

Rejoinder to Discussions of “Distributional Independent Component Analysis for Diverse Neuroimaging Modalities”

Ben Wu¹, Subhadip Pal², Jian Kang^{3,*} and Ying Guo^{4,**}

¹Center for Applied Statistics, School of Statistics, Renmin University of China, Beijing, 100872, CN

²Department of Biostatistics and Bioinformatics, University of Louisville, Louisville, KY, 40292, USA

³Department of Biostatistics, University of Michigan, Ann Arbor, MI, 48109, USA

⁴Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, 30322, USA

**email*: jiankang@umich.edu

***email*: yguo2@emory.edu

SUMMARY: We thank the editors for organizing the discussions and the discussants for insightful comments. Our rejoinder provides results and comments to address the questions raised in the discussions. Specifically, we present results showing DICA largely demonstrates better or comparable stability as compared with standard ICA. We also validate the DICA in real fMRI application by showing DICA generally shows higher reliability in reproducibly recovering major brain functional networks as compared with the standard ICA. We provide details on the computational complexity of the method. The computational cost of DICA is very reasonable with the analysis of the fMRI and DTI data easily implementable on a PC or laptop. Finally, we include discussions on several directions for extending the DICA framework in the future.

This paper has been submitted for consideration for publication in *Biometrics*

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/biom.13588

This article is protected by copyright. All rights reserved.

Author Manuscript

1. Introduction

We thank the editor and associate editor for inviting discussions of our article, and the discussants ([Keeratimahat and Nichols, 2021](#); [Mejia, 2021](#); [Moerkerke and Seurinck, 2021](#); [Shappell and Simpson, 2021](#)) for insightful conversations about our work. We summarize the major comments and questions raised by the discussants as follows: (1) the stability and robustness of the proposed DICA, (2) the computation time and memory requirements of DICA, (3) method validation and further comparison with standard ICA in real fMRI data, and (4) potential directions for extending the DICA framework. In the rejoinder, we provide results and discussions to address these comments and questions.

2. Stability of the DICA

The discussants have raised the question about the stability of results from DICA, considering the EM algorithm used for mixture modeling at stage one of DICA is known to be sensitive to the initial values ([Mejia, 2021](#); [Shappell and Simpson, 2021](#)); and the initial random initialization of the mixing matrix of the ICA at stage two of DICA also leads to variation in ICA results ([Keeratimahat and Nichols, 2021](#)). To address the question, we conduct additional analysis to investigate the stability of DICA.

As [Keeratimahat and Nichols \(2021\)](#), we conduct stability analysis by running a method multiple times using the example imaging data sets and evaluate the overlap in the extracted spatial independent components (ICs) maps across runs with the Dice coefficient. We consider four variations of methods: (1) DICA-FI1: DICA with fixed initialization for EM at stage one, and random initialization of ICA at stage two; (2) DICA-FI2: DICA with random initialization for EM at stage one and fixed initialization for ICA at stage two; (3) DICA-RI: DICA with random initialization for both EM and ICA; and (4) the standard ICA with random initialization. As per the discussants’ request ([Mejia, 2021](#)), we provide additional

details for the implementation of DICA. In the DICA R package, we use clustering methods such as k-means to generate initial values for the EM algorithm at stage one. For the ICA analysis at stage two, the DICA R package uses the Infomax algorithm implemented via the function “icaimax” from R package “ica” where random initiations can be specified for the mixing matrix. Since the initial release of our DICA R package which was used in Keeratimahat and Nichols (2021), we have implemented an update for the package which fixes a numerical underflow issue in the original code when dealing with extremely small posterior weights. The updated R package is available on Github at <https://github.com/benwu233/DICA>.

To obtain a more reliable assessment on the stability, we conduct the numerical experiments with 50 replications. In contrast, Keeratimahat and Nichols (2021) draw their conclusions based on only 5 replications. We match the ICs across different replications using the same greedy matching algorithm adopted by Keeratimahat and Nichols (2021). The overlap between matched ICs is measured with the multi-class Dice coefficient. As shown in the Figure 1, for fMRI data, all the three variations of DICA have clearly better stability than the standard ICA, which is consistent with the findings in Keeratimahat and Nichols (2021). For DTI data, DICA-FI1 provides more stable results than the standard ICA while DICA-FI2 and DICA-RI show comparable or slightly lower stability as compared with standard ICA (Figure 1). It is worth noting the Dice coefficients of the DICA methods in DTI data have demonstrated an obvious improvement over those reported in Keeratimahat and Nichols (2021) after we fix the numerical underflow issue in the original DICA R package. Among the three DICA variants, DICA-FI1 and DICA-FI2 which have fixed initiation at either stage one or stage two have more stable results than DICA-RI which has random initiation for both the mixture modeling and ICA. In their discussions, Keeratimahat and Nichols (2021) expressed worry that random initialization in mixture modeling and ICA may lead to

