

Neuroconductor: An R Platform for Medical Imaging Analysis

John Muschelli

https://johnmuschelli.com/neuroc_talk/smi_2021.html

Johns Hopkins Bloomberg School of Public Health

(Bio)Statisticians were **LATE** to the
neuroimaging game

- Engineers got there first, a lot MRI started in engineering (also, there are more of them)
- MATLAB reigned supreme
 - Cost, closed source had issues, also...the language (ever tried to read a CSV?)
- Python is stepping in: Nipy (also great logos)

Movie: Margin Call

There are three ways to make a living in this business: be first; be smarter; or cheat.



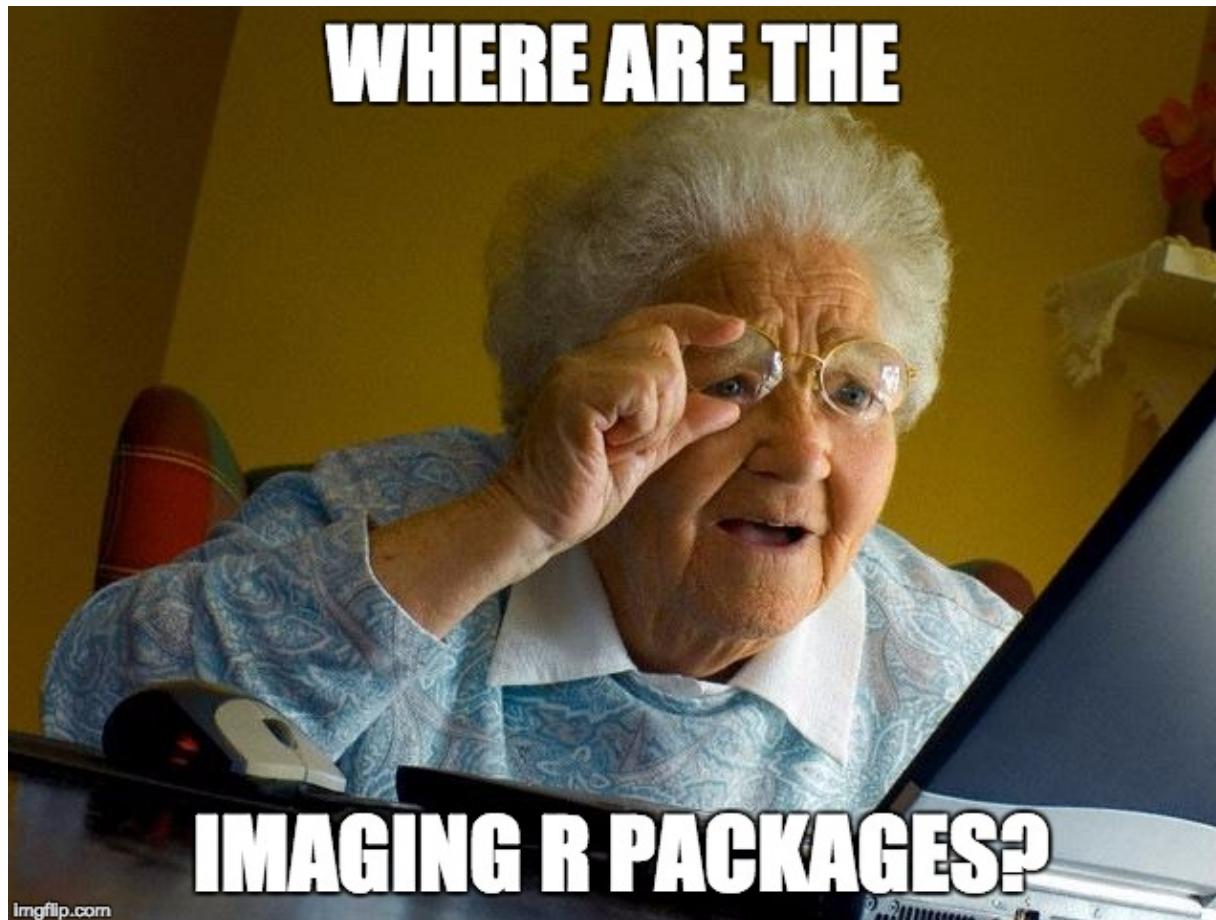
If you can't be first, be different.
Also ... steal

R is a language and environment
for statistical computing
and graphics.

