

Selection of multiscale functional brain connections for early diagnosis of Alzheimer's disease in multicentric studies

Christian Dansereau

Université de Montréal

5 décembre 2014

Overview

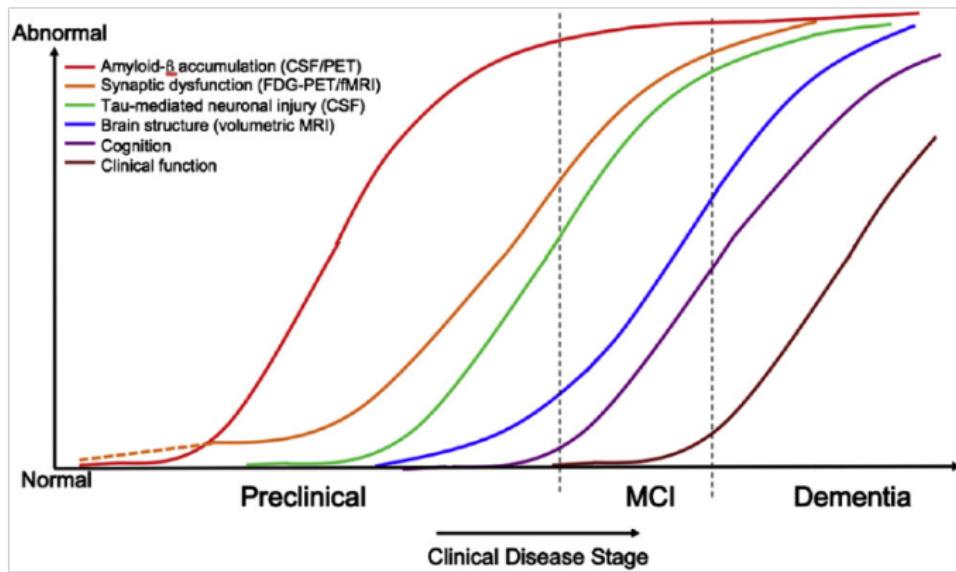
- 1 Contexte général
- 2 Optimisation des étapes de prétraitement
- 3 Étude de connectivité multicentrique en IRMf
- 4 Classification et analyse multivariée

Definition de la maladie d'Alzheimer

- Maladie neurodégénérative incapacitante.
- Cause la plus fréquente de démence chez l'homme (36M).
- Bases physiologiques encore mal comprises.

Progression de la maladie d'Alzheimer

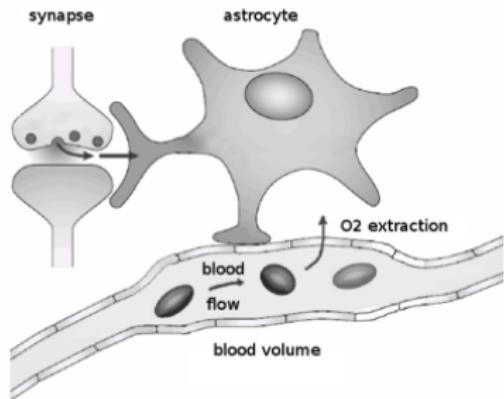
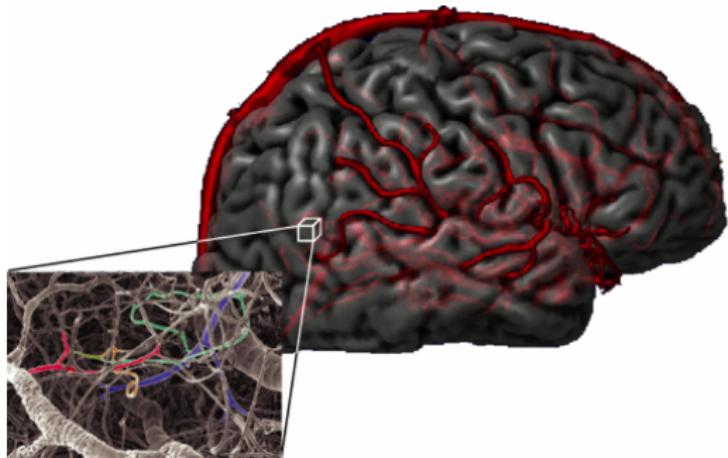
R.A. Sperling et al. / Alzheimer's & Dementia 7 (2011) 280–292



Imagerie par résonance magnétique (IRM)

