Effect of Methotrexate on Progression and Survival in Patients with Primary Biliary Cirrhosis

BioStat 517, Group #1
University of Washington, Seattle, WA

Abstract

Primary biliary cirrhosis (PBC) is a rare, progressive, often fatal liver disease that occurs mainly in women. This study investigates whether the immunosuppressant methotrexate changes disease progression-free survival and overall survival rates for PBC patients. We analyzed data from a double-blind, placebo controlled trial of methotraxate and placebo, randomized by severity of liver disease. There was no statistically significant difference between methotrexate and placebo on either disease progression or survival at 3, 6, or 9 years after start of treatment. At the end of the observation period (9 years), patients on methotrexate were an absolute 8% more likely to show progression of disease; the 95% confidence interval suggests this result would be typical of a true difference ranging from a 4% decrease to 19% increase, and is not statistically significant (P=0.21). Similarly, the patients on methotrexate were on average an absolute 3% more likely to die, but this result would be typical of a true difference ranging from a 5% decrease to 10% increase in death, and is not statistically significant (P=0.50). The randomization across disease stage was not exactly balanced, with more subjects in the placebo group than in the methotrexate group showing stage 1 and stage 4 disease, but in a subgroup analysis disease stage did not appear to modify consistently the association of treatment type with survival or progression-free survival. There is thus no evidence of benefit from treatment with methotrexate.

Comment: data on ??? subjects (and I would usually give the number on each arm)

Comment: conducted at 11 geographical centers

Comment: with randomization stratified

Comment: how many subjects do you have at risk at these times? Some description of the length of follow-up might be helpful

Comment: Need to watch about choosing a time period with relatively little precision. How many patients could you have observed for a nine year period?

Comment: It would be useful to tell us what the survival probability was in each group. Is this 100% vs 92%; 8% vs 0%, or somewhere in between. This has ramification re the generalizability of the results

Comment: I was not particularly impressed with there being substantial imbalance in this regard, but this is in fact a judgement call.

Comment: what is this?

Background:

Primary biliary cirrhosis (PBC) is a rare, progressive, often fatal liver disease that occurs mainly in women. Symptoms of PBC overlap with those of other autoimmune diseases such as rheumatoid arthritis, Sjögren's syndrome, and CREST syndrome. Antimitochondrial and antinuclear antibodies, markers of autoimmune disease, are found in most patients with PBC (1). The only recommended treatment for PBC is ursodeoxycholic acid (UDCA), a synthetic bile acid, which has been shown to improve survival (2). Because of the autoimmune nature of this disease, it has been speculated that treatment with immunosuppressive agents such as methotrexate might decrease liver damage, slow progression, and increase survival. The current study aimed to determine difference in survival between subjects with PBC receiving UDCA plus either methotrexate or placebo.

Questions of Interest:

The initial questions were: (a) do subjects treated with methotrexate experience a different survival rate than those treated with a placebo?, and (b) do subjects treated with methotrexate experience a different progression-free survival rate than those treated with placebo?

Since death from PCB often occurs within five to 10 years, and since treatment with a medication might have different consequences at each of these stages, we reframed the question relative to survival at clinically meaningful time periods to ask: what are the differences in survival and progression-free survival based on treatment with methotrexate or placebo, as estimated at 3, 6, and 9 years?

Comment: would this estimate still hold after restricting study entry to those with fairly good laboratory values?

Sources of Data:

The study was a randomized, double-blind placebo controlled trial. Subjects were screened to determine the diagnosis of PBC using consensus guidelines, and excluded because of liver disease of another etiology, current immunosuppressive use, or other significant medical comorbidity. Initially, patients were treated with only UDCA and observed for 6 months. During this time, each participant underwent a liver biopsy for disease staging.

Subjects were randomized accounting for disease stage, so that subjects with early (stage 1-2) or advanced (stage 3-4) disease would be distributed evenly between the two treatment arms. One group received UDCA plus methotrexate (132 patients), and the other received UDCA plus placebo (133 patients). The methotrexate dose was adjusted to account for side effects. A record of medication use was kept, and unused medications were counted to monitor compliance.

At baseline, measures were recorded for age, sex, height, weight, time between initial PBC diagnosis and randomization, disease stage, the presence or absence of an enlarged spleen (splenomegaly), and laboratory measurements associated with liver function (blood serum measurements of bilirubin, albumin, alkaline phosphatase, ALT, and cholesterol; blood clotting (prothrombin) time, and platelet counts). There were two main outcome measures: time until death (or departure from study), and time until disease progression (or departure from study). Progression was defined by the development of worsening clinical disease (ascites, hepatic encephalopathy, bleeding varices), liver transplantation, or death, and progression-free survival was defined as the absence of any of these events.

