6th May 2020

Analysis of the Clinical Trial on D-penicillamine efficiency to treat Primary Biliary Cirrhosis performed by the Mayo Clinic in the 70's

Cyril Basquin

INTRODUCTION

Primary biliary cirrhosis (PBC), a specific form of cirrhosis, is a chronic progressive cholestatic granulomatous disease, with destructive inflammatory lesion of small intralobular and septal bile ducts, which is likely to be caused by an autoimmune mechanism with the presence of serum antimitochondrial antibodies (1).

The aim of the clinical trial performed in the Mayo Clinic between 1974 and 1984 was to test the efficiency of D-penicillamine (a chelator of heavy metals, and more specifically copper) which was proposed to be a good candidate to treat PBC in a small clinical trial (2). At the entry in the trial, different biomarker from patients were measured. During this analysis, these markers will be used to build a model for survival prediction of the disease.

A total of 424 PBC patients referred to Mayo Clinic during that ten-year interval (1974-1984), met eligibility criteria for the randomized placebo-controlled trial of the drug D-penicillamine. The first 312 cases in the data set participated in the randomized trial and contain largely complete data. The additional 112 cases did not participate in the clinical trial but consented to have basic measurements recorded and to be followed for survival. Six of those cases were lost to follow-up shortly after diagnosis, so the data here are on an additional 106 cases as well as the 312 randomized participants (3).

The database is constituted of some general information about patients (Age, Sex, Time since entry in the cohort (Time)), informations about the disease itself (The patient survived at the end of the trial, was transplanted or die (status), the stage of the PBC (ranked from mild to severe: 1 to 4) and different marker of cirrhosis gravity sign (Presence of ascites, hepatomegaly, blood vessel malformations in the skin, increase coagulation time (protime), presence of edema, biomarker of liver deficiency (alkaline phosphatase, aspartate aminotransferase, bilirubin, albumin), copper concentration (copper is a target of D-penicillamine). Triglyceride and cholesterol were also monitored. A basic description of these parameters in the dataset (for Placebo and D-penicillamine group) is presented in figure 1.

Note: the code in R to access easily to this dataset, as well as all the tests and construction performed in this report is accessible here: **https://tinyurl.com/ybupqf5g**

```
       id             time            status        trt               age            sex     ascites hepato  spiders edema
 Min.   :  1.00   Min.   :  1.347   0:187   drug     :158   Min.   :26.28   m: 36   0:288   0:152   0:222   0  :263
 1st Qu.: 78.75   1st Qu.: 39.129   1:125   Not_rand:  0   1st Qu.:42.24   f:276   1: 24   1:160   1: 90   0.5: 29
 Median :156.50   Median : 60.435           placebo :154   Median :49.79                                   1  : 20
 Mean   :156.50   Mean   : 65.917                          Mean   :50.02
 3rd Qu.:234.25   3rd Qu.: 88.616                          3rd Qu.:56.71
 Max.   :312.00   Max.   :149.684                          Max.   :78.44

      bili             chol            albumin          copper          alk.phos            ast
 Min.   : 0.300   Min.   : 120.0   Min.   :1.96   Min.   :  4.00   Min.   :  289.0   Min.   : 26.35
 1st Qu.: 0.800   1st Qu.: 249.5   1st Qu.:3.31   1st Qu.: 41.25   1st Qu.:  871.5   1st Qu.: 80.60
 Median : 1.350   Median : 309.5   Median :3.55   Median : 73.00   Median : 1259.0   Median :114.70
 Mean   : 3.256   Mean   : 369.5   Mean   :3.52   Mean   : 97.65   Mean   : 1982.7   Mean   :122.56
 3rd Qu.: 3.425   3rd Qu.: 400.0   3rd Qu.:3.80   3rd Qu.:123.00   3rd Qu.: 1980.0   3rd Qu.:151.90
 Max.   :28.000   Max.   :1775.0   Max.   :4.64   Max.   :588.00   Max.   :13862.4   Max.   :457.25
                  NA's   :28                      NA's   :2
      trig           platelet        protime       stage
 Min.   : 33.00   Min.   : 62.0   Min.   : 9.00   1: 16
 1st Qu.: 84.25   1st Qu.:199.8   1st Qu.:10.00   2: 67
 Median :108.00   Median :257.0   Median :10.60   3:120
 Mean   :124.70   Mean   :261.9   Mean   :10.73   4:109
 3rd Qu.:151.00   3rd Qu.:322.5   3rd Qu.:11.10
 Max.   :598.00   Max.   :563.0   Max.   :17.10
 NA's   :30       NA's   :4
```

