

Examining the Progression of Primary Biliary Cirrhosis in a Longitudinal Cohort of Mayo Clinic Patients, 1974-1984

Joe LaRocca

P8157: Analysis of Longitudinal Data

Professor Zhonghua Liu

December 12, 2025

Abstract

Primary biliary cirrhosis (PBC) is a chronic autoimmune condition that leads to substantial liver damage. We used a linear mixed effects model to analyze a longitudinal cohort of 312 Mayo Clinic patients from 1974 to 1984 and determine whether the severity of PBC or the progression of PBC over time was influenced by treatment with D-penicillamine, a copper-chelating agent used to eliminate copper-containing bile in the liver. Adjusting for age, sex, and other relevant biomarkers, we found that there was no statistically significant association between D-penicillamine treatment and serum bilirubin levels or the change in serum bilirubin over time, indicating that the treatment was not effective within the study cohort. Recent meta-analyses have also arrived at the conclusion that D-penicillamine does not significantly affect health outcomes in patients with PBC, and alternative treatments have become gold standards today as a result. Limitations include the fact that serum bilirubin is an imperfect predictor of PBC severity as well as small amounts (< 5%) of missing data for some relevant covariates.

Keywords

longitudinal data analysis, linear mixed effects models, random effects, primary biliary cirrhosis, D-penicillamine

Introduction

Primary biliary cirrhosis (PBC), often known as primary biliary cholangitis, is a chronic autoimmune disease that can cause significant liver damage if left untreated.¹ One of the primary indicators of PBC is cholestasis, a condition where bile flow is reduced or stopped, leading to a build-up of bile in the liver. Mechanistically, PBC often occurs through the destruction of biliary epithelia of interlobular bile ducts, which can in turn damage the bile ducts themselves and lead to cholestasis. According to the American Association for the Study of Liver Diseases, PBC can be diagnosed if two of the following criteria are met: biochemical evidence of cholestasis, presence of PBC-specific antibodies, or histologic evidence of nonsuppurative destructive cholangitis and the destruction of interlobular bile ducts.² PBC can be difficult to diagnose because its primary symptoms (pruritus and fatigue) are nonspecific and symptoms are often not easily attributable to the disease.

A recent meta-analysis conducted on 59 studies between 1976 and 2024 found that the global prevalence and incidence rate of PBC are approximately 18.1 cases per 100,000 people and 1.8 cases per 100,000 person-years, respectively.³ PBC is much more common in women than in men, and is most prevalent in women between 60 and 79 years old. While historically, PBC has been most prevalent in North America and Europe, PBC rates have risen fastest in the Asia-Pacific region since about 2000.

In the 1970s, the D-penicillamine was proposed as a potential treatment for PBC. D-penicillamine acts as a copper-chelating compound that can bind to copper in the liver and therefore reduce liver copper concentrations, which are often elevated in patients with PBC due to the buildup of copper-containing bile in the liver.⁴ The objective of our study is to evaluate whether D-penicillamine was associated with improved outcomes on patients from the 1970s and early 1980s. Specifically, we are investigating whether there was a statistically significant difference in the change in serum bilirubin levels, a strong indicator of the severity of PBC, over time between the treatment and control groups, independent of other biomarker concentrations and measures of liver health.

¹ Tanaka, A., Ma X., Takahashi A., and Vierling JM. 2024. "Primary Biliary Cholangitis." *The Lancet* 404 (10457): 1053–66. [https://doi.org/10.1016/S0140-6736\(24\)01303-5](https://doi.org/10.1016/S0140-6736(24)01303-5)

² Lindor et al., "Primary Biliary Cholangitis: 2018 Practice Guidance."

³ Tan et al., "Global Epidemiology of Primary Biliary Cholangitis."

⁴ Jain et al., "Controlled Trial of D-Penicillamine Therapy."

Materials and Methods

The Mayo Clinic Primary Biliary Cirrhosis (PBC) Study was run between 1974 and 1984 and sought to investigate the effect of D-penicillamine treatment on the progression of PBC over time. A total of 1,945 data points were collected on 312 patients, 158 of which were assigned to the D-penicillamine group and the other 154 of which were assigned to the control group. The primary outcome was serum bilirubin levels; high levels of bilirubin in the blood indicate that the liver is not properly clearing it, implying impaired liver function.

