

Imaging Statistics in R

https://github.com/muschellij2/Neuroimaging_in_R

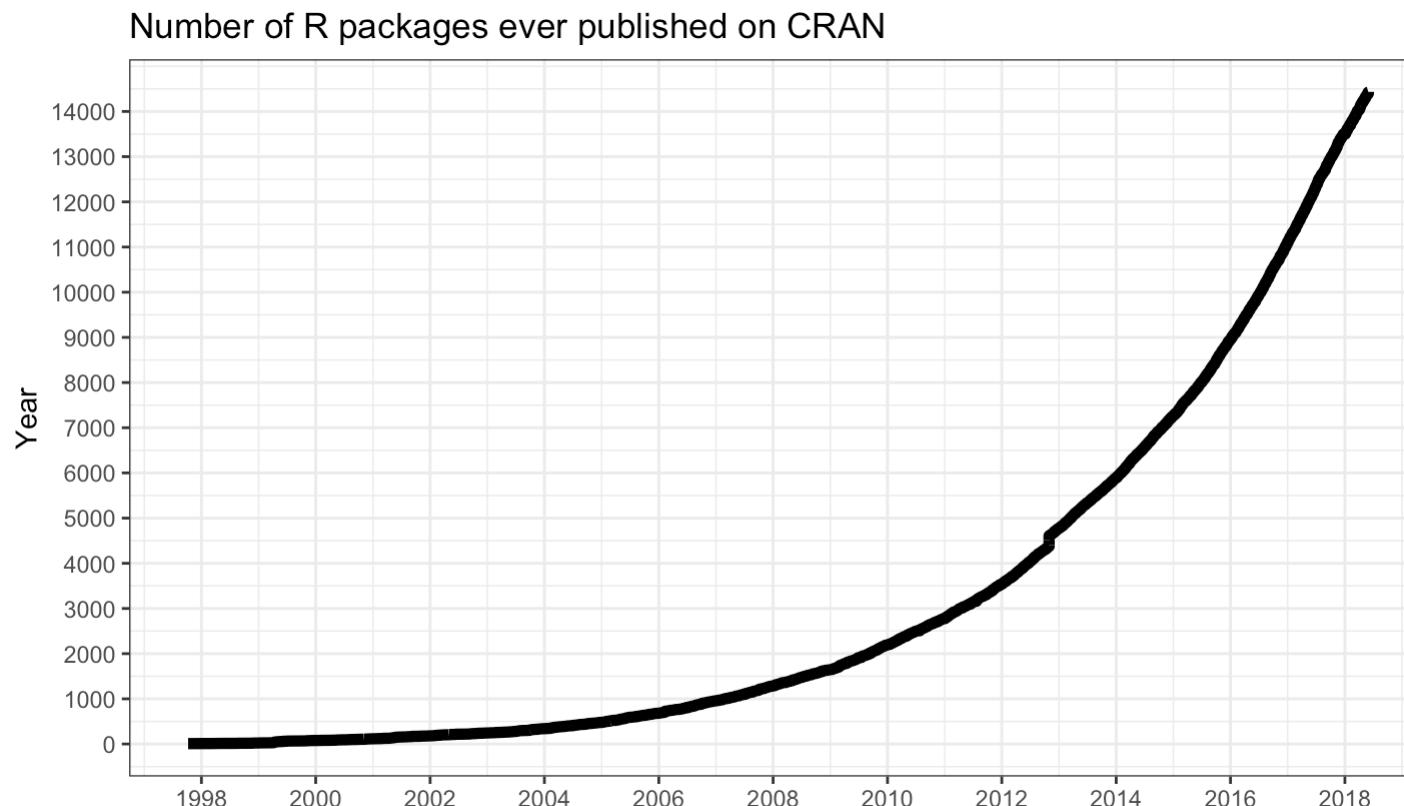
John Muschelli http://johnmuschelli.com/smi_2018

CRAN Packages 5-10 years ago vs. today

- BioC+GitHub 15258

(<https://www.rdocumentation.org/>)

Script from <https://gist.github.com/daroczig/3cf06d6db4be2bbe3368>



Imaging in R: 5 years ago

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)

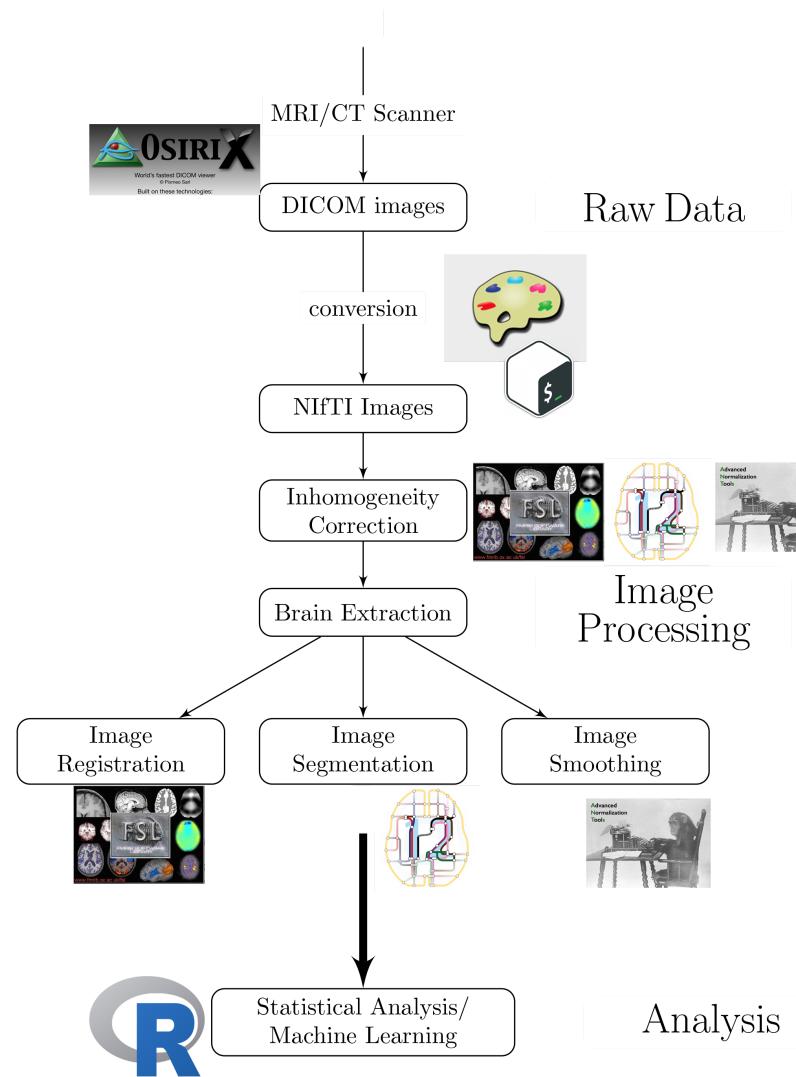
Diffusion Tensor Imaging (DTI)