<u>Figure 1:</u> Basic description of the data in the dataset. Time: number of months between registration and the earlier of death ('1') or transplantion / end of study analysis in July, 1986 ('0'), **status**: status at endpoint, for censored ('0'), dead ('1') participants, **trt**: treatment D-penicillmain (drug), placebo, not randomised (Not_rand), age in years**, sex:** Female/Male, ascites: presence of ascites (1 for presence, 0 for absence)**, edema**: no edema ('0'), untreated or successfully treated ('0.5'), presence despite diuretic therapy ('1'), **hepato:**hepatomegaly (yes: 1, no: 0), **spiders:** blood vessel malformations in the skin (presence: 1, absence: 0), **albumin**: serum albumin (g/dl), **alk.phos**: alkaline phosphotase (U/liter), **ast:** aspartate aminotransferase (U/ml). **bili:** serum bilirunbin (mg/dl), **chol**: serum cholesterol (mg/dl), **copper**: urine copper (ug/day). **trig:** triglycerides (mg/dl). **platelet:** platelet count, **protime:** standardised blood clotting time (equivalent of Prothombine time, i.e coagulation).

RESULTS

<u>A / D-penicillamine as a treatment of PBC?</u>

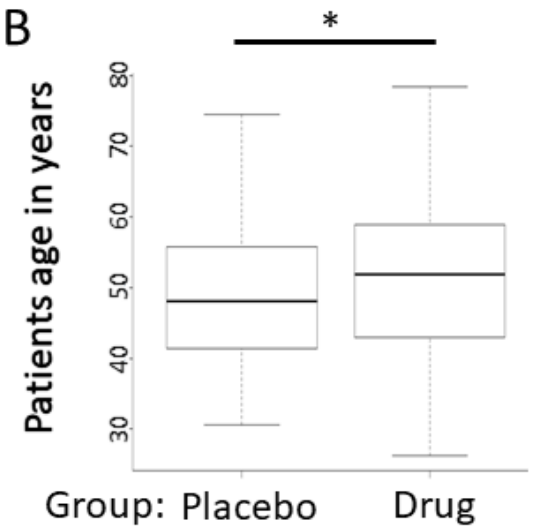
The Dataset give us the age of patient when they entered the trial, their sex, the received treatment (Placebo, Drug, not treated) and the status at endpoint: censored, transplanted, dead. For the purpose on this study, transplanted patients were pooled in the censored group.

PBC is a disease that target female predominantly. Because of this, the male population is outnumbered in the cohort (13% of male, Figure 2A) but, they are equally distributed between the placebo and the D-penicillamin group. Chi-square, p=0.3774 (figure 2A). We will keep them for the study.

PBC has been histologically differentiated in 4 stage. As expected, advance form of the disease lead to lower survival (Cox regression model, $p < 10^{-10}$). In the cohort, patients that reached the different stage are fairly distributed in the 2 groups.

