Multimodal IVA fusion for detection of linked neuroimaging biomarkers

Rogers Silva[a,*], Eswar Damaraju[a], Xinhui Li[a,b], Peter Kochonov[d], Aysenil Belger[e], Judith M. Ford[f,g], Sarah McEwen[h], Daniel H. Mathalon[f,g], Bryon A. Mueller[i], Steven G. Potkin[j], Adrian Preda[j], Jessica A. Turner[k], Theo G.M. van Erp[l,m], Tulay Adali[c], Vince D. Calhoun[a,b]

[a] *Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Atlanta, GA, USA*
[b] *Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA*
[c] *Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD, USA*
[d] *Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD, USA*
[e] *Department of Psychiatry, University of North Carolina School of Medicine, Chapel Hill, NC, USA*
[f] *Veterans Afairs San Francisco Healthcare System, San Francisco, CA, USA*
[g] *Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, CA, USA*
[h] *Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, USA*
[i] *Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minneapolis, MN, USA*
[j] *Department of Psychiatry and Human Behavior, University of California, Irvine, Irvine, CA, USA*
[k] *Psychology Department, Georgia State University, Atlanta, GA, USA*
[l] *Clinical Translational Neuroscience Laboratory, Department of Psychiatry and Human Behavior, University of California Irvine, Irvine, CA, USA*
[m] *Center for the Neurobiology of Learning and Memory, University of California Irvine, Irvine, CA, USA*

Abstract

With the increasing availability of large-scale multimodal neuroimaging datasets, it is necessary to develop data fusion methods which can extract cross-modal features. A general framework, multidataset independent subspace analysis (MISA), has been developed to encompass multiple blind source separation approaches and identify linked cross-modal components in multiple datasets. In this work we utilized the multimodal independent vector analysis model in MISA to directly identify meaningful linked features across three neuroimaging modalities — structural magnetic resonance imaging (MRI), resting state functional MRI and diffusion MRI — in two large independent datasets, one comprising of healthy subjects and the other including patients with schizophrenia. Results show several linked subject profiles (the sources/components) that capture age-associated reductions, schizophrenia-related biomarkers, sex effects, and cognitive performance.

1. Introduction

Multimodal neuroimaging data can provide rich information to better understand brain structures and functions, and boost biomarker detection (Uludag and Roebroeck, 2014). Although analysis of each data modality separately can yield important insights into the structural or functional integrity of the brain, the relationship between different views from multimodal neuroimaging data is often complex and unknown. Therefore, data-driven approaches are required to discover brain structure-function relationships, which may not occur in the same regions and may covary among subjects in complex ways. As more research institutions are participating in open science data sharing practices, large-scale neuroimaging datasets (>1000 subjects) with multimodal data are widely becoming available, giving researchers opportunities to develop novel approaches to multimodal fusion analysis that can offer insights into cross-modal (joint) associations and identify important missing links in brain development and mental disorders (Calhoun and Sui, 2016).

*Corresponding author
Email address: rsilva@gsu.edu (Rogers Silva)

Blind source separation (BSS) techniques, in particular independent component analysis (ICA) (Comon, 1994; Bell and Sejnowski, 1995) have gained popularity in neuroimaging analysis due to their minimal assumptions about the latent sources, availability, and interpretability of unmixed latent sources. However, ICA is only suitable for single modality data. Several other methods have been developed to extend ICA to extract multimodal features, including canonical correlation analysis (Hotelling, 1936), joint ICA (Calhoun et al., 2006), parallel ICA (Liu et al., 2008), independent vector analysis (IVA, a generalization of ICA to multiple modalities) (Adali et al., 2014a; Kim et al., 2006). Recently, we proposed a data driven blind source separation model called multidataset independent subspace analysis (MISA) (Silva et al., 2021) that generalizes many existing BSS technique such as IVA to multiple datasets and also incorporates subspaces to identify linked unimodal or linked cross-modal components. MISA utilizes the Kotz distribution (Kotz, 1975) to model source distributions, thus leveraging all-order statistics (second- and higher-order) to model the underlying source subspaces. Another advantage of MISA is that it allows comparison across many different types of BSS. In this work, we use MISA to implement the multimodal IVA (MMIVA) model (see section 2.3) and demonstrate its applicability on two large neuroimaging datasets: one including typical adults (UK Biobank) and the other including pooled data from multiple studies/sites that investigate psychosis and age-matched controls.

We evaluated the model on standard derivative data from three neuroimaging modalities: 1) Gray matter segmentations (GM) from structural magnetic resonance imaging (sMRI) data whose intensities markedly convey regional variations in gray matter densities with age, sex and other neurobiological factors, 2) Amplitude of low frequency fluctuations (ALFF) maps computed from resting state functional MRI (rfMRI) scans that inform about the strength of local connectivity as well as potential for long-range associations, and 3) Fractional anisotropy (FA) maps obtained from diffusion MRI (dMRI) scans which characterize the degree of directional water diffusion in a brain white matter bundles. The results show cross-modal associations in the two independent datasets and several of these linked sources show covariation with age, sex and also discriminate patients with psychosis.

In the following, Section II describes the data, preprocessing, and methodology utilized in this work. Section III presents our results, which are further discussed in Section IV before presenting our final conclusions.

2. Methods

2.1. Data

We used two large independent multi-site datasets to evaluate the MMIVA model. For the first dataset, we used imaging data from a subset of 3500 subjects participating in the UK Biobank study (Miller et al., 2016a), a prospective epidemiological study with a large imaging database. Specifically, we utilized multivariate features (Calhoun and Adali, 2008) from sMRI, dMRI and rfMRI data extracted from each subject. All data included in our analysis were collected in two of the three participating locations in the United Kingdom. All participants provided informed consent from their respective institutional review boards.

T1-weighted structural MRI images were acquired using a 3D MPRAGE sequence at 1mm^3 isotropic sagittal slices with acquisition parameters: $208 \times 256 \times 256$ matrix, R=2, TI/TR=880/2000 ms. Diffusion MRI images were acquired using a standard Stejskal-Tanner spin-echo sequence at 2mm^3 isotropic resolution at two different b values (b = 1,000 and 2,000 s/mm^2) and 50 distinct diffusion-encoding directions each acquired with a multi-band (MB) factor of 3. Resting functional MRI were acquired axially at 2.4mm^3 resolution while the subjects fixated at a cross with the following acquisition parameters: $88 \times 88 \times 64$ matrix, TE/TR=39/735ms, MB=8, R=1, flip angle 52°. A pair of spin echo scans of opposite phase encoding direction in the same imaging resolution as rfMRI scans were acquired to estimate and correct distortions in rfMRI echo-planar images, and a single

band high resolution reference image was acquired at the start of the rfMRI scan to ensure good realignment and normalization.

The second dataset includes pooled data from 4 studies that collected imaging data from patients with schizophrenia, schizo-affective disorder, bipolar disorder and age-matched healthy controls. These studies are Center of Biomedical Research Excellence in brain function and mental illness (COBRE) (Aine et al., 2017), function biomedical informatics network (fBIRN) (Keator et al., 2016), Maryland Psychiatric Research Center (MPRC), and Bipolar and Schizophrenia Network for Intermediate Phenotypes (BSNIP) (Tamminga et al., 2014). The dataset information is summarized in Table 1.

Table 1: Dataset information.

Dataset	Site	N (M,F)	Mean Age (std) in Years	Age Range in Years
UK Biobank Dataset				
UKB	1	2379 (1201,1178)	63.1 (7.2)	48-79
UKB	2	528 (251,277)	61.8 (7.3)	46-79
Patient Datasets				
COBRE	1	160 (123,37)	37.79 (12.8)	18-65
MPRC	1	60 (44,16)	38.67 (13.2)	16-62
MPRC	2	106 (43,63)	40.28 (13.1)	16-64
MPRC	3	1 (1,0)	55.00 (0.0)	55-55
BSNIP	1	237 (133,104)	37.62 (14.7)	16-65
BSNIP	2	192 (103,89)	40.13 (13.2)	15-65
FBIRN	1	24 (19,5)	36.00 (9.9)	22-51
FBIRN	2	17 (14,3)	40.88 (9.9)	23-57
FBIRN	3	53 (41,12)	42.81 (12.5)	19-60
FBIRN	4	54 (44,10)	36.04 (11.4)	20-58
FBIRN	5	16 (9,7)	38.62 (9.4)	22-53
FBIRN	6	44 (21,23)	34.93 (10.6)	19-58
FBIRN	7	35 (30,5)	37.91 (10.2)	18-60

2.2. Preprocessing

We processed each of the three imaging data modalities to obtain GM, ALFF, and FA feature maps, which were then used for multimodal fusion analysis.