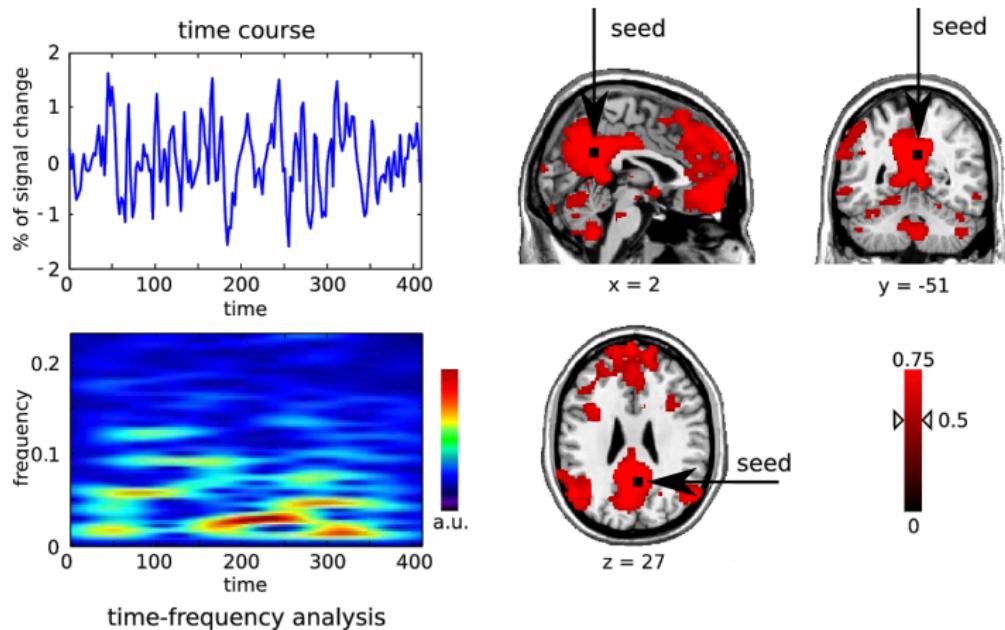


IRM fonctionnelle (IRMf)



Adapté de Heeger 2002.

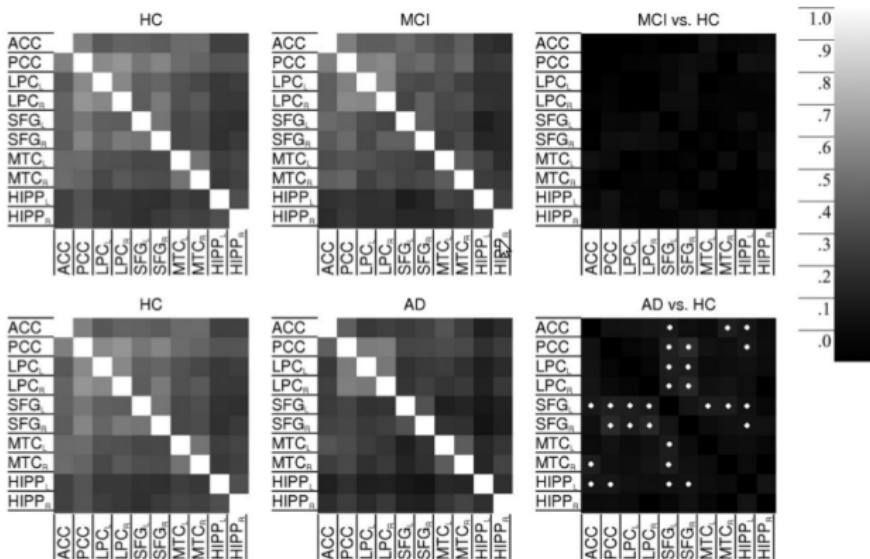
Connectivité au repos (resting-state)



Tiré de (Bellec, Springer 2011).

Le cortex cingulaire postérieur est utilisé comme cible pour une carte de connectivité fonctionnelle individuelle en IRMf.

Comparison between healthy subjects, patients with MCI and patients with a dementia of the AD type



A two sample *t*-test between healthy elderly subjects ($n = 21$), patients with MCI ($n = 17$) patients with a DAT ($n = 15$), using various connexions within the default-mode network. Statistics corrected for multiple comparisons with the false-discovery rate ($q \leq 0.05$).

From Koch et al., 2012, Neurobiology of aging, 466-478.

Objectifs

- ① *L'analyse de données en IRMf n'est pas optimisée pour des populations âgées (CNE, pMCI, pDAT).*

Objectif 1: Optimiser la chaîne de traitements des données d'IRMf pour ces populations.

- ② *La collecte de grands échantillons passe par l'acquisition de données dans plusieurs centres d'imagerie.*

Objectif 2: Evaluer la faisabilité d'études multicentriques en IRMf.

- ③ *De nombreuses mesures peuvent être générées en IRMf.*

Objectif 3: Élaborer un algorithme d'identification de biomarqueurs de la maladie d'Alzheimer.

