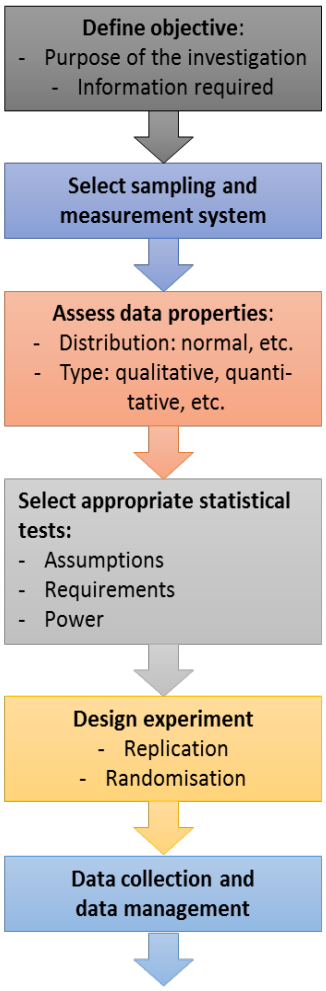
STATISTICAL GUIDANCE

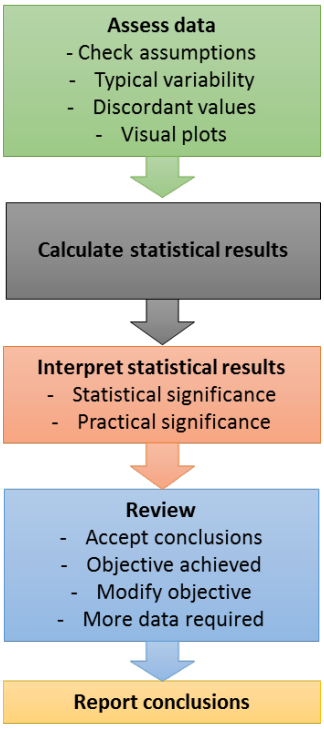
Statistics is the science of using data to increase the probability of making correct decisions. Generally, inferences are made about populations based on data obtained by sampling from that population.

Statistical testing should be used to give confidence that the study outcome in the context of its biological, clinical or physical relevance is unlikely to be due to chance. In addition, knowledge of the magnitude and variance of the measured response can help to define the size of the experimental study sample to ensure the study has sufficient power.

In a guideline of this kind, it is not possible to provide a complete treatise on the selection of appropriate statistical methods for the analysis of data obtained in a wide variety of study types. Selection of relevant statistical tests must be based on a knowledge of the scale of measurement, the variability of the data and the normality of the observations or data. It is important to use a statistical method to analyse the data that is appropriate to the purpose of the analysis, to the data type and to the data independency: using an incorrect technique will mean the conclusions drawn are unlikely to be sound. If in doubt, refer to a suitable text or seek assistance from a suitably qualified person.

The flowchart below summarizes the steps you should go through to ensure an effective statistical approach to data analysis:





General principles:

1) Sources of variation in data

All measurements are subject to variation. There are two types: special cause or common cause, either of which may be systematic or random. They have different properties. Special causes of variation are factors known to affect the measurement e.g. concentration of reagent. These effects can be estimated or eliminated by good experimental design. Common causes of variation are random, uncontrolled or uncontrollable effects e.g. measured value is different from true value because of the variability inherent in the measurement method. If variation is systematic this will introduce bias in the data which may make it impossible to derive sound conclusions from your results. All measurements are subject to random error. Random errors cause the measured values to vary without any particular pattern of deviation.

2) Study design

Before the study can be designed, you must define the study objective, what information is required to test it and how you wish to analyse the data. You may need to loop round the flowchart iteratively until the design is optimised. There are many possible designs that could be considered – it is important to choose the design that is most appropriate to address the study objective. Also when designing a study, it is important to minimise possible bias. Randomisation, pairing and blocking are techniques to minimise this.

3) Sample Size & Power

The size of study required will depend on the magnitude of the effect you wish to detect, the variability of the data and the power of the study. In general, the smaller an effect you wish to detect, the larger your study needs to be (all other factors holding constant). The power of a study is the probability that e.g. it will detect a difference of the magnitude specified if it

truly exists. It is typical to size studies based on 80% or 90% power. However, for exploratory or pilot studies a smaller power can be chosen.

The number of subjects /size of a study should always be large enough to provide a reliable answer to the questions addressed (i.e. have sufficient power). The number may be determined by the primary objective of the study through a formal sample size calculation or by a justification based on statistical and/or methodological expertise (background data, former study, etc.).

4) Data Management

Poor data collection and recording can affect the results of the analysis. Processes must be in place covering data entry, data manipulation and data transfer to ensure high data quality. Data should be recorded to adequate levels of resolution required for analysis. Check that data is not truncated or rounded before recording and record it to appropriate statistically significant figures.

5) Making decisions

During the study design phase, you will have generated an hypothesis (e.g. null hypothesis: no difference between treatments versus alternative hypothesis: there is a difference between treatments) that you wish to test. Now you have generated your study data, we ascertain whether there is sufficient evidence from the data to reject the null hypothesis in favour of the alternative hypothesis. The decision whether to reject the Null hypothesis or not, is based on the value of an appropriate test statistic calculated from the data and compared with a critical value of the statistic and this results in a p-value. A p-value is the probability of obtaining the value observed or one more extreme when there is in fact no difference.

Typically a significance level of 5% is chosen (2.5% in case of one-sided). This is the benchmark against which the p-value generated from the hypothesis test is compared. Obtaining results with p-values below 0.05 indicate that the risk of these differences having happened by chance alone is small i.e. less than 5%.

It is also good practice to calculate confidence intervals for your results to present with the p-values. A confidence interval gives an indication of the reliability with which the statistic based on the sample, estimates the true value from the population. Typically 95% confidence intervals are presented.

It is important to appreciate that you may obtain results that are statistically significant i.e. with p-values less than 0.05, but the results may not be of practical or clinical significance because for example the difference you have detected is so small to be of no practical or clinical relevance.

6) Statistical Method

There are many different statistical techniques. To analyse the data it is important to use a statistical method which is appropriate to the purpose of the analysis, to the data type and to the data interdependency. Using the incorrect technique will mean the conclusions drawn are not sound.

References

Statistical Methods in Medical Research. Fourth Edition. P Armitage, G Berry and JNS Matthews (2001)

Practical Statistics for Medical Research, DG Altman (1991)

Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building. Box GEP, Hunter WG & Hunter JS (1978)

ICH E9 Statistical Principles for Clinical Trials