greater instability in DICA relative to standard ICA. Our results show that DICA-RI actually demonstrates better stability than standard ICA in fMRI data and has only slightly lower Dice coefficients in DTI data. These findings alleviate the concern raised in [Keeratimahat and Nichols \(2021\)](#). Furthermore, we want to point out that the standard ICA of DTI data is using ICA to decompose the six values of the tensors ([Keeratimahat and Nichols, 2021](#)), which means the standard ICA of DTI can only extract up to six independent components. This restriction would limit the ability of the standard ICA in discovering fine-scale latent components in DTI data. In comparison, the proposed DICA potentially provides a more powerful and flexible tool for DTI decomposition.

Our results show that different initial values in the EM mixture modeling contribute to the variations in DICA results. Therefore, we recommend using clustering methods to generate informative initial values for the EM algorithm to improve its convergence. Additionally, we agree with the discussants ([Mejia, 2021](#); [Shappell and Simpson, 2021](#)) to consider multiple sets of initial values to improve the performance of the EM. Specifically, one may consider multiple initialization strategies such as the short EM, the multiple-repeated k-means, and the REBIMX algorithm ([Panić et al., 2020](#)) in practice.

3. Computational complexity

Another question from the discussants ([Mejia, 2021](#); [Shappell and Simpson, 2021](#)) is related to the computational cost of DICA when applied to the imaging data and when the number of ICs is large. In the Supporting Information of our paper, we have included details and discussions on DICA computation. Specifically, the user CPU time was around 150 seconds for fMRI and 585 seconds for DTI on a MacBook Pro laptop with a 3.1 GHz Dual-Core Intel Core i5 processor and 8 GB memory. The most computationally expensive and memory-demanding component in DICA is the mixture modeling at stage one since it involves the original imaging data. The computational complexity in each iteration of the EM algorithm

is $O(JKT)$, where T is the dimension of the imaging measurements at each voxel, J is the number of voxels and K is the number of components in the mixture distribution. That is, the computation time increases linearly with the number of mixture components. In terms of memory requirements, since we utilize an iterative algorithm for the mixture modeling, DICA does not require a large amount of memory to implement. In addition to the original imaging data, we only need to store the posterior weights and the distribution parameters from the mixture modeling, which only need a small memory space. DICA of the fMRI data and DTI data in the paper can be easily implemented on a personal laptop. [Mejia \(2021\)](#) has a question about DICA's scalability to a large number of ICs. The mixture modeling at stage one of DICA significantly reduces the dimension of the inputs to the ICA which essentially decomposes the posterior weights from stage one. Therefore, the computational cost for DICA does not increase dramatically for large number of ICs.

4. Method validation in imaging applications

[Moerkerke and Seurinck \(2021\)](#) and [Mejia \(2021\)](#) raise very good points on method validation of DICA in neuroimaging applications. [Moerkerke and Seurinck \(2021\)](#) suggests more evaluation criteria could be considered depending on the ultimate goal of a study. [Mejia \(2021\)](#) comments that DICA should be validated in real imaging data in comparison with the standard ICA which is widely applied to fMRI. In neuroimaging studies, ICA is mainly used to decompose observed images to reveal underlying brain networks. Therefore, an important evaluation criterion for validating DICA in neuroimaging applications is whether the method can generate reliable and reproducible findings regarding brain networks. To this end, we evaluate the reliability of the DICA in recovering functional networks using rs-fMRI data obtained from the Philadelphia Neurodevelopmental Cohort (PNC) study. Based on the discussants' comments, we perform a comparison with the standard ICA. Specifically, we apply DICA and the standard ICA to the rs-fMRI data 50 times and evaluate

the reproducibility of the two methods in recovering brain functional networks. We note that Keeratimahat and Nichols (2021) perform 5 runs of DICA and standard ICA on the rs-fMRI data and presented the estimated visual network and auditory network from the 5 runs. For more reliable and comprehensive assessment, we present reproducibility results from 50 runs for six major resting-state functional networks. In Figure 2, we show the proportion of runs that regions in the six resting-state functional networks are reproducibly recovered by the methods. DICA generally shows a noticeably higher reliability in recovering key regions in the brain networks as compared with the standard ICA, especially for the occipital pole visual, lateral visual, sensorimotor, auditory, and executive control networks. In particular, our analysis shows DICA demonstrates an improved reliability in consistently recovering the auditory network, which is different from the conclusion in Keeratimahat and Nichols (2021). This is due to the difference in the criterion and number of runs used in assessing reliability in the two papers. Our finding is obtained by evaluating the reproducibility rate which quantifies the reliability of results across 50 runs, while Keeratimahat and Nichols (2021) draw their conclusions based on visual inspections of results from 5 runs. For the default mode network, the performance of DICA and the standard ICA is comparable with each method generating slightly better results in different regions of the networks.

