Biost 517: Applied Biostatistics I
Emerson, Fall 2005

Homework #5 Key
November 4, 2005

A file containing the annotated Stata commands I used to solve this homework is available on the class web pages.

Written problems: To be handed in at the beginning of class on Friday, November 4, 2005.

> *On this (as all homeworks) unedited Stata output is **TOTALLY** unacceptable. Instead, prepare a table of statistics gleaned from the Stata output. The table should be appropriate for inclusion in a scientific report, with all statistics rounded to a reasonable number of significant digits. (I am interested in how statistics are used to answer the scientific question.)*

The written problems all refer to the DFMO data set as stored on the class web pages. My guess is that you will find this problem easiest to do using the "wide" format for the data, but it does not make too much of a difference either way.

Some overall comments about this problem:

In lecture we are covering one-sample tests, so I only asked you to analyze each dose group separately. This is NOT the way a clinical trial should really be analyzed. As noted below in some of the answers, we would really want to compare the results for the placebo group to the results for the high dose group to account for any natural trends over time in polyamine measurements. That is, the placebo group might naturally have decreased or increased polyamine measurements due to aging (they are a year older at the end of the study), due to environmental effects (who knows, maybe pollution causes changes in polyamine levels), due to secular trends in behavior (maybe exercise is increasing or decreasing), or due to seasonal effects in diet (although we are less worried about this when the treatment period is 12 months).

It would not be acceptable in a proper scientific report (and this is the definition of "proper") to merely report the analyses given below and to conclude a treatment effect because significant differences were (at times) observed in the high dose group but not the placebo group. (We could have done that if there were nonoverlapping CI, but since the CI for the two dose groups do overlap, we would have to perform a proper test comparing the two groups directly.)

I also note that I had you do several different analyses. In real life you would have *a priori* chosen one of the analyses (mean difference, mean ratio, proportion decreasing, ratio of geometric means) as your primary analysis. Otherwise, there is a multiple comparison issue. This exercise was meant to have you gain experience in performing and interpreting the tests, and to gain a little insight into how they behave relative to each other. In real life, I looked at geometric means, and I really used the spermidine : spermine ratio.

1. Perform an analysis to assess whether the dose 0 group had a change in mean spermidine level after 12 months of treatment. Use both the difference between measurements and

the ratio between measurements as a measure of comparison. Provide relevant point estimates, 95% confidence intervals, and P values. Do the same for the dose 0.4 group. Make clear the interpretation of your confidence interval and P values, including the scientific relevance of your results.

Ans: **Thirty-two (32) subjects were randomized to receive placebo, but only 28 of those subjects had spermidine measurements available at 12 months. The distribution of age, sex, and polyamine measurements at the time of randomization did not differ systematically between the 28 subjects for whom 12 month biopsy measurements were available and the 4 subjects missing measurements at 12 months.** *(The annotated Stata log file contains the descriptive statistics that I looked at for this latter comparison. I might also have looked at the values of the polyamines at 6 months for those patients who dropped out relative to those who continued. Nothing can tell us whether the biopsies we would have obtained were unusual, but had we seen trends in the baseline or intermediate outcome data, we would certainly worry more about nonignorable missing data, i.e., we would have worried that the patients for whom we were missing data might have had very different values than we observed in the other patients.)*

Analysis of the change in spermidine levels in the placebo group is restricted to those 28 patients with data available at 12 months. At the time of randomization, the average spermidine measurement in that group was 3.30 μmol / mg protein (SD 0.276 μmol / mg). After the 12 month treatment period, the average spermidine measurement was 3.26 μmol / mg protein (SD 0.248 μmol / mg). *(Note my careful wording "after the 12 month treatment period" rather than "after 12 months of treatment". Fact is, there are almost always some patients who do not take the treatment the full time, but we still analyze their data with the rest of the patients. This is called an "intent to treat" analysis, and that forms the standard for reporting clinical trial results. Ideally we would have obtained biopsies on the patients who dropped out.)*

