Single-Subject Independent Component Analysis-Based Intensity Normalization in Non-Quantitative Multi-Modal Structural MRI

Sebastian Papazoglou ^{1,*} Jens Würfel,² Friedemann Paul,^{1,3,4}
Alexander U. Brandt,¹ and Michael Scheel¹

¹NeuroCure Clinical Research Center, NeuroCure Cluster of Excellence,
Charité – Universitätsmedizin Berlin, Berlin, Germany

²Medical Image Analysis Center (MIAG AG), Basel, Switzerland

³Department of Neurology, Charité – Universitätsmedizin Berlin, Berlin, Germany

⁴Experimental and Clinical Research Center, Max Delbrück Center for Molecular Medicine
and Charité, Charité – Universitätsmedizin Berlin, Berlin, Germany

Abstract: Non-quantitative MRI is prone to intersubject intensity variation rendering signal intensity level based analyses limited. Here, we propose a method that fuses non-quantitative routine T1-weighted (T1w), T2w, and T2w fluid-saturated inversion recovery sequences using independent component analysis and validate it on age and sex matched healthy controls. The proposed method leads to consistent and independent components with a significantly reduced coefficient-of-variation across subjects, suggesting potential to serve as automatic intensity normalization and thus to enhance the power of intensity based statistical analyses. To exemplify this, we show that voxelwise statistical testing on single-subject independent components reveals in particular a widespread sex difference in white matter, which was previously shown using, for example, diffusion tensor imaging but unobservable in the native MRI contrasts. In conclusion, our study shows that single-subject independent component analysis can be applied to routine sequences, thereby enhancing comparability in-between subjects. Unlike quantitative MRI, which requires specific sequences during acquisition, our method is applicable to existing MRI data. *Hum Brain Mapp* 38:3615–3622, 2017. © 2017 Wiley Periodicals, Inc.

Key words: multi-modal MRI; structural MRI; sexual dimorphism; intensity normalization; independent component analysis

INTRODUCTION

Non-quantitative MRI contrasts feature limited intersubject comparability due to scanner properties, field inhomogeneity and non-existing standards for sequence parameters. Even for data acquired with the same scanner high signal variability impedes statistical tests based on native MR image intensity. Moreover, non-quantitative MRI contrasts are inherently mixed signals. They do not represent distinct relaxation properties such as, for example, T1 or T2 relaxation times, but mixtures or weightings of relaxation properties and other physical parameters, depending on the details of the used MR sequence [Glover, 2005]. Their non-quantitative character

A.U.B. and M.S. contributed equally to this work and share senior authorship.

*Correspondence to: Sebastian Papazoglou, NeuroCure Clinical Research Center, Charité - Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. E-mail: sebastian.papazoglou@charite.de
Received for publication 31 October 2016; Revised 3 April 2017; Accepted 5 April 2017.

DOI: 10.1002/hbm.23615

Published online 22 April 2017 in Wiley Online Library (wileyonlinelibrary.com).

is reflected using the terms T1-weighted (T1w) or T2-weighted (T2w), respectively. Although intensity normalization techniques were developed for enhancing the intrasite and intersite comparability of MR data [Ghassemi et al., 2015; Nyúl and Udupa, 1999; Shah et al., 2011], non-quantitative MRI data is commonly evaluated in a qualitative and individual manner in clinical routine, while in research one very often resorts to second level data such as tissue probability maps for further statistical processing. While tissue probability or density maps have proven to be very efficient particularly for analyzing global and local changes in gray matter (GM), for example, in voxel-based morphometry analyses [Ashburner and Friston, 2000; Leemput et al., 2003] they feature different information than the relaxation time weighted original MRI data.

For several years, there is a growing interest in fusing several MR contrasts to elucidate otherwise hidden cross-information [Sui et al., 2012]. Due to the mixed character of non-quantitative MRI contrasts, data fusion methods have the potential to detect subtle information common among all contrasts. It is hypothesized that data fusion has the potential to transform this information into new contrasts or components, in which this subtle information is enhanced. A unified post-processing of several contrast channels requires mathematical methods capable of decomposing the measured non-quantitative MR contrasts into meaningful new contrasts in a consistent way.

