Web-based Supplementary Materials for "Multiple Testing for Neuroimaging via Hidden Markov Random Field" by Hai Shu, Bin Nan, and Robert Koeppe

1. Web Appendix A: Interpretations of the Ising Model Parameters

For the two-parameter Ising model defined in (1) in the main paper, we can show that

$$\log\left\{\frac{P(\Theta_s=1,\Theta_t=1|\boldsymbol{\theta}_{S\setminus\{s,t\}})}{P(\Theta_s=1,\Theta_t=0|\boldsymbol{\theta}_{S\setminus\{s,t\}})} \times \frac{P(\Theta_s=0,\Theta_t=0|\boldsymbol{\theta}_{S\setminus\{s,t\}})}{P(\Theta_s=0,\Theta_t=1|\boldsymbol{\theta}_{S\setminus\{s,t\}})}\right\} = \begin{cases}\beta, & t\in\mathcal{N}(s),\\ 0, & \text{otherwise.}\end{cases}$$

Therefore, if s and t are neighbors, β is equal to a log odds ratio that describes the association between Θ_s and Θ_t conditional on all the other state variables being withheld. We can see that β reflects how likely the same-state voxels are clustered together. Similarly,

$$\log\left\{\frac{P(\Theta_s=1|\sum_{t\in\mathcal{N}(s)}\Theta_t=0)}{P(\Theta_s=0|\sum_{t\in\mathcal{N}(s)}\Theta_t=0)}\right\}=h,$$

which is the log odds for $\Theta_s=1$ given that $\boldsymbol{\Theta}_{\mathcal{N}(s)}$ are all zero. Thus, $\beta\geqslant 0$ and $h\leqslant 0$ imply the nonnegative dependency of state variables at neighboring voxels. In addition, for a voxel s with m nearest neighbors, we have

$$\begin{aligned}\log\left\{\left(\frac{P(\Theta_s=1|\sum_{t\in\mathcal{N}(s)}\Theta_t=n)}{P(\Theta_s=0|\sum_{t\in\mathcal{N}(s)}\Theta_t=n)}\right)\right.&\\ \left.\Big/\left(\frac{P(\Theta_s=0|\sum_{t\in\mathcal{N}(s)}\Theta_t=m-n)}{P(\Theta_s=1|\sum_{t\in\mathcal{N}(s)}\Theta_t=m-n)}\right)\right\}&\\ =m\beta+2h,&\end{aligned}$$

where n is an integer satisfying $0\leqslant n\leqslant m$, which reflects the log ratio of the cluster effect of signals (nonnulls) relative to the cluster effect of noises (nulls).

2. Web Appendix B: Theoretical Results of the Oracle LIS-Based Procedures for HMRF

In this section, we show the theoretical results of the oracle LIS-based procedures originally for HMC model in Sun and Cai (2009) (Theorems 1 to 4 and Corollary 1) and Wei et al. (2009) (Theorems 1 and 2), including the validity and optimality of the procedures, also hold for our HMRF model. Here, an FDR procedure is called *valid* if it controls FDR at a prespecified level α, and is called *optimal* if it minimizes marginal FNR (mFNR) while controlling marginal FDR (mFDR) at the level α.

Unless stated otherwise, the notation in this section is the same as in Sun and Cai (2009) to which readers are referred. Define $\pi_{ij} = P(\Theta_i = j), i \in S, j = 0, 1$. The model homogeneity, i.e., $\pi_{ij} = \pi_j^{(k)}$ for all i in k-th HMC, is required in Sun and Cai (2009) and in Wei et al. (2009) but fails to hold for HMRF because the boundary voxels and interior voxels have different numbers of neighbors. However, the theory of the oracle procedures still holds for HMRF if we redefine the average conditional cumulative distribution functions (CDFs) of the test statistic $\boldsymbol{T}(\boldsymbol{x}) = \{T_i(\boldsymbol{x}) : i \in S\}$ by

$$G^j(t) = \frac{\sum_{i \in S} \pi_{ij} G_i^j(t)}{\sum_{i \in S} \pi_{ij}}, \tag{B.1}$$

where $G_i^j(t) = P(T_i < t | \Theta_i = j)$.

