

Neuroconductor: Medical Image Analysis in R

John Muschelli
http://johnmuschelli.com/Pitt_2017.html
Johns Hopkins Bloomberg School of Public Health

What language do you use for analysis?

- R
- SAS / Stata
- Matlab
- Python

What language do you use for imaging analysis?

- R
- Matlab (e.g. SPM)
- Python
- bash (e.g. FSL)

(note SAS/Stata do not exist in this domain)

Why R?

- Free, open-source
- Cross-platform
- Over 10000 packages
- Taught by stat/biostat departments
- Packaging system is good (well-documented, tested, doable)
- RStudio
- (I know R)

Inspiration



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

Inspiration



- team of developers/maintainers
- multiple grants of support

Inspiration



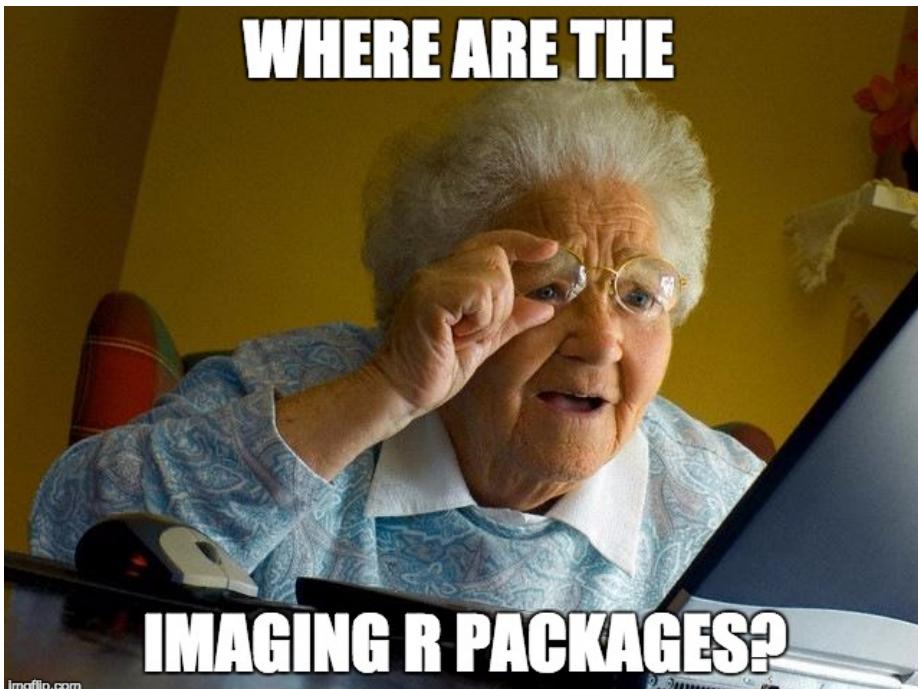
Nipype:
Neuroimaging in Python
Pipelines and Interfaces

- integrates multiple neuroimaging software
- standardizes the syntax
- provides workflows
- allows the user to take advantage of all of Python
 - pandas, scikit-learn, Jupyter notebooks

Inspiration: SPM/MATLAB has GUI + courses!

The image shows two windows of the SPM12 software. The left window is titled "SPM12 (6225): Menu" and contains buttons for "Real...", "Slice timing", "Smooth", "Cor...", "Nor...", "Segment", "Specify 1st-level", "Review", "Specify 2nd-level", "Estimate", and "Results". Below these are buttons for "Dynamic Causal Modelling" and "SPM for functional MRI". The right window is titled "SPM12 (6225): Graphics" and shows a "Welcome to SPM12" message. It includes instructions to refer to the "SPM12" version in papers and communications, links to the "Manual" and "Release Notes", information about "Updates" and the "SPM mailing list", and a contact email address "fil.spm@ucl.ac.uk". A footer note states "SPM is developed under the auspices of the Functional".

What did R have?



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

What did R have?

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,



An R Platform for
Medical Imaging Analysis

What really is Neuroconductor?

