contrôle local ou de survie globale pour des tumeurs œsophagiennes et pulmonaires [42,45,46], soit prédictives de la réponse histologique après chimiothérapie néoadjuvante pour des cancers du sein [44].

Aujourd'hui, les principaux paramètres étudiés comme potentiels facteurs prédictifs de récurrence sur une TEP sont le SUVmax, le SUVpeak, le MTV et le TLG. La Figure 1.5 présente un tableau tiré d'une revue systématique de la littérature synthétisée par Moon *et al.* [47] en 2013. Cette étude montre que les paramètres MTV et TLG sont les paramètres ressortant le plus souvent comme étant prédictif dans de nombreuses localités de cancer, notamment dans les cancers ORL et pulmonaire. Figure 1.6 présente un tableau de revue systématique de la littérature de 2014 focalisé sur les cancers de la tête et du cou [48]. Pour cette localité particulièrement, le TLG et le MTV semblent être des facteurs fortement pronostic du devenir des patients.

**Table 1. Summary of Literature on Use of Volume-Based PET Parameters in Predicting Prognosis**

Authors (Reference)	Modality	Subject No.	Tumor	Study Design	Used PET Parameters	Target Lesion	Significant Prognostic Predictor
Arslan N, et al. (30)	PET	25	Small cell lung cancer	Retrospective	SUV <sub>max</sub> , SUV <sub>mean</sub> , MTV, TLG	Tumor sites with visibly increased FDG uptake	TLG
Oh JR, et al. (61)	PET/CT	106	Small cell lung cancer	Retrospective	SUV <sub>max</sub> , MTV	All malignant lesions	MTV
Zhu D, et al. (62)	PET/CT	98	Small cell lung cancer	Retrospective	SUV <sub>max</sub> , SUV <sub>mean</sub> , MTV, TLG	All malignant lesions	MTV, TLG
Liao S, et al. (31)	PET/CT	169	Non-small cell lung cancer	Retrospective	SUV <sub>max</sub> , SUV <sub>mean</sub> , MTV, TLG	Primary tumor, nodal metastases, distant metastases	MTV and TLG
Kim K, et al. (25)	PET/CT	91	Non-small cell lung cancer	Retrospective	SUV <sub>max</sub> , MTV, TLG	Primary tumor	MTV
Chu KP, et al. (33)	PET/CT	51	Head-and-neck cancer	Retrospective	SUV <sub>max</sub> , MTV, Velocity (change in MTV or SUV <sub>max</sub> over time)	Primary tumor and nodal metastasis	MTV velocity
Choi K, et al. (36)	PET/CT	56	Head-and-neck cancer	Retrospective	SUV <sub>peak</sub> , MTV	Primary tumor and nodal metastasis	MTV
La TH, et al. (34)	PET/CT	85	Head-and-neck cancer	Retrospective	SUV <sub>max</sub> , MTV	Primary tumor and nodal metastasis	MTV
Chan SC, et al. (32)	PET	196	Nasopharyngeal carcinoma	Prospective	SUV <sub>max</sub> , MTV, TLG	Primary tumor and nodal metastasis	TLG
Xie P, et al. (35)	PET/CT	41	Nasopharyngeal carcinoma	Retrospective	MTV, TLG	Not available	TLG
Moon SH, et al. (9)	PET/CT	69	Tonsil cancer	Retrospective	SUV <sub>max</sub> , SUV <sub>mean</sub> , MTV, TLG	Primary tumor	TLG
Hatt M, et al. (27)	PET/CT	50	Esophageal cancer	Retrospective	SUV <sub>max</sub> , SUV <sub>peak</sub> , MTV, SUV <sub>mean</sub> , TLG	Primary tumor	MTV and TLG
Hyun SH, et al. (17)	PET	151	Esophageal cancer	Retrospective	SUV <sub>max</sub> , MTV	Primary tumor	MTV
Lee HY, et al. (44)	PET/CT	13	Malignant pleural mesothelioma	Retrospective	SUV <sub>max</sub> , SUV <sub>mean</sub> , MTV, TLG	Primary tumor	MTV and TLG
Chung HH, et al. (39)	PET/CT	55	Ovarian cancer	Retrospective	SUV <sub>max</sub> , SUV <sub>mean</sub> , MTV, TLG	Tumor sites with visibly increased FDG uptake	MTV and TLG
Kim BS, et al. (40)	PET/CT	45	Cervical cancer	Retrospective	SUV <sub>max</sub> , MTV	Primary tumor	MTV
Costelloe CM, et al. (42)	PET/CT	31	Osteosarcoma	Retrospective	SUV <sub>max</sub> , TLG	Primary tumor	SUV <sub>max</sub> and TLG
Yoo J, et al. (43)	PET/CT	44	Gallbladder cancer	Retrospective	SUV <sub>max</sub> , SUV <sub>mean</sub> , MTV, TLG	Primary tumor	TLG
Summary		1376					

**Note.**— PET = positron emission tomography, SUV<sub>max</sub> = maximum standardized uptake value, SUV<sub>mean</sub> = mean SUV, SUV<sub>peak</sub> = peak SUV, MTV = metabolic tumor volume, TLG = total lesion glycolysis, FDG = fluorodeoxyglucose

**Figure 1.5** Tableau résumant les paramètres TEP pronostic du devenir du patient dans différentes localités de cancer. Ce tableau est extrait de la revue systématique de la littérature publié par Moon *et al.* [47]

Study	Year	Country	Study design	Quality score (%)	No. of patients	Disease	TNM staging	Treatment	Endpoints	Volumetric parameters	VOI	Tumor delineation (thresholds)	Determination of cutoff values	Cutoff values	
														MTV (cm <sup>3</sup> )	TLG
Chung et al. (16)	2009	Korea	R	80	82	N/O/H	I-IV	RTx/CCRTx	DFS	MTV	T + LN	SUV2.5	Minimum P value method	40	
Xie et al. (8)	2009	China	R	72.5	41	N	I-IV	RTx/CCRTx	DFS/OS	MTV/TLG	T	SUV2.5	ROC	30	130
Kim et al. (15)	2011	Korea	R	72.5	69	Or/O/H/L	III/IV	op + RTx/op + CCRTx	DFS/OS	MTV	T + LN	SUV2.5	Minimum P value method	41	
Chan et al. (19)	2011	Taiwan	P	85	196	N	III/IV	CCRTx	DFS/OS	MTV/TLG	T	SUV2.5	Minimum P value method	45	330
Park et al. (2)	2013	Korea	R	82.5	81	H/L	III/IV	op/op + RTx/op + CTx	LRC/OS	MTV	T	SUV2.5	ROC	18	
Kao et al. (17)	2012	Taiwan	R	80	64	O/H	II-IV	CCRTx	DFS	MTV	T	SUV2.5	Median	13.6	
Chang et al. (18)	2012	Taiwan	P	80	108	N	I-IV	RTx/CCRTx	DFS/OS	TLG	T	SUV2.5	Median	65	
Dibble et al. (20)	2012	United States	R	75	45	Or/O	I-IV	op/CCRTx/RTx/CTx	OS	MTV/TLG	T	Gradient segmentation	Median	7.7	55
Lim et al. (21)	2012	United States	R	82.5	176	O	I-IV	CCRTx	OS	MTV/TLG	T	SUVmax42%	Median	9.6	79.8
Lee et al. (22)	2012	Korea	R	85	57	Or	I-IV	op/op + CTx	OS	MTV/TLG	T	SUV2.5	Minimum P value method	7.78	
Abd El-Hafez et al. (23)	2013	Taiwan	P	85	126	Or	II-IV	op/op + RTx/op + CCRTx	DFS	MTV/TLG	T	SUV3.0	Median	11.2	71.4
Ryu et al. (24)	2013	Korea	R	82.5	49	S	I-IV	op/op + RTx/op + CCRTx	PFS/OS	MTV/TLG	T + LN	SUVmax30%	ROC	17.7	56.3
Garsa et al. (25)	2013	United States	R	70	86	O	I-IV	RTx	DFS/OS	MTV/TLG	T + LN	SUVmax50%	ROC	20.5	

R = retrospective; N = nasopharynx; O = oropharynx; H = hypopharynx; RTx = radiotherapy; CCRTx = concurrent chemoradiotherapy; DFS = disease-free survival; T = tumor; Or = oral cavity; L = larynx; op = surgery; P = prospective; LRC = locoregional control; CTx = chemotherapy; S = salivary gland; PFS = progression free survival.

Figure 1.6 Tableau des valeurs pronostics des paramètres MTV et TLG dans les cancers de la tête et du cou extrait d'une revue systématique de la littérature de Pak *et al.* [48].

### 1.1.5 L'UTILISATION DE LA TEP AU 18F-FDG DANS LA PREDICTION DES CANCERS DU COL UTERIN

Concernant la localité du col, différents paramètres ont été étudiés à partir de la TEP au 18F-FDG, réalisée lors du diagnostic, afin de rechercher une corrélation avec le devenir des patientes : volume de la tumeur cervicale, site d'hyperactivité métabolique (ganglions pelviens, ganglions lombo-aortiques, à distance), intensité de la captation du glucose quantifiée par le SUV maximal (SUVmax) au niveau de la tumeur ou des ganglions, glycolyse totale de la lésion (TLG), etc. [49–53].

Les études publiées dans la littérature concernant l'analyse de l'images TEP au 18F-FDG pour la prédiction de récurrence post-traitement des cancers du col de l'utérus sont résumées dans le Tableau 1.4.

Dès 2002, Miller et Grigsby ont montré chez 15 patientes présentant un cancer du col localement avancé, une différence significative en survie sans récurrence et en survie globale, si le volume tumoral (MTV), calculé par seuillage automatique à 40% de l'intensité maximale (40% du SUVmax), était supérieur à 60 cm<sup>3</sup> [51].

Plusieurs études ont démontré que l'intensité de fixation du glucose, traduite par le SUVmax, reflète l'agressivité tumorale. Xue *et al.* ont ainsi montré qu'un SUVmax cervical inférieur à 10,2 sur la TEP au 18F-FDG initiale était associé à un meilleur pronostic chez les patientes irradiées [52].

Dans une étude de 2007, Kidd *et al.* ont confirmé l'impact du SUVmax tumoral sur le risque de récurrence locale et de décès [50]. À partir d'une analyse rétrospective de 287 patientes traitées pour un cancer du col utérin de stade IA2 à IVB, Kidd *et al.* ont déterminé trois groupes pronostiques en fonction du SUVmax tumoral initial. La survie globale à 5 ans était de 95 % pour un SUVmax ≤ 5,2, 70 % pour un SUVmax entre 5,2 et 13,3 et de 44 % pour un SUVmax > 13,3 respectivement.

L'impact du SUVmax mesuré cette fois au niveau ganglionnaire a été rapporté par Yen *et al.* en 2008, dans une série rétrospective de 70 patientes prises en charge pour un cancer du col utérin avec des adénopathies pelviennes ou lombo-aortiques, révélées par IRM ou TDM [53]. La présence d'un SUVmax supérieur à 3,3 au niveau des adénopathies lombo-aortiques était associé avec une baisse significative de la survie sans récurrence à 5 ans et de la survie globale à 5 ans.

Dans leur étude rétrospective incluant 75 patientes traitées pour un cancer du col utérin allant du stade IB au stade IIA, Chung *et al.* en 2010 ont montré qu'un SUVmax ≥ 7,3 mesuré au sein de la tumeur initiale était associé à un risque de récurrence accrue.

Les SUVmax initiaux intratumoral et ganglionnaire semblent donc refléter l'agressivité tumorale et fournir des facteurs précieux pour la prédiction, cependant il est difficile de retenir un seuil précis utilisable par d'autres équipes. En effet, le SUV est une variable dépendante de la machine utilisée et de nombreux paramètres : poids, activité métabolique injectée, ce qui explique les difficultés à définir un seuil applicable d'un centre à l'autre [54]. Il est donc nécessaire pour chaque équipe de rechercher le seuil de SUVmax pouvant être prédictif du devenir des patients et tenant compte des caractéristiques de l'équipement utilisé.

Pour limiter les difficultés de reproductibilité de mesure du SUV, Miller *et al.* ont proposé une évaluation visuelle simple, basée sur la présence de ganglions pelviens ou lomboaortiques et sur des critères morphologiques de la tumeur (hétérogénéité, taille, forme de la fixation) pondérés dans un score pronostique [55]. Un score supérieur à 4 était associé avec une baisse significative de la survie globale.

Dans une étude rétrospective concernant 101 patientes, l'analyse visuelle simplement basée sur les sites d'hyperactivité métabolique a montré que la présence de ganglions lomboaortiques hyperfixants en TEP au 18F-FDG était le facteur pronostique le plus puissant, en analyse multivariée, de survie sans récurrence, après exclusion des patientes métastatiques [49].

L'importance de l'évaluation ganglionnaire a été confirmée par Singh *et al.* chez 47 patientes traitées pour un cancer de stade IIIB [56]. La survie à 3 ans était de 73 % chez les patientes sans métastase ganglionnaire, de 58 % chez celles ayant des métastases pelviennes, 29 % pour les patientes ayant à la fois des métastases lomboaortiques et pelviennes et de 0 % chez les patientes ayant des métastases sus-claviculaires. L'analyse visuelle simplifiée, basée sur des critères morphologiques tumoraux et sur les sites ganglionnaires hyperfixants, semble prédictive de la survie globale et de la survie sans récurrence et permet de s'affranchir des problèmes de reproductibilité liés à l'utilisation du SUV.

Cependant, les mesures visuelles nécessitent une certaine expertise et peuvent être source d'une variabilité inter-observateur, notamment en ce qui concerne les études multicentriques. Plus récemment, d'autres équipes se sont donc penchées sur l'analyse de paramètres TEP, extraits au travers de différentes méthodes de segmentation plus automatiques, afin de permettre une évaluation reproductible entre les centres et robuste aux variations inter-patients.

En utilisant un seuillage d'intensité  $SUV > 2,5$ , Chung *et al.* en 2011 [57] ont évalué la valeur pronostic du MTV et du SUVmax pour prédire la récurrence. Basé sur une cohorte de 63 patientes (stade FIGO IB-IIA), l'étude a montré que le groupe de patientes avec un  $MTV \geq 23.4$  connaissait un risque de récurrence plus important ( $p=0.004$ ). Parallèlement, l'étude montrait également que le  $SUVmax \geq 9.5$  était également un facteur pronostic de récurrence ( $p=0.040$ ). En analyse multivariée, seul le MTV et l'âge des patientes se révélaient être des facteurs indépendants pour la prédiction de récurrence.

En utilisant cette fois une segmentation à 40% du SUVmax sur une cohorte de 45 patientes (stade FIGO IA-IIIB), Kim *et al.* [58] ont montré qu'un  $MTV \geq 20\text{cm}^3$  était associé à un risque plus fort de récurrence. Le MTV était encore une fois le seul facteur indépendant en analyse multivariée.

L'intérêt du MTV, extraits au travers d'une segmentation à 40% du SUVmax, a ensuite été confirmé en 2012 par Kidd *et al.* [59] dans une étude incluant 234 patientes traitées pour un cancer du col utérin (stage FIGO IB1-IVA) analysé au travers de leur examen TEP/TDM pré-thérapeutique. Dans cette étude, les auteurs ont proposé 3 nomogrammes permettant d'estimer, à 3 et 5 ans, les chances de survie sans récurrence, de survie sans maladie et de survie globale de chaque patiente en fonction des paramètres de SUVmax, MTV et d'une stadification de ganglions lymphatiques (stade défini par la même équipe dans [60]).

En parallèle de la mesure du volume métabolique, diverses études ont investigué l'utilisation de paramètres mesurant l'intensité totale au sein de la région tumorale tels que le SUVmean ou le TLG.

En 2011, Yoo *et al.* [61] ont étudié la valeur prédictive des paramètres SUVmax, SUVmean, MTV et TLG. Ces quatre paramètres ont été extraits des images TEP pré-thérapeutique de 73 patientes atteintes d'un cancer du col utérin de stade FIGO allant de I à IV. Pour chacune de ces patientes, la tumeur initiale était segmentée à l'aide d'un seuillage calculé à partir d'une mesure de SUV au sein de la région médiastinale. En analyse univariée, le MTV (supérieur à  $82\text{cm}^3$ ) et le TLG (supérieur à 7600 g) et la présence ou non de ganglions lymphatiques (mesurée par TEP ou IRM/TDM) étaient des facteurs pronostiques de récurrence. En analyse multivariée, seuls le TLG et la présence de ganglions lymphatiques étaient des facteurs indépendants pour prédire la récurrence. Dans cette étude, le SUVmax et le SUVmean n'étaient pas prédictifs de la récurrence.

Récemment, Miccò *et al.* ont effectué une étude sur la valeur prédictive des paramètres SUVmax, MTV et TLG [62]. L'étude se basait sur une cohorte de 49 patientes (stade FIGO IB-IVB) dont chaque tumeur initiale avait été segmentée sur image TEP à l'aide d'un seuillage à 42% du SUVmax. Les événements à prédire étaient la survie sans récurrence ainsi que la survie globale. Les résultats de leur étude montraient que le TLG était prédictif de la récurrence et que le MTV était prédictif de la récurrence et de la survie globale. Dans cette étude, le SUVmax ne montrait pas de corrélation significative avec ces événements.

À l'inverse, dans une cohorte de 89 patientes de stade FIGO IB1 et IIA, Crivellaro *et al.* [63] n'ont pas constaté de corrélation statistique entre le SUVmax, le SUVmean, le MTV et le TLG et l'apparition de récurrence post-traitement. Cette différence avec les études précédemment citées peut s'expliquer par des caractéristiques de population différentes et l'utilisation d'un seuillage adaptatif (PETVCAR, GE Healthcare) pour le contourage des lésions.

À la lumière de ces études, les images TEP montrent une importante capacité à prédire les événements de récurrence post-traitement ou à évaluer les chances de survie des patientes traitées pour un cancer du col de l'utérus. Cependant, les mesures d'hyperfixation tumorale au sein des images peuvent afficher des résultats contradictoires entre les études du fait d'un nombre de patientes et d'événements différents, d'un temps de recul trop court, du stade FIGO considéré, ou encore de la méthode de segmentation utilisée.

Tableau 1.4 Etudes publiées dans la littérature concernant l'analyse de l'images TEP pour la prédiction de récidence post-traitement des cancers du col de l'utérus. (DFS), survie sans maladie ; (OS), survie globale

Étude	Année	Nb patientes	Stade FIGO	Méthode de segmentations	Paramètres extraits	Évènement	Paramètre prédictif
Miller <i>et al.</i> [51]	2002	51	IB1-IA1	Threshold at 40% SUVmax	MTV	DFS, OS	MTV>60cm <sup>3</sup>
Miller <i>et al.</i> [55]	2003	47	IB1-IA1	visual analyse	Visual positive uptake	DFS, OS	Visual positive uptake
Ryu <i>et al.</i> [64]	2003	249	Ib and Ila	visual analyse	Visual positive uptake	DFS	Visual positive uptake
Singh <i>et al.</i> [56]	2003	47	IIIB	visual analyse	Visual positive uptake	OS	Visual positive uptake
Grigsby <i>et al.</i> [65]	2004	152	IB-IV	visual analyse	abnormal FDG uptake	DFS	abnormal FDG uptake
Xue <i>et al.</i> [52]	2006	96	IB-IVB	manually segmented	SUVmax	DFS, OS	SUVmax>1 0.2
Kidd <i>et al.</i> [50]	2007	287	IVB-IA2	Threshold at 40% SUVmax	SUVmax, MTV, Lymph Nodes Metastasis	OS	SUVmax
Yen <i>et al.</i> [53]	2008	70	I-IV	Spherical ROI manually placed	SUVmax	DFS, OS	SUVmax>3, 3
Chung <i>et al.</i> [66]	2010	75	IB to IIA	manually segmented	SUVmax	DFS	SUVmax
Chung <i>et al.</i> [57]	2011	63	IB to IIA	Threshold at 2.5	SUVmax, MTV	DFS	SUVmax, MTV, Age
yoo <i>et al.</i> [61]	2011	73	I-VI	SUV of mediastinal	SUVmax, SUVmean, MTV, TLG	DFS	MTV, TLG
kim <i>et al.</i> [58]	2011	45	IA–IIB	Threshold at 40% SUVmax	MTV	DFS	MTV
Kidd <i>et al.</i> [59]	2012	234	Ib1–IVa	Threshold at 40% SUVmax	SUVmax, MTV, Highest lymph nodes	DFS, OS	SUVmax, MTV, Highest lymph nodes
Crivellaro <i>et al.</i> [63]	2012	49	IB1 and IIA	adaptive threshold PETVCAR (GE Healthcare)	SUVmax, SUVmean, MTV, TLG	DFS	None
Kidd <i>et al.</i> [67]	2013	25	Ib1-IVa	Threshold at 40% SUVmax	SUVmax, FDGhetero, MTV	DFS	SUVmax, FDGhetero, MTV
Micco <i>et al.</i> [62]	2014	49	IB-IVB	Threshold at 42% SUVmax	SUVmean, MTV, TLG	DFS, OS	MTV, TLG
Renier <i>et al.</i> [68]	2014	41	≥1B	Threshold at 2.5, Threshold at 40% SUVmax and FLAB	SUVmax, SUVpeak, MTV, TLG, and texture features	DFS, OS	MTV, TLG40, entropy, intensity variability,
Chung <i>et al.</i> [69]	2015	48	IB-IIA	SUV-based Automated delineation of Syngo.na (Siemens Healthcare)	intratumoural heterogeneity(IH), SUVmax, MTV, TLG	DFS	IH, MTV, TLG

## 1.2 L'ANALYSE DE L'IMAGERIE TEP POUR LA PREDICTION DE RECIDIVE

D'un point de vue général, l'analyse de l'image dans une visée de prédiction peut se modéliser par la transformation de données d'acquisitions « brutes » – apportées par les images et des informations tierces – vers une information quantifiée, utilisable pour prédire l'apparition d'évènements futurs. Dans le cadre de la prédiction d'un résultat thérapeutique, l'objectif est d'obtenir des paramètres corrélés statistiquement à un évènement clinique en vue de pouvoir, *in fine*, utiliser ces résultats en clinique et diminuer la probabilité de la survenue de l'évènement en question.

Notre problématique d'analyse des images TEP pour la prédiction de récurrence s'inscrit donc dans ce schéma. La procédure se présente sous la forme d'une étude d'apprentissage sur population, utilisant des examens TEP de patientes traitées ainsi que leurs résultats thérapeutiques, en vue de trouver des marqueurs permettant de discriminer les patientes ayant présenté une récurrence des patientes en rémission.

Ces dernières décennies, la littérature s'est énormément développée sur l'analyse des images TEP et de nombreuses méthodologies ont été proposées pour aborder les problématiques liées au traitement de celles-ci.

Dans notre objectif de prédiction de récurrence ou de survie, le processus de l'analyse des images se décomposera donc en 3 problématiques :

1. Le prétraitement des images : améliorer, préparer et transformer les images ;
2. La segmentation des images : isoler le signal d'intérêt du reste de l'image ;
3. La quantification : extraire des paramètres des images TEP caractérisant le métabolisme tumoral et son évolution au cours du traitement.

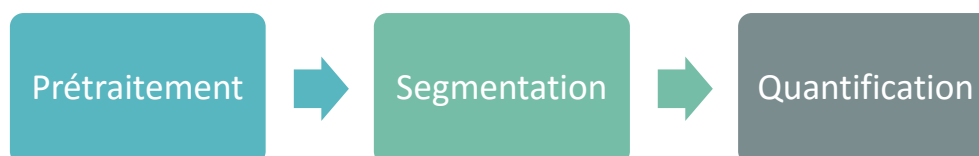


Figure 1.7 Schéma de traitement des images TEP

Nous dresserons ci-dessous un état de l'art succinct des méthodologies proposées par la communauté scientifique.

### 1.2.1 PRETRAITEMENTS D'IMAGE DANS L'IMAGERIE TEP

Le prétraitement regroupe l'ensemble des processus visant à améliorer ou transformer les images TEP, à plusieurs niveaux, dans le but de permettre une analyse plus fine du signal d'acquisition brut. Ces applications sont multiples, souvent combinées, et déterminent la conduite des étapes de traitement suivantes (segmentation et quantification). Parmi ces applications les plus courantes nous pouvons citer : la définition de régions d'intérêt pour l'analyse, le débruitage, la correction de l'effet de volume partiel, le recalage avec d'autres images et la création d'images paramétriques (voir Figure 1.6).



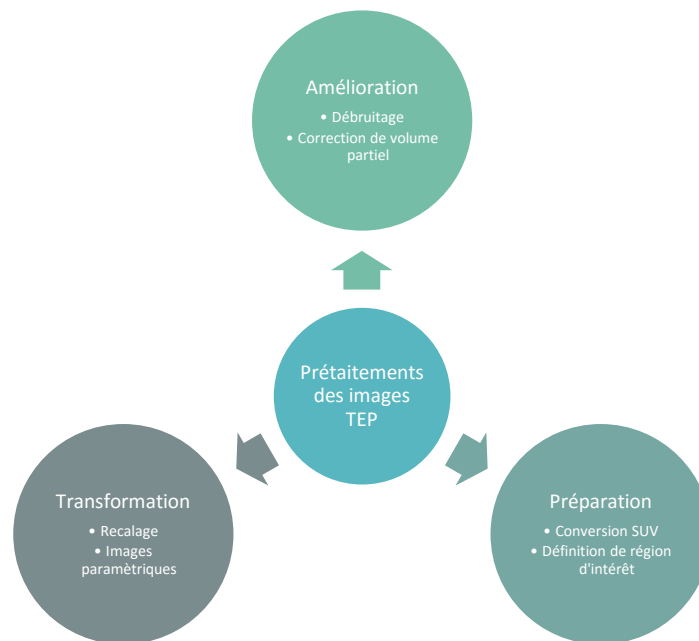


Figure 1.8 Exemples de pré-traitements des images dans l'analyse de la TEP au 18F-FDG

### Débruitage des images TEP

Le bruit est un facteur rendant la segmentation difficile et pouvant biaiser la quantification des images. En imagerie TEP, le bruit est modélisé le plus couramment par un bruit poissonien ou par un bruit gaussien, voire les deux.

Les méthodes de débruitage du bruit gaussien les plus utilisées sont le filtrage bilatéral [70], le filtrage par diffusion anisotrope [71] et le filtrage par ondelettes [72]. Dans [73], Trukheimer *et al.* ont proposé d'ajouter les informations structurelles apportées par les images TDM ou IRM pour renforcer le rapport signal sur bruit en utilisant un modèle multi-résolution par décomposition sur base d'ondelettes.

Bagci *et al.* en 2013 [74] ont proposé une approche considérant le bruit poissonien. L'algorithme se base sur une transformation du bruit multiplicatif en bruit additif par la transformation par stabilisation de variance d'Anscombe [75], puis par troncation de la décomposition en valeurs singulières. Nous noterons que cette méthode, comparée au filtrage par diffusion anisotropique, permet de mieux conserver les mesures de SUVmax et de SUVmean (valeur moyenne de l'hyperfixation au sein de la tumeur).

Les méthodes de débruitage, dans la littérature, sont généralement utilisées pour augmenter les performances des méthodes de segmentation [76,77] et en complément de méthode de déconvolution pour la correction de volume partiel [78].

### Correction de l'effet de volume partiel des images TEP

La correction de l'effet de volume partiel a pour but de corriger le flou et l'entremêlement des signaux de structures inhérents à l'acquisition TEP. Le but est ici d'inverser l'effet de la convolution d'une image TEP idéale infiniment résolue avec la PSF dans le but afin de corriger la mesure de la distribution du 18F-FDG dans le corps.

Globalement, il existe deux stratégies distinctes :

- Les méthodes basées et appliquées durant la reconstruction pour améliorer la résolution spatiale de l'image ;
- Les méthodes de déconvolution, appliquées sur les images volumiques pour restaurer le signal d'origine.

La communauté est très active sur le sujet et de nombreuses stratégies ont été mise en place pour résoudre ces problèmes [79–81]. Concernant les méthodes de déconvolutions dans l'imagerie TEP, Teo *et al.* et Thoka *et al.* ont proposé d'utiliser l'algorithme itératif de Van Clittert [82,83]. Tohka *et al.*, quant à eux, ont proposé l'utilisation de l'algorithme de Richardson-Lucy [83]. L'utilisation de la méthode d'estimation par maximum de vraisemblance a également été proposée par Kirov *et al.* et Barbee *et al.* [84,85].

### Définition de la région d'intérêt pour l'analyse

La définition de la région d'intérêt, en prétraitement, a pour objectif d'établir un périmètre dans lequel s'effectuera le reste de l'analyse de l'imagerie TEP. Généralement, la région d'intérêt dans laquelle sont appliquées les méthodes de segmentations est définie par une boîte englobante 3D placée autour de la tumeur [86–89]. Cette boîte englobante doit être suffisamment large pour pouvoir entourer tout le signal tumoral mais suffisamment compact pour éviter les structures environnantes pouvant induire une variabilité inter- et intra- observateur.

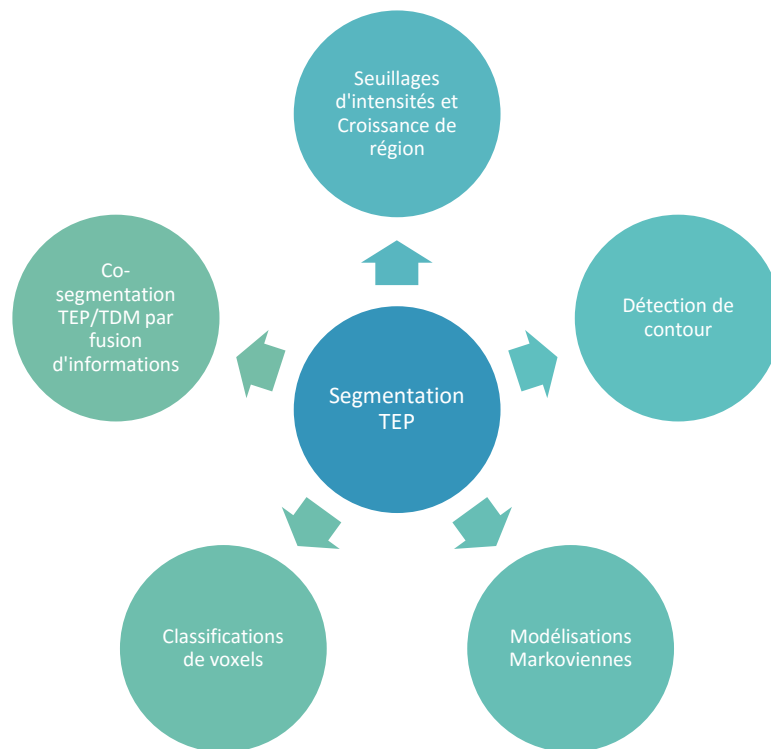
### Recalage des images TEP et création d'images paramétriques

En 2011, Necib *et al.* [90] ont proposé l'utilisation 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'intervalle et à identifier ainsi les zones dans lesquelles l'activité métabolique a varié de manière significative entre les examens. Elle permet d'aboutir à une image 3D paramétrique qui décrit la réponse de la maladie à la thérapie et d'obtenir des mesures quantitatives décrivant les modifications métaboliques relatives aux tumeurs. À ce jour, peu d'études se sont penchés sur la création, l'utilisation et l'analyse d'images paramétriques dans la prédiction.

#### 1.2.2 SEGMENTATION DE L'IMAGERIE TEP

La segmentation est l'un des problèmes phares du traitement d'images. Elle consiste à partitionner l'image en différentes zones, en vue d'en extraire une ou plusieurs régions d'intérêts. Dans l'analyse d'images TEP, la segmentation vise à isoler le signal tumoral du reste de l'examen dans le but de pouvoir étudier son métabolisme. Dans leur revue sur les méthodes de segmentation TEP, Foster *et al.* ont montré en 2014 [24] un nombre de publication sur le sujet en très forte progression entre 2008 et 2011 (étude basé sur la période 1983-2012), étudiant plus d'une centaines de méthodes. D'autres revues ont été publiées, et permettent d'obtenir une vision globale sur les méthodes de segmentation appliquées aux images TEP [24,37,91,92].

La Figure 1.9 schématise les différentes approches méthodologiques utilisées dans la littérature pour la segmentation du signal tumoral.



**Figure 1.9 Les différentes approches méthodologiques utilisées dans la littérature pour segmenter le signal métabolique tumoral dans l'imagerie TEP**

### Méthodes de seuillage d'intensité et croissance de région

Une des premières techniques de segmentation automatique à avoir été utilisée en routine clinique consiste à réaliser un seuillage de l'image TEP. Les méthodes de seuillage proposées dans la littérature ont été résumé par Dewalle et son équipe en 2010 [93]. Parmi celles les plus connues, nous pouvons citer le seuil fixe de SUV, le seuillage en pourcentage du SUVmax et les méthodes de seuillage adaptatifs.

Pour le seuillage fixe de SUV, la valeur seuil de 2,5 a été proposée pour la première fois en 1993 par Patz *et al.* dans la segmentation des images TEP pour la différenciation d'anomalies pulmonaires bénignes et malignes [94]. Il a ensuite été régulièrement utilisé dans des applications pulmonaires [95–97], mais aussi dans des objectifs de prédiction [98–100] et en comparaison avec d'autres méthodes de segmentation [36,101–103]. Ce seuil est toutefois très discuté notamment à cause de la non-reproductibilité du SUV, intrinsèquement affecté par de nombreux paramètres comme la glycémie du patient, les mouvements physiologiques ou encore le temps d'acquisition de l'examen [104,105].

Le seuillage le plus utilisé en imagerie TEP reste celui basé sur un pourcentage fixe du SUVmax. La valeur appliquée varie généralement entre 30% et 70% mais les valeurs de références restent autour de 40-42%. La valeur de 42% a originellement été retenue au travers de l'étude de Erdi et son équipe en 1995 sur la tomographie d'émission monophotonique (TEMP) [106], puis reproduite en 1997 pour la TEP [107]. Cette dernière étude montrait, entre autres, que pour des lésions TEP très fixantes, un seuillage à 42% permettait d'obtenir des volumes métaboliques très proches des volumes anatomiques obtenus sur TDM. Cette valeur a ensuite été reprise par la communauté scientifique comme une référence. Cette méthodologie ne permet cependant pas de s'adapter au bruit de l'image.

Dans leur étude de 2003, Daisne *et al.* [108] présentaient une méthode permettant de trouver le seuil adapté à chaque image en le calculant en fonction du contraste de l'image. Cette méthode vise à établir

une relation entre le seuillage et le rapport signal sur bruit (RSB) à l'aide d'une étude sur fantômes sphériques utilisant différentes tailles de sphères et différents paramètres d'acquisition. Partant de cette relation, l'utilisateur peut alors établir le seuillage nécessaire pour la segmentation de l'image en fonction du contraste (RSB) mesuré dans l'image. L'intervention du RSB dans le calcul du seuil requiert cependant la définition de plusieurs volumes d'intérêt (signal et fond), difficilement reproductible et pouvant faire varier le résultat. Dans la même optique de trouver un seuil adapté à chaque image, Black *et al.* ont proposé en 2004 [109], également au travers d'une étude de calibration sur fantômes sphériques, d'établir une relation linéaire entre le seuil optimal et le SUV moyen au sein de la tumeur (SUVmean). Ces méthodes présentent l'avantage d'être adaptées à chaque caméra TEP et aux variations d'acquisition. En revanche, la calibration de la caméra TEP étant difficilement reproductible, les relations seuillage-RSB et seuillage-SUVmean doivent être calculées pour chaque caméra, induisant une faible reproductibilité inter-centres.

Le désavantage premier des méthodes uniquement basées sur les intensités est qu'elle ne prend pas en compte la distribution spatiale des voxels. Dans l'optique de prendre en compte cette information dans la formation de la région, les seuillages présentés ci-dessus sont aujourd'hui appliqués en croissance de région – au sein d'une région d'intérêt – en vue de garder un volume compact. La croissance de région permet, à partir d'un point de départ nommé graine, de faire grandir la région en sélectionnant comme nouvelles graines tous les voxels voisins partageant une caractéristique commune.

Green *et al.* ont proposé d'améliorer la croissance de région tridimensionnelle pour la segmentation des hyperfixations, avec des critères d'agrégation définis par l'utilisateur [110]. Day *et al.* en 2009 [111] ont proposé l'application d'un seuillage en croissance de région itératif, s'adaptant aux statistiques de moyenne et de variance des valeurs de voxels au sein de la région en croissance. Les résultats montraient que, comparée aux seuillages fixes (2,5 et 42% du SUVmax), cette méthode permettait de réduire les erreurs de classification.

### Méthodes de détection de contours

Les méthodes de segmentation basées sur la détection de contours consistent à détecter la région d'intérêt au travers du contour de l'objet plutôt que l'objet lui-même, utilisant l'information du gradient de l'image au lieu de l'information d'intensité directe.

La détection de contours pour l'imagerie TEP a été proposée au travers des modèles de contours actifs, également appelée modèle de *snakes*, dans [112,113]. La méthode se base sur un contour, positionné autour de l'objet à l'initialisation, se déformant et se déplaçant pour venir se « coller » au contour de l'objet recherché (symbolisé par un gradient). La déformation du contour actif est gérée par une fonction d'énergie interne et externe ; l'énergie interne cherchant à garantir un contour continu et dérivable (pas de cassure brutale dans le contour) et l'énergie externe poussant le contour à se déformer autour de l'objet.

Pour détecter les contours de l'objet, il a également été proposé d'utiliser des méthodes de ligne de niveaux, utilisant l'algorithme de partage des eaux (ou *watershed* en anglais) [76,114,115]. L'algorithme de *watershed* est un outil de segmentation utilisé en morphologie mathématique basé sur l'analogie de l'image à une surface topologique se remplissant d'eau à partir de sources données.

Les méthodes dérivatives sont néanmoins sensibles au bruit présent dans l'image et nécessitent de renforcer le gradient naturellement faible de la TEP. Ces méthodes nécessitent donc la mise en place de prétraitements de type débruitage ou de déconvolution avant d'être en mesure d'appliquer les algorithmes de segmentations sur les images. Par exemple, Abdoli *et al.* ont combiné l'utilisation d'un contour actif avec un filtrage par diffusion anisotropique et une transformation en ondelettes de contours [77]. Également, dans [76], Geets *et al.* ont proposé, préalablement à l'utilisation du *watershed*, d'appliquer un filtre bilatéral

puis l'algorithme de déconvolution itérative de Van Clitter, afin de réduire le bruit et d'augmenter le contraste de l'image.

### Modélisations markoviennes

Le formalisme des champs de Markov permet d'effectuer une segmentation de l'image en prenant en compte les interactions avec les pixels voisins. Cette méthodologie repose sur une modélisation probabiliste de l'image observée ; l'image est une observation d'un champ aléatoire constitué de plusieurs régions. Chaque voxel est donc une réalisation de la variable aléatoire correspondant à sa région d'appartenance. La théorie des champs de Markov considère, par ailleurs, que la probabilité d'apparition de chaque voxel n'est dépendant que de ses propres voisins. Cette théorie permet donc de prendre en compte l'information de relation spatiale de chaque voxel de l'image tout en modélisant un système stochastique bruité.

En 2001, Chen *et al.* ont présenté une application des champs de Markov pour la segmentation TEP en incorporant ce modèle dans les itérations d'un algorithme de maximisation de l'espérance (*Expectation-Maximization* ou EM, en anglais) [116]. Dans leurs études de 2007, Montgomery *et al.* ont comparé l'utilisation du modèle Markovien avec une version de ce modèle en multi-échelle incluant une transformée en ondelettes et l'algorithme des k-moyens [117]. Les résultats montraient une sous-estimation globale de l'algorithme des k-moyens et une surestimation dans le cas des champs de Markov traditionnel. L'application multi-échelle donnait, quant à elle, les résultats les plus encourageants, sous la contrainte d'un temps de calcul plus long.

L'utilisation des chaînes de Markov pour la segmentation TEP ont été proposées par Hatt *et al.* en 2007 afin de réduire ce temps de calcul [118]. L'algorithme repose sur le parcours de l'image (3D) en une marche de voxel en voxel (1D). En incluant une mesure floue d'appartenance des voxels aux classes « signal d'intérêt » et « fond », la chaîne de transition markovienne est ensuite estimée à l'aide d'un algorithme EM stochastique. Cette méthode, comparée à un seuillage à 42% du SUVmax, montrait une meilleure robustesse aux variations de qualité et de contrastes entre les images. Nous noterons que la méthode présentait des quantifications de volumes similaires au seuillage à 50% du SUVmax.

### Modélisations stochastiques et classification de voxels

D'autres méthodologies se sont basées sur une modélisation probabiliste de l'image en cherchant à classer les voxels de l'image en fonction de leur appartenance à des classes ; par exemple en partitionnant l'image en deux classes « tumeur » et « non-tumeur ». La mixture de gaussiennes [86,119], le partitionnement flou de l'image [120,121] et d'autres modélisations dérivées ont été proposées [38,39,122] pour déterminer la classe d'appartenance de chaque voxel.

L'algorithme de partitionnement flou des c-moyens (Fuzzy-c-mean ou FCM, en anglais) est une méthode très connue en segmentation d'image, basée sur une classification non-supervisée utilisant la logique floue dans la définition des partitions de l'image en assignant à chaque voxel une probabilité d'appartenance à chaque classe. L'algorithme repose sur un processus itératif d'optimisation d'une fonction de coût basée sur la différence de chaque voxels au barycentre de chaque classe. Elle a été suggérée pour l'imagerie TEP pour la segmentation des tumeurs cérébrales par Zhu *et al.* en 2003 [120]. Cependant, cette méthode se base uniquement sur les valeurs d'intensité des voxels et reste très sensible au bruit et à l'hétérogénéité au sein des structures de l'image. Belhassen et Zaidi ont proposé en 2010 [121] d'améliorer les résultats de cette méthode en incorporant la corrélation spatiale des voxels et en réduisant l'hétérogénéité de l'image par i) l'utilisation au préalable d'un filtrage anisotrope et ii) la transformation de l'image en ondelettes. Comparé à un algorithme FCM simple, cette méthode améliorée réduisait substantiellement les erreurs de classification des voxels appartenant à la tumeur avec une tendance à la surestimation du volume réel.

L'algorithme *Fuzzy Locally Adaptive Bayesian* (FLAB) développé par Hatt *et al.* en 2009 [86] se base sur une classification statistique des voxels basée sur l'établissement d'un modèle de mélange de gaussiennes résolue par inférence bayésienne. Cette méthode prend en compte une assignation floue de l'appartenance des voxels à chaque classe. L'initialisation des classes de l'image est effectuée par l'algorithme FCM. L'algorithme espérance-maximisation stochastique est ensuite utilisé afin de d'estimer les paramètres des distributions mélangées et estimer la probabilité d'appartenance de chaque voxel aux classes. Cette méthode a été originellement développée pour deux classes [86] et a été étendue à trois classes dans [87], afin de permettre la segmentation de tumeurs hétérogènes. L'équipe du LATIM a démontré à maintes reprises la robustesse, la reproductibilité et la précision de la méthode [123–125].

En complément de la théorie des probabilités, la théorie des possibilités et la théorie de l'évidence ont été abordées dans la segmentation des tumeurs en TEP afin de prendre en compte la nature imprécise de l'image.

La théorie de l'évidence proposé par Dempster et Shafer en 1976 [126] est une théorie mathématique basée sur l'utilisation de notion de croyance et de plausibilité apportées par différentes sources d'informations afin d'établir le degré d'évidence d'un état. Lelandais *et al.* en 2014 [127] ont proposé une méthode de segmentation basée sur les fonction de croyances dans le cadre d'une étude sur le *dose-painting* multi-traceur en TEP. Cette méthode propose l'utilisation de l'algorithme évidentiel des c-moyens [128] amélioré pour prendre en compte la corrélation spatiale de chaque voxel (SECM). L'algorithme SECM est basé sur le même principe d'optimisation que le FCM, en remplaçant la logique floue par des fonctions de croyance. La méthode a été comparée à FLAB et a un seuillage adaptatif présenté par Vauclin *et al.* dans [129]. Les résultats montraient une meilleure estimation du volume tumoral, avec une légère surestimation du volume dans les données réelles. L'avantage de leur méthode réside principalement dans la possibilité de fusionner les différentes sources d'information (différents radio-traceurs TEP dans l'étude) en vue de réduire les imprécisions et incertitudes de chaque source prise séparément.

La théorie des possibilités, suggérée par Lotfi Zadeh [130] est une approche mathématique pour la représentation de l'information incomplète. Celle-ci se base sur l'utilisation de deux évaluations duales, dites de possibilités et de nécessité permettant de représenter et de quantifier l'incertitude d'un état. Dewalle-Vignion *et al.* ont présenté en 2011 une méthodologie basée sur l'utilisation de cette approche en combinant les degrés de possibilité obtenus à partir de plusieurs projections d'intensité maximale (Maximum Intensity Projection and Aggregation by Possibility Theory, ou MIPAPT en anglais) [131]. Cette méthode surperformait les méthodes de seuillage (42% et de Daisne) et l'algorithme FCM.

### Méthodes par fusion de modalités d'images

Afin de résoudre les problèmes d'imprécisions inhérentes à l'imagerie TEP, plusieurs équipes ont récemment proposé inclure dans le processus de segmentation l'information apportée par d'autres modalités d'imagerie, telles que la TDM ou l'IRM [132–137]. La fusion d'images fonctionnelles et anatomiques (TEP/TDM ou TEP/IRM) est en effet très utilisée dans l'analyse visuelle des examens TEP car elle permet de réduire les incertitudes de localisation du métabolisme tout en maximisant l'information pertinente.

El Naqa *et al.* en 2007 [132], ont proposé d'utiliser une co-segmentation TEP/TDM au travers d'un modèle déformable basé sur la méthode des surfaces de niveaux (topologie surfacique des niveaux d'intensité de l'image) permettant de combiner les résultats de segmentation des deux images prises séparément.

Dans [134] et [137], D. Han *et al.* et Q. Song *et al.* ont proposé une méthode de segmentation conjointe des images TEP/TDM basée sur l'utilisation des champs de Markov encodant l'information de deux modalités. La résolution du champ Markovien est résolue par construction et élagage d'un graphe. Nous noterons

cependant que cette méthode nécessite le placement de graines d'initialisation et ne permet pas de s'affranchir d'une variabilité inter- et intra-observateur.

Plus récemment, Bagci *et al.* [136] ont également proposé une méthode basée sur une marche aléatoire markovienne permettant une segmentation conjointe des images TEP/TDM. Cette méthode a l'avantage d'être purement automatique et utilisable sur les images prises séparément ou conjointement. Les résultats montraient une précision croissante à l'utilisation conjointe des deux modalités, surpassant la méthode de Han [134].

### 1.2.3 QUANTIFICATION ET EXTRACTION DE PARAMETRES DE L'IMAGERIE TEP

L'étape de quantification des images TEP a pour objectif d'extraire des paramètres permettant de caractériser l'activité tumorale.

La problématique réside dans la difficulté à calculer des paramètres invariants aux variations statistiques des images (bruit, contraste, etc.) comparables entre les examens mais suffisamment pertinent pour évaluer les spécificités tumorales de chaque patient.

*A contrario* des images de TDM ou d'IRM fournissant une information anatomique des structures du corps humain, l'imagerie TEP au 18F-FDG reflète, elle, l'activité glucidique au sein des cellules de ce même corps. L'hyperfixation en FDG des cellules cancéreuses étant connue pour refléter l'agressivité de la tumeur, cette information supplémentaire nous permet d'obtenir une dimension nouvelle dans la caractérisation tumorale.

Les paramètres extraits des images TEP pour caractériser le métabolisme tumoral se divisent en 3 catégories :

- les mesures d'intensité du métabolisme ;
- les mesures d'hétérogénéités du métabolisme tumoral ;
- les mesures de formes de la région métabolique.

#### Mesures d'intensité du métabolisme tumoral

Comme vu précédemment, les paramètres les plus couramment évalués dans les images TEP en routine clinique [22] sont :

- le SUVmax ;
- le pic de SUV (peak SUV ou SUVpeak en anglais) ;
- le SUV moyen (mean SUV ou SUVmean en anglais) ;
- le volume métabolique tumoral (*Metabolic Tumor Volume* ou MTV en anglais) ;
- la glycolyse totale de la lésion (*Total Lesion Glycolysis* ou TLG en anglais).

Il existe cependant d'autres paramètres basés sur l'intensité des voxels encore très peu étudiés dans leur capacité à caractériser l'hyperfixation tumorale. Dans leur étude sur la réponse tumorale du traitement des cancers de l'œsophage, Tixier *et al.* [138] ont proposé l'utilisation de paramètres d'intensités liés à l'histogramme de l'image, en complément de paramètres de texture décrits dans la prochaine section. Ces paramètres sont présentés dans le Tableau 1.5.

#### Mesures de l'hétérogénéité du métabolisme tumoral

La mesure d'indices de texture des images TEP à récemment été proposée pour compléter l'information d'intensité donnée par le SUV [99,138–141].

L'exploration des paramètres de forme et d'hétérogénéité est motivée par le fait que la présence d'hétérogénéité dans le métabolisme tumoral pourrait refléter certaines caractéristiques des tumeurs, comme la présence de nécrose, de fibroses, zones avec angiogénèses, de hautes prolifération cellulaire ou la présence de récepteurs spécifiques.

Dès 2003, Miller *et al.* [55] ont proposé une mesure de cette hétérogénéité à des fins pronostiques dans les cancers du col. Cette mesure se base une analyse visuelle et consiste à cumuler un score selon différents critères tels que la taille de la tumeur par rapport aux dimensions du pelvis, sa sphéricité et son hétérogénéité.

Dans la localité du cancer pulmonaire non à petites cellules, Hicks *et al.* en 2004 [142] ont visuellement évalué l'hétérogénéité des tumeurs pour évaluer la réponse tumorale. Kalff *et al.* en 2006 [143] ont également proposé l'utilisation de critères basés sur la reconnaissance de forme de la fixation tumorale de FDG pour caractériser la réponse thérapeutique de carcinomes rectaux.

Dans le cadre d'une étude exploratoire sur la prédiction des récives des cancers du col de l'utérus, El Naqa *et al.* en 2009 [140] ont proposé d'extraire des paramètres quantitatifs basés sur les caractéristiques de textures de Haralick [144] ainsi que des descripteurs de formes. De plus, les auteurs ont proposé l'utilisation d'histogramme Intensité-Volume – un analogue à l'histogramme Dose-Volume utilisé en radiothérapie [145] – permettant de mesurer l'intensité totale en fonction du volume considéré. Les résultats de cette étude montraient l'intérêt des paramètres de texture dans la caractérisation de la tumeur. Cependant la valeur prédictive de ceux-ci restait en générale en dessous du MTV. Les auteurs ont conclu sur la nécessité d'investigations plus profondes sur l'utilisation des textures pour l'analyse TEP.

Dans la localité des cancers de l'œsophage, Tixier *et al.* [138], en 2011, ont également proposé d'extraire des paramètres de texture pour prédire la réponse thérapeutique. Dans cette étude, les paramètres de texture étaient extraits de l'histogramme, des matrices de co-occurrences de niveaux de gris (*Gray-Level Co-occurrence Matrix* ou GLCM en anglais) de Haralick *et al.* [144], des matrices d'isolongueur de nuances de gris proposé (Run-Length Matrix ou RLM en anglais) par Xiaoou Tang en 1998 [146] et des matrices de taille de zone de niveaux de gris (*Gray-Level Size Zone Matrix* ou GLSZM en anglais) de Thibault *et al.* [147]. Les résultats statuaient que l'analyse de texture du métabolisme permettaient de mieux prédire la réponse au traitement que les mesures standard d'intensité de SUV.

