UNIVERSITE DE BOURGOGNE
UFR Sciences et Techniques
LE2I - UMR 5158
Ecole doctorale E2S

THESE

Pour obtenir le grade de

Docteur de l'Université de Bourgogne
Spécialité : Instrumentation et Informatique de l'Image
présentée et soutenue publiquement
par
Sébastien PARFAIT
le 6 décembre 2010 à DIJON (21)

Classification de spectres et recherche de biomarqueurs en spectroscopie par résonance magnétique nucléaire du proton dans les tumeurs prostatiques

Directeur de thèse : Professeur **Johel MITERAN**

Coencadrant : Docteur **Paul WALKER**

JURY

M. Pr. Christophe DUCOTTET	Université Jean Monnet, Saint Etienne	Rapporteur
M. Dr. Dominique SAPPEY-MARINIER	Université Claude Bernard, Lyon	Rapporteur
M. Pr. Franck MARZANI	Université de Bourgogne, Dijon	Examinateur
Mme. Pr. Edwige PISSALOUX	Université Pierre et Marie Curie, Paris	Examinateur
M. Pr. Johel MITERAN	Université de Bourgogne, Dijon	Examinateur
M. Dr. Paul WALKER	Université de Bourgogne, Dijon	Examinateur
M. Dr. Xavier TIZON	Oncodesign SA, Dijon	Membre invité

A mon Papa et à Stéphanie,

Ainsi qu'à toutes les autres personnes que j'aimais qui sont disparues durant ces quatre années.

Remerciements

Je tiens à remercier le Professeur Johel MITERAN pour m'avoir encadré durant cette thèse, pour sa rigueur scientifique, ses remarques avisées et la liberté de mouvement qu'il m'a laissé durant ce travail. Je tiens à le remercier également pour son soutien sans faille pendant les heures les plus sombres que j'ai connues au cours de ces quatre années.

Je remercie également le Docteur Paul WALKER pour le co-encadrement de cette thèse, pour les innombrables fois où il a dû m'expliquer et me réexpliquer les subtilités de la résonance magnétique et pour toutes les corrections qu'il a dû faire lorsque je saccageais sa belle langue (et la nôtre aussi d'ailleurs). Je le remercie également pour les longues discussions, principalement sportives, que nous avons eues et pour la découverte du Cricket.

Je tiens également à remercier le Professeur Christophe DUCOTTET et le Docteur Dominique SAPPEY-MARINIER qui m'ont fait l'honneur d'être les rapporteurs de ce travail.

Je remercie les Professeurs Edwige PISSALOUX et Franck MARZANI d'avoir participé au jury sanctionnant ces quatre années de travail.

Je remercie la société Oncodesign pour le financement de cette thèse et notamment son Président, le Docteur Philippe GENNE et Olivier DUCHAMP pour la conduite du projet au sein de la société et la gestion des protocoles expérimentaux. Je suis reconnaissant au Docteur Xavier TIZON pour les acquisitions IRM, ses conseils toujours pertinents et sa participation active à cette thèse par le comité de suivi et enfin sa présence en tant que membre invité de ce jury. Je remercie également le Docteur Peggy PROVENT pour les acquisitions et ses connaissances du signal de RMN ainsi que tout le personnel de la société Oncodesign pour son accueil toujours très chaleureux.

Je remercie aussi le Conseil Régional de Bourgogne et le Fond Social Européen pour le soutien financier apporté à cette thèse.

Je remercie également tous les membres du LE2I que j'ai pu côtoyer, pour leur sympathie et plus particulièrement le Professeur François BRUNOTTE pour m'avoir accueilli au sein de son équipe d'imagerie médicale ainsi que le Docteur Alain LALANDE et Marie MAGNIER pour les débats d'idées (sur nos travaux ou autres…) parfois autour d'un café.

Je tiens aussi à remercier l'équipe de médecine nucléaire du Centre Georges François Leclerc pour les informations échangées lors des colloques hebdomadaires.

Je remercie les laboratoires d'anatomopathologie du CHU et du CGFL pour l'étude histologique des prostates et pour les images fournies, ainsi que le personnel du service d'imagerie par résonance magnétique du CHU pour sa gentillesse et son accueil lors de mes quelques visites notamment en tant que "cobaye".

Je remercie également tout particulièrement le Docteur Gilles CREHANGE pour tous les éclaircissements médicaux qu'il m'a donnés sur le cancer de la prostate et pour m'avoir associé à ses propres travaux de recherche.

Je tiens également à remercier la Société YMAG pour m'avoir permis de reprendre ces études qui m'ont mené ici et notamment ses dirigeants Patrick AYROULET, Didier MACK et Frédéric CHOLLET. Merci aussi à tous les employés pour tous les bons moments passés ensemble, les repas de Noël ou encore les foots, les paint balls et autres activités que nous partageons toujours, vous êtes un peu trop nombreux pour que je vous cite tous.

Ensuite je tiens à remercier tous mes amis (et Dieu seul, à part moi, sait qu'ils sont nombreux) : je suis désolé de ne pas pouvoir vous adresser un petit mot à chacun, mais cela demanderait un autre volume à cette thèse. J'espère que vous vous reconnaitrez et que vous savez à quel point je vous aime et vous suis reconnaissant : sans vous, je ne serais peut être pas là. Et que ce soit, à la chasse, au sport, au collège, au lycée, à l'université ou encore devant un bon match, la WII ou tout simplement autour d'une bonne table merci pour tous ces moments passés et surtout futurs.

Ensuite je tiens à remercier tous les membres de ma famille pour le soutien sans faille qu'ils m'accordent depuis le début de ce projet, pour tout l'amour qu'ils m'ont témoigné depuis ma naissance et dont ils m'apportent encore des preuves tous les jours : Notamment ma Maman, Jean-Eric, Christophe, Catherine, Thierry, Céline, mes frères et sœurs, mes beaux frères et belles sœurs, et enfin mes neveux et nièces : Coralie, Solène, Quentin, Pierrick, Lilian et Axel.

Une mention particulière pour deux personnes qui appartiennent aux deux précédentes catégories : Francis et Samuela, je les remercie tout particulièrement pour tout ce qu'ils ont fait et feront et je les félicite pour leur petit Maxence et leur nouvelle vie.

A ces remerciements à ma famille, je souhaiterais associer Monsieur et Madame MARTINEZ : nous avons vécu des moments très difficiles ensemble. Ils m'ont toujours soutenu même dans la difficulté et je suis très content que nous soyons toujours un peu plus qu'amis, que ce lien ne s'éteigne jamais.

Comment ne pas remercier également toutes les personnes qui ont participés à mon éducation de l'enseignement primaire, à l'enseignement supérieur, toutes ces personnes qui ont crues en moi et m'ont parfois poussé malgré moi dans la bonne voie.

Avant de terminer, je tiens à remercier Philippe, Damien, Hélène et Jacqueline pour leur soutien et la confiance qu'ils m'accordent : j'espère me montrer digne d'eux.

Enfin, "last but not least" comme dirait notre très cher Paul, un grand merci à Sophie. Tu m'as évité un grand nombre de fautes en relisant scrupuleusement ce manuscrit à plusieurs reprises. Merci pour tout ce que tu as fait pour m'aider pour cette thèse et tout ce que tu fais tous les jours, les mots me manque pour t'exprimer ma gratitude … Je voudrais encore te remercier pour m'avoir fait le plus beau cadeau du monde : notre petit Arthur.(enfin merci à lui de me laisser dormir pendant ces dernières nuits de labeurs et de stress).

A toutes ces personnes et à toutes celles que j'aurai oubliées et auprès desquelles je m'excuse, je dis un grand **MERCI.**

Résumé

Le cancer de la prostate est le cancer le plus fréquent chez l'homme de plus de 50 ans. Actuellement, les méthodes de dépistage manquent soit de sensibilité, soit de spécificité ou sont désagréables pour le patient.

La spectroscopie par résonance magnétique (SRM) permet l'étude du métabolisme in vivo. L'utilisation d'appareil haut champ (≥3T) permet dorénavant d'analyser la prostate sans antenne endorectale.

L'objectif de cette thèse est de créer un système automatique de dépistage de ce cancer en mettant au point une méthode de classification automatique permettant de traiter les données obtenues grâce à la SRM.

La SRM est un phénomène complexe, très sensible aux conditions d'acquisition. Nous avons donc étudié comment améliorer l'acquisition de ce signal. Cependant, même avec une acquisition de très bonne qualité, le signal de résonance magnétique doit subir quelques traitements pour être analysable automatiquement par une méthode de classification. La suite du travail a donc consisté à rechercher les traitements à appliquer pour optimiser les spectres en vue d'une classification. Nous avons alors recherché la méthode de classification optimale pour ce problème. Cet ensemble d'étapes (acquisition du signal, traitement des spectres puis classification des données obtenues) nous permet de mettre en évidence la présence de tumeurs de la prostate avec un taux d'erreur global de moins de 12%.

Dans un second temps, nous avons cherché de nouveaux biomarqueurs dans les spectres. Ces biomarqueurs pouvaient être un métabolite précis ou une plage de fréquence correspondant à plusieurs métabolites. Nous n'avons pas trouvé d'attributs plus significatifs que la choline ou le citrate, cependant quelques bandes de fréquence semblent participer à l'amélioration des taux d'erreurs.

Enfin, nous avons élargi notre champ d'investigation en tentant d'appliquer ces techniques chez le rat. Des contraintes liées à l'acquisition ne nous ont pas permis d'obtenir suffisamment de spectres dans le cas pré-clinique. Nous avons cependant pu valider la faisabilité de la SRM chez le rongeur et sa pertinence dans le cerveau. La technique doit cependant être améliorée pour pouvoir être validée dans le cas du cancer de la prostate chez le rat.

Abstract

Prostate cancer is the most common cancer in men over 50 years. Current detection methods either lack sensitivity or specificity or are unpleasant for the patient.

Magnetic resonance spectroscopy allows the study of metabolism in vivo. The use of a high field machine (≥3T) has allowed us to dispense with the use of an endorectal coil, which is particularly uncomfortable for the patient.

The objective of this thesis is to create an automatic method to detect cancer by processing data obtained through magnetic resonance spectroscopy

MRS is a complex phenomenon, very sensitive to acquisition conditions. Firstly, we have studied how to improve and optimise signal acquisition. However, even with a very good quality signal, it must still undergo further post-processing to be analysed automatically by a classification method. Further work was therefore needed to investigate which post-processing steps were required in order to optimize the spectra for classification. We then investigated the optimal classification method for this problem. A particular set of steps (signal acquisition, processing and spectral classification data) allows us to highlight the presence of prostate tumors with an overall error rate of less than 12%.

In a second step, we searched for new biomarkers within the spectra. These biomarkers could be a metabolite or a specific frequency range corresponding to several metabolites. We did not find any additional significant attributes other than choline and citrate, however, some frequency bands seem to participate in improving the error rate.

Finally, we expanded our investigation by attempting to apply these techniques to the rat. Technical constraints related to acquisition did not allow us to obtain a sufficient number of spectra in the pre-clinical cases. Nonetheless, we have validated the feasibility of MRS in rodents and its relevance in the brain. The technique, however, must be improved in order to be validated in the case of prostate cancer in rats.

Table des matières

Abréviations 21
Notations mathématiques 23
Définitions 25
Liste des figures 27
Liste des tables 31
Introduction 33
Chapitre 1 : Le cancer de la prostate : origine et méthodes d'analyses 39
1.1 La prostate 39
1.1.1 Anatomie 39
1.1.1.1 Cas de l'homme 39
1.1.1.2 Cas du rat 41
1.1.2 Métabolisme 42
1.1.3 Pathologies 43
1.1.3.1 L'hyperplasie bénigne de la prostate 43
1.1.3.2 Le cancer 43
1.1.4 Dépistage 44
1.1.4.1 Toucher rectal 44
1.1.4.2 Dosage de l'antigène spécifique de la prostate 45
1.1.4.3 Biopsie 46
1.1.4.4 Imagerie de résonance magnétique 46
1.2 Etat de l'art en spectroscopie par résonance magnétique du proton et classification de spectres 47
1.2.1 Spectroscopie de la prostate 47
1.2.2 Classification de spectres 48
1.2.2.1 Cas du cerveau 48
1.2.2.2 Cas de la prostate 52
1.2.3 Conclusion 56
Chapitre 2 : La spectroscopie de résonance magnétique et l'optimisation des paramètres d'acquisition pour l'obtention des données. 59
2.1 La résonance magnétique 59
2.1.1 Principe de la résonance magnétique 60

2.1.1.1 Notions de physique 60
2.1.1.2 La précession 64
2.1.1.3 Le signal de résonance magnétique 65
2.1.2 La spectroscopie de résonance magnétique 71
2.1.2.1 Les séquences 76
2.1.2.2 Localisation 79
2.2 Paramètres des séquences 82
2.2.1 Le couplage spin-spin 82
2.2.2 Influence du temps de répétition 85
2.2.3 Influence du temps d'écho 87
2.2.4 Influence des paramètres d'échantillonnage 87
2.3 Choix des paramètres de l'acquisition 88
2.3.1 Clinique 88
2.3.2 Pré-clinique 89
2.4 Ensembles de données 90
2.4.1 Cas clinique 90
2.4.2 Cas préclinique 91
Chapitre 3 : Prétraitement des données 93
3.1 Introduction 93
3.2 Traitements mis en œuvre 94
3.2.1 Correction de phase 94
3.2.2 Suppression de la ligne de base 97
3.2.3 Normalisation 100
3.2.4 Estimation des concentrations des métabolites 102
3.2.4.1 LCModel 102
3.2.4.2 Modélisation par des fonctions élémentaires 104
Chapitre 4 : Classification 117
4.1 Les machines à vecteurs de support 119
4.1.1 Cas simple : SVM linéaire à deux classes. 120
4.1.2 Cas non linéairement séparable 123
4.1.3 Cas multiclasse 129
4.2 Les Réseaux de neurones 129
4.3 Validation croisée 133
4.4 Résultats 135

Chapitre 5 : Sélection de biomarqueurs....141
5.1 Analyse en composantes principales....142
5.1.1 Principe de l'ACP....142
5.1.2 Application de l'ACP à la réduction de dimension des spectres....143
5.1.3 L'ACP pour la sélection d'attributs : application basique....145
5.1.4 L'ACP pour la sélection d'attributs : Principal Feature Analysis....148
5.2 Sequential Forward Selection....153
5.3 Sélection par plage....155
5.4 Conclusion....166
Chapitre 6 : Données pré-cliniques....167
6.1 Expériences sur le cerveau....167
6.2 Expériences sur la prostate....169
6.3 Conclusion....172
Conclusion....173
Références....177
Annexes....187
Annexe 1....187
Annexe 2....189
Annexe 3....190

Abréviations

ACME : Automated phase Correction based on Minimization of Entropy.

ACP : Analyse en Composantes Principales.

BA : Base d'Apprentissage.

CSI : Chemical Shift Imaging.

EOPS : Exempt d'Organisme Pathogène Spécifique.

FFT : Fast Fourier Transform (transformée de Fourier rapide).

FID : Free Induction Decay (signal de précession libre).

FN : Faux Négatifs.

FP : Faux Positifs.

GE : Graph Embedding.

HBP : Hyperplasie (ou Hypertrophie) Bénigne de la Prostate.

IRM : Imagerie par Résonance Magnétique.

LDA : Linear Discriminant Analysis (analyse discriminante linéaire).

LLE : Locally Linear Embedding.

MLP : MultiLayer Perceptron.

NAA : N-acétyl-aspartate.

PRESS : Point REsolved SpectroScopy.

PSA : Prostate Specific Antigen (antigène specifique de prostate).

RBF : Radial Basis Function.

RF : (onde) RadioFréquence.

RMN : Résonance Magnétique Nucléaire.

SFS : Sequential Forward Selection.

SRM : Spectroscopie par Résonance Magnétique.

STEAM : STimulated Echo Acquisition Mode.

SVM: Support Vector Machine (machines d'apprentissage à vecteurs de support).

TE : Temps d'Echo.

TM : Temps de Mélange.

TMS : TétraMéthylSilane.

TR : Temps de Répétition.

Outre ces abréviations usuelles, nous utiliserons les acronymes suivants pour nos algorithmes :

CLB : Correction de la Ligne de Base.

DM : Détection de Métabolites.

DSF : Décomposition du spectre en une Somme de Fonctions.

Notations mathématiques

p(x) : densité de probabilité de la variable x.

P(x) : Probabilité de l'événement x.

$\hat{F}$: Transformée de Fourier de la fonction F.

$\|\cdots\|$: Norme euclidienne.

$\tanh$: Tangente hyperbolique.

$\langle .,. \rangle$: Produit scalaire.

$\exp(x)$: Fonction exponentielle.

Définitions

Adénocarcinome : (grec : adên, glande ; karkinos, crabe) Tumeur maligne développée aux dépens d'un épithélium glandulaire.

Adénome : (grec : adên, glande) Tumeur bénigne développée aux dépens d'une glande et dont la structure rappelle celle de la glande normale dont elle dérive.

Astrocytome : Variété de gliome bénin, solide ou kystique, constituée aux dépens des astrocytes (variété de cellules gliales).

Carcinome : (grec : karkinos, crabe) Tumeur épithéliale maligne.

Epithélium : Tissu de recouvrement de la surface et des cavités internes de l'organisme.

Gliales (cellules) : Le système nerveux est composé de deux types de cellules : les neurones et les cellules gliales.

Glioblastome : Tumeur primitive du cerveau la plus fréquente et la plus agressive.

Gliome : Terme générique désignant les tumeurs primitives du système nerveux central.

Méningiome : Tumeur habituellement bénigne développée dans les méninges.

Oligodendrogliome : Tumeur cérébrale individualisée parmi les autres tumeurs gliales en raison d'anomalies génétiques particulières et d'une plus grande sensibilité à la chimiothérapie.

Orthotopique : Se dit d'un élément se trouvant dans sa localisation normale. L'implantation d'un cancer de la prostate dans celle-ci, est dite orthotopique contrairement à l'implantation de la même tumeur en sous cutané.

Liste des figures

Figure 1 : Tumeur de la prostate en imagerie T2 (a), diffusion (b), perfusion (c) et coupe histologique(d). 34
Figure 2 : Exemple de spectres de prostate saine (gauche) et de tumeur (droite) analysés avec le logiciel LCModel....34
Figure 3 : Méthodologie de la classification supervisée. 35
Figure 4 : Coupe sagittale de l'appareil génital masculin. 39
Figure 5 : Représentation de la prostate (vue de 3/4). 40
Figure 6 : Anatomie zonale de la prostate. 41
Figure 7 : Prostate de rat. 42
Figure 8 : Schéma de dépistage habituel d'un cancer de la prostate. 45
Figure 9 : Biopsie transrectale sous contrôle échographique. 46
Figure 10 : Niveaux d'énergie des populations de spins avec et sans champ magnétique pour un noyau avec un nombre de spins de ½. 61
Figure 11 : Energie des spins en fonction de leur orientation par rapport au champ magnétique $B_{0.}$ 62
Figure 12 : Variation de la différence d'énergie entre les deux niveaux en fonction de l'intensité du champ magnétique $B_{0.}$ 63
Figure 13 : Représentation d'un noyau comme un élément présentant une aimantation avec son moment angulaire (L) et son moment d'aimantation (u). 64
Figure 14 : Les deux rotations d'un noyau en présence d'un champ magnétique. 64
Figure 15 : Mouvement de précession et aimantation globale. 65
Figure 16 : Le signal de résonance magnétique nucléaire. 66
Figure 17 : Bascule de l'aimantation globale et sa décomposition. 67
Figure 18 : La relaxation transversale. 67
Figure 19 : Signal de précession libre. 68
Figure 20 : Séquence d'écho de spin. 69
Figure 21 : Le processus de la séquence d'écho de spin. 70
Figure 22 : La séquence d'écho de spin multiple permet d'obtenir un signal en décroissance T2. 70
Figure 23 : Structure chimique de la molécule de TMS. 72
Figure 24 : Influence du temps T2* sur la largeur des pics. 73

Figure 25 : Influence du temps T2* sur la séparation des pics. 73
Figure 26 : Couplage spin-spin. 74
Figure 27 : Spectres de l'éthanol en résonance magnétique nucléaire du proton en solution dans du Chloroforme (CHCL3). 75
Figure 28 : Schéma de la séquence d'écho de spin. 76
Figure 29 : Séquence STEAM. 78
Figure 30 : Séquence PRESS. 79
Figure 31 : Principe de sélection d'un voxel avec la technique mono-voxel. 80
Figure 32 : Séquence d'imagerie spectroscopique (2D) basée sur la séquence PRESS. 81
Figure 33 : Schémas de codage de la matrice CSI. 82
Figure 34 : Formule semi développée et modélisation 3D de la molécule de citrate. 83
Figure 35 : Variation des pics de citrate en SRM du proton : simulation (a et c) et in vitro (b et d). 84
Figure 36 : Pics de citrate obtenus sur la machine Siemens en fonction du TE. 85
Figure 37 : Influence du TR sur le signal appliqué à la choline. 87
Figure 38 : IRM de prostate avec antenne endorectale et grille de spectroscopie. 88
Figure 39 : Les deux modes d'un pic. 94
Figure 40: Spectre avant (a) et après (b) l'application d'une correction de phases de zéro et de premier ordre. 96
Figure 41 : Variation de la fonction d'entropie en fonction des phases de zéro et de premier ordre. 96
Figure 42 : Ligne de base obtenue par la méthode de Devos en fonction du paramètre β. 98
Figure 43 : Processus de détection de la ligne de base. 99
Figure 44 : Erreur quadratique entre la ligne de base simulée et la ligne de base estimée pour les deux méthodes évaluées. 100
Figure 45 : Intensité du pic d'eau par coupe (acquisition de 8 coupes de 16x16 voxels). 101
Figure 46 : Spectre analysé avec LCModel. 103
Figure 47 : Approximation du spectre de l'eau par une gaussienne et une lorentzienne. 105
Figure 48 : Simulation d'un spectre et représentation de sa dérivée première et seconde. 107
Figure 49 : Progression de la détection des pics. 109
Figure 50 : Spectre et estimation constituée de la somme des gaussiennes détectées. 109
Figure 52 : Exemples de détection de métabolites par Lorentzienne (vert) comparés à LCModel (rouge). 113

Figure 53 : Représentation de l'aire sous la courbe lorentzienne versus la concentration estimée par LCModel pour le citrate et la choline. 114
Figure 54 : Exemple de détection par la méthode DSF sur le spectre d'une tumeur de rat. 115
Figure 55 : Principe des SVM : trouver l'hyperplan séparant les groupes d'éléments en maximisant la marge. 121
Figure 56 : Exemple d'hyperplans séparateurs. 122
Figure 57 : Deux classes non linéairement séparables. 124
Figure 58 : Principe de "l'outlier". 125
Figure 59 : Variation de l'hyperplan séparateur calculé par les SVM avec un noyau RBF Gaussien en fonction du paramètre γ 128
Figure 60 : Schéma d'un neurone selon la définition de McCulloch-Pitts. 130
Figure 61 : Schéma du perceptron. 131
Figure 62 : Exemple de perceptron à 2 couches de neurones. 133
Figure 63 : Schéma de validation croisée d'ordre c. 134
Figure 64 : Organisation des méthodes de sélection d'attributs. 141
Figure 65 : Variance expliquée en fonction du nombre d'attributs sélectionnés. 144
Figure 66 : Variations des taux d'erreur des SVM en validation croisée d'ordre 5 en fonction du paramètre du noyau RBF gaussien. 145
Figure 67 : Représentation graphique du coefficient α associé à chaque attribut dans la combinaison linéaire de la nouvelle variable C_1. 146
Figure 68 : Coefficients de la combinaison linéaire de la variable C_1 triés par ordre décroissant. 147
Figure 69 : Variation des taux d'erreur en fonction du noyau RBF avec 120 attributs en entrée des SVM en validation croisée d'ordre 5. 147
Figure 70 : Attributs sélectionnés par la méthode PFA en conservant 95 et 99 pourcents de la variance et positionnés sur un spectre. 149
Figure 71 : Variation des taux d'erreur avec les attributs sélectionnés par la méthode PFA en conservant 95% (23 attributs) et 99% (96 attributs) de la variance utilisée en entrée des SVM avec un noyau RBF gaussien. 150
Figure 72 : Bande de l'eau et bande "des graisses" supprimées. 151
Figure 73 : Attributs sélectionnés par la méthode PFA en conservant 95% et 99% de la variance. 152
Figure 74 : Algorithme SFS. 154

Figure 75 : Les 12 attributs sélectionnés par la méthode SFS positionnés sur un spectre moyen et numérotés par leur ordre de sélection par l'algorithme. 155
Figure 76 : Variation des taux d'erreur par classe et total en fonction du noyau RBF gaussien utilisé. 155
Figure 77 : Fenêtre glissante de 200 points (≈2 ppm). 156
Figure 78 : Variation du meilleur taux d'erreur totale atteint en fonction de la taille de la plage. 156
Figure 79 : Plage de 10 points soit environ 0,1 ppm par pas de 5 points : représentation des deux principales plages entrainant une baisse du taux d'erreur. 157
Figure 80 : Pas de 5, plage de 2% (20 points) 158
Figure 81 : Pas de 5, plage de 5% (51 points) 159
Figure 82 : Pas de 5, plage de 10 % (102 points). 160
Figure 83 : Pas de 5, plage de 15% (153 points). 161
Figure 84 : Pas de 5, plage de 20% (204 points). 162
Figure 85 : Evolution du taux d'erreur pour les plages de 307 (a), 409 (b) et 512 points (c). 163
Figure 86 : Représentation des plages permettant d'obtenir le meilleur taux de classification pour les plages de 307, 409, et 512 points. 164
Figure 87 : Plages permettant d'obtenir individuellement une amélioration notable du taux d'erreur totale. 165
Figure 88 Spectres de cerveaux de rats sains, de tumeurs CGL9 et U87-MG avec et sans traitement (RT : radiothérapie, TMZ : Témozolomide). 168
Figure 89 : Variations du rapport NAA/choline avec les deux modeles de gliome en fonction du temps. 169
Figure 90 : Spectres obtenus pour chaque rat de chaque groupe au trois temps d'imagerie. 170
Figure 91 : Pour chaque position du spectre étudié (3.9, 3.65, 3.21, 2, 1.3 et 0.9 ppm), variation de l'aire sous la courbe de la lorentzienne détectée pour chaque rat de chaque groupe (CTRL : contrôle CSTR : castré). 171
Figure 92 : Concentrations en choline estimées par LCModel dans des tumeurs PC3-MM2 orthotopiques. 172
Figure 93 : Schéma de classification optimale. 174
Figure 94 : Schéma de traitement optimal envisagé. 175

Liste des tables

Tableau 1 : Réceptivité relative des principaux noyaux utilisés en résonance magnétique...62
Tableau 2 : Durée de l'examen et quantité de signal mesuré en fonction du TR et du nombre de répétitions pour la choline. 86
Tableau 3 : Fréquence de résonance des principaux métabolites détectables dans la prostate en RMN (en ppm). 110
Tableau 4 : Taux d'erreur de classification (%) des SVM avec uniquement les cas sains et pathologiques en entrée. 138
Tableau 5 : Taux d'erreur de classification(%) des SVM et des MLP avec les cas sains, pathologiques et indéterminés en entrée. 138
Tableau 6 : Nombre d'attributs des ensembles de données suivant la sélection d'attributs utilisée et l'ensemble de départ. 152
Tableau 7 : Taux d'erreur (%) obtenus avec les SVM (noyau RBF gaussien) en validation croisée d'ordre 5....152
Tableau 8 : Tumeur primitive (T)....187
Tableau 9 : Ganglions régionaux (N)....188
Tableau 10 : Métastases (M)....188

Introduction

Cette thèse est le fruit de la collaboration entre Oncodesign et le Le2i dans le cadre du projet Pharmimage et avec le soutien financier de la région Bourgogne et du Fond Social Européen. Oncodesign est une société de biotechnologie spécialisée dans la recherche de nouvelles molécules pour la thérapie anticancéreuse et le Le2i est le Laboratoire Electronique, Informatique et Image, UMR CNRS 5158 de l'Université de Bourgogne, tous deux sont situés à Dijon.

Le thème central de cette étude est le cancer de la prostate : il s'agit du cancer le plus fréquent chez l'homme de plus de 50 ans. Chaque année, plus de 40 000 diagnostics sont effectués. Ce cancer représente 3,4% des causes de décès chez l'homme et 10,7 % des causes de décès par cancer. Il est actuellement dépisté par toucher rectal et dosage dans le sang d'une protéine : l'antigène spécifique de la prostate, plus connu sous le nom de PSA (Prostate Specific Antigen). Cependant l'élévation de PSA dans le sang n'est pas liée spécifiquement au cancer de la prostate, mais peut être due à d'autre pathologie comme l'Hyperplasie Bénigne de la Prostate (HBP). Le diagnostic est alors réalisé par les médecins anatomopathologistes qui analysent les biopsies prostatiques effectuées sous contrôle échographique. Cette technique est très fiable si les biopsies ont ciblé les zones tumorales, or au début de la maladie, les lésions peuvent être de petite dimension et il est possible que le praticien fasse ses prélèvements en dehors du ou des foyers tumoraux.

L'Imagerie par Résonance Magnétique (IRM) présente plusieurs avantages dans le cas de la prostate :

- L'IRM permet d'obtenir un très bon contraste dans les tissus mous contrairement aux rayons X,
- la résolution spatiale est aussi excellente (infra millimétrique),
- il existe des marqueurs spécifiques pouvant être suivis en IRM,
- la spectroscopie permet d'évaluer in vivo la représentation (on parle souvent de concentration) de certaines molécules spécifiques du métabolisme (les métabolites).

Nous avons représenté sur la Figure 1, un exemple de prostate en imagerie de résonance magnétique ainsi qu'une coupe histologique.

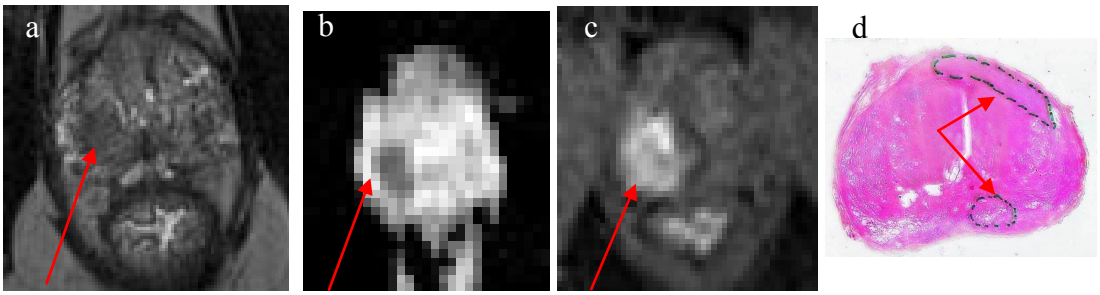


Figure 1 : Tumeur de la prostate en imagerie T2 (a), diffusion (b), perfusion (c) et coupe histologique(d).

La flèche rouge pointe sur la tumeur. Elle apparait avec un signal légèrement moins intense sur l'image T2(a) ainsi que sur l'image de diffusion (b). L'image (c) est issue de la séquence de perfusion lors de l'arrivée du traceur dans la tumeur. Sur la coupe histologique (d), le praticien a entouré la tumeur (patient différent des images précédentes).

De plus, avec l'arrivée des aimants haut champ (≥3T), cette technique qui nécessitait auparavant l'emploi d'une antenne endo-rectale pour obtenir un signal de qualité suffisante (au moins pour la spectroscopie), peut être maintenant réalisée en routine avec une antenne en réseau phasé (de surface).

Malgré ces avantages, l'IRM est très peu employée en urologie car les appareils haut champ sont relativement peu répandus.

L'équipe du service d'imagerie par résonance magnétique et de médecine nucléaire du CHU de Dijon a donc choisi de s'investir dans l'étude de la prostate. En moyenne, cinq patients sont examinés par semaine. En routine clinique, ces examens sont utilisés pour vérifier l'extension tumorale ou pour orienter une seconde série de biopsies lorsque la première s'est révélée négative malgré un taux de PSA élevé ou en augmentation.

Plusieurs études [1-5] ont montré qu'il existait un lien entre les taux de choline et de citrate mesurés en spectroscopie par résonance magnétique (SRM) et la présence d'une tumeur (voir Figure 2). Ces constatations ont permis de définir deux axes d'étude.

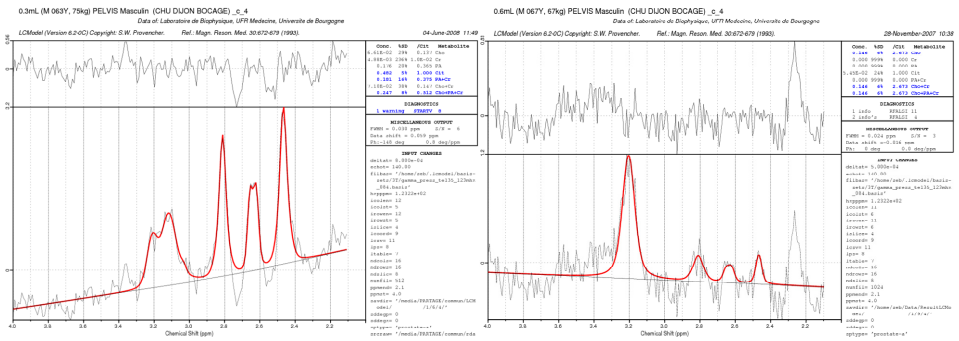


Figure 2 : Exemple de spectres de prostate saine (gauche) et de tumeur (droite) analysés avec le logiciel LCModel.

Le premier axe inclut des patients du Centre anticancéreux Georges François Leclerc (CGFL) et du CHU de Dijon et consiste en l'évaluation de techniques de détection

et de classification automatique du cancer de la prostate à partir de données issues de SRM. Le laboratoire Le2i travaillant depuis plusieurs années dans le domaine de la classification, nous avons utilisé cette expertise afin de mettre au point une méthodologie de classification automatique des spectres de la prostate, qui pourra s'avérer une aide précieuse au diagnostic.

La méthodologie couramment employée en classification supervisée (avec avis d'un expert) peut être résumée par le schéma représenté Figure 3. Après différents traitements destinés à améliorer les données en vue de leur classification (normalisation, suppression d'informations inutiles ou extraction de l'information principale, …), l'ensemble des données est séparé en deux sous ensembles. Le premier servira à la phase d'apprentissage, c'est-à-dire la création d'un modèle permettant au classifieur de séparer au mieux les données. Le second sera utilisé pour valider le modèle en vérifiant qu'avec celui-ci, le classifieur affecte la bonne étiquette à chaque élément.

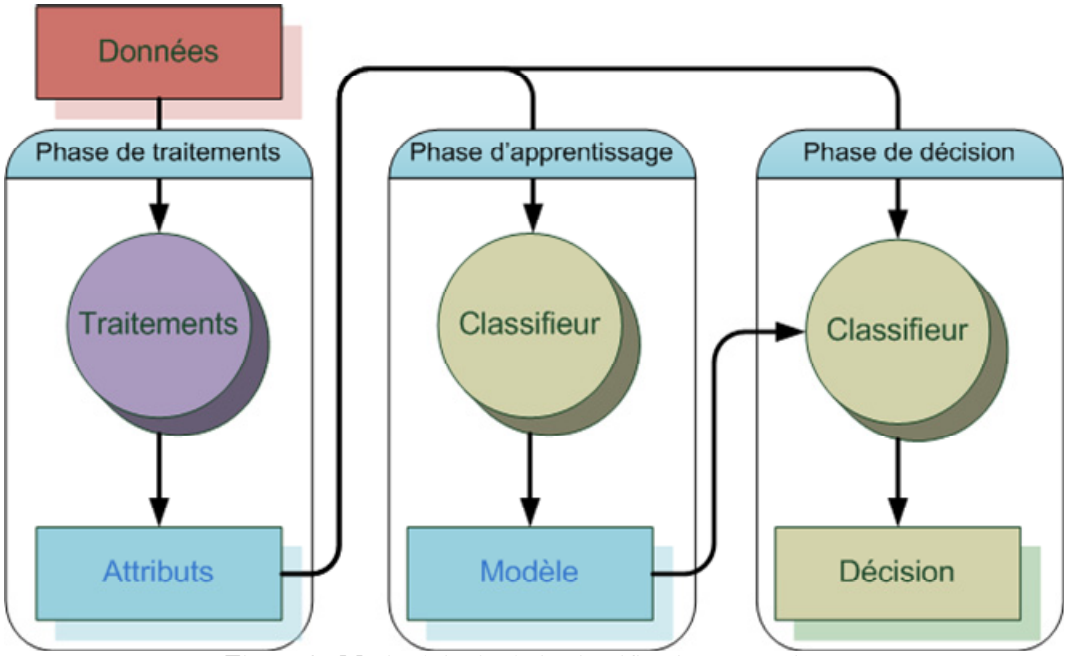


Figure 3 : Méthodologie de la classification supervisée.

Le second axe, en partenariat avec la société Oncodesign, s'intéresse à l'évaluation du traitement anticancéreux chez le rat. La société Oncodesign collabore avec différents laboratoires à la mise au point de molécules anticancéreuses en les testant sur des rongeurs (phase pré-clinique). Pour ce faire, l'entreprise dispose d'une série de procédures permettant de valider l'action des molécules. Pour élargir son offre de méthodes de test, Oncodesign a acquis un appareil IRM spécifique au petit animal et souhaitait donc pouvoir

juger du potentiel de la spectroscopie dans le cadre de l'évaluation de la réponse des tumeurs prostatiques aux traitements.

Pour cette étude, bien que nous disposions des données acquises avec toutes les modalités de la résonance magnétique, nous nous sommes concentrés sur la spectroscopie pour essayer de discerner les tumeurs malignes des autres pathologies de la prostate (il sera bien entendu intéressant dans de futures études de coupler plusieurs modalités telles que l'imagerie de diffusion, l'imagerie de perfusion et la spectroscopie).

Comme nous allons tenter d'en faire la démonstration au cours de ce mémoire, nos contributions portent essentiellement sur les points suivants : l'optimisation des conditions d'acquisition et de la qualité des données, l'évaluation des performances en classification automatique, la recherche de biomarqueurs, et ce, dans la mesure du possible, dans les cas clinique et préclinique.

La classification de spectre de résonance magnétique étant un axe de recherche nouveau à Dijon, nous avons commencé par réaliser une étude de l'état de l'art des différents domaines abordés. A partir de ces informations, nous avons donc cherché à optimiser différentes phases de ce processus en commençant par l'acquisition des données. En effet des données de bonne qualité nécessitent moins de traitements et la plupart des travaux de la littérature ont été réalisés sur des appareils à résonance magnétique à 1,5T avec antenne endo-rectale. Cette antenne est inconfortable pour le patient, et déforme la prostate empêchant toute mise en correspondance avec d'autre modalité d'imagerie. Nous avons donc travaillé à obtenir un signal de bonne qualité, sans antenne endorectale, ce qui a été rendu possible notamment grâce à l'utilisation d'un appareil à 3T.

Dans le but d'optimiser les résultats de classification automatique, nous avons ensuite implémenté et parfois amélioré une série de traitements usuels en SRM. Dans la pratique, nous avons dû évaluer régulièrement les performances de classification pour déterminer les meilleures données et les meilleurs prétraitements. Seuls les résultats les plus significatifs seront présentés dans ce mémoire. Cette recherche nous a notamment menés au développement d'un algorithme d'évaluation des concentrations des métabolites, algorithme spécifique à la prostate humaine.

Nous avons enfin travaillé à la sélection des attributs utilisés, toujours en cherchant à minimiser le taux d'erreur de classification, ou en recherchant les plages d'attributs contenant le maximum d'informations.

Nous présenterons ces différents points dans ce document organisé en six chapitres.

Le premier chapitre sera tout d'abord consacré à une présentation du sujet en commençant par des rappels sur la prostate, ses pathologies et les techniques de dépistage. Nous poursuivrons par un résumé de l'état de l'art qui sera traité en deux parties : la première traitera des études basées sur la SRM du proton de la prostate, qui ont principalement inspiré cette étude, et la seconde présentera les méthodes existantes permettant la classification automatique de spectres de résonance magnétique du proton.

Le chapitre deux sera dédié à la résonance magnétique : ce phénomène est très complexe et le signal peut subir d'importantes variations en fonction des paramètres d'acquisition. Il est donc très utile de connaitre les effets de ces paramètres sur le signal pour pouvoir l'analyser convenablement. Nous en verrons le principe général, puis nous approfondirons le problème de la SRM. Nous examinerons ensuite les principaux phénomènes qui ont une influence importante sur la qualité du signal et les traitements correctifs à appliquer au spectre pour pouvoir l'étudier. Nous présenterons enfin les paramètres que nous avons choisis pour cette étude et l'ensemble des données produites, qui représente également une des contributions de ce travail.

Dans le chapitre trois, nous discuterons des traitements nécessaires à mettre en œuvre sur les spectres issus de la résonance magnétique pour qu'ils soient analysables par un système automatisé. Après avoir présenté la modélisation mathématique du signal que nous étudions, nous verrons comment nous avons corrigé les principaux problèmes que sont la ligne de base, le problème de phase et les problèmes de normalisation. Nous terminerons par un aperçu des techniques d'estimation de concentrations utilisées.

Le chapitre quatre sera consacré à l'évaluation des performances en classification supervisée. Nous présenterons alors les deux techniques que nous avons utilisées : les machines à vecteurs de support (SVM) et les réseaux de neurones. Nous étudierons notamment l'effet des traitements présentés en chapitre trois sur l'efficacité de la classification.

Dans le chapitre cinq, nous parlerons des techniques de sélection d'attributs mises en œuvre pour détecter la présence de biomarqueurs différents de ceux connus ou pouvant présenter un intérêt pour la classification.

Le sixième chapitre sera consacré au cas des données précliniques, qui, étant donné le faible nombre d'expériences disponibles, ne peut être traité de la même manière que le cas clinique.

Enfin nous conclurons par un rappel des principaux résultats et les perspectives ouvertes par ces travaux.

Chapitre 1 : Le cancer de la prostate : origine et méthodes d'analyses

1.1 La prostate

1.1.1 Anatomie

1.1.1.1 Cas de l'homme

La prostate est une glande de l'appareil génital masculin ressemblant à une châtaigne (3 cm de haut et 2,5 cm de profondeur). Sa masse est estimée entre 15 et 25 g pour un adulte. A la naissance, la prostate est très petite puis elle connait deux périodes de croissance : la première qui l'amène à sa taille normale lors de la puberté, la seconde à partir de 60 ans (dite hyperplasie bénigne de la prostate).

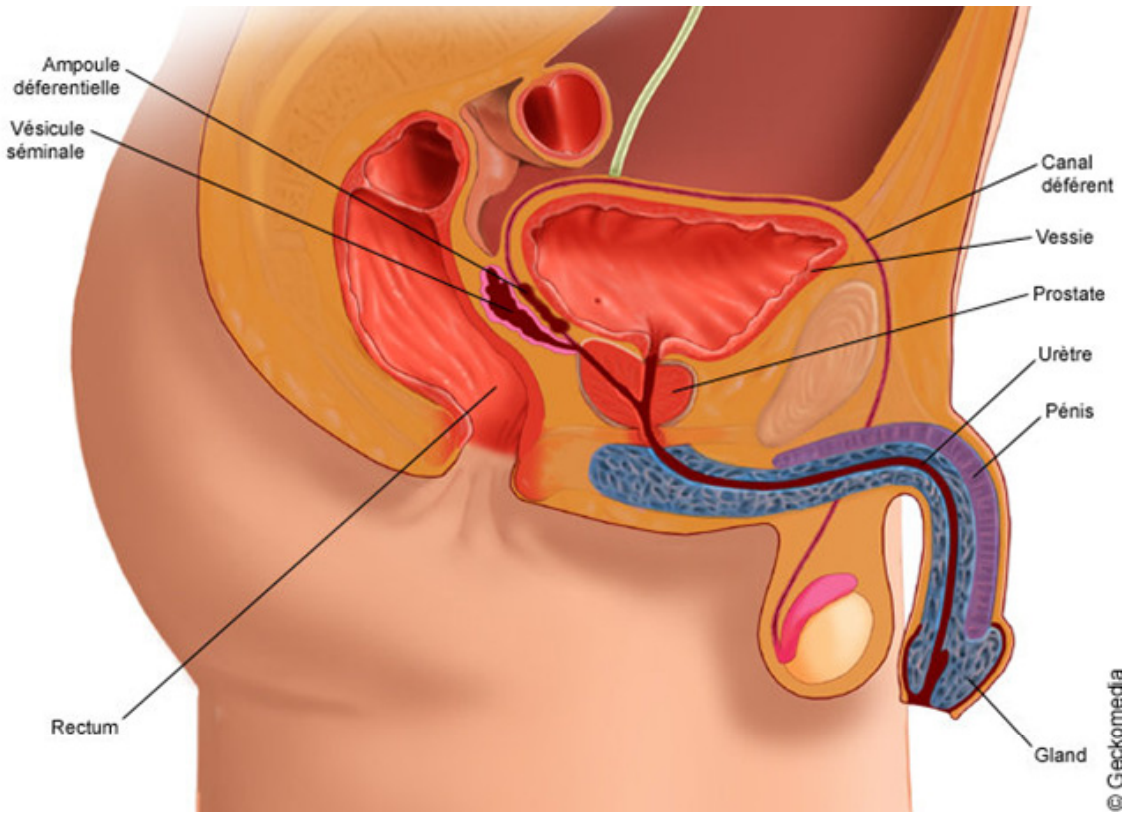


Figure 4 : Coupe sagittale de l'appareil génital masculin.
(source : Logiciel Natom de Geckomedia)

Elle se situe juste sous la vessie, devant le rectum et est traversée par l'urètre (Figure 4). Elle est de forme pyramidale avec une base adossée à la vessie et un sommet, dit apex,

se prolongeant par l'urètre. Dans son prolongement supérieur, entre la vessie et le rectum, se trouvent les vésicules séminales. C'est à l'intérieur de la prostate que se fait la jonction entre l'urètre venant de la vessie, les canaux déférents et les vésicules séminales (Figure 5). Ces dernières se jettent d'abord dans les canaux déférents qui prennent alors le nom de canaux éjaculateurs. Ces canaux vont ensuite se déverser dans l'urètre prostatique. L'ensemble de la glande est entouré d'une membrane fibro-élastique appelée capsule.

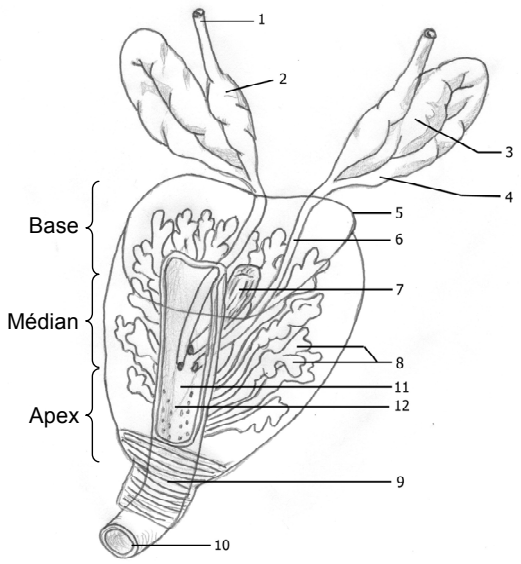


Figure 5 : Représentation de la prostate (vue de 3/4).
1. Conduit déférent, 2. Ampoule du conduit déférent, 3. Vésicule séminale, 4. Conduit excréteur de la vésicule séminale, 5. Contour de la prostate, 6. Conduit éjaculateur, 7. Utricule prostatique, 8. Substance glandulaire, 9. Sphincter de l'urètre, 10. Urètre, 11. Colliculus séminal, 12. Crête urétrale (Dessin de Raziel sur Wikipedia).

Sous la prostate, autour de l'urètre, se trouve le sphincter strié qui, en se contractant, permet la vidange des glandes. La prostate contient également un muscle qui aide à expulser le sperme pendant l'éjaculation.

La prostate et les vésicules séminales participent à la production du fluide séminal, un des constituants du sperme : la prostate à hauteur de 10 à 30%, le reste par les vésicules séminales. Pour fonctionner correctement, la prostate a besoin d'hormones masculines (les androgènes).

Depuis Mac Neal [6], trois zones glandulaires sont ainsi identifiées dans la prostate (Figure 6) :

- Zone centrale
- Zone de transition
- Zone périphérique

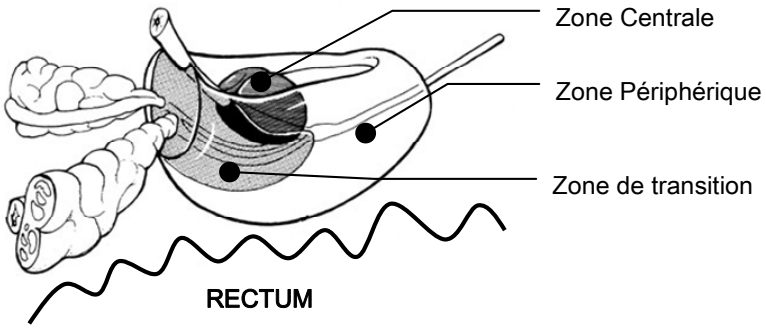


Figure 6 : Anatomie zonale de la prostate.
(extrait de Greene DR, Shabsigh R, Scardino P T.Urologic ultrasonography. In: Walsh P C, Retik A B, Stamey T A et al. eds. Campbell's Urology, 6th edn.Philadephia: WB Saunders, 1992; 342–393).

La zone centrale constitue environ 20 à 25% du tissu glandulaire prostatique. Elle est de forme conique avec le sommet pointant vers l'apex de la prostate. Les canaux séminaux et les ampoules déférentielles pénètrent par sa base et convergent en son sein pour former les canaux éjaculateurs. Elle est peu développée et verra sa part diminuer progressivement avec l'âge et le développement concurrent de la zone de transition.

La zone de transition est constituée de 2 lobes (droit et gauche) symétriques, positionnés de part et d'autre de l'urètre. Elle représente 5% de la prostate chez le sujet jeune mais augmente avec l'âge et avec le développement de pathologie telle l'hyperplasie bénigne de la prostate.

En IRM, il est parfois difficile de distinguer la zone centrale et la zone de transition, c'est pourquoi nous emploierons, dans la suite de ce document, le terme de glande centrale à la fois pour la zone centrale et la zone de transition.

La zone périphérique occupe la plus grande partie du tissu glandulaire et représente 70 à 75% du tissu prostatique chez le sujet jeune.

1.1.1.2 Cas du rat



La prostate d'un rat est en forme de papillon (Figure 7). Contrairement à la prostate humaine, elle est multilobée avec un lobe ventral, un lobe dorsal et un lobe latéral. En IRM, le lobe latéral est indiscernable du lobe dorsal.

Dans le cadre de ses études, Oncodesign utilise des rats dit "nude" caractérisés par une déficience immunitaire. Contrairement à ce qu'il se passe pour certains cancers et certaines espèces animales, ces rats ne développent pas spontanément de tumeur : il faut donc leur implanter chirurgicalement. Compte tenu de la taille de l'implant, cette implantation est réalisée dans le lobe ventral qui a un volume plus important et est plus facile d'accès.

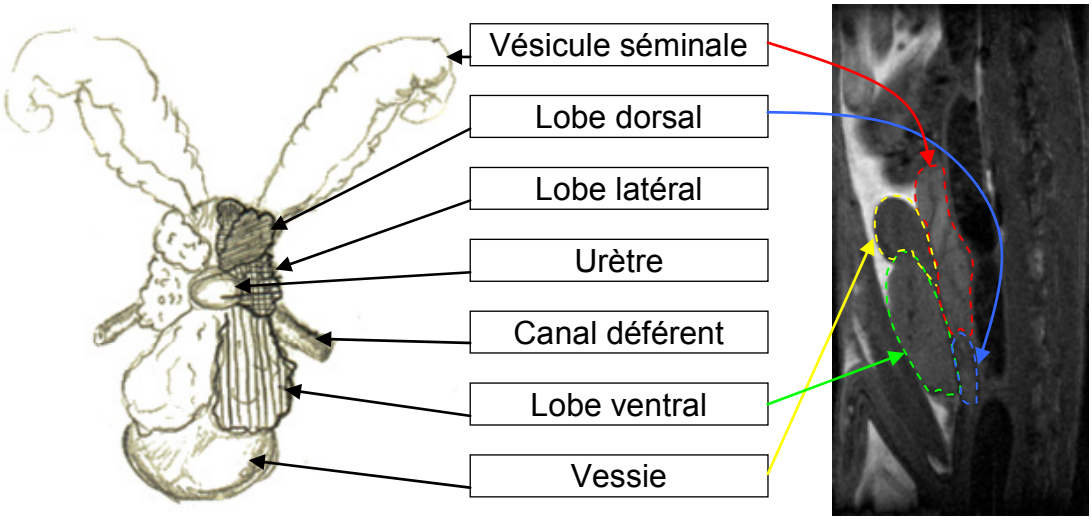


Figure 7 : Prostate de rat.
(dessin de Oliver Putz, PhD Université de l'Illinois, Chicago)

1.1.2 Métabolisme

La prostate est une glande hormono-dépendante qui produit principalement du liquide spermatique.

Le principal anion du liquide prostatique normal est le citrate. La composition ionique du liquide prostatique varie entre les individus, en fonction de l'activité sécrétoire de la glande et de la présence ou de l'absence de maladie inflammatoire prostatique.

En cas de tumeur, les cellules cancéreuses remplacent les cellules prostatiques qui ne remplissent donc plus leurs rôles et la production de citrate diminue fortement, voire disparait. Dans le même temps, un cancer est caractérisé par une forte prolifération cellulaire et donc une forte production de membranes cellulaires dont l'un des constituants de base est la choline (sous sa forme phosphatidylcholine). On observe donc une augmentation de la concentration en choline dans les tumeurs.

Le métabolisme est similaire chez le rat mais nous verrons que la teneur en citrate n'a pas pu être détectée en spectroscopie du proton in vivo avec les appareils que nous avons utilisés.

1.1.3 Pathologies

1.1.3.1 L'hyperplasie bénigne de la prostate

On trouve également les termes d'hypertrophie[1] bénigne de la prostate et d'adénome prostatique. Cette pathologie, fréquente chez l'homme de plus de 60 ans (fréquence de 50% à 60 ans et de 90% à plus de 90 ans), se traduit par une augmentation du volume de la prostate qui entraine principalement des troubles de la miction car en augmentant de volume, la prostate a tendance à comprimer l'urètre passant en son sein. Il s'agit d'une pathologie bénigne, comme son nom l'indique, ne mettant pas en jeu la vie du patient.

1.1.3.2 Le cancer

Le cancer (synonyme : tumeur maligne) est une maladie caractérisée par une prolifération cellulaire indéfinie dont l'évolution spontanée menace la survie de l'individu atteint. Les cellules cancéreuses sont toutes des clones d'une cellule initiatrice du cancer qui a acquis certaines caractéristiques lui permettant notamment donc de se diviser indéfiniment. Au cours de l'évolution de la maladie, certaines cellules peuvent migrer de leur lieu de production et former des métastases. Pour ces raisons, le dépistage du cancer doit être le plus précoce possible.

Les cancers de la prostate se développent préférentiellement dans la zone périphérique (70%) mais peuvent également se développer dans la zone de transition (30%). Les cancers de la zone de transition sont habituellement asymptomatiques et diagnostiqués le plus souvent fortuitement sur du matériel de résection.

Le cancer de la prostate est généralement pluri-focal et correspond, sur le plan histologique, le plus souvent à un adénocarcinome.

Il existe principalement deux méthodes pour évaluer un cancer de la prostate : la classification TNM et le score de Gleason.

La classification TNM (Annexe 1) est le système adopté par l'OMS pour classer les cancers en fonction de leur extension anatomique. Elle a été proposée par Pierre Denoix

[1] Le terme « hyperplasie » désignant une augmentation de volume dûe à une augmentation du nombre de cellules d'un tissu alors que le terme « hypertrophie » désigne une augmentation de volume quel qu'en soit la cause.

dans les années 1950 et est basée sur l'observation clinique. Chaque lettre T, N et M correspond à un site de propagation possible de la maladie cancéreuse : T pour la tumeur initiale, N pour les ganglions (Node) et M pour les métastases. A chacune de ces localisations est affectée une cotation :

- Pour la lettre T : de 0 pour une absence de tumeur à 4 pour une tumeur initiale très étendue.
- Pour la lettre N : de 0 pour une absence d'envahissement ganglionnaire à 2 pour une atteinte de plusieurs ganglions.
- Pour la lettre M : 0 pour une absence de métastase et 1 pour une ou des métastases.

Il existe différentes variantes de cette classification adaptées à chaque cancer (Cf. Annexe 2 pour la prostate).

Le score de Gleason est lui spécifique de la prostate. Il a été décrit pour la première fois par le docteur Donald Gleason et est basé sur l'observation microscopique de la tumeur. Les biopsies ou pièce de prostatectomie sont observées et gradées de 1 à 5 en fonction de la ressemblance entre les cellules prostatiques normales et les cellules cancéreuses. Le score de Gleason est ensuite la somme des deux grades les plus représentés. Le grade le plus représenté est dit primaire et le deuxième plus représenté, secondaire. Un score de Gleason de 7 peut donc être obtenu par plusieurs combinaisons (5+2, 4+3, 3+4 ou 2+5). Plus le score est élevé, plus le cancer est considéré comme agressif et possède la capacité de se développer et de s'étendre rapidement en dehors de la prostate. La combinaison des grades primaire et secondaire permet de nuancer le diagnostic.

1.1.4 Dépistage

1.1.4.1 Toucher rectal

L'hyperplasie bénigne et le cancer entrainent tous deux une augmentation du volume de la prostate. La proximité de la prostate et du rectum peuvent permettre de palper cette augmentation lors d'un examen rectal. Dans certains cas, le médecin peut également sentir la présence de nodule plus dur dans la prostate. Cet examen ne permet en aucun cas un diagnostic de cancer ou d'hyperplasie mais il est simple et peut être pratiqué directement au cabinet du praticien. Il permet d'orienter la suite des investigations à réaliser.

1.1.4.2 Dosage de l'antigène spécifique de la prostate

L'antigène prostatique spécifique (communément appelé PSA pour *Prostate Specific Antigen*) est une protéine sécrétée exclusivement par la prostate. Elle sert principalement à liquéfier le sperme pour faciliter le déplacement des spermatozoïdes. Il semblerait qu'elle aide également à dissoudre la muqueuse cervicale pour favoriser l'entrée du sperme. Elle est présente dans le sang de tous les hommes et l'on peut donc en faire un dosage sérique. Un taux très élevé est associé à une anomalie de la prostate : soit une hyperplasie bénigne de la prostate, soit un cancer de la prostate. Cependant certains facteurs peuvent également engendrer une élévation de ce taux : infection de la prostate, irritation, éjaculation récente, toucher rectal, etc.