Comment: Arguably, the outcomes were time until death or progression. The censoring is just a technical measurement problem.

Statistical Methods:

We first assessed the baseline characteristics of the two treatment groups. In order to produce a clinically useful measure of body habitus, we calculated body mass index (BMI) from height and weight. For those measures where the arithmetic mean might be skewed by outliers, we calculated geometric means. We compared baseline distributions of these variables across the treatment groups to verify that no clinically relevant differences existed, and to look for potential confounders.

Our primary analysis was to determine whether methotrexate and placebo groups differed in survival or progression-free survival at clinically meaningful time points. We evaluated survival and progression-free survival at three time periods corresponding roughly to early, middle, and late disease: 3 years, 6 years, and 9 years after the start of the treatment phase of the study. This design was to allow a distinction between short-, medium-, and long-term difference in treatment effect, which a single comparison of hazard of dying throughout the whole study (such as a Cox proportional hazards regression) would not allow.

Comment: Well motivated description. I do note that the PH regression would be the standard analysis, though perhaps not the best.

Since patients dropped out of the study in both groups, we do not have time to death or time to progression data on all patients. Hence, direct measurement of the percentage of patients who had not died or progressed at specific time-points was not an accurate measure of survival. We conducted a survival analysis using the Kaplan-Meier method to estimate the probability that subjects would survive or remain without progression of disease at the three time points. The Kaplan-Meier method accounts for missing data by assuming that subjects who dropped out would have been have been similar in outcome to all those who remained in the study at the time they dropped out. Using this approach, we compared survival estimates in the methotrexate and placebo groups at each time point by taking the absolute difference in estimated survival between the groups, and testing whether the observed difference was unequal to zero at a 95% confidence level. We calculated confidence intervals and P-values using the standard normal Z statistic.

Both because disease stage has clinical relevance with regard to survival and progression in PBC, and because unequal randomization by disease stage might confound the overall survival analysis, we conducted a subgroup analysis by histological disease stage. We compared survival and progression-free survival for the two treatment groups using the approach defined above, but stratified by histological stages 1 – 4.

Comment: Careful here. A “stratified analysis” merely adjusts for the third variable. A “subgroup analysis” might try to answer the question separately for each stratum. In assessing “effect modification”, we would be interested in “proving” that there might be differences in treatment effect across subgroups.

We analyzed the data using Stata version 9 (StataCorp, College Station, Texas).

Results

Baseline Characteristics

The 265 subjects admitted to the study were randomized so that 133 received UCDA plus placebo, and 132 received UCDA plus methotrexate. Table 1 shows the baseline characteristics of these groups. The treatment groups were reasonably equivalent in baseline demographic statistics as well as relevant laboratory measurements. While there were few males overall, they were evenly distributed across treatment groups (10 each). Treatment groups had slightly different mean ages at randomization, with the placebo group 1.8 years older. The placebo group had a slightly higher BMI – 27.32 vs 26.09. The placebo group had more Stage 1 patients than did the methotrexate group (18.1% vs 9.9%), and more stage 4 patients (16.5% vs 12.1%), but less stage 2 and stage 3 (30.1% vs 35.6% for stage 2; 35.3% vs 42.4% for stage 3).

Comment: Do you think either of these matter? That is, are such differences predictive of differential survival?

	Placebo		Methotrexate	
	Missing	Mean (SD)	Missing	Mean (SD)
Demographic				
Age (years)	0	52.19 (8.64)	0	50.38 (8.67)
% male	0	8.00%	0	8.00%
BMI (kg/m^2)	1	27.32 (5.41)	0	26.09 (4.67)
Disease history				
Disease duration (years) **	0	2.47 (1.08)	0	2.29 (1.08)
Histological Stage:				
1	0	18.1%	0	9.9%
2	0	30.1%	0	35.6%
3	0	35.3%	0	42.4%
4	0	16.5%	0	12.1%
Clinical findings				
Splenomegaly (% observed)	0	10.53%	1	8.40%
Laboratory				
Bilirubin (mg/dl) **	0	0.62 (1.05)	0	0.55 (1.06)
Albumin (g/dl)	0	4.00 (0.34)	0	4.00 (0.35)
Alk. Phosphatase (U/l) **	0	197.88 (1.06)	0	207.71 (1.05)
ALT (U/l) **	1	40.32 (1.06)	1	42.99 (1.06)
INR (sec)	0	11.37 (1.08)	4	11.24 (1.13)
Serum Chol. (mg/dl)	3	235.77 (58.81)	2	239.18 (58.18)
Platelets (1000/cc)	0	234.71 (83.18)	0	243.53 (88.60)

Comment: So are these means or geometric means? If geometric means, what do the SD mean?

