Supplementary File: Multi-object Data Integration in the Study of Primary Progressive Aphasia

February 11, 2023

Abstract

This file has two sections. Section 1 details out posterior sampling for BMRR with full conditional posterior distributions. Section 2.1 provides additional figures and results corresponding to the simulation study in Section 5 of the main article. Section 2.2 offers additional results for data analysis which could not be included in Section 6 of the main article due to space constraint.

1 Posterior Sampling

In this section we elaborate the posterior sampling procedure of BMMR. As preeliminary notation, for any matrix $\boldsymbol{M}$ let $\boldsymbol{M}_{i,\cdot}$ be the i-th row of $\boldsymbol{M}$, $\boldsymbol{M}_{\cdot,j}$ be the j-th column of $\boldsymbol{M}$, $\boldsymbol{M}_{-i,\cdot}$ be the $\boldsymbol{M}$ matrix without the i-th column, $\boldsymbol{M}_{\cdot,-j}$ be the $\boldsymbol{M}$ matrix without the jth column, $\boldsymbol{M}_{i,-j}$ the ith column of $\boldsymbol{M}$ without the jth entry, $\boldsymbol{M}_{-i,j}$ be the jth column of $\boldsymbol{M}$ without the ith entry. Similarly for a vector $\boldsymbol{v}$ let $\boldsymbol{v}_{-i}$ be vector $\boldsymbol{v}$ without the ith entry. Also, let $\otimes$ denote the Kronecker product operator and "$\cdot$" the entry-wise product operator for matrices.

Bayesian estimation of the model is performed through Gibbs sampling which cycles through the following steps.

1. For each $p \in \{1, ..., P\}$ sample from $\boldsymbol{\psi}^a_{p,\cdot}$, from the full conditional $p(\boldsymbol{\psi}^a_{p,\cdot}|\boldsymbol{\psi}^a_{-p,\cdot}, \boldsymbol{\Theta}, \sigma^2, \boldsymbol{A})$.

2. For each $p \in \{1, ..., P\}$ sample from $\boldsymbol{\psi}_{p,.}^g$, from the full conditional $p(\boldsymbol{\psi}_{p,.}^g | \boldsymbol{\beta}_p, \sigma^2, \boldsymbol{g}_p)$.

3. For each $p \in \{1, ..., P\}$ sample jointly ξ_p, $\boldsymbol{\Theta}_{-p,p}$ and $\boldsymbol{\beta}_p$ from

$$p(\xi_p, \boldsymbol{\Theta}_{-p,p}, \boldsymbol{\beta}_p | \nu, \tau^2, \sigma_\theta^2, \eta_p^2, \boldsymbol{\Lambda}_{-p,p}, \phi_p^2, \boldsymbol{A}, \boldsymbol{G})$$
$$= p(\boldsymbol{\Theta}_{-p,p}, \boldsymbol{\beta}_p | \boldsymbol{\xi}, \tau^2, \sigma_\theta^2, \eta_p^2, \boldsymbol{\Lambda}_{-p,p}, \phi_p^2, \boldsymbol{A}, \boldsymbol{G})$$
$$\times p(\xi_p | \nu, \boldsymbol{\xi}_{-p}, \tau^2, \sigma_\theta^2, \eta_p^2, \boldsymbol{\Lambda}_{-p,p}, \phi_p^2, \boldsymbol{A}, \boldsymbol{G})$$

4. Sample τ^2 from the full conditional

$$p(\tau^2 | \boldsymbol{\xi}, \boldsymbol{\Theta}, \boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_P, \sigma_\theta^2, \eta_1^2, ..., \eta_P^2, \boldsymbol{\Lambda}, \phi_1^2, ..., \phi_P^2, \boldsymbol{A}, \boldsymbol{G})$$

5. Sample ν from the full conditional $p(\nu | \boldsymbol{\xi})$.

6. Sample the horseshoe parameters using the latent variable approach as in Makalic and Schmidt (2016).

Next, we describe all the relevant conditional distributions in steps 1-6. Note that,

$$\boldsymbol{\psi}_{p,.}^a | \boldsymbol{\psi}_{-p,.}^a, \boldsymbol{\Theta}, \sigma^2, \boldsymbol{A} \sim N\left(\hat{\boldsymbol{\psi}}_{p,.}^a, \tau^2(\boldsymbol{W}'\boldsymbol{W})^{-1}\right)$$

where

$$\hat{\boldsymbol{\psi}}_{p,.}^a = (\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{R}$$
$$\boldsymbol{R} = [(\boldsymbol{R}_1)'_{-p,p}, ..., (\boldsymbol{R}_n)'_{-p,p}]'$$
$$\boldsymbol{W} = [\boldsymbol{W}'_1, ..., \boldsymbol{W}'_n]'$$
$$(\boldsymbol{R}_i)_{-p,p} = (\boldsymbol{A}_i)_{-p,p} - \boldsymbol{\Theta}_{-p,p}y_i \quad \forall i \in \{1, ..., n\}$$
$$\boldsymbol{W}_i = \boldsymbol{\psi}_{-p,.}^a \cdot (\boldsymbol{1}_P \otimes \boldsymbol{x}'_i) \quad \forall i \in \{1, ..., n\}.$$

$$\boldsymbol{\psi}_{p,.}^g | \boldsymbol{\beta}_p, \sigma^2, \boldsymbol{g}_p \sim N\left(\hat{\boldsymbol{\psi}}_{p,.}^g, \tau^2(\boldsymbol{W}'\boldsymbol{W})^{-1}\right)$$

where

$$\hat{\boldsymbol{\psi}}^g_{p,.} = (\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{R}$$

$$\boldsymbol{R} = [(\boldsymbol{R}_1)'_{-p,p}, ..., (\boldsymbol{R}_n)'_{-p,p}]'$$

$$\boldsymbol{W} = [\boldsymbol{W}'_1, ..., \boldsymbol{W}'_n]'$$

$$(\boldsymbol{R}_i)_{-p,p} = \boldsymbol{g}_{i,p} - \boldsymbol{\beta}_p y_i \quad \forall i \in \{1, ..., n\}$$

$$\boldsymbol{W}_i = \boldsymbol{1}_{V_p} \otimes \boldsymbol{x}'_i \quad \forall i \in \{1, ..., n\}.$$

