Application of item response theory to the study of mixed neuropathologies in the aging brain

Yuriko Katsumata, David W. Fardo,
Peter T. Nelson, Erin L. Abner

Sanders-Brown Center on Aging
University of Kentucky

Acknowledgment

NACC

THE NIA ALZHEIMER'S DISEASE RESEARCH CENTERS PROGRAM
National Alzheimer's Coordinating Center

2022-2023 New Investigator Awards

Pathway-specific polygenic risk scores on mixed neuropathologies

Mentors

Dr. Dave Fardo

Dr. Erin Abner

Dr. Pete Nelson

PI & Director

Dr. Linda Van Eldik

Motivation of this study

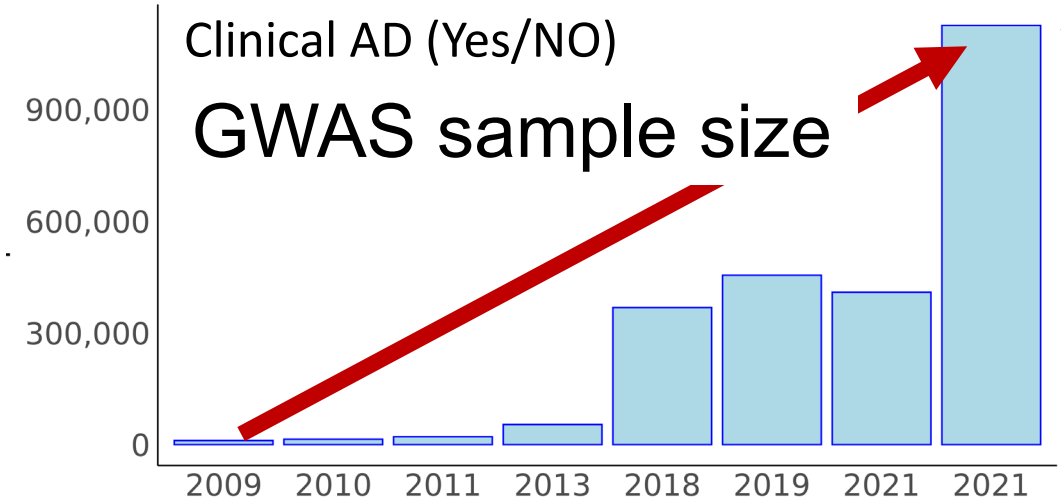
Genetics

Mixed pathology

We mainly have two issues

1. Sample size for genetic analysis on mixed pathology
2. Definition of mixed pathology

Sample size for genetic analysis on mixed pathology



n = **1,126,563**

ADRCs

ADSP WGS

ADGC Genotype

Linked

NACC NP
(March 2022 freeze data)
n = 6,980

ADGC = Alzheimer's Disease Genetics Consortium
ADSP = Alzheimer's Disease Sequencing Project

WGS + NP in Europeans n = **1,126**	ADGC + NP in Europeans n = **2,453**
WGS + NP in others n = **140**	ADGC + NP in others n = **134**

Definition of mixed pathology

Neurobiology of Disease 174 (2022) 105880

Contents lists available at ScienceDirect

Neurobiology of Disease

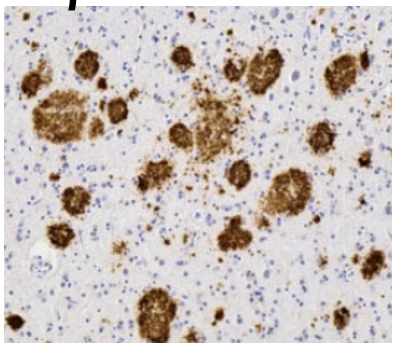
journal homepage: www.elsevier.com/locate/ynbdi

Multiple gene variants linked to Alzheimer's-type clinical dementia via GWAS are also associated with non-Alzheimer's neuropathologic entities

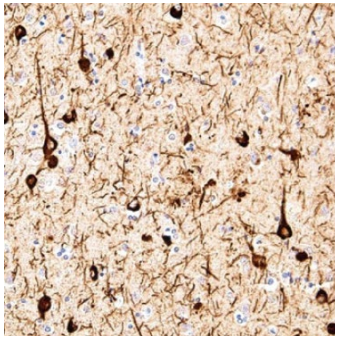
Yuriko Katsumata [a,b], Lincoln M. Shade [a], Timothy J. Hohman [c], Julie A. Schneider [d,e,f], David A. Bennett [d,e,f], Jose M. Farfel [d,e,f], Alzheimer's Disease Genetics Consortium, Walter A. Kukull [g], David W. Fardo [a,b], Peter T. Nelson [b,h,*]

Genetics

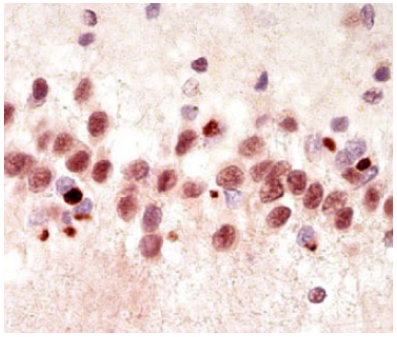
Aβ



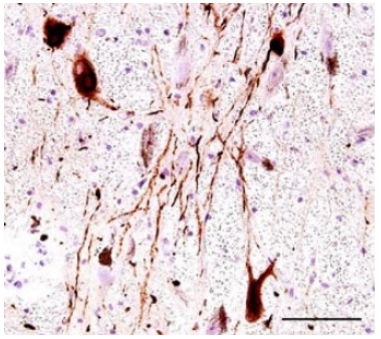
Tau



TDP-43



α-Synuclein



NACC Neuropathology (NP) Data Set

Neuropathology	NACC variable	0	1	2	3
A score	NPTHAL	Phase 0	Phase 1/2	Phase 3	Phase 4/5
B score	NACCBRAA	Stage 0	Stage 1/2	Stage 3/4	Stage 5/6
C score	NACCNEUR	No	Sparse	Moderate	Frequent
TDP-43 in amygdala	NPTDPB	No	Yes		
TDP-43 in limbic	NPTDPC/D	No	Yes		
TDP-43 in Neocortex	NPTDPE	No	Yes		
Lewy bodies	NACCLEWY NPLBOD	No	Others	Neocortex	
HS	NPHIPSCL	No	Yes		
CAA	NACCAMY	None	Mild	Moderate	Severe
Arteriolosclerosis	NACCARTE	None	Mild	Moderate	Severe
Atherosclerosis	NACCAVAS	None	Mild	Moderate	Severe
Infarct and lacunes	NACCINF	No	Yes		
Microinfarcts	NACCMICR	No	Yes		

