ADNI preliminary results:

A First Look at the

Longitudinal Data

Chicago, April 2008

ADNI Biostatistics Core:

Laurel Beckett, PhD, Danielle Harvey, PhD

Qian Weng, MS, Hao Zhang, MS

UC Davis

Goals for preliminary analysis:

1. To describe data available on March 15, 2008: clinical data, imaging summaries at LONI.
2. To summarize clinical changes in first year of follow-up: diagnostic category, functional and neuropsychological assessments.
3. To characterize patterns of change in imaging summary measurements.
4. To give preliminary estimates of imaging summary performance in study design.
5. To give preliminary estimates of correlation of imaging summary change with clinical change.

Datasets available for public access on LONI:

A bulk clinical data download is available containing demographics, diagnosis, and neuropsych test results (819 people). You get everything for all people; no selection is possible as yet.

Instructions on how to download data are posted on the ADNI Biostat web site; also available from our Biostat Core.

Neuroimaging summaries are available by lab.

Documentation is also online. READ the fine print!

About 3/4 have completed first year follow-up.
Summary of month 12 diagnostic changes:

Normal: 4/198 (2.0%) became MCI, no AD.

MCI: 54/302 (17.9%) became AD.

Some backward reclassification:

MCI: 7/302 (2.3%) reclassed as normal.

AD: 2/142 (1.4%) reclassed as MCI, no normal.

Note: not every participant has 12-month follow-up yet. Data are still preliminary: interpret with care!

Change over time in neuropsychological and functional assessment: a reference standard for comparison.

We present summaries for 4 key measures (3 general, one more specific to memory):

- MMSE
- ADAS-COG Total 11
- CDR Sum of boxes
- Verbal learning: RAVLT Sum of 5 trials

What features would show that a measure is a good marker for AD progression?

- At baseline, Normal better than MCI better than AD.
- Rate of decline: Normal less than MCI less than AD.
- Variation in rate of change within each group: we hope to be able to see clinical differences between people with similar diagnosis! (some progress, some do not.) Likely greater in MCI than normal.
- Low level of noise: paths are generally straight, not zig-zag.

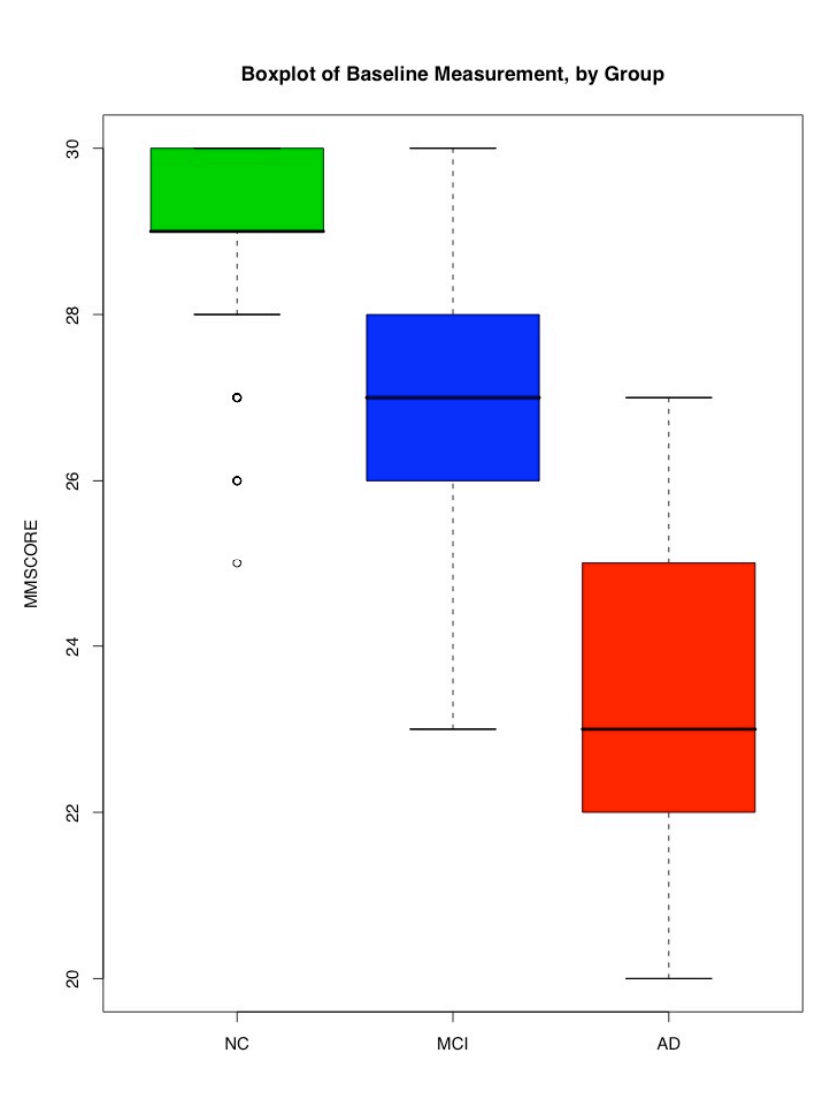
How do our clinical measures perform?

MMSE scores at baseline: Note the ceiling effect.

Green=NC

Blue=MCI

Red=AD



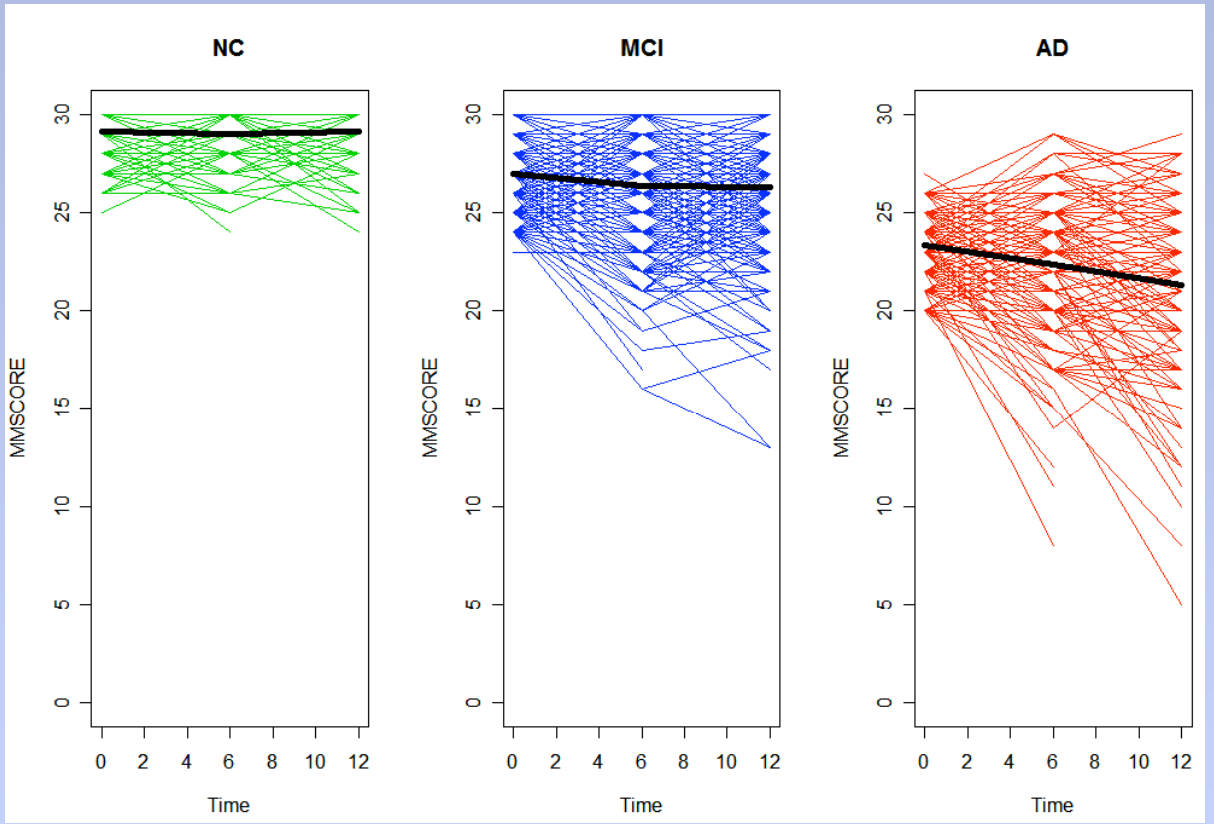
MMSE change over first year of follow-up.

Green=NC

Blue=MCI

Red=AD

Note: decided trend, but lots of scatter, lots of zig-zag!

