An Approach to Automatically Label & Order Brain Activity/ Component Maps

Mustafa S. Salman
Tor D. Wager
Eswar Damaraju
Vince D. Calhoun

Joanne Wardell

The Problem of Noise in fMRI Data

Meaningful Changes are Often Small and Hidden

- BOLD signal change vs total MR signal intensity
- BOLD signal change vs total spatial and temporal variability across images

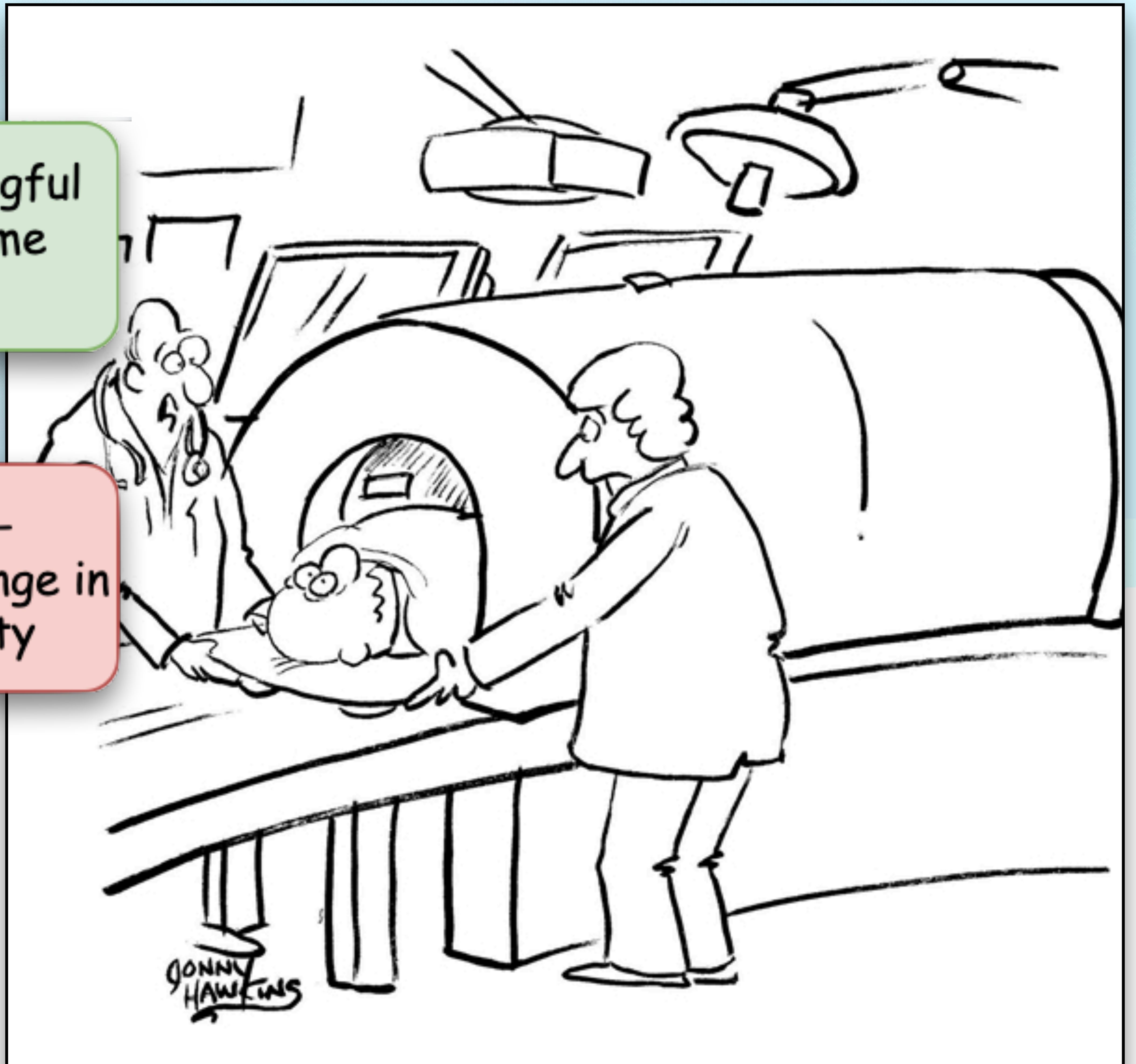
Signal: Meaningful change in some quantity

Noise: Non-Meaningful change in some quantity

Physiological Processes

Scanner Interference

Task Performance Variations

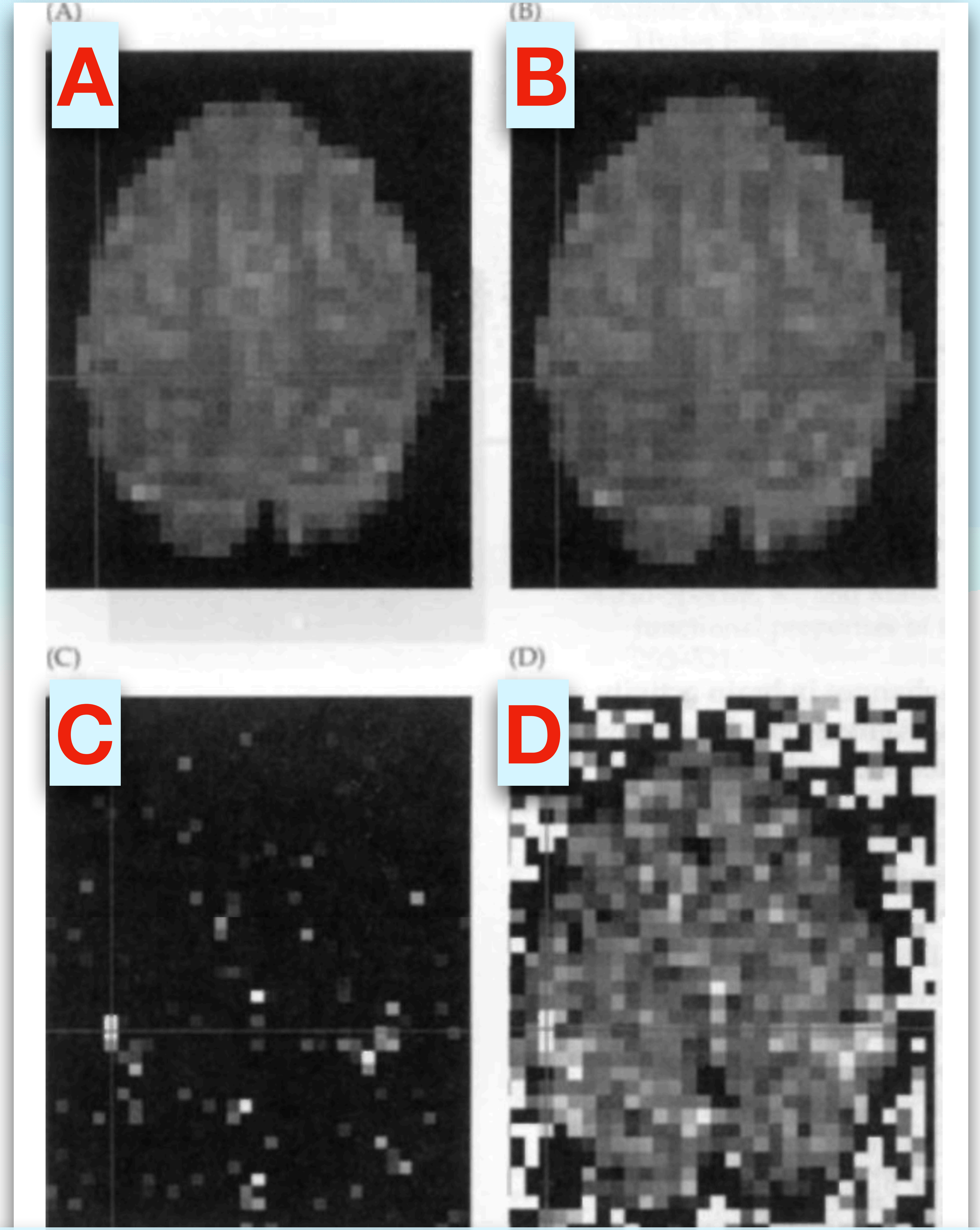


"Now just close your eyes and envision being the inside of a Twinkie."

