

Imagerie par résonance magnétique et estimation volumique cérébrale : comparaison de quatre techniques

*Olivier Salomon¹, Agnès Dorion², Michel Zanca¹,
Michel Duyme² et Christiane Capron^{2,3}*

Résumé. Les images, issues de techniques d'Imagerie par Résonance Magnétique en trois dimensions, d'un échantillon de cerveaux de 10 participants féminins sains (20,15 ans \pm s : 1,14 ans) sont soumises à des mesures d'estimation volumique cérébrale. Quatre techniques sont utilisées : une par délimitation manuelle et trois de seuillage/érosion/dilatation (deux semi-automatiques et une automatique). Les deux techniques semi-automatiques fournissent des moyennes volumiques non différentes statistiquement ($1233 \pm s : 86,2 \text{ cm}^3$ et $1218 \pm s : 110,5 \text{ cm}^3$); il en est de même pour la technique automatique et celle par délimitation manuelle, qui fournissent des valeurs volumiques significativement plus élevées que les deux précédentes ($1305 \pm s : 100,6 \text{ cm}^3$ et $1311 \pm s : 82,3 \text{ cm}^3$). Les coefficients de corrélation, pour évaluer la fidélité inter cotateurs, sont tous supérieurs à 0.97 ($p < .001$), quelle que soit la technique. Trois des coefficients de corrélation entre techniques sont

Correspondance à Christiane Capron, INSERM Unité 155, Laboratoire d'Anthropologie Biologique et d'Epidémiologie Génétique, Université Paris VII, 2 Place Jussieu, Tour 16, 3ème étage, 75005 Paris, France (e-mail: capron@ccr.jussieu.fr).

1. M.I.R, Hôpital Gui de Chauliac, CHU Montpellier, France.

2. INSERM U 155, Université Paris VII, 75005 Paris, France.

3. Université Paul Valéry - Montpellier III, France.

inférieurs à 0.66 et ne sont pas statistiquement significatifs, indiquant ainsi que, malgré la bonne fidélité de leurs estimations volumiques, ces techniques ne sont pas toutes interchangeables. La technique "automatique" et celle par délimitation manuelle présentent une corrélation satisfaisante bien qu'imparfaite ($r = 0.92$).

Mots clés : IRM, volume cérébral, techniques de mesure, fidélité.

Key words: MRI, cerebral volume, measurement techniques, reliability.

INTRODUCTION

Le développement récent des techniques d'Imagerie anatomique par Résonance Magnétique (IRM) et d'imagerie métabolique comme le PET (Tomographie d'Emission de Positons) a engendré un intérêt croissant pour l'étude des relations entre des anomalies de différentes structures cérébrales et des pathologies caractérisées, entre autre, par des troubles comportementaux, affectifs et cognitifs. Ces essor a généré une augmentation de publications dans ce domaine. Leurs résultats souvent contradictoires sont à la base de nombreuses revues critiques (pour la schizophrénie par exemple, voir Elkis, Friedman, Wise et Meltzer, 1995 ; Lawrie et Suheib, 1998 ; Videbech, 1997 ; Ward, Friedman, Wise et Schulz, 1996, parmi d'autres) qui souvent, elles-mêmes, exigeraient une analyse méthodologique interne sérieuse (Fox, Parsons et Lancaster, 1998 ; Weinberger, 1996). Néanmoins, ces techniques d'imagerie ont, également, contribué à modifier notre système de pensée relatif à l'évaluation des fonctions cérébrales (Botez, 1996 ; Fox et al., 1998).

Ainsi, dans l'objectif de comprendre l'organisation du fonctionnement psychologique, la neuropsychologie cognitive, dans son secteur neuromorphométrique, ne s'intéresse pas seulement à découvrir comment des atrophies cérébrales importantes pourraient expliquer des modifications comportementales ou l'altération d'une fonction comme le font, par exemple, Pozzilli, Bastianello, Padovani, Passafiume et al. (1991) pour l'étude de la variation entre les atrophies du corps calleux (CC) et des performances de fluidité verbale chez des femmes atteintes de sclérose

en plaques. Elle s'intéresse aussi aux variations cérébrales quantitatives, plus fines, et à leurs relations avec des variations comportementales ou les variations d'une fonction psychologique. C'est dans ce cadre que des auteurs comme Goldberg, Torrey, Berman et Weinberger (1992) étudient les relations éventuelles entre des performances neuropsychologiques, particulièrement des fonctions mnésiques, la morphologie du cerveau et des mesures physiologiques chez des jumeaux monozygotes discordants pour la schizophrénie. D'autres (Giedd, Castellanos, Casey, Kozuch, King et al., 1992 ; Hynd, Semrud-Clikeman, Lorys, Noverly et Eliopoulos, 1991) se centrent sur les relations possibles entre les aires des différentes régions du CC, des mesures cognitives et des comportements sociaux, à partir d'échantillons à la fois d'enfants hyperactifs présentant des troubles de l'attention (ADHD) et de sujets témoins. L'objectif, alors, est celui d'une analyse différentielle quantitative qui permettrait de décrire des phénomènes dans la marge de variation pathologique pour comprendre l'étiologie d'un trouble. A partir de sujets situés dans la marge de variation normale, l'intérêt peut être une tentative de concevoir des lois générales de l'organisation psychologique (voir Hines, Chiu, McAdams, Bentler et Lipcamon, 1992, pour une relation entre les aires de la partie postérieure du CC [le splénium] et la variation des performances aux tests de fluidité verbale chez des femmes volontaires). L'intérêt peut être, également, à travers une analyse différentielle et développementale, celui de cerner les lois générales du développement ontogénétique de structures du système nerveux central (par exemple, Giedd, Castellanos, Rajapakse, Vaituzis et Rapoport, 1997, pour l'étude du lien entre sexe et développement de structures du cerveau chez des enfants) et permettre l'élaboration de normes. Pour ces propos, les techniques d'IRM apparaissent comme un outil puissant pour étudier volumes et surfaces du cerveau et de différentes structures cérébrales.

Certes une approche telle que celle de l'imagerie fonctionnelle par résonance magnétique, pour visualiser les modifications des régions cérébrales lors d'un comportement, aura probablement une valeur heuristique plus grande, grâce à son aspect dynamique. Il s'agit aujourd'hui, pour ce but précis, d'une technique encore en cours d'élaboration et applicable à des comportements simples. Ses difficultés majeures proviennent des artéfacts liés à la récupération et à l'interprétation des

signaux informatifs, propres à l'effet de l'opération mentale sur les indicateurs biologiques cérébraux.

Pour le neuropsychologue, un questionnement sur les relations entre la variation d'un comportement et celle d'une, ou de structures cérébrales peut donc être considéré comme une étape nécessaire. Il s'agit d'une analyse statique. L'exigence des qualités psychométriques, validité et fidélité, des outils nécessaires à l'évaluation des comportements ainsi que la standardisation des conditions d'évaluations sont des impératifs scientifiques communs. Ces mêmes qualités sont à exiger pour les mesures biométriques (volumes et surfaces) des structures cérébrales et pour les conditions d'acquisition et de traitement des IRM. Etant donné l'évolution et donc les changements rapides des techniques d'imagerie anatomique par résonance magnétique, l'aspect différentiel des procédures menant à l'obtention de surfaces et de volumes doit être considéré.

