Transformations

Summary

Many statistical methods require the data to fit assumptions of Normal distribution and uniform variance. When data do not fit these, one approach is to make them do so by a mathematical transformation. The most frequently used are the square root, the logarithm, and the reciprocal. These all reduce positive skewness and the extent to which variation increases with magnitude, the square root having the least effect and the reciprocal the greatest for each. The logarithm is the most often used. For a single, simple sample all will give us interpretable confidence intervals after transformations back to the original scale. For comparisons of means, only the logarithm can do this. Concentrations in blood are often analysed on the logarithmic scale, counts on the square root scale. There are other transformations in use, but they are seen rarely in the health research literature. Some data cannot be transformed satisfactorily and some, such as cost data, should not be. If data cannot be transformed, there are other strategies available, which do not require assumptions of Normal distribution or uniform variance.

1. The need for transformations

In Week 4 I described statistical methods in which we have to assume that data follow a Normal distribution with uniform variance. Later we shall meet other methods, regression and correlation, which require similar assumptions to be made about the data. Most analyses of continuous data in the health research literature are of this type. We should always check these assumptions. If the data meet the assumptions we can analyse the data as described. If they are not met, we have two possible strategies: we can use a method which does not require these assumptions, such as a rank-based method, or we can transform the data mathematically to make them fit the assumptions more closely. In this lecturer I describe the second approach. Instead of analysing the data as observed, we carry out a mathematical transformation first.

For example, Figure 1 shows serum cholesterol in stroke patients. As we have noted before, this does not follow a Normal distribution closely. This is shown by both the shape of the histogram and the Normal plot, in which the points should be close to the straight line. Figure 2 shows the same plots for the logarithms of cholesterol measurements. (See separate document *Logarithms*.) The logarithm follows a Normal distribution more closely than do the cholesterol measurements themselves. We could analyse the logarithm of serum cholesterol using methods which required the data to follow a Normal distribution. We call the logarithm of the cholesterol a **logarithmic transformation** of the data, or **log transformation** for short. We call the data without any transformation the **raw data**.

Even if a transformation does not produce a really good fit to the Normal distribution, it may still make the data much more amenable to analysis. Figure 3 shows a histogram and Normal plot for the area of venous ulcer at recruitment to the VenUS I trial, with the same for the log transformed area. The raw data have a very skew distribution and the small number of very large ulcers might lead to problems in analysis. Although the log transformed data are still skew, the skewness is much less and the data much easier to analyse.

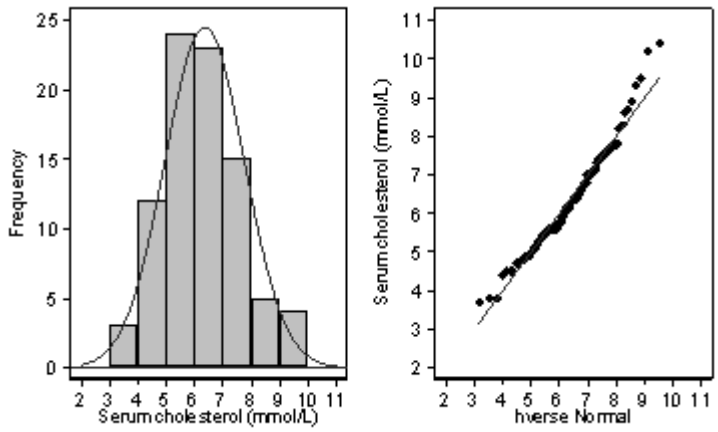

Figure 1. Histogram with Normal distribution curve and Normal plot for serum cholesterol in 86 stroke patients (data of Markus *et al.*, 1995)

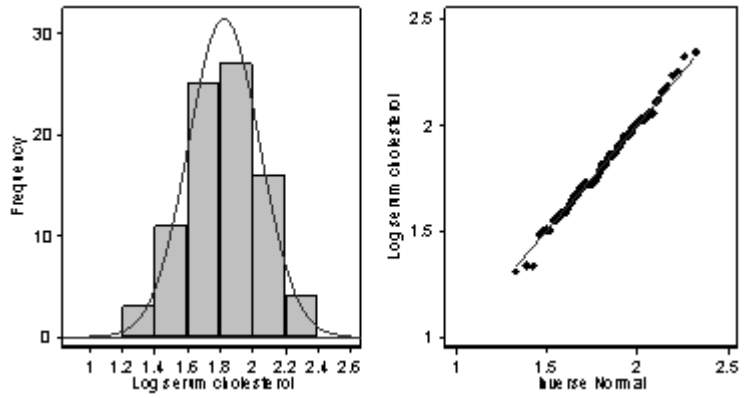

Figure 2. Histogram with Normal distribution curve and Normal plot for log transformed serum cholesterol in 86 stroke patients