1. A community of developers and users of R packages for imaging
2. A website <https://neuroconductor.org/>.
 - with tutorials and help
3. A team helping developers and users (John, Adi Gherman, Ciprian Crainiceanu, Brian Caffo)
4. A centralized repository of maintained packages

Goal: Centralize the packages (currently 44)

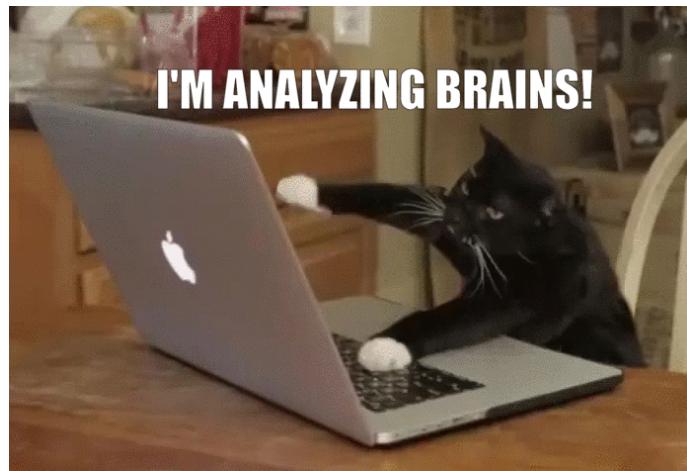
List Packages

List Packages					
View Dependency Graph			View Pending Packages		
Show 50 entries			Search:		
Package Name	Version	Package Title	Maintainer(s)		
Last updated	GitHub URL				
ANTsR	0.4.0	ANTs in R: quantification tools for biomedical images	Brian B. Avants	2017-03-18	stnava/ANTsR
ANTsRCore	0.0.0	ANTsRCore: core software infrastructure for ANTsR	Brian B. Avants	2017-03-18	stnava/ANTsRCore.git
brainR	1.4.2.1	Helper Functions to Misc3d and rgl Packages for Brain Imaging	John Muschelli	2017-05-26	muscelli/j2/brainR
cifti	0.4.2	Toolbox for Connectivity Informatics Technology Initiative ('CIFTI') Files	John Muschelli	2017-05-26	muscelli/j2/cifti
dcemriS4	0.57.1.2	A Package for Image Analysis of DCE-MRI (S4 Implementation)	Brandon, Whitcher	2017-05-26	bjw34032/dcemriS4
dcm2nir	0.5	Conversion of 'DICOM' to 'NifTI' Imaging Files Through R	John Muschelli	2017-02-24	muscelli/j2/dcm2nir
divest	0.3.0.1	Get Images Out of DICOM Format Quickly	Jon Clayden	2017-05-25	jonclayden/divest
EveTemplate	0.99.14.2	JHU-MNI-ss (Eve) template	Jean-Philippe Fortin	2017-05-26	JFortin1/EveTemplate
extrantsr	2.17.2.3	Extra Functions to Build on the ANTsR Package	John Muschelli	2017-05-26	muscelli/j2/extrantsr.git
freesurfer	1.6.6	Wrapper Functions for 'Freesurfer'	John Muschelli	2017-05-26	muscelli/j2/freesurfer
fslr	2.12.6	Wrapper Functions for FSL ('FMRIB' Software Library) from Functional MRI of the Brain ('FMRIB')	John Muschelli	2017-05-26	muscelli/j2/fslr
gifti	0.7	Reads in Neuroimaging 'GIFTI' Files with Geometry Information	John Muschelli	2016-11-09	muscelli/j2/gifti
ITKR	0.0.1	ITK in R	Brian B. Avants	2017-02-24	stnava/ITKR
itksnapr	2.1.6	Package of ITK-SNAP	John Muschelli	2017-05-26	muscelli/j2/itksnapr
kirby21.asl	1.5.1	Example ASL Data from the Multi-Modal MRI Reproducibility Resource	John Muschelli	2017-05-03	muscelli/j2/kirby21.asl

Neuroconductor

Goal:

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

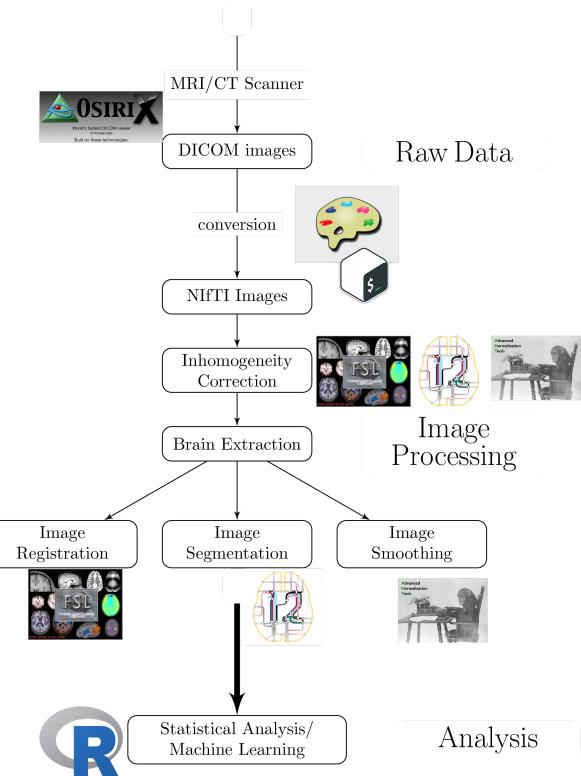


From <http://i.imgur.com/0Y1xISa.gifv>.

- [http://johnmusc
helli.com/neuroc
/](http://johnmusc
helli.com/neuroc
/)

Workflow for an Analysis

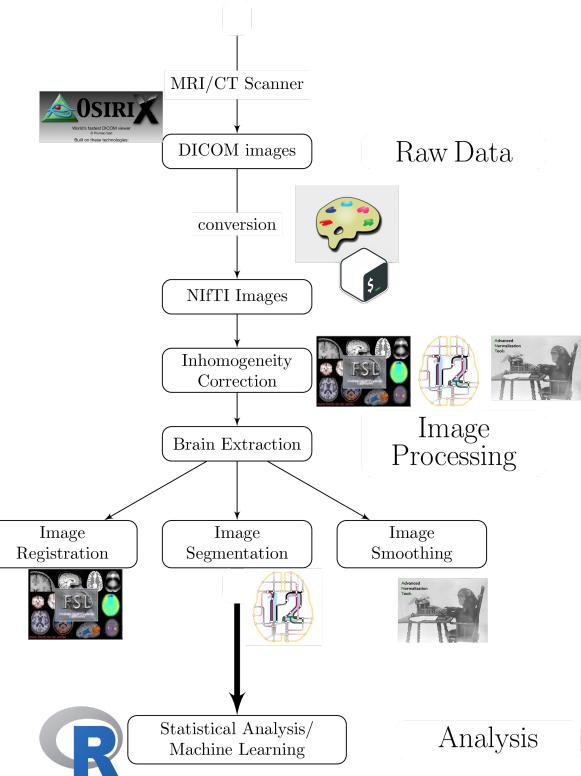
- bash 
- FSL 
- ANTs 
- MRICroGL 
- OsiriX 
- SPM 12 



