Applications to AD with Sample SAS Codes

Chengjie Xiong

Washington University

October 15, 2016

Database Reality in ADCs

1. Subjects have different number/length of follow-ups;
2. The time interval between two adjacent visits differs among subjects;
3. Longer cognitive follow-ups, much shorter biomarkers follow-ups;
4. Missing data may NOT be random.

Shared Longitudinal Statistics

1. To estimate the longitudinal rates of change (the slope);
2. To compare the slopes across groups;
3. To adjust for the effect of covariates in the estimation/comparison of slopes;
4. To estimate/compare nonlinear pattern over time

Linear Mixed Models for Longitudinal Data (in English)

Outcome at Time t =

Fixed Effects+Random Effects+Errors

SAS implements the model by specifying all three terms in PROC MIXED.

A Scientific Example

To compare the rate of cognitive change in autosomal dominant AD from the Dominantly Inherited Alzheimer Network (DIAN) and late onset AD (LOAD)?

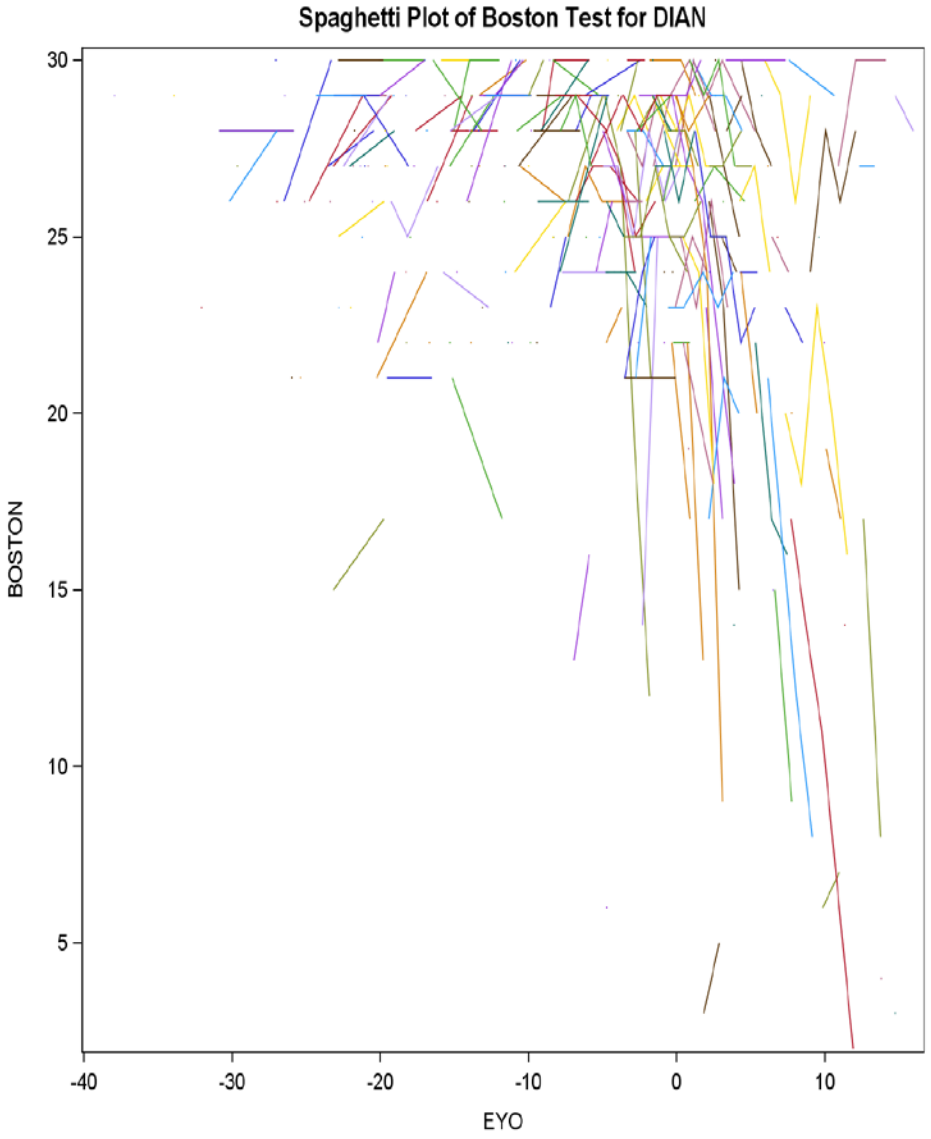
DIAN mutation carriers (amyloid precursor protein, presenilin 1 and 2);

