No Reduction in Diffusion Capacity after treatment with Methotrexate in Primary Biliary Cirrhosis

Abstract:

Context: Methotrexate (MTX) has been used in the treatment of Primary Biliary Cirrhosis (PBC) but unclear of its effect on diffusion capacity.

Objective: To assess the effect of methotrexate on diffusion capacity in treatment of subjects with primary biliary cirrhosis.

Design, Settings, and Patients: The study is a randomized, double blind, placebo controlled clinical trial involving patients with primary biliary cirrhosis from 12 clinical centers in the U.S.A. Between January 1994 to March 1998, a total of 265 patients ages 20 to 69 years old with primary biliary cirrhosis were enrolled.

Interventions: Patients were randomly assigned to receive ursodeoxycholic acid (UDCA) plus MTX (n=132) or UDCA plus placebo (n=133).

Main Outcome Measures: Diffusion capacity (DLCO) at baseline and last recorded DLCO for each subject was assessed.

Results: At baseline, comparing the demographic factors the two groups of subjects appear to be well matched. Stratifying by height, which is directly related to vital capacity and therefore DLCO, there was no statistical difference when comparing DLCO difference in height <162cm (p=0.49 95% CI -0.89, 1.84) and height ≥162cm (p=0.47, 95% CI -2.12, 0.98). There was no statistical difference between mean baseline DLCO and final DLCO when stratified by months on study drug: ≤60 months (p=0.88, 95% CI -2.23, 2.59), 60-75 months (p=0.89 95% CI-1.88, 2.16), 75-95 months (p=0.96 95% CI -1.94, 2.04) and ≥95 months (p=0.72, 95% CI -2.34, 1.63). No statistical significant change in DLCO when comparing groups by stage of liver disease; Stage 1(p=0.44, 95% CI -2.37, 5.20) Stage 2 (p=0.3, 95% CI -0.84, 2.73) Stage 3 (p=0.99, 95% CI -1.74, 1.73) and Stage 4 (p=0.34, 95% CI -4.15,1.48). Using a >20% decrease in DLCO between baseline and final as clinical significant, there was no significant difference between treatment groups (p=0.73, 95% CI -0.09, 0.13)

Conclusions: Treatment with methotrexate does not impair diffusion capacity in patients with primary biliary cirrhosis.

Comment: This is a potential toxicity. You would want to make that clear.

Also "diffusion capacity" would not be at all clear. We are dealing with potential lung toxicity.

Comment: Careful here. You are doing a subgroup analysis, and losing a lot of power. If you wanted to stratify, you should have done some sort of average across strata. Also, you needed to have presented a point estimate. And given units. And tell whether you subtracted MTX from Plc or vice versa.

Comment: You should still be comparing the treatment groups. Were you?

Comment: Why is this of interest? Does liver function affect lung function?

Comment: What are your point estimates?

Comment: Did we have sufficient precision to rule out all meaningful difference? For instance would a 13% absolute increase in the number having a 20% decrease in DLCO be meaningful?

Introduction:

Primary Biliary Cirrhosis (PBC) is an auto-immune disease of the liver where the body's own immune system causes damage to intrahepatic bile ducts and leads to eventual blockage. Blockage of bile ducts leads to damage to liver cells and impairment of liver function, including protein synthesis, carbohydrate metabolism, fat metabolism and detoxification of toxins. In advanced stages this is known as cirrhosis. As with most auto-immune disease, PBC occurs more frequently in women than men and tends to affect people in the fourth to seventh decade of life. In some patients, they may remain asymptomatic but for patients who do become symptomatic, death usually occurs within 5-10 years of initial diagnosis.