Specifically, the sMRI images underwent segmentation and normalization to MNI space using the SPM12 toolbox, yielding GM, white matter (WM), and cerebro-spinal fluid (CSF) tissue probability maps. The normalized GM segmentations were spatially smoothed using a 10mm FWHM Gaussian filter. The smoothed images were resampled to 3mm^3. We defined a group mask to restrict the analysis to GM voxels as follows. First, an average GM segmentation map from all subjects was obtained from normalized segmentation images at 1mm^3 resolution. This map was binarized at a 0.2 threshold (any voxel with a group average probability > 0.2 for being classified as gray matter) and resampled to 3mm^3 resolution, which resulted in 44,318 in-brain voxels.

We used distortion corrected, FIX-denoised (Griffanti et al., 2014), normalized rfMRI data provided by the UK Biobank data resource. We computed ALFF maps, defined as the area under the low frequency band [0.01-0.08 Hz] power spectrum of each voxel time course. We then obtained mean scaled ALFF maps (mALFF), ALFF maps divided by the global mean ALFF value, as this scaling has been shown to result in greater test-retest reliability of ALFF maps (Zhao et al., 2018). The mALFF maps were smoothed using a 6mm FWHM Gaussian filter and resampled to 3mm isotropic

voxels. We used the same group mask learned from GM features for mALFF maps in the subsequent fusion analysis.

For dMRI data, we used the FA maps provided by the UK Biobank consortium. The preprocessing steps that raw dMRI images underwent are thoroughly described in (Alfaro-Almagro et al., 2018). The FA maps were then spatially smoothed using a 6mm FWHM Gaussian filter and resampled to 3mm^3 voxels. For dMRI data, we computed a group mask similar to the approach described above for sMRI data. However, the group average WM segmentation was binarized at a threshold of 0.4, resulting in 18,684 in-brain voxels in the group mask.

Prior to running our fusion model, site was regressed out voxelwise for both datasets to account for mean differences in feature maps due to scanner effects.

2.3. Multimodal IVA (MMIVA) fusion model

Here we present a general IVA approach for direct analysis of heterogeneous multimodal data. As mentioned earlier, IVA is a natural extension of ICA. While ICA operates on a single dataset to obtain statistically independent source signals via estimation of *one* linear unmixing matrix, IVA performs joint estimation of *many* unmixing matrices simultaneously across multiple datasets (Adali et al., 2014b).

Briefly, ICA is a blind source separation model that assumes linear mixing of C statistically independent sources $\mathbf{s}$, yielding the observed data $\mathbf{x}$:

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n),\ 1 \leq n \leq N, N \in \mathbb{Z},\ \mathbf{x}(n), \mathbf{s}(n) \in \mathbb{R}^C, \tag{1}$$

where $\mathbf{A}$ is the mixing matrix, and N is the number of observations (here, the number of subjects). The ICA algorithm seeks to identify the sources $\hat{\mathbf{s}}(n) = \mathbf{W}\mathbf{x}(n)$ by estimating an unmixing matrix $\mathbf{W}$, according to certain properties of the sources such as higher-order statistics and non-Gaussianity. A typical strategy for solving the ICA problem involves minimizing the mutual information among the sources defined as:

$$\mathcal{I}^{\text{ICA}}(\mathbf{W}) = \sum_{i=1}^{C} H(\hat{s}_i) - \log|\det\mathbf{W}|, \tag{2}$$

where $H(\hat{s}_i)$ is the differential entropy, given by $H(\hat{s}_i) = -\,\mathbb{E}\{\log p_{\hat{s}_i}(w_i^\top x)\}$.

IVA extends the ICA model to multiple (K) datasets, assuming a linear mixture of C independent sources for each dataset:

$$\mathbf{x}^{[k]}(n) = \mathbf{A}^{[k]}\mathbf{s}^{[k]}(n), 1 \leq k \leq K,\ 1 \leq n \leq N, K, N \in \mathbb{Z} \tag{3}$$

additionally assuming statistical *dependence* (i.e., linkage) of corresponding sources. This collection of linked sources is defined as the source component vector (SCV) $\mathbf{s}_i(n) = [s_i^{[1]}(n), s_i^{[2]}(n), \ldots, s_i^{[k]}(n)]^\top \in \mathbb{R}^K$. Here, $K = 3$, such that each SCV spans across three modalities.

Solving the IVA problem comes from minimizing the following mutual information:

$$\mathcal{I}^{\text{IVA}}(\mathbf{W}) = \sum_{i=1}^{C}\left(\sum_{k=1}^{K} H(\hat{s}_i^{[k]}) - \mathcal{I}(\hat{\mathbf{s}}_i)\right) - \sum_{k=1}^{K}\log|\det(\mathbf{W}^{[\text{k}]})| \tag{4}$$

The second term $\mathcal{I}(\hat{\mathbf{s}}_i)$ in the equation above is mutual information accounting for dependence among sources in each SCV. Altogether, the terms in the big parentheses correspond to the joint entropy of an SCV $H(\hat{\mathbf{s}}_i)$, simply indicating that IVA identifies independence *among* SCVs while taking into account the dependence across datasets. See Adali et al. (2014b) for a general discussion on ICA and IVA algorithms and Adali et al. (2015b,a) for their application to data fusion (particularly our choice of transposed IVA). See Silva et al. (2021) for details on the MISA implementation we utilized to estimate the IVA model in this work.

We performed MMIVA fusion of the GM, mALFF, and FA features by treating each modality as one of the K datasets in the IVA model above. Each subject's feature set was z-scored per modality and then reduced to 20 principal directions using multimodal group principal component analysis (MGPCA). Unlike standard PCA that finds orthogonal directions of maximal variation for each modality separately, MGPCA finds directions of maximal *common* variation, i.e., eigenvectors are computed based on the average of the scaled covariance matrices ($\mathbf{\Sigma}^{[k]}$) of all three modalities. The scaling factor used is trace($\mathbf{\Sigma}^{[k]}$)/N, which is the ratio of the variance in the modality to the number of observations (subjects). The MGPCA-reduced data then underwent an ICA estimation using the Infomax objective (Bell and Sejnowski, 1995) to obtain 20 *common* independent sources.

We improved upon the Infomax estimation by configuring and running MISA as an ICA model, in which case it assumes source distributions to follow a *univariate* Kotz distribution. The final combined MGPCA+ICA estimates of $\mathbf{W}^{[k]}$ were then utilized to initialize the MMIVA model after reconfiguring and running MISA as an IVA model to obtain the final joint decomposition. As discussed earlier, MMIVA accounts for dependence within corresponding sources across modalities. For both MISA and MMIVA models, the source distributions are assumed to take a multivariate Kotz distribution, as this has been shown to generalize well across multivariate Gaussian, multivariate Laplace, and multivariate power exponential source distributions (Anderson et al., 2013). All methods were implemented using the MISA toolbox (Silva et al., 2021).

2.4. Statistics

UK Biobank provides extensive phenotype information for each subject including age, sex, lifestyle measures, cognitive scores, etc. We used a subset of the subject measures (SM) reported in (Miller et al., 2016b) to identify associations between the subject demographics and the MMIVA sources obtained from our decomposition. We computed a multivariate MANCOVA model using the MANCOVAN toolbox, which implements multivariate stepwise regression, to identify associations between subject demographics and MMIVA sources for each modality separately. Subsequently, univariate tests are performed for each component with the terms of design matrix that survive the multivariate Lawley-Hotelling trace test. The p-values were then corrected for multiple comparison correction at false discovery rate of 0.05.

Following the approach in Smith et al. (2015), we dropped subject scores with more than 4% missing data. This resulted in 2907 subject scores out of 3497 for MANCOVAN analysis. Of 64 SMs, we dropped 10 columns which had extreme values. Extreme values are identified in 2 steps. First sum of square of absolute median deviations (ssqamdn) for each SM is computed. If there is any max(ssqamdn) >100*mean(ssqamdn), then the SM has subjects with extreme outliers which can influence statistical analysis and so it was dropped. This resulted in 54 phenotypes that include age, sex, fluid intelligence, a set of phenotypes covering amount and duration of physical activity, frequency of alcohol intake, and cognitive test scores (see Miller et al. (2016b) for details).

