Intro to the
Human Connectome Project

Greg Burgess, Ph.D.

Staff Scientist

https://wiki.humanconnectome.org/download/attachments/29130878/HCP_intro_brainhack.pdf

Overview of today's talk

- What was the goal of the HCP?

- What data is available?

- How can data be accessed?

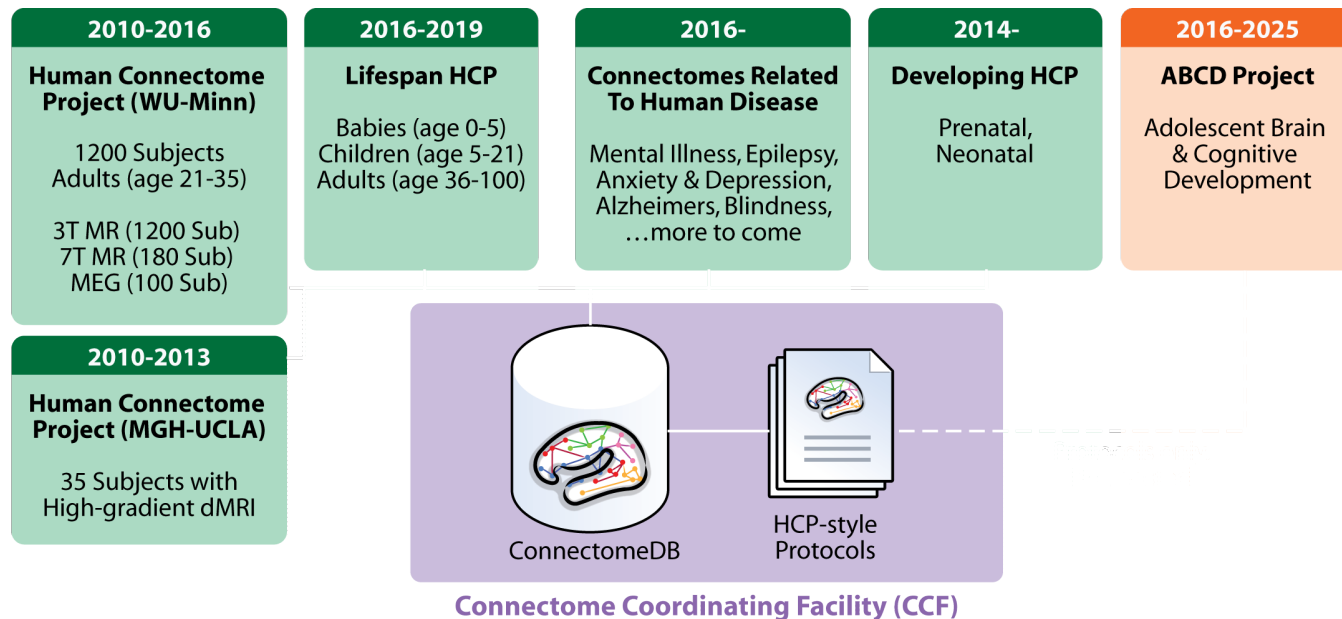- What tools exist for analysis?

- How can I learn more?

What was the goal of the HCP?

- Produce comprehensive map of brain connectivity in healthy adults (macro-level)
  - Leverage recent technical advances to collect highest quality data possible
- Understand factors related to individual differences in brain connectivity
  - Link to heritability, genetics
  - Sub-clinical psychological symptoms

What was the goal of the HCP?

- Provide baseline for future projects
  - Connectomes of disease: Alzheimer's Disease, Depression, Psychosis, Visual Impairment
  - Connectomes across the lifespan: Development, Aging

2010-2016	2016-2019	2016-	2014-	2016-2025
Human Connectome Project (WU-Minn)	**Lifespan HCP**	**Connectomes Related To Human Disease**	**Developing HCP**	**ABCD Project**
1200 Subjects Adults (age 21-35)	Babies (age 0-5) Children (age 5-21) Adults (age 36-100)	Mental Illness, Epilepsy, Anxiety & Depression, Alzheimers, Blindness, …more to come	Prenatal, Neonatal	Adolescent Brain & Cognitive Development
3T MR (1200 Sub) 7T MR (180 Sub) MEG (100 Sub)				

2010-2013

Human Connectome Project (MGH-UCLA)

35 Subjects with High-gradient dMRI

ConnectomeDB

HCP-style Protocols

Connectome Coordinating Facility (CCF)

What was the goal of the HCP?