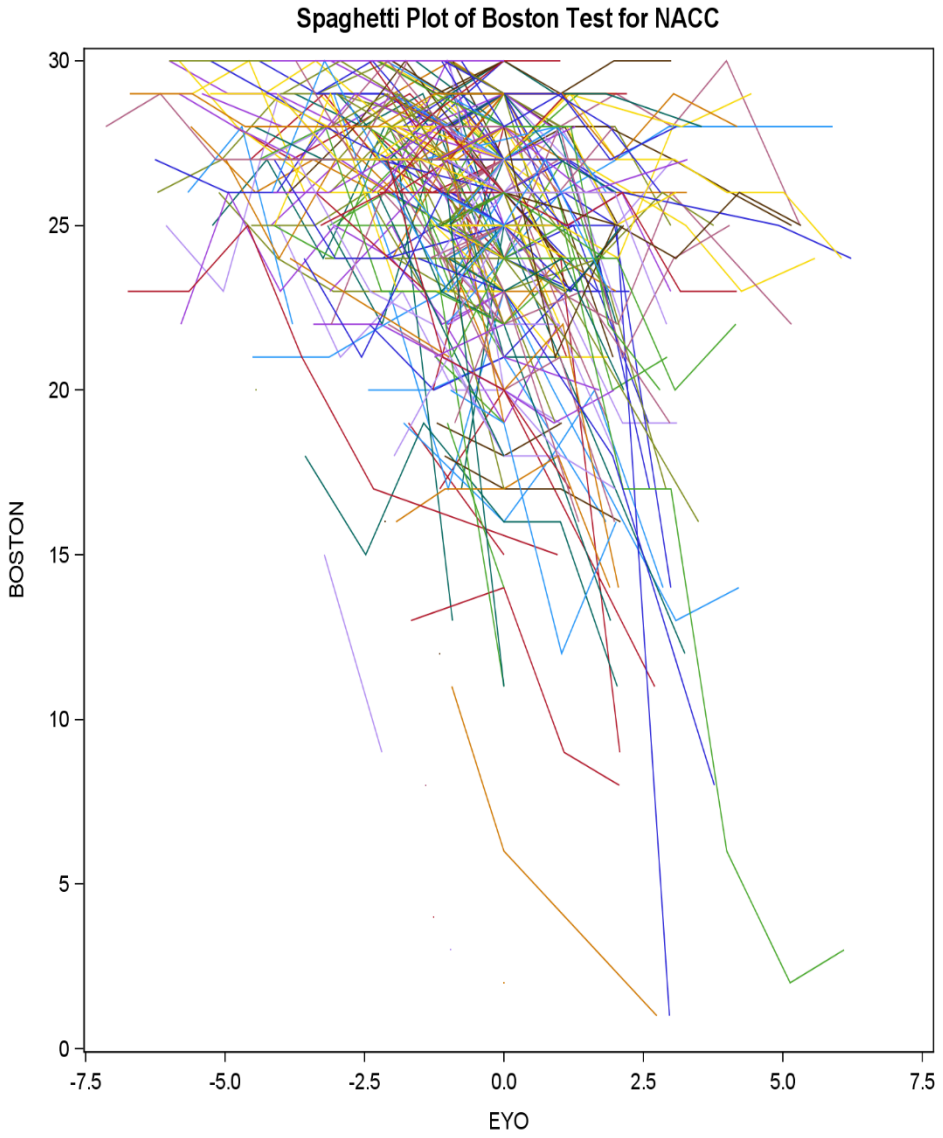
LOAD: NACC pathologically confirmed AD.