A

	Male	Female
Drug	21	137
Placebo	15	139
Total	36	276



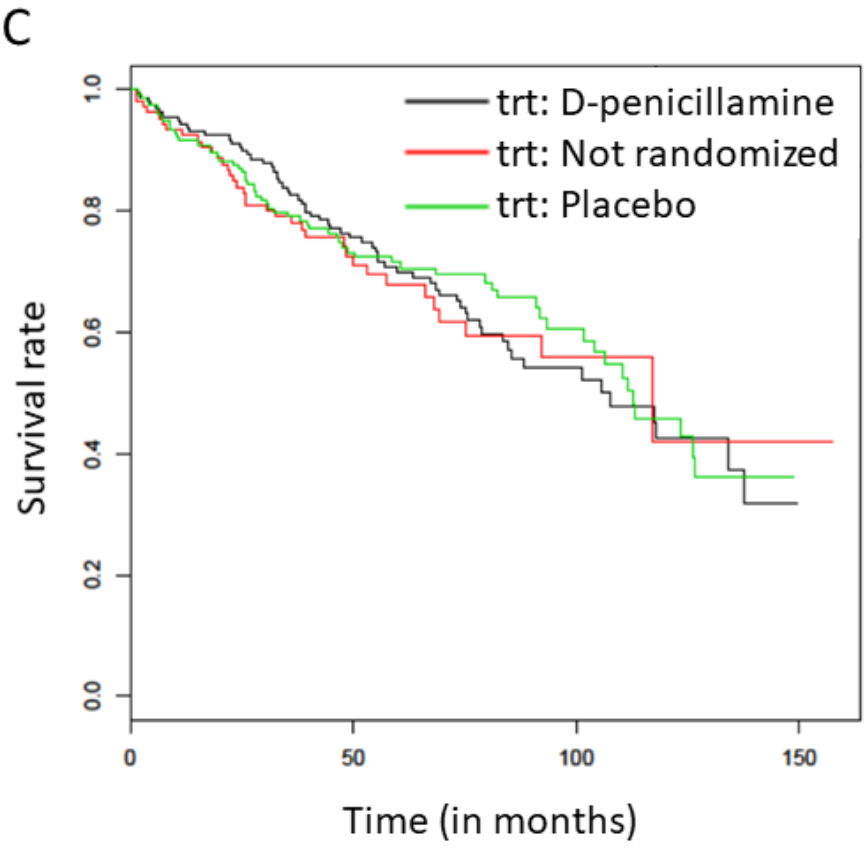


Figure 2: Effect of D-penicillamine on PBC. A) Table that's shows the repartition of female and male in placebo group and drug group (chi²; p>0.37). B) Age repartition between the 2 groups. t.test, p=0.017. C) Survival rate depending on time using Kaplan-Meier estimator.

The placebo group is younger (median=48.11) that the drug group (median= 51.93). t-test p=0.017, Figure 2B. As age is a risk factor (p-value < 10^{-6} with Cox Regression over the whole dataset) , to overcome this

unfair distribution, age of patients was stratified into 3 Groups depending on 1st and 3rd Quartile (group 1 < 43 years, group2 is between 43 years and 56 years, and group3: > 56 years).

Finally, we build a Cox regression model with treatment, sex and stratified age groups as covariates. In this model, no statistical differences where seen neither between the treatment (treated vs placebo p = 0.883) nor the sex (p = 0.174).

The absence of effect of D-penicillamine is confirmed by the plot of Kaplan-Meier estimators that shows no differences between the placebo group, the drugged group and the not randomized group (figure1). The logrank test using treatment as unique feature is also not significant (p=0.7). Altogether, those results show that D-penicillamine is inefficient to treat PBC.

B/ Building a model to predict patient survival

In order to evaluate survival rate of the patient, we build a model using the different monitored measures. As we saw previously, D-penicillamine is inefficient to treat PBC. As the cohort is small, we decided to build 2 models. One using only the placebo group (in case of the drug is not as inefficient as we think), and another one in which we pooled the data from Placebo and drug groups to augment the size of the dataset.

To construct the Cox regression model, a backward stepwise selection was performed using AIC as criterion. Shortly, the model with all covariates was build, it's Akaike information criterion (AIC) was calculated and the covariate with the higher p-value was removed if the AIC of the null model is higher than the one of the current models. All data provided in the dataset were taken into consideration except stage of the disease was excluded due to a too small number of patients in stage 1. The variable 'edema' was also remove because in the category '0.5' patients in which cured edema patients were pooled with untreated edema patients. From my point of view, this pool was not relevant and should not appear in the model because it did not differentiate really patients with or without edema. To construct the model, a log function was applied to all covariates that are expressed as a concentration.