Les paramètres de textures extraits des images TEP dans ces études sont présentés dans :

- le Tableau 1.5 pour les paramètres de l'histogramme des hyperfixations au sein de la tumeur
- le Tableau 1.6 pour les paramètres de textures basés sur les GLCM de Haralick *et al.* ;
- le Tableau 1.7 pour les paramètres de textures basés sur les RLM de Tang ;
- le Tableau 1.8 pour les paramètres de textures basés sur les GLSZM de Thibault *et al.*

La même équipe a également évalué la reproductibilité de ces paramètres de texture en fonction de la discrétisation appliquée [148]. Sur un total de 24 paramètres testés, les auteurs ont préconisé l'utilisation de 5 d'entre eux pour leur reproductibilité (Homogénéité, Entropie, Dissimilarité, SZNU et GLNUz, cf. Tableau 1.6 et Tableau 1.8).

En 2014, Orhac *et al.* ont publié une étude sur les corrélations existantes entre les paramètres de texture, les paramètres d'histogrammes les mesures de SUV et les paramètres MTV et TLG [149]. 41 paramètres ont été extraits, au travers de différentes méthodes de segmentation, pour évaluer leur robustesse et leur reproductibilité. Les résultats de l'étude montraient qu'aucun paramètre d'histogramme et seulement 17 paramètres de texture (sur les 31 testés) étaient robustes à la segmentation utilisée. Les auteurs préconisaient également d'utiliser une discrétisation de niveau de gris d'au moins 32 niveaux pour éviter



une corrélation trop forte avec les mesures de SUV standard. Enfin, les auteurs soulignaient que certains paramètres de textures (notamment GLNUr, GLNUz, ZLNU et TLG) étaient très corrélés aux paramètres MTV. Il existe donc aujourd'hui encore de nombreux questionnement concernant l'utilisation de ces paramètres dans l'analyse des images TEP et des études sont nécessaires pour évaluer leur valeur dans la prédiction.

Tableau 1.5 Paramètres extraits de l'histogramme des valeurs de voxels. Les paramètres Maximum et Mean correspondent respectivement au SUVmax et SUVmean.

Nom du paramètre	Equation	Définition
Minimum	$\min(I(x_i))$	Valeur minimale d'hyperfixation de la lésion
Maximum	$\max(I(x_i))$	Valeur maximale d'hyperfixation de la lésion
Moyenne	$\mu = \mathbb{E}[X]$ $\mu = \frac{1}{N} \cdot \sum_i^N I(x_i)$	Valeur moyenne d'hyperfixation de la lésion
Variance	$\sigma^2 = \mathbb{E}[(X - \mu)^2]$ Estimé par : $\sigma^2 = \frac{1}{N-1} \cdot \sum_i^N (I(x_i) - \mu)^2$	Variance de la distribution des voxels
Ecart type (Standard Deviation, SD)	$SD = \sqrt{\sigma^2}$	Ecart type des valeurs de voxels
Skewness	$\gamma_1 = \mathbb{E} \left[ \left( \frac{X - \mu}{\sigma^2} \right)^3 \right]$ Estimé par : $\gamma_1 = \frac{N^2}{(N-1) \cdot (N-2)} \cdot \frac{\frac{1}{N} \sum_i^N (I(x_i) - \mu)^3}{(\sigma^2)^{3/2}}$	Mesure d'asymétrie de la distribution des voxels
Kurtosis	$\gamma_2 = \mathbb{E} \left[ \left( \frac{X - \mu}{\sigma^2} \right)^4 \right]$ Estimé par : $\gamma_2 = \frac{N(N+1)}{(N-1) \cdot (N-2) \cdot (N-3)} \cdot \frac{\frac{1}{N} \sum_i^N (I(x_i) - \mu)^4}{(\sigma^2)^2} - 3 \frac{(n-1)^2}{(N-2) \cdot (N-3)}$	Mesure d'aplatissement de la distribution des voxels
Avec : $X = \{x_1, x_2, \dots, x_i, x_N\}$ les N voxels dans le volume d'intérêt ; $I(x_i)$ l'intensité du voxel $x_i$		

**Tableau 1.6 Paramètres de texture de Haralick extraits des matrices de co-occurrences**  
 Les paramètres correspondent aux moyennes des caractéristiques dans toutes les directions.

Nom du paramètre	Equation	Interprétation
Energie, Homogénéité ou Second moment angulaire	$\sum_i^M \sum_j^M GLCM_{i,j}^2$	Mesure l'homogénéité : une forte valeur caractérise une texture homogène.
Entropie	$\sum_i^M \sum_j^M GLCM_{i,j} \times \log_2(GLCM_{i,j})$	Mesure le degré d'organisation de la texture (varie de façon opposée au second moment angulaire).
Inertie ou contraste	$\sum_i^M \sum_j^M (i-j)^2 \times GLCM_{i,j}$	Mesure l'étalement des valeurs de la matrice normalisée autour de la diagonale principale. Une forte valeur caractérise une texture fine.
Dissimilarité	$\sum_i^M \sum_j^M  i-j  \times GLCM_{i,j}$	Mesure de dissimilarité des valeurs de voxels. S'interprète de la même manière que le contraste.
Moment inverse de différence	$\sum_i^M \sum_j^M \frac{(i-j)^2}{GLCM_{i,j}}$	Une forte valeur indique que les éléments texturaux sont de grande taille.
Cluster shade	$\sum_i^M \sum_j^M (i - \mu_i + j - \mu_j)^3 \times GLCM_{i,j}$	Mesure de dissymétrie des valeurs de la matrice GLCM (skewness).
Cluster Prominence	$\sum_i^M \sum_j^M (i - \mu_i + j - \mu_j)^4 \times GLCM_{i,j}$	Mesure d'aplatissement des valeurs de la matrice GLCM (kurtosis).

**Tableau 1.7 Paramètres de texture extraits des matrices d'isolongueurs de nuances de gris (RLM)**  
Les paramètres correspondent aux moyennes des caractéristiques dans toutes les directions.

Nom du paramètre	Equation	Interprétation
<b>Courtes isolongueurs (Short Run Emphasis, SRE)</b>	$SRE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{RLM_{i,j}}{j^2}$	Mesure accentuant la distribution des courtes isolongueurs. Une forte valeur de SRE est attendue pour les textures fines.
<b>Grandes isolongueurs (Long Run Emphasis, LRE)</b>	$LRE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N RLM_{i,j} \times j^2$	Mesure accentuant la distribution des longues isolongueurs. Une forte valeur de LRE est attendue pour une texture à structure grossière.
<b>Isolongueurs des faibles niveaux de gris (Low Gray-level Run Emphasis, LGRE)</b>	$LGRE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{RLM_{i,j}}{i^2}$	Mesure accentuant la distribution des faibles niveaux de gris. Une forte valeur de LGRE est attendue pour une image de faible intensité moyenne.
<b>Isolongueurs des hauts niveaux de gris (High Gray-level Run Emphasis, HGRE)</b>	$HGRE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N RLM_{i,j} \times i^2$	Mesure accentuant la distribution des hauts niveaux de gris. Une forte valeur de HGRE est attendue pour une image de haute intensité moyenne.
<b>Courtes isolongueurs des faibles niveaux de gris (Short Run Low Gray-level Emphasis, SRLGE)</b>	$SRLGE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{RLM_{i,j}}{i^2 \cdot j^2}$	Mesure accentuant conjointement les courtes isolongueurs et faibles niveaux de gris. Une forte valeur de SRLGE indique une texture dont les structures sont essentiellement composées de courtes isolongueurs à faible intensité.
<b>Courtes isolongueurs des hauts niveaux de gris (Short Run High Gray-level Emphasis, SRHGE)</b>	$SRHGE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{RLM_{i,j} \times i^2}{j^2}$	Mesure accentuant conjointement les courtes isolongueurs et hauts niveaux de gris. Une forte valeur de SRHGE indique une texture dont les structures sont essentiellement composées de courtes isolongueurs à haute intensité.
<b>Longues isolongueurs des faibles niveaux de gris (Long Run Low Gray-level Emphasis, LRLGE)</b>	$LRLGE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{RLM_{i,j} \times j^2}{i^2}$	Mesure accentuant conjointement les longues isolongueurs et les faibles niveaux de gris. Une forte valeur de LRLGE indique une texture dont les structures sont essentiellement composées de longues isolongueurs à faible intensité.
<b>Longues isolongueurs des hauts niveaux de gris (Long Run High Gray-level Emphasis, LRHGE)</b>	$LRHGE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{RLM_{i,j}}{i^2 \times j^2}$	Mesure accentuant conjointement les longues isolongueurs et les faibles niveaux de gris. Une forte valeur de LRHGE indique une texture dont les structures sont essentiellement composées de longues isolongueurs à haute intensité.
<b>Non-Uniformité des niveaux de gris (Gray-Level Non-Uniformity, GLNUr)</b>	$GLNUr = \frac{1}{Nb_{voxel}} \sum_i^M \left( \sum_j^N RLM_{i,j} \right)^2$	Mesure la non-uniformité des niveaux de gris. Une faible valeur de GLNUr indique une homogénéité dans les intensités.
<b>Non-Uniformité des isolongueurs (Run-Length Non-Uniformity, RLNU)</b>	$RLNU = \frac{1}{Nb_{voxel}} \sum_j^N \left( \sum_i^M RLM_{i,j} \right)^2$	Mesure la non-uniformité des isolongueurs. Une faible valeur de RLNU indique une homogénéité dans les isolongueurs.

---

<b>Pourcentage d'isolongueurs (Run Percentage, RP)</b>	$RP = \frac{Nb_{voxel}}{RLM_{i,j} \times j}$	Mesure l'homogénéité de la matrice. RP est maximale quand la longueur des isolongueurs est égale à 1 pour tous les niveaux de gris.
--	--	---

Tableau 1.8 Paramètres de texture extraits des matrices de taille de zone de niveaux de gris (GLSZM). Les paramètres sont calculés de la même manière que les paramètres de RLM à l'exception qu'aucune moyenne en fonction des directions n'est nécessaire car les GLSZM ne sont pas dépendantes d'une direction.

Nom du paramètre	Equation	Interprétation
Petites zones (Short Zone Emphasis, SZE)	$SZE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{GLSZM_{i,j}}{j^2}$	Mesure accentuant la distribution des petites zones. Une forte valeur de SZE est attendue pour les textures fines.
Grandes zones (Long Zone Emphasis, LZE)	$LZE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N GLSZM_{i,j} \times j^2$	Mesure accentuant la distribution des grandes zones. Une forte valeur de LZE est attendue pour une texture à structure grossière.
Zones des faibles niveaux de gris (Long Gray-level Zone Emphasis, LGZE)	$LGZE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{GLSZM_{i,j}}{i^2}$	Mesure accentuant la distribution des zones de faibles niveaux de gris. Une forte valeur de LGZE est attendue pour une image de faible intensité moyenne.
Zones des hauts niveaux de gris (High Gray-level Zone Emphasis, HGZE)	$HGZE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N GLSZM_{i,j} \times i^2$	Mesure accentuant la distribution des zones de hauts niveaux de gris. Une forte valeur de HGZE est attendue pour une image de haute intensité moyenne.
Petites zones des faibles niveaux de gris (Short Zone Low Gray-level Emphasis, SZLGE)	$SZLGE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{GLSZM_{i,j}}{i^2 \cdot j^2}$	Mesure accentuant conjointement les petites zones et faibles niveaux de gris. Une forte valeur de SZLGE indique une texture dont les structures sont essentiellement composées de petites zones à faible intensité moyenne.
Petites zones des hauts niveaux de gris (Short Zone High Gray-level Emphasis, SZHGE)	$SZHGE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{GLSZM_{i,j} \times i^2}{j^2}$	Mesure accentuant conjointement les petites zones et hauts niveaux de gris. Une forte valeur de SZHGE indique une texture dont les structures sont essentiellement composées de petites zones à haute intensité.
Grandes zones des faibles niveaux de gris (Large Zone Low Gray-level Emphasis, LZLGE)	$LZLGE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N \frac{GLSZM_{i,j} \times j^2}{i^2}$	Mesure accentuant conjointement les grandes zones et les faibles niveaux de gris. Une forte valeur de LZLGE indique une texture dont les structures sont essentiellement composées de grandes zones à faible intensité moyenne.
Grandes zones des hauts niveaux de gris (Large Zone High Gray-level Emphasis, LZHGE)	$LZHGE = \frac{1}{Nb_{voxel}} \sum_i^M \sum_j^N GLSZM_{i,j} \times i^2 \times j^2$	Mesure accentuant conjointement les grandes zones et les faibles niveaux de gris. Une forte valeur de LZHGE indique une texture dont les structures sont essentiellement composées de grandes zones à haute intensité moyenne.
Non-uniformité des niveaux de gris (Gray-Level Non-Uniformity, GLNUz)	$GLNUz = \frac{1}{Nb_{voxel}} \sum_i^M \left( \sum_j^N GLSZM_{i,j} \right)^2$	Mesure la non-uniformité des niveaux de gris. Une faible valeur de GLNUz indique une texture homogène dans ces intensités.
Non-uniformité de tailles de zones (Size-Zone Non-Uniformity, SZNU)	$SZNU = \frac{1}{Nb_{voxel}} \sum_j^N \left( \sum_i^M GLSZM_{i,j} \right)^2$	Mesure la non-uniformité des zones. Une faible valeur de SZNU indique une texture de zones homogène.
Pourcentage de zone (Zone Percentage, ZP)	$ZP = \frac{Nb_{voxel}}{GLSZM_{i,j} \times j}$	Mesure l'homogénéité de la matrice. ZP est maximale quand la longueur des zones est égale à 1 pour tous les niveaux de gris.

## 1.3 ENJEUX METHODOLOGIQUES DANS L'ANALYSE DES IMAGES TEP POUR LA PREDICTION DE RECIDIVE

### 1.3.1 FAIBLE SPECIFICITE DU RADIOTRACEUR FDG

La première limitation dans l'analyse de la TEP 18F-FDG est sa spécificité, car l'augmentation du métabolisme glucidique est aussi marquée lors de phénomènes inflammatoires ou infectieux. L'inflammation au sein des tissus sains, après irradiation est donc une source potentielle de perte de spécificité du signal. Ainsi, l'analyse des images per-thérapeutiques nécessite la prise en compte de nouveaux paradigmes en plus de ceux utilisés pour l'analyse des images de diagnostic ; la présence d'artefacts inflammatoires et la fonte tumorale induit un signal métabolique faible, noyé dans le bruit de fond de l'image, voire inexistant. L'analyse doit donc prendre en compte ces différences avec les examens pré-thérapeutiques en modifiant ces paramètres pour prendre en compte cette diminution du signal au sein d'un terrain lui-même artefacté (ex : débruitage, modification du seuillage, alignement de l'image sur la tumeur initiale, etc.).

Par ailleurs, l'hyperfixation physiologique des organes environnants peut également être la source d'une perte de spécificité dans l'analyse du signal. Adjacente au col utérin, la vessie peut fixer le 18F-FDG et créer un continuum entre l'hyperfixation vésicale et tumorale [150]. Plusieurs solutions ont été apportées pour éviter la présence de cet artefact dans l'image. La première solution est d'éviter que la vessie soit pleine durant l'acquisition, en invitant les patientes à uriner avant l'examen [22]. Dans certains cas, l'utilisation d'un cathéter urinaire peut être nécessaire, mais celui-ci est parfois difficile à mettre en place et peut induire des risques non-négligeables d'irradiation [151]. Une autre solution consiste à supprimer l'artefact vésical directement sur l'image en utilisant une délinéation de la vessie effectuée sur TDM. Cependant, la délinéation manuelle – méthode utilisée en routine clinique – reste laborieuse, subjective et très peu reproductible entre les opérateurs [24]. Une dernière solution consiste à placer une boîte englobante autour de la tumeur, afin de ne pas prendre en compte les voxels extérieurs à cette région. Évidemment, cette solution, bien que simple à mettre en œuvre, ne permet pas d'éviter une variabilité intra- et inter-observateur, induisant des variations de mesures. Il est à noter, par ailleurs, que ce genre d'hyperfixation physiologique est difficilement prévisible car très variable d'un patient à l'autre, limitant de surcroît la mise en place d'algorithme de suppression automatique des artefacts.

### 1.3.2 IMPACT DE LA SEGMENTATION POUR LA PREDICTION

Une autre limitation dans l'évaluation des lésions en TEP est l'isolation de la zone tumorale du reste de l'image. En effet, l'imagerie TEP souffre d'un fort bruit d'acquisition et de reconstruction. Une des premières techniques à avoir été utilisée en routine clinique consistait à réaliser un seuillage de l'image TEP. Cette méthodologie est cependant très sensible au seuillage choisi et ne permet pas de s'adapter au bruit de l'image. Biehl *et al.* en 2006, ont démontré que l'utilisation d'un unique seuil pour différentes images ne permettait pas d'obtenir une définition précise des volumes tumoraux, notamment pour les tumeurs les plus petites [152].

Hatt *et al.* en 2010 ont comparé la reproductibilité des seuillages fixes (50% et 42%), du seuillage adaptatif de Daisne [108], de l'algorithme FCM et de FLAB dans le cadre de lésion de l'œsophage et du sein [124]. L'équipe a montré que le seuillage adaptatif de Daisne n'était pas plus reproductible que les seuillages fixes et que ces méthodes étaient sujettes à une faible reproductibilité due aux variations de qualité des acquisitions TEP et du SUVmax. Cette méthodologie présente donc une limite dans sa capacité à s'adapter à la taille des tumeurs et aux variations de contraste entre les images.

Dans la section 1.2.1, nous avons vu que de nombreuses méthodes, plus évoluées, ont été proposées dans la littérature pour prendre en compte ces limitations, en utilisant, entre autres, des approches par classifications, par contours, par modèles stochastiques, par fusion d'informations, etc. Néanmoins, malgré les avancées sur ce domaine, aucune méthode ne fait aujourd'hui consensus. L'utilisation de méthodes plus évoluées nécessite généralement i) une instanciation de la méthode au problème présent (pré-traitement, optimisation, apprentissage, etc.), ii) un coût de calcul plus important et iii) l'ajout éventuel d'informations tierces (caractéristiques des patients, informations *a priori*, calibration de la caméra, fusion avec l'imagerie TDM ou IRM, etc.). Le seuillage fixe lié à une valeur du SUV reste une segmentation très couramment utilisée sur ces images car il est simple à mettre en œuvre et est rapidement applicable sur toute image. La segmentation des images TEP en routine clinique oscille donc entre l'utilisation de méthodes génériques simples, introduisant des biais de mesure, et l'utilisation de méthodes plus précises mais plus spécifiques, complexifiant la mise en place de ladite mesure. Il est à noter que les méthodes de segmentations TEP sont généralement développées dans le but de définir très précisément le volume tumoral. Néanmoins, *a contrario* de l'objectif de ciblage thérapeutique, la recherche de paramètres prédictifs n'impose pas une délimitation précise de la zone tumorale mais une recherche de la localisation de l'information d'intérêt.

L'importance de la segmentation TEP dans la prédiction de récurrence est une question encore très peu étudiée et son impact sur les paramètres extraits reste encore à établir.

### 1.3.3 LIMITES DE LA QUANTIFICATION DE L'IMAGERIE TEP

La prédiction de la réponse thérapeutique, au travers de l'évaluation de la captation glucidique exprimée dans les images TEP est, aujourd'hui encore, généralement focalisée sur des paramètres simples – et rapides à mettre en œuvre – comme l'hyperfixation maximale (SUVmax, SUVpeak), le volume métabolique (MTV), la glycolyse totale de la lésion (TLG) ou encore la valeur moyenne des hyperfixations au sein de la lésion (SUVmean). Il existe cependant de nombreuses autres approches permettant de quantifier le signal d'une image. Par exemple, les moments statistiques ou les indices d'histogrammes (variance, coefficient d'asymétrie, coefficient d'aplatissement, etc.) permettent de caractériser la distribution des valeurs de voxels. Les paramètres de texture mesurent, quant à eux, l'organisation spatiale de la distribution des valeurs de voxels. Enfin les paramètres de forme s'attachent à décrire la distribution spatiale des coordonnées des voxels au sein de la tumeur dans l'espace. Ces nouvelles mesures, appliquées à la TEP, permettraient de quantifier l'hétérogénéité de l'activité tumorale au sein de la tumeur et sont de plus en plus étudiées [99,138–141,149,153,154]. La pertinence de ces paramètres reste cependant à démontrer dans l'objectif de prédiction.

Une autre limitation de la TEP est la faible résolution spatiale et l'effet de volume partiel qu'elle induit dans l'image. Ces limitations sont induites par la nature stochastique de l'acquisition (parcours des positons, non colinéarité des photons lors de l'annihilation), la caméra TEP (ex : taille, nature et espacement des capteurs à photons), et l'algorithme de reconstruction utilisé [79,155]. En pratique, l'effet de volume partiel, et donc le flou de l'image, est modélisé par reconstruction tomographique d'une image source par une fonction d'étalement de point (PSF). L'image TEP est donc, par nature, une image sans contours précis, aux frontières incertaines entre les structures. Se pose alors la question de la pertinence de quantifications extraites au travers des méthodes de partitionnement binaire des voxels de l'image.





## 2 OBJECTIFS DE LA THESE

Dans le cadre de la radiothérapie des cancers du col utérin, les objectifs de cette thèse sont :

- l'exploration de la valeur prédictive sur le risque de récurrence des TEP pré- et per-thérapeutique ;
- l'exploration de l'utilisabilité de divers paramètres, extraits des images TEP pré- et per-thérapeutique prises séparément et prises conjointement ;
- l'identification de paramètres d'intérêt pour la prédiction de récurrence post-traitement ;
- l'exploration de l'impact de la segmentation TEP sur les paramètres extraits ;
- le développement de méthodes reproductibles pour analyser les images TEP prenant en compte :
  - les erreurs inhérentes à l'utilisation de méthodes de segmentations,
  - la nature floue de l'image TEP,
  - les artefacts liés à l'hyperfixation à vessie.

La Figure 2.1 illustre le schéma de traitement et d'analyse des images TEP pré- et per-thérapeutiques pour la prédiction de récurrence et les méthodes mises en œuvre dans ces travaux de thèse pour répondre aux objectifs.

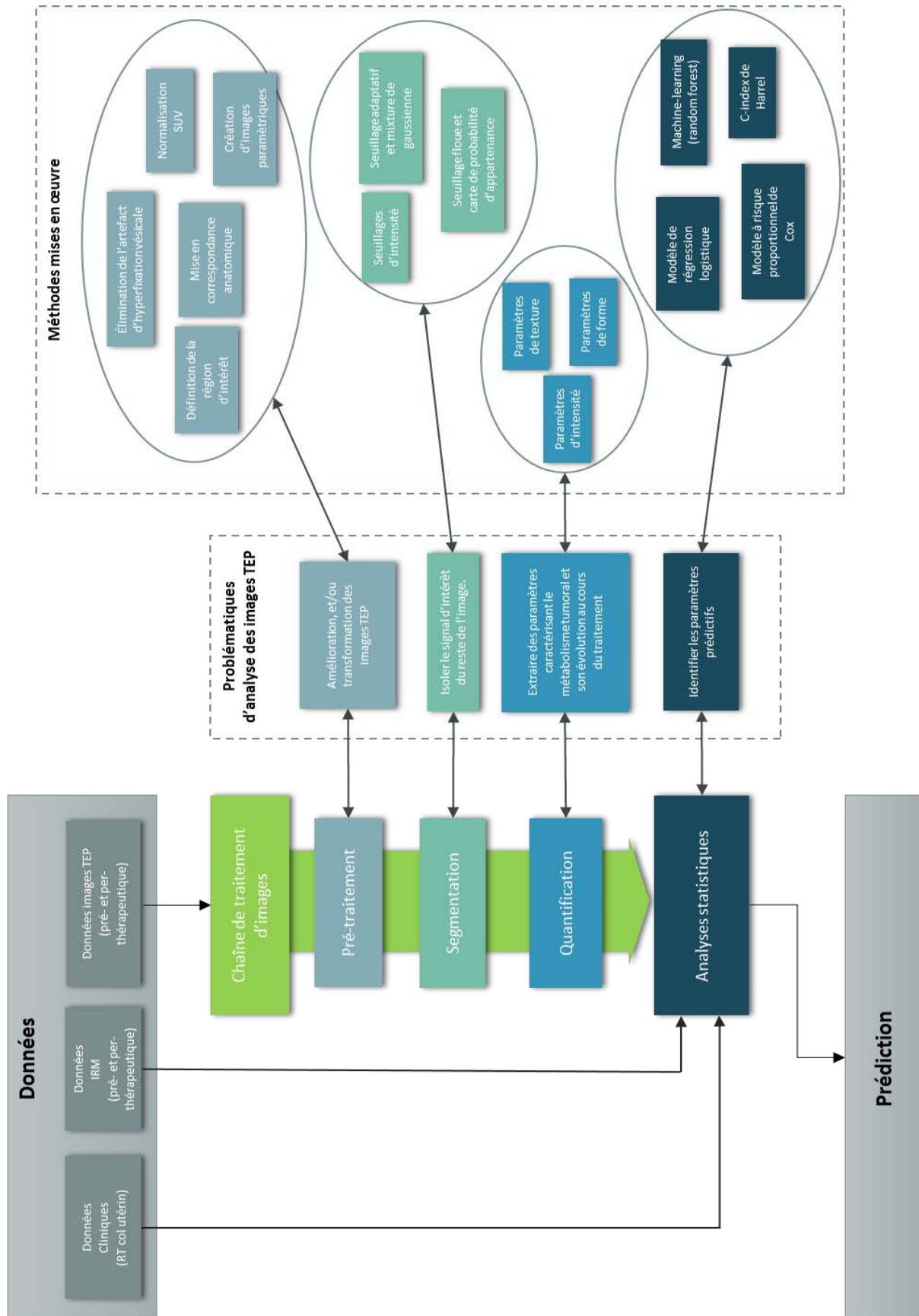


Figure 2.1 Schéma de traitement et d'analyse des images TEP pré- et per-thérapeutique pour la prédiction de récurrence et les méthodes mises en œuvre dans ces travaux

### 3 ETAPE DE PRETRAITEMENT : SUPPRESSION DES ARTEFACTS D'HYPERFIXATION VESICALE DANS LES IMAGES TEP AU 18F-FDG POUR L'ANALYSE DU METABOLISME DES TUMEURS DU COL UTERIN (ARTICLE AU JOURNAL CBM 2015)

*Ce chapitre présente une méthode semi-automatique de détection de l'hyperfixation vésicale dans les images TEP pour permettre une analyse non-parasitée du métabolisme des tumeurs. La contribution de cette méthode se situe dans la problématique de prétraitement des images TEP, en vue d'augmenter la spécificité de l'analyse TEP et réduire les incertitudes sur la présence ou non d'un artefact vésical. La méthode est basée sur un double algorithme des k-moyens utilisant l'information de la TEP, l'information de la TDM ainsi que des informations « a priori » sur la position de la tumeur et de la vessie. La méthode a été testée sur un total de 52 images TEP pré-thérapeutiques de patientes traitées pour un cancer du col de l'utérus. Les résultats montrent que la méthode proposée est capable de supprimer la majeure partie de l'artefact vésicale tout en préservant l'activité tumorale.*

*Le contenu de ce chapitre a été publié au journal « Computers in Biology and Medicine » (doi : 10.1016/j.combiomed.2016.02.002).*

---

## CONTENU

<b>3</b>	<b>ETAPE DE PRETRAITEMENT : SUPPRESSION DES ARTEFACTS D'HYPERFIXATION VESICALE DANS LES IMAGES TEP AU 18F-FDG POUR L'ANALYSE DU METABOLISME DES TUMEURS DU COL UTERIN (ARTICLE AU JOURNAL CBM 2015)</b> .....	<b>51</b>
3.1	Abstract.....	53
3.2	Introduction .....	53
3.3	Materials and Methods.....	55
3.3.1	PET/CT dataset.....	55
3.3.2	Detection of bladder metabolic artefact .....	57
3.3.3	Validation.....	60
3.4	Results.....	62
3.5	Discussion .....	65
3.6	Conclusion.....	67

---



## Detection of bladder metabolic artefact in 18F-FDG PET imaging

Geoffrey Roman-Jimenez<sup>a,b,d,\*</sup>, Renaud De Crevoisier<sup>a,c</sup>, Julie Leseur<sup>c</sup>, Anne Devillers<sup>c</sup>, Juan David Ospina<sup>a,b</sup>, Antoine Simon<sup>a,b</sup>, Pierre Terve<sup>d</sup>, Oscar Acosta<sup>a,b</sup>

<sup>a</sup>INSERM, U1099, Rennes, F-35000, France

<sup>b</sup>Université de Rennes 1, LTSI, Rennes, F-35000, France.

<sup>c</sup>Département de Radiothérapie, Centre Eugène Marquis, Rennes, F-35000, France.

<sup>d</sup>Keosys Medical Imaging, Saint-herblain, F-44815, France.

### 3.1 ABSTRACT

Positron emission tomography using 18F-fluorodeoxyglucose (18F-FDG-PET) is a widely-used imaging modality in oncology. It enables significant functional information to be included in analyses of anatomical data provided by other image modalities. Although PET offers high sensitivity in detecting suspected malignant metabolism, 18F-FDG uptake is not tumor-specific and can also be fixed in surrounding healthy tissue, which may consequently be mistaken as cancerous. PET analyses may be particularly hampered in pelvic-located cancers by the bladder's physiological uptake potentially obliterating the tumor uptake. In this paper, we propose a novel method for detecting 18F-FDG bladder artifacts based on a multi-feature double-step classification approach. Using two manually-defined seeds (tumor and bladder), the method consists of a semi-automated double-step clustering strategy that simultaneously takes into consideration standard uptake values (SUV) on PET, Hounsfield values on computed tomography (CT), and the distance to the seeds. This method was performed on 52 PET/CT images from patients treated for locally-advanced cervical cancer. Manual delineations of the bladder on CT images were used in order to evaluate bladder uptake detection capability. Tumor preservation was evaluated using a manual segmentation of the tumor, with a threshold of 42% of the maximal uptake within the tumor. Robustness was assessed by randomly selecting different initial seeds. The classification averages were  $0.94 \pm 0.09$  for sensitivity,  $0.98 \pm 0.01$  specificity, and  $0.98 \pm 0.01$  accuracy. These results suggest this method is able to detect most 18F-FDG bladder metabolism artifacts while preserving tumor uptake, and could thus be used as a pre-processing step for further non-parasitized PET analyses.

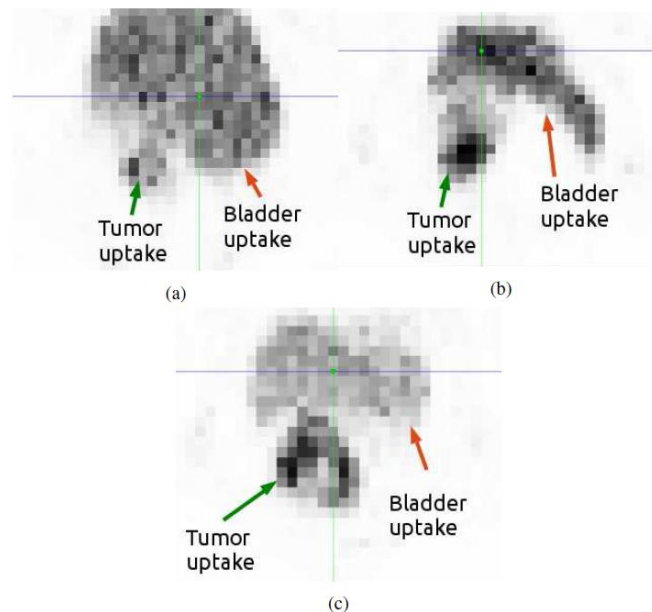
### 3.2 INTRODUCTION

In the last few decades, positron emission tomography/computed tomography (PET/CT) has been increasingly employed in cancer management, enabling the simultaneous assessment of tumor metabolic and morphological characteristics [156,157]. Of the different techniques, <sup>18</sup>F-fluorodeoxyglucose (<sup>18</sup>F-FDG) PET imaging is now widely used to improve target delineation accuracy [158], diagnostic accuracy [159,160], as well as early outcome prediction in forms of cancer [100,161–163] that require accurate delimitation and precise quantification of tumor metabolism. PET imaging techniques' weak points are low spatial resolution, partial volume effect, low contrast, and high noise features. Also, despite PET being highly sensitive for detecting malignant metabolism, FDG uptake is not specific to tumor activity. <sup>18</sup>F-FDG fixation can be observed both in healthy tissue and benign diseases like inflammation, which can thereby be mistaken for cancer [164].

Today, numerous methodologies for <sup>18</sup>F-FDG PET segmentation have been proposed in the scientific literature using region-growing thresholds, edge detection, clustering, stochastic models, and deformable models, among other approaches [24,37,91,165]. However, despite the accuracy achievable with these

recent methodologies, most are based solely on the PET data, leading to the possibility of including non-tumor  $^{18}\text{F}$ -FDG uptake in the resulting segmentations.

When particularly considering cancers located in the pelvic region, PET analyses may be especially hampered by physiological uptake in the bladder, which could be a confounder for tumor uptake [150]. Due to the bladder's natural flow mechanics of filling and emptying,  $^{18}\text{F}$ -FDG uptake can vary, thereby generating a similar PET signal for the bladder to that of a tumor. Figure 3.1 presents three cases of locally-advanced cervical tumors with continuum uptake within the bladder.



**Figure 3.1** Three examples of PET acquisition of locally-advanced cervical cancer, demonstrating a continuum uptake between tumor and bladder hyperfixation.

To avoid  $^{18}\text{F}$ -FDG bladder hyperfixation, patients are invited to empty their bladder before commencing  $^{18}\text{F}$ -FDG PET/CT scanning [22]. In rare cases, urinary catheters can be used to ensure the bladder is empty for the PET acquisition. Nevertheless, urinary catheter positioning is known to be difficult and carries a non-negligible risk of increased irradiation damage [151].

Bladder uptake may also be suppressed by using expert manual bladder CT delineation. However, manual delineations are known to be laborious, subjective, and inter-observer dependent [24].

Current clinical practice involves manually positioning a 3D-box encompassing the lesion to avoid adjacent structures [86–89]. Nevertheless, the 3D-box needs to be large enough to contain the entire tumor metabolism but also as small as possible to reduce high signals damaging neighboring tissues. Evidently, these procedures involve inter- and intra-user variability and can include voxel artifacts depending on the proximity and shape of the bladder.

To the best of our knowledge, no other automatic or semi-automatic method for detecting metabolism artifacts with the aim of improving and simplifying pelvic PET analysis have previously been proposed. In this paper, we present a semi-automatic method for detecting  $^{18}\text{F}$ -FDG bladder artifacts using a double-step clustering approach that simultaneously exploits multimodal PET-CT data and *a priori* spatial information.

The k-means algorithm was used to ensure unsupervised clustering and low computational cost. As bladder uptake is not systematically found in these procedures, our detection method could be considered a pre-processing step for PET imaging in order to enable further non-parasitized tumor quantification. From this point of view, the issue we wish to address herein is how to distinguish the bladder signal while preserving the information arising from the tumor uptake. This methodology was applied to real clinical data from a standardized clinical protocol where bladder uptake was observed in several cases.

The paper is structured as follows: Section 2 introduces the overall framework of the study; Section 2.2 exposes each step of the proposed methodology, followed by validation based on the clinical data; Section 3 presents the results; Section 4 presents our discussion and conclusions.

### 3.3 MATERIALS AND METHODS

The overall framework of our proposed methodology for semi-automatically detecting bladder metabolic artifacts is presented in Figure 3.2. This method is divided into two major clustering steps. First, k-means clustering was performed using only PET intensity to roughly select the voxels with the highest values, likely to be those belonging to the tumor or bladder (2.2.1). The largest connected component in the resulting cluster was then considered for the next step. Manual inspection was conducted to verify if tumor metabolism could be visually detected in the resulting area (2.2.2). If so, a second step of clustering was performed using PET/CT data and *a priori* spatial information (2.2.3). Morphological correction was then applied to ensure the topological compactness of the detected region (2.2.4). Validation (2.3) was conducted on a dataset of 52 PET/CT images by: i) comparing the detected bladder voxel artifacts using expert manual bladder CT delineation; ii) comparing tumor metabolism preservation with respect to a segmentation on PET images using the commonly-used threshold of 42% of the maximum uptake within the tumor ( $T_{42}$ ).

The clustering algorithms and segmentations were carried out using the C++ library Insight Segmentation and Registration Toolkit (ITK).

#### 3.3.1 PET/CT DATASET

We initially included 52 patients (median age: 52.44 years [32.15 - 84.62]) with locally-advanced cervical cancer treated at the *Centre Eugene Marquis* (CEM), Rennes, France. All were treated using external beam radiation therapy with concurrent cisplatin chemotherapy followed by brachytherapy.

Each patient underwent an  $^{18}\text{F}$ -FDG PET/CT scan prior to treatment. Patients were invited to void before imaging. No intravenous contrast was used during the CT acquisitions.

Each voxel from the PET scan was converted into a standardized uptake value (SUV) for comparisons across patients. The SUV is a standardized decay-corrected value of  $^{18}\text{F}$ -FDG activity per unit volume of body weight (MBq/kg) [166]. The full-body PET/CT images were cropped to the pelvic region based on the coxal bone structures visualized on CT imaging. The box boundaries were delimited according to well-defined landmarks that could be easily reproducible. Figure 3.3 depicts an example of a cropped pelvic region.



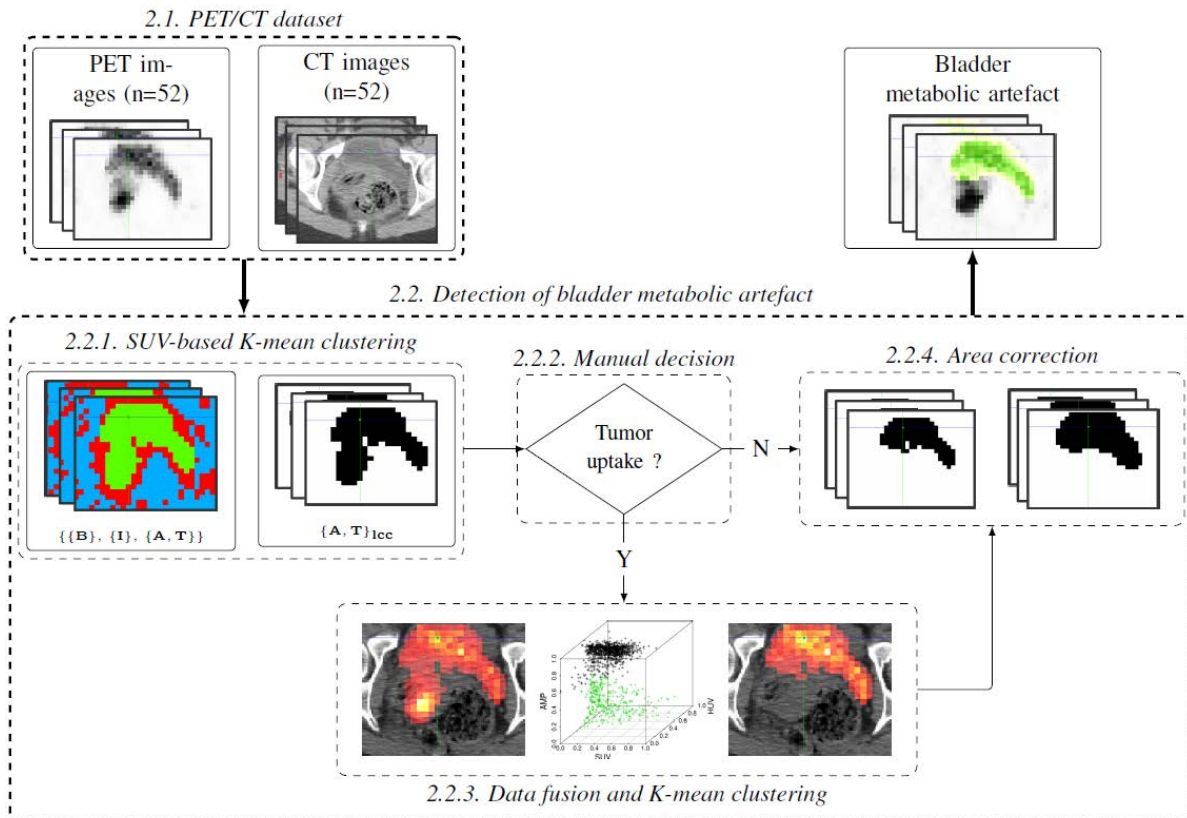


Figure 3.2 Overall framework of the proposed methodology for semi-automatically detecting the bladder metabolic artifacts. First, k-means clustering was performed using only PET intensity to roughly select the voxels with the highest values, likely belonging to the tumor or bladder (2.2.1). Following manual verification (2.2.2), an optional second step of clustering using PET/CT data and *a priori* spatial information was performed (2.2.3). Finally, morphological correction was applied to ensure the topological compactness of the detected region (2.2.4)

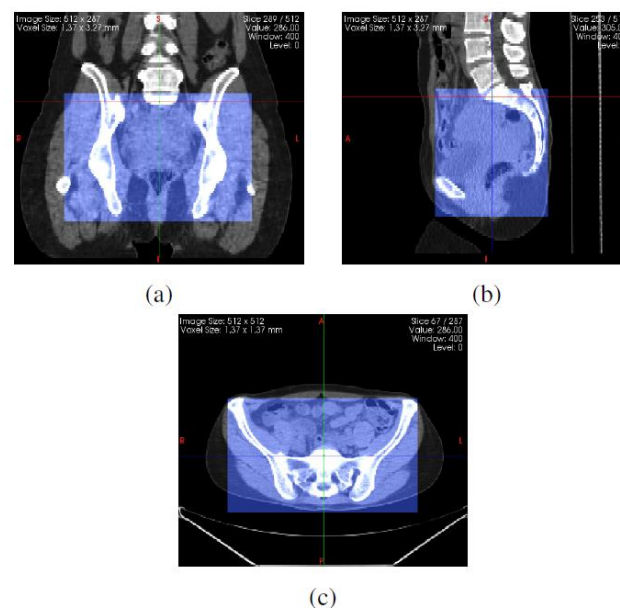


Figure 3.3 Example of pelvic region cropped according to coxal bone structures on CT images

### 3.3.2 DETECTION OF BLADDER METABOLIC ARTEFACT

We defined bladder metabolic artifact (BMA) as areas where the  $^{18}\text{F}$ -FDG bladder activity was highest and comparable to that originating from the tumor. We defined the calculation of BMA detection as a classification formula where each voxel of the PET image was assigned to one of the three clusters of Set  $\mathbf{S}$ :

$$\mathbf{S} = \{\{B\}, \{I\}, \{A, T\}\}$$

where  $\{B\}$  and  $\{I\}$  refer to the voxels producing very low and intermediate uptake values, respectively. Cluster  $\{A, T\}$  refers to very high uptake voxels, which may belong to either tumor  $\{T\}$  or non-tumor  $\{A\}$  classes, respectively.

The BMA detection sought to accurately identify the voxels belonging to  $\{A\}$ . It should be noted that  $\{T\}$  does not correspond to tumor segmentation but only to voxels with high uptake, which thus likely belong to a tumor and should be preserved for further non-parasitized quantitative analyses.

Due to the bladder shapes observed, we assumed that the BMA revealed in PET imaging was a topologically compact object. Two voxels roughly representing the bladder and tumor barycenter were selected, denoting  $s_b$  and  $s_t$ , respectively.

#### SUV-based k-means clustering

In our calculations,  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ , the set of  $n$  voxels on the PET image. In general terms, the k-means algorithm aims to partition the  $n$  voxels of  $\mathbf{X}$  into  $k$  clusters by minimizing the within-cluster sum of squares difference:

$$\min_{\mathbf{C}} \sum_{j=1}^k \sum_{x_i \in c_j} \|x_i - \mu_j\|^2$$

where  $\mathbf{C} = \{c_1, c_2, \dots, c_k\}$  is the set of clusters and  $\mu_j$  denotes the mean of voxels in  $c_j$ . With cluster set  $\mathbf{S}$ , the objective function becomes:

$$\min_{\mathbf{S}} \left( \sum_{x_i \in \{B\}} \|x_i - \mu_B\|^2 + \sum_{x_i \in \{I\}} \|x_i - \mu_I\|^2 + \sum_{x_i \in \{A, T\}} \|x_i - \mu_{A, T}\|^2 \right)$$

Figure 3.4 presents an example of an SUV-based  $k$ -means clustering method with cluster set  $\mathbf{S}$ .

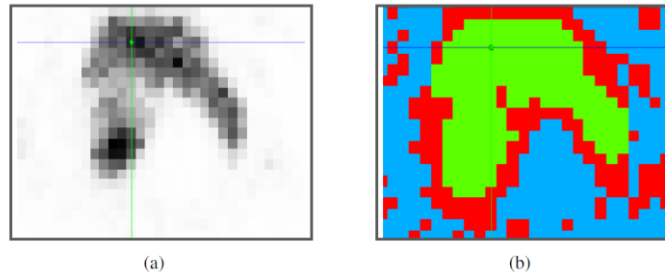


Figure 3.4 SUV-based k-means clustering example. (a) Original PET exam. (b) k-means clustering result: (green) corresponds to Cluster  $\{A, T\}$ , (red) to Cluster  $\{I\}$ , and (blue) to Cluster  $\{B\}$

As the SUV-based k-means clustering took no spatial information into consideration, we selected the largest connected component ( $lcc$ ) of Cluster  $\{A, T\}$  connected to seed  $s_b$  to avoid distant regions and ensure that the resulting area was only related to the BMA. The resulting cluster was labeled  $\{A, T\}_{lcc}$ . Figure 5.4a presents an example of Cluster area  $\{A, T\}_{lcc}$ .

#### Manual decision

The potential presence of tumor metabolism in Cluster  $\{A, T\}_{lcc}$  was determined visually. If no tumor metabolism was present in  $\{A, T\}_{lcc}$ , the cluster was denoted as just  $\{A\}$ . If tumor metabolism was visually detected in  $\{A, T\}_{lcc}$ , a second step of clustering using CT information and *a priori* information was performed with the intention of splitting  $\{A\}$  from  $\{T\}$ .

#### Data Fusion and k-means clustering

In the second clustering step, a k-means algorithm was applied for further classification of Set  $\{A, T\}_{lcc}$ , based on a simultaneous exploitation of CT Hounsfield units, PET SUVs, and normalized distance maps. The distance map was built as an artifact membership probability (AMP) as follows:

$$AMP(x_i) = 1 - \frac{d^2(x_i, s_b)}{d^2(x_i, s_t) + d^2(x_i, s_b)}$$

where  $d^2(x_i, s_t)$  and  $d^2(x_i, s_b)$  are the squared Euclidean distances of voxel  $x_i$  from seed  $s_t$  and seed  $s_b$ , respectively.

The hypothesis behind the AMP relies on the concept that the further the voxel is from the tumor seed, the lower the probability of it belonging to the artifacts. Similarly, we assumed that the closer the voxel was to the bladder seed, the higher the probability of it belonging to the bladder artifacts.

An example of AMP computed for each voxel in  $\{A, T\}_{lcc}$  is illustrated in Figure 5.4b.

From each  $x_i$  clustered in  $\{A, T\}_{lcc}$ , three features were considered: the SUV from the PET, labeled  $U_i$ ; the Hounsfield unity value (HUV) from the CT, labeled  $H_i$ ; the AMP, labeled  $P_i$ . The SUVs and HUVs were normalized between 0 and 1. We denoted that  $V\{A, T\}_{lcc} = \{v_i\}$  the group of vector  $v_i = (U_i H_i P_i)^T$ .

Each  $v_i$  was assigned a label stating if it belonged to cluster  $\{T\}$  or  $\{A\}$ , following the objective function below:

$$\min_{\{A,T\}_{icc}} \left( \sum_{v_i \in \{T\}} \|x_i - \mu_B\|^2 + \sum_{v_i \in \{A\}} \|x_i - \mu_B\|^2 \right)$$

Figure 3.6b represents an example of the k-means clustering in the tri-parametric space  $[SUV, HUV, \text{and } AMP]$ .

In accordance with the compactness hypothesis on the BMA shape, the largest connected component of the cluster connected with seed  $s_b$  was selected to be  $\{A\}$ .

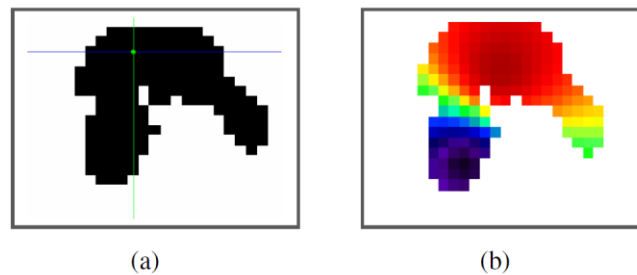


Figure 3.5 (a) Cluster  $\{A, T\}_{icc}$  and (b) the corresponding AMP for each voxel in  $\{A, T\}_{icc}$ : (red) corresponds to AMP=1 and (black) to AMP=0

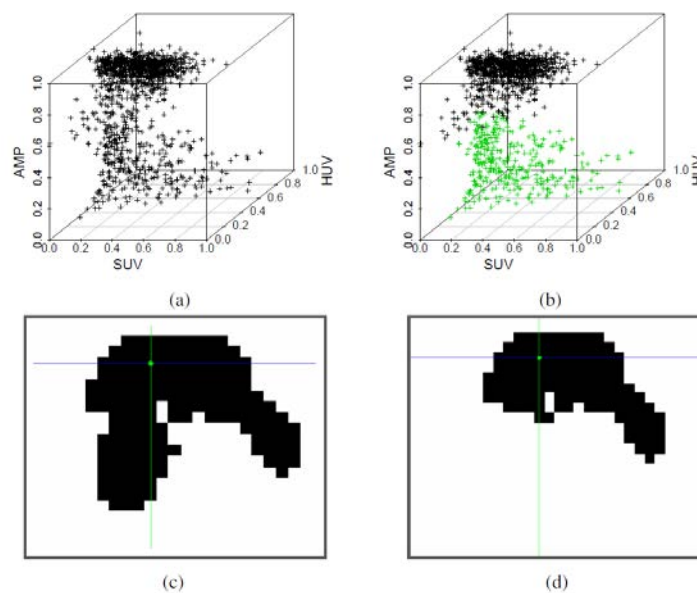


Figure 3.6 Data fusion and k-means clustering example. (a) voxels belonging to  $\{A, T\}_{icc}$  placed in the tri-parametric clustering space  $[SUV, HUV, AMP]$ : SUV value (SUV), Hounsfield unity value (HUV), and the artifact membership probability (AMP). (b) k-means clustering result: cluster  $\{A\}$  in black and cluster  $\{T\}$  in green. (c) Cluster  $\{A, T\}_{icc}$  and (d) the resulting cluster  $\{A\}$

### Area correction

Topological transformations, namely closing and opening, were applied to ensure the compactness of the resulting area formed by  $\{A\}$ . The resulting corrected area was labeled BMA. Figure 3.7 presents an example of an area correction for obtaining the BMA.

