

EDA

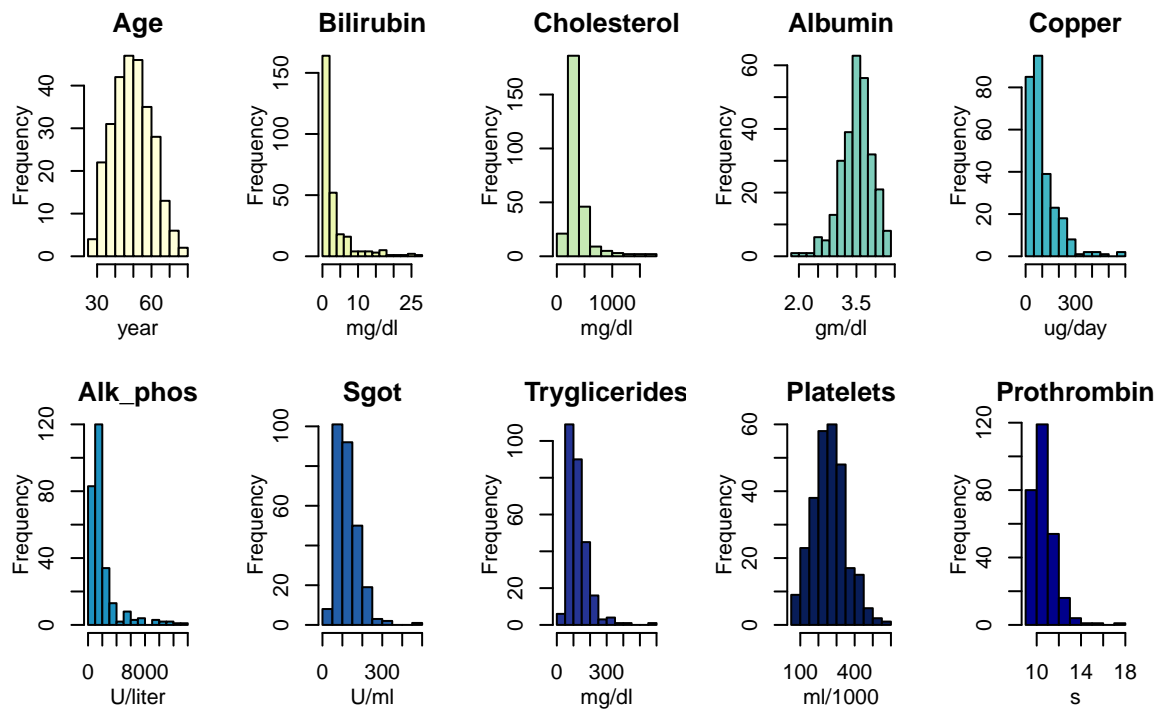
Chen Liang

2024-11-16

Load data

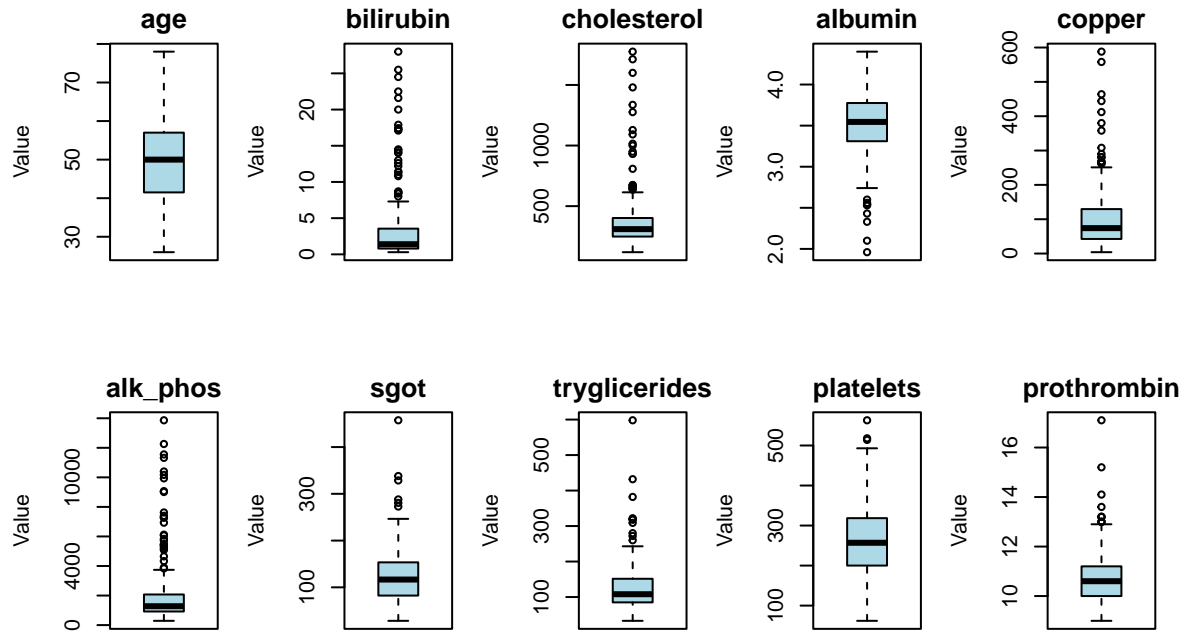
```
## Rows: 418 Columns: 20
## -- Column specification -----
## Delimiter: ","
## chr (7): Status, Drug, Sex, Ascites, Hepatomegaly, Spiders, Edema
## dbl (13): ID, N_Days, Age, Bilirubin, Cholesterol, Albumin, Copper, Alk_Phos...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

Histogram Plots for continuouse variables

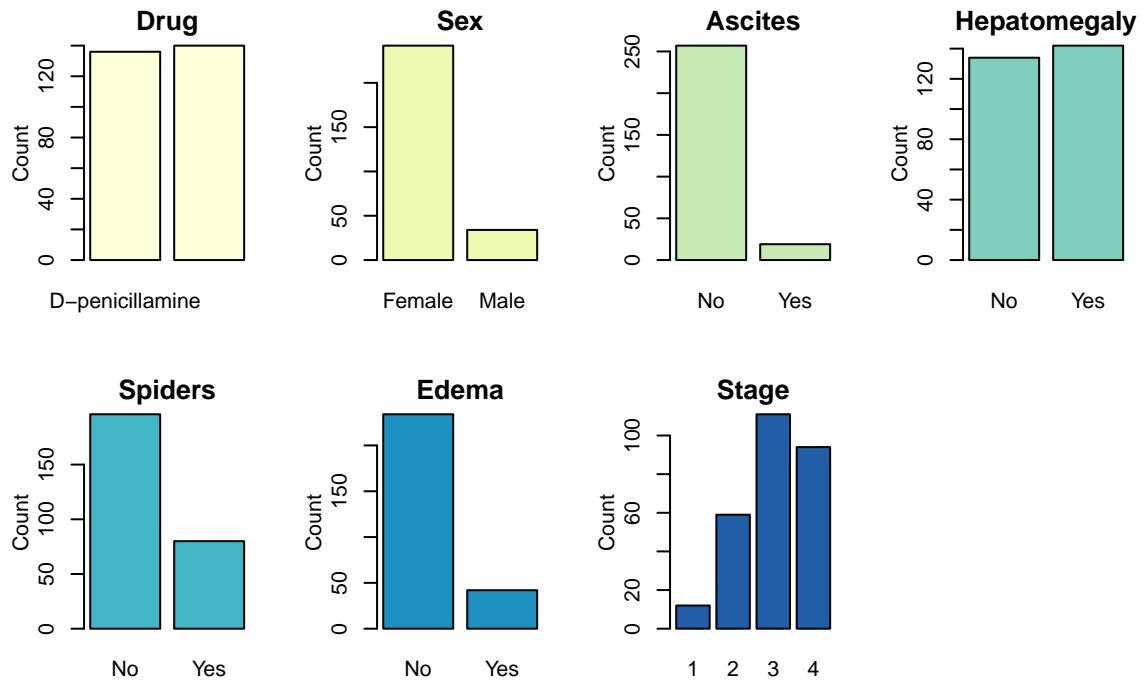


Boxplot for continuous variables

Boxplots for Continuous Variables



Bar Plots for categorical variables



Correlation Plot

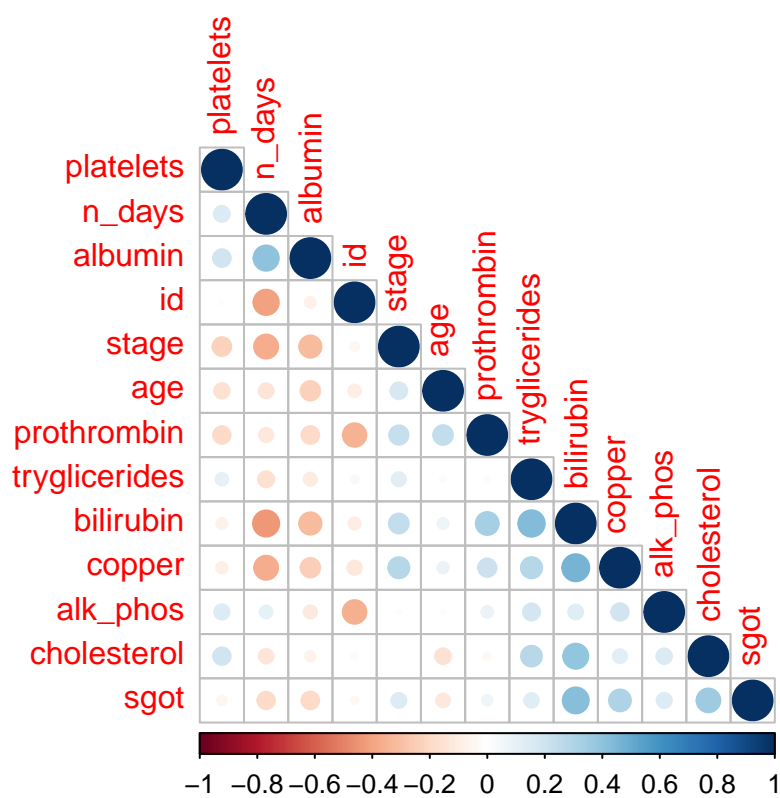


Table 1: Baseline Characteristics

Setting theme ‘New England Journal of Medicine‘

Table 1: Baseline Characteristics