13 neuropathologies

NACC Neuropathology (NP) Data Set

Neuropathology	NACC variable	0	1	2	3
A score	NPTHAL	Phase 0	Phase 1/2	Phase 3	Phase 4/5
B score	NACCBRAA	Stage 0	Stage 1/2	Stage 3/4	Stage 5/6
C score	NACCNEUR	No	Sparse	Moderate	Frequent
TDP-43 in amygdala	NPTDPB	No	Yes		
TDP-43 in limbic	NPTDPC/D				
TDP-43 in Neocc					
Lewy bodies				rtex	
HS					
CAA	NACCAMY	None	Mild	Moderate	Severe
Arteriolosclerosis	NACCARTE	None	Mild	Moderate	Severe
Atherosclerosis	NACCAVAS	None	Mild	Moderate	Severe
Infarct and lacunes	NACCINF	No	Yes		
Microinfarcts	NACCMICR	No	Yes		

# of combinations of pathologies

786,432 ways

13 neuropathologies

Motivation of this study

Genetics

Mixed pathology

Summarizing variables

- Principal component analysis
- Correspondence analysis
- Linear discriminant Analysis
- T-distributed Stochastic Neighbor Embedding (t-SNE)
- Uniform manifold approximation and projection (UMAP)

Dimensionality reduction

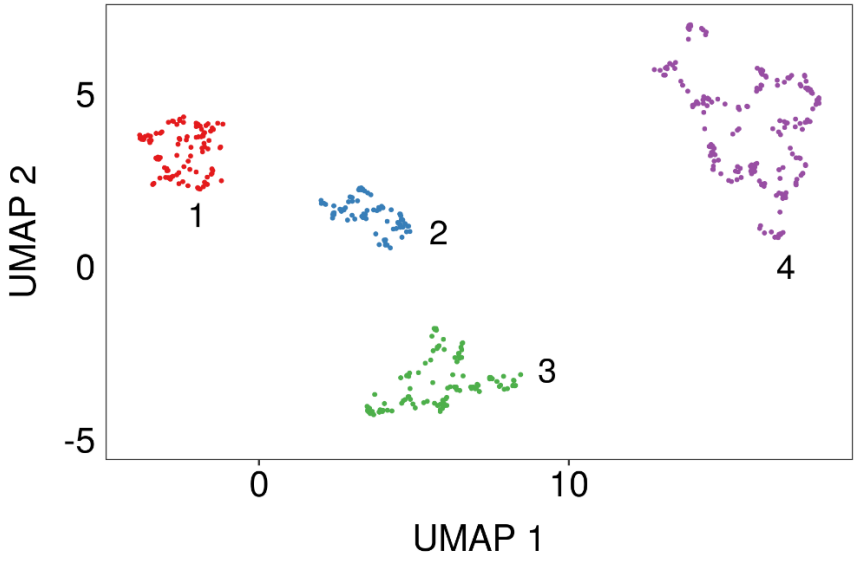
Acta Neuropathologica (2020) 140:659–674
https://doi.org/10.1007/s00401-020-02211-0

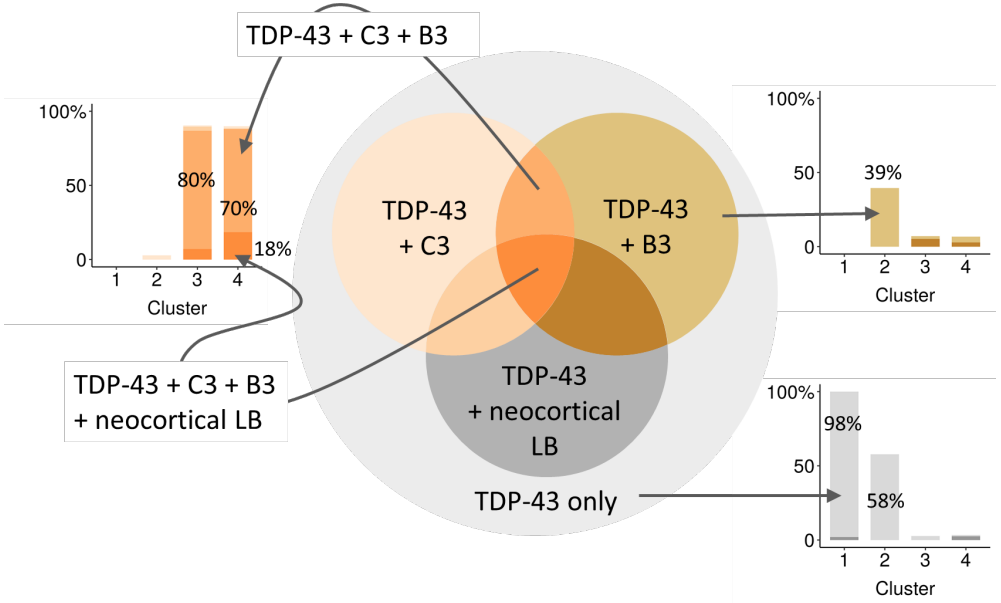
ORIGINAL PAPER

Distinct clinicopathologic clusters of persons with TDP-43 proteinopathy

Yuriko Katsumata[1,2] · Erin L. Abner[2,3] · Shama Karanth[3] · Merilee A. Teylan[4] · Charles N. Mock[4,5] · Matthew D. Cykowski[6] · Edward B. Lee[7] · Kevin L. Boehme[8] · Shubhabrata Mukherjee[9] · John S. K. Kauwe[10,11] · Richard J. Kryscio[2,12] · Frederick A. Schmitt[2,13] · David W. Fardo[1,2] · Peter T. Nelson[2,14]

Received: 2 June 2020 / Revised: 7 August 2020 / Accepted: 8 August 2020 / Published online: 14 August 2020





Motivation of this study

Genetics

Spec Aim 2

Spec Aim 1

Mixed pathology

Summarizing variables

Dimensionality reduction

Item Response Theory

- Math ability = not directly observed

➡ **Latent variable**

- Test score = directly observed

Difficulty ↓

Problem	Correct = 1 incorrect = 0	
	Student 1	Student 2
Item 1	1	1
Item 2	0	0
Item 3	1	0
Item 4	1	1
Item 5	0	**1**
Total	3	3