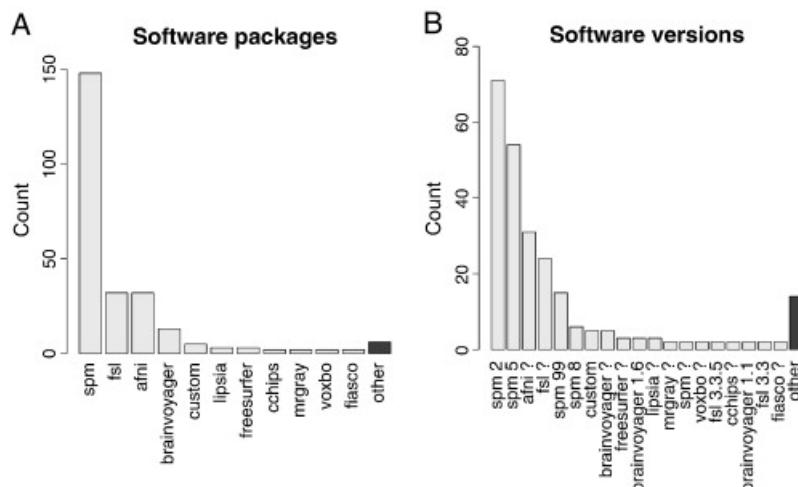
Workflow for an Analysis

Multiple pieces of software used

- all different syntax



It's typical to have lots of software choices



Carp, Joshua. "The secret lives of experiments: methods reporting in the fMRI literature." Neuroimage 63.1 (2012): 289-300.

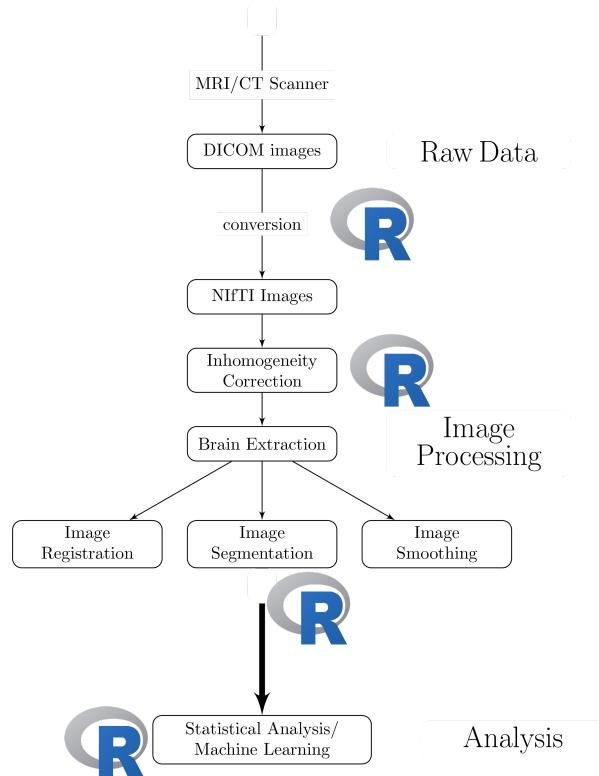
Goal:

Lower the bar to entry

- all R code
 - pipeline tool
 - "native" R code

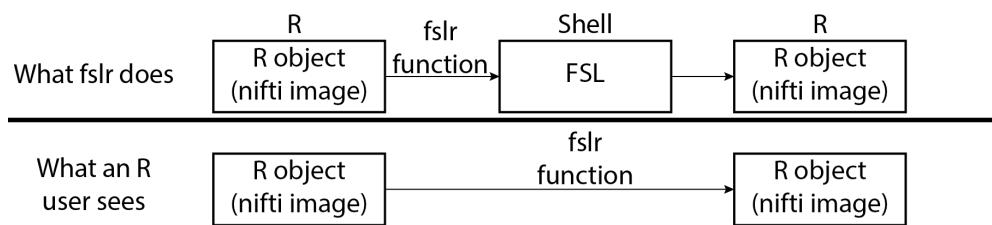
Complete pipeline

- preprocessing and analysis



Using R as a Pipeline Tool: fslr

- `fslr` package - call FSL from R
- Requires FSL to be installed and in PATH
 - FSL only available on Unix-style systems



Why not CRAN: The story of ANTsR

ANTs: Advanced Normalization Tools

- State-of-the-art image processing pipelines
- Built at UPenn under Brian Avants
 - Group has won challenges for imaging analysis
- Still actively maintained and developed
- Depends on the Insight ToolKit (ITK) medical image processing library

The story of ANTsR: Porting ANTs to R

- Group tried to wrap C++ code with Rcpp for it to work "seamlessly" with R
- [Had some issues with CRAN](#)
- Most developers response: "just make it work"
- **NOT** like `fslr` - this built the libraries and could be called using `.Call`
 - possible (not yet) to use on Windows