Dans les travaux qui prennent en compte uniquement une ou des structures particulières pour un comportement donné, la question fondamentale de la variation de la valeur volumétrique du cerveau se pose. En effet, il est admis que le volume du cerveau varie d'un individu à l'autre et qu'il existerait, entre autres, des différences morphométriques liées au sexe (Blatter et al., 1995, parmi d'autres). Si une association, statistiquement significative, est observée entre des mesures d'une structure cérébrale particulière (hippocampe, amygdales, corps calleux, lobe temporal etc.) et le volume cérébral, alors toute approche rigoureuse exige que ce lien soit considéré (Soininen et al., 1994). Traditionnellement, quand une covariation statistiquement significative existe entre deux variables, les données brutes sont converties en mesures standardisées en utilisant le(s) modèle(s) statistique(s) approprié(s) permettant d'ajuster pour un (ou des) covariate(s). De telles analyses offrent la possibilité d'annuler la part de variance non pertinente pour la question posée. Les préalables à de telles corrections reposent sur les qualités de la mesure. Préciser ces qualités, qui sont dépendantes des protocoles d'acquisitions des IRM, est non seulement essentiel pour opérer ces corrections, mais également indispensable pour effectuer des comparaisons entre études.

La fidélité d'un ensemble de mesures peut être logiquement définie comme la proportion de variance vraie, c'est-à-dire celle due aux différences individuelles, et non de variance due à l'erreur de mesure. Cette

erreur de mesure peut être imputable (1) à la variabilité dans la répétition de la même mesure par un même cotateur, ou (2) à la variabilité dans la reproduction d'une même mesure entre différents cotateurs.

La validité se réfère, ici, à la question de savoir si la mesure d'une structure cérébrale (ou d'un ensemble de structures) obtenue par une technique donnée, représente bien ce qui est prétendu être mesuré. Il est question plus précisément de validité externe. Evaluer cette propriété dans le domaine biométrique n'est pas simple. Il faudrait vérifier que la technique de mesure employée estime un volume cérébral identique au volume réel *in vivo*. La réponse directe à cette interrogation nécessiterait donc d'accéder au volume cérébral entier *in vivo*. Actuellement, ceci n'est scientifiquement pas possible et l'objectif de cette étude n'est pas de répondre à cette question. Il est considéré que les procédures IRM fournissent une représentation opérationnelle du cerveau.

Il est connu que toute mesure cérébrale *post-mortem* dépend étroitement du mode d'extraction (méninges plus ou moins bien enlevées, détérioration de certaines zones du cerveau, etc.) et du mode de conservation post-opératoire. Soulignons que les techniques d'analyse d'image présentent des difficultés d'isolement analogues. Par exemple, dans les techniques de seuillage/érosion/dilatation¹, selon le seuil et le nombre d'érosion/dilatation décidé par l'opérateur, le cerveau sera partiellement ou totalement isolé des enveloppes cutanéograsseuses et méningées. Il faut alors s'assurer que le volume cérébral ne variera pas, pour un même sujet, selon le seuil initial choisi. Les techniques par délimitations manuelles, où la région d'intérêt de chacune des coupes du cerveau est dessinée manuellement par l'opérateur, nécessitent des connaissances anatomiques approfondies et des prises de décisions dépendantes de l'opérateur.

Des techniques d'imagerie médicale très diverses ont été publiées pour estimer le volume cérébral (Jernigan et al., 1991 ; DeCarli et al., 1992 ; Kohn et al., 1992 ; Murphy et al., 1992 ; Andreassen et al., 1993 ; Cline et al., 1993). Cependant ces travaux incluent de petits nombres de sujets sains. Deux équipes ont récemment utilisé les techniques

1. Pour une définition, voir la partie Méthode.

en trois dimensions (3D) pour estimer le volume cérébral sur un échantillon plus important. Blatter et al. (1995) ont mis au point une technique de segmentation multispectrale en utilisant une décomposition en densité de protons, T1 et T2. Leurs indices de fidélité, portant sur un échantillon de 15 cerveaux, sont supérieurs à 0,95 pour un même cotateur, et varient de 0,76 à 0,84 entre des cotateurs différents. La fidélité inter-cotateurs n'apparaît pas très élevée (de 58% à 70% de variance commune). Hasboun et al. (1996) présentent une technique assez rapide fondée sur des opérations successives de seuillage, érosion et dilatation. Pour un échantillon de 16 cerveaux, les auteurs obtiennent un indice de fidélité entre cotateurs de 0,98 ($p < 0.0001$). Dans leur revue de travaux antérieurs (Hasboun et al., 1996, p. 10), les auteurs considèrent que le volume moyen des cerveaux varie en fonction de la technique utilisée (technique de délimitation manuelle, technique multispectrale ou technique *post-mortem*). Mais, comme dans les publications citées, les indicateurs statistiques indispensables aux comparaisons des volumes obtenus par les différentes techniques (taille de l'échantillon et écart type) ne sont pas toujours fournis, une comparaison rigoureuse, entre ces résultats particuliers, reste impossible.

Dans le présent travail, quatre techniques différentes d'estimation volumique sont appliquées à un échantillon de cerveaux appartenant à 10 participants sains. Il s'agit d'une technique de délimitation manuelle et de trois techniques de seuillage/érosion/dilatation. L'objectif est de connaître (1) s'il y a une variation importante du volume cérébral selon les techniques, (2) si les fidélités des estimations volumiques sont satisfaisantes, d'un point de vue morphométrique et volumétrique, et (3) si ces techniques sont interchangeables.

MÉTHODE

Echantillon d'étude

L'étude porte sur un échantillon de cerveaux de 10 sujets de sexe féminin. Ces sujets sont des étudiantes de première et deuxième années de psychologie. Elles ont été sélectionnées dans le cadre d'une recherche portant sur l'étude des liens éventuels entre des aires du corps calleux et

des performances de fluidité verbale. La sélection des sujets ayant subi un examen IRM s'est fondée sur leurs performances aux différentes épreuves de fluidité verbale. Il s'agissait de retenir les étudiantes représentant les extrêmes de la distribution des performances observées. L'âge moyen des sujets au moment de l'ensemble des épreuves était de 20,15 ans (18,5-22 ans ; $s = 1,14$ ans). Les sujets retenus pour l'étude ne présentaient aucun antécédent neurologique connu. Ces participants ont tous donné leur consentement écrit informé.

Acquisitions IRM

Les examens IRM ont été réalisés sur une unité SIGNA 1,5 Tesla (Société "General Electric", GE). Chaque sujet est placé dans l'aimant en décubitus dorsal. La tête est discrètement basculée en hyperextension de manière à ce que la ligne, réunissant la commissure labiale et le conduit auditif externe, soit perpendiculaire à l'axe horizontal de la table. L'indicateur axial de l'IRM est ainsi parallèle à cette ligne. Une séquence préalable en spin-écho 600/11/1 (TR/TE/NEX) permet de positionner sur une coupe sagittale médiane la future acquisition volumétrique. Cette dernière est obtenue au moyen d'une séquence d'écho de gradient 3D de type SPGR (Spoiled Gradient Recalled acquisition at the steady state) avec les paramètres suivant: 24/5/35°/1 (TR/TE/angle de Bascule/NEX). La taille de la matrice est de 256x192 interpolée en 256x256 et le champs de vue réglé à 24 cm. Cette séquence fournit 124 coupes de 1,5 mm d'épaisseur. Tous les examens sont sauvegardés sur cassettes Digital Audio Tape.