- Mixed pathology condition

  = not directly observed

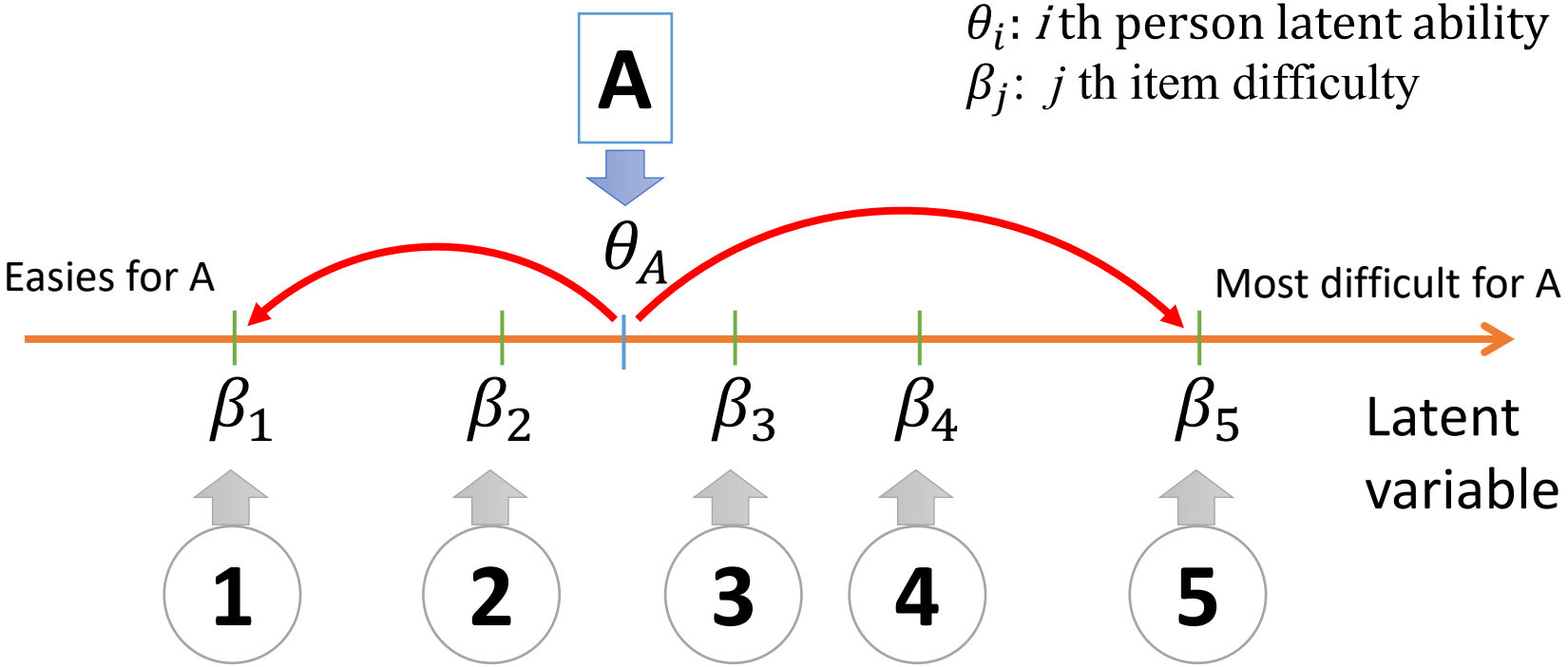
  ➡ **Latent variable**

- Individual pathology = directly observed at autopsy

Severity ↓

Pathology	Presence = 1 absence = 0	
	Case 1	Case 2
Item 1	1	1
Item 2	0	0
Item 3	1	0
Item 4	1	1
Item 5	0	**1**
Total	3	3

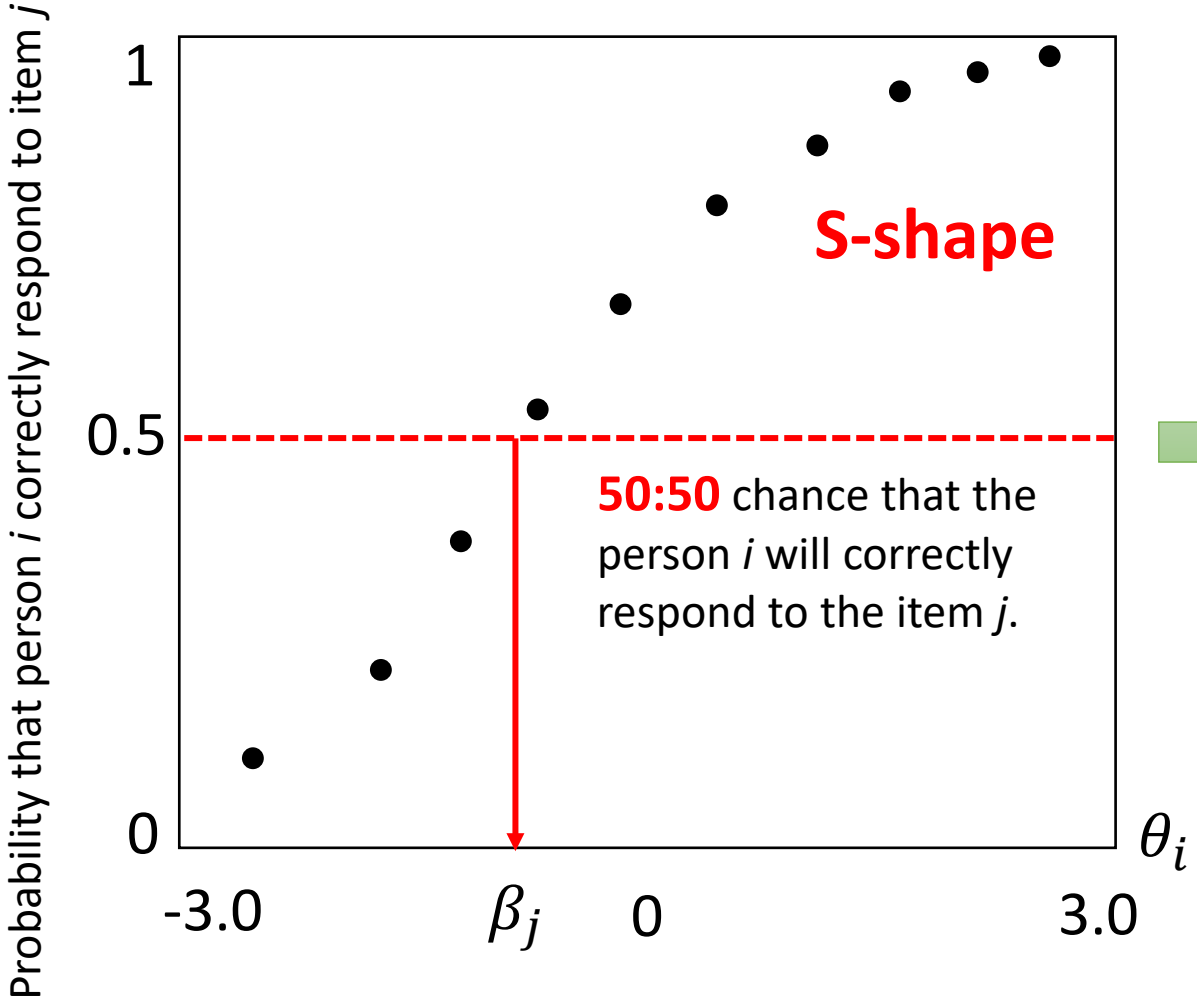
Both items and persons are located on the same latent variable continuum.