Section 2

Optimisation des étapes de prétraitement

Correction des artefacts de mouvement

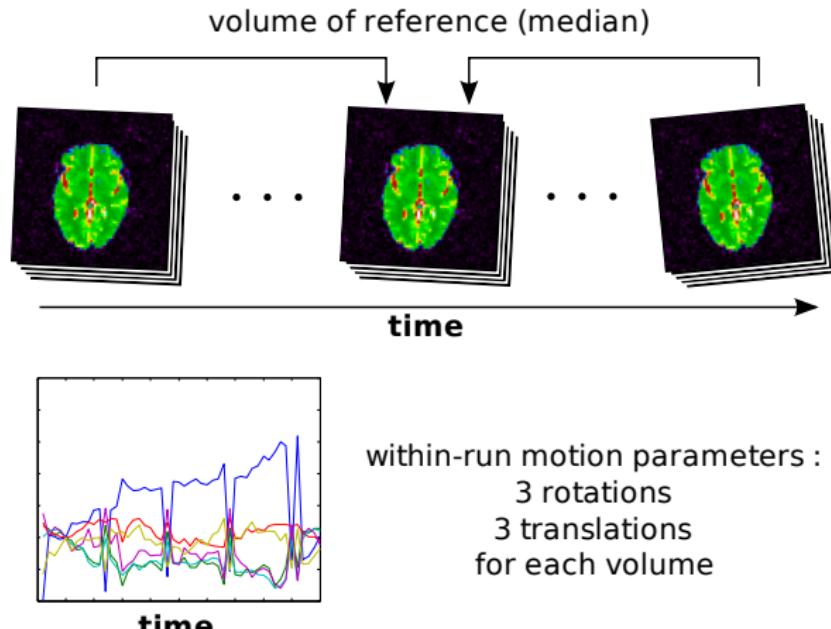
Les mouvements de la tête sont inévitables et présentent l'un des plus grands problèmes en IRMf.

- Possibilité de réaligner les déplacements de la tête.
- Ne corrige pas les artefacts induits dans le signal par des inhomogénéités du champ magnétique.
- Le mouvement est plus important/fréquent chez les populations âgées.

Objectifs étude 1

- Évaluer la quantité de mouvement chez les populations âgées ainsi que celles avec un début de démence.
- Évaluer l'impact du prétraitement sur la connectivité.
- Évaluer l'impact du prétraitement sur notre capacité à détecter un effet.

Estimation du mouvement



(from Chaine de prétraitement du "Neuroimaging Analysis Kit" (NIAK)¹).

¹<http://www.nitrc.org/projects/niak/>

Résumé de la quantité de mouvement

Mesure dite de “frame displacement” (FD).

$$FD_i = |\Delta d_x(t)| + |\Delta d_y(t)| + |\Delta d_z(t)| + |\Delta r_x(t)| + |\Delta r_y(t)| + |\Delta r_z(t)|, \quad (1)$$

$$r_x(t) = 50 \left(\frac{2\pi\alpha_x(t)}{360} \right), \quad (2)$$

avec

- $(d_x(t), d_y(t), d_z(t))$ paramètres de translation (mm),
- $(\alpha_x(t), \alpha_y(t), \alpha_z(t))$ paramètres de rotation (degrés),
- Δ la différence entre l'instant t et $t - 1$.

Prétraitement

- Prétraitement standard (sans scrubbing)
- Avec scrubbing ($FD > 0.5$)
- Avec scrubbing ($FD > 0.2$)

Minimum de 50 volumes restant.

Jeux de données

Ce projet provient de l'agrégation de plusieurs jeux de données (5 études différentes)

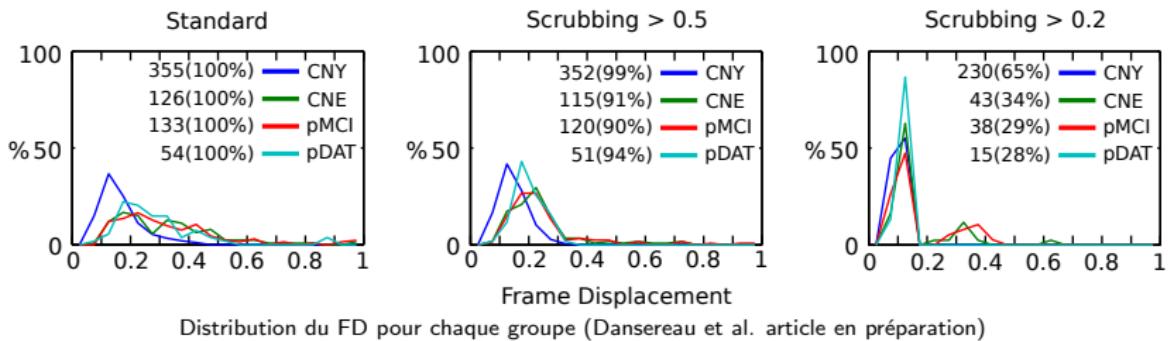
ADNI2 et 4 études basées à Montréal (total: 313 sujets)

- 126 CNE participants (51M, âge = 57-94)
- 133 pMCI (70M, âge = 55-89)
- 54 pDAT (22M, âge = 55-88)

