Alzheimer's Disease Survival Analysis

By Jace Higa, Paxton Jones, Landon Waters

Background & Data

Background Information

- Alzheimer's is a neurodegenerative condition that affects memory, thinking, and behavior
- It is becoming more and more common and is the leading cause of dementia
  - Dementia is a fever, while Alzheimer's is the flu
- Understanding factors that contribute most to the disease is important for both early intervention and long-term care planning

https://www.alz.org/

Data and Methodology

- Data was collected by the CDC through an observational study
- Contains 2149 patients, aged 60-90 years old with 35 covariates
  - Demographic info, lifestyle factors, medical histories, cognitive tests, functional assessments, symptoms
  - Diagnosis
- Allowed us to explore many factors that could be associated with the onset of Alzheimer's disease
- Right censoring, this study ended but patients will still develop Alzheimer's

---

https://www.kaggle.com/datasets/rabieelkharoua/alzheimers-disease-dataset?resource=download

Questions of Interest

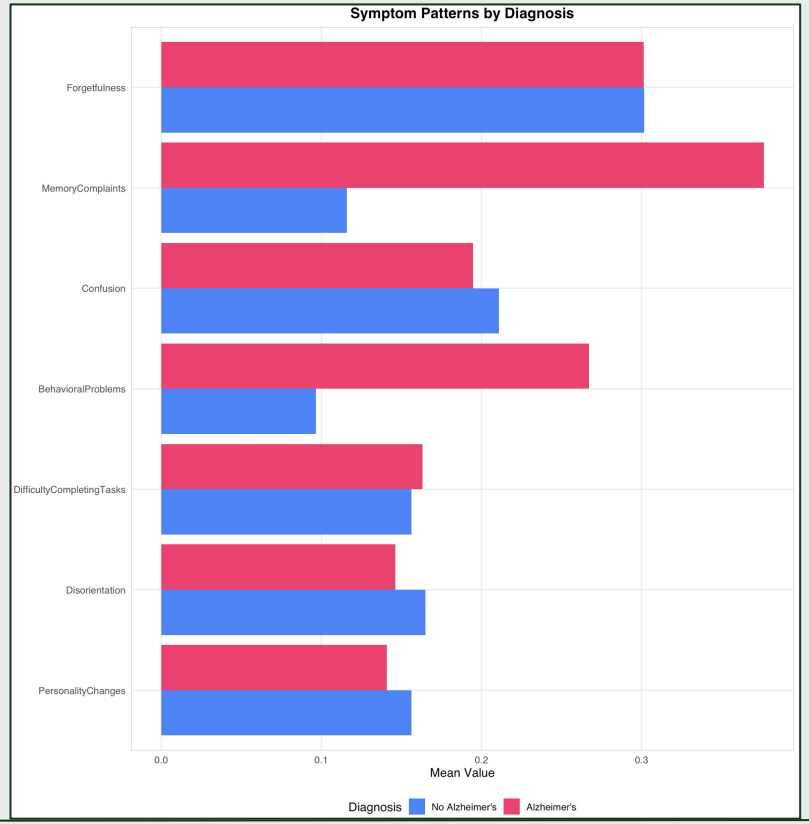
- What covariates are most closely tied to Alzheimer's at time(t)?
- Using those dominant covariates which models are most effective in modeling and fitting the data?

Important Covariates

Covariate	Description	Range / Interpretation
Functional Assessment	Functional assessment score	0–10 (lower = more impaired)
ADL	Activities of Daily Living score	0–10 (lower = more impaired)
Memory Complaints	Presence of memory complaints	0 = No, 1 = Yes
Behavioral Problems	Presence of behavioral problems	0 = No, 1 = Yes
MMSE	Mini-Mental State Exam score	0–30 (lower = more impaired)
CholesterolLDL	LDL cholesterol level	50–200 mg/dL