*Table 1: Summary statistics at baseline for both placebo and methotrexate treatment groups. Measurements marked ** were compared using geometric means.*

Survival Analysis

The survival analysis is presented in Figure 1 and Table 2. The calculated differences are estimates of the differences in survival between the two groups, with inference based on the likelihood of finding by chance differences at least as extreme as those observed. These tests did not demonstrate any statistically significant differences at a 95% confidence level between the groups as measured at 3, 6, and 9 years, for either survival or progression-free survival. This involved a total of 6 comparisons, with no correction for multiple comparisons.

Comment: What was the length of follow-up on the patients? Did we follow most subjects for all 9 years?

Comment: Good to note.

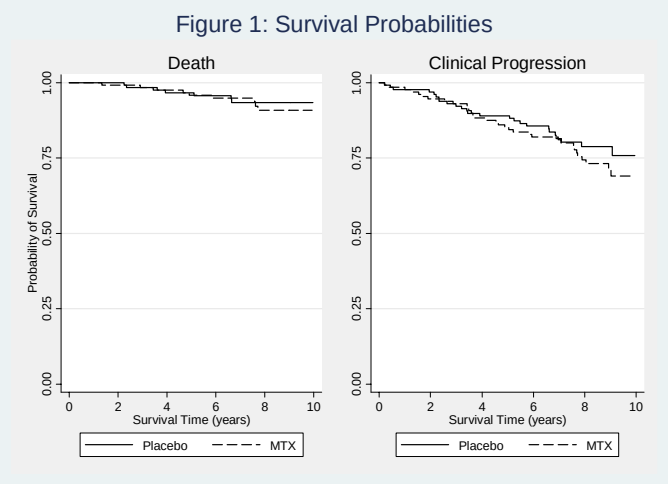

Figure 1: Probability of progression free survival (left) and of survival (right) by treatment group, estimated using Kaplan-Meier methods.

Progression Free Survival					
Years	**Group**	**Estimate (SD)**	**95% CI**		**P-value**
3	Placebo	92% (2%)	86%	96%	
	MTX	93% (2%)	87%	96%	
	Difference	1% (3%)	-6%	7%	0.79
6	Placebo	86% (3%)	78%	91%	
	MTX	82% (3%)	74%	88%	
	Difference	-4% (5%)	-13%	5%	0.43
9	Placebo	79% (4%)	70%	85%	
	MTX	71% (5%)	61%	79%	
	Difference	-8% (6%)	-19%	4%	0.21
Survival					
Years	**Group**	**Estimate**	**95% CI**		**P-value**
3	Placebo	98% (1%)	94%	100%	
	MTX	98% (1%)	94%	100%	
	Difference	0% (2%)	-3%	3%	0.99
6	Placebo	96% (2%)	90%	98%	
	MTX	95% (2%)	89%	98%	
	Difference	-1% (3%)	-6%	5%	0.75
9	Placebo	93% (2%)	87%	97%	
	MTX	91% (3%)	83%	95%	
	Difference	-3% (4%)	-10%	5%	0.50

Table 2: Progression free survival and survival at 3, 6, and 9 years by treatment group, estimated using Kaplan-Meier techniques.

Subgroup Analysis:
Table 3 represents the subgroup analysis, with survival and progression-free survival estimated for each of the histological stages at baseline. No significant differences between the methotrexate and placebo groups were seen except that among subjects classified at baseline with Stage 4 disease the methotrexate group had a higher survival at 6 years than the placebo group (100% vs 79%, p = 0.024, 95% CI 2.7% - 27.7%). This was not seen in the Stage 4 group at 3 or 9 years.

