Final Model (for publication)

randy

2024-04-11

Contents

0.1 Goal of the analysis

- ☒ To separate the `site` and `ethnicity` as dummy variables
- ☒ Fit the final joint model (without rate) with `JMbayes2` based on model selection results
- ☒ Dynamic prediction for the joint model with plots
- ☒ Calculate the AUC and ROC for the joint model for $Dt = 1, 2, 3, 4, 5$ (year)

0.2 Fit the joint model

0.2.1 Longitudinal model

$lme1: \quad (fixed = lamh \sim time,\ random = \sim time|id)$

Summary for the longitudinal model:

effect	group	term	estimate	std.error	df	statistic	p.value
fixed	NA	(Intercept)	7.5166321	0.1042432	2725	72.10669	0
fixed	NA	time	-0.5304830	0.0116235	2725	-45.63885	0
ran_pars	id	sd_(Intercept)	1.6842278	NA	NA	NA	NA
ran_pars	id	cor_time.(Intercept)	-0.3595600	NA	NA	NA	NA
ran_pars	id	sd_time	0.1039758	NA	NA	NA	NA
ran_pars	Residual	sd_Observation	1.2991322	NA	NA	NA	NA

Here is the latex code for the longitudinal model:

```
% latex table generated in R 4.2.2 by xtable 1.8-4 package
% Wed Apr 10 20:46:26 2024
\begin{table}[ht]
\centering
\begin{tabular}{rlllrrrrr}
  \hline
 & effect & group & term & estimate & std.error & df & statistic & p.value \\
  \hline
1 & fixed &  & (Intercept) & 7.52 & 0.10 & 2725.00 & 72.11 & 0.00 \\
  2 & fixed &  & time & -0.53 & 0.01 & 2725.00 & -45.64 & 0.00 \\
  3 & ran\_pars & id & sd\_(Intercept) & 1.68 &  &  &  &  \\
  4 & ran\_pars & id & cor\_time.(Intercept) & -0.36 &  &  &  &  \\
  5 & ran\_pars & id & sd\_time & 0.10 &  &  &  &  \\
  6 & ran\_pars & Residual & sd\_Observation & 1.30 &  &  &  &  \\
   \hline
\end{tabular}
\caption{Longitudinal model}
\end{table}
```

0.2.2 Survival model

$surv16: \quad coxph(Surv(etime, event) \sim married + bc\_pills + bmi + factor(site) + factor(ethnic) + factor(smoke))$

Summary for the survival model:

term	estimate	std.error	statistic	p.value
marriedMarried/Partnered	0.2174655	0.0946349	2.2979424	0.0215651
bc_pillsEver use BC pills	-0.2268031	0.0988301	-2.2948798	0.0217400
bmi	-0.0030690	0.0067362	-0.4555921	0.6486833
ethnic_black	-0.1763593	0.1171415	-1.5055240	0.1321894
ethnic_chine	-0.0461925	0.2082844	-0.2217760	0.8244883
ethnic_hispa	-0.2032159	0.4114351	-0.4939197	0.6213629
ethnic_japan	-0.1775519	0.2024076	-0.8771994	0.3803784

term	estimate	std.error	statistic	p.value
site_c	0.1102747	0.1602660	0.6880732	0.4914067
site_m	0.2698247	0.1525156	1.7691616	0.0768669
site_nj	0.1970137	0.3524136	0.5590411	0.5761337
site_p	0.5676597	0.1606871	3.5327014	0.0004113
site_ucd	0.2299409	0.1971701	1.1662055	0.2435314
site_ucla	0.2625519	0.2056597	1.2766325	0.2017320
smoke_current_smoker	0.5935254	0.1361619	4.3589687	0.0000131
smoke_past_only	0.0682581	0.1042874	0.6545193	0.5127773

Here is the latex code for the survival model:

```
% latex table generated in R 4.2.2 by xtable 1.8-4 package
% Wed Apr 10 20:49:12 2024
\begin{table}[ht]
\centering
\begin{tabular}{rlrrrr}
  \hline
 & term & estimate & std.error & statistic & p.value \\
  \hline
1 & marriedMarried/Partnered & 0.22 & 0.09 & 2.30 & 0.02 \\
  2 & bc\_pillsEver use BC pills & -0.23 & 0.10 & -2.29 & 0.02 \\
  3 & bmi & -0.00 & 0.01 & -0.46 & 0.65 \\
  4 & ethnic\_black & -0.18 & 0.12 & -1.51 & 0.13 \\
  5 & ethnic\_chine & -0.05 & 0.21 & -0.22 & 0.82 \\
  6 & ethnic\_hispa & -0.20 & 0.41 & -0.49 & 0.62 \\
  7 & ethnic\_japan & -0.18 & 0.20 & -0.88 & 0.38 \\
  8 & site\_c & 0.11 & 0.16 & 0.69 & 0.49 \\
  9 & site\_m & 0.27 & 0.15 & 1.77 & 0.08 \\
  10 & site\_nj & 0.20 & 0.35 & 0.56 & 0.58 \\
  11 & site\_p & 0.57 & 0.16 & 3.53 & 0.00 \\
  12 & site\_ucd & 0.23 & 0.20 & 1.17 & 0.24 \\
  13 & site\_ucla & 0.26 & 0.21 & 1.28 & 0.20 \\
  14 & smoke\_current\_smoker & 0.59 & 0.14 & 4.36 & 0.00 \\
  15 & smoke\_past\_only & 0.07 & 0.10 & 0.65 & 0.51 \\
   \hline
\end{tabular}
\caption{Survival model}
\end{table}
```

0.2.3 Joint model

$$
\begin{aligned}
&jmbayes2.116: \\
&\quad lme1: \quad \boldsymbol{m}(time) = (fixed = lamh \sim time,\ random = \sim time|id) \\
&\quad surv16: \quad Surv(etime, event) \sim married + bc\_pills + bmi + factor(site) + \\
&\qquad\qquad factor(ethnic) + factor(smoke) + \boldsymbol{m}(time)
\end{aligned}
$$

For the `lme` sub-model:

	Mean	StDev	2.5%	97.5%	P	Rhat
(Intercept)	7.9135455	0.1085692	7.7084534	8.1336243	0	1.012414
time	-0.5953003	0.0125513	-0.6211691	-0.5721701	0	1.015866
sigma	1.3077694	0.0189601	1.2711295	1.3436842	0	1.002651

```
% latex table generated in R 4.2.2 by xtable 1.8-4 package
% Wed Apr 10 21:31:42 2024
\begin{table}[ht]
\centering
\begin{tabular}{rrrrrrr}
  \hline
 & Mean & StDev & 2.5\% & 97.5\% & P & Rhat \\
  \hline
(Intercept) & 7.91 & 0.11 & 7.71 & 8.13 & 0.00 & 1.01 \\
  time & -0.60 & 0.01 & -0.62 & -0.57 & 0.00 & 1.02 \\
  sigma & 1.31 & 0.02 & 1.27 & 1.34 & 0.00 & 1.00 \\
   \hline
\end{tabular}
\caption{Longitudinal model}
\end{table}
```

For the `surv` sub-model:

	Mean	StDev	2.5%	97.5%	P	Rhat
marriedMarried/Partnered	0.1266179	0.1490185	-0.1617231	0.4137407	0.3875000	1.006580
bc_pillsEver use BC pills	-0.0263297	0.1322543	-0.2854227	0.2389329	0.8211667	1.004448
bmi	-0.0482663	0.0090790	-0.0666461	-0.0309416	0.0000000	1.016472
ethnic_black	-0.3653539	0.1533250	-0.6638584	-0.0677044	0.0150000	1.000042
ethnic_chine	-0.2258512	0.2759631	-0.7497783	0.3270956	0.4053333	1.014657
ethnic_hispa	-0.6071708	0.6053575	-1.7585837	0.6166765	0.3263333	1.051092
ethnic_japan	-0.0380953	0.2651336	-0.5569021	0.4735115	0.8910000	1.008927
site_c	0.2423351	0.2020243	-0.1562895	0.6222388	0.2498333	1.011504
site_m	0.3299295	0.1904628	-0.0359440	0.7099669	0.0785000	1.017660
site_nj	0.6440294	0.4900872	-0.3486416	1.5668388	0.2110000	1.017173
site_p	0.5092241	0.2002082	0.1304235	0.9174151	0.0095000	1.020442
site_ucd	0.1383832	0.2553424	-0.3751667	0.6395364	0.5673333	1.017673
site_ucla	0.4199383	0.2649476	-0.0869673	0.9488132	0.1006667	1.017816
smoke_current_smoker	0.1722077	0.1751894	-0.1870096	0.5279286	0.3048333	1.005363
smoke_past_only	-0.0686700	0.1346129	-0.3293920	0.1994389	0.6005000	1.004055
value(lamh)	-1.0086003	0.0684292	-1.1549865	-0.8799643	0.0000000	1.059476

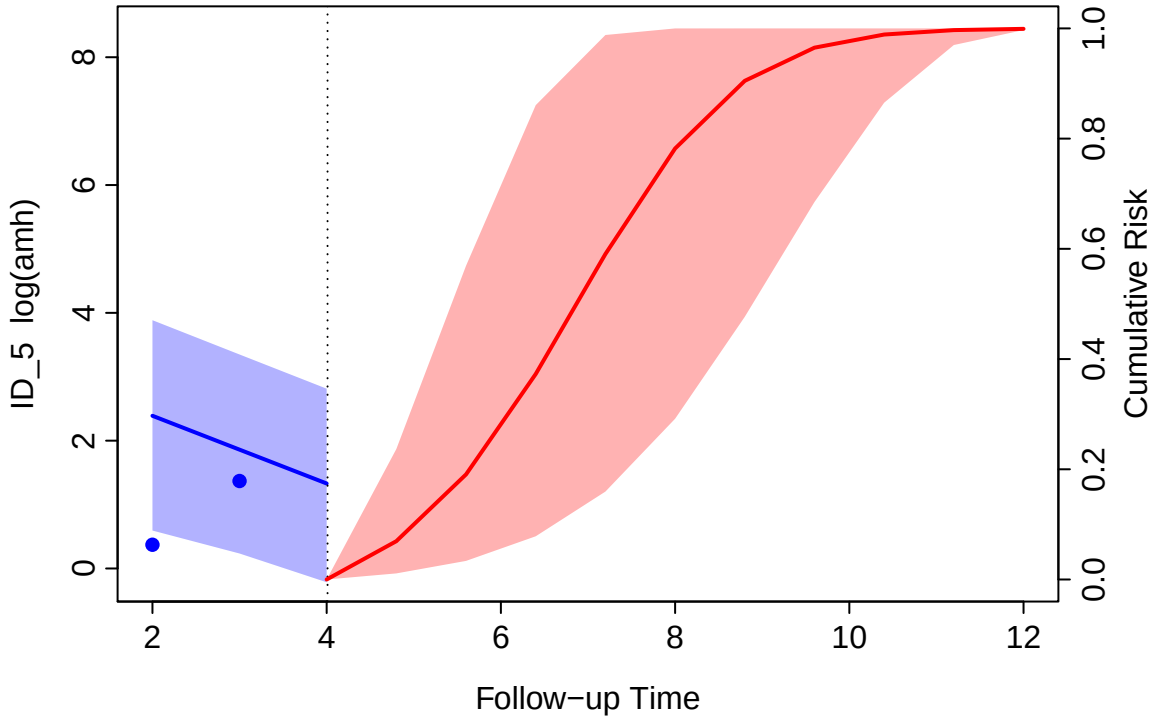
```
% latex table generated in R 4.2.2 by xtable 1.8-4 package
% Wed Apr 10 21:30:38 2024
\begin{table}[ht]
\centering
\begin{tabular}{rrrrrrr}
  \hline
 & Mean & StDev & 2.5\% & 97.5\% & P & Rhat \\
  \hline
marriedMarried/Partnered & 0.13 & 0.15 & -0.16 & 0.41 & 0.39 & 1.01 \\
```

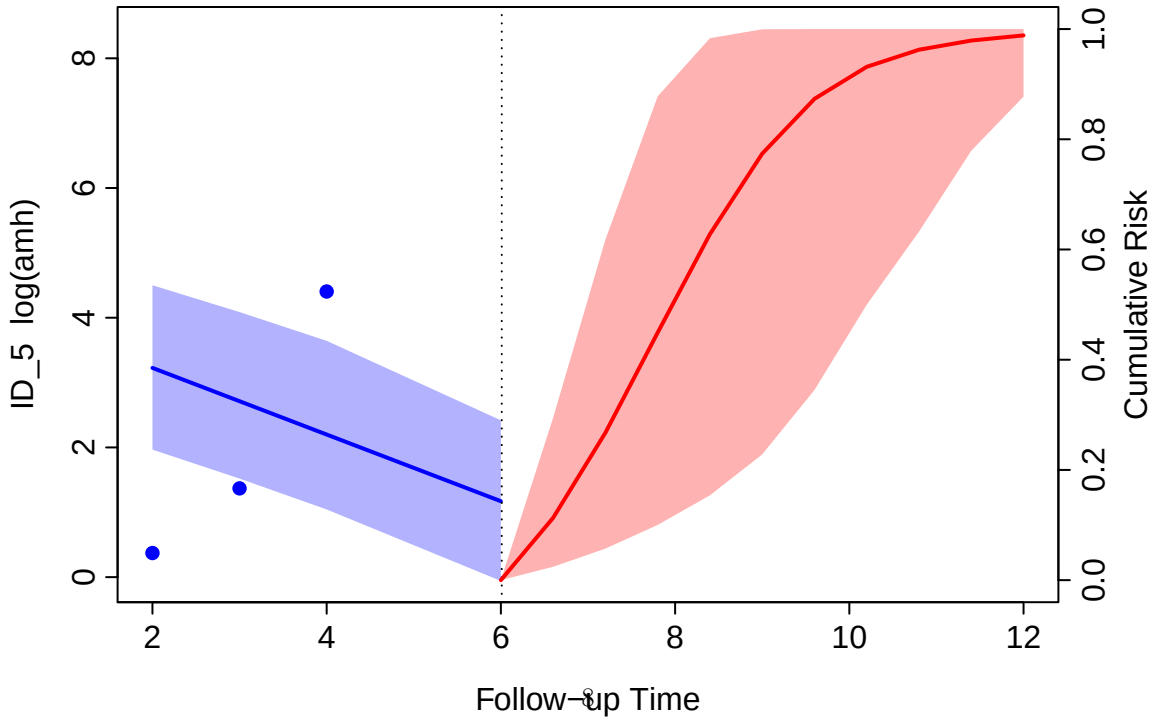
```
  bc\_pillsEver use BC pills & -0.03 & 0.13 & -0.29 & 0.24 & 0.82 & 1.00 \\
  bmi & -0.05 & 0.01 & -0.07 & -0.03 & 0.00 & 1.02 \\
  ethnic\_black & -0.37 & 0.15 & -0.66 & -0.07 & 0.01 & 1.00 \\
  ethnic\_chine & -0.23 & 0.28 & -0.75 & 0.33 & 0.41 & 1.01 \\
  ethnic\_hispa & -0.61 & 0.61 & -1.76 & 0.62 & 0.33 & 1.05 \\
  ethnic\_japan & -0.04 & 0.27 & -0.56 & 0.47 & 0.89 & 1.01 \\
  site\_c & 0.24 & 0.20 & -0.16 & 0.62 & 0.25 & 1.01 \\
  site\_m & 0.33 & 0.19 & -0.04 & 0.71 & 0.08 & 1.02 \\
  site\_nj & 0.64 & 0.49 & -0.35 & 1.57 & 0.21 & 1.02 \\
  site\_p & 0.51 & 0.20 & 0.13 & 0.92 & 0.01 & 1.02 \\
  site\_ucd & 0.14 & 0.26 & -0.38 & 0.64 & 0.57 & 1.02 \\
  site\_ucla & 0.42 & 0.26 & -0.09 & 0.95 & 0.10 & 1.02 \\
  smoke\_current\_smoker & 0.17 & 0.18 & -0.19 & 0.53 & 0.30 & 1.01 \\
  smoke\_past\_only & -0.07 & 0.13 & -0.33 & 0.20 & 0.60 & 1.00 \\
  value(lamh) & -1.01 & 0.07 & -1.15 & -0.88 & 0.00 & 1.06 \\
   \hline
\end{tabular}
\caption{Survival model}
\end{table}
```