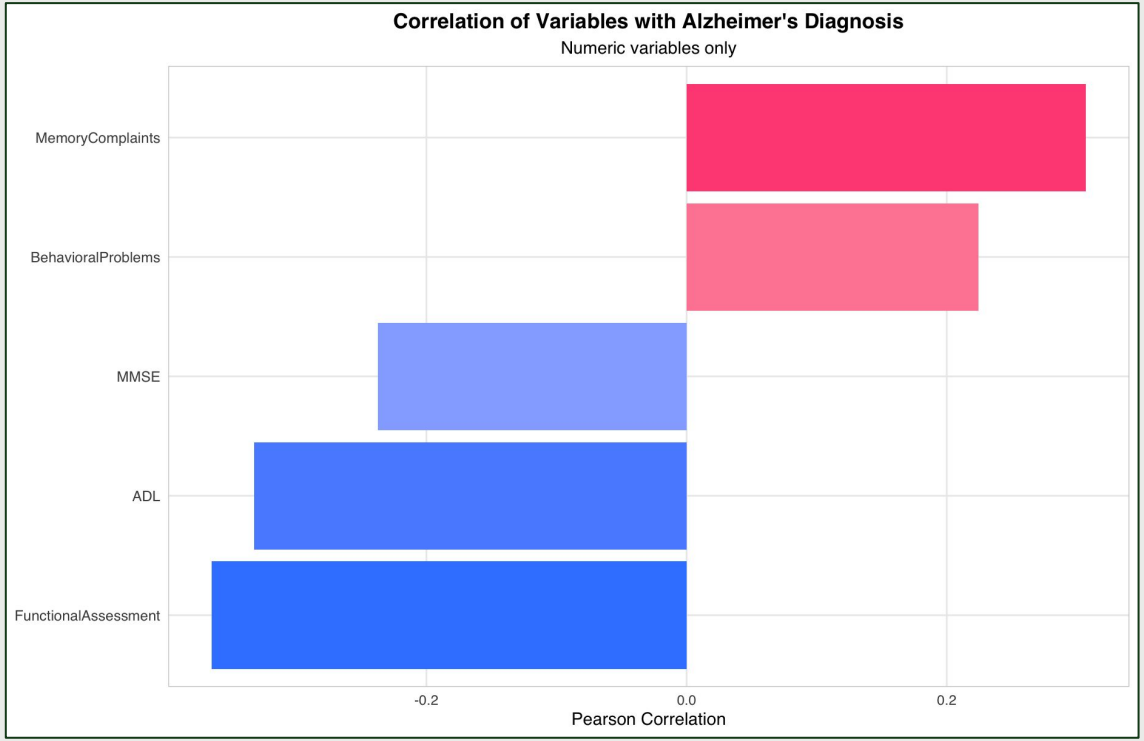
Results & Analysis

Patterns in Symptoms by Diagnosis

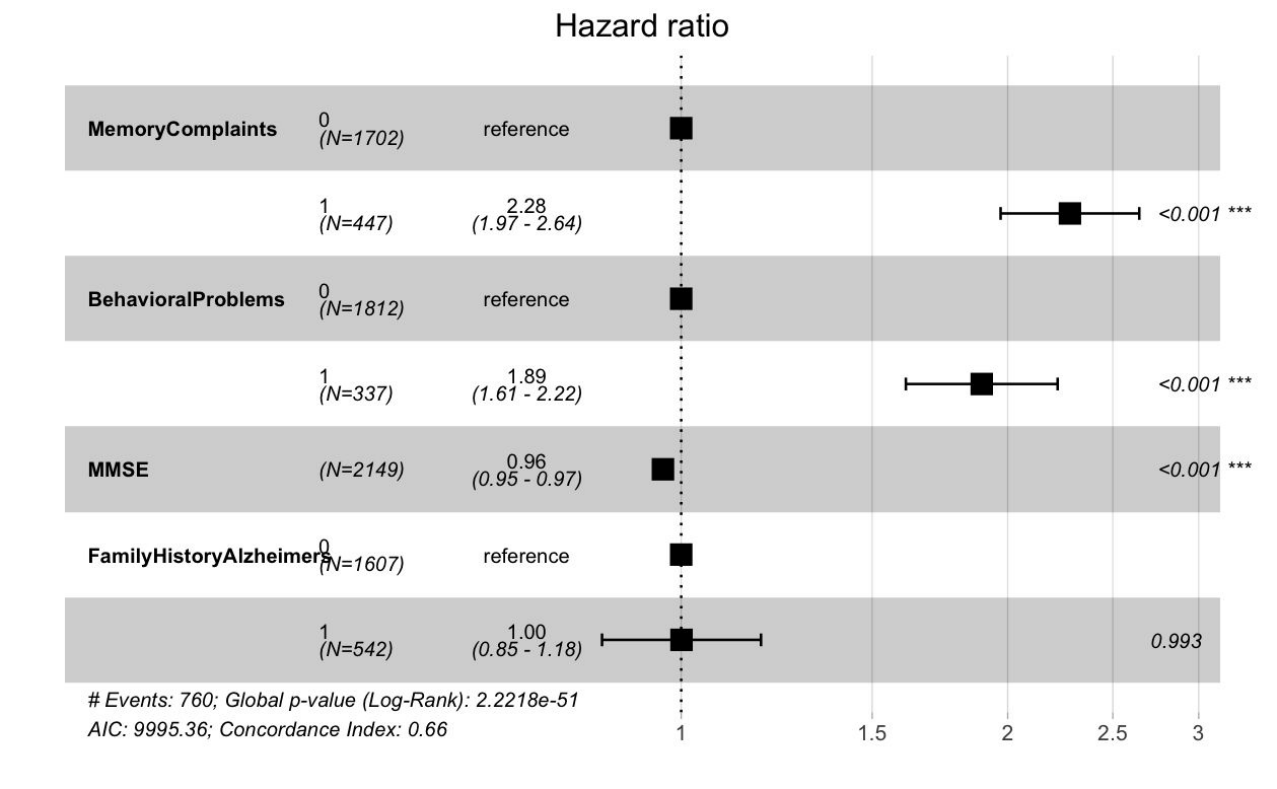


- Symptoms are binary variables
- Mean value in this case measures proportion of individuals with the symptom
- Greatest differences seen are eas

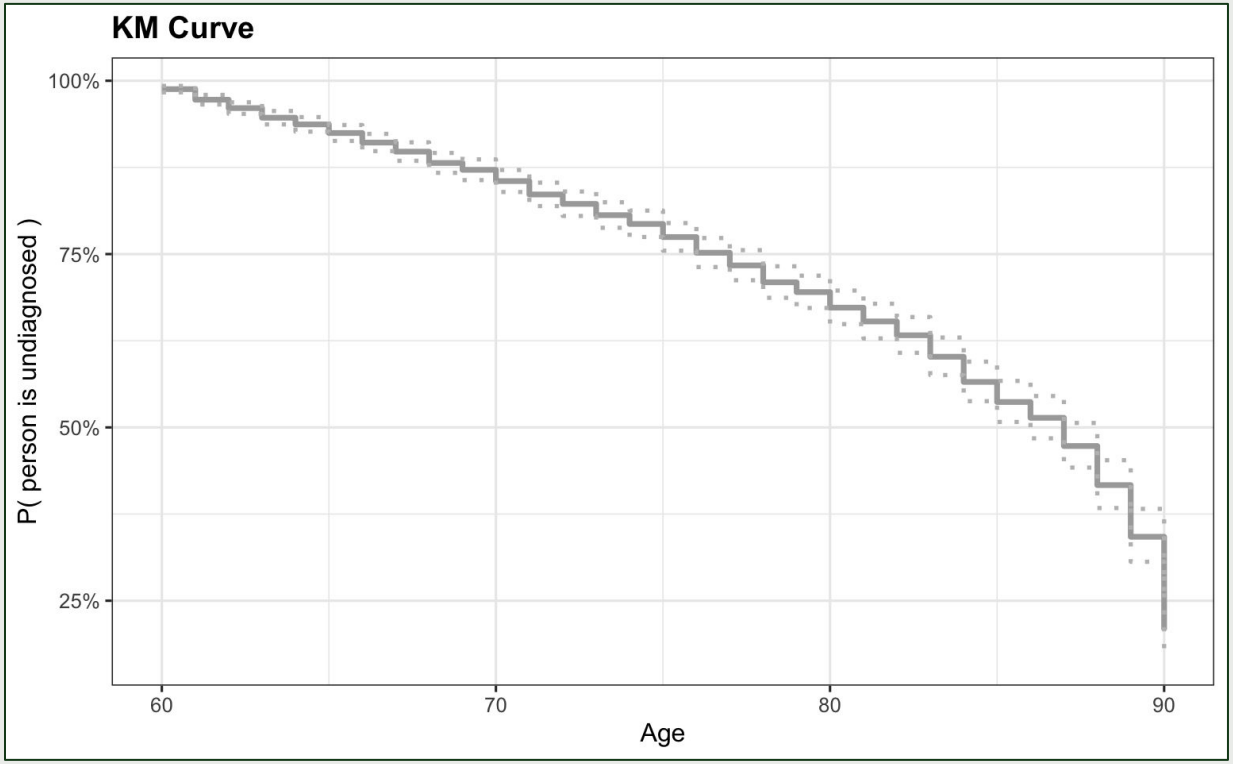
Correlation Plot



Forest Plot


Hazard ratio
MemoryComplaints
0 (N=1702)
reference
1 (N=447)
2.28 (1.97 - 2.64)
<0.001 ***
BehavioralProblems
0 (N=1812)
reference
1 (N=337)
1.89 (1.61 - 2.22)
<0.001 ***
MMSE
(N=2149)
0.96 (0.95 - 0.97)
<0.001 ***
FamilyHistoryAlzheimers
0 (N=1607)
reference
1 (N=542)
1.00 (0.85 - 1.18)
0.993
Events: 760; Global p-value (Log-Rank): 2.2218e-51
AIC: 9995.36; Concordance Index: 0.66
1
1.5
2
2.5
3