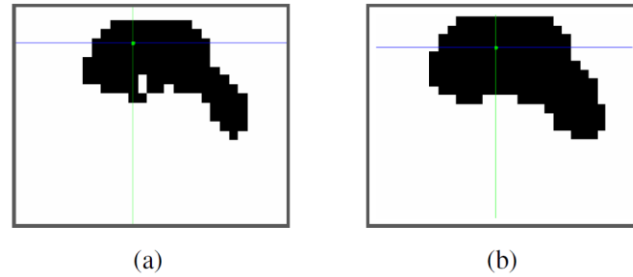


Figure 3.7 Area correction example. (a) Cluster {A}. (b) Bladder metabolic artifact (BMA) following area correction

### 3.3.3 VALIDATION

#### Measures of performance of BMA detection

We measured the performance of the BMA detection carried out by the proposed algorithm in two phases. First, we considered solely an expert manual bladder CT delineation as ground-truth (true-positives) and computed the sensitivity, specificity, and accuracy of the BMA detection. Secondly, we considered a tumor segmentation mask as the background (true-negatives) and computed the related sensitivity, specificity, and accuracy in order to assess the algorithm's tumor preservation capability. The bladder was manually segmented by an expert on each CT image using the Phillips Pinnacle<sup>3</sup> Version 8.0m treatment planning software. The CT images were then linearly down-sampled to the PET resolution so that each PET voxel corresponded to only one voxel of the CT, and likewise for the bladder delineations. Tumor metabolic activities were segmented on PET images by region-growing using a threshold of 42% of the maximum uptake within the tumor ( $T_{42}$ ). Residual bladder uptake in the  $T_{42}$  segmentation mask was suppressed by the corresponding manual bladder CT delineation.

The voxels outside the bladder mask were considered true negatives (TN), whereas the voxels within the bladder mask were the true positives (TP). The number of voxels originally located in the bladder and not detected as artifacts were labeled false negatives (FN). Lastly, the number of voxels originally located outside the bladder mask and detected as artifacts were labeled false positives (FP).

In parallel, the detection was evaluated a second time using the  $T_{42}$  segmentation mask. The number of voxels detected as artifacts yet originally located inside the  $T_{42}$  mask were this time considered false positives ( $FP_{T_{42}}$ ), whereas the number of voxels in the  $T_{42}$  mask that were not detected as artifacts were labeled true negatives ( $TN_{T_{42}}$ ).

Figure 3.8 illustrates the PET and CT delineations used to evaluate the detection of bladder metabolic artifacts performed by the proposed methodology.

For the sake of clarity, we describe below how we computed the measures of performance of the detection method for each PET image, namely sensitivity (SEN), specificities (SPE and  $SPE_{T_{42}}$ ), and accuracies (ACC and  $ACC_{T_{42}}$ ):

$$SEN = \frac{TP}{TP + FN}$$

$$SPE = \frac{TN}{TN + FP}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$SPE_{T42} = \frac{TN_{T42}}{TN_{T42} + FP_{T42}}$$

$$ACC_{T42} = \frac{TP + TN_{T42}}{TP + TN_{T42} + FP_{T42} + FN}$$

### Reproducibility evaluation

Reproducibility was assessed by randomly selecting 30 different seed pairs ( $s_b$  and  $s_t$ ) within a  $2 \times 2$ -cm region around the initial manually-selected seeds. Thus, 30 different BMA detections were performed for each image.

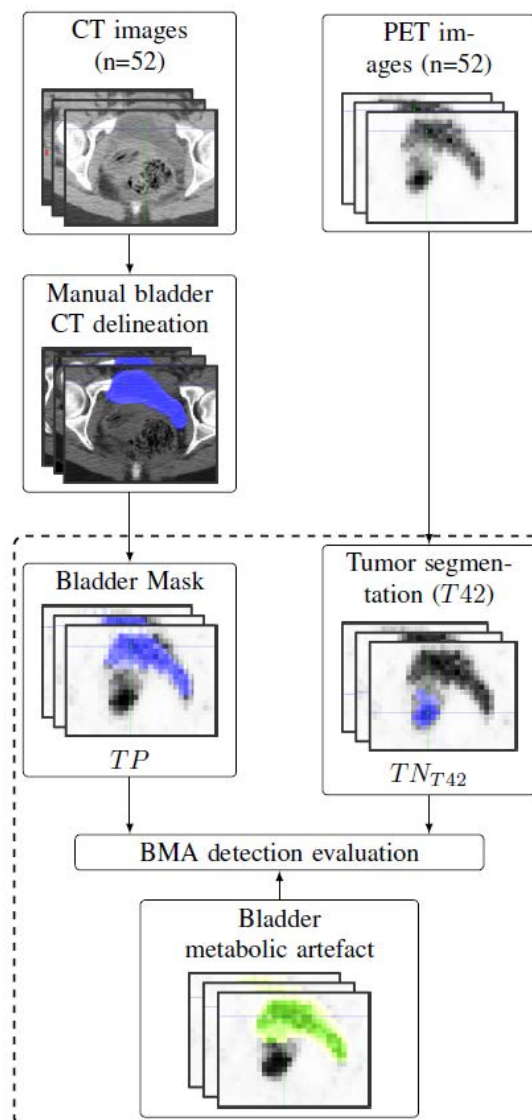


Figure 3.8 PET and CT delineations used for evaluating the BMA detection method. Bladder Mask: Expert manual bladder CT delineations linearly down-sampled to the PET resolution. Tumor Segmentation (T42): Metabolic activities segmented on PET images by region-growing using a threshold of 42% of the maximum uptake within the tumor

### 3.4 RESULTS

In this study, 47 out of 52 PET images presented tumor uptake in region  $\{A, T\}_{lcc}$  and were considered in the second clustering step. Figure 3.9 presents four examples of BMA detection using the proposed method.

The values attesting to the proposed detection method's performance (SEN, SPE,  $SPE_{T42}$ , ACC, and  $ACC_{T42}$ ) for all PET images of our dataset have been presented using a boxplot representation in Figure 3.10. For all 52 BMA detections evaluated, the averaged results were  $0.94 \pm 0.09$  SEN,  $0.98 \pm 0.01$  SPE, and  $0.98 \pm 0.01$  ACC. For the  $T42$  segmentation, the averaged results were  $0.97 \pm 0.05$   $SPE_{T42}$  and  $0.94 \pm 0.08$   $ACC_{T42}$ .

In the 30 seed pairs randomly selected for each image, a total of 1,560 BMA detections ( $52 \times 30$ ) were obtained to enable assessment of the method's reproducibility.

For these BMA detections, the average SEN, SPE, and ACC values were  $0.91 \pm 0.13$ ,  $0.98 \pm 0.02$ , and  $0.98 \pm 0.02$ , respectively. For the  $T42$  mask, the average SEN values were the same, and the  $SPE_{T42}$  and  $ACC_{T42}$  averages were  $0.96 \pm 0.07$  and  $0.92 \pm 0.12$ , respectively.

The overall robustness of the results for the whole cohort is depicted in Figure 3.11. This demonstrates that for most of the seed locations, the same segmentation was obtained. However, if the seed was located outside the bladder, the regions could naturally be erroneously detected.

In some cases, SEN=0 due to the random placement of seed  $s_b$  outside the bladder.

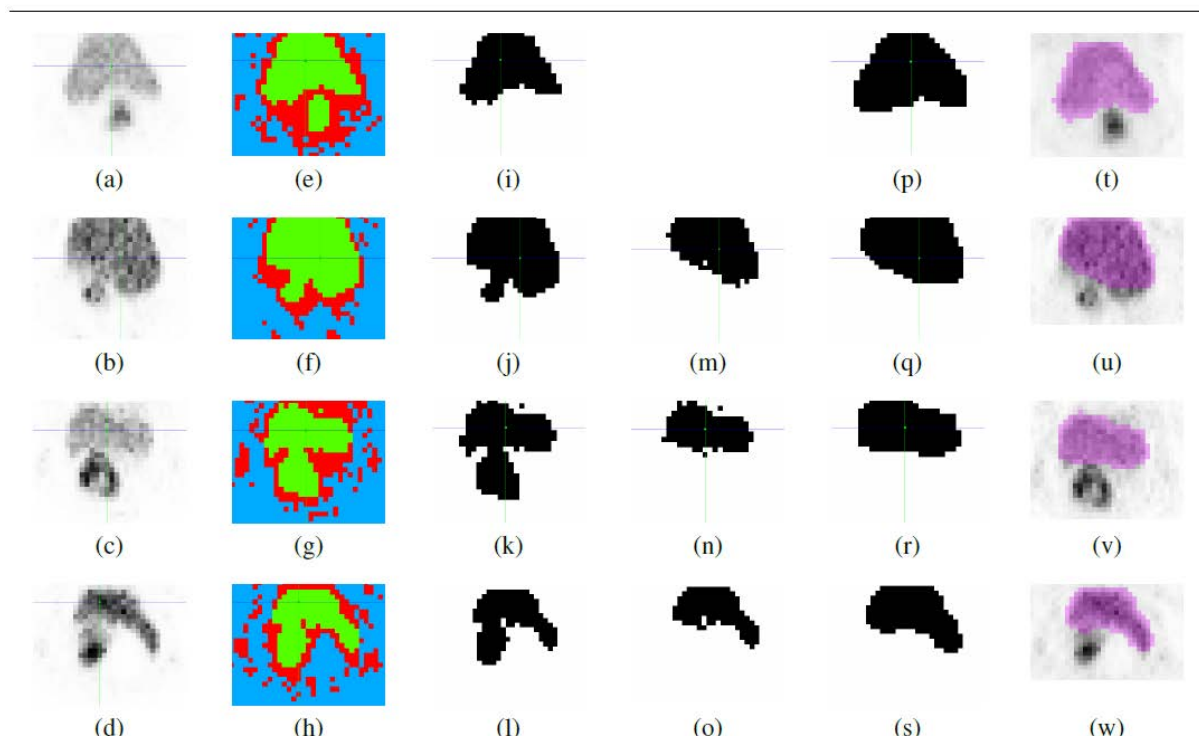


Figure 3.9 Example of BMA detection results for four patients. (a)-(d) Original PET exams. (e)-(h) 2.2.1. SUV-based k-means clustering results. (i)-(l) Clusters  $\{A, T\}_{lcc}$ . (m)-(o) Cluster  $\{A\}$  following 2.2.3. data fusion and k-means clustering. (p)-(s) BMA regions following 2.2.4. area correction. (t)-(w) PET image with BMA highlighted

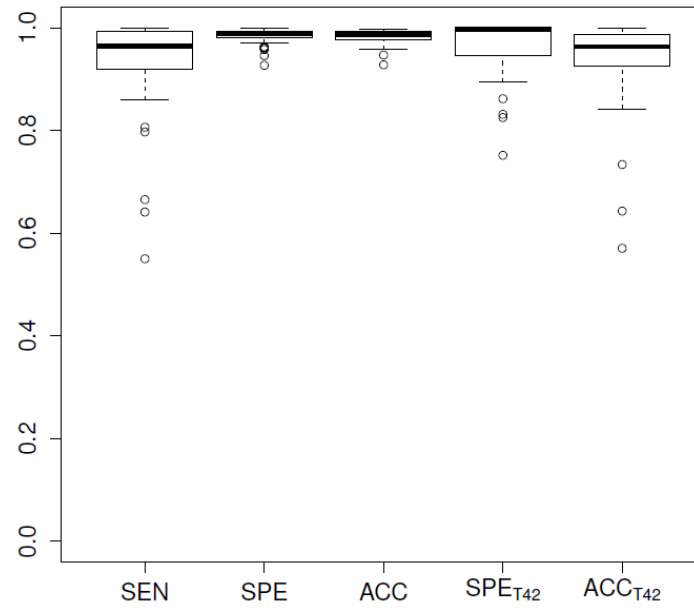


Figure 3.10 Values of the proposed algorithm's performance for the 52 PET images. SEN, SPE, and ACC averages corresponding to the evaluation of the BMA detection method, only considering bladder mask. SPET<sub>T42</sub> and ACC<sub>T42</sub> correspond to the evaluation of the BMA detection method considering both bladder mask and T42 mask



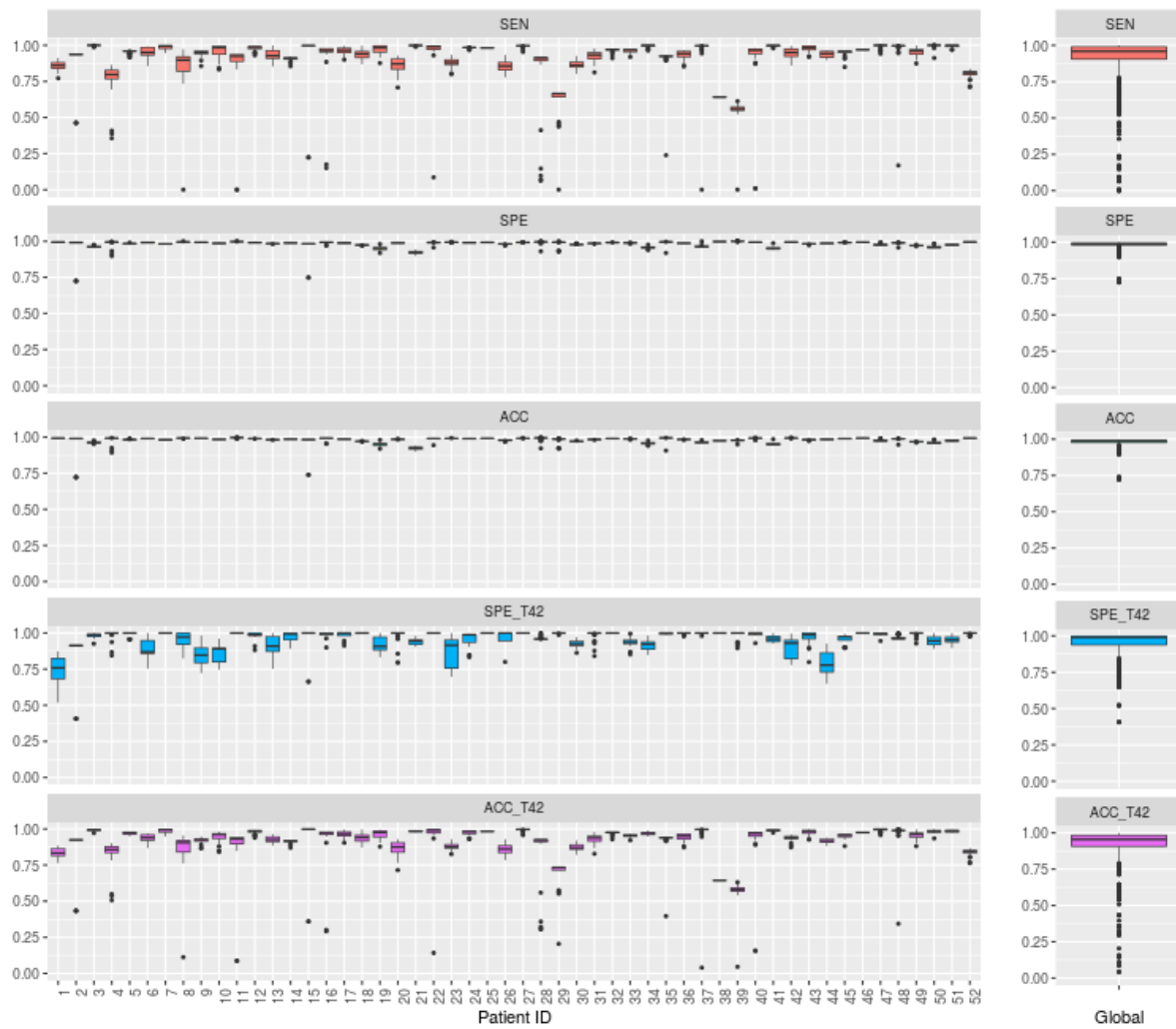


Figure 3.11.: Overall results for the 52 patients for different seed locations. SEN, SPE, and ACC values correspond to the evaluation of the BMA detection method only considering bladder mask. SPE<sub>T42</sub> and ACC<sub>T42</sub> correspond to the evaluation of BMA detection performance using both bladder mask and T42 mask

### 3.5 DISCUSSION

In PET images,  $^{18}\text{F}$ -FDG fixation can be observed in healthy tissue and benign disease, which are consequently mistaken for cancer [164]. PET analysis can be particularly hampered in cancers involving the pelvic region by the physiological uptake of the bladder, which can obliterate the tumor uptake of lesions located in this area [150].

In this paper, we propose a semi-automatic method to detect  $^{18}\text{F}$ -FDG bladder artifacts in PET imaging of pelvic cancer using a k-means clustering approach by combining PET and CT information. Artifact detection was considered to be a means of preprocessing PET voxel classification, and particular attention was paid to the preservation of tumor metabolism to enable further non-parasitized PET analyses.

The problem of detection was tackled by a two-step clustering relying on two assumptions. 1) That only the voxels with high intensity would hamper the analyses of tumor metabolism in PET imaging. To select these voxels, we applied an SUV-based clustering algorithm with three classes (low, intermediate, and high intensities). 2) That the bladder artifacts would be the sole artifacts detected among the previously-selected voxels. Therefore, the high-intensity voxels could belong to either artifacts or tumors. A second clustering step using CT and *a priori* information was then performed to separate voxels belonging to the artifacts from those belonging to tumor activity. As the number of classes was fixed, we used the k-means algorithm to ensure unsupervised clustering and low computational cost. Nonetheless, for one given patient, we visually decided if the second classification step was necessary. Numerous other clustering algorithms have also been published in the literature [167]. For instance, the affinity propagation algorithm [168] or density-based spatial clustering [169] could be used to automatically determine the number of clusters needed and, therefore, detect bladder artifacts in a one-step clustering process. Further investigation will be required to evaluate the use of such algorithms for the current challenge of reducing user dependency by automating the decision procedure.

Our proposed bladder artifact detection algorithm was tested on 52 PET/CT images of patients with locally-advanced cervical cancer. We believe that the cancer's localization was relevant to our detection problem, as the bladder often presents an uptake continuum with the cervical tumor. However, in order to be applied to different pelvic cancer localizations, the algorithm would need to be adapted accordingly. To achieve this, further investigations should be carried out to evaluate the proposed method in other cancer localizations.

On our dataset of 52 PET images, only five were considered well-clustered following the first k-means clustering. This proves the interest of exploiting multimodal PET-CT data combined with *a priori* spatial information to distinguish bladder artifacts from tumor uptake.

We evaluated this proposed method firstly by comparing the detected bladder voxel artifacts with expert manual bladder CT delineation. Our results demonstrated the ability of the proposed method to classify voxels belonging to the bladder, with good average sensitivity, specificity, and accuracy ( $0.94\pm 0.09$ ,  $0.98\pm 0.01$ , and  $0.98\pm 0.01$ , respectively). We observed a lower sensitivity corresponding to a general underestimation of the bladder artifacts, which enabled the preservation of tumor activity. It should be noted that the  $^{18}\text{F}$ -FDG uptake of the bladder did not systematically correspond to the bladder's anatomical shape. The bladder actually fills during PET image acquisition, as it takes several minutes, while CT data is acquired in a matter of seconds. Manual CT delineations used for evaluation are thus not strictly representative of the bladder's real metabolism.

The method was further evaluated using a threshold method often used in clinical practice (T42) as a region to spare. Specificity and accuracy were recomputed according to the T42 mask. It should be noted that the T42 mask was not considered here as a ground truth of tumor metabolism, but rather as a reference

enabling evaluation of the method's tumor-preservation capability. With its average specificity ( $SPE_{T42}$ ) of  $0.97\pm 0.05$  and accuracy ( $ACC_{T42}$ ) of  $0.94\pm 0.08$ , it demonstrated good conservation of tumor metabolism.

In order to initialize the second step of clustering to separate bladder voxels from tumor voxels, an artifact membership probability (AMP), derived from the squared Euclidean distance of the *a priori* tumor and bladder locations, was computed. We assumed, however, that the seeds were positioned near the barycenter of the bladder and tumor metabolism, positions that can lead to result variability. The robustness of the proposed algorithm was assessed by randomly selecting 30 different initial seeds for each of the 52 PET images. The overall averages indicate that the method is globally precise and robust for the seed initialization, achieving  $0.91\pm 0.13$  SEN,  $0.98\pm 0.02$  SPE,  $0.98\pm 0.02$  ACC,  $0.96\pm 0.07$   $SPE_{T42}$ , and  $0.92\pm 0.12$   $ACC_{T42}$ . To reduce the variability, an alternative method that could be employed in future studies would be to include the possibility for users to select more seed points, thus incorporating more *a priori* information.

In conclusion, this proposed semi-automated method for detecting bladder metabolic artifacts on PET images of the pelvic region offers good preservation of tumor metabolism. This method can be used as a preprocessing step for further non-parasitized analyses.

### 3.6 CONCLUSION

Dans cette étude nous avons proposé une méthode de détection de l'hyperfixation vésicale visible sur l'imagerie TEP et pouvant gêner l'analyse du métabolisme tumorale dans la localité du col utérin. L'algorithme est considéré ici comme un prétraitement, permettant de marquer les voxels de haute intensité, ne correspondant pas à l'activité tumorale, et ayant un impact négatif sur la quantification TEP.

La méthode se base sur une double étape de classification ; la première étape permettant de se focaliser sur les voxels de haute intensité, et la deuxième étape permettant de discerner les voxels de l'activité tumorale et les voxels de l'artefact vésical.

Testé sur 52 images TEP pré-thérapeutiques, la méthode a montré de très bon résultats de détection de l'hyperfixation vésicale et de préservation de l'activité tumorale ( $0.94 \pm 0.09$  de sensibilité,  $0.98 \pm 0.01$  de spécificité et  $0.98 \pm 0.01$  de précision en ne prenant en compte que le masque vésical). La méthode pourrait donc permettre d'augmenter la spécificité de l'analyse de l'imagerie TEP, notamment dans l'optique d'extraire des paramètres quantitatifs uniquement lié à l'activité tumorale.

Il est à noter que cette méthode nécessite le placement de graines (ou « seeds », en anglais) pour marquer le positionnement *a priori* de la classe « Tumeur » et la classe « Vessie ». Evidemment, ce positionnement manuel peut engendrer une variabilité inter- et intra-opérateur. Toutefois, les méthodes d'évitement actuelles (placement d'une boîte englobante autour de la tumeur, délinéation manuelle de la vessie) sont également sujettes à ce genre de variabilité et nécessite généralement une mise en place plus longue et fastidieuse. Pour mesurer la variabilité de l'algorithme à l'initialisation, nous avons sélectionné, de manière aléatoire, 30 couples de *seeds* pour chacune des 52 images. Les résultats ont montré de très bons résultats moyens, démontrant la robustesse de l'algorithme à l'initialisation.

L'utilisation de l'algorithme de partitionnement des k-moyens se justifie par le fait que, pour chaque étape de classification, le nombre de classe était défini préalablement. En effet, initialisé de manière aléatoire, le nombre de classe est l'unique paramètre de l'algorithme des k-moyens. Toutefois, la méthode ne considère actuellement que l'artefact vésicale, sans prendre en compte la présence possible de digestifs, d'inflammations ou d'infections, qui peuvent également fixer le 18F-FDG dans cette localité et être confondu avec le signal tumoral. Une évolution possible de la méthode consisterait à prendre en compte un nombre variable d'artefact afin de rendre la méthode adaptable à différentes situations. D'une part celle-ci pourrait permettre d'unifier les deux étapes de classification actuelles. D'autre part cette amélioration permettrait l'application de cette méthode à toute localité du cancer.

Concernant la contribution de cette étude aux objectifs des travaux de thèse, plusieurs points sont à retenir :

- A notre connaissance, ceci est la première contribution qui aborde le problème de la détection automatique (ou semi-automatique) d'artefact dans l'imagerie TEP ;
- La méthode a montré de bonnes performances sur 52 examens TEP ;
- Parmi ces 52 examens, 47 présentaient un continuum d'hyperfixation entre la tumeur et la vessie) ;
- Afin de restreindre l'algorithme au sein de la région pelvienne et de garder les résultats comparables entre les différents examen TEP, nous avons défini une région d'intérêt aisément reproductible et dont les limitations sont des marqueurs anatomiques osseux définies sur l'acquisition TDM associé à l'image TEP.

À ce jour, l'impact de l'algorithme sur la valeur prédictive des paramètres n'a pas encore été évalué. De plus, comme vu précédemment, la méthode nécessite encore des améliorations et une plus forte validation avant de pouvoir être utilisée dans des études cliniques d'analyse TEP.



## 4 SEGMENTATION ET CARACTERISATION DU METABOLISME TUMORAL PRE- ET PER-THERAPEUTIQUE DES IMAGES TEP 18F-FDG POUR LA PREDICTION DE RECIDIVE (*COMMUNICATION AU WORKSHOP IGRT DE MICCAI 2012*)

*Ce chapitre présente une étude exploratoire menée sur l'évaluation de la valeur prédictive des images TEP pré- et per-thérapeutique au travers d'une méthode développée spécifiquement pour la localité du col (évitant l'hyperfixation vésicale). Pour cette étude, les images per-thérapeutique ont été ramenées dans l'espace anatomique des images pré-thérapeutiques pour être analysées au sein de la région de la tumeur initiale. Les paramètres extraits de ses deux signaux portaient sur des mesures d'intensités et de textures. La valeur prédictive des paramètres a été évaluée au travers de courbe de survie sans récurrence de Kaplan-Meier. Les résultats de l'étude ont montré que les mesures d'intensités de SUV (notamment et TLG) et le MTV en TEP pré- et per-thérapeutique ont permis de discriminer les patientes à risque de récurrence.*

*Le contenu de ce chapitre a été présenté au workshop « Image-Guidance and Multimodal Dose Planning in Radiation Therapy » à la conférence internationale « Medical Image Computing and Computer Assisted Interventions » (MICCAI) de 2012.*

---

## CONTENU

<b>4</b>	<b>SEGMENTATION ET CARACTERISATION DU METABOLISME TUMORAL PRE- ET PER-THERAPEUTIQUE DES IMAGES TEP 18F-FDG POUR LA PREDICTION DE RECIDIVE (<i>COMMUNICATION AU WORKSHOP IGRT DE MICCAI 2012</i>)</b> .....	<b>69</b>
4.1	Abstract .....	70
4.2	Introduction .....	70
4.3	Materials and Methods .....	71
4.3.1	Data and Clinical Protocol .....	71
4.3.2	Segmentation & Characterization of the metabolic tumor region .....	71
4.3.3	Tumor recurrence prediction .....	75
4.4	Results .....	75
4.4.1	FGMC evaluation .....	75
4.4.2	Explanatory features identification .....	75
4.5	Discussion .....	78
4.6	Conclusion .....	79

---

## Segmentation and characterization of tumors in 18F-FDG PET-CT for outcome prediction in cervical cancer radio-chemotherapy

Geoffrey Roman-Jimenez<sup>1,2</sup>, Julie Leseur<sup>2,3</sup>, Anne Devillers<sup>1,2,3</sup>, Juan David Ospina<sup>1,2</sup>, Guillaume Louvel<sup>2,3</sup>, Pascal Haigrón<sup>1,2</sup>, Antoine Simon<sup>1,2</sup>, Renaud de Crevoisier<sup>1,2,3</sup>, and Oscar Acosta<sup>1,2</sup>

<sup>1</sup> INSERM, U 1099, Rennes, F-35000, France

<sup>2</sup> Université de Rennes 1, LTSI, F-35000, France

<sup>3</sup> Département de Radiothérapie, Centre Eugène Marquis, Rennes, F-35000, France

### 4.1 ABSTRACT

Cervical cancer is one of the most common cancer to affect women worldwide. Despite the efficiency of radiotherapy treatment, some patients present recurrence. Early unfavorable outcomes prediction could help oncologist to adapt the treatment. Several studies suggest that tumor characteristics visible with 18F-FDG PET imaging before and during the treatment could be used to predict post-treatment recurrence. We present a framework for segmentation and characterization of metabolic tumor activity aimed at exploring the predictive value of pre-treatment and per-treatment 18F-FDG PET images. Thirty-five patients with locally advanced cervix cancer treated by chemoradiotherapy were considered in our study. For each patient, a coregistered PET/CT scan was acquired before and during the treatment and was segmented and characterized with our semi-automated framework. A segmentation process was applied on the baseline acquisition in order to find the metabolic tumor region (MTR). This MTR was propagated to the follow-up acquisition using a rigid registration step. For every patient, 40 features from the two MTRs were extracted to characterize the tumor changes between the two observation points. We identified explanatory characteristics by exploring the threshold which minimizes the p-value computed from the Kaplan-Meier free-disease survival curves. Seven features were identified as potentially correlated with cancer recurrence ( $p$ -value $<0.05$ ). Results suggest that our method can compute early meaningful features that are related with tumor recurrence.

### 4.2 INTRODUCTION

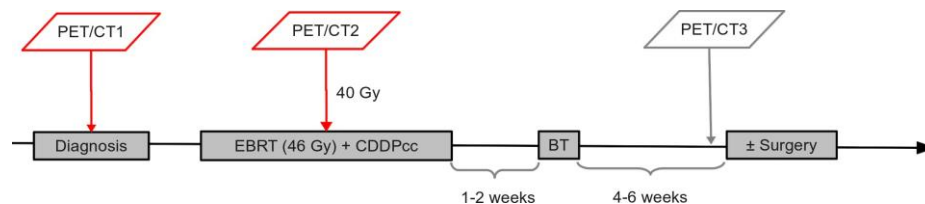
Nowadays, cervical cancer is the third most common cancer and the fourth cause of cancer death in females worldwide with 529,800 cases diagnosed and 275,000 cancer deaths in 2008 [170]. The standard treatment of locally advanced cervical cancer is based on chemoradiotherapy and brachytherapy, inducing a risk of acute and late irreversible toxicity [171]. Because recurrence in cervical cancer significantly increase the risk of death [10], prediction of such events is crucial. Medical imaging can provide different markers not only to plan the therapy but also to help oncologist to adapt the ongoing treatment thereby increasing the chance of patients' survival. During the last decade, several studies have shown that co-registered PET/CT improve the diagnostic accuracy in several cancers [159,160]. As opposed to anatomic imaging such as CT or MRI, PET provides metabolic information of the tumor. Since metabolic changes of the tumor precede the morphologic modifications, PET-based studies could provide early prediction outcomes. Several studies have demonstrated that metabolic tumor changes occurring between PET scans allow to quantitatively assess tumor response [40,89]. Further, some studies have shown the predictive value of the information extracted in PET [161,172].

In this study, we present a semi-automated framework for segmentation and characterization of metabolic tumor activity in cervical cancer from pre and per-treatment PET/CT acquisitions. We demonstrate the utility of our framework in predicting recurrence on a cohort of 35 patients treated for locally advanced cervix cancer. In the proposed framework the pre-treatment PET/CT fused information is exploited to first isolate the tumor region from the rest of the image, then several characteristics are extracted from the PET at two time points during the treatment to find features that may explain recurrence. The identification of the recurrence correlated features was performed using Kaplan-Meier survival curves and the log-rank test.

### 4.3 MATERIALS AND METHODS

#### 4.3.1 DATA AND CLINICAL PROTOCOL

Thirty-five patients (median age 52.44 years [32.15 - 84.62]) with locally advanced cervix cancer treated at the Centre Eugene Marquis (CEM), Rennes, France, were considered. The patients were treated with external beam radiation therapy (EBRT) with concurrent chemotherapy (CDDPcc) followed by brachytherapy (BT). With median follow-up of 29.21 months [7.44 - 52.64], eight patients developed tumor recurrence and three patients died. As shown in Figure 4.1, each patient underwent three 18F-FDG PET/CT scans performed with a DISCOVER ST scanner each observing the same time between injection and acquisition. In our study, we will only consider the pre- and per-treatment acquisitions. The two pre-treatment images are denoted PET1 and CT1. At 40 Gy of the EBRT, the acquisitions were denoted PET2 and CT2. In order to have the images comparables, each PET image was converted into standardized uptake value (SUV) which is a standardized decay-corrected value of 18F-FDG activity per unit volume of body weight (MBq/kg) [173].



**Figure 4.1 Acquisitions during the treatment protocol: A baseline exam, PET/CT1, is acquired before the treatment. A second acquisition, PET/CT2, is performed at 40 Gy of the chemoradiotherapy. The last acquisition, PET/CT3, is performed 4-6 weeks after the brachytherapy**

#### 4.3.2 SEGMENTATION & CHARACTERIZATION OF THE METABOLIC TUMOR REGION

The framework of the proposed method is illustrated in Figure 4.2. The first step aims to isolate the tumor area in PET1, avoiding the bladder hyperfixation. Because the tumor uptake in PET2 is less visible in responder patients due to the treatment, a rigid registration is carried out to map the region found in PET1 on PET2 thereby accessing to the same tumor region. In step three, the two regions found in the previous stages are used to characterize the tumor before and during the treatment through features computation.



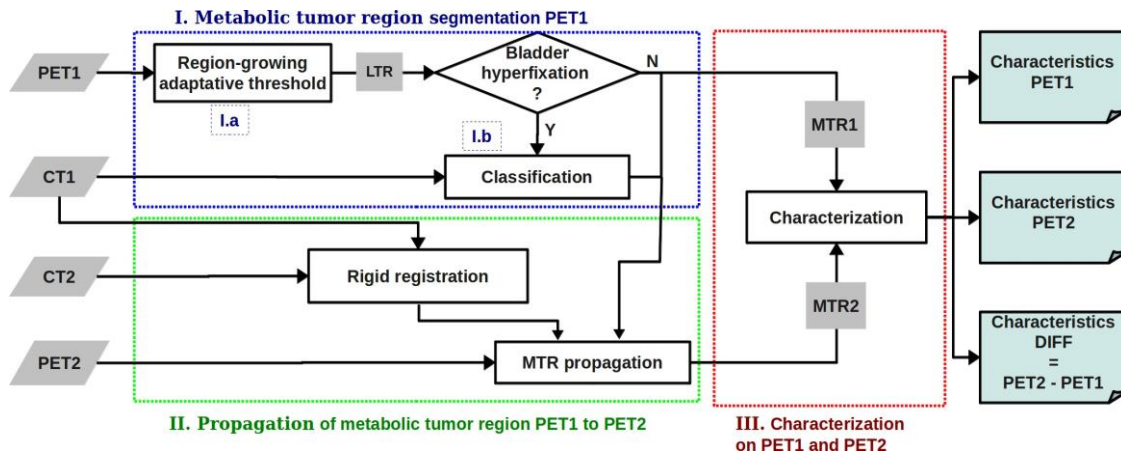


Figure 4.2 Overall framework for the characterization study.

### I - Metabolic tumor region segmentation in PET1:

The metabolic tumor region (MTR) is denoted as the region where the acquired  $^{18}\text{F}$ -FDG uptake corresponds to the tumor activity. The purpose of this step is to determine the metabolic tumor region in PET1 (MTR1). We considered here that the  $^{18}\text{F}$ -FDG hyperfixation of the tumor before the treatment is topologically compact (one connected object). In that case, a region-growing threshold allows to select the voxels belonging to the MTR1. Region-growing was used to limited the inclusion of adjacent intense structures, such as the bladder, lymph nodes, and bowel [51].

#### I.a - Region-growing adaptative threshold (RGAT):

Let's denote likely tumoral region (LTR) the region wherein the  $^{18}\text{F}$ -FDG hyperfixation is high enough to be considered as tumor metabolism. In order to extract the LTR, an adaptive threshold as the one proposed by Daisne [108] was implemented. This method aimed to adapt the threshold needed to segment the LTR according to the signal-to-background (S/B) measured in the image. The relationship between S/B and the threshold required was set up through a physical phantom study. Thus, because of the compactness hypothesis, the threshold was applied using a region-growing in order to keep a compact LTR.

#### I.b - Fusion and Gaussian mixture based classification (FGMC):

Due to natural filling and emptying of the bladder,  $^{18}\text{F}$ -FDG can be fixed in an extremely variable way leading to part of the bladder having intensities comparable to tumor metabolism. It was necessary to visually determine if voxels in the bladder were selected in the LTR. In such case, a second step of classification was required to separate bladder and tumor. Firstly, the CT images were down-sampled to the PET resolution so that each voxel in PET correspond to only one voxel in CT. A voxel seed was selected as being likely within the tumor. From each voxel selected in the LTR, three features were considered: the standard uptake value on PET (SUV), the Hounsfield unity value on CT (HU) and a tumor membership probability (TMP). The group of voxels in LTR was thus denoted as  $X = \{x_i | i \in I\}$ . With  $x_i$ , a three-feature vector  $[SUV \ HU \ TMP]$  at the voxel  $i$  and  $I$ , the 3D coordinates in the image. We computed the TMP as:

$$TMP(x_i) = 1 - \frac{d(x_i, s)}{\max_i d(x_i, s)}$$

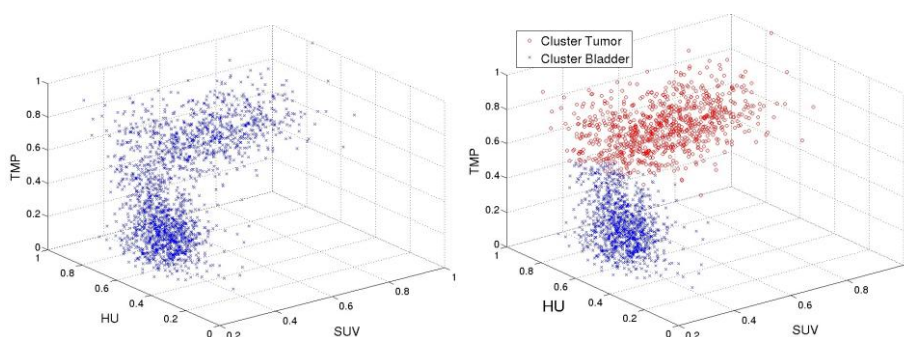
where  $d(x_i, s)$  is the Euclidean distance of the voxel  $x_i$  from the seed  $s$ , and  $\max_i d(x_i, s)$  denote the Euclidean distance of the furthest voxel in the LTR from the seed. It was assumed that the further the voxel is from the reference the lower is the probability to belong to the tumor. SUV and HU were normalized between 0 and 1. The three normalized features were then projected in the tri-parametric space represented in Figure 4.3. Voxels belonging to the tumor were identified by fitting a Gaussian mixture model

(GMM). In general terms, the GMM expresses that the distribution of points in this tri-parametric space is a sum of Gaussian functions. Two clusters were considered: a cluster  $T$  representing voxels in the MTR and a cluster  $B$  for voxels in the bladder:

$$g(x_i, \theta) = p_T \cdot f_T(x_i, \mu_T, \Sigma_T) + p_B \cdot f_B(x_i, \mu_B, \Sigma_B),$$

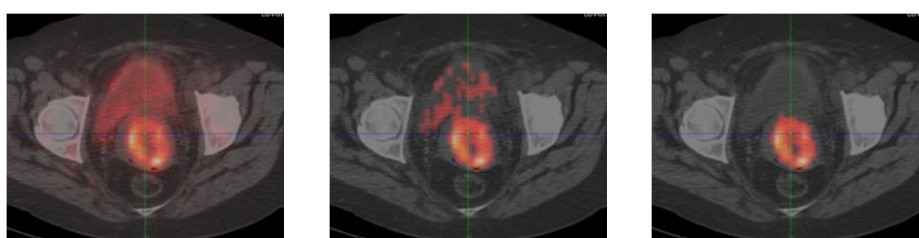
where  $g$  denotes the mixture density of each LTR-voxel  $x_i$ . The constants  $p_T$  and  $p_B$  are the mixing proportions of the two gaussian distributions  $f_T$  and  $f_B$  respectively characterized by the means  $\mu_T$  and  $\mu_B$  and variance matrix  $\Sigma_T$  and  $\Sigma_B$ .  $\theta$  represents the model parameters of the gaussians mix  $\theta = [p_T, p_B, \mu_T, \mu_B, \Sigma_T, \Sigma_B]$  need to be estimated.

An expectation-maximization algorithm (EM) was used to calculate the maximum likelihood estimates of  $\theta$ . Given these estimates, each voxel was assigned a label stating if it belongs to the cluster  $T$  or  $B$ . Figure 4.3 gives an example of clustering in the classification space.



**Figure 4.3 Fusion and gaussian mixture based classification. (left) Tri-parametric classification space : SUV value, Hounsfield value and Tumor Membership Probability. (right) resulting EM estimation**

The largest connected component in the cluster "Tumor" was selected to be the MTR. Figure 4.4 summarizes an example of metabolic tumor region segmentation in the pre-treatment PET.



**Figure 4.4 Example of metabolic tumor region segmentation in PET1. (left) PET/CT1 with a high bladder uptake. (middle) LTR after RGAT. (right) MTR1 after the FGMC.**

## II - MTR1 to MTR2 registration:

This step aims to propagate the MTR extracted in the PET1 to PET2. The propagated MTR was denoted MTR2. Since PET are co-registered with CT, these CT images were used to determine the transformation between PET images. A two-step registration using a block-matching algorithm [174] was adopted. Firstly, CT1 to CT2 were registered in order to match the whole body. Secondly, a rigid registration was applied locally to align the cervix. In practice, this is obtained by registering a VOI around the cervix. Note that local

rigid registration was assumed because of the restricted VOI considered. This allows to keep the same shape between MTR in PET1 and PET2 and thus keep the same tumor area.

### III - Characterization:

Fourteen features were extracted from the two PET MTR to characterize patient's tumor before and during the treatment. In addition to SUVmax, metabolic tumor volume (MTV) and total lesion glycolysis (TLG) often reported in literature [16,175,176], we explored intensity and texture features [140,177]. The features are summarized in Table 4.1 and Table 4.2.

Note that, in PET1, the total activity by voxel volume (TAVV), is equivalent to the total lesion glycolysis (TLG) calculated as  $SUV_{mean} \times MTV$  when  $MTV$  is well defined. Since we estimated the  $MTR_2$  by registration, the TAVV cannot be denoted as TLG because MTV in PET2 is roughly delineated. However, for a good responder, it was expected that if the activity outside the true MTV was neglected, the TAVV will give a good approximation of the TLG.

Texture features, originally proposed by Haralick [144], are based on the co-occurrence matrix  $\Phi$ , reflecting spatial grey-level dependencies. The intensity range into the MTR was quantified with 16 bins. The texture features calculated shown in Table 4.2 were computed for every direction covering the 26-connected neighborhood and were averaged to keep a restricted number of characteristics. Finally, to characterize the changes between PET1 and PET2, the difference (DIFF) was calculated between features evaluated in MTR1 and MTR2.

**Table 4.1 Evaluated intensity Features**

Features	Formula	Definition
SUVmax	$max(X)$	Maximal uptake in the MTR
SUVpeak	$\frac{1}{K} \sum_k x_k, k \in max_{neighbor}$	SUVmax averaged by its 26 neighbors
SUVmean	$\mu = \mathbb{E}[X]$	Average of SUV in the MTR
SUVvariance	$\Sigma = \mathbb{E}[(X - \mu)^2]$	Variance of SUV in the MTR
SUVskewness	$\gamma_1 = \mathbb{E} \left[ \left( \frac{X - \mu}{\Sigma} \right)^3 \right]$	Asymmetry measure of the MTR activity distribution
SUVkurtosis	$\gamma_2 = \mathbb{E} \left[ \left( \frac{X - \mu}{\Sigma} \right)^4 \right]$	Peakedness measure of the MTR activity distribution
MTV	$Volume_{voxel} \times Nb_{Voxel}$	Metabolic Tumor Volume of the MTR
TAVV	$Volume_{voxel} \times \sum_n x_n$	Total Activity by Voxel Volume of the MTR

**Table 4.2 Evaluated textures features**

Features	Formula
Energy	$\sum_{i,j} \Phi(i,j)^2$
Entropy	$\sum_{i,j} \Phi(i,j) \times \log_2(\Phi(i,j))$
Inertia	$\sum_{i,j} (i-j)^2 \times \Phi(i,j)$
InverseDifferentMoment (IDM)	$\sum_{i,j} (i-j)^2 \times \Phi(i,j)$
Cluster Shade	$\sum_{i,j} (i+j-\mu_j-\mu_i)^3 \times \Phi(i,j)$
Cluster Prominence	$\sum_{i,j} (i-\mu_i+j-\mu_j)^4 \times \Phi(i,j)$

### 4.3.3 TUMOR RECURRENCE PREDICTION

A total of 40 features were extracted from 35 patients. Fourteen features were extracted from PET1, 13 features from PET2 and 13 from DIFF. The Kaplan-Meier method was used to evaluate the disease-free survival curves of the group splitted by a threshold. The feature assessment was performed following this method: for each feature, a threshold was found in order to minimize the p-value of the log-rank test from the comparison of the two survival curves generated when splitting the group using this threshold. The level of significance for the log-rank test was fixed to p-value<0.05.

## 4.4 RESULTS

### 4.4.1 FGMC EVALUATION

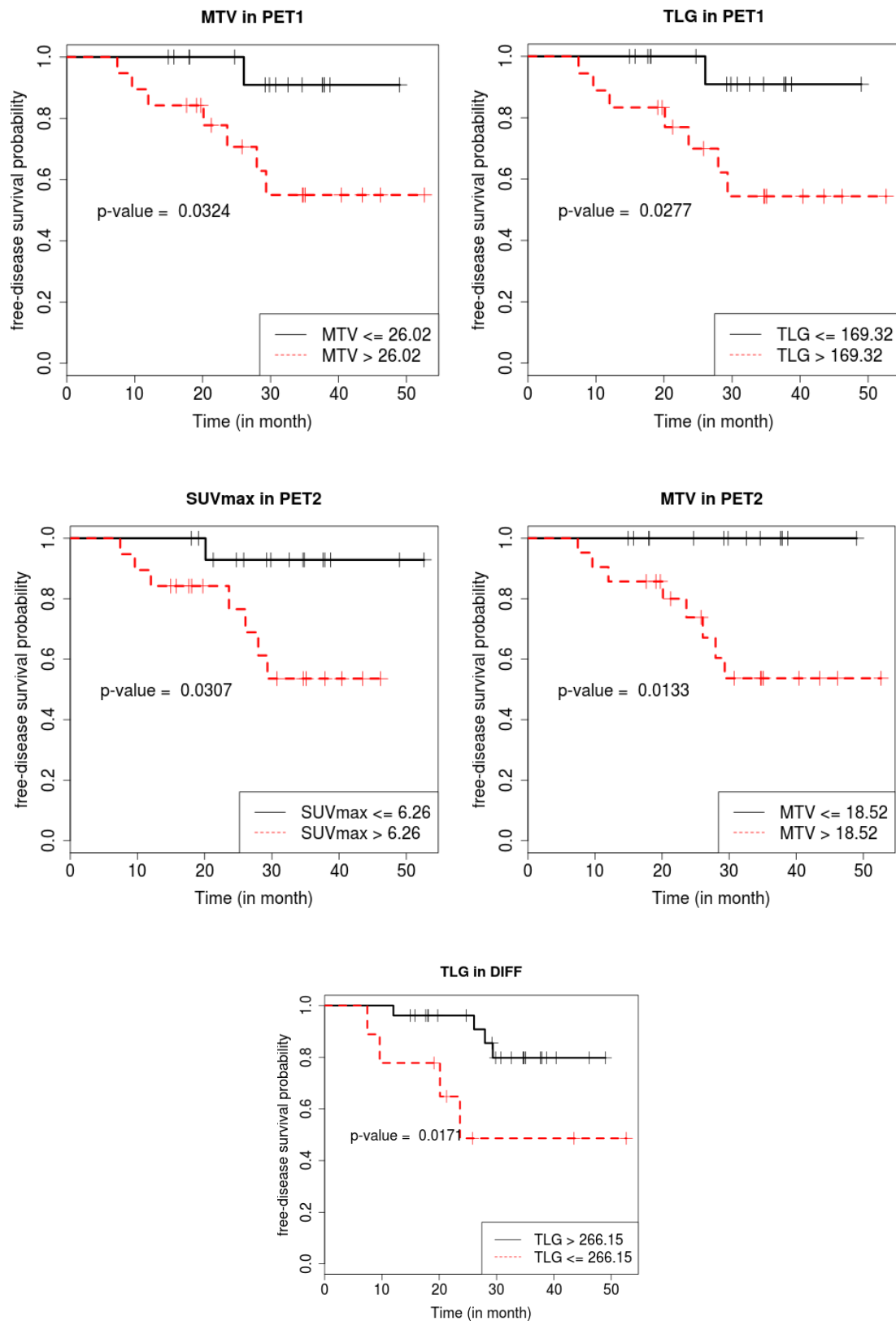
Among the 35 PET1 segmented in step I.a, 12 visually presented high bladder uptake and were clustered in step I.b. In order to evaluate the deletion of bladder voxels performed by the FGMC, we used expert manual cervix and bladder CT-segmentation available for six patients. After I.b, the number of voxels remained in the cervix mask was considered as true positive (TP) whereas false positive (FP) was the number of those in the bladder mask. For the six clustering evaluated, the averaged results were  $0.80 \pm 0.17$  sensitivity,  $0.97 \pm 0.08$  specificity and  $0.85 \pm 0.11$  accuracy. We can observe a lower sensitivity due to the worst patient which presented a classification with 0.53 sensitivity, 1.00 specificity and 0.63 accuracy. In this case, the patient presented a small tumor and as the number of voxels in the cervix mask in LTR was limited, any suppressed voxel during the FGMC had an important impact on the resulting sensitivity.

### 4.4.2 EXPLANATORY FEATURES IDENTIFICATION

Table 4.3 summarizes the significant features extracted in step III. Results show that the explanatory features found are mainly based on SUV intensity. The four characteristics found from PET1 might suggest that tumor recurrence could be predicted before the treatment. Moreover, the TAVV evolution between the two exams (DIFF) might be a recurrence correlated feature that could be used by oncologists to adapt the ongoing treatment. Five example of a Kaplan-Meier curve are displayed in Figure 4.5.

**Table 4.3 Assessed features by the Kaplan-Meier method with p-value <0.05 and the corresponding threshold.**

exam	feature	p-value	threshold
PET1	SUVmax	0.0335	10.52
	SUVpeak	0.0478	7.27
	SUVmean	0.0335	6.30
	MTV	0.0324	26.02
	TLG	0.0277	169.32
PET2	SUVmax	0.0307	6.26
	MTV	0.0133	18.52
	TAVV	0.0365	74.18
DIFF	TAVV	0.0171	-266.15



**Figure 4.5 Example of feature evaluation with Kaplan-Meier method**

(top left) MTV in PET1 ( $p=0.0344$ ); (top right) TLG in PET1 ( $p=0.0227$ ); (middle left) SUVmax in PET2 ( $p=0.0307$ ); (middle right) MTV in PET2 ( $p=0.0133$ ); (bottom) TLG in DIFF ( $p=0.0171$ )

## 4.5 DISCUSSION

In this study, we presented a whole framework to characterize cervical cancer tumor from  $^{18}\text{F}$ -FDG PET imaging and to predict its response to radiotherapy. One of the challenges was to isolate the tumor region from the bladder in PET1. With the use of CT information, we proposed a semi-automated method to segment and characterize PET imaging based on a RGAT followed by a FGMC approach. Since metabolic information can be hidden in lower uptake than the applied threshold, an intensity-based segmentation can lead to an underestimation of the true LTR extracted. Also, an adaptive threshold suffers from poor reproducibility [92] and a more evolved segmentation method could enhance this. The FGMC has demonstrated the ability to classifying voxels between tumor and bladder with good specificity, sensitivity and accuracy (respectively 0.80, 0.97 and 0.85). Nonetheless, for one given patient, we visually decided if the classification step was necessary. An integration of the FGMC into the segmentation step is in progress in order to reduce user dependency. The characterization step was performed by evaluating intensity and texture features on MTR1 and MTR2. Also, we characterized the tumor changes by computing the difference between the features in MTR1 and MTR2. In future work, shape metrics could be extracted to describe the tumor thereby providing complimentary information. We identified seven features that might be correlated with cervical cancer recurrence by assessing free-disease survival curves. This is an exploratory study so, in future work, the predictive capabilities of the found thresholds will be assessed through a larger cohort of patients.

## 4.6 CONCLUSION

Dans cette étude, nous avons exploré la valeur prédictive des images TEP pré- et per-thérapeutiques au travers de toute une chaîne de traitement et d'analyse d'images à visée de prédiction : prétraitement, segmentation, quantification, analyses statistiques.