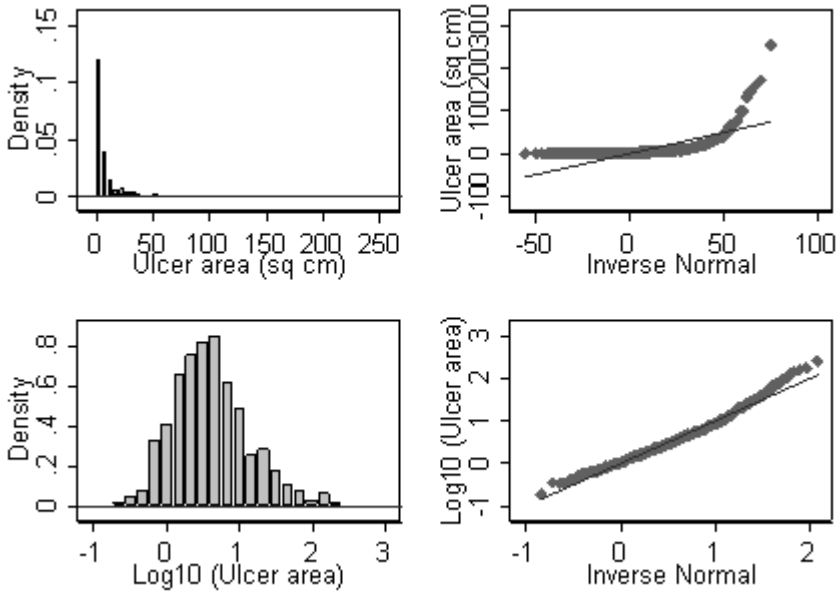

Figure 3. Histogram and Normal plot for area of venous ulcer at recruitment and log transformed ulcer area, VenUS I trial

Making a distribution more like the Normal is not the only reason for using a transformation. Figure 4 shows prostate specific antigen (PSA) for three groups of prostate patients: with benign conditions, with prostatitis, and with prostate cancer. One very high value makes it very difficult to see the structure in the rest of the data, although, as we would expect, we can see that the cancer group have the highest PSA values. A log transformation of the PSA gives a much clearer picture, shown in Figure 5. The variability is now much more similar in the three groups. Figure 6 shows a histogram and Normal plot for the raw data and the log transformed data. This shows the distribution of the within-group residuals, the difference between observed values and the mean of the group. The log transformation not only makes the variability more uniform but also makes the distribution closer to the Normal, thus meeting both assumptions: Normal distribution and uniform variance.

If we want to use the transformation only to make the scatter diagram easier to see, we sometimes use a logarithmic scale rather than the actual logarithms of the data (see *Logarithms*). Figure 7 shows this for the PSA data. The picture looks like the scatter plot for the log transformed PSA in Figure 5, but the vertical scale shows the original units.

The logarithm is not the only transformation used in the analysis of continuous data. Figure 8 shows arm lymphatic flow in patients with and without rheumatoid arthritis and oedema. The distribution is positively skew and the variability is clearly greater in the groups with greater lymphatic activity. A square root transformation has the effect of making the data less skew and making the variation more uniform. In these data, a log transformation proved to have too great an effect, making the distribution negatively skew, and so the square root of the data was used in the analysis (Kiely *et al.*, 1995).

2. Commonly used transformations for quantitative data

There are three commonly used transformations for quantitative data: the logarithm, the square root, and the reciprocal. (The **reciprocal** of a number is one divided by that number, hence the reciprocal of 2 is ½.) There are good mathematical reasons for these choices, Bland (2000) discusses them. They are based on the need to make variances uniform. If we have several groups of subjects and calculate the mean and variance for each group, we can plot variability against mean. We might have one of these situations:

- Variability and mean are unrelated. We do not usually have a problem and can treat the variances as uniform. We do not need a transformation.
- Variance is proportional to mean. A square root transformation should remove the relationship between variability and mean.
- Standard deviation is proportional to mean. A logarithmic transformation should remove the relationship between variability and mean.
- Standard deviation is proportional to the square of the mean. A reciprocal transformation should remove the relationship between variability and mean.

We call these transformations **variance-stabilising**, because their purpose is to make variances the same. For most data encountered in healthcare research, the first or third situation applies.

