



HAL
open science

Analyse statistique d'images médicales : étude et utilisation du logiciel SPM

Marie-Laure Tabary, Nicholas Ayache, Jacques Darcourt, Grégoire Malandain

► **To cite this version:**

Marie-Laure Tabary, Nicholas Ayache, Jacques Darcourt, Grégoire Malandain. Analyse statistique d'images médicales : étude et utilisation du logiciel SPM. RR-3802, INRIA. 1999. inria-00072857

HAL Id: inria-00072857

<https://inria.hal.science/inria-00072857>

Submitted on 24 May 2006

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

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



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Analyse statistique d'images médicales : étude et utilisation du logiciel SPM

Marie-Laure Tabary — Nicholas Ayache — Jacques Darcourt — Grégoire Malandain

N° 3802

Novembre 1999

THÈME 3

A large blue rectangular area containing the text 'Rapport de recherche' in a white serif font. To the left of the text is a large, light grey 'R' logo. A horizontal grey brushstroke is positioned below the text.

*Rapport
de recherche*

Analyse statistique d'images médicales : étude et utilisation du logiciel SPM

Marie-Laure Tabary, Nicholas Ayache, Jacques Darcourt, Grégoire Malandain

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Epidaure

Rapport de recherche n° 3802 — Novembre 1999 — 24 pages

Résumé : Les médecins ont besoin de données objectives et quantifiées pour les aider dans le diagnostic de certaines pathologies. L'analyse statistique des images médicales de populations de patients est un moyen de leur fournir ce type de données. Dans ce rapport nous étudions le logiciel SPM, son utilisation sur des images SPECT.

Mots-clés : imagerie médicale 3D, analyse statistique d'images, SPM, imagerie fonctionnelle

Statistical analysis of medical images : using the SPM software

Abstract: Physicians need objective and quantified data to help them in the diagnosis of some diseases. the statistical analysis of 3D medical images is a way to provide them with such data. In this report, we study the software SPM used on SPECT images.

Key-words: 3D medical imaging, statistical image analysis, SPM, functional imaging

1 Introduction

Pour proposer aux médecins un outil diagnostique fondé sur des images médicales, il faut avoir préalablement déterminé comment cette pathologie se manifeste dans les images considérées.

Dans cet objectif nous souhaitons pouvoir comparer les images acquises sur des patients atteints par la maladie à celles de sujets sains, afin de mettre en évidence les différences significatives entre les deux populations et de les quantifier.

2 L'outil statistique SPM

Afin d'effectuer les comparaisons statistiques entre les populations de malades et de sujets sains, nous avons besoin d'un outil statistique adapté. Le simple calcul de la différence entre ces images est généralement insuffisant. En particulier cette méthode est très sensible à une mauvaise résolution ou un faible rapport signal sur bruit (voir l'image de différence figure 1). D'autre part l'exploitation d'une image de différence peut poser des problèmes tout d'abord dans le choix des seuils à appliquer pour déterminer les zones à retenir et également parce qu'elle ne fournit pas de critère de confiance sur les activations obtenues. Il faut donc envisager des outils statistiques plus puissants. Nous avons dans un premier temps choisi d'étudier de façon approfondie le logiciel SPM¹ [FHW⁺95] et la méthode qu'il implémente, étant donné que c'est la plus reconnue des méthodes actuellement utilisées pour l'analyse statistique des images médicales.

SPM a été développé par les membres du département *Wellcome* de Neurologie Cognitive de l'Institut de Neurologie de Londres. L'implémentation a été faite sous Matlab.

2.1 Buts de SPM

SPM a pour objet la comparaison par analyse statistique de deux groupes d'images, afin de déterminer les zones qui diffèrent significativement entre ces deux groupes.

Une des ambitions principales de cet outil est la prise en compte des corrélations spatiales dans les images considérées, afin notamment de s'affranchir autant que possible des effets du bruit.

L'autre objectif important est de parvenir à fournir une mesure de validité des résultats obtenus (en l'occurrence sous la forme d'une probabilité d'erreur sur ces résultats).

2.2 Jeu de données synthétiques

Pour nous aider dans l'étude de ce logiciel, nous avons constitué un jeu de données synthétiques (voir figure 1), dont nous connaissons les régions activées et leurs tailles respectives. Ceci nous a permis de mieux comprendre le comportement de la méthode dans différents cas.

Nous nous sommes inspirés d'une étude faite dans [FWF⁺94], où les caractéristiques de ces images sont décrites comme reflétant celles d'images acquises sur des sujets humains :

- Ce sont des images 2D, 128x128.
- La base correspond à des tirages d'une distribution de Poisson de paramètre 10.
- L'activation appliquée correspond à 6 puits de tailles différentes tirés suivant une distribution de Poisson de paramètre 11. Le fond des images activées correspond toujours à une distribution de Poisson de paramètre 10.
- Toutes ces images sont filtrées par un noyau gaussien, $\sigma = 3$.

1. Statistical Parametric Mapping.

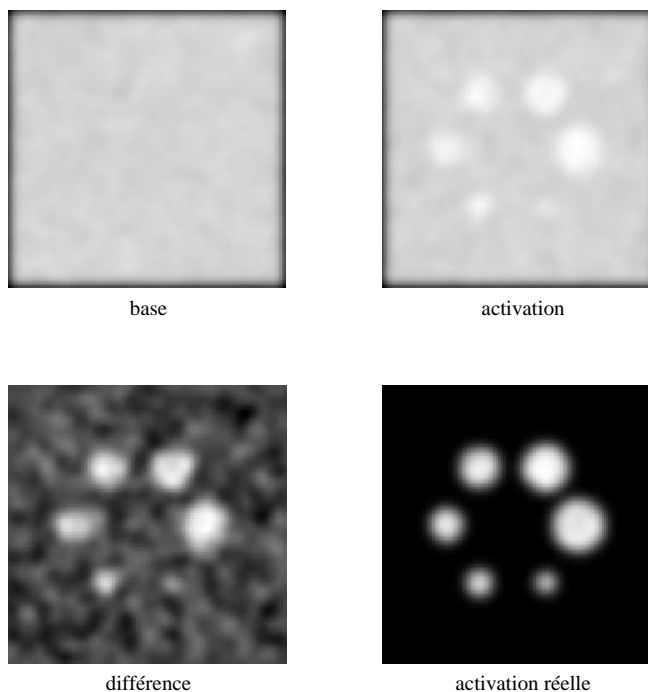


FIG. 1 – *Images synthétiques utilisées*

2.3 Test statistique

La comparaison proprement dite des images se fait par le moyen d'un test statistique effectué voxel par voxel.

Soit S le volume des images (autrement dit le nombre de voxels ou pixels contenus dans les images), on effectue indépendamment S tests statistiques correspondant à chaque position dans les images.

Le test le plus couramment utilisé est le *t-test*, ou *test de Student*.

2.3.1 Caractéristiques du t-test

- **But** : comparer les moyennes de deux distributions
- **Contrainte** : valable pour des distributions gaussiennes de même variance
- **Robustesse** : valable également dans le cas de distributions non gaussiennes sans contrainte sur les variances, à condition que les échantillons considérés soient "assez grands" (au moins 10 mesures par échantillon) [Sap90],[Pap91]

2.3.2 Distribution de Student

Soient :

U variable aléatoire suivant une loi gaussienne $\mathcal{N}(0;1)$

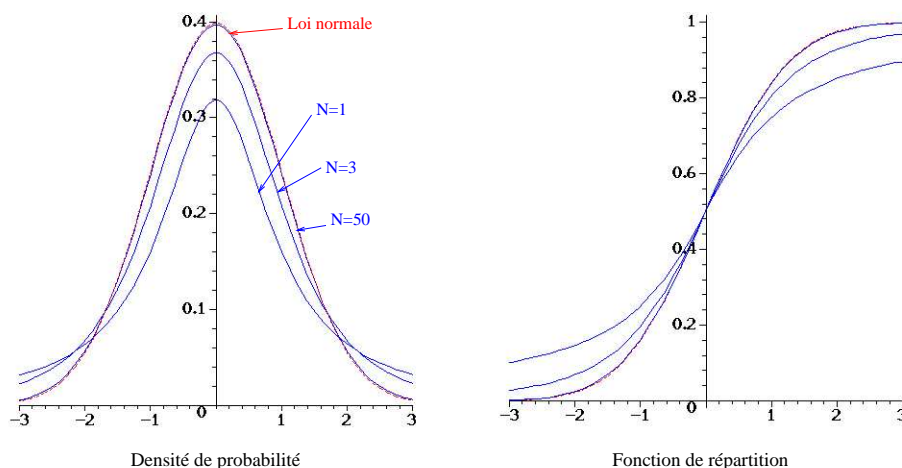
X variable aléatoire suivant une loi du χ^2 à N degrés de liberté

On définit la variable de Student à N degrés de liberté :

$$T_N = \frac{U}{\sqrt{\frac{X}{N}}}$$

La loi de Student à N degrés de liberté est la loi de la variable aléatoire T_N .