A widely applied method for mixed signal separation is independent component analysis (ICA) [Comon, 1994; Hyvärinen and Oja, 2000]. ICA is a data-driven unsupervised demixing technique. It was used in previous studies for identifying spatially independent structures, for example using source-based morphometry, joint source-based morphometry and linked ICA [Groves et al., 2011; Steenwijk et al., 2016; Xu et al., 2009a, 2009b]. The majority of these approaches take data from many subjects as input and estimate independent structures common among different groups of subjects. These approaches rely on GM and white matter (WM) probability maps combined with quantitative data or functional data as input. While group-based approaches benefit from increased statistical power, they are less suited for use on individual subjects. In particular, in clinical routine, it is necessary that multi-channel data are combined in a data fusion framework on a single subject level. Approaches using multi-channel data in conjunction with ICA on an individual level use non-quantitative T1w, T2w, PDw, or fluid-attenuated inversion recovery (FLAIR) data as input to ICA for estimating structurally independent components [Chai et al., 2010; Mitra et al., 2014; Nakai et al., 2004]. These components were shown to feature enhanced tissue separation, thereby supporting subsequent tissue classification steps, for example, for GM and WM discrimination and tumor or lesion delineation. However, aside from enhancing contrasts between tissue classes, the statistical independence of the components is accompanied by altered intersubject intensity variation compared to the original MR scans.

Here, we propose a method that processes a combination of T1w, T2w, and FLAIR scans using ICA that can be

applied on a single subject level (S2ICA). Our hypothesis is that the proposed method leads to a set of consistent structurally independent components with significantly reduced intersubject coefficient-of-variation (CoV) in distinct tissue classes, for example, GM and WM. We apply this method in a group of healthy subjects and demonstrate at this method detects sex differences that are not detectable with a conventional segmentation approach.

METHODS

Study Cohort

MRI data of healthy controls were taken from the local database at the NeuroCure Clinical Research Center, Charité – Universitätsmedizin Berlin, Germany. The data were previously acquired in the course of clinical studies at the NCRC. The ethics committee approved all studies and written informed consent of all participants was ensured. Our study cohort encompassed a total of 14 subjects: 7 male (mean age 36.1 years, standard deviation 15 years) and 7 female (mean age 29.9 years, standard deviation 13.2 years). The group's mean age (male and female) was 33 years with standard deviation 14 y. The mean age of females and males did not differ significantly ($P = 0.4243$). Furthermore, subjects were manually chosen according to two quality criteria: first, all image volumes must feature an optimal quality, i.e. in particular they must not feature any visible motion artifacts, and second, subjects must not show any T2 WM-hyperintensities.

MR Imaging

All MRI was performed on a 3T scanner (Magnetom TIM Trio, Siemens Healthcare, Erlangen, Germany). Imaging data acquired by the following MR sequences applied to healthy controls was used: (1) T1-weighted 3D-MPRAGE, sagittal slice selection, 176 slices, FOV: 256×256 , resolution: $1 \times 1 \times 1$ mm, TR: 1,900 ms, TE: 3.03 ms, TI: 900 ms, and flip-angle: 9, TA: 4:26 min, (2) T2w-weighted 3D-SPACE, sagittal slice selection, 176 slices, FOV: 256×256 , resolution: $1 \times 1 \times 1$ mm, TR: 5,000 ms, TE: 502 ms, and flip-angle: 120, TA: 5:52 min (3) FLAIR, sagittal slice selection, 176 slices, FOV: 256×256 , resolution: $1 \times 1 \times 1$ mm, TR: 6,000 ms, TE: 388 ms, TI: 2,100 ms, and flip-angle: 120, TA: 7:44 min.