Dans cette étude, nous avons proposé une méthode de segmentation, spécifique à la localité du col, permettant d'éviter l'hyperfixation vésicale. Celle-ci se base sur la combinaison d'un seuillage adaptatif (proposé par Daisne dans [108]) et d'une étape de classification par mixture de gaussiennes après fusion des informations TEP/TDM. Cette deuxième étape a montré de résultats satisfaisants en termes de classification. On notera cependant que celle-ci n'a été évaluée que sur six patients dont les délimitations manuelles étaient disponibles. Cette segmentation n'a été appliquée qu'aux images pré-thérapeutiques.

Afin de pouvoir analyser les images per-thérapeutiques, les images ont été projetées par recalage rigide dans l'espace anatomique des images pré-thérapeutiques. Le calcul des caractéristiques de la tumeur avant et en cours de traitement s'est effectué au sein d'une seule région d'intérêt calculée pour chaque tumeur initiale. L'évolution du signal a été évaluée en soustrayant les paramètres extraits séparément des images pré- et per-thérapeutiques.

Au total, 40 paramètres ont été analysés au regard de la valeur prédictive sur l'apparition de récurrence ; 14 paramètres pour les images TEP pré-thérapeutiques (7 paramètres d'intensité, 6 paramètres de texture et le volume tumoral), 13 pour les images per-thérapeutiques et 13 pour les différences entre ces paramètres (mêmes paramètres excepté pour le volume de la région tumorale calculé sur l'examen TEP initial).

La valeur pronostique des paramètres a été évaluée en minimisant la valeur-p associée au test log-rang calculé entre les deux courbes de survies de Kaplan-Meier construites à l'aide d'un seuil de séparation déterminé sur les paramètres.

Les paramètres discriminants étaient :

- le SUVmax, le SUVpeak, le SUVmean, le MTV et le TLG pour les images pré-thérapeutiques ;
- le SUVmax, le MTV et le TLG (calculés au sein de la région tumorale en initial) pour les images per-thérapeutiques ;
- la différence d'activité tumorale (TLG) entre les deux temps d'acquisitions des images TEP.

Certaines limitations doivent être prises en compte. En effet, avec une cohorte de 35 patientes, 9 événements de récurrence et un recul clinique médian de 29,21 mois, les analyses statistiques ne permettent pas de conclure sur une valeur seuil pour chaque paramètre, qui puisse être utilisée de manière fiable par d'autres centres.

Ces analyses nous permettent toutefois d'évaluer l'intérêt de certains paramètres dans notre objectif de prédiction. Nous noterons que le calcul de caractéristiques de texture proposées par Haralick (calculées sur une matrice de co-occurrence avec une quantification sur 16 niveaux de gris) ne nous a pas permis d'extraire des paramètres discriminants. Notre conclusion est donc que les paramètres de texture sont difficilement utilisables seuls, et qu'une investigation plus approfondie, notamment par association avec les mesures d'intensité identifiées comme prédictives, pourrait permettre d'évaluer leur apport dans la problématique de prédiction.

Par ailleurs, la méthode de segmentation proposée ici a l'avantage de s'adapter au contraste de l'image (le seuillage adaptatif de Daisne se calcule en fonction du contraste mesuré dans l'image) et d'éviter l'hyperfixation vésicale. Toutefois, l'impact de cette segmentation sur la valeur prédictive des paramètres



nécessite une étude comparative aux méthodes de seuillages plus conventionnelles afin de dégager les méthodes pertinentes dans notre objectif de prédiction.

## 5 IMPACT DE LA METHODE DE SEGMENTATION SUR LA VALEUR PRONOSTIC DU SIGNAL TUMORAL 18F-FDG TEP ET DE SON EVOLUTION AU COURS DU TRAITEMENT (ARTICLE PUBLIE AU JOURNAL IRBM 2013)

*Ce chapitre présente une étude explorant l'impact de la segmentation sur la valeur prédictive de paramètres extraits d'images TEP pré- et per-thérapeutiques et de l'évolution du signal tumoral entre ces deux examens. Les deux images ont été recalées entre elles à l'aide de leurs images TDM respectives. Trois signaux ont alors été analysés : l'image pré-thérapeutique, l'image per-thérapeutique, et l'image de l'évolution voxel-à-voxel entre les deux images (définie par la différence absolue entre chaque correspondance de voxel). Les paramètres calculés étaient le SUVmax, le SUVpeak, le MTV et le TLG. Sept seuillages fixes (de 20% à 80% du SUVmax, par pas de 10%) et une méthode de segmentation spécifique au col utérin (présentée au chapitre précédent) ont été testés pour évaluer leur impact sur la prédiction. La valeur prédictive des paramètres TEP a été analysée au travers de l'indice de concordance de Harrel, analogue de l'aire sous la courbe ROC pour les modèles actuariels. Les résultats ont montré premièrement, que la méthode de segmentation a un impact majeur sur la valeur pronostique des paramètres extraits, et deuxièmement, que le signal d'évolution peut fournir des paramètres possiblement plus prédictifs que le signal des images d'origine prises séparément.*

*Le contenu de ce chapitre a été publié au journal IRBM Volume 34, Issues 4–5, November 2013, Pages 274–277 Numéro spécial - RITS 2013 (doi : [10.1016/j.irbm.2013.09.002](https://doi.org/10.1016/j.irbm.2013.09.002)).*

### CONTENU

#### 5 IMPACT DE LA METHODE DE SEGMENTATION SUR LA VALEUR PRONOSTIC DU SIGNAL TUMORAL 18F-FDG TEP ET DE SON EVOLUTION AU COURS DU TRAITEMENT (ARTICLE PUBLIE AU JOURNAL IRBM 2013)

81

5.1	Abstract.....	83
5.2	Introduction .....	83
5.3	Materials and Methods.....	84
5.3.1	Data and clinical protocol .....	84
5.3.2	PET image analysis .....	84
5.4	Results.....	86
5.5	Discussion .....	87
5.6	Conclusion.....	88



# Investigating the contribution of pre and per-treatment 18F-FDG PET-CT segmentation methodologies for post-treatment tumor recurrence prediction in cervical cancer

Geoffrey Roman-Jimenez<sup>a,b,d,\*</sup>, Juan David Ospina<sup>a,b,c</sup>, Julie Leseur<sup>c</sup>, Anne Devillers<sup>c</sup>, Antoine Simon<sup>a,b</sup>, Pierre Terve<sup>d</sup>, Oscar Acosta<sup>a,b</sup> and Renaud De Crevoisier<sup>a,b,c</sup>

<sup>a</sup>INSERM, U1099, Rennes, F-35000, France

<sup>b</sup>Université de Rennes 1, LTSI, Rennes, F-35000, France.

<sup>c</sup>Département de Radiothérapie, Centre Eugène Marquis, Rennes, F-35000, France.

<sup>d</sup>Keosys Medical Imaging, Saint-herblain, F-44815, France.

## 5.1 ABSTRACT

Cervical cancer is one of the most common cancers to affect women worldwide. Despite the efficiency of radiotherapy treatment, some patients present post-treatment tumor recurrence, which increases the risk of death. Several studies suggest that tumor characteristics visible with PET imaging before and during the treatment could be used to predict post-treatment recurrence. We evaluate the contribution of pre- and per-treatment 18F-FDG PET images by exploring the predictive value of features extracted through several segmentation methods. Forty-one patients with locally advanced cervix cancer treated by chemoradiotherapy were considered. For each patient, two coregistered PET/CT scan were acquired before and during the treatment. A non-rigid registration was used to match the two PET acquisitions and evaluate the tumor metabolism inside the same area. Maximum and peak standardized uptake value (SUVmax and SUVpeak), the metabolic tumor volume (MTV) and the total lesion glycolysis (TLG) were evaluated. The predictive value of the extracted features was assessed through the Harrel's C-index. Results suggest that accurate segmentation can compute early meaningful features that are related with tumor recurrence. TLG seems to be strongly informative in prediction of tumor recurrence in cervical cancer.

## 5.2 INTRODUCTION

Nowadays, cervical cancer is one of the most common types of malignancy in women worldwide with 529,800 cases diagnosed and 275,000 cancer deaths in 2008 [170].

Locally advanced cervical cancer is treated with a combination of pelvic radiotherapy with concurrent cisplatin chemotherapy and followed by brachytherapy, with a 5-years survival rate of 70% and a 5-years overall recurrence rate of 30%; post-treatment tumor recurrences occur within the first 2 years after completing the chemoradiotherapy [178]. As recurrence in cervical cancer significantly increase the risk of death [10], early prediction of such events is crucial to place or correct the in course treatment. In the last two decades, positron emission tomography (PET) with 18F-Fluorodeoxyglucose (18F-FDG) has dramatically increase the accuracy in diagnosis of numerous cancers, and has significantly improved the radiotherapy planning the assessment of cancer patients at different stages of their disease [159,160]. Several studies have shown the predictive value of information extracted from PET acquired during the treatment [172,179,180]. These studies based their statistical analysis on features extracted through fixed thresholds segmentation methodologies. Thresholds techniques are widely used in clinical routine due to their ease to implement and use. However, thresholding is known to be significantly sensitive to noise and contrast

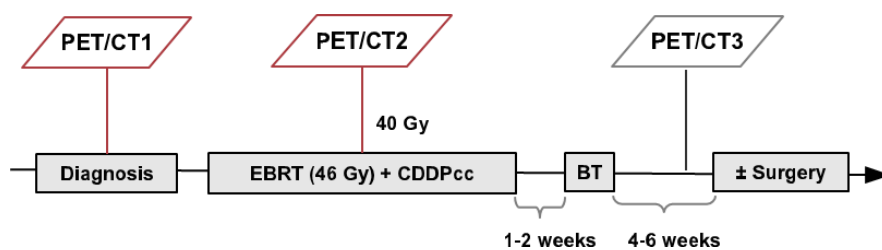
variation and often leads to the inclusion of adjacent intense structures. To our knowledge, no studies have demonstrated the effect of the segmentation accuracy in post-treatment outcomes prediction.

In this paper, we present a framework for segmentation and characterization of metabolic tumor activity in cervical cancer from pre and per-treatment PET/CT acquisitions. For the segmentation step, we proposed the use of a specific method in cervix segmentation avoiding bladder uptake [181]. Several fixed thresholds were also tested to compare their relevance in prediction of post-treatment tumor recurrence. From those segmentation methods, we extracted features often reported in literature: maximum and peak standardized uptake value (SUV<sub>max</sub> and SUV<sub>peak</sub>), the metabolic tumor volume (MTV) and the total lesion glycolysis (TLG). We evaluate the predictive value of these features by computing Harrell's C-index.

## 5.3 MATERIALS AND METHODS

### 5.3.1 DATA AND CLINICAL PROTOCOL

Forty-one patients (median age 52.38 years [26.16 - 84.56]) with locally advanced cervix cancer treated at the Centre Eugene Marquis (CEM), Rennes, France, were considered. The patients were treated with external beam radiation therapy (EBRT) with concurrent chemotherapy (CDDPcc) followed by brachytherapy (BT). With median follow-up of 22.79 months [4.79 - 55.39], nine patients developed tumor recurrence and seven patients died. As shown in Figure 5.1, each patient underwent three 18F-FDG PET/CT scans performed with a DISCOVER ST scanner each observing the same time between injection and acquisition. In our study, we will only consider the pre and per-treatment acquisitions. The two pre-treatment images are denoted PET1 and CT1. At 40 Gy of the EBRT, the acquisitions were denoted PET2 and CT2. In order to have comparable images, each PET image was converted into standardized uptake value (SUV) which is a standardized decay-corrected value of 18F-FDG activity per unit volume of body weight (MBq/kg) [173].



**Figure 5.1 Acquisitions during the treatment protocol: A baseline exam, PET/CT1, is acquired before the treatment. A second acquisition, PET/CT2, is performed at 40 Gy of the chemoradiotherapy. The last acquisition, PET/CT3, is performed 4-6 weeks after the brachytherapy**

### 5.3.2 PET IMAGE ANALYSIS

The framework to characterize the metabolic tumor along the treatment is illustrated in Figure 5.2. The segmentation step aims to isolate the tumoral area in PET1 to define the tumor region. The registration step aims to match PET2 to PET1 and characterize the tumor change inside the tumor region. The features extraction step aims to compute features characterizing the tumor metabolism along the treatment.

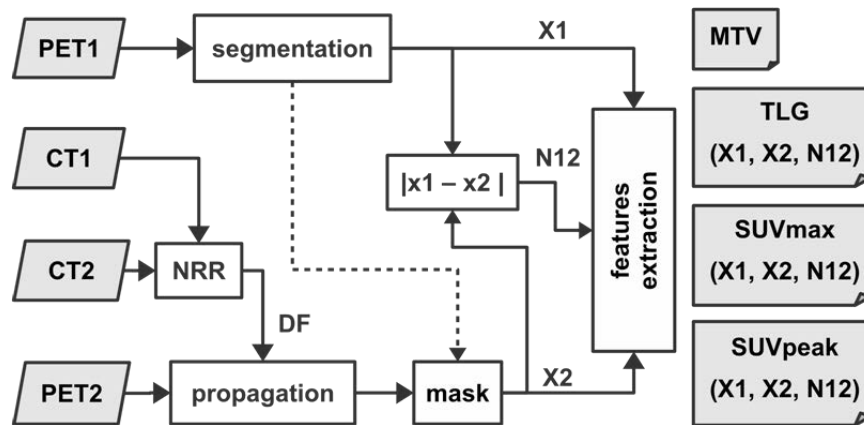


Figure 5.2 Overall framework to characterize the metabolic tumor along the treatment.

### Segmentation methods

We considered several methods to compare their relevance in predicting tumor recurrence. We segmented these images through 7 fixed thresholds (TH) in percentage of the SUVmax found inside the tumor (TH20%-TH80%). We also implemented a specific method in segmentation of tumor region in cervical cancer, proposed in [181], based on the use of a region-growing adaptive threshold [108] improved through an optional step called fusion and Gaussian mixture classification (RGAT+FGMC).

### Non-rigid registration

Since PET are coregistered with CT, these CT images were used to determine the transformation between PET images. A non-rigid registration (NRR) using a demon algorithm [182] was adopted. The deformation field (DF) computed between CT2 and CT1 was then applied to register PET2 to PET1.

### Features extraction

Three signals were defined to characterize the tumor before and during the treatment and the evolution between them:

- X1, the voxels in PET1 inside the tumor region;
- X2, the voxels in PET2 inside the tumor region;
- N12, the absolute difference between the voxels in PET1 and PET2 inside the tumor region.

From each signal, two different intensity based measurements were extracted: the maximum SUV found (SUVmax) and the SUVpeak, the SUVmax averaged by its 26 neighbors. These features are independent on the segmentation method. We computed the metabolic tumor volume (MTV) and the total lesion glycolysis (TLG) as the multiplication of the MTV and the average SUV inside the tumor region. TLG and MTV are dependent of the segmentation method.

### Harrel's concordance index

The C-index was used to evaluate the predictive value of a feature computed through each segmentation methods [183]. The C-index is an estimation of the probability that the predictions and the outcomes are concordant. The probability of concordance  $P_c$  is the probability that two randomly chosen patients  $i$  and  $j$  are ordered in the same way according both the marker and the outcome. This probability is defined as  $P_c = P(M_i < M_j | T_i > T_j)$ , where  $M$  is the marker (feature value) for the patient considered and  $T$

represents the time of occurrence of tumor recurrence or the last time of clinical follow-up. Note that the C-index is interpreted relatively to 0.5,  $c = 0.5$  corresponds to a non-informative feature and  $c = 1$  indicates perfect predictive discrimination.

## 5.4 RESULTS

For X1, X2 and N12 respectively, SUV based features were moderately discriminative with  $c=0.63$ ,  $0.68$  and  $0.65$  for SUVmax and  $c=0.67$ ,  $0.67$  and  $0.63$  for SUVpeak. Figure 5.3 show the predictive values of the MTVs found according to every segmentation method. The MTV found with RGAT+FGMC was the most correlated with post-treatment tumor recurrence with a C-index of  $0.68$ . Best MTV for threshold were found with TH50 ( $c=0.67$ ). Figure 5.4 show the predictive values of TLG found over each signal for each segmentation method implemented. TLG computed with RGAT+FGMC presents good predictive value with a C-index of  $0.73$ ,  $0.72$  and  $0.74$  for X1, X2 and N12 respectively. C-index observed by thresholding were comparatively lower and best results were found with TH50 ( $0.69$ ,  $0.68$  and  $0.70$  for X1, X2 and N12 respectively).

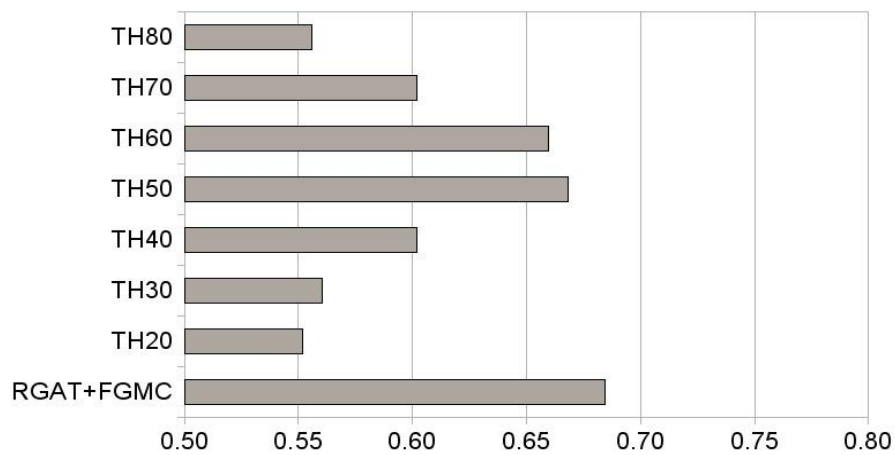


Figure 5.3 C-index of MTV with the associated segmentation method

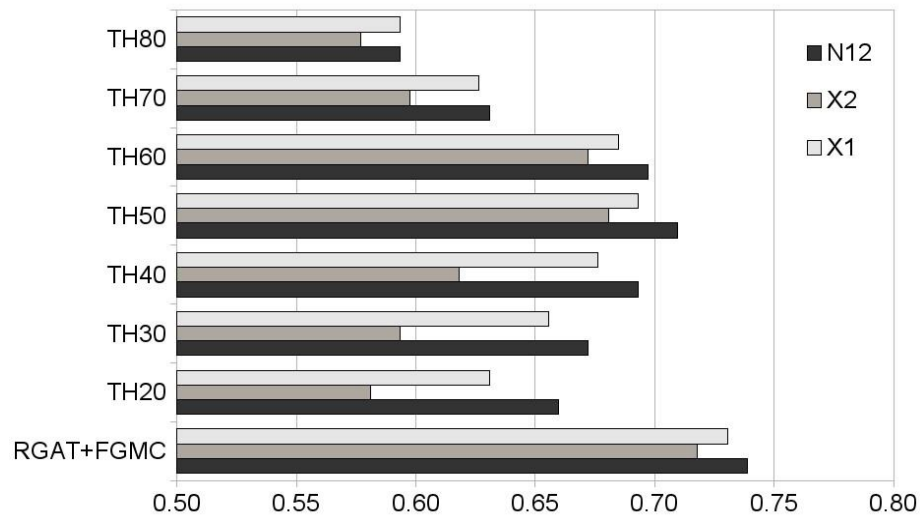


Figure 5.4 C-index of TLG found on X1, X2 and N12 with the associated segmentation method

## 5.5 DISCUSSION

In this study, we assessed the predictive value of features extracted from pretreatment and per-treatment PET images depending on the segmentation used. We implemented thresholds which are often used in clinical practice. A specific method for cervical tumor segmentation was also proposed to investigate the role of accurate segmentation in tumor recurrence prediction. We extracted features before the treatment (X1), during the treatment (X2) and the evolution between these two points (N12) to characterize the tumor metabolism along the treatment.

Maximum and peak standardized uptake value (SUVmax and SUVpeak) were extracted over X1, X2 and N12 to evaluate the tumor metabolism along the treatment; however, those features were moderately discriminative. Metabolic tumor volume (MTV) was computed through several segmentation methods. These results highlight the significant role of the segmentation method used in tumor recurrence prediction. Total lesion glycolysis (TLG) was extracted over X1, X2 and N12 through several segmentation methods. TLG computed over N12 outperformed TLG computed over X1 and X2, regardless the segmentation method used. This shows that the tumor evolution along the treatment is strongly informative and could be used in prediction of post-treatment tumor recurrence. C-index computed over TLG through fixed threshold were more discriminative with a threshold at 50% of the SUVmax (TH50). TLG extracted with RGAT+FGMC and computed over N12 gave the best results with a C-index of 0.74. These results show the feasibility of using the TLG as a predictor of post-treatment tumor recurrence in cervical cancer, provided that an appropriate method of segmentation is used.



## 5.6 CONCLUSION

Dans ce chapitre, nous avons présenté une étude portant sur l'impact de la méthode de segmentation sur la valeur prédictive de caractéristiques extraites du signal TEP pré- et per-thérapeutique ainsi que son évolution au cours du traitement.

Cette étude nous a permis de mettre en évidence les résultats suivants :

- la méthode de segmentation a un impact majeur sur la valeur prédictive des paramètres TEP ;
- parmi les paramètres testés, le TLG (ou son analogue sur les images paramétriques d'évolution du signal) était le paramètre le plus performant ;
- concernant les paramètres MTV et TLG, notre méthode de segmentation spécifique à la localité du col donnait les meilleurs résultats pour chacun des signaux analysés. Le seuillage à 50% permettait toutefois d'obtenir des résultats similaires ;
- les images paramétriques d'évolution du signal (différence absolue) permettaient d'obtenir des paramètres qui surpassaient, en termes de prédiction, les paramètres extraits des images d'origine.

L'algorithme de segmentation RGAT+FGMC se base sur le seuillage adaptatif de Daisne et nécessite une étude préalable sur fantôme physique pour estimer la relation entre le rapport signal sur bruit (RSB) et le seuillage optimal ( $S_{opt}$ ). Cependant, cette méthode reste difficilement applicable en étude multicentrique car la relation  $S_{opt}(RSB)$  nécessite d'être recalculée pour chaque caméra TEP, à chaque variation de calibration et pour différents niveaux de RSB.

De plus, nous avons pu montrer dans cette étude que les seuillages d'intensité (appliqués par croissance de région) permettent d'obtenir des paramètres prédictifs similaires à notre méthode de segmentation plus sophistiquée. Et, contrairement aux problématiques liées au diagnostic du cancer ou au ciblage radiothérapeutique, la recherche de paramètres prédictifs n'impose pas une délimitation très précise du métabolisme tumoral réel, mais se focalise sur l'extraction d'informations discriminantes. Ainsi, malgré les limitations connues quant à l'utilisation des seuillages pour la délimitation des volumes tumoraux, ceux-ci ont l'avantage d'être très simples à mettre en œuvre et parfaitement reproductibles sur une même image. De plus, les seuillages d'intensité étant très utilisés en routine clinique, la détermination d'un seuil optimal pour l'extraction de paramètres permettrait d'accumuler plus rapidement des preuves multicentriques.

Au regard de ces constatations, la pertinence des seuillages pour la prédiction de récurrence est investiguée plus en profondeur dans la suite du manuscrit.

## 6 VALEUR PREDICTIVE DES PARAMETRES EXTRAITS DES IMAGES TEP AU 18F-FDG POUR LA RECIDIVE DES TUMEURS DU COL UTERIN (COMMUNICATION ESTRO 2014 ET SOUMISSION ARTICLE AU JOURNAL IJROBP EN FEVRIER 2016)

*Ce chapitre présente une étude sur l'évaluation de la valeur prédictive de paramètres TEP, testés au travers de différents seuillages d'intensité (de 30% à 70% du SUVmax, par pas de 1%). Le but de cette étude était double : i) analyser l'évolution des paramètres entre les deux examens TEP et ii) identifier les paramètres les plus discriminants pour la prédiction de récurrence et pour quels seuillages.*

*L'étude s'est basée sur une cohorte de 53 patientes ayant subi deux examens, une TEP de diagnostic et une TEP à 40 Gy du traitement radio-thérapeutique. Les paramètres étudiés étaient le SUVmax, le SUVmean, le MTV et le TLG. L'indice de concordance de Harrel a été utilisé pour déterminer le seuil d'intensité optimal pour la segmentation TEP. La significativité des paramètres TEP extraits a été analysée au travers d'analyses univariées et multivariées basées sur des modèles de survie à risque proportionnel de Cox. Enfin, la capacité des paramètres à discriminer les patientes ayant présentées, ou non, une récurrence post-traitement, a été évaluée en utilisant des courbes de survie de Kaplan-Meier.*

*Les études statistiques ont montré que les deux images pouvaient fournir des informations pour prédire la récurrence, avec un seuillage autour de 55% du SUVmax pour les images pré-thérapeutiques et un seuillage de 32% du SUVmax pour les images per-thérapeutiques. Les paramètres les plus prédictifs étaient le MTV (avec un seuillage à 55% du SUVmax) pour la TEP pré-thérapeutique et TLG (avec un seuillage à 32% du SUVmax) pour la TEP per-thérapeutique. En analyse multivariée, ces deux paramètres étaient indépendants, offrant ainsi la possibilité d'être combinés pour améliorer la prédiction.*

*Le contenu de ce chapitre a été présenté à la conférence internationale « European Society for radiotherapy and Oncology » (ESTRO) de 2014 et sera soumis au journal « International Journal of Radiation Oncology Biology and Physics » (IJROBP).*

---

## CONTENU

<b>6</b>	<b>VALEUR PREDICTIVE DES PARAMETRES EXTRAITS DES IMAGES TEP AU 18F-FDG POUR LA RECIDIVE DES TUMEURS DU COL UTERIN (<i>COMMUNICATION ESTRO 2014 ET SOUMISSION ARTICLE AU JOURNAL IJROBP EN FEVRIER 2016</i>)</b> .....	<b>89</b>
6.1	Abstract .....	91
6.2	Introduction .....	91
6.3	Material and Methods .....	92
6.3.1	Patients and tumor .....	92
6.3.2	Treatment and follow-up.....	92
6.3.3	18F-FDG PET/CT acquisitions.....	93
6.3.4	18F-FDG PET/CT image analysis.....	93
6.3.5	Statistical analysis .....	94
6.4	Results.....	94
6.4.1	Assessment of PET response to treatment (at 40 Gy) .....	95
6.4.2	Predictive value of 18F-FDG PET/CT parameters.....	95
6.5	Discussion.....	96
6.5.1	Feasibility and timing.....	96
6.5.2	Predictive value compared with literature .....	97
6.5.3	Optimal Cut-off .....	97
6.5.4	Limits.....	97
6.6	Conclusion .....	104

---

## Pre- and per-treatment 18F-FDG PET/CT parameters to predict recurrence and survival in cervical cancer

J Leseur <sup>a</sup>, G Roman-Jimenez <sup>b-c</sup>, A Devillers <sup>d</sup>, O Acosta <sup>b</sup>, JD Ospina-Arango <sup>b-e</sup>, D Guillaume <sup>a</sup>, A Simon <sup>b</sup>, P Terve <sup>c</sup>, V Lavoue <sup>f</sup>, R De Crevoisier <sup>a-b</sup>

<sup>a</sup> Radiotherapy Department, Center Eugene Marquis, Rennes, F-35000, France

<sup>b</sup> INSERM, U1099, Rennes, F-35000, France; and with Université de Rennes 1, LTSI, Rennes, F-35000, France.

<sup>c</sup> Keosys medical imaging, 1 impasse Augustin Fresnel, Saint-herblain, F-44815, France

<sup>d</sup> Nuclear Medecine Department, Center Eugene Marquis, F-35000, Rennes, France

<sup>e</sup> School of Statistics, Universidad Nacional de Colombia, Campus Medellin, Colombia

<sup>f</sup> Gynecology and Obstetrics Department, CHU Anne-de-Bretagne - Rennes, F-35000, France

### 6.1 ABSTRACT

**Purpose:** The purposes of our study were first to describe the evolution of metabolic parameters derived from two 18F-fluorodeoxyglucose positron emission tomography (FDG-PET) images acquired before and during concomitant chemoradiation therapy (CCRT) (at 40 Gy) for locally advanced cervical cancer (LACC), and secondly, to identify the most predictive parameters, testing different SUV thresholds of segmentation.

**Material and methods:** 53 patients treated for LACC by CCRT, followed by brachytherapy underwent FDG PET/CT before treatment (PET1) and at 40 Gy (PET2). PET analyzed parameters were: visual metabolic response, maximum standardized uptake value (SUVmax), metabolic tumor volume (MTV), and total lesion glycolysis (TLG). MTV were automatically segmented (region-growing) using various SUVmax (from 30% to 70%) and fixed SUV thresholds (all voxels > SUV value from 1 to 20). Predictive capabilities of PET parameters on disease free (DFS) and overall survivals (OS) were assessed by the Harrel's C-index and the Cox model.

**Results:** The most predictive threshold segmentations were 55% of SUVmax at the planning (PET1) and 32% of SUVmax at 40 Gy (PET2). MTV1(55% of SUVmax) ( $p=0.021$ ; OR= 1.06 (95%CI: 1.01-1.11)) and TLG2(32% of SUVmax) ( $p=0.050$ ; OR= 1.01 (95%CI: 1.00-1.02)) were independent predictor of DFS in multivariate analysis ( $c=0.82$ ). MTV1(55%) and TLG1(55%) were also predictive of OS ( $p<0.05$ ). Visual metabolic response and parameters calculated from fixed SUV threshold segmentations were less predictive than MTV or TLG calculated from percentage of SUVmax.

**Conclusions:** In case of cervix cancer irradiation, MTV and TLG calculated from both PET imaging (performed before and during treatment) can be used to predict recurrence and survival, at an optimal threshold around 55% for PET1 and 32% for PET2.

### 6.2 INTRODUCTION

The standard treatment of bulky and locally advanced cervical cancers (IB2-IVA) is concurrent chemoradiotherapy followed by brachytherapy, and sometimes surgery [184]. Using this therapeutic strategy, overall five year survival rates reported for locally advanced cervical cancer ranges from 70 to 80% [178]. Prognostic factors are mainly clinical stage, lymph node involvement and response to treatment assessed clinically or by MRI [6,185]. [18F]fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) is a more recent imaging used routinely for initial staging and follow-up in cervical cancer [8]. Prognostic impact

of pre-treatment FDG-PET has also been demonstrated, either by the semi-quantitative standardized uptake value (SUV) of the primary tumor or by assessing the lymph node status [49,53]. Recently, the prognostic value of quantitative parameters such as the metabolic tumor volume (MTV) or the total lesion glycolysis (TLG) have been investigated at diagnosis of cervix cancer with controversial results [57,61,62,186]. The clinical impact of early assessment of tumor response during chemoradiation by FDG-PET is still scarce and conflicting [67,180,187,188]. In first studies, no correlation was found with SUVmax, MTV or visual analysis performed during radiochemotherapy and patient's outcome [180,187,189]. Recently, Oh *et al.* demonstrated in a population of 60 cervix cancer, that diminution of more than 60% of SUVmax between pre-treatment and at 40 Gy was correlated with 5 year disease free survival (DFS) ( $p = 0.045$ ) [188]. SUVmax estimated at the 4th week of radiochemotherapy was significantly associated with post-treatment PET response ( $p = 0.037$ ) [67]. In these two studies, MTV quantified during treatment was however not predictive and the prognostic value of TLG during radiochemotherapy for cervix cancer was not assessed.

Moreover, the quantification of MTV or TLG related to the choice of the optimal threshold value of SUVmax to more accurately predict the risk of recurrence is also matter of debate: threshold of 40% [186] or 42% of SUVmax [62] or a fixed threshold of SUV > 2.5 [57] or SUV > 2 [61]? Indeed, to our knowledge, no study compared different thresholds of SUVmax segmentation, in FDG-PETs performed both before and during treatment, for cervix cancer recurrence prediction.

The purposes of our study were therefore first to describe the evolution of PET parameters during radiochemotherapy (RTCT) for locally advanced cervical cancers, and secondly to identify the most predictive parameters of recurrence and survival, testing different SUVmax and fixed SUV thresholds of segmentation.

## 6.3 MATERIAL AND METHODS

### 6.3.1 PATIENTS AND TUMOR

Between January 2007 and December 2011, 53 consecutive patients treated for locally advanced cervical cancer in a curative intent were included in this prospective trial and analysed. Eligibility criteria were: histological confirmed invasive cervix carcinoma, stage IB2-IVA according to the International Federation of Gynecology and Obstetrics (FIGO), complete staging comprising clinical staging, MRI and 18F-FDG PET scan. A para aortic lymph node dissection was performed in all the patients, except in case of hyperfixation in this area at the PET imaging. A pelvic lymph node dissection could also be performed.

The median age was 52 years (range, 26 - 81 years). The stage distribution according to the FIGO classification was as follows: stage IB2 in 21 patients (40%), stage IIA1 in 9 patients (17%), stage IIA2 in 3 patients (6%), stage IIB in 14 patients (26%), stage IIIA in 2 patients (4%), stage IIIB in 1 patient (2%) and stage IVA in 3 patients (5%). Histology was squamous cell carcinoma for 40 patients and adenocarcinoma for 18 patients. Thirteen patients presented an involvement of pelvic lymph nodes. Eight patients presented pathological hyperfixation of para-aortic lymph nodes on PET and 5 an involvement of para-aortic lymph nodes on lymph nodes dissection.

### 6.3.2 TREATMENT AND FOLLOW-UP

All patients were treated with a combination of three-dimensional (3D) conformal external beam radiation therapy (EBRT) and pulse-dose-rate brachytherapy. EBRT delivered 45 Gy in the pelvis at 1.8 Gy per fraction,

using an IMRT technique with 5 to 8 - 18 MV photon beams. A simultaneous integrated boost up to 55 Gy (at 2.2 Gy/fraction) was delivered in the positive pelvic nodes when shown in imaging. The target volume was extended to the para-aortic lymph nodes in case of lymph node involvement in the para-aortic lymphadenectomy (5 patients) or at the PET imaging (8 patients). Pulse dose rate (PDR) brachytherapy was performed two weeks after the end of EBRT using a standard Nucletron® ring applicator delivering a dose of 15 Gy into the intermediate risk CTV according to the GEC ESTRO recommendations [190,191].

In case of parametrial involvement a lateropelvic external radiation boost was performed after brachytherapy to deliver a total dose of 60 Gy (4 patients). Fifty-one patients received concurrent weekly cisplatin chemotherapy (40 mg/m<sup>2</sup> weekly for 4-5 cycles). Two patients did not receive concurrent chemotherapy because of co-morbidities or advanced age ( $\geq 80$  years). The total treatment duration did not extend 7 weeks.

For the 18 patients with stage IB2 and IIA-B having residual disease at clinical examination, or biopsy or PET or MRI 4 to 6 weeks after brachytherapy, a total hysterectomy was performed (8 patients with stage IB2, 3 patients with stage IIA1 and 7 patients with stage IIB). The follow-up consisted in clinical examination, MRI and FDG PET 4 to 6 weeks after the end of brachytherapy, at 3 months and then every 6 months.

### 6.3.3 *18F-FDG PET/CT ACQUISITIONS*

The 18F-FDG PET/CT were realized first before the beginning of radiation therapy (PET1), the second one after a mean dose of 40 Gy (range, 32-46 Gy) (PET2) and then in the follow-up. 18F-FDG PET/CT scans were acquired using a PET/CT device (Discovery ST; GE Healthcare), which consisted of a PET scanner (Advance NXi; GE Healthcare) and an 2-slice CT scanner (Light- Speed Plus; GE Healthcare). Approximately 60 minutes after the intravenous injection of 4 - 5 MBq of 18F-FDG (Flucis), acquisitions were performed until the base of the skull to the upper part of the thighs. Scanning parameters were 120 kV, 80 mA, 0.8-s tube rotation, and 3.75 mm slices spacing. PET images were displayed with an ordered-subset expectation maximization iterative reconstruction algorithm (OSEM, 2 iterations, 28 subsets) and a reconstruction iterative fully 3D algorithm (Discovery ST). Patients fasted for at least 4 hours, a premedication with diazepam was administered systematically, as well as intravenous injection of furosemide (0.02 ml / kg) 15 min prior acquisitions.

### 6.3.4 *18F-FDG PET/CT IMAGE ANALYSIS*

Both qualitative and quantitative analyses were performed to assess the response to treatment in the primary cervical cancer at 40 Gy.

The qualitative analysis was realized by a single experiment nuclear medicine physician, comparing the cervical tumor FDG uptake in PET1 (before treatment) and PET2 (at 40 Gy) and according to the recommendations [192]. Visual complete metabolic response (CMR) was defined as a decrease of the primary cervical tumor FDG uptake similar to the background level. Partial metabolic response (PMR) was defined as a residual activity in the primary cervical tumor corresponding to at least a decrease of 25% of the initial SUVmax. Stability was defined as persistent activity in the primary tumor with change in SUVmax less than 25% of the initial SUVmax (increase or decrease). Progression was defined as an increase of SUVmax in cervical tumor more than 25% of the initial SUVmax or as new foci of abnormal FDG uptake.

For the quantitative analysis, the full body PET/CT images were restricted to an anatomical 3D box. The bounds of the box were defined according to bony landmarks: inferiorly corresponding to the inferior edge of the ischium, laterally corresponding to the right and left edge of the ilium, superiorly defined by the bottom edge of the fifth lumbar vertebra (L5), posteriorly defined by the posterior border of the sacrum

and anteriorly defined by the anterior edge of the ischial pubic. The bladder was manually segmented. Voxels inside the bladder mask were suppressed on PET images for the analyses.

The following PET parameters were calculated in PET1 and PET2: the maximum standardized uptake values (SUVmax1, SUVmax2), the metabolic tumour volumes (MTV1, MTV2), and the total lesion glycolysis (TLG1, TLG2). The MTV was defined as the volume of the group of voxels having a SUV greater than a given threshold value. The TLG was defined as the MTV  $\times$  average SUV of voxels inside the MTV. The response to treatment at 40 Gy was quantified using SUVmax, MTV and TLG.

To calculate the impact of different SUVmax thresholds for MTV1 and TLG1, segmentations were performed in PET1 by region-growing, using first different SUVmax percentage thresholds (from 30% to 70% of SUVmax1 by step of 1%). Other segmentations were secondly performed defined on fixed SUV thresholds including all voxels with a SUV superior to a given value between 1 and 20 in the cervical area defined by the anatomical box in PET1.

In order to analyze the cervical area in PET2, regions formed by threshold of 40% in PET1 were first propagated from PET1 to PET2, using a rigid registration based on the block matching algorithm [193]. Segmentations on PET2 were then performed by region-growing using first different thresholds of SUVmax2 (from 30% to 70% of SUVmax2) and secondly defined on fixed SUV thresholds including all voxels with a SUV superior to a given value between 1 and 20, restricted in the propagated PET2 area. The corresponding MTV2 and TLG2 parameters were calculated.

To evaluate the interest of combining both PET to improve the recurrence prediction, the relative change in SUVmax, MTV and TLG between the two PET/CT images were calculated for each patient and tested for prediction.

The image analyses have been performed using C++ language and ITK libraries.

### 6.3.5 STATISTICAL ANALYSIS

Disease free survival (DFS) was the endpoint of the study. The duration of the DFS was defined as the time measured from the beginning of radiation therapy to the date of the recurrence event when detected, or to the date of the last follow-up visit. The recurrence event could be local and/or nodal and/or metastatic, detected by clinical examination and/or MRI and/or FDG PET. To estimate the predictive capabilities of PET parameters for DFS, Harrel's C-index were used to compare different models (c-index) (c-index  $\approx$  0.5  $\rightarrow$  not predictive, c-index  $\approx$  1  $\rightarrow$  predictive) [183]. The c-index was used to determinate the optimal SUVmax threshold giving the most predictive parameters for MTV and TLG.

The prognostic significance of the PET parameters estimated with optimal threshold were then analyzed in univariate and multivariate analyses using the Cox proportional hazards regression model. The actuarial DFS rate and survival curves were estimated using the Kaplan–Meier method, and the difference between survival curves was assessed by the log-rank test. MTV, TLG and SUVmax were analyzed as continuous variables. All statistical analyses were performed using R software.

## 6.4 RESULTS

The median follow-up was of 52 months [range: 27-74 months] for the 42 alive patients at the last visit, 36 of them being without evidence of disease. Sixteen patients had recurrence and 10 of them died from

cervical cancer. One patient died from cardiovascular disease (in complete response). The patterns of recurrences were: local recurrence in 10 patients (3 patients with local recurrences only and 7 patients had local recurrences with metastases), 3 patients had lymph nodes recurrences only (2 patients in the pelvis and para-aortic area and 1 patient in the para aortic area only) and 3 patients had metastatic disease without local recurrence. Five-year disease free survival and overall survival rates were 69% (CI95%: 56-82%) and 77% (CI95%: 65-89%), respectively. The FIGO stage was not statistically predictive of DFS ( $p=0.07$ ).

#### 6.4.1 ASSESSMENT OF PET RESPONSE TO TREATMENT (AT 40 Gy)

The visual analysis shown that CMR was observed in 19 patients and PMR in 34 patients.

Concerning the quantitative analysis, the median SUVmax1 (before treatment) and SUVmax2 (at 40 Gy) were 14.4 (3.5- 43.2) and 5.8 (1.7 – 20.9), respectively. Only 4 patients presented an increase of SUVmax during RTCT, two of them had a recurrence at the follow-up. Figure 6.1 represents the decrease of MTV (Figure 6.1A) and TLG (Figure 6.1B) estimated on PET/CT before treatment (in blue) and at 40 Gy (in red), according to the different thresholds of SUVmax (from 30% to 70%) used to estimate the two parameters.

The median MTV estimated with a threshold of 55% of SUVmax before treatment (MTV1(55%)) was 9.8 cc (0.4 – 44.9). The median MTV estimated with a threshold of 32% of SUVmax at 40 Gy (MTV2(32%)) was 11.8 cc (0.29 – 53.9). The median TLG estimated with a threshold of 55% of SUVmax before treatment (TLG1(55%)) was 74.9 (3.3 – 504.1). The median TLG estimated with a threshold of 32% of SUVmax at 40 Gy (TLG2(32%)) was 30.4 (1.2– 198.7).

#### 6.4.2 PREDICTIVE VALUE OF 18F-FDG PET/CT PARAMETERS

##### Predictive value of the visual metabolic response at 40 Gy

Two of the 19 patients with CMR and 14 patients among the 34 patients with PMR relapsed at the last follow-up time. The 3-year DFS rates for patients with CMR and PMR on PET2 were 94% (CI95%: 84-100%) and 61% (CI95%: 44-78%), respectively. CMR at PET2 was a significant predictor of DFS ( $p=0.04$ ; OR= 4.90 (95%CI: 1.11-21.56) ( $c=0.66$ ) and not of OS ( $p=0.08$ ).

##### Identification of the most predictive SUVmax (in %) and fixed threshold SUV segmentations for MTV and TLG (in PET1 and PET2)

The Figure 6.2 represents the C-index values of MTV and TLG depending on different thresholds of SUVmax (from 30% to 70%) in the two PET images (before treatment for PET1 and at 40 Gy for PET2), for DFS (Figure 6.2A) and OS (Figure 6.2B).

For DFS, MTV1 has a higher prediction capability than MTV2, especially for thresholds between 52% and 65%. TLG1 and TLG2 are both strongly correlated with DFS. The thresholds of SUVmax presenting the highest DFS prediction are, at diagnosis (in PET1): 55% for MTV1 ( $c=0.80$ ) and TLG1 ( $c=0.78$ ), and at 40 Gy (in PET2): 32% for MTV2 ( $c=0.71$ ) and TLG2 ( $c=0.78$ ). SUVmax2 has a higher DFS prediction capability than SUVmax1 ( $c=0.66$  and  $c=0.58$ , respectively), however smaller than MTV and TLG.

For OS, the thresholds of SUVmax presenting the highest DFS prediction are, at diagnosis (in PET1): 55% for MTV1 ( $c=0.71$ ) and TLG 1 ( $c=0.73$ ), and at 40 Gy (in PET2): 63% for MTV2 ( $c=0.67$ ) and TLG2 ( $c=0.73$ ).

MTV or TLG calculated from fixed SUV threshold segmentations (all voxels > SUV value from 1 to 20) were less predictive than MTV or TLG calculated from percentage of SUVmax, in PET1 and PET2 for DFS and OS. For DFS prediction, the c-index were, with a fixed threshold SUV > 2.5, at diagnosis: MTV1 ( $c=0.68$ ) and TLG1 ( $c=0.72$ ), and at 40Gy: MTV2 ( $c=0.62$ ) and TLG2 ( $c=0.64$ ).



### Predictive value of the PET parameters from PET1 (before treatment) and PET2 (at 40 Gy) considered separately

- In univariate analysis considering PET1 only, MTV1(55%) was highly predictive of DFS ( $p= 1.10^{-4}$ ; OR= 1.08 (95%CI: 1.04-1.13) and OS ( $p=0.01$ ; OR= 1.06 (95%CI: 1.01-1.11)). Figure 6.3A and Figure 6.3A show the DFS curves for patients with MTV1 < 9.8 cc (median value) and MTV1 > 9.8 cc ( $p= 6.10^{-5}$ ). TLG1(55%) was predictive of DFS ( $p= 0.001$ ; OR= 1.005 (95%CI: 1.002-1.008) and OS ( $p=0.03$ ; OR= 1.004 (95%CI: 1.000-1.008)). Figure 6.3B and Figure 6.3B show the DFS curves between patients with TLG1 < 74.9 (median value) and TLG1 > 74.9 ( $p= 1.10^{-4}$ ). SUVmax1 was not predictive of DFS ( $p=0.62$ ).

- In univariate analysis considering PET2 only, MTV2(32%) was significantly predictive of DFS ( $p= 3.10^{-4}$ ). Figure 6.4A shows the DFS curves for patients with MTV2 < 11.8 (median value) and MTV2 > 11.8 ( $p= 0.01$ ). TLG2(32%) was significantly predictive of DFS ( $p= 9.10^{-5}$ ). Figure 6.4B shows the DFS curves for patients with TLG2 < 30.4 (median value) and TLG2 > 30.4 ( $p= 0.001$ ). SUVmax2 was predictive of DFS ( $p= 0.02$ ).

### Predictive value of PET combined parameters (PET1 and PET2) in uni- and multi-variate analyses

The relative and absolute changes in SUVmax, MTV and TLG between the PET/CT images acquired before (PET1) and during treatment (PET2) estimated with all SUVmax thresholds were not significantly predictive of DFS ( $p>0.05$ ).

Selecting the significant variables from univariate analysis, MTV1(55%) ( $p=0.021$ ; OR= 1.06 (95%CI: 1.01-1.11)) and TLG2(32%) ( $p=0.050$ ; OR= 1.01 (95%CI: 1.00-1.02)) were independent predictor of DFS in multivariate analysis ( $c=0.82$ ).

## 6.5 DISCUSSION

This study investigated the changes of PET parameters during radiochemotherapy (RTCT) and evaluated the predictive value of PET parameters using different segmentation thresholds for tumor segmentation. The most predictive thresholds of SUVmax for the segmentation were 55% at diagnosis and 32% at 40 Gy. The combination of MTV1(55%) and TLG2(32%) provided the highest DFS prediction in multivariate analysis. MTV1(55%) and TLG1(55%) were predictive of OS, while FIGO stage, visual metabolic response and parameters calculated from fixed SUV threshold segmentations were less/not significantly predictive of patient outcome.

### 6.5.1 FEASIBILITY AND TIMING

Contrary to previous recommendations to wait a few weeks after the end of radiotherapy to perform a PET [192], we investigated the capacity of PET during treatment to predict patient outcome. For each patient, visual metabolic analysis, measurements of SUVmax, MTV and TLG were assessed, not finally altered by radiation-induced inflammation artifacts. In our study, visual metabolic response at PET2 was a significant predictor of DFS in univariate analysis ( $p= 0.04$ ), requiring however a very strong experience of the nuclear medicine physician. The goal of our study was therefore to evaluate the benefit of a simple and reproducible method of segmentation, using both pre- and pre-treatment FDG/PET, to predict patient outcome. Indeed, we found that the second PET was also justified since TLG2(32%) combined with MTV1(55%) increased significantly the DFS predictive capability in multivariate analysis ( $c=0.82$ ) compared to the parameters

derived from PET1 only ( $c \leq 0.8$ ). Quantitative analysis appears also more predictive than metabolic visual analysis.

### 6.5.2 PREDICTIVE VALUE COMPARED WITH LITERATURE

Studies have investigated the prognostic impact of several per-treatment PET parameters, which were mainly visual metabolic response, SUVmax, metabolic tumor volume and the percentage of change in SUVmax [67,187–189]. Schwarz *et al.*, first described PET changes during RTCT for 36 patients with stage Ib1 to IIIB cervical cancer without finding however a correlation between visual response or SUVmax and survival [189]. The schedule of PET imaging was nevertheless different than ours, because PET was performed at three time intervals during RTCT (days 4-28; days 19-43 and days 33-56). Bjurberg *et al.*, evaluating early PET response after a mean irradiation dose of 23 Gy in 32 patients with stage Ib1 to IVa identified a small group of patients with excellent prognosis (patient with visual complete response) [187]. The authors suggested therefore that treatment response should be evaluated at later time points. Two recent studies confirmed that the best time for response prediction is at 4 weeks [67,188]. Kidd *et al.* [67], in a cohort of 25 patients with stage Ib1-IVa cervical cancer, found a correlation between patient outcome and SUVmax assessed at 4 weeks of RTCT and 3-months after treatment, ( $p = 0.037$ ). Moreover, Oh *et al.*, demonstrated that a decrease of SUVmax more than 60% at 40 Gy was a significant predictor factor of PFS in a population of 60 patients treated for stage Ib1 to IVa cervix cancer ( $p=0.045$ ) [188]. Our data confirmed these time point as we found that SUVmax2, MTV2 and TLG2 estimated at 40 Gy were significantly predictive of DFS. Kidd *et al.* also reported a correlation between pretreatment MTV and post-treatment PET response ( $P=0.008$ ), however not with MTV assessed during treatment [67].

Prognostic value of quantitative parameters such as metabolic tumor volume (MTV) or total lesion glycolysis (TLG) have been investigated at diagnosis of cervix cancer with controversial results [57,61,62,186]. Cervical tumor with high MTV and TLG at diagnosis showed an increase risk of recurrence and decrease OS in two studies [57,62]. Contradictory results have been however reported by Crivellaro and Akkas [63,186]. Concerning cervix cancer, no data have been yet published, to our knowledge, concerning the predictive impact of TLG estimated during RTCT. Our results had a longer follow up than previous study and shows that MTV and TLG estimated during RTCT were more predictive than the reports in the literature.

### 6.5.3 OPTIMAL CUT-OFF

Another important issue in using TEP for prediction is the identification of the optimal method to evaluate the metabolic response. Our study was the first to test a wide range of threshold values of both percentage of SUVmax, by step of 1% from 30% to 70% and fixed thresholds of SUVmax (from 1 to 20), for MTV and TLG calculation. In previous study evaluating the use of PET/CT during radiation therapy planning, a threshold of 50% of maximum SUV was found to be correlated with tumor volume measured by CT [194,195]. Still in cervix cancer, Chung *et al.*, shown that metabolic tumor volume estimated with a fixed threshold of SUVmax of 2.5 on PET performed before treatment was correlated with DFS ( $p=0.04$ ) [57]. Lin *et al.*, were the first to demonstrate the feasibility of estimating MTV using the threshold of 40% of SUVmax, before and during treatment, in a series of 32 patients with stage Ib1 to IIIB cervical cancer [180]. No correlation with patient outcome was however reported. Using a threshold of 40% of SUVmax in our study decreases the prediction capability (MTV1 ( $c=0.62$ ,  $p=0.13$ ), TLG1 ( $c=0.67$ ,  $p=0.09$ ), MTV2 ( $c=0.60$ ,  $p=0.12$ ) and TLG2 ( $c=0.68$ ,  $p=0.05$ )) compared to 55% of SUVmax threshold ( $c=0.8$ ).