- 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,

Workflow: 5

years ago

- bash 
- FSL 
- ANTs 
- MRIcroGL 
- OsiriX 
- SPM 12 

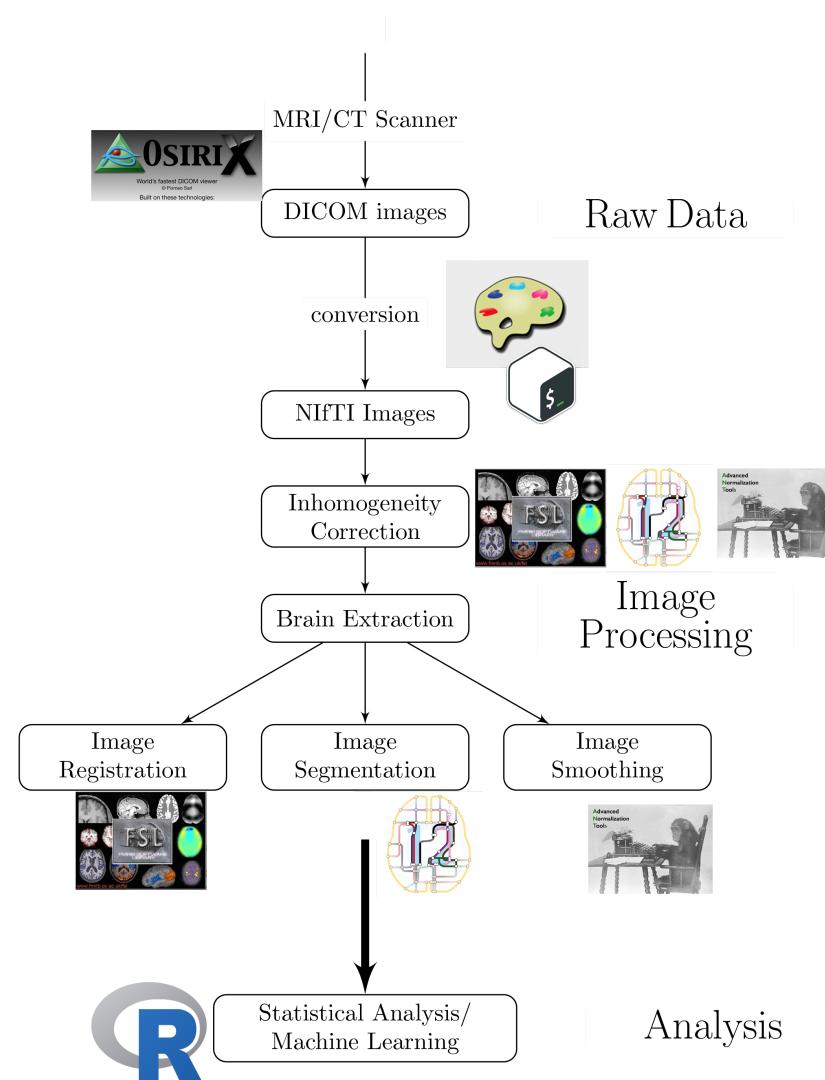


Workflow: 5

years ago

Multiple pieces
of software used

- all different syntax



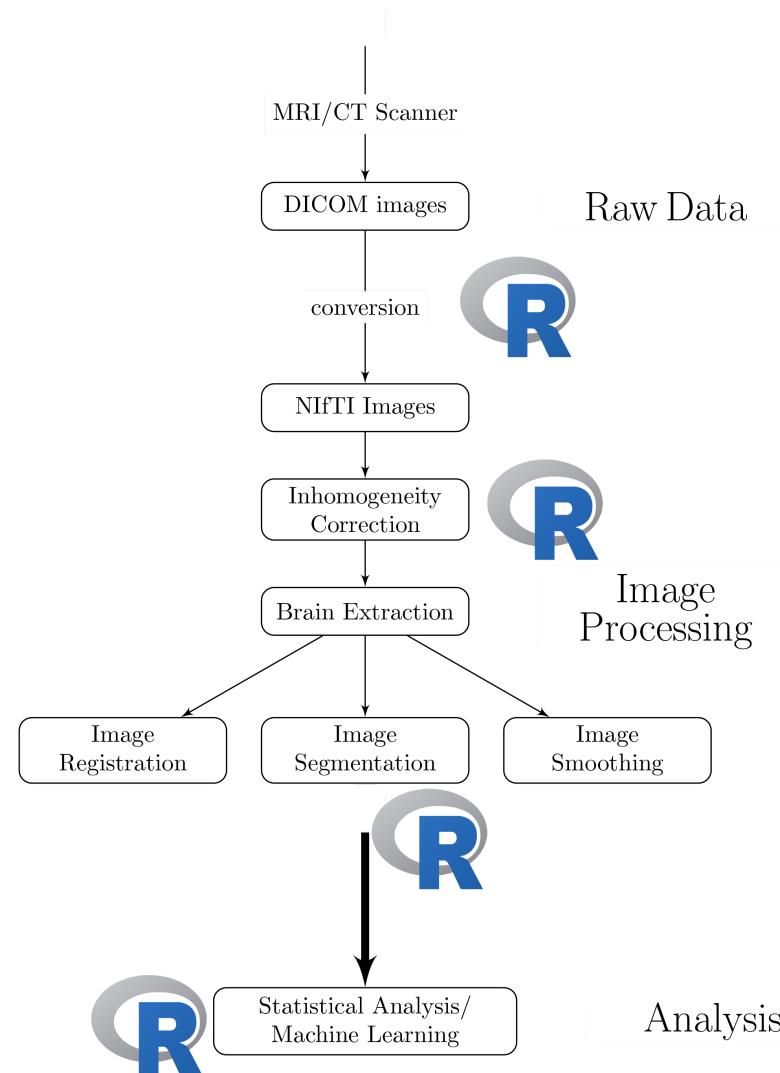
R: Programming and Interface Language

Workflow: Now

- all R code
 - interface/pipeline tool
 - “native” R code

Complete pipeline

- preprocessing and analysis



Imaging in R: Now

Install Submit Package **List Packages** Help ▾ About ▾ 

Home / Neuroconductor Packages

Neuroconductor Packages

[View Dependency Graph](#) | [Browse releases](#)

Show **100** entries

Package Name	Version	Package Title	Maintainer(s)	GitHub URL	Last updated
afnir	0.4.5	Wrapper Functions for 'AFNI' (Analysis of Functional 'NeuroImages')	John Muschelli	muschellij2/afnir	2018-04-26
ANTsR	0.6.1	ANTs in R: Quantification Tools for Biomedical Images	Brian B. Avants	stnava/ANTsR	2017-10-05
ANTsRCore	0.5.0	Core Software Infrastructure for ANTsR	Brian B. Avants	stnava/ANTsRCore.git	2018-04-20
brainKCCA	0.0.0.9001	Region-level Connectivity Network Construction via Kernel Canonical Correlation Analysis	Jian Kang	xuboyue/brainKCCA_v4	2018-03-23
brainR	1.5.2	Helper Functions to 'misc3d' and 'rgl' Packages for Brain Imaging	John Muschelli	muschellij2/brainR	2017-11-29
cfma	1.0	Causal Functional Mediation Analysis	Yi Zhao	zhaoyi1026/cfma	2018-05-25
cifti	0.4.5	Toolbox for Connectivity Informatics Technology Initiative ('CIFTI') Files	John Muschelli	muschellij2/cifti	2018-04-18
dcmriS4	0.57.1	A Package for Image Analysis of DCE-MRI (S4 Implementation)	Brandon Whitcher	bjw34032/dcmriS4	2017-10-08
dcm2niir	0.6.1	Conversion of 'DICOM' to 'NIFTI' Imaging Files Through R	John Muschelli	muschellij2/dcm2niir	2017-12-04
dcmsort	0.2.4	Sort DICOM Images	John Muschelli	muschellij2/dcmsort	2018-04-21
DensParcorr	1.0	Dens-Based Method for Partial Correlation Estimation in Large Scale Brain Networks	Yikai Wang	johnzonwyk/DensParcorr	2017-10-09
divest	0.4.1	Get Images Out of DICOM Format Quickly	Jon Clayden	jonclayden/divest	2017-10-06
dti	1.2.6.2	Analysis of Diffusion Weighted Imaging (DWI) Data	Karsten Tabelow	WIAS-BERLIN/dti	2017-10-20
EveTemplate	0.99.14	JHU-MNI-ss (Eve) template	Jean-Philippe Fortin	Jfortin1/EveTemplate	2017-10-06
extrantrs	3.7.5.9002	Extra Functions to Build on the ANTsR Package	John Muschelli	muschellij2/extrantrs.git	2018-05-02