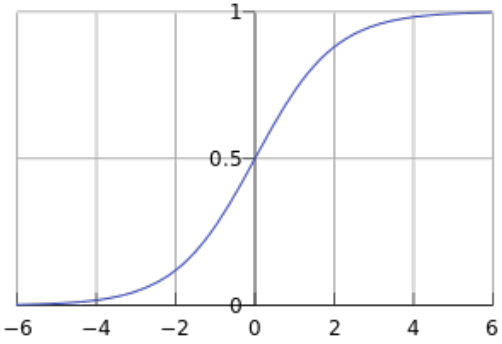


The **distance** between the person and item locations ➡ $\theta_i - \beta_j$

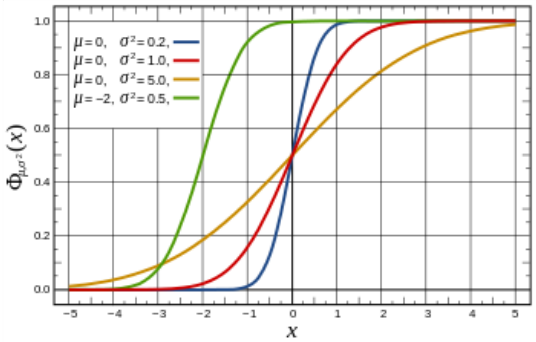
The probability that the person *i* will correctly respond to the item *j* ➡ $f(\theta_i - \beta_j)$

What does $f(\theta_i - \beta_j)$ look like?





Logistic function



CDF of normal distribution

Logistic function $P\left(y_{ij}=1\middle|\theta_i,\beta_j\right)=\dfrac{1}{1+e^{-(\theta_i-\beta_j)}}$

Rasch model

In general, the Rasch model can be written as

$$L(\boldsymbol{\theta}, \boldsymbol{\beta}|\boldsymbol{y}) = \prod_{i=1}^{N}\prod_{j=1}^{J}\frac{e^{y_{ij}(\theta_i-\beta_j)}}{1+e^{(\theta_i-\beta_j)}} = \frac{\exp\left\{\sum_{i=1}^{N}\sum_{j=1}^{J} y_{ij}\left(\theta_i-\beta_j\right)\right\}}{\prod_{i=1}^{N}\prod_{j=1}^{J}\left\{1+e^{(\theta_i-\beta_j)}\right\}}$$

Model can be extended

Binary 0, 1 ➡ Ordered 0, 1, 2, 3,…. $y_{ijk} = \begin{cases} 1 \text{ if } u_{ij} = k \\ 0 \text{ otherwise} \end{cases}$

Unidimensional ➡ Multidimensional

One parameter ➡ Two or more parameters $\frac{\gamma_j}{1+e^{-\alpha_j(\theta_i-\beta_j)}}$

Multidimensional Partial Credit Model

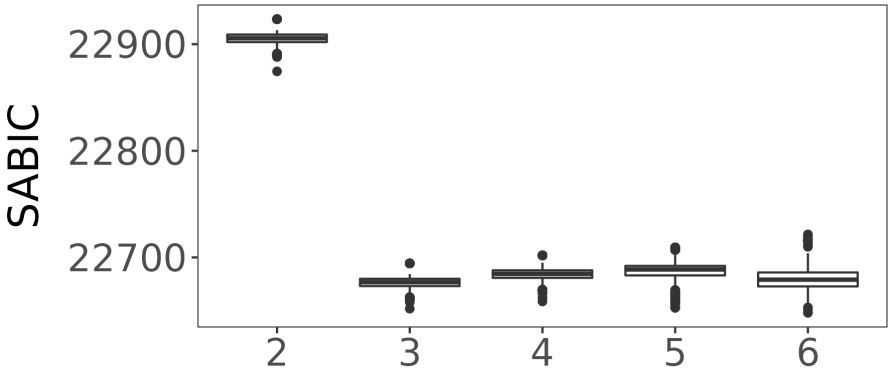
implemented by the “mirt” R package

https://cran.r-project.org/web/packages/mirt/mirt.pdf

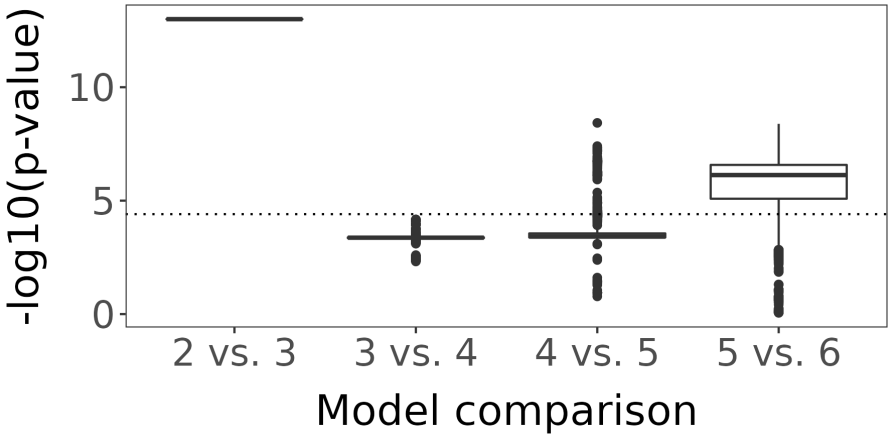
NACC Neuropathology (NP) Data Set --- 3/31/2022 freeze data