To sample $\boldsymbol{\Theta}_{-p,p}, \boldsymbol{\beta}_p$ we set them to $\boldsymbol{0}$ if $\xi_p = 0$, if $\xi = 1$ we sample from

$$\boldsymbol{\Theta}_{-p,p}, \boldsymbol{\beta}_p | \boldsymbol{\xi}_{-p}, \tau^2, \sigma^2_\theta, \eta^2_p, \boldsymbol{\Lambda}_{-p,p}, \phi^2_p, \boldsymbol{A}, \boldsymbol{G} \sim N(\hat{b}, \tau^2 \text{diag}(1/(S^2_{yy} + 1/\boldsymbol{L})))$$

where

$$\hat{b} = \boldsymbol{S}_{xy}/(S^2_{yy} + 1/\boldsymbol{L})$$

$$\boldsymbol{S}_{xy} = [(\boldsymbol{S}_{ay})_{p,-p}[\boldsymbol{\xi}_{-p} = 1]', \boldsymbol{S}'_{gy,p}]'$$

$$S^2_{yy} = \boldsymbol{y}'\boldsymbol{y}$$

$$\boldsymbol{L} = [\sigma^2_\theta \boldsymbol{\Lambda}_{p,-p}[\boldsymbol{\xi}_{-p} = 1]', \eta^2_p(\phi^2_p)']'$$

$$(\boldsymbol{S}_{ay})_{p,p'} = \sum_{i=1}^n (R^a_i)_{p,p'} y_i$$

$$(\boldsymbol{S}_{gy,p})_v = \sum_{i=1}^n R^g_{i,v,p} y_i$$

$$(R^a_i)_{p,p'} = (\boldsymbol{A}_i)_{p,p'} - \sum_{h=1}^H (\psi^a)_{p,h}(\psi^a)_{p',h} x_{i,h}$$

$$R^g_{i,v,p} = \boldsymbol{g}_{i,(v,p)} - \sum_{h=1}^H \psi^g_{p,h} x_{i,h}$$

3

We sample ξ_p from

$$\xi_p|\nu, \boldsymbol{\xi}_{-p}, \tau^2, \sigma_\theta^2, \eta_p^2, \boldsymbol{\Lambda}_{-p,p}, \phi_p^2, \boldsymbol{A}, \boldsymbol{G} \sim Bernoulli(\hat{v}_p)$$

where

$$\hat{v}_p = o_p/(1 + o_p)$$

$$o_p = \exp(\sum_{i=1} c_i)$$

$$c_i = -\frac{1}{2}\left(\log(\boldsymbol{L}_i) + \log\left(S_{yy}^2 + \frac{1}{\boldsymbol{L}_i}\right)\right) + \hat{b}_i \frac{S_{yy}^2 + \frac{1}{\boldsymbol{L}_i}}{2\tau^2}$$

where $\boldsymbol{L}_i$ and $\hat{b}$ are as in the sampling of $\boldsymbol{\Theta}_{-p,p}, \boldsymbol{\beta}_p$.

Finally, we sample τ^2 from

$$\tau^2|\boldsymbol{\xi}, \boldsymbol{\Theta}, \boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_P, \sigma_\theta^2, \eta_1^2, ..., \eta_P^2, \boldsymbol{\Lambda}, \phi_1^2, ..., \phi_P^2, \boldsymbol{A}, \boldsymbol{G} \sim IG(\hat{a}_\tau, \hat{b}_\tau)$$

where

$$\hat{b}_\tau = b_\tau + \frac{\sum_{i=1}^n \sum_{p<p'} R_i^{(}a_{p,p'}) + \sum_{i=1}^n \sum_{p=1}^P R_{i,p}^g}{2}$$
$$+ \frac{\sum_{p<p'} \boldsymbol{\Theta}_{\boldsymbol{\xi}=1,\boldsymbol{\xi}=1}^2/\boldsymbol{\Lambda}_{\boldsymbol{\xi}=1,\boldsymbol{\xi}=1}/\sigma_\theta^2}{2} + \frac{\sum_{p=1}^P \xi_p \boldsymbol{\beta}_p^2/(\eta_p \phi_p^2)}{2}$$
$$\hat{a}_\tau = a_\tau + \frac{\frac{np(p-1)}{2} + n\sum_{p=1}^P V_p + \frac{nq(q-1)}{2} + n\sum_{p=1}^P V_p \xi_p}{2}$$
$$R_i^a = \left(\boldsymbol{A}_i - \sum_{h=1}^H \boldsymbol{\psi}_h^a(\boldsymbol{\psi}_h^a)' x_{i,h} - \boldsymbol{\Theta}\, y_i\right)^2$$
$$R_{i,p}^g = \left(\boldsymbol{g}_{i,p} - \sum_{h=1}^H \boldsymbol{1}_{V_p} \psi_{p,h}^g x_{i,h} - \boldsymbol{\beta}_p\, y_i\right)^2$$

2   Additional Empirical Results

2.1   Additional Simulation Results

The simulation results in Section 5.3 in the main article presents TPR and TNR for all the competitors in terms of identifying influential regions. Since there are trade-offs between the TPR and the TNR performances, Figure 1 in this document presents a single performance measure, F1-score for influential region identification. Given that the two most competing methods in terms of the F1-score are Horseshoe and Regional MUA, we show the plots for F1-score corresponding to BMRR, Horseshoe and Regional MUA. All other competitors show worse performance than these three. As expected, the performance of BMRR along with all other competitors deteriorate for higher dimensional simulation cases. We also observe notable improvement in the performance of Horseshoe, especially in smaller dimensional cases, when it is used for joint model fitting as opposed to using Horseshoe for fitting structural and network object models separately. In contrast, Regional MUA improves only moderately when both set of objects are considered for multiplicity correction, as opposed to employing them separately. BMRR outperforms both its competitors in all simulation cases. The performance gap between BMRR and Horseshoe is more in low sparsity cases than in high sparsity cases.