Reconstruction 3D des images cérébrales

Chaque examen est transféré sur une station de travail SUN fonctionnant sous l'environnement Advantages Windows de la Société GE. Les reconstructions tridimensionnelles et les mesures de volumes sont réalisées grâce à l'option 3D développée par GE. Le problème d'anisotropie volumique (due à la matrice non carrée et à l'épaisseur de coupe différente de la taille du pixel) est compensé, lors de la reconstruction, par une interpolation adaptée, incluse dans le logiciel.

Le modèle reconstruit est visualisé dans quatre fenêtres de visualisation synchronisée : 3D, axiale, sagittale et coronale. Tout déplacement du curseur, dans une de ces fenêtres, est immédiatement répercuté dans les trois autres. La largeur et le niveau de l'échelle de gris des quatre fenêtres sont ajustés visuellement.

Techniques de mesure du volume cérébral

Dans tous les cas, le volume cérébral est reconstruit selon la technique précédemment décrite.

1. Délimitation manuelle (DEL). Cette technique procède en trois étapes : restauration de la tête, section de la moelle épinière et extraction manuelle de l'encéphale.

Restauration de la tête : la tête est reconstruite à partir des 124 coupes d'origine. La première opération consiste en un seuillage éliminant les voxels non pertinents, c'est-à-dire le bruit de fond. Les valeurs des niveaux de gris s'étendent sur une échelle de 0 à 4095 unités et sont fixées empiriquement en pourcentage sur le critère d'efficacité d'élimination visuelle du bruit de fond. Les valeurs appliquées sont de 0,5% (seuil inférieur) et de 61% (seuil supérieur).

Section de la moelle épinière : la moelle épinière est sectionnée au niveau de la base du cervelet sur chaque coupe où elle est visible (Figure 1). On utilise à cette fin une ligne perpendiculaire à l'axe médian de la moelle.

Extraction de l'encéphale : le logiciel développé par GE permet, par l'intermédiaire de différentes "peintures", d'isoler des autres structures un organe ou une structure hétérogène. Il s'agit, ici, de l'encéphale. Il suffit de délimiter la structure en s'aidant de l'option du logiciel, permettant une définition guidée du tracé, par reconnaissance des séparations entre pixels de différentes densités. La délimitation est effectuée, coupe par coupe, sur une vue 2D sagittale à l'aide d'un pinceau de taille 1 voxel. Elle est reportée de manière synchronisée sur les vues axiales et coronales. Cela permet le contrôle tridimensionnel du travail effectué. La surface délimitée est en outre simultanément colorée en rouge, ce qui permet de s'assurer de la prise en compte de tous les voxels. Une fois la

délimitation des 124 coupes réalisée, il faut reconstruire le volume de la structure isolée, c'est-à-dire la "zone d'intérêt". Une visualisation de la structure encéphalique en trois dimensions est alors obtenue et affichée. Le volume cérébral est directement calculé en cm^3 par le logiciel qui convertit le nombre de voxels grâce à un étalonnage dimensionnel.

Cette technique exige de très bonnes connaissances en neuroanatomie pour la sélection de la région d'intérêt. Elle est réalisable en 2 heures.

Figure 1. Plan de coupe de la moelle épinière sur une image sagittale médiane.

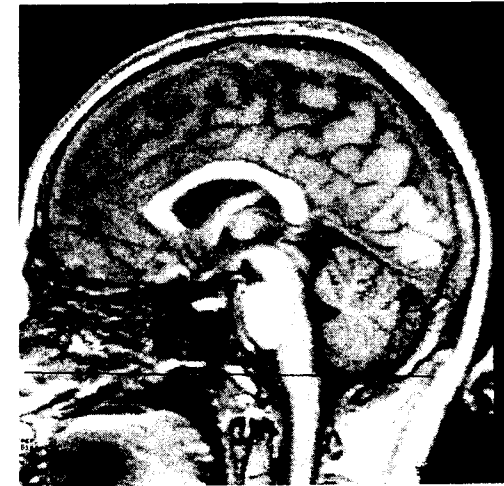


Figure 1. Slice plane of the spinal cord on the midsagittal image.

2. Seuillage automatique [AUTO]. Cette technique ne procède qu'en deux étapes : restauration de la tête et extraction automatique de l'encéphale.

Restauration de la tête : un seuil inférieur de 0,5% et un seuil supérieur de 61% du maximum sont appliqués. Il n'y a pas, ici, d'étape spécifique de section de la moelle épinière. Elle est en effet éliminée automatiquement lors des opérations (érosion) qui vont suivre.

Extraction de l'encéphale : elle consiste à éliminer les enveloppes osseuses, cutanéograsseuses et méningées par des opérations successives d'érosion/dilatation. L'application de N érosions élimine N couches de voxels, tandis que N dilatations restituent N couches de voxels. Le nombre d'érosions, pour isoler l'encéphale des autres structures cérébrales, est déterminé par appréciation visuelle sur la vue 3D. Vu les seuils choisis lors de la première étape, 9 érosions sont requises pour éliminer tous les voxels non pertinents. Ensuite, la région d'intérêt est sélectionnée et les voxels non contigus sont éliminés. Cette deuxième opération doit être suivie par un nombre identique (9) de dilatations (idempotence) afin de restaurer le volume cérébral correct. Il est, alors, directement calculé en cm³ par le logiciel.

Cette technique est qualifiée d'"automatique" car, hormis le seuillage initial faisant intervenir l'opérateur, la sélection de la région d'intérêt et l'estimation du volume sont effectuées par l'ordinateur. Il n'y a donc pas de variation dans la mesure imputable au cotateur. De très courte durée (3 minutes), cette technique n'exige pas de connaissance en neuroanatomie.

3. Seuillage semi-automatique, 1ère technique [SMAUTOA, Hasboun et al., 1996]. Elle procède en trois étapes, comme pour la technique 1, mais en diffère par les seuils utilisés pour la restauration de la tête et par la technique d'extraction de l'encéphale.

Restauration de la tête : un seuil inférieur de 0,6% et un seuil supérieur de 9,8% du maximum sont appliqués. Ces valeurs sont données dans Hasboun et al. (1996, p. 45).

Section de la moelle épinière : le plan de coupe de la moelle épinière est défini par la ligne de Malgaigne, tracée sur la vue médio-sagittale. Ce plan passe par le bord postérieur du foramen magnum en arrière et le palais osseux en avant. Il est tracé jusqu'à la face dorsale du bulbe rachidien pour éviter de sectionner plus en arrière la face inférieure du cervelet. Une opération de coupe à l'infini supprime la moelle épinière et les structures situées sous ce plan de coupe.

