Improving current normalization approaches to detect longitudinal changes in gray and white matter using DTI.

1851

Florencia Jacobacci[a] , Jorge Jovicich[b] , Gonzalo Lerner[a] , Arnaud Boré[c] , Márcia Renata Hidalgo-Marques[d] , Khallil Taverna Chaim[d] , Edson Amaro[d] , Jorge Armony[e] , Julien Doyon[c,f] , Valeria Della-Maggiore[a]

a IFIBIO-Houssay - CONICET - Department of Physiology, School of Medicine, U. of Buenos Aires, Buenos Aires, Argentina; ***b*** Center for Mind/Brain Sciences (CIMEC), University of Trento, Rovereto, Italy; ***c*** Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal, Montréal, Canada; ***d*** PISA, Faculdade de Medicina FMUSP, Universidade de Sao Paulo, Sao Paulo, Brazil; ***e*** McGill University, Montreal, Canada; ***f*** BIC, MNI, McGill University, Montreal, Quebec

florenciajacobacci@gmail.com

Background

- Scalar DTI measures such as MD and FA are increasingly being used to evaluate longitudinal changes in brain microstructure induced by learning[1,2], development[3,4] or neurodegenerative disease[5,6].
- These studies require pre-processing scalar DTI volumes to align them to a template in standard stereotaxic space (a process called normalization).
- There is no unanimity regarding the optimal registration approach to draw valid conclusions from voxelwise analysis conducted on multisession DWI data.

Traditional pipeline

- FSL[7] based
- Tract-based spatial statistics[8,9] (TBSS)
- Normalization tool: flirt + fnirt
- Normalization target: FMRIB58 FA template
- Moving image: FA
- Normalization strategy: Direct registration of FA images to FMRIB58 FA template
- Focuses on evaluating differences along white-matter tracts
- Not optimized for longitudinal studies
- Not optimized for evaluating differences in gray matter

Proposed pipeline

- Advanced Normalization Tools (ANTs)[10] is a toolbox that has gained substantial attention in the last years.
- It is reported to be more sensitive for detecting true changes than TBSS[7,8,9], and leads to lower type I error rate caused by registration[11,12].

Objective

In this work, we used a voxel-based approach (VBA) to compare different normalization pipelines based on ANTs.

Our goal was to look for a volumetric normalization pipeline that optimized detection of longitudinal changes in Diffusion Tensor Images (DTI).

We aimed at minimizing across-session test-retest reproducibility error in the following traits:

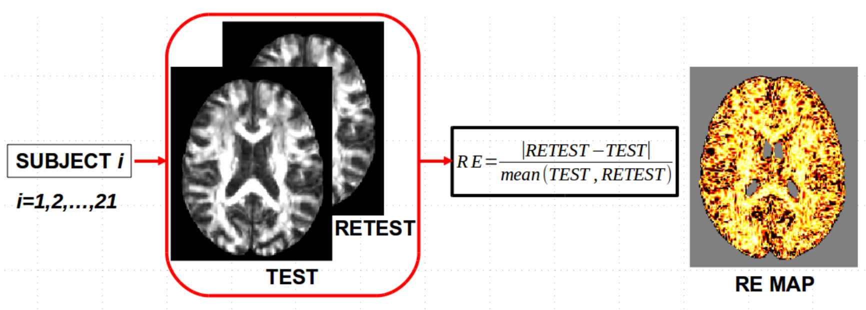
1) Normalization tool (FSL vs ANTs)
2) Normalization target (FMRIB58 FA template vs MNI152 T1 template)
3) Moving image to bring to standard space (MD, FA or B0)
4) Normalization strategy (direct normalization vs intermediate template)

Methods

- 21 normal subjects (11 female, ages 18-31, mean 23.6, std dev. 3.1) underwent two scans obtained 24 hours apart
- Preprocessing steps (common to all pipelines):
  - correction of susceptibility-induced distortions using topup[13]
  - correction of eddy currents induced distortions, motion correction and b-vectors corrections were done using eddy[14]
  - FA and MD images were created using DTIfit[7]
- Registration of DTI images to standard stereotaxic space was carried out according to the parameters defined in each analytical approach.
- In all the pipelines, transformations that map the moving image from each subject to the target image are then applied to the other DTI images in the same space.