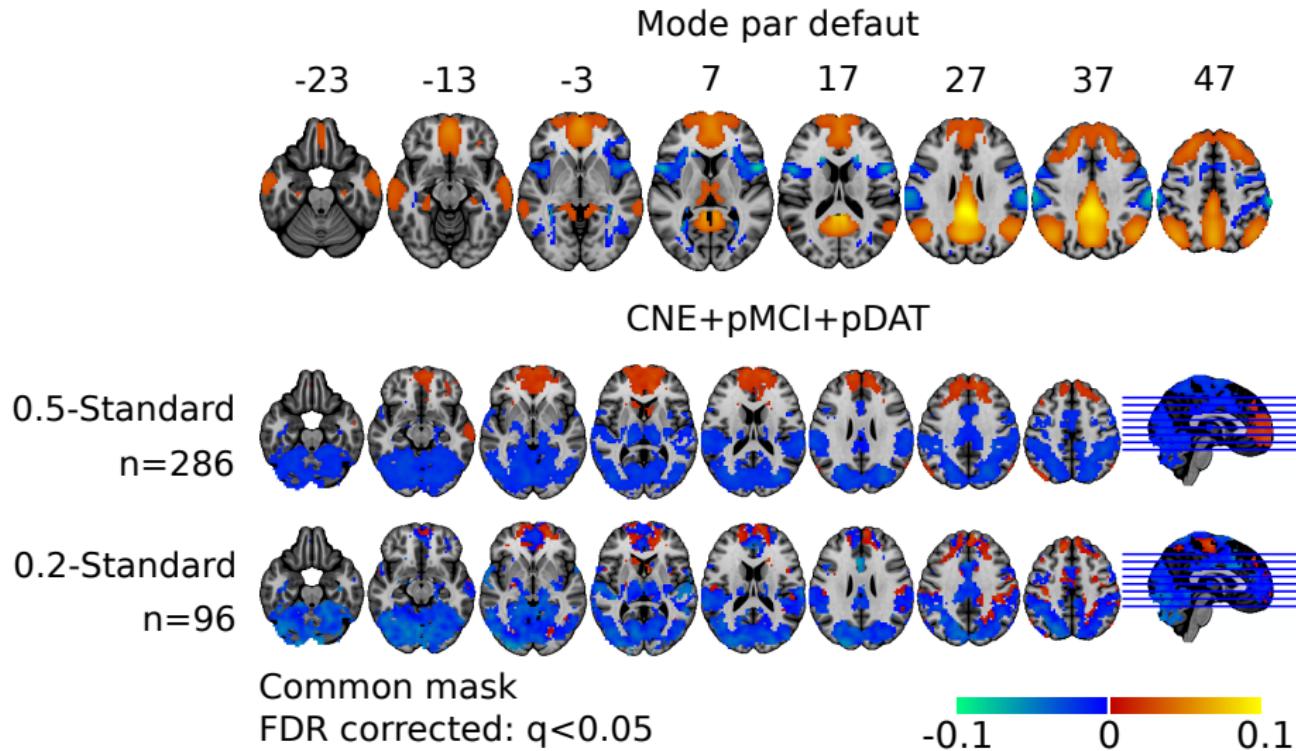
Jeux de données de référence: “1000 functional connectome project”

- 355 jeunes adultes sains (CNY) (150M, âge = 18-46)

Distribution du FD



Impact du prétraitement



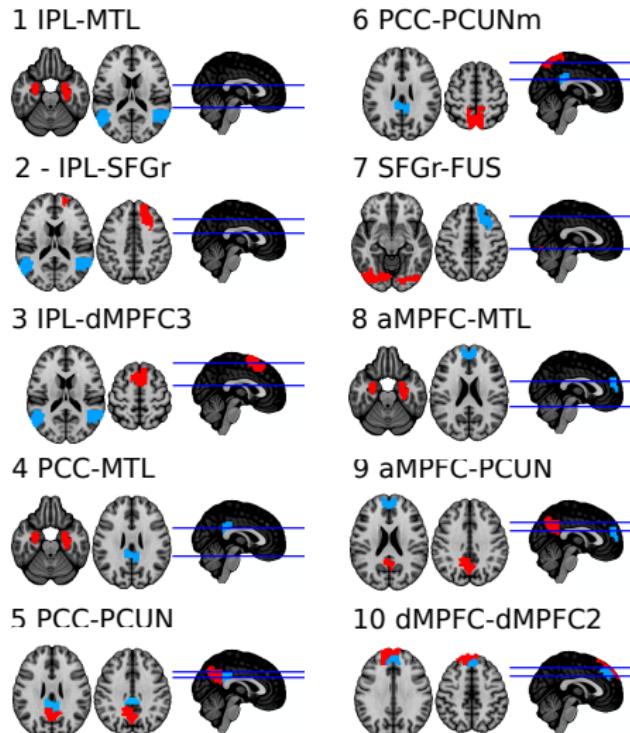
Différence de connectivité entre stratégies de prétraitement (Dansereau et al., manuscript en préparation)