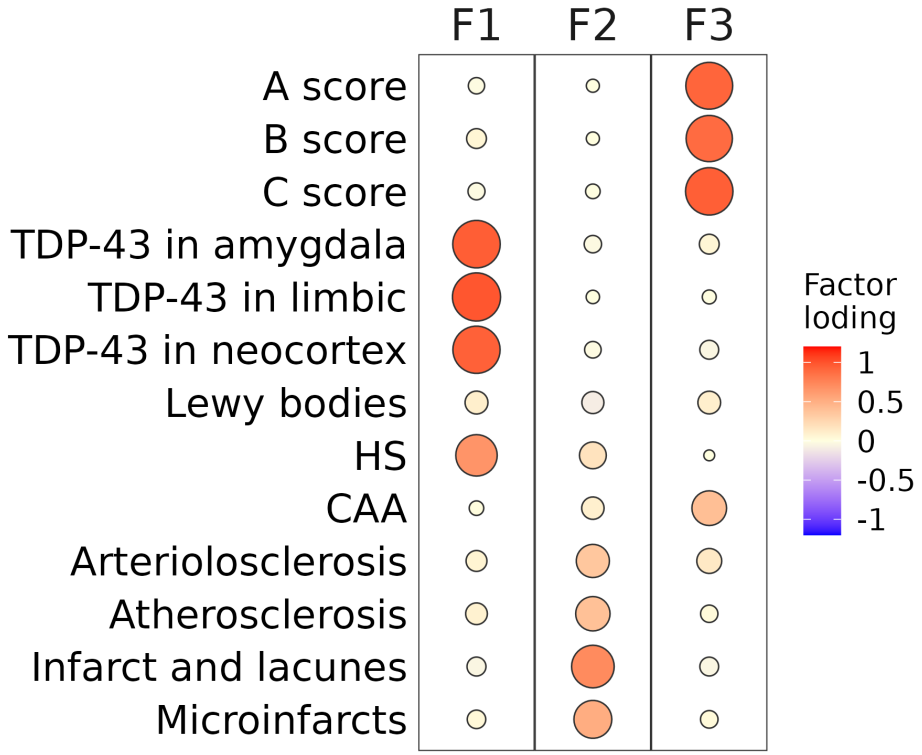
Neuropathology	NACC variable	0	1	2	3
A score	NPTHAL	Phase 0	Phase 1/2	Phase 3	Phase 4/5
B score	NACCBRAA	Stage 0	Stage 1/2	Stage 3/4	Stage 5/6
C score	NACCNEUR	No	Sparse	Moderate	Frequent
TDP-43 in amygdala	NPTDPB	No	Yes		
TDP-43 in limbic	NPTDPC/D	No	Yes		
TDP-43 in Neocortex	NPTDPE	No	Yes		
Lewy bodies	NACCLEWY NPLBOD	No	Others	Neocortex	
HS	NPHIPSCL	No	Yes		
CAA	NACCAMY	None	Mild	Moderate	Severe
Arteriolosclerosis	NACCARTE	None	Mild	Moderate	Severe
Atherosclerosis	NACCAVAS	None	Mild	Moderate	Severe
Infarct and lacunes	NACCINF	No	Yes		
Microinfarcts	NACCMICR	No	Yes		

n = 1,222 who died after 60 years or older (no missing data)