Weaknesses of Imaging in R

- Speed, but Rcpp, not always ported to Rcpp
 - A lot of work to interface with large libraries
- Memory issues
- MATLAB is more popular and has GUIs
- System Dependencies are not handled well in packages

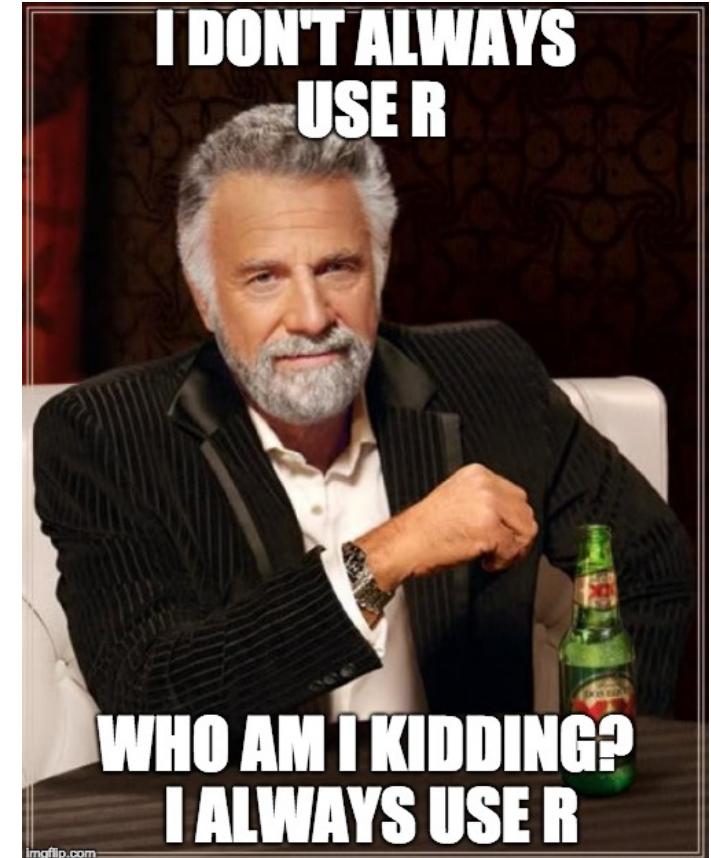


<https://giphy.com/gifs/top-gun-need-for-speed-i-feel-the-26AHLNr8en8J3ovOo>

- e.g. (fsl, FSL, install FSL at
<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FsllnstaIIon>)

Strengths of Imaging in R

- Free
- Large/strong R community
- RStudio
- Validated statistical models
- Package development system
- Neuroconductor
- Shiny



<https://imgflip.com/i/2blglk>

Threats of Imaging in R

- Not much different than Python
 - scikit-learn/ Nipype
- Neural Networks
- BME have tools for advanced statistical analysis
- **Study design is not as valued in imaging (neuro)**
 - one stronghold of statisticians
- Credit/Incentives for software

Opportunities of Imaging in R

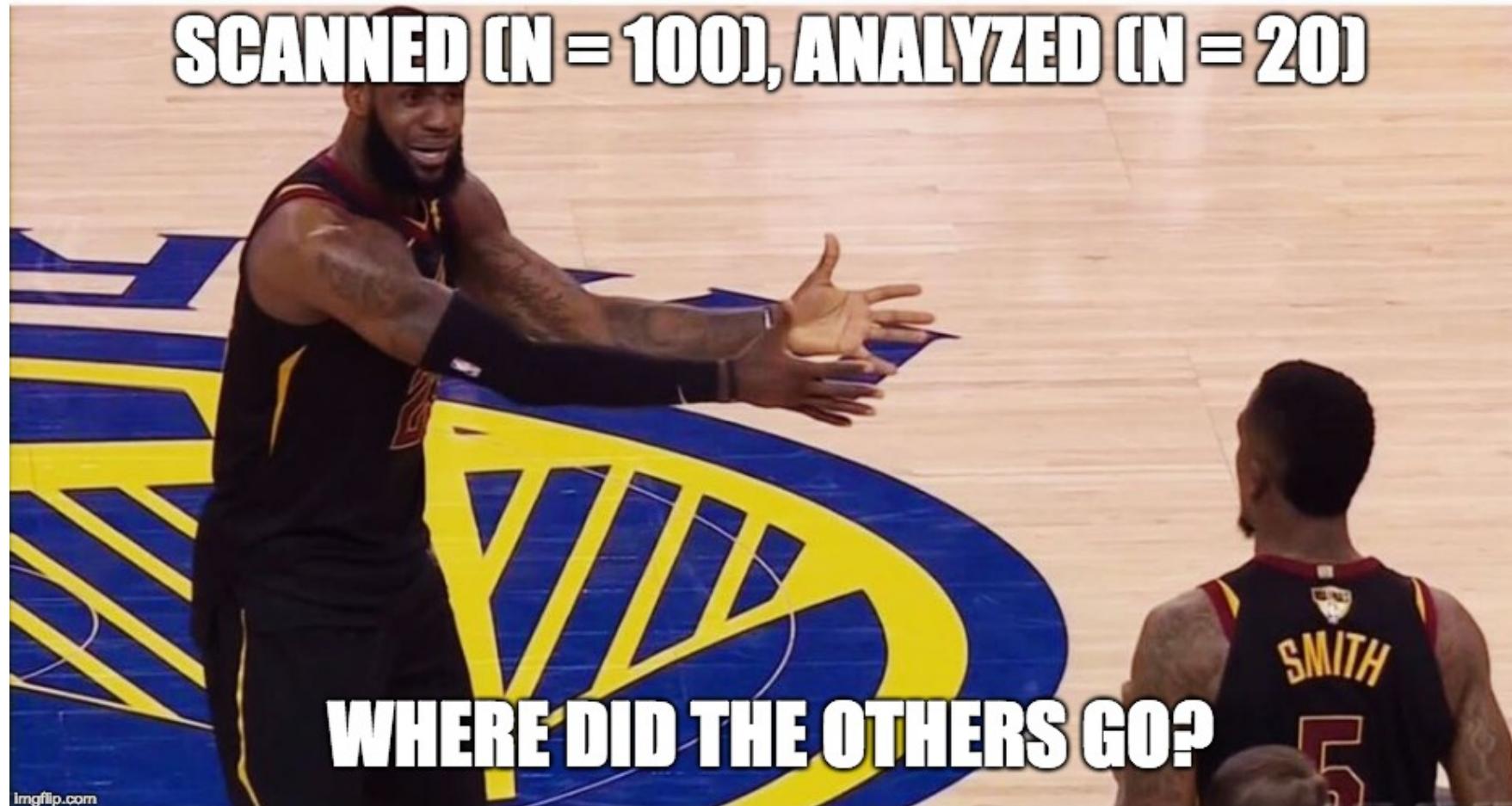
- Neural Networks: Keras and Tensorflow
- Most analyses still do general linear models
- Genomic analyses are very similar
 - use Bioconductor tools in a new way

Opportunities of Imaging in R

- Neural Networks: Keras and Tensorflow
- Most analyses still do general linear models
- Genomic analyses are very similar
 - use Bioconductor tools in a new way