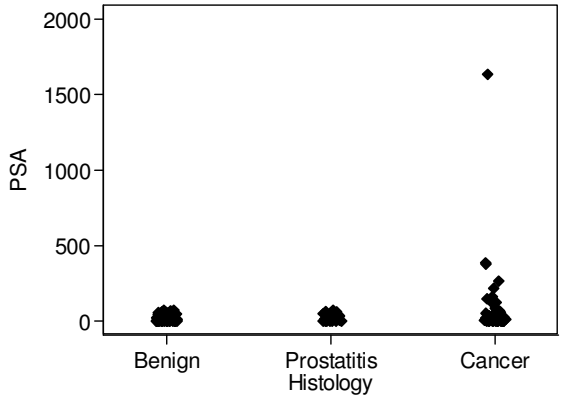

Figure 4. Prostate specific antigen (PSA) by prostate diagnosis (data of Cutting *et al.*, 1999)

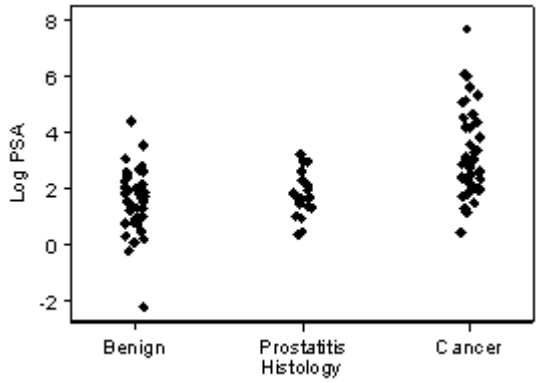

Figure 5. Log transformed PSA by prostate diagnosis

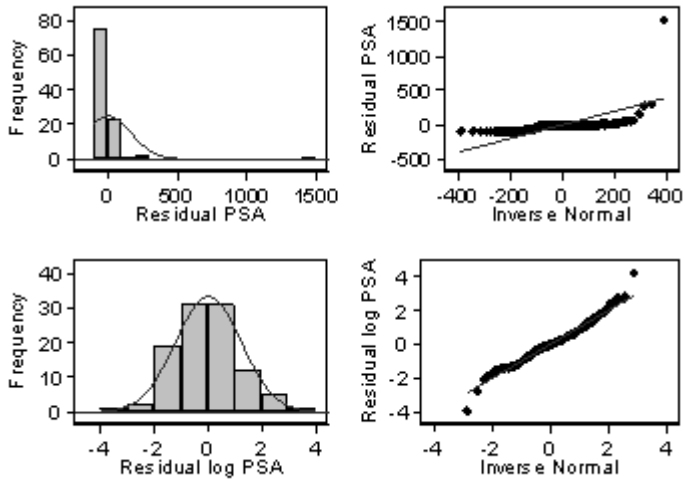

Figure 6. Histograms and Normal plots for the within-group residuals for the raw PSA data and the log transformed PSA data

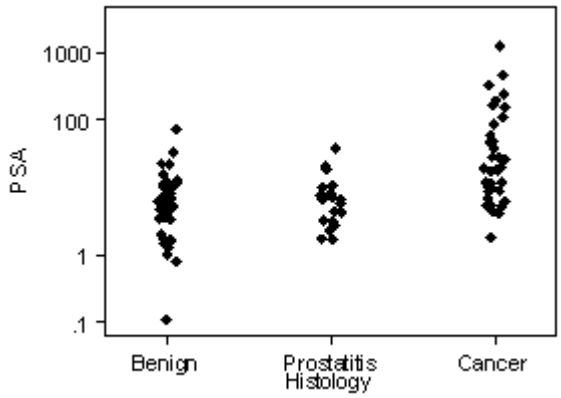

Figure 7. PSA by prostate diagnosis, shown on a logarithmic scale.