<https://cran.r-project.org/>

Should read “R started as ...”

What does R have for medical imaging?



<https://imgflip.com/memegenerator/Grandma-Finds-The-Internet>

What does R have for medical imaging?

CRAN Task View: Medical Image Analysis

Maintainer: Brandon Whitcher

Contact: bwhitcher at gmail.com

Version: 2016-12-30

URL: <https://CRAN.R-project.org/view=MedicalImaging>

Data Input/Output

DICOM

The industry standard format, for data coming off a clinical imaging device, is [DICOM](#) (Digital Imaging and Communications in Medicine). The DICOM "standard" is very broad and very complicated. Roughly speaking each DICOM-compliant file is a collection of fields organized into two four-byte sequences (group,element) that are represented as hexadecimal numbers and form a *tag*. The (group,element) combination announces what type of information is coming next. There is no fixed number of bytes for a DICOM header. The final (group,element) tag should be the "data" tag (7FE0,0010), such that all subsequent information is related to the image(s).

- The packages [oro.dicom](#), [divest](#), [fmri](#) and [tractor.base](#) (part of the [tractor](#) project) provide R functions that read DICOM files and facilitate their conversion to ANALYZE or NIfTI format.

ANALYZE and NIfTI

Although the industry standard for medical imaging data is DICOM, another format has come to be heavily used in the image analysis community. The [ANALYZE](#) format was originally developed in conjunction with an image processing system (of the same name) at the Mayo Foundation. An Anlayze (7.5) format image is comprised of two files, the "hdr" and "img" files, that contain information about the acquisition and the acquisition itself, respectively. A more recent adaption of this format is known as [NIfTI-1](#) and is a product of the Data Format Working Group (DFWG) from the Neuroimaging Informatics Technology Initiative (NIfTI). The NIfTI-1 data format is almost identical to the ANALYZE format, but offers a few improvements: merging of the header and image information into one file (.nii), re-organization of the 348-byte fixed header into more relevant categories and the possibility of extending the header information.

- The packages [RNifti](#), [AnalyzeFMRI](#), [fmri](#), [tractor.base](#) (part of the [tractor](#) project), [oro.nifti](#), and [neuroim](#) all provide functions that read/write ANALYZE and NIfTI files.

Magnetic Resonance Imaging (MRI)

Diffusion Tensor Imaging (DTI)

- The R package [dti](#) provides structural adaptive smoothing methods for the analysis of diffusion weighted data in the context of the DTI model. Due to its edge preserving properties these smoothing methods are capable of reducing noise without compromizing significant structures (e.g., fibre tracts). The package also provides functions for DTI data processing from input,

Envy: Bioconductor



- centralized bioinformatics packages (> 1300)
- large community/developer team
- published tutorials and workflows
- additional requirements to CRAN (e.g. packages need vignettes)

Why should you care?

- Students see resources - “I can learn genomics with Bioconductor easier than imaging”
- Researchers see analyses - “That’s the analysis I want to do, ready made”
- Papers with software get **cited more**
 - Usually 1 paper for analysis/demonstration and 1 for software
- All of these lead to **funding**

There is Enough Space for Python
and R, but we need our niche(s)

Where is the niche for R/(Bio)statisticians?

- Medical imaging, not just neuroimaging
 - lungs, livers, oh my!
- Not just MRI: CT, PET, EEG, etc.
- Mixed effects and longitudinal modeling
- Data Harmonization
 - Steal (aka adapt) from genomics/BioC
- Label fusion statistical properties
- BAYES

[HTML] Removing inter-subject technical variability in magnetic resonance imaging studies

[JP Fortin](#), [EM Sweeney](#), [J Muschelli](#), [CM Crainiceanu](#)... - *NeuroImage*, 2016 - Elsevier

Magnetic resonance imaging (MRI) intensities are acquired in arbitrary units, making scans non-comparable across sites and between subjects. Intensity normalization is a first step for the improvement of comparability of the images across subjects. However, we show that ...