For HMC model, Sun and Cai (2009) proved the optimality of oracle LIS procedure in their Theorems 1 to 3 and Corollary 1, and its validity in their Theorem 4; Wei et al. (2009) showed the validity of oracle SLIS procedure in their Theorem 1, and both validity and optimality of oracle PLIS procedure in their Theorem 2. Let us keep all the statements in these theorems and corollary by

(i) replacing HMM by HMRF;

(ii) in Corollary 1 of Sun and Cai (2009), replacing the definition of $G^j(t)$ by (B.1) and the equation $g^1(t)/g^0(t) = (1/t)\pi_0/\pi_1$ by $g^1(t)/g^0(t) = (1/t)\sum_{i \in S} \pi_{i0} / \sum_{i \in S} \pi_{i1}$;

(iii) in Theorem 2 of Wei et al. (2009), more precisely stating the optimality of oracle PLIS procedure based on mFDR and mFNR.

For simplicity, we omit all these statements and only provide their proofs in the following.

2.1 *Theorem 1 of Sun and Cai (2009) for HMRF*

Proof. Following the proof of Proposition 1 in Sun and Cai (2007), we have

$$g^0(c)G^1(c) - G^0(c)g^1(c) > 0 \tag{1}$$

and

$$g^0(c)[1 - G^1(c)] - g^1(c)[1 - G^0(c)] < 0. \tag{2}$$

Additionally, by (B.1),

$$\begin{aligned} \mathrm{mFDR}(c) &= \frac{E(N_{10})}{E(R)} = \frac{\sum_{i\in S} P(T_i < c, \Theta_i = 0)}{\sum_{i\in S} P(T_i < c)} \\ &= \frac{\sum_{i\in S} \pi_{i0} G_i^0(c)}{\sum_{i\in S}(\pi_{i0} G_i^0(c) + \pi_{i1} G_i^1(c))} \\ &= \frac{G^0(c)\sum_{i\in S} \pi_{i0}}{G^0(c)\sum_{i\in S} \pi_{i0} + G^1(c)\sum_{i\in S} \pi_{i1}}, \end{aligned}$$

and

$$\begin{aligned} \mathrm{mFNR}(c) &= \frac{E(N_{01})}{E(S)} = \frac{\sum_{i\in S} P(T_i \geqslant c, \Theta_i = 1)}{\sum_{i\in S} P(T_i \geqslant c)} \\ &= \frac{\sum_{i\in S} \pi_{i1}[1 - G_i^1(c)]}{\sum_{i\in S}(\pi_{i0}[1 - G_i^0(c)] + \pi_{i1}[1 - G_i^1(c)])} \\ &= \frac{[1 - G^1(c)]\sum_{i\in S} \pi_{i1}}{[1 - G^0(c)]\sum_{i\in S} \pi_{i0} + [1 - G^1(c)]\sum_{i\in S} \pi_{i1}}. \end{aligned}$$

Then,

$$
\begin{aligned}
&\frac{d(\mathrm{mFDR}(c))}{dc} \\
&= \left(g^0(c) \sum_{i \in S} \pi_{i0} \left[G^0(c) \sum_{i \in S} \pi_{i0} + G^1(c) \sum_{i \in S} \pi_{i1} \right] \right. \\
&\quad \left. - G^0(c) \sum_{i \in S} \pi_{i0} \left[g^0(c) \sum_{i \in S} \pi_{i0} + g^1(c) \sum_{i \in S} \pi_{i1} \right] \right) \\
&\quad \Bigg/ \left[G^0(c) \sum_{i \in S} \pi_{i0} + G^1(c) \sum_{i \in S} \pi_{i1} \right]^2 \\
&= \frac{[g^0(c)G^1(c) - G^0(c)g^1(c)](\sum_{i \in S} \pi_{i0})(\sum_{i \in S} \pi_{i1})}{[G^0(c) \sum_{i \in S} \pi_{i0} + G^1(c) \sum_{i \in S} \pi_{i1}]^2} \\
&> 0
\end{aligned}
$$