### 6.5.4 LIMITS

The current study has several limitations. First, it was a single institution study needing therefore to be validated in a multi-institutional study. Indeed, PET interpretation depends on physician experience and acquisition protocol may also influence SUV, MTV and TLG values. Secondly, a larger multi-institutional study appears necessary to test the PET contribution in a multivariate analysis integrating other parameters such as from MRI.

Despite these limitations, our study suggests that the combination of TLG and MTV estimated at 55% of SUVmax at diagnosis and at 32% of SUVmax during RTCT may predict treatment outcomes in patients with cervical cancer. Such approach may finally help selecting a subset of patients at higher risk of recurrence who might benefit from “intensive” adjuvant treatment such as surgery or chemotherapy.

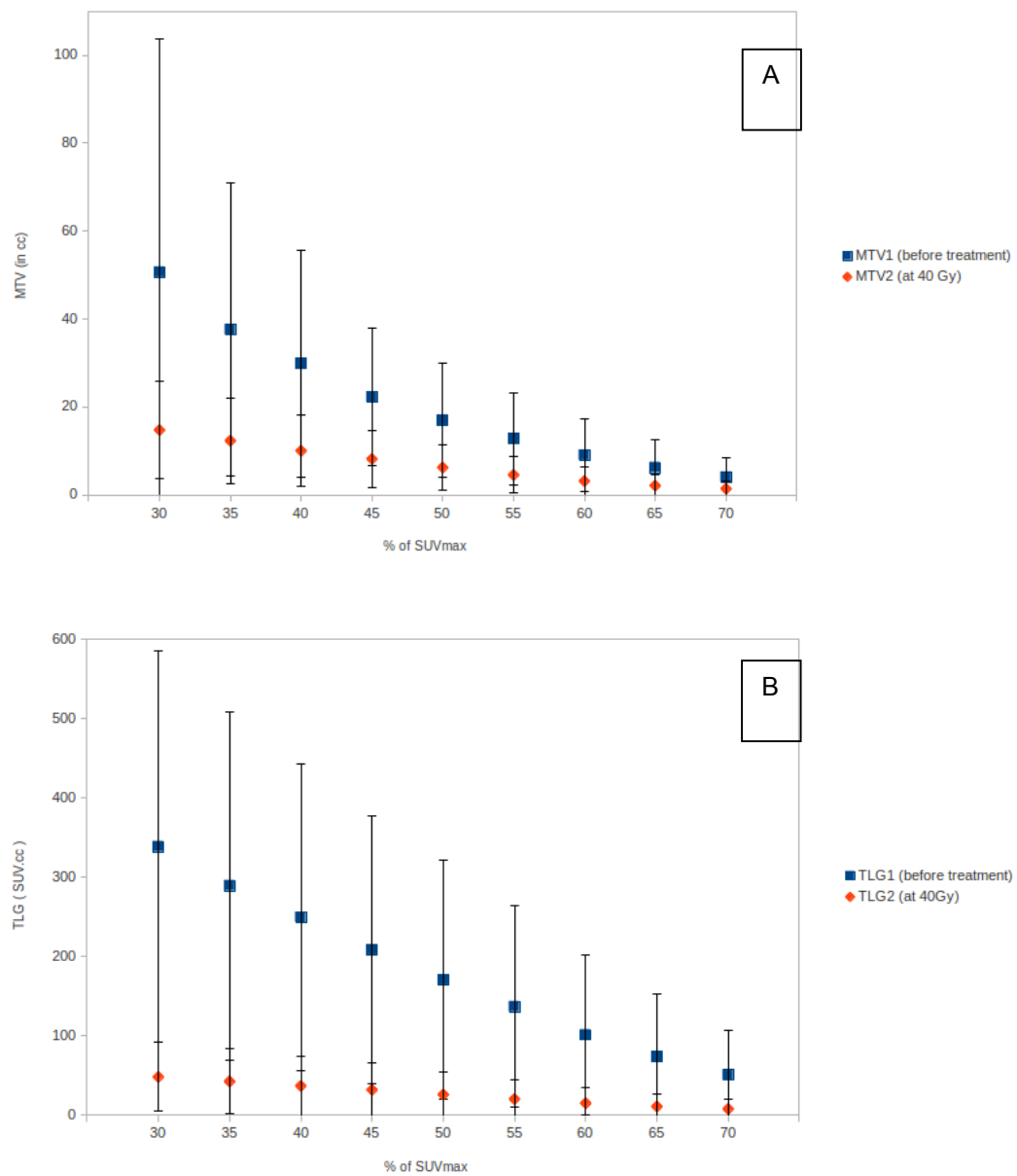


Figure 6.1 Evolution of values of MTV (A) and TLG (B) quantified before treatment (in blue) and during RTCT at 40 Gy (in red) according to the different threshold of SUVmax (from 30% to 70%).

Errors bars are standard deviations.

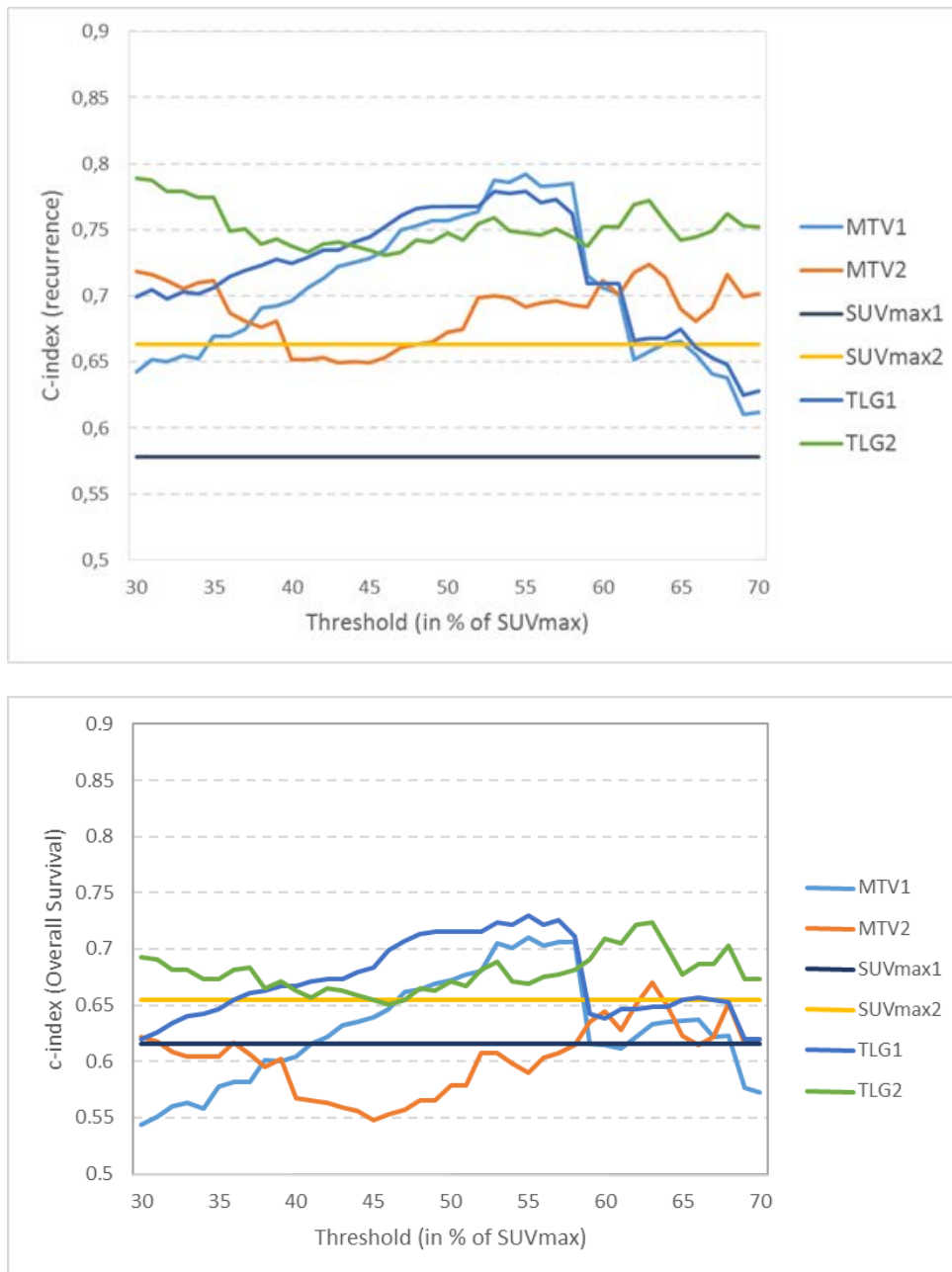


Figure 6.2 C-index values for SUVmax, MTV and TLG estimated on two PET images (PET1: before treatment and PET2: during treatment at 40 Gy), depending on different threshold of SUV max (from 30% to 70% of SUV max) to predict DFS (A) or OS (B)

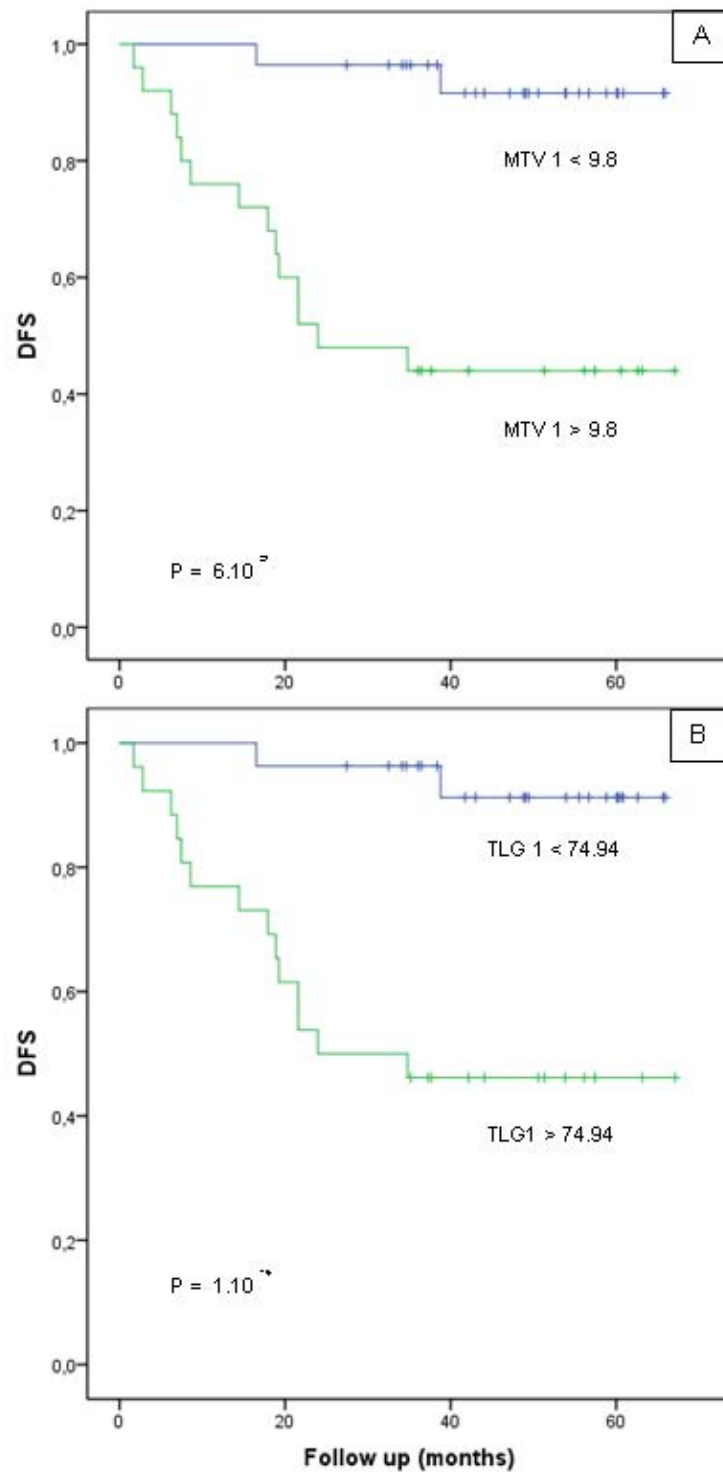


Figure 6.3 Disease Free Survival (DFS) curves among patients with cervical cancer stratified by metabolic tumor volume (MTV1) (A) and total lesion glycolysis (TLG1) (B) estimated on PET1 pre-treatment

The MTV1 and TLG1 have been calculated in a segmentation of 55% of SUVmax. The population has been divided in 2 groups according to the median values of MTV1 (=9.8cc) and TLG1 (=74.94)

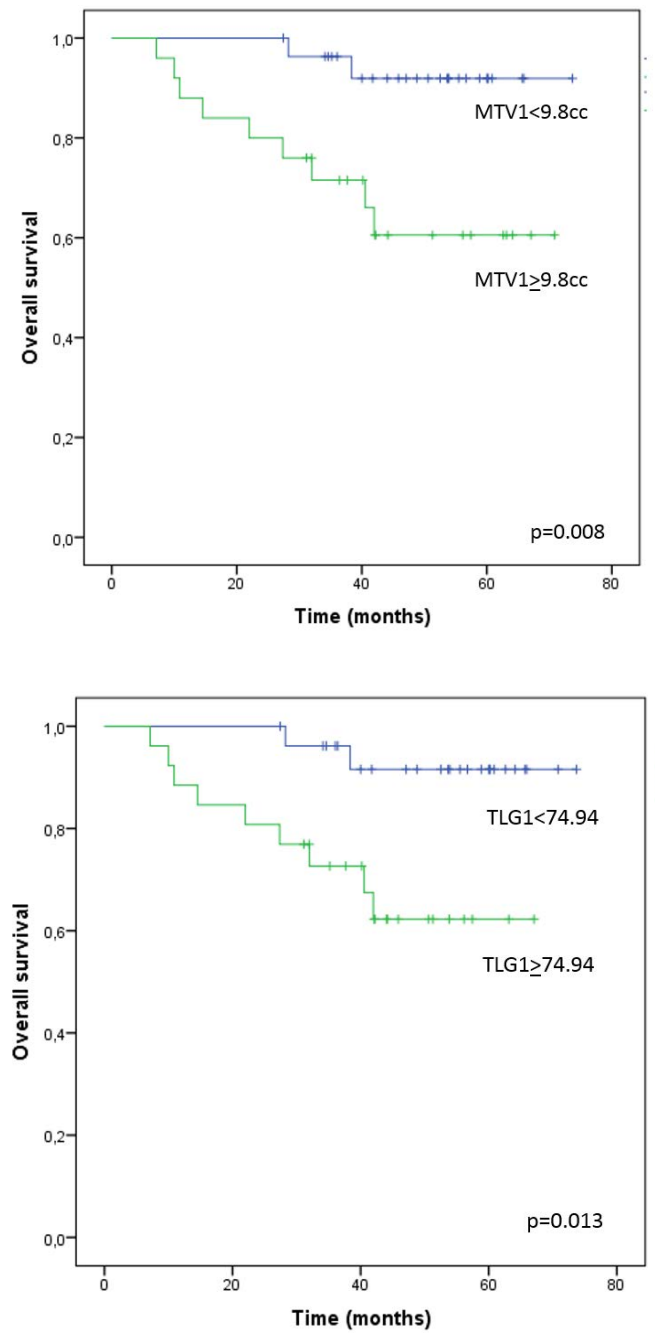


Figure 6.4 Overall survival (OS) curves among patients with cervical cancer stratified by metabolic tumor volume (MTV1) (A) and total lesion glycolysis (TLG1) (B) estimated on PET1 pre-treatment

The MTV1 and TLG1 have been calculated in a segmentation of 55% of SUVmax. The population has been divided in 2 groups according to the median values of MTV1 (=9.8cc) and TLG1 (=74.94)

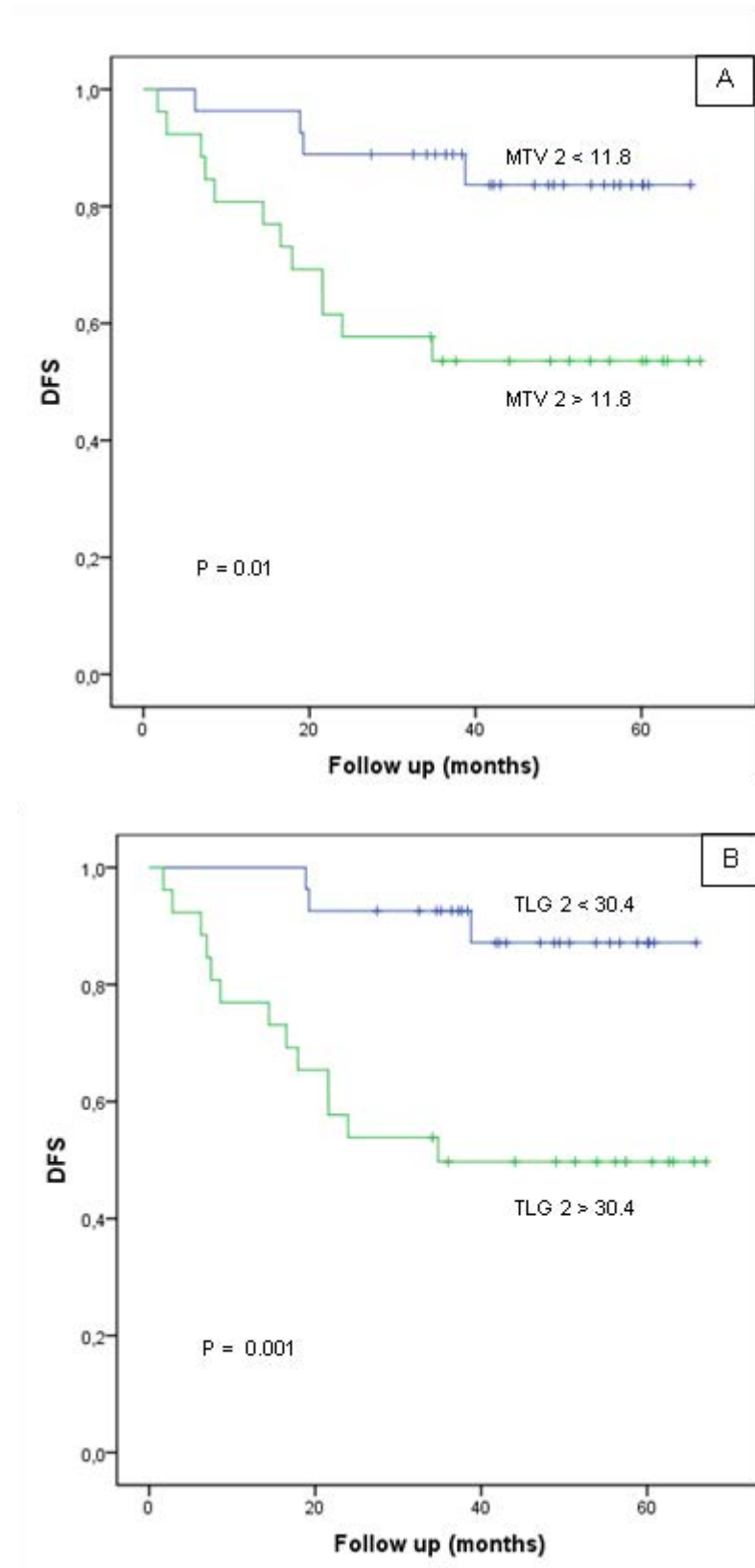


Figure 6.5 Disease Free Survival (DFS) curves among patients with cervical cancer stratified by metabolic tumor volume (MTV2) (A) and total lesion glycolysis (TLG2) (B) estimated on PET2 during treatment at 40 Gy

The MTV2 and TLG2 have been calculated in a segmentation of 32% of SUVmax. The population has been divided in 2 groups according to the median values of MTV2 (=11.8cc) and TLG2 (=30.4)



## 6.6 CONCLUSION

Dans ce chapitre, nous avons présenté une étude portant sur l'évaluation de paramètres TEP extraits avant et en cours de traitement radio-chimiothérapeutique ainsi que sur l'identification des paramètres les plus prédictifs pour prédire la récurrence post-traitement, testant différents seuillages.

Les principaux résultats de cette étude sont :

Tout d'abord, les deux images TEP permettent individuellement d'obtenir des paramètres très discriminants pour prédire la récurrence mais également la survie globale des patientes atteintes d'un cancer du col utérin localement avancé.

De plus, nous avons montré que pour les images per-thérapeutiques, les analyses quantitatives permettent d'obtenir une meilleure prédiction par rapport à l'analyse visuelle.

Dans les images pré-thérapeutiques, le paramètre le plus prédictif de la récurrence et de la survie était le MTV avec un seuillage de 55%. À partir de l'identification de ce paramètre, nous avons pu établir qu'un seuil de MTV à  $9,8 \text{ cm}^3$  permettait de discriminer significativement les patientes à risques de récurrence ( $p=0,008$ ).

Dans les images per-thérapeutiques, le TLG avec un seuillage de 32% a été identifié comme paramètre très prédictif de la récurrence et de la survie. De la même manière, les patientes à risques de récurrence étaient significativement discriminées en utilisant un seuil de TLG à 30,4 ( $p=0,001$ ). Par ailleurs, en analyse multivariée, ces deux paramètres étaient indépendants, offrant ainsi la possibilité d'être combinés pour améliorer la prédiction ( $c=0,82$ ).

Nous pouvons constater que les paramètres prédictifs identifiés en pré- et en per-thérapeutiques sont d'une part de nature différente, et d'autre part extraits avec des niveaux de seuil différents. Ainsi, contrairement à la pratique courante, qui consiste à utiliser un même paramètre tout au long du suivi des patientes, l'utilisation adaptée de différents paramètres permet une analyse plus fine des caractéristiques de la tumeur et de sa réponse au traitement. Plus précisément, en initial, une mesure du volume de la tumeur permet de distinguer les patientes à risques, indépendamment de l'intensité du signal. En revanche, à 40 Gy du traitement radio-thérapeutique, la mesure de volume seule n'est plus suffisamment pertinente et la prise en compte de l'intensité totale résiduelle du métabolisme tumoral est nécessaire pour une bonne discrimination des patientes.

D'autre part, nos résultats montrent l'importance d'adapter la méthode de segmentation au stade du traitement auquel les images ont été acquises. Plus concrètement, en initial, un seuil de 55% du SUVmax permet d'extraire un sous volume particulièrement actif et de discriminer les patientes ayant un haut risque de récurrence. Au cours du traitement, la fonte tumorale induisant une réduction de l'hyperfixation, la prise en compte d'une plus large région permet d'augmenter la sensibilité de la détection du signal résiduel. Dans notre cohorte, un seuil à 32% du SUVmax s'est révélé particulièrement pertinent.

Ces observations mettent en évidence l'importance d'un changement de paradigme dans l'analyse des images TEP pour la prédiction de récurrence. Cependant, les méthodes que nous avons proposées reposent encore beaucoup sur l'établissement d'un seuil précis, qui dépend beaucoup de la population étudiée, et dont la pertinence sur des études multicentriques doit encore être évaluée. Afin de s'affranchir de l'impact de la segmentation dans l'extraction de paramètres prédictifs, et dans l'optique de rendre l'analyse plus robuste à la population étudiée, deux stratégies peuvent être envisagées :

- modifier la méthode d'extraction de paramètres afin de réduire sa dépendance à la méthode de segmentation ;

- modifier la méthode d'analyse de manière à utiliser la totalité de l'information disponible en combinant plusieurs seuils.

Ces deux stratégies sont l'objet des deux prochains chapitres.



## 7 QUANTIFICATION PAR PONDERATION D'APPARTENANCE DES IMAGES TEP AU 18F-FDG POUR LA PREDICTION DE RECIDIVE (COMMUNICATION EMBC 2015 ET ESTRO 2016 ET ARTICLE SOUMIS AU JOURNAL JNM FEVRIER 2016)

*Ce chapitre présente une étude sur la valeur prédictive de paramètres TEP quantifiés par pondération d'appartenance au signal tumoral. Ce type de quantification, a l'avantage de prendre en compte la nature imprécise de l'imagerie TEP, en pondérant la valeur de chaque voxel par sa probabilité d'appartenance à la classe tumeur. Ce calcul est donc une généralisation du calcul de paramètres, classiquement extraits des images TEP, au travers d'une classification floue des voxels de l'image. La fonction de Zadeh a été utilisée pour pondérer la probabilité d'appartenance en fonction du seuil. Dans cette étude, nous avons définies quatre paramètres pondérés par appartenance, analogues aux paramètres standards utilisés préalablement SUVmax, SUVpeak, MTV et TLG. Pour cette étude, ces paramètres ont été extraits des images TEP pré-thérapeutique de 53 patientes traités pour un cancer de col de l'utérus localement avancé. Nous avons ensuite mesuré l'impact sur la prédiction au travers de différents seuillages et niveaux de flous. Les résultats ont montré que ses paramètres étaient discriminants pour une large gamme de niveaux de flous et a permis de démontrer que ce type de quantification permet d'atténuer la dépendance de la méthode de segmentation dans l'analyse TEP à visée prédictive.*

*Le contenu de ce chapitre a été présenté à la 37ème conférence international de « IEEE Engineering in Medicine an Biology Society » (EMBC) de 2015, sera présenté à l'ESTRO 2016 et une version complétée de validation a été soumis au journal « Journal of Nuclear Medicine ».*

---

## CONTENU

<b>7</b>	<b>QUANTIFICATION PAR PONDERATION D'APPARTENANCE DES IMAGES TEP AU 18F-FDG POUR LA PREDICTION DE RECIDIVE (<i>COMMUNICATION EMBC 2015 ET ESTRO 2016 ET ARTICLE SOUMIS AU JOURNAL JNM FEVRIER 2016</i>)</b> .....	<b>107</b>
7.1	Abstract .....	109
7.2	Introduction .....	110
7.3	Material and Methods .....	110
7.3.1	Patients and tumor .....	111
7.3.2	Treatment and follow-up.....	111
7.3.3	18F-FDG PET/CT acquisitions .....	112
7.3.3.1	18F-FDG PET image analyses.....	112
7.3.4	Statistical analysis .....	113
7.4	Results.....	114
7.4.1	Predictive value of hard-thresholding on recurrence.....	114
7.4.2	Predictive value of fuzzy-thresholding.....	114
7.5	Discussion.....	115
7.5.1	State of the Art .....	115
7.5.2	Fuzzy-thresholding and fuzzy-weighted quantification .....	115
7.5.3	Limitations .....	116
7.6	Conclusion.....	124

---

## Impact of fuzzy-weighted parameters in 18F-FDG PET imaging in the prediction of cervical cancer recurrence

Geoffrey Roman Jimenez<sup>1,2,4</sup>, Anne Devillers<sup>3</sup>, Julie Leseur<sup>3</sup>, Juan D. Ospina<sup>1,2</sup>, Henri Der Sarkissian<sup>4</sup>, Oscar Acosta<sup>1,2</sup>, Renaud De Crevoisier<sup>1,2,3</sup>.

<sup>1</sup>INSERM, U1099, Rennes, F-35000, France

<sup>1</sup>Université de Rennes 1, LTSI, Rennes, F-35000, France

<sup>3</sup>Département de Radiothérapie, Centre Eugène Marquis, Rennes, F-35000, France

<sup>4</sup>Keosys Medical Imaging, Saint-herblain, F-44815, France

### 7.1 ABSTRACT

**Objective:** Using pre-treatment 18F-FDG PET in locally advanced cervical cancer, the goal of the study was to assess the recurrence capability prediction of fuzzy-weighted standard uptake value (SUV) parameters extracted through a tumor membership probability map, compared to “standard” fixed or hard threshold SUV based parameters.

**Methods:** 53 patients treated for locally advanced cervical cancer had a 18F-FDG PET/CT before treatment. Tumor membership probability maps were extracted from PET images through a fuzzy-thresholding defined by a given threshold  $Th$  (41%, 50% and 70% of SUVmax) combined with a level of fuzziness  $\Delta Th$  (from 0 to 40% of SUVmax,  $\Delta Th=0\%$  corresponding to hard-thresholding) using a Zadeh's standard function. Four weighted parameters were calculated within these maps: the maximum uptake averaged by its 26 neighbors (SUVpeak), the average SUV inside the tumor region (SUVmean), the metabolic tumor volume (MTV) and the total lesion glycolysis (TLG). The area under the receiver operating curve (AUC) of the logistic regression model were used to quantify and compared the capability of the PET parameters to predict 3-year recurrence.

**Results:** Median follow-up was 52 months and 16 patients presented a recurrence within the 3 years after treatment. With a fixed-threshold of 41% and 50% (and not 70%), MTV and TLG were significantly predictive of recurrence with AUC values ranging from 0.72 to 0.77. With fuzzy-thresholding, at  $Th$  being equal to 41%, 50% and 70%, the most significant  $\Delta th$  for MTV and TLG were 7%, 16% and 36%, respectively. The corresponding AUC ranged from 0.73 to 0.80, the highest AUC value being with TLG at  $Th=70\%$  and  $\Delta th=36\%$  (AUC=80). SUVmax, SUVpeak and SUVmean were not significantly predictive of recurrence.

**Conclusion:** The fuzzy-weighted MTV and TLG parameters were found to be highly predictive of recurrence in cervical cancer, and less sensitive to the choice of the SUVmax thresholds, compared to the “standard” hard threshold SUV based parameters.

Fuzzy-weighted SUV parameters are highly predictive of tumor recurrence in cervical cancer.

## 7.2 INTRODUCTION

Locally advanced cervical cancer treatment is based on external beam radiation therapy with concurrent cisplatin chemotherapy followed by brachytherapy, leading to a 5-year survival rate of 70% [10][178]. Indeed, around 30% of the patients will present a recurrence within the first 2 years after treatment. The identification of new predictors of recurrence, in addition of the known FIGO stage, is crucial to target “intensified” adjuvant treatment to these high risk patients. In this context, [18F]-2-fluoro-2-deoxy-D-glucose positron emission tomography (18F-FDG PET) has been widely investigated in the literature to provide meaningful early markers of recurrence [50,51,59,61–63,68,196]. Basically, these studies evaluated the predictive capability of PET images, firstly by extracting 18F-FDG PET parameters through manual or automatic segmentation and secondly by performing statistical analyses to evaluate the predictive value of the extracted features.

One of the challenges of the quantitative 18F-FDG PET analyses is the isolation of the metabolic tumor area from the rest of the image due to the low spatial resolution, partial volume effect (PVE), low contrast and high noise characteristics of the PET images. In clinical practice, segmentation of tumor metabolism is performed either by applying a threshold between 40% and 70% of the maximum standardized uptake values (SUVmax) or a fixed SUV threshold (most frequently a value of 2.5), depending on the study [22]. If threshold methods are relatively simple to implement, they are however sensitive to noise and contrast variation [38,152]. Moreover, PET imaging suffer from partial volume effect, due to the acquisition and the reconstruction process, inducing a natural blur in the PET images [79]. Therefore, because segmentation is basically a voxel dichotomization, the use of this standard techniques can lead to lose relevant information for analyses requiring accurate characterization of metabolic tumor activity.

In the literature of pattern recognition, the fuzzy-logic have been therefore introduced to refine the segmentation by adding uncertainty in the classification process [197–199]. As opposed to classical segmentation approaches assuming that a voxel belongs or not belong to a class, the fuzzy-logic allows a gradual transition from membership to non-membership, providing a probability for each voxel to belong to the signal of interest. In other terms, instead of hardly delineate the tumor activity, fuzzy-segmentation aims at finding a map of tumor membership probability indicating where the tumor metabolism is likely located. Several fuzzy-based methodologies have been proposed to deal with the natural blur of PET signal and determinate therefore PET metabolic tumor area [86,120–122]. For the need of standard PET parameters extractions, these methodologies always conclude with a label assignation of each voxel to only one of the classes used. Therefore, despite the use of evolved methodologies including the natural PET uncertainty, the extraction of standard PET parameters induces a hardly defined tumor delineation in contradiction with the nature of the image. To our knowledge, no study has tested the use of PET parameters including a tumor membership probability of each voxel in the computation.

In case of locally advanced cervical cancer, the first goal of this study was therefore to generate a fuzzy-weighted quantification of the 18F-FDG PET tumor metabolic activity, extracted through a tumor membership probability map provided by fuzzy-thresholding. The second goal was to assess the predictive value of the extracted fuzzy-weighted parameters and to compare their prediction capability with the “standard” non-fuzzy-weighted (or hard thresholding) parameters.

## 7.3 MATERIAL AND METHODS

The workflow of the study is briefly described. Using pre-treatment FDG-PET for cervix carcinoma, fuzzy-thresholding segmentations were first generated based on the Zadeh's standard function [200] [199], defined by a SUVmax threshold ( $T_h$ ) and a level of fuzziness ( $\Delta T_h$ ). Different level of fuzziness (0 to 40%)

were applied to the three thresholds (41%, 50% and 70%) defined in the 2015 EANM procedure guidelines [22]. Four fuzzy-weighted “analogs” of clinical used PET parameters were then calculated, namely, the SUVmax averaged by its 26 neighbors (SUVpeak), the average SUV inside the tumor region (SUVmean), the metabolic tumor volume (MTV) and the total lesion glycolysis (TLG). The predictive capabilities for recurrence of the features were finally evaluated using the area under the receiver operating characteristic curve (AUC) and the p-value of the logistic regression model.

### 7.3.1 PATIENTS AND TUMOR

Between January 2007 and December 2011, 53 consecutive patients treated for locally advanced cervical cancer in a curative intent were included in this prospective trial and analysed. Eligibility criteria were: histological confirmed invasive cervix carcinoma, stage IB2-IVA according to the International Federation of Gynecology and Obstetrics (FIGO), complete staging comprising clinical staging, MRI and 18F-FDG PET scan. A para aortic lymph node dissection was performed in all the patients, except in case of hyperfixation in this area at the PET imaging. A pelvic lymph node dissection could also be performed.

The median age was 52 years (range, 26 - 81 years). The stage distribution according to the FIGO classification was as follows: stage IB2 in 21 patients (40%), stage IIA1 in 9 patients (17%), stage IIA2 in 3 patients (6%), stage IIB in 14 patients (26%), stage IIIA in 2 patients (4%), stage IIIB in 1 patient (2%) and stage IVA in 3 patients (5%). Histology was squamous cell carcinoma for 40 patients and adenocarcinoma for 18 patients. Thirteen patients presented an involvement of pelvic lymph nodes. Eight patients presented pathological hyperfixation of para-aortic lymph nodes on PET and 5 an involvement of para-aortic lymph nodes on lymph nodes dissection.

### 7.3.2 TREATMENT AND FOLLOW-UP

All patients were treated with a combination of three-dimensional (3D) conformal external beam radiation therapy (EBRT) and pulse-dose-rate brachytherapy. EBRT delivered 45 Gy in the pelvis at 1.8 Gy per fraction, using an IMRT technique with 5 to 8 - 18 MV photon beams. A simultaneous integrated boost up to 55 Gy (at 2.2 Gy/fraction) was delivered in the positive pelvic nodes when shown in imaging. The target volume was extended to the para-aortic lymph nodes in case of lymph node involvement in the para-aortic lymphadenectomy (5 patients) or at the PET imaging (8 patients). Pulse dose rate (PDR) brachytherapy was performed two weeks after the end of EBRT using a standard Nucletron® ring applicator delivering a dose of 15 Gy into the intermediate risk CTV according to the GEC ESTRO recommendations [190,191].

In case of parametrial involvement a lateropelvic external radiation boost was performed after brachytherapy to deliver a total dose of 60 Gy (4 patients). Fifty-one patients received concurrent weekly cisplatin chemotherapy (40 mg/m<sup>2</sup> weekly for 4-5 cycles). Two patients did not receive concurrent chemotherapy because of co-morbidities or advanced age (≥80 years). The total treatment duration did not extend 7 weeks.

For the 18 patients with stage IB2 and IIA-B having residual disease at clinical examination, or biopsy or PET or MRI 4 to 6 weeks after brachytherapy, a total hysterectomy was performed (8 patients with stage IB2, 3 patients with stage IIA1 and 7 patients with stage IIB). The follow-up consisted in clinical examination, MRI and FDG PET 4 to 6 weeks after the end of brachytherapy, at 3 months and then every 6 months.



### 7.3.3 18F-FDG PET/CT ACQUISITIONS

The 18F-FDG PET/CT were realized before the beginning of radiation therapy and in the follow-up. 18F-FDG PET/CT scans were acquired using a PET/CT device (Discovery ST; GE Healthcare), which consisted of a PET scanner (Advance NXi; GE Healthcare) and a 2-slice CT scanner (Light-Speed Plus; GE Healthcare). Approximately 60 minutes after the intravenous injection of 4-5 MBq of 18F-FDG (Flucis), acquisitions were performed until the base of the skull to the upper part of the thighs. Scanning parameters were 120 kV, 80 mA, 0.8-s tube rotation, and 3.75 mm slices spacing. PET images were displayed with an ordered-subset expectation maximization iterative reconstruction algorithm (OSEM, 2 iterations, 28 subsets) and a reconstruction iterative fully 3D algorithm (Discovery ST). Patients fasted for at least 4 hours, a premedication with diazepam was administered systematically, as well as intravenous injection of furosemide (0.02 ml / kg) 15 min prior acquisitions.

#### 7.3.3.1 18F-FDG PET IMAGE ANALYSES

##### Image preparation

The full body PET/CT images were cropped keeping the pelvic region according to coxal bone structures on CT images. The bound of the box was defined with respect to well defined landmarks which may be easily reproducible. The landmarks were the bottom edge of lumbar 5, the bottom edge of ischium, the right edge of iliac bone, the left edge of iliac bone, the anterior edge of ischium and the posterior edge du sacrum. A mask of a manual bladder CT delineation was finally subtracted to the resulting area to ensure the non-inclusion of bladder 18F-FDG uptake.

##### 18F-FDG PET Fuzzy-thresholding

Let denote  $X = \{x_1, x_2, \dots, x_n\}$  the set of  $n$  voxels of the PET image. The problem of fuzzy-thresholding is to assign to each voxel  $x_i$  of  $X$  a tumor membership probability according to its SUV value. In this study, we performed fuzzy-thresholding methods using a Zadeh's standard function [200] as in Murthy *et al.* [199] defined by a threshold and a level of fuzziness. The function  $f(x)$  is defined as:

$$f(x) = \begin{cases} 0 & \text{if } I(x) < (Th - \Delta Th) \\ 2 \left( \frac{I(x) - (Th - \Delta Th)}{\Delta Th} \right)^2 & \text{if } \left( Th - \frac{\Delta Th}{2} \right) \leq I(x) < Th \\ 1 - 2 \left( \frac{I(x) - Th}{\Delta Th} \right)^2 & \text{if } (Th - \Delta Th) \leq I(x) < \left( Th - \frac{\Delta Th}{2} \right) \\ 1 & \text{if } I(x) \geq Th \end{cases}$$

where  $I(x)$  is the SUV of the voxel  $x$ ,  $Th$  is the threshold used and  $\Delta Th$  the level of fuzziness.  $Th$  and  $\Delta Th$  were expressed in % of the SUVmax. Figure 7.1 shows a graphical representation of the function Zadeh's standard depending on  $Th$  and  $\Delta Th$ .  $f(x)$  is thus defined as a combination of a hard-thresholding at  $Th$  and a gradual transition between  $Th$  and  $\Delta Th$ . In order to keep the probability map compact, we applied a region-growing at  $Th - \Delta Th$  and kept only the voxels inside the resulting area.

To evaluate the contribution and the relevance of the fuzziness, several fuzzy-thresholding were tested by applying  $Th=41\%$ ,  $50\%$  and  $70\%$  and  $\Delta Th$  equal from  $0\%$  to  $40\%$ .

Note that if  $\Delta\text{th}=0\%$ ,  $f(x)$  is binary and the tumor membership probability corresponds to a standard region-growing.

For example, with a segmentation at  $\text{Th}=41\%$ , a fuzzy-level  $\Delta\text{Th}=0\%$  will correspond to a clinically used 41% region-growing. With the same threshold and a fuzzy-level  $\Delta\text{Th}=20\%$ , the resulting probability map will include voxels from 21% to 41% of the SUVmax.

### 18F-FDG PET fuzzy-weighted quantification

Usually, 18F-FDG PET parameters are computed using only voxels labeled as belonging to a class “tumor”, other voxels being ignored. In the fuzzy-logic, the voxel labelization is generalized to a tumor membership probability assignation, being binary in hard-segmentation and gradual in fuzzy-segmentation. In this context, we can generalize the computation of a parameter by weighting each voxel in the PET image by its tumor membership probability (instead of counting only voxels labeled as belonging to the class “tumor”).

The fuzzy-weighted analogs of the SUVpeak, SUVmean, MTV and TLG were defined and computed as:

$$SUV_{peak} = \frac{1}{\sum_i^{N_{max}} f(x_i)} \cdot \sum_i^{N_{max}} f(x_i) \cdot I(x_i)$$

$$SUV_{mean} = \frac{1}{\sum_i^N f(x_i)} \cdot \sum_i^N f(x_i) \cdot I(x_i)$$

$$MTV = Vol_{voxel} \cdot \sum_i^N f(x_i)$$

$$TLG = Vol_{voxel} \cdot \sum_i^N f(x_i) \cdot I(x_i)$$

where  $I(x_i)$  and  $f(x_i)$  are respectively the SUV values and the tumor membership probability of voxel  $x_i$ .  $N$  is the Voxels of the PET image,  $N_{max}$  corresponds to the area including the SUVmax and its 26 neighbors and  $Vol_{voxel}$  is the volume of a voxel.

Note that we considered here the tumor membership probability as a percentage of tumor metabolism inside a voxel. For instance, a couple of voxels  $x_i$  and  $x_j$  with  $I(x_i)=10$ ,  $I(x_j)=5$  and  $f(x_i)=0.9$ ,  $f(x_j)=0.5$ , gave a total SUV of 11.5 in 1.4 voxels.

Note that if  $f(x_i)$  is binary ( $\Delta\text{Th}=0\%$ ), the parameters computed are equivalent to the standard parameters computed with fixed threshold.

### 7.3.4 STATISTICAL ANALYSIS

The overall three-year overall recurrence rate was the main end point of the study. The recurrence events could be local and/or nodal and/or metastatic, detected by clinical examination and/or MRI and/or FDG PET. The prognostic significance of the extracted PET parameters was analyzed using the area under the curve (AUC) and the p-value calculated from the logistic regression model. MTV, TLG and SUVmax (both fuzzy and non-fuzzy-weighted parameters.) were analyzed as continuous variables in univariate and multivariate analyses.

Disease free survival (DFS) and overall survival were also calculated. The duration of the DFS was defined as the time measured from the beginning of radiation therapy to the date of the recurrence event when detected, or to the date of the last follow-up visit. The actuarial DFS and OS rates and survival curves were estimated using the Kaplan–Meier method, and the difference between survival curves was assessed by the log-rank test. The prognostic significance on survivals of the extracted PET parameters were tested by the Cox proportional hazards regression model.

## 7.4 RESULTS

The median follow-up was of 52 months [range: 27-74 months] for the 42 alive patients at the last visit, 36 of them being without evidence of disease. Sixteen patients had recurrence within the three years following treatment, and 10 of them died from cervical cancer. One patient died from cardiovascular disease (in complete response). The patterns of recurrences were: local recurrence in 10 patients (3 patients with local recurrences only and 7 patients had local recurrences with metastases), 3 patients had lymph nodes recurrences only (2 patients in the pelvis and para-aortic area and 1 patient in the para aortic area only) and 3 patients had metastatic disease without local recurrence. Five-year disease free survival and overall survival rates were 69% (CI95%: 56-82%) and 77% (CI95%: 65-89%), respectively. The FIGO stage was not statistically predictive of DFS ( $p=0.07$ ).

The predictive capabilities of the extracted parameters (SUVpeak, SUVmean, MTV and TLG) on overall recurrence, depending on the three threshold Th (41%, 50% and 70%) and the fuzzy-level  $\Delta$ th (0 to 40%), are represented on Figure 7.2 and Table 7.1. The SUVmax, independent from the performed segmentation, was not predictive (AUC=0.58,  $p=0.5$ ).

### 7.4.1 PREDICTIVE VALUE OF HARD-THRESHOLDING ON RECURRENCE

Predictive values of MTV and TLG of hard-thresholding are presented on Figure 7.2 (column  $\Delta$ Th=0%) and in Table 7.1. MTV and TLG were predictive with a fixed-threshold of 41% and 50% (AUC values ranging from 0.72 to 0.77) but not 70%. SUVpeak and SUVmean were not predictive.

### 7.4.2 PREDICTIVE VALUE OF FUZZY-THRESHOLDING

Predictive values of MTV and TLG of fuzzy-thresholding are presented on Figure 7.2 (columns  $\Delta$ Th=2% to 40%) and in Table 7.1.

At Th=41%, the MTV was significant with  $\Delta$ th = [1% - 12%] and TLG with  $\Delta$ th = [0% - 22%], the most significant  $\Delta$ th being 7%. At Th=50%, the MTV was significant with  $\Delta$ th = [1% - 26%] and TLG with  $\Delta$ th = [0% - 32%], the most significant  $\Delta$ th being 16%. At Th=70%, MTV was significant with  $\Delta$ th = [6% - 40%] and TLG with  $\Delta$ th = [8% - 40%], the most significant  $\Delta$ th being 36%. With a fuzzy-thresholding, the AUC of MTV and TLG ranged from 0.73 to 0.80. The highest AUC value was observed with TLG at Th=70% and  $\Delta$ th=36% (AUC=80). Figure 7.3 and Figure 7.4 present the DFS curves for MTV and TLG at Th=70% and  $\Delta$ th=36% by dividing the population into two groups according to the median values of MTV and TLG ( $p<0.01$ ).

Fuzzy-weighted SUVpeak and SUVmean were not predictive for none of Th and  $\Delta$ th tested. Multivariate analysis didn't find any significant combined parameters.

## 7.5 DISCUSSION

The objective of this study was to propose the use fuzzy-logic in the segmentation and the quantification of tumor metabolism in 18F-FDG PET imaging to predict recurrence in locally advanced cervical cancer.

### 7.5.1 STATE OF THE ART

During the last decade, several studies evaluating the predictive value of 18F-FDG PET in cervical cancer extracted parameters through manual or automatic segmentation (see Table 7.2). Two kind of FDG parameters can be distinguished. Those such as SUVmax and SUVpeak correspond to a direct measurement of common direct measurement of SUV and other parameters such as MTV, TLG and SUVmean requires a first step of tumor segmentation. The predictive value of SUVmax is controversial. Indeed, if some studies have demonstrated SUVmax as a predictor of recurrence [50,66] or overall survival (OS) [67], other studies, including our, didn't find any significant predictive value of SUVmax [61–63]. The explanation can be that SUVmax, being based on the measure of only one voxel in the whole tumor metabolism, is highly sensitive to the image noise [21] depending therefore on the analysis protocols [17].

Besides, studies have suggested that MTV or TLG may provide better prognostic information. In 2002, Miller *et al.* highlighted the interest of using MTV as a prognostic factor predicting DFS and OS in cervix carcinoma [51]. In 2011, Yoo *et al.* [61] and [58] shown that TLG was also a significant prognostic parameter for DFS. In 2012, Kidd *et al.* published nomograms using both SUVmax and MTV to predict DFS, disease specific survival (DSS) and OS [59]. More recently, Renier *et al.* in 2014 [68] [68] tested the predictive value of the SUVmax, SUVpeak, SUVmean, MTV and TLG through three segmentation methods: an absolute SUV threshold of 2.5, a threshold at 40% of the SUVmax and the fuzzy locally adaptive Bayesian [86]. MTV (2.5, 40%, and FLAB) and TLG (40%) were predictive of RFS and OS. With the used of parameters including the whole tumor signal, the issue of the isolation of the metabolic tumor area from the rest of the image is therefore raised. In the studies resumed in Table 7.2, the most common delineation method was to apply a region-growing at 40-42% of the SUVmax.

Biehl *et al.* in 2006, demonstrated however, that the use of single threshold do not provide an accurate delineation for tumor of different size [152]. Besides, Hatt *et al.* in 2010 have shown a lake of reproducibility of the delineation commonly used methods [124]. Moreover, we demonstrated previously that the choice of the used threshold has an important impact on the prognostic value of the extracted parameters [201,202].

### 7.5.2 FUZZY-THRESHOLDING AND FUZZY-WEIGHTED QUANTIFICATION

18F-FDG PET parameters are conventionally computed among voxels selected during the segmentation step. Despite major advances and achievements in PET segmentation during the last decades, the algorithms proposed in the literature are designed to dichotomize PET voxels into two class hardly defined, "tumor" and "non-tumor". Nevertheless, PET imaging are known to suffer from partial volume effect, due to the acquisition and the reconstruction process, inducing a natural blur in PET images [198]. Hence, any attempts to find hardly-defined frontier between PET structures having uncertain boundaries will be arbitrary, leading to a bias in qualitative and quantitative measurement [79]. To the extent that frontiers between metabolic structures in PET imaging are uncertain, we propose to integrate the fuzzy-logic in the segmentation and the characterisation of the tumor in order to take into account the gradual transition of the tumor area and consider consequently the totality of the tumor metabolism information.

In this study, we assigned a tumor membership probability to each voxel of the PET image by using fuzzy-based segmentations, and generalized the computation of PET parameters to take into account each voxel according to its membership probability. First, we performed fuzzy-thresholding segmentations using three thresholds ( $Th$ ) commonly used in clinical practice (namely 41%, 50% and 70% of the SUVmax) and different level of fuzziness ( $\Delta Th$  from 0% to 40% of the SUVmax). Resulting tumor membership probability map were kept compact by using a region-growing at  $Th-\Delta Th$ . Fuzzy-thresholding is a generalization of hard-thresholding and was chosen to be compared to commonly used thresholds. Second, the calculation of PET parameters was generalized by injecting the tumor membership probability in the computation process. The tumor membership probability was therefore considered as a percentage of tumor metabolism inside a voxel. Fuzzy-weighted analogs of SUVpeak, SUVmean, MTV and TLG were then extracted. TLG and MTV were finally found as the most interesting parameters among the features tested, SUVpeak and SUVmean being not discriminative in our study. As found in the literature, standard parameters ( $\Delta th=0\%$ ) were discriminative of  $Th=41\%$  and  $Th=50\%$ , but not with  $Th=70\%$ , TLG being the highest predictive parameter at  $Th=50\%$  (AUC=0.77,  $p=0.006$ ). The contribution of the study is that, for each  $Th$  tested threshold, predictive parameters were found in a large range of  $\Delta th$  fuzzy-level. Indeed, fuzzy-parameters were either equal or better than standard quantification when  $Th-\Delta th=34\%$ . In particular, inclusion of voxels  $>34\%$ , coupled with a gradual transition between tumor membership and non-membership, produced good prognostic parameters for all the thresholds tested, indicating that the prognostic values of fuzzy-weighted parameters are robust to the choice of threshold.

The TLG was the highest prognostic parameter with  $Th=70\%$  and  $\Delta th=36\%$  (AUC=0.80,  $p=0.005$ ). Indeed, a threshold at 70% aims at keeping the higher uptakes, believed to reflect the tumor aggressiveness [203]. On the other hand, threshold around 40% is the most common used threshold to estimate the tumor volume in cervical cancer [51,204]. Thus, this parameter can be interpreted as the quantification of the TLG, computed with all voxels supposedly belonging to the tumor while focusing on the voxels reflecting the tumor aggressiveness.