☆ 59 [Cited by 51](#) Related articles All 15 versions »

[HTML] Harmonization of multi-site diffusion tensor imaging data

[JP Fortin](#), [D Parker](#), [B Tunç](#), [T Watanabe](#), [MA Elliott](#)... - *NeuroImage*, 2017 - Elsevier

Diffusion tensor imaging (DTI) is a well-established magnetic resonance imaging (MRI) technique used for studying microstructural changes in the white matter. As with many other imaging modalities, DTI images suffer from technical between-scanner variation that hinders ...

☆ 59 [Cited by 245](#) Related articles All 10 versions »

[HTML] Harmonization of cortical thickness measurements across scanners and sites

[JP Fortin](#), [N Cullen](#), [YI Sheline](#), [WD Taylor](#), [I Aselcioglu](#)... - *NeuroImage*, 2018 - Elsevier

With the proliferation of multi-site neuroimaging studies, there is a greater need for handling non-biological variance introduced by differences in MRI scanners and acquisition protocols. Such unwanted sources of variation, which we refer to as "scanner effects", can ...

☆ 59 [Cited by 205](#) Related articles All 13 versions »



An R Platform for
Medical Imaging Analysis

What is Neuroconductor?

1. A centralized repository of packages (N = 115)
2. A community of developers (N = 29) and users
3. A website <https://neuroconductor.org/>.
 - with tutorials and help
4. A team helping developers and users
5. Set of packages to allow R to perform processing.