Estimation de la sensibilité de détection

- Différences inter-groupes (pDAT-CNE).
- 10 paires de connexions reproductibles et impactées par la maladie sont sélectionnées sur la base de la littérature.
- Modélisation (FD, âge, sexe) dans un GLM. Combinaison inter-sites par METAL.
- Significativité des résultats obtenue avec un test t de Student.
- La sensibilité du test est évaluée en sous-échantillonnant 70% du jeu de données ($B = 10^4$ échantillons aléatoire).
- Pour chaque échantillon b , on a une p -value p_b^* .
- La sensibilité de la détection est donc estimée par la probabilité que p_b^* soit inférieur à 0.05.

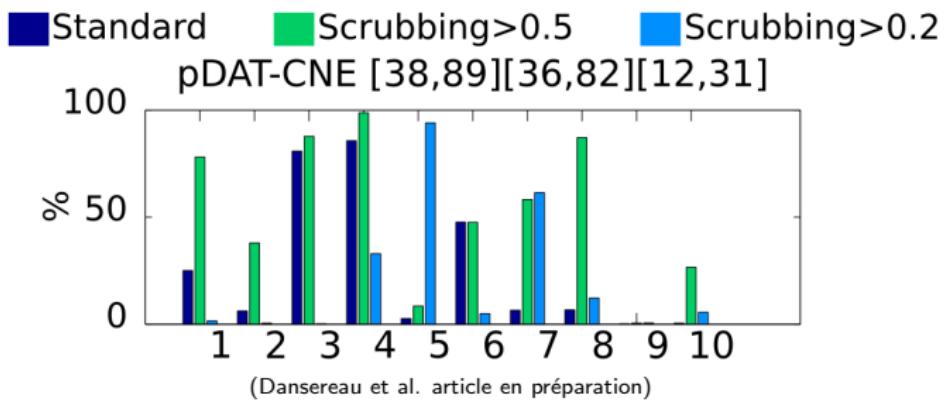
$$\frac{1}{B} \sum_{b=1}^B (p_b^* \leq 0.05). \quad (3)$$

Connexions sélectionnées



Paires de connexions revue de littérature (Dansereau et al. article en préparation)

Pouvoir de détection avec scrubbing



Conclusions étude 1

- Moins de données de meilleure qualité sont préférables à plus de données bruitées.
- Prétraitement avec scrubbing à 0.5 pour les populations âgées, plutôt que 0.2 (seuil recommandé pour jeunes adultes).
- Les techniques actuelles permettent de réduire mais pas d'effacer l'impact du mouvement.
- Travail en cours: (1) précision de la prédiction plutôt que sensibilité.
(2) approche METAL plus agressive.

Section 3

Étude de connectivité multicentrique en IRMf

Multi-sites

Le nombre de sujets augmente mécaniquement la puissance statistique (sensibilité) d'une étude, ce qui peut être réalisé rapidement au moyen d'une étude multicentrique.

- Académique (ex. 1000 functional connectome, ADNI).
- Pharmaceutique (essai clinique).

Source de variance:

- variance d'acquisition (marque et modèle du scanner)
- contexte de l'expérience (instruction au participant, caractéristiques des populations locales)
- variance de l'environnement (son, température).

Faisabilité d'étude IRMf multicentrique

But: quantifier l'impact des différences inter-sites sur la connectivité fonctionnelle.