### 7.5.3 LIMITATIONS

This study has several limitations. The number of patients was relatively small so it is not possible to recommend a cut-off values of the PET parameters to identify different patient outcome. Another limitation is the fact that the study was focused only on the recurrence analysis. Because of the limited number of observed deaths, we were therefore unable to perform statistical analyses on OS maintaining the stability of the regression model [205]. We performed fuzzy-thresholding to evaluate the prognostic value of fuzzy-weighted quantifications, being comparable with commonly used parameters. Although, as the objective was to investigate the use of fuzzy-logic in outcome prediction, we did not assess the performance of the fuzzy-segmentation. For that reason, further investigations should be done to evaluate the relevance of more evolved fuzzy-segmentation methods, such as the fuzzy c-mean algorithm [121,122] or the fuzzy locally adaptive Bayesian [86].

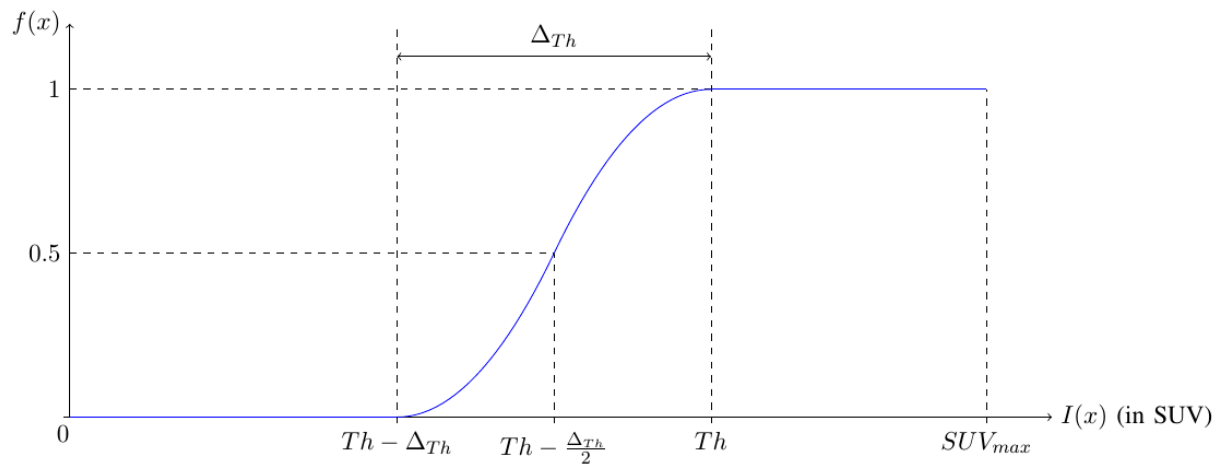
To conclude, we proposed a computation of  $^{18}F$ -FDG PET SUV parameters based on a tumor membership probability provided by fuzzy-segmentation.

The fuzzy-weighted MTV and TLG parameters were found to be highly predictive of recurrence in cervical cancer, and less sensitive to the choice of the SUVmax threshold, compared to the “standard” hard threshold SUV based parameters.



Figure 7.1 Graphical representation of the Zadeh's standard function

$f(x)$  was used to compute tumor membership probability of a voxel depending on its SUV value. The threshold  $Th$  values were 41%, 50% and 70% of  $SUV_{max}$ . The level of fuzziness  $\Delta Th$  ranged from 0% to 40% of  $SUV_{max}$ , by step of 1%.



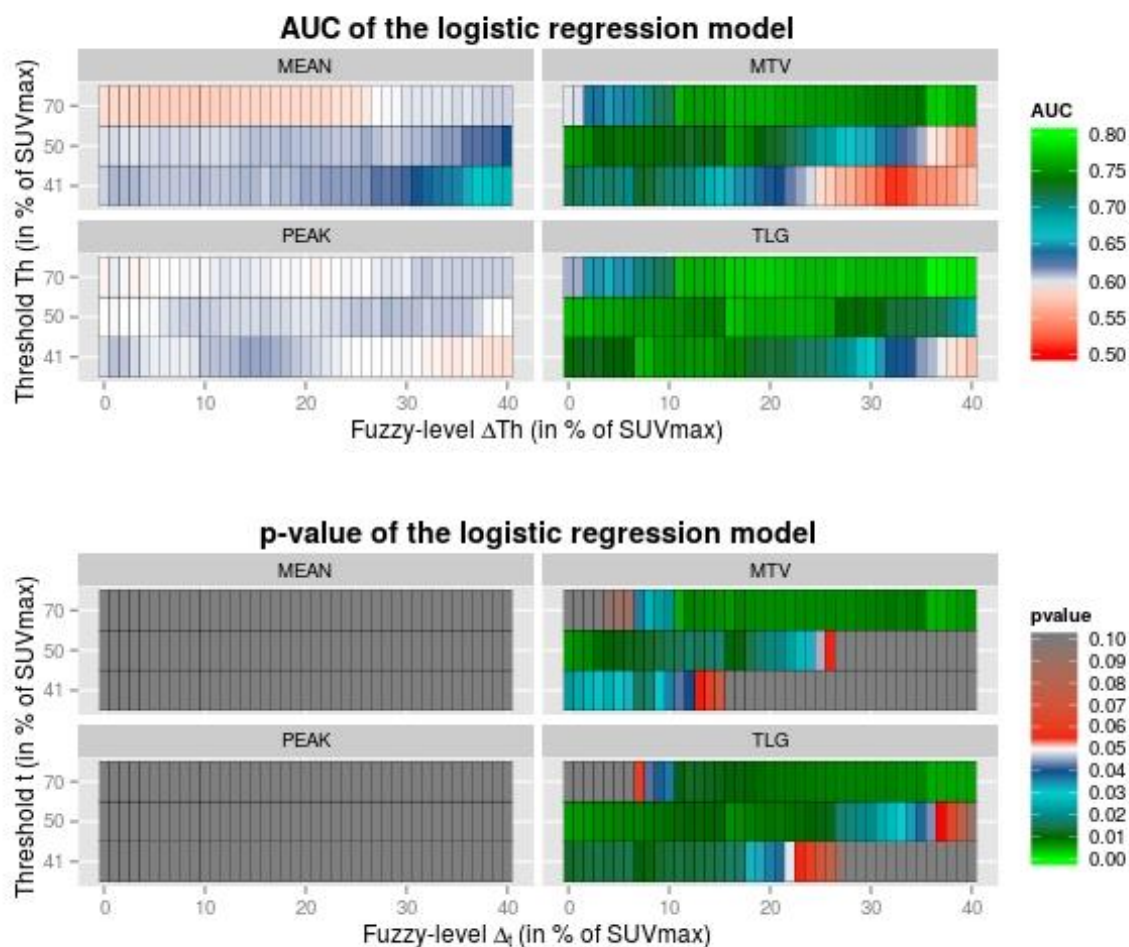


Figure 7.2 Predictive capability of the extracted PET parameters regarding 3-year tumor recurrence using logistic regression

MEAN = SUVmean; PEAK = SUVpeak, MTV = metabolic tumor volume; TLG = total lesion glycolysis. The thresholds of SUVmax (Th) were 41%, 50% and 70% (in ordinate). The level of fuzziness ( $\Delta th$ ) ranged from 0% to 40% (in abscissa).



**Table 7.1 Highest predictive fuzzy-weighted parameters ( $\Delta th > 0\%$ ) compared with parameters computed with fixed thresholds ( $\Delta th = 0\%$ )**

Threshold (Th)	Fuzziness ( $\Delta th$ )	MTV				TLG			
		Parameters values (in cc)	AUC	p-value	OR (95% CI)	Parameters values	AUC	p-value	OR (95% CI)
41	0	26.5 ± 17.8 (3.5 - 77.1)	0.72	0.022	1.016 (1.001 - 1.045)	239.1 ± 190.3 (17.6 - 818.9)	0.74	0.012	1.004 (1.001 - 1.008)
	7	32.0 ± 20.2 (4.5 - 92.3)	0.73	0.016	1.011 (0.992 - 1.036)	239.1 ± 190.3 (17.6 - 921.9)	0.77	0.01	1.004 (1.001 - 1.008)
50	0	17.1 ± 13.2 (2.2 - 49.3)	0.76	0.006	1.070 (1.021 - 1.129)	172.9 ± 149.7 (12.8 - 589.6)	0.77	0.006	1.006 (1.002 - 1.011)
	16	26.49 ± 17.24 (3.5 - 74.6)	0.76	0.01	1.019 (0.994 - 1.050)	237.1 ± 182.4 (17.5 - 786.9)	0.78	0.007	1.005 (1.002 - 1.009)
70	0	4.2 ± 4.4 (0.3 - 18.9)	0.60	0.22	1.082 (0.949 - 1,239)	53.6 ± 60.1 (3.1 - 234.6)	0.61	0.209	1.006 (0.997 - 1.016)
	36	17.2 ± 11.9 (2.2 - 44.6)	0.78	0.006	1.082 (1.000 - 1.092)	169.5 ± 135.4 (12.3 - 550.4)	0.80	0.005	1.007 (1.002 - 1.013)

Logistic regression is performed regarding the 3-year tumor recurrence. AUC = area under the receiver operating curve; OR = odd ratio; CI = confidence interval.  
Parameters values are mean ± SD (range).

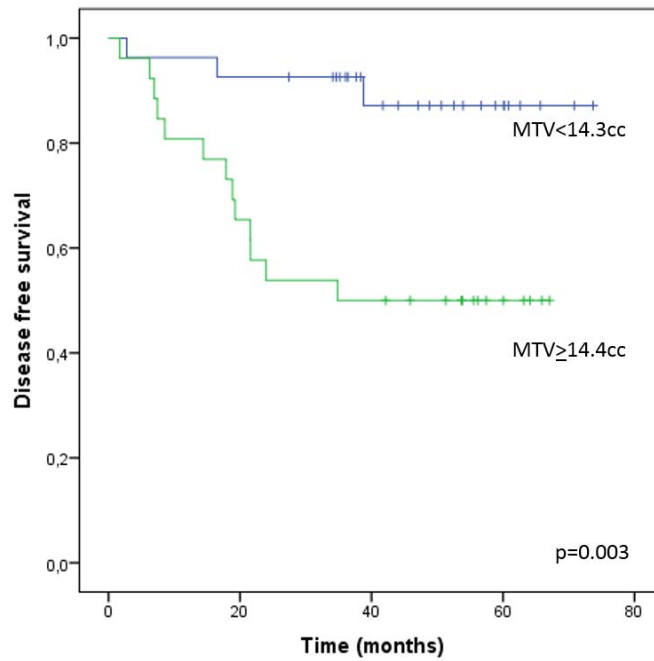


Figure 7.3 Disease Free Survival (DFS) curves among patients with cervical cancer stratified by metabolic tumor volume (MTV) with a threshold of 70% and a fuzziness of 36% of SUVmax

The population has been divided in two groups according to the median values of MTV (=14.4 cc).

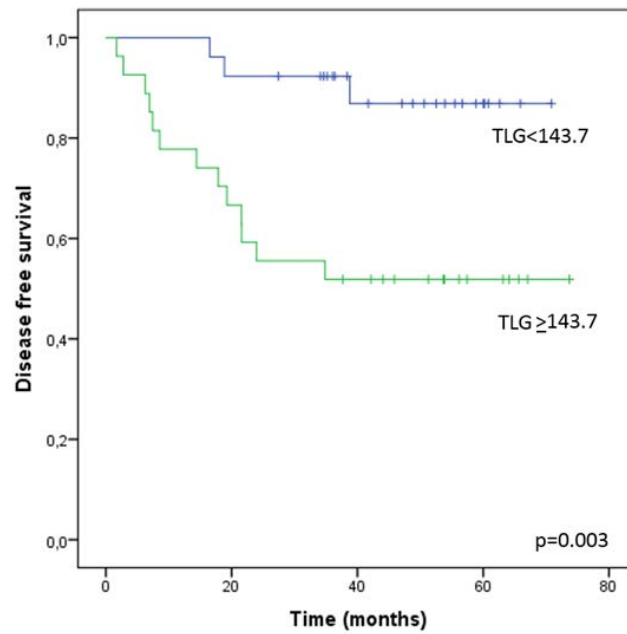


Figure 7.4 Disease Free Survival (DFS) curves among patients with cervical cancer stratified by total lesion glycolysis (TLG) with a threshold of 70% and a fuzziness of 36% of SUVmax

The population has been divided in two groups according to the median values of TLG (=143.7).

Table 7.2 Published results regarding the predictive value of 18F-FDG PET parameters in cervical cancer

Study	Number of patients	FIGO stage	Tumor delineation method	Extracted Parameters	Endpoints	Predictive parameters
Miller <i>et al.</i> [51]	51	IB1-IA1	Threshold at 40% SUVmax	MTV	PFS, OS	MTV
Miller <i>et al.</i> [55]	47	IB1-IA1	visual analyse	Visual positive uptake	PFS, OS	Visual positive uptake
Ryu <i>et al.</i> [64]	249	Ib and Ila	visual analyse	Visual positive uptake	DFS	Visual positive uptake
grigsby <i>et al.</i> [65]	152	IB-IV	visual analyse	abnormal FDG uptake (in the irradiated region)	DFS	abnormal FDG uptake
Kidd <i>et al.</i> [50]	287	IVB-IA2	Threshold at 40% SUVmax	SUVmax, MTV, Lymph Nodes Metastasis	OS	SUVmax
Chung <i>et al.</i> [66]	75	IB to IIA	manually segmented	SUVmax	DFS	SUVmax
Chung <i>et al.</i> [57]	63	IB to IIA	Threshold at 2.5	SUVmax, MTV	DFS	SUVmax, MTV
yoo <i>et al.</i> [61]	73	I-VI	SUV of mediastinal	SUVmax, SUVmean, MTV, TLG	DFS	MTV, TLG
kim <i>et al.</i> [58]	45	IA–IIB	Threshold at 40% SUVmax	MTV	DFS	MTV
Kidd <i>et al.</i> [59]	234	Ib1–IVa	Threshold at 40% SUVmax	SUVmax, MTV, Highest lymph nodes	DFS and OS	SUVmax, MTV, Highest lymph nodes
Crivellaro <i>et al.</i> [63]	89	IB1 and IIA	adaptive threshold, PETVCAR (GE Healthcare)	SUVmax, SUVmean, MTV, TLG	Lymph nodes metastases, relapse after primary treatment	MTV, TLG
Kidd <i>et al.</i> [67]	25	Ib1-IVa	Threshold at 40% SUVmax	SUVmax, FDGhetero, MTV	3-month DFS	SUVmax, FDGhetero, MTV
Micco <i>et al.</i> [62]	49	IB-IVB	Threshold at 42% SUVmax	SUVmean, MTV, TLG	DFS, OS	MTV, TLG
Renier <i>et al.</i> [68]	41	≥1B	Threshold at 2.5, Threshold at 40% SUVmax and FLAB	SUVmax, SUVpeak, MTV, TLG and texture features	DFS, OS	MTV, TLG (with threshold at 40%), entropy, intensity, variability
Chung <i>et al.</i> [69]	48	IB-IIA		intratumoural heterogeneity (IH), SUVmax, MTV, TLG	DFS	intratumoural heterogeneity (IH), MTV, TLG

OS = overall survival; DFS = disease-free survival; PFS = progression-free survival; TLG = total lesion glycolysis; MTV = metabolic tumor volume; FLAB = fuzzy locally adaptive Bayesian proposed in [86]

## 7.6 CONCLUSION

Dans ce chapitre, nous avons présenté une méthode originale pour quantifier le métabolisme tumoral en pondérant chaque voxel de l'image par sa probabilité d'appartenance au signal TEP de la tumeur. Ainsi, tous les voxels contribuent à la classification produisant un résultat moins dépendant de la méthode de segmentation utilisée.

Conventionnellement, les paramètres TEP sont calculés en utilisant les voxels appartenant à la région tumorale définie lors de l'étape de segmentation. Malgré les avancées majeures dans le domaine de la segmentation TEP depuis ces dernières années, les algorithmes proposés dans la littérature ont été développés dans le but de séparer les voxels de l'image en deux classes distinctes : « tumeur » et « non tumeur ». Cependant, les images TEP souffrant naturellement des limitations intrinsèques à la nature de l'acquisition et de la reconstruction tomographique, induisant une résolution spatiale limitée, ne permettent pas une délimitation précise des structures et par conséquent une caractérisation précise du métabolisme tumoral. Ainsi, la définition de frontières distinctes entre les structures observées par imagerie TEP sera nécessairement biaisée.

Afin de rendre la quantification du signal tumoral moins dépendante de la segmentation, nous avons proposé d'utiliser la logique floue au sein même du processus de quantification, pour prendre en compte les transitions lisses de la région tumorale. De cette manière, nous considérons la totalité de l'information incluse au sein de la délimitation et ainsi que de son voisinage.

Une probabilité d'appartenance à la tumeur est attribuée à chaque voxel de l'image TEP en utilisant différents seuillages flous, et en généralisant le calcul des paramètres TEP proportionnellement à la contribution de chacun des voxels au signal tumoral.

Des seuils intégrant différents niveaux de flou (tous deux calculés en pourcentage du SUVmax) ont été générés. En particulier, la probabilité d'appartenance au signal était interprétée comme la fraction de métabolisme tumorale au sein d'un voxel.

Nous avons ensuite comparé la pertinence des paramètres extraits (SUVpeak, SUVmean, MTV et TLG) avec les paramètres TEP standards par régression logistique en calculant l'aire sous la courbe ROC et la valeur-p associée au modèle. Les résultats montrent le caractère prédictif des paramètres MTV et TLG pour un large intervalle de niveaux de flou, en accordant plus d'importance aux hautes intensités connues pour décrire l'agressivité tumorale, tout en intégrant la totalité du signal métabolique actif. Cette étude démontre donc que l'utilisation de la quantification par pondération d'appartenance permet de réduire l'impact de la méthode de segmentation dans l'analyse TEP à visée prédictive.

De plus, les résultats ont montré que l'utilisation de la logique floue permettait d'extraire des paramètres statistiquement plus prédictifs que ceux extraits de manière conventionnelle. Les deux paramètres les plus prédictifs étaient le MTV et le TLG extraits par un seuillage à 70% du SUVmax et un niveau de flou à 36% du SUVmax. Dans cette configuration précise, optimale vis-à-vis de notre étude, nous avons pu établir pour chaque paramètre un seuil permettant de discriminer de manière significative les patientes à risques de récurrence.

La stratégie proposée dans ce chapitre se basait sur la modification de la méthode d'extraction des paramètres en vue d'augmenter leur valeur prédictive. Dans le prochain chapitre, nous allons adopter une nouvelle stratégie basée sur une méthode d'apprentissage non paramétrique permettant une analyse plus exhaustive de l'information contenue dans les images TEP.

## 8 RANDOM FOREST POUR LA PREDICTION DE RECIDIVE DES CANCERS DU COL UTERIN A PARTIR DE LA TEP AU 18F-FDG

*Ce chapitre porte sur une étude analysant un grand nombre de paramètres d'intensité, de forme et de texture calculés sur chaque image TEP pré- et per-thérapeutique et extraits pour différents seuillages (de 10% à 90% du SUVmax, par pas de 5%). Les objectifs de cette étude étaient i) d'identifier les paramètres les plus prédictifs de la récurrence à 3 ans chez les patientes atteintes d'un cancer du col utérin, et ii) d'augmenter la prédiction en combinant les paramètres par l'utilisation d'une méthode d'apprentissage supervisée, les forêts aléatoires.*

*Au total, 1026 paramètres ont été analysés par régressions logistiques uni- et bivariées, et parallèlement par une sélection de variables basée sur la construction de forêts aléatoires.*

*Les résultats de l'étude ont permis de montrer que 36 paramètres étaient hautement prédictifs en analyse uni- et bivariable. En particulier, le MTV en TEP1 avec un seuillage à 55% du SUVmax, le TLG en TEP2 avec un seuillage à 30% du SUVmax et l'inertie des matrices de co-occurrence en TEP2 avec un seuillage à 70% du SUVmax, confirment la capacité des images TEP pré- et per-thérapeutique à fournir de nombreux paramètres fortement prédictifs. Parallèlement, la forêt aléatoire utilisant la totalité des paramètres a donné un taux d'erreur de classification de 28,3%. En éliminant récursivement les variables les moins importantes, la combinaison de neuf paramètres a permis d'obtenir une forêt aléatoire avec un taux d'erreur de classification convergent à 13,21%, soulignant un très fort potentiel de prédiction. L'utilisation des forêts aléatoires a montré leur capacité à traiter un grand nombre de variables, supérieur au nombre d'observations, permettant d'identifier les paramètres les plus discriminants tout en augmentant la prédiction de récurrence.*

---

## CONTENU

<b>8</b>	<b>RANDOM FOREST POUR LA PREDICTION DE RECIDIVE DES CANCERS DU COL UTERIN A PARTIR DE LA TEP AU 18F-FDG.....</b>	<b>125</b>
8.1	Abstract.....	127
8.2	Introduction.....	127
8.3	Material and Methods.....	128
8.3.1	Patients and tumor.....	128
8.3.2	Treatment and follow-up.....	129
8.3.3	18F-FDG PET/CT acquisitions.....	129
8.3.4	18F-FDG PET analyses.....	129
8.3.5	statistical analyses.....	130
8.4	Results.....	131
8.5	Discussion.....	132
8.6	Appendix.....	140
8.7	Conclusion.....	142

---

## Random forests to predict tumor recurrence following cervical cancer therapy using pre- and per-treatment 18F-FDG PET parameters

Geoffrey Roman Jimenez<sup>1,3</sup>, Anne Devillers<sup>2</sup>, Julie Leseur<sup>2</sup>, Juan D. Ospina<sup>1,4</sup>, Henri Der Sarkissian<sup>3</sup>, Oscar Acosta<sup>1,3</sup>, Renaud De Crevoisier<sup>1,2,3</sup>.

<sup>1</sup>Laboratoire de Traitement du Signal et de l'Image (LTSI), INSERM 1099, Université de Rennes 1, F-35000, Rennes, France

<sup>2</sup>Centre Eugène Marquis, Department of Radiotherapy, Rennes, France

<sup>3</sup>Keosys medical imaging, Department of medical imaging, Saint-Herblain, France

<sup>4</sup>Escuela de Estadística, Universidad Nacional de Colombia, Medellín, Colombia

### 8.1 ABSTRACT

The ability to predict tumor recurrence after chemoradiotherapy of locally advanced cervical cancer is a crucial clinical issue to intensify the treatment of the most high-risk patients. The objective of this study was to investigate tumor metabolism characterizations from pre- and per-treatment 18F-FDF PET images to predict 3-year overall recurrence (OR). A total of 53 locally advanced cervical cancer patients underwent pre- and per-treatment 18F-FDG PET (respectively PET1 and PET2). Tumor metabolism was characterized through several delineations using different thresholds, based on a percentage of the maximum uptake (SUVmax), and applied by region-growing. The SUV distribution in PET1 and PET2 within each segmented region was characterized through 7 intensity and histogram-based parameters, 9 shape descriptors and 16 heterogeneity features for a total of 1026 parameters. Predictive capability of the extracted parameters was assessed using the area under the receiver operating curve (AUC) associated to uni- and multi-variate logistic regression models and random forest (RF) classifier. In univariate analyses, 36 parameters were highly significant predictors of 3-year OR ( $p < 0.01$ ), AUC ranging from 0.72 to 0.83. With RF, the Out-of-Bag (OOB) error rate using the totality of the extracted parameters was 28.3% (AUC=0.70). By recursively eliminating the less important variable, OOB error rate of the RF classifier using the nine most important parameters was 13.21% (AUC=0.90). Results suggest that both pre- and per-treatment 18F-FDG PET exams provides meaningful information to improve the prediction of post-treatment tumor recurrence. RF classifier is able to handle a very large number of extracted features and allows the identification of the most predictive parameters.

### 8.2 INTRODUCTION

Cervical cancer is the second most commonly cancer and the third leading cause of cancer death worldwide [206]. Locally advanced cervix tumors are classically treated with concurrent chemo-radiotherapy followed by brachytherapy and sometimes surgery [184]. Using this therapeutic strategy, 5-year overall survival for locally advanced cervical cancer ranges from 70% to 80%, and about 20% to 30% of patients will present tumor recurrence within the 3 years following the treatment [178] [10]. Prognostic factors or early markers of recurrence are crucial to identify high-risk patients and to target "intensified" adjuvant treatments to these patients. The known prognostic factors are clinical FIGO stage, lymph node involvement and response to treatment assessed clinically or by MRI [6,185].



In the last few decades, [<sup>18</sup>F]-2-fluoro-2-deoxy-D-glucose positron emission tomography (<sup>18</sup>F-FDG PET) has provided meaningful markers to early predict post-treatment outcomes [50,51,64,161,180,196,207]. In these studies, <sup>18</sup>F-FDG tumor uptake has been quantitatively assessed to evaluate the recurrence predictive capability of PET images by extracting features from <sup>18</sup>F-FDG PET signal through manual or automatic segmentations. In these studies, tumors are often characterized regarding only one or a few specified measurements of the standardized uptake value (SUV) intensity, such as the maximum uptake (SUV<sub>max</sub>), the mean SUV (SUV<sub>mean</sub>), the metabolic tumor volume (MTV) or the total lesion glycolysis (TLG). Recently, an increasing number of parameters has been explored, such as texture features to describe the heterogeneity of tumor metabolism or shape descriptors to characterize the 3D-object formed by the metabolic tumor region [68,138,140]. Moreover the PET analysis can be performed on both pre- and post-treatment PET images.

Whilst these studies have started to explore the benefits of this increasing amount of features for predicting outcomes, drawing definite conclusions is difficult as each study relies upon different tumor segmentation methodologies and considers different PET features in their predictive models. Consequently, a new statistical analysis challenge is arising to identify the most prognostic/predictive parameters among a large number of image features, and to improve therefore the prediction in the daily practice.

The random forests (RF) method is a machine-learning technique proposed by Breiman [208] which intensively uses intuitive classification and regression trees (CART) [209]. Used for classification and regression, RF does not require a priori information concerning the relationship between input and output variables. Indeed, it provides a framework to assess the importance of input variables at predicting the outcome. RF presents several advantages over other predictive methods. First, in contrast to other machine-learning methods, such as artificial neural network, RF does not overfit data, as demonstrated by Breiman [208]. Second, RF can deal with data sets containing more variables than observations, in contrast to more standard logistic regression or Cox proportional hazards regression models. Finally, RF can handle correlated/redundant variables. To our knowledge, RF has never been used to predict post-treatment tumor recurrence using FDG-PET in cervical cancer.

The objective of this work is two-fold. First, we sought to identify the most predictive parameters extracted from <sup>18</sup>F-FDG PET imaging to predict 3-year tumor recurrence in cervical cancer using logistic regression. Second, we set out to improve the prediction using the most predictive variables amount this considerable number of variables extracted from <sup>18</sup>F-FDG PET imaging, using RF.

## 8.3 MATERIAL AND METHODS

The overall workflow of the study is presented in Figure 8.1

### 8.3.1 PATIENTS AND TUMOR

Between January 2007 and December 2011, 53 consecutive patients treated for locally advanced cervical cancer in a curative intent, were included in this prospective trial and analyzed. Eligibility criteria were: histologically confirmed invasive cervix carcinoma, stage IB2-IVA according to the International Federation of Gynecology and Obstetrics (FIGO), complete staging comprising clinical staging, MRI and <sup>18</sup>F-FDG PET scan. A para aortic lymph node dissection was performed in all the patients, except in case of hyperfixation in this area at the PET imaging. A pelvic lymph node dissection could also be performed.

The median age was 52 years (range, 26 - 81 years). The stage distribution according to the FIGO classification was as follows: stage IB2 in 21 patients (40%), stage IIA1 in 9 patients (17%), stage IIA2 in 3 patients (6%), stage IIB in 14 patients (26%), stage IIIA in 2 patients (4%), stage IIIB in 1 patient (2%) and stage IVA in 3 patients (5%). Histology was squamous cell carcinoma for 40 patients and adenocarcinoma for 18 patients. Thirteen patients presented an involvement of pelvic lymph nodes. Eight patients presented pathological hyperfixation of para-aortic lymph nodes on PET and 5 an involvement of para-aortic lymph nodes on lymph nodes dissection.

### 8.3.2 TREATMENT AND FOLLOW-UP

All patients were treated with a combination of three-dimensional (3D) conformal external beam radiation therapy (EBRT) and pulse-dose-rate brachytherapy. EBRT delivered 45 Gy in the pelvis at 1.8 Gy per fraction, using an IMRT technique with 5 to 8 - 18 MV photon beams. A simultaneous integrated boost up to 55 Gy (at 2.2 Gy/fraction) was delivered in the positive pelvic nodes when shown in imaging. The target volume was extended to the para-aortic lymph nodes in case of lymph node involvement in the para-aortic lymphadenectomy (5 patients) or at the PET imaging (8 patients). Pulse dose rate (PDR) brachytherapy was performed two weeks after the end of EBRT using a standard Nucletron® ring applicator delivering a dose of 15 Gy into the intermediate risk CTV according to the GEC ESTRO recommendations [190,191].

In case of parametrial involvement a lateropelvic external radiation boost was performed after brachytherapy to deliver a total dose of 60 Gy (4 patients). Fifty-one patients received concurrent weekly cisplatin chemotherapy (40 mg/m<sup>2</sup> weekly for 4-5 cycles). Two patients did not receive concurrent chemotherapy because of co-morbidities or advanced age (≥80 years). The total treatment duration did not extend 7 weeks.

For the 18 patients with stage IB2 and IIA-B having residual disease at clinical examination, or biopsy or PET or MRI 4 to 6 weeks after brachytherapy, a total hysterectomy was performed (8 patients with stage IB2, 3 patients with stage IIA1 and 7 patients with stage IIB). The follow-up consisted in clinical examination, MRI and FDG PET 4 to 6 weeks after the end of brachytherapy, at 3 months and then every 6 months.

### 8.3.3 18F-FDG PET/CT ACQUISITIONS

The 18F-FDG PET/CT were realized first before the beginning of radiation therapy (PET1), the second one after a mean dose of 40 Gy (range, 32-46 Gy) (PET2) and then in the follow-up. 18F-FDG PET/CT scans were acquired using a PET/CT device (Discovery ST; GE Healthcare), which consisted of a PET scanner (Advance NXi; GE Healthcare) and a 2-slice CT scanner (Light-Speed Plus; GE Healthcare). Approximately 60 minutes after the intravenous injection of 4 - 5 MBq of 18F-FDG (Flucis), acquisitions were performed until the base of the skull to the upper part of the thighs. Scanning parameters were 120 kV, 80 mA, 0.8-s tube rotation, and 3.75 mm slices spacing. PET images were displayed with an ordered-subset expectation maximization iterative reconstruction algorithm (OSEM, 2 iterations, 28 subsets) and a reconstruction iterative fully 3D algorithm (Discovery ST). Patients fasted for at least 4 hours, a premedication with diazepam was administered systematically, as well as intravenous injection of furosemide (0.02 mL/kg) 15 min prior acquisitions.

### 8.3.4 18F-FDG PET ANALYSES

### Image pre-processing

Each voxel from PET scan was converted into a standardized uptake value (SUV) for comparisons across patients. The SUV is a standardized decay-corrected value of  $^{18}\text{F}$ -FDG activity per unit volume of body weight (MBq/kg) [19]. The full-body PET/CT images were cropped to the pelvic region based on the coxal bone structures visualized on CT imaging. The box boundaries were delimited according to well-defined landmarks that could be easily reproducible. The bounds of the box were: the inferior edge of the ischium, the right and left edge of the ilium, the bottom edge of the fifth lumbar vertebra (L5), the posterior border of the sacrum and the anterior edge of the ischial pubic. In addition, a mask of a manual bladder CT delineation was subtracted to the resulting area to ensure the non-inclusion of bladder  $^{18}\text{F}$ -FDG uptake in the placed 3D-box.

### Image segmentations

Segmentations were performed in PET1 by region-growing (RG), using different threshold values computed from a percentage of the maximum uptake within the region ( $\text{SUV}_{\text{max}1}$ ) (between 10% and 90% of  $\text{SUV}_{\text{max}1}$ , by step of 5%). In order to analyze the cervical area in PET2, regions formed by a 40% threshold in PET1 were propagated from PET1 to PET2, using a rigid registration based on the block matching algorithm [20]. Segmentations on PET2 were then performed by region-growing using different thresholds of the maximum uptake within the region (from 10% to 90% of  $\text{SUV}_{\text{max}2}$ ), restricted in the propagated PET2 area restricted in the propagated area.

### $^{18}\text{F}$ -FDG PET feature extraction

The SUV distribution within each segmentation was characterized through 7 intensity and histogram-based parameters, 9 shape descriptors and 16 heterogeneity features. The extracted features are summarized in Table 8.1.

The following SUV measurements were extracted:  $\text{SUV}_{\text{max}}$ ; peak SUV ( $\text{SUV}_{\text{peak}}$ ), defined as the mean of the  $\text{SUV}_{\text{max}}$  and its 26 neighbors; mean SUV ( $\text{SUV}_{\text{mean}}$ ); total lesion glycolysis (TLG), defined as the sum of SUVs; variance of the SUV distribution ( $\text{SUV}_{\text{variance}}$ ); skewness of the SUV distribution ( $\text{SUV}_{\text{skewness}}$ ); and the kurtosis of the SUV distribution ( $\text{SUV}_{\text{kurtosis}}$ ). Note that the measurements  $\text{SUV}_{\text{max}}$  and  $\text{SUV}_{\text{peak}}$  are independent of the segmentation performed.

The shape of the 3D-object formed by each segmentation was first described by the metabolic tumor volume (MTV). In addition, eight shape descriptors were extracted. These descriptors were based on the consideration of i) the oriented bounding box encompassing the 3D-object, ii) the eigen values of the covariance matrix of the voxels' coordinates within the considered region and iii) the convex hull of the 3D-object.

To analyze the heterogeneity of the SUV distribution, we extracted six texture features proposed by Haralick [144] based on the gray-level co-occurrence matrix (GLCM) reflecting spatial gray-level dependencies, and 10 features based on Run Length Matrix (RLM) reflecting alignment of voxels with the same intensity [210–212]. To compute GLCM and RLM, intensity range into segmentation regions was quantified with 32 bins. The texture features calculated were computed for every direction covering the 26 neighborhood directions and were averaged to keep a restricted number of features.

Considering the measurements of  $\text{SUV}_{\text{max}}$  and  $\text{SUV}_{\text{peak}}$  and the 30 features extracted within the 17 segmentations performed on each PET1 and PET2, a total of 1026 features ( $2 \times (2 + 17 \times 30)$ ) were obtained for each of the 52 patients.

### 8.3.5 STATISTICAL ANALYSES

The 3-year overall recurrence (OR) was the endpoint of this study. The OR binary event was defined as “recurrence during the three years following treatment”. The recurrence event could be local and/or nodal and/or metastatic, detected by clinical examination and/or MRI and/or FDG PET.

To test the predictive capabilities of the extracted PET features we proceed as follows. Firstly, we defined the binary event “recurrence during the three years following the treatment”. Secondly, we used univariate logistic regression to test the association among this binary event and the 18F-FDG PET parameters. Significant parameters were combined pairwise in bivariate logistic regression. Thirdly, a RF was constructed to predict 3-year tumor recurrence using the totality of the extracted parameters. Features were classified according to their importance in the RF construction. Finally, we performed a backward variables selection as in [213] based on a recursively elimination of the less important variables.

The area under the receiver operating curve (AUC) was used to evaluate the predictive capability of the different models. Note that for RF, the AUC was computed using the same cohort used to construct the model.

The RF is based on classification and regression trees (CART), aggregating the different predictions coming from independent trees [208]. The mechanism to build a random forest can be summarized in three steps: 1) generating new databases by re-sampling with replacement from the original database; 2) building a classification tree on each one of the new databases; 3) finally, aggregating the predictions given by each classification tree. Thus, the RF is a classifier consisting of a collection of tree-structured classifiers, and each tree casts a unit vote for the most popular class.

In our calculation, the number of trees to train a RF was set to 100,000 and the rest of the function parameters were set to the default values of the *randomForest* function in R [214]. The quality of a RF was measured using the out-of-bag (OOB) error rate and the importance of variables in the RF construction were assessed by the variable importance (VI) measures, namely Mean Decrease Accuracy (MDA) and Mean Decrease Gini (MDG) provided by the R implementation of RF [26] and described in Appendix. Note that OOB error rate can be interpreted as the opposite of the AUC.

All calculations were performed using R environment and *randomForest* R package [214].

## 8.4 RESULTS

In univariate analysis, 36 out of the 1026 features were highly significant predictors of 3-year OR ( $p < 0.01$ ). The AUC used to measure the performance of the univariate models ranged from 0.72 to 0.83. Considering the significant variables from univariate analysis, 29 variables combinations were independent predictor of 3-year OR in multivariate analysis ( $p < 0.05$ ), with AUC ranging from 0.77 to 0.89. Table 8.2 and Table 8.3 present the most significant parameters with univariate logistic regression and the most predictive combinations found in multivariate analysis ( $AUC \geq 0.85$ ), respectively. These combinations combined couples of parameters from both pre- and per-treatment PET.

Using RF classifier, OOB error rate of the RF constructed with the totality of the extracted features was 28.3% ( $AUC = 0.71$ ). By recursively eliminating the less important variables, OOB error rate of the RF classifier using the nine most important parameters was 13.21% coming from both PET exams ( $AUC = 0.90$ ).

Figure 8.2 and Figure 8.3 present respectively, OOB error rate and AUC evolutions depending on the number of features used to construct the RF. Figure 8.4 presents the MDA and MDG measures for the RF constructed with the nine most important parameters.

## 8.5 DISCUSSION

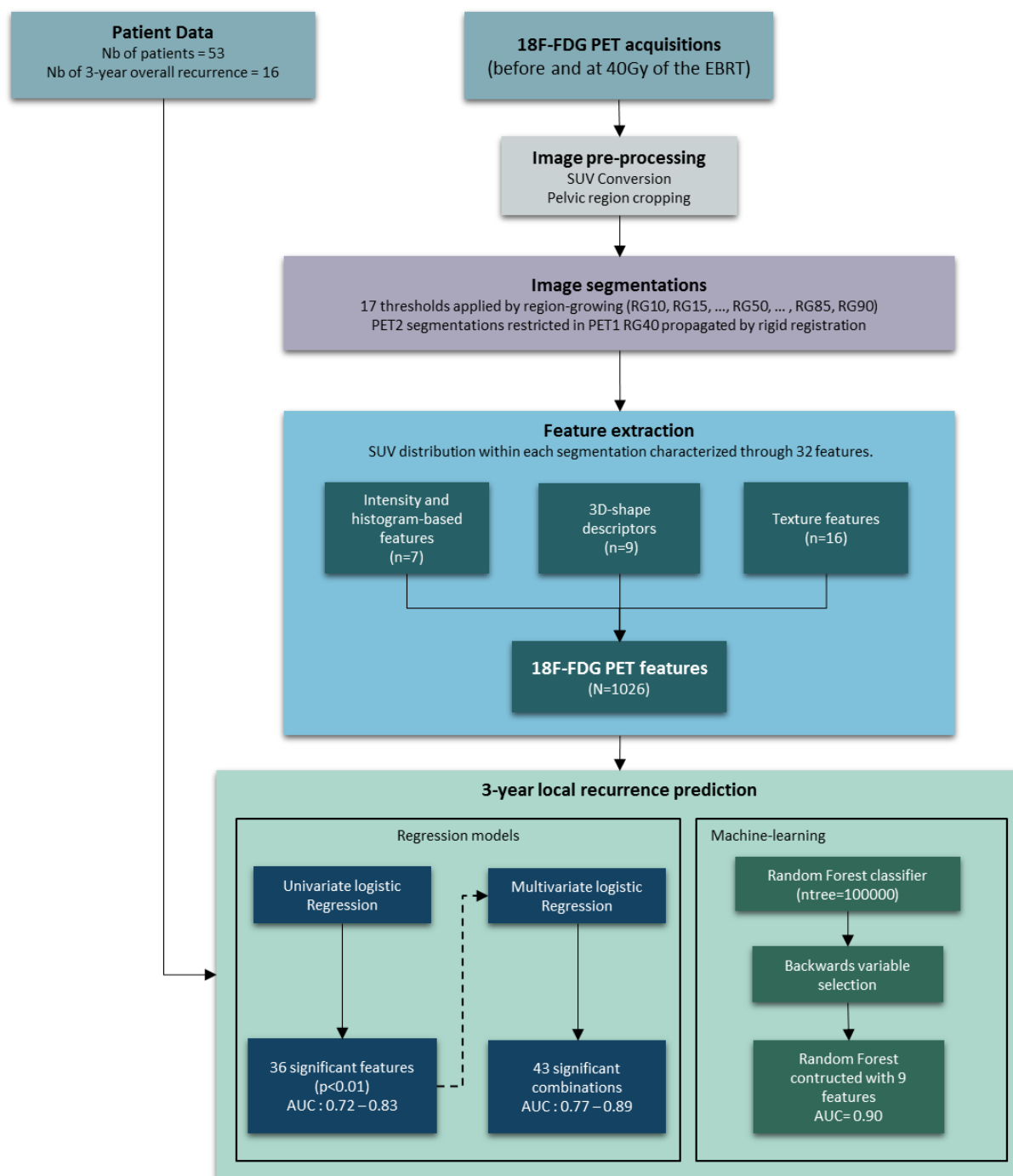
The goal of the study was to investigate if parameters derived from 18F-FDG PET imaging, singularly or in combination, could reliably improve the prediction of recurrence in cervical cancer. We included 53 patients with locally advanced cancer, treated by external beam radiotherapy concomitantly with cisplatin, followed by brachytherapy. Each patient was evaluated by 18F-FDG PET, before (PET1) and during the treatment (PET2). Every PET image was segmented using several region-growing thresholds. Finally, we characterized the SUV distribution within each segmentation by computing 7 histogram-based parameters, 9 shape descriptors and 16 heterogeneity features. A total of 1026 18F-FDG PET parameters were evaluated using logistic regression models and RF classifier in order to i) find the most predictive parameters between them and ii) improve the prediction of 3-year overall recurrence.

Using univariate logistic regression models, 36 parameters were highly significant predictors ( $p < 0.01$ ) with AUC ranging from 0.72 to 0.83 and 29 parameters were independent predictors in multivariate analyses (range AUC: 0.77-0.89). These results demonstrate the ability of the two 18F-FDG PET exams (PET1 and PET2) to provide meaningful information to predict post-treatment tumor outcomes.

Using the RF classifier, the OOB error rate of the model to predict 3-year overall recurrence using the totality of the extracted features was 28.3%. Evidently, because several similar thresholds were used to extract the 18F-FDG features, numerous variables were correlated and/or redundant. Thus, some features may have been used interchangeably in the decision trees of the random forest models leading to underestimated VI measures.

To identify the most relevant features, we recursively constructed RFs using the  $n$ th variables giving the best VI measures. Among all the computed RFs, the best OOB error rate (13.21%) was obtained with the RF constructed using the nine most important variables (AUC=0.90).

Our results suggest that both pre- and per-treatment 18F-FDG PET exams provide meaningful parameters to predict the post-treatment tumor recurrence. RF classifier was able to handle a very large number of imaging features and provided good predictive capability by combining 9 parameters.



**Figure 8.1 Overall workflow of the study**  
EBRT, external beam radiation therapy; SUV, standard uptake value; AUC, area under the receiver operating curve.  
Extracted features are described in Table 8.1

Table 8.1 Type, name and formula of the 32 features extracted to characterize the SUV distribution of PET images within each segmented region

Type	Name	Formula
Intensity and histogram-based features	SUVmax	$\max(X)$
	SUVpeak	$\frac{1}{K} \sum_k x_k$
	SUVmean	$\mu = \mathbb{E}[X]$
	SUVvariance	$\Sigma = \mathbb{E}[(X - \mu)^2]$
	SUVskewness	$\gamma_1 = \mathbb{E} \left[ \left( \frac{X - \mu}{\Sigma} \right)^3 \right]$
	SUVkurtosis	$\gamma_2 = \mathbb{E} \left[ \left( \frac{X - \mu}{\Sigma} \right)^4 \right]$
	TLG	$V_{\text{voxel}} \times \sum_n x_n$
<p><math>X = x_1, x_2, \dots, x_i, x_N</math> corresponds to the <math>N</math> voxels within the volume of interest;  <math>V_{\text{voxel}}</math> is the volume of a voxel;  <math>K</math> are the voxels</p>		
3D-shape descriptors	MTV	$V_{\text{voxel}} \times N$
	Eccentricity	$\sqrt{\frac{\lambda_1 - \lambda_3}{\lambda_1}}$
	Sphericity	$\sqrt[3]{\frac{\lambda_2 \cdot \lambda_3}{\lambda_1^2}}$
	Aspect Ratio	$\frac{L}{I}$
	Elongation	$\frac{L}{S}$
	Flatness	$\frac{S}{I}$
	Shape Factor	$\frac{S}{\sqrt{L \cdot I}}$
	Extent	$\frac{V_{\text{Object}}}{V_{\text{BoundingBox}}}$
Solidity	$\frac{V_{\text{Object}}}{V_{\text{ConvexHull}}}$	
<p><math>\lambda_1, \lambda_2</math> and <math>\lambda_3</math> are the eigen values of the voxels coordinates covariance matrix inside the object;  <math>L, I</math> and <math>S</math> are respectively the length, the width and the thickness of the oriented bounding box encompassing the object;  <math>V_{\text{Object}}</math> corresponds the volume of the object (MTV);  <math>V_{\text{ConvexHull}}</math> is the volume of the convex hull of the object.</p>		
Haralick texture features	Energy or Second angular moment	$\sum_{i,j} GLCM_{i,j}^2$
	Entropy	$\sum_{i,j} GLCM_{i,j} \times \log_2(GLCM_{i,j})$
	Inertia or contrast	$\sum_{i,j} (i - j)^2 \times GLCM_{i,j}$
	InverseDifferentMoment (IDM)	$\sum_{i,j} \frac{(i-j)^2}{GLCM_{i,j}}$
	Cluster Shade (CS)	$\sum_{i,j} (i + j - \mu_j - \mu_i)^3 \times GLCM_{i,j}$
	Cluster Prominence (CP)	$\sum_{i,j} (i - \mu_i + j - \mu_j)^4 \times GLCM_{i,j}$
<p><math>GLCM</math> is the Gray-Level Co-occurrence Matrix computed for one of the 26 neighborhood directions;  <math>\mu_i</math> and <math>\mu_j</math> are the mean value of the <math>i^{th}</math> row and the <math>j^{th}</math> column of the <math>GLCM</math>, respectively.</p>		
Run-Length Matrix texture features	Short Run Emphasis (SRE)	$\frac{1}{N_r} \sum_{i,j} \frac{RLM_{i,j}}{j^2}$
	Long Run Emphasis (LRE)	$\frac{1}{N_r} \sum_{i,j} RLM_{i,j} \times j^2$
	Low Gray-level Run Emphasis (LGRE)	$\frac{1}{N_r} \sum_{i,j} \frac{RLM_{i,j}}{i^2}$
	High Gray-level Run Emphasis (HGRE)	$\frac{1}{N_r} \sum_{i,j} RLM_{i,j} \times i^2$
	Short Run Low Gray-level Emphasis (SRLGE)	$\frac{1}{N_r} \sum_{i,j} \frac{RLM_{i,j}}{i^2 \times j^2}$
	Short Run High Gray-level Emphasis (SRHGE)	$\frac{1}{N_r} \sum_{i,j} \frac{RLM_{i,j} \times i^2}{j^2}$
	Long Run Low Gray-level Emphasis (LRGRE)	$\frac{1}{N_r} \sum_{i,j} \frac{RLM_{i,j} \times j^2}{i^2}$
	Long Run High Gray-level Emphasis (LRHGE)	$\frac{1}{N_r} \sum_{i,j} RLM_{i,j} \times i^2 \times j^2$
	Gray-Level Non-Uniformity (GLNU)	$\frac{1}{N_r} \sum_i \left( \sum_j RLM_{i,j} \right)^2$
	Run-Length Non-Uniformity (RLNU)	$\frac{1}{N_r} \sum_j \left( \sum_i RLM_{i,j} \right)^2$
<p><math>RLM</math> is the Run-Length Matrix computed for one of the 26 neighborhood directions;</p>		

Table 8.2 Significant predictors of 3-year overall recurrence using logistic regression (univariate analyses)

Exam	Type	Feature	Threshold	Values	AUC	p-value	OR	CI	
PET1	Haralick Textural features	Entropy	RG50	6.25 ± 1.52	0.78	0.0051	2.6716	[ 1.4708 - 5.9408 ]	
			RG55	5.64 ± 1.92	0.8	0.0031	2.5845	[ 1.4885 - 5.3103 ]	
	Intensity	TLG	RG50	170.9 ± 151.45	0.81	0.0036	1.0069	[ 1.0026 - 1.0121 ]	
			RG55	136.9 ± 128.23	0.82	0.0026	1.0084	[ 1.0033 - 1.0145 ]	
	Run-Length Textural Features	GLNU	RG50	9.29 ± 6.76	0.76	0.0073	1.1543	[ 1.0483 - 1.2959 ]	
			RG55	7.13 ± 5.27	0.83	0.0015	1.2716	[ 1.1113 - 1.5011 ]	
		RLNU	RG50	175.23 ± 137.24	0.79	0.0038	1.0076	[ 1.0028 - 1.0134 ]	
			RG55	132.08 ± 111.61	0.83	0.0016	1.0105	[ 1.0045 - 1.0179 ]	
	Shape	MTV	RG50	17.01 ± 13.2	0.79	0.0025	1.0897	[ 1.0353 - 1.1586 ]	
			RG55	12.74 ± 10.65	0.83	0.001	1.1266	[ 1.0558 - 1.2197 ]	
PET2	Haralick Textural features	Entropy	RG10	6.48 ± 1.21	0.78	0.0073	3.8405	[ 1.6174 - 11.7881 ]	
			RG15	6.44 ± 1.25	0.78	0.0073	3.6292	[ 1.5854 - 10.6235 ]	
			RG20	6.29 ± 1.46	0.78	0.0079	3.3546	[ 1.5259 - 9.2633 ]	
			RG65	2.19 ± 1.89	0.73	0.0095	1.615	[ 1.1499 - 2.4037 ]	
			RG70	1.53 ± 1.68	0.77	0.0052	1.786	[ 1.2212 - 2.7918 ]	
			Inertia	RG70	41.69 ± 74.83	0.82	0.0092	1.0187	[ 1.0063 - 1.0347 ]
	Intensity	Peak	///	3.69 ± 2.11	0.74	0.0096	1.6796	[ 1.1827 - 2.597 ]	
			TLG	RG10	67.73 ± 55.75	0.8	0.0023	1.0204	[ 1.0083 - 1.0354 ]
				RG15	65.3 ± 52.95	0.8	0.0017	1.0236	[ 1.0102 - 1.0407 ]
				RG20	60.81 ± 50	0.81	0.0015	1.0275	[ 1.0124 - 1.0476 ]
				RG25	54.99 ± 46.94	0.82	0.0015	1.0304	[ 1.0139 - 1.0529 ]
				RG30	48.12 ± 43.96	0.83	0.0015	1.0333	[ 1.0152 - 1.0582 ]
				RG35	42.63 ± 41.09	0.81	0.002	1.0347	[ 1.0153 - 1.0614 ]
				RG40	36.95 ± 37.79	0.76	0.0044	1.0323	[ 1.0126 - 1.0587 ]
				RG45	31.93 ± 33.74	0.76	0.0064	1.0341	[ 1.0125 - 1.0633 ]
				RG50	25.99 ± 28.74	0.77	0.0087	1.0384	[ 1.0133 - 1.0732 ]
				Run-Length Textural Features	GLNU	RG30	8.82 ± 5.97	0.78	0.0047
	RG35	7.2 ± 5.21	0.74			0.0091	1.214	[ 1.0662 - 1.4355 ]	
	RLNU	RG20	221.3 ± 157.89		0.74	0.0088	1.006	[ 1.0019 - 1.0111 ]	
		RG25	187.25 ± 136.33		0.72	0.0063	1.0079	[ 1.0028 - 1.0143 ]	
		RG30	152.45 ± 117.07		0.74	0.0049	1.0094	[ 1.0035 - 1.0169 ]	
		RG35	127.89 ± 102.27		0.73	0.0068	1.0098	[ 1.0033 - 1.0179 ]	
	Shape	MTV	RG20	21.75 ± 15.5	0.73	0.0076	1.0687	[ 1.0226 - 1.1289 ]	
			RG25	18.25 ± 13.17	0.74	0.0053	1.0956	[ 1.035 - 1.1784 ]	
			RG30	14.74 ± 11.21	0.75	0.0038	1.1148	[ 1.044 - 1.2116 ]	
			RG35	12.3 ± 9.75	0.74	0.0054	1.1185	[ 1.0419 - 1.2221 ]	

Only highly significant predictors are shown (p<0.01).