Methotrexate is a structural analogue of folate and competitively inhibits dihydrofolatereductase (DHFR). Methotrexate inhibits the synthesis of purine nucleotides and therefore prevents DNA and RNA synthesis. Methotrexate is mostly common used as an anti-neoplastic and immunosuppressive agent. Methotrexate can be used in low, moderate or high doses. Toxicities associated with methotrexate use include bone marrow suppression, dermatitis, mucositis and interstitial lung disease. Interstitial lung disease is a rare and serious toxicity seen with methotrexate use, diagnosed with pulmonary function tests and specifically, diffusion specific capacity (DLCO). DLCO measures the ability of gases to diffuse from the alveoli into the capillaries. In subjects with interstitial lung disease, their DLCO are reduced.

In this study, we studied the effect of methotrexate in subjects with primary biliary cirrhosis and the effect on diffusion capacity. Our hypothesis is that subjects on methotrexate will have decreased diffusion capacity compared to the placebo group.

Comment: So here you noted it is a toxicity we were looking at.

Methods:

Patient Selection:

The study was a randomized, double-blind, placebo controlled clinical trial conducted at 12 clinical centers in the U.S.A. Patients were selected from subjects ages 20-69 years old with moderately advanced liver disease at entry. To document sufficiently advanced disease, patients have had a diagnosis

of chronic cholestatic liver disease of at least 6 months duration, on ursodeoxycholic acid (UDCA) for at least 6 months, documented positive anti-mitochondrial antibody test and alkaline phosphatase levels of at least 1.5 times the upper limit of normal and a liver biopsy within 6 months prior to randomization with histological findings compatible with diagnosis of PBC. Asymptomatic patients must have histologic stage greater than Stage I using Ludwig classification. Exclusion criteria for the study included: advanced PBC (defined as history of bilirubin greater than 3 mg, serum albumin less than 3 mg, or history of ascites, hepatic encephalopathy or variceal bleeding), liver disease of other etiology, history of alcohol abuse, history of treatment with immunosuppressive agents, rifampin or dilatin, history of major illness that limit lifespan including HIV, malignant diseases, pregnancy or unwillingness to use birth control.

Randomization was stratified by the Ludwig histologic stage of disease and as read by the pathologist at the respective centers. Of the 126 reported to be stage 1-2, 62 randomized to methotrexate group and 64 to receive placebo while among the 139 reported as stage 3-4. 70 randomized to methotrexate group and 69 to the placebo group.

Drug Treatments:

All patients received UDCA in 300mg capsule taken orally at bedtime. Methotrexate or placebo were provided as 2.5mg tablets and administered at bedtime once a week. The initial dose was half the maximum dose and increased every month by 2.5 mg to a maximum dose of 15mg per 1.73 m^2 BSA, with a maximum dose of 20mg per week. Subjects taking cholestyramine or colestipol were asked to take medications at least 2 hours before or after UDCA, methotrexate or placebo. Patient continued in their treatment groups until closure of study, death, liver transplantation, drug toxicity induced withdrawal, development of cancer or voluntary withdrawal.

Modification of Methotrexate dose:

No current evidence that UDCA affects bone marrow or any other significant side effects. Therefore, the development of myelosuppression, mucositis, nausea, or anorexia was considered to be related to methotrexate and the methotrexate dose was modified based on the severity of toxicity rating: mild (acceptable-no change), moderate (requires dose change) or severe (requires discontinuation of therapy). Moderate toxicity decrease weekly dose by a quarter or third and monitored weekly until resolved. The dose of methotrexate as increase by 2.5 mg until dose was 2.5 mg less than the original toxic dose and that symptoms did not reappear. For severe toxicity, the methotrexate was stopped during management of symptoms. Once resolved, methotrexate was restarted at half toxic dose and increased by 2.5 mg per week at monthly intervals if no toxicity recurred to a weekly dose of 2.5 mg less than original toxic dose. If severe toxicity did not improve within 1-2 weeks, leucovorin was administered at a dose of 5 mg PO or IV every 12 hours for at least 48 hours. Other reasons to stop decrease or stop methotrexate include: allergic symptoms, severe skin rash, pulmonary symptoms consistent with pulmonary fibrosis, severe liver exacerbation, worsening renal function, development of cancer, subject became pregnant or unwilling to use birth control.

