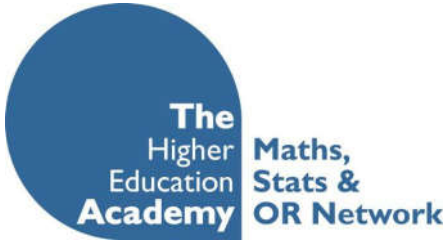
Guide to statistics

Supporting Statistics in Medicine



Resources to support the learning of mathematics, statistics and OR in higher education.
www.mathstore.ac.uk

The Statistical Education through Problem Solving (STEPS) glossary
www.stats.gla.ac/steps/glossary

www.rsscse.org.uk

www.medev.ac.uk

Email: admin@rsscse.org.uk

Contents

Figure 11: Kaplan-Meier curves for time until return to work for participants in a randomised controlled trial of treatments for lower back pain

Cox proportional hazards regression

The log rank test can be extended to adjust for multiple risk factors. This is called Cox proportional hazards regression. For the OT example above, it may be that in comparing the return to work times of the two groups, it is necessary to adjust for age, severity of original back pain, and whether the pain was of an acute onset, or chronic as all may affect healing. The outcome measure from a Cox regression analysis is the hazard ratio which may be interpreted as a relative risk of survival between the two groups adjusted for survival times.

10 Sample size

When planning a study it is important to have a good idea of how many individuals are needed to take part. The justification for a chosen sample size can range from a formal calculation, to what is feasible. A sample size calculation is also called a **power calculation** as it involves answering the question '*How many patients are needed in order to detect a particular size of difference for a pre-determined power (usually at least 80%) at a particular level of statistical significance (usually 5%)*?'. In order to calculate the size of the sample required, several quantities need to be known or estimated a priori as listed below:

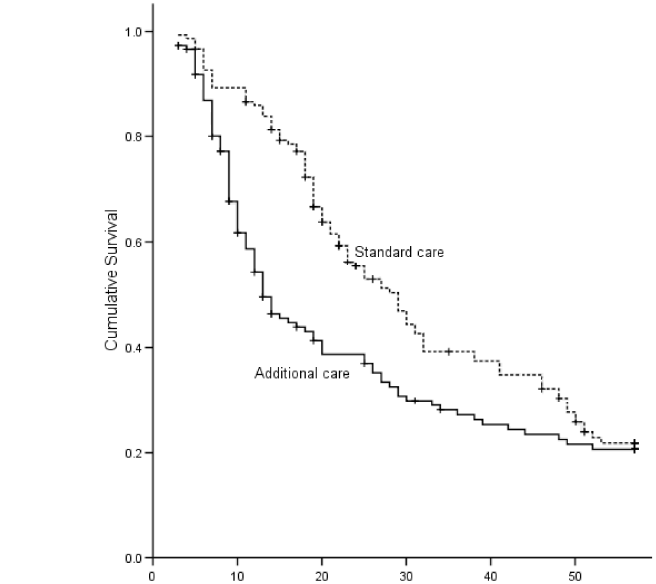
1. The number of groups to be compared
2. The outcome variable (and its distribution)
3. Standard deviation of the outcome variable
4. Minimum Clinically Important Difference and the size of difference likely to be detected. The latter should be at least as great as the former. If the size of difference that is likely to be detected is smaller than the minimally important difference it is arguable whether it is worth doing the study at all.
5. The acceptable level for the type 1 error rate (α). Usually set at 5%
6. The power ($1 - \beta$). Usually set at either 80% or 90%.

In addition, it is worth considering drop-outs when calculating sample size as the number calculated is the number necessary for analysis, the analysable sample. To account for potential dropouts this should be scaled up in order to give the sample size to be recruited.

and Cox proportional hazards regression have been developed to analyse survival data as they can explicitly deal with censored observations. Survival data can be plotted using Kaplan-Meier curves. The Kaplan-Meier curve plots average patient survival over time for the outcome of interest. Survival curves for several groups can be plotted together to show the survival of the groups relative to each other. The plot below shows the survival curves until return to work for two groups of patients with lower back pain, one randomised to specialised treatment by an occupational therapist and the other randomised to receive usual care. It can be seen that patients in the occupational therapy group returned to work more quickly over the 1 year follow-up period than patients in the control group.