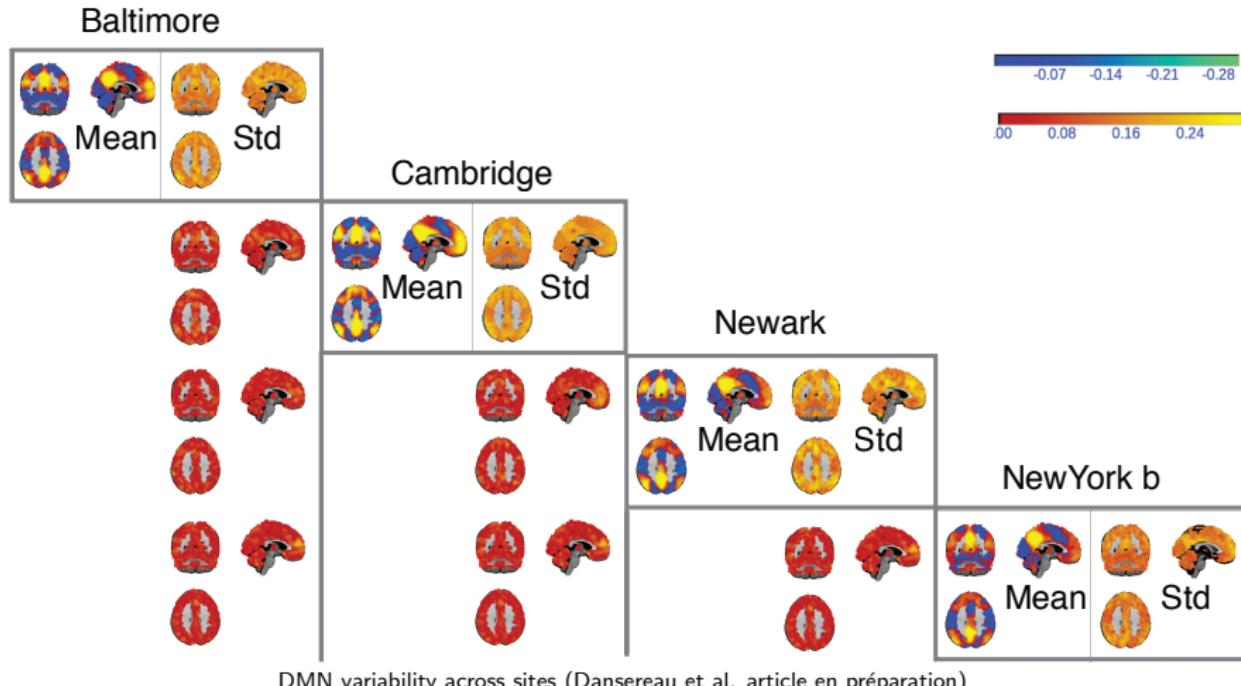
- ① Caractériser l'amplitude des différences (biais) inter-sites.
- ② Quantifier l'impact du biais inter-sites sur le pouvoir de détection d'un effet en IRMf au repos.
- ③ Évaluation de différents modèles statistiques pour capturer le biais inter-sites.

Simulation: jeux de donnée

1000 Functional Connectomes Project (150M, age = 18-46)

- 1 gros site ~ 200 (Cambridge)
- 7 petit sites ~ 20 /site

Résultats: variance inter site du DMN



Simulation: correction multi-site (methode 1)

Introduction de variables dites «dummy» dans le GLM.

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{V}\boldsymbol{\gamma} + \mathbf{E}, \quad (4)$$

- \mathbf{Y} : $N \times 1$, valeurs de connectivité pour la paire (i, j) ,
- \mathbf{X} : $N \times K$, variables explicatives,
- $\boldsymbol{\beta}$: $1 \times K$, valeurs de régression pour chaque variable explicative,
- \mathbf{V} : $N \times S$, chaque colonne code pour un site (0/1),
- $\boldsymbol{\gamma}$: $1 \times S$, moyennes de connectivité par site,
- \mathbf{E} : $N \times 1$, résidus de la régression,

avec N nombre de sujets, K nombre de variables explicatives, S nombre de sites.

Simulation: correction multi-site (methode 2 METAL)

β_v , effect size estimate for site v . (5)

σ_v , standard error for site v . (6)

$w_v = \frac{1}{\sigma_v^2}$, weight estimate for site v . (7)

$\beta = \frac{\sum_v \beta_v w_v}{\sum_v w_v}$, global β . (8)

$\sigma = \sqrt{\frac{1}{\sum_v w_v}}$, global standard error. (9)

$Z = \frac{\beta}{\sigma}$, global Z score. (10)

$p = 2(1 - \phi(|Z|))$, p-value. (11)

Simulation: cohen's d

Pour chaque site, un effet est ajouté à la connectivité de 50% des sujets, sélectionnés aléatoirement (groupe «pathologique»):