Comment: This, of course, is the most relevant to this paper

Schedule of Patient Visits and Investigations:

Subjects were seen and blood drawn at weeks 2 and 4, then monthly for first 6 months, bimonthly for next 6 months and then every 3 months for the rest of the study. Blood was drawn the day of the next dose, before dose was given. At each visit, symptoms of liver disease and toxicity were obtained. At each visit, CBC was obtained; monthly visit ALT, AST and alkaline phosphatase; at 3 month intervals total protein and albumin and at 6 month intervals prothrombin time. Complete history, physical examinations, chest x-rays and pulmonary function tests including diffusion capacity were performed annually. Liver biopsy and upper endoscopy were performed after 2 years and every 2 year intervals.

Comment: And this of most relevance for the availability of the data.

Subjects were given a known quantity of medicine and compliance was checked by monitoring their logs and unused medication counted. Throughout this paper, “baseline DLCO” will refer to the earliest pre-dosing measurement made on each subject, and “final DLCO” will refer to the last available DLCO measurement made on each subject before discontinuation of study drug, regardless if discontinuation is due to withdrawal from study or administrative censoring.
Subjects reported adverse effects and grouped into categories defined by organ system.

Monitoring of Clinical Trial:

The data was monitored on semi-annual basis by an independent Data Safety and Monitoring

Board (DSMB). The data was reviewed by DSMB for safety as well as making recommendations for early termination of trials.

Statistical Methods:

To make statistical comparisons between our methotrexate and placebo samples as described below, we utilize Student's T-test on DLCO measurements while assuming inequality of variances between groups. This is a method of comparing means of distributions of the measurement in question between groups in relation to their variances, allowing us to reject the standard null hypothesis (that the groups are similar) or fail to reject this hypothesis and conclude that we do not have enough evidence to show statistically significant differences between groups. It must be noted that the loose assumptions required to use the T-test (approximate normality of the distribution of the measurement in question, and independence of our two samples) are satisfied in each scenario detailed below, and the test is satisfactorily powerful for samples sizes considered here.

We will consider three DLCO measurements: final DLCO, difference from baseline DLCO to final DLCO, and minimum DLCO while on study drug. In order to obtain a complete assessment of methotrexate safety as it relates to lung function, we consider potential factors on DLCO measurements by creating a number of subgroups in which DLCO is hypothesized to have a significant effect. These subgroups include groups defined by size (weight and height, which are both strongly correlated with lung function), time on study drug, stage of liver disease, age, and gender.

Statistical computations were performed with Intercooled Stata Version 9.1 (College Station, TX).

Comment: This is not a requirement. Instead we need the sample means to be approximately normal.

Comment: I think I disagree on two counts: First, we have a small sample size. Second, the way you broke the analysis into subgroups decreased the power even more.

Comment: Say, what? We would be interested in subgroups that would be likely to have a difference in DLCO, not DLCO having an effect.

Comment: Why? Tell us why you would want to look at this.

Statistical Results:

At baseline, it appears that our sample of patients randomized to receive methotrexate are similar in all pertinent demographic factors to the placebo group, so we are confident that none of these factors contribute to confounding in this randomized controlled trial. See Table 1. Note that we have some missing data (1missing DLCO observation at baseline for a patient on methotrexate; 8 placebo patients and 12 methotrexate patients with missing observations both final DLCO and minimum DLCO while on study drug,), but it does not appear to create a problem with our analysis. For two-sample tests for which a difference between baseline and final DLCO is required, and at least one of the measurements is missing, this patient is excluded from the statistical test. The distribution of demographic characteristics among missing patients was similar to those for patients with complete data.