The dataset includes several biological covariates that are potentially associated with serum bilirubin levels. These variables include other biomarker concentrations in the blood, including those of cholesterol, other biological indicators of liver disease (ascites, hepatomegaly, spider angiomas, and edema), and blood clotting measures (platelet count per microliter and prothrombin time). Additionally, each patient's age, sex, treatment status, and time from enrollment (for post-baseline study visits) were also included.

In order to isolate the effect of D-penicillamine in the presence of relevant covariates, we used a linear mixed-effects (LME) model to account for correlation between data points (since nearly all participants recorded more than one study visit). We first evaluated multiple candidate models, accounting for different combinations of random effects and variance structures, and then selected a final model using minimum AIC. Finally, we ran model diagnostics to assess whether our model could be used to accurately predict outcomes of other PBC patients treated with D-penicillamine.

Results

Exploratory Data Analysis

Out of the 312 participants in the study, 276 (88.5%) were female. The mean age of all participants was 50.0 years, with a standard deviation of 10.2 years. 109 (34.9%) had stage 4 PBC, while 120 (38.5%) had stage 3 PBC; the remaining participants had either stage 1 or stage 2 PBC.

The distributions of serum bilirubin levels among the treatment and control groups are very similar; both are right-skewed with medians and interquartile ranges of 1.300 (0.700, 3.615) and 1.400 (0.800, 3.750) respectively (Figure 1). Visually, there is no clearly defined difference between the trajectory of patients in the treatment and control groups; we see, however, that serum bilirubin levels generally increase over time and sharply increase for several patients, with strong evidence of nonlinear progression over time (Figure 2). The distribution of the number of study visits was slightly right-skewed, with a mode of 4, a minimum of 1, and a maximum of 15 (Figure 3).