Log Rank test

The log rank test is used to compare the survival curves for two groups. It assumes that the survival times are ordinal or measurable and that the risk of an event in one group relative to the risk in the other group does not change over time (proportional hazards assumption). The null hypothesis assumes that the median survival is the same in the two groups and the test examines the difference between the observed and expected number of events in each group given the survival experience of the two groups. The p-value obtained from the comparison between the treatment groups in the occupational therapy trial was 0.003 indicating that the two groups differed significantly in their time to return to work.



Number at risk:

Additional care	149	91	48	35	27	23
Standard care	150	133	91	54	42	29

1 What is Evidence Based Medicine?

> "Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients." See Sackett DL et al. http://www.cebm.net/

The practice of Evidence Based Medicine (EBM) involves making critical use of up-to-date medical research in order to make informed decisions with respect to the treatment and care of individual patients.
The procedure followed in EBM closely mirrors the statistical problem solving cycle (see below) and consists of the following stages:

- Posing a focused clinical question;
- Searching for the evidence;
- Appraising the evidence;
- Applying the evidence.

Statistics is an essential tool to aid the practice of EBM. It provides the relevant techniques required in order to analyse and interpret data from research and thus draw appropriate conclusions (inferences) from the results. Inferences are generally based on a sample of individuals drawn from the population of interest, since it is not usually feasible to study the whole population. Provided the sample is representative, these inferences will relate to the population of interest. An understanding of statistics also allows us to critically appraise the literature and apply the results of medical research to clinical decision-making.

The statistical problem solving cycle

Data are numbers in context and the goal of statistics is to get information from those data, usually through problem solving.

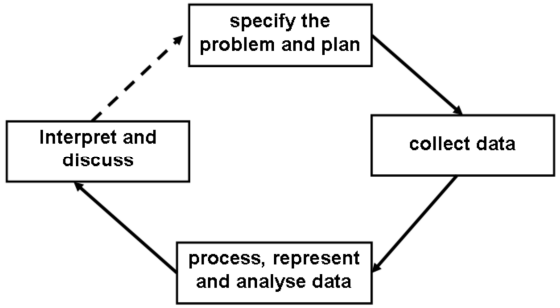


Figure 1: The statistical problem solving cycle

A procedure or paradigm for statistical problem solving and scientific enquiry is illustrated in the diagram. The dotted line means that, following discussion, the problem may need to be reformulated and at least one more iteration completed.

Population and sample: definitions

Population: the entire group of interest e.g. people who are obese.

Parameter: a particular characteristic of the population that we are interested in e.g. the population mean or proportion. This is usually unknown and so needs to be estimated from sample data.

Sample: a subset of the population of interest taken in order to estimate the population parameter. Measurements or values (*observations*) are taken from the sample in order to learn about the population. E.g. we could take a sample of people from a particular city in order to estimate the proportion of people in the city who are obese. It is important to ensure that the sample is a good and unbiased representation of the population of interest.

Statistic: a quantity calculated from the sample and used to estimate a population parameter e.g. the sample mean, or proportion.

Variable: in order to estimate the prevalence of obesity we would measure weight and height. These would be described as variables as their values differ from one individual to the next.
The table below provides some simple examples of population parameters with the corresponding sample statistics (estimate of the population parameter).

Table 1: Parameters and their corresponding sample statistics

Name	population parameter	sample statistic
Mean	μ	$\bar{x}$
Standard deviation	σ	s
Proportion	π	p

2 What are data?

How data are recorded, analysed and displayed will depend upon their type. Thus, in order to apply statistical methods we need to be able to classify the data that we collect into specific types. One simple classification is shown in the table below:

Table 2: a simple classification of data types

Data type	Sub-type	Example
Numerical	measurable	blood pressure; height; weight
	Count	number of visits to general practitioner in a year; parity
Categorical	Binary (two categories)	Gender; presence/absence of disease
	Nominal (no natural ordering)	Ethnic group; area of residence
	Ordinal (ordered)	Duke's stages in cancer development; Frequency of symptoms (never, rarely, sometimes, often, always)

Teenage pregnancy example:

The example below shows the relationship between Deprivation and Teenage pregnancy rates for 40 local authorities in the England for the years 1999-2001. In this example the equation for the regression line is:

Pregnancy rate = 13.04 + 6*deprivation score

Thus, when deprivation is 0 the teenage pregnancy rate is 13.04 and for every additional increase in deprivation of 1 unit, the pregnancy rate increases by 6 per 1000 women aged 15-17.