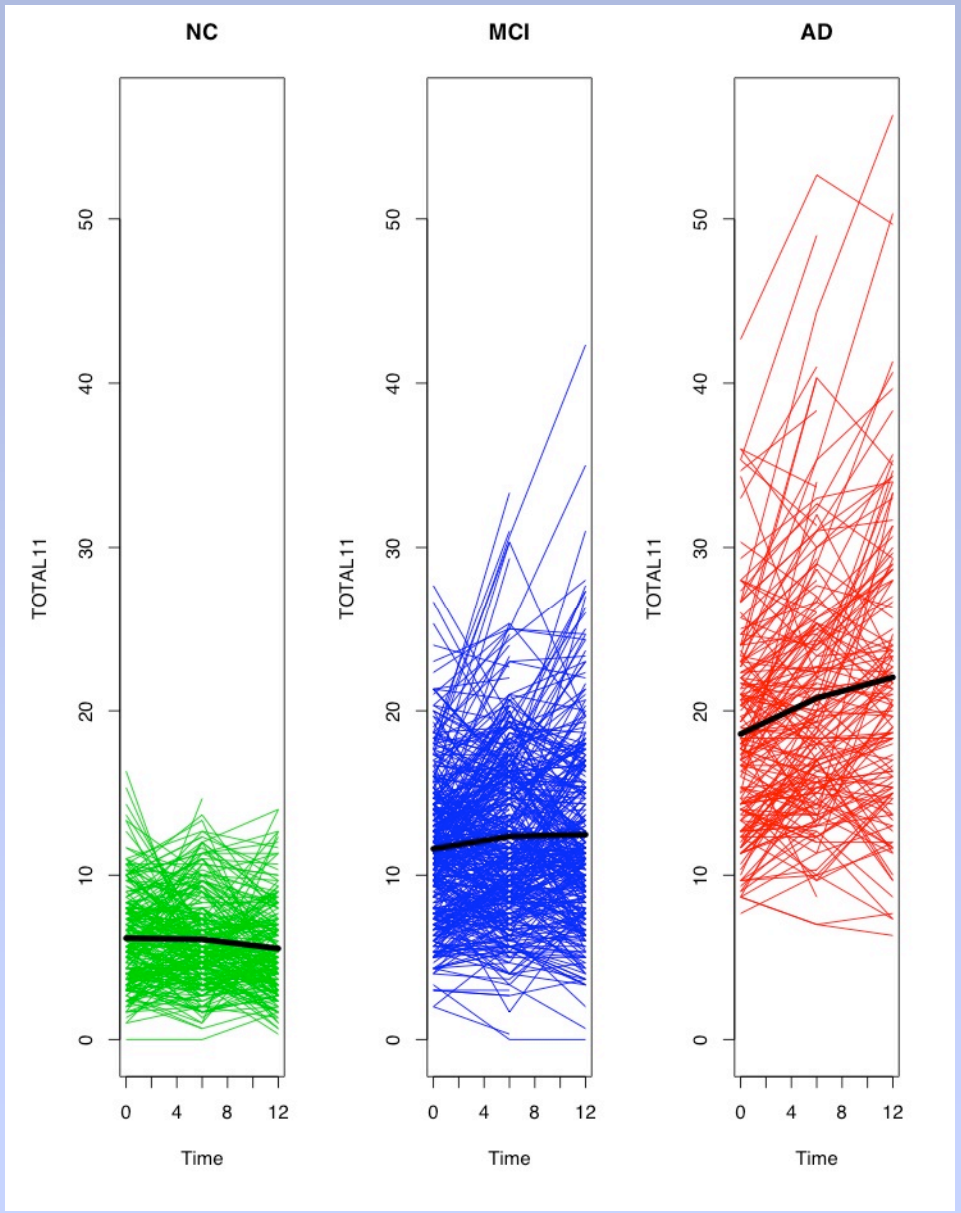


ADAS-COG Total 11 change over first year.

Green=NC

Blue=MCI

Red=AD

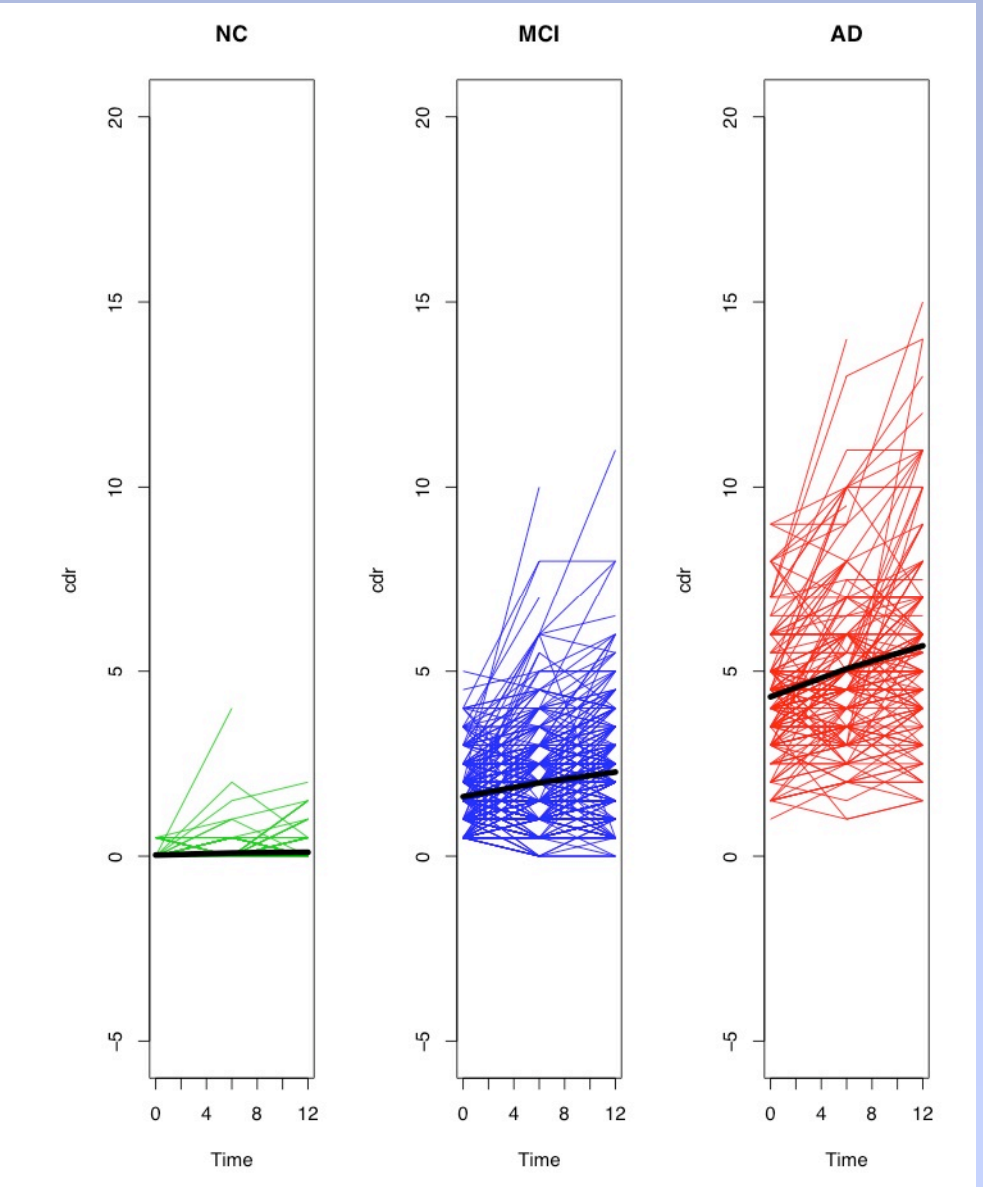


CDR Change over first year of follow-up.