Opportunities: CONSORT diagrams

<https://imgflip.com/i/2bltg8>



Opportunities of Stats in Imaging

- Task fMRI: compare baseline to 1+ conditions
- Smoothed to account for neighborhood
- Run LM on time-series voxelwise with AR-type correlation with HRF-convolved design matrix (X) + motion/other factors (Z)

$$Y_{iv} = X_i \beta_{iv} + Z_i \theta_{iv} + \varepsilon_{iv} \quad \varepsilon_{iv} \sim N(0, \Sigma_{iv}) \quad \Sigma_{iv} \sim AR$$

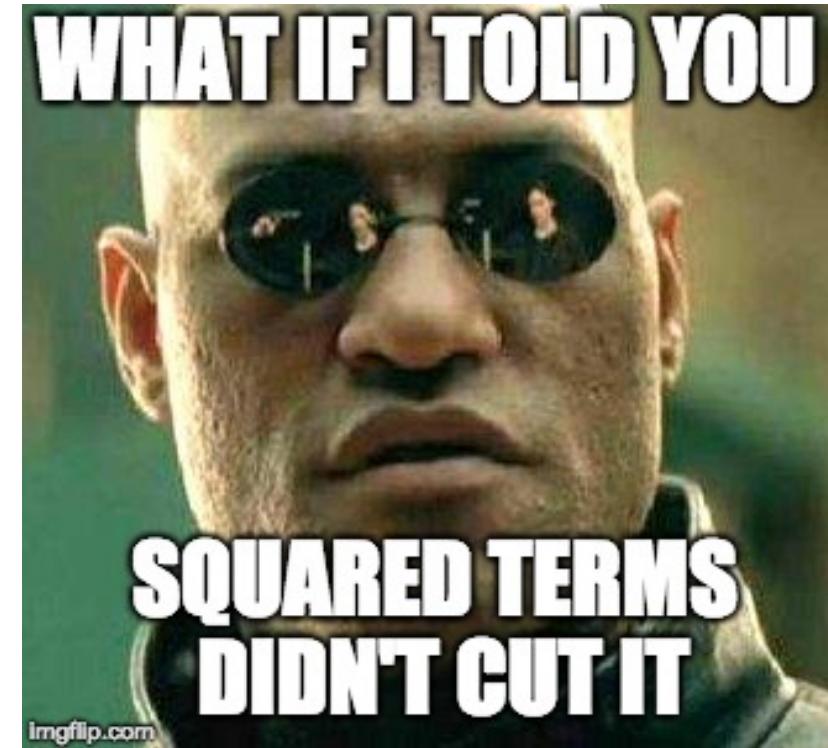
Z_i can be different for each i (scrubbing), but usually isn't.

Opportunities of Stats in Imaging

- Get β_i map in MNI space
- Run t-test, random effects analysis (hard to find sometimes), or permutation test
- Perform multiple comparisons correction (maybe with spatial component)

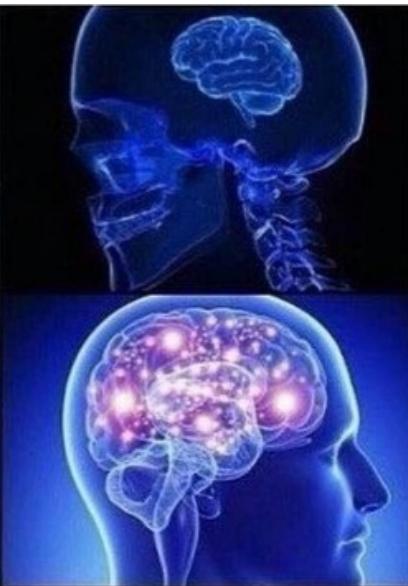
Opportunities of Stats in Imaging

- What if the motion had a non-linear effect?
 - currently: use derivatives and squared effects
 - why not GAMs? (speed)_{voxel} package
- fully Bayes (spatial) hard because of data size (maybe empirical)



<https://imgflip.com/i/2bltw2>

**WHEN YOUR BRAIN
DOESN'T REGISTER
CORRECTLY**



**WHAT PEOPLE
THINK
FMRI IS**



**WHAT PEOPLE
THINK RESTING
STATE IS**



**AND...LASER
BRAIN**

imgflip.com

Opportunities of Stats in Imaging: Laser Brain

<https://imgflip.com/i/2blg14>

Opportunities of Stats in Imaging

- Multi-site/multi-scanner data from consortium
 - batch effects abound



<https://imgflip.com/i/2blid6>

What we need: tutorials

Guides for Developers

- Frequently Asked Questions (FAQ)
- Installation Guides
- Required Readings
- Data
- General Tutorials
- Disease-specific Tutorials

Neuroconductor

2018-04-19

Guides for Developers

1. [Preparing Your Package for Submission](#)
2. [Changes to your Package](#)

Frequently Asked Questions (FAQ)

Please visit the [FAQ](#) for information on how to begin.

Installation Guides

1. [Installing devtools](#)
2. [Installing ANTsR](#)

Required Readings

1. [NIfTI Basics](#)

Data

1. [Downloading Human Connectome Project Data](#)
2. [Downloading Functional Connectomes Projects](#)

What we really need: tutorials on fMRI

- preprocess fMRI from ANTsR
- spm12r - <https://neuroconductor.org/neuroc-help-fmri-analysis-spm12r>
- fslr - <https://neuroconductor.org/neuroc-help-fmri-analysis-fslr>

What we need: Shiny GUIs

(Maybe) What we need: challenges

https://grand-challenge.org/all_challenges/

Grand-Challenges Why Challenges? All Challenges Host A Challenge Contributors



All Challenges

Here is an overview of all challenges that have been organized within the area of medical image analysis that we are aware of. If you know any study that would fit in this overview, or want to advertise your challenge, please leave a message in the [forum](#) or send mail to support@grand-challenge.org and we will add the challenge to the list on this page.

Showing 165 projects of 165

Filter by:

Open for submissions (98) Data download (101) Hosted on Grand-challenge (26)

2018



EndoVis

As an endoscopic vision CAI challenge at MICCAI, our aim is to provide a formal framework for evaluating the current state of the art, gather researchers in the field and provide high quality data with protocols for validating endoscopic vision algorithms.

Workshop: Sep 16, 2018

Associated with: [MICCAI 2018](#)

Hosted on: [grand-challenge.org](#)



ICIP2018-Challenge

Can you develop a method for automatic detection of cancerous regions in breast cancer histology images?