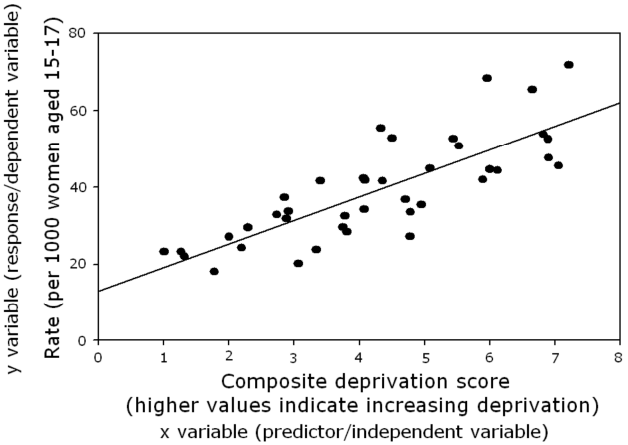


Figure 10: Scatter plot of teenage pregnancy rate against deprivation score for 40 English local authorities, with fitted regression line.

Ref: www.empho.org.uk/whatsnew/teenage-pregnancy-presentation.ppt

The method can be extended to adjust for other risk factors. In this case it is called **multiple regression**. Logistic regression is used when the outcome of interest has only two possible values (e.g. event/no event). In this case the outcome is expressed in terms of an odds ratio.

9 Survival – prognosis

Time to an event (e.g. recovery, death, etc.) is one of the main outcomes of interest in medicine. Data on the time to an event are called survival data, irrespective of whether the event is death or not, as it represents the time that the patient survives until they experience the event of interest e.g. time to pregnancy for women undergoing fertility treatment. Typically there will be some patients who do not experience the event of interest for example because they are lost to follow-up or they experience some other event. These are known as 'censored' observations and they require special handling in the data analysis. In particular methods such as the log rank test

Scatter plot: a visual representation of the direction and strength of the relationship between two variables. If it is known or suspected that one variable (known as the independent/explanatory/predictor variable) influences the value of the other variable (known as the dependent/response variable), the independent variable should be plotted on the horizontal axis and the dependent variable on the vertical axis. When undertaking either a correlation or simple linear regression analysis it is important to construct a scatter plot of the data as this will reveal how the two variables co-vary. It may be that the relationship is not monotonic and thus neither correlation nor simple linear regression analysis would be appropriate (e.g. Figure 6).

Pearson's correlation coefficient: r is used to quantify the strength and direction of the linear relationship between two variables.

Spearman's rank correlation coefficient: r_s is used if one or both variables are ordinal or we are interested in whether the two variables are increasing or decreasing in general together rather than in a straight line.

Positive correlation: $r > 0$, both variables increase simultaneously.

Negative correlation: $r < 0$, one variable increases as the other decreases.

No correlation: $r = 0$, no linear association (Figure 7).

Perfect correlation: $r = 1$ or $r = -1$, all points lie on a straight line (Figures 8 & 9).

Simple Linear Regression: A technique for describing quantitatively the linear relationship between a dependent variable *Y* and an independent variable *X*. It enables prediction of the value of *Y* from a known value of *X*.

Regression line: A straight-line equation that is used to model the relationship between the dependent (response) variable and the independent (predictor) variable. Note that the regression line should not be used to make predictions for *X* values outside the range of values in the observed data. For simple linear regression where there is a single response variable and a single predictor variable the equation of the regression line is given by:

$$Y = a + bX$$

Where:

- ***Y*** = dependent/response variable
- ***X*** = independent/predictor variable
- ***a*** = intercept: the value of the *Y* variable when the *X* variable is zero
- ***b*** = regression coefficient or slope. It shows the change in *Y* for a unit change in *X*. When the value of *Y* increases as *X* increases this will be positive. Conversely when the value of *Y* decreases as the value of *X* increases this coefficient will be negative.

The proportion of the total variability of the dependent variable, *Y*, explained by the regression on *X* is called r^2 and is often quoted as a measure of goodness of fit of the regression line to the data. Note that this is equal to the square of the correlation coefficient *r*.

Descriptive statistics for samples of size n

For numerical data:

Sample mean: the sum of the observations divided by the sample size (n).

Sample variance: the sum of the squared distances from the sample mean divided by the sample size minus one i.e. $n-1$. The divisor $(n-1)$ is called the **degrees of freedom (df)**.

Sample standard deviation (s): the square root of the sample variance. The standard deviation is useful as it provides a measure of the spread of the data that is in the same units as the mean, unlike the variance which is the square of the standard deviation.