Le PSA peut se trouver sous deux formes dans le sang : libre (~10 %) ou lié à une protéine (~90%). Il existe donc deux dosages possibles : PSA libre ou PSA total.

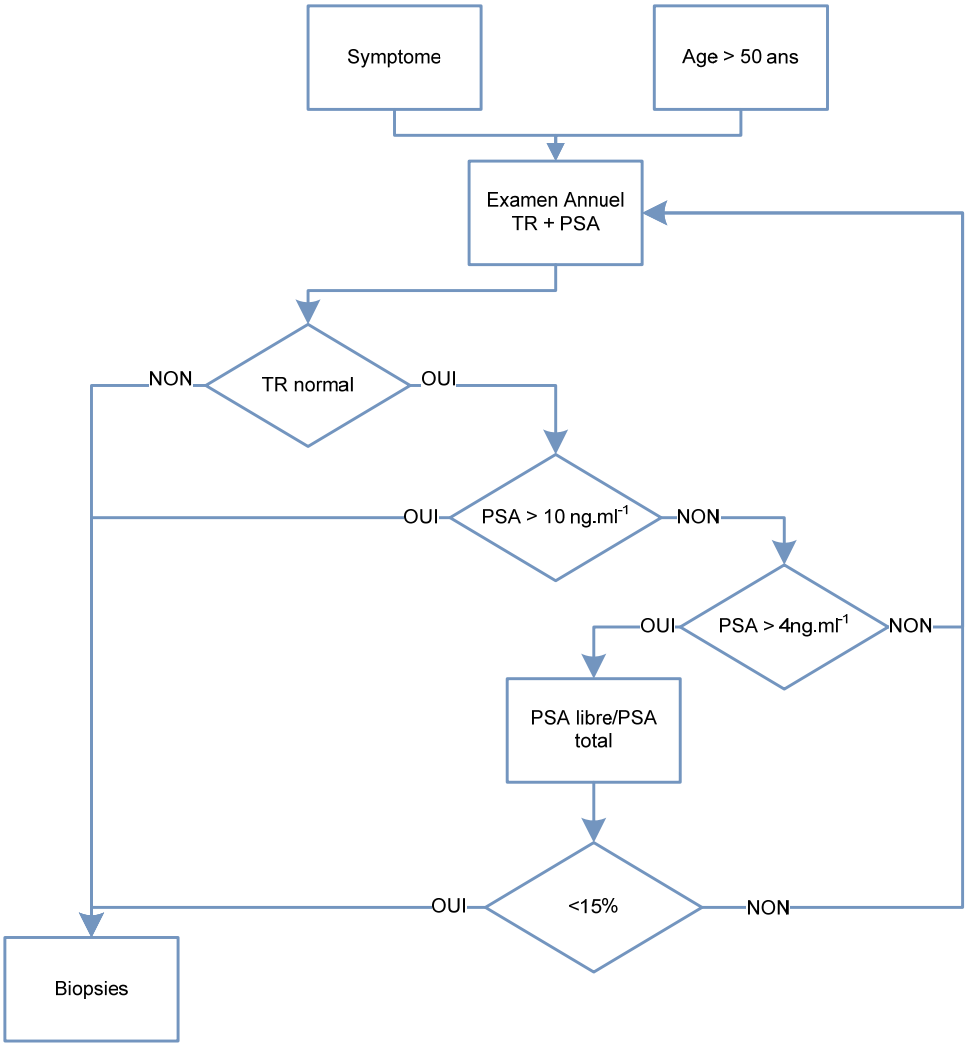


Figure 8 : Schéma de dépistage habituel d'un cancer de la prostate.

Généralement un taux de PSA total est considéré comme pathologique lorsqu'il est supérieur à 10 $ng.mL^{-1}$ et suspect de 4 à 10 $ng.mL^{-1}$. Un taux de 2 $ng.mL^{-1}$ en constante augmentation peut également suggérer un problème. Un schéma permet de résumer le processus de diagnostic en fonction du taux de PSA (Figure 8). Comme le toucher rectal, le dosage du PSA, même si c'est un très bon indicateur, ne suffit pas pour poser le diagnostic.

1.1.4.3 Biopsie

La proximité de la prostate avec le rectum permet de réaliser des biopsies transrectales sous contrôle échographique (TRUS : TransRectal Ultra Sound) (Figure 9).

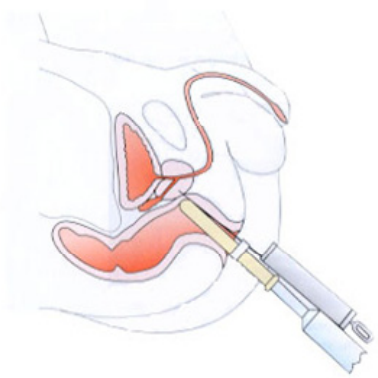


Figure 9 : Biopsie transrectale sous contrôle échographique.
(Illustration extraite du site www.uropage.com)

Cet examen ne nécessite ni hospitalisation ni anesthésie générale ; seule une anesthésie locale peut éventuellement être réalisée, l'examen n'étant pas douloureux mais très désagréable. La prostate est alors divisée artificiellement en 6 zones (apex, médian et base ; gauche/droite pour chaque niveau) et un ou deux prélèvements sont réalisés dans chaque zone. Le médecin anatomopathologiste pose alors le diagnostic définitif en étudiant ces échantillons et peut ainsi donner le score de Gleason.

Les biopsies sont généralement prescrites pour confirmer ou infirmer une suspicion de cancer décelé par toucher rectal ou devant une élévation (ou un fort taux initial) du PSA.

1.1.4.4 Imagerie de résonance magnétique

L'imagerie de résonance magnétique permet d'obtenir de très bons contrastes dans les tissus mous ce qui est le cas de la prostate. Cependant la France manque d'appareils d'IRM, ils étaient donc réservés jusqu'à maintenant à des applications dites plus nobles (cerveau, cœur, …).

A Dijon, une vacation d'IRM est dorénavant consacrée à la prostate. Le grand avantage de l'IRM comparé aux autres méthodes est sa résolution qui permet de voir

l'extension de la tumeur. En cas d'extension extra-capsulaire, le traitement par chirurgie sera évité, n'apportant plus de bénéfice, en termes de survie, au patient.

La résonance magnétique permet d'accéder à différentes modalités d'imagerie : l'imagerie T2 permet d'obtenir des images dont le contraste dépend entre autres de la densité de protons. Dans ce cas, la prostate apparait en gris (plus clair pour la zone périphérique) alors que la tumeur s'avère légèrement plus sombre. L'imagerie de diffusion permet d'obtenir un indice sur le degré de liberté des molécules d'eau : ces molécules étant "moins libres" dans les tumeurs, ces dernières apparaissent plus sombres que le reste de la prostate. L'imagerie de perfusion utilise un agent de contraste : après injection de celui-ci, on capture des images à intervalle régulier, puis on étudie la variation de contraste au cours du temps. Les tumeurs étant hyper vascularisées, on note une hausse plus rapide du signal dans cette zone que dans la zone saine et une élimination plus rapide de l'agent de contraste. Enfin la spectroscopie, comme nous allons le voir dans la suite, permet d'étudier la concentration des molécules participant au métabolisme et comme celui-ci varie entre le tissu sain et le tissu tumoral, on peut distinguer des variations entre les deux tissus. La résolution de la spectroscopie est beaucoup moins fine que les autres modalités utilisées en résonance magnétique (5-6 mm contre 1 mm). Par contre, elle est beaucoup plus spécifique, car d'autres pathologies peuvent provoquer une baisse de signal en imagerie T2 et en diffusion et, avoir les mêmes variations de contraste en perfusion alors que les modifications du signal observées en spectroscopie sont bien spécifiques d'une tumeur.

1.2 Etat de l'art en spectroscopie par résonance magnétique du proton et classification de spectres

1.2.1 Spectroscopie de la prostate

Les premiers essais d'imagerie du cancer de la prostate en résonance magnétique nucléaire ont été réalisés au début des années 1980 [7-10] et cette technique s'est vite révélée comme la technique idéale pour analyser les lésions de la prostate grâce à la très bonne résolution atteinte dans les tissus dits "mous". Les techniques basées sur les rayons X ont un mauvais contraste dans ces zones et les autres techniques (échographie ou médecine nucléaire) ont une résolution inférieure ne permettant pas de détecter les petites lésions.

En 1988, Sillerud [11] a réalisé une des premières spectroscopies par résonance magnétique de la prostate humaine in vivo. Son but était de vérifier la détectabilité du citrate dans la prostate. La spectroscopie permet d'obtenir des informations sur l'activité

métabolique du tissu alors que l'imagerie se contente de donner des informations sur la structure tissulaire, ce qui peut être plus discriminant pour la détection d'un cancer de la prostate.

Plusieurs travaux, menés par Lynch[12], Kurhanewicz [4], Garcìa-Segura [13] et Ackerstaff [1] ont démontré la variation de certains métabolites en fonction des lésions de la prostate. Différentes techniques ont été utilisées : in vivo avec antenne endorectale[4, 13], in vitro à partir d'échantillons de fluide prostatique [12] ou in vitro sur des cultures cellulaires [1]. Des travaux sont toujours en cours concernant la recherche de marqueurs du cancer de la prostate en SRM comme en attestent les publications de Wang, Nagarajan, Balique, et Schmuecking [14-17].

Toutes les publications s'entendent pour établir que les principaux métabolites détectables dans la prostate sont la choline, les polyamines (principalement la spermine), la créatine et le citrate. Des études réalisées in vitro à très haut champ (11,7T) ont prouvé l'existence d'autres marqueurs statistiquement significatifs : Lynch [12] en travaillant sur des échantillons de fluide prostatique a mis en évidence le myo-inositol tandis que Tessem [18] a mis en avant le lactate et l'alanine en étudiant des biopsies de la prostate. Cependant, compte tenu des paramètres d'acquisition que nous avons utilisés, nous ne nous sommes pas placés dans de bonnes conditions pour quantifier ces molécules.

A partir de données issues de SRM de la prostate, un certain nombre de chercheurs a travaillé à classer automatiquement ces spectres, suivant en cela les démarches entamées auparavant dans le cadre de la classification de spectres du cerveau.

Les méthodes employées étant similaires, nous commencerons par un bref survol des méthodes mises en œuvre pour le cerveau avant de détailler les quelques travaux existant qui concernent la prostate.

1.2.2 Classification de spectres

1.2.2.1 Cas du cerveau

De 2000 à 2002, une grande base de données multicentrique de spectres cérébraux appelée INTERPRET [19] a été constituée. Le but de cette expérience était de faciliter l'utilisation de la SRM afin d'améliorer le diagnostic et le traitement des cancers du cerveau, car même si les tumeurs présentent des profils spectroscopiques significativement différenciés, l'interprétation de ces données par les radiologues/spectroscopistes reste très

difficile. Cette expérience regroupe plusieurs institutions européennes[2]. Cette base de données de spectres est caractérisée par l'hétérogénéité des données qu'elle contient : ils ont été acquis avec des matériels de différentes marques (General Electric, Siemens ou Philips), avec des séquences différentes (STEAM ou PRESS) et des paramètres d'acquisition variés (temps d'écho, temps de répétitions …). Nous verrons en détail ces différents paramètres dans une prochaine partie (Chapitre 1). Plus de 70 publications sont issues des travaux associés à ce projet. Nous avons choisi de résumer les méthodes qui nous ont paru les plus significatives.

En 2001, Tate et al [20] ont comparé deux méthodes de correction de phase et deux méthodes de sélection d'attributs. Pour cela, ils ont choisi dans la base de données INTERPRET les examens de 51 patients tous acquis sur la même machine (1,5T GE SIGNA). Les patients étaient atteints soit de méningiomes (12), soit de tumeurs dites de "bas grade" (7 oligodendrogliomes et 6 astrocytomes de grade II) soit de glioblastomes (26). Les examens ont tous été réalisés avec deux temps d'écho (20 et 136 ms). Pour chaque examen, les auteurs ont d'une part calculé le module des spectres et d'autre part corrigé la phase de ces spectres avec l'algorithme QUALITY [21]. Cet algorithme est basé sur la déconvolution du signal temporel par une fonction de référence, permettant d'obtenir après application de la transformée de Fourier, un signal avec des pics respectant la forme lorentzienne théorique. Ensuite, avec les spectres issus de chacun des deux traitements précédents, les auteurs ont appliqué deux méthodes de sélection d'attributs : la première [22] consiste à choisir les attributs dont le coefficient de corrélation entre leur valeur et l'index de la classe correspondante est maximum. La seconde consiste à calculer l'analyse en composantes principales (ACP) [23] sur les spectres. Cette technique repose sur la transcription des valeurs des spectres dans un nouveau référentiel où chaque axe maximise la variance des données (le principe sera détaillé dans la partie 5.1). Les auteurs ont conservé les 4 premières composantes principales (les coordonnées selon les 4 premiers axes) et ont mesuré les taux d'erreur obtenus avec l'analyse discriminante linéaire (LDA :

[2] Universitat Autònoma de Barcelona (Barcelone, Espagne), Saint George's Hospital Medical School (Londres, Royaume Uni), Institut de Diagnòstic per la Imatge, Centre Bellvitge, L'Hospitalet de Llobregat (Barcelone Espagne), Centre Diagnòstic Pedralbes, Cetir Grup Mèdic (Barcelone, Espagne), Unité mixte Université Joseph Fourier/INSERM U594 (Grenoble, France), Akademische Ziekenhuis Nijmegen. Katholieke Universiteit Nijmegen (Nijmegen, Pays-bas), Katholieke Universiteit Nijmegen (Nijmegen, Pays-bas), University of Sussex (Brighton, Royaume Uni), SARL PRAXIM (La Tronche, France) et Siemens AG, Medical Engineering (Erlangen, Allemagne)

Linear Discriminant Analysis) et ce pour différentes combinaisons de traitements. Le meilleur taux de classification est obtenu avec les spectres à 136ms avec les quatre composantes principales pour séparer les tumeurs de bas grade des méningiomes avec un taux de réussite de 96%. Les auteurs obtiennent un taux de réussite de seulement 74% en classant les glioblastomes versus les tumeurs de bas grade avec les quatre composantes principales, quel que soit le temps d'écho. En conclusion, dans le cadre de cette étude, le module permet d'obtenir de meilleurs taux de classification que la correction de phase. Les deux temps d'écho donnent des résultats comparables.

En 2003, Tate et al. ont publié une nouvelle étude [24] dans laquelle ils ont démontré la relative robustesse de la classification aux variations des paramètres d'acquisition. Les auteurs ont utilisé pour cela 144 spectres correspondant à des tumeurs (méningiome, astrocytome de bas grade, glioblastome et métastase) confirmées par histopathologie. Ces spectres ont été acquis dans des conditions variables en termes d'appareil (GE et Philips), de séquence (STEAM et PRESS), de temps d'écho (20, 30, 31 et 32 ms) et de temps de répétition (1600 et 2000 ms). Avant de calculer les spectres, les signaux acquis ont d'abord subit une correction des effets des courants de Foucault en utilisant l'algorithme QUALITY [21] et une suppression du pic d'eau résiduel réalisée grâce l'algorithme HLSVD (Hankel-Lanczos Singular Value Decomposition [25]). Cette technique est basée sur la décomposition en valeurs singulières qui permet de paramétriser les pics du signal de RMN. L'algorithme de décomposition en valeurs singulières est optimisé en utilisant l'algorithme de Lanczos et la structure des matrices de Hankel. Après passage dans le domaine fréquentiel, les auteurs ont calculé le module des spectres. A noter que pour ceux issus de machine Philips, présentant une plus faible résolution spectrale, une simple interpolation linéaire a été utilisée pour ramener la résolution au niveau de celle des appareils GE. Tous les spectres ont alors été normalisés en les divisant par la norme L2 (somme des carrés de chaque intensité). Les spectres ont alors été séparés en deux groupes : le premier, contenant les spectres issus des centres de Barcelone, a été utilisé pour l'apprentissage et le second, contenant les spectres acquis à Londres, a servi pour la phase de test. Le classifieur utilisé était la LDA et 48 sur les 50 spectres de l'ensemble de test ont été correctement classés. Les auteurs en ont conclu que leur système de classification était robuste aux variations des paramètres d'acquisition.

Les membres du laboratoire SCD-SISTA de la Katholieke Universiteit Leuven, (Louvain, Belgique) ont réalisé différentes expériences [26-28] sur la base de données INTERPRET pour vérifier l'incidence des prétraitements et du choix du classifieur sur les

performances de classification. A chaque fois, les auteurs ont sélectionné un ensemble de données parmi la base INTERPRET soit à temps d'écho court (≈30 ms), soit à temps d'écho long (≈135ms) pour des temps de répétition variant de 1500 ms à 2000 ms. Les différents types de tumeurs à classer étaient des glioblastomes, des méningiomes, des métastases ou des astrocytomes. Les différents classifieurs testés étaient la LDA et les machines à vecteurs de support (SVM [29]). Le principe de base des SVM consiste à maximiser la distance entre les ensembles de points de chaque classe mais il sera approfondi dans le chapitre 4.1. Les auteurs ont utilisé deux versions des SVM : la première, linéaire, et la seconde, non linéaire avec un noyau de type Radial Basis Function (RBF). Les différents traitements du spectre avant son introduction dans les classifieurs sont :

- La normalisation : les auteurs ont comparé la normalisation par l'eau (chaque valeur du spectre est divisée par une estimation de la concentration en eau) à la normalisation L2.
- La correction de phase : les auteurs ont employé le module du signal et l'ont comparé à la partie réelle du spectre rephasée par la méthode de Klose [30]
- La correction de la ligne de base : les auteurs ont utilisé un algorithme basé sur les propriétés de la transformée de Fourier en extrayant le signal basse fréquence du spectre. (Cf. 3.2.2)

Les auteurs ont aussi utilisé différents ensembles d'attributs en entrée des classifieurs : le spectre complet, certaines portions du spectre connues pour correspondre aux métabolites significatifs du cerveau. Enfin, pour le dernier ensemble, les auteurs ont calculé l'intégrale (par la méthode des trapèzes) de chaque zone de l'ensemble précédent (13 attributs).

Les auteurs en concluent que les SVM avec un noyau RBF obtiennent dans la plupart des cas de meilleurs résultats que les techniques linéaires. Cependant les techniques linéaires comme la LDA peuvent aussi donner de très bons résultats à leur problème ce qui amènent les auteurs à penser que seule une petite partie des données n'est pas linéairement séparable.

Dans leurs conclusions, les auteurs notent également que les SVM permettent d'avoir de très bons taux de classification avec des données de grande dimension alors que la LDA nécessite une réduction de dimensions (en l'occurrence, l'ACP). De plus les deux techniques de réduction de dimensions employées, bien que pouvant donner de bons

résultats (moins de 10% d'erreur), fournissaient en moyenne de moins bons résultats de classification que le spectre complet.

Enfin les auteurs constatent que le meilleur taux de classification est obtenu avec le module du spectre sans correction de ligne de base avec la normalisation L2. Concernant le temps d'écho, bien que les spectres à temps d'écho court contiennent plus d'informations, ce qui est dû à la disparation du signal des métabolites à T2 long avec l'augmentation du temps d'écho, les taux de classifications sont similaires.

Ceci n'est qu'un petit aperçu des travaux qui ont été réalisés sur la classification de spectres de résonance magnétique du proton du cerveau. Cependant il est possible d'en retirer quelques points clefs. D'une part, les techniques linéaires comme la LDA associée à l'ACP peuvent donner de très bons résultats, cependant les SVM sont souvent plus efficaces notamment si les données appartiennent à un espace de grande dimension. D'autre part, la normalisation des spectres par l'intégrale du pic d'eau ne permet pas d'obtenir de meilleurs résultats de classification que la normalisation par la norme L2 du spectre. Ensuite la suppression de la ligne de base, avec la méthode utilisée, n'améliore pas les résultats de classification. Enfin l'utilisation du module offre d'aussi bons, voir de meilleurs taux de classification que la partie réelle avec une correction de phase.

On peut noter que dans les études précédentes, les auteurs effectuaient systématiquement une suppression du signal résiduel de l'eau soit dans le domaine temporel, soit dans le domaine fréquentiel. Cependant avec les évolutions du matériel, ce signal est dorénavant très bien supprimé dès l'acquisition.

Toutefois il faut noter que ces différentes conclusions ont représenté pour nous des indications ou tendances générales, mais ne sont pas forcément directement applicables dans le cas de la prostate.

1.2.2.2 Cas de la prostate

Beaucoup d'études tendent à prouver qu'il existe une corrélation entre la concentration des métabolites détectés en SRM du proton et la malignité du voxel étudié, mais peu d'équipes ont tenté de développer un système de classification automatique.

En 2007, Kelm et al. [31] ont réalisé une des premières études de classification automatique de spectres de résonance magnétique du proton de la prostate in vivo. Les auteurs ont travaillé sur une cohorte de 36 patients auxquels ils ont fait passer une spectroscopie de la prostate sur un scanner 1,5T avec une antenne endorectale. Sur ces 36 examens, 12 ont été éliminés car les spectroscopistes ont considéré que les données étaient

de mauvaise qualité (champ magnétique pas assez homogène, suppression des graisses inefficace, problème de positionnement de l'antenne endorectale, …). Le classement de chaque spectre a été déterminé selon des règles standards basées sur les pics de choline, de créatine et de citrate [32, 33]. Les intitulés des différentes classes étaient "healthy", "undecided" et "tumor".

Les classifieurs évalués sont la régression logistique [23], la GPLS (Generalized Partial Least Square) [34] et la PSR (P-spline signal regression) [35] pour les techniques linéaires et les SVM, les "Gaussian processes" [23] et "random forest"[36] pour les méthodes non linéaires. La régression PLS maximise la variance des prédicteurs tout en maximisant la corrélation entre ceux-ci et la variable à expliquer. Cette technique s'inspire de l'ACP et des méthodes de régression. La GPLS permet de calculer simultanément une régression logistique et la PLS. La PSR est une technique conçue spécialement pour l'étude de spectres. La méthode est basée sur la connaissance *a priori* que deux valeurs du spectre voisines sont corrélées et qu'il est possible de les modéliser par des fonctions "spline" cubiques.

La méthode des "Gaussian process" consiste à effectuer un changement d'espace de représentation des données, par l'intermédiaire d'une fonction noyau, comme nous le verrons dans le chapitre 4.1.2. pour les SVM non linéaires. Les probabilités de distribution sont alors estimées dans le nouvel espace.

"Random forest" consiste à générer plusieurs arbres de décision légèrement différents. La décision est prise par la majorité des votes de chaque arbre.

Les auteurs ont voulu comparer des méthodes d'estimation de concentrations et des méthodes de "reconnaissance de forme". Les techniques d'estimation de concentrations relatives utilisées sont : QUEST [37], AMARES [38] et VARPRO [39], toutes trois implémentées dans le logiciel jMRUI [40]. Ces trois méthodes effectuent une modélisation du signal temporel par différentes fonctions (signaux des métabolites recherchés simulés selon un algorithme basé sur la mécanique quantique pour QUEST, sinusoïdes amorties par une fonction exponentielle décroissante pour AMARES) et avec différents algorithmes de résolution du système ainsi obtenu.

Pour les méthodes de "reconnaissance de forme", les attributs utilisés en entrée des classifieurs sont le module et la partie réelle des spectres entre 2,36 et 3,34 ppm après correction de phase de zéro ordre. Les données sont alors normalisées par la somme des valeurs absolues de chaque attribut du spectre (norme L1).

Les auteurs ont évalué quatre méthodes de réduction de dimension appliquées à ces données : l'ACP, les "partial least square" (PLS) [41], l'analyse en composantes indépendantes (Independent Component Analysis : ICA) [42] et la "nonnegative matrix factorization" (NMF) [43]. L'ICA a pour objectif de transposer les données dans une nouvelle base où chaque composante est aussi linéairement indépendante que possible des autres.

Les tests de classification ont été réalisés deux à deux (healthy vs tumor, healthy vs undecided et Tumor vs undecided) et la performance en classification a été mesurée par l'aire sous la courbe ROC (Receiver Operator Characteristic) obtenue en traçant le complément à 1 de la spécificité en fonction de la sensibilité. Les expériences ont été réalisées en validation croisée "leave one patient out", c'est-à-dire qu'à chaque itération, les données d'un patient servent d'ensemble de test alors que les données de tous les autres patients servent d'ensemble d'apprentissage.

Les auteurs ont donc étudié une grande partie des combinaisons possibles entre toutes ces méthodes. Il en ressort que :

- Les méthodes d'estimation de concentrations sont très dépendantes des paramètres utilisés.
- Le meilleur résultat est atteint avec les classifieurs non linéaires appliqués au module du spectre.

Matulewicz a présenté ses travaux en 2009 à l'ISMRM (International Society of Magnetic Resonance in Medecine)[44]. Les données sont issues de dix patients ayant eu une SRM du proton avec antenne endorectale à 1,5T, suivie d'une prostatectomie totale. Un spectroscopiste a attribué à chaque voxel une étiquette entre sain et tumeur. Les voxels classés comme « tumeur » ont été confirmés par histopathologie en faisant correspondre chaque coupe spectroscopique avec les coupes d'anatomopathologie. Etant donné la déformation des coupes anatomopathologiques, les voxels sont localisés au sextant près, c'est-à-dire en divisant chaque coupe en secteur gauche/droite et apex/médian/base.

Après une suppression du bruit par un filtrage gaussien, les spectres sont recalés en fréquence. Les spectres sont alors restreints à la plage 0,6-3,6 ppm puis traités avec la méthode "Orthogonal Signal Correction" (OSC) [45] basée sur l'ACP et qui doit supprimer l'information la moins significative des spectres. Une classification automatique est alors réalisée sur ces données en validation croisée. Les auteurs obtiennent un taux de

réussite de 80,4% mais remarquent que le résultat dépend fortement de la qualité du recalage effectué *a priori*.

Depuis 2007, Tiwari a réalisé un important travail sur la classification de spectres en résonance magnétique de la prostate concrétisé par plusieurs publications [46-49]. Nous ne considérons ici que le dernier article où l'auteur obtient ses meilleurs résultats. Les données sont issues d'une base multicentrique initiée par l'American College of Radiology Imaging Network (ACRIN) et accessible à ses membres. Le souhait de l'auteur était de ne pas utiliser d'algorithme d'estimation de concentration des métabolites car il considère cet exercice très difficile étant donné la qualité des spectres (faible rapport signal sur bruit, ligne de base irrégulière, recouvrement de pics, …). L'algorithme est décomposé en deux phases, la première consistant à séparer les spectres obtenus dans la prostate des spectres extérieurs à la prostate et la seconde devant classer les spectres entre "normal", "suspicieux" et "indéterminé". La "vérité de terrain" a été définie statistiquement à partir de biopsies, du volume estimé de la tumeur et du quadrant prostatique dont le voxel était issu.

Dans les deux phases, l'auteur utilise un algorithme de classification non supervisé, appelé "replicated clustering", basé sur les k-means [50]. La méthode des k-means permet de séparer un ensemble de données en k sous ensembles en les regroupant selon leur distance au barycentre de chaque sous ensemble. La méthode du "replicated clustering" consiste à lancer plusieurs fois l'algorithme des k-means avec une initialisation aléatoire et différente à chaque fois. A l'issue de cela, le regroupement conservé est celui qui minimise la variance intra-groupe.

La première étape consistant à sélectionner les voxels localisés dans la prostate est réalisée à partir des spectres, dans le sous espace généré par l'algorithme "graph embedding" (GE) [51]. La méthode "replicated clustering" est alors appliquée pour séparer l'ensemble de données en deux groupes. Sachant que la plupart des voxels de l'acquisition se situent en dehors de la prostate, l'ensemble contenant le moins d'éléments est conservé. Ce processus est itéré jusqu'à ce que l'ensemble résiduel contienne autant de spectres qu'il peut y en avoir dans une prostate (seuil généralement atteint en 3 itérations). L'auteur atteint un taux de réussite de 89,29% (sensibilité = 97,66% et spécificité = 98,87%). Cet ensemble de spectres reconnus comme étant dans la prostate est alors traité au second niveau pour détecter les voxels tumoraux.

Dans la seconde phase, l'auteur a utilisé trois techniques de réduction de dimension qui sont l'ACP, la "Locally linear embedding" (LLE) [52] et GE. La LLE est une technique qui consiste dans un premier temps à estimer les coordonnées de chaque point dans la base

de ses plus proches voisins puis à rechercher un nouveau système de coordonnées de plus faible dimension préservant au maximum les distances avec les proches voisins. Après avoir appliqué ces techniques de réduction de dimension aux spectres, l'auteur a utilisé sa méthode "replicated clustering" pour séparer les voxels en trois groupes ("normal", "suspicieux" et "indéterminé"). Chaque groupe est alors considéré, tour à tour comme étant celui des voxels tumoraux, les deux autres groupes étant considérés comme sains. L'auteur a alors calculé le taux de vrais positifs (VP) et de faux négatifs (FN). Le groupe retenu comme étant celui des voxels tumoraux est alors celui qui maximise la proportion de VP, tout en minimisant les FN. L'auteur obtient une sensibilité allant de 79,53% avec l'ACP à 89,33% avec les GE et une spécificité allant de 79,79% pour les GE à 83,97% pour l'ACP.

L'auteur a également comparé ses résultats avec une classification manuelle réalisée par un spectroscopiste qualifié. Les spectres ont été étiquetés avec une valeur comprise entre 1 (normal) et 5 (très suspicieux). Etant donné que la tache est fastidieuse, la base de données a été réduite aux spectres de six patients. Pour la comparaison avec le système automatique, les spectres considérés comme normaux étaient ceux marqués 1 et 2, les autres étant considérés comme suspicieux. Les valeurs de sensibilité vont alors de 76,68% pour l'ACP à 81,36 % pour la LLE. Les valeurs de spécificité vont de 48,26% pour l'ACP à 65,49% pour le GE. Pour cette dernière étude, l'auteur a également calculé pour chaque spectre le z-score [53] qui donne une estimation pour chaque spectre, son degré de déviation par rapport à la moyenne des spectres. La classification se fait alors en fonction d'un seuil prédéfini. Cette méthode est beaucoup moins performante puisque la sensibilité n'est que de 62,70% et la spécificité de 50,73%.

1.2.3 Conclusion

Bien que le cancer de la prostate soit l'un des plus fréquents chez l'homme et qu'il existe beaucoup d'études tendant à prouver la présence d'un lien fort entre la SRM du proton de la prostate et la survenue d'un cancer, il n'y a à l'heure actuelle que peu d'études sur la classification automatique de ces données. Pourtant l'étude manuelle, par un spectroscopiste, des voxels un par un est très fastidieuse et les biopsies sont désagréables et ne sont pas fiables à 100%.

De nombreuses études ont montré qu'une classification automatique pouvait être réalisée sur des spectres de cerveau avec une très bonne fiabilité, puisque les auteurs de ces études arrivent même à distinguer certaines variétés de tumeur.

Cependant la spectroscopie de la prostate n'est pas aussi simple que celle du cerveau : l'environnement est beaucoup moins homogène ce qui perturbe la SRM et provoque des chevauchements de pics, des décalages en fréquence du spectre …

Une autre difficulté, relevée notamment par Tiwari [48] est d'obtenir la "vérité de terrain". Les deux méthodes qui pourraient être considérées comme telles sont les biopsies et les coupes anatomopathologiques. La localisation des biopsies est imprécise et il est compliqué de faire correspondre un voxel avec une biopsie. La correspondance est encore plus difficile dans le cas des coupes anatomopathologiques : lors de la prostatectomie, la prostate est déformée puisqu'elle n'est plus contenue par tous les tissus l'environnant. Il est alors très difficile pour l'anatomopathologiste de réaliser des coupes parallèles aux coupes d'imagerie et il devient alors compliqué de faire correspondre une coupe histologique avec une coupe spectroscopique. Une des solutions les plus simples est de se baser sur l'expertise humaine : un spectroscopiste entrainé peut, en utilisant des critères précis, classer chaque spectre mais cela se révèle très fastidieux.

Les travaux de classification de la prostate étant relativement récents, nous nous sommes principalement appuyés sur les travaux réalisés sur le cerveau bien que, comme nous le verrons par la suite, les résultats ne sont pas toujours transposables. En effet dans notre cas, une infime variation des paramètres d'acquisition peut avoir des effets importants sur le spectre et donc sur les résultats de la classification.

Chapitre 2 : La spectroscopie de résonance magnétique et l'optimisation des paramètres d'acquisition pour l'obtention des données.

Dans [24], Tate affirme que les techniques de classification sont robustes aux variations de paramètres d'acquisition. Cependant cela doit être nuancé car pour être performantes, les méthodes de classification ont besoin de données de très bonne qualité et homogènes. Or dans la prostate, une petite variation des paramètres d'acquisition peut avoir des conséquences très importantes sur les spectres. Pour obtenir des spectres de qualité optimale, nous devons donc comprendre les phénomènes mis en jeu lors de l'acquisition des données en résonance magnétique du proton. C'est un phénomène très complexe expliqué rigoureusement par la physique quantique. Une explication est cependant possible en se restreignant à la physique classique. Cette deuxième voie est largement suffisante pour comprendre les variations du signal que nous voulons étudier.

2.1 La résonance magnétique

La résonance magnétique est un phénomène quantique portant sur le couplage entre le moment magnétique du noyau de certains atomes et un champ magnétique externe. Dès 1938, Isidor Isaac Rabi découvre qu'un noyau atomique absorbe le rayonnement électromagnétique à une fréquence spécifique en présence d'un fort champ magnétique mais le phénomène n'est décrit qu'en 1946 par Felix Bloch et Edward Mills Purcell (Prix Nobel de Physique 1952). A partir de là, les premiers spectrographes vont être développés pour l'analyse physico-chimique de petits échantillons.

Dans les années 1970, la résonance magnétique nucléaire (RMN) va connaitre un important développement d'applications en imagerie médicale. En 1969, Raymond Vahan Damadian propose d'utiliser la spectroscopie de RMN pour la détection des cancers. En 1973, Paul Lauterbur et Peter Mansfield mettent au point, simultanément mais indépendamment, une technique d'imagerie à partir de la RMN en s'inspirant des méthodes de reconstruction d'images utilisées en tomodensitométrie (Prix Nobel de Physiologie et de Médecine 2003). Lauterbur baptise sa technique la "zeugmatographie" (du grec "zeugma" : ce qui relie) qui lui permet d'obtenir la première image en coupe de deux tubes à essai

remplis d'eau [54]. Mansfield réalisera la première image de tissus humains en 1975. En 1977, Damadian produira la première coupe du corps humain. Ce bond en avant dans les années 1970 a été favorisé par les progrès réalisés en informatique : augmentation de puissance des processeurs et développement de nouvelles méthodes numériques permettant d'améliorer les temps de calcul. Ainsi en 1975, Richard E. Ernst utilise la Transformée de Fourier Rapide (FFT : Fast Fourier Transform) pour analyser le codage en phase et en fréquence du signal de RMN.

Dans les années 1980, alors que la spectroscopie était, jusque là, effectuée exclusivement in-vitro, les premières expériences de spectroscopie in-vivo vont être réalisées avec la montée en puissance des aimants (bobines supraconductrices).

Aujourd'hui, l'IRM est une technique d'imagerie très employée du fait de l'absence de rayons ionisants, de la finesse de résolution obtenue et de la variation des contrastes qu'elle permet d'obtenir. La spectroscopie commence à être utilisée en routine pour l'étude in situ du métabolisme des organes.

2.1.1 Principe de la résonance magnétique

2.1.1.1 Notions de physique

La résonance magnétique nucléaire est un phénomène complexe qui nécessite quelques notions de mécanique quantique notamment la notion de "spin". Le spin est une propriété quantique des particules subatomiques (électrons, protons et neutrons) que l'ont peut se représenter comme la rotation autour d'un axe de ces particules. Dans la plupart des atomes, comme le carbone 12 (^{12}C), ces spins sont associés par paires et s'annulent. Dans certains atomes, comme l'hydrogène (^{1}H) ou le carbone 13 (^{13}C), la symétrie est brisée et il apparait alors un spin global au noyau.

Les noyaux avec un nombre de spins (noté I) non nul peuvent absorber et émettre des radiations électromagnétiques. Ce nombre de spins dépend du nombre de masse (A) et du numéro atomique (Z) de l'élément X (noté alors : ${}^{A}_{Z}X$)

- Si les deux sont pairs alors le nombre de spins est nul (ex : ${}^{12}_{6}C$, ${}^{16}_{8}O$)
- Si le nombre de masse est pair et que le numéro atomique est impair alors le nombre de spins est un entier (ex : ${}^{14}_{7}N$, ${}^{2}_{1}H$)
- Si le nombre de masse est impair alors le nombre de spins est un demi-entier (ex ${}^{1}_{1}H$, ${}^{13}_{6}C$, ${}^{15}_{7}N$, ${}^{31}_{15}P$ …)

L'état du spin d'un noyau (m) peut être quantifié. La mécanique quantique nous dit qu'un noyau ayant un nombre de spins I a 2I+1 orientations possibles :

m=I,(I-1), (I-2), … (de I à –I)

Ainsi un noyau ayant un nombre de spins de 1/2 aura deux états de spin différents possibles m=1/2 et m=-1/2 (Figure 10). C'est le cas des principaux noyaux utilisés en résonance magnétique et notamment de celui d'hydrogène, nous ne nous intéresserons donc par la suite qu'à ce cas.

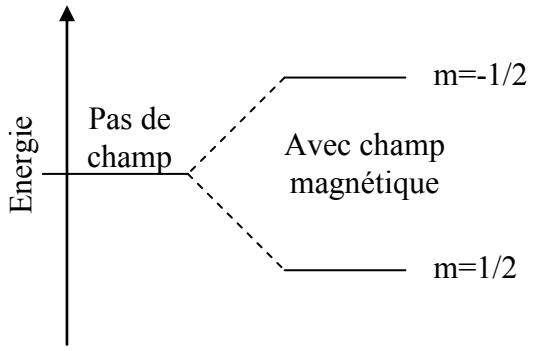


Figure 10 : Niveaux d'énergie des populations de spins avec et sans champ magnétique pour un noyau avec un nombre de spins de ½.

En l'absence de champ magnétique externe, ces deux orientations ont la même énergie. En présence d'un champ magnétique, tous les spins s'orientent selon ce champ (ce n'est pas exact mais nous reviendrons sur cette notion plus loin). La population de spins correspondant à m=1/2 s'oriente dans le même sens que le champ magnétique alors que les spins correspondant à m=-1/2 s'orienteront dans le sens inverse. Chaque orientation correspondra alors à un niveau d'énergie différent et le niveau d'énergie le plus faible contient légèrement plus de noyaux. Il est possible d'exciter la population d'énergie la plus faible pour la faire passer au niveau d'énergie supérieur avec une onde électromagnétique. La fréquence de cette onde est déterminée par la différence d'énergie entre les deux niveaux.

Chaque population de spins différents possède un moment magnétique μ donné par :

$$\mu = \frac{\gamma I \hbar}{2\pi}$$

où h est la constante de Planck et γ le rapport gyromagnétique spécifique à chaque espèce nucléaire. En fonction du rapport gyromagnétique et de l'abondance naturelle (C) de chaque noyau (pourcentage en nombre de noyaux de chaque isotope d'un élément), on peut calculer sa réceptivité relative, c'est-à-dire un indice de la quantité de signal fournie par

ce noyau en résonance magnétique (Tableau 1). La réceptivité d'un noyau est donnée par la formule :

$$D = \left|\gamma^3 C\right|$$

La réceptivité relative d'un isotope est obtenue en divisant sa réceptivité par celle du proton (^1_1H). Le noyau d'hydrogène est celui qui permet d'obtenir le plus de signal en résonance magnétique.

Tableau 1 : Réceptivité relative des principaux noyaux utilisés en résonance magnétique.

Isotope	γ (10^7 rad $T^{-1}s^{-1}$)	Abondance naturelle [C]	Réceptivité relative [D]	Fréquence (MHz) à 1,5T
^{1}H	26,752	99,985	1,0000	63,8
^{13}C	6,728	1,108	0,000176	16,0
^{19}F	25,181	100	0,834	60,1
^{31}P	10,841	100	0,0665	25,9
^{23}Na	7,080	100	0,0185	16,9

Le moment magnétique est un vecteur qui donne la direction et l'intensité de l'aimantation du noyau. Chaque noyau possède donc une aimantation différente. Ainsi l'énergie du spin dans un champ magnétique dépend de ce champ magnétique (noté B_0) et de μ. Pour le cas qui nous intéresse (I=1/2), il y a donc deux orientations possibles de μ : soit μ est dans le sens de B_0, soit μ est opposé à B_0 (Figure 11). L'énergie correspondant à chaque orientation est donnée par $E = -\mu.B_0$

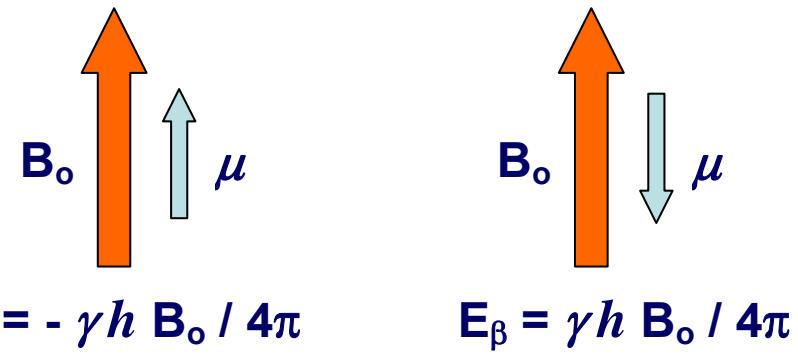


Figure 11 : Energie des spins en fonction de leur orientation par rapport au champ magnétique B_0.

La différence d'énergie entre ces deux niveaux, notée α et β, est donc $\Delta E = \frac{\gamma h B_0}{2\pi}$ et est croissante avec l'intensité du champ (Figure 12).

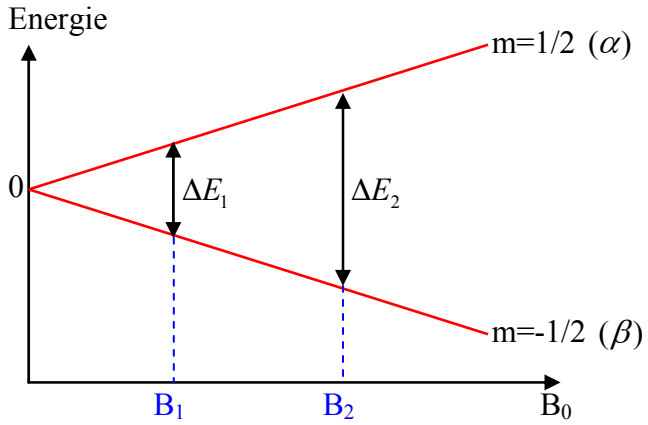


Figure 12 : Variation de la différence d'énergie entre les deux niveaux en fonction de l'intensité du champ magnétique B_0.

Donc plus B_0 est grand, plus la différence d'énergie entre les deux populations est grande. C'est cette différence d'énergie qui sera finalement à l'origine du signal mesuré dans les appareils d'imagerie à résonance magnétique.

On peut remarquer que même avec un champ magnétique de très forte intensité, le signal mesuré est très faible. Considérons un champ magnétique de 9,4T (200 000 fois plus intense que le champ magnétique terrestre), la différence d'énergie entre les deux niveaux α et β est de 4.10^{-5} kcal/mol soit 0,1672 J/mol. Le rapport entre le nombre de noyaux du niveau α et le nombre de noyaux du niveau β suit une distribution de Boltzmann et est donné par $\frac{N_\alpha}{N_\beta} = e^{\frac{\Delta E}{RT}}$ avec R la constante des gaz parfaits et T la température absolue (310 K pour le corps humain). On obtient donc un rapport de 1,000064, c'est-à-dire une différence de 64 noyaux pour un million.

2.1.1.2 La précession

La précession est le phénomène de base de la résonance nucléaire magnétique. On peut représenter, dans le modèle classique, les noyaux comme de petits aimants tournant autour d'un axe (Figure 13).

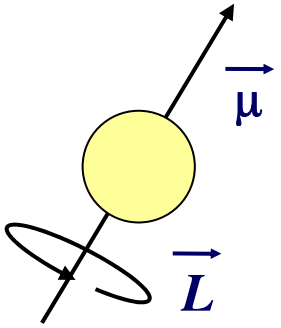


Figure 13 : Représentation d'un noyau comme un élément présentant une aimantation avec son moment angulaire (L) et son moment d'aimantation (u).

En présence d'un champ magnétique, deux forces interviennent sur le noyau : la première essaie d'orienter le moment magnétique selon le champ B_0 et la seconde tente de maintenir le moment angulaire. La résultante de ces deux forces fait que le noyau tourne comme une toupie (Figure 14).

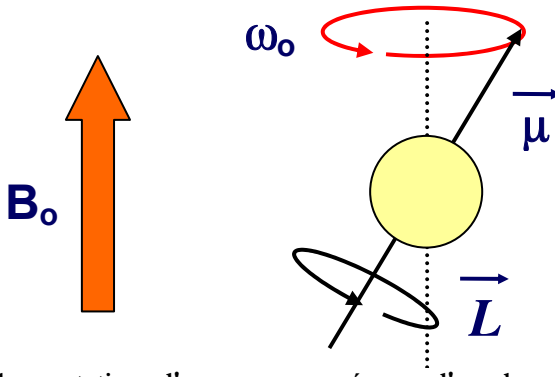


Figure 14 : Les deux rotations d'un noyau en présence d'un champ magnétique.
La première se fait autour de son axe et la seconde autour des lignes de champ magnétique.

Cette rotation autour de l'axe du champ magnétique s'appelle la précession. La fréquence de cette rotation appelée fréquence de précession ou fréquence de Larmor est donnée par la formule :

$$\omega_0 = \gamma B_0$$

$$\nu_0 = \frac{\gamma B_0}{2\pi}$$

où ω_0 est la fréquence angulaire et s'exprime en radians.s^{-1} et ν_0 est la fréquence et s'exprime en Hz.

Contrairement à ce qui est souvent dit pour expliquer la résonance magnétique nucléaire, les spins ne s'alignent pas avec B_0 mais ils précessent autour de l'axe du champ magnétique selon un angle bien défini de 54,7° (Figure 15).

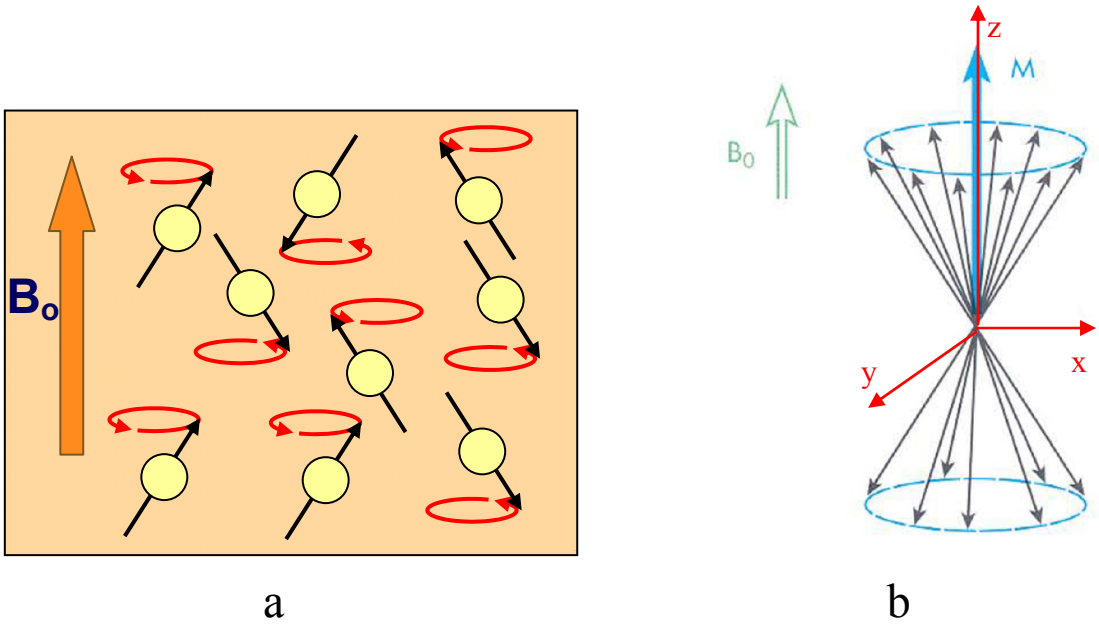


Figure 15 : Mouvement de précession et aimantation globale.
(a)Mouvement de précession des noyaux dans un champ magnétique. (b) Réunion de tous les spins à l'origine du référentiel et aimantation résultante (M).

Nous avons vu précédemment (2.1.1.1) que les spins s'orientant dans le sens du champ magnétique présentaient un niveau énergétique plus bas et qu'ils étaient légèrement plus nombreux. En ramenant tous les spins à l'origine du système de coordonnées (Figure 15 b), les projections dans le plan <x,y> s'annulent, alors que les projections selon l'axe z, du fait d'un nombre légèrement plus important de noyaux orientés dans le sens du champ magnétique, donnent un vecteur d'aimantation M_0 aligné avec le champ B_0.

2.1.1.3 Le signal de résonance magnétique

Le vecteur d'aimantation n'est pas mesurable tel quel : il faut l'exciter pour le faire quitter son état d'équilibre et ainsi obtenir un signal mesurable.

Cela est réalisé en appliquant une onde électromagnétique de fréquence angulaire ω_0 : dans ce cas, les moments d'aimantation interagissent et provoquent une rotation du vecteur M_0 autour de l'axe y. On notera par la suite M, le vecteur d'aimantation à l'instant t, M_0 étant le vecteur à l'état d'équilibre. L'angle de cette rotation dépend de la puissance et de la durée de l'onde radio fréquence émise : ainsi, il est courant de parler

d'onde 90° ou 180° pour parler de l'onde qui permet le basculement du vecteur d'aimantation de 90° ou de 180°. A l'arrêt de l'émission de l'onde radiofréquence, le vecteur M retourne à son état d'équilibre en décrivant une spirale (Figure 16). La rotation de la projection du vecteur M dans le plan <x,y> (M_{xy}) génère un champ magnétique fluctuant qui peut induire une tension électrique dans une bobine : c'est cette tension qui est mesurée (Figure 16). Ce signal dit de précession libre est couramment appelé par sont acronyme anglais : FID (Free Induction Decay).

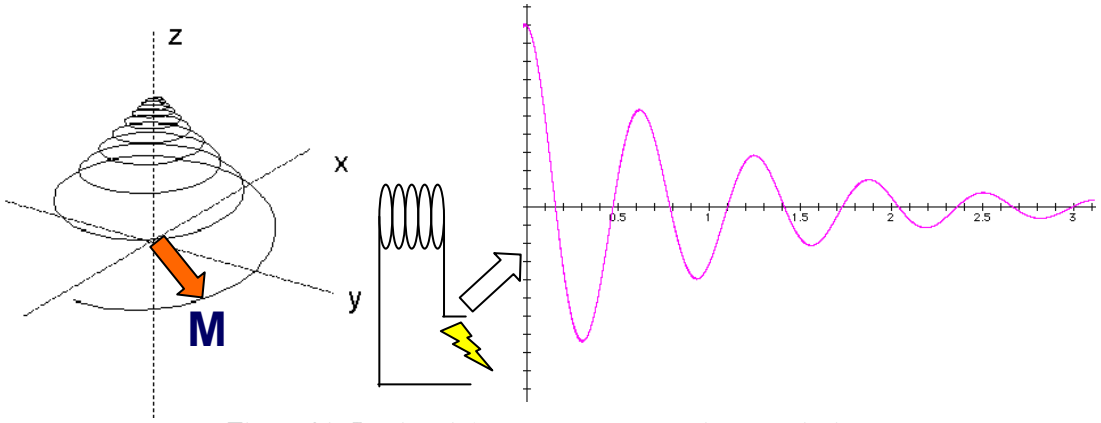


Figure 16 : Le signal de résonance magnétique nucléaire.
Le retour à l'équilibre de M se fait selon un parcours en spirale. La mesure du courant induit dans une bobine fournit le signal de RMN : le FID.

Cependant le système de coordonnées utilisé ci-dessus, dit référentiel du laboratoire, n'est pas le plus adapté au phénomène qui nous intéresse (cela revient à essayer de lire l'étiquette d'un disque vinyle pendant qu'il tourne sur la platine). La solution consiste donc à adopter un référentiel tournant à la fréquence angulaire ω_0. Dans ce nouveau référentiel, M_{xy} ne tourne plus si nous sommes exactement à la condition de résonance (i.e. ω la fréquence de l'onde radiofréquence d'excitation est exactement celle du noyau c'est à dire ω_0).

Le phénomène qui conduit au retour à l'équilibre du vecteur d'aimantation M à l'arrêt de l'émission de l'onde radiofréquence est appelé la relaxation. Dans le référentiel tournant, elle peut se décomposer en deux processus suivant une fonction exponentielle en fonction du temps (Figure 17) : la relaxation longitudinale caractérisée par le temps T1 et la relaxation transversale caractérisée par le temps T2.

La relaxation longitudinale correspond au retour à l'équilibre de la projection du vecteur M sur l'axe z notée M_z. Elle provient surtout de la perte d'énergie dans l'environnement par émission de chaleur notamment et au couplage bipolaire avec les

autres spins ou aux interactions avec des particules paramagnétiques. Le retour à l'équilibre est donné par la formule :

$$M_z(t) = M_0\left(1 - e^{-t/T1}\right)$$

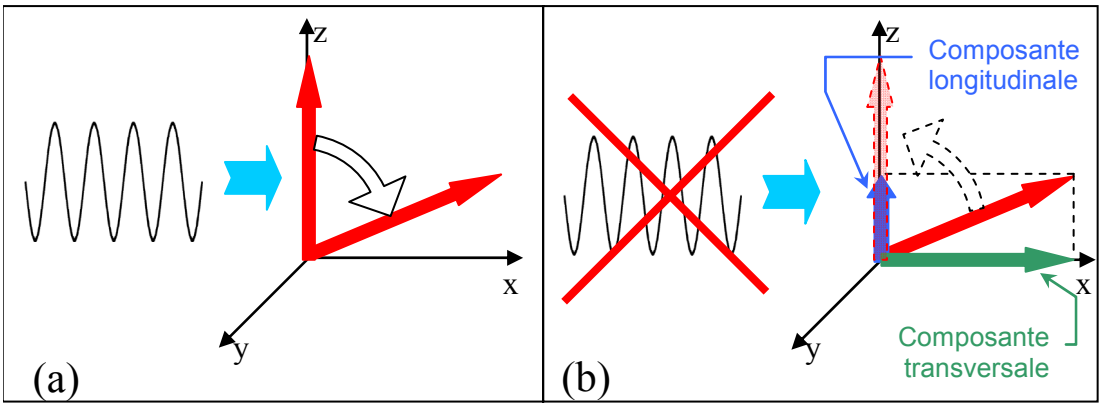


Figure 17 : Bascule de l'aimantation globale et sa décomposition.
(a) Sous l'effet d'une onde radiofréquence, le vecteur d'aimantation globale du proton bascule selon un angle dépendant de l'onde. (b) A l'arrêt de l'excitation, le vecteur revient vers sa position d'équilibre. Le vecteur d'aimantation globale peut être décomposé en deux vecteurs : une composante longitudinale selon l'axe z(relaxation longitudinale) et une composante transversale dans le plan <x,y> (relaxation transversale).

La relaxation transversale correspond au retour à l'équilibre de la projection du vecteur d'aimantation M dans le plan <x,y> appelé M_y dans le référentiel tournant. Lors de l'émission de l'onde radiofréquence, les spins se retrouvent tous en phase et ne forment plus qu'un vecteur dans le référentiel tournant. Cependant, rapidement, les couplages spin-spin (couplage J) et les imperfections d'homogénéité du champ magnétique entrainent leur déphasage puisqu'ils ne tournent plus à la même vitesse et cela entraine une diminution du vecteur d'aimantation globale (leur somme tend progressivement vers 0) (Figure 18).

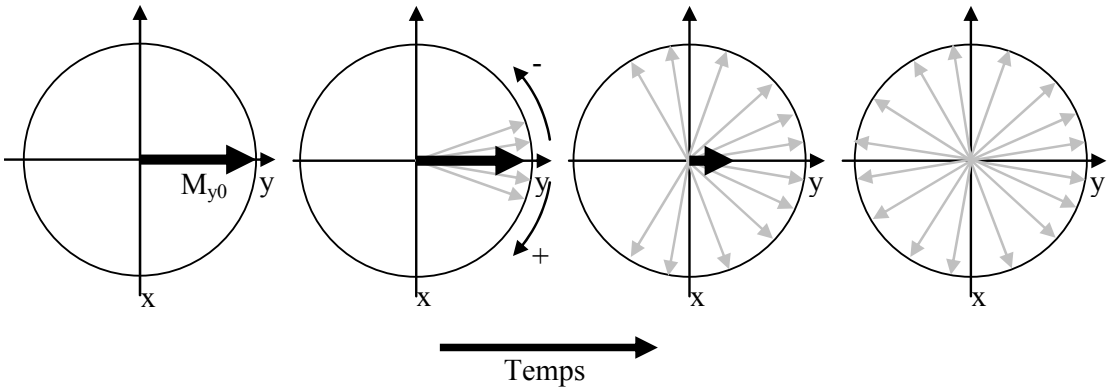


Figure 18 : La relaxation transversale.
Les spins se déphasent progressivement entrainant la repousse de l'aimantation transversale.

Cette décroissante suit une exponentielle caractérisée par le temps T2* dépendant principalement de l'hétérogénéité du champ magnétique :

$$M_y(t) = M_{y0} e^{-t/T2*}$$

La relaxation T2 est caractérisée à la fois par l'hétérogénéité du champ magnétique et par les fluctuations locales de champ magnétique produites par les noyaux eux-mêmes. On a donc :

$$\frac{1}{T2*} = \frac{1}{T2} + \gamma.\Delta\omega_0 \qquad (1)$$

où $\Delta\omega_0$ représente l'hétérogénéité locale du champ magnétique B_0.

