What we are trying to do here

Jessica Turner & Tom Nichols

Data Harmonization workshop, June 19 2015

Data sharing is happening

- Neuroimage issue, “Sharing the Wealth: Brain Imaging Repositories, 2015”
  - 25 papers currently in press

PING

Pain and Interoception Imaging Network (PAIN)

NeuroVau

A public repository of unthresholded brain act

NUNDA

Northwestern University Neuroimaging Data Archive

NiDB

Neuroinformatics Database

INS

Store Everything

Cimbi

THE LUNDBECK FOUNDATION

XNAT

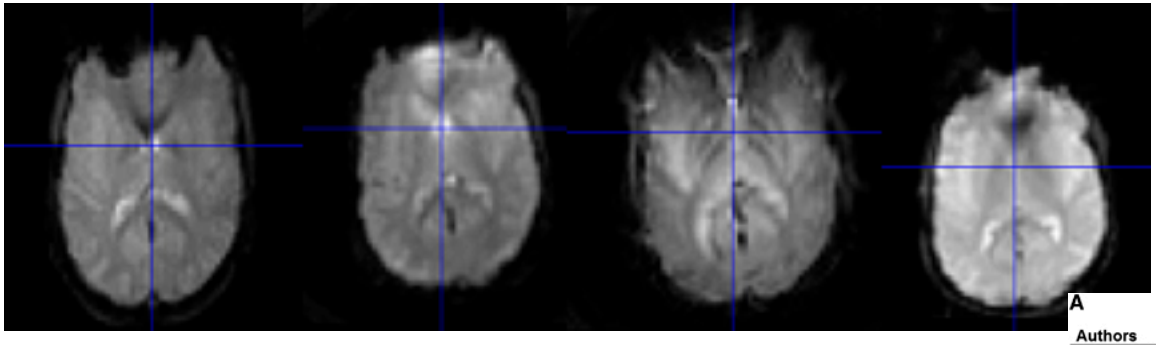
Openf

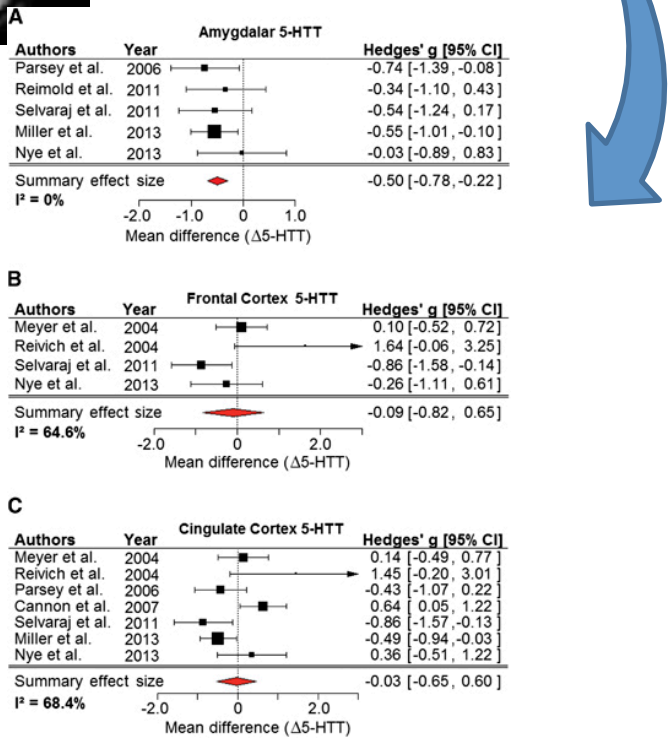
Harvard Aging Brain Study

Athinoula A. Martinos Center for Biomedical Imaging

IEEG.org

What is the problem?





Do differences across scanners and images in T1, T2, fMRI, ASL, DTI, etc.
limit our sensitivity to true effects?

If so, how do we reduce the variability
For the analyses we want to do?

Goal today

- Hear from people who have worked with legacy data
- Consider what needs to be done and approaches to do it

- Aim for a white paper with recommendations for using these datasets in different ways

Key question

- What are the limits to combining data across protocols (within a single modality)?
  - How bad is it?
  - How is that affected by the analysis goals?
- What are possible approaches to minimizing those issues?
  - Pros and cons
- How do we minimize this going forward?
  - Including education for data collectors and big data analysts

Panel Discussion

- Someone hands you an aggregated dataset of samples of different sizes and modalities. What do you do?

- Is there a minimum set of characteristics we can agree on?

Roundtable Discussion

- What do we share?
- Where do we share it?
- How do we continue to make available new analytic techniques for these data?