

Primary biliary cirrhosis

Matteo Fasulo, Simone Flavio Paris, Matteo Sivoccia

27/12/2021

1 Abstract

Nell'affrontare l'analisi dei PBC data ci siamo posti alcuni quesiti e abbiamo applicato delle metodologie nel tentativo di rispondere a tali domande. L'interesse primario è la sopravvivenza del paziente e come questa possa essere prolungata dal trattamento in un trial clinico. In secondo le correlazioni tra il genere dei pazienti, l'età e le analisi cliniche effettuate dai singoli individui nella stima della probabilità di sopravvivenza. Obiettivo finale dell'analisi è raggruppare gli individui che hanno caratteristiche simili tra loro e modellare la loro transizione nel tempo.

2 Introduction

La Colangite Primaria Sclerosante è una malattia autoimmune che porta al malfunzionamento di piccoli dotti biliari nel fegato. La progressione è lenta ma inesorabile portando spesso a cirrosi e deterioramento funzionale del fegato stesso. In letteratura è indicata come Cirrosi Biliare Primitiva sebbene la cirrosi occorra solo in stadi avanzati della patologia e per questo fu proposto un cambiamento nel 2014 in Colangite Primaria Biliare.

I dati provengono da un trial clinico tra il 1974 e il 1984 presso la Mayo Clinic su un totale di 424 pazienti. Di questi solo 312 parteciparono al trial sottoponendosi a cure di D-penicillina e placebo, mentre i restanti acconsentirono a fornire misurazioni base per studiare la sopravvivenza di tali individui. Sei di questi casi furono persi nel “follow-up” poco dopo la diagnosi.

3 Methods

I PBC data provenienti dalla Mayo Clinic presentano dei valori mancanti in corrispondenza dei biomarker analizzati che sono stati imputati con **Multivariate Imputation** (*mice*). Con la **PCA** si sono visualizzate le componenti principali e le caratteristiche di individui e variabili come considerazioni per le metodologie successive. Viene poi proposta un'analisi sul tempo di sopravvivenza tramite **Cox Proportional-Hazards model** e **Stimatore di Kaplan-Meier** (*survival*), cercando di migliorare i risultati dei due precedenti metodi attraverso l'implementazione **ensembled** di un **Random Forest** (*randomForestSRC*). Infine si è cercato di analizzare le transizioni dei pazienti nel tempo individuando degli stati di appartenenza attraverso **Latent Markov Models** (*Lmest*).

4 Results

Le analisi sviluppate mostrano che non c'è un effetto statisticamente significativo tra i pazienti trattati con D-penicillina e placebo, inoltre sembra esserci un'incidenza maggiore della malattia sulle donne che però

rispondono in maniera migliore al decorso della patologia. Allo stesso modo (intuitivamente) l'età incide sulla risposta immunitaria ed è stato notato un effetto non del tutto lineare con questa. La feature più significativa per distinguere tra i pazienti è sicuramente la *bilirubina* che, a causa della rottura dei dotti biliari, si riversa nel sangue. E' stato possibile, confrontando lo status originario con l'ultimo stato del modello di Markov, interpretare ognuno degli stati con un livello di prognosi differente. Lo stato dei pazienti con la prognosi migliore è il *secondo stato*.

5 Discussion

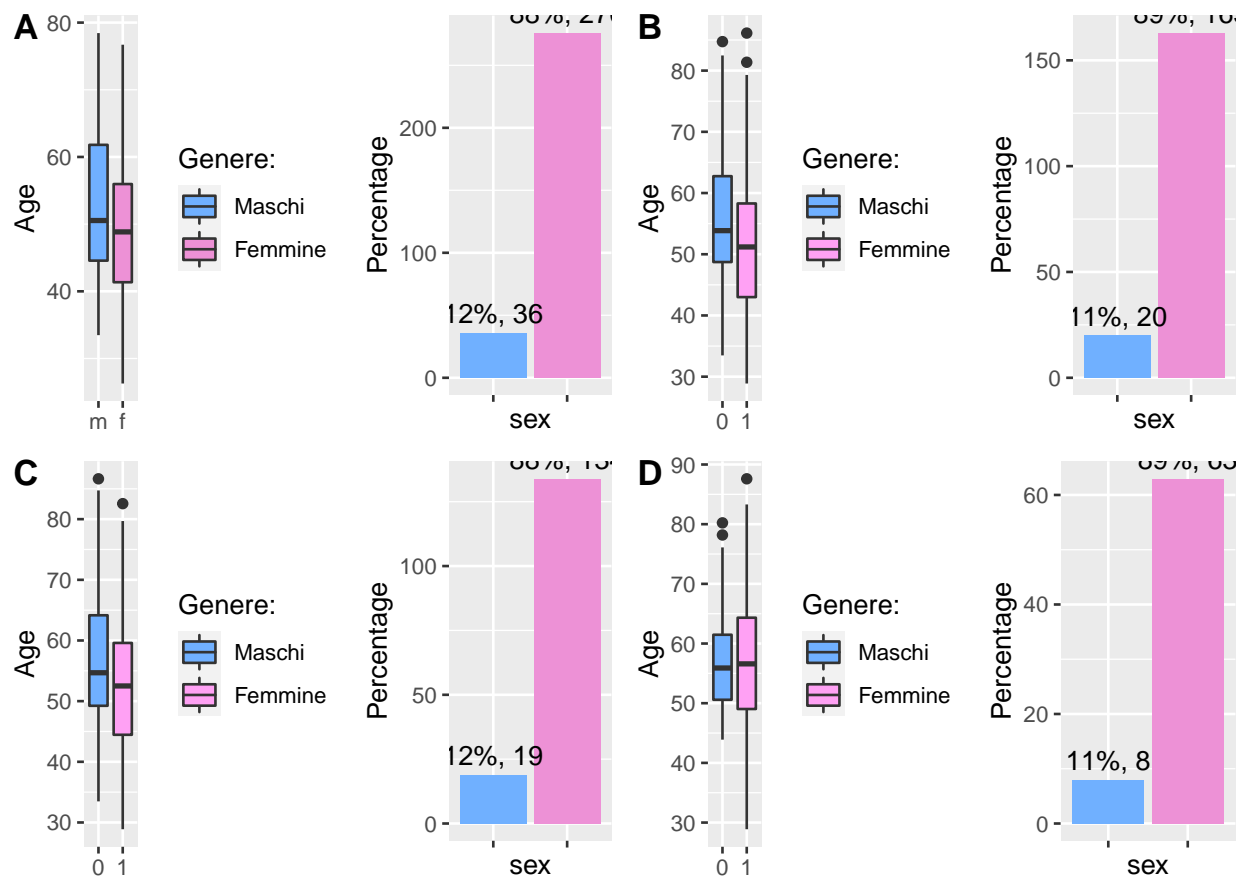
Forniamo per la riproducibilità della ricerca la lista dei pacchetti utilizzati per computare tutte le analisi di seguito.

```
## [1] "LMest"           "pec"             "prodlm"          "randomForestSRC"
## [5] "survminer"       "ggpubr"          "factoextra"      "FactoMineR"
## [9] "gamlss"          "nlme"            "gamlss.dist"     "MASS"
## [13] "gamlss.data"     "forcats"         "stringr"         "dplyr"
## [17] "purrr"           "readr"           "tidyr"           "tibble"
## [21] "tidyverse"       "magrittr"        "mice"            "survival"
## [25] "ggplot2"
```

Le statistiche descrittive relative alle principali variabili di risposta sono riassunte nella seguente tabella

```
##      age      bili      chol      albumin
## Min.   :26.28  Min.    : 0.100  Min.    : 55.0  Min.    :1.17
## 1st Qu.:44.63  1st Qu.: 0.800  1st Qu.: 235.0  1st Qu.:3.11
## Median :52.52  Median : 1.400  Median : 281.0  Median :3.44
## Mean   :52.40  Mean    : 3.672  Mean    : 320.5  Mean    :3.39
## 3rd Qu.:59.78  3rd Qu.: 3.900  3rd Qu.: 349.2  3rd Qu.:3.70
## Max.   :84.65  Max.    :41.000  Max.    :1775.0  Max.    :8.01
##
##      alk.phos      ast      platelet      protime
## Min.    : 73      Min.    : 6.2      Min.    : 40.0  Min.    : 9.0
## 1st Qu.: 737      1st Qu.: 72.0      1st Qu.:165.0  1st Qu.:10.1
## Median :1072      Median :107.0      Median :228.0  Median :10.8
## Mean    :1382      Mean    :122.7      Mean    :233.7  Mean    :11.0
## 3rd Qu.:1636      3rd Qu.:155.0      3rd Qu.:290.2  3rd Qu.:11.5
## Max.    :13862     Max.    :1205.0     Max.    :991.0  Max.    :36.0
## NA's    :60              NA's    :73
```

Tra i 312 individui l'età media delle **donne** è più bassa (49 anni) rispetto a quella degli **uomini** (53 anni). Ogni individuo viene visitato di media 6 volte, indipendentemente dal genere. Le visite sono fatte sei mesi dopo la prima, 1 anno dopo la prima e poi 1 anno una dall'altra.



Nell'analisi dei PBC data forniti sono stati trovati molti valori nulli in corrispondenza delle analisi cliniche dei pazienti. Tra le possibili soluzioni in materia alla presenza di valori nulli, vi è l'**imputazione multipla**. Attraverso la funzione *mice* si sono imputati i valori mancanti sia delle variabili numeriche che di quelle categoriche con i metodi disponibili nel pacchetto *mice*. Per le variabili numeriche è stato utilizzato il metodo **predictive mean matching** mentre con le variabili categoriche la **logistic regression**. Poiché l'imputazione non è deterministica, sono stati imputati dieci diversi set di dati a partire dai valori mancanti e si è scelto di utilizzare la media tra tutti questi come valore finale.