KM Curve



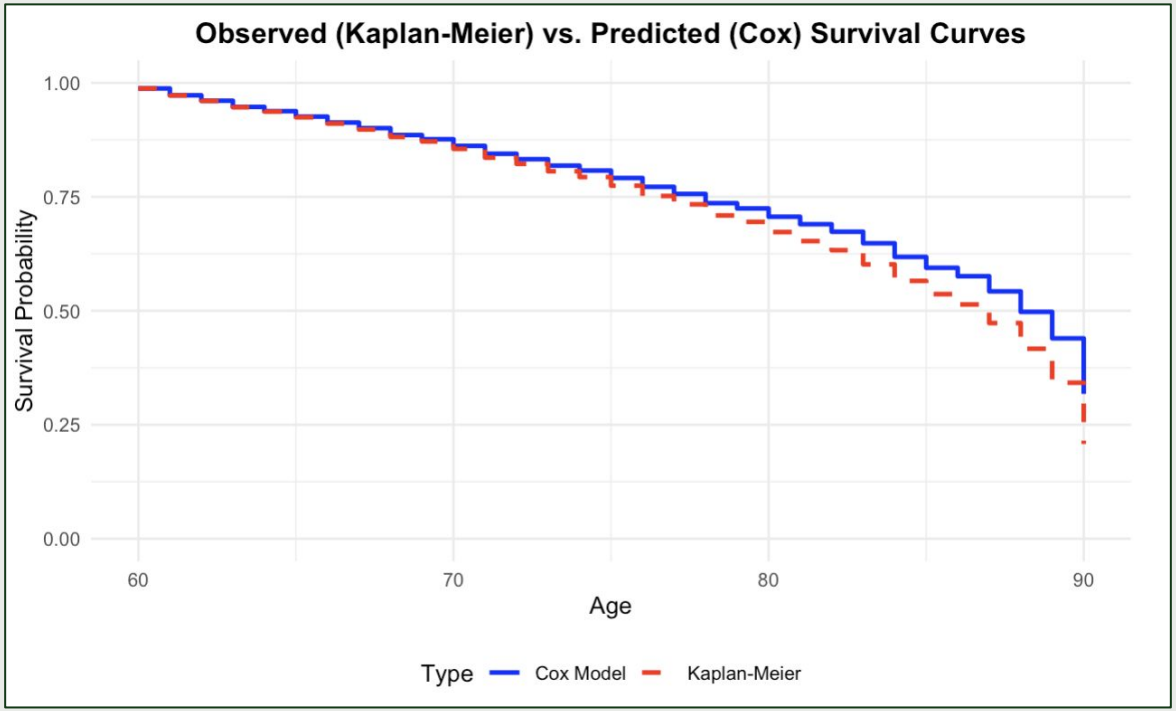
Minimum Survival Time:
- 60 years

Maximum Survival Time:
- 90 years

Median Survival Time:
- 75 years

Mean Survival Time:
- 75 years

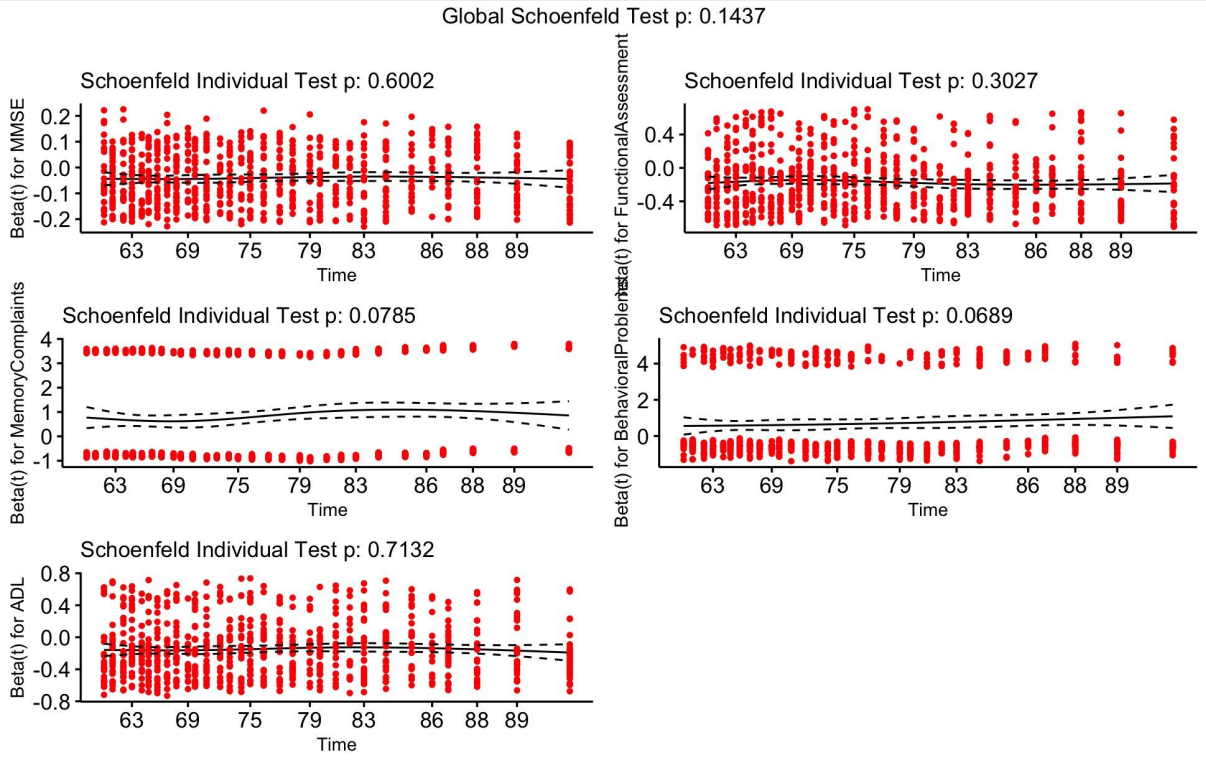
Cox Model



