HAL
open science

Méthodes d'analyse de volumes d'images multimodales pour le suivi longitudinal en oncologie

Jean-Marc Tacchella

To cite this version:

Jean-Marc Tacchella. Méthodes d'analyse de volumes d'images multimodales pour le suivi longitudinal en oncologie. Autre [cs.OH]. Université Pierre et Marie Curie - Paris VI, 2015. Français. NNT : 2015PA066049 . tel-01146051

HAL Id: tel-01146051

https://theses.hal.science/tel-01146051

Submitted on 27 Apr 2015

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 DE DOCTORAT DE
l'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité
Informatique

École doctorale Informatique, Télécommunications et Électronique (EDITE de Paris)

Présentée par

Jean-Marc Tacchella

Pour obtenir le grade de
DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

Méthodes d'analyse de volumes d'images multimodales pour le suivi longitudinal en oncologie

soutenance prévue le 02 avril 2015

devant le jury composé de :

Mme. Frédérique Frouin	Directrice de thèse
Mme. Elodie Roullot	Encadrante de thèse
Mme. Julie Delon	Rapporteur
M. Laurent Najman	Rapporteur
Mme. Mireille Garreau	Examinatrice
Mme. Aurélie Kas	Examinatrice
M. Nicolas Rougon	Examinateur

Table des matières

Listes des acronymes et des notations **6**

Introduction **8**

1 Contexte clinique **12**

1.1 Suivi de patients atteints de gliomes de haut grade et sous traitement antiangiogénique 12

1.1.1 Les gliomes de haut grade 13

1.1.2 Les traitements conventionnels et le traitement antiangiogénique 14

1.1.3 Modalités d'imagerie médicale utlisées pour le diagnostic et le suivi 15

1.1.4 Critères d'évaluation de l'évolution tumorale 16

1.1.5 Etude clinique sur les glioblastomes 21

1.2 Etude clinique sur le suivi des métastases hépatiques 22

1.3 Conclusion 23

2 Etat de l'art **24**

2.1 Recalage : méthodologie 25

2.1.1 Les approches géométriques 26

2.1.2 Les approches iconiques 28

2.1.3 Les approches hybrides 33

2.1.4 Les modèles de transformation 34

2.1.5 Les méthodes d'interpolation 38

2.1.6 La stratégie d'optimisation 40

2.2 Evaluation du recalage 43

2.2.1 Critères d'évaluation 44

2.2.2 Estimation d'une transformation de référence ("Gold Standard") 50

2.3 Segmentation 51

2.3.1 Méthodologie 51

2.3.2 Segmentation des gliomes sur les IRM 55

2.3.3 Segmentation des gliomes sur les données TEMP 57

2.3.4 Validation des méthodes de segmentation 59

2.4 Perspectives pour l'évaluation tumorale 60

2.5 Conclusion . . . 62

3 Stratégie d'analyse de l'évolution tumorale 63

3.1 Stratégie de recalage . . . 64

3.1.1 Principe de la stratégie de recalage . . . 64

3.1.2 Mise en oeuvre des méthodes de recalage . . . 65

3.1.3 Classement et évaluation . . . 67

3.1.4 Validation visuelle . . . 67

3.2 Segmentation et analyse de l'évolution des tumeurs . . . 68

3.2.1 Segmentation de la zone tumorale . . . 68

3.2.2 Critères de l'évaluation de l'évolution tumorale . . . 68

4 Application de la stratégie d'analyse de l'évolution tumorale pour des patients atteints de gliomes de haut grade 70

4.1 Spécificités de cette application . . . 71

4.1.1 Rappel du contexte médical . . . 71

4.1.2 Intérêt de la stratégie pour le recalage TEMP-IRM . . . 72

4.1.3 Base de données . . . 73

4.2 Mise en œuvre de la stratégie pour le recalage TEMP-IRM . . . 74

4.2.1 Choix des différentes approches de recalage . . . 74

4.2.2 Evaluation quantitative et classement . . . 81

4.2.3 Validation visuelle . . . 95

4.3 Mise en œuvre de la stratégie pour le recalage longitudinal IRM-IRM . . . 96

4.3.1 Choix des différentes approches de recalage . . . 97

4.3.2 Evaluation quantitative et classement . . . 98

4.3.3 Validation visuelle . . . 99

4.4 Segmentation des tumeurs sur l'IRM . . . 99

4.4.1 Segmentation des tumeurs sur la TEMP . . . 100

4.5 Critères d'analyse . . . 101

4.5.1 Critères d'analyse globale . . . 101

4.5.2 Critères d'analyse locale . . . 102

4.6 Conclusion . . . 103

5 Résultats et discussion relatifs à la stratégie de recalage appliquée au protocole IRM-TEMP pour les patients atteints de gliomes de haut grade 109

5.1 Résultats du recalage TEMP-IRM . . . 110

5.1.1 Validation du critère d'évaluation UC . . . 110

5.1.2 Performance de la stratégie . . . 116

5.1.3 Limites et perspectives . . . 121

5.2 Résultats du recalage longitudinal IRM-IRM . . . 122

5.2.1 Validation du critère d'évaluation Dice moyen . . . 122

5.2.2 Performance de la stratégie . . . 124
5.2.3 Conclusion . . . 125
5.3 Conclusion . . . 125

6 Résultats de l'analyse de l'évolution tumorale 127
6.1 Notations utiles au chapitre résultats . . . 127
6.2 Résultats obtenus avec les critères utilisés en routine clinique . . . 132
6.3 Comparaisons globales et locales des IRM après injection de Gadolinium avant et après le début du traitement . . . 137
6.3.1 Etude globale et comparaison avec les indices PFS et OS . . . 137
6.3.2 Etude locale et comparaison avec les indices PFS et OS . . . 139
6.3.3 Conclusion . . . 141
6.4 Comparaisons globales et locales des TEMP avant et après le début du traitement . . . 142
6.4.1 Analyse des volumes et intensités avant et après traitement . . . 142
6.4.2 Comparaison des différentes méthodes de segmentation . . . 157
6.4.3 Comparaison des données TEMP précoces et tardives . . . 164
6.4.4 Comparaison entre les informations extraites des TEMP avec les indices PFS et OS . 175
6.4.5 Discussion et conclusion relatives aux études TEMP . . . 182
6.5 Comparaisons globales et locales entre TEMP et IRM . . . 184
6.5.1 Etude locale . . . 187
6.6 Conclusion . . . 192

7 Mise en œuvre de la stratégie de recalage sur des données tomodensitométriques abdominales 194
7.1 Spécificités de cette application . . . 195
7.1.1 Contexte médical . . . 195
7.1.2 Base de données . . . 196
7.2 Mise en œuvre de la stratégie pour le recalage rigide longitudinal de TDM abdominales acquises pour le suivi de patients atteints de tumeurs hépatiques . . . 196
7.2.1 Choix des différentes approches de recalage . . . 196
7.2.2 Evaluation quantitative et classement . . . 199
7.2.3 Validation visuelle . . . 199
7.3 Mise en œuvre de la stratégie pour le recalage d'études de TDM abdominales de perfusion au premier passage . . . 200
7.3.1 Choix des différentes approches de recalage . . . 200
7.3.2 Evaluation quantitative et classement . . . 203
7.3.3 Validation visuelle . . . 203
7.4 Résultats du recalage TDM-TDM pour le suivi longitudinal . . . 205
7.4.1 Performance de la stratégie . . . 205
7.4.2 Discussion . . . 210

7.4.3 Limites et perspectives 210
7.4.4 Conclusion 212
7.5 Résultats du recalage TDM-TDM sur les études dynamiques de perfusion 213
7.5.1 Performance de la stratégie 213
7.5.2 Discussion 215
7.5.3 Conclusion 215
7.6 Conclusion 216

Conclusion et perspectives **217**

Bibliographie **223**

A Publications **237**
A.1 Publications originales dans des revues à comité de lecture 237
A.2 Articles publiés dans des actes de conférences 237
A.3 Résumés publiés dans des actes de conférences 237
A.4 Autres publications 238

Listes des acronymes et des notations

Liste des acronymes

CCN Corrélation Croisée Normalisée
DC Distance de Chanfrein
ECC Entropie du Coefficient de Corrélation
eRWT extended Regression Without Truth
FDA Food and Drug Administration
FDG FluoroDésoxyGlucose
FFD Free Form Deformation
FLAIR FLuid Attenuated Inversion Recovery
FLE Fiducials Localization Evaluation
ICP Iterative Closest Point
IFT Iterative Forest Transform
IM Information Mutuelle
IMN Information Mutuelle Normalisée
IRM Imagerie par Résonnance Magnétique
MIBI Methoxy-IsoButyl-Isonitrile
OS Overall Survival
PDC Points De Contrôle
PFS Progression Free Survival
RANO (Response Assessment in Neuro-Oncology) évaluation de la réponse tumorale en neuro-oncologie
RBF (Radial Basis Function) fonction à bases radiales
RC Ratio de Corrélation
RECIST (Response Evaluation Criteria In Solid Tumors) critères d'évaluation de la réponse au traitement dans les tumeurs solides
SCD Somme des Carrés des Différences
SDA Somme des Différences Absolues
TEMP Tomographie par Emission MonoPhotonique
TEP Tomographie par Emission de Positons
TPS (Thin Plate Spline) fonction à plaques minces

TRE Target Registration Evaluation
UC (Uptake Criteria) critère de fixation

Liste des notations

I_{ref} image de référence
I_{src} image source à recaler
f_1 fonction d'extraction des informations de l'image de référence
f_2 fonction d'extraction des informations de l'image source
M_m Méthode de recalage numéro m
d jeu de données numéro d
UC (Uptake Criteria) critère de fixation

Introduction

L'oncologie est la spécialité médicale qui se consacre au diagnostic et à la prise en charge thérapeutique des cancers. De façon schématique, les cancers se caractérisent par une prolifération cellulaire incontrôlée et anormalement élevée au sein d'un tissu de l'organisme, pouvant entraîner sa mort. Ce mécanisme anormal s'oppose à la prolifération cellulaire contrôlée dans les tissus sains, n'intervenant qu'en cas de pertes cellulaires accidentelles ou liées au vieillissement, dans le but de réparer les tissus endommagés. D'après le site de l'INSERM, les cancers représenteraient la première cause de mortalité en France devant les maladies cardiovasculaires. Il existe un nombre important de types de cancers car cette maladie peut toucher n'importe quel organe comme les poumons, la prostate, le sein, le foie, le cerveau, etc. Au cours de l'évolution de la maladie, les cellules d'une tumeur primaire développée dans un tissu ou un organe peuvent migrer vers une autre région via les voies sanguines ou lymphatiques pour former des tumeurs secondaires appelées métastases. Bien que de nombreux facteurs de risques restent inconnus à ce jour, il est admis que l'apparition des tumeurs est due soit à une cause directement reliée à l'organisme (causes endogènes) comme un gène transmissible de manière héréditaire, soit à des facteurs dits "environnementaux", extérieurs à l'organisme (exogènes) tels que la consommation d'alcool, le tabagisme, l'alimentation, les polluants, les radiations, certains virus, etc. Certains symptômes sont spécifiques à un type de cancer, et facilement identifiables tandis que d'autres sont difficiles à déceler.

Ces symptômes cliniques n'apparaissent souvent qu'une fois que la tumeur a déjà atteint un stade avancé. Pour mieux prendre en charge le patient, il est nécessaire de procéder à un diagnostic précoce. C'est la raison pour laquelle des campagnes de dépistage des cancers sont proposées à la population.

L'imagerie médicale joue un rôle important dans la détection d'une tumeur et, bien évidemment, pour la localiser précisément et faire une première estimation de sa stadification. Les modalités d'imagerie utilisées dépendent de la localisation anatomique et de la nature des symptômes. La mammographie par rayons X constitue l'examen de référence pour les cancers du sein ; l'IRM cérébrale est généralement réalisée pour repérer les tumeurs cérébrales ; le scanner est souvent utilisé pour identifier les cancers des poumons ; l'échographie est utilisée en première intention sur les régions abdominales. Dans la plupart des cas, une biopsie est ensuite indispensable pour établir de façon certaine un diagnostic et décider du traitement à adopter. La biopsie est une étape invasive qui consiste à prélever des échantillons de tissus supposés cancéreux (repérés au préalable à l'aide des examens d'imagerie) dans le but de déterminer la nature des cellules et le stade du cancer dans le cas où celles-ci se révèleraient cancéreuses. La localisation de la tumeur, l'état général du patient ainsi que le stade de la tumeur sont autant de critères à prendre en compte lors de la définition du protocole

de traitement à adopter pour soigner le patient ou ralentir les effets du cancer lorsqu'aucune guérison n'est envisageable. Ce protocole implique généralement l'utilisation d'un ou plusieurs types de traitements, qui seront alors appliqués de façon complémentaire. Dans les cas où le cancer est accessible, l'exérèse chirurgicale de la zone tumorale peut être effectuée, mais l'utilisation d'autres traitements comme la radiothérapie, la chimiothérapie, et dernièrement les traitements anti-angiogéniques sont à considérer.

Dans le but d'optimiser la stratégie thérapeutique en fonction de l'évolution de la tumeur, la mise en place d'un suivi du patient est essentielle. Ce suivi est rendu possible en partie grâce à l'imagerie médicale en réalisant des examens à différents temps au cours du traitement pour juger de l'efficacité de celui-ci. Plusieurs modalités d'imagerie peuvent alors être utilisées conjointement afin de combiner des informations complémentaires d'ordre anatomique et d'ordre fonctionnel de la tumeur.

L'essentiel de mes travaux de thèse a été motivé par un projet de recherche clinique mis en place par les services de neurologie et de médecine nucléaire de l'hôpital de la Pitié-Salpêtrière pour étudier l'apport des examens de tomographie par émission monophotonique (TEMP) au Sestamibi marqué au Technétium-99m en complément de l'imagerie par résonance magnétique (IRM) pour le suivi de patients atteints de gliomes de haut grade, pour lesquels un traitement antiangiogénique supplémentaire a été proposé.

Le chapitre 1 présente le contexte médical des deux applications cliniques considérées, la première exposée ci dessus et la seconde application concernant le suivi de patients atteints de métastases hépatiques. **Le chapitre 1** est principalement orienté sur la premère application concernant le suivi de patients atteints de gliomes, en expliquant dans un premier temps ce que sont les gliomes, quelles sont les thérapies possibles en mettant l'accent sur les traitements antiangiogéniques et quelles sont les modalités d'imagerie qui permettent le suivi des patients. Dans un second temps, certains critères d'évaluation conventionnels de l'évolution tumorale basés essentiellement sur des modalités d'imagerie anatomique sont présentés. Enfin la dernière partie s'attache à présenter le protocole clinique mis en place à la Pitié-Salpêtrière pour déterminer si la modalité TEMP au Sestamibi-Tc99m peut apporter des informations complémentaires de l'IRM pour le suivi de patients atteints de gliomes de haut grade et traités par antiangiogénique.

Afin de répondre à cette problématique de suivi de patients, nous avons identifié puis cherché à résoudre plusieurs problèmes. Premièrement, les images obtenues dans les protocoles de suivi de patients atteints de cancer n'ont pas été acquises aux mêmes dates et peuvent provenir d'appareils différents (multi-modalités), elles nécessitent une étape de recalage consistant à les réaligner afin de pouvoir les analyser conjointement et les comparer localement. Cette étape de recalage peut poser problème surtout dans des cas de recalages multimodaux comme par exemple le recalage TEMP/IRM où les informations des deux images peuvent être très différentes en fonction de la distribution du radiotraceur utilisé dans la modalité TEMP. Dans un second temps, pour réaliser le suivi, il est nécessaire d'identifier des biomarqueurs permettant d'évaluer de manière plus précoce l'évolution tumorale. L'extraction de ces biomarqueurs peut nécessiter des étapes de segmentation et notamment la segmentation de la zone tumorale sur les examens de chacune des modalités d'imagerie impliquées dans le protocole. Or le choix d'une méthode adaptée pour la segmentation de la zone tumorale peut se révéler très complexe en fonction de la modalité d'imagerie utilisée, de la résolution et de

l'environnement autour de la zone à segmenter.

Le chapitre 2 constitue un état de l'art général sur deux domaines essentiels en traitement d'images dans le but de proposer une stratégie d'analyse de la pertinence de biomarqueurs pour le suivi de l'évolution tumorale. : le recalage d'images et la segmentation d'images.

Dans la première partie du chapitre 2, le principe général du recalage est exposé afin de bien comprendre l'impact quant au choix de chaque élément d'un algorithme de recalage, puis plusieurs méthodes conventionnelles sont présentées dont certaines ont été utilisées pour résoudre le problème du recalage multimodal des données TEMP/IRM. En fin de première partie de ce chapitre, des stratégies d'évaluation mettant en œuvre des critères spécifiques aux applications sont décrites afin de comprendre comment choisir une méthode de recalage adaptée aux spécificités des données. La seconde partie du chapitre 2 se concentre sur la présentation de méthodes de segmentation pour extraire la région correspondant aux zones tumorales visibles sur deux modalités : l'IRM pondérée en T1 après injection ou non d'un produit de contraste et la TEMP. Dans le contexte de la segmentation des gliomes sur les images IRM, les méthodes de contours actifs du type snake sont généralement utilisées tandis que sur les images TEMP, ce sont les méthodes de seuillage qui paraissent les plus efficaces.

Le chapitre 3 présente la stratégie que j'ai proposée pour suivre l'évolution tumorale à partir d'examens multi-modalités. Cette approche permet également de comparer les informations provenant de diverses modalités. La stratégie proposée se compose de trois étapes, à savoir, le recalage des données, l'extraction des volumes tumoraux sur les différentes modalités et enfin le calcul d'indices basés sur les segmentations précédentes en vue de quantifier l'évolution tumorale. Cette stratégie générique doit adapter chacune de ses étapes aux spécificités de l'application concernée. Ainsi le recalage d'images médicales constitue un problème complexe qui dépend des spécificités des données impliquées et pour lesquelles de nombreuses solutions ont été proposées. Le choix d'une méthode adaptée au problème posé est critique et peut se révéler complexe. De plus, pour un même type d'images, les performances des différentes méthodes peuvent varier d'un jeu de données à un autre. Ceci m'a amené à proposer une stratégie mettant en œuvre plusieurs méthodes de recalage puis à définir des critères afin de sélectionner automatiquement la meilleure.

Les chapitres 4, 5 et 6 concernent l'application de la stratégie d'évaluation de l'évolution tumorale à l'étude de patients atteints de gliomes de haut grade et sous traitement antiangiogénique.

Le chapitre 4 décrit comment la stratégie d'analyse a été adaptée aux données TEMP/IRM pour le suivi des patients atteints de gliomes.

Dans un premier temps, nous avons appliqué la stratégie de recalage à deux types de recalage : le recalage multimodal des données TEMP/IRM, et le recalage longitudinal des données IRM. Pour faciliter la sélection de la meilleure méthode parmi un ensemble de méthodes classiques testées, un critère qui repose sur les fixations physiologiques du radiotraceur utilisé (Sestamibi) a été développé. La mise en œuvre de ce critère demande notamment la segmentation de structures anatomiques particulières comme les globes oculaires, les

muscles oculomoteurs et l'hypophyse. Dans un second temps, les méthodes proposées pour la segmentation des volumes tumoraux sur les images TEMP et IRM sont présentées. Une méthode de levelset 2D a été appliquée pour la segmentation des données IRM tandis que des approches de seuillage ont été testées sur les données TEMP. Enfin, des critères globaux et locaux, s'appuyant sur les données TEMP/IRM segmentées et recalées, sont définis pour l'évaluation de l'évolution tumorale et la comparaison des deux modalités.

Le chapitre 5 présente les résultats des stratégies de recalage décrites au chapitre 4 sur la base de données disponible. Il est montré, dans un premier temps, que les critères d'évaluation proposés sont pertinents pour sélectioner la meilleure méthode. Les résultats obtenus en applicant la stratégie sont comparés aux résultats obtenus par les meilleures méthodes montrant ainsi l'intérêt de l'approche proposée.

Le chapitre 6 présente les résultats des indices globaux et locaux calculés à partir des segmentations des zones tumorales extraites sur les images TEMP et IRM. Des études de corrélation impliquant ces indices ont permis de comparer les différentes méthodes de segmentation utilisées en imagerie TEMP pour extraire la zone tumorale, de comparer les images TEMP précoces et tardives et enfin de comparer les images des modalités TEMP et IRM afin de déterminer s'il existait une quelconque complémentarité entre ces deux modalités. Enfin les valeurs de ces indices pour les deux modalités ont été comparées aux indices de durée de survie PFS et OS pour déterminer si les informations extraites de l'une ou l'autre des modalités pouvaient constituer un biomarqueur de l'évolution tumorale. PFS correspond à la durée entre le début du traitement et l'aggravation du cancer alors que l'OS correspond à la durée entre le début du traitement et le décès du patient.

Le chapitre 7 illustre l'application de la stratégie de recalage dans le contexte de suivi de métastases hépatiques grâce à des examens TDM dynamiques. Deux types de recalage ont été mis en œuvre : le recalage dynamique des séquences TDM et le recalage longitudinal des examens TDM. Le critère d'évaluation proposé pour la sélection de la meilleure méthode de recalage correspond à un indice de recouvrement basé sur des structures osseuses comme les côtes et les vertèbres.

Une synthèse des travaux réalisés et des résultats obtenus par la stratégie d'analyse est présentée en fin de manuscrit. La perspective d'optimisation de la stratégie de recalage est discutée en proposant des solutions pour diminuer le nombre de méthodes à tester pour chaque jeu de données. Enfin l'application de la stratégie d'analyse de l'évolution tumorale à d'autres études cliniques est proposée.

Chapitre 1

Contexte clinique

Sommaire

1.1 Suivi de patients atteints de gliomes de haut grade et sous traitement anti-angiogénique . . . 12

1.1.1 Les gliomes de haut grade . . . 13

1.1.2 Les traitements conventionnels et le traitement antiangiogénique . . . 14

1.1.3 Modalités d'imagerie médicale utlisées pour le diagnostic et le suivi . . . 15

1.1.4 Critères d'évaluation de l'évolution tumorale . . . 16

1.1.5 Etude clinique sur les glioblastomes . . . 21

1.2 Etude clinique sur le suivi des métastases hépatiques . . . 22

1.3 Conclusion . . . 23

Le but de ce premier chapitre est d'introduire le contexte clinique des deux applications sur lesquelles nous avons travaillé. La première application concerne le suivi de patients atteints de gliomes de haut grade et sous traitement antiangiogénique et la seconde, le suivi de patients atteints de métastases hépatiques. Le contexte clinique de la seconde application est décrit de manière succinte. En dernière partie de ce chapitre, sont décrits des critères conventionnels pour évaluer l'évolution tumorale.

1.1 Suivi de patients atteints de gliomes de haut grade et sous traitement antiangiogénique

Ce premier projet a été mis en place dans le cadre d'une collaboration entre les services de neurologie, neuro-radiologie et de médecine nucléaire de la Pitié- Salpêtrière dans le but d'améliorer la prise en charge et le suivi de patients atteints de gliomes de haut grade. L'objectif était d'évaluer l'apport de la tomodensitométrie à émission monophotonique (TEMP), examen réalisé après l'injection de Sestamibi marqué au Technétium 99m, en complément des images de résonance magnétiques (IRM).

1.1.1 Les gliomes de haut grade

Les gliomes sont des tumeurs cérébrales issues des cellules gliales qui constituent un tissu de soutien des neurones. Il existe plusieurs types de cellules gliales, les plus importantes étant les astrocytes, les oligodendrocytes et les cellules microgliales. Les astrocytes, cellules en forme d'étoile assurent notamment l'approvisionnement nutritionel et le métabolisme énergétique des neurones et ont un rôle important dans la cicatrisation du cerveau à la suite d'une lésion. Elles participent également au maintien de la barrière hémato-encéphalique. Les oligodendrocytes jouent quant à eux un rôle important dans la formation de la myéline qui constitue une gaine protectice entourant les fibres nerveuses et permettant de faciliter la conduction de l'influx nerveux. L'organisation mondiale de la santé (OMS) a proposé une classification basée sur des caractéristiques histologiques des tumeurs du système nerveux central [1] en 4 grades corrélés à leur degré de malignité :

- Grade I : ce sont des tumeurs bénignes à croissance lente, n'impliquant aucune néovascularisation ni nécrose.
- Grade II : ce sont des tumeurs bénignes à croissance lente mais avec une division cellulaire plus rapide. Les oligoastrocytomes, oligodendrogliomes et astrocytomes sont des gliomes de grade II.
- Grade III : ce sont des tumeurs malignes impliquant une division cellulaire très rapide. Ces tumeurs proviennent généralement de l'évolution de gliomes de grade inférieur et évoluent vers des grades supérieurs. Les oligodendrogliomes anaplasiques, les oligodendrocytomes anaplasiques et les astrocytomes anaplasiques sont des gliomes de grade III.
- Grade IV : ce sont des tumeurs malignes à évolution très rapide impliquant une néovascularisation et des nécroses. Les glioblastomes multiformes sont les seuls gliomes de grade IV.

Les glioblastomes multiformes constituent non seulement les gliomes de plus haut grade mais également le type de gliomes le plus fréquent (plus de 50% chez l'adulte avec un taux de l'ordre de 3 cas sur 100 000 personnes par an [2]). On distingue deux types de glioblastomes multiformes, les glioblastomes primaires (de novo) et les glioblastomes secondaires qui proviennent de l'évolution de gliomes de grade inférieur pour lesquels des études récentes ont découvert des différences génétiques [3, 4]. Les glioblastomes primaires ont plutôt tendance à toucher des populations agées de plus de 50 ans alors que les glioblastomes secondaires touchent des populations plus jeunes [5].

Bien que l'incidence des gliomes semble être à la hausse [6], les causes d'apparition des gliomes ne sont pas aujourd'hui clairement établies. De nombreux facteurs environnementaux ont été suggérés, comme l'exposition à des produits chimiques cancérigènes dans des professions liées à la fabrication de caoutchouc, la production de pétrole, l'exposition au chlorure de vinyle, aux pesticides, les travaux forestiers et les services de nettoyage, l'exposition au tabagisme passif ainsi que l'utilisation de téléphone portable, même si aucun de ces facteurs n'est une cause réellement établie [7, 8, 9]. Le seul facteur de risque environnemental prouvé est l'exposition aux rayonnements ionisants, comme cela peut être le cas pour des patients atteints de leucémie et traités par radiothérapie [10]. Certaines études ont mis en évidence que des mutations dans les gènes des enzymes IDH1 et IDH2 étaient fréquentes dans les gliomes, suggérant une cause potentielle non environnementale des gliomes [4].

1.1.2 Les traitements conventionnels et le traitement antiangiogénique

Bien que les thérapies proposées aux patients dépendent de nombreux facteurs tels que la localisation de la tumeur et le degré de malignité, les traitements conventionnels pour les gliomes comprennent la chirurgie, la radiothérapie, et éventuellement la chimiothérapie en cas de récidive tumorale.

En premier lieu, une intervention chirurgicale consistant en l'ablation d'une partie de la tumeur la plus large possible est généralement effectuée (après une biopsie ayant permis de confirmer le diagnostic) si le gliome ne se trouve pas trop en profondeur. Cependant, l'ablation complète du glioblastome n'est généralement pas possible du fait du caractère infiltrant de la tumeur dans le parenchyme rendant l'intervention chirurgicale impossible sans causer de lésions cérébrales importantes. Certaines études ont montré qu'une ablation de plus de 98% du volume tumoral permettait de doubler le taux de survie avec un maximum de 11 à 12 mois par rapport à une biopsie seule en l'absence d'exérèse chirurgicale [11, 12].

En deuxième intention, après un délai de 2 semaines après la chirurgie pour laisser guérir la blessure, une radiothérapie est appliquée sur le lit tumoral. La radiothérapie est généralement répartie sur 30 séances (environ 6 semaines) avec des doses irradiantes de 2 Gy et peut permettre de prolonger le taux de survie de 9 à 10 mois après chirurgie [13]. En complément de la radiothérapie et en cas de récidive tumorale, une chimiothérapie peut être proposée avec notamment de la Témozolomide (Temodal) (qui est une molécule autorisée en France depuis mai 2000)[14, 15].

Ces 3 thérapies constituent aujourd'hui une ligne de traitement conventionnelle utilisée en routine clinique et dont le bénéfice a été démontré. Récemment de nouveaux traitements s'attaquant à des mécanismes d'évolution de la tumeur comme l'angiogénèse ont été suggérés et testés.

L'angiogénèse correspond à la formation de nouveaux vaisseaux sanguins à partir de vaisseaux préexistants (néovascularisation). Ce processus est régulé par des signaux chimiques proangiogéniques (VEGF), qui peuvent stimuler la réparation de vaisseaux endommagés ou la formation de nouveaux vaisseaux, et les signaux inhibiteurs de l'angiogénèse qui interfèrent avec la formation de nouveaux vaisseaux sanguins. Normalement, les effets de stimulation et d'inhibition de ces signaux chimiques sont équilibrés de manière à ce que de nouveaux vaisseaux sanguins ne se forment qu'en cas de besoin, comme pour une cicatrisation. L'angiogénèse correspond donc à un processus naturel, cependant ce phénomène joue un rôle crucial dans le développement et la propagation du cancer. En effet, pour se développer et atteindre des volumes supérieurs à quelques millimètres, les tumeurs nécessitent un appport plus important en oxygène et nutriments comme illustré par la figure 1.1. Pour ce faire, les tumeurs produisent des molécules de signalisation proangiogéniques telles que le facteur de croissance vasculaire (VEGF) pour stimuler les cellules endothéliales, qui tapissent les parois des vaisseaux sanguins existants, pour qu'elles rompent la membrane basale endothéliales existante dans un premier temps puis prolifèrent et forment la membrane basale du nouveau vaisseau.

Pour stimuler l'angiogénèse, il est nécessaire que les molécules de signalisation comme le VEGF se lient aux récepteurs des cellules endothéliales. Pour limiter l'angiogénèse, des inhibiteurs ont été proposés. Par exemple, le bevacizumab est un anticorps monoclonal qui se lie spécifiquement au VEGF ; une fois cette liaison effectuée, le VEGF ne peut plus activer les récepteurs des cellules endothéliales. D'autres inhibiteurs comme le sunitinib et le sorafénib se lient aux récepteurs en surface des cellules endothéliales ou à d'autres protéines en amont permettant de bloquer l'angiogénèse. Dans le cadre du traitement des glioblastomes, le

bevacizumab a été approuvé par l'agence américaine des produits alimentaires et médicamenteux (FDA : Food and Drug Administration).

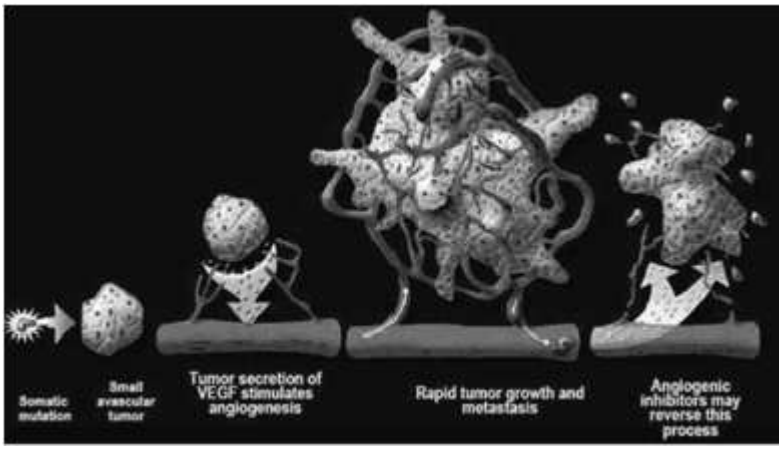


FIGURE 1.1: Principe de l'angiogénèse et action du traitement antangiogénique [A. Griffoen Laboratoire (VUMC)].

Malgré les progrès des nouvelles thérapies, le pronostic de patients atteints de gliomes de haut grade reste péjoratif [16]. L'étude d'Ohgaki et al. [17] rapporte 17 à 30% de survie à 1 an et moins de 5% à 2 ans.

De plus, ces nouveaux traitements comme les antiangiogéniques induisent des effets secondaires importants comme des saignements, des caillots dans les artères et de l'hypertension. L'évaluation précoce de la réponse tumorale est donc essentielle pour sélectionner la stratégie de traitement la plus appropriée au patient. La neuroimagerie structurelle et fonctionnelle permet l'évaluation des gliomes de manière non invasive et se révèle cruciale pour le suivi des patients.

1.1.3 Modalités d'imagerie médicale utlisées pour le diagnostic et le suivi

Les symptômes les plus communs chez les patients atteints de glioblastomes sont les maux de tête, les déficits neurologiques ainsi que des crises de convulsions. Même si les progrès de l'imagerie médicale permettent de détecter davantage de tumeurs bénignes asymptomatiques, la détection des glioblastomes est le plus souvent fortuite. L'examen de routine clinique utilisé pour le diagnostic et le suivi des gliomes est l'imagerie par résonance magnétique incluant des séquences pondérées en T1 sans et avec agent de contraste (chélates de Gadolinium), ainsi que des séquences pondérées en T2 et "FLAIR". Dans le cas où les patients ne peuvent bénéficier d'IRM, l'examen de tomodensitométrie avec produit de contraste iodé est utilisé, cependant il est inférieur à l'IRM et il a l'inconvénient d'être irradiant.

Sur les images IRM pondérées en T1 suite à l'injection d'un produit de contraste, les gliomes apparaissent en hypo ou iso-intense avec un anneau de réhaussement de Gadolinium en périphérie. Sur les examens IRM pondérés en T2 et sur les examens FLAIR les gliomes apparaissent en hyper-intense comme le montre la figure 1.2. Des critères conventionnels ont été définis à partir d'informations extraites de ces modalités pour l'évaluation de l'évolution tumorale.

Cependant l'IRM a un rôle limité pour différencier les modifications bénignes induites par le traitement, tels que l'œdème ou les radionécroses, des récidives tumorales [18]. La spectroscopie par résonance magnétique (SRM) ou des examens de médecine nucléaire peuvent se révéler être des modalités complémentaires afin d'étudier le métabolisme de ces tumeurs. Comme le Sestamibi marqué au Technetium 99m (MIBI) s'accumule dans les tumeurs malignes [19], la TEMP s'appuyant sur ce traceur pourrait permettre de fournir des informations complémentaires à celles apportées par l'IRM. Cette hypothèse a été suggérée dans une étude sur le suivi de patients atteints de gliomes de haut grade ayant subi une chimiothérapie [20, 21]. Ce traceur a permis notamment de distinguer les radionécroses des récidives tumorales quelle qu'en soit l'origine, les tumeurs apparaissent en hypo signal sur les IRM T1 suite à l'injection d'un produit de contraste. Sur les TEMP au Sestamibi les récidives tumorales apparaissent en hyper signal et les radionécroses en hypo signal.

Afin de quantifier l'évolution tumorale, différents critères extraits des modalités d'imagerie ont été proposés. Ceux ci sont essentiellement basés sur les modalités IRM bien que d'autres modalités puissent apporter une information complémentaire comme nous venons de le signaler.

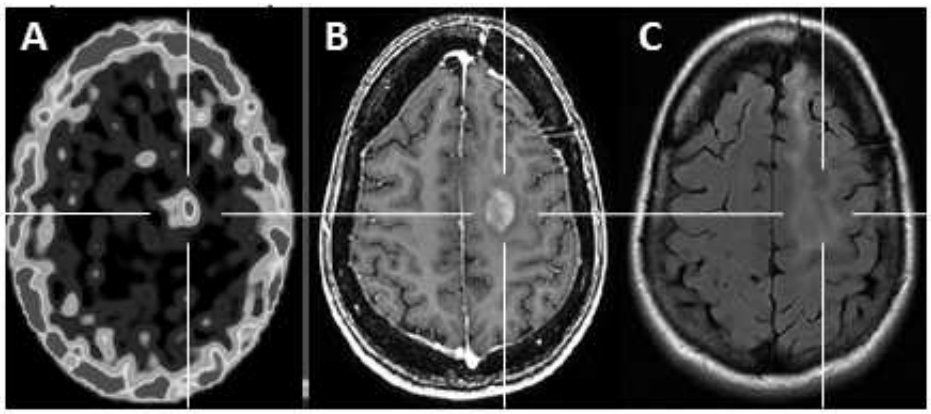


FIGURE 1.2: Détection d'un gliome sur plusieurs modalités d'imagerie médicale. Les gliomes apparaissent en A) hyper intense sur les TEMP au Sestambi marqué au Technétium-99m, B) en iso intense avec un anneau de réhaussement sur les IRM T1 suite à l'injection de chélates de Gadolinium, et C) en hyper intense sur les IRM FLAIR [22].

1.1.4 Critères d'évaluation de l'évolution tumorale

La question de l'évaluation de l'évolution tumorale pour le suivi de patients est une question cruciale pour laquelle plusieurs critères basés essentiellement sur les modalités IRM (séquence T1 suite ou non à l'injection d'un produit de contraste, séquences FLAIR) ont été proposés et sont aujourd'hui classiquement utilisés [23].

1.1.4.1 Critères RECIST

Dans le but de simplifier et standardiser les critères WHO [24], basés notamment sur une mesure bidimensionnelle des lésions, proposés en 1981 pour l'évaluation de la réponse tumorale par l'organisation mondiale

de la santé, une collaboration internationale regroupant la "European Organisation for Research and Treatment of Cancer" (EORTC), le "National Cancer Institute" aux Etats Unis et le "National Cancer Institute" du Canada ont publié en 2000 les critères RECIST [25]. Ces critères permettant d'évaluer l'évolution de la réponse tumorale suite à un traitement dans des tumeurs solides ont été affinés en 2009 [26]. Ils incluent notamment les définitions de la taille minimale des lésions mesurables et du nombre maximum de lésions à considérer pour la mesure. L'ensemble de ces règles est décrit ci dessous :

Les critères RECIST nécessitent de sélectionner les lésions et les ganglions pour lesquels on peut mesurer de manière précise la dimension (diamètre) la plus longue dans un plan de vue, et cette mesure doit être supérieure à une valeur donnée. Cette mesure sera considérée par la suite comme la taille de la lésion (respectivement du ganglion). Pour qu'une lésion soit mesurable, sa dimension la plus longue doit être :

- supérieure ou égale à 10 mm au scanner ou supérieur ou égal à 2 fois la résolution dans la direction spatiale de mesure, si cette résolution spatiale est supérieure à 5 mm.
- supérieure ou égale à 20 mm par radiographie du thorax

Pour qu'un ganglion lymphatique soit mesurable, celui-ci doit avoir un petit axe supérieur ou égal à 15 mm, sachant que le petit axe correspond à l'axe perpendiculaire à la plus grande dimension du ganglion.

Les autres lésions et ganglions seront considérés comme non mesurables.

On effectue ensuite une deuxième sélection consistant à ne garder que 5 lésions que l'on nommera "lésions cibles" avec un maximum de 2 lésions cibles par organe parmi les lésions mesurables. En général, on choisit les lésions de dimensions importantes dont on pourra suivre l'évolution. La même procédure est utilisée pour les ganglions. La somme des diamètres de ces lésions et ganglions cibles permettra par la suite d'évaluer la réponse tumorale. Bien que n'entrant pas dans le calcul de la somme des diamètres, les autres lésions mesurables mais non-cibles seront également prises en compte dans l'évaluation de la réponse tumorale.

A partir de ces mesures, la réponse tumorale (ou réponse de la tumeur au traitement) sur les lésions cibles est alors classifiée dans l'une des quatre catégories suivantes :

- Réponse complète (Complete Response (CR)) : si toutes les lésions cibles ont disparu et si la taille de tous les ganglions est $\leq$ 10 mm dans leur petit axe.
- Réponse partielle (Partial Response (PR)) : si une diminution supérieure ou égale à 30% de la somme des diamètres des lésions cibles est constatée sur l'examen après traitement le plus récent par rapport à l'examen avant traitement.
- Progression (Progressive Disease (PD)) : si une augmentation supérieure ou égale à 20% de de la somme des diamètres des lésions cibles est constatée sur l'examen après traitement le plus récent par rapport à la plus petite somme des lésions cibles observée durant l'étude (Nadir). Cette augmentation doit dépasser 0,5 cm.
- Stabilisation (Stable Disease (SD)) : tous les autres cas.

Remarque : si on observe une réponse partielle par rapport à l'examen avant traitement mais une progression par rapport à l'un des examens de l'étude (Nadir : l'examen Nadir correspond à l'examen sur lequel on a constaté la meilleure réponse au traitement), c'est l'information de progression qui est retenue.

La réponse tumorale sur les lésions non-cibles est classifiée en 3 catégories :

Lésions cibles	Lésions non cibles	Nouvelles lésions	Réponse tumorale globale
CR	CR	Non	CR
CR	Non-CR/Non-PD	Non	PR
CR	Non évaluées	Non	PR
PR	Non-PD ou Non évaluées	Non	PR
SD	Non-PD ou Non évaluées	Non	SD
Non évaluées	Non-PD	Non	Non évaluée
PD	/	/	PD
/	PD	/	PD
/	/	Oui	PD

TABLE 1.1: Réponse tumorale globale. **"CR"** Réponse complète ; **"PR"**, Réponse partielle ; **"SD"**, Stabilisation ; **"PD"**, Progression.

- CR : si toutes les lésions non-cibles ont disparu et si la taille de tous les ganglions non-cibles n'excède pas 10 mm dans leur petit axe.
- SD : s'il reste au moins une lésion non-cible.
- PD : si les lésions non cibles ont une forte augmentation en taille ou si de nouvelles lésions apparaissent.

Les critères RECIST sont entièrement basés sur des modalités anatomiques, cependant d'autres modalités peuvent apporter des informations complémenatires. Afin de mieux caractériser la présence de nouvelles lésions, la TEP au fluorodéoxyglucose (FDG) marqué au fluor 18 a été suggérée en complément de l'imagerie anatomique. Si une hyper-fixation de FDG apparait au cours du traitement, cela signifie qu'il y a progression de la maladie. Dans le cas d'absence d'examen TEP avant traitement mais d'observation d'une fixation de FDG durant le suivi, trois cas de figure se présentent :

- si le lieu de fixation du FDG correspond à une nouvelle lésion visible sur l'examen anatomique, on considère qu'il y a progression.
- si le lieu de fixation du FDG n'apparaît pas sur l'examen anatomique en tant que nouvelle lésion, un nouvel examen anatomique est réalisé pour déterminer s'il y a réellement progression.
- si le lieu de de fixation du FDG correspond à une lésion pré-existante visible sur l'examen anatomique mais qui a disparu lors du suivi, on considère qu'il n'y a pas de progression.

A partir des classications des lésions cibles et non-cibles on peut définir la réponse tumorale globale comme le montre le tableau 1.1.

1.1.4.2 Critères MacDonald et RANO

Pour faire face aux spécificités des tumeurs cérébrales, les critères de MacDonald [27] sont devenus des critères de référence pour l'évaluation de la réponse thérapeutique pour des patients atteints de gliomes. Ils s'appuient sur une mesure 2D de la taille de la zone de réhaussement de contraste détectée sur l'image anatomique et caractérisant la zone tumorale, couplée avec l'état clinique et les doses de corticostéroïdes données aux patients. Cependant, avec l'apparition de traitements anti-angiogéniques qui sont censés bloquer la néovascularisation, ces critères peuvent se montrer insuffisants. En effet, il se peut que sur un examen IRM pondéré en T1 avec injection de gadolinium, la prise de constraste diminue sans refléter pour autant un véritable effet antitumoral. En effet, pour certains patients, malgré une diminution de contraste dans une zone de l'image IRM pondérée en T1 après injection de gadolinium, cette même zone peut apparaître en hypersignal sur des IRM pondérées en T2 ou sur des séquences en inversion-récupération pondérées en T2 (FLuid Attenuated Inversion Recovery (FLAIR)), allant à l'encontre de l'hypothèse de diminution de la tumeur.

C'est pourquoi les critères RANO [28] ont été définis, pour affiner les critères de MacDonald en rajoutant les modalités FLAIR/T2 dans le protocole d'évaluation des gliomes.

Les séquences requises pour les critères RANO sont :

- les séquences IRM pré-contraste pondérées en T1 et pondérées en T2/FLAIR ,
- les séquences IRM post-contraste pondérées en T1, avec au moins 2 plans orthogonaux (ou l'acquisition d'un volume).

La séquence FLAIR correspond à une séquence IRM pondérée en T2 dans laquelle le signal provenant du liquide céphalo-rachidien (LCR) est supprimé. L'épaisseur des coupes doit de préférence être inférieure à 5 mm.

Les critères RANO se définissent comme indiqué ci-dessous :

On commence par sélectionner les lésions mesurables. Sont considérées comme mesurables, les lésions qui apparaissent réhaussées sur l'IRM pondérée en T1 après injection de gadolinium et dont 2 diamètres perpendiculaires de la lésion (définissant sa taille) sont supérieurs à 10 mm (ou supérieurs à 2 fois l'épaisseur de coupe si celle-ci est supérieure à 5 mm). Les autres lésions, dont la taille est trop petite ou qui apparaissent seulement en T2/FLAIR sont considérées comme non-mesurables.

Parmi ces lésions mesurables sur l'IRM, on va sélectionner 5 lésions cibles au maximum. De préférence, les lésions les plus grandes et les plus faciles à reproduire sont choisies. On calcule ensuite le produit des diamètres maximaux de chaque lésion cible et on les ajoute pour former la somme des produits des diamètres (SPD) comme indiqué sur la figure 1.3.

Comme pour les critères RECIST, la réponse des lésions cibles peut être classée en 4 catégories en fonction de leur taille :

- Réponse complète (Complete Response (CR)) : si toutes les lésions cibles ont disparu.
- Réponse partielle (Partial Response (PR)) : si une diminution supérieure ou égale à 50% de la somme des produits des diamètres des lésions cibles (SPD) est constatée sur l'examen après traitement le plus récent par rapport à l'examen avant traitement.

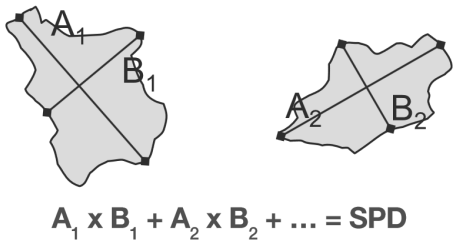


FIGURE 1.3: Calcul de la somme des produits diamètres (SPD) dans les critères RANO. (Image du site www.iconplc.com)

- Stabilisation (Stable disease (SD)) : si la diminution de la SPD entre l'examen avant traitement et celui après traitement est inférieure ou égale à 50% ou si l'augmentation de SPD est supérieure ou égale à 25%.
- Progression (Progression Disease (PD)) : si une augmentation supérieure ou égale à 25% de de la somme des diamètres (SPD) des lésions cibles sur l'examen après traitement le plus récent est constatée par rapport à la plus petite observée durant l'étude (Nadir).

La réponse complète (CR) ou la réponse partielle (PR) doivent être confirmées 4 semaines plus tard, sinon la réponse est considérée comme une stabilisation (SD). En effet, dans le cas d'un traitement anti-angiogénique, les effets du traitement peuvent ressembler à une réponse au traitement alors que le cancer continue de se développer. Ce phénomène s'appelle "pseudo-réponse". D'autre part, dans les 12 semaines après la fin d'une radiothérapie, il se peut qu'on ait une impression de grossissement de lésions existantes ou d'apparition de nouvelles lésions [29]. Ce phénomène peut être dû au traitement plutôt qu'au développement de la tumeur et s'appelle "pseudo-progression". Si les lésions continuent à grossir, le grossissement initialement observé correspond à une vraie progression. Dans le cas où les lésions se stabilisent ou réduisent, il s'agit d'une pseudo-progression et la lésion concernée ne doit plus être incluse dans le calcul de la somme des diamètres.

Pour distinguer une pseudo-progression d'une vraie progression, l'intérêt de l'imagerie par IRM de diffusion et de la spectroscopie sont à l'étude [30].

En ce qui concerne les lésions non cibles, on peut en trouver 2 types : les lésions non mesurables qui se réhaussent en T1 après injection de gadolinium, et les lésions uniquement visibles en T2/FLAIR. La réponse de ces lésions peut être classée en trois catégories :

- Réponse complète (Complete Remission (CR)) : si toutes les lésions non-cibles ont disparu ;
- Stabilisation (Stable disease (SD)) : si des lésions non cibles sont présentes ; si leur taille diminue ou reste stable ;
- Progression (Progression Disease (PD)) : si leur taille augmente notablement.

La réponse des lésions visibles uniquement sur les séquences T2/FLAIR peut être :

- Améliorée : si le signal anormal d'intensité importante diminue ;
- Inchangée : si le signal reste le même par rapport à une image antécédente ;

– Dégradée : si le signal anormal d'intensité importante augmente.

A partir des réponses des lésions cibles et non-cibles, on peut établir la réponse globale comme le montre le tableau 1.2 :

	CR	PR	SD	PD
T1-Gd	Rien	$\downarrow \geq 50\%$	$\downarrow \geq 50\%$ et $\uparrow \leq 50\%$	$\uparrow \geq 50\%$
T2/FLAIR	Inchangée ou Améliorée	Inchangée ou Améliorée	Inchangée ou Améliorée	Dégradée
Nouvelle lésions	Non	Non	Non	Oui
Corticostéroïdes	Non	Stable ou $\downarrow$	Stable ou $\downarrow$	/
Etat clinique	Stable ou $\uparrow$	Stable ou $\uparrow$	Stable ou $\uparrow$	$\downarrow$
besoin d'éléments pour la réponse	Tous	Tous	Tous	Un seul

TABLE 1.2: Réponse tumorale globale. Explication des critères RANO. **"CR"** Réponse complète ; **"PR"**, Réponse partielle ; **"SD"**, Stabilisation ; **"PD"**, Progression ; sur la ligne T1-Gd $\downarrow$ signifie diminution et $\uparrow$ signifie augmentation de la taille des lésions entre 2 examens ; sur la ligne Etat clinique $\downarrow$ signifie que l'état clinique est dégradé et $\uparrow$ signifie que l'état clinique est amélioré.

Par exemple, pour que la réponse globale soit une réponse complète, il faut que tous les éléments de la première colonne du tableau soient vérifiés : aucune lésion ne doit être visible sur l'IRM T1-Gd, la réponse sur la T2/FLAIR est inchangée ou améliorée, il n'y a pas de nouvelles lésions, il n'y a pas besoin de corticostéroïdes et l'état l'état clinique est soit stable soit amélioré.

1.1.5 Etude clinique sur les glioblastomes

L'évaluation précoce de la réponse tumorale s'avère primordiale pour le choix d'une stratégie de traitement adapté aux patients. Afin de mener cette évaluation, plusieurs modalités complémentaires peuvent être nécessaires. C'est pourquoi l'étude clinique monocentrique ayant débuté en avril 2010 et résultant de la collaboration entre les services de neuro-oncologie (J-Y. Delattre, G. Petrirena), de médecine nucléaire (M-O. Habert, A. Kas, N. Yeni), et de neuro-radiologie (R. Guillevin) du CHU de la Pitié-Sapêtrière a eu pour objectif, de tester si la TEMP au Sestamibi-Tc99m, en complément de la modalité d'imagerie IRM, peut permettre une évaluation thérapeutique chez des patients atteints de gliomes de haut grade recevant un traitement anti-angiogénique.

Le protocole mis en place dans cette étude a été appliqué à seize patients atteints de gliomes de haut grade sous traitement anti-angiogénique. Avant le début du traitement antiangiogénique, chaque patient a subi une IRM pondérée en T1 obtenue après l'injection de gadolinium-DOTA acquise sur un appareil Signa Hdxt 3T (GE), deux examens TEMP respectivement réalisés 15 min et 3h après l'injection de Sestamibi 99mTc (MIBI) sur un tomographe Irix (Philips). Les examens IRM et TEMP ont été réalisés à moins de 72 heures d'écart. Le même protocole a été répété un mois plus tard, pour une évaluation des volumes tumoraux

après le début du traitement.

Idéalement l'analyse doit être faite après recalage des images fonctionnelles (TEMP) sur les images anatomiques obtenues par IRM pour combiner les informations morphologiques et fonctionnelles. Pour suivre l'évolution de la réponse tumorale, les examens réalisés un mois après le début du traitement anti-angiogénique doivent également être recalés sur les examens réalisés avant le début du traitement. Cette dernière étape peut être effectuée en recalant les examens anatomiques IRM entre eux, puis en appliquant la transformation trouvée aux images TEMP préalablement recalées sur les IRM respectives. Une fois ces recalages effectués, la segmentation des zones tumorales sur chaque modalité sera nécessaire pour comparer les informations des différentes modalités et évaluer la réponse tumorale au traitement antiagiogénique.

1.2 Etude clinique sur le suivi des métastases hépatiques

Durant ces dernières années, différents traitements anticancéreux ont été développés dans le but d'améliorer la prise en charge du cancer. Dans ce contexte, de plus en plus de patients atteints de métastases hépatiques sont traités avec des traitements anti-angiogéniques [31]. Ces médicaments visent à réduire le développement des vaisseaux sanguins autour des tumeurs et ainsi à réduire voire stopper leur approvisionnement en oxygène et en nutriments. L'évaluation précoce de la réponse tumorale à un traitement anti-angiogénique est cruciale pour décider de poursuivre ou de changer de traitement compte-tenu du coût de ces thérapies et de leurs effets secondaires. Des critères définis au niveau international pour l'évaluation de l'évolution des tumeurs solides ont été proposés comme les critères RECIST (Response Evaluation Criteria in Solid Tumors) [26]. Ces caractéristiques sont basées sur des critères anatomiques et plus précisément sur des mesures de taille de la tumeur. Cependant, les critères RECIST sont purement morphologiques et ne tiennent pas compte de l'hétérogénéité tumorale et des changements physiologiques, tels que la modification de la vascularisation, qui peuvent survenir plus tôt que les changements de taille dans la réponse au traitement. Pour surmonter ces limitations, une adaptation des critères RECIST, appelée critères RECIST modifiés, a été conçue spécifiquement pour le carcinome hépatocellulaire, où la lésion cible n'est plus la globalité de la lésion tumorale, mais seulement la partie avec réhaussement de contraste de la lésion sur des données dynamiques de perfusion [32]. La tomodensitométrie (TDM) est largement utilisée pour évaluer la réponse des tumeurs hépatiques à un traitement, en raison de sa grande disponibilité dans un environnement clinique.

Pour une évaluation plus objective et quantitative de la thérapie, les volumes d'examens TDM pré et post-traitement doivent être recalés avant d'évaluer la réponse tumorale. Il s'agit d'une étape nécessaire pour accéder à des déformations locales de la tumeur en réponse au traitement et pour comparer les modifications de paramètres fonctionnels comme la perfusion. Pour tenir compte des déformations locales, l'utilisation de méthodes de recalage non rigides comme l'algorithme de Free Form Deformation (FFD) semble pertinente [33]. Cependant ces algorithmes non rigides nécessitent généralement une bonne initialisation par une méthode de recalage rigide pour fournir des résultats satisfaisants.

Le premier objectif de cette étude a été de trouver la meilleure initialisation possible par recalage rigide pour chaque jeu de données permettant un recalage des données longitudinales.

En raison de la durée de chaque examen de perfusion (2 à 5 min), les acquisitions ont été réalisées en

respiration libre. Il est alors nécessaire de corriger ces mouvements respiratoires sur chacun des examens. Le second problème de cette étude a concerné la compensation de ces mouvements respiratoires.

1.3 Conclusion

Ce premier chapitre a présenté le contexte clinique de deux applications de suivi de patients en oncologie, le suivi de patients atteints de gliomes de haut grade sous traitement antiangiogénique et le suivi de patients atteints de métastases hépatiques, mettant en évidence l'importance d'établir des biomarqueurs précoces de l'évolution tumorale afin d'adapter le traitement en conséquence. Dans le cadre du suivi en oncologie, l'imagerie médicale joue un rôle prépondérant. L'utilisation conjointe de modalités d'imagerie anatomique et fonctionnelle peut se révéler d'une grande aide dans le suivi des patients, cependant leur exploitation requiert certaines étapes pour profiter pleinement de la complémentarité des informations des différentes modalités. Ainsi le recalage des données, de même que la segmentation des volumes tumoraux se révèlent être des étapes nécessaires. Le prochain chapitre s'attache , dans un premier temps, à présenter un état de l'art sur le recalage d'images afin d'identifier la stratégie à adopter pour résoudre les problèmes du recalage longitudinal et du recalage de données provenant de modalités différentes ; puis, dans un second temps, s'intéresse au domaine de la segmentation d'images afin de déterminer quelle méthode utiliser pour extraire les zones tumorales de chacune des modalités.

Chapitre 2

Etat de l'art

Sommaire

2.1 Recalage : méthodologie **25**

2.1.1 Les approches géométriques 26

2.1.2 Les approches iconiques 28

2.1.3 Les approches hybrides 33

2.1.4 Les modèles de transformation 34

2.1.5 Les méthodes d'interpolation 38

2.1.6 La stratégie d'optimisation 40

2.2 Evaluation du recalage **43**

2.2.1 Critères d'évaluation 44

2.2.2 Estimation d'une transformation de référence ("Gold Standard") 50

2.3 Segmentation **51**

2.3.1 Méthodologie 51

2.3.2 Segmentation des gliomes sur les IRM 55

2.3.3 Segmentation des gliomes sur les données TEMP 57

2.3.4 Validation des méthodes de segmentation 59

2.4 Perspectives pour l'évaluation tumorale **60**

2.5 Conclusion **62**

Pour identifier et/ou conclure sur la pertinence de biomarqueurs de l'évolution tumorale dans le cadre du suivi de patients en oncologie, il nous a semblé pertinent d'organiser ce chapitre d'état de l'art en 2 grandes parties qui sont essentielles à ce type d'analyse :

- le recalage d'images pour réaligner les examens réalisés à différents temps du traitement ou provenant de différentes modalités ;
- la segmentation pour extraire les volumes tumoraux de chaque modalité d'image

2.1 Recalage : méthodologie

La première partie de ce chapitre ne prétend pas faire une liste exhaustive de méthodes de recalage mais vise à introduire le principe général d'une méthode de recalage. Elle mettra aussi en évidence qu'étant donné la grande diversité des méthodes décrites dans la littérature, le choix des méthodes à utiliser pour une application donnée peut se révéler complexe.

Le recalage d'images médicales est une problématique qui a beaucoup évolué depuis son introduction vers la fin des années 1970. Une littérature importante sur le recalage d'images a abouti à la publication de plusieurs états de l'art généraux [34, 35, 36] ainsi que d'autres plus focalisés sur l'imagerie médicale [37, 38, 39, 40, 41, 42, 43]. Maintz [37] a introduit une classification des algorithmes de recalage basée sur les caractéristiques des applications (les modalités impliquées, le sujet, le type de transformation, etc.). Cette classification montre que la problématique du recalage dépend des données à recaler. Jannin [40] reprend cette idée pour définir 4 types de recalage : le recalage intra-sujet monomodal, le recalage intra-sujet multi-modalités, le recalage inter-sujets monomodal, le recalage inter-sujets multi-modalités.

D'un point de vue algorithmique, le recalage consiste à trouver la transformation T_{opti} permettant de réaligner une image I_{src} sur une image I_{ref}. Dans un premier temps, des informations sont extraites de l'image I_{ref} via la fonction f_1 et de l'image I_{src} via la fonction f_2. Ces informations, $f_1(I_{ref})$ et $f_2(I_{src})$, sont alors utilisées dans le calcul d'un critère de similarité S permettant de quantifier la ressemblance entre deux images. En fonction de sa définition, le critère S doit être maximisé ou minimisé pour atteindre une ressemblance maximale entre les images. Dans la suite nous considérerons pour des raisons de simplicité que plus S est grand, plus la ressemblance entre les images est importante. Le recalage revient alors à chercher les paramètres de la transformation qui permettent de trouver le maximum global de S. Ce principe se traduit par l'équation suivante :

$$T_{opti} = \underset{T}{argmax}\ S(f_1(I_{ref}), f_2(I_{src} \circ T)) \tag{2.1}$$

Pour décrire le fonctionnement d'un algorithme de recalage nous avons retenu les états de l'art proposés dans les thèses de Commowick [44] et Rubeaux [45]. Rubeaux reprend notamment la classification introduite par Brown [34] qui décompose une méthode de recalage en 4 phases : l'extraction des attributs, le calcul du critère de similarité, l'application du modèle de transformation utilisé, la mise en œuvre de l'algorithme d'optimisation. Ces caractéristiques sont décrites plus en détail dans les paragraphes suivants. La méthode permettant de résoudre l'équation (2.1) peut être représentée par le schéma de la Figure 2.1 [45].

Les attributs sont les informations extraites à l'aide des fonctions f ; des images à recaler et sur lesquelles l'algorithme de recalage s'appuiera pour aligner les images. Il existe deux grandes approches de recalage qui dépendent du type d'attributs utilisés dans le calcul du critère de similarité : l'approche géométrique s'appuyant sur des attributs géométriques (tels que des points, des surfaces, des courbes etc.) correspondant à des informations de haut niveau, et l'approche iconique s'appuyant sur l'intensité des voxels qui correspond à une information bas niveau. Dans un souci de synthèse, la littérature étant particulièrement abondante dans ce domaine, seuls les principaux attributs et critères de similarité de ces approches sont décrits dans les parties suivantes.

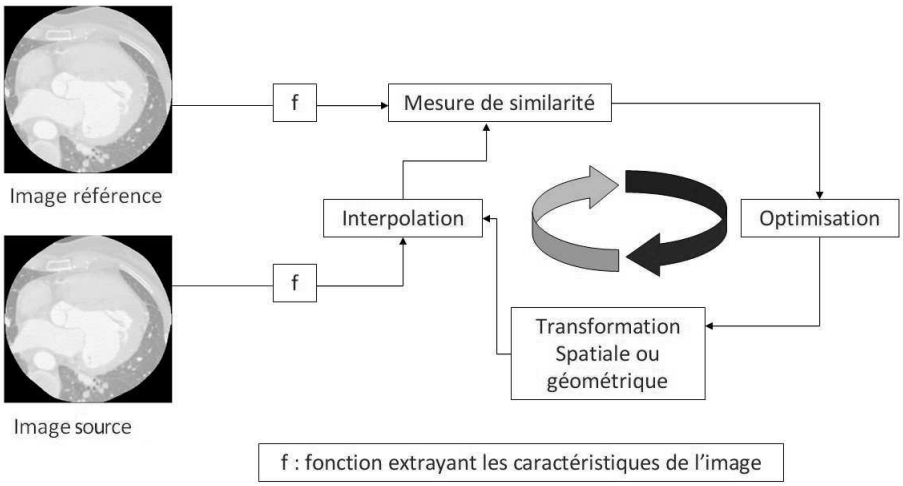


FIGURE 2.1: Schéma de principe d'un algorithme de recalage [45]

2.1.1 Les approches géométriques

Historiquement, les méthodes de recalage géométrique ont été les premières à voir le jour, sans doute car leur fonctionnement est très similaire à la manière dont l'oeil humain cherche à mettre en correspondance 2 images, à savoir la recherche d'attributs géométriques (points, lignes, etc.) bien visibles et communs aux deux images. La mesure de similarité entre images pour cette classe de méthodes correspond à une distance entre les informations géométriques extraites des images à recaler. Son calcul inclut 3 étapes : l'extraction des attributs géométriques, l'appariement de ceux-ci et enfin l'estimation de la transformation minimisant une distance définie à partir de ces attributs géométriques.

2.1.1.1 Les attributs géométriques

Les attributs géométriques constituent des informations de haut niveau comme des points, des lignes, des surfaces qu'il faut extraire au préalable sur les images. Ces attributs constituent en quelque sorte un résumé de l'information contenue dans l'image sur laquelle ledit algorithme de recalage se fonde pour mettre en correspondance les images. Il est donc essentiel que ces attributs à prendre en compte soient présents et détectables dans toutes les images à recaler et non simplement dans l'une ou l'autre des images. De plus, la qualité du recalage obtenu dépend énormément de la qualité d'extraction de ces attributs. Bien qu'on pense essentiellement à des primitives géométriques évidentes comme des points, des courbes ou des plans, il existe plusieurs types d'attributs pris en compte dans les approches géométriques tels que des attributs topologiques (composantes connexes, cavités, etc.), des primitives morphologiques (squelettes ou enveloppe convexe) ou des primitives différentielles (lignes de crête, discontinuités). Ces attributs permettent comme dit précédemment de condenser l'information nécessaire à la mise en correspondance des images. Cependant leur extraction peut se révéler complexe. Dans le cas où des marqueurs externes comme des points lors de l'utilisation d'un cadre stéréotaxique sont visibles sur l'image, l'extraction est relativement simple. En revanche dans les cas où aucun marqueur externe n'est visible sur les images, l'extraction des attributs géométriques nécessite

généralement une intervention manuelle rendant le processus semi-automatique. Dans les cas où le processus de recalage est entièrement automatisé, il s'agit généralement de l'étape la plus compliquée.

2.1.1.2 Les critères de similarité des approches géométriques

Ils représentent des distances entre des attributs géométriques extraits. Généralement ces attributs doivent être appariés avant de minimiser le critère de similarité, c'est-à-dire qu'il faut établir une correspondance entre les éléments détectés sur l'image de référence et sur l'image source. Cette étape d'appariement peut être faite manuellement ou de manière automatique. Pour rendre automatique cet appariement certaines méthodes se basent sur l'intensité des pixels voisins ou d'autres caractéristiques des attributs géométriques à apparier. Cette étape d'appariement peut également être automatisée en la combinant à l'estimation de la transformation lors du calcul du critère de similarité, comme c'est le cas pour des algorithmes tels que l' Iterative Closest Point (ICP) [46].

Certaines méthodes de recalage géométriques ne requièrent pas nécessairement cette étape d'appariement. C'est le cas notamment de la méthode de recalage par distance de chanfrein qui cherche à minimiser une distance entre contours en utilisant des cartes de distance [47, 48]. Dans de tels cas, la carte de distance est créée à partir des contours extraits sur l'image de référence. L'intensité de chaque voxel de cette carte correspond alors à la distance entre le voxel considéré et le voxel le plus proche appartenant au contour de référence. Le recalage consistant à minimiser la distance entre les points du contour de l'image source et ceux de l'image de référence revient alors à bouger le contour de l'image source et à le reporter sur la carte de distance afin de calculer directement la distance moyenne (cf figure 2.2). La carte de distance permet d'approximer la distance euclidienne entre contours tout en évitant d'avoir à calculer à chaque itération cette distance entre chaque point du contour source et le point le plus proche du contour de référence.

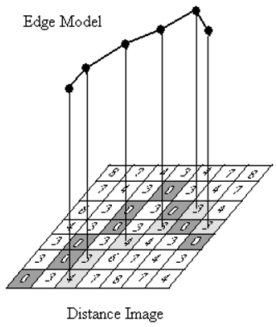


FIGURE 2.2: Illustration de la distance de chanfrein (figure extraite du cours de Caroline Petitjean)

Les distances utilisées (et la nécessité d'appariement) dépendent du type d'attributs extraits ; une simple distance euclidienne pour des points ou des critères comme la distance de chanfrein pour des contours peuvent être utilisés.

Les approches géométriques sont généralement utilisées dans le cadre de transformations rigides ou affines ne possédant que peu de degrés de liberté et apparaissent moins bien adaptées pour des transformations

non rigides, qui nécessiteraient un nombre important d'attributs géométriques à extraire. Leur principal inconvénient vient de la nécessité d'un prétraitement pour extraire les attributs géométriques, tâche qui se révèle souvent complexe. De plus, pour les applications multimodales ne faisant pas apparaître le même type d'information, la recherche d'attributs communs aux différents types d'images se révèle généralement compliquée voir impossible. En revanche, elles ont l'avantage de se baser sur des attributs de haut niveau très descriptifs qui permettent de condenser l'information, et ainsi de diminuer grandement le temps de calcul par rapport à une approche iconique.

2.1.2 Les approches iconiques

2.1.2.1 Les attributs iconiques

On parle généralement d'attributs iconiques pour qualifier l'intensité des pixels d'une image, qui est une information de bas niveau, directement disponible sur les images. Contrairement aux approches géométriques, les attributs iconiques ne nécessitent pas d'étape d'extraction ou d'appariement. Certains prétraitements, comme le filtrage pour limiter le bruit dans les images, peuvent néanmoins être appliqués aux images, ce qui a pour conséquence de modifier l'intensité des pixels des images à recaler.

2.1.2.2 Critères de similarité iconiques

Les critères de similarité iconiques prennent en compte l'intensité de tous les voxels de l'image. Ils ne demandent aucune étape d'extraction d'information ni d'étape d'appariement et sont donc généralement entièrement automatiques. Cependant, il est à noter que le temps de calcul pour ce type de critère est généralement plus important que celui des critères basés sur des attributs géométriques. Le choix d'un critère iconique adapté à une application dépend bien évidemment des modalités impliquées mais plus généralement de la relation liant les intensités des voxels des images à recaler.

La relation la plus simple de **conservation de l'intensité** correspond à l'hypothèse selon laquelle, lorsque les images sont correctement recalées, les voxels des images ont les mêmes intensités pour une même position. La Somme des Différences Absolues (SDA) correspond à une mesure triviale respectant cette relation. Cependant le critère le plus utilisé de cette classe est la Somme des Carrés des Différences (SCD) définie par la formule suivante :

$$SCD(I_{ref}, I_{src} \circ T) = \sum_p [I_{ref}(p) - I_{src}(T(p))]^2 \tag{2.2}$$

Plus la (SCD) est faible, meilleur est considéré le recalage. Ce critère doit donc être minimisé. Ce critère suppose que les deux images à recaler ne diffèrent que d'un bruit gaussien, ce qui est rarement le cas en pratique, même dans le cadre d'un recalage monomodal intra-sujet, pour lequel ce critère a été défini.

Souvent il existe **une relation linéaire ou affine** entre les intensités des pixels d'images provenant d'une même modalité, exprimée par :

$$I_1(p) = a.I_2(p) + b$$

avec $I_1(p)$ intensité de l'image 1 au pixel p, $I_2(p)$ intensité de l'image 2 au voxel p, a et b étant des constantes réelles, b étant nul dans le cas d'une relation linéaire. C'est pourquoi certains critères comme le coefficient de corrélation [49] (CC) ont été définis :

$$CC(I_{ref}, I_{src} \circ T) = \frac{\sum_p [I_{ref}(p) - E(I_{ref})].[(I_{src} \circ T(p)) - E(I_{src} \circ T)]}{\sqrt{\sum_p [I_{ref}(p) - E(I_{ref})]^2 \sum_p [(I_{src} \circ T(p)) - E(I_{src} \circ T)]^2}} \tag{2.3}$$

$$= \frac{E[(I_{ref} - E(I_{ref})).(I_{src} \circ T - E(I_{src} \circ T))]}{\sqrt{E[(I_{ref} - E(I_{ref})]^2 E[(I_{src} \circ T - E(I_{src} \circ T)]^2}} \tag{2.4}$$

$$= \frac{Cov(I_{ref}, I_{src} \circ T)}{\sqrt{\sigma_{I_{ref}} \sigma_{I_{src} \circ T}}} \tag{2.5}$$

avec $Cov(I_{ref}, I_{src} \circ T)$ la covariance entre les images I_{ref} et $I_{src} \circ T$; $\sigma_{I_{ref}}$ et $\sigma_{I_{src} \circ T}$ les variances des images I_{ref} et $I_{src} \circ T$. Contrairement à la SCD vue précédemment, la valeur absolue ou le carré du CC est un critère à maximiser pour recaler différentes images. Sa valeur est comprise entre -1 et 1. Une valeur absolue de CC proche de 1 signifie qu'il y a une faible dispersion des couples d'intensités $(I_{ref}(p), T(I_{src}(p)))$ autour d'une droite. Dans ce cas, une relation affine permet d'approximer la vraie relation entre les intensités des pixels des images I_{ref} et $I_{src} \circ T$. Ce critère permet notamment de prendre en compte les différences de contraste entre images comme cela est souvent le cas pour des images d'une même modalité. Il s'agit d'une mesure de référence pour le recalage monomodal mais qui reste sensible dans les cas où les images ne représentent pas les mêmes champs de vue (c'est-à-dire lorsqu'une zone apparaît dans une image mais pas dans les autres).

Dans le cas de recalages multimodaux, la relation affine ne permet plus d'expliquer le lien entre les intensités des différentes images. Ainsi, l'hypothèse plus générale d'une dépendance fonctionnelle entre les intensités d'images de différentes modalités à recaler a été émise. Cette relation, moins restrictive que la relation affine, s'appuie sur l'hypothèse selon laquelle la connaissance des niveaux de gris d'une image permet de prédire ceux de la seconde et qu'ainsi il existe une fonction g telle que $I_2 = g(I_1)$, qui à toute valeur de I_1 n'associe qu'une seule et unique valeur de I_2. Cette fonction g n'étant pas nécessairement bijective, la réciproque n'est pas forcément vraie. En pratique la fonction g pouvant approximer la relation entre les intensités des pixels de I_1 et I_2 n'associe que rarement une seule valeur de I_2 à chaque intensité de l'image I_1 mais l'important est que le nombre de valeurs associées reste faible. Ainsi les critères se basant sur cette hypothèse peuvent être asymétriques, contrairement à ceux présentés jusqu'à présent. Dans un recalage multimodal, il faut, pour que cette relation fonctionnelle soit acceptable, retrouver globalement (en dehors des zones pathologiques) à peu près les mêmes structures anatomiques dans les différentes modalités d'images considérées. L'un des premiers critères à avoir été défini pour ce type de relation est le critère de Woods [50] dans le contexte du recalage de données cérébrales TEP/IRM. Son idée principale est de décomposer la première image en classes d'intensités proches puis de reporter ces classes sur la seconde image et de calculer en chacune d'elle, l'écart-type normalisé sur la moyenne représentant la dispersion d'intensités des voxels au sein de chaque classe de la deuxième image. Avant de donner l'expression du critère de Woods, il convient de définir certains termes et d'apporter certaines précisions.

Dans un premier temps, les intensités des images sont converties en entiers allant de 0 jusqu'à $C_{ref} - 1$ avec C_{ref} le nombre de classes d'intensités proches voulu. Pour tout voxel i avec une intensité j dans I_{ref},

notons $I_{src} \circ T | I_{ref} = j$ les intensités des voxels correspondants de l'image I_{src} sur laquelle la transformation T a été appliquée. Soient $m_{I_{src} \circ T | I_{ref} = c}$ la moyenne et $\sigma_{I_{src} \circ T | I_{ref} = j}$ l'écart type de $I_{src} \circ T | I_{ref} = j$ pour tout voxel i ayant une intensité j dans I_{ref}. Soient n_j le nombre de voxels ayant une valeur j dans I_{ref} et $N = \sum_j n_j$, le critère de Woods s'écrit alors :

$$Woods(I_{ref}, I_{src} \circ T) = \frac{1}{N} \sum_{j=0}^{C_{ref}-1} \frac{\sigma_{I_{src} \circ T | I_{ref} = j}}{m_{I_{src} \circ T | I_{ref} = j}} . n_j \tag{2.6}$$

$$= E[\frac{\sigma_{I_{src} \circ T | I_{ref}}}{m_{I_{src} \circ T | I_{ref}}}] \tag{2.7}$$

Le critère de Woods est un critère asymétrique à minimiser, c'est pourquoi il est important de choisir comme image de référence celle qui décrit le mieux la deuxième image. Dans l'étude de Woods [50] qui s'intéressait au recalage multimodal de TEP et IRM cérébrales, l'image IRM a été choisie comme référence. Il est à noter que comme le traceur utilisé en TEP ne se fixait pas sur les structures extra-cérébrales, une étape de segmentation pour extraire ces zones de l'IRM a été nécessaire afin d'obtenir une information anatomique proche dans les deux images. En temps normal les critères iconiques ne nécessitent pas d'étape de segmentation ; aussi rappelons que cette étape n'a été nécessaire que parce que certains lieux de fixation du traceur ne correspondaient pas à l'information anatomique de l'IRM. Ainsi, pour un autre traceur cette étape ne sera pas forcément nécessaire.

Le rapport de corrélation [51], apparu après l'information mutuelle (critère de référence pour du recalage multimodal dont nous parlons plus loin), reste également un critère très utilisé pour le recalage multimodal, fondé sur cette relation de dépendance fonctionnelle. En reprenant les notations introduites pour le critère de Woods, le rapport de corrélation a pour expression :

$$RC(I_{ref}, I_{src} \circ T) = 1 - \frac{1}{\sigma_{I_{src} \circ T} . N} \sum_{j=0}^{C_{ref}-1} \sigma_{I_{src} \circ T | I_{ref} = j} . n_j \tag{2.8}$$

$$= 1 - \frac{E[\sigma_{I_{src} \circ T | I_{ref}}]}{\sigma_{I_{src} \circ T}} \tag{2.9}$$

Ce critère doit être maximisé dans le cas d'un recalage. Comme pour le critère de Woods, il s'agit d'un critère asymétrique basé sur la minimisation de variance, il est donc important de choisir comme référence, l'image qui représente le mieux la seconde. La variance n'étant pas robuste aux points aberrants, ce critère sera sensible si des éléments apparaissent dans une image et pas dans l'autre. L'utilisation d'un estimateur robuste peut permettre de limiter l'impact de ces points aberrants [52]. D'autres mesures moins connues, basées sur une dépendance fonctionnelle des intensités des images à recaler, comme le critère d'Alpert [53], peuvent également être utilisées dans un contexte de recalage multimodal.

Cependant, il se peut que dans certains cas de recalage multimodal, la relation de dépendance fonctionnelle entre images soit discutable. Par exemple, si certains tissus ont des intensités proches uniquement sur l'image 1 et pas sur l'image 2 et que pour d'autres tissus, ce soit l'inverse, alors la relation fonctionnelle n'est plus adaptée pour expliquer le lien entre les intensités des voxels des deux images. Dans ce contexte il existe un

dernier type de relation plus générale que la dépendance fonctionnelle qui a permis une véritable avancée dans le domaine du recalage multimodal. Il s'agit de considérer une relation statistique entre les intensités des images à recaler. Le concept d'histogramme conjoint introduit par Hill [54] a permis l'apparition de plusieurs critères de similarité. L'histogramme conjoint correspond à une généralisation de l'histogramme 1D d'une image qui associe à chaque intensité le nombre d'occurences que cette valeur d'intensité apparaît, ce qui revient à associer à chaque intensité sa densité de probabilité lorsque l'histogramme est normalisé par le nombre de voxels de l'image.

L'histogramme conjoint ou 2D "h", se base quant à lui sur le nombre d'occurrences d'un couple d'intensités. Pour un même voxel p, si $I_{ref}(p) = i$ et $I_{src}(p) = j$, cela correspond à une occurrence du couple d'intensités $(I_{ref} = i, I_{src} = j)$. L'histogramme conjoint "h" associe à chaque couple d'intensité $(I_{ref} = i, I_{src} = j)$ le nombre d'occurrences de ce couple, normalisé par le nombre d'occurrences de tous les couples d'intensités, ce qui correspond à la densité de probabilité conjointe $p_{i,j}$ que ce couple apparaisse. Les densités de probabilité marginales p_i de l'image I_{ref} et p_j de l'image I_{src} s'obtiennent en sommant les lignes, respectivement les colonnes de l'histogramme conjoint. La figure 2.3 représente un schéma de l'histogramme conjoint.

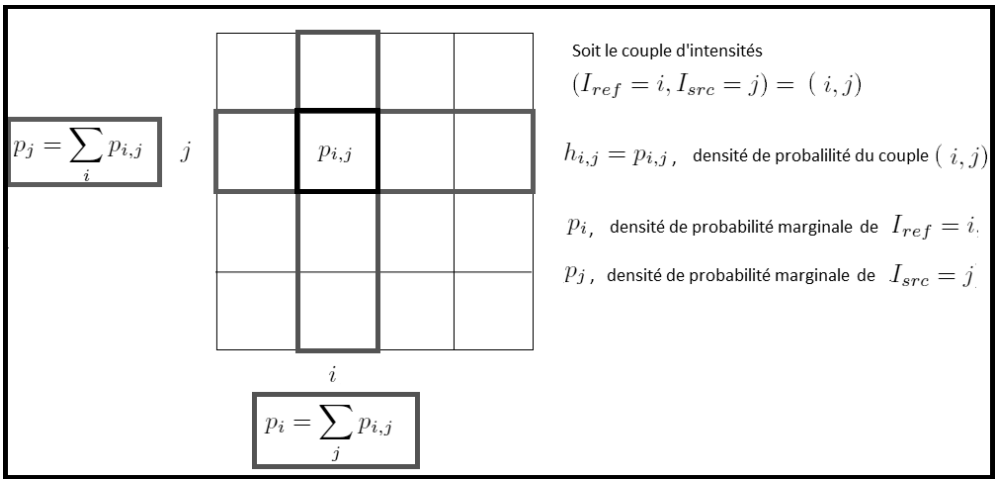


FIGURE 2.3: Histogramme conjoint

A partir de cet histogramme conjoint, plusieurs critères appartenant au domaine de la théorie de l'information et fondés sur l'entropie de Shanon ont été proposés. L'entropie de Shanon permet de quantifier la quantité d'information ou d'incertitude liée à une source d'information. Plus la source émet d'informations différentes, plus l'entropie aura une valeur importante, ce qui, du point de vue du récepteur, correspond à une incertitude plus élevée sur l'information qu'il va recevoir. L'entropie conjointe H est l'un des premiers critères basé sur cette notion d'entropie :

$$H(I_{ref}, I_{src} \circ T) = -\sum_{i,j} p_{i,j} log(p_{i,j}) \tag{2.10}$$

Un recalage de qualité implique une faible dispersion des couples $(I_{ref} = i, I_{src} = j)$ et donc une valeur d'entropie conjointe plus faible. Ce critère est à minimiser. Ce premier critère a en revanche un défaut

assez important : si on fait sortir complètement l'image source du champ de vue de l'image référence via une transormation quelconque, résultant en une image source remplie de zéros, on obtient un minimum de l'entropie conjointe entre les deux images qui ne correspond pas à une superposition.

Pour pallier cet inconvénient, l'information mutuelle (IM)[55, 56] a été proposée et elle reste actuellement l'un des critères les plus utilisés pour le recalage multimodal. Il existe plusieurs formulations de l'information mutuelle qui sont toutes équivalentes mais ont des interprétations un peu différentes. La première formulation reprend l'utilisation de l'entropie conjointe entre images comme le montre l'expression suivante :

$$IM(I_{ref}, I_{src} \circ T) = H[I_{ref}] + H[I_{src} \circ T] - H[I_{ref}, I_{src} \circ T] \tag{2.11}$$

Pour recaler deux images, l'information mutuelle entre les images doit être maximisée. Par rapport à la simple utilisation de l'entropie conjointe, deux nouveaux termes sont présents. $H[I_{ref}]$ est un terme constant qui représente la quantité d'information contenue dans l'image I_{ref}. $H[I_{src} \circ T]$ représente la quantité d'information présente dans l'image I_{src} après transformation par T. Contrairement à $H[I_{ref}]$, le terme $H[I_{src} \circ T]$ varie en fonction des transformations T appliquées sur l'image I_{src}, ce qui permet d'éviter de trouver un optimum de l'IM lorsque l'image source I_{src} sort complètement du champ de vue de l'image de référence I_{ref}. Le terme $H[I_{ref}]$ permet de garder une valeur de IM positive ou nulle car $H[I_{ref}] + H[I_{src} \circ T] \leq H[I_{ref}, I_{src} \circ T]$ même lorsque $H[I_{src} \circ T] = 0$ où l'IM sera nulle.

La deuxième expression fait intervenir la notion d'entropie conditionnelle :

$$IM(I_{ref}, I_{src} \circ T) = H[I_{src} \circ T] - H[I_{src} \circ T / I_{ref}] \tag{2.12}$$

$$IM(I_{ref}, I_{src} \circ T) = H[I_{ref}] - H[I_{ref} / I_{src} \circ T] \tag{2.13}$$

Avec cette formulation, on voit que l'IM peut être calculée à partir de l'entropie conditionnelle $H[I_{src} \circ T / I_{ref}]$, $H[I_{ref} / I_{src} \circ T]$ expliquant ainsi le terme d'"information mutuelle". Dans ces formulations, la maximisation de l'IM est fondée sur la minimisation de l'entropie conditionnelle. En effet, lorsque 2 images sont correctement recalées, on a une meilleure prédiction des intensités des voxels d'une image en connaissant les intensités de l'autre image, ce qui a pour conséquence de baisser la valeur de l'entropie conditionnelle par rapport à celle obtenue pour des images moins bien recalées.

La dernière expression fait intervenir les densités de probabilité marginales p_i et p_j et la densité de probabilité $p_{i,j}$ et permet d'appréhender comment, généralement, l'information mutuelle est calculée en pratique.

$$IM(I_{ref}, I_{src} \circ T) = -\sum_{i,j} p_{i,j} log(\frac{p_{i,j}}{p_i . p_j}) \tag{2.14}$$

Pour résumer, l'information mutuelle est un critère de similarité symétrique très largement utilisé en recalage multimodal. Il permet contrairement à la simple utilisation de l'entropie conjointe d'éviter que l'image source sorte complètement du champ de vue de l'image de référence. Cependant, comme pour les critères basés sur une dépendance fonctionnelle, l'IM est sensible au phénomène de recouvrement partiel,

lorsqu'une partie de l'information d'une première image n'est pas présente sur l'autre et/ou vice versa. Pour limiter ce problème, l'information mutuelle normalisée a été définie [57] par :

$$IMN(I_{ref}, I_{src} \circ T) = \frac{H[I_{ref}] + H[I_{src} \circ T]}{H[I_{ref}, I_{src} \circ T]} \tag{2.15}$$

Cette mesure à maximiser reste l'une des plus utilisées en recalage multimodal et souvent préférée à l'utilisation de l'IM. Il existe d'autres critères dérivés de l'IM, comme l'information exclusive (IE), l'information mutuelle modifiée, etc. que nous ne détaillerons pas dans cette partie.