following from (1), and

$$
\begin{aligned}
&\frac{d(\mathrm{mFNR}(c))}{dc} \\
&= \left\{ - g^1(c) \sum_{i \in S} \pi_{i1} \left([1 - G^0(c)] \sum_{i \in S} \pi_{i0} + [1 - G^1(c)] \sum_{i \in S} \pi_{i1} \right) \right. \\
&\quad \left. - \left([1 - G^1(c)] \sum_{i \in S} \pi_{i1} \right) \left(-g^0(c) \sum_{i \in S} \pi_{i0} - g^1(c) \sum_{i \in S} \pi_{i1} \right) \right\} \\
&\quad \Bigg/ \left([1 - G^0(c)] \sum_{i \in S} \pi_{i0} + [1 - G^1(c)] \sum_{i \in S} \pi_{i1} \right)^2 \\
&= \frac{(g^0(c)[1 - G^1(c)] - g^1(c)[1 - G^0(c)])(\sum_{i \in S} \pi_{i0})(\sum_{i \in S} \pi_{i1})}{([1 - G^0(c)] \sum_{i \in S} \pi_{i0} + [1 - G^1(c)] \sum_{i \in S} \pi_{i1})^2} \\
&< 0
\end{aligned}
$$

following from (2). Hence we obtain part (a) and (b) of the theorem.

For part (c), the classification risk with the loss function

$$
L_\lambda(\boldsymbol{\Theta}, \boldsymbol{\delta}) = \frac{1}{N} \sum_{i \in S} \{ \lambda(1 - \Theta_i)\delta_i + \Theta_i(1 - \delta_i) \}
$$

is

$$\begin{aligned} E[L_\lambda(\boldsymbol{\Theta}, \boldsymbol{\delta})] &= \frac{1}{N} \sum_{i \in S} \{\lambda P(\Theta_i = 0, T_i < c) + P(\Theta_i = 1, T_i \geqslant c)\} \\ &= \frac{1}{N} \sum_{i \in S} \{\lambda \pi_{i0} G_i^0(c) + \pi_{i1}[1 - G_i^1(c)]\} \\ &= \frac{1}{N} \left\{ \lambda G^0(c) \sum_{i \in S} \pi_{i0} + [1 - G^1(c)] \sum_{i \in S} \pi_{i1} \right\}. \end{aligned}$$

The optimal cutoff c^* that minimizes this risk satisfies

$$\lambda = \frac{g^1(c^*) \sum_{i \in S} \pi_{i1}}{g^0(c^*) \sum_{i \in S} \pi_{i0}}.$$

Since $\boldsymbol{T} \in \mathcal{T}$, we have $g^1(c^*)/g^0(c^*)$ is monotonically decreasing in c^*. Thus, $\lambda(c^*)$ is monotonically decreasing in c^*.

2.2 *Theorem 2 in Sun and Cai (2009) for HMRF*

Proof. Suppose there are v_L hypotheses from the null and k_L hypotheses from the nonnull among the r rejected hypotheses when the decision rule $\boldsymbol{\delta}(\boldsymbol{L}, c_L)$ is applied with test statistic $\boldsymbol{L}$ and cutoff c_L. We have $v_L = \sum_{i \in S} P(\Theta_i = 0, L_i < c_L)$ and $k_L = \sum_{i \in S} P(\Theta_i = 1, L_i < c_L)$, and the classification risk

$$\begin{aligned} R_{\lambda(\alpha)} &= E[L_{\lambda(\alpha)}(\boldsymbol{\Theta}, \boldsymbol{\delta}(\boldsymbol{L}, c_L))] \\ &= \frac{1}{N} \sum_{i \in S} \{\lambda(\alpha) P(\Theta_i = 0, L_i < c_L) + P(\Theta_i = 1, L_i \geqslant c_L)\} \\ &= \frac{1}{N} \left\{ \sum_{i \in S} \pi_{i1} + \lambda(\alpha) v_L - k_L \right\}. \end{aligned} \tag{C.4}$$

Then following the proof of Theorem 1 in Sun and Cai (2007) using the expression (C.4) for the classification risk $R_{\lambda(\alpha)}$, we complete the proof.

2.3 *Theorems 3 and 4 in Sun and Cai (2009) for HMRF*

Proof. The proofs are the same as those of Theorems 3 and 4 in Sun and Cai (2009), thus omitted.

