Group 10

Summary

The question addressed by the following analysis relates to whether or not the administration of the study drug, methotrexate, leads to impaired lung function as measured by DLCO. The study is a prospective, double blinded, randomized study of methotrexate as a potential treatment for primary biliary cirrhosis, involving 265 patients divided between two treatment arms, one receiving placebo (n=133 patients) and one receiving methotrexate (n=132 patients). The goal of the overall study was to investigate whether methotrexate would prolong the survival or progression free survival of patients with auto-immune cirrhosis. The main outcome measure analyzed was the average change in the DLCO, defined as the difference between the final DLCO measurement and the baseline DLCO. The difference in the methotrexate group was 1.59 (95% CI from .85 to 2.34) and the difference in the placebo group was 1.51 (95% CI from .78 to 2.24). The absolute difference between these two groups was .082 (95% CI from -1.12 to .95) and this difference was not statistically significant (p = 0.562).

Comment: A potential toxicity of MTX treatment

Comment: multicenter

Comment: Treated for how long?

Comment: units? And are these diff Final – baseline? Did they tend to go up in DLCO?

Background

Primary biliary cirrhosis (PBC) is thought to be an autoimmune liver disease of uncertain etiology characterized by scarring and blockage of the intrahepatic bile ducts ultimately leading to advanced liver failure or cirrhosis. There is a striking female predominance to the disease and there are a number of studies that have linked PBC to serum antibodies to mitochondrial antigens. Although convincing epidemiologic studies have not investigated PBC, in most cases the disease results in progressive liver failure requiring liver transplantation.

Comment: convincing whom?

Because of blockage of normal egress of bile from the liver the clinical presentation of patient with PBC is with elevated serum bilirubin and a clinical picture consistent with jaundice. Chronic exposure to elevated bilirubin damages the liver leading to some of the hallmarks of liver failure. These hallmarks include derangements of the fat metabolism and liver excretory function leading to elevated serum cholesterol, triglycerides, bilirubin, estrogen and copper, breakdown of liver cells leading to elevated liver function tests and sequelae of portal hypertension including splenomegaly, ascites and low platelet count. This study investigated a number of these serum studies and clinical variables as markers of severity of liver disease.

Because of the presumed autoimmune nature of PBC, methotrexate has been considered as a potential treatment. Methotrexate is an antimetabolite drug that has been used in a number of diseases including cancer and disease processes associated with rapid cell growth. Other autoimmune diseases like rheumatoid arthritis and psoriasis, have been successfully managed with methotrexate and therefore it seems a natural extension to consider it for treatment of PBC. This study considered the effect of methotrexate as a treatment for PBC as compared to a placebo group. The primary outcomes of interest were survival and progression of disease.

This particular analysis does not consider these primary outcomes but rather the association between methotrexate and pulmonary toxicity, an association which has been previously described in previous trials of methotrexate. Although this toxicity is thought to be progressive and related to pulmonary fibrosis, there is certainly a discrete group of patients who have acute lung toxicity that is reversible and abates when the drug is discontinued or if the dose is reduced. The test that is utilized in this analysis to study lung toxicity is DLCO or the diffusion capacity of the lung for carbon monoxide.

DLCO is utilized to determine how well oxygen is able to pass from the alveolar sacs to the blood and therefore this study investigates lung pathology at the alveolar capillary membrane. The test is performed by inhaling a single breath from a volume of gas containing a known small quantity of carbon monoxide. The patient holds their breath for 10 seconds, and then rapidly exhales. The exhaled gas is then analyzed to determine the amount of carbon monoxide that was absorbed during the breath. A decline in DLCO from a baseline value would be considered a sign of lung disease. Considerable within-subject variability—up to 8mL/min/mmHg can—can occur in healthy subjects, primarily because the test is effort dependent to some degree. This test is usually reported as the percent of predicted amount of carbon monoxide inhaled that should be absorbed, according to the age, sex, and height of the person. Despite the limitations associated with the test, DLCO is considered a reasonably informative and non- invasive test to evaluate pulmonary toxicity in patients who are prescribed methotrexate.