Bien que ces mesures aient été définies pour la plupart en se basant sur l'utilisation de l'histogramme conjoint pour estimer les densités de probabilité conjointes et marginales, de nouvelles méthodes ont été ensuite développées pour estimer les probabilités de densité sans calcul de l'histogramme conjoint [45]. En effet, le calcul de l'histogramme conjoint représente généralement 90% du temps de calcul. De plus, le nombre de classes choisi, ou la largeur des noyaux dans le cas de l'estimation de l'histogramme conjoint à l'aide d'estimateurs du type Parzen [58] peuvent conduire à des densités de probabilité très différentes.

2.1.3 Les approches hybrides

Dans cette partie, nous voulons mettre en évidence que les critères hybrides se développent pour améliorer la robustesse du recalage dans les cas où des critères classiques sont insuffisants.

Les approches géométriques et iconiques ont chacune des avantages et des inconvénients. Les méthodes géométriques ont l'avantage de s'appuyer sur des attributs très informatifs qui permettent de résumer l'information et ainsi de réduire les temps de calcul. En revanche, l'extraction de ces attributs nécessite une étape de segmentation ou de placement de points qui implique souvent des interventions manuelles d'un opérateur. Ce type d'approche est généralement limité aux cas de recalages multimodaux où il est difficile d'identifier des structures communes aux deux images. De plus, la qualité du recalage obtenu dépend principalement des attributs extraits. Au contraire, les méthodes de recalage iconiques sont généralement entièrement automatiques et permettent, avec des critères comme l'IMN, d'obtenir de bons résultats dans des applications de recalages multimodaux. En revanche, le temps de calcul nécessaire à leur exécution est plus long que celui des méthodes géométriques. De plus, elles s'appuient sur les intensités des voxels qui correspondent à des attributs peu informatifs. Des critères iconiques comme l'information mutuelle ne prennent pas en compte l'information spatiale contenue dans les images.

Lorsque les images source et référence sont "éloignées" l'une de l'autre, les méthodes géométriques sont généralement plus précises que les méthodes iconiques alors que lorsque l'image source est proche de la solution attendue, les méthodes iconiques obtiennent généralement de meilleurs résultats. Il peut paraître intéressant de combiner les 2 approches. Par exemple, dans des cas de recouvrement partiel, où des informations apparaissent sur une image mais pas sur l'autre et vice versa, l'IM peut aboutir à un recalage de qualité insatisfaisante si l'image source se trouve trop éloignée de l'image de référence et donc de la solution attendue. Dans de tels cas, il est nécessaire d'initialiser les paramètres de transformation en utilisant, par exemple, les centres de gravité des images pour les translations et les axes d'inertie pour les angles dans le cas de transformations rigides. Parfois, ce type d'initialisation n'est pas suffisant et il peut être intéressant de

commencer par un recalage géométrique, si l'extraction des attributs n'est pas trop complexe, pour initialiser les paramètres de transformation, puis de procéder à un recalage iconique. Il est donc possible de combiner des approches géométriques et iconiques en les utilisant l'une après l'autre dans le but d'améliorer la qualité du recalage obtenu.

Une autre manière de combiner les deux types d'approches consiste à définir des méthodes de recalage "hybrides" couplant un critère de similarité iconique et des informations géométriques afin de guider le recalage. Une des raisons de l'intérêt pour ce type d'approches vient de problèmes rencontrés en recalage non rigide où le critère de similarité peut posséder de nombreux minima locaux pouvant faire diverger l'algorithme. En effet, malgré l'utilisation d'approches pyramidales, l'algorithme peut converger vers un minimum local différent du minimum global. L'ajout d'information spatiale par rapport à la simple utilisation de critère iconique comme l'IM peut permettre d'aider à résoudre ce problème.

Dans le cadre du recalage d'IRM cérébrales avant et après opération, Pluim et al ont décidé de rajouter une information liée au gradient de chaque image [59]. En effet, de fortes valeurs du gradient traduisent des différences importantes d'intensité entre différents voxels appartenant généralement à des structures différentes. Il s'agit donc d'un attribut très informatif. Le critère défini dans cette étude correspond à la multiplication d'un terme lié au gradient dépendant des images source et référence et de l'IM. Dans le cadre de recalage de données cérébrales TEMP/IRM [60] ou de recalage TDM/IRM [61], l'information de gradient couplée à l'IM a aussi montré un intérêt avec certains traceurs.

Dans le cadre de recalage TEMP/IRM de données cérébrales pour des patients atteints de la maladie de Parkinson [62], l'addition d'un terme spatial à l'IM correspondant à la moyenne d'intensité de l'image TEMP au sein de zones segmentées sur l'IRM (puis reportées sur la TEMP) permet également d'améliorer la robustesse du recalage.

Certains auteurs ont quant à eux décidé d'incorporer directement l'information spatiale dans la définition même de l'IM. Ainsi, dans [63], l'IM est définie non seulement en fonction de l'intensité des voxels des deux images à recaler mais également à partir des intensités des voxels du voisinage nécessitant la définition d'un histogramme 4D. L'IMR (Information Mutuelle Régionale) [64] et l'IMC (Information Mutuelle Conditionnelle) [65] sont également des exemples où une information spatiale est ajoutée à la définition de l'IM. Dans ces critères, en plus des intensités des images source et reférence, une troisième composante spatiale est ajoutée dans le calcul de l'IM, nécessitant un histogramme 3D. Bien que ces mesures semblent intéressantes pour améliorer la robustesse du recalage, le calcul d'histogramme d'ordre supérieur à 2 reste un problème complexe (et coûteux en temps de calcul).

2.1.4 Les modèles de transformation

Le choix d'un modèle de transformation est un élément crucial dans un algorithme de recalage d'images. En effet, c'est le modèle de transformation qui définit les déplacements ou déformations possibles que l'on peut appliquer sur l'image source pour pouvoir l'aligner avec l'image référence. Bien évidemment, le choix d'un modèle de transformation va dépendre des caractéristiques de l'application concernée. En effet, par exemple dans le cas de recalage intra-patient de structures rigides comme des os, un modèle rigide peut être choisi, alors que dans des cas où les structures à recaler ont tendance à changer de forme au cours du temps,

par exemple les poumons lors de la respiration, ce modèle est inadapté et l'utilisation d'un modèle moins contraint est recommandée. Les modèles moins contraints, dits de transformations non linéaires permettent de mieux prendre en compte les déformations locales grâce à leur nombre important de degrés de liberté. Les degrés de liberté correspondent à la dimension de l'espace de recherche, autrement dit au nombre de paramètres du modèle à chercher. Un modèle ayant plus de degrés de liberté est donc généralement considéré comme plus "réaliste" qu'un modèle ayant peu de degrés de liberté dans le cas de déformations locales. Les paramètres optimaux du modèle de transformation doivent correspondre aux antécédants de l'optimum du critère de similarité choisi. Il paraît donc logique de faire l'hypothèse que, plus un modèle a de degrés de liberté, plus le critère de similarité aura d'optima locaux et plus la recherche des paramètres optimaux du modèle sera complexe, ce qui peut conduire l'algorithme à diverger. De plus, les modèles ayant un nombre important de degrés de liberté ont généralement du mal à prendre en compte les transformations globales de grande amplitude. C'est la raison pour laquelle on les initialise souvent à l'aide de modèles possédant moins de degrés de liberté comme les modèles rigides ou affines. Globalement on peut classer les modèles en 2 classes : les modèles linéaires, comprenant les modèles rigides et affines, et les modèles non linéaires.

Les transformations sont ensuite appliquées aux coordonnées des images. Les nouvelles coordonnées n'étant pas forcément entières, un algorithme d'interpolation est nécessaire pour récupérer l'intensité des voxels de l'image transformée.

2.1.4.1 Les modèles linéaires

Les modèles linéaires possèdent un nombre de degrés de liberté faible et fixe contrairement aux modèles non linéaires. Ce type de transformations peut se modéliser, pour du recalage 3D, sous la forme d'une matrice de coordonnées homogènes $A = M.Tr$, de taille 4 x 4, avec $M = \begin{pmatrix} M_{11} & M_{12} & M_{13} & 0 \\ M_{21} & M_{22} & M_{23} & 0 \\ M_{21} & M_{22} & M_{23} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

et la matrice de translation $Tr = \begin{pmatrix} 1 & 0 & 0 & Tr_x \\ 0 & 1 & 0 & Tr_y \\ 0 & 0 & 1 & Tr_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Il existe principalement 2 types de transformations linéaires : les transformations dites rigides et celles dites affines.

- **Les transformations rigides** correspondent à des transformations qui préservent les distances et les angles, n'autorisant que des mouvements de translation et de rotation. Ce type correspond au modèle de transformation le plus contraint ne possédant que 6 degrés de liberté en 3D, 3 paramètres de translation (Tr_x, Tr_y, Tr_z) et 3 paramètres de rotation (θ_x, θ_y, θ_z). Ce type de transformation est généralement utilisé pour du recalage intra-patient si la zone étudiée est supposée peu déformable.
  Dans ce cadre de transformation, la matrice M correspond à une matrice de rotation $R = R_x.R_y.R_z$ avec R_x la matrice de rotation autour de l'axe x, R_y la matrice de rotation autour de l'axe y et R_z la matrice de rotation autour de l'axe z.

$$R_x = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & cos(\theta_x) & -sin(\theta_x) & 0 \\ 0 & sin(\theta_x) & cos(\theta_x) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, R_y = \begin{pmatrix} cos(\theta_y) & 0 & -sin(\theta_y) & 0 \\ 0 & 1 & 0 & 0 \\ sin(\theta_y) & 0 & cos(\theta_y) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, R_z = \begin{pmatrix} cos(\theta_z) & -sin(\theta_z) & 0 & 0 \\ sin(\theta_z) & cos(\theta_z) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Si la taille des voxels des images source et référence n'est pas la même, il faut multiplier la matrice de transformation rigide par une matrice de mise à l'échelle des tailles des voxels Me lors du processus de recalage : $Me = \begin{pmatrix} Me_x & 0 & 0 & 0 \\ 0 & Me_y & 0 & 0 \\ 0 & 0 & Me_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ avec Me_x, Me_y et Me_z rapport de la taille des voxels de l'image référence à celle des voxels de l'image source dans les direction x, y et z. Dans le cadre d'une transformation rigide, Me_x, Me_y et Me_z sont supposés constants et sont en général connus par ailleurs.
La transformation rigide s'écrit au final $A = R_x.R_y.R_z.Tr$ ou $A = R_x.R_y.R_z.Tr.Me$ avec la prise en compte de la taille des voxels.

- **Les transformations affines** peuvent être utilisées comme pour les modèles rigides dans des cas de recalage intra-patient si la zone d'intérêt est peu déformable. Elles permettent, par rapport au modèle rigide, des déformations supplémentaires telles que des changements d'échelle ou des cisaillements. Ces déformations supplémentaires rendent ce modèle idéal pour prendre en compte de fortes déformations globales et initialiser le recalage pour poursuivre avec un modèle non rigide dans le cas d'un recalage inter-patients par exemple. La transformation affine possède, par rapport au modèle rigide, 3 paramètres de facteur d'échelle (Ce_x, Ce_y et Ce_z) dans les 3 directions et 6 paramètres de cisaillement (Cis_{xy}, Cis_{yx}, Cis_{xz}, Cis_{zx}, Cis_{yz}, Cis_{zy}) en plus, soit 15 degrés de liberté en tout, permettant les déformations supplémentaires décrites précédemment.
La transformation affine sous forme matricielle correspond à $A = R_x.R_y.R_z.Tr.Ce.Cis$ avec Ce matrice de facteur d'échelle et Cis matrice de cisaillement.

$$Ce = \begin{pmatrix} Ce_x & 0 & 0 & 0 \\ 0 & Ce_y & 0 & 0 \\ 0 & 0 & Ce_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ et } Cis = \begin{pmatrix} 1 & Cis_{xy} & Cis_{xz} & 0 \\ Cis_{yx} & 1 & Cis_{yz} & 0 \\ Cis_{zx} & Cis_{zy} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Généralement les paramètres Cis_{yx}, Cis_{zx}, Cis_{zy} sont mis à 0, ramenant la transformation affine à un modèle ne possédant que 12 degrés de liberté.

2.1.4.2 Les modèles non linéaires

Comme dans le cadre de la thèse nous avons principalement travaillé avec les modèles affine et rigide, nous ne présentons ici que succinctement différents modèles non linéaires. Pour plus de précisions sur ces modèles, le lecteur pourra se référer à l'état de l'art de Holden [36].

Les modèles non linéaires imposent des déformations plus complexes que les modèles linéaires et sont surtout focalisés sur les déformations locales de petite amplitude pour tenir compte de phénomènes comme la respiration, le mouvement cardiaque ou les différences morphologiques entre individus. C'est à dire que 2 points éloignés l'un de l'autre ne subissent pas la même déformation. Ils possèdent donc plus de degrés de liberté que les modèles rigides et affines mais nécessitent néanmoins une initialisation par ces derniers pour prendre en compte les déplacements de grande amplitude. Les transformations non linéaires induites sont représentées par exemple par des champs de vecteurs de déformation. Ces modèles généralement utilisés avec des critères de similarité iconiques peuvent être classés en 2 catégories :

- **Les modèles de transformations paramétriques :** ils estiment le champ de déformations comme une somme de fonctions polynômiales définies à partir de grilles de Points De Contrôle (PDC) répartis sur l'image de référence. Nommons P_k le $k^{ème}$ PDC avec K le nombre total de PDC. Les PDC peuvent être répartis de manière uniforme ou non suivant les déformations recherchées, c'est à dire que plus il y a de PDC dans une zone, plus la déformation sera censée être précise, à condition de ne pas tomber dans un optimum local. Le nombre de PDC est directement lié au nombre de degrés de liberté de la transformation non linéaire. A chaque PDC P_k, est associé un vecteur de déformation V_k de dimension 3 pour un recalage 3D. A partir de ces vecteurs V_k sont calculées les valeurs du champ dans les autres points de l'image par interpolation à l'aide d'une fonction polynômiale que nous nommerons Fp. Ainsi la transformation T_{local} au voxel p s'écrit :

$$T_{local}(p) = \sum_{k=1}^{K} V_k \times Fp(p, P_k) \tag{2.16}$$

  En fonction du type de fonction polynômiale choisie, la zone d'influence de chaque point de contrôle P_k, sur les autres points peut être plus ou moins grande. Les (Radial Basis Function) fonction à bases radiales (RBF) et les (Thin Plate Spline) fonction à plaques minces (TPS), sont par exemple des fonctions polynômiales qui auront pour conséquence que chaque PDC aura une influence sur tous les points de l'image alors que les B-Splines et les RBF localisées [66] impliquent que le vecteur de déformation en un point p sera déterminé uniquement par les PDC de son voisinage. Autrement dit, les PDC n'ont une zone d'influence que dans leur voisinage lorsqu'on utilise ces 2 dernières fonctions. L'algorithme Free Form Deformation (FFD) présenté par Rueckert [67] utilise par exemple des fonctions de type B-Splines. Le type de modèle de déformation présenté ici est dit paramétrique. Il est contraint de par sa définition et ne nécessite donc généralement pas de terme de régularisation.

– **Les modèles de transformations denses (ou libres) :** ils ne sont plus contraints par des paramètres de transformation. Dans ces modèles, à chaque voxel de l'image de référence est associé un vecteur de déplacement. Ainsi, pour contraindre le modèle, il est indispensable de rajouter un terme de régularisation afin d'obtenir des transformations "physiquement possibles". Les principaux termes de régularisation se basent soit sur le flux optique [68], qui fait l'hypothèse d'une invariance d'intensité au cours du temps, soit sur une contrainte élastique [69], précise sur les petits déplacements mais incapable d'appréhender de grands déplacements, soit sur une contrainte fluide [70], capable d'appréhender de grands déplacements.
Une première transformation globale affine est souvent utilisée pour initialiser des recalages nécessitant des modèles de transformation élastiques.

Il est à noter que les différents modèles présentés ci-dessus peuvent être utilisés de manière locale ou globale sur toute l'image comme le montre la Figure 2.4. Généralement les modèles non linéaires sont appliqués de manière locale et les modèles affines de manière globale. Il existe néanmoins des approches qui utilisent un modèle affine localement comme l'approche par appariement de blocs que nous ne décrirons pas ici.

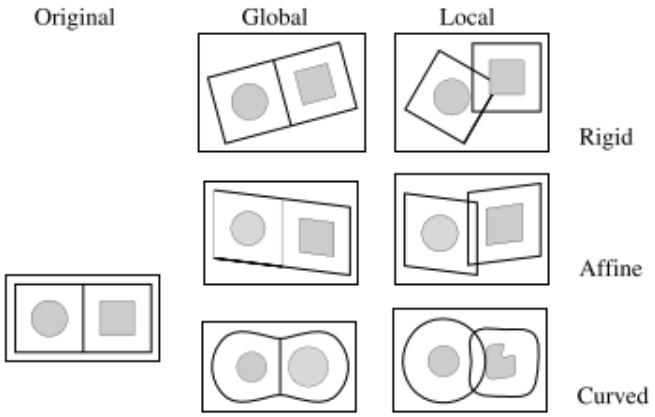


FIGURE 2.4: Exemples de transformations [37]

2.1.5 Les méthodes d'interpolation

Pour déterminer les intensités des voxels de l'image source déformée $\mathrm{T}(I_{src})$ par la transformation T, il est nécessaire d'appliquer la transformation T^{-1} aux coordonnées des voxels de $\mathrm{T}(I_{src})$ pour connaître l'intensité des voxels correspondant sur l'image source I_{src} avant transformation. Or lorsqu'on applique cette transformation T^{-1} sur les voxels de l'image $I_{src} \circ T$, il est probable que les coordonnées obtenues soient non entières, ce qui empêche ainsi de récupérer directement l'intensité du voxel correspondant. L'estimation de l'intensité en cette coordonnée peut alors se faire à l'aide de méthodes d'interpolation. Il existe plusieurs

méthodes d'interpolation pour résoudre ce problème, chacune prenant en compte un voisinage plus ou moins large autour de la coordonnée non entière considérée, en pondérant l'influence de chaque voxel de manière différente. Nous n'en citerons que quelques unes des plus usuelles.

La méthode la plus rapide nommée **"méthode des plus proches voisins"** consiste à conserver l'intensité du voxel le plus proche de la position trouvée. L'intensité de l'image $I_{src} \circ T$ estimée au voxel p de coordonnées $[p_x, p_y, p_z]$ par la méthode des plus proches voisins peut s'exprimer par la formule suivante :

$$I_{src} \circ T(p) = I_{src}(round[T^{-1}(p)]) \tag{2.17}$$

Avec "round" une fonction qui associe à chaque coefficient réel x_n d'un vecteur $[x_1, x_2, .., x_N]$, sa valeur arrondie à l'entier le plus proche.

Cette méthode d'interpolation peut être utilisée pour appliquer une transformation sur une image binaire comme c'est le cas pour l'image de contour dans le cas d'un recalage par la méthode de la distance de chanfrein. En revanche, bien que très rapide, elle est moins utilisée pour des images non binaires car l'estimation de l'intensité ne prend pas en compte les voxels du voisinage, ce qui implique une estimation grossière de l'intensité au voxel considéré.

L'interpolation trilinéaire permet d'éviter ce problème. Cette méthode permet de prendre en compte l'intensité du voisinage direct de la coordonnée considérée, ce qui augmentera le temps de calcul par rapport à la méthode des plus proches voisins mais garantira une meilleure estimation. L'interpolation trilinéaire de l'intensité de l'image $I_{src} \circ T$ estimée au voxel p de coordonnées $[p_x, p_y, p_z]$ calcule la somme des intensités des 8 voisins de $T^{-1}(p)$, a_i de coordonnées $[a_{i,x}, a_{i,y}, a_{i,z}]$ (dans le cadre d'images 3D) pondérée linéairement par rapport à leur distance à cette coordonnée d'après l'expression 2.18.

$$I_{src} \circ T(p_x, p_y, p_z) = \sum_{i=1}^{8} (1 - abs(T^{-1}(p_x) - a_{i,x})) \times (1 - abs(T^{-1}(p_y) - a_{i,y})) \times (1 - abs(T^{-1}(p_z) - a_{i,z})) I_{src}(a_{i,x}, a_{i,y}, a_{i,z}) \tag{2.18}$$

D'autres méthodes d'interpolation, que nous ne détaillerons pas ici, comme le sinus cardinal, les B-Splines etc. sont également très utilisées dans le cadre du recalage d'images médicales. Concernant notre application, ces méthodes censées obtenir une meilleure estimation, aboutissaient à des résultats similaires à ceux obtenus par l'utilisation de l'interpolation trilinéaire en augmentant le temps de calcul. Pour notre stratégie, nous avons donc essentiellement utilisé l'interpolation trilinéaire qui représente un bon compromis entre qualité d'estimation et temps de calcul.

2.1.6 La stratégie d'optimisation

Le recalage consiste à rechercher les paramètres, dont le nombre peut varier en fonction du modèle de transformation choisi, conduisant à l'optimum global (minimum ou maximum) du critère de similarité sélectionné. Etant donné que les espaces de recherche sont de dimension 6 ou plus, parfois plus de 1000 pour les modèles non linéaires, il paraît évident qu'on ne peut pas effectuer une recherche exhaustive dans cet espace pour trouver ces paramètres optimaux. Il convient donc d'utiliser une stratégie d'optimisation mettant en jeu un algorithme d'optimisation ainsi qu'une approche de recherche pyramidale que nous expliquerons.

Le but de l'algorithme d'optimisation est de rechercher les paramètres de transformation conduisant à l'optimum (minimum ou maximum) du critère de similarité tout en minimisant les temps de calcul. De manière idéale, l'optimisation permet de trouver le chemin le plus court (dans l'espace des transformations possibles) menant à l'optimum global de la fonction de coût, en choisissant une direction et un pas de recherche adaptés afin de ne pas rester bloqué dans un optimum local. Dans la suite, on considèrera l'optimum comme étant un minimum, ainsi si une fonction f est à maximiser on cherchera le minimum de $-f$. Des méthodes stochastiques comme celle du "recuit simulé" [71] permettent de trouver les minima globaux, cependant ce type de méthodes nécessite généralement un temps de calcul beaucoup plus long que des méthodes déterministes qui s'attachent à trouver un minimum local. Dans cette partie, on s'intéressera donc aux algorithmes déterministes. On peut modéliser un algorithme d'optimisation comme un processus itératif d'estimation du vecteur de paramètres μ par l'expression suivante :

$$\mu_{i+1} = \mu_i + a_i.d_i \tag{2.19}$$

avec μ_i le vecteur de paramètres estimé à la $i^{ème}$ itération, a_i le pas de recherche et d_i la direction de recherche à la $i^{ème}$ itération. Le processus peut s'arrêter lorsqu'on a atteint un certain nombre d'itérations ou lorsqu'on estime que la différence entre les paramètres estimés à l'itération i et ceux estimés à l'itération $i+1$ est "suffisamment faible".

Hormis la méthode du simplex [71] qui permet d'estimer conjointement les directions et le déplacement, les algorithmes d'optimisation multidimensionnelle nécessitent, en plus du processus de choix de la direction à privilégier, un algorithme d'optimisation 1D pour rechercher le déplacement optimal dans cette direction.

Le choix du pas est un élément primordial dans la recherche d'un minimum de la fonction de coût puisqu'un pas trop grand diminuera le nombre de calcul mais risque d'éviter le minimum global alors qu'un pas trop petit risque de focaliser la recherche sur un minimum local. Cette recherche de pas revient à trouver la position a_i du minimum de la fonction de coût dans la direction d_i et correspond donc à une procédure d'optimisation 1D. La "recherche linéaire inexacte" [71] peut être utilisée à ces fins. Elle se compose de 2 étapes : une étape de "mise entre parenthèses" où l'on recherche les bornes d'un bassin de convergence, zone plus ou moins convexe contenant le minimum, et une étape de "mise entre sections" consistant à rechercher le minimum entre les bornes du bassin de convergence. Pour la deuxième étape, la méthode de Brent [71] utilisant des interpolations paraboliques pour calculer la dérivée et trouver la position du minimum peut être

utilisée. La figure 2.5 représente l'enchaînement de ces deux phases permettant de trouver le minimum de la fonction de coût dans la direction choisie. Des méthodes utilisant le calcul de la dérivée peuvent également être utilisées pour cette seconde phase.

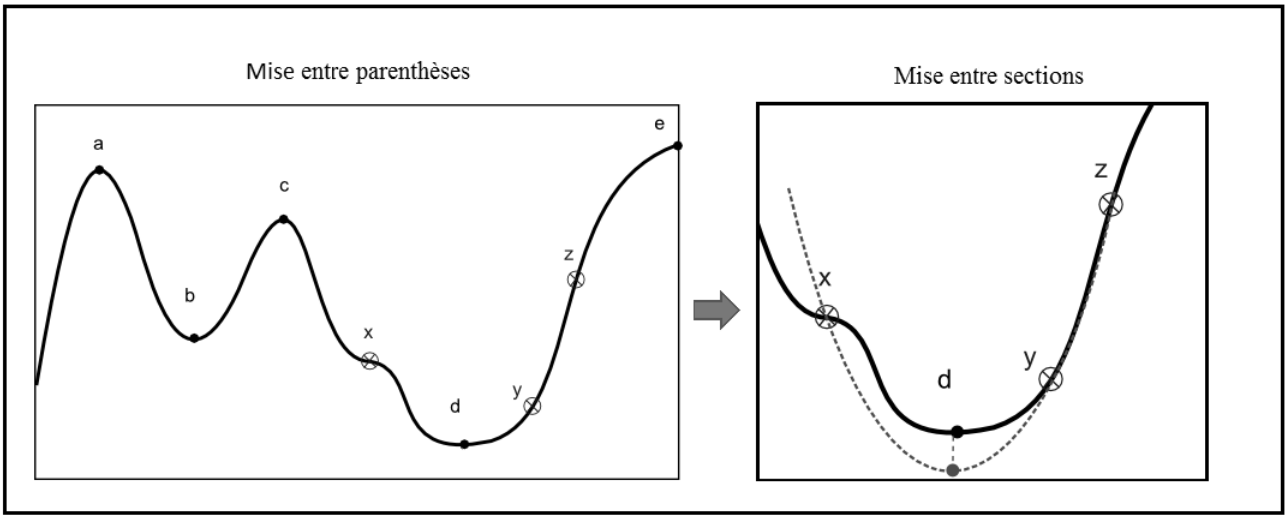


FIGURE 2.5: A gauche : Phase de mise entre parenthèses où l'on peut voir que la fonction de coût f possède plusieurs optimum locaux. Finalement, la phase de mise entre parenthèses permet de trouver les bornes x et z du bassin de convergence contenant un minimum car $f(x) > f(y)$ et $f(z) > f(y)$ avec $x < y < z$. A droite : Phase de mise entre sections en utilisant la méthode de Brent qui consiste à contruire une parabole à partir des bornes du bassin de convergence et d'un point contenu dans celui-ci. La position du minimum de cette parabole est ensuite déterminée et reportée en tant que nouvelle borne sur la fonction de coût. On répète ensuite le même processus à partir des points x, y et du minimum de l'ancienne parabole jusqu'à convergence. Ici seule une parabole est représentée, car dans cet exemple le minimum de la parabole est déjà très proche du minimum global de f.

Quant au **choix de la direction** d_i à privilégier, il existe plusieurs types de méthodes que l'on peut classer en fonction de la nécessité ou non du calcul des dérivées partielles.

Les méthodes ne nécessitant pas le calcul du gradient comme la méthode de Powell [71] ont l'avantage d'être simples à mettre en œuvre mais peuvent poser problème dans les cas où un grand nombre de paramètres est à estimer à cause de leur recherche "aléatoire" des directions de recherche. Dans le cadre du recalage affine, cette méthode convient à l'estimation des paramètres.

Les méthodes nécessitant le calcul du gradient, comme la méthode de descente de gradient [71], permettent d'avoir une information supplémentaire quant aux directions de recherche à privilégier. La direction de recherche privilégiée sera l'opposée de celle du gradient comme le montre l'expression suivante :

$$\mu_{i+1} = \mu_i - a_i.grad(f(\mu_i)) \tag{2.20}$$

Avec $grad(f(\mu_i)$ le gradient de la fonction de coût f pour les paramètres μ_i. La figure 2.6 montre un

exemple de choix de direction à privilégier en utilisant la descente de gradient pour une fonction de coût simple.

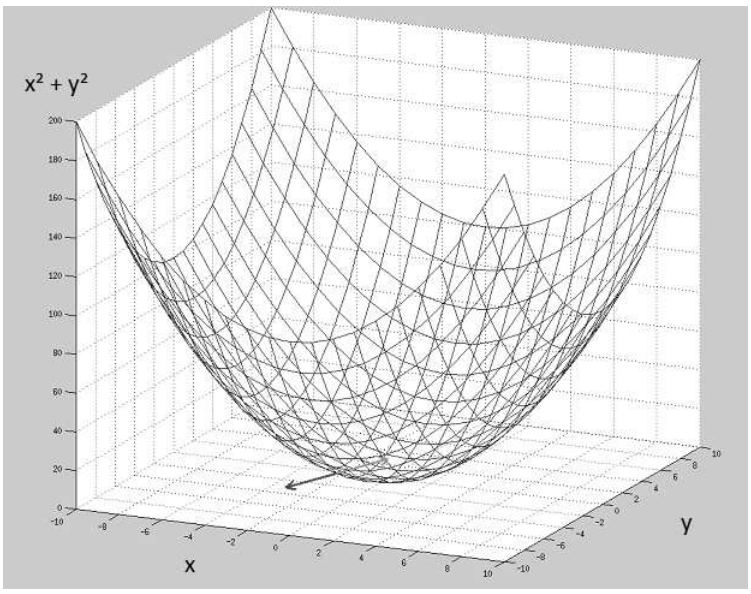


FIGURE 2.6: En rouge, vecteur du gradient de la fonction de coût $f(x, y) = x^2 + y^2$ au point (-1,-3). En vert, vecteur de direction privilégiée, opposée au vecteur de gradient. Le vecteur vert pointe bien en direction du minimum de la fonction f.

Il faut néanmoins faire attention aux erreurs d'estimation du gradient et de la fonction de coût lorsqu'elles ne peuvent pas être calculées directement. Le gradient conjugué non linéaire [72] et la descente de gradient stochastique à pas adaptatif [73] correspondent à d'autres méthodes ayant montré leur efficacité pour des problèmes liés au recalage d'images.

Enfin, **certaines méthodes s'appuient sur les dérivées partielles du second ordre** imposant le calcul de la matrice hessienne d'après l'algorithme de Newton-Raphson :

$$\mu_{i+1} = \mu_i - [Hes(f(\mu_i))]^{-1}.grad(f(\mu_i)) \tag{2.21}$$

Avec $Hes(f(\mu_i))$ et $grad(f(\mu_i))$ la matrice Hessienne et le gradient de la fonction de coût f pour les paramètres μ_i. En revanche, la matrice hessienne ne peut pas être calculée directement et doit être estimée pour éviter des temps de calcul trop importants.

Le choix de la méthode à adopter dépend des caractéristiques de la fonction à optimiser (convexité, nombre d'optima locaux, etc.). Dans le cas du recalage, ces caractéristiques sont liées au critère de similarité choisi mais également au modèle de transformation retenu. De manière générale, plus la solution recherchée est proche de la position de départ, plus la stratégie d'optimisation sera efficace. C'est pourquoi l'initialisation de la méthode d'optimisation est également une étape importante. Par exemple, pour des modèles

de transformation rigides, l'alignement des centres de gravité des images pour initialiser les paramètres de translation et la recherche des axes d'inertie pour initialiser les angles sont classiquement utilisés pour donner un jeu de paramètres de départ à l'algorithme d'optimisation. De plus, un recalage affine peut permettre d'initialiser les paramètres pour un recalage non linéaire. Au-delà de l'algorithme d'optimisation utilisé, une approche pyramidale consistant à estimer les paramètres de transformation pour des résolutions croissantes est souvent utilisée afin d'éviter certains optima locaux éloignés de la solution recherchée, et de diminuer les temps de calcul comme le montre l'image 2.7. Pour les méthodes non linéaires paramétriques qui nécessitent des PDC, une démarche similaire peut être utilisée en plaçant des grilles de PDC de plus en plus fines. Ces deux approches combinées permettent de diminuer les temps de calcul tout en augmentant la robustesse de l'algorithme en évitant des minimum locaux.

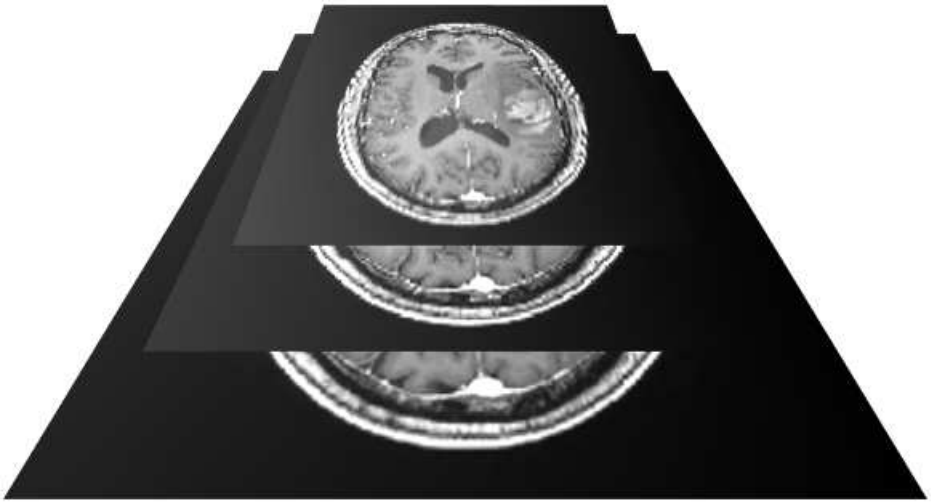


FIGURE 2.7: 3 niveaux de résolution de la même image, de la plus grossière à la plus fine en partant de haut en bas. Lors d'un processus de recalage, on commence par recaler les images à basse résolution puis on affine aux meilleures résolutions en repartant de la position trouvée sur les images de résolution grossière.

2.2 Evaluation du recalage

Etant donné la multitude de choix quant aux éléments constitutifs d'un algorithme de recalage, conduisant à autant de méthodes, il apparaît que la sélection d'une méthode adaptée à une application ainsi que son paramétrage ne sont pas triviaux. Les articles présentant de nouveaux algorithmes de recalage comportent souvent une comparaison avec des algorithmes classiques pour présenter l'intérêt de leur méthode mais cette comparaison porte généralement sur des images privées (ne provenant pas de bases de données publiques) et la comparaison est loin d'être exhaustive [60, 62]. Afin de guider les utilisateurs, des études d'évaluation de méthodes de recalage sur des bases de données publiques ont été menées [74, 75, 76, 77, 78]. Bien que l'évaluation du recalage ait toujours posé problème aux concepteurs d'algorithmes de recalage, l'intérêt pour ce domaine de recherche est assez récent.

2.2.1 Critères d'évaluation

Maintz propose dans son état de l'art [37] plusieurs caractéristiques pour la validation d'algorithmes de recalage, à savoir :

1. **la précision** correspond à l'erreur systématique obtenue dans un recalage avec des entrées parfaites. Prenons l'exemple d'un algorithme de recalage ne prenant en compte que les translations suivant une direction avec un pas de 2 voxels, on s'attend à avoir une erreur de moins de 2 voxels en prenant des images de référence et source identiques. Ce critère prend en compte toute la procédure de recalage.

2. **l'exactitude** consiste à quantifier l'erreur entre la position d'un voxel trouvée par l'algorithme de recalage et "la véritable position" de ce voxel.

3. **la robustesse/stabilité** correspond au fait que pour des positions de départ légèrement différentes de l'image source l'algorithme doit aboutir à la fin du processus de recalage à des positions trouvées légèrement différentes également.

4. **la fiabilité** correspond au fait que l'algorithme trouve une solution cohérente lorsque l'image source n'est pas trop éloignée de l'image de référence.

5. **les ressources requises** concernent le matériel nécessaire pour la mise en œuvre de l'algorithme.

6. **la complexité de l'algorithme** concerne les temps de calcul qui doivent être en accord avec les contraintes de l'application.

7. **la vérification des hypothèses** concerne le réalisme du modèle vis-à-vis de l'application.

8. **l'utilisation clinique** consiste à savoir si l'algorithme correspond et répond bien à un problème clinique.

L'exactitude est souvent la seule caractéristique considérée dans les études d'évaluation. L'estimation de l'exactitude d'une méthode revient à comparer la transformation trouvée par cette méthode avec la vraie transformation ("Gold Standard") permettant de réaligner parfaitement les images. Or cette vraie transformation qui pourrait servir de référence pour l'évaluation est généralement indisponible ou inconnue pour des données réelles.

2.2.1.1 Critères d'évaluation nécessitant l'utilisation d'une transformation de référence "Gold Standard"

Pour juger des performances de méthodes de recalage et notamment estimer leur exactitude, différents critères d'évaluation ont été développés. Certains de ces critères nécessitent la connaissance d'une transforma-

tion de référence, comme le Target Registration Evaluation (TRE) [39, 79], classiquement utilisé. Ce critère TRE consiste à placer des points dans des régions d'intérêt de l'image à recaler puis à calculer la distance euclidienne moyenne entre les coordonnées de ces points, obtenues après application de la transformation trouvée par une méthode de recalage, et celles obtenues par la transformation de référence :

$$TRE(p) = |T(p) - T_g(p)| \tag{2.22}$$

avec T transformation trouvée par l'algorithme d'intérêt et T_g transformation "gold standard".

Ce critère dérive du critère Fiducials Localization Evaluation (FLE) qui permet, dans le cas où des marqueurs externes sont visibles sur les images à recaler, de comparer la position des marqueurs de l'image source après recalage avec celle des marqueurs de l'image de référence. Le problème de ce dernier critère est que le recalage n'est alors évalué qu'au niveau des marqueurs et pas forcément dans les régions d'intérêt, contrairement au TRE.

Généralement la valeur moyenne ou maximale du TRE est prise en compte pour résumer l'information. Les valeurs de TRE sont plus importantes sur les bords de l'image qu'en son centre à cause des rotations. Les valeurs de TRE en chaque voxel dépendent donc de leur position dans l'image. De plus, bien que le TRE soit classiquement utilisé pour juger de la qualité de recalages rigides, il ne permet pas de déterminer si une erreur de recalage est plutôt due aux translations ou aux rotations. Afin de résoudre ce problème, d'autres mesures consistant à comparer directement les paramètres de transformation trouvés par un algorithme de recalage et ceux de la transformation de référence peuvent également servir à estimer l'exactitude d'une méthode. Ainsi, pour la validation d'un algorithme de recalage TEMP/IRM sur des données cérébrales, Loi [80] a introduit 2 critères permettant de quantifier individuellement les erreurs de rotation et de translation pour des algorithmes de recalage avec un modèle rigide. Soient $Tr = (Tr_x, Tr_y, Tr_z)$ et $Tr_g = (Tr_{gx}, Tr_{gy}, Tr_{gz})$ les vecteurs de translation des transformations de l'algorithme d'intérêt et du Gold Standard, le critère quantifiant les erreurs de translation s'écrit alors :

$$TE = \sqrt{(Tr_x - Tr_{gx})^2 + (Tr_y - Tr_{gy})^2 + (Tr_z - Tr_{gz})^2} \tag{2.23}$$

Soient R et R_g les matrices de rotations des transformations de l'algorithme d'intéret et du Gold Standard et u_1 un vecteur unité normalisé, le critère quantifiant les erreurs de rotation s'écrit :

$$ME = (R \times u_1).(R_g \times u_1) = |R \times u_1| \times |R_g \times u_1| \times cos(\phi) \tag{2.24}$$

ME correspond donc au produit scalaire, du vecteur normalisé $u1$ transformé par la rotation R avec le vecteur $u1$ transformé par la rotation R_g. ϕ correspond à l'angle entre les vecteurs transformés. Idéalement TE doit tendre vers 0 et ME vers 1.

Ces précédents critères d'évaluation permettent de mesurer l'exactitude sans pour autant déterminer la source d'éventuelles erreurs. En effet, si un recalage échoue, cela peut être dû soit au modèle utilisé, soit au critère de similarité, soit à l'algorithme d'optimisation etc. Skerl [81] s'est intéressé plus particulièrement à l'influence de 9 critères de similarité différents au sein d'algorithmes de recalage rigide. Afin d'établir un

processus d'évaluation indépendant des algorithmes d'optimisation utilisés dans chaque algorithme, il a défini un espace de paramètres normalisé et échantillonné.

Cet espace, caractérisé par les K paramètres du modèle de transformation considéré, permettant de recaler 2 images, est normalisé de telle sorte qu'un même changement de valeur sur n'importe quel paramètre du modèle aboutisse à un déplacement moyen de voxels similaire sur toute l'image. Cette normalisation est faite de telle sorte qu'on puisse calculer une distance entre 2 transformations. Par exemple, on peut calculer la distance entre une transformation de référence (Gold Standard) et une autre transformation quelconque. Considérons X_0 comme origine de cet espace normalisé de dimension K en tant que position de la transformation de référence. De manière pratique, l'espace est échantillonné en $N \times M + 1$ points $X_{n,m}$, n=1,...,N ; m =-M/2,...,M/2 répartis sous forme de N lignes de M+1 points régulièrement espacés le long de ces lignes. Chaque point $X_{n,m}$ représente une transformation quelconque de K paramètres $[x_1, x_2, ..., x_K]$ où n représente le numéro de la ligne sur laquelle se trouve le point et m sa position sur la ligne. Plus m est proche de zéro, plus le point $X_{n,m}$ est proche de X_0 indiquant que la transformation quelconque est proche de la transformation de référence. Chaque ligne est créée en plaçant aléatoirement un point de départ $X_{n,-M/2}$ situé à une distance R de la position X_0 et le point $X_{n,M/2}$ qui lui est diamétralement opposé par rapport à X_0. Ensuite M+1 points sont répartis régulièrement sur cette droite. L'ensemble des points de départ $X_{n,M/2}$ uniformément et aléatoirement distribués forment une hypersphère de rayon R et d'origine X_0 comme le montre la figure 2.8 en 2D (avec K = 2 dans cet exemple). $SM(X_{n,m})$ est calculé en chaque point de cet espace et représente la valeur du critère de similarité entre l'image de référence et l'image source transformée par $X_{n,m}$. La figure 2.8 représente la courbe de la mesure de similarité le long d'une ligne de l'espace normalisé.

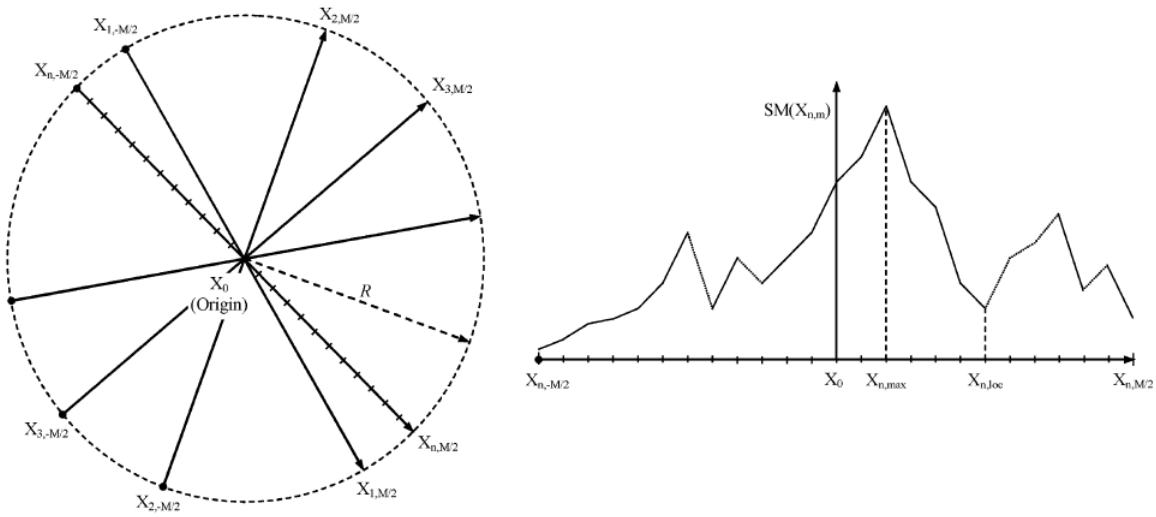


FIGURE 2.8: Représention 2D de l'espace normalisé des paramètres [81] à gauche. A droite, courbe de la mesure de similarité le long d'une ligne de l'espace normalisé.

Dans cet espace, cinq critères ont été définis :

1. **l'exactitude** (ACC) qui représente la racine de la distance au carré entre la position du maximum de chaque ligne $X_{n,max}$ et la position X_0.

$$ACC = \sqrt{\frac{1}{N}\sum_{n=1}^{N} ||X_{n,max} - X_0||^2} \tag{2.25}$$

2. **la distinctivité du maximum global** (DO) qui est une fonction de $r = m.\delta$ avec $\delta = \frac{2R}{M}$. Ce critère indique si les valeurs de similarité chutent fortement autour de la position du maximum, ce qui permet de bien distinguer un pic, ou si les valeurs chutent faiblement, ce qui rend la position du maximum difficile à localiser.

$$DO(r) = \frac{1}{2rN}\sum_{n=1}^{N} 2SM(X_{n,max}) - SM(X_{n,max-k}) - SM(X_{n,max+k}) \tag{2.26}$$

3. **la portée de capture** (CR) qui est définie par la plus petite des N distances entre le maximum global $X_{n,max}$ et son minimum le plus proche $X_{n,min}$ de chaque ligne.

$$CR = min_n(||X_{n,max} - X_{n,min}||) \tag{2.27}$$

4. **le nombre de minima** (NOM) qui correspond au nombre de minima de chaque ligne à une distance r de la position du maximum de chaque ligne.

5. **le risque de non convergence** (RON) qui est une propriété traduisant le comportement du critère de similarité au voisinage des N maxima globaux des N lignes. Une valeur importante de (RON) signifie qu'il y a de nombreux maxima locaux sur lesquels un algorithme d'optimisation risquerait de converger.

$$RON(r) = \frac{1}{2rN}\sum_{n=1}^{N}\sum_{m=max-k}^{max+k} d_{n,m} \tag{2.28}$$

avec

$$d_{n,m} = \begin{cases} SM(X_{n,m-1}) - SM(X_{n,m-1}), \text{ si } m < max \text{ et } SM(X_{n,m-1}) > SM(X_{n,m}) \\ SM(X_{n,m-1}) - SM(X_{n,m-1}), \text{ si } m > max \text{ et } SM(X_{n,m+1}) > SM(X_{n,m}) \\ 0, \text{ sinon} \end{cases} \tag{2.29}$$

Plus les valeurs de ACC, NOM et RON sont petites et plus les valeurs de CR et DO sont grandes, plus le critère de similarité est considéré comme adapté à l'application. Dans [82], Skerl a étendu ces critères pour

évaluer des méthodes de recalage non rigides en adaptant l'espace normalisé des paramètres.

Tout comme pour l'évaluation des critères de similarité, certains auteurs se sont intéressés à l'évaluation d'algorithmes d'optimisation. Par exemple, Klein [83] a comparé plusieurs algorithmes d'optimisation dans le cadre de recalage non rigide à l'aide d'une métrique pour comparer le champ de transformation obtenu par l'algorithme de recalage et un champ de déformation de référence simulé. Dans cette étude, il utilise également une métrique de recouvrement des structures segmentées sur l'image de référence et l'image source recalée. Ce dernier critère ne nécessite pas nécessairement de connaître la vraie transformation pour pouvoir évaluer un algorithme de recalage. La valeur de l'indice de recouvrement choisi entre l'image de référence et l'image recalée est généralement normalisée et permet donc d'estimer la qualité d'un recalage.

2.2.1.2 Critères d'évaluation ne nécessitant pas l'utilisation d'un "Gold Standard"

Dans le cas de données réelles, donc en l'absence de transformation de référence, des indices de recouvrement de surfaces, de volumes (comme les indices de Dice [84] et Jaccard [85]) ou de distances entre contours sur des structures préalablement segmentées sont classiquement utilisés [75, 76]. Ces critères dépendent de l'application puisqu'ils nécessitent pour la plupart l'extraction d'informations spécifiques comme le contour d'un organe par exemple. Dans [75], les auteurs s'intéressent à l'évaluation de 14 méthodes de recalage pour des donnéees d'IRM T1 cérébrales sur une base de données de 80 acquisitions ayant été manuellement étiquetées en 56 ou 128 régions suivant 4 protocoles différents, pour pouvoir évaluer les 14 méthodes à l'aide de 7 critères d'évaluation. Les cinq premiers critères correspondent à trois mesures de recouvrement et deux mesures d'erreur de recouvrement entre volumes. Soient $I_{ref,k}$, le volume correspondant à l'étiquette k de l'image de référence, et $(I_{src} \circ T)_k$, le volume correspondant à l'étiquette k de l'image source recalée par la transformation T, les cinq premiers critères sont les suivants :

1. **"Target Overlap"** (TO) défini par :

$$TO = \frac{\sum_k |I_{ref,k} \cap (I_{src} \circ T)_k|}{\sum_k |I_{ref,k}|} \tag{2.30}$$

2. **"Mean Overlap"** (MO) qui correspond à l'indice de Dice moyenné sur plusieurs régions définies par les étiquettes. L'indice de Dice ayant pour expression

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|} \text{ avec A et B deux volumes à superposer} \tag{2.31}$$

le **Mean Overlap** s'écrit :

$$MO = 2\frac{\sum_k |I_{ref,k} \cap (I_{src} \circ T)_k|}{\sum_k (|I_{ref,k}| + |(I_{src} \circ T)_k|} \tag{2.32}$$

3. **"Union Overlap"** (UO) qui correspond à l'indice de Jaccard moyenné sur plusieurs régions définies par les étiquettes :

$$UOO = \frac{\sum_k |I_{ref,k} \cap (I_{src} \circ T)_k|}{\sum_k (|I_{ref,k} \cup (I_{src} \circ T)_k|} \tag{2.33}$$

4. **"False Negative"** (FN) qui correspond au rapport de la moyenne des voxels appartenant seulement à $I_{ref,k}$ et pas à $(I_{src} \circ T)_k$ sur la moyenne du nombre de voxels de $I_{ref,k}$:

$$FN = \frac{\sum_k |I_{ref,k} \backslash (I_{src} \circ T)_k|}{\sum_k |I_{ref,k}|} \tag{2.34}$$

5. **"False Positive"** (FP) qui correspond au rapport de la moyenne des voxels appartenant seulement à $(I_{src} \circ T)_k$ et pas à $I_{ref,k}$ sur la moyenne du nombre de voxels de $(I_{src} \circ T)_k$:

$$FN = \frac{\sum_k |I_{ref,k} \backslash (I_{src} \circ T)_k|}{\sum_k |I_{ref,k}|} \tag{2.35}$$

Ces cinq critères ont été définis en premier lieu pour des volumes mais ont également été calculés sur les surfaces des étiquettes des images source et référence.

Le sixième critère **"Volume Similarity"** (VO) ne permet pas à proprement parler d'indiquer si le recalage est correct ou non mais est utilisé pour comparer la taille des volumes :

$$VS = 2\frac{\sum_k (|(I_{src} \circ T)_k| - |I_{ref,k}|}{\sum_k (|(I_{src} \circ T)_k| + |I_{ref,k}|)} \tag{2.36}$$

Enfin le dernier critère **"Distance Error"** (DE) correspond à un calcul de distance entre les contours des étiquettes de l'image de référence et ceux de l'image source recalée. Il se calcule en pratique à l'aide d'une carte de distance qui permet de calculer la distance minimum entre un point du contour de l'étiquette k de l'image source recalée $C((I_{src} \circ T)_k)(p)$ et le contour de l'étiquette k de l'image de référence $C(I_{ref,k})$.

$$DE_k = \frac{1}{P}\sum_{p=1}^{P} mindist[C(I_{ref,k}), C((I_{src} \circ T)_k)(p)] \tag{2.37}$$

Plus (TO), (MO) et (UO) tendent vers 1 et plus (FP), (FN), (VS) et (DE) tendent vers 0, plus une méthode de recalage est considérée comme bonne.

Dans l'étude "EMPIRE" [77], qui porte sur l'évaluation de méthodes de recalage pour des données de TDM thoracique, certains critères de recouvrement ressemblant à (FN) et (FP) ont été utilisés pour estimer la qualité de l'alignement des poumons en définissant des étiquettes au sein de ces structures.

Certains critères se basent uniquement sur les paramètres de la transformation et peuvent être utilisés pour tout type d'application. Dans cette même étude, le signe de la matrice jacobienne du champ de transformation qui permet de déterminer si une transformation non rigide est "physiquement acceptable" est utilisé. Il s'agit d'un critère d'évaluation qui ne nécessite pas d'étape d'extraction d'information comme des contours, ou des points et donc n'est pas spécifique à l'application étudiée.

Dans l'étude "NIREP", Christensen [74], en plus d'utiliser des indices de recouvrement classiques, introduit 2 critères d'évaluation de méthodes de recalage indépendants de l'application :

1. **"Cumulative Inverse Consistency Error"** (CICE) qui est un critère permettant de quantifier la différence entre les transformations trouvées en utilisant les 2 sens de transformation, en recalant M images i sur une image de référence j par $T_{i->j}$ ou en recalant une image de référence j sur M images i par $T_{j->i}$.

$$CICE_j(p) = \frac{1}{M}\sum_{i=1}^{M} |T_{j->i}(T_{i->j}(p)) - p| \tag{2.38}$$

   Plus $CICE_j(p)$ est faible plus l'algorithme sera considéré comme "robuste" ; cependant ce critère ne permet pas de quantifier l'exactitude d'un algorithme.

2. **"Cumulative Transitivity Error"** (CTE) qui est un critère permettant de quantifier une erreur de transition lorsqu'on combine des transformations successives.

$$CTE_k(p) = \frac{1}{(M-1)(M-2)}\sum_{\substack{i=1\\ i\neq k}}^{M}\sum_{\substack{j=1\\ j\neq i\\ j\neq k}}^{M} |T_{k->i}(T_{i->j}(T_{j->k}(p))) - p| \tag{2.39}$$

2.2.1.3 Evaluation visuelle

L'évaluation visuelle qualitative par un expert reste le plus souvent le critère préféré en clinique pour la validation des méthodes de recalage.

2.2.2 Estimation d'une transformation de référence ("Gold Standard")

Pour des données réelles, la transformation de référence est généralement indisponible ou inconnue. Dans un contexte de recalage rigide, un expert peut effectuer un recalage de façon manuelle pour obtenir une transformation qui sera considérée comme un "Gold Standard". Cependant cette étape est chronophage pour des données de grandes dimensions et se révèle quasiment impossible dans un contexte de recalage "non rigide".

Pour contourner la difficulté liée à l'absence de transformation de référence dans l'évaluation de méthodes de recalage sur des données réelles, il est possible de "prendre le problème à l'envers" en appliquant une transformation connue sur des données réelles ou simulées [86]. Cette transformation sert de matrice de

référence ("Gold Standard") et est alors comparée à la transformation trouvée par la méthode de recalage pour l'évaluer. La question se pose alors sur le degré de réalisme des images simulées et notamment sur les artefacts liés aux acquisitions des images qui sont des éléments complexes à simuler. Il est possible d'utiliser des "fantômes" [87] qui sont des objets physiques spécialement conçus pour réagir de manière similaire à un tissu ou un organe lors de l'acquisition d'une image. Certains fantômes possèdent des marqueurs externes qui pourront servir de repères pour l'évaluation. En revanche, les "fantômes" restent des objets simplifiée qui ne permettent pas de complètement prendre en compte la forme de l'organe.
Des études récentes s'intéressent également à l'estimation d'une transformation de référence à partir de données réelles en utilisant un *"a priori"* introduit par l'utilisateur [88] ou à partir de transformations trouvées par plusieurs méthodes de recalage dans l'approche "'Bronze Standard"' [89].

2.3 Segmentation

Dans le cadre du suivi de l'évolution tumorale, l'étape de recalage permet notamment de réaligner spatialement des examens à plusieurs instants du traitement ainsi que des examens de plusieurs modalités afin de pouvoir analyser conjointement les données de manière locale. Dans le but d'analyser l'évolution de la tumeur, il est ensuite nécessaire d'extraire la zone tumorale visible sur les examens préalablement recalés et impliqués lors du suivi du patient. La segmentation manuelle par un expert reste le plus souvent la méthode de référence, cependant cette tâche se révèle chronophage avec potentiellement de fortes variations intra et inter-opérateurs. Afin de faciliter cette tâche en limitant le nombre d'opérations que doit effectuer l'utilisateur, de nombreuses méthodes semi-automatiques et entièrement automatiques ont été proposées [90]. Il n'existe pas de méthode générique pour résoudre tous les problèmes de segmentation ; le choix de la méthode à utiliser dépend des zones anatomiques concernées ainsi que des modalités d'imagerie utilisées. Dans le cadre de cette thèse, nous ne nous sommes pas attachés à appronfondir le domaine de la segmentation des gliomes, mais simplement à choisir une méthode adéquate pour notre application. La littérature dans ce domaine étant particulièrement abondante, nous ne prétendons pas fournir dans cette partie une liste exhaustive de méthodes, mais proposons de présenter dans un premier temps les grandes approches de segmentation possibles en nous appuyant sur l'état de l'art de Pham [90], puis de nous concentrer sur la segmentation des gliomes sur les modalités IRM et TEMP, en nous appuyant sur les états de l'art d'Angelini [91], de Bauer [92] et de Dewalle [93]. Puis nous aborderons le problème de la validation des méthodes de segmentation, avant de présenter les critères d'évaluation de l'évolution tumorale.

2.3.1 Méthodologie

La segmentation consiste à partitionner l'image en plusieurs sous régions ne se chevauchant pas et qui respectent un critère d'appartenance, comme l'homogénéité d'intensité dans chaque région par exemple. Mathématiquement, cette idée peut se traduire par l'expression suivante :

$$I = \cup_{k=1}^{K} R_k \tag{2.40}$$

avec $R_k \cap R_j = \varnothing$ pour $k \neq j$; I étant une image ou un volume d'images, R_k la région numéro k de l'image I et K le nombre de régions dans l'image.

Cette segmentation peut se faire suivant différentes approches :

- **le seuillage :** le cas du seuillage simple ne faisant intervenir qu'un seul seuil constitue l'approche la plus simple, puisqu'elle s'appuie directement sur les intensités des voxels de l'image et consiste à choisir un seuil d'intensité permettant de diviser l'image en 2 régions, l'une avec des intensités supérieures au seuil et les autres inférieures à ce seuil.

$$I(x,y,z) = \begin{cases} b_0, \text{ si } I(x,y,z) \leq seuil \\ b_1, \text{ si } I(x,y,z) > seuil \end{cases} \tag{2.41}$$

  En généralisant cette approche, le seuillage multiple consiste à appliquer N seuils pour diviser l'image en N+1 sous régions. La principale difficulté de cette méthode concerne la recherche du seuil adéquat à appliquer. Le choix du seuil peut se faire de manière interactive par un utilisateur, en validant le choix du seuil par une visualisation du résultat de la segmentation. Afin d'automatiser cette sélection, certaines méthodes s'appuient sur l'histogramme de l'image afin d'en repérer les modes. Des approches comme la méthode des modes, la méthode d'Otsu ou la méthode d'analyse des concavités de l'histogramme [94] font partie de ces méthodes. Le seuillage est souvent utilisé comme une étape initiale pour obtenir une segmentation grossière. Les principales limitations de cette approche sont qu'elle est sensible au bruit dans les images et qu'elle ne prend pas en compte d'information spatiale.

- **la croissance de région** : les approches de croissance de région permettent d'extraire une région dont les voxels respectent un critère d'appartenance à la région, comme une contrainte sur l'intensité des voxels de l'image par exemple. Généralement, cette approche nécessite un point de départ placé manuellement par un utilisateur [90]. La croissance de région commence alors en aggrégeant les voxels directement connectés à ce point qui respectent le critère d'appartenance à la région. Ce processus se répète alors avec chaque voxel nouvellement inclus jusqu'à ce que plus aucun voxel voisin ne respecte la condition d'appartenance à la région. L'une des principales contraintes de ce type d'approche vient de la nécessité d'une interaction pour placer le point d'initialisation de l'algorithme. De plus, si on veut segmenter plusieurs régions avec cette approche, il est nécessaire de placer autant de points d'initialisation que de régions à segmenter. Cette approche est également sensible au bruit dans les images. On peut remarquer que l'algorithme "Split and merge" (division et fusion), bien que considéré comme algorithme de croissance de région, ne nécessite pas de point d'initialisation [95].

- **les modèles déformables** : ces modèles correspondent à des courbes paramétriques fermées ou des surfaces, qui se déforment sous l'influence de force externes et internes pour délimiter les contours de l'objet d'intérêt [96]. Ils s'appuient donc sur un processus itératif visant à minimiser une énergie

globale imposée par les forces externes et internes du modèle. Les forces internes définissent les propriétés de la courbe et assurent notamment un certain lissage de la courbe durant sa déformation. Les forces externes sont directement liées à des propriétés caractéristiques de l'image comme le gradient de l'intensité par exemple. Les méthodes de type "Snake" [97] utilisent des modèles paramétriques tandis que les approches de type "Level Set" [98] correspondent à des modèles géométriques que l'on peut considérer comme un ensemble de niveaux d'une fonction scalaire de dimension supérieure. Ces approches sont très largement utilisées en imagerie médicale car elles permettent d'obtenir directement des contours fermés, et que la contrainte de lissage imposée dans le modèle permet de les rendre moins sensibles au bruit. En revanche, les approches paramétriques nécessitent bien souvent une initialisation manuelle qui consiste à placer la courbe paramétrique de base proche de la structure à segmenter.

- **les classificateurs** : ce sont des méthodes de reconnaissance de formes supervisées qui ont pour but de partitionner un espace de caractéristiques extraites de l'image en utilisant des données d'entraînement préalablement segmentées manuellement et étiquetées, permettant ainsi de segmenter automatiquement de nouvelles données. Les caractéristiques sur lesquelles s'appuient les algorithmes incluant souvent l'intensité des voxels et des textures locales [99]. La méthode des "k plus proches voisins" (K-NN) constitue l'un des classificateurs non paramétriques les plus simples et consiste, dans le cadre de la segmentation d'images, à associer à chaque pixel de l'image la même classe que la majorité des k données d'entraînement possédant les intensités les plus proches. La méthode du fenêtrage de Parzen [58] constitue une autre méthode de classification non paramétrique. Parmi les classificateurs paramétriques, on peut retrouver la méthode du maximum de vraissemblance (maximum-likelihood) ou le classificateur de Bayes. Le fait que les classificateurs ne soient pas des processus itératifs les rend efficaces en termes de temps de calcul. Ils ont également l'avantage de travailler sur des vecteurs de caractéristiques, ce qui les rend utilisables pour la segmentation de données multimodales. En revanche, ils nécessitent des données d'apprentissage et ne font généralement aucune hypothèse de répartition spatiale, bien que cet aspect ait tendance à changer dans des approches plus récentes.

- ***le clustering*** : cette approche est relativement similaire aux classificateurs dans le sens où les méthodes de "clustering" se basent sur des vecteurs de caractéristiques pour associer à chaque pixel de l'image une classe. Cependant elles n'utilisent pas de données d'entraînement, il s'agit donc d'une approche non supervisée. Pour compenser l'absence de données d'entraînement, cette approche alterne itérativement les phases de classification et la redéfinition des propriétés caractérisant chaque classe. L'algorithme des "K-moyennes" [100] constitue l'une des méthodes de clustering les plus connues. Son but est de partitionner une image en k régions en alternant itérativement le calcul de la moyenne dans chaque classe et la segmentation de l'image, en associant à chaque voxel la classe dont la valeur moyenne est la plus proche.

  L'algorithme des "C moyennes floues" généralise le principe des K-moyennes en autorisant un pixel à appartenir à plusieurs classes pendant le processus de classification. A la fin de ce processus la classe la plus probable lui est associée. [99]. L'algorithme EM (Expectation Maximisation) [101, 102] reprend ce

principe qui permet d'associer à un pixel plusieurs classes durant le processus de classification mais se fonde sur l'hypothèse selon laquelle les données suivent un modèle mixte gaussien. Il s'agit donc d'une méthode paramétrique. Durant le processus itératif, la méthode alterne entre le calcul du maximum de vraissemblance et le calcul des probabilités d'appartenance aux classes. Les méthodes de clustering ont l'avantage de ne pas nécessiter de données d'apprentissage, mais requièrent tout de même une initialisation de la segmentation. En revanche, tout comme les classificateurs, elles ne font pas l'hypothèse d'un modèle spatial et donc peuvent être sensibles au bruit dans les images. Pour prendre en compte cet aspect, les champs de Markov aléatoires peuvent être utilisés [103].

- **les réseaux de neurones artificiels :** ils constituent une approche de "machine learning" directement inspirée du fonctionnement biologique des neurones. Un réseau de neurones artificiel est constitué de nœuds qui peuvent effectuer des opérations simples telles qu'associer une valeur binaire en sortie, en fonction de plusieurs valeurs en entrée. Dans un réseau, les neurones sont organisés en couches où généralement les sorties des neurones d'une couche représentent les entrées de la couche suivante. Le mécanisme d'apprentissage est alors réalisé en adaptant le poids associé aux connections entre neurones. Cette étape peut être faite soit en utilisant des données d'apprentissage, soit en utilisant d'autres algorithmes. On comprend donc que dans le cadre de la segmentation, les réseaux de neurones peuvent être utilisés comme des classificateurs [99] ou comme des méthodes de clustering [104]. De plus, il est possible d'ajouter une information spatiale dans le processus de classification.

- **la segmentation par atlas** : les atlas ou "template" correspondent à une moyenne d'images d'une même zone obtenue pour des patients différents. Ils sont en général constitués de plusieurs images "moyennes" pour lesquelles une segmentation a été effectuée ou validée par des experts. Il s'agit donc d'images de référence. La segmentation par atlas consiste alors à recaler les données d'un patient sur l'atlas puis à reporter les segmentations de l'atlas sur les données du patient en leur appliquant la transformation trouvée lors du recalage. Pour l'étape de recalage, un modèle de transformation non linéaire est souvent choisi pour prendre en compte les différences d'anatomie entre les images du patient et celles de l'atlas. Ces approches ressemblent à des classificateurs puisqu'elles associent à chaque pixel de l'image du patient une classe présente sur l'atlas, qui constitue ici les données d'entraînement. Cependant les approches de type atlas fonctionnent sur le domaine spatial alors que les classificateurs se basent sur l'espace des caractéristiques pour effectuer la classification.

L'approche à utiliser dépend des spécificités de l'application, et certains prétraitements comme une étape de filtrage pour lisser ou extraire de potentielles caractéristiques de l'image peuvent être nécessaires avant d'utiliser l'une des approches citées ci-dessus. Dans la suite, nous nous concentrons sur la segmentation des gliomes à partir des modalités IRM et TEMP.

2.3.2 Segmentation des gliomes sur les IRM

Comme l'a exposé Angelini dans son état de l'art, on peut diviser les méthodes de segmentation des gliomes sur les images d'IRM en 2 grandes catégories : les approches déterministes regroupant le seuillage, les approches régions et les modèles déformables, et les approches statistiques regroupant les classificateurs et les méthodes de clustering [91].

Les approches déterministes

En 1996, Gibbs [105] a proposé une technique de détection de contours en utilisant des méthodes de morphologie mathématique combinées à une approche de croissance de région dans le but de segmenter des tumeurs se réhaussant sur des données IRM pondérées en T1. Partant d'un échantillon initial au niveau du signal réhaussé dans la tumeur et d'un autre dans les tissus non réhaussés, défini manuellement par l'utilisateur, une première segmentation grossière est effectuée par seuillage. Une carte de contours obtenue par un filtrage de Sobel est ensuite calculée. La région tumorale est ensuite extraite en appliquant un algorithme de croissance de région respectant la contrainte d'appartenance suivante : les voxels appartiennent à la région s'ils ont une intensité comprise entre 4 déviations standards autour de la valeur moyenne de la région et s'ils ne dépassent pas des limites supposées de la tumeur définies par la carte de contours.

En 2004 Letteboer [106] a proposé une méthode de segmentation intéractive pour 3 types de tumeurs : entièrement réhaussée, réhaussement en forme d'anneau, non réhaussé. Aprés une segmentation initiale réalisée manuellement sur une coupe, un ensemble d'opérations morphologiques sont effectuées pour partitionner le volume IRM en régions homogènes. Ces opérations sont effectuées à différents échantillonnage de l'image. Le processus complet de segmentation est guidé par l'utilisateur via une interface adaptée.

Les modèles déformables constituent également des approches intéressantes pour la segmentation des gliomes. En général, le modèle déformable évolue jusqu'aux contours de la tumeur en recherchant les fortes valeurs du module du gradient ou en se basant sur des propriétés locales de régions. En 2005, Droske [107] a proposé une approche de level set pour segmenter l'IRM en régions possédant des propriétés similaires basées sur l'intensité des voxels et la vraissemblance que ceux-ci appartiennent à une région tumorale. Un modèle de grille spatiale adaptative a été utilisé pour le modèle d'optimisation en augmentant la résolution de la grille dans les zones non homogènes. Certains ajustements manuels ont pu être effectués pour corriger les erreurs de segmentation.

En 2009, Wang [108] a proposé d'utiliser un modèle d'écoulement de fluide pour faire évoluer les contours du modèle déformable en direction des limites de la tumeur sur des IRM pondérées en T1. En 2012, Sachdeva [109] a proposé d'utiliser l'intensité des voxels et des motifs de texture pour guider un modèle déformable vers les bords de la tumeur dans différentes modalités IRM.

Les approches statistiques

Les méthodes présentées dans cette partie s'appuient essentiellement sur les approches de classification et de clustering. En effet, de nombreuses méthodes de segmentation se basent sur ces approches car elles permettent de prendre en compte diverses informations dans le processus de segmentation et notamment des informations provenant de jeux de données multimodaux, en les considérant comme des composantes d'un vecteur de caractéristiques. L'approche clustering pour la segmentation de tumeurs cérébrales a été initiée par Schad [110] en 1993 alors qu'il s'intéressait aux motifs de texture de différents tissus.

En 1995, Phillips [111] a proposé d'utiliser une méthode de "C moyennes floues" et la même année Vaiddynathan [112] a comparé cette approche à la méthode de classification supervisée des "k plus proches voisins". Pour cette comparaison, 9 classes de tissus ont été prises en compte : le fond, le liquide céphalo-rachidien, la matière blanche, la matière grise, la graisse, le muscle, la tumeur, l'œdème et la nécrose. Les mêmes auteurs ont également testé une approche de croissance de région pour segmenter uniquement le tissu tumoral sur des IRM pondérées en T1 réhaussées après la définition d'un ensemble de points par l'utilisateur.

En 1998, Clark [113] a introduit une méthode automatique de segmentation pour les glioblastomes basée sur des connaissances a priori définissant un modèle "tumoral". Cette méthode s'appuie sur les informations combinées des modalités IRM pondérée en T1 réhaussée, IRM pondérée en T2 et IRM pondérée en densité de protons. Une étape d'entraînement a été effectuée sur les données de 7 patients en extrayant les tailles de tumeurs et les caractéristiques de réhaussement dans celles-ci. Dans un premier temps, les coupes sont caractérisées comme étant normales ou anormales en utilisant une méthode de "C moyennes floues" et en soumettant le résultat à un système d'analyse expert. Ce système d'analyse se base sur plusieurs a priori tels que le fait que dans une image de cerveau, tous les tissus sains sont sensés être symétriques par rapport à l'axe rostro-caudal. Une fois les coupes anormales repérées, un masque du cerveau est calculé et une segmentation initiale est effectuée en seuillant les histogrammes des 3 modalités. Les voxels n'appartenant pas à la tumeur sont alors éliminés en considérant certains a priori comme 1) les voxels appartenant à la tumeur occupent la partie des plus hauts niveaux de gris de l'histogramme de l'IRM T1 réhaussée ; 2) les voxels de la tumeur apparaissent également dans la partie des plus hauts niveaux de gris de l'histogramme de l'IRM pondérée en densité de protons ; 3) les voxels de la tumeur apparaissent dans la partie des niveaux de gris moyens de l'histogramme de l'IRM pondérée en T2. Il est important de noter que les auteurs considèrent dans leur approche la nécrose et l'œdème comme faisant partie intégrante de la tumeur. Finalement les auteurs ont préféré se concentrer sur l'élimination des voxels n'appartenant pas à la tumeur car ils étaient plus faciles à caractériser que la tumeur elle-même, qui peut avoir des formes et réhaussements très différents d'un cas à un autre.

En 2001, Fletcher [114] reprend cette idée de combiner une méthode de "c moyennes floues" avec des a priori pour la segmentation de tumeurs ne présentant pas de réhaussement. Cette approche fut utilisée sur des données multimodales (IRM pondérée en T1, T2 et en densité de proton). Une des difficultés pointée par les auteurs concerne la séparation entre le liquide céphalo-rachidien et le signal au sein de la tumeur.

En 2001, Kaus [115] a proposé d'utiliser un atlas pour guider une classification en utilisant la méthode des "k plus proches voisins" pour la segmentation de tumeurs en IRM pondérée en T1 réhaussée. Cette approche itérative alterne entre les étapes de classification et de recalage non linéaire sur un atlas. Cela permet à la méthode des "k plus proches voisins" de s'appuyer sur l'intensité des pixels mais également sur l'information spatiale apportée par l'atlas.

Les atlas peuvent également être utilisés pour apporter des informations probabilistes sur les tissus en plus de contraintes spatiales. Par exemple Moon [116] en 2002 et Prastawa [117] en 2003 ont utilisé un modèle probabiliste de tissus et un algorithme de maximisation d'espérance qui permet de segmenter les images en modifiant un atlas avec des informations spécifiques au patient concernant la tumeur et provenant de différentes modalités IRM. En 2004 Prastawa [118] a étendu son travail au cas monomodal.

En 2001, Moonis [119] puis Liu en 2005 [120] ont proposé d'utiliser un processus de segmentation basé sur un algorithme de "connections floues" qui permet de partitionner une image en régions de forte connectivité.

D'après cette littérature, on comprend donc qu'il existe de nombreuses méthodes pour segmenter les tumeurs cérébrales et que l'ajout de contraintes spatiales ou d'informations probabilistes concernant les modèles de tissus permet de mieux contourner certaines difficultés de segmentation.

2.3.3 Segmentation des gliomes sur les données TEMP

De par la faible résolution spatiale des images TEMP et étant donné le bruit élevé qu'elles contiennent, l'extraction du contour précis de la tumeur reste un élément complexe. Il existe de nombreuses approches de segmentation, cependant il semble que pour ce type de modalité, l'approche la plus efficace soit la segmentation par seuillage. Dans cette partie, nous nous appuierons sur l'état de l'art proposé par Dewalle [93] qui présente les méthodes de seuillage utilisées en TEP concernant la segmentation de tumeurs.

La méthode la plus simple consiste à utiliser un seuillage fixe. Afin d'appliquer cette méthode, l'image doit être normalisée au préalable en faisant appel à la notion de "Standardized Uptake Value" (SUV) qui est définie comme le rapport entre la fixation dans le tissu à un instant t (en MBq/kg), et l'activité injectée (MBq) au même instant t divisée par le poids du patient en kg.

$$SUV = \frac{fixation}{\text{activité injectée / poids}} \tag{2.42}$$

La segmentation de la zone tumorale revient alors à sélectionner les voxels d'intensité supérieure à une valeur de 2,5 de l'image normalisée [121]. Cette méthode très simple d'utilisation a l'inconvénient de ne pas être reproductible d'un examen à un autre car le SUV est influencé par plusieurs paramètres comme l'algorithme de reconstruction, les effets de volume partiels, etc.

En 1998, Zasadny [122] a également proposé une méthode de segmentation utilisant un seuil fixe mais avec une expression différente :

$$Ts = m_{fond} + 3\ \sigma_{fond} \tag{2.43}$$

avec m_{fond} et σ_{fond} la moyenne et l'écart type au sein d'une région d'intérêt d'environ 250 pixels prise dans le fond de l'image mais à proximité de la zone tumorale. Ce seuil est ensuite appliqué au sein d'une région d'intérêt englobant la tumeur et tous les voxels supérieurs à ce seuil sont alors sélectionnés comme appartenant à la tumeur. Cette méthode a été validée en comparant les volumes tumoraux détectés en TEP avec ceux trouvés en TDM sur les données de 14 patients. Une forte corrélation sur les données TEP et TDM avant chimiothérapie a été observée. En revanche, d'importantes différences ont été observées sur les données après traitement, dues, pour les auteurs, à la différence entre volume résiduel et volume tumoral.

L'une des méthodes les plus largement utilisées aujourd'hui correspond à l'application d'un seuil basé sur un pourcentage de la valeur maximum de SUV. En routine clinique la valeur de 40% du SUV max est classiquement utilisée [123] comme seuil au-dessus duquel les voxels sont considérés comme appartenant à la tumeur. Ce seuil est appliqué sur une région d'intérêt englobant la tumeur. Bien que cette méthode soit très simple d'utilisation et que la valeur de 40% soit classiquement utilisée, il semblerait que pour certains types de tumeur le seuil à imposer ne doive pas dépasser 20%. Le pourcentage correct à utiliser semble donc dépendre de l'application.

Les méthodes précédentes s'appuient sur l'utilisation d'un seuil fixe. Les méthodes décrites dans la suite sont basées sur l'application d'un seuil adaptatif dépendant notamment de caractéristiques comme le rapport signal sur bruit (S/B) dans les images. En 1997, Erdi [124] a proposé un seuillage basé sur le niveau de bruit dans l'image. Cette méthode a nécessité une calibration en utilisant un fantôme de type Jaszczak contenant des sphères de différentes tailles connues. Chacune des sphères a été remplie avec différentes concentrations de solution de ^{18}F pour simuler différents niveaux de rapport signal sur bruit. En déterminant de manière itérative le seuil optimal à appliquer pour chaque sphère en fonction du rapport signal sur bruit, des courbes de calibrage ont été obtenues. A partir de ces courbes, une expression du seuillage adapté a pu être approximée :

$$Ts(\%) = a.e^{-c} \times volume \tag{2.44}$$

avec $Ts(\%)$ seuil optimal en pourcentage de l'intensité maximale et a et c des paramètres liés au rapport signal sur bruit. A chaque rapport signal sur bruit on associe un couple de valeur (a,c) déterminé à partir des courbes de regression. Pour des sphères de plus de 4 ml, le seuil optimal varie entre 36 et 44% en fonction du rapport signal sur bruit. Ce résultat permet d'expliquer le seuil de 40% classiquement utilisé en routine clinique. En pratique, pour appliquer le seuil, le paramètre de volume est déterminé à partir de données TDM. Cette méthode a l'avantage de prendre en compte le rapport signal sur bruit ainsi que la taille de la tumeur pour sélectionner le seuil idéal. Une forte corrélation fut observée entre les volumes tumoraux extraits en TDM et en TEP.

En 2007, Jentzen [125] a proposé une version améliorée de l'approche précédente qui permet de s'affranchir de l'utilisation de la TDM ; la valeur du paramètre du volume de la tumeur est déterminée à l'aide d'une approche itérative. La détermination de l'expression du seuil optimal nécessite comme précédemment de passer par une étape de calibrage sur fantôme. L'expression du seuil est la suivante :

$$Ts(\%) = \frac{m}{volume} + t \tag{2.45}$$

Ts correspond au seuil optimal en % et m et t constituent des paramètres liés au rapport signal sur bruit qui sont déterminés en fonction des courbes de calibrage.

En 2003, Daisne [126] a proposé une une méthode où l'expression du seuil ne dépend pas du volume tumoral mais seulement du rapport signal sur bruit S/B de l'image. L'expression du seuil déterminé après calibrage est la suivante :

$$Ts(\%) = a + \frac{b}{S/B} \tag{2.46}$$

De la même manière d'autres auteurs comme Black en 2004 [127] et Nestle en 2005 [128, 129] ont proposé des méthodes de segmentation dont l'expression du seuil ne dépend plus du volume tumoral mais s'appuie sur le SUV moyen.

2.3.4 Validation des méthodes de segmentation

Comme pour le recalage, la validation et la comparaison de méthodes de segmentation est essentielle pour choisir la plus adaptée aux données. Pour évaluer les performances d'une méthode, il est intéressant de comparer les résultats de celle-ci avec la vérité terrain [114, 112, 107, 118]. Seulement, dans le cas de données réelles, une telle vérité terrain n'est pas disponible. L'approche la plus classique pour approximer cette vérité terrain est de demander à plusieurs experts d'extraire les contours de la zone à segmenter afin de pouvoir comparer le résultat à celui de la méthode à tester. Cependant, cette approche peut avoir plusieurs limitations. En effet, cette étape se révèle généralement chronophage et difficile, comme dans le cadre de la segmentation de la zone tumorale sur des IRM T1 réhaussées où le réhaussement dépend de la vascularisation et de la perméabilité des vaisseaux au Gadolinium. De plus, dans le cadre de la segmentation manuelle, se pose toujours la question de la variation intra-observateur et inter-observateurs. Mazzara [130] a reporté une variabilité intra-opérateur de $20 \pm 15\%$ et inter-opérateurs de $28 \pm 12\%$ sur les volumes pour la segmentation de tumeurs cérébrales.

Afin de limiter ce problème, on peut tenter de définir une segmentation de référence à partir des segmentations manuelles d'experts en considérant par exemple un point comme faisant partie de la segmentation de référence si la majorité des opérateurs l'a inclus dans sa segmentation [115]. Mazzara [130] a proposé d'utiliser une vérité terrain statistique à partir de segmentations manuelles où chaque pixel possède une probabilité d'appartenir à la tumeur. Cette probabilité servira alors de poids dans le processus d'évaluation des méthodes

de segmentation. Archip [131] a proposé d'utiliser, dans le contexte de l'évaluation de segmentation de tumeurs cérébrales, l'algorithme STAPLE [132] qui permet de créer une vérité terrain en combinant plusieurs segmentations manuelles d'experts. L'utilisation de fantômes comme BrainWeb phantoms [133] peut également permettre d'évaluer des méthodes de segmentation en connaissant cette fois-ci directement la vérité terrain. On comprend donc que la définition d'une vérité terrain reste bien souvent compliquée. De plus, l'un des problèmes majeurs auquel se confronte l'utilisateur quant aux choix de la méthode de segmentation à utiliser vient du fait que dans la plupart des articles, les méthodes sont testés sur des données non publiques. L'apparition de base de données publiques tend néanmoins à limiter ce problème et un véritable effort a été fait dans ce sens avec le BraTS (Brain Tumor image Segmentation) challenge lors de la conférence MICCAI en 2012.

Les méthodes de segmentation sont généralement évaluées en comparant leurs résultats avec une vérité terrain en utilisant divers indices de recouvrement comme l'indice de Dice, les faux positifs, etc. L'évaluation de méthodes de segmentaton nécéssite généralement la connaissance d'une vérité terrain néanmoins certaines méthodes comme l'approche eRWT (extended Regression Without Truth) [134] qui vise à classer des méthodes de segmentation en l'absence de vérité terrain ont montré une certaine pertinence.

2.4 Perspectives pour l'évaluation tumorale

Les critères RANO et RECIST introduits au chapitre 1 sont aujourd'hui classiquement utilisés et constituent une vraie aide à la décision quant à la mise en place de la thérapie des patients. Ils s'appuient principalement sur des modalités IRM et se basent sur des propriétés simples à calculer comme le diamètre de la tumeur, malgré certaines difficultés de définition dans des cas compliqués, et ont pour objectif de classer l'évolution tumorale en 4 catégories. Ces critères ne permettent pas en revanche d'évaluer la tumeur de manière quantitative, ni de s'intéresser à l'hétérogénéité de la tumeur. De plus, l'information tumorale n'est pas toujours visible ou distinguable sur les modalités d'IRM. Or de nombreux efforts sont faits en recherche afin de déterminer des biomarqueurs fiables de l'évolution tumorale.

On peut notamment citer le travail de Moffat [30] en IRM de diffusion qui propose une carte de diffusion fonctionnelle comme biomarqueur précoce de l'évolution tumorale dans le contexte de patients atteints de tumeurs cérébrales traitées par radiothérapie et chimiothérapie.

