ISCTM
THE INTERNATIONAL SOCIETY FOR CNS CLINICAL TRIALS AND METHODOLOGY

A data driven approach to developing composite endpoints

Clint Hagen, MS

Mayo Clinic

February 20, 2019

Disclosures

- No disclosures
- My work presented here was funded through unrestricted research grants from:
  - Biogen - Pre Clinical AD
  - Lundbeck – MCI/Early AD

Overview

Introduction

Background

Statistical methodology to derive PACC-R

Evaluating PACC-R in Pre-clinical AD

Extending to MCI/Early AD - EMACC

Introduction

Per the FDA Guidance for Industry, composite endpoints will be considered in early Alzheimer's disease (AD) clinical trials

Composites already exist, but could be less than ideal for pre clinical AD trials

- Waterfall problems from ceiling effects
- Developed on symptomatic populations
- Comprised of measures thought to be 'best' might not be sensitive in pre clinical AD

What if we could develop a composite of cognitive tests which is sensitive amyloid related cognitive decline in pre clinical AD patients?

- A composite derived from real-world data on pre clinical individuals
- Replicated across multiple datasets in parallel
- Driven by the data, not opinion
- Useable in global clinical trial settings

PACC-R* is just that

*Hassenstab J, Hagen C, Han B, Lim Y, Maruff P, Mielke M, O'Gorman J, Stricker N, Jaeger J. Reliability and Reproducibility of Cognitive Composites for Alzheimer's Disease Secondary Prevention Trials: The Power-PACC. Poster Presented at: The International Conference on Alzheimer's and Parkinson's Diseases; 2017 March 29-April 2; Vienna, Austria

PACC-R

- A data driven approach building on the ADCS-PACC*
- Collaborative effort, datasets from four studies included
  - Charles F. & Joanne Knight Alzheimer's Disease Research Center (Knight ADRC)
  - Alzheimer's Disease Neuroimaging Initiative (ADNI)
  - Australian Imaging, Biomarkers and Lifestyle Study (AIBL)
  - Mayo Clinic Study of Aging (MCSA)
- Curated by Cognition Metrics

*Donohue, M. C., Sperling, R. A., Salmon, D. P., Rentz, D. M., Raman, R., Thomas, R. G., . . . Alzheimer's Disease Cooperative, S. (2014). The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol, 71*(8), 961-970. doi:10.1001/jamaneurol.2014.803

PACC-R

Instead of a handful of NP tests thought to be most closely associated with AD , we started with a large list of cognitive tests

- MCSA dataset started with 32 potential tests
- No composites, diagnostics, or potential waterfall effects
- No paragraph recall or other tests which pose significant challenge in translating to multiple languages and administering across cultures
- After applying these criteria the Mayo dataset had 13 remaining candidate measures for the composite