Cohort Description

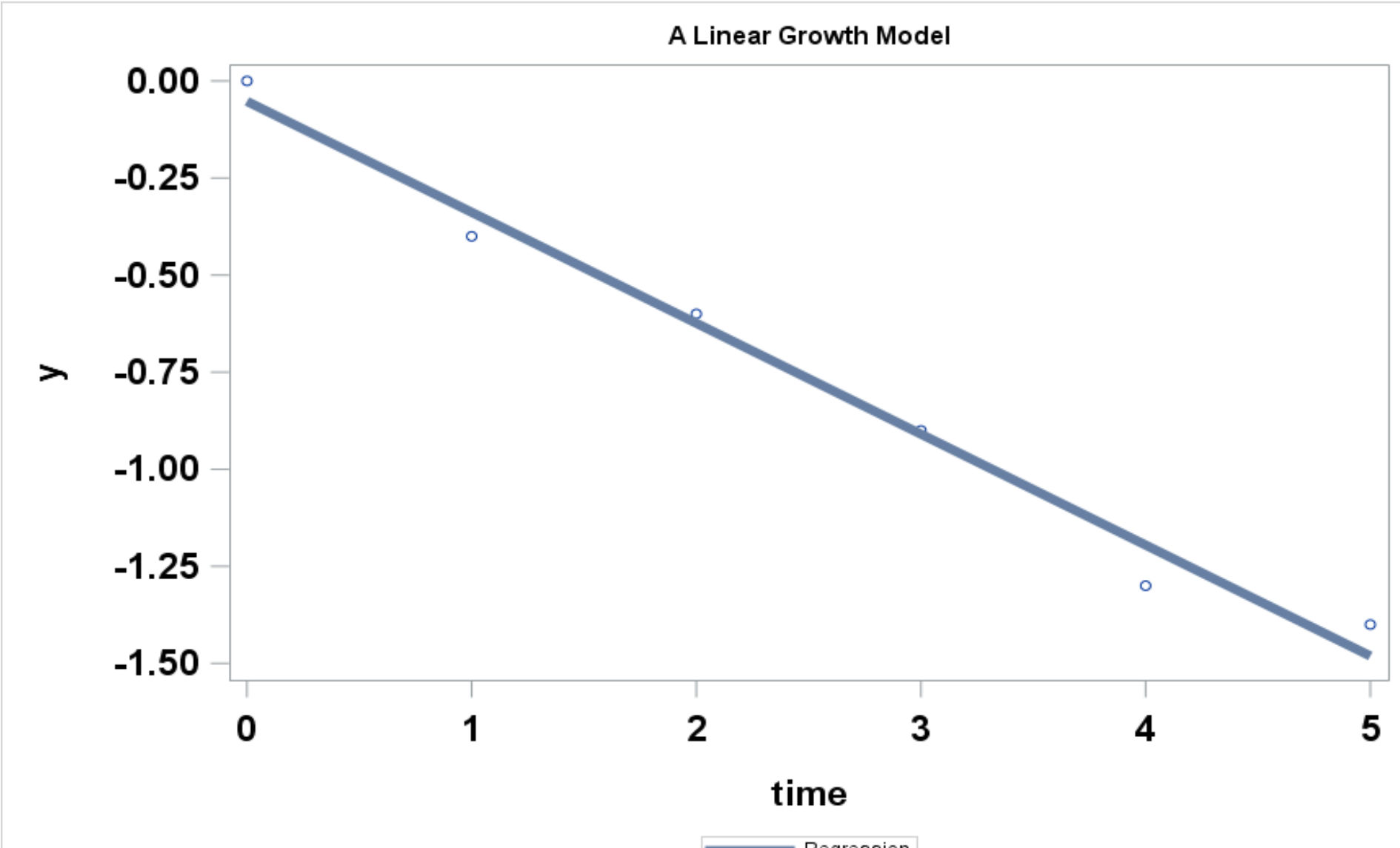
	NACC Participants N=239	DIAN Mutation Carriers N=129
Mean Age at Baseline (SD)	84.2 (8.9)	44.2 (8.0)
Mean Age at Symptom Onset (SD)	86.6 (9.01)	45.1 (9.7)
Mean Years of follow up (SD)	4.1 (1.8)	2.3 (1.2)
Gender (% Female)	61.5	62.0
Mean Years Education (SD)	15.4 (2.7)	13.6 (2.7)
Mutation (PSEN1:PSEN2:APP)	0	103 : 3 : 23
CDR at Baseline		
CDR 0 (%)	239 (100)	56 (43.4)
CDR 0.5 (%)	0	48 (37.2)
CDR 1 (%)	0	16 (12.4)
% ApoE4+ 1 E4	28.4	31.8
% ApoE4+ 2 E4	0.5	5.4