BOLD Signal Variation Illustrated

Small but Meaningful changes in Brain Activity Can be Masked by Noise

- A => Resting State
- B => Performing Task
- C => A - B
- D => Percent Signal Change



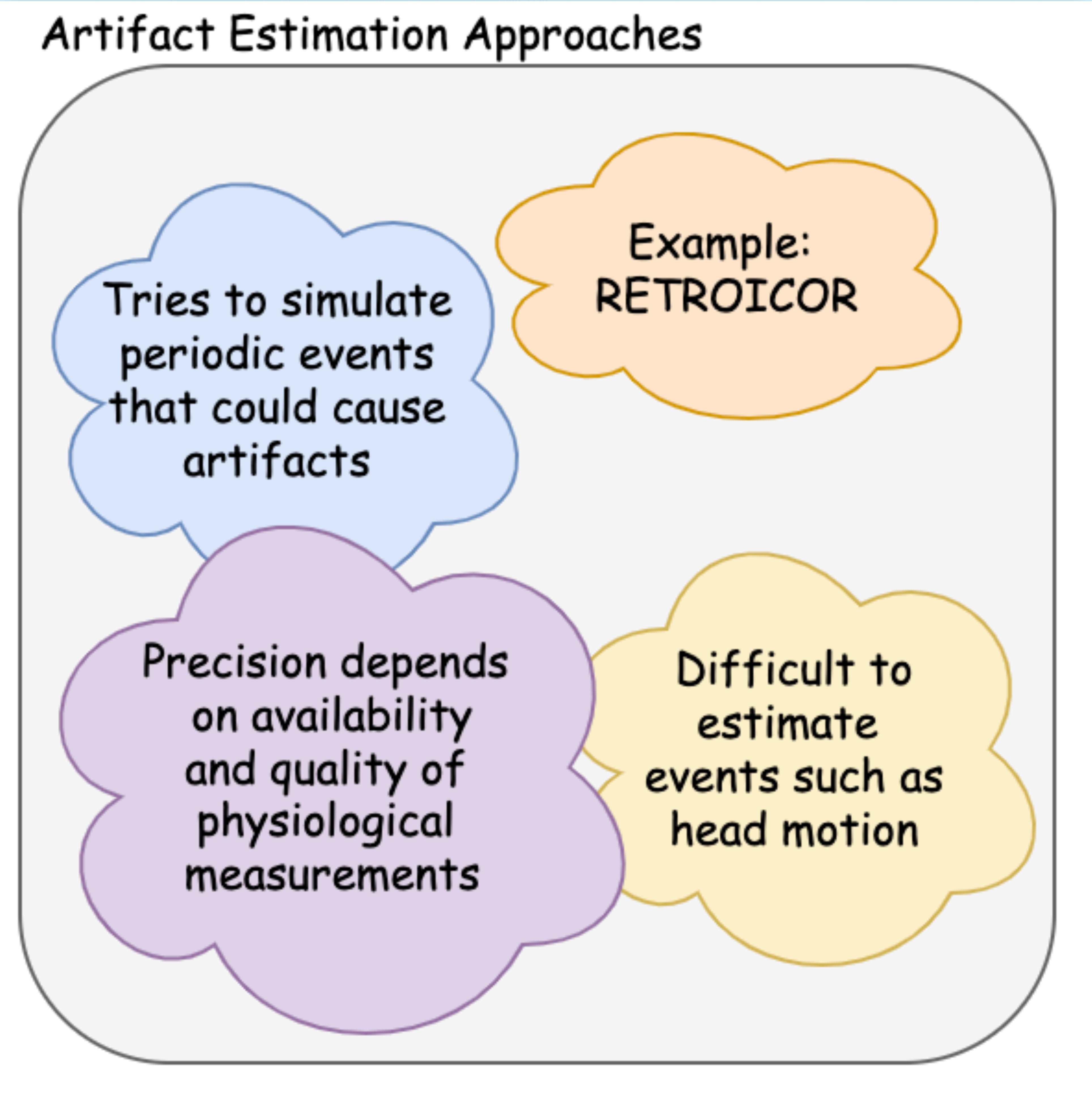
Functional Magnetic Resonance Imaging : Scott A. Huettel, Allen W. Song, Gregory McCarthy. Pg. 218

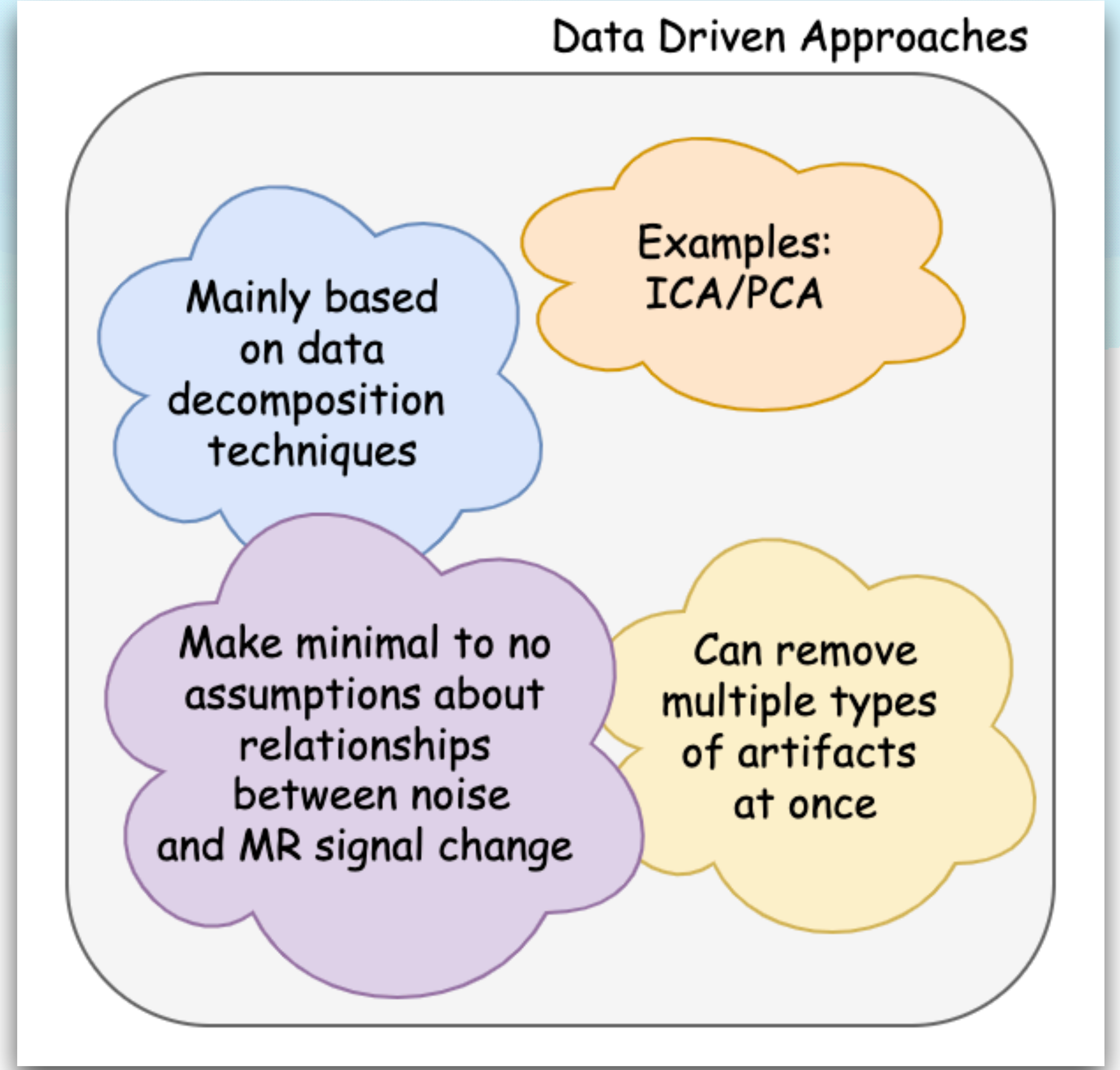
Which changes are due to the experiment?

Which changes are noise?