Domain	Construct/ Paradigm	HABS		Mayo		AIBL		ADNI		Standard WU ADRC Battery for participants enrolled before	
		Test	Variable	Test	Variables	Test	Variable	Test	Variable	Test	Variable
Memory	Word list	FCSRT (Picture or word?) Memory Capacity Test (Bushke) 6-trial selective reminding test	total free recall (could consider cued recall or total recall)	Rey Auditory Verbal Learning Test + 30 min delay and recognition	30 min delayed recall; Learning 1: Trials 1-5 Total (0-75), Learning 2: Trials 1-5 Total + List B (0-90) Learning 3: Trials 1-3 Total Delayed	CVLT	T-score learning (1–5) Short delay free recall Z-score L ong delay free recall Z-score;	RAVLT	Trials 1-5 (could consider trials 1-3); Delayed recall, recognition (if given), composite of immediate and	FCSRT – SRT Free (Picture or word?)	total free recall (3 trials) (Should we look at cued also?)
	Verbal Paired									WMS-Associate	raw score (ACSS
	Paragraph Recall	Logical Memory	?	WMS R LM	WMS R Logical Memory IIa (0-25) (Current PACC Donahue; different than what we typically include in Mayo studies, only the Anna Thompson	Logical Memory : Anna Thompson	Recall immed, raw score Recall Delayed raw score, Pass/Fail	WMS R LM	Same as for Mayo	WMS R Logical Memory Story A, Craft Story 21 Recall Immediate, Delayed , Time elapsed, Cued	for WMS-R, same as Mayo. Jason can you advise on Craft story 21?
	Design (spatial) memory	?FNAME		Wms-R Vis Repro II	WMS-R VR Immediate WMS-R VR Delay WMS-R VR						
Language	Naming	Boston Naming test	raw score no cue; Even	Boston Naming Test	raw score no cue	Boston Naming Test, no cue AU	raw score, no cue	Boston Naming Test	raw score no cue	Boston Naming Test	raw score, no cue
Executive	Verbal Fluency	FAS Categories (3)		Category fluency-Animals, fruits, vegetables??	Total animals + fruits + vegetables Animals only	FAS; Animals and Boys names; Category switching	FAS total Z-score, Animals and Boys names—sum of two,	Category Fluency: Animals	number correct z-score	Category fluency-Verbal Fluency:	Categories: z-score animals only, z-score animals plus vegetabl
	Trailmaking	Trails		Trails A & B	time to complete A; Bminus/V B+A			Trailmaking Test	Same as Mayo	Trailmaking A &B	Same as Mayo
	Coding	DSST		WAIS-R DSST	N correct- z transformed	WAIS-R DSST	Same as Mayo	DSST (only available in ADNI-1,	Same as Mayo	WAIS-R DSST	Same as Mayo
	Stroop	Computerized Flanker	version?			Stroop with colored dots, Black words, Colored	Dots Z-score Words Z-score C olors Z-score C /D Z-			Color only stroop, Stroop Switch	Jason is your version same as AIBL? Would you plan
	Letter number sequencing	Letter Number Sequencing + Computerized								WAIS III Letter Number Sequencing	Total correct responses
Attention	Digit Span	Digit Span				Digit span-SS	One combined score of digits forward	Digit Span forward / Back (ADNI-1	z score digits forward , z-score digits	WMS-R Digits forward	
Visuospatial	Block Design	x	x	WAIS-R BD	Raw score	x	x	x	x	Block Design	raw score
	Picture Completion			WAIS-R Pic Completion	Raw score						
	Complex figure	x	x	x	x	Rey Complex Figure test:	Copy Z-score S hort Recall Z-score L ong Recall Z-score R	x	x	Benson Complex Figure Copy Immediate and delayed	Jason can you pls specify variables ideally to match
	Clock Drawing					Clock draw score		Clock drawing			
	Line orientation										
	Form Discrimination	Benton Visual Form Discrim Test									
Motor										Tapping Task	
Global		AMNART MMSE		Short test of mental status (Proxy MMSE	Kokmen total score (label MMSE but is Kokmen short test). Also Kokmen Orientation subscale (0-8	WTAR estimated IQ (Wechsler test of adult reading, US norms)		MMSE ADAS-	Total, Orientation items alone	WAIS-R Information	z-total raw score?
				Blessed							
	CDR	CDR		CDR Sum				CDR		CDR	
Computerized		Cogstate??, FNAME, Color CAP testing-computerized		Cogstate – started in Feb 2012 on the 50-69 year olds;		Detection task reaction time Identification task reaction					

ISCTM
THE INTERNATIONAL SOCIETY FOR CNS CLINICAL TRIALS AND METHODOLOGY

	Cognitive Measures			
Cognitive Domain	AIBL	ADNI	Knight ADRC	MCSA
Global	MMSE	MMSE	MMSE	Kokmen Short Test of Mental Status
		ADAS-COG Orientation		
		Montreal Cognitive Assessment		
Episodic Memory	WMS-R Logical Memory	WMS-R Logical Memory	WMS-R Logical Memory	WMS-R Logical Memory
	California Verbal Learning Test, Second Edition	Rey Auditory Verbal Learning Test	Free and Cued Selective Reminding	Rey Auditory Verbal Learning Test
	Rey Complex Figure Test, Delayed Recall	ADAS-COG Word Recall	WMS Associate Learning	WMS-R Visual Reproduction
Semantic Memory	Boston Naming Test	Boston Naming Test	Boston Naming Test	Boston Naming Test
	Letter Fluency	Animal Naming	Animal Naming	Animal Naming
	Animal Naming	Vegetable Naming	Vegetable Naming	Vegetable Naming
	Fruit Naming		Letter Fluency	Fruit Naming
	Boy's Names		WAIS Information	
	Furniture Naming			
Processing Speed	WAIS-R Digit Symbol	WAIS-R Digit Symbol	WAIS-R Digit Symbol	WAIS-R Digit Symbol
		Trailmaking Test Part A	Trailmaking Test Part A	Trailmaking Test Part A
		Trailmaking Test Part B	Trailmaking Test Part B	Trailmaking Test Part B
Visuospatial	Rey Complex Figure Test, Copy		WAIS Block Design	WAIS-R Block Design
	Clock Drawing			WAIS-R Picture Completion
Working Memory/Executive	WMS-R Digit Span Forward		WMS-R Digit Span Forward	
	WMS-R Digit Span Backward	WMS-R Digit Span Forward	WMS-R Digit Span Backward	
	Stroop	WMS-R Digit Span Backward	WAIS-III Letter-Number Sequencing	