ANALYTICAL APPROACH	Pipeline	Normalization tool	Normalization target	Moving image	Normalization strategy
1: Normalization tool	1	*FSL (flirt+fnirt)*	FMRIB 58 FA template	FA image	Direct
	2	*ANTs (SyN algorithm)*			
2: Normalization target	1	ANTs (SyN algorithm)	*FMRIB58 FA template*	FA image	Direct
	2		*MNI152 T1 template*		
3: Moving image	1	ANTs (SyN algorithm)	MNI152 T1 template	*FA image*	Direct
	2			*MD image*	
	3			*b0 image*	
	4			*FA + b0 images*	
	5			*MD + b0 images*	
4: Normalization strategy	1	ANTs (SyN algorithm)	MNI152 T1 template	FA image	*Direct*
	2				*via an individual FA template*
	3				*via a group FA template*

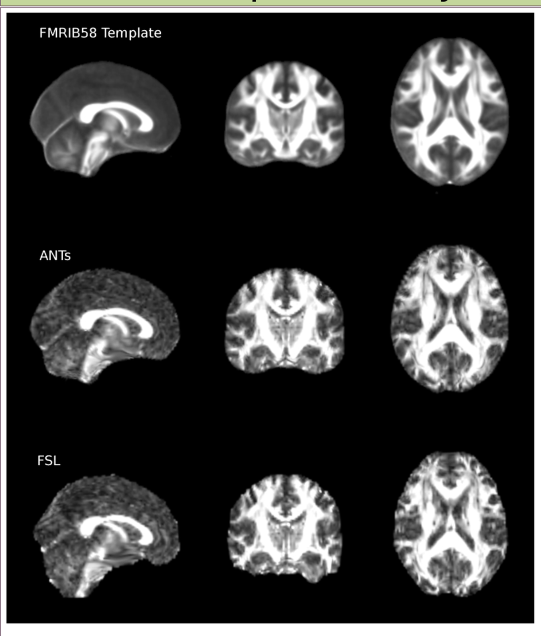
- Across-session test-retest reproducibility errors (RE) of FA and MD were defined in a voxelwise fashion for each subject (and each pipeline) as the absolute difference between the test and the retest DTI measures and divided by their mean value in the two sessions[15].



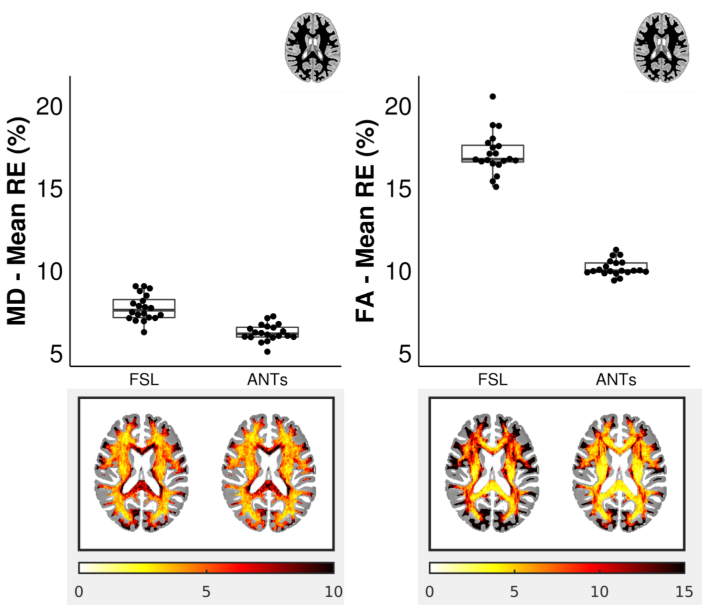
- Atlas gray matter (GM, upper pannel) and atlas white matter (WM, lower pannel) masks were used to assess mean reproducibility error on different tissue types. Voxels belonging to each mask are shown in black.

Results

1) Normalization tool: ANTs nonlinear transformation yields lower reproducibility error in WM for both MD and FA



Registration to FMRIB58 standard space using two nonlinear registration toolboxes: FSL or ANTs. FMRIB58 FA template is shown in the upper pannel. FA registered to template using ANTs (middle pannel) or FSL (lower pannel).