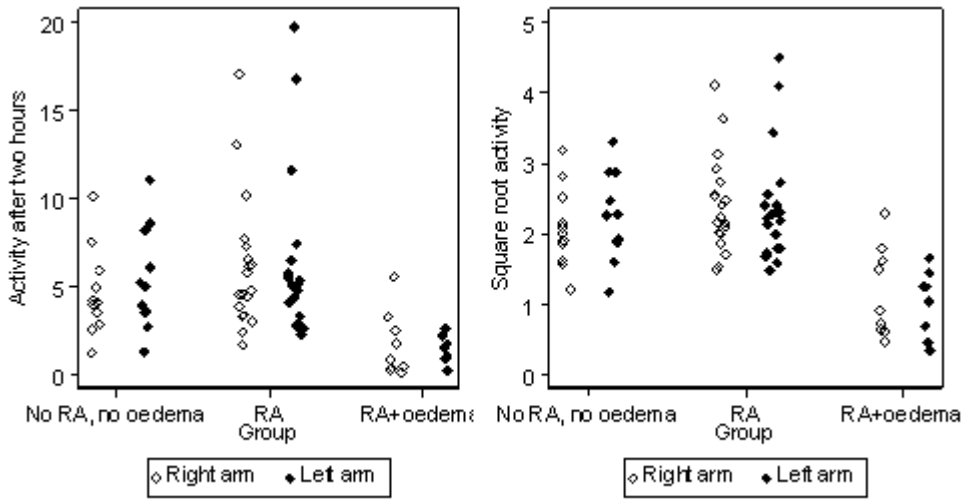

Figure 8. Arm lymphatic flow in rheumatoid arthritis with oedema (data of Kiely *et al.*, 1995)

Table 1. Biceps skinfold thickness (mm) in two groups of patients

Crohn's Disease				Coeliac Disease	
1.8	2.8	4.2	6.2	1.8	3.8
2.2	3.2	4.4	6.6	2.0	4.2
2.4	3.6	4.8	7.0	2.0	5.4
2.5	3.8	5.6	10.0	2.0	7.6
2.8	4.0	6.0	10.4	3.0	

Variance-stabilising transformations also tend to make distributions Normal. There is a mathematical reason for this, as for so much in statistics. It can be shown that if we take several samples from the same population, the means and variances of these samples will be independent if and only if the distribution is Normal. This means that uniform variance tends to go with a Normal Distribution. A transformation which makes variance uniform will often also make data follow a Normal distribution and *vice versa*.

There are many other transformations which could be used, but you see them very rarely. We shall meet one other, the logistic transformation used for dichotomous data, in Week 7. By far the most frequently used is the logarithm. This is particularly useful for concentrations of substances in blood. The reason for this is that blood is very dynamic, with reactions happening continuously. Many of the substances we measure are part of a metabolic chain, both being synthesised and metabolised to something else. The rates at which these reactions happen depends on the amounts of other substances in the blood and the consequence is that the various factors which determine the concentration of the substance are multiplied together. Multiplying and dividing tends to produce skew distributions. If we take the logarithm of several numbers multiplied together we get the sum of their logarithms. So log transformation produces something where the various influences are added together and addition tends to produce a Normal distribution. The square root is best for fairly weak relationships between variability and magnitude, i.e. variance proportional to mean or standard deviation proportional to the square root of the mean. The logarithm is next, for standard deviation proportional to the mean, and the reciprocal is best for very strong relationships, where the standard deviation is proportional to the square of the mean. In the same way, the square root removes the least amount of skewness and reciprocal the most.

The square root can be used for variables which are greater than or equal to zero, the log and the reciprocal can only be used for variables which are strictly greater than zero, because neither the logarithm nor the reciprocal of zero are defined. We shall look at what to do with zero observations in Section 6.

Which transformation should we use for what kind of data? For physical body measurements, like limb length or peak expiratory flow, we often need use only the raw data. For concentrations measured in blood or urine, we usually try the log first, then if this is insufficient try the reciprocal. For counts, the square root is usually the first thing to try. There are methods to determine which transformation will best fit the data, but trial and error, with scatter plots, histograms and Normal plots to check the shape of the distribution and relationship between variability and magnitude, are usually much quicker because the computer can produce them almost instantaneously.

3. Are transformations cheating?

At about this point, someone will ask ‘Aren’t transformations cheating?’. Data transformation would be cheating if we tried several different transformations until we found the one which gave the result we wanted, just as it would be if we tried several different tests of significance and chose the one which gave the result nearest to what we wanted, or compared treatment groups in a clinical trial using different outcome variables until we found one which gave a significant difference. Such approaches are cheating because the P values and confidence intervals we get are wrong. However, it is not cheating if we decide on the analysis we want to use before we see its result and then stick to it.

It should be remembered that the linear scale is not the only scale which we use for measurements. Some variables are always measured on a log scale. Well-known examples are the decibel scale for measuring sound intensity and the Richter scale for measuring earthquakes. The reason we use logarithmic scales for these is that the range of energy involved is huge and a difference which is easily perceived at low levels would not be noticed at high. If you are sitting in a library and someone said 'Hello' to you, you would certainly notice it, but if you were standing next to a jumbo jet preparing for take-off or in a disco you would not. Many healthcare professionals see several measurements of acidity every working day, but they do worry about pH being a logarithmic scale, and a logarithm with its minus sign removed at that. It is simply the most convenient scale on which to measure.