For the measures that were retained, any missing values were imputed with K-Nearest Neighborhood (MATLAB's knnimpute) method (Cunningham and Delany, 2021). Stepwise regression approach was used to eliminate insignificant terms ($\alpha < 0.01$) at each step. Univariate tests were performed and corrected for multiple comparisons at Bonferroni threshold (0.05/30 for 30 sources). In addition to the SMs, the following nuisance covariates were added to the design matrix:

- sMRI: correlation of warped subject segmentation map to mean segmentation map,
- dMRI: correlation of warped subject FA map to mean FA map,
- rs-fMRI: correlation of warped subject ALFF map to mean ALFF map,
- mean framewise displacement (mFD) computed from rigid body movement estimates from resting fMRI scan realignment step.

Any variables with fewer than 8 levels were modeled as categorical variables and the rest were modeled as continuous variables. Only age by sex interaction was considered.

2.5. Brain-age delta modeling

To further evaluate the significant age-related SCVs identified with MANCOVA, we evaluated the difference between the predicted brain age and the actual chronological age (the brain-age delta) in two steps, as described in Smith et al. (2020). In the first step, all three multimodal subject expression levels from five SCVs with significant age association (based on the multivariate MANCOVA assessment in 2.4) were used to predict age. Thus, the initial estimate of brain-age delta was evaluated as:

$$\delta_1 = X\beta_1 - Y, \tag{5}$$

where Y is the actual age after removing the mean age across subjects, X is the matrix of significant age-related SCVs with size $N \times \bar{C}$ (N = number of subjects, $\bar{C} = 5$ components $\times 3$ modalities $= 15$), and β_1 is the $\bar{C} \times 1$ vector of regression parameters. As such, δ_1 shows the aging effects that cannot be captured by the imaging modalities.

In the second step, the model is refined to identify aspects of δ_1 that cannot be captured by age terms (notice that δ_1 is by definition dependent on age) nor confounds, such as sex. Thus, δ_2 is evaluated as:

$$\delta_2 = \delta_1 - Y_2\beta_2, \tag{6}$$

where the regression matrix Y_2 includes linear, quadratic and cubic age terms, confound variables (sex and framewise displacement), the interaction between the sex variable and each of three age terms, the framewise displacement variable after regressing out the age and sex effects. Effects among zero-mean variables were also regressed out, for example, the (zero-mean) linear age effect was removed from the (zero-mean) quadratic and (zero-mean) cubic age terms. After that, the mean was removed from each term again.

The β regression coefficients can be estimated as usual by $\beta_1 = (X^\top X)^{-1}X^\top Y$ and $\beta_2 = (Y_2^\top Y_2)^{-1}Y_2^\top \delta_1$. Note that in the equations above all three multimodal subject expressions from all five SCVs are fit in the same model. Source-specific contributions to the overall brain aging delta can be described by their individual delta vectors. These source-specific contributions can be combined to recover the original Y (here, $\frac{1}{\bar{C}}$ is arbitrary and irrelevant since it is removed in the following step):

$$\sum_{i=1}^{\bar{C}} \frac{1}{\bar{C}} Y = \sum_{i=1}^{\bar{C}} (X_i\beta_{1i} - \delta_{1i}). \tag{7}$$

The source-specific delta vector at the first stage, δ_{1i}, can be calculated by:

$$\delta_{1i} = X_i\beta_{1i} - \frac{Y}{\bar{C}}. \tag{8}$$

In the next step, we compute the regression coefficients $\beta_{2i} = (Y_2^\top Y_2)^{-1}Y_2^\top \delta_{1i}$ and the individual delta vector at the second stage is obtained:

$$\delta_{2i} = \delta_{1i} - Y_2\beta_{2i}. \tag{9}$$

We repeated the above individual brain-age delta estimation process (8) and (9) for each source in each SCV.

3. Results

The MANCOVA analysis revealed significant effects of age, psychosis, sex, and fluid intelligence across the three imaging modalities. Table 2 below summarizes the strength of the multimodal link captured by the proposed multimodal IVA approach within the SCVs selected by the MANCOVA.

Table 2: Pearson correlation of SCVs between modalities.

SCV	sMRI-fMRI	fMRI-dMRI	sMRI-dMRI
SCVs in UK Biobank dataset			
3	0.720	0.840	0.667
5	0.639	0.764	0.613
8	0.643	0.920	0.634
10	0.740	0.700	0.703
16	0.773	0.908	0.746
17	0.728	0.785	0.693
SCVs in patient datasets			
19	0.591	0.675	0.552

3.1. Associations with aging in UK Biobank

For the UK Biobank data, five SCVs (3, 5, 8, 10, and 16) showed significant age effects for all three modalities after correcting for multiple comparisons ($p_{corrected} < 0.05$). The sources from all SCVs with age effects had their sign oriented in the direction of decline with age. Figure 1 presents the linked subject profile source (SCV 8) with the most significant age association, and the corresponding mixing weights (spatial maps), which highlight regions associated with age reductions for each of the three modalities. Age-associated reductions can be observed in 1) caudate, thalamus, insular regions, anterior and posterior cingulate cortex and lingual gyrus in sMRI, 2) the visual cortex in fMRI, 3) superior and posterior thalamic radiation in dMRI. The remaining age-associated components are shown in the supplementary section Figure 6.

The subject expression level sources from each modality in these five SCVs (15 sources in total) were subsequently used to predict the brain age and estimate the brain-age delta. The mean absolute error of the overall brain-age delta estimation is 2.80 years when using all fifteen sources in one model. The mean absolute error are 2.77, 2.45 and 2.73 years from sMRI, fMRI and dMRI respectively. Figure 2 (from top to bottom) shows 1) the predicted expression levels of each source as a function of a cubic age model which includes the linear, quadratic and cubic age terms; 2) the correlation between age and the expression levels of each source, the regression coefficient and regression significance at the first stage; 3) the standard deviation of the estimated brain-age delta at the second stage, its mean absolute error, and the standard deviation of predicted age at the first stage. Seven out of the fifteen sources have a β coefficient with significant age dependence ($p < 0.05/15$, Bonferroni correction). The bottom panel indicates that the sMRI sources from SCVs 8 and 16 explain the largest amount of the age variance, with smaller but significant contributions observed in the fMRI and dMRI modalities (SCVs 3 and 5).

3.2. Associations with schizophrenia identification

Similarly, a linked subject expression pattern (SCV 19, see Figure 3) can be identified in the patient dataset. The SCV not only shows areas with age-related decline, but also demonstrates group differences between age-matched controls and patients with schizophrenia. The patients show significantly lower source intensities in all three modalities ($p_{corrected} < 0.001$). Schizophrenia-related effects can be observed from the indicated areas in the multimodal spatial maps, specifically the temporal lobe in sMRI, the occipital and the parietal lobes in fMRI, and the corpus callosum and lateral ventricle areas in dMRI.