L'utilisation de l'imagerie fonctionnelle TEMP a également démontré sa pertinence dans certains cas, notamment pour distinguer une récidive tumorale d'une radionécrose, comme l'a montré Prigent [20] dans son travail portant sur l'utilité de la modalité TEMP avec le radio-traceur Sestamibi en complément de l'IRM dans le contexte de patients atteints de gliomes et traités par chimiothérapie. L'imagerie paramétrique à partir d'images TEP au FDG semble également prometteuse pour le suivi de patients atteints de tumeurs cérébrales [135]. En IRM, les travaux d'Angelini [136] portant sur une analyse différentielle après normalisation des intensités sur les IRM paraissent également prometteurs.

En plus de la quantification, la modélisation de l'évolution tumorale [137, 138, 139, 140] constitue également un domaine de recherche important pour mieux appréhender les processus mis en jeu, afin de pouvoir choisir des traitements adaptés de manière précoce.

Enfin certains travaux comme celui de Gal [141] visent à améliorer la représentation de l'information tumorale en résumant cette information dans un espace de plus faible dimension tout en gardant une quantification précise afin de faciliter le suivi de patients atteints de tumeurs cérébrales.

2.5 Conclusion

Dans ce chapitre d'état de l'art il a été présenté deux étapes de prétraitement nécessaires à l'identification de biomarqueurs de l'évolution tumorale que sont le recalage longitudinal et/ou multimodal de données de patients atteints de tumeurs et la segmentation de la zone tumorale sur des données d'IRM et d'imagerie nucléaire. Dans un premier temps, nous avons présenté le principe du recalage d'images et avons montré que pour accomplir cette tâche il existait un grand nombre de méthodes. Or le choix d'une méthode adaptée à une application spécifique peut se révéler complexe d'autant plus que, suivant les jeux de données d'un même type, une méthode peut réussir ou échouer. Pour faciliter ce choix, des critères d'évaluation ont été proposés. D'après la littérature, les méthodes les plus appropriées pour du recalage multimodal sont basées sur la notion d'entropie comme l'information mutuelle, tandis que pour du recalage monomodal des méthodes, s'appuyant sur des contraintes plus fortes comme une relation affine entre les intensités des voxels des images à recaler, paraissent plus adaptées. Une fois le recalage des données effectué, la segmentation de la zone tumorale sur chaque modalité doit faciliter l'identification des biomarqueurs de l'évolution tumorale. Au final, pour la segmentation de tumeurs comme les gliomes, des approches de contours actifs de type snake ou level set semblent adaptées pour l'extraction de la zone tumorale sur les données IRM tandis que sur les données TEMP les approches de type seuillage semblent plus intéressantes. Le prochain chapitre présente comment ces différentes étapes peuvent être incorporées dans une stratégie d'analyse de l'évolution tumorale.

Chapitre 3

Stratégie d'analyse de l'évolution tumorale

Sommaire

3.1 Stratégie de recalage 64

3.1.1 Principe de la stratégie de recalage 64

3.1.2 Mise en oeuvre des méthodes de recalage 65

3.1.3 Classement et évaluation 67

3.1.4 Validation visuelle 67

3.2 Segmentation et analyse de l'évolution des tumeurs 68

3.2.1 Segmentation de la zone tumorale 68

3.2.2 Critères de l'évaluation de l'évolution tumorale 68

Ce chapitre présente le principe général de la stratégie d'analyse de l'évolution tumorale proposée en utilisant plusieurs modalités d'imagerie. Cette stratégie a pour but d'aider à répondre à plusieurs questions, premièrement de déterminer si les différentes modalités d'imagerie impliquées dans le protocole apportent des informations complémentaires et dans un second temps, de pouvoir suivre l'évolution tumorale d'une manière locale en déterminant par exemple si, malgré une diminution globale de la taille de la tumeur, celle ci ne poursuit pas son évolution dans d'autres territoires. Afin de pouvoir comparer l'information tumorale contenue dans chaque modalité, une première étape de recalage des images est nécessaire. Cette première étape se compose de 2 types de recalages :

- recalage mutimodal : les images des différentes modalités sont recalées sur celles d'une modalité choisie comme référence
- recalage monomodal : les différents examens acquis à des temps différents de la modalité de référence sont recalés sur l'examen considéré comme étant l'examen de référence, généralement il s'agit de l'examen avant le début d'un traitement. Les transformations ainsi trouvées sont alors réappliquées aux examens des autres modalités préalablement recalés sur ceux de la modalité de référence à leurs temps respectifs. Ainsi tous les examens se trouvent dans un référentiel commun. Pour chaque recalage, une

stratégie a été mise en place pour s'assurer d'obtenir les résultats ayant la meilleure qualité possible. Cette stratégie est décrite dans la partie suivante.
La deuxième étape de la stratégie d'analyse comprend la segmentation des volumes tumoraux sur les différentes modalités et le calcul d'indices locaux et globaux sur ces volumes extraits. Bien que l'analyse globale puisse se faire complètement indépendamment de l'étape de recalage, le calcul des indices locaux va fortement varier en fonction de la qualité des recalages obtenus. La figure 3.1 présente le schéma global d'analyse de l'évolution tumorale.

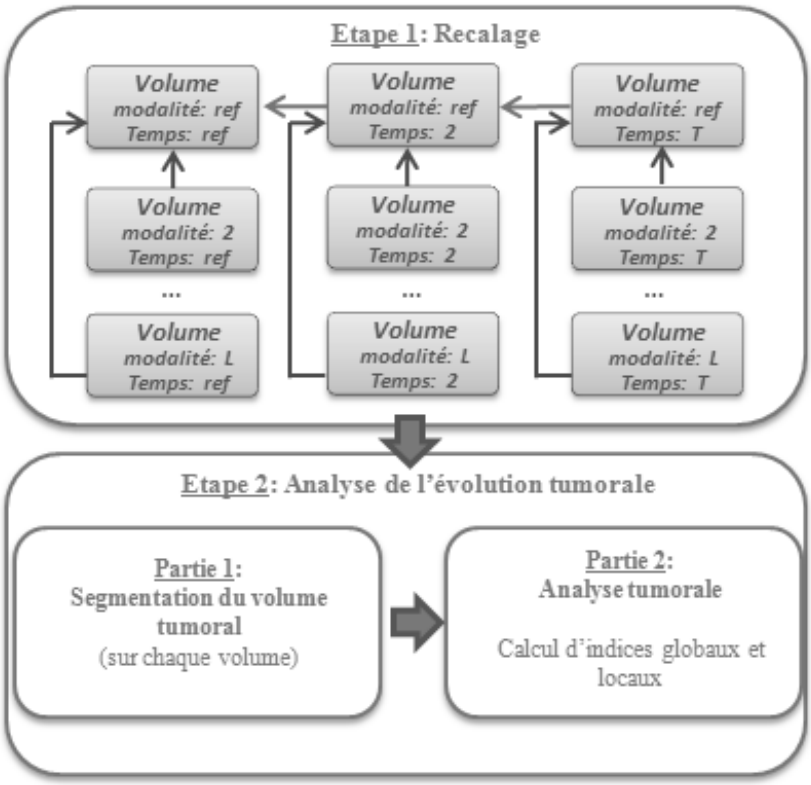


FIGURE 3.1: Schéma de la stratégie d'analyse de l'évolution tumorale.

3.1 Stratégie de recalage

3.1.1 Principe de la stratégie de recalage

Le choix d'une méthode de recalage adaptée aux types de données à traiter est complexe et nous avons vu dans le chapitre 2 que, pour guider les utilisateurs dans leur choix, des études d'évaluation de méthodes de recalage [74, 75, 76, 77, 78] ont été menées sur des bases de données publiques. L'évaluation est classiquement utilisée pour identifier la méthode permettant de recaler correctement le plus de données sur une base de données particulière. Cependant, cette "meilleure" méthode peut échouer ponctuellement et certaines données de la base peuvent donc ne pas être recalées de façon satisfaisante en utilisant cette "meilleure"

méthode. Or il est possible que ces mêmes données puissent être recalées correctement par une autre méthode moins performante globalement. Par exemple dans l'étude "EMPIRE10" [77] qui cherche à évaluer des algorithmes de recalage pour des données thoraciques, on peut observer dans la partie "résultats" (sur le tableau VII) que pour un jeu de données (le jeu de données 25) l'algorithme classé en troisème position obtient un meilleur recalage que les deux meilleures méthodes. Cet exemple met en évidence qu'il peut exister une certaine "complémentarité" des méthodes.
L'idée générale de notre stratégie est de s'appuyer sur cette notion de "complémentarité" entre méthodes pour améliorer globalement les résultats de recalage sur une base de données. Contrairement à l'approche classique qui utilise uniquement la meilleure méthode globale pour recaler tous les jeux de données d'une base, notre stratégie consiste à sélectionner la meilleure méthode de recalage pour chaque jeu de données. Cette stratégie se décompose en trois phases appliquées à chaque jeu de données comme le montre la figure 3.2 :

- la mise en œuvre de plusieurs méthodes de recalage ;
- la définition d'un critère d'évaluation quantitatif et la sélection de la meilleure méthode après classement à l'aide de ce critère ;
- la validation visuelle.

3.1.2 Mise en oeuvre des méthodes de recalage

Cette première phase consiste à utiliser plusieurs (N) méthodes (M_m, $m \in [1, N]$) de recalage sur un même jeu de données pour obtenir les images recalées (I_{srcMm}). En fonction du type de problème, les méthodes à utiliser dans cette phase ne seront pas les mêmes. Par exemple, pour du recalage d'images sur lesquelles il faut tenir compte de la respiration, il vaudra mieux choisir des méthodes utilisant un modèle de transformation non rigide, alors que sur une étude mono-patient au niveau du cerveau on peut se permettre de choisir des méthodes utilisant un modèle de transformation rigide. Il est important d'apporter des précisions sur le terme "méthode" que nous utilisons ici, pour définir une mise en œuvre particulière d'algorithmes de recalage. Ainsi, dans ce contexte, deux paramétrages différents d'un même algorithme de recalage aboutissent à deux méthodes de recalage différentes. De même, une méthode peut combiner l'utilisation de plusieurs algorithmes de recalage, en utilisant par exemple le premier pour initialiser le second.
Une méthode sera donc définie par les paramètres suivants :

- le modèle de transfomation (rigide, affine, élastique, etc.)
- l'initialisation des paramètres du modèle de transformation choisi ;
- l'algorithme d'optimisation ;
- le critère de similarité ;
- les prétraitements appliqués aux images référence et source avant recalage ;
- le sens de transformation utilisé (direct ou indirect).

Le sens de transformation est *"direct"* si on estime directement la transformation permettant de recaler l'image source sur l'image de référence. Le sens de transformation est *"indirect"* si on estime dans un premier temps la transformation permettant de recaler l'image de référence sur l'image source avant d'utiliser la transformation inverse pour réaligner l'image source sur l'image de référence. Certaines méthodes auront

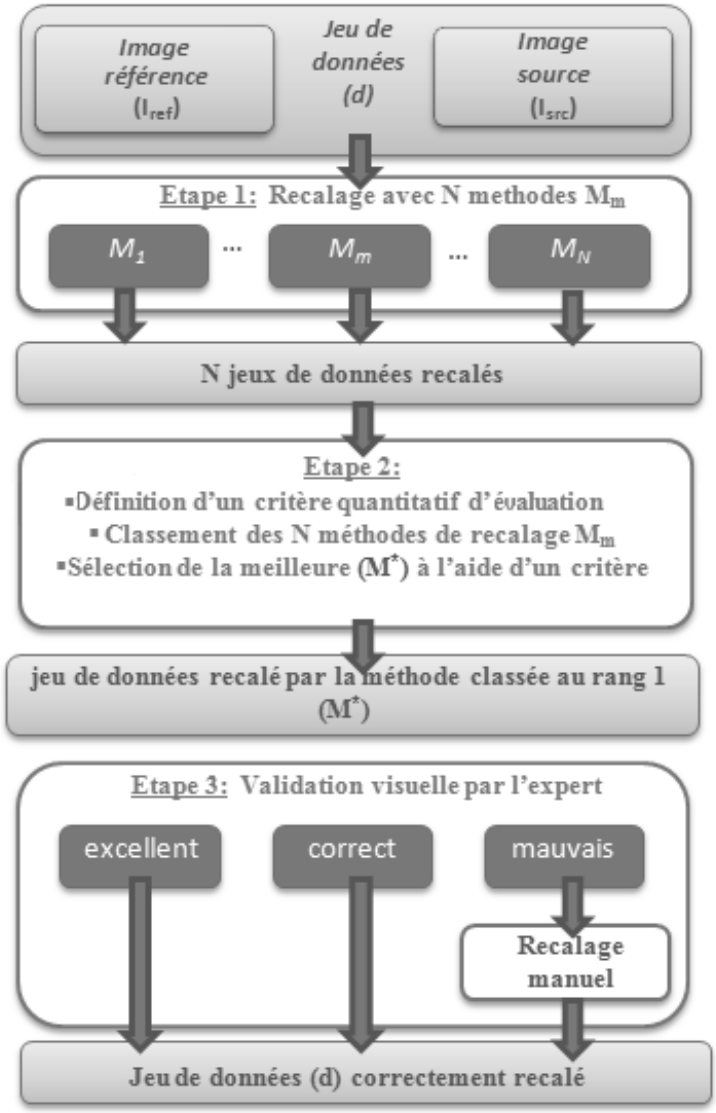


FIGURE 3.2: Schéma de la stratégie de recalage

des étapes en commun comme l'extraction d'informations utiles au critère de similarité utilisé.

3.1.3 Classement et évaluation

La sélection de la meilleure méthode de recalage pour un jeu de données particulier se fait à l'aide d'un critère d'évaluation que l'on calcule sur les image recalées (I_{srcMn}). Ce critère d'évaluation est spécifique de la base de données considérée pour tenir compte des particularités des données, comme les lieux de fixation physiologique d'un traceur radioactif par exemple. Plusieurs critères peuvent être nécessaires et combinés pour l'évaluation de méthodes de recalage. Dans la plupart des cas, un critère représente une propriété nécessaire mais non suffisante pour permettre d'affirmer qu'une méthode est meilleure qu'une autre. Le classement des méthodes, qui doit prendre en compte l'ensemble des critères, est donc difficile. Dans le cas de différents critères de même nature, comme des indices de recouvrement de différentes régions, il est parfois possible de se ramener à la définition d'un critère unique en établissant une relation prenant en compte l'influence de chacun. En revanche, dans le cas de critères de nature différente, il est nécessaire de s'appuyer sur des approches d'aide à la décision pour effectuer le classement. Dans la stratégie que nous avons utilisée pour la première application clinique, les critères d'évaluation permettent d'évaluer les performances "relatives" d'une méthode par rapport aux autres pour un jeu de données et donc de sélectionner l'image recalée I_{srcM^*} obtenue par la méthode de recalage optimale M^*. En revanche, ces critères ne permettent pas systématiquement de déterminer la performance absolue d'une méthode sur un jeu de données, c'est-à-dire de savoir si le recalage obtenu par la méthode M^* est satisfaisant ou non.

Remarque : Pour un même jeu de données il se peut que plusieurs méthodes possédant des valeurs très proches du critère d'évaluation aient des performances similaires. Dans un tel cas, il faudra utiliser d'autres critères comme celui des (TRE) pour confirmer cette équivalence. Cependant dans la stratégie que nous avons définie, seule la méthode ayant obtenue la plus forte valeur du critère est nécessaire et sera validée visuellement par l'expert.

3.1.4 Validation visuelle

Si l'on considère que les critères d'évaluation établis sont pertinents, la méthode M^* a toujours une meilleure performance que les autres méthodes. Il n'est alors pas nécessaire de vérifier visuellement les résultats obtenus par les autres méthodes. Néanmoins, rien ne garantit que le recalage obtenu par M^* soit satisfaisant. La phase de validation visuelle est donc une étape essentielle qui permet de juger de la qualité du recalage obtenu par la meilleure méthode (M*) sélectionnée lors de l'étape d'évaluation. Cette étape de validation consiste généralement en une évaluation qualitative en s'appuyant sur des repères anatomiques. Dans les cas où les critères d'évaluation utilisés dans la phase de classement permettent d'estimer la qualité "absolue" d'une méthode de recalage, l'étape de validation visuelle reste recommandée mais n'est pas systématiquement indispensable pour confirmer la qualité du recalage obtenu.

3.2 Segmentation et analyse de l'évolution des tumeurs

3.2.1 Segmentation de la zone tumorale

La segmentation de la zone tumorale constitue une étape importante dans le cadre de l'évaluation de la réponse tumorale. En effet, l'analyse de l'évolution tumorale comporte, dans la recherche d'un biomarqueur pertinent, le calcul d'indices et de caractéristiques basés sur ces zones segmentées et dont les valeurs peuvent varier de manière significative en fonction de la qualité des segmentations obtenues. L'étape de segmentation se doit donc d'être aussi précise que possible. Comme pour le recalage, les performances d'une méthode de segmentation dépendent des caractéristiques de l'application concernée comme la modalité des images impliquées, la localisation ainsi que la forme de la zone tumorale, etc. Par exemple, pour les modalités TEP et TEMP, la segmentation de la zone tumorale est obtenue principalement en utilisant des méthodes basées sur du seuillage [93] alors que pour des modalités ayant une résolution spatiale plus élevée comme l'IRM, la segmentation de la zone tumorale implique généralement des techniques plus complexes [91].

3.2.2 Critères de l'évaluation de l'évolution tumorale

La question de l'évaluation de l'évolution tumorale pour le suivi de patients est une question cruciale pour laquelle plusieurs critères ont été proposés et sont aujourd'hui classiquement utilisés [23], notamment les critères RANO et RECIST présentés dans le chapitre 1.

Dans la suite de la thèse cette stratégie d'analyse de l'évolution tumorale a été appliquée totalement ou en partie à deux applications cliniques :

1. le suivi de patients atteints de gliomes de haut grade et traités par antiangiogénique ;
2. le suivi de patients atteints de métastases hépatiques.

Ces critères sont aujourd'hui classiquement utilisés mais peuvent être limités dans certains cas. Ils peut donc être intéressant d'utiliser d'autres modalités qui pourraient apporter des informations complémentaires afin de mieux caractériser la réponse tumorale. A partir des segmentations de la zone tumorale vues dans la partie précédente, des indices globaux peuvent être calculés, comme le volume de la zone tumorale, l'intensité moyenne et l'écart-type de l'intensité au sein de la zone tumorale, etc. qui permettent de comparer globalement des examens provenant de différentes modalités ou de temps différents. En particulier, les variations relatives de volume ou de SUV max sont couramment utilisées. Cependant de tels indices ne tiennent pas compte du caractère spatial de l'information, et c'est pourquoi il peut être utile de calculer des indices plus locaux, comme des indices de recouvrement par exemple, permettant de quantifier l'information apportée par les différentes études.

Finalement, l'utilisation de tests statistiques visant à comparer les indices précédents à des mesures de taux de survie des patients tels que le taux de survie global (OS) et le taux de survie sans progression du cancer (PFS) peuvent aider à déterminer leur pertinence et à répondre à la question "Est ce que tels indices permettent de rendre compte de l'évolution tumorale ?" Des tests statistiques sur plusieurs indices peuvent

également permettre de chercher des liens entre eux et de répondre aux questions du type "Est ce que les indices de la modalité A apportent une information complémentaire à ceux de la modalité B ?".

Chapitre 4

Application de la stratégie d'analyse de l'évolution tumorale pour des patients atteints de gliomes de haut grade

Sommaire

4.1 Spécificités de cette application 71

4.1.1 Rappel du contexte médical 71

4.1.2 Intérêt de la stratégie pour le recalage TEMP-IRM 72

4.1.3 Base de données 73

4.2 Mise en œuvre de la stratégie pour le recalage TEMP-IRM 74

4.2.1 Choix des différentes approches de recalage 74

4.2.2 Evaluation quantitative et classement 81

4.2.3 Validation visuelle 95

4.3 Mise en œuvre de la stratégie pour le recalage longitudinal IRM-IRM 96

4.3.1 Choix des différentes approches de recalage 97

4.3.2 Evaluation quantitative et classement 98

4.3.3 Validation visuelle 99

4.4 Segmentation des tumeurs sur l'IRM 99

4.4.1 Segmentation des tumeurs sur la TEMP 100

4.5 Critères d'analyse 101

4.5.1 Critères d'analyse globale 101

4.5.2 Critères d'analyse locale 102

4.6 Conclusion 103

Ce chapitre correspond à la mise en œuvre de la stratégie d'analyse de l'évolution tumorale dans le contexte de patients atteints de gliomes de haut grade. Il se décompose en cinq parties :

- les spécificités de cette application qui rappellent le contexte médical de l'application et notamment les difficultés liées au recalage des données impliquées dans le protocole d'acquisition ;
- la mise en œuvre de la stratégie de recalage qui détaille comment a été implémentée la stratégie de recalage dans le contexte du recalage multimodal TEMP/IRM et du recalage longitudinal IRM/IRM ;
- la segmentation des tumeurs sur l'IRM ;
- la segmentation des tumeurs sur les données TEMP ;
- la présentation des critères d'analyse globaux et locaux mis en œuvre à partir des données segmentées et recalées

4.1 Spécificités de cette application

4.1.1 Rappel du contexte médical

Les gliomes sont les tumeurs du système nerveux central les plus fréquentes chez l'adulte. Les traitements conventionnels comprennent la chirurgie, la radiothérapie, et éventuellement la chimiothérapie en cas de récidive tumorale. Des études récentes portent également sur l'apport de nouvelles thérapies émergentes telles que les traitements anti-angiogéniques qui visent à réduire la néovascularisation impliquée dans la croissance de la tumeur. Cependant, malgré les progrès des nouvelles thérapies, le pronostic de patients atteints de gliomes de haut grade reste péjoratif et ces nouveaux traitements induisent des effets secondaires importants. L'évaluation précoce de la réponse tumorale est donc essentielle pour sélectionner la stratégie de traitement la plus adaptée au patient. La neuroimagerie structurelle et fonctionnelle permet l'évaluation des gliomes de manière non invasive et se révèle cruciale pour le suivi des patients [20]. Actuellement, l'évaluation de la réponse aux traitements anti-angiogéniques se fait en neuro-oncologie à l'aide des critères RANO (Response Assessment in Neuro-Oncology)[142]. Ils sont basés sur les indices cliniques et l'imagerie par résonance magnétique (IRM), incluant des séquences pondérées en T1 avec injection de gadolinium et des séquences FLAIR. Cependant, l'IRM a un potentiel limité pour différencier les modifications bénignes induites par certains traitements, tels que l'œdème ou les radionécroses des récidives tumorales. Comme le Sestamibi marqué au Technétium 99m s'accumule dans les tumeurs malignes [19], la TEMP s'appuyant sur ce traceur pourrait permettre de fournir des informations complémentaires à celles apportées par l'IRM, comme cela a été suggéré dans une étude sur le suivi de patients atteints de gliomes de haut grade ayant subi une chimiothérapie [20, 21]. Ce traceur permet notamment de distinguer les radionécroses des récidives tumorales. L'étude clinique monocentrique résultant de la collaboration entre les services de neuro-oncologie (J. Delattre, G. Petrirena), de médecine nucléaire (M. Habert, A. Kas, N. Yeni), et de neuro-radiologie (R. Guillevin) du CHU de la Pitié-Sapêtrière a pour objectif de tester si la Tomographie par Emission Mono-Photonique (TEMP) au Sestamibi-Tc99m peut permettre une évaluation thérapeutique chez des patients atteints de gliomes de haut grade recevant un traitement anti-angiogénique. Idéalement cette analyse doit être faite après recalage des images fonctionnelles (TEMP) sur des images anatomiques obtenues par Imagerie par Résonnance Magnétique (IRM) pour combiner les informations morphologiques et fonctionnelles. Pour suivre l'évolution de la réponse tumorale, les examens réalisés un mois après le début du traitement anti-angiogénique doivent également être recalés sur les examens réalisés avant le début du traitement. Cette

dernière étape peut être effectuée en recalant les examens anatomiques acquis par IRM entre eux, puis en appliquant la transformation trouvée aux images TEMP préalablement recalées sur leurs IRM respectives.

4.1.2 Intérêt de la stratégie pour le recalage TEMP-IRM

Pour éviter la tâche fastidieuse du recalage manuel et pour réduire la variabilité intra-et inter-opérateur associée, des algorithmes de recalage automatiques [143, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43] peuvent être utilisés. Le choix optimal et le paramétrage de chaque composante d'un algorithme de recalage dépendent des caractéristiques de l'application : le type d'appariement (données inter- ou intra-patients), la localisation de la zone étudiée, les modalités d'imagerie utilisées. Pour les problèmes de recalage impliquant plusieurs modalités d'imagerie cérébrale, de nombreux critères de similarité ont été proposés, tels que le Ratio de Corrélation (RC) [51], la Corrélation Croisée Normalisée (CCN) [144] , l'Entropie du Coefficient de Corrélation (ECC) [145] , l' Information Mutuelle (IM) [55, 56], et l'Information Mutuelle Normalisée (IMN) [57]. Bien que moins utilisés, certains critères de similarité basés sur des attributs géométriques tels que la Distance de Chanfrein (DC) [48] se sont avérés utiles. Toutefois, même la meilleure méthode globale pour une base de données peut ne pas recaler correctement certains jeux de données spécifiques, pour lesquels d'autres méthodes moins bien classées peuvent potentiellement réussir. Par exemple, dans notre étude, malgré le nombre élevé d'algorithmes de recalage disponibles pour les données cérébrales multimodales, une première évaluation faite par les experts a montré qu'aucune méthode individuelle ne parvient à recaler correctement suffisamment de jeux de données. En effet, la distribution physiologique du Sestamibi (peu de fixation dans le cerveau et hyper-fixations au niveau des plexus choroïdes, des muscles oculomoteurs et de l'hypophyse) peut mettre les méthodes de recalage classiques en échec. La figure 4.1 montre un exemple d'échec de recalage obtenu à l'aide de l'approche de recalage par information mutuelle mis en œuvre dans le logiciel SPM. Sur cette coupe spécifique, des fixations importantes et inattendues de Sestamibi sont observées au niveau de l'œil gauche ainsi qu'à l'extérieur des muscles oculomoteurs de l'œil droit, comme le montrent les flèches blanches. Pour ce jeu de données spécifique, un algorithme basé sur la distance de chanfrein permet d'obtenir un recalage correct. Le logiciel SPM et le recalage par IM obtient de meilleurs résultats sur d'autres jeux de données. Partant de ce constat, la stratégie de recalage présentée au chapitre précédent consistant à utiliser plusieurs méthodes de recalage a été proposée pour tenter de résoudre ce problème de recalage multimodal.

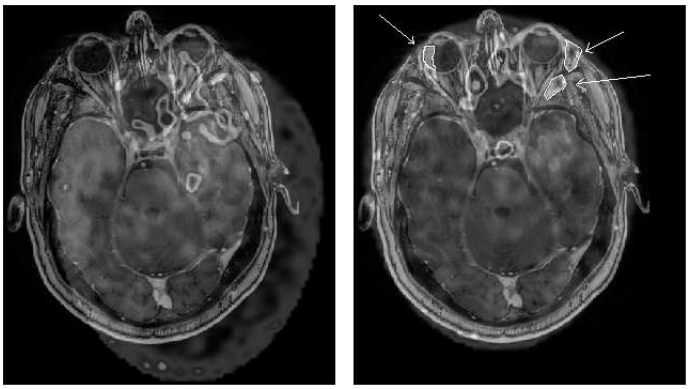


FIGURE 4.1: Fusion des images TEMP et IRM avant recalage (à gauche) et après recalage (à droite). Bien que les contours du crâne de l'image TEMP semblent correspondre à ceux de l'image IRM, le recalage obtenu est insatisfaisant car on peut observer une fixation anormale du traceur dans le globe oculaire (situé à gauche sur l'image) et en dehors des muscles oculomoteurs associés à l'œil droit (zones entourées en blanc et indiquées par les flèches). Dans cet exemple, la TEMP apparaît décalée d'une rotation dans le plan axial (xy) par rapport à l'image IRM.

4.1.3 Base de données

Le protocole comprend seize patients atteints de gliomes de haut grade sous traitement anti-angiogénique. Avant le début du traitement antiangiogénique, chaque patient a subi une IRM pondérée en T1 obtenue après l'injection de gadolinium-DOTA acquise sur un appareil Signa Hdxt 3T (GE), deux examens TEMP respectivement réalisés 15 min et 3 h après l'injection de 740 MBq de Sestamibi marqué au Technétium 99m sur un tomographe Irix (Philips). Les examens IRM et TEMP ont été réalisés à moins de 72 heures d'écart. Le même protocole a été répété un mois plus tard, pour une évaluation des volumes tumoraux un mois après le début du traitement. La base de données sur laquelle j'ai travaillé comporte 62 jeux de données multimodaux (l'IRM après le début du traitement étant manquante pour un patient), chacun étant constitué d'une image TEMP (acquise 15 min ou 3 h après l'injection de Sestamibi) et de son IRM respective. Les volumes TEMP ont été post-filtrés avec un filtre passe-bas d'ordre 4 et une fréquence de coupure de 0,5 cm, puis corrigés de l'atténuation en utilisant la méthode de Chang [146] avec $\mu = 0,12$ cm^{-1}. Pour la correction de l'atténuation, un modèle d'ellipsoïde de la tête a été considéré. Chacun de ces volumes TEMP est constitué de 66 coupes axiales (voxels isotropes de 2,3 mm dans chaque direction). Les volumes IRM sont constitués de 248 coupes axiales (épaisseur 0,69 mm, taille de pixel 0,48 x 0,48 mm^2). La figure 4.2 montre un exemple de jeu de données composé d'une IRM pondérée en T1 après injection de Gadolinium et d'une image TEMP acquise 15 min après l'injection de Sestamibi.

Afin de suivre la réponse tumorale, les examens réalisés un mois après le début du traitement antiangiogénique ont été recalés sur les examens avant le début du traitement en utilisant les images IRM. La transformation ainsi trouvée a alors été appliquée aux images TEMP préalablement recalées sur les IRM correspondantes. On peut donc considérer que la base contient également 15 jeux de données monomodaux IRM-IRM (une IRM après traitement étant manquante pour un patient).

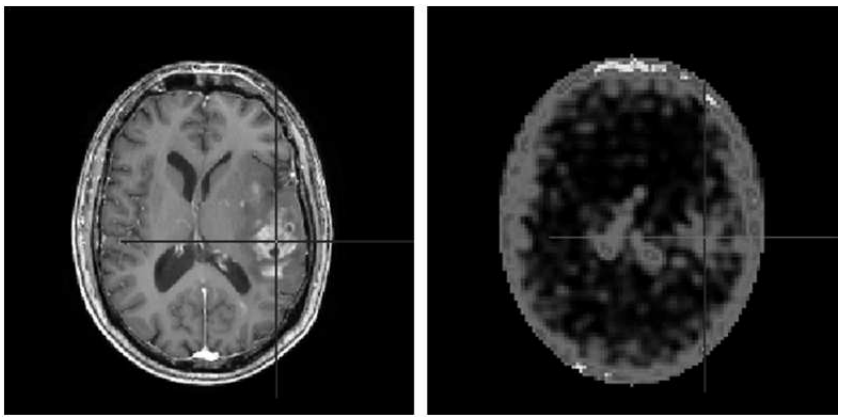


FIGURE 4.2: Exemple de jeu de données. A gauche : coupe axiale du volume 3D IRM pondéré en T1 après l'injection de chélates de Gadolinium montrant une augmentation de contraste dans la tumeur repérée par le curseur. A droite : coupe axiale du volume TEMP (acquis 15 min après l'injection de Sestamibi) après recalage montrant une forte fixation du traceur dans la tumeur et dans les plexus choroïdes.

4.2 Mise en œuvre de la stratégie pour le recalage TEMP-IRM

La mise en œuvre de la stratégie de recalage s'appuie sur trois éléments :

- le choix de méthodes de recalage adaptées à la problématique de recalage (intra-sujet, inter-sujets, intramodale, multimodale, régions déformables, etc.) ;
- le choix de critères d'évaluation quantitatifs adaptés aux données impliquées dans le recalage ;
- la définition de critères visuels qualitatifs adaptés aux données impliquées dans le recalage.

4.2.1 Choix des différentes approches de recalage

Pour cette étude, plus de 30 méthodes différentes ont été testées au cours de cette thèse. Cependant, certaines méthodes n'obtenant que peu de recalages de qualité satisfaisante ont été retirées tandis que d'autres ont été rajoutées au cours du temps. Ainsi les méthodes décrites dans ce chapitre peuvent être différentes d'articles que nous avons publié sur ce travail [147, 148]. Au final, pour la mise en œuvre de la stratégie de recalage nous avons sélectionné 20 méthodes de recalage qui s'appuient sur différentes possibilités quant au choix des paramètres suivants :

- le modèle et le sens de transfomation (rigide, sens direct ou indirect) ;
- le critère de similarité ;
- les prétraitements appliqués aux images référence et source avant recalage ;
- l'algorithme d'optimisation et l'initialisation des paramètres du modèle de transformation choisi.

Le tableau 4.2.1 résume les jeux de paramètres de chaque méthode. Toutes les méthodes ainsi que la stratégie de recalage ont été implémentées dans la version utilisateur du logiciel Anatomist/BrainVisa (*http : //brainvisa.info/index_f.html*). L'ensemble de ces méthodes a été défini à partir d'algorithmes de recalage déjà présents dans les logiciels BrainVisa 4.2.0 [149], SPM 8 [150] et AIR 5.3.0 [50]. Nous avons fait le choix de rendre entièrement automatique toute la procédure qui permet d'exécuter toutes les méthodes

de recalage afin que l'utilisateur n'ait pas à fournir de paramètres.

	M	transfo.	opti.	init.	critère de similarité	prétraitements		sens de transfo.
						IRM	TEMP	
Brainvisa 4.2.0	1	rigide	Powell	cgd	IM	IRM/8	ca	i
	2	rigide	Powell	cdg	IM	IRM/8	ca-skm	i
	3	rigide	Powell	cdg	RC	IRM/8	ca-skm	i
	4	rigide	Powell	cdg	RC	IRM/8	ca	i
	5	rigide	Powell	cdg	DC	IRM/TEMP-s_op_gm	ca-s_gm	d
	6	rigide	Powell	cdg	DC	IRM/TEMP-s_op_gm	ca-s_gm	i
	7	rigide	Powell	cdg	DC	IRM/TEMP-s_op_gm	a-D_sb_gcc	d
	8	rigide	Powell	cdg	DC	IRM/TEMP-s_op_gm	a-D_sb_gcc	i
	9	rigide	Powell	cdg	DC	IRM/TEMP-s_op_gm	ca-D_sb_gcc	d
	10	rigide	Powell	cdg	DC	IRM/TEMP-s_op_gm	ca-D_sb_gcc	i
SPM 8	11	rigide	Powell		IM	IRM/8	ca	d
	12	rigide	Powell		IM	IRM/8	ca	i
	13	rigide	Powell		IMN	IRM/8	ca	d
	14	rigide	Powell		IMN	IRM/8	ca	i
	15	rigide	Powell		ECC	IRM/8	ca	d
	16	rigide	Powell		ECC	IRM/8	ca	i
	17	rigide	Powell		CCN	IRM/8	ca	d
	18	rigide	Powell		CCN	IRM/8	ca	i
AIR 5.3.0	19	rigide	Powell		W	IRM/8	ca	d
	20	rigide	Powell		W	IRM/8	ca	i

4.2.1.1 Modèle de transformation

Comme le problème du recalage est intra-sujet et concerne une région peu déformable, nous avons opté pour un modèle de transformation rigide (3 paramètres de translation T_x, T_y, T_z et 3 angles de rotation (θ_x, θ_y, θ_z).

4.2.1.2 Critères de similarité

Les algorithmes de recalage implémentés dans SPM utilisent des critères iconiques classiques tels que l'information mutuelle (IM), l'information mutuelle normalisée (IMN), l'entropie du coefficient de corrélation (ECC), et la corrélation croisée normalisée (CCN). Les différents algorithmes utilisant ces critères conduisent aux méthodes M11 à M18. Un algorithme basé sur l'information mutuelle (IM), critère iconique basé sur la minimisation de l'entropie conjointe entre 2 images, est également disponible sur BrainVisa avec une implémentation légèrement différente. Cet algorithme a été utilisé pour définir les méthodes M1 et M2. Pour ces méthodes, on a choisi un histogramme conjoint de 128 classes. Les méthodes M3 et M4 sont basées sur le ratio de corrélation RC [51] (disponible dans BrainVisa), critère de similarité iconique développé pour les recalages multimodaux. Le critère de Woods [50] constitue un autre critère de similarité iconique développé

pour recaler des données provenant de différentes modalités. Ce critère implémenté dans le logiciel AIR conduit aux méthodes M19 et M20. Enfin, la distance de chanfrein (DC), critère géométrique mis en œuvre dans BrainVisa et basé sur la minimisation d'une distance entre contours en utilisant une carte de distance (pour minimiser les temps de calcul), a été testé dans six configurations différentes conduisant aux méthodes M5 à M10.

4.2.1.3 Prétraitements appliqués aux images IRM et TEMP avant recalage

Les différents prétraitements que nous présentons dans cette section sont les suivants :
- correction d'atténuation des images TEMP ;
- sous-échantillonnage des données IRM ;
- seuillage des images TEMP ;
- extraction des contours du crâne sur les images TEMP et IRM.

4.2.1.3.1 Correction d'atténuation

Les images TEMP ont été corrigées de l'atténuation à la suite des acquisitions à l'aide du logiciel constructeur en utilisant la méthode de Chang [146] avec $\mu = 0,12$ cm^{-1}. Dans le cadre de cette correction, un masque ellipsoïdal a été placé manuellement par les opérateurs pour distinguer les voxels appartenant à la tête de ceux correspondant à du bruit en dehors de la tête. Par la suite, les images TEMP avant correction d'atténuation **(a)** et après correction d'atténuation **(ca)** sont considérées.

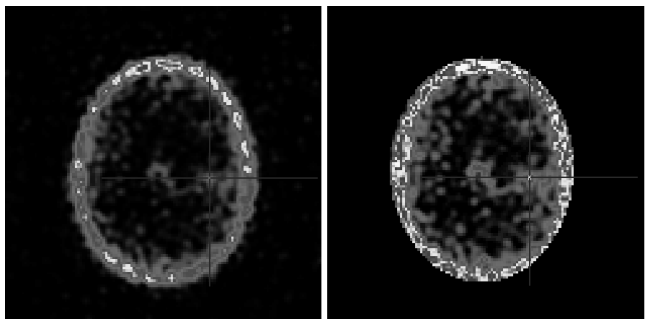


FIGURE 4.3: A gauche : coupe axiale du volume TEMP sans correction d'atténuation. A droite : coupe axiale du volume TEMP après correction d'atténuation par la méthode de Chang ; on aperçoit la troncature liée au modèle ellipsoïdal.

4.2.1.3.2 Sous-échantillonnage et seuillage

Les IRM ont été sous-échantillonnées, soit d'un facteur 8 **(IRM/8)** (2 dans chaque direction) pour les méthodes M1 à M4 et M11 à M20 afin de diminuer les temps de calcul, soit sous-échantillonnées aux dimensions et résolutions de l'image TEMP **(IRM/TEMP)** à l'aide d'un algorithme d'interpolation trilinéaire pour les méthodes basées sur DC afin d'obtenir des contours du crâne de même dimension sur les IRM et

les TEMP. Suréchantillonner l'image TEMP aboutirait à déformer le contour du crâne extrait des images TEMP (en rajoutant de l'information inconnue).

Dans les méthodes M2 et M3, les images TEMP après correction d'atténuation **(ca)** sont seuillées **(s)** avant le recalage en utilisant un algorithme de "k-moyennes" (skm) implémenté dans Brainvisa pour récupérer uniquement des zones de forte fixation du Sestamibi qui correspondent à des zones apparaissant en intense sur les IRM afin que les images des deux modalités se ressemblent un peu plus. De manière pratique, on utilise l'algorithme de "k-moyennes" sur les images TEMP corrigées de l'atténuation en définissant 2 classes (sans compter le fond de l'image) et on récupère les voxels appartenant à la classe ayant la plus forte intensité comme le montre la figure 4.4.

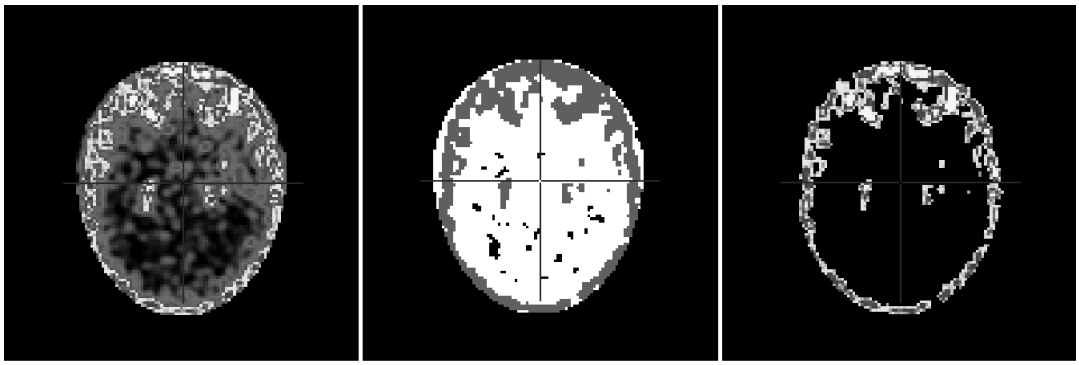


FIGURE 4.4: A gauche : coupe axiale du volume TEMP corrigé de l'atténuation. Au milieu : résultat de la segmentation par k-moyennes avec 2 classes (en bleu la classe d'intensité maximale). A droite : coupe axiale du volume TEMP sur lequel on n'a gardé que les voxels appartenant à la classe de plus forte intensité.

4.2.1.3.3 Extraction des contours du crâne sur l'IRM

Les méthodes M5 à M10 s'appuyant sur la Distance de Chanfrein (DC) nécessitent l'extraction des contours du crâne sur les images TEMP et les IRM après sous-échantillonnage aux dimensions de la TEMP **(IRM/TEMP)**.

La méthode choisie pour l'extraction du contour du crâne sur l'IRM **(s_op_gm)** se décompose en deux étapes. Tout d'abord, le masque du crâne est extrait sur l'image IRM en appliquant un seuillage binaire (dont la valeur du seuil 1000 a été déterminée empiriquement sur les données) pour séparer le fond de l'image et le crâne, puis en procédant à un processus itératif automatique d'opérations morphologiques (ouverture, fermeture, etc.) à l'aide de la fonction VipGetHead implémenté dans Brainvisa [149] pour remplir les trous liés au seuillage et lisser le masque. Le contour du crâne sur l'IRM est ensuite obtenu en appliquant un gradient morphologique (image originale - image originale érodée avec un élément structurant de taille 2 fois supérieure à la résolution du voxel) sur le masque précédemment appliqué (Figure 4.5).

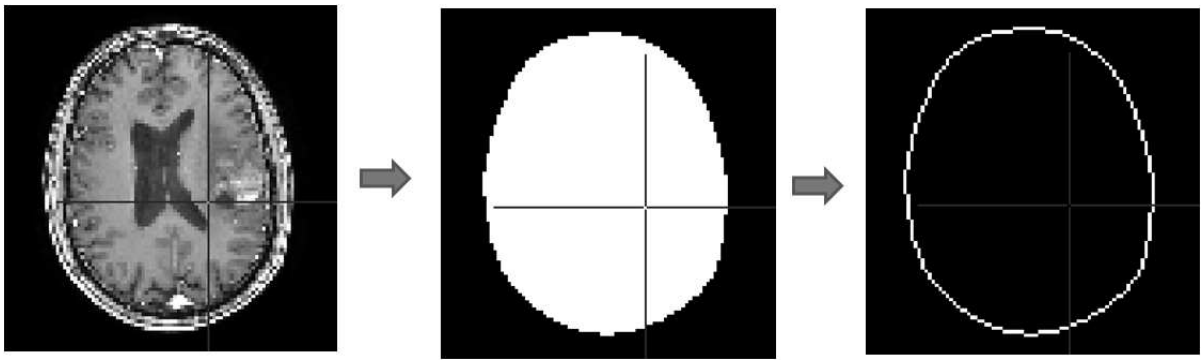


FIGURE 4.5: Processus d'extraction du contour du crâne sur les images IRM. De gauche à droite les étapes sont les suivantes : seuillage de l'image suivi d'un processus itératif d'opérations morphologiques pour récupérer le masque du crâne, et récupération du contour du crâne par un gradient morphologique.

4.2.1.3.4 Extraction des contours du crâne sur la TEMP

Deux méthodes d'extraction des contours du crâne à partir des images TEMP ont été mises en œuvre. Pour les méthodes M5 et M6, les contours sont obtenus par seuillage (> 0) suivi d'un gradient morphologique sur l'image TEMP avec correction d'atténuation **(s_gm)**.

Pour les méthodes M7 et M8 **(D_sb_gcc)**, l'extraction des contours du crâne se décompose en 4 étapes :

- l'utilisation du filtre de Deriche permet de récupérer une image de gradient de l'image TEMP sans correction d'atténuation ;
- un seuillage (≥ 2) est appliqué à l'image de gradient pour ne garder que les contours des structures du crâne ;
- on récupère ensuite les contours externes de l'image de gradient à l'aide d'une méthode de contours actifs sans contrainte de forme (implémentée dans BrainVisa) initialisée aux bords de l'image ;
- on sélectionne la plus grande composante connexe pour récupérer le contour du crâne.

Pour les méthodes M8 et M9 la méthode d'extraction est similaire à celle utilisée pour les méthodes M7 et M8, si ce n'est que le filtre de Deriche est appliqué sur l'image TEMP après correction d'atténuation.

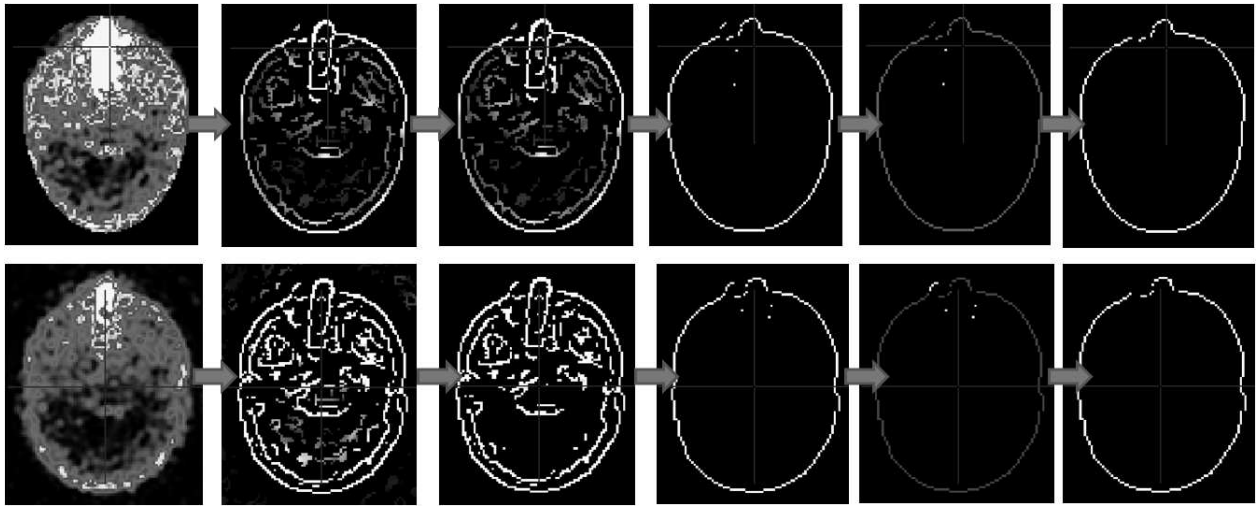


FIGURE 4.6: Processus d'extraction du contour du crâne sur les images TEMP en utilisant le filtrage de Deriche sur l'image corrigée de l'atténuation (haut) et sur l'image sans correction d'atténuation (bas). L'image TEMP est dans un premier temps filtrée à l'aide d'un filtre de Deriche pour obtenir une image de gradient. Celle-ci est ensuite seuillée (≥ 2) pour retirer un certain nombre de structures internes ou externes au crâne. On récupère ensuite les contours externes du crâne à l'aide d'une méthode de contour actif sans contrainte de forme implémentée dans BrainVisa et on extrait la composante connexe la plus importante.

Une dernière étape consiste à supprimer une partie du contour situé au niveau de la base du crâne, sur les premières coupes axiales. En effet, cette partie du contour liée au bord de l'image plutôt qu'à la structure à segmenter risque de biaiser le recalage et d'aboutir à un mauvais résultat. Cette partie du contour est donc supprimée automatiquement en éliminant les deux premières coupes axiales pour ne pas être prise en compte lors du recalage avec les algorithmes basés sur les distances entre contours.

L'ensemble du processus d'extraction de ces contours est entièrement automatique.

4.2.1.4 Sens de transformation

Les méthodes M5, M7, M9, M11, M13, M15, M17, M19 utilisent un sens de transformation direct TEMP->IRM **(d)**, tandis que les autres méthodes utilisent un sens de transformation indirect **(i)** IRM->TEMP qui requiert l'inversion de la transformation trouvée pour estimer la matrice de transformation "TEMP->IRM". La matrice "TEMP->IRM" ainsi trouvée permet de recaler l'image TEMP sur l'image IRM en utilisant un algorithme d'interpolation trilinéaire. Les méthodes M1 à M4 basées sur des critères iconiques n'utilisent que le sens de transformation indirect car l'algorithme implémenté dans BrainVisa conduisait à des temps de calcul trop longs dans le sens direct sans que le résultat soit de meilleure qualité. En théorie, si le critère de similarité est symétrique, les deux sens conduisent à la même matrice de transformation. Cependant, dans la pratique, les deux sens de transformation peuvent conduire à des résultats de recalage très différents (comme

le montre la figure 4.7) ; les deux sens de transformation ont donc été pris en compte dans la stratégie de recalage. Cet élément peut être dû au fait que les algorithmes d'optimisation permettent en général de trouver un extremum local et non global. Cet extremum trouvé dépend alors du chemin emprunté par l'algorithme pour y parvenir. Il est donc cohérent de trouver, même pour un critère de similarité symétrique, des résultats différents si on utilise un sens de transformation ou son inverse.

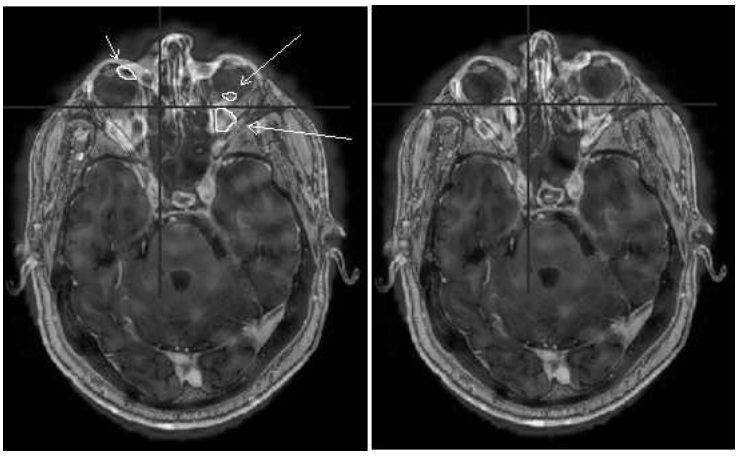


FIGURE 4.7: Exemple de recalages obtenus par 2 méthodes basées sur la distance de chanfrein qui diffèrent uniquement de par le sens de transformation utilisé. A gauche : mauvais recalage obtenu par la méthode M_9 utilisant un sens de transformation direct. Les flèches blanches montrent des fixations anormales dans l'œil gauche et en-dehors du muscle oculomoteur droit. A droite : bon recalage obtenu par la méthode M_{10} utilisant le sens de transformation indirect.

4.2.1.5 Algorithme d'optimisation et initialisation

Pour les méthode de BrainVisa, les paramètres de translation (T_x, T_y, T_z) de la matrice de transformation ont été initialisés à partir du centre de gravité des images **(cdg)** et les angles de rotation (θ_x, θ_y, θ_z) ont été initialisés à une valeur de 0 pour toutes les méthodes. Pour SPM et AIR ces paramètres ne sont pas précisés. L'algorithme d'optimisation de Powell a été systématiquement utilisé pour trouver les extrema des fonctions de coût.

4.2.1.6 Précisions sur le développement de la mise en œuvre des méthodes de recalage

Pour appeler les méthodes provenant d'environnements logiciels différents, nous avons développé une surcouche dans BrainVisa en adaptant les entrées-sorties aux spécificités de chaque environnement. Par exemple, le format des matrices de transformation ainsi que les référentiels dans lesquels elles étaient exprimées sont différents entre les différents environnements et ont du être adaptés à Brainvisa. Nous avons codé cette surcouche logiciel en respectant un aspect de base de données afin de gérer un nombre conséquent de données. De plus, pour rendre le processus de recalage entièrement automatique, tous les paramètres requis ont été

fixés et notamment ceux nécessaires à l'extraction des contours sur crâne sur les examens des modalités IRM et TEMP. Au final, l'utilisateur n'a qu'à entrer les images dans la base de données à l'aide d'une interface graphique Brainvisa et lancer le processus de recalage via une deuxième interface. La gestion des paramètres et des différents environnements pour l'appel des méthodes est complètement transparente pour l'utilisateur.

4.2.2 Evaluation quantitative et classement

Comme la performance d'une méthode de recalage varie d'un jeu de données à un autre, un critère d'évaluation quantitatif a été développé afin de classer les méthodes et déterminer la méthode la plus adaptée à chaque jeu de données. Afin de ne pas biaiser l'évaluation et le classement des méthodes, il était important de choisir un critère d'évaluation indépendant des critères de similarité utilisés dans les méthodes de recalage choisies pour la stratégie, comme l'IM. Pour cette raison, nous avons développé un nouveau critère pour évaluer les performances de chaque méthode M_n. Il s'appuie sur les lieux de fixation physiologique du traceur dans certaines zones anatomiques utilisées visuellement par les experts pour juger de la qualité d'un recalage. En effet, pour procéder à cette évaluation les experts ne peuvent pas se baser sur le recouvrement entre la prise de Gadolinium et la prise de Sestamibi au sein de la tumeur, puisque celles-ci pourraient être différentes en terme d'étendue et que cette différence pourrait s'accroître en raison du traitement anti-angiogénique. C'est pourquoi l'évaluation doit se baser sur des structures en dehors de la zone tumorale mais visibles sur les deux images : l'hypophyse, les muscles oculomoteurs, le nez et les plexus choroïdes ont une prise de Sestamibi importante sur l'image TEMP et sont facilement observables sur l'image IRM. Les globes oculaires sont aussi visibles sur l'image IRM et correspondent à des zones d'hypofixation du Sestamibi sur l'image TEMP. Le contour du crâne est aussi censé être une zone d'hyperfixation du Sestamibi. Bien que l'observation du recouvrement entre les lieux de fixation du Sestamibi et ces zones anatomiques en 3D soit délicate, ces critères visuels semblent robustes et fiables pour juger de la qualité des recalages. C'est pourquoi il paraissait intéressant de traduire ces critères visuels qualitatifs en un critère quantitatif semi-automatique. Les structures d'intérêt citées ci-dessus comme les globes oculaires, les muscles oculomoteurs et l'hypophyse ont donc dû être segmentées en premier lieu.

Une première étape d'extraction des zones anatomiques correspondant aux lieux d'hypofixation et d'hyperfixation du Sestamibi a donc été réalisée. Du fait de la faible résolution spatiale des images TEMP, la segmentation de ces zones sur ces images s'est avéree très compliquée, voire impossible. Nous avons donc choisi d'extraire uniquement ces zones sur les images IRM.

4.2.2.1 Segmentation des globes oculaires sur l'IRM par transformée de Hough

Les globes oculaires ont été segmentés semi automatiquement en utilisant une transformée de Hough 3D sphérique. Cette segmentation se décompose en trois étapes :

- définition d'une région d'intérêt parallélépipédique incluant le globe oculaire ;
- application d'un filtre de Deriche pour récupérer un ensemble de points appartenant au globe oculaire ;
- utilisation de la transformée de Hough pour déterminer le contour de la sphère correspondant au globe oculaire ;

.

4.2.2.1.1 Définition de la région d'intérêt

Une région d'intérêt parallélépipédique incluant le globe oculaire, de taille (18,5 mm en x, 18,5 mm en y et 26,6 mm en z) et centrée sur un point défini manuellement par l'utilisateur a été tout d'abord délimité autour de chaque œil.

4.2.2.1.2 Application du filtre de Deriche

Afin de récupérer les contours grossiers du globe oculaire, un filtre de Deriche [151] ($\alpha = 1$) a été appliqué sur la région d'intérêt. L'extraction de contour avec le filtre de Deriche reprend le principe de l'extraction de contours utilisant un filtre de Canny [152] qui consiste à réduire le bruit dans l'image à l'aide d'un filtre lisseur avant de calculer le gradient de l'image puis d'extraire les maxima locaux du gradient et de distinguer les points qui appartiennent au contour. La différence entre les deux vient de l'utilisation de filtres lisseurs et dérivateurs différents pour débruiter l'image puis récupérer les contours. Le filtre 1D lisseur de Deriche a pour expression :

$$L(n) = \frac{(1-\mathrm{e}^{-\alpha})^2}{(1+2\alpha\mathrm{e}^{-\alpha}-\mathrm{e}^{-2\alpha})}(\alpha|n|+1)\mathrm{e}^{-\alpha|n|} \tag{4.1}$$

Le paramètre α définit la largeur du filtre : plus α est petit, plus on distingue facilement la présence de contours, mais moins la localisation de ces contours est précise. Au contraire, plus α est grand, plus la détection de contours est difficile, mais plus leur localisation est précise. Cette expression a été obtenue par intégration du filtre dérivateur 1D de Deriche (qui correspond à la solution exacte de l'équation de Canny) :

$$D(n) = -\frac{(1-\mathrm{e}^{-\alpha})^2}{\mathrm{e}^{-\alpha}} n\mathrm{e}^{-\alpha|n|} \tag{4.2}$$

La figure 4.8 montre la réponse impulsionnelle de ces filtres.

Ces filtres correspondent à des filtres à Réponse Impulsionnelle Infinie (RII) et peuvent donc être implémentés de manière récursive. Par exemple, on peut décomposer le filtre dérivateur (4.2) en 2 parties causale $\mathrm{d}^+(\mathrm{n})$ et anticausale $\mathrm{d}^-(\mathrm{n})$:

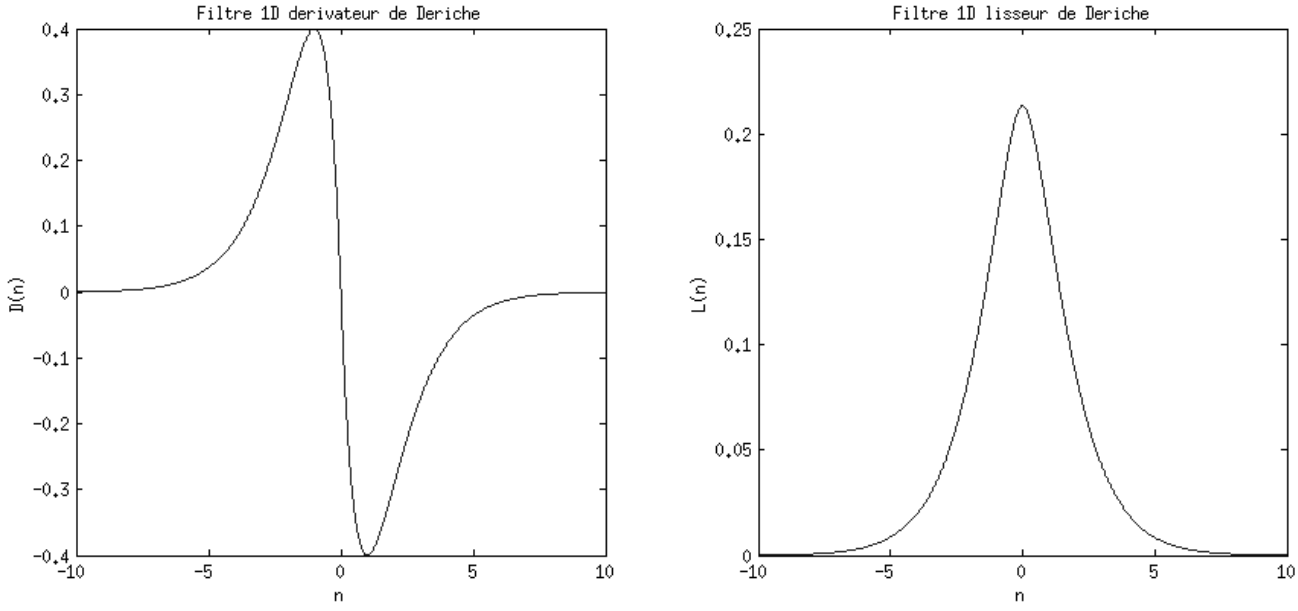
$$\begin{cases} d^+(n) = -\frac{(1-\mathrm{e}^{-\alpha})^2}{\mathrm{e}^{-\alpha}} n\mathrm{e}^{-\alpha n} \text{ si } n \geq 0 \\ d^-(n) = -\frac{(1-\mathrm{e}^{-\alpha})^2}{\mathrm{e}^{-\alpha}} n\mathrm{e}^{\alpha n} \text{ si } n \leq 0 \end{cases} \tag{4.3}$$

Les parties causale $\mathrm{d}^+(\mathrm{n})$ et anticausale $\mathrm{d}^-(\mathrm{n})$ ont alors respectivement pour transformée en z, $\mathrm{D}^+(\mathrm{z})$ et $\mathrm{D}^-(\mathrm{Z})$:

$$\begin{cases} D^+(z) = \frac{Y^+(z)}{X(z)} = -\frac{(1-\mathrm{e}^{-\alpha})^2}{\mathrm{e}^{-\alpha}} \sum_{k=0}^{+\infty} \frac{\mathrm{e}^{-\alpha} z^{-1}}{(1-\mathrm{e}^{-\alpha} z^{-1})^2} \\ D^-(z) = \frac{Y^-(z)}{X(z)} = -\frac{(1-\mathrm{e}^{-\alpha})^2}{\mathrm{e}^{-\alpha}} \sum_{k=0}^{+\infty} \frac{\mathrm{e}^{-\alpha} z}{(1-\mathrm{e}^{-\alpha} z)^2} \end{cases} \tag{4.4}$$

Ces transformées en z conduisent aux équations de récurrence d'ordre 2 suivantes :

$$\begin{cases} y_k^+ = (1-\mathrm{e}^{-\alpha})^2 x_{k-1} + 2\mathrm{e}^{-\alpha} y_{k-1}^+ - \mathrm{e}^{-2\alpha} y_{k-2}^+ \\ y_k^- = -(1-\mathrm{e}^{-\alpha})^2 x_{k+1} + 2\mathrm{e}^{-\alpha} y_{k+1}^+ - \mathrm{e}^{-2\alpha} y_{k+2}^+ \\ y_k = y_k^+ + y_k^- \end{cases} \tag{4.5}$$



(a) Filtre 1D dérivateur de Deriche

(b) Filtre 1D lisseur de Deriche

FIGURE 4.8: Réponse impulsionnelle des filtres 1D dérivateur (a) et lisseur (b) de Deriche avec $\alpha = 1$

4.2.2.1.3 Définition : Transformée de Hough La transformée de Hough est une technique de reconnaissance des formes inventée initialement par PVC Hough [153] dans les années 60 pour détecter des alignements de points. Elle est très utilisée pour rechercher des courbes paramétriques simples comme des lignes ou des cercles dans les images. Cependant plusieurs travaux tels que ceux de Richard Duda et Peter Hart ont permis de généraliser cette méthode pour pouvoir reconnaître des formes quelconques [154, 155]. De manière pratique, la transformée de Hough correspond à une fonction définie dans l'espace des paramètres d'une famille spécifique de courbes paramétriques (droite ou cercle, etc.) dont les coordonnées du maximum de cette fonction correspondent aux paramètres de la courbe paramétrique passant par le plus de points de l'image prétraitée (généralement une image de gradient binarisée). Plus cette fonction a une valeur importante, plus la courbe paramétrée par les coordonnées de cette valeur va passer par un nombre important de points de l'image prétraitée. En résumé, il existe pour chaque famille de courbes au moins une transformée de Hough, et rechercher toutes les courbes d'une famille spécifique revient à chercher les valeurs importantes de la transformée de Hough. Dans sa revue [156], Henri Maître a éclairci la terminologie en proposant une définition mathématique rigoureuse de la transformée de Hough :

"Soit $\mathbb{R}^n$ l'espace image et ε un ensemble de N points sélectionnés par un prétraitement :

$\varepsilon = \{M_i, i = 1...N\} \in \mathbb{R}^n$. Un point M de $\mathbb{R}^n$ est repéré par ses coordonnées $\mathbf{x}$. Soit $\Omega \subset \mathbb{R}^p$ un espace de paramètres et F une famille de courbes dans $\mathbb{R}^n$ paramétrées par $\mathbf{a}$:

$F = \{\{x : f(\mathbf{x},\mathbf{a}) = 0, \mathbf{x} \in \mathbb{R}^n\}, \mathbf{a} \in \Omega\}$

On appelle transformée de Hough associée à la famille F une transformation qui fait correspondre à l'ensemble

ϵ une fonction g définie sur Ω."

D'après cette définition, il existe donc plusieurs transformées de Hough dont H. Maître décrit les deux principales :

- *Les transformées 1 à m* qui s'appuient sur le fait que" par tout point M_i de $\mathbb{R}^n$ passent m courbes de *F*. Soit A_i l'ensemble des valeurs de $\mathbf{a}$ telles que $f(\mathbf{x}_i, \mathbf{a}) = 0$:
  $A_i = \{\mathbf{a}_i : f(\mathbf{x}_i, \mathbf{a}_k) = 0\}$. La transformation de Hough de 1 à m est définie par :
  $g(\mathbf{a}) = \sum_{M_i \in \varepsilon} \sum_{a_k \in A_i} C(a - a_k)$
  avec C(a) la fonction caractéristique de $\mathbb{R}^p$.
  Dans le cas d'une transformée de Hough circulaire, cela revient à chercher tous les cercles qui passent par un point de l'image prétraitée (binaire) sur l'image de base et à répéter cette opération pour tous les points de l'image prétraitée. L'équation cartésienne d'un cercle est la suivante (4.6) :

$$(x - a)^2 + (y - b)^2 - R^2 = 0 \tag{4.6}$$

  avec R le rayon du cercle, a et b les coordonnées du centre du cercle, x et y les coordonnées des points du cercle. L'espace des paramètres est donc un espace de dimension 3. Pour un rayon R fixé, en fixant les coordonnées x et y, et en considérant les coordonnées du centre du cercle a et b variables, on démontre que l'image de tous les cercles passant par un point de coordonnées (x,y) correspond bien à un cercle dans l'espace des paramètres (cf 4.9) et donc à un cône lorsque R varie. Donc chaque point de l'image prétraitée va engendrer un cercle dans l'espace des paramètres. Les coordonnées du point de l'espace des paramètres, commun au plus de cercles dans cet espace correspondront aux paramètres du cercle passant par le plus de points de l'image prétraitée.

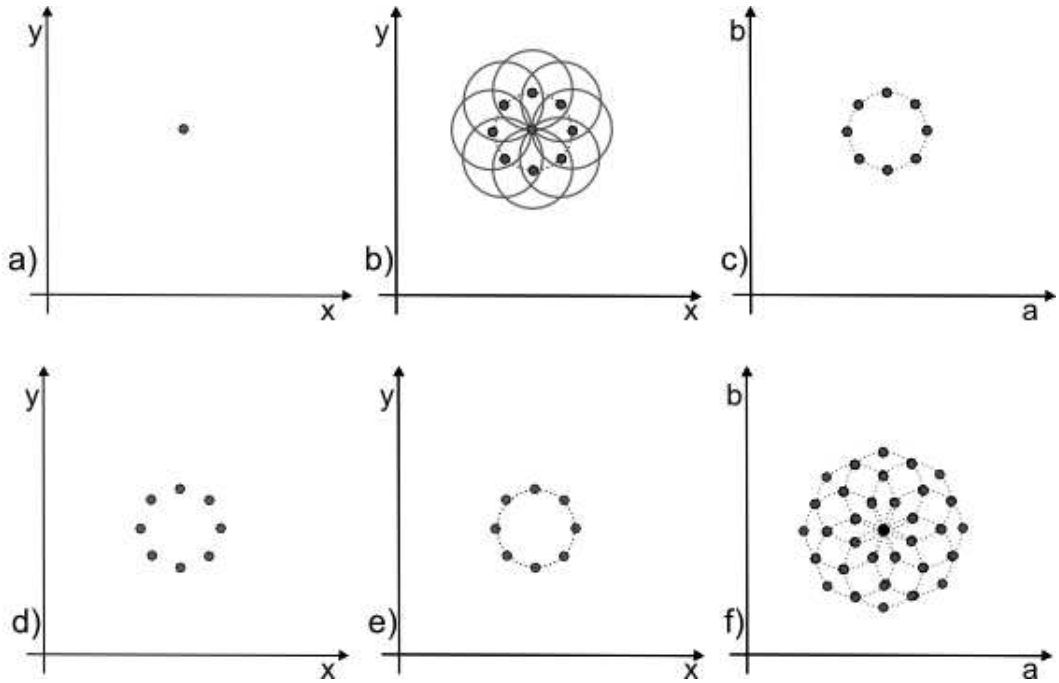


FIGURE 4.9: Transformée 1 à m circulaire. La figure a) représente un point dans l'espace image. En b) on peut voir en bleu un ensemble de cercles passant par ce point et en rouge les centres de tous ces cercles (d'un rayon fixé). En pointillé, on peut voir un cercle formé par tous les centres des cercles (d'un rayon donné) passant par le point de départ. En c) on peut observer le résultat de la transformée de Hough dans l'espace des paramètres (pour un rayon donné) à savoir un cercle constitué de l'ensemble des centres des cercles passant par le point défini dans l'espace image. La figure d) représente un ensemble de 8 points pouvant potentiellement former un cercle. En f) on peut observer la transformée de Hough dans l'espace des paramètres qui correspond à 8 cercles (en pointillés rouges) engendrés par les 8 points de l'espace image. En noir, on peut voir le point de l'espace des paramètres qui est commun au plus de cercles. Ces coordonnées correspondent donc aux paramètres du cercle passant par le plus de points de l'espace image comme on peut le voir en pointillés sur la figure e).

– *Les transformées m à 1.* "Soit m le nombre minimal de points de $\mathbb{R}^n$ définissant une courbe de F. Soit g(m) l'ensemble de tous les m-uples issus de $\varepsilon^{(m)}$ avec $\varepsilon^{(m)} = \{M_i^{(m)} = \{M_{i,1}, M_{i2}, ..., M_{im} : M_{ik} \in \varepsilon\},\}$ avec $Card(\varepsilon^{(m)}) = C_m^n$. A tout m-uple $M_i^{(m)}$ de $\varepsilon^{(m)}$ est associée une courbe de F de paramètre a_i. Soit C(a) la fonction caractéristique de $\mathbb{R}^p$. La transformation de Hough de m à 1 est définie par : $g(\mathbf{a}) = \sum_{M_i^{(m)} \in \varepsilon^{(m)}} C(a - a_i)$."
Dans le cas d'une transformée de Hough circulaire, cela revient à chercher le cercle qui passe par une combinaison de 3 points de l'image prétraitée (binaire) sur l'image de base et de répéter cette opération pour toutes les combinaisons de 3 points possibles de l'image prétraitée. Chacune de ces combinaisons de 3 points dans l'espace image engendre un point dans l'espace des paramètres. Ainsi, de même que pour la transformée de 1 à m, les coordonnées du point de l'espace des paramètres engendré par le plus de combinaisons de 3 points de l'espace de l'image prétraitée correspond aux paramètres du cercle regroupant le plus de points de l'image prétraitée.

.

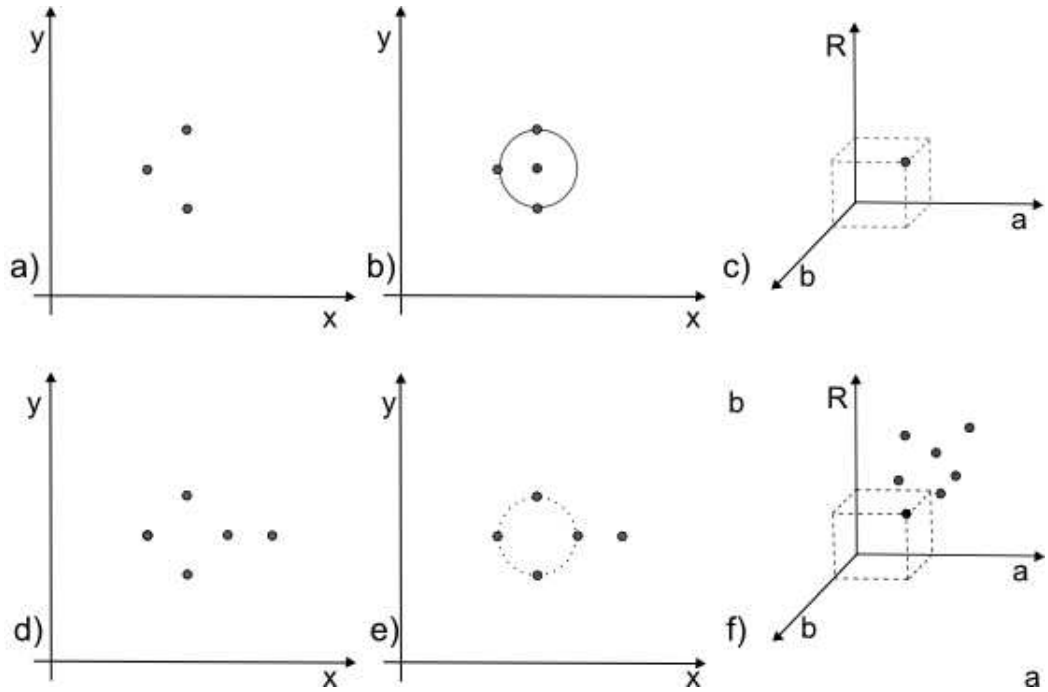


FIGURE 4.10: Transformée m à 1 circulaire. La figure a) représente 3 points dans l'espace image. En b) on peut voir en rouge un cercle passant par ces 3 points point et en rouge le centre de ce cercle. En c) on peut observer le résultat de la transformée de Hough dans l'espace des paramètres à savoir un point correspondant au centre du cercle passant par les 3 points définis dans l'espace image. La figure d) représente un ensemble de 4 points pouvant potentiellement former un cercle. En f) on peut observer la transformée de Hough dans l'espace des paramètres qui correspond à 10 points dont 4 sont superposés (point noir) engendrés par les combinaisons de 3 points de l'espace image. En noir, on peut voir le point de l'espace des paramètres qui est engendré par le plus de combinaison de 3 points. Ces coordonnées correspondent donc aux paramètres du cercle passant par le plus de points de l'espace image comme on peut le voir en pointillés sur la figure e).

De manière pratique, la transformée de Hough de 1 à m conduit généralement à moins de calcul que la seconde et c'est pourquoi quand on utilise le terme générique de transformée de Hough, on fait référence à ce type de transformation.

Dans le cadre de l'extraction des yeux nous avons utilisé une troisième définition différente des deux précédentes :

Soient $\mathbb{R}^n$ l'espace image et ε un ensemble de N points sélectionnés par un prétraitement :
$\varepsilon = \{M_i, i = 1...N\} \in \mathbb{R}^n$.
Soit $\Omega \subset \mathbb{R}^p$ un espace de paramètres et F une famille de courbes dans $\mathbb{R}^n$ paramétrées par $\mathbf{a}$:
$F = \{\{x : f(\mathbf{x},\mathbf{a}) = 0, \mathbf{x} \in \mathbb{R}^n\}, : \mathbf{a} \in \Omega\}$
Soit $\varphi_\mathbf{a}$ l'ensemble des P points de la courbe de la famille F paramétrées par $\mathbf{a}$: $\varphi_\mathbf{a} = \{M_{\mathbf{a},k}, k = 1...P\}$
Soit $\psi_\mathbf{a}$ un ensemble de $\mathbb{R}^n$ de J points dans lequel sont inclus les ensembles $\varphi_{(\mathbf{a})}$ et ε :
$\psi_\mathbf{a} = \{M_{\mathbf{a},j}, j = 1...J\}$ et tel que $\varphi_\mathbf{a} \cup \varepsilon \subset \psi_\mathbf{a}$
La transformée de Hough s'écrit alors :

$$g(a) = \sum_{M_{\mathbf{a},j} \in \psi_\mathbf{a}} S(M_{\mathbf{a},j})$$

avec

$$S(M_{\mathbf{a},j}) = \begin{cases} 1 & \text{si } M_{\mathbf{a},j} \in \varphi_\mathbf{a} \cap \varepsilon \\ 0 & \text{sinon.} \end{cases}$$

Cette dernière définition revient à remplir l'espace des paramètres avec $\mathbf{a} \in \Omega$ et donc dans le cas d'une transformée de Hough sphérique à déterminer par combien de points de l'image prétraitée passe chaque sphère possible de paramètres $\mathbf{a}$ quelconques de Ω (respectivement tous les cercles pour une transformée de Hough circulaire). A priori cela peut paraître peu pertinent d'utiliser une telle tranformation qui semble bien moins optimale que les deux types de transformation cités précédemment.

Cependant cette dernière définition permet la mise en oeuvre d'une méthode de calcul de la transformée de Hough par convolution [157] qui permet d'éviter de passer par des accumulateurs utilisés dans le calcul de transformée de Hough de 1 à m, qui peuvent poser des problèmes de mémoire ou de temps de calcul.

4.2.2.1.4 Mise en œuvre de la transformée de Hough pour détecter une sphère

Une sphère de rayon R avec un centre de coordonnées cartésiennes (a,b,c) a pour équation :

$$(x-a)^2 + (y-b)^2 + (z-c)^2 - R^2 = 0 \tag{4.7}$$

L'espace des paramètres de toutes les sphères possibles est un sous ensemble de $\mathbb{R}^4$. Soient V_1 le volume d'image prétraitée binaire (1 pour les points et 0 ailleurs) de taille (I,J,K) et V_2 un volume binaire de sphère de taille (2R+1,2R+1,2R+1) de rayon R et de centre (R,R,R). Le volume V_2 ayant une infinité de plans de symétrie, la valeur g(a,b,c,R) correspond à la somme de l'intensité de tous les voxels de la multiplication du volume V_1 et du volume V_2 translaté du vecteur (a-R, b-R, c-R). Donc pour un R donné, remplir l'espace des paramètres revient à effectuer une corrélation 3D entre les volumes V_1 et V_2.

La corrélation entre deux volumes d'images 3D I_1 de taille (N_x, N_y, N_z) et I_2 de taille (M_x, M_y, M_z) est décrite par l'équation suivante :

$$I_1 o I_2(x,y,z) = \sum_{i=\frac{-N_x-M_x}{2}}^{\frac{N_x+M_x}{2}-1} \sum_{j=\frac{-N_y-M_y}{2}}^{\frac{N_y+M_y}{2}} \sum_{i=\frac{-N_z-M_z}{2}}^{\frac{N_x+M_x}{2}} I_1(i,j,k) I_2(x+i, y+j, z+k)$$

Remarque : le résultat de la corrélation correspond à un volume d'image de taille ($N_x + M_x - 1, N_y + M_y - 1, N_z + M_z - 1$). Or dans le cadre du remplissage de l'espace des paramètres, seul le volume d'images de taille (N_x, N_y, N_z) correspondant à la partie centrale du volume de corrélation nous intéresse comme le montre la figure pour un exemple de corrélation 2D.

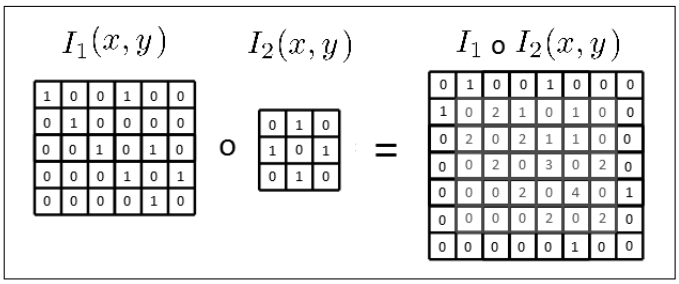


FIGURE 4.11: Corrélation de 2 images. La matrice en rouge correspond à la partie centrée du produit de corrélation de même taille que l'image I_1 et où le premier coefficient est égal à la somme du produit des voxel de l'image I_1 et de l'image I_2 translatée de 1 voxel vers la gauche et vers le haut.

La convolution entre deux volumes d'images 3D I_1 de taille (N_x, N_y, N_z) et I_2 de taille (M_x, M_y, M_z) est décrit par l'équation suivante :

$$I_1 * I_2(x,y,z) = \sum_{i=\frac{-N_x-M_x}{2}}^{\frac{N_x+M_x}{2}-1} \sum_{j=\frac{-N_y-M_y}{2}}^{\frac{N_y+M_y}{2}} \sum_{i=\frac{-N_z-M_z}{2}}^{\frac{N_x+M_x}{2}} I_1(i,j,k) I_2(x-i, y-j, z-k)$$

La convolution entre un volume I_1 et un volume I_2 correspond donc à une corrélation entre ce même volume I_1 et le volume I_2 inversé en x, y et en z, $(I_{2inv}(x,y,z) = I_2(X-x, Y-y, Z-z)$ avec I_{2inv} et I_2 des volumes de taille (X,Y,Z)).

Comme V_2 possède les trois plans de symétries x = R , y = R et Z = R, la corrélation des volumes V_1 et V_2 est égale à la convolution de ces deux mêmes volumes.

En pratique, l'espace des paramètres n'est pas continu et il faut donc choisir un pas pour chaque paramètres. Pour les coordonnées du centre de la sphère (a,b,c) nous avons choisi des pas égaux à la résolution du voxel pour chacune des directions et des valeurs possibles allant de 0 à la taille de l'image -1 dans la direction choisie. Pour le rayon de la sphère, nous avons choisi un pas égal à la résolution en z de l'image (0.69 mm) pour des valeurs comprises entre 10 et 12 mm pour tenir compte des variations de la taille de l'oeil.

Afin de remplir l'espace des paramètres (a,b,c,R) pour un rayon R donné, il est nécessaire de construire V_2, le volume de la sphère à convoluer avec le volume V_1 de l'image prétraitée. Comme la résolution en x et

y est différente de celle en z, le volume V_2 correspond à un volume ellipsoïdal plutôt qu'à une sphère. Pour respecter ces conditions le volume V_2 est construit de la manière suivante :

- Choix de la dimension de V_2 : Soient r_x , r_y et r_z les résolutions des voxels de V_1 en x, y et z. Soit ceil la fonction qui associe à un nombre réel le nombre entier directement supérieur. Pour un R en mm multiple de r_z compris entre 10 mm et 12 mm, la dimension de V_2 est de (2 ceil(R / r_x)+1, 2 ceil(R / r_y)+1, 2 ceil(R / r_z)+1)
- Construction d'un volume de distance au carré de même dimension que V_2 : l'intensité en chaque voxel de ce volume de distance est définie par l'équation :

$$D(x,y,z) = (x - ceil(R/r_x) \text{ x } r_x)^2 + (y - ceil(R/r_y) \text{ x } r_y)^2 + (z - ceil(R/r_z) \text{ x } r_z)^2 \tag{4.8}$$

- Récupération des contours de l'ellipsoïde à partir de la carte de distance D par un double seuillage binaire (qui associe la valeur 1 si l'intensité est entre les deux seuils fixés et 0 sinon).

$$V_2(x,y,z) = ((R - r_z)^2 < D(x,y,z) < R^2)$$

Pour chaque valeur de R on effectue la convolution entre V_1 et V_2 pour remplir l'espace des paramètres. Pour chaque R les valeurs de convolution sont divisées par R pour éviter d'accorder plus d'influence aux cercles de grand rayon qui possèdent généralement plus de points communs avec l'image prétraitée sans que ces points n'appartiennent à la forme recherchée. Les coordonnées du maximum de l'espace des paramètres nous permettent de déterminer les paramètres de l'ellipsoïde qui correspond aux contours du globe oculaire recherché.

