DISCUSSION

Discussion on “distributional independent component analysis for diverse neuroimaging modalities” by Ben Wu, Subhadip Pal, Jian Kang, and Ying Guo

Kan Keeratimahat1 | Thomas E. Nichols2

1 Department of Computer Science, Parks Road, University of Oxford, Oxford, UK
2 Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Correspondence
Thomas E. Nichols, Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, OX1 2JD, UK.
Email: thomas.nichols@bdi.ox.ac.uk

Funding information
UK Research and Innovation, Grant/Award Number: EP/S02428X/1; National Institutes of Health, Grant/Award Number: R01DA048993

Abstract
Wu et al. have made an important contribution to the methodology for data-driven analysis of MRI data. However, we wish to challenge the authors on new potential applications of their approach beyond diffusion tensor imaging data, and to think carefully about the impact of random initialization implicit in their method. We illustrate the variability found from re-analyzing the supplied demonstration data multiple times, finding that the discovered independent components have a wide range of reliability, from nearly perfect overlap to no overlap at all.

KEYWORDS
DTI, fMRI, independent component analysis, multimodality neuroimaging

1 | INTRODUCTION

We congratulate Wu et al. (2021) on their work that advances probabilistic approaches for Independent Components Analysis (ICA). ICA has long played an essential role in the analysis of resting Functional Magnetic Resonance Imaging (fMRI) data (Jung et al., 2001) but it is now being used with other types of MRI data (Xu et al., 2009). They propose a Distributional ICA (DICA) that uses a generic mixture model to allow ICA to be applied on structured data, like that from Diffusion Tensor Imaging (DTI). We have a number of comments on the work and ideas for future directions.

The Wishart mixture for the DTI data is well-motivated, as the diffusion tensors are exactly the covariance of a Gaussian diffusion model. However, we were hoping to see some other examples where this approach could be deployed. Have they considered other types of highly structured imaging data that would be suitable for this method? For example, the current trend in diffusion MRI are macroscopic models that account for the heterogeneous diffusion in a voxel, for example, Neurite Orientation Dispersion and Density Imaging (NODDI) (Zhang et al., 2012). NODDI produces multiple measures with complex non-Gaussian interdependence, but not in a manner consistent with a standard parametric model. Do they have ideas for mixtures models when physics does not dictate the stochastic structure?

We wondered if another potential application is multisubject binary brain data, whether representing white matter hyperintensities (WMH) (Wardlaw et al., 2015), multiple sclerosis lesions (Bakshi et al., 2008), or stroke lesions (Karnath et al., 2018). After intersubject alignment, an ICA decomposition would be incredibly useful to
discover distinct patterns of lesions over subjects. For
example, for WMH there is a average pattern of anterior
and posterior periventricular lesions, but what if it is the
case that half of subjects mainly have anterior lesions
and the other half have mainly posterior lesions; an
ICA decomposition ideally would discover these as two
distinct modes of variation. However, univariate Bernoulli
mixtures are not identifiable and so we do not see how
the authors’ DICA strategy would be applicable to this
important use case.

An eternal question for all work with mixture models
and ICA is: how many components are needed? We were
expecting the authors to dispense with at least one tuning
parameter by avoiding the PCA dimension reduction for
fMRI and directly estimating mixtures with the original
fMRI data. Although perhaps impractical computationally,
will not direct mixture modeling of the full-size fMRI data
perhaps produce a more robust method?

And finally, we were disappointed that the issue of sta-
bility and robustness were not addressed more thoroughly.
Both ICA and mixture models are notoriously sensitive
to initial conditions, and with random initialization some-
times producing quite different results (Himberg et al.,
2004). Although it is not discussed explicitly in the paper,
the ICA tool used in the authors R software (icaimax from
the ica package) uses a fixed initialization of the mixing
matrix. In contrast, it is common to use a random initializa-
tion, as is in FSL’s MELODIC and “fastica” in R and Mat-
lab. In particular, we worry that the extra source of random
initialization (mixture model + ICA) may lead to greater
instability in DICA relative to usual ICA.

In the remainder of this comment, we use the author’s
software and example data sets to examine the stability of
the estimated ICA results, comparing DICA to standard
ICA. Also, we explore the similarity and interpretability of
the discovered DICA fMRI components in comparison to
usual ICA.