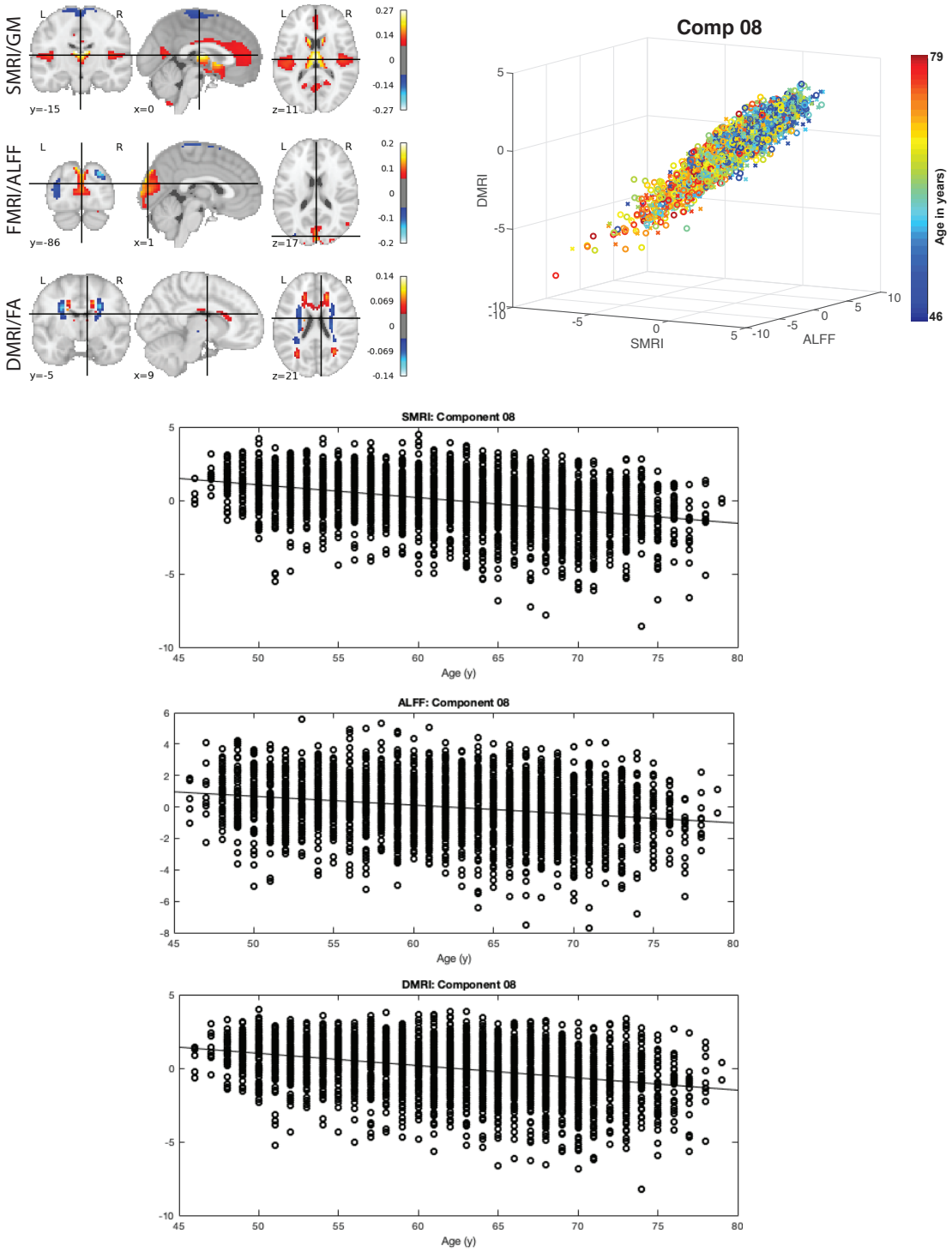

Figure 1: **Age-related SCV**. Top right: A 3D scatter plot of SCV 8 illustrates the association between the identified multimodal subject expression levels and age for UK Biobank data. Each point represents a subject, color coded by the age (circles indicate females; crosses indicate males). Top left: The spatial maps corresponding to the mixing weights for each modality are also shown. Bottom: The same data as in the top right panel is depicted separately for each modality and plotted against subject age.

3.3. Associations with sex effects in UK Biobank

Three SCVs (10, 16 and 17) show significant sex effects in UK Biobank data. Females show significantly higher source intensities than males in SCV 16 (see Figure 4) while males show significantly higher source intensities than females in SCVs 10 and 17 (see supplementary section, Figures 7 and 8).

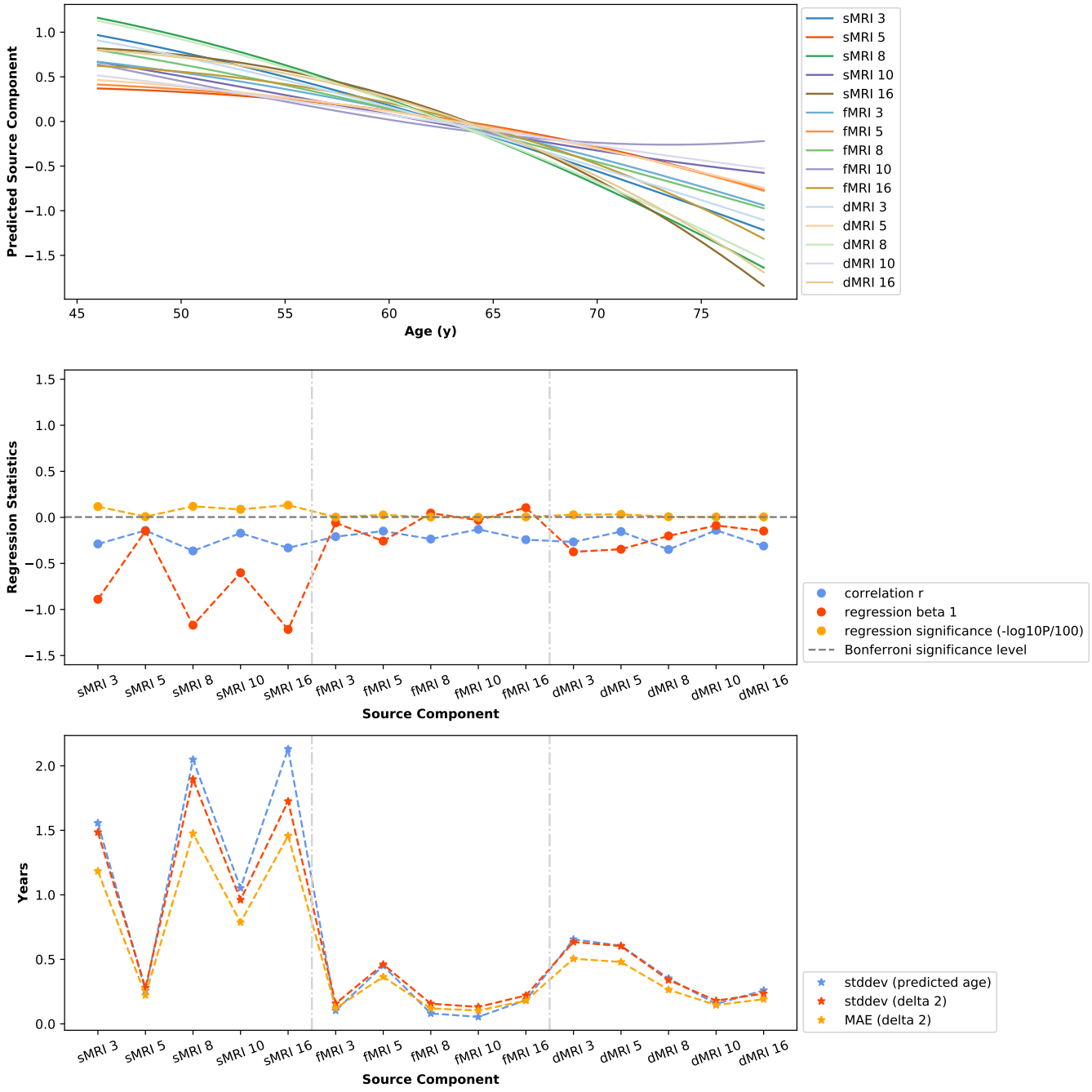

Figure 2: **Brain age delta modeling**. Top: SCV as a function of brain age. Middle: Regression statistics at the first brain-age delta modeling stage. Bottom: Standard deviation of predicted age, standard deviation of estimated brain-age delta, and mean absolute error of estimated brain-age delta at the second stage.

For SCV 16, in both sexes, notice the linear covariation with age and similar trajectory of decline with age observed in cerebellar regions (sMRI/GM), parietal cortex (fMRI/ALFF), and parietal cortico-pontine tracts (dMRI/FA).

Components 10 and 17 have negative weights in the parietal lobe and the cerebellum in sMRI respectively. In fMRI, component 16 shows negative weights in the temporal lobe, components 10 and 17 show positive weights in the frontal lobe. In dMRI, components 10 and 16 have positive weights in caudate nucleus while component 17 shows negative weights in the posterior cingulate cortex. As the age increases, regions with positive weights experience decline (conversely, increase in regions with negative weights).