Comment: Good to note. Probably better to put this sentence in Stat Methods

Table 1: Sample demographic characteristics and Laboratory Measurements

	Patients (N, %)	
	Placebo (133, 50.2%)	***Methotrexate (132, 49.8%)***
	Mean (SD), min-max	***Mean (SD), min-max***
Age (years)	52.19 (8.64), 25-67	50.38 (8.67), 31-70
Female (%)	92.48	92.42
Height (cm)	163.69 (8.07), 149.5-191.9	163.71 (7.92), 142.8-181.6
Weight (kg)	73.35 (16.17), 46.0-149.1	70.12 (14.42), 42.7-114.2
Laboratory Measurements		
Total Bilirubin (mg/dl)	0.72 (.39), 0.1-2.3	0.66 (.44), -0.1-2.8
Albumin (g/dl)	4.00 (.34), 3.1-4.9	4.00 (.35), 2.0-5.9
Alkaline Phosphatase (U/l)	244.93 (187.44), 63-1127	242.85 (145.67), 50-930
ALT (U/l)	50.17 (41.61), 9-199	54.29 (41.43), 9-314
Prothrombin Time (sec)	11.37 (1.08), 8.6-13.9	11.24 (1.13), 8.5-14.6

Serum Cholesterol (mg/dl)	235.77 (58.81), 128-560	239.18 (58.18), 137-472
Platelets (1000 cells / cu mm)	234.71 (83.18), 77-561	243.53 (88.60), 86-619
Duration of Study Drug Administration (months)	69.8 (28.2),	71.6 (30.1),
	Stage of Liver Disease (%)	
Stage 1	18.1	9.9
Stage 2	30.1	35.6
Stage 3	35.3	42.4
Stage 4	16.5	12.1
	Splenomegaly (%)	
Present	89.47	90.91
Absent	10.53	8.33
Unknown	0	0.76

Since this clinical trial was designed to study survival in patients with primary biliary cirrhosis, we have a number of laboratory measurements made at baseline. Summary measures are presented in Table 1. None of these laboratory measurements were distributed differently amongst treatment groups. Some outlying values exist in the upper tails of the distribution of the lab measurements on alkaline phosphatase, serum cholesterol, and platelet count, but they appear to affect both treatment groups approximately equally, and do not appear to pose a serious problem, so we will include these patients in our analysis.

Comment: This is an absolute necessity. Deleting outliers is a very bad thing to do in this situation. You should not even be considering anything but including those patients in your analysis.

Comment: It would have been interest to see the minimum and maximum, because individual patients could have had toxicities even if the group did not on average

Table 2: DLCO as primary outcome

	Patients (N, %)			
	Placebo (133, 50.2%)	***Methotrexate (132, 49.8%)***	***95% CI for Difference in Means***	***Two-sided p-value for test of difference from 0***
	Mean (SD)	***Mean (SD)***		
DLCO at Baseline (ml/min/mmHg)	19.74 (5.02)	20.29 (5.05)	(-1.775, 0.667)	0.3723
DLCO Minimum on Study Drug (ml/min/mmHg)	16.71 (4.17)	16.96 (4.37)	(-1.325, 0.8262)	0.6485
DLCO at patient-specific study end date (ml/min/mmHg)	18.37 (4.96)	18.70 (4.81)	(-1.560, 0.9027)	0.5994
DLCO at baseline minus final DLCO (ml/min/mmHg)	1.51 (4.11)	1.59 (4.09)	(-1.117, 0.9531)	0.8760
	N (%)	***N (%)***	***95% CI for Difference in proportions***	***Two-sided p-value for test of difference in proportions of 0***
Proportion of patients exhibiting a 20% drop in DLCO from baseline to final	35 (28%)	31 (26%)	(-0.0919, 0.1309)	0.7318

First, we compare our three primary measures of DLCO in these patients. No significant differences in mean DLCO between treatment groups were found at baseline [95% CI (-1.775, 0.667), p-value 0.3723]. From our 95% confidence interval, we can say that our results are consistent with a difference in mean DLCO at baseline of -1.775 to 0.667. Throughout, p-values and confidence intervals are interpreted similarly. No significant differences were exhibited in the minimum DLCO while on study drug [95% CI (-1.325, 0.8262), p-value 0.6485] or at the patient-specific study end date [95% CI (-1.560, 0.9027), p-value 0.5994]. This is illustrated graphically in Figure 1.

