Evaluating the ANTs longitudinal cortical thickness pipeline

Andrew Holbrook

Alzheimer's Disease Research Center
Institute for Memory Impairments and Neurological Disorders
University of California, Irvine

Department of Statistics
University of California, Irvine

Department of Human Genetics
University of California, Los Angeles

SMI 2019

The ANTs Longitudinal Cortical Thickness Pipeline

Nicholas J. Tustison, Andrew J. Holbrook, Brian B. Avants, Jared M. Roberts, Philip A. Cook, Zachariah M. Reagh, Jeffrey T. Duda, James R. Stone, Daniel L. Gillen, Michael A. Yassa, for the Alzheimer's Disease Neuroimaging Initiative

doi: https://doi.org/10.1101/170209

Contents

Contents

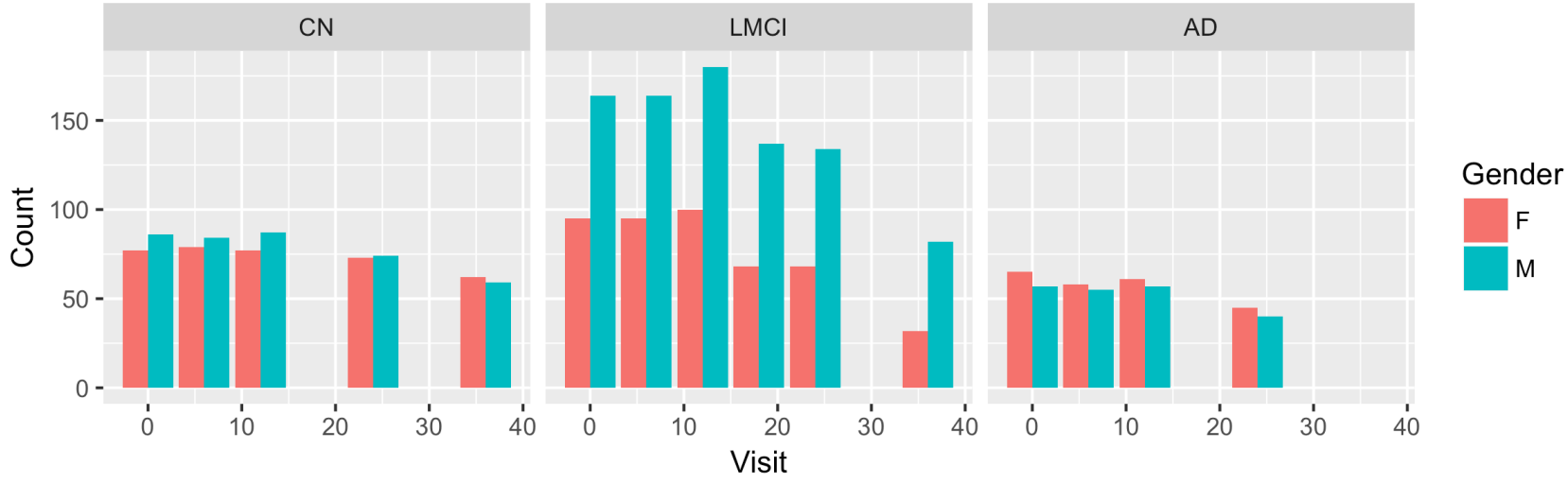
Contents

Contents

1. The data

Data collection

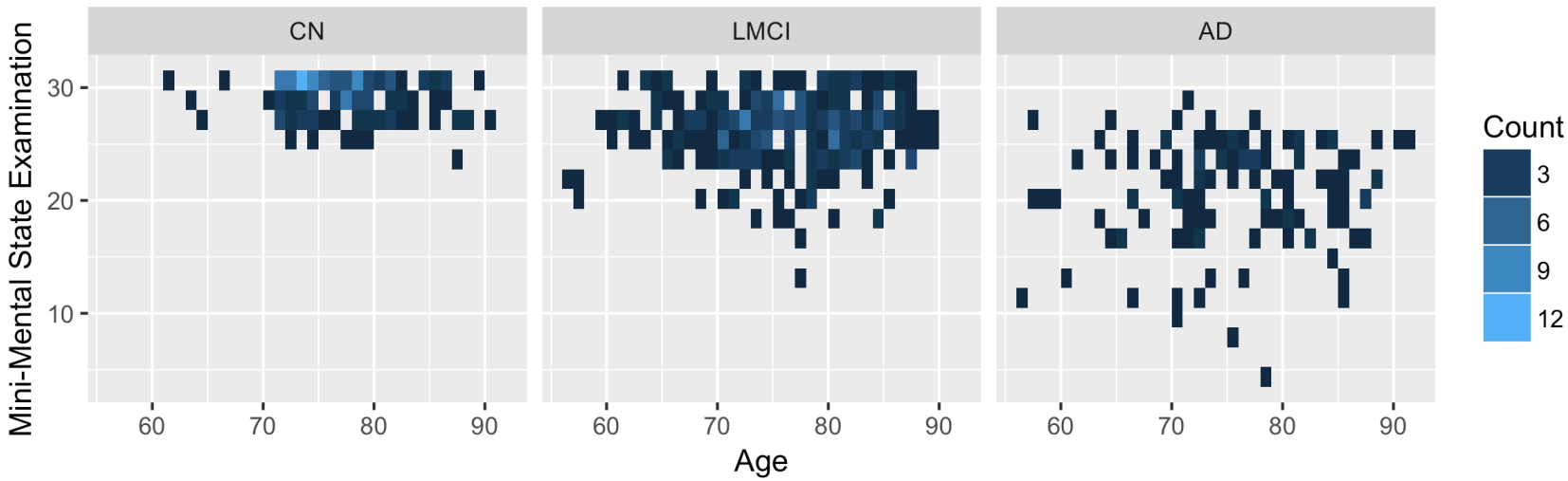
We use ADNI-1 data from 663 elderly adults: 197 cognitively normal (CN), 324 late mildly cognitively impaired (LMCI) and 142 with Alzheimer's disease (AD).



Data collected at 0, 6, 12, 18, 24 and 36 months. Images total 4399.

Health and age

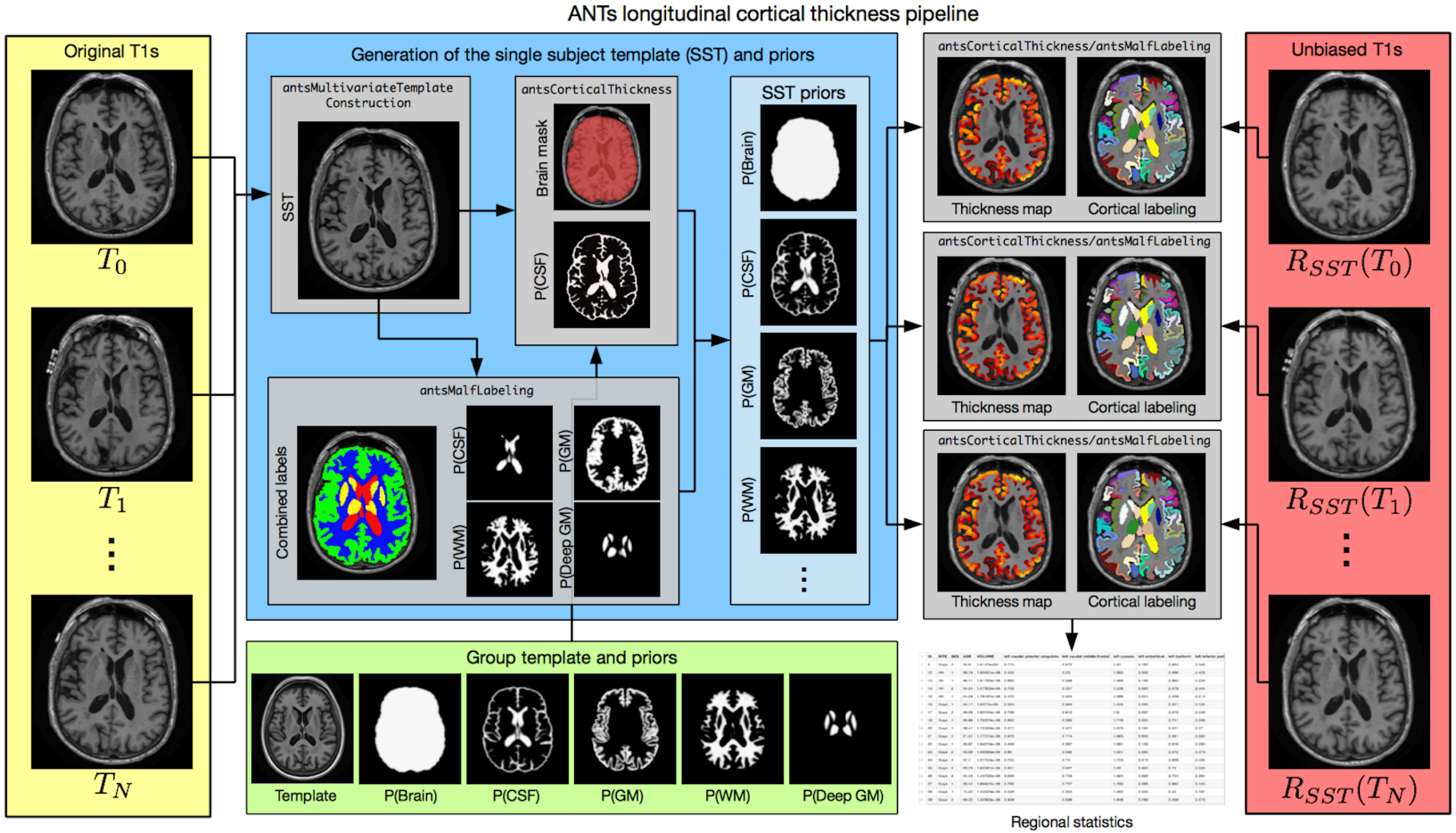
Mini-mental state exam (MMSE) score distributions vary by diagnostic status.