- Advance more-sophisticated neuroimaging analysis techniques
  - Advanced distortion correction
  - Surface-based registration and analysis
  - MSM: Multimodal Surface Matching
  - FIX: Automated ICA denoising (resting state fMRI)
  - EDDY: motion correction with outlier detection (diffusion MRI)

What was the goal of the HCP?

- Advance more-sophisticated statistical techniques
  - Permutation Analysis of Linear Models (PALM):
    - https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/PALM
    - Account for family relationships in statistical testing
    - TFCE cluster-correction for multiple comparisons in HCP data
  - MegaTrawl:
    - Compute individual, subject-level "netmats" using partial correlation with L2 regularization
    - Predict each behavioral measure using variance in netmat edges across participants
  - Parcellation
    - Determine subject-specific, multi-modal cortical parcellation using machine-learning classifier

What was the goal of the HCP?

- Make the data freely available to the public!

What data is available?

- Recruited 1200 healthy adult participants
  - Heritability: Recruited from families containing twins and non-twin siblings
- Genotyping (DNA)
  - From blood or saliva
  - Available soon in dbGaP

What data is available?

- Extensive behavioral testing
  - NIH Toolbox
    - Cognition, Emotion, Motor, Sensory
      - https://www.neuroscienceblueprint.nih.gov/factSheet/toolbox.htm
  - Computerized Neuropsychological Testing (Penn CNP)
    - Personality, Psychiatric, and additional cognitive and emotion measures
      - https://penncnp.med.upenn.edu
  - Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA)
    - Psychiatric / Substance use questionnaires
  - Additional measures:
    - Demographics, health measures, sleep, drug screening, etc.
  - https://wiki.humanconnectome.org/display/PublicData/HCP+Data+Dictionary+Public-+500+Subject+Release

What data is available?

- Multimodal neuroimaging data
  - Structural images: 3T (WUSTL, 1200 subjects)
    - T1-weighted, T2-weighted: 0.7 mm isotropic voxels
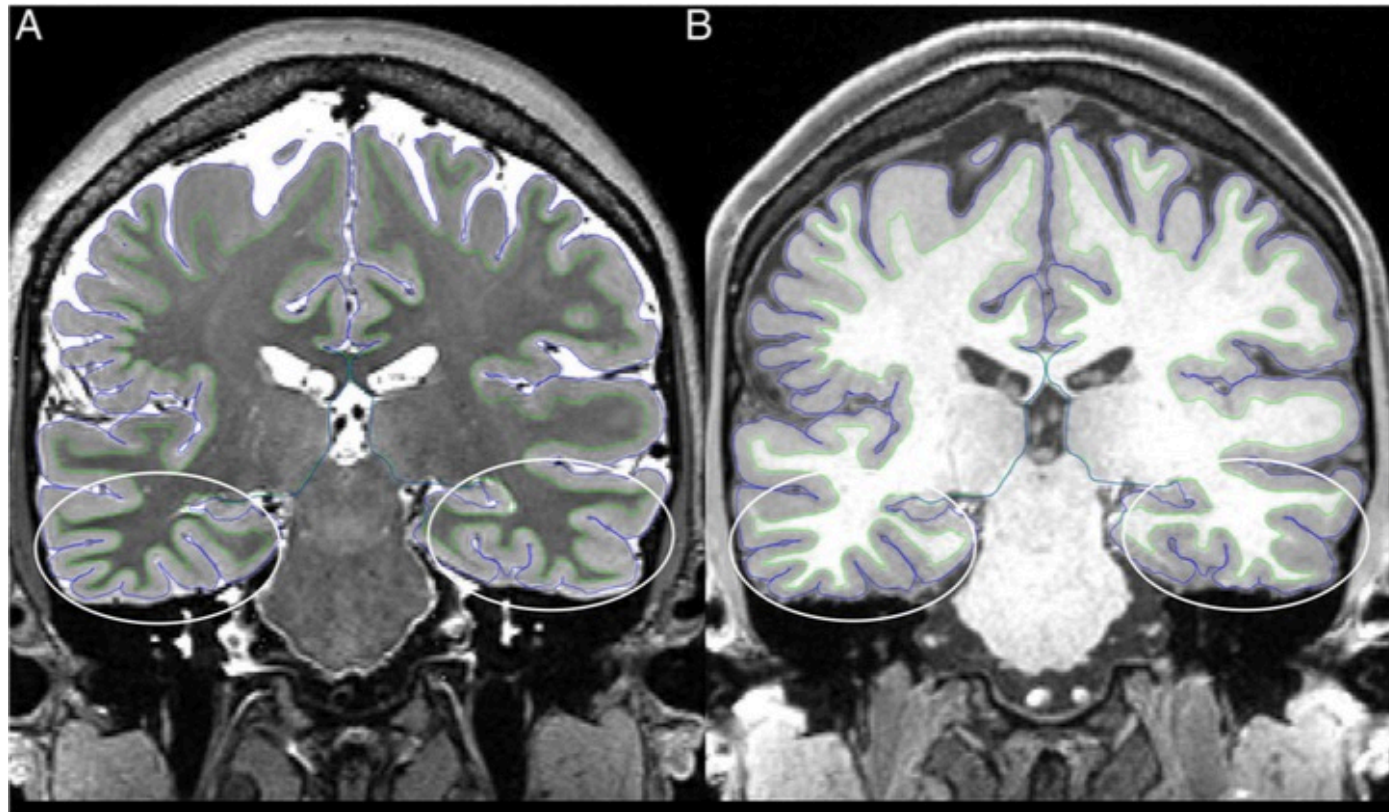
What data is available?

