Is Dementia predictable?

F. Di Filippo, E. Manfrin, E. Musiari, E. Palli

17 december 2021

Dataset

Dataset Dementia and Alzheimer longitudinal

Subject.ID	MRI.ID	Group	Visit	MR.Delay	M.F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
OAS2_0001	OAS2_0001_MR1	Nondemented	1	0	M	R	87	14	2	27	0.0	1987	0.696	0.883
OAS2_0001	OAS2_0001_MR2	Nondemented	2	457	M	R	88	14	2	30	0.0	2004	0.681	0.876
OAS2_0002	OAS2_0002_MR1	Demented	1	0	M	R	75	12	*NA*	23	0.5	1678	0.736	1.046
OAS2_0002	OAS2_0002_MR2	Demented	2	560	M	R	76	12	*NA*	28	0.5	1738	0.713	1.010
OAS2_0002	OAS2_0002_MR3	Demented	3	1895	M	R	80	12	*NA*	22	0.5	1698	0.701	1.034
OAS2_0004	OAS2_0004_MR1	Nondemented	1	0	F	R	88	18	3	28	0.0	1215	0.710	1.444
OAS2_0004	OAS2_0004_MR2	Nondemented	2	538	F	R	90	18	3	27	0.0	1200	0.718	1.462
OAS2_0005	OAS2_0005_MR1	Nondemented	1	0	M	R	80	12	4	28	0.0	1689	0.712	1.039
OAS2_0005	OAS2_0005_MR2	Nondemented	2	1010	M	R	83	12	4	29	0.5	1701	0.711	1.032

where SES is Socioeconomic Status, MMSE is Mini Mental State Examination, CDR is Clinical Dementia Rating, eTIV is Estimated Total Intracranial Volume, nWBV is Normalize Whole Brain Volume and ASF is Atlas Scaling Factor.

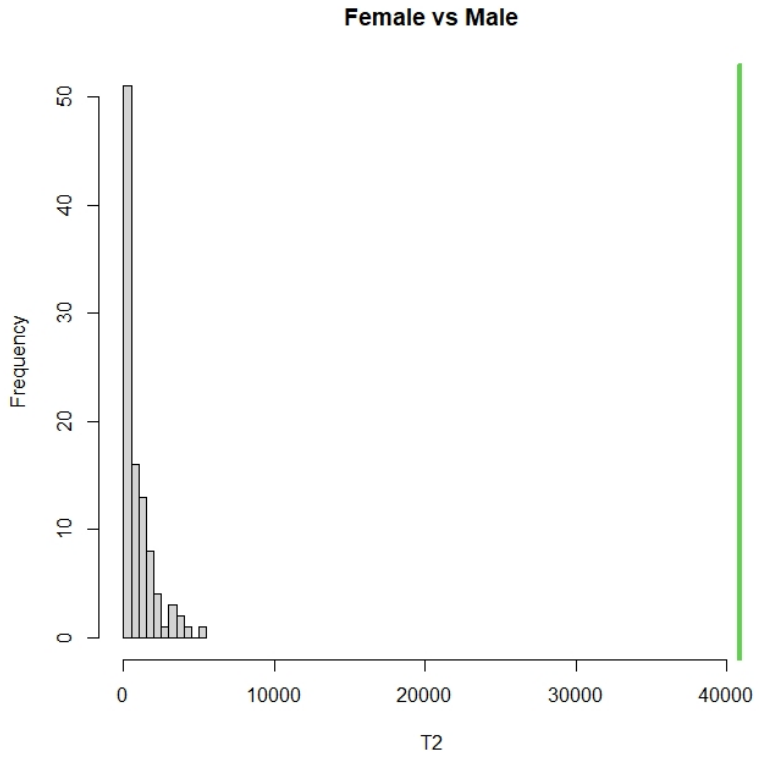
Source: Kaggle

We tried to solve the problem of correlation using the PCA method.