MRI Data Pre-Processing

All data pre-processing was performed using SPM12 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London, UK) using VBM8 (Voxel-Based Morphometry, University of Jena, Jena, Germany) with Matlab (Matlab R2013a, Mathworks, Bethesda). All MRI data were noise-corrected using the non-local means filter as implemented in VBM8. In addition, VBM8 includes segmentation of T1w data yielding GM/WM/CSF-probability maps

and partial volume estimates (PVE label) [Tohka et al., 2004]. PVE label images encode segmentation results in terms of numerical values ranging continuously from one to three. Voxels featuring integer numbers were classified as single tissue voxels by the segmentation algorithm (1-CSF, 2-GM, and 3-WM), while non-integer values suggest tissue mixtures in the corresponding voxels. Binary CSF, GM, and WM masks were built from the PVE label by setting all voxels zero featuring numerical values unequal to 1, 2, and 3, respectively. The PVE label images were further used as binary brain masks by setting all non-zero voxels equal to one. The brain mask was used for skull stripping noise-corrected T1w images. Next, the noise-corrected T2w data were coregistered to the T1w brains, to prevent the coregistration algorithm from registering non-brain structures. Finally, the coregistered T2 data were skull stripped using the brain masks. Warping to MNI space (McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Toronto, Canada) was applied to all image data for voxelwise comparisons and for building group mean templates as heat map background for statistical results presentation. Tissue masks in MNI space for voxelwise statistics were built averaging all warped corresponding individual tissue probability maps and thresholding their average at 0.9. Total intracranial volume (TIV) was determined in subject space by the number of non-zero voxels in the binary brain masks.

Independent Component Analysis of Multi-Modal Structural MRI Data

In the following, the basic notation required to describe the proposed multi-modal ICA method are introduced. ICA is based on a linear model

$$\mathbf{m} = \mathbf{A}\mathbf{s}, \quad (1)$$

where the vector $\mathbf{m}$ ($N \times 1$) contains N measured mixed signals, $\mathbf{A}$ is the mixing matrix and $\mathbf{s}$ are the sources or independent components. In the following, $N = 3$ and the mixed signals are given by the three MR modalities $m_1 = \text{T1w}$, $m_2 = \text{T2w}$, and $m_3 = \text{Flair}$. Both the mixing matrix and the vector of independent components are unknown. In ICA, a degree of uniqueness of the solution to this undetermined linear system of equations is enforced estimating a set of mixing coefficients a_{ij} ($i, j = 1 \dots 3$) such that the obtained components m are statistically independent. Hence, ICA yields both the ICs $\mathbf{s}$ and the mixing matrix $\mathbf{A}$, with at most as many ICs as input mixed signals, that is, three ICs can be estimated on basis of the aforementioned three MR modalities. The mixed signal image volumes of size $N_1 \times N_2 \times N_3$ are rearranged into $(1 \times M)$ row vectors, where $M = N_1 \times N_2 \times N_3$. Then Eq. (1) is to be understood voxelwise and hence the complete concatenated set of equations reads

$$\begin{bmatrix} m_1(1) & \dots & m_1(M) \\ m_2(1) & \dots & m_2(M) \\ m_3(1) & \dots & m_3(M) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} s_1(1) & \dots & s_1(M) \\ s_2(1) & \dots & s_2(M) \\ s_3(1) & \dots & s_3(M) \end{bmatrix}. \quad (2)$$

Note that the entries of $\mathbf{A}$ are the same for each voxel, that is, the ICs are estimated based on the whole volumes rearranged as one-dimensional signals. Also note that the ordering of the mixed signals is irrelevant. Any other solution to Eq. (1) also satisfies the condition of statistical independence if they differ from $\mathbf{s}$ and $\mathbf{A}$ only by scaling and permutation of the column vectors of $\mathbf{A}$. Hence, the sign and order in which independent components are estimated is intrinsically indeterminate. This sorting and scaling problem requires a consistent and automatic way of ordering the components in such a way that the order of components in subsequent estimations is the same [Hyvärinen, 2011]. Here, this problem is tackled by sorting the estimated ICs according to the tissue structure featuring the highest mean of absolute intensity into three ICs: IC1 (GM), IC2 (WM), and IC3 (CSF). In each IC, the corresponding tissue class is the one that was selected and enhanced by ICA and the sign is chosen so that this tissue class's intensity is positive. To verify the consistency of the estimated and subsequently sorted ICs across subjects, we used a similarity measure $\gamma_{ij,kl}$ introduced in [Hyvärinen, 2011], which is based on the column vectors of the individual mixing matrices $\mathbf{A}_k$. Here, $\mathbf{A}_k = (\mathbf{a}_{1k} \ \mathbf{a}_{2k} \ \mathbf{a}_{3k})$ is the mixing matrix of subject k and the column vectors $\mathbf{a}_{1k}$ should not be confused with the matrix elements a_{ij} in Eq. (1) but instead $\mathbf{a}_{1k} = (a_{11} \ a_{21} \ a_{31})^T$ for the k -th subject and the T in the exponent denotes transpose. The measure quantifies the similarity between column vector i of subject k and column vector j of subject l . The measure similarity takes values between zero and unity, where zero means distinct and unity indicates identical column vectors. As we are interested in the