Characteristic	Censored, N = 147 ¹	Censored due to liver tx, N = 18 ¹	Death, N = 111 ¹
N_days	2,391.8 / 2,224.0 (984.3)	1,511.6 / 1,368.0 (754.4)	1,508.5 / 1,191.0 (1,110.4)
Drug			
D-penicillamine	70 (48%)	9 (50%)	57 (51%)
Placebo	77 (52%)	9 (50%)	54 (49%)
Age	48.3 / 48.0 (10.3)	40.7 / 40.5 (6.0)	53.4 / 53.0 (10.0)
Sex			
Female	137 (93%)	15 (83%)	90 (81%)
Male	10 (6.8%)	3 (17%)	21 (19%)
Ascites	1 (0.7%)	0 (0%)	18 (16%)
Hepatomegaly	55 (37%)	12 (67%)	75 (68%)
Spiders	29 (20%)	5 (28%)	46 (41%)
Edema	8 (5.4%)	2 (11%)	32 (29%)
Bilirubin	1.6 / 0.9 (1.8)	3.2 / 3.3 (2.0)	5.7 / 3.3 (6.2)
Cholesterol	326.9 / 293.0 (168.1)	439.5 / 343.5 (335.5)	418.9 / 344.0 (277.9)
Albumin	3.6 / 3.6 (0.3)	3.6 / 3.6 (0.4)	3.4 / 3.4 (0.5)
Copper	68.1 / 52.0 (58.7)	123.3 / 101.0 (102.9)	140.3 / 121.0 (100.9)