Green=NC

Blue=MCI

Red=AD

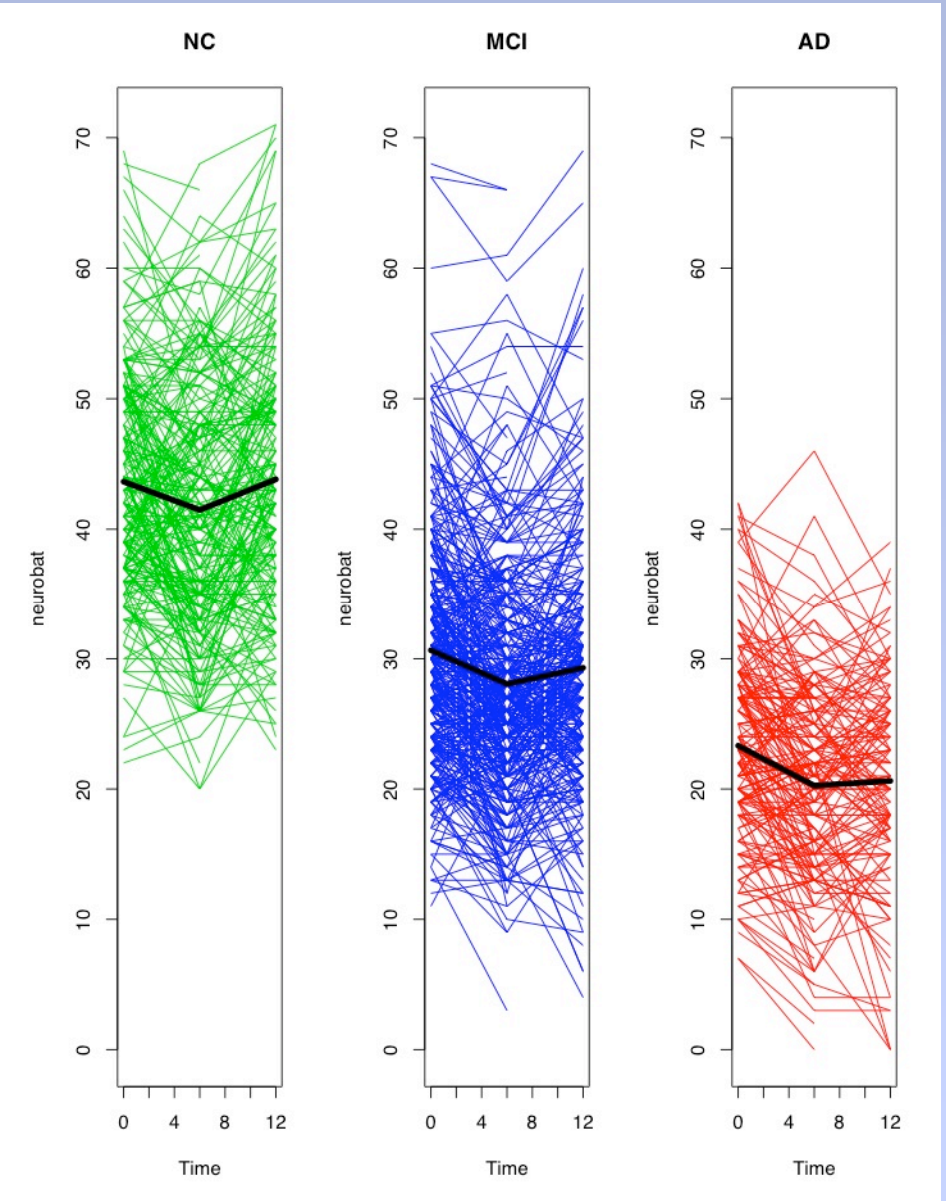


Word learning - RAVLT sum of five trials - one-year change.

Green=NC

Blue=MCI

Red=AD



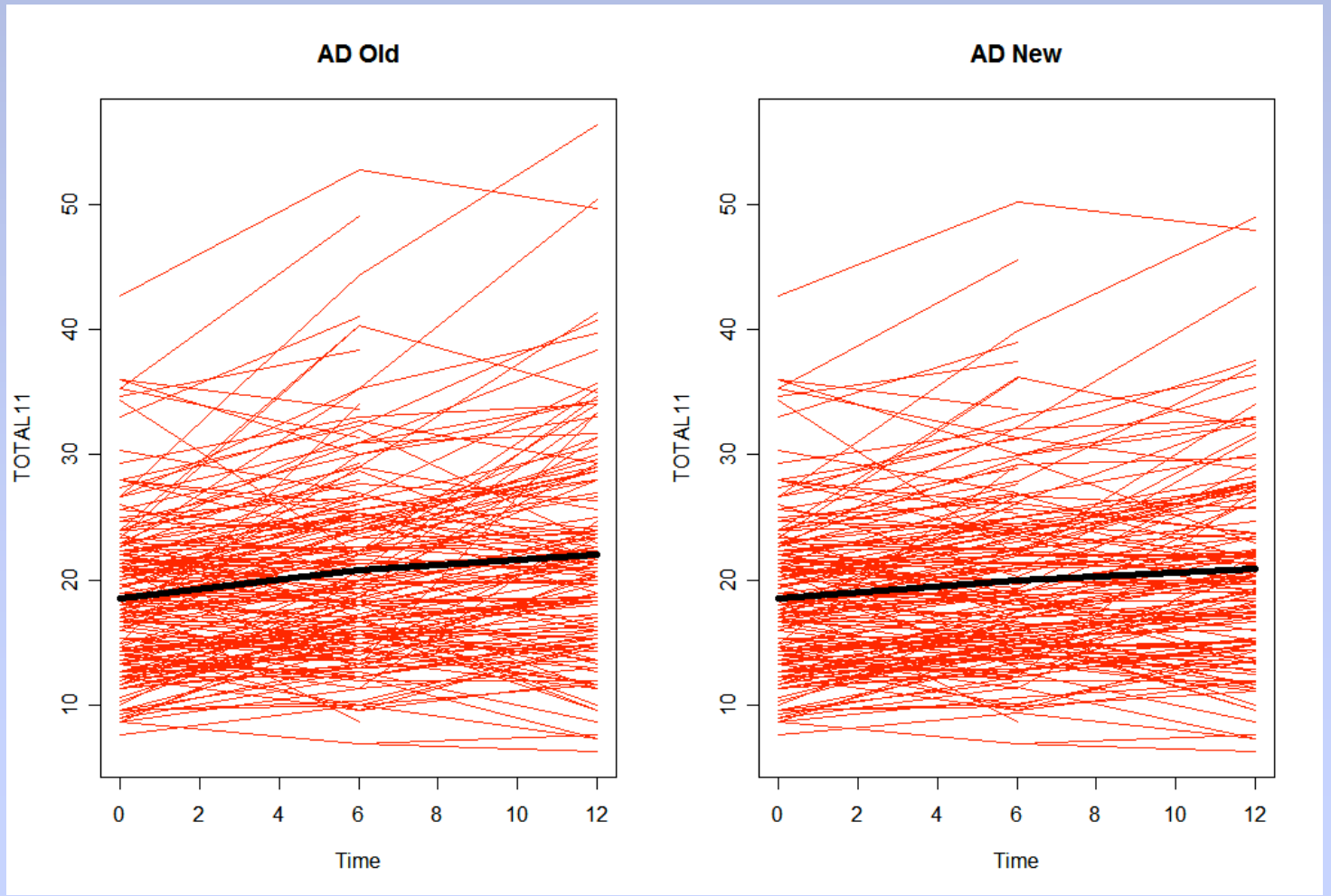
Implications for clinical trial design: how many people would you need to detect a 25% reduction in rate of change in the AD group? (for example, 3 per year instead of 4 per year on ADAS-Cog.)

Let's assume 80% power, two arms, equal numbers per arm, measurements at 0 and 12 months, 2-sided test, level 0.05.

Test	MMSE	ADAS COG	CDR Sum	Word learning
Number per arm	824	529	503	678

(even bigger samples needed for MCI studies!)

Why are sample sizes so big? Picture compares original ADAS-COG data for AD (Left) with modified (Right). We reduced score exactly 25% for m 6 and 12. It's hard to see difference!



These numbers are the key motivation for ADNI!

We need better, faster, cleaner ways to detect AD progression (and reduction in progression).

Let's take a **preliminary** look at the neuro-imaging summaries.

CAUTION!

1. Incomplete data: not all people imaged at 12 months, not all images QC and processed.
2. Different numbers/ different people per lab.
3. Analyses still preliminary; data just 3 wks old.
4. Some analyses not adjusted for multiple comparison.

Qualities we are checking for in imaging summaries:

- Cross sectional differences by baseline diagnosis: Normal better than MCI better than AD.
- Rate of decline: Normal less than MCI less than AD.
- Variation in rate of change within each group: differences in imaging change between people with similar diagnosis! (some progress, some do not.) Likely greater in MCI than normal.
- Low level of noise: paths are generally straight, not zig-zag.
- Promising summaries: correlate with clinical?

Longitudinal data analyzed for this meeting: MRI labs

We asked each lab to identify key metrics for this analysis; many more on line. Number of baseline and month 12 images processed for analysis varies by lab.

Lab	# NC	# MCI	# AD	Metrics
Alexander	100	143	56	16
Dale	53	51	24	6
DeCarli* (WMH)	124	156	71	1
Fox	103	148	52	2
Schuff	108	158	65	2
Studholme	111	154	61	1
Thompson	40	40	20	1

Let's see some pictures for imaging measures:

Dale lab: Whole brain volume.