[Data download](#)

Workshop: Jun 27, 2018

Associated with: [ICIP 2018](#)

Hosted on: [grand-challenge.org](#)



IDRiD

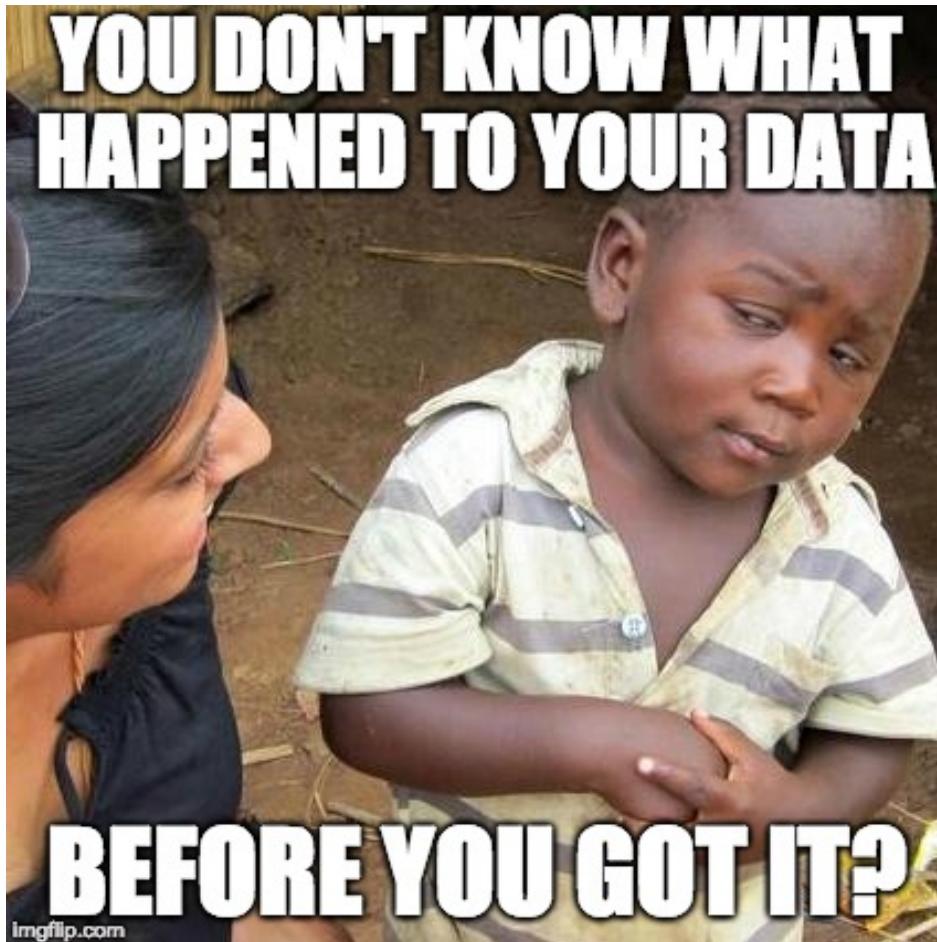
This challenge evaluates automated techniques for analysis of fundus photographs. We target segmentation of retinal lesions like exudates, microaneurysms, and hemorrhages and detection of the optic disc and fovea. Also, we seek grading of fundus images according to the severity level of DR and DME.

[Data download](#)

Associated with: [ISBI 2018](#)

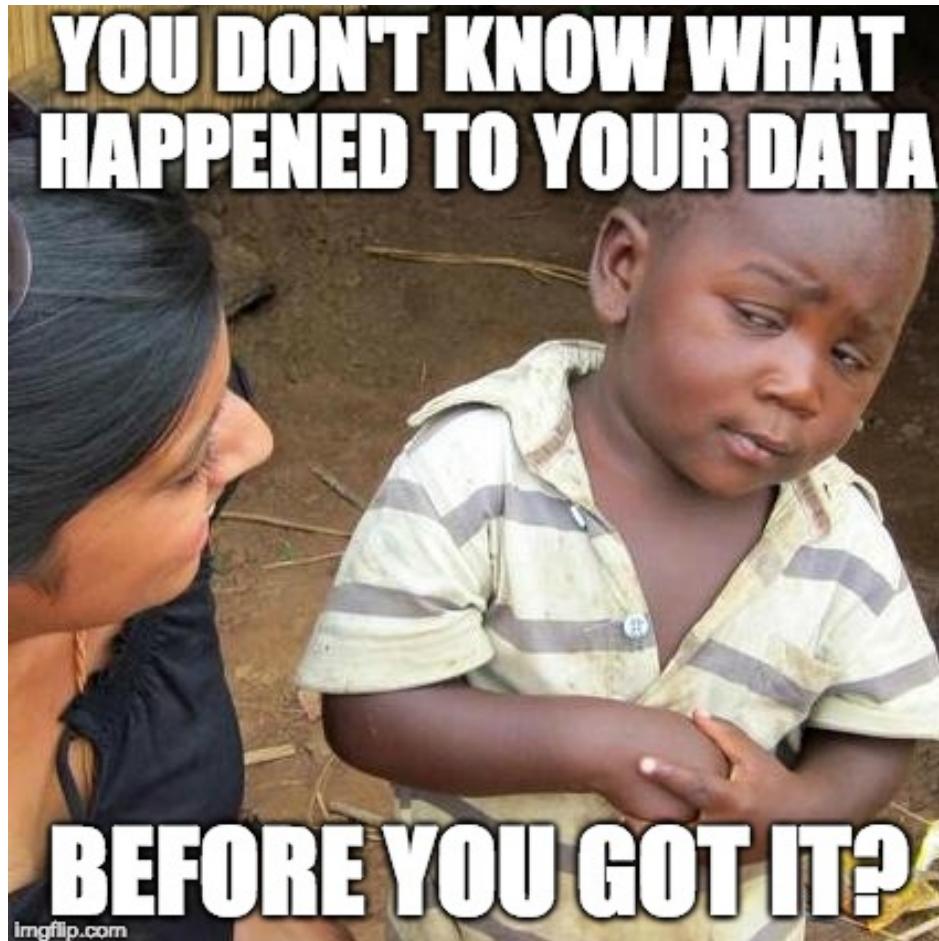
Hosted on: [grand-challenge.org](#)

“Pre”-processing is Important



<https://imgflip.com/i/2blh8f>

“Pre”-processing is Hard



<https://imgflip.com/i/2blh8f>



<https://imgflip.com/i/2bhsc>

Package Showcase

Data

- EveTemplate/MNI Template - templates
- kirby21 series
- sri24 - SRI24 MRI Atlas
- malf.templates - templates for label fusion
- neurohcp - download data from HCP/INDI
- nitrcbot - download data from NITRC
- neurovault - download data from
<https://neurovault.org/>