Figure 2 presents the identification of influential regions in the large dimensional simulation example. Here black spots correspond to the regions which are truly influential and color in a cell is more dark if the posterior probability for the corresponding region being influential is higher. The results show excellent identification of influential regions. We did not offer the posterior probabilities of being influential as the plot looks too clumsy to be understood in that case.

2.2   Additional Data Analysis Results

Figures 3-5 present the coefficients for gender, age and cognitive decline corresponding to the GM response across ROIs, respectively. Figures 6-8 present the coefficients for gender, age and cognitive decline corresponding to the network response matrix, respectively. The purpose of adjusting for the gender, age, and cognitive decline covariates is to allow

Figure 1: F1 Score for the Horseshoe, the Regional MUA approach and BMRR for different simulation scenarios. Here "Big" and "Small" refer to high-dimensional and smaller dimensional examples. The top and bottom row correspond to the true sparsity level of 85% and 70%, respectively.
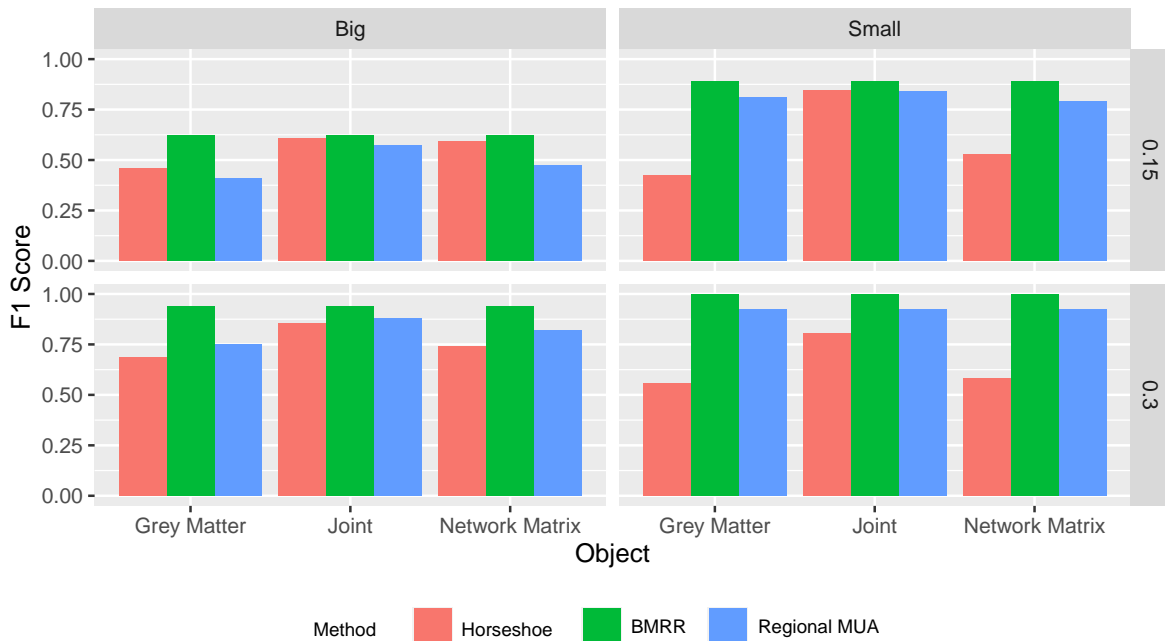
for an adjusted estimate of the association between the speech rate measure and the two image outcomes. The three covariates were selected due to their well known association with both natural and accelerated brain atrophy and network disruption in aging adults and are frequently included as adjustment variables (Mandelli *et al.*, 2018). Thus, they are purposefully included as adjustment variables to reduce potential bias in estimates of our primary predictor of interest, the speech rate measure. Therefore, we discuss the coefficients for gender, age, and cognitive decline to ensure their estimated effects are biologically sensible but note that they are not the primary focus of our inference. These coefficient maps are interpreted in a similar manner to the speech rate coefficient maps in Section 6.2 of the main article, albeit without the hierarchical regularization priors and in terms of ROI-level rather than voxel-level effects due to modeling restraints discussed in Section 3.1 of the main article. In Figure 3, we see that on average males generally have higher grey matter probability throughout the brain as evidenced by the ubiquity of positive associations between gray matter probability and gender status male which is consistent with previous findings