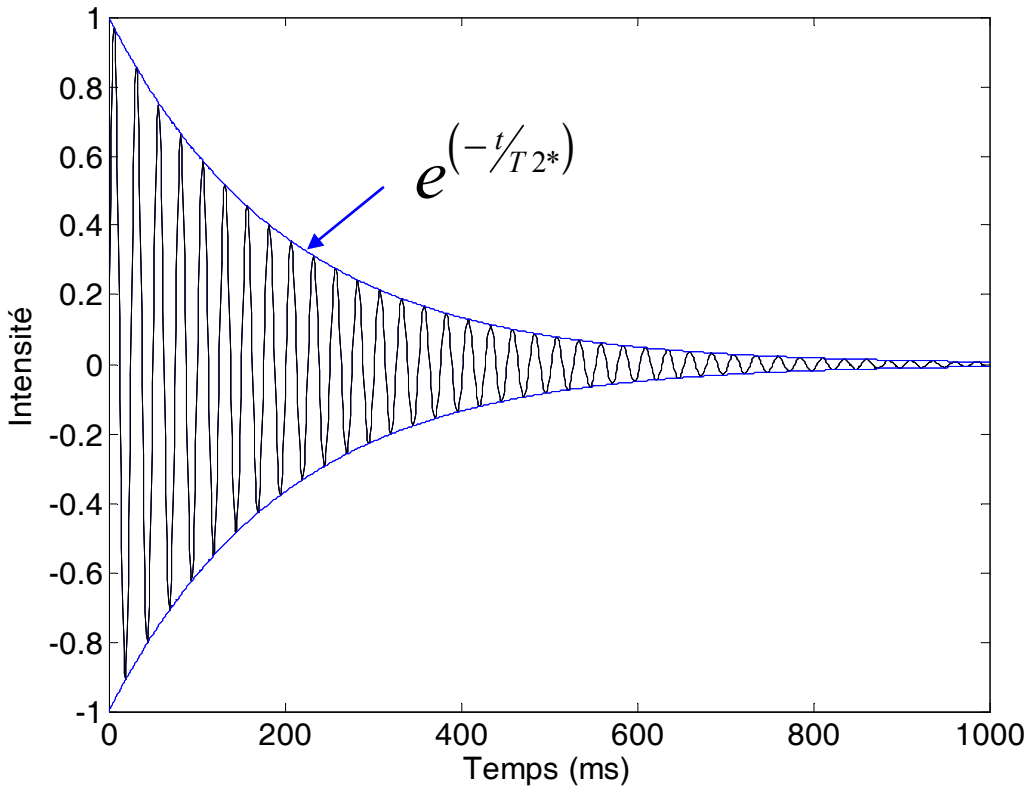


Figure 19 : Signal de précession libre.
Le signal mesuré par l'antenne de réception suit une décroissance exponentielle de paramètre T2*.

L'expérience la plus simple en résonance magnétique consiste à appliquer une impulsion 90° à un ensemble uniforme de protons placé dans un champ magnétique parfaitement homogène. Le vecteur d'aimantation bascule alors dans le plan <x,y> et précesse avec la fréquence υ_0. La relaxation en T2* entraine la décroissance du signal (Figure 19).

Dans le but de faciliter l'acquisition du signal de résonance magnétique nucléaire et de réduire l'influence des hétérogénéités du champ magnétique, une séquence plus

complexe a été mise au point : l'écho de spin (couramment appelée "Spin echo") (Figure 20). Cette séquence est basée sur l'application de 2 impulsions radiofréquences : une impulsion 90° suivie d'une impulsion 180° (parfois appelée impulsion de refocalisation).

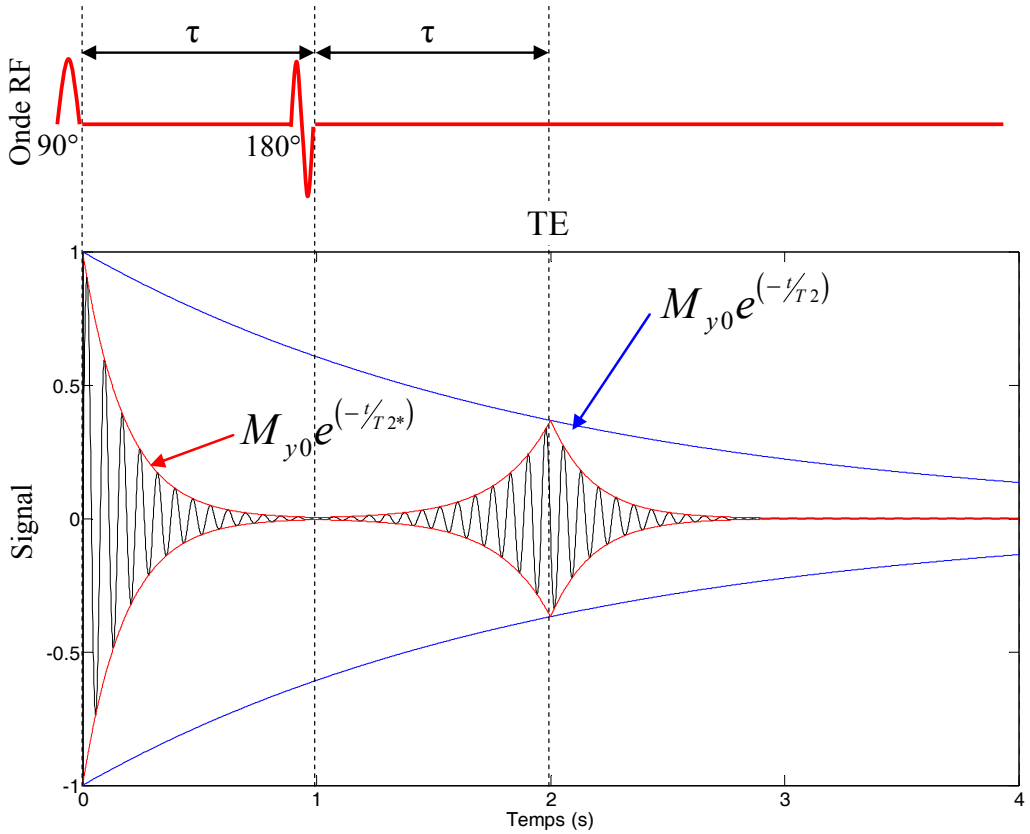


Figure 20 : Séquence d'écho de spin.
Une première onde RF génère un premier écho qui n'est pas mesuré, une seconde impulsion RF à 180° provoque la refocalisation des spins et un nouvel écho au temps 2τ qui est alors mesuré.

L'application d'une onde radiofréquence 90° entraine la bascule du vecteur d'aimantation dans le plan <x,y> et l'apparition d'un écho qui n'est pas mesuré. Rapidement les spins se déphasent et le signal subit une décroissance en T2*. Au temps τ, une impulsion radiofréquence 180° est appliquée provoquant la bascule de tous les spins de 180° mais tous les spins conservent leur fréquence de précession, provoquant le rephasage de ceux-ci. Au temps 2τ, les spins sont refocalisés, provoquant l'apparition d'un écho dont l'amplitude a subi uniquement la décroissance en T2 qui est échantillonnée. Le temps 2τ est appelé temps d'écho et est noté TE. Ce processus est illustré sur la Figure 21.

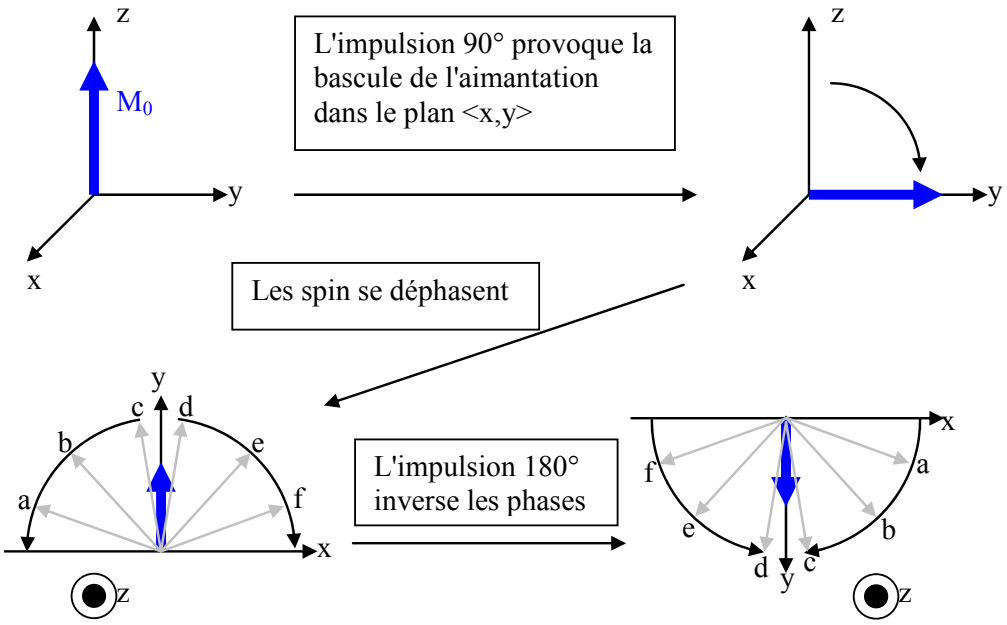


Figure 21 : Le processus de la séquence d'écho de spin.
L'impulsion à 180° provoque l'inversion de tous les spins : les spins les plus rapides se retrouvent "derrière" et les plus lents "devant". Ainsi au temps d'écho, ils se retrouveront tous synchronisés.

L'utilisation de la séquence d'écho de spin avec plusieurs impulsions radiofréquence permet d'avoir un signal en décroissance T2 (Figure 22).

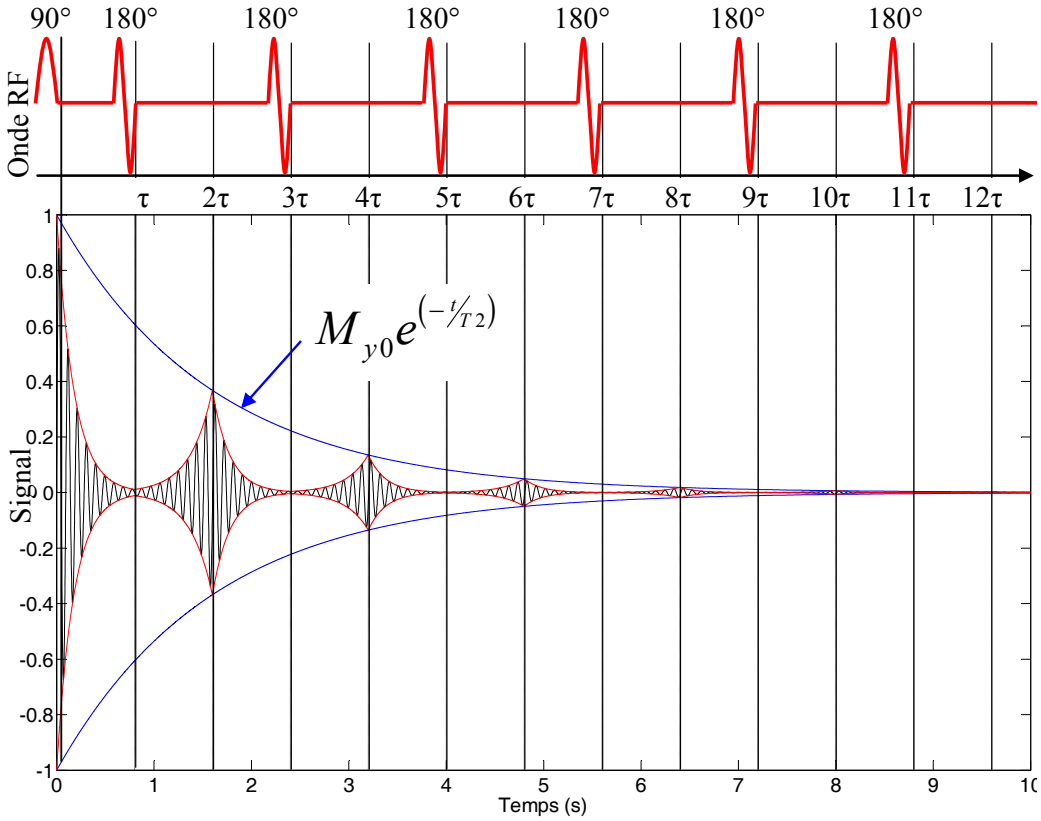


Figure 22 : La séquence d'écho de spin multiple permet d'obtenir un signal en décroissance T2.

2.1.2 La spectroscopie de résonance magnétique

En spectroscopie, comme en imagerie, on excite sélectivement certains noyaux (en général l'hydrogène mais on peut aussi travailler avec le phosphore, le carbone, le fluor ou le sodium) et l'on mesure le signal généré par ces noyaux en réponse à l'excitation. C'est un signal temporel et on obtient sa représentation fréquentielle par une transformée de Fourier. Le résultat est un spectre qui donne des informations sur la composition chimique du volume étudié.

Nous avons vu que la fréquence de Larmor d'un noyau est proportionnelle au champ magnétique externe B_0 ressenti par celui-ci. Cependant le champ magnétique externe ressenti peut être différent pour chaque noyau en fonction de son environnement chimique. En présence du champ B_0, le mouvement des électrons est tel qu'un champ magnétique supplémentaire de faible intensité s'opposant au champ B_0 est créé. En fonction de la densité électronique et des liaisons chimiques formées à proximité du noyau, ce dernier ressent un champ magnétique légèrement différent de B_0. L'équation de Larmor devient donc :

$$\nu = \frac{\gamma\left(B_0 - \sigma B_0\right)}{2\pi} = \frac{\gamma\left(1-\sigma\right)B_0}{2\pi}$$

où σ est la constante d'écran et correspond à l'environnement de chaque noyau.

Donc si l'environnement électronique varie d'un noyau à l'autre, on aura un champ B_{eff} et donc une fréquence de résonance différente pour chaque noyau. Ce phénomène, appelé "déplacement chimique" et noté δ, s'exprime en p.p.m. (partie par million) de façon à le rendre indépendant de la valeur du champ B_0.

$$\delta = \frac{\nu - \nu_{ref}}{\nu_{ref}}.10^6\ ppm \qquad (2)$$

où ν_{ref} est la fréquence de résonance choisie comme 0 de l'échelle.

Cependant cette fréquence de référence est très proche de la fréquence nominale du système ν_0 d'où on utilise couramment :

$$\delta = \frac{\nu - \nu_{ref}}{\nu_0}.10^6\ ppm$$

On utilise traditionnellement le tétraméthylsilane (TMS) comme référence car sa structure chimique est composée de 12 protons symétriquement répartis dans la molécule (Figure 23).

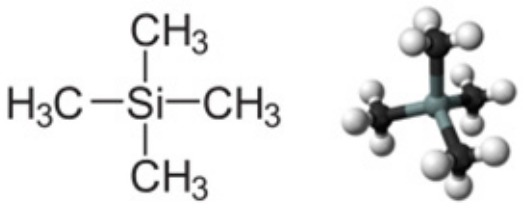


Figure 23 : Structure chimique de la molécule de TMS.
Elle présente 12 noyaux d'hydrogène (blanc) symétriquement répartis dans la molécule.

Cette structure fait que ces 12 protons sont équivalents et que la molécule a ainsi une fréquence de résonance unique et précise. Le choix de cette référence vient du fait que la spectroscopie était initialement réalisée in vitro et que les échantillons doivent être placés en solution : le TMS a donc été choisi car la majorité des composants étudiés en SRM ont une fréquence de résonance supérieure à celle du TMS permettant qu'il n'y ait habituellement pas d'interférence entre la référence et l'échantillon. Depuis, la position de ce pic sert de référence 0 en SRM quelque soit les conditions.

Le signal brut capté par l'antenne (le FID) provient de la combinaison de toutes les molécules contenant des protons, présentes dans l'élément de volume (voxel). Etant donné que les fréquences de résonance de chaque molécule sont légèrement différentes, l'analyse du spectre obtenu par la transformée de Fourier permet d'identifier les molécules présentes et d'en déterminer leur concentration.

Pour rappel, la transformée de Fourier d'une fonction f intégrable sur $\mathbb{R}$ notée $\hat{f}$ est donnée par :

$$\hat{f}(\nu) = \int_{-\infty}^{+\infty} f(t) e^{-i2\pi\nu t} dt \qquad (3)$$

C'est la transformée de Fourier du FID qu'on appelle spectre.

Chaque pic du spectre correspond donc à un groupement de protons particuliers, on peut ainsi associer chacun des pics à une molécule présente dans l'échantillon analysé. Certaines molécules présentant différents groupements de protons peuvent comporter différentes fréquences de résonance et donc être représentées dans le spectre par plusieurs pics.

Parmi les paramètres caractéristiques d'un tissu en résonance magnétique, le T2* influe directement sur la largeur des pics (Figure 24). Ce paramètre dépend d'une part du T2 et d'autre part de l'homogénéité du champ (cf. équation(1)). Etant donné que le paramètre T2 est relativement fixe pour un tissu, il faut que le champ B_0 soit le plus

homogène possible pour que les pics soient les plus fins possibles et donc plus facilement discernables (Figure 25).

A l'opposé, quelque soit la valeur du T2*, le pic représente toujours la même quantité de protons, donc l'intégrale du pic est toujours la même.

L'amélioration de l'homogénéité du champ permet d'améliorer la séparation des pics et facilite l'estimation de la concentration relative des métabolites.

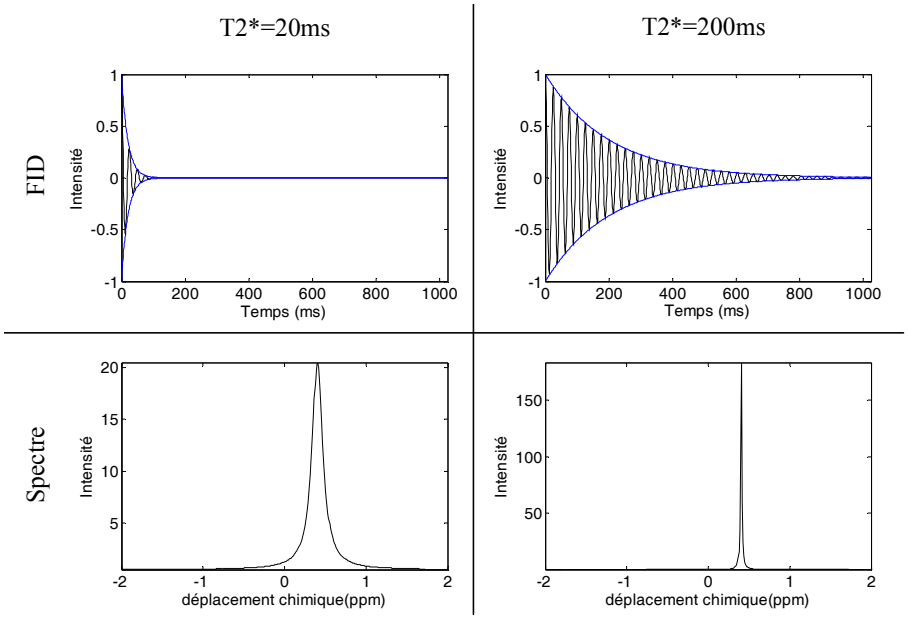


Figure 24 : Influence du temps T2* sur la largeur des pics.

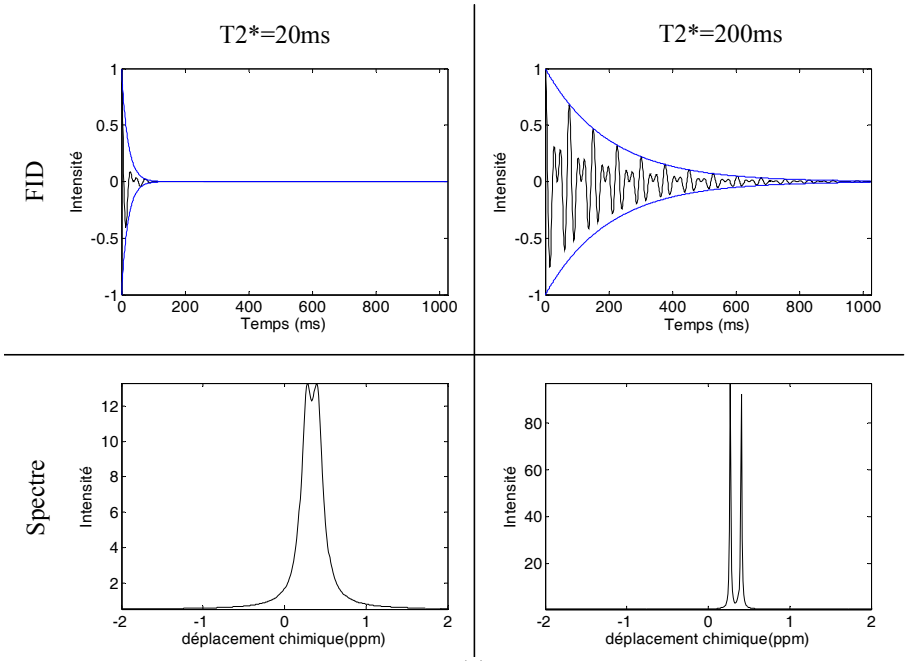


Figure 25 : Influence du temps T2* sur la séparation des pics.

En plus de l'environnement nucléaire, certaines molécules peuvent présenter des interactions entre les groupes de noyaux la composant. C'est ce qu'on appelle le couplage nucléaire. Ces couplages peuvent être homonucléaires (entre noyaux de même espèce) ou hétéronucléaires (entre noyaux d'espèce différente : ${}^{1}_{1}H$ et ${}^{13}_{6}C$ par exemple). Le cas le plus important en spectroscopie du proton est le couplage homonucléaire. C'est un phénomène très complexe donc il existe plusieurs variantes (couplage J, couplage AB,…). Nous allons présenter dans la suite un cas simple de couplage J. Il est caractérisé par la constante de couplage J qui exprime, en Hertz, la distance entre 2 raies (Figure 26) indépendamment de l'intensité du champ B_0. Ce phénomène déforme les pics en fonction du temps d'écho : cette déformation pouvant aller jusqu'à l'inversion complète des pics.

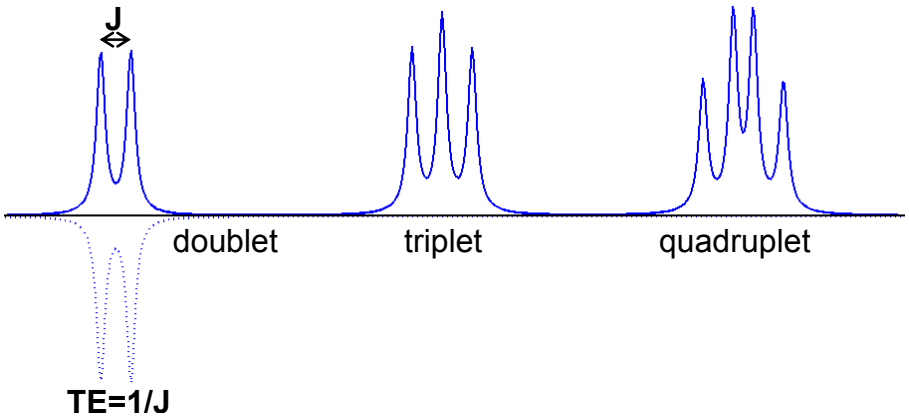


Figure 26 : Couplage spin-spin.
L'interaction entre les groupements chimiques d'une même molécule provoque l'apparition de pics multiples : J définit la distance entre 2 pics d'une même molécule. Avec un TE de 1/J, les pics sont inversés.

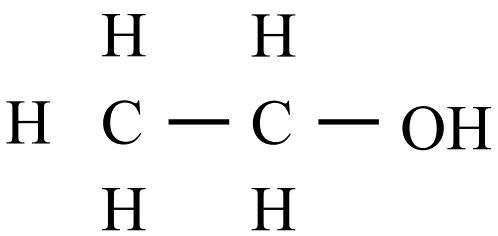
Le nombre de pics de résonance d'un noyau est donné par la formule suivante :

$$2nI+1$$

où n est le nombre de noyaux voisins participant à l'interaction et I le nombre de spins (1/2 pour ${}^{1}H$) donc dans notre cas la formule se résume à n+1 pics.

Donc pour un noyau ayant une interaction avec un seul autre noyau, on aura deux pics, trois pics pour une interaction avec deux autres noyaux …

Prenons en exemple le cas de l'éthanol très pur (CH_3CH_2OH) :



Le proton du groupement OH est voisin de deux protons (ceux du groupement méthylène : CH_2) donc il sera représenté par trois pics. Chacun des protons du groupement CH_3 (méthyle), qui sont équivalents et n'interagissent pas entre eux, sont également affectés par les deux protons du groupement CH_2 et seront donc représentés par 3 pics. Enfin les deux protons du groupement méthylène sont affectés d'une part par le groupement méthyle ce qui entraine l'apparition de 4 pics mais également par le proton du groupement OH ce qui va entrainer le dédoublement du quadruplet (Figure 27)

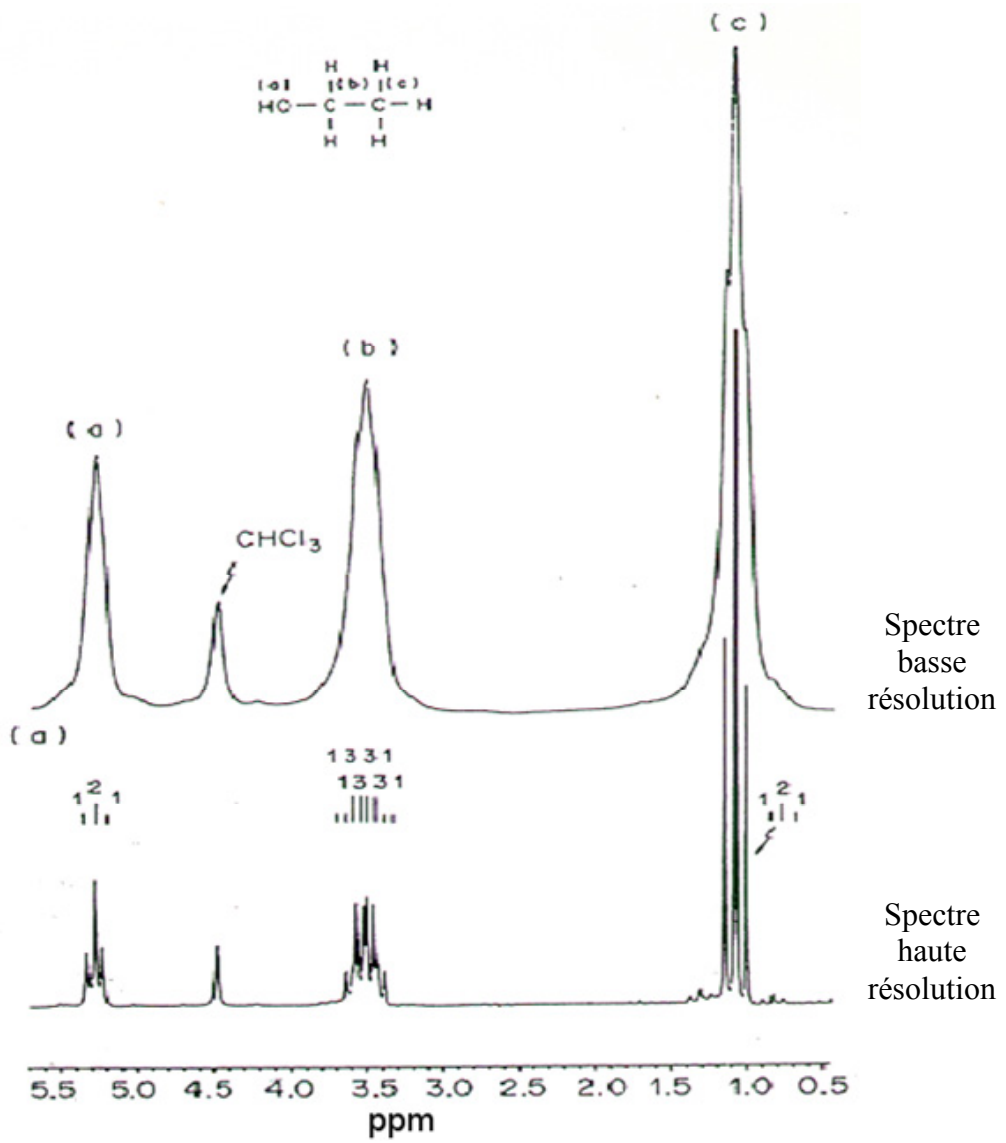


Figure 27 : Spectres de l'éthanol en résonance magnétique nucléaire du proton en solution dans du Chloroforme (CHCL3).

Avec un champ magnétique faible ou peu homogène, les pics sont larges et on parle de spectre basse résolution et avec un champ magnétique élevé ou très homogène, les pics sont bien séparés et on parle de spectre haute résolution.

Cette explication est la théorie "simple", d'autres effets peuvent s'ajouter. Ainsi en fonction du diluant, par exemple, les pics correspondant au groupement hydroxyle peuvent se déplacer ou se singulariser. De même le dédoublement du quadruplet du groupement CH_2 n'est pas toujours visible car l'effet du groupement OH est beaucoup plus faible que celui du groupement méthyle. Le citrate présent dans la prostate présente un tel couplage (bien que plus complexe) donnant un quadruplet à haute résolution.

2.1.2.1 Les séquences

L'acquisition d'un signal en RMN est soumise à un programme bien particulier définissant le moment d'émission des ondes radiofréquences (RF), leurs durées, la puissance des gradients à appliquer au champ magnétique … Ce programme s'appelle une séquence. Il existe différentes séquences prédéfinies permettant chacune de mettre en évidence un phénomène particulier ou de s'affranchir de certains problèmes. Les deux séquences principalement utilisées pour la SRM sont appelées STEAM (Stimulated Echo Acquisition Mode) et PRESS (Point REsolved SpectroScopy).

Pour chaque séquence, il existe une série de paramètres qui permettent de s'adapter au cas courant. Pour présenter les principaux paramètres d'une séquence, nous allons faire référence à une séquence de base de l'IRM : la séquence d'écho de spin (cf. 2.1.1.3).

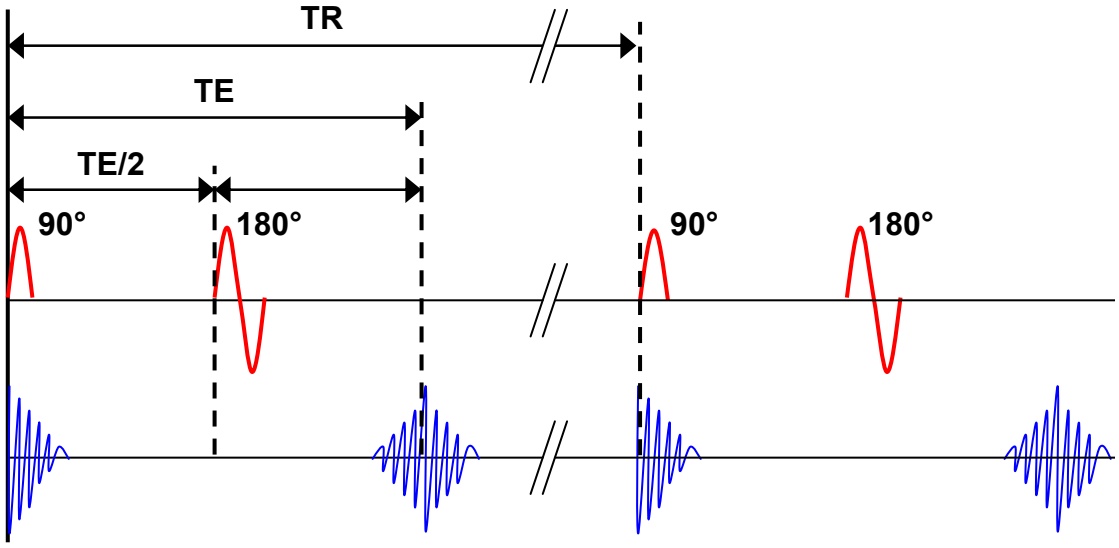


Figure 28 : Schéma de la séquence d'écho de spin.

Pour différentes raisons, il est nécessaire de reproduire ce schéma plusieurs fois. En spectroscopie, cette répétition de l'acquisition sur le même volume d'intérêt, permet d'améliorer le rapport signal sur bruit (S/N) d'un rapport $\sqrt{2}$ à chaque fois. Entre chaque répétition, il faut laisser un temps, dit de répétition, s'écouler (Figure 28).

La mesure du signal ne peut s'effectuer que dans le plan transversal après la bascule de l'aimantation suite à une impulsion RF. Cette aimantation notée $\overrightarrow{M_0}$, tournant autour de l'axe du champ B_0, va subir une décroissance exponentielle en fonction du T2*. Le signal électrique mesuré dans l'antenne est donc une fonction sinusoïdale amortie par une exponentielle en T2*.

On peut représenter $\overrightarrow{M_0}$ sur un cercle trigonométrique en le projetant sur les axes x et y. On obtient les deux composantes M_x et M_y, appelées respectivement signal réel et signal imaginaire. Ce sont ces deux composantes qui sont mesurées et forment le FID. C'est un vecteur composé de nombres complexes dont le nombre de composantes dépend des paramètres d'acquisition : fréquence d'échantillonnage et nombre de mesures effectuées (généralement 512, 1024 ou 2048 sur des machines cliniques). C'est donc un signal temporel. On le convertit en signal fréquentiel en appliquant une transformée de Fourier discrète pour obtenir le spectre correspondant.

a STEAM

La séquence STEAM est basée sur une série de trois impulsions RF 90°. Deux impulsions successives à 90° permettent d'obtenir un écho de spin comme avec un couple 90°-180°, mais d'intensité moindre. L'utilisation d'une troisième impulsion permet d'obtenir un écho stimulé. Cependant dans cette série de trois impulsions RF, chaque combinaison génère un écho : Il en résulte l'apparition de quatre échos où seul celui issu de la combinaison des trois impulsions est échantillonné. Dans cette séquence, le temps d'écho est la somme du temps séparant les deux premières impulsions et de celui entre la dernière impulsion et l'écho mesuré. Le temps entre la deuxième et la troisième est appelé temps de mélange (TM). Nous avons représenté le chronogramme de cette séquence sur la Figure 29.

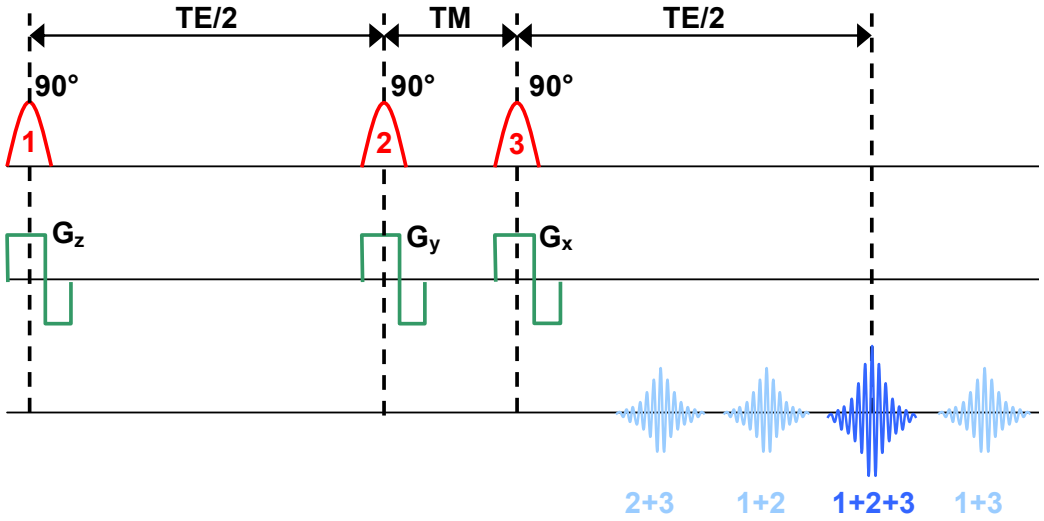


Figure 29 : Séquence STEAM.
Le signal enregistré provient d'un écho stimulé généré par la succession d'impulsions RF 90°. Trois impulsions génèrent quatre échos correspondant aux différentes combinaisons possibles. Seul le troisième, correspondant à l'écho stimulé est mesuré. Le temps séparant les deuxième et troisième impulsions est appelé temps de mélange.

Cette séquence est particulièrement adaptée aux temps d'écho courts mais l'intensité du signal est plus faible que l'écho de spin (90°-180°) dû au fait que seules des impulsions 90° sont utilisées.

b PRESS

La séquence PRESS (Figure 30) est, quant à elle, basée sur un train d'impulsions 90°-180°-180° permettant d'obtenir un écho de spin. Sur certains appareils, il est possible de régler le temps entre les impulsions 90° et 180° (τ_1) et entre les 2 impulsions 180° (τ_2). On a alors $TE = \tau_1 + \tau_2 + (\tau_2 - \tau_1) = 2\tau_2$. Deux échos sont générés, toujours symétriques par rapport aux impulsions 180° : le premier à $2\tau_1$ et le second à $2\tau_2$ (ou TE) mais seul le deuxième est échantillonné puisqu'il correspond au volume sélectionné par l'intersection des trois coupes définies par les gradients de champ magnétique appliqués avec les ondes RF (cf. 2.1.2.2). La variation des durées τ_1 et τ_2 permet de modifier la réponse des couplages spin-spin et donc la configuration des pics. Sur la machine Siemens TRIO TIM 3T, le temps τ_1 est fixe (6,2 ms), en jouant sur le TE, on modifie donc uniquement τ_2.

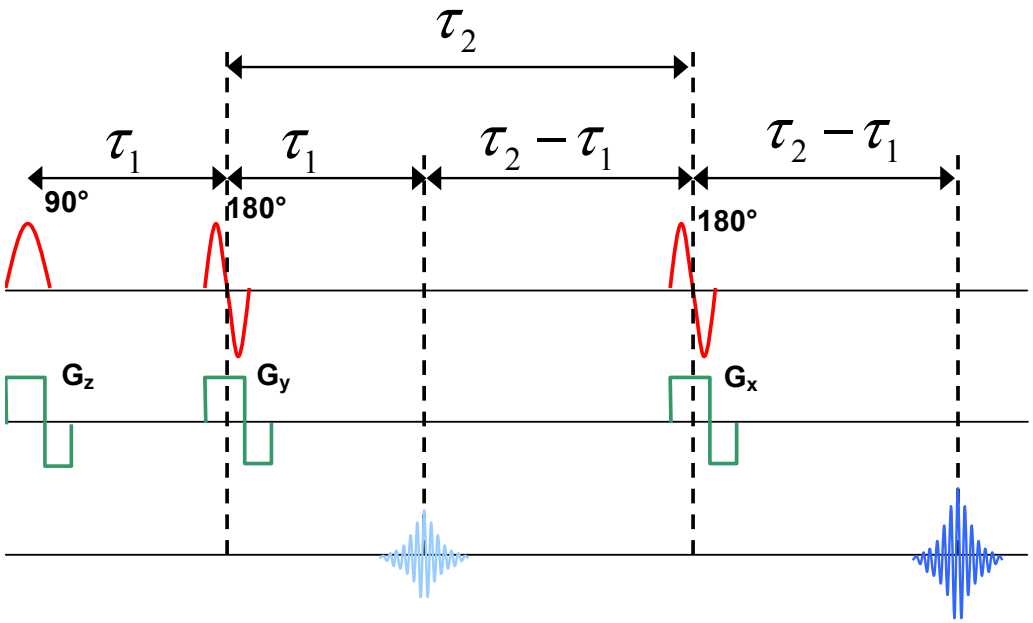


Figure 30 : Séquence PRESS.
Elle est basée sur un train d'ondes RF 90°-180°-180°. Les trois impulsions sont associées à des gradients dans chaque axe pour sélectionner le volume d'intérêt. Seul le deuxième écho, après sélection du volume par les trois plans de coupe, est mesuré. Certaines machines permettent de faire varier le temps entre les deux premiers échos (τ_1).

Cette méthode permet d'obtenir un signal plus important en intensité que la méthode STEAM (environ 2 fois) mais ne permet pas d'atteindre des temps d'écho aussi courts. Les études envisagées ne nécessitant pas d'utiliser des TE très courts, nous utiliserons donc la méthode PRESS.

2.1.2.2 Localisation

Les deux techniques de localisation principalement utilisées pour la SRM du proton sont la spectroscopie mono-voxel et l'imagerie spectroscopique ou multi-voxels. Il est important de mentionner que la technique multi-voxels n'est pas une juxtaposition de méthodes mono-voxel, mais une technique de codage spatial qui permet d'acquérir tous les spectres d'un volume donné en une seule acquisition de données (séquence).

a Spectroscopie mono-voxel

La sélection d'un élément de volume se fait par l'intersection de trois plans orthogonaux. La sélection d'un plan se fait par l'application simultanée d'une onde RF et d'un gradient magnétique dans l'axe de la coupe à sélectionner (Figure 31). Chaque

application d'un gradient et d'une onde RF génère en réponse l'émission d'un signal : seul le dernier est échantillonné pour obtenir le FID du voxel.

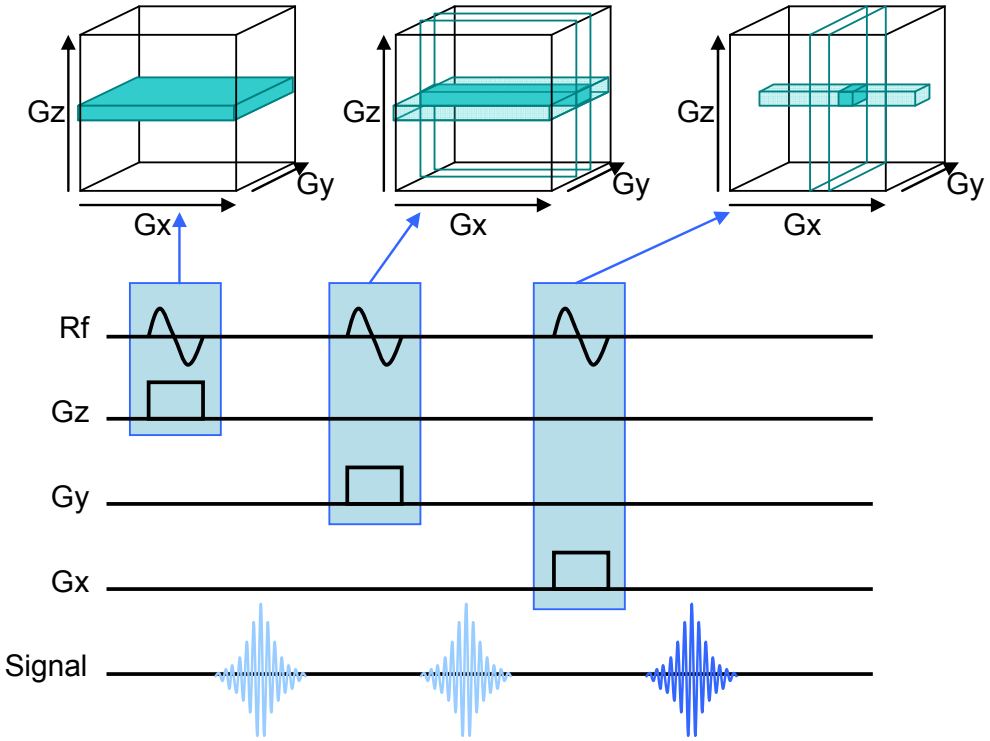


Figure 31 : Principe de sélection d'un voxel avec la technique mono-voxel.

b Spectroscopie multi-voxels ou CSI.

Cette technique porte également les noms de "technique de localisation multi-voxels", "d'imagerie du déplacement chimique" et, en anglais, "chemical shift imaging" (CSI). Elle permet d'obtenir en une seule fois les spectres provenant de tout un plan de coupe (Figure 32) ou de tout un volume.

La séquence utilisée se base sur les mêmes séquences que pour l'acquisition mono-voxel mais des gradients de phase permettant de coder l'information dans l'espace sont ajoutés.

Un volume d'analyse est alors défini par trois gradients de sélection (comme pour la spectroscopie mono-voxel). Les gradients de phase sont appliqués pour définir les voxels (généralement 16x16 par plan de coupe et dans le cas 3D 8 ou 16 plans de coupe).

Les signaux sont alors mesurés dans l'espace de Fourier et le codage spatial de "l'image spectroscopique" est retrouvé par l'utilisation de la transformée de Fourier bi ou tridimensionnelle. Comme en spectroscopie mono-voxel, une dernière transformation de

Fourier permet de passer l'information dans l'espace spectral, c'est pourquoi on trouve parfois le terme de CSI 4D en parlant de spectroscopie multi-voxels d'un volume.

Cette technique présente l'intérêt de fournir beaucoup plus de signal que la technique mono-voxel(on acquiert le signal de tout le volume d'intérêt au lieu du signal d'un seul voxel) et donc un meilleur rapport signal sur bruit. Par contre, il est plus difficile d'obtenir un champ parfaitement homogène (on parle de "shim") sur tout le volume.

Un autre avantage de cette technique est de faire de l'échantillonnage elliptique : les données du plan de Fourier sont acquises selon un schéma elliptique en partant du centre et parfois pondérées en nombre d'acquisitions, au détriment des "coins" qui ne contiennent que peu d'informations (Figure 33). Les données manquantes du plan de Fourier sont alors complétées par des zéros (zero filling) et l'information totale est alors inférée par les transformations de Fourier.

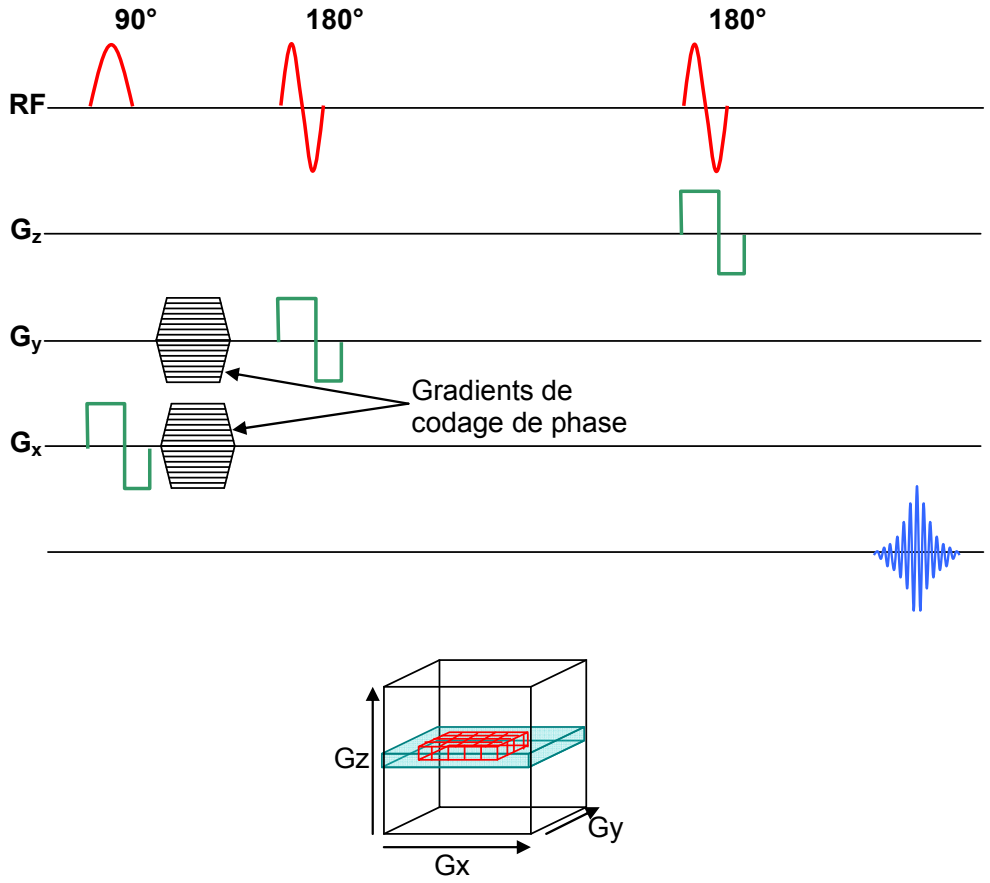


Figure 32 : Séquence d'imagerie spectroscopique (2D) basée sur la séquence PRESS.
Le gradient selon z sélectionne la coupe, les gradients selon x et y définissent la zone de la coupe à analyser et les deux gradients de phase selon x et y permettent de coder l'espace pour le découper en voxels.

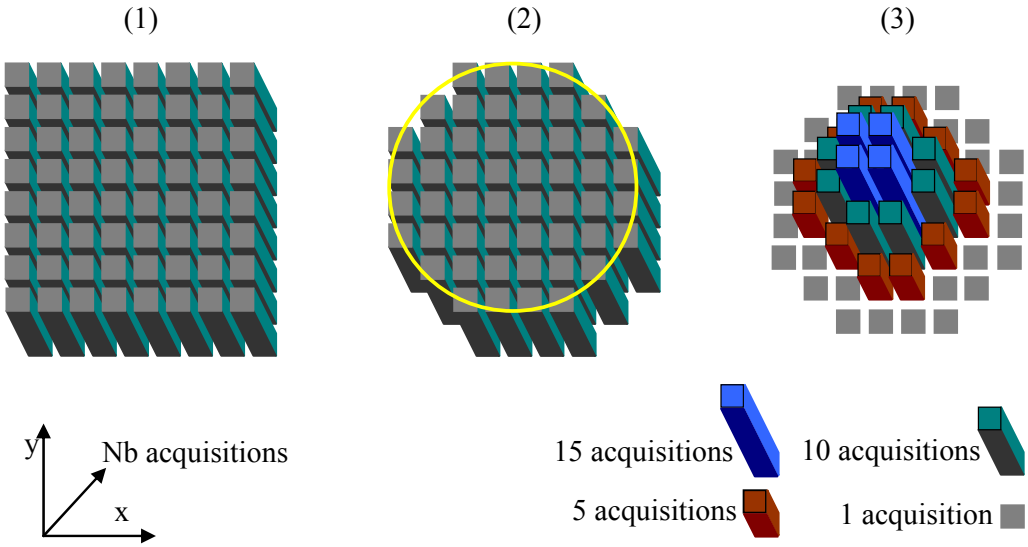


Figure 33 : Schémas de codage de la matrice CSI.
(1) Echantillonnage uniforme du plan de Fourier (2) Echantillonnage elliptique (3) Echantillonnage elliptique pondéré.

2.2 Paramètres des séquences

2.2.1 Le couplage spin-spin

Comme nous l'avons vu dans la partie 2.1.2, il y a un phénomène de couplage spin-spin qui peut apparaitre avec certaines molécules. Dans le cas de ces molécules, le choix du temps d'écho est très important car la configuration des pics change avec celui-ci. Le citrate ($C_6H_8O_7$ cf. Figure 34) présente deux groupements CH_2 qui ont un fort couplage et apparaissent donc, suivant la puissance de séparation du spectromètre, avec trois ou quatre pics. La configuration de ces pics variant en fonction du temps d'écho [55], il a donc fallu trouver la valeur de celui ci permettant d'obtenir la configuration optimale. Scheenen a effectué ce travail en 2005 [55] à partir de spectres simulés et de fantômes contenant une solution de citrate (Figure 35).

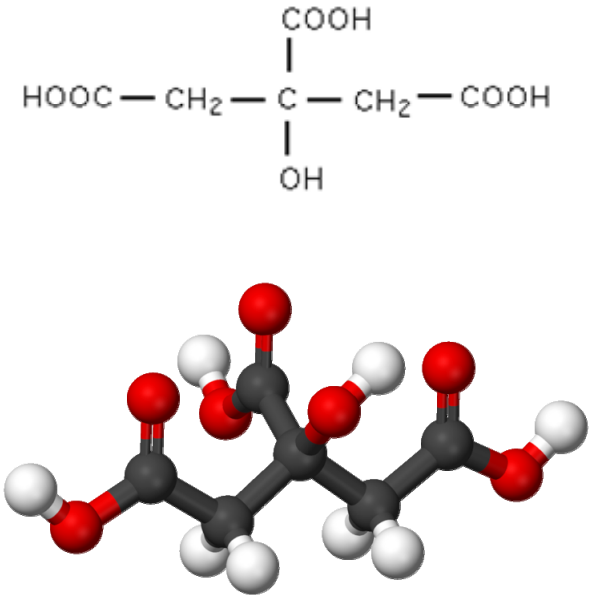


Figure 34 : Formule semi développée et modélisation 3D de la molécule de citrate.
(Hydrogène en blanc, Carbone en noir et Oxygène en rouge)

Nous avons donc réalisé le même type d'expérience sur notre IRM 3T avec des fantômes contenant une solution de citrate. Cet appareil ne permet pas de faire varier le paramètre τ_1 : il est fixé à 6,2ms. Nous nous sommes donc contentés de faire varier le TE (donc τ_2) entre 82 ms (plus petit TE possible avec la séquence PRESS multi-voxels fournie par Siemens) à 400 ms. Nous avons représenté sur la Figure 36 les spectres que nous avons obtenus pour les différents temps d'écho.

Le spectre recherché doit représenter une configuration facilement identifiable, avec des pics principalement positifs pour pouvoir estimer la concentration. De plus le citrate a un T2 moyennement long (170 ms) donc pour conserver suffisamment de signal par rapport aux autres métabolites, le TE utilisé ne doit pas être trop long. La solution proposée par Siemens présente l'inconvénient d'avoir deux pics, non symétriques avec une forte composante négative. En termes de simplicité de forme, la solution idéale aurait été d'utiliser un TE à 400ms : on a alors un seul pic positif. La solution retenue est d'utiliser un TE à 140 ms qui permet d'avoir 3 pics bien définis (avec un champ très homogène, le pic central se dédouble, on observe ainsi 4 pics) et majoritairement positifs.

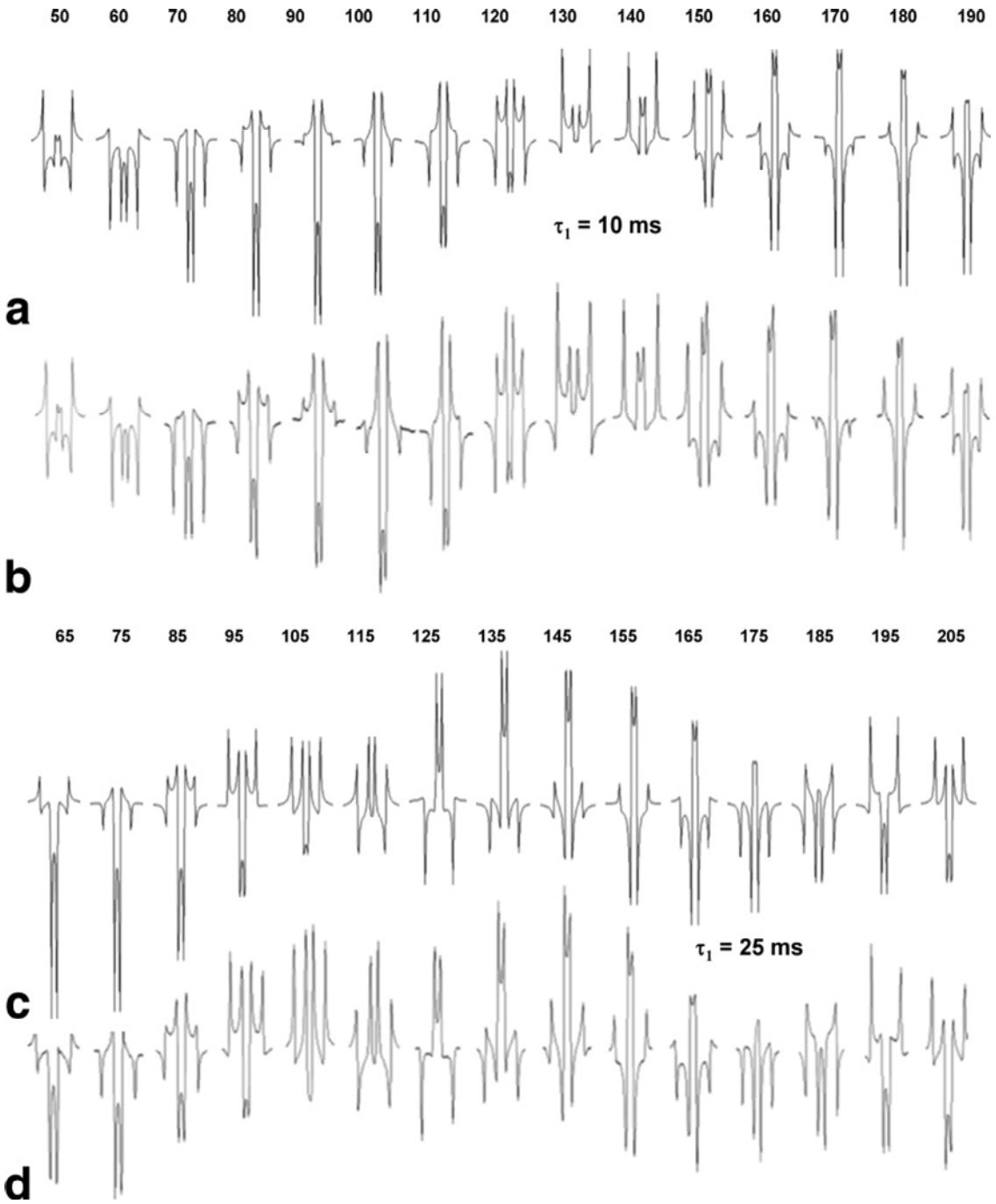


Figure 35 : Variation des pics de citrate en SRM du proton : simulation (a et c) et in vitro (b et d).

Les séries a et b sont effectuées avec τ_1 fixé à 10 ms et TE variant de 50 à 190 ms. Les séries c et d sont effectuées avec τ_1 fixé à 25 ms et TE variant de 65 à 205 ms (Images extraites de [55]).