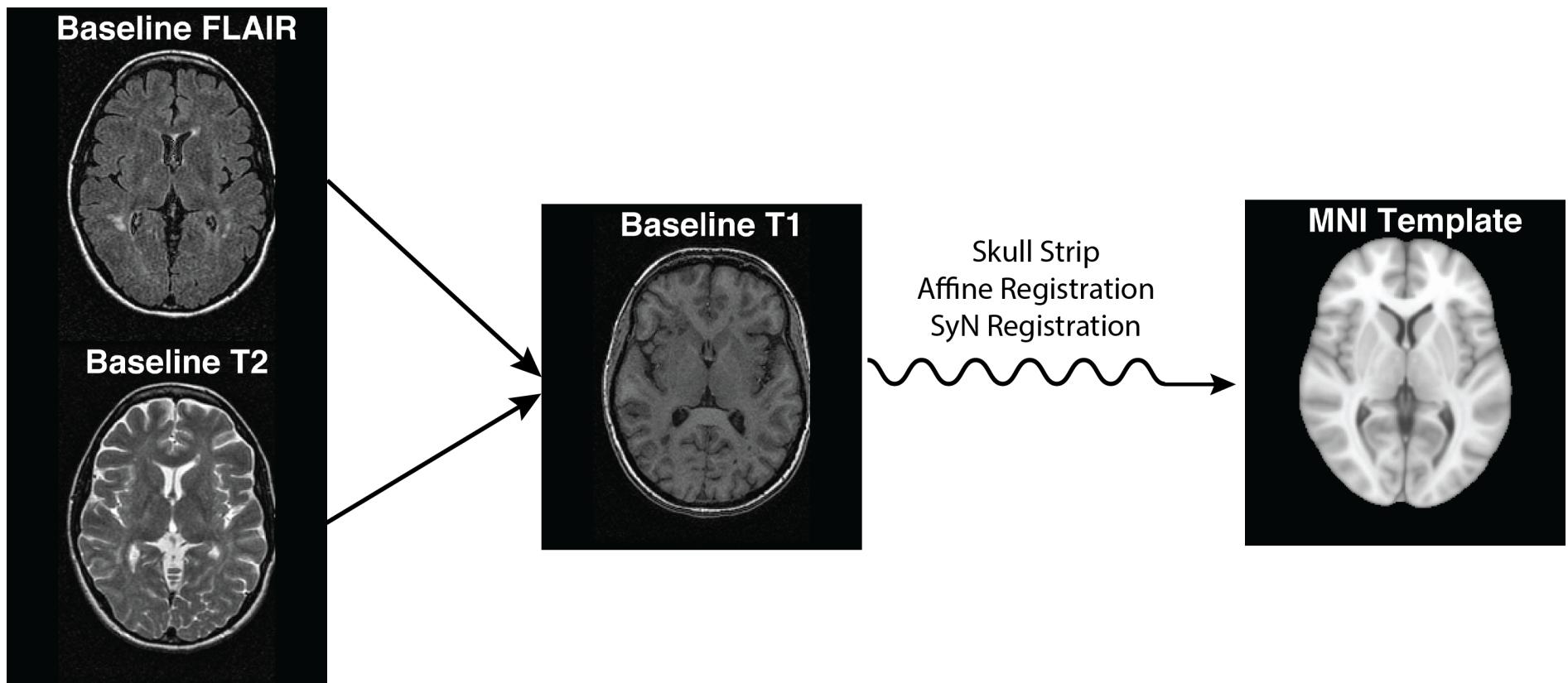
Benefits of Neuroconductor

Allow imaging to use all R has to offer:

- **Statistics and Machine Learning**
(tensorflow)
- **Versioning** and testing (GitHub)
- **Reproducible** reports and analyses
- **Shiny** (web applications)
- Genomics/Imaging analysis in one platform
 - Bioconductor

Package Features: Image Registration

- ANTsR, RNiftyReg, fslr



Package Features: Inhomogeneity Correction

- ANTsR, fslr

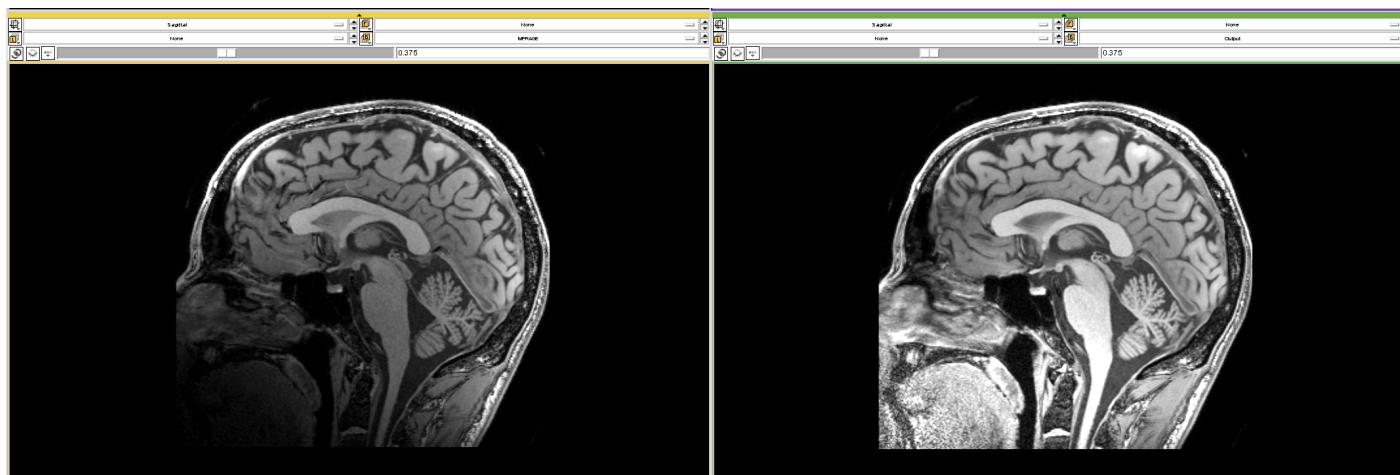


Image From https://www.slicer.org/w/images/7/77/MRI_Bias_Field_Correction_Slicer3_close_up.png

Package Features: Intensity Normalization

- WhiteStripe -
<https://github.com/muschellij2/whitestripe>
- RAVEL -
<https://github.com/Jfortin1/RAVEL>
- ComBat -
<https://github.com/Jfortin1/ComBatHarmonization>

