# Time Interval and Analysis Block Considerations in Repeated Measures Analysis

Due to advances in technology, the ability to collect continuous measurement (such as heart rate, blood pressure, etc) on animals during a study has become increasingly common over the last few years. BioSTAT makes use of an area of statistics known as ***repeated measures analysis*** to analyze this type of data. Repeated measures analysis relies on the underlying assumption that responses measured repeatedly over time are correlated. Typically responses that are closer together in time are more highly correlated with each other than responses that are farther apart in time. Repeated measures analysis models this correlation in order to get a more precise estimate of underlying variability used in the test for treatment effects.

It is not uncommon for a safety pharmacology study to include 24 hours (or more) of continuous measurements on a parameter such as heart rate. This data is then 'binned' into various time intervals for summary and analysis. These time intervals are determined by the study scientist and range from 5 minute, 15 minute, 30 minute or even 1 hour averages. This binned data is then sent to BioSTAT for analysis. How this analysis proceeds should involve some discussion between the scientists and BioSTAT to determine the primary objective of statistical analysis.

## Analysis Blocks

BioSTAT recommends that the scientist consults with the statistician when choosing analysis blocks. The discussion with BioSTAT should involve the size, in terms of time and number of time points, in the individual analysis blocks. The term analysis block refers to the number of repeats (time points) that are included in a repeated measures analysis. For instance, an analysis block may include the first 4 post-dose hours with data binned every 15 minutes resulting in 16 repeated measures in the analysis block. The determination of the size of the analysis block is important in several ways.

It is important that environmental conditions be taken into consideration when setting analysis blocks. It is essential that data within an analysis block be highly correlated. Therefore, it is recommended that data from light and dark cycles be analyzed in separate analysis blocks. If meals are given by lab technicians the intervals should be split to include pre-meal times in one analysis block and post-meal times in a separate analysis block. Other known study conduct items should be used in determining analysis blocks. This will ensure that data within an analysis block remains highly correlated for statistical analysis purposes.

1. *Determination of blocks over the collection interval*

When available, known study drug characteristics should be used to help determine binning and analysis blocks. For instance, if the Tmax of the study drug is 30 minutes post-dose, the binning for the first hour might be 5 minute intervals with the first analysis block defined as the first hour post-dose. This would create an analysis block with 12 highly correlated time points. After the first hour, one might expand

the binning to include every 15 minutes and expand the analysis block to include the next four hours or until an environmental condition changes during the study.

2. *Number of time points in an analysis block*

Another consideration for analysis block size should be the amount of detail one hopes to garner from the statistical analysis. If the study is designed in such a fashion as to collect continuous data for 24 hours, then it can be inferred that the small details are of interest. When an analysis block includes too many time points the small details of the drug's effect on study parameters can be diluted or buried amongst the volume of data.

For example, if an analysis block includes 12 hours of data binned every 15 minutes, there are 48 time points in the repeated measures analysis. If the primary effect of a drug occurs over a short time frame (eg. 30-60 minutes post-dose or 3-4 time points) it can easily be diluted, from a statistical analysis perspective, by the remaining 44 unremarkable time points included in same analysis block. That is, the statistical analysis of the 48 time points may result in a false negative due to the inclusion of an overwhelming proportion of unremarkable time points. With this in mind, consideration should be given to the time frame of interest and the number of time points in an analysis block. Vickers (2003) notes that, although repeating measurements (time intervals) can have dramatic effects on the power of a statistical analysis, the marginal benefit of an additional measure rapidly decreases as the number of measures rises. And there is little value in increasing the number of post-treatment time points beyond seven when baseline measurements are taken. **With this in mind, BioSTAT suggests limiting the number of time points in an individual analysis block to no more than 12**.

3. *Illustrative example*

Figure 1, below illustrates an experiment where the initial block spanned 8 hours with data binned every 10 minutes for a total of 48 time points. Note that the data displayed have been converted to differences from baseline for illustrative purposes. When the data are analyzed using all 48 time points, there is no significant dose by time interaction. Table 1 displays the results of Dunnett's test to evaluate the differences from control.

Table 1: 8 hour, 48 time point analysis

| Dose Comparison | Dunnett p-value |
|---|---|
| 2 vs. 1 | 0.066 |
| 3 vs. 1 | 0.003 |
| 4 vs. 1 | 0.06 |

