Nonclinical Statistics & Computing

Outliers

Date: April 18, 2013
Prepared by: Dhammika Amaratunga
With contributions from: Bill Pikounis, Ken Goldberg, Al Barron, Mariusz Lubomirski, Jessica Behrle, Kjell Johnson

Data will occasionally include unusual observations ("outliers") which are not consistent with the general tendency or variability of the data. Such observations could have a significant impact on the analysis and interpretation of the data. Therefore, we strongly recommend that an appropriate procedure be established and documented in the study protocol prior to any data collection to clearly describe what would be done in the event that an outlier occurs in your data. Ad hoc elimination of outliers is not prudent as this may introduce bias into the interpretation of the data.

Recommended approach

As stated above, prior to data collection, a procedure should be established that documents how outliers would be dealt with. One possible general approach (which is outlined in the "Outliers" flowchart) is as follows (but please also see the "Alternatives" section below):

1. Check the data for outliers using an appropriate outlier detection rule. The rule to be used should be specified as part of the established procedure[1].
2. If an observation is identified as an outlier, determine whether there is an assignable cause.
3. If such a cause is available and is defendable, delete the outlier and redo the analysis, documenting clearly the action taken and the rationale.
4. If such a cause is not available, the data may be analyzed without the outliers provided that this is stated as the established procedure prior to conduct of the experiment; the detection and exclusion of the outliers should be documented.
5. If such a procedure has not been established, analyze the data with and without the outliers.
6. If the two results are the same, report the analysis with all the data.
7. If the results differ, both should be reported and a management decision would have to be taken as to which would be used as the primary result; care should be taken that any exclusion of outliers is done in an unbiased manner; all details should be clearly documented.

Recommended outlier detection method

For a single batch of data, the following outlier detection method is recommended:

- To identify multiple outliers in samples of size 5 or more, use the Median-MAD test.

- To identify a single outlier in samples of size less than 5, use Grubbs test.

Notes regarding outlier detection methods

1. Two points to keep in mind regarding outlier detection methods are (i) they rely on an underlying model and assumptions (ii) they do not work well with small sample sizes.

2. Tests based on outlier-resistant statistics such as the Median-MAD test (symmetric rule) or the Boxplot test (asymmetric rule) are able to identify multiple outliers provided that the sample size is adequate. However, almost all other outlier tests (including commonly used ones like Grubbs test which is available in GraphPad Prism) fail when there are multiple outliers.

3. Visual inspection techniques are useful to assess the distribution of the data and identify potential outliers, but they do not offer any probability statements to judge removal of an outlier.

Some alternatives to consider

1. In certain cases, the data may appear to contain outliers when actually the situation is that the data are being analyzed on an incorrect scale. In such cases, a simple change in the scaling of the data, such as a log transformation, may be sufficient to take care of what seem to be outliers. In fact, failure to properly assess the scaling of the data is likely to weaken any analysis of the data whether or not there are outliers. For regularly performed assays, any transformation should be consistently applied across analyses.

2. Sometimes it is possible to include more data in outlier assessments by modeling the larger set of data and examining residuals (or standardized residuals if variances do not appear to be homogenous) from a resistant fit for outliers. This can be done with analysis of variance models, linear regression models, nonlinear models (e.g., the models used for IC50 estimation), etc.

3. Reference ranges constructed from historical control data could be useful indicators of how extreme values can be for a particular experiment, provided there is reasonable comparability of assay values over time.

4. Use robust statistical methods! We recommend the use of modern "outlier accommodation” methods rather than older "outlier rejection" methods. These encompass so-called resistant and robust methods that work well whether or not potential outliers are present. Rather than decide if each individual point is an outlier and thus should be excluded or not, all data are weighted, with some of the more unduly influential points downweighted. This removes the human burden of deciding to exclude or keep points in the statistical analysis. The inherent theory and algorithms for these methods have been well researched the past 40 years. Unfortunately they have not been widely implemented in mainstream data analysis software. We at Nonclinical Statistics & Computing can develop software for robust analysis of data and wish to make them available to all at Janssen R&D.