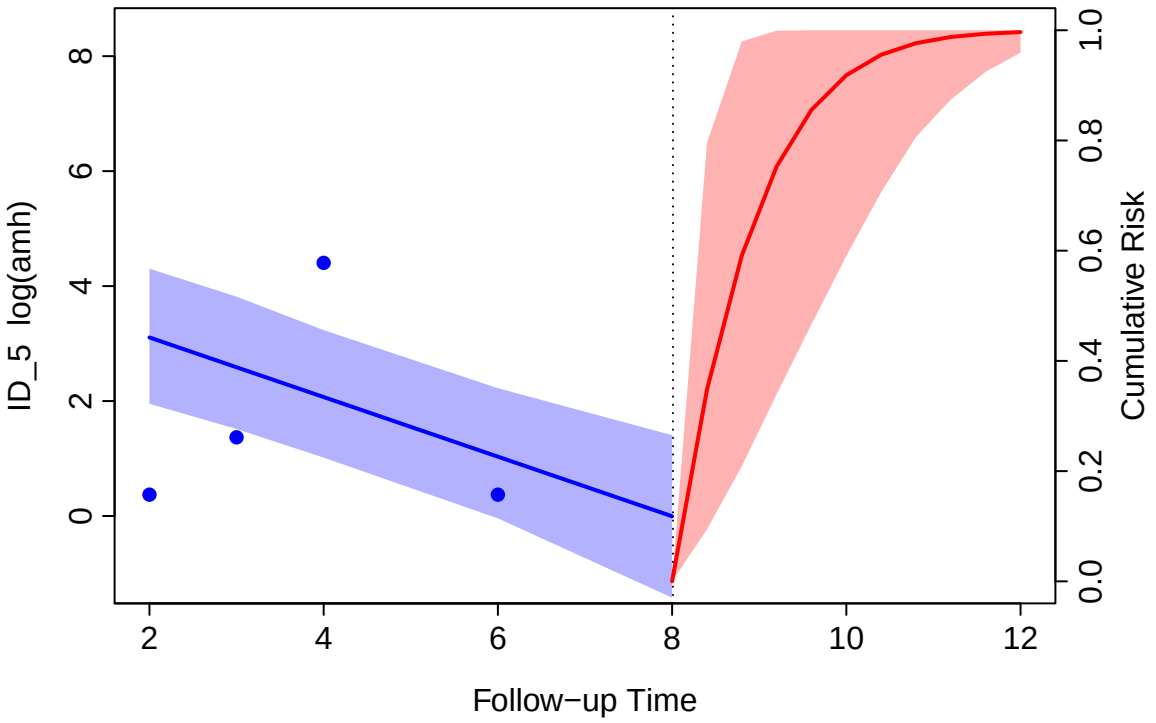
Notice: ask EJC for which computer to use for long chain running

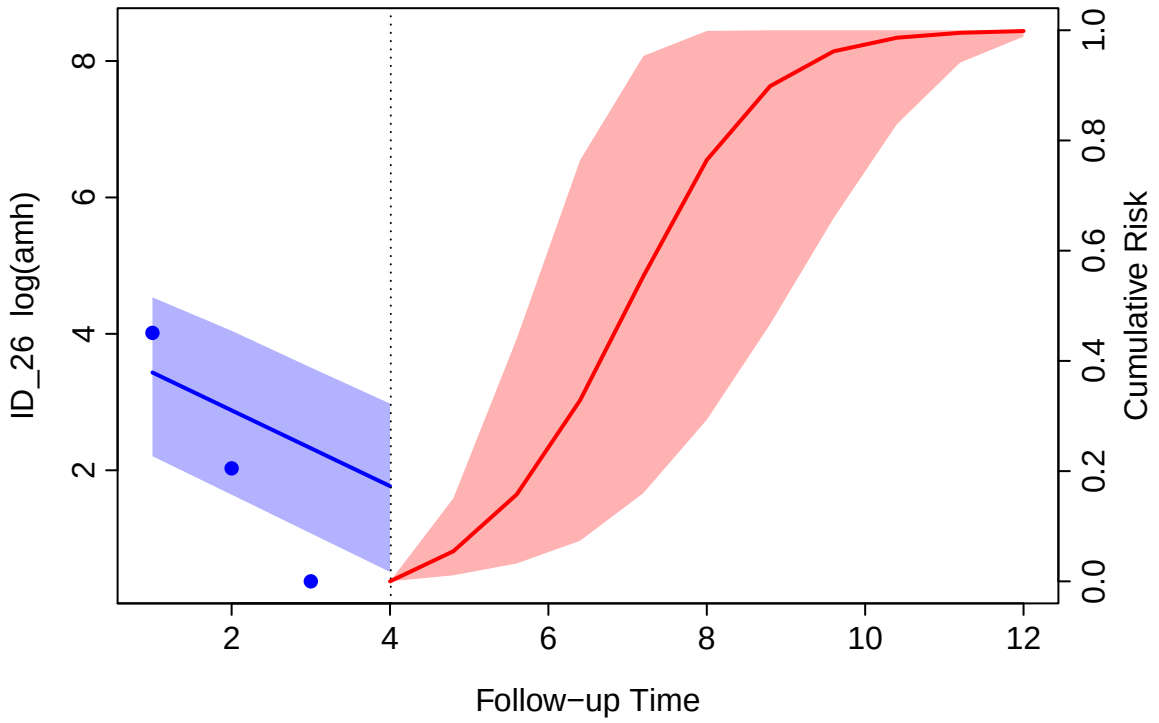
0.3 Dynamic prediction

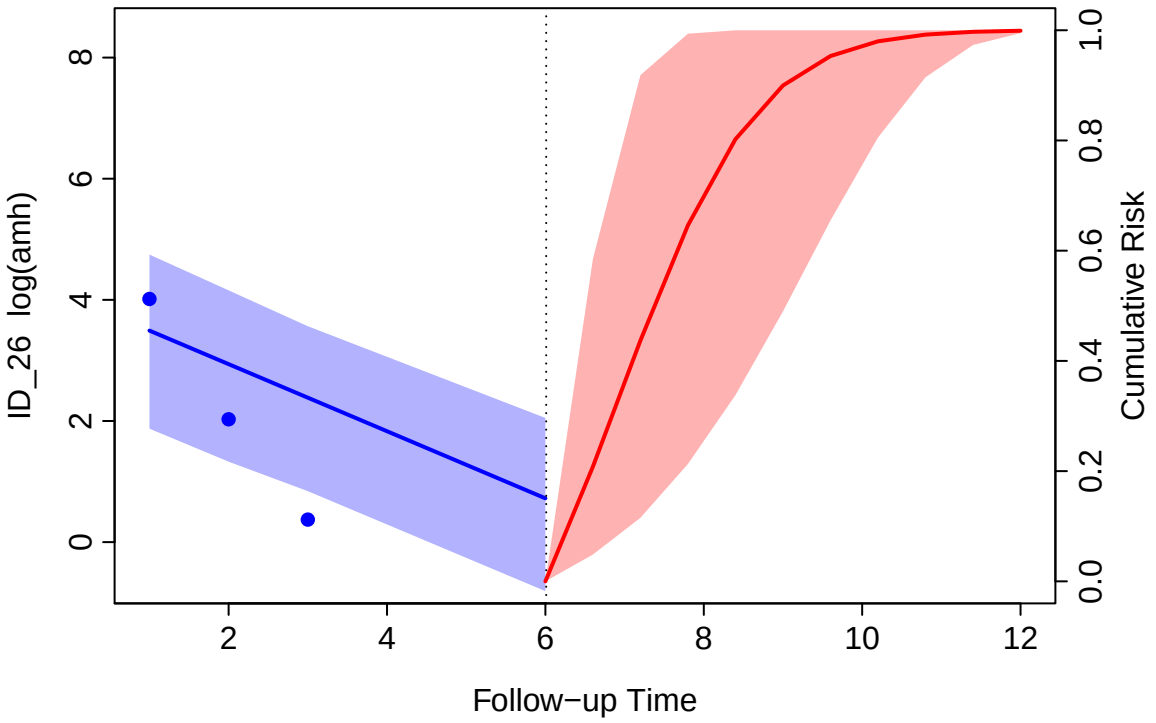
0.3.1 Individual prediction plots

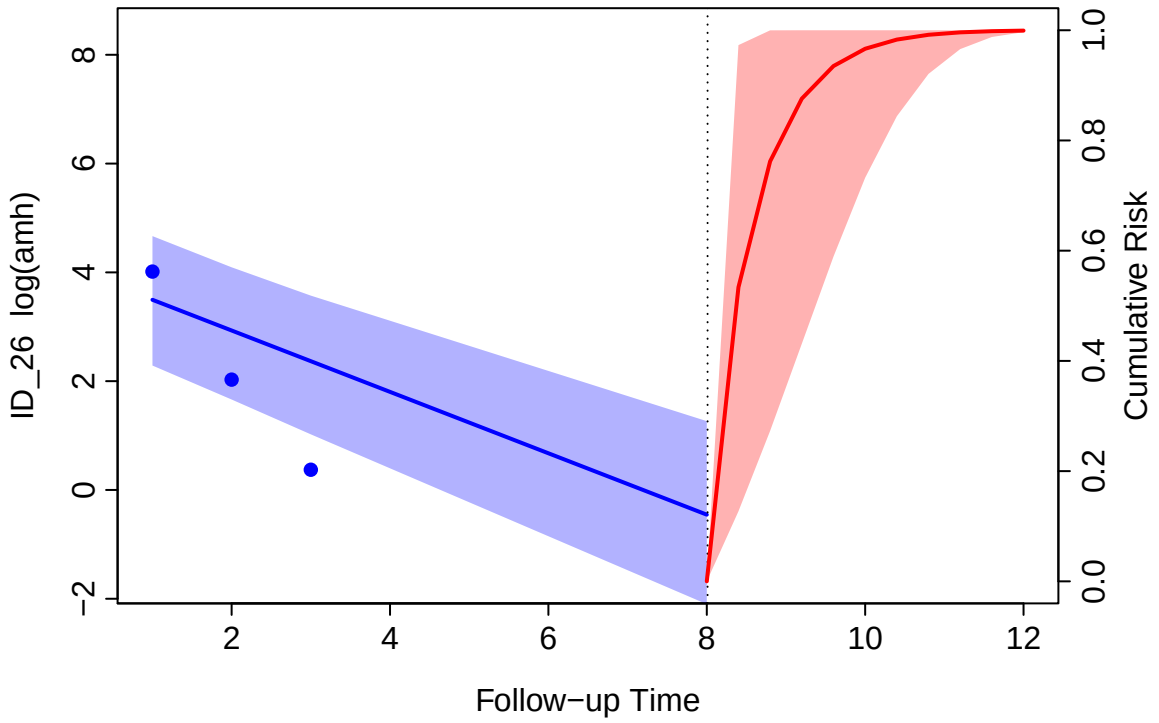


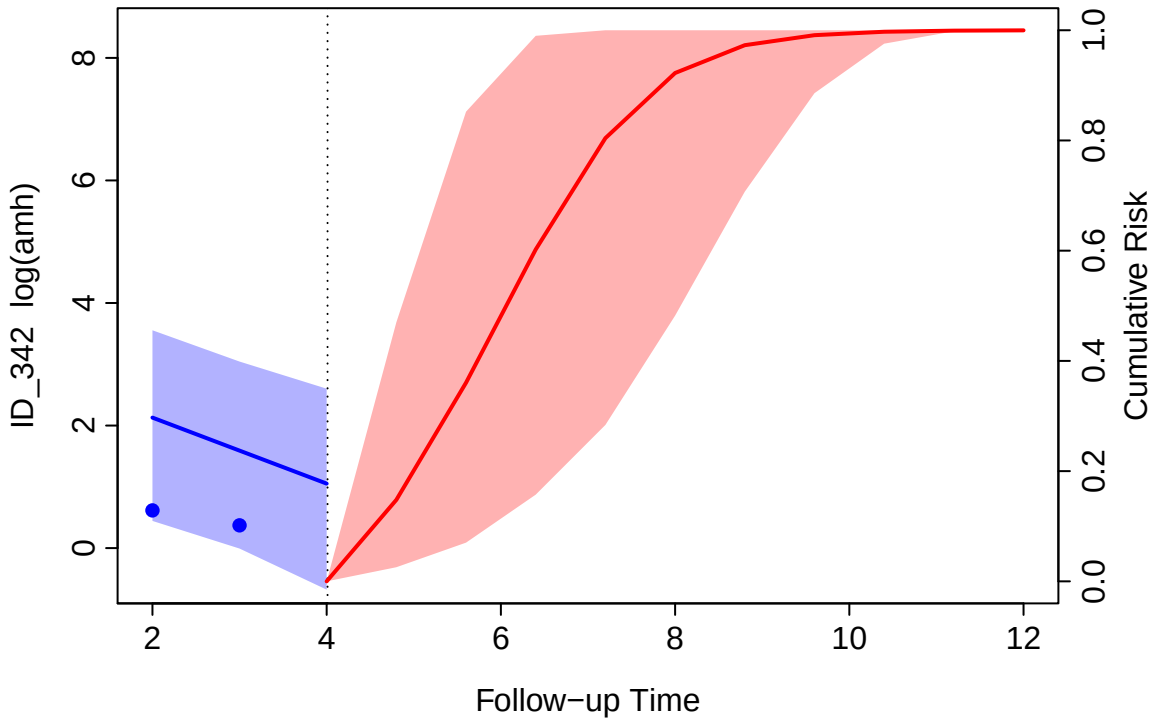


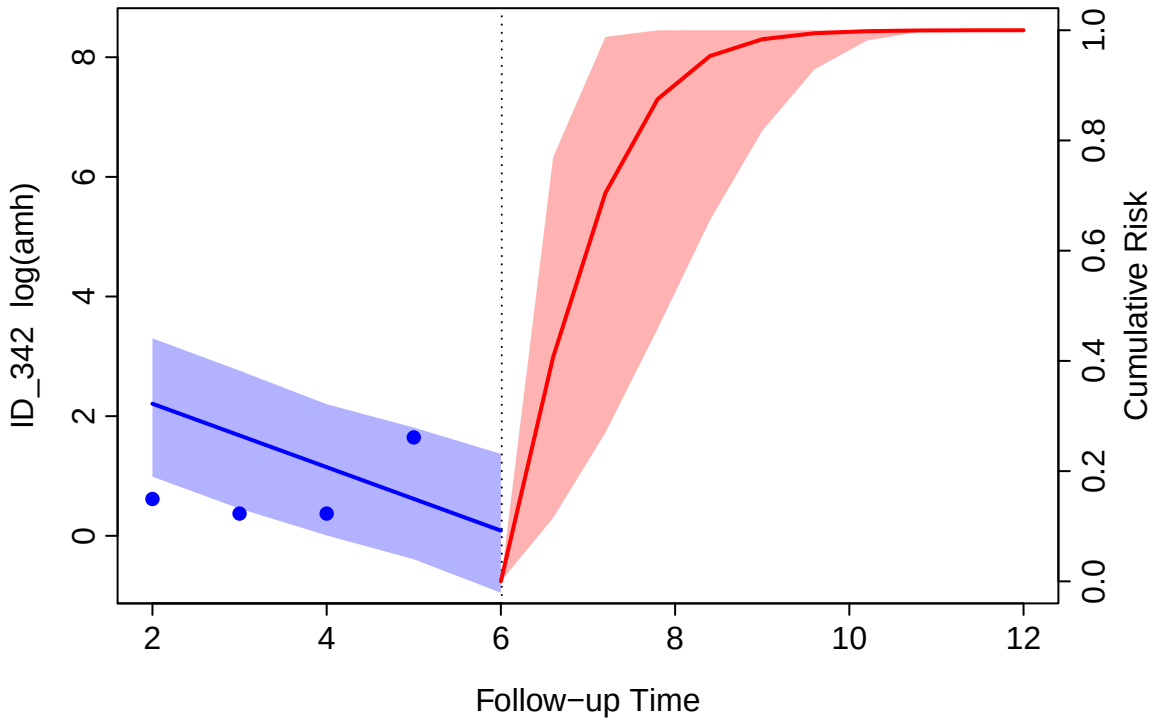

ID_5 log(amh)
0
2
4
6
8
2
4
6
8
10
12
Follow−up Time
0.0
0.2
0.4
0.6
0.8
1.0
Cumulative Risk