Figure 2: The figure presents performance for influential node identification by BMRR under the two simulation cases in bigger dimensional example. Each row corresponds to a simulation case. The black square in each cell indicate the truly influential regions. A cell is more dark if the posterior probability of being influential for the corresponding region is high.
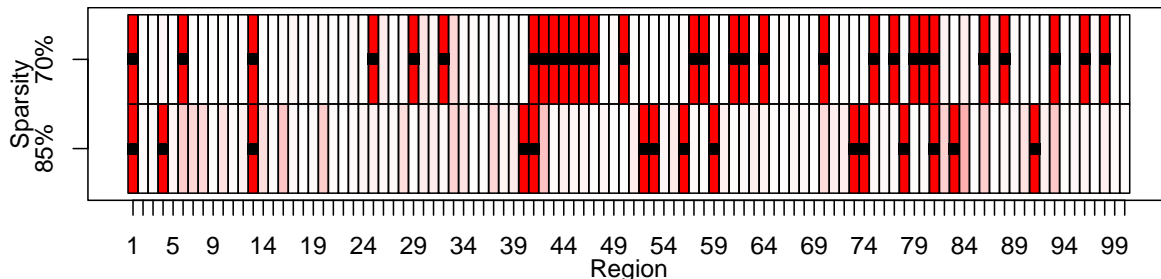
(Good *et al.*, 2001). In Figure 4, we see that there is consistent negative association between age and grey matter content which reflects increasing brain atrophy as humans age (Hafkemeijer *et al.*, 2014). In Figure 5, we see both negative and positive association between cognitive function and grey matter content and note that the effect sizes are small and not clinically meaningful for interpretation. For Figures 6-8, the existing literature on resting state functional connectivity describes a heterogeneous landscape of associations and thus we limit ourselves to characterizing the associations observed but do not consistently try to interpret their direction. In Figure 6, we see a mixture of positive and negative associations between gender and functional connectivity among brain ROIs which is consistent with summaries of the extant literature (Gong *et al.*, 2011). Similarly, Figures 7 and 8 show a mixture of positive and negative association between age and cognitive function with functional connectivity, respectively, which is consistent with observed heterogeneous functional reorganization of brain connectivity that occurs due to aging and atrophy (Zonneveld *et al.*, 2019). In summary, the associations captured in our set of adjustment variables, gender, age, and cognitive function are plausible and their inclusion helps bolster our primary inference related to the speech rate measure.

Finally, Figure 9 and 10 show the point estimates of regression coefficients for GM voxels and for the network matrix, respectively, corresponding to regional MUA. The results show very little signal identified by regional MUA. This is expected since the simulation studies show lower TPR observed in regional MUA compared to BMRR.