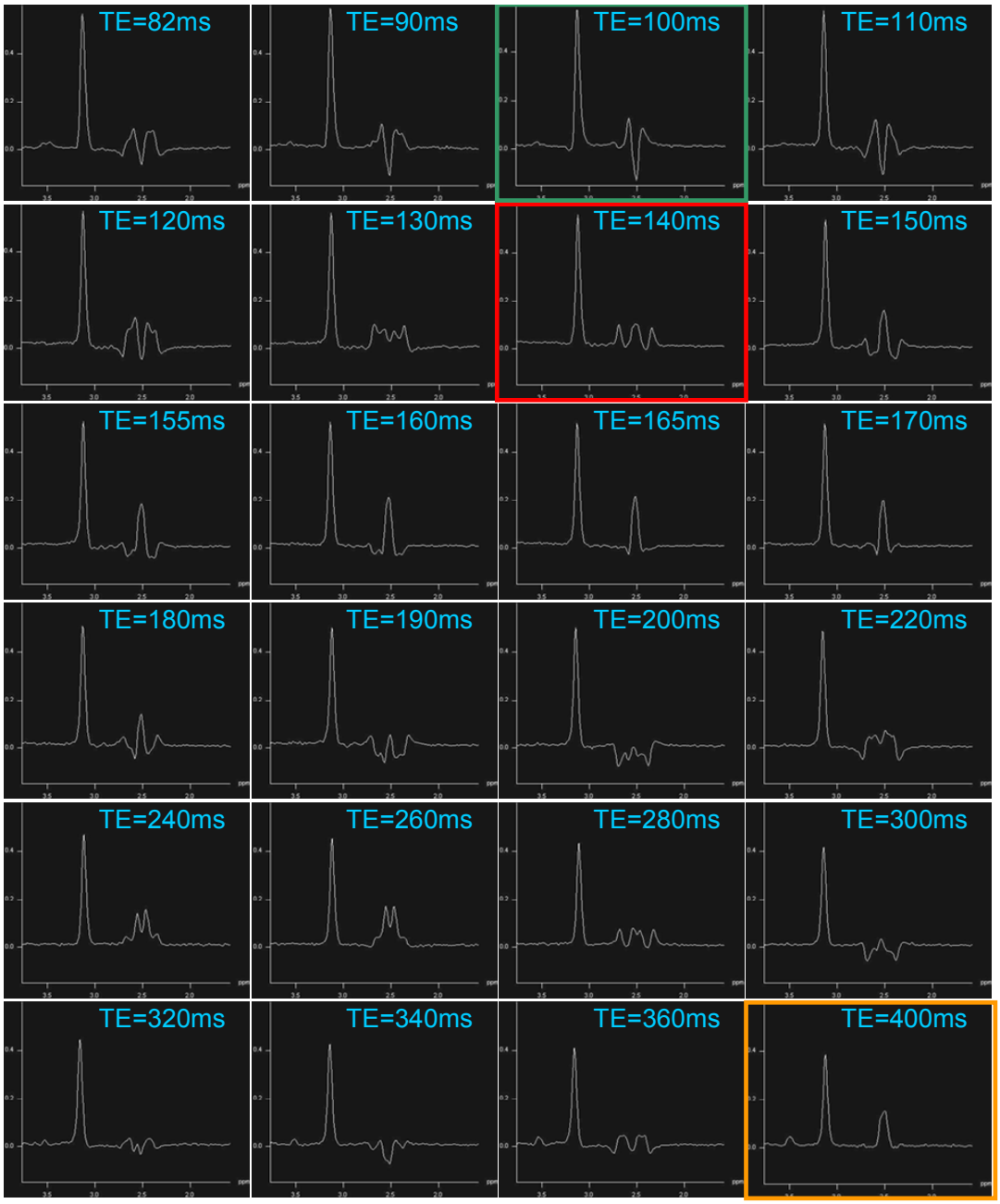


Figure 36 : Pics de citrate obtenus sur la machine Siemens en fonction du TE.
En vert, le spectre obtenu avec le TE recommandé par Siemens, en orange le spectre qui nous semble optimal et en rouge le spectre obtenu avec le TE retenu.

2.2.2 Influence du temps de répétition

Le temps de répétition (TR) définit le temps entre deux trains d'impulsions RF générant un écho mesuré (cf. Figure 30). Il est donc nécessairement supérieur au temps d'écho pour pouvoir acquérir le signal. Durant ce temps de répétition, l'aimantation longitudinale de chaque molécule repousse en fonction de leur T1 pour revenir à son état

initial. La formule donnant l'intensité du signal en fonction du temps pour une molécule est : $I = I_0\left(1 - e^{-t/T1}\right)$.

La repousse est donc de 63% au bout du temps T1 et atteint 98% au bout de 4 fois le temps T1. En diminuant le temps de répétition, on a donc une perte de signal (Figure 37), il faut donc qu'il soit le plus long possible. Cependant, en augmentant le temps de répétition, on augmente aussi le temps des examens.

Une autre méthode, pour obtenir plus de signal, consiste à réaliser plus d'acquisitions. La durée d'un examen étant donnée par le nombre de répétitions multiplié par le temps de répétition, il faut trouver un compromis pour obtenir le maximum de signal dans un minimum de temps.

Nous avons choisi un temps de répétition de 720 ms et nous supposons que nous utilisons un nombre de répétitions de 500. La durée de l'examen est donc de $500 \times 720 = 360000 ms$ soit 6 minutes. Si l'on tient compte uniquement de T1, à chaque itération, le signal reçu est de 38,12% de l'intensité maximale, soit en considérant un signal d'intensité maximale 1 à chaque itération, le signal mesuré sera d'intensité 190,6 (au lieu de 500).

Considérons le même examen avec un temps de répétition de 2500ms : à chaque itération, on capte 81,11% du signal mais avec un nombre d'acquisitions de 500, la durée d'acquisition passe à plus de 20 minutes. Pour atteindre une durée d'examen de 6 minutes, il faut réduire le nombre d'acquisitions et le signal mesuré est alors inférieur (Tableau 2) à celui mesuré dans les mêmes conditions avec un temps de répétition de 720 ms.

Tableau 2 : Durée de l'examen et quantité de signal mesuré en fonction du TR et du nombre de répétitions pour la choline.

	TR de 720 ms		TR de 2500 ms	
Nb de répétitions	Durée d'examen	Quantité de signal reçu	Durée d'examen	Quantité de signal reçu
500	6 min.	190,60	20 min. 50 s	405,56
144	1 min. 26 s.	54,89	6 min.	116,80

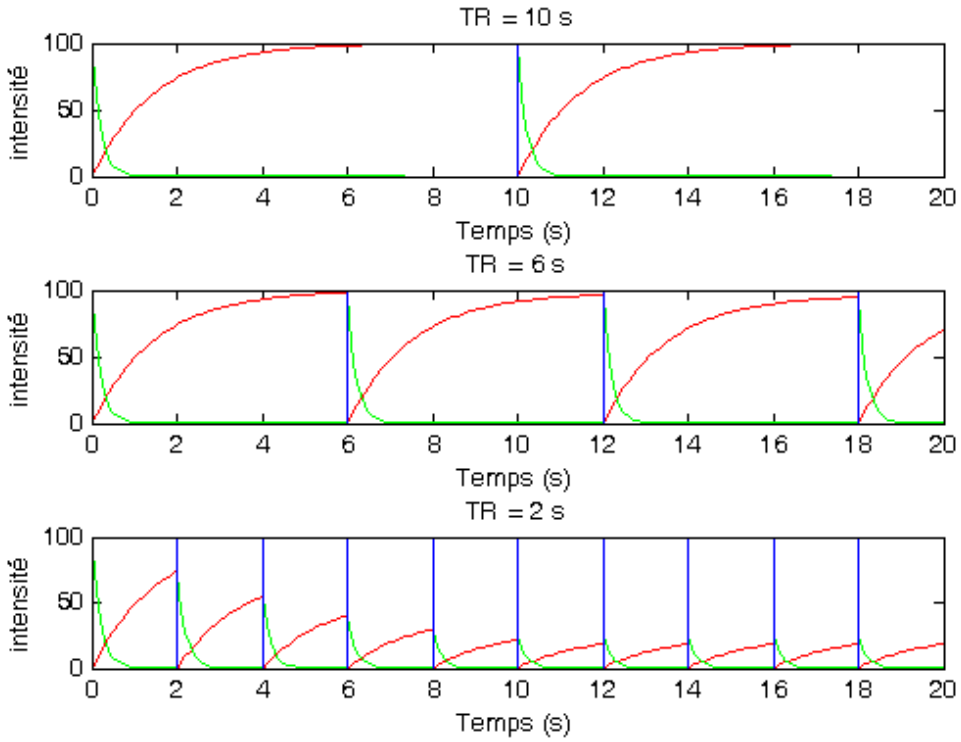


Figure 37 : Influence du TR sur le signal appliqué à la choline.
Le T1 de la choline est de 1500ms et le T2 de 300ms.Pour un TR de 10s et de 6s (4xT1), la repousse de l'aimantation longitudinale a le temps de s'effectuer (courbe rouge) et l'aimantation transversale (courbe verte) s'annule. Par contre pour un TR < T1 la repousse n'est pas suffisante et l'on mesure donc un signal inférieur au signal mesurable avec un TR suffisamment long.

2.2.3 Influence du temps d'écho

La quantité de signal mesuré est directement liée au temps T2. Ainsi, si on choisit un temps d'écho supérieur à 4 fois le temps T2, on aura perdu 98% du signal. De plus étant donné que chaque molécule possède un T2 différent, l'intensité du signal de chaque molécule sera donc différente. Idéalement, il faut donc choisir un temps d'écho très court pour capter le plus de signal possible et surtout minimiser la différence de signal entre deux métabolites. Cependant nous avons vu précédemment (2.2.1) que la configuration spectrale d'un métabolite pouvait être liée au temps d'écho à cause du couplage de spins, donc choisir un temps d'écho le plus court possible n'est pas toujours la meilleure solution. Il faut trouver le bon compromis pour obtenir le spectre le plus aisément identifiable, tout en conservant un maximum de signal.

2.2.4 Influence des paramètres d'échantillonnage

Lors de la phase d'échantillonnage du signal proprement dite, deux paramètres sont configurables : le nombre de points à acquérir (N : généralement 512 ou 1024) et la durée pendant laquelle se fait l'acquisition (Δ).

La durée entre chaque échantillonnage de points, appelé dwell time est alors : $Dw = \frac{\Delta}{N}$ et la bande passante ou sweep width $Sw = \frac{1}{Dw} = \frac{N}{\Delta}$.

Ainsi pour un échantillonnage sur 512 points pendant un temps de 256 ms, on a un dwell time de 0,5 ms et une bande passante de 2000 Hz : il y aura donc une différence de 2000Hz entre la plus basse fréquence du spectre et la plus haute. Si cette acquisition est réalisée sur un appareil à 3T donc avec une fréquence de résonance de 127,7 MHz, le spectre couvrira donc les fréquences allant de 127 699 000 Hz à 127 701 000 Hz; ce qui représente une plage de 2000/127,7 ≈15,7ppm.

2.3 Choix des paramètres de l'acquisition

2.3.1 Clinique

L'appareil utilisé en clinique est un Siemens Trio Tim 3 T (Erlangen, Allemagne). La plupart des études concernant la prostate utilisent une antenne endorectale permettant de se rapprocher au maximum de la prostate et ainsi d'obtenir un très bon gain. Cependant il se dégrade fortement en s'éloignant de l'antenne, on ne peut donc pas imager la prostate dans son ensemble (Figure 38) avec une bonne qualité, notamment en spectroscopie.

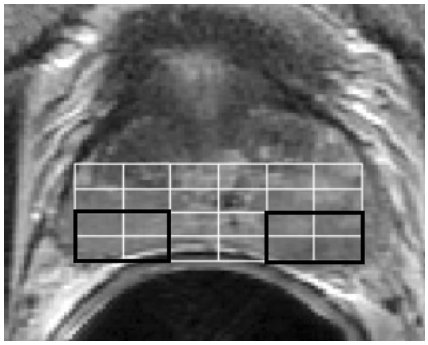


Figure 38 : IRM de prostate avec antenne endorectale et grille de spectroscopie.
La décroissance rapide du signal avec l'éloignement de l'antenne ne permet pas de faire de la spectroscopie sur toute la prostate (image extraite de [56]).

Les études utilisant ce type d'antenne se basent sur le fait que les cancers de la prostate se développent préférentiellement dans la zone périphérique donc la plus proche de l'antenne et par conséquent l'analyse de cette zone est suffisante.

Cette antenne devant être gonflée (notamment pour empêcher la présence de gaz entre l'antenne et la prostate), elle présente en outre l'inconvénient de déformer la prostate (Figure 38) rendant ainsi impossible toute comparaison avec d'autres modalités d'imagerie médicale. Or dans le cadre d'autres études effectuées parallèlement à celle-ci, nous devions avoir cette possibilité. Notamment dans le cadre de la curiethérapie, le calcul du volume de

la prostate, nécessaire pour connaitre le nombre de grains radioactifs scellés à utiliser, se fait par une échographie. Une étude est donc en cours pour évaluer la possibilité de faire ce calcul directement avec les données de l'IRM qui est faite dans tous les cas : cela permettrait d'économiser un examen.

Enfin l'utilisation de ce type d'antenne est très inconfortable pour le patient. Pour ces raisons nous avons choisi de ne pas l'utiliser au profit de l'antenne pelvienne.

La prostate étant une glande disposant d'un volume, il était naturel de faire de l'acquisition 3D et donc du CSI, d'autant plus que cette technique permet d'avoir un meilleur rapport signal sur bruit, et ainsi d'avoir une meilleure résolution. Le volume de la prostate restant limité, il est encore possible de réaliser une bonne correction de l'homogénéité du champ dans ce volume.

Le citrate présente des groupements chimiques possédant un couplage spin-spin, ce qui fait que la configuration du spectre de ce métabolite change avec le TE. Le fait que les premiers spectres étaient analysés visuellement par les spectroscopistes ont conduit à choisir une configuration du spectre aisément identifiable tout en conservant un temps d'écho le plus court possible pour conserver un maximum de signal. C'est pourquoi le TE de 140 ms a été retenu sur la machine Siemens 3T. Le temps de répétition devait quand à lui être supérieur à ces 140 ms et surtout le plus long possible pour ne pas trop perdre de signal sur les métabolites à T1 long comme la choline (1500ms) mais ne pas être trop long pour augmenter le nombre d'acquisitions et ainsi améliorer le rapport signal sur bruit. La valeur de 720 ms a donc été choisie : elle est bien inférieure au T1 de la choline et engendre donc un tassement de son signal mais permet d'avoir un nombre d'acquisitions suffisant (25) et donc un bon rapport signal sur bruit dans un temps d'examen total supportable pour le patient (13 mn uniquement pour la spectroscopie, l'examen total durant environ 45 mn).

L'acquisition se fait dans une matrice de 8 coupes, 12 lignes et 16 colonnes et l'appareil réalise une interpolation pour nous fournir 8 coupes, 16 lignes, et 16 colonnes.

Les spectres sont échantillonnés sur 512 points avec une bande passante de 1250 Hz (soit 1250 Hz / 123,218MHz ≈10,14 ppm).

2.3.2 Pré-clinique

En pré-clinique, l'appareil utilisé est un Bruker Biospin Pharmascan à 4,7T (200MHz). Il est dédié à l'étude du petit animal (diamètre maximal efficace de l'antenne : 7,2 cm). Cet appareil permet d'utiliser toutes les modalités d'imagerie de la résonance magnétique. Par contre, la spectroscopie multi-voxels souffre de faiblesses de

l'implémentation du logiciel d'acquisition et n'est pas utilisable en routine. Nous avons donc dû nous contenter de spectroscopies mono-voxels. Etant donné l'importance d'avoir un nombre significatif de spectres pour pouvoir utiliser les algorithmes de classification, l'absence de cette antenne a été préjudiciable puisque nous n'avons pas pu réaliser de classification sur ce type de données.

Pour être introduits dans le tunnel de la machine, les rats sont anesthésiés par un flux constant de gaz isoflurane (2-3%). La durée maximale d'un examen est donc limitée à une heure : au-delà, la température corporelle peut atteindre une valeur trop basse et les rats souffriraient. Cette durée maximale d'examen laisse donc une marge conséquente dans le choix des paramètres.

N'ayant pas les contraintes de la spectroscopie 3D, nous avons donc opté pour le temps d'écho le plus court (11 ms) disponible sur le système afin de maximiser le rapport signal sur bruit. Le temps de répétition étant, lui, de 2500 ms. Les spectres cérébraux ne présentant pas de contraintes interdisant ces paramètres, ont aussi été acquis avec ces valeurs.

2.4 Ensembles de données

2.4.1 Cas clinique

Bien que chaque semaine, cinq examens de la prostate sont réalisés et ce depuis 2007, la base de cette étude ne comporte que 22 patients. La mise au point du protocole a pris beaucoup de temps et il n'a été finalisé qu'en janvier 2008. Ce n'est qu'à partir de cette date que nous avons pu commencer à constituer une base de données fiable et homogène. De plus pour appliquer une méthodologie de classification supervisée nous avons besoin de connaitre la "vérité de terrain", c'est-à-dire de connaitre pour chaque voxel défini par l'IRM, son état d'envahissement par la tumeur. Il était initialement prévu d'obtenir cette référence par l'anatomopathologie : il s'agit de la spécialité médicale chargée d'analyser les pièces chirurgicales afin d'établir le diagnostic définitif et de vérifier si l'exérèse de la tumeur est complète. Pour cela, nous avons sélectionné une trentaine de patients devant subir une ablation totale de la prostate et nous leur avons fait passer une spectroscopie de résonance magnétique. Cependant la mise en correspondance des données histologiques et d'IRM s'est avérée plus difficile que prévue. Cela est dû à la déformation de la prostate lors de son prélèvement et au fait que les anatomopathologistes ont un protocole de découpe qui n'est pas compatible avec les coupes d'IRM.

Nous avons donc finalement décidé de définir la "vérité de terrain" par un expert observant les spectres. Ce travail est lui aussi fastidieux mais le résultat est plus fiable. M. Walker a donc classé les spectres de 22 patients. Pour chaque voxel, il a dû définir sa localisation (hors prostate, zone périphérique ou glande centrale) et son état (indéterminé, sain, pathologique) en ne se fiant qu'au spectre. Les critères étant :

- Indéterminé : le bruit ne permet d'identifier ni le pic de choline ni les pics de citrate.
- Pathologique : le pic de choline dépasse en amplitude les trois pics de citrate (ses expériences l'ayant amené à constater que les spectres pathologiques présentaient un rapport choline/citrate supérieur à 0,4).
- Sain : sont considérés comme sains tous les autres spectres.

Parmi tous ces spectres, nous n'avons conservé que ceux qui étaient localisés en zone périphérique où se situe la majorité des tumeurs. Notre base se compose donc de 1062 spectres répartis ainsi :

- 286 indéterminés,
- 636 sains,
- 140 pathologiques.

Les patients étaient âgés de 56 à 81 ans (âge moyen : 67,9 ; écart type : 6,7). Leur taux de PSA s'échelonnait de 4 à 61 ng/mL (moyenne de 11,3 et écart type de 11,9) et chacun a au moins une biopsie positive. Les examens de résonance magnétique se sont déroulés au moins six semaines après la dernière biopsie pour éliminer les risques d'artéfact dus aux épanchements de sang.

2.4.2 Cas préclinique

Les acquisitions en résonance magnétique sur les rats sont faites au sein de l'entreprise Oncodesign, sur un appareil Bruker Pharmascan 4.7 T. Cette machine est située dans un local EOPS (Exempt d'Organismes Pathogènes Spécifiques), pour éviter de contaminer les animaux puisque leurs défenses immunitaires sont amoindries. L'IRM et la SRM étant de nouvelles méthodes d'acquisition au sein de la société, il a donc fallu se familiariser avec le matériel puis établir les protocoles d'acquisition.

L'anatomie cérébrale est mieux connue que celle de la prostate et la littérature nous fournit les concentrations standards des différents métabolites. De plus les tissus sont

beaucoup plus homogènes dans le cerveau (il n'y a "que le cerveau") que dans la prostate (qui est entourée de différents organes), il est donc plus simple d'y obtenir un champ magnétique homogène. Pour ces raisons, les premiers étalonnages de l'appareil ont donc été réalisés dans le cerveau. Des souches de gliomes humains (modèle CGL 9 et U87) ont été implantées dans un des lobes cérébraux des rats.

Chez le rat, la prostate est soumise aux mouvements respiratoires, donc avant de faire des acquisitions dans cette zone, nous avons effectué une série de tests avec des rats auxquels une greffe de cellules tumorales (CGL 9 et U87) a été effectuée en sous-cutané au niveau des côtes. Ainsi les acquisitions se faisaient dans un tissu connu mais elles étaient soumises aux mouvements respiratoires.

Après avoir maitrisé la technique sur les gliomes implantés en sous cutané, ces modèles tumoraux ont été remplacés par des modèles de tumeurs prostatiques (PC3 MM2 et PAC 120). Le modèle PAC 120 n'a présenté aucun pic significatif.

Connaissant le profil spectroscopique de ces deux modèles de tumeur, une dernière série d'acquisitions a été réalisée en orthotopique (i.e. dans la prostate). Pour ces acquisitions, différents groupes ont été créés pour tester différents traitements et ainsi observer leurs effets en SRM.

Pour le cerveau, des acquisitions ont été réalisées systématiquement d'une part dans la tumeur et d'autre part en zone cérébrale saine. Bien évidement, en sous cutané, seule l'acquisition dans la tumeur est réalisée puisque il n'y a pas de tissu sain à proprement parler : la tumeur n'est pas greffée dans un organe.

Pour la prostate, nous espérions pouvoir réaliser des acquisitions 3D pour multiplier les données obtenues et ainsi avoir à la fois des spectres de zones saines et de zones pathologiques. Malheureusement, cette technique s'est révélée impossible à mettre en place sur cet appareil.

Les acquisitions se faisaient généralement par groupe de 3 à 5 rats. Pour les expériences avec traitements thérapeutiques, un groupe de contrôle ne recevait pas de traitement et un groupe recevait le traitement.

Chapitre 3 : Prétraitement des données

3.1 Introduction

Un spectre de résonance magnétique est par nature un signal complexe (au sens mathématique du terme), mais seule la partie réelle est habituellement analysée par les spectroscopistes. Même avec cette restriction, le signal reste difficile à analyser notamment à cause de l'impossibilité d'avoir un champ parfaitement homogène dans le volume d'étude de la prostate. Cela entraîne un certain nombre de phénomènes qui perturbent les spectres obtenus et les rendent inexploitables en l'état, la phase de classification pouvant être sensible entre autre à la dispersion des données.

Il est donc nécessaire de prétraiter les spectres de SRM. Nous avons pour cela utilisé une modélisation couramment employée en analyse de spectres [57, 58] :

$$\hat{Y}(\upsilon_k) = S(\upsilon_k)\mathrm{e}^{(-i\Phi(\upsilon_k))}$$

Avec $S(\upsilon_k) = \rho M(\upsilon_k) + B(\upsilon_k) + \varepsilon(\upsilon_k)$

Dans cette expression :

- $M(\upsilon_k)$ est le spectre corrigé que l'on recherche,
- $\mathrm{e}^{(-i\Phi(\upsilon_k))}$ est un facteur de correction de phase,
- ρ est un facteur de normalisation,
- $B(\upsilon_k)$ est dit "ligne de base", sorte de composante continue à éliminer,
- $\varepsilon(\upsilon_k)$ est un bruit blanc gaussien.

Le but des traitements qui vont suivre est d'extraire le signal $M(\upsilon_k)$ du signal d'origine $\hat{Y}(\upsilon_k)$.

Il est à noter qu'au cours des différentes améliorations apportées à la technique d'acquisition, nous avons pu remarquer que le rapport signal sur bruit des spectres obtenus était très bon : le bruit blanc, représenté par le terme $\varepsilon(\upsilon_k)$, a donc été éliminé par un simple filtre Gaussien.

3.2 Traitements mis en œuvre

3.2.1 Correction de phase

Pour être interprétable, il est d'usage en spectroscopie de travailler avec un spectre dit en mode d'absorption (Figure 39).

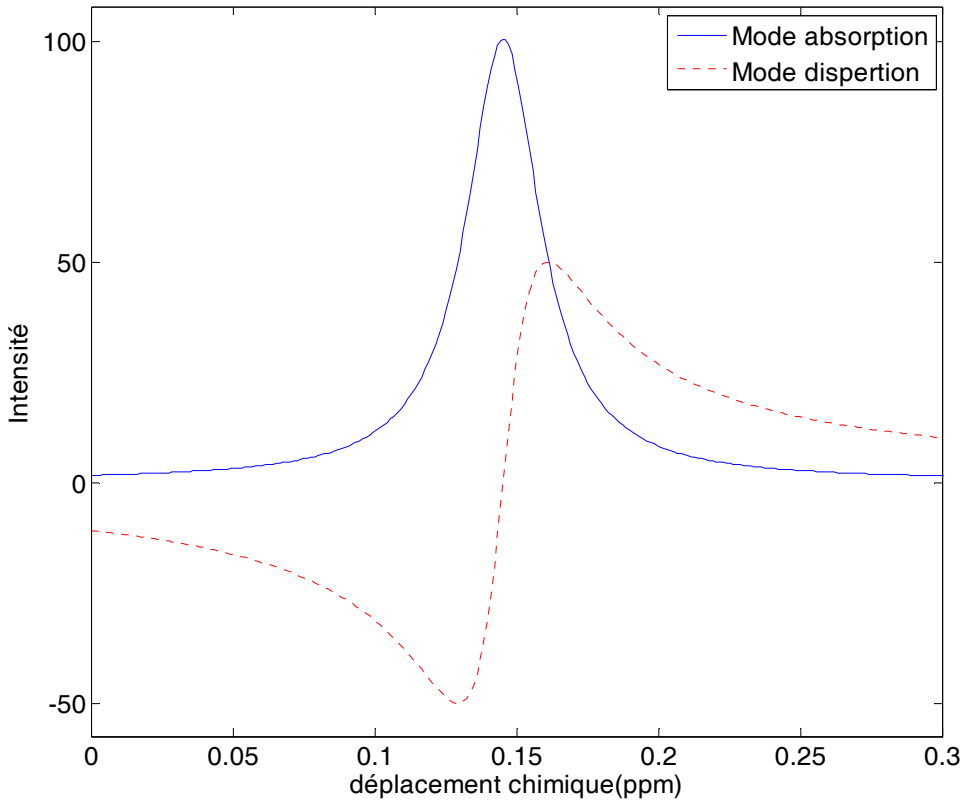


Figure 39 : Les deux modes d'un pic.

Plusieurs méthodes ont été développées pour se rapprocher de ce mode. Nous pouvons citer notamment les algorithmes de correction automatique de phase de Neff [59], l'algorithme DISPA (DISPersion-versus Absorption) par Marshall [60, 61], celui de Brown [62] basé sur l'optimisation de la ligne de base, ceux de Van Vaals [63] en 1990, et Heuer [64] qui se fondent sur des paramètres de symétrie. Plus récemment, la méthode ACME (Automated phase Correction based on Minimization of Entropy) a été développée par Chen en 2002 [65] utilisant une fonction d'entropie; la plus récente étant la méthode de Larry Bretthorst [66] en 2008.

Le but de ces méthodes est de trouver les paramètres de phase qui permettent de se rapprocher de manière optimale du mode d'absorption.

Si on reprend l'expression du spectre $\hat{Y}(\upsilon_k) = S(\upsilon_k)\mathrm{e}^{(-i\Phi(\upsilon_k))}$

et soit n le nombre de points dans le spectre alors $\Phi(\upsilon_k)$, la phase totale est donnée par :

$$\Phi(\upsilon_k) = \varphi_0 + \frac{\upsilon_k}{n}\varphi_1$$

où φ_0 et φ_1 sont les valeurs de correction de zéro ordre et de premier ordre.

Nous avons choisi, pour notre étude, la méthode donnant les meilleurs résultats au moment où nous avons mis en place les prétraitements, à savoir la méthode ACME proposée par Chen [65].

L'algorithme ACME tente donc de trouver les meilleures valeurs pour les paramètres φ_0 et φ_1 qui minimiseront une fonction d'objectif E. Cette fonction est composée de deux termes. Le premier, noté H, est basé sur la dérivée normalisée de la partie réelle du spectre et correspond à une mesure de l'entropie de l'information du type Shannon [67]. Le second, noté P, est une fonction de pénalité qui permet de minimiser les parties négatives du spectre.

Soit $R_{\hat{Y}}(\upsilon_k)$ et $I_{\hat{Y}}(\upsilon_k)$ les parties réelle et imaginaire de $\hat{Y}(\upsilon_k)$ alors H est donné par :

$$H = \sum_{i=1}^{n-3} -D_i \ln\left(D_i\right)$$

avec

$$D_i = \frac{d\mathrm{R}_{\hat{Y}}(\upsilon_i)}{\sum_{k=1}^{n-3} d\mathrm{R}_{\hat{Y}}(\upsilon_k)}$$

et

$$d\mathrm{R}_{\hat{Y}}(\upsilon_i) = \left|\frac{R_{\hat{Y}}(\upsilon_{i+2}) - R_{\hat{Y}}(\upsilon_i)}{2}\right|$$

P la fonction de pénalité est donnée par :

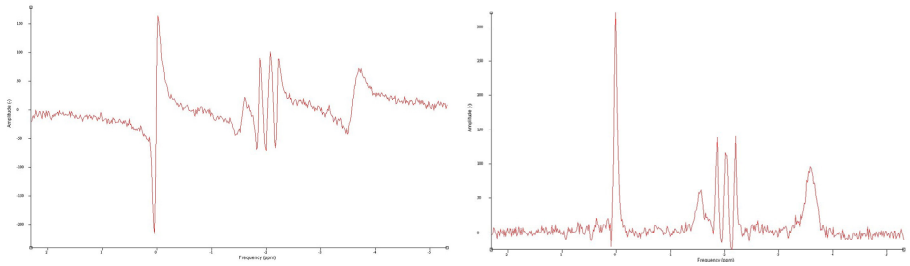
$$P = \gamma \sum F\left(R_{\hat{Y}}(\upsilon_i)\right)\left(R_{\hat{Y}}(\upsilon_i)\right)^2$$

γ est un facteur de pénalité : sa valeur doit être ajustée pour équilibrer la contribution des termes H et P. La fonction F est définie par :

$$F(x) = \begin{cases} 0, x \geq 0 \\ 1, x < 0 \end{cases}$$

La fonction $E = H + P$ est alors minimisée par une méthode du simplex [68, 69].

Nous avons représenté sur la Figure 40 un spectre avant et après sa correction par l'algorithme ACME.



a b

Figure 40: Spectre avant (a) et après (b) l'application d'une correction de phases de zéro et de premier ordre.

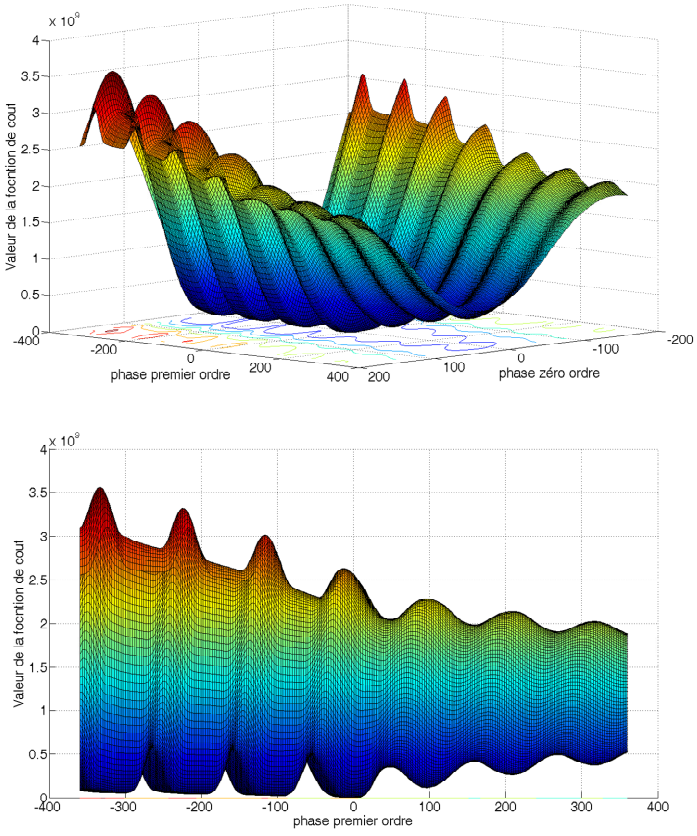


Figure 41 : Variation de la fonction d'entropie en fonction des phases de zéro et de premier ordre.

Nous avons représenté Figure 41 un exemple de variation de la fonction d'entropie en fonction des phases de zéro et de premier ordre. On constate bien l'existence d'un

minimum global, situé ici en -6 pour la phase de zéro ordre et 0 pour la phase de premier ordre.

Une autre méthode parfois employée, pour sa simplicité, consiste à calculer le module du spectre complexe. Cette solution est une bonne approximation dans les cas où la correction de premier ordre nécessaire est faible et ce quelle que soit la correction de zéro ordre. Cependant dans les cas où la correction de premier ordre est plus importante, cette approximation n'est plus valable et des algorithmes plus perfectionnés comme celui mentionné ci-dessus sont nécessaires. Le calcul du module présente également l'inconvénient d'élargir les pics ce qui peut d'une part rendre difficile la distinction de deux pics proches et d'autre part fausser l'estimation de concentration des métabolites réalisée en calculant l'aire sous la courbe.

3.2.2 Suppression de la ligne de base

Les macromolécules sont de très grosses molécules qui ont un spectre de résonance très large qui vient s'additionner à tout le spectre. A cela peut s'ajouter une mauvaise correction de la phase de premier ordre, ce qui fait que la base des pics n'est pas proche de zéro ; ceci peut être très gênant pour la quantification de la concentration des métabolites.

La détermination de la ligne de base est un problème dit "mal posé", où seules des solutions approximatives existent [26, 70]. Cependant plusieurs techniques ont été mises au point pour l'estimer au plus juste. Devos [26] propose de multiplier le signal temporel par une fonction exponentiellement décroissante

$$e^{-\beta n \Delta t}, \quad n = 0,...,N-1$$

où N est le nombre de points du signal et β un paramètre permettant de régler la hauteur de la ligne de base.

Si β est trop petit, la ligne de base sera trop plate et certains pics seront surestimés, mais s'il est trop grand, la ligne de base risque de prendre une partie du signal de certains pics qui seront alors sous-estimés : Devos suggère d'utiliser $\beta = 0.15$. En appliquant la transformée de Fourier à ce signal, il obtient alors un modèle de la ligne de base du spectre. Cette méthode donne, à notre sens, une ligne de base trop élevée (Figure 42).

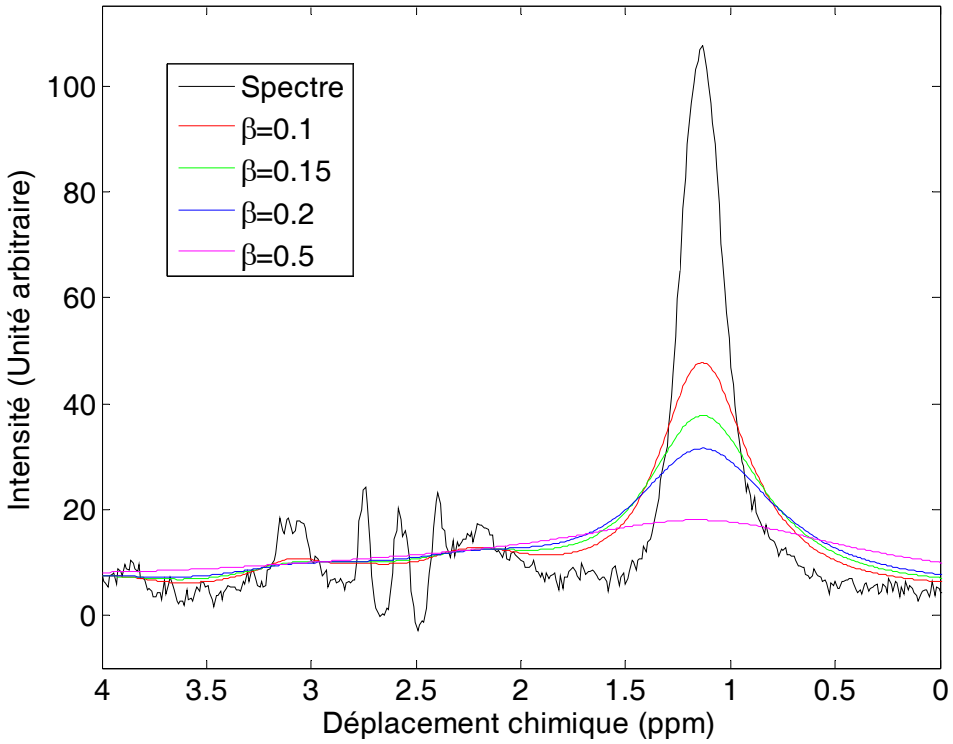


Figure 42 : Ligne de base obtenue par la méthode de Devos en fonction du paramètre β.

De nombreuses autres méthodes ont été développées[58, 71, 72], mais elles sont intégrées à la méthode de quantification, or nous voulons justement extraire la ligne de base sans faire d'estimation de concentrations.

La technique de calcul de la ligne de base que nous avons employée est basée sur la technique de Lieber et al [73] et consiste à appliquer un filtre passe bas de manière itérative. A chaque itération, nous éliminons du spectre tous les points supérieurs au spectre lissé par le filtre passe bas (Figure 43). Nous avons vérifié expérimentalement que l'algorithme converge la plupart du temps avec environ 200 itérations (Figure 44) sans pénaliser l'ensemble de l'algorithme en termes de temps de calcul. La ligne de base est ensuite soustraite du spectre d'origine. Dans la suite de ce manuscrit, nous nommerons cette technique de Correction de la Ligne de Base : CLB.

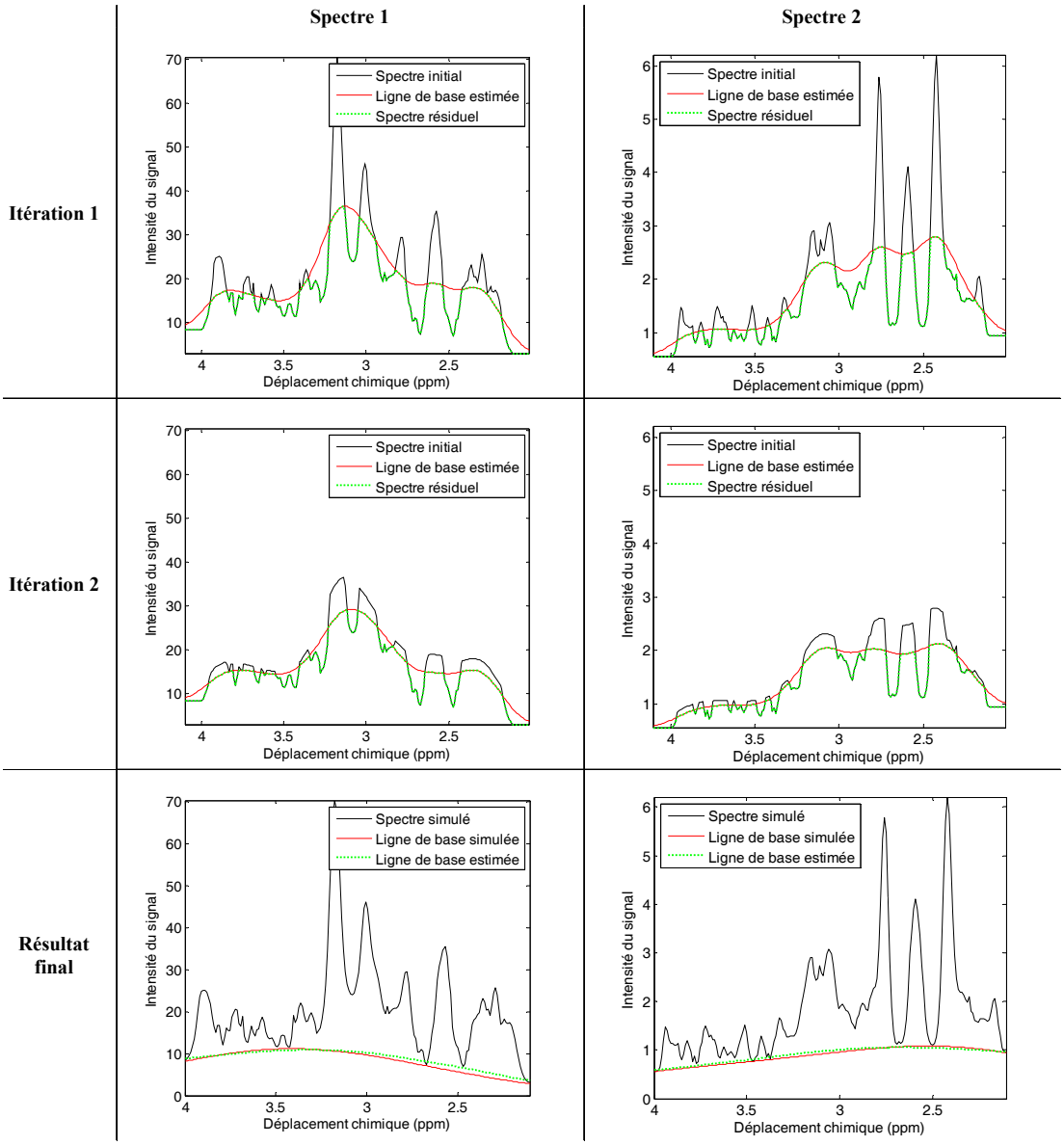


Figure 43 : Processus de détection de la ligne de base.

A chaque itération, le spectre (en noir) subit un filtrage passe bas (convolution avec signal gaussien) ce qui donne le signal représenté en rouge. Seule la partie inférieure (en vert) est conservée pour la prochaine itération. Dans le résultat final la ligne de base simulée est en rouge et la ligne de base estimée est en vert.

Pour évaluer ces algorithmes, nous avons sélectionné 100 spectres aléatoirement dans notre base de données. Chaque spectre a été approximé par une somme de Lorentziennes (cf. 3.2.4.2b). Le spectre a alors été reconstitué et une ligne de base simulée a été rajoutée. Cette ligne de base est constituée par la somme de 2 Lorentziennes dont la fréquence est tirée aléatoirement dans la plage du spectre et l'amplitude est également tirée aléatoirement entre 0 et 20% de l'amplitude maximale du spectre. La largeur à mi-hauteur est fixée à 5 ppm. Nous avons alors mesuré la différence quadratique entre la ligne de base simulée et la ligne de base estimée (différence normalisée par l'énergie du spectre simulé),

en fonction du nombre d'itérations pour la méthode de Lieber et du paramètre β de la méthode de Devos. Les résultats sont reportés Figure 44.

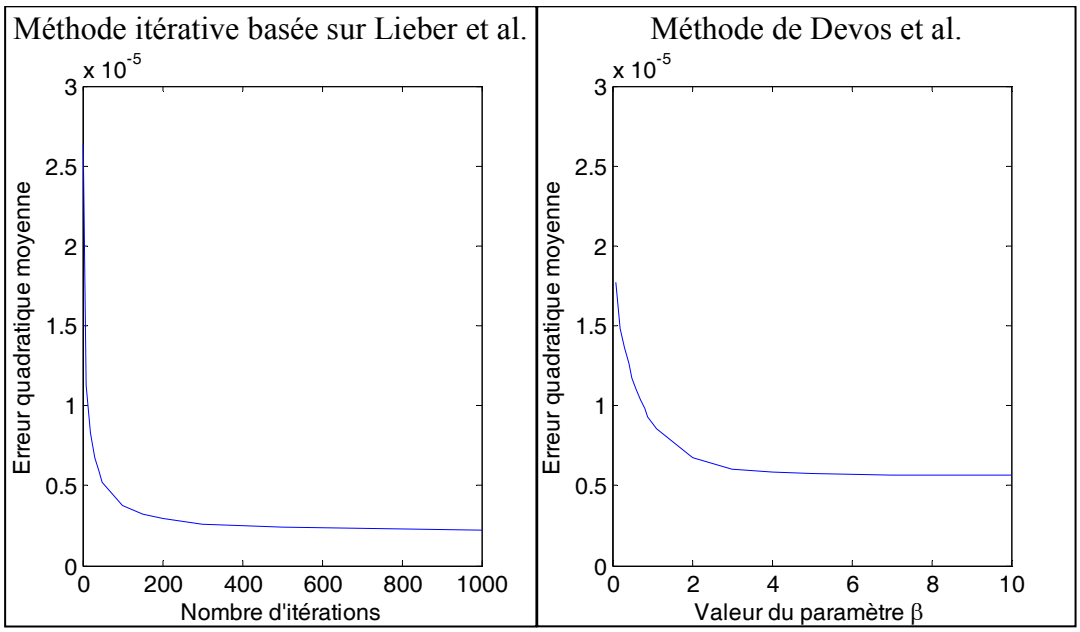


Figure 44 : Erreur quadratique entre la ligne de base simulée et la ligne de base estimée pour les deux méthodes évaluées.

Les deux méthodes convergent bien, mais l'erreur obtenue par la méthode de Lieber est finalement plus faible (quasiment d'un facteur 2). C'est donc cette méthode que nous avons retenue.

3.2.3 Normalisation

Il existe en SRM du proton principalement deux facteurs modifiant l'intensité du signal et donc l'énergie du spectre. Ces deux facteurs sont le volume du voxel (plus le voxel contient de protons, plus il renvoie de signal) et la forme de l'onde radiofréquence utilisée pour l'excitation des protons.

Le volume d'acquisition est cadré sur la prostate du patient, mais comme aucun patient n'a la même prostate (variation de volume de 10 à 120 cm^3) alors que la grille d'acquisition reste fixe (8 coupes de 16x16), le volume des voxels n'est pas le même pour chaque examen.

De plus, la forme de l'onde d'excitation permettant de faire basculer l'angle de précession des protons n'est pas parfaite et est plus efficace au centre du volume d'acquisition que sur les bords. Il en résulte un signal d'intensité plus forte au centre du volume d'acquisition, comme l'illustre la Figure 45 pour le pic de l'eau.

Il est donc nécessaire de normaliser les spectres pour que l'écart type d'un attribut donné (une abscisse ou valeur en ppm), pour l'ensemble des spectres, soit le plus faible possible, pour des tissus de même catégorie.

Comme nous l'avons vu dans la partie 1.2.2, plusieurs méthodes de normalisation existent comme la normalisation par l'eau (division de chaque valeur du spectre soit par l'intégrale du pic d'eau, soit plus simplement par l'intensité maximale de ce pic), la normalisation L1 (division de chaque attribut par la somme de la valeur absolue de chaque attribut du spectre) ou encore la normalisation L2 (division de chaque attribut par la somme des carrés de chaque attribut). Comme cela est suggéré notamment par le laboratoire SCD-SISTA de la Katholieke Universiteit Leuven [26, 28], la normalisation par la norme L2 est celle qui leur a permis d'obtenir les taux d'erreurs les plus faibles et c'est donc celle que nous avons retenue.

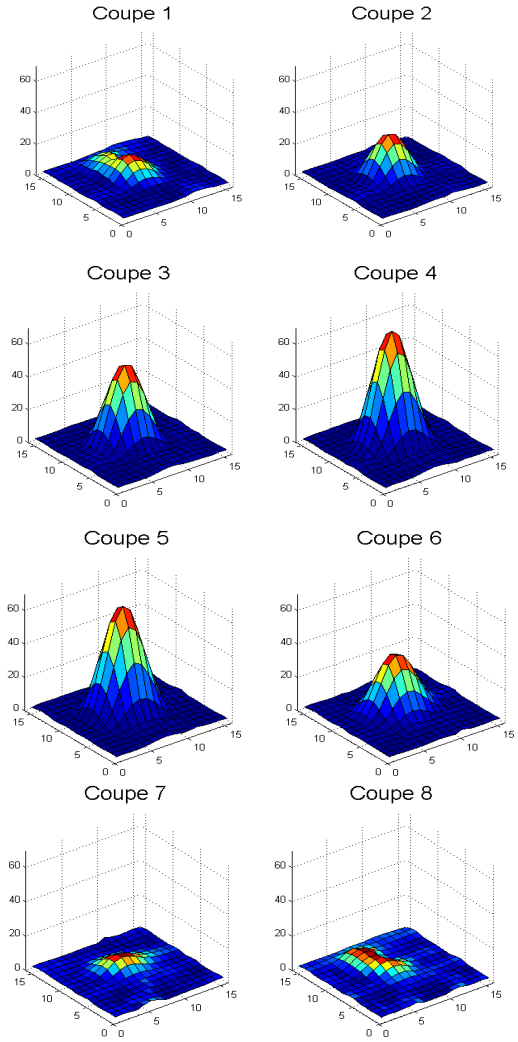


Figure 45 : Intensité du pic d'eau par coupe (acquisition de 8 coupes de 16x16 voxels).
L'intensité est plus forte au centre du volume d'acquisition.

3.2.4 Estimation des concentrations des métabolites

Un spectre est composé de pics. L'aire sous chaque pic est proportionnelle au nombre de protons résonnant à la fréquence correspondant au maximum du pic. Donc l'aire sous un pic est proportionnelle à la concentration de la molécule dont les protons résonnent à cette fréquence. En estimant cette aire, il est donc possible de réaliser un dosage des molécules correspondantes. Pour cela, il est nécessaire d'une part de caractériser chaque pic et d'autre part d'identifier automatiquement le métabolite auquel il correspond.

On trouve dans la littérature de nombreuses méthodes pour réaliser cette estimation. La méthode développée par S. Provencher [74] dans les années 1990 et implémentée dans le logiciel LCModel est certainement la plus utilisée actuellement dans le milieu médical, c'est pourquoi nous l'avons étudiée dans un premier temps. Toutefois, l'algorithme présentant certaines faiblesses dans le cas des spectres de la prostate, nous avons développé notre propre méthode, plus spécifique.

3.2.4.1 LCModel

LCModel est donc un outil logiciel commercial d'analyse de spectres largement utilisé dans le domaine médical et cité comme référence dans de nombreuses études scientifiques concernant le cerveau. Il réalise une estimation des concentrations de chaque métabolite recherché. Comme son nom l'indique, il considère tout spectre comme une combinaison linéaire (Linear Combination Model) des spectres de chaque métabolite constituant l'échantillon d'étude.

$$\hat{Y}(\upsilon_k) = \exp(-i\Phi(\upsilon_k))\left[\sum_{j=1}^{N_B}\beta_j B_j(\upsilon_k) + \sum_{l=1}^{N_M} C_l \sum_{n=-N_s}^{N_s} \Gamma_n M_l(\upsilon_{k-n};\xi_l,\delta_l)\right]$$

Avec les contraintes :

$$C_l \geq 0,\ \xi_l \geq 0,\ \sum_{n=-N_s}^{N_s} \Gamma_n = 1$$

Et où :

- La ligne de base est représentée par $\sum_{j=1}^{N_B}\beta_j B_j(\upsilon_k)$, une somme de N_B B-splines cubiques uniformes, $B_j(\upsilon_k)$,
- C_l sont les concentrations de chaque métabolite,

- $M_l(\upsilon_{k-n};\xi_l,\delta_l)$ modélise le spectre du métabolite l : ξ_l est un paramètre permettant de tenir compte du temps T2* plus court in-vivo et δ_l est un paramètre pour tenir compte des possibles petits décalages en fréquence,
- Γ_n est un facteur de normalisation.

L'algorithme cherche à minimiser la différence entre ce modèle et le spectre réel. Initialement LCModel utilisait comme modèle de spectre de métabolite (M_l), un spectre obtenu avec une solution du métabolite in vitro. Il a ensuite été modifié à notre demande pour traiter les spectres de prostate : la base de spectres n'est plus obtenue in vitro mais est simulée par un logiciel (intégré à LCModel) tenant compte des paramètres d'acquisition de l'échantillon à étudier. Un exemple de sortie LCModel est représenté sur la Figure 46.

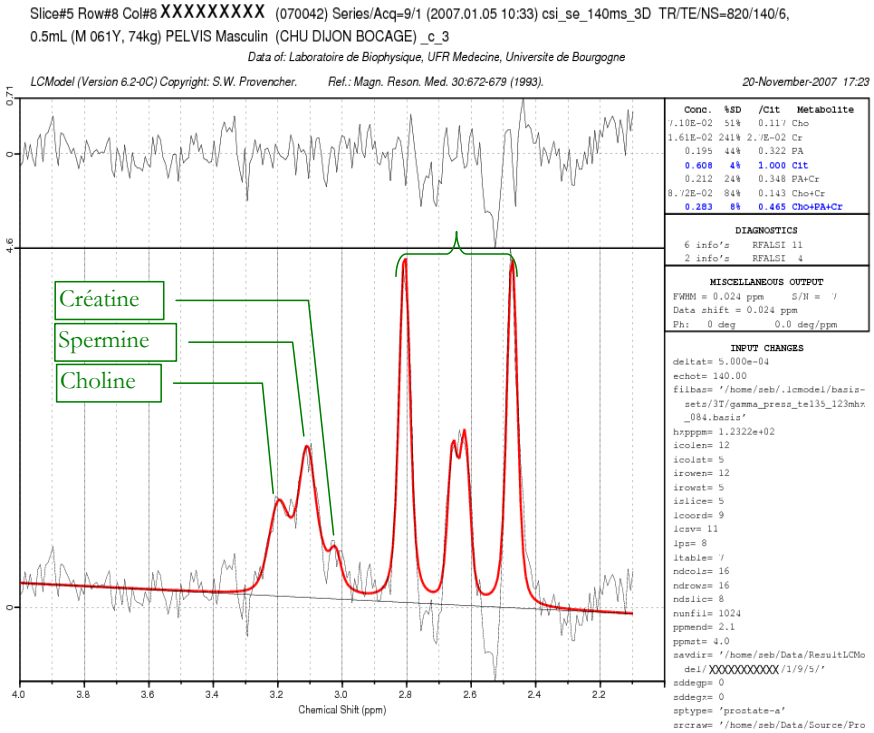


Figure 46 : Spectre analysé avec LCModel.
En rouge, apparaît le spectre estimé à partir du spectre d'origine (noir) et à droite apparaît, pour chaque métabolite reconnu, sa concentration estimée.

Les spectres du cerveau contiennent beaucoup plus d'informations que ceux de la prostate, puisque qu'on peut y détecter une vingtaine de métabolites. Ce grand nombre d'éléments détectables dans les spectres cérébraux permettent à LCModel d'avoir beaucoup de points de repère facilitant la phase d'optimisation qui détermine le spectre estimé.

En outre, pour le cerveau, pour lequel il est très efficace, LCModel dispose d'une base de spectres pour modéliser les spectres du cerveau sain et d'une autre base pour

modéliser les spectres des tumeurs du cerveau. Le choix du modèle n'est pas automatique, il faut donc connaitre l'état, sain ou pathologique, du spectre pour pouvoir lui appliquer le bon modèle. Dans le cas de la prostate, on ne dispose que d'un modèle pour traiter à la fois les spectres sains et les spectres pathologiques, qui sont pourtant très différents : la prostate saine produit du citrate alors que dans la tumeur, la production de citrate est bloquée et l'on observe une apparition de choline. Ce schéma n'est donc pas tout à fait adapté à notre utilisation.

De plus, les temps d'analyse de LCModel sont relativement longs (de l'ordre de 6 à 10 secondes par spectre, soit une quinzaine de minutes par prostate. Ces différentes limitations nous ont conduit à développer notre propre algorithme plus spécialisé mais aussi plus rapide.

3.2.4.2 Modélisation par des fonctions élémentaires

Dans l'introduction de ce chapitre (3.1), nous avons décomposé un spectre de la manière suivante :

$$\hat{Y}(\upsilon_k) = S(\upsilon_k)\exp\left(-i\Phi(\upsilon_k)\right) \quad avec \quad S(\upsilon_k) = \rho M(\upsilon_k) + B(\upsilon_k) + \varepsilon(\upsilon_k)$$

où $M(\upsilon_k)$ est le signal que l'on veux étudier. Provencher propose de décomposer ce signal par la somme des spectres (simulés ou acquis in vitro) des métabolites censés composer l'échantillon. Nous proposons de considérer que ce signal est composé d'une somme de signaux modélisés par une fonction élémentaire. Plusieurs fonctions ont déjà été étudiées dans la littérature, la plupart du temps basées sur des fonctions de Gauss, de Lorentz ou de Voigt.

a La fonction élémentaire

Nous avons vu dans la partie 2.1.1.3 que le signal de RMN suivait une décroissance exponentielle et que le spectre était issu de la transformée de Fourier discrète de celui-ci. Or la transformée de Fourier d'une fonction en décroissance exponentielle est une fonction de Lorentz (Figure 47) d'équation :

$$L(x) = \frac{\Gamma}{2\pi}\frac{1}{\left(\frac{\Gamma}{2}\right)^2 + (x-\mu)^2} \tag{4}$$

avec μ abscisse du maximum et Γ la largeur à mi-hauteur.

Cependant plusieurs paramètres viennent perturber le signal (notamment les filtres électroniques appliqués au signal par la machine), et la forme des pics se rapproche alors d'une gaussienne (Figure 47) (elle-même transformée de Fourier d'une gaussienne).

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (5)$$

avec μ abscisse du maximum et σ l'écart-type.

Ainsi, il est communément admis [75] que la forme réelle des pics de RMN respecte une fonction de Voigt qui est la convolution d'une fonction gaussienne et d'une fonction lorentzienne de même sommet.

$$V(x) = (G \circ L)(x) = \int_{-\infty}^{\infty} L(t)G(x-t)dt \quad (6)$$

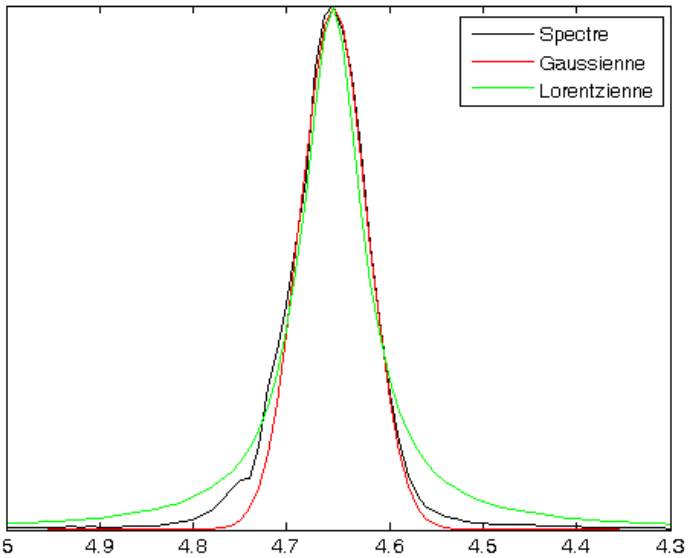


Figure 47 : Approximation du spectre de l'eau par une gaussienne et une lorentzienne.

Dans le cadre de cette étude, nous nous sommes restreints à l'utilisation de fonctions gaussiennes ou lorentziennes exprimées de la manière suivante :

$$L_{\mu,\alpha,\Gamma}(x) = \alpha \frac{1}{1 + 4\frac{(x-\mu)^2}{\Gamma^2}}$$

où $L_{\mu,\alpha,\Gamma}(x)$ est une lorentzienne ayant pour maximum le point (μ,α) et pour largeur à mi-hauteur Γ, et

$G_{\alpha,\beta,\mu}(x) = \alpha e^{-\frac{(x-\mu)^2}{\beta}}$, en posant $\alpha = \frac{1}{\sigma\sqrt{2\Pi}}$ et et $\beta = 2\sigma^2$

où $G(x)$ est l'équation d'une fonction gaussienne ayant pour sommet le point (μ,α).