Comment: very good to note

Question of Interest

The overall goal of this analysis is to evaluate lung toxicity using DLCO in the two treatment arms in this prospective randomized trial. The primary question addressed by this analysis is whether the use of methotrexate, at maximum dosages of 15mg/1.73m^2 or 20mg/week, results in pulmonary toxicity, as measured by DLCO measurements obtained at baseline and annually during the treatment phase of a trial of methotrexate for the treatment of PBC.

Source of Data

This analysis was performed within the context of a larger, randomized, double blind, placebo controlled trial assessing the efficacy of ursodeoxycholic acid (UDCA) followed by methotrexate or placebo on transplant free survival and disease progression in patients with primary biliary cirrhosis (PBC). Inclusion criteria was age 20-69 years, diagnosis of PBC made at a minimum of 6 months prior to study enrollment, a liver biopsy confirming the diagnosis of PBC obtained at a minimum of 6 months prior to enrollment, positive anti-mitochondrial antibody test, and alkaline phosphatase level 1.5 times the upper limit of normal. For asymptomatic patients, a liver biopsy 6 months prior to randomization indicating disease stage greater than 1 was required. Patients with advanced PBC were excluded based on levels of serum bilirubin > 3mg or albumin < 3mg and a history of ascites, hepatic encephalopathy, or variceal bleeding. Patients with liver disease not due to PBC were excluded, as were patients with a history of alcohol abuse within 2 years of study enrollment, a history of cancer or other life-threatening illness, HIV, immunosuppressive medication use, use of rifampin or dilantin, current pregnancy, or refusal to use birth control. Patients were randomized to methotrexate or placebo by disease stage as determined by biopsy results. Baseline measurements of variables known to be associated with disease severity, including disease duration prior to randomization, alkaline phosphatase, ALT, platelet count, bilirubin, prothrombin time, albumin, and cholesterol were obtained on each patient at the time of randomization. Age, height, and weight were recorded for each patient at baseline. Pulmonary toxicity was monitored by DLCO measurements at baseline and annually (while the subjects were still receiving placebo or methotrexate) during the study period (Tables 1).

Comment: multicenter

Two hundred sixty five patients were enrolled and followed at 12 different medical centers within the US. Minimum and final DLCO measurements were missing for 8 participants in the placebo group and for 13 participants in the treatment group

Comment: okay, now you tell me. But how long were they followed?

Table 1: Participant Baseline Characteristics (n=265)

Comment: This is usually put in Results.

	Placebo (n=133)	**Treatment (n=132)**
Female – n (%)	123 (92.5)	122 (92.4)
Age - mean (± sd)	52 (8.6)	50 (8.7)
Splenomegaly Status – n (%)		
Present	12 (10.5)	11 (8.3)
Unknown	0	1 (0.8)
Stage – n (%)		
1	24 (18)	13 (10)
2	40 (30)	47 (36)
3	47 (35)	56 (42)
4	22 (17)	16 (12)
Participant Characteristics – median (IQR)*		
Weight (kg)	73 (63-84)	68 (59-77)
Height (cm)	163 (157-168)	162 (158-170)
Disease Duration (years)	2.4 (1-5.4)	2.7 (0.9-4.5)
Laboratory Values – mean (± sd)		
Alkaline phosphatase (U/I)	245 (187)	243 (146)
ALT (U/I)	50 (42)	54 (41)
Platelet count (1000 cells/cu mm)	235 (83)	244 (89)
Bilirubin (mg/dl)	0.72 (0.39)	0.66 (0.44)
Prothrombin time (sec)	11 (1)	11 (1)
Albumin (g/dl)	4 (0.34)	4 (0.35)
Cholesterol (mg/dl)	236 (59)	237 (58)
Pulmonary Function Testing – mean (± sd)		
DLCO (ml/min/mmHg)	20 (5)	20 (5)

Comment: Why medians here? I would argue that we are probably more interested in means than medians for all of these variables. It is confounding we are afraid of.

Statistical Methods

The results are reported using point estimates, 95% confidence intervals, and one-sided p values. The point estimates considered were mean difference, geometric mean difference, and proportion of persons with decreased lung function in each group. In each case, these represent the best estimate we expect to see for each treatment

Comment: I hope not. I hope it was geometric mean ratio. (Negative values cannot be summarized with a geometric mean.)

based on the data collected. The 95% confidence interval indicates that if the true value were in this interval the point estimate would be not be atypical for the data. Finally, the p value represents the probability that under the null hypothesis of equal statistics between the treatment groups, of observing this data would occur. Low values of p values indicate that we should reject the null hypothesis that the statistic under analysis will be equal across treatment groups; with one-sided p values (because we are concerned about a decrease in lung function), the critical value is 0.025.