If the sample data are ordered from smallest to largest then the:

Minimum (Min) is the smallest value;

Lower quartile (LQ) is the $\frac{1}{4}(n+1)^{th}$ value; it is the value below which the lowest 25% of data values lie; n.b. 75% of data values lie above it;

Median (Med) is the middle or the $\frac{1}{2}(n+1)^{th}$ value; it is the value which exactly divides the data in half; 50% of data values lie below and 50% of data values lie above it;

Upper quartile (UQ) is the $\frac{3}{4}(n+1)^{th}$ value; it is the value below which 75% of data values lie. N.b. 25% of values lie above it;

Maximum (Max) is the largest value.

These five values constitute a **five-number summary** of the data. They can be represented diagrammatically by a *box-and-whisker plot*, commonly called a *boxplot*. Note that the distance between the lower and upper quartiles is known as the **interquartile range** (IQR) and represents the region within which the middle 50% of the data lie. The distance between the minimum and maximum is known as the **range**. N.b. It is good practice to include the sample size.

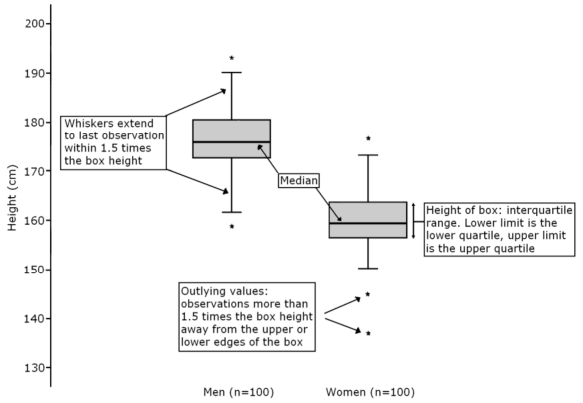


Figure 2: Boxplots for the heights of 200 randomly chosen men and women

We summarise numerical data that are symmetrically distributed using the mean and standard deviation and data that are not symmetrically distributed around the mean (skewed) using the median and interquartile range.

For categorical data:

We summarise data that are categorical by calculating the **proportion** that are in each category. Categorical data can be displayed using a **barchart**.

3 Sampling distributions and confidence intervals

Normal distribution: one of the most important distributions in Statistics. It is also known as the Gaussian distribution and has the following properties:

- it is bell shaped
- it is symmetrical
- any point on the horizontal axis can be expressed in terms of the number of standard deviations from the mean, for example 95% of the values lie within 1.96 standard deviations either side of the mean. This interval (mean ± 1.96xstandard deviations) is called often the reference or normal range.

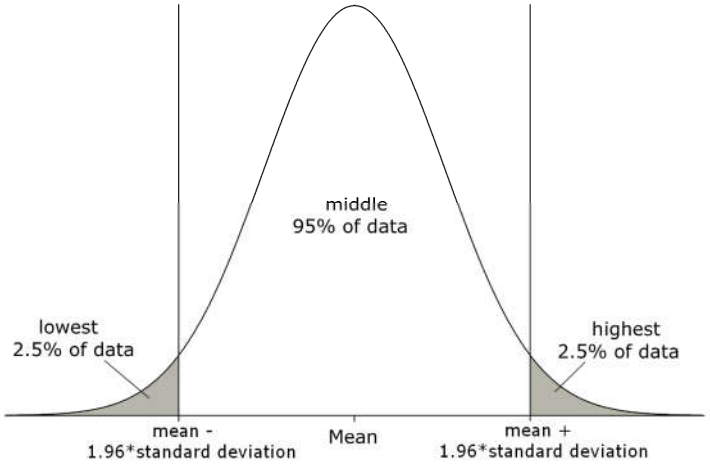


Figure 3: The Normal distribution

The Normal distribution with mean of 0 and standard deviation of 1 is called the **standard Normal** distribution.

Sampling distributions

The observed value of a statistic will in general vary from sample to sample. For example the proportion of individuals classified as obese will vary from one random sample to another. If we take many samples from the population of interest then the values of all these individual sample statistics

8 General concepts of correlation and simple linear regression

In medicine interest often centres on how two measurable variables taken from the same group of individuals relate to each other, in effect how they co-vary. E.g. the relationship between birthweight and maternal alcohol consumption. Two common techniques for analysing such data are correlation and regression.

Correlation: Measures the strength of the association between paired measurable data. Causation should not be inferred from a correlation coefficient as it simply measures the degree of association between the two. In addition, just because two variables are correlated at a particular range of values, it should not be assumed that the same relationship holds for a different range. Figures 6-9 illustrate ways in which two measurable variables can co-vary.