Comment: This is of no interest in a RCT. By randomization the populations were equivalent at baseline. It is of interest to see the descriptive statistics, just don't do a p value and CI

Not unexpectedly, the measurements of DLCO at baseline are highly correlated with those at the final examination date (α= 0.659), so we expect to see patients with low DLCO at baseline exhibit low DLCO at final examination, and those with high DLCO at baseline exhibit a high DLCO at final examination. Hence, we must take this into account when comparing samples. We first calculate the patient-specific difference in DLCO from baseline to final, and then utilize the two-sample T-test to compare these samples of differences with respect to their means. We find no significant difference (CI for difference of mean methotrexate difference minus mean placebo difference (-1.117, 0.9531), two-sided p-value 0.8760).

Comment: Give point estimates!!! Also, this is the best summary of your results. It should have been in your abstract instead of the subgroup analyses.

In order to consider whether there is toxicity related to methotrexate causing impaired lung function, as measured by DLCO, we must consider a variety of aspects that may contribute to this adverse outcome. Methotrexate is being administered to these patients at levels lower than that which is known to cause toxicity yielding impairment of lung function, yet these patients are on study drug for an average of over 5 years (69.8 months in the placebo arm, 71.6 months in the methotrexate arm, or 5.8 years and 6.0 years, respectively). Thus, we are concerned with buildup of methotrexate that leads to toxicity.

Comment: You could have used this for your primary analysis as well. The key point: Donn't do the subgroups first, which is what you did in the absrtract.

Comparisons between placebo and methotrexate at 4 levels of on-drug experience (<60 months, 60-75 months, 75-95 months, > 95 months) show no statistically significant differences between the placebo arm and the methotrexate arm in the final DLCO measurements made on each patient, even without controlling our Type-I error rate for the multiple comparisons we have made. See Table 3 (A). The same lack of statistical significance is found when considering stages of liver disease separately. See Table 3 (B).

Comment: This was a good thing in concept. Of course, you are comparing the final DLCO for patients who received this much drug. While we do have data for subjects at 4 years even if they took drug for 7 years, say, you did not have that data.

Table 3: Comparison of mean difference of Baseline DLCO minus Final DLCO by Treatment Group

Strata	***Patients (N)***		***Mean of difference***	***95% CI for mean baseline to final DLCO difference in Placebo group minus mean final to baseline DLCO difference in methotrexate group (ml/min/mmHg)***	***Two-sided p-value for test of difference from 0***
	Placebo	***Methotrexate***			
			(A) Months on Study Drug		
≤ 60	24	20	0.1775	(-2.233, 2.588)	0.8822
60 – 75	38	28	0.1417	(-1.876, 2.159)	0.8887
75 – 95	33	37	.04906	(-1.944, 2.042)	0.9610
≥ 95	30	34	-0.3520	(-2.335, 1.631)	0.7239
			(B) Stage of Liver Disease		
Stage 1	24	11	1.411	(-2.373, 5.195)	0.4354
Stage 2	38	42	0.9442	(-0.8405, 2.729)	0.2952
Stage 3	45	54	-0.009629	(-1.744, 1.725)	0.9912
Stage 4	18	12	-1.333	(-4.146, 1.479)	0.3364
			(C) Height Strata		
<162 cm	57	58	0.4783	(-0.888, 1.8447)	0.4888
≥162 cm	68	61	-0.5760	(-2.118, 0.9841)	0.4685
			(D) Weight Strata		
< 70 kg	62	66	0.01716	(-1.270, 1.304)	0.9790
≥ 70 kg	62	53	0.1358	(-1.545, 1.816)	0.8731

Comment: Randomization was stratified by stage. For that reason a stratified (not subgroup) analysis would have been preferred by some. But you did not motivate why we would suspect worse lung function with worse liver disease. I am OK with month on drug and height, for the reasons you mentioned. You did not motivate weight subgroups, but most people can usually see why weight is of interest in a drug trial.)