An Example of Spaghetti Plot





An Assumed Linear Growth Model



SAS Data Set Format

Data set format: (Time=EYO=expected years to onset, i.e., CDR>=0.5)

ID	Cohort	Time	Cognition
1	NACC	0.6	20
1	NACC	1.8	18
2	DIAN	0.9	22

…

SAS Codes to ESTIMATE Slopes/Intercepts

```
PROC MIXED DATA=A;
CLASS Cohort  ID;
MODEL Cog=Cohort Cohort*Time/noint s
ddfm=satterthwaite;
RANDOM Int Time/sub=ID type=un;
REPEATED  /sub=ID type=AR(1);
RUN;
```

SAS Outputs—2 Slopes and 2 Intercepts

Solution for Fixed Effects

Effect	Cohort	Est	SE	DF	t-value	Pr>|t|
Cohort	NACC					
Cohort	DIAN					
Cohort*Time	NACC					
Cohort*Time	DIAN					

SAS Codes to COMPARE Slopes/Intercepts

```
PROC MIXED DATA=A;
CLASS Cohort ID;
MODEL Cog=Cohort Time Cohort*Time/ddfm=satterthwaite;
RANDOM Int Time/sub=ID type=un;
REPEATED /sub=ID type=AR(1);
RUN;
```

Outputs—P-values for Testing Equality of Slopes (Intercepts)

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F value**	**Pr>F**
Cohort				
Time				
Cohort* Time				

Adding APOE4: To ESTIMATE Slopes/Intercepts

```
PROC MIXED DATA=a;
CLASS Cohort  ID E4;
MODEL Cog=Cohort*E4
Cohort*E4*Time/noint s ddfm=;
RANDOM Int Time/type=un sub=ID;
REPEATED /type=AR(1) sub=ID;
ESTIMATE ‘comparing slopes for E4-’
Cohort*E4*Time 1 -1 0 0/e;
RUN;
```

Outputs—4 Slopes+4 Intercepts

Solution for Fixed Effects							
Effect	**Cohort**	**E4**	**Est**	**SE**	**DF**	**t-value**	**Pr>|t|**
Cohort*E4	NACC	0					
Cohort*E4 …	NACC …	1 …					
Cohort*E4 *Time	NACC	0					
Cohort*E4 *Time ….	NACC	1					

Adding APOE4: To COMPARE Slopes/Intercepts

```
PROC MIXED DATA=A;
CLASS  Cohort  ID  E4;
MODEL Cog=Cohort   E4   Cohort*E4  Time
Cohort*Time   E4*Time
Cohort*E4*Time/ddfm=;
RANDOM  Int Time/type=un sub=ID;
REPEATED /type=AR(1) sub=ID;
RUN;
```