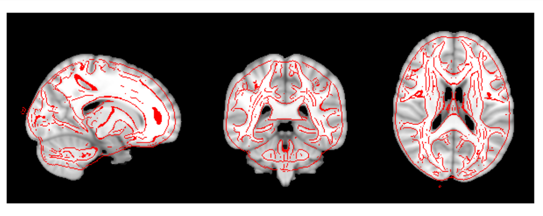
Statistics: One-way repeated measures ANOVA

MD-Mean RE: $F(1,19)=126.41$; $p<0.0001$

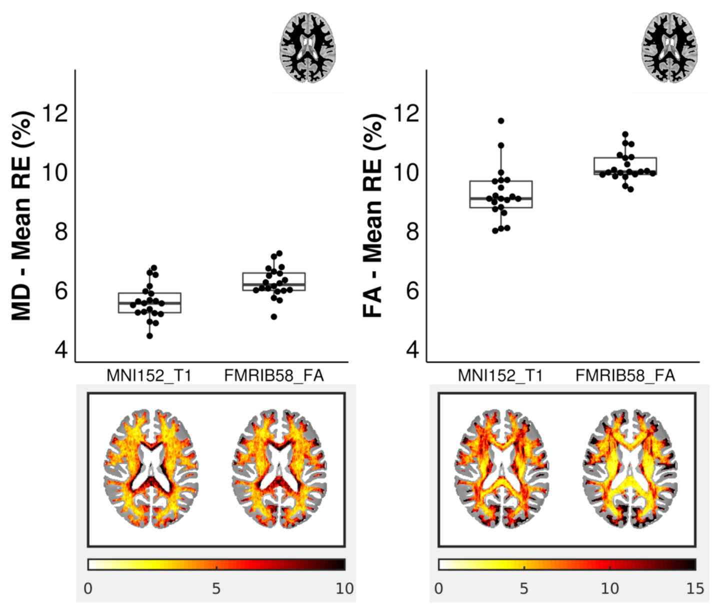
FA-Mean RE: $F(1,19)=1103.83$; $p<0.0001$

Mean reproducibility error (RE) of MD and FA in white matter (WM) mask for two normalization tools, FSL and ANTs. The inset shows the mask used. Note that reproducibility error is larger in FA than in MD.
Voxelwise maps for the mean RE in WM across subjects are shown (axial slice coordinate z=19 mm).

2) Target image: MNI152 T1 template yields lower reproducibility error than FMRIB58 FA template



Comparison of the two templates in stereotaxic space commonly used as targets of normalization. FMRIB58 FA template (red edges) is shown overlaid over MNI152 T1 template.



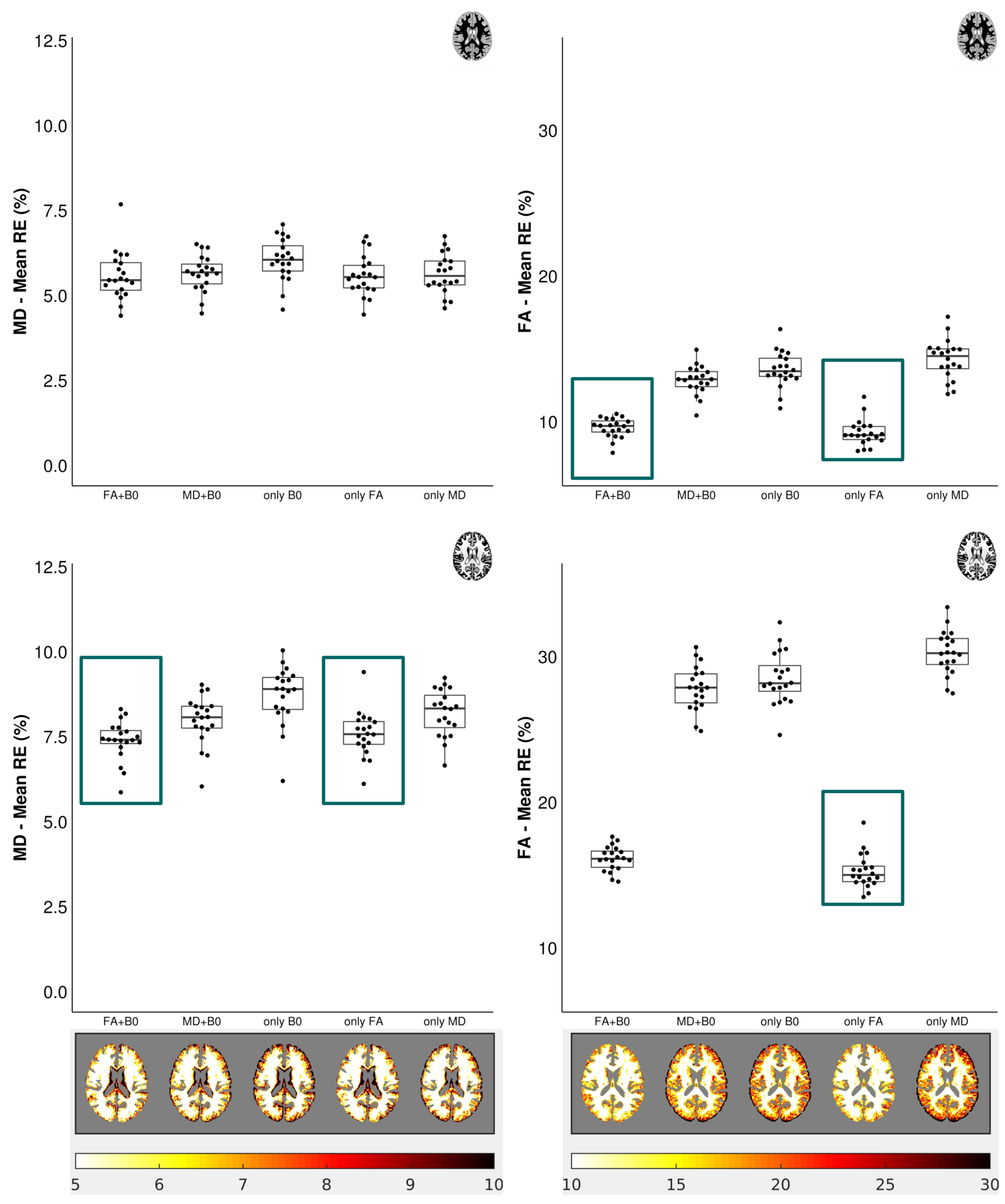
Statistics: One-way repeated measures ANOVA