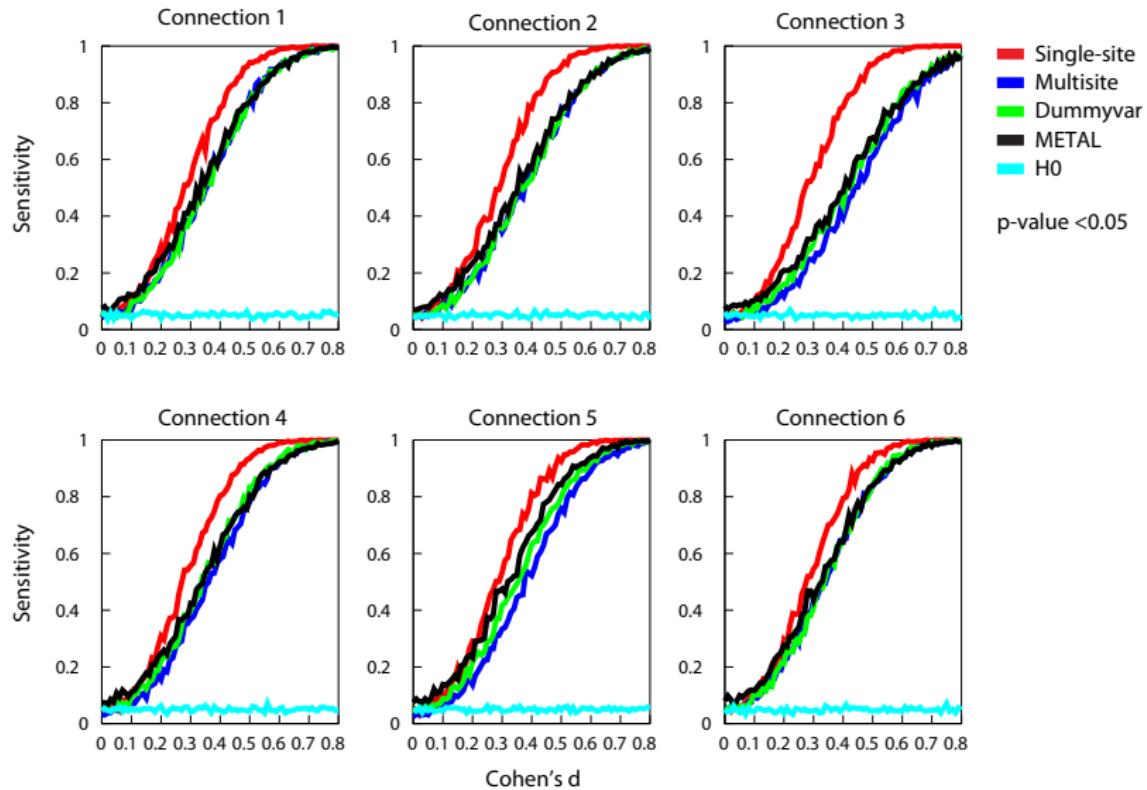
$$y_{i,j} = y_{i,j} + \mu. \quad (12)$$

Le paramètre μ est choisi pour réaliser une certaine taille d'effet (mesuré par le d de Cohen):

$$d = \frac{\mu}{s_{i,j}}, \quad (13)$$

où $s_{i,j}$ est l'écart-type de la connectivité entre les régions i et j pour la population de référence (mono-site, Cambridge). Différentes tailles d'effet sont simulées (de 0 à 0.8).

Résultats: pouvoir de détection avec les deux corrections



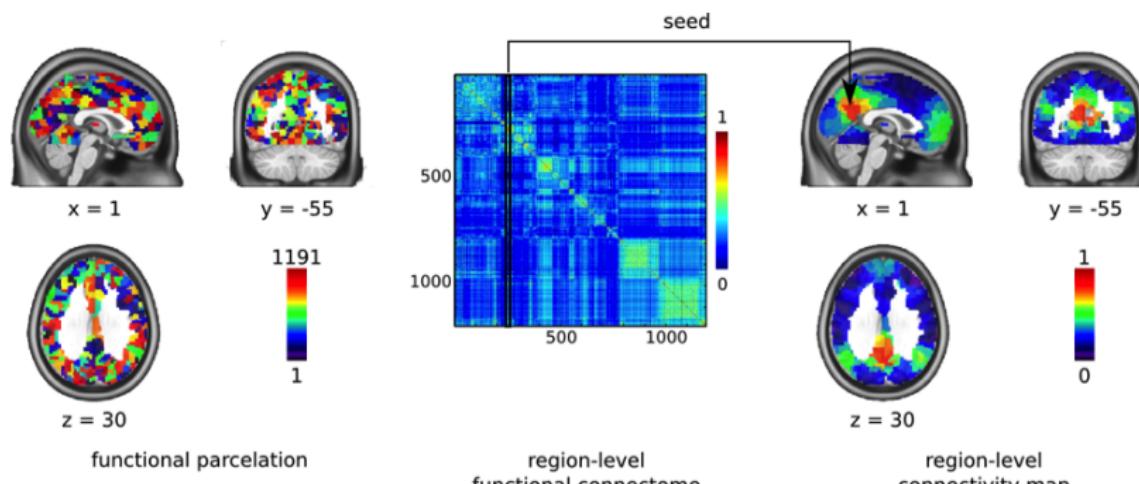
Étude multi-sites conclusion

- Existence d'un biais multi-sites.
- Selon la répartition des sujets, on peut avoir une petite diminution du pouvoir statistique.
- A compléter: exploration de scénarios plus complexes (sites débalancés, petits sites).