$$\text{similarity}(i, j) = \left\{ \gamma_{ij,kl} \right\} \text{ for all } k, l \quad (3)$$

where $i, j = 1 \dots 3$ and $k, l = 1 \dots 14$. The fastICA package [Hyvärinen and Oja, 2000] was used for estimating the independent components hidden among the multi-modal data composed of T1w, T2, and fluid suppressed T2w MR data. In fastICA, data are first demeaned and whitened. After that independent components are estimated by maximizing an approximate expression for negentropy as measure for statistical independence using fixed-point iteration.

Statistics

All non-voxelwise group comparisons were performed using Student's t -test implemented in Matlab. The CoV was computed according to

Figure 1.

Brain extracted and noise-corrected MRI and the corresponding estimated ICs for a single subject.

$$\text{CoV} = \frac{\text{standard deviation of the individual mean intensity}}{\text{absolute value of the individual mean intensity}} \quad (4)$$

and served as an indicator for the comparability of intensity levels across subjects [Weiskopf et al., 2013]. Boxplots were created using Graphpad Prism (GraphPad Software, CA). Voxelwise comparisons were performed using FSL's randomise (FSL 5.0.8, University of Oxford, Oxford, UK) performing 5,000 permutations for each test and using threshold-free cluster enhancement. An uncorrected $P < 0.001$ was considered significant in all cases. Voxelwise tests were performed for sex differences in GM and WM in MRIs and ICs.

RESULTS

A sample result of a S2ICA run on T1w, T2w, and Flair combined with the sorting explained in the methods section is shown in Figure 1. Each IC shows predominantly a single tissue class: GM (IC1), WM (IC2), and CSF (IC3). The

histograms of MR images and ICs are shown in Figure 2. In each independent component, a single tissue class is well separated while other tissue's mean intensity is centered close to zero. The intersubject consistency of the estimated ICs according to the sorting method described in the methods section is illustrated in Figure 3. The columns of the individual mixing-matrix show high similarity (close to unity) for the elements corresponding to presumed intra-component pairings (j, j) and low similarities for presumed intercomponent pairings, that is, (j, k) with $j \neq k$. Also, the variation is lowest for the intracomponent pairings. The intensity normalization property of the proposed method is supported by the substantially reduced intersubject CoV shown in Figure 4. While all MRI data show a moderate CoV around 0.15, the ICs feature a significantly lower CoV for the enhanced tissue class, that is, GM in IC1, WM in IC2, and CSF in IC3 ($P < 0.001$). Voxelwise group comparisons between sexes are illustrated in Figure 5. None of the original MRIs showed any differences at the given significance level ($P < 0.001$). Significantly larger IC1 values in GM were found in women in the basal ganglia, hippocampus and cerebellum. An unpaired t -test revealed significantly smaller

Figure 2.

Histograms of the MRIs and ICs. Total histograms are shown as solid black lines. Dashed lines correspond to: gray matter (blue), white matter (red), and cerebrospinal fluid (green). [Color figure can be viewed at wileyonlinelibrary.com]

($P < 0.001$) TIV in the women enrolled in this study compared to men. After adjusting for TIV mainly putamen and thalamus showed significant differences. Most striking is the widespread difference observed in WM in IC2. Although this can largely be explained by TIV, the remaining significant regions are the largest and most widespread observed in this study. Voxels with lower IC values in women compared to men were found in particular in IC3 GM and WM. After correcting for TIV, the number of significant voxels

was strongly reduced. IC1 shows significant differences in WM after accounting for TIV and similarly, IC2 GM shows significant differences after TIV correction.

DISCUSSION

In this study, we show that ICA based on single-subject MRI (S2ICA) may serve as automatic, data-based intensity

Figure 3.

Pairwise similarity of the column vectors of the mixing matrix A as estimated by ICA for all subjects. Note the different scale for intercolumn and intracolumn vector pairings.