Why not CRAN: Dependencies

- ANTsR depends on CMake:  **CMake**
- ANTsR also depends on ITKR - large code base
- External dependencies are not always welcome
- CRAN response:  **CM~~X~~ke** (more or less)

Why not CRAN: Checks

- We would like more stringent checks than CRAN (like vignettes) like Bioconductor
- But **also** more lenient ones
- Examples may take > 5 seconds to compile
- Time to build package may take a while
- May not work on Windows if not applicable

Why not CRAN: Data

Medical images are big (> 5 mb) compared to code

CRAN doesn't like this

- Example data packages are harder to get passed
- We got `kirby21.t1` and `kirby21.fmri`, but they have an implicit downloader function in the code
- Necessary for testing code examples in imaging
- **STANDARDIZED** file location (`system.file`)

**Solution: Build a Bioconductor-like
Solution using Current Tools**

Git and GitHub

- Git version control system: stores changes of files



- GitHub is an **online** server of repositories
- Distribute packages and install them via
`devtools::install_github`



Continuous Integration: Travis and Appveyor

- Builds and checks R packages on Windows (Appveyor) and Linux/OS X (Travis CI)
- Works well with GitHub



Travis CI

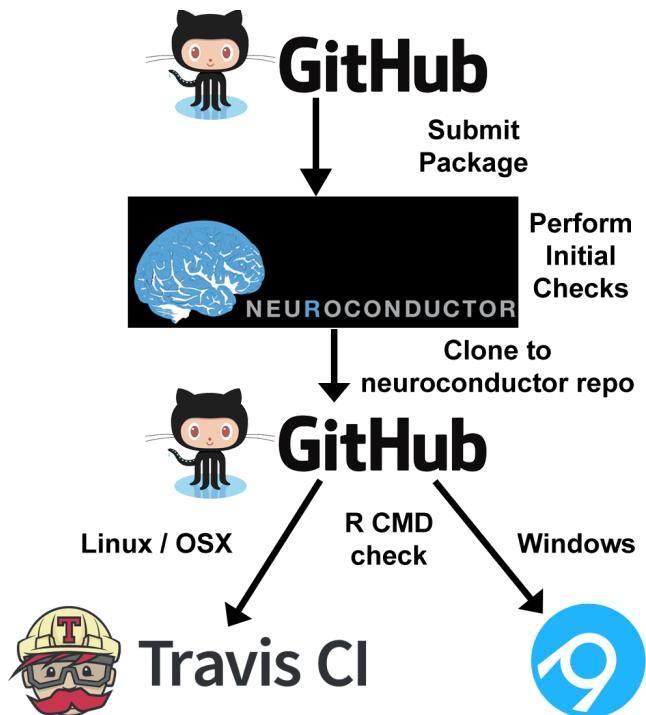


Development

Pipeline:

Check the package
for stability

- check against other imaging software (e.g. FSL)



ANTsR Revisited: Helping Developers

- GitHub allows the Neuroconductor team to help fix issues
- Standardized ANTsR checking for NeuroC
- Pull Requests to Group ICA package (Brian Caffo)
- Remove unnecessary hurdles for developers

Neuroconductor Capabilities

Capabilities	Packages
DICOM Images	oro.dicom, dcm2niir, divest, ANTsR
NIfTI Images	oro.nifti, RNifti, ANTsR
Image Registration	spm12r, fslr, ANTsR, freesurfer
Inhomogeneity Correction	spm12r, fslr, ANTsR
Brain Extraction	spm12r, fslr, ANTsR, extractsr
Structure Segmentation	spm12r, fslr, ANTsR, extractsr, freesurfer
Intensity Normalization	WhiteStripe, neurobase, ANTsR
3D Smoothing	ANTsR, spm12r, fslr
Temporal Filtering	spm12r, fslr, ANTsR
Slice-timing correction	spm12r, fslr
DTI models	rcamino, oro.dti, fslr