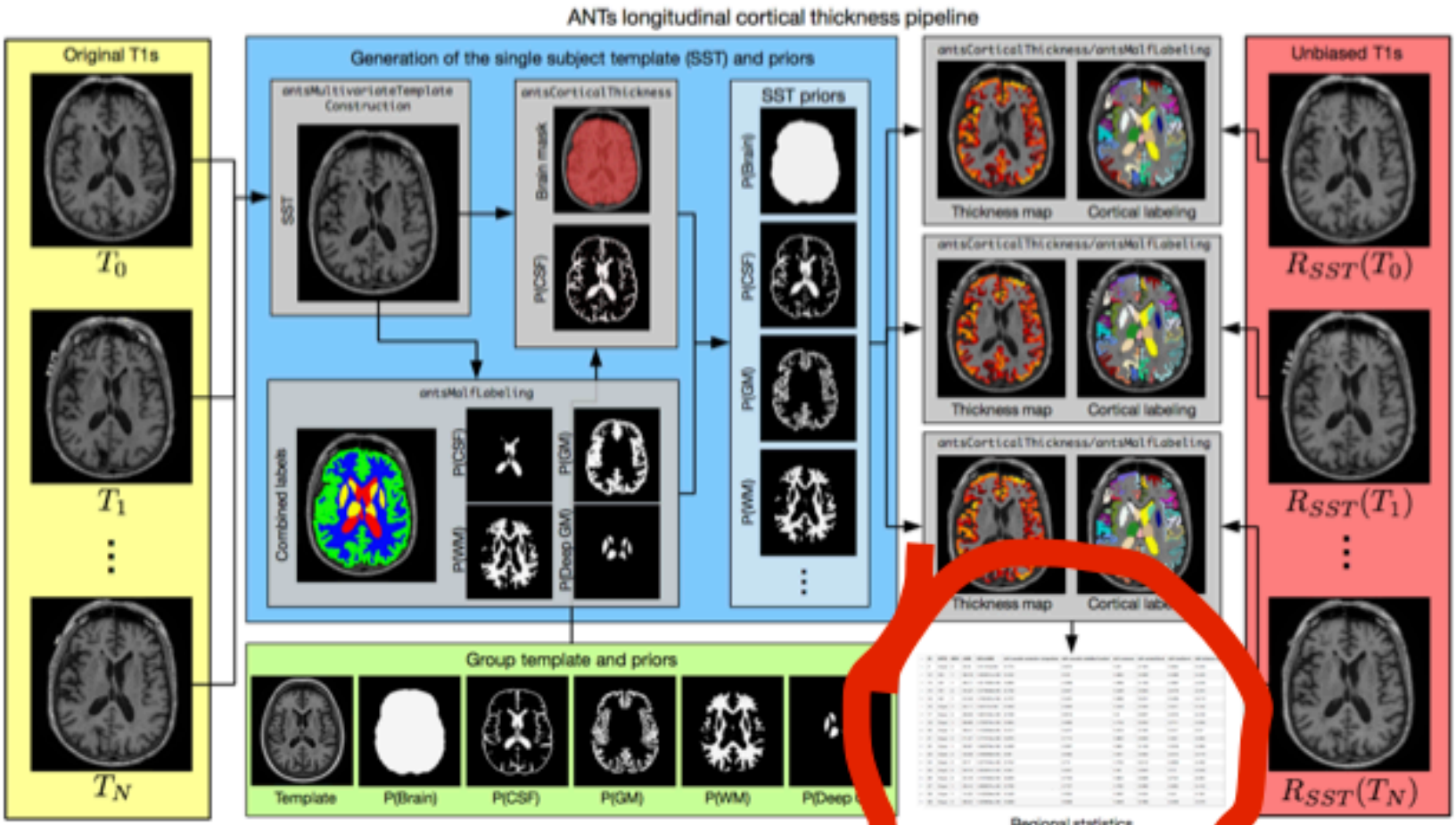
Ages range from late 50s to early 90s.