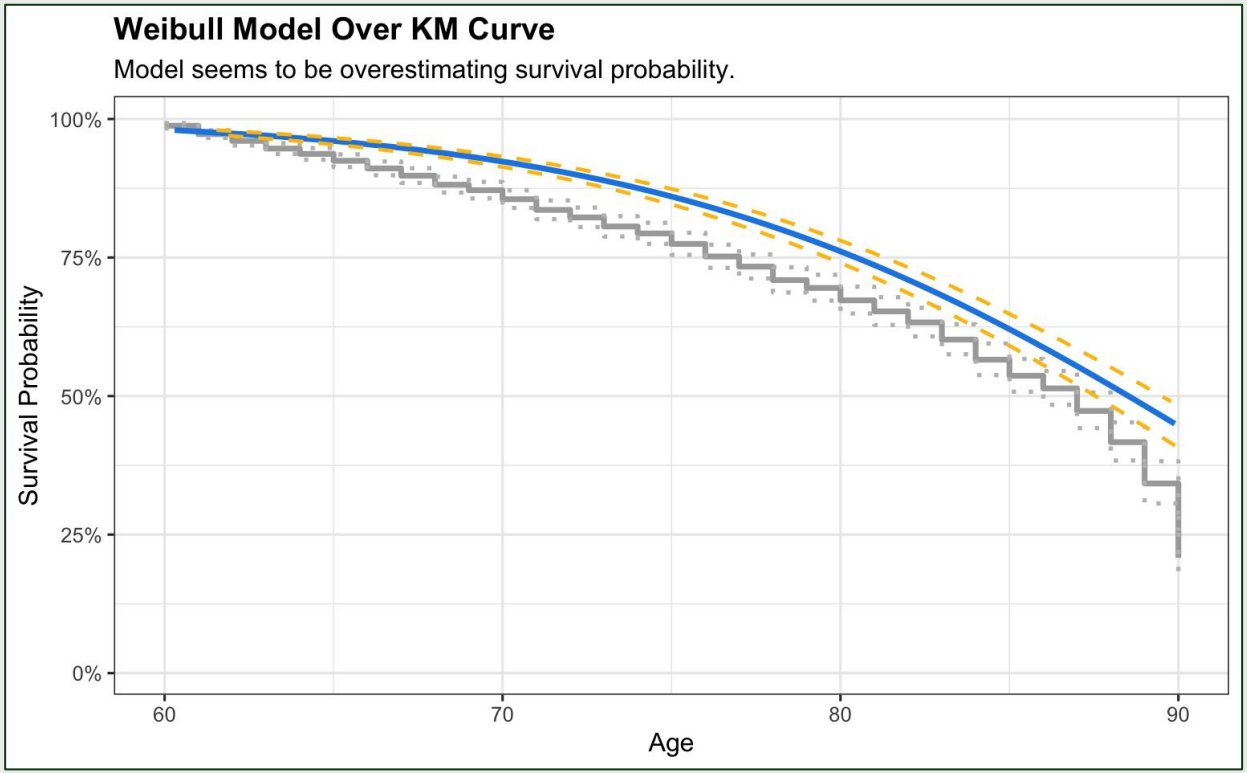
Concordance of 0.718

Schoenfeld Test

- None of our variables violate the assumption



Weibull Model

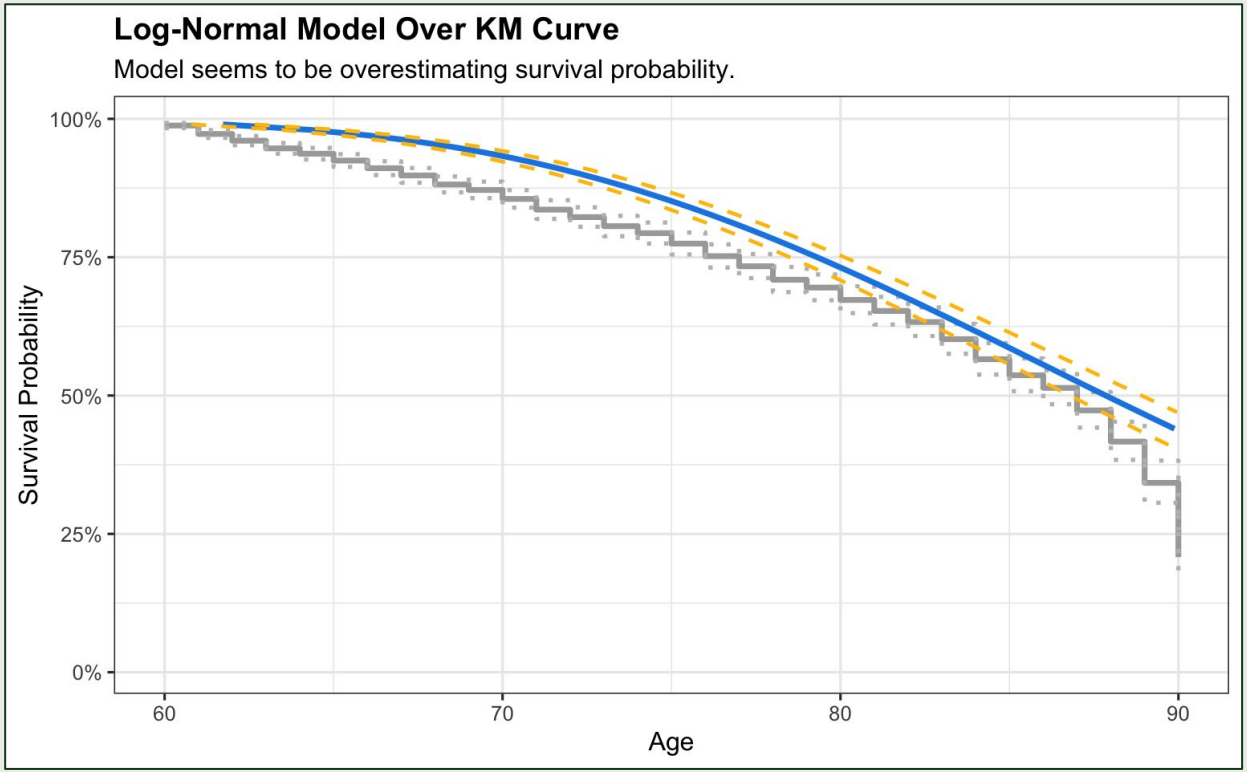


β (shape): 9.206289
α (scale): 74.95807

Interpretation: Most people avoid diagnosis until around age 70, but then the chances of being diagnosed increase rapidly, around age 75.

survreg(Surv(Age,Diagnosis) ~ MMSE + FunctionalAssessment + MemoryComplaints + BehavioralProblems + ADL, data = ds)