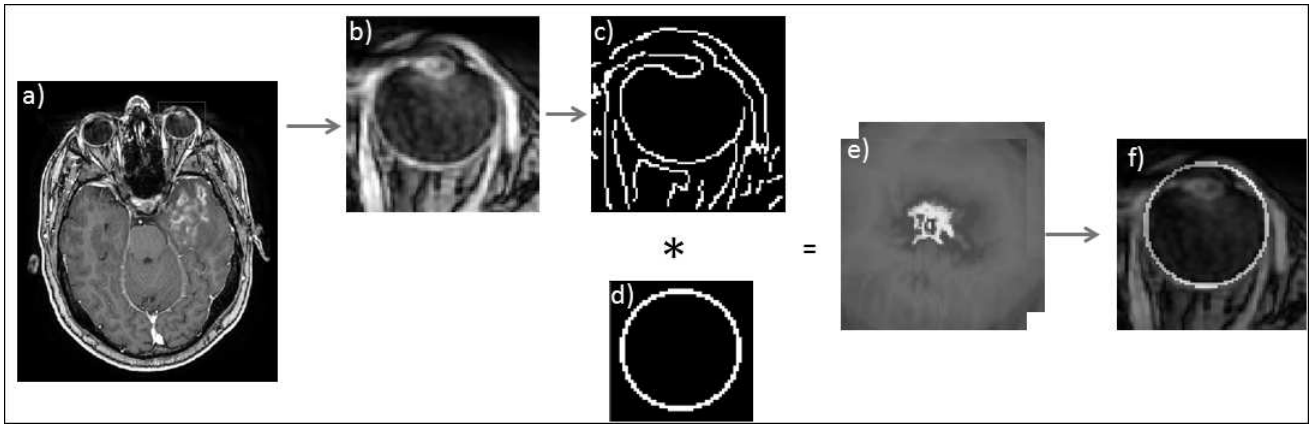


FIGURE 4.12: Extraction des globes oculaires. L'image a) représente une coupe axiale de l'IRM sur laquelle l'encadré rouge représente les limites de la région d'intérêt défini pour la recherche du globe oculaire. En appliquant un filtre de Deriche à la région d'intérêt b) et en seuillant le résultat on obtient l'image prétraitée c). Pour chaque rayon possible, on crée une image de sphère d) et on effectue la convolution entre cette image d) et l'image prétraitée c). Le résultat de ces convolutions correspond à la figure e). A partir des coordonnées du maximum du résultat de convolution on en déduit les paramètres de la sphère correspondant aux contours du globe oculaire comme le montre la figure f).

Dans le but de réduire le temps de calcul de la convolution, nous avons passé ces deux volumes dans

l'espace de Fourrier en utilisant un algorithme de transformée de Fourrier rapide [158]. En effet la convolution entre deux volumes dans l'espace natif correspond à une multiplication dans l'espace de Fourrier.

4.2.2.2 Segmentation des muscles oculomoteurs par "ligne de partage des eaux" (IFT)

Les muscles oculomoteurs ont été segmentés en utilisant une méthode de "ligne de partage des eaux" nécessitant le positionnement manuel par l'utilisateur de marqueurs pour son initialisation [159, 160] sur une région d'intérêt parallélépipédique manuellement définie. Sa mise en œuvre est basée sur un algorithme appelé "image foresting transform" (IFT) qui permet de trouver le chemin le plus court entre un voxel quelconque et le voxel correspondant au marqueur le plus proche.
Remarque : Le coût d'un chemin ne correspond pas à une distance euclidienne entre 2 points mais se base sur l'intensité des voxels entres ces 2 points. Le voxel considéré prend alors la même étiquette que son marqueur le plus proche.

4.2.2.2.1 Description de l'algorithme IFT R. Lotufo et A.Falcao ont décrit l'algorithme de "ligne de partage des eaux" en utilisant l'algorithme d'IFT comme indiqué ci dessous[159] :

Afin de décrire l'algorihme il est nécessaire dans un premier temps de définir quelques éléments. Considérons un volume d'image V de dimension X,Y,Z comme un graphe G(N,A) non orienté avec N un jeu de $X.Y.Z$ nœuds et A un jeu de $X.Y.(Z-1)+X.Z.(Y-1)+Y.Z.(X-1)$ arcs définis par un voisinage de connexité 6 (en 3D). Soit w(p,q) le poids de l'arc entre les noeuds p et q, défini par $w(p,q)=I(q)$ avec I(q) l'intensité du volume d'image D au nœud q. Le coût d'un chemin allant d'un nœud p vers un nœud q est défini par le vecteur $C(p,q)=(L,d)$. L correspond au maximum des poids des arcs des nœuds adjacents qui composent le chemin (p,q). En décomposant le chemin (p,q) en un ensemble d'arcs de nœuds adjacents (voisins en 6 connexités), $(p,q)=(n_1,n_2,...,n_m)$ avec $n_1=p$, $n_m=q$ et $n_i \neq n_j$ pour $i \neq j$, L est alors défini par $L=L(m)=max(w(n_i,n_j))$. d correspond à la distance minimale entre le dernier nœud du chemin et le nœud le plus proche ayant un coût de passage plus faible. Autrement dit si d a une valeur supérieure à 1 cela signifie qu'on a plusieurs nœuds successifs reliés par des arcs de mêmes poids ce qui indique que l'algorithme a atteint un plateau. d est défini par $d=min\{j : L(i)=L(i-j), j=0,...,i-1\}$. Le paramètre L a une priorité plus importante que le paramètre d, c'est à dire que lorsqu'on compare deux chemins on commence par comparer leurs valeurs de L et si celles-ci sont égales, on compare les valeurs de d de chacun des chemins pour pouvoir déterminer lequel a le poids le plus important. Le paramètre L peut alors être vu comme la hauteur d'un mur entre 2 nœuds et le paramètre d comme un phénomène d'étalement de l'eau lorsque le chemin atteint un plateau. d a une valeur nulle en dehors des plateaux. Le principe de l'algorithme "Image Foresting Transform" est de trouver le chemin le plus court qui relie chaque voxel à son nœud de départ le plus proche afin de leur attribuer la même étiquette. Dans cette application les nœuds de départ sont définis par des marqueurs placés manuellement par l'utilisateur.

Soit $C(p)$ le coût du chemin allant du nœud p au marqueur le plus proche, $E(p)$ l'étiquette du nœud p, $DP(p)$ un booléen qui indique si le nœud a été étiqueté définitivement ($FAIT$) ou si l'étiquetage peut encore changer ($TEMP$). Soit liste(p), une liste dans laquelle les noeuds possédant les coûts les plus faibles de chemin sont traités en priorité avec une configuration FIFO (premier élément entré, premier élément sorti) qui

permet de prendre en compte intrinsèquement le paramètre d sans avoir à le calculer. Soit $E(p)$ l'étiquette du nœud p. Soient les fonctions EntreListe$(p, C(p))$ qui permet d'entrer un nouveau nœud p dans la liste à une position déterminée en fonction de son coût $C(p)$, RetireListe(p) qui permet de retirer un nœud p de la liste des priorités, RetireListeMin() qui permet de renvoyer le premier élément de la liste des priorités et de le retirer de la liste, EstDansListe(p) qui permet de vérifier si un nœud se trouve dans la liste, ListeNonVide() qui permet de savoir si la liste de priorités est vide. L'algorithme "image foresting transform" est défini [159] par :

1. Initialisation
   (a) Pour tout nœud p :
   $DP(p) = TEMP$;
   (b) Pour tout nœud p non marqueur :
   $C(p) =$ infini ; $E(p) = 0$;
   (c) Pour tout nœud p marqueur :
   $C(p) = 0$; EntreListe(p,0) ; $E(p) =$ étiquette du marqueur ;

2. Propagation

   Tant que ListeNonVide() :
   (a) $v =$ RetireListeMin() ;
   (b) $DP(v) = FAIT$;
   (c) Pour tout nœud p dans le voisinage (6 connexités) de v avec DP(p) == TEMP :
   Si max($C(v)$, w(v, p)) $< C(p)$:
      i. $C(p) =$ max($C(v)$, w(v, p)) ; $E(p) = E(v)$;
      ii. Si EstDansListe(p) :
      RetireListe(p) ;
      iii. Entreliste$(p, C(p))$;

Durant l'étape d'initialisation, on définit pour tous les nœuds un étiquetage temporaire (a), on fixe le coût de chaque nœud non marqueur à l'infini et on leur donne le label 0 (b) puis on fixe le coût des marqueurs à 0, on initialise leurs étiquettes et on entre les marqueurs dans la liste des priorités (c).

Durant l'étape de propagation, on repète un certain nombre d'étapes tant que la liste des priorités n'est pas vide. On commence par retirer de la liste des priorité le nœud p ayant le coût le plus faible (a). Comme la liste a une organisation FIFO, cela revient à retirer le premier élément de la liste (b). Une fois le nœud v retiré, on indique qu'il a désormais son étiquetage permanent (b). Cela signifie, qu'un nœud v a son étiquetage définitif lorsqu'il sort de la liste des priorités. Ensuite pour chaque nœud voisin p ayant un étiquetage

temporaire du nœud v sélectionné, on vérifie que le coût du nœud p soit supérieur à la valeur maximale entre le coût du nœud v et le poids de l'arc $w(v,p)$ liant les nœuds v et p. Si tel est le cas, le nouveau cout du nœud p prend la valeur maximale entre le coût du nœud v et le poids de l'arc $w(v,p)$ liant les nœuds v et p (c). Si le nœud p était déjà dans la liste des priorités, on le retire. Enfin, on place le nœud p dans la liste des priorités en fonction de son coût, avant ceux possédant un coût plus élevé et après ceux possédant un coût plus faible ou égal. Cela permet de garder la structure FIFO de la liste des priorités, et le fait de placer le nouveau nœud après ceux possédant un coût égal permet de prendre en compte le paramètre d sans avoir à le calculer.

4.2.2.2.2 Etape de la segmentation des muscles oculomoteurs Pour segmenter les muscles oculomoteurs, il faut dans un premier temps définir une région d'intéret parallélépipédique englobant les muscles oculomoteurs. En pratique, l'utilisateur définit cette boite englobante sur les coupes axiales en plaçant 2 points correspondant à deux sommets du parallélépipède. Le premier point doit être placé sur une coupe axiale au dessus des yeux en haut à gauche de l'œil gauche et le second point doit être placé sur une coupe axiale en dessous des yeux en bas à droite de l'œil droit à hauteur de l'hypophyse de manièrre à définir un parallélépipède englobant totalement les muscles oculomoteurs comme le monte la figure 4.13 a).

Ensuite il faut définir les marqueurs pour initialiser l'algorithme de "ligne de partage des eaux". Une étiquette ayant pour valeur 0 est automatiquement placé dans un voxel du fond de l'image. Chaque œil possède 4 muscles oculomoteurs droits (au dessus, en dessous, à gauche et à droite de l'œil concerné). L'utilisateur doit placer au minimum un marqueur dans chacun de ces muscles. A chaque fois qu'un utilisateur clique dans une zone, il place un marqueur, lui attribut une étiquette (1 ou 2) (cf figure 4.13 b)) et déclenche l'application de l'algorithme de "ligne de partage des eaux". Les 4 muscles oculomoteurs droits d'un œil sont censés avoir la même étiquette. Si l'utilisateur n'est pas satisfait du résultat de segmentation ou n'a pas placé de marqueur dans chaque muscle, il peut définir de nouveaux marqueurs ce qui réappliquera l'algorithme en prenant en compte les anciens marqueurs plus ceux nouvellement définis. De même, l'utilisateur peut retirer d'anciens marqueurs.

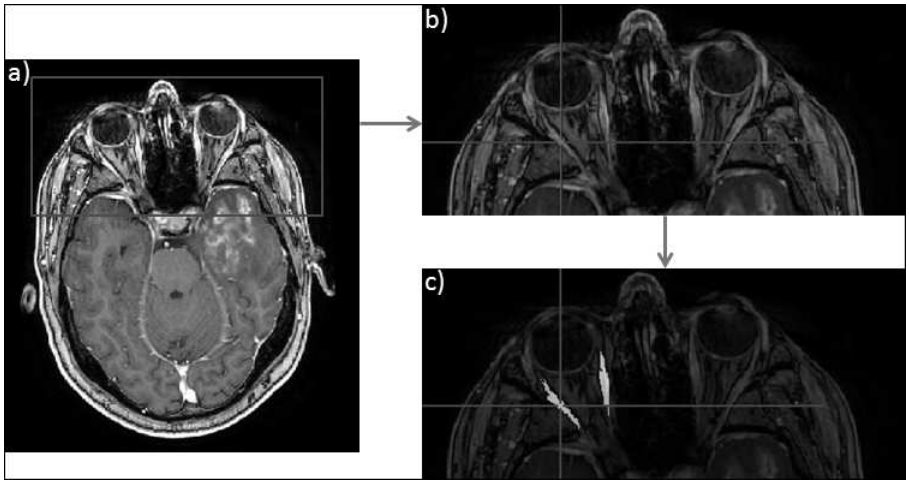


FIGURE 4.13: Extraction des muscles oculomoteurs. L'image a) représente une coupe axiale de l'IRM sur laquelle l'encadré rouge représente les limites de la région d'intérêt définie pour la recherche des muscles oculomoteurs. Le curseur sur l'image de région d'intérêt b) pointe à la position du marqueur pour l'initialisation de l'algorithme de "ligne d partage des eaux" par IFT. La figure c) représente le résultat pour les 4 muscles oculomoteurs droits associés à l'œil gauche.

4.2.2.3 Segmentation de l'hypophyse

L'hypophyse a été segmentée de manière semi-automatique en positionnant manuellement une région d'intérêt cylindrique incluant l'hypophyse.

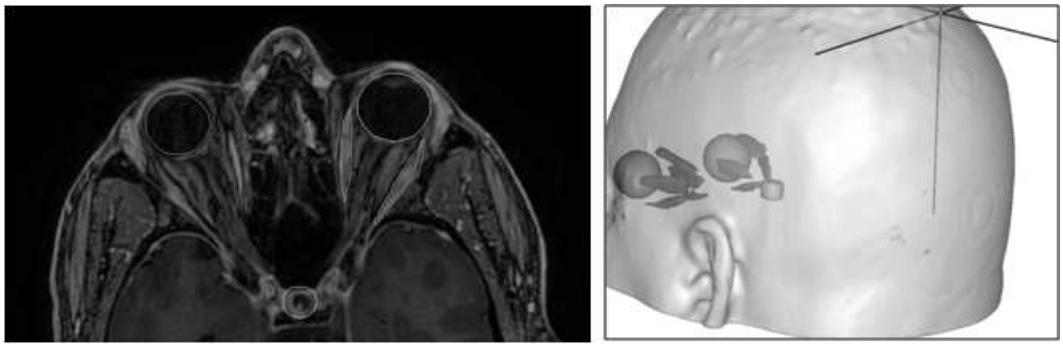


FIGURE 4.14: Segmentation des globes oculaires (en vert), des muscles oculomoteurs (en rouge et orange) et de l'hypophyse (en jaune) sur une image IRM. A gauche : vue axiale. A droite : visualisation 3D

Les différentes régions segmentées ont été étiquetées et stockées dans un masque. Ce masque a ensuite été sous-échantillonné aux dimensions de l'image TEMP en utilisant une interpolation selon les plus proches voisins.

4.2.2.4 Critère quantitatif d'évaluation

Les structures anatomiques citées précédemment ont pu être segmentées sur l'IRM avec une précision satisfaisante, cependant, à cause de la distribution spécifique du Sestamibi et de la faible résolution de la modalité TEMP, la segmentation des lieux d'hyper et d'hypofixation du traceur est très compliquée voire impossible. En effet, la captation de Sestamibi peut être partielle ou s'étendre en dehors des structures anatomiques dans d'autres cas. Pour ces raisons, les critères de recouvrement entre structures ou de distance entre contours classiquement utilisés pour l'évaluation d'un recalage monomodal ne sont pas adaptés au problème. C'est pourquoi il a été nécessaire d'exploiter un autre type d'information comme l'intensité des pixels de l'image TEMP dans les zones correspondant à celles délimitées sur l'IRM pour définir un critère d'évaluation nommé UC pour "Uptake Criterion". Ainsi pour traduire l'hyperfixation et l'hypofixation au sein des structures anatomiques précédemment segmentées, nous avons fait l'hypothèse suivante : une image TEMP correctement recalée présente une intensité moyenne élevée dans les zones d'hyperfixation et une intensité moyenne faible dans les zones d'hypofixation. Ces zones d'hyperfixation et d'hypofixation correspondent aux régions segmentées sur l'IRM et présentes sur le masque étiqueté. Ainsi, pour chaque jeu de données d, le score $UC(d, M_m)$ de chaque méthode de recalage M_n a été calculé d'après l'équation (4.9) :

$$UC(d, Mn) = \frac{1}{Card(V_{hu}(d))} \sum_{p \in V_{hu}(d)} I(d, M_m, p) - \frac{1}{Card(V_{lu}(d))} \sum_{p \in V_{lu}(d)} I(d, M_m, p) \tag{4.9}$$

- $I(d, M_m, p)$ étant l'intensité du voxel p de l'image TEMP recalée du jeu de données d ;
- $Vhu(d)$ le volume correspondant aux régions d'hyperfixation (High Uptake) segmentées sur l'IRM et reportées sur l'image TEMP recalée par la méthode M_n du jeu de données d. $Vhu(d)$ comprend la segmentation de l'hypophyse et des muscles oculomoteurs.
- $Vlu(d)$ le volume correspondant aux régions d'hypofixation (Low Uptake) segmentées sur l'IRM et reportées sur l'image TEMP recalée par la méthode M_n du jeu de données d. $Vlu(d)$ comprend la segmentation des globes oculaires.

Le critère $UC(d, M_m)$ est supposé atteindre une valeur maximale pour l'image TEMP la mieux recalée pour le jeu de données d.

4.2.2.5 Classement des méthodes de recalage

Pour chaque jeu de données d, les valeurs de UC ont été calculées pour les 18 méthodes de recalage afin de pouvoir les classer. La méthode ayant obtenu la valeur la plus élevée de UC a été classée au rang 1 et est considérée comme la meilleure méthode $M^*(d)$ de recalage à utiliser (pour le jeu de données d considéré) :

$$M^*(d) = argmax_{M_m}(UC(d, M_m) \tag{4.10}$$

- avec d le jeu de données considéré ;
- M_m la méthode de recalage numéro m ($1 \leq m \leq 18$) ;
- $UC(M_m)$ la valeur du critère d'évaluation UC ("Uptake Criterion") obtenue par la méthode M_m pour le jeu de donnée d.

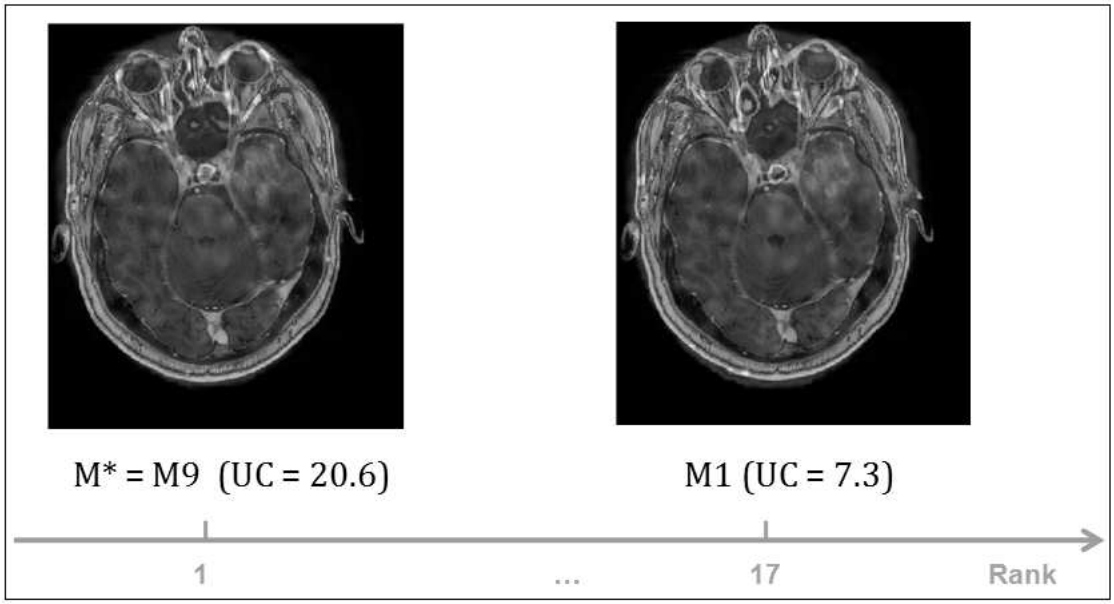


FIGURE 4.15: Résultats visuels et score $UC(d, M_m)$ du recalage obtenu par la méthode M_9 considérée comme la meilleure méthode $M^*(d)$ pour le jeu de données d et de la méthode M_1 classée 17ème pour ce jeu de données.

Cependant, le classement n'assure pas que la meilleure méthode permette d'obtenir un recalage satisfaisant, il assure juste que la méthode M^* obtient un meilleur recalage que les autres méthodes. On peut considérer UC comme étant un critère d'évaluation des performances "relatives" d'une méthode par rapport à d'autres.

Remarque : il se peut qu'une autre méthode que celle classée au rang 1 obtienne un score de UC proche de celui obtenu par cette dernière pour un même jeu de données. Dans un tel cas, des indicateurs présentés au chapitre suivant dans la partie portant sur la validation du critère UC pourront être utilisés pour déterminer si les méthodes sont équivalentes. Cependant, dans la mise en œuvre de la stratégie, seule la validation de la méthode classée au rang 1 est nécessaire.

4.2.3 Validation visuelle

La validation visuelle par les experts consiste en une évaluation qualitative du recalage obtenu par la méthode M^* pour le jeu de données d considéré. Les critères visuels utilisés pour cette évaluation incluent les critères décrits plus haut mais également le recouvrement entre d'autres zones d'hyperfixation du Sestamibi et des zones anatomiques comme les plexus choroïdes, le contour du crâne et le nez. A chaque recalage obtenu par la méthode M^*, un indice de qualité a été attribué par un expert : excellent, correct, ou mauvais. La figure 4.16 montre un exemple de recalage excellent et un exemple de recalage de qualité insatisfaisante.

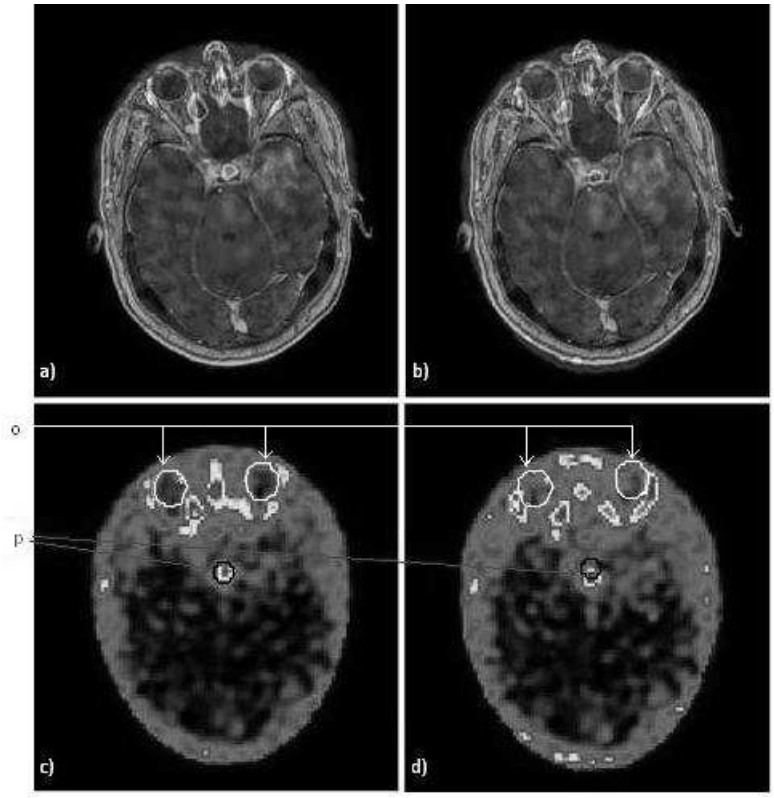


FIGURE 4.16: Recalage TEMP-IRM sur un jeu de données spécifique qui montre : a) un recalage excellent obtenu par la méthode M* (M9 ici) b) un recalage de qualité insatisfaisante (obtenu par M1 ici). Les zones d'hyperfixation (en noir) et d'hypofixation (en blanc) segmentées sur l'image IRM et reportées sur l'image TEMP recalée sont représentées en utilisant c) M9 et d) M1 (p : hypophyse, o : globes oculaires)

4.3 Mise en œuvre de la stratégie pour le recalage longitudinal IRM-IRM

Cette partie consiste à recaler les images IRM acquises un mois après le début du traitement sur les images IRM acquises avant le début du traitement. Les transformations ansi trouvées seront réappliquées aux images TEMP après traitement, déjà recalées sur les IRM après traitement dans le but de les comparer avec les images TEMP avant traitement comme sur la figure 4.17.

Le recalage monomodal IRM-IRM pour les données cérébrales en intra patient est un problème bien connu pour lequel de nombreuses méthodes de recalage ont été proposées. D'après la littérature [56], des méthodes classiques comme celles basées sur la corrélation croisée normalisée ou l'information mutuelle par exemple sont supposées obtenir de très bons résultats pour ce type d'application. Néanmoins il est possible que l'évolution de la tumeur puisse biaiser le recalage, c'est la raison pour laquelle la stratégie de recalage a été utilisée. Afin d'adapter la stratégie aux spécificités de cette application nous avons défini un nouveau

	M	transfo.	opti.	init.	critère de similarité	prétraitements		sens de transfo.
						IRM avant traitement	IRM après traitement	
Brainvisa 4.2.0	1	rigide	Powell	cdg	DC	IRM/TEMP-s_op_gm	IRM/TEMP-s_op_gm	d
	2	rigide	Powell	cdg	DC	IRM/TEMP-s_op_gm	IRM/TEMP-s_op_gm	i
SPM 8	3	rigide	Powell		IM	IRM/8	IRM/8	d
	4	rigide	Powell		IM	IRM/8	IRM/8	i
	5	rigide	Powell		IMN	IRM/8	IRM/8	d
	6	rigide	Powell		IMN	IRM/8	IRM/8	i
	7	rigide	Powell		ECC	IRM/8	IRM/8	d
	8	rigide	Powell		ECC	IRM/8	IRM/8	i
	9	rigide	Powell		CCN	IRM/8	IRM/8	i
	10	rigide	Powell		CCN	IRM/8	IRM/8	i

TABLE 4.1: Definition des 10 méthodes de recalage.
"cdg" une initalisation consistant à aligner les *centres de gravité* des deux images ;
"IM", l'*information mutuelle* ;
"DC", la distance de chanfrein ;
"RC", le *ratio de corrélation* ;
"IMN" l'*information mutuelle normalisée* ;
"ECC" l'*entropie du ratio de corrélation* ;
"CCN" la *corrélation croisée normalisée* ;
"IRM/8" correspondant à *un sous-échantillonnage de l'image IRM d'un facteur 8* et **"IRM/TEMP"** d'un *sous-échantillonnage aux dimensions de l'image TEMP* ;
"s_op_gm", signifiant qu'une étape de *seuillage* suivie d'*opération morphologique* et d'un *gradient morphologique* ont été nécessaires pour estimer les contours du crâne sur l'IRM, **"d"** et **"i"**, sens de transformation *direct* et *indirect*.
Pour les prétraitements, les acronymes sont combinés en utilisant "-".

critère d'évaluation quantitatif ainsi que les critères visuels mais le principe de la stratégie reste le même. En ce qui concerne les méthodes de recalage utilisées, nous avons repris 10 des 20 méthodes définies pour le recalage TEMP-IRM. Les diverses adaptations de la stratégie au recalage IRM-IRM sont détaillées dans les parties suivantes.

4.3.1 Choix des différentes approches de recalage

Dans le cadre de cette application de recalage IRM-IRM nous avons testé 10 des 20 méthodes définies pour le recalage TEMP-IRM (car ces méthodes obtenaient des résultats satisfaisants sur un petit échantillon de la base de données lors d'un premier test) dont nous rappelons le paramétrage dans le tableau 4.1.

4.3.1.1 Modèle de transformation

Comme pour le problème du recalage TEMP-IRM, l'application est intra-sujet et concerne une région peu déformable ; nous avons donc opté pour un modèle de transformation rigide (3 paramètres de translation T_x, T_y, T_z et 3 angles de rotation θ_x, θ_y, θ_z).

4.3.1.2 Critères de similarité

La distance de chanfrein (DC), critère géométrique mis en œuvre dans BrainVisa et basé sur la minimisation d'une distance entre contours en utilisant une carte de distance (pour minimiser les temps de calcul), a été testée dans deux configurations différentes conduisant aux méthodes M1 et M2. Parmi les algorithmes implémentés dans SPM, nous avons utilisé ceux basés sur l'information mutuelle (IM) et l'information mutuelle normalisée (IMN), bien que ceux-ci soient plutôt réservés au recalage multimodal, ainsi que ceux basés sur l'entropie du coefficient de corrélation (ECC) et la corrélation croisée normalisée (CCN). Les différents algorithmes utilisant ces critères conduisent aux méthodes M3 à M10.

4.3.1.3 Prétraitements appliqués aux images IRM

Afin de réduire les temps de calcul, les IRM ont été sous échantillonnées d'un facteur 2 dans chaque direction ; autrement dit, la dimension a été divisée par un facteur 8 (IRM/8). Pour les méthodes M1 et M2 basées sur la distance de chanfrein (DC), les contours du crâne ont été extraits en utilisant la méthode décrite dans la section 4.2.1.3.3 pour le recalage TEMP-IRM.

4.3.1.4 Sens de transformation

Les 2 sens de transformation "IRM avant traitement vers IRM après traitement et IRM après traitement vers IRM avant traitement" ont été pris en compte. Le recalage de l'IRM après traitement sur l'IRM avant traitement est considéré comme le sens direct (d) et l'autre comme le sens indirect (i). Dans le cas de la transformation dans le sens indirect, la matrice est inversée pour pouvoir recaler l'IRM après traitement sur l'IRM avant traitement. L'interpolation trilinéaire est utilisée pour effectuer ces transformations.

4.3.1.5 Algorithme d'optimisation et initialisation

Les paramètres de translation (T_x, T_y, T_z) de la matrice de transformation et les angles de rotation (θ_x, θ_y, θ_z) ont été initialisés à une valeur de 0 pour les méthodes M3 à M10 tandis que pour les méthodes M1 et M2 les paramètres de translation (T_x, T_y, T_z) ont été initialisés d'après le centre de gravité (cdg) . L'algorithme d'optimisation de Powell a été systématiquement utilisé pour trouver les extrema des fonctions de coût.

4.3.2 Evaluation quantitative et classement

Afin de classer et sélectionner la meilleure méthode de manière automatique ou semi-automatique, la définition d'un critère quantitatif d'évaluation s'avère indispensable. Dans le cas du recalage d'IRM cérébrales, plusieurs critères d'évaluation comme des indices de recouvrement entre différentes régions du cerveau

segmentées [161] peuvent être utilisés. Pour juger de la qualité des recalages nous avons utilisé également un indice de recouvrement classique, l'indice de Dice moyen, traduisant le recouvrement entre différentes régions. Cet indice est défini par l'expression suivante :

$$Dice(d,m) = \frac{1}{N}\sum_{i=1}^{N}\frac{|A_i(d,m) \cap B_i(d)|}{|A_i(d,m) + B_i(d)|} \tag{4.11}$$

- avec N le nombre de structures, $A_i(d,m)$ la structure i de l'image IRM après traitement du jeu de données d recalée par la méthode m ;
- $B_i(d)$ la structure i de l'image IRM avant traitement du jeu de données d.

La méthode obtenant l'indice de Dice moyen le plus important est alors considérée comme la meilleure méthode pour ce jeu de données.

Afin d'éviter une étape de segmentation supplémentaire pour définir les régions, nous avons choisi comme structures pour cet indice de Dice moyen, les segmentations des globe oculaires, de l'hypophyse et des muscles oculomoteurs (dont l'extraction a été expliquée dans la section 4.2.1.3.3 pour le recalage TEMP-IRM).

4.3.3 Validation visuelle

Afin de s'assurer que le résultat obtenu par la meilleure méthode pour un jeu de données est correct, un expert a procédé à une évaluation visuelle. Dans le cadre de cette application, l'évaluation visuelle se base sur la superposition des contours du crâne, des contours du cerveau, des globes oculaires. Cette évaluation visuelle est faite en fusionnant les deux images IRM et en faisant varier le taux de mélange. On utilise un seuillage consistant à ne garder que les valeurs les plus importantes d'intensité, faisant ressortir le contour du crâne, la tumeur et la matière blanche notamment, puis on applique des tables de couleurs distinctes sur les 2 images pour faciliter l'évaluation visuelle. Pour chaque jeu de données, un score est attribué au résultat du recalage : mauvais, correct ou excellent.

4.4 Segmentation des tumeurs sur l'IRM

La segmentation des zones tumorales constitue une étape essentielle dans le suivi de patient pour extraire des critères comme l'évolution de la taille de la tumeur tel que cela est demandé par les critères RECIST et RANO. Comme nous l'avons vu dans le chapitre 2, il existe un grand nombre de méthodes pour réaliser la segmentation des gliomes de manière semi ou entièrement automatique : des approches déterministes s'appuyant souvent sur des modèles déformables et des modèles statistiques comprenant des approches de classification. Dans le cadre de notre projet, nous nous sommes appuyés sur une approche déterministe.

Pour la segmentation des gliomes nous n'avons pas développé de nouvelles méthodes mais utilisé la méthode de level set 2D dont la théorie est décrit dans le livre de Sethian sur la théorie des approches "Level set" [98] et implémentée dans le logiciel MIPAV. De manière pratique, nous avons demandé à des experts d'utiliser la méthode de level set 2D pour segmenter coupe par coupe les gliomes sur les IRM pondérées en T1 réhaussées. Grâce à l'interface graphique du logicel MIPAV 5.4.4, les experts ont pu initialiser le contour sur la coupe d'intérêt et reprendre le résultat de la segmentation manuellement. Le but était d'obtenir des

segmentations réalisées par des experts sans que ceux-ci aient à segmenter entièrement la tumeur de manière manuelle car cette tâche se révèle chronophage et complexe. Cette méthode nous a donc permis d'obtenir un bon compromis entre le nombre d'interactions utilisateurs et la précision de la segmentation comme le montre la figure 4.19.

4.4.1 Segmentation des tumeurs sur la TEMP

Afin d'analyser conjointement les informations sur la tumeur extraites des modalités IRM et TEMP, la segmentation de la zone tumorale sur l'image TEMP est également nécessaire. Dans le chapitre 2, nous avons vu qu'en raison de la faible résolution et du bruit présent dans ce type d'image, la segmentation se faisait généralement en utilisant une méthode de seuillage. Cependant, avant de procéder à cette étape de seuillage, il peut être nécessaire de normaliser les données TEMP par rapport à une activité de référence définie sur l'image afin de pouvoir comparer des examens à différents temps.

4.4.1.1 Normalisation de la TEMP

Cette normalisation s'est faite en divisant la valeur d'intensité de tous les voxels de l'image TEMP, par la valeur moyenne d'intensité des voxels dans un région d'intérêt sphérique placée dans le lobe frontal mais en en dehors de la zone tumorale. On peut observer cette zone sur l'image 4.20.

4.4.1.2 Segmentation de la TEMP par seuillage

Une fois l'image TEMP normalisée, on peut procéder à la segmentation de la partie tumorale à l'aide d'une méthode basée sur le seuillage des intensités, la tumeur apparaissant avec des intensités élevées sur l'image TEMP. Cependant, afin d'exclure des zones de fixation physiologique qui apparaissent également avec des intensités élevées, l'utilisateur doit définir manuellement une région d'intérêt englobant la tumeur comme le montre la figure 4.21.

Une fois cette région d'intérêt définie, la segmentation par seuillage peut être effectuée à l'intérieur de celle-ci. La partie tumorale sera extraite en gardant tous les voxels de cette région d'intensité supérieure à un certain seuil. Le choix de la valeur du seuil adéquat est une question délicate. En routine clinique, le seuil de 40% de l'intensité maximale dans la région d'intérêt est conventionnellement utilisé. Ce seuil a été fixé d'après les travaux d'Erdi [124]. Cependant, il semble que ce seuil ne soit pas toujours optimal dans le cas de suivi longitudinal. En particulier, nous avons testé l'utilisation d'un même seuil pour les examens d'un patient donné à différents temps. De plus, la région d'intérêt étant définie pour chaque examen, nous avons voulu voir si l'union des régions d'intérêt des examens avant et après traitement avait une influence par rapport à l'utilisation d'une région d'intérêt différente pour les examens d'un même patient à différents temps. Au total nous avons défini cinq méthodes différentes pour la segmentation des gliomes sur les données TEMP :

1. Seg1 : appliquer un seuil égal à 40% de l'intensité maximale des voxels, au sein de régions d'intérêt définies pour chaque examen, et au dessus duquel les voxels sont considérés comme faisant partie inté-

grante de la tumeur.

2. Seg2 : appliquer un seuil égal à 40% de l'intensité maximale des voxels, au sein d'une région d'intérêt qui correspond désormais à l'union des régions d'intérêt des examens avant et après traitement précédemment définies, et au dessus duquel les voxels sont considérés comme faisant partie intégrante de la tumeur.

3. Seg3 : appliquer un seuil égal à 40% de l'intensité maximale des voxels sur l'examen avant traitement au sein d'une région d'intérêt qui correspond à l'union des régions d'intérêt des examens avant et après traitement précédemment définis. Les voxels d'intensité supérieure à ce seuil sont considérés comme faisant partie de la tumeur. Le seuil est défini sur l'examen avant traitement puis réutilisé sur l'examen après traitement.

4. Seg4 : appliquer un seuil égal à 40% de l'intensité maximale des voxels sur l'examen après traitement au sein d'une région d'intérêt qui correspond à l'union des régions d'intérêt des examens avant et après traitement précédemment définis. Les voxels d'intensité supérieure à ce seuil sont considérés comme faisant partie de la tumeur. Le seuil est défini sur l'examen après traitement puis réutilisé sur l'examen avant traitement.

5. Seg5 : appliquer un seuil fixe égal à 6 au sein d'une région d'intérêt qui correspond à l'union des régions d'intérêt des examens avant et après traitement précédemment définis. Les voxels d'intensité supérieure à ce seuil sont considérés comme faisant partie de la tumeur. Ce seuil a été déterminé de manière empirique en observant les données après plusieurs seuillages.

Un exemple de seuillage sur une image TEMP normalisée est visible sur la figure 4.22

4.5 Critères d'analyse

Les zones tumorales ayant été définies lors de la partie segmentation, il est maintenant nécessaire de trouver des indices permettant de caractériser l'évolution tumorale chez chaque patient afin de déterminer si le traitement apporte une amélioration ou non. Ces critères peuvent être globaux ou définis de manière plus locale. Ces derniers nécessitent alors un recalage de qualité entre les examens avant et après traitement et entre les examens des diverses modalités impliquées dans le protocole.

4.5.1 Critères d'analyse globale

Les indices globaux que nous avons utilisés ΔV et ΔI correspondent à des différences relatives de volume ou d'intensité entre examen avant et après traitement.

$$\Delta V = \frac{|V_1 - V_2|}{V_1} \tag{4.12}$$

avec V_1 et V_2 les volumes tumoraux en cm^3 détectés sur les images 1 et 2.

$$\Delta I = \frac{|I_1 - I_2|}{I_1} \tag{4.13}$$

Avec I_1 et I_2 les intensités moyennes au sein des voxels considérés comme faisant partie de la tumeur pour les images 1 et 2.

Remarques : Le critère ΔI a uniquement été calculé sur les images TEMP et non sur les données IRM. Le critère ΔV a également été utilisé pour comparer les volumétries sur les données TEMP et IRM.

4.5.2 Critères d'analyse locale

Comme précisé précédemment, les critères locaux permettent de mieux appréhender l'évolution de la tumeur d'une manière locale et nécessite donc au préalable un recalage de qualité entre les données à analyser. Ces indices peuvent être utilisés par exemple pour déterminer si, malgré la régression apparente d'une tumeur (diminution global du volume tumoral), la tumeur continue de se développer localement.

Dans ce but, deux critères ont été définis. Le premier correspond à un critère décomposé en 5 classes qui permet non seulement de quantifier la superposition entre 2 volumes mais également la distance à laquelle se trouvent les voxels d'un volume V_1 par rapport à un second volume V_2 et vice versa.

Les 5 classes appartenant au premier critère sont définies comme ci-suit :

- **k1 :** nombre de voxels appartenant au volume V_1 et n'appartenant pas au volume V_2 et qui sont distants du volume V_2 de 2 voxels ou plus.
- **k2 :** nombre de voxels appartenant au volume V_1 et n'appartenant pas au volume V_2 et qui sont distants du volume V_2 de 1 (à 2) voxels.
- **k3 :** nombre de voxels appartenant à la fois au volume V_1 et au volume V_2
- **k4 :** nombre de voxels appartenant au volume V_2 et n'appartenant pas au volume V_1 et qui sont distants du volume V_1 de 1 (à 2) voxels.
- **k5 :** nombre de voxels appartenant au volume V_2 et n'appartenant pas au volume V_1 et qui sont distants du volume V_1 de 2 voxels ou plus.

Chacune de ces classes a ensuite été normalisée par le nombre de voxels total que comprend le volume V_1 conduisant aux indices normalisés C1, C2, C3, C4 et C5. La figure 4.23 nous montre un exemple d'utilisation de ce critère.

Dans cet exemple où V_1 correspond au volume tumoral chez un patient détecté sur la TEMP avant traitement et V_2 au volume tumoral chez le même patient détecté sur la TEMP après traitement, la somme des indices C1 et C2 est supérieure à celle des indices C4 et C5. Cela indique une diminution globale du volume tumoral au cours du traitement. En revanche, les indices C4 et C5, même s'ils sont faibles, semblent indiquer que la tumeur continue d'évoluer localement malgré la diminution globale du volume. Cet élement se confirme en observant la visualisation 3D des 2 volumes.

Ce premier critère permet donc de connaître la proportion de voxels d'un volume située à une distance plus ou moins éloignée d'un second volume. Dans le cas d'une forte proportion de voxels appartenant au volume tumoral après traitement mais éloigné du volume tumoral détecté avant traitement, on peut faire l'hypothèse que la tumeur continue d'évoluer. Cependant il serait intéressant de connaître également l'activité au sein de ces voxels éloignés, ce que ne nous permet pas ce premier indice. Ainsi nous avons développé un second critère permettant de prendre en compte l'intensité des voxels dans ces zones. Il est composé de six indices définis en fonction de la distance minimale entre, les voxels appartenant à un volume V_1, et ceux appartenant au volume V_2. Au lieu d'associer le nombre de voxels se trouvant à la distance caractérisant la classe, on lui associe la somme des intensités des voxels se trouvant à cette distance. Ainsi les six classes de ce second critère sont définies comme ci-après :

- **k1 :** somme des intensités des voxels de l'image I_1 appartenant au volume V_1 et n'appartenant pas au volume V_2 et qui sont distants du volume V_2 de 2 voxels ou plus.
- **k2 :** somme des intensités des voxels de l'image I_1 appartenant au volume V_1 et n'appartenant pas au volume V_2 et qui sont distants du volume V_2 de 1 à 2 voxels.
- **k3 :** somme des intensités des voxels de l'image I_1 appartenant au volume V_1 et au volume V_2
- **k3a :** somme des intensités des voxels de l'image I_2 appartenant au volume V_1 et au volume V_2
- **k4 :** somme des intensités des voxels de l'image I_2 appartenant au volume V_2 et n'appartenant pas au volume V_1 et qui sont distants du volume V_1 de 1 à 2 voxels.
- **k5 :** somme des intensités des voxels de l'image I_2 appartenant au volume V_2 et n'appartenant pas au volume V_1 et qui sont distants du volume V_1 de 2 voxels ou plus.

Les 3 premières classes ont été normalisées en divisant k1, k2 et k3 par la valeur de k3 aboutissant aux valeurs L1, L2 et L3. Les 3 dernières classes ont été normalisées en divisant k3a, k4 et k5 par la valeur de k3a aboutissant aux valeurs L3a, L4 et L5. En pratique, si L1 est supérieur à 0,2, on peut considérer que les voxels appartenant à V_1 et distants de plus de 2 voxels de V_2 ont une activité importante. Si L5 est supérieur à 0,2, on peut considérer que les voxels appartenant à V_2 et distants de plus de 2 voxels de V_1 ont une activité importante.

4.6 Conclusion

Ce chapitre a présenté la mise en œuvre de la stratégie d'analyse de l'évolution tumorale dans le contexte de patients atteints de gliomes de haut grade. Cette stratégie comporte l'application de la stratégie de recalage sur les données des modalités TEMP et IRM, la segmentation de la zone tumorale détectée sur les images de chaque modalité et la définition de critères d'évaluation de l'évolution tumorale.
L'étape de recalage comporte 2 applications : le recalage multimodal des données TEMP/IRM et le recalage longitudinal des données IRM.

Pour le recalage multimodal TEMP/IRM, 20 méthodes de recalage rigides ont été utilisées. Afin de sélectionner la meilleure méthode parmi les 20 testées, nous avons proposés un critère d'évaluation nommé UC basé sur les fixations physiologiques du Sestamibi (dans muscles oculomoteurs et l'hypophyse). La mise en œuvre de ce critère nécessite la segmentation sur l'IRM de certaines structures anatomiques. Les globes oculaires, qui correspondent à des zones d'hypofixation du Sestamibi, ont été segmentés en utilisant

une méthode de Hough 3D. Les muscles oculomoteurs, qui correspondent à des zones d'hyperfixation du Sestamibi, ont été segmentés en utilisant un algorithme de ligne de partage des eaux. Enfin l'hypophyse, qui correspond à une zone d'hyperfixation, a été segmentée en définissant une boîte englobante cylindrique.

Pour le recalage longitudinal IRM/IRM, 10 méthodes de recalage rigides ont été utilisées. Afin de sélectionner la meilleure méthode parmis les 10 testées, nous avons utilisé l'indice de recouvrement de Dice moyen basé sur les structures (globes oculaires, muscles oculomoteurs et hypophyse) segmentées sur l'IRM.

Pour l'étape de segmentation des zones tumorales sur l'IRM, une approche de level set a été utilisée tandis que les volumes tumoraux ont été extraits à l'aide de méthodes de seuillage sur les images de la modalité TEMP.

Enfin, deux critères permettant d'évaluer globalement l'évolution tumorale ont été proposés :

- les variations au cours du traitement des volumes des zones tumorales ΔV ;
- les variations d'intensité moyenne ΔI au sein des zones tumorales.

Afin de mener une analyse locale, nous avons proposé deux nouveaux indicateurs :

- un histogramme à 5 classes (C1, C2, C3, C4, C5) traduisant le nombre de voxels d'un volume en fonction de la distance à un autre volume.
- un histogramme à 5 classes (L1, L2 ,L3, L3a, L4, L5) d'intensités en fonction des distances entre volumes.

Ces deux indicateurs se basent sur l'hypothèse suivante : au cours du temps, il peut se produire des évolutions régionales de la zone tumorale active (apparaissant en hyper ou hypo signal suivant les modalités d'imagerie considérées) et qu'une zone détectée tardivement se trouve en dehors de celle détectée sur un examen acquis à un temps antérieur, et ce, même en cas de diminution globale du volume tumoral.

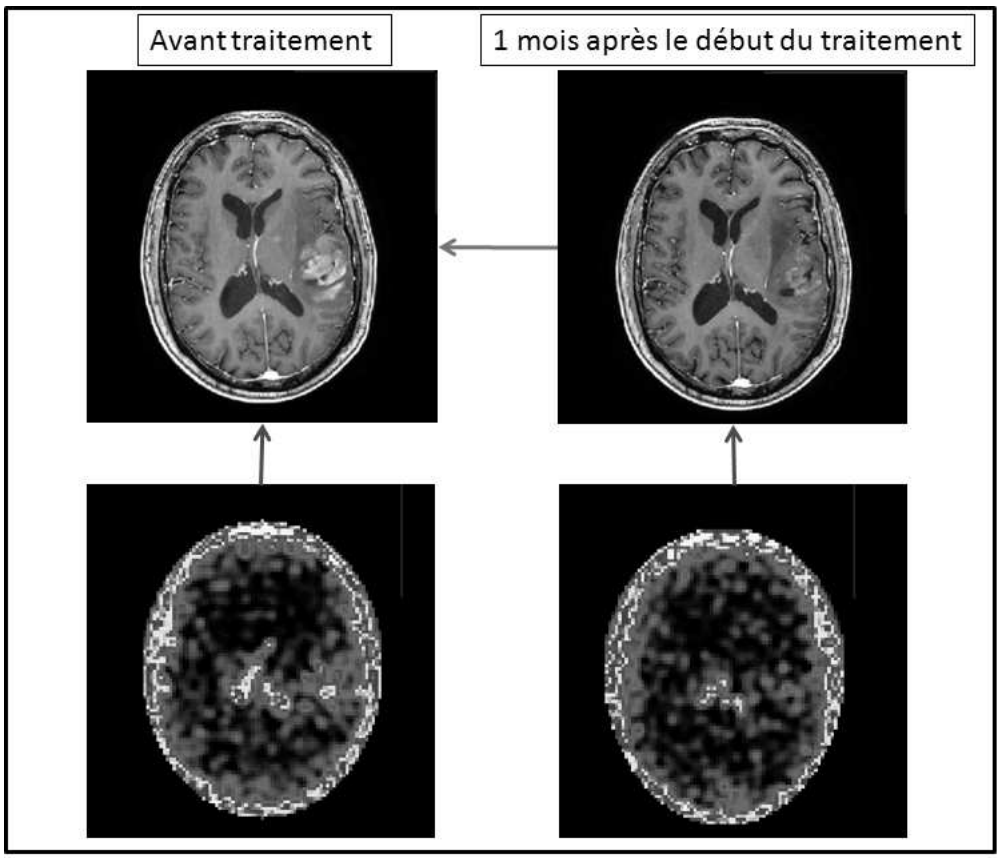


FIGURE 4.17: Recalage longitudinal. La flèche bleue représente le recalage monomodal de l'IRM après le début du traitement sur l'IRM avant traitement et les flèches rouges représentent les recalages multimodaux des TEMP vers les IRM correspondantes. Ainsi pour comparer les examens TEMP avant et après traitement d'un point de vue local, il faut appliquer à la TEMP après traitement (précoce ou tardive) une succession de deux transformations, celle trouvée pour le recalage multimodal sur l'IRM associée puis celle trouvée pour recaler les deux IRM ; les examens TEMP avant traitement doivent également être recalés sur l'IRM avant traitement.

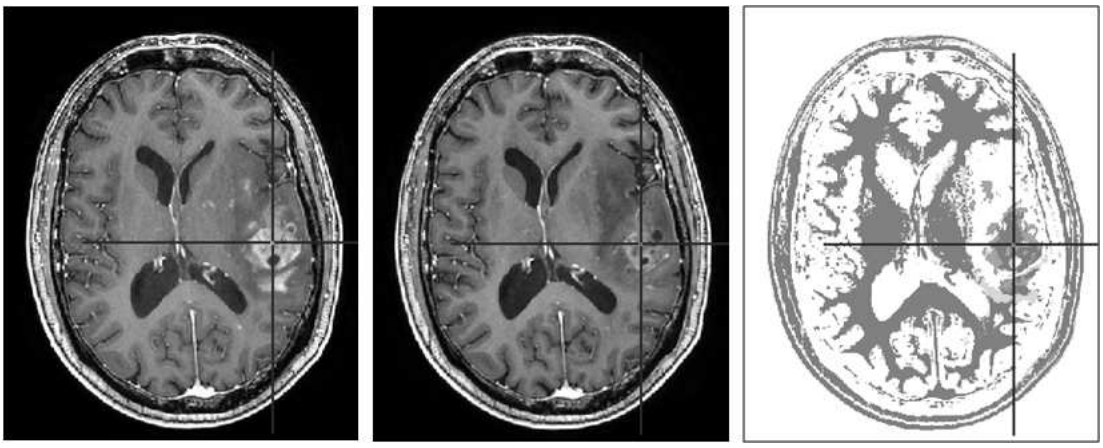


FIGURE 4.18: Exemple d'un recalage considéré comme excellent. A gauche : coupe axiale de l'IRM avant traitement. Au milieu : coupe axiale de l'IRM après traitement recalée sur l'IRM avant traitement. A droite : fusion des 2 IRM avec en gris (où se trouve le curseur) les zones ayant des intensités similaires, en cyan les zones apparaissant plus intenses sur l'IRM recalée après traitement, et en rouge les zones aparaissant plus intenses sur l'IRM avant traitement.

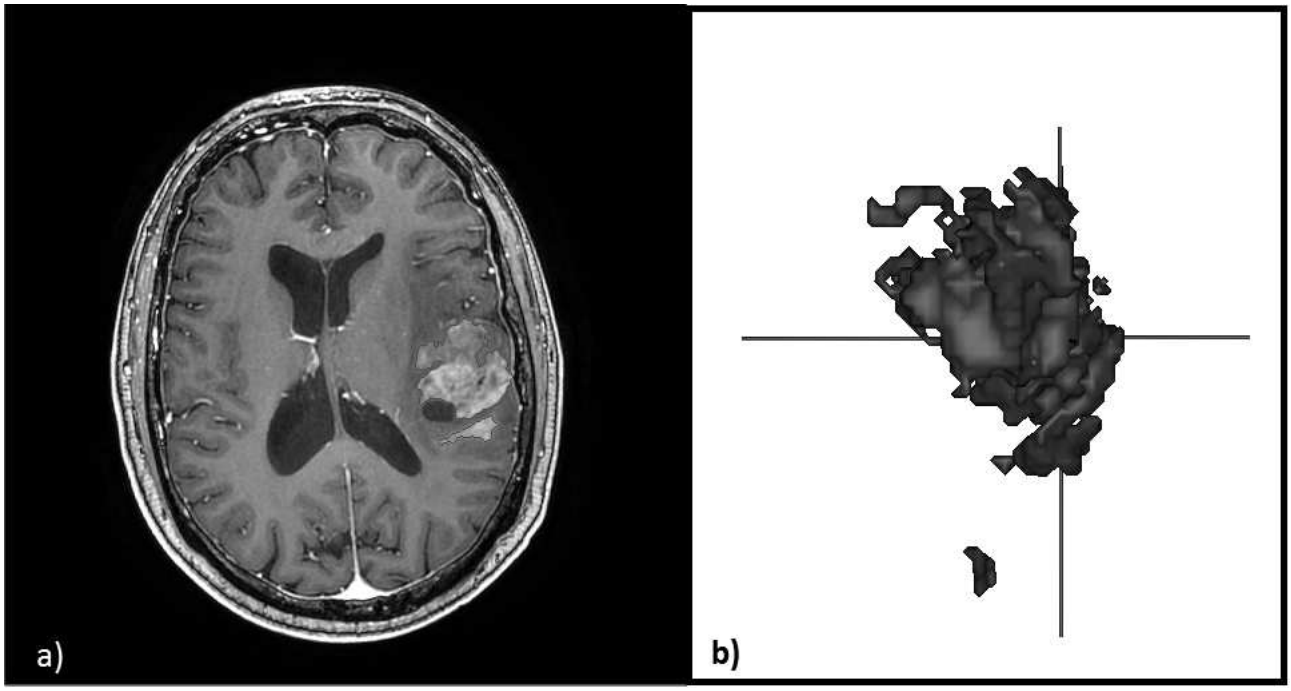


FIGURE 4.19: Segmentation des gliomes sur une IRM pondérée en T1 réhaussée. A gauche : résultat de la segmentation sur une coupe d'IRM à l'aide de la méthode de level set. A droite : visualisation 3D du volume segmenté.

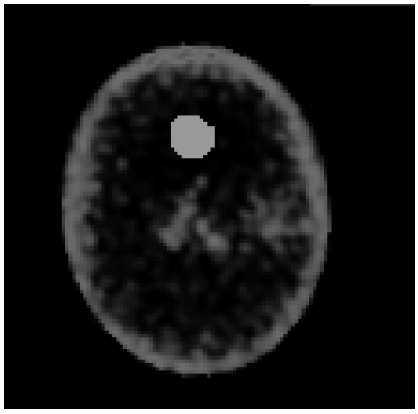


FIGURE 4.20: Normalisation de l'image TEMP par rapport à la valeur mesurée dans le lobe frontal (région d'intérêt sphérique représentée en bleu sur l'image).

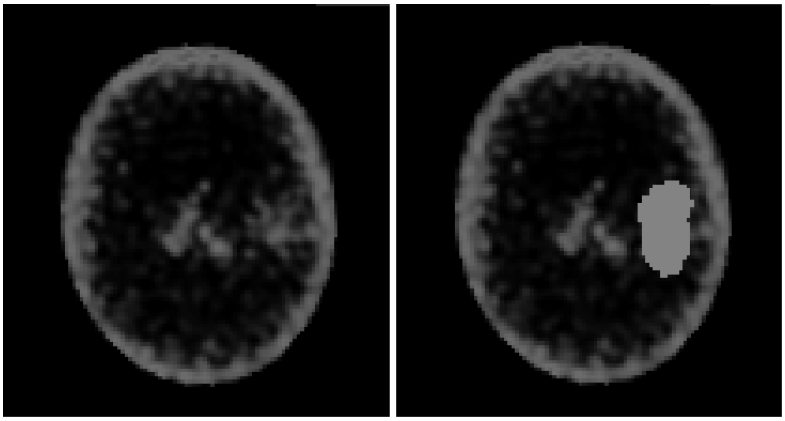


FIGURE 4.21: A gauche : image TEMP normalisée au niveau de la tumeur. A droite : Image TEMP normalisée montrant la région d'intérêt englobant la tumeur en vert.

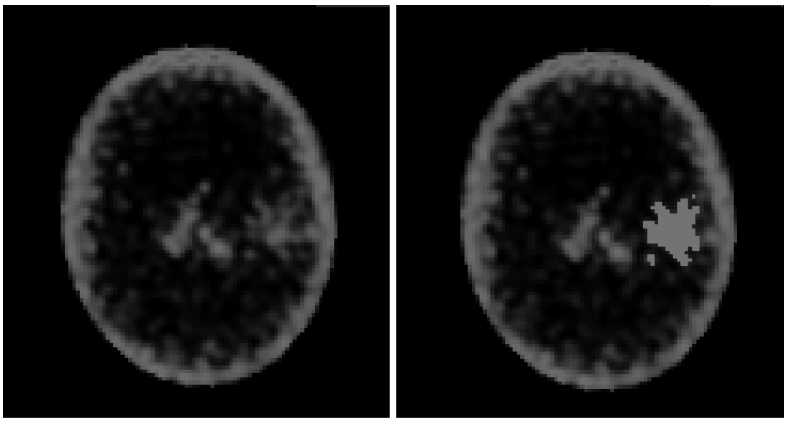


FIGURE 4.22: A gauche : image TEMP normalisée. A droite : Image TEMP normalisée segmentée en utilisant un seuil de 40% de l'intensité maximale, la partie correspondant à la tumeur apparaît en rouge.

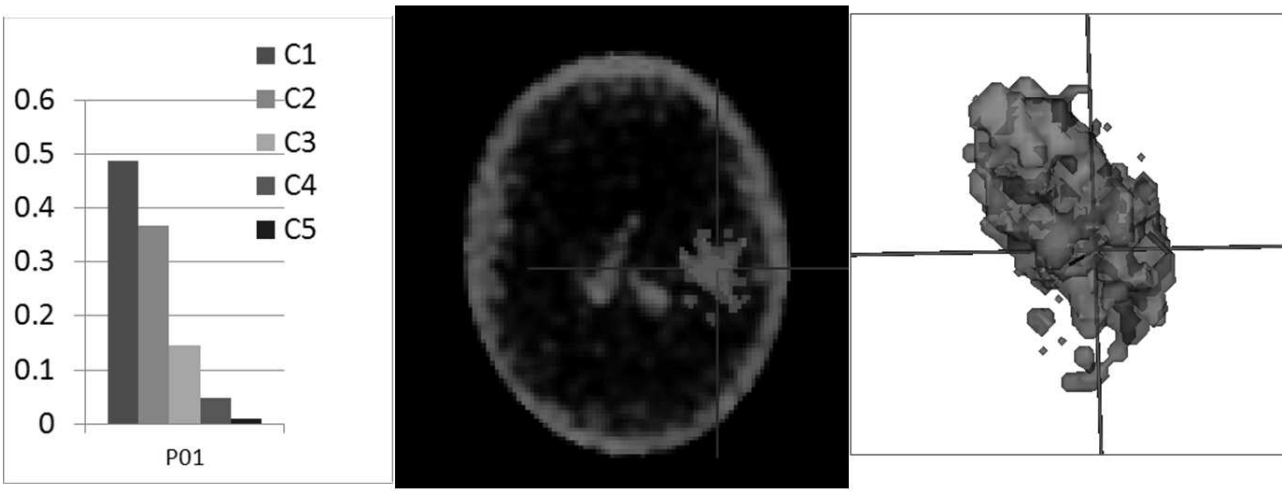


FIGURE 4.23: Exemple de l'utilisation de l'indice à 5 classes en prenant V_1 le volume tumoral détecté sur l'image TEMP avant traitement et V_2 le volume tumoral chez le même patient détecté sur l'image TEMP après traitement. A gauche : les indices montrent une nette regression du volume tumoral au cours du temps. Cependant, les indices C4 et C5 indiquent qu'une petite partie tumorale semble apparaître en dehors de la localisation initiale. Cet élément peut être observé sur l'image du milieu et la visualisation 3D de droite avec le volume V_1 en rouge et le volume V_2 en bleu.

Chapitre 5

Résultats et discussion relatifs à la stratégie de recalage appliquée au protocole IRM-TEMP pour les patients atteints de gliomes de haut grade

Sommaire

5.1 Résultats du recalage TEMP-IRM . . . 110

5.1.1 Validation du critère d'évaluation UC . . . 110

5.1.2 Performance de la stratégie . . . 116

5.1.3 Limites et perspectives . . . 121

5.2 Résultats du recalage longitudinal IRM-IRM . . . 122

5.2.1 Validation du critère d'évaluation Dice moyen . . . 122

5.2.2 Performance de la stratégie . . . 124

5.2.3 Conclusion . . . 125

5.3 Conclusion . . . 125

Dans le chapitre 4, nous avons présenté deux mises en œuvre particulières de la stratégie de recalage dans le contexte du suivi des patients atteints de gliomes de haut grade et traités par antiangiogéniques en utilisant les modalités IRM après injection de Gadolinium et TEMP avec injection de Sestamibi. La première mise en œuvre correspond à une application de recalage multimodal IRM/TEMP tandis que la seconde concerne le recalage longitudinal des IRM avant le début du traitement antiangiogénique sur les IRM acquises un mois après le début du traitement. Pour rappel, notre stratégie de recalage consiste à sélectionner, pour chaque jeu de données, la meilleure méthode de recalage parmi plusieurs à l'aide d'un critère d'évaluation quantitatif afin de soumettre le résultat de recalage par cette méthode à un expert pour validation. Chacune des deux mises en œuvre utilise des méthodes de recalage et des critères d'évaluation différents. Ce chapitre présente

la performance de ces deux mises en œuvre sur la base de données présentée au chapitre 4 après avoir validé la robustesse et la pertinence des critères d'évaluation.

5.1 Résultats du recalage TEMP-IRM

5.1.1 Validation du critère d'évaluation UC

Le critère d'évaluation se trouve au cœur de la stratégie de recalage puisque que c'est l'élément qui permet pour chaque jeu de données de sélectionner la meilleure méthode. Il est donc particulièrement important que ce critère soit en accord avec l'observation des résultats de recalage par des experts, qu'il soit reproductible et ne nécessite pas d'intervention trop complexe de la part de l'opérateur. Dans le cas du recalage TEMP(Sestamibi)/IRM, nous avons proposé le critère UC en tant que critère d'évaluation.

5.1.1.1 Présentation d'indicateurs utilisés pour cette validation

Avant d'exposer les résultats permettant la validation du critère UC, il convient de présenter un certain nombre d'indices permettant cette validation.

Afin de démontrer la pertinence du critère d'évaluation UC, des recalages ont été effectués manuellement par des experts sur 9 jeux de données (sur les 62 que la base de données compte) pour les comparer aux recalages obtenus par les méthodes automatiques utilisées dans la stratégie. Dans le but de faciliter cette comparaison, nous avons introduit le critère UR qui correspond au rapport entre la valeur de UC obtenue par une méthode automatique et celle obtenue par le recalage manuel (méthode $M_{manuelle}$) pour un jeu de données d comme le montre la formule suivante :

$$UR(d, M_m) = \frac{UC(d, M_m)}{UC(d, M_{manuelle})} \tag{5.1}$$

Comme défini au paragraphe 4.2.2.4, le calcul du critère d'évaluation UC s'appuie sur la segmentation de structures anatomiques spécifiques : les globes oculaires supposés être des zones d'hypofixation, l'hypophyse et les muscles oculomoteurs supposés être des zones d'hyperfixation. La segmentation des globes oculaires et de l'hypophyse reste simple et ne nécessite que la définition de régions d'intérêt grossières parallélipédiques contrairement à la segmentation des muscles oculomoteurs pour laquelle les points d'initialisation de l'algorithme de "ligne de partage des eaux" doivent être entrés par un utilisateur quelque peu entrainé. Il est donc légitime de déterminer quelles structures anatomiques sont nécessaires dans le calcul d'UC pour obtenir un classement cohérent. Afin d'étudier l'influence des structures anatomiques prises en compte dans le calcul du critère UC sur le classement général et notamment sur la sélection de la meilleure méthode pour chaque jeu de données, plusieurs variantes de UC ont été définies. UC_{pe} correspond au calcul de UC en ne prenant en compte que l'hypophyse et les globes oculaires, UC_{pm} au calcul de UC en prenant en compte uniquement l'hypophyse et les muscles oculomoteurs, enfin UC_{me} au calcul d'UC en ne prenant en compte que les muscles oculomoteurs et les globes oculaires. Pour chacune des variantes de UC et pour chaque jeu de données d, la meilleure méthode M^*_{test} est identifiée (M^*_{pe} pour UC_{pe}, M^*_{pm} pour UC_{pm} et M^*_{me} pour

UC_{me}) et comparée à celle obtenue pour la configuration de UC prenant en compte toutes les structures anatomiques M^*. Cette comparaison se fait à l'aide de plusieurs indices définis ci-dessous.

- l'indice I1 consiste à compter le nombre de fois que M^*_{test} est différente de M^*. On nommera cet indice M^* inversions. En cas de M^* inversions, deux indices quantitatifs I2 et I3 sont calculés et un indice qualitatif est évalué.
- l'indice I2 correspond à la variation relative de UC, ΔUC trouvée entre la méthode M^*_{test} et M^*. Cet indice est défini par la formule :

$$\Delta UC = \frac{|UC(d, M^*_{test}(d)) - UC(d, M^*(d))|}{|UC(d, M^*(d))|} \tag{5.2}$$

  En dessous d'une valeur de 2% les méthodes M^* et M^*_{test} peuvent être considérées comme équivalentes. Au dessus de cette valeur, une visualisation est nécessaire pour savoir si les deux méthodes sont équivalentes.
- l'indice I3 nommé "Target Registration Evaluation" (TRE) [39] est un critère classiquement utilisé pour définir une distance entre deux matrices de transformation à partir des coordonnées des pixels de l'image source. Si la valeur de TRE entre deux méthodes est inférieure à la taille des voxels dans la direction de plus mauvaise résolution de l'image de moins bonne résolution (2.33 mm pour la TEMP), les deux méthodes sont considérées comme équivalentes.

Enfin l'indice I4 n'est calculé que si $\Delta UC \geq 2\%$ ou si $TRE \geq 2.33mm$. l'indice I4 correspond au score d'une évaluation visuelle pour définir si les méthodes sont équivalentes ou non.

Ces indices sont également utilisés pour étudier la reproductibilité et la robustesse de la méthode de segmentation des muscles oculomoteurs en étudiant l'influence de cette segmentation sur le classement final des méthodes pour chaque jeu de données. De la même manière que pour étudier l'influence des zones anatomiques sur le classement, plusieurs variantes de UC ont été définies en fonction de la manière dont ont été segmentés les muscles oculomoteurs. Deux experts op_{m1} et op_{m2} ont effectué des segmentations des muscles oculomoteurs, en s'aidant d'une méthode de Level Set2D et en reprenant à la main quand nécessaire, conduisant aux 2 variantes de UC, $UC_{op_{m1}}$ et $UC_{op_{m2}}$. Deux opérateurs entrainés mais non experts op_{w1} et op_{w2} ont également procédé à la segmentation des muscles oculomoteurs mais en utilisant la méthode de "ligne de partage des eaux" conduisant aux variantes de UC, $UC_{op_{w1}}$ et $UC_{op_{w2}}$. Pour chaque jeu de données, chaque variante de UC conduit à la sélection d'une meilleure méthode : $M^*_{op_{m1}}$ pour $UC_{op_{m1}}$, $M^*_{op_{m2}}$ pour $UC_{op_{m2}}$, $M^*_{op_{w1}}$ pour $UC_{op_{w1}}$ et $M^*_{op_{w2}}$ pour $UC_{op_{w2}}$.

5.1.1.2 Validation de la cohérence de UC

Comme indiqué dans le paragraphe 5.1.1.1, des recalages "manuels" ont été effectués par des experts en utilisant le logiciel Anatomist sur un sous ensemble de 9 jeux de données dans le but de valider la pertinence du critère d'évaluation UC. Le recalage manuel est considéré comme le "gold standard" et doit conduire à des valeurs de UC plus élevées que celles obtenues par les méthodes automatiques. Afin de faciliter cette comparaison, nous avons défini l'indice UR qui correspond au rapport de la valeur UC obtenue par une méthode automatique sur la valeur obtenue par le recalage manuel. Ainsi, si le critère est cohérent, les méthodes automatiques devraient obtenir des valeurs de UR inférieures à 1. La Figure 5.1 regroupe l'ensemble

des valeurs de UR obtenues par les méthodes automatiques sur les 9 jeux de données.

FIGURE 5.1: Valeur de UR sur 9 jeux de données pour la validation de la pertinence du critère UC. Le critère UC a été calculé à partir des globes oculaires, de l'hypophyse, et des muscles oculomoteurs segmentés à l'aide d'un algorithme de "ligne de partage des eaux" sur l'IRM.

La figure5.1 montre que dans 8 cas, toutes les méthodes automatiques ont obtenu une valeur de UR inférieure ou égale à 1, traduisant le fait que pour chacun de ces cas, le recalage manuel est identifié par le critère d'évaluation UC comme étant la meilleure méthode de recalage. Dans le cas du jeu de données D004, la méthode M20 obtient une valeur de UR supérieure à 1 et donc peut être considérée comme meilleure que le recalage manuel suivant le critère UC. Cependant, la valeur de TRE entre les matrices de transformation obtenues par la méthode M20 et par la méthode de recalage manuel est inférieure à la résolution du voxel ($<$ 2,33 mm) ce qui nous permet d'affirmer que pour ce jeu de données ces deux méthodes sont équivalentes. Cette équivalence a été confirmée par l'expert après visualisation des résultats du recalage. Comme les méthodes automatiques obtiennent pour chacun des jeux de données testés un résultat moins bon ou équivalent aux recalages manuels au sens du critère UC, on peut affirmer que le critère UC est pertinent pour sélectionner la meilleure méthode de recalage pour chaque jeu de données.

La figure5.1 montre aussi que dans 7 cas, il existe une méthode automatique ayant une valeur de UR très proche de 1 traduisant, d'après la pertinence du critère démontrée précédemment, que dans de nombreux cas, la stratégie permet d'obtenir de très bons résultats de recalage. Après expertise visuelle, les recalages obtenus par les meilleurs méthodes au sens de UC sur ces 7 cas ont été jugés d'excellente qualité. On remarque également que la meilleure méthode au sens de UC dépend de chaque cas et qu'une méthode excellente sur un jeu de données peut avoir une performance moyenne sur un autre jeu de données. Par exemple, la méthode M20 obtient des résultats excellents sur D003, D004, D005 et de mauvais résultats sur D006, D007 et D008. La performance d'une méthode varie donc d'un jeu de données à un autre. Ces remarques indiquent l'intérêt

de l'utilisation de la stratégie de recalage par rapport à l'utilisation classique d'une seule méthode pour recaler efficacement l'ensemble des examens présents dans la base de données.

Enfin, au vu des valeurs très proches de UR dans certains cas, il apparaît que plusieurs méthodes peuvent obtenir des performances équivalentes pour un jeu de données considéré. Cependant les méthodes équivalentes sur un jeu de données ne le sont pas forcément sur les autres. Par exemple les méthodes M20 et M7 sont équivalentes sur le jeu de données D005 et non équivalentes sur le jeu de données D007.

Remarque : Dans cette partie, nous avons validé la pertinence du critère UC en utilisant la méthode de "ligne de partage des eaux" pour la segmentation des muscles oculomoteurs. Dans l'article que nous avons publié [162], nous avons montré la pertinence du critère UC mais en utilisant une segmentation manuelle faite coupe à coupe des muscles oculomoteurs. Cette étude ne comportait alors que 9 méthodes de recalage dont 4 ont été finalement retirées de l'étude pour cause de mauvaises performances.

5.1.1.3 Influence des zones utilisées dans le calcul de UC sur le classement

La segmentation de l'hypophyse consiste en la définition d'une région d'intérêt englobant cette zone et celle des globes oculaires ne requiert qu'un point de positionnement dans chaque œil pour définir ensuite une région d'intérêt centrée autour de l'oeil (et éviter de calculer la transformée de Hough sur toute l'image). L'interventon manuelle concernant la segmentation des muscles oculomoteurs est moins facile. Elle consiste au placement de points dans les muscles oculomoteurs pour initialiser l'algorithme de "ligne de partage des eaux". De ces points vont dépendre la qualité de la segmentation finale obtenue. L'interface graphique mise en place pour la segmentation des muscles oculomoteurs a permis de faciliter l'intervention de l'opérateur en permettant de replacer les points au cours du processus en cas de mauvaise segmentation. Bien qu'un expert ne soit pas requis pour cette tâche, il est préférable que cette étape soit réalisée par un opérateur entrainé.

La segmentation de l'hypophyse et des globes oculaires est donc triviale alors que celle des muscles oculomoteurs est un peu plus complexe. Cependant chacune de ces segmentation nécessite l'intervention d'un opérateur. Il est donc légitime de se poser la question de l'utilité de chacune de ces zones dans le calcul d'UC pour garder un critère pertinent. Pour étudier l'influence de chaque région sur la pertinence du critère d'évaluation, il est nécessaire de comparer, pour chaque jeu de données, les classements obtenus et plus particulièrement la méthode sélectionnée et considérée comme la meilleure selon des variantes de UC avec celle sélectionnée selon un critère de référence dont la pertinence a déjà été démontré. La pertinence du critère UC calculé à partir des segmentations des globes oculaires, de l'hypophyse et des muscles oculomoteurs a été démontrée au paragraphe 5.1.1.2 donc UC sera considéré comme référence. Chacune des variantes (UC_{pm}, UC_{me} et UC_{pe}) a été comparée à UC sur l'ensemble des 62 jeux de données en utilisant les indices introduits au paragraphe 5.1.1.1. Le tableau 5.1 présente l'ensemble de ces résultats.

Il faut indiquer que le nombre de cas où les meilleures méthodes identifiées par UC et une variante de UC sont, différentes et considérées comme non équivalentes, et le nombre de fois que leur TRE est supérieur à la résolution du voxel peut être différent. Il se peut que dans certains cas, comme le suggère la validation visuelle, malgré une distance importante entre 2 transformations, ces 2 transformations puissent être assez proches de la transformation inconnue dite "vérité terrain" permettant de recaler parfaitement les 2 images.

	UC_{pm}	UC_{me}	UC_{pe}
$M^{*}inversions$	18	15	43
$\Delta UC \geq 2\%$	3	4	37
$TRE \geq 2.33mm$	9	9	26
non équivalents visuellement	6	2	22

TABLE 5.1: Comparaison des variantes de UC, UC_{pm}, UC_{me} et UC_{pe} pour déterminer quelles structures anatomiques sont réellement nécessaires dans le calcul de UC pour obtenir un classement pertinent. Chaque indice est donné en nombre d'occurences avec un maximum de 62 (nombre de jeux de données dans la base)

Concernant le second indice ΔUC, deux méthodes différentes mais ayant un $\Delta UC \leq 2\%$ (valeur déterminée empiriquement) sont en principe équivalentes mais cette condition n'est pas suffisante pour conclure sur l'équivalence des méthodes. D'autre part, une valeur de $\Delta UC \geq 2\%$ n'exclue pas le fait que deux méthodes puissent être équivalentes, notamment pour de faibles valeurs de UC où la différence relative ΔUC sera plus importante.

Ainsi, dans les cas où aucune des conditions, $\Delta UC \leq 2\%$ ou TRE inférieur à la résolution du voxel, n'est remplie les méthodes sont considérées comme non équivalentes. Dans les cas où au moins une des conditions précédentes est remplie, une validation visuelle est nécessaire pour conclure sur l'équivalence ou la non équivalence. Enfin, lorsque deux transformations ont une distance "TRE" inférieure à la résolution du voxel et que $\Delta UC \leq 2\%$, elles sont considérées comme équivalentes sans besoin de validation visuelle.

En observant le tableau 5.1, on remarque que, malgré un nombre assez important d'inversions de M^{*} pour les variantes UC_{pm} et UC_{me}, dans la majorité des cas, les méthodes ainsi sélectionnées sont équivalentes à celles retenues en utilisant UC. Ainsi, en comparant UC et UC_{pm}, parmi les 18 cas où la meilleure méthode diffère, les transformations trouvées ne sont réellement différentes que dans 6 cas après validation visuelle. La même remarque s'applique pour la variante UC_{me}. En comparant UC et UC_{me}, parmi les 15 cas où la meilleure méthode diffère, les transformations trouvées ne sont réellement différentes que dans 2 cas après validation visuelle. Pour un nombre important de cas, l'hypophyse et les globes oculaires ne paraissent pas si indispensables au calcul de UC pour obtenir un classement cohérent. Cependant, étant donné la facilité de segmentation de ces zones, et pour limiter au maximum le nombre d'erreurs de classement, nous avons préféré garder ces zones dans le calcul de UC.

L'utilité de la segmentation des muscles oculomoteurs est par contre indéniable. En comparant UC_{pe} et UC, on constate 43 inversions concernant le choix de la meilleure méthode dont 22 sont considérées comme non équivalentes après validation visuelle. On peut donc conclure que si l'utilité de l'hypophyse et des globes oculaires dans le calcul de UC peut être discutable, la segmentation des muscles oculomoteurs est en revanche indispensable pour sélectionner la meilleure méthode pour chaque jeu de données.

5.1.1.4 Robustesse du critère vis à vis des segmentations

Après avoir montré que la segmentation des muscles oculomoteurs était indispensable pour obtenir un classement cohérent et sélectionner la meilleure méthode pour chaque jeu de données, nous nous sommes

	op_{m1}/op_{m2}	op_{m1}/op_{w1}	op_{m1}/op_{w2}	op_{w1}/op_{w2}
Dice	$0,64 \pm 0,15$	$0,61 \pm 0,15$	$0,62 \pm 0,15$	$0,89 \pm 0,13$

TABLE 5.2: Indice de Dice (moyenne + écart-type) entre les segmentations des muscles oculomoteurs réalisées par plusieurs opérateurs

	op_{m1}/op_{m2}	op_{m1}/op_{w1}	op_{m1}/op_{w2}	op_{w1}/op_{w2}
M^*	14	11	10	1
$\Delta UC \geq 2\%$	3	2	2	0
$TRE > 2.33$ mm	5	3	3	0
non équivalentes	0	0	0	0

TABLE 5.3: Influence de la segmentation des muscles oculomoteurs par différents opérateurs sur le calcul de UC et sur le classement

intéressés à la reproductibilité de cette segmentation. La segmentation de l'hypophyse et des globes oculaires étant facile, reproductible et peu sujette à erreur, nous nous sommes focalisés sur l'influence de la segmentation des muscles oculomoteurs par différents opérateurs et différentes approches sur le classement final.

La pertinence du classement trouvé en utilisant $UC_{op_{w1}}$ a été démontrée au paragraphe 5.1.1.2 et celle du classement trouvé en utilisant $UC_{op_{m1}}$ avait été démontrée dans un premier travail publié [162].

Pour comparer des segmentations, des critères de recouvrement comme l'indice de Dice sont très souvent utilisés. Les valeurs moyennes et écarts types de l'indice de Dice entre les différentes segmentations des muscles oculomoteurs (sur l'ensemble des 62 jeux de données de la base) sont répertoriés dans le tableau 5.2. Pour le calcul de l'indice de Dice, nous avons considéré les quatre muscles de chaque œil comme une région, et avons fait la moyenne des scores de Dice de ces deux régions.

En observant les scores de Dice sur le tableau 5.2, le recouvrement entre les segmentations manuelles des mucles oculomoteurs des opérateurs op_{m1} et op_{m2}, sont moyens. Le résultat est similaire lorsqu'on compare les segmentations manuelles de l'opérateur op_{m1} avec celles obtenues par les opérateurs op_{w1} et op_{w2} utilisant un algorithme de "ligne de partage des eaux". En revanche, en comparant les segmentations faites par les deux opérateurs op_{w1} et op_{w2} avec l'algorithme de "ligne de partage des eaux", le score de Dice est nettement plus élevé. La segmentation des muscles oculomoteurs est donc plus reproductible en utilisant l'algorithme de "ligne de partage des eaux" qu'en effectuant la segmentation manuellement.

Cependant dans le contexte de la stratégie de recalage, ce ne sont pas les différences entre les segmentations des muscles oculomoteurs obtenues par les différents opérateurs qui nous intéressent mais leur impact sur le classement obtenu et la sélection de la meilleure méthode pour chaque jeu de données. Pour étudier cette influence, nous avons utilisé les mêmes indices que ceux retenus que pour connaître l'influence de chaque structure anatomique sur la sélection de la meilleure méthode. Ces résultats apparaissent dans le tableau 5.3.

Lorsque l'on s'intéresse à l'influence de la segmentation des muscles oculomoteurs par différents opérateurs sur le classement, on peut constater que malgré un certain nombre d'inversions dans la sélection de la meilleure méthode, les méthodes ainsi sélectionnées sont équivalentes dans tous les cas.

On peut donc conclure, que malgré un score de Dice moyen entre la segmentation manuelle de l'opérateur op_{m1} et celles des opérateurs op_{w1} et op_{w2} suggérant une qualité de segmentation intermédiaire, toutes les segmentations permettent d'obtenir un classement pertinent et reproductible. En pratique, l'algorithme de "ligne de partage des eaux" est plus rapide à mettre en œuvre que la segmentation coupe à coupe des muscles oculomoteurs, c'est la raison pour laquelle ce choix apparaît le meilleur.

En conclusion, le critère UC est un critère d'évaluation pertinent pour sélectionner la meilleure méthode de recalage des données (Sestamibi)TEMP et IRM pour chaque jeu de données dans le contexte de suivi de patients atteints de gliomes de haut grade sous traitement antiangiogénique. Ce critère s'appuie sur trois sructures anatomiques segmentées sur l'IRM : les globes oculaires, l'hypophyse et les muscles oculomoteurs. La segmentation de ces structures nécessite une intervention manuelle assez simple pour les deux premières structures. La segmentation des muscles oculomoteurs se révélant plus complexe, son utilité dans le calcul d'UC pour obtenir un classement cohérent a été dans un premier temps démontrée. Dans un second temps, nous avons montré que l'utilisation d'un algorithme de "ligne de partage des eaux" permettait de faciliter les interventions d'un opérateur entrainé pour la segmentation des muscles oculomoteurs tout en obtenant un résultat reproductible et en préservant un classement cohérent. De plus, par rapport à la segmentation manuelle, l'utilisation de l'algorithme de "ligne de partage des eaux" permet de réduire le temps de segmentation d'un facteur 4 au minimum (soit un temps total de segmentation d'environ cinq minutes par IRM).

5.1.2 Performance de la stratégie

La pertinence et la reproductibilité du critère d'évaluation ayant été démontrées (au §5.1.1.4), on peut analyser les performances de la stratégie pour le recalage de données (Sestamibi)TEMP sur les données IRM. Dans cette partie, pour chaque jeu de données d, la meilleure méthode $M^*(d)$ a été sélectionnée suite au classement des méthodes de recalage en utilisant le critère d'évaluation UC dérivant de la segmentation de l'opérateur op_w1. Lors de l'étude préliminaire de la pertinence du critère UC sur 9 jeux de données, nous avons pu observer sur la figure 5.1 que les valeurs d'UR associées à une méthode de recalage pouvaient varier énormément d'un jeu de données à l'autre ce qui indique que la performance d'une méthode varie en fonction des jeux de données. Par exemple, la figure 5.2 illustre un exemple de recalage de bonne qualité obtenu par la meilleure méthode M^* (avec $M^* = M5$ pour ce jeu de données) et un exemple de mauvais recalage obtenu par la méthode M1 classée 19ème pour ce même jeu de données . A l'inverse, la figure 5.3 illustre, sur un autre jeu de données, un recalage de bonne qualité obtenu par la meilleure méthode M^*, avec ette fois $M^* = M1$, et un recalage de mauvaise qualité obtenu par M5 classée 12ème sur ce même jeu de données.