Figure 1: Distributions of Serum Bilirubin Level by Treatment Status

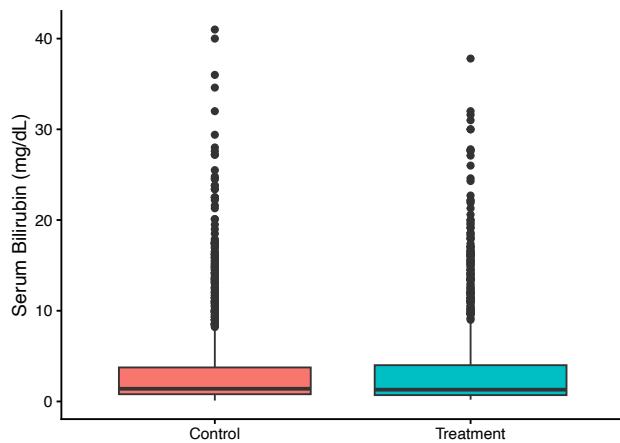


Figure 2: Trends in Serum Bilirubin Level Over Time

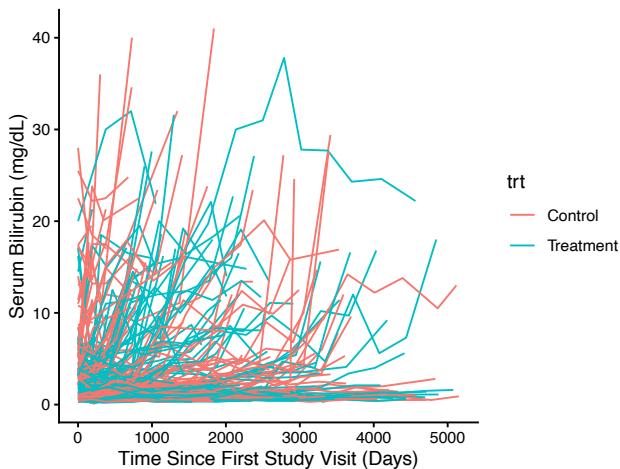
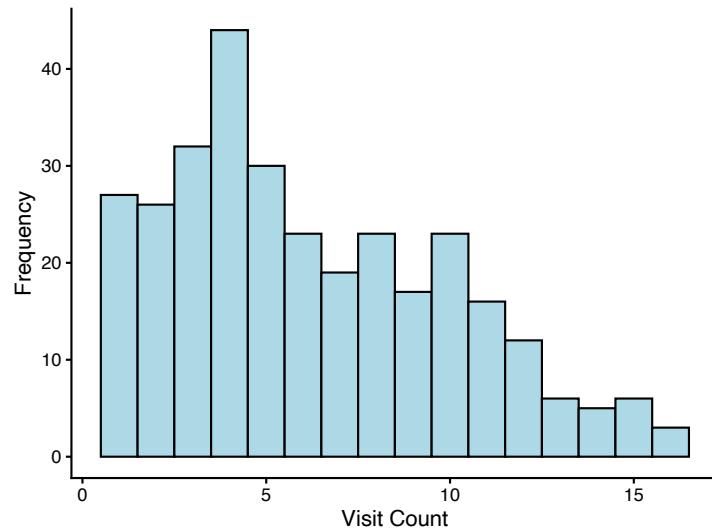


Figure 3: Distribution of Number of Study Visits Per Patient



Missing Data

Data were complete for all variables used in the model except for ascites, hepatomegaly, spider angiomas, alkaline phosphatase, and platelet count. For each of the five variables for which missing data was present, missingness was between 3% and 4%, with 60, 61, 58, 60, and 73 out of 1,945 total observations missing for ascites, hepatomegaly, spider angiomas, alkaline phosphatase, and platelet count, respectively. For the first three variables, all of which were categorical, we added a missing indicator variable for missing values in order to maintain sample size and observe potential differences in data points for which each variable was missing. For alkaline phosphatase and platelet count, we imputed the value from an individual subject's previous study visit, if available; otherwise, we imputed the median value for that individual. We decided to use imputation due to the low proportion of missing data and in order to keep as many data points as possible.

Model Selection

We decided to log-transform serum bilirubin levels since the distribution of serum bilirubin was highly right-skewed; the distribution of log-transformed serum bilirubin much more closely resembled Normal behavior. For the sake of interpretation and the lack of a theoretical reason, we did not transform any predictor variables.

Since there were a very large number of potential candidate models, we decided to first select a model based on combinations of fixed and random effects. Once we selected a candidate model, we then selected a covariance structure while holding both fixed and random effects constant. We decided to include treatment status, age, sex, and time from follow-up for all candidate models since we hypothesized that all three variables would have a clear association with the response. We also included an interaction term between treatment status and time in all models since we sought to investigate whether the rate of change of serum albumin levels over time differed based on treatment status.

In one set of models, we included only a set of “basic” variables, which included treatment status, age, sex, time from follow-up, and the treatment/time interaction term. In the second set, we also included relevant markers of liver health (albumin, glutamic-oxaloacetic transaminase, edema, ascites, hepatomegaly, spider angiomas, alkaline phosphatase). Finally, in the third set, we added measures of blood clotting (platelet count, prothrombin time). Within each set, we tested a model with fixed effects only, a random intercept model, and a random intercept/random slope model. We fit each model using maximum likelihood (ML) in order to compare models with different sets of fixed effects. Out of the nine candidate models, we selected the model with basic and liver health variables, but without blood clotting variables, with both a random intercept and random slope (Model 6), by virtue of minimum AIC (Table 1).