Au fur et à mesure que N augmente, la distribution de Student se rapproche de la distribution normale (voir figure 2).

FIG. 2 – *Distribution de Student et fonction de répartition associée*

2.3.3 t-test

Soient $A(N_A, \bar{A}, S_A)$ et $B(N_B, \bar{B}, S_B)$, deux échantillons.

nombres de mesures : N_A, N_B

moyennes : \bar{A}, \bar{B}

écarts types : S_A, S_B

Sous hypothèse nulle H_0 (hypothèse selon laquelle les deux distributions sont identiques), le test de Student (t-test) entre ces deux échantillons vaut :

$$T_{N_A+N_B-2} = \frac{(\bar{A} - \bar{B})}{\sqrt{\frac{(N_A S_A^2 + N_B S_B^2) \left(\frac{1}{N_A} + \frac{1}{N_B} \right)}{\sqrt{N_A + N_B - 2}}}} \quad (1)$$

Il suit une loi de Student à $N_A + N_B - 2$ degrés de liberté.

2.3.4 Cas extrême

Dans le cadre médical il peut être utile, à des fins diagnostiques notamment, de comparer **une image unique** résultant de l'examen d'un seul patient à un groupe d'images de contrôle.

On a vu que pour que le t-test ait un sens, il faut soit que l'on considère deux distributions gaussiennes de même variance, soit que l'on ait des échantillons de taille suffisante.

La situation qui nous intéresse ne répond pas à ces critères, on peut cependant l'analyser, en imposant alors l'hypothèse forte suivante :

La population dont est issu le patient unique a une variance comparable à celle de la population de contrôle

La question qui se pose est donc de savoir si cette hypothèse forte est légitime. Il convient, en collaboration avec les praticiens, de vérifier sa validité suivant la pathologie considérée.

NB: Le t-test consistant en une comparaison des moyennes, il faut aussi être conscient du fait qu'une seule

acquisition (pour le patient unique) constitue une très mauvaise estimation de la moyenne de la distribution correspondante.

2.4 Cartes statistiques

2.4.1 carte-t

Un t-test calculé voxel par voxel donne une image de mêmes dimensions que les images initiales, comportant en chaque voxel le résultat de ce t-test (voir figure 3).

Cette image est la carte statistique sur laquelle l'analyse va maintenant porter, et s'appelle la carte-t (*t-map*).

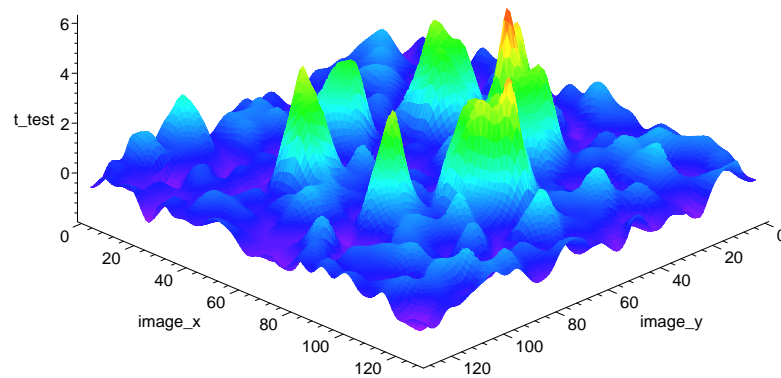


FIG. 3 – *carte-t* calculée sur les données synthétiques

2.4.2 Gaussianisation de la carte-t en carte-z

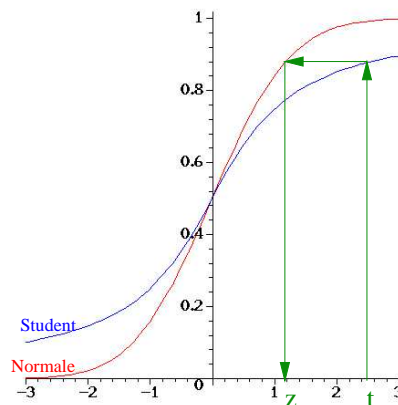


FIG. 4 – *Transformation de t en z*

Pour des raisons de “confort” on modifie voxel par voxel cette carte-t en une carte-z (*z-map*) par la transformation (voir figure 4) :

$$z = \Phi^{-1}(\Psi(t))$$

Φ fonction de répartition pour une distribution normale

Ψ fonction de répartition pour une distribution de Student (avec le bon nombre de degrés de liberté).

On s'assure ainsi qu'au niveau monovarié z suit une distribution normale en chaque voxel.

2.5 Seuils de confiance

2.5.1 Seuillage en intensité de la carte- z

▷ Principe

On est sous hypothèse nulle H_0 (à savoir : les distributions considérées sont identiques), la valeur du t-test est donc celle donnée par l'équation 1.

H_0 implique notamment que les moyennes des deux distributions soient identiques, autrement dit le t-test est censé être nul en tout voxel vérifiant H_0 .

La "gaussianisation" ne change pas cette propriété, en effet Φ est continue, bijective, et $\Phi(0) = 0$.

Ainsi, les points où z est loin de 0 violent H_0 et constituent donc des points où les distributions diffèrent.

▷ Seuillage

La loi normale suivie par z en chaque voxel nous donne pour tout seuil u la probabilité :

$$p(z > u)$$

On se fixe une p_1_value , par exemple $p_1 = 0,01$, on en déduit le seuil u à appliquer à la carte- z , ici 2,33 (voir figure 5).

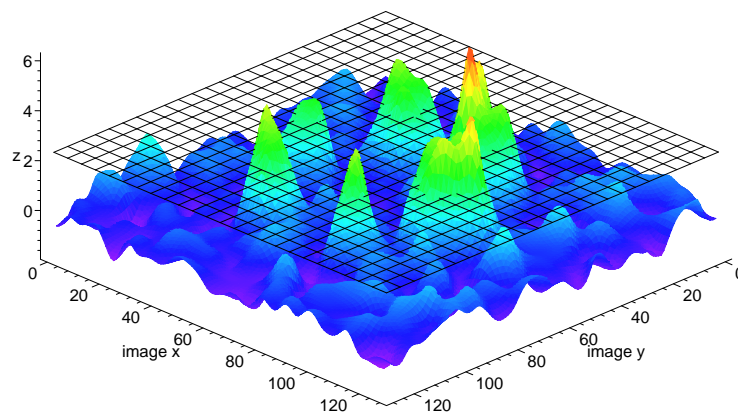


FIG. 5 – *carte- z seuillée à $u = 2.33$, $p_1 = 0.01$*

p_1 est la probabilité de se tromper en disant que les voxels pour lesquels $z > u$ contredisent H_0 .

Par la suite on ne s'intéressera plus qu'aux composantes connexes au dessus du seuil u choisi.

2.5.2 Seuillage de taille des composantes connexes

▷ Champs gaussiens

Grossièrement un champ gaussien est une collection de variables aléatoires multivariées gaussiennes.

Une variable aléatoire X sur \mathbb{R}^n est **multivariée gaussienne** si pour tout n -uplet de réels $(\alpha_1, \dots, \alpha_n)$ la variable $Y = \sum_{i=1}^n \alpha_i X_i$ est gaussienne.

La densité de probabilité de X est alors donnée par :

$$f_X(x) = \frac{1}{\sqrt{2\pi^n |V|}} \exp \left[-\frac{1}{2} (x - \mu)^t V^{-1} (x - \mu) \right]$$

μ vecteur de \mathbb{R}^n de coordonnées $\mu_i = E\{X_i\}$
 V matrice de covariance définie positive, $v_{ij} = E\{(X_i - \mu_i)(X_j - \mu_j)\}$

La théorie des champs gaussiens développée par Adler [Adl81] permet d'analyser plus avant la carte- z et les composantes connexes résultant du seuillage précédent.