SABIC = sample size adjusted Bayesian information criterion





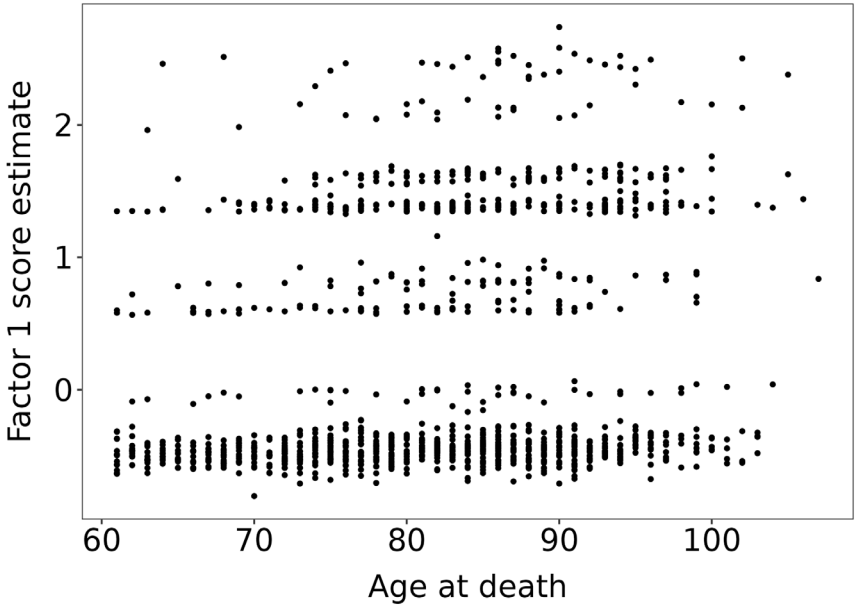
Rotated SS loadings

F1 = 3.220
F2 = 1.107
F3 = 2.721

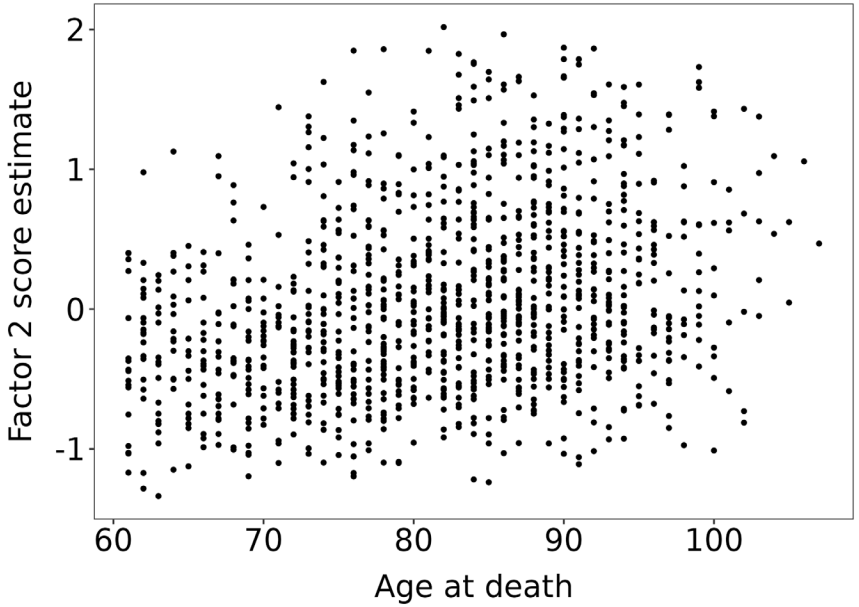
Factor correlations

	F1	F2
F2	0.068	
F3	0.281	-0.167

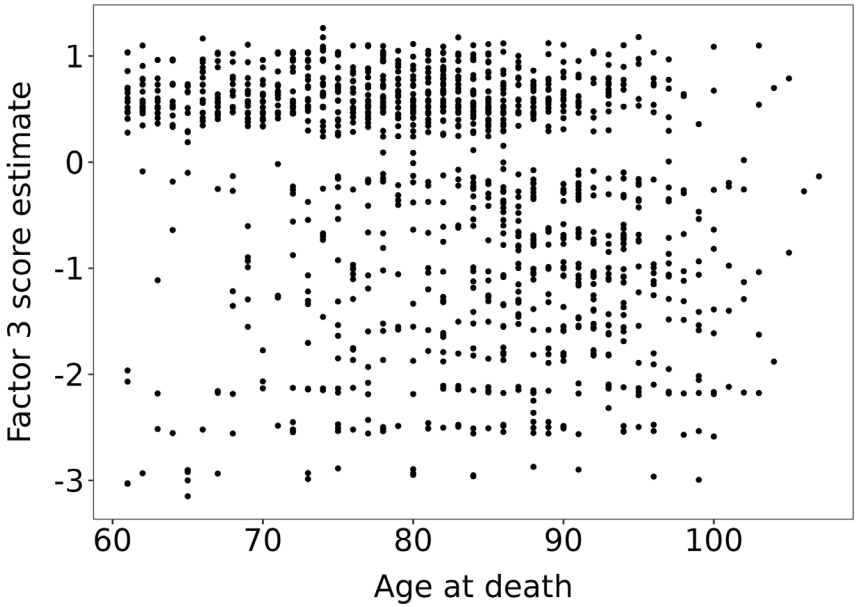
TDP-43/LATE



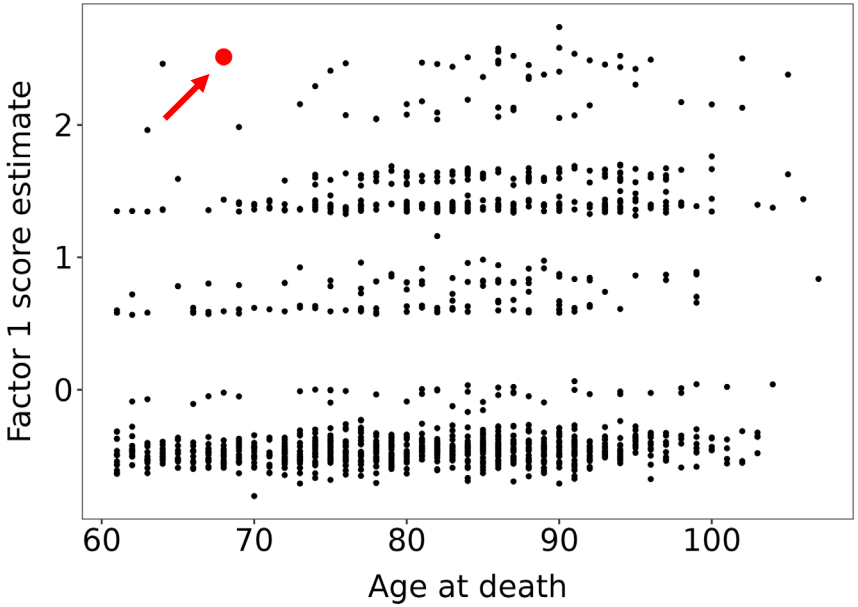
CVD



AD

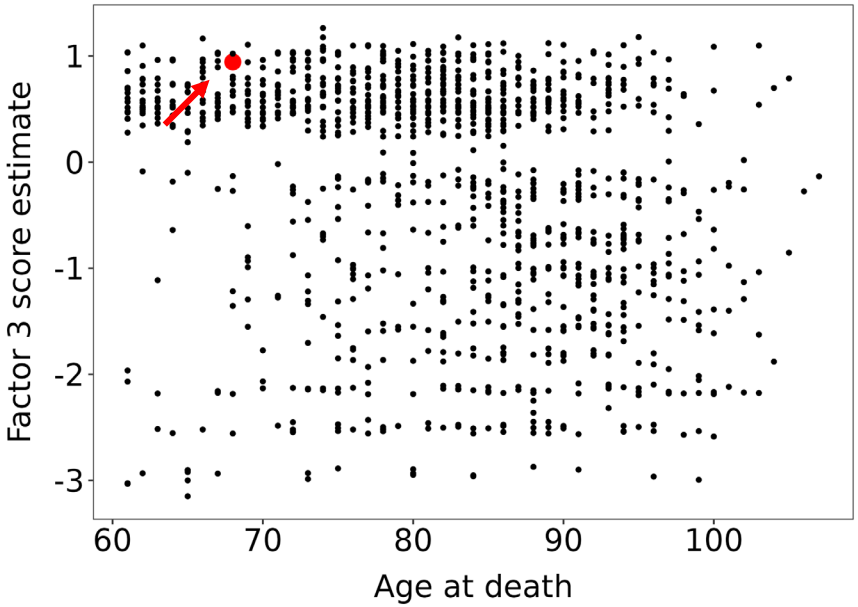


TDP-43/LATE



CVD

AD

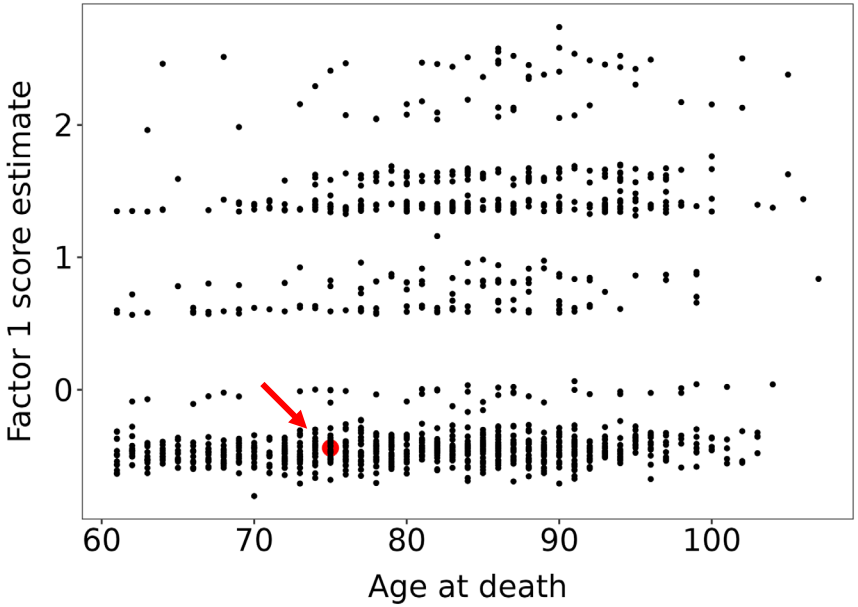


Mixed pathology pattern (NACC031468)