Table 1: Results from Initial Model Selection Procedure

Model	Complexity	Effects	Log_Likelihood	AIC
Model 1	Basic	Fixed Only	-2,929	5,872
Model 2	Basic	RI	-1,883	3,783
Model 3	Basic	RI + RS	-1,525	3,069
Model 4	+ Liver Variables	Fixed Only	-2,166	4,369
Model 5	+ Liver Variables	RI	-1,537	3,112
Model 6	+ Liver Variables	RI + RS	-1,361	2,764
Model 7	+ Liver and Blood Clotting Variables	Fixed Only	-2,183	4,404
Model 8	+ Liver and Blood Clotting Variables	RI	-1,533	3,106
Model 9	+ Liver and Blood Clotting Variables	RI + RS	-1,364	2,772

After selecting fixed and random effects for the model, we tested five different covariance structures: independent, exchangeable, exponential spatial (with and without a nugget) and independent but allowing for heteroskedasticity by treatment group. We compared models using restricted maximum likelihood (REML), since fixed and random effects were the same for each of the five models fit. We found that the models fit with an exponential spatial correlation structure had a much lower AIC than those fit by other structures. Since the AIC of the model with the nugget (Model IV) was slightly lower, we picked Model IV as our final model (Table 2).

Table 2: Comparison of Residual Covariance Structures for Candidate Models

Residual_Correlation	Log_Likelihood	AIC
Independent	-1,412	2,864
Exchangeable	-1,412	2,866
Exponential Spatial	-1,005	2,052
Exponential Spatial (+ nugget)	-1,000	2,043
Independent/Heteroskedastic	-1,412	2,866

Fixed Effects Interpretation

Serum bilirubin levels did not significantly differ by age ($P = 0.498$), but did by female sex ($P = 0.200$) and time ($P < 0.001$); interestingly, holding other variables constant, female sex was associated with lower serum bilirubin levels, running contrary to the fact that women are generally at higher risk of developing PBC (Table 3). Adjusting for the other variables in the model, a 1-year difference in time from follow-up was associated with a 0.05-unit increase in log(serum bilirubin), implying that serum bilirubin gradually increased over time for the average participant.

Out of the liver health variables, edema ($P < 0.001$ for those responding to diuretics, $P = 0.005$ for those not responding), AST ($P < 0.001$), hepatomegaly ($P = 0.016$) and spider angiomas ($P = 0.037$) also had a statistically significant, positive relationship with serum bilirubin adjusting for other predictors. Therefore, holding all else constant, our model predicts that patients with edema, hepatomegaly, spider angiomas, or elevated AST to generally have higher serum bilirubin and therefore a higher severity of PBC. Serum albumin ($P = 0.298$) and ascites ($P = 0.302$), however, did not have statistically significant associations with the response, adjusting for other variables.

Interestingly, neither the treatment variable ($P = 0.350$) nor the treatment/time interaction variable ($P = 0.800$) is statistically significant, implying that the relationship between serum bilirubin levels and time does not significantly differ by treatment status, adjusting for relevant covariates.

Table 3: Fixed Effects Table for Final Model

Term	Value	SE	P-Value
Intercept	1.3044	0.3927	0.001
Treatment	-0.1224	0.1307	0.350
Age	-0.0041	0.0061	0.498
Female Sex	-0.4419	0.1969	0.026
Time (years)	0.0524	0.0103	0.000
Edema / Resp. to Diuretics	0.1081	0.0249	0.000
Edema / Not Resp. to Diuretics	0.1188	0.0425	0.005
Albumin	-0.0182	0.0175	0.298
AST	0.0006	0.0001	0.000
Ascites	0.0336	0.0325	0.302
Ascites - Missing	0.0576	0.1194	0.629
Hepatomegaly	0.0410	0.0171	0.016
Hepatomegaly - Missing	-0.0803	0.1265	0.526
Spider Angiomas	0.0435	0.0209	0.037
Spider Angiomas - Missing	0.2880	0.1763	0.102
Time : Treatment Interaction	-0.0036	0.0142	0.800