Ces deux figures montrent effectivement que les méthodes peuvent avoir des performances complètement différentes en fonction des jeux de données. La question qui se pose désormais est de savoir dans quelle

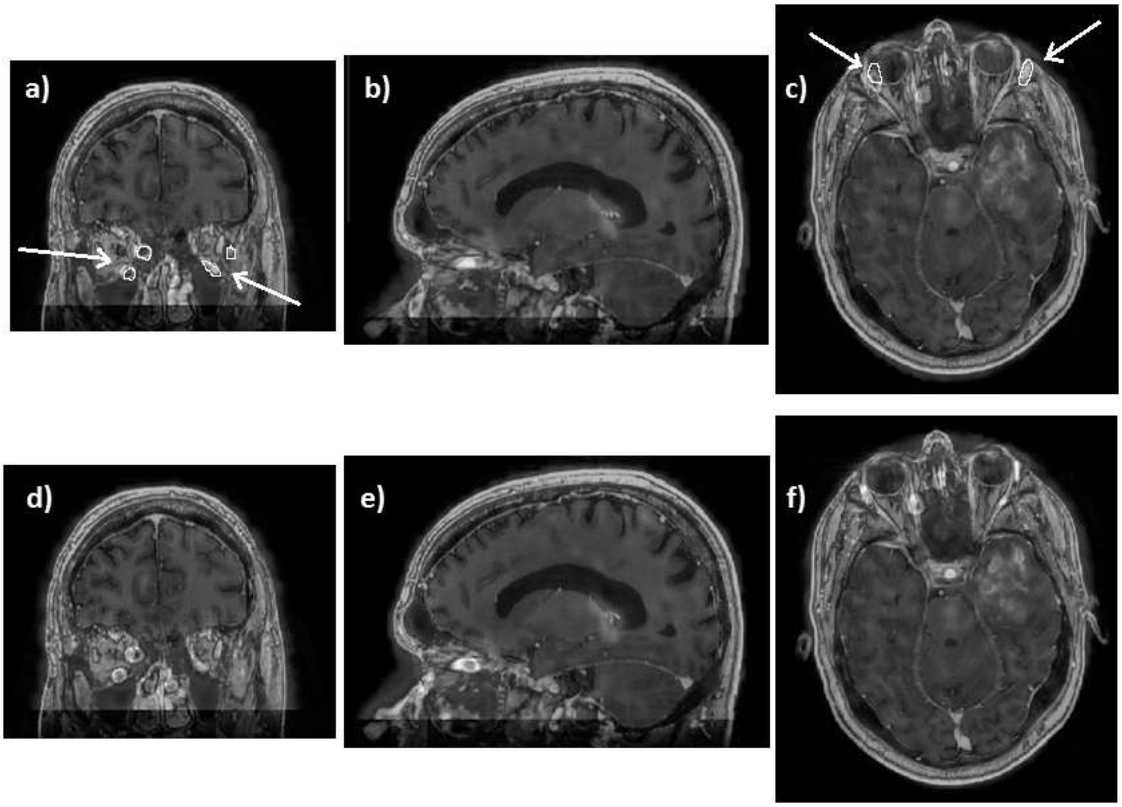


FIGURE 5.2: Exemple d'un recalage de mauvaise qualité obtenu par la méthode M1, classée au 19ème rang (a,b,c) et d'un recalage d'excellente qualité obtenu par la méthode M5 considérée comme la meilleure méthode pour ce jeu de données (d,e,f). Les figures a et d représentent une vue coronale sur laquelle les muscles oculomoteurs sont visibles. Les figures b et e représentent une vue coronale sur laquelle une partie des muscles oculomoteurs est visible. Enfin les figures c et f représentent une vue axiale sur laquelle les globes oculaires, l'hypophyse et les muscles oculomoteurs sont visibles. Sur la figure a, on observe une fixation anormale du radiotraceur légèrement en dehors des muscles oculomoteurs de l'œil gauche et complètement en dehors des muscles oculomoteurs de l'œil droit ; la figure c montre quant à elle une fixation anormale au niveau du globe oculaire gauche et une complètement en dehors du muscle oculomoteur droit. Ces zones de fixation anormales du traceur entourées en blanc et indiquées par des flèches traduisent un recalage de mauvaise qualité. En revanche sur les figures d et f la fixation est importante au niveau des muscles oculomoteurs et faible au niveau des globes oculaires ce qui correspond à un recalage de bonne qualité. Le classement obtenu pour chacune de ces 2 méthodes est en accord avec la validation visuelle.

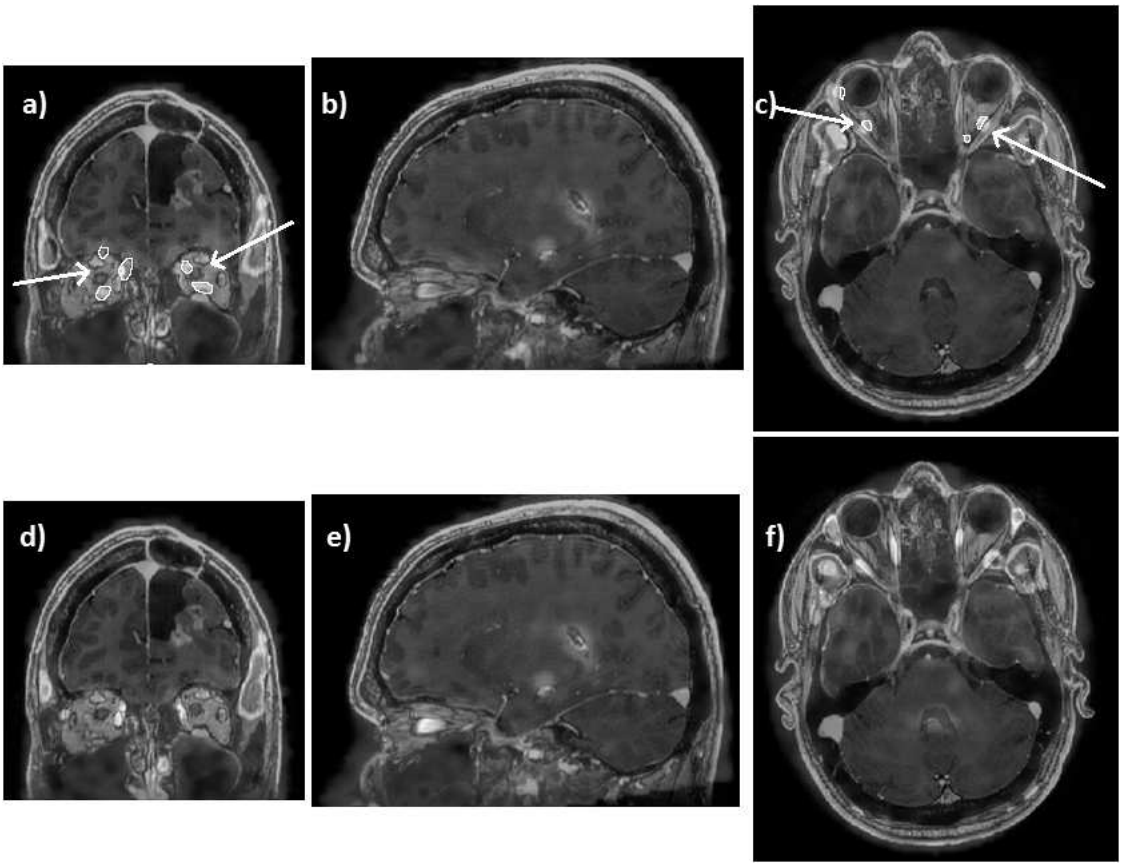


FIGURE 5.3: Exemple d'un recalage de mauvaise qualité obtenu par la méthode M5, classée au 12ème rang (a,b,c) et d'un recalage d'excellente qualité obtenu par la méthode M1 considérée comme la meilleure méthode pour ce jeu de données (d,e,f). Les figures a et d représentent une vue coronale sur laquelle les muscles oculomoteurs sont visibles. Les figures b et e représentent une vue coronale sur laquelle une partie des muscles oculomoteurs est visible. Enfin les figures c et f représentent une vue axiale sur laquelle les globes oculaires et les muscles oculomoteurs sont visibles. Sur la figure a, on observe une fixation anormale du radiotraceur complètement en dehors des muscles oculomoteurs de l'œil gauche et légèrement en dehors des muscles oculomoteurs de l'œil droit ; la figure c montre quant à elle une fixation anormale au niveau du globe oculaire gauche et une fixation complètement en dehors du muscle oculomoteur droit. Ces zones de fixation anormales du traceur entourées en blanc et indiquées par des flèches traduisent un recalage de mauvaise qualité. En revanche sur les figures d et f la fixation est importante au niveau des muscles oculomoteurs et faible au niveau des globes oculaires ce qui correspond à un recalage de bonne qualité. Le classement obtenu pour chacune des 2 méthodes est en accord avec la validation visuelle.

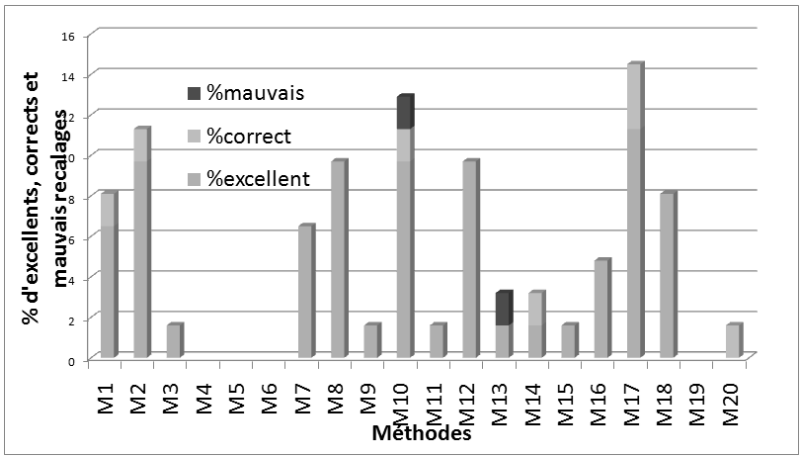


FIGURE 5.4: Fréquence à laquelle chacune des 20 méthodes est classée en premier et qualité du recalage associé

proportion une méthode arrive en premier au classement et de savoir dans un tel cas, si le résultat de recalage est de bonne qualité.

La figure 5.4 montre le pourcentage de cas où chacune des méthodes testées arrive en premier puis la qualité des recalages obtenus avec cette méthode.

D'après la figure 5.4, aucune méthode n'est classée en premier dans plus de 15% des cas. De plus, au moins huit méthodes arrivent en premier dans plus de 6% des cas. Après visualisation par les experts, la qualité globale du recalage obtenue par la stratégie (c'est à dire l'évaluation des méthodes lorsque celles-ci sont classées en première position), montre 85% de recalages d'excellente qualité, 12% de recalages de qualité correcte et seulement 3% de recalages de mauvaise qualité. Ces résultats soulignent l'intérêt de la stratégie de recalage. Cette évaluation ne prend pas en compte l'équivalence entre différentes méthodes de recalage. Ainsi même si une méthode n'est pas classée en premier, elle peut tout de même obtenir un recalage de qualité comparable à la méthode sélectionnée. Notre approche ne permet pas dévaluer chaque méthode individuellement.C'est pourquoi, pour démontrer l'intérêt de la stratégie, nous avons ensuite comparé les résultats de recalage obtenus avec la stratégie et ceux obtenus par les "meilleures" méthodes lorsqu'elles sont utilisées individuellement. Pour identifier les 4 "meilleures" méthodes de la base nous avons repris l'approche proposé par Murphy [77] , en calculant la somme des rangs obtenue par les différentes méthodes sur les 62 jeux de données et en sélectionnant celles ayant obtenu les sommes les plus faibles.

Le tableau 5.4 montre que les quatre meilleures méthodes individuelles identifiées par la somme des rangs sont M14, M16, M17 et M18. La qualité des recalages obtenus sur l'ensemble de la base de données par chacune de ces méthodes a été comparée aux résultats obtenus par la stratégie de recalage comme le montre la figure 5.5.

Ainsi, la méthode individuelle qui obtient le moins de recalages de mauvaise qualité sur la base de données est M18 avec 13% de recalage de mauvaise qualité et la méthode obtenant le plus de recalages d'excellente qualité est M17 avec 61% de recalages d'excellente qualité. Comparée à ces méthodes individuelles, la stratégie de recalage permet d'améliorer grandement la qualité globale du recalage sur la base de données avec 85%

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
somme des rangs	616	728	959	922	975	965	647	650	643	545
rang final	10	14	18	17	20	19	12	13	11	9

	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
somme des rangs	488	446	467	**428**	481	**402**	**431**	**429**	903	893
rang final	8	5	6	2	7	1	4	3	16	15

TABLE 5.4: Somme des rangs obtenue sur les 62 jeux de données par chacune des méthodes

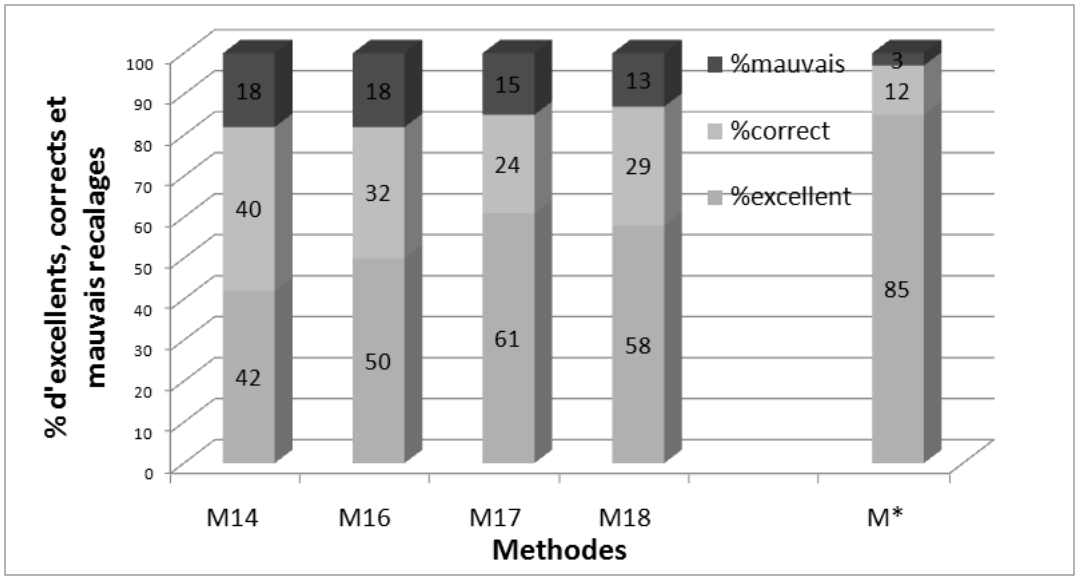


FIGURE 5.5: Comparaison entre la qualité des recalages obtenus par les 4 meilleures méthodes individuelles de la base et celle des recalages obtenus par la stratégie.

de recalages d'excellente qualité et seulement 3% de recalages de mauvaise qualité. Ces derniers résultats confirment l'intérêt de la mise en œuvre de la stratégie dans le contexte du recalage des données TEMP sur les données IRM.

Dans le cas des deux jeux de données mal recalés (représentant 3% des 62 jeux de données de la base), l'extraction des structures non cérébrales pour garder uniquement le cerveau sur l'IRM a été testée pour déterminer si cela pouvait avoir un impact positif sur le recalage. Cette proposition avait été suggérée par l'un des relecteurs de notre article [148]. Les 20 méthodes ont donc été testées sur ces deux jeux de données en retirant les structures non cérébrales. La qualité du recalage obtenu par la meilleure méthode a été jugée moins bonne par les experts que le résultat de la meilleure méthode sélectionnée en utilisant toutes les structures. Cette remarque reste valable pour les deux méthodes basées sur le critère de Wood, pour lesquelles il est conseillé de ne garder que les voxels situés dans le cerveau. Ce résultat peut s'expliquer par le fait que le Sestamibi se fixe essentiellement sur des zones extra cérébrales comme le scalp, les muscles oculomoteurs, le nez. La fixation au niveau du cerveau (en dehors de l'hypophyse et des plexus choroïdes) reste faible et est souvent liée à la pathologie tumorale. De plus, comme les fixations du Sestamibi et le réhaussement du Gadolinium dans la tumeur ne se superposent pas parfaitement, comme nous le montrerons dans le chapitre suivant, cette discordance peut également poser problème pour le recalage limité aux seules structures cérébrales.

5.1.3 Limites et perspectives

La stratégie proposée n'a pas permis de recaler correctement deux jeux de données. Dans de tels cas, le recalage manuel est envisageable mais cette tâche reste complexe et chronophage ($>$ 25 minutes). Pour améliorer les résultats de recalage, des méthodes additionnelles, s'appuyant sur des critères hybrides [62, 60] mêlant informations géométriques et iconiques, dont la pertinence a déjà été démontrée pour d'autres études de recalage TEMP/IRM, pourraient être testées. En effet, l'ajout de nouvelles méthodes dans la stratégie peut permettre d'améliorer la qualité des recalages obtenus, comme nous avons pu le montrer en comparant les résultats exposées dans cette partie et publiés dans [148] avec les résultats obtenus lors d'un travail préliminaire où seulement neuf méthodes étaient utilisées [162]. Cependant ces ajouts augmentent forcément le temps total de calcul. Dans l'implémentation actuelle de la stratégie, un jeu de données IRM/TEMP est recalé en quelques secondes par les méthodes basées sur l'information géométrique (Distance de chanfrein : méthodes M5 à M10) alors qu'un recalage avec une méthode iconique prend environ une minute. Donc la stratégie prend environ 15 minutes pour recaler un jeu de données dans une exécution séquentielle. Une implémentation parallèle pourrait permettre bien évidemment de réduire ce temps d'exécution. Une autre piste intéressante consisterait à déterminer le nombre minimum de méthodes complémentaires à combiner pour garder une qualité de recalage optimale tout en réduisant le temps de calcul. Par exemple, sur cette base de données, les méthodes M5, M6 et M19 n'arrivent jamais en premier et même si elles pouvaient avoir un intérêt pour de nouveaux jeux de données, elles devraient pouvoir être retirées de la stratégie. Bien évidemment, trouver la combinaison idéale de méthodes reste un problème complexe. Pour résoudre ce problème il est essentiel de s'intéresser à la notion d'équivalence entre méthodes sur des jeux de données comme nous avons initié la réflexion dans la partie 5.1.1 en introduisant les critères ΔUC et TRE. Cependant

ces seuls critères ne permettent pas pour le moment d'identifier tous les cas d'équivalences entre méthodes et la visualisation par un expert reste nécessaire pour examiner les cas difficiles. Enfin, nous pensons que cette stratégie de recalage pourrait être utilisée dans d'autres contextes en adaptant les méthodes de recalage et les critères d'évaluation en fonction des spécificités de l'application de recalage concernée. Les études d'évaluation de méthodes de recalage [77, 161, 163] constituent d'autres applications intéressantes où la stratégie pourrait être appliquée directement puisque des critères d'évaluation ont déjà été définis dans ces contextes.

5.2 Résultats du recalage longitudinal IRM-IRM

5.2.1 Validation du critère d'évaluation Dice moyen

Cette validation a été faite à partir des segmentations de l'opérateur op_w1 et en utilisant toutes les structures anatomiques (globes oculaires, hypophyse et muscles oculomoteurs).

Dans le cadre de la validation du critère d'évaluation choisi qui est l'indice de Dice moyen, on réutilisera les 4 indices introduits au §5.1.1 en remplaçant UC par l'indice de Dice moyen (Ainsi ΔUC devient $\Delta Dice$). Des recalages manuels ont également été effectués sur un sous ensemble de jeux de données pour démontrer la pertinence du critère de Dice moyen proposé pour le classement des méthodes.

5.2.1.1 Validation de la cohérence du Dice moyen

Comme pour la validation du critère UC au §5.1.1.2, des recalages manuels ont été effectués sur sept jeux de données. Les recalages manuels sont considérés comme étant le "gold standard". Pour faciliter la visalisation des résultats, nous avons utilisé l'indice de Dice moyen normalisé, défini par la formule 5.3 :

$$ID(d, M_m) = \frac{Dice(d, M_m)}{Dice(d, M_{manuelle})} \tag{5.3}$$

La figure 5.6 montre les résultats obtenus sur ces 7 jeux de données.

Sur cette figure, on peut constater que les 8 dernières méthodes (M3 à M10) obtiennent des valeurs d'indice du Dice moyen normalisé très similaires, laissant supposer qu'elles sont équivalentes. De plus, comme elles obtiennent des scores très proches de 1, on peut en déduire que les recalages obtenus par ces 8 méthodes sont équivalents aux recalages manuels. Cette équivalence a été confirmée en montrant que pour ces 7 jeux de données ($TRE \leq 0,67$ mm et $\Delta Dice \leq 2\%$). En revanche les deux premières méthodes (M1 et M2) obtiennent des scores d'indice de Dice moyen normalisé plus faibles que les 8 autres méthodes. Après visualisation, ces deux méthodes M1 et M2 fournissent des recalages de moins bonne qualité que les 8 autres méthodes. Le critère proposé apparaît donc pertinent pour classer les méthodes.

5.2.1.2 Robustesse du critère vis à vis des segmentations

L'étude de la robustesse de l'indice de Dice moyen vis à vis des segmentations a été effectuée sur toute la base de données. Comme pour étudier l'influence des segmentations des muscles oculomoteurs dans le calcul d'UC, cette partie étudie l'influence de ces segmentations sur l'indice de Dice moyen et sur le classement en

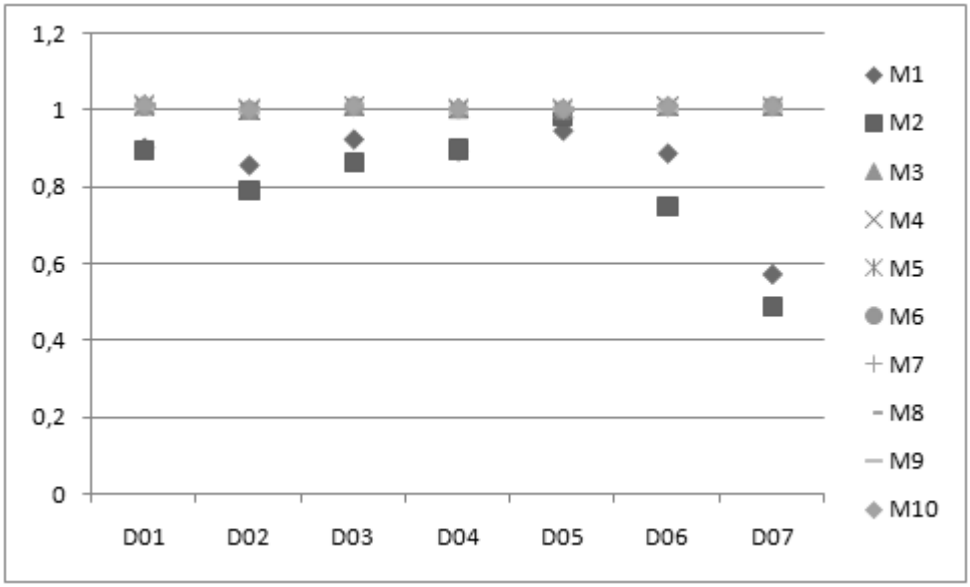


FIGURE 5.6: Indice de Dice moyen normalisé sur 7 jeux de données pour comparer les 10 méthodes de recalage

	op_{m1}/op_{m2}	op_{m1}/op_{w1}	op_{m1}/op_{w2}	op_{w1}/op_{w2}
M^*	9	7	5	4
$\Delta UC \geq 2\%$	0	0	0	0
$TRE > 2.33mm$	0	0	0	0
non équivalentes	0	0	0	0

TABLE 5.5: Influence de la segmentation des muscles oculomoteurs par différents opérateurs sur le calcul de l'indice de Dice moyen et sur le classement

découlant. A ce titre, les segmentations faites par les opérateurs op_m1, op_m2, op_w1 et op_w2 ont été reprises. Dans cette partie, la segmentation de chaque structure a été utilisée mais, contrairement à l'étude menée pour UC, l'influence de chaque zone (des globes oculaires, hypophyse et muscles oculomoteurs) sur le Dice moyen n'a pas été étudiée puisque ces zones sont déjà nécessaires pour le calcul d'UC. Le tableau 5.5 montre les valeurs des 4 indices lorsqu'on étudie l'influence des segmentations des 4 opérateurs sur le classement.

Dans ce tableau, on observe un grand nombre d'inversions quant à la sélection de la meilleure méthode en utilisant la segmentation de chaque opérateur. En effet, utiliser les segmentations de op_m1 ou op_m2 conduit à 9 inversions, celles de op_m1 ou op_w1 à 7 inversions, celles de op_m1 ou op_w2 à 6 inversions et celles de op_w1 ou op_w2 à 4 inversions alors que la base ne comporte que 15 jeux de données. En revanche, malgré ces inversions, TRE est toujours inférieur à la taille du voxel dans sa résolution la moins élevée (0,69 mm) et $\Delta Dice \leq 2\%$ entre les meilleures méthodes sélectionnées à partir des segmentations des 4 différents opérateurs pour tous les jeux de données. Ce résultat est dû à l'existence de méthodes équivalentes pour de nombreux cas comme cela avait déjà été suggéré au §5.2.1.1. Les méthodes sélectionnées à l'aide du critère de Dice moyen calculé

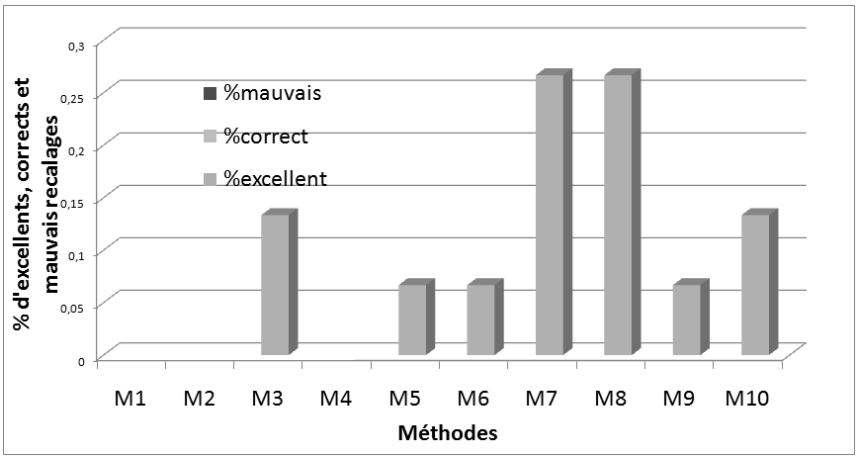


FIGURE 5.7: Fréquence à laquelle chacune des dix méthodes est classée en première position et qualité du recalage associée

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
somme des rangs	141	144	**61**	**63**	**63**	64	**47**	**60**	99	83
rang final	9	10	3	4	4	6	1	2	8	7

TABLE 5.6: Somme des rangs obtenue sur les 15 jeux de données par chacune des dix méthodes de recalage

à partir des segmentations de différents opérateurs sont souvent différentes mais toujours équivalentes, donc le critère permet de sélectionner la meilleure méthode, tout en étant robuste à la segmentation.

5.2.2 Performance de la stratégie

Le critère de Dice moyen ayant été validé, l'intérêt de la stratégie dans ce contexte de recalage longitudinal IRM/IRM peut être étudié. Comme pour le critère UC, nous nous sommes intéressés à la répartition des méthodes M* sur l'ensemble des jeux de données entre les méthodes testées dans la stratégie. La figure 5.7 montre le pourcentage pour lequel une méthode arrive en premier et la qualité des recalages obtenus.

D'après cette figure, tous les recalages obtenus par la stratégie sont excellents. Bien que les méthodes M1, M2 et M4 puissent être retirées sans changer la qualité de recalage globale puisque elles n'arrivent jamais en premier, il semble, néanmoins, que toutes les autres méthodes arrivent au moins une fois en tête. Cependant nous avons vu au §5.2.1.2 et au §5.2.1.1 que certaines méthodes étaient équivalentes sur plusieurs jeux de données. Pour étudier l'intérêt de la stratégie dans ce contexte, il convient donc de comparer les résultats de la stratégie à celui des quatre meilleures méthodes de la base. Pour les sélectionner, la somme des rangs a été calculée et les quatre méthodes obtenant la somme la plus faible ont été considérées. Le tableau 5.6 répertorie la somme des rangs obtenue pour chaque méthode.

D'après ce tableau les cinq meilleures méthodes sont M7, M8, M3, M4 et M5. La figure 5.8 montre le nombre de recalages de bonne qualité obtenus par les 4 méthodes M7, M8, M3 et M4 après validation par

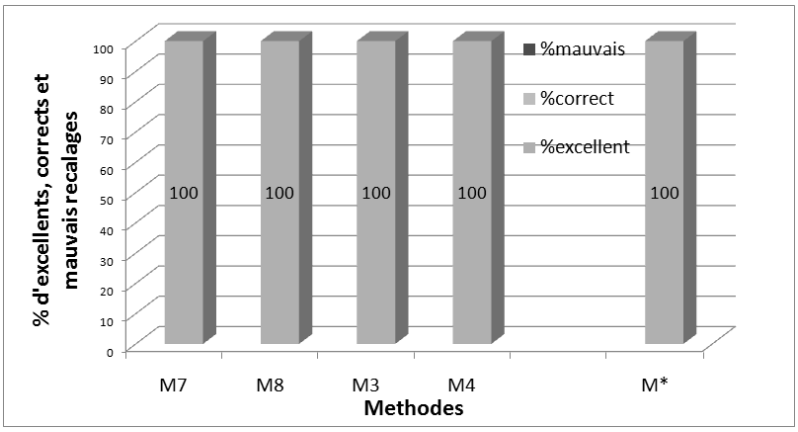


FIGURE 5.8: Comparaison entre les 4 meilleures méthodes et M* définie par la stratégiede recalage.

l'expert, quand on les utilise individuellement et celui obtenu par la stratégie.

D'après cette figure, chacune des quatre méthodes permet d'obtenir 100% de bons recalages sur la base, soit le même score que la stratégie. La stratégie de recalage n'a donc pas d'intérêt pratique par rapport à l'utilisation de l'une de ces quatre méthodes dans ce contexte de recalage longitudinal IRM/IRM. De plus, ces quatre méthodes sont équivalentes sur chaque jeu de données.

5.2.3 Conclusion

Dans le contexte du recalage cérébral intra-patient et monomodal IRM/IRM pour des patients atteints de gliomes de haut grade la stratégie de recalage n'a pas montré d'intérêt par rapport à l'utilisation classique d'une seule méthode comme M7, M8, M3 ou M4.

5.3 Conclusion

Ce chapitre a présenté les résultats de mise en œuvre de la stratégie de recalage appliquée dans le contexte de patients atteints de gliomes de haut grade.
Cette étape de recalage comportait 2 applications : le recalage multimodal des données TEMP/IRM et le recalage longitudinal des données IRM.

Pour la mise en œuvre du recalage multimodal TEMP/IRM, le critère d'évaluation UC nécessitait la segmentation sur l'IRM de structures anatomiques (les globes oculaires, les muscles oculomoteurs et l'hypophyse). La pertinence du critère en UC utilisant toutes les structures anatomiques a été prouvée pour sélectionner la meilleure méthode parmis les 20 testées, en montrant que, sur un sous échantillon de 9 jeux de données, les méthodes automatiques obtenaient des scores de UC inférieurs ou similaires à ceux obtenus par des recalage manuels effectués par un expert. Il a été également démontré sur l'ensemble de la base de données que, pour garder un classement cohérent, la segmentation des muscles oculomoteurs était indispensable dans la mise en œuvre du critère UC. Bien que l'utilité de l'hypophyse et des globes oculaires soit plus

discutable, nous avons décidé de garder ces structures dans la mise en œuvre du critère UC. Enfin l'intéret de la stratégie de recalage pour cette application a clairement été démontré puisqu'elle obtient 85% de recalages d'excellente qualité, 12% de recalages de qualité correcte et seulement 3% de recalages de mauvaise qualité alors que les meilleures méthodes utilisées individuellement n'obtiennent au mieux que 61% de recalages d'excellente qualité et 13% de recalages de mauvaise qualité.

Pour la mise en œuvre du recalage longitudinal des données IRM, la pertinence du critère d'évaluation du Dice moyen a également été démontré pour sélectionner la meilleure méthode parmis les 10 testées. En revanche pour cette application, la stratégie de recalage n'a pas montré d'intérêt par rapport à l'utilisation classique d'une seule méthode. En effet, il a été démontré que les 8 méthodes basées sur des critères iconiques étaient équivalentes pour cette application et que chacune obtenait 100% de recalages d'excellente qualité.

Chapitre 6

Résultats de l'analyse de l'évolution tumorale

Sommaire

6.1 Notations utiles au chapitre résultats 127

6.2 Résultats obtenus avec les critères utilisés en routine clinique 132

6.3 Comparaisons globales et locales des IRM après injection de Gadolinium avant et après le début du traitement 137

6.3.1 Etude globale et comparaison avec les indices PFS et OS 137

6.3.2 Etude locale et comparaison avec les indices PFS et OS 139

6.3.3 Conclusion 141

6.4 Comparaisons globales et locales des TEMP avant et après le début du traitement 142

6.4.1 Analyse des volumes et intensités avant et après traitement 142

6.4.2 Comparaison des différentes méthodes de segmentation 157

6.4.3 Comparaison des données TEMP précoces et tardives 164

6.4.4 Comparaison entre les informations extraites des TEMP avec les indices PFS et OS 175

6.4.5 Discussion et conclusion relatives aux études TEMP 182

6.5 Comparaisons globales et locales entre TEMP et IRM 184

6.5.1 Etude locale 187

6.6 Conclusion 192

6.1 Notations utiles au chapitre résultats

Ce premier paragraphe a pour objectif de définir les notations qui seront utiles pour la présentation des résultats. Pour l'ensemble des indices présentés, nous avons utilisé la norme suivante :

$$X^{M}_{it,ic}$$

Avec

- **X** : nom du critère
- **M** : nom de la modalité de l'image sur laquelle porte le critère. Cet indice peut prendre les valeurs "IRM" ou "TEMP"
- **it** : indice indiquant à quel moment l'image a été acquise après l'injection du traceur radioactif sur les images TEMP. Cet indice n'est utilisé que pour les images TEMP et peut prendre les valeurs 'p' pour "précoce" (acquisition réalisée 15 min après l'injection) ou 't' pour "tardif" (acquisition réalisée 3 h après l'injection).
- **ic** : indice qui indique les types d'images sur lesquels porte ce critère dans le cas d'une comparaison. Cet indice peut prendre les valeurs suivantes :
  - "b-a" pour indiquer que la comparaison porte sur les examens acquis avant ("b" comme before) et après le début du traitement antiangiogénique ("a" comme after) ;
  - "p-t" pour indiquer que la comparaison porte sur les examens TEMP précoces (p) et tardifs (t) ;

  Remarque : "b-a" peut être utilisé pour les deux modalités TEMP et IRM alors que "p-t" est uniquement utilisé pour la modalité TEMP.

Remarque : Les critères PFS ou OS ne sont pas concernés par cette nomenclature.

La liste suivante donne les notations des critères présentés au chapitre 4 triés par ordre d'apparition :

- ΔS^{IRM}_{b-a} : variations relatives de la surface tumorale délimitée sur les IRM après injection de gadolinium au cours du traitement antiangiogénique
- **PFS** : "Progression Free Survival", temps de survie du patient sans aggravation de son état de santé à partir du début du traitement antiangiogénique
- **OS** : "'Overall Survival"', temps de survie globale du patient à partir du début du traitement antiangiogénique
- V^{IRM}_{b} : volume tumoral délimité sur les IRM pondérées en T1 après injection de gadolinium avant traitement antiangiogénique
- V^{IRM}_{a} : volume tumoral délimité sur les IRM pondérées en T1 après injection de gadolinium après traitement antiangiogénique
- ΔV^{IRM}_{b-a} : variations relatives du volume tumoral délimité sur les IRM pondérées en T1 après injection de gadolinium au cours du traitement antiangiogénique
- $Dice^{IRM}_{b-a}$: critère de Dice calculé entre les volumes tumoraux détectés sur les IRM pondérées en T1 avant et après traitement.
- $C1^{IRM}_{b-a}$: la valeur affectée à l'indice C1 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur l'IRM pondérée en T1 avant traitement et situés à 2 ou plus de 2 voxels du volume tumoral après traitement
- $C5^{IRM}_{b-a}$: la valeur affectée à l'indice C5 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur l'IRM pondérée en T1 après

traitement et situés à 2 ou plus de 2 voxels du volume tumoral avant traitement

- V_p^{TEMP} : volumes tumoraux détectés sur les TEMP précoces
- V_t^{TEMP} : volumes tumoraux détectés sur les TEMP tardives
- I_p^{TEMP} : intensité moyenne normalisée au sein des volumes tumoraux détectés sur les TEMP précoces
- I_t^{TEMP} : intensité moyenne normalisée au sein des volumes tumoraux détectés sur les TEMP tardives
- $Imax_p^{TEMP}$: intensité maximale normalisée au sein des volumes tumoraux détectés sur les TEMP précoces
- $Imax_t^{TEMP}$: intensité maximale normalisée au sein des volumes tumoraux détectés sur les TEMP tardives
- ΔV_{pb-a}^{TEMP} : variations relatives du volume tumoral délimité sur les TEMP précoces au cours du traitement antiangiogénique
- ΔV_{tb-a}^{TEMP} : variations relatives du volume tumoral délimité sur les TEMP tardives au cours du traitement antiangiogénique
- ΔI_{pb-a}^{TEMP} : variations relatives de l'intensité moyenne normalisée au sein du volume tumoral délimité sur les TEMP précoces au cours du traitement antiangiogénique
- ΔI_{tb-a}^{TEMP} : variations relatives de l'intensité moyenne normalisée au sein du volume tumoral délimité sur les TEMP tardives au cours du traitement antiangiogénique
- $\Delta Imax_{pb-a}^{TEMP}$: variations relatives de l'intensité maximale normalisée au sein du volume tumoral délimité sur les TEMP précoces au cours du traitement antiangiogénique
- ΔI_{tb-a}^{TEMP} : variations relatives de l'intensité maximale normalisée au sein du volume tumoral délimité sur les TEMP tardives au cours du traitement antiangiogénique
- $Dice_{pb-a}^{TEMP}$: critère de Dice calculé entre les volumes tumoraux détectés sur les TEMP précoces avant et après traitement.
- $C1_{pb-a}^{TEMP}$: la valeur affectée à l'indice C1 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur la TEMP précoce avant traitement et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur la TEMP précoce après traitement
- $C5_{pb-a}^{TEMP}$: la valeur affectée à l'indice C5 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur la TEMP précoce après traitement et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur la TEMP précoce avant traitement
- $Dice_{tb-a}^{TEMP}$: critère de Dice calculé entre les volumes tumoraux détectés sur les TEMP tardives avant et après traitement.
- $C1_{tb-a}^{TEMP}$: la valeur affectée à l'indice C1 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur la TEMP tardive avant traitement et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur la TEMP tardive après traitement
- $C5_{tb-a}^{TEMP}$: la valeur affectée à l'indice C5 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur la TEMP tardive après traitement et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur la TEMP tardive avant

traitement

- $L3a_{pb-a}^{TEMP}$: la somme des intensités des voxels de la TEMP précoce après traitement appartenant à la classe $C3_{pb-a}^{TEMP}$ normalisée par la somme des intensités des voxels de la TEMP tardive avant traitement appartenant à la classe $C3_{pb-a}^{TEMP}$.
- $L5_{pb-a}^{TEMP}$: la somme des intensités des voxels de la TEMP précoce après traitement appartenant à la classe $C5_{pb-a}^{TEMP}$ normalisée par la somme des intensités des voxels de la TEMP tardive avant traitement appartenant à la classe $C3_{pb-a}^{TEMP}$
- $L3a_{tb-a}^{TEMP}$: la somme des intensités des voxels de la TEMP tardive après traitement appartenant à la classe $C3_{tb-a}^{TEMP}$ normalisée par la somme des intensités des voxels de la TEMP tardive avant traitement appartenant à la classe $C3_{tb-a}^{TEMP}$.
- $L5_{tb-a}^{TEMP}$: la somme des intensités des voxels de la TEMP tardive après traitement appartenant à la classe $C5_{tb-a}^{TEMP}$ normalisée par la somme des intensités des voxels de la TEMP tardive avant traitement appartenant à la classe $C3_{tb-a}^{TEMP}$
- $Dice_{p-t}^{TEMP}$: critère de Dice calculé entre les volumes tumoraux détectés sur les TEMP précoces et sur ceux détectés sur les TEMP tardives (avant ou après traitement).
- $C1_{p-t}^{TEMP}$: la valeur affectée à l'indice C1 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur la TEMP précoce et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur la TEMP tardive (avant ou après traitement)
- $C5_{p-t}^{TEMP}$: la valeur affectée à l'indice C5 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur la TEMP tardive et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur la TEMP précoce (avant ou après traitement)
- $L1_{p-t}^{TEMP}$: la somme des intensités des voxels de la TEMP précoce appartenant à la classe $C1_{p-t}^{TEMP}$ normalisée par la somme des intensités des voxels de la TEMP précoce avant traitement appartenant à la classe $C3_{p-t}^{TEMP}$ (avant ou après traitement)
- $L3a_{p-t}^{TEMP}$: la somme des intensités des voxels de la TEMP tardive appartenant à la classe $C3_{p-t}^{TEMP}$ normalisée par la somme des intensités des voxels de la TEMP précoce appartenant à la classe $C3_{p-t}^{TEMP}$ (avant ou après traitement)
- $L5_{p-t}^{TEMP}$: la somme des intensités des voxels de la TEMP tardive appartenant à la classe $C5_{p-t}^{TEMP}$ normalisée par la somme des intensités des voxels de la TEMP précoce appartenant à la classe $C3_{p-t}^{TEMP}$ (avant ou après traitement)
- $Dice_{p}^{IRM-TEMP}$: critère de Dice calculé entre les volumes tumoraux détectés sur les TEMP précoces et sur ceux détectés sur les IRM pondérées en T1 (avant ou après traitement)
- $C1_{p}^{IRM-TEMP}$: la valeur affectée à l'indice C1 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur l'IRM pondérée en T1 et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur la TEMP précoce (avant ou après traitement)
- $C3_{p}^{IRM-TEMP}$: la valeur affectée à l'indice C3 du critère local de comparaison des volumes qui traduit le nombre de voxels présents à la fois dans le volume tumoral détecté sur l'IRM pondérée en T1 et de volume tumoral détecté sur la TEMP précoce (avant ou après traitement)
- $C5_{p}^{IRM-TEMP}$: la valeur affectée à l'indice C5 du critère local de comparaison des volumes qui traduit

le nombre de voxels présents uniquement dans le volume tumoral détecté sur la TEMP précoce et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur l'IRM pondérée en T1 (avant ou après traitement)

- $L1_p^{IRM-TEMP}$: la somme des intensités des voxels de l'IRM pondérée en T1 appartenant à la classe $C1_p^{IRM-TEMP}$ normalisée par la somme des intensités des voxels de l'IRM pondérée en T1 appartenant à la classe $C3_p^{IRM-TEMP}$ (avant ou après traitement)
- $L5_p^{IRM-TEMP}$: la somme des intensités des voxels de la TEMP précoce appartenant à la classe $C5_p^{IRM-TEMP}$ normalisée par la somme des intensités des voxels de la TEMP précoce appartenant à la classe $C3_p^{IRM-TEMP}$ (avant ou après traitement)
- $Dice_t^{IRM-TEMP}$: critère de Dice calculé entre les volumes tumoraux détectés sur les TEMP tardives et sur ceux détectés sur les IRM pondérées en T1. (avant ou après traitement)
- $C1_t^{IRM-TEMP}$: la valeur affectée à l'indice C1 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur l'IRM pondérée en T1 et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur la TEMP tardive (avant ou après traitement)
- $C3_t^{IRM-TEMP}$: la valeur affectée à l'indice C3 du critère local de comparaison des volumes qui traduit le nombre de voxels présents à la fois dans le volume tumoral détecté sur l'IRM pondérée en T1 et le volume tumoral détecté sur la TEMP tardive (avant ou après traitement)
- $C5_t^{IRM-TEMP}$: la valeur affectée à l'indice C5 du critère local de comparaison des volumes qui traduit le nombre de voxels présents uniquement dans le volume tumoral détecté sur la TEMP tardive et situés à 2 ou plus de 2 voxels du volume tumoral détecté sur l'IRM pondérée en T1 (avant ou après traitement)
- $L1_t^{IRM-TEMP}$: la somme des intensités des voxels de l'IRM pondérée en T1 appartenant à la classe $C1_t^{IRM-TEMP}$ normalisée par la somme des intensités des voxels de l'IRM pondérée en T1 appartenant à la classe $C3_t^{IRM-TEMP}$ (avant ou après traitement)
- $L5_t^{IRM-TEMP}$: la somme des intensités des voxels de la TEMP tardive appartenant à la classe $C5_t^{IRM-TEMP}$ normalisée par la somme des intensités des voxels de la TEMP tardive appartenant à la classe $C3_t^{IRM-TEMP}$ (avant ou après traitement)

6.2 Résultats obtenus avec les critères utilisés en routine clinique

Afin de juger de l'efficacité d'un traitement et/ou de chercher "des biomarqueurs" de l'évolution tumorale, on étudie généralement le lien existant entre ces critères et des indices de durée de survie et/ou de qualité de vie. Les durées de survie PFS (Progression Free Survival) et OS (Overall Survival) sont généralement utilisées[164], la qualité de vie étant plus difficilement quantifiable.

L'OS correspond à la durée de survie globale du patient, c'est dire la durée entre le début du traitement juqu'au décès du patient. Cet indice a l'avantage d'être facile à mesurer cependant, de par sa nature, il n'est pas adapté au suivi précoce de patient mais est utilisé dans le cadre de protocoles pour juger l'efficacité de traitements.

Plus récemment proposé, la PFS correspond à la durée de survie sans aggravation de l'état du patient ; il s'agit de la durée durant laquelle la tumeur n'évolue pas selon certains critères d'imagerie et cliniques. Parmi ces critères, on peut citer les critères WHO [24], RECIST [26], les critères de MacDonald [27] ou encore les critères RANO [28]. Ces derniers ont notamment été utilisés pour définir la PFS dans cette étude. L'indice PFS est assez largement utilisé depuis ces dernières années dans le cadre de la validation de nouveaux traceurs et de l'étude de nouveaux traitements [165, 166, 167]. Cependant certains auteurs remettent en question cet indice [168] car s'il n'est pas corrélé à OS, il n'est représentatif ni de la survie réelle du patient ni de la qualité de vie du patient. En effet, le fait que la tumeur semble regresser au cours du traitement d'après l'imagerie et certains critères cliniques n'implique pas nécessairement que le traitement est efficace ; la tumeur peut régresser pendant une période et devenir plus aggressive par la suite, et cela sans que le patient perçoive une amélioration de sa qualité de vie.

Ainsi les indices PFS et OS ont été systématiquement utilisés dans ce chapitre car ils ne sont pas nécessairement corrélés.

Pour rappel, les critères RANO [28] (cf chapitre 1) permettent d'évaluer la réponse tumorale à un traitement en définissant quatre classes :

- PD : progression de la tumeur
- SD : tumeur stable
- PR : réponse partielle au traitement
- CR : réponse complète

Pour la population étudiée, aucun patient n'entre dans la catégorie "CR". Pour évaluer la réponse tumorale, les critères RANO s'appuient sur plusieurs éléments définis au §1.1.4.2 comme l'évolution des volumes après prise de contraste sur les IRM pondérées en T1 (suite à l'injection de Gadolinium). Cette évolution des volumes est alors estimée à partir de la somme du produit des diamètres de chaque zone de prise de Gadolinium correspondant à une lésion tumorale sur l'IRM pondérée en T1.

Il est intéressant de vérifier le lien entre les critères RANO et les indices PFS (Progression Free Survival) et dans un second temps d'étudier le lien entre les critères RANO et l'indice OS (Overall Survival).

Le tableau 6.1 apporte un résumé de certaines caractéristiques cliniques des seize patients étudiés, le traitement utilisé, ainsi que les durées de survie PFS et OS. Les termes utilisés dans ce tableau sont définis ci-dessous.

Les termes relatifs aux traitements administrés aux patients sont les suivants : Bev (Bévacizumab) qui

correspond au traitement antangiogénique par un inhibiteur de l'angiogénèse et Ir (Ionizing radiation) qui correspond au traitement par radiothérapie.

KPS (Karnofsky Performance Status) est un indice qui correspond en quelque sorte à une mesure de "qualité de vie" du patient. Ce dernier indice peut prendre des valeurs de 0 à 100% :

- 100% : Normal, pas de signes de maladie
- 90% : Peut mener une activité normale, symptômes ou signes mineurs de la maladie. Totalement autonome.
- 80% : Peut mener une activité normale mais avec effort, symptômes ou signes mineurs. Totalement autonome.
- 70% : Peut se prendre en charge, incapable de mener une activité normale. Autonome, mais à stimuler.
- 60% : Nécessite une aide occasionnelle, mais peut prendre en charge la plupart de ses besoins. Semi-autonome.
- 50% : Nécessite une aide suivie et des soins médicaux fréquents. Semi-autonome.
- 40% : Handicapé, nécessite une aide et des soins particuliers, dépendant.
- 30% : Sévèrement handicapé, dépendant.
- 20% : Très malade, soutien actif, absence totale d'autonomie.
- 10% : Moribond, processus fatal progressant rapidement.

Dans le tableau 6.1, il manque un certain nombre d'informations pour les patients P13 et P15 car ceux-ci ont été exclus de l'étude en cours de protocole. Le patient P13 a été exclu pour une complication digestive liée au Bévacizumab et motivant son arrêt. Le patient P15 a été exclu parce qu'il a fait une hydrocéphalie motivant également une interruption transitoire du traitement pour la dérivation du LCR, rendant difficile la détermination précise de la PFS. La dernière colonne indique les symboles qui seront utilisés dans les graphiques suivants pour faire référence aux données des patients. Chaque combinaison de couleur et de symbole correspond donc à un patient distinct. Les patients représentés par des symboles rouges sont des patients dont la tumeur progresse (PD selon les critères RANO), ceux représentés par des symboles bleus sont des patients stables (SD), ceux représentés par des symboles verts sont des patients qui répondent au traitement (PR) et ceux avec des symboles noirs n'appartiennent à aucune de ces classes car les données PFS sont manquantes pour ces patients.

Patient/ Sexe/Age	Diagnostic initial	Nombre de récurrences avant traitement	KPS	Stéroïdes	Traitement	RANO	PFS (jours)	OS (jours)	Symbole patients
P01/M/39	AA	1	80	Non	Bev+ Ir	PR	78	177	○
P02/M/72	GBM	2	70	Oui	Bev	PD	41	160	○
P03/F/71	GBM	1	70	Oui	Bev+Ir	SD	34	104	✱
P04/M/29	AA	2	90	Non	Bev	SD	83	477	○
P05/F/33	OA II	2	70	Oui	Bev	PD	8	445	◇
P06/F/66	GBM	1	90	Oui	Bev	SD	112	448	◇
P07/M/51	OAA	1	70	Oui	Bev	SD	94	152	+
P08/F/61	AA	2	60	Oui	Bev+Ir	PR	163	235	◇
P09/F/67	GBM	1	70	Oui	Bev	PD	20	104	☆
P10/M/59	GBM	1	70	Oui	Bev+Ir	SD	75	193	★
P11/M/65	GBM	1	80	Oui	Bev+Ir	PR	103	337	×
P12/M/62	GBM	1	60	Oui	Bev	PD	15	127	×
P13/M/65									○
P14/M/62	GBM	1	80	Oui	Bev+Ir	PD	29	66	□
P15//M/66									+
P16/M/46	GBM	1	80	Non	Bev	PD	61	249	✱

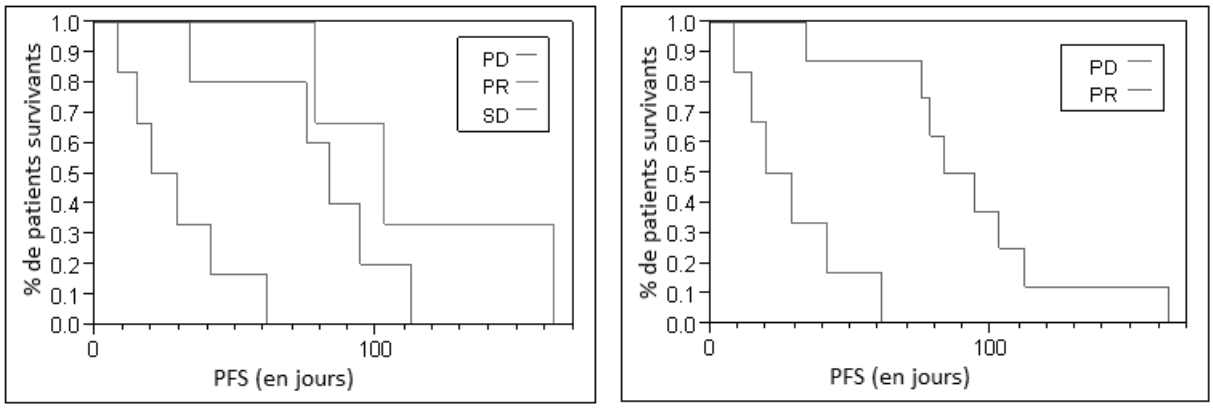
TABLE 6.1: Tableau des critères RANO et des indices PFS et OS. Avec M : Masculin, F : Féminin, AA : Astrocytomes Anaplasiques, OA : oligodendrogliomes Anaplasiques, OAA : Oligoastrocytomes Anaplasiques, GBM : Glioblastomes, KPS : Karnofsky Performance Status, .

Pour étudier le lien entre les critères RANO et les indices PFS et OS, nous avons étudié les courbes de Kaplan-Meier représentées sur la figure 6.1 puis effectué un test de rang logarithmique. Dans un second temps nous avons rassemblé les classes SD et PD conduisant à l'étude de survie sur seulement 2 classes. Enfin nous avons regroupé les classes SD et PR . Les résultats des tests logarithmiques par rangs sont indiqués dans le tableau 6.2.

	PFS	OS
3 classes (PD, SD, PR)	p = 0,0016	p = 0,4167
2 classes (PD et SD, PR)	p = 0,0787	p = 0,9858
2 classes (PD, SD et PR)	p = 0,0005	p = 0,2264

TABLE 6.2: Résultat des tests logarithmiques par rang. Ces tests ont été effectués sur 14 points, les données des critères RANO et les indices PFS et OS manquant pour les 2 patients P13 et P15. La valeur de p correspond à la probabilité d'obtenir une valeur plus extrème du test statistique si l'hypothèse nulle est vraie, l'hypothèse nulle étant dans ce cas qu'il n'y a pas de différence entre les populations comparées en terme de probabilité d'évènements (nombre de patients décédés pour OS et nombre de patients dont la maladie s'est aggravée) à un temps t. Pour interpréter le test statistique, on compare la valeur de p obtenue à un seuil en dessous duquel on peut considèrer que l'hypothèse nulle peut être rejetée. Ce seuil est défini par l'utilisateur et est généralement fixé à une valeur de 5%. Dans notre cas une valeur de p inférieure à 5% signifie que les populations comparées sont différentes et que les critères qui ont permis de définir les population sont prédictifs des durées de survie. Dans le cas contraire, à savoir une valeur de p supérieure à 5%, le résultat n'est pas significatif et on ne peut rien conclure.

Les critères RANO sont prédictifs de la durée de survie sans aggravation de l'état de santé des patients PFS ($p < 0{,}05$) en 3 classes, ou en 2 classes en regroupant SD et PR comme le montrent également les courbes de survie représentées sur les figures 6.1a et 6.1b qui sont éloignées les unes des autres et permettent de distinguer des populations bien distinctes. Par contre, les valeurs prises par les critères RANO ne semblent pas prédictives de la durée de survie globale des patients OS ($p > 0{,}05$).



(a) Analyse de Kaplan Meier traduisant le PFS pour trois classes de patients PD, SD, PR.

(b) Analyse de Kaplan Meier traduisant le PFS pour 2 classes de patients PD, (SD et PR).

FIGURE 6.1: Analyse de Kaplan Meier traduisant PFS pour des classes de patients définies à partir des critères RANO.

6.3 Comparaisons globales et locales des IRM après injection de Gadolinium avant et après le début du traitement

6.3.1 Etude globale et comparaison avec les indices PFS et OS

Nous avons étudié, au cours du traitement antiangiogénique, les variations des volumes correspondant à la partie tumorale "visible" sur les IRM après injection de Gadolinium. Cet élément constitue un des quatre indices sur lequel repose la définition des critères RANO à la différence près que le critère RANO s'appuie sur une variation de surface (cf paragraphe 1.1.4.2). Pour se rapprocher de cet indice, nous avons considéré le volume tumoral comme une sphère pour en déduire un diamètre puis une surface équivalente sur la coupe passant par le centre de la sphère et pouvoir calculer les variations de cette surface ($\pi\lambda^2$) entre les examens avant et après traitement. Cet indice sera nommé ΔS_{ba}^{IRM}. Les volumes V_b^{IRM} et V_a^{IRM} ainsi que les variations relatives de volume ΔV_{b-a}^{IRM} et de surface équivalente ΔS_{ba}^{IRM} sont donnés dans le tableau 6.3.

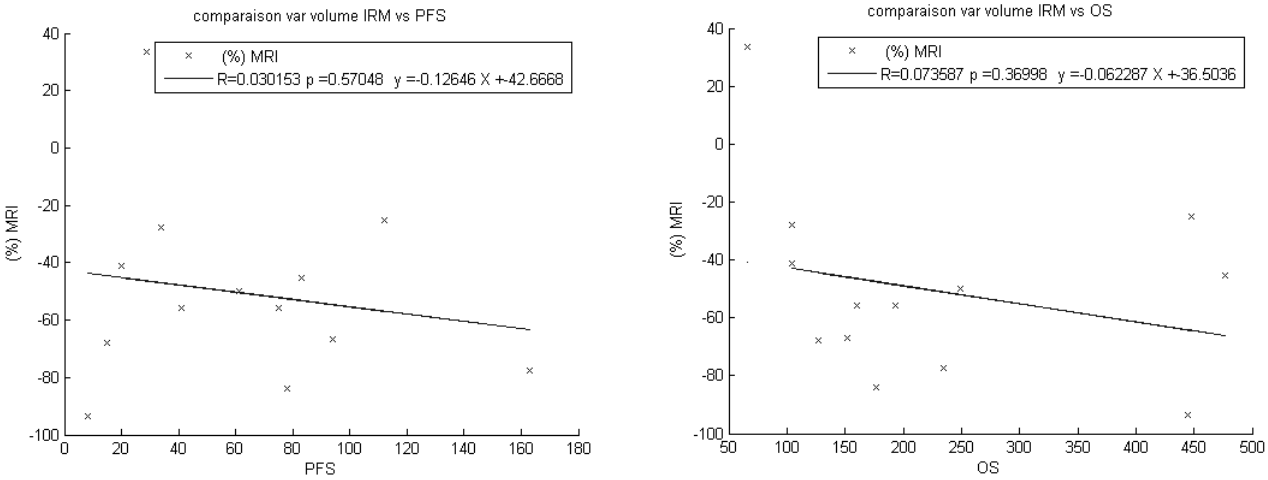
	V_b^{IRM}	V_a^{IRM}	ΔV_{b-a}^{IRM} (en %)	ΔS_{ba}^{IRM}
P01	40,9	6,61	-83,8	-70,3
P02	29,6	13,1	-55,9	-42,1
P03	14,3	10,3	-27,8	-19,5
P04	7,73	4,23	-45,3	-33,1
P05	6,08	0,40	-93,4	-83,7
P06	4,16	3,11	-25,2	-17,6
P07	15,5	5,14	-66,8	-52,0
P08	30,2	6,81	-77,4	-62,9
P09	18,2	10,7	-41,3	-29,8
P10	26,6	11,7	-56,0	-42,1
P11	29,5	/	/	/
P12	19,9	6,36	-68,0	-53,2
P13	3,44	0,09	-97,5	-91,4
P14	33,6	44,8	33,4	21,2
P15	6,04	3,27	-45,8	-33,5
P16	8,79	4,40	-50,0	-37,0

TABLE 6.3: Volumes et variations relatives de volume et de surface équivalente correspondant à la partie tumorale détectée sur les IRM pondérées en T1 après l'injection de Gadolinium au cours du traitement antiangiogénique

Etant donné qu'il manque l'IRM après traitement du patient P11, les variations de volume des zones tumorales sur l'IRM au cours du traitement n'ont pu être calculées que sur 15 points. Le tableau 6.3 montre que sur les IRM, la partie tumorale détectée diminue pour 14 patients et augmente seulement pour le patient

P14. Si l'on considère les seuils indiqués par les critères RANO, le volume tumoral ne croît que pour le patient P14, il est stable pour 8 patients (P02, P03, P04, P06, P09, P10, P15, P16) et il diminue largement pour 6 patients (P01, P05, P07, P08, P12, P13).

Afin de déterminer si ces variations au cours du traitement antiangiogénique ont un lien avec les indices PFS et OS (et si tel est le cas de déterminer une valeur prédictive), nous avons testé s'il existait une corrélation entre ces variations relatives de volumes et les durées PFS (figure 6.2a) et OS (figure 6.2b)



(a) Corrélation entre ΔV_{b-a}^{IRM} et PFS

(b) Corrélation entre ΔV_{b-a}^{IRM} et OS

FIGURE 6.2: Corrélation entre ΔV_{b-a}^{IRM} avec PFS et OS

L'étude de corrélation a porté sur 13 points car il manque l'IRM après traitement pour le patient P11 et les données PFS et OS pour les patients P13 et P15.

La figure 6.2a montre qu'il n'existe pas de corrélation entre ΔV_{b-a}^{IRM} et PFS bien que ce critère intervienne dans la définition des critères RANO. Il n'existe aucune corrélation entre ΔV_{b-a}^{IRM} et la durée de survie globale des patients OS comme le montre la figure 6.2b.

Cela prouve que l'évaluation de l'évolution du volume tumoral observée sur l'IRM après injection de Gadolinium ne suffit ni à prédire la durée de survie globale (OS) des patients, ni à prédire la durée sans aggravation de l'état du patient (PFS) en réponse au traitement antiangiogénique. Cela montre bien les limites des critères de Mac Donald [27]. Ce critère de variations de volume doit être associé à d'autres critères (FLAIR, status clinique, nouvelles lésions ...).

6.3.2 Etude locale et comparaison avec les indices PFS et OS

Nous avons relevé dans la partie précédente une diminution quasi systématique du volume tumoral sous l'effet du traitement antiangiogénique. Nous avons voulu étudier dans un second temps si malgré une décroissance globale du volume de la tumeur, il existait une augmentation locale de la zone tumorale. Pour répondre à cette question, nous avons utilisé les indices locaux proposés au §4.5 comparant les volumes délimités sur l'IRM après l'injection de Gadolinium avant traitement et ceux déterminés après traitement. Les résultats associés aux indices C1 et C5 (critères introduits au §4.5) ainsi que l'indice de Dice entre les deux volumes permettent de comparer localement ces volumes comme le montre le tableau 6.5.

	$Dice_{b-a}^{IRM}$	$C1_{b-a}^{IRM}$	$C5_{b-a}^{IRM}$
Seg	min = 0,05 max = 0,87 moyenne = 0,49 écart-type =0,23	min = 0,05 max = 0,97 moyenne = 0,47 écart-type = 0,27	min = 0 max = 0,54 moyenne = 0,06 écart-type = 0,14

TABLE 6.4: Indices locaux pour comparer les examens avant et après traitement en IRM

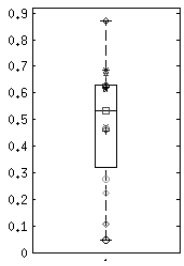
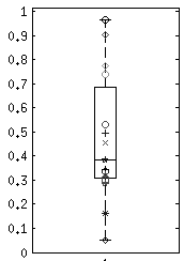
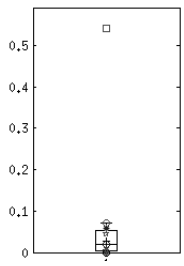
	$Dice_{b-a}^{IRM}$	$C1_{b-a}^{IRM}$	$C5_{b-a}^{IRM}$
Seg			

TABLE 6.5: Indices locaux pour comparer les examens avant et après traitement en IRM. Ces indices $Dice_{b-a}^{IRM}$, $C1_{b-a}^{IRM}$ et $C5_{b-a}^{IRM}$ ont été calculés sur 15 points (l'IRM après traitement du patient P11 étant manquante).

Le critère de Dice est en moyenne peu élevé (moyenne = 0,49) indiquant que les volumes tumoraux détectés sur l'IRM avant et après traitement ne se superposent pas parfaitement. En effet, un indice de Dice faible indique une faible concordance entre les deux volumes. La principale raison est due à la diminution du volume après traitement. En effet, si le volume après traitement est entièrement inclus dans le volume avant traitement mais est par exemple de taille deux fois moindre, cela aboutit à un indice de Dice égal à 2/3. L'avantage des indices C1 à C5 est de pouvoir préciser si, malgré une mauvaise superposition, un volume est entièrement inclus dans l'autre, ou, dans le cas contraire, à quelle distance les voxels qui n'appartiennent qu'à un seul volume se situent du second. La valeur de $C1_{b-a}^{IRM}$ est élevée (>0,20) pour 13 points. Les deux patients P03 et P06 pour lesquels $C1_{b-a}^{IRM}$ est faible sont considérés comme stables selon les critères RANO.

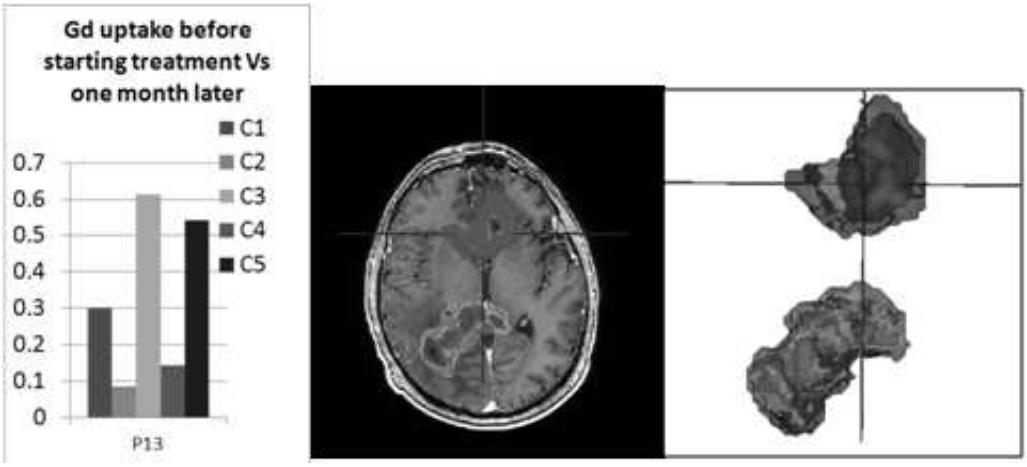


FIGURE 6.3: Comparaison des examens IRM avant et après traitement du patient P14. A gauche, les indices locaux calculés pour le patient P14 avec V1 = volume IRM avant traitement et V2 = volume IRM après traitement. Les classes en rouge et orange C1 et C2 montrent que malgré une décroissance globale (comparées aux classes en bleu et bleu sombre) il existe une décroissance locale de la tumeur. Au milieu, la visualisation 2D d'une coupe IRM, avec le volume tumoral détecté avant traitement en rouge et après traitement en bleu corrobore ce résultat. A droite, les volumes tumoraux avant traitement en rouge et après traitement en bleu sont affichés en 3D.

Il n'y a donc pas de diminution notable du volume tumoral au cours du traitement dans ces deux cas. Ces deux patients ont également des valeurs de l'indice de Dice élevées. La valeur de $C5_{b-a}^{IRM}$ est faible (<0,08) pour 14 sujets (tous sauf P14), le volume tumoral détecté sur l'IRM après injection de Gadolinium après traitement est donc dans sa plus grande partie inclus dans le volume avant traitement pour ces patients, indiquant l'absence d'augmentation locale du volume tumoral. Le patient P14 présente une valeur de $C1_{b-a}^{IRM}$ non négligeable (0,30) et une valeur de $C5_{b-a}^{IRM}$ importante (0,54) indiquant que malgré une augmentation globale du volume tumoral, il y a eu une diminution locale d'une partie de la tumeur au cours du traitement comme le montre la figure 6.3.

Les corrélations entre l'indice $C5_{b-a}^{IRM}$ et les indices de survie PFS et OS ont ensuite été calculées et sont reportées dans le tableau 6.6.

	$C5_{b-a}^{IRM}$ et PFS	$C5_{b-a}^{IRM}$ et OS
Seg IRM	NS	$R^2 = 0,16 \| p = 0,17 -> NS$ a = -0,04 ;b = 15,8

TABLE 6.6: Corrélation entre l'indice local C5 (utilisé ici pour comparer les examens IRM avant et après traitement) avec PFS et OS (pour 13 points). Avec "NS" : non significatif.

Aucune corrélation n'a été constatée entre l'indice $C5_{b-a}^{IRM}$ et PFS d'une part et OS d'autre part.

6.3.3 Conclusion

Nous avons vu que sur la base de données considérée pour notre étude, hormis pour le patient P14, le volume tumoral estimé sur l'IRM après injection de gadolinium diminuait au cours du traitement antiangiogénique. Plus précisément, suivant les seuils utilisés sur les variations de volumes pour établir les critères RANO, un seul patient est considéré comme ne réagissant pas au traitement, 8 patients sont stables et 6 patients montrent une réponse partielle au traitement. En revanche ces variations ne sont pas liées aux différentes durées de survie car ni les indices globaux ΔV ni les indices locaux $C5_{b-a}^{IRM}$ ne sont corrélés avec les indices PFS et OS. Les variations du volume mesurées sur les IRM pondérées en T1 après injection de Gadolinium ne sont donc pas prédictives d'une durée de survie.

6.4 Comparaisons globales et locales des TEMP avant et après le début du traitement

6.4.1 Analyse des volumes et intensités avant et après traitement

6.4.1.1 Etude globale

Afin de suivre l'évolution de la tumeur avec les données TEMP, nous avons étudié les variations des volumes tumoraux et les variations d'intensité au sein de ces volumes de manière globale en utilisant les 5 méthodes de segmentation (Seg1, Seg2, Seg3, Seg4, Seg5) décrites au paragraphe 4.4.1.2 sur les données TEMP précoces (15 minutes après injection du Sestamibi) et TEMP tardives (3 h après injection du Sestamibi). Les seuils utilisés pour les quatre premières méthodes de segmentation sont indiqués pour chaque jeu de données dans le tableau 6.7. Pour les segmentations utilisant la méthode Seg3, les seuils utilisés sont ceux de la méthode Seg2 pour les images avant traitement, que ce soit pour les images précoces ou tardives. Pour les segmentations utilisant la méthode Seg4, les seuils utilisés sont ceux de la méthode Seg2 pour les images après traitement, que ce soit pour les images précoces ou tardives.

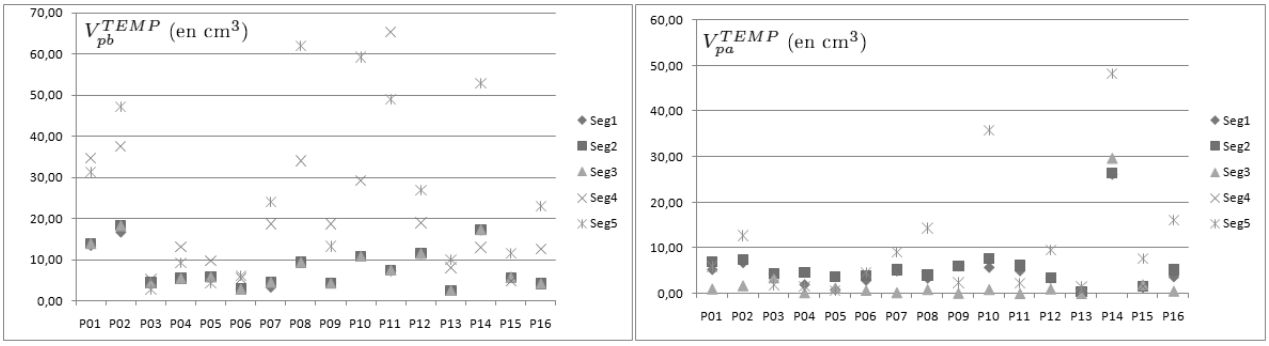
	Seg1				Seg2			
					Seg3	Seg4	Seg3	Seg4
	15 min (p)		3h (t)		15 min (p)		3h (t)	
	avant (b)	après (a)	avant (b)	après (a)	avant (b)	après (a)	avant (b)	après (a)
P01	9,3	5,6	8,1	6,2	9,3	5,6	8,1	7,1
P02	12,5	7,8	14,5	8,6	12,5	7,8	14,5	8,6
P03	5,2	4,8	7,3	5,9	5,2	4,8	8,2	5,9
P04	8,4	4,1	11,6	4,0	8,4	4,1	11,6	4,0
P05	5,3	2,7	5,2	2,3	5,3	4,1	5,2	2,6
P06	12,8	6,8	13,9	6,0	12,8	6,8	13,9	6,0
P07	16,4	7,7	15,6	6,6	16,4	7,7	15,6	6,6
P08	18,0	10,8	18,7	13,8	18,0	10,8	18,7	13,8
P09	10,1	4,6	11,3	6,7	10,1	4,6	11,3	6,7
P10	37,4	20,4	37,4	34,6	37,4	20,4	37,4	34,6
P11	16,8	4,2	8,2	4,4	16,8	4,2	8,2	4,4
P12	11,9	8,6	10,5	11,1	11,9	8,6	10,5	11,1
P13	22,1	8,4	22,1	4,6	22,1	8,4	22,1	7,8
P14	12,3	13,9	14,7	19,8	12,3	13,9	14,7	19,8
P15	15,6	17,1	24,5	12,9	15,6	17,1	24,5	12,9
P16	22,9	13,1	13,2	8,8	22,9	13,1	13,2	8,8

TABLE 6.7: Seuils utilisés pour les segmentations Seg1, Seg2, Seg3 et Seg4

Le seuil pour la cinquième méthode de segmentation Seg5 est de 6 quel que soit le jeu de données considéré. Ce seuil a été choisi sur les TEMP normalisées de manière à pouvoir distinguer l'activité tumorale des autres zones.

A partir de ces seuils, les volumes tumoraux ont été extraits des images TEMP normalisées permettant de déterminer l'intensité moyenne au sein de ces volumes. Les variations de volumes et variations d'intensité moyenne au sein de ces volumes au cours du traitement pour les images TEMP précoces et tardives selon les cinq méthodes de segmentation ont ainsi pu également être calculées.

La figure 6.4 montre la mesure des volumes détectés sur les images TEMP précoces avant traitement (V_{pb}^{TEMP}) et après traitement (V_{pa}^{TEMP}) pour les cinq méthodes de segmentation.



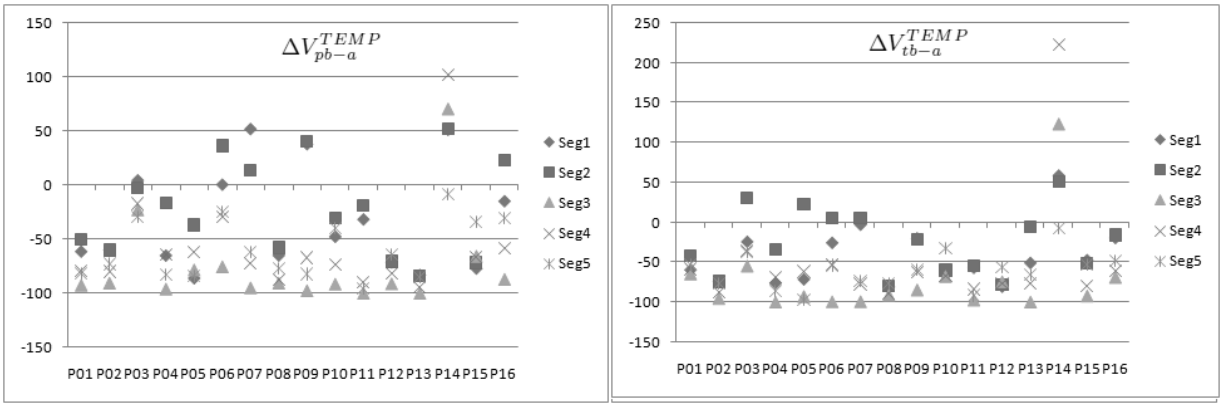
(a) V_{pb}^{TEMP} (en cm^3) (b) V_{pa}^{TEMP} (en cm^3)

FIGURE 6.4: Volumes détectés sur les images TEMP (a) avant traitement V_{pb}^{TEMP} et (b) après traitement V_{pa}^{TEMP}

La figure 6.4 montre qu'en utilisant la segmentation Seg3, il existe des cas où quasiment aucun volume tumoral n'est détecté sur l'examen après traitement. Pour les patients P11 et P13, aucun volume tumoral n'est détecté sur la TEMP après traitement et pour les patients P04, P07 et P09 le volume tumoral détecté sur la TEMP après traitement est inférieur à 0,3 cm^3. Ceci implique que pour cette segmentation, il existe des cas où la variation relative de volume est proche de -100% du fait que le volume détecté a totalement disparu lors du traitement. Cela est dû à la valeur élevée du seuil utilisé pour cette segmentation (cf tableau 6.7 et figure 6.4).

La figure 6.5 montre les variations de volumes au cours du traitement sur les images TEMP précoces ΔV_{pb-a}^{TEMP} et tardives ΔV_{tb-a}^{TEMP} obtenues avec les différentes méthodes de segmentation.

Pour une première analyse sur les variations de volumes, nous avons réutilisé les seuils définis dans les critères RANO, bien que ceux-ci aient été définis uniquement pour les volumes détectés en IRM après injection de Gadolinium. Si $\Delta V_{pb-a}^{TEMP} <$ -50% on considère que le volume diminue, si 25% $< \Delta V_{pb-a}^{TEMP} <$ -50% on considère qu le volume est stable et si $\Delta V_{pb-a}^{TEMP} >$ 25% on considère que le volume augmente. Les mêmes seuils sont à considérer pour les TEMP tardives.



(a) ΔV_{pb-a}^{TEMP} (b) ΔV_{tb-a}^{TEMP}

FIGURE 6.5: Variations relatives de volumes (exprimées en %) au cours du traitement pour les images TEMP précoces ΔV_{pb-a}^{TEMP} et tardives ΔV_{tb-a}^{TEMP}

Avec la segmentation Seg1, sur les images précoces, le volume diminue au cours du traitement pour les 9 patients P01, P02, P04, P05, P08, P10, P12, P13, P15 ; le volume est stable pour les 4 patients P03, P06, P11 et P16 ; enfin le volume augmente pour les 3 patients P07, P09 et P14. Avec la segmentation Seg1, sur les images tardives, le volume diminue au cours du traitement pour les 10 patients P01, P02, P04, P05, P08, P10, P11, P12, P13, P15 ; le volume est stable pour les 5 patients P03, P06, P07, P09 et P16 ; enfin le volume augmente pour le patient P14.

Avec la segmentation Seg2, sur les images précoces, le volume diminue au cours du traitement pour les 6 patients P01, P02, P08, P12, P13, P15 ; le volume est stable pour les 6 patients P03, P04, P05, P07, P10, P11 ; enfin le volume augmente pour les 4 patients P06, P09, P14 et P16. Sur les images tardives, le volume diminue au cours du traitement pour les 7 patients P01, P02, P08, P10, P11, P12, P15 ; le volume est stable pour les 6 patients P04, P06, P07, P09, P13 et P16 ; enfin le volume augmente pour les 3 patients P03, P05, P14.

Avec la segmentation Seg3, sur les images précoces, le volume diminue au cours du traitement pour les 14 patients P01, P02, P04, P05, P06, P07, P08, P09, P10, P11, P12, P13, P15, P16 ; le volume est stable pour le patient P03 ; enfin le volume augmente pour le patient P14. Sur les images tardives, le volume diminue au cours du traitement pour tous les patients excepté le patient P14 pour lequel le volume augmente.

Avec la segmentation Seg4, sur les images précoces, le volume diminue au cours du traitement pour les 13 patients P01, P02,P04, P05, P07, P08, P09, P10, P11, P12, P13, P15, P16 ; le volume est stable pour les 2 patients P03, P06 ; enfin le volume augmente pour le patient P14. Sur les images tardives, le volume diminue au cours du traitement pour les 14 patients P01, P02, P04, P05, P06, P07, P08, P09, P10, P11, P12, P13, P15, P16 ; le volume est stable pour le patient P03 ; enfin le volume augmente pour le patient P14.

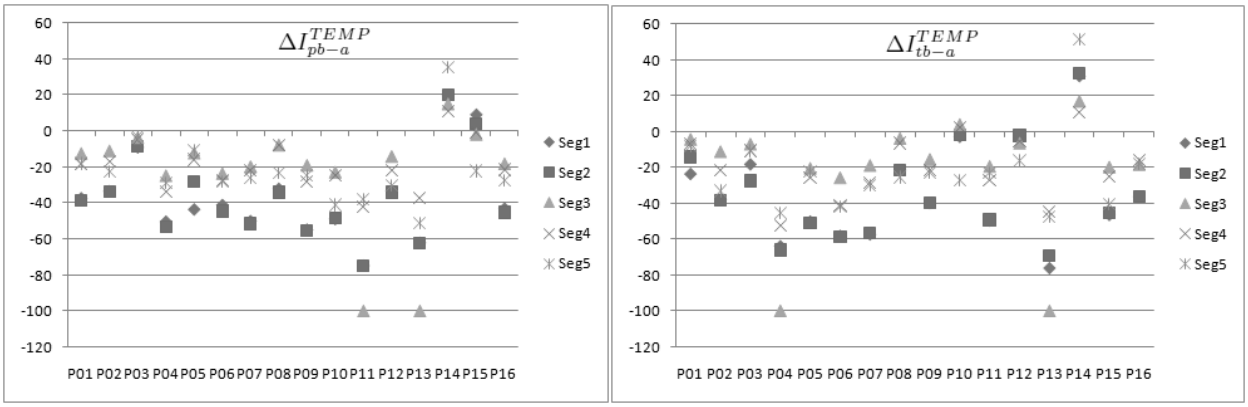
Avec la segmentation Seg5, sur les images précoces, le volume diminue au cours du traitement pour les 10 patients P01, P02, P04, P05, P07, P08, P09, P11, P12, P13 ; le volume est stable pour les 6 patients P03, P06, P10, P14, P15, P16. Sur les images tardives, le volume diminue au cours du traitement pour les 13 patients P01, P02, P04, P05, P06, P07, P08, P09, P11, P12, P13, P15, P16 ; le volume est stable pour les 3 patients P03, P10, P14.

Les variations de volume relatif peuvent varier et modifier largement la classification des patients en fonction des segmentations. Par exemple, la figure 6.5a montre pour le patient P14, une forte augmentation du volume tumoral au cours du traitement avec les segmentations Seg1, Seg2, Seg3 et Seg4 tandis que pour la segmentation Seg5 le volume évolue très peu. Les variations de volumes sur les images TEMP précoces pour le patient P07 montre également une contradiction entre les résultats obtenus par les méthodes Seg1 et Seg2 et ceux obtenus par Seg3, Seg4 et Seg5 : une augmentation du volume tumoral allant jusqu'à 50% pour Seg1 et une diminution du volume tumoral dépassant les 50% avec les 3 dernières segmentations.

D'autre part, suivant le type d'image utilisé, précoce ou tardif, les résultats relatifs aux variations de volumes ne sont pas en concordance. Par exemple pour le patient P07, la segmentation Seg1 sur l'image précoce indique une forte augmentation du volume tumoral au cours du traitement (>50%) tandis que sur l'image tardive le volume est plutôt stable.
A priori, les variations de volume sur les données TEMP ne semblent donc pas être un critère robuste et ces deux éléments seront analysés de façon plus approfondie dans la suite du chapitre en étudiant les scores de corrélation entre les différentes segmentations, puis en étudiant les corrélations entre les images précoces et tardives.

La figure 6.6 montre les variations d'intensité moyenne au sein des volumes détectés au cours du traitement sur les images TEMP précoces ΔI_{pb-a}^{TEMP} et tardives ΔI_{tb-a}^{TEMP} pour les cinq méthodes de segmentations.

