An Introduction to Methods for Analyzing Large Neuroimaging Datasets

Robert Whelan, PhD Organizer
Trinity College Dublin
Dublin, Ireland

Herve Lemaitre Co-Organizer
IMN-GIN
Bordeauz, FM

Educational Course

Timeliness and importance of the topic: Human brain imaging is in a period of profound change. There is growing recognition that sample size must drastically increase to achieve adequate statistical power and reproducibility. Accordingly, several large neuroimaging databases have been established recently. For example, Adolescent Brain and Cognitive Development (ABCD: https://abcdstudy.org/) is a 10-year, $300 million, neuroimaging project that will recruit 10,000 people (currently mid-way through). Importantly, ABCD data will be fully open access, available with minimal restrictions. Similar open-access databases include Alzheimer's Disease neuroimaging initiative (ADNI: http://adni.loni.usc.edu/) and Open Access Series of Imaging Studies (OASIS; https://www.oasis-brains.org/), sharing platforms (e.g., openneuro.org) or belonging to large consortia (IMAGEN: https://imagen-europe.com/). The UK Biobank (https://www.ukbiobank.ac.uk/) will collect neuroimaging data from 100,000 people (currently, 50,000 have been scanned), and these data are available to researchers for a very modest access fee. These datasets will open up neuroimaging to a new generation of scientists.

Timeliness and importance of the desired learning outcomes: Despite the recent proliferation of large, easily available, neuroimaging datasets there are few resources that are specifically oriented towards the specialized methods required to analyze them. Rather, extant resources focus on methods appropriate to analysis of a relatively small number of subjects that a single laboratory would obtain. The aim of this course will be to provide researchers with the skills to interrogate these large datasets, and with best practice recommendations for managing and curating data.

Learning Objectives

Having completed this course, participants will be able to:
-Access and manage large neuroimaging datasets available to the scientific community
-Work with very large datasets: preprocessing and statistical/algorithmic analyses to examine between-group or individual differences
-Apply best practice methods for ensuring reproducible workflows and for data curation and annotation

Target Audience

The target audience is researchers who have some prior experience working with structural and/or functional MRI (and/or EEG/MEG) data and who are interested in analyzing large neuroimaging datasets. Some level of coding knowledge would be advantageous, although some of the toolboxes can be used entirely through a user interface.

OHBM, Glasgow, June 19th, 2022

An Introduction to Methods for Analyzing Large Neuroimaging Datasets

Human brain imaging is in a period of profound change, with a growing recognition that sample size must drastically increase to achieve adequate statistical power and reproducibility. Fortunately, several large neuroimaging databases are now widely accessible (e.g., ABCD, ADNI, OASIS, UK Biobank) plus open-data platforms such as Openneuro.

Do you want the tools to work on these large neuroimaging datasets?
If so, this workshop is for you!

Experts (talk list on next page) will show you how to deploy *scalable methods* to process neuroimaging data from structural to functional modalities and how to apply approaches such as *machine learning* and *graph theory* to large datasets. The course will also cover important general topics, including setting up a *reproducible* and *sustainable* workflow, using *cloud-based* software and *data annotation*.

The educational course will showcase step-by-step concrete examples so you can translate the knowledge gained to your own research.

All the materials – data and links to software – will be collated in a single location on OSF for attendees.

For more information, email: robert.whelan@tcd.ie or herve.lemaitre@u-bordeaux.fr

Coffee available from 7.30

Presenter(s)	Title	Time
Jivesh Ramduny	**Getting started**	**8.00-8.30**
Tara Madhyastha	**Neuroimaging workflow in the cloud**	**8.30-9.10**
Felix Hoffstaedter	**From dozens to thousands: important lessons when scaling up structural MRI processing using CAT**	**9.10-9.50**
	BREAK	**9.50-10.10**
Oscar Esteban	**NeuroImaging PREProcessing toolS (NiPreps)**	**10.10-10.50**
Lea Waller	**Harmonizing data analysis for diverse neuroimaging datasets using HALFpipe**	**10.50-11.30**
Sage Hahn	**Machine learning for Neuroimaging in Python with the Brain Predictability toolbox**	**11.30-12.10**
	LUNCH	**12.10-1.20**
Yihe Weng & Rory Boyle	**Studying the connectome at a large scale**	**1.20-2.00**
Emin Serin	**Combining graph theory and machine learning**	**2.00-2.40**
	BREAK	**2.40-3.00**
Scott Makeig	**Annotating the timeline of neuroimaging time series data using Hierarchical Event Descriptors (HED)**	**3.00-3.40**
Mélanie Garcia & Clare Kelly	**Establishing a reproducible and sustainable analysis workflow**	**3.40-4.10**
	WRAP-UP	