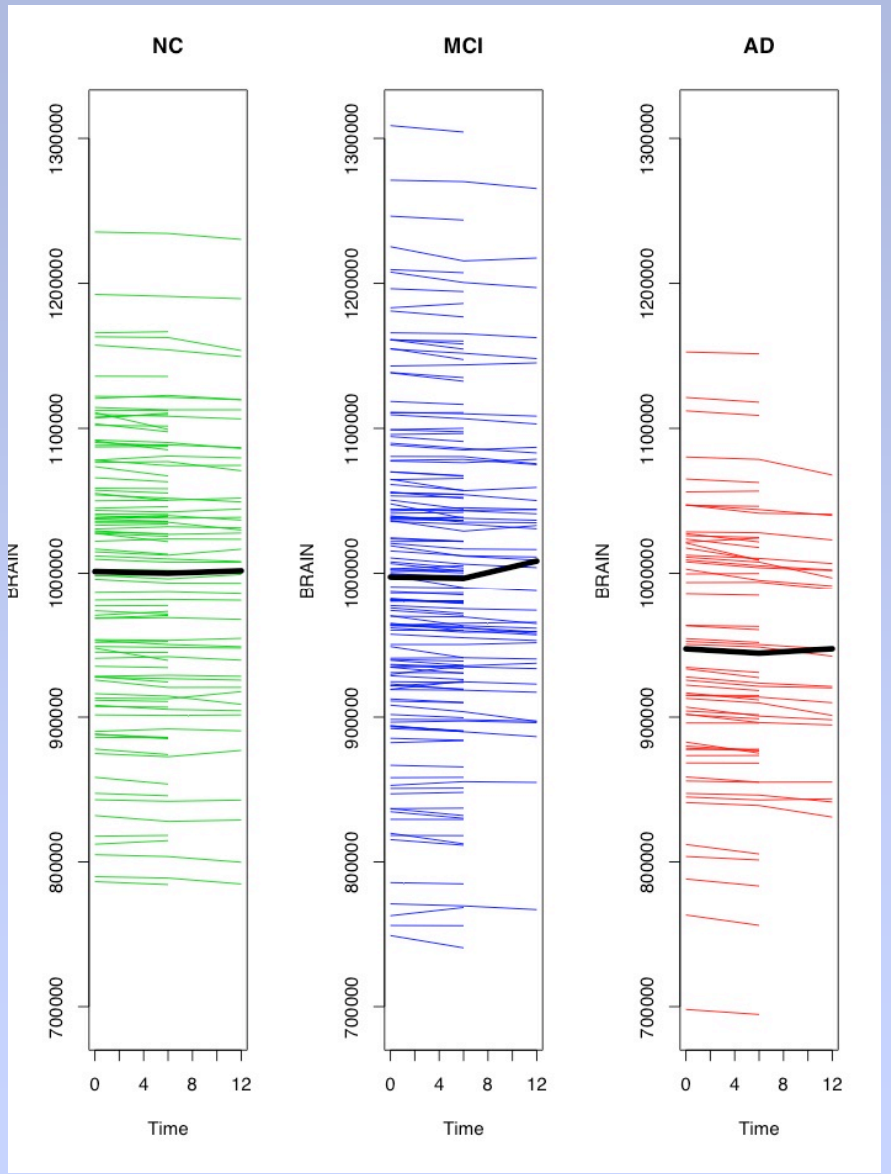
Green = NC

Blue = MCI

Red = AD

Note: much less zigzag pattern!

Easier to detect a shift in slope.

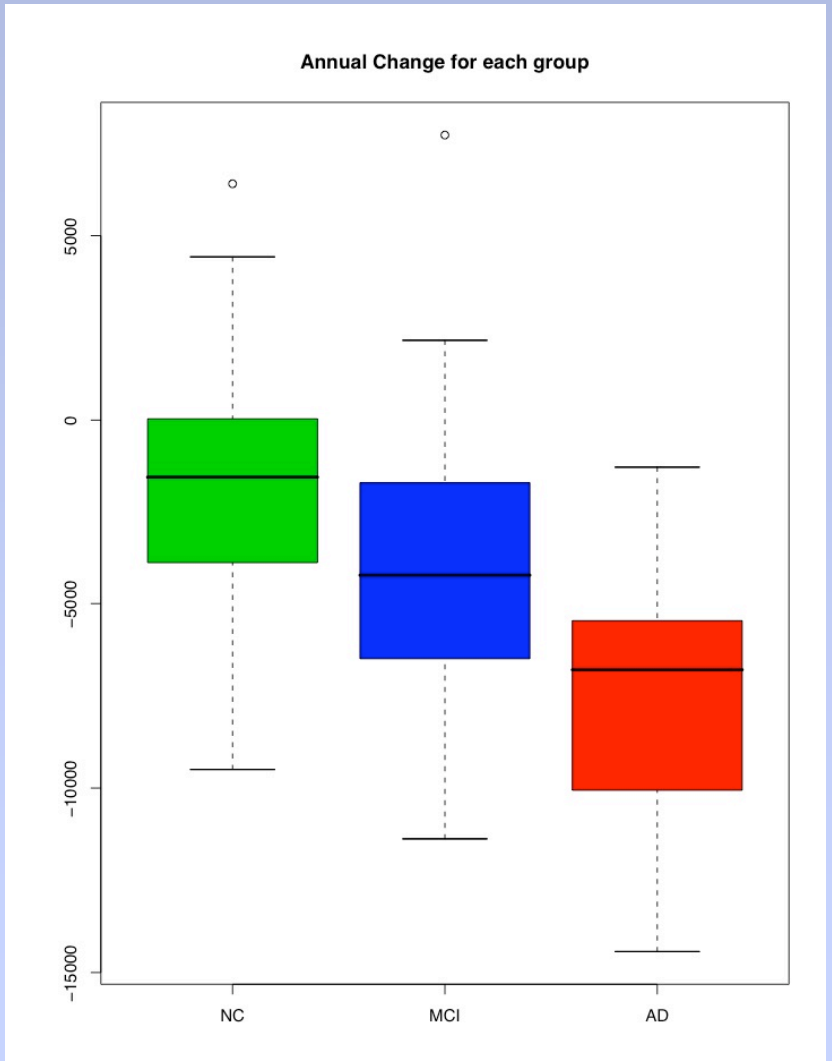


The slopes look quite parallel but there is some heterogeneity: box plots for Dale lab 12-month change in whole brain volume:

Green=NC

Blue=MCI

Red=AD

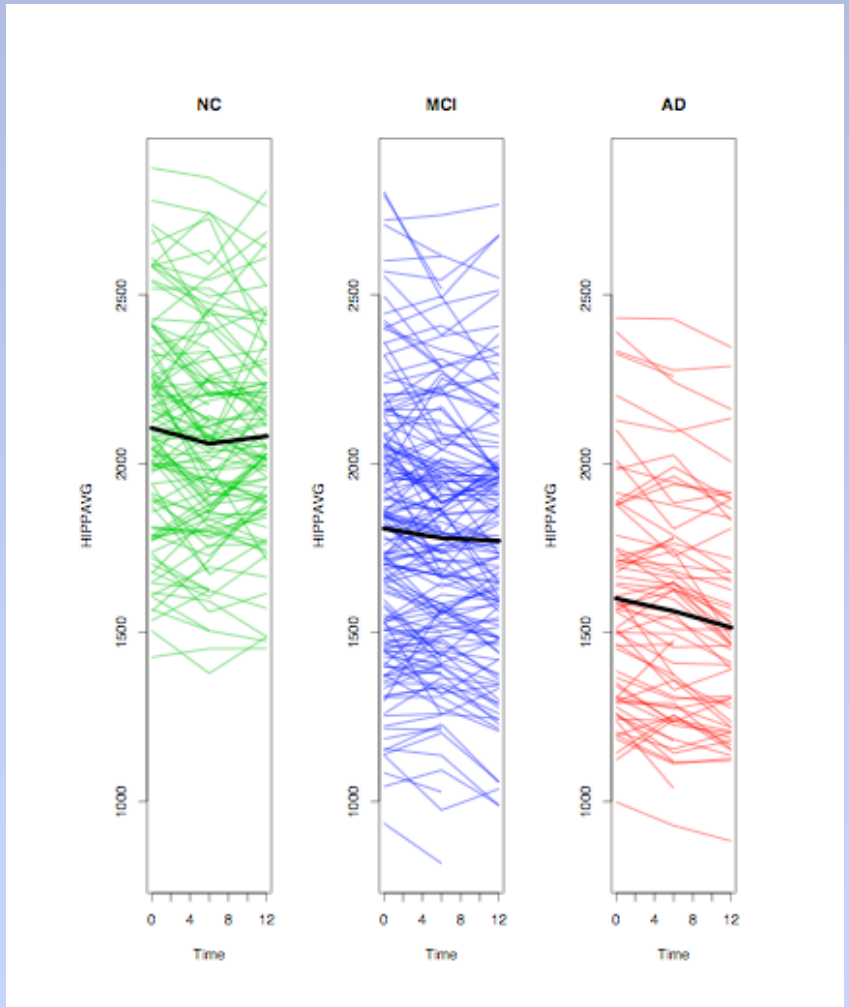


Schuff lab: average hippocampal volume. Trajectories show clear trends, especially in AD group.