Denoising BOLD fMRI Signal

Artifact Estimation and Data Driven Approaches





High Level Overview of the Paper

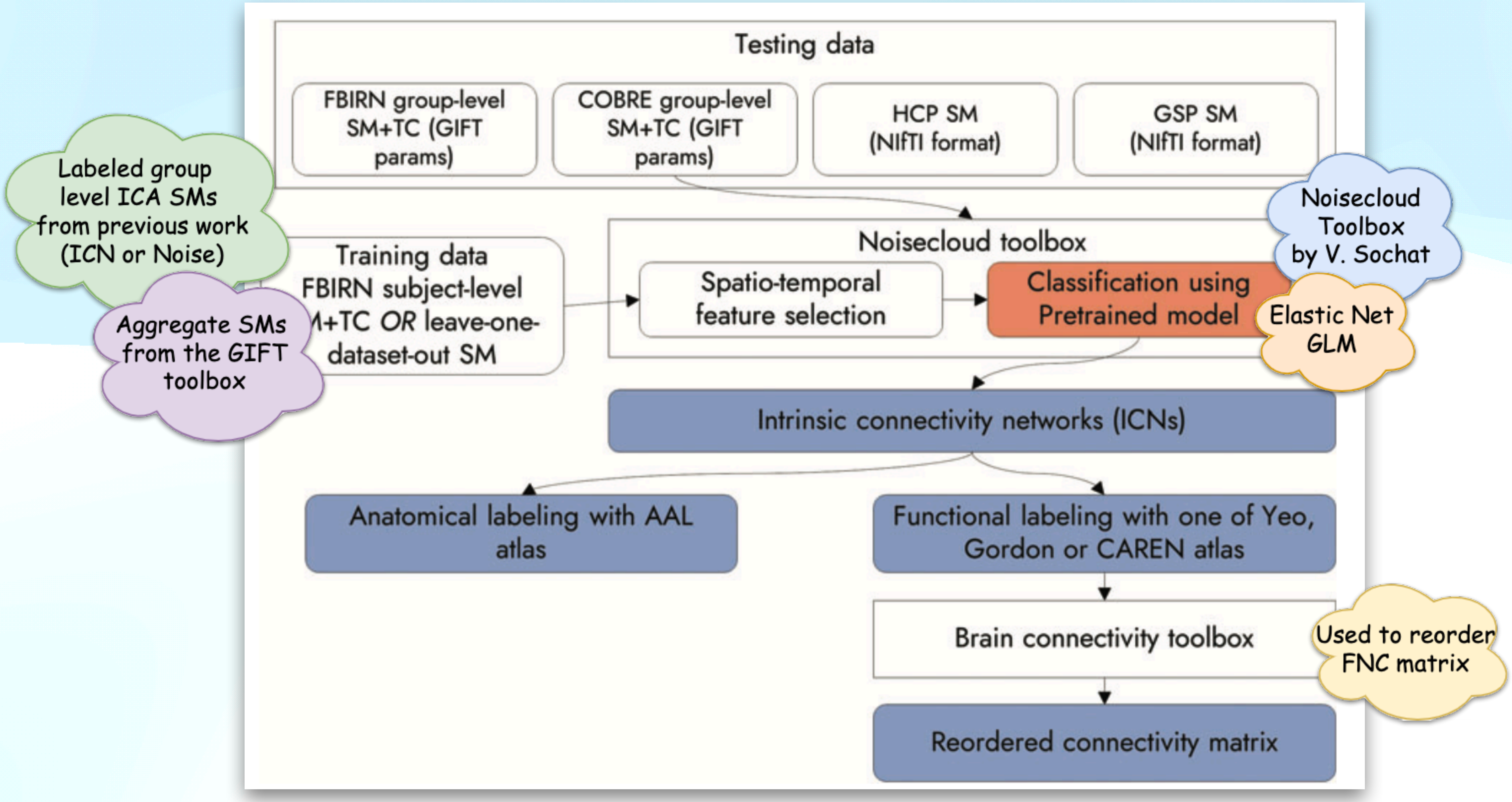
What Are the Unique Contributions

- Advent of an Autolabeler Toolbox Software Package

- This Toolbox Offers 3 Distinct Advantages:
    - Automates & unifies tasks associated with Group ICA for fMRI data
    - Provides weights of a pre-trained classification model
    - Offers the capability of being a stand alone software