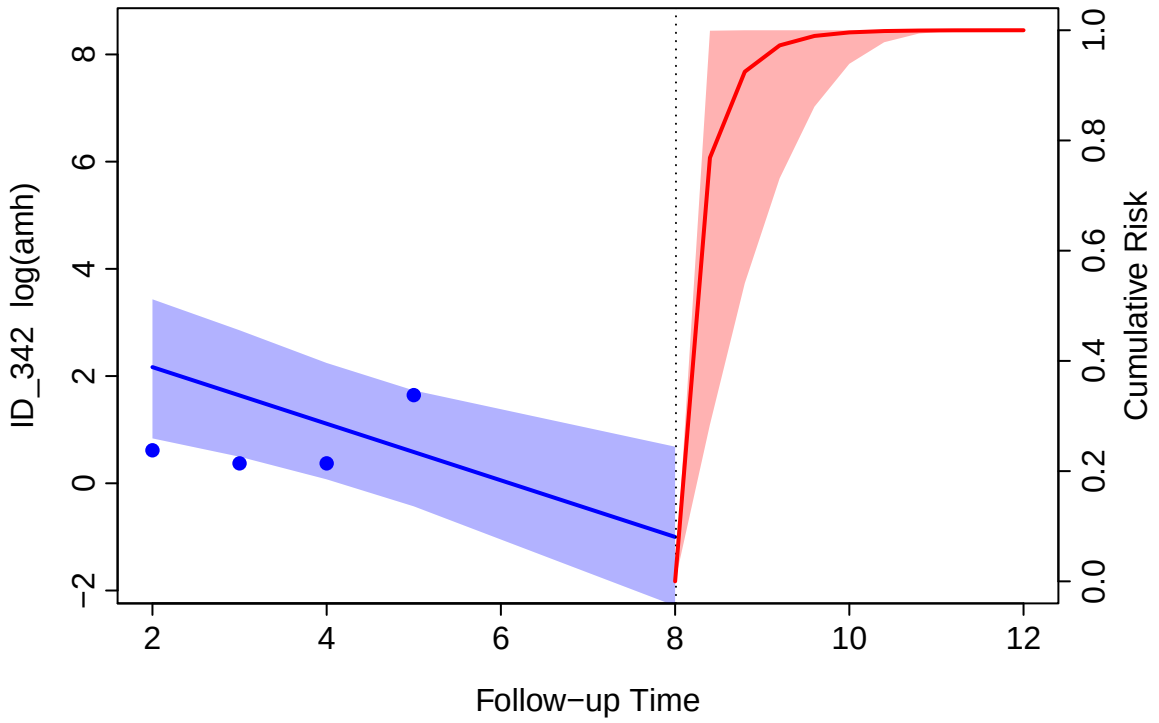

ID_26 log(amh)
2
4
6
8
2
4
6
8
10
12
Follow−up Time
Cumulative Risk
0.0
0.2
0.4
0.6
0.8
1.0

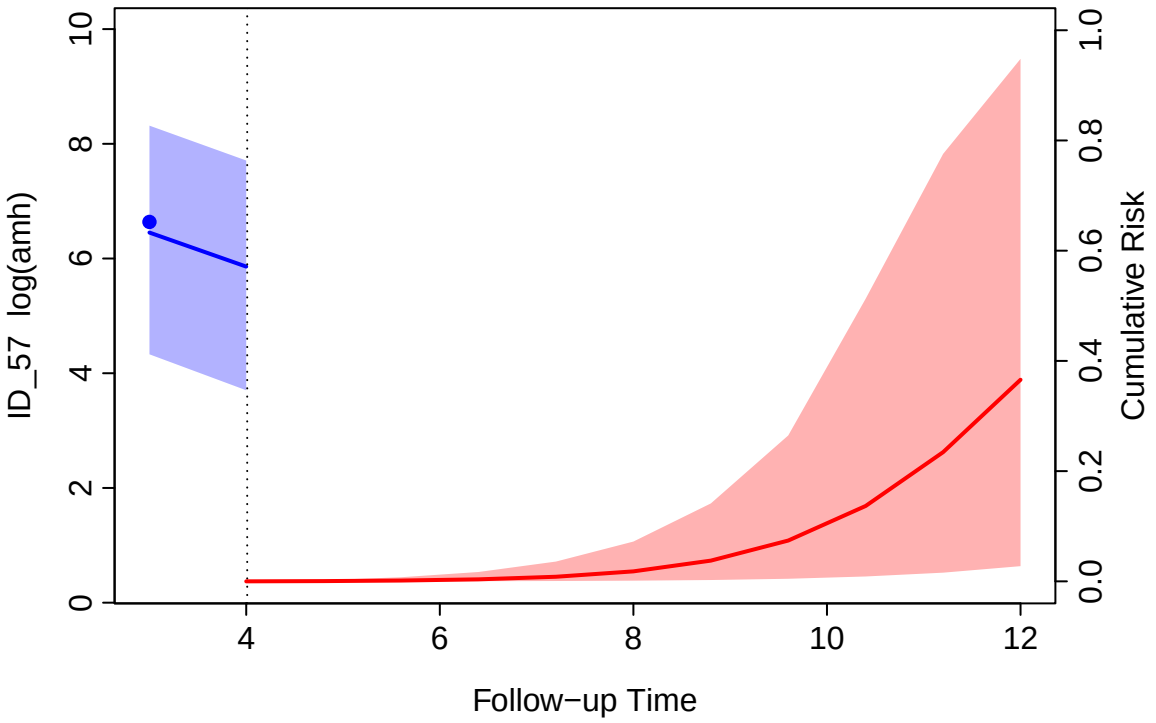

ID_26 log(amh)
0
2
4
6
8
2
4
6
8
10
12
Follow−up Time
0.0
0.2
0.4
0.6
0.8
1.0
Cumulative Risk

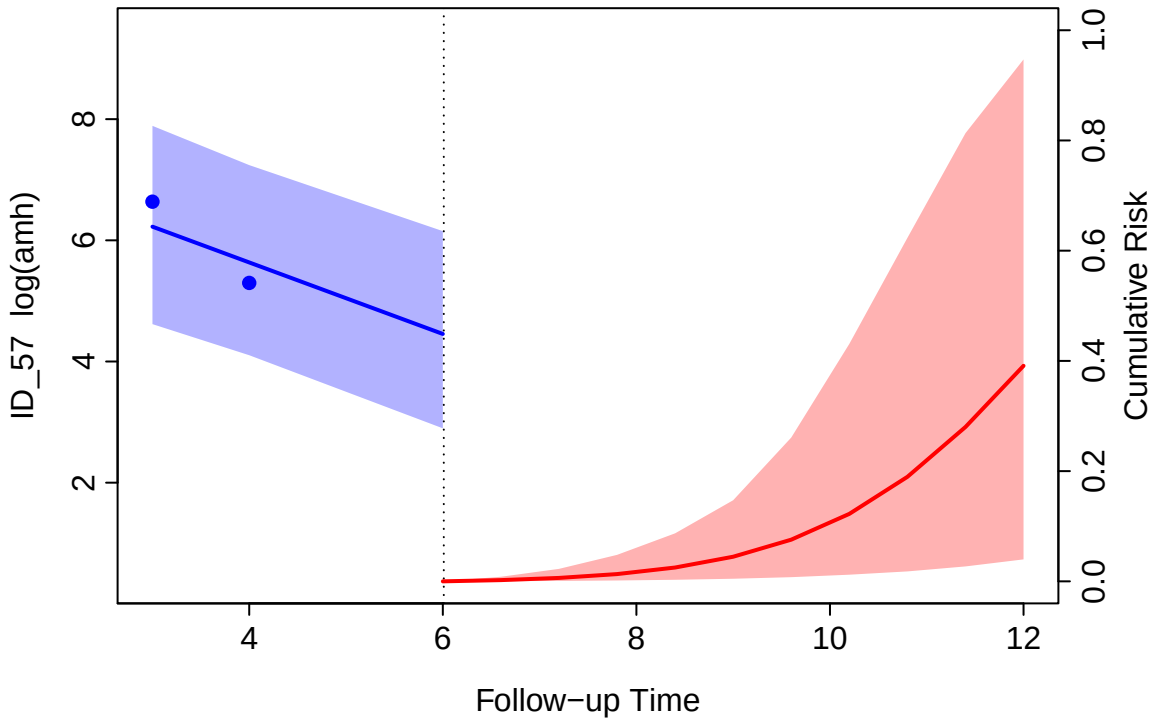

ID_26 log(amh)
Cumulative Risk
Follow−up Time
−2
0
2
4
6
8
0.0
0.2
0.4
0.6
0.8
1.0
2
4
6
8
10
12

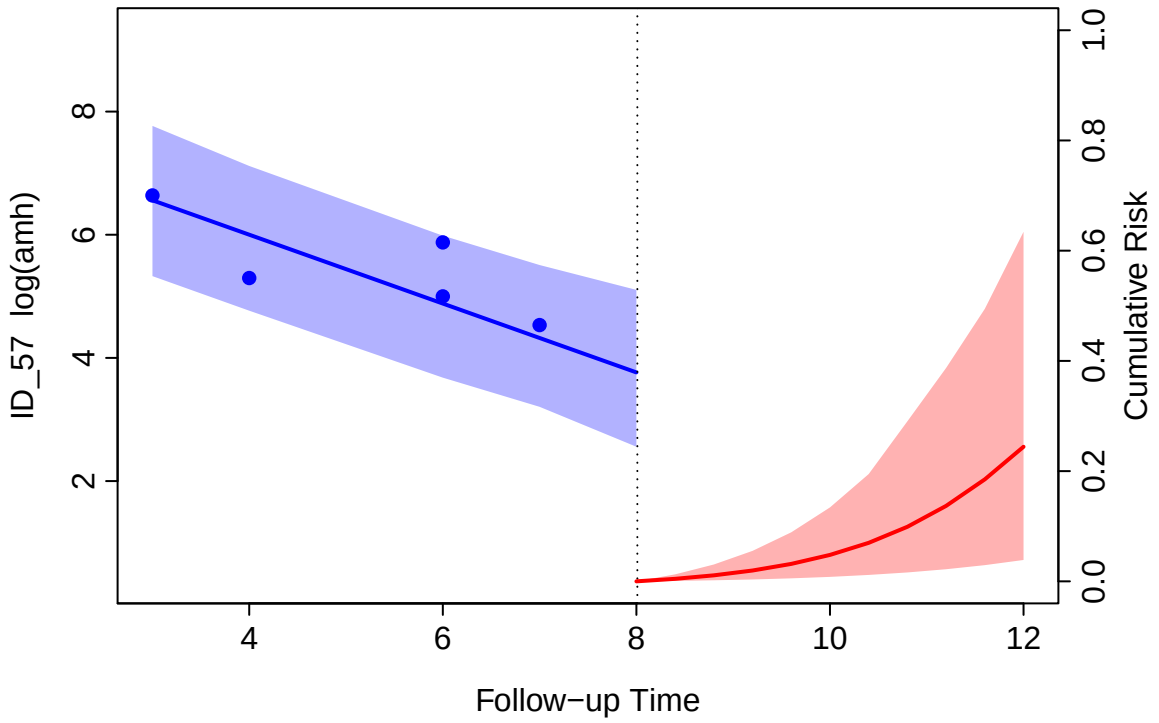

ID_342 log(amh)
0
2
4
6
8
2
4
6
8
10
12
Follow−up Time
0.0
0.2
0.4
0.6
0.8
1.0
Cumulative Risk

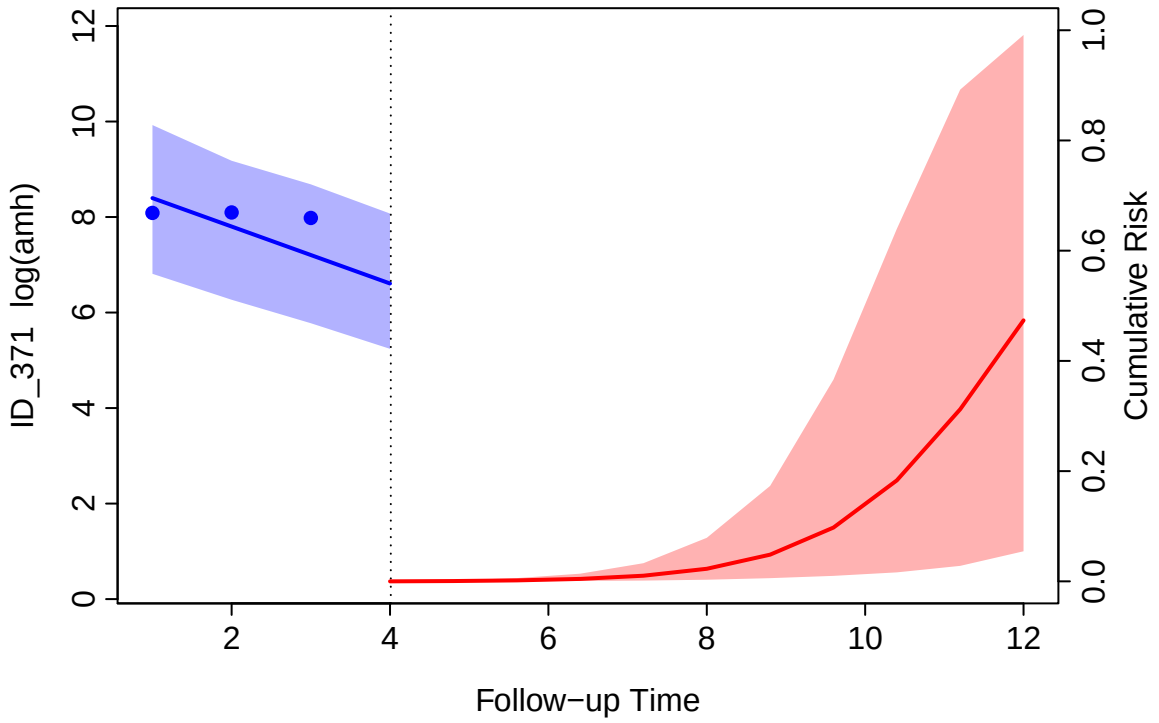

ID_342 log(amh)
0
2
4
6
8
2
4
6
8
10
12
Follow−up Time
Cumulative Risk
0.0
0.2
0.4
0.6
0.8
1.0