▷ principe

On assimile la carte- z à la réalisation d'un champ gaussien (on s'est assurés que c'est valide au niveau monovarié par la transformation de t en z , et cela reste acceptable au niveau multivarié à condition d'avoir suffisamment d'images dans chaque population).

On s'intéresse aux composantes connexes au dessus du seuil u choisi précédemment.

On détermine :

$p(n > k)$ distribution des tailles (n nombre de voxels) des composantes connexes au dessus du seuil u .

En se fixant alors une $p2_value$, on détermine la taille k minimale (en nombre de voxels) pour retenir une composante connexe.

$p2$ est alors la probabilité de se tromper en considérant une activation de taille au moins k comme significative.

▷ Nombre et taille des composantes connexes

On retient trois grandeurs importantes :

- N nombre de voxels au dessus d'un seuil u
- m nombre de régions activées
- n nombre de voxels dans chacune de ces régions

Etant donné un seuillage à $z = u$, en dimension D , pour un volume S , on s'intéresse plus particulièrement à la probabilité pour qu'il existe au moins une composante connexe au dessus du seuil u qui soit de taille au moins k (pour $k \in \mathbb{N}$ nombre de voxels). Cela revient à la probabilité pour que la plus grande composante connexe au dessus du seuil u soit de taille au moins k : $p(n_{max} \geq k)$.

▷ Espérances de N , m et n

Ces trois variables sont liées :

$$E\{N\} = E\{m\} \cdot E\{n\} \tag{2}$$

De la "gaussianisation" de la carte- t en carte- z on tire :

$$E\{N\} = S \cdot \Phi(-u) = \frac{S}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{x^2}{2}} dx \tag{3}$$

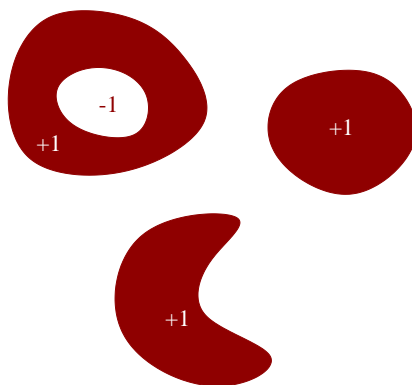


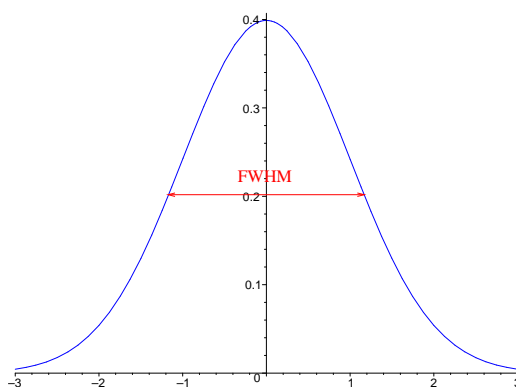
FIG. 6 – caractéristique d'Euler

D'après la théorie des champs gaussiens [PHWF97], on estime $E\{m\}$ par la caractéristique d'Euler χ (voir figure 6) :

$$E\{m\} \approx \chi = S \cdot 2^{-\frac{D+1}{2}} \cdot W^{-D} \cdot u^{D-1} \cdot e^{-\frac{u^2}{2}} \quad (4)$$

χ est en fait le nombre de composantes connexes moins le nombre de trous, au dessus du seuil u . D'où $\chi \leq E\{m\}$ (converge pour $u \rightarrow \infty$)

W est une mesure de régularité liée à la valeur $FWHM$ (full width at half maximum, $FWHM = 2\sigma\sqrt{2\ln 2}$, voir figure 7) de la résolution du champ gaussien [Wor96].

FIG. 7 – $FWHM$

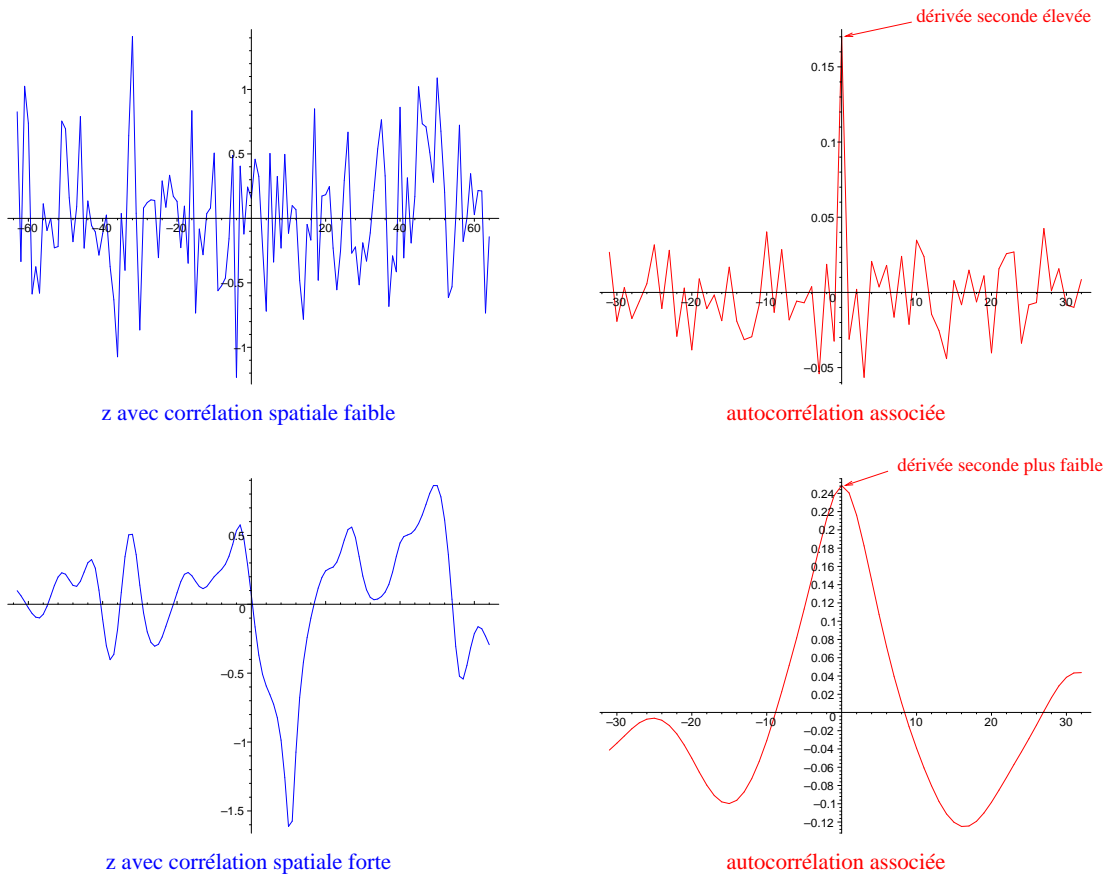
W est défini par :

$$W = \prod_{i=1}^D \{Var z'_i\}^{-1/2D} = \prod_{i=1}^D (-\rho''_i(0))^{-1/2D}$$

Où $\rho = (\rho_1, \dots, \rho_D)$ est l'autocorrélation spatiale de la carte-z.

$$\rho_i(x) = \int_{-\infty}^{+\infty} z_i(t)z_i(x-t)dt.$$

W est le terme qui constitue la prise en compte de la corrélation spatiale du champ gaussien (voir figure 8).

FIG. 8 – z et ρ suivant la corrélation spatiale

En pratique on estime W par :

$$W = \frac{1}{\sqrt{4 \ln 2}} \prod_{i=1}^D FWHM_i^{1/D}$$

Cela correspond à la valeur de W dans le cas d'un champ gaussien ayant pour $FWHM$ $FWHM_i$ suivant chaque coordonnée i .

▷ **Approximations de $p(m = x)$ et $p(n = x)$**

Il est prouvé dans [Adl81] que m suit une distribution de Poisson de paramètre $E\{m\}$, pour des seuils u élevés, d'où :

$$p(m = x) \approx \frac{1}{x!} E\{m\}^x e^{-E\{m\}}$$

Pour approcher $p(n = x)$ on utilise une forme que l'on sait asymptotiquement correcte d'après [Adl81], et l'on détermine les paramètres ensuite.