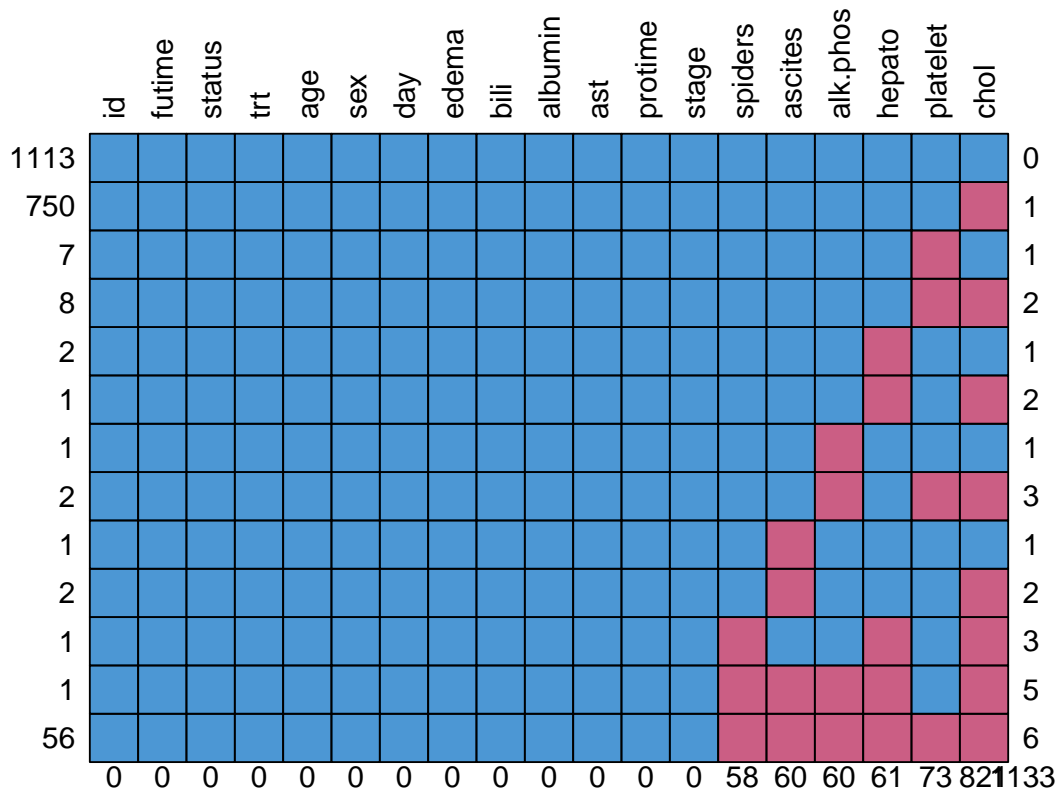


Figure 1: Missing Data pattern for PBC data where x axis represents the number of values missing of that specific variable and y axis the number of rows with missing values of those variables

Per il density plot delle variabili imputate rispetto alle originali si veda Figura 10

La **Principal Component Analysis** è tra i più famosi metodi esplorativi di analisi dei dati e viene adoperata per racchiudere le informazioni di set di dati con tante variabili in **componenti principali** che spiegano un certa percentuale della varianza totale. Poiché i PBC data presentano diverse variabili (i biomarker) sia numeriche che categoriche, si è cercato di riassumere il contenuto di queste attraverso la PCA.

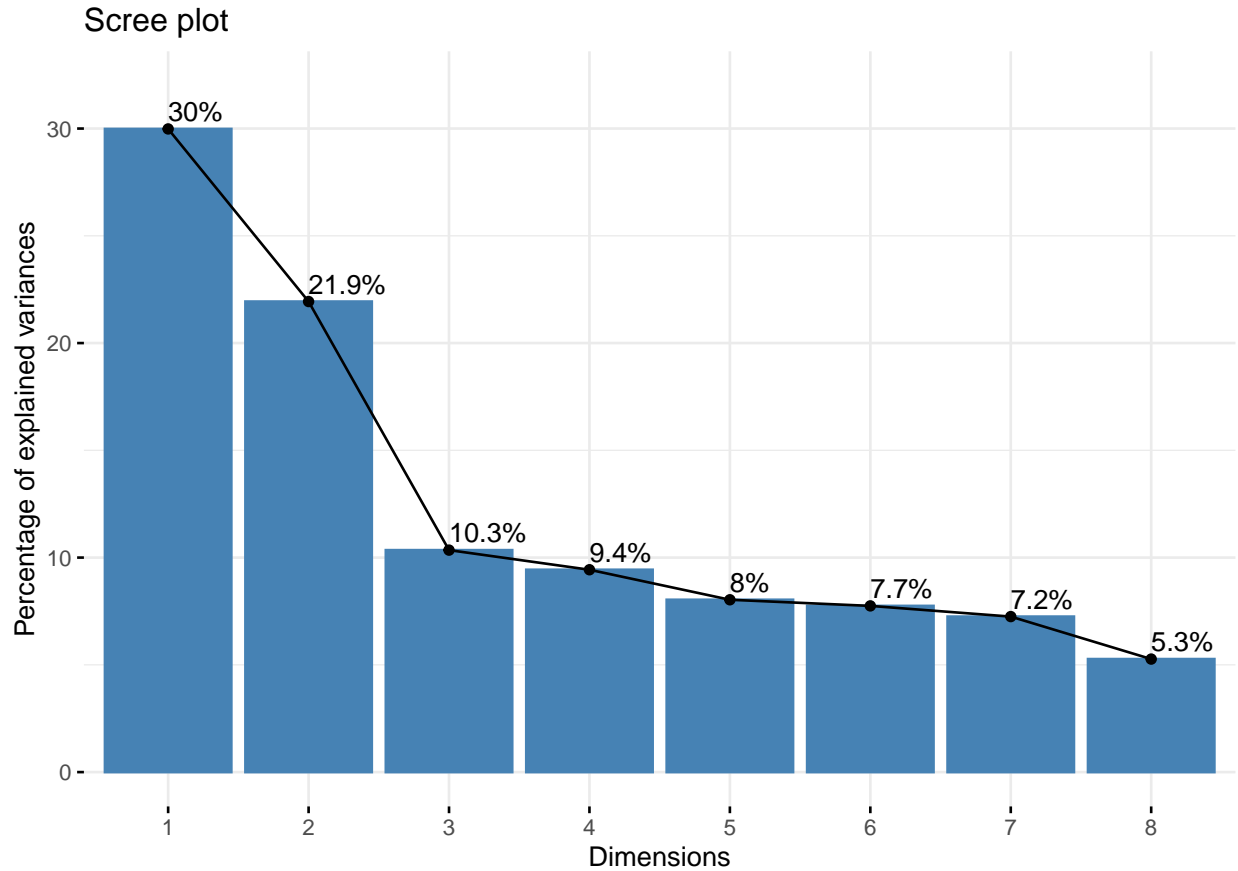


Figure 2: Scree Plot of PCA on PBC data

Dallo **screeplot** si nota come le prime due dimensioni rappresentano oltre il 50% della varianza totale e dunque sono di notevole rilievo per l'interpretazione dei dati.

Nella funzione PCA di **FactoMineR** viene consentito l'uso di variabili quantitative e categoriche di supporto per l'interpretabilità dei risultati ottenuti. Si è scelto dunque di utilizzare come variabili di supporto quantitative l'età dei pazienti e il tempo tra una visita e l'altra mentre le qualitative sono il trattamento ricevuto, il genere, lo stadio della malattia, presenza di ascite, malformazione dei vasi sanguigni nella pelle, presenza di epatomegalia o fegato ingrossato, presenza o meno di edema sia esso non tratto, parzialmente trattato o presente nonostante una terapia diuretica.

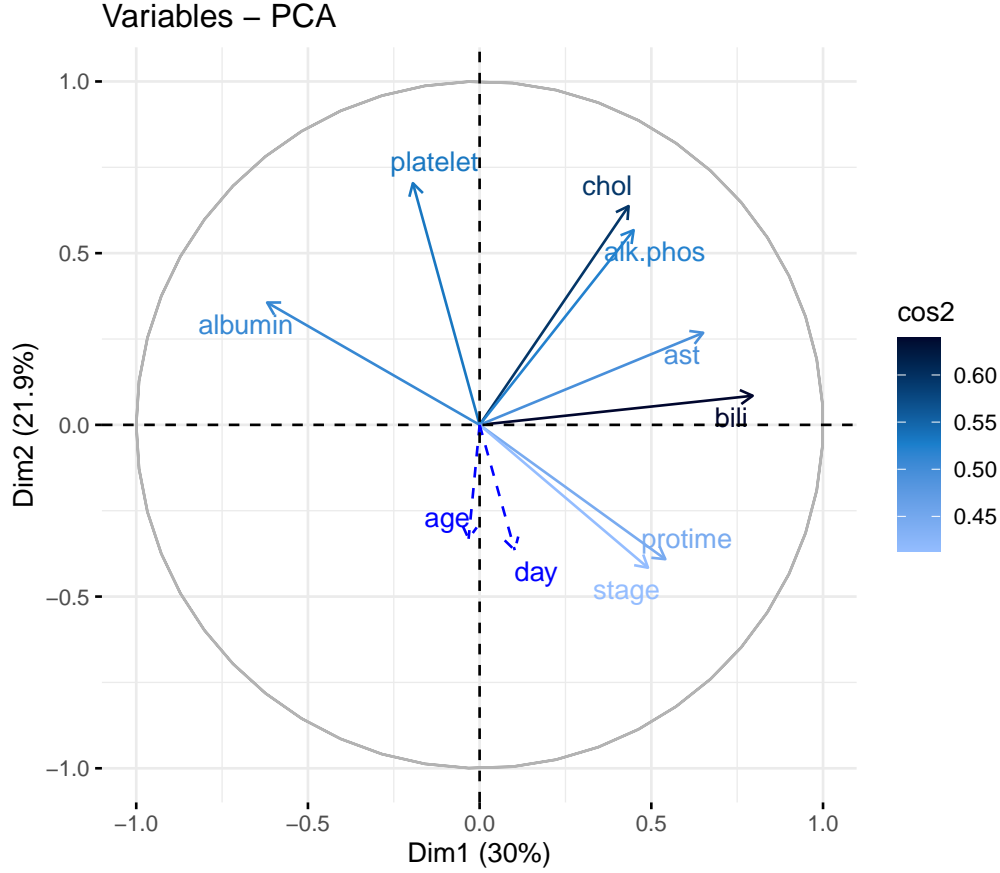


Figure 3: Top 10 most important variables for cos2

Le due dimensioni risultanti sono suddivisibili in:

- Dimensione 1: *bili, protime, stage, ast, albumin*
- Dimensione 2: *platelet, chol, alk.phos*

Le prime considerazioni riguardano l'età dei pazienti che segue il semiasse negativo della dimensione 2 fornendo l'informazione che gli individui nel semiasse positivo della dimensione 2 siano tutti giovani. Inoltre le variabili *protime* e *stage* sono quasi sovrapposte indicando che all'aumentare della gravità della malattia, il tempo di coagulazione tende ad aumentare, questo perché negli stadi avanzati si perde la capacità di controllare la normale coagulazione del sangue. A conferma di ciò nella parte opposta troviamo *albumin* che infatti rappresenta la densità dei globuli rossi nel sangue. Per quanto riguarda le altre variabili, esse sono tutte strettamente legate alle funzionalità epatiche di base. Quindi interpretiamo la dimensione 1 come la funzionalità epatica e la dimensione 2 come la capacità di coagulazione del sangue.

Per il risultato della **Cluster Analysis** sulla PCA si veda Figura 11

La nostra analisi prosegue con i **modelli di sopravvivenza**. Come prima metodologia abbiamo usato lo stimatore di Kaplan-Meier per calcolare la funzione di sopravvivenza. Un importante vantaggio della curva di Kaplan-Meier è che il metodo può tenere conto di alcuni tipi di dati censurati, particolarmente della censura a destra, che si verifica se un paziente si ritira da uno studio. Lo stimatore di Kaplan-Meier è la stima non parametrica della massima verosimiglianza di $S(t)$:

$$\hat{S}(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i}$$

Quando non vi è censura, n_i è solo il numero di sopravvissuti appena prima del tempo t_i . Con la censura, n_i è il numero di sopravvissuti meno il numero di perdite (casi censurati).

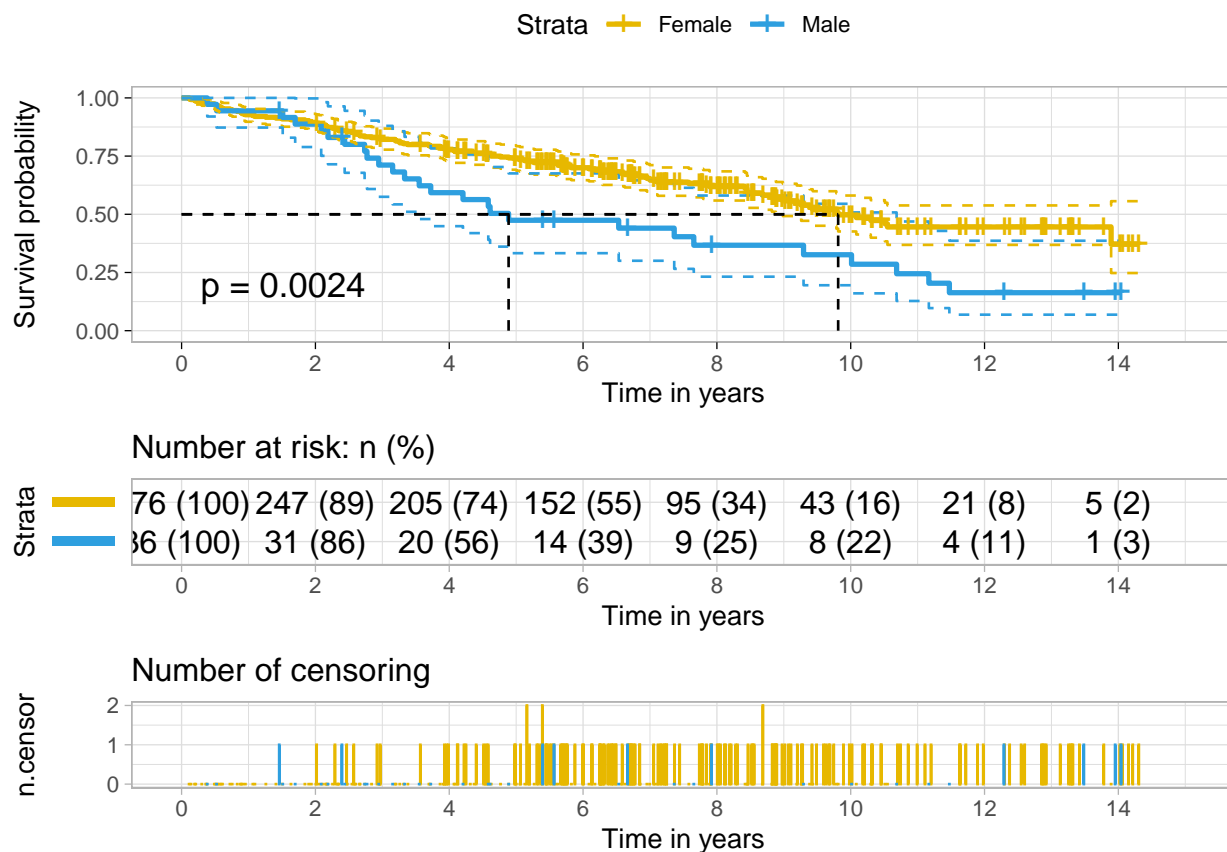


Figure 4: Gender Survival probability

Possiamo notare che la probabilità di sopravvivenza delle donne è maggiore di quella degli uomini e che questo effetto è statisticamente significativo. Nella seconda tabella è mostrato il numero di soggetti rimanenti ogni 2 anni e in percentuale rispetto a t_0 stratificato per genere. Nell'eventualità di analizzare l'efficacia dei trattamenti sui nostri soggetti, abbiamo rappresentato la probabilità di sopravvivenza come funzione del trattamento ricevuto (penicillina o placebo) disponibile in Figura 12

Attraverso la funzione *survreg* abbiamo sviluppato un **Accelerated Failure Time Model**. Gli *AFT* sono modelli parametrici che forniscono un'alternativa ai **proportional hazards models**. Mentre un *proportional hazards model* assume che l'effetto delle covariate moltiplichi il rischio di qualche costante, gli *AFT* interpretano le covariate come un'accelerazione o decelerazione del corso di vita di una malattia di qualche costante. Questa tecnica, infatti, è indicata per malattie che sono formate da sequenze di stadi intermedi. Essendo appunto un modello parametrico, assume una distribuzione sottostante ai dati. Nella nostra analisi abbiamo provato ad utilizzare 4 diverse distribuzioni, osservando quale fosse la "migliore" come fit dei dati (Weibull).

Per verificare quale modello tenere, considerando anche modelli complessi dove magari l'età del paziente non segue un andamento lineare con la probabilità di sopravvivenza, abbiamo effettuato un test **anova** tra il semplice modello lineare con la **Weibull** e un modello con **effetto quadratico** dell'età.

Terms Resid. Df -2*LL Test Df Deviance Pr(>Chi)

```
## 1          trt + sex + age      307 989.4596      NA      NA      NA
## 2 (trt + sex) * (age + I(age^2)) 302 984.8790      = 5 4.580586 0.4691742
```

Da qui ricaviamo che il modello più semplice è quello da considerare.

Abbiamo inoltre analizzato l'effetto di utilizzo del farmaco a seconda del genere del paziente notando che c'è una minima differenza (seppur su scala logaritmica) tra coloro a cui è stata somministrata la penicillina e chi ha ricevuto il placebo.