AUC = area under the receiver operating curve; OR = odd ratio; CI = confidence interval.

\* Most important variables found using random forest classifier. Values are mean ± SD.



Table 8.3 Multivariate logistic regression with combinations of independent predictors of 3-year overall recurrence

1st feature	2nd feature	AUC	p-value 1	OR 1	CI95% 1	p-value 2	OR 2	CI95% 2
MTV RG55 PET1*	Inertia RG70 PET2*	0.89	0.0028	1.1267	[ 1.0492 - 1.2305 ]	0.0298	1.014	[ 1.0038 - 1.0293 ]
GLNU RG55 PET1	Inertia RG70 PET2*	0.89	0.0043	1.2575	[ 1.0914 - 1.5038 ]	0.0374	1.0141	[ 1.0036 - 1.0301 ]
TLG RG55 PET1	Inertia RG70 PET2*	0.88	0.0063	1.0081	[ 1.0027 - 1.0146 ]	0.0265	1.0147	[ 1.004 - 1.0301 ]
Entropy RG55 PET1	Inertia RG70 PET2*	0.88	0.0057	2.9402	[ 1.5086 - 7.1085 ]	0.0175	1.0135	[ 1.0043 - 1.0282 ]
RLNU RG55 PET1	Inertia RG70 PET2*	0.88	0.0046	1.0101	[ 1.0036 - 1.018 ]	0.0345	1.0139	[ 1.0036 - 1.0295 ]
MTV RG50 PET1	Inertia RG70 PET2*	0.87	0.0074	1.088	[ 1.0282 - 1.1657 ]	0.0323	1.014	[ 1.0038 - 1.0295 ]
TLG RG50 PET1	Inertia RG70 PET2*	0.87	0.009	1.0066	[ 1.002 - 1.0122 ]	0.0294	1.0145	[ 1.0039 - 1.03 ]
Entropy RG50 PET1	Inertia RG70 PET2*	0.86	0.0126	2.8144	[ 1.3937 - 7.2629 ]	0.0297	1.0129	[ 1.0036 - 1.028 ]
GLNU RG50 PET1	Inertia RG70 PET2*	0.86	0.023	1.1355	[ 1.0253 - 1.2827 ]	0.0376	1.0148	[ 1.0036 - 1.0311 ]
RLNU RG50 PET1	Inertia RG70 PET2*	0.86	0.0121	1.007	[ 1.0019 - 1.0132 ]	0.0385	1.014	[ 1.0035 - 1.0298 ]
TLG RG30 PET2*	TLG RG40 PET2	0.85	0.0105	1.1323	[ 1.042 - 1.2655 ]	0.0445	0.8977	[ 0.7963 - 0.9867 ]
TLG RG30 PET2*	TLG RG45 PET2	0.85	0.0044	1.1278	[ 1.0484 - 1.2424 ]	0.0225	0.8923	[ 0.7984 - 0.9763 ]
GLNU RG30 PET2	Peak PET2	0.85	0.005	1.2198	[ 1.0777 - 1.4316 ]	0.0126	1.7715	[ 1.1932 - 2.9387 ]

Only the combinations of independent predictors ( $p < 0.05$ ) with  $AUC \geq 0.85$  are shown.

AUC = area under the receiver operating curve; OR = odd ratio; CI = confidence interval.

\* Most important variables found using random forest classifier.

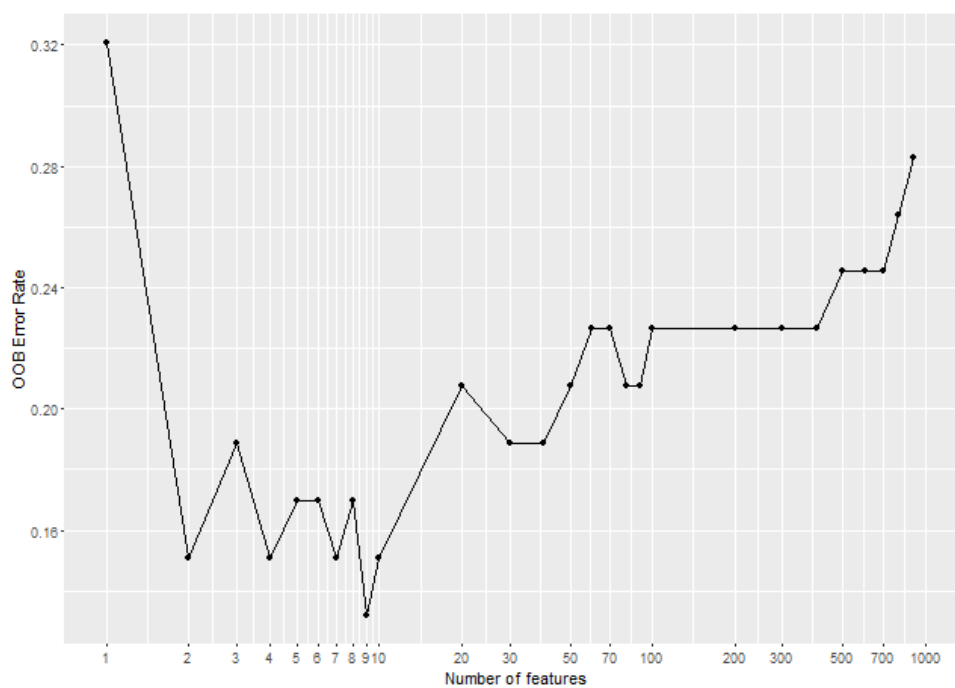


Figure 8.2 Evolution of the OOB error rate depending on the number of features used to construct the RF

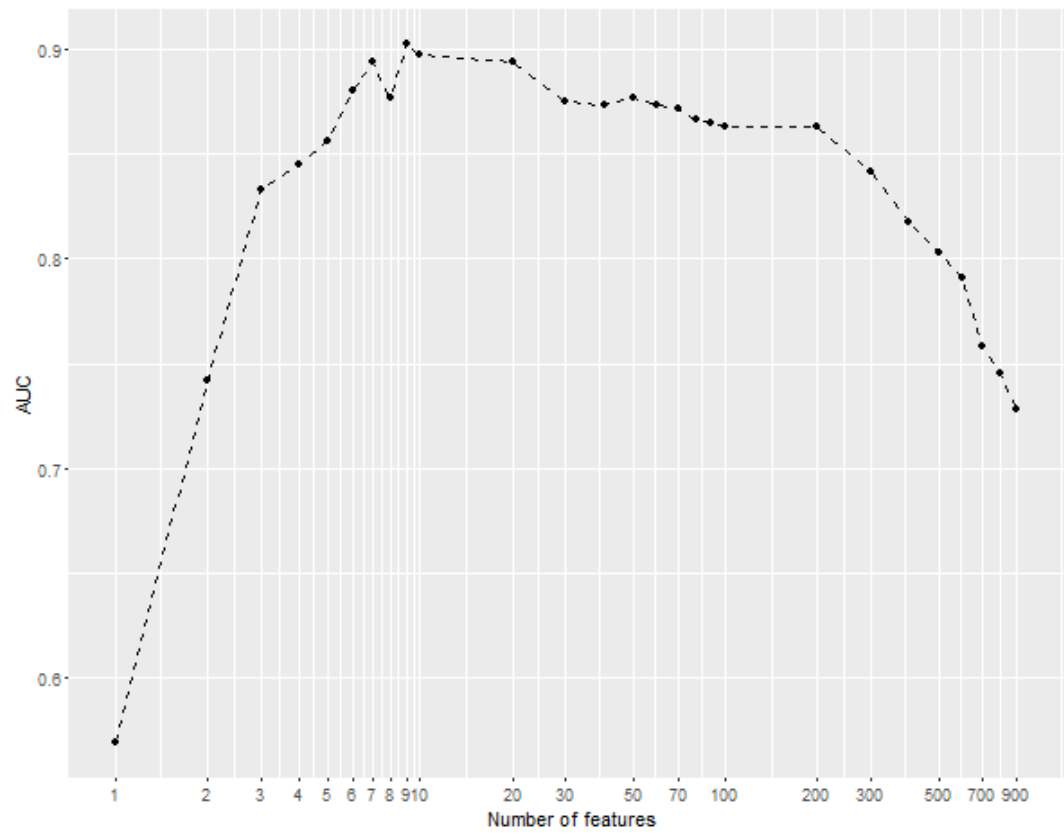
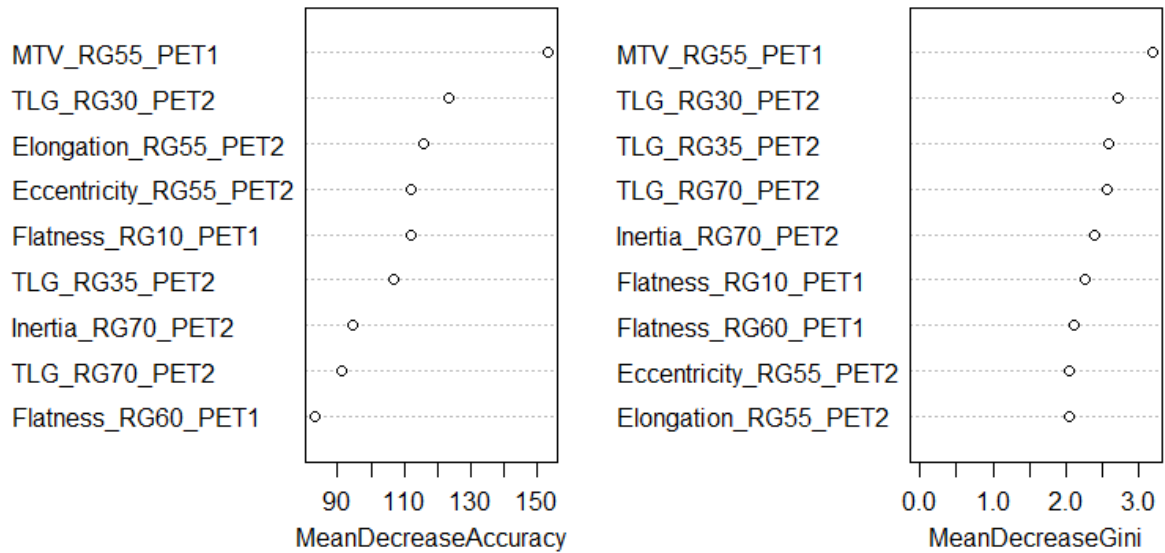


Figure 8.3 Evolution of the AUC depending on the number of features used to construct the RF



**Figure 8.4 Variable-importance measures for the RF constructed using the nine most important parameters to predict 3-year overall recurrence in locally advanced cancer**

Mean Decrease Accuracy is defined as the average difference between OOB errors before and after permuting the variable for all trees. A high value indicates that the variable is important in the RF construction.

Mean Decrease Gini is defined as the average Gini gain for each variable. A high value indicates that the variable is important in discriminating patients.

## 8.6 APPENDIX

Random forest calculation is divided into three steps: 1) Generating  $q$  new databases by re-sampling with replacement from the original database; 2) building a classification tree for each of the new  $q$  databases; finally, 3) aggregating the predictions given by each classification tree.

The first step is usually carried out using bootstrap (described in (28)). A bootstrap sample is a sample with replacement taken from the original database. For example, if a database with five patients is available, namely  $DB = \{Pt_1, Pt_2, Pt_3, Pt_4, Pt_5\}$ , five bootstrap samples could be  $DB^{(1)} = \{Pt_2, Pt_2, Pt_1, Pt_1, Pt_1\}$ ,  $DB^{(2)} = \{Pt_4, Pt_1, Pt_4, Pt_4, Pt_3\}$ ,  $DB^{(3)} = \{Pt_3, Pt_4, Pt_5, Pt_5, Pt_3\}$ ,  $DB^{(4)} = \{Pt_3, Pt_1, Pt_4, Pt_4, Pt_3\}$  and  $DB^{(5)} = \{Pt_4, Pt_3, Pt_2, Pt_3, Pt_2\}$ . In this case,  $q = 5$ . Of course, this is only an example and a bigger database is required in order to use the bootstrap methodology.

The methodology to generate a classification tree could be modified if the analyst wants to add more randomness during the model-building process. In traditional classification, a tree is created by partitioning the input variable space, which is usually referred to as *growing a tree*. For example, if the input space consists of three continuous variables,  $V_1$ ,  $V_2$  and  $V_3$ , to set up the  $k$ -th node of the tree, a threshold for each input variable is calculated, namely  $v_{01}^{(k)}$ ,  $v_{02}^{(k)}$  and  $v_{03}^{(k)}$ . Using these thresholds, *if-then* classification rules are set. When individuals are classed in two categories, namely  $\{0,1\}$  these rules are of the type “if  $V_i > v_{0i}^{(k)}$  then the individual is classed as 0, otherwise as 1.” The thresholds  $v_{0i}^{(k)}$  are chosen in such a way that they minimize some impurity measure, such as the Gini index (10), that may be interpreted as a classification error. The threshold with the lowest impurity measure becomes the  $k$ -th node of the tree. The same procedure is followed within each partition until a stopping criterion is met, mostly associated with a minimum number of observations within each partition. At this point, the tree has reached its maximal size and should be *pruned* to avoid overfitting. In RF with random inputs, a subset of inputs is randomly chosen to generate each node. This introduces another source of randomness. In addition, no pruning is performed in RF. These last two features give two characteristics to the tree ensemble: 1) each tree is different from the others; and 2) each tree is a good predictor on its training data, both characteristics essential for good results when aggregating individual outcomes.

The aggregation step depends on the type of problem, i.e. classification or regression. For regression, the outcome of the RF could be defined as the average of all tree outcomes, whereas concerning classification the RF outcome could be defined as the most voted category. Moreover, the portion of trees voting each category could be used as an estimate of the probabilities that a given individual belongs to each class, which in our work would be interpreted as a probability of normal tissue complication.

The quality of a RF can be measured using the out-of-bag (OOB) error. It consists of predicting each individual's outcome using only the trees where that individual was not present during the training step. In the illustrative example above, the OOB error for the  $Pt_2$  individual should be calculated using the predictions from the trees built using  $DB^{(2)}$ ,  $DB^{(3)}$  and  $DB^{(4)}$ . For a given individual, it is expected that approximately 37% of the trees do not make use of it during training. The OOB subsample associated to the  $DB^{(i)}$  is defined as the set of patients that do not belong to it. For example, the OOB subsample associated to the  $DB^{(1)}$  is  $DB_{OOB}^{(1)} = \{Pt_3, Pt_4, Pt_5\}$ . The OOB error associated to  $DB_{OOB}^{(1)}$  is the sum of the OOB errors obtained for patients  $Pt_3$ ,  $Pt_4$  and  $Pt_5$ .

The logic behind measuring the importance of a variable can be stated as follows: if an important variable is disturbed, then predictions are expected to be worse, whereas disturbing a non-important variable should

not have a significant impact on prediction. Variable importance assessment is then carried out using OOB errors. To assess the importance of variable  $V_j$ , first all OOB subsamples are obtained; secondly, the values of  $V_j$  are randomly permuted among all individuals, leading this to perturbed OOB subsamples denoted by  $DB_{OOB,V_j}^{(i)}$ ; finally, the OOB error associated with each  $DB_{OOB,V_j}^{(i)}$ , namely  $errorOOB_{V_j}^{(i)}$ , is compared with its correspondent associated to  $DB_{OOB,V_j}^{(i)}$ , namely  $errorOOB^{(i)}$ . If the variable  $V_j$  is important, then it is expected to have  $errorOOB_{V_j}^{(i)} > errorOOB^{(i)}$ . The importance of the variable  $V_j$ ,  $VI(V_j)$ , can be thus defined as the average difference between OOB errors before and after permuting the values of the variable  $V_j$  for all trees, as is shown in equation (5). It is thus called the mean decrease accuracy,

$$VI(V_j) = \frac{1}{q} \sum_{i=q}^q (errorOOB_{V_j}^{(i)} - errorOOB^{(i)}). \quad (1)$$

Another variable importance measure can be defined using the Gini index, which is used in classification trees to assess the impurity of a node. Suppose that an individual can be classed in two categories. For the  $k$ -th node, let  $p_{k0}$  and  $p_{k1}$  be the proportion of individuals classed as 0 and 1, respectively. The Gini index of the  $k$ -th node is calculated as in equation (6),

$$imp(k) = 1 - p_{k0}^2 - p_{k1}^2. \quad (2)$$

Note that when the node achieves perfect separation among the two classes, the Gini index equals 0. To split the  $k$ -th node, all the variables and several thresholds are tested. For instance, assuming that the  $k$ -th node is split using variable  $V_j$  and threshold  $v_{0j}^{(k)}$ , the Gini index for the left (L) and right (R) nodes can be calculated. Let's represent those by  $imp(V_j, v_{0j}^{(k)}, L)$  and  $imp(V_j, v_{0j}^{(k)}, R)$ , respectively, and let  $p_{(k,L),0}$  and  $p_{(k,R),1}$  represent the portion of the data in the  $k$ -th node that lies on the left and on the right side, respectively, after the splitting. The Gini gain for variable  $V_j$  is then calculated as in equation (7),

$$\begin{aligned} \Delta imp(k, V_j) = & imp(k) - p_{(k,L),0} imp(V_j, v_{0j}^{(k)}, L) \\ & - p_{(k,R),1} imp(V_j, v_{0j}^{(k)}, R) \end{aligned} \quad (3)$$

Equation (7) is minimized to choose both the variable and its threshold. To calculate the importance of a variable during the construction of the tree, one can aggregate all the Gini gains associated to this variable. Thus, if  $N_j$  represents all the node splits using the variable  $V_j$ , the total Gini gain for variable  $V_j$  can be defined as in equation (8),

$$Gini(V_j) = \frac{1}{|N_j|} \sum_{i \in N_j} \Delta imp(i, V_j). \quad (4)$$

A high value for  $Gini(V_j)$  indicates that variable  $V_j$  is important in discriminating the individuals. In a RF, it is possible to have the total Gini gain for each variable. Another measure of variable importance is defined in equation (9), which is called the mean decreasing Gini,

$$VI_G(V_j) = \frac{1}{q} \sum_{i=q}^q Gini_i(V_j) \quad (5)$$

The former  $VI_G(\cdot)$  is to be used in classification rather than in regression problems. These two measures,  $VI(\cdot)$  and  $VI_G(\cdot)$ , help to quantify the performance of a variable in terms of prediction results and discrimination capabilities, respectively. These two variable importance measures will be used as they are furnished by the R implementation of RF (see [26]).

## 8.7 CONCLUSION

Dans ce chapitre nous avons exploré l'utilisation combinée d'un grand nombre de paramètres d'intensité, de forme et de texture dans l'objectif d'étudier de manière exhaustive l'information fournie par les images TEP pré- et per-thérapeutiques pour la prédiction de récurrence.

Notre analyse univariée a permis d'identifier 36 paramètres hautement significatifs dans la prédiction de récurrence à 3 ans ( $p < 0,01$ ). Considérant ces paramètres, les AUC obtenues étaient comprises entre 0,72 et 0,83. Il est important de souligner, qu'à la différence de notre précédente étude présentée au chapitre 4, plusieurs caractéristiques de texture hautement prédictifs ont été mises en évidence notamment l'inertie et les paramètres basés sur les matrices d'isolongueur. Les paramètres les plus importants étaient le MTV en TEP1 avec un seuillage à 55% du SUVmax, le TLG en TEP2 avec un seuillage à 30% du SUVmax et l'inertie des matrices de co-occurrence en TEP2 avec un seuillage à 70% du SUVmax.

En analyse bivariée, nous avons mis en évidence 43 combinaisons de 29 de ces paramètres permettant d'augmenter fortement la prédiction, avec des AUC allant de 0,77 à 0,89. Nous avons observé que les combinaisons deux-à-deux de paramètres de nature distincte et extraits à différents temps du traitement radio-thérapeutique permettaient d'augmenter significativement la prédiction des modèles construits. En particulier, l'association du paramètre MTV à 55% du SUVmax en pré-thérapeutique et de la mesure de l'inertie (également appelée contraste) à 70% du SUVmax en per-thérapeutique donnait les meilleurs résultats de prédiction. Étant donné qu'un contraste important traduit fréquemment une forte hétérogénéité de l'activité tumorale résiduelle, une mesure élevée de l'inertie peut ainsi être interprétée comme une mauvaise réponse au traitement. Le volume de la tumeur initiale mesuré à 55% du SUVmax indique, quant à lui, l'existence d'un sous-volume particulièrement actif de l'activité métabolique réelle.

Ces régressions conventionnelles sont cependant limitées par le nombre de paramètres pouvant être intégrés au modèle. Les nouveaux modèles d'apprentissage non-paramétrique tels que les forêts aléatoires permettent quant à eux d'intégrer un grand nombre de variables en s'affranchissant des phénomènes de sur-apprentissages et, concomitamment, d'identifier les paramètres les plus pertinents.

En utilisant la totalité des 1026 paramètres extraits, la forêt aléatoire construite sur ce modèle a permis d'obtenir un taux d'erreur de classification de 28,3% et une AUC de 0,71. Ces scores ont pu être améliorés en procédant à l'élimination récursive des variables les moins importantes, pour arriver à une forêt aléatoire avec un taux d'erreur de classification convergeant à 13,21% et une AUC de 0,90.

Nous soulignons que la nature des paramètres utilisés pour la construction de cette forêt aléatoire est différente selon le stade pré- ou per-thérapeutique de l'étude. En effet, nous constatons que les paramètres pertinents en pré-thérapeutique correspondent à des descripteurs de forme (MTV et aplatissement) tandis que ceux identifiés en per-thérapeutique sont issus d'une combinaison de nature variée, à savoir d'intensité (TLG), de forme (élongation et excentricité) et de texture (inertie).

En conclusion, les résultats obtenus confirment bien que l'utilisation des forêts aléatoires permet de combiner un grand nombre de paramètres pour augmenter la prédiction, tout en identifiant les variables les plus pertinentes.

## 9 CONCLUSION ET PERSPECTIVES

Ces travaux de thèse s'inscrivent dans le contexte de prédiction de la récurrence post-traitement des cancers du col de l'utérus traités par radio-chimiothérapie. La problématique générale se base sur l'analyse du métabolisme tumoral, reflété dans les images TEP-FDG pré- et per-thérapeutiques, en vue d'identifier des paramètres pronostiques de la récurrence du cancer (et *a fortiori* de la survie) des patientes.

Nos travaux se sont focalisés sur l'extraction de paramètres des images TEP, pré- et per-thérapeutiques, d'une cohorte de patientes traitées pour un cancer du col de l'utérus localement avancé et sur l'évaluation de leur capacité à prédire les événements de récurrence.

La méthodologie proposée pour extraire les paramètres des images TEP repose principalement sur un schéma d'analyse des images constitué de trois familles de traitements :

1. Le prétraitement des images : améliorer, préparer et transformer les images ;
2. La segmentation des images : isoler le signal d'intérêt du reste de l'image ;
3. La quantification : extraire des paramètres TEP caractérisant le métabolisme tumoral et son évolution au cours du traitement.

Ces traitements constituent de véritables challenges méthodologiques car de nombreux facteurs limitent l'analyse de l'imagerie TEP comme la faible résolution spatiale, le faible contraste, l'effet de volume partiel ainsi que la présence de bruit et d'artefacts.

Nous avons proposé plusieurs solutions depuis le traitement d'images permettant de répondre, à ces problématiques dans notre objectif de prédiction de la récurrence. Nous présenterons ici les contributions, résultats et limites de ces travaux.

### 9.1 PRETRAITEMENT : CONTRIBUTIONS ET LIMITES

#### 9.1.1 SUPPRESSION DE L'HYPERFIXATION VÉSICALE

La première des limitations que nous avons abordée concernait la non-spécificité du radiotracer 18F-FDG aux cellules cancéreuses et, notamment, à la présence d'artefacts pouvant biaiser l'analyse du signal métabolique tumoral en TEP. Dans la localité du cancer du col utérin, la vessie est une des principales sources de gêne dans l'analyse de l'image TEP car celle-ci peut présenter une hyperfixation d'intensité comparable à celle de la tumeur. La vessie étant adjacente au col de l'utérus, un continuum d'hyperfixation peut alors se former entre les deux structures, rendant l'analyse du signal tumoral plus difficile. Les stratégies mises en place en clinique pour s'affranchir de cette problématique reposent sur un protocole lors de l'acquisition, en demandant par exemple aux patients d'aller uriner avant l'examen, ou au cours de l'analyse de l'image, en plaçant une boîte englobante autour de la tumeur ou en supprimant l'hyperfixation vésicale au travers d'une délimitation manuelle sur TDM. Ces solutions ne permettent pas de se prémunir d'une variabilité intra- et inter patient et observateur.

Afin de réduire cette variabilité, nous avons proposé un algorithme de détection semi-automatique de l'hyperfixation vésicale. S'appuyant sur une localisation *a priori* de l'hyperfixation de la vessie et de la tumeur sur les images TEP (placement de graines par l'utilisateur), l'algorithme utilise l'information TDM et TEP afin de discerner les voxels appartenant à la vessie tout en tentant de préserver l'activité tumorale. Testée sur un total de 52 images TEP/TDM pré-thérapeutiques, les résultats de classification ont montré



que la méthode était précise et robuste et qu'elle permettait de détecter la majeure partie de l'hyperfixation vésicale tout en préservant l'activité tumorale dans la grande majorité des cas. Cette méthode semble donc être en mesure d'augmenter la spécificité de l'analyse TEP en détectant les voxels n'appartenant pas au signal d'intérêt.

Il est à noter que cette méthode n'est aujourd'hui adaptée qu'à la localité du col et que certaines évolutions sont à prévoir afin de rendre la méthode plus adaptative à diverses localités. Par ailleurs, l'interaction de l'utilisateur avec la méthode est très forte (placement de graines, évaluation manuelle de la présence ou non d'hyperfixation vésicale) et une automatisation de la procédure est à envisager afin d'augmenter la reproductibilité des résultats. La reproductibilité de la méthode indépendamment du placement initial de la graine a été toutefois validée. Enfin, une mesure de l'impact de notre approche sur le reste du processus d'analyse des images et, notamment, sur la prédiction des paramètres extraits est également nécessaire.

## 9.2 SEGMENTATION : CONTRIBUTIONS ET LIMITES

### 9.2.1 L'IMPACT DE LA SEGMENTATION SUR LA PREDICTION DE RECIDIVE

Une des grandes problématiques dans l'analyse du signal métabolique tumoral est liée à la difficulté d'établir une segmentation du signal tumoral pertinente et reproductible. De nombreuses méthodes ont été proposées afin d'obtenir des segmentations très précises du métabolisme tumoral réel de chaque image. Cependant, aujourd'hui aucune ne fait consensus et la méthode la plus utilisée en routine clinique reste le seuillage d'intensité. De plus, les contributions scientifiques sont généralement tournées autour de la problématique de la délimitation précise de la totalité de la région métabolique tumorale et peu d'études ont abordé le problème de l'impact de la segmentation sur la valeur prédictive des paramètres TEP extraits. Cet aspect a été évalué au cours de deux études (cf. chapitres 5 et 6).

L'étude du chapitre 5 a permis de confirmer que la méthode de segmentation joue un rôle important dans l'extraction de paramètres prédictifs. Comparée à différents seuillages d'intensité (de 20% à 80% du SUVmax par pas de 10%, appliqués par croissance de région), notre méthode de segmentation développée pour la localité du col (RGAT+FGMC, présentée dans l'étude du chapitre 4) permettait d'obtenir des paramètres plus prédictifs. Cependant, au cours de cette étude, nous avons également pu montrer que certaines valeurs de seuillage d'intensité permettaient d'obtenir des résultats satisfaisants (notamment à 50% du SUVmax). L'avantage des seuillages d'intensité est, que pour une image donnée, la mesure est parfaitement reproductible d'opérateur en opérateur, avec un coût de mise en œuvre plus faible que les méthodes plus évoluées.

À la lumière de ces résultats, l'extraction de paramètres TEP prédictifs n'impose pas nécessairement une délimitation très précise de la région du métabolisme tumoral réel et l'analyse des images TEP doit se focaliser sur la recherche d'un signal d'intérêt, appartenant au métabolisme tumoral, dont la segmentation doit être aisément reproductible. Dans ce contexte, les seuillages d'intensité, malgré leurs limitations à s'adapter à la taille des tumeurs ou aux variations statistiques d'acquisition, peuvent fournir un cadre suffisamment robuste et reproductible à l'extraction de paramètres prédictifs de la récurrence.

Une seconde étude, exposée au chapitre 6, a permis d'évaluer l'impact de ces seuillages avec plus de précision en testant les valeurs de seuil allant de 30% à 70% du SUVmax par pas de 1%. Les résultats ont permis d'obtenir des valeurs de seuil qui produisent des paramètres plus prédictifs pour les acquisitions pré- et per-thérapeutiques. Pour les images pré-thérapeutiques, le seuil le plus performant se situe aux alentours de 55% du SUVmax. En routine clinique, la valeur de 42% du SUVmax est généralement reconnue

pour permettre une délimitation du volume métabolique relativement proche du volume tumoral obtenu sur TDM, pour les tumeurs inférieures à 30 mL selon [152]. Nous pouvons donc en déduire que, en TEP initiale, le seuillage à 55% permet d'extraire un sous-volume du métabolisme initial, particulièrement actif, dont la mesure de l'activité nous permet de discriminer au mieux les patientes à risques de récurrence.

Pour les images per-thérapeutiques, le seuillage d'intérêt pour l'extraction de paramètres prédictifs était à 32% du SUVmax. L'analyse du métabolisme en cours de traitement nécessitait donc une délimitation avec un seuil plus faible que lors de l'évaluation du métabolisme sur l'image initiale. Ceci peut s'expliquer par le fait que, évaluée à 40 Gy du traitement radio-thérapeutique, la fonte tumorale montre une activité très faible, parfois nulle, produisant une image d'intensité moyenne basse et difficilement détectable avec les seuils traditionnellement utilisés.

Ces valeurs de seuil ont prouvé leur pertinence dans nos études pour la prédiction de récurrence des cancers du col. Il est toutefois important de noter que celles-ci ne sont destinées qu'à l'extraction de paramètres dans une cohorte de patientes, sans pouvoir pour autant être considérées comme des méthodes de segmentation pertinentes pour la délimitation précise des volumes tumoraux des patientes considérées individuellement.

## 9.3 QUANTIFICATION : CONTRIBUTIONS ET LIMITES

### 9.3.1 LES PARAMETRES IDENTIFIES COMME PREDICTIFS DE LA RECURRENCE POST-TRAITEMENT DES CANCERS DU COL DE L'UTERUS

L'évaluation de la captation glucidique des images TEP en routine clinique est généralement focalisée sur des paramètres d'intensité maximale comme le SUVmax ou le SUVpeak. Ces mesures sont très utilisées pour leur relation suspectée avec l'agressivité tumorale. En parallèle, d'autres paramètres sont généralement étudiés comme le volume métabolique (MTV), la glycolyse totale de la lésion (TLG) ou encore la valeur moyenne des hyperfixations au sein de la lésion (SUVmean). Ces paramètres ont l'avantage de prendre en compte la totalité du métabolisme tumoral mais impliquent une délimitation de la tumeur dans l'image. Dans la littérature, un nombre croissant d'études ont proposé de prendre en compte de nouvelles mesures, notamment liées à l'hétérogénéité de l'hyperfixation au sein de la tumeur, afin de caractériser plus finement le métabolisme tumoral.

Dans une première étude, présentée au chapitre 4, nous avons exploré la valeur prédictive de plusieurs paramètres TEP basés sur des mesures d'intensité et de texture d'Haralick extraits des images TEP pré- et per-thérapeutiques. Le groupe analysé était alors de 35 patientes, avec un recul clinique médian de 29 mois et 8 événements de récurrence. Au total, 40 paramètres étaient analysés au regard de la prédiction de récurrence. La capacité discriminante des paramètres était explorée au travers de la recherche d'une valeur de seuil minimisant la valeur-p du test log-rang calculé sur des courbes de survies de Kaplan-Meier. Les paramètres donnant les meilleurs résultats étaient le TLG en TEP initial ( $p=0,0277$ , test log-rang), le SUVmax en cours de traitement ( $p=0,0107$ , test log-rang) et la différence de TLG entre les images pour l'évolution du signal en cours de traitement ( $p=0,0421$ , test log-rang). Ces analyses sont cependant intrinsèquement biaisées par la recherche exhaustive d'un seuil de séparation, rendant les valeurs seuil difficilement réutilisables par d'autres équipes. Dans cette étude, les paramètres de textures d'Haralick n'apportaient pas d'information majeure dans notre problématique de prédiction de récurrence.

Au cours de notre étude sur l'impact des méthodes de segmentation présentée au chapitre 5, le SUVmax, le SUVpeak, le MTV et le TLG ont été évalués. Avec un groupe de 41 patientes, un temps de recul clinique

médian de 22,79 mois et 9 évènements de récurrence, les résultats ont montré que le SUVmax et le SUVpeak n'étaient pas discriminants. Le paramètre TEP le plus discriminant était l'analogie du TLG calculé au sein des images paramétriques avec un seuillage à 50% du SUVmax ( $c=0,70$ ) et la méthode de segmentation spécifique à la localité du col RGAT+FGMC ( $c=0,74$ ). Le TLG, calculé au sein des images pré-thérapeutiques seules, donnait des résultats très similaires ( $c=0,69$  pour le seuillage à 50% et  $c=0,73$  avec RGAT+FGMC).

L'étude présentée au chapitre 6 a ensuite confirmé la pertinence du MTV et du TLG dans la prédiction de récurrence. Celle-ci était basée sur une cohorte de 53 patientes avec un recul clinique médian de 42 mois et 16 évènements de récurrence post-traitement. En pré-thérapeutique, le MTV et le TLG à 55% (MTV<sub>155%</sub> et TLG<sub>155%</sub>) du SUVmax étaient les paramètres les plus performants avec des c-index respectifs de  $c=0,80$  et  $c=0,78$ . En per-thérapeutique, le TLG et le MTV extraits avec un seuillage à 32% du SUVmax étaient les plus performants ( $c=0,71$  pour MTV<sub>232%</sub> et  $c=0,78$  pour TLG<sub>232%</sub>). Cette étude a également permis de dégager des seuils sur ces paramètres permettant de discriminer les patientes à risques. Notamment, les critères présentant une discrimination optimale des patients à risques de récurrence sont pour nos études :

- MTV<sub>155%</sub> > 9,8 cm<sup>3</sup> ( $p=6.10^{-5}$ , test log-rang);
- TLG<sub>155%</sub> > 70,4g ( $p=1.10^{-4}$ , test log-rang);
- MTV<sub>232%</sub> > 11,8cm<sup>3</sup> ( $p=0,01$ , test log-rang);
- TLG<sub>232%</sub> > 30,4g ( $p=0,001$ , test log-rang).

Nous noterons que ces valeurs de seuil ont été établies à l'aide d'une cohorte monocentrique, et une réévaluation au travers d'un groupe de patientes plus important et multicentrique est nécessaire.

### 9.3.2 L'ÉVALUATION DE L'ÉVOLUTION DU SIGNAL TUMORAL AU COURS DU TRAITEMENT POUR LA PREDICTION

Il existe différentes manières d'évaluer l'évolution du signal tumoral au cours du traitement. Il s'agit d'un signal multidimensionnel dynamique, dont les caractéristiques peuvent représenter des marqueurs importants de la récurrence. Une manière de faire est de mesurer l'évolution des caractéristiques de la tumeur à différents temps d'acquisition. C'est la méthode la plus utilisée car la plus simple à mettre en œuvre. À titre d'exemple, les critères de l'EORTC, RECIST et PERCIST se basent généralement sur le taux d'évolution (ou différence relative) de paramètres entre différents examens TEP – généralement SUVmax, SUVpeak ou SUVmean.

Dans notre étude du chapitre 4, nous avons évalué la valeur prédictive de la différence absolue de plusieurs paramètres extraits des images pré- et per-thérapeutiques. Dans cette étude, les images per-thérapeutiques étaient recalées dans l'espace anatomique des images pré-thérapeutiques et analysées sur la région d'intérêt définie par la segmentation de la tumeur initiale. Les analyses statistiques ont montré que seule la différence entre les TLG permettait d'obtenir une séparation des patientes dont les courbes de survie étaient significativement différentes.

Par la suite, dans notre étude du chapitre 6, la différence absolue et la différence relative des paramètres entre les deux temps d'acquisition ne permettaient pas de discerner les patientes à risques de récurrence. Malgré un plus grand nombre de patientes, un temps de recul clinique plus important et plusieurs méthodes de segmentation, nous pouvons conclure que sur cette cohorte, l'évolution des paramètres n'a pas permis d'évaluer l'apport supplémentaire dans la prédiction de la récurrence.

Une seconde approche pour évaluer l'évolution du signal tumoral consiste à créer et quantifier des images paramétriques. Sur ces images, chaque voxel représente l'évolution du signal à deux instants différents. Une cartographie de différences peut alors être construite et permettre d'identifier des régions répondant ou

non à la thérapie. Cette méthode nécessite cependant un recalage des images dans un repère commun afin de garantir la correspondance anatomique.

Dans l'étude du chapitre 5, nous avons étudié la valeur prédictive de caractéristiques extraites de ce type d'images paramétriques. Les images TEP pré- et per-thérapeutiques étaient recalées entre elles en propageant le champ de déformation obtenu par un recalage non-rigide des images TDM. L'évolution de chaque voxel au cours du temps a été évaluée par la différence absolue entre les valeurs de SUV des images TEP. Ainsi, les paramètres extraits de ces images surpassaient, pour chaque méthode de segmentation testée, la valeur prédictive des paramètres extraits des images considérées séparément les unes des autres. Nous pouvons cependant noter que les images pré-thérapeutiques seules ont produit des résultats similaires avec un coût de calcul et de mise en place moins important. Des études supplémentaires seront donc nécessaires pour évaluer l'apport des images paramétriques pour la prédiction de récurrence. Pour ces analyses, deux pistes pourront être considérées : i) l'impact de la qualité du recalage des images entre elles et ii) l'impact de l'apport de chaque paramètre de mesure d'évolution du signal entre chaque temps d'acquisition (différence relative, différence absolue, etc.).

Donc notre dernière contribution, nous avons exploré l'utilisation combinée d'un grand nombre de paramètres, incluant des mesures d'intensité, de forme et de texture, en utilisant des modèles de régression classiques et des modèles d'apprentissage supervisé basés sur les forêts aléatoires. Tout d'abord, les résultats ont permis de confirmer la capacité prédictive des paramètres identifiés au cours des études précédentes. Nous avons de plus montré l'émergence de paramètres de forme et de texture hautement prédictifs, qui n'avaient pas été identifiés au cours de l'étude du chapitre 4. Enfin, l'utilisation des forêts aléatoires a permis d'augmenter la valeur prédictive du modèle construit par l'association d'un ensemble de paramètres extraits aux différents stades thérapeutiques et de natures différentes – principalement de forme en pré-thérapeutique et de tous types pour les images per-thérapeutiques.

## 9.4 PERSPECTIVES

Les travaux présentés dans cette thèse ouvrent la voie à des études approfondies, ou nouvelles, concernant les méthodes utilisées pour l'analyse des images TEP pour la prédiction de récurrence des cancers du col de l'utérus.

Parmi les nombreuses perspectives de ces travaux les points clés sont :

### **Évaluer l'impact de la suppression d'artefact dans l'analyse du métabolisme tumoral**

Dans le but d'améliorer l'analyse du signal tumoral au sein des images TEP, nous avons proposé une méthode de suppression d'hyperfixation vésicale. Il est toutefois nécessaire d'évaluer l'impact d'un tel traitement sur la qualité des paramètres extraits par la suite et, notamment, vis-à-vis de leur valeur prédictive. Par ailleurs, cette méthode peut être améliorée en automatisant la procédure décisionnelle de présence, ou non, d'un artefact vésical.

### **Analyser la pertinence et l'interprétation des seuillages d'intérêt et des paramètres identifiés comme prédictifs**

Aux cours de ces travaux, les paramètres ressortant comme étant les plus prédictifs sont le TLG et le MTV avec un seuillage à 55% du SUVmax pour les images pré-thérapeutiques et à 32% pour les images per-thérapeutiques. Une analyse sur l'interprétation clinique de ces paramètres est nécessaire afin de comprendre et d'établir leur pertinence dans la prédiction de récurrence des cancers du col de l'utérus

localement avancés. Une autre perspective serait d'évaluer ces mêmes paramètres (en se laissant la possibilité de les adapter) pour différents stades tumoraux ou même d'autres localités.

### **Investiguer plus en profondeur l'impact des images paramétriques et de l'évolution du signal tumoral pour la prédiction**

L'évolution du signal tumoral tout le long du traitement a été évaluée par deux méthodes : i) l'évolution des paramètres TEP entre les deux examens (pré- et per-thérapeutiques) et ii) la création et l'analyse des images paramétriques d'évolution du signal. Malgré leur intérêt dans l'évaluation de la réponse au traitement, l'analyse de l'évolution ne nous a pas permis d'obtenir des paramètres prédictifs de la récurrence. Les images paramétriques ont, quant à elles, permis d'obtenir des paramètres d'intérêt pour la prédiction. Toutefois, la construction de celles-ci nécessite une mise en correspondance précise des deux examens dans le même espace anatomique. Des études sont donc nécessaires pour évaluer la validité et la pertinence des recalages mis en œuvre pour cette mise en correspondance ainsi que pour analyser l'utilisation d'images paramétriques reflétant l'évolution du signal tumoral.

### **Evaluer la pertinence et les perspectives apportées par l'extraction de paramètres calculés par pondération d'appartenance**

Nous avons proposé l'utilisation de paramètres pondérés par la probabilité d'appartenance au signal métabolique tumoral. L'idée repose sur la prise en compte de la nature incertaine du processus d'acquisition en TEP en généralisant le calcul des paramètres à l'utilisation d'une segmentation floue – là où cette segmentation est généralement binaire – afin de rendre l'extraction des paramètres plus pertinente et moins sensible au seuil utilisé. Nos résultats montrent que ces paramètres permettent d'obtenir des valeurs prédictives similaires aux paramètres classiques tout en réduisant l'impact de la segmentation.

### **Approfondir la contribution des forêts aléatoires pour la prédiction d'évènements post-thérapeutiques**

La tomographie par émission de positons ouvre la voie à l'extraction d'un grand nombre de paramètres permettant de caractériser de plus en plus finement les tumeurs et leur métabolisme (intensité, forme et texture). Parallèlement, les nombreuses méthodes de segmentation proposées dans la littérature augmentent d'autant plus le nombre de paramètres à évaluer. L'utilisation de méthodes statistiques récentes telles que les forêts aléatoires nous permettent aujourd'hui d'analyser simultanément un nombre important et varié de paramètres TEP de manière non-paramétrique. Nous avons montré que son utilisation permettait d'augmenter la prédiction et d'identifier les paramètres les plus pertinents. Des études plus approfondies considérant un volume de données plus important sont nécessaires pour valider et affiner ces résultats.

## VALORISATION SCIENTIFIQUE DU TRAVAIL DE THESE

### ARTICLES DE JOURNAUX PUBLIES

G Roman-Jimenez, J D Ospina, J Leseur, A Devillers, A Simon, P Terve, O Acosta and R De Crevoisier – *“Investigating the contribution of pre and per-treatment 18F-FDG PET-CT segmentation methodologies for post-treatment tumor recurrence prediction in cervical cancer”* – journal IRBM Volume 34, Issues 4-5, November 2013, Pages 274–277 Numéro spécial - RITS 2013 (doi : 10.1016/j.irbm.2013.09.002)

G Roman-Jimenez, R De Crevoisier, J Leseur, A Devillers, J D Ospina, A Simon, P Terve, O Acosta – *“Detection of bladder metabolic artefact in 18F-FDG PET imaging”* – Journal Computer in Biology and Medicine (CBM) - 2016

### ARTICLES DE JOURNAUX SOUMIS

J Leseur, G Roman-Jimenez, A Devillers, O Acosta, JD Ospina-Arango, D Williaume, A Simon, P Terve, V Lavoue, R De Crevoisier – *“18F-FDG PET/CT parameters during chemoradiotherapy to predict tumor recurrence in cervical cancer”* – soumis au journal « International Journal of Radiation Oncology and Biology and Physics » - 2016

G Roman-Jimenez, A Devillers, J Leseur, JD Ospina, H Der Sarkissian, O Acosta, R De Crevoisier – *“Impact of fuzzy-weighted parameters in 18F-FDG PET imaging in the prediction of cervical cancer recurrence”* – soumis au journal « Journal of Nuclear Medicine » - 2016

### COMMUNICATIONS EN CONFERENCE AVEC COMITE DE LECTURE

G Roman-Jimenez, J Leseur, A Devillers, JD Ospina, G Louvel, P Haignon, A Simon, R De Crevoisier, and O Acosta – *“Segmentation and characterization of tumors in 18F-FDG PET-CT for outcome prediction in cervical cancer radio-chemotherapy”* - MICCAI 2012

G Roman-Jimenez, JD Ospina, J Leseur, A Devillers, A Simon, P Terve, O Acosta and R De Crevoisier – *“Investigating the contribution of pre and per-treatment 18F-FDG PET-CT segmentation methodologies for post-treatment tumor recurrence prediction in cervical cancer”* – RITS 2013

G Roman-Jimenez, J Leseur, A Devillers, J D Ospina, O Acosta, M Gobeli, N Gautier, P Terve, V Lavoué, D Williaume, R De Crevoisier – *“18F-FDG PET parameters during radio-chemotherapy to predict tumor recurrence in cervical cancer”* – ESTRO 2014

G Roman-Jimenez, O Acosta, J Leseur, A Devillers, J Le Gouestre, J D Ospina, A Simon, P Terve and R De Crevoisier – *“Weighted quantification of 18F-FDG tumor metabolism activity using fuzzy-thresholding to predict post-treatment tumor recurrence”* – EMBC 2015

G Roman-Jimenez, A. Devillers, J. Leseur, J-D Ospina, H. Der Sarkissian, O. Acosta, R. De Crevoisier – *“Impact of fuzzy-thresholding of 18F-FDG PET images for cervical cancer recurrence prediction”* – ESTRO 2016



## RÉFÉRENCES

- [1] S. Zircirkeser, E. Şahin, M. Halac, S. Sager, M. Soret, S.L. Bacharach, et al., Global cancer statistics, *J. Nucl. Med.* 16 (2012) 351–369. doi:10.1007/s00259-011-1834-9.
- [2] W.T. Creasman, New gynecologic cancer staging., *Obstet. Gynecol.* 75 (1990) 287–288.
- [3] J.A. Green, J.M. Kirwan, J.F. Tierney, P. Symonds, L. Fresco, M. Collingwood, et al., Survival and recurrence after concomitant chemotherapy and radiotherapy for cancer of the uterine cervix: a systematic review and meta-analysis, *Lancet.* 358 (2001) 781–786.
- [4] P.G. Rose, B.N. Bundy, E.B. Watkins, J.T. Thigpen, G. Deppe, M.A. Maiman, et al., Concurrent cisplatin-based radiotherapy and chemotherapy for locally advanced cervical cancer, *N. Engl. J. Med.* 340 (1999) 1144–1153.
- [5] A. Singnurkar, S.B. Solomon, M. Gönen, S.M. Larson, H. Schöder, 18F-FDG PET/CT for the prediction and detection of local recurrence after radiofrequency ablation of malignant lung lesions, *J. Nucl. Med.* 51 (2010) 1833–1840.
- [6] J.M. Classe, P. Rauch, J.F. Rodier, P. Morice, E. Stoeckle, S. Lasry, et al., Surgery after concurrent chemoradiotherapy and brachytherapy for the treatment of advanced cervical cancer: morbidity and outcome: results of a multicenter study of the GCCLCC (Groupe des Chirurgiens de Centre de Lutte Contre le Cancer), *Gynecol. Oncol.* 102 (2006) 523–529.
- [7] R.M. Lanciano, B.W. Corn, The role of surgical staging for cervical cancer, in: *Semin. Radiat. Oncol.*, 1994: pp. 46–51.
- [8] J.K. Schwarz, B.A. Siegel, F. Dehdashti, P.W. Grigsby, Association of posttherapy positron emission tomography with tumor response and survival in cervical carcinoma, *Jama.* 298 (2007) 2289–2295.
- [9] A.H. Russell, W.-J. Koh, K. Markette, K.J. Russell, J.M. Cain, H.K. Tamimi, et al., Radical reirradiation for recurrent or second primary carcinoma of the female reproductive tract, *Gynecol. Oncol.* 27 (1987) 226–232.
- [10] G.M. Sommers, P.W. Grigsby, C.A. Perez, H.M. Camel, M.S. Kao, A.E. Galakatos, et al., Outcome of recurrent cervical carcinoma following definitive irradiation, *Gynecol. Oncol.* 35 (1989) 150–155. doi:10.1016/0090-8258(89)90033-4.
- [11] T. Mochizuki, E. Tsukamoto, Y. Kuge, K. Kanegae, S. Zhao, K. Hikosaka, et al., FDG uptake and glucose transporter subtype expressions in experimental tumor and inflammation models, *J. Nucl. Med.* 42 (2001) 1551–1555.
- [12] I. Buvat, Les limites du SUV, *M{é}decine Nucl{é}aire.* 31 (2007) 165–172.
- [13] J.W. Keyes Jr, SUV: standard uptake or silly useless value?, *J. Nucl. Med.* 36 (1995).
- [14] J.A. Thie, Understanding the standardized uptake value, its methods, and implications for usage, *J. Nucl. Med.* 45 (2004) 1431–1434.
- [15] H. Young, R. Baum, U. Cremerius, K. Herholz, O. Hoekstra, A.A. Lammertsma, et al., Measurement of clinical and subclinical tumour response using [18 F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations, *Eur. J. Cancer.* 35 (1999) 1773–1782.
- [16] R.L. Wahl, H. Jacene, Y. Kasamon, M.A. Lodge, From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors., *J Nucl Med.* 50 Suppl 1 (2009) 122S–150S. doi:10.2967/jnumed.108.057307.