7

Figure 3: Point estimates (median) of the regression coefficients for Gender in selected grey matter voxels across horizontal slices using BMRR.
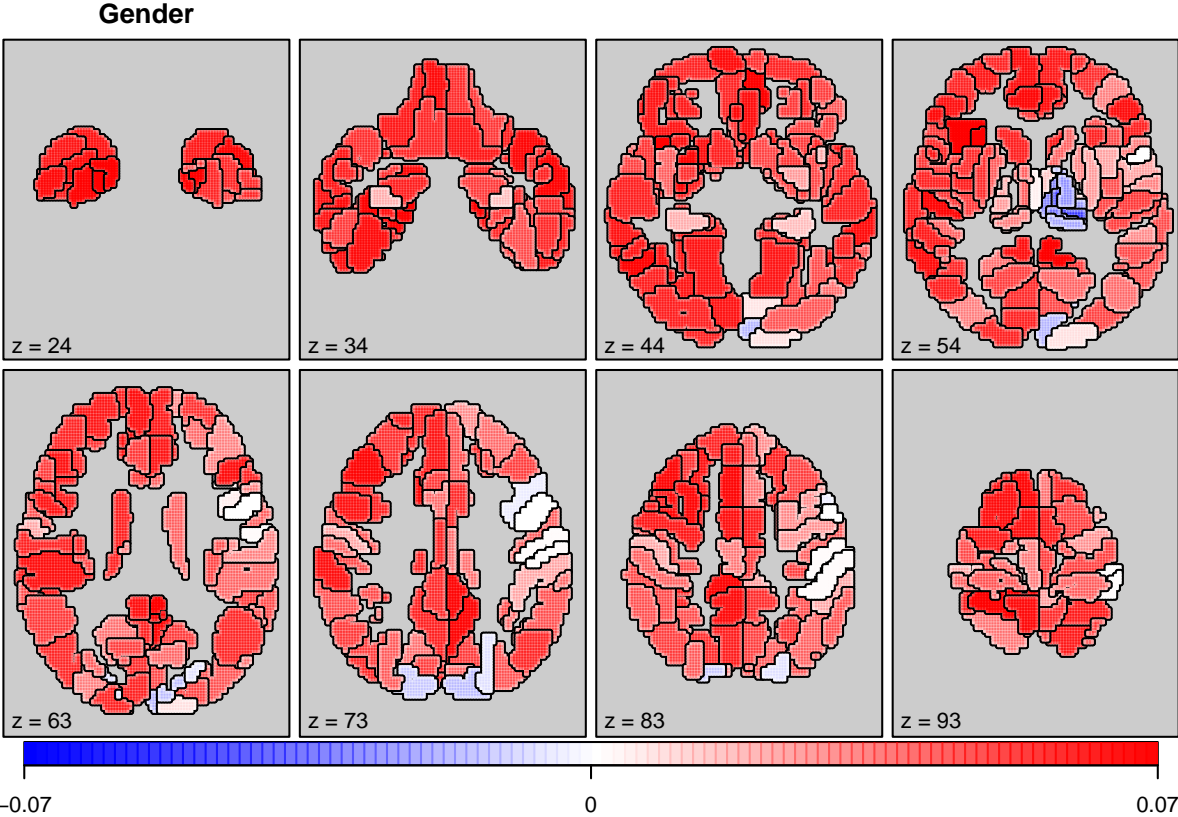
Figure 4: Point estimates (median) of the regression coefficients for Age in selected grey matter voxels across horizontal slices using BMRR.
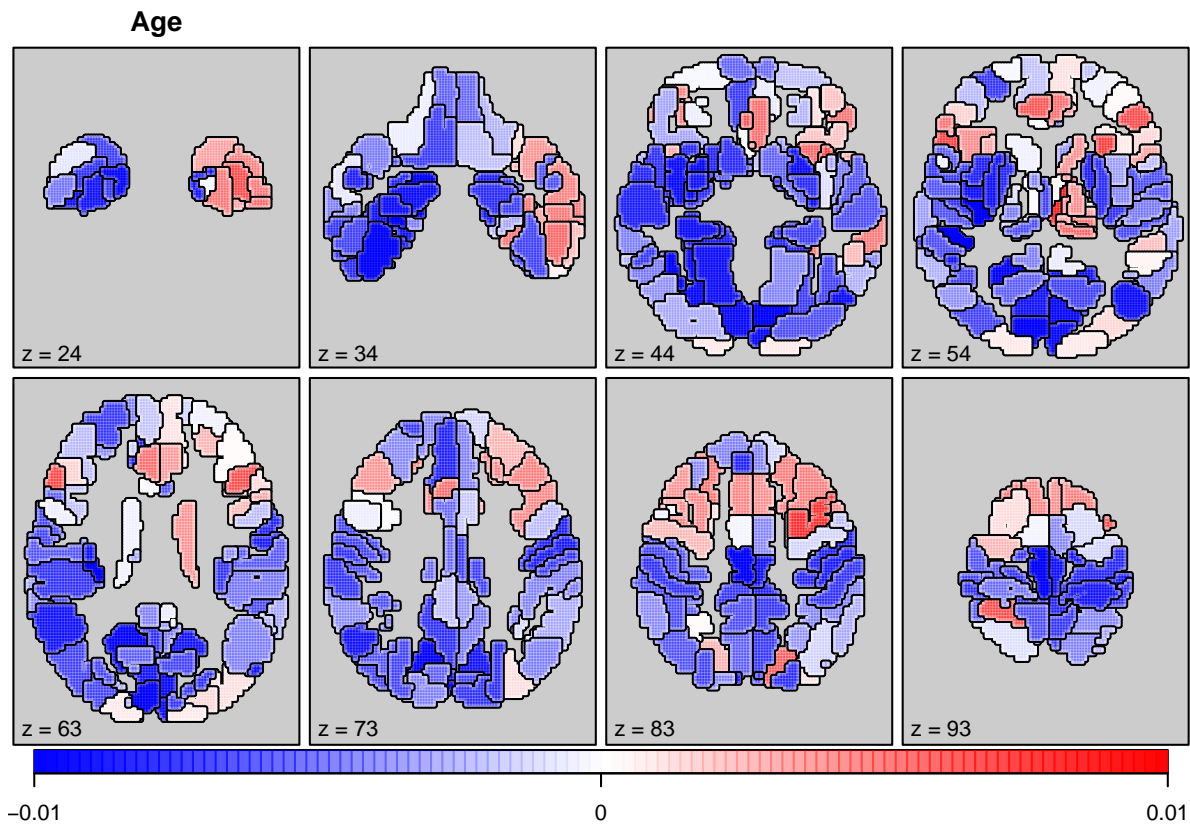
Figure 5: Point estimates (median) of the regression coefficients for Cognitive Decline in selected grey matter voxels across horizontal slices using BMRR.
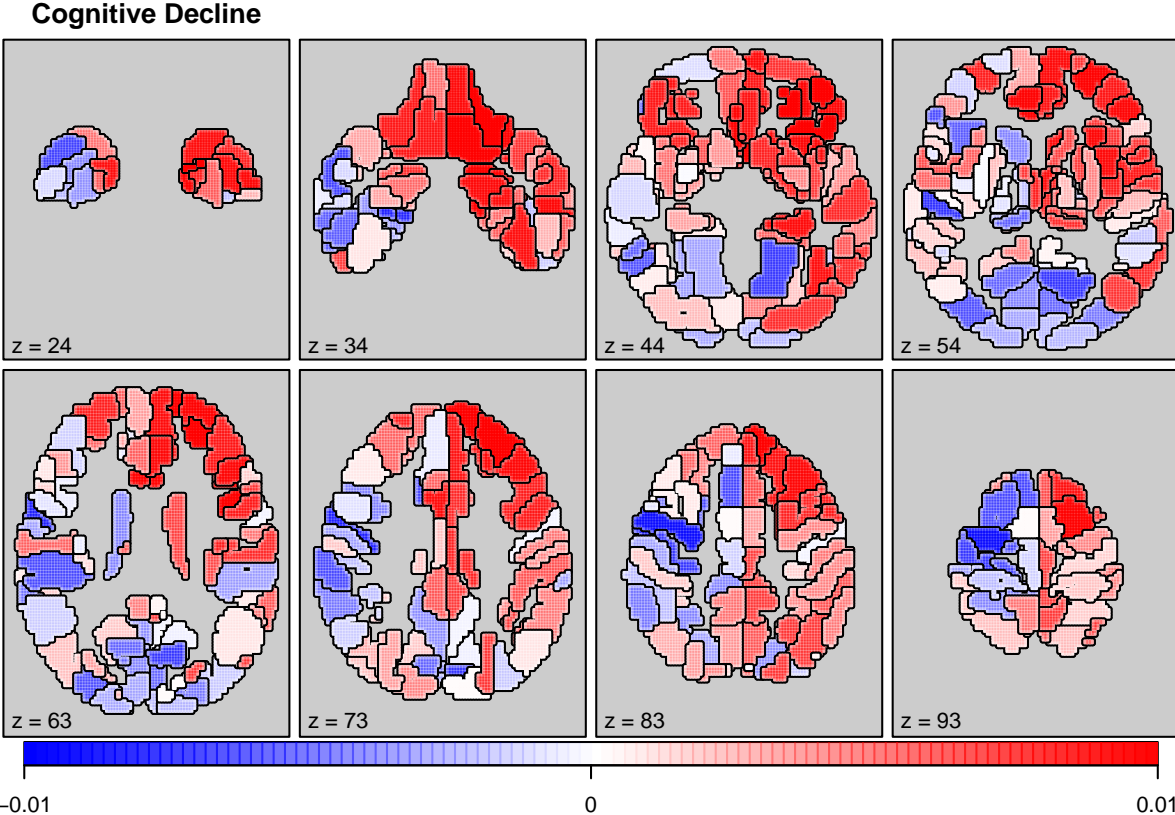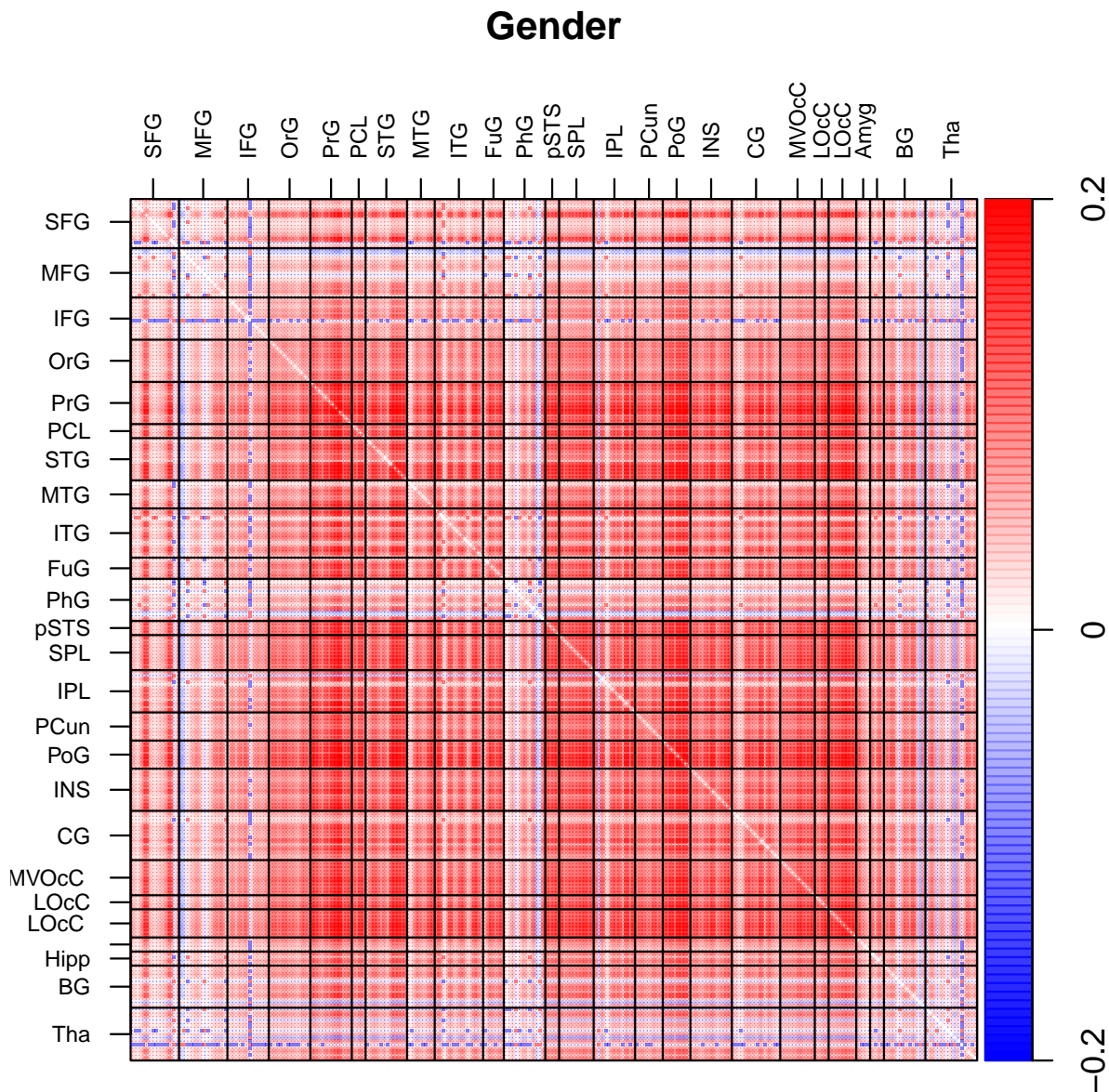
Figure 6: Point estimates (median) of the regression coefficients for Gender for the Network Matrix Edges using BMRR.