Benefits of Neuroconductor:

Allow imaging to use all R has to offer:

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



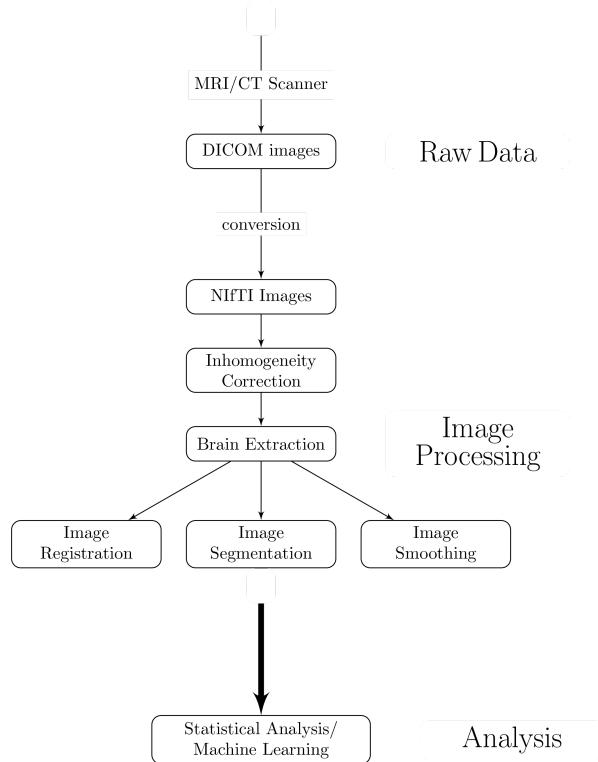
Neuroconductor Downsides

1. More control over the workflow = more work
2. Users need external software (versions/installation)
3. No control over external software
 - if maintainer changes something, not much recourse
4. Need the content (buy-in from the community)

Potential Neuroconductor Downsides:

Enabling statisticians
to do preprocessing
also **enables** imagers
to do advanced
statistics.

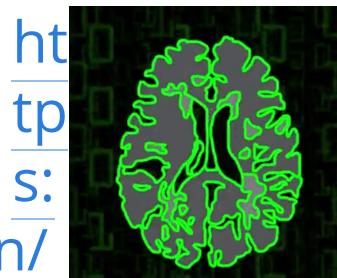
- Why enlist us for
only this part?



Training we are providing

Coursera Course:
Introduction to
Neurohacking In R

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



Neuroimaging analysis within R (ISBI, Melbourne, April 2017)

ENAR 2018

**Without Knowing the Processing,
Analyses Cannot be Trusted**

No Code = No Method

Current Neuroconductor Packages Overviews

dcm2niir and divest: Converting DICOM data

- `dcm2nii` by Dr. Chris Rorden
 - jackknife for almost any DICOM type to NIfTI
- `dcm2niir` wraps a binary executable of `dcm2nii`
 - calls the command prompt (not always Windows-friendly)
- `divest` wraps the background C++ code of `dcm2nii`
 - runs in "native" R

neurohcp: Human Connectome Project

- Allows you to download data from [Human Connectome Project](#)
- The 1200 Subjects release: behavioral and 3T MR imaging data from 1206 healthy young adult participants. Standardized protocol.
- Tutorial:
<http://johnmuschelli.com/neuroc/neurohcp>

spm12r: Wrapper Functions for SPM

- Wraps some MATLAB code to call SPM scripts
- Built from SPM batch commands
- spm12_slice_timing, spm12_realign
- spm12_coregister, spm12_segment
- spm12_normalize, spm12_smooth
- http://johnmuschelli.com/neuroc/fMRI_analysis_spm12r/index.html

rcamino: Port of Camino Software

- Wraps [Camino Diffusion MRI Toolkit](#)
- Takes in b-values, b-vectors, and tensors
- Fits models for DTI data
- http://johnmuschelli.com/neuroc/DTI_analysis_rcamino/index.html