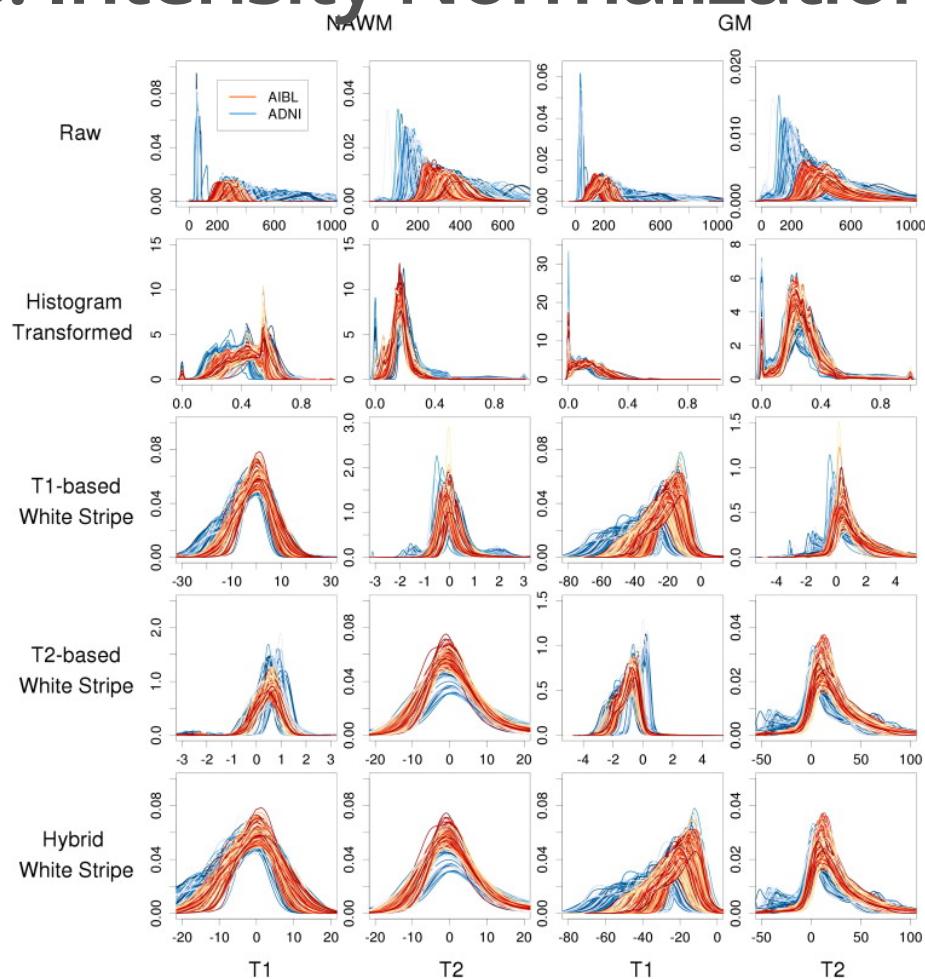
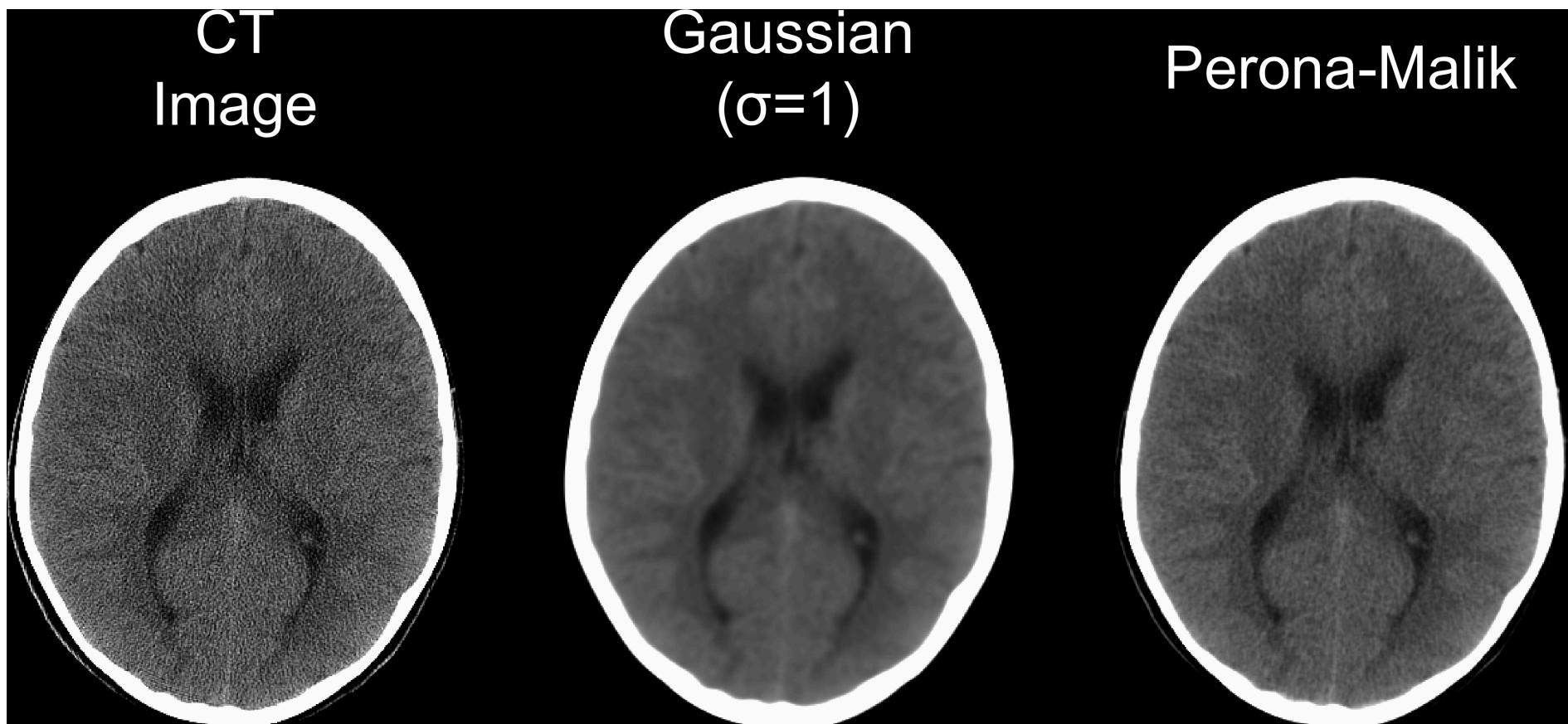