Green=NC

Blue=MCI

Red=AD



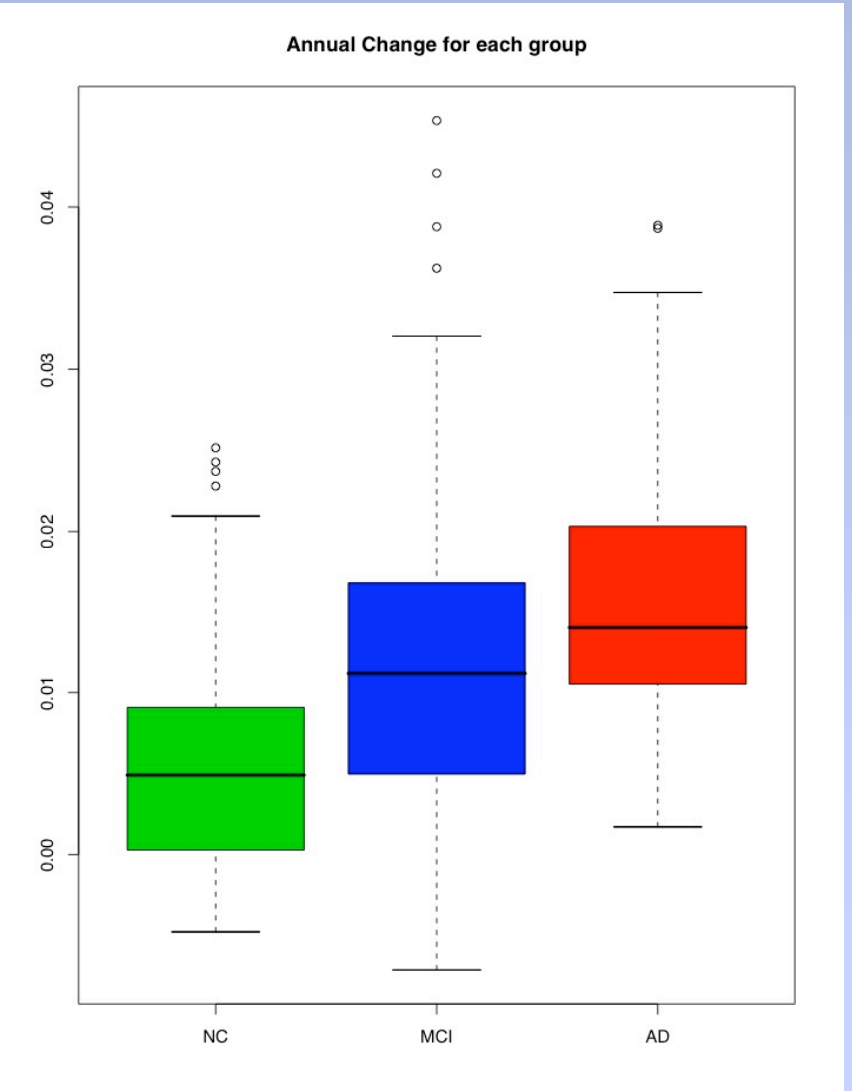
Some labs generate change rather than measure at each time point. Fox lab: differential bias-corrected boundary shift integral (% change from baseline to m 12)

Green=NC

Blue=MCI

Red=AD

Note: spread small compared to mean and to no change; easier to detect 25% reduction.

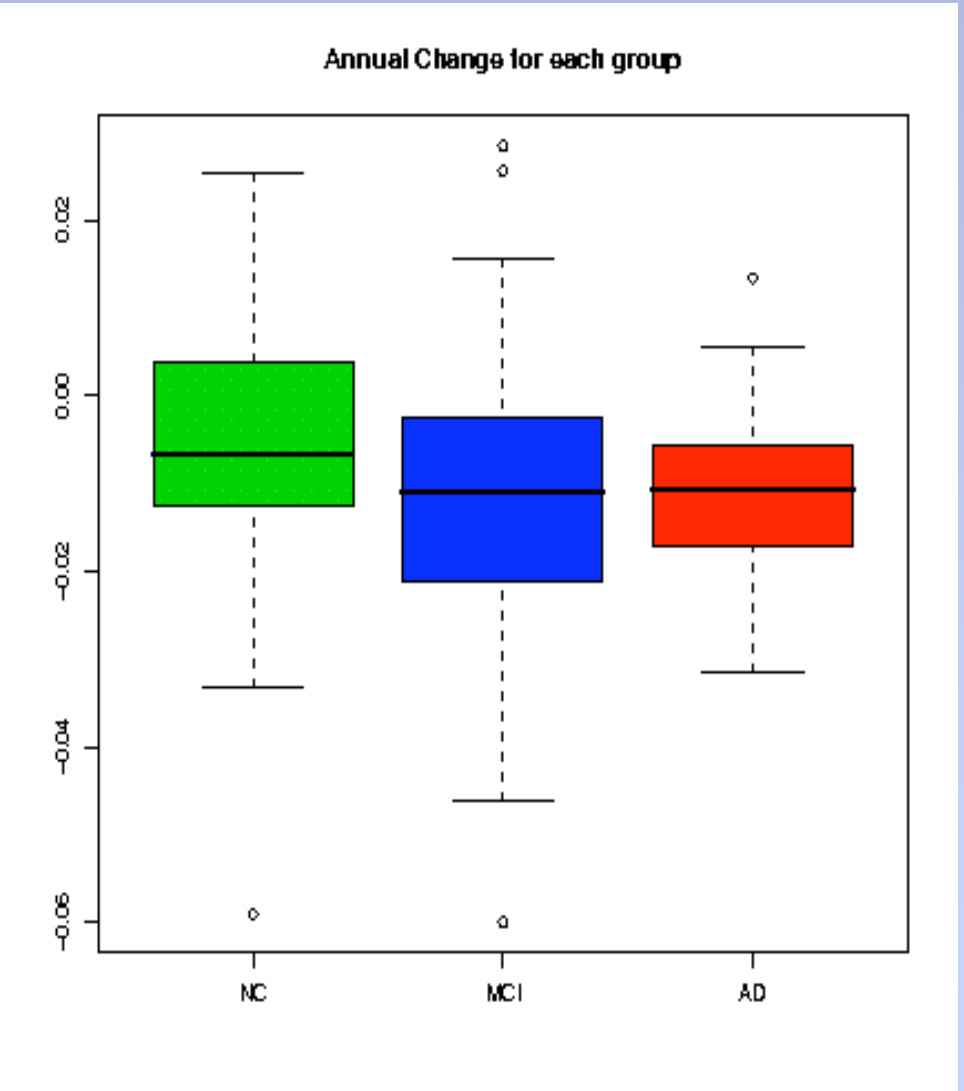


Another approach: Alexander lab, VBM Lt Hippocampal Formation, looking at differences from baseline to month 12.