References

Gong, G., He, Y., and Evans, A. C. (2011). Brain connectivity: gender makes a difference. *The Neuroscientist*, **17**(5), 575–591.

Good, C. D., Johnsrude, I., Ashburner, J., Henson, R. N., Friston, K. J., and Frackowiak, R. S. (2001). Cerebral asymmetry and the effects of sex and handedness on brain structure:

a voxel-based morphometric analysis of 465 normal adult human brains. *Neuroimage*, **14**(3), 685–700.

Hafkemeijer, A., Altmann-Schneider, I., de Craen, A. J., Slagboom, P. E., van der Grond, J., and Rombouts, S. A. (2014). Associations between age and gray matter volume in anatomical brain networks in middle-aged to older adults. *Aging cell*, **13**(6), 1068–1074.

Makalic, E. and Schmidt, D. F. (2016). A Simple Sampler for the Horseshoe Estimator. *IEEE Signal Processing Letters*, **23**(1), 179–182. Conference Name: IEEE Signal Processing Letters.

Mandelli, M. L., Welch, A. E., Vilaplana, E., Watson, C., Battistella, G., Brown, J. A., Possin, K. L., Hubbard, H. I., Miller, Z. A., Henry, M. L., Marx, G. A., Santos-Santos, M. A., Bajorek, L. P., Fortea, J., Boxer, A., Rabinovici, G., Lee, S., Deleon, J., Rosen, H. J., Miller, B. L., Seeley, W. W., and Gorno-Tempini, M. L. (2018). Altered topology of the functional speech production network in non-fluent/agrammatic variant of PPA. *Cortex*, **108**, 252–264.

Zonneveld, H. I., Pruim, R. H., Bos, D., Vrooman, H. A., Muetzel, R. L., Hofman, A., Rombouts, S. A., van der Lugt, A., Niessen, W. J., Ikram, M. A., *et al.* (2019). Patterns of functional connectivity in an aging population: The rotterdam study. *Neuroimage*, **189**, 432–444.