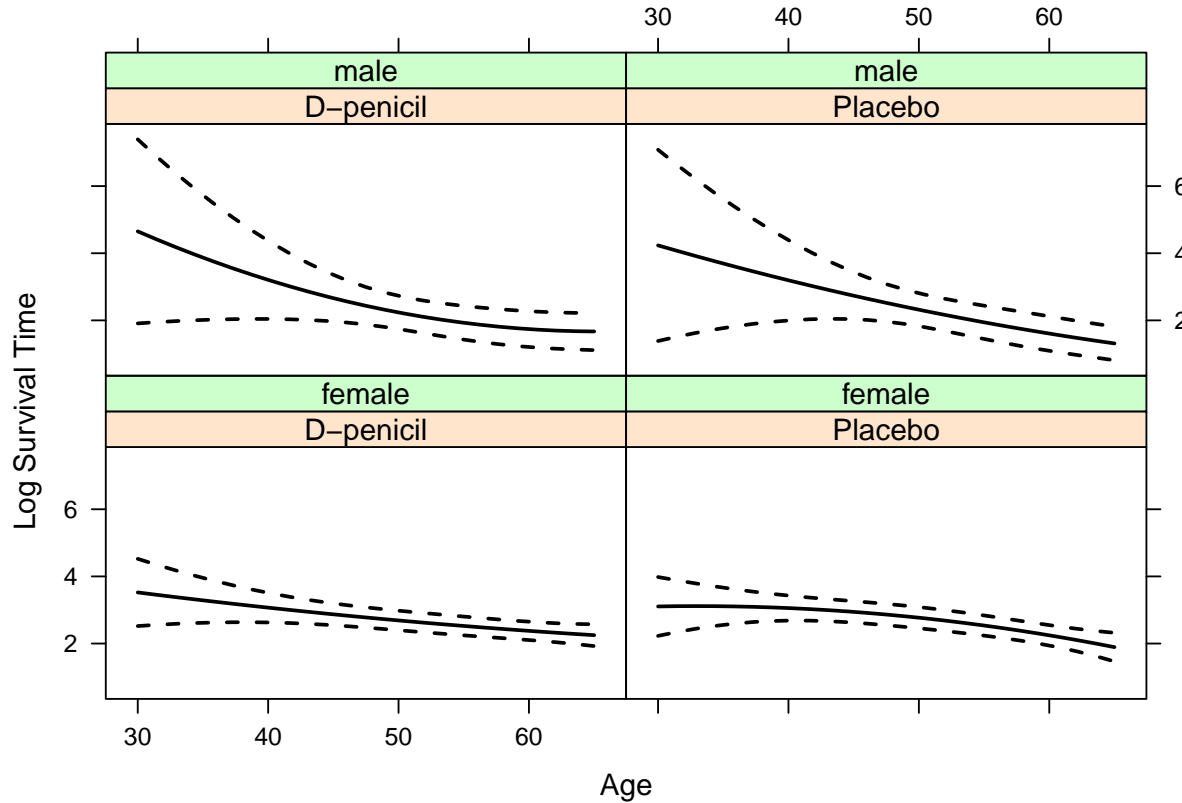


Figure 5: Treatment stratified by gender

Per visualizzare quale fosse il fit migliore di un modello che tenesse conto dell'interazione tra trattamento e genere, abbiamo rappresentato il modello con *Weibull*, *log-Normal* e *Log-logistic* in Figura 13.

Utilizzando invece il **Cox proportional hazards model** spesso abbreviato con **Cox**, abbiamo ripercorso l'esperienza precedente e confermato che il trattamento non altera significativamente la probabilità di sopravvivenza come in Figura 14.

Poiché la patologia è strettamente legata alla quantità di bilirubina nel sangue, ci sembrava opportuno rappresentare le probabilità di sopravvivenza divise per età e intervalli di bilirubina.

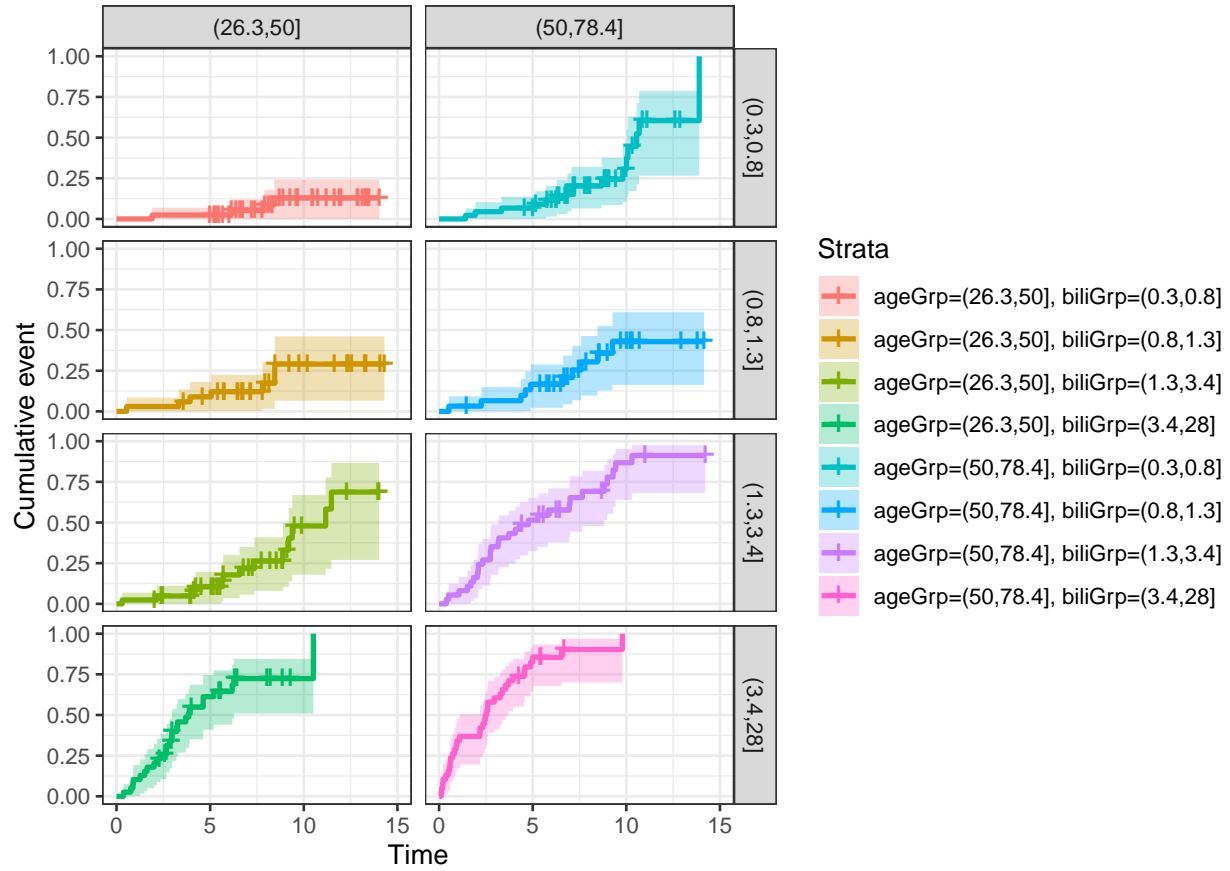


Figure 6: Cumulative event by age and bili (first occurrence)

Il grafico relativo all'ultima visita è disponibile in Figura 15.

Inoltre, modellando l'incidenza cumulativa tra pazienti che hanno subito trapianto e coloro che invece hanno sperimentato l'evento, notiamo che il follow-up di coloro che hanno subito un trapianto segue un andamento logaritmico contro il quasi-lineare dei morti.

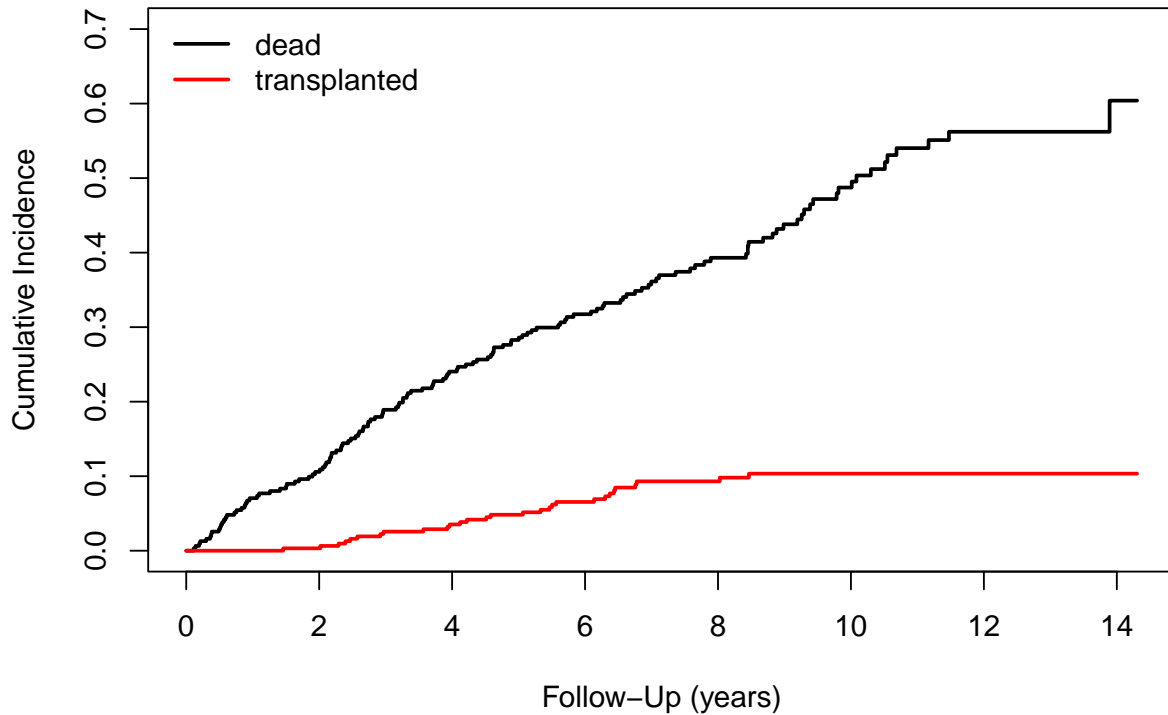


Figure 7: Cumulative incidence for transplanted and dead

Documentandoci sulle metodologie per i modelli di sopravvivenza, abbiamo notato che spesso si preferisce utilizzare il **Random Forest** rispetto al modello **Cox**. Abbiamo dunque messo a confronto l'errore di predizione dei due diversi modelli al variare del tempo. Il **Random Forest** è particolarmente indicato perché utilizza un approccio **ensemble** creando diversi **decision tree** e sviluppando parallelamente n modelli diversi. Il modello del Random Forest assicura che tutti gli alberi siano de-correlati l'uno dall'altro e ad ogni split, valuta la miglior feature con cui suddividere le istanze a seconda di criteri come *LogRank*, *BS-Gradient* e *LogRankScore*. L'implementazione sviluppata è su una foresta di 500 alberi mentre il resto degli **hyper-parameters** è stato trovato attraverso **grid-search**. Per minimizzare l'errore di predizione, il *random forest* presentato è basato su **bootstrap-cross-validation** con 50 iterazioni.