By using a Cox regression model, with a backward stepwise selection on data from placebo group only, we end-up with a model with 4 covariates (age + spiders + log(bili) + ascites). Ascites was not significant and was removed from the model. The final model contains 3 covariates (age + spiders + log(bili)). All of them are positively correlated (positive beta coefficient) with diminution of survival time. The concordance of our model is 0.843 (se 0 0.023) which confirm its relevance. To check the coherence of this model, martingales residuals were calculated and plotted for each covariate (Figure 3). Except for 'extreme' values (age > 60 years and bilirubin > 15 mg/dl) for which we have less data, the residuals are linear enough to accept the model. In addition, to detect the potential over-influence of some specific observation, we calculated the Deltas-Betas

residuals. The idea is to calculate how coefficients change if each individual observation is omitted. We do not identify any patient out of range from the other using this analysis (Figure 3).

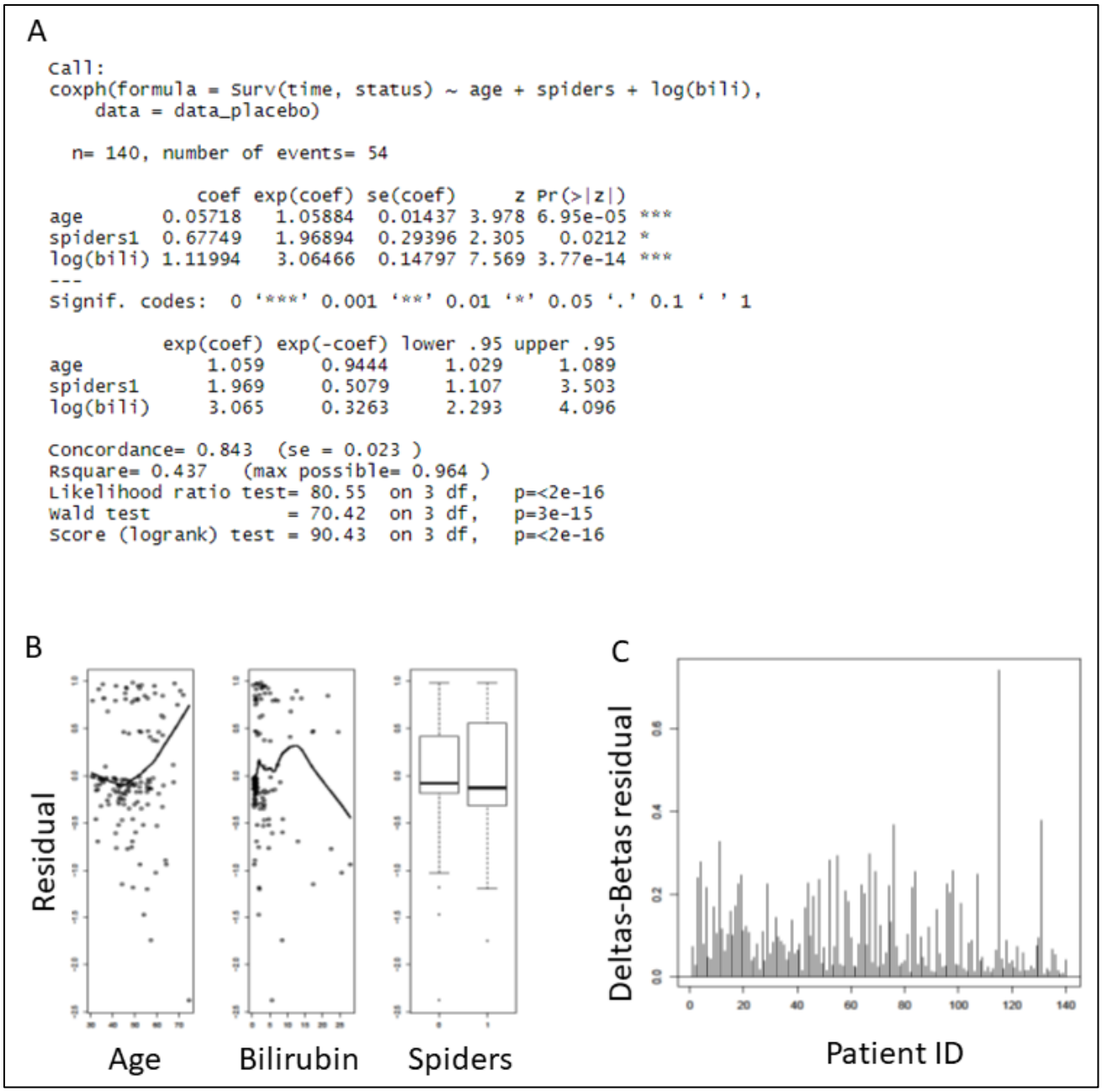


Figure 3: Cox regression Model 1. A) Summary of the model construct using only the Placebo group. B) Martingales residuals for the 3 covariates of the model. C) Deltas-Betas residual for each patient of the dataset.

The same approach was performed with the full dataset (Placebo + Drug group). Following the same reasoning that in the first model that we built, we end-up with a model with 5 covariates (Age + log(bili) + log(albumin) + log(copper) + protime). The concordance of this second model is similar to the first model (0.843 se = 0.02). In this model, martingales residuals are correct, and no patients were observed out of range in the Deltas-Betas method (Figure 4). In order to "improve" Martingales residual in our model, we applied a