Analysis Male VS Female

Variable considered: $[Age, EDUC, MMSE, eTIV, nWBV, ASF]$

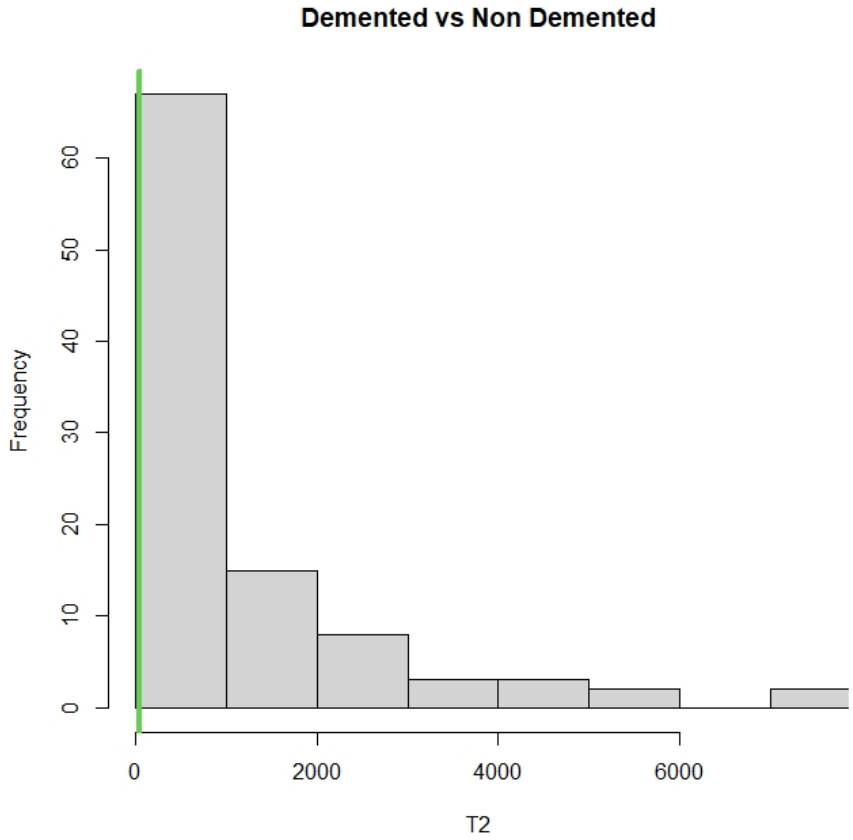
$$H_0 : \mathbf{Y}_{female} \overset{d}{=} \mathbf{Y}_{male} \ vs \ H_1 : \mathbf{Y}_{female} \overset{d}{\neq} \mathbf{Y}_{male}$$



$pvalue = 0$

Demented VS Nondemented

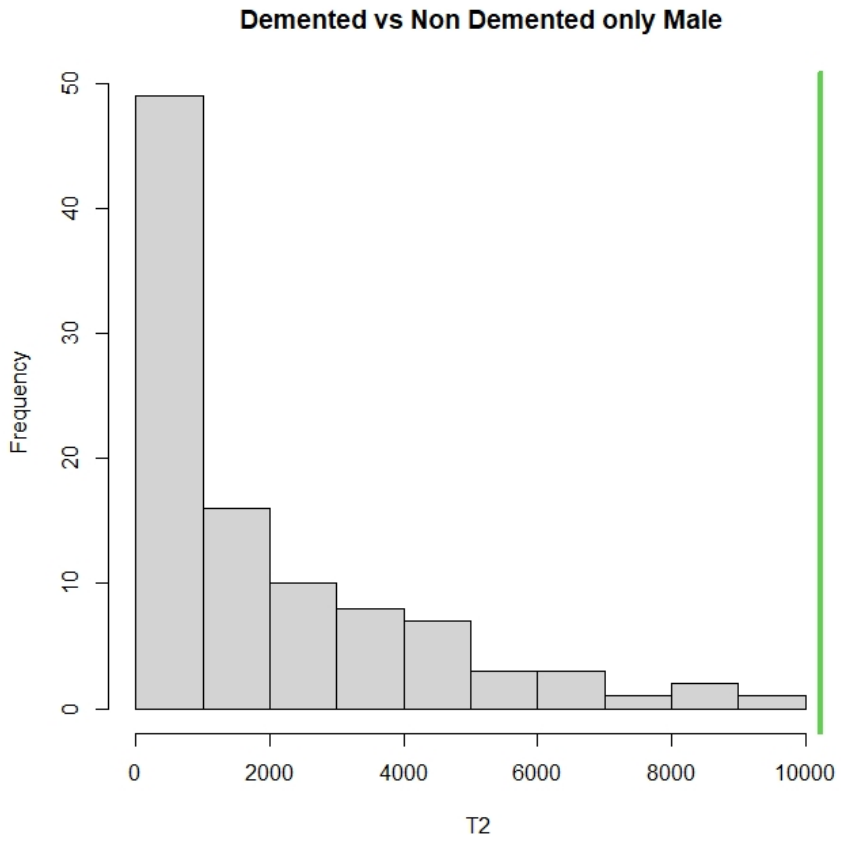
$$H_0 : \mathbf{Y}_{Demented} \overset{d}{=} \mathbf{Y}_{NonDemented} \ vs \ H_1 : \mathbf{Y}_{Demented} \overset{d}{\neq} \mathbf{Y}_{NonDemented}$$