The idea was to set up each dataset to be as close to a pre-clinical AD clinical trial where amyloid is the target of interest as possible

- Required cognitively unimpaired (CU) at baseline with known amyloid status (AB)
  - Each site used their own cognitive diagnosis (also required CDR = 0) and PET AB measure
- Required at least one follow-up visit after baseline with cognitive testing
- CU and AB+ (SUVR 1.4+) are the pre clinical AD group, the individuals who will experience AB related cognitive decline
- CU and AB- are the “responders”, the individuals who don’t have AB related cognitive decline

	AIBL		ADNI		Knight ADRC		MCSA	
	AB Negative	AB Positive	AB Negative	AB Positive	AB Negative	AB Positive	AB Negative	AB Positive
n	207	88	85	48	171	92	373	240
Age, y	70.9 (4.7)	73.5 (5.4)	74.75(4.2)	75.6(4.1)	71.0 (5.8)	73.4 (5.2)	74.6 (5.1)	76.4 (5.2)
Sex (% Female)	48.8	46.6	48.2	43.7	63.5	52.2	42.9	43.3
Education, y	13-15 years	13-15 years	16.2(2.7)	15.9(2.5)	15.5 (2.7)	15.7 (2.6)	14.8 (2.7)	14.6 (2.7)
APOE e4 alleles, n, %								
0	174 (84.1)	43 (48.9)	71(83.5)	28(58.3)	118 (74.2)	28 (41.8)	303 (81.5)	149 (62.3)
1	33 (15.9)	42 (47.7)	13(15.3)	19(39.6)	39 (24.5)	32 (47.8)		
2	0 (0.0)	3 (3.4)	1(1.2)	1(2.1)	2 (1.3)	7 (10.4)	69 (18.5)*	90 (37.7)*
MMSE	28.9 (1.2)	28.8 (1.2)	29.1(1.0)	29.1(1.1)	29.0 (1.2)	28.7 (1.6)	N/A	N/A
CDR Sum of Boxes	0.0 (0.2)	0.0 (0.1)	0.02(0.10)	0.01(0.07)	0.0 (0.1)	0.1 (0.2)	0.0 (0.2)	0.0 (0.2)
Amyloid PET, SUVR	1.15 (0.10)	1.93 (0.37)	1.00(0.06)	1.34(0.7)	1.01 (0.2)	2.46 (0.7)	1.3 (0.0)	1.8 (0.4)

*Either or both e4 alleles

Variable Selection

Variable selection methodology has evolved with the availability of increased computing power at out fingertips. Least absolute shrinkage selection operator (Lasso)* is one option. Lasso is a form of penalized regression very useful in selecting variables for model inclusion.

- I assured everyone we could make this work
- By the way, lasso doesn't work to select dependent variables
- Furthermore, lasso doesn’t work on longitudinal data

"If somebody offers you an amazing opportunity but you are not sure you can do it, say yes – then learn how to do it later!" – Richard Branson

*Tibshirani, Robert. 1996. “Regression Shrinkage and Selection via the lasso”. Journal of the Royal Statistical Society. Series B (methodological) 58 (1). Wiley: 267–88. http://www.jstor.org/stable/2346178

Why Lasso?

- Lasso is a data driven variable selection method
- It is tolerant of multicollinearity
- Very effective in avoiding overfitting

To make lasso work we needed to flip the regression equation. What would have been amyloid status predicting the composite outcome, we now have the list of cognitive tests predicting amyloid status as a logistic regression type problem.