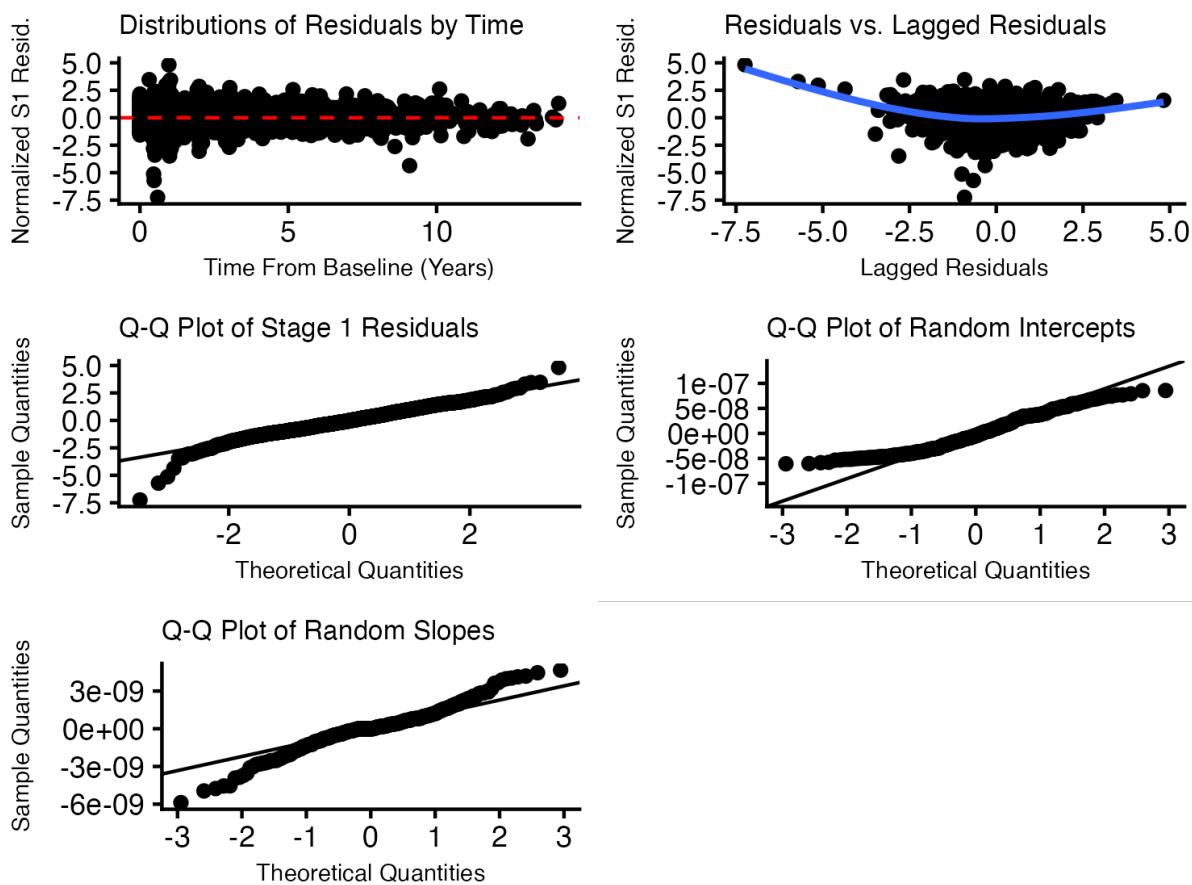
Random Effects Interpretation

The standard deviations of the random intercept and slope were 2.050×10^{-4} and 1.441×10^{-5} , respectively, indicating that most of the total variance is explained by within-subject differences as opposed to between-subject differences. Additionally, the correlation between random intercepts and random slopes was 0.000, implying no relationship between serum bilirubin at baseline and change over time.

Diagnostics

In order to assess how effectively our model fit the data, we created five diagnostic plots: a residuals vs. time plot, a residuals vs. lagged residuals plot, and three Q-Q plots for the stage 1 residuals, random intercepts, and random slopes (Figure 4). From the residual plot, we see that the residuals are approximately evenly distributed around zero, with some evidence of heteroskedasticity (specifically, lower variance at a higher time from baseline). From the residuals vs. lagged residuals plot, we see very little association between lagged residuals of -2.5 and 2.5; however, there are some data points with strongly positive residuals and strongly negative lagged residuals. From the Q-Q plots, we can see that the distributions of the stage 1 residuals, random intercepts, and random slopes each reasonably approximate Normal behavior.