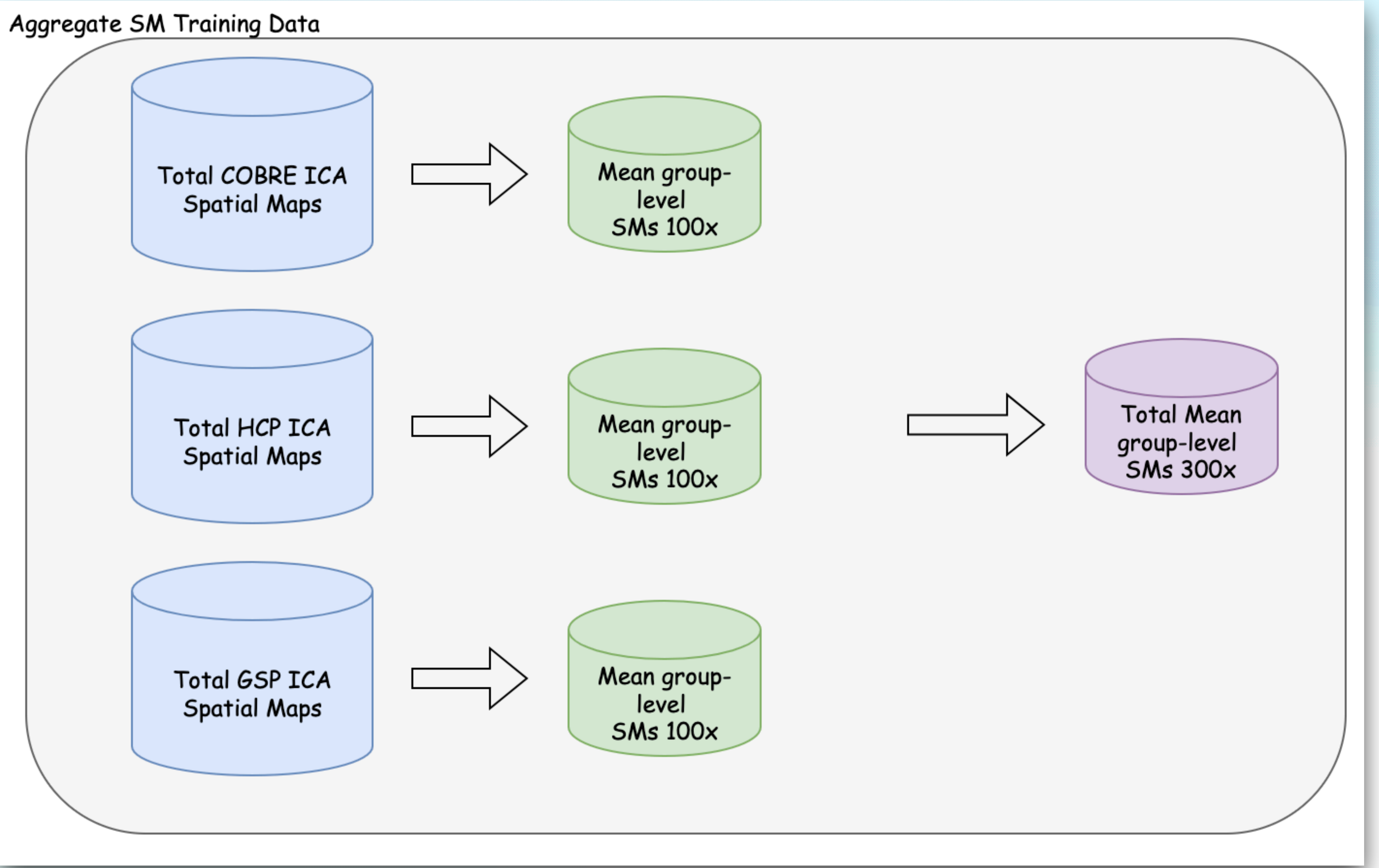
Overview of Methods

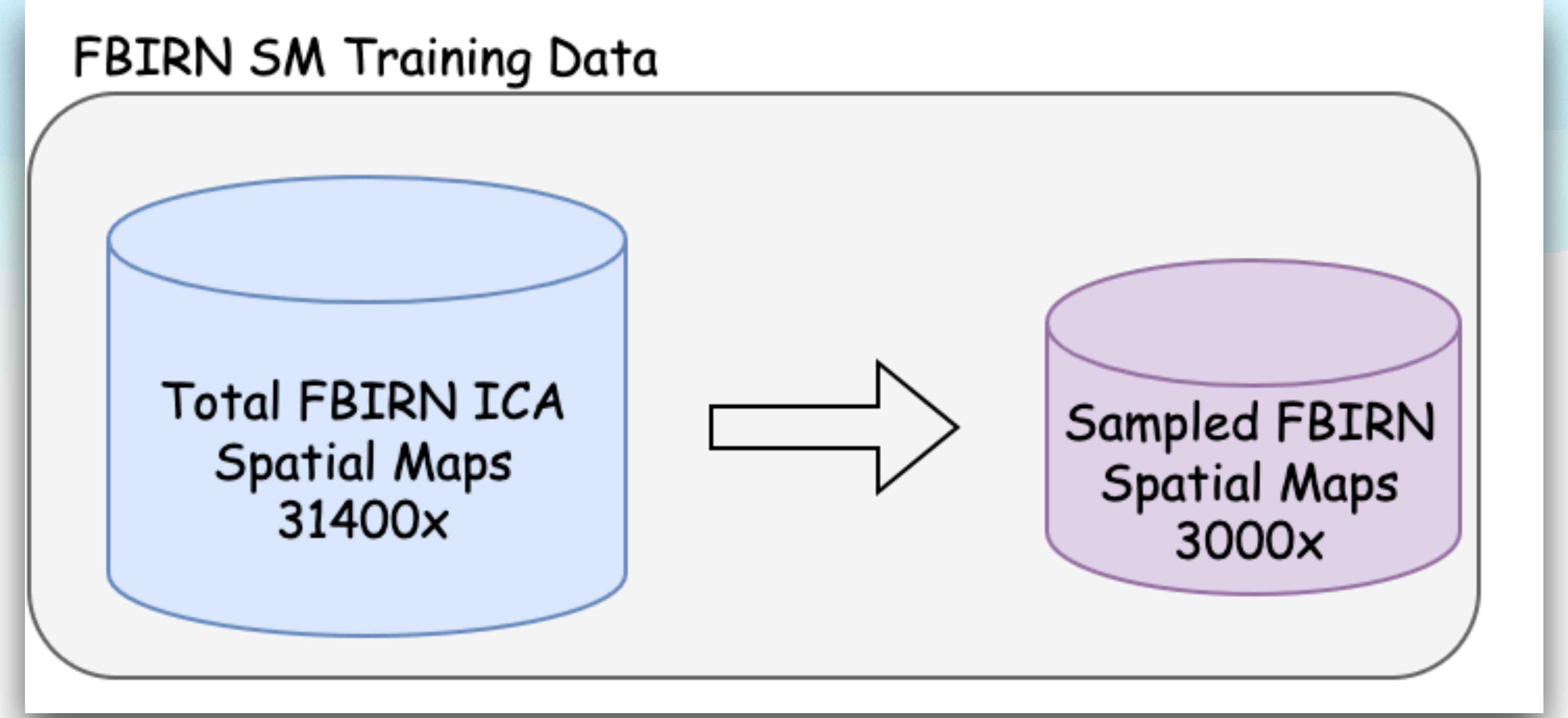
Flowchart Explaining what was Done



Training Data

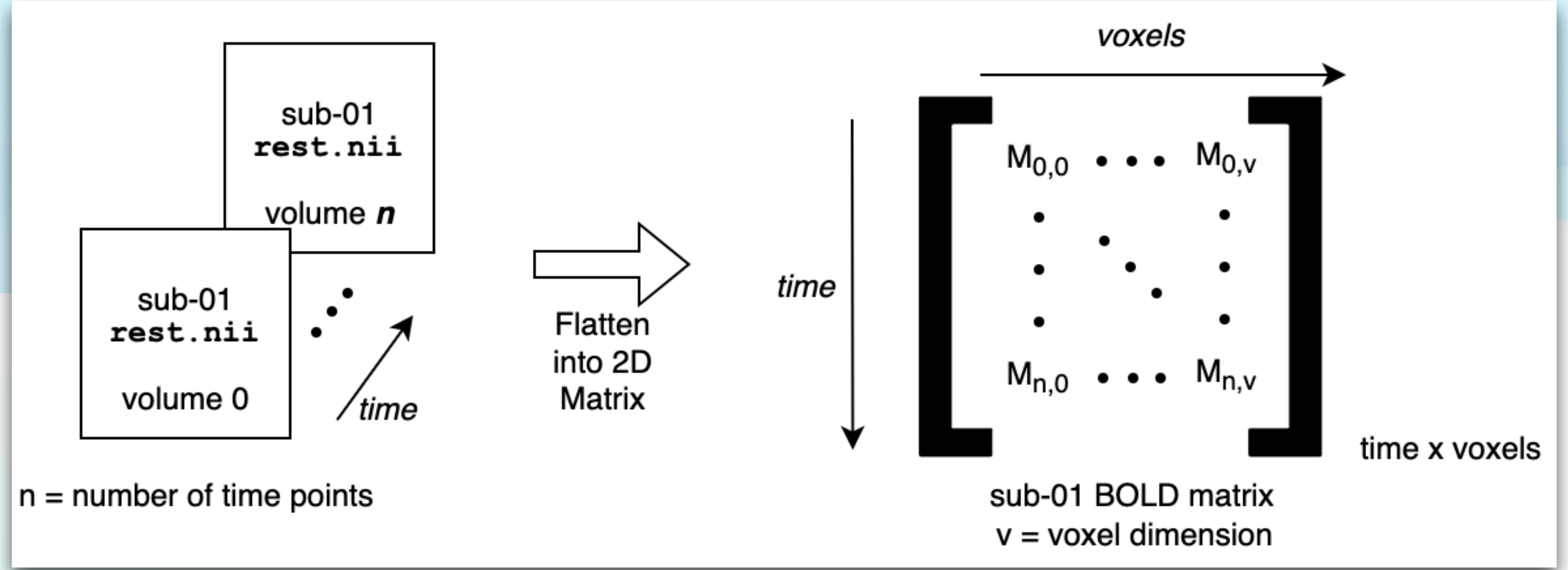
How the training data was assembled





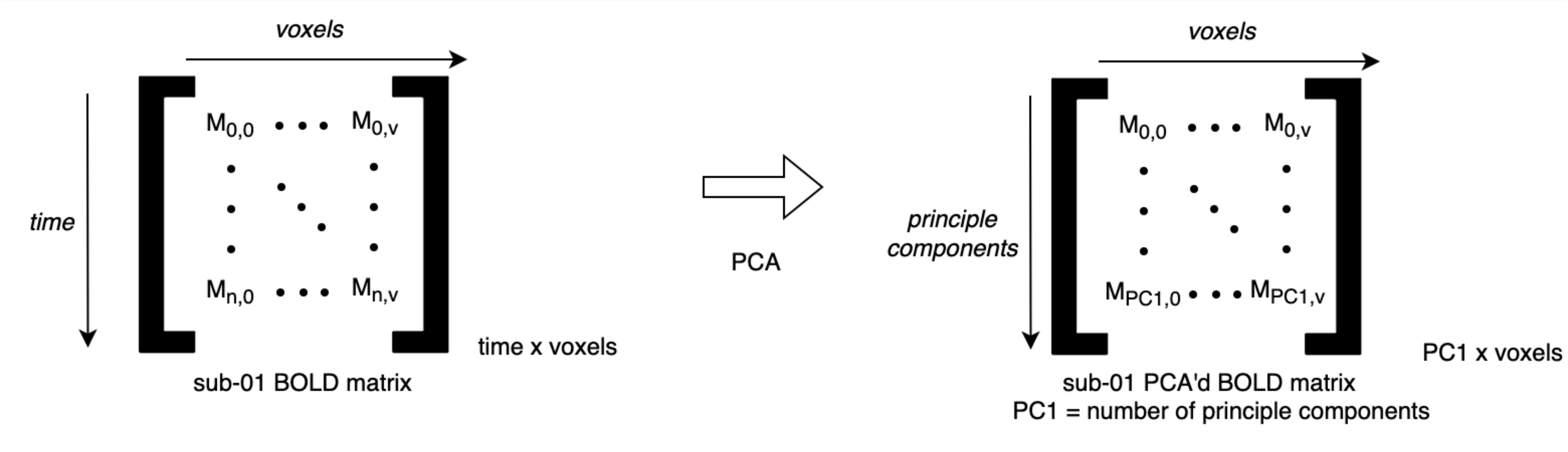
Group Independent Component Analysis ICA

Decomposing the Data into independent signals



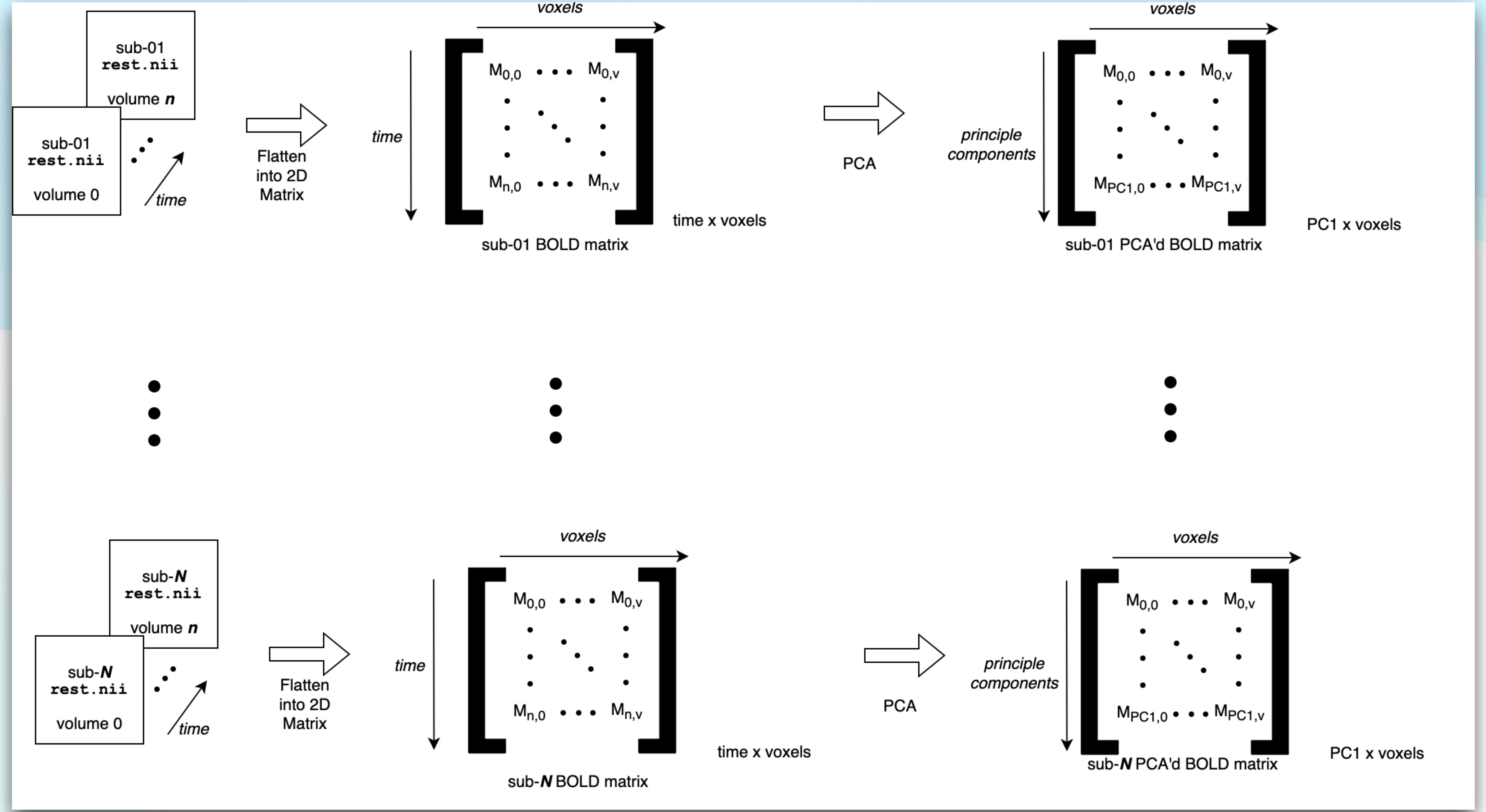
Group Independent Component Analysis ICA