Each dataset standardized candidate cognitive measures into a z-scores

- Baseline mean and standard deviation used for follow-up visit z-score calculations

Three desirable trial lengths: 36, 48, and 60 months

- Annualized rate of change calculated for each cognitive measure z-score, at each of three time points
- MCSA visits are at 15 month intervals
  - (30 month – baseline) / years since baseline
  - (45 month – baseline) / years since baseline
  - (60 month – baseline) / years since baseline

Missing data were imputed prior to lasso using random forest imputation*

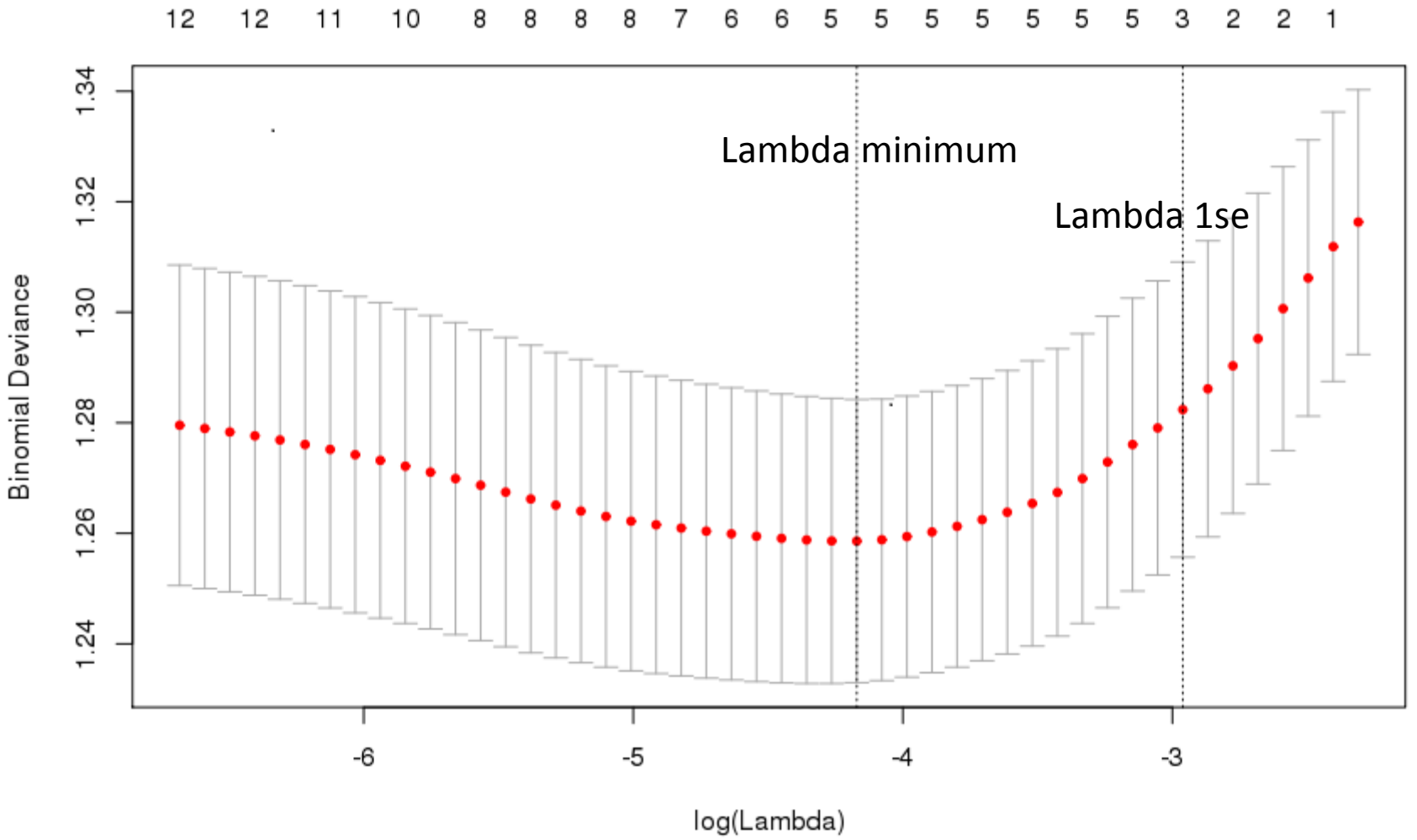
- This routine doesn't accept missing data, missing data weren't the driving factor for imputation
- Random forests (RF) perform well in high dimensions, complex interactions, and are non-parametric
- Trains a RF on the existing data then predicts the missing iteratively