Figure from Shinohara, Russell T., et al. "Statistical normalization techniques for magnetic resonance imaging." *NeuroImage: Clinical* 6 (2014): 9-19.

Package Features: Smoothing

- fslr, ANTsR, AnalyzeFMRI

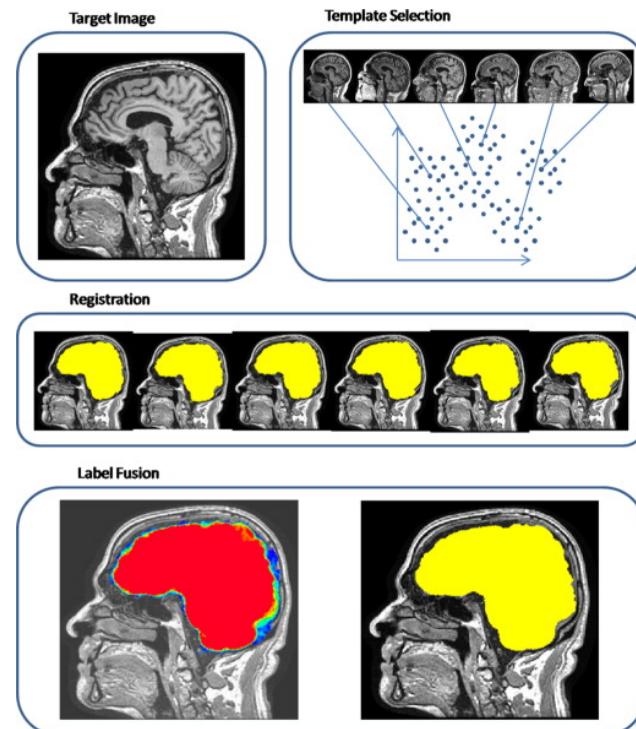


Package Features: Data

- neurohcp - [Human Connectome Project](#)
- kirby21.t1, kirby21.fmri - <https://www.nitrc.org/projects/multimodal>
- nitrcbot - <https://www.nitrc.org/>
- Rxnat - XNAT databases
- <https://www.templateflow.org/>