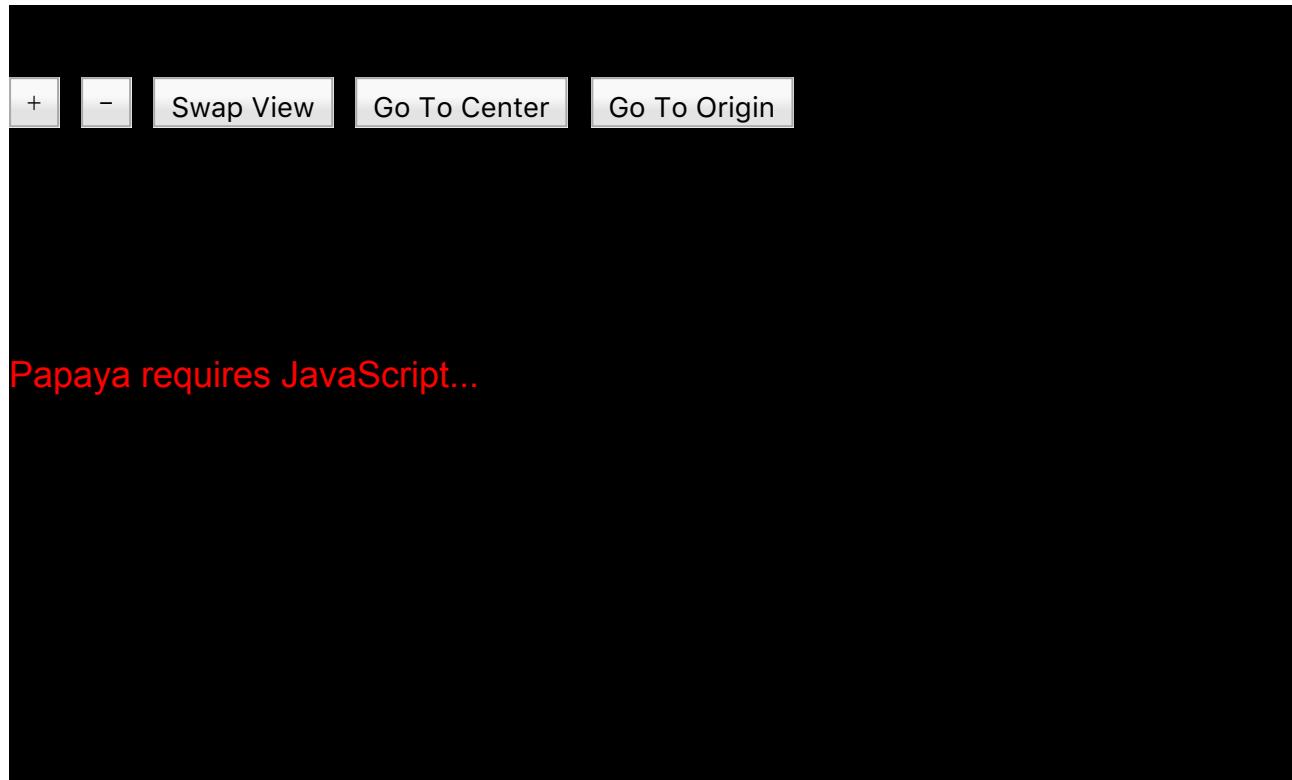
General Imaging Tools

- EBImage - image processing and analysis
- magick - Bindings to ImageMagick
- OpenCV: [swarm-lab/ROpenCVLite](#) or
[ropenscilabs/opencv](#)
- ANTsR - a large imaging suite (mainly neuro)

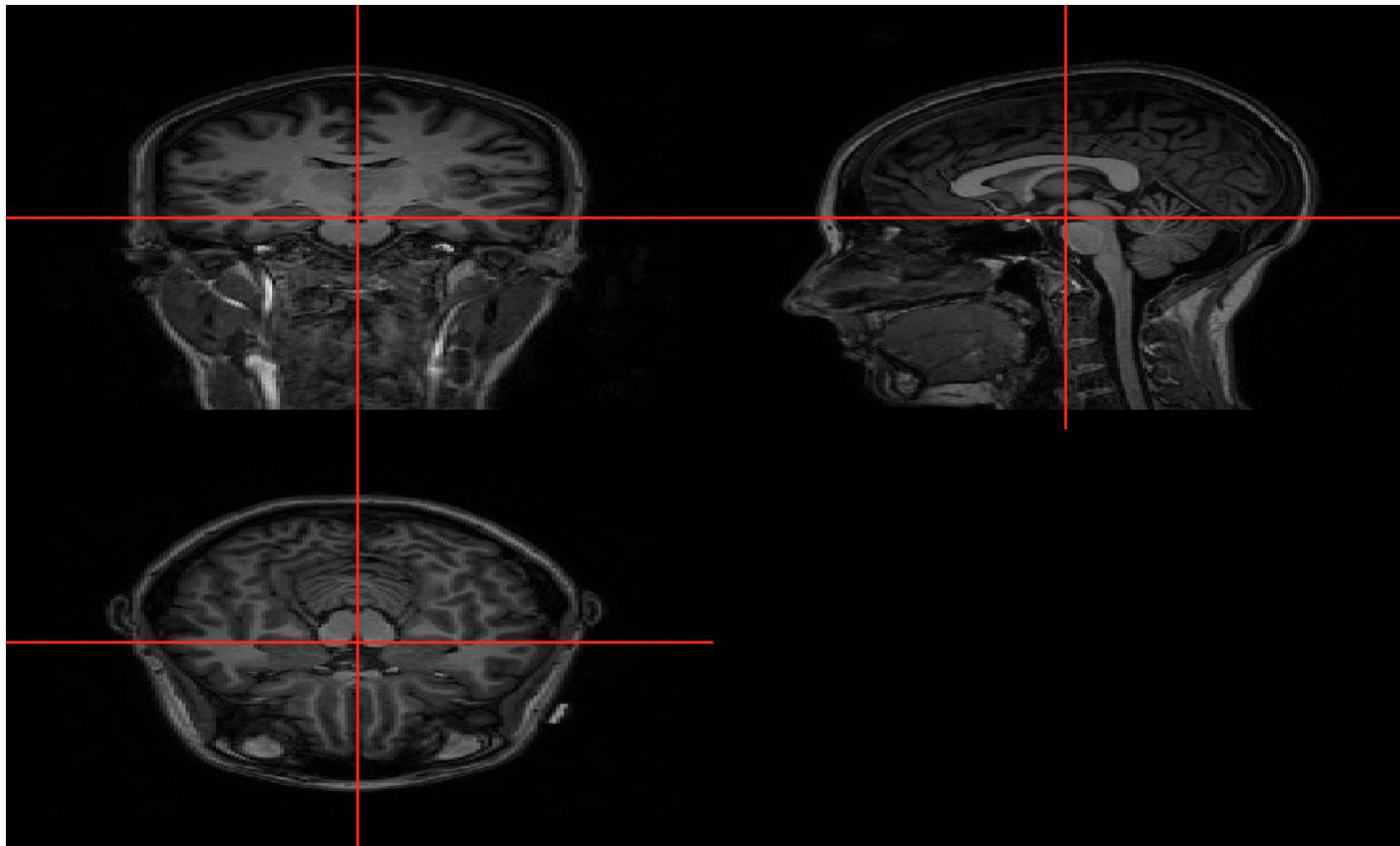
Starting from Raw Data/DICOM

- `oro.dicom` - read/write DICOM data
- `dcm2niir` - uses `dcm2niix` from Chris Rorden
- `divest` - Rcpp wrapped `dcm2niix`
- `dcmtk` - interface package for DCMTK
- `matlabr` - could use `dicomread` MATLAB code and execute through R

Interactive Visualization using papayaWidget



ggneuro visualization: ggplot2 object



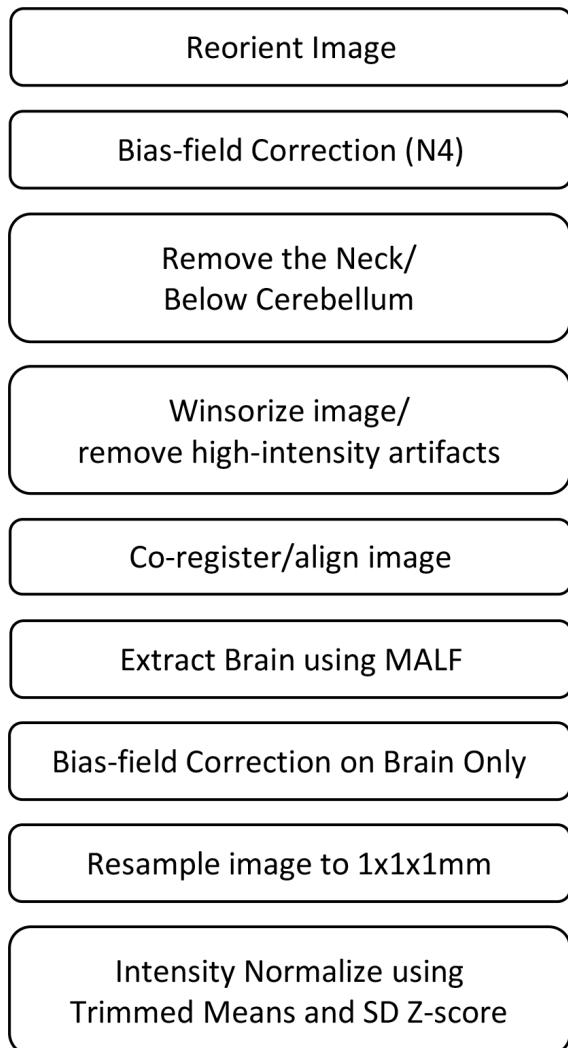
Intensity Normalization

- whitestripe - WhiteStripe (Shinohara et al. 2014)
- Quantile transform `quantile_img` (in neurobase)
- Whole brain z-scoring: `zscore_img` in neurobase
- RAVEL - Fortin et al. (2016)
- Histogram matching (in RAVEL)
- General standardization methods