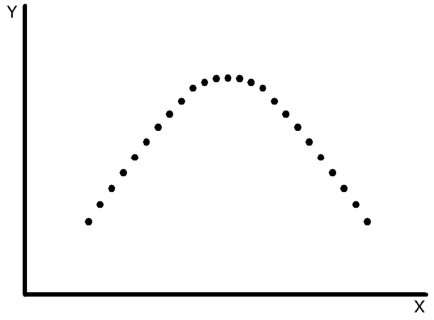


Figure 6: Correlation not appropriate

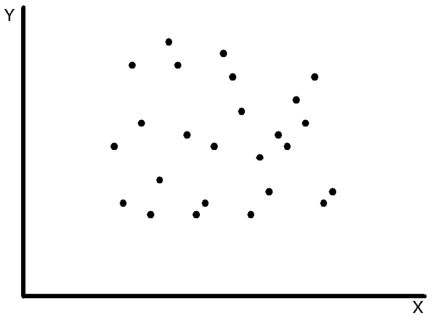


Figure 7: No correlation $r = 0$

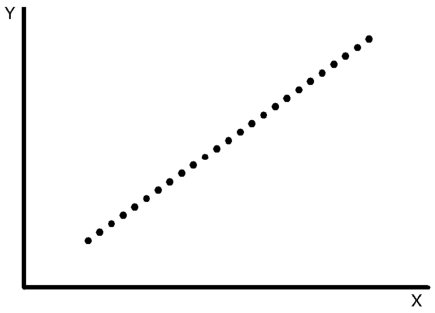


Figure 8: Perfect positive linear correlation $r = 1$

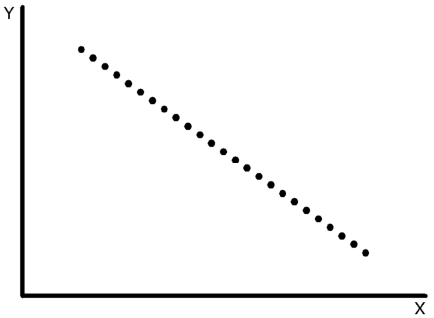


Figure 9: Perfect negative linear correlation $r = -1$

The Sensitivity and Specificity are useful for developing a test and explain the relative performance of the test. Theoretically they are not affected by the disease prevalence. The positive and negative predictive values are useful to a clinician and a patient using the new test in practice. They are affected by the prevalence of the disease in the population of interest and are therefore only applicable to the population for which the values were calculated.

A statistic that combines the extra information provided by the test result is the likelihood ratio.

Likelihood ratio: the ratio of the probability that a given test result would occur in a patient with the disease to the probability that the same result would occur in a patient without the disease.

For a positive test result:

LR(+) = Sensitivity/(1- Specificity) = a/(a + c) / b/(b + d).

For a negative test result:

LR(-) = (1- Sensitivity)/ Specificity=c/(a + c) / d/(b + d).

Clinical example:

Children with suspected measles were screened using a new laboratory test which was compared against the existing laboratory test (Gold Standard). In this study, the authors recruited children from primary care who were suspected of having measles.

Table 8:

	Gold standard:				
	Positive	Negative			
New laboratory test:			Sensitivity	=43/(43+59)	=0.42
			Specificity	=190/(190+8)	=0.96
			PPV	=43/(43+8)	=0.84
	43	8	NPV	=190/(190+59)	=0.76
Positive					
Negative	59	190	LR (+)	= 0.42/0.04	=10.5
			LR (-)	=0.58/0.96	=0.604

This test has a high specificity; i.e. there are very few false positive results. The test would therefore be useful to "rule in" the disease, if a positive result is found. The following mnemonics have been suggested:

S_p Pin: for a test with high specificity (S_p) a positive test rules the disease *in.*

S_n Out: for a test with high sensitivity (S_n) a negative result rules the disease *out*.

will form a sampling distribution. With a large enough sample the sampling distribution will be Normal. We can use this property of the sampling distribution to help us draw conclusions about the population of interest.

Standard Error

The standard error is the standard deviation of the sampling distribution of a sample statistic. It gives an estimate of precision of the statistic and is a measure of the uncertainty associated with that statistic. For example, given a single sample the standard error of the sample mean is estimated as $\frac{s}{\sqrt{n}}$, where s is the standard deviation of the sample data and n is the number of observations in the sample.