Figure 4: Diagnostic Plots for Final Model



Discussion and Conclusion

We used a linear mixed-effects model to determine whether treatment with the copper-chelating drug D-penicillamine led to improved outcomes, via serum bilirubin levels, in a longitudinal cohort of 312 patients with primary biliary cirrhosis. Through our analysis, we found that serum bilirubin gradually increased over time and was positively correlated with several markers of liver health, such as edema, hepatomegaly, spider angiomas, and AST. However, our model showed no statistically significant association between treatment status and either serum bilirubin or the change in serum bilirubin over time. Therefore, we conclude that adjusting for other markers of liver health and demographic factors, D-penicillamine did not significantly ameliorate participants' PBC severity over time.

Our observation that D-penicillamine treatment status does not have a statistically significant effect on serum bilirubin levels aligns with more recent literature, such as a 2006 meta-analysis reporting that not only does D-penicillamine not reduce the risk of mortality or morbidity in patients with PBC, but also increases the risk of adverse events, likely through harmful side effects.⁵ As a result of these and other findings, D-penicillamine has been largely discontinued as a treatment for PBC in practice. Treatments used today include ursodeoxycholic acid (UDCA) and obeticholic acid (OCA), which work by replacing damaging bile acids and reducing bile salt levels within the liver, respectively.⁶ Additionally, with advances in modern medicine, liver transplantation has become an option for patients who do not respond well to current treatments.

One of the key limitations of our model is that it is built entirely on data from more than 40 years ago (study visits were logged between 1974 and 1984) and is limited to Mayo Clinic patients. Additionally, while certainly predictive, serum bilirubin levels are not a perfect measure of PBC severity. Imputation for five different predictor variables (ascites, hepatomegaly, spider angiomas, alkaline phosphatase, and platelet count) may have affected model estimates despite being used for less than 5% of data points.

⁵ Gong, Klingenberg, and Gluud, "D-Penicillamine vs. Placebo/No Intervention."

⁶ British Liver Trust, "Treating Primary Biliary Cholangitis."

References

- 1) Tanaka, A., Ma X., Takahashi A., and Vierling JM. 2024. "Primary Biliary Cholangitis." *The Lancet* 404 (10457): 1053–66. [https://doi.org/10.1016/S0140-6736\(24\)01303-5](https://doi.org/10.1016/S0140-6736(24)01303-5)
- 2) Lindor, Keith D., Christopher L. Bowlus, James Boyer, Cynthia Levy, and Marlyn Mayo. "Primary Biliary Cholangitis: 2018 Practice Guidance From the American Association for the Study of Liver Diseases." *Clinical Liver Disease* 15, no. 1 (2020): 1–2.
- 3) Tan, Jarell Jie-Rae, Ambrose Hon-Lam Chung, Jing Hong Loo, Joo Wei Ethan Quek, Sagar Sharma, Corrine Lee Singh, Roe Xin Jacqueline Yap, Wei Xuan Tay, Matthew K. Smith, Ellina Lytvyak, Andrew Mason, Aldo J. Montano-Loza, and Yu Jun Wong. "Global Epidemiology of Primary Biliary Cholangitis: An Updated Systematic Review and Meta-analysis." *Clinical Gastroenterology and Hepatology* (May 19, 2025).
- 4) Jain, S., P. J. Scheuer, S. Samourian, J. O'D. McGee, and Sheila Sherlock. "A Controlled Trial of D-Penicillamine Therapy in Primary Biliary Cirrhosis." *The Lancet* 309, no. 8016 (16 April 1977): 831–834.
- 5) Gong, Y., S. L. Klingenberg, and C. Gluud. "Systematic Review and Meta-Analysis: D-Penicillamine vs. Placebo/No Intervention in Patients With Primary Biliary Cirrhosis — Cochrane Hepato-Biliary Group." *Alimentary Pharmacology & Therapeutics* 24, no. 11-12 (December 2006): 1535–1544.
- 6) British Liver Trust. "Treating Primary Biliary Cholangitis." Last modified November 2024. Accessed November 29, 2025. <https://britishlivertrust.org.uk/information-and-support/liver-conditions/primary-biliary-cholangitis/treatment/>