Alternatively, we entertain the possibility that age might play a role in DLCO toxicity from methotrexate. It is not unreasonable to assume that older patients have decreased lung function, as measured by DLCO, than young patients. We choose an age cutoff that is appropriate to study this possible relationship, 50 years. We have 42% of individuals age <50 years at baseline, and 58% age greater than 50 years at baseline. Analysis reveals that the less than 50 year old group does not show a significant decrease in DLCO from baseline to final examination date [95% CI for mean placebo minus mean methotrexate (-1.077, 2.094), p-value 0.5257]. The same is true for the greater than 50 year old group [95% CI for mean placebo minus mean methotrexate (-1.472, 1.370), p-value 0.9438], indicating a lack of evidence to prove that DLCO difference from baseline to final is affected by age at baseline.

When stratified by gender, tests of final DLCO in placebo patients minus final DLCO in methotrexate patients show that these differences are not significantly different from zero [95% CI for mean difference (-7.478, 0.5672), p-value 0.4638 for males, (-0.973, 1.1155), p-value 0.8931 for females].

Dosing amounts were ultimately variable by body surface area, so we do not expect the data to exhibit differences in mean DLCO between when stratified by height category, as all patients were scheduled to receive the same dose of methotrexate or placebo, proportional to their size. For verification, we examine the t-test for differences between baseline and final DLCO between treatment categories presented in Table 3 (C). We see that the mean difference between treatment groups in the less than 162 cm group is 0.4783 ml/min/mmHg with a 95%CI of (-0.888, 1.8447), p-value 0.4888. From Table 3 (C), we see that the mean difference between treatment groups in the greater than 162 cm group is -0.5760 ml/min/mmHg, with a 95% CI of (-2.118, 0.9841), p-value 0.4685. This relationship is confirmed by examining tests of mean baseline to final DLCO difference conducted in two weight categories of less than 70 kg, and 70 kg and above, which is presented in table 3 (D). Hence, we can conclude that there is no significant effect on mean difference of DLCO from baseline to final between treatment groups by either height or weight.

It is also of interest to investigate whether DLCO in patients with low baseline DLCO are significantly impacted by treatment with methotrexate, as these patients might be considered "at risk" for increased toxicity. We do so by testing for a difference between treatment arms in the difference between DLCO at baseline and at final in the patients who achieve the lowest minimum DLCO while on study drug. Here, we consider the patients with lowest 5% DLCO, DLCO < 11 ml/min/mmHg, in each treatment group. There are 8 placebo patients with minimum DLCO <11 ml/min/mmHg, and 7 such methotrexate patients. The 95% CI for the mean difference from baseline to final examination in DLCO (placebo minus methotrexate) is (-4.898, 4.7731), with a two-sided p-value of 0.9796. For the remainder of patients (DLCO ≥11 ml/min/mmHg) the 95% CI is (-1.347, 0.7908), with a two-sided p-value of 0.8494.

DLCO are abnormal when below 80% of the predicted value. Without predicted DLCO values, we assumed that the subjects had normal DLCO prior to the study and therefore a decline in DLCO by 20% would be significant and abnormal. When comparing the proportion of patients exhibiting such a drop between treatment groups, we find 35 (28%) such patients in the placebo arm, and 31 (26%) such in patients in the methotrexate arm. Thus we obtain a 95% CI for the difference in proportions of (-0.09192, 0.1309) and a two-sided p-value of 0.7318. We conclude that there is not a statistically significant difference between the proportion of patients exhibiting a significant drop in DLCO in the methotrexate group and the placebo group.