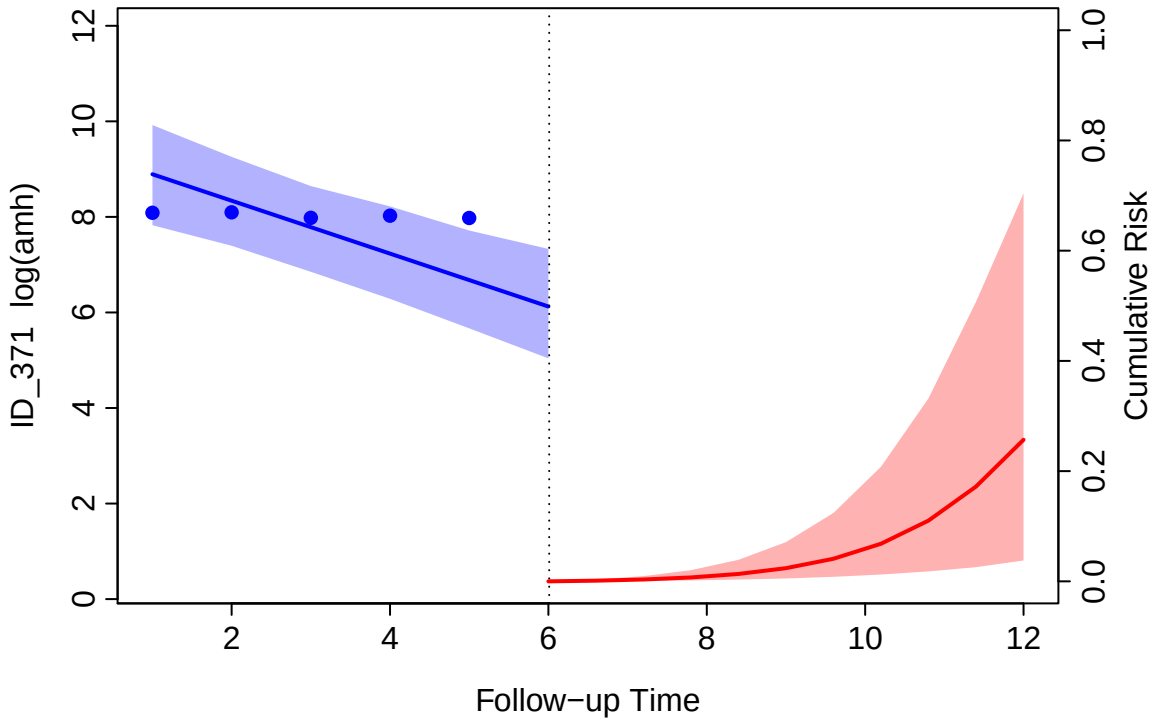

ID_342 log(amh)
Cumulative Risk
Follow−up Time
−2
0
2
4
6
8
0.0
0.2
0.4
0.6
0.8
1.0
2
4
6
8
10
12

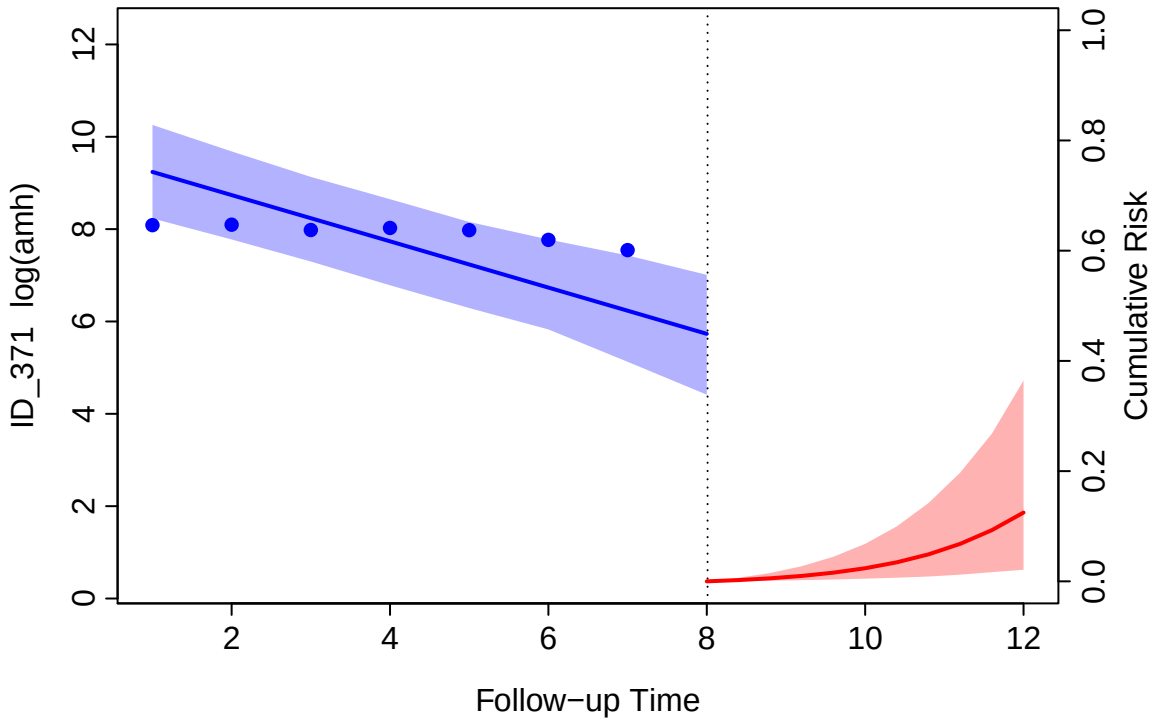

ID_57 log(amh)
0
2
4
6
8
10
4
6
8
10
12
Follow−up Time
0.0
0.2
0.4
0.6
0.8
1.0
Cumulative Risk

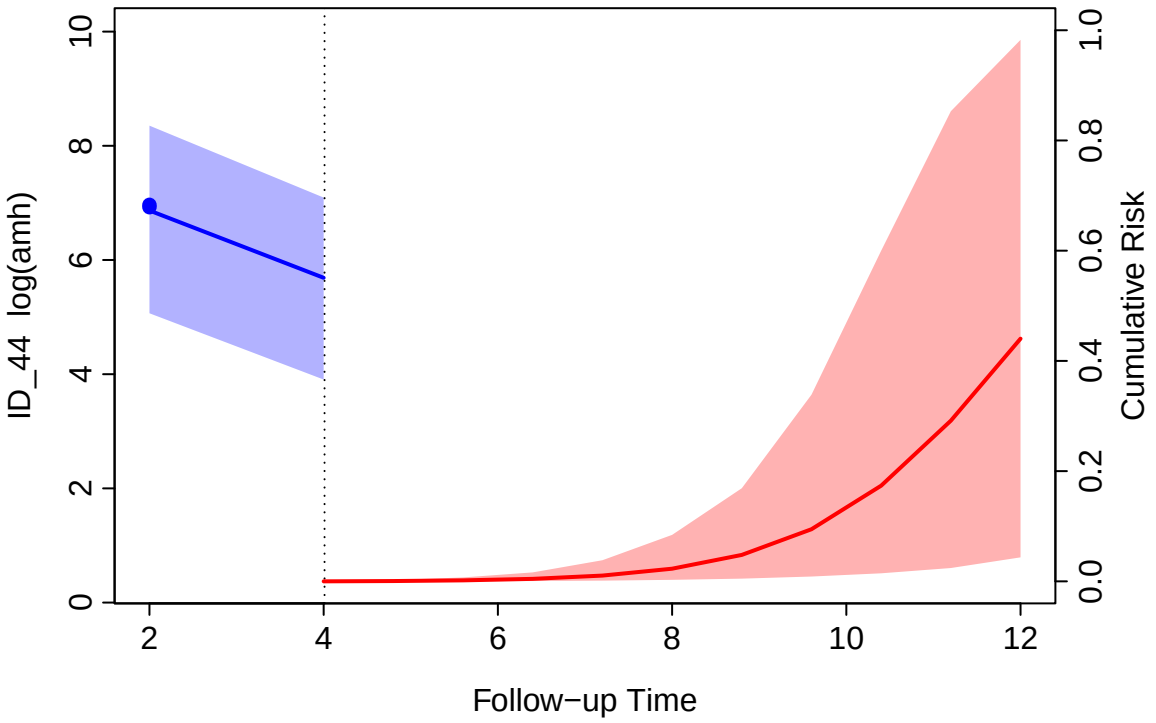

ID_57 log(amh)
2
4
6
8
4
6
8
10
12
Follow−up Time
0.0
0.2
0.4
0.6
0.8
1.0
Cumulative Risk

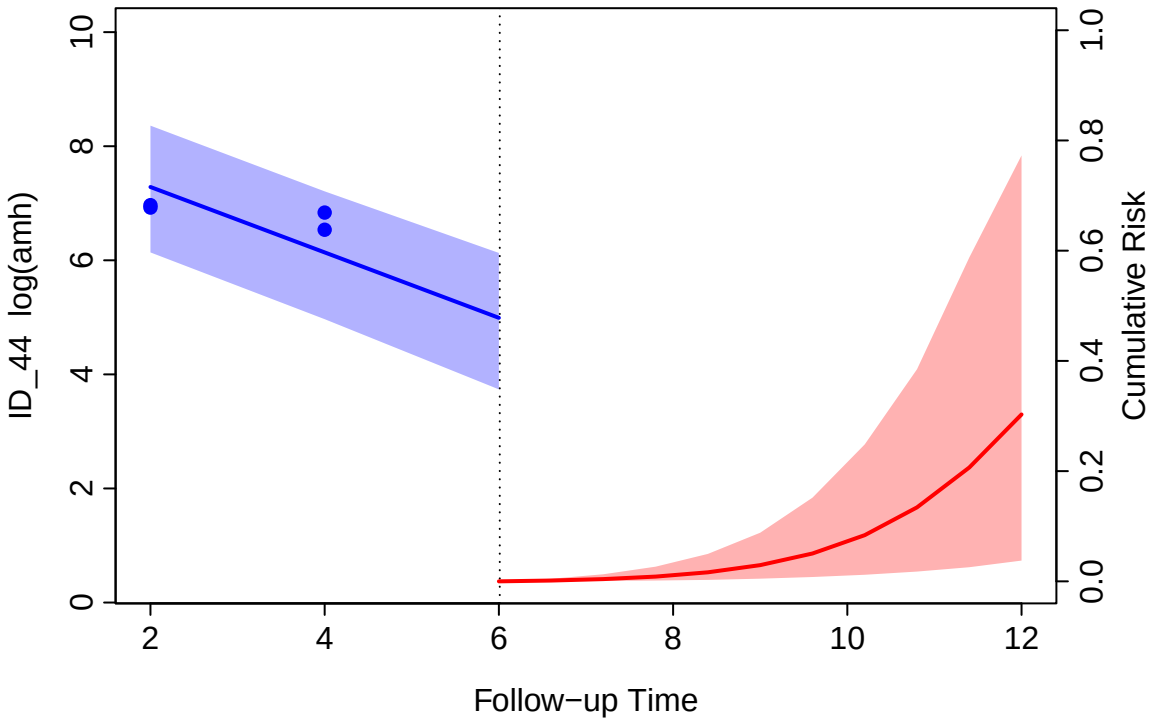

ID_57 log(amh)
2
4
6
8
4
6
8
10
12
Follow−up Time
0.0
0.2
0.4
0.6
0.8
1.0
Cumulative Risk

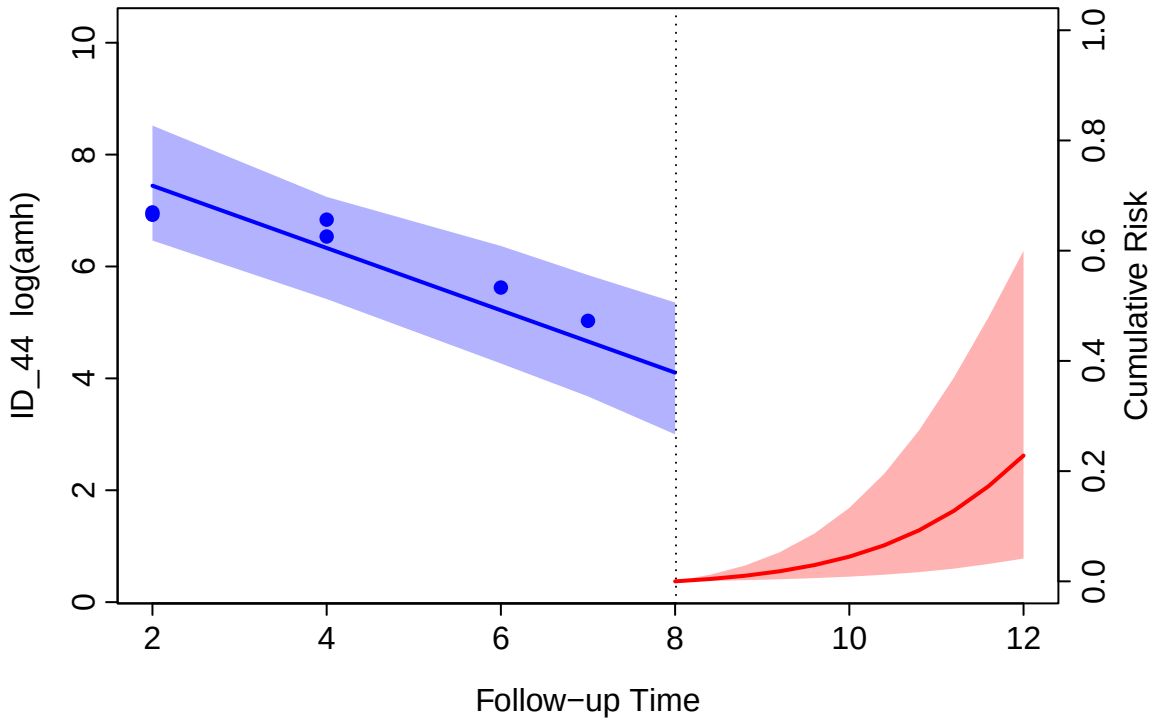

ID_371 log(amh)
Cumulative Risk
Follow−up Time
0
2
4
6
8
10
12
0.0
0.2
0.4
0.6
0.8
1.0


ID_371 log(amh)
Cumulative Risk
Follow−up Time
0
2
4
6
8
10
12
0.0
0.2
0.4
0.6
0.8
1.0


ID_371 log(amh)
Cumulative Risk
Follow−up Time
0
2
4
6
8
10
12
0.0
0.2
0.4
0.6
0.8
1.0


ID_44 log(amh)
Cumulative Risk
Follow−up Time
0
2
4
6
8
10
0.0
0.2
0.4
0.6
0.8
1.0
2
4
6
8
10
12


ID_44 log(amh)
Cumulative Risk
Follow−up Time
0
2
4
6
8
10
0.0
0.2
0.4
0.6
0.8
1.0
2
4
6
8
10
12


ID_44 log(amh)
Cumulative Risk
Follow−up Time
0
2
4
6
8
10
0.0
0.2
0.4
0.6
0.8
1.0
2
4
6
8
10
12

0.4 AUC and ROC

For `JMbayes2`, the function `tvAUC()` is used to estimate the time-dependent AUC only for all the subjects (there is no function for individual AUC). Here is an error message for the AUC $Tstart = 2$ $Dt = 1$

```
Error in tvROC.jm(object, newdata, Tstart, Thoriz, Dt, ...) :
  it seems that there are no events in the interval [Tstart, Thoriz).
```

```
#> Error in tvROC.jm(object, newdata, Tstart, Thoriz, Dt, ...) :
#>   it seems that there are no events in the interval [Tstart, Thoriz).
```