Figure 7: Point estimates (median) of the regression coefficients for Age for the Network Matrix Edges using BMRR.
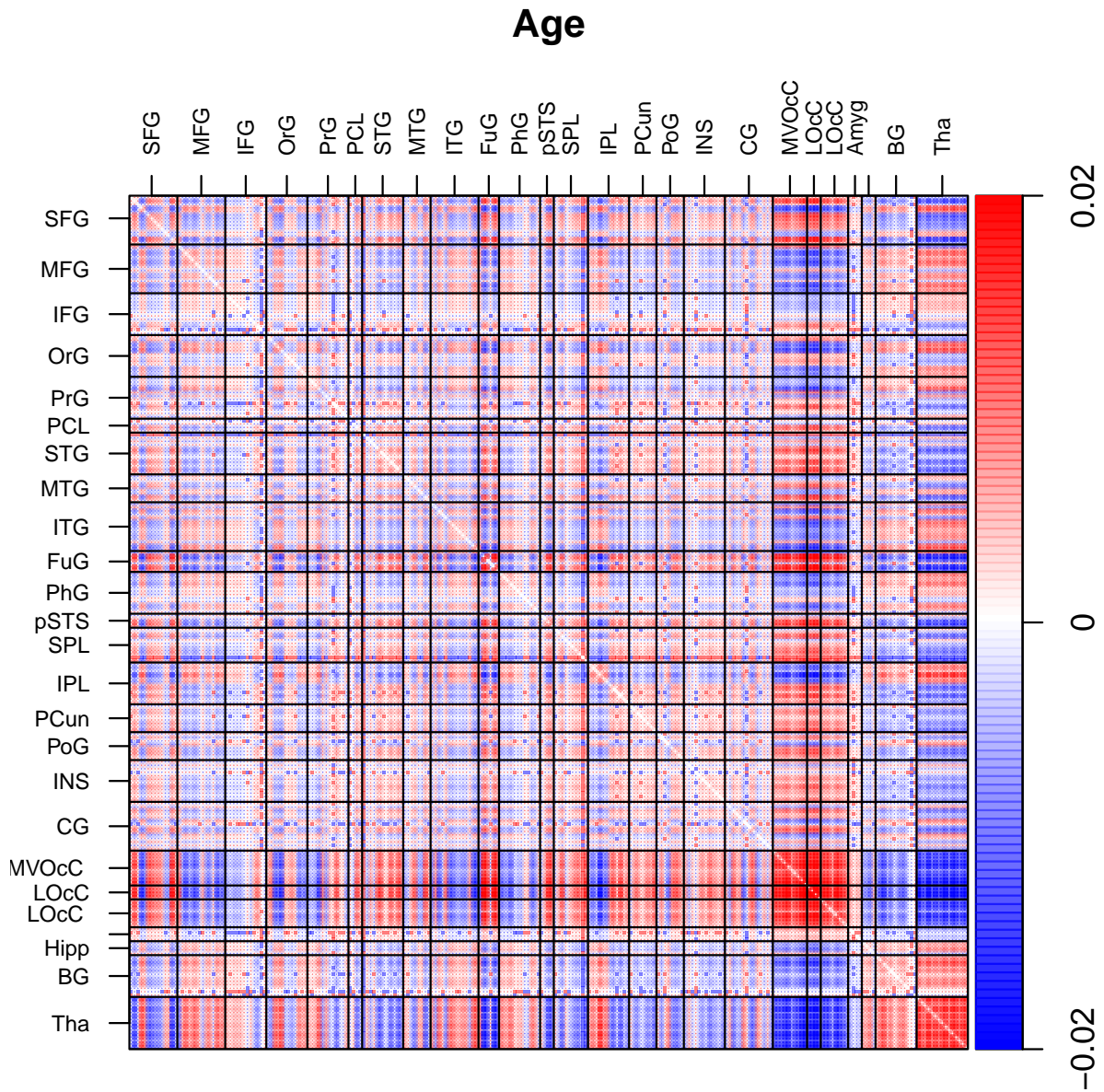