On suppose donc que $p(n = x)$ est de la forme :

$$p(n = x) \approx \frac{2\beta}{D} x^{\frac{2}{D}-1} e^{-\beta \frac{2}{D}}$$

ce qui donne entre autres $E\{n\} = \Gamma(D/2 + 1)\beta^{-D/2}$, d'où l'on déduit par l'équation 2 :

$$\beta = \left(\Gamma(D/2 + 1) \frac{E\{m\}}{E\{N\}} \right)^{2/D}$$

Γ est ici la fonction Gamma de Euler définie pour tout $\alpha > 0$ par :

$$\Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt$$

▷ Détermination de la p_value

• p_value non corrigée :
de l'approximation précédente, on déduit :

$$p(n \geq k) \approx e^{\beta k^2/D}$$

Cette valeur constitue la *p_value non corrigée*. C'est la probabilité pour une région activée donnée au dessus du seuil u d'avoir k voxels ou plus. Elle présuppose l'existence d'une activation à une certaine localisation et donne sa distribution de taille. Elle ne prend donc pas en compte le volume total des images.

En pratique on se réfère à cette valeur quand l'étude que l'on mène consiste à vérifier une hypothèse sur l'existence d'une activation à un endroit donné dans les images.

• p_value corrigée :
de $p(m = x)$ et $p(n = x)$ on déduit par sommation :

$$p(n_{max} \geq k) = \sum_{i=1}^{\infty} p(m = i)(1 - p(n < k)^i) = 1 - \exp(-E\{m\}p(n \geq k))$$

d'où :

$$p(n_{max} \geq k) = 1 - \exp\left(-E\{m\}e^{-\beta k^2/D}\right)$$

Cette valeur constitue la *p_value corrigée*. C'est la probabilité qu'il existe dans le volume S au moins une région activée de k voxels ou plus. Elle ne présuppose pas l'existence des activations et prend en compte le volume total des images. Elle est dite corrigée en ce sens qu'elle rapporte les statistiques au nombre S de voxels étudiés.

En pratique si l'on ne sait pas à l'avance où rechercher les activations, et que le but de l'étude est de déterminer où elles vont apparaître, on se réfère à cette valeur.

2.5.3 Résultats et problèmes soulevés par le logiciel

Nous avons utilisé notre jeu de données synthétiques ainsi que des données réelles pour caractériser le comportement de la méthode SPM dans différents cas.

▷ choix du seuillage en intensité de la carte-z

Il est nécessaire (voir section 2.5.1) de commencer par effectuer un seuillage en intensité de la carte-z. Le problème réside dans le caractère arbitraire du choix de ce seuil.

Les concepteurs de SPM préconisent "par expérience" d'appliquer un seuil u entre 2 et 3,5 (pour mémoire, un seuil $u = 2$ correspond à une p_value associée de $2,27 \cdot 10^{-2}$ et un seuil $u = 3,5$ à une p_value de $2,33 \cdot 10^{-4}$).

On a pu constater que même dans cet intervalle, le choix de u n'est pas anodin. Ainsi sur un échantillon de nos données synthétiques comprenant d'un côté 20 images de base et de l'autre 20 images d'activation, nous avons appliqué successivement un seuil $u = 2,50$ puis $u = 3,30$. La figure 9 montre les composantes connexes résultant de ce seuillage.

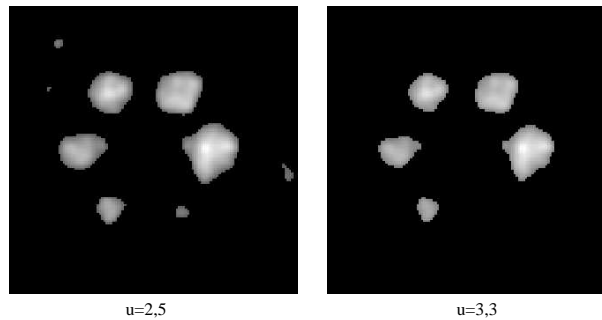


FIG. 9 – *Composantes connexes résultant de différents seuils u*

▷ prétraitement des images : filtrage gaussien

Même s'il ne fait pas partie intégrante de la méthode statistique proprement dite, un filtrage des images par un noyau gaussien est le plus souvent nécessaire. C'est pourquoi un module du logiciel SPM propose la réalisation de ce lissage. Cette étape a deux objectifs :

- Augmenter le rapport signal sur bruit.
- Accroître la corrélation spatiale dans les images.

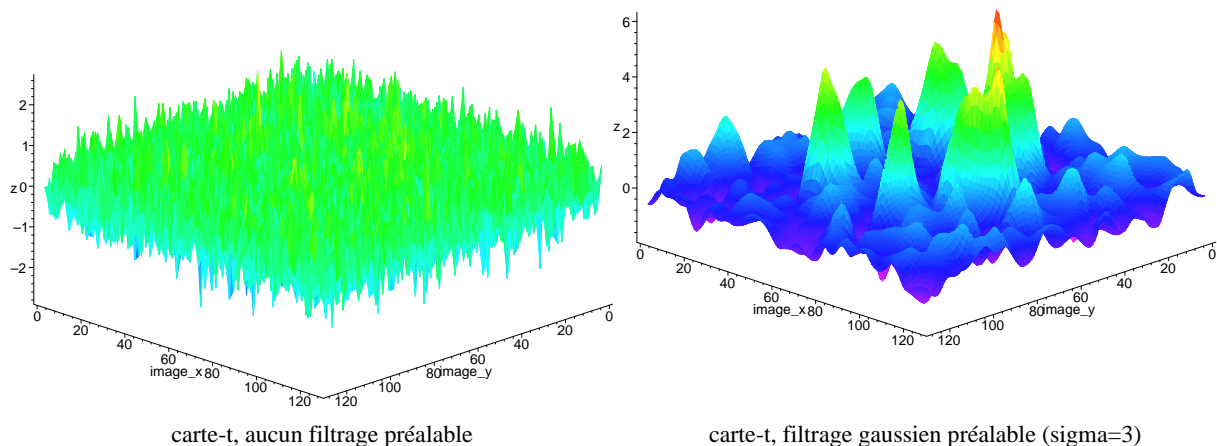


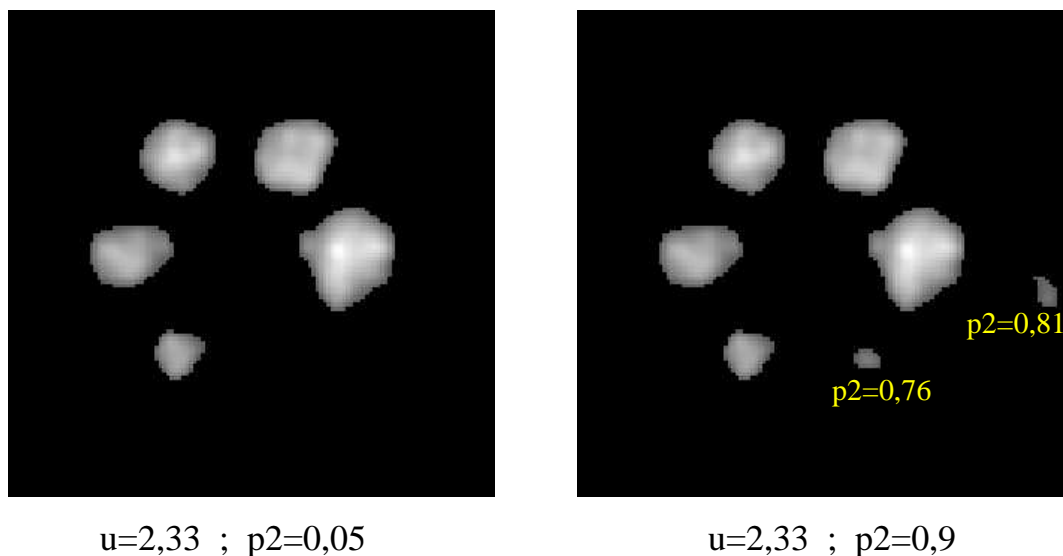
FIG. 10 – *Effets du filtrage gaussien*

Ainsi, si l'on observe les cartes-t obtenues sur un échantillon de nos données, avec ou sans filtrage gaussien (voir figure 10), on constate bien ces deux effets.

La carte-t obtenue sans filtrage gaussien préalable sera beaucoup plus difficile à interpréter et à exploiter.