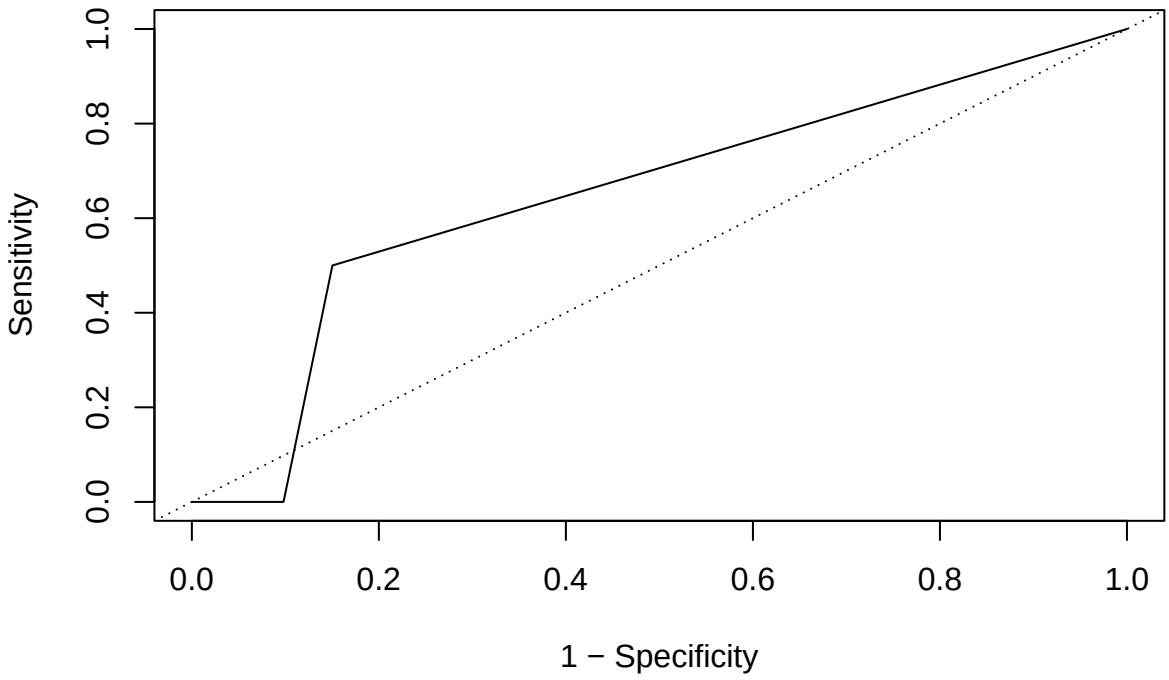
	Dt1	Dt2	Dt3	Dt4	Dt5
Tstart2	NA	0.6503	0.8988	0.8214	0.8471
Tstart3	0.6741	0.9011	0.8650	0.8105	0.8115
Tstart4	0.9590	0.9172	0.8310	0.8070	0.8070
Tstart5	0.8834	0.8459	0.8244	0.8196	0.8302
Tstart6	0.8256	0.8258	0.8140	0.8278	0.8014
Tstart8	0.7984	0.8063	0.7879	0.7810	0.8306

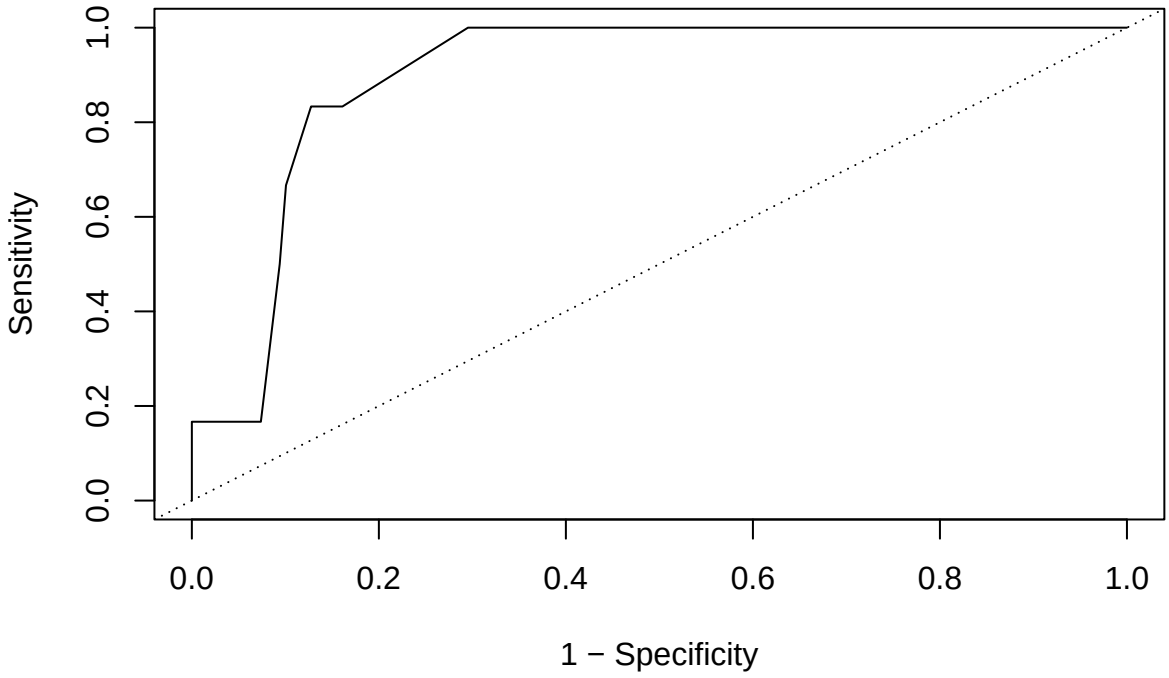
```
#> % latex table generated in R 4.2.2 by xtable 1.8-4 package
#> % Thu Apr 11 09:50:46 2024
#> \begin{table}[ht]
#> \centering
#> \begin{tabular}{rrrrrr}
#>   \hline
#>  & Dt1 & Dt2 & Dt3 & Dt4 & Dt5 \\
#>   \hline
#> Tstart2 &  & 0.65 & 0.90 & 0.82 & 0.85 \\
#>   Tstart4 & 0.96 & 0.92 & 0.83 & 0.81 & 0.81 \\
#>   Tstart6 & 0.83 & 0.83 & 0.81 & 0.83 & 0.80 \\
#>   Tstart8 & 0.80 & 0.81 & 0.79 & 0.78 & 0.83 \\
#>    \hline
#> \end{tabular}
#> \caption{AUC Table}
#> \end{table}
```

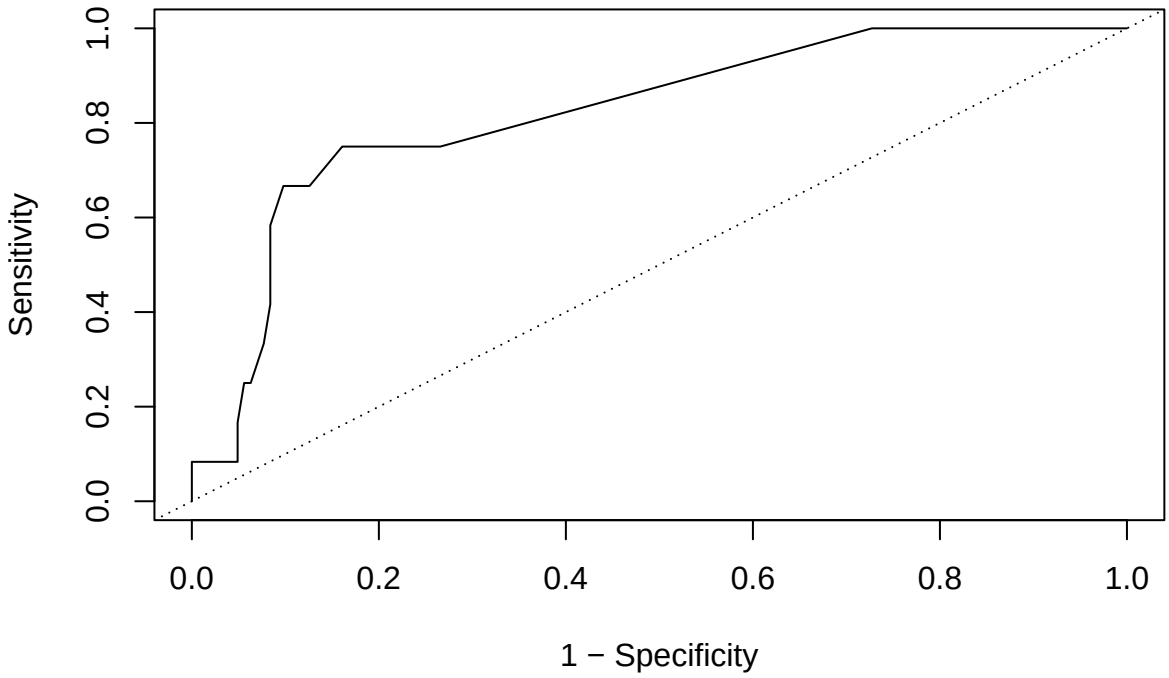
The ROC curves for the joint model for $Tstart = 2, 3, 4, 5, 6, 8$ by rows,

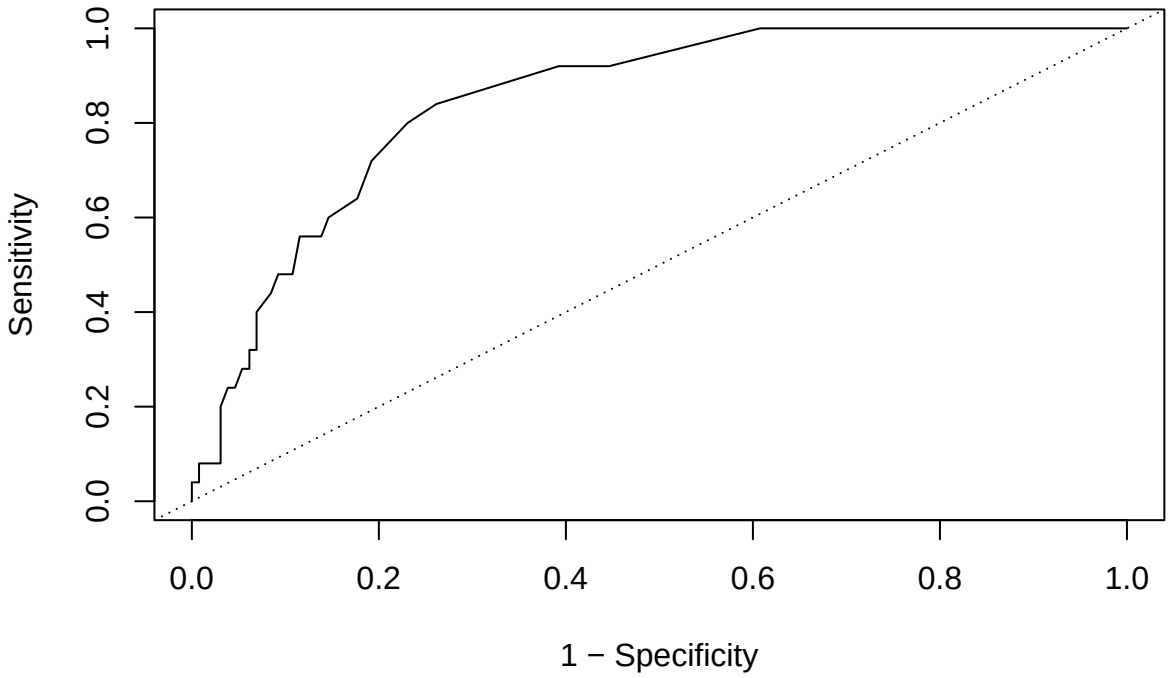
and $Dt = 2, 3, 4, 5$ by columns are shown below:

0.4.1 Start Time 2

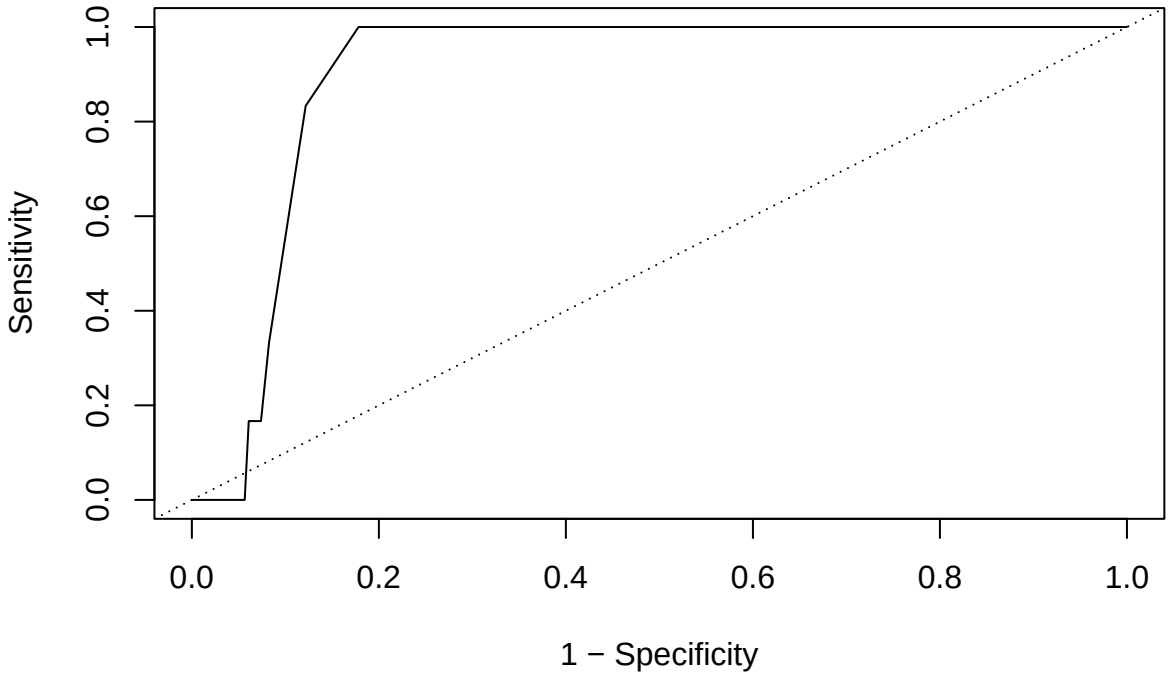


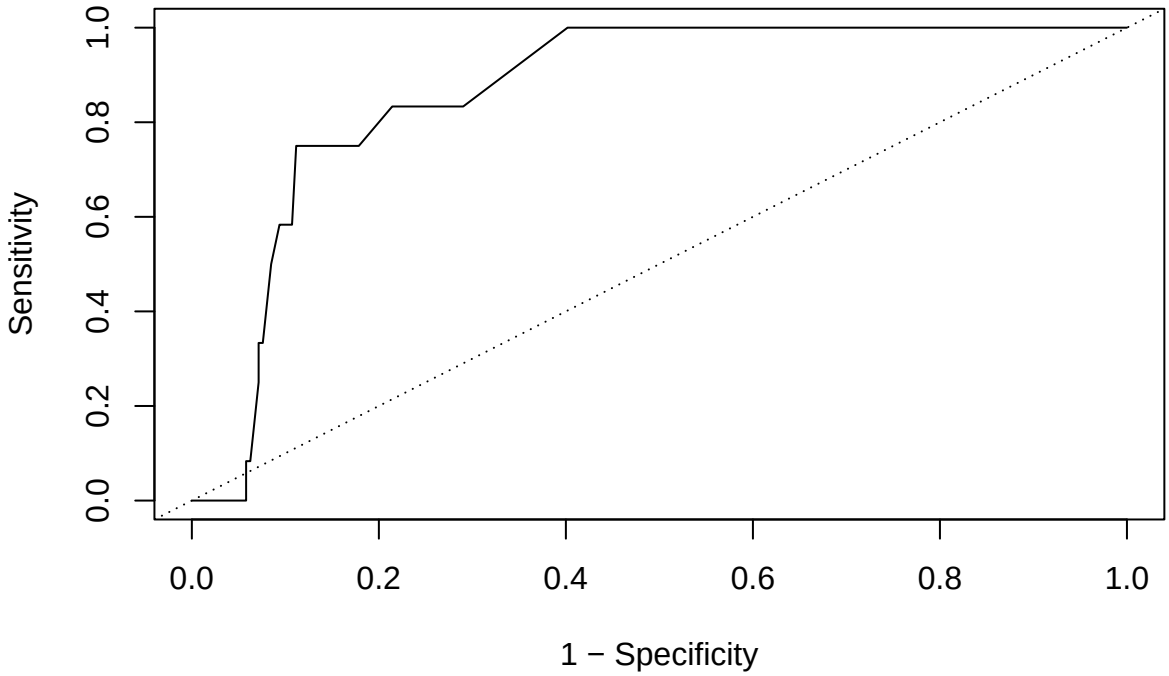


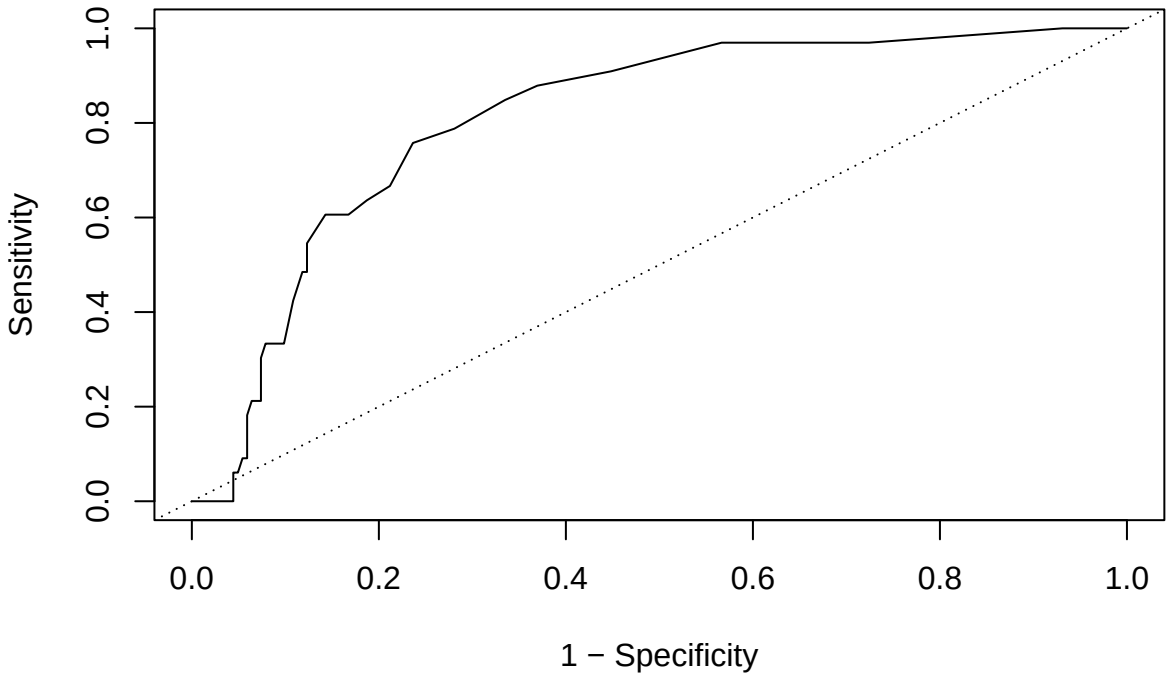

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

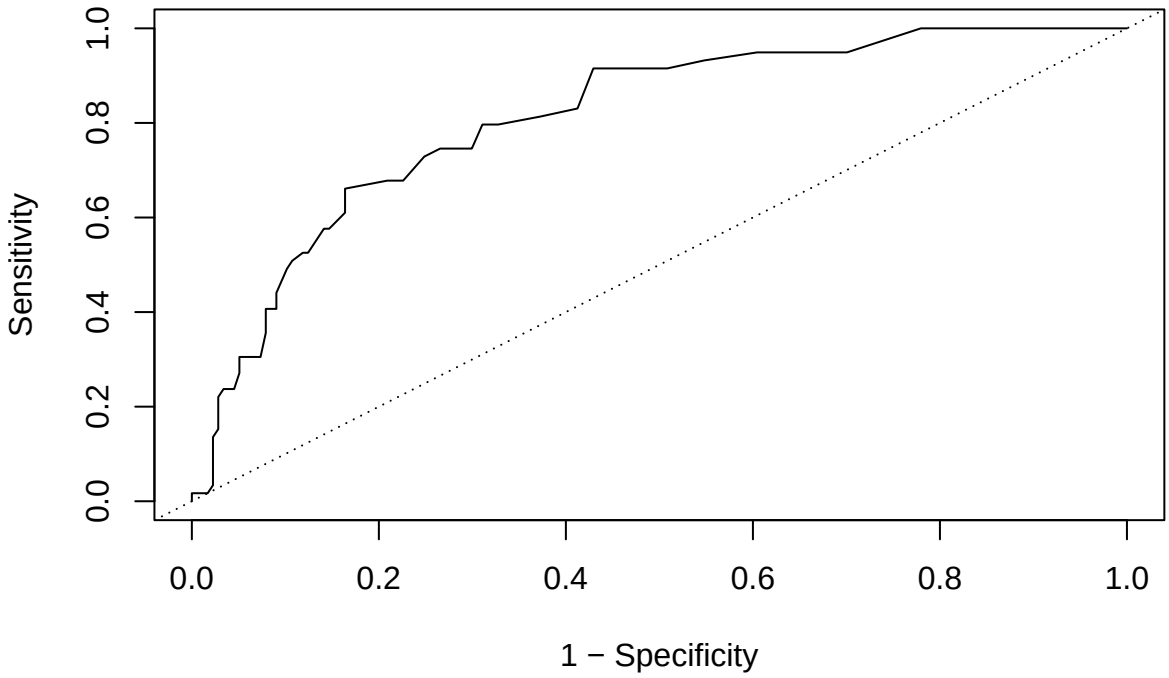

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

0.4.2 Start Time 3

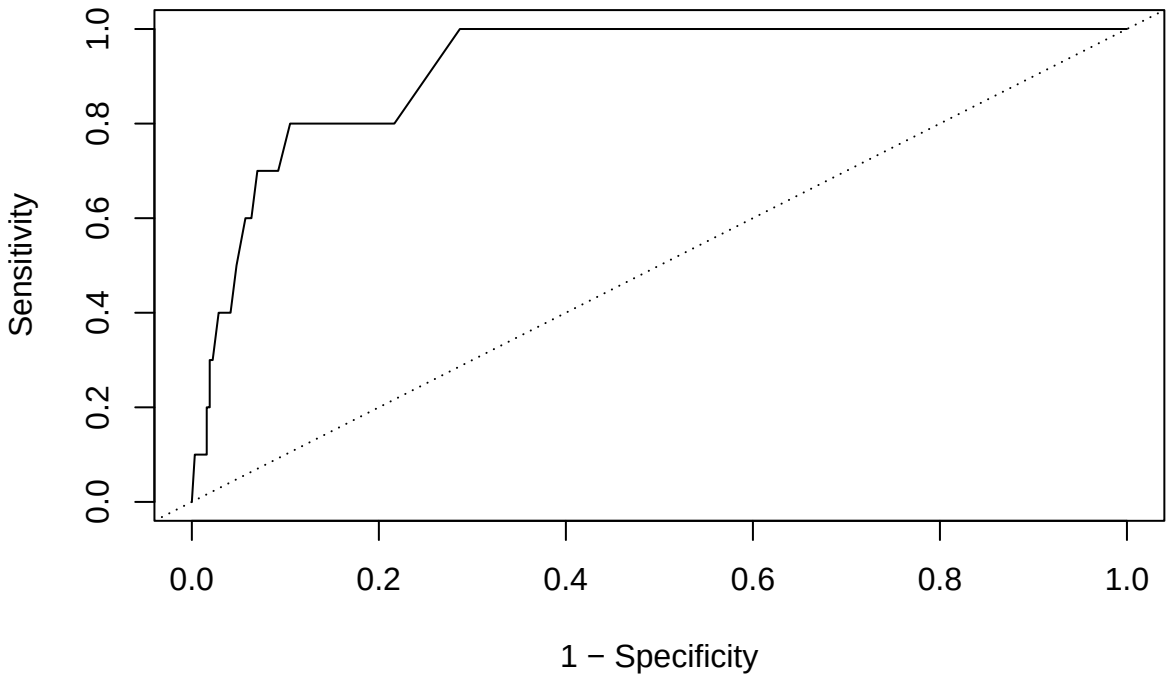


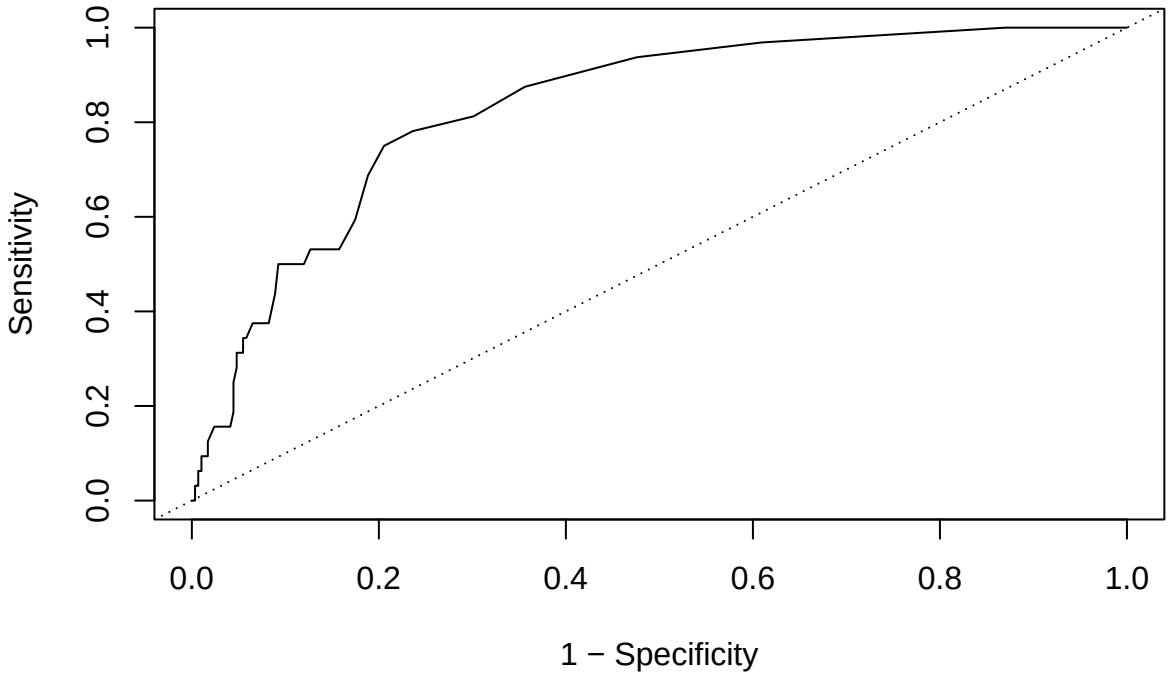


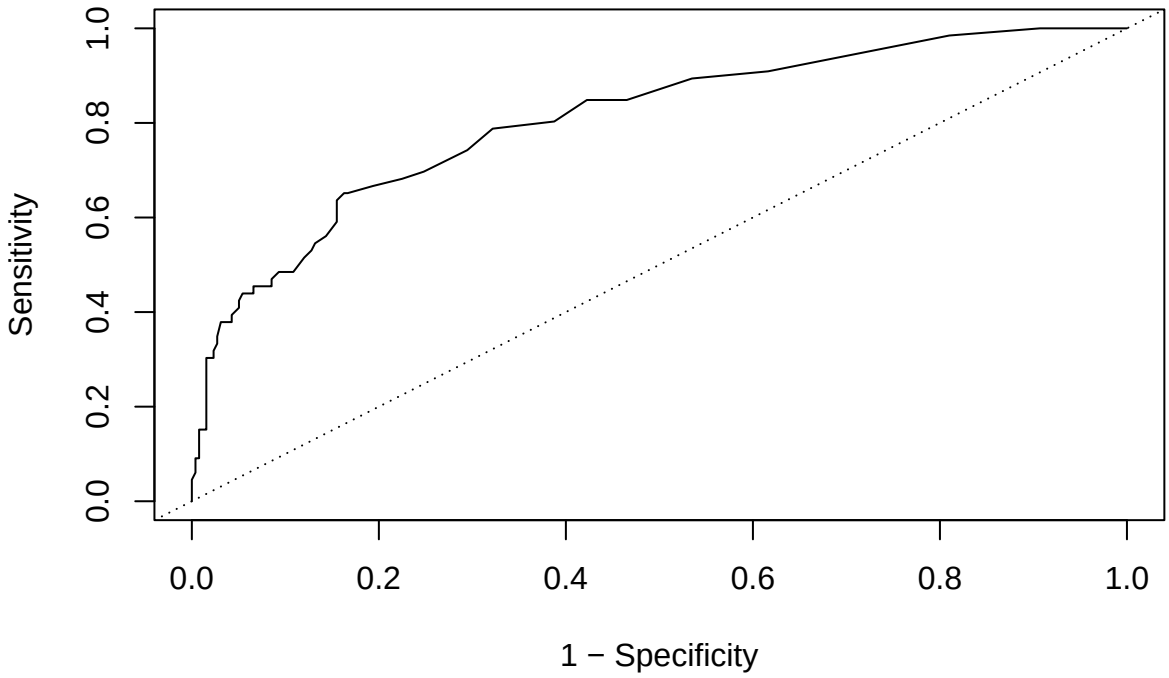

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

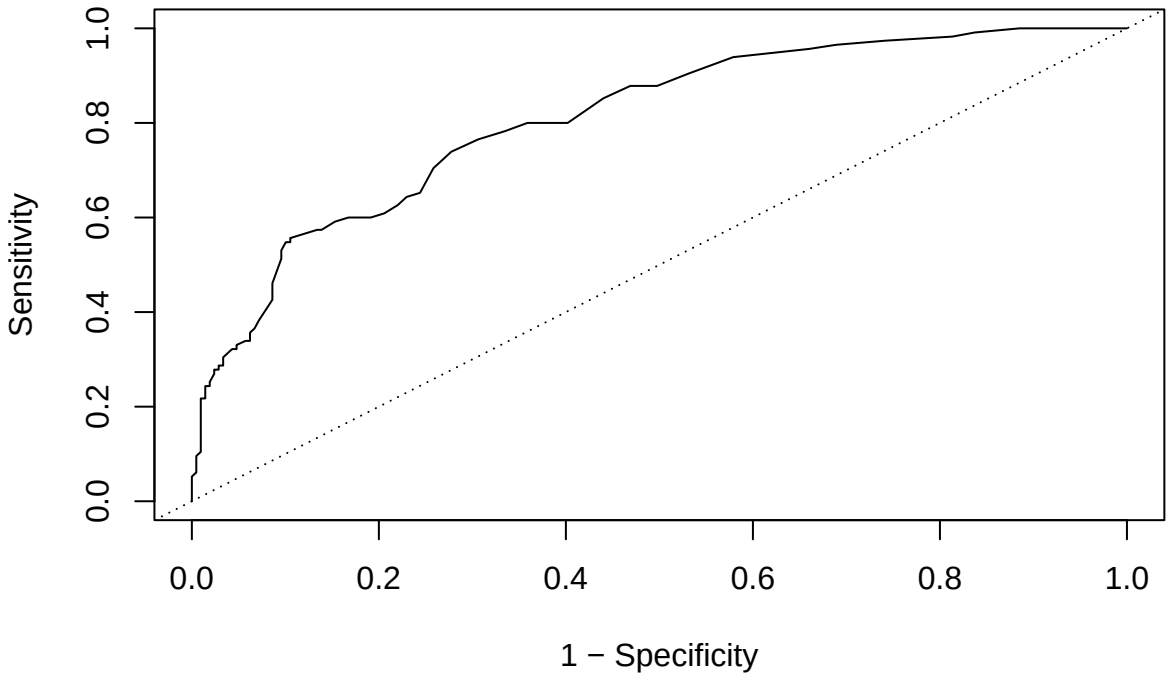

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

0.4.3 Start Time 4

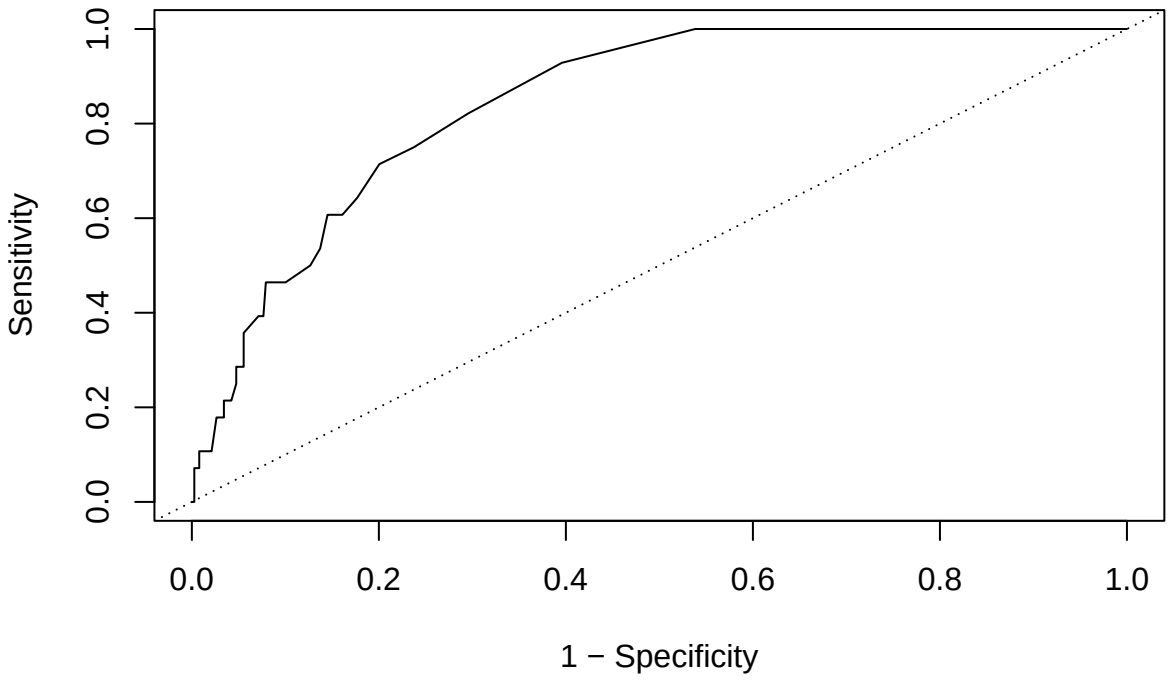


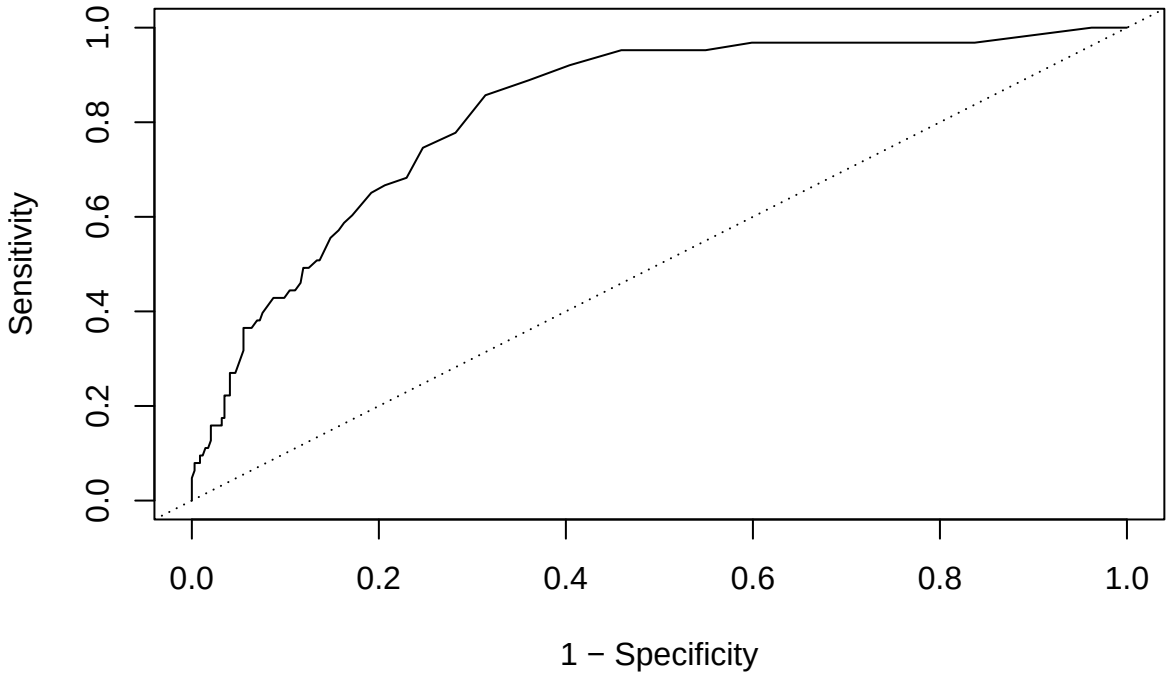


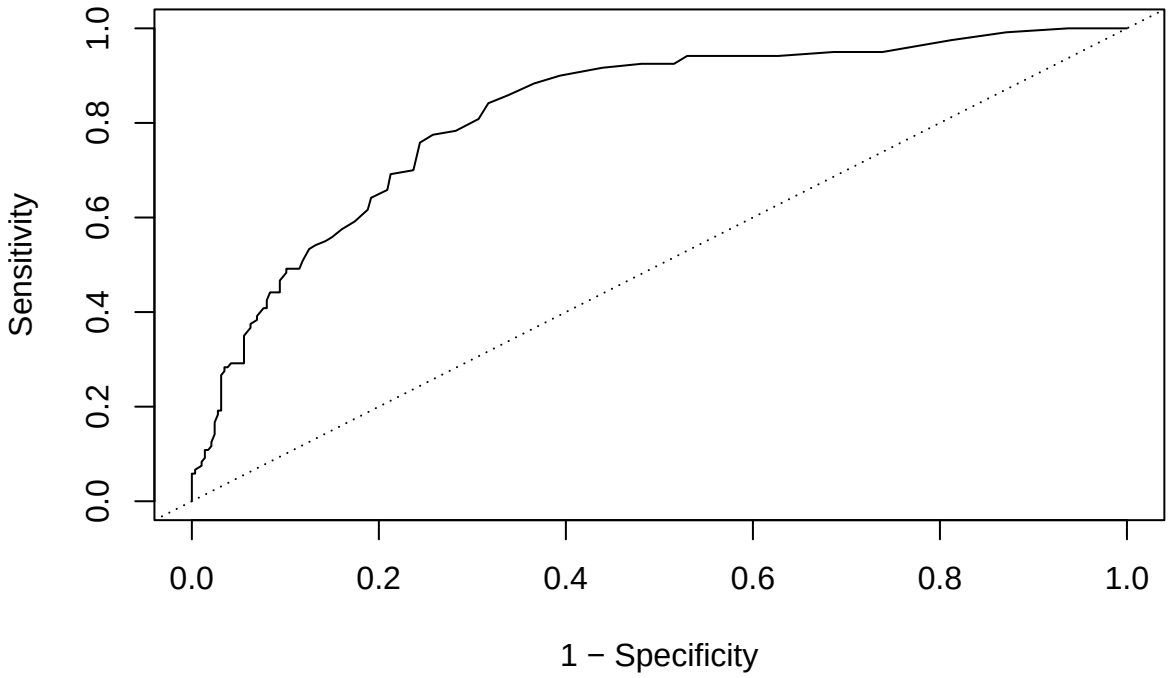

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

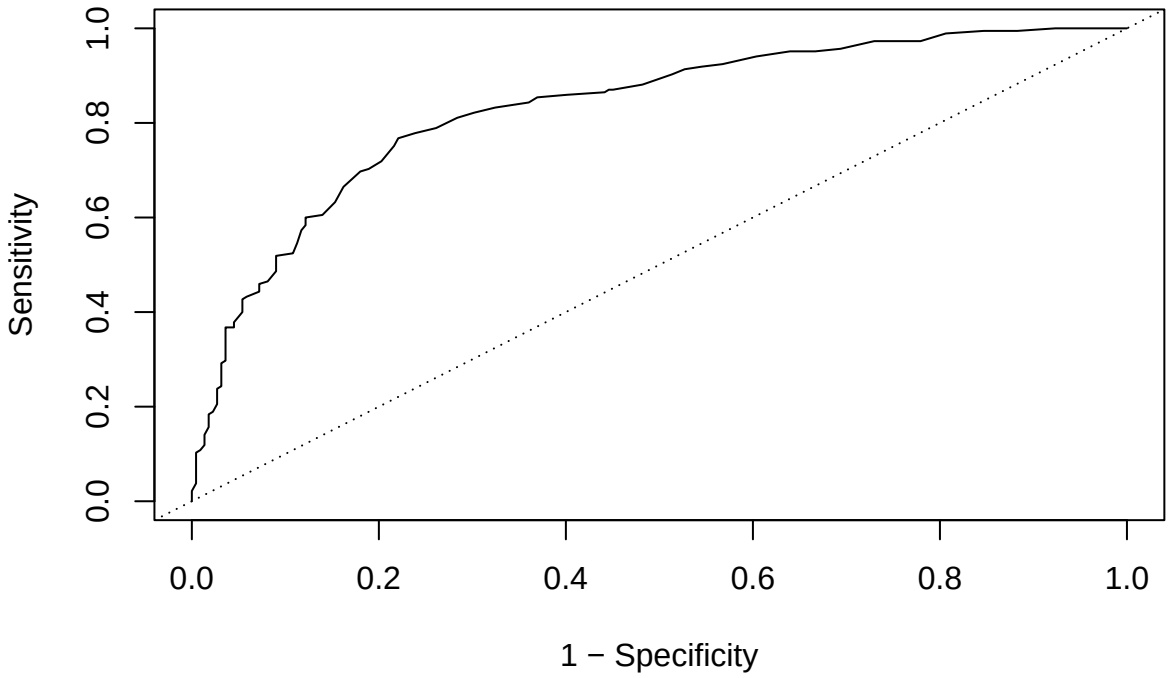

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

0.4.4 Start Time 5

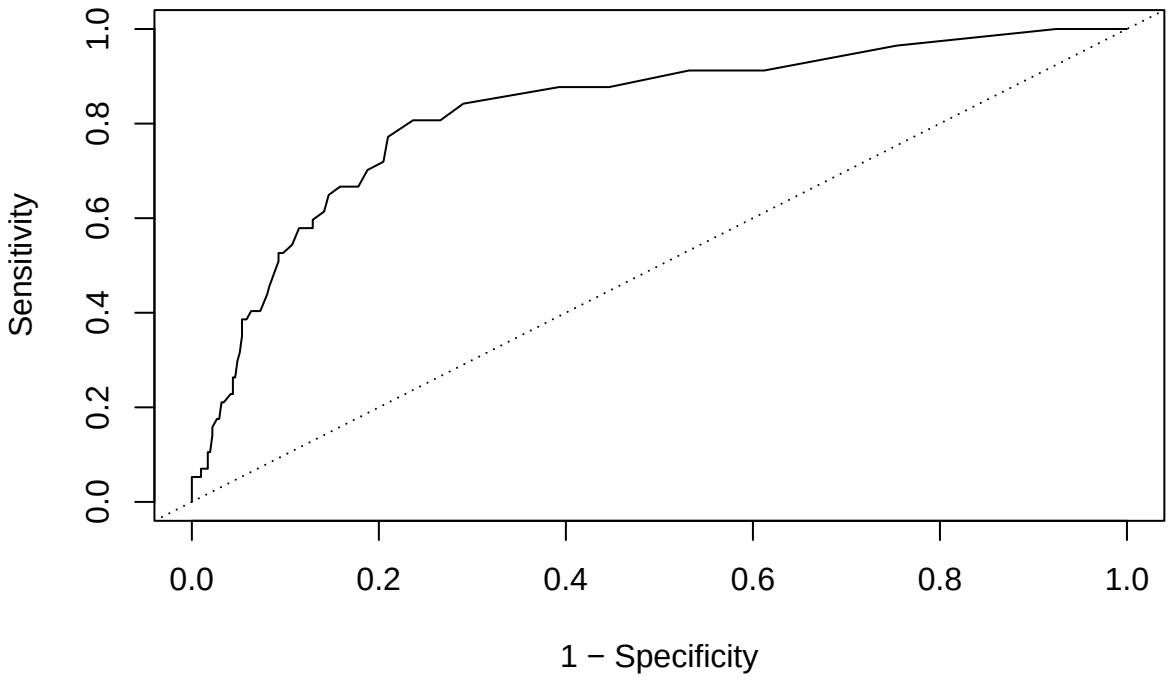


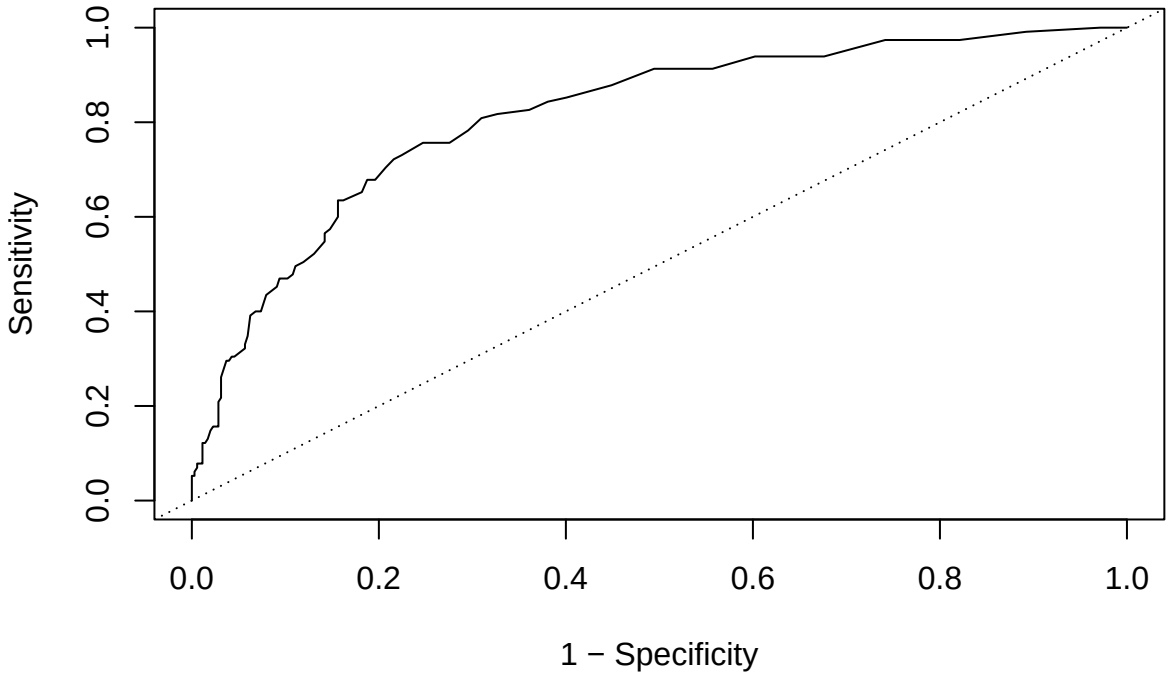


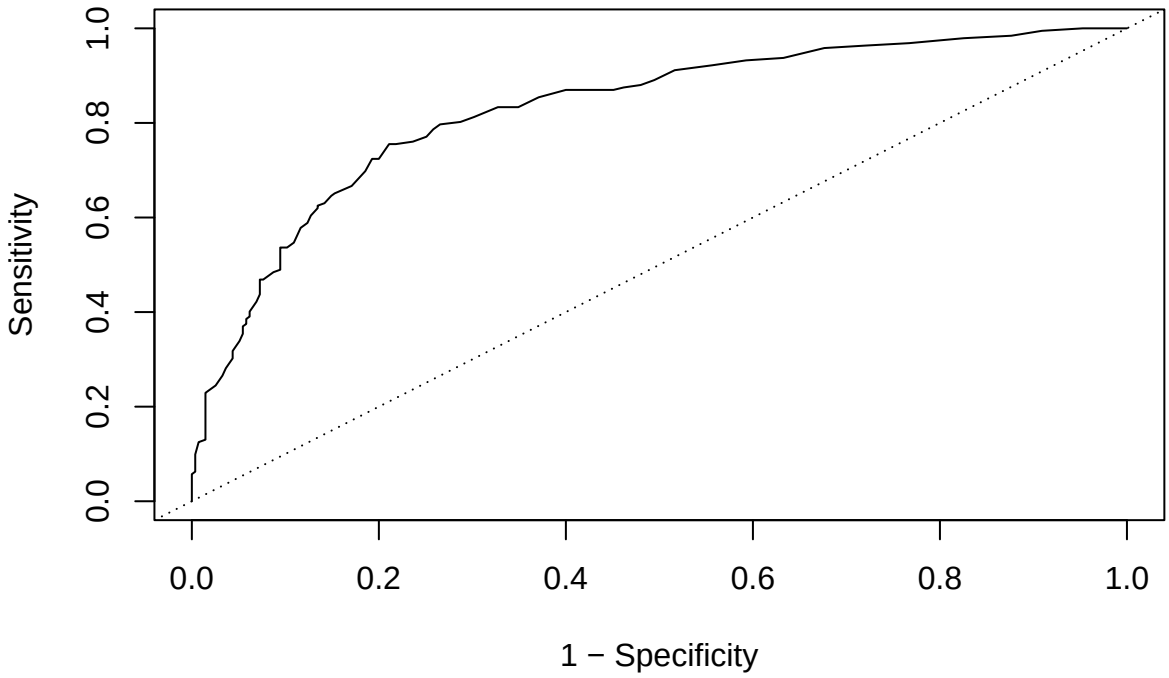