Green=NC

Blue=MCI

Red=AD

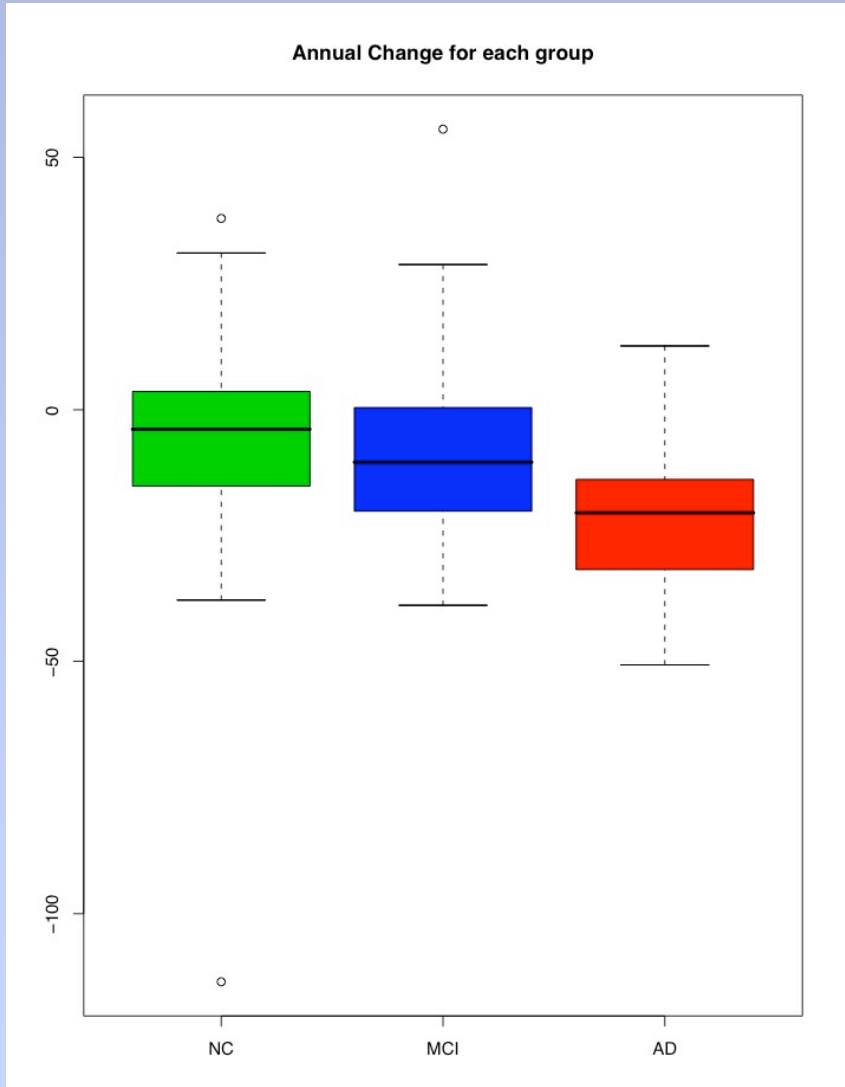


Thompson lab: a tensor based morphometry (TBM) approach: average Jacobian in a ROI (temporal lobe) .

Green=NC

Blue=MCI

Red=AD



If you designed our same study with imaging measures, how many subjects per arm to detect a 25% reduction in the change per year? (Based on differences; numbers differ a little using estimates based on mixed models.)

Lab (PI):	Measure (ROI)	Subjects per arm
Alexander	Lt Hippocampal Formation	149
Dale	Whole brain volume	43
Fox	% change in whole brain volume (using BSI)	76
Schuff	Hippocampus (total)	151
Studholme	Average Jacobian (temporal lobe)	78
Thompson	Average Jacobian (temporal lobe)	123

PET labs: We asked each lab to identify key regions of interest or summaries for us to analyze for this meeting.

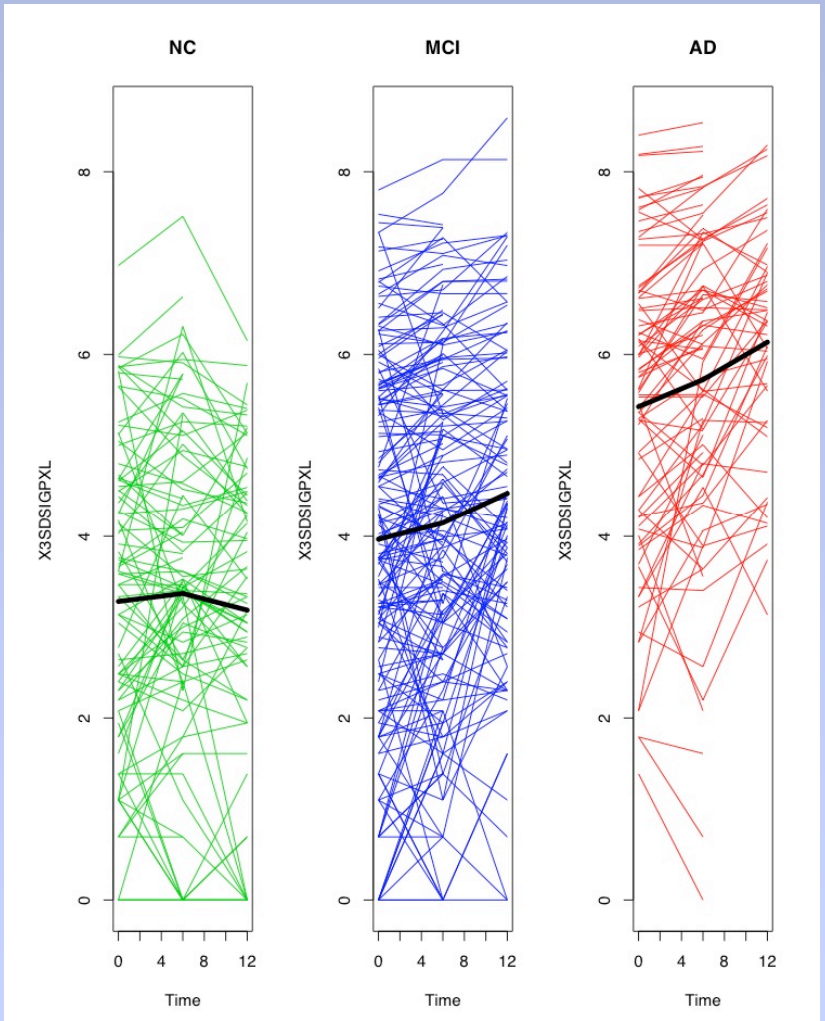
PET lab (PI)	# NC	# MCI	# AD	# ROI
Foster	77	118	53	4
Jagust	79	130	57	5
Reiman	--	112	49	6 AD, 4 MCI

Change plots, Foster lab: # pixels > 3SD below normal control mean glucose metabolism.

Green=NC

Blue=MCI

Red=AD

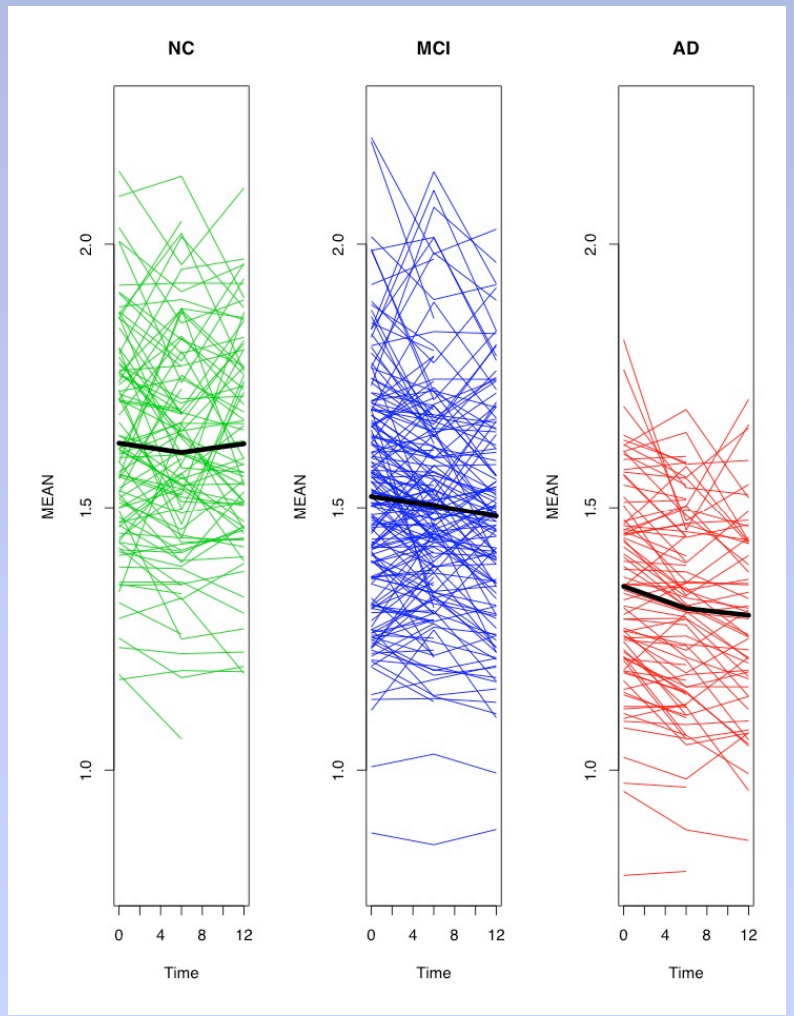


Change plots, Jagust lab: bilateral posterior cingulate (mean activity).