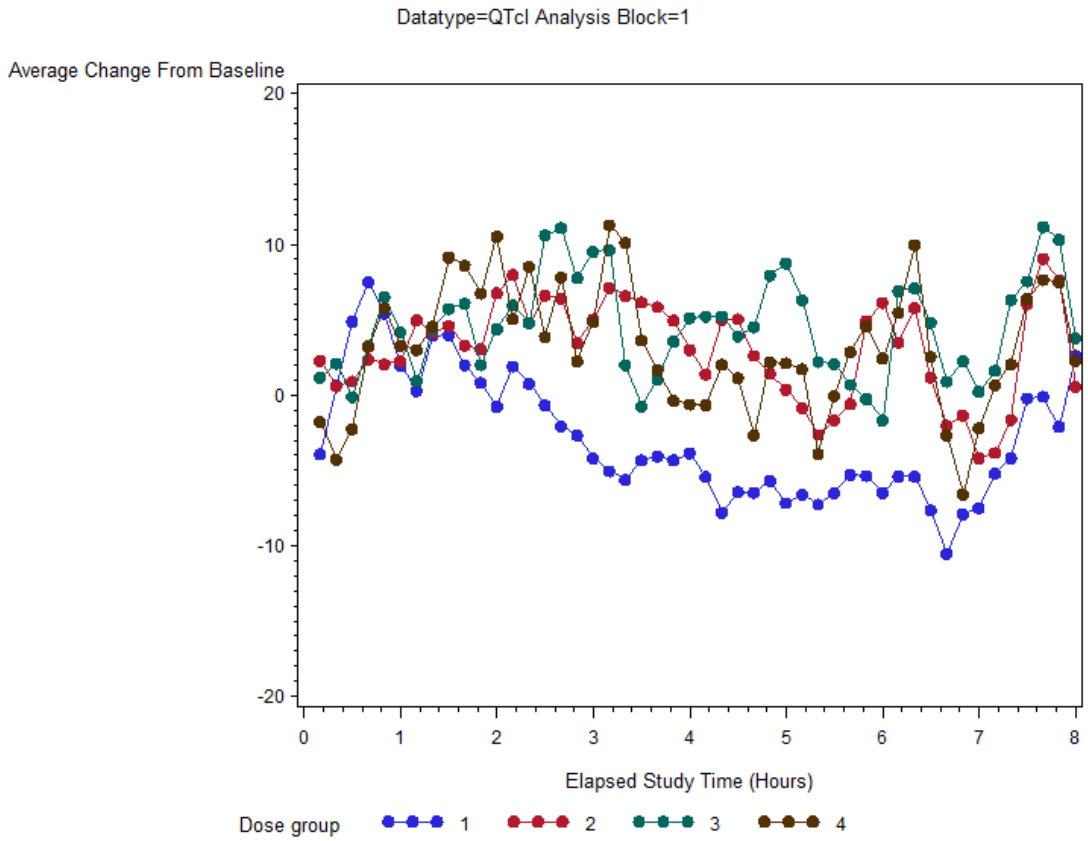
Contrast those results with Table 2 where the same data have been divided into 4 blocks, each spanning 2 hours with 8 time points per block. Note that in the first block there are no significant differences while the second block results show that all the treated groups significantly differ from control. Those

conclusions appear to better reflect what can be seen in Figure 1. The same amount of data has been analyzed as in Table 1 but the strategic selection of blocks allows the scientist to see the beginning of treatment effects, all of which are significant from control during the block 2 through 4 hours. Additionally, Block 4 shows that the effects are diminishing especially in dose groups 2 and 4. In the analysis of one large block of 48 time points, these effects are diluted by the sheer number of time points. The selection of blocks with 8 time points each meets the number of time points suggested by BioSTAT.

Table 2: 2 hour, 8 time point analysis

| Block 1 (0-2 hrs) | | Block 2 (2-4 hrs) | | Block 3 (4-6 hrs) | | Block 4 (6-8 hrs) | |
|---|---|---|---|---|---|---|---|
| Dose Comparison | Dunnett p-value | Dose Comparison | Dunnett p-value | Dose Comparison | Dunnett p-value | Dose Comparison | Dunnett p-value |
| 2 vs. 1 | 0.60 | 2 vs. 1 | 0.01 | 2 vs. 1 | 0.08 | 2 vs. 1 | 0.27 |
| 3 vs. 1 | 0.54 | 3 vs. 1 | 0.008 | 3 vs. 1 | 0.01 | 3 vs. 1 | 0.03 |
| 4 vs. 1 | 0.47 | 4 vs. 1 | 0.002 | 4 vs. 1 | 0.07 | 4 vs. 1 | 0.12 |

Figure 1



## Super Intervals

There have been recent discussions about the use of a "super interval" (Sivarajah, et al., 2010). This methodology takes average values for each animal over large intervals of time. The time intervals used in the reference noted above range from 6 to 8 hours. This means that a particular animal will have only 1 value for statistical analysis in the 6-8 hour interval. The suggested statistical analysis is performed for each of these 6-8 hour "super intervals".

There are several issues associated with this methodology. First, if the "super intervals" are not appropriately chosen, one could miss significant changes in the response. Referring to Figure 1, if a "super interval" was chosen to be 0-6 hours, an initial period of no drug effect would be averaged together with a period of very significant drug effect, leading to a lack of a statistically significant dose effect. Table 3 contains the comparisons of each dose with control using the same statistical model as in Sivarajah, A., et al. (2010).

Table 3: 6 hour super interval analysis

| Dose Comparison | Dunnett p-value |
|---|---|
| 2 vs. 1 | 0.28 |
| 3 vs. 1 | 0.14 |
| 4 vs. 1 | 0.80 |

Secondly, taking averages of large periods of time artificially reduces the amount of variation associated with a particular animal, known to be present in ECG data. This could result in artificial gains in statistical power. The use of the statistical technique of repeated measures with 12 or fewer time points in each block avoids those errors.

While the "super interval" technique may be useful as a quick and dirty analysis of the data, it naively ignores that the data are correlated and makes no use of that information for the analysis. It only gives a snapshot analysis for one point in time. Repeated measures analysis allows for the inference of a given interval of time because of the modeling of the correlation of the data between time points.

## Recommendations

- Repeated measures analysis methodology should be employed when a response is measured multiple times on the same subject.
- When determining the size of individual analysis blocks, consideration should be given to environmental conditions (e.g., light/dark cycles, feed cycles ) and drug characteristics (e.g. , tmax)

- An individual analysis block should consist of no more than 12 time points.
- The use of a “super interval” approach to the statistical analysis is useful only when more sophisticated analyses are not available and the reviewer is content with a quick, big picture review of the data. Otherwise, repeated measures analysis methodology provides a far superior evaluation of multiple responses from a subject.

Reference

Sivarajah, A., et al., Cardiovascular safety assessments in the conscious telemetered dog: Utilisation of super-intervals to enhance statistical power. J. Pharm. and Tox Meth. (2010).

Vickers, A.J. How many repeated measures in repeated measures designs? Statistical issues for comparative trials. BMC Medical Research Methodology, 3:22, 2003.