2. The *data*

A data making machine



A data making machine



data

The analyst hat

Suppose I want to use cortical thickness (CT) measurements for science. I have 5 black boxes I call:

1. [redacted]
2. [redacted]
3. [redacted]
4. [redacted]
5. [redacted]

Each provides regional CT data. Which to choose? How to choose?

The analyst hat

Suppose I want to use cortical thickness (CT) measurements for science. I have 5 black boxes I call:

1. ANTs Cross-sectional (ANTs Cross)
2. ANTs Longitudinal-SST (ANTs SST)
3. ANTs Longitudinal-native (ANTs Native)
4. FreeSurfer Cross-sectional (FS Cross)
5. FreeSurfer Longitudinal (FS Long)

Each provides regional CT data. Which to choose? How to choose?

The analyst hat

Suppose I want to use cortical thickness (CT) measurements for science. I have 5 black boxes I call:

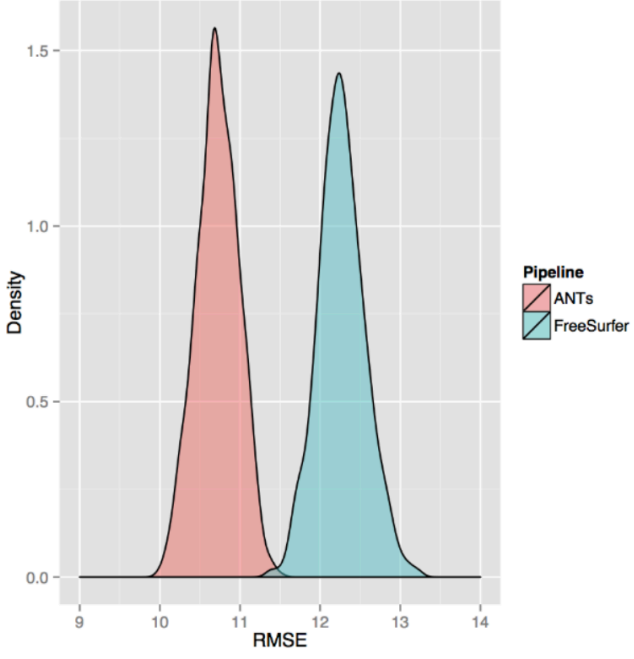
1. [redacted]
2. [redacted]
3. [redacted]
4. [redacted]
5. [redacted]

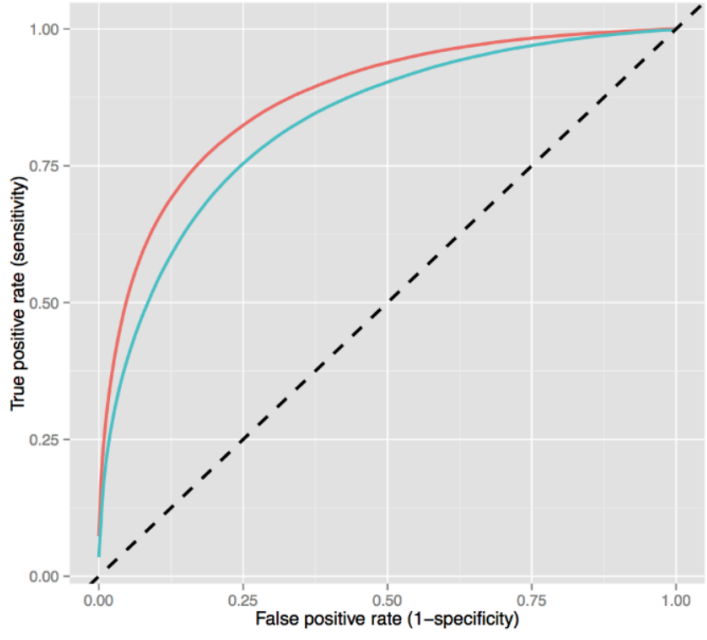
Each provides regional CT data. Which to choose? How to choose?