Confidence intervals for population parameters

The sampling distribution of a statistic is approximately Normally distributed. This fact can be used to provide a measure of how precisely the corresponding population parameter has been estimated. This is known as a confidence interval (CI). A 95% confidence interval takes the form:

$$\text{sample statistic} \pm 1.96 \times \text{standard error}$$

Example of mean, standard error and confidence interval

The mean height of a randomly selected group of 100 men was 174.8cm with a standard deviation of 5.07cm. Thus the 95% confidence interval for the sample mean was given by

$$174.8 \pm 1.96 \times \frac{5.07}{\sqrt{100}} cm = 174.8 \pm 1.0 cm$$

that is from 173.8cm to 175.8cm. This is usually taken to mean that we are 95% confident that, if we use 174.8cm for the mean height of men, then the worst mistake we are likely to make, is 1cm.

4 Hypothesis tests

Statistics is useful as it enables us to answer questions of interest, for example whether a new treatment is better than the current treatment. We start with a null hypothesis and then examine evidence to see if it can be sustained. A **hypothesis test** involves testing a claim, or **null hypothesis** H_0, against an alternative, H_1. A decision to **reject** H_0 or not reject H_0 uses sample evidence to calculate a **test statistic** which is used to obtain a **p-value**. The p-value is a useful reformulation of the test statistic and is the probability of obtaining the test results or results more extreme if the null hypothesis is true. H_0 is maintained unless it is made untenable by sample evidence i.e. the p-value is less than or equal to (≤) some pre-specified **critical value**. On the basis of the sample evidence the null hypothesis is either *rejected* or *not rejected*. Table 3 shows the different situations that can arise.

Rejecting H_0 when we should not is a **Type I error**. The probability of making a Type I error is called the **significance level**, α. Not rejecting H_0 when we should is a **Type II error**, which has probability β. The acceptable levels of committing a Type I error and a Type II error are specified before an analysis is conducted and this acceptable Type I error rate provides the critical value mentioned above. The **power** of a hypothesis test is the probability of rejecting the null hypothesis when it is actually false (power = 1-β).

Table 3: Testing hypotheses

The probability of rejecting the null hypothesis when it is actually false is called the **POWER** of the study (Power =1-β)

		The null hypothesis is actually: False	The null hypothesis is actually: True
Based on the test results you decide to:	**Reject the null hypothesis** (i.e. the p-value is less than or equal to α, usually set to 0.05)	Conclusion is correct. Have rejected the null when it is in fact false (1-β)	False positive / type I error / α
	Not reject the null hypothesis (i.e. the p-value is greater than α, usually set to 0.05)	False negative / type II error / β	Conclusion is correct. Have not rejected the null hypothesis when it is in fact true

The p-value is the probability of obtaining your results or results more extreme **if the null hypothesis is true**. It is the probability of committing a false positive error i.e. of rejecting the null when it is in fact true.

One-sided vs two-sided testing

A two-sided test is one in which the alternative hypothesis does not state a particular direction for the effect or difference. Conversely a one-sided test is one in which the alternative hypothesis is that an effect or difference is in a particular direction (e.g. greater than zero). It should be either theoretically plausible or interest only lies in one direction. For example, suppose that a new technology or treatment has been developed that is much cheaper than the existing treatment and it may be that interest lies in proving only that it is no worse. Provided the new treatment is at least as good as the old treatment then it will be used, as it is much cheaper. This is one of the few occasions when a one-sided test is justifiable in Medicine. If a one-sided test is to be used it should be stated at the design stage.

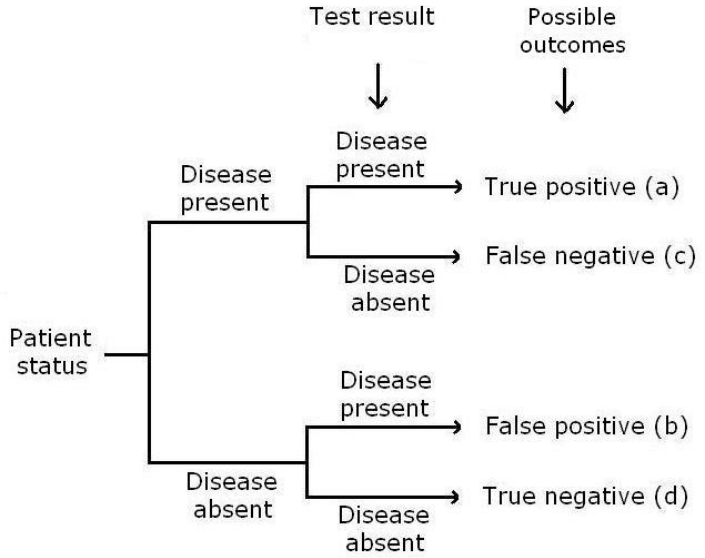


Figure 5: Possible outcomes from a diagnostic test

and these may be organised in a table as follows:

Table 7: Relationship between gold standard diagnosis and the results of a diagnostic test

	Gold standard diagnosis: Disease present	Disease absent	
Test:			
Disease present	a	b	a+b
Disease absent	c	d	c+d
	a+c	b+d	a+b+c+d=n

Tests are usually evaluated using the following statistical measures:

Sensitivity: the proportion of people with disease who have a positive test. In the 2x2 table, calculate as: a/(a + c).

Specificity: the proportion of people free of disease who have a negative test. In the 2x2 table, calculate as: d/(b + d).

Positive predictive value: the proportion of patients who test positive who actually have the disease. a/(a + b).

Negative predictive value: the proportion of patients who test negative who genuinely don't have the disease. d/(c + d).

Odds of an event occurring is the ratio of the probability of the event occurring to the probability of the event not occurring:

odds of an event given exposure = a/c = 45/905 = 0.05

odds of an event given not exposed = b/d= 25/1098 = 0.02

Odds ratio is the ratio of the odds of an event in the exposed group compared to the unexposed group = (a/c)/(b/d) = ad/bc = (45*1098)/(25*905) = 2.18. *Note that when the event of interest is rare the odds ratio approximates to the relative risk and is often interpreted as a relative risk, as is the case with this example.*

Absolute risk difference is the absolute additional risk of an event due to a particular exposure. It is calculated as the risk in the exposed group minus the risk in the unexposed group. If the risk is harmful, so that the risk is increased by the exposure this difference is called the **absolute risk excess (ARE)** (for example the absolute risk excess of an adverse event for treatment 1 is 0.05-0.02=0.03). If the risk is decreased by the exposure (for example using sunscreen to reduce the risk of melanoma) then this difference is called the **absolute risk reduction (ARR)**

The **number needed to treat** is a measure of the impact of a particular risk on patients often used in clinical practice. It is the additional number of people that would need to be given a new treatment in order to cure one extra person compared to the old treatment = 1/ARR. Alternatively, for a harmful exposure the number needed to treat is referred to as the **number needed to harm** and is calculated in the same way as the number needed to treat, but ignoring the sign. For adverse event example above the number needed to harm is 1/0.03=33.3. Thus approximately 33 people would need to on treatment 1 compared to treatment 2 for one additional adverse event.

Note that as with any estimated quantity it is possible to construct confidence intervals for these measures.

Points to consider when communicating risk

Individuals who do not deal with numbers and data regularly can often struggle to understand measures of risk, and this case it can be useful to express risks in terms of natural frequencies rather than percentages. Thus if we assume that the success rate following a single cycle of IVF is about 33% then it is more easily understood by stating that of 100 women undergoing treatment 33 will become pregnant.

7 Screening and diagnostic tests

Studies which evaluate diagnostic tests aim to compare their performance against a 'gold standard'. The gold standard is regarded as the true diagnosis of a condition. However it may not always be possible to know absolutely the true diagnosis and in this case the gold standard represents diagnosis using the best currently available method. Figure 5 shows the four possible outcomes when performing a diagnostic test:

Simple statistical tests

When comparing two groups it is important to distinguish between independent groups and paired groups. Two groups are considered to be *independent* when subjects are either randomly sampled from two distinct populations or randomly assigned to one of two groups. Two groups are considered to be *paired* when they consist of observations made within the same individual or between individuals who are explicitly paired.