		PLACEBO	METHOTREXATE				
Stage	**Yrs**	% Without Progression (SD)	% Without Progression (SD)	Estimated Difference (SD)	P	95% CI of Difference	
1	**3**	95.5% (4.4%)	91.7% (8.0%)	-3.8% (9.1%)	0.679	-21.7%	27.0%
	6	95.5% (4.4%)	91.7% (8.0%)	-3.8% (9.1%)	0.679	-21.7%	27.0%
	9	95.5% (4.4%)	80.2% (12.8%)	-15.2% (13.5%)	0.260	-41.8%	40.1%
2	**3**	100.0% (0.0%)	97.8% (2.2%)	-2.2% (2.2%)	0.313	-6.5%	6.5%
	6	94.7% (3.7%)	88.7% (4.8%)	-6.0% (6.0%)	0.322	-17.8%	17.8%
	9	91.4% (4.8%)	78.6% (8.2%)	-12.8% (9.5%)	0.179	-31.5%	28.2%
3	**3**	86.6% (5.1%)	94.6% (3.0%)	8.0% (5.9%)	0.175	-3.6%	17.5%
	6	79.7% (6.1%)	80.3% (5.3%)	0.6% (8.1%)	0.939	-15.2%	23.8%
	9	73.6% (7.0%)	70.5% (6.6%)	-3.0% (9.6%)	0.753	-21.8%	28.4%
4	**3**	85.9% (7.6%)	75.0% (10.8%)	-10.9% (13.2%)	0.409	-36.8%	39.1%
	6	71.6% (9.8%)	62.5% (12.1%)	-9.1% (15.6%)	0.560	-39.6%	46.1%
	9	50.6% (12.5%)	46.9% (13.2%)	-3.8% (18.2%)	0.837	-39.4%	53.9%

		PLACEBO	METHOTREXATE				
Stage	**Yrs**	% Surviving (SD)	% Surviving (SD)	Estimated Difference (SD)	P	95% CI of Difference	
1	**3**	100.0% (0.0%)	91.7% (8.0%)	-8.3% (8.0%)	0.297	-24.0%	23.6%
	6	100.0% (0.0%)	91.7% (8.0%)	-8.3% (8.0%)	0.297	-24.0%	23.6%
	9	100.0% (0.0%)	80.2% (12.8%)	-19.8% (12.8%)	0.122	-44.9%	37.9%
2	**3**	100.0% (0.0%)	97.8% (2.2%)	-2.2% (2.2%)	0.313	-6.5%	6.5%
	6	100.0% (0.0%)	95.4% (3.2%)	-4.6% (3.2%)	0.148	-10.9%	9.4%
	9	96.6% (3.4%)	95.4% (3.2%)	-1.2% (4.7%)	0.803	-10.3%	13.8%
3	**3**	97.6% (2.4%)	100.0% (0%)	2.4% (2.4%)	0.311	-2.2%	7.0%
	6	97.6% (2.4%)	94.1% (3.3%)	-3.6% (4.1%)	0.384	-11.5%	12.1%
	9	94.0% (4.2%)	90.8% (4.5%)	-3.2% (6.2%)	0.607	-15.3%	18.3%
4	**3**	94.7% (5.1%)	100.0% (0.0%)	5.3% (5.1%)	0.304	-4.8%	15.2%
	6	79.0% (9.4%)	100.0% (0.0%)	21.1% (9.4%)	0.024	2.7%	27.7%
	9	79.0% (9.4%)	87.5% (11.7%)	8.6% (15.0%)	0.568	-20.8%	44.3%

Table 3: Progression-Free Survival and Survival by Baseline Disease Stage

Discussion:

Methods:
A Kaplan-Meier survival analysis assumes that subjects who left the study early, and whose survival or death was not observed, were similar to subjects who remained in the study at that point (this is known as "non-informative censoring"). There were 26 such subjects after 3 years (15 on placebo, 11 on methotrexate), and 54 subjects dropped out by year 6 (29 from the placebo group, 25 on methotrexate).

After 9 years, 181 subjects had dropped out; combined with 16 deaths; this left 68 subjects in observation after 9 years. Subjects may have dropped out for reasons material to the analysis performed, such as becoming more ill or more healthy. We examined the subjects who dropped, and found that they did not differ in material ways from other subjects in their baseline measurements and sociodemographic characteristics, but this does not fully determine that censoring was not informative, since the comparison should be made to the subset of subjects remaining in the study at time of censoring. We did not have data available to do this, so for our analysis we assumed non-informative censoring.

In addition to censoring, some patients in each group stopped taking medication during the study but remained available for follow up measurements. After 3 years, there were 9 such subjects in the placebo group, and 12 in the methotrexate group; after 6 years, a there were a total of 29 from the placebo group and 27 from the methotrexate group. For the main analysis of survival we retained these subjects, conducting an intent-to-treat analysis without adjusting for the time the subjects took the drug. A meaningful analysis using true time on methotrexate, rather than simply methotrexate group, would require additional data about medication starting and stopping, dose, and dose adjustments.