Cependant un des inconvénients majeurs de cette étape de filtrage gaussien des images est de faire considérablement chuter la résolution effective des données. Ainsi dans notre cas, en filtrant par un noyau gaussien tel que $\sigma = 3$, on arrive à une *FWHM* de 7 pixels (ou voxels) dans chaque direction. Cela correspond à une résolution effective de 7^D où D est la dimension des images. Toute activation de taille voisine sera assimilée à du bruit par l'analyse statistique de SPM.

Sur le même échantillon de données que pour l'exemple précédent nous avons obtenu le résultat montré figure 11. Nous avons dans les deux cas appliqué un seuil $u = 2,33$ (à savoir $p1 = 0,01$), puis dans le premier cas une $p2_value$ corrigée $p2 = 0,05$ (5% de chances de se tromper en disant que les composantes connexes retenues constituent bien des activations), et dans un deuxième temps $p2 = 0,9$.

FIG. 11 – *Activations de petite taille*

On constate que la plus petite des zones d'activation a une p_2 _value de 0,76 alors qu'une composante provenant du bruit dans les images a une p_2 _value de 0,81. Or par construction, nous avons fixé le diamètre de la plus petite activation comme étant la *FWHM* de nos images. La perte de résolution résultant du filtrage gaussien pose donc le problème de la détection des activations de taille modérée.

► prétraitement des images : normalisation

Un module permettant la normalisation des images en intensité est également fourni dans le logiciel SPM. Il semble en effet important, si l'on souhaite faire une analyse statistique sur des images, que les intensités au sein de ces images soient comparables. Or, malgré un protocole d'acquisition identique, d'infimes variations de dose de traceur injecté ou de temps d'attente entre l'injection et l'acquisition peuvent occasionner des variations non négligeables de l'intensité des images (remarque : ce problème ne touche pas notre jeu de données synthétiques puisque par construction toutes les images ont les mêmes distributions d'intensité).

Ainsi l'exemple figure 12 montre deux images affichées toutes les deux sur 256 niveaux de gris de 0 à 256. Elles correspondent toutes les deux à des sujets de contrôle, et ne doivent donc pas présenter de différence majeure. Elles ont les paramètres suivants :

	<i>intensité min</i>	<i>moyenne</i>	<i>intensité max</i>
<i>image 1</i>	0	15	256
<i>image 2</i>	0	7	105

Il est donc bien nécessaire d'effectuer une normalisation des intensités dans les images avant leur analyse statistique.

Le module que propose le logiciel SPM effectue cette normalisation conjointement dans toutes les images en fonction d'une valeur moyenne commune fixée par l'utilisateur. Ceci semble à première vue être une solution satisfaisante, qui convient parfaitement dans le cas de patients normaux, ou dont l'atteinte est faible.

Cependant, et en particulier dans le cas de la maladie d'Alzheimer, un patient atteint aura une hypoperfusion cérébrale globale (les zones normales seront perfusées de manière identiques à celles d'un sujet de contrôle et les zones atteintes le seront moins). Normaliser les images suivant une même valeur moyenne aura alors tendance à augmenter artificiellement l'intensité dans les zones normales et donnera alors l'impression d'une hyperperfusion dans ces zones.

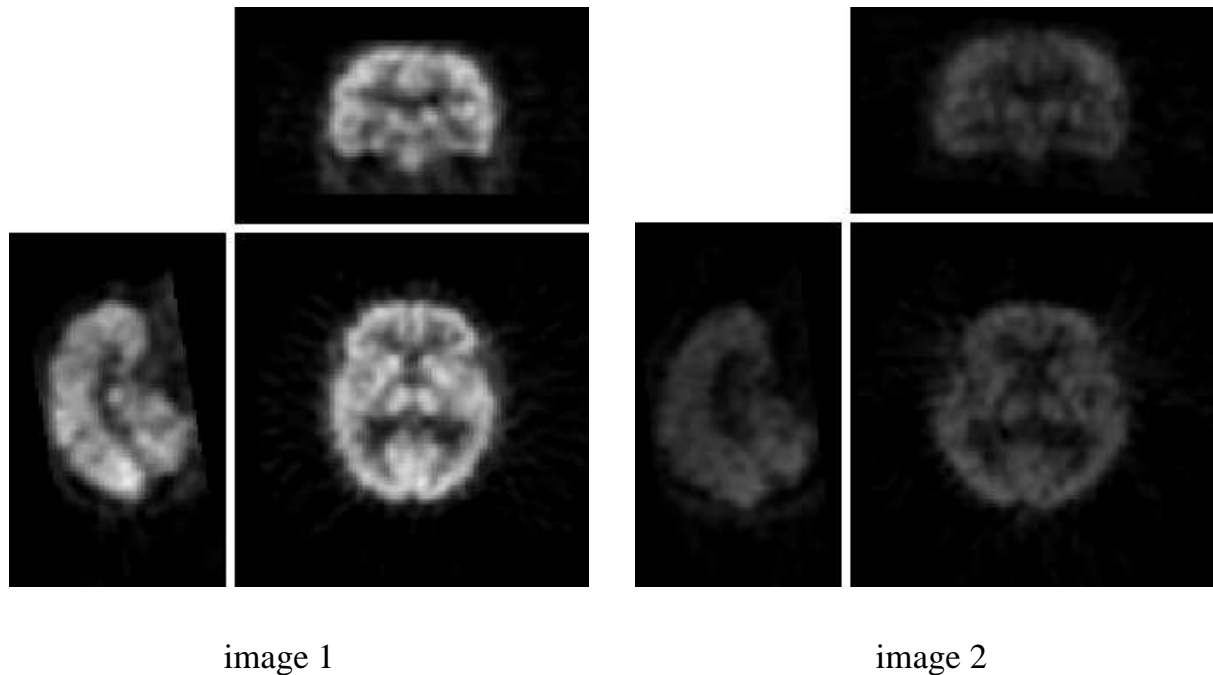


FIG. 12 – *Variations d'intensité entre les images*

Ce défaut a été mis en évidence dans [SPQH98]. Ils proposent comme alternative de normaliser chaque image par rapport aux 1% de voxels ayant les plus fortes intensités. De cette façon on n'augmente pas artificiellement les intensités dans certaines zones. De plus le fait de ne pas normaliser par rapport au voxel d'intensité maximale seul, mais à un groupe de voxels permet une normalisation plus robuste.

▷ **prétraitement des images : recalage**

Un autre module est fourni avec le logiciel SPM. Il permet la mise en correspondance des images que l'on souhaite analyser avec un modèle de même modalité, lui-même recalé sur l'atlas de Talairach et Tournoux [TT88].

Il est en effet indispensable si l'on souhaite faire une analyse statistique sur des images qu'elles soient toutes dans le même référentiel. Or l'acquisition des images (notamment les mouvements des patients) et les variations morphologiques entre les patients ne permettent pas de respecter une telle contrainte. C'est pourquoi une étape de recalage est nécessaire.

Par ailleurs le fait que SPM utilise l'atlas de Talairach permet après l'analyse statistique d'identifier les zones anatomiques auxquelles correspondent les activations trouvées.

Cependant le recalage que propose le logiciel SPM est un recalage affine (transformation rigide + facteur d'échelle + non conservation des angles), qui n'est pas réellement adapté au cas de patients multiples. Nous nous proposons donc d'étudier les possibilités d'utiliser des méthodes de recalage élastique qui sont elles adaptées à cette situation. Nous verrons à la section 3.2.3 les effets de telles modifications sur les résultats de l'analyse statistique.

3 Applications

3.1 Images cardiaques

3.1.1 Contexte

La société Focus Imaging (Sophia Antipolis) a développé un outil (cardiomatch) fournissant aux cardiologues une aide au diagnostic des déficiences cardiaques.

Dans ce cadre, elle souhaite améliorer l'utilisation qui est faite des bases de données de référence suivant le

patient considéré. Pour cela il faut déterminer quels paramètres influent de manière significative sur la perfusion du ventricule gauche des individus.

Nous avons participé à cette étude en comparant, au moyen du logiciel SPM, la différence de perfusion du ventricule gauche à l'effort entre hommes et femmes.

3.1.2 Données

La société Focus Imaging nous a fourni les images SPECT cardiaques (voir figure 13) à l'effort de 41 femmes et 34 hommes, préalablement recalées entre elles et normalisées en intensité.