Section 4

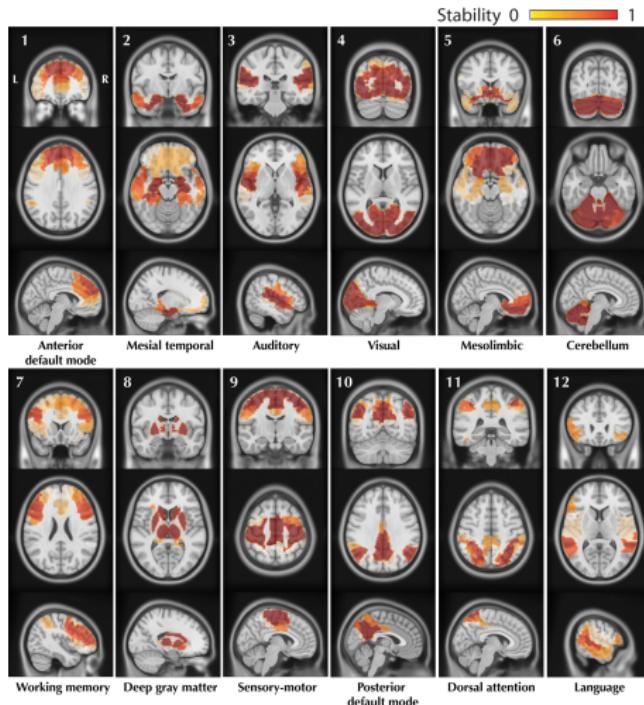
Classification et analyse multivariée

Connectivité au repos



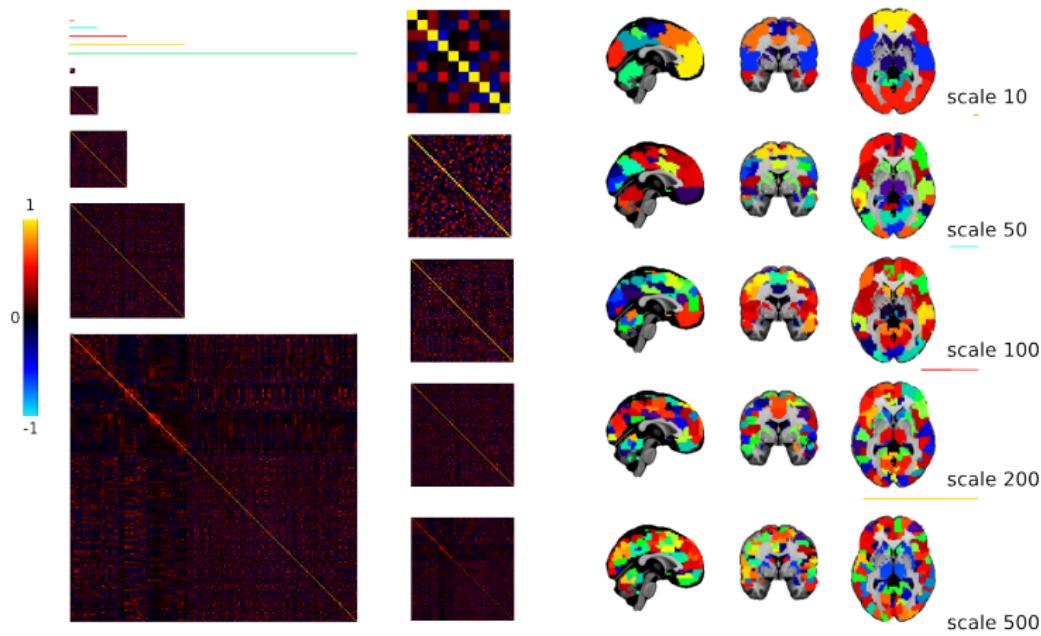
Bellec et al. Neuroimage 2006, connectome fonctionnel.

Décomposition en réseaux



Dansereau et al. 2014 Front. Neurosci. in press.

Connectomes multiéchelles



Bellec et al., International Conference on Resting-state Connectivity, Magdeburg, 2012

Évaluation du pipeline et plan d'analyse

- ① Cobre (74 controles, 72 schizophrénes).
 - Prédiction de la maladie.
- ② Enhanced Nathan Kline Institute-Rockland Sample (NKI) (> 500 subjects).
 - Prédiction de l'âge.
 - Multi protocoles.

Implication dans l'industrie

- Consultation pour des sous-traitants de pharmaceutique (NeuroRx et Biospective).
- Preuve de faisabilité et analyses pour des essais clinique de phase II sur la maladie d'Alzheimer pour le compte de grand joueur de l'industrie pharmaceutique.
- Identification de biomarqueur avec le Canadian Consortium on Neurodegeneration in Aging (CCNA).

Timeline

J'ai identifié trois questions que je compte adresser dans mon PhD:

- Impact du mouvement sur la connectivité (posters présentés à OHBM2013 et AAIC2014 article en préparation soumis avant la fin de l'hiver 2015).
- Faisabilité d'étude multi-sites (poster présentés à OHBM2013 et AAIC2013 article en préparation soumis avant l'été 2015).
- Prediction
 - Expérimenter avec les outils d'apprentissage et identifier des métriques à utiliser.
 - Faire une validation de la stabilité.
 - Application sur données réelles avec Cobre et NKI-RS.
 - Évaluation du pipeline sur un jeu de données final (5 sites) en cross sectionnel et 1 site en longitudinal (prendre fin à l'hiver 2016).
- Je planifie soumettre ma thèse à la fin 2016.

Merci