MD-Mean RE: $F(1,19)=108.15$; $p<0.0001$

FA-Mean RE: $F(1,19)=58.75$; $p<0.0001$

Mean reproducibility error (RE) of MD and FA in white matter (WM) mask for two normalization targets, MNI152 T1 template and FMRIB58 FA template. Note that reproducibility error is larger in FA than in MD.
Voxelwise maps for the mean RE in WM across subjects are shown (axial slice coordinate z=19 mm).

3) Moving image: Using FA as moving image reduces reproducibility error of MD and FA in GM and WM



Statistics: One-way repeated measures ANOVA

MD-Mean RE in WM
$F(1.309,36.272)=9.483$; $p<0.001$
Post-hoc Tukey test: *"only b0"* pipeline differs from the rest of the pipelines ($p<0.01$)

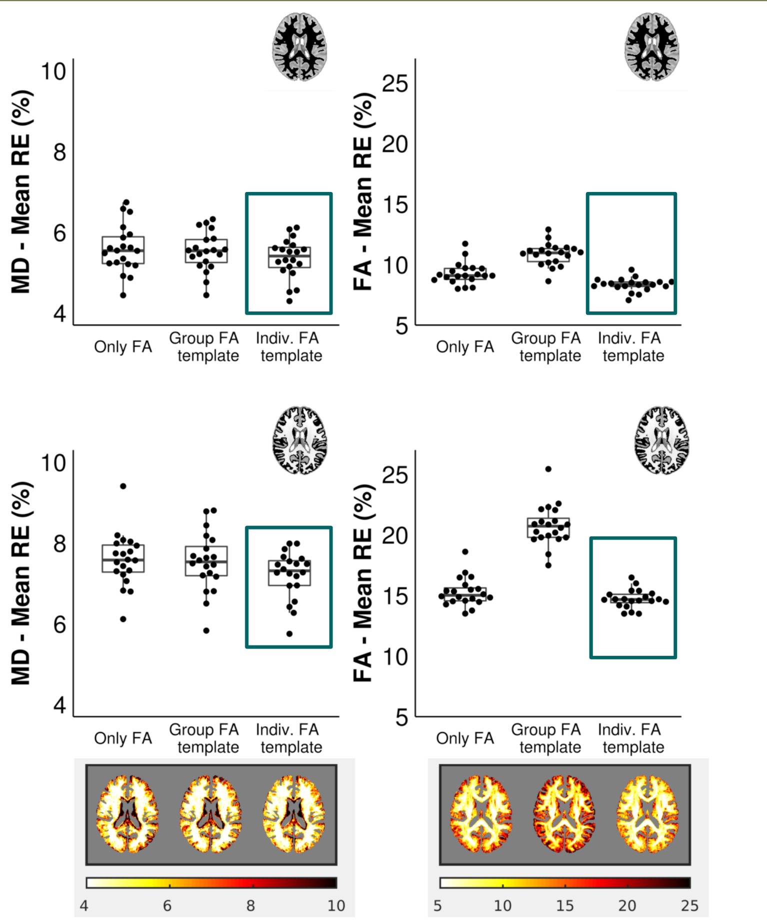
FA-Mean RE in WM
$F(2.406,45.714)=199.158$; $p<0.0001$
Post-hoc Tukey test: *"FA+b0"* and *"only FA"* pipelines show the lowest RE, and differ from the rest of the pipelines ($p<0.0001$)