Decomposing the Data into independent signals



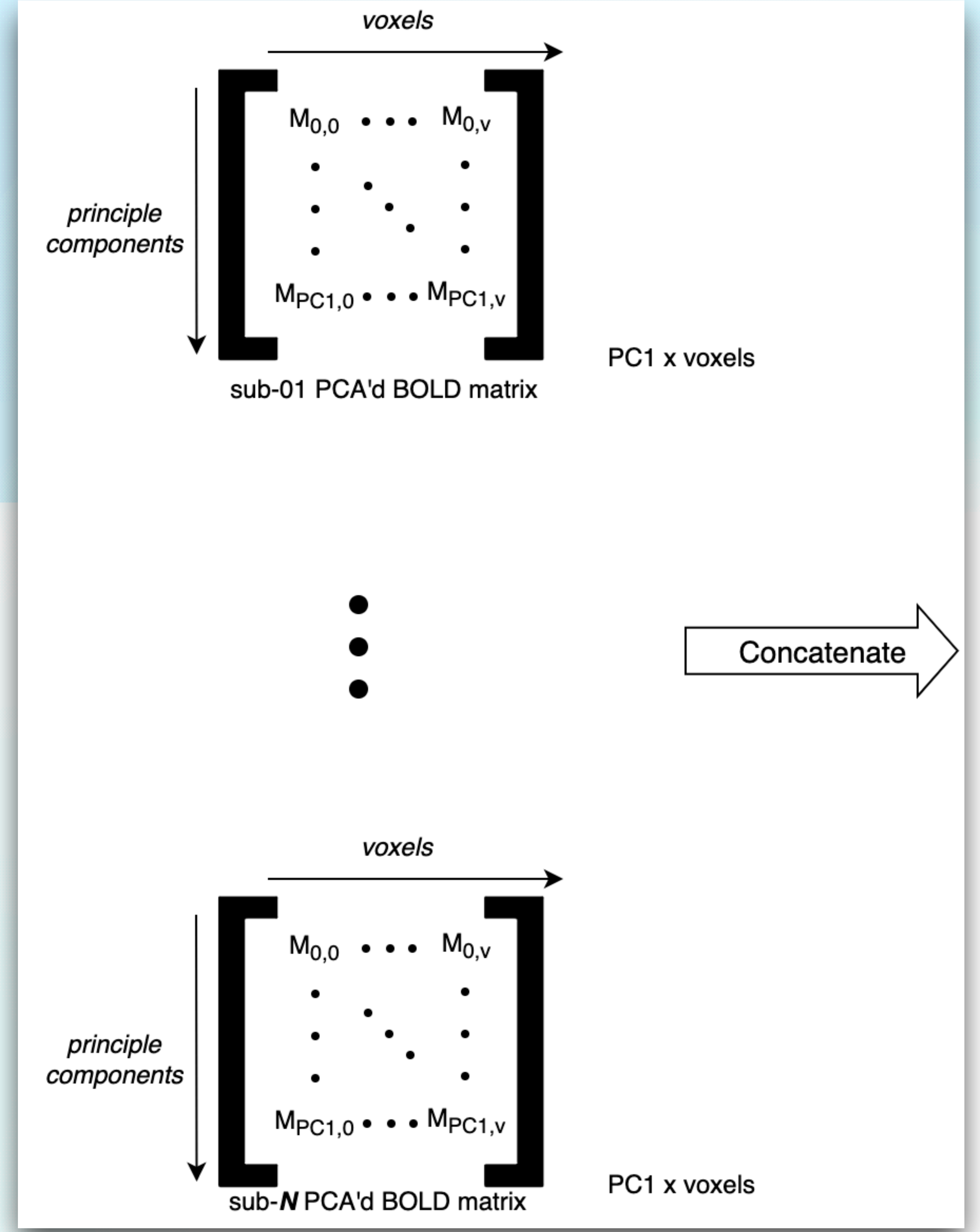
Group Independent Component Analysis ICA

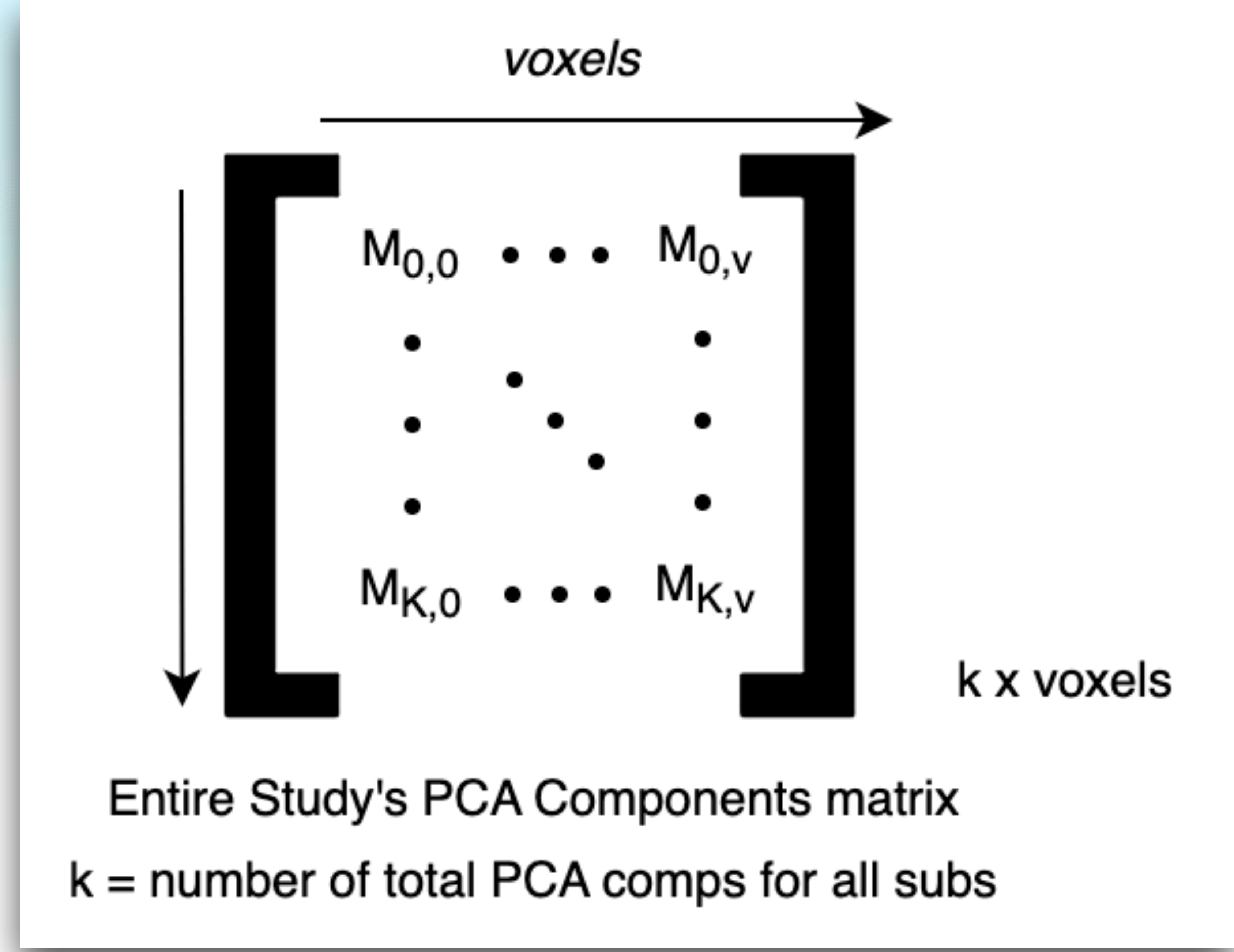
Decomposing the Data into independent signals



Group Independent Component Analysis ICA

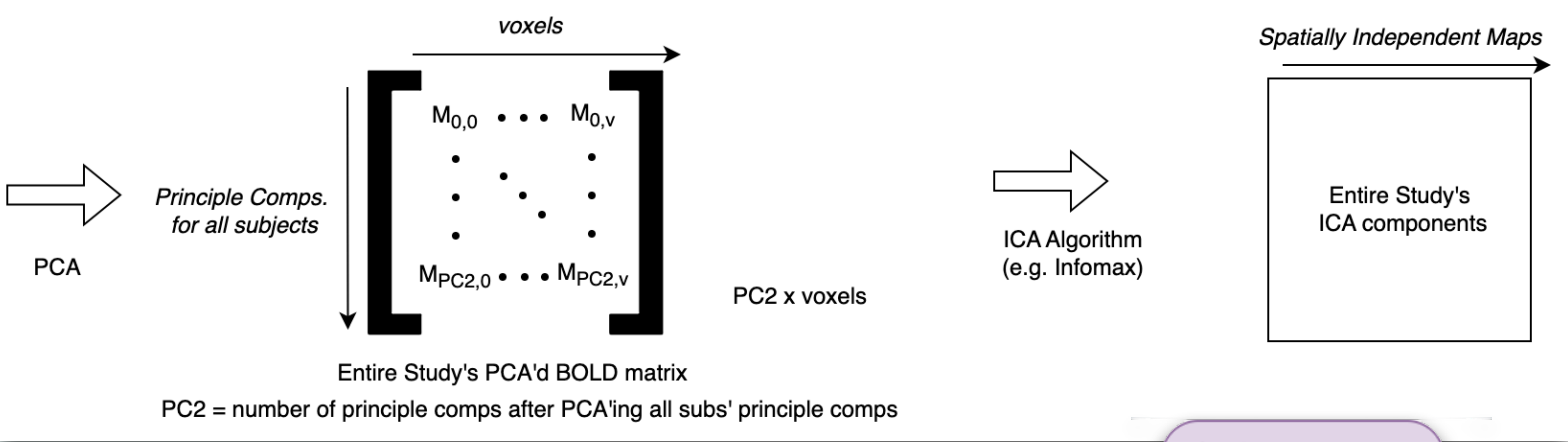
Decomposing the Data into independent signals





Group Independent Component Analysis ICA

Decomposing the Data into independent signals



Result: Set of noise and brain functional networks.