Figure 4.

Coefficient of variation (across subjects) in (a) GM, (b) WM, and (c) CSF. Horizontal bars denote mean and 95% CI.

normalization using a set of T1w, T2w, and FLAIR image volumes in a group of healthy controls. One of the main challenges in ICA is identification and sorting of the individual components. A mathematically principled approach to solving the sorting problem is given in [Hyvärinen,

2011]. The approach to the sorting problem developed in this study was based on a similarity measure for the column vectors of the demixing matrix. This similarity measure follows a Student-*t* distribution enabling for parametric statistical tests on rigorous mathematical

Figure 5.

Sex differences (red: $f > m$, green: $m > f$) found by voxelwise permutation testing (uncorrected $P < 0.001$). [Color figure can be viewed at wileyonlinelibrary.com]

grounds. Here, we propose sorting by the tissue class featuring the highest absolute mean value and used the similarity measure to verify the consistency of the derived set of ICs for each processed subject. The intersubject CoV was significantly reduced in the ICs pronounced tissue class (see Fig. 4). This feature enhances intersubject comparability globally but also on a voxel-by-voxel basis [Weiskopf et al., 2013]. The voxelwise comparisons further support our hypothesis, that the ICs allow a more quantitative comparison across subjects on a group level. While the sex differences found in IC1 GM ($f > m$) are possibly due to the restricted size of the cohort of healthy controls enrolled in this study (Fig. 5), the effect observed in IC2 WM ($f > m$) is rather unlikely due to chance. This huge sex difference in the entire WM can be explained to great extent by including the TIV as covariate. In fact, an ROC analysis suggested that both IC1 GM, IC2 WM and TIV were able to excellently discriminate between males and females with an area-under-curve (AUC) $AUC = 0.9988$ for the ICs and $AUC = 0.9242$ for TIV. Nevertheless, our study cohort is too small for drawing any conclusions for a larger population. As the TIV was significantly smaller in women compared to men in our study cohort, we also tested using TIV as covariate of no interest. The remaining significant regions suggest larger free water content in putamen and thalamus and higher myelin density in some regions in WM of women. In contrast, significant voxels where $m > f$ appear rather random and are presumably due to the restricted size of the study cohort and multiple comparisons. The observed WM difference is in accordance with the findings in [Paus and Toro, 2009], where a larger myelin content observed in women was explained by an increased q -ratio. Other studies on sex differences have focused mainly on diffusion properties such fractional anisotropy on basis of much larger study cohorts than in our study [Dunst et al., 2014; Gur et al., 1999; Ingallhalikar et al., 2014; Kanaan et al., 2012]. This suggests that our approach potentially has an enhanced sensitivity and therefore could contribute to the ongoing discussion regarding the sexually dimorphic view of human brains [Joel et al., 2015]. This clearly suggests that ICA based on native structural MR scans needs to account for TIV. The fact, that this difference remained insignificant in the native MRI contrasts clearly suggests that impeding random intensity fluctuations have been effectively removed by ICA. ICA appears to remove these fluctuations at the cost of a significantly increased CoV in the other two tissue classes. In [Nakai et al., 2004] it was argued that ICA separates clusters of similar T1 and T2 values according to (1) short T1 component, (2) relatively short T2 component, (3) long T2 component and sorts and combines them into the ICs. Nevertheless, one has to be aware that ICA is an unsupervised statistical method and the resulting ICs do not represent distinct relaxation maps. ICA is unaware of the underlying physics of the utilized contrasts and it also cannot be excluded that aside from unwanted random

fluctuations ICA also removes interesting features. However, independent components estimated on basis of non-quantitative MRI data still feature relaxation time weighting related information in contrast to, for example, tissue probability maps derived using Gaussian mixture models [Leemput et al., 2003]. This may be of particular interest for mapping pathologic differences in WM, where tissue density oriented voxelwise approaches may not be optimal and diffusion data is unavailable.