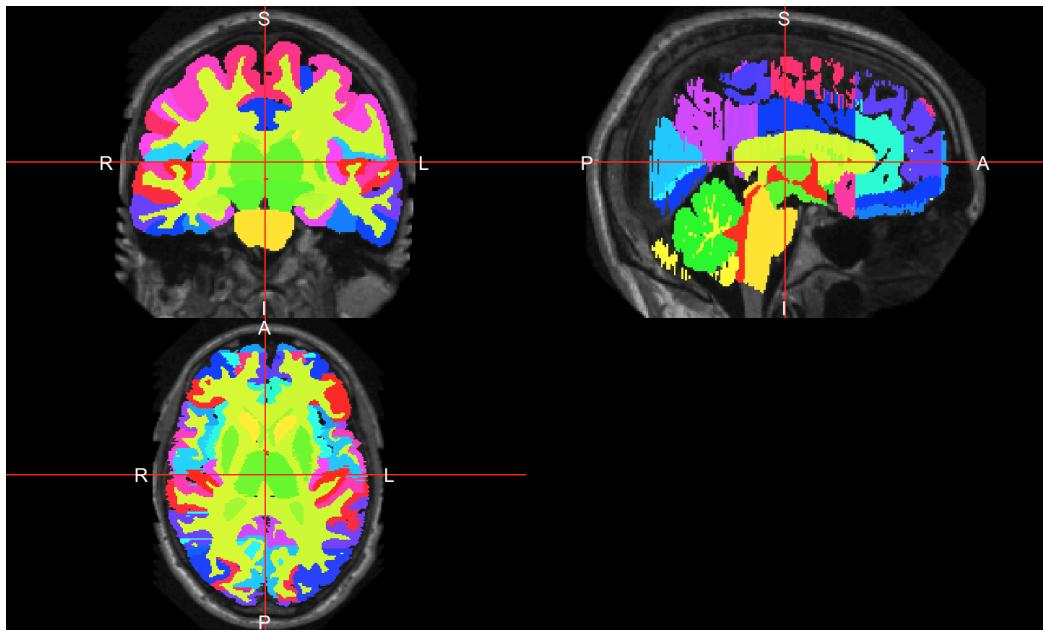
extrantsr: Extra functions for ANTsR

- Wraps ANTsR functions for nifti objects (from `oro.nifti`)
- `malf` function - performs a basic MALF algorithm
- `within_visit_registration` - performs within-visit registration for structural MRI
- `preprocess_mri_within` - pipeline to process within-visit sMRI

malf.templates: Segmented T1-weighted Images

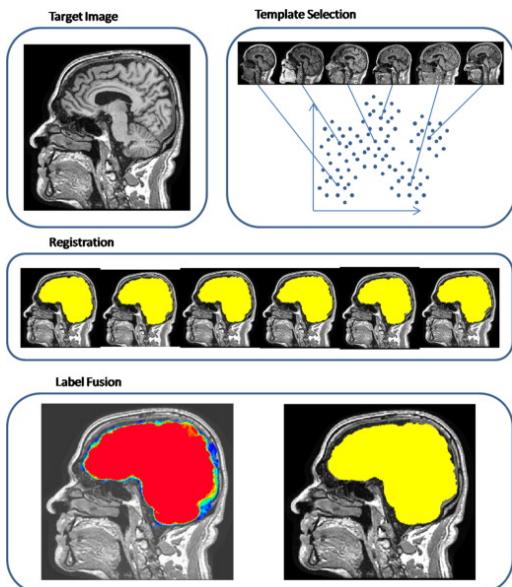
- Data from the MICCAI 2012 Challenge on Multi-atlas Labelling Data
- From OASIS project and the labeled data as provided by Neuromorphometrics, Inc. (<http://Neuromorphometrics.com/>)
- Have T1-weighted image, brain mask, image with mask applied, substructure segmentation
 - Hand segmented structures - useful for multi-atlas label fusion (MALF)

malf.templates: Segmented T1-weighted Images



MALF: Skull Stripping Example

From (Doshi et al. 2013):