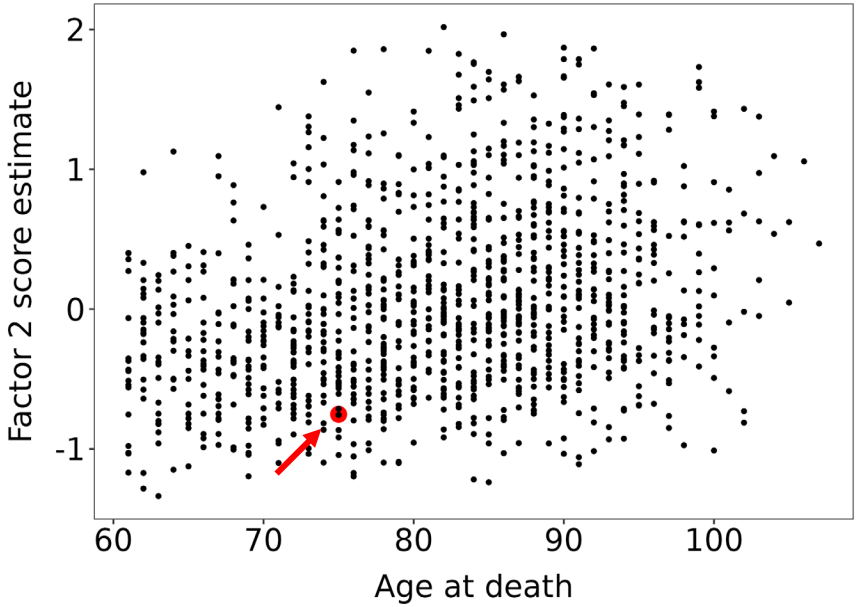
A score	3
B score	3
C score	3
TDP-43 in amygdala	1
TDP-43 in limbic	1
TDP-43 in Neocortex	1
Lewy bodies	2
HS	1

CAA	3
Arteriolosclerosis	1
Atherosclerosis	1
Infarct and lacunes	0
Microinfarcts	0

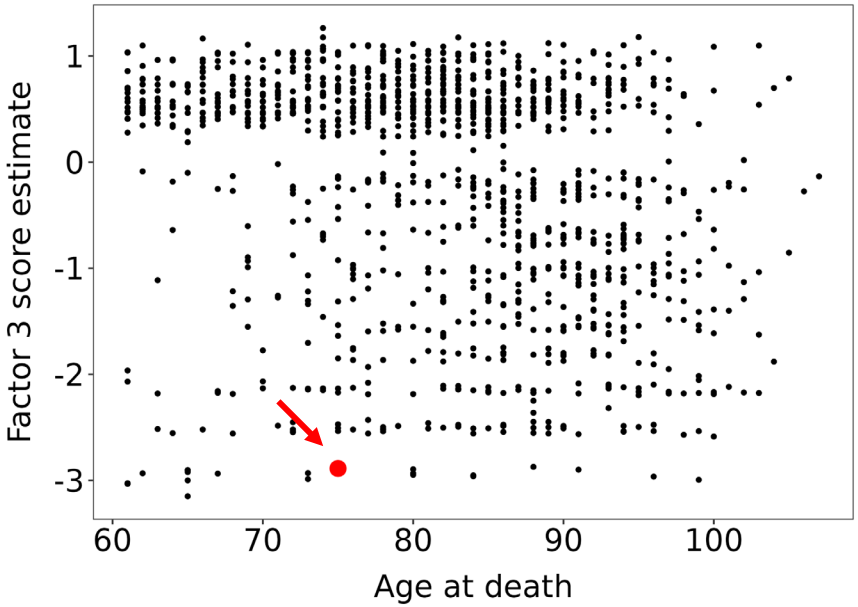
TDP-43/LATE



CVD



AD

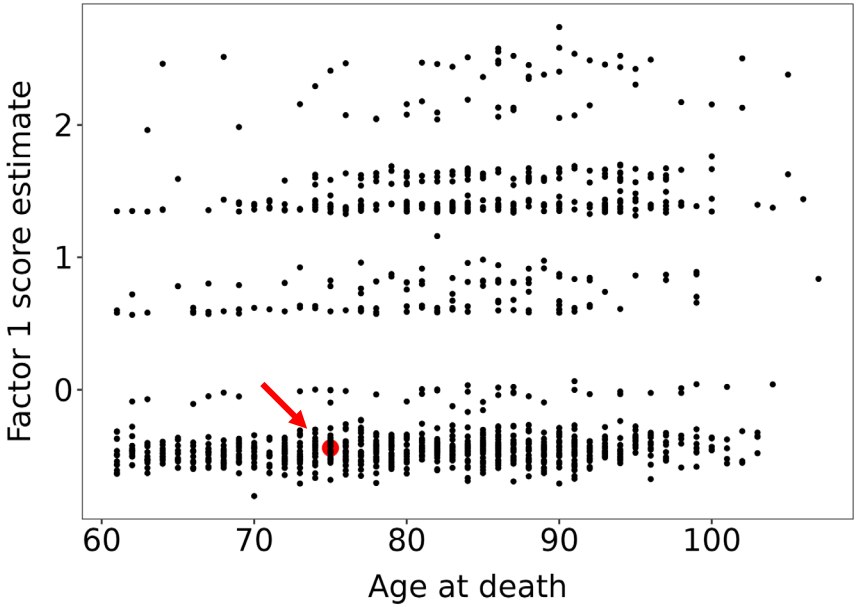


Mixed pathology pattern (NACC017944)

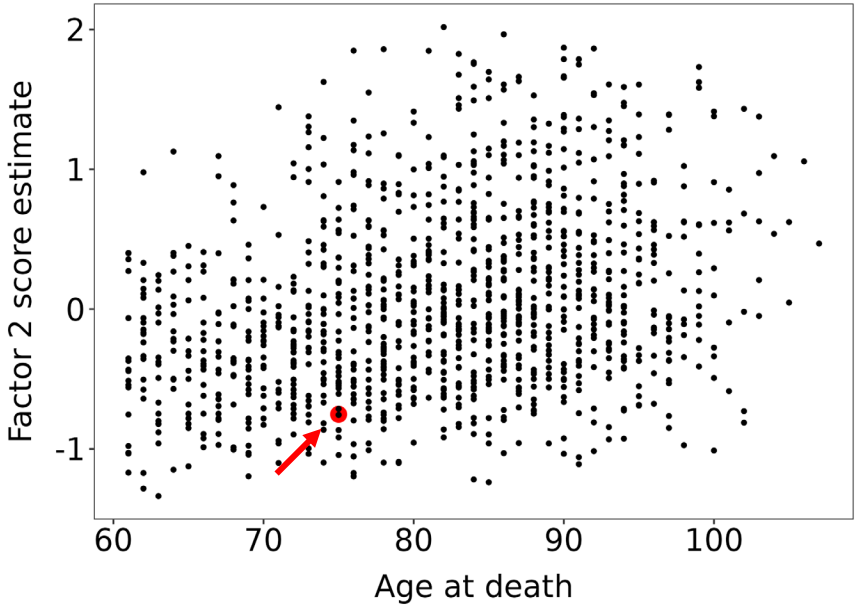
A score	0
B score	0
C score	0
TDP-43 in amygdala	0
TDP-43 in limbic	0
TDP-43 in Neocortex	0
Lewy bodies	2
HS	0

CAA	1
Arteriolosclerosis	0
Atherosclerosis	1
Infarct and lacunes	0
Microinfarcts	0