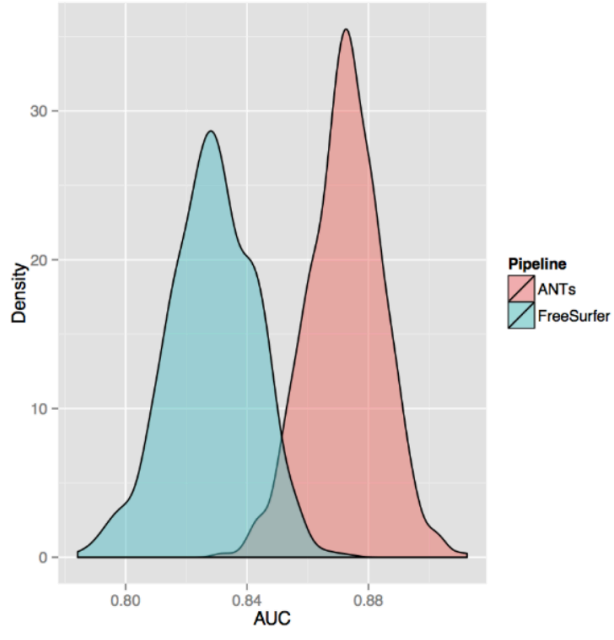
3. Supervised data selection

Predictive criteria

Use CT to predict age or sex. Look at the MSE or AUC.







Inferential criteria

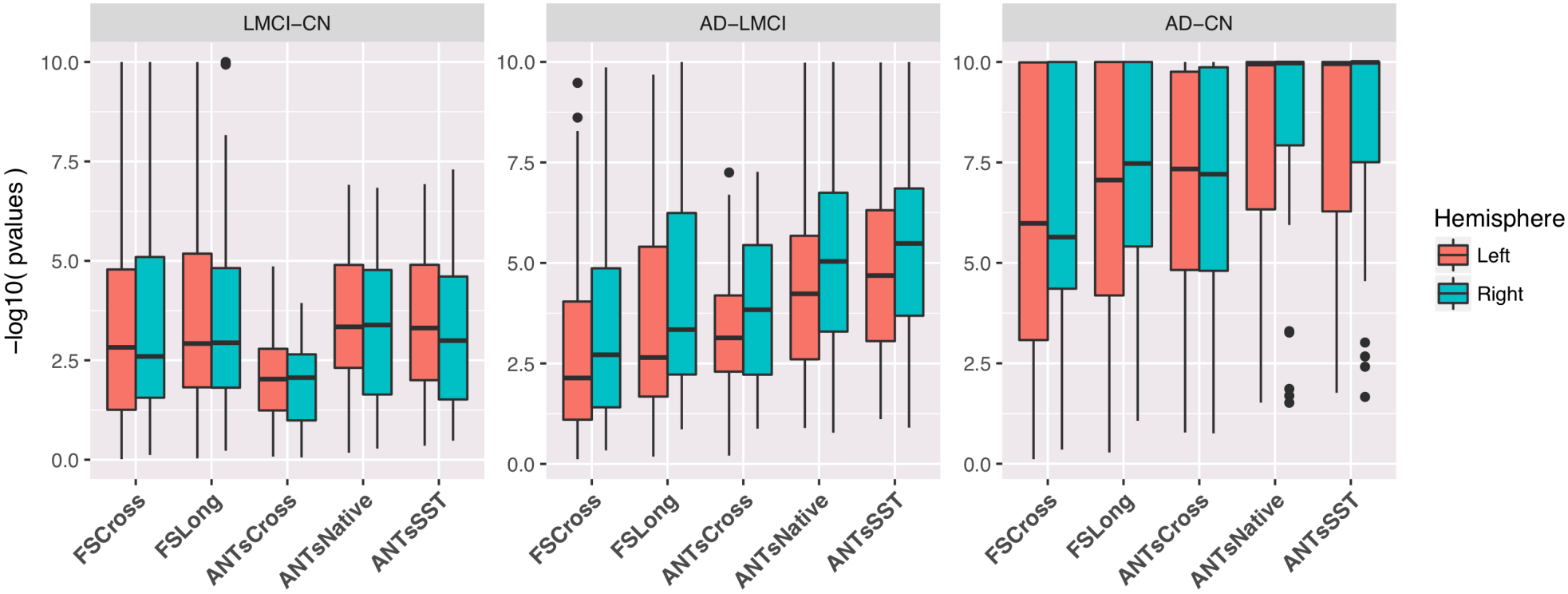
Model association between diagnosis and CT through time. LME helps control for confounders and correlations.