MS Lesion Segmentation

- `sublime` - E. Sweeney et al. (2013)
- `oasis` - Sweeney et al. (2013)
- `mimosa` - Valcarcel et al. (2018)
- `smri.process` - my package on sMRI processing

Step 1: Image Processing: Workflow

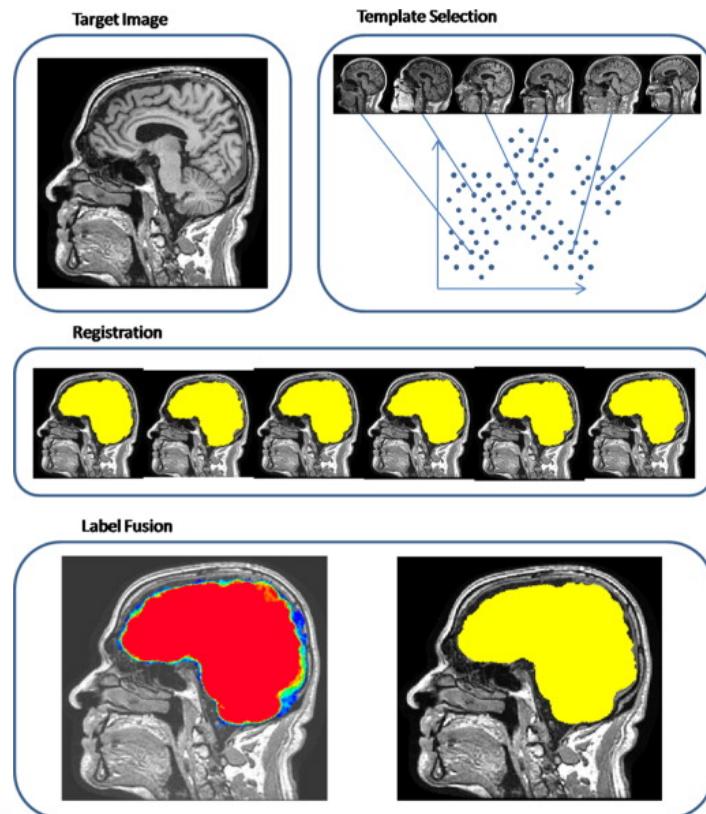


The N4 (Tustison et al. 2010)
EM-style model assumed is:
$$\log(x(v)) = \log(u(v)) + \log(f(v))$$

- x : given image
- u : uncorrupted image
- f : bias field
- v : location in the image

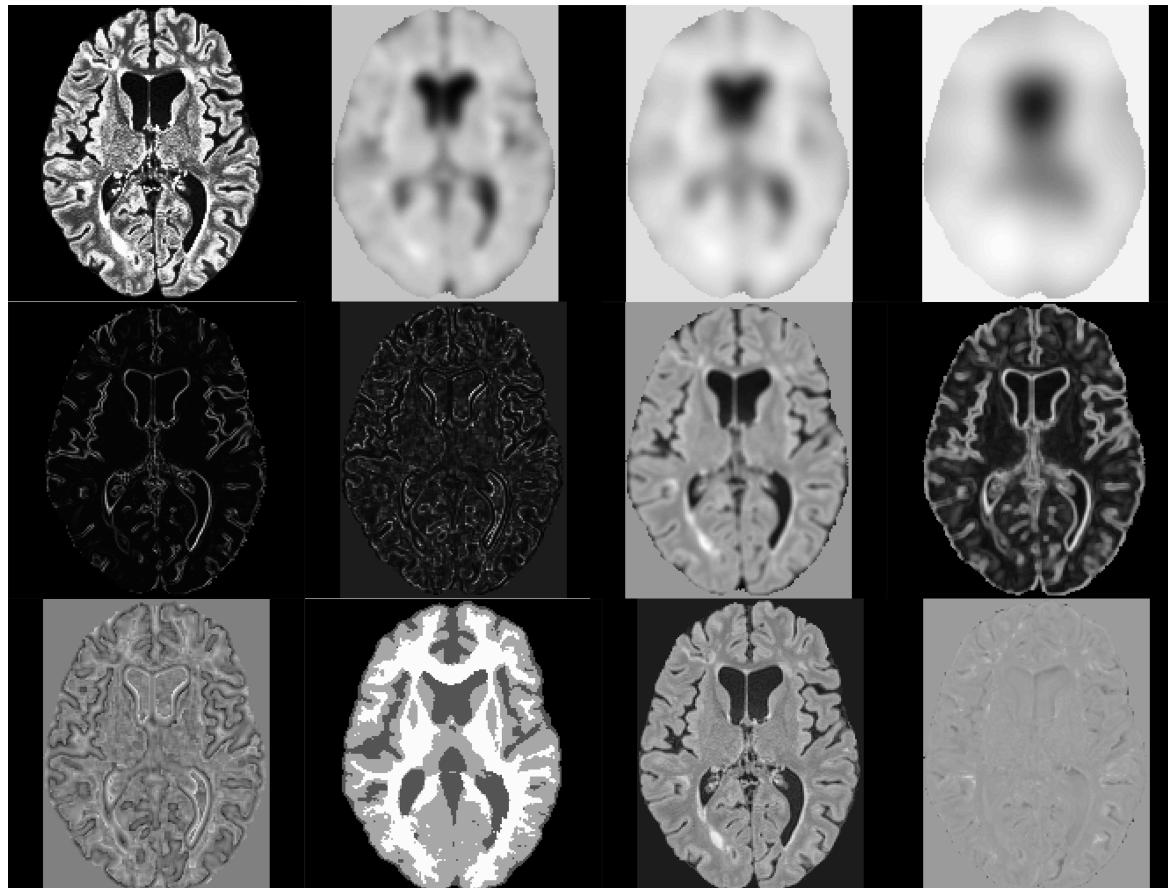
Image Processing: MALF

Figure from MASS paper (Doshi et al. 2013):



- Register templates to an image using the T1 for that subject
- Apply transformation to the label/mask
- Average each voxel over all templates
 - there are “smarter” (e.g. weighted) ways

Step 2: Create Predictors for each Sequence



- Predictors created with intensity-normalized data
 - Quantile images, smoothers, local moments
- Tissue class probability with local moments: MALF and FAST (Zhang, Brady, and Smith 2001)
- Z-score to a population template