$pvalue = 0.9$

Demented VS Nondemented only Male

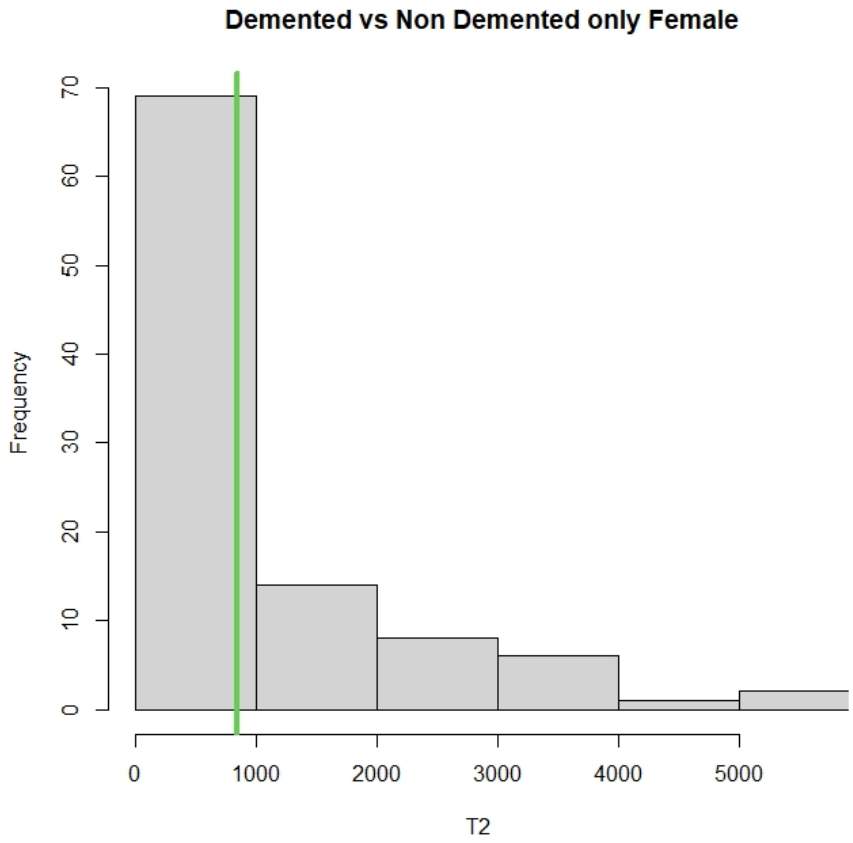
$$H_0 : \mathbf{M}_{Demented} \overset{d}{=} \mathbf{M}_{NonDemented} \ vs \ H_1 : \mathbf{M}_{Demented} \overset{d}{\neq} \mathbf{M}_{NonDemented}$$



$pvalue = 0$

Demented VS Nondemented only Female

$$H_0 : \mathbf{F}_{Demented} \overset{d}{=} \mathbf{F}_{NonDemented} \; vs \; H_1 : \mathbf{F}_{Demented} \overset{d}{\neq} \mathbf{F}_{NonDemented}$$