Extraction de l'encéphale : l'opérateur apprécie le niveau de gris moyen en évaluant le niveau de gris de deux fois 10 pixels : dix représentatifs de la substance grise et 10 des espaces sous-arachnoïdiens dans

une coupe axiale de référence pour chaque cerveau. Un seuil inférieur est alors déterminé. Il correspond, d'après Hasboun et collègues (1996), aux 2/3 de la distance séparant la moyenne de 10 niveaux de gris de la substance grise et celle de 10 autres appartenant aux espaces sous-arachnoïdiens. La valeur du seuil supérieur est, quant à elle, fixée à 20 unités au-dessus du double de la valeur du seuil inférieur. Les résidus, externes à la région d'intérêt à conserver, sont éliminés par 5 opérations successives d'érosion/dilatation et 18 opérations successives de dilatation/érosion. L'encéphale est alors affiché et son volume calculé en cm³.

Cette technique demande des connaissances en neuroanatomie. Elle est réalisable en 30 minutes.

4. Seuillage semi-automatique, 2ème technique [SMAUTOB]. Cette technique est analogue à celle d'Hasboun et al. (1996), pour la section de la moelle épinière et pour le seuil de restauration de la tête, et à la technique "automatique" (pour la segmentation de l'encéphale). Elle procède donc en trois étapes, comme les techniques 1 et 3.

Restauration de la tête : la reconstruction 3D est traitée en utilisant les mêmes seuils que la technique SMAUTOA (Hasboun et al.), c'est-à-dire un seuil inférieur de 0,6% et un seuil supérieur de 9,8%.

Section de la moelle épinière : cette étape est réalisée comme décrite dans la technique SMAUTOA (voir Hasboun et al.). Elle n'est cependant pas indispensable dans la mesure où l'étape suivante le fait automatiquement.

Extraction de l'encéphale : neuf érosions, un masque et 9 dilatations sont effectués, comme dans la technique AUTO. Le cerveau est affiché directement par sélection de la touche 3D. Le volume est ensuite calculé en cm³. Plus simple que la technique SMAUTOA, cette technique est réalisable en 15 minutes.

Fidélité de la mesure

Pour toutes les mesures, l'ordre dans lequel les 10 cerveaux sont mesurés est établi selon une procédure aléatoire. La fidélité inter-cotateurs a été estimée pour chacune des trois techniques de mesures

non-automatiques. Deux cotateurs (c1 et c2) évaluent une fois (c1.1 et c2.1) les volumes de dix cerveaux, à l'aide des trois techniques non automatiques (SMAUTOA, SMAUTOB, DEL).

La variance attribuable aux erreurs dues à la répétition de la mesure, pour un même cotateur, pour la technique par délimitation manuelle, a également été estimée. Ainsi, les deux cotateurs, c1 et c2, ont mesuré une seconde fois (c1.2 et c2.2) le volume des dix cerveaux par la technique manuelle (DEL).

Corrélation entre techniques

Il s'agit de savoir si les techniques utilisées dans ce travail sont interchangeable. Pour cet objectif, la valeur volumique moyenne – calculée à partir des mesures obtenues par les cotateurs, c1 et c2, pour chaque cerveau et par chacune des techniques – est utilisée (puisque les mesures inter-cotateurs ne sont pas statistiquement différentes ; voir résultats).

ANALYSES STATISTIQUES

Comparaisons intra et inter-cotateurs pour l'évaluation du volume cérébral en fonction des différentes techniques

(1) Pour tester l'hypothèse nulle selon laquelle il n'y aurait pas de différence significative entre les mesures des 10 volumes cérébraux – selon la technique d'isolement des cerveaux, selon le cotateur et la fidélité intra-cotateur – une ANOVA (Statistica, 1991) à mesures répétées a été utilisée. Dix conditions pour la variable dépendante ont ainsi été comparées. (2) Comme le F était significatif (voir résultats), une comparaison multiple *post-hoc* des 10 conditions a été effectuée en utilisant l'approche de Scheffé (Maxwell et Delaney, 1990). Enfin, le coefficient de variation pour chaque technique a été calculé (s/m) afin d'obtenir un paramètre de dispersion portant sur la précision de la mesure.

Analyse des composantes de la variance pour les mesures intra- et inter-cotateurs pour la technique par délimitation manuelle

Avec la technique de délimitation manuelle, deux différents cotateurs (facteur cotateurs) mesurent manuellement deux fois (facteur essais) chaque cerveau (facteur cerveaux).

Le modèle utilisé [ANOVA - Statistica (1991), CSS: Process analysis, p. 794] permet d'estimer la part de variance attribuable à chacun de ces facteurs.

Ce modèle suppose que chaque mesure est la résultante des différentes composantes suivantes :

- une composante due aux caractéristiques de chacun des cerveaux (facteur cerveaux) ;
- une composante due aux caractéristiques de la mesure de chacun des cotateurs (facteur cotateur) ;
- une composante due à la répétition de la mesure par le même cotateur (facteur essai).

De plus, le modèle comprend des composantes dues aux effets d'interactions :

- l'effet de l'interaction du cotateur pour chacun des cerveaux ;
- l'effet d'interaction du cotateur pour chacun des essais ;
- l'effet d'interaction cotateur \times cerveau \times essai.

Dans ce modèle, les erreurs aléatoires ne sont pas séparables des autres composantes.

Fidélité inter-cotateurs pour les techniques de seuillage et de délimitation manuelle

L'indice de fidélité est estimé à partir d'une ANOVA à un facteur de classification. Le facteur cotateur ($k = 2$) est considéré comme un facteur fixé et l'unité d'analyse est une évaluation du volume du cerveau ($n = 10$) par cotateur. La corrélation intraclasse (CIC) est calculée par le rapport suivant : (Variance Inter-Variance Intra) / Variance Inter + ($k-1$) (Variance Intra) – où k est le nombre de cotateurs (Shrout et Fleiss, 1979). La probabilité pour la CIC est fournie par celle du F . Ce

coefficient a été choisi parce qu'il ne traduit pas uniquement un ordre quantitatif pour les volumes des cerveaux, comme cela le serait par une simple corrélation de Bravais-Pearson (Bartko, 1976). Il prend en compte la valeur des écarts à la moyenne. Un seuil de probabilité prenant en compte le nombre de tests statistiques ($\alpha = 0.05/3$) a été choisi.

Corrélations entre techniques pour les estimations des volumes du cerveau

Les corrélations intraclassées entre les quatre techniques, considérées deux à deux, ont été estimées à partir d'une ANOVA. Six estimations sont obtenues [$4(4-1)/2 = 6$]. Un seuil de $\alpha = 0.05/6$ a donc été choisi.

RÉSULTATS

Comparaison des volumes moyens du cerveau pour les techniques de seuillage et de délimitation manuelle

Comme la valeur du $F(9, 81) = 20.80$ est significative ($p < 0.0001$), cela signifie qu'au moins deux groupes de mesures diffèrent significativement. Les moyennes des techniques de seuillage dites semi-automatiques sont très proches, elles varient de 1218 cm^3 (écart type (s) = $\pm 110,5 \text{ cm}^3$) à 1233 cm^3 ($s = \pm 86,2 \text{ cm}^3$). Les moyennes volumiques évaluées sur les mesures obtenues par la technique de seuillage dite automatique et celles issues de la technique par délimitation manuelle sont également peu différentes (tableau 1 et figure 2). Elles varient, respectivement, de 1305 cm^3 ($s = \pm 100.6 \text{ cm}^3$) à 1311 cm^3 ($s = \pm 84.48 \text{ cm}^3$). La valeur du coefficient de variation est dépendante de la technique (tableau 1 et figure 2).