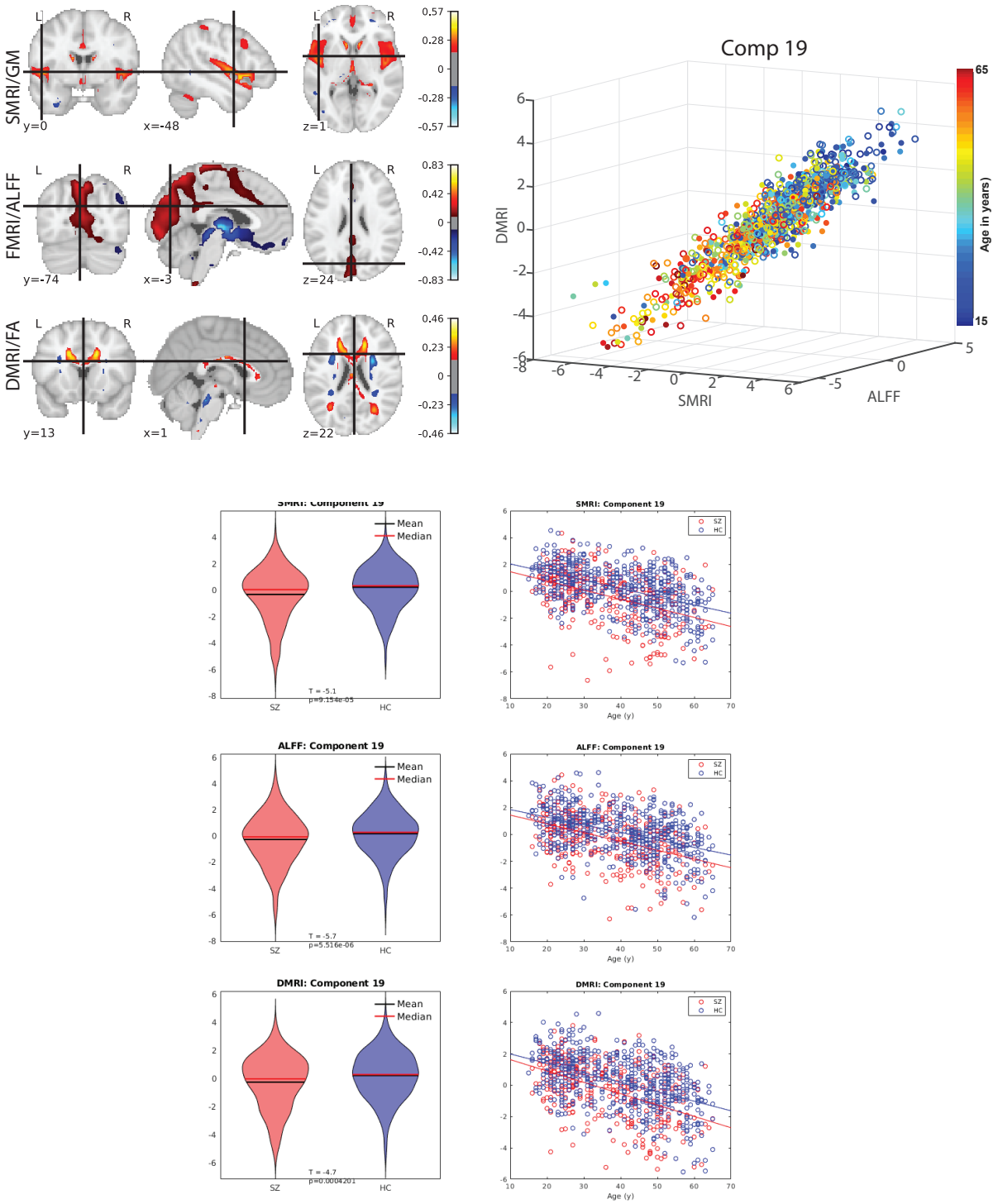

Figure 3: **Schizophrenia-related SCV**. The SCVs corresponding to component 19 for each of the three modalities for the patient data color coded by subject age. The component maps correspond to the mixing weights of the component for each modality are also shown. The SCVs are also plotted by control (HC) and schizophrenia patient group (SZ) as bar plots and by age as scatter plots to demonstrate consistent reduction in source intensities in SZ patients across all ages.

3.4. Associations with cognitive performance in UK Biobank

As shown in Figure 5, one linked SCV (10) shows significant linear association with fluid intelligence score across all three modalities. The higher the cognitive score is, the higher associated component intensities can be observed.

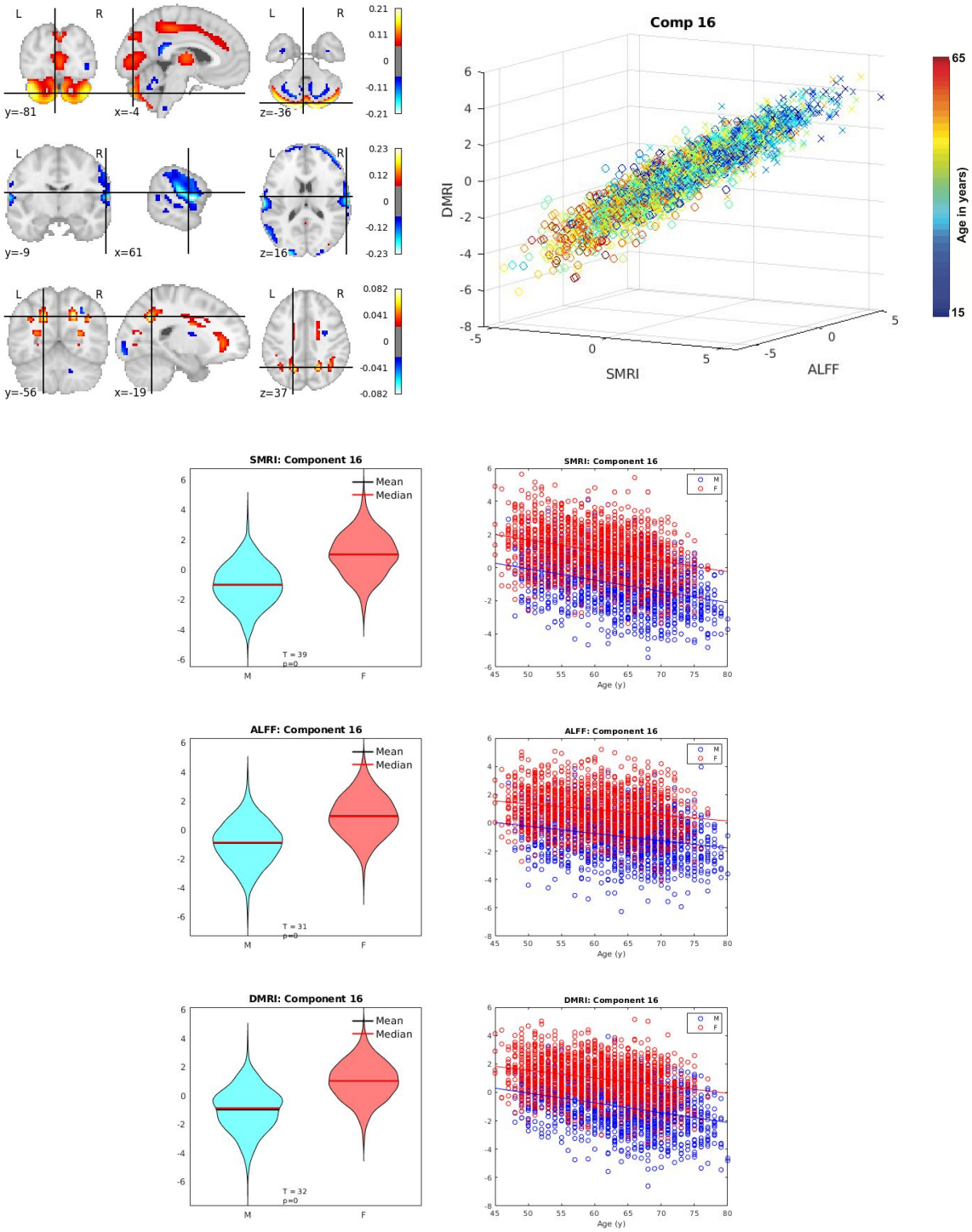

Figure 4: **Sex-related SCV**. Illustration of SCV 16. Top left: Spatial maps. Top right: Multimodal subject expression level sources, depicting all subjects (colored by age). Bottom: Source intensity differences related to sex effects.

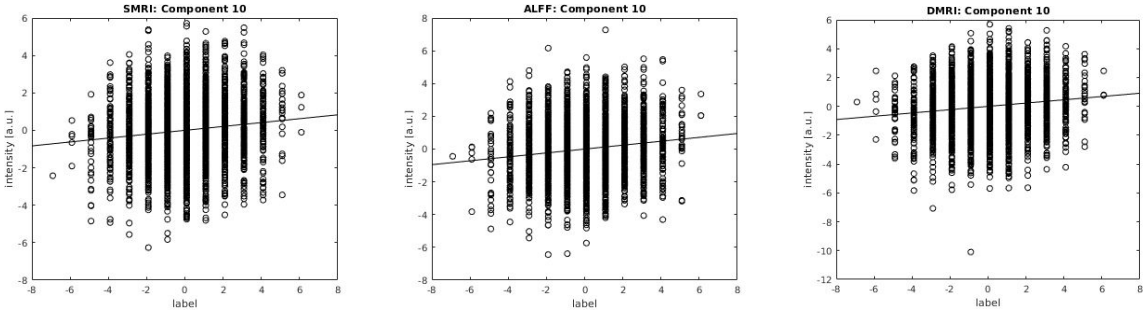

Figure 5: **Fluid intelligence-related SCV**. The SCVs corresponding to component 10 for each of the three modalities. Component 10 shows significant positive correlation between the cognitive score and the source intensities.