$pvalue = 0.37$

$$EDUC = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon$$

$i = male, female \;\; j = Demented, NonDemented$
$\alpha = sex, \;\; \beta = diagnostic, \;\; \gamma = interaction$

$$H_0 : \gamma_{ij} = 0 \;\; vs \; H_1 : \gamma_{ij} \neq 0$$

TEST STATISTIC: $T0 = F - STATISTICS \quad p - value = 0.082$ at level of confidence 95% there's no evidence to reject H_0 so we reduce the model

$$EDUC = \mu + \alpha_i + \beta_j + \epsilon$$

$$H_0 : \beta_j = 0 \;\; vs \; H_1 : \beta_i \neq 0$$

$p - value = 0.069$ at level of confidence 95% there's no evidence to reject H_0 there's no evidence to reject H_0

$$EDUC = \mu + \alpha_i$$

$$H_0 : \alpha_i = 0 \ \ vs \ H_1 : \alpha_i \neq 0$$

$p - value = 0.08$ we could say that's neither of the grouping is significant at 95%
while with parametric test at least the diagnostic division is significant

$$MMSE = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon$$

$i = male, female$ $j = Demented, NonDemented$

$$H_0 : \gamma_{ij} = 0 \quad vs \; H_1 : \gamma_{ij} \neq 0$$

TEST STATISTIC: $T0 = F - STATISTICS \quad p - value = 0.875$ there's no evidence to reject H_0 so we reduce the model

$$EDUC = \mu + \alpha_i + \beta_j + \epsilon$$

$$H_0 : \beta_j = 0 \quad vs \; H_1 : \beta_j \neq 0$$

$p - value = 0.446$ a there's no evidence to reject H_0 there's no evidence to reject H_0