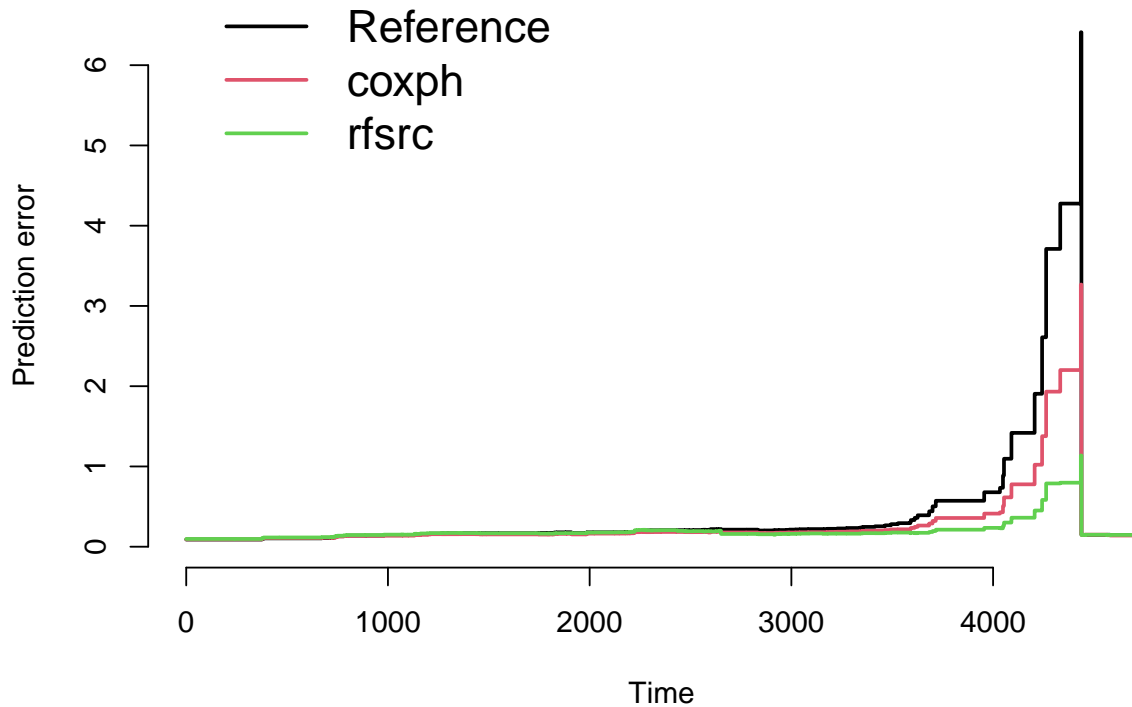


Figure 8: Prediction error between Cox-Proportional Hazard and Random Forest

La tabella con i valori sull'errore di predizione è disponibile nella sezione Tables

Poiché esistono diverse regole per lo splitting ad ogni nodo decisionale nell'albero, abbiamo confrontato gli errori di predizione usando tutte le tecniche per scegliere il miglior criterio di splitting. Il risultato è che il *LogRank* è il miglior criterio tra quelli disponibili (Tables). Nei decision-tree è fondamentale considerare anche l'altezza alla quale si effettua lo splitting. Le feature più vicine alla radice sono quelle in cui la maggior parte delle istanze viene suddivisa e possono essere considerate come le principali per suddividere nettamente tra le osservazioni. Siccome solamente l'altezza non basta a garantire che una determinata feature sia centrale per lo splitting, abbiamo riassunto le migliori features di tutti e 500 gli alberi sia per profondità minima nell'albero che per importanza attraverso il criterio **VIMP**. Il **VIMP** o Variable Importance attraverso la misurazione dell'Entropia sceglie tra le migliori feature quelle che permettono di avere la minore incertezza nella classificazione di istanze. Si tratta dunque di una metrica il cui valore deve tendere ad essere il più grande possibile.

##	md	vimp1	vimp2
## bili	2.06	18.011	16.6090
## protime	3.08	10.053	9.5696
## ascites	3.65	13.255	12.1683
## edema	3.70	8.388	10.4212
## alk.phos	3.78	2.229	3.6670
## ast	3.80	0.853	1.9357
## albumin	3.89	2.016	0.6952
## chol	3.97	1.475	0.8865

```
## age      4.16  1.412  1.8538
## platelet 4.36  0.422  0.1399
## sex      5.28  4.548  3.9736
## spiders  6.16  1.466  0.9597
## hepato   6.55  0.737  1.0930
## trt      7.14 -0.134 -0.0381
```

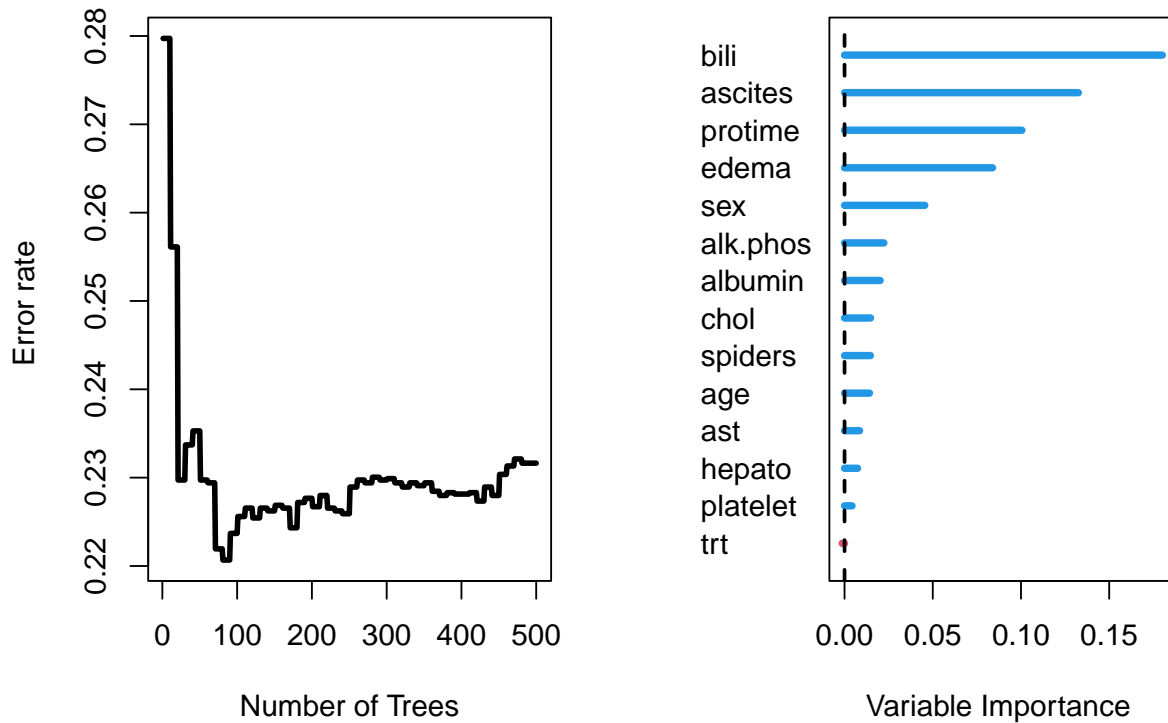


Figure 9: Error rate and Variable Importance

Per modellare la transizione nel tempo dei pazienti abbiamo utilizzati un **Latent Markov Model**. La libreria *Lmest* permette attraverso una procedura deterministica e stocastica combinata di trovare il numero ottimale di stati in un intervallo definito. Abbiamo dunque computato per prima cosa un modello senza covariate sia nel modello osservato che nel modello latente per individuare il numero ottimale di stati; poi si è scelto il migliore per il criterio **BIC** (Figura 16) e l'analisi è proseguita su tale modello. Si sarebbero potuti scegliere anche 6 stati ma abbiamo preferito l'interpretabilità del risultato ad un **BIC** migliore. Di seguito sono riassunte le medie condizionate dei biomarker (su scala logaritmica per evitare valori anomali) in ogni stato:

```
##          state
##          1      2      3      4
## Bilirubin 0.8387033 -0.5049485 0.06231439 1.965247
## Albumin   1.1838997 1.3105085 1.27515541 1.184774
## Alkaline   6.8677097 6.3378522 7.29784634 7.710743
## Cholesterol 5.5283870 5.5101531 5.70820350 6.175918
```

```
## Transaminase 4.7472262 4.0405023 4.77568151 5.199302
## Platelets 4.7544615 5.4971288 5.58286355 5.612687
## Prothrombin 2.4172307 2.3465657 2.33745190 2.422196
```

Per la visualizzazione grafica della matrice di transizione si veda Figura 17.

Ci è sembrato opportuno unire con un'unica tabella sia le probabilità iniziali di appartenenza ad uno stato che la matrice di transizione

```
##          Good  Best  Bad  Worst
## init  0.1190 0.2490 0.5296 0.1023
## Good  0.7702 0.0889 0.0824 0.0585
## Best  0.0696 0.8030 0.0745 0.0529
## Bad   0.0734 0.0848 0.7860 0.0558
## Worst 0.0925 0.1069 0.0991 0.7015
```

La maggior parte dei pazienti (53%) all'inizio del trial vengono osservati nel 3 stato che definiremo come **Bad**; il restante è suddiviso nel secondo stato (**Best** 25%), primo stato **Good** (12%) e quarto stato **Worst** (10%).