SAS Outputs

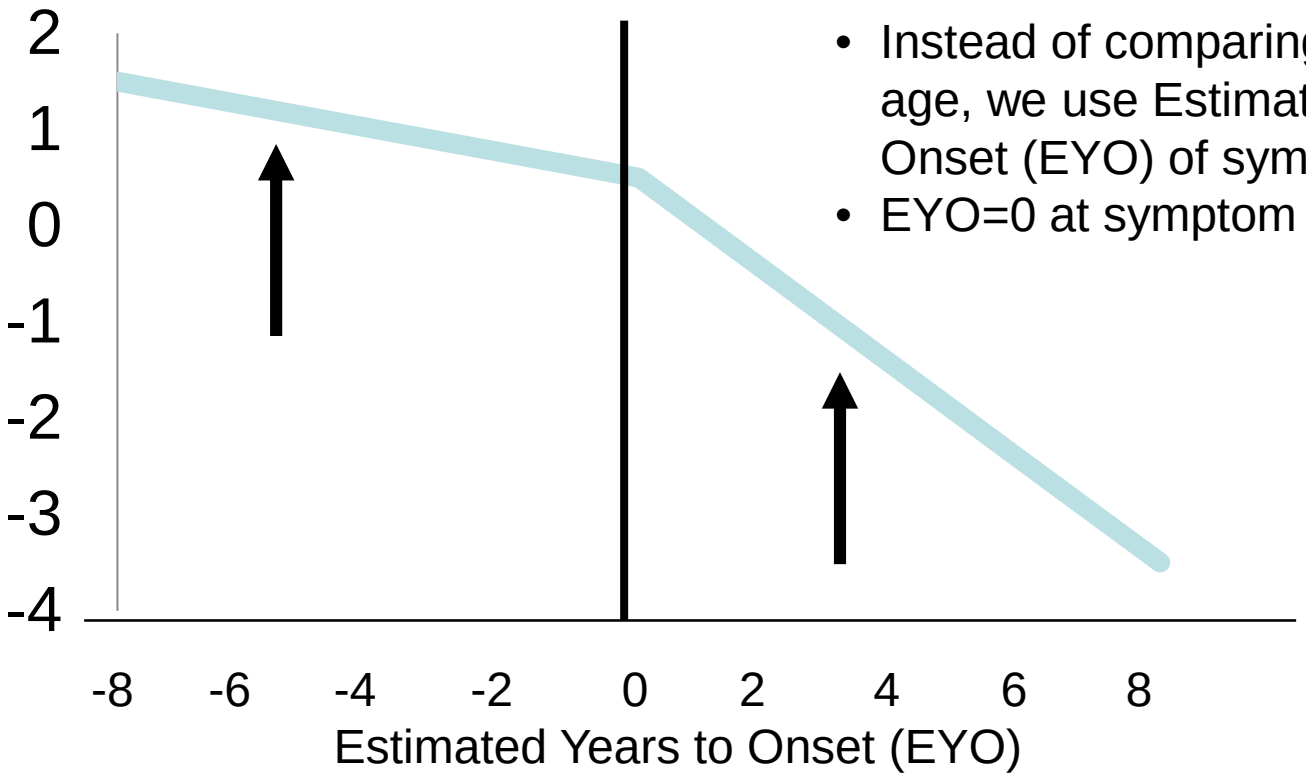
Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F value**	**Pr>F**
Cohort				
E4				
Cohort*E4				
Time				
Cohort*Time				
E4*Time				
Cohort*E4*Time				

Nonlinear Changes

Subjects in the NACC data set progressed to a higher CDR, justifying the nonlinear growth;

Subjects in the DIAN started at baseline with either EYO<0 or EYO>0.

One Possible Analytic Approach



- Instead of comparing groups on age, we use Estimated Years to Onset (EYO) of symptoms
- EYO=0 at symptom onset

Nonlinear Changes—Preparing SAS Data Set

Time1=EYO; and 0 if EYO>0;

Time2=EYO; and 0 if EYO<0;

Data set format:

ID Cohort Time1 Time2 Cog

….