Puisque le F est significatif, les séries de mesures du volume cérébral ont été comparées deux à deux. Les valeurs des probabilités *post-hoc* estimées pour le test de Scheffé sont significatives ($p < .005$) pour 24 paires de moyennes sur 45. Les deux techniques semi-automatiques ne fournissent pas des mesures du volume du cerveau statistiquement diffé-

rentes. Il en est de même pour la technique automatique et celle par délimitation manuelle. En revanche, la technique de délimitation manuelle et celle dite automatique procurent, toutes les deux, des mesures du volume cérébral significativement différentes de celles obtenues à partir des deux autres techniques. Ces volumes sont plus importants pour ces dernières techniques que pour celles dites "semi-automatiques" (voir tableau 1). Quelle que soit la technique considérée, notons qu'il n'y a pas de différence *inter-* ou *intra-*cotateurs pour l'évaluation des volumes cérébraux.

Figure 2. Moyenne (m), écart type (s), erreur type (s/n) pour le volume cérébral (cm^3) pour quatre techniques en fonction des cotateurs c1 et c2 et de la double cotation 1 et 2, pour la technique par délimitation manuelle.

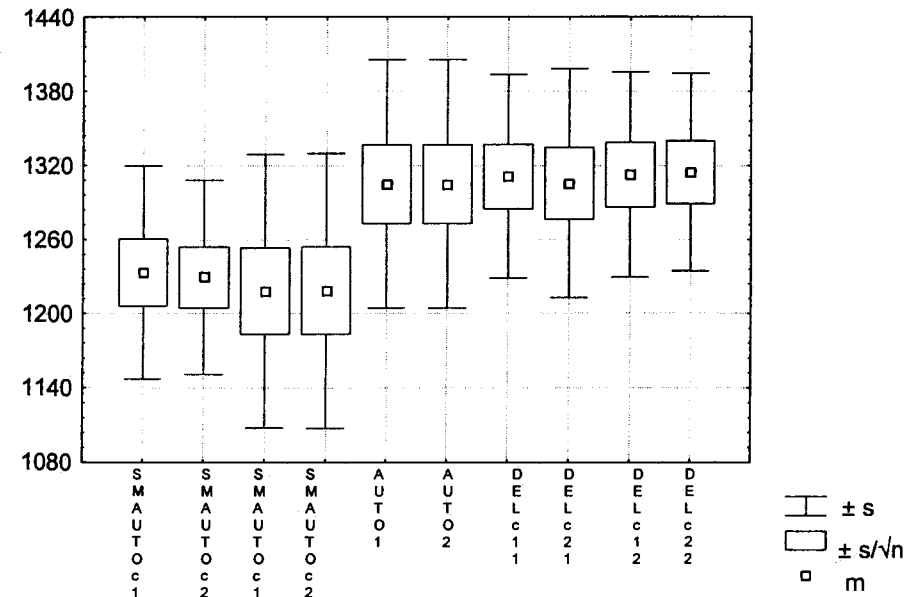


Figure 2. Mean (m), standard-deviation (s), standard error (s/n) for cerebral volume (cm^3) from four techniques according to the c1 or c2 cotateurs and the double cotation 1 or 2, for the manual contour technique.

Tableau 1. Moyenne, valeur minimale et maximale, écart type et coefficient de variation pour les estimations du volume cérébral (cm³) par les différentes techniques à partir d'IRM de cerveaux (N = 10).

Technique	Moyenne	Minimum	Maximum	Ecart type	Coefficient de variat.
SMAUTO Ac1	1233	1104	1351	86.2	0.0699
SMAUTO Ac2	1229	1101	1351	78.5	0.0638
SMAUTO Bc1	1218	1024	1400	110.5	0.0907
SMAUTO Bc2	1219	1023	1398	111.2	0.0912
AUTO1	1305	1134	1469	100.6	0.0770
AUTO2	1305	1134	1469	100.6	0.0770
DELC1.1	1311	1204	1404	82.3	0.0627
DELC2.1	1313	1206	1416	83.1	0.0632
DELC1.2	1305	1173	1427	92.6	0.0709
DELC2.2	1315	1202	1422	79.9	0.0607

Table 1. Mean, minimal and maximal values, standard deviation and coefficient of variation for cerebral volume evaluations (cm³) for brain MRI obtained from different techniques (N = 10).

Décomposition des sources de la variance pour les mesures du cerveau avec la technique par délimitation manuelle

Le tableau 2 présente la décomposition de la variance en fonction des différentes sources de variation. Les résultats indiquent que 97% de la variance sont attribuables aux différences vraies entre cerveaux. L'ensemble de la variance erreur n'est donc que de 3%.

Fidélité inter-cotateurs pour chacune des techniques non automatiques

Les coefficients de corrélation intraclass inter-cotateurs (tableau 3) pour les techniques semi-automatiques sont élevés : 0.98 et 0.99 ($p < 10^{-6}$). Celui pour la technique par délimitation manuelle est de 0.99 ($p < 10^{-6}$).

Tableau 2. Décomposition de la variance pour la méthode par délimitation manuelle : Cotateurs × Cerveaux × 1ère ou 2ème cotation.

Source de la variance	Somme des carrés	% du total	dl
cotateurs	34,4	0,0013	1
cerveaux	252288,5	97,68	9
cotations	278,2	0,107	1
cotateurs × cerveaux	2885,8	1,117	9
cotateurs × cotations	150,2	0,058	1
cerveaux × cotations	1672,5	0,648	9
cotateurs × cerveaux × cotations	970,7	0,376	9
Total	254032,0	100,0000	39

Table 2. Variance components for the manual contour technique: Rater × Brain × First or second cotation.

Tableau 3. Indice de fidélité inter-cotateurs pour les techniques semi-automatiques A et B.

	Somme des carrés Effet	dl inter	Somme des carrés Erreur	dl intra	F	Corrélation intra-classe	P
SEM A	13480.89	9	115.82	10	116.39	0.98	$< 10^{-6}$
SEM B	24575.86	9	1.416	10	17319.14	0.99	$< 10^{-6}$
DEL	128248.6	9	712.96	10	199.87	0.99	$< 10^{-6}$

Table 3. Inter-rater reliability coefficient for A and B "semi-automatic" techniques.

Corrélations entre les différentes techniques

Les valeurs des coefficients de corrélation intraclasse (tableau 4), entre les deux techniques "semi-automatiques" ($r = 0.81$) ou entre la technique "automatique" et de délimitation manuelle ($r = 0.92$), sont élevées et statistiquement significatives au seuil de signification de $\alpha/6$ ($0.05/6 = 0.00083$; correction de Bonferroni) prenant en compte les 6

Tableau 4. Corrélations intra-classes pour les estimations volumiques en fonction des quatre techniques.

(* Significatif au seuil: $\alpha/6 = 0,05/6 = 0,0083$; méthode de Bonferroni.)