Age

Figure 8: Point estimates (median) of the regression coefficients for Cognitive Decline for the Network Matrix Edges using BMRR.
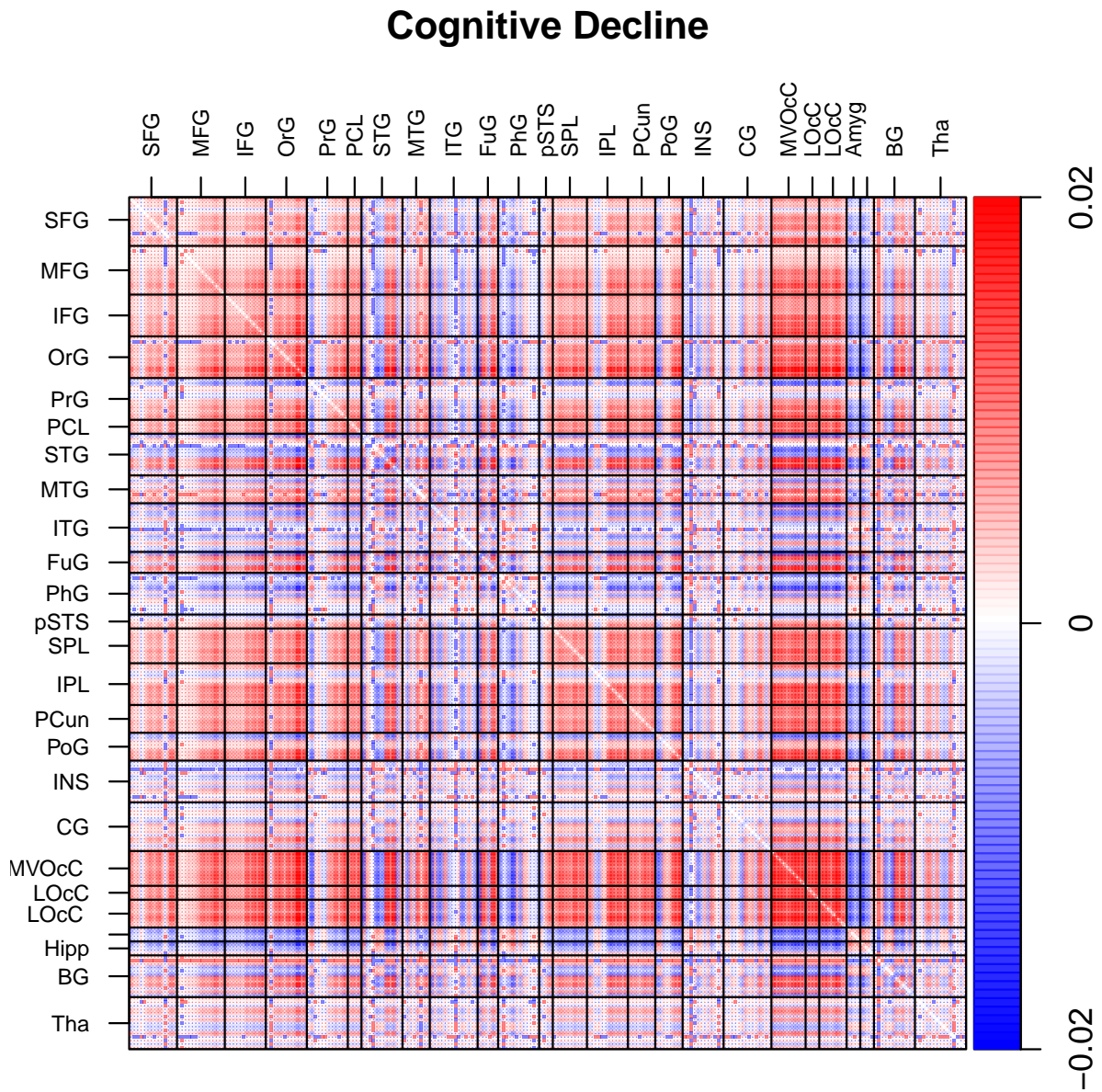
Figure 9: Point estimates (median) of the regression coefficients for grey matter probability for selected voxels across horizontal slices using MUA with speech rate as outcome.
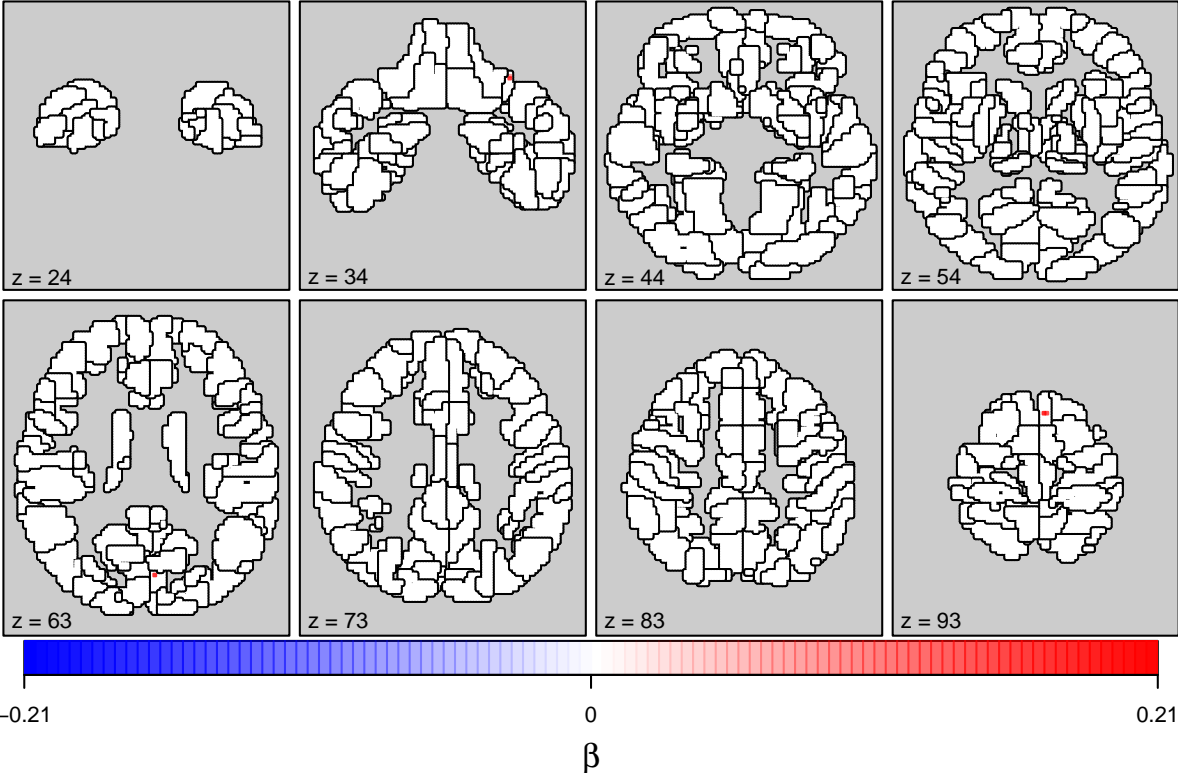
Figure 10: Point estimates (median) of the regression coefficients for Network Matrix z-Scores using MUA with speech rate as outcome. Each cell displays the ROI level network coefficient maps organized by anatomical brain region where the top and bottom half of each anatomical region indicated by a horizontal dash corresponds to the left hemisphere and right hemispheres, respectively.