5. Extension of the DICA framework

As the discussants mentioned, there are several directions we can extend the DICA framework. In the Discussion section of our paper, we have pointed out various strategies that can be taken to generate DICA to multi-subject imaging data. For example, we can consider concatenation of the imaging data and the posterior weights across subjects in the two stages of DICA. Alternatively, we may follow the hierarchical ICA framework (Guo and Tang, 2013; Shi and Guo, 2016; Wang and Guo, 2019) to develop a multilevel modeling extension for DICA to first perform the individual-level ICA decomposition on the first

level of the group DICA and then model the individual-level source signals in terms of population sources and individual effects on the second level of the group DICA. Another major extension of DICA is to adapt the method for other imaging modalities such as EEG. The main task in extending to other imaging modalities lies in identifying appropriate mixture distributions to model the imaging data at stage one of DICA. The distribution model should be suitable for the data characteristics (e.g. dimension, scale) of the specific imaging modality and can effectively capture the variability in the imaging. Some good-of-fit methods could potentially be applied to evaluate the fit of various mixture models and choose a desirable distribution. Finally, to fuse information across imaging modalities, it is of interest to extend the DICA to jointly analyzing multiple imaging modalities. Since DICA represents a unified framework to decompose different imaging modalities, it provides a great platform for multimodality analysis. We could apply modality-specific mixture distributions at stage one of DICA for dimension reduction. After obtaining the modality-specific posterior weights from the mixture modeling, we can then develop a joint DICA method to simultaneously decompose the weights across the imaging modalities. There is a rich literature on joint ICA decomposition across imaging modalities, which provides a solid foundation for developing a joint DICA method.

6. Other questions and comments

The discussants have some questions regarding the tuning parameters in DICA, such as the effects of the number of mixtures K at stage one. In Section 3 of the Supporting Information of our DICA paper, we have included sensitive analyses in both the simulation studies and real data applications. The analyses show results from DICA remain fairly stable for various choices of K within a reasonable range. In [Keeratimahat and Nichols \(2021\)](#), the discussants suggest dispensing with the tuning parameter in the PCA step by avoiding the PCA dimension reduction for fMRI and directly estimating the mixture distribution

for the full-size fMRI data. We respectfully disagree with this suggestion. Based on our experience, direct mixture modeling of the original fMRI time series increases the challenges and instability in the mixture modeling due to the large dimension of the full-size fMRI and the high noise level in the original data. The issue is especially serious for recent fMRI studies which can acquire around a thousand of volumes. PCA dimension reduction on the fMRI time series has proven to be an effective dimension reduction and denoising technique prior to standard ICA of fMRI. PCA has also shown to improve the stability and robustness of the mixture modeling for DICA. Hence, its inclusion is beneficial to improving the performance of the method.

ACKNOWLEDGEMENTS

Research reported in this publication was supported by the National Institute of Health under Award Number R01MH105561, R01MH118771, R01DA048993, R01GM124061 and R01CA249096. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DATA AVAILABILITY STATEMENT

The example imaging data is publicly available to download with the DICA R package on Github at <https://github.com/benwu233/DICA>.

REFERENCES

- Guo, Y. and Tang, L. (2013). A hierarchical model for probabilistic independent component analysis of multi-subject fmri studies. *Biometrics* **69**, 970–981.
- Keeratimahat, K. and Nichols, T. E. (2021). Discussion on “distributional independent component analysis for diverse neuroimaging modalities”. *Biometrics* .
- Mejia, A. F. (2021). Discussion on “distributional independent component analysis for diverse neuroimaging modalities”. *Biometrics* .

- Moerkerke, B. and Seurinck, R. (2021). Discussion on “distributional independent component analysis for diverse neuroimaging modalities”. Biometrics .
- Panić, B., Klemenc, J., and Nagode, M. (2020). Improved initialization of the em algorithm for mixture model parameter estimation. Mathematics **8**, 373.
- Shappell, H. and Simpson, S. L. (2021). Discussion on “distributional independent component analysis for diverse neuroimaging modalities”. Biometrics .
- Shi, R. and Guo, Y. (2016). Investigating differences in brain functional networks using hierarchical covariate-adjusted independent component analysis. The Annals of Applied Statistics **10**, 1930–1957.
- Wang, Y. and Guo, Y. (2019). A hierarchical independent component analysis model for longitudinal neuroimaging studies. NeuroImage **189**, 380–400.