Green=NC

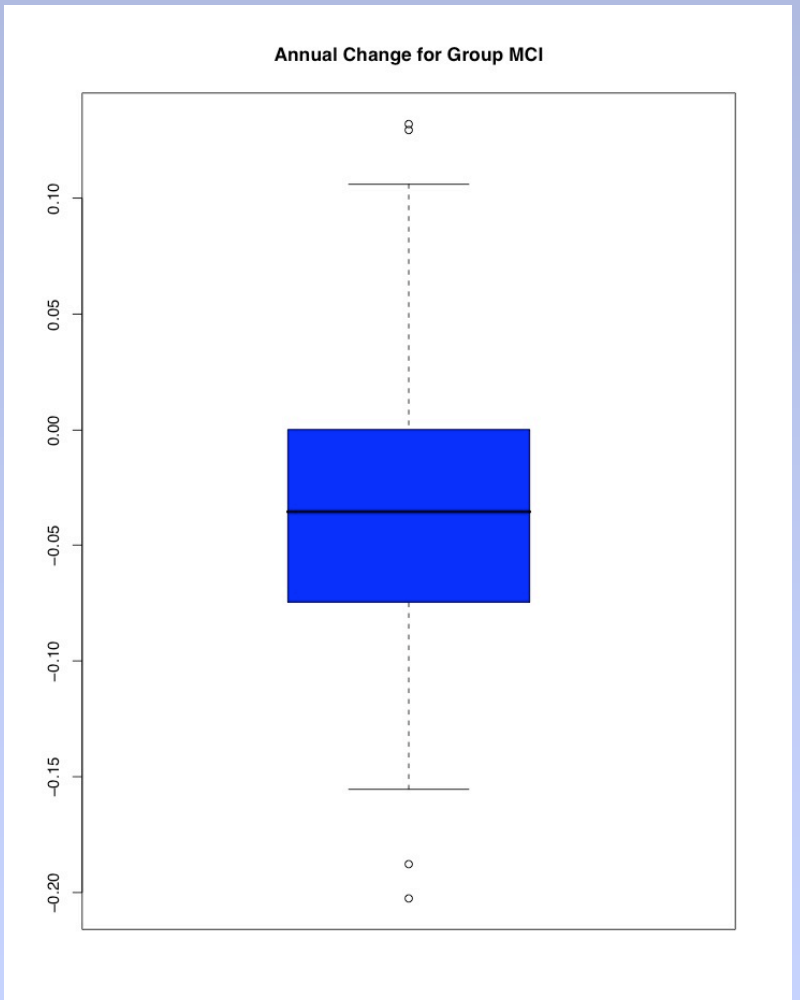
Blue=MCI

Red=AD



Reiman lab: Calculated change scores (bl to m 12) for posterior cingulate in the MCI group (BLUE).

Note that most people show a decline in activity.



Sample size calculations for selected PET measures: number to detect 25% reduction in annual change in AD

Lab PI	ROI/ Variable	Available data	Required n (per arm)
Foster	#pixels > 3 SD below NC mean	53	384
Jagust	Bilateral Posterior Cingulate	57	503
Reiman	Right Temporal Pole Superior	35	107

Most of the summary measures show some heterogeneity of changes within diagnostic group.

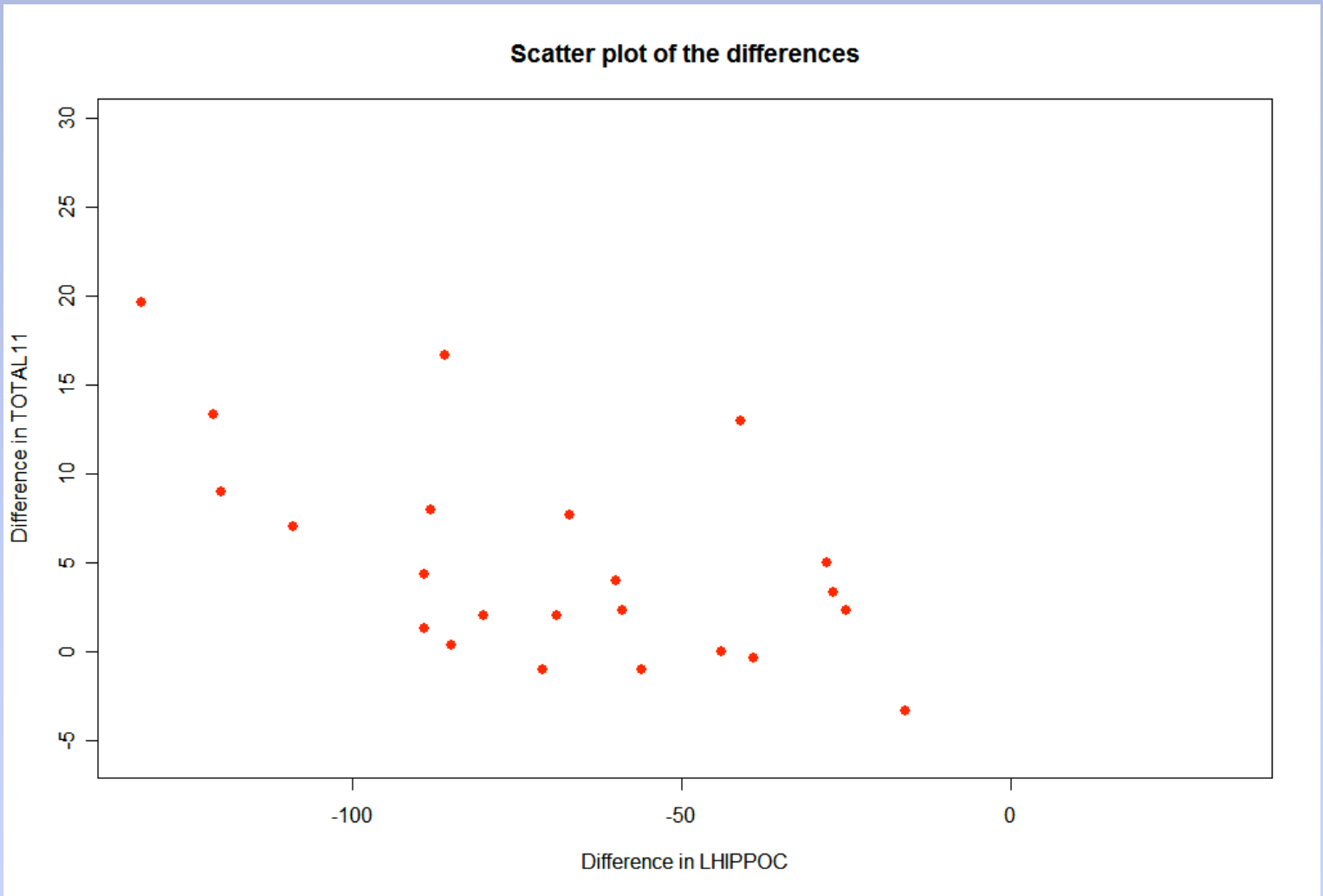
Does this correlate with heterogeneity in neuropsychological assessment change?

For example, people with MCI who decline most rapidly might show most change in imaging.

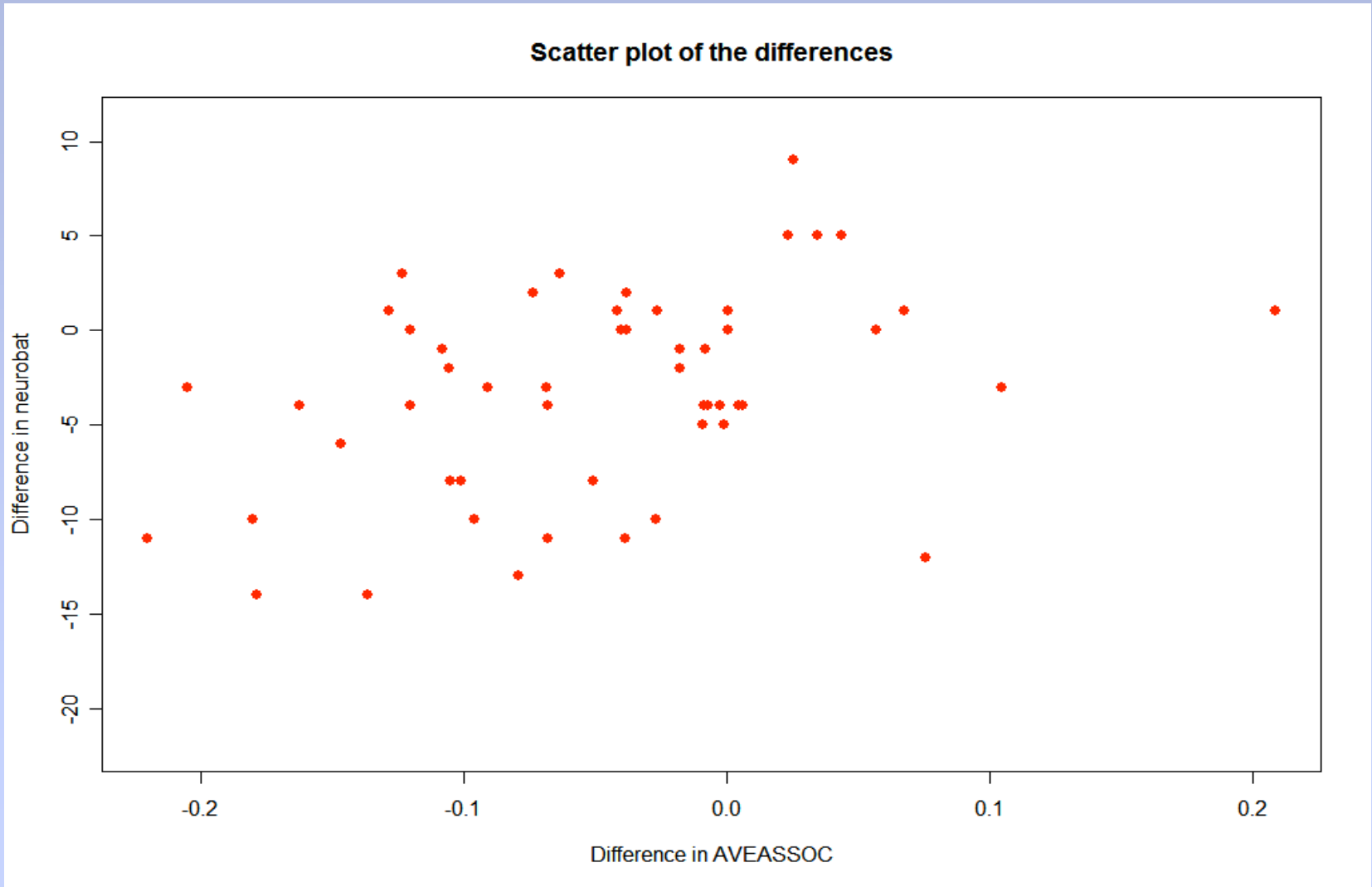
(CAUTION: correlations are attenuated by within-person variation.)