Les paramètres à déterminer automatiquement lors de la modélisation seront donc systématiquement le couple (μ,α), Γ pour les fonctions de Lorentz et β pour les fonctions de Gauss.

b *Approximation du spectre par une somme de fonctions*

La décomposition du spectre en une somme de fonctions élémentaires n'est pas unique. Nous avons choisi une stratégie privilégiant les pics de plus forte amplitude, ceux-ci étant réputés plus stables et moins sensibles au bruit que les pics de faible amplitude. Nous détectons donc tout d'abord le pic de plus grande amplitude, nous le modélisons par la fonction élémentaire, et soustrayons ce modèle au spectre. Le processus est itéré jusqu'à ce qu'un critère d'arrêt soit satisfait.

Soit le spectre exprimé sous forme d'une somme de fonctions élémentaires :

$$M(\upsilon_k) \approx \sum_{i=1}^{n} f_i(\upsilon_k) \tag{7}$$

Nous faisons alors une détection des maxima locaux du spectre en relevant les points où la dérivée s'annule et est décroissante. La dérivée du spectre est approximée par :

$$df(M(\upsilon_k)) = M(\upsilon_k) - M(\upsilon_{k-1})$$

Ensuite les maxima sont détectés par la condition suivante :

$$df(M(\upsilon_k)) \cdot df(M(\upsilon_{k+1})) < 0 \text{ et } df(M(\upsilon_k)) > 0$$

On obtient alors une liste Ps de m points, chacun étant le sommet d'un pic.

$Ps_i \quad i \in [1...m]$ désignant le sommet du ième pic de la liste.

Pour estimer de manière précise les paramètres de la fonction élémentaire qui modélise le pic d'amplitude maximale, nous avons utilisé la méthode du simplex de Nelder-Mead [69], avec comme fonction de coût, la distance entre le pic réel et le pic modélisé.

Afin d'éviter l'influence de pics voisins du pic d'amplitude maximale sur l'estimation des paramètres, nous avons choisi de restreindre la portion du spectre étudiée à la région située entre les points d'inflexion situés de part d'autre du sommet (Figure 48). Ce point d'inflexion est caractérisé par l'annulation de la dérivée seconde :

$$
\begin{aligned}
df_2(M(\upsilon_k)) &= df(M(\upsilon_k)) - df(M(\upsilon_{k-1})) \\
&= \left(M(\upsilon_k) - M(\upsilon_{k-1})\right) - \left(M(\upsilon_{k-1}) - M(\upsilon_{k-2})\right) \\
&= M(\upsilon_k) - 2M(\upsilon_{k-1}) + M(\upsilon_{k-2})
\end{aligned}
$$

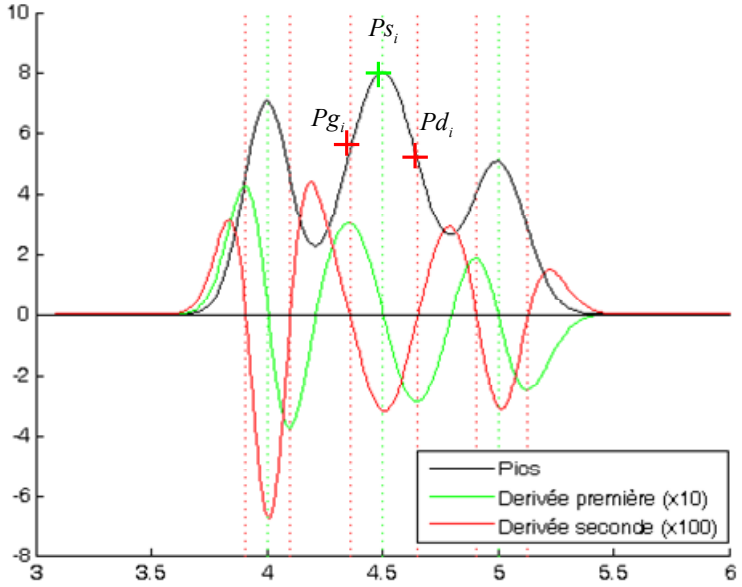


Figure 48 : Simulation d'un spectre et représentation de sa dérivée première et seconde.
L'annulation de la dérivée première marque les maxima du spectre et l'annulation de la dérivée seconde correspond aux points d'inflexion du spectre.

Il est à noter que le spectre a été au préalable sur-échantillonné (8 fois) afin de garantir une bonne précision dans l'estimation des paramètres de la fonction élémentaire.

Notons Pg_i et Pd_i les deux points d'inflexion de chaque coté de Ps_i. Pour chaque pic, nous disposons donc de trois points. Nous pouvons alors définir la fonction de cout par :

$$C(i) = \sum_{k=Pg_i}^{Pd_i} M(\upsilon_k) - f_i(\upsilon_k)$$

où f_i est la fonction choisie pour la modélisation (de Gauss, de Lorentz ou autre).

L'algorithme peut se résumer de la manière suivante :

1. Rechercher les maxima du spectre : Ps
2. Trier Ps par ordre décroissant d'intensité. Ps_1 est alors le sommet de plus grande intensité du spectre.
3. Rechercher les points d'inflexion Pg_1 et Pd_1 correspondant au sommet Ps_1

4. Estimer les paramètres de la fonction choisie (Gaussienne ou Lorentzienne) approchant le mieux le spectre entre Pg_1 et Pd_1.
5. Soustraire la fonction estimée au spectre
6. Reprendre au point 1 jusqu'à ce que le spectre résiduel soit négligeable.

L'évolution de cet algorithme est illustrée sur la Figure 49. On vérifie sur la Figure 50 que la recomposition du spectre en additionnant toutes les fonctions, est une bonne approximation du spectre. La base des courbes de Lorentz étant plus large que celle des courbes de Gauss (Figure 47), la surestimation (Figure 50 à 3,1 ppm) est plus fréquente dans le premier cas que dans le second. Cela est dû au fait que l'approximation n'est évaluée qu'à partir du sommet du pic, mais ce phénomène n'apparait que dans le cas de pics très proches.

Nous obtenons ainsi une liste Pr de m triplets de paramètres (α, μ et Γ avec des courbes de Lorentz, α, μ et β avec des courbes de Gauss) des fonctions permettant d'approximer au mieux le spectre.

Afin d'estimer les concentrations des métabolites, il est alors nécessaire d'identifier les pics, c'est-à-dire de déterminer les métabolites qui leur correspondent. Dans la suite de document nous nommerons cet algorithme DSF pour Décomposition du spectre en une Somme de Fonctions.



Figure 49 : Progression de la détection des pics.

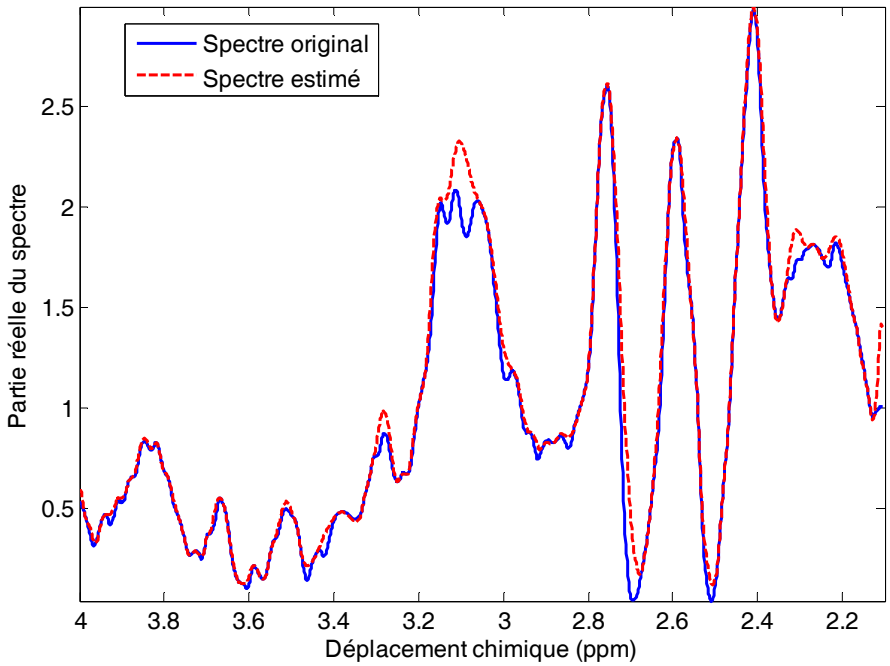


Figure 50 : Spectre et estimation constituée de la somme des gaussiennes détectées.

c Détection des métabolites

Nous disposons d'une liste de paramètres Pr où chaque pic est modélisé par trois valeurs : son amplitude, sa fréquence et une dernière liée à sa largeur à mi hauteur. Nous allons maintenant trouver parmi ces pics ceux pouvant être issus des métabolites qui nous intéressent.

Nous avons à notre disposition la position de référence en ppm des pics des 4 métabolites usuellement quantifiés pour la prostate. Nous noterons ces positions théoriques P_{th}

Tableau 3 : Fréquence de résonance des principaux métabolites détectables dans la prostate en RMN (en ppm).

	Choline	Spermine	Créatine	Citrate		
ppm	3,22	3,11	3,02	2,81	2,64	2,47

Cependant pour de multiples raisons, notamment l'hétérogénéité locale de champs, il peut y avoir une translation de quelques centièmes de ppm, voire même dans des cas extrêmes, de quelques dixièmes de ppm du spectre, il n'est donc pas forcément pertinent d'affecter à chaque métabolite le pic dont la fréquence se rapproche le plus de sa fréquence théorique.

Nous avons mesuré ce décalage sur 451 spectres. Deux opérateurs ont effectué ce travail indépendamment. Nous avons mesuré un taux de corrélation de 0,94. L'amplitude moyenne de la valeur absolue de ce décalage est de 0,21 ppm avec un écart type de 0,22 et une médiane à 0,15.

Nous avons donc essayé de mettre au point une méthode, en se basant sur nos connaissances *a priori* du spectre de la prostate, pour sélectionner le plus pertinemment possible la fonction en cloche correspondant à chaque métabolite.

Plutôt que de chercher à corriger ce décalage en une seule fois, ce qui pourrait parfois nous amener à certaines aberrations, nous avons choisi de décomposer cette recherche en plusieurs étapes (2 à 3 suivant les cas) : la première étape réalise une sélection grossière des pics qui est affinée dans les étapes suivantes ("coarse to fine" : grossière à fine).

Pour optimiser la reconnaissance des métabolites parmi les pics, nous allons utiliser une série de connaissances *a priori* sur les spectres de la prostate : cet algorithme sera donc spécifique. Voici la liste des 6 règles utilisées dans les algorithmes suivants :

(1) Le décalage moyen est de 0.21 ppm et le médian de 0.15 ppm donc en faisant une première recherche dans une plage de 0,2 ppm, nous pouvons corriger plus de la moitié des cas.

(2) Chaque spectre de notre ensemble de données contient au moins un pic correspondant soit à la choline soit aux pics de citrate.

(3) La distance entre les pics ne présente que peu de variation.

(4) Si tous les pics ont une amplitude négligeable par rapport à un autre et que ce dernier se situe dans la zone 3 ppm - 3,4 ppm, on le considère comme de la choline.

(5) Les pic d'amplitude plus forte sont considérés comme les plus fiables et servent de base à la correction du décalage.

(6) Les spectres pathologiques possèdent au moins un pic de choline et les spectres sains présentent un trident caractéristique du citrate. Dans les spectres de transition, on doit détecter tous les métabolites.

L'algorithme de recherche des pics correspondant aux métabolites d'intérêts (choline, spermine, créatine et trois pics de citrate) sera développé en annexe (Annexe 3) mais peut se résumer par :

1) Pour chaque métabolite, une recherche dans une plage de 0.4 ppm autour de sa fréquence de résonnance théorique du pic le plus intense est réalisée et, en fonction du résultat, une estimation du décalage est calculée. Dans la recherche des pics correspondant aux métabolites, les pics de plus forte amplitude sont privilégiés (règle 5).

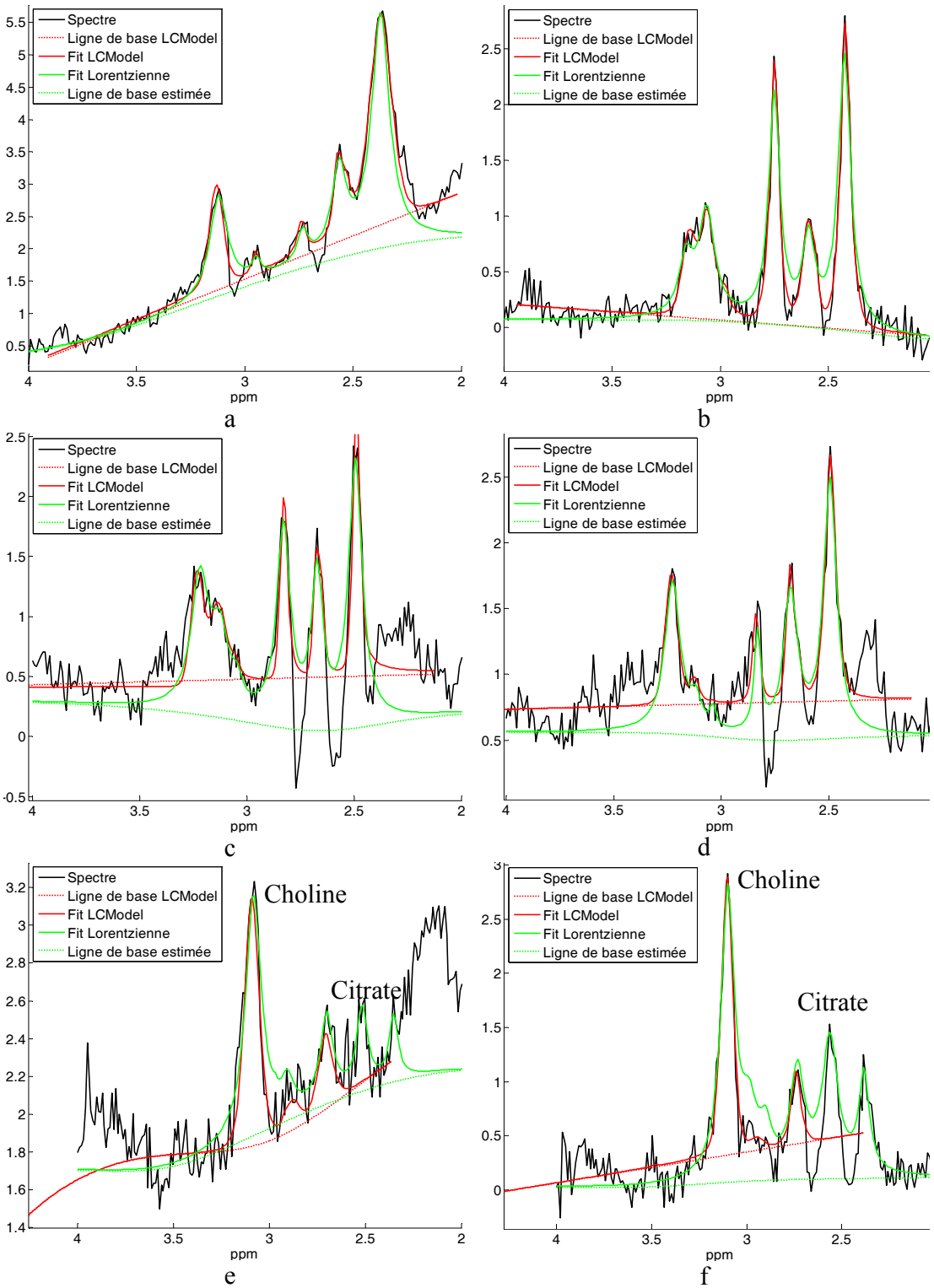
2) Après la correction de décalage du spectre de l'étape précédente, la liste des pics est réinitialisée avec l'algorithme DSF. Pour chaque métabolite, une recherche est effectuée parmi la nouvelle liste dans une plage de 0,2 ppm autour de sa position de référence. Le critère de sélection retenu est que la distance entre les pics sélectionnés soit la plus proche de la distance de référence. En fonction de cette nouvelle sélection, le décalage est ré-estimé.

3) En fonction des résultats des deux premières étapes, la liste des pics peut être réévaluée par l'algorithme DSF, puis la sélection des pics réévaluée avec le même algorithme que l'étape précédente et le décalage est réévalué.

A l'issue de ces trois étapes, nous obtenons donc une liste contenant six pics modélisés par trois paramètres et une estimation du décalage.

Nous noterons dans la suite de cette thèse cet algorithme DM pour Détection de Métabolites. Figure 51, sont représentés six exemples de détection par cette méthode ainsi que par LCModel.

La méthode DM est très rapide : moins de 0,5 s par spectre contre plus de 6 secondes pour LCModel (sur le même PC) et présente l'avantage de mieux détecter la choline que LCModel. Près de 30% des cas pathologiques, qui présentaient des défauts de détection de la choline par LCModel, ont correctement été analysés par notre méthode. La méthode DM reste sensible à la contamination lipidique des pics de citrate (augmentation artificielle de la concentration) et dans le cas de très forts décalages dus à un champ magnétique très hétérogène, notre méthode peut confondre le pic de créatine et le premier pic de citrate (dans ce cas LCModel donne également la plupart du temps une réponse erronée). Nous avons représenté Figure 52 la corrélation entre le résultat fourni par LCModel et celui donné par la méthode DM. Cette corrélation est en général bonne, sauf bien entendu pour les erreurs de LCModel concernant la choline, qui sont particulièrement remarquables.



Cho	5,71x10-4
Cr	0
PA	0
Cit	0,280

Cho	0
Cr	0
PA	0
Cit	0,198

Figure 51 : Exemples de détection de métabolites par Lorentzienne (vert) comparés à LCModel (rouge).

(a, b) Les 2 techniques donnent sensiblement le même résultat, (c, d) la ligne de base de LCModel tronque les pics de citrate, (e, f) les courbes sont semblables cependant le pic de choline est identifié comme du citrate par LCModel (les aires estimées par LCModel sont indiquées sous les courbes et sont à 0 pour la choline).

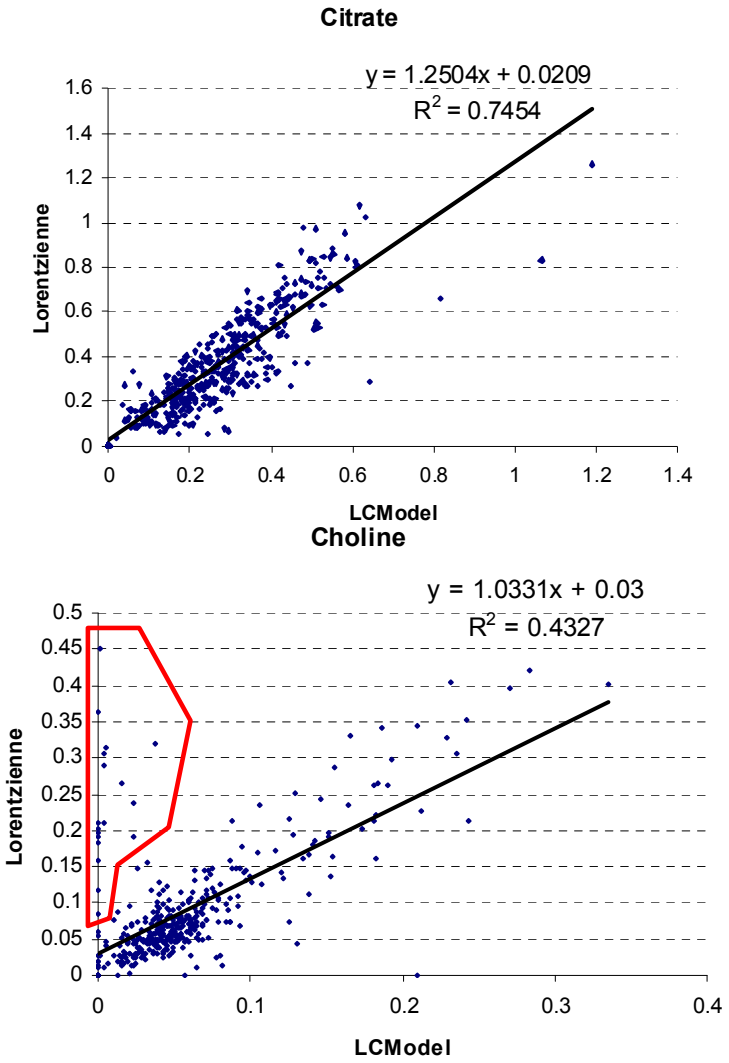


Figure 52 : Représentation de l'aire sous la courbe lorentzienne versus la concentration estimée par LCModel pour le citrate et la choline.
Sur le graphique de la choline sont entourés en rouge les points pour lesquels LCModel a mal estimé la choline (après suppression de ces points, on obtient R^2=0,7647).

Il est à noter que cette technique utilise beaucoup de connaissances *a priori* sur les pics et plus précisément sur l'interdépendance entre ces pics, ce qui fait qu'elle ne peut être appliquée que dans le cas de la prostate humaine. Elle est donc inapplicable directement aux spectres de prostate de rat en raison du très faible taux de citrate, d'une forte présence de lipides dans cette zone (voir Figure 53), et de l'absence de connaissances sur d'éventuelles interdépendances entre pics.

Il est cependant possible dans ce cas d'utiliser un algorithme simplifié, ne tenant pas compte des interdépendances entre pics, en faisant une estimation du spectre par la

méthode DSF et en sélectionnant les pics se rapprochant des fréquences de résonnance des métabolites recherchés. Cette adaptation reste sensible aux éventuels décalages du spectre.

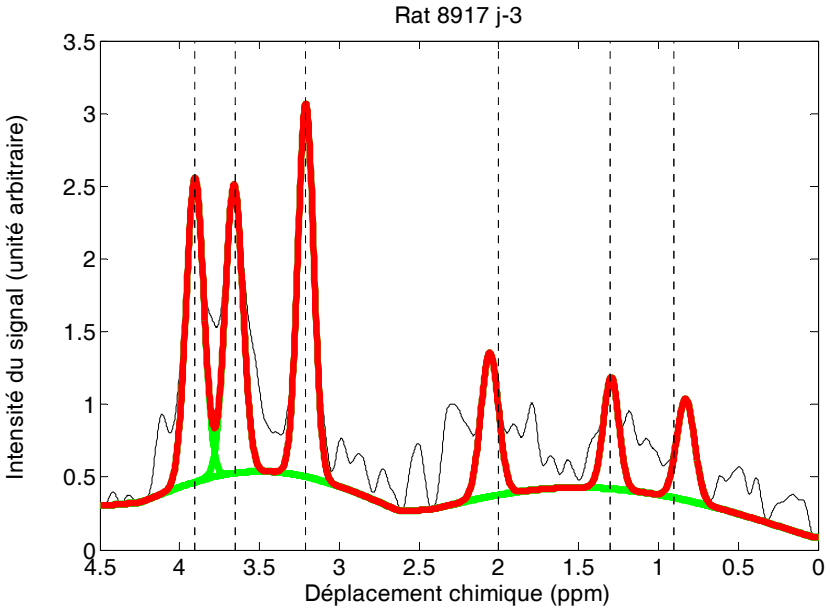


Figure 53 : Exemple de détection par la méthode DSF sur le spectre d'une tumeur de rat.
Cette tumeur était issue d'un modèle prostatique (PAC-120) implantée en sous cutanée. L'examen à eu lieu trois jour avant castration. L'objectif était de quantifier les métabolites résonnants à 0.9, 1.3, 1.99, 3.21, 3.6 et 3.9 ppm.

Chapitre 4 : Classification

Après une brève présentation de quelques notions utilisées en classification, quelques méthodes vont être présentées. Ce ne sera pas une liste exhaustive de toutes les méthodes de classification existantes : seules celles qui nous ont semblé les plus adaptées à notre problème seront présentées. Une description plus complète pourra être trouvée en se référant aux livres de Bishop et Duda & Hart [23, 76].

La classification consiste à attribuer à un objet une classe le caractérisant : dans le monde industriel, les classes peuvent être "pièce avec défaut" versus "pièce sans défaut" par exemple. Dans le monde médical, les classes sont généralement "échantillon sain" versus "échantillon pathologique" mais on peut aussi avoir des classes non binaires en distinguant les défauts ou les pathologies. L'objet à classer est représenté par des attributs qui ont chacun une valeur : l'ensemble de ces valeurs forme un vecteur dans l'espace de description. Ainsi nous pouvons parler soit de l'objet (monde réel) soit de sa représentation mathématique x ($\mathrm{x} \in \mathbb{R}^d$, d est le nombre d'attributs de description de l'objet, c'est la dimension de l'espace de description). Dans un monde idéal, tous les objets de la même classe présenteraient des valeurs d'attributs identiques, les seules variations seraient entre attributs d'objet de classe différente, permettant ainsi de distinguer rapidement la classe de chaque objet (variation inter-classe). Malheureusement, les attributs définissant les objets subissent des variations à l'intérieur même de la classe. Ces variations intra-classe peuvent être dues d'une part à la variation naturelle des objets : aucun être humain n'a les mêmes caractéristiques, le même métabolisme … Il existe toujours une différence même entre deux jumeaux alors qu'ils appartiennent tous à la classe "être humain". D'autre part, ces variations peuvent également être dues aux imprécisions de mesure. Pour ces deux raisons, deux représentations ponctuelles dans l'espace de description de deux objets de la même classe peuvent être totalement différentes. Ainsi l'objectif des méthodes de classification est de réaliser une distinction entre chaque nuage pour leur affecter une classe.

Il existe ainsi deux grandes familles de méthodes de classification :

- Classification supervisée : Dans ce cas, un expert est capable pour chaque objet qui lui est présenté de définir sa classe. Le classifieur doit alors être capable de prédire la classe d'appartenance ω_j d'un objet de vecteur de description x_i avec

un minimum d'erreur. Pour cela, il estimera la probabilité d'appartenance à la classe ω_i du prototype x_i notée $P(\omega_i \mid x_i)$. On considère une machine devant classer L objets d'un ensemble D. Chaque échantillon est donc représenté par un doublet $\{x_i, y_i\}$ où x_i est le vecteur représentant le i^{ème} objet dans l'espace de description (typiquement $\mathbb{R}^d$) et y_i est l'étiquette attribuée à la classe d'appartenance de l'objet (i.e. si $x_i \in \omega_j$ alors $j = y_i$). On cherche donc à définir la loi de probabilité $P(x, y)$ qui pour tout x donné retourne la probabilité qu'il appartienne à la classe d'étiquette y.

- Classification non supervisée : Dans ce cas il n'existe pas d'expert capable de définir la classe de chaque objet. Les méthodes de classification ne peuvent pas se baser sur l'estimation de probabilités puisqu'elles ne sont ni définissables ni estimables dans la mesure où on ne peut connaitre l'appartenance d'un objet à une classe. Ces méthodes essayent donc de découper la représentation ponctuelle dans l'espace de description en parties homogènes dites cluster par le biais d'agrégations. La principale difficulté de ces méthodes réside dans la détermination du critère d'homogénéité permettant de former les clusters.

Que la classification soit supervisée ou non, les méthodes nécessitent un grand ensemble de données pour être statistiquement valables. Cet ensemble de données doit être divisé en deux parties appelées :

- Base d'apprentissage : Cet ensemble d'objets x_i noté $\mathcal{B}$ permettra soit de déterminer le modèle dans le cas d'une classification supervisée, soit les clusters pour une classification non supervisée. $\mathcal{B}$ doit être le plus représentatif possible de l'ensemble de la population finale d'objets qui sera observée, aussi bien d'un point de vue quantitatif en respectant au mieux les probabilités de chaque classe que d'un point de vue qualitatif (ne pas se restreindre aux meilleures mesures par exemple).
- Base de test : C'est un ensemble noté T d'éléments à partir duquel sera menée la phase de validation. Sa construction doit être aussi rigoureuse que celle de l'ensemble$\mathcal{B}$, car c'est à partir de T que les performances de l'algorithme seront mesurées. Ainsi tout biais introduit dans T pourra orienter l'analyste vers un algorithme de classification non performant.

Pour évaluer la performance de chaque méthode, on se base sur la mesure de taux d'erreur. Ceux-ci peuvent être définis différemment selon que l'on prenne en compte deux ou plusieurs classes. Dans le cas classique d'échantillons médicaux, deux cas se présentent : soit l'échantillon est sain, soit il est pathologique (cas le plus simple, ne tenant pas compte des phases de transition ou des "mélanges"). Dès lors on distingue deux types d'erreur : les éléments réellement sains qui se retrouvent classés comme pathologiques, on parle alors de "Faux Positif" (FP) et les éléments pathologiques classés comme sains : les "Faux Négatif" (FN). On peut alors affecter un poids différent à ces deux taux d'erreur. Parfois le choix des poids est aisé : si on se place dans le cas d'une production de pièces industrielles pour des moteurs d'avion, pour privilégier la vie des passagers, il vaut mieux rejeter beaucoup de pièces même si elles ne sont pas défectueuses, par contre pour certaines industries où la qualité est moins importante, on peut alors privilégier la production en négligeant les FP et en privilégiant les FN. Dans le cas du cancer, la question est plus philosophique : si on a beaucoup de FP et peu de FN, on va faire subir des traitements très lourds (chirurgie, radiothérapie ou chimiothérapie) à des personnes qui n'en avaient pas besoin mais peu de cancers ne seront pas détectés. Inversement si on a beaucoup de FN et peu de FP, on traitera moins de personnes qui n'en ont pas besoin mais on laissera partir plus de patients avec des cancers sans les traiter.

Dans le cas de plus de 2 classes, cette terminologie ne s'applique plus et l'on parle d'erreur globale. Elle s'évalue en calculant le rapport entre le nombre d'erreurs de classification total (i.e. le nombre total d'objets mals classés) et le nombre total d'éléments pris en compte dans la phase de test (i.e. $Card(T)$).

4.1 Les machines à vecteurs de support

En 1995,Vladimir Vapnik a proposé une nouvelle technique de classification : les machines d'apprentissage à vecteurs de support (Support Vector Machine : SVM)[29]. Depuis cette époque, les SVM ont été largement utilisées pour la reconnaissance de forme, la régression et l'estimation de densité. Le principe de base de cette technique est de trouver l'hyperplan optimal séparant deux classes dans l'espace de description. Cet hyperplan est celui qui maximise la distance entre les deux classes qu'on appelle la marge. C'est pourquoi, dans la littérature francophone, on trouve parfois le terme de "Séparateur à

Vaste Marge". Nous allons ici en rappeler les principes élémentaires basés sur la notion de minimisation du risque empirique.

Rappelons que dans tout problème par apprentissage, nous cherchons à classer correctement les futures observations (ici les attributs mesurés sur les voxels, attributs issus de la spectroscopie) grâce à la connaissance apprise (les expertises des médecins).

On cherche à définir la loi de probabilité P(x, y) qui, pour tout x (observation) donné, donne la probabilité qu'il appartienne à la classe d'étiquette y. Pour cela, il faut définir une fonction f(x,α) telle que x→ f(x,α). α est un ensemble de paramètres dépendant du problème et permettant une réponse optimale.

L'erreur de classification donnée par une telle machine est appelée le "risque attendu" ou simplement le "risque" et est définie par :

$$R(\alpha) = \frac{1}{2}\int \left| y - f(\mathrm{x},\alpha) \right| dP(\mathrm{x},\alpha) \tag{8}$$

On remarque que si la densité de probabilité $p(\mathrm{x},\alpha)$ existe, alors la quantité $dP(\mathrm{x},\alpha)$ peut s'écrire $dP(\mathrm{x},\alpha) = p(\mathrm{x},\alpha)d\mathrm{x}dy$.

On peut alors définir la quantité qui nous intéresse, appelée le "risque empirique" et qui est la moyenne des erreurs de mauvaise classification :

$$R_{emp}(\alpha) = \frac{1}{2p}\sum_{i=1}^{n} \left| y_i - f(\mathrm{x}_i,\alpha) \right| \tag{9}$$

où n est le nombre d'échantillons. Nous remarquons qu'elle ne fait pas intervenir de distribution de probabilité et que $R_{emp}(\alpha)$ est un scalaire donné pour un choix particulier de α et pour un ensemble de couples $\{\mathrm{x}_i, y_i\}$ donnés.

Après cette définition du risque, on peut définir les SVM à proprement dit, cependant il existe différents cas possibles que nous allons décrire ci-dessous.

4.1.1 Cas simple : SVM linéaire à deux classes.

On considère une base d'apprentissage B de n objets à deux classes linéairement séparables. Chaque objet est alors décrit par le couple $\{\mathrm{x}_i, y_i\}, \mathrm{x}_i \in \mathbb{R}^d, y_i \in \{-1,1\}$ où x_i est le vecteur décrivant chaque objet dans l'espace de description de dimension d et y_i l'étiquette de la classe de cet objet. (Figure 54).

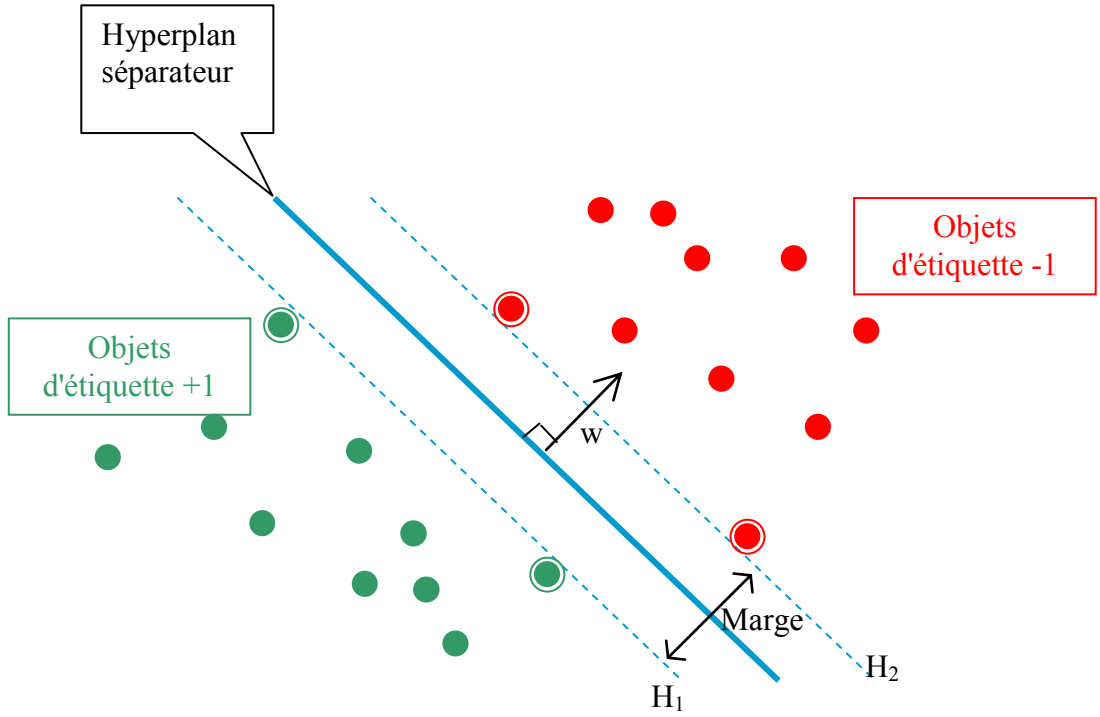


Figure 54 : Principe des SVM : trouver l'hyperplan séparant les groupes d'éléments en maximisant la marge.

On suppose qu'il existe un hyperplan séparant les points de chaque classe : l'équation d'un tel hyperplan s'écrit alors :

$$\langle w, \mathrm{x} \rangle + b = 0 \tag{10}$$

où w est une normale à l'hyperplan.

Il existe une infinité d'hyperplans respectant cette équation, il faut trouver le meilleur. La notion de meilleur hyperplan est introduite en considérant deux hyperplans parallèles (H_1 et H_2) passant chacun par les points les plus proches de la frontière des classes qu'ils limitent et n'ont aucun point de la base d'apprentissage B entre eux (Figure 55). Leurs équations respectives sont :

$$H_1 : \langle w, \mathrm{x} \rangle + b = 1 \qquad H_2 : \langle w, \mathrm{x} \rangle + b = -1$$

Pour garantir qu'il n'y ait aucun point de B entre eux, tout point de B doit satisfaire l'inégalité suivante :

$$\forall i \in [1..n], \quad y_i \left(\langle w, \mathrm{x}_i \rangle + b \right) \geq 1 \tag{11}$$

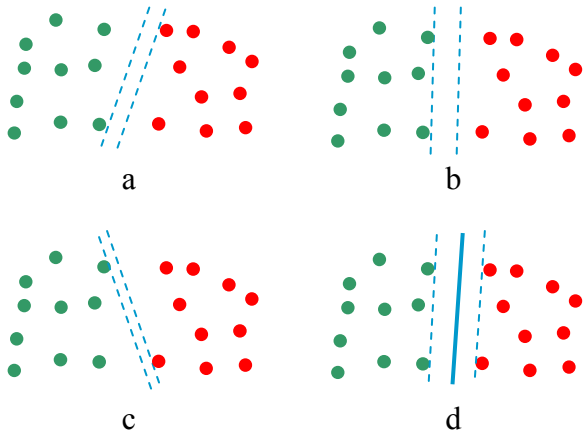


Figure 55 : Exemple d'hyperplans séparateurs.
Pour définir l'hyperplan séparateur optimal, il faut prendre deux hyperplans parallèles sans élément de la base d'apprentissage entre eux et maximiser la distance entre eux. L'hyperplan optimal est alors au milieu de ces deux hyperplans (d). a, b, et c ne maximisent pas la marge.

La distance entre ces deux hyperplans (la marge) est $\frac{2}{\|w\|}$ donc pour maximiser sa valeur, il faut minimiser $\|w\|$. Ceci est un problème d'optimisation quadratique. Le meilleur hyperplan séparateur est alors celui qui est à égale distance de H_1 et H_2. Pour résoudre ce problème, on utilise la théorie de Lagrange pour reformuler le problème de minimisation sans la contrainte d'inégalité. On obtient alors la fonction de Lagrange définie par :

$$L(w,b,\alpha) = \tfrac{1}{2}\|w\| - \sum_{i=1}^{n} \alpha_i y_i \left(\langle w_i, \mathrm{x}_i \rangle + b\right) + \sum_{i=1}^{n} \alpha_i, \quad \alpha_i \geq 0 \tag{12}$$

où les α_i sont appelés les coefficients de Lagrange. Cette fonction est alors minimisée par rapport à w et b.

En prenant la dérivée partielle de $L(w,b,\alpha)$ par rapport à w et b, on obtient :

$$\frac{\partial L(w,b,\alpha)}{\partial w} = w - \sum_{i=1}^{n} \alpha_i y_i \mathrm{x}_i \tag{13}$$

et

$$\frac{\partial L(w,b,\alpha)}{\partial b} = -\sum_{i=1}^{n} \alpha_i y_i \tag{14}$$

En annulant les deux dérivées partielles, on obtient le système suivant :

$$\sum_{i=1}^{n} \alpha_i y_i \mathrm{x}_i = w \tag{15}$$

et

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \tag{16}$$

On peut alors obtenir un problème dual en prenant en considération les éléments de la base d'apprentissage B par couple x_i et y_i. Ce problème s'écrit alors :

$$L(w,b,\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \langle \mathrm{x}_i, \mathrm{x}_j \rangle \tag{17}$$

$L(w,b,\alpha)$ est alors maximisé selon les α_i en respectant les contraintes $\sum_{i=1}^{n} \alpha_i y_i = 0$ et $\alpha_i \geq 0$.

Une fois l'optimisation effectuée, on peut aisément déduire b.

La décision de classification est alors obtenue pour tout x n'appartenant pas à l'ensemble d'apprentissage par :

$$y = signe(\langle w, \mathrm{x} \rangle + b)$$

En remplaçant par w sa valeur de l'équation(15), on obtient la fonction de décision suivante :

$$f(\mathrm{x}) = \sum_{i=1}^{n} \alpha_i y_i \langle \mathrm{x}_i, \mathrm{x} \rangle + b \tag{18}$$

Chaque coefficient multiplicateur de Lagrange α_i correspond à un élément x_i de la base d'apprentissage B. Les x_i qui ont contribué à la maximisation de la marge sont associés à des α_i non nuls et sont donc appelés vecteurs supports. Les autres vecteurs de la base d'apprentissage ne participent donc pas à la décision finale et l'équation (18) devient :

$$f(\mathrm{x}) = \sum_{i \in SV} \alpha_i y_i \langle \mathrm{x}_i, \mathrm{x} \rangle + b \tag{19}$$

où SV est l'ensemble des indices des vecteurs de support (vecteurs pour lesquels les coefficients de Lagrange sont non nuls).

4.1.2 Cas non linéairement séparable

Le cas non linéairement séparable se présente lorsqu'il n'existe pas d'hyperplan d'équation linéaire séparant les données sans erreur (Figure 56).

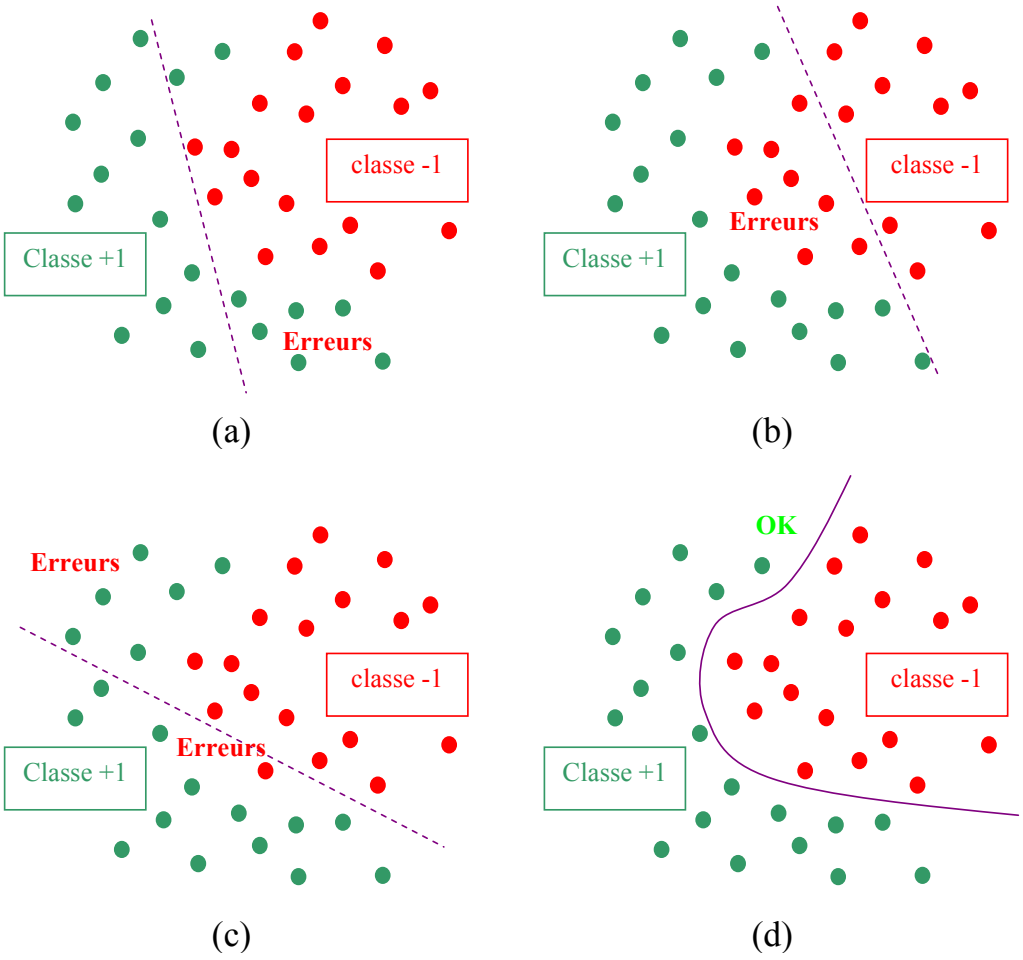


Figure 56 : Deux classes non linéairement séparables.
Il n'existe pas d'hyperplan séparant les deux classes sans erreur (a,b,c), seul un hyperplan d'équation non linéaire le permet (d).

Pour apporter une solution à ce problème, on introduit la notion de marge souple (soft margin), qui autorise un certain nombre d'éléments d'une classe à se trouver du mauvais coté de la frontière lors de l'apprentissage (Figure 57). Ces éléments sont appelés des "outliers".

On applique donc une marge d'erreur ξ_i à l'équation (11) pour obtenir :

$$\forall i \in [1..n], \quad y_i \left(\langle w, \mathrm{x}_i \rangle + b \right) \geq 1 - \xi_i \tag{20}$$

Pour maximiser la marge, il ne suffit plus de minimiser $\|w\|$ mais la fonction :

$$\|w\| + C \sum_{i=1}^{n} \xi_i$$

où C est le coût attribué à un échantillon mal classé.

De la même façon que dans le cas de données linéairement séparables, on détermine les vecteurs supports par résolution quadratique. Dans ce cas, la fonction de décision est identique à celle définie par l'équation(19).

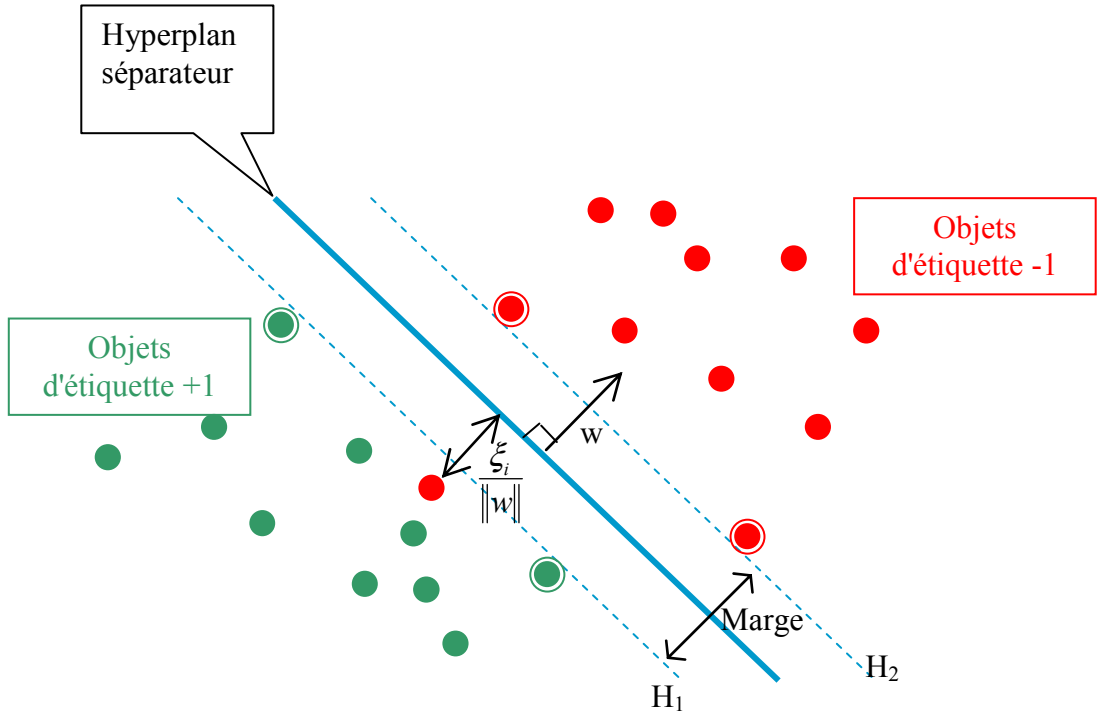


Figure 57 : Principe de "l'outlier".
Certains éléments sont "autorisés" à se trouver du mauvais coté de la frontière lors de la phase d'apprentissage.

La frontière définie est là aussi linéaire mais décalée en fonction du coût fixé. L'erreur due à ce type de séparation est souvent plus importante que l'erreur générée par une frontière non linéaire, c'est pourquoi une modélisation non linéaire a été intégrée, toujours en vue de minimiser le risque empirique. Cette méthode dite "astuce du noyau" (Kernel trick) a été introduite par Aizerman en 1964 [77].

Les équations de base font intervenir les données d'apprentissage (les échantillons et leur classe) sous la forme d'un produit scalaire. Il a été démontré qu'il est possible que les données initiales (de l'espace $\mathbb{R}^d$) soient linéairement séparables dans un espace euclidien $H = \mathbb{R}^e$ (avec e>d fini ou non), dans lequel on projette l'espace de description $\mathbb{R}^d$ à l'aide d'une fonction ϕ telle que :

$$\phi : \mathbb{R}^d \rightarrow \mathbb{R}^e, \phi(\mathrm{x}) = \tilde{\mathrm{x}} \qquad (21)$$

En effet, plus la dimension de l'espace de description est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les classes est élevée. En transformant l'espace d'entrée en un espace de redescription H de très grande dimension, éventuellement infinie, il devient donc possible d'envisager d'utiliser la méthode de résolution quadratique précédente même dans le cas de classes non linéairement séparables.

Si ces espaces sont munis d'un produit scalaire, noté (.), on peut définir sur $\mathbb{R}^d$ une fonction noyau telle que :

$$K : \mathbb{R}^d \times \mathbb{R}^e \to \mathbb{R}^e \times \mathbb{R}^e, K(\mathrm{x},\mathrm{y}) = \langle \phi(\mathrm{x}), \phi(\mathrm{y}) \rangle = \langle \tilde{\mathrm{x}}, \tilde{\mathrm{y}} \rangle \tag{22}$$

L'algorithme des SVM est alors le même que dans le cas linéaire : il suffit de remplacer tous les produits scalaires par une fonction noyau non linéaire. Avec cette astuce, les SVM peuvent définir le meilleur hyperplan séparateur linéaire dans le nouvel espace $\mathbb{R}^e$ en utilisant la fonction noyau appropriée.

En appliquant la technique à noyau à la fonction de Lagrange, on obtient la fonction suivante :

$$L(w,b,\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(\mathrm{x}_i, \mathrm{x}_j) \tag{23}$$

Et la fonction de décision devient donc :

$$y = signe\left(\sum_{i \in SV} \alpha_i y_i K(\mathrm{x}_i, \mathrm{x}) + b \right) \tag{24}$$

où les SV est l'ensemble des indices des vecteurs supports issus de la phase d'apprentissage.

Toute fonction satisfaisant au critère de Mercer peut être utilisée comme noyau. Ce critère est le suivant : soit k une fonction symétrique, on définit alors la matrice $\left(k(\mathrm{x}_i, \mathrm{x}_j) \right)_{i,j}$. Si cette matrice est définie positive, c'est-à-dire que pour chaque vecteur colonne χ de cette matrice $\chi^T M \chi > 0$, alors il existe un espace F et une fonction ϕ tels que $k(\mathrm{x}, y) = \langle \phi(\mathrm{x}), \phi(y) \rangle$.

On peut ainsi définir notre propre noyau, cependant il existe plusieurs noyaux prédéfinis dont les plus employés sont :

- Le noyau polynomial (d'ordre p) : $K(\mathrm{x}, y) = \left(a\langle \mathrm{x}, y\rangle + c\right)^p$,
- Le noyau RBF : $K(\mathrm{x}, y) = \exp\left(-\gamma \|\mathrm{x} - y\|^2\right), \quad avec\ \gamma > 0$,
- Le noyau RBF gaussien : $K(\mathrm{x}, y) = \exp\left(-\frac{\|\mathrm{x} - y\|}{2\sigma^2}\right)$,
- Le noyau Tangente hyperbolique : $K(\mathrm{x}, y) = Tanh\left(k\langle \mathrm{x}, y\rangle - \theta\right)$.

Malgré l'apparente complexité de la méthode, sa mise en œuvre est relativement simple. La prédiction d'appartenance à une classe d'un élément s'effectue à partir d'un calcul d'une somme de produits dans laquelle intervient la fonction noyau. La complexité de la fonction de décision est donc essentiellement liée au nombre de vecteurs supports et à la complexité de K. Dans les différentes expériences décrites par la suite, nous avons utilisé le noyau RBF gaussien qui apporte un bon pouvoir de généralisation tout en ayant l'avantage de ne présenter qu'un paramètre à fixer. Ce paramètre permet de régler la qualité d'apprentissage de la méthode (Figure 58). Une valeur faible correspond à une forte capacité de généralisation, mais avec un pouvoir discriminant proche de celui du modèle linéaire (Figure 58 b). Inversement, en augmentant la valeur de ce paramètre, l'hyperplan séparateur se ressert sur une classe, en augmentant le nombre de vecteurs supports. Le système a alors tendance à faire de l'apprentissage "par cœur", c'est-à-dire qu'il classe sans erreur les points qui ont servi à l'apprentissage mais il commet beaucoup d'erreurs avec de nouveaux points. (Figure 58 k, l et m).

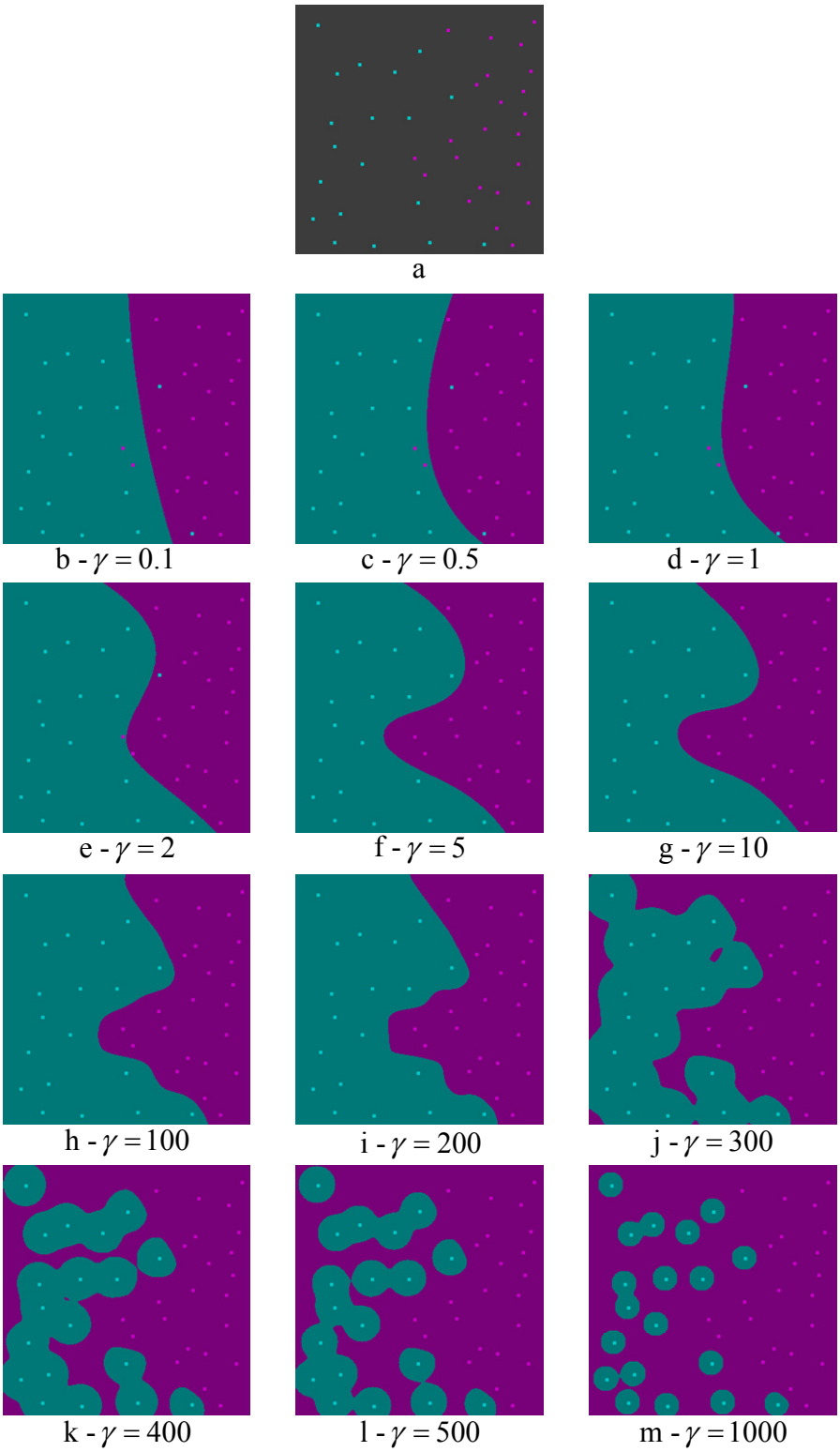


Figure 58 : Variation de l'hyperplan séparateur calculé par les SVM avec un noyau RBF Gaussien en fonction du paramètre γ .

Au delà d'une certaine valeur (ici 200 environ), le système ne fait plus de généralisation, mais de l'apprentissage "par cœur".

4.1.3 Cas multiclasse

Les méthodes précédentes ne considèrent que le cas où on est en présence que de 2 classes. Dans les autres cas (z classes, $z>2$), on utilise la technique du "un contre tous". On définit alors, n classifieurs chacun permettant de discriminer une classe parmi toutes les autres. On obtient donc n hyperplans séparateurs. Lors de la phase de décision, on affecte à l'élément à classer la classe pour laquelle on obtient la plus grande distance positive entre l'élément et l'hyperplan séparateur. Il existe une autre technique que nous n'emploierons pas : "un contre un". Il faut alors définir tous les classifieurs à deux classes possibles soit $\frac{z(z-1)}{2}$ classifieurs. Lors de la phase de décision, le nouveau point est classé par tous les classifieurs et le résultat est la classe la plus souvent retournée par chaque classifieur.

Il semblerait que la méthode "un contre tous " donne de meilleurs résultats que la méthode "un contre un" [78].

4.2 Les Réseaux de neurones

Les réseaux de neurones (ou réseaux neuronaux) sont une très grande famille de techniques d'intelligence artificielle dont une branche très importante permet de faire de la classification. Les premières notions de réseaux de neurones sont apparues dans les années 1950 lorsque les chercheurs ont voulu simuler le fonctionnement du cerveau.

C'est une technique toujours très employée en classification, notamment dans le milieu médical [79-89].

Les approches neuronales sont basées sur la combinaison d'éléments fonctionnels simples appelés "neurones", acceptant une ou plusieurs entrées et générant une seule sortie. Les neurones sont donc connectés entre eux (les connexions sont appelées synapses) reproduisant ainsi la structure d'un cerveau.