- Multimodal neuroimaging data
  - Diffusion-weighted images:
    - 3T (WUSTL, 1200 subjects):
      - 1.25mm isotropic voxels
      - 3 shells of b=1000, 2000, and 3000 s/mm^2
      - 90 diffusion-weighting directions acquired with RL and LR phase encoding
    - 7T (Uminn, 200 subjects)
      - 1.05 mm isotropic voxels
      - 2 shells of b=1000 and 2000 s/mm^2
      - 130 diffusion-weighting directions acquired with opposing phase encoding

What data is available?

- Multimodal neuroimaging data
  - Resting-state fMRI images:
    - 3T (WUSTL, 1200 subjects)
      - 2 mm isotropic voxels
      - Four runs of 1200 volumes at 720ms TR (14 min 33 sec each)
    - 7T (Uminn, 200 subjects)
      - 1.6 mm isotropic voxels

What data is available?

- Multimodal neuroimaging data
  - Task fMRI images:
    - Original tasks at 3T (WUSTL, 1200 subjects)
      - 2 mm isotropic voxels
      - 720ms TR
      - WM (Working Memory), GAMBLING, MOTOR, LANGUAGE, RELATIONAL, SOCIAL, EMOTION
    - Different tasks at 7T (Uminn, 200 subjects)
      - 1.6 mm isotropic voxels
      - Retinotopy, passive movie viewing

What data is available?

- Multimodal neuroimaging data
  - MEG:
    - SLU, 100 subjects
    - Resting MEG
    - Task MEG
      - MOTOR, LANGUAGE, and WM (Working Memory)

How can the data be accessed?

- ConnectomeDB
  - https://db.humanconnectome.org
- Open Access data use terms:
  - All imaging & most behavioral data
- Restricted Access data use terms:
  - Identifying information (e.g., family structure, age, height, weight, race)
  - Sensitive information (e.g., drug test / history, DSM symptom scores)

How can the data be accessed?

Subject Dashboard: WU-Minn HCP Data - 1200 Subjects

- Behavioral data
  - Create custom group: Filter subject list based on data availability, data values, etc.
  - Export spreadsheet of scores (comma-separated)

How can the data be accessed?

- Imaging data
  - Download "packages" for the selected group
  - Aspera protocol speeds up data downloads
  - Still… downloading enough data to analyze will take tons of time and space

How can the data be accessed?

- Connectome in a Box
  - Set of hard drives containing full dataset: ~ $2500
  - "Expansion set" (update from 500 or 900 subjects) may be available at a reduced price
  - https://store.humanconnectome.org

How can the data be accessed?

- Amazon Web Services (AWS)
  - HCP data is hosted on Simple Storage Service (S3)
    - You don't need to download and store it locally!
    - You don't need to buy Connectome in a Box!
  - HCP processing pipelines are preconfigured on Elastic Compute Cloud (EC2)
    - You don't need to install and configure the pipelines locally.
    - You don't need to have local computing resources to process large quantities of data.
  - Not free (but it can be pretty cheap)
    - https://wiki.humanconnectome.org/display/PublicData/How+To+Connect+to+Connectome+Data+via+AWS

What tools exist for analysis?