STATA 9.0 was used to complete all analyses on the data. Several different statistics, including the mean differences, difference in geometric means, and the proportion of persons with decreased lung function in the two treatment groups were considered and standard techniques for comparisons were completed. While each of these analyses provided alternate vantage points to look at the data, all reached the same conclusion with similar precision and support. We therefore chose the mean difference of DLCO levels (baseline DLCO level was subtracted from the final DLCO level obtained at discontinuation of drug or placebo) to determine the change in DLCO during the course of the study. While minimum DLCO values were also reported for each patient, we elected not to include this in our analysis as isolated low DLCO values may reflect the effort-dependence nature of the test or reversible changes. For analysis of the mean difference, a t test assuming unequal variances was performed, and one-sided p-values and 95% confidence intervals were examined. Finally in order to gain better precision, linear regression of DLCO levels at the conclusion of treatment was completed over treatment groups with adjustment for baseline DLCO levels as a covariate. With linear regression, we are looking at the first order trend between the treatment group and final DLCO levels – because of the influence on the final DLCO levels, we include baseline DLCO levels and adjust according to these levels. This analysis is likely to be more precise than the t test; however, in this case the results were essentially equivalent. Multiple comparison was not a factor in the analysis, and missing data was excluded (22 cases out of 265 seen to be sufficiently small enough and evenly distributed to be ignored).

Comment: How did you do this? I bet you looked at ratio of geometric means

Comment: How defined?

Comment: Generally just tell us what you used as primary outcome. And choose it prior to looking at the data.(Though in this case, I have a little sympathy for trying to show that you could not detect any toxicity any way you looked at it. Of course if something had been stat significant, you would have been in trouble.

Comment: Yes, but that is why you have a placebo group. Ignoring this is allowing some toxicity to show up (though if it did not persist, it might not be clinically important)

Comment: If you are good enough to do this analysis, it should have been your headline result: It is the best one.

Comment: I think it certainly was, given the number of different analyses you did. Of course, since they were all not significant, you don't have the problem of interpretation that you might have had. But you need to be aware that if you did all these analyses and then selected which you reported, the mult comparison issue is very real.

Results

Descriptive Statistics

At the completion of the study, DLCO levels were available for 119 of the 131 subjects randomized to receive the study drug methotrexate, and 125 of the 133 subjects randomized to receive the placebo drug. In the placebo group, the change in DLCO levels range between a decrease of 15.3 to an increase of 15.9 ml/min/mmHg, with an average decrease of 1.51 ml/min/mmHg (Table 2). In the group randomized to receive the study drug, methotrexate, the change in DLCO levels ranged from a decrease of 15.3 to an increase of 11.5 ml/ min/ mmHg, with an average decrease of 1.59 ml/min/mmHg. The overall difference in DLCO level changes between treatment arms is 0.08 ml/min/mmHG, with those receiving methotrexate having a slightly greater decrease in DLCO levels.

Table 2: Change in DLCO Measurements by Treatment Group

Change in DLCO levels (final levels – baseline) *(244 observations, 21 missing observations)*									**Correlation, Slope & Variance** *(DLCO difference & Time on Drug)*		
	n	***mean***	***std dev***	***min***	***25th %ile***	***median***	***75th %ile***	***max***	***Correlation (r)***	***Slope (ml/min /mmHg /year)***	***Variance***
Placebo	125	-1.51	4.11	-15.3	-4.20	-1.6	0.600	15.9	-0.18	-0.377	16.5
Drug	119	-1.59	4.09	-14.4	-4.5	-1.1	1.20	11.5	-0.10	-0.207	16.7
TOTAL	**244**	**-1.55**	**4.10**	**-15.3**	**-4.30**	**-1.3**	**1.05**	**15.9**	**-0.14**	-0.295	**16.5**

Comment: You have not told us how long the patients took the drug. This is very much of interest.

Comment: You should also include the final DLCO measurements. From change we cannot tell whether anyone got to some toxic level. And it might have been just one or two patients, so you do need the min and max as you supplied here.