Pour analyser les variations d'intensité, il n'existe pas de seuil prédéfini qui permette de dire si la tumeur progresse, reste stable, ou si le patient réagit au traitement. Nous avons simplement considéré les augmentations et les diminutions de ces variations relatives d'intensité. La figure 6.6 montre que quelle que soit la segmentation considérée, l'intensité diminue au cours du traitement pour tous les patients sauf les patients P14 et P15 sur les images précoces. Sur les images tardives, l'intensité diminue au cours du traitement pour tous les patients sauf le patient P14. Ainsi les variations relatives d'intensité semblent plus stables que les



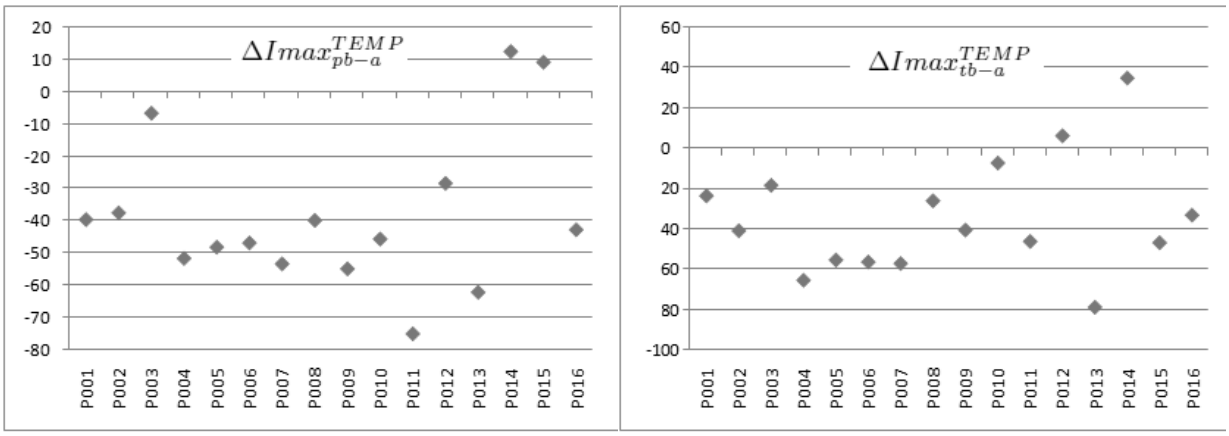
(a) ΔI_{pb-a}^{TEMP} (b) ΔI_{tb-a}^{TEMP}

FIGURE 6.6: Variations relatives d'intensité moyenne (exprimées en %) au cours du traitement pour les images TEMP précoces ΔI_{pb-a}^{TEMP} et tardives ΔI_{tb-a}^{TEMP}

variations relatives de volume et sont beaucoup moins dépendantes de la méthode de segmentation choisie. De plus les résultats entre images précoces et tardives sont quasiment similaires. Il est à noter qu'il existe, pour certains sujets, des variations importantes d'intensité pour la segmentation Seg3 qui sont directement liées au fait que le volume de la tumeur est très réduit après traitement pour ces sujets.

La figure 6.7 montre les variations d'intensité maximale au sein des volumes détectés au cours du traitement sur les images TEMP précoces $\Delta Imax_{pb-a}^{TEMP}$ et tardives $\Delta Imax_{tb-a}^{TEMP}$.

Les deux figures 6.7a et 6.7b montrent que les variations relatives d'intensité maximale suivent à peu près les mêmes variations que les variations relatives d'intensité moyenne.



(a) $\Delta Imax_{pb-a}^{TEMP}$ (b) $\Delta Imax_{tb-a}^{TEMP}$

FIGURE 6.7: Variations relatives d'intensité maximale (exprimées en %) au cours du traitement pour les images TEMP précoces $\Delta Imax_{pb-a}^{TEMP}$ et tardives $\Delta Imax_{tb-a}^{TEMP}$

6.4.1.2 Etudes locales des volumes et intensités avant et après traitement

6.4.1.2.1 Etudes locales des volumes et intensités avant et après traitement sur les TEMP précoces Les tableaux 6.8 et 6.9 donnent les résultats des indices locaux $Dice_{pb-a}^{TEMP}$, $C1_{pb-a}^{TEMP}$ et $C5_{pb-a}^{TEMP}$ pour les TEMP précoces.

	$Dice_{pb-a}^{TEMP}$	$C1_{pb-a}^{TEMP}$	$C5_{pb-a}^{TEMP}$
Seg1	min = 0,02 max = 0,58 moyenne = 0,33 écart-type =0,16	min = 0,08 max = 0,93 moyenne = 0,39 écart-type = 0,24	min = 0 max = 0,88 moyenne = 0,17 écart-type = 0,23
Seg2	min = 0,04 max = 0,66 moyenne = 0,39 écart-type =0,15	min = 0,02 max = 0,67 moyenne = 0,29 écart-type = 0,17	min = 0 max = 0,61 moyenne = 0,18 écart-type = 0,17
Seg3	min = 0,23 max = 0,48 moyenne = 0,18 écart-type =0,16	min = 0,02 max = 0,87 moyenne = 0,62 écart-type = 0,22	min = 0 max = 0,46 moyenne = 0,06 écart-type = 0,13
Seg4	min = 0,09 max = 0,71 moyenne = 0,34 écart-type =0,15	min = 0,08 max = 0,76 moyenne = 0,42 écart-type = 0,20	min = 0 max = 0,57 moyenne = 0,06 écart-type = 0,15
Seg5	min = 0,03 max = 0,77 moyenne = 0,40 écart-type =0,24	min = 0,05 max = 0,82 moyenne = 0,40 écart-type = 0,23	min = 0 max = 0,20 moyenne = 0,03 écart-type = 0,06

TABLE 6.8: Indices locaux liés aux volumes pour comparer les examens TEMP précoces avant et après traitement

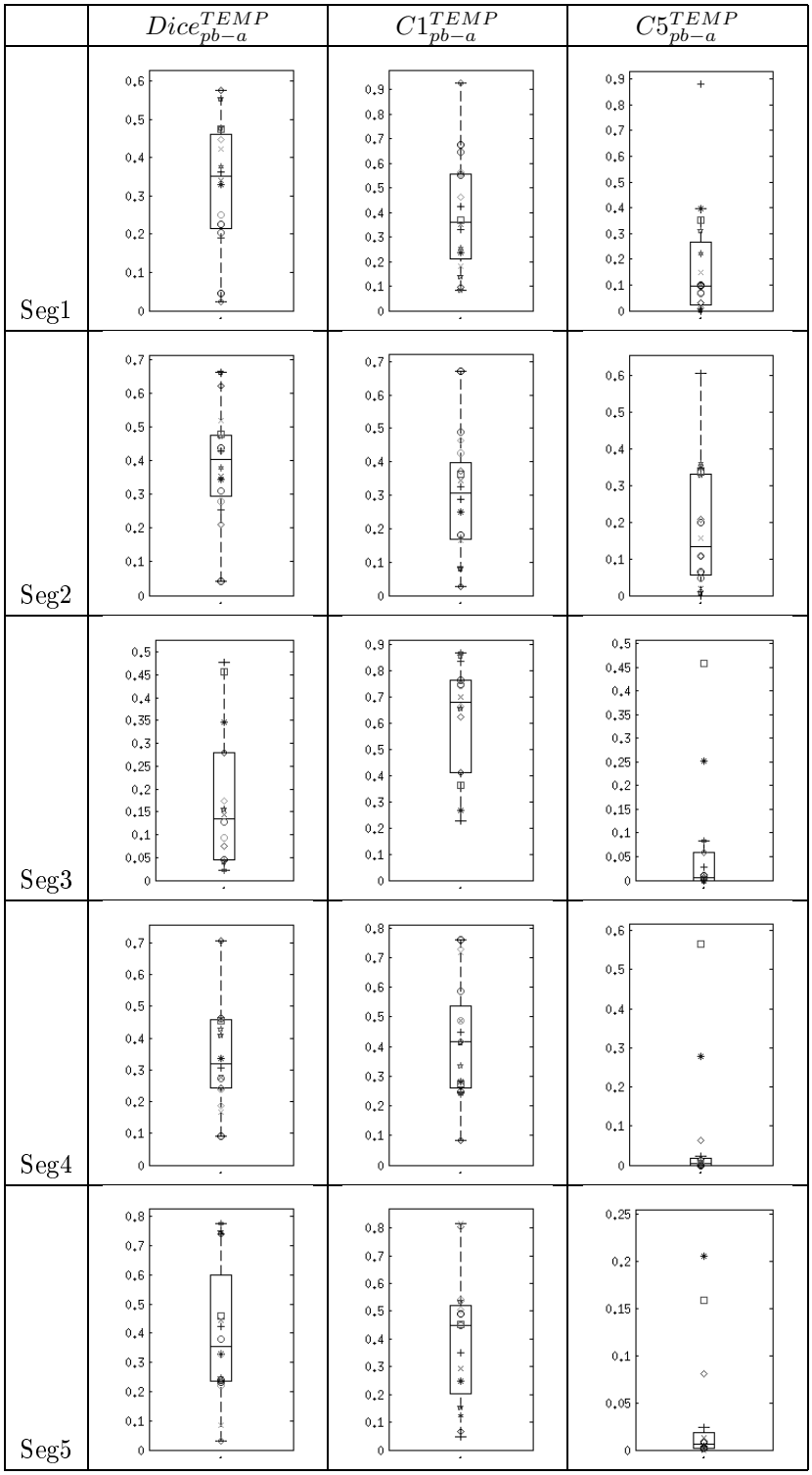


TABLE 6.9: Indices locaux liés à la comparaison des volumes entre les TEMP précoces avant et après traitement (calculés sur 16 points)

Le tableau 6.9 montre que les résultats dépendent des segmentations. Cependant, plusieurs tendances se dessinent. L'indice de Dice est assez proche en moyenne pour les 4 segmentations Seg1, Seg2, Seg4 et Seg5. Il est en revanche plus faible pour Seg3. Ceci peut s'expliquer par la réduction importante de la taille des tumeurs avec cette segmentation. Quoiqu'il en soit, l'indice de Dice est plutôt faible pour toutes les segmentations ($< 0{,}6$) à quelques exceptions près. L'indice $C1_{pb-a}^{TEMP}$ est élevé ($> 0{,}2$) et $C5_{pb-a}^{TEMP}$ est faible ($< 0{,}2$) dans un grand nombre de cas (tous les patients sauf P14 et P03) pour les segmentations Seg3, Seg4 et Seg5. Ainsi, le volume après traitement est souvent compris presque entièrement dans le volume avant traitement sur les TEMP précoces. Par contre, pour les 2 premières segmentations, $C5_{pb-a}^{TEMP}$ se révèle important ($> 0{,}2$) dans la moitié des cas où le volume délimité après traitement n'est pas entièrement inclus dans le volume délimité avant traitement.

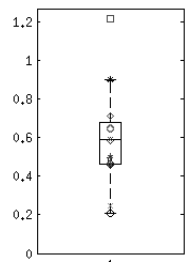
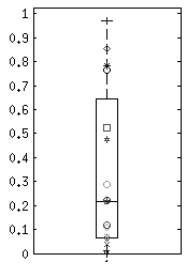
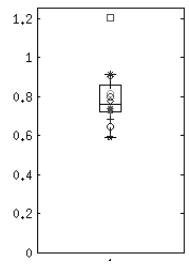
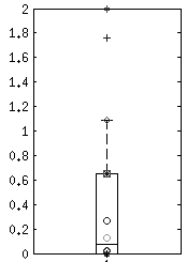
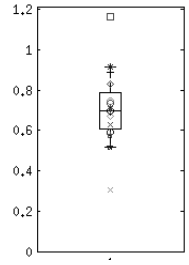
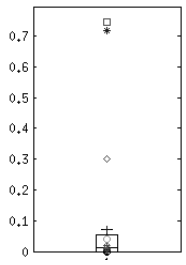
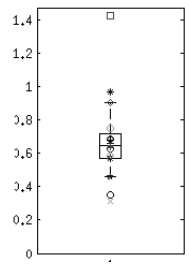
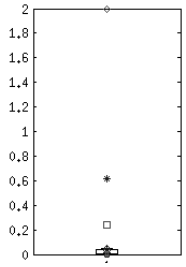
	$L3a_{p-ba}^{TEMP}$	$L5_{p-ba}^{TEMP}$
Seg2		
Seg3		
Seg4		
Seg5		

TABLE 6.10: Indices locaux liés à la comparaison des intensités entre les examens TEMP précoces avant et après traitement (calculés sur 16 points).

Le tableau 6.10 montre que les valeurs de $L3a_{p-ba}^{TEMP}$ et $L5_{p-ba}^{TEMP}$ dépendent également des segmentations. L'indice $L5_{p-ba}^{TEMP}$ est important (>0,2) dans la moitié des cas pour les segmentations Seg2 et Seg3. Pour les autres méthodes de segmentation, Seg4 et Seg5, $L5_{p-ba}^{TEMP}$ est relativement faible ($< 0{,}2$) sauf dans 3 cas (P14, P03 et P05). Le patient P14 correspond à un cas pour lequel le volume tumoral global augmente au cours du traitement. Dans les deux autres cas en revanche, le volume tumoral diminue pour ces méthodes de segmentation. Pour le patient P03, malgré la diminution globale du volume, l'indice $C5_{p-ba}^{TEMP}$ est important ($> 0{,}25$) et l'intensité dans cette zone semble élevée également comme le suggère la valeur de $L5_{p-ba}^{TEMP}$ ($> 0{,}6$). Pour le patient P05, l'indice $C5_{p-ba}^{TEMP}$ est plutôt faible ($< 0{,}1$) mais l'intensité dans cette zone est élevée comme l'indique la valeur de $L5_{p-ba}^{TEMP}$ ($> 0{,}25$). Cela suggère dans les deux cas une évolution locale de la tumeur comme l'illustre la figure 6.8. Cette figure montre qu'une partie du volume tumoral détectée sur l'image TEMP précoce avant traitement (en rouge) du patient P03 se trouve en dehors du volume tumoral détecté sur l'image TEMP précoce après traitement (en bleu), laissant supposer une modification de la localisation tumorale.

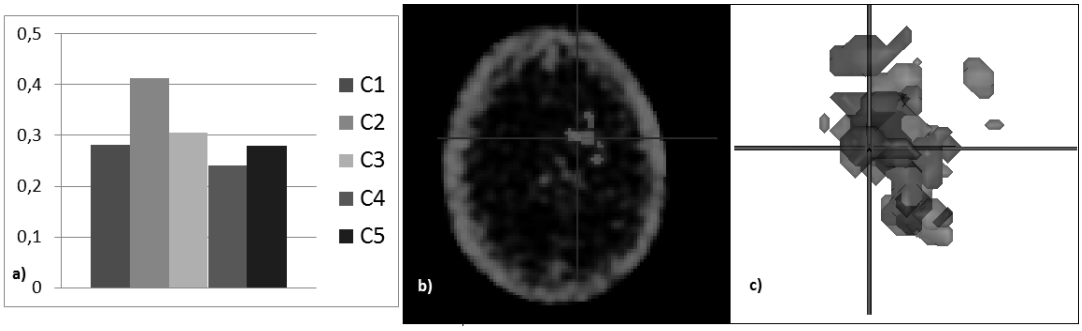


FIGURE 6.8: Comparaison des images TEMP précoces avant et après le début du traitement pour le patient P03 qui suggèrent, malgré une diminution globale du volume tumoral, une augmentation locale de la tumeur. Sur les images de coupe b) et de représentation 3D c), le volume rouge représente la zone tumorale détectée sur la TEMP précoce avant traitement et le volume bleu la zone tumorale détectée sur la TEMP précoce après traitement avec la segmentation Seg4. A gauche a) sont affichées les valeurs des indices de comparaison de volume.

6.4.1.2.2 Etudes locales des volumes et intensités avant et après traitement sur les TEMP précoces Les tableaux 6.11 et 6.12 donnent les résultats des indices locaux pour les TEMP tardives.

	$Dice_{tb-a}^{TEMP}$	$C1_{tb-a}^{TEMP}$	$C5_{tb-a}^{TEMP}$
Seg1	min = 0,01 max = 0,58 moyenne = 0,31 écart-type =0,15	min = 0,13 max = 0,91 moyenne = 0,41 écart-type = 0,22	min = 0 max = 0,72 moyenne = 0,14 écart-type = 0,29
Seg2	min = 0,10 max = 0,57 moyenne = 0,35 écart-type =0,15	min = 0,08 max = 0,58 moyenne = 0,30 écart-type = 0,17	min = 0 max = 0,69 moyenne = 0,21 écart-type = 0,24
Seg3	min = 0 max = 0,50 moyenne = 0,19 écart-type =0,17	min = 0,17 max = 1 moyenne = 0,64 écart-type = 0,27	min = 0 max = 0,63 moyenne = 0,07 écart-type = 0,17
Seg4	min = 0,17 max = 0,49 moyenne = 0,31 écart-type =0,12	min = 0,06 max = 0,70 moyenne = 0,40 écart-type = 0,17	min = 0 max = 1,21 moyenne = 0,10 écart-type = 0,30
Seg5	min = 0,00 max = 0,79 moyenne = 0,38 écart-type =0,20	min = 0,09 max = 0,97 moyenne = 0,38 écart-type = 0,23	min = 0 max = 0,19 moyenne = 0,03 écart-type = 0,06

TABLE 6.11: Indices locaux liés à la comparaison des volumes entre les examens TEMP tardifs avant et après traitement

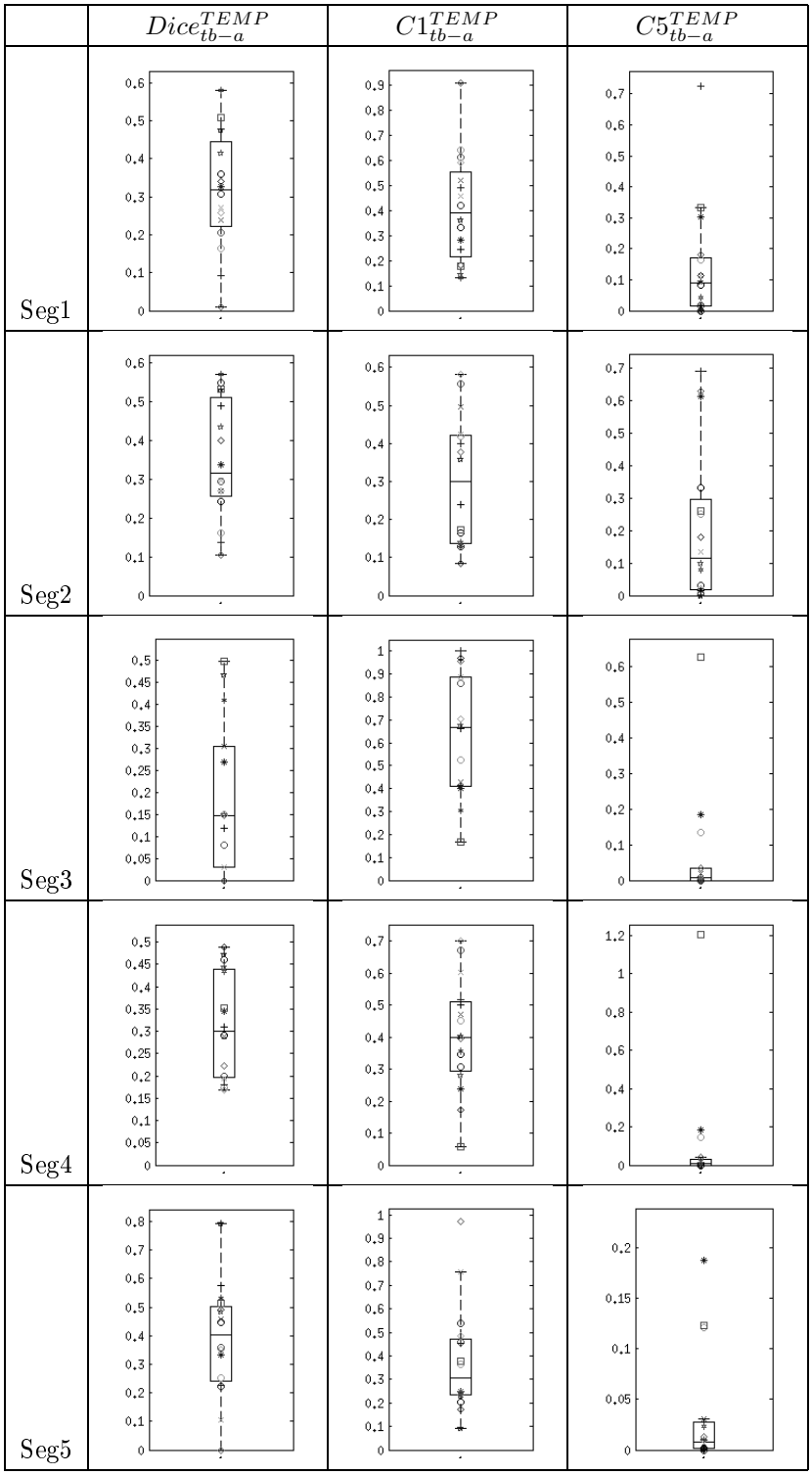


TABLE 6.12: Indices locaux liés à la comparaison des volumes entre les examens TEMP tardifs avant et après traitement (calculés sur 16 points)

Le tableau 6.12 montre, comme pour les examens TEMP précoces, que les résultats dépendent des segmentations. L'indice $C1_{pb-a}^{TEMP}$ est important ($> 0{,}2$) et l'indice $C5_{pb-a}^{TEMP}$ est faible ($< 0{,}2$) pour tous les patients sauf P14, P03 et P01 pour les segmentations Seg3, Seg4 et Seg5. Dans les autres cas, le volume après traitement est presque entièrement compris dans le volume avant traitement sur les TEMP tardives. Par contre, pour les 2 premières segmentations, $C5_{pb-a}^{TEMP}$ se révèle important dans un grand nombre de jeux de données (P14, P03, P07, P05 et P01), ce qui indique que les volumes tumoraux détectés après traitement sont en dehors de la tumeur initialement détectée avec ces segmentations. Globalement les remarques faites précédemment sur les études précoces s'appliquent également aux études tardives.

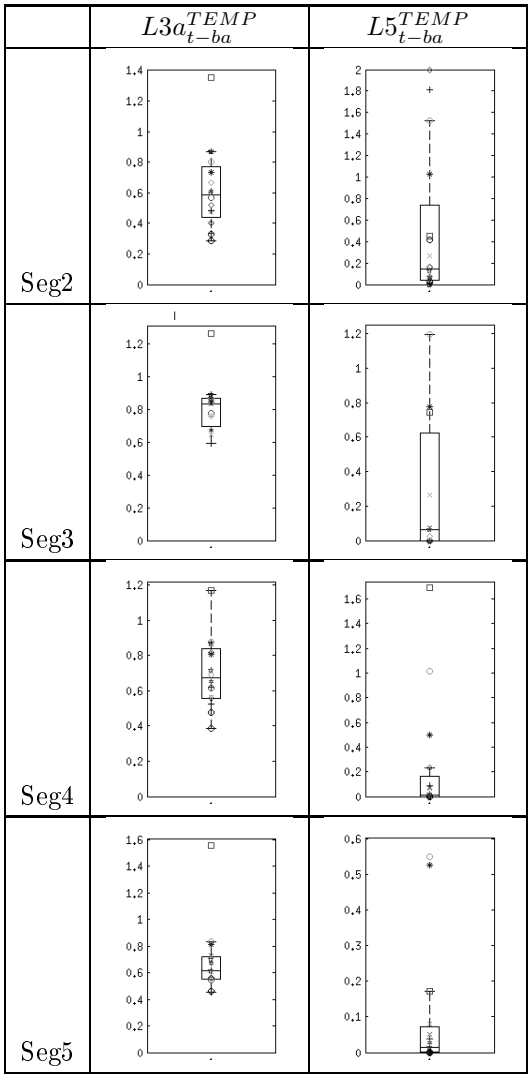


TABLE 6.13: Indices locaux liés à la comparaison des intensités entre les examens TEMP tardifs avant et après traitement (calculés sur 16 points)

Le tableau 6.13 montre que les résultats de $L3a_{t-ba}^{TEMP}$ et $L5_{t-ba}^{TEMP}$ dépendent également des segmentations. L'indice $L5_{t-ba}^{TEMP}$ est important ($>$ 0,2) dans la moitié des cas pour la segmentation Seg2. Pour les autres méthodes de segmentation Seg3, Seg4 et Seg5, l'indice $L5_{t-ba}^{TEMP}$ est faible ($<$ 0,2) dans tous les cas sauf dans 3 cas (P14, P03 et P01). Le patient P14 correspond à un cas pour lequel le volume tumoral global augmente au cours du traitement. Dans les deux autres cas en revanche, le volume global tumoral diminue globalement. Pour le patient P03, malgré la diminution de volume, l'indice $C5_{p-ba}^{TEMP}$ est très élevé ($>$ 1) et l'intensité dans cette zone est élevée également, comme le montre la valeur de $L5_{p-ba}^{TEMP}$ ($>$ 0,4). Pour le patient P01, l'indice $C5_{p-ba}^{TEMP}$ a une valeur supérieure à 0,15 et l'intensité dans cette zone est élevée comme l'indique la valeur de $L5_{p-ba}^{TEMP}$ ($>$ 1). Ceci suggère qu'il y a une évolution locale de la tumeur pour ces patients. La figure 6.9 montre qu'une partie du volume tumoral détecté sur l'image TEMP tardive avant traitement (en rouge) du patient P03 se trouve en dehors du volume tumoral détecté sur l'image TEMP tardive après traitement (en bleu), laissant supposer le développement de la tumeur.

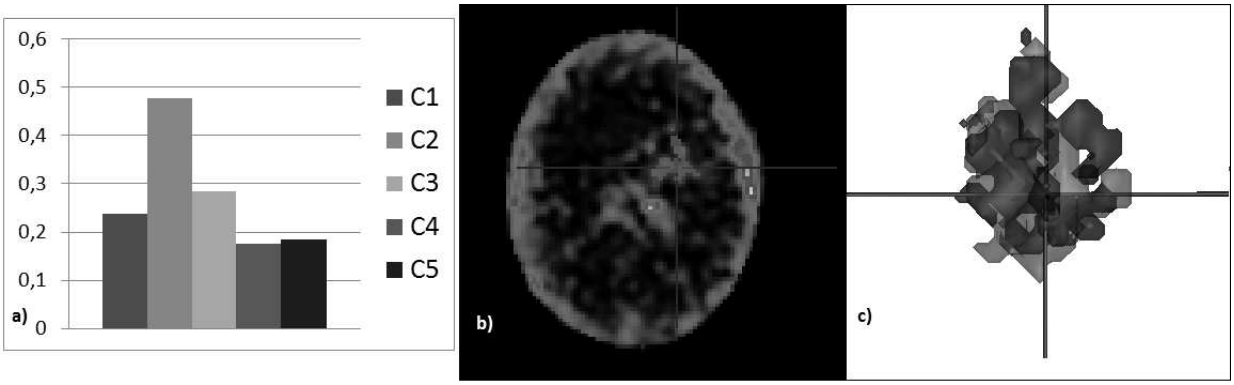


FIGURE 6.9: Résultats issus de la comparaison des images TEMP tardives avant et après le début du traitement pour le patient P03 qui suggèrent, malgré une diminution globale du volume tumoral, une augmentation locale de la tumeur. Sur les images de coupe b) et de représentation 3D c), le volume rouge représente la zone tumorale détectée sur la TEMP tardive avant traitement, et le volume bleu la zone tumorale détectée sur la TEMP tardive après traitement avec la segmentation Seg4. A gauche a) sont affichées les valeurs de l'indice de comparaison de volumes.

6.4.1.3 Conclusion

Ces premières analyses montrent que les résultats varient fortement en fonction de la méthode de segmentation. Malheureusement, la validation de méthodes de segmentation en TEMP est un problème complexe en l'absence de vérité absolue. C'est la raison pour laquelle, pour la suite des analyses, nous avons gardé les cinq méthodes de segmentation. En revanche, nous avons pu constater des similitudes entre les segmentations Seg1 et Seg2 ainsi qu'entre les méthodes Seg3, Seg4 et Seg5. Il convient donc de vérifier cette hypothèse en étudiant les corrélations possibles entre les différentes segmentations (cf §6.4.2).
L'utilisation des images TEMP précoces ou tardives peuvent également mener à des résultats quelque peu différents. C'est pourquoi une comparaison a été réalisée entre les deux modalités (cf §6.4.3).

6.4.2 Comparaison des différentes méthodes de segmentation

La segmentation de volume tumoraux sur des données TEMP dans le cadre du suivi de patients est un problème ouvert. Dans la littérature, plusieurs méthodes ont été proposées pour la segmentation de volumes tumoraux : des méthodes basées sur le seuillage [93], d'autres sur le gradient et s'appuyant sur de l'information morphologique [169], ou encore des approches par contours actifs [170]. Dans le cadre de cette étude, nous nous sommes focalisés sur des méthodes de segmentation par seuillage utilisant la valeur maximale des données TEMP normalisées. Contrairement à une étude d'évaluation cherchant à déterminer la meilleure méthode à utiliser (qui reste un problème difficile en l'absence de vérité terrain), nous voulions déterminer si ces méthodes menaient à l'extraction d'indices similaires ou non, comme le volume tumoral, l'intensité moyenne au sein de ce volume ainsi que les variations relatives de ces indices au cours du traitement.

Nous nous proposons donc dans cette partie de comparer les résultats obtenus par les différentes segmentations en fonction des indices globaux V^{TEMP}, ΔV^{TEMP}, I^{TEMP} et ΔI^{TEMP}, pour les images précoces et tardives.

Les tableaux 6.14, 6.15, 6.16, et 6.17 indiquent les scores de corrélation suivant les indices globaux V_p^{TEMP}, V_t^{TEMP}, ΔV_{pb-a}^{TEMP} et ΔV_{tb-a}^{TEMP} obtenus avec deux méthodes parmi les cinq méthodes de segmentation.

	Seg1	Seg 2	Seg3	Seg4	Seg5
Seg1	$R^2 = 1$	$R^2 = 0,98$ a = 0,98 ; b = 0,78	$R^2 = 0,90$ a = 1,16 ; b = -2,07	$R^2 = 0,30$ a = 1,40 ; b = 4,34	$R^2 = 0,53$ a = 2,52 ; b = 2,70
Seg2		$R^2 = 1$	$R^2 = 0,87$ a =1,15 ; b = -2,75	$R^2 = 0,29 \vert p = 0,002$ a =1,39 ; b = 3,52	$R^2 = 0,51$ a =2,52 ; b = 1,13
Seg3			$R^2 = 1$	$R^2 = 0,35$ a =1,25 ; b = 6,60	$R^2 = 0,53$ a =2,06 ; b = 7,80
Seg4				$R^2 = 1$	$R^2 = 0,59$ a =1,04 ; b = 5,09
Seg5					$R^2 = 1$

TABLE 6.14: Comparaison des volumes précoces V_p^{TEMP} pour les différentes segmentations (sur 32 points). R^2 est le score de corrélation entre les volumes V_p^{TEMP} segmentés par les différentes méthodes de segmentation, "a" et "b" respectivement le coefficient directeur et l'ordonnée à l'origine de la droite de corrélation et la valeur p qui en dessous de la valeur de 0,05 indique que les données sont significativement corrélées.

	Seg1	Seg 2	Seg3	Seg4	Seg5
Seg1	$R^2 = 1$	$R^2 = 0,96$ a = 0,97 ; b = 0,82	$R^2 = 0,90$ a = 1,28 ; b = -2,48	$R^2 = 0,57$ a = 1,51 ; b = 1,38	$R^2 = 0,65$ a = 2,90 ; b = 0,12
Seg2		$R^2 = 1$	$R^2 = 0,87$ a = 1,28 ; b = -3,20	$R^2 = 0,58$ a = 1,55 ; b = 0,21	$R^2 = 0,62$ a = 2,87 ; b = -1,40
Seg3			$R^2 = 1$	$R^2 = 0,52$ a = 1,07 ; b = 4,96	$R^2 = 0,65$ a = 2,14 ; b = 6,40
Seg4				$R^2 = 1$	$R^2 = 0,33$ a = 1,03 ; b = 7,58
Seg5					$R^2 = 1$

TABLE 6.15: Comparaison des volumes tardifs V_t^{TEMP} pour les différentes segmentations (sur 32 points). R^2 est le score de corrélation entre les volumes V_t^{TEMP} segmentés par les différentes méthodes de segmentation, "a" et "b" respectivement le coefficient directeur et l'ordonnée à l'origine de la droite de corrélation et la valeur p qui en dessous de la valeur de 0,05 indique que les données sont significativement corrélées.

	Seg1	Seg 2	Seg3	Seg4	Seg5
Seg1	$R^2 = 1$	$R^2 = 0,78$ a = 0,80 ; b = -4,65	$R^2 = 0,22$ a = 0,42 ; b = -61,9	$R^2 = 0,34$ a = 0,58 ; b = -39,7	$R^2 = 0,20$ a = 0,26 ; b =-51,7
Seg2		$R^2 = 1$	$R^2 = 0,19$ a = 0,44 ; b = -66,1	$R^2 = 0,39$ a = 0,68 ; b = -43,8	$R^2 = 0,22$ a = 0,29 ; b = -53,7
Seg3			$R^2 = 1$	$R^2 = 0,93$ a = 1,05 ; b = 21,2	$R^2 = 0,47$ a = 0,43 ; b = -27,5
Seg4				$R^2 = 1$	$R^2 = 0,51$ a = 0,41 ; b = -36,0
Seg5					$R^2 = 1$

TABLE 6.16: Comparaison des variations relatives de volumes précoces ΔV_{pb-a}^{TEMP} pour les différentes segmentations (sur 16 points). R^2 est le score de corrélation entre les variations relatives de volume ΔV_{pb-a}^{TEMP} obtenues par les différentes méthodes de segmentation, "a" et "b" respectivement le coefficient directeur et l'ordonnée à l'origine de la droite de corrélation et la valeur p qui en dessous de la valeur de 0,05 indique que les données sont significativement corrélées.

En terme de volumes, les segmentation Seg1, Seg2 et Seg3 fournissent des valeurs très corrélées à la fois pour les TEMP précoces et les TEMP tardives (cf tableaux 6.14 et 6.15). Les segmentations Seg4 et Seg5 sont corrélées plus faiblement avec les autres segmentations. Concernant les variations relatives de volumes, les segmentations Seg1 et Seg2 sont toujours très corrélées, mais la segmentation Seg3 est alors très corrélée avec la segmentation Seg4 à la fois pour les TEMP précoces et tardives (cf tableaux 6.16 et 6.17). La corrélation entre les segmentations Seg1 et Seg2 est logique puisque seule "la boîte englobante" dans laquelle est segmentée la tumeur est modifiée entre ces deux segmentations, les seuils étant quasiment les mêmes dans tous les cas pour tous les patients. La corrélation entre les segmentations Seg3 et Seg4 apparaissant sur les variations relatives de volume s'explique certainement par le fait qu'avec ces deux segmentations des seuils identiques mais dépendants de chaque patient sont appliqués sur les examens avant et après traitement rendant la comparaison plus pertinente.

.

Les tableaux 6.18, 6.19, 6.20 et 6.21 indiquent les scores de corrélation entre les indices globaux liés aux intensités et aux variations d'intensités pour les TEMP précoces et tardives en prenant toutes les combinaisons de deux méthodes parmi les cinq méthodes de segmentation.

	Seg1	Seg 2	Seg3	Seg4	Seg5
Seg1	$R^2 = 1$	$R^2 = 0,55$ a = 0,83 ; b = 10,75	$R^2 = 0,53$ a = 1,07 ; b = -26,3	$R^2 = 0,61$ a = 1,58 ; b = 16,84	$R^2 = 0,41$ a = 0,41 ; b =-42,9
Seg2		$R^2 = 1$	$R^2 = 0,24 \| p = 0,06 -> NS$ a = 0,64 ; b = -56,4	$R^2 = 0,36$ a = 1,10 ; b = -23,9	$R^2 = 0,11$ a = 0,19 ; b = -55,8
Seg3			$R^2 = 1$	$R^2 = 0,95$ a = 1,34 ; b = 45,7	$R^2 = 0,55$ a = 0,32 ; b = -37,1
Seg4				$R^2 = 1$	$R^2 = 0,45$ a = 0,21 ; b = -49,6
Seg5					$R^2 = 1$

TABLE 6.17: Comparaison des variations relatives de volumes tardifs ΔV_{tb-a}^{TEMP} pour les différentes segmentations (sur 16 points). R^2 est le score de corrélation entre les variations relatives de volume ΔV_{tb-a}^{TEMP} obtenues par les différentes méthodes de segmentation, "a" et "b" respectivement le coefficient directeur et l'ordonnée à l'origine de la droite de corrélation et la valeur p qui en dessous de la valeur de 0,05 indique que les données sont significativement corrélées.

	Seg1	Seg 2	Seg3	Seg4	Seg5
Seg1	$R^2 = 1$	$R^2 = 1,00$ a = 1,00 ; b = 0,00	$R^2 = 0,86$ a = 1,01 ; b = 1,06	$R^2 = 0,84$ a = 0,67 ; b = 2,63	$R^2 = 0,93$ a = 0,37 ; b = 5,19
Seg2		$R^2 = 1$	$R^2 = 0,86$ a =1,01 ; b = 1,05	$R^2 = 0,84$ a =0,67 ; b = 2,66	$R^2 = 0,93$ a =0,37 ; b = 5,19
Seg3			$R^2 = 1$	$R^2 = 0,79$ a =0,60 ; b = 3,12	$R^2 = 0,80$ a =0,31 ; b = 5,72
Seg4				$R^2 = 1$	$R^2 = 0,85$ a =0,48 ; b = 4,69
Seg5					$R^2 = 1$

TABLE 6.18: Comparaison des intensités précoces I_p^{TEMP} pour les différentes segmentations (sur 32 points).R^2 est le score de corrélation entre les intensités moyennes I_p^{TEMP} obtenues par les différentes méthodes de segmentation, "a" et "b" respectivement le coefficient directeur et l'ordonnée à l'origine de la droite de corrélation et la valeur p qui en dessous de la valeur de 0,05 indique que les données sont significativement corrélées.

	Seg1	Seg 2	Seg3	Seg4	Seg5
Seg1	$R^2 = 1$	$R^2 = 1,00$ a = 1,00 ; b = 0,20	$R^2 = 0,91$ a = 0,98 ; b = 1,42	$R^2 = 0,90$ a = 0,87 ; b = 0,18	$R^2 = 0,91$ a = 0,35 ; b = 5,39
Seg2		$R^2 = 1$	$R^2 = 0,91$ a = 0,98 ; b = 1,33	$R^2 = 0,90$ a = 0,88 ; b = -0,02	$R^2 = 0,90$ a = 0,36 ; b = 5,33
Seg3			$R^2 = 1$	$R^2 = 0,84$ a = 0,83 ; b = 0,10	$R^2 = 0,77$ a = 0,32 ; b = 5,64
Seg4				$R^2 = 1$	$R^2 = 0,77$ a = 0,36 ; b = 6,10
Seg5					$R^2 = 1$

TABLE 6.19: Comparaison des intensités tardives I_t^{TEMP} pour les différentes segmentations (sur 32 points). R^2 est le score de corrélation entre les intensités moyennes I_t^{TEMP} obtenues par les différentes méthodes de segmentation, "a" et "b" respectivement le coefficient directeur et l'ordonnée à l'origine de la droite de corrélation et la valeur p qui en dessous de la valeur de 0,05 indique que les données sont significativement corrélées.

	Seg1	Seg 2	Seg3	Seg4	Seg5
Seg1	$R^2 = 1$	$R^2 = 0,97$ a = 0,97 ; b = -1,00	$R^2 = 0,53$ a = 0,93 ; b = 10,40	$R^2 = 0,89$ a = 0,53 ; b = -0,06	$R^2 = 0,60$ a = 0,60 ; b =-0,46
Seg2		$R^2 = 1$	$R^2 = 0,54$ a = 0,95 ; b = 11,2	$R^2 = 0,91$ a = 0,54 ; b = 0,44	$R^2 = 0,67$ a = 0,64 ; b = 0,98
Seg3			$R^2 = 1$	$R^2 = 0,63$ a = 0,35 ; b = -11,2	$R^2 = 0,45$ a = 0,41 ; b = -12,8
Seg4				$R^2 = 1$	$R^2 = 0,67$ a = 1,13 ; b = -0,51
Seg5					$R^2 = 1$

TABLE 6.20: Comparaison des variations d'intensité précoces ΔI_{pb-a}^{TEMP} pour les différentes segmentations (sur 16 points). R^2 est le score de corrélation entre les variations relatives d'intensité moyenne ΔI_{pb-a}^{TEMP} obtenues par les différentes méthodes de segmentation, "a" et "b" respectivement le coefficient directeur et l'ordonnée à l'origine de la droite de corrélation et la valeur p qui en dessous de la valeur de 0,05 indique que les données sont significativement corrélées.

	Seg1	Seg 2	Seg3	Seg4	Seg5
Seg1	$R^2 = 1$	$R^2 = 0,98$ a = 0,98 ; b = -0,46	$R^2 = 0,56$ a = 0,89 ; b = 8,20	$R^2 = 0,88$ a = 0,59 ; b = 0,47	$R^2 = 0,68$ a = 0,69 ; b =1,49
Seg2		$R^2 = 1$	$R^2 = 0,54$ a = 0,87 ; b = 7,58	$R^2 = 0,89$ a = 0,60 ; b = 0,62	$R^2 = 0,69$ a = 0,70 ; b = 1,63
Seg3			$R^2 = 1$	$R^2 = 0,72$ a = 0,46 ; b = -9,76	$R^2 = 0,38$ a = 0,44 ; b = -12,6
Seg4				$R^2 = 1$	$R^2 = 0,57$ a = 1,01 ; b = -2,33
Seg5					$R^2 = 1$

TABLE 6.21: Comparaison des variations d'intensité tardives ΔI_{tb-a}^{TEMP} pour les différentes segmentations (sur 16 points). R^2 est le score de corrélation entre les variations relatives d'intensité moyenne ΔI_{tb-a}^{TEMP} obtenues par les différentes méthodes de segmentation, "a" et "b" respectivement le coefficient directeur et l'ordonnée à l'origine de la droite de corrélation et la valeur p qui en dessous de la valeur de 0,05 indique que les données sont significativement corrélées.

Les tableaux 6.18, 6.19, 6.20 et 6.21 montrent que les intensités moyennes dans les zones segmentées sont très corrélées ($R^2 \geq 0,77$) que ce soit pour les TEMP précoces ou tardives, et ce pour toutes les

segmentations. Les segmentations Seg1, Seg2, Seg3 et Seg4 mènent à des intensités moyennes équivalentes. L'intensité moyenne au sein des structures segmentées par Seg5 est moins importante que pour les autres segmentations car les volumes détectés sont plus grands. Concernant les variations relatives d'intensité, les résultats produits par les segmentations Seg1, Seg2 et Seg4 sont fortement corrélés ($R^2 \geq 0,89$). Les résultats obtenus par les segmentations Seg3 et Seg5 sont moins corrélés. Ces résultats peuvent être dus au fait que ces deux méthodes de segmentation obtiennent pour les données de certains patients des volumes tumoraux très faibles .

6.4.2.1 Conclusion

Afin de segmenter les volumes tumoraux sur les données TEMP, cinq méthodes de segmentation basées sur un seuillage utilisant l'intensité maximale normalisée au sein d'une région d'intérêt ont été utilisées. Ces cinq méthodes ont été comparées entre elles en termes d'indices globaux comme V^{TEMP}, ΔV^{TEMP}, I^{TEMP} et ΔI^{TEMP}, pour les images précoces et tardives, afin de déterminer si elles apportaient les mêmes informations. Les segmentations Seg1 et Seg2 sont très corrélées en terme de volumes V^{TEMP} ($R^2 \geq 0,96$) et variations de volumes ΔV^{TEMP} ($R^2 \geq 0,78$). Seg3 et Seg4 sont très corrélées en terme de variations de volumes ΔV^{TEMP} ($R^2 \geq 0,93$). En-dehors de ces corrélations observées, chaque segmentation semble apporter une information différente en terme de volumes. Lorsqu'on s'intéresse à l'indice d'intensité I^{TEMP}, toutes les méthodes sont très corrélées entre elles. **L'intensité** moyenne dans les volumes segmentés apparaît donc être un indice **plus fiable que le volume** pour analyser les données. Il est bon de noter néanmoins que malgré une forte corrélation entre les valeurs de I^{TEMP} obtenues par Seg5 et les valeurs de I^{TEMP} obtenues par les autres méthodes de segmentation, les intensités moyennes extraites à l'aide de Seg5 sont plus faibles (a $<$ 0,40). En terme de variations d'intensité ΔI^{TEMP}, les résultats produits par Seg1, Seg4 et Seg5 sont fortement corrélés. En revanche, les résultats obtenus par les segmentations seg3 et Seg5 semblent moins corrélés, ce qui peut s'expliquer par le fait que les volumes obtenus par ces segmentations peuvent être très faibles sur les données après traitement pour Seg3 et dans certains cas pour Seg5.

6.4.3 Comparaison des données TEMP précoces et tardives

6.4.3.1 Etude globale

La deuxième question que nous nous sommes posée était de savoir si les données TEMP à 15 minutes apportaient la même information que les données TEMP à 3 heures. Nous avons donc observé les corrélations entre les volumes, les variations relatives de volume au cours du traitement, les intensités et les variations d'intensité au cours du traitement, à partir des images TEMP précoces et tardives suivant les cinq méthodes de segmentation utilisées (tableaux 6.22 et 6.24). Pour le calcul de la droite et du coefficient de corrélation, les valeurs des indices pour les images précoces se trouvent en abscisse et celles des images tardives en ordonnées. En plus de la valeur du coefficient de corrélation, nous avons indiqué le biais "d" obtenu par une analyse de Bland Altman pour la comparaison des volumes (cf tableau 6.23).

	V_p^{TEMP} et V_t^{TEMP}	ΔV_{pb-a}^{TEMP} et ΔV_{tb-a}^{TEMP}
Seg1	$R^2 = 0,79$ a = 0,87 ; b = 1,00	$R^2 = 0,73$ a = 0,66 ; b = -22,2
Seg2	$R^2 = 0,75$ a = 0,85 ; b = 1,20	$R^2 = 0,39$ a = 0,58 ; b = -12,7
Seg3	$R^2 = 0,88$ a= 1,01 ; b = 0,50	$R^2 = 0,85$ a = 1,16 ; b = 14,7
Seg4	$R^2 = 0,82$ a = 0,70 ; b = 2,05	$R^2 = 0,91$ a = 1,5 ; b = 36,0
Seg5	$R^2 = 0,90$ a = 0,96 ; b = 0,98	$R^2 = 0,66$ a = 0,71 ; b = -18,2

TABLE 6.22: Comparaison entre les TEMP précoces et tardives en terme de volumes et de variations de volumes segmentés. Les mesures V_p^{TEMP} et V_t^{TEMP} ont été calculées sur 32 points, les mesures ΔV_{pb-a}^{TEMP} et ΔV_{tb-a}^{TEMP} ont été calculées sur 16 points.

Le tableau 6.22 montre, pour toutes les méthodes de segmentation, une forte corrélation entre les volumes précoces et tardifs ($R^2 \geq 0{,}75$). De plus, les volumes précoces et tardifs sont de taille sensiblement similaire pour les 3 premières segmentations et Seg5 (|biais| $< 0{,}58$ ml et a $> 0{,}85$). Pour la segmentation Seg4 les volumes précoces et tardifs sont corrélés mais de taille un peu différente. Pour cette segmentation les volumes tardifs sont généralement un peu plus importants que les volumes précoces (a = 1,20). Le tableau 6.22 montre également une corrélation importante entre les variations relatives de volumes précoces et tardifs. Cette corrélation est moindre pour la segmentation Seg2 ($R^2 = 0,39$).

Comme pour les volumes et variations de volumes, nous avons étudié les corrélations entre intensité moyenne et variations d'intensité moyenne calculées sur les volumes détectés sur les TEMP précoces et tardives (tableau 6.24).

	V_p^{TEMP} et V_t^{TEMP}
Seg1	d = -0,14 $b_{inf} = -5,20$; $b_{sup} = 4,92$
Seg2	d = -0,12 $b_{inf} = -5,53$; $b_{sup} = 5,29$
Seg3	d = -0,58 $b_{inf} = -5,58$; $b_{sup} = 4,42$
Seg4	d = 1,99 $b_{inf} = -10,2$; $b_{sup} = 14,1$
Seg5	d = -0,26 $b_{inf} = -12,5$; $b_{sup} = 11,9$

TABLE 6.23: Analyse de Bland Altman entre les TEMP précoces et tardives en terme de volumes et de variations de volumes segmentés. Les résultats sont donnés en cm^3 ou en ml. d correspond à la moyenne des différence entre volumes, b_{sup} à la moyenne des différences plus 2 fois l'écart type et b_{inf} à la moyenne des différences moins 2 fois l'écart type

	I_p^{TEMP} et I_t^{TEMP}	ΔI_{pb-a}^{TEMP} et ΔI_{tb-a}^{TEMP}
Seg1	$R^2 = 0,73$ a = 0,91 ; b = 2,03	$R^2 = 0,35$ a = 0,66 ; b = -10,0
Seg2	$R^2 = 0,74$ a = 0,91 ; b = 2,11	$R^2 = 0,33$ a = 0,64 ; b = -10,8
Seg3	$R^2 = 0,79$ a = 0,89 ; b =2,34	$R^2 = 0,32$ a = 0,58 ; b = -8,44
Seg4	$R^2 = 0,80$ a = 1,20 ; b = -1,45	$R^2 = 0,39$ a = 0,79 ; b = -4,58
Seg5	$R^2 = 0,84$ a = 0,95 ; b = 0,69	$R^2 = 0,73$ a = 1,03 ; b = 0,76

TABLE 6.24: Comparaison entre les TEMP précoces et tardives en termes d'intensité moyenne et de variation d'intensité moyenne à l'intérieur des volumes segmentés. Les mesures I_p^{TEMP} et I_t^{TEMP} ont été calculées sur 32 points et les mesures ΔI_{pb-a}^{TEMP} et ΔI_{tb-a}^{TEMP} ont été calculées sur 16 points.

Le tableau 6.24 montre que, pour les 5 segmentations, les intensités moyennes mesurées dans les volumes tumoraux segmentés sur les images TEMP entre les examens précoce et tardif sont fortement corrélées ($R^2 \geq 0,73$). Les variations relatives d'intensité moyenne montrent une corrélation nettement moins importante ($R^2 \leq 0,39$) sauf pour Seg5 ($R^2 = 0,73$) laissant supposer l'existence d'une différence d'information entre les images TEMP précoces et tardives.

Nous avons également étudié les corrélations entre les intensités maximales normalisées au sein de la tumeur entre les images TEMP précoces $Imax_p^{TEMP}$ et tardives $Imax_t^{TEMP}$ (tableau 6.25). Nous avons procédé de même pour les variations relatives de cette intensité maximale entre les images précoces $\Delta Imax_{pb-a}^{TEMP}$ et tardives $\Delta Imax_{tb-a}^{TEMP}$.

	$Imax_p^{TEMP}$ et $Imax_t^{TEMP}$	$\Delta Imax_{pb-a}^{TEMP}$ et $\Delta Imax_{tb-a}^{TEMP}$
Seg	$R^2 = 0,73$ a = 0,97 ; b = 2,06	$R^2 = 0,37$ a = 0,72 ; b = -7,47

TABLE 6.25: Comparaison entre les TEMP précoces et tardives en terme d'intensité maximale et de variations d'intensité maximale. Les mesures $Imax_p^{TEMP}$ et $Imax_t^{TEMP}$ ont été calculées sur 32 points et les mesures $\Delta Imax_{pb-a}^{TEMP}$ et $\Delta Imax_{tb-a}^{TEMP}$ ont été calculées sur 16 points.

Le tableau 6.25 montre des résultats très similaires à ceux obtenus pour l'intensité moyenne, à savoir des intensités maximales normalisées très corrélées ($R^2 = 0,73$) et une corrélation moins importante des variations relatives d'intensité maximales normalisées au cours du traitement entre les TEMP précoces et tardives.

Ces résultats s'expliquent par la corrélation importante entre les intensités maximales et les intensités moyennes normalisées (cf tableau 6.26). Ce résultat se retrouve pour la plupart des segmentations. Les exceptions concernent les variations d'intensité sur les images TEMP précoces pour Seg3 et Seg5 et les variations d'intensité sur les images TEMP tardives pour Seg3.

	$Imax_p^{TEMP}$ et I_p^{TEMP}	$\Delta Imax_{pb-a}^{TEMP}$ et ΔI_{pb-a}^{TEMP}	$Imax_t^{TEMP}$ et I_t^{TEMP}	$\Delta Imax_{tb-a}^{TEMP}$ et ΔI_{tb-a}^{TEMP}
Seg1	$R^2 = 0,99$ a = 0,57 ; b = -0,39	$R^2 = 0,98$ a = 1,00 ; b = 1,60	$R^2 = 0,99$ a = 0,54 ; b = 0,55	$R^2 = 0,99$ a = 0,94 ; b = 1,22
Seg2	$R^2 = 0,99$ a = 0,57 ; b = -0,41	$R^2 = 0,94$ a = 0,97 ; b = 0,40	$R^2 = 0,99$ a = 0,53 ; b = 0,76	$R^2 = 0,97$ a = 0,93 ; b = -1,66
Seg3	$R^2 = 0,86$ a = 0,58 ; b = 0,55	$R^2 = 0,50$ a = 0,92 ; b = 11,2	$R^2 = 0,91$ a= 0,53 ; b = 1,92	$R^2 = 0,55$ a = 0,83 ; b = 6,99
Seg4	$R^2 = 0,81$ a = 0,38 ; b = 2,50	$R^2 = 0,83$ a = 0,52 ; b = 0,31	$R^2 = 0,90$ a = 0,47 ; b = 0,58	$R^2 = 0,84$ a = 0,55 ; b = -0,53
Seg5	$R^2 = 0,90$ a = 0,21 ; b = 5,11	$R^2 = 0,52$ a = 0,57 ; b = -0,84	$R^2 = 0,91$ a = 0,19 ; b = 5,57	$R^2 = 0,70$ a = 0,67 ; b = -1,08

TABLE 6.26: Corrélation entre intensité moyenne I^{TEMP} et intensité maximale $Imax^{TEMP}$. Les mesures $Imax_p^{TEMP}$ et I_p^{TEMP}, $Imax_t^{TEMP}$ et I_t^{TEMP} ont été calculées sur 32 points et les mesures $\Delta Imax_{pb-a}^{TEMP}$ et ΔI_{pb-a}^{TEMP}, $\Delta Imax_{tb-a}^{TEMP}$ et ΔI_{tb-a}^{TEMP} sur 16 points.

6.4.3.2 Etude locale

Ce paragraphe a pour objectif de déterminer si les volumes TEMP précoces et tardifs se trouvent aux mêmes positions. Les résultats de l'étude locale comparant les volumes TEMP précoces et tardifs donnant les indices $Dice_{p-t}^{TEMP}$, $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ sont indiqués dans les tableaux 6.27 et 6.28.

	$Dice_{p-t}^{TEMP}$	$C1_{p-t}^{TEMP}$	$C5_{p-t}^{TEMP}$
Seg1	min = 0 max = 0,75 moyenne = 0,47 écart-type =0,19	min = 0 max = 1,00 moyenne = 0,19 écart-type = 0,22	min = 0 max = 1,45 moyenne = 0,28 écart-type = 0,36
Seg2	min = 0,05 max = 0,75 moyenne = 0,47 écart-type =0,19	min = 0 max = 0,67 moyenne = 0,17 écart-type = 0,16	min = 0 max = 2,52 moyenne = 0,30 écart-type = 0,47
Seg3	min = 0 max = 0,83 moyenne = 0,42 écart-type =0,23	min = 0 max = 1,00 moyenne = 0,23 écart-type = 0,29	min = 0 max = 8,75 moyenne = 0,67 écart-type = 1,64
Seg4	min = 0,05 max = 0,75 moyenne = 0,50 écart-type =0,20	min = 0 max = 0,67 moyenne = 0,17 écart-type = 0,15	min = 0 max = 2,52 moyenne = 0,25 écart-type = 0,47
Seg5	min = 0 max = 0,82 moyenne = 0,51 écart-type =0,24	min = 0 max = 0,98 moyenne = 0,16 écart-type = 0,19	min = 0,01 max = 1,69 moyenne = 0,31 écart-type = 0,48

TABLE 6.27: Comparaison locale des volumes tumoraux détectés sur les TEMP précoces et tardives.

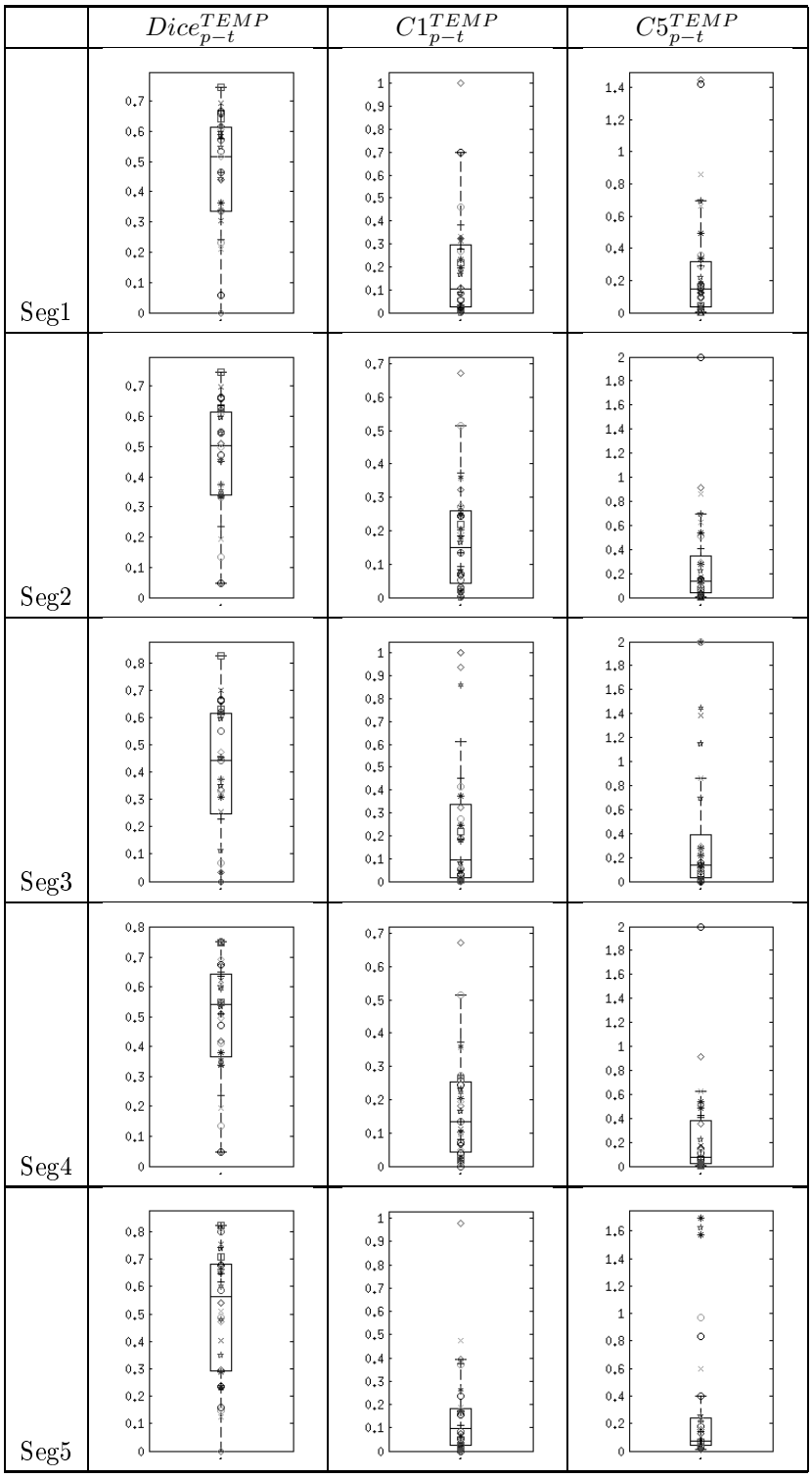


TABLE 6.28: Comparaison locale des volumes tumoraux détectés sur les TEMP précoces et tardives. Les indices $Dice_{p-t}^{TEMP}$ $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ ont été calculés pour 32 points.

D'après le tableau 6.28, les indices $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ ont des valeurs importantes ($>$ 0,15) dans plus de la moitié des cas pour toutes les segmentations.

Pour Seg1, dans 8 cas (P01 après traitement, P03 avant et après traitement, P05 après traitement, P07 après traitement, P11 après traitement, P12 après traitement et P13 après traitement), $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ ont des valeurs importantes ($>$ 0,15) , dans 6 cas (P01 avant traitement, P05 avant traitement, P10 après traitement, P14 avant traitement, P15 avant traitement et P16 avant traitement) $C1_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15), dans 8 cas (P04 avant et après traitement, P07 avant traitement, P08 avant et après traitement, P09 avant et après traitement, P11 avant traitement) $C5_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15).

Pour Seg2, dans 8 cas (P01 après traitement, P03 avant et après traitement, P05 après traitement, P07 après traitement, P11 après traitement, P12 après traitement et P13 après traitement), $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ ont des valeurs importantes ($>$ 0,15) , dans 8 cas (P01 avant traitement, P05 avant traitement, P08 après traitement, P10 après traitement, P14 avant traitement, P15 avant traitement et P16 avant et après traitement) $C1_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15), dans 5 cas (P04 avant traitement, P08 avant traitement, P09 avant et après traitement, P11 avant traitement) $C5_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15).

Pour Seg3, dans 4 cas (P01 après traitement, P03 avant et après traitement, P16 après traitement), $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ ont des valeurs importantes ($>$ 0,15) , dans 10 cas (P01 avant traitement, P02 après traitement, P05 avant et après traitement, P06 après traitement, P07 après traitement, P14 avant traitement, P15 avant et après traitement et P16 avant traitement) $C1_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15), dans 8 cas (P04 avant traitement, P08 avant et après traitement, P09 après traitement, P10 après traitement, P11 avant traitement, P12 après traitement) $C5_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15).

Pour Seg4, dans 8 cas (P01 après traitement, P03 après traitement, P05 avant et après traitement, P07 après traitement, P11 après traitement, P12 après traitement et P13 après traitement), $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ ont des valeurs importantes ($>$ 0,15) , dans 7 cas (P01 avant traitement, P08 après traitement, P10 avant t après traitement, P14 avant traitement, P16 avant et après traitement) $C1_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15), dans 4 cas (P03 avant traitement, P06 avant traitement, P09 après traitement, P15 avant traitement) $C5_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15).

Pour Seg5, dans 9 cas (P01 après traitement, P03 après traitement, P04 après traitement, P05 avant traitement, P07 après traitement, P08 après traitement, P11 après traitement, P12 après traitement et P13 après traitement), $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ ont des valeurs importantes ($>$ 0,15) , dans 6 cas (P01 avant traitement, P05 après traitement, P11 avant traitement, P15 après traitement, et P16 avant et après traitement) $C1_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15), dans 4 cas (P03 avant traitement, P04 avant traitement, P09 avant et après traitement) $C5_{p-t}^{TEMP}$ a des valeurs importantes ($>$ 0,15).

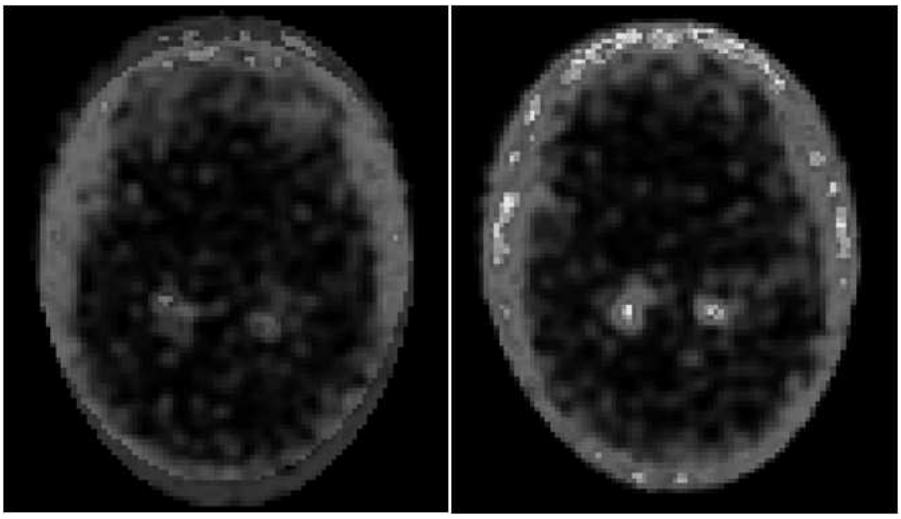
Ainsi les volumes tumoraux segmentés sur les TEMP précoces et tardives ne se trouvent pas exactement aux mêmes positions et dans certains cas l'un est inclus dans l'autre. De façon complémentaire, nous avons également calculé la distance moyenne entre la position de la valeur maximale sur les images TEMP précoces et tardives. Sur l'ensemble de ces jeux de données (cf tableau 6.29), la valeur moyenne de cette distance est de 14,3 $\pm$ 15,9 mm (6 $\pm$ 7 voxels) ce qui est un argument supplémentaire pour indiquer que les volumes tumoraux ne se trouvent pas exactement aux mêmes positions sur les images TEMP précoces et tardives. Il

est bon de rappeler que ces distances ne sont pas dues à un mauvais recalage comme le montre la figure 6.10 qui représente le recalage correct obtenu pour les données précoces et tardives après traitement du patient P01. En effet après recalage, les deux volumes paraissent correctement alignés et ne peuvent au maximum être décalés que d'un voxel. Un décalage de 2 ou de plus de 2 voxels ne peut donc pas être du à un mauvais recalage.

P01 bt	P01 at	P02 bt	P02 at	P03 bt	P03 at	P04 bt	P04 at
78,3	21,3	6,6	3,3	24,5	27,4	3,3	5,2
P05 bt	P05 at	P06 bt	P06 at	P07 bt	P07 at	P08 bt	P08 at
41,4	20,4	5,2	7,4	5,7	2,3	9,6	8,4
P09 bt	P09 at	P10 bt	P10 at	P11 bt	P11 at	P12 bt	P12 at
20,5	7,0	3,3	6,6	31,9	3,3	22,0	30,7
P13 bt	P13 at	P14 bt	P14 at	P15 bt	P15 at	P16 bt	P16 at
0,0	9,6	17,9	3,3	0,0	4,0	4,0	22,0

TABLE 6.29: Distance en mm entre la position du maximum trouvé sur l'examen TEMP précoce et celui sur l'examen TEMP tardif, avec "bt" pour les examens avant traitement et "at" pour les examens après traitement.

Le tableau 6.30 indique les résultats de la comparaison locale en intensité des zones tumorales détectées sur les TEMP précoces et tardives.



(a) Fusion avant recalage (b) Fusion après recalage

FIGURE 6.10: Fusion des examens TEMP précoces et tardifs avant traitement du patient P01 avant recalage (a) et après recalage (b).

	$L1_{p-t}^{TEMP}$	$L3a_{p-t}^{TEMP}$	$L5_{p-t}^{TEMP}$
Seg2			
Seg3			
Seg4			
Seg5			

TABLE 6.30: Comparaison locale des intensités dans les volumes segmentés entre les images TEMP précoces et les images TEMP tardives

D'après le tableau 6.30, les valeurs de $L1_{p-t}^{TEMP}$ et $L5_{p-t}^{TEMP}$ sont importantes (>0,2) dans plus de la moitié des cas pour chaque segmentation. Ceci prouve que les voxels des classes $C1_{p-t}^{TEMP}$ et $C5_{p-t}^{TEMP}$ ne peuvent pas être négligés. Même si ces valeurs varient avec les segmentations, la remarque reste valable quelle que soit la méthode de segmentation choisie. Ainsi les volumes tumoraux détectés sur les TEMP précoces et tardives ne se trouvent pas exactement aux mêmes positions comme le montre la figure 6.11 sur laquelle les volumes tumoraux détectés sur la TEMP précoce (en rouge) et sur la TEMP tardive (en bleu) pour le patient P01 ne se superposent pas.

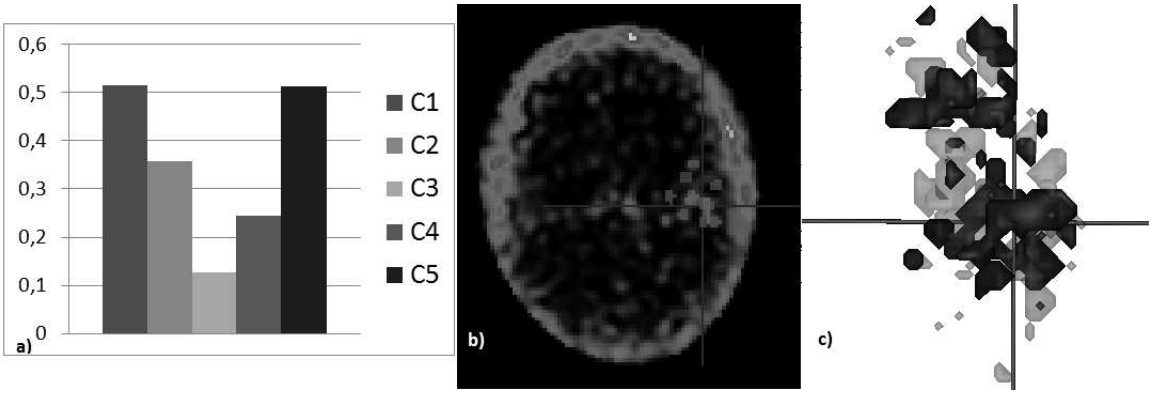


FIGURE 6.11: Résultats issus de la comparaison du volume tumoral détecté sur la TEMP précoce (en rouge) et du volume tumoral détecté sur la TEMP tardive (en bleu) (avec la méthode Seg4) pour le patient P03 qui suggère que les volumes tumoraux issus de ces 2 types d'images ne se trouvent pas aux mêmes endroits. Sur les images de coupe b) et de représentation 3D c), le volume rouge représentant la zone tumorale détectée sur la TEMP précoce est éloigné du volume bleu représentant la zone tumorale détectée sur la TEMP tardive avec la segmentation Seg4. A gauche a) sont affichées les valeurs de l'indice de comparaison de volume qui montrent clairement que les deux volumes ne se superposent pas.

Cette étude locale confirme donc que **les images TEMP précoces et tardives n'apportent pas exactement la même information**.

6.4.4 Comparaison entre les informations extraites des TEMP avec les indices PFS et OS

6.4.4.1 Etude globale

Ce paragraphe donne les résultats de comparaison des indices de volumes tumoraux et d'intensité moyenne à l'intérieur de ces volumes avec les durées de survie PFS et OS pour déterminer si les informations des données TEMP peuvent être pertinentes pour le suivi de patients.

Le tableau 6.31 regroupe les scores de corrélation entre les variations des volumes tumoraux extraits des TEMP précoces et tardives, et les indices PFS et OS.

	ΔV_{pb-a}^{TEMP} et PFS	ΔV_{pb-a}^{TEMP} et OS	ΔV_{tb-a}^{TEMP} et PFS	ΔV_{tb-a}^{TEMP} et OS
Seg1	NS	NS	NS	NS
Seg2	NS	NS	NS	NS
Seg3	NS	NS	NS	$R^2 = 0,23 \| p = 0,08 -> NS$
Seg4	NS	NS	NS	NS
Seg5	NS	NS	NS	$R^2 = 0,37 \| p = 0,02$ a = -0,11 ; b = -34,6

TABLE 6.31: Etude des corrélations entre les informations extraites des études TEMP précoces et tardives et les indices PFS et OS. Toutes ces mesures sont calculées sur 14 points car les valeurs des indices PFS et OS ne sont pas disponibles pour les patients P13 et P15. NS signifie "non significatif".

Le tableau 6.31 montre qu'il n'existe pas en général de corrélation entre la variation du volume tumoral détecté sur les TEMP précoces ou tardives avec les indices PFS ou OS. Néanmoins, il existe une corrélation entre les variations de volumes sur les images TEMP tardives pour la segmentation Seg5 et la durée de survie globale OS comme le montre la figure 6.12.

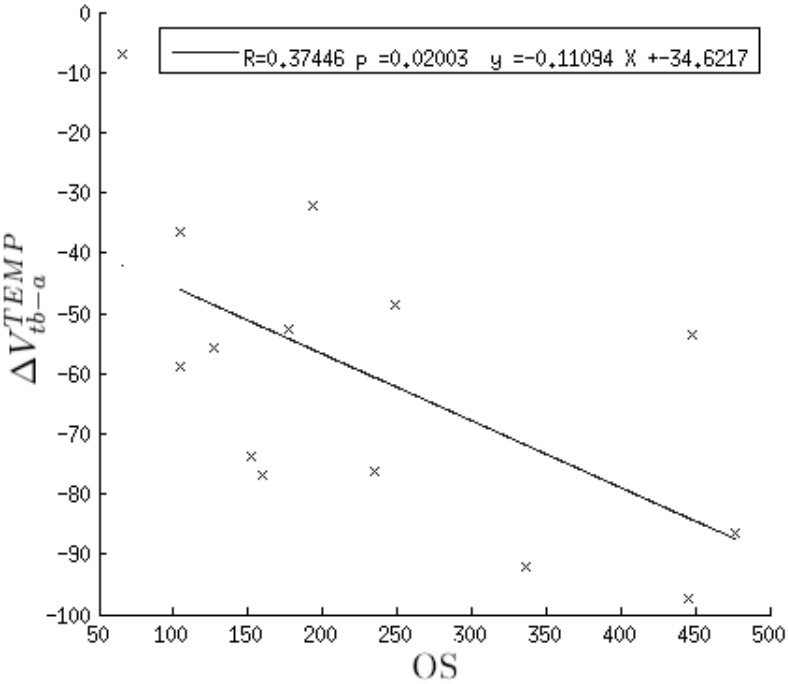


FIGURE 6.12: Corrélation entre ΔV_{tb-a}^{TEMP} et OS pour la segmentation Seg5.

	ΔI_{pb-a}^{TEMP} et PFS	ΔI_{pb-a}^{TEMP} et OS	ΔI_{tb-a}^{TEMP} et PFS	ΔI_{tb-a}^{TEMP} et OS
Seg1	NS	NS	NS	$R^2 = 0,47$ a = -0,13 ; b = 0,13
Seg2	NS	NS	NS	$R^2 = 0,46$ a = -0,13 ; b = -0,05
Seg3	NS	NS	NS	$R^2 = 0,48$ a = -0,13 ; b = 14,2
Seg4	NS	$R^2 = 0,27$ | p=0,06 -> NS	NS	$R^2 = 0,56$ a = -0,09 ; b = 3,69
Seg5	NS	NS	NS	$R^2 = 0,32 \| p = 0,04$ a = -0,09 ; b = -2,59

TABLE 6.32: Corrélation entre les informations liées aux intensités moyennes des volumes extraits des TEMP et les indices PFS et OS. Toutes ces mesures ont été calculées sur 14 points car les valeurs des indices PFS et OS ne sont pas disponibles pour les patients P13 et P15. NS signifie "non significatif".

Le tableau 6.32 regroupe les scores de corrélation entre les variations d'intensité dans les zones tumorales extraites des TEMP précoces et tardives, et les indices PFS et OS.

Le tableau 6.32 montre que les variations relatives d'intensité sur les TEMP précoces et tardives ne sont pas corrélées à la durée de survie sans progression PFS. De même, il n'existe pas de corrélation entre les variations relatives d'intensité détectées sur les TEMP précoces (même si une tendance se dessine à la limite de la signicativité pour la segmentation Seg4). Par contre, les variations relatives d'intensité dans les zones tumorales détectées sur les TEMP tardives sont systématiquement corrélées avec l'indice de survie global OS quelle que soit la méthode de segmentation considérée. Le résultat de cette corrélation pour la segmentation Seg4 est visible sur la figure 6.15a.

Nous avons également effectué une analyse de Kaplan-Meier en utilisant la médiane des valeurs ΔI_{tb-a}^{TEMP} comme valeur seuil pour définir les classes "R" comme réactif au traitement et "NR" comme non réactif au traitement représentée sur la figure 6.13. Le résultat des tests logarithmiques par rangs suite à cette analyse de Kaplan Meier donne une valeur de p = 0,0078 et est donc significatif.

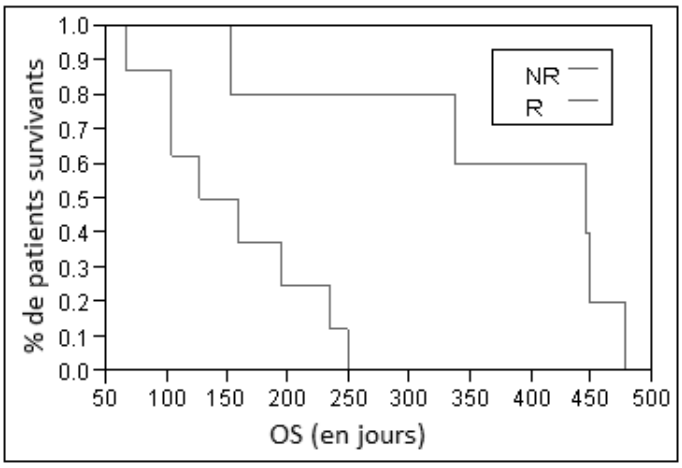


FIGURE 6.13: Analyse de Kaplan-Meier en fonction de OS et des classes "R" et "NR" définies à partir des valeurs de ΔI_{tb-a}^{TEMP} obtenues par la segmentation Seg4. Cette analyse a été effectuée sur 14 points (p = 0,0078).

Comme l'intensité maximale est très corrélée à l'intensité moyenne (cf tableau 6.26) nous avons également étudié les corrélations entre les variations de l'intensité maximale dans les zones tumorales extraites des TEMP précoces et tardives, et les indices de survie PFS et OS. Le tableau 6.33 regroupe ces scores.

	$\Delta Imax_{pb-a}^{TEMP}$ et PFS	$\Delta Imax_{pb-a}^{TEMP}$ et OS	$\Delta Imax_{tb-a}^{TEMP}$ et PFS	$\Delta Imax_{tb-a}^{TEMP}$
Seg	NS	$R^2 = 0,30 \| p = 0,044$ a = -0,08 ; b = -20,1	NS	$R^2 = 0,47 \| p = 0,007$ a = -0,13 ; b = 1,26

TABLE 6.33: Corrélation entre $\Delta Imax_{pb-a}^{TEMP}$,$\Delta Imax_{tb-a}^{TEMP}$, et PFS et OS. Toutes ces mesures sont calculées sur 14 points car les valeurs des indices PFS et OS ne sont pas disponibles pour les patients P13 et P15.

Nous aboutissons aux mêmes résultats que ceux obtenus avec les variations d'intensité moyenne. Il existe une corrélation entre les variations relatives d'intensité maximale dans les zones tumorales détectées sur les images TEMP précoces et tardives avec la durée de survie globale OS comme le montre la figure 6.15b. La corrélation est faible pour les temps précoces et plus importante pour les temps tardifs.

Nous avons également effectué une analyse de Kaplan-Meier en utilisant la médiane des valeurs $\Delta Imax_{tb-a}^{TEMP}$ comme valeur seuil pour définir les classes "R" comme réactif au traitement et "NR" comme non réactif au traitement représentée sur la figure 6.14. Le résultat des tests logarithmiques par rangs suite à cette analyse de Kaplan Meier donne une valeur de p = 0,0182 et est donc significatif.
Par contre il n'existe pas de corrélation entre les variations relatives d'intensité maximale normalisée dans les zones tumorales détectées sur les images TEMP précoces et tardives avec la durée de survie sans aggravation PFS.

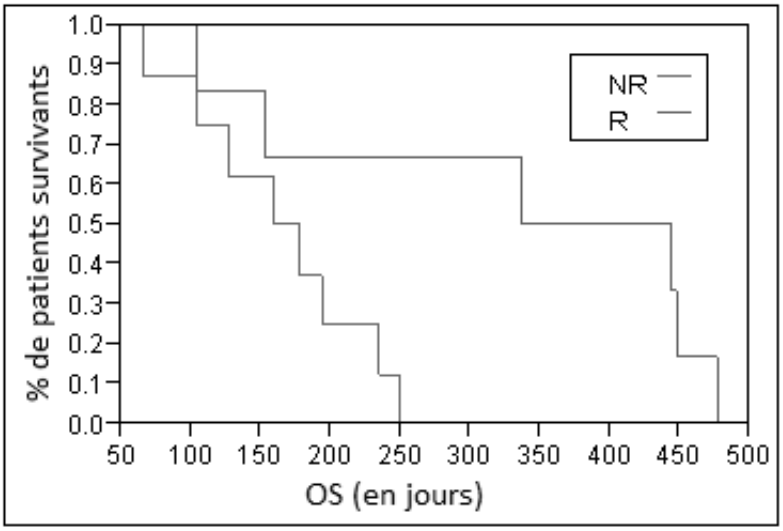
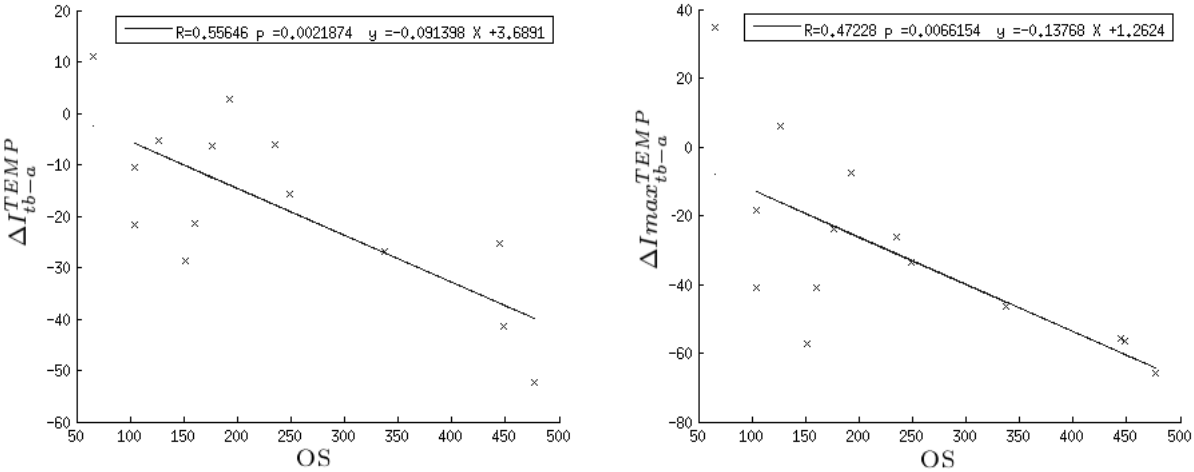


FIGURE 6.14: Analyse de Kaplan-Meier en fonction de OS et des classes "R" et "NR" définies à partir des valeurs de $\Delta Imax_{tb-a}^{TEMP}$ (obtenues par la segmentation Seg4). Cette analyse a été effectuée sur 14 points (p = 0,0182).



(a) Corrélation entre ΔI_{tb-a}^{TEMP} et OS

(b) Corrélation entre $\Delta Imax_{tb-a}^{TEMP}$ et OS

FIGURE 6.15: Corrélation entre ΔI_{tb-a}^{TEMP} avec OS (a) et $\Delta Imax_{tb-a}^{TEMP}$ avec OS (b)

Comme les variations relatives de la valeur maximale et les variations relatives d'intensité moyenne à l'intérieur des tumeurs aboutissent au même type de résultats, les variations relatives de la valeur maximale

pourraient être utilisées en priorité, car cela permettrait de s'affranchir de la segmentation de la zone tumorale sur les images TEMP.

6.4.4.2 Etude locale

Ce paragraphe donne les résultats de corrélation entre les indices locaux permettant de comparer les volumes avant et après traitement avec les indices de survie PFS et OS.

Nous avons vu que d'un point de vue local, les indices $C5^{TEMP}_{pb->a}$ et $C5^{TEMP}_{tb->a}$ variaient fortement en fonction des segmentations. En comparant les indices $C5^{TEMP}_{pb->a}$ et $C5^{TEMP}_{tb->a}$ avec les indices PFS et OS, nous allons pouvoir déterminer si ces indices se révèlent prédictifs de l'évolution tumorale ou pas.

Les corrélations entre $C5^{TEMP}_{pb-a}$, $C5^{TEMP}_{tb-a}$ avec PFS et OS ont été calculées sur 14 points car les valeurs de PFS et OS sont indisponibles pour les patients P13 et P15

Après analyse, les indices locaux de volumes $C5^{TEMP}_{pb-a}$, extraits des TEMP précoces, et $C5^{TEMP}_{tb-a}$, extraits des TEMP tardives, ne sont corrélés ni avec PFS ni avec OS, quelle que soit la méthode de segmentation choisie.

De même, les indices locaux d'intensité $L5^{TEMP}_{pb-a}$, extraits des TEMP précoces, et $L5^{TEMP}_{tb-a}$, extraits des TEMP tardives, ne sont corrélés ni avec PFS ni avec OS quelle que soit la méthode de segmentation choisie.

6.4.4.3 Conclusion

Dans ce paragraphe, nous avons cherché à savoir si des indices extraits des données TEMP avaient des valeurs prédictives de PFS et/ou OS. Après analyse nous avons montré qu'aucun des critères locaux ne possédait de corrélation avec les durées de survie PFS et OS. Aucun indice global n'est corrélé avec la durée de survie PFS. En revanche, les indices ΔI^{TEMP}_{tb-a} extraits des images TEMP tardives pour chaque segmentation et $\Delta Imax^{TEMP}_{tb-a}$ extraits des TEMP tardives sont fortement corrélés avec la durée de survie OS ($R^2 = 0,56$ et $R^2 = 0,47$). Etant donné que $\Delta Imax^{TEMP}_{tb-a}$ ne nécessite pas d'étapes de segmentation du volume tumoral sur les images TEMP, il est préférable de l'utiliser plutôt que ΔI^{TEMP}_{t-a}.

6.4.5 Discussion et conclusion relatives aux études TEMP

Le but principal de cette étude était de déterminer s'il existait un lien entre des indices globaux ou locaux extraits des images TEMP et la durée de survie sans aggravation PFS et la durée de survie globale OS dans l'objectif du suivi de patient atteints de gliomes de haut grade sous traitement antiangiogénique. Afin de calculer ces indices, cinq méthodes de segmentation ont été utilisées pour extraire les volumes tumoraux des images TEMP.