quadratic term and/or logarithm term on copper and/or age. But no improvement was observed on Martingales (data not shown). Stratification of these covariates could be a way to explore to improve the model.

A

```
call:
coxph(formula = Surv(time, status) ~ age + log(bili) + log(albumin) +
    log(log(copper)) + protime, data = data_model)

  n= 276, number of events= 111

                      coef exp(coef)  se(coef)      z Pr(>|z|)
age                0.030402  1.030869  0.009047  3.360 0.000778 ***
log(bili)          0.787412  2.197700  0.119322  6.599 4.14e-11 ***
log(albumin)      -3.034878  0.048081  0.735726 -4.125 3.71e-05 ***
log(log(copper))   1.710813  5.533459  0.643237  2.660 0.007821 **
protime            0.292090  1.339223  0.092062  3.173 0.001510 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                   exp(coef) exp(-coef) lower .95 upper .95
age                  1.03087     0.9701   1.01275    1.0493
log(bili)            2.19770     0.4550   1.73941    2.7767
log(albumin)         0.04808    20.7984   0.01137    0.2033
log(log(copper))     5.53346     0.1807   1.56845   19.5219
protime              1.33922     0.7467   1.11812    1.6040

Concordance= 0.843  (se = 0.02 )
Rsquare= 0.455   (max possible= 0.981 )
Likelihood ratio test= 167.6  on 5 df,   p=<2e-16
Wald test            = 159.5  on 5 df,   p=<2e-16
Score (logrank) test = 200.8  on 5 df,   p=<2e-16
```