Optimal tuning parameter/penalty determined using cross validation (10 fold) in the glmnet R package+

- Ran the three trial scenarios separately
- Optimal penalty applied and lasso selected the subset of cognitive measures for inclusion in the composite
- Lambda 1se was used

*Stekhoven, D.J. and Buehlmann, P. (2012), 'MissForest - nonparametric missing value imputation for mixed-type data', Bioinformatics, 28(1) 2012, 112-118, doi:10.1093/bioinformatics/btr597.

+ "CRAN - Package glmnet". *r-project.org*.


12 12 11 10 8 8 8 8 7 6 6 5 5 5 5 5 5 5 3 2 2 1
Lambda minimum
Lambda 1se
Binomial Deviance
1.24
1.26
1.28
1.30
1.32
1.34
-6
-5
-4
-3
log(Lambda)

Creation of a Composite

- A lot of reasons exist to include or exclude MMSE
  - We did run lasso selection with and without
  - Ultimately we examine the effect of MMSE explicitly
- Lasso selected variables across datasets show remarkable overlap, but not a perfect match
- Selected variables synthesized to satisfy the same paradigms in all datasets
  - Small amount of human intervention to “massage” into a single list of variables for each dataset

		FINAL PACC-R MEASURES IN EACH COHORT							
		MCSA		AIBL		ADNI		WUSTL	
Domain	Paradigm	Test	Measure	Test	Measure	Test	Measure	Test	Measure
Episodic Memory	Word list learning	RAVLT	Trials 1-5	CVLT	Trials 1-5	RAVLT	Trials 1-5	FCSRT	Free Recall
			30 min delay		Delayed recal	RAVLT	30 min Delay		
			Rcognition hits		True Pos		Recognition hits		
Semantic Memory	Fluency	Category fluency	Animals+fruits+ vegetables	Category fluency	Animals+Boys names	Category fluency	Animals+ Vegetables	Category fluency	Animals+ Vegetables
Executive	Trailmaking	TMT	Trails B (secs)			TMT	Trails B (secs)	TMT	Trails B (secs)
	Conflict Inhibition			Stroop	Color naming speed (secs)				
	Coding	DSST	Total score	DSST	Total score	DSST	Total score	DSST	Total score

Calculate the Composite(s)

The PACC-R composite is simply the mathematical mean of the available measures

$$\text{PACC-R} = \frac{Sum\ of\ test\ z{-}scores}{Number\ of\ tests\ available}$$

MMSE remains a highly regarded measure, we created an additional composite including it for comparison purposes