(An analysis based on the mean difference in spermidine measurements)
The placebo group thus averaged a decrease of 0.041 μmol / mg protein (SD 0.248 μmol / mg) over the 12 month treatment period. Such a decrease was not statistically significantly different from 0 (lower one-sided P = 0.44). A 95% confidence interval suggests that the observed results were not unusual if the true average change in spermidine over the 12 month period in a population treated with placebo were anywhere between a decrease of 0.635 μmol / mg protein to an increase of 0.554 μmol / mg protein. *(Note that I used a one-sided P value, because our goal was to decrease polyamine levels. I do not think, however, that it is wrong to report two-sided P values in this setting. Just make clear which you use. When I use one-sided P values, I will only declare it statistically significant if the $P < 0.025$, while with a two-sided P value I would call it statistically significant if $P < 0.05$. In either case, I still report a 95% confidence interval, which by its very nature is two-sided.)*

(An analysis based on the mean ratio of spermidine measurements)
When analyzed as a proportionate change, after the 12 month treatment period the placebo group averaged a 10.3% increase in spermidine measurements relative to each patient's measurement at baseline. Such a measurement would not allow rejection of the null hypothesis of no change in spermidine measurements (lower one-sided P= 0.872). A 95% confidence interval suggests that the observed results were not unusual if the true average percentage change in individual spermidine levels were anywhere between a 7.9% decrease or a 28.5% increase. *(Note that the estimated average ratio was 1.10, with a 95% confidence interval for the ratio of 0.921 to 1.285. I chose to express these results as a percentage change, rather than as a ratio. Had the average ratio been greater than 2, I would be more likely to use*

wording based on the ratio. For this data, my wording would have been "When analyzed as a proportionate change, after the 12 month treatment period the placebo group averaged spermidine measurements that were 1.10 times higher than each patient's measurement at baseline. Such a measurement would not allow rejection of the null hypothesis of no change in spermidine measurements (lower one-sided P= 0.872). A 95% confidence interval suggests that the observed results were not unusual if the true average proportionate change in individual spermidine levels were anywhere between measurements that were only 0.921 as high as the measurement made at randomization to 1.28 times higher.")

Twenty-eight (28) subjects were randomized to receive a dose of 0.4 mg/m^2/day, but only 20 of those subjects had spermidine measurements available at 12 months. The distribution of age, sex, and polyamine measurements at the time of randomization did not differ systematically between the 20 subjects for whom 12 month biopsy measurements were available and the 8 subjects missing measurements at 12 months.

Analysis of the change in spermidine levels in the high dose group is restricted to those 20 patients with data available at 12 months. At the time of randomization, the average spermidine measurement in that group was 3.71 μmol / mg protein (SD 0.423 μmol / mg). After the 12 month treatment period, the average spermidine measurement was 1.95 μmol / mg protein (SD 0.179 μmol / mg).

(An analysis based on the mean difference in spermidine measurements)
The high dose group thus averaged a decrease of 1.76 μmol / mg protein (SD 0.485 μmol / mg) over the 12 month treatment period. Such a decrease was highly statistically significantly different from 0 (lower one-sided P = 0.009). A 95% confidence interval suggests that the observed results were not unusual if the true average change in spermidine over the 12 month period in a population treated with a DFMO dose of 0.4 mg/m^2/day were anywhere between a decrease of 0.74 μmol / mg protein to an decrease of 2.77 μmol / mg protein.

(An analysis based on the mean ratio of spermidine measurements)
When analyzed as a proportionate change, after the 12 month treatment period the high dose group averaged a 24.5% decrease in spermidine measurements relative to each patient's measurement at baseline. Such a measurement would not allow rejection of the null hypothesis of no change in spermidine measurements (lower one-sided P= 0.058). A 95% confidence interval suggests that the observed results were not unusual if the true average percentage change in individual spermidine levels were anywhere between a 55.6% decrease or a 6.69% increase. *(It is not unusual for a comparison of mean ratios to not be statistically significant even when the analysis based on mean differences does suggest a decrease. In this data, I believe it is the very low measurement values that creates a lot of variability in the estimated ratios. Generally it is better to analyze data on a log scale if you are interested in proportionate change, but when we do that, we have shifted our focus to geometric means instead of means.)*

2. Perform an analysis to assess the proportion of the dose 0 group that had lower spermidine levels after 12 months of treatment than they did at the time of randomization. Provide relevant point estimates, 95% confidence intervals, and P values. Do the same for the dose 0.4 group. Make clear the interpretation of your confidence interval and P values, including the scientific relevance of your results.