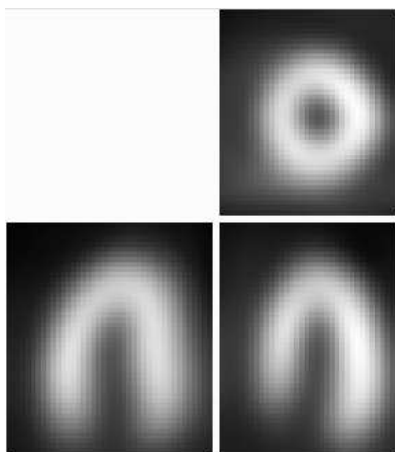
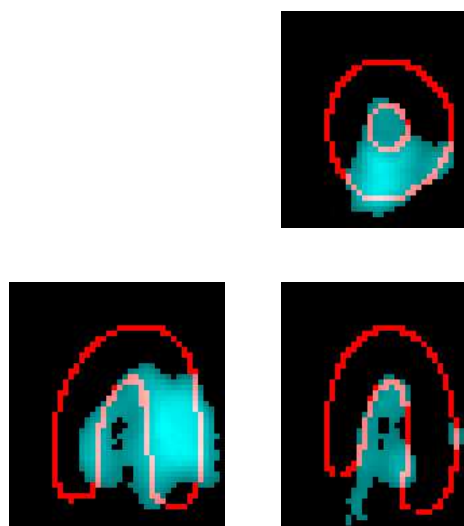


FIG. 13 – Image SPECT cardiaque

3.1.3 Résultats

Nous avons effectué la comparaison statistique de ces deux groupes d'images (hommes/femmes) à l'aide de SPM, et avons obtenu les résultats exposés figure 14.



$p_1=0,001$ $p_2=0,001$

FIG. 14 – Différences significatives hommes/femmes à l'effort par SPM

Le contour du ventricule gauche calculé sur la moyenne des images utilisées y est représenté en rouge, et en vert les zones considérées comme significativement différentes entre hommes et femmes à l'effort, pour un seuillage en intensité à $p_1 = 0,001$ ($u = 3,09$) et un seuillage sur les composantes connexes avec $p_2 = 0,001$ corrigée.

De leur côté les membres de l'équipe Recherche & Développement de Focus Imaging ont effectué la comparaison des mêmes images par la méthode suivante :

- calcul du test de Student sur les images.
- seuillage de la carte-t ainsi obtenue à $p = 0,001$.
- exclusion des voxels isolés restants.

Ils ont obtenu les résultats exposés figure 15, où l'on note en blanc les zones retenues comme significativement différentes entre les deux populations.

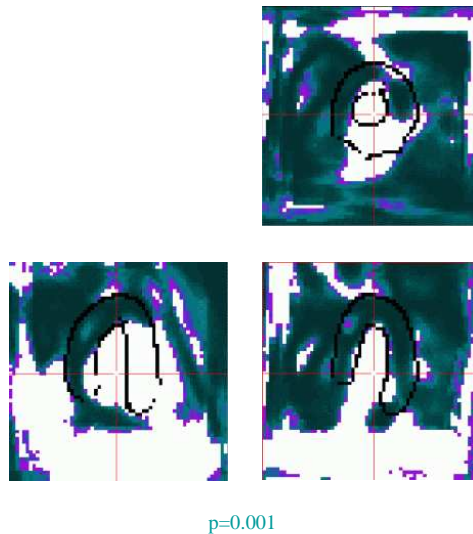


FIG. 15 – *Différences significatives hommes/femmes à l'effort, résultats de Focus*

On constate la très nette amélioration de la localisation des activations apportée par l'utilisation du logiciel SPM. Une simple exclusion (par un choix arbitraire de la taille du voisinage à considérer) des voxels isolés ne suffit pas à évacuer des activations retenues un grand nombre de voxels qui sont en dehors du myocarde et qui correspondent donc à du bruit sur le fond des images.

On peut également observer le résultat obtenu par seuillage sur l'image de différence des moyennes des deux populations, exposé figure 16.

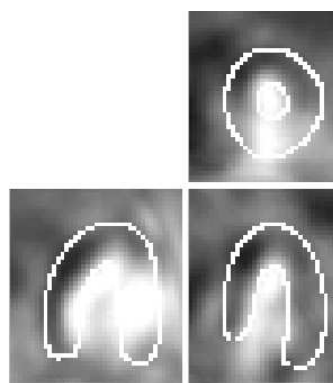


FIG. 16 – *Image seuillée de différence hommes/femmes*

On retrouve bien globalement les mêmes zones d'activation, mais leur localisation est plus délicate et très sensible au bruit. Par ailleurs le seuil choisi sur cette image de différence ne présente pas de justification statistique et ne permet pas de donner le degré de significativité des zones retenues.

Ces résultats ont par ailleurs été approuvés par les cardiologues avec lesquels la société Focus Imaging collabore pour le développement de son produit.

3.2 Étude de l'apathie

3.2.1 Contexte

Le service de médecine nucléaire du Centre Antoine Lacassagne (Nice) pratique régulièrement, à la demande des psychiatres prescripteurs, des explorations de la perfusion cérébrale de certains patients.

Dans ce cadre, ce service a recueilli un certain nombre d'examen se rapportant à l'apathie. En collaboration avec le Dr. Philippe Robert, psychiatre au CAL, il a été décidé d'utiliser ces images pour une étude de cette pathologie. Un des objectifs était de vérifier que la zone cingulaire (voir figure 17) présente une hypoperfusion chez les patients apathiques.

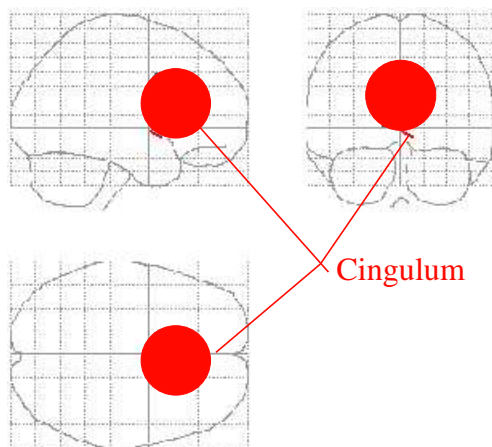


FIG. 17 – Localisation du cingulum

3.2.2 Données

Pour cette étude nous avons exploité 41 images SPECT cérébrales de patients répartis de la façon suivante :

- 21 apathiques, dont 13 déments et 8 non-déments.
- 20 non-apatiques, donc 14 déments et 6 non-déments.

3.2.3 Résultats

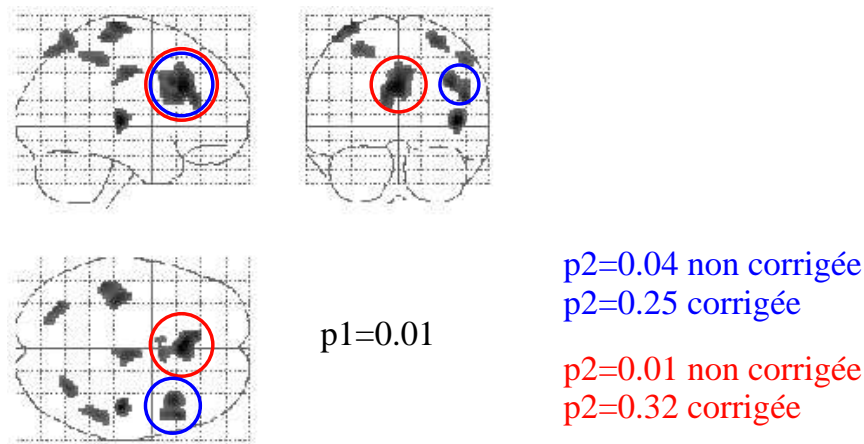
Nous avons profité de ces données pour observer les effets du recalage sur les résultats de l'analyse statistique.

▷ première phase : avec le recalage affine proposé par SPM

Nous nous sommes contentés dans un premier temps de recalcr les 41 images sur le modèle dans l'espace de Talairach à l'aide du module de recalage du logiciel SPM.

Nous avons obtenu les résultats exposés figure 18.

On vérifie bien l'hypothèse d'hypoperfusion du cingulum dans le cas des patients apathiques (zone rouge). En effet, p_2 n'est significative que non corrigée, mais on a vu section 2.5.2 que si l'on se plaçait dans le cadre de la vérification d'une hypothèse, on pouvait se contenter de la p -value non corrigée.