Each component includes a 3D brain map that shows the location of the network OR artifact

Visualization of Labeled FBRIN Group ICA

Used as ground truth for training Noisecloud classifier

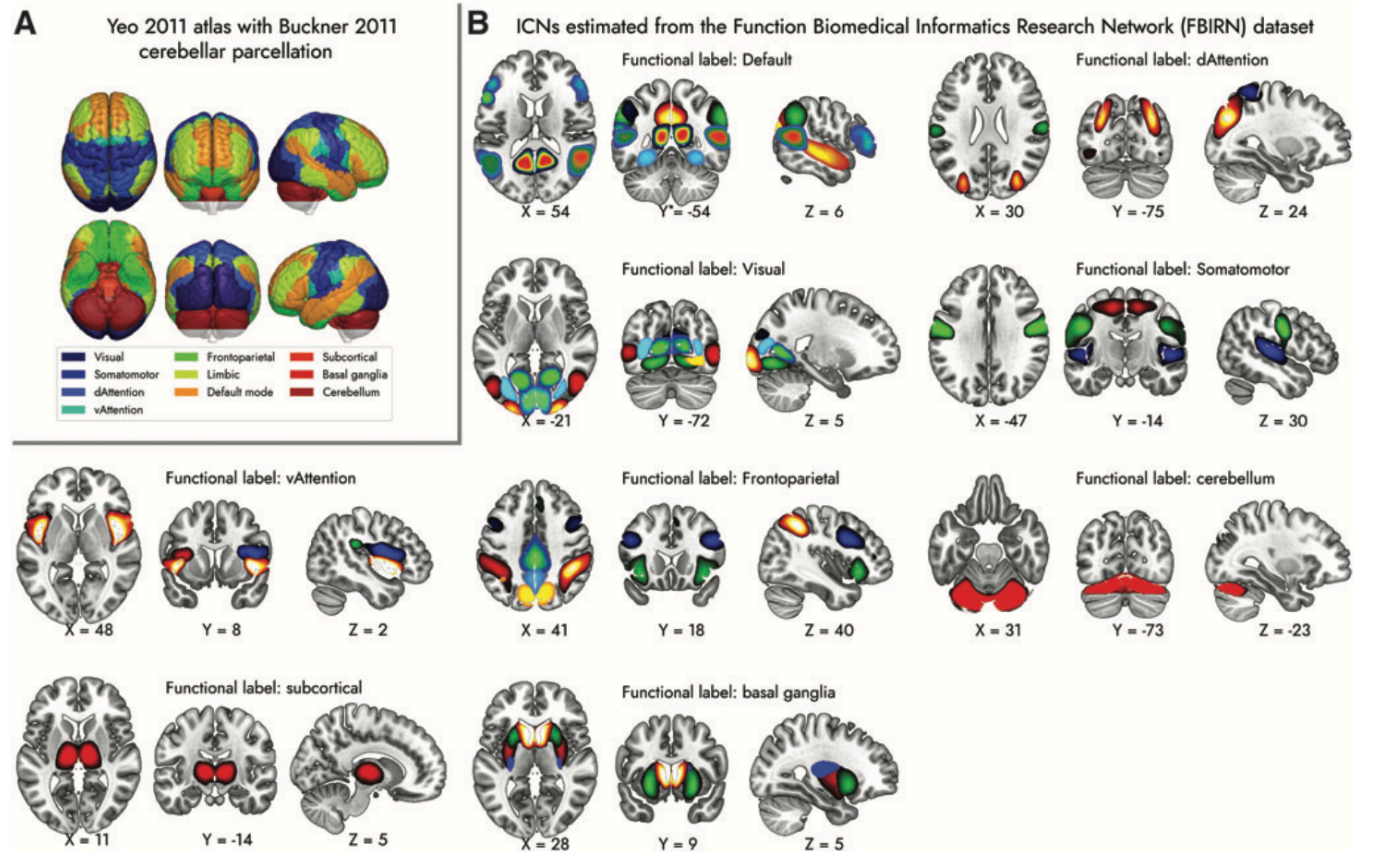


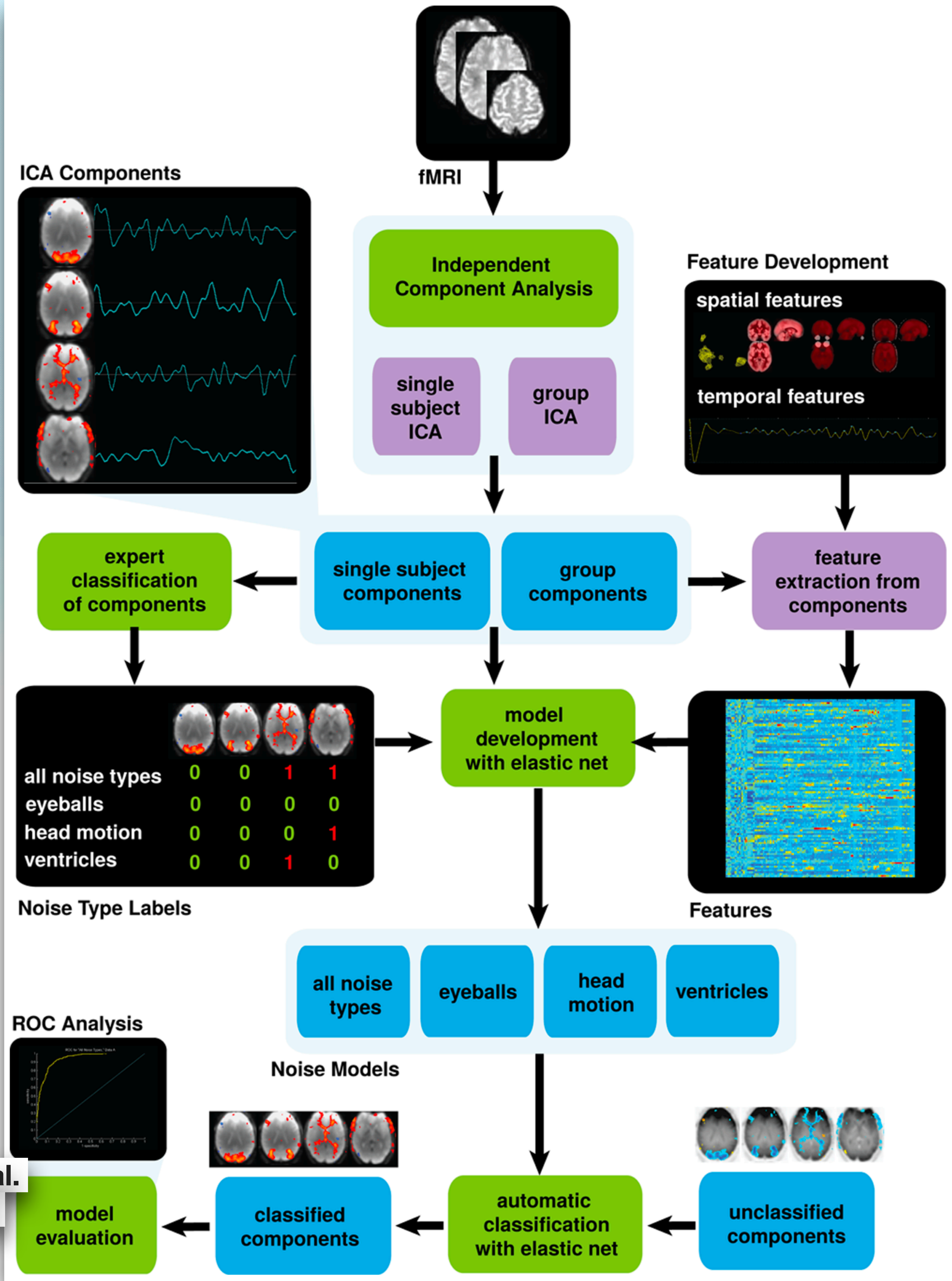
FIG. 2. **(A)** Mosaic view of the Yeo et al. (2011) atlas with Buckner et al. (2011) cerebellum parcellation of the brain (Buckner et al., 2011; Yeo et al., 2011). **(B)** ICNs estimated from the FBIRN data set, grouped into functional domains based on the Yeo et al. (2011) atlas. See Supplementary Table S2 and prior work for ICN labels and peak coordinates (Damaraju et al., 2014). FBIRN, Function Biomedical Informatics Research Network; ICN, intrinsic connectivity network. Color images are available online.

Noisecloud Toolbox

Classifies ICN from Noise

Develops features from group ICA components

Uses Elastic Net GLM to classify noise from ICN components



Sochat V, Supekar K, Bustillo J, Calhoun V, Turner JA, et al. (2014) A Robust Classifier to Distinguish Noise from fMRI Independent Components.