2 | METHODS

The DICA method consists of two stages, first a mixture
model fitting and then ICA on the mlogit-transformed mix-
ture weights. In the current implementation, the initial-
ization of the mixture model via k-means is random and
leads to different results with each run, but the icaimax
ICA initializes the mixing matrix to the identity; other ICA
implementations initialize the mixing matrix with inde-
pendent and identically distributed (i.i.d.) random Gauss-
ian entries. We consider measuring both the stability of
the results over multiple re-runs of the algorithm, and sim-
ilarity between the DICA and usual ICA results. We con-
sider three variants of methods: DICA as originally imple-
mented (k-means-only randomness), DICA with random
initialization of the ICA mixing matrix as well, and the
usual ICA with random initialization.

Interpretation of spatial independent components (ICs)
are based on thresholded maps; we follow the DICA exam-
ple code that thresholded the ICs at the 95th percentile
in absolute value. As correlation of thresholded maps would
largely be driven by overlap, we instead measure similarity
as the spatial overlap on the basis of thresholded IC signs
(-1,0,1) at each voxel. We measure overlap with the multi-
class Dice coefficient, taking the better of the original com-
parison and a comparison with one map sign flipped. The
ICs from each run are unordered and may not corre-
spond at all between different runs or between methods.
We use a greedy matching algorithm, where we find the
pair of ICs with the best Dice coefficient, then remove that
pair from consideration, and repeat until all IC pairs
are matched. Each IC then has a Dice value for an optim-
ally matched IC. To assess stability, we use five runs of
one method on a data set, producing \( L \times (5 - \text{choose} - 2) \) Dice values; for similarity, we use five runs for each
method and then have \( L \times 5^2 \) Dice values for each pair of
methods.

We use the data provided with the authors’ software,
one fMRI data set with 71 volumes and 271,633 voxels in
mask, and one DTI data set provided as eigenvalues (3 vol-
umes) and eigenvectors (3 3-volume images) with 195,589
voxels in mask. For fMRI, we ran DICA as illustrated by
the authors (40 principal components (PCs), \( K = 20 \) mix-
ture components, \( L = 14 \) ICs), and ran ICA similarly (40
PCs, \( L = 14 \) ICs). For DTI, we likewise ran DICA as illus-
trated (\( K = 20, L = 14 \)) and with reduced dimension (\( K =
20, L = 6 \); for ICA used the six unique values of the tensor
\( K = 6 \) ICs).

Finally, we manually identified the primary sensory net-
works, visual, motor and auditory, in each of the five
runs of DICA (full random initialization). These are the
strongest of resting state fMRI networks and should be
identified with some consistency.

3 | RESULTS

The stability results are presented in Figure 1, where “RI”
stands for fixed initialization of ICA and “FI” for random
initialization of ICA. The Dice values are sobering: Simi-
larity between re-runs on the very same data produces
Dice similarity values that range fully from zero to one. For
fMRI (Figure 1 left), DICA-FI has greater stability (higher
Dice) than DICA-RI, as expected; remarkably, even DICA-
RI’s degraded similarity, median Dice around 0.4, is bet-
ter than ICA-RI. This suggests that the mixture model-
ing is providing a better latent representation than the
PCA dimension reduction alone. For DTI (Figure 1 right),
with DICA and the suggested \( L = 14 \) components, RI has
FIGURE 1  Stability of DICA and ICA methods on fMRI (left) and DTI (right) data among five re-runs for $L$ ICs. For fMRI, DICA with its original fixed-initialization ICA (DICA-FI) has better Dice similarity between runs than with random initialization (DICA-RI), but this is still better than standard ICA (ICA-RI). For DTI, there are two sets of results, one with $L = 14$ as recommended by the authors, and one with $L = 6$ to facilitate comparison with standard ICA. DICA shows an expansive range of Dice similarities, from 0 to 1, with median Dice slightly better for $L = 6$; standard ICA is only slightly better. Overall, the Dice similarities are disappointing; while some ICs may be reliable, the near zero Dice values indicate that often ICs will be obtained on some runs that will simply not correspond to ICs found on other runs.

FIGURE 2  Similarity between DICA and ICA methods on fMRI (left) and DTI (right) data over five re-runs for $L$ ICs. For fMRI, DICA with its original fixed-(DICA-FI) and random-initialization (DICA-RI) had comparable similarity with standard ICA (ICA-RI), but low median Dice of about 0.3. For DTI, DICA-FI and DICA-RI were equally (dis)similar to ICA-RI. For both fMRI and DTI, there were positive outliers above Dice of 0.8 (fMRI) and 0.9 (DTI), suggesting that some ICs are reliably discovered even if many others are not.

little impact; when running with $L = 6$ components to allow comparison with usual ICA, median Dice values are higher and if anything ICA is more reliable.