4. Discussion

In this work, we demonstrate that the proposed multimodal IVA model, initialized with a multimodal group projection (see Methods 2.3), can extract independent multimodal subspaces that show significant covariation with subject phenotypes. The proposed approach directly leverages multimodal associations and is applied to two large datasets independently. The results highlight several linked multimodal patterns of expression over subjects that are significantly associated with aging, sex, cognition, and psychosis indicators.

Age associated decline in gray matter density (hot spatial areas of sMRI SCV 8 weights) was primarily seen in caudate, thalamus, insular regions, anterior and posterior cingulate cortex, and lingual gyrus, consistent with earlier findings (Brickman et al., 2007). ALFF maps corresponding to SCV 8 suggest reductions in parietal and visual regions of the brain that covary with the structural changes. In dMRI, the identified spatial map weights show FA intensity reductions in periventricular regions, including superior and posterior thalamic radiation, are associated with the age.

After identifying five SCVs significantly associated with age using a MANCOVA approach, the brain-age delta was estimated using all 15 individual sources (3 per SCV) across all three modalities. Each individual source from each of the three modalities within an SCV describes a pattern of subject-specific expression levels that is linked (covaries) between modalities within that SCV, while being statistically independent from sources in other SCVs. The mean absolute error of overall age prediction from all 15 sources in one model is 2.80 years, and the mean absolute errors are 2.77, 2.45, and 2.73 years when using only the 5 individual sources from a single modality (sMRI, fMRI, and dMRI, respectively) — both cases based on the same multimodal IVA decomposition. The association between each individual component and the age for each subject was also computed. As shown in Figure 2, four sMRI sources (SCVs 3, 8, 10, 16), one fMRI source (SCV 5), and two dMRI sources (SCVs 3, 5) show significant age regression beta weights. Though the sMRI source from SCV 5 does not contribute significantly to the age regression model, the corresponding fMRI and dMRI sources from SCV 5 do show significant contribution. Thus, all five SCVs display significant age regression coefficients in at least one modality, and we conclude that each modality contributes complementary information not captured in other modalities. Moreover, the brain-age delta result using only the 5 sources from individual modalities supports that each modality is valuable and equally predictive of age, which is largely a consequence of the inherent multimodal link (correlation) among expression levels within SCVs. The advantage of such multimodal analysis is not only that it leverages hidden covariation across modalities to recover linked subject expression patterns, but also that it captures unique information from each modality by allowing these patterns to differ across modalities. Note that in our brain-age delta analysis, confound variables head size, scanner table position and scan-date-related slow drifts described in Smith et al. (2020) were not included.

Schizophrenia associated effects (SCV 19 from the patient dataset) were found as sMRI gray matter intensity reduction in the temporal lobe, consistent with previous findings (Karlsgodt et al., 2010). Schizophrenia associated reductions can also be observed in the medial parietal lobe, including the posterior cingulate cortex in fMRI ALFF maps, also aligned with previous literature (Venkataraman et al., 2012). The dMRI component map shows decreased FA values in the corpus callosum area, which is close to the superior longitudinal fasciculus and cingulate bundle reported in Kyriakopoulos et al. (2008). As is the nature of the proposed multimodal IVA approach, all of these regional changes covary at similar levels over subjects across all three modalities.

Sex differences were mainly reflected in SCV 10, 16 and 17. Females show significantly higher source intensities than males in SCV 16 while males show higher intensities than females in SCV 10 and 17. Consistent with previous studies (Liu et al., 2020; Ritchie et al., 2018), males have significantly higher source intensities than females in the parietal lobe (SCV 10) and the occipital lobe (SCV 17) from sMRI, the frontal lobe in both SCV 10 and 17 from fMRI, the caudate nucleus (SCV 10) and the posterior cingulate cortex (SCV 17) from dMRI. As shown in Figure 4 for SCV 16, females have higher intensities than males in the occipital lobe from sMRI, the temporal lobe from fMRI and the

caudate nucleus from dMRI. The decline with aging association can be found in the sMRI and dMRI modalities while the opposite association can be observed in the temporal lobe from ALFF maps.

In summary, we demonstrate the ability of multimodal independent vector analysis to extract linked multimodal modes of subject variations that capture different aspects of phenotypical information including aging effects, schizophrenia-related biomarkers, sex effects and cognitive performance across two large independent datasets. With the increasing demand of multimodal neuroimaging data analysis, the MMIVA fusion model can be applied to identify linked components with associated phenotypes across multiple neuroimaging modalities and multiple datasets.

5. Acknowledgments

This work was supported by NIH grant R01MH118695 and NSF grant #2112455.

References

Adali, T., Anderson, M., Fu, G.-S., 2014a. Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging. IEEE Signal Processing 31, 18–33.

Adali, T., Anderson, M., Fu, G.-S., 2014b. Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging. IEEE Signal Processing Magazine 31 (3), 18–33.

Adali, T., Levin-Schwartz, Y., Calhoun, V. D., 2015a. Multimodal data fusion using source separation: Application to medical imaging. Proceedings of the IEEE 103 (9), 1494–1506.

Adali, T., Levin-Schwartz, Y., Calhoun, V. D., 2015b. Multimodal data fusion using source separation: Two effective models based on ica and iva and their properties. Proceedings of the IEEE 103 (9), 1478–1493.

Aine, C. J., Bockholt, H. J., Bustillo, J. R., Cañive, J. M., Caprihan, A., Gasparovic, C., Hanlon, F. M., Houck, J. M., Jung, R. E., Lauriello, J., Liu, J., Mayer, A. R., Perrone-Bizzozero, N. I., Posse, S., Stephen, J. M., Turner, J. A., Clark, V. P., Calhoun, V. D., 2017. Multimodal neuroimaging in schizophrenia: Description and dissemination. Neuroinformatics 15 (4), 343–364.

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L., Griffanti, L., Douaud, G., Sotiropoulos, S. N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., et al., 2018. Image processing and quality control for the first 10,000 brain imaging datasets from uk biobank. Neuroimage 166, 400–424.

Anderson, M., Fu, G.-S., Phlypo, R., Adalı, T., 2013. Independent vector analysis, the Kotz distribution, and performance bounds. In: Proc IEEE ICASSP 2013. Vancouver, Canada, pp. 3243–3247.

Bell, A., Sejnowski, T., 1995. An information-maximization approach to blind separation and blind deconvolution. Neural Computation 7 (6), 1129–1159.

Brickman, A. M., Habeck, C., Zarahn, E., Flynn, J., Stern, Y., 2007. Structural mri covariance patterns associated with normal aging and neuropsychological functioning. Neurobiology of aging 28 (2), 284–295.

Calhoun, V. D., Adali, T., 2008. Feature-based fusion of medical imaging data. IEEE Transactions on Information Technology in Biomedicine 13 (5), 711–720.

Calhoun, V. D., Adali, T., Kiehl, K. A., Astur, R., Pekar, J. J., Pearlson, G. D., 2006. A method for multitask fmri data fusion applied to schizophrenia. Human brain mapping 27 (7), 598–610.

Calhoun, V. D., Sui, J., 2016. Multimodal fusion of brain imaging data: A key to finding the missing link(s) in complex mental illness. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 1 (3), 230 – 244, brain Connectivity in Psychopathology.

Comon, P., 1994. Independent component analysis, A new concept? In: SIGNAL PROCESSING 1994. pp. 287–314.

Cunningham, P., Delany, S. J., Jul 2021. k-nearest neighbour classifiers - a tutorial. ACM Computing Surveys 54 (6), 1–25.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., Zsoldos, E., Ebmeier, K. P., Filippini, N., Mackay, C. E., et al., 2014. Ica-based artefact removal and accelerated fmri acquisition for improved resting state network imaging. Neuroimage 95, 232–247.

Hotelling, H., 1936. Relations between two sets of variates. Biometrika 28 (3/4), 321–377.