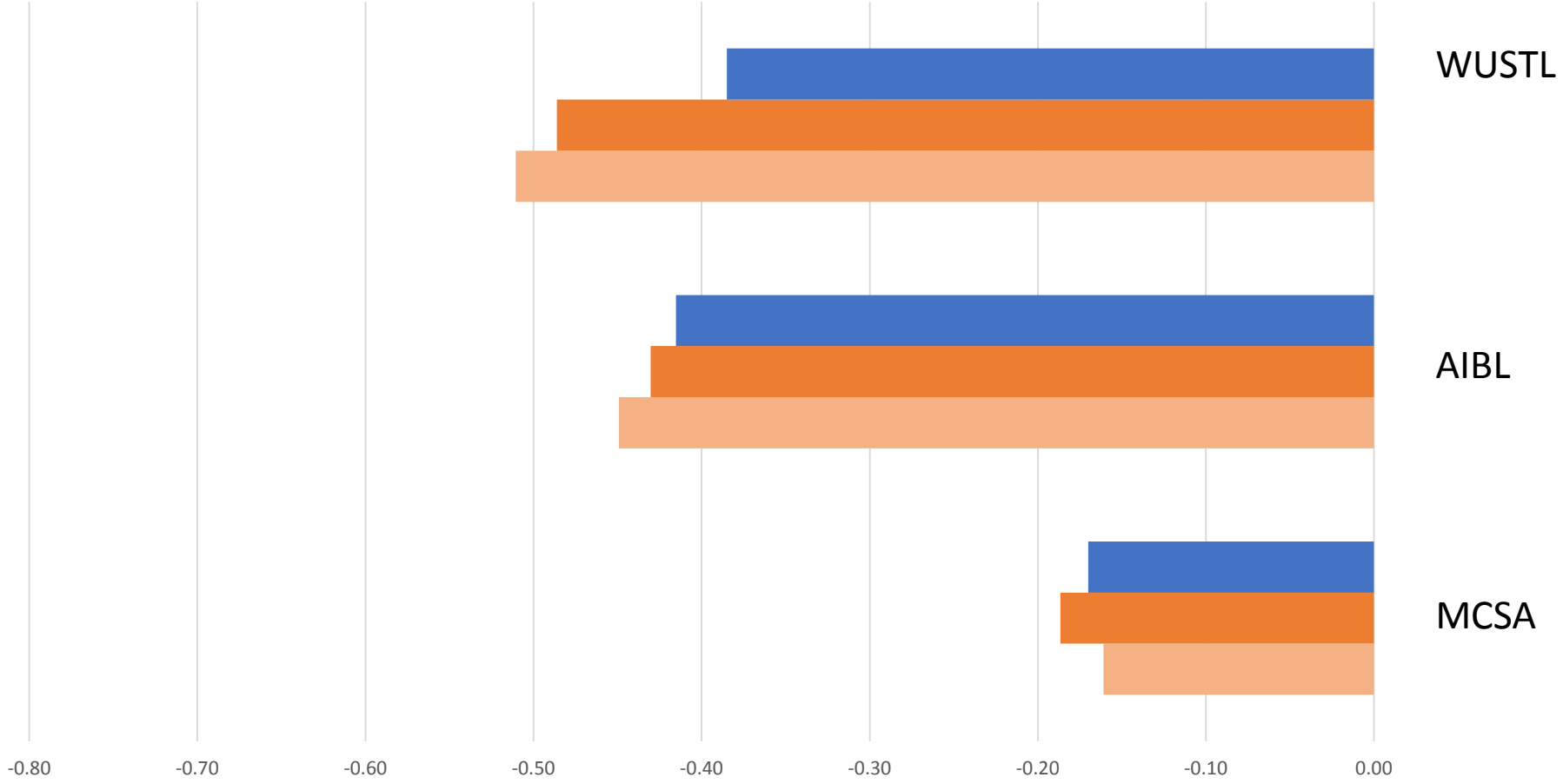
$$\text{PACC-R+MMSE} = \frac{Sum\ of\ test\ z{-}scores\ and\ MMSE\ z{-}score}{Number\ of\ tests\ available}$$