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

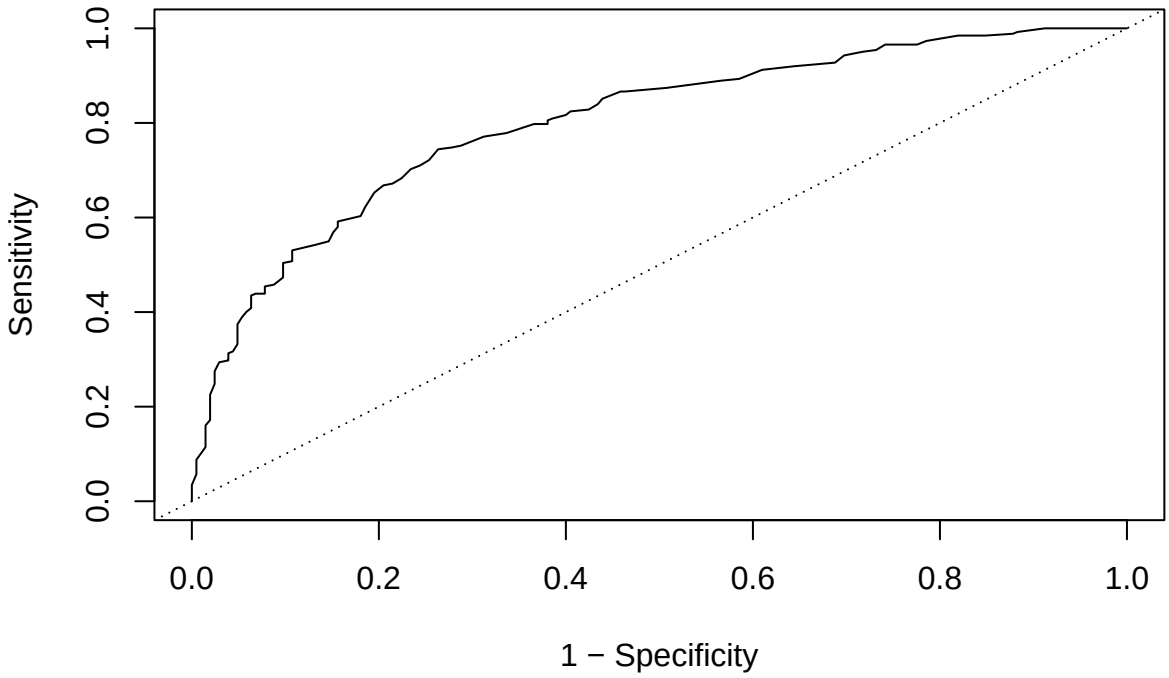

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

0.4.5 Start Time 6

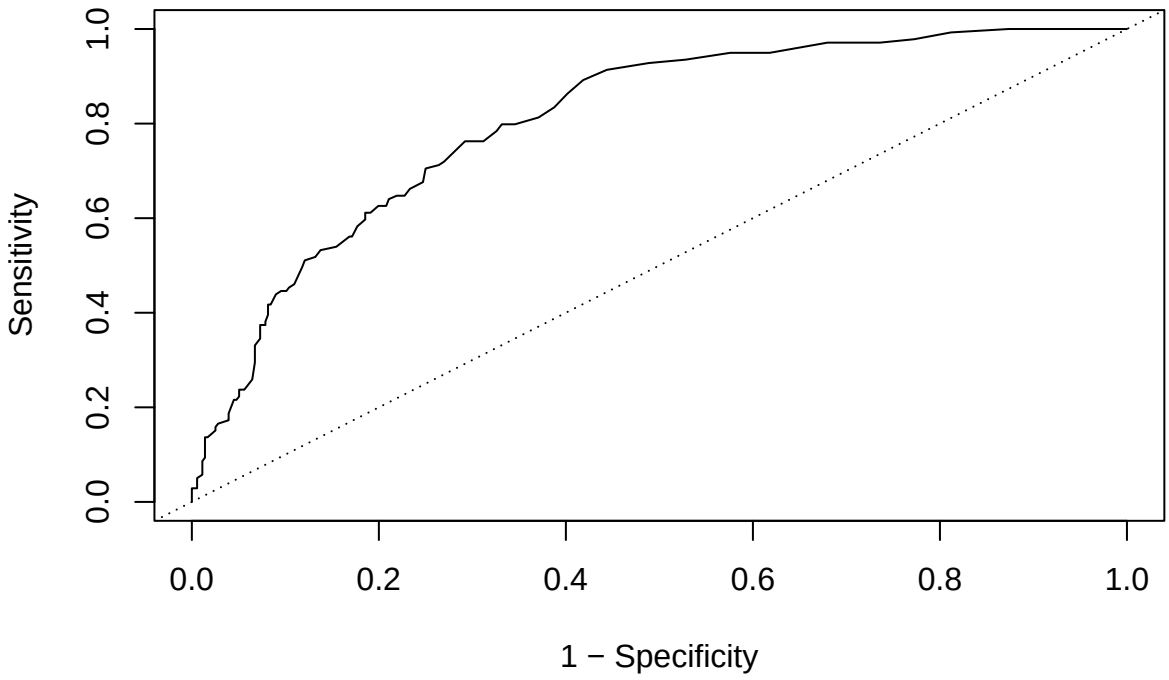


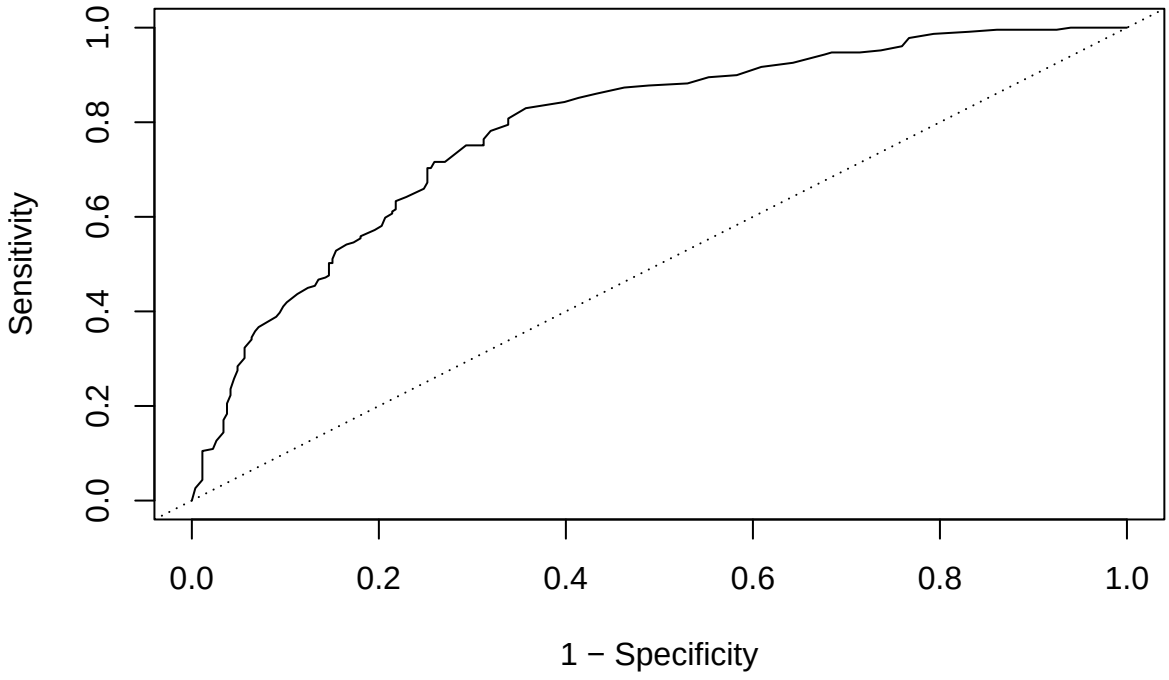


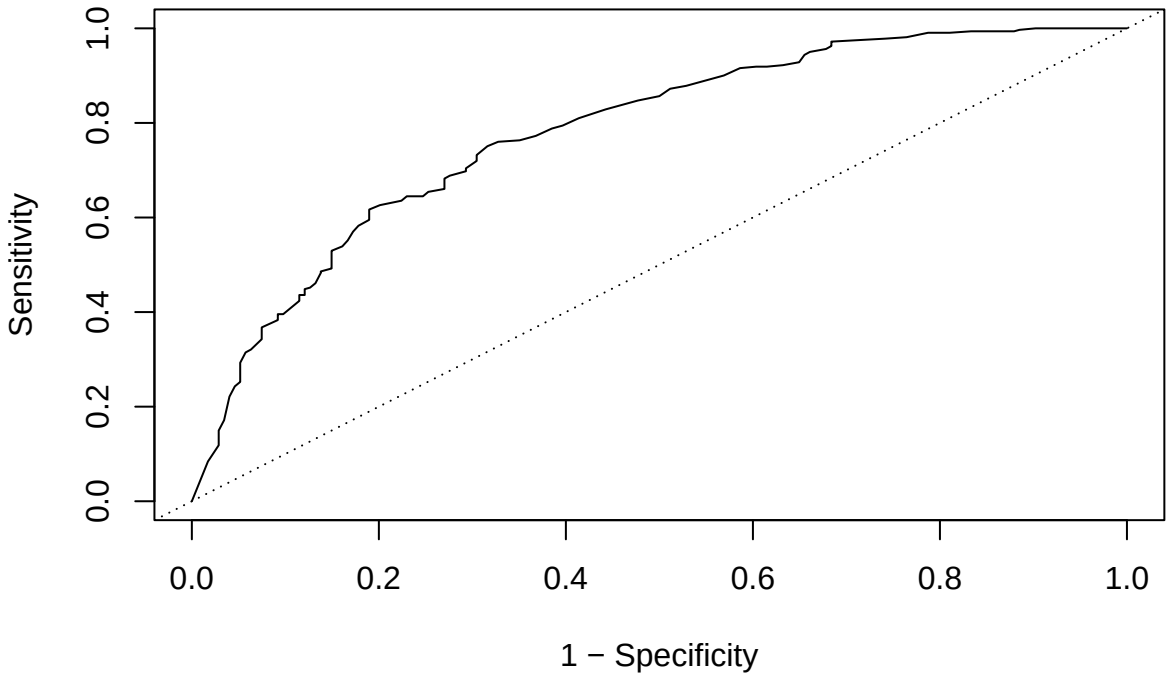

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

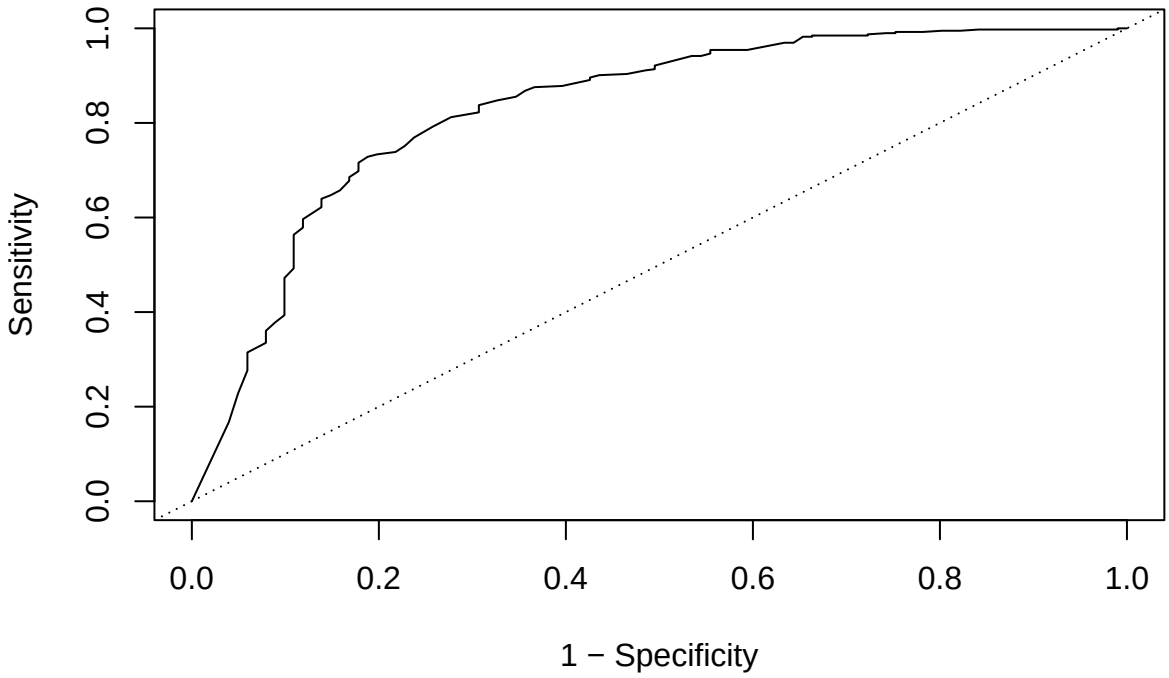

Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

0.4.6 Start Time 8






Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity


Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1 − Specificity

0.5 Brier Score

The warning massage for the Brier score is shown below:

```
Error in br(Thoriz) :
  it seems that there are no events in the interval [2.000001, 3).

Warning: there are fewer than 5 subjects with an event in the interval [2.000001, 4).
```

```
#> Error in br(Thoriz) :
#>   it seems that there are no events in the interval [2.000001, 3).
```

```
#> Warning in br(Thoriz): there are fewer than 5 subjects with an event in the interval [2.000001, 4).