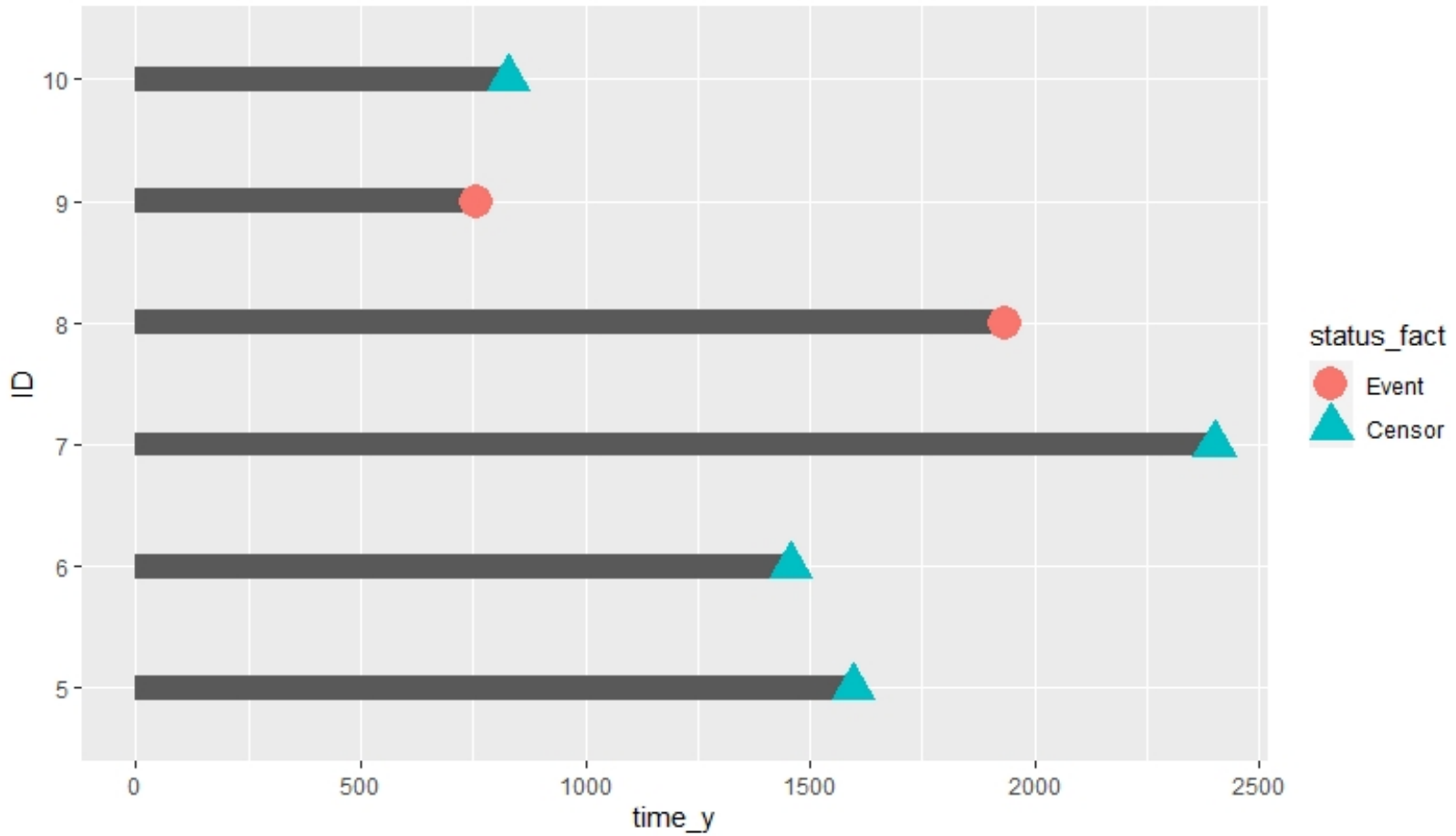
$$EDUC = \mu + \alpha_i$$

$$H_0 : \alpha_i = 0 \ \ vs \ H_1 : \alpha_i \neq 0$$

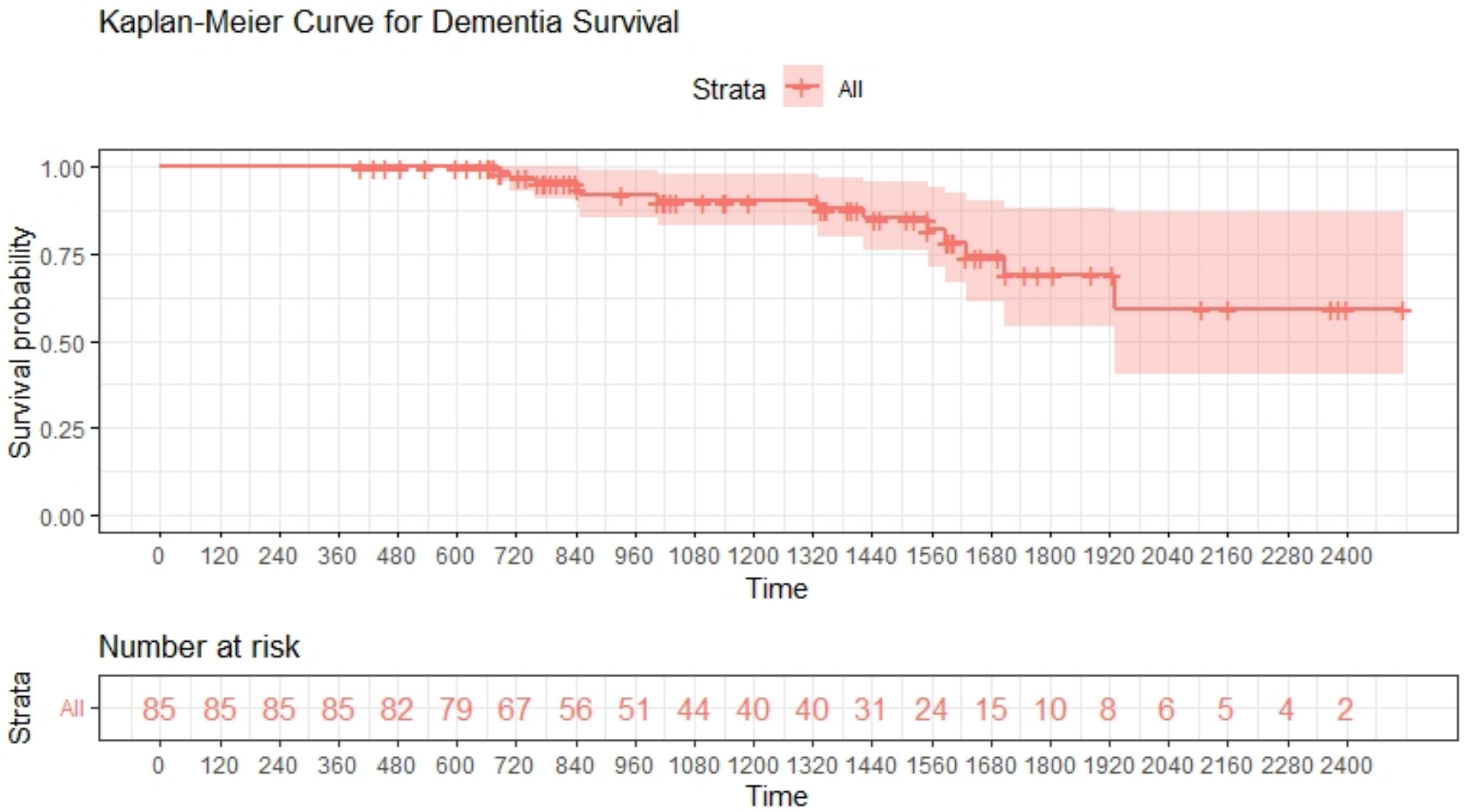
$p - value = 0$ there's evidence to reject H_0 so the most significative model is $MMSE \sim Diagnostic$ where Diagnostic is the division between Demented and Non Demented

