# Is your data normally distributed?

Authors: Christian J. Raue, Sandra Mükusch, Harald Seitz
Fraunhofer Institute of Cell Therapy and Immunology, Branch Bioanalytics and Bioprocesses

For statistical evaluation and determining the significance of experimental data with parametric tests (e.g. Student's *t*-test) a normal distribution of the data is necessary and essential. This kind of tests requires values like arithmetic mean and standard deviation (sd) for calculating the significance. Parametric tests are often way more precise compared to non-parametric ones. The Luminex® FlexMap 3D® Software computes mean and sd from the trimmed raw data. Therefore the question came up whether the data meets the prerequisite of normality. Our measurements (antibody antigen interaction using self-coupled microspheres as well as a commercial Cytokine Kit) showed that in many instances this is not the case. For that reason the application of parametric tests is not recommended here. We propose an option to achieve a normal distribution due to data transformation. Furthermore a robust alternative for the trimming is suggested. In cytokine quantification studies with a commercially available kit no normal distribution could be observed for the predominant amount of data. When using the mean as a descriptive value, care should be taken. Deviations up to 50 % occurred between the means and the medians of the datasets. In summary it can be stated that in these cases either non-parametric tests should be used or one should transform the data for the use of parametric tests. Besides, the removal of outliers is more robust in order to obtain normal distributed data than trimming the data.