```

```
#> Warning in br(Thoriz): there are fewer than 5 subjects with an event in the interval [3.000001, 4).
```

```
#> Warning in br(Thoriz): there are fewer than 5 subjects with an event in the interval [4.000001, 5).
```

	Dt1	Dt2	Dt3	Dt4	Dt5
Tstart2	NA	0.0128	0.0349	0.0649	0.1177
Tstart3	0.0084	0.0244	0.0470	0.1146	0.1545
Tstart4	0.0110	0.0261	0.0794	0.1243	0.1764
Tstart5	0.0133	0.0577	0.1089	0.1600	0.1768
Tstart6	0.0433	0.0893	0.1416	0.1700	0.1832
Tstart8	0.0949	0.1589	0.1932	0.1835	0.1219

```
#> % latex table generated in R 4.2.2 by xtable 1.8-4 package
#> % Thu Apr 11 09:41:16 2024
#> \begin{table}[ht]
#> \centering
#> \begin{tabular}{rrrrrr}
#>   \hline
#>  & Dt1 & Dt2 & Dt3 & Dt4 & Dt5 \\
#>   \hline
#> Tstart2 &  & 0.01 & 0.03 & 0.06 & 0.12 \\
#>   Tstart4 & 0.01 & 0.03 & 0.08 & 0.12 & 0.18 \\
#>   Tstart6 & 0.04 & 0.09 & 0.14 & 0.17 & 0.18 \\
#>   Tstart8 & 0.09 & 0.16 & 0.19 & 0.18 & 0.12 \\
#>    \hline
#> \end{tabular}
#> \caption{BS Table}
#> \end{table}
```