Techniques	Somme des carrés Effet	dl inter	Somme des carrés Erreur	dl intra	F	Corréla- tion intra- classe	P
SEM A SEM B	17134.43	9	1788.14	10	9.58	0.81	0.0007*
SEM A AUTO	16178.15	9	3309.25	10	4.89	0.66	0.0100
SEM A DEL	12649.49	9	4151.77	10	3.05	0.51	0.0400
SEM B AUTO	21850.99	9	4230.76	10	5.16	0.68	0.0080*
SEM B DEL	17915.66	9	5516.90	10	3.25	0.53	0.0400
AUTO DEL	16394.72	9	669.83	10	24.47	0.92	0.00001*

Table 4. Intraclass correlation for volumic estimations according to four techniques.

(* Level of significance: $\alpha/6 = 0.05/6 = 0,083$; Bonferroni's method.

tests statistiques appliqués (voir Dunn, 1961). Les coefficients de corrélations entre chacune des techniques "semi-automatiques" et les autres techniques sont faibles et deux de ces quatre coefficients ne sont pas statistiquement significatifs (tableau 4).

DISCUSSION

L'objectif de ce travail était de comparer quatre techniques d'estimation du volume du cerveau à partir de mesures d'images produites par des procédures IRM. Il s'agissait de savoir (1) si les volumes cérébraux estimés étaient identiques quelle que soit la technique employée, (2) si la fidélité des mesures était satisfaisante d'un point de vue biométrique pour chaque technique. Si elle l'était, alors il était légitime de s'interroger sur le fait que ces quatre techniques puissent être interchangeables. Il s'agissait enfin d'estimer la variance commune à ces techniques.

Les résultats obtenus indiquent que l'évaluation du volume cérébral peut varier significativement selon la technique considérée. Ainsi, si les techniques considérées par groupe de deux sous-ensembles, c'est-à-dire deux techniques "semi-automatiques" d'une part, celle par délimitation manuelle et celle automatique d'autre part, donnent des moyennes volumiques non différentes statistiquement, les deux dernières techniques fournissent des valeurs qui varient significativement des valeurs obtenues par les deux premières techniques. Elles sont, en moyenne, supérieures à celles fournies par les techniques "semi-automatiques". La valeur du volume cérébral est donc dépendante de la technique de mesures employée.

Cependant, pour chacune de ces techniques, les coefficients de corrélation intraclasse estimant la fidélité des mesures, entre cotateurs, sont supérieurs à 0,97. Ils sont tous statistiquement significatifs. De plus, pour la technique par délimitation manuelle, la variance due aux erreurs dans la répétition d'une mesure, par un même cotateur, est de 0,107, ce qui est peu important. L'ensemble de ces estimations indique qu'indépendamment du cotateur et/ou de la technique utilisée, du point de vue de la biométrie cérébrale, l'évaluation du volume du système nerveux central est satisfaisante pour la fidélité des mesures. En d'autres termes,

la variance due aux erreurs de mesures varie peu, entre deux cotateurs, pour une même technique. Elle reste très faible quand un même cotateur duplique ses propres mesures au moyen d'une même technique. Pour les trois techniques de mesures utilisées dans ce travail (la technique automatique étant exclue puisque sa fidélité est de 1), les variations dans les mesures ne sont que très faiblement dépendantes du cotateur. Chacune de ces techniques est donc reproductible.

Les différences trouvées dans les volumes cérébraux moyens, leurs variances et leurs coefficients de variation semblent donc devoir être attribuées aux dissemblances dans la conception des techniques. Il est clair que le seuillage préliminaire, pour la reconstruction des cerveaux, élimine plus de voxels dans les techniques "semi-automatiques" que dans les autres. Ce phénomène pourrait être la cause de la valeur plus faible des volumes cérébraux estimés à partir de ces techniques. Notons que la technique dite "automatique" et celle par délimitation manuelle fournissent des résultats proches de ceux trouvés par Blatter et al. (1995) qui ont utilisé une technique multispectrale sur les cerveaux de femmes de leur échantillon. Cette dernière exige deux types d'acquisition volumique en IRM 3D et elle est supposée – sous réserve d'études ultérieures – être plus proche des structures anatomiques. Néanmoins, la différence la plus importante observée, ici, dans les volumes cérébraux moyens, est de 195 cm³ (entre les techniques DELc2.1 et SMAUB1), la différence la plus faible étant de 72 cm³ (entre les techniques AUTO2 ou DELc1.2 ou c2.2 et SMAUA1). Ces différences de volumes associées aux techniques pourraient ne pas être négligeables d'un point de vue anthropométrique et/ou neuropsychologique. Effectivement, Hasboun et al. (1996, p. 48) trouvent une différence de 92 cm³ statistiquement significative, entre des cerveaux des participants féminins et masculins de leur échantillon, en utilisant une technique proche de la SEMAUA1 (voir Méthode). Dans leur étude, et sous réserve des erreurs d'échantillonnage, les volumes des cerveaux des femmes sont significativement plus faibles que ceux des hommes. Ces 92 cm³ sembleraient donc pouvoir être interprétés comme une différence en relation avec la variation du sexe et suggèrent qu'une telle différence pourrait avoir une valeur biologique. Ce constat implique que les inégalités, dans les estimations des volumes cérébraux, dépendantes des disparités dans les techniques de mesures, devraient être

considérées avec attention. L'étude des corrélations entre techniques apporte des précisions numériques sur l'ampleur de leur variation.

Les valeurs des corrélations entre techniques indiquent (tableau 4), que la technique "automatique" et celle par délimitation manuelle peuvent être considérées comme interchangeables ($r = 0.92$). Il en est de même pour les deux techniques semi-automatiques ($r = 0.81$), tandis que les méthodes "semi-automatiques" ne sont pas interchangeables avec les autres ($r = 0.66, 0.51$ et 0.53), exceptée la technique "semi-automatique B" avec celle automatique ($r = 0.68$). Bien que ce dernier coefficient soit statistiquement significatif, ces deux techniques ne présentent que 68% de variance commune, ce qui n'est pas très satisfaisant. Elles ne devraient donc pas être utilisées indifféremment. Les deux techniques "semi-automatiques" ont en commun 81% de variance, ce qui n'est pas suffisant pour une utilisation indifférenciée. D'un point de vue biométrique, seule la technique automatique et celle par délimitation manuelle offrent une corrélation intraclasse, donc une similitude dans les mesures, satisfaisante bien qu'imparfaite (92% de variance commune). De fait, si chacune de ces techniques permet une mesure opérationnelle reproductible du volume du cerveau, les volumes obtenus ne sont pas identiques selon la technique adoptée, et leur précision (ou leur sensibilité) est variable puisque les coefficients de variation ne sont pas identiques.