FIG. 18 – *Apathie - recalage affine*

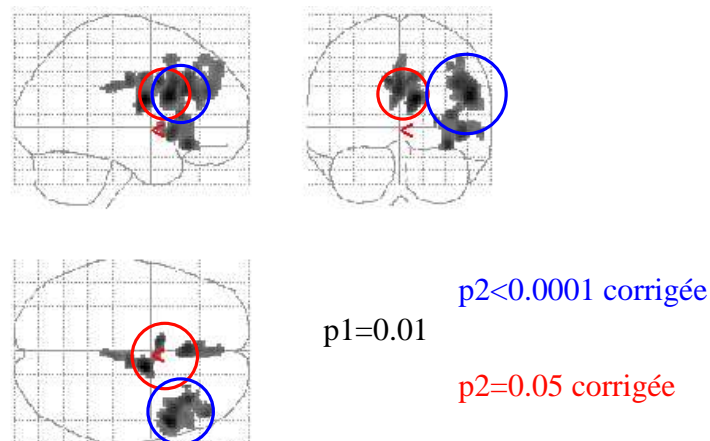
On remarque une autre zone, correspondant au gyrus frontal droit, qui semble être à peu près de même importance que le cingulum, mais on ne peut pas exploiter ce phénomène d'avantage puisque nous n'avons pas de connaissance a priori sur la perfusion dans cette zone dans le cadre de notre expérience. Or a posteriori, le Dr. Robert était satisfait de voir apparaître le gyrus frontal droit.

▷ seconde phase : avec une méthode de recalage élastique

Étant donnés ces résultats, nous nous sommes demandés si un recalage plus performant des images pourrait permettre d'accroître le degré de significativité des zones détectées.

Dans cet objectif nous avons employé une méthode de recalage élastique, développée au sein du projet Epidaure par JP.Thirion [Thi95], pour recaler toutes les images sur celle d'un des sujets normaux de la base. Nous avons ensuite recalé ce sujet sur le modèle de SPM dans l'espace de Talairach et appliqué la même transformation aux autres images.

Nous avons alors obtenu les résultats exposés figure 19.

FIG. 19 – *Apathie - recalage élastique*

On constate que la localisation des zones a été légèrement modifiée, mais reste similaire à ce que l'on trouvait précédemment. En revanche, le degré de significativité des activations a augmenté de façon spectaculaire. Outre l'accroissement du degré de confiance dans les activations détectées, cela permet également de s'affranchir des taches provenant du bruit présentes figure 18.

Il est donc indéniable que la qualité du recalage des images est un facteur influant de manière non négligeable sur

les résultats de l'analyse statistique ultérieure. De plus un recalage élastique semble (d'après les commentaires effectués par le Dr. Robert) donner des résultats plus satisfaisants que le recalage affine proposé par le logiciel.

3.3 La maladie d'Alzheimer

3.3.1 contexte

La maladie d'Alzheimer est l'étiologie la plus fréquente des syndromes démentiels et sa prévalence progresse exponentiellement avec l'âge de 3% vers 65 ans à 30% après 90 ans [Ber97]. Ce problème de santé a stimulé les recherches pharmacologiques qui ont permis de proposer des traitements améliorant les performances cognitives des patients. Cependant, ces thérapeutiques doivent être prescrites tôt dans l'évolution de la maladie pour avoir une efficacité. Or le diagnostic de certitude de la maladie d'Alzheimer ne peut être obtenu qu'en post-mortem, par la constatation des stigmates histopathologiques classiques que sont la dégénérescence neurofibrillaire et les plaques séniles. Le diagnostic précoce repose actuellement sur des critères cliniques qui permettent de parler de démence de type Alzheimer ou de maladie d'Alzheimer probable. Il devient particulièrement important d'étayer ce diagnostic à la phase précoce par des données objectives et quantifiées.

Dans cet objectif nous souhaitons comparer les images fonctionnelles cérébrales de patients atteints par la maladie à celles de sujets sains, afin de mettre en évidence les différences significatives entre les deux populations et de les quantifier.

3.3.2 Données - Acquisition

Pour l'étude de la maladie d'Alzheimer, la constitution d'une base de données d'images SPECT de patients atteints et de sujets sains (servant de population de contrôle) est nécessaire. Cette base servira à la comparaison statistique des images fonctionnelles des deux populations.

Nous avons donc tout d'abord effectué une recherche bibliographique visant à déterminer les critères à utiliser lors du recrutement des patients de la base en question.

• Les sujets malades

▷ Diagnostic clinique

Pour l'établissement du diagnostic clinique de maladie d'Alzheimer probable, le critère utilisé est celui du NINCDS² et de l'ADRDA³ [MDF⁺84].

▷ Différents stades de la maladie

Les patients chez lesquels on a diagnostiqué cliniquement la maladie d'Alzheimer sont ensuite classés en fonction du degré de maladie dont ils sont atteints.

Selon les études différents moyens de gradations ont été utilisés :

Clinical Dementia Rating (CDR) défini dans [HBD⁺82].

Mini Mental State Examination (MMSE) défini dans [FFM75].

▷ Autres critères d'inclusion

Dans certains articles on trouve des critères d'inclusion complémentaires :

être droitier [BMH⁺97]

l'examen clinique de tous les patients a été effectué par le même neurologue (pour homogénéiser le diagnostic) [OIC⁺94].

2. National Institute of Neurological and Communicative Disorders and Stroke.

3. Alzheimer's Disease and Related Disorders Association.

▷ Tests complémentaires

On vérifie que les patients répondant au critère NINCDS-ADRDA ne comportent pas d'autre problème cérébral. Le plus souvent cette vérification est effectuée au moyen d'exams IRM⁴ et/ou radiologiques [MFK⁺95].

Une confirmation du diagnostic de maladie d'Alzheimer probable est le plus souvent effectuée, afin de contrôler que :

Les patients atteints de maladie d'Alzheimer probable continuent à satisfaire le critère NINCDS-ADRDA pendant un an [MFK⁺95].

Les patients très faiblement atteints au départ vérifient ce critère au cours d'un suivi ultérieur [BMH⁺97].

▷ Données recueillies

Outre les paramètres propres à la maladie d'Alzheimer, d'autres données sont collectées pour permettre de mieux caractériser la base de données des patients, en particulier l'âge et le sexe.

• Les sujets sains

▷ Critères d'exclusion

Selon les études, on retrouve différents critères d'exclusion pour les sujets sains servant de contrôles. Cependant, d'une manière générale, le but de ces critères est d'éliminer tous les patients pouvant présenter des anomalies cérébrales :

pertes de connaissance, commotions [HKM94]

traumatismes crâniens [OIC⁺94]

troubles neurologiques ou psychiatriques [HKM94, MFK⁺95]

toxicomanie [OIC⁺94]

plongée, boxe [HKM94]

Bien entendu ces sujets suivent également les examens cliniques de diagnostic de la maladie d'Alzheimer, afin de contrôler qu'ils sont négatifs.

▷ Répartition de la population

De même que pour les patients atteints par la maladie d'Alzheimer, on note la répartition de la population de contrôle suivant l'âge et le sexe afin de mieux caractériser la base de données, et autant que possible d'avoir des répartitions similaires dans les deux populations.

▷ Prise en compte d'autres types d'atteintes cérébrales

On peut envisager d'utiliser ultérieurement un groupe de patients atteints de problèmes vasculaires cérébraux non liés à la maladie d'Alzheimer pour contrôler la spécificité de la méthode diagnostique. [MFK⁺95].

3.3.3 Premières images

Nous avons participé, avec le service de médecine nucléaire du Centre hospitalier Antoine Lacassagne de Nice, à l'acquisition des premières images de patients atteints par la maladie d'Alzheimer qui viendront constituer notre base.

Il est à noter que le centre n'a malheureusement pas l'habilitation pour recruter des volontaires sains afin de constituer la base d'images de la population de contrôle. Ces images sont également en cours d'acquisition dans le cadre du projet européen SID⁵, au MRC⁶ d'Edimbourg.

4. Imagerie par Résonance Magnétique

5. SPECT In Dementia - improving cost effectiveness by using novel methods of image analysis for the early detection of Alzheimer's disease.

6. Medical Research Center.

4 Perspectives

4.1 Outil d'analyse statistique et prétraitement des images

En nous appuyant sur notre étude du logiciel SPM, nous prévoyons d'en modifier certains aspects afin de fournir un outil plus proche des besoins que nous rencontrerons. Nous avons déjà recodé (en langage C) les principales étapes de l'analyse statistique afin de les étudier et de les modifier plus facilement.