Log-Normal Model



σ (sigma): 0.1516862
μ (mu): 4.247985

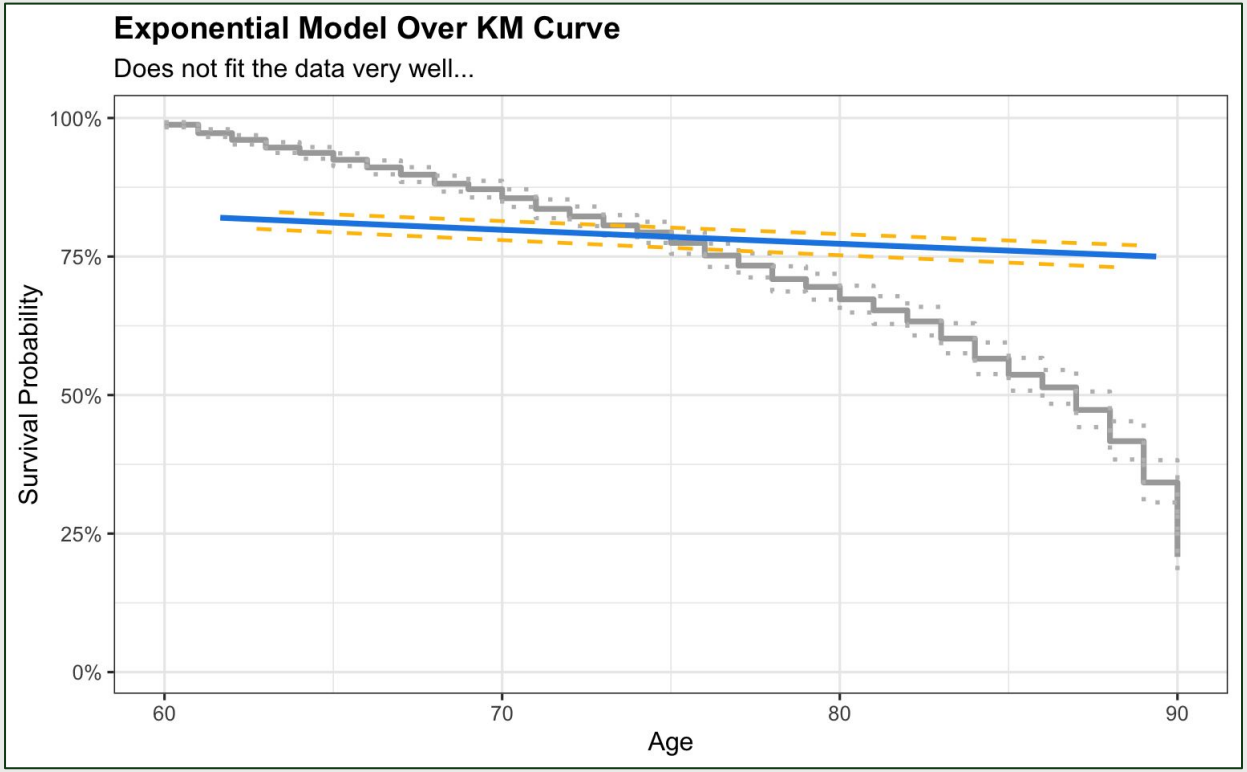
Interpretation: With a narrow spread in survival times, most diagnoses happen around age ~70 ($e^{4.248}$)

survreg(Surv(Age,Diagnosis) ~ MMSE + FunctionalAssessment + MemoryComplaints + BehavioralProblems + ADL, data = ds, dist = "lognormal")

Exponential Model

λ (rate) = 0.021

Interpretation: Risk of AD doesn't change with much with age...



survreg(Surv(Age,Diagnosis) ~ MMSE + FunctionalAssessment + MemoryComplaints + BehavioralProblems + ADL, data = ds, dist = "exp")

AIC Score Comparison

- Weibull Model = 6582.198
- Log-Normal Model = 6586.154
- Exponential Model = 9097.189

- The Weibull and Log-Normal Models seems to be comparable

Comparing Between Groups

Alzheimer's Rates by Gender

- “**Two-thirds** of clinically diagnosed cases of **dementia** and **AD** are **women**”
  - Source: “Differences Between Women and Men in Incidence Rates of Dementia and Alzheimer's Disease”
- Does this show up in our data?