Comment: I can't imagine what you think the correlation is telling us here. The slope is of far greater interest.

A scatterplot visualization of the change in DLCO levels for both the methotrexate and placebo arms as they relate to the amount of time on the drug shows a slight decrease over time in DLCO levels for the entire population (0.295 ml/min/mmHg/year), as well as for each treatment arm (Figure 1, Table 2). The rate of decrease for the placebo arm (0.377 ml/min/mmHG/year) is slightly greater than that for the methotrexate group (0.207 ml/min/mmHG/year).

Comment: This is very nice (and does give me some idea of time on drug)

Comment: Can you give CI and inference re a difference between groups?

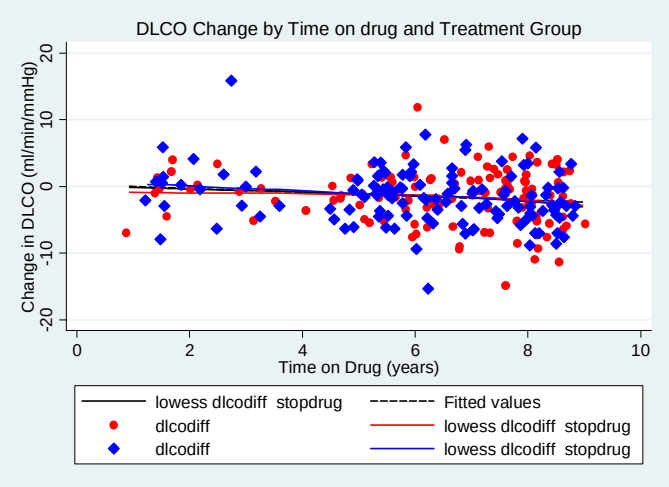

Figure 1: Scatterplot of Change in DLCO vs. Time on Drug
(black = total population, red = methotrexate, blue = placebo)

The variance of DLCO changes and the time on the drug for the entire population (16.5) is the same as the value for the placebo (16.5) and only slightly less than that for methotrexate (16.7). While this high value suggests variability in measurements, the similar values suggest that the spread for each group is consistent.

Comment: SD would be better because the units are then correct

The correlation values for the change in DLCO levels and the time on drug are similar among the entire population (r=-0.14) and both the placebo (r=-0.18) and the methotrexate arm (r=-0.10). Because these values are close to zero, they suggest that there is minimal correlation between the change in DLCO levels and the amount of time that each participant remained on either the placebo or the study drug.

Comment: Correlations are very hard to interpret scientifically. High variability leads to low correlation even when there are quite significant and clinically important trends. I strongly urge that you stay away from correlation as a scientific summary measure. It is not one.

Statistical Inference

As noted above, two different approaches were completed in this analysis: 1) comparison of the change in DLCO levels between treatment groups from baseline to end of study, and 2) comparison of DLCO levels between treatment groups at the end of the study with baseline DLCO levels as a covariate.

Comment: All you really needed. Numerically both would be giving the same answer on average, but statistically the ANCOVA is more precise.

Statistical Inference on Mean Change in DLCO Levels

In the placebo group, the average decrease in DLCO levels was 1.51 ml/min/mmHg. The 95% confidence interval for the true mean change in DLCO levels in a population receiving the placebo is from a decrease of 2.24 to a decrease of 0.78 ml/min/mmHg. In the group randomized to receive the study drug, methotrexate, the change in DLCO levels averaged a 1.59 ml min/mmHg decrease. The 95% confidence interval for the true mean change in DLCO levels in a population receiving the study drug methotrexate is from a decrease of 2.34 to a decrease of 0.85 ml/min/mmHg. Based on these data, it can be estimated that prescription of the study drug is associated with an average decrease in DLCO level that is 0.08 ml/min/mmHg more than any change in the absence of treatment with methotrexate (95% CI 0.95 ml/min/mmHg higher to 1.12 ml/ min/ mmHg lower). These results are not statistically significant (one-sided p = 0.562), and thus, these data are consistent with results that might be observed by random chance in the absence of treatment effect (Table 3).

Comment: Does the CI rule out all clinically important differences? We can never prove equivalence, but that is what we wish we could establish when talking about toxicity.