MD-Mean RE in GM
$F(2.413,45.838)=39.856$; $p<0.0001$
Post-hoc Tukey test: *"FA+b0"* and *"only FA"* pipelines show the lowest RE, and differ from the rest of the pipelines ($p<0.0001$)

FA-Mean RE in GM
$F(2.395,45.500)=1078.86$; $p<0.0001$
Post-hoc Tukey test: *"only FA"* pipeline shows the lowest RE, and differs from the rest of the pipelines ($p<0.0001$)

Mean reproducibility error (RE) of MD and FA in white matter (WM, upper pannels) and gray matter (GM, lower pannels) masks using different moving images. The inset shows the mask used. Note that different scales are used for pannels showing MD and FA results.
Voxelwise maps for the mean RE in WM across subjects are shown (axial slice coordinate z=19 mm).

4) Normalization strategy: Using an intermediate individual FA template for normalization reduces reproducibility error of MD and FA in both GM and WM. The reduction in reproducibility error of MD is associated with higher SNR.



Statistics: One-way repeated measures ANOVA

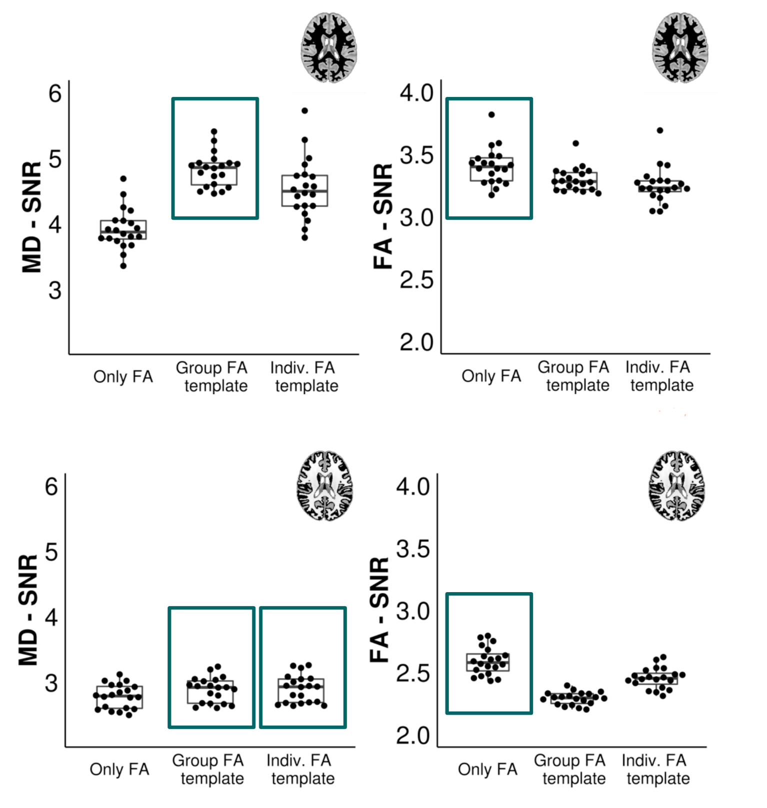
MD-Mean RE in WM:
$F(2,38)=10.188$; $p<0.0001$
Post-hoc Tukey test: *"indiv. FA template"* pipeline shows the lowest RE, and differs from the rest of the pipelines ($p<0.03$)