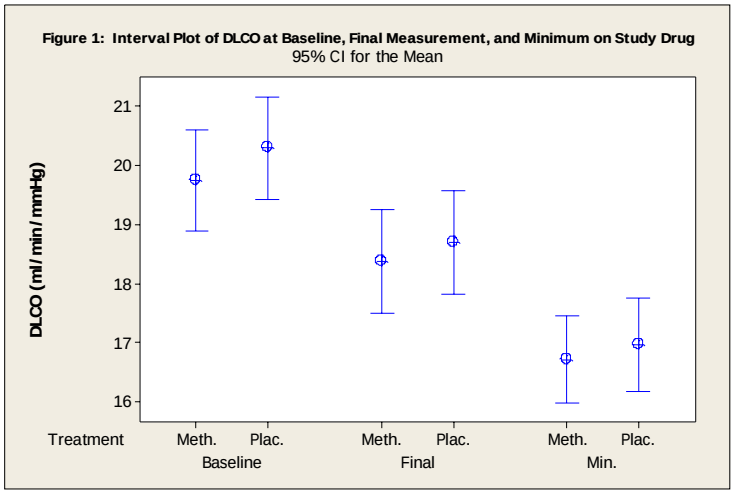
Comment: Okay to explore all these subgroups. Just recognize that you lose substantial power as you get a smaller and smaller sample size.

Comment: This is a very good thing to look at

Comment: Why no point estimate?

Comment: I don't think you had to invoke that patients were normal at baseline. The drop of 20% is still of interest.

Comment: Go ahead and give the point estimate too



Discussion:

Because high methotrexate concentrations have documented lung toxicity, the patients' lung function was also monitored in the duration of the study. We describe lung function as DLCO, which is a measurement of diffusion of CO in the lungs from air to blood. Low DLCO is correlated with interstitial lung disease. To determine significant differences between the placebo and methotrexate groups, we chose to compare means of DLCO and employ T-tests to determine statistical significance. From the statistical analysis, we found that our hypothesis, that the methotrexate group would have lower DLCO than the placebo group, is not supported. Instead, we discovered that there appears to be no statistical difference between the two groups. We compared the mean baseline and final DLCO of the methotrexate and placebo groups but found no significant difference. The baseline DLCO measurements were 19.74 ml/min/mmHg and 20.29 ml/min/mmHg and the final DLCO measurements were 16.71 ml/min/mmHg and 16.96 ml/min/mmHg for the placebo and methotrexate groups, respectively. To determine if a possible statistical difference occurred in the duration of the study, we also investigated the difference in mean differences at different time intervals of the study between the two groups (baseline DCLO - final DLCO). According to table 3, there appeared to be no statistical difference between the two groups at the times ≤60 months, 60-75 months, 75-95 months or ≥ 95 months. These observations suggest that the experimental methotrexate dosages were not sufficient enough to affect patient DLCO.

Comment: Is "no statistical difference" also "no clinically important difference"?

To elucidate possible factors affecting DLCO, we stratified the groups based on age, sex, weight and phase of liver disease. Because DLCO is proportional to lung size, it is important to account for variables that would affect body size (height, weight and gender), as larger people tend to have larger lungs and would consequently have a higher DLCO. In table 1, we investigated the demographics of the sample population and found very little difference between the methotrexate and placebo groups, which allowed us to continue with our analysis without concern for confounding due to these factors. Comparison by gender would not have been very compelling as the sample of males was be too small to make accurate gender specific inferences (n =9). However, we find that the proportion of females in each group is approximately equal; therefore, allowing comparison without concern for confounding due to imbalance of gender. Phase of liver disease and age may affect DLCO due to effect of decreased general health. Stage of liver disease may specifically affect DLCO when patients with severe liver disease can develop a condition called hepatopulmonary syndrome (HPS). HPS is has been typically described in patients with severe liver disease, awaiting liver tranplantation. Hepatopulomonary syndrome is a condition characterized by the development of intrapulmonary vacular dilatations, leading to right to left shunting of blood. Patients typically present with hypoxemia and complain of dysnea, platypnea and orthopnea. DLCO will typically be decreased in patients with HPS. However, we have no clinical data that leads us to believe that any of subjects had HPS. There seemed to be no statistically significant differences between the placebo and methotrexate groups in each division, which reinforces our confidence in our findings to be without confounding due to these factors.