- Structure of CIFTI data
  - Cortical data mapped from volume to surface
  - Subcortex and cerebellum remain in volume
  - Data stored in single column / vector
  - Adjacent values in file not adjacent in space
  - XML header explains mapping of values to location
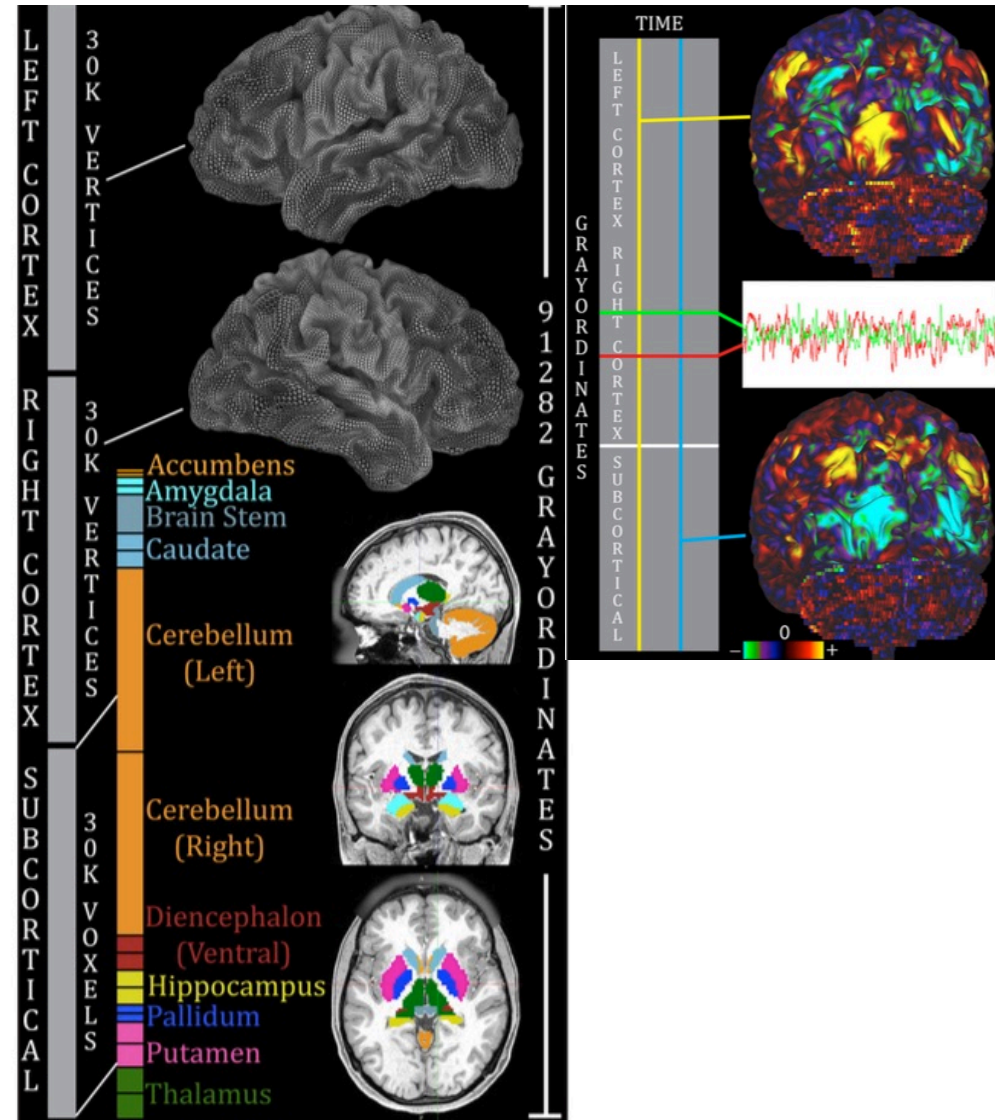
What tools exist for analysis?