2.4 *Corollary 1 in Sun and Cai (2009) for HMRF*

Proof. Following the proof of Corollary 1 in Sun and Cai (2009) with the expression of the risk R replaced by

$$\begin{aligned} R &= \frac{1}{N}\sum_{i\in S}\left\{\frac{1}{t}\pi_{i0}G_i^0(t^*)+\pi_{i1}[1-G_i^1(t^*)]\right\} \\ &= \frac{1}{N}\left\{\frac{1}{t}G^0(t^*)\sum_{i\in S}\pi_{i0}+[1-G^1(t^*)]\sum_{i\in S}\pi_{i1}\right\} \end{aligned}$$

and their equation $g^1(t^*)/g^0(t^*) = (1/t)\pi_0/\pi_1$ substituted by the new equation $g^1(t^*)/g^0(t^*) = (1/t)\sum_{i\in S}\pi_{i0}/\sum_{i\in S}\pi_{i1}$, we complete the proof.

2.5 *Theorems 1 and 2 in Wei et al. (2009) for HMRF*

Proof. For Theorem 1 and the validity of oracle PLIS procedure in Theorem 2, the proofs are the same as those in Wei et al. (2009). For the optimality of oracle PLIS procedure in Theorem 2, the proof is the same as the proof of the optimality of oracle LIS procedure given above.

3. Web Appendix C: Unbounded likelihood of HMRF

For any voxel $t \in S$, define a specific configuration of $\boldsymbol{\Theta}$ by $\boldsymbol{\theta}_{\{t\}} = (\theta_s)_{s\in S}$ with $\theta_t = 1$ and $\theta_s = 0$ if $s \neq t$. Then the observed likelihood function of HMRF

$$\begin{aligned} L(\boldsymbol{\Phi}|\boldsymbol{x}) = P_{\boldsymbol{\Phi}}(\boldsymbol{x}) &= \sum_{\boldsymbol{\Theta}} P_{\boldsymbol{\phi}}(\boldsymbol{x}|\boldsymbol{\Theta})P_{\boldsymbol{\varphi}}(\boldsymbol{\Theta}) \\ &\geqslant P_{\boldsymbol{\phi}}(\boldsymbol{x}|\boldsymbol{\Theta}=\boldsymbol{\theta}_{\{t\}})P_{\boldsymbol{\varphi}}(\boldsymbol{\Theta}=\boldsymbol{\theta}_{\{t\}}) \\ &= P_{\boldsymbol{\phi}}(x_t|\Theta_t=1)\prod_{s\in S\setminus\{t\}}P_{\boldsymbol{\phi}}(x_s|\Theta_s=0)P_{\boldsymbol{\varphi}}(\boldsymbol{\Theta}_{S\setminus\{t\}}=\mathbf{0},\Theta_t=1) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left\{-\frac{(x_t-\mu_1)^2}{2\sigma_1^2}\right\}+\sum_{l=2}^{L}N(x_t;\mu_l,\sigma_l^2)\right) \\ &\qquad \times(2\pi)^{-\frac{N-1}{2}}\exp\left\{-\frac{1}{2}\sum_{s\in S\setminus\{t\}}x_s^2\right\}\frac{e^h}{Z(\beta,h)} \\ &\to \infty \end{aligned}$$

if $\mu_1 = x_t$ and $\sigma_1^2 \to 0$ with other parameters fixed. Thus the observed likelihood function is unbounded. The similar unbounded-likelihood phenomenon for Gaussian hidden Markov chain model has been shown in Ridolfi (1997) and Chen, Huang, and Wang (2014).