Comment: This justification for your subgroup analysis should have been in your methods section.

DLCO are deemed to abnormal when below 80% of the predicted value. Without predicted DLCO values, we assumed that the subjects had normal DLCO prior to the study and therefore a decline in DLCO by 20% would be significant. Using a 20% decrease in DLCO as a cutoff, we did not find a significant difference in the proportion of patients who had a significant decrease between the placebo and methotrexate. Final DCLO levels seem to be most correlated to the baseline DLCO, as patients with higher baseline DLCO levels have higher final DLCO, while those with lower baseline DLCO levels have lower final levels. We entertained the possibility that methotrexate may affect those already at a low DCLO. In our study, we defined patients having the minimum DLCO to be patients having the lowest 5% DLCO values, which was DLCO < 11 ml/min/mmHg. We compared the mean difference from baseline to final of patients with DLCO <11 ml/min/mmHg (treatment minus placebo) with the same statistic from patients with a DLCO ≥ 11 ml/min/mmHg. There were no statistically significant differences between the lower and higher DLCO groups. This finding reinforces our conclusion that the methotrexate dosages were not enough to affect DLCO levels in patients.

We also consider the limitations to the study. One such limitation of these findings is the skewed proportion of females to males, which is similar to the prevalence of this disease in the population as a whole. While it is possible that no gender specific differences exist, the sample population of males (n = 9) was too small to make any meaningful inferences about males in general. Because the sample population was chosen from hospital patients, we must also mention the possibility of Berkson's bias, which describes the situation where hospitalized patients differ from the population at large. The presence of such a bias would also decrease the generalizability of our findings. It is also probable that the duration of the study was not an adequate amount of time for adequate drug bioaccumulation. If the treatment was to continue longer, we may observe a more significant drop in DCLO in methotrexate patients. The presence of missing data should also be noted, as there is the possibility that missing data may be significant in itself. The missing values do not seem to be problematic in our analysis, but there is a chance that there may be a significant pattern or reason that these patients are missing data.

Future studies might take into account the higher doses of methotrexate that are administered to patients for other illnesses, in efforts to examine their potential effect on DLCO. Data collection methods could be improved by recording an estimate of the total amount of drug taken over time, since this amount is likely to be very variable, and not consistent among patients. For instance, it is possible that more patients on the methotrexate arm experienced toxicities, and hence had their dose amount reduced. From these data, it could not be determined if this was the case. Additionally, lung toxicity could be examined by more sensitive methods such as high resolution CT, or other imaging technologies.

Investigating potential confounding factors of this study could be a part of future studies as well. It would be of interest to ascertain tobacco smoke exposure, history of restrictive or obstructive lung disease, history of defects in coagulation, history of anemia, history of structural cardiac disease, and history of carboxyhemoglobinemia.

Comment: So this is testing for an interaction. A good thing to do, but you will of course have very little statistical power, so lack of statistical significance does not necessarily mean we do not have a problem.

Comment: So why is this a limitation. Generally we like to have representative patient samples.

Comment: No, these were all outpatients.

Comment: We had 6 years average follow-up!!

Conclusion:

From the statistical analysis presented above, we find that there is no significant difference in DLCO between the methotrexate and placebo groups. Therefore, we can conclude that the experimental doses of methotrexate do not cause lung toxicity as measured by DLCO in this sample population in the duration of this study.

Comment: The major limitation was the lack of precision due to the low sample size and the highly variable DLCO measurements.