Dans ce but, nous comptons notamment faire une étude systématique de différents modules de SPM. Ainsi pour le prétraitement des images, il nous semble nécessaire d'envisager d'autres méthodes de lissage, linéaires (filtrage médian par exemple) ou même non-linéaires [KMA97], qui nous permettraient de nous affranchir du bruit sans perdre les activations de faible volume.

De même nous comptons étudier plus précisément différentes méthodes de recalage des images et leurs effets sur les résultats afin de déterminer celle qui sera la plus adaptée à nos images. Plusieurs possibilités s'offrent à nous en matière de recalage non-rigide. Tout d'abord certaines méthodes s'appuient sur la recherche d'invariants géométriques pour la classe de transformations recherchées [FDMA97, SL96]. D'autres méthodes s'appuient directement sur les intensités des points des images comme la méthode des démons que nous avons utilisée dans le cas de l'apathie [Thi96, CPA99], ou des méthodes de maximisation de certains critères de similarité [RMPA98]. Pour le choix de la méthode que nous utiliserons, nous pourrions également nous appuyer sur les travaux de X.Pennec concernant les incertitudes en matière de recalage [PT97].

Nous explorerons également la possibilité d'appliquer d'autres types de seuillages aux cartes statistiques obtenues [Wor94], qui pourraient aussi permettre une meilleure discrimination entre le bruit et les petites activations.

4.2 Phase expérimentale

Nous prévoyons d'appliquer ce travail à l'étude de la maladie d'Alzheimer dans le but de fournir un outil de diagnostic précoce de cette pathologie au moyen d'images fonctionnelles. Ces travaux seront effectués en collaboration avec le Centre Antoine Lacassagne, dans le cadre du projet européen SID⁷.

Nous envisageons également d'effectuer une étude similaire dans le cadre du projet européen QAMRIC⁸, afin d'analyser la maladie de Creutzfeldt Jakob, en utilisant des IRM cérébrales.

7. Spect In Dementia - improving cost effectiveness by using novel methods of image analysis for the early detection of Alzheimer's disease.

8. Quantitative Analysis of MR Scans in Creutzfeldt Jakob Disease.

Références

- [Adl81] R.J. Adler. *The geometry of random fields*. Wiley, 1981.
- [Ber97] C. Berr. 100 000 Nouveaux cas de démence par an. *La Recherche*, pages 68–71, November 1997.
- [BMH⁺97] P. Barenstein, S. Minoshima, C. Hirsh, K. Buch, F. Willoch, D. Mösch, D. Schad, M. Schwaiger, and Kurz A. Quantitative assesment of cerebral blood flow in patients with Alzheimer's disease by SPECT. *The Journal of Nuclear Medicine*, 38(7):1095–1101, July 1997.
- [CPA99] P. Cachier, X. Pennec, and N. Ayache. Fast non-rigid matching by gradient descent: Study and Improvements of the "Demons" algorithm. Technical Report 3706, INRIA, 1999.
- [FDMA97] J. Feldmar, J. Declerck, G. Malandain, and N. Ayache. Extension of the ICP algorithm to non-rigid intensity-based registration of 3d volumes. *Computer Vision and Image Understanding*, 66(2):193–206, 1997.
- [FFM75] MF. Folstein, SE. Folstein, and PR. McHugh. "Mini mental state." A practical method for grading the cognitive state of patients from the clinician. *Journal of Psychiatr. Res.*, 12:189–198, 1975.
- [FHW⁺95] K.J. Friston, AP. Holmes, K.J. Worsley, JP. Poline, and RSJ. Frith, CD. Frakowiak. Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping*, 2:189–210, 1995.
- [FWF⁺94] K.J. Friston, K.J. Worsley, RSJ. Frackowiak, JC. Mazziotta, and AC. Evans. Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1:214–220, 1994.
- [HBD⁺82] CP. Hughes, L. Berg, WL. Danziger, LA. Coben, and RL. Martin. A new clinical scale for the staging of dementia. *Brit. Journal of Psychiatry*, 140:566–572, 1982.
- [HKM94] AS. Houston, PM. Kemp, and MA. Macleod. A method for assessing the significance of abnormalities in HMPAO brain SPECT images. *The Journal of Nuclear Medicine*, 35(2):239–244, February 1994.
- [KMA97] K. Krissian, G. Malandain, and N. Ayache. Directional anisotropic diffusion applied to segmentation of vessels in 3d images. In *First Conference on Scale-space Theory in Computer Vision*, Utrecht, the Netherlands, July 1997.
- [MDF⁺84] G. McKhann, Drachman D., M. Folstein, R. Katzman, D. Price, and EM. Stadlan. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, 34:939–944, 1984.
- [MFK⁺95] S. Minoshima, KA. Frey, RA. Koeppe, NL. Foster, and DE. Kuhl. A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *The Journal of Nuclear Medicine*, 36(7):1238–1248, July 1995.
- [OIC⁺94] A. Osimani, M. Ichise, DG. Chung, Pogue JM., and M. Freedman. SPECT for Differential Diagnosis of Dementia and Correlation of rCBF with Cognitive Impairment. *Canadian Journal of Neurological Sciences*, 21(2):104–111, May 1994.
- [Pap91] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill International, third edition, 1991.
- [PHWF97] JB. Poline, AP. Holmes, KJ. Worsley, and KJ. Friston. Statistical Inference and the theory of Random Fields. Ch. 4 of SPM Course Notes, 1997.
- [PT97] X. Pennec and J.P. Thirion. A framework for uncertainty and validation of 3D registration methods based on points and frames. *Int. Journal of Computer Vision*, 25(3):203–229, 1997.
- [RMPA98] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration. In *First international conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI'98*, Boston, Massachusetts, USA, October 1998.
- [Sap90] G. Saporta. *Probabilités Analyse des données et Statistique*. Technip, 1990.
- [SL96] R. Szeliski and S. Lavall e. matching 3-d anatomical surfaces with non rigid deformations using octree-spline. *the International Journal of Computer Vision*, 18(2):171–186, 1996.
- [SPQH98] P. Saxena, DG. Pavel, JC. Quitana, and B. Horwitz. An Automatic Threshold-Based Scaling Method for Enhancing the Usefulness of Tc-HMPAO SPECT in the Diagnosis of Alzheimer's Disease. In WM. Well, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98*, pages 623–630, October 1998.
- [Thi95] JP. Thirion. Fast Non-Rigid Matching of 3D Medical Images. Technical Report 2547, INRIA, May 1995.

-
- [Thi96] JP. Thirion. Non rigid matching using demons. In *Computer Vision and Pattern Recognition, CVPR'96*, San Francisco, California, USA, June 1996.
- [TT88] J. Talairach and P. Tournoux. *Coplanar stereotaxic atlas of the human brain*. New York: Thieme Medical, 1988.
- [Wor94] KJ. Worley. Local Maxima and the expected Euler characteristic of excursion sets of χ^2 , F and t fields. *Advances Appl. Prob.*, 26:13–42, 1994.
- [Wor96] KJ. Worsley. The Geometry of Random Images. *Chance*, 9(1):27–40, 1996.

Table des matières

1	Introduction	3
2	L'outil statistique SPM	3
2.1	Buts de SPM	3
2.2	Jeu de données synthétiques	3
2.3	Test statistique	4
2.3.1	Caractéristiques du t-test	4
2.3.2	Distribution de Student	4
2.3.3	t-test	5
2.3.4	Cas extrême	5
2.4	Cartes statistiques	6
2.4.1	carte-t	6
2.4.2	Gaussianisation de la carte-t en carte-z	6
2.5	Seuils de confiance	7
2.5.1	Seuillage en intensité de la carte-z	7
2.5.2	Seuillage de taille des composantes connexes	7
2.5.3	Résultats et problèmes soulevés par le logiciel	11
3	Applications	14
3.1	Images cardiaques	14
3.1.1	Contexte	14
3.1.2	Données	15
3.1.3	Résultats	15
3.2	Étude de l'apathie	17
3.2.1	Contexte	17
3.2.2	Données	17
3.2.3	Résultats	17
3.3	La maladie d'Alzheimer	19
3.3.1	contexte	19
3.3.2	Données - Acquisition	19
3.3.3	Premières images	20
4	Perspectives	21
4.1	Outil d'analyse statistique et prétraitement des images	21
4.2	Phase expérimentale	21



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)

<http://www.inria.fr>

ISSN 0249-6399