$$\begin{aligned}\Delta CT_{ij}^{k} = \beta_0 &+ CT_{i0}^{k}\,\beta_1 + AGE_{i0}\,\beta_2 + ICV_{i0}\,\beta_3 + APOE_i\,\beta_4 \\ &+ GENDER_i\,\beta_5 + DIAGNOSIS_i\,\beta_6 + VISIT_{ij}\,\beta_7 \\ &+ VISIT_{ij} \times DIAGNOSIS_i\,\beta_8 + \alpha_i^k + \gamma_s^k + \epsilon_{ij}^k\end{aligned}$$

Let k, i, s and j index regions, individuals, sites and months (from baseline), respectively.

Inferential criteria

Collapse the p-values over all 62 regions.



Pros and cons

Pro: easy to communicate

Pro: choice is tailored to application

Con: choice is tailored to application

Con: beware double dipping

Con: choice depends on prediction algorithm

Con: associations don’t prove accuracy

4. Unsupervised data selection

Use known structure

We model the CT trajectory with subject specific intercepts and slopes.

$$CT_{ij}^k \sim N(\alpha_i^k + \beta_i^k t, \sigma_k^2)$$

$$\alpha_i^k \sim N(\alpha_0^k, \tau_k^2) \qquad \beta_i^k \sim N(\beta_0^k, \rho_k^2)$$

$$\alpha_0^k, \beta_0^k \sim N(0, 10) \qquad \sigma_k, \tau_k, \rho_k \sim \text{Cauchy}^+(0, 5)$$