A package to do all this: `smri.process`

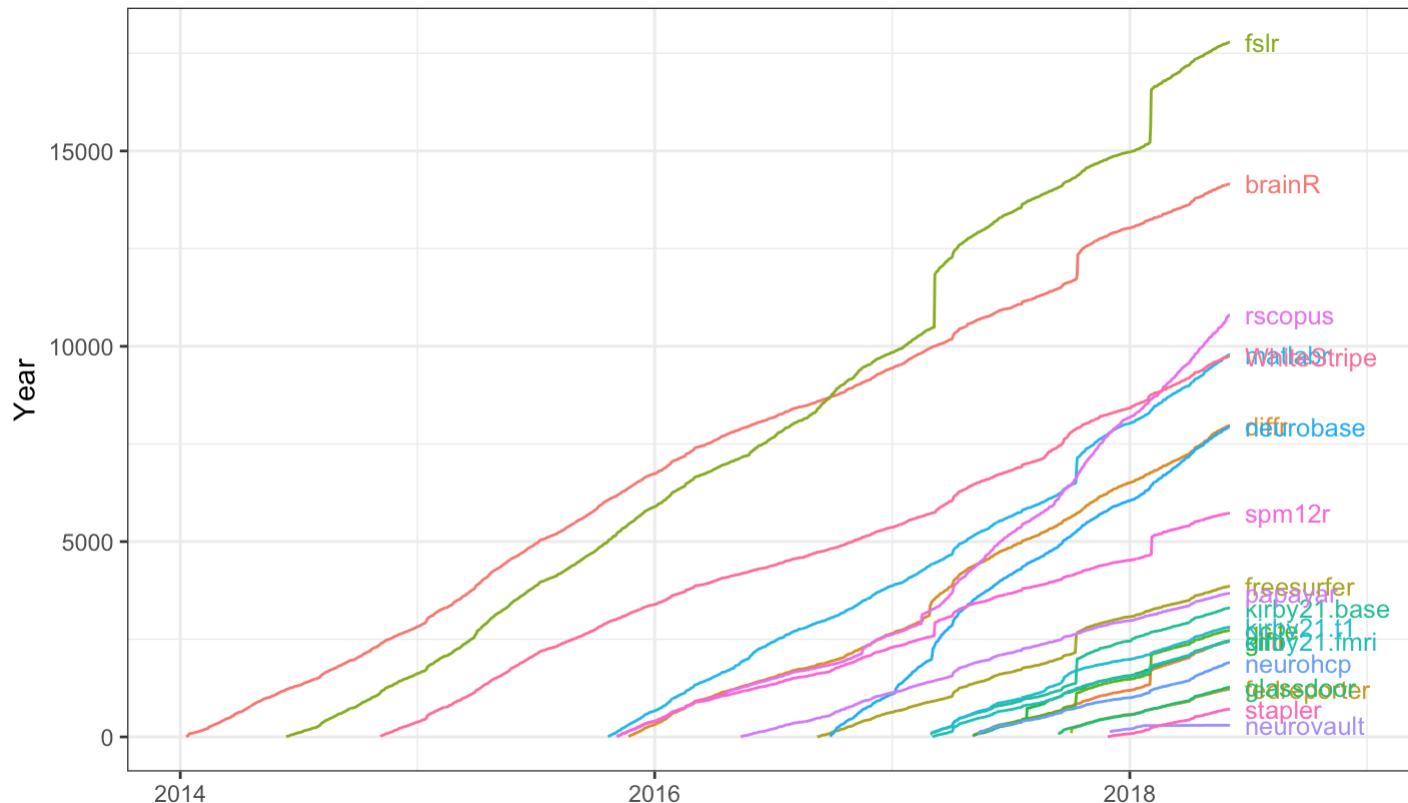
- GitHub package ([muschelliij2/smri.process](https://github.com/muschelliij2/smri.process))

```
65  processed = smri_prenormalize(  
66    x = files,  
67    outdir = idf$proc_dir,  
68    gold_standard = gold_standard,  
69    gs_space = "FLAIR",  
70    reg_space = "FLAIR",  
71    malf_transform = "SyN",  
72    verbose = 2,  
73    outprefix = outprefix,  
74    probs = c(0, 0.995),  
75    num_templates = 15,  
76    force_registration = FALSE)  
77  
78  outprefix = file.path(  
79    idf$malf_dir,  
80    "T1")  
81  
82  all_resampled = seg_normalize(  
83    prenormalize = processed,  
84    template = "none",  
85    verbose = TRUE,  
86    force_registration = FALSE,  
87    outprefix = outprefix  
88  )  
89  normalized = all_resampled$normalized  
90  
91  pred = norm_predictors(  
92    normalized = normalized,  
93    normalization = "trimmed_z"  
94  )
```

Publishing software

- R Journal - see `rticles::rjournal_article()`
(only accept if in CRAN)
- F1000 -
<https://f1000research.com/collections/Neuroconductor>
- JOSS - <https://joss.theoj.org/>
- JSS - see `rticles::jss_article()`
- cranlogs track downloads
- neuroconductor API tracks downloads

Tracking downloads



Conclusions

- Many methods are being developed for processing neuroimaging in R
- Analysis tools exist but need adaptation
- Develop more standardization like BioConductor
 - standard data structures
- GitHub and Neuroconductor

Where else can statisticians lead?

Do we need spatial results? If not,
why register to the template?

Registration

ANTsR/extrantsr

- antsRegistration - rigid/affine/non-linear diffeomorphic
- extrantsr::registration - wraps antsRegistration to use nifti objects

fslr

- flirt - linear/affine registration
- fnirt - non-linear registration (need affine first)
- fnirt_with_affine - wraps above 2

Bibliography

- Doshi, Jimit, Guray Erus, Yangming Ou, Bilwaj Gaonkar, and Christos Davatzikos. 2013. "Multi-Atlas Skull-Stripping." 20 (12). Elsevier:1566–76.
- Fortin, Jean-Philippe, Elizabeth M Sweeney, John Muschelli, Ciprian M Crainiceanu, Russell T Shinohara, Alzheimer's Disease Neuroimaging Initiative, and others. 2016. "Removing Inter-Subject Technical Variability in Magnetic Resonance Imaging Studies." 132. Elsevier:198–212.
- Shinohara, Russell T., Elizabeth M. Sweeney, Jeff Goldsmith, Navid Shiee, Farrah J. Mateen, Peter A. Calabresi, Samson Jarso, Dzung L. Pham, Daniel S. Reich, and Ciprian M. Crainiceanu. 2014. "Statistical Normalization Techniques for Magnetic Resonance Imaging." 6:9–19. <https://doi.org/10.1016/j.nicl.2014.08.008>.
- Sweeney, Elizabeth M, Russell T Shinohara, Navid Shiee, Farrah J Mateen, Avni A Chudgar, Jennifer L Cuzzocreo, Peter A Calabresi, Dzung L Pham, Daniel S Reich, and Ciprian M Crainiceanu. 2013. "OASIS Is Automated Statistical Inference for Segmentation, with Applications to Multiple Sclerosis Lesion Segmentation in MRI." 2. Elsevier:402–13.
- Sweeney, EM, RT Shinohara, CD Shea, DS Reich, and CM Crainiceanu. 2013. "Automatic Lesion Incidence Estimation and Detection in Multiple Sclerosis Using Multisequence Longitudinal MRI." 34 (1). Am Soc Neuroradiology:68–73.
- Tustison, Nicholas J., Brian B. Avants, Philip A. Cook, Yuanjie Zheng, Alexander Egan, Paul A. Yushkevich, and James C. Gee. 2010. "N4ITK: Improved N3 Bias Correction." 29 (6):1310–20.
<https://doi.org/10.1109/TMI.2010.2046908>.
- Valcarcel, Alessandra M, Kristin A Linn, Simon N Vandekar, Theodore D Satterthwaite, John Muschelli, Peter A Calabresi, Dzung L Pham, Melissa Lynne Martin, and Russell T Shinohara. 2018. "MIMoSA: An Automated Method for Intermodal Segmentation Analysis of Multiple Sclerosis Brain Lesions." Wiley Online Library.
- Zhang, Yongyue, Michael Brady, and Stephen Smith. 2001. "Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm." 20 (1):45–57.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=906424.