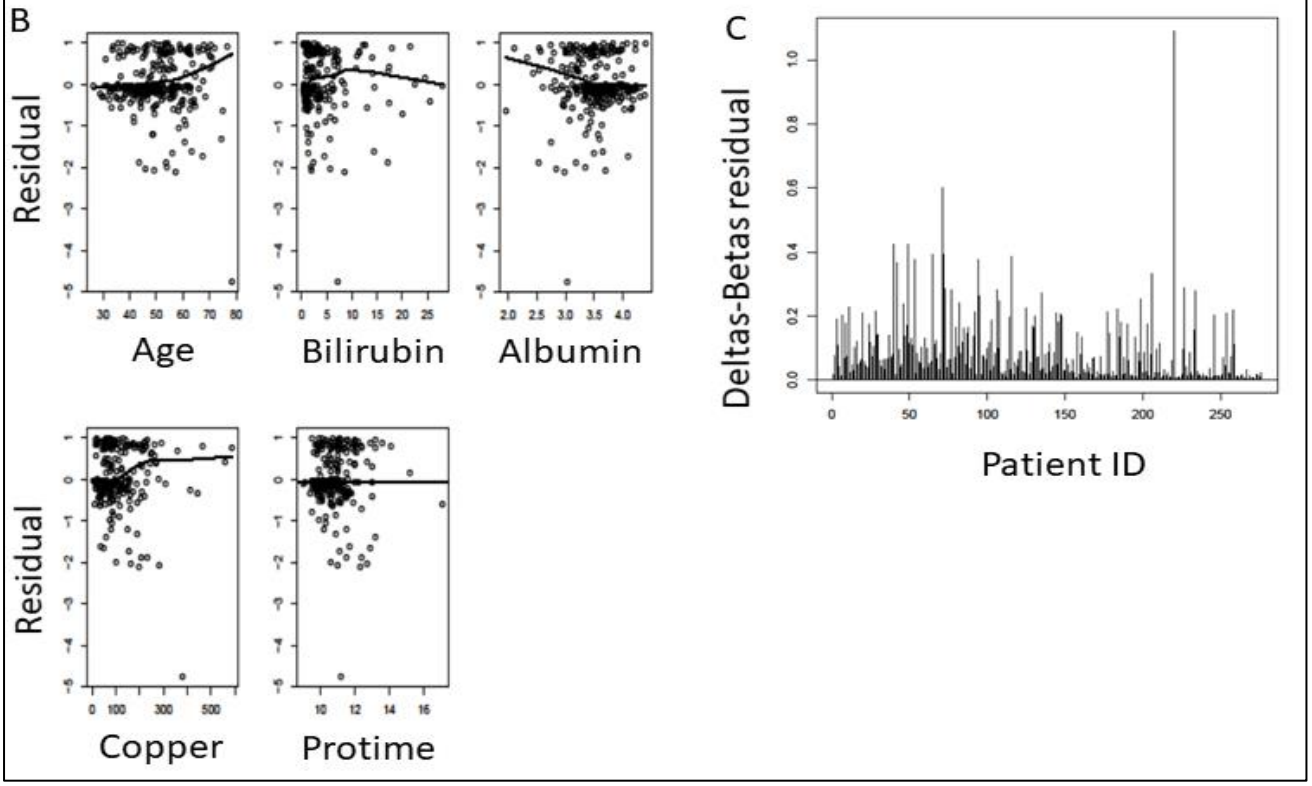


Figure 4: Cox regression Model 1. A) Summary of the model construct using only the Placebo group. B) Martingales residuals for the 3 covariates of the model. C) Deltas-Betas residual for each patient of the dataset.

Conclusion

Our analysis did not allow us to see any effects of D-penicillamine over PBC. This was confirmed by the cox regression model that do not include treatment as a covariate in model 2. The D-penicillamine drug targets serum copper and interestingly we found in model 2 that copper (at least in the urine) can be used to predict survival of patients. Consequently, it seems coherent to try to reduce copper concentration to treat the disease, but alone this is not enough. This assumption was kind of confirm by model 1 in which copper was not included.

Age and high concentration of bilirubin are correlated with a decrease in life expectancy in both models. An increasing age is associated with a lot of disease and it's not surprising to identify it here. Bilirubin is a yellow pigment produce by the degradation of red blood cells (4) and liver defects (as in cirrhosis) is often associated with high concentration of this molecule. It's also coherent to identify it as a factor of risk in PBC.

Unfortunately, due to the lack of data, we cannot test the strongness of our model on the not randomized group since, presence of ascites (model 1) and urine copper quantity (model 2) was not monitored in this group. Further study will be required to use our model in the future to predict of survival time of patients with PBC.

Is our model sufficient to predict PBC disease? The response is probably not. In general, to screen for cirrhosis, and not specifically Primary biliary cirrhosis, the Child-Pugh score is perfomed in routine. This score is composed of 5 criteria: concentration of bilirubin and albumin, Prothombin Time (=protime), presence of ascites and hepatic encephalopathy. Strikingly, except for hepatic encephalopathy that was not monitored in this study, the 4 others criteria are present in model 1, model 2 or both. This suggest that the models that we built to predict survival time might be extended to other form of cirrhosis.

Bibliography

(1) Reshetnyak, 2015. Primary biliary cirrhosis: Clinical and laboratory criteria for its diagnosis.

(2) Jain et al., 1977. A controlled trial of D-Penicillamine therapy in primary biliary cirrhosis

(3) https://rdrr.io/cran/survival/man/pbc.html

(4) https://en.wikipedia.org/wiki/Bilirubin