(An analysis based on the proportion of patients showing a decrease in spermidine measurements)

Over the 12 month treatment period, 46.4% of the patients in the placebo group were observed to have a decrease in spermidine levels. This observation was not statistically significantly different from a 50% rate of observed decreases that might be expected if there were no systematic trend toward lower or higher spermidine levels over time (P= 0.714 in a one-sided test looking for rates higher than 50%). A 95% confidence interval suggests that the observed results were not unusual if the true percentage of patients expected to have observed decreases in a population treated with placebo were anywhere between 27.5% and 66.1%. *(Note that this approach based on testing the proportion decreasing against a null hypothesis of 50% relies on a belief that spermidine measurements would not naturally increase or decrease with age or time. Of course, the reason we do controlled clinical trials is to be able to assess what would naturally happen in the absence of treatment.)*

Over the 12 month treatment period, 80.0% of the patients in the high dose group were observed to have a decrease in spermidine levels. This observation was statistically significantly higher than a 50% rate of observed decreases that might be expected if there were no systematic trend toward lower or higher spermidine levels over time (P= 0.006 in a one-sided test looking for rates higher than 50%). A 95% confidence interval suggests that the observed results were not unusual if the true percentage of patients expected to have observed decreases in a population treated with a DFMO dose of 0.4 mg/m^2/day were anywhere between 56.3% and 94.3%.

3. Perform an analysis to assess whether the dose 0 group had a change in geometric mean spermidine level after 12 months of treatment. Use the ratio of geometric means as a measure of comparison. (Note: Inference on the geometric mean is easily obtained by taking the log transform of your data, and then comparing means using differences. When you exponentiate the resulting estimates, you will have inference based on the geometric means and the ratios of geometric means. There is a handout on the class web pages which deals with the interpretation of log transformed data.) Provide relevant point estimates, 95% confidence intervals, and P values. Do the same for the dose 0.4 group. Make clear the interpretation of your confidence interval and P values, including the scientific relevance of your results.

(An analysis based on the ratio of geometric means of spermidine measurements, which would also be the geometric mean of ratios computed for each individual.)

Analysis of the change in spermidine levels in the placebo group is restricted to those 28 patients with data available at 12 months. At the time of randomization, the geometric mean of spermidine measurements in that group was 3.00 μmol / mg protein. After the 12 month treatment period, the geometric mean of spermidine measurement was 3.01 μmol / mg protein. Thus, after the 12 month treatment period the geometric mean of spermidine levels in the placebo group was 0.17% higher than the geometric mean for those patients at randomization. Such a measurement would not allow rejection of the null hypothesis of no change in spermidine measurements (lower one-sided P= 0.508). A 95% confidence interval suggests that the observed results were not unusual if the true average percentage change in geometric means were anywhere between a 16.2% decrease or a 19.7% increase.

Analysis of the change in spermidine levels in the placebo group is restricted to those 20 patients with data available at 12 months. One of those patients had a measured spermidine level at 12 months that was below the limit of detection, and a value of one-half the lowest

observed nonzero value was used for that patient. At the time of randomization, the geometric mean of spermidine measurements in that group was 3.21 μmol / mg protein. After the 12 month treatment period, the geometric mean of spermidine measurement was 1.71 μmol / mg protein. Thus, after the 12 month treatment period the geometric mean of spermidine levels in the placebo group was 46.8% lower than the geometric mean for those patients at randomization. Such an observation highly statistically significant, allowing rejection of the null hypothesis of no change in spermidine measurements (lower one-sided P= 0.004). A 95% confidence interval suggests that the observed results were not unusual if the true average percentage change in geometric means were anywhere between a 17.2% decrease or a 65.8% decrease.

DATA ANALYSIS

To be discussed in discussion section November 2-7, 2005.

We will discuss descriptive and inferential analyses appropriate for the scientific question posed in the documentation for the FEV data set.