La lecture de la littérature publiée révèle que les méthodes employées pour l'acquisition des IRM et pour l'évaluation des structures cérébrales (Constant et Ruther, 1996, par exemple pour le CC dans l'étude du dimorphisme sexuel) ne sont pas identiques, selon les équipes de recherche, bien que certaines aient une même problématique. Prenons pour exemple les travaux centrés sur l'étude du volume des ganglions de la base (noyau caudé, putamen et pallidum) chez des patients présentant des troubles bipolaires. L'étude des lésions dans les ganglions de la base indiquerait que cette structure est impliquée dans des fonctions motrices et dans des processus cognitifs (voir par exemple, Austin et Mitchell, 1995). Les variations volumiques, chez ces patients, sont considérées comme l'indicateur du niveau de lésions, liées à la pathologie, par comparaison à un groupe de sujets non atteints. Deux études (Aylward et al., 1994 ; Swayze et al., 1992) sont présentées comme fournissant des

résultats contradictoires (voir par exemple, Videbech, 1997), Aylward et al. trouvant, pour leurs échantillons d'hommes, des estimations volumiques du noyau caudé significativement plus importantes chez leurs sujets atteints que chez leurs sujets témoins. Ces résultats n'avaient pas été observés par Swayze et al. (1992). Or, dans les tentatives d'explication de ces différences, les techniques d'IRM et d'analyses d'image ne sont pas considérées. Pourtant, (1) l'intensité du champs magnétique est différente (0,5 Tesla [Swayze et al., 1992] ; 1,5 Tesla [Aylward et al., 1994]), (2) les coupes sont faites selon un plan coronal [Aylward et al.], ou selon un plan axial [Aylward et al.], (3) l'épaisseur des coupes est différente, (4) les techniques de mesures ne sont pas identiques (l'une avec un planimètre [Aylward et al.] à partir de projections d'images agrandies, l'autre à l'aide d'un ordinateur et par délimitations manuelles), (5) alors que les structures sont mesurées en deux parties différentes (gauche et droite) pour les deux études, les volumes fournis sont donnés pour chacune des parties par Swayze et al., pour la structure en entier par Aylward et al. Des remarques analogues (nombre de Teslas, épaisseur des coupes et techniques de segmentation du CC différents) peuvent être émises pour des études sur la maladie d'Alzheimer et la variation de la surface du CC (Biegon et al., 1994 ; Janowsky et al., 1996). Certes, des auteurs comme Woodruff et al. (1995) introduisent ces variables (et d'autres) en tant que variables indépendantes dans des analyses de régressions multiples. Effectuant des analyses secondes à partir de données publiées sur la schizophrénie et différentes mesures du CC, ils regroupent les résultats de onze études dans ce domaine. Comme, actuellement, il n'y a pas de consensus quant au lien entre schizophrénie et anomalies du CC, les auteurs pensent par ce regroupement de travaux obtenir un meilleur pouvoir statistique dans leurs analyses.

Sans s'interroger en détail, ici, sur la validité de la compilation de données issues d'études méthodologiquement si différentes – (Woodruff et al., 1995, par exemple p. 458, tableau 1, les tests d'égalité des moyennes entre les groupes de patients atteints de Matthew et al. [1985] et de Kelsoe et al. [1988] ou entre ces derniers et ceux de Hauser et al. [1989] indiquent qu'ils diffèrent significativement et donc que ces groupes de patients ne sont pas issus de la même population parente : ces résultats ne devraient pas être traités statistiquement ensemble) – il doit

être souligné qu'une telle pratique ne remplacera jamais la valeur scientifique fournie par une duplication rigoureuse (hypothèse, plan et méthode) des études.

Il est clair que les différences, associées aux techniques elles-mêmes, doivent impérativement être prises en compte quand les résultats entre études sont comparés. Actuellement de telles comparaisons demeurent difficiles.

Parmi les quatre techniques présentées et en l'absence d'un indice de validité externe scientifiquement satisfaisant, si l'une d'entre elles devait être sélectionnée, pour une future étude, ce serait la technique automatique qui serait retenue. Bien qu'un peu moins précise que la technique par délimitation manuelle et que la "semi-automatique A" (son coefficient de variation [7,7 %] est plus faible que celui de la technique "semi-automatique B" [9,1 %]), sa fidélité est parfaite (1) et son temps d'exécution très rapide (3mn).

Les techniques d'imagerie médicale ouvrent des possibilités et des objectifs à l'anthropométrie et la neuropsychologie que la craniométrie traditionnelle ne pouvait pas atteindre. Cette dernière a cependant l'avantage de bénéficier d'une métrique internationale, donc d'une mesure qui rend possible les comparaisons entre travaux. Pour notre propos, si une référence externe *in vivo* du volume du cerveau était accessible, cela permettrait de définir la validité externe des mesures fournies par les différentes techniques d'IRM et de choisir celle qui pourrait être acceptée par tous. Dans l'état présent des connaissances, une telle référence n'est pas disponible et apparaît comme difficile à concevoir avec les seuls outils existants.

ABSTRACT

Recent developments in Three Dimensional (3D) Magnetic Resonance Imaging (MRI) open new research horizons, allowing further work in cerebral biometry. Using MRI, 124 contiguous slices of the brain can be obtained and the brain can then be reconstructed in 3D for surface area or volume evaluations. In this study, images of brains from 10 healthy

females aged 20,15 years ($s \pm 1,14$ years) were analyzed to perform cerebral volume evaluations. The results of four techniques were compared. One is based on manual contour of the brain, two are semi-automatic and one is automatic. The cerebral volumes obtained by the two semi-automatic techniques did not differ statistically ($1233 \pm s: 86,2 \text{ cm}^3$ and $1218 \pm s: 110,5 \text{ cm}^3$), they nor did differ when the manual and automatic techniques were compared ($1305 \pm s: 100,6 \text{ cm}^3$ and $1311 \pm s: 82,3 \text{ cm}^3$). On the other hand, the latter two techniques gave volumes statistically different from those given by the first two. The intraclass correlation coefficients (ICC) in assessing interrater reliability were over 0,97. For correlations between techniques, three ICC were below 0,66 and not statistically significant. Only the manual contour and automatic techniques showed a significant correlation, giving a proportion of common variance (92%) that was satisfactory but not biometrically perfect. Although reliable, these techniques are not interchangeable. More generally, the results of studies using MRI and other techniques of brain measurement should not be compared. A single and relevant technique should be found and used for scientific and clinical purposes.