Another important benefit of our S2ICA intensity normalization is its independence of prior knowledge about the pathology, such as, for example, the intensity normalization used in voxel-based intensitometry [Hartung et al., 2014]. There, prior knowledge about the anatomical regions affected by amyotrophic lateral sclerosis were removed from the WM using atlas tools, to generate a pathology free WM template for intensity normalization. Also, other more automated intensity normalizations are available [Nyúl and Udupa, 1999], but their usefulness in presence of pronounced pathologies such as, for example, T2 lesions or so-called dirty appearing WM in multiple sclerosis (MS), appears to be questionable [Ghassemi et al., 2015; Shah et al., 2011]. In contrast, S2ICA provides a convenient way of learning how to intensity normalize data of healthy controls. From a suitable set of controls, a mean (across subjects) demixing matrix could be learned, that could readily be applied to patient data, without need to a priori identifying normal appearing tissue in patient MR images. The similarity in the demixing matrix column vectors further supports this, as it reflects the independence of the estimated ICs with respect to individual head geometry and variations in spatial distribution of tissues. Aside from the demixing matrices, the ICs derived from healthy controls could be used to build a mean “healthy template” that will be used to compute statistical z-maps in analogy to [Hartung et al., 2014]. Particularly in presence of WM hyperintensities, we expect significantly higher z-values in these voxels than in similar maps derived on basis of the individual native scans. A promising clinical application for our method will be intensity normalization for non-quantitative clinical routine MR data for enhancing comparability between subjects. It will be interesting to apply this technique to neurodegenerative diseases that appear normal on clinical routine MRI, for example, normal appearing WM in MS and neuromyelitis optica and evaluate if subtle differences can be detected. A future improvement of our method could be to include data from different MR scanners. This would require more control of the SNR behavior in the ICs than in the present study, where all scans are filtered as implemented in VBM8. This is feasible here, as the ICs are linear combinations of the original MR scans with approximately constant mixing coefficients. We can, therefore, expect the SNR behavior in the ICs to be approximately the same across subjects for a given IC. This will not be the case for data from different MR scanners and a more careful SNR treatment will be

necessary in these scenarios. Furthermore, the proposed method is not restricted to the combination of T1, T2, and FLAIR. Other sequences could be included, adding information that enable the demixing of other features into separate ICs. From an ICA point of view, one would have to take care not to include redundant information, as then automatic dimension reduction would reduce the number of estimated independent components. Further investigations are required using appropriate patient data for finding the optimal combinations and processing pipelines. While quantitative MRI is potentially powerful as it opens the perspective of physical modeling of the underlying signal and hence modeling of the underlying tissue microstructure, it is not yet established in clinical routine. Moreover, a huge amount of non-quantitative data is already available. We, therefore, believe that S2ICA could serve as a complementary technique to quantitative MRI and tissue probability maps.

In summary, our results suggest that S2ICA can be used to estimate a set of calibrated structurally independent components, which could serve as basis for voxelwise comparisons on a group level. It is a promising method complementary to other multi-modal and also quantitative imaging as it can readily applied to data typically acquired in clinical routine and retrospectively to elucidate potentially overseen features in already existing databases.