To ESTIMATE Slopes and Intercepts

```
PROC MIXED DATA=A;
CLASS Cohort ID;
MODEL Cog=Cohort   Cohort*Time1 Cohort*Time2/noint s ddfm=;
RANDOM Int Time1 Time2/type=un sub=ID;
REPEATED  /TYPE=AR(1) SUB=ID;
RUN;
```

Nonlinear Changes: To COMPARE Slopes/Intercepts

```
PROC MIXED DATA=A;
CLASS  Cohort  ID ;
MODEL Cog=Cohort  Time1  Cohort*TIME1
TIME2   Cohort*TIME2/ddfm=;
RANDOM Int Time1 Time2/type=UN
sub=ID;
REPEATED  /type=AR(1) sub=ID;
RUN;
```

Results:
DIAN-NACC Comparison

To compare cognitive performance and rate of change between DIAN MC and LOAD from NACC with pathologically confirmed AD.

Slope Prior to EYO = 0

(Preclinical)

Cognitive Test (Mean annual change & SE)	NACC (n=239)	DIAN MC (n=129)	p-value
Animal Fluency	-1.07 (0.12)	-0.64 (0.21)	0.0752
Boston Naming			**0.0073**
Digits Backward			0.5139
Digits Forward			0.2803
Logical Memory			0.1017
Logical Memory			0.1737
MMSE			0.1466
Trails A			**0.0212**
Trails B			**0.0030**
WAIS-R Digit Sy			0.2758
Composite (All above)	-0.13 (0.01)	-0.07 (0.02)	**0.0050**
Non-speed Composite *	-0.09 (0.01)	-0.06 (0.02)	0.0709

*Non-speed Composite includes:

- Boston Naming
- Digits F
- Digits B
- Logical Memory Immediate
- Logical Memory Delayed
- MMSE

Score at EYO = 0 (Symptom Onset)

Cognitive Test (Mean & SE)	NACC (n=239)	DIAN MC (n=129)	p-value
Animal Fluency	14.60 (0.35)	17.22 (0.70)	**0.0009**
Boston Naming	24.41 (0.33)	26.14 (0.60)	**0.0119**
Digits Backward	5.67 (0.13)	5.91 (0.26)	0.4255
Digits Forward	7.70 (0.15)	7.72 (0.29)	0.9300
Logical Memory – Imm.	11.35 (0.30)	10.75 (0.58)	0.3689
Logical Memory – Del.	9.61 (0.33)	8.24 (0.63)	0.0539
MMSE	26.77 (0.27)	25.78 (0.52)	0.0928
Trails A	59.81 (2.17)	37.79 (4.02)	**<.0001**
Trails B	186.40 (5.21)	117.54 (10.80)	**<.0001**
WAIS-R Digit Symbol	30.10 (0.96)	44.45 (1.72)	**<.0001**
Composite (All above)	-0.24 (0.04)	0.13 (0.07)	**<.0001**
Non-speed Composite	-0.05 (0.04)	-0.09 (0.08)	0.6830

Slope After EYO = 0 (Progression)

Cognitive test (mean annual change & SE)	NACC (n=239)	DIAN MC (n=129)	P-value
Animal fluency	-1.56 (0.16)	-1.07 (0.24)	0.0884
Boston naming	-1.22 (0.15)	-0.77 (0.24)	0.1114
Digits backward	-0.33 (0.06)	-0.43 (0.08)	0.3232
Digits forward	-0.23 (0.07)	-0.45 (0.10)	0.0726
Logical memory – imm.	-1.40 (0.13)	-1.24 (0.19)	0.4936
Logical memory – del.	-1.46 (0.13)	-0.81 (0.20)	**0.0058**
Mmse	-1.84 (0.21)	-1.64 (0.30)	0.5868
Trails A	14.86 (1.87)	7.89 (2.53)	**0.0284**
Trails B	25.47 (3.42)	18.98 (4.77)	0.2717
WAIS-R digit symbol	-3.68 (0.56)	- 4.00 (0.80)	0.7452
Composite (all above)	-0.25 (0.02)	-0.16 (0.03)	**0.0279**
Non-speed composite	-0.26 (0.02)	-0.23 (0.04)	0.4827

Bear In Mind

LOAD:DIAN =Old:Young

=Many:Few comorbidities

=Without:With EYO

=…

Thanks

NACC investigators/staff;

DIAN investigators/staff;

Dr. Lena McCue from Biostatistics Core
Washington University Knight ADRC