4. Web Appendix D: Gibbs Sampler Approximations

This section presents the approximations of quantities of interest in GEM. Let Ω be the set of all possible configurations of $\boldsymbol{\Theta}$: $\Omega = \{\boldsymbol{\theta} = (\theta_s)_{s\in S} : \theta_s \in \{0,1\}, s \in S\}$. By the ergodic theorem of the Gibbs sampler (See Lemma 1 and Theorem 1 in Roberts and Smith (1994)), for any Gibbs distribution (See definition (4.3) in Geman and Geman (1984)) $\pi(\boldsymbol{\theta})$ and any real-valued function $f(\boldsymbol{\theta})$ on Ω, with probability one,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{\theta}^{(i)}) = \int_{\Omega} f(\boldsymbol{\theta}) d\pi(\boldsymbol{\theta}) = E[f(\boldsymbol{\Theta})],$$

where $\boldsymbol{\theta}^{(i)}, i = 1, ..., n$ are samples successively generated using the Gibbs sampler by $\pi(\boldsymbol{\theta})$. For our HMRF, it is easy to see that both the Ising model probability distribution $P_{\boldsymbol{\varphi}}(\boldsymbol{\theta})$ and the conditional probability distribution $P_{\boldsymbol{\Phi}^{(t)}}(\boldsymbol{\theta}|\boldsymbol{x})$ are Gibbs distributions. Thus by the ergodic theorem, the following quantities can be approximated using Monte Carlo averages via Gibbs sampler:

$$\begin{aligned}
\boldsymbol{U}^{(t+1)}(\boldsymbol{\varphi}) &= E_{\boldsymbol{\Phi}^{(t)}}[\boldsymbol{H}(\boldsymbol{\Theta})|\boldsymbol{x}] - E_{\boldsymbol{\varphi}}[\boldsymbol{H}(\boldsymbol{\Theta})] \\
&\approx \frac{1}{n}\sum_{i=1}^{n}\left(\boldsymbol{H}(\boldsymbol{\theta}^{(t,i,\boldsymbol{x})}) - \boldsymbol{H}(\boldsymbol{\theta}^{(i,\boldsymbol{\varphi})})\right), \\
\boldsymbol{I}(\boldsymbol{\varphi}) &= Var_{\boldsymbol{\varphi}}[\boldsymbol{H}(\boldsymbol{\Theta})] \\
&= E_{\boldsymbol{\varphi}}\left[(\boldsymbol{H}(\boldsymbol{\Theta}) - E_{\boldsymbol{\varphi}}[\boldsymbol{H}(\boldsymbol{\Theta})])^{\otimes 2}\right] \\
&\approx \frac{1}{n-1}\sum_{i=1}^{n}\left(\boldsymbol{H}(\boldsymbol{\theta}^{(i,\boldsymbol{\varphi})}) - \frac{1}{n}\sum_{j=1}^{n}\boldsymbol{H}(\boldsymbol{\theta}^{(j,\boldsymbol{\varphi})})\right)^{\otimes 2}, \\
\gamma_s^{(t)}(i) &= P_{\boldsymbol{\Phi}^{(t)}}(\Theta_s = i|\boldsymbol{x}) = E_{\boldsymbol{\Phi}^{(t)}}[\mathbf{1}(\Theta_s = i)|\boldsymbol{x}] \\
&= E_{\boldsymbol{\Phi}^{(t)}}[\mathbf{1}(\Theta_s = i)\mathbf{1}(\boldsymbol{\Theta} \in \Omega)|\boldsymbol{x}] \\
&\approx \frac{1}{n}\sum_{k=1}^{n}\mathbf{1}(\theta_s^{(t,k,\boldsymbol{x})} = i),
\end{aligned}$$

$$\begin{aligned}\frac{C}{Z(\boldsymbol{\varphi})} &= E_{\boldsymbol{\varphi}}[\exp\{-\boldsymbol{\varphi}^T\boldsymbol{H}(\boldsymbol{\Theta})\}] \\ &\approx \frac{1}{n}\sum_{i=1}^{n}\exp\{-\boldsymbol{\varphi}^T\boldsymbol{H}(\boldsymbol{\theta}^{(i,\boldsymbol{\varphi})})\},\end{aligned}$$