Riteniamo importante mostrare anche la seguente matrice dove le varianze stimate sono sulla diagonale principale, le covarianze sulla triangolare inferiore e le correlazioni sulla triangolare superiore, possiamo osservare che:

- La **bilirubina** ha con il tempo di **protrombina** una buona correlazione **positiva**, mentre entrambe sono correlate negativamente con l'**Albumina**.
- La **fosfatasi alcalina** è correlata **positivamente** con le **transaminasi** e le **piastrine**.

```
##          Bili  Albu  Alka  Chol  Trans  Plat  Proth
## Bilirubin  0.256 -0.106 0.148 0.103 0.242 -0.100 0.110
## Albumin   -0.007 0.017 -0.015 0.089 0.061 0.072 -0.171
## Alkaline   0.038 -0.001 0.258 0.223 0.304 0.130 0.000
## Cholesterol 0.018 0.004 0.039 0.119 0.147 0.228 -0.162
## Transaminase 0.046 0.003 0.058 0.019 0.141 0.000 -0.030
## Platelets  -0.016 0.003 0.021 0.025 0.000 0.101 -0.106
## Prothrombin 0.005 -0.002 0.000 -0.005 -0.001 -0.003 0.008
```

6 Limitations

Non è stato possibile considerare uno stato aggiuntivo di **dropout** nella stima del modello di Markov e tale considerazione è da ritenersi come spunto per future analisi. Una volta raggiunto tale stato, l'individuo non ha più la possibilità di effettuare alcun tipo di transizione verso uno stato successivo o precedente in quanto esso corrisponde alla cessazione delle analisi da parte del paziente nel trial per tutti i possibili motivi.

7 Tables and Figures

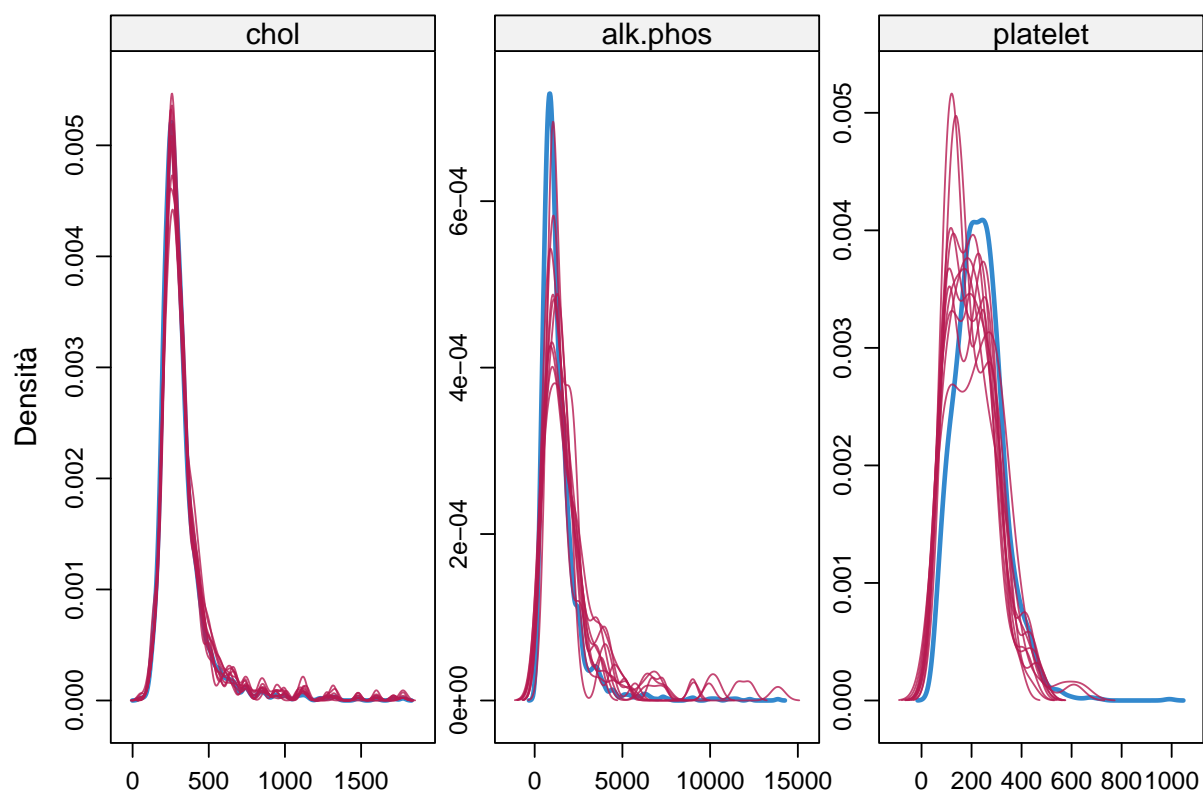


Figure 10: Density plot of imputed values vs non missing values.

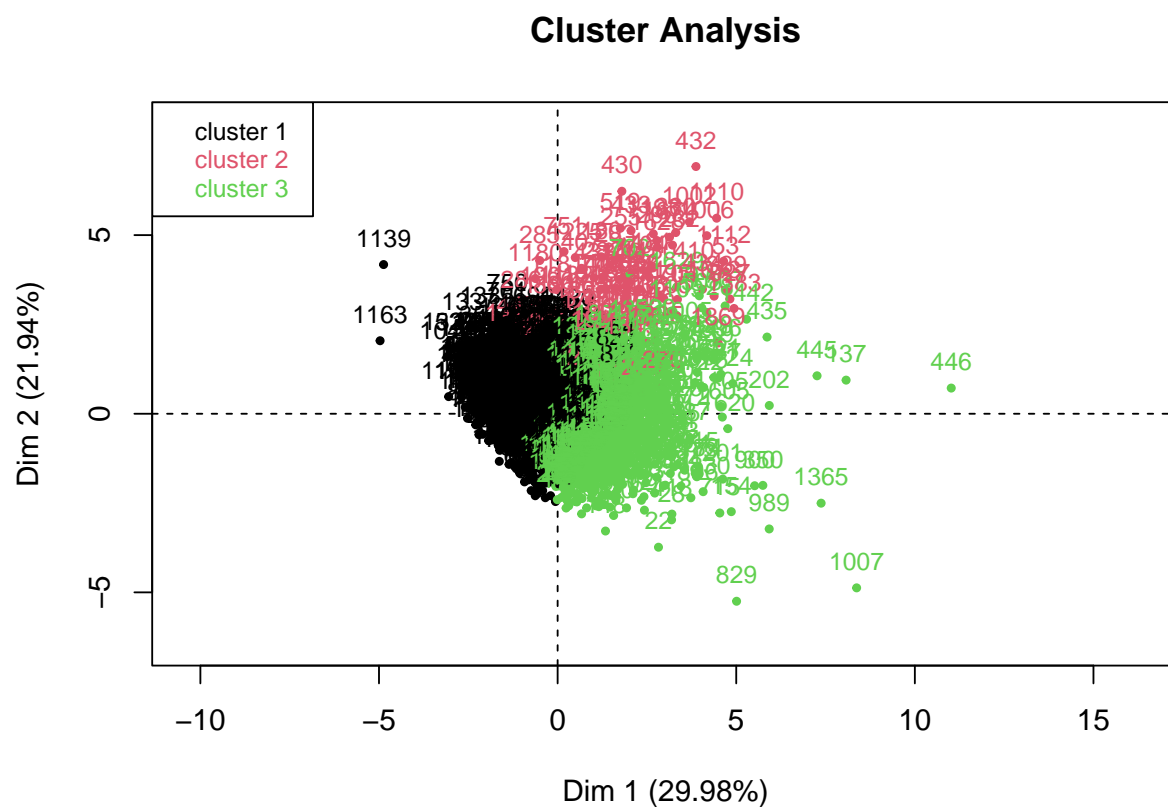


Figure 11: Cluster Analysis on PCA components

Da qui è visibile come il gruppo *nero* rappresenti le persone con la prognosi migliore (che comunque hanno problemi con la concentrazione di piastrine e con la densità dei globuli rossi), il *rosso* coloro che hanno valori di colesterolo, bilirubina molto alti e di alcalino fosfati mentre il *verde* coloro che hanno lo stadio più avanzato della malattia che già sta intaccando le funzionalità epatiche del fegato e i dotti biliari trasformandosi in cirrosi.