Multi-Atlas Label Fusion (MALF)

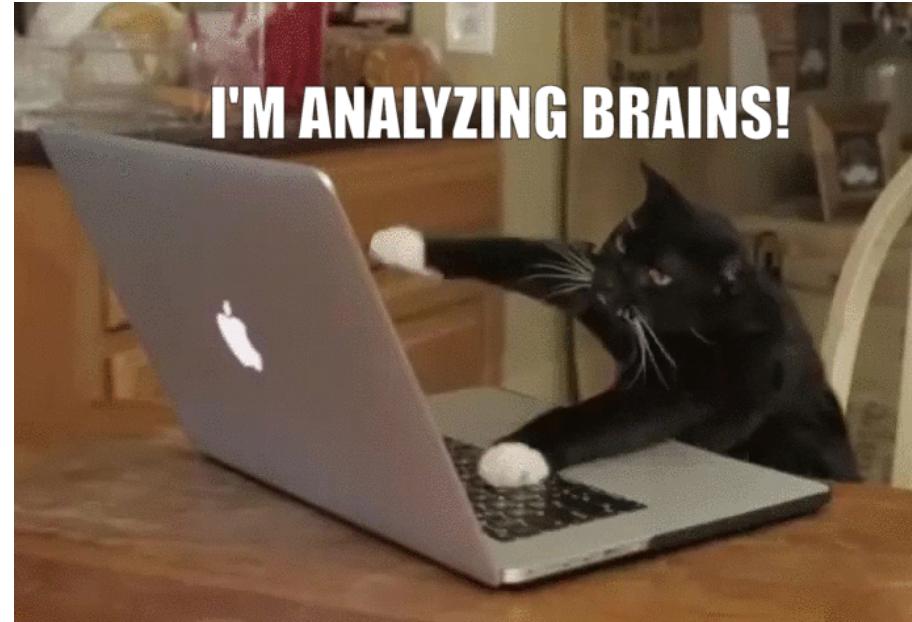
- malf.templates - <https://github.com/muschellij2/malf.templates>



Neuroconductor

Goal:

Detailed **tutorials**
on how to actually
perform an
analysis



From <http://i.imgur.com/0Y1xISa.gifv>.

<http://johnmuschelli.com/neuroc>

Training we are providing

Coursera Course:
Introduction to
Neurohacking In R

<https://www.coursera.org/learn/neurohacking/>



http://johnmuschelli.com/imaging_in_r/

Things in the Works/Other work

- <https://github.com/ANTsXNet> - Wide array of deep learning applications for R
- <https://github.com/msharrock/deepbleed> - CNN for ICH Segmentation
- EEG -
<https://github.com/craddm/eegUtils>

github.com/ANTsXNet

BrainSegmentationPatchBased

Three tissue (CSF, GM, WM) patch-based brain segmentation

● R ⚡ 2 ★ 0 ⓘ 2 ⏺ 0 Updated on Oct 8, 2019

BrainSegmentation

Brain segmentation (6-tissue) ANTsRNet app

● R ⚡ 3 ★ 2 ⓘ 2 ⏺ 1 Updated on Oct 4, 2019

ProtonMRI_LungSegmentation

Proton MRI lung segmentation ANTsRNet app

● R ⚡ 0 ★ 1 ⓘ 0 ⏺ 0 Updated on Sep 12, 2019

FunctionalLungSegmentation

Functional hyper polarized gas MRI lung segmentation ANTsRNet app

● R ⚡ 0 ★ 1 ⓘ 0 ⏺ 0 Updated on Sep 8, 2019

PigLungSegmentation

● R ⚡ 0 ★ 0 ⓘ 0 ⏺ 0 Updated on Aug 28, 2019

BrainAgeGender

Predict age and gender from gray matter probability images

● R ⚡ 0 ★ 1 ⓘ 0 ⏺ 0 Updated on Feb 11, 2019

Questions?

Submit Packages at:

<https://neuroconductor.org/submit-package>

Funding: NIH 2R01NS060910 and U24HG010263