Table 4: Simple statistical methods for comparing two groups

Comparison	Data type	Assumptions	Method
Difference between two independent groups	Numerical:		
	Measurable	Normally distributed	Independent samples t-test
		Not Normally distributed	Mann-Whitney U test
	Count		Mann Whitney U
	Categorical:		
	Binary	Large sample, most expected frequencies > 5	Chi-squared test
		Small sample, at least 1 expected frequency < 5	Fisher's exact test
	Nominal	More than two categories Most expected frequencies > 5	Chi-squared test
	Ordinal		Mann-Whitney U
Difference between paired groups:	Numerical:		
	Measurable	Differences Normally distributed	Paired t-test
		Differences not Normally distributed	Wilcoxon matched pairs test
	Count		Wilcoxon matched pairs test
	Categorical:		
	Binary		McNemar's test
	Nominal	More than two categories	No simple test available, consult a statistician
	Ordinal		Wilcoxon matched pairs test or sign test

5 Estimation, hypothesis tests and clinical significance

Estimation (using confidence intervals)

Assuming that the null hypothesis is true, a p-value indicates the likelihood of obtaining results at least as extreme as those found in the study. It can only be used to decide whether the results are statistically significant or not, it does not give any information about the likely size of the difference. Much more information, such as whether the result is likely to be of clinical importance can be gained by calculating the difference and its **confidence interval**. As confidence intervals are so informative there is a growing

consensus that only the estimate of the effect size and its confidence intervals should be reported for studies. However, it is unlikely that p-values will ever be eliminated as a way to quantify differences.

For example consider a large study comparing two treatments for high blood pressure; the results suggest that there is a statistically significant difference (p=0.023) in the amount by which blood pressure is lowered. This p-value relates to a difference of 3mmHg between the two treatments. Whilst the difference is statistically significant, it could be argued that a difference of 3mm Hg is not clinically important. This is supported by the 95% confidence interval of 2.3mm Hg to 3.7mm Hg.

Clinical and statistical significance

A clinically significant, or important, difference is one that is large enough to make a difference to patients or patient management. It should be noted that this is usually subjective and may differ depending upon who is making the judgement. Whilst a result may be statistically significant, it may not be clinically significant (relevant/important) and conversely an estimated difference that is clinically important may not be statistically significant: absence of evidence does not equate to evidence of absence. Figure 4 illustrates this as it displays the results of 7 theoretical studies comparing two treatments. It shows a range of possible estimates of the treatment difference and their confidence intervals.

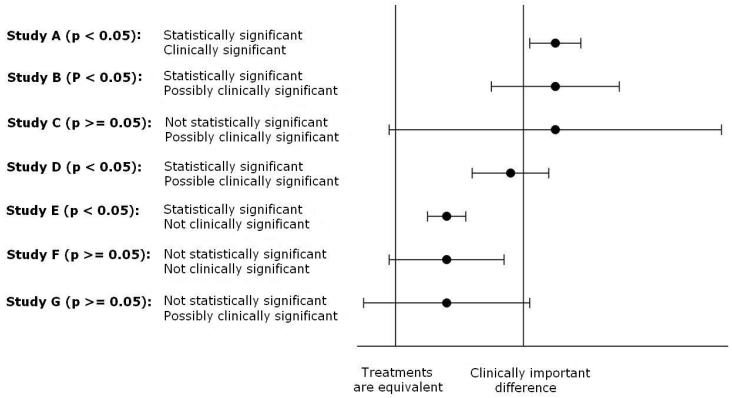


Figure 4: Relationship between statistical and clinical significance

6 Handling risk

It is often of interest to know about the risks associated with particular events or exposures. For example the risk of an adverse event for participants in a clinical trial comparing two treatments for wound infections.

Consider the following table where:
a = number of individuals who are exposed and have the event of interest
b = number of individuals not exposed who have the event of interest
c = number of individuals exposed who do not have the event of interest
d = number of individuals not exposed who do not have the event

Table 5:2 x 2 table illustrating the calculation of risk

	Exposure: Yes	Exposure: No	Total
Event:			
Yes	a	b	a+b
No	c	d	c+d
	a+c	b+d	n

Risk of event for exposed = a/(a+c)
Risk of event for unexposed = b/(b+d)

Table 6: Adverse event rates by treatment group

	Exposure: Treatment 1	Exposure: Treatment 2	Total
Event:			
Adverse event	45	25	70
No adverse event	905	1098	2003
	950	1123	2073

Risk of adverse event on treatment 1 = a/(a+c) = 45/950 =0.05
= 5 per 100 or 5%
Risk of adverse event on treatment 2 = b/(b+d) = 25/1123 = 0.02
= 2 per 100 or 2%

Absolute risk of an event is the probability of the event occurring (usually within a stated time period for a defined population):

Relative risk of a particular event for a given exposure is the ratio of the risk of the event occurring in the exposed group divided by the risk of the event occurring in the unexposed group:

$$\text{Relative risk} = \frac{a/(a+c)}{b/(b+d)} = \frac{a(b+d)}{b(a+c)} = (45*1123)/(25*950) = 2.13$$