Elastic Net GLM in Noisecloud

Noisecloud uses an automated classifier to detect noisy fMRI components

Uses Logistic Regression with the elastic net penalty

$$h(X)=p(y=1|x')=\frac{e^{\beta 0+\beta' x'}}{1+e^{\beta 0+\beta' x'}} \quad \text{(C)}$$

2.5 Distinguishing Noise-related from Network-related Components using Sparse Logistic Regression

We developed an automated classifier to detect noisy rsfMRI components using logistic regression with the elastic net penalty, a cost function added to the optimization step of regression [54]. For N as the number of components to classify, y_i as a vector of class labels, λ and α as parameters determined by cross validation and grid search (explained later), β as the vector of weights to optimize, and x_i a feature matrix with 246 spatial and temporal features (columns) for each of N components, the elastic net optimization problem is defined as follows:

$$\min_{\beta_0 \beta}\left(\frac{1}{2N}\sum_{i=1}^{N}(y_i-\beta_0-x_i^T\beta)^2+\lambda P_\alpha(\beta)\right) \quad \text{(A)}$$

where

$$P_\alpha(\beta)=\frac{1-\alpha}{2}\|\beta\|_2^2+\alpha\|\beta\|_1\sum_{j=1}^{p}\left(\frac{1-\alpha}{2}\beta_j^2+\alpha|\beta_j|\right) \quad \text{(B)}$$

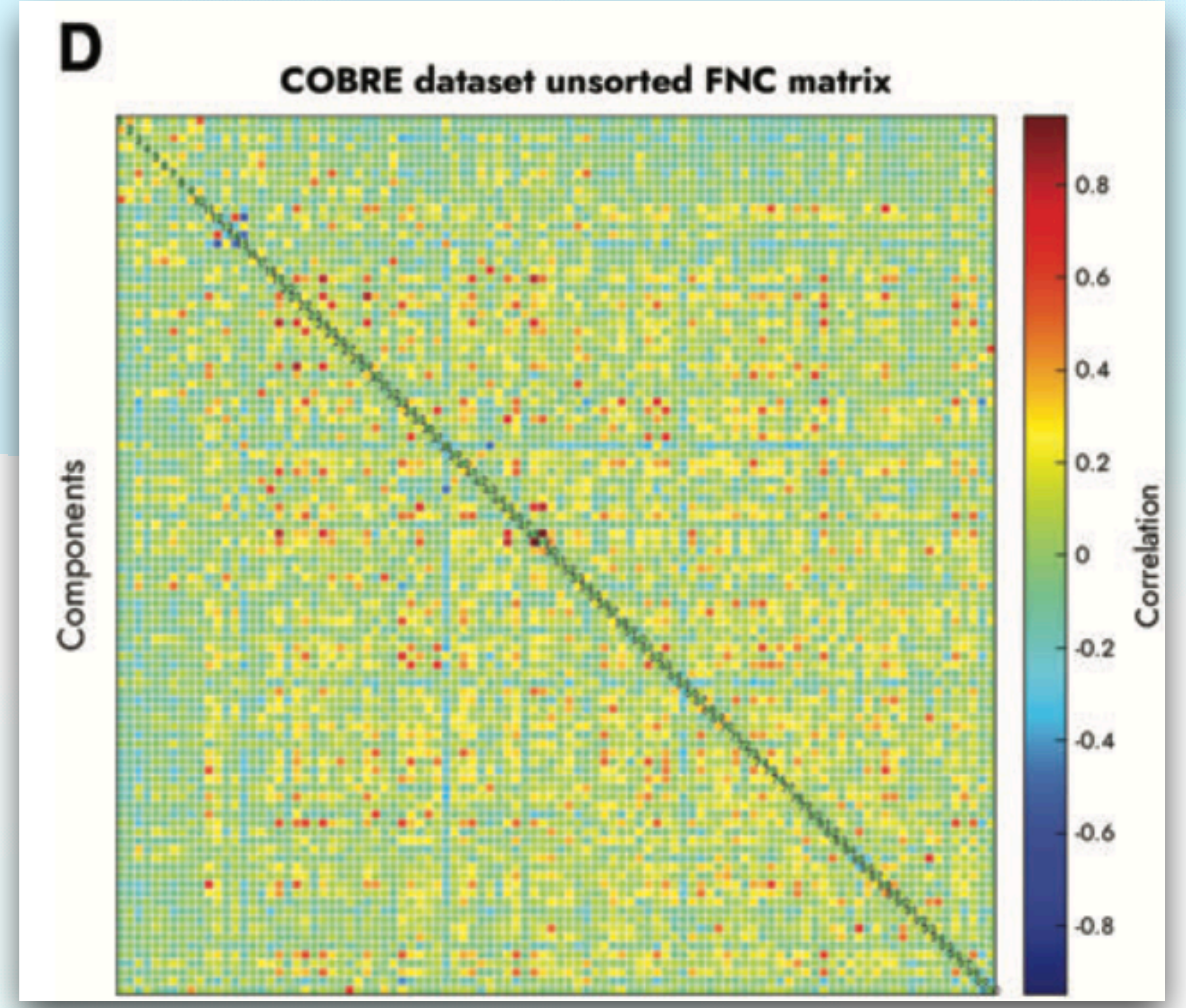
Sochat V, Supekar K, Bustillo J, Calhoun V, Turner JA, et al. (2014) A Robust Classifier to Distinguish Noise from fMRI Independent Components.