and

$$\begin{aligned}&Q_2(\boldsymbol{\varphi}^{(t+1,m)}|\boldsymbol{\Phi}^{(t)}) - Q_2(\boldsymbol{\varphi}^{(t)}|\boldsymbol{\Phi}^{(t)}) \\ &= E_{\boldsymbol{\Phi}^{(t)}}[\log P_{\varphi^{(t+1,m)}}(\boldsymbol{\Theta}) - \log P_{\varphi^{(t)}}(\boldsymbol{\Theta})|\boldsymbol{x}] \\ &= E_{\boldsymbol{\Phi}^{(t)}}[(\boldsymbol{\varphi}^{(t+1,m)} - \boldsymbol{\varphi}^{(t)})^T\boldsymbol{H}(\boldsymbol{\Theta})|\boldsymbol{x}] + \log\left(\frac{Z(\varphi^{(t)})}{Z(\varphi^{(t+1,m)})}\right) \\ &\approx \frac{1}{n}(\boldsymbol{\varphi}^{(t+1,m)} - \boldsymbol{\varphi}^{(t)})^T\sum_{i=1}^{n}\boldsymbol{H}(\boldsymbol{\theta}^{(t,i,\boldsymbol{x})}) \\ &\quad + \log\left(\frac{\sum_{i=1}^{n}\exp\{-\boldsymbol{\varphi}^{(t+1,m)^T}\boldsymbol{H}(\boldsymbol{\theta}^{(i,\boldsymbol{\varphi}^{(t+1,m)})})\}}{\sum_{i=1}^{n}\exp\{-\boldsymbol{\varphi}^{(t)^T}\boldsymbol{H}(\boldsymbol{\theta}^{(i,\boldsymbol{\varphi}^{(t)})})\}}\right),\end{aligned}$$

where $\{\boldsymbol{\theta}^{(1,\boldsymbol{\varphi})},...,\boldsymbol{\theta}^{(n,\boldsymbol{\varphi})}\}$ and $\{\boldsymbol{\theta}^{(t,1,\boldsymbol{x})},...,\boldsymbol{\theta}^{(t,n,\boldsymbol{x})}\}$ are large n samples successively generated using the Gibbs sampler by $P_{\boldsymbol{\varphi}}(\boldsymbol{\theta})$ and $P_{\boldsymbol{\Phi}^{(t)}}(\boldsymbol{\theta}|\boldsymbol{x})$ respectively, and C is the cardinality of set Ω.

5. Web Appendix E: ADNI FDG-PET Imaging Data Analysis

Alzheimer's disease (AD) is the most common cause of dementia in the elderly population. The worldwide prevalence of Alzheimer's disease was 26.6 million in 2006 and is predicted to be 1 in 85 persons by 2050 (Brookmeyer et al., 2007). Much progress has been made in the diagnosis of AD including clinical assessment and neuroimaging techniques. One such extensively used neuroimaging technique is ^{18}F-Fluorodeoxyglucose positron emission tomography (FDG-PET) imaging, which can be used to evaluate the cerebral metabolic rate of glucose (CMRgl). Numerous FDG-PET studies (Nestor et al., 2003; Mosconi et al., 2005; Langbaum et al., 2009) have demonstrated significant reductions of CMRgl in brain regions in patients with AD and its prodromal stage mild cognitive impairment (MCI), compared with normal control (NC) subjects. These reduction can be used for the early

detection of AD. Voxel-level multiple testing methods are common approaches to identify voxels with significant group differences in CMRgl (Alexander et al., 2002; Mosconi et al., 2005; Langbaum et al., 2009). We focus on the comparison between MCI and NC for such a purpose.

The motivating FDG-PET imaging data are obtained from Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`adni.loni.usc.edu`). These are the baseline FDG-PET images of 102 NC subjects and 206 MCI patients. Each subjects baseline FDG-PET image has been reoriented into a standard $160 \times 160 \times 96$ voxel image grid with 1.5 mm cubic voxels and the anterior-posterior axis of the subject is parallel to the line connecting the anterior and posterior commissures, so-called AC-PC line. Each image is normalized by the average of voxel values in pons and cerebellar vermis, which are well preserved regions in Alzheimers patients. In human brain, the cerebral cortex is segregated into 43 Brodmann areas (BAs) based on the cytoarchitectural organization of neurons (Garey, 2006). We consider 30 of them after removing the BAs that are either too small or not always reliably registered. We also investigate 9 subcortical regions, including hippocampus, which are commonly considered in AD studies. A region is further divided into two if its bilateral parts in the left and right hemispheres are separated completely without a shared border in the middle of the brain. We have considered combining neighboring regions to potentially increase accuracy, but failed to find any pair with similar estimated HMRF model parameters. Finally, 61 regions of interest (ROIs) are included in the analysis, where the number of voxels in each region ranges from 149 to 20,680 with a median of 2,517. The total number of voxels of these 61 ROIs is $N = 251,500$.