- Connectome Workbench
  - wb_view: Data viewer for CIFTI, GIFTI, and NIFTI files
    - https://humanconnectome.org/software/connectome-workbench.html
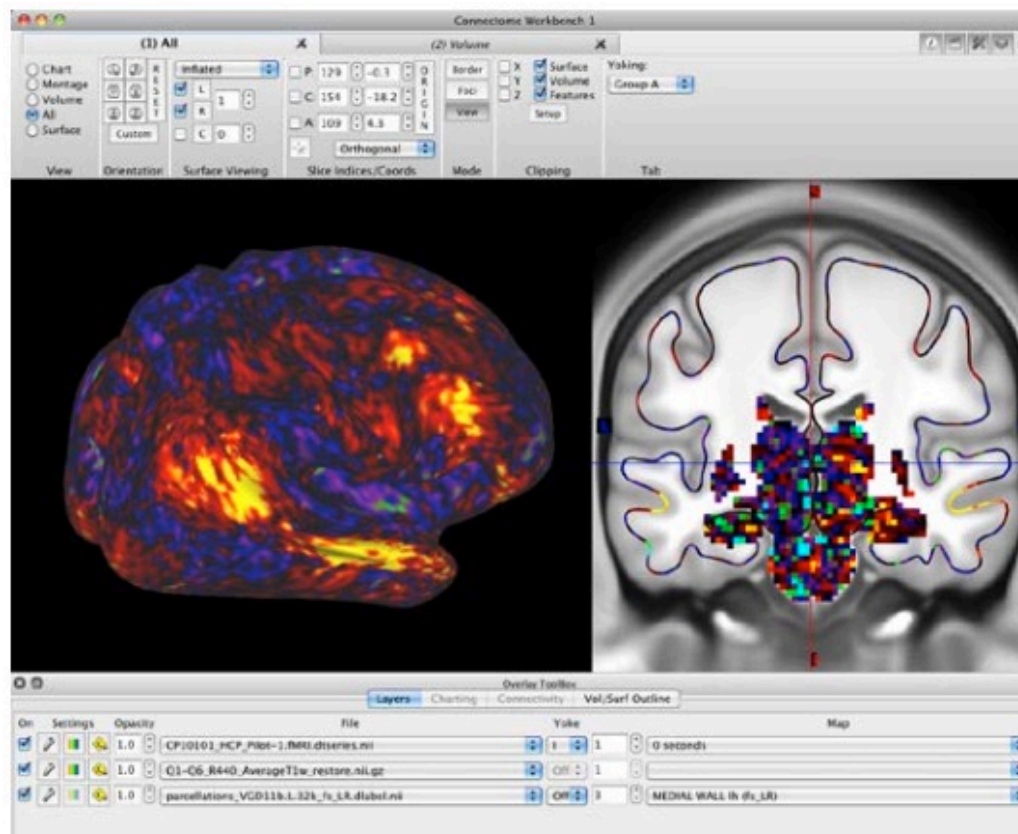
What tools exist for analysis?

- Connectome Workbench
  - wb_command: perform operations on CIFTI files
  - https://humanconnectome.org/software/workbench-command.php

```
Version: 1.2.3
Commit Date: 2016-08-23 19:08:10 -0500
Operating System: Linux

Information options:
    -help                      show this help info
    -arguments-help            explain the format of subcommand help info
    -cifti-help                explain the cifti file format and related terms
    -gifti-help                explain the gifti file format (metric, surface)
    -version                   show extended version information
    -list-commands             list all processing subcommands
    -list-deprecated-commands  list deprecated subcommands
    -all-commands-help         show all processing subcommands and their help
                                   info - VERY LONG

Global options (can be added to any command):
    -disable-provenance        don't generate provenance info in output files
    -logging           set the logging level, valid values are:
            SEVERE
            WARNING
            INFO
            CONFIG
            FINE
            FINER
            FINEST
            ALL
            OFF

To get the help information of a processing subcommand, run it without any
    additional arguments.

If the first argument is not recognized, all processing commands that start
    with the argument are displayed
```

What tools exist for analysis?

- FieldTrip
  - MATLAB toolbox for analysis of MEG data
    - http://www.humanconnectome.org/documentation/tutorials/field-trip-training.html

What tools exist for analysis?

- MATLAB cifti functions:
  - https://wiki.humanconnectome.org/display/PublicData/HCP+Users+FAQ#HCPUsersFAQ-2.HowdoyougetCIFTIfilesintoMATLAB?
  - Reads values from CIFTI files into matrix
    - Does not parse XML for spatial information