Composite Performance

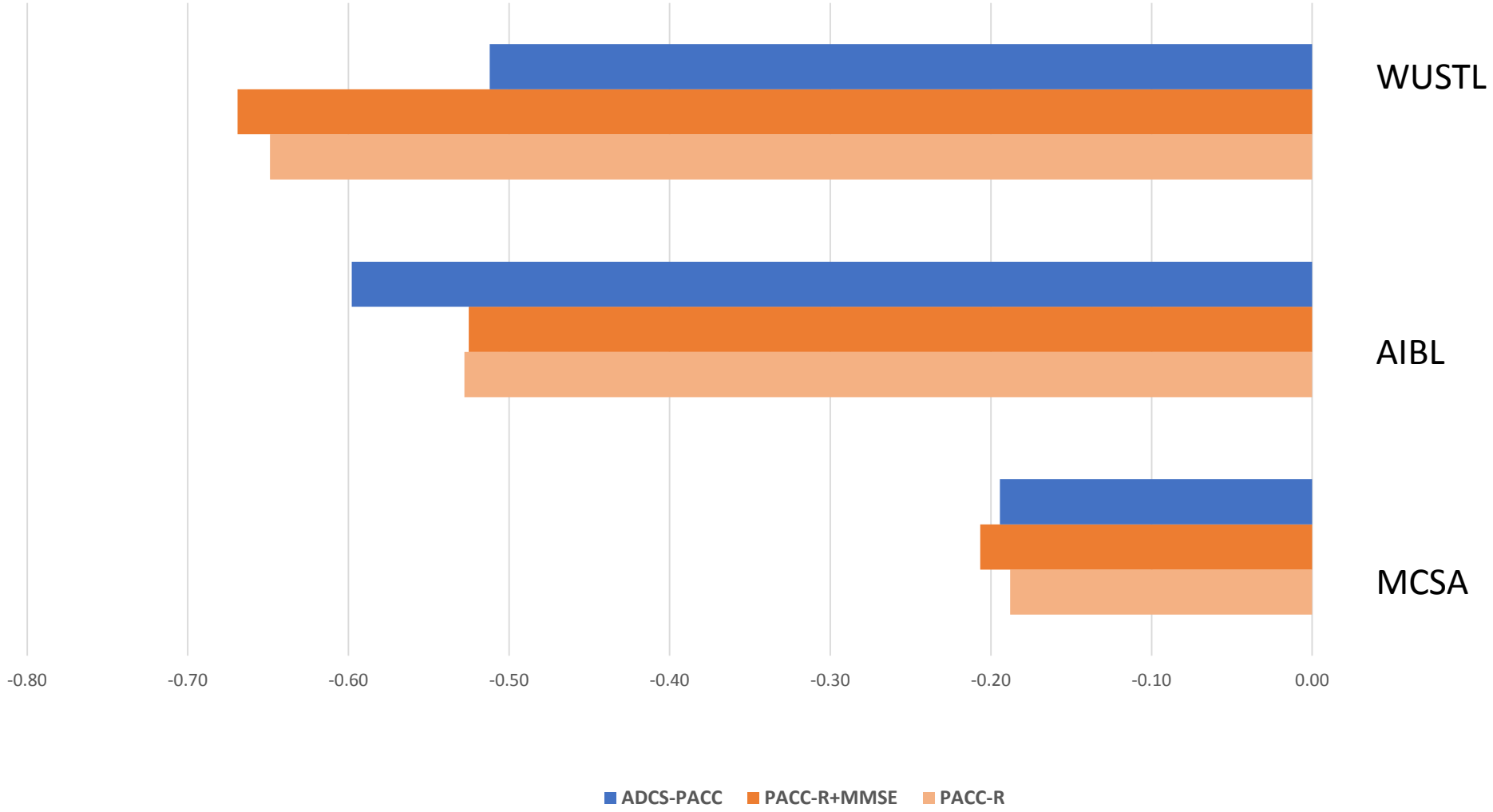
- Compared performance of PACC-R, PACC-R+MMSE and ADCS-PACC in the AB+ and AB-
  - ADCS-PACC calculated as described in the literature
- Hierarchical models (linear mixed effects models) used on longitudinal data
- No imputed data, random intercepts, random slopes
- Effect size calculated, difference between slopes of AB+ and AB-
  - Modified Cohen's D

$$\text{Effect Size} = \frac{Slope\ Difference}{Standard\ Deviation\ of\ Slope\ Difference}$$