BIBLIOGRAPHIE

- Andreassen, N.C., Cizadlo, T., Harris, G., Swayze, V., O'Leary, D.S., Cohen, G., Ehrhardt, J., & Yuh, W.T. (1993). Voxel processing techniques for the antemortem study of neuroanatomy using magnetic resonance imaging. *Journal of Neuropsychiatry Clinical Neuroscience*, 5 (2), 121-130.
- Austin, M.P., & Mitchell, P. (1995). The anatomy of melancholia: Does frontal-subcortical pathophysiology underpin its psychomotor and cognitive manifestations? *Psychological Medicine*, 25, 665-672.
- Aylward, E.H., Roberts-Twillie, J.V., Barta, P.E., Kumar, A.J., Harris, G.J., Geer, M., Peyser, C.E., & Pearlson, G.D. (1994). Basal ganglia volumes and white matter hyperintensities in patients with bipolar disorder. *American Journal of Psychiatry*, 151, 687-693.
- Bartko, J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin*, 83, 762-765.
- Biegón, A., Eberling, J.L., Richardson, B.C., Roos, M.S., Wong, S.T.S., Reed, B.R., & Jagust, W.J. (1994). Human corpus callosum in aging and Alzheimer's disease: A Magnetic Resonance Imaging study. *Neurobiology of Aging*, 15 (4), 393-397.
- Blatter, D.D., Bigler, E.D., Gale, S.D., Johnson, S.C., Anderson, C.V., Burnett, B.M., Parker, N., Kurth, S., & Horm, S.D. (1995). Quantitative volumetric analysis of brain MR: Normative database spanning 5 decades of life. *American Journal of Neuroradiology*, 16 (2), 241-251.
- Botez, M.I. (1996). *Neuropsychologie clinique et neurologie du comportement* (2ème ed.). Paris: les Presses de l'Université de Montréal, Masson.
- Cline, H.E., Lorenzen, W.E., Kikinis, R., & Jolesz, F. (1990). Three-dimensional segmentation of MR images of the head using probability and connectivity. *Journal of Computer Assisted Tomography*, 14 (6), 1037-1045.
- Constant, D., & Ruther, H. (1996). Sexual dimorphism in the human corpus callosum? A comparison of methodologies. *Brain Research*, 727, 99-106.
- DeCarli, C., Maisog, J., Murphy, D.G.M., Tiechberg, D., Rapoport, S.I., & Horwitz, B. (1992). Method for quantification of brain, ventricular, and subarachnoid CSF volumes from MR images. *Journal of Computer Assisted Tomography*, 16, 274-284.
- Dunn, O.J. (1961). Multiple comparison procedure among means. *Journal of the American Statistical Association*, 56, 52-64.
- Elkis, H., Friedman, L., Wise, A., & Meltzer, H.Y. (1995). Meta-analyses of studies of ventricular enlargement and cortical sulcal prominence in mood disorders. Comparisons with controls or patients with schizophrenia. *Archives of General Psychiatry*, 52, 735-746.
- Fox, P.T., Parsons, L.M., & Lancaster, J.L. (1998). Beyond the single study: function/location metaanalysis in cognitive neuroimaging. *Current Opinion in Neurobiology*, 8, 178-187.
- Giedd, J., Castellanos, F.X., Rajapakse, J.C., Vaituzis, A.C., & Rapoport, J. (1997). Sexual dimorphism of the developing human brain. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 21, 1185-1201.
- Giedd, J.N., Castellanos, F.X., Casey, B.J., Kozuch, P., King, A.C., Hamburger, S.D., & Rapoport, J.L. (1994). Quantitative morphology of the corpus callosum in attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 151, 665-669.
- Goldberg, T.E., Torrey, E.F., Berman, K.F., & Weinberger, D.R. (1993). Relations between neuropsychological performance and brain morphological and physiological measures in monozygotic twins discordant for schizophrenia. *Psychiatry Research: Neuroimaging*, 55, 51-61.
- Hasboun, D., Chantôme, M., Zouaoui, L.M., Sahel, M., Baulac, M., Marceault, C., Duyme, M., & Dormont, D. (1996). Evaluation du volume cérébral : Reproductibilité et Précision d'une technique 3D IRM. *Bulletin et Mémoire de la Société d'Anthropologie de Paris*, n.s.t. 8, 1-2, 43-56.

- Hauser, P., Dauphinais, D., Berrettini, W., DeLisi, L.E., Gelernter, J., & Post, R.M. (1989). Corpus callosum dimensions measured by magnetic resonance imaging in bipolar affective disorder and schizophrenia. *Biological Psychiatry*, *26*, 659-668.
- Hines, M., Chiu, L., McAdams, L. A., Bentler, P. M., & Lipcamon, J. (1992). Cognition and the corpus callosum: Verbal fluency, visuospatial ability, and language lateralization related to midsagittal surface areas of callosal subregions. *Behavioral Neuroscience*, *106* (1), 3-14.
- Hynd, G.W., Semrud-Clikeman, M., Lorys, A.R., Noverly, E.S., & Eliopoulos, D. (1991). Corpus callosum morphology in attention deficit-hyperactivity disorder (ADHD): morphometric analysis of MRI. *Journal of Learning Disabilities*, *24*, 141-146.
- Janowsky, J.S., Kaye, J.A., & Carper, R.A. (1996). Atrophy of the corpus callosum in Alzheimer's disease versus healthy aging. *Journal of the American Geriatrics Society*, *44*, 798-803.
- Jernigan, T.L., Press, G.A., & Hesselink, J.R. (1990). Methods for measuring brain morphologic features on magnetic resonance images: validation and normal aging. *Archives of Neurology*, *47*, 27-32.
- Kelsoe, J.R., Cadet, J.L., Pickar, D., & Weinberger, D.R. (1988). Quantitative neuroanatomy in schizophrenia. *Archives of General Psychiatry*, *45*, 533-541.
- Kohn, M.I., Tanna, M.K., Herman, G.T., Resnick, S.M., Mozley, R.E., Gur, R.E., Alavi, A., Zimmerman, R.A., & Gur, R.C. (1991). Analysis of brain and cerebrospinal fluid volumes with MR imaging, I: Methods, reliability, and validation. *Radiology*, *178* (1), 115-122.
- Lawrie, S.M., & Abukmeil, S.S. (1998). Brain abnormality in schizophrenia. *British Journal of Psychiatry*, *172*, 110-120.
- Matthew, R.J., Partain, C.L., Prakash, R., Kulkarni, M.V., Logan, T.P., & Wilson, W.H. (1985). A study of the septum pellucidum and corpus callosum in schizophrenia: a preliminary study. *Acta Psychiatrica Scandinavica*, *72*, 414-421.
- Maxwell, S.E., & Delaney, H.D. (1990). *Designing experiments and analyzing data*. Belmont: Wadsworth, Inc.
- Murphy, D.G.M., DeCarly, C., Shapiro, M.B., Rapoport, S.I., & Horwitz, B. (1992). Age related differences in volumes of subcortical nuclei, brain matter, and cerebrospinal fluid in healthy men as measured with magnetic resonance imaging. *Archives of Neurology*, *49*, 839-845.
- Pozzilli, C., Bastianello, S., Padovani, A., Passafiume, D., Millefiorini, E., Bozzao, L., & Fieschi, C. (1991). Anterior corpus callosum atrophy and verbal fluency in multiple sclerosis. *Cortex*, *27*, 441-445.
- Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *2*, 420-428.

- Soininen H.S., Partanen, K., Pitkanen, A., Vainio, P., Hanninen, T., Hallikainen, M., Koivisto, K., & Riekkonen, P.J. (1994). Volumetric MRI analysis of the amygdala and the hippocampus in subjects with age-associated memory impairment: Correlation to visual and verbal memory. *Neurology*, *44*, 1660-1668.
- Statistica (CSS). (1991). New York: Statsoft.
- Swayze, V.W., Andreasen, N.C., Alliger, R.J., Yuh, W.T.C., & Ehrhardt, J.C. (1992). Subcortical and temporal structure in affective disorder and schizophrenia: A Magnetic Resonance Imaging study. *Biological Psychiatry*, *31*, 221-240.
- Videbech, P. (1997). MRI findings in patients with affective disorder: a meta-analysis. *Acta Psychiatrica Scandinavica*, *96*, 157-168.
- Ward, K.E., Friedman, L., Wise, A., & Schulz, S.C. (1996). Meta-analysis of brain and cranial size in schizophrenia. *Schizophrenia Research*, *22*, 197-213.
- Weinberger, D.R. (1996). On the plausibility of "the neurodevelopmental hypothesis" of schizophrenia. *Neuropsychopharmacology*, *14*, 1S-11S.
- Woodruff, P.W.R., McManus, I.C., & David, A.S. (1995). Meta-analysis of corpus callosum size in schizophrenia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *58*, 457-461.