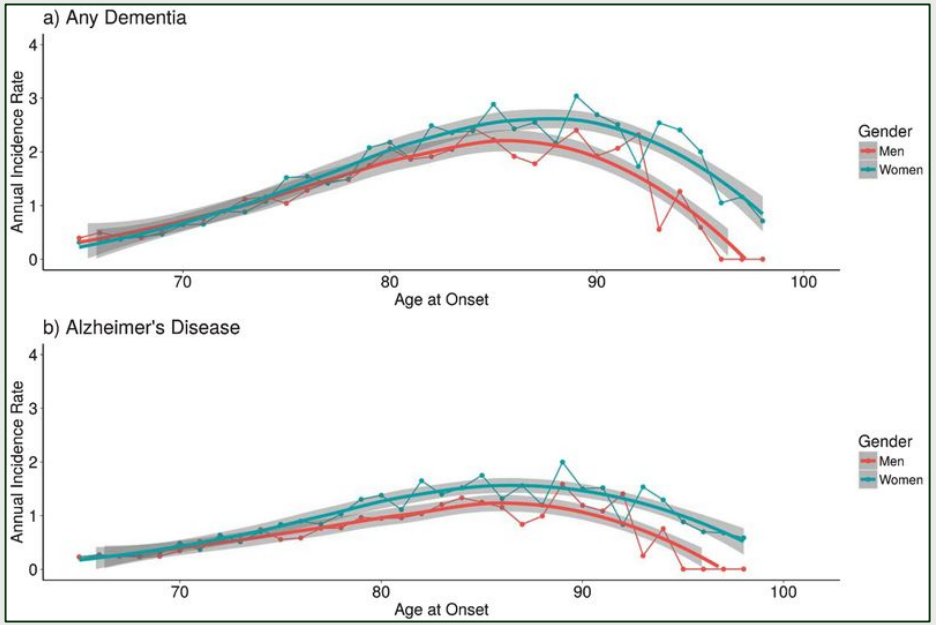
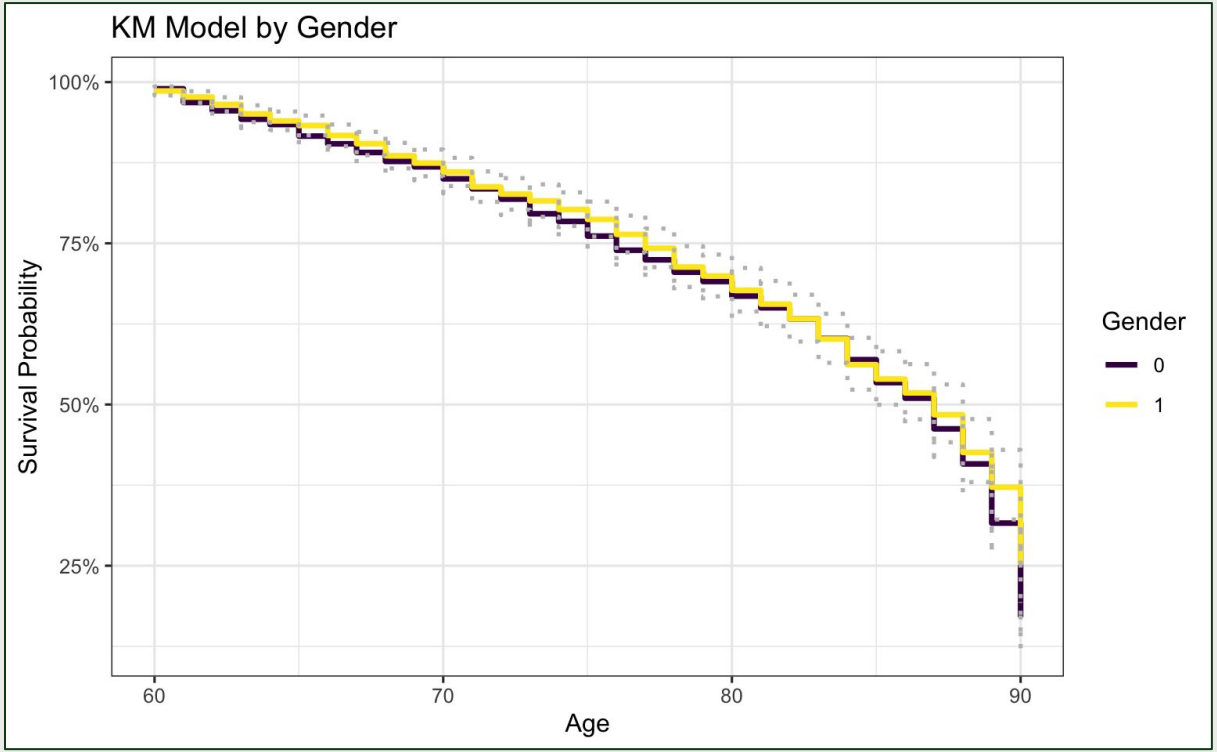


Image via same source

Log Rank Test on Gender



H_0: No statistically significant difference in survival between the genders

- p-value = 0.3

We **fail to reject** the null hypothesis.

NOTE: 0 is Male, 1 is Female

Conclusions

Key Takeaways/Future Directions

- Functional Assessment, Activities of Daily Living, and Mini-Mental State Exam scoring alongside high cholesterol, memory complaints, and behavioral problems are predictive of AD diagnosis
- Parametric models consistently underestimate age of diagnosis
- Data did not show a significant difference between gender groups for AD diagnosis
    - Worth exploring further!
- Evaluating this data in a different context could have been really interesting

Questions?