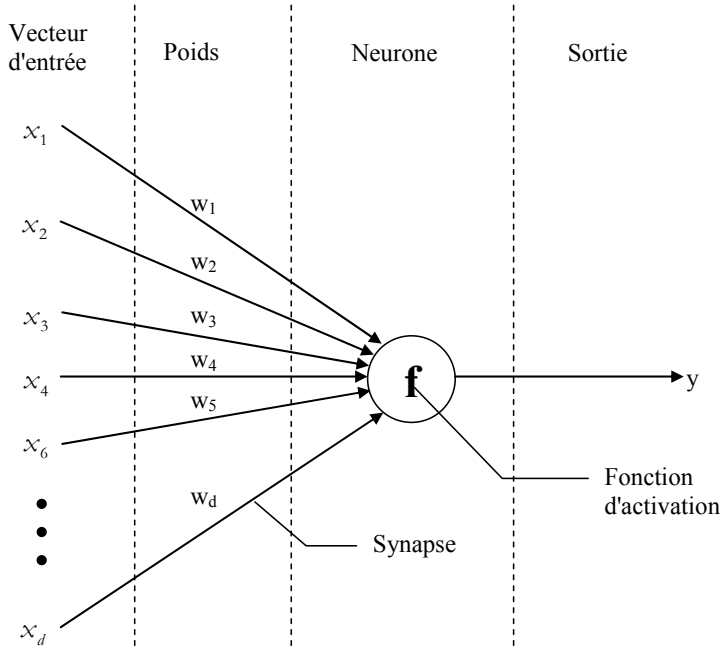


Figure 59 : Schéma d'un neurone selon la définition de McCulloch-Pitts.

Les entrées d'un neurone sont les composantes du vecteur $\mathrm{x} = (x_1,...,x_d)$ caractérisant l'échantillon (Figure 59). Le neurone calcule alors la somme pondérée de ces composantes :

$$v = \sum_{i=1}^{d} x_i w_i$$

où $w_1,...,w_d$ sont les poids synaptiques.

La valeur de sortie du neurone est alors calculée en appliquant la fonction d'activation f à v. Les fonctions sigmoïde $\left(\frac{1}{1+e^{-v}}\right)$, tangente hyperbolique $\left(th(v) = \frac{e^{v} - e^{-v}}{e^{v} + e^{-v}}\right)$ ou Heaviside $\left(\begin{cases} 0 & si\ v < 0 \\ 1 & si\ v \geq 0 \end{cases}\right)$ sont couramment employées pour la fonction f.

Cette structure a été définie dès 1943 par McCulloch et Pitts [90] et reste la plus courante. A partir de ce neurone, l'idée générale des réseaux est d'en interconnecter plusieurs. Cette idée mime le haut niveau d'interconnexion existant entre les neurones du cerveau qui est supposé expliquer les capacités de résistance, de mémorisation et de restitution du cerveau humain.

Différents paramètres définissant les interactions entre les neurones existent : ils permettent de spécialiser le réseau dans une activité :

- Les liaisons entre les neurones : est-ce que chaque neurone est relié à tous les autres ou n'est il relié qu'à un sous ensemble ?
- L'organisation des neurones : organiser par couche, hiérarchiquement …
- La dynamique du neurone : séquentielle (les neurones changent d'état à un instant t en fonction de leurs entrées) ou parallèle (tous les neurones changent d'état simultanément).

Nous allons donc voir maintenant une architecture mise au point en 1962 par Rosenblatt [91] : le Perceptron. C'est cette structure, améliorée depuis en Perceptron multi-couche (MLP : Multi Layer Perceptron), qui est la plus utilisée en classification.

Dans ces réseaux, les neurones sont organisés en couches. Le perceptron de base ne possède lui qu'une seule couche de neurones avec d entrées (de x_1 à x_d), m neurones (de N_1 à N_m) et donc m sorties (de y_1 à y_m) puisqu'un neurone n'a qu'une seule sortie (Figure 60). Toutes les entrées sont reliées à tous les neurones de la première couche, chacune de ces liaisons appelées synapse dispose donc d'un poids noté $w_{1,1}$ à $w_{d,m}$. Chaque couche attend le résultat de la couche précédente pour travailler (feed forward). La phase d'apprentissage se fait par rétro-propagation (back-propagation), algorithme développé par Rumelhart et McClelland en 1986 [92].

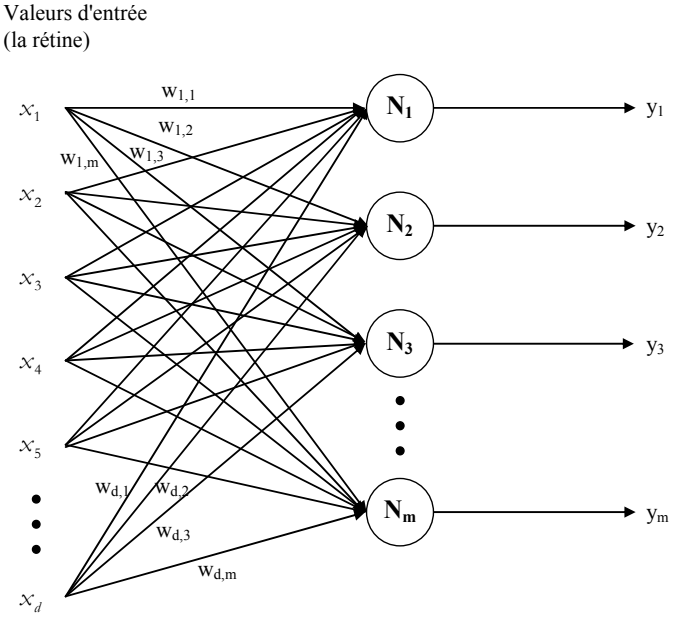


Figure 60 : Schéma du perceptron.

Dans le cas d'un perceptron mono-couche, le principe de la rétro-propagation est relativement simple : il consiste à comparer la sortie calculée du réseau avec la sortie attendue. Une erreur est alors calculée en faisant la somme des carrés des différences entre sortie calculée et valeur attendue. Le but de l'algorithme consiste donc à minimiser cette erreur en faisant varier les poids synaptiques.

Notons x_j^i la j-ème composante du i-ème vecteur d'entrée, y_j^i la j-ème valeur du ième vecteur de sortie calculé et t_j^i la ième valeur du j ème vecteur de sortie désiré, l'erreur est alors donnée par :

$$E = \sum_i \sum_j \left(y_j^i - t_j^i \right)^2$$

Le poids à l'itération k+1 est alors déterminé en fonction de sa valeur à l'itération k par :

$$w_{i,j}(k+1) = w_{i,j}(k) - \varphi \frac{\partial E}{\partial w_{i,j}}$$

où φ est appelé la constante d'apprentissage.

Cet algorithme peut s'appliquer quelque soit le nombre de couches du réseau. Le processus de convergence peut être très lent.

En 1988, Irie et Miyake [93] ont prouvé que le perceptron à 2 couches (Figure 61) avait toujours une solution, cependant rien ne prouve que l'algorithme arrive à converger vers cette solution.

De nombreux article ont été publiés concernant des tentatives d'amélioration de processus de convergence, mais il n'y a à l'heure actuelle toujours aucun algorithme à convergence sûre et optimale.

Dans le cas du perceptron monocouche, le calcul des poids synaptiques est :

$$w_{i,j}(k+1) = w_{i,j}(k) - \varphi(t_j - y_j)v_i$$

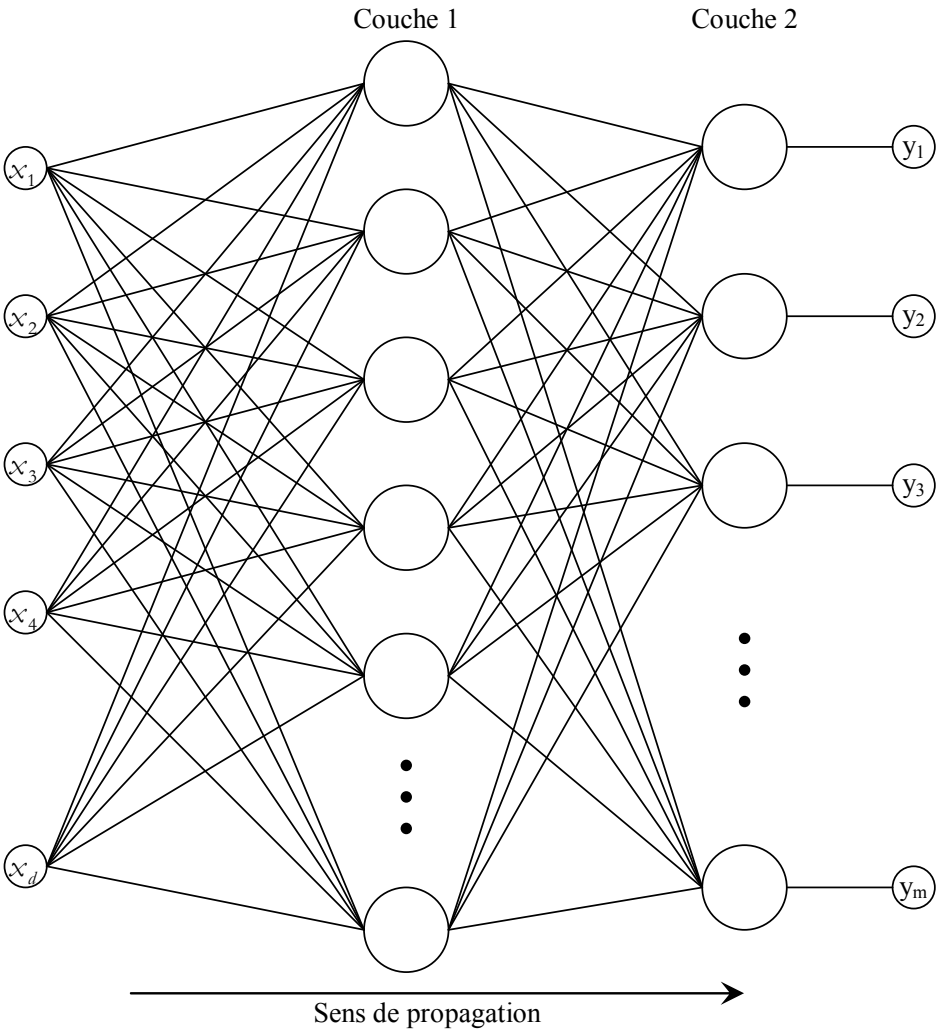


Figure 61 : Exemple de perceptron à 2 couches de neurones.

4.3 Validation croisée

En classification supervisée, il faut une base d'apprentissage et une base de test pour éviter l'apprentissage "par cœur" c'est-à-dire d'utiliser en test des échantillons qui ont servi à l'apprentissage. Or les performances des méthodes de classification dépendent généralement du nombre d'échantillons d'apprentissage : plus ce nombre est élevé, plus fiables seront les règles de classification. En même temps, il est nécessaire de conserver un nombre significatif d'échantillons de test pour que l'évaluation de ces performances soit significative. La technique de la validation croisée est fréquemment utilisée pour répondre à ces deux besoins : elle consiste à diviser l'ensemble de départ en un certain nombre de sous ensembles de taille égale, chaque sous ensemble étant alors utilisé comme base de test, alors

que l'union de tous les autres sous ensembles est utilisée comme base d'apprentissage (Figure 62). Soit *r* ce nombre de sous ensembles, on parle alors de validation croisée d'ordre *r* (en anglais on parle de *r* folds cross validation – *r*-CV). Naturellement *r* est supérieur ou égal à 2.

Soit D notre ensemble de données de *n* échantillons (dont on connaît la classe), chacun comportant *d* attributs. L'ensemble D est alors découpé en *r* sous-ensembles disjoints de taille identique $(E\left(\frac{n}{r}\right)\pm 1)$, la représentation statistique de chaque classe étant préservée par rapport à celle de l'ensemble D.

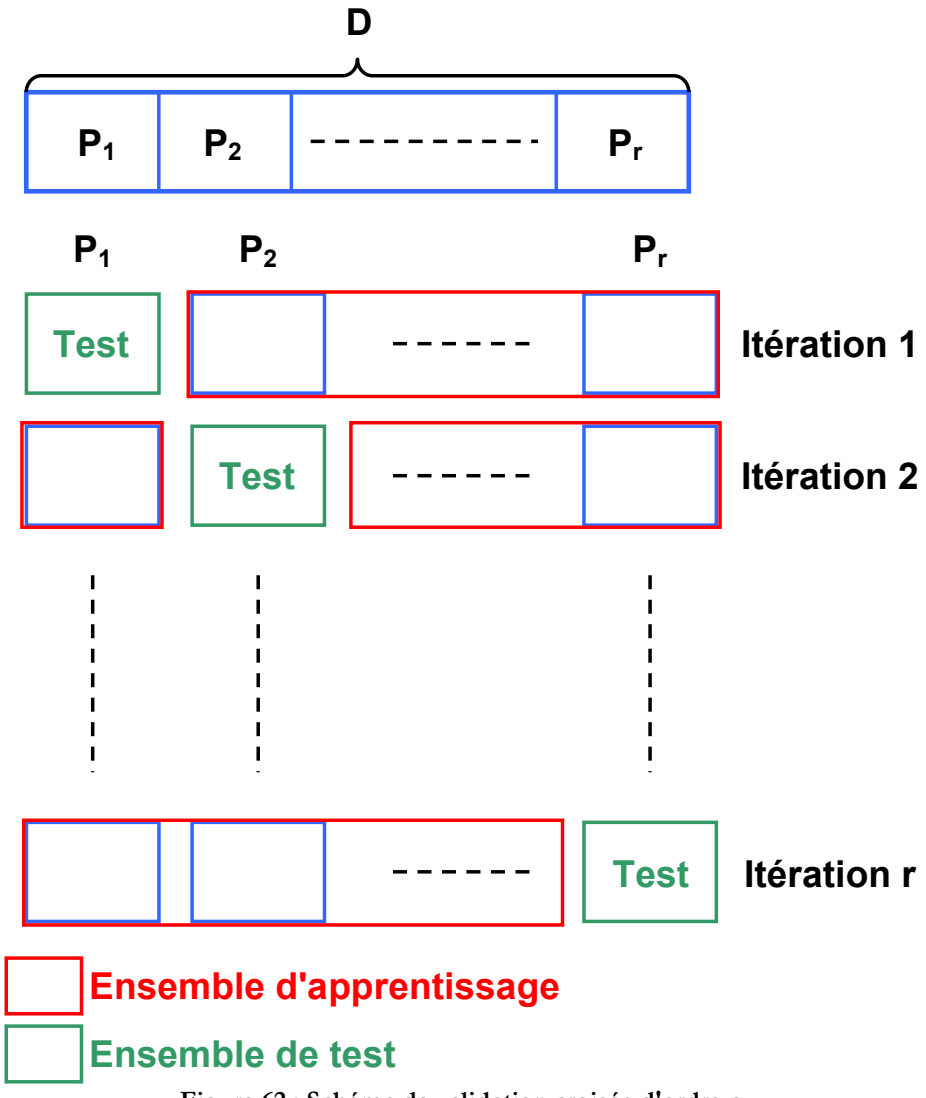


Figure 62 : Schéma de validation croisée d'ordre c.

A chaque itération, le nombre d'erreurs de classification réalisées est comptabilisé et le taux d'erreur de classification total est obtenu en divisant ce nombre d'erreurs par le nombre d'échantillons *n*.

Etant donné que la classe de chaque échantillon est connue, nous pouvons également produire la matrice de confusion qui recense pour chaque classe réelle, le nombre d'échantillons que le classifieur a classé dans chaque classe ce qui permet de détecter les éventuelles confusions que pourrait faire le système. Si c est le nombre de classes de D, la matrice de confusion est une matrice de taille $c \times c$ où la cellule à l'intersection de la ligne i et de la colonne j est le nombre d'éléments de classe réelle i classés par le système dans la classe j.

On appelle "leave one out" le cas particulier pour lequel le nombre de sous ensembles de la validation croisée est égal au nombre de mesures de l'ensemble (r=n). L'apprentissage se fait alors sur la base complète sauf 1 échantillon et le test est effectué sur l'élément restant. Cela permet, dans le cas d'un faible ensemble de données, de maximiser l'ensemble d'apprentissage mais présente l'inconvénient de se rapprocher de l'apprentissage "par-cœur". La difficulté de cette méthode est de trouver la bonne valeur de d en fonction de la taille de l'ensemble de départ et de la complexité du problème : plus le problème est complexe, plus le système a besoin d'exemples pour apprendre.

- Si r est petit, l'ensemble d'apprentissage sera trop faible et le système de classification ne généralisera pas assez, entrainant une augmentation du taux d'erreur. Par contre, il y a moins de phase d'apprentissage à réaliser et l'obtention des résultats est plus rapide. Cette solution est intéressante quand on dispose de beaucoup de mesures.
- Si r est grand, l'ensemble de test sera plus faible par rapport à l'ensemble d'apprentissage et on risque le phénomène de sur-apprentissage (cas extrême : le "leave one out").

Il est courant dans la littérature de trouver une validation croisée d'ordre 10, cependant comme nous disposons d'un ensemble de données relativement important, nous nous sommes contentés d'une validation croisée d'ordre 5. Quelques tests en VC-10 nous ont permis de constater que le taux d'erreur n'était pas amélioré par rapport à la VC-5.

4.4 Résultats

Une première série d'expériences a été menée pour essayer de trouver la meilleure combinaison entre les prétraitements appliqués aux spectres et la méthode de classification.

Nous nous sommes restreint, dans un premiers temps à la plage 2-4ppm car c'est celle contenant la choline et le citrate qui est régulièrement utilisée dans la littérature et que

cette partie du travail consiste à trouver le meilleur couple prétraitements/classifieur. Dans la partie suivante (Chapitre 1), nous utiliserons le spectre complet pour chercher de nouveaux biomarqueurs et les résultats seront validés par la technique qui aura donné les meilleurs résultats dans cette partie.

Nous avons donc évalué les performances en classification des SVM et d'un Perceptron multicouche sur 6 ensembles d'attributs :

- La partie réelle du spectre complet sur 1024 points avec une correction de phase automatique réalisée selon l'algorithme ACME.
- La partie réelle du spectre corrigée par ACME mais limitée à la plage 2,0-4,0 ppm soit 192 points.
- La partie réelle du spectre corrigée par ACME, limitée à la portion 2,0-4,0 ppm et avec la ligne de base corrigée avec CLB.
- Le module du spectre limité à la plage 2,0-4,0 ppm.
- Les concentrations en choline, polyamine, créatine et citrate fournies par LCModel.
- Les paramètres (hauteur, largeur à mi-hauteur et position dans le spectre) des gaussiennes détectées avec l'algorithme DM.

Pour les SVM, nous avons utilisé la librairie LIBSVM[94] implantée dans un logiciel développé au Le2i. Ce logiciel réalise le tirage aléatoire précédant la validation croisée et donne les taux d'erreur ainsi que les matrices de confusion.

Les résultats des perceptrons multicouches ont été obtenus à l'aide du logiciel commercial NeuroSolutions© de la société NeuroDimension Incorporated mis au point par Curt Lefebvre et Jose Principe. Il faut lui fournir en entrée l'ensemble d'apprentissage et l'ensemble de test et l'on récupère en sortie son estimation pour l'ensemble de test. La validation croisée doit alors être faite manuellement.

Dans tous les cas, les spectres ont au préalable été normalisés selon le volume d'acquisition, filtrés à l'aide d'un léger filtre gaussien pour atténuer le bruit et enfin normalisés par la norme L2. Dans chaque cas, nous avons fait varier le paramètre γ du noyau RBF entre 10^{-5} et 10^{4} da façon logarithmique (valeurs de γ testées : 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 5, 10, 50, 10^{2}, 5.10^{2}, 10^{3} et 10^{4}). Nous avons utilisé le noyau RBF car il est réputé ([26, 28, 80, 95]) pour fournir de bons résultats, posséder une bonne capacité de généralisation et être simple d'emploi du fait qu'il ne dépende que d'un seul paramètre (γ). Des tests préliminaires nous ont conforté dans notre choix : en le comparant à un noyau

polynomial, un noyau sigmoïde et aux SVM linéaires, il nous a donné les meilleurs taux de classification. A chaque fois le taux d'erreur a été mesuré à partir des résultats d'une validation croisée d'ordre 5.

Dans un premier temps, nous avons évalué la capacité du système de classification à séparer les voxels de la zone saine, des voxels de la zone tumorale. Dans un deuxième temps les spectres réputés inclassables par l'expert on été ajoutés en entrée du système pour évaluer son aptitude à les séparer des autres spectres.

Un premier tableau présente les résultats de la classification en utilisant uniquement deux classes (Tableau 4). Un second tableau présente les résultats avec 3 classes (Tableau 5).

La répartition des classes dans l'ensemble de données est la suivante :

- 286 spectres inclassables (26,9%)
- 636 spectres sains (59,9% avec 3 classes et 82% avec 2 classes)
- 140 spectres pathologiques (13,2% avec 3 classes et 18% avec 2 classes)

On a donc une forte prépondérance des spectres sains : cette répartition est normale compte tenu de la population d'étude. En effet tous les patients de l'étude ont déjà eu un dépistage positif du cancer de la prostate par dosage du PSA, généralement vérifié par biopsies. Cependant les campagnes de dépistage réalisées en France ont permis d'avoir des dépistages très précoces, toutes les prostates vues en IRM sont donc porteuses d'une tumeur mais souvent à un stade très précoce de développement et sont donc petites. Ce qui explique que bien que l'étude soit basée sur une population de patients porteurs d'une tumeur, les spectres « pathologiques » sont en plus faible proportion que les autres. Nous avons donc considéré que les proportions étaient représentatives des probabilités *a priori* de chacune des classes.

Cette forte prépondérance de représentation de la classe saine incite les classifieurs lors de leur phase d'apprentissage à privilégier cette classe. Ainsi en classant tous les éléments dans cette classe, le classifieur ne fait pas plus de 18% d'erreur avec 2 classes et 40,1% d'erreur avec 3 classes.

Tableau 4 : Taux d'erreur de classification (%) des SVM avec uniquement les cas sains et pathologiques en entrée.

	SVM			Perceptron multicouche		
	Sains	Pathologiques	Erreur totale	Sains	Pathologiques	Erreur totale
Spectre complet	3,14	21,43	6,44	4,66	8,37	5,80
2-4 ppm + ACME	2,36	15,71	4,77	3,30	20,71	6,44
2-4 ppm + ACME + CLB	**1,89**	**16,43**	**4,51**	2,99	19,29	5,93
Module 2-4 ppm	1,26	20,71	4,77	0,8	85,7	3,23
Concentrations LCModel	1,73	37,14	8,12	1,6	42,9	8,97
DM	2,67	17,14	5,28	2,67	33,57	8,25

Tableau 5 : Taux d'erreur de classification(%) des SVM et des MLP avec les cas sains, pathologiques et indéterminés en entrée.

	SVM				Perceptron multicouche			
	Indéterminés	Sains	Pathologiques	Erreur totale	Indéterminés	Sains	Pathologiques	Erreur totale
Spectre complet	22,03	7,08	32,86	14,50	40,56	4,56	60,71	21,66
2-4 ppm + ACME	23,42	7,39	30,00	14,69	27,62	9,28	24,29	16,20
2-4 ppm + ACME + CLB	**17,83**	**5,66**	**25,00**	**11,49**	19.,93	9,12	27,86	14,50
Module 2-4 ppm	20,28	6,76	27,86	13,18	24,13	6,45	40,00	15,63
Concentrations LCModel	52,45	7,08	47,14	24,58	57,34	9,75	52,86	28,25
DM	29,72	10,06	28,57	17,79	48,25	9,43	53,57	25,71

La première série d'expériences avec seulement deux classes consistait à vérifier la séparabilité des spectres sains et pathologiques en l'absence des spectres inclassables. Quelle que soit la méthode utilisée, on ne peut faire mieux que 15% d'erreurs sur les spectres pathologiques. Si on considère qu'on acquiert dans une prostate environ 50 spectres (la moyenne de cette ensemble de données est de 1062/22 = 48,3). En respectant les proportions de l'ensemble de données, on en déduit que 9 sont potentiellement pathologiques parmi ces 50. Un taux d'erreur de 15% représente donc 1,35 spectre pathologique non détecté, ce qui permet d'avoir de 7 à 8 spectres pathologiques détectés et donc de poser le bon diagnostic. Même avec 20% d'erreurs il est toujours potentiellement possible d'en détecter 7. Donc même si le taux d'erreur sur les spectres pathologiques peut sembler élevé, il n'est pas rédhibitoire.

La seconde série d'expérience consistait à vérifier que les classifieurs étaient capables de classer n'importe quel spectre de la prostate sans vérifier sa qualité (il faut néanmoins qu'il soit situé dans la zone périphérique de la prostate).

Nous pouvons cependant remarquer que les prétraitements que sont la correction de phase et la correction de la ligne de base améliorent la classification alors que les méthodes d'extraction de paramètres que nous avons utilisé (LCModel et estimation par gaussienne) dégradent les performances par rapport à l'utilisation du spectre brut. La comparaison des résultats nous permet cependant de valider notre technique par rapport à LCModel : le taux d'erreur total est plus faible (5,28% au lieu de 8,12% avec 2 classes et 17,79% au lieu de 24,58% avec 3 classes) et surtout le taux d'erreur sur les spectres pathologiques est bien meilleur (37,14% contre 17,14% avec 2 classes et 47,14% contre 28,57% avec 3 classes). Cela valide donc notre choix de développer cette technique puisque nous avions remarqué que la majeure partie des erreurs de LCModel était sur la détection de la choline en l'absence de citrate.

Bien que les MLP soient une méthode de classification éprouvée et reconnue notamment dans la communauté médicale, les SVM donnent ici de meilleurs résultats. Dans le cas à deux classes et de l'utilisation du module du spectre, le taux d'erreur global pour les MLP est inférieur à celui des SVM, mais ceci au prix d'une erreur beaucoup trop élevée sur les spectres pathologiques (plus de 85% d'erreur).

Globalement, nous pouvons considérer que le meilleur résultat est obtenu en utilisant le spectre réduit à la partie 2 à 4 ppm, corrigé en phase et dont la ligne de base a été supprimée. L'erreur globale dans le cas à deux classes est alors de 4,51%, avec un taux d'erreur de 1,89% sur les spectres sains et de 16,43% sur les spectres pathologiques. Nous

pouvons également remarquer que l'effet de la suppression de la ligne de base ne se fait sentir de manière significative que dans le cas à 3 classes (l'erreur globale étant alors de 11,49%, ce qui représente un gain de plus de 1,5% par rapport à toutes les autres méthodes).

Chapitre 5 : Sélection de biomarqueurs

Le chapitre précédent nous a permis de montrer qu'il était possible de classer automatiquement les spectres SRM de la prostate avec un taux d'erreur compatible avec une utilisation clinique. Toutefois, nous avions pour cela présélectionné la plage 2-4 ppm des spectres en nous basant sur les études antérieures mais il reste à explorer d'autres plages afin de trouver l'optimum. De plus, un des buts de notre étude est de déterminer la pertinence de biomarqueurs connus et éventuellement d'en trouver de nouveaux. Nous avons donc d'une part mis en œuvre des méthodes de sélection d'attributs (ici l'attribut est la valeur du déplacement chimique exprimée en ppm) ou de plages d'attributs à l'aide de méthodes standards utilisées habituellement en classification. D'autre part, nous avons mis en œuvre notre propre méthode de recherche de plage optimum d'attributs, méthode adaptée au problème de l'analyse des spectres.

D'une manière générale, la sélection d'attributs, en classification, consiste à tenter de réduire la taille de l'espace de travail initial en sélectionnant les attributs les plus significatifs (minimisation du risque empirique). De nombreuses méthodes ont été développées pour remplir cette tâche : Jain [96] les a répertoriées. On peut résumer ses travaux par la figure suivante (Figure 63):

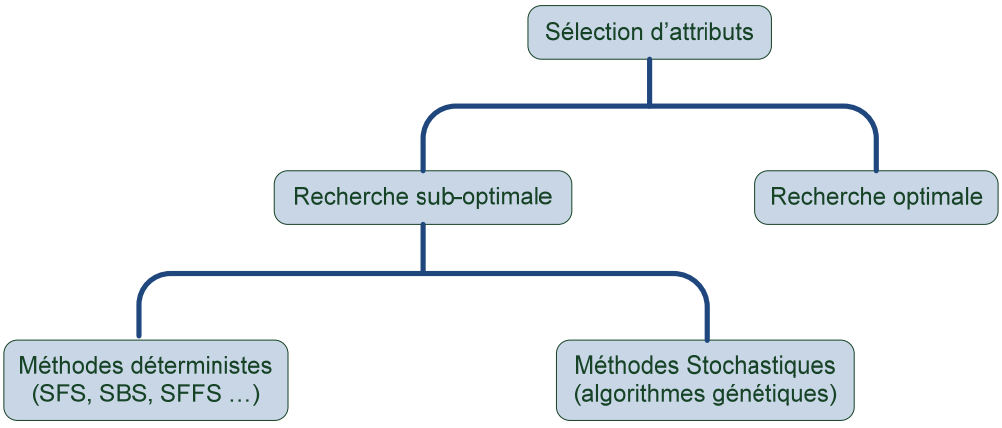


Figure 63 : Organisation des méthodes de sélection d'attributs.

La première distinction se fait entre méthode optimale et méthode sub-optimale : les méthodes optimales fournissent le meilleur ensemble d'attributs pour optimiser la

classification mais sont beaucoup plus gourmandes en temps de calcul, du fait qu'elles parcourent habituellement l'ensemble des possibilités ou combinaisons d'attributs. Or, le nombre de combinaisons possibles peut rapidement être rédhibitoire, les temps de calculs devenant incompatibles avec les contraintes d'une étude scientifique. Si nous prenons, par exemple, les données restreintes à la bande de fréquence 2-4 ppm, soit 192 attributs, et que l'on souhaite restreindre la dimension de l'espace final à 10, le nombre total de combinaisons est :

$$C_{192}^{10} = \frac{192!}{(192-10)!10!} = \frac{\prod_{i=183}^{192} i}{\prod_{i=1}^{10} i} > 1,478 \times 10^{16}$$

Si l'on considère que le temps de calcul pour une itération (détermination de l'erreur de classification avec 10 attributs) est de l'ordre de 1ms, explorer toutes ces combinaisons prendrait environ 500 000 ans… C'est pour cette raison que les méthodes sub-optimales sont généralement privilégiées quand le nombre d'attributs est important. L'inconvénient étant que rien ne nous assure d'avoir le meilleur sous-ensemble.

Parmi les méthodes sub-optimales, on peut également distinguer les méthodes basées sur les algorithmes de réduction de dimension, qui eux-mêmes peuvent déboucher sur une sélection des attributs d'origine, et les méthodes basées sur des parcours particuliers des combinaisons d'attributs.

Nous allons donc dans un premier temps explorer les possibilités offertes par l'Analyse en Composantes Principales, qui reste une des méthodes les plus efficaces en terme de réduction de dimension [97], et dans un deuxième temps, nous évaluerons l'efficacité de la méthode dite SFS (Sequential Forward Selection).

5.1 Analyse en composantes principales

5.1.1 Principe de l'ACP

L'ACP est une technique dite de réduction de dimensions. Le principe est de construire de nouvelles variables synthétiques qui sont des combinaisons linéaires des attributs initiaux de plus grande variance possible.

Ces nouvelles variables, appelées composantes principales, captureront la plus grande part de variabilité des données. Mathématiquement, cela se traduit dans un premier

temps par la construction d'une nouvelle variable $C_i = \alpha_{i,1}X_{i,1} + \alpha_{i,2}X_{i,2} + ... + \alpha_{i,n}X_{i,n}$ telle que la variance de C soit maximale.

$$Var(C_i) = Var\left(\sum_{j=1}^{n} \alpha_{i,j}X_{i,j}\right)$$

En notant $\mathrm{A}_i = (\alpha_{i,1}, \cdots, \alpha_{i,n})$ et V la matrice de taille n x n, dite de variance-covariance des variables $X_1, \cdots, X_n : V_{ij} = Cov(X_i, X_j)$, on a $Var(C_i) = \mathrm{A}_i' V \mathrm{A}_i$

Les poids doivent vérifier la relation : $\alpha_{i,j} : \sum_{j=1}^{n} {\alpha_{i,j}}^2 = 1$

En général, on s'intéresse aux variations par rapport à l'individu moyen. Ceci revient à centrer les variables $X_1, \cdots, X_n$. Alors la variable C_i est nécessairement aussi de moyenne nulle.

On itère ensuite la procédure en cherchant une nouvelle variable C_2 qui soit non corrélée avec la première variable C_1 et de variance maximale.

Le résultat de cette recherche est l'ensemble des vecteurs propres de la matrice de variance-covariance V. Le vecteur A_i de taille n est vecteur propre de la matrice V associé à la valeur propre λ_i si $VA_i = \lambda_i A_i$. Il existe n vecteurs propres $\mathrm{A}_1, \cdots, \mathrm{A}_n$ de la matrice V associés à des valeurs propres positives (ou nulles) rangées par ordre décroissant :

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_j \geq \cdots \geq \lambda_n \geq 0$$

Chaque valeur propre λ_i est la variance de la variable C_i. Les composantes des vecteurs propres A_j sont les facteurs de la combinaison linéaire qui permettent de passer des variables initiales X aux variables finales C de variances maximales.

5.1.2 Application de l'ACP à la réduction de dimension des spectres.

L'ACP utilisée en réduction de dimension ne réalise pas directement une sélection d'attributs, puisqu'ils interviennent tous dans la combinaison linéaire finale. Toutefois, les performances de classification sont parfois meilleures dans l'espace de dimension réduite, c'est pourquoi nous avons mis en œuvre l'ACP à ce niveau de l'étude.

Nous avons calculé la variance totale et nous avons sélectionné deux ensembles d'attributs parmi ceux de la nouvelle base. Le premier ensemble, soit 18 composantes,

correspond à 95% de la variance totale. Le deuxième ensemble, soit 91 composantes, correspond à 99% de la variance totale (Figure 64).

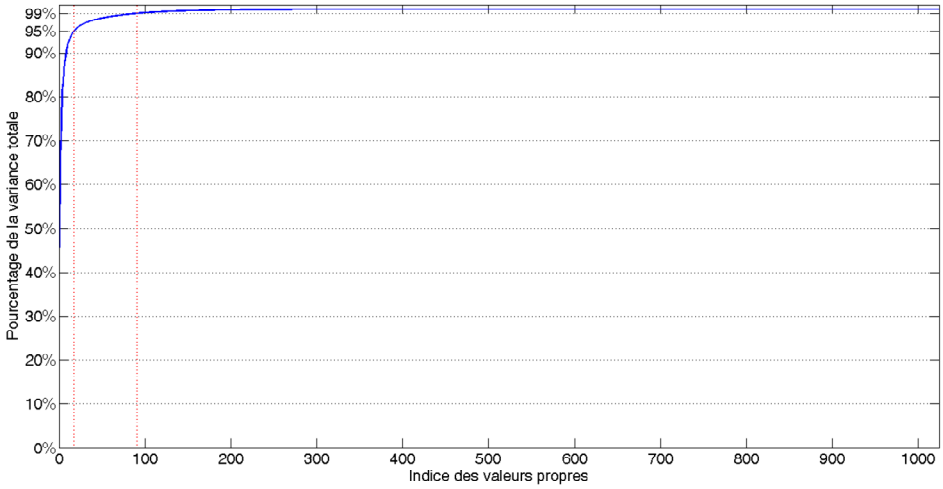


Figure 64 : Variance expliquée en fonction du nombre d'attributs sélectionnés.
95% de la variance est expliquée avec les 18 premiers attributs. On atteint 99% avec 91 attributs.

Nous avons donc projeté les 1062 spectres (rephasés, ligne de base corrigée et attributs normalisés) dans ces deux nouveaux référentiels et nous avons utilisé ces ensembles de données en entrée d'un classifieur SVM avec un noyau RBF gaussien dont nous avons fait varier le paramètre γ de 10^{-5} à 10^{8}.

Avec 18 attributs, on atteint, dans le meilleur des cas, un taux d'erreur totale de 23,07%. Avec 91 attributs, ce taux descend à 21,56% (Figure 65).

Ces résultats sont moins bons que ceux obtenus avec les SVM sur la plage de spectres habituellement étudiée en SRM de la prostate, qui étaient de 11,49 %.

La simple réduction de dimension basée sur la variance n'est donc pas adaptée à notre cas. Une explication peut venir du fait que l'ACP cherche à retenir le maximum de la variance des données. Or dans notre cas, l'utilisation d'une séquence de suppression de l'eau et des graisses génère de grandes variations du signal dans ces zones (suivant l'efficacité de celle-ci) qui sont donc prisent en compte par l'ACP alors que cela n'apporte pas d'information sur l'état sain ou pathologique du spectre.

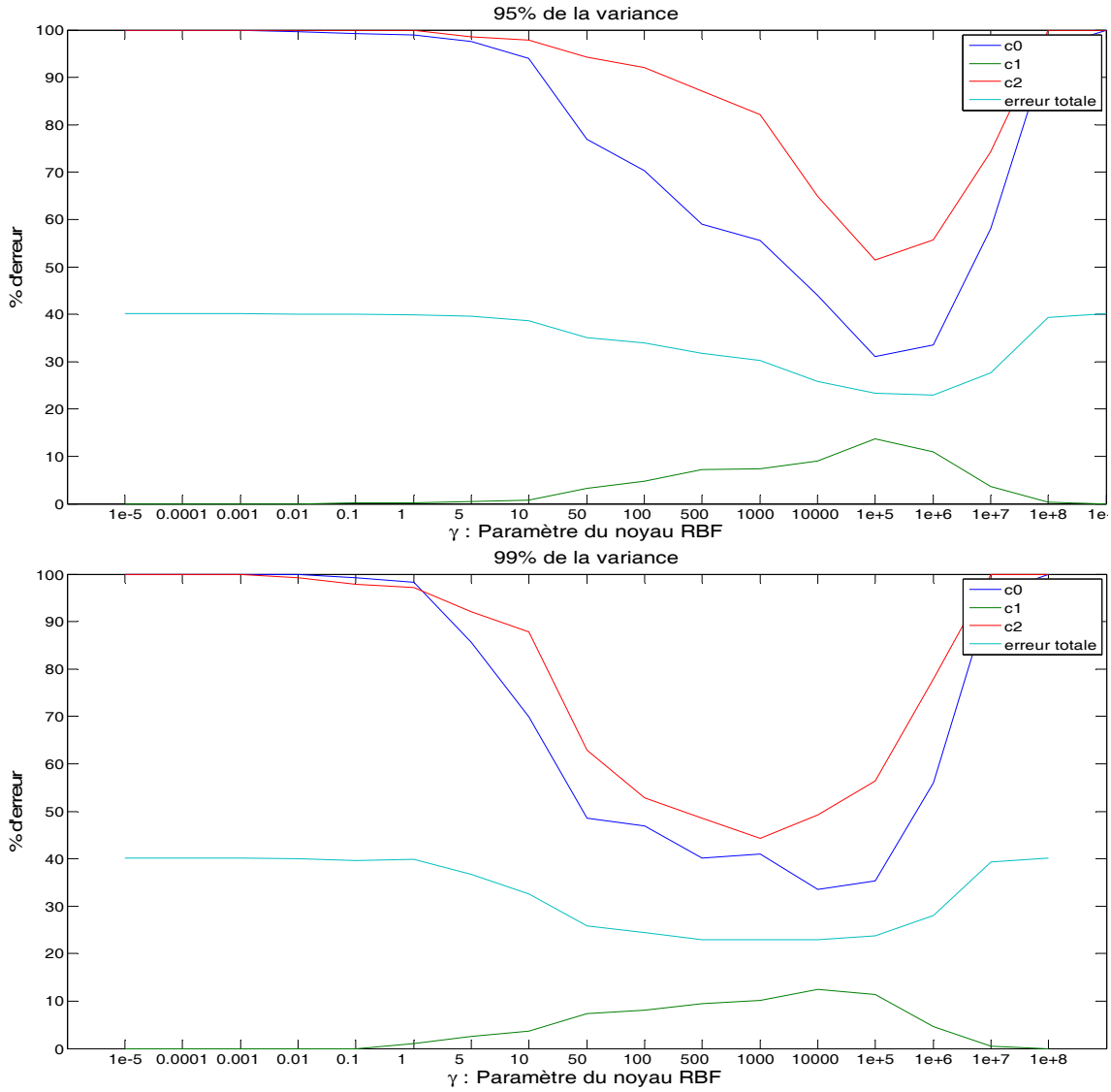


Figure 65 : Variations des taux d'erreur des SVM en validation croisée d'ordre 5 en fonction du paramètre du noyau RBF gaussien.

5.1.3 L'ACP pour la sélection d'attributs : application basique

Nous avons vu que l'ACP permet de maximiser la variance de l'ensemble de données en le projetant dans un nouvel espace où chaque dimension est définie par une combinaison linéaire des attributs de l'espace d'origine. Si on considère le premier vecteur propre C_1 associé à la valeur propre λ_1, celle-ci représente la variante associée à la variable C_1 et qui est la variable de variance maximum du nouvel ensemble de données.

$$C_1 = A_1 X_1 = \alpha_{1,1} X_{1,1} + \alpha_{1,2} X_{1,2} + ... + \alpha_{1,n} X_{1,n}$$

Les coefficients $(\alpha_{1,1},...,\alpha_{1,n})$ sont associés aux variables d'origine $X_1,\cdots,X_n$. Les variables correspondant au coefficient de plus grandes valeurs absolues et donc apportant

la plus grande participation à la valeur de C_1, peuvent être considérées comme participant à la variance de l'ensemble d'origine et peuvent être sélectionnées comme attributs pouvant avoir un sens pour la classification des objets. Cette méthode est utilisée de manière courante en analyse de données et en classification.

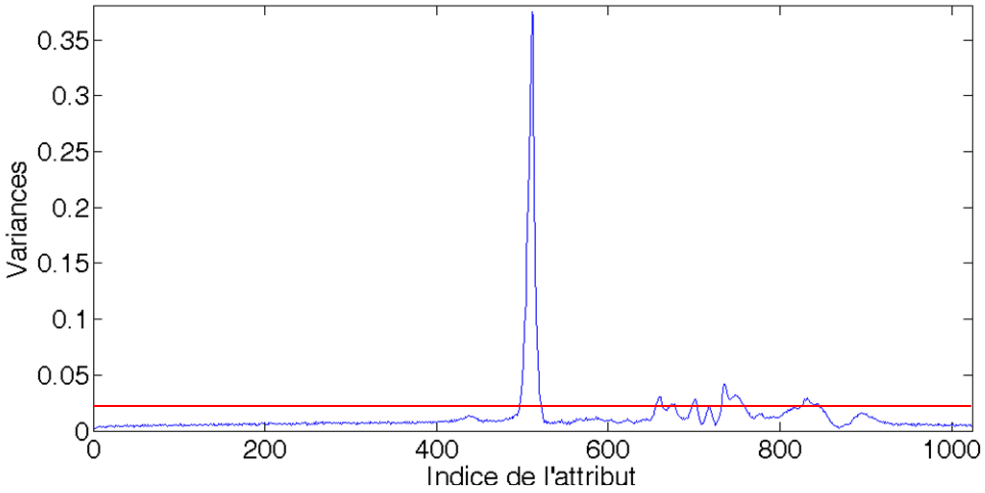


Figure 66 : Représentation graphique du coefficient α associé à chaque attribut dans la combinaison linéaire de la nouvelle variable C_1.
La ligne rouge représente le seuil de 0,02.

Si on représente la valeur de chaque coefficient α sur un graphique (Figure 66), on retrouve la forme globale d'un spectre moyen. Les plus fortes variations correspondent aux attributs liés à l'eau, on trouve ensuite ceux correspondant au citrate, à la choline, à la créatine et enfin aux lipides. On veut donc sélectionner les attributs correspondant aux plus grands coefficients cependant aucune méthode ne nous permet de savoir combien d'attributs doivent être sélectionnés.

Une première méthode consiste à classer les attributs par ordre des coefficients α décroissants (Figure 67). La net rupture de courbure à partir de 20 coefficients nous a incité à sélectionner les 20 attributs correspondant à ces coefficients.

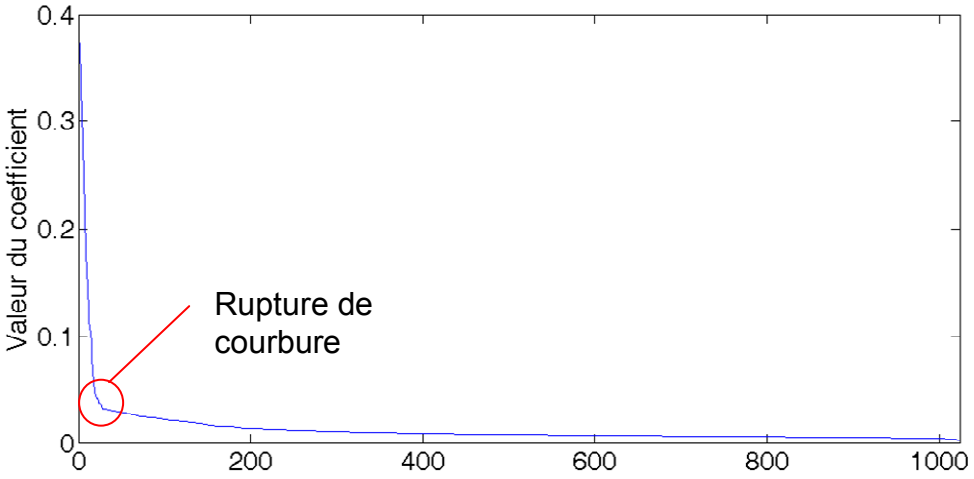


Figure 67 : Coefficients de la combinaison linéaire de la variable C_1 triés par ordre décroissant.

Toutefois, la validation croisée d'ordre 5 avec les SVM donne un taux d'erreur de 25,52%, ce qui n'est pas satisfaisant.

Une autre solution est donc de placer un seuil Th sur la valeur des coefficients pour conserver les variations significatives. Au vu de la Figure 68, nous avons arbitrairement choisi Th=0,02, ce qui permet de sélectionner 120 attributs, avec lesquels nous réalisons à nouveau une validation croisée d'ordre 5 avec les SVM à noyau RBF. Le taux d'erreur total obtenu est alors de 16,67% (Figure 68), toujours supérieur au meilleur des résultats obtenus précédemment.

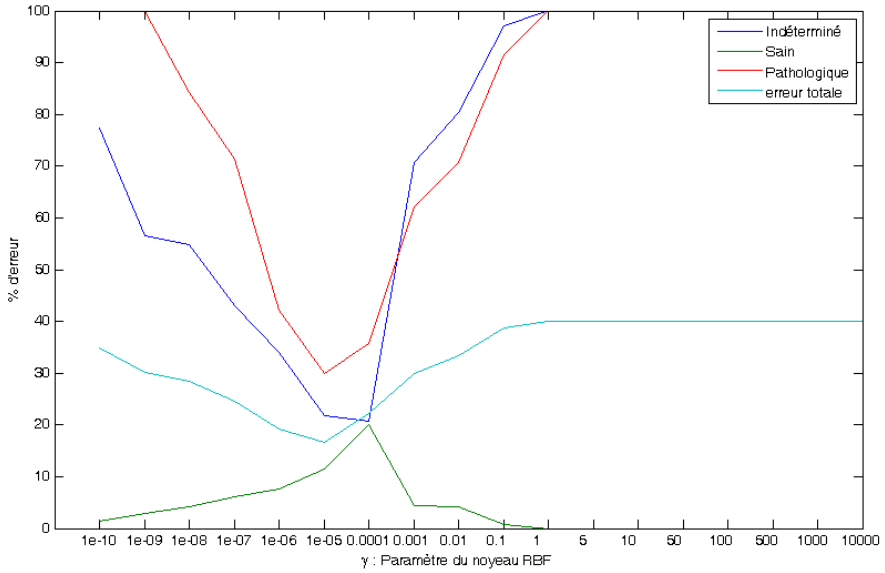


Figure 68 : Variation des taux d'erreur en fonction du noyau RBF avec 120 attributs en entrée des SVM en validation croisée d'ordre 5.

5.1.4 L'ACP pour la sélection d'attributs : Principal Feature Analysis

La PFA est une méthode de sélection d'attributs basée sur l'ACP, soumise par Cohen et al. [98].

La méthode consiste à considérer qu'il est possible de regrouper les attributs de l'espace de départ selon leur contribution à l'espace d'arrivée. Ainsi, deux attributs étant affectés de coefficients similaires dans la combinaison linéaire qui permet de changer d'espace, seront affectés au même groupe d'attributs. Parmi les différents attributs du groupe, on ne sélectionnera finalement que le représentant le plus proche de la moyenne de ce groupe. La méthode des K-means [50] est utilisée pour déterminer automatiquement les groupes, de manière non supervisée. Seul le nombre d'attributs final, donc de clusters dans la méthode des K-Means, doit être fixé de manière arbitraire, par exemple en se basant sur la proportion de variance souhaitée.

Cette méthode se décompose donc en 5 étapes :

- Etape 1 : Calculer le vecteur des valeurs propres et la matrice des vecteurs propres de la matrice de variance-covariance.
- Etape 2 : Choisir la taille q du sous espace de projection en fixant la variance à retenir (cf 5.1.2).
- Etape 3 : Construire la matrice

$$Q = \left(A_1 \cdots A_q \right) = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{q1} \\ \vdots & \ddots & \vdots \\ \alpha_{1n} & \cdots & \alpha_{qn} \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix}$$

  Chaque ligne de cette matrice notée de L_1 à L_n correspond aux coefficients appliqués à un attribut de l'ensemble de départ.

- Etape 4 : En utilisant l'algorithme des K-Means, regrouper les vecteurs L_1 à L_n en p ensembles (p>q). Dans leur article, Cohen et al. conseillent $q + 1 \le p \le q + 5$.
- Etape 5 : Pour chaque ensemble ainsi constitué, trouver le vecteur L_i le plus proche de la moyenne de l'ensemble. i est alors l'indice de l'attribut à sélectionner. On sélectionne donc ainsi p attributs.

Nous avons réalisé l'expérience à deux reprises en conservant 95% puis 99% de la variance soit des valeurs de q de 18 et 91 (respectivement). A l'étape 4, nous avons choisi p=q+5, nous avons donc obtenu 23 et 96 attributs que nous avons placés en entrée du classifieur (SVM, noyau RBF Gaussien).

Nous avons représenté sur la Figure 69 les attributs sélectionnés par cette méthode.

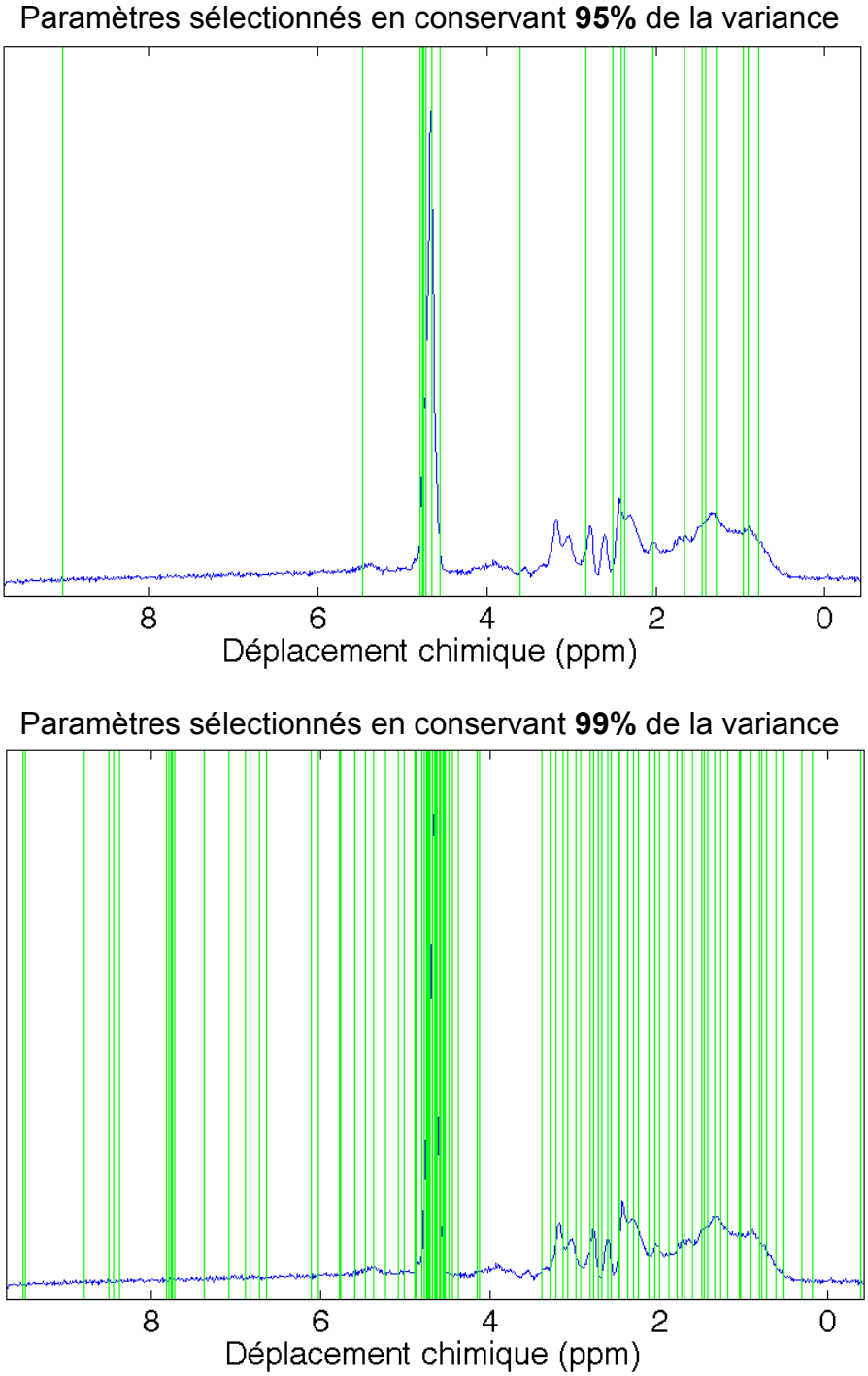


Figure 69 : Attributs sélectionnés par la méthode PFA en conservant 95 et 99 pourcents de la variance et positionnés sur un spectre.

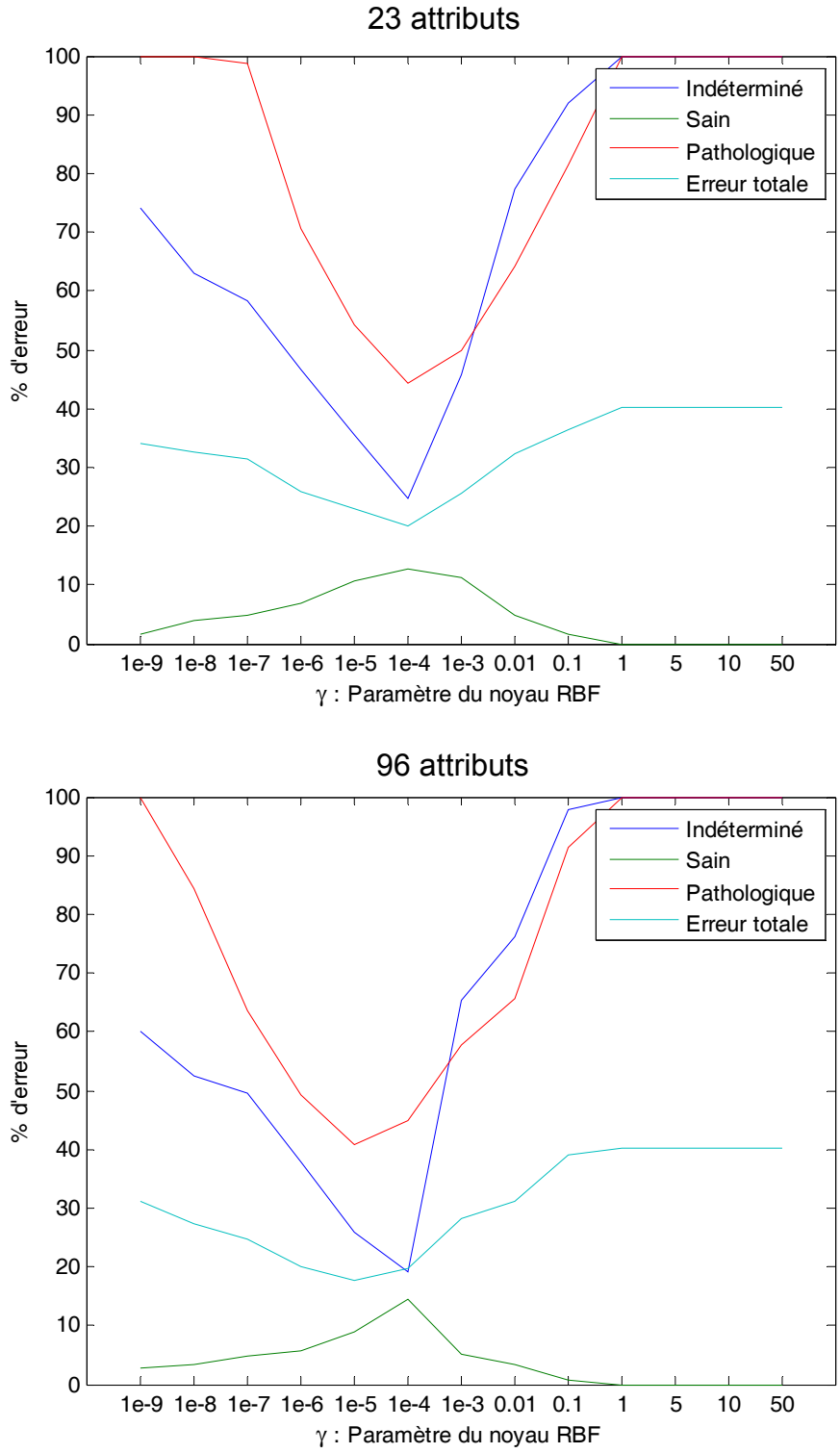


Figure 70 : Variation des taux d'erreur avec les attributs sélectionnés par la méthode PFA en conservant 95% (23 attributs) et 99% (96 attributs) de la variance utilisée en entrée des SVM avec un noyau RBF gaussien.

Nous obtenons avec cette technique un taux d'erreur total de 20,15% avec 23 attributs et de 17,61% avec 96 attributs (Figure 70).