Prediction

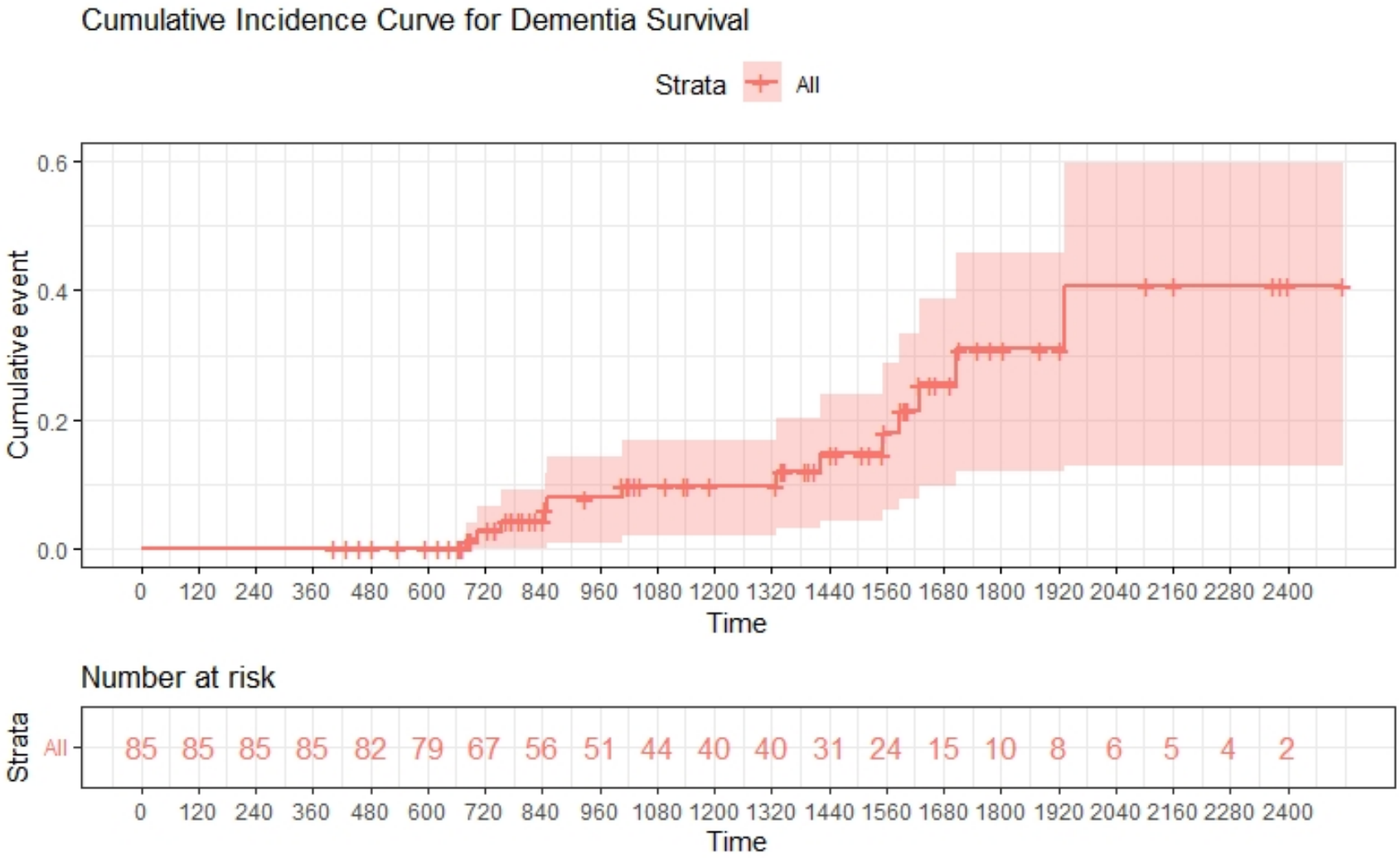
Event: disease occurred



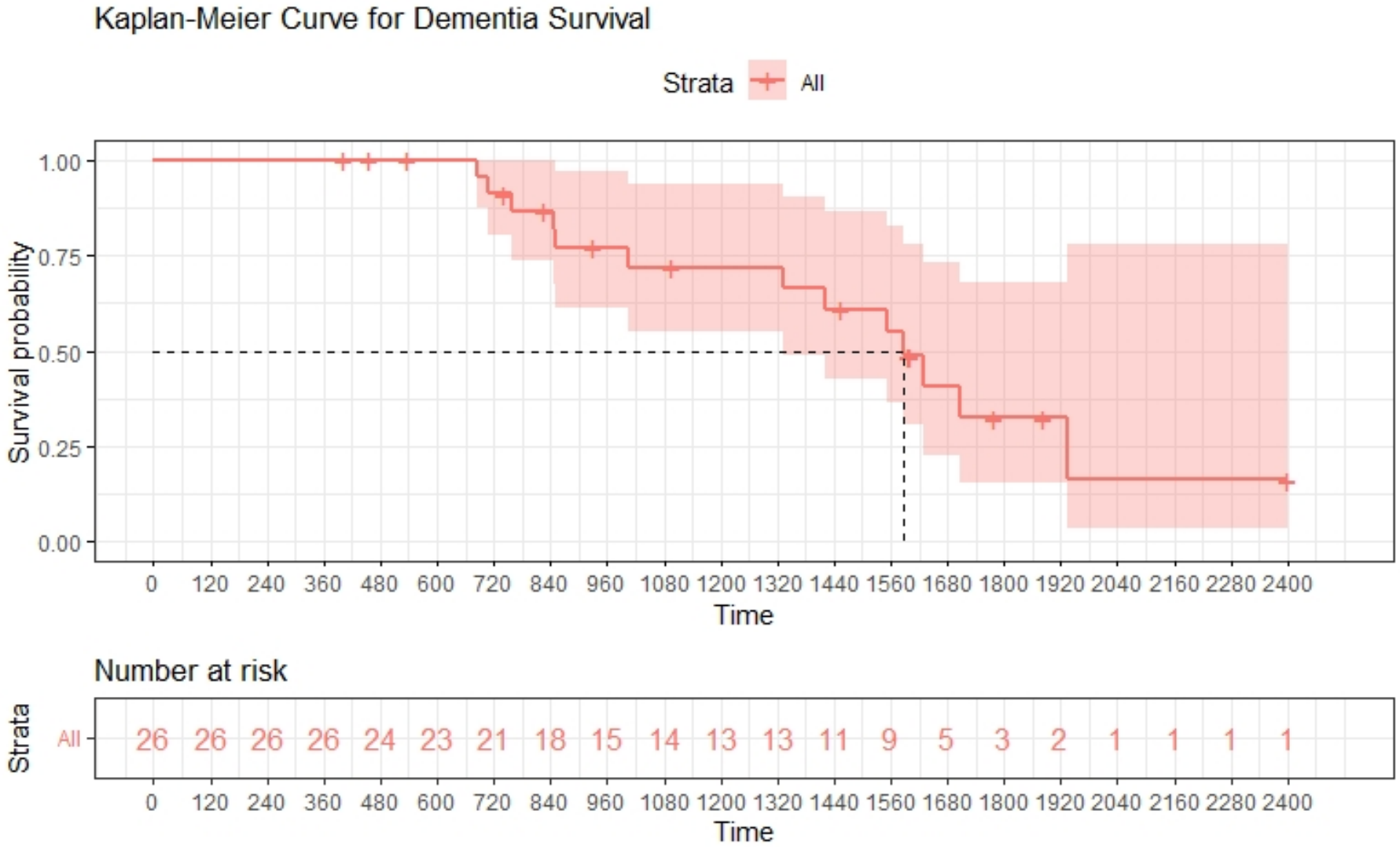
Survival Analysis (II)