Baseline Characteristics:

The two groups were well-matched at baseline for all the measured characteristics except for stage. There were slight differences in BMI and age, but we did not consider these to be sufficiently causally related to survival or progression-free survival for them to qualify as potential confounders. The placebo group had more stage 1 and stage 4 patients, and the methotrexate group had more stage 2 and 3 patients. This might be expected because the methotrexate and placebo groups were randomized by stage category, not absolute stage (1/2 vs. 3/4). We considered this to be a possible source of confounding. For instance, if the stage 1 patients had much improved survival, then the group with more of these patients might be expected to have a greater survival estimate. This was the motivation for performing a subgroup analysis.

Main Result:

None of the three time points showed a significant difference between methotrexate and placebo with regard to survival and progression-free survival. This was a consistent result in both analyses and at all time periods. No correction was made for multiple comparisons, but we determined such a correction was not necessary in the setting of no results that demonstrated a statistically significant difference.

Subgroup Analysis:

The groups were not well-randomized by the 4 separate stage categories, especially Stage 1. Less severe disease at baseline might be expected confer a survival advantage, as described above. We thus conducted a subgroup analysis of survival and progression-free survival based on disease stage. The only significant difference that appeared in this analysis was that patients in the methotrexate group with histological stage 4 showed a greater survival estimate at 6 years than patients in placebo group (100% vs. 79%, p = 0.024, 95% CI 2.7% - 27.7%). This amounted to 3 of 22 patients in the placebo group and zero of 16 patients in the methotrexate group who had died by that time. This finding was in the setting of other multiple comparisons, since a total of 12 comparisons were made for the subgroup analysis, and we did not apply a formal correction. Even correcting for 3 comparisons with a Bonferroni correction would result in the finding being considered non-significant at a 95% confidence level. Several other observations argue against this being an important finding. First, we would expect to see a similar difference in survival for earlier-stage patients when they are later in the disease (e.g. stage 3 patients at 9 years), but this was not observed. Second, no difference was observed in the stage 4 group at 3 or 9 years. Third, the sample sizes were small for this subgroup. From this single finding, therefore, it would be quite imprudent to conclude that methotrexate confers improved survival in the mid-time range for those with stage 4 disease

Conclusions:

In our analysis of patients with PBC treated with UDCA plus either methotrexate or placebo, there was no evidence of increased survival or progression-free survival with the drug at early (3 years), middle (6

Comment: Are these KM estimates of dropout rates? They should be. That is, differential rates of progression or death would mean that fewer subjects might be observed to dropout on one arm. Usin KM estimates protects you in this regard.

Comment: Turns out few subjects had been enrolled more than 9 years prior to the time of data analysis. This is often called “administrative censoring”.

Comment: Ideally, some aspect of compliance would be presented in the Results section. Given the data you had, I might have considered the PH regression on time to d/c of study drug. This is subject to the competing risk of death, however, so this is at best a descriptive analysis, and prett vague at that.

Comment: And even that would not be very meaningful, for the reasons given above, as well as due to the possibility that toxicity precluded taking the drug further.

Comment: Your wording seems to suggest you were more worried about confounding than effect modification. In that case, you would have wanted to combine survival estimates across stage groups, rather than presenting all the different analyses. This would have been somewhat involved, but not impossible. Probably combining on the odds scale would have been best. Alternatively, going with an adjusted PH analysis might be easiest. This latter approach would only be appropriate, however, if you were satisfied that a HR would capture your concept of an important summary measure of treatment effect.

years), or late (9 years) time periods. A subgroup analysis by baseline disease stage did not show any trends to increased survival in the group treated with methotrexate, except in Stage 4 patients at 6 years, but this finding is tempered by a rather small sample, multiple comparisons, no similar finding in other groups, and no difference between treatment groups at 9 years for these same patients. Further work could attempt to determine if methotrexate might confer increased survival in later-stage patients, although the current study implies that methotrexate is most likely an ineffective treatment to modify progression in PBC, and that its use should not be recommended.

Comment: Very good discussion of an exploratory analysis.

[1] Kaplan, MM. Primary biliary cirrhosis. *N Engl J Med* 1996; 335:1570

[2] Poupon RE, Lindor KD, Cauch-Dudek K, Dickson ER, Poupon R, Heathcote EJ. Combined analysis of randomized controlled trials of ursodeoxycholic acid in primary biliary cirrhosis. *Gastroentorology* 1997;113(3):884-890.