The similarity results are presented in Figure 2, where DICA is compared to ICA, without (FI) and with (RI) random ICA initialization. The Dice values are worse here for both fMRI (Figure 2 left) with a median around 0.3 and DTI (Figure 2 right) with a median of 0.1; however, for this $L = 6$ decomposition, there are a few outlier components that show high agreement.

Figure 3 shows manually identified components for five re-runs (columns), the visual network (top two rows, DICA then ICA) and auditory network (bottom two rows). For the visual network, close examination between columns shows some differences in the lateral extent of the thresholded regions, although there is more variability in ICA than DICA. For auditory, there is more evident variation, with the third run for DICA picking up little of auditory cortex, while other runs adding areas quite distant (e.g., the first run of DICA includes lateral occipital cortex), and there is variable negative weighting on the lingual gyrus (blue area near the midline). The motor network runs had variability similar to the auditory network’s (not shown).
FIGURE 3 Hand-identified resting state networks found with DICA and ICA over five re-runs. In each row, the results for five re-runs are shown, where the $L = 14$ IC maps were inspected to find the visual (top 2 rows) and auditory (bottom 2 rows) networks. DICA with random ICA initialization was used (DICA-RI). For the visual networks, the five re-runs are fairly consistent, although DICA appears to exhibit better consistency than ICA. For auditory, both exhibit variation with perhaps DICA having greater variability. In particular, considering the auditory network on the third and fourth runs of DICA, a user might make quite different conclusions about the involvement of this subject’s superior temporal gyrus in auditory processing depending on which run’s results they saw.

4 | CONCLUSIONS

We have highlighted the important issue of stability of solutions obtained from mixture model and ICA fits. On fMRI, DICA was more stable than usual ICA while on DTI, for the case of $L = 6$ ICs, ICA was more stable. Visual comparisons of hand-identified fMRI ICs look grossly similar, although close inspection reveals some appreciable differences between re-runs on the same data, and these differences are reflected quantitatively in quite low Dice coefficients. For fMRI, though, the Dice results suggest DICA may deliver greater inter-run stability than ICA alone.

These findings are, however, at best illustrative. We used but five re-runs and used a single data set that may not reflect current data quality. For example, typical resting state fMRI data have much longer time series and are usually subjected to extensive cleaning. However, we hope this
exploration helps throw a spotlight on the issue of stability in the development of data-driven methodology for fMRI.

ACKNOWLEDGMENTS
KK and TEN are supported by NIH grant R01DA048993; KK is a part of EPSRC Health Data Science Centre for Doctoral Training, EP/S02428X/1.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT
All analyses are based on freely available data, in fact, that distributed with the software of Dr Wu & co.

ORCID
Kan Keeratimahat © https://orcid.org/0000-0002-5110-2547
Thomas E. Nichols © https://orcid.org/0000-0002-4516-5103

REFERENCES
Bakshi, R., Thompson, A.J., Rocca, M.A., Pelletier, D., Dousset, V., Barkhof, F. et al. (2008) MRI in multiple sclerosis: current status and future prospects. Lancet Neurology, 7, 615–25.

Himberg, J., Hyvärinen, A. & Esposito, F. (2004) Validating the independent components of neuroimaging time series via clustering and visualization. NeuroImage, 22, 1214–1222.

Jung, T.-P., Makeig, S., McKeown, M., Bell, A., Lee, T.-W. & Sejnowski, T. (2001) Imaging brain dynamics using independent component analysis. Proceedings of the IEEE, 89, 1107–1122.

Karnath, H.O., Sperber, C. & Rorden, C. (2018) Mapping human brain lesions and their functional consequences. NeuroImage, 165, 180–189.

Wardlaw, J.M., Valdés Hernández, M.C. & Muñoz-Maniega, S. (2015) What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. Journal of the American Heart Association, 4, 001140.

Wu, B., Pal, S., Kang, J. & Guo, Y. (2021) Distributional independent component analysis for diverse neuroimaging modalities. Biometrics, 1–23.

Xu, L., Groth, K.M., Pearlson, G., Schretlen, D.J. & Calhoun, V.D. (2009) Source-based morphometry: the use of independent component analysis to identify gray matter differences with application to schizophrenia. Human Brain Mapping, 30, 711–724.

Zhang, H., Schneider, T., Wheeler-Kingshott, C.A. & Alexander, D.C. (2012) NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. NeuroImage, 61, 1000–1016.

How to cite this article: Keeratimahat, K., Nichols, T.E. (2022) Discussion on “distributional independent component analysis for diverse neuroimaging modalities” by Ben Wu, Subhadip Pal, Jian Kang, and Ying Guo. Biometrics, 78, 1113–1117. https://doi.org/10.1111/biom.13591