Afin de comparer les résultats de chacune de ces méthodes de segmentation, nous avons étudié les corrélations entre les indices globaux liés aux volumes tumoraux ainsi que ceux liés aux intensités au sein de ces volumes extraits à partir de chacune de ces méthodes. Bien que les valeurs des indices puissent fortement varier d'une méthode de segmentation à une autre, certaines tendances se dessinent. Les segmentations Seg1 et Seg2 apportent des résultats très similaires en terme de volumes, ce qui paraît logique étant donné que ces deux méthodes utilisent les mêmes seuils pour la segmentation, et que le seul élément changeant entre ces deux méthodes vient de la définition de la région d'intérêt sur laquelle est appliqué le seuil. Les segmentations Seg3 et Seg4, bien que moins corrélées dans la comparaison directe des volumes extraits par chacune de ces méthodes, montrent une corrélation importante des variations relatives de volumes ainsi que des variations relatives d'intensité. Ceci est lié au fait que, pour ces deux méthodes, le seuil appliqué aux données d'un patient avant traitement est le même que celui utilisé pour les données après traitement de ce même patient. En termes d'intensité moyenne, toutes les segmentations apportent des résultats beaucoup plus similaires ce qui fait que ce critère semble pertinent. Ces valeurs ont été comparées dans la suite avec les variations d'intensité maximale. La segmentation des images TEMP reste un problème ouvert et nous n'avons pas pu conclure quant à l'équivalence de certaines méthodes ni quant à la validation de la meilleure méthode à utiliser, ce problème n'étant pas facilement soluble en l'absence de gold standard. Nous avons donc gardé les résultats de toutes ces méthodes dans l'analyse des données TEMP. Pour valider la pertinence d'un critère dans la suite, il faut que les résultats suivant les différentes segmentations mènent à la même conclusion.

Dans le cadre de cette étude clinique, deux types d'images TEMP ont été proposés, les premières acquises juste après injection de Sestamibi (précoces) et les secondes 3 h après injection (tardives). Nous avons donc cherché à déterminer si ces deux types d'images apportaient la même information ou non. En terme, de volumes et de variations relatives de volume globaux, les images TEMP précoces et tardives apportent des informations globalement similaires. En revanche, après analyse locale, les volumes détectés sur les TEMP précoces et tardives ne se trouvent pas exactement aux mêmes localisations, suggérant une différence entre ces deux types d'images. Même si les intensités moyennes au sein des volumes tumoraux extaits sur les TEMP précoces et tardives sont fortement corrélées, les variations relatives d'intensité moyenne ainsi que les variations relatives d'intensité maximale sont différentes entre les TEMP précoces et tardives. L'indice local d'intensité confirme cette différence. De plus ce résultat est conforté par le fait que la position du voxel possédant l'intensité maximale n'est pas la même entre les études précoces et tardives. Les images TEMP précoces et tardives n'apportent donc pas exactement les mêmes informations. La comparaison avec les durées de survie PFS et OS a donc été faite d'une part avec les indices extraits des images TEMP précoces et

d'autre part avec ceux extraits des TEMP tardives.

Cette analyse montre que quelle que soit la segmentation choisie, aucun indice global ou local n'est corrélé à la durée de survie sans aggravation, PFS. En revanche, le résultat non attendu de cette étude est que les variations relatives d'intensité moyenne au sein des volumes détectés ainsi que les variations relatives d'intensité maximale sur les TEMP tardives sont fortement corrélées avec la durée de survie globale OS quelles que soient les segmentations. Les corrélations sont bien moins importantes en utilisant les images TEMP précoces même si une certaie tendance se dessine. Il paraît donc possible d'utiliser les variations relatives d'intensité maximale sur les TEMP tardives, puisque cet indice robuste ne nécessite que la définition d'une boîte englobante autour de la tumeur et permet d'éviter l'étape de segmentation. Ceci est conforté par les analyses de Kaplan Meier qui montrent que les variations relatives d'intensité moyenne ou maximale constituent des "biomarqueurs" prédicteurs de la durée de survie globale OS. En revanche, plusieurs éléments sont à prendre en compte. Premièrement, dans cette étude tous les patients ont été traités par antiangiogénique empêchant la comparaison avec un groupe "placebo" pour évaluer l'avantage du traitement antiangiogénique. De plus, la durée de survie globale OS correspond au temps de survie globale du patient depuis le début du traitement jusqu'à son décès or, pendant cette durée, le traitement antiangiogénique a pu être interrompu et/ou remplacé par un autre traitement. Donc le critère OS permet de juger de l'efficacité de la stratégie globale de traitement et non pas seulement de l'efficacité du traitement antiangiogénique. Enfin, bien que l'indice des variations relatives d'intensité maximale soit fortement corrélé à la durée de survie globale du patient, cette étude ne permet pas de conclure sur la pertinence de cet indice pour le suivi du patient dans le cadre du traitement antiangiogénique. Aussi, il serait intéressant d'appliquer cette approche à d'autres études en prenant en compte les point cités plus hauts.

6.5 Comparaisons globales et locales entre TEMP et IRM

Dans un premier temps nous avons calculé la corrélation directe entre les volumes tumoraux détectés en IRM et ceux détectés en TEMP suivant les 5 segmentations utilisées (tableau 6.34). En plus de la valeur du coefficient de corrélation, nous avons indiqué le biais "d" obtenu par une analyse de Bland Altman pour la comparaison des volumes (tableau 6.35).

	V^{IRM} et V_p^{TEMP}	V^{IRM} et V_t^{TEMP}
Seg1	$R^2 = 0,78$ a = 0,40 ; b = 0,97	$R^2 = 0,83$ a = 0,41 ; b = 0,98
Seg2	$R^2 = 0,75$ a = 0,38 ; b = 1,75	$R^2 = 0,79$ a = 0,39 ; b = 1,82
Seg3	$R^2 = 0,77$ a = 0,48 ; b = -1,05	$R^2 = 0,79$ a = 0,53 ; b = -1,12
Seg4	$R^2 = 0,60$ a = 0,90 ; b = 1,26	$R^2 = 0,42$ a = 0,42 ; b = 3,56
Seg5	$R^2 = 0,67$ a = 1,26 ; b = 1,82	$R^2 = 0,67$ a = 1,33 ; b = 1,86

TABLE 6.34: Comparaison des volumes tumoraux détectés sur les IRM et images TEMP suivant les 5 segmentations. Cette comparaison a porté sur 31 jeux de données IRM-TEMP, l'IRM après traitement étant manquante pour le patient 11.

Le tableau 6.35 montre les caractéristiques de l'analyse de Bland-Altman ("d" étant moyenne des différences, "b_{sup}" et "b_{inf}" les bornes supérieures et inférieurs correspondant à la moyenne plus ou moins deux fois l'écart type des différences) entre V^{IRM} et V_p^{TEMP} puis entre V^{IRM} et V_t^{TEMP}.

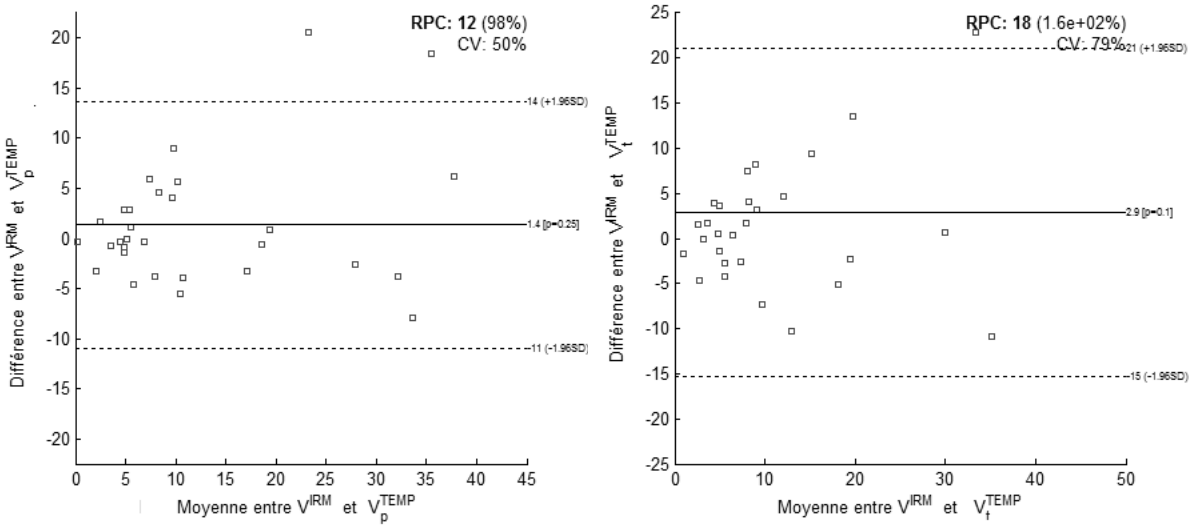
Le tableau 6.34 montre que les volumes détectés sur l'IRM et sur la TEMP précoce sont fortement corrélés pour les 5 segmentations ($R^2 \geq 0,60$). Les volumes détectés sur l'IRM et sur la TEMP tardive sont fortement corrélés pour les 4 segmentations ($R^2 \geq 0,67$) Seg1, Seg2, Seg3 et Seg5. Ils le sont moins pour Seg4 ($R^2 = 0,60$ et $R^2 = 0,42$). En revanche, pour sélectionner la meilleure segmentation en médecine nucléaire, il est classique de choisir celle qui fournit le volume le plus proche de celui observé sur une modalité anatomique [93]. Dans le cas de notre étude, il semble que ce soit Seg4 qui permette d'obtenir les volumes les plus proches de ceux mesurés en IRM ((a = 0,90 et b=1,26 cm^3) et le biais le plus faible (d $<$ 3 cm^3) et écart-type réduit). La figure 6.16 montre le graphique d'analyse de Bland-Altman entre V^{IRM} et V_p^{TEMP} (a) puis entre V^{IRM} et V_t^{TEMP} (b), V_p^{TEMP} et V_t^{TEMP} étant extraits à l'aide de Seg4.

Dans un second temps, les variations relatives des volumes tumoraux détectés sur les TEMP suivant les 5 segmentations ont été comparées aux variations relatives des volumes tumoraux détectés en IRM (tableau 6.36).

Les variations de volumes de la zone tumorale détectée sur l'IRM et sur les images TEMP sont également corrélées surtout pour la segmentation Seg4 sur les volumes TEMP précoces, un peu moins sur les TEMP

	V^{IRM} et V_p^{TEMP}	V^{IRM} et V_t^{TEMP}
Seg1	d = 6,74 $b_{inf} = -7,71$; $b_{sup} = 21,2$	d = 6,90 $b_{inf} = -7,94$; $b_{sup} = 21,7$
Seg2	d = 6,12 $b_{inf} = -8,68$; $b_{sup} = 20,9$	d = 6,32 $b_{inf} = -9,14$; $b_{sup} = 21,8$
Seg3	d = 7,69 $b_{inf} = -5,52$; $b_{sup} = 20,9$	d = 0,37 $b_{inf} = -18,6$; $b_{sup} = 19,3$
Seg4	d = 1,35 $b_{inf} = -11,0$; $b_{sup} = 13,7$	d = 2,89 $b_{inf} = -15,4$; $b_{sup} = 21,1$
Seg5	d = -5,30 $b_{inf} = -27,1$; $b_{sup} = 16,5$	d = -6,22 $b_{inf} = -29,5$; $b_{sup} = 17,1$

TABLE 6.35: Analyse de Bland Altman des volumes tumoraux détectés sur les IRM et images TEMP suivant les 5 segmentations. Toutes les valeurs sont données en cm^3 ou ml.



(a) Analyse de Bland-Altman entre V^{IRM} et V_p^{TEMP} (b) Analyse de Bland-Altman entre V^{IRM} et V_p^{TEMP}

FIGURE 6.16: Analyse de Bland-Altman entre V^{IRM} et V_p^{TEMP} (a) puis entre V^{IRM} et V_t^{TEMP} (b) avec V_p^{TEMP} et V_t^{TEMP} extraits à l'aide de Seg4

tardives. Ainsi il semble que les données TEMP précoces donnent des résultats plus proches de l'IRM que ceux obtenus sur les TEMP tardives.

	ΔV_{b-a}^{IRM} et ΔV_{pb-a}^{TEMP}	ΔV_{b-a}^{IRM} et ΔV_{tb-a}^{TEMP}
Seg1	$R^2 = 0,44 \| p = 0,007$ a = 1,02 ; b = 22	$R^2 = 0,56$ a = 0,88 ; b = 4,48
Seg2	$R^2 = 0,50$ a = 0,98 ; b = 31,3	$R^2 = 0,20$ | p = 0,10 -> NS a = 0,57 ; b = 7,49
Seg3	$R^2 = 0,66$ a = 1,10 ; b = -14,7	$R^2 = 0,55$ a = 1,28 ; b = -2,37
Seg4	$R^2 = 0,78$ a = 1,31 ; b = 14,0	$R^2 = 0,61$ a = 1,85 ; b = 49,5
Seg5	$R^2 = 0,59$ a =0,62 ; b = -24,0	$R^2 = 0,50$ a = 0,50 ; b = -31,8

TABLE 6.36: Comparaison des variations relatives de volumes de la tumeur sur les images TEMP précoces et tardives suivant les 5 segmentations avec les variations de volumes de la tumeur détectées sur l'IRM. Ces mesures ont été calculées pour 15 points.

6.5.1 Etude locale

Comme nous l'avons vu au paragraphe précédent, les volumes tumoraux détectés sur les IRM après injection de Gadolinium et ceux segmentés sur les images TEMP précoces à l'aide de la méthode de segmentation Seg 4 sont corrélés et de taille assez proche. Afin de savoir si ces volumes se trouvent aux mêmes endroits, nous avons utilisé les indices locaux de volumes et d'intensité.

Les valeurs des indices locaux de volumes pour comparer les volumes tumoraux détectés sur les IRM et les TEMP précoces (Seg4) $Dice_p^{IRM-TEMP}$, $C1_p^{IRM-TEMP}$ et $C5_p^{IRM-TEMP}$ sont disponibles dans les tableaux 6.37 et 6.38.

	$Dice_p^{IRM-TEMP}$	$C1_p^{IRM-TEMP}$	$C5_p^{IRM-TEMP}$
Seg4	min = 0,02 max = 0,74 moyenne = 0,44 écart-type =0,18	min = 0,00 max = 0,76 moyenne = 0,24 écart-type = 0,20	min = 0,00 max = 7,62 moyenne = 0,42 écart-type = 1,35

TABLE 6.37: Indices locaux de volumes pour comparer V^{IRM} et V_p^{TEMP}. Les mesures $Dice_p^{IRM-TEMP}$, $C1_p^{IRM-TEMP}$ et $C5_p^{IRM-TEMP}$ ont été calculées sur 31 points car l'IRM après traitement est manquante pour le patient P11.

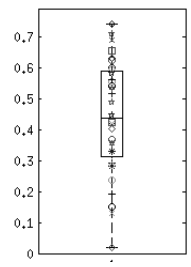
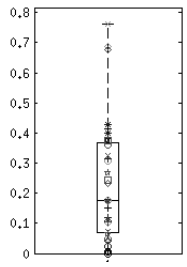
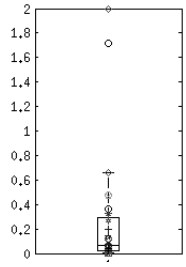
	$Dice_p^{IRM-TEMP}$	$C1_p^{IRM-TEMP}$	$C5_p^{IRM-TEMP}$
Seg4			

TABLE 6.38: Indices locaux de volumes pour comparer V^{IRM} et V_p^{TEMP}. Les mesures $Dice_p^{IRM-TEMP}$, $C1_p^{IRM-TEMP}$ et $C5_p^{IRM-TEMP}$ ont été calculées sur 31 points car l'IRM après traitement est manquante pour le patient P11.

Malgré une valeur moyenne de l'indice de Dice assez faible (m = 0,44) indiquant que les volumes ne se superposent pas parfaitement, les cas où $C1_p^{IRM-TEMP}$ et $C5_p^{IRM-TEMP}$ ont simultanément des valeurs assez fortes (supérieures à 0,2 indiquant des volumes éloignés l'un de l'autre) sont au nombre de 4 : les données après traitement du patient P01, les données avant et après traitement du patient P05 et les données après traitement du patient P07. La figure 6.17 montre la superposition des volumes tumoraux détectés sur la TEMP précoce et sur l'IRM avant traitement antiangiogénique pour le patient P05. Le volume détecté sur l'image TEMP précoce est inclus dans celui détecté sur l'IRM dans 10 cas ($C1_p^{IRM-TEMP} > 0{,}2$ et $C5_p^{IRM-TEMP} < 0{,}2$) et le volume tumoral détecté sur l'IRM est inclus dans celui détecté sur la TEMP précoce dans 5 cas ($C1_p^{IRM-TEMP} < 0{,}2$ et $C5_p^{IRM-TEMP} > 0{,}2$). Enfin, les deux volumes sont relativement bien superposés dans les 12 cas restants ($C1_p^{IRM-TEMP} < 0{,}2$ et $C5_p^{IRM-TEMP} < 0{,}2$).

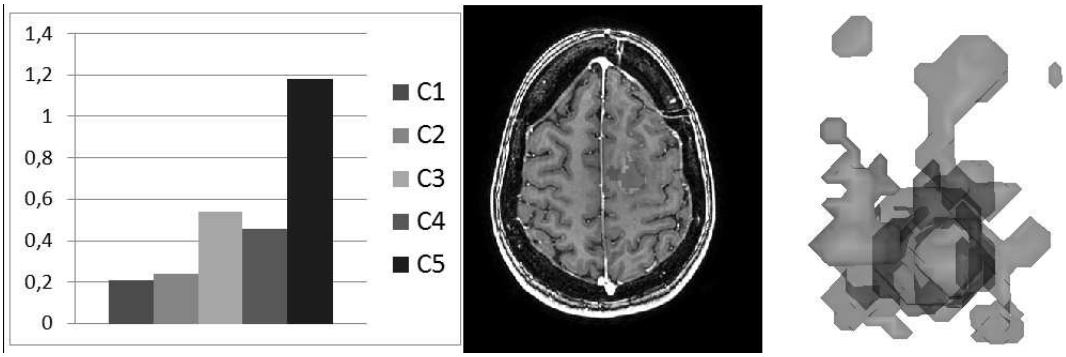


FIGURE 6.17: Comparaison entre le volume tumoral détecté sur l'IRM avant traitement (en rouge) et le volume détecté sur la TEMP précoce avant traitement (en bleu) à l'aide de la méthode de segmentation Seg 4 pour le patient P05.

Les valeurs des indices locaux d'intensité pour comparer les volumes tumoraux détectés sur les IRM et les TEMP précoces (Seg4) $L1_p^{IRM-TEMP}$ et$L5_p^{IRM-TEMP}$ sont disponibles dans le tableau 6.39.

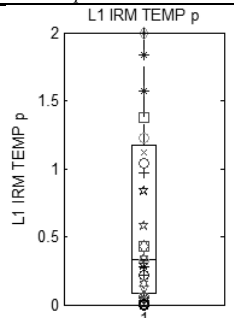
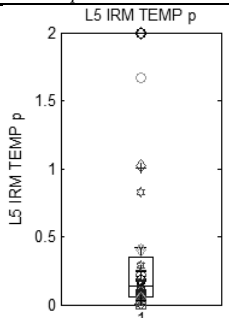
	$L1_p^{IRM-TEMP}$	$L5_p^{IRM-TEMP}$
Seg4		

TABLE 6.39: Indices locaux d'intensité $L1_p^{IRM-TEMP}$ et $L5_p^{IRM-TEMP}$ pour comparer les volumes tumoraux détectés sur les TEMP précoces et les IRM.

Le tableau 6.39 confirme que les données après traitement du patient P01, les données avant et après traitement du patient P05 et les données après traitement du patient P07, les valeurs de $L1_p^{IRM-TEMP}$ et $L5_p^{IRM-TEMP}$ produisent des valeurs élevées de $L1_p^{IRM-TEMP}$ ($> 0{,}4$) et de $L5_p^{IRM-TEMP}$ ($> 0{,}4$) ce qui confirme le fait que les volumes détectés sur l'IRM et sur la TEMP précoce sont éloignés pour ces jeux de données. Six cas supplémentaires présentent des valeurs de $L1_p^{IRM-TEMP}$ et $L5_p^{IRM-TEMP}$ élévées ($> 0{,}2$) impliquant que, malgré un nombre de voxels restreint en-dehors de la zone commune entre les deux volumes, les intensités dans ces voxels sont assez importantes, ce qui laisse supposer une forte activité de la tumeur dans ces zones. Dans les 22 autres cas, les volumes détectés sur les TEMP précoces sont relativement bien superposés ou inclus dans les volumes détectés sur les IRM.

Les valeurs des indices locaux de volumes pour comparer les volumes tumoraux détectés sur les IRM et les TEMP tardives (Seg4) $Dice_t^{IRM-TEMP}$, $C1_t^{IRM-TEMP}$ et $C5_t^{IRM-TEMP}$ sont disponibles dans les tableaux 6.40 et 6.41.

	$Dice_t^{IRM-TEMP}$	$C1_t^{IRM-TEMP}$	$C5_t^{IRM-TEMP}$
Seg4	min = 0,00	min = 0,00	min = 0,00
	max = 0,71	max = 0,76	max = 18,00
	moyenne = 0,40	moyenne = 0,27	moyenne = 1,07
	écart-type =0,18	écart-type = 0,23	écart-type = 3,74

TABLE 6.40: Indices locaux de volumes pour comparer V^{IRM} et V_t^{TEMP}

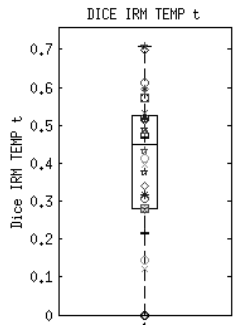
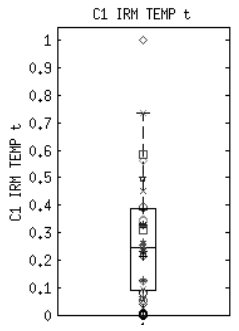
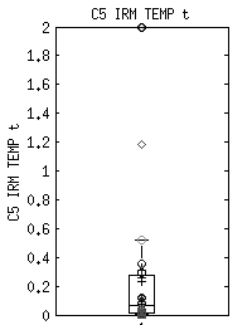
	$Dice_t^{IRM-TEMP}$	$C1_t^{IRM-TEMP}$	$C5_t^{IRM-TEMP}$
Seg4			

TABLE 6.41: Indices locaux de volumes pour comparer V^{IRM} et V_t^{TEMP}

Malgré une valeur moyenne de Dice assez faible (m = 0,40) indiquant que les volumes ne se superposent pas parfaitement, les cas où $C1_t^{IRM-TEMP}$ et $C5_t^{IRM-TEMP}$ ont simultanément des valeurs assez fortes (supérieures à 0,2 indiquant des volumes éloignés l'un de l'autre) sont au nombre de 5 en tout : les données après traitement de patient P01, les données avant et après traitement du patient P05, les données après

traitement du patient P07 et les données avant traitement du patient P15. La figure 6.18 montre que les volumes tumoraux détectés sur les IRM et images TEMP tardives ne se superposent pas parfaitement pour le patient P07. Le volume détecté sur l'image TEMP précoce est inclus dans celui détecté sur l'IRM dans 14 cas ($C1_t^{IRM-TEMP} > 0{,}2$ et $C5_t^{IRM-TEMP} < 0{,}2$) et le volume tumoral détecté sur l'IRM est inclus dans celui détecté sur la TEMP précoce dans 5 cas ($C1_t^{IRM-TEMP} < 0{,}2$ et $C5_t^{IRM-TEMP} > 0{,}2$). Enfin, les deux volumes sont relativement bien superposés dans les 7 cas restants ($C1_t^{IRM-TEMP} < 0{,}2$ et $C5_t^{IRM-TEMP} < 0{,}2$).

Les valeurs des indices locaux d'intensité pour comparer les volumes tumoraux détectés sur les IRM et les TEMP tardives (obtenus à l'aide de la méthode de segmentation Seg4) $L1_t^{IRM-TEMP}$ et $L5_t^{IRM-TEMP}$ sont reportées dans le tableau 6.42.

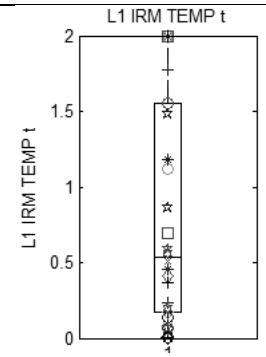
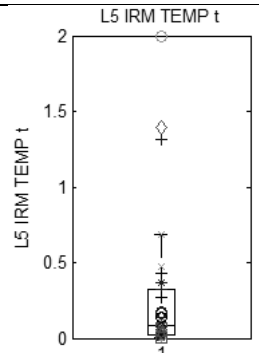
	$L1_t^{IRM-TEMP}$	$L5_t^{IRM-TEMP}$
Seg4		

TABLE 6.42: Indices locaux d'intensité pour comparer V^{IRM} et V_t^{TEMP}

Le tableau 6.42 montre que pour les données après traitement du patient P01, les données avant et après traitement du patient P05, les données après traitement du patient P07 et les données après traitement du patient P15, les valeurs de $L1_t^{IRM-TEMP}$ et $L5_t^{IRM-TEMP}$ sont élevées ($L1_t^{IRM-TEMP} > 0{,}<4$ et $L5_t^{IRM-TEMP} < 0{,}4$) ce qui confirme le fait que les volumes détectés sur l'IRM et la TEMP tardive soient éloignés pour ces jeux de données. A ces 5 cas, s'ajoutent 3 autres cas avec des valeurs de $L1_t^{IRM-TEMP} > 0{,}2$ et $L5_t^{IRM-TEMP} > 0{,}2$ élévées, impliquant que malgré un nombre de voxels plus faible (puisque $C1_t^{IRM-TEMP} < 0{,}2$ et $C5_t^{IRM-TEMP} < 0{,}2$) en-dehors de la zone commune entre les deux volumes, les intensités de ces voxels sont assez importantes, ce qui laisse supposer une forte activité de la tumeur dans ces zones (P03 après traitement, P07 avant traitement et P15 avant traitement). Pour les 23 autres cas, les volumes détectés sur les TEMP précoces sont relativement bien superposés ou inclus dans les volumes détectés sur les IRM.

D'après les critères locaux, deux tiers des volumes tumoraux détectés sur les TEMP (aussi bien précoces

que tardives) sont bien superposés ou inclus dans les volumes détectés sur l'IRM. Dans un tiers des cas, les volumes détectés sur l'IRM sont soit inclus soit éloignés des volumes détectés sur les images TEMP, ce qui représente une portion non négligeable. Il semble donc que magré des tailles de volumes globalement similaires, les volumes tumoraux détectés sur les IRM et les TEMP ne se trouvent pas nécessairement aux mêmes positions.

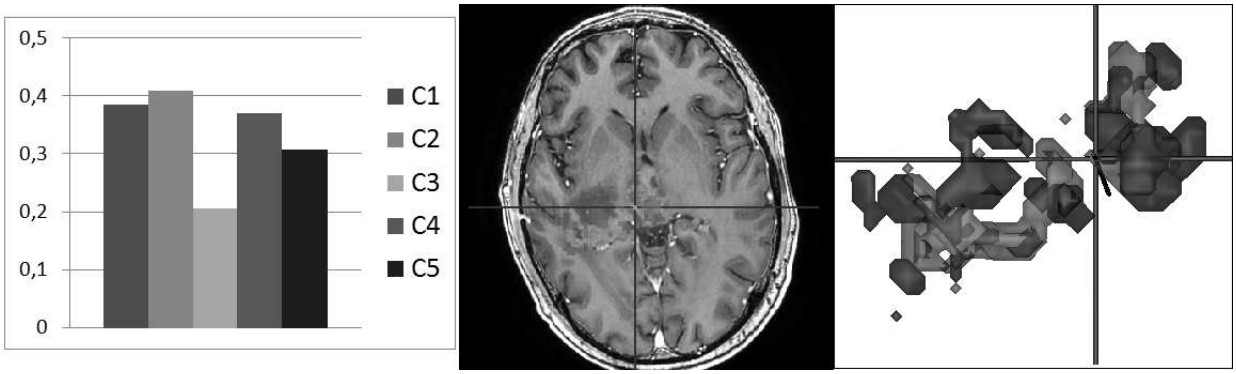


FIGURE 6.18: Comparaison entre le volume tumoral détecté sur l'IRM après traitement (en rouge) et le volume détecté sur la TEMP tardive après traitement (en bleu) avec la méthode de segmentation Seg4 pour le patient P07.

6.6 Conclusion

Dans ce chapitre, nous avons présenté les résultats de notre méthodologie d'analyse globale et locale des volumes tumoraux sur deux modalités ; l'IRM après injection de gadolinium et la TEMP avec injection de Sestamibi pour le suivi longitudinal de patients atteints de gliomes de haut grade. La partie 6.4 a montré les difficultés d'analyse des volumes tumoraux sur les images TEMP liées au choix relatif à la méthode de segmentation. Nous avons montré aux paragraphes 6.3 et 6.4 que ni les variations relatives de volumes en IRM ni celles observées en TEMP n'étaient liées à la durée de survie sans aggravation de l'état des patients PFS.

On peut supposer que les changements de la perméabilité de la barrière hématoencéphalique pendant le traitement peuvent expliquer ces résultats. [22]. En effet, en plus de son action d'inhibition de la croissance tumorale, le Bevacizumab entraîne des modifications durables de la fonction vasculaire et réduit la densité et la perméabilité des microvaisseaux [171]. Ces modifications vasculaires peuvent réduire la diffusion du traceur dans les cellules tumorales. En conséquence, la réponse à la thérapie antiangiogénique évaluée par l'examen TEMP marqué au Sestamibi ne devrait pas être interprétée comme traduisant uniquement l'efficacité anti-tumorale, mais aussi comme une normalisation sang-tumeur de la perméabilité de la barrière. Ce phénomène est une limitation bien connue de l'évaluation de l'évolution tumorale à l'aide d'examens avec injection de Gadolinium pour des patients traités par antiangiogéniques. Au cours des dernières années, l'adoption généralisée du Bevacizumab comme thérapie de sauvetage a mis en évidence le phénomène de "pseudo-réponse" [172] à opposer à une réelle régression de la tumeur. Cette limitation particulière des critères de Macdonald [27] a suscité la proposition de nouveaux critères d'évaluation de la réponse pour les gliomes de haut grade par le groupe de travail RANO, qui incluent désormais des séquences T2 / FLAIR [28]. Dans notre étude, les variations au cours du traitement antiangiogénique des volumes tumoraux déterminés à partir des TEMP Sestamibi et trouvés en IRM après injection de gadolinium sont fortement corrélés (cf partie 6.5). Ainsi les relations entre les variations relatives de volumes, qu'elles soient observées sur l'IRM ou sur la TEMP et le devenir des patients renforcent l'hypothèse que l'utilisation de ces seuls critères est insuffisante pour évaluer les gliomes de haut grade traités par antiangiogéniques. Après le recalage des données IRM/TEMP à l'aide de notre stratégie de recalage, nous avons néanmoins montré, à l'aide d'un indice local, pouvant prendre en compte une erreur d'alignement d'un voxel, que les volumes de Gadolinium et de Sestamibi liés à la captation tumorale ne se trouvaient pas nécessairement exactement aux mêmes endroits, ce qui suggère une possible complémentarité entre les deux examens.

De plus, nous avons montré que les variations de l'intensité maximale ou de l'intensité moyenne au sein des volumes tumoraux détectés sur les TEMP tardives suivant l'une des 5 méthodes de segmentation proposées étaient fortement liées à la durée de survie globale des patients. Cet aspect est nettement moindre lorsqu'on s'intéresse aux variations de l'intensité maximale au sein des volumes tumoraux détectés sur les TEMP précoces. Donc, les variations d'intensité maximale des TEMP normalisées tardives (dont la mise en œuvre ne nécessite que la définition d'une boîte englobante) paraît être un "biomarqueur" prédictif de la durée de survie globale OS et reste plus simple à utiliser que la valeur moyenne au sein des volumes tumoraux (qui requiert la segmentation de ces dits volumes, ce qui reste une étape complexe à valider).

Malgré l'identification de ce "biomarqueur" de OS, conclure sur la pertinence du traitement antiangio-

génique pour des patients atteints de gliomes de haut grade n'est pas chose aisée pour plusieurs raisons. Premièrement, dans cette étude tous les patients ont été traités par antiangiogénique, empêchant la comparaison avec un groupe placebo pour déduire l'avantage de ce type de traitement. De plus, la durée de survie globale OS correspond au temps de survie globale du patient depuis le début du traitement jusqu'à son décès or, pendant cette durée, le traitement antiangiogénique a pu être interrompu et/ou remplacé par un autre traitement. Donc OS permet de juger de l'efficacité de la stratégie globale de traitement et non pas seulement de l'efficacité du traitement antiangiogénique. Le faible nombre de patients inclus reste une limite de l'étude réalisée. Il serait intéressant de poursuivre cette étude pour d'autres types de traitement.

Chapitre 7

Mise en œuvre de la stratégie de recalage sur des données tomodensitométriques abdominales

Ce projet qui a été mené dans l'équipe 2 du LIF [173] constitue une preuve de concept pour démontrer la généricité et l'intéret de la stratégie générale de recalage définie au chapitre 3. La personne qui a développé le processus de recalage est Mike-Ely Cohen ; mon implication dans ce projet a permi d'utiliser la stratégie de recalage dans ce type d'étude en proposant des idées pour sa mise en œuvre.

Sommaire

7.1 Spécificités de cette application . . . 195

7.1.1 Contexte médical . . . 195

7.1.2 Base de données . . . 196

7.2 Mise en œuvre de la stratégie pour le recalage rigide longitudinal de TDM abdominales acquises pour le suivi de patients atteints de tumeurs hépatiques 196

7.2.1 Choix des différentes approches de recalage . . . 196

7.2.2 Evaluation quantitative et classement . . . 199

7.2.3 Validation visuelle . . . 199

7.3 Mise en œuvre de la stratégie pour le recalage d'études de TDM abdominales de perfusion au premier passage . . . 200

7.3.1 Choix des différentes approches de recalage . . . 200

7.3.2 Evaluation quantitative et classement . . . 203

7.3.3 Validation visuelle . . . 203

7.4 Résultats du recalage TDM-TDM pour le suivi longitudinal . . . 205

7.4.1 Performance de la stratégie . . . 205

7.4.2 Discussion . . . 210

7.4.3 Limites et perspectives 210
7.4.4 Conclusion 212
7.5 Résultats du recalage TDM-TDM sur les études dynamiques de perfusion . . 213
7.5.1 Performance de la stratégie 213
7.5.2 Discussion 215
7.5.3 Conclusion 215
7.6 Conclusion 216

7.1 Spécificités de cette application

7.1.1 Contexte médical

Depuis plusieurs années, différents traitements anticancéreux ont été développés dans le but d'améliorer la prise en charge du cancer. Dans ce contexte, de plus en plus de patients atteints de métastases hépatiques sont traités avec des traitements anti-angiogéniques [31]. Ces médicaments visent à réduire le développement des vaisseaux sanguins autour des tumeurs et ainsi à réduire voire stopper leur approvisionnement en oxygène et en nutriments. L'évaluation précoce de la réponse tumorale à un traitement anti-angiogénique est cruciale pour décider de poursuivre ou de changer de traitement compte-tenu du coût de ces thérapies et de leurs effets secondaires. Pour l'évaluation de l'évolution des tumeurs solides, les critères RECIST (Response Evaluation Criteria in Solid Tumors) [26] ont été proposés. Cependant, les critères RECIST sont purement morphologiques et ne tiennent pas compte de l'hétérogénéité tumorale et des changements physiologiques, tels que la modification de la vascularisation, qui peuvent survenir plus tôt que les changements de taille dans la réponse au traitement. Pour surmonter ces limitations, une adaptation des critères RECIST, appelée critères RECIST modifiés, a été conçue spécifiquement pour le carcinome hépatocellulaire, où la lésion cible n'est plus la globalité de la lésion tumorale, mais seulement la partie avec réhaussement de contraste de la lésion sur des données dynamiques de perfusion [32]. La tomodensitométrie (TDM) est largement utilisée pour évaluer la réponse des lésions hépatiques à un traitement, en raison de sa grande disponibilité dans un environnement clinique.

Pour une évaluation plus précise de la thérapie, les volumes d'examens TDM pré et post-traitement doivent être recalés avant d'évaluer la réponse tumorale. Il s'agit d'une étape nécessaire pour accéder à des déformations locales de la tumeur et pour comparer les paramètres fonctionnels comme la perfusion tissulaire. Pour tenir compte des déformations non rigides, l'utilisation de méthodes de recalage non rigides comme l'algorithme de Free Form Deformation (FFD) semble pertinente [33]. Cependant ces algorithmes non rigides nécessitent une initialisation de bonne qualité. Cette initialisation peut être apportée par une méthode de recalage rigide.

Le premier objectif de l'étude qui suit a été de trouver la meilleure initialisation possible par recalage rigide pour chaque jeu de données permettant un recalage des données longitudinales.

En raison de la durée de chaque examen de perfusion (2 à 5 min), les acquisitions ont été réalisées en respiration libre. Il est alors nécessaire de corriger les mouvements respiratoires sur chacun des examens

en vue d'améliorer la quantification. Le second problème de cette étude a concerné la compensation de ces mouvements respiratoires.

Pour résoudre ces deux problèmes, nous avons appliqué la stratégie de recalage exposée au chapitre 3. Ce chapitre décrit la mise en œuvre de cette stratégie en adaptant le choix des méthodes de recalage à utiliser et les critères d'évaluation aux spécificités de l'application.

7.1.2 Base de données

La base de données est constituée de dix-neuf paires de volumes TDM de patients atteints de métastases hépatiques. Ces patients faisaient partie d'un protocole de phase 2 dans lequel était impliqué le service de radiologie centrale de la Pitié-Salpétrière. Ce protocole clinique a été mis en place par le Professeur Olivier Lucidarme. Les acquisitions ont été réalisées soit avec un scanner GE VCT Pro 32 (S1), soit avec un scanner Philips iCT256 (S2). Les courants de tubes étaient de 100 mA pour S1 et 242 mA pour S2, et les temps d'exposition étaient de 1 s pour S1, 330 ms pour S2. Chaque volume était constitué de huit à vingt-deux coupes axiales avec une épaisseur de coupe égale soit à 2,5 mm, soit à 5 mm, et une taille de pixel variant entre 0,68x0,68 mm^2 et 0,80x0,80 mm^2. La présence d'artéfacts, comme un implant métallique, sur certains jeux de données a rendu le recalage difficile pour ces données. D'autre part le recalage longitudinal est délicat car il n'existe que peu de coupes communes sur les volumes acquis aux différents temps du suivi longitudinal.

7.2 Mise en œuvre de la stratégie pour le recalage rigide longitudinal de TDM abdominales acquises pour le suivi de patients atteints de tumeurs hépatiques

La stratégie de recalage consiste à utilise plusieurs méthodes de recalage et à sélectionner la meilleure méthode pour chaque jeu de données à l'aide d'un critère quantitatif. Le résultat obtenu par la méthode de recalage retenue sera alors validé ou non visuellement par un expert en dernière étape. Ce chapitre explique la mise en œuvre de cette stratégie dans le contexte du recalage rigide longitudinal de TDM abdominales acquises pour le suivi de patients atteints de tumeurs hépatiques.

7.2.1 Choix des différentes approches de recalage

Le but du recalage longitudinal est de pouvoir superposer les données TDM avant le début du traitement anti-angiogénique avec celles acquises après le début du traitement anti-angiogénique.

De plus, chaque séquence TDM comporte plusieurs volumes sur lesquels on suit la distribution du produit de contraste au cours du temps. Or cette distribution de produit de contraste peut être différente entre les volumes pré et post traitement du fait de l'évolution de la tumeur, des variation liées à l'injection du produit de contraste et de l'état général du patient. Pour le recalage longitudinal, il a donc été décidé de choisir le premier volume sur l'examen dynamique de perfusion pré-traitement comme volume de référence, ce volume étant acquis avant l'injection du produit de contraste, et de choisir comme volume source à recaler le premier

volume de l'examen post traitement acquis avant l'injection du produit de contraste. Dans cette partie nous ne cherchons pas à recaler parfaitement les données mais à fournir la meilleure initialisation possible en vue d'un éventuel recalage non rigide adapté. En effet le recalage non rigide est complexe à mettre en œuvre dans la mesure où la tumeur se déforme au cours du temps. Pour répondre à cette problématique d'initialisation, les méthodes de recalage définies dans cette section s'appuient sur différents paramétrages d'algorithmes basés sur les critères de similarité IM (Information Mutuelle) [56] et DC (Distance de Chanfrein) [47, 48] dont le paramétrage est résumé dans le tableau 7.1. Toutes les méthodes exposées ici ont été implémentées dans l'environnement Brainvisa [149].

M	transfo.	opti.	init.	similarité	prétraitements	sens de transfo
					TDM	
M_{IM}	rigide	powell	T_{z1}-cdg	IM	-	d
	rigide	powell	T_{z1}-cdg	IM	-	i
	...	...	...	...	...	...
	rigide	powell	T_{zn}-cdg	IM	-	d
	rigide	powell	T_{zn}-cdg	IM	-	i
M_{IMS}	rigide	powell	T_{z1}-cdg	IM	s	d
	rigide	powell	T_{z1}-cdg	IM	s	i
	...	...	...	...	...	...
	rigide	powell	T_{zn}-cdg	IM	s	d
	rigide	powell	T_{zn}-cdg	IM	s	i
M_{DC}	rigide	powell	cdg	DC	s_gm	d
	rigide	powell	cdg	DC	s_gm	i

TABLE 7.1: Définition des 3 types de méthodes de recalage. Avec **"cdg"** correspondant à une initalisation consistant à aligner les *centres de gravité* des coupes axiales des images ; T_{z1}, une valeur initial pour le paramètre T_z de la transformation ; **"IM"**, l'*information mutuelle* ; **"DC"**, la *distance de chanfrein* ;**"s_gm"**, indiquant qu'un *seuillage* suivi d'un *gradient morphologique* ont été nécessaires pour récupérer les contours des côtes et des vertèbres ; **"s"** un *seuillage* permettant de ne récupérer que les côtes et les vertèbres sur les images TDM ; **"d"** et **"i"**, les sens de transformation *direct* et *indirect*.

7.2.1.1 Modèle de transformation

Nous avons opté pour un modèle de transformation rigide : 3 paramètres de translation T_x, T_y, T_z et 3 angles de rotation θ_x, θ_y, θ_z.

7.2.1.2 Critères de similarité

Les algorithmes de recalage basés sur l'information mutuelle ont été appliqués sur deux types d'images : les images TDM originales ou les images TDM seuillées **(s)**, conduisant à deux types de méthodes que l'on nommera M_{IM} et M_{IMS} (pour Information Mutuelle Seuillée).

Le dernier type de méthodes utilisé que l'on nommera M_{DC}, s'appuie sur la distance de chanfrein appliquée aux contours des structures osseuses extraits de l'image.

7.2.1.3 Prétraitements

Dans le cadre de l'Information Mutuelle Seuillée (IMS), un seuillage avec une valeur empirique de seuil de 1295 a été appliqué aux images TDM afin de ne récupérer que les structures osseuses : les vertèbres et les côtes qui apparaissent en intense sur les images TDM. Ce seuillage a consisté à mettre à 0 tous les voxels en-dessous de ce seuil et à garder les niveaux de gris des voxels d'intensité supérieure à ce seuil.

Pour les méthodes basées sur la méthode de chanfrein, on s'est intéressé aux contours des structures osseuses donc aux vertèbres et aux côtes. Un seuillage binaire (seuil identique à celui utilisé pour l'IMS) permet de récupérer les régions osseuses et l'application d'un gradient morphologique **(s_gm)** permet d'extraire les contours.

7.2.1.4 Sens de transformation

Le recalage est réalisé systématiquement dans les sens direct **(d)** et indirect **(i)**.

7.2.1.5 Algorithmes d'optimisation et initialisation

L'algorithme d'optimisation de Powell a été systématiquement utilisé pour trouver les extrema des fonctions de coût. Les paramètres de translation (T_x, T_y, T_z) de la matrice de transformation ont été initialisés à partir des centres de gravité des volumes TDM et les angles de rotation (θ_x, θ_y, θ_z) ont été initialisés à une valeur de 0 pour toutes les méthodes dans un premier temps. Puis, il a été remarqué que pour les types de méthodes de recalage M_{IM} et M_{IMS}, l'initialisation correcte du T_z (correspondant au niveau de coupe) était indispensable pour obtenir un résultat satisfaisant. L'alignement des centres de gravité des images ne permettant pas de retrouver les bons niveaux de coupe étant donné le nombre limité de coupes considérées, il a été choisi d'initialiser le paramètre de transformation T_z avec autant de valeurs qu'il y a de coupes dans le plus petit des deux volumes TDM (T_{zn}). On peut considérer que chaque initialisation génère une méthode de recalage différente. Il n'a pas été jugé nécessaire d'utiliser le même processus d'initialisation de la valeur de T_z pour les méthodes basées sur la distance de chanfrein (M_{DC}).

7.2.2 Evaluation quantitative et classement

Le critère d'évaluation qui a été utilisé dans cette application est un critère de recouvrement classique, à savoir l'indice de Dice.

Les structures utilisées pour le calcul de l'indice de Dice moyen correspondent aux côtes et aux vertèbres.

Ces différentes régions ont été extraites par seuillage binaire dans un premier temps en utilisant le même seuil que pour l'IMS. Une fermeture morphologique a ensuite été effectuée sur cette image binaire pour combler les trous à l'intérieur des structures osseuses. Un seuillage sur les tailles des structures restantes a alors été appliqué pour éliminer les structures de petite taille.

Chaque structure restante a ensuite été étiquetée, fournissant des régions connexes 3D. Afin d'attribuer les mêmes étiquettes aux structures sur les examens pré et post-traitement, la stratégie suivante a été utilisée : la structure de taille la plus importante a été identifiée comme étant les vertèbres ; les autres structures ont ensuite été étiquetées en fonction de leur distance à la structure correspondant à la colonne vertébrale et de leur positionnement à gauche ou à droite de celle-ci. Etant donné l'épaisseur des coupes, l'ensemble des vertèbres constitue une structure unique. Ces étapes d'étiquetage et d'appariement entre les structures de l'image source post-traitement I_{post} recalée et de l'image de référence I_{pre} pré-traitement a été nécessaire pour calculer des indices de Dice locaux comme le montre la figure 7.1. La moyenne de ces indices permet ensuite de définir un score de Dice Global (DG), défini par l'équation (7.1), utilisé ensuite pour classer les méthodes et estimer la meilleure méthode suivant le jeu de données considéré.

$$DG = \frac{1}{N} \sum_{i=1}^{N} Dice(A_i, B_i) \tag{7.1}$$

avec N nombre de structures (côtes et vertèbres) présentes dans les images I_{post} recalée et I_{pre} ; A_i la structure numéro i dans l'image I_{pre} et B_i la structure numéro i de l'image I_{post} recalée.

7.2.3 Validation visuelle

Selon les valeurs de l'indice de Dice Global (DG), il est possible de déduire qu'un recalage est de bonne ou mauvaise qualité. Cependant une validation visuelle a été faite dans tous les cas. Cette validation s'appuie notamment sur les vertèbres pour un même niveau de coupe et sur le recouvrement des côtes, en admettant que le recouvrement au niveau des côtes situées vers l'avant de l'abdomen peut être très faible même pour un recalage rigide de qualité, à cause des déformations liées à la respiration.

En pratique, toutes ces étapes on été développées dans Brainvisa/Anatomist. Pour la validation du recalage, les volumes à comparer sont affichés dans deux fenêtres séparées pour vérifier si les deux volumes sont sur le même niveau de coupe. Ensuite les deux volumes sont fusionnés en affichant en couleur les structures osseuses d'un des deux volumes. Cette fusion permet de mieux juger la superposition des deux volumes sur le plan axial comme le montre la figure 7.1.

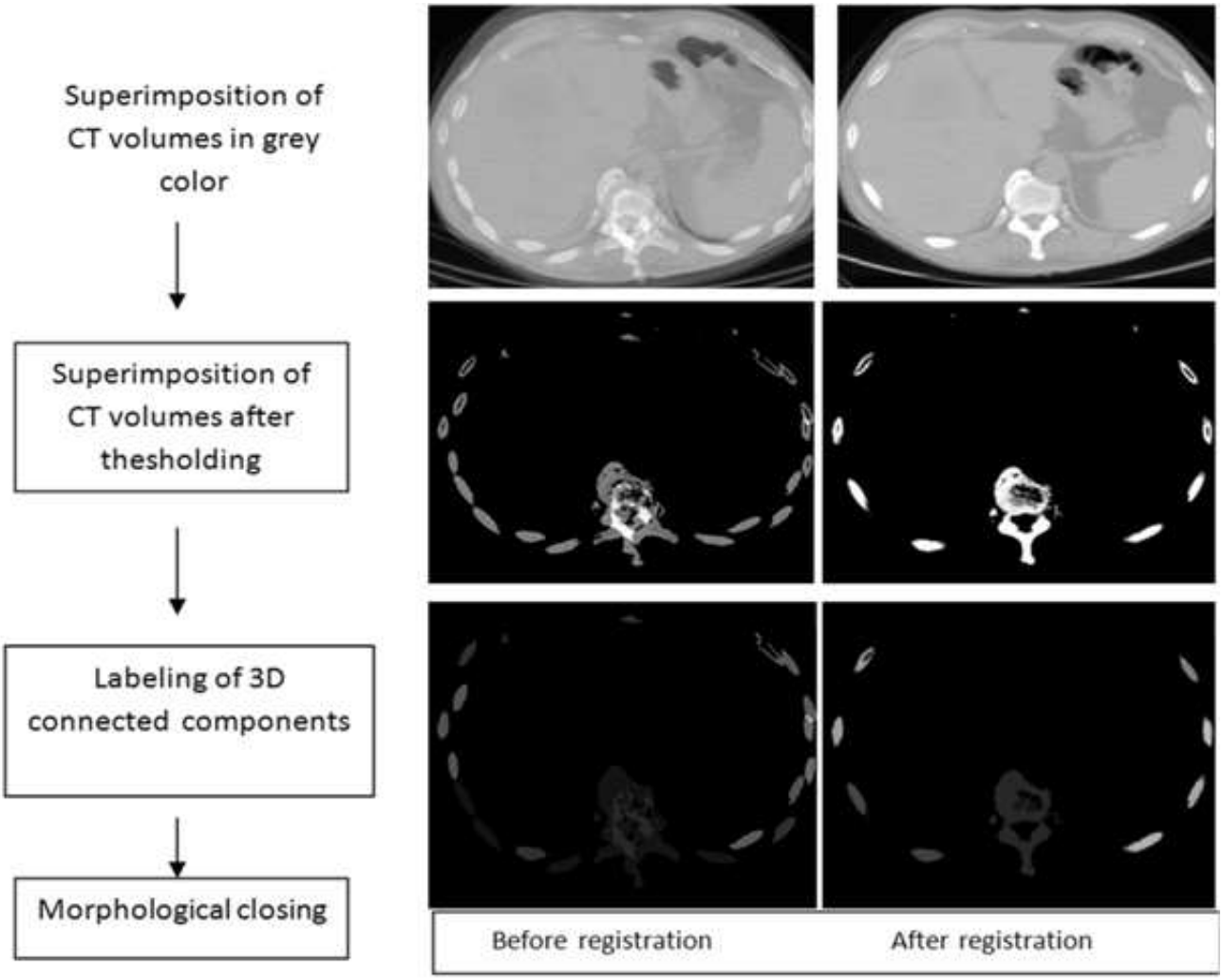


FIGURE 7.1: Les étapes du processus permettant le calcul de l'indice de Dice global DG pour l'évaluation quantitative du recalage.

7.3 Mise en œuvre de la stratégie pour le recalage d'études de TDM abdominales de perfusion au premier passage

7.3.1 Choix des différentes approches de recalage

Pour suivre l'évolution tumorale au cours du traitement, chaque patient est soumis à un examen TDM avant le début du traitement et à un ou plusieurs examens à différents temps au cours du traitement. L'acquisition de chacun de ces examens se fait après injection de produit de contraste iodé et dure entre 2 et 5 minutes pendant lesquelles le patient respire librement. Les volumes ainsi acquis sont donc des volumes 4D (3D + t) et sont composés d'autant de volumes 3D que la séquence acquise comporte de temps (environ 50, un volume par seconde puis un volume toutes les 10 secondes). Afin de compenser les mouvements respiratoires

que l'on peut observer dans une séquence dynamique, il convient de recaler ces différents volumes 3D sur un des volumes 3D de la séquence, qu'on choisit alors comme référence.

Pour répondre à cette problématique, nous avons comparé dans cette étude préliminaire deux algorithmes de recalage basés sur l'Information Mutuelle (IM) [56] mais l'un utilisant un modèle de transformation "rigide" et l'autre un modèle de transformation "élastique" (cf tableau 7.2).

M	modèle de transformation	opti.	initialisation	critère de similarité	sens de transformation
M_1	rigide	powell	cdg	IM	d-prev
M_2	élastique (FFD)	descente de gradient	cdg	IM	d-ref

TABLE 7.2: Définition des 2 types de méthodes de recalage. Avec **"cdg"** correspondant à une initalisation consistant à aligner les *centres de gravité* des coupes axiales des images ; **"IM"**, l'*information mutuelle* ; **"d-prev"** sens de transformation *direct* vers le volume précédent et **"d-ref"** sens de transformation *direct* vers le volume de référence ; **FFD**, "Free Form Deformation".

7.3.1.1 Modèles de transformation

La première méthode de recalage M_1 se base sur un modèle de transformation rigide (3 paramètres de translation T_x, T_y, T_z et 3 angles de rotation (θ_x, θ_y, θ_z)
La deuxième méthode de recalage M_2 se base sur un modèle élastique de déformations libres (FFD) [67]. Ce modèle de transformation Tr est composé de deux parties comme indiqué dans l'expression suivante :

$$Tr = Tr_{globale}(x, y, z) + Tr_{locale}(x, y, z) \tag{7.2}$$

En effet, ce modèle de transformation est composé d'une transformation globale affine qui permet de prendre en compte les grands déplacements globaux et d'une partie locale permettant de prendre en compte les petits déplacements de manière plus locale. Contrairement à la partie globale, la partie locale ne constitue pas une transformation paramétrique mais s'appuie sur une grille (G) de n_x x n_y x n_z points de contrôle qui permet de déformer l'objet. Une translation locale 3D est calculée en chacun de ces points de contrôle ($G_{i,j,k}$), puis la déformation en un point quelconque de l'image est calculée comme une combinaison de B-splines 1D [174], selon l'expression suivante [63] :

$$T_{locale}(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(w) G_{i+l,j+m,k+n} \tag{7.3}$$

avec $\lfloor x \rfloor$ signifiant partie entière de x, i = $\lfloor x/n_x \rfloor$ - 1, j = $\lfloor y/n_y \rfloor$ - 1, k = $\lfloor z/n_z \rfloor$ - 1, u = x/n_x - $\lfloor x/n_x \rfloor$, v = y/n_y - $\lfloor y/n_y \rfloor$, w = z/n_z - $\lfloor z/n_z \rfloor$, et B_i la i ème fonction B-spline.

Les déformations dépendent du nombre de points dans la grille. Lorsque la grille contient peu de points, le modèle prend en compte les grandes déformations "pseudo globales" non rigides alors que lorsqu'elle contient beaucoup de points, le modèle de transformation peut prendre en compte des déformations locales beaucoup plus précises. Afin de prendre en compte au mieux ce type de déformations, il est utile de procéder par une approche pyramidale en faisant varier le nombre de points de la grille au cours du processus de recherche des déformations. Il faut commencer par une grille avec un faible nombre de points de contrôle puis augmenter au fur à mesure le nombre de points de la grille en partant des transformations trouvées à partir des grilles précédentes. Dans cette optique nous avons fait varier le nombre de points de 4x4x2 à 10x10x6. L'algorithme de FFD que nous avons utilisé a été implémenté par Thierry Delzescaux et ses collègues dans l'environnement Brainvisa [149].

7.3.1.2 Critères de similarité

Les deux méthodes de recalage utilisées dans ce contexte sont basées sur l'information mutuelle **(IM)**. Dans l'article de Rueckert [67], pour la méthode qui utilise un modèle de transformation non rigide, un terme de contrainte supplémentaire C_{lisse} est rajouté au critère de similarité pour imposer au modèle de FFD de trouver des transformations plus lisses (cf : équation (7.4))

$$C_{lisse} = \frac{1}{V}\int_0^X \int_0^Y \int_0^Z \left[\left(\frac{\delta^2 Tr}{\delta x^2}\right) + \left(\frac{\delta^2 Tr}{\delta y^2}\right) + \left(\frac{\delta^2 Tr}{\delta z^2}\right) + \left(\frac{\delta^2 Tr}{\delta xy}\right) + \left(\frac{\delta^2 Tr}{\delta xz}\right) + \left(\frac{\delta^2 Tr}{\delta yz^2}\right)\right] \tag{7.4}$$

Donc la fonction de coût f à minimiser est la suivante :

$$f = -IM + \lambda\, C_{lisse} \tag{7.5}$$

avec λ un paramètre qui permet de régler l'influence du critère de lissage.

A la différence de Rueckert [67] qui propose une fonction de coût avec un terme de lissage, l'algorithme de FFD implémenté dans Brainvisa (qui correspond à la méthode $M2$) maximise simplement l'IM.

7.3.1.3 Sens de transformation

Comme chaque séquence dynamique acquise comporte plusieurs volumes, plusieurs possibilités quant au choix du volume de référence pouvaient être envisagées. Il a été décidé de prendre le premier volume de la séquence, celui avant l'injection du produit de contraste, comme volume de référence. Pour recaler les autres volumes de la séquence sur ce volume de référence, il est possible soit de recaler directement chaque volume sur le volume de référence (d-ref) comme cela est le cas pour la méthode M_2, soit de recaler chaque volume sur le volume précédent (d-prev) puis, par combinaison des transformations trouvées, de retrouver celle permettant de recaler chaque volume de la séquence sur le volume de référence. Cette dernière démarche a été adoptée pour l'utilisation de la méthode M_1.

7.3.1.4 Algorithmes d'optimisation et initialisation

Les paramètres de translation (T_x, T_y, T_z) de la matrice de transformation de la méthode M_1 ont été initialisés à partir des centres de gravité des volumes TDM et les angles de rotation (θ_x, θ_y, θ_z) ont été initialisés à une valeur de 0 dans un premier temps. La même initialisation a été utilisée pour la partie globale du modèle de transformation de la méthode M_2.

L'algorithme d'optimisation de Powell a été utilisé pour trouver les extrema de l'IM pour la méthode M_1. Pour la méthode M_2, l'algorithme de descente de gradient était celui proposé dans l'implémentation testée.

7.3.2 Evaluation quantitative et classement

Pour l'évaluation du recalage dynamique, nous avons utilisé le même critère que celui décrit pour le recalage longitudinal. L'indice de Dice global (DG) a donc été utilisé pour juger du recouvrement entre les structures osseuses de chaque volume. Il est bon de noter que l'extraction des structures osseuses sur les volumes acquis après injection du produit de contraste peut poser problème. En effet, sur ces volumes, le seuillage sélectionne non seulement les structures osseuses mais également les zones de réhaussement du produit de contraste et les gros vaisseaux (aorte). Pour éliminer ces structures, un masque convexe reliant les structures osseuses extraites sur le volume de référence avant injection du produit de contraste a été défini. Ce masque a ensuite été érodé en utilisant un élément structurant sphérique de 11 voxels de rayon. Ce masque a ensuite été reporté sur les volumes après injection de produit de contraste pour supprimer les structures se trouvant à l'intérieur de ce masque. Les étapes d'étiquetage et d'appariemment des strutures restantes sont alors les mêmes que celles décrites au §7.2.2.

Ainsi, une valeur de l'indice DG a pu être calculée entre chaque volume de la séquence et le volume choisi comme référence, conduisant pour chaque séquence à autant de mesures de DG que de volumes que comprend la séquence. La méthode obtenant les valeurs les plus importantes de DG pour la séquence est alors considérée comme la meilleure.

7.3.3 Validation visuelle

Afin de pouvoir juger de la qualité du recalage dynamique, la séquence recalée a été visualisée de deux manières différentes. Premièrement, les séquences recalées ont été visualisées en mode "ciné", c'est-à-dire que les volumes des différents temps sont affichés dans une même fenêtre à une cadence ajustable afin de juger de l'alignement des structures anatomiques. La deuxième méthode consiste à créer une carte représentant les profils "coronaux" de l'intensité au cours du temps. En pratique, une ligne est délimitée sur une coupe du volume TDM et l'évolution de l'intensité au cours du temps des voxels appartenant à cette ligne est représentée en colonne. Cette ligne a été placée sur la partie haute des volumes dans la vue axiale car c'est dans cette région que les déplacements dus à la respiration sont de plus grande amplitude (cf figure 7.2).

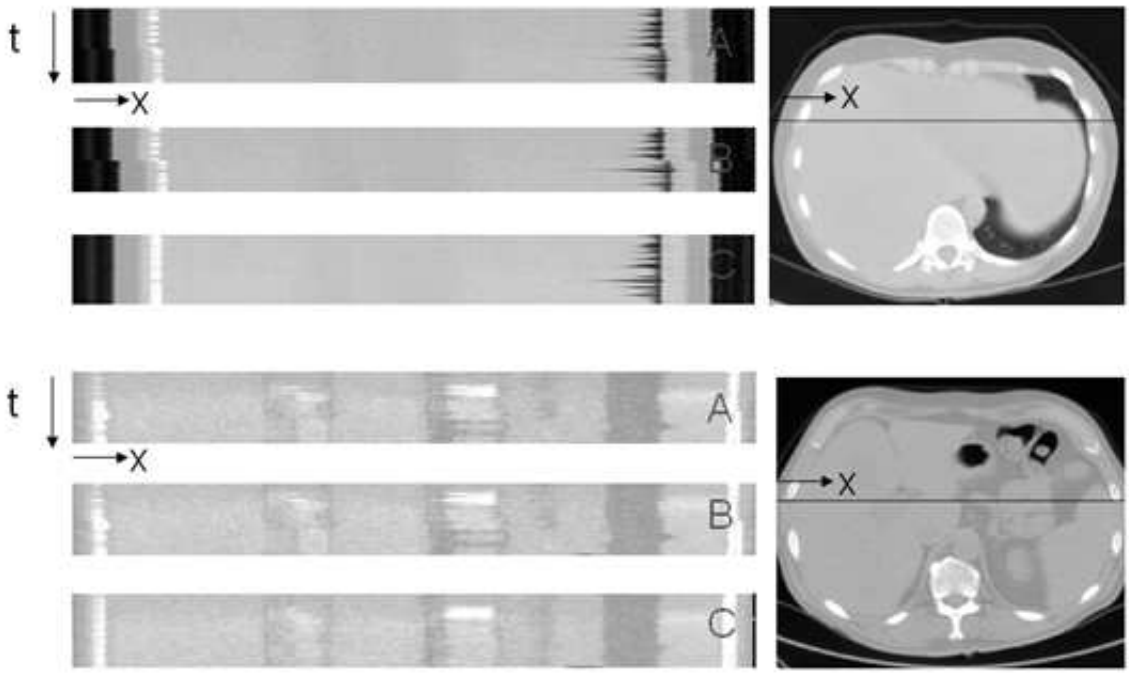


FIGURE 7.2: Profils d'intensité au cours du temps pour 2 jeux de données [173]. Le profil A correspond au profil avant recalage, le profil B au profil après recalage par la méthode rigide M1 et le profil C au profil obtenu après recalage par la méthode non rigide M2. Sur le premier jeu de données, le recouvrement des structures semble moins bon après recalage par M1 qu'avant recalage. En revanche le résultat est meilleur après recalage par la méthode M2. Pour le second jeu de données (bas), les résultats obtenus avant et après recalage par la méthode M1 sont similaires. La méthode M2 obtient un meilleur recouvrement des structures anatomiques.

7.4 Résultats du recalage TDM-TDM pour le suivi longitudinal

7.4.1 Performance de la stratégie

Le mauvais alignement entre les volumes acquis pour le suivi a été vérifié visuellement dans un premier temps (cf figure 7.3), puis quantitativement. Avant le recalage, les scores de l'indice de Dice Global (DG) calculés pour les 19 paires de volumes TDM varient entre 0 et 0,13. Les scores proches de 0 confirment en effet le mauvais alignement entre les volumes. La figure 7.4 illustre les valeurs prises par l'indice de Dice Global suivant les différentes initialisations de la composante de translation selon l'axe z, testées pour la méthode M_{TMI} ; elle souligne clairement la nécessité de tester toutes les superpositions possibles entre les niveaux de coupes axiales des volumes avant et après traitement pour trouver la valeur initiale de T_z la plus satisfaisante et montre que sinon la méthode de recalage aboutit à un minimum local. Les scores de Dice Global (DG) obtenus après recalage en utilisant la valeur optimale de T_z varient entre 0,08 et 0,84 et sont toujours plus élevés que les scores avant recalage. Ces valeurs ont confirmé que certains recalages sont de mauvaise qualité (DG proche de 0) alors que d'autres recalages sont corrects (DG proche de 1). Un exemple de recalage de bonne qualité est illustré sur la figure 7.3. La figure 7.5 représente les valeurs de l'indice de Dice Global (DG) des 19 paires de volumes TDM obtenues par les trois méthodes de recalage dans les modes directs et indirects. Les indices des recalages considérés comme de bonne qualité après vérification visuelle sont marqués en rouge alors que les autres sont marqués en bleu. Dans 97% des cas, l'indice DG est cohérent avec l'observation visuelle, montrant des valeurs plus élevées pour les recalages de bonne qualité que pour les recalages de mauvaise qualité. En outre, les recalages qui ont été jugés visuellement corrects ont des scores de DG allant de 0,26 à 0,84. Les recalages ayant été jugés visuellement de mauvaise qualité ont des scores de DG allant de 0,10 à 0,72. Au final, au delà d'un seuil de 0,7 de la valeur de DG, les recalages peuvent être considérés comme de bonne qualité et en-dessous d'un seuil de 0,25 on considère que les recalages sont de mauvaise qualité. En revanche lorsqu'un recalage obtient un score de DG compris entre 0,25 et 0,7 une vérification visuelle est obligatoire pour juger de la qualité du recalage.

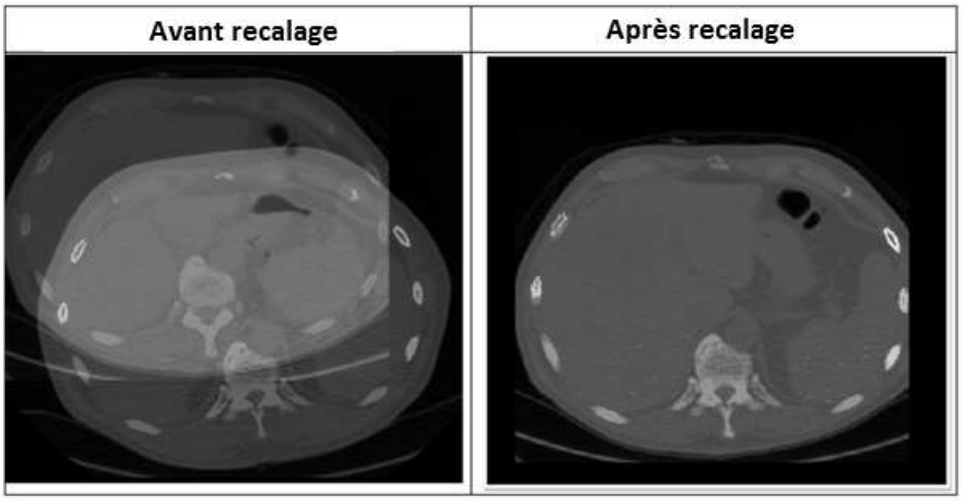


FIGURE 7.3: Superposition des coupes axiales d'une paire de volumes avant recalage (à gauche) et après recalage (à droite). Les images sont affichées en niveaux de gris et les structures osseuses du volume acquis après traitement anti-angiogénique sont colorées. Les voxels qui apparaissent en rouge correspondent à du bruit dans le foie et dans les reins et ne sont pas étiquetés comme des structures osseuses par l'algorithme. Le score de DG avant recalage est de 0,08 et il est de 0,78 après recalage.

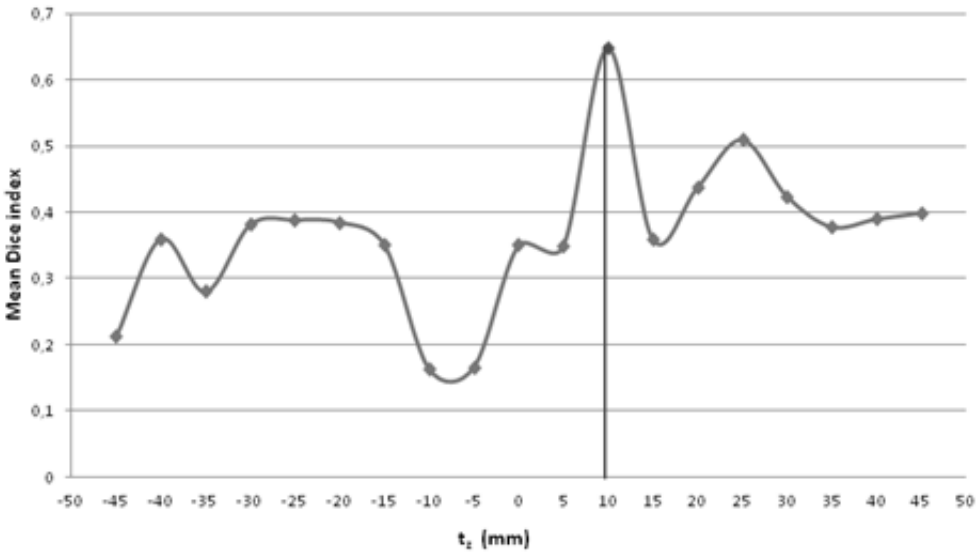


FIGURE 7.4: Valeurs de DG en fonction de la valeur initiale de T_z obtenue pour une paire de volumes utilisant la méthode M_{TMI}

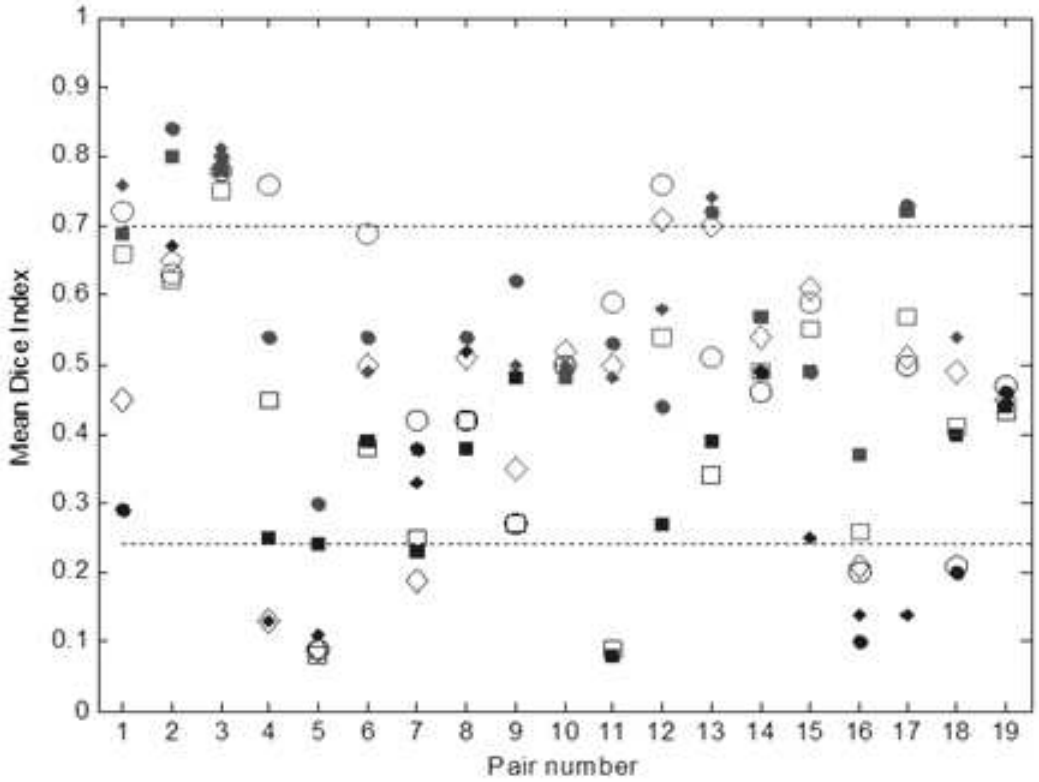


FIGURE 7.5: Valeurs de l'indice de Dice Global (DG) obtenues par différentes méthodes de recalage sur 19 paires de volumes TDM. Les résultats correspondant à un recalage de mauvaise qualité selon la validation visuelle sont représentés en bleu alors que les résultats correspondant à des recalages de bonne qualité sont affichés en rouge. Les carré pleins et les carrés vides correspondent aux scores de DG obtenus par les méthodes basées sur l'IM (M_{IM}) respectivement dans le sens direct et indirect. Les cercles pleins et les cercles vides correspondent aux scores de DG obtenus par les méthodes basées sur l'IM seuillée (M_{IMS}) respectivement dans le sens direct et indirect. Les losanges pleins et les losanges vides correspondent aux scores de DG obtenus par les méthodes basées sur la distance de chanfrein DC (M_{DC}) respectivement dans les sens direct et indirect.

Le tableau 7.3 donne le résumé du nombre de recalages de qualité correcte obtenus par chaque méthode selon les critères visuels. Les résultats des modes direct et indirect sont présentés séparément. Dans certains cas ils ne donnent pas les mêmes résultats. Les méthodes basées sur la maximisation de l'information mutuelle utilisant les volumes originaux (M_{IM}) obtiennent les scores les plus faibles, avec 42% et 47% de recalages de qualité correcte pour les sens direct et indirect respectivement. Les méthodes basées sur la maximisation de l'information mutuelle des volumes seuillés (M_{IMS}) et les méthodes basées sur la distance de chanfrein (M_{DC}) obtiennent des résultats assez similaires (68% et 53% de recalages corrects pour M_{IMS} et 53% et 68% pour M_{DC} dans les sens direct et indirect respectivement). Deux paires de volumes TDM ont été correctement recalées par les trois méthodes. Dans la plupart des cas, un recalage de qualité correcte est obtenu par une ou deux méthodes, mais pour deux paires de volumes TDM (P7 et P19), aucun recalage n'a été satisfaisant. Il n'est pas possible de prédire quelle méthode il faut utiliser pour obtenir un recalage correct : non seulement le choix de la méthode adéquate pour le recalage dépend des données, mais le résultat dépend également du sens de transformation utilisé. En effet, certains jeux de données sont correctement recalés pour un sens de transformation et non pour l'autre. Cela démontre clairement qu'il n'y a pas de méthode qui fonctionne pour toutes les situations. Les scores quantitatifs DG nous ont permis de confirmer cette tendance et de classer les différentes méthodes de recalage. Les méthodes M_{IMS} et M_{DC} sont classées en premier tandis que les méthodes M_{IM} arrivent en dernier. Cependant, pour un seul cas, M_{IM} obtient un recalage de qualité correcte, ce qui confirme l'intérêt de combiner toutes les méthodes testées. La sélection des meilleurs résultats selon les six possibilités de recalage (M_{IM}, M_{IMS} et M_{DC} dans les deux modes direct et indirect) a permis d'augmenter le taux de réussite de recalage de 68% (pour le meilleur des cas avec une seule méthode) à 89%.

	M_{IM}		M_{IMS}		M_{DC}		Sélection de la
	direct	indirect	direct	indirect	direct	indirect	meilleure méthode
NRC	8	9	13	10	10	13	17
PRC	42%	47%	68%	53%	53%	68%	89%

TABLE 7.3: Nombre (NRC) et pourcentage (PRC) de recalages de bonne qualité obtenus par chaque méthode sur la base de données des 19 paires de volumes TDM.

Enfin, le tableau 7.4 présente la moyenne des valeurs absolues et les écarts-types des paramètres de transformation qui ont été obtenus pour le meilleur recalage, c'est-à-dire celui qui fournit la valeur la plus élevée de l'indice Dice Global (DG). Les angles de rotation estimés obtenus après recalage n'étaient que de quelques degrés. Les translations estimées ont des amplitudes de quelques centimètres et sont du même ordre de grandeur dans le plan (x, y) que suivant l'axe z. Cela correspond à un écart moyen de cinq à six coupes entre paires de volumes.

T_x (mm)	T_y (mm)	T_z (mm)	θ_x ($^\circ$)	θ_y ($^\circ$)	θ_z ($^\circ$)
19 ± 16	36 ± 28	32 ± 23	4,5 ± 4,1	1,3 ± 1,2	2,6 ± 2,1

TABLE 7.4: Moyennes et écarts-types des paramètres de transformation estimés après sélection de la meilleure méthode de recalage pour chaque jeu de données.

7.4.2 Discussion

Trois méthodes de recalage de volumes TDM abdominaux ont été proposées avec différents types d'initialisation pour le suivi de patients atteints de lésions hépatiques et sous traitement anti-tumoral. Deux méthodes sont basées sur une approche iconique utilisant l'information mutuelle comme critère de similarité tandis que la troisième méthode est basée sur une approche de minimisation de distance entre contours utilisant la distance de chanfrein. Ces méthodes de recalage rigide ont été appliquées dans les deux sens de transformation (direct et indirect). Une évaluation visuelle a été réalisée ainsi qu'une évaluation quantitative basée sur l'indice de Dice Global (DG) appliqué à des structures osseuses (vertèbres et côtes) de manière à rendre cette évaluation reproductible. Les résultats quantitatifs ont été corroborés par l'évaluation visuelle. En effet, l'étude expérimentale effectuée sur une base de 19 paires de volumes a montré que l'indice DG est robuste et pertinent pour l'évaluation du recalage longitudinal. La combinaison de l'analyse visuelle avec l'utilisation de l'indice DG a mis en évidence trois groupes d'examens. Les recalages du groupe 1 ont obtenu des valeurs de DG supérieures à 0,7 et ont été classés comme visuellement corrects (18 cas répartis sur les données de 7 patients). Les recalages du groupe 2 ont obtenu des valeurs de DG inférieures à 0,25 et correspondent à des recalages de mauvaise qualité (15 cas répartis sur les données de 7 patients). Les recalages du groupe 3 ont obtenu des valeurs intermédiaires de DG comprises entre 0,25 et 0,7 (81 cas répartis sur les données des 19 patients). Pour ce dernier groupe, les recalages doivent être vérifiés visuellement pour établir leur qualité. La classification des méthodes basées sur l'évaluation quantitative a montré que M_{IMS} obtient en général des recalages de meilleure qualité que M_{IM}, ce qui suggère que l'étape de pré-traitement (seuil) a une grande influence sur le recalage. La valeur de seuil a été définie empiriquement sur la base de données. Il a été choisi de garder la même valeur de seuil pour chaque examen plutôt qu'un seuil adaptatif, pour éviter d'avoir à sélectionner de nouveau un seuil pour chaque examen. La pertinence de ce choix pourrait être étudiée, même si ce choix paraît adapté dans le cadre de cette étude. L'évaluation quantitative a également montré que M_{IMS} et M_{DC} sont équivalentes en nombre de bons recalages obtenus sur la base, mais ne sont pas nécessairement toujours efficaces dans les mêmes situations. La performance de chaque méthode dépend ainsi grandement des données d'entrée et une méthode pourrait échouer pour une paire et réussir pour une autre où d'autres méthodes qui fonctionnent habituellement échouent. La comparaison entre les recalages bidirectionnels a également indiqué que les paramètres de transformation trouvés peuvent différer en fonction du sens de la transformtion estimée. En effet, les algorithmes d'optimisation peuvent être piégés dans un minimum local dans un sens et pas nécessairement dans l'autre. Par conséquent, tester les recalages dans les deux sens permet d'augmenter la probabilité de trouver une transformation correcte.

7.4.3 Limites et perspectives

Malgré la stratégie de recalage mise en place visant à sélectionner la meilleure méthode de recalage pour chaque jeu de données, deux paires de TDM restent mal recalées. Ces recalages de mauvaise qualité pourraient venir d'artéfacts présents dans les images ou du fait qu'il n'existe que très peu de coupes communes entre les deux volumes TDM, compliquant ainsi grandement le recalage. Il convient de noter que la base de données étudiée comprenait tous les cas vus en clinique sans pré-sélection en fonction de la mise en œuvre des acquisitions ou de la qualité des données. En particulier, la taille du volume imagé a été limitée par les

directives RECIST qui demandent à ce que la dose de rayonnement reste en-dessous d'un seuil recommandé. Certaines difficultés viennent du nombre restreint de coupes acquises qui ont réduit fortement les possibilités d'adaptation dans la direction de l'axe z et de l'anisotropie des voxels avec une épaisseur de coupe importante, environ 3 à 6 fois plus grande que la taille des pixels dans les coupes axiales. Pour les paires de volumes associés aux patients P7 et P19 pour lesquelles les trois approches de recalage ont échoué, les meilleurs scores de DG étaient de 0,42 et 0,47 respectivement. Pour P19, la difficulté est liée à des artefacts de balayage, en raison d'un implant métallique situé sur les vertèbres (figure 7.6) aboutissant à une distorsion des formes osseuses obtenues après l'opération de seuillage. Cependant, il est intéressant de noter que la stratégie de recalage a permis de recaler une autre paire d'examens (P18) sur laquelle le même type d'artéfacts était présent. Dans ces cas difficiles, une solution pratique pourrait consister en un recalage manuel des examens de TDM. La paire P7 était un autre cas compliqué à recaler car les examens pré et post-traitements sont très différents. Tout d'abord, une différence dans la forme du foie et une augmentation de la graisse ont pu être observées entre les deux examens. De plus, une angulation de la colonne vertébrale a été constatée dans l'une des deux acquisitions. Ces deux problèmes ont rendu le modèle de recalage rigide très approximatif. Ces deux cas soulignent également la nécessité de vérifier visuellement les examens lorsque des scores de DG intermédiaires ($> 0{,}25$ et $< 0{,}7$) sont obtenus, afin d'essayer d'identifier les causes du recalage de mauvaise qualité.

Les trois méthodes de recalage rigides testées dans cette étude ont des temps d'exécution très différents. Les méthodes les plus chronophages sont M_{IM} et M_{IMS} avec un temps d'exécution environ 50 fois plus élevé que la méthode M_{DC}, pour les volumes TDM de plus de 20 coupes. Cette différence significative dans le temps de traitement est principalement due à l'étape d'initialisation consistant à tester toutes les valeurs initiales de T_z possibles pour les méthodes basée sur l'IM ; comme indiqué précédemment, cette étape est une condition essentielle pour éviter d'être pris au piège dans un minimum local lors de la procédure de recalage. Le temps d'exécution a été considérablement réduit pour M_{DC} puisque cette méthode ne nécessite pas ce type d'initialisation et parce qu'elle fait appel à un critère non iconique. Une solution efficace pour réduire le temps de calcul induit par la stratégie proposée serait de commencer systématiquement les recalages avec la méthode M_{DC}. Le critère quantitatif éventuellement assisté par une évaluation visuelle pourrait valider le recalage ou décider de lancer la méthode de recalage M_{IMS}, en raison de son taux élevé de succès dans l'étude présentée. L'utilité du test systématique de toutes les coupes dans le processus d'initialisation de T_z devrait être étudiée. Pour réduire le temps de calcul, un test consistant à sous-échantillonner les volumes TDM a été mené mais n'a pas été concluant. Enfin, M_{IM} pourrait être testé uniquement en cas de défaillance de M_{IMS}.

Bien sûr, le temps d'exécution de la stratégie proposée est le prix à payer pour augmenter le taux de bons recalages dans un contexte clinique. Ce type d'approche serait bien adapté pour une mise en œuvre sur un "cluster" d'ordinateurs afin de réduire les temps de calcul.

Le recalage rigide est une première étape essentielle qui imite l'expertise du radiologue. C'est une étape d'initialisation essentielle en vue d'un alignement ultérieur précis à l'aide de méthodes de recalage non rigides, en particulier pour les applications de scanner abdominal où les tissus mous subissent différentes déformations élastiques. Par conséquent, dans un travail futur, des transformations plus complexes comme des transformations affines et / ou des transformations élastiques seront testées pour affiner les résultats

obtenus à partir des méthodes rigides. Cet algorithme de recalage élastique devra être spécialement conçu pour l'application du suivi de patient atteints de tumeur hépatiques où le nombre de coupes communes peut être très faible entre les volumes pré et post-traitement, et où les tumeurs peuvent changer considérablement en taille et en forme. Les lésions tumorales ne peuvent pas constituer des repères pour le recalage. La méthode générique présentée dans ce document pourrait être appliquée sans modification à ces modèles de recalage plus complexes.