- Register templates to an subject T1
- Apply transformation to the label/mask, average over voxels
 - there are "smarter" (e.g. weighted) ways

Neuroconductor installer:

Provide an installer to download the current packages:

```
source("https://neuroconductor.org/neurocLite.R")
neurocLite("kirby21.fmri")
neurocLite("neurobase")
neurocLite("ANTSr")
```

kirby21.fmri: fMRI data from Kirby21

```
library(kirby21.t1) # load T1 data
library(kirby21.fmri) # load fMRI data
library(kirby21.base) # helper package to download data
library(neurobase) # neuroconductor base package
fnames = get_image_filenames_list(
  modalities = c("fMRI", "T1"), id = 113, visit = 1)
print(fnames)
t1_fname = fnames$T1
fmri_fname = fnames$fMRI
```

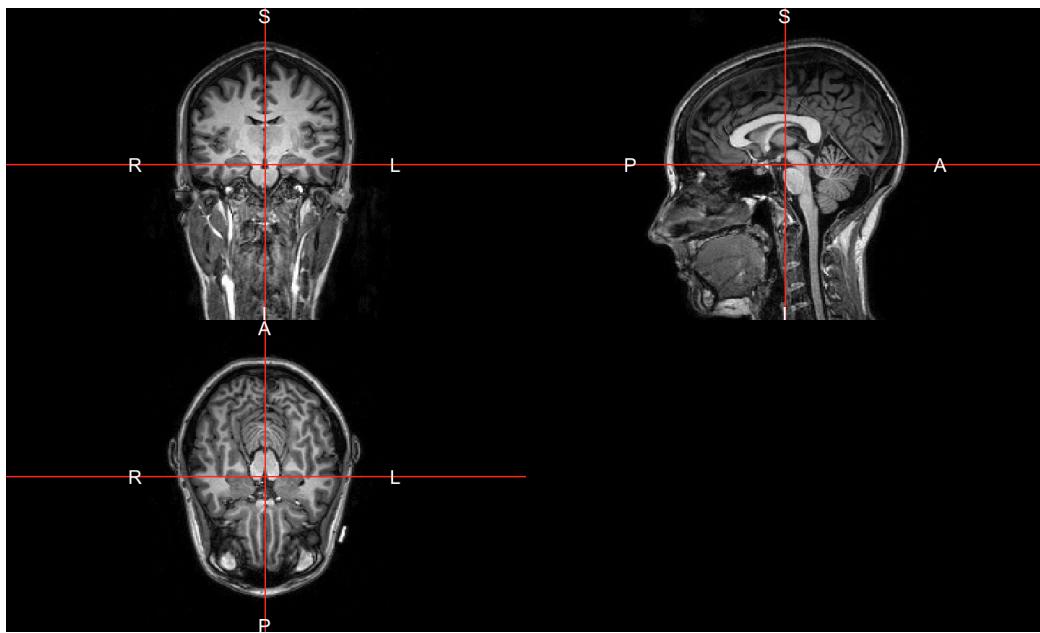
fMRI

```
"library/kirby21.fmri/visit_1/113/113-01-fMRI.nii.gz"
```

T1

```
"library/kirby21.t1/visit_1/113/113-01-T1.nii.gz"
```

Plot the T1



Future work

- F1000 R packages channel - get publication for workflow
- Removing duplicate packages from CRAN - like Bioconductor
- Docker image for Neuroconductor:
 - have all dependencies/external software in one container

Go to

http://johnmuschelli.com/imaging_in_r/index.html

Example MPRAGE T1 image can be found at
<http://bit.ly/pittimg>

Download the .R file <http://bit.ly/pittscript>

fMRI in Shiny <http://bit.ly/shinybrain>

Papaya Example <http://bit.ly/pittpapaya>

References

Doshi, Jimit, Guray Erus, Yangming Ou, Bilwaj Gaonkar, and Christos Davatzikos. 2013. "Multi-Atlas Skull-Stripping." 20 (12). Elsevier: 1566–76.