Table 3: Results of Statistical Analysis

Mean Difference DLCO levels					
	Mean Diff	*Std Error*	*95% Conf Interval*		*P Value (1 Sided)*
placebo	-1.512	0.368	-2.240	-0.784	
methotrexate	-1.594	0.375	-2.337	-0.851	
difference	**0.082**	**0.525**	**-0.953**	**1.117**	***0.562***
Regression on DLCO final over treatment groups with DLCO baseline as Covariate					
	Coefficient	*Robust Std Error*	*95% Conf Interval*		*P Value*
Treatment	0.092	0.471	-0.835	1.019	0.845
DLCO Baseline	0.636	0.6	0.518	0.754	<.0001

Comment: Probably don't care about this, but OK. Do note the narrower CI from ANCOVA than you obtained from the analysis of the differences.

Linear Regression of Final DLCO Across the Treatment Groups with Baseline DLCO as a Covariate

The mean DLCO level of the placebo group at the conclusion of the study was 0.09 ml/ min/ mmHg higher than that of the methotrexate study drug group at the conclusion of the study with the same baseline DLCO level (Table 3). The 95% confidence interval for the true mean DLCO level for the placebo group at the conclusion of the study was observed to be between 0.83 ml/ min/ mmHg lower and 1.01 mm/ min/ mmHg higher than subjects in the study drug group with similar baseline DLCO levels. These results are not statistically significant (p = 0.845), and thus, the data are not atypical of what we might expect to see with no true difference in mean DLCO levels between treatment groups of the same baseline DLCO levels. The mean DLCO level at the conclusion is 0.636 ml/ min/ mmHg higher for each ml/ min/ mmHg of baseline DLCO difference between two groups with similar treatment status (95% CI 0.518 ml/ min/ mmHg higher to 0.754 ml/ min/ mmHg higher).

Discussion:

There does not appear to be a significant effect of methotrexate on pulmonary function, as measured by DCLO, in this select group of patients with PBC. These findings were consistent across each statistical analysis performed, including difference in mean change, geometric mean change, proportion of subjects with decreased DLCO and linear regression of final DLCO measurement using baseline DLCO as a covariate. This relationship was not affected by the duration of exposure to methotrexate (Figure 1). The minimum DCLO was not utilized in this analysis because it was not considered to be scientifically relevant as isolated low DLCO values may reflect the effort dependent nature of the test or reversible changes in pulmonary function.

Comment: OK

This analysis is limited by a number of factors. Because information was not provided regarding the dose of methotrexate per patient, we were unable to analyze the dataset for a dose relationship between methotrexate and DLCO and we are also unable to determine the proportion of participants on methotrexate dosages significant enough to be expected to result in toxicity. There were also a number of missing values in both the placebo (8) and the methotrexate (13) groups, and these patients were not part of the analysis. DLCO was the only measure of lung toxicity used in this analysis and, as noted in the background section, it is effort dependent and subject to variability. Therefore, because there was only one measure of DLCO reported at baseline and one reported at the end of the treatment, these single measurements may not be accurate representations of true lung function at the different time points. Additional measures of lung function such as FEV1 or arterial blood gas may have supplemented this analysis. Finally, this is a select group of patients enrolled in a clinical trial for PBC treatment we therefore cannot make statements regarding the generalizability of these conclusions to the use of methotrexate for the treatment of other disorders.

Given our above findings, we can conclude that the use of methotrexate, in this specific patient population and at the specified dosages, does not result in significant pulmonary toxicity as measured by DLCO measurements obtained at baseline and just prior to discontinuation of the medication. Given the limitations of our analysis, we would recommend further exploration of this relationship by expanding the analysis to include information regarding methotrexate dosing and would also recommend performing similar analyses of the effects of methotrexate on pulmonary function in patients with different comorbidities (ie renal disease, which would effect excretion of the medication) which may result in different susceptibilities to methotrexate toxicity.

Comment: Be afraid. Be very afraid. If you are equating lack of statistical significance with lack of clinical significance you may well kill patients regularly. Comment on the precision and whether you have ruled out all differences that are clinically important.

And that would just be on average. You would also need to consider anecdotally any patients whose DLCO dropped a lot (you gave me the min and max of these) as well as any patients whose DLCO dropped to some low level (you never told me this)