Alk_phos	1,501.1 / 1,120.0 (1,376.8)	1,509.7 / 1,253.5 (854.4)	2,731.8 / 1,794.0 (2,765.3)
SGOT	110.2 / 97.0 (54.4)	130.2 / 123.5 (38.0)	141.5 / 134.9 (57.7)
Tryglicerides	111.1 / 103.0 (47.8)	133.9 / 124.0 (70.5)	141.8 / 124.0 (79.3)
Platelets	267.0 / 265.0 (86.4)	294.8 / 297.5 (79.9)	249.5 / 236.0 (102.1)
Prothrombin	10.4 / 10.2 (0.9)	10.4 / 10.2 (0.6)	11.2 / 11.0 (1.0)
Stage			
1	11 (7.5%)	0 (0%)	1 (0.9%)
2	42 (29%)	3 (17%)	14 (13%)
3	62 (42%)	8 (44%)	41 (37%)
4	32 (22%)	7 (39%)	55 (50%)

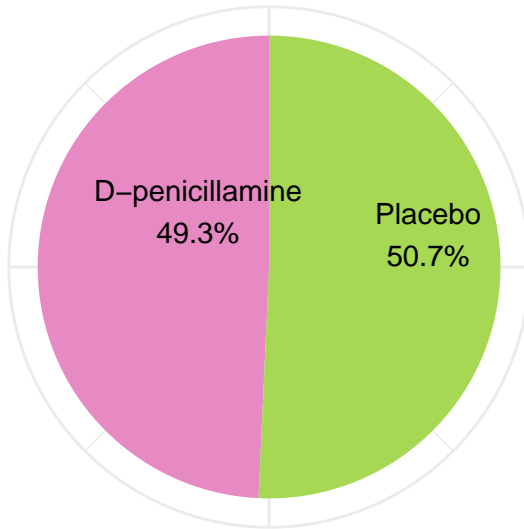
¹Mean / Median (SD); n (%)

Fraction of Treatment & Outcomes

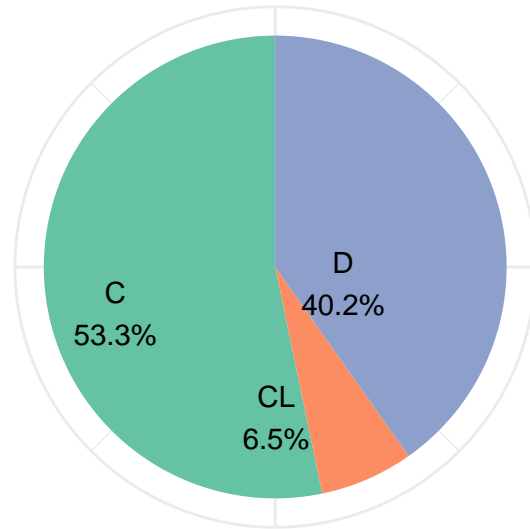
Distribution of Patient Outcomes and Drug Treatment

Total Patients: 276

Drug Treatment



Patient Status



Category C CL D D-penicillamine Placebo
53.3% 6.5% 40.2% 49.3% 50.7%

Multivariate analysis

```
## Warning: There was 1 warning in 'mutate()'.  
## i In argument: 'status = case_when(...)'.  
## Caused by warning:  
## ! NAs introduced by coercion  
  
## Table printed with 'knitr::kable()', not {gt}. Learn why at  
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html  
## To suppress this message, include 'message = FALSE' in code chunk header.
```

Table 2: Multivariate Cox Proportional Hazards Analysis

Characteristic	HR	95% CI	p-value
age	1.02	1.00 to 1.04	0.049
sex			
Female	—	—	
Male	1.30	0.72 to 2.34	0.38
bilirubin	1.12	1.08 to 1.16	<0.001
albumin	0.38	0.22 to 0.63	<0.001
copper	1.00	1.00 to 1.01	0.002
prothrombin	1.28	1.06 to 1.53	0.009
stage	1.50	1.14 to 1.98	0.003