Karlsgodt, K. H., Sun, D., Cannon, T. D., 2010. Structural and functional brain abnormalities in schizophrenia. Curr Dir Psychol Sci 19 (4), 226–231.

Keator, D. B., van Erp, T., Turner, J. A., Glover, G. H., Mueller, B. A., Liu, T. T., Voyvodic, J. T., Rasmussen, J., Calhoun, V. D., Lee, H. J., Toga, A. W., McEwen, S., Ford, J. M., Mathalon, D. H., Diaz, M., O'Leary, D. S., Jeremy Bockholt, H., Gadde, S., Preda, A., Wible, C. G., Stern, H. S., Belger, A., McCarthy, G., Ozyurt, B., Potkin, S. G., FBIRN, 2016. The function biomedical informatics research network data repository. NeuroImage 124, 1074–1079.

Kim, T., Eltoft, T., Lee, T.-W., 2006. Independent vector analysis: An extension of ica to multivariate components. In: International conference on independent component analysis and signal separation. Springer, pp. 165–172.

Kotz, S., 1975. Multivariate distributions at a cross road. In: Patil, G. P., Kotz, S., Ord, J. K. (Eds.), A Modern Course on Statistical Distributions in Scientific Work. Springer, Netherlands, pp. 247–270.

Kyriakopoulos, M., Vyas, N. S., Barker, G. J., Chitnis, X. A., Frangou, S., 2008. A diffusion tensor imaging study of white matter in early-onset schizophrenia. Biological psychiatry 63 (5), 519–523.

Liu, J., Demirci, O., Calhoun, V. D., 2008. A parallel independent component analysis approach to investigate genomic influence on brain function. IEEE signal processing letters 15, 413–416.

Liu, S., Seidlitz, J., Blumenthal, J. D., Clasen, L. S., Raznahan, A., 2020. Integrative structural, functional, and transcriptomic analyses of sex-biased brain organization in humans. Proceedings of the National Academy of Sciences of the United States of America 117 (31), 18788–18798.

Miller, K., Alfaro-Almagro, F., Bangerter, N., et al., 2016a. Multimodal population brain imaging in the uk biobank prospective epidemiological study. Nature Neuroscience 19, 1523–1536.

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., Bartsch, A. J., Jbabdi, S., Sotiropoulos, S. N., Andersson, J. L., et al., 2016b. Multimodal population brain imaging in the uk biobank prospective epidemiological study. Nature neuroscience 19 (11), 1523–1536.

Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., Harris, M. A., Alderson, H. L., Hunter, S., Neilson, E., Liewald, D. C. M., Auyeung, B., Whalley, H. C., Lawrie, S. M., Gale, C. R., Bastin, M. E., McIntosh, A. M., Deary, I. J., 2018. Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants. Cerebral Cortex 28 (8), 2959–2975.

Silva, R. F., Plis, S. M., Adali, T., Pattichis, M. S., Calhoun, V. D., 2021. Multidataset independent subspace analysis with application to multimodal fusion. IEEE Transactions on Image Processing 30, 588–602.

Smith, S. M., Elliott, L. T., Alfaro-Almagro, F., McCarthy, P., Nichols, T. E., Douaud, G., Miller, K. L., 2020. Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. eLife 9, e52677.

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E., Glasser, M. F., Ugurbil, K., Barch, D. M., Van Essen, D. C., Miller, K. L., 2015. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nature neuroscience 18 (11), 1565–1567.

Tamminga, C. A., Pearlson, G., Keshavan, M., Sweeney, J., Clementz, B., Thaker, G., 2014. Bipolar and schizophrenia network for intermediate phenotypes: outcomes across the psychosis continuum. Schizophrenia bulletin 40, S131–S137.

Uludag, K., Roebroeck, A., 2014. General overview on the merits of multimodal neuroimaging data fusion. NeuroImage 102, 3 – 10, multimodal Data Fusion.

Venkataraman, A., Whitford, T. J., Westin, C. F., Golland, P., Kubicki, M., 2012. Whole brain resting state functional connectivity abnormalities in schizophrenia. Schizophrenia research 139 (1-3), 7–12.

Zhao, N., Yuan, L.-X., Jia, X.-Z., Zhou, X.-F., Deng, X.-P., He, H.-J., Zhong, J., Wang, J., Zang, Y.-F., 2018. Intra-and inter-scanner reliability of voxel-wise whole-brain analytic metrics for resting state fmri. Frontiers in neuroinformatics 12, 54.

6. Supplementary Materials

6.1. Aging Effects

Four other SCVs (component 3, 5, 10, 16) were significantly associated with age identified by the MANCOVA analysis.

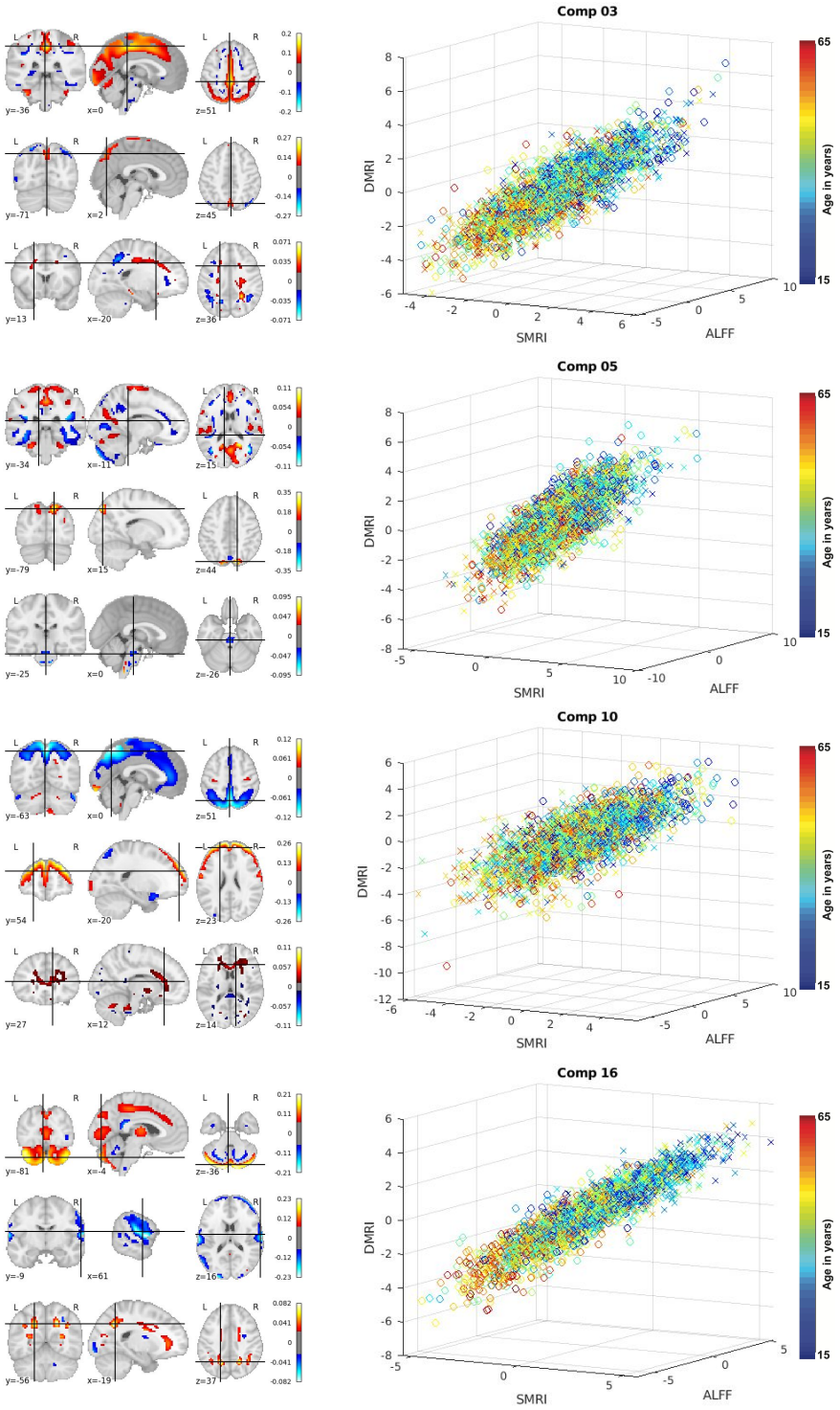

Figure 6: **Age-related SCV**. Right: 3D scatter plots of SCVs 3, 5, 10, 16 illustrate the association between the identified multimodal subject expression levels and age for UK Biobank data. Each point represents a subject, color coded by the age (circles indicate females; crosses indicate males). Left: The spatial maps corresponding to the mixing weights for each modality are also shown.