Quelle que soit la méthode utilisée, soit pour réduire la dimension, soit pour sélectionner des attributs, les résultats en classification sont moins bons qu'en utilisant la plage d'étude habituelle des spectroscopistes sur la prostate. Une explication à cela est la

conception même de l'ACP : cette méthode cherche à maximiser la variance, elle sélectionne de ce fait les attributs qui ont la plus grande amplitude de variation parmi les attributs. Or nous avons ici un exemple où la plus grande variation des attributs ne permet pas de les distinguer. En effet nous avons pu remarquer que les attributs sélectionnés étaient en grande partie situés dans la région de l'eau. C'est effectivement dans cette plage de fréquence qu'on observe les plus grandes variations mais même si on ne peut nier qu'il y ait un lien entre l'amplitude du pic d'eau et la classe pathologique du voxel correspondant, cette variation est principalement due à l'efficacité de la méthode de suppression du signal de l'eau de l'appareil RMN. Pour valider cette hypothèse, nous avons effectué la même expérience mais en supprimant la portion du spectre correspondant à l'eau, c'est-à-dire les attributs 494 à 525 (4,83 à 4,52 ppm) (Figure 71). Pour poursuivre dans la même idée, étant donné que la séquence utilise également un schéma de suppression dit "des graisses" nous avons également évalué la méthode en supprimant les attributs à partir de 780 (2 ppm).

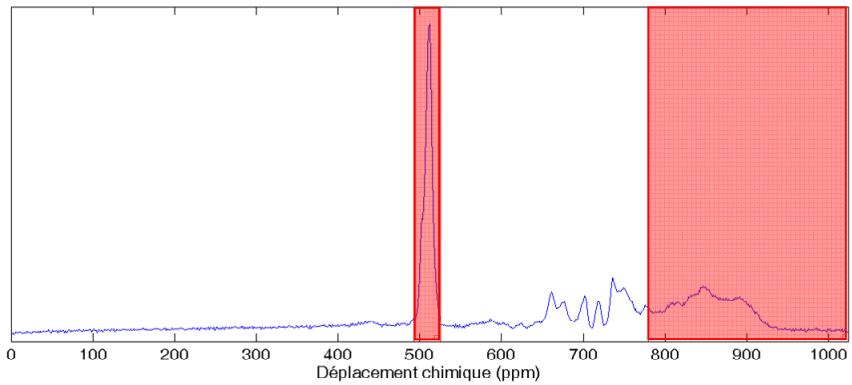


Figure 71 : Bande de l'eau et bande "des graisses" supprimées.

Les deux bandes supprimées comprennent une certaine partie de la variance du spectre initial donc plus on supprime d'attributs de l'ensemble de départ, plus l'algorithme sélectionne d'attributs pour conserver le pourcentage de variance souhaité (95% et 99%). On peut voir dans [Tableau 6] que l'eau est responsable d'une grande partie de la variance de l'ensemble de départ puisqu'il faut sélectionner beaucoup plus d'attributs pour conserver le même pourcentage de variance.

Tableau 6 : Nombre d'attributs des ensembles de données suivant la sélection d'attributs utilisée et l'ensemble de départ.

	Spectre complet	Sans la bande de l'eau	Sans les bandes d'eau et de graisses
Nombre d'attributs	1024	993	749
95% de la variance	23	70	78
99% de la variance	96	165	170

Le Tableau 7 récapitule les taux d'erreur obtenus à chaque expérience et la Figure 72 illustre les attributs sélectionnés dans chaque cas. On peut noter qu'ils baissent à chaque expérience et étant donné que nous avons supprimé les attributs correspondant à l'eau puis aux graisses, on peut en déduire qu'ils ne sont pas significatifs pour la classification. Cependant dans le même temps, le nombre d'attributs sélectionnés a augmenté. La question de savoir si l'amélioration de la classification provient de l'augmentation du nombre d'attributs ou de l'élimination d'attributs détériorant les performances de classification reste alors ouverte.

Tableau 7 : Taux d'erreur (%) obtenus avec les SVM (noyau RBF gaussien) en validation croisée d'ordre 5.

	Spectre complet	Sans la bande de l'eau	Sans les bandes d'eau et de graisses
95% de la variance	20,15	18,36	17,61
99% de la variance	17,61	16,76	14,5

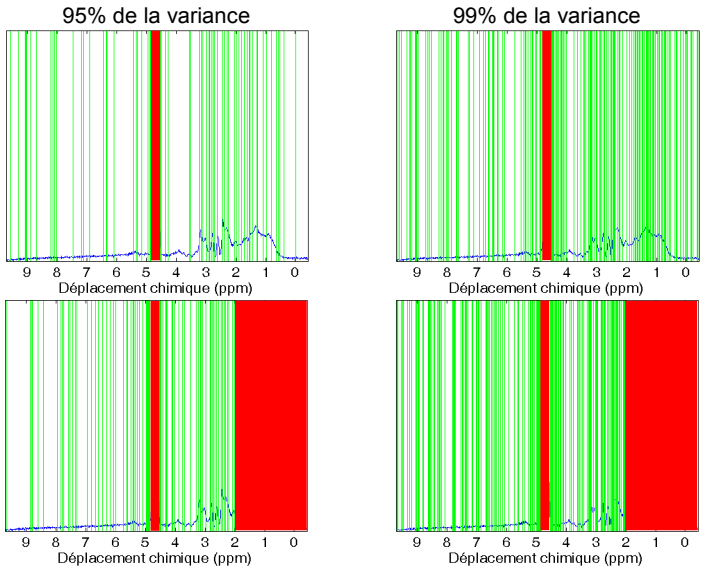


Figure 72 : Attributs sélectionnés par la méthode PFA en conservant 95% et 99% de la variance.
Sur la première ligne sont représentés les attributs sélectionnés en ayant supprimé la bande de l'eau (en rouge) et sur la deuxième sont représentés les attributs sélectionnés après suppression des bandes de l'eau et "des graisses" (en rouge).

5.2 Sequential Forward Selection

Les méthodes de sélection par parcours de l'arbre des possibilités sont généralement constituées de 3 éléments :

- Une procédure de génération des combinaisons : elle sélectionne un sous-ensemble d'attributs noté X selon certains paramètres. Ce sous-ensemble X est ensuite évalué.
- Une fonction d'évaluation J(X) : elle permet de noter l'efficacité du sous ensemble d'attributs généré par la procédure précédente. On utilise généralement le taux d'erreur du classifieur, qui doit donc être le plus faible possible pour que le sous ensemble X soit considéré comme pertinent.
- Un critère d'arrêt : il s'agit habituellement d'un seuil sur la fonction d'évaluation, d'un taux de variation de celle-ci ou encore d'un nombre d'itérations maximum.

Etant donné un ensemble Y de D attributs, la sélection d'attributs consiste à trouver un ensemble X de d attributs avec $d<D$ qui optimisera le classifieur, c'est-à-dire pour lequel $J(X) \leq J(Y)$. On peut, dans certains cas, accepter $J(X)>J(Y)$ si la taille du sous ensemble (X) est très largement inférieure à la taille de l'ensemble initial (Y) et que la différence des fonctions d'évaluation est négligeable. Cela peut être utile dans le cas où la rapidité de décision est importante.

Il existe deux types principaux de méthodes sub-optimales : les méthodes dites "forward", où la fonction de génération commence par sélectionner un seul élément ($d=1$) et en rajoute progressivement durant le déroulement de l'algorithme et les méthodes dites "backward" qui partent de l'ensemble des attributs Y pour en supprimer successivement jusqu'à arriver à l'ensemble X final.

La SFS (Sequential Forward Selection) est la plus simple des méthodes de sélection d'attributs. C'est une méthode déterministe à une seule solution sub-optimale. On part d'un ensemble vide et on lui rajoute l'attribut qui donne le meilleur résultat avec la fonction d'évaluation J. On réitère alors l'opération en rajoutant un nouvel attribut. La condition d'arrêt est alors soit un taux d'erreur à atteindre, soit le fait que le taux d'erreur ne diminue plus, soit que le nombre d'itérations a atteint le nombre d'attributs maximum à sélectionner (illustré Figure 73).

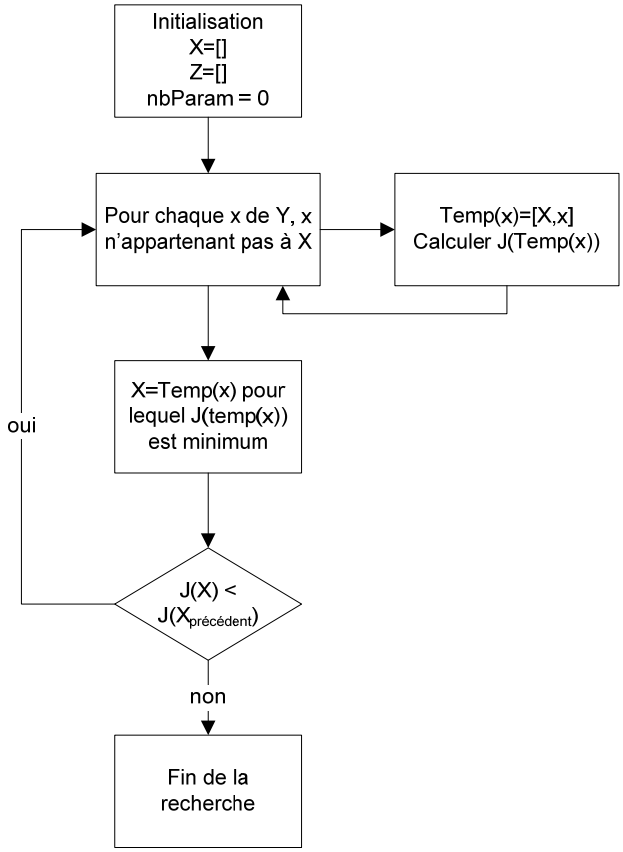


Figure 73 : Algorithme SFS.

Nous avons donc appliqué cet algorithme à notre base de spectre complet (1024 attributs), avec les SVM comme classifieur, la fonction d'évaluation étant le taux d'erreur total. Nous avons choisi comme condition d'arrêt le fait que le taux d'erreur augmente (ou que la taille de l'ensemble des attributs sélectionnés soit égale à la taille de l'ensemble de départ). Le résultat est un ensemble de 12 attributs (Figure 74) atteint avec un taux d'erreur de 12,2%.

Comme pour la méthode de sélection par plage (voir paragraphe suivant), les attributs sélectionnés se situent principalement dans la plage du citrate, de la choline, de l'eau et des lipides. Un attribut entre l'eau et la choline apparait également.

Pour valider cette sélection d'attributs, nous avons donc restreint notre base à cet ensemble puis nous avons fait une validation croisée d'ordre 5 avec les SVM en faisant varier la taille du noyau RBF gaussien de 0,0001 à 1000. Le meilleur taux d'erreur atteint est alors de 20,4% (Figure 75). Cette valeur est supérieure à celle fournie par la SFS et à celle obtenue par la sélection de plages mais elle est atteinte avec seulement 12 attributs.

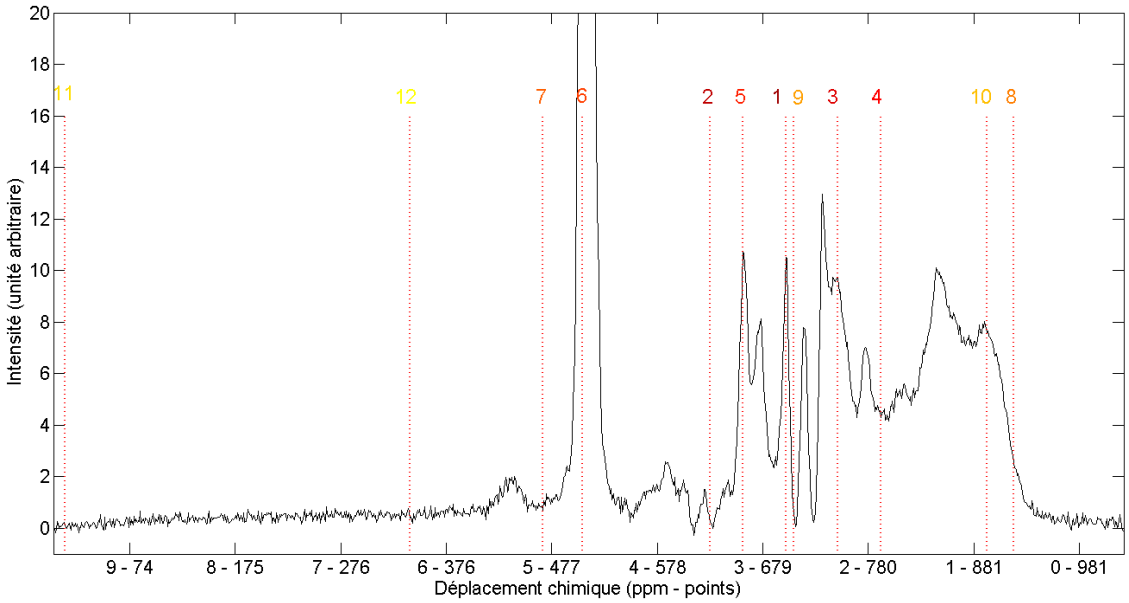


Figure 74 : Les 12 attributs sélectionnés par la méthode SFS positionnés sur un spectre moyen et numérotés par leur ordre de sélection par l'algorithme.

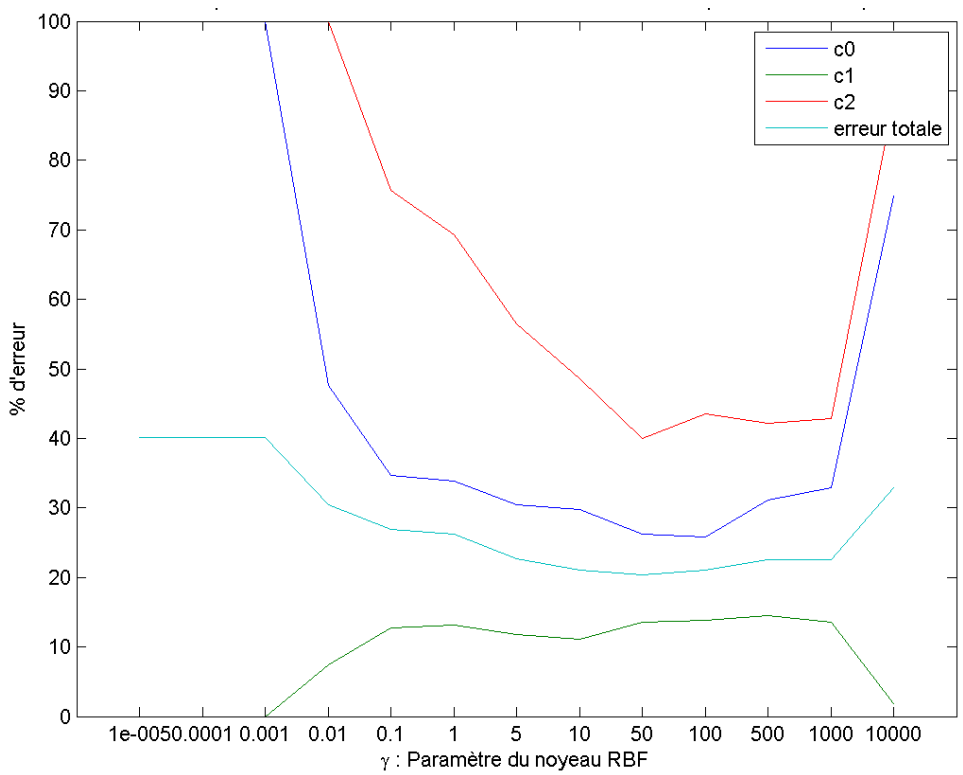


Figure 75 : Variation des taux d'erreur par classe et total en fonction du noyau RBF gaussien utilisé.

5.3 Sélection par plage

Les spectres de la prostate sont habituellement étudiés sur la bande des 2-4 ppm, car c'est dans cette plage que se situe la fréquence de résonance des métabolites présentant un intérêt dans la détection du cancer de la prostate. Nous avons voulu voir s'il pouvait

exister d'autres métabolites présents dans le spectre pouvant améliorer les taux de classification. Pour cela, nous avons défini différentes tailles de fenêtre (Figure 76) que nous avons fait glisser sur le spectre et à chaque fois nous avons évalué le taux d'erreur total obtenu en utilisant les SVM.

Les tests ont été réalisés avec 9 tailles de fenêtre différente de 1% du spectre (10 points soit environ 0,1 ppm) à 100% du spectre (1024 points soit environ 10,15 ppm). Les tailles de fenêtre intermédiaires sont de 2%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60% , 70%, 80% et 90%. Nous avons utilisé un pas de déplacement de la fenêtre de 5 points (Figure 76 : le pas est de 100 points).

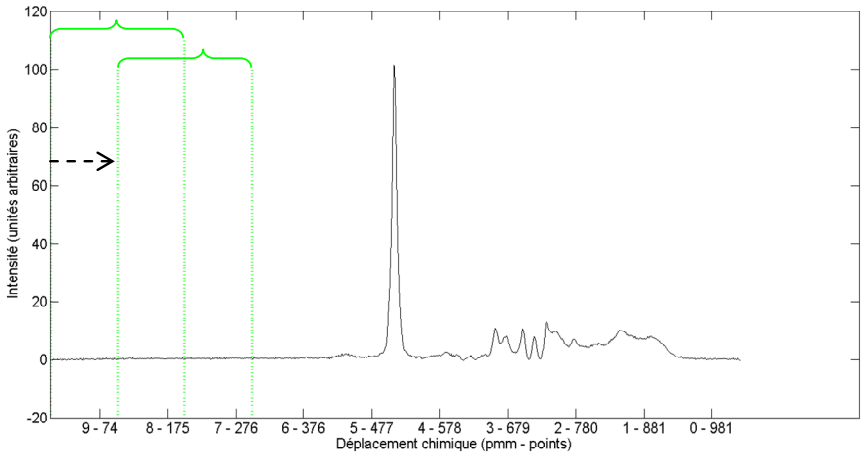


Figure 76 : Fenêtre glissante de 200 points (≈2 ppm).

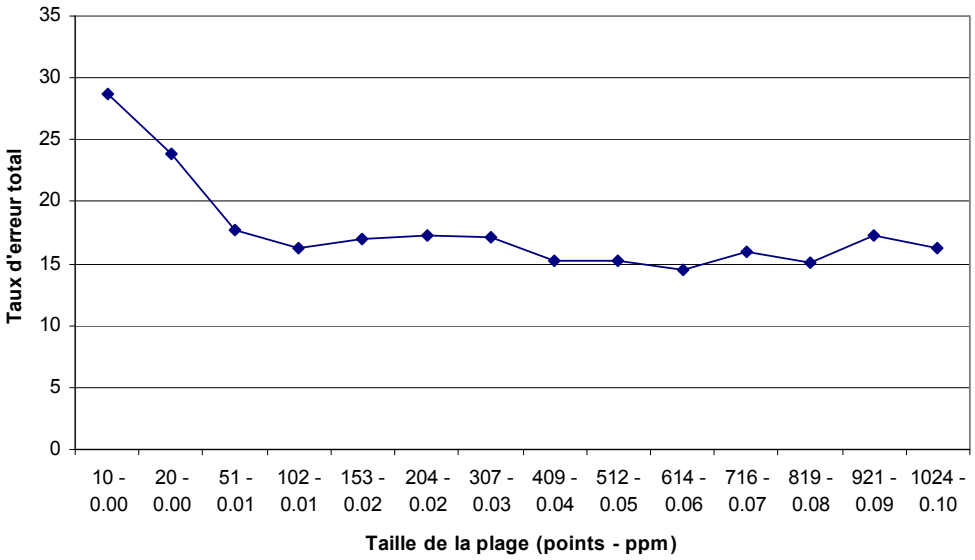


Figure 77 : Variation du meilleur taux d'erreur totale atteint en fonction de la taille de la plage.

Pour chaque taille de fenêtre, nous avons relevé le taux d'erreur minimum atteint et nous l'avons représenté sur la Figure 77. On remarque que ce minimum décroit rapidement dès les premières augmentations du nombre d'attributs utilisés pour se stabiliser entre 14 et 17 % d'erreur dès la plage de 51 points.

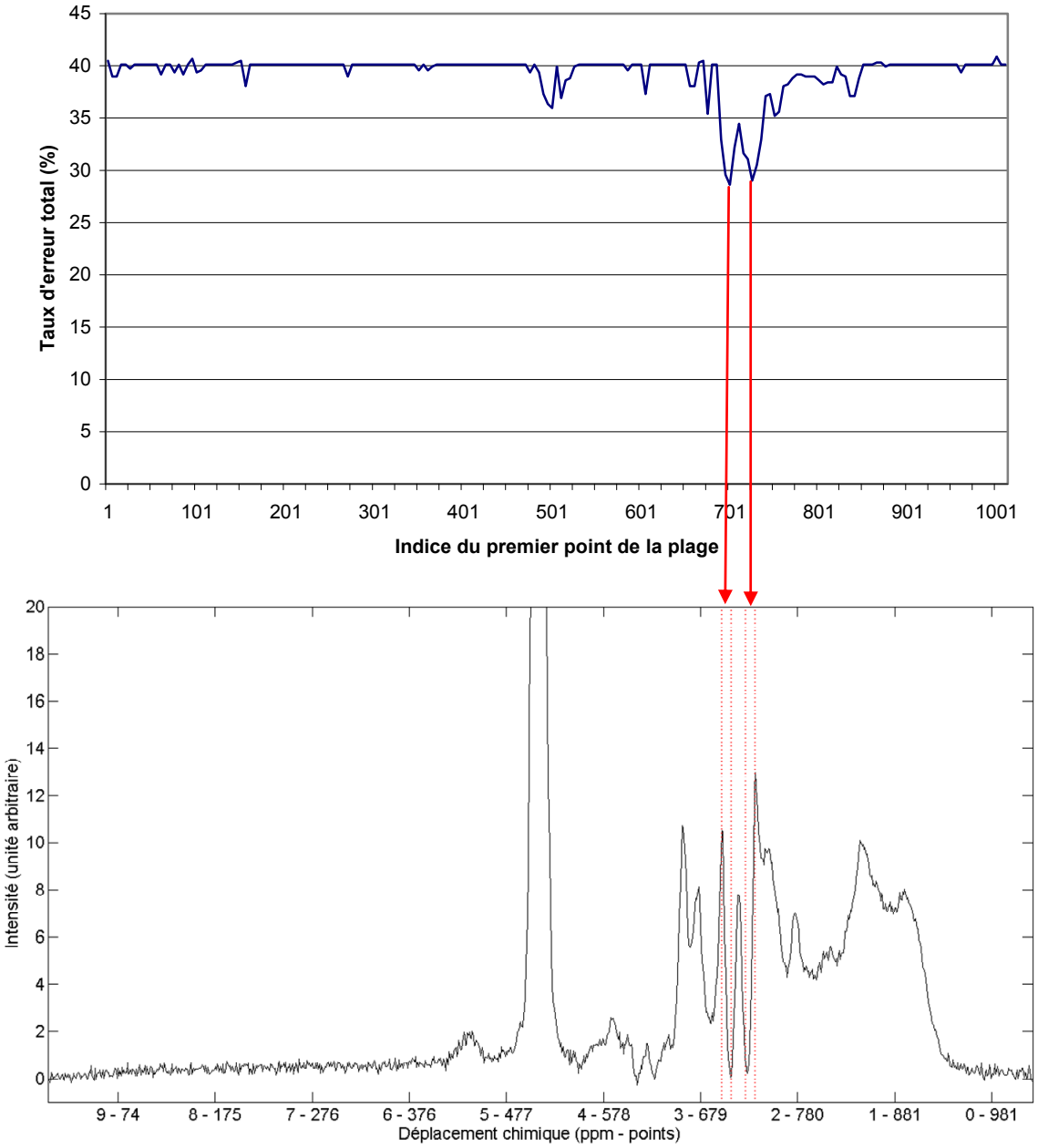


Figure 78 : Plage de 10 points soit environ 0,1 ppm par pas de 5 points : représentation des deux principales plages entrainant une baisse du taux d'erreur.

.

Avec une taille de fenêtre de 10 points, les taux d'erreur totale ne sont pas excellents puisque nous atteignons au mieux un résultat de 28,63%. Deux plages (Figure 78) permettent d'atteindre un taux d'erreur inférieur à 29% : elle se situe dans la plage du

citrate sur le premier et le troisième pic. On peut également noter deux légères baisses sur les plages correspondant à l'eau (4,7 ppm) et dans la région des lipides (< 2 ppm).

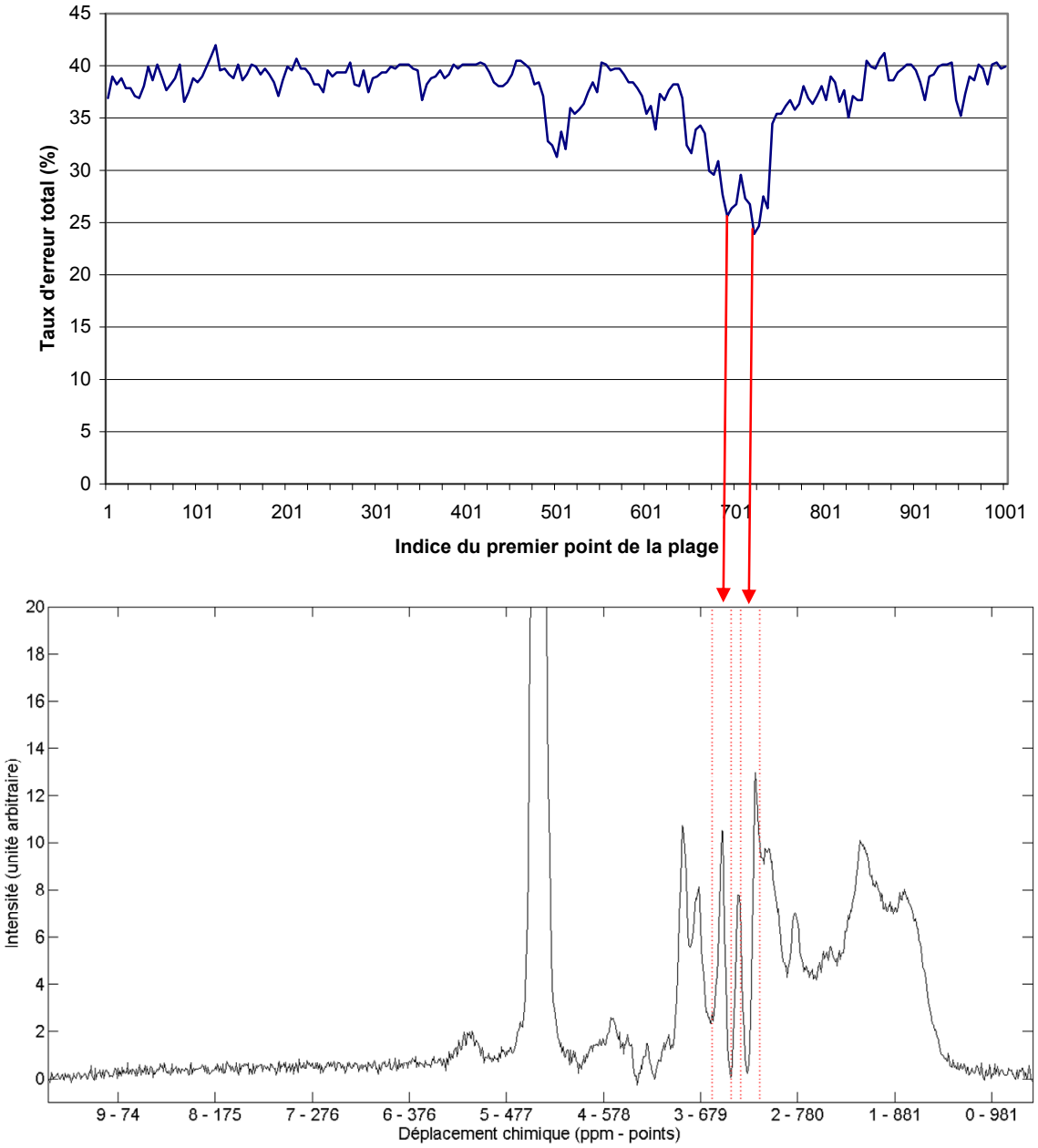
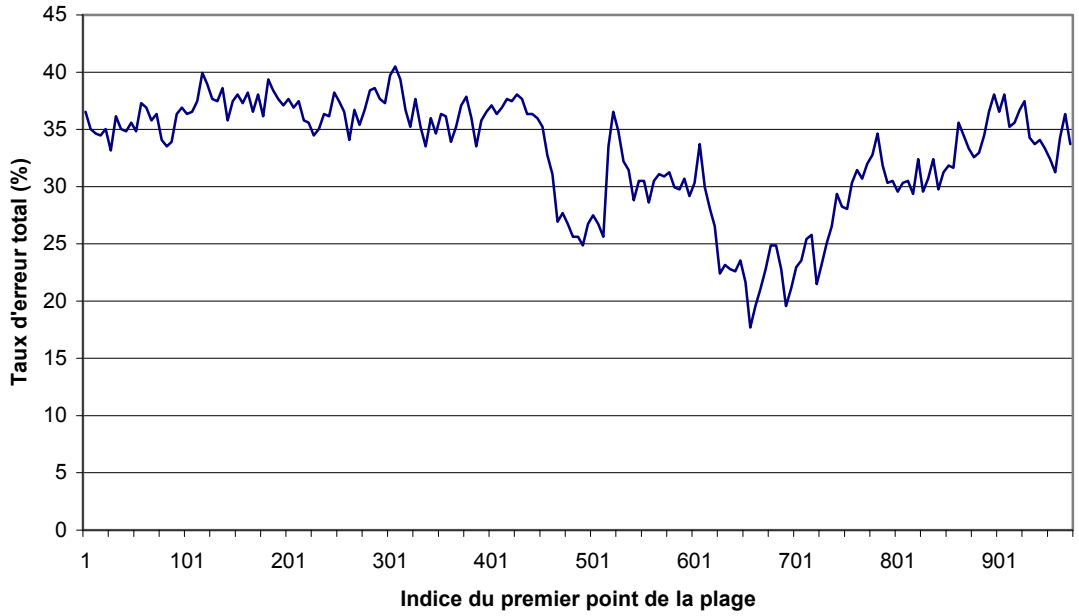


Figure 79 : Pas de 5, plage de 2% (20 points).

L'extension de la taille de la plage à 20 points permet d'améliorer nettement les résultats puisqu'on atteint une erreur de 23,92% (Figure 79). Là encore, deux plages proches situées dans la zone du citrate permettent d'atteindre un taux d'erreur inférieur à 26%. Des "sursauts" apparaissent également dans la zone de l'eau, de la choline et des lipides.



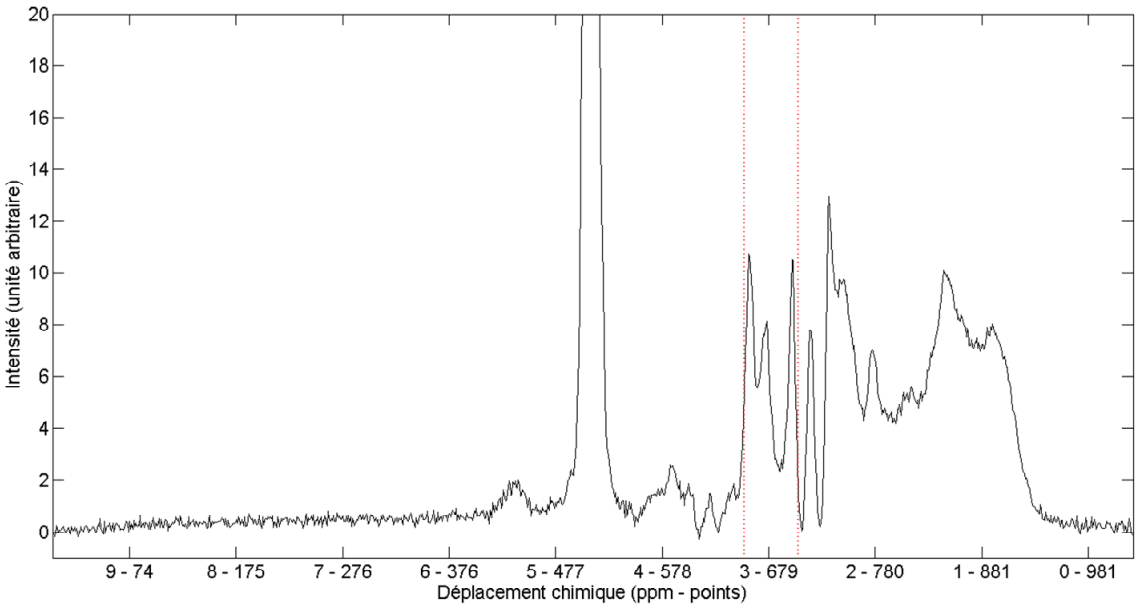
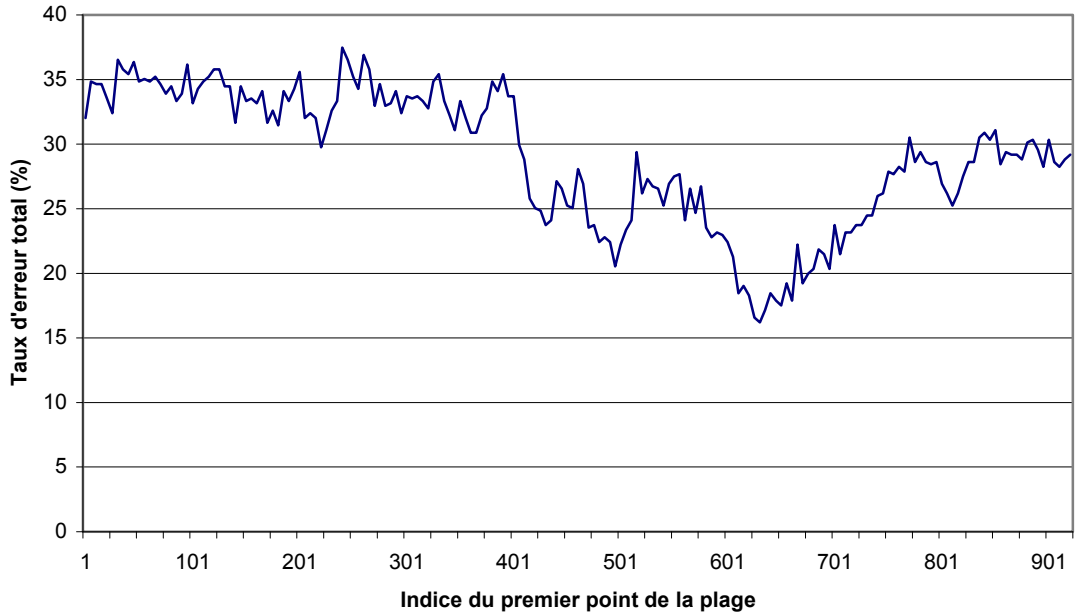


Figure 80 : Pas de 5, plage de 5% (51 points).



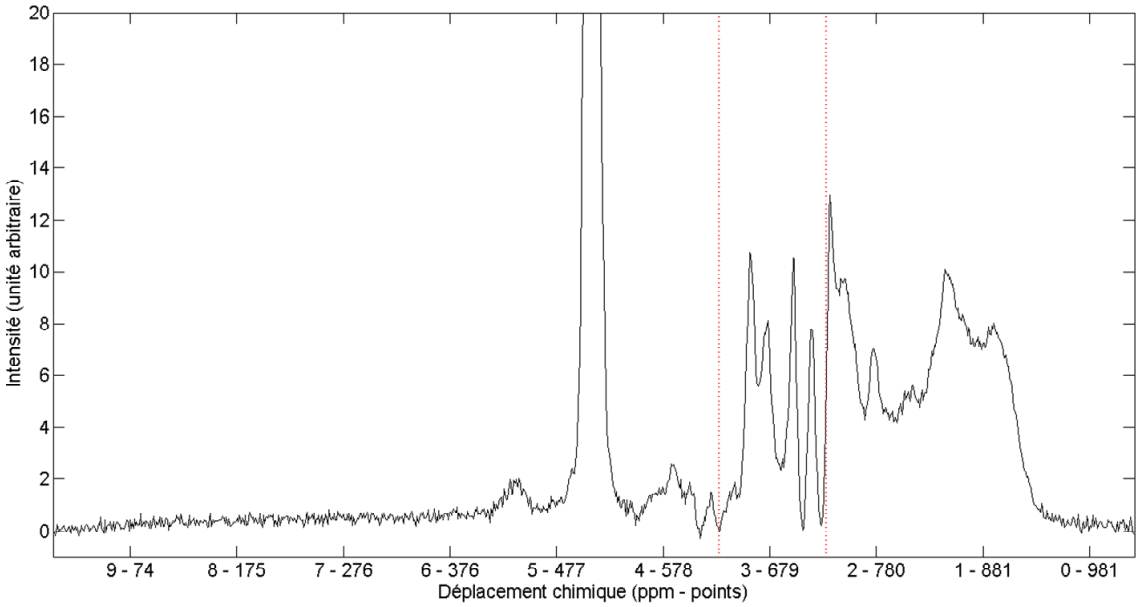


Figure 81 : Pas de 5, plage de 10 % (102 points).

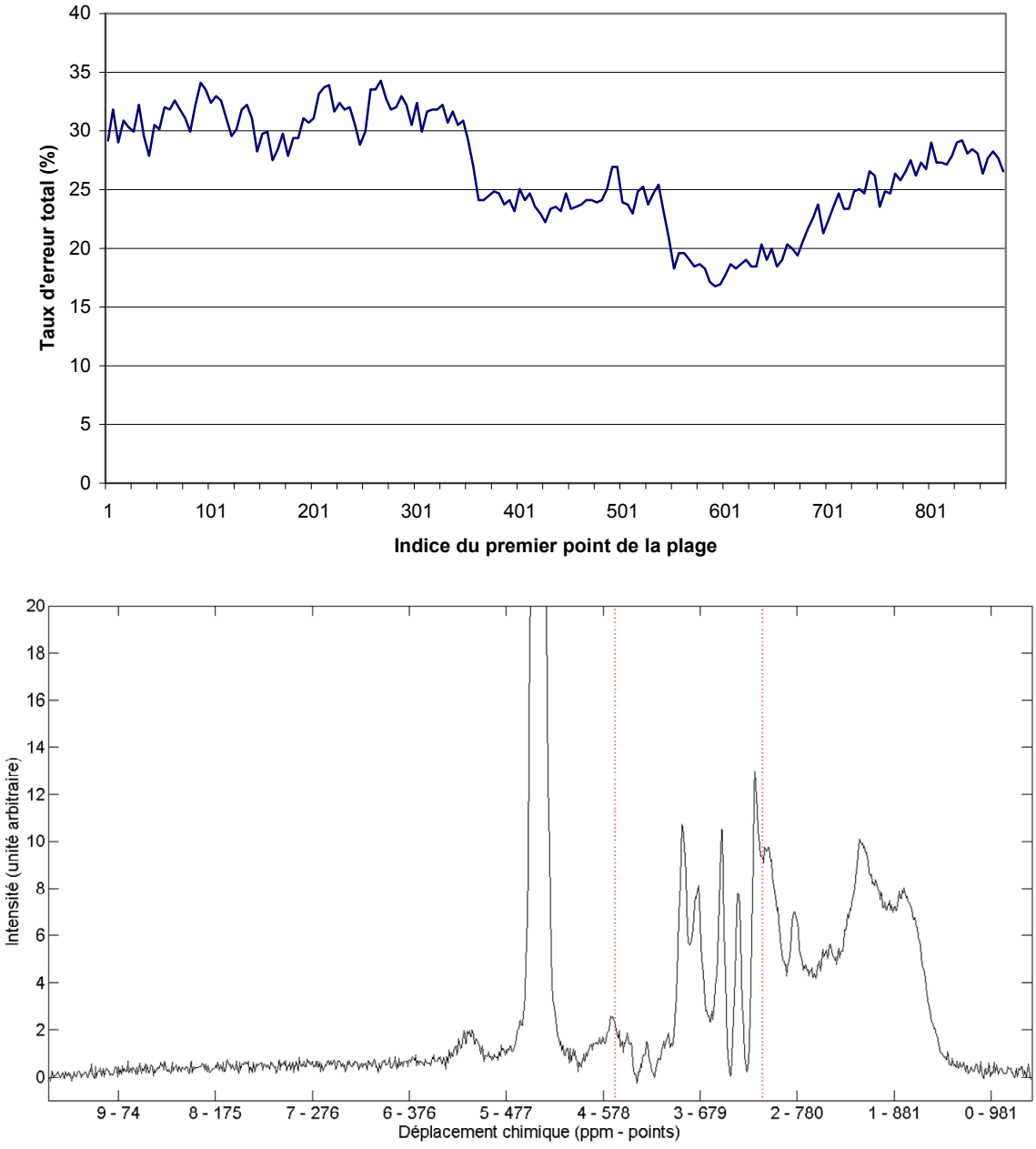


Figure 82 : Pas de 5, plage de 15% (153 points).

Les plages de 5 (Figure 80), 10 (Figure 81) et 15 % (Figure 82) permettent d'améliorer le taux de d'erreur de classification pour approcher les 15%. Le meilleur taux est toujours atteint sur une plage englobant la choline et le premier pic de citrate. Ce résultat confirme les études précédentes recommandant de se baser sur ces métabolites pour diagnostiquer un cancer de la prostate. On peut toujours noter des améliorations du taux d'erreur avec des zones englobant l'eau ou les lipides.

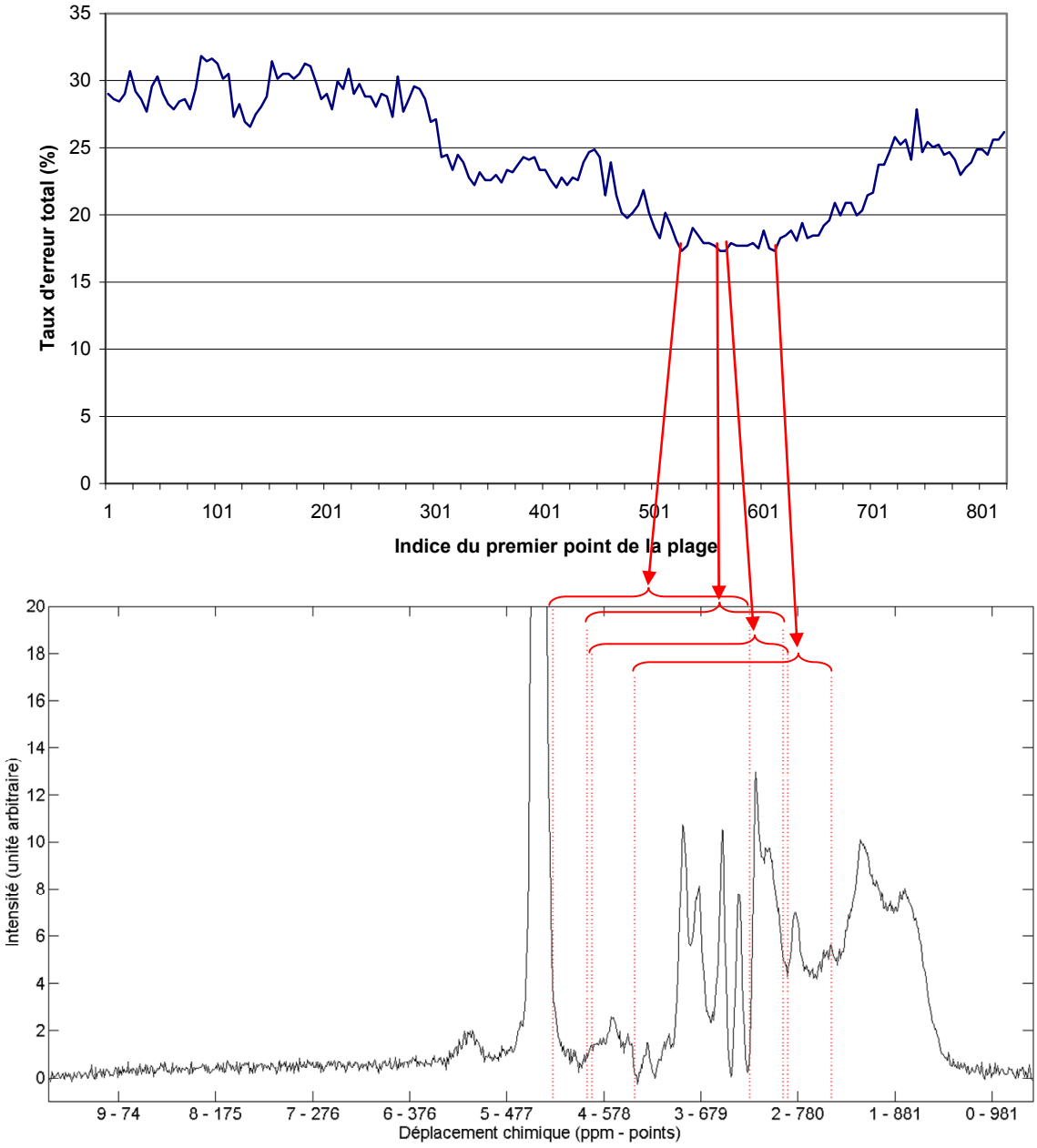


Figure 83 : Pas de 5, plage de 20% (204 points).

A partir de la plage de 20%, la courbe a tendance à se lisser (Figure 83) car la fenêtre englobe rapidement l'intégralité des fréquences apportant une amélioration du taux d'erreur. Notamment sur la plage de 20%, 4 plages offrent le même meilleur taux de classification (17,33%). Ces 4 plages englobent la choline et le premier pic de citrate et les plages voisines présentent un taux d'erreur très proche.

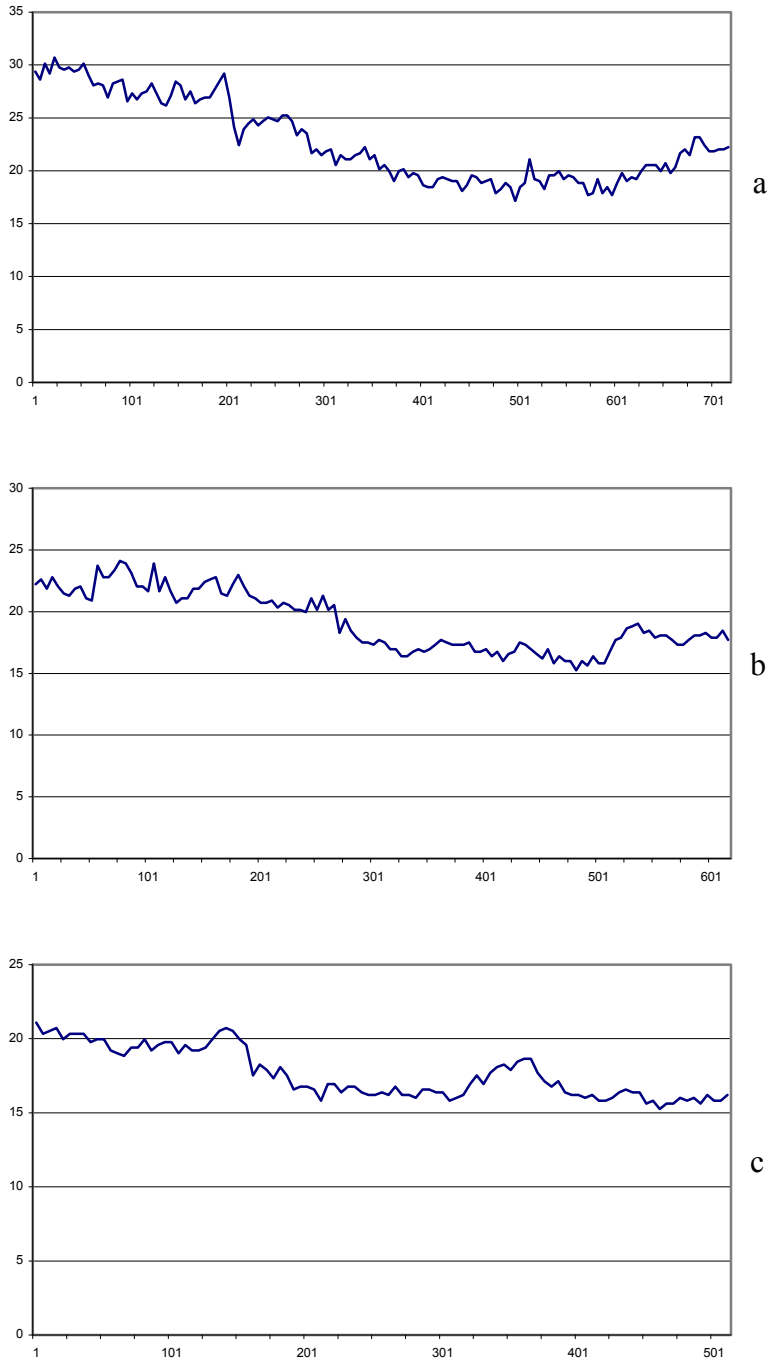


Figure 84 : Evolution du taux d'erreur pour les plages de 307 (a), 409 (b) et 512 points (c).

Au delà de 30% du spectre, le taux d'erreur minimum ne varie que du fait du tirage aléatoire des échantillons lors de la validation croisée. La meilleure plage englobe toujours l'eau, la choline, le citrate et une partie des lipides (Figure 84 et Figure 85).

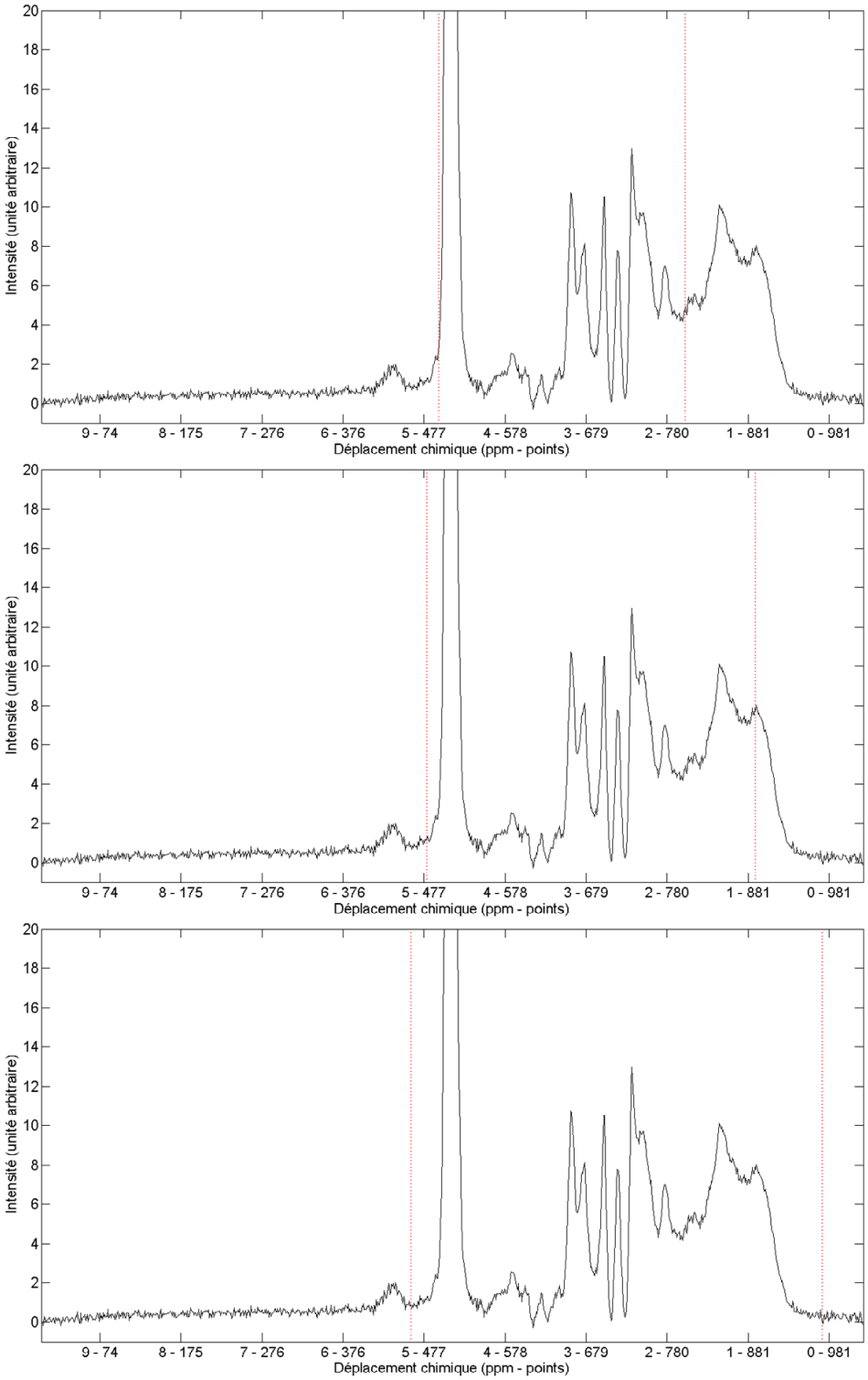


Figure 85 : Représentation des plages permettant d'obtenir le meilleur taux de classification pour les plages de 307, 409, et 512 points.

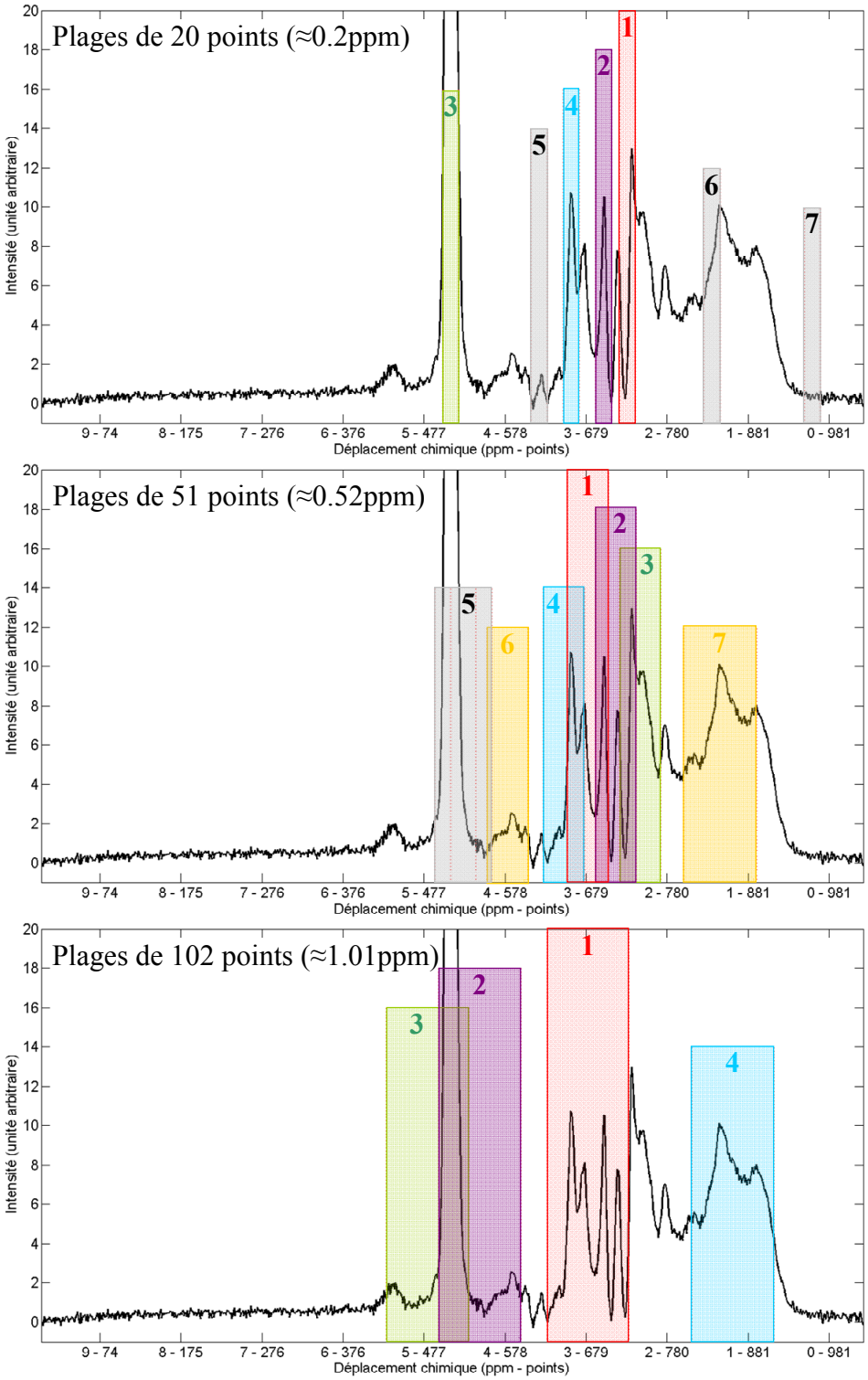


Figure 86 : Plages permettant d'obtenir individuellement une amélioration notable du taux d'erreur totale.

Les chiffres sur les plages représentent le classement par taux d'erreur croissant.

Une étude plus attentive des plages de 20, 51 et 102 points montre qu'elles présentent des minimums locaux d'erreur (Figure 86). Par ordre croissant des taux d'erreur, les plages sur le citrate apparaissent en premier, suivies de l'eau et de la choline qui présentent des taux d'erreur, en moyenne, similaires. Un fait plus intéressant dans notre

recherche de nouveaux marqueurs du cancer de la prostate, est l'apparition de bandes sur des pics intermédiaires à l'eau et à la choline ainsi que dans la plage des lipides. Ces plages de fréquence correspondent certainement à des molécules ayant un rôle dans le métabolisme de la prostate ou du cancer ou simplement issues de ce métabolisme cependant elles n'améliorent pas les résultats en classification quand elles sont associées aux plages donnant les meilleurs résultats.

5.4 Conclusion

Nous avons évalué différentes techniques de sélection d'attributs. Aucune ne nous a permis de détecter, dans le spectre, des attributs plus significatifs que la plage allant de 2 à 4ppm. Les métabolites les plus pertinents restent la choline et le citrate. On peut cependant noter que la créatine, la spermine et le myo-inositol (ou une autre molécule résonnant à une fréquence proche) participent à l'amélioration de la classification, même s'ils ne sont pas discriminants utilisés seuls. Enfin avec certaines méthodes, le pic d'eau et les pics de lipides (aux alentours de 2 ppm) présentent une certaine pertinence dans le résultat de classification. Cependant, étant donné que nous utilisons une séquence de suppression du signal de l'eau et de la graisse (CHESS), nous pouvons nous demander dans quelle mesure ces résultats sont valables.