REFERENCES

- Ashburner J, Friston KJ (2000): Voxel-based morphometry—The methods. *NeuroImage* 11:805–821.
- Chai J-W, Chi-Chang Chen C, Chiang C-M, Ho Y-J, Chen H-M, Ouyang Y-C, Yang C-W, Lee S-K, Chang C-I (2010): Quantitative analysis in clinical applications of brain MRI using independent component analysis coupled with support vector machine. *J Magn Reson Imaging* 32:24–34.
- Comon P (1994): Independent component analysis, a new concept. *Signal Process* 36:287–314.
- Dunst B, Benedek M, Koschutnig K, Jauk E, Neubauer AC (2014): Sex differences in the IQ-white matter microstructure relationship: A DTI study. *Brain Cogn* 91:71–78.
- Ghassemi R, Brown R, Narayanan S, Banwell B, Nakamura K, Arnold DL (2015): Normalization of white matter intensity on T1-weighted images of patients with acquired central nervous system demyelination. *J Neuroimaging* 25:184–190.
- Glover GH (2005): Handbook of MRI pulse sequences, M. A. Bernstein, K. F. King and X. J. Zhou. Elsevier Academic Press, 2004, ISBN: 0-12-092861-2. *NMR Biomed* 18:202–203.
- Groves AR, Beckmann CF, Smith SM, Woolrich MW (2011): Linked independent component analysis for multimodal data fusion. *NeuroImage* 54:2198–2217.
- Gur RC, Turetsky BI, Matsui M, Yan M, Bilker W, Hughett P, Gur RE (1999): Sex differences in brain gray and white matter in healthy young adults: Correlations with cognitive performance. *J Neurosci* 19:4065–4072.
- Hartung V, Prell T, Gaser C, Turner MR, Tietz F, Ilse B, Bokemeyer M, Witte OW, Grosskreutz J (2014): Voxel-based MRI intensitometry reveals extent of cerebral white matter pathology in amyotrophic lateral sclerosis. *PLoS One* 9: e104894.
- Hyvärinen A (2011): Testing the ICA mixing matrix based on inter-subject or inter-session consistency. *NeuroImage* 58: 122–136.
- Hyvärinen A, Oja E (2000): Independent component analysis: Algorithms and applications. *Neural Netw* 13:411–430.
- Ingallhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K, Hakonarson H, Gur RE, Gur RC, Verma R (2014): Sex differences in the structural connectome of the human brain. *Proc Natl Acad Sci USA* 111:823–828.
- Joel D, Berman Z, Tavor I, Wexler N, Gaber O, Stein Y, Shefi N, Pool J, Urchs S, Margulies DS, Liem F, Hänggi J, Jäncke L, Assaf Y (2015): Sex beyond the genitalia: The human brain mosaic. *Proc Natl Acad Sci USA* 112:15468–15473.
- Kanaan RA, Allin M, Picchioni M, Barker GJ, Daly E, Shergill SS, Woolley J, McGuire PK (2012): Gender differences in white matter microstructure. *PLoS One* 7:e38272.
- Leemput KV, Maes F, Vandermeulen D, Suetens P (2003): A unifying framework for partial volume segmentation of brain MR images. *IEEE Trans Med Imaging* 22:105–119.
- Mitra J, Bourgeat P, Fripp J, Ghose S, Rose S, Salvado O, Connolly A, Campbell B, Palmer S, Sharma G, Christensen S, Carey L (2014): Lesion segmentation from multimodal MRI using random forest following ischemic stroke. *NeuroImage* 98:324–335.
- Nakai T, Muraki S, Bagarinao E, Miki Y, Takehara Y, Matsuo K, Kato C, Sakahara H, Isoda H (2004): Application of independent component analysis to magnetic resonance imaging for enhancing the contrast of gray and white matter. *NeuroImage* 21:251–260.
- Nyúl LG, Udupa JK (1999): On standardizing the MR image intensity scale. *Magn Reson Med* 42:1072–1081.
- Paus T, Toro R (2009): Could sex differences in white matter be explained by g ratio? *Front Neuroanat* 3:14.
- Shah M, Xiao Y, Subbanna N, Francis S, Arnold DL, Collins DL, Arbel T (2011): Evaluating intensity normalization on MRIs of human brain with multiple sclerosis. *Med Image Anal* 15: 267–282.
- Steenwijk MD, Geurts JJG, Daams M, Tijms BM, Wink AM, Balk LJ, Tewarie PK, Uitdehaag BMJ, Barkhof F, Vrenken H, Pouwels PJW (2016): Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. *Brain* 139: 115–126.
- Sui J, Adali T, Yu Q, Calhoun VD (2012): A review of multivariate methods for multimodal fusion of brain imaging data. *J Neurosci Methods* 204:68–81.
- Tohka J, Zijdenbos A, Evans A (2004): Fast and robust parameter estimation for statistical partial volume models in brain MRI. *NeuroImage* 23:84–97.
- Weiskopf N, Suckling J, Williams G, Correia MM, Inkster B, Tait R, Ooi C, Bullmore ET, Lutti A (2013): Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: A multi-center validation. *Front Neurosci* 7:95.
- Xu L, Groth KM, Pearson G, Schretlen DJ, Calhoun VD (2009a): Source-based morphometry: The use of independent component analysis to identify gray matter differences with application to schizophrenia. *Hum Brain Mapp* 30:711–724.
- Xu L, Pearson G, Calhoun VD (2009b): Joint source based morphometry identifies linked gray and white matter group differences. *NeuroImage* 44:777–789.