We apply the PLIS procedure with HMRFs to the analysis of ADNIs FDG-PET imaging data, which is compared with BH, q-value and CLfdr procedures. Since the FDG-PET scans were normalized to the average of pons and cerebellar vermis, areas of the brain known to

be least affected in AD, it was not surprising that almost all the signal voxels are found with decreased CMRgl. Both PLIS and CLfdr procedures discovered significant metabolic reduction, with a regional proportion of signals $> 50\%$, in brain regions preferentially affected by AD, including the posterior cingulate (BAs 23, 31; Mosconi et al., 2008; Langbaum et al., 2009), parietal cortex (BAs 7, 37, 39, 40; Minoshima et al., 1995; Matsuda, 2001), temporal cortex (BAs 20 to 22; Alexander et al., 2002; Landau et al., 2011), medial temporal cortex (BAs 28, 34; Karow et al., 2010), frontal cortex (BAs 8 to 11, and 44 to 47; Mosconi, 2005), insular cortex (Perneczky et al., 2007), amygdala (Nestor et al., 2003) and hippocampus (Mosconi et al., 2005). In regions also typically affected in AD, such as anterior cingulate (BAs 24, 32; Fouquet et al., 2009) and occipital cortex (BAs 17 to 19; Langbaum et al., 2009), the proportions of signals found by PLIS are 49.6% and 39.0%, respectively, compared with 35.4% and 11.6% found by CLfdr, 12.2% and 0.94% by q-value, as well as only 1.24% and 0.87% by BH.

With respect to the regions that are relatively spared from AD (Benson et al., 1983; Matsuda, 2001; Ishii, 2002) or rarely reported in the literature of the disease, caudate, thalamus and putamen are found with high proportions of signals by PLIS ($> 45\%$) and CLfdr ($> 25\%$) in each of these regions; signals in medulla, midbrain, cerebellar hemispheres, pre-motor cortex (BA 6) and primary somatosensory cortex (BAs 1, 2, 3, 5) are each claimed with a proportion greater than 20% by PLIS, but very sparse found by the other three procedures. Since MCI as a group consists of a mix of patients, many of them will progress to AD but some will not which may include subjects with corticobasal degeneration (Ishii, 2002), frontotemporal dementia (Jeong et al., 2005), or Parkinsonism (Huang et al., 2007; Zeman, Carpenter, and Scott, 2011; Ishii, 2013), it is not surprising that some areas not typical of AD patients were found to be abnormal in the MCI group.

References

Alexander, G. E., Chen, K., Pietrini, P., Rapoport, S. I., and Reiman, E. M. (2002). Longitudinal PET evaluation of cerebral metabolic decline in dementia: A potential outcome measure in Alzheimer's disease treatment studies. *American Journal of Psychiatry* **196,** 738-745.

Benson, D. F., Kuhl, D. E., Hawkins, R. A., Phelps, M. E., Cummings, J. L., and Tsai, S. Y. (1983). The fluorodeoxyglucose 18F scan in Alzheimer's disease and multi-infarct dementia. *Archives of Neurology* **40,** 711-714.

Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). Forecasting the global burden of Alzheimers disease. *Alzheimer's & Dementia* **3,** 186-191.

Chen, J., Huang, Y., and Wang, P. (2014). Composite likelihood under hidden Markov model. *Statistica Sinica* [Preprint], doi:10.5705/ss.2013.084t.

Fouquet, M., Desgranges, B., Landeau, B., Duchesnay, E., Mezenge, F., De La Sayette, V., et al. (2009). Longitudinal brain metabolic changes from amnestic mild cognitive impairment to Alzheimer's disease. *Brain* **132,** 2058-2067.

Garey, L. J. (2006). *Brodmann's Localisation in the Cerebral Cortex*. New York: Springer.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6,** 721-741.