- [17] R. Boellaard, N.C. Krak, O.S. Hoekstra, A.A. Lammertsma, Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study, *J. Nucl. Med.* 45 (2004) 1519–1527. <http://www.ncbi.nlm.nih.gov/pubmed/15347719>.
- [18] C. Nahmias, L.M. Wahl, Reproducibility of standardized uptake value measurements determined by 18F-FDG PET in malignant tumors, *J. Nucl. Med.* 49 (2008) 1804–1808.
- [19] Y. Nakamoto, K.R. Zasadny, H. Minn, R.L. Wahl, Reproducibility of common semi-quantitative parameters for evaluating lung cancer glucose metabolism with positron emission tomography using 2-deoxy-2-[18 F] fluoro-D-glucose, *Mol. Imaging Biol.* 4 (2002) 171–178.
- [20] M. Vanderhoek, S.B. Perlman, R. Jeraj, Impact of the definition of peak standardized uptake value on quantification of treatment response, *J. Nucl. Med.* 53 (2012) 4–11.
- [21] M.A. Lodge, M.A. Chaudhry, R.L. Wahl, Noise considerations for PET quantification using maximum and peak standardized uptake value, *J. Nucl. Med.* 53 (2012) 1041–1047.
- [22] R. Boellaard, R. Delgado-Bolton, W.J.G. Oyen, F. Giammarile, K. Tatsch, W. Eschner, et al., FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0, *Eur. J. Nucl. Med. Mol. Imaging.* 42 (2015) 328–354.
- [23] S.M. Larson, Y. Erdi, T. Akhurst, M. Mazumdar, H.A. Macapinlac, R.D. Finn, et al., 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 (1999) 159–171.
- [24] B. Foster, U. Bagci, A. Mansoor, Z. Xu, D.J. Mollura, A review on segmentation of positron emission tomography images, *Comput. Biol. Med.* 50 (2014) 76–96.
- [25] P. Bourguet, de T.S.O.R. Groupe, 2002 Standards, Options and Recommendations for the use of [18F]-FDG (PET-FDG) in cancerology, *Bull. Cancer.* 90 (2003) S1.
- [26] J. Czernin, M. Allen-Auerbach, H.R. Schelbert, Improvements in cancer staging with PET/CT: literature-based evidence as of September 2006, *J. Nucl. Med.* 48 (2007) 78S–88S.
- [27] H.J. Choi, W. Ju, S.K. Myung, Y. Kim, Diagnostic performance of computer tomography, magnetic resonance imaging, and positron emission tomography or positron emission tomography/computer tomography for detection of metastatic lymph nodes in patients with cervical cancer: Meta-analysis, *Cancer Sci.* 101 (2010) 1471–1479.
- [28] H.-H. Chou, T.-C. Chang, T.-C. Yen, K.-K. Ng, S. Hsueh, S.-Y. Ma, et al., Low value of [18F]-fluoro-2-deoxy-D-glucose positron emission tomography in primary staging of early-stage cervical cancer before radical hysterectomy, *J. Clin. Oncol.* 24 (2006) 123–128.
- [29] H.-H. Chou, H.-P. Chang, C.-H. Lai, K.-K. Ng, S. Hsueh, T.-I. Wu, et al., 18F-FDG PET in stage IB/IIB cervical adenocarcinoma/adenosquamous carcinoma, *Eur. J. Nucl. Med. Mol. Imaging.* 37 (2010) 728–735.
- [30] M. Signorelli, L. Guerra, C. Pirovano, C. Crivellaro, R. Fruscio, A. Buda, et al., Detection of nodal metastases by 18F-FDG PET/CT in apparent early stage ovarian cancer: a prospective study, *Gynecol. Oncol.* 131 (2013) 395–399.
- [31] M.L. Berman, H. Keys, W. Creasman, P. DiSaia, B. Bundy, J. Blessing, Survival and patterns of recurrence in cervical cancer metastatic to periaortic lymph nodes: A Gynecologic Oncology Group Study, *Gynecol. Oncol.* 19 (1984) 8–16.
- [32] G. Michel, P. Morice, D. Castaigne, M. Leblanc, A. Rey, P. Duvillard, Lymphatic spread in stage Ib and II cervical carcinoma: anatomy and surgical implications., *Obstet. Gynecol.* 91 (1998) 360–363.

- [33] M.J. Reinhardt, C. Ehrhrit-Braun, D. Vogelgesang, C. Ihling, S. Högerle, M. Mix, et al., Metastatic Lymph Nodes in Patients with Cervical Cancer: Detection with MR Imaging and FDG PET 1, *Radiology*. 218 (2001) 776–782.
- [34] H.J. Choi, J.W. Roh, S.-S. Seo, S. Lee, J.-Y. Kim, S.-K. Kim, et al., Comparison of the accuracy of magnetic resonance imaging and positron emission tomography/computed tomography in the presurgical detection of lymph node metastases in patients with uterine cervical carcinoma, *Cancer*. 106 (2006) 914–922.
- [35] C. Rossier, V. Dunet, O. Matzinger, J.O. Prior, TEP/TDM en radiothérapie: indications et perspectives, *Cancer/Radiothérapie*. 16 (2012) 152–163.
- [36] H. Veas, S. Senthamizhchelvan, R. Miralbell, D.C. Weber, O. Ratib, H. Zaidi, Assessment of various strategies for 18F-FET PET-guided delineation of target volumes in high-grade glioma patients, *Eur. J. Nucl. Med. Mol. Imaging*. 36 (2009) 182–193.
- [37] H. Zaidi, I. El Naqa, PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques., *Eur J Nucl Med Mol Imaging*. 37 (2010) 2165–2187. doi:10.1007/s00259-010-1423-3.
- [38] M. Hatt, D. Visvikis, Defining radiotherapy target volumes using {18F-fluoro-deoxy-glucose} positron emission tomography/computed tomography: still a Pandora’s box?: in regard to Devic et al. (Int J Radiat Oncol Biol Phys 2010)., *Int J Radiat Oncol Biol Phys*. 78 (2010) 1605. doi:10.1016/j.ijrobp.2010.08.002.
- [39] B. Lelandais, I. Gardin, L. Mouchard, P. Vera, S. Ruan, Segmentation of biological target volumes on multi-tracer PET images based on information fusion for achieving dose painting in radiotherapy, in: *Med. Image Comput. Comput. Interv. 2012, Springer, 2012*: pp. 545–552.
- [40] H. Necib, C. Garcia, A. Wagner, B. Vanderlinden, P. Emonts, A. Hendlisz, et al., Detection and characterization of tumor changes in {18F-FDG} {PET} patient monitoring using parametric imaging., *J Nucl Med*. 52 (2011) 354–361. doi:10.2967/jnumed.110.080150.
- [41] M. Hatt, D. Visvikis, O. Pradier, C.C. Rest, Baseline {<sup>18</sup>F-FDG} {PET} image-derived parameters for therapy response prediction in oesophageal cancer., *Eur J Nucl Med Mol Imaging*. 38 (2011) 1595–1606. doi:10.1007/s00259-011-1834-9.
- [42] W.A. Weber, V. Petersen, B. Schmidt, L. Tyndale-Hines, T. Link, C. Peschel, et al., Positron emission tomography in non-small-cell lung cancer: prediction of response to chemotherapy by quantitative assessment of glucose use, *J. Clin. Oncol*. 21 (2003) 2651–2657.
- [43] E. Eisenhauer, P. Therasse, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford, et al., New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1), *Eur. J. Cancer*. 45 (2009) 228–247.
- [44] C. Rousseau, A. Devillers, C. Sagan, L. Ferrer, B. Bridji, L. Campion, et al., Monitoring of early response to neoadjuvant chemotherapy in stage II and III breast cancer by [18F] fluorodeoxyglucose positron emission tomography, *J. Clin. Oncol*. 24 (2006) 5366–5372.
- [45] M. Schelling, N. Avril, J. Nährig, W. Kuhn, W. Römer, D. Sattler, et al., Positron emission tomography using [18F] fluorodeoxyglucose for monitoring primary chemotherapy in breast cancer, *J. Clin. Oncol*. 18 (2000) 1689–1695.
- [46] W.A. Weber, K. Ott, K. Becker, H.-J. Dittler, H. Helmberger, N.E. Avril, et al., Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging, *J. Clin. Oncol*. 19 (2001) 3058–3065.
- [47] S.H. Moon, S.H. Hyun, J.Y. Choi, Prognostic significance of volume-based PET parameters in cancer patients, *Korean J. Radiol*. 14 (2013) 1–12.

- [48] K. Pak, G.J. Cheon, H.-Y. Nam, S.-J. Kim, K.W. Kang, J.-K. Chung, et al., Prognostic value of metabolic tumor volume and total lesion glycolysis in head and neck cancer: a systematic review and meta-analysis, *J. Nucl. Med.* 55 (2014) 884–890.
- [49] P.W. Grigsby, B.A. Siegel, F. Dehdashti, Lymph node staging by positron emission tomography in patients with carcinoma of the cervix, *J. Clin. Oncol.* 19 (2001) 3745–3749.
- [50] E.A. Kidd, B.A. Siegel, F. Dehdashti, P.W. Grigsby, The standardized uptake value for F-18 fluorodeoxyglucose is a sensitive predictive biomarker for cervical cancer treatment response and survival, *Cancer.* 110 (2007) 1738–1744.
- [51] T.R. Miller, P.W. Grigsby, Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy, *Int. J. Radiat. Oncol.* 53 (2002) 353–359. doi:10.1016/S0360-3016(02)02705-0.
- [52] F. Xue, L.L. Lin, F. Dehdashti, T.R. Miller, B.A. Siegel, P.W. Grigsby, F-18 fluorodeoxyglucose uptake in primary cervical cancer as an indicator of prognosis after radiation therapy., *Gynecol Oncol.* 101 (2006) 147–151. doi:10.1016/j.ygyno.2005.10.005.
- [53] T.-C. Yen, L.-C. See, C.-H. Lai, C.-S. Tsai, A. Chao, S. Hsueh, et al., Standardized uptake value in para-aortic lymph nodes is a significant prognostic factor in patients with primary advanced squamous cervical cancer, *Eur. J. Nucl. Med. Mol. Imaging.* 35 (2008) 493–501.
- [54] R.J. Hicks, Role of 18F-FDG PET in assessment of response in non--small cell lung cancer, *J. Nucl. Med.* 50 (2009) 315–425.
- [55] T.R. Miller, E. Pinkus, F. Dehdashti, P.W. Grigsby, Improved prognostic value of 18F-FDG PET using a simple visual analysis of tumor characteristics in patients with cervical cancer, *J. Nucl. Med.* 44 (2003) 192–197.
- [56] A.K. Singh, P.W. Grigsby, F. Dehdashti, T.J. Herzog, B.A. Siegel, FDG-PET lymph node staging and survival of patients with FIGO stage IIIB cervical carcinoma, *Int. J. Radiat. Oncol. Biol. Phys.* 56 (2003) 489–493.
- [57] H.H. Chung, J.W. Kim, K.H. Han, J.S. Eo, K.W. Kang, N.-H. Park, et al., Prognostic value of metabolic tumor volume measured by FDG-PET/CT in patients with cervical cancer, *Gynecol. Oncol.* 120 (2011) 270–274.
- [58] B.S. Kim, I.J. Kim, S.-J. Kim, H.-Y. Nam, K.J. Pak, K. Kim, et al., The Prognostic Value of the Metabolic Tumor Volume in FIGO stage IA to IIB Cervical Cancer for Tumor Recurrence: Measured by F-18 FDG PET/CT., *Nucl. Med. Mol. Imaging* (2010). 45 (2011) 36–42. doi:10.1007/s13139-010-0062-8.
- [59] E.A. Kidd, I. El Naqa, B.A. Siegel, F. Dehdashti, P.W. Grigsby, FDG-PET-based prognostic nomograms for locally advanced cervical cancer, *Gynecol. Oncol.* 127 (2012) 136–140.
- [60] E.A. Kidd, B.A. Siegel, F. Dehdashti, J.S. Rader, D.G. Mutch, M.A. Powell, et al., Lymph node staging by positron emission tomography in cervical cancer: relationship to prognosis, *J. Clin. Oncol.* 28 (2010) 2108–2113.
- [61] J. Yoo, J.Y. Choi, H.J. Lee, S. Bin Park, M. Cheon, S.K. Cho, et al., Prognostic significance of volume-based metabolic parameters by 18F-FDG PET in uterine cervical cancer, *J. Nucl. Med.* 52 (2011) 143.
- [62] M. Miccò, H.A. Vargas, I.A. Burger, M.A. Kollmeier, D.A. Goldman, K.J. Park, et al., Combined pre-treatment MRI and 18F-FDG PET/CT parameters as prognostic biomarkers in patients with cervical cancer, *Eur. J. Radiol.* 83 (2014) 1169–1176.
- [63] C. Crivellaro, M. Signorelli, L. Guerra, E. De Ponti, A. Buda, C. Dolci, et al., 18F-FDG PET/CT can predict nodal metastases but not recurrence in early stage uterine cervical cancer, *Gynecol. Oncol.* 127 (2012) 131–135.

- [64] S.-Y. Ryu, M.-H. Kim, S.-C. Choi, C.-W. Choi, K.-H. Lee, Detection of early recurrence with 18F-FDG PET in patients with cervical cancer, *J. Nucl. Med.* 44 (2003) 347–352.
- [65] P.W. Grigsby, B.A. Siegel, F. Dehdashti, J. Rader, I. Zoberi, Posttherapy [18F] fluorodeoxyglucose positron emission tomography in carcinoma of the cervix: response and outcome, *J. Clin. Oncol.* 22 (2004) 2167–2171.
- [66] H.H. Chung, B.-H. Nam, J.W. Kim, K.W. Kang, N.-H. Park, Y.-S. Song, et al., Preoperative [18F] FDG PET/CT maximum standardized uptake value predicts recurrence of uterine cervical cancer, *Eur. J. Nucl. Med. Mol. Imaging.* 37 (2010) 1467–1473.
- [67] E.A. Kidd, M. Thomas, B.A. Siegel, F. Dehdashti, P.W. Grigsby, Changes in cervical cancer FDG uptake during chemoradiation and association with response, *Int. J. Radiat. Oncol. Biol. Phys.* 85 (2013) 116–122.
- [68] F. Renier, C. Habran, C. Bernard, F. Tixier, M. Hatt, D. Visvikis, et al., Intensity, volume-based and texture analyses of FDG PET/CT for predicting the outcome of patients with locally advanced cervical cancer treated by concomitant radiochemotherapy, *J. Nucl. Med.* 55 (2014) 631–631.
- [69] H.H. Chung, G.J. Cheon, Prognostic value of preoperative intratumoral FDG uptake heterogeneity in early stage uterine cervical cancer, *J. Nucl. Med.* 56 (2015) 346.
- [70] F. Hofheinz, J. Langner, B. Beuthien-Baumann, L. Oehme, J. Steinbach, J. Kotzerke, et al., Suitability of bilateral filtering for edge-preserving noise reduction in PET, *EJNMMI Res.* 1 (2011) 1–9.
- [71] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, *Pattern Anal. Mach. Intell. IEEE Trans.* 12 (1990) 629–639.
- [72] F.E. Turkheimer, J.A.D. Aston, R.B. Banati, C. Riddell, V.J. Cunningham, A linear wavelet filter for parametric imaging with dynamic PET, *Med. Imaging, IEEE Trans.* 22 (2003) 289–301.
- [73] F.E. Turkheimer, N. Boussion, A.N. Anderson, N. Pavese, P. Piccini, D. Visvikis, PET image denoising using a synergistic multiresolution analysis of structural (MRI/CT) and functional datasets, *J. Nucl. Med.* 49 (2008) 657–666.
- [74] U. Bagci, D.J. Mollura, Denoising PET images using singular value thresholding and stein’s unbiased risk estimate, in: *Med. Image Comput. Comput. Interv.* 2013, Springer, 2013: pp. 115–122.
- [75] J. Lee, X. Geets, V. Grégoire, A. Bol, others, Edge-preserving filtering of images with low photon counts, *Pattern Anal. Mach. Intell. IEEE Trans.* 30 (2008) 1014–1027.
- [76] X. Geets, J.A. Lee, A. Bol, M. Lonnew, V. Grégoire, A gradient-based method for segmenting FDG-PET images: methodology and validation, *Eur. J. Nucl. Med. Mol. Imaging.* 34 (2007) 1427–1438.
- [77] M. Abdoli, R. Dierckx, H. Zaidi, Contourlet-based active contour model for PET image segmentation, *Med. Phys.* 40 (2013) 82507.
- [78] N. Boussion, C.C. Le Rest, M. Hatt, D. Visvikis, Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging, *Eur. J. Nucl. Med. Mol. Imaging.* 36 (2009) 1064–1075.
- [79] M. Soret, S.L. Bacharach, I. Buvat, Partial-volume effect in PET tumor imaging, *J. Nucl. Med.* 48 (2007) 932–945.
- [80] O. Rousset, A. Rahmim, A. Alavi, H. Zaidi, Partial volume correction strategies in PET, *PET Clin.* 2 (2007) 235–249.
- [81] K. Erlandsson, I. Buvat, P.H. Pretorius, B.A. Thomas, B.F. Hutton, A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and

- oncology, *Phys. Med. Biol.* 57 (2012) R119.
- [82] B.-K. Teo, Y. Seo, S.L. Bacharach, J.A. Carrasquillo, S.K. Libutti, H. Shukla, et al., Partial-volume correction in PET: validation of an iterative postreconstruction method with phantom and patient data, *J. Nucl. Med.* 48 (2007) 802–810.
- [83] J. Tohka, A. Reilhac, Deconvolution-based partial volume correction in Raclopride-PET and Monte Carlo comparison to MR-based method, *Neuroimage.* 39 (2008) 1570–1584.
- [84] A.S. Kirov, J.Z. Piao, C.R. Schmidlein, Partial volume effect correction in PET using regularized iterative deconvolution with variance control based on local topology, *Phys. Med. Biol.* 53 (2008) 2577.
- [85] D.L. Barbee, R.T. Flynn, J.E. Holden, R.J. Nickles, R. Jeraj, 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 (2010) 221.
- [86] M. Hatt, C. Cheze le Rest, A. Turzo, C. Roux, D. Visvikis, C. le Rest, et al., A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET, *Med. Imaging, IEEE Trans.* 28 (2009) 881–893.
- [87] M. Hatt, C. Cheze le Rest, P. Descourt, A. Dekker, D. De Ruyscher, M. Oellers, et al., Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications, *Int. J. Radiat. Oncol. Biol. Phys.* 77 (2010) 301–308.
- [88] F. Yang, P.W. Grigsby, Delineation of FDG-PET tumors from heterogeneous background using spectral clustering, *Eur. J. Radiol.* 81 (2012) 3535–3541.
- [89] S. David, D. Visvikis, C. Roux, M. Hatt, Multi-observation {PET} image analysis for patient follow-up quantitation and therapy assessment., *Phys Med Biol.* 56 (2011) 5771–5788. doi:10.1088/0031-9155/56/18/001.
- [90] H. Necib, C. Garcia, A. Wagner, B. Vanderlinden, P. Emonts, A. Hendlisz, et al., Detection and characterization of tumor changes in 18F-FDG PET patient monitoring using parametric imaging., *J Nucl Med.* 52 (2011) 354–361. doi:10.2967/jnumed.110.080150.
- [91] S. Vauclin, I. Gardin, K. Doyeux, S. Hapdey, A. Edet-Sanson, P. Vera, Segmentation des images TEP au  $^{18}\text{F}$ -FDG. Principe et revue de la littérature, *Médecine Nucléaire.* 34 (2010) 358–369. doi:10.1016/j.mednuc.2010.03.005.
- [92] M. Hatt, N. Bousson, C.C.-L. Rest, D. Visvikis, O. Pradier, [Metabolically active volumes automatic delineation methodologies in {PET} imaging: review and perspectives]., *Cancer Radiother.* 16 (2012) 70–81; quiz 82, 84. doi:10.1016/j.canrad.2011.07.243.
- [93] A.-S. Dewalle-Vignion, A. El Abiad, N. Betrouni, C. Hossein-Foucher, D. Huglo, M. Vermandel, Les méthodes de seuillage en TEP: un état de l'art, *Médecine Nucléaire.* 34 (2010) 119–131.
- [94] E.F. Patz Jr, V.J. Lowe, J.M. Hoffman, S.S. Paine, P. Burrowes, R.E. Coleman, et al., Focal pulmonary abnormalities: evaluation with F-18 fluorodeoxyglucose PET scanning., *Radiology.* 188 (1993) 487–490.
- [95] A. Al-Sugair, R.E. Coleman, Applications of PET in lung cancer, in: *Semin. Nucl. Med.*, 1998: pp. 303–319.
- [96] A. Matthies, M. Hickeson, A. Cuchiara, A. Alavi, Dual time point 18F-FDG PET for the evaluation of pulmonary nodules, *J. Nucl. Med.* 43 (2002) 871–875.
- [97] K. Macdonald, J. Searle, I. Lyburn, The role of dual time point FDG PET imaging in the evaluation of solitary pulmonary nodules with an initial standard uptake value less than 2.5, *Clin. Radiol.* 66 (2011)

- 244–250.
- [98] H. Yu, C. Caldwell, K. Mah, I. Poon, J. Balogh, R. MacKenzie, et al., Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images, *Int. J. Radiat. Oncol. Biol. Phys.* 75 (2009) 618–625.
- [99] S. Tan, S. Kligerman, W. Chen, M. Lu, G. Kim, S. Feigenberg, et al., Spatial-temporal [18 F] FDG-PET features for predicting pathologic response of esophageal cancer to neoadjuvant chemoradiation therapy, *Int. J. Radiat. Oncol. Biol. Phys.* 85 (2013) 1375–1382.
- [100] S.-H. Shim, D.-Y. Kim, D.-Y. Lee, S.-W. Lee, J.-Y. Park, J.J. Lee, et al., Metabolic tumour volume and total lesion glycolysis, measured using preoperative 18F--FDG PET/CT, predict the recurrence of endometrial cancer, *BJOG An Int. J. Obstet. Gynaecol.* 121 (2014) 1097–1106.
- [101] U. Nestle, S. Kremp, A. Schaefer-Schuler, C. Sebastian-Welsch, D. Hellwig, C. Rube, et al., 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 (2005) 1342–1348.
- [102] S. Zhang, J. Xin, J. Ma, Q. Ma, H. Sun, Q. Guo, Defining PET tumor volume in cervical cancer with hybrid PET/MR: A comparative study, *J. Nucl. Med.* 55 (2014) 170.
- [103] Y.-C. Wang, T.-C. Hsieh, C.-Y. Yu, K.-Y. Yen, S.-W. Chen, S.-N. Yang, et al., The clinical application of 4D 18F-FDG PET/CT on gross tumor volume delineation for radiotherapy planning in esophageal squamous cell cancer, *J. Radiat. Res.* 53 (2012) 594–600.
- [104] T.C. Kwee, G. Cheng, M.G.E.H. Lam, S. Basu, A. Alavi, SUVmax of 2.5 should not be embraced as a magic threshold for separating benign from malignant lesions, *Eur. J. Nucl. Med. Mol. Imaging.* 40 (2013) 1475–1477.
- [105] Y. Hashimoto, T. Tsujikawa, C. Kondo, M. Maki, M. Momose, A. Nagai, et al., Accuracy of PET for diagnosis of solid pulmonary lesions with 18F-FDG uptake below the standardized uptake value of 2.5, *J. Nucl. Med.* 47 (2006) 426–431.
- [106] Y.E. Erdi, B.W. Wessels, M.H. Loew, A.K. Erdi, Threshold estimation in single photon emission computed tomography and planar imaging for clinical radioimmunotherapy, *Cancer Res.* 55 (1995) 5823s–5826s.
- [107] Y.E. Erdi, O. Mawlawi, S.M. Larson, M. Imbriaco, H. Yeung, R. Finn, et al., Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding, *Cancer.* 80 (1997) 2505–2509. doi:10.1002/(SICI)1097-0142(19971215)80:12+<2505::AID-CNCR24>3.0.CO;2-F [pii].
- [108] J.-F. Daisne, M. Sibomana, A. Bol, T. Doumont, M. Lonneux, V. Grégoire, Tri-dimensional automatic segmentation of {PET} volumes based on measured source-to-background ratios: influence of reconstruction algorithms., *Radiother Oncol.* 69 (2003) 247–250.
- [109] Q.C. Black, I.S. Grills, L.L. Kestin, C.-Y.O. Wong, J.W. Wong, A.A. Martinez, et al., Defining a radiotherapy target with positron emission tomography, *Int. J. Radiat. Oncol. Biol. Phys.* 60 (2004) 1272–1282.
- [110] A.J. Green, R.J. Francis, S. Baig, R.H.J. Begent, Semiautomatic volume of interest drawing for 18F-FDG image analysis—method and preliminary results, *Eur. J. Nucl. Med. Mol. Imaging.* 35 (2008) 393–406.
- [111] E. Day, J. Betler, D. Parida, B. Reitz, A. Kirichenko, S. Mohammadi, et al., A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients, *Med. Phys.* 36 (2009) 4349–4358.
- [112] C.-Y. Hsu, C.-Y. Liu, C.-M. Chen, Automatic segmentation of liver PET images, *Comput. Med. Imaging Graph.* 32 (2008) 601–610.

- [113] H. Li, W.L. Thorstad, K.J. Biehl, R. Laforest, Y. Su, K.I. Shoghi, et al., A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours, *Med. Phys.* 35 (2008) 3711–3721.
- [114] C. Riddell, P. Brigger, R.E. Carson, S.L. Bacharach, The watershed algorithm: a method to segment noisy PET transmission images, *Nucl. Sci. IEEE Trans.* 46 (1999) 713–719.
- [115] P. Tylski, G. Bonniaud, E. Decenciere, J. Stawiaski, J. Coulot, D. Lefkopoulos, et al., 18 F-FDG PET images segmentation using morphological watershed: a phantom study, in: *Nucl. Sci. Symp. Conf. Rec.* 2006. IEEE, 2006: pp. 2063–2067.
- [116] J.L. Chen, S.R. Gunn, M.S. Nixon, R.N. Gunn, Markov random field models for segmentation of PET images, in: *Inf. Process. Med. Imaging*, 2001: pp. 468–474.
- [117] D.W.G. Montgomery, A. Amira, H. Zaidi, Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model, *Med. Phys.* 34 (2007) 722–736.
- [118] M. Hatt, F. Lamare, N. Boussion, A. Turzo, C. Collet, F. Salzenstein, et al., Fuzzy hidden Markov chains segmentation for volume determination and quantitation in PET, *Phys. Med. Biol.* 52 (2007) 3467.
- [119] M. Aristophanous, B.C. Penney, M.K. Martel, C.A. Pelizzari, A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography, *Med. Phys.* 34 (2007) 4223–4235.
- [120] W. Zhu, T. Jiang, Automation segmentation of PET image for brain tumors, in: *Nucl. Sci. Symp. Conf. Rec.* 2003 IEEE, IEEE, 2003: pp. 2627–2629.
- [121] S. Belhassen, H. Zaidi, A novel fuzzy C-means algorithm for unsupervised heterogeneous tumor quantification in PET, *Med. Phys.* 37 (2010) 1309–1324.
- [122] J. Lapuyade-Lahorgue, D. Visvikis, M. Hatt, SPEQTACLE: A new fuzzy clustering algorithm for fully automatic delineation of tumors in PET images, *Soc. Nucl. Med. Annu. Meet. Abstr.* 55 (2014) 261.
- [123] A. Le Maitre, M. Hatt, O. Pradier, C. Cheze-le Rest, D. Visvikis, Impact of the accuracy of automatic tumour functional volume delineation on radiotherapy treatment planning, *Phys. Med. Biol.* 57 (2012) 5381.
- [124] M. Hatt, C. Cheze-Le Rest, E.O. Aboagye, L.M. Kenny, L. Rosso, F.E. Turkheimer, et al., Reproducibility of 18F-FDG and 3'-deoxy-3'-18F-fluorothymidine PET tumor volume measurements., *J. Nucl. Med.* 51 (2010) 1368–1376. doi:10.2967/jnumed.110.078501.
- [125] M. Hatt, C.C. Le Rest, N. Albarghach, O. Pradier, D. Visvikis, PET functional volume delineation: a robustness and repeatability study, *Eur. J. Nucl. Med. Mol. Imaging.* 38 (2011) 663–672.
- [126] G. Shafer, others, *A mathematical theory of evidence*, Princeton university press Princeton, 1976.
- [127] B. Lelandais, S. Ruan, T. Denoeux, P. Vera, I. Gardin, Fusion of multi-tracer PET images for dose painting, *Med. Image Anal.* 18 (2014) 1247–1259.
- [128] M.-H. Masson, T. Denoeux, ECM: An evidential version of the fuzzy c-means algorithm, *Pattern Recognit.* 41 (2008) 1384–1397.
- [129] S. Vauclin, K. Doyeux, S. Hapdey, A. Edet-Sanson, P. Vera, I. Gardin, 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 (2009) 6901.
- [130] C. Negoita, L. Zadeh, H. Zimmermann, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets Syst.* 1 (1978) 3–28.
- [131] A.-S. Dewalle-Vignion, N. Betrouni, R. Lopes, D. Huglo, S. Stute, M. Vermandel, A new method for volume segmentation of PET images, based on possibility theory, *Med. Imaging, IEEE Trans.* 30

- (2011) 409–423.
- [132] I. El Naqa, D. Yang, A. Apte, D. Khullar, S. Mutic, J. Zheng, et al., Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning, *Med. Phys.* 34 (2007) 4738–4749.
- [133] H. Yu, C. Caldwell, K. Mah, D. Mozeg, Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning., *IEEE Trans Med Imaging.* 28 (2009) 374–383. doi:10.1109/TMI.2008.2004425.
- [134] D. Han, J. Bayouth, Q. Song, A. Taurani, M. Sonka, J. Buatti, et al., Globally optimal tumor segmentation in PET-CT images: A graph-based co-segmentation method, in: *Inf. Process. Med. Imaging*, 2011: pp. 245–256.
- [135] U. Bagci, J.K. Udupa, J. Yao, D.J. Mollura, Co-segmentation of functional and anatomical images, in: *Med. Image Comput. Comput. Interv.* 2012, Springer, 2012: pp. 459–467.
- [136] U. Bagci, J.K. Udupa, N. Mendhiratta, B. Foster, Z. Xu, J. Yao, et al., Joint segmentation of anatomical and functional images: Applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images, *Med. Image Anal.* 17 (2013) 929–945.
- [137] Q. Song, J. Bai, D. Han, S. Bhatia, W. Sun, W. Rockey, et al., Optimal co-segmentation of tumor in PET-CT images with context information, *Med. Imaging, IEEE Trans.* 32 (2013) 1685–1697.
- [138] F. Tixier, C.C. Le Rest, M. Hatt, N. Albarghach, O. Pradier, J.-P. Metges, et al., Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer, *J. Nucl. Med.* 52 (2011) 369–378.
- [139] M. Vaidya, K.M. Creach, J. Frye, F. Dehdashti, J.D. Bradley, I. El Naqa, Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer, *Radiother. Oncol.* 102 (2012) 239–245.
- [140] I. El Naqa, P. Grigsby, A. Apte, E. Kidd, E. Donnelly, D. Khullar, et al., Exploring feature-based approaches in {PET} images for predicting cancer treatment outcomes., *Pattern Recognit.* 42 (2009) 1162–1171. doi:10.1016/j.patcog.2008.08.011.
- [141] G.J.R. Cook, C. Yip, M. Siddique, V. Goh, S. Chicklore, A. Roy, et al., Are Pretreatment 18F-FDG PET Tumor Textural Features in Non--Small Cell Lung Cancer Associated with Response and Survival After Chemoradiotherapy?, *J. Nucl. Med.* 54 (2013) 19–26.
- [142] R.J. Hicks, M.P. Mac Manus, J.P. Matthews, A. Hogg, D. Binns, D. Rischin, et al., Early FDG-PET imaging after radical radiotherapy for non--small-cell lung cancer: inflammatory changes in normal tissues correlate with tumor response and do not confound therapeutic response evaluation, *Int. J. Radiat. Oncol. Biol. Phys.* 60 (2004) 412–418.
- [143] V. Kalff, C. Duong, E.G. Drummond, J.P. Matthews, R.J. Hicks, 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 (2006) 14–22.
- [144] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural Features for Image Classification, *IEEE.* 3 (1973) 610–621. doi:10.1109/TSMC.1973.4309314.
- [145] R.E. Drzymala, R. Mohan, L. Brewster, J. Chu, M. Goitein, W. Harms, et al., Dose-volume histograms, *Int. J. Radiat. Oncol. Biol. Phys.* 21 (1991) 71–78.
- [146] X. Tang, Texture information in run-length matrices, *Image Process. IEEE Trans.* 7 (1998) 1602–1609.
- [147] G. Thibault, B. Fertil, C. Navarro, S. Pereira, P. Cau, N. Levy, et al., Texture indexes and gray level size zone matrix application to cell nuclei classification, (2009).



- [148] F. Tixier, M. Hatt, C.C. Le Rest, A. Le Pogam, L. Corcos, D. Visvikis, Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET, *J. Nucl. Med.* 53 (2012) 693–700.
- [149] F. Orlhac, M. Soussan, J.-A. Maisonobe, C.A. Garcia, B. Vanderlinden, I. Buvat, Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis, *J. Nucl. Med.* 55 (2014) 414–422.
- [150] W.-Y. Lin, K.-B. Wang, S.-C. Tsai, S.-S. Sun, Unexpected accumulation of F-18 FDG in the urinary bladder after bladder irrigation and retrograde filling with sterile saline: a possible pitfall in PET examination, *Clin. Nucl. Med.* 34 (2009) 560–563.
- [151] F.O. Roberts, D.H. Gunawardana, K. Pathmaraj, A. Wallace, T. Mi, S.U. Berlangieri, et al., Radiation dose to PET technologists and strategies to lower occupational exposure, *J. Nucl. Med. Technol.* 33 (2005) 44–47.
- [152] K.J. Biehl, F.-M. Kong, F. Dehdashti, J.-Y. Jin, S. Mutic, I. El Naqa, et al., 18F-FDG PET Definition of Gross Tumor Volume for Radiotherapy of Non--Small Cell Lung Cancer: Is a Single Standardized Uptake Value Threshold Approach Appropriate?, *J. Nucl. Med.* 47 (2006) 1808–1812.
- [153] F. Yang, M.A. Thomas, F. Dehdashti, P.W. Grigsby, Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer, *Eur. J. Nucl. Med. Mol. Imaging.* 40 (2013) 716–727.
- [154] X. Dong, L. Xing, P. Wu, Z. Fu, H. Wan, D. Li, et al., Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma: relationship between tumor 18F-fluorodeoxyglucose uptake heterogeneity, maximum standardized uptake value, and tumor stage, *Nucl. Med. Commun.* 34 (2013) 40–46.
- [155] A. Rahmim, H. Zaidi, PET versus SPECT: strengths, limitations and challenges, *Nucl. Med. Commun.* 29 (2008) 193–207.
- [156] T.Z. Belhocine, P.W. Grisby, FDG PET and PET-CT in uterine cancers, *Cancer Ther.* 3 (2005) 201–218.
- [157] B.J. Krause, S. Schwarzenböck, M. Souvatzoglou, FDG PET and PET/CT, *Recent Results Cancer Res.* 187 (2013) 351–369. doi:10.1007/978-3-642-10853-2\_12.
- [158] M. MacManus, U. Nestle, K.E. Rosenzweig, I. Carrio, C. Messa, O. Belohlavek, et al., Use of PET and PET/CT for radiation therapy planning: IAEA expert report 2006--2007, *Radiother. Oncol.* 91 (2009) 85–94.
- [159] R. Bar-Shalom, N. Yefremov, L. Guralnik, D. Gaitini, A. Frenkel, A. Kuten, et al., Clinical performance of {PET/CT} in evaluation of cancer: additional value for diagnostic imaging and patient management., *J Nucl Med.* 44 (2003) 1200–1209.
- [160] V. Kumar Dhingra, P. Kand, S. Basu, V.K. Dhingra, P. Kand, S. Basu, et al., Impact of FDG-PET and PET/CT imaging in the clinical decision-making of ovarian carcinoma: an evidence-based approach., *Women's Heal.* 8 (2012) 191–203. doi:10.2217/whe.11.91.
- [161] J. Leseur, A. Devillers, J.O. Arango, D. Williaume, I. Lecouillard, J. Léveque, et al., Metabolic Monitoring by 18F-FDG PET during Radio-chemotherapy for Locally Advanced Cervical Cancer: Predicting Outcome, *Int. J. Radiat. {Oncology\*Biography\*Physics}*. 81 (2011) S47 – S48. doi:10.1016/j.ijrobp.2011.06.096.
- [162] J. George, P. Claes, K. Vunckx, S. Tejpar, C.M. Deroose, J. Nuyts, et al., A textural feature based tumor therapy response prediction model for longitudinal evaluation with PET imaging, in: *Biomed. Imaging (ISBI), 2012 9th IEEE Int. Symp., 2012: pp. 1048–1051.* doi:10.1109/ISBI.2012.6235738.
- [163] M. Machtay, F. Duan, B.A. Siegel, B.S. Snyder, J.J. Gorelick, J.S. Reddin, et al., Prediction of survival

- by [18F]fluorodeoxyglucose positron emission tomography in patients with locally advanced non-small-cell lung cancer undergoing definitive chemoradiation therapy: results of the ACRIN 6668/RTOG 0235 trial., *J. Clin. Oncol.* 31 (2013) 3823–3830. doi:10.1200/JCO.2012.47.5947.
- [164] S.J. Rosenbaum, T. Lind, G. Antoch, A. Bockisch, False-positive FDG PET uptake- the role of PET/CT, *Eur. Radiol.* 16 (2006) 1054–1065.
- [165] M. Hatt, N. Bousson, C. Cheze-Le Rest, D. Visvikis, O. Pradier, Méthodologies de définition automatique des volumes métaboliquement actifs en TEP: évaluation et perspectives, *Cancer/Radiothérapie.* 16 (2012) 70–81. doi:10.1016/j.canrad.2011.07.243.
- [166] L.G. Strauss, P.S. Conti, The applications of PET in clinical oncology., *J Nucl Med.* 32 (1991) 623–48.
- [167] R. Xu, D. Wunsch, others, Survey of clustering algorithms, *Neural Networks, IEEE Trans.* 16 (2005) 645–678.
- [168] B.J. Frey, D. Dueck, Clustering by passing messages between data points, *Science (80-. ).* 315 (2007) 972–976.
- [169] H.-P. Kriegel, P. Kröger, J. Sander, A. Zimek, Density-based clustering, *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 1 (2011) 231–240.
- [170] A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, *{CA} A Cancer J. Clin.* 61 (2011) 69–90. doi:10.3322/caac.20107.
- [171] J.M. Kirwan, P. Symonds, J.A. Green, J. Tierney, M. Collingwood, C.J. Williams, A systematic review of acute and late toxicity of concomitant chemoradiation for cervical cancer., *Radiother Oncol.* 68 (2003) 217–226.
- [172] H.H. Chung, H. Jo, W.J. Kang, J.W. Kim, N.-H. Park, Y.-S. Song, et al., Clinical impact of integrated PET/CT on the management of suspected cervical cancer recurrence., *Gynecol Oncol.* 104 (2007) 529–534. doi:10.1016/j.ygyno.2006.09.009.
- [173] L.G. Strauss, P.S. Conti, The applications of {PET} in clinical oncology., *J Nucl Med.* 32 (1991) 623–650.
- [174] S. Ourselin, A. Roche, G. Subsol, X. Pennec, N. Ayache, Reconstructing a {3D} structure from serial histological sections, *Image Vis. Comput.* 19 (2001) 25–31. doi:10.1016/S0262-8856(00)00052-4.
- [175] M. Hatt, A. Le Pogam, D. Visvikis, O. Pradier, C.C. Le Rest, 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 (2012) 12–20. doi:10.2967/jnumed.111.092775.
- [176] J. Leseur, A. Devillers, J.O. Arango, D. Williaume, I. Lecouillard, J. Lèveque, et al., Metabolic Monitoring by {18F-FDG} {PET} during Radio-chemotherapy for Locally Advanced Cervical Cancer: Predicting Outcome, *Int. J. Radiat. {Oncology\*Biology\*Physics}.* 81 (2011) S47 – S48. doi:10.1016/j.ijrobp.2011.06.096.
- [177] H. Yu, C. Caldwell, K. Mah, D. Mozeg, Coregistered {FDG} {PET/CT-based} textural characterization of head and neck cancer for radiation treatment planning., *{IEEE} Trans Med Imaging.* 28 (2009) 374–383. doi:10.1109/TMI.2008.2004425.
- [178] P.J. Eifel, K. Winter, M. Morris, C. Levenback, P.W. Grigsby, J. Cooper, et al., Pelvic irradiation with concurrent chemotherapy versus pelvic and para-aortic irradiation for high-risk cervical cancer: an update of radiation therapy oncology group trial ({RTOG}) 90-01, *J. Clin. Oncol.* 22 (2004) 872–880.
- [179] J. Leseur, A. Devillers, D. Williaume, E. Le Prisé, C. Fougrou, C. Bouriel, et al., 18F-fluorodeoxyglucose PET/CT in cervix cancer: lymph node assessment and prognostic/predictive value of primary tumour analysis., *Cancer Radiother. J. La Soc. Fr. Radiother. Oncol.* 15 (2011) 699–

- 708.
- [180] L.L. Lin, Z. Yang, S. Mutic, T.R. Miller, P.W. Grigsby, {FDG-PET} imaging for the assessment of physiologic volume response during radiotherapy in cervix cancer, *Int. J. Radiat. Oncol. Biol. Phys.* 65 (2006) 177–181.
- [181] G. Roman-Jimenez, J. Leseur, A. Devillers, J. David, Segmentation and characterization of tumors in 18F-FDG PET-CT for outcome prediction in cervical cancer radio-chemotherapy, *Image-Guidance Multimodal Dose Plan. Radiat. Ther.* (2012) 17.
- [182] J.-P. Thirion, Image matching as a diffusion process: an analogy with Maxwell’s demons, *Med. Image Anal.* 2 (1998) 243–260. doi:10.1016/S1361-8415(98)80022-4.
- [183] F.E. Harrell, K.L. Lee, D.B. Mark, Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, *Stat. Med.* 15 (1996) 361–387.
- [184] D.K. Gaffney, B.A. Erickson-Wittmann, A. Jhingran, N.A. Mayr, A.A. Puthawala, D. Moore, et al., ACR Appropriateness Criteria{\textregistered} on Advanced Cervical Cancer Expert Panel on Radiation Oncology—Gynecology, *Int. J. Radiat. Oncol. Biol. Phys.* 81 (2011) 609–614.
- [185] N.A. Mayr, T. Taoka, W.T.C. Yuh, L.M. Denning, W.K. Zhen, A.C. Paulino, et al., Method and timing of tumor volume measurement for outcome prediction in cervical cancer using magnetic resonance imaging, *Int. J. Radiat. Oncol. Biol. Phys.* 52 (2002) 14–22.
- [186] B.E. Akkas, B.B. Demirel, A. Dizman, G.U. Vural, Do clinical characteristics and metabolic markers detected on positron emission tomography/computerized tomography associate with persistent disease in patients with in-operable cervical cancer?, *Ann. Nucl. Med.* 27 (2013) 756–763.
- [187] M. Bjurberg, E. Kjellén, T. Ohlsson, P.-O. Bendahl, E. Brun, Prediction of patient outcome with 2-deoxy-2-[[18F]fluoro-D-glucose-positron} emission tomography early during radiotherapy for locally advanced cervical cancer., *Int J Gynecol Cancer.* 19 (2009) 1600–1605. doi:10.1111/IGC.0b013e3181c00359.
- [188] D. Oh, J.E. Lee, S.J. Huh, W. Park, H. Nam, J.Y. Choi, et al., Prognostic significance of tumor response as assessed by sequential 18 F-fluorodeoxyglucose-positron emission tomography/computed tomography during concurrent chemoradiation therapy for cervical cancer, *Int. J. Radiat. Oncol. Biol. Phys.* 87 (2013) 549–554.
- [189] J.K. Schwarz, L.L. Lin, B.A. Siegel, T.R. Miller, P.W. Grigsby, 18-F-Fluorodeoxyglucose--Positron Emission Tomography Evaluation of Early Metabolic Response During Radiation Therapy for Cervical Cancer, *Int. J. Radiat. Oncol. Biol. Phys.* 72 (2008) 1502–1507.
- [190] C. Haie-Meder, R. Pötter, E. Van Limbergen, E. Briot, M. De Brabandere, J. Dimopoulos, et al., Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group☆(I): concepts and terms in 3D image based 3D treatment planning in cervix cancer brachytherapy with emphasis on MRI assessment of GTV and CTV, *Radiother. Oncol.* 74 (2005) 235–245.
- [191] R. Pötter, C. Haie-Meder, E. Van Limbergen, I. Barillot, M. De Brabandere, J. Dimopoulos, et al., Recommendations from gynaecological (GYN) GEC ESTRO working group (II): concepts and terms in 3D image-based treatment planning in cervix cancer brachytherapy—3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics, radiobiology, *Radiother. Oncol.* 78 (2006) 67–77.
- [192] M. Allen-Auerbach, W.A. Weber, Measuring response with FDG-PET: methodological aspects, *Oncologist.* 14 (2009) 369–377.
- [193] S. Ourselin, A. Roche, G. Subsol, X. Pennec, N. Ayache, Reconstructing a 3D structure from serial histological sections, *Image Vis. Comput.* 19 (2001) 25 – 31. doi:10.1016/S0262-8856(00)00052-4.

- [194] I.F. Ciernik, E. Dizendorf, B.G. Baumert, B. Reiner, C. Burger, J.B. Davis, et al., Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study, *Int. J. Radiat. Oncol. Biol. Phys.* 57 (2003) 853–863.
- [195] T.H. La, E.J. Filion, B.B. Turnbull, J.N. Chu, P. Lee, K. Nguyen, et al., Metabolic tumor volume predicts for recurrence and death in head-and-neck cancer, *Int. J. Radiat. Oncol. Biol. Phys.* 74 (2009) 1335–1341.
- [196] R. Lim, A. Eaton, N.Y. Lee, J. Setton, N. Ohri, S. Rao, et al., 18F-FDG PET/CT metabolic tumor volume and total lesion glycolysis predict outcome in oropharyngeal squamous cell carcinoma, *J. Nucl. Med.* 53 (2012) 1506–1513.
- [197] E.E. Kerre, M. Nachttegael, *Fuzzy techniques in image processing*, Springer Science & Business Media, 2000.
- [198] C. V Jawahar, P.K. Biswas, A.K. Ray, Analysis of fuzzy thresholding schemes, *Pattern Recognit.* 33 (2000) 1339–1349.
- [199] C.A. Murthy, S.K. Pal, Fuzzy thresholding: mathematical framework, bound functions and weighted moving average technique, *Pattern Recognit. Lett.* 11 (1990) 197–206.
- [200] L.A. Zadeh, others, *Calculus of fuzzy restrictions*, Electronics Research Laboratory, University of California, 1975.
- [201] G. Roman-Jimenez, J.-D. Ospina, J. Leseur, A. Devillers, J. Castelli, A. Simon, et al., Investigating the contribution of pre-and per-treatment  $^{18}\text{F}$ -FDG PET-CT segmentation methodologies for post-treatment tumor recurrence prediction in cervical cancer, *IRBM.* 34 (2013) 274–277.
- [202] G. Roman-Jimenez, J. Leseur, A. Devillers, J.D. Ospina, O. Acosta, P. Terve, et al., {18F-FDG PET} parameters during radio-chemotherapy to predict tumor recurrence in cervical cancer, *ESTRO 33 - Poster Discuss. 6 - Phys. Imaging.* (2014).
- [203] K. Higashi, Y. Ueda, Y. Arisaka, T. Sakuma, Y. Nambu, M. Oguchi, et al., 18F-FDG uptake as a biologic prognostic factor for recurrence in patients with surgically resected non--small cell lung cancer, *J. Nucl. Med.* 43 (2002) 39–45.
- [204] T.N. Showalter, T.R. Miller, P. Huettner, J. Rader, P.W. Grigsby, 18F-fluorodeoxyglucose-positron emission tomography and pathologic tumor size in early-stage invasive cervical cancer, *Int. J. Gynecol. Cancer.* 19 (2009) 1412–1414.
- [205] P. Peduzzi, J. Concato, E. Kemper, T.R. Holford, A.R. Feinstein, A simulation study of the number of events per variable in logistic regression analysis, *J. Clin. Epidemiol.* 49 (1996) 1373–1379. doi:10.1016/S0895-4356(96)00236-3.
- [206] R. Siegel, D. Naishadham, A. Jemal, Cancer statistics, 2012, *CA. Cancer J. Clin.* 62 (2012) 10–29.
- [207] H.H. Chung, H. Jo, W.J. Kang, J.W. Kim, N.-H. Park, Y.-S. Song, et al., Clinical impact of integrated {PET/CT} on the management of suspected cervical cancer recurrence., *Gynecol Oncol.* 104 (2007) 529–534. doi:10.1016/j.ygyno.2006.09.009.
- [208] L. Breiman, Random forests, *Mach. Learn.* 45 (2001) 5–32.
- [209] L. Breiman, J. Friedman, C.J. Stone, R.A. Olshen, *Classification and regression trees*, CRC press, 1984.
- [210] M.M. Galloway, Texture analysis using gray level run lengths, *Comput. Graph. Image Process.* 4 (1975) 172–179.
- [211] A. Chu, C.M. Sehgal, J.F. Greenleaf, Use of gray value distribution of run lengths for texture analysis, *Pattern Recognit. Lett.* 11 (1990) 415–419.

- [212] B. V Dasarathy, E.B. Holder, Image characterizations based on joint gray level—run length distributions, *Pattern Recognit. Lett.* 12 (1991) 497–502.
- [213] R. Diaz-Uriarte, S.A. de Andrés, Variable selection from random forests: application to gene expression data, *arXiv Prepr. Q-bio/0503025*. (2005).
- [214] A. Liaw, M. Wiener, Classification and regression by randomForest, *R News.* 2 (2002) 18–22.



## **Analyse des images de tomographie par émission de positons pour la prédiction de récurrence du cancer du col de l'utérus**

### **Résumé**

Ces travaux de thèse s'inscrivent dans le contexte de la prédiction de la récurrence en radiothérapie du cancer de l'utérus. L'objectif était d'analyser les images de tomographie par émission de positons (TEP) au 18F-fluorodésoxyglucose (<sup>18</sup>F-FDG) en vue d'en extraire des paramètres quantitatifs statistiquement corrélés aux événements de récurrence. Six études ont été réalisées afin de répondre aux différentes problématiques soulevées par l'analyse des images <sup>18</sup>F-FDG TEP telles que la présence d'artefact, l'isolation du métabolisme tumoral ou l'évaluation du signal en cours de traitement. Les études statistiques ont porté sur l'analyse de paramètres reflétant l'intensité, la forme et la texture du métabolisme tumoral avant, et en cours de traitement. À l'issue de ces travaux, le volume métabolique tumoral pré-thérapeutique ainsi que la glycolyse totale de la lésion per-thérapeutique apparaissent comme les paramètres les plus prometteurs pour la prédiction de récurrence de cancers du col de l'utérus. De plus, il apparaît que la combinaison de ces paramètres avec d'autres caractéristiques de texture ou de forme, à l'aide de modèles statistiques d'apprentissage supervisé ou de modèles de régression plus classiques, ont permis d'augmenter la prédiction des événements de récurrence.

---

## **Analysis of positron emission tomography images for recurrence prediction of cervical cancer**

### **Abstract**

This thesis deals with the issue of predicting the recurrence within the context of cervical cancer radiotherapy. The objective was to analyze positron emission tomography (PET) with 18F-fluorodeoxyglucose (<sup>18</sup>F-FDG) to extract quantitative parameters that could show statistical correlation with tumor recurrence. Six study were performed to address <sup>18</sup>F-FDG PET imaging issues such as the presence of bladder uptake artifacts, tumor segmentation impact, as well as the analysis of tumor evolution along the treatment. Statistical analyses were performed among parameters reflecting intensity, shape and texture of the tumor metabolism before, and during treatment. Results show that the pre-treatment metabolic tumor volume and the per-treatment total lesion glycolysis are the most promising parameters for cervical cancer recurrence prediction. In addition, combinations of these parameters with shape descriptors and texture features, using machine-learning methods or regression models, are able to increase the prediction capability.