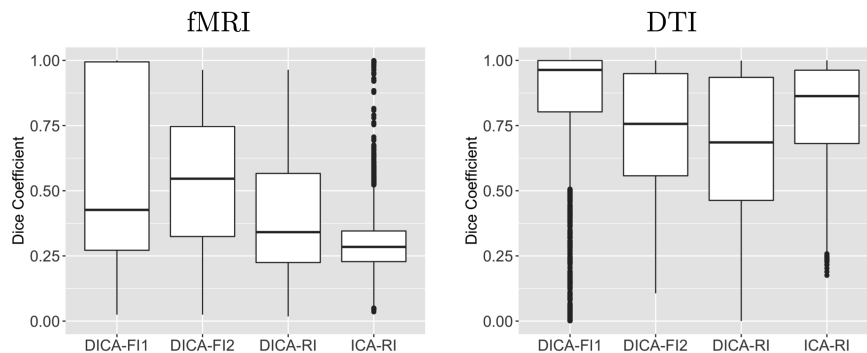


Figure 1 Stability of four variations of methods across 50 replicates for the fMRI and DTI example datasets. DICA-FI1: DICA with fixed and random initialization at stage one and stage two, respectively; DICA-FI2: DICA with random and fixed initialization at stage one and stage two, respectively; DICA-RI: DICA with random initialization at both stages; ICA-RI: standard ICA with random initialization.

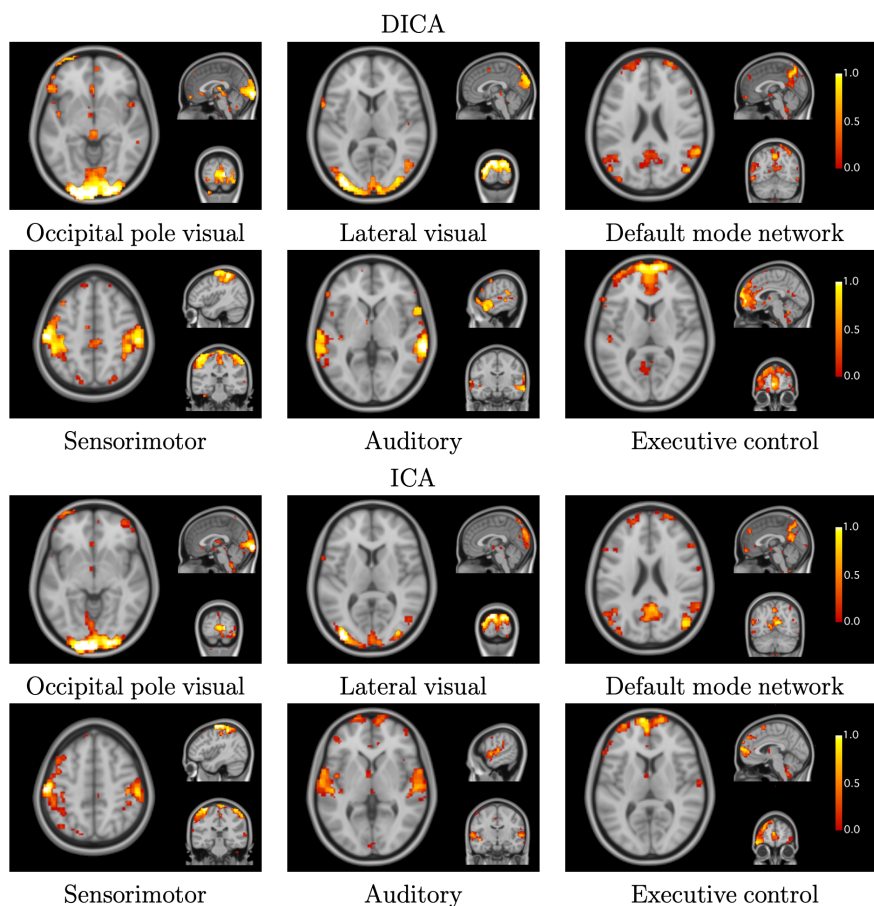


Figure 2 Reliability of DICA and standard ICA in recovering brain functional networks from a subject’s resting state fMRI data. The presented values are the reproducibility rate across 50 runs.

Author Manuscript