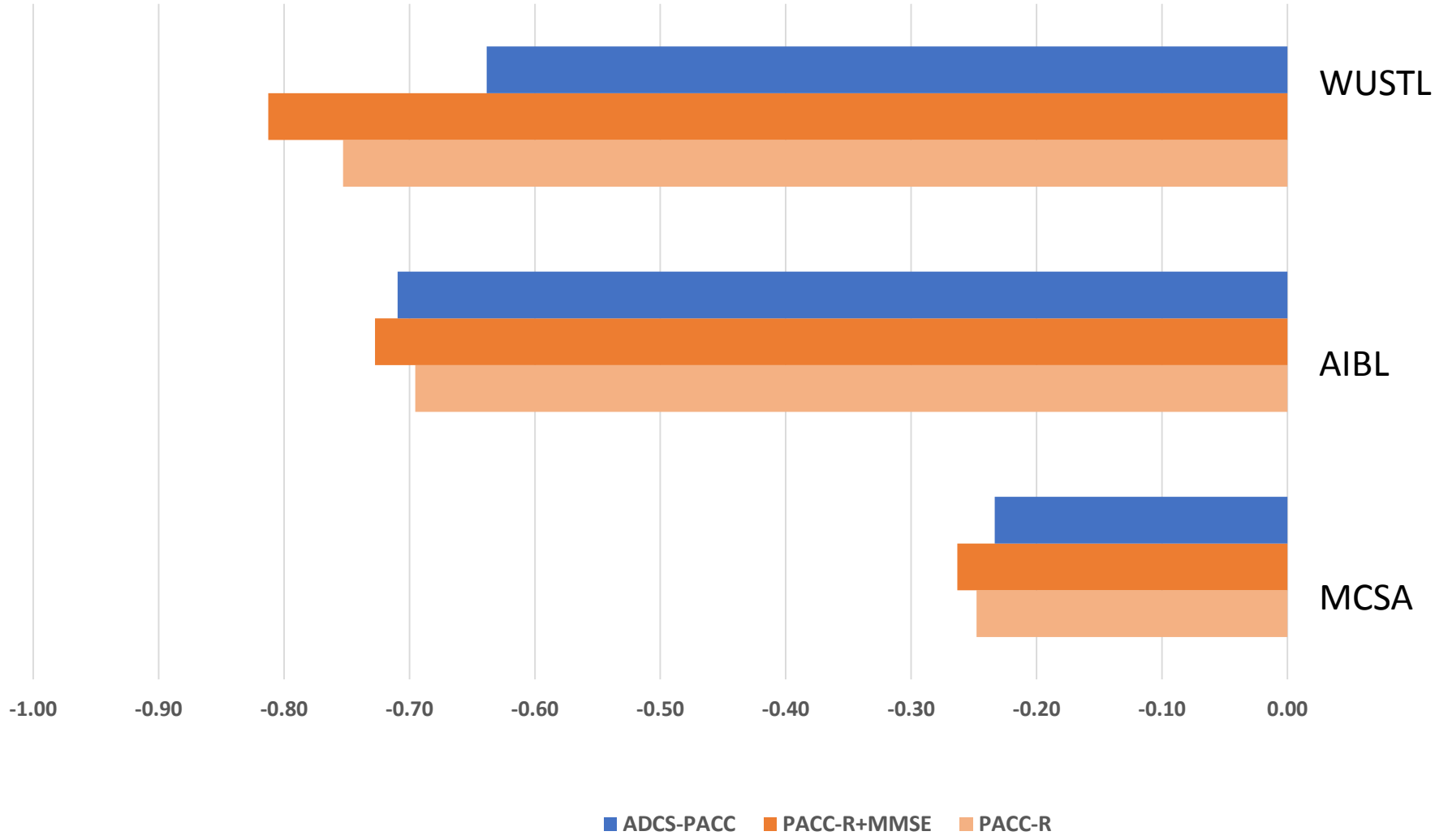
Slope Difference Effect Size in PACC at Year 3



Slope Difference Effect Size in PACC at Year 4



Slope Difference Effect Size in PACC at Year 5



PACC-R Conclusion

- Replication!
- Easy to simulate outcome groups
  - AB+ is normal AB related cognitive trajectory
  - AB- would be responders to an AB drug
- Data driven approach to assess AB related cognitive decline

- Does this extend to MCI/Early AD?

MCI and Early AD

The same four datasets in a separate effort tried extending the lasso methodology to an MCI/Early AD sample

- Met with significant challenges to using lasso
  - In early AD no "responder" group exists to simulate an effective AB drug since early AD w/o AB+ isn't the right pathway, had to use CU AB-
  - Lasso couldn't discriminate these two groups with the same subset of cognitive tests
  - Floor effects now where MCI/EAD aren't declining in some measures
- Instead we modeled all permutations of the available cognitive tests (2-7 in a model) in the MCI/EAD AB+ group in LMM
- This work was presented 2018 ISCTM and the result is now known as EMACC*

*Jaeger J, Hagen C, Loft H, Lim Y, Aschenbrenner A, Segerdahl M, Tong G, Mielke M, Hassenstab J, Stricker N. Cognitive Endpoints for Early Alzheimer's Disease Trials: Development of the Early AD/ MCI Alzheimer's Cognitive Composite (EMACC). Poster presented at: International Society of CNS Clinical Trials and Methodology; 2018 February 21-23; Washington, DC