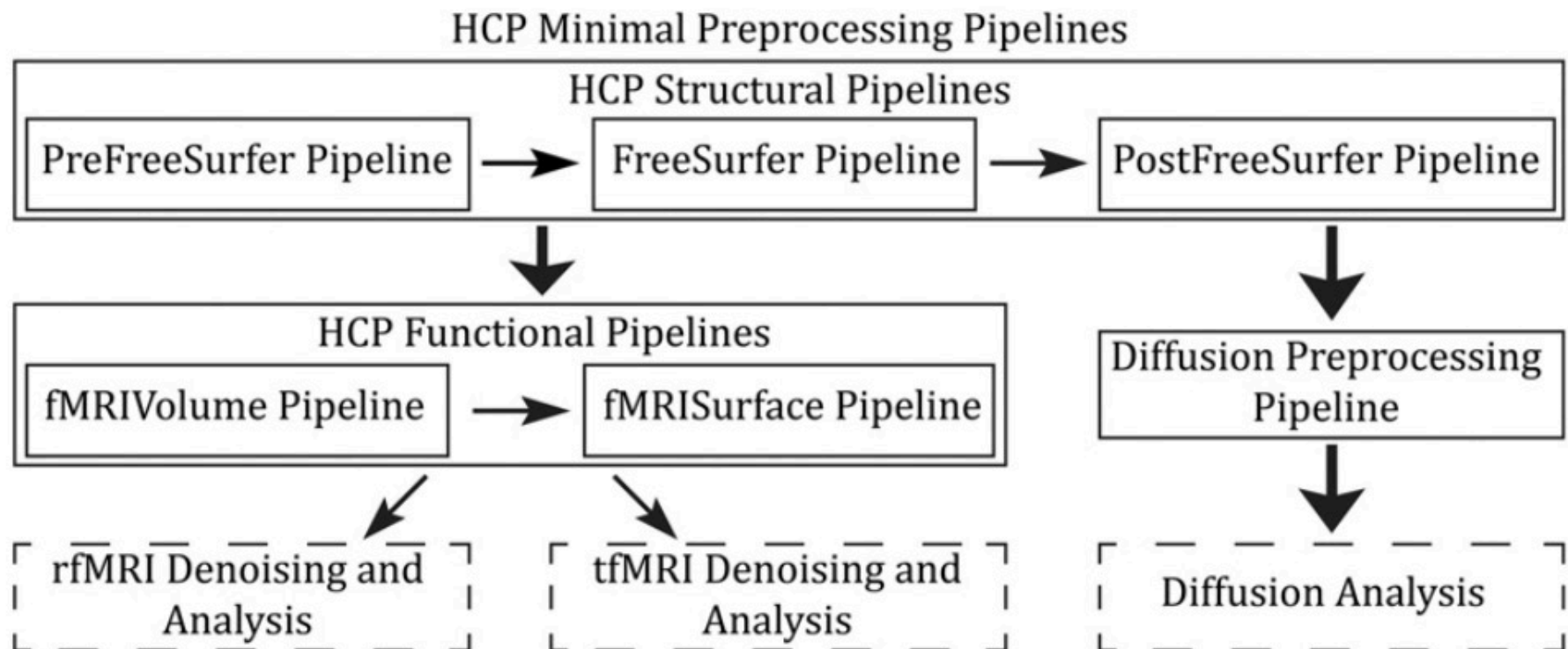# What tools exist for analysis?

- Using non-CIFTI tools with CIFTI data
  - List of CIFTI compliant tools:
    https://wiki.humanconnectome.org/display/PublicData/List+of+CIFTI-compliant+tools
  - Use wb_command to extract data
    - Split CIFTI into two separate GIFTI hemispheres and one NIFTI subcortical volume
    - Extract data values into NIFTI format file: 1 x 1 x N x t

What tools exist for analysis?

- HCP Pipelines:
  - Pipelines used to preprocess and analyze HCP data
  - https://github.com/Washington-University/Pipelines

HCP Minimal Preprocessing Pipelines

HCP Structural Pipelines

PreFreeSurfer Pipeline → FreeSurfer Pipeline → PostFreeSurfer Pipeline

HCP Functional Pipelines

fMRIVolume Pipeline → fMRISurface Pipeline

Diffusion Preprocessing Pipeline

rfMRI Denoising and Analysis

tfMRI Denoising and Analysis

Diffusion Analysis

How can I learn more?

- Documentation
  - https://humanconnectome.org/documentation/S1200

- Neuroimage Special Issue: Mapping the Connectome
  - http://www.sciencedirect.com/science/journal/10538119/80

- Wiki
  - https://wiki.humanconnectome.org/display/PublicData/Home

- HCP Users Mailing List
  - https://www.humanconnectome.org/contact/hcp-users-request.php

- 2017 HCP Course in Vancouver, BC
  - https://store.humanconnectome.org/courses/