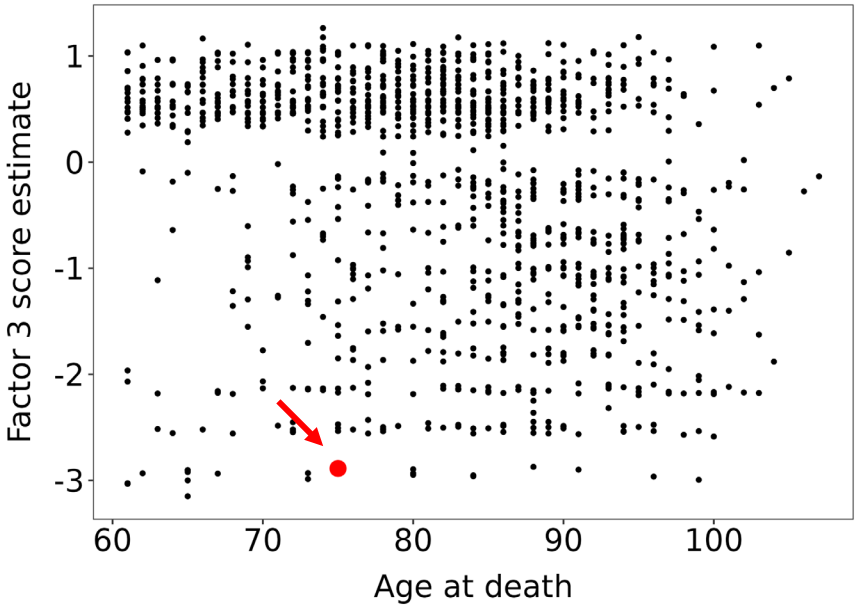
TDP-43/LATE



CVD



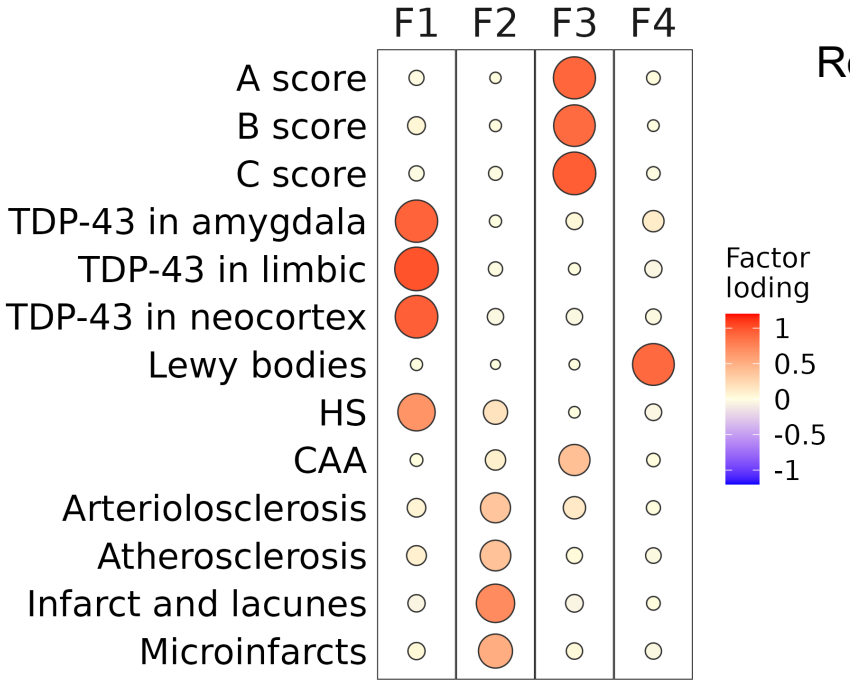
AD



Mixed pathology pattern (NACC017944)

A score	0
B score	0
C score	0
TDP-43 in amygdala	0
TDP-43 in limbic	0
TDP-43 in Neocortex	0
Lewy bodies	**2**
HS	0

CAA	1
Arteriolosclerosis	0
Atherosclerosis	1
Infarct and lacunes	0
Microinfarcts	0



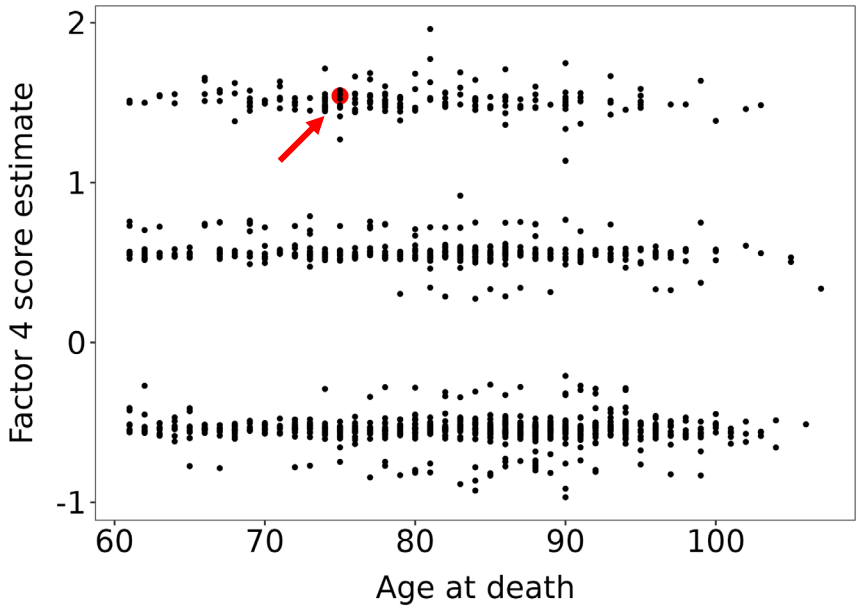
Rotated SS loadings

F1 = 3.216
F2 = 1.101
F3 = 2.708
F4 = 0.826

Factor correlations

	F1	F2	F3
F2	0.068		
F3	0.278	-0.166	
F4	0.172	-0.111	0.236

LB



Mixed pathology pattern (NACC017944)

A score	0
B score	0
C score	0
TDP-43 in amygdala	0
TDP-43 in limbic	0
TDP-43 in Neocortex	0
Lewy bodies	**2**
HS	0

CAA	1
Arteriolosclerosis	0
Atherosclerosis	1
Infarct and lacunes	0
Microinfarcts	0

Summary

- We detected four dimensions with TDP-43/LATE, CVD, AD, and LB predominant pathologies
- Age at death is an important factor to be taken into account
- We estimated individual **quantitative** risk scores with four dimensions
  - More powerful than binary and categorical outcome variables in genetic analyses
- We will examine the associations between the four dimensional risk scores and genetics in future