FA-Mean RE in WM:
$F(1.510,28.684)=98.573$; $p<0.0001$
Post-hoc Tukey test: *"indiv. FA template"* pipeline shows the lowest RE, and differs from the rest of the pipelines ($p<0.0001$)

MD-Mean RE in GM:
$F(1.447,27.498)=3.950$; $p=0.043$
Post-hoc Tukey test: *"indiv. FA template"* pipeline shows the lowest RE, and differs from the rest of the pipelines ($p<0.001$)

FA-Mean RE in GM:
$F(1.393,26.472)=230.70$; $p<0.0001$
Post-hoc Tukey test: *"indiv. FA template"* pipeline shows the lowest RE, and differs from the rest of the pipelines ($p<0.02$)

Mean reproducibility error (RE) of MD and FA in white matter (WM, upper pannels) and gray matter (GM, lower pannels) masks using different normalization strategies. Note that different scales are used for pannels showing MD and FA results. Voxelwise maps for the mean RE in WM across subjects are shown (axial slice coordinate z=19 mm).



Statistics: One-way repeated measures ANOVA

MD-Mean RE in WM:
$F(1.460,27.737)=167.543$; $p<0.0001$
Post-hoc Tukey test: *"group FA template"* pipeline yields the highest SNR and is different from the rest ($p<0.0001$)

FA-Mean RE in WM:
$F(2,38)=19.056$; $p<0.0001$
Post-hoc Tukey test: *"only FA"* pipeline yields the highest SNR and is different from the rest ($p<0.005$)

MD-Mean RE in GM:
$F(1.440,27.358)=22.518$; $p<0.0001$
Post-hoc Tukey test: *"indiv. FA template" and "group FA template"* pipelines yield the highest SNR and are different from *"only FA"* pipeline ($p<0.0001$)

FA-Mean RE in GM:
$F(1.521,28.904)=168.093$; $p<0.0001$
Post-hoc Tukey test: *"only FA"* pipeline yields the highest SNR and is different from the rest ($p<0.0001$)

Signal-to-noise ratio (SNR) of MD and FA in white matter (WM, upper pannels) and gray matter (GM, lower pannels) masks using different normalization strategies. Note that different scales are used for pannels showing MD and FA results.

Conclusions

1) Using ANTs improves registration reproducibility for longitudinal studies.
2) Using the MNI152 T1 template improves reproducibility of MD and FA in WM. It is the best option if one aims at evaluating changes in FA, but is of particular importance if one is interested in assessing changes in MD or wishes to study cortical gray matter.
3) Using FA as moving image yields the best reproducibility of FA in GM. Combining FA and b0 gives the lowest reproducibility error of MD in GM but it does not significantly improve reproducibility over using FA alone. Using MD as moving image deteriorates reproducibility.
4) Normalizing via an individual FA Template yields the best reproducibility.

REFERENCES:

1.Landi, S.M. et al. (2011), Journal of Neuroscience, 31(33), 11808–11813.
2. Sagi, Y. Et al. (2012), Neuron, 73(6), 1195–1203.
3. Sexton, C. E. et al. (2014), Journal of Neuroscience, 34(46), 15425–15436.
4. Krogsrud, S.K. et al. (2016), NeuroImage, 124, 473–486.
5. Keihaninejad, S. et al. (2013), NeuroImage, 72, 153–163.
6. Lam, B.Y.K. et al. (2014), Human Brain Mapping, 35(7), 3547–3557.
7. Smith S. M. et al. (2004), NeuroImage, 23(S1):208-219.
8.Smith S. M. et al. (2006), NeuroImage, 31:1487-1505.
9. Smith S. M. et al. (2007), Nat
10. Avants, B.B. et al. (2011), NeuroImage, 54(3), 2033–2044.
11. Klein, A. et al. (2009), NeuroImage, 46(3), 786–802.
12. Schwarz, C.G. et al. (2014), NeuroImage, 94, 65–78.
13. Andersson J.L.R. et al. (2003), NeuroImage, 20(2):870-888, 2003.
14. Andersson J.L.R. and Sotiropoulos S.N. (2016), NeuroImage, 125:1063-1078.
15. Papinutto, N.D. et al. (2013), Magnetic Resonance Imaging, 31(6), 827–839.

ACKNOWLEDGMENTS:

- RBIQ/QBIN Grant RI-05
- ANPyC FONCYT: PICT 2015-0844
- MINDEF PIDDEF 2014/2017 #17