Chapitre 6 : Données pré-cliniques

Les études menées par la société Oncodesign emploient généralement des cohortes de 10 rats (5 dans le groupe contrôle et 5 dans le groupe test). L'étude d'un rat en IRM durant couramment une heure ou plus, en prenant en compte les temps de manipulation et l'accès à l'appareil, qui n'est pas aisé de par le statut EOPS, il est difficile d'avoir des groupes plus importants. Les dates auxquelles sont réalisés les examens d'imagerie sont précisément définies dans les protocoles de recherche et ne peuvent être modifiées car les tumeurs "poussent" tellement vite que pour être comparées les unes aux autres, elles doivent être imagées en même temps : on ne peut donc pas réaliser les acquisitions sur plusieurs jours.

Nous espérions compenser ce faible nombre d'individus grâce à la spectroscopie multi-voxels 3D. Malheureusement, cette technique s'est révélée impossible à mettre en œuvre sur la machine dans son état actuel. Nous ne pouvions donc acquérir à chaque examen qu'un ou deux spectres par rat. Ce nombre de données se révèle largement insuffisant pour espérer réaliser un apprentissage fiable avec un classifieur. Les données ont donc été analysées avec des études statistiques classiques.

Comme nous l'avons dit dans le paragraphe 2.4.2, étant donné que nous ne connaissions pas l'appareil, nous avons choisi de procéder par étape pour nous familiariser avec celui-ci ainsi qu'avec les protocoles d'imagerie du rongeur. Nous avons commencé par des acquisitions faciles, dans le cerveau qui est connu et sur lequel on trouve une abondante littérature, pour finir progressivement par l'étude de tumeurs de la prostate et de leurs réactions en fonction des traitements.

Dans la suite de ce chapitre, nous allons donc présenter les principaux résultats obtenus au cours de ces expériences, en commençant par le cerveau et en terminant par la prostate.

6.1 Expériences sur le cerveau

La première étude préclinique réalisée dans le cadre de cette thèse comprenait des rats auxquels des modèles de tumeurs du cerveau (U-87 MG et CGL9) ont été implantés en position orthotopique. En plus de l'imagerie RMN et de la SRM, des examens de diffusion ont également été réalisés sur ces rats. Les aires des pics de choline et de N-acetyl-aspartate

(NAA) ont été mesurées avec l'aide de LCModel. Un exemple de l'analyse par LCModel de chaque type de spectre à été représenté sur la Figure 87.

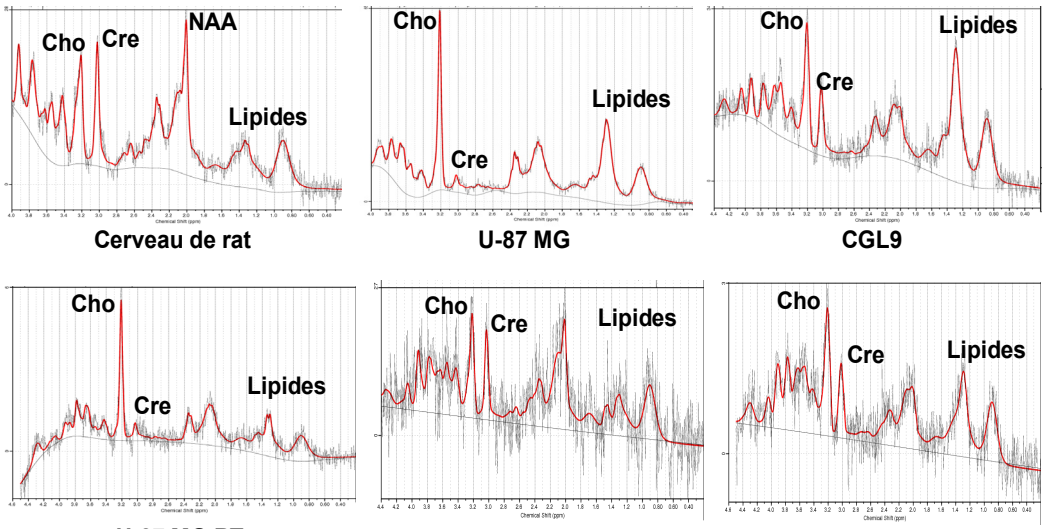


Figure 87 Spectres de cerveaux de rats sains, de tumeurs CGL9 et U87-MG avec et sans traitement (RT : radiothérapie, TMZ : Témozolomide).
En abscisse, le déplacement chimique (de 0,4 ppm à 4ppm, de droite à gauche) et en ordonnée, l'intensité du signal en unité abstraite.

Cinq rats U-87 ont eu un traitement par radiothérapie, cinq rats U-87 et cinq rats CGL9 ont eu un traitement par chimiothérapie (Témolozomide : TMZ) selon les protocoles standards. Les résultats biologiques obtenus par la société Oncodesign ont fait l'objet d'un poster au congrès de l'American Association for Cancer Reseach 2008 [99], sur la réponse de ces modèles de tumeurs à la radiothérapie et à une molécule de chimiothérapie. Un exemple des spectres obtenus a été représenté sur la Figure 89.

La choline est comme nous l'avons déjà vu un marqueur de la prolifération cellulaire alors que le N-acétyl-aspartate (NAA) est un marqueur de l'activité neuronale (il est produit par les neurones fonctionnant correctement). La variation du rapport entre la concentration en choline et en NAA a donc été mesurée (Figure 88) : l'effet du Témolozomide apparait nettement avec le modèle U87-MG avec un retour à la normale et un peu moins nettement avec le modèle CGL9 où les rapports restent plus élevés avec le traitement que sans. On n'observe aucune variation significative en SRM avec la radiothérapie si ce n'est une baisse globale des pics due à la destruction du métabolisme par le rayonnement. Cette étude nous a en outre permis de valider la faisabilité de la SRM sur le modèle animal au niveau cérébral ainsi que la possibilité de détecter des changements métaboliques en SRM en fonction du traitement.

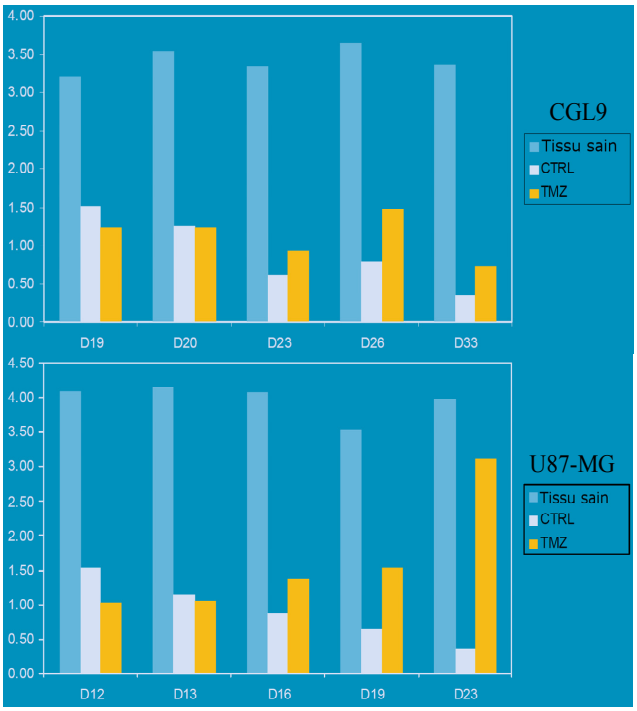


Figure 88 : Variations du rapport NAA/choline avec les deux modeles de gliome en fonction du temps.

6.2 Expériences sur la prostate

Un jeu de données issu de rats greffés en sous cutané avec une tumeur PAC 120 a également été analysé avec l'algorithme DSF. L'étude comportait 3 rats contrôles et 3 rats castrés. Le but de cette étude était, en effet, de constater l'effet de la castration des rats et donc de la suppression de la sécrétion d'androgènes sur les tumeurs prostatiques. Une spectroscopie a été réalisée dans la tumeur, 3 jours avant la castration puis 11 jours après et enfin au bout de 25 jours. Les spectres obtenus ont été représentés sur la Figure 89.

Sur un si faible ensemble de données, il est impossible d'utiliser des méthodes de classification. Nous nous sommes donc restreints, de nouveau, à des analyses classiques sur les données extraites à partir d'une décomposition en Lorentziennes par l'algorithme DSF.

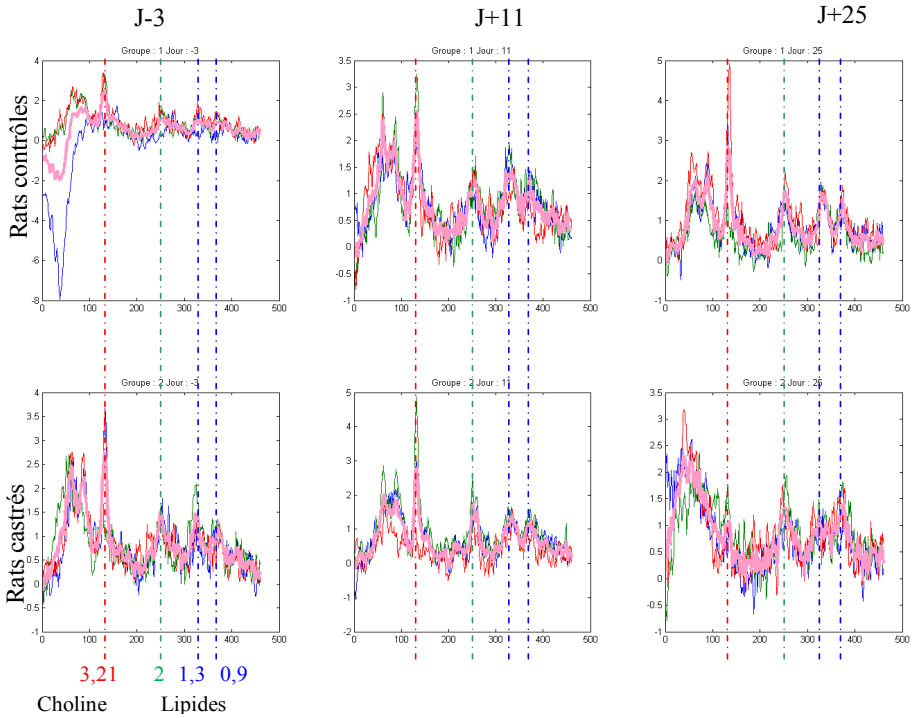


Figure 89 : Spectres obtenus pour chaque rat de chaque groupe au trois temps d'imagerie.
Chaque rat est représenté par une couleur, le spectre rose représente la moyenne des trois.

Cet exemple (Figure 89) nous montre les limites apportées par le travail réalisé sur des groupes de seulement trois rats : au premier jour, nous avons obtenu avec un rat du groupe contrôle (bleu), une spectroscopie de très mauvaise qualité alors que le protocole avait été parfaitement respecté. Les données de ce rat sont donc perdues pour toute la suite de l'étude. En observant ces spectres, il semblerait que la castration entraine une baisse du taux de choline. A partir de la décomposition obtenue par l'algorithme DSF, nous avons estimé les concentrations relatives de certains métabolites en calculant l'aire sous la courbe des pics. Les variations, regroupées par métabolite et par groupe de rats (castrés (CSTR) et contrôle (CTRL)), ont été représentées sur la Figure 90. La baisse moyenne du taux de choline observée (Figure 89), peut difficilement être confirmée par la mesure de l'aire sous la courbe, compte tenu des variations entre chaque rat du même groupe.

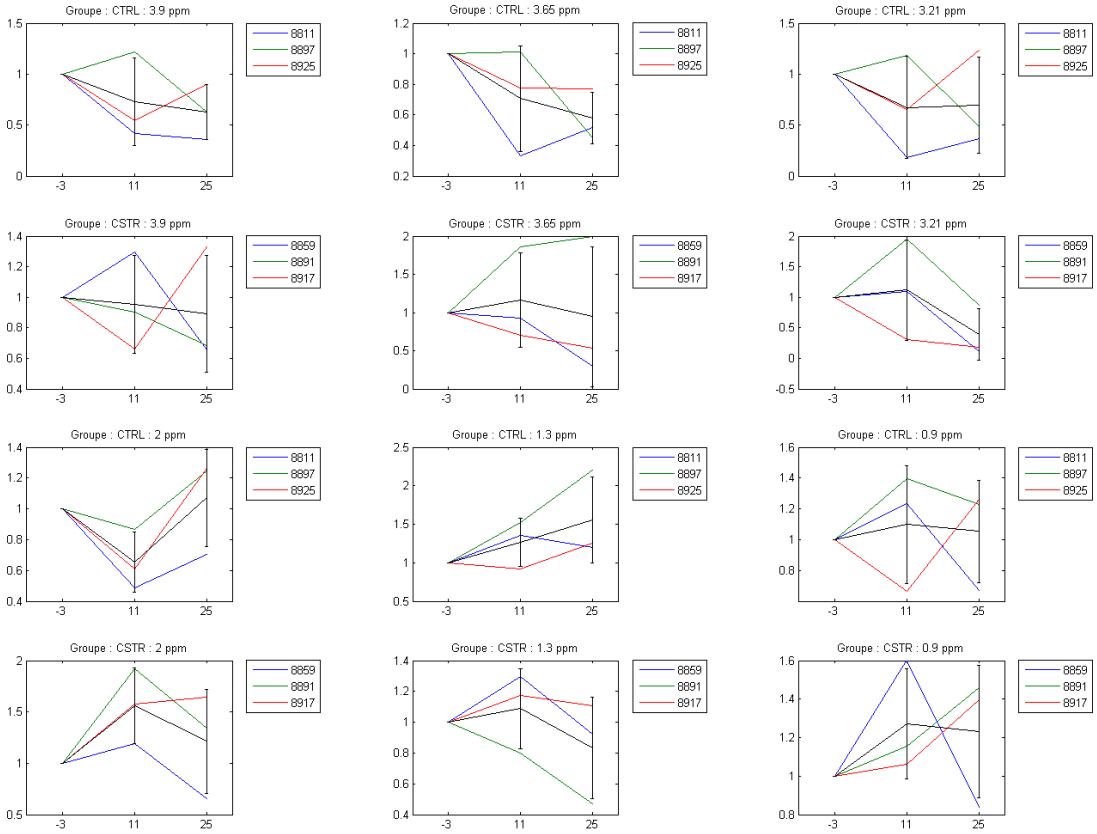


Figure 90 : Pour chaque position du spectre étudié (3.9, 3.65, 3.21, 2, 1.3 et 0.9 ppm), variation de l'aire sous la courbe de la lorentzienne détectée pour chaque rat de chaque groupe (CTRL : contrôle CSTR : castré).

Pour la suite de l'étude, le modèle de tumeur PC3-MM2 greffé en position orthotopique a été utilisé. Les concentrations ont été estimées par LCModel qui donne une dispersion équivalente (Figure 91). En outre, ce logiciel ne possède pas de modèle pour estimer les concentrations des métabolites dans les prostates de rats qui ont un métabolisme différent des prostates humaines.

Nous observons donc dans le cas des tumeurs prostatiques chez le rat, une grande variation des données issues de la SRM. Il est par conséquent difficile d'obtenir des résultats fiables avec cette technique. Il faudrait déterminer si ces variations sont naturelles chez le rat ou si elles proviennent d'une trop grande sensibilité de la technique aux conditions d'acquisition (ou plus vraisemblablement aux deux phénomènes dans des proportions à déterminer). Les études ont été poursuivies en mesurant les marqueurs habituels que sont le volume tumoral et la survie des animaux. Ces données sont plus homogènes et montrent bien une amélioration en cas de traitement (soit un allongement de la survie, soit une stagnation voire une diminution du volume tumoral).

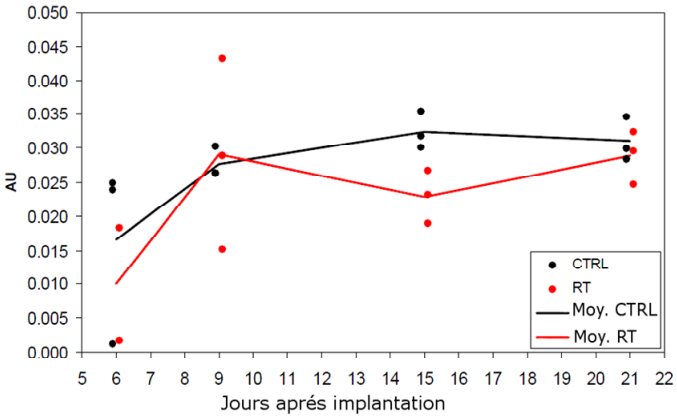


Figure 91 : Concentrations en choline estimées par LCModel dans des tumeurs PC3-MM2 orthotopiques.

Les résultats de cette étude ont fait l'objet d'un poster au congrès de l'American Association for Cancer Research 2009 [100] et d'un autre au congrès de l'International Society of Magnetic Resonance in Medecine [101]. Dans ces deux publications, nous avons principalement montré que le métabolisme prostatique du rat est différent de celui de l'homme notamment par l'absence de citrate dans les spectres de l'animal, mais que certaines similarités sont présentes, notamment pour la choline. Il existe une très grande variance dans les données obtenues en SRM. Il est donc nécessaire de poursuivre la mise au point de la technique, mais nous avons montré que cela était techniquement possible.

6.3 Conclusion

Nous n'avons pu mener l'intégralité des recherches que nous souhaitions faire dans le cas pré-clinique. Néanmoins, nous avons quand même pu vérifier la faisabilité de la SRM dans le cerveau du rat et surtout vérifier qu'il existait une variation des spectres en fonction du traitement. Le résultat est beaucoup plus nuancé en ce qui concerne la pathologie prostatique.

Nous avons pu remarquer, comme dans le chapitre précédent, une variation non négligeable des pics de lipides dans les différents tests qui ont été réalisés. Dans ce cas, les données ont été acquises sans utiliser de séquence de suppression du signal des lipides, cette voie reste donc à explorer.

Conclusion

Dans ce mémoire, nous avons abordé l'acquisition et les traitements des spectres issus de la résonance magnétique du proton, puis nous avons étudié leur classification à l'aide d'algorithmes reconnus. Enfin, nous avons recherché des marqueurs parmi les attributs des spectres permettant d'améliorer cette classification.

Au début de cette thèse, il n'existait pas de littérature concernant la classification automatique de spectre de résonance magnétique de la prostate puisque les premières publications ont été éditées en 2007. Il existait toutefois une abondante littérature traitant d'une part, de la pertinence de l'utilisation de la SRM du proton dans le dépistage du cancer de la prostate, principalement basée sur l'analyse des pics de citrate et de choline et d'autre part de la classification automatique de SRM du proton du cerveau.

Les résultats obtenus sur les spectres cérébraux étaient très satisfaisants puisque les auteurs arrivaient à distinguer les principales catégories de tumeurs avec quelques pourcents d'erreurs en moyenne. Nous nous sommes donc grandement inspirés de leurs méthodes dans un premier temps avant de nous apercevoir que toutes n'étaient pas transposables à l'étude de la prostate. Ainsi dans le cerveau, les auteurs supprimaient le pic d'eau résiduel, alors que dans le cas de la prostate, il est inutile de s'en préoccuper. Inversement les auteurs travaillant sur le cerveau ont obtenu de très bons résultats en utilisant le module du spectre alors que pour la prostate, nous avons obtenus les meilleurs résultats avec la partie réelle du spectre avec une correction de phase automatique. Nous avons utilisé un algorithme de correction de phase qui a été développé en 2002 pendant que les études concernant le cerveau étaient menées. Cet algorithme mériterait d'être comparé avec les algorithmes employés dans ces publications pour vérifier s'il permet d'obtenir de meilleurs résultats.

Il ressort de notre étude que le meilleur taux de classification a été obtenu par les SVM, avec comme prétraitement une correction de phase par l'algorithme ACME suivi d'une correction de la ligne de base avec l'algorithme CLB. L'ensemble des spectres a ensuite été restreint à la portion 2-4ppm avant d'être normalisé par la norme L2. Cet enchainement est résumé sur la Figure 92.

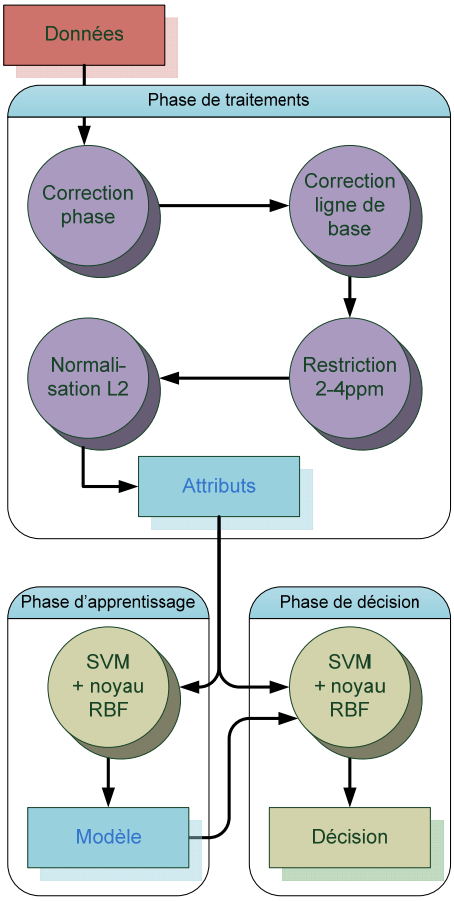


Figure 92 : Schéma de classification optimale.

Grace à cette méthodologie, nous obtenons des résultats comparables à l'état de l'art. Nous envisageons dorénavant d'améliorer le système en associant aux données de SRM, l'imagerie T2, la diffusion et la perfusion. Nous espérons ainsi pourvoir conjuguer la spécificité de la SRM avec la précision spatiale et la sensibilité des trois autres techniques d'imagerie (cf. Figure 93). Nous espérons ainsi pouvoir réaliser une détection automatique de la prostate (les voxels intra-prostatiques sont actuellement définis manuellement) et obtenir une cartographie paramétrique de la tumeur.

Il sera également certainement profitable d'intégrer une analyse prenant en compte la connexité des voxels. En effet, un voxel tumoral est théoriquement entouré de voxels plus ou moins tumoraux. Des algorithmes utilisés en segmentation 3D pourraient alors être appliqués.

Nous avons également défini un algorithme de détection des pics de choline, créatine, spermine et citrate dans le cas bien particulier des spectres de RMN du proton acquis avec un temps d'écho de 140ms, permettant d'obtenir un triplet pour le citrate. Cet algorithme est légèrement plus efficace que LCModel pour la détection des pics de choline

dans les spectres issus de tumeur et se révèle équivalent dans les autres cas. Malgré cette supériorité, l'expérience nous a montré que la classification portant sur des spectres complets était plus performante. La méthode pourra néanmoins fournir au praticien une information sur les concentrations que le praticien pourra utiliser pour affiner son propre diagnostic.

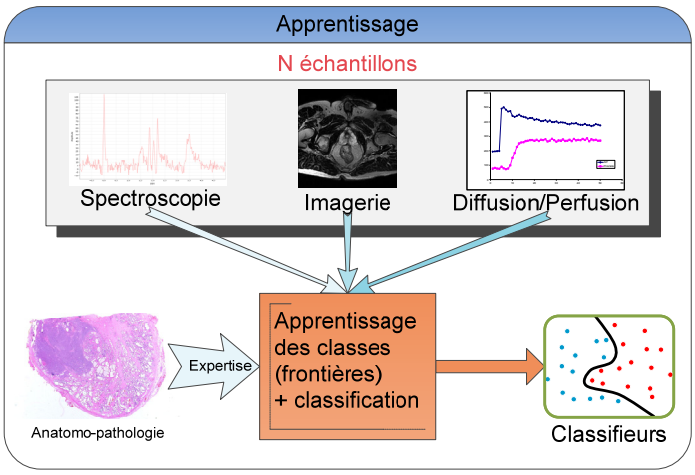


Figure 93 : Schéma de traitement optimal envisagé.

Le deuxième volet de cette thèse concernait la recherche d'éventuels nouveaux biomarqueurs de l'activité tumorale. Les différentes techniques employées ne nous ont permis de détecter, ni molécules plus discriminantes que la choline et le citrate, ni zones plus significatives que la zone regroupant ces deux métabolites. Cependant il semblerait que certains métabolites, même s'ils ne permettent pas d'obtenir de meilleurs taux de classification que l'utilisation de la zone 2-4ppm, influent sur la classification lorsqu'ils sont utilisés. On peut ainsi noter le myo-inositol ou tout autre métabolite ayant une fréquence de résonance proche.

Les lipides ressortent aussi lors des tests de sélection de paramètres, et ce, à la fois en clinique et en pré-clinique. Cependant dans le cas clinique, la séquence d'acquisition utilise une technique de suppression des graisses (Méthode CHESS), donc on ne peut avoir aucune certitude quant à la validité de ces données et il est inenvisageable de ne pas utiliser cette technique, car sans elle, les spectres seraient "pollués" par des pics d'eau et de lipides de forte intensité et les autres pics deviendraient difficilement détectables. Dans le cas pré-clinique, on observe également une variation des lipides qui peut sembler liée à la présence de tumeur, cependant des données obtenues sur trois rats ne peuvent suffire à généraliser.

Enfin des calculs d'estimations absolues de concentrations [102] semblent indiquer que l'inversion d'amplitude des pics de citrate et de choline n'est pas due à une hausse

importante de la choline comme nous le pensions au début de cette thèse mais à une inhibition du métabolisme de la prostate. Plus aucun métabolite n'apparait alors entrainant une hausse relative du pic de choline. Nous avons observé depuis que les spectres tumoraux présentaient généralement un rapport signal sur bruit plus faible. Cela explique en partie pourquoi les classifieurs ont beaucoup de difficultés à séparer les spectres pathologiques des spectres dit "indéterminés".

Ce travail ouvre donc de nombreuses perspectives, parmi lesquelles nous pouvons citer les possibilités suivantes :

- Obtenir un indice de qualité sur la spectroscopie pour d'une part améliorer encore l'acquisition (notamment dans le cas pré clinique) et réussir à éliminer les spectres de mauvaise qualité (certains spectres indéterminés).
- Réaliser de la spectroscopie 3D en préclinique.
- Réussir une comparaison entre les images des coupes histologiques et de l'IRM pour obtenir une vérité de terrain qui ne serait pas issue de l'observation des spectres eux-mêmes.
- Effectuer une classification multi-modalité (spectroscopie, perfusion, diffusion, imagerie). Nous avons déjà imaginé un tel système, mais il reste à définir les poids à attribuer à chaque modalité et la manière de gérer les différences de résolution spatiale.
- Définir un système de classification comportant plus de classes, car, dans le cas des grosses tumeurs (volume supérieur à celui de plusieurs voxels) nous sommes obligés de prendre une décision pour les voxels qui sont "à cheval" entre la tumeur et la zone saine, ce qui manque de précision. D'autre part, le cancer de la prostate se présente parfois sous forme de micro-tumeurs disséminées dans la prostate : il faudrait donc réussir à caractériser ce cas.

Nous espérons donc pouvoir évaluer prochainement certaines de ces perspectives, que les résultats obtenus pourront être validés sur des bases plus larges, et que finalement la classification automatique pourra apporter une aide significative au diagnostic et à l'évaluation des traitements anticancéreux.

Références

1. Ackerstaff E., Pflug B.R., Nelson J.B. *et al.*, ***Detection of Increased Choline Compounds with Proton Nuclear Magnetic Resonance Spectroscopy Subsequent to Malignant Transformation of Human Prostatic Epithelial Cells***. Cancer Res, 2001. 61(9): p. 3599-3603.

2. Bourne R., Katelaris P., Danieletto S. *et al.*, ***Detection of prostate cancer by magnetic resonance imaging and spectroscopy in vivo***. ANZ J Surg, 2003. 73(8): p. 666-8.

3. Costello L.C. and Franklin R.B., ***Citrate metabolism of normal and malignant prostate epithelial cells***. Urology, 1997. 50(1): p. 3-12.

4. Kurhanewicz J., Vigneron D.B., Hricak H. *et al.*, ***Three-dimensional H-1 MR spectroscopic imaging of the in situ human prostate with high (0.24-0.7-cm3) spatial resolution***. Radiology, 1996. 198(3): p. 795-805.

5. Kurhanewicz J., Vigneron D.B., Hricak H. *et al.*, ***Prostate cancer: metabolic response to cryosurgery as detected with 3D H-1 MR spectroscopic imaging***. Radiology, 1996. 200(2): p. 489-496.

6. McNeal J.E., ***The zonal anatomy of the prostate***. Prostate, 1981. 2(1): p. 35-49.

7. Steyn J.H. and Smith F.W., ***Nuclear magnetic resonance imaging of the prostate***. Br J Urol., 1982. 54(6): p. 726-8.

8. Hricak H., William R.D., Spring D.B. *et al.*, ***Anatomy and pathology of the male pelvis by magnetic resonance imaging***. AJR Am J Roentgenol., 1983. 141(6): p. 1101-10.

9. Poon P.Y., McCallum R.W., Henkelman M.M. *et al.*, ***Magnetic resonance imaging of the prostate***. Radiology, 1985. 154(1): p. 143-9.

10. Smith F.W., Cherryman G.R. and Steyn J.H., ***A comparison of T1 measurements at 1.7 and 3.4 MHz in the diagnosis of prostatic carcinoma***. Magn Reson Med, 1985. 2(4): p. 350-4.

11. Sillerud L.O., Halliday K.R., Griffey R.H. *et al.*, ***In vivo 13C NMR spectroscopy of the human prostate***. Magn Reson Med, 1988. 8(2): p. 224-30.

12. Lynch M.J. and Nicholson J.K., ***Proton MRS of human prostatic fluid: correlations between citrate, spermine, and myo-inositol levels and changes with disease***. The Prostate, 1997. 30(4): p. 248-55.

13. Garcìa-Segura J.M., Sánchez-Chapado M., Ibarburen C. *et al.*, ***In vivo proton magnetic resonance spectroscopy of diseased prostate: spectroscopic features of malignant versus benign pathology***. Magnetic Resonance Imaging, 1999. 17(5): p. 755-765.

14. Wang X.Z., Wang B., Gao Z.Q. *et al.*, ***1H-MRSI of prostate cancer: the relationship between metabolite ratio and tumor proliferation***. Eur J Radiol. 73(2): p. 345-51.

15. Nagarajan R., Gomez A.M., Raman S.S. *et al.*, ***Correlation of endorectal 2D JPRESS findings with pathological Gleason scores in prostate cancer patients***. NMR Biomed, 2010. 23(3): p. 257-61.

16. Balique J., Cuilleron M. and Veyret C., ***Spectro-IRM de prostate avec antennes externes : faisabilité en pratique clinique***. J Radiol, 2010. 91(2): p. 207-12.

17. Schmuecking M., Boltze C., Geyer H. *et al.*, ***Dynamic MRI and CAD vs. choline MRS: where is the detection level for a lesion characterisation in prostate cancer?*** Int J Radiat Biol, 2009. 85(9): p. 814-24.

18. Tessem M.-B., R. K.K., D. J. *et al.*, ***Lactate and alanine as metabolic biomarkers for prostate cancer in TRUS- guided biopsies measured by ^{1}H HR-MAS spectroscopy***. 2007.

19. INTERPRET. 2002 [cited 2010 20 Septembre]; Available from: http://gabrmn.uab.es/interpret/.

20. Tate A.R., Griffiths J.R., Howe F.A. *et al.*, ***Differentiating types of human brain tumours by MRS. A comparison of pre-processing methods and echo times***. in *ISMRM01*. 2001. Glasgow, Scotland.

21. de Graaf A.A., van Dijk J.E. and Bovee W.M., ***QUALITY: quantification improvement by converting lineshapes to the Lorentzian type***. Magn Reson Med, 1990. 13(3): p. 343-57.

22. Tate A.R., Griffiths J.R., Martinez-Perez I. *et al.*, ***Towards a method for automated classification of 1H MRS spectra from brain tumours***. NMR Biomed, 1998. 11(4-5): p. 177-91.

23. Bishop C.M., ***Pattern Recognition and Machine Learning***. Information Science and Statistics, ed. Jordan M., Kleinberg J., and Schölkopf B. Vol. 1. 2007, Springer. 738.

24. Tate A.R., Majós C., Moreno A. *et al.*, ***Automated classification of short echo time in in vivo ^{1}H brain tumor spectra: A multicenter study***. Magnetic Resonance in Medicine, 2003. 49(1): p. 29-36.

25. Pijnappel W.W.F., van den Boogaart A., de Beer R. *et al.*, ***SVD-based quantification of magnetic resonance signals***. Journal of Magnetic Resonance (1969), 1992. 97(1): p. 122-134.

26. Devos A., Lukas L., Suykens J.A.K. *et al.*, ***Classification of brain tumours using short echo time ^{1}H MR spectra***. Journal of Magnetic Resonance, 2004. 170(1): p. 164-175.

27. Devos A., Simonetti A.W., van der Graaf M. *et al.*, ***The use of multivariate MR imaging intensities versus metabolic data from MR spectroscopic imaging for brain tumour classification***. Journal of Magnetic Resonance, 2005. 173(2): p. 218-228.

28. Lukas L., Devos A., Suykens J.A.K. *et al.*, ***Brain tumor classification based on long echo proton MRS signals***. Artificial Intelligence In Medicine, 2004. 31(1): p. 73-89.

29. Vapnik V., ***The Nature of Statistical Learning Theory***. Springer-Verlag ed., ed. Springer-Verlag. 1995, Springer-Verlag.

30. Klose U., ***In vivo proton spectroscopy in presence of eddy currents***. Magn Reson Med, 1990. 14(1): p. 26-30.

31. Kelm B.M., Menze B.H., Zechmann C.M. *et al.*, ***Automated estimation of tumor probability in prostate magnetic resonance spectroscopic imaging: Pattern recognition vs quantification***. Magnetic Resonance in Medicine, 2007. 57(1): p. 150-159.

32. Noworolski S., Henry R., Vigneron D. *et al.*, ***Dynamic contrast-enhanced MRI in normal and abnormal prostate tissues as defined by biopsy, MRI, and 3D MRSI***. Magnetic Resonance in Medicine, 2005. 53(2): p. 249-255.

33. Swindle P., McCredie S., Russell P. *et al.*, ***Pathologic characterization of human prostate tissue with proton MR spectroscopy***. Radiology, 2003. 228(1): p. 144-51.

34. Marx and B D., ***Iteratively reweighted partial least squares estimation for generalized linear regression***. Vol. 38. 1996, Alexandria, VA, ETATS-UNIS: American Society for Quality and the American Statistical Association.

35. Marx, B D., Eilers *et al.*, ***Generalized linear regression on sampled signals and curves : A P-spline approach***. Vol. 41. 1999, Alexandria, VA, ETATS-UNIS: American Society for Quality and the American Statistical Association.

36. Breiman L., ***Random Forests***. in *Machine Learning*. 2001. p. 5-32.

37. Ratiney H., Sdika M., Coenradie Y. *et al.*, ***Time-domain semi-parametric estimation based on a metabolite basis set***. NMR Biomed, 2005. 18(1): p. 1-13.

38. vanhamme L., van den Boogaart A. and Van Huffel S., ***Improved Method for Accurate and Efficient Quantification of MRS Data with Use of Prior Knowledge***. 1997.

39. van der Veen J.W., de Beer R., Luyten P.R. *et al.*, ***Accurate quantification of in vivo 31P NMR signals using the variable projection method and prior knowledge***. Magn Reson Med, 1988. 6(1): p. 92-8.

40. Stefan D. and et al., ***Quantitation of magnetic resonance spectroscopy signals: the jMRUI software package***. 2009. p. 104035.

41. Wold S., Ruhe A., Wold H. *et al.*, ***The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses***. SIAM Journal on Scientific and Statistical Computing, 1984. 5(3): p. 735-743.

42. Hyvärinen A. and Oja E., ***Independent Component Analysis : Algorithms and Applications***. Neural Networks, 2000. 13(4-5): p. 411-430.

43. Sajda P., Du S., Brown T.R. *et al.*, ***Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain***. IEEE Trans Med Imaging, 2004. 23(12): p. 1453-65.

44. Matulewicz L., Zakian K.L., Shukla-Dave A. *et al.*, ***A pattern recognition model for automatic classification of 1H MRSI voxels in the prostate***. in *ISMRM 2009*. 2009. Honolulu, Hawaii, USA.

45. Wold S., Antti H., Lindgren F. *et al.*, ***Orthogonal signal correction of near-infrared spectra***. Chemometrics and Intelligent Laboratory Systems, 1998. 44(1-2): p. 175-185.

46. Tiwari P., Madabhushi A. and Rosen M., ***A hierarchical unsupervised spectral clustering scheme for detection of prostate cancer from magnetic resonance spectroscopy (MRS)***. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv, 2007. 10(Pt 2): p. 278-86.

47. Tiwari P., Rosen M. and Madabhushi A., ***Consensus-locally linear embedding (C-LLE): application to prostate cancer detection on magnetic resonance spectroscopy***. Med Image Comput Comput Assist Interv, 2008. 11(Pt 2): p. 330-8.

48. Tiwari P., Rosen M. and Madabhushi A., ***A hierarchical spectral clustering and nonlinear dimensionality reduction scheme for detection of prostate cancer from magnetic resonance spectroscopy (MRS)***. Medical Physics, 2009. 36(9): p. 3927-3939.

49. Madabhushi A., Feldman M.D., Metaxas D.N. *et al.*, ***Automated detection of prostatic adenocarcinoma from high-resolution ex vivo MRI***. Ieee Transactions on Medical Imaging, 2005. 24(12): p. 1611-1625.

50. Macqueen J.B., ***Some methods of classification and analysis of multivariate observations***. in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. 1967.

51. Shi J. and Malik J., ***Normalized cuts and image segmentation***. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. 22(8): p. 888-905.

52. Roweis S.T. and Saul L.K., ***Nonlinear dimensionality reduction by locally linear embedding***. Science, 2000. 290(5500): p. 2323-+.

53. Zakian K.L., Sircar K., Hricak H. *et al.*, ***Correlation of proton MR spectroscopic imaging with gleason score based on step-section pathologic analysis after radical prostatectomy***. Radiology, 2005. 234(3): p. 804-14.

54. Lauterbur P.C., ***Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance***. Nature, 1973. 242(5394): p. 190-191.

55. Scheenen T.W.J., Gambarota G., Weiland E. *et al.*, ***Optimal timing for in vivo 1H-MR spectroscopic imaging of the human prostate at 3T***. 2005. p. 1268-1274.

56. Jung J.A., Coakley F.V., Vigneron D.B. *et al.*, ***Prostate Depiction at Endorectal MR Spectroscopic Imaging: Investigation of a Standardized Evaluation System***. Radiology, 2004. 233(3): p. 701-708.

57. Antoniadis A., Bigot J. and S. L.-L., ***Peaks detection and alignment for mass spectrometry data***. Journal de la Société Française de Statistique, 2010. 151(1): p. 17-37.

58. Provencher S.W., ***Estimation of metabolite concentrations from localized in vivo proton NMR spectra***. Magnetic Resonance in Medicine, 1993. 30(6): p. 672-9.

59. Neff B.L., Ackerman J.L. and Waugh J.S., ***Fully automatic software correction of fourier transform NMR spectra***. Journal of Magnetic Resonance (1969), 1977. 25(2): p. 335-340.

60. Marshall A.G., ***Dispersion vs. absorption (DISPA): A magic circle for spectroscopic line shape analysis***. Chemometrics and Intelligent Laboratory Systems, 1988. 3(4): p. 261-275.

61. Marshall A.G. and D.C. R., ***Dispersion versus Absorption:Spectral Line Shape Analysis for Radiofrequencyand Microwave Spectrometry***. Anal Chem, 1978. 50: p. 756-763.

62. Brown D.E., Campbell T.W. and Moore R.N., ***Automated phase correction of FT NMR spectra by baseline optimization***. Journal of Magnetic Resonance (1969), 1989. 85(1): p. 15-23.

63. van Vaals J.J. and van Gerwen P.H.J., ***Novel methods for automatic phase correction of NMR spectra***. Journal of Magnetic Resonance (1969), 1990. 86(1): p. 127-147.

64. Heuer A., ***A new algorithm for automatic phase correction by symmetrizing lines***. Journal of Magnetic Resonance (1969), 1991. 91(2): p. 241-253.

65. Chen L., Weng Z., Goh L. *et al.*, ***An efficient algorithm for automatic phase correction of NMR spectra based on entropy minimization***. Journal of Magnetic Resonance, 2002. 158(1-2): p. 164-168.

66. Larry Bretthorst G., ***Automatic phasing of MR images. Part I: Linearly varying phase***. Journal of Magnetic Resonance, 2008. 191(2): p. 184-192.

67. Shannon C.E., ***A Mathematical Theory of Communication***. A Mathematical Theory of Communication, 1948. 27: p. 379-423,623-656.

68. Lagarias J.C., Reeds J.A., Wright M.H. *et al.*, ***Convergence Properties of the Nelder--Mead Simplex Method in Low Dimensions***. SIAM Journal on Optimization, 1998. 9(1): p. 112-147.

69. Nelder J.A. and Mead R., ***A simplex method for function minimization***. Computer Journal, 1965. 7: p. 308-313.

70. Mazet V., ***Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies***, in *Automatique, traitement du signal et génie informatique*. 2005, Henry Poincaré, Nancy 1: Nancy. p. 164.

71. Hofmann L., Slotboom J., Jung B. *et al.*, ***Quantitative 1H-magnetic resonance spectroscopy of human brain: Influence of composition and parameterization of the basis set in linear combination model-fitting***. Magnetic Resonance in Medicine, 2002. 48(3): p. 440-453.

72. Bartha R., Drost D.J. and Williamson P.C., ***Factors affecting the quantification of short echo in-vivo 1H MR spectra: prior knowledge, peak elimination, and filtering***. NMR Biomed, 1999. 12(4): p. 205-16.

73. Lieber C.A. and Mahadevan-Jansen A., ***Automated method for subtraction of fluorescence from biological Raman spectra***. Appl Spectrosc, 2003. 57(11): p. 1363-7.

74. Provencher S.W., ***Estimation of metabolite concentrations from localized in vivo proton NMR spectra***. Magn Reson Med, 1993. 30(6): p. 672-9.

75. Marshall I., Higinbotham J., Bruce S. *et al.*, ***Use of Voigt lineshape for quantification of in vivo 1H spectra***. Magn Reson Med, 1997. 37(5): p. 651-7.

76. Duda R.O., Hart P.E. and Stork D.G., ***Pattern Classification***. second ed. Vol. 1. 2001, John Wiley & Sons. 654.

77. Aizerman A., Braverman E.M. and Rozoner L.I., ***Theoretical foundations of the potential function method in pattern recognition learning***. Automation and Remote Control, 1964. 25: p. 821-837.

78. Guigue V., Rakotomamonjy A. and Canu S., ***SVM et k-ppv pour la reconnaissance d'émotions***. in *19° Colloque sur le traitement du signal et des images*. 2003. France: GRETSI, Groupe d'Etudes du Traitement du Signal et des Images.

79. Bakken I.J., Axelson D., Kvistad K.A. *et al.*, ***Applications of neural network analyses to in vivo 1H magnetic resonance spectroscopy of epilepsy patients***. Epilepsy Research, 1999. 35(3): p. 245-252.

80. Chaplot S., Patnaik L.M. and Jagannathan N.R., ***Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network***. Biomedical Signal Processing and Control, 2006. 1(1): p. 86-92.

81. Hagberg G., ***From magnetic resonance spectroscopy to classification of tumors. A review of pattern recognition methods***. NMR Biomed, 1998. 11(4-5): p. 148-56.

82. Hilario M., Kalousis A., Muller M. *et al.*, ***Machine learning approaches to lung cancer prediction from mass spectra***. Proteomics, 2003. 3(9): p. 1716-9.

83. Hiltunen Y., Kaartinen J., Pulkkinen J. *et al.*, ***Quantification of Human Brain Metabolites from in Vivo1H NMR Magnitude Spectra Using Automated Artificial Neural Network Analysis***. Journal of Magnetic Resonance, 2002. 154(1): p. 1-5.

84. Lindon J.C., Holmes E. and Nicholson J.K., ***Pattern recognition methods and applications in biomedical magnetic resonance***. Progress in Nuclear Magnetic Resonance Spectroscopy, 2001. 39(1): p. 1-40.

85. Hudson B.D., Whitley D.C., Browne A. *et al.*, ***Extraction of comprehensible logical rules from neural networks. application of TREPAN in bio and chemoinformatics***. Croatica chemica acta, 2005. 78(4): p. 557-561.

86. Kaartinen J., Mierisova S., Oja J.M.E. *et al.*, ***Automated Quantification of Human Brain Metabolites by Artificial Neural Network Analysis fromin VivoSingle-Voxel1H NMR Spectra***. Journal of Magnetic Resonance, 1998. 134(1): p. 176-179.

87. Sordo M., Buxton H. and Watson D., ***A hybrid approach to breast cancer diagnosis***. Practical applications of computational intelligence techniques. 2001: Kluwer Academic Publishers. 299-330.

88. Valenzuela O., Rojas I., Rojas F. *et al.*, ***Automatic classification of prostate cancer using pseudo-gaussian radial basis function neural network***. in *ESANN 2005*. 2005. Bruges, Belgium.

89. Ye C.Z., Yang J., Geng D.Y. *et al.*, ***Fuzzy rules to predict degree of malignancy in brain glioma***. Med Biol Eng Comput, 2002. 40(2): p. 145-52.

90. McCulloch W.S. and Pitts W., ***A logical calculus of ideas immanent in nervous activity***. Bulletin of Mathematical Biophysics, 1943. 5: p. 115-133.

91. Rosenblatt F., ***Principles of Neurodynamics : perceptrons and the theory of brain mechanisms*** 1962, Spartan Books, Washington DC.

92. Rumelhart D. and McClelland J., ***Parallel distributed Processing***. Vol. 1. 1986, Cambridge, MA: MIT Press.

93. Irie B. and Miyake S., ***Capabilities of Three-Layer Perceptrons***. in *IEEE Int. Conf. Neural Networks*. 1988.

94. Chang C.-C. and Lin C.-J., ***LIBSVM : a library for support vector machines***, in *Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm*. 2001.

95. Smach F., Miteran J., Atri M. *et al.*, ***An FPGA-based accelerator for Fourier Descriptors computing for color object recognition using SVM***. Journal of Real-Time Image Processing, 2007. 2(4): p. 249-258.

96. Jain A. and Zongker D., ***Feature selection: Evaluation, application, and small sample performance***. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997. 19(2): p. 153-158.

97. Journaux L., Destain M.-F., Miteran J. *et al.*, ***Texture Classification with Generalized Fourier Descriptors in Dimensionality Reduction Context: An Overview Exploration***. in *Artificial Neural Networks in Pattern Recognition*. 2008, Springer Berlin / Heidelberg. p. 280-291.

98. Lu Y., Cohen I., Zhou S.X. *et al.*, ***Feature Selection Using Principal Feature Analysis***. in *ACM Multimedia*. 2007. Augsburg, Germany.

99. Walker P., Tizon X., Parfait S. *et al.*, ***Evaluation of Early Response to Temozolomide and Radiotherapy with Magnetic Resonance Imaging and Proton Magnetic Resonance Spectroscopy in Human Glioma Models in Nude Rats.*** in *AACR*. 2008. Chicago.

100. Provent P., Walker P., Tizon X. *et al.*, ***Prostate cancer metabolism assessment by proton magnetic resonance spectroscopy (1H-MRS) and high resolution magic angle spinning spectroscopy (1H-HRMAS) in experimental rat models***. in *AACR 2009*. 2009. Denver.

101. Tizon X., Provent P., Parfait S. *et al.*, ***In vivo characterisation of orthotopic prostate tumor and healthy rat prostate metabolism using 1H-MRS at 4.7T***. in *ISMRM 2010*. 2010. Stockholm, Sweden.

102. Walker P., Créhange G., Parfait S. *et al.*, ***Absolute Quantification in 1H MRSI of the Prostate at 3T***. in *ISMRM 2010*. 2010. Stockholm, Sweden.

Annexes

Annexe 1

Classification TNM pour le cancer de la prostate (Réseau régional de cancérologie de Lorraine : http://www.oncolor.org/referentiels/genital_masculin/prost_class.htm)

Tableau 8 : Tumeur primitive (T)

Tx		Non évalué
T0		Absence de Tumeur
T1		Tumeur non palpable et non visible à l'imagerie
	Tx	Tumeur primitive non évaluée
	T0	Tumeur primitive non retrouvée
	T1a	≤ 5 % des copeaux de RTUG[3] pour HBP
	T1b	> 5 % des copeaux de RTUG pour HBP
	T1c	Découverte par biopsie en raison d'une augmentation de PSA
NB	*Le grade de Gleason n'est pas pris en compte dans la classification de l'UICC. la classification de l'American Joint Committee intègre le score de Gleason pour différencier le T1a du T1b.*	
	T1a	*< 5% du tissu réséqué avec un score de Gleason < 7 ou absence de grade 4 ou 5*
	T1b	*> 5% du tissu réséqué et/ou score de Gleason > 7 ou présence de grade 4 ou 5*
T2		Tumeur limitée à la prostate (apex et capsule compris)
	T2a	Atteinte de la moitié d'un lobe ou moins
	T2b	Atteinte de plus de la moitié d'un lobe mais sans atteindre les 2 lobes
	T2c	Atteinte des deux lobes
T3		Extension au-delà de la capsule
	T3a	Extension extra-capsulaire uni ou bilatérale
	T3b	Extension aux vésicules séminales uni ou bilatérale
T4		Extension aux organes adjacents (col vésical, sphincter, rectum, paroi pelvienne) ou tumeur fixée

[3] Résection Trans-Urétrale de la Glande.

Tableau 9 : Ganglions régionaux (N)

Nx	Ganglions régionaux non évalués
N0	Absence de métastase ganglionnaire
N1	Atteinte ganglionnaire régionale

Tableau 10 : Métastases (M)

Mx	Métastases à distance non évaluées
M0	Absence de métastase à distance
M1	Métastase à distance
M1a	Ganglions non régionaux
M1b	Os
M1c	Autres sites avec ou sans atteinte osseuse

Annexe 2

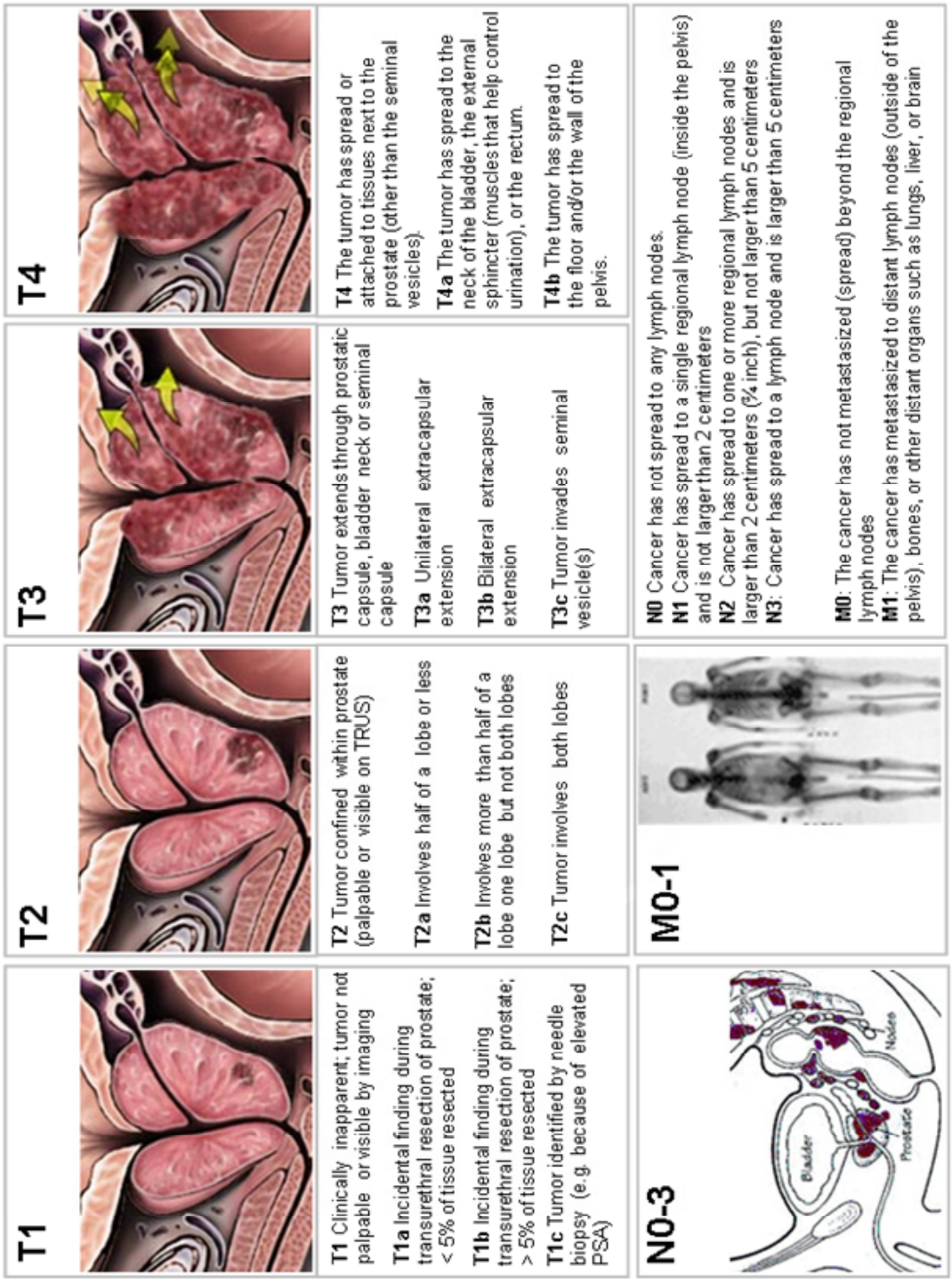


Image issue du serveur de document de l'université de Berlin : http://edoc.hu-berlin.de/dissertationen/kaiser-simone-2004-06-10/HTML/kaiser2_html_7fe064af.png

Annexe 3

Algorithme de sélection des pics parmi les fonctions composant le spectre.

P_{th} est la position de référence du pic considéré.

Voici la liste des 6 règles utilisées dans les algorithmes suivants :

(1) Le décalage moyen est de 0.21 ppm et le médian de 0.15 ppm donc en faisant une première recherche dans une plage de 0,2 ppm, nous pouvons corriger plus de la moitié des cas.

(2) Chaque spectre de notre ensemble de données contient au moins un pic correspondant soit à la choline soit aux pics de citrate.

(3) La distance entre les pics ne présente que peu de variation.

(4) Si tous les pics ont une amplitude négligeable par rapport à un autre et que ce dernier se situe dans la zone 3 ppm - 3,4 ppm, on le considère comme de la choline.

(5) Les pics d'amplitude plus forte sont considérés comme les plus fiables et servent de base à la correction du décalage.

(6) Les spectres pathologiques possèdent au moins un pic de choline et les spectres sains présentent un trident caractéristique du citrate. Dans les spectres de transition, on doit détecter tous les métabolites.

Première étape :

Dans cette partie, il s'agit de faire une détection grossière des pics, l'essentiel étant de sélectionner des pics majeurs pour effectuer un recalage en fréquence du spectre. Pour cela, on affecte à chaque métabolite à détecter, le pic d'amplitude la plus importante dans la plage $[P_{th}-0,2..P_{th}+0,1]$ (règles 1 et 6). On obtient une liste de 6 pics notée Lp.

Tant que cette liste contient des doublons, c'est-à-dire qu'un pic est affecté à deux métabolites, le pic concerné est associé au métabolite le plus proche en fréquence et le second pic en amplitude de la plage $[P_{th}-0,2..P_{th}+0,1]$ est affecté à l'autre métabolite.

La liste Lp contient donc pour chaque métabolite les paramètres d'une fonction modélisant le pic correspondant et ces pics sont tous différents.

A partir de cette liste, on estime une première correction de décalage à réaliser.

A partir de la règle 4, si un pic associé à l'un des métabolites choline/spermine/créatine, est très supérieur en amplitude aux pics de citrate, on considère ce pic comme étant la choline et on corrige les fréquences du spectre en fonction.

Si les deux groupes de pics (choline, créatine, spermine d'un coté et les 3 pics de citrate de l'autre) sont équivalents en amplitude, on décale le spectre du médian des décalages mesurés.

Si les trois pics de citrate sont dominants en amplitude, la valeur de la correction du décalage est le médian des décalages des pics de citrate sinon cette correction de décalage sera le médian des décalages mesurés sur les pics de choline, spermine et créatine.

Cette recherche sur une plage relativement large nous permet de corriger des décalages importants. Une seconde étape basée sur le même principe (approximation du spectre, sélection des pics correspondant aux métabolites et estimation du décalage) peut être réalisée.

Deuxième étape :

Après avoir corrigé le décalage avec la valeur trouvée précédemment, une nouvelle approximation du spectre par une somme de fonctions est effectuée : il en ressort une nouvelle liste de paramètres Lp.

Parmi cette liste, pour chaque métabolite, la liste des pics dont la fréquence est dans une plage de taille 0.1 ppm autour de la position de résonance théorique du métabolite est sélectionnée. Nous obtenons donc une liste de pics pour chaque métabolite. La liste contenant toutes les combinaisons de 6 pics possibles parmi cette présélection est alors générée et notée Lc.

Pour chaque combinaison de Lc, si un pic est affecté à deux métabolites, tous les paramètres correspondant au métabolite le plus éloigné en fréquence de ce pic sont passés à zéro.

L'algorithme sélectionne alors la combinaison de pics qui maximise la hauteur des pics, tout en minimisant la différence quadratique entre la distance inter-pics théorique et la distance inter-pics de la combinaison (règle 3).

Nous disposons alors d'une nouvelle liste de six pics correspondant aux quatre métabolites recherchés.

Si à la fin de la première étape, le décalage a été estimé sur la choline en se basant sur la règle 4 alors on considère la recherche terminée : on retourne la liste de pic parmi Lc et le décalage estimé.

Sinon, la somme des amplitudes des pics de choline, créatine et spermine est calculée, ainsi que la somme des amplitudes des trois pics de citrate. Si la somme des trois pics de citrate est la plus élevée, l'estimation du nouveau décalage est réalisée en calculant la moyenne des trois fréquences de ces pics, sinon c'est la moyenne des fréquences des pics de choline, spermine et créatine qui est prise en compte.

Cette étape a permis d'affiner le décalage en fréquence, une troisième étape est effectuée si nécessaire, c'est-à-dire si la première étape n'a pas détecté un pic prééminent de choline.

Troisième étape :

Le nouveau décalage est appliqué au spectre et l'approximation de celui-ci par une somme de fonctions est ré-estimée.

Une sélection des pics correspondant aux métabolites est effectuée sur le même principe que pour l'étape 2. La liste des pics sélectionnés est retournée et le décalage est estimé en faisant la moyenne des différences entre les fréquences des pics sélectionnés et les fréquences théoriques de chaque pic.