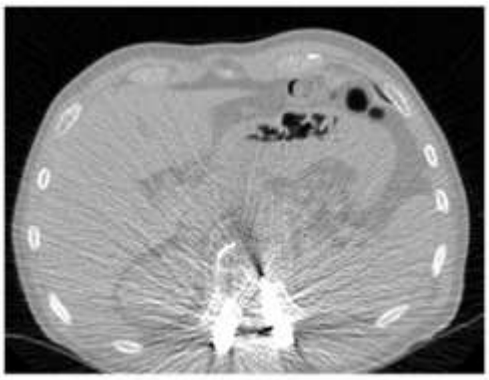


FIGURE 7.6: Image TDM d'un patient avec un implant métallique provoquant un artefact sur l'image et rendant difficile la segmentation des structures osseuses (notamment des vertèbres).

7.4.4 Conclusion

La stratégie générale pour comparer et choisir la méthode de recalage la plus adaptée à chaque jeu de données a été appliquée pour le suivi longitudinal de patients atteints de tumeurs hépatiques réalisé au moyen d'examens de TDM abdominale acquis avant et quelques semaines après le début d'une thérapie anticancéreuse. Dans cette application, trois méthodes de recalage rigide ont été comparées, à la fois quantitativement et qualitativement. Un bon accord entre les deux types d'évaluation (qualitative et quantitative) a été observé. Le critère quantitatif, basé sur l'indice de Dice Global (DG) calculé sur les structures osseuses, a montré sa capacité à évaluer les méthodes de recalage et peut aider dans le choix de la meilleure méthode. L'application de cette stratégie sur une base de données des 19 paires de volumes a démontré que, même si M_{IMS} a tendance à obtenir de meilleurs résultats que les autres, il existe des situations où une méthode moins efficace peut réussir là où la meilleure méthode a échoué. Ce travail monte que l'utilisation d'une méthode de recalage unique n'est pas suffisante pour assurer un bon taux de réussite sur toutes les données cliniques. Elle a également montré qu'il subsiste des cas difficiles lorsque deux volumes ont peu de coupes en commun ou lorsque des artéfacts d'amplitude importante sont présents dans les données. Pour cette application de suivi longitudinal, seules les méthodes de recalage rigide ont été prises en considération pour servir d'initialisation à une méthode non rigide. Pour ce type de données, l'utilisation de méthodes élastiques doit être adaptée afin de ne pas sous-estimer les modifications des lésions tumorales. En revanche, n'importe quel type de méthodes de recalage ou de combinaisons de méthodes peut être inclus dans la stratégie de recalage et par conséquent les méthodes élastiques pourraient également intervenir dans la stratégie sans modifier le

critère quantitatif.

7.5 Résultats du recalage TDM-TDM sur les études dynamiques de perfusion

7.5.1 Performance de la stratégie

Les scores de Dice Globaux (moyennés sur l'ensemble des images d'une séquence) ont été calculés pour huit examens (tableau 7.5). Avant recalage, la valeur moyenne de DG était de 0,78. Après recalage rigide, les valeurs de DG sont globalement plus faibles avec une moyenne de 0,73, alors qu'après recalage élastique, DG est largement supérieur à celui avant recalage, avec une moyenne de 0,90. Par conséquent, le recalage élastique améliore significativement les scores de DG, ce qui indique une amélioration de l'alignement des structures osseuses. Les scores de Dice Global obtenus après recalage rigide montre clairement que le recalage rigide est inefficace pour réaligner les structures osseuses dans le contexte du recalage dynamique. En plus des mouvements respiratoires, des déplacements fortuits ont été identifiés sur certains volumes. Ce type de déplacement peut être identifié sur la courbe représentant DG au cours du temps (figure 7.7). Sur cette courbe, les faibles variations autour d'un score moyen représentent le mouvement respiratoire. Une baisse importante des scores indique un mouvement de plus grande amplitude, correspondant à un mouvement non physiologique du patient sur la table d'examen. Ces interprétations ont été confirmées par une inspection visuelle. Après recalage rigide, une correction partielle des mouvements a été observée pour certains volumes temporels, avec des scores plus élevés qu'avant recalage. Cependant, les résultats globaux montrent l'incapacité de la méthode rigide à aligner correctement les structures indépendamment du type de déplacement à compenser. Le recalage élastique est généralement efficace, même pour des déplacements de grande amplitude (figure 7.7). Sur l'ensemble de la base de données considérée, les volumes ont des valeurs de DG plus élevées après recalage élastique qu'auparavant dans 89% des volumes. Par ailleurs, les variations de DG d'un volume à l'autre sont réduites, ce qui montre une bonne compensation des mouvements respiratoires. Les cartes visuelles de profils d'intensité au cours du temps (figure 7.2) ont confirmé qu'un mauvais alignement des structures osseuses est obtenu après recalage rigide, tandis qu'un bon alignement est obtenu après recalage élastique. Ces cartes montrent aussi un meilleur alignement des structures internes après recalage élastique. Cependant certaines structures peuvent rester mal recalées localement.

	E1	E2	E3	E4	E5	E6	E7	E8
Avant traitement	0,74 ±0,14	0,33 ±0,28	0,77±0,15	0,87±0,06	0,87±0,05	0,93±0,02	0,84±0,07	0,90±0,03
Après recalage rigide	0,66±0,16	0,27±0,29	0,67±0,25	0,84±0,10	0,81±0,10	0,91±0,04	0,81±0,09	0,91±0,03
Après recalage élastique	0,86±0,04	0,77±0,19	0,92±0,04	0,96±0,03	0,93±0,03	0,93±0,03	0,88±0,03	0,94±0,01

TABLE 7.5: Scores de Dice Global pour les 8 examens (E1 à E8)

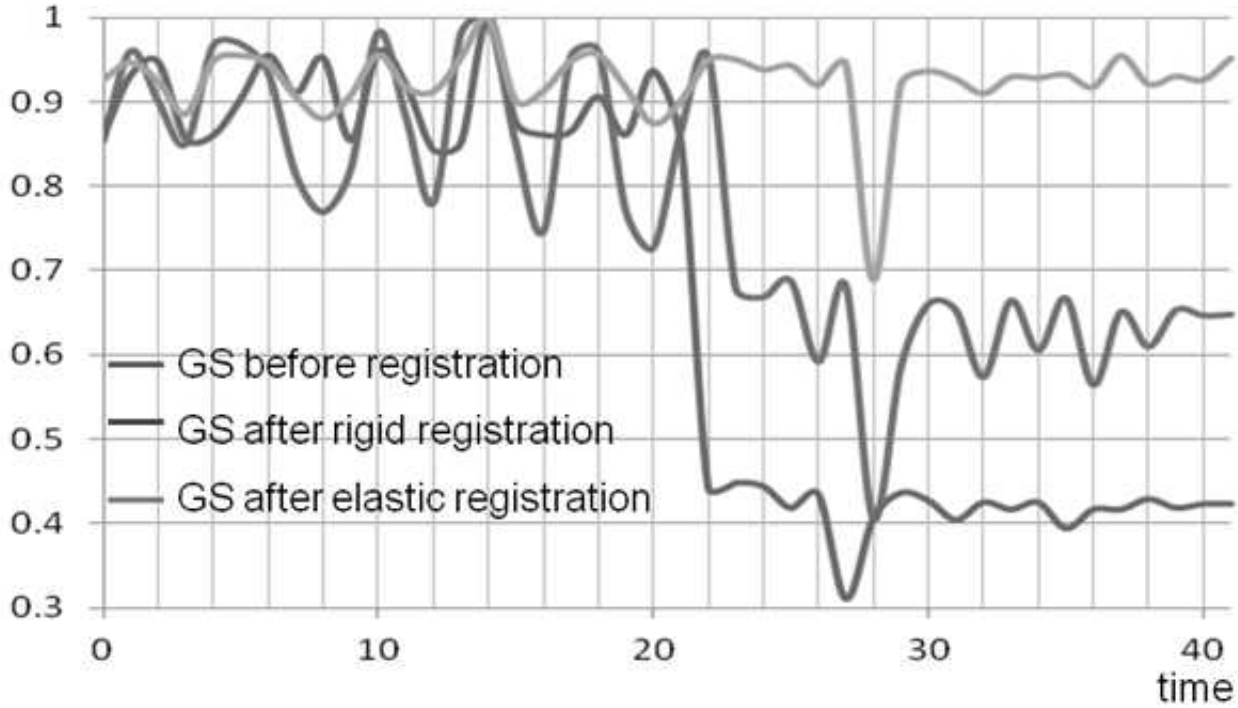


FIGURE 7.7: Scores de DG pour l'ensemble des volumes dynamiques correspondant à un examen de perfusion

7.5.2 Discussion

Une méthode de recalage rigide basée sur la maximisation de l'information mutuelle ainsi qu'une méthode de recalage élastique basée sur un algorithme FFD ont été appliquées sur des volumes de perfusion TDM de tumeurs hépatiques pour corriger les mouvements entre les images qui composent la séquence TDM. Une méthode d'évaluation quantitative basée sur la mesure de l'indice de Dice Global (DG) basé sur les structures osseuses (vertèbres et côtes) a été proposée pour évaluer la qualité de l'alignement des structures osseuses et comparer les méthodes de recalage. Une évaluation visuelle basée sur les cartes de profils d'intensité au cours du temps a permis de confirmer ces résultats quantitatifs et évaluer plus localement l'alignement interne des structures. La méthode quantitative a montré son intérêt pour l'évaluation de l'alignement des structures osseuses et la distinction entre plusieurs types de mouvement. Elle permet de confirmer ou d'invalider la correction des mouvements respiratoires et les déplacements de forte amplitude. Les côtes sont des structures représentatives du mouvement pour le recalage, en particulier les côtes les plus éloignées des vertèbres. Il serait intéressant de rechercher une pondération des critères en fonction de la position des structures osseuses, par rapport aux vertèbres, avec l'affectation d'un poids supérieur aux structures testées éloignées des vertèbres par exemple. Les cartes des profils d'intensité au cours du temps fournissent plus d'informations locales sur l'alignement des structures internes. Cette approche a également été utilisée pour évaluer le recalage non rigide de données de perfusion en IRM abdominale [175]. La quantification de l'alignement des structures internes reste à développer. Nous avons montré que, comme prévu, les méthodes de recalage rigide ne sont pas adaptées pour la correction des mouvements respiratoires de volumes abdominaux. L'équipe avait déjà démontré l'intérêt d'un recalage rigide dynamique limité à des régions d'intérêt centrées sur les tumeurs hépatiques en utilisant un algorithme d'appariement de blocs régularisé [176]. Il serait intéressant de comparer cette approche au recalage basé sur l'algorithme FFD restreint à des régions d'intérêt. En effet, dans cette situation l'approximation des mouvements respiratoires par une transformation rigide est plausible. Un recalage dynamique élastique sur l'ensemble du champ de vue en ajustant une méthode de correspondance de blocs 3D a récemment été proposé par Romain et al [177]. Contrairement à notre approche de FFD, les structures osseuses sont masquées pour le recalage. En raison de ce masquage, la comparaison des deux approches est difficile en utilisant le critère proposé. En revanche, une quantification basée sur les cartes de profils d'intensité pourrait permettre de comparer objectivement ces deux approches.

Rueckert et al [63] ont montré que le recalage FFD élastique pouvait améliorer la correction des mouvements respiratoires par rapport aux méthodes rigides ou affines en IRM de perfusion au niveau du sein. Leurs résultats ont également montré que la méthode élastique était plus efficace avec un nombre accru de points de contrôle. Ainsi, notre méthode de recalage élastique doit être optimisée en étudiant l'impact de différents paramètres sur la précision du recalage (nombre de points de contrôle, nombre de «classes» pour l'histogramme conjoint utilisé dans l'IM, etc.). L'approche quantitative que nous avons proposée pourrait être utilisée pour tester les avantages potentiels de ces nouveaux paramètres.

7.5.3 Conclusion

Ce travail propose une méthode d'évaluation automatique pour contrôler la qualité des recalages dynamiques TDM, dans le cas de l'imagerie de perfusion abdominale. Cette évaluation s'est basée sur la définition

d'un score global mesurant la superposition de structures osseuses (vertèbres et côtes). Il a été appliqué avec succès à un premier ensemble de huit examens dynamiques, pour lesquels deux modèles de recalage différents ont été testés. Les méthodes élastiques sont plus efficaces que les méthodes rigides pour le recalage des volumes dynamiques abdominaux TDM, lorsque l'on considère l'ensemble du champ de vue.

7.6 Conclusion

Nous avons présenté dans ce chapitre deux nouvelles applications de mise en œuvre de la stratégie de recalage exposée au chapitre 3. La stratégie de recalage consiste à utiliser plusieurs méthodes de recalage pour un même jeu de données et à sélectionner la plus adaptée pour chaque jeu de données à l'aide d'un critère quantitatif semi ou entièrement automatique.

La première application concerne le recalage d'examens de TDM abdominale de perfusion acquis à différents temps du traitement. En pratique, comme l'évolution de la tumeur complexifie la tâche du recalage, nous nous sommes focalisés sur l'utilisation de méthodes rigides en vue d'initialiser le recalage pour l'utilisation potentielle d'un algorithme non rigide. Les méthodes que nous avons utilisées reposent sur la distance de chanfrein, l'information mutuelle et l'information mutuelle "seuillée". Le critère quantitatif utilisé pour évaluer ces différentes méthodes se base sur le recouvrement des structures osseuses telles que les côtes et les vertèbres. Les méthodes basées sur l'information mutuelle seuillée ont montré de meilleures performances que toutes les autres méthodes cependant leur initialisation requiert un temps de calcul important, contrairement à la méthode de chanfrein qui obtient également des résultats corrects et qui pourrait être utilisée pour initialiser les méthodes basées sur l'information mutuelle seuillée. De plus, nous avons montré que l'utilisation de la stratégie de recalage permettait d'obtenir plus de recalages corrects que l'utilisation unique des méthodes de recalage basées sur l'information mutuelle seuillée.

La seconde application concerne le recalage de données TDM abdominales dynamiques de perfusion. Le recalage consiste à recaler tous les volumes de la séquence dynamique sur un volume de référence. Pour ce faire, la stratégie de recalage a été appliquée sur ces données en utilisant deux méthodes de recalage : une méthode de recalage rigide, et une basée sur l'algorithme de FFD. Ces deux méthodes utilisent l'information mutuelle en tant que critère de similarité. Afin de classer les méthodes, deux critères d'évaluation ont été utilisés. Le premier critère d'évaluation est le même que celui utilisé pour la première application. Le second consiste en l'étude d'un "profil coronal" de l'intensité au cours du temps qui permet de juger de l'alignement des structures internes. Après évaluation, la méthode non rigide obtient de meilleurs résultats que la méthode rigide dans chaque cas. Dans ce contexte, la stratégie a permi d'identifier la meilleure méthode à utiliser.

En conclusion, nous avons démontré que la stratégie proposée au chapitre 3 était pertinente pour plusieurs applications de recalage.

Conclusion et perspectives

Conclusion

Le suivi longitudinal en oncologie consiste à évaluer l'évolution tumorale afin d'adapter au mieux le traitement pour chaque patient, dans le but d'accroître la durée de vie et/ou la qualité de vie. Ce suivi est rendu possible notamment par l'imagerie médicale. Plusieurs choix se posent : celui des modalités d'imagerie à utiliser, celui des informations à extraire contenues dans les images pour évaluer l'évolution tumorale et celui des méthodes à utiliser pour l'extraction fiable de cette information. L'objectif de cette thèse était de développer une méthodologie pour aider à répondre à ces questions. Ainsi nous avons proposé au chapitre 3 un système d'analyse de l'évolution tumorale s'appuyant sur l'estimation de critères spatio-temporels permettant une analyse globale (liée aux volumes des tumeurs) et locale (étude du positionnement relatif des volumes les uns par rapport aux autres). Ce système est composé de trois étapes :

- le recalage des données sur un examen de référence permettant de réaligner les examens acquis à différents temps du traitement ou provenant de différentes modalités ;
- la segmentation des volumes tumoraux détectés sur les images des différentes modalités impliquées dans le protocole ;
- le calcul d'indices, basés sur les segmentations précédentes, pour évaluer l'évolution tumorale de manière globale et locale.

En fonction des modalités d'imagerie impliquées dans le processus de suivi longitudinal, l'étape de recalage peut s'avérer complexe. Afin d'améliorer le nombre de recalages de qualité satisfaisante, une nouvelle stratégie consistant à utiliser plusieurs méthodes de recalage et à sélectionner la plus adaptée à chaque jeu de données a été proposée. Cette stratégie se décompose elle-même en trois étapes appliquées à chaque jeu de données :

- la mise en œuvre de plusieurs méthodes de recalage ;
- la sélection de la "meilleure" méthode après classement à l'aide d'un critère d'évaluation quantitatif ;
- la validation visuelle par un expert du recalage proposé.

Dans cette stratégie, le terme de "méthode de recalage" définit une mise en œuvre particulière d'un algorithme de recalage. Ainsi un même algorithme de recalage utilisant deux paramétrages différents conduit à deux méthodes de recalage. Le choix des méthodes de recalage à mettre à œuvre et la définition des critères d'évaluation quantitatifs dépendent du type de données considéré, la stratégie proposée reste générique. Cette approche s'est avérée efficace pour un certain nombre d'applications envisagées dans cette thèse.

Pour la segmentation des zones tumorales, des méthodes classiques sont utilisées.

Enfin, les variations au cours du traitement des volumes des zones tumorales et des variations d'intensité moyenne au sein de ces zones permettent d'évaluer globalement l'évolution tumorale. Afin de mener une analyse locale, nous avons proposé deux nouveaux indicateurs, l'un prenant la forme d'un histogramme traduisant le nombre de voxels d'un volume en fonction de la distance à un autre volume et l'autre d'un histogramme d'intensités en fonction des distances entre volumes. Ces deux indicateurs se basent sur l'hypothèse suivante : au cours du temps, il peut se produire des évolutions régionales de la zone tumorale active (apparaissant en hyper ou hypo signal suivant les modalités d'imagerie considérées) et qu'une zone détectée tardivement se trouve en dehors de celle détectée sur un examen acquis à un temps antérieur, et ce, même en cas de diminution globale du volume tumoral.

La stratégie d'analyse a été appliquée à l'étude de patients atteints de gliomes de haut grade traités par antiangiogénique. Pour le suivi de patients atteints de gliomes, la modalité d'imagerie utilisée en routine clinique est l'IRM, cependant celle-ci ne permet que difficilement de faire la distinction entre récidive tumorale et radionécrose. Dans le cadre du suivi de patients atteints de gliomes et traités par radiothérapie, l'utilisation de la modalité Sestamibi TEMP a été suggérée en complément de l'IRM pondérée en T1 acquise après injection de produit de contraste. Afin d'établir réellement la complémentarité de ces deux modalités dans le cadre de la thérapie par antiangiogénique, le système que nous avons proposé a été appliqué aux données IRM et TEMP. Pour les données TEMP deux examens étaient disponibles, le premier acquis 15 minutes après l'injection du radiotraceur et le second plus tardif acquis 3 heures après l'injection.

Comme il s'agit d'une étude impliquant plusieurs modalités d'imagerie, l'étape de recalage comporte plusieurs applications :

- le recalage multimodal de chaque examen TEMP sur l'examen IRM correspondant ;
- le recalage monomodal des examens IRM après le début du traitement sur les examens acquis avant le début du traitement.

En raison de la distribution du radiotraceur (Sestamibi marqué au Technétium) utilisé pour la TEMP, le recalage multimodal est difficile et les méthodes de recalage classiques peuvent échouer. Pour résoudre ce problème, la stratégie de recalage définie au chapitre 3 a été utilisée. Vingt méthodes de recalage utilisant un modèle de transformation rigide et basées sur des critères de similarité iconiques (IM, IMN, RC, etc.) ou de distance entre contours (DC) ont été utilisées. Afin de sélectionner la meilleure méthode pour chaque jeu de données TEMP/IRM, un critère quantitatif, UC, a été défini. Il se calcule en estimant l'intensité moyenne de l'image TEMP au sein de structures anatomiques, détectées sur l'IRM après l'injection de Gadolinium, qui correspondent à des lieux d'hypofixation ou d'hyperfixation physiologiques du Sestamibi. Les globes oculaires, correspondant à des zones d'hypofixation du traceur, ont été segmentés à l'aide d'une méthode de Hough, tandis que les muscles oculomoteurs, qui correspondent à des zones d'hyperfixation du traceur, ont été segmentés en utilisant un algorithme de ligne de partage des eaux 3D. L'hypophyse, qui correspond également à une zone d'hyperfixation, a été segmentée en définissant une boîte englobante cylindrique.

La pertinence du critère d'évaluation UC a été validée sur un ensemble de 9 jeux de données en montrant que le recalage manuel obtenait toujours des valeurs de UC équivalentes ou supérieures à celles obtenues par les méthodes automatiques. Après visualisation par des experts des résultats sur 62 jeux de données, la stratégie consistant à sélectionner le meilleur recalage à l'aide de UC a permis d'obtenir 85% de recalages d'excellente qualité et seulement 3% de recalages non satisfaisants. Les 4 "meilleures" méthodes de la base de données utilisées individuellement obtiennent au mieux 61% de recalages d'excellente qualité et 13% de recalages de mauvaise qualité. La stratégie de recalage que nous avons mise en place dans le cadre du recalage TEMP/IRM permet donc une nette amélioration des résultats à l'échelle individuelle.

Dans le cadre du recalage monomodal longitudinal IRM-IRM, la stratégie de recalage a également été utilisée. Dix méthodes de recalage basées sur un modèle rigide et des critères iconiques classiques ou des critères de distance ont été testées. L'indice de Dice défini sur les structures précédemment segmentées : globes oculaires, muscles oculomoteurs et hypophyse, a permis d'évaluer les méthodes de recalage. Chacune des méthodes de recalage utilisant un critère iconique a obtenu des scores de Dice très proches sur l'ensemble des 15 jeux de données testés, suggérant une équivalence des méthodes utilisées pour ce problème. Cette équivalence a été confirmée par l'évaluation visuelle et par le fait que la valeur de TRE entre les méthodes était inférieure à la résolution du voxel.

La segmentation des zones tumorales sur les données IRM après injection de Gadolinium a été conduite par un expert. Pour la segmentation des zones tumorales sur les données TEMP, différentes méthodes de seuillage ont été appliquées, conduisant à autant de résultats de segmentation. A partir de ces segmentations les variations de volume et d'intensité au sein des zones tumorales ainsi que les indicateurs locaux ont été calculés sur les données TEMP précoces et tardives et sur les données IRM. Les critères locaux ont également été mis en œuvre pour comparer les données TEMP et les données IRM.

L'ensemble de ces indices a été comparé aux durées de survie sans aggravation (PFS) et aux durées de survie globale (OS). Aucune corrélation n'a été trouvée entre les durées de survie et les indices globaux et locaux extraits des volumes tumoraux estimés sur les IRM après injection de Gadolinium. Bien que les segmentations des zones tumorales sur les données TEMP soient assez différentes en fonction des seuils utilisés, nous avons pu mettre en évidence un certain nombre de résultats. Quelle que soit la méthode de segmentation utilisée, aucune corrélation n'a été trouvée entre les durées de survie PFS et OS, et les indices globaux et locaux de variations de volume. En revanche, les variations relatives d'intensité moyenne ou maximale au sein de la zone tumorale détectée sur les données TEMP sont assez similaires quelle que soit la méthode de segmentation considérée et sont de plus corrélées à la durée de survie globale OS. En pratique, ces variations d'intensité maximale ne nécessitent pas de segmentation mais uniquement la définition d'une boîte englobante autour de la zone tumorale. Ces variations mesurées sur les examens TEMP tardifs constituent donc un indice robuste et facile à calculer.

Lorsque la barrière hémato-encéphalique est reconstruite suite à un traitement antiangiogénique, il a été décrit que le Gadolinium ne peut plus accéder à la zone tumorale. La forte corrélation entre volumes tumoraux détectés sur les TEMP et les IRM suggère que le Sestamibi est confronté au même problème, cependant notre étude montre que la corrélation entre les volumes tumoraux détectés sur les IRM est plus forte avec les volumes extraits des images TEMP précoces qu'avec ceux extraits des images TEMP tardives. Pour expliquer ce résultat, notre hypothèse est que les examens TEMP précoces reflètent davantage un mécanisme de perfusion, alors que les examens TEMP tardifs pourraient être plus liés à l'activité mitochondriale augmentée dans les régions tumorales.

Ainsi la TEMP tardive semble apporter une information complémentaire à l'IRM de perfusion, puisque les variations d'intensité maximale détectées sur les données TEMP tardives sont prédictives de la durée de survie globale. En revanche, cette étude ne permet pas de conclure sur l'efficacité du traitement antiangiogénique puisque l'étude qui a été menée n'était pas randomisée.

La stratégie de recalage que nous avons établie au cours de cette thèse a également été appliquée à une seconde étude menée au LIB concernant le suivi de patients atteints de lésions hépatiques sous traitement antiangiogénique à l'aide de l'imagerie par tomodensitométrie. Pour réaliser ce suivi, deux types de recalage ont été nécessaires :

- le recalage longitudinal des données TDM ;
- une correction du mouvement au sein des séquences dynamiques TDM.

Pour le recalage longitudinal, nous nous sommes intéressés à la phase d'initialisation du processus à l'aide de méthodes de recalage rigide, en vue d'un recalage ultérieur non rigide. Dans ce but, trois types de méthodes basées sur l'information mutuelle, l'information mutuelle seuillée et la distance de chanfrein ont été utilisées. Afin de sélectionner la meilleure méthode pour chaque jeu de données, l'indice de Dice a été défini sur les structures osseuses (côtes et vertèbres).

Après analyse, la méthode basée sur l'information mutuelle seuillée a été identifiée comme étant la méthode obtenant le meilleur taux de succès soit 68% des cas. En sélectionnant pour chaque jeu de données, la "meilleure" méthode, la stratégie de recalage permet d'obtenir 89% de recalages de bonne qualité. Ainsi, notre approche s'avère pertinente pour la problématique posée. Cependant des difficultés importantes subsistent pour réussir parfaitement cette étape de recalage longitudinal. Elles sont liées notamment au faible nombre de coupes mis en correspondance et à la non isotropie des données ce qui rend difficile la mise en œuvre de méthodes non rigides.

Pour le recalage dynamique des données TDM, deux méthodes de recalage basées sur l'information mutuelle ont été appliquées, l'une utilisant un modèle de transformation rigide, l'autre un modèle non rigide faisant appel à l'algorithme FFD (Free Form Deformation). Afin d'évaluer quelle était la meilleure méthode pour chaque jeu de données, le même indice que pour l'application du recalage longitudinal a été utilisé. La méthode de recalage rigide s'est avérée non adéquate, alors que la méthode basée sur la FFD a permis d'obtenir des recalages de bonne qualité.

Perspectives

Nous avons montré, en nous appuyant sur deux applications cliniques différentes, que la stratégie de recalage proposée dans cette thèse permettait d'améliorer la qualité des recalages à l'échelle individuelle, par rapport à l'utilisation d'une seule méthode, si performante soit-elle. Cependant, avec l'utilisation de plusieurs méthodes, le temps de calcul nécessaire augmente. Dans la stratégie actuellement établie, nous utilisons pour chaque jeu de données, un nombre élevé de méthodes. Il serait intéressant à terme de limiter au maximum le nombre de méthodes de recalage à utiliser pour recaler les jeux de données. Idéalement il faudrait tester les méthodes les unes après les autres jusqu'à obtention d'un résultat de recalage satisfaisant.

Pour pouvoir appliquer ce processus, il faudrait en premier lieu normaliser le critère d'évaluation de manière à établir une correspondance absolue entre la valeur de ce critère et la qualité du recalage obtenu. En pratique, une telle normalisation sur le critère UC a été tentée mais n'a pas donné de résultats satisfaisants. Cette approche est plus prometteuse pour le critère développé dans la seconde application pour laquelle deux valeurs seuils du critère ont été établies : une valeur de seuil haute au delà de laquelle un recalage est considéré comme étant de bonne qualité et une valeur de seuil basse en dessous de laquelle un recalage est considéré de qualité insatisfaisante. Dans le cas où le critère d'évaluation obtient une valeur entre ces deux seuils une évaluation visuelle est requise pour déterminer la qualité du recalage.

La deuxième étape de ce processus concerne l'ordre dans lequel utiliser les méthodes. Il semble logique d'utiliser la méthode qui obtient le taux le plus élevé de résultats d'excellente qualité puis les suivantes en fonction de leur taux de réussite. Une autre stratégie serait d'utiliser en priorité les plus rapides. On peut même envisager d'initialiser le recalage avec des méthodes rapides comme celles basées sur la distance de chanfrein qui ne prennent que quelques secondes pour obtenir un résultat. Il serait aussi pertinent de tester en priorité des méthodes qui soient complémentaires, c'est à dire des méthodes qui ont tendance à obtenir des résultats satisfaisants pour des jeux de données différents. En effet, l'utilisation de méthodes qui obtiendraient des taux de réussite élevés mais pour les mêmes cas n'aurait qu'un intérêt limité dans le contexte de la stratégie. Ces premières pistes de réflexion montrent que l'optimisation de la stratégie n'est donc pas un problème trivial. Enfin quelle que soit la stratégie choisie, une partie d'apprentissage pour déterminer les taux de réussite et les temps de calcul de chacune des méthodes est nécessaire pour déterminer l'ordre optimal dans lequel les tester. Une telle approche prend tout son sens dans le cadre d'applications impliquant un nombre de cas à traiter conséquents.

La définition d'une nouvelle méthode de recalage adaptée reste également une alternative qui pourrait apparaître plus simple que la proposition précédente. Cependant, en pratique, la réussite d'une telle méthode dépendra fortement de la forme de la fonction de coût. Si une fonction de coût possède de nombreux optimum locaux, le processus de recalage ne sera pas forcément pertinent.

Nous pensons que la partie stratégie de recalage que nous avons proposée pourrait être utilisée dans d'autres applications cliniques.

Par exemple, pour confirmer le diagnostic de la maladie de Parkinson, l'imagerie DaTSCAN utilisant comme radiotraceur l'ioflupane marqué à l'iode 123 qui se fixe de manière spécifique sur les transporteurs

présynaptique de la dopamine (DAT), peut être utilisée en complément de l'IRM. Chez les sujets sains le traceur se fixe sur les putamens et les noyaux caudés alors que chez des patients atteints de Parkinson, le traceur a tendance à moins se fixer sur les putamens. Donc la localisation du traceur dans des structures spécifiques est modifiée avec la pathologie. Dans le cadre d'une collaboration avec le Centre d'Acquisition et de Traitement automatisé de l'Image (CATI) pour le recalage multimodal de données TEMP(DATSCAN)/IRM pour des patients atteints de la maladie Parkinson, nous avons proposé d'utiliser la stratégie de recalage décrite dans cette thèse. Mickaël Labit, ingénieur du CATI travaille actuellement à sa mise en œuvre. A la différence de l'exemple des tumeurs où on s'est appuyé sur des régions qui restaient normales, la difficulté réside, dans la définition d'un critère robuste permettant de classer les différentes méthodes de recalage.

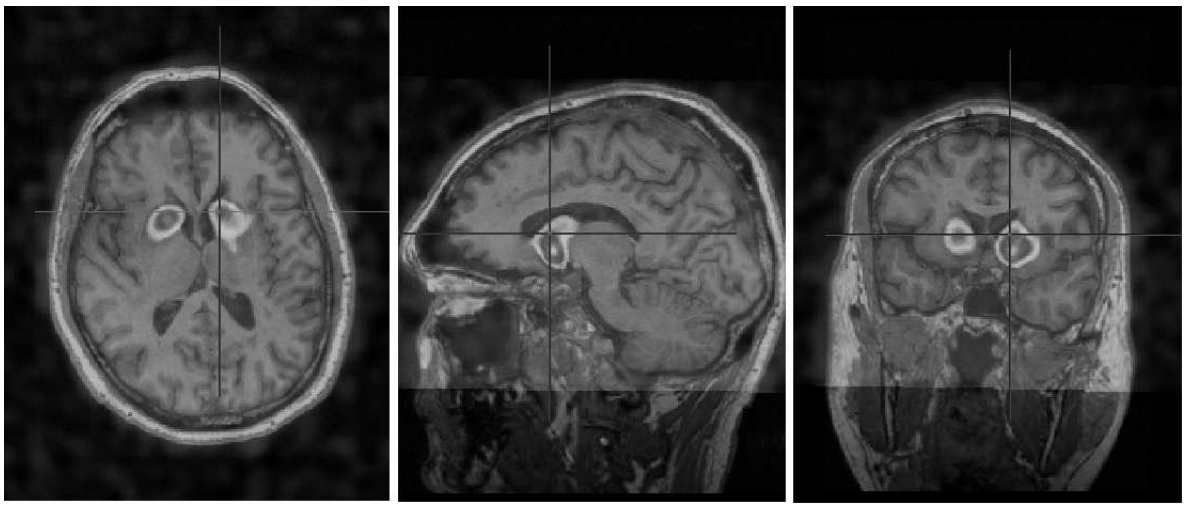


FIGURE 7.8: Exemple d'un recalage de bonne qualité de données DaTSCAN/IRM d'un patient obtenu par la stratégie de recalage.

Il est également important de poursuivre l'analyse des données de type TDM dynamiques pour des patients atteints de lésions hépatiques. La segmentation tumorale n'a pas encore été abordée dans cette étude. Cependant, l'utilisation d'approches de type "GVF snake" [97] peut permettre d'extraire les lésions hépatiques. Il faudait aussi exploiter l'information de perfusion dans les tissus pour définir de nouveaux critères d'évaluation tumorale plus spécifiques que les simples variations de volumes ou d'intensité des zones tumorales au cours du traitement. Des applications similaires en IRM dynamique sont également en cours de développement dans l'équipe.

Bibliographie

[1] D. N. Louis, H. Ohgaki, O. D. Wiestler, W. K. Cavenee, P. C. Burger, A. Jouvet, B. W. Scheithauer, and P. Kleihues, "The 2007 WHO classification of tumours of the central nervous system.," *Acta neuropathologica*, vol. 114, pp. 97–109, Aug. 2007.

[2] CBTRUS, "Statistical report : primary brain tumors in the United States, 2007-2008 Chicago : Central Brain Tumor Registry of the United States," tech. rep., 2008.

[3] D. W. Parsons, S. Jones, X. Zhang, J. C.-H. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, I.-M. Siu, G. L. Gallia, A. Olivi, R. McLendon, B. A. Rasheed, S. Keir, T. Nikolskaya, Y. Nikolsky, D. A. Busam, H. Tekleab, L. A. Diaz, J. Hartigan, D. R. Smith, R. L. Strausberg, S. K. N. Marie, S. M. O. Shinjo, H. Yan, G. J. Riggins, D. D. Bigner, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V. E. Velculescu, and K. W. Kinzler, "An integrated genomic analysis of human glioblastoma multiforme.," *Science (New York, N.Y.)*, vol. 321, pp. 1807–12, Sept. 2008.

[4] H. Yan, D. W. Parsons, G. Jin, R. McLendon, B. A. Rasheed, W. Yuan, I. Kos, I. Batinic-Haberle, S. Jones, G. J. Riggins, H. Friedman, A. Friedman, D. Reardon, J. Herndon, K. W. Kinzler, V. E. Velculescu, B. Vogelstein, and D. D. Bigner, "IDH1 and IDH2 mutations in gliomas.," *The New England journal of medicine*, vol. 360, pp. 765–73, Feb. 2009.

[5] H. Ohgaki, "Genetic pathways to glioblastomas.," *Neuropathology : official journal of the Japanese Society of Neuropathology*, vol. 25, pp. 1–7, Mar. 2005.

[6] K. R. Hess, K. R. Broglio, and M. L. Bondy, "Adult glioma incidence trends in the United States, 1977-2000.," *Cancer*, vol. 101, pp. 2293–9, Nov. 2004.

[7] M. Wrensch, Y. Minn, T. Chew, M. Bondy, and M. S. Berger, "Epidemiology of primary brain tumors : current concepts and review of the literature.," *Neuro-oncology*, vol. 4, pp. 278–99, Oct. 2002.

[8] I. Deltour, A. Auvinen, M. Feychting, C. Johansen, L. Klaeboe, R. Sankila, and J. Schüz, "Mobile phone use and incidence of glioma in the Nordic countries 1979-2008 : consistency check.," *Epidemiology (Cambridge, Mass.)*, vol. 23, pp. 301–7, Mar. 2012.

[9] M. Z. Braganza, P. Rajaraman, Y. Park, P. D. Inskip, N. D. Freedman, A. R. Hollenbeck, A. B. de González, and C. M. Kitahara, "Cigarette smoking, alcohol intake, and risk of glioma in the NIH-AARP Diet and Health Study.," *British journal of cancer*, vol. 110, pp. 242–8, Jan. 2014.

[10] J. L. Fisher, J. A. Schwartzbaum, M. Wrensch, and J. L. Wiemels, "Epidemiology of brain tumors.," *Neurologic clinics*, vol. 25, pp. 867–90, vii, Nov. 2007.

[11] M. Lacroix, D. Abi-Said, D. R. Fourney, Z. L. Gokaslan, W. Shi, F. DeMonte, F. F. Lang, I. E. McCutcheon, S. J. Hassenbusch, E. Holland, K. Hess, C. Michael, D. Miller, and R. Sawaya, "A multivariate analysis of 416 patients with glioblastoma multiforme : prognosis, extent of resection, and survival.," *Journal of neurosurgery*, vol. 95, pp. 190–8, Aug. 2001.

[12] J. Simpson, J. Horton, C. Scott, W. Curran, P. Rubin, J. Fischbach, S. Isaacson, M. Rotman, S. Asbell, J. Nelson, A. Weinstein, and D. Nelson, "Influence of location and extent of surgical resection on survival of patients with glioblastoma multiforme : Results of three consecutive radiation therapy oncology group (RTOG) clinical trials," *International Journal of Radiation Oncology*Biology*Physics*, vol. 26, pp. 239–244, May 1993.

[13] K. Kristiansen, S. Hagen, T. Kollevold, A. Torvik, I. Holme, R. Nesbakken, R. Hatlevoll, M. Lindgren, A. Brun, S. Lindgren, G. Notter, A. P. Andersen, and K. Elgen, "Combined modality therapy of operated astrocytomas grade III and IV. Confirmation of the value of postoperative irradiation and lack of potentiation of bleomycin on survival time : a prospective multicenter trial of the Scandinavian Glioblastoma Study Grou," *Cancer*, vol. 47, pp. 649–52, Feb. 1981.

[14] M. G. Hart, R. Garside, G. Rogers, K. Stein, and R. Grant, "Temozolomide for high grade glioma.," *The Cochrane database of systematic reviews*, vol. 4, p. CD007415, Jan. 2013.

[15] S. Parisi, P. Corsa, and A. Raguso, "Temozolomide and Radiotherapy versus Radiotherapy Alone in High Grade Gliomas : A Very Long Term Comparative Study and Literature Review," *BioMed Research International*, 2014.

[16] N. Ahmadloo, A.-A. Kani, M. Mohammadianpanah, H. Nasrolahi, S. Omidvari, A. Mosalaei, and M. Ansari, "Treatment outcome and prognostic factors of adult glioblastoma multiforme.," *Journal of the Egyptian National Cancer Institute*, vol. 25, pp. 21–30, Mar. 2013.

[17] H. Ohgaki, P. Dessen, B. Jourde, S. Horstmann, T. Nishikawa, P.-L. Di Patre, C. Burkhard, D. Schüler, N. M. Probst-Hensch, P. C. Maiorka, N. Baeza, P. Pisani, Y. Yonekawa, M. G. Yasargil, U. M. Lütolf, and P. Kleihues, "Genetic pathways to glioblastoma : a population-based study.," *Cancer research*, vol. 64, pp. 6892–9, Oct. 2004.

[18] J. Bolcaen, B. Descamps, K. Deblaere, T. Boterberg, F. De Vos Pharm, J.-P. Kalala, C. Van den Broecke, E. Decrock, L. Leybaert, C. Vanhove, and I. Goethals, "(18)F-fluoromethylcholine (FCho), (18)F-fluoroethyltyrosine (FET), and (18)F-fluorodeoxyglucose (FDG) for the discrimination between high-grade glioma and radiation necrosis in rats : A PET study.," *Nuclear medicine and biology*, vol. 42, pp. 38–45, Jan. 2015.

[19] C. Soler, P. Beauchesne, K. Maatougui, T. Schmitt, F. G. Barral, D. Michel, F. Dubois, and J. Brunon, "Technetium-99m sestamibi brain single-photon emission tomography for detection of recurrent gliomas after radiation therapy," *Eur. J. Nucl. Med.*, vol. 25, no. 12, pp. 1649–1657, 1998.

[20] F. Prigent-Le Jeune, F. Dubois, S. Perez, S. Blond, and M. Steinling, "Technetium-99m sestamibi brain SPECT in the follow-up of glioma for evaluation of response to chemotherapy : first results.," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 31, pp. 714–9, May 2004.

[21] S. Bleichner-Perez, F. Le Jeune, F. Dubois, and M. Steinling, "99mTc-MIBI brain SPECT as an indicator of the chemotherapy response of recurrent, primary brain tumors.," *Nucl. Med. Commun.*, vol. 28, pp. 888–94, Dec. 2007.

[22] N. Yeni, G. Petrirena, J.-M. Tacchella, R. Guillevin, J.-B. Martini, G. Alain, M.-O. Habert, J.-Y. Delattre, and A. Kas, "Early assessment of bevacizumab response using 99mTc-sestamibi single-photon emission tomography in recurrent high grade glioma.," *J. Nucl. Med.*

[23] J. Gállego Pérez-Larraya, M. Lahutte, G. Petrirena, G. Reyes-Botero, A. González-Aguilar, C. Houillier, R. Guillevin, M. Sanson, K. Hoang-Xuan, and J.-Y. Delattre, "Response assessment in recurrent glioblastoma treated with irinotecan-bevacizumab : comparative analysis of the Macdonald, RECIST, RANO, and RECIST + F criteria.," *Neuro-oncology*, vol. 14, pp. 667–73, May 2012.

[24] A. B. Miller, B. Hoogstraten, M. Staquet, and A. Winkler, "Reporting results of cancer treatment.," *Cancer*, vol. 47, pp. 207–14, Jan. 1981.

[25] P. Therasse, S. G. Arbuck, E. A. Eisenhauer, J. Wanders, R. S. Kaplan, L. Rubinstein, J. Verweij, M. V. Glabbeke, T. V. Oosterom, M. C. Christian, and S. G. Gwyther, "New Guidelines to Evaluate the Response to Treatment," *J Natl Cancer Inst*, vol. 92, pp. 205–216, Feb. 2000.

[26] E. A. Eisenhauer, P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, and J. Verweij, "New response evaluation criteria in solid tumours : revised RECIST guideline (version 1.1).," *Eur. J Cancer*, vol. 45, pp. 228–47, Jan. 2009.

[27] D. R. Macdonald, T. L. Cascino, S. C. Schold, and J. G. Cairncross, "Response criteria for phase II studies of supratentorial malignant glioma.," *J. Clin. Oncol.*, vol. 8, pp. 1277–80, July 1990.

[28] P. Y. Wen, D. R. Macdonald, D. A. Reardon, T. F. Cloughesy, A. G. Sorensen, E. Galanis, J. Degroot, W. Wick, M. R. Gilbert, A. B. Lassman, C. Tsien, T. Mikkelsen, E. T. Wong, M. C. Chamberlain, R. Stupp, K. R. Lamborn, M. A. Vogelbaum, M. J. van den Bent, and S. M. Chang, "Updated response assessment criteria for high-grade gliomas : response assessment in neuro-oncology working group.," *J. Clin. Oncol.*, vol. 28, pp. 1963–72, Apr. 2010.

[29] L. C. Hygino da Cruz, I. Rodriguez, R. C. Domingues, E. L. Gasparetto, and A. G. Sorensen, "Pseudoprogression and pseudoresponse : imaging challenges in the assessment of posttreatment glioma.," *AJNR. American journal of neuroradiology*, vol. 32, pp. 1978–85, Dec. 2011.

[30] B. A. Moffat, T. L. Chenevert, T. S. Lawrence, C. R. Meyer, T. D. Johnson, Q. Dong, C. Tsien, S. Mukherji, D. J. Quint, S. S. Gebarski, P. L. Robertson, L. R. Junck, A. Rehemtulla, and B. D. Ross, "Functional diffusion map : a noninvasive MRI biomarker for early stratification of clinical brain tumor response.," *PNAS*, vol. 102, pp. 5524–9, Apr. 2005.

[31] P. M. Wilson, M. J. LaBonte, and H.-J. Lenz, "Assessing the in vivo efficacy of biologic antiangiogenic therapies.," *Cancer chemotherapy and pharmacology*, vol. 71, pp. 1–12, Jan. 2013.

[32] J. Edeline, E. Boucher, Y. Rolland, E. Vauléon, M. Pracht, C. Perrin, C. Le Roux, and J.-L. Raoul, "Comparison of tumor response by Response Evaluation Criteria in Solid Tumors (RECIST) and

modified RECIST in patients treated with sorafenib for hepatocellular carcinoma.," *Cancer*, vol. 118, pp. 147–56, Jan. 2012.

[33] O. Camara, G. Delso, and I. Bloch, "Evaluation of a thoracic elastic registration method using anatomical constraints in oncology," in *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology*, vol. 2, pp. 1011–1012, IEEE, 2002.

[34] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, pp. 325–376, Dec. 1992.

[35] B. Zitová and J. Flusser, "Image registration methods : a survey," *Image Vis. Comput.*, vol. 21, pp. 977–1000, Oct. 2003.

[36] M. Holden, "A review of geometric transformations for nonrigid body registration.," *IEEE Trans. Med. Imag.*, vol. 27, pp. 111–28, Jan. 2008.

[37] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Imag. Anal.*, vol. 2, no. 1, pp. 1–36, 1998.

[38] H. Lester and S. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognit.*, vol. 32, no. 1, pp. 129–149, 1999.

[39] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration.," *Phys. Med. Biol.*, vol. 46, pp. R1–45, Mar. 2001.

[40] P. Jannin, C. Grova, and B. Gibaud, "Fusion de données en imagerie médicale : revue méthodologique basée sur le contexte clinique," *ITBM-RBM*, vol. 22, no. 4, pp. 196–215, 2001.

[41] W. R. Crum, "Non-rigid image registration : theory and practice," *Br. J. Radiol.*, vol. 77, pp. S140–S153, Dec. 2004.

[42] A. Gholipour, N. Kehtarnavaz, R. Briggs, M. Devous, and K. Gopinath, "Brain functional localization : a survey of image registration techniques.," *IEEE Trans. Med. Imag.*, vol. 26, pp. 427–51, Apr. 2007.

[43] F. P. M. Oliveira and J. a. M. R. S. Tavares, "Medical image registration : a review.," *Comput. Methods Biomech. Biomed. Engin.*, vol. 17, pp. 1–21, Mar. 2012.

[44] O. Commowick, *"Création et utilisation d'atlas anatomiques numériques pour la radiothérapie"*. Thèse informatique, Université de Nice - Sophia Antipolis, 2007.

[45] M. Rubeaux, *"Approximation de l'Information Mutuelle basée sur le développement d'Edgeworth : application au recalage d'images médicales."*. Thèse traitement du signal et télécommunications, Université de Rennes, 2011.

[46] P. Besl and H. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[47] G. Borgefors, "Hierarchical chamfer matching : a parametric edge matching algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 6, pp. 849–865, 1988.

[48] J. Mangin, V. Frouin, I. Bloch, B. Bendriem, and J. Lopez-Krahe, "Fast nonsupervised 3D registration of PET and MR images of the brain," *J. Cereb. Blood Flow Metab.*, vol. 14, no. 5, pp. 749–762, 1994.

[49] M. Svedlow, C. McGillem, and P. Anuta, "Experimental examination of similarity measures and preprocessing methods used for image registration," *LARS Symposia*, 1976.

[50] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, "MRI-PET registration with automated algorithm.," *Journal of computer assisted tomography*, vol. 17, no. 4, pp. 536–46, 1993.

[51] G. M. Alexis Roche, "The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration," *Medical Image Computing and Computer-Assisted Intervention : MICCAI*, vol. 1496, pp. 1115–1124, 1998.

[52] C. Nikou, F. Heitz, I.-j. Namers, J.-p. Armspach, and B. S. Brant, "Single and multimodal subvoxel registration of dissimilar medical images using robust similarity measures," vol. 3338, no. February, 1998.

[53] N. M. Alpert, D. Berdichevsky, Z. Levin, E. D. Morris, and A. J. Fischman, "Improved methods for image registration.," *NeuroImage*, vol. 3, pp. 10–8, Feb. 1996.

[54] D. Hill, "Voxel similarity measures for automated image registration," *Visualization in Biomedical Computing*, vol. 2359, 1994.

[55] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multimodality image registration based on information theory," *Inf. Process. Med. Imag.*, vol. 3, pp. 263 – 274, 1995.

[56] P. Viola and W. M. W. Iii, "Alignment by Maximization of Mutual Information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, 1997.

[57] C. Studholme, D. J. Hawkes, and D. L. Hill, "Normalized entropy measure for multimodality image alignment," in *Medical Imaging '98* (K. M. Hanson, ed.), pp. 132–143, International Society for Optics and Photonics, June 1998.

[58] E. Parzen, "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, vol. 33, pp. 1065–1076, Sept. 1962.

[59] J. P. Pluim, J. B. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information.," *IEEE transactions on medical imaging*, vol. 19, pp. 809–14, Aug. 2000.

[60] S. Bullich, D. Ros, J. Pavía, A. Cot, N. López, and A. M. Catafau, "Neurotransmission SPECT and MR registration combining mutual and gradient information.," *Med. Phys.*, vol. 36, pp. 4903–10, Nov. 2009.

[61] R. Gan, A. C. S. Chung, and S. Liao, "Maximum distance-gradient for robust image registration.," *Medical image analysis*, vol. 12, pp. 452–68, Aug. 2008.

[62] J.-D. Lee, C.-H. Huang, Y.-H. Weng, K.-J. Lin, and C.-T. Chen, "An automatic MRI/SPECT registration algorithm using image intensity and anatomical feature as matching characters : application on the evaluation of Parkinson's disease.," *Nucl. Med. Biol.*, vol. 34, pp. 447–57, May 2007.

[63] D. Rueckert, M. J. Clarkson, D. L. G. Hill, and D. J. Hawkes, "Non-rigid registration using high-order mutual information," in *Medical Imaging 2000* (K. M. Hanson, ed.), pp. 438–447, International Society for Optics and Photonics, Mar. 2000.

[64] C. Studholme, C. Drapaca, B. Iordanova, and V. Cardenas, "Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change.," *IEEE transactions on medical imaging*, vol. 25, pp. 626–39, May 2006.

[65] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens, "Nonrigid image registration using conditional mutual information.," *IEEE transactions on medical imaging*, vol. 29, pp. 19–29, Jan. 2010.

[66] G. K. Rohde, A. Aldroubi, and B. M. Dawant, "The adaptive bases algorithm for intensity-based nonrigid image registration.," *IEEE transactions on medical imaging*, vol. 22, pp. 1470–9, Nov. 2003.

[67] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations : application to breast MR images.," *IEEE transactions on medical imaging*, vol. 18, pp. 712–21, Aug. 1999.

[68] B. K. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, Aug. 1981.

[69] C. Broit, "Optimal registration of deformed images," *Phd thesis*, Jan. 1981.

[70] G. Christensen, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.

[71] W. . Teukolsky, W., Vetterling, S., Flannery, *Numerical Recipes in C.* Cambridge University Press : Press, Cambridge University, second ed. ed., 1992.

[72] S. Klein, M. Staring, and J. P. W. Pluim, "Evaluation of Optimization Methods for Nonrigid Medical Image Registration Using Mutual Information and B-Splines," *IEEE Transactions on Image Processing*, vol. 16, pp. 2879–2890, Dec. 2007.

[73] S. Klein, J. P. W. Pluim, M. Staring, and M. a. Viergever, "Adaptive Stochastic Gradient Descent Optimisation for Image Registration," *International Journal of Computer Vision*, vol. 81, pp. 227–239, Aug. 2008.

[74] G. Christensen, X. Geng, and J. Kuhl, "Introduction to the non-rigid image registration evaluation project (NIREP)," *Bio. Imag. Reg.*, pp. 128–135, 2006.

[75] A. Klein, J. Andersson, B. a. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration.," *NeuroImage*, vol. 46, pp. 786–802, July 2009.

[76] K. K. Brock, "Results of a multi-institution deformable registration accuracy study (MIDRAS).," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 76, pp. 583–96, Feb. 2010.

[77] K. Murphy, B. van Ginneken, J. M. Reinhardt, S. Kabus, K. Ding, X. Deng, K. Cao, K. Du, G. E. Christensen, V. Garcia, T. Vercauteren, N. Ayache, O. Commowick, G. Malandain, B. Glocker, N. Paragios, N. Navab, V. Gorbunova, J. Sporring, M. de Bruijne, X. Han, M. P. Heinrich, J. a. Schnabel, M. Jenkinson, C. Lorenz, M. Modat, J. R. McClelland, S. Ourselin, S. E. a. Muenzing, M. a. Viergever, D. De Nigris, D. L. Collins, T. Arbel, M. Peroni, R. Li, G. C. Sharp, A. Schmidt-Richberg, J. Ehrhardt,

R. Werner, D. Smeets, D. Loeckx, G. Song, N. Tustison, B. Avants, J. C. Gee, M. Staring, S. Klein, B. C. Stoel, M. Urschler, M. Werlberger, J. Vandemeulebroucke, S. Rit, D. Sarrut, and J. P. W. Pluim, "Evaluation of registration methods on thoracic CT : the EMPIRE10 challenge.," *IEEE Trans. Med. Imag.*, vol. 30, pp. 1901–20, Nov. 2011.

[78] B. B. Avants, N. J. Tustison, G. Song, P. a. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ANTs similarity metric performance in brain image registration.," *NeuroImage*, vol. 54, pp. 2033–44, Feb. 2011.

[79] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. Hill, D. J. Hawkes, C. Studholme, J. B. Maintz, M. A. Viergever, G. Malandain, and R. P. Woods, "Comparison and evaluation of retrospective intermodality brain image registration techniques.," *Journal of computer assisted tomography*, vol. 21, no. 4, pp. 554–66, 1997.

[80] G. Loi, M. Dominietto, I. Manfredda, E. Mones, A. Carriero, E. Inglese, M. Krengli, and M. Brambilla, "Acceptance test of a commercially available software for automatic image registration of computed tomography (CT), magnetic resonance imaging (MRI) and 99mTc-methoxyisobutylisonitrile (MIBI) single-photon emission computed tomography (SPECT) brain images.," *Journal of digital imaging : the official journal of the Society for Computer Applications in Radiology*, vol. 21, pp. 329–37, Sept. 2008.

[81] D. Skerl, B. Likar, and F. Pernus, "A protocol for evaluation of similarity measures for rigid registration.," *IEEE transactions on medical imaging*, vol. 25, pp. 42–54, Feb. 2006.

[82] D. Skerl, B. Likar, and F. Pernus, "A protocol for evaluation of similarity measures for non-rigid registration.," *Medical image analysis*, vol. 12, pp. 42–54, Feb. 2008.

[83] S. Klein, M. Staring, and J. P. W. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines.," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 16, pp. 2879–90, Dec. 2007.

[84] L. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297 – 302, 1945.

[85] P. Jaccard, "Étude comparative de la distribution florale dans une portion des Alpes et des Jura," *Bulletin del la Société Vaudoise des Sciences Naturelles*, vol. 37, pp. 547 – 579, 1901.

[86] B. Aubert-Broche, C. Grova, A. Reilhac, A. C. Evans, and D. L. Collins, "Realistic simulated MRI and SPECT databases. Application to SPECT/MRI registration evaluation.," *Med. Image Comput. Comput. Assist. Interv.*, vol. 9, pp. 330–7, Jan. 2006.

[87] N. Kirby, C. Chuang, U. Ueda, and J. Pouliot, "The need for application-based adaptation of deformable image registration.," *Medical physics*, vol. 40, p. 011702, Jan. 2013.

[88] K. Murphy, B. van Ginneken, S. Klein, M. Staring, B. J. de Hoop, M. a. Viergever, and J. P. W. Pluim, "Semi-automatic construction of reference standards for evaluation of image registration.," *Med. Imag. Anal.*, vol. 15, pp. 71–84, Feb. 2011.

[89] T. Glatard, X. Pennec, and J. Montagnat, "Performance evaluation of grid-enabled registration algorithms using bronze-standards.," *Medical image computing and computer-assisted intervention : MICCAI 2006 International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 9, pp. 152–60, Jan. 2006.

[90] D. Pham, C. Xu, and J. Prince, "Current methods in medical image segmentation," *Annu Rev Biomed Eng*, pp. 315–37, 2000.

[91] E. Angelini, O. Clatz, E. Mandonnet, E. Konukoglu, L. Capelle, and H. Duffau, "Glioma Dynamics and Computational Models : A Review of Segmentation, Registration, and In Silico Growth Algorithms and their Clinical Applications," *Curr. Med. Imaging Rev.*, vol. 3, pp. 262–276, Nov. 2007.

[92] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies.," *Physics in medicine and biology*, vol. 58, pp. R97–129, July 2013.

[93] a. S. Dewalle-Vignion, a. El Abiad, N. Betrouni, C. Hossein-Foucher, D. Huglo, and M. Vermandel, "Les méthodes de seuillage en TEP : un état de l'art," *Médecine Nucléaire*, vol. 34, pp. 119–131, Feb. 2010.

[94] P. Sahoo, S. Soltani, and A. Wong, "A survey of thresholding techniques," *Computer Vision, Graphics, and Image Processing*, vol. 41, pp. 233–260, Feb. 1988.

[95] I. N. Manousakas, P. E. Undrill, G. G. Cameron, and T. W. Redpath, "Split-and-merge segmentation of magnetic resonance medical images : performance evaluation and extension to three dimensions.," *Computers and biomedical research, an international journal*, vol. 31, pp. 393–412, Dec. 1998.

[96] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis : a survey.," *Medical image analysis*, vol. 1, pp. 91–108, June 1996.

[97] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow.," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 7, pp. 359–69, Jan. 1998.

[98] R. Malladi, J. Sethian, and B. Vemuri, "Shape modeling with front propagation : a level set approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.

[99] J. C. Bezdek, L. O. Hall, and L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition.," *Medical physics*, vol. 20, no. 4, pp. 1033–48.

[100] G. Coleman and H. Andrews, "Image segmentation by clustering," *Proceedings of the IEEE*, vol. 67, pp. 773–785, May 1979.

[101] T. Lei and W. Sewchand, "Statistical approach to X-ray CT imaging and its applications in image analysis. II. A new stochastic model-based image segmentation technique for X-ray CT image.," *IEEE transactions on medical imaging*, vol. 11, pp. 62–9, Jan. 1992.

[102] Z. Liang, J. R. Macfall, and D. P. Harrington, "Parameter estimation and tissue segmentation from multispectral MR images.," *IEEE transactions on medical imaging*, vol. 13, pp. 441–9, Jan. 1994.

[103] J. C. Rajapakse, J. N. Giedd, and J. L. Rapoport, "Statistical approach to segmentation of single-channel cerebral MR images.," *IEEE transactions on medical imaging*, vol. 16, pp. 176–86, Apr. 1997.

[104] L. Hall and A. Bensaid, "A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain," *IEEE Trans Neural Netw*, vol. 41, pp. 672–82, 1992.

[105] P. Gibbs, D. L. Buckley, S. J. Blackband, and A. Horsman, "Tumour volume determination from MR images by morphological segmentation.," *Physics in medicine and biology*, vol. 41, pp. 2437–46, Nov. 1996.

[106] M. M. J. Letteboer, O. F. Olsen, E. B. Dam, P. W. A. Willems, M. A. Viergever, and W. J. Niessen, "Segmentation of tumors in magnetic resonance brain images using an interactive multiscale watershed algorithm.," *Academic radiology*, vol. 11, pp. 1125–38, Oct. 2004.

[107] M. Droske, B. Meyer, M. Rumpf, and C. Schaller, "An adaptive level set method for interactive segmentation of intracranial tumors.," *Neurological research*, vol. 27, pp. 363–70, June 2005.

[108] T. Wang, I. Cheng, and A. Basu, "Fluid vector flow and applications in brain tumor segmentation.," *IEEE transactions on bio-medical engineering*, vol. 56, pp. 781–9, Mar. 2009.

[109] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "A novel content-based active contour model for brain tumor segmentation.," *Magnetic resonance imaging*, vol. 30, pp. 694–715, June 2012.

[110] L. R. Schad, S. Blüml, and I. Zuna, "MR tissue characterization of intracranial tumors by means of texture analysis.," *Magnetic resonance imaging*, vol. 11, pp. 889–96, Jan. 1993.

[111] W. E. Phillips, R. P. Velthuizen, S. Phuphanich, L. O. Hall, L. P. Clarke, and M. L. Silbiger, "Application of fuzzy c-means segmentation technique for tissue differentiation in MR images of a hemorrhagic glioblastoma multiforme.," *Magnetic resonance imaging*, vol. 13, pp. 277–90, Jan. 1995.

[112] M. Vaidyanathan, L. P. Clarke, R. P. Velthuizen, S. Phuphanich, A. M. Bensaid, L. O. Hall, J. C. Bezdek, H. Greenberg, A. Trotti, and M. Silbiger, "Comparison of supervised MRI segmentation methods for tumor volume determination during therapy.," *Magnetic resonance imaging*, vol. 13, pp. 719–28, Jan. 1995.

[113] M. C. Clark, L. O. Hall, D. B. Goldgof, R. Velthuizen, F. R. Murtagh, and M. S. Silbiger, "Automatic tumor segmentation using knowledge-based techniques.," *IEEE transactions on medical imaging*, vol. 17, pp. 187–201, Apr. 1998.

[114] L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, and F. R. Murtagh, "Automatic segmentation of non-enhancing brain tumors in magnetic resonance images.," *Artificial intelligence in medicine*, vol. 21, no. 1-3, pp. 43–63, 2001.

[115] M. R. Kaus, S. K. Warfield, A. Nabavi, P. M. Black, F. A. Jolesz, and R. Kikinis, "Automated segmentation of MR images of brain tumors.," *Radiology*, vol. 218, pp. 586–91, Feb. 2001.

[116] N. Moon, E. Bullitt, K. V. Leemput, and G. Gerig, "Automatic brain and tumor segmentation," *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002*, pp. 372–379, 2002.

[117] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "Robust estimation for brain tumor segmentation," *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, pp. 530–537, 2003.

[118] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "A brain tumor segmentation framework based on outlier detection.," *Medical image analysis*, vol. 8, pp. 275–83, Sept. 2004.

[119] G. Moonis, J. Liu, J. K. Udupa, and D. B. Hackney, "Estimation of tumor volume with fuzzy-connectedness segmentation of MR images.," *AJNR. American journal of neuroradiology*, vol. 23, pp. 356–63, Mar. 2002.

[120] J. Liu, J. K. Udupa, D. Odhner, D. Hackney, and G. Moonis, "A system for brain tumor volume estimation via MR imaging and fuzzy connectedness.," *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, vol. 29, pp. 21–34, Jan. 2005.

[121] A. S. Bryant and R. J. Cerfolio, "The maximum standardized uptake values on integrated FDG-PET/CT is useful in differentiating benign from malignant pulmonary nodules.," *The Annals of thoracic surgery*, vol. 82, pp. 1016–20, Sept. 2006.

[122] K. R. Zasadny, P. V. Kison, I. R. Francis, and R. L. Wahl, "FDG-PET Determination of Metabolically Active Tumor Volume and Comparison with CT.," *Clinical positron imaging : official journal of the Institute for Clinical P.E.T*, vol. 1, pp. 123–129, Mar. 1998.

[123] T. R. Miller and P. W. Grigsby, "Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy.," *International journal of radiation oncology, biology, physics*, vol. 53, pp. 353–9, June 2002.

[124] Y. E. Erdi, O. Mawlawi, S. M. Larson, M. Imbriaco, H. Yeung, R. Finn, and J. L. Humm, "Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding.," *Cancer*, vol. 80, pp. 2505–9, Dec. 1997.

[125] W. Jentzen, L. Freudenberg, E. G. Eising, M. Heinze, W. Brandau, and A. Bockisch, "Segmentation of PET volumes by iterative image thresholding.," *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, vol. 48, pp. 108–14, Jan. 2007.

[126] J.-F. Daisne, M. Sibomana, A. Bol, T. Doumont, M. Lonneux, and V. Grégoire, "Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios : influence of reconstruction algorithms," *Radiotherapy and Oncology*, vol. 69, pp. 247–250, Dec. 2003.

[127] Q. C. Black, I. S. Grills, L. L. Kestin, C.-Y. O. Wong, J. W. Wong, A. A. Martinez, and D. Yan, "Defining a radiotherapy target with positron emission tomography.," *International journal of radiation oncology, biology, physics*, vol. 60, pp. 1272–82, Nov. 2004.

[128] U. Nestle, S. Kremp, A. Schaefer-Schuler, C. Sebastian-Welsch, D. Hellwig, C. Rübe, and C.-M. Kirsch, "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.," *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, vol. 46, pp. 1342–8, Aug. 2005.

[129] U. Nestle, A. Schaefer-Schuler, S. Kremp, A. Groeschel, D. Hellwig, C. Rübe, and C.-M. Kirsch, "Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer.," *European journal of nuclear medicine and molecular imaging*, vol. 34, pp. 453–62, Apr. 2007.

[130] G. P. Mazzara, R. P. Velthuizen, J. L. Pearlman, H. M. Greenberg, and H. Wagner, "Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation.," *International journal of radiation oncology, biology, physics*, vol. 59, pp. 300–12, May 2004.

[131] N. Archip, F. A. Jolesz, and S. K. Warfield, "A validation framework for brain tumor segmentation.," *Academic radiology*, vol. 14, pp. 1242–51, Oct. 2007.

[132] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE) : an algorithm for the validation of image segmentation.," *IEEE transactions on medical imaging*, vol. 23, pp. 903–21, July 2004.

[133] C. Cocosco, V. Kollokian, Kwan, and A. Evans, "BrainWeb : Online Interface to a 3D MRI Simulated Brain Database," *NeuroImage*, vol. 5, no. 4, 1997.

[134] J. Lebenberg, I. Buvat, A. Lalande, P. Clarysse, C. Casta, A. Cochet, C. Constantinides, J. Cousty, A. de Cesare, S. Jehan-Besson, M. Lefort, L. Najman, E. Roullot, L. Sarry, C. Tilmant, M. Garreau, and F. Frouin, "Nonsupervised ranking of different segmentation approaches : application to the estimation of the left ventricular ejection fraction from cardiac cine MRI sequences.," *IEEE transactions on medical imaging*, vol. 31, pp. 1651–60, Aug. 2012.

[135] H. Necib, C. Garcia, A. Wagner, B. Vanderlinden, P. Emonts, A. Hendlisz, P. Flamen, and I. Buvat, "Detection and characterization of tumor changes in 18F-FDG PET patient monitoring using parametric imaging.," *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, vol. 52, pp. 354–61, Mar. 2011.

[136] E. D. Angelini, J. Delon, A. B. Bah, L. Capelle, and E. Mandonnet, "Differential MRI analysis for quantification of low grade glioma growth.," *Medical image analysis*, vol. 16, pp. 114–26, Jan. 2012.

[137] S. Sanga, H. B. Frieboes, X. Zheng, R. Gatenby, E. L. Bearer, and V. Cristini, "Predictive oncology : a review of multidisciplinary, multiscale in silico modeling linking phenotype, morphology and growth.," *NeuroImage*, vol. 37 Suppl 1, pp. S120–34, Jan. 2007.

[138] C. Hogea, C. Davatzikos, and G. Biros, "An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects.," *Journal of mathematical biology*, vol. 56, pp. 793–825, June 2008.

[139] E. Konukoglu, O. Clatz, B. H. Menze, B. Stieltjes, M.-A. Weber, E. Mandonnet, H. Delingette, and N. Ayache, "Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic eikonal equations.," *IEEE transactions on medical imaging*, vol. 29, pp. 77–95, Jan. 2010.

[140] O. Clatz, M. Sermesant, P.-Y. Bondiau, H. Delingette, S. K. Warfield, G. Malandain, and N. Ayache, "Realistic simulation of the 3-D growth of brain tumors in MR images coupling diffusion with biomechanical deformation.," *IEEE transactions on medical imaging*, vol. 24, pp. 1334–46, Oct. 2005.

[141] Y. Gal, N. Dowson, P. Bourgeat, O. Salvado, P. Thomas, M. Fay, S. Rose, R. L. Jeffery, A. Fazlollahi, and S. Crozier, "Amorphous regions-of-interest projection method for simplified longitudinal comparison of dynamic regions in cancer imaging.," *IEEE transactions on bio-medical engineering*, vol. 61, pp. 264–72, Feb. 2014.

[142] M. J. van den Bent, J. S. Wefel, D. Schiff, M. J. B. Taphoorn, K. Jaeckle, L. Junck, T. Armstrong, A. Choucair, A. D. Waldman, T. Gorlia, M. Chamberlain, B. G. Baumert, M. A. Vogelbaum, D. R. Macdonald, D. A. Reardon, P. Y. Wen, S. M. Chang, and A. H. Jacobs, "Response assessment in

neuro-oncology (a report of the RANO group) : assessment of outcome in trials of diffuse low-grade gliomas.," *The lancet oncology*, vol. 12, pp. 583–93, June 2011.

[143] P. Hellier, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D. L. Collins, A. Evans, G. Malandain, N. Ayache, G. E. Christensen, and H. J. Johnson, "Retrospective evaluation of intersubject brain registration.," *IEEE Trans. Med. Imag.*, vol. 22, pp. 1120–30, Sept. 2003.

[144] J. Sarvaiya, S. Patnaik, and S. Bombaywala, "Image Registration by Template Matching Using Normalized Cross-Correlation," in *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, pp. 819–822, IEEE, Dec. 2009.

[145] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information.," *IEEE Trans. Med. Imag.*, vol. 16, pp. 187–98, Apr. 1997.

[146] L.-T. Chang, "A Method for Attenuation Correction in Radionuclide Computed Tomography," *IEEE Trans. Nucl. Sci.*, vol. 25, pp. 638–643, Feb. 1978.

[147] J.-M. Tacchella, E. Roullot, M. Lefort, M.-E. Cohen, R. Guillevin, G. Petrirena, J.-Y. Delattre, M.-O. Habert, N. Yeni, A. Kas, and F. Frouin, "A new strategy to improve coregistration of spect and MR images in patients with high grade glioma.," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2013, pp. 4002–5, July 2013.

[148] J.-M. Tacchella, E. Roullot, M. Lefort, M.-E. Cohen, R. Guillevin, G. Petrirena, J.-Y. Delattre, M.-O. Habert, N. Yeni, A. Kas, and F. Frouin, "An efficient strategy based on an individualized selection of registration methods. Application to the coregistration of MR and SPECT images in neuro-oncology.," *Physics in medicine and biology*, vol. 59, pp. 6997–7011, Nov. 2014.

[149] D. Rivière, D. Geffroy, I. Denghien, N. Souedet, and Y. Cointepas, "BrainVISA : an extensible software environment for sharing multimodal neuroimaging data and processing tools," *NeuroImage*, vol. 47, p. S163, 2009.

[150] J. Ashburner, "SPM : A history," *NeuroImage*, vol. 62, no. 2, pp. 791–800, 2012.

[151] R. Deriche, "Using Canny's criteria to derive a recursively implemented optimal edge detector," *International journal of computer vision*, vol. 187, pp. 167–187, 1987.

[152] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, 1986.

[153] P. Hough, "Method and means for recognizing complex patterns," *United States Patent*, vol. 3, no. 069, 654, 1962.

[154] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, pp. 11–15, Jan. 1972.

[155] D. Ballard, "Generalizing the Hough Transform to Detect Arbitrary Shapes.," *Pattern Recognition*, vol. 6, no. 2, pp. 111–122, 1981.

[156] H. Maitre, "Un panorama de la transformation de Hough," *Traitement du signal*, vol. 2, no. 4, pp. 111–122, 1985.

[157] C. Hollitt, "A convolution approach to the circle Hough transform for arbitrary radius," *Machine Vision and Applications*, vol. 24, pp. 683–694, Apr. 2012.

[158] "Fast Fourier Transform," *Numerical Recipes in C*, pp. 1046–57, 1988.

[159] R. Lotufo and A. Falcao, "The Ordered Queue and the Optimality of the Watershed Approaches," *Mathematical Morphology and its Applications to Image and Signal Processing Computational Imaging and Vision*, vol. 18, pp. 341–350, 2000.

[160] P. Felkel, M. Bruckwschwaiger, and R. Wegenkittl, "Implementation and complexity of the watershed-from-markers algorithm computed as a minimal cost forest," *Comput. Graph. Forum*, vol. 20, pp. 26–35, June 2002.

[161] A. Klein, S. S. Ghosh, B. Avants, B. T. T. Yeo, B. Fischl, B. Ardekani, J. C. Gee, J. J. Mann, and R. V. Parsey, "Evaluation of volume-based and surface-based brain image registration methods.," *NeuroImage*, vol. 51, pp. 214–20, May 2010.

[162] J.-M. Tacchella, E. Roullot, M. Lefort, M.-E. Cohen, R. Guillevin, G. Petrirena, J.-Y. Delattre, M.-O. Habert, N. Yeni, A. Kas, and F. Frouin, "A new strategy to improve coregistration of spect and MR images in patients with high grade glioma.," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2013, pp. 4002–5, July 2013.

[163] A. Sarkar, R. J. Santiago, R. Smith, and A. Kassaee, "Comparison of manual vs. automated multimodality (CT-MRI) image registration for brain tumors.," *Med. Dosim.*, vol. 30, pp. 20–4, Jan. 2005.

[164] K. Han, M. Ren, W. Wick, L. Abrey, A. Das, J. Jin, and D. a. Reardon, "Progression-free survival as a surrogate endpoint for overall survival in glioblastoma : a literature-based meta-analysis from 91 trials.," *Neuro-oncology*, vol. 16, pp. 696–706, May 2014.

[165] M. Nowosielski, W. Recheis, G. Goebel, O. Güler, G. Tinkhauser, H. Kostron, M. Schocke, T. Gotwald, G. Stockhammer, and M. Hutterer, "ADC histograms predict response to anti-angiogenic therapy in patients with recurrent high-grade glioma.," *Neuroradiology*, vol. 53, pp. 291–302, Apr. 2011.

[166] L. B. Saltz, S. Clarke, E. Díaz-Rubio, W. Scheithauer, A. Figer, R. Wong, S. Koski, M. Lichinitser, T.-S. Yang, F. Rivera, F. Couture, F. Sirzén, and J. Cassidy, "Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer : a randomized phase III study.," *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, vol. 26, pp. 2013–9, Apr. 2008.

[167] B. M. Ellingson, H. J. Kim, D. C. Woodworth, W. B. Pope, J. N. Cloughesy, R. J. Harris, A. Lai, P. L. Nghiemphu, and T. F. Cloughesy, "Recurrent glioblastoma treated with bevacizumab : contrast-enhanced T1-weighted subtraction maps improve tumor delineation and aid prediction of survival in a multicenter clinical trial.," *Radiology*, vol. 271, pp. 200–10, Apr. 2014.

[168] C. M. Booth, E. A. Eisenhauer, N. Clinical, and T. Group, "Progression-Free Survival : Meaningful or Simply Measurable ?," *Journal of Clinical Oncology*, vol. 30, no. 10, pp. 1030–1033, 2012.

[169] P. Tylski, S. Stute, N. Grotus, K. Doyeux, S. Hapdey, I. Gardin, B. Vanderlinden, and I. Buvat, "Comparative assessment of methods for estimating tumor volume and standardized uptake value in

(18)F-FDG PET.," *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, vol. 51, pp. 268–76, Feb. 2010.

[170] H. Li, W. L. Thorstad, K. J. Biehl, R. Laforest, Y. Su, K. I. Shoghi, E. D. Donnelly, D. a. Low, and W. Lu, "A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours," *Medical Physics*, vol. 35, no. 8, p. 3711, 2008.

[171] H.-P. Gerber and N. Ferrara, "Pharmacology and pharmacodynamics of bevacizumab as monotherapy or in combination with cytotoxic therapy in preclinical studies.," *Cancer research*, vol. 65, pp. 671–80, Feb. 2005.

[172] D. Brandsma and M. J. van den Bent, "Pseudoprogression and pseudoresponse in the treatment of gliomas.," *Current opinion in neurology*, vol. 22, pp. 633–8, Dec. 2009.

[173] M.-E. Cohen, C. Pellot-Barakat, J.-M. Tacchella, M. Lefort, A. De Cesare, J. Lebenberg, N. Souedet, O. Lucidarme, T. Delzescaux, and F. Frouin, "Quantitative evaluation of rigid and elastic registrations for abdominal perfusion imaging with X-ray computed tomography," *IRBM*, vol. 34, no. 4, pp. 283–286, 2013.

[174] S. Lee, G. Wolberg, and S. Shin, "Scattered data interpolation with multilevel B-splines," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 3, pp. 228–244, 1997.

[175] Z. Li, M. Caan, and M. Ziech, "3D non-rigid motion correction of free-breathing abdominal DCE-MRI data," *Abdominal Imaging : Computational and Clinical Applications*, pp. 44–50, 2012.

[176] B. Romain, O. Lucidarme, J. Dauguet, S. Mulé, N. Souedet, Y. Chenoune, a. Guibal, T. Delzescaux, and F. Frouin, "Registration and functional analysis of CT dynamic image sequences for the follow-up of patients with hepatic tumors undergoing antiangiogenic therapy," *Irbm*, vol. 31, pp. 263–270, Oct. 2010.

[177] B. Romain, V. Letort, and O. Lucidarme, "Registration of free-breathing abdominal 3D contrast-enhanced CT," *IEEE Transactions on Visualization and Computer Graphics*, vol. 7601, pp. 274–282, 2012.

Annexe A

Publications

A.1 Publications originales dans des revues à comité de lecture

- **J.M. Tacchella**, E. Roullot, M. Lefort, M.E. Cohen, R. Guillevin, G. Petrirena, J.Y.Delattre, M.O. Habert, N. Yeni, A. Kas, F. Frouin, "An efficient strategy based on an individualized selection of registration methods. Application to the coregistration of MR and SPECT images in neuro-oncology.Phys Med Biol, 21 Nov. 2014 ;59(22) :6997-7011

- M.E. Cohen, C. Pellot-Barakat, **J.M. Tacchella**, M. Lefort, A. De Cesare, J.Lebenberg, N. Souedet, O. Lucidarme, T. Delzescaux, F. Frouin, Quantitative evaluation of rigid and elastic registrations for abdominal perfusion imaging with X-ray Computed Tomography. IRBM, 2013, 34(4-5) :283-286.

A.2 Articles publiés dans des actes de conférences

- **J.M. Tacchella**, N. Yeni, E. Roullot, M. Lefort, M.E. Cohen, R. Guillevin, G. Petrirena, J.Y. Delattre, M.O. Habert, A. Kas, F. Frouin, A Framework using Multimodal Imaging for Longitudinal Monitoring of Patients in Neuro-oncology. Application to a SPECT/MRI study. Conf Proc IEEE Eng Med Biol Soc (EMBS), Chicago, 2014, pp. 1905-8 (Présentation orale).

- **J.M. Tacchella**, E. Roullot, M. Lefort, M.E. Cohen, R. Guillevin, G. Petrirena, J.Y. Delattre, M.O. Habert, N. Yeni, A. Kas, F. Frouin, A new strategy to improve coregistration of spect and MR images in patients with high grade glioma.Conf Proc IEEE Eng Med BiolSoc, Osaka, 2013 :4002-4005. (Présentation orale).

A.3 Résumés publiés dans des actes de conférences

- N. Yeni, G. Petrirena, **J.M. Tacchella**, R. Guillevin, J.Y. Delattre, A. Kas Usefulness of brain 99mTc-MIBI SPECT in the early assessment of the response to anti-angiogenic therapy in patients with recur-

rent high-grade glioma. EUROPEAN JOURNAL OF NUCLEAR MEDICINE AND MOLECULAR IMAGING, OCT 2013, 40(2) :S395-S396

- **J.M. Tacchella**, E. Roullot, M.E. Cohen, M. Lefort, M. Reynal, J.B. Martini, R. Guillevin, J.Y. Delattre, G. Petrirena, M.O. Habert, N. Yeni, A. Kas, F. Frouin, Robust Coregistration of Sestamibi SPECT and MR Images in Patients with High Grade Glioma : a revisited approach. MAGMA, 2013, 26(Suppl 1) :447-448.

A.4 Autres publications

- **J.M. Tacchella**, N. Yeni, G. Petrirena, M.E. Cohen, M. Lefort, R. Guillevin, M.O. Habert, E. Roullot, A. Kas, F. Frouin, Usefulness of Tc-99m Sestamibi studies for monitoring response to therapy in patients with high grade gliomas : a preliminary study, Journées RITS de Recherche en Imagerie et Technologies de la Santé, Mars 2015
- **J. M. Tacchella**, E. Roullot, M. Lefort, M. E. Cohen, R. Guillevin, G. Petrirena, J. Delattre, M. Habert, N. Yeni, A. Kas, F. Frouin, "Recalage TEMP / IRM chez des patients atteints de gliomes de haut grade . Optimisation de la mise en œuvre .," Journées RITS de Recherche en Imagerie et Technologies de la Santé, Avril 2013, Bordeaux

- M. E. Cohen, C. Pellot-Barakat, **J. M. Tacchella**, M. Lefort, A. De Cesare, J. Lebenberg, O. Lucidarme and F. Frouin, "Evaluation quantitative de recalages rigides et élastiques en imagerie CT de perfusion abdominale", Journées RITS de Recherche en Imagerie et Technologies de la Santé, Avril 2013, Bordeaux

Résumé

Le suivi longitudinal en oncologie consiste à évaluer régulièrement l'évolution tumorale afin de définir un protocole de traitement adapté à chaque patient. Il est alors nécessaire d'identifier des biomarqueurs pertinents, qui permettent de quantifier précocement l'évolution tumorale. Les examens d'imagerie permettent d'assurer un suivi non invasif et de proposer des biomarqueurs complémentaires de ceux préexistants. Cette recherche de nouveaux biomarqueurs nécessite de s'appuyer sur des études cliniques et d'effectuer une étude comparative des examens réalisés à différents temps, dans différentes modalités, et ce tant au niveau global que local. Dans cette thèse, nous nous sommes attachés à développer une approche intégrée de traitement d'images pour analyser la pertinence d'indices pour le suivi tumoral. Cette stratégie se décompose en trois étapes : le recalage des données acquises à différents temps et provenant de modalités différentes, la segmentation des lésions tumorales sur les examens de chaque modalité, et le calcul d'indices globaux et locaux traduisant l'évolution spatio-temporelle de la tumeur.

L'aspect le plus innovant réside dans l'étape de recalage pour laquelle, au vu des difficultés rencontrées avec les approches classiques, nous avons proposé une nouvelle approche consistant en l'utilisation de plusieurs méthodes de recalage et la sélection de la meilleure pour chaque jeu de données, grâce à un critère quantitatif défini en fonction des applications considérées. L'apport de cette approche a été démontré dans le cadre de deux études cliniques : 1) le suivi de patients atteints de gliomes de haut grade traités par un médicament antiangiogénique nécessitant la mise en correspondance de données de tomographie d'émission monophotonique (TEMP) obtenues après injection de Sestamibi marqué au Technétium-99m avec des images de résonnance magnétique (IRM) pondérées en T1 et acquises après l'injection d'un produit de contraste ; 2) le suivi de patients atteints de lésions hépatiques et traités par différents médicaments anticancéreux, nécessitant le recalage de données de perfusion en tomodensitométrie (TDM).

La stratégie complète d'analyse a été appliquée à la première étude. Les zones tumorales ont été segmentées sur les IRM en utilisant une méthode classique de Level Set 2D et sur les données TEMP à l'aide de cinq méthodes de seuillage qui diffèrent par les options de choix du seuil. Malgré une forte corrélation en terme de volumes globaux, les indices locaux ont permis de montrer qu'une partie des volumes tumoraux détectés sur les examens TEMP précoces (réalisés 15 minutes après injection) et tardifs (réalisés 3 heures après injection) pouvaient être localisés en-dehors des volumes détectés sur les IRM après injection de Gadolinium. La corrélation élevée constatée entre, d'une part, les variations d'intensité au sein des volumes tumoraux au cours du traitement sur les examens TEMP tardifs et, d'autre part, l'indice de survie globale, suggère que cette variation relative d'intensité pourrait être prédictive de la durée de survie globale des patients, ce qui n'est pas le cas des indices utilisant les données IRM sur notre série limitée de patients. Ainsi ces résultats montrent que l'imagerie TEMP, avec un examen réalisé 3 heures après l'injection de Sestamibi marqué au Technétium-99m, peut être complémentaire de l'IRM, pour évaluer l'évolution tumorale des glioblastomes.

Les perspectives de ce travail de thèse sont d'appliquer la stratégie d'analyse à d'autres études cliniques. Néanmoins chaque étape doit être adaptée aux spécificités de l'application concernée, notamment aux modalités d'imagerie impliquées et à la zone anatomique considérée. La principale difficulté à résoudre réside dans l'automatisation et la robustesse des différentes étapes de traitement d'images envisagées.

Mots clés : recalage d'images multimodales, **TEMP**, **IRM**, **suivi longitudinal en oncologie**