Maximizing open data & data sharing during COVID-19

Sook-Lei Liew, PhD, OTR/L

Assistant Professor & Director, Neural Plasticity and Neurorehabilitation Lab

University of Southern California

sliew@usc.edu | https://chan.usc.edu/npnl/

Open Science: What is it?

- <u>Open science movement</u>: Sharing (published & unpublished) data, code, resources

- <u>Why do it?</u> To improve scientific reproducibility and build the capacity of the scientific community (especially trainees)

- <u>What's involved?</u> Usually free to download, with some agreement you won't abuse/sell the data. That's it!

Open Data: What can I do with it?

- Can be used for your thesis or to maintain productivity

- Learn/develop/validate a method/analysis pipeline you'll use later when you collect your proposed dataset

- Test (part of) your hypothesis in a related dataset

- Examine and identify new covariates of interest

- Create your own open data archive/repository (!!)

Open Data: What types of data are there?

- Data types: Surveys, behavioral measures, demographics, kinematic data, videos, physiological data (e.g., brain imaging)

- Health services / medical records

- Prospective data collections (protocol is set prior to data collection)

- Retrospective data archives (usually study-specific data)

Open Data: What types of data are there?

- **Rehabilitation-Related Data Archives (NCMRR-funded)**
  - CLDR: https://www.utmb.edu/cldr/
    - Center for Large Data Research and Data Sharing in Rehabilitation
    - Many types including health services research (e.g., medical records) and retrospective study-specific rehabilitation data
  - ICPSR/ADDEP: https://www.icpsr.umich.edu/web/pages/ADDEP/index.html
    - Archive of Data on Disability to Enable Policy and research
    - Retrospective study-specific rehabilitation data
  - OpenSim: https://opensim.stanford.edu/
    - Free motion simulation toolbox and trained models for different populations: http://simtk.org/

Open Data: What types of data are there?

- **Prospective/Coordinated Brain Imaging, Clinical/Behavior**
  - Human Connectome Project: https://www.humanconnectome.org/
    - Lifespan, young adult, clinical populations, with harmonized behavior
  - UK Biobank: https://www.ukbiobank.ac.uk/
    - UK health records data including brain imaging, genetics, clinical variables
    - Working up to 100,000 individuals
  - All of Us: https://allofus.nih.gov/
    - On beta release; will be US health records data including brain imaging, genetics, clinical variables and questionnaires
    - Working up to 1 million individuals

Open Data: What types of data are there?

- **Community (Study-Specific) Brain Imaging**
  - Open Neuro: https://openneuro.org/
    - 372 MRI, MEG, EEG, ECoG datasets
  - INDI: http://fcon_1000.projects.nitrc.org/
    - International Neuroimaging Data-Sharing Initiative: Prospective and retrospective data
    - Resting state fMRI, structural MRI, diffusion MRI with behavioral measures
  - NITRC: https://www.nitrc.org/
    - Neuroimaging Tools and Resources Collaboratory: Atlases, data, and tons of software/tools

Open Data: But I want something specific?

NPNL at USC
@NPNLatUSC

If you're a student/researcher who can't collect data right now but who needs data to analyze to support your thesis/grant/project, let me know what type of data you're looking for, and I'll try to find an openly shared source! Will be doing an @ASNRehab webinar on this soon!

- Myelin water fraction MRI with behavior
- EEG data during FES-evoked movements
- Walking data with EMG, kinematics
- Resting state EEG with motor learning

Open Data: Collaborative data sharing

- If you have a specific need, you may consider reaching out to someone who has published a dataset that you'd like to utilize

- General guidelines:
  - Collaborate on the data (including authorship)
  - Receive useful insight on the data wrt how you use it
  - No one's data is perfect!
  - Maybe help organize their data into a data archive that you both can also publish (see journals like *Scientific Data*, *GigaScience*) or cite

Open Data: I want to share data

- You can apply for a CLDR rehabilitation-specific data sharing grant to cover your time/effort to archive existing data: https://www.utmb.edu/cldr/pilot-projects/data-sharing-studies

- Learn more about FAIR principles and reproducible methods for open science: https://www.repronim.org/index.html

- Happy to discuss best place to archive or other questions: sliew@usc.edu

Next Up:
Meta-Analysis with Keith Lohse, PSTAT, PhD