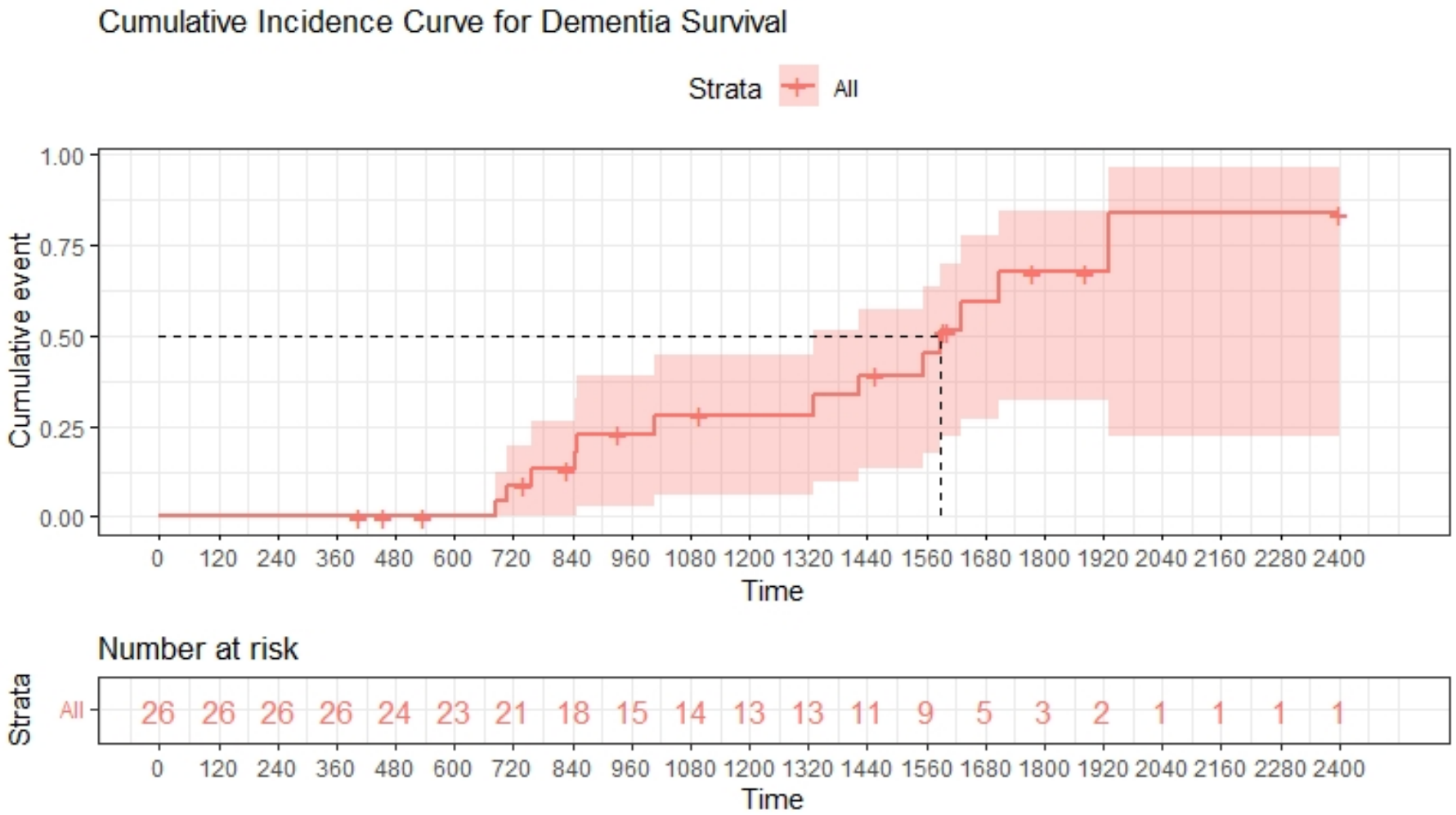
Survival Analysis(III)



Survival Analysis - Balanced (I)



Survival Analysis - Balanced (II)



Future questions