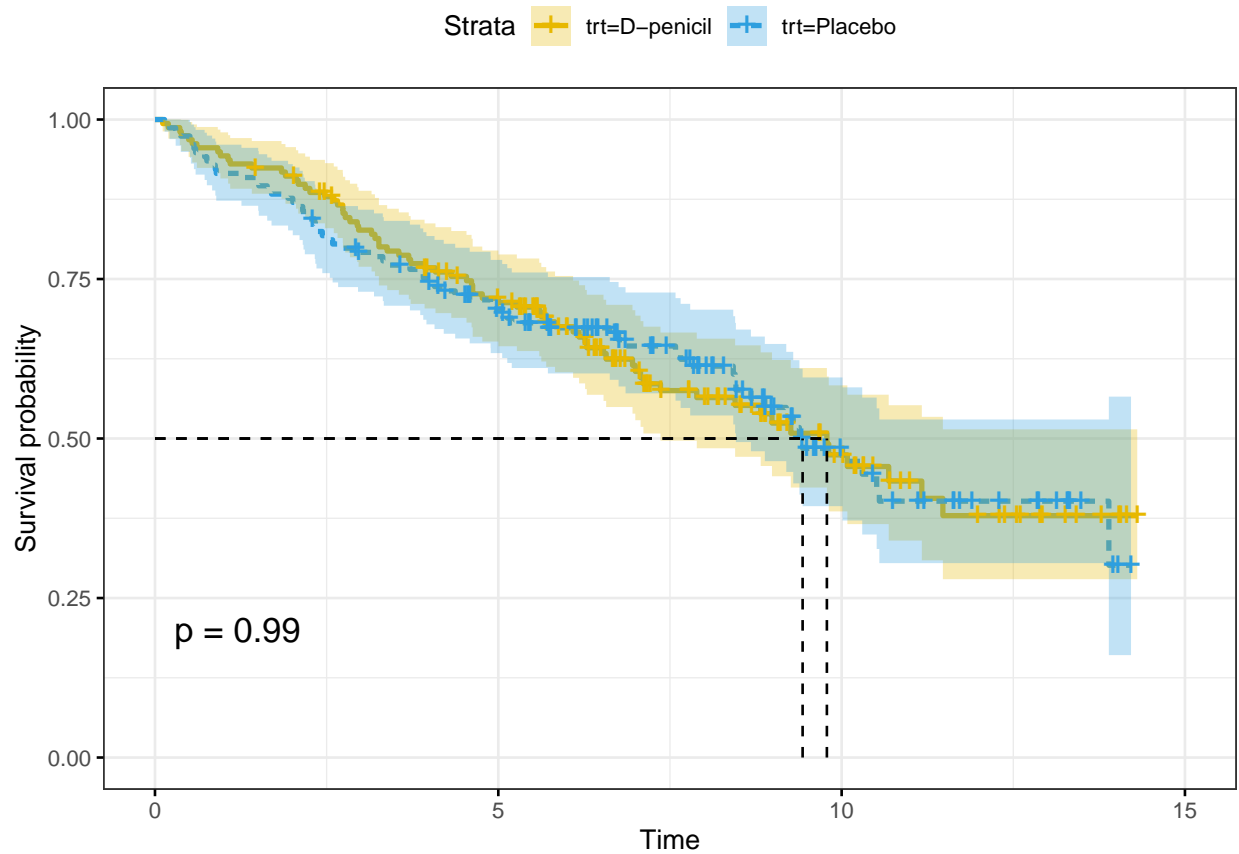


Figure 12: Treatment Survival probability

```
##
## Call:
## survreg(formula = Surv(years, status2) ~ trt + sex + age, data = pbcseq.survival.head)
##           Value Std. Error      z      p
## (Intercept)  4.71378    0.46090 10.23 <2e-16
## trtPlacebo   -0.13046    0.15576  -0.84  0.402
## sexmale      -0.44472    0.20147  -2.21  0.027
## age          -0.03874    0.00793  -4.89 1e-06
## Log(scale)   -0.10223    0.07423  -1.38  0.168
##
## Scale= 0.903
##
## Weibull distribution
## Loglik(model)= -494.7   Loglik(intercept only)= -511.8
##  Chisq= 34.23 on 3 degrees of freedom, p= 1.8e-07
## Number of Newton-Raphson Iterations: 5
## n= 312
##
## Call:
## survreg(formula = Surv(years, status2) ~ trt + sex + age, data = pbcseq.survival.head,
##         dist = "exponential")
##           Value Std. Error      z      p
```



```

## (Intercept)  4.9535      0.4750 10.43 < 2e-16
## trtPlacebo  -0.1438      0.1721 -0.84   0.40
## sexmale     -0.4816      0.2217 -2.17   0.03
## age         -0.0420      0.0084 -5.00 5.7e-07
##
## Scale fixed at 1
##
## Exponential distribution
## Loglik(model)= -495.6   Loglik(intercept only)= -512.3
##  Chisq= 33.35 on 3 degrees of freedom, p= 2.7e-07
## Number of Newton-Raphson Iterations: 4
## n= 312

##
## Call:
## survreg(formula = Surv(years, status2) ~ trt + sex + age, data = pbcseq.survival.head,
##         dist = "lognormal")
##               Value Std. Error      z      p
## (Intercept)  4.92969    0.52473  9.39 < 2e-16
## trtPlacebo  -0.21196    0.18775 -1.13   0.26
## sexmale     -0.30242    0.27761 -1.09   0.28
## age         -0.04841    0.00944 -5.13 2.9e-07
## Log(scale)   0.34711    0.06473  5.36 8.2e-08
##
## Scale= 1.41
##
## Log Normal distribution
## Loglik(model)= -498.9   Loglik(intercept only)= -515.2
##  Chisq= 32.72 on 3 degrees of freedom, p= 3.7e-07
## Number of Newton-Raphson Iterations: 3
## n= 312

##
## Call:
## survreg(formula = Surv(years, status2) ~ trt + sex + age, data = pbcseq.survival.head,
##         dist = "loglogistic")
##               Value Std. Error      z      p
## (Intercept)  4.52995    0.48921  9.26 < 2e-16
## trtPlacebo  -0.14532    0.17387 -0.84 0.40325
## sexmale     -0.38446    0.24161 -1.59 0.11155
## age         -0.04151    0.00884 -4.69 2.7e-06
## Log(scale)  -0.26129    0.07326 -3.57 0.00036
##
## Scale= 0.77
##
## Log logistic distribution
## Loglik(model)= -497.4   Loglik(intercept only)= -513
##  Chisq= 31.08 on 3 degrees of freedom, p= 8.2e-07
## Number of Newton-Raphson Iterations: 4
## n= 312