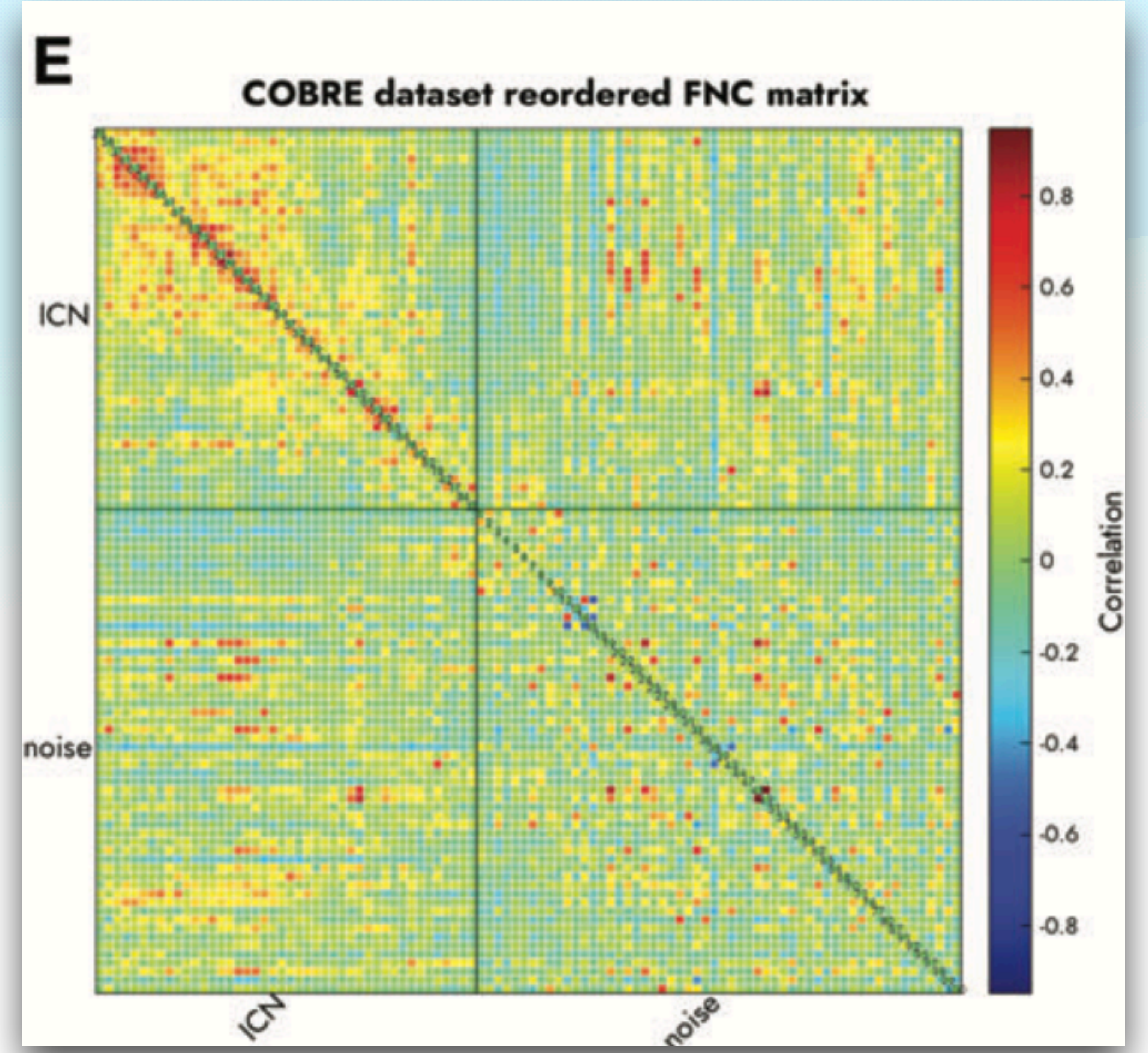
Classified ICs from FBIRN Dataset

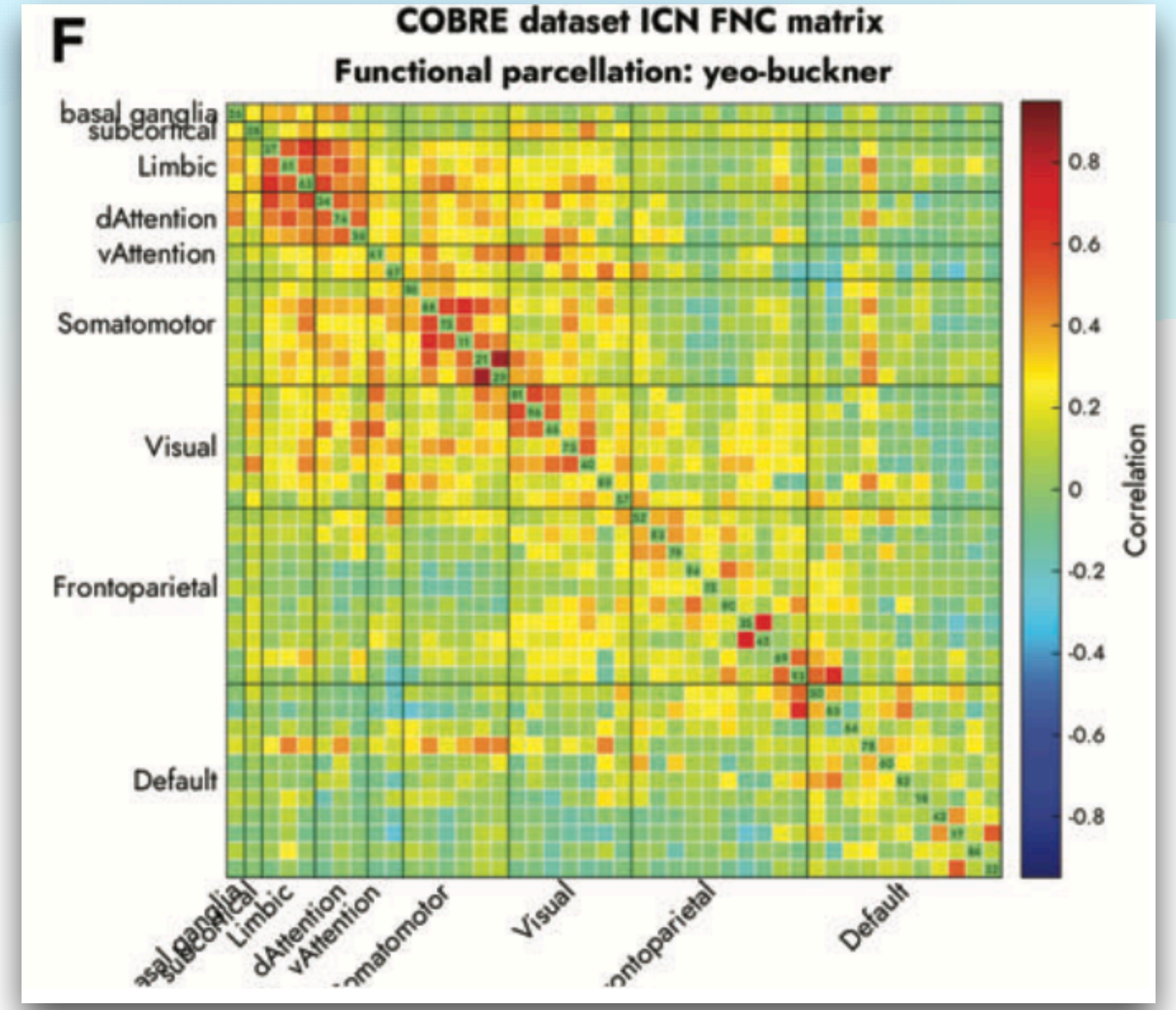
ICN Unsorted and Sorted using Brain Connectivity Toolbox (BCT)

Classified ICs from COBRE Dataset

ICN Unsorted and Sorted using Brain Connectivity Toolbox (BCT)







Visualization of Classified ICs

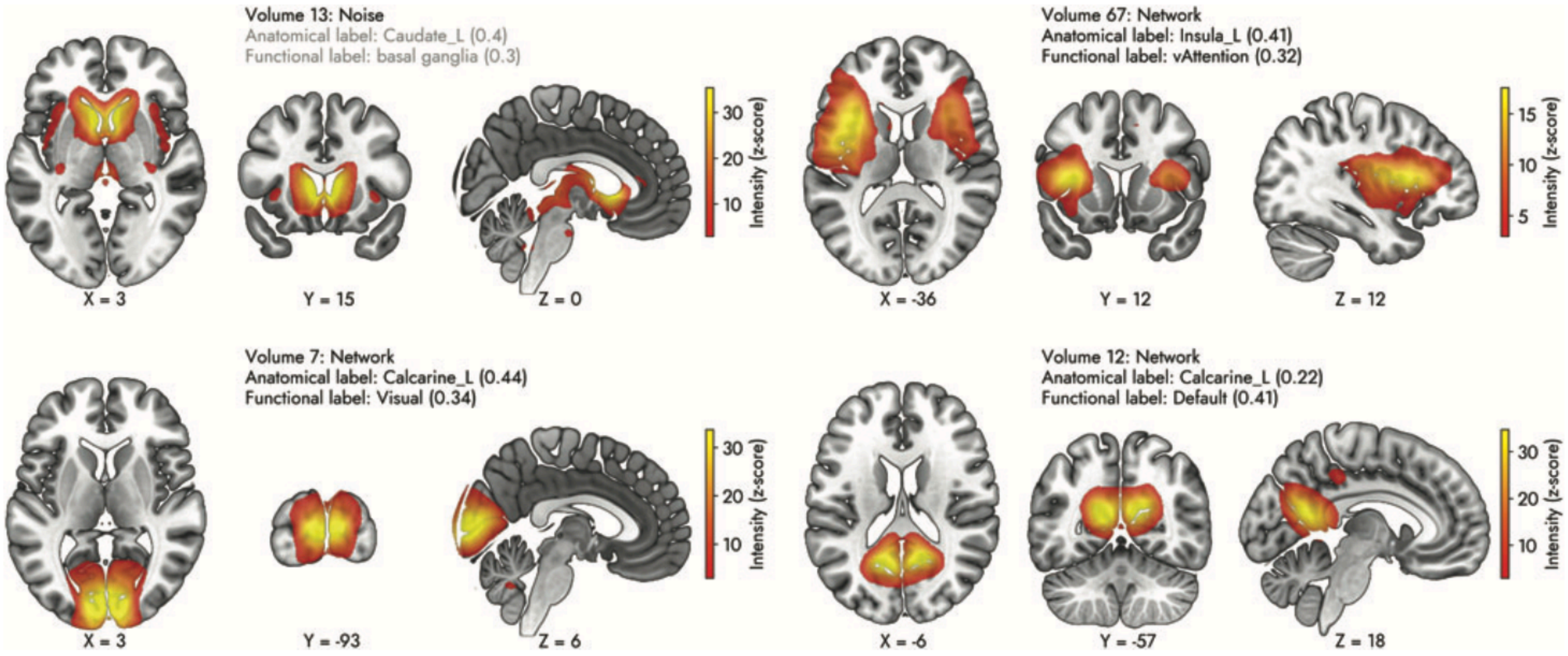


FIG. 4. Examples of predictions generated by the pretrained model of the proposed method in FBIRN data set. The labels indicate the output determined by the pretrained noisecloud model as well as the highest correlated anatomical and functional ROI labels for each IC. By default, the three highest correlated ROIs with each IC can be found in Autolabeler output. IC, independent component; ROI, region of interest. Color images are available online.

Classification Results

TABLE 1. PERFORMANCE IN DETECTING INTRINSIC CONNECTIVITY NETWORK VERSUS NOISE USING A PRETRAINED MODEL

Testing data	*Training data*	*Training N*	*Training accuracy*	*Testing accuracy*	*Precision*	*Recall*
FBIRN mean SM+TC	FBIRN subject SM+TC	3000	81.55	87[a]	80.85	90.48
COBRE mean SM+TC		3000	81.55	86	91.67	75
FBIRN aggregate SM	Leave-one-data set-out SM	300	72.67	71	63.83	71.43
COBRE aggregate SM		300	73.67	78	88.89	64
HCP aggregate SM		300	73	68	54.9	75.68
GSP aggregate SM		300	84	77	66.67	85

COBRE, Centers of Biomedical Research Excellence; FBIRN, Function Biomedical Informatics Research Network; GSP, Genomic Superstruct Project; HCP, Human Connectome Project; SM, spatial map; TC, time course.

The proposed approach developed using the Autolabeler toolbox incorporates a pretrained cross-validated elastic-net regularized GLM for separating ICNs from artifacts. Table 1 demonstrates the performance of the pretrained models applied to different data sets using two different training samples. The accuracy, precision, and recall values for recognizing ICNs are all in percentages. The model predicted ICN labels with 87% accuracy in the primary (FBIRN) data set and 86% accuracy in a validation (COBRE) data set using both the SMs and TCs as features. A 91.67% precision in the validation data set indicates that the majority of …

… the true ICNs were correctly labeled. ICN labels were predicted with between 68% and 77% accuracy in the validation data sets using only the SMs as features. The asterisk (*) indicates the case when the model was trained with FBIRN subject-level IC features and the testing data included FBIRN mean ICA components, indicating a biased result.