Generally, yes; some are strong even without adjusting for attenuation.

Scatterplot: left hippocampal change (Dale) vs. ADAS-COG Total 11 change, in AD, R=-0.58 (n= 24)



Scatterplot: Change in association cortex glucose metabolism (Foster) vs. change in word learning trials (AD), R=0.39, n=53.



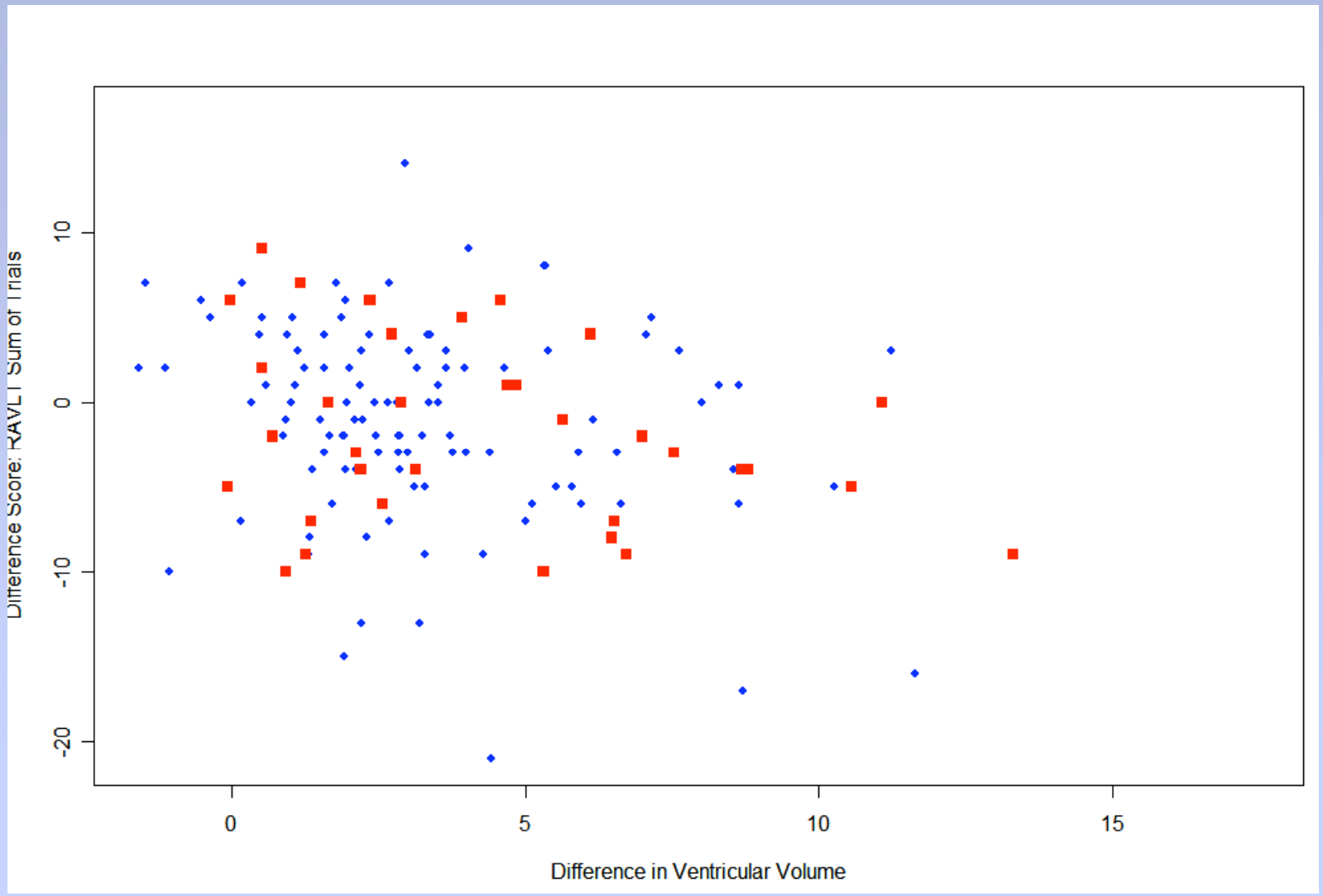
Selected correlations between change in image and change in neuropsych measure (not adjusted for within-person noise):

Diagnostic group	Neuropsych measure	Image measure	Correlation
MCI	CDR	Fox Δ Ventricle vol. (% whole brain)	0.38
		Ventricle vol.	0.35
		Jagust Rt Temporal	-0.32
	Word learning	Dale Lt Mid Temp	0.23
		Reiman Rt Inf. Parietal	0.28
AD	ADAS Cog	Dale Lt Hippocampus	-0.58
		Fox Δ Ventricle vol. (% whole brain)	0.45
	CDR	Dale Lt Mid Temp	-0.52
		Fox Δ Ventricle vol. (% whole brain)	0.42

An early look at MCI -> AD. Example: Longitudinal mixed effects model, testing whether rate of change in ventricular volume (Fox lab) differed for those who converted. Based on first n=237.

Variable	Estimate	P value
Baseline ventricular volume, non converters	46.9	<0.001
Annualized change, non converters	3.1	<0.001
Difference in baseline, converters	4.1	0.34
Difference in annualized change, converters	1.4	0.002

Scatter plot showing change in ventricular volume (Fox lab) vs. change in word learning in MCI group, with converters to AD highlighted.



Remarks about statistical analyses:

1. This is only a sample; more measures, more analyses, more pictures.
2. We have fitted some random effects models, results similar; 12 months still short to detect overall non linearity or difference in slopes (swamped by noise).
3. Correlations are attenuated by noise in clinical measurements. More sophisticated analysis to come.
4. Reminder: data for first year not complete yet; premature to compare across measures, labs!
5. Cross-validation is set up for labs that use data-driven approach to identify ROI, but not yet implemented.

The Biostatistics Core has been eager to get started with our analytic plan to assess the relationship between imaging and clinical data.

Imaging summary data were slow to come in:

But we now have vast quantities of data!

Summary of longitudinal analysis progress and challenges:

1. Clinical change is evident but with substantial within-person noise, some floor and ceiling effects. Hard to confirm between-person differences in rate of change.

    Should we consider composite clinical measures? If so, how should we decide which groups of measures we might combine?

    Simultaneous models will have better power because they adjust for measurement variation, but they take more data.

2. We have an abundance of imaging data now!

It's important for statistical validity to have guided hypotheses for the imaging summary approaches with many measurements.

It's also critical to work closely with labs as the measures and summaries are complex.

3. We have only had data for 3 weeks. We already have hundreds of pages of computer output.

Communicating with our colleagues is critical. We want to find efficient, clear ways to summarize and communicate results.

These results are still changing as the data are edited and the database grows. Versioning will be critical for comparing over time and across research groups.

The complexity of the data and the analyses should encourage us to work very closely together and share results, like this meeting!

ADNI is a big job, but we can do it together!

With thanks to:

Danielle, Qian and Hao at UC Davis,

John Kornak at UCSF,

Anthony Gamst and Mike Donohue at UCSD,

The ADCS group and LONI group for amazing data flow,

The Biostatistics group for discussions on Tuesdays,

Our colleagues in MRI and PET groups for sharing data and advice and interpretations,

Mike Weiner for support, feedback, and encouragement.