6.2. Sex Effects

SCVs (10, 17) were significantly associated with sex identified by the MANCOVA analysis.

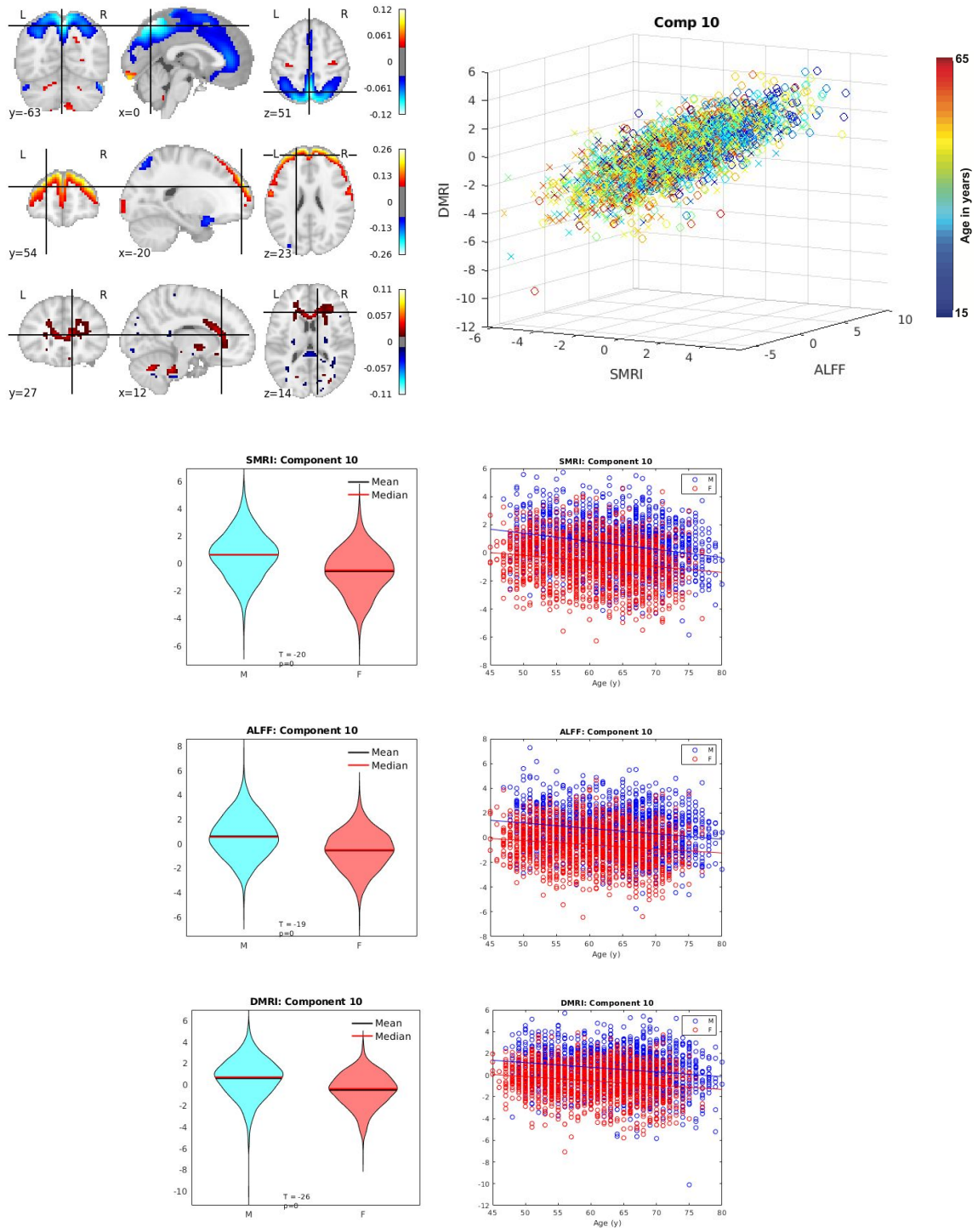

Figure 7: **Sex-related SCV**. Illustration of SCV 10. Top left: Spatial maps. Top right: Multimodal subject expression level sources, depicting all subjects (colored by age). Bottom: Source intensity differences related to sex effects.

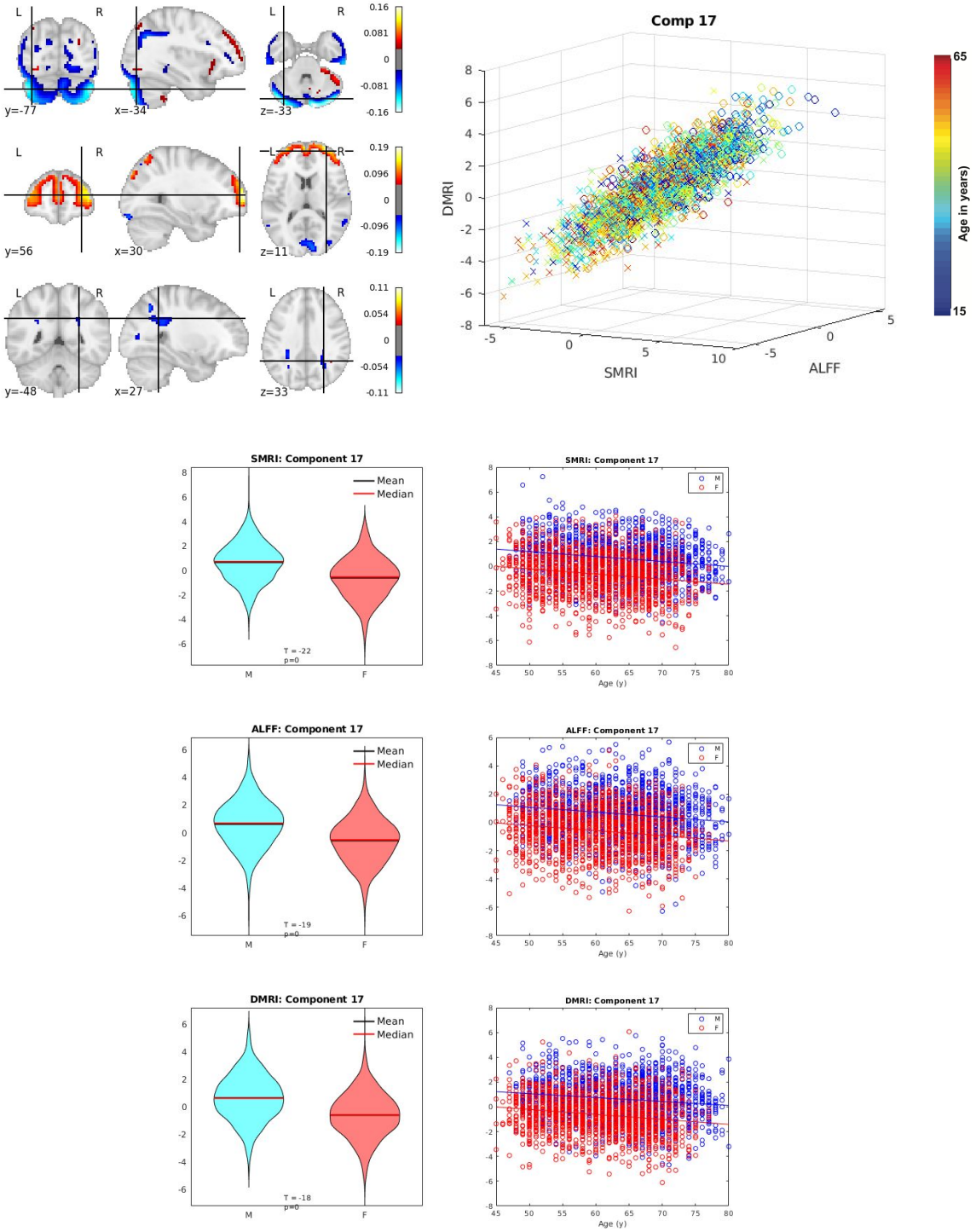

Figure 8: **Sex-related SCV**. Illustration of SCV 17. Top left: Spatial maps. Top right: Multimodal subject expression level sources, depicting all subjects (colored by age). Bottom: Source intensity differences related to sex effects.