```

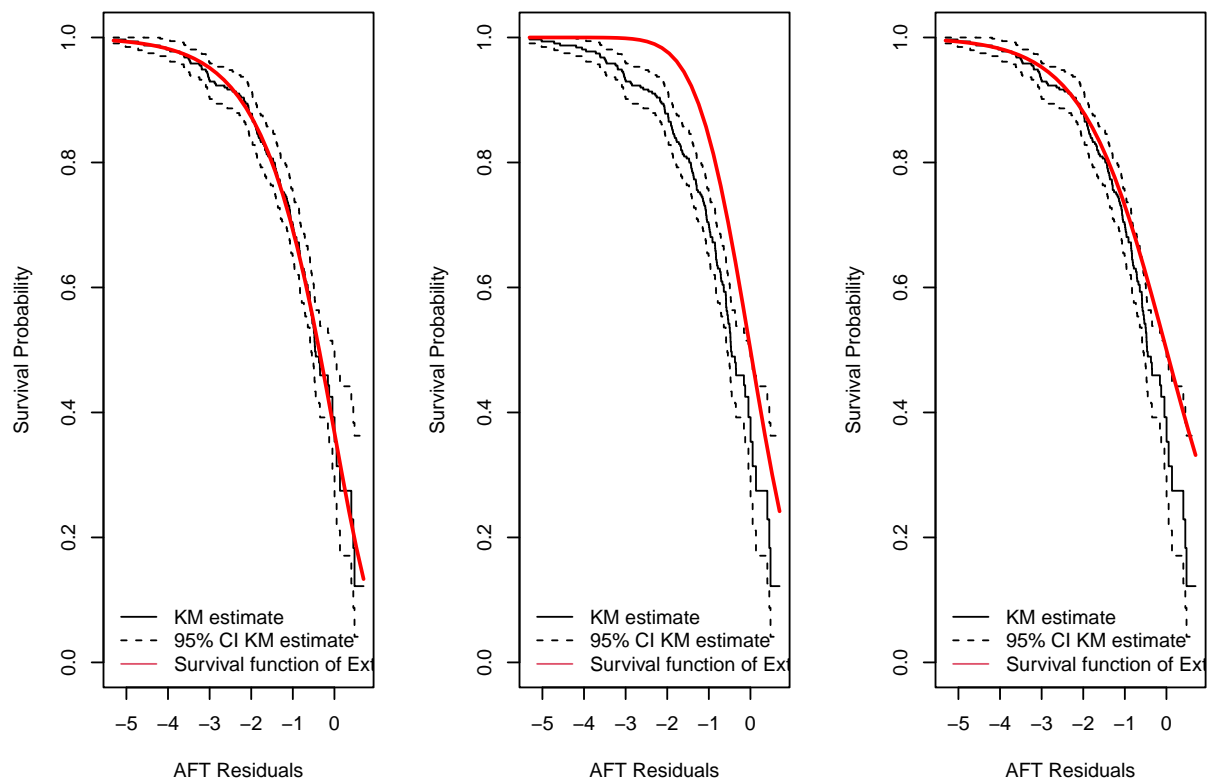


Figure 13: Fit on PBC data

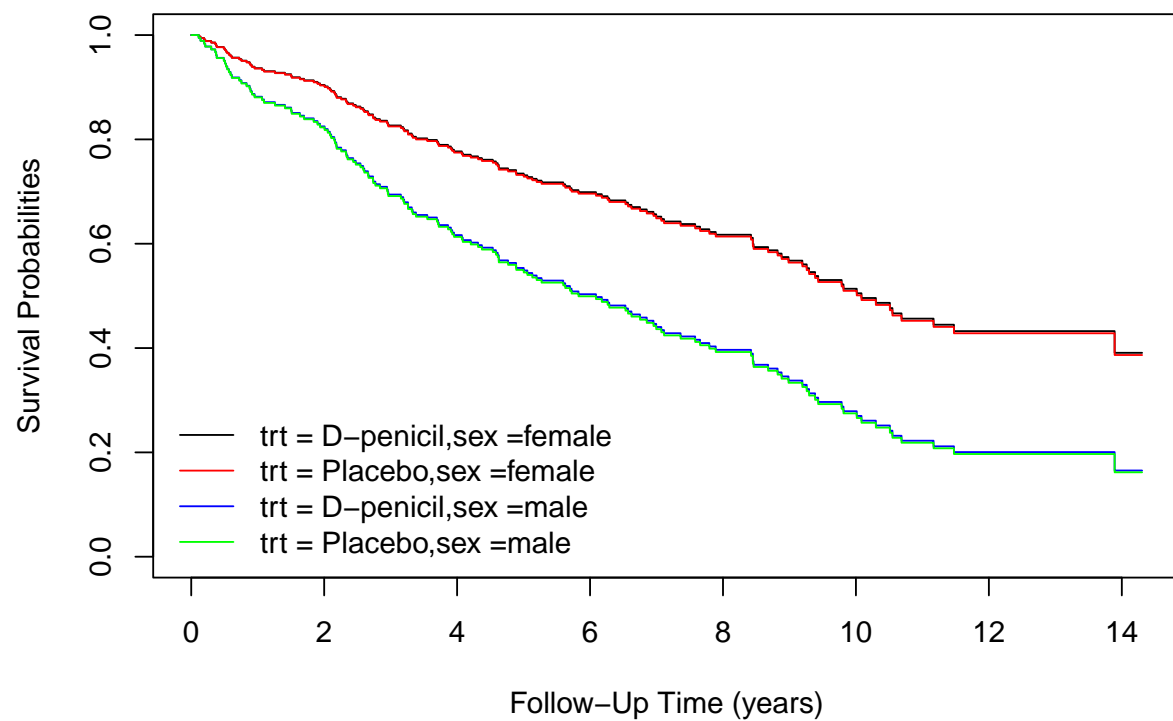


Figure 14: Cox survival model stratified by gender and treatment

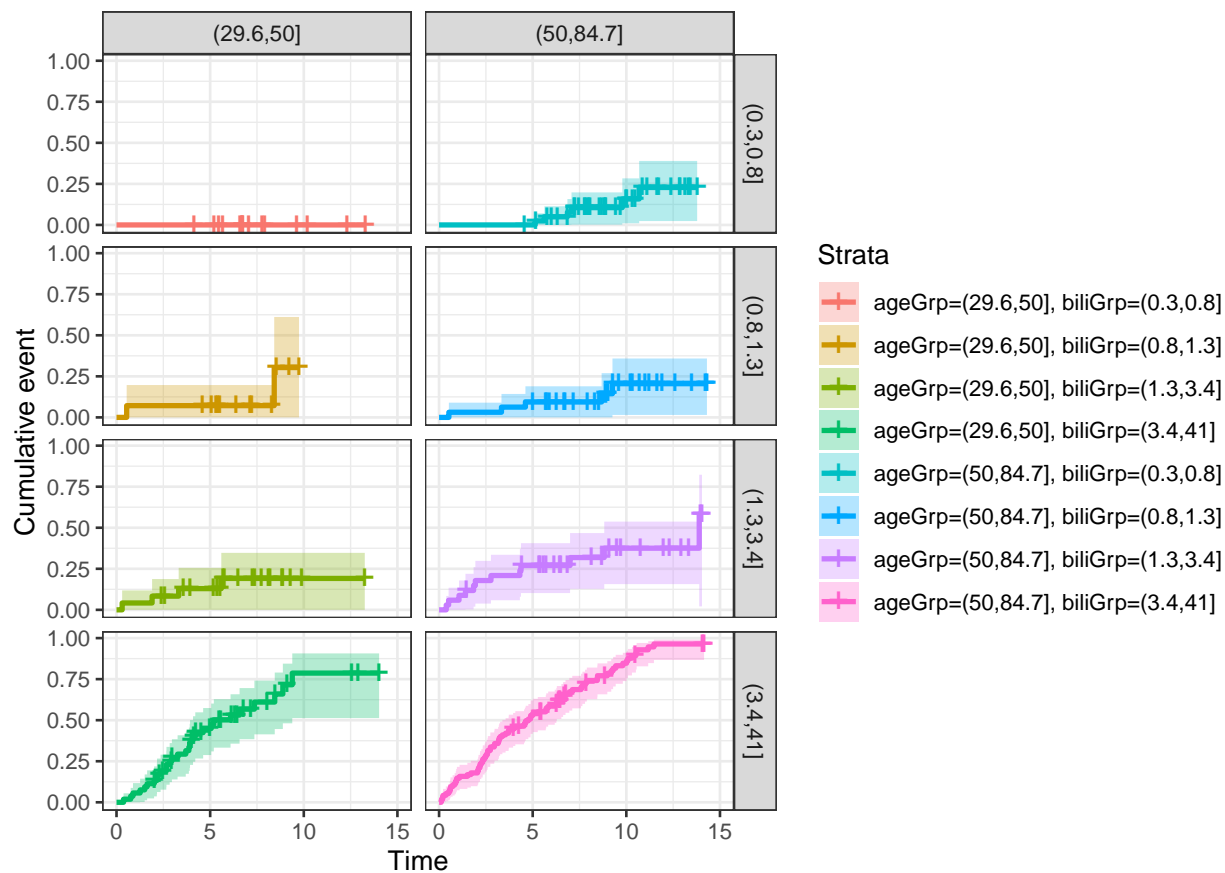


Figure 15: Cumulative event by age and bili (first occurrence)

```
##
## Prediction error curves
##
## Prediction models:
##
## Reference      coxph      rfsrc
## Reference      coxph      rfsrc
##
## Right-censored response of a survival model
##
## No.Observations: 185
##
## Pattern:
##           Freq
## event      52
## right.censored 133
##
## IPCW: cox model
##
## Method for estimating the prediction error:
##
## Bootstrap cross-validation
##
```

```

## Type: resampling
## Bootstrap sample size: 185
## No. bootstrap samples: 50
## Sample size: 185
##
## Cumulative prediction error, aka Integrated Brier score (IBS)
## aka Cumulative rank probability score
##
## Range of integration: 0 and time=4714 :
##
##
## Integrated Brier score (crps):
##
##          IBS[0;time=4714)
## Reference          0.415
## coxph              0.276
## rfsrc              0.195

##      logrank bs.gradient logrankscore
## mean  0.231      0.233      0.242
## std   0.003      0.003      0.005

## Call:
## lmestSearch(responsesFormula = lbili + lalbumin + lalk.phos +
##      lchol + lsgot + lplatelet + lprotime ~ NULL, data = pbcseq11,
##      index = c("id", "times"), k = 1:8, version = "continuous",
##      nrep = 20, out_se = TRUE, seed = 1234)
##
##      states      lk  np      AIC      BIC
##      1 -1772.3906  35 3614.781 3720.847
##      2 -1402.8766  53 2911.753 3072.366
##      3 -1189.3055  81 2540.611 2786.076
##      4  -979.3459 119 2196.692 2557.314
##      5  -777.7574 167 1889.515 2395.598
##      6  -613.3923 225 1676.785 2358.633
##      7  -489.9700 293 1565.940 2453.858
##      8  -389.6264 371 1521.253 2645.545

```

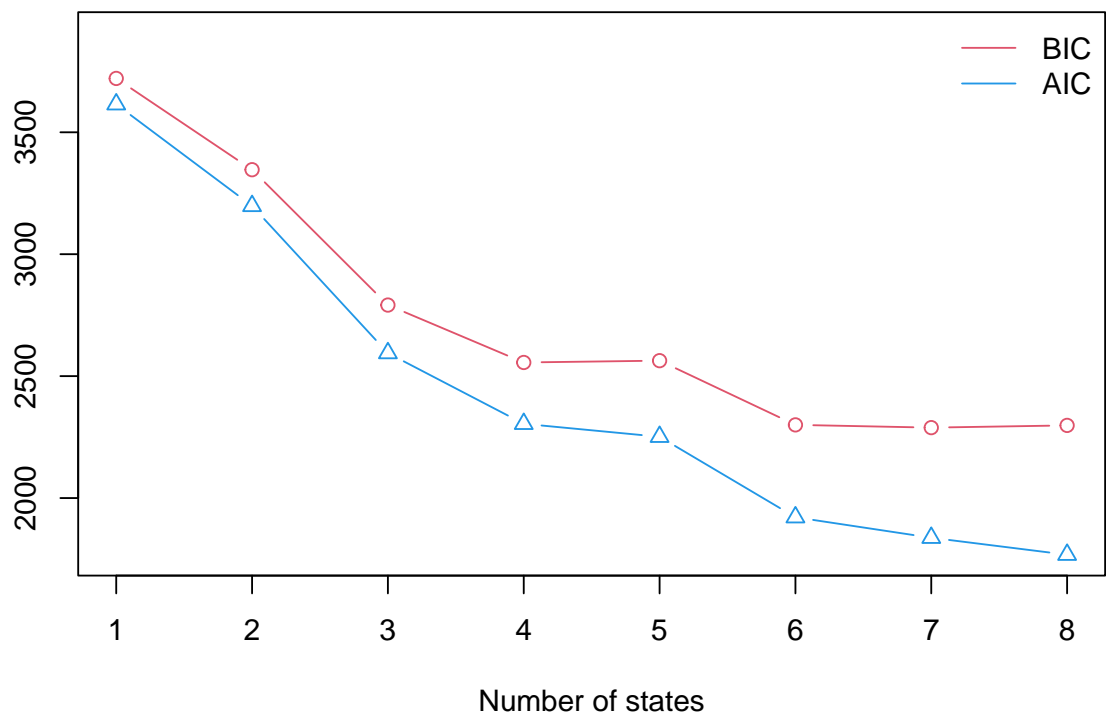


Figure 16: Best model for interpretability and BIC Criterion

Averaged transition probabilities

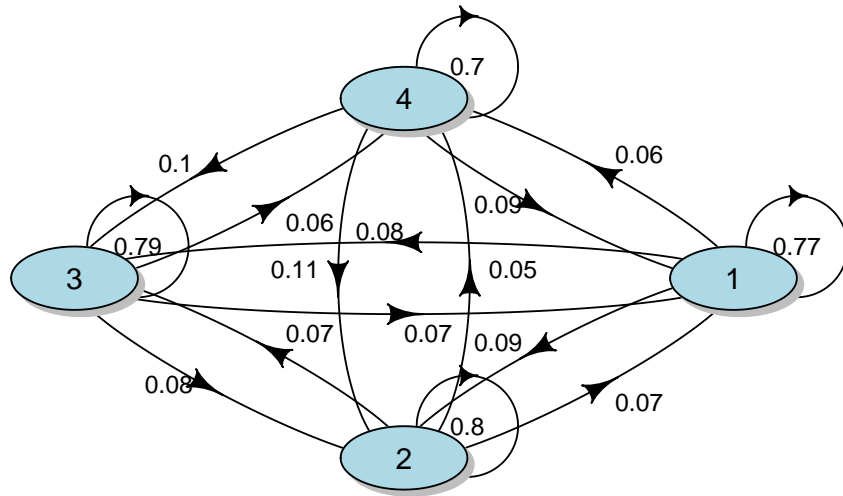


Figure 17: Transitions Plot