Huang, C., Tang, C., Feigin, A., Lesser, M., Ma, Y., Pourfar, M., et al. (2007). Changes in network activity with the progression of Parkinsons disease. *Brain* **130,** 1834-1846.

Ishii, K. (2002). Clinical application of positron emission tomography for diagnosis of dementia. *Annals of Nuclear Medicine* **16,** 515-525.

Ishii, K. (2013). PET approaches for diagnosis of dementia. *American Journal of Neuroradiology* [online], DOI: 10.3174/ajnr.A3695.

Jeong, Y., Cho, S. S., Park, J. M., Kang, S. J., Lee, J. S., Kang, E., et al. (2005). 18F-FDG PET findings in frontotemporal dementia: An SPM analysis of 29 patients. *Journal of Nuclear Medicine* **46,** 233-239.

Karow, D. S., McEvoy, L. K., Fennema-Notestine, C., Hagler, D. J., Jennings, R. G., Brewer, J. B., et al. (2010). Relative capability of MR imaging and FDG PET to depict changes associated with prodromal and early Alzheimer disease. *Radiology* **256,** 932-942.

Landau, S. M., Harvey, D., Madison, C. M., Koeppe, R. A., Reiman, E. M., Foster, N. L., et al. (2011). Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. *Neurobiology of Aging* **32,** 1207-1218.

Langbaum, J. B. S., Chen, K., Lee, W., Reschke, C., Bandy, D., Fleisher, A. S., et al. (2009). Categorical and correlational analyses of baseline fluorodeoxyglucose positron emission tomography images from the Alzheimer's Disease Neuroimaging Initiative (ADNI). *NeuroImage* **45,** 1107-1116.

Matsuda, H. (2001). Cerebral blood flow and metabolic abnormalities in Alzheimers disease. *Annals of Nuclear Medicine* **15,** 85-92.

Minoshima, S., Frey, K. A., Koeppe, R. A., Foster, N. L., and Kuhl, D. E. (1995). A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of Fluorine-18-FDG PET. *Journal of Nuclear Medicine* **36,** 1238-1248.

Mosconi, L. (2005). Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease. *European Journal of Nuclear Medicine and Molecular Imaging* **32,** 486-510.

Mosconi, L., Tsui, W. H., De Santi, S., Li, J., Rusinek, H., Convit, A., et al. (2005). Reduced hippocampal metabolism in MCI and AD: Automated FDG-PET image analysis. *Neurology* **64,** 1860-1867.

Mosconi, L., Tsui, W. H., Herholz, K., Pupi, A., Drzezga, A., Lucignani, G., et al.

(2008). Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias. *Journal of Nuclear Medicine* **49,** 390-398.

Nestor, P. J., Fryer, T. D., Smielewski, P., and Hodges, J. R. (2003). Limbic hypometabolism in Alzheimer's disease and mild cognitive impairment. *Annals of Neurology* **54,** 343-351.

Perneczky, R., Drzezga, A., Diehl-Schmid, J., Li, Y., and Kurz, A. (2007). Gender differences in brain reserve: An (18)F-FDG PET study in Alzheimers disease. *Journal of Neurology* **254,** 1395-1400.

Ridolfi, A. (1997). Maximum likelihood estimation of hidden Markov model parameters, with application to medical image segmentation. Politecnico di Milano, Milan, Italy.

Roberts, G. O. and Smith A. F. M. (1994). Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic Processes and their Applications* **49,** 207-216.

Sun, W. and Cai, T. T. (2007). Oracle and adaptive compound decision rules for false discovery rate control. *Journal of the American Statistical Association* **102,** 901-912.

Sun, W. and Cai, T. T. (2009). Large-scale multiple testing under dependence. *Journal of the Royal Statistical Society, Series B* **71,** 393-424.

Wei, Z., Sun, W., Wang, K., and Hakonarson, H. (2009). Multiple testing in genome-wide association studies via hidden Markov models. *Bioinformatics* **25,** 2802-2808.

Zeman, M. N., Carpenter, G. M., and Scott, P. J. (2011). Diagnosis of dementia using nuclear medicine imaging modalities. In *12 Chapters on Nuclear Medicine*. Croatia: InTech, pp. 199-230.