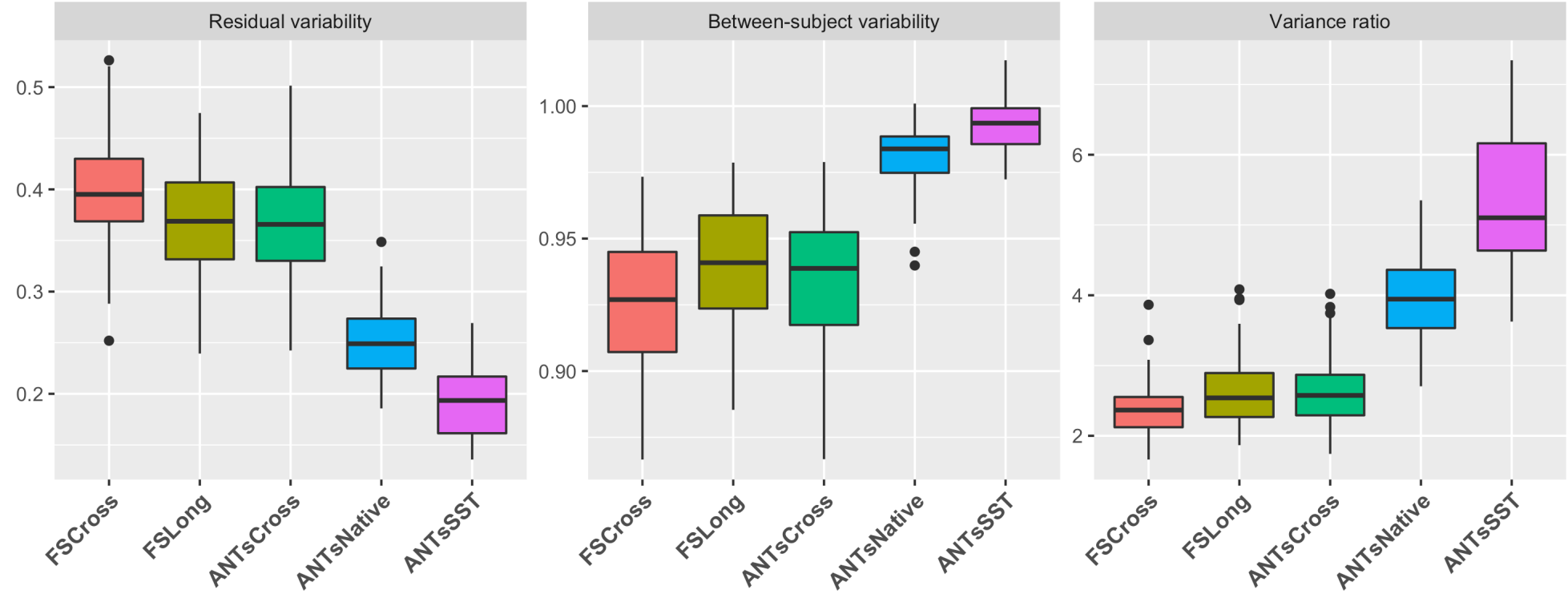
Choose data with higher

$$r^k = \frac{\tau_k}{\sigma_k}$$

and avoid the "double whammy": attenuation bias and increased variance (Carroll, et al. 2006).

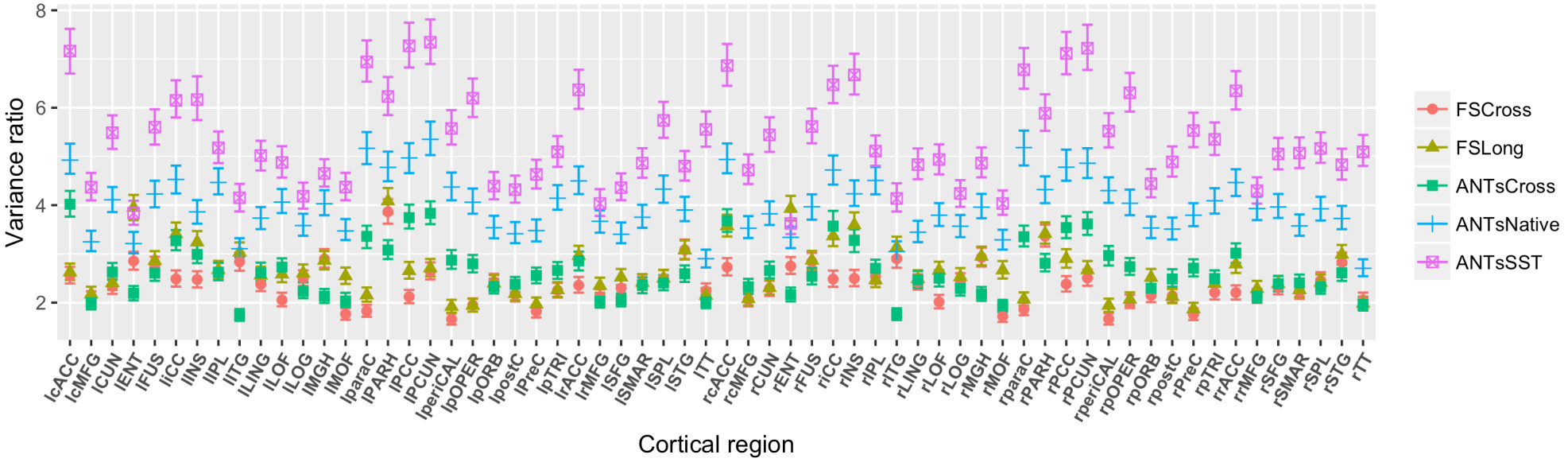
F-test, revisited

We can average empirical posteriors across the 62 regions.



Higher is better

We can compare 95% credible intervals across regions.



Left
Right
Superior
Frontal
Posterior
FreeSurfer Cross-sectional
FreeSurfer Longitudinal
ANTs Cross-sectional
ANTs Long-Native
ANTs Long-SST
8.0
2.0

Pros and cons

Pro: ranking depends on data alone

Pro: application independent

Con: structure must be specified correctly

Conclusions

Supervised data selection:

- good for specific applications
- easy to understand
- hard to extrapolate to different scenarios

Unsupervised data selection:

- only needs data of interest
- leverages known hierarchical structure
- application independent

Combination: evidence for ANTs SST performance

References

Tustison, et al. "The ANTs Longitudinal Cortical Thickness Pipeline" (2017)

Carroll, et al. "Measurement Error in Nonlinear Models: A Modern Perspective" (2006)