Should we measure the power of spectacle lenses by focal length or in dioptres? We use dioptres in ophthalmology, which is the reciprocal transformation of the focal length in metres. Concentrations are measured in units of solute in contained in one unit of solvent, but this is an arbitrary choice. We could measure in units of solvent required to contain one unit of solute — the reciprocal. Similarly, we measure car speed in miles or kilometres per hour, but we could just as easily use the number of hours or minutes required to go one mile. We measure fuel consumption like this, in miles per gallon or kilometres per litre rather than gallons per mile or litres per kilometre.

We often choose scales of measurement for convenience, but they are just that, choices. There is often no overwhelming reason to use one scale rather than another. In the same way, when we use a transformation, we are choosing the scale for ease of statistical analysis, not to get the answer we want.

4. Transformations for a single sample

As we have seen, for the serum cholesterol in stroke patients data, the log transformation gives a good fit to the Normal. What happens if we analyse the logarithm of serum cholesterol then try to transform back to the natural scale?

For the raw data, serum cholesterol: mean = 6.34, SD = 1.40.

For log (base e) serum cholesterol: mean = 1.82, SD = 0.22.

If we take the mean on the transformed scale and back-transform by taking the antilog, we get exp(1.82) = 6.17. This is less than the mean for the raw data. The antilog of the mean log is not the same as the untransformed arithmetic mean.

In fact, it is the **geometric mean**, which is found by multiplying all the observations and taking the n'th root. (It is called geometric because if we have just two numbers we could draw a rectangle with those two numbers as the lengths of the long and short sides. The geometric mean is the side of a square which has the same area as this rectangle.) Now, if we add the logs of two numbers we get the log of their product. Thus when we add the logs of a sample of observations together we get the log of their product. If we multiply the log of a number by a second number, we get the log of the first raised to the power of the second. So if we divide the log by n, we get the log of the n'th root. Thus the mean of the logs is the log of the geometric mean.

What about the units for the geometric mean? If cholesterol is measured in mmol/L, the log of a single observation is the log of a measurement in mmol/L. The sum of n logs is the log of the product of n measurements in mmol/L and is the log of a

measurement in mmol/L to the power n. The n'th root is thus again the log of a number in mmol/L and the antilog is back in the original units, mmol/L.

The antilog of the standard deviation is not measured in mmol/L. To find a standard deviation, we calculate the differences between each observation and the mean, square and add. On the log scale, we take the difference between each log transformed observation and subtract the log geometric mean. We have the difference between the log of two numbers each measured in mmol/L, giving the log of their ratio, which is the log of a dimensionless pure number. We cannot transform the standard deviation back to the original scale.

If we want to use the standard deviation, it is easiest to do all calculations on the transformed scale and transform back, if necessary, at the end. For example, to estimate the 95% confidence interval for the geometric mean, we find the confidence interval on the transformed scale. On the log scale the mean is 1.8235 with standard error of 0.0235. This standard error is calculated from the standard deviation, which is a pure number without dimensions, and the sample size, which is also a pure number. It, too, is a pure number without dimensions. The 95% confidence interval for the mean is

$$1.8235 - 1.96 \times 0.0235 \text{ to } 1.8235 + 1.96 \times 0.0235 = 1.777 \text{ to } 1.870.$$

If we antilog these limits we get 5.91 to 6.49. To get the confidence limits we took the log of something in mmol/L, the mean, and added or subtracted the log of a pure number, the standard error multiplied by 1.96. On the natural scale we have taken something in mmol/L and multiplied or divided by a pure number. We therefore we still have something in mmol/L. The 95% confidence interval for the geometric mean is therefore 5.91 to 6.49 mmol/L.

For the arithmetic mean, using the raw, untransformed data we get 6.04 to 6.64 mmol/L. This interval is slightly wider than for the geometric mean. In highly skew distributions, unlike serum cholesterol, the extreme observations have a large influence on the arithmetic mean, making it more prone to sampling error and the confidence interval for the arithmetic mean is usually quite a lot wider.

In the same way we can estimate centiles on the transformed scale and then transformed back. In the a Normal distribution the central 95% of observations are within 1.96 standard deviations from the mean. For log serum cholesterol, this is 1.396 to 2.251. The antilog is 4.04 to 9.50 mmol/L.

We can do this for square root transformed and reciprocal transformed data, too. If we do all the calculations on the transformed scale and transform back only at the end, we will be back in the original units. The mean calculated in this way using a reciprocal transformations also has a special name, the **harmonic mean**.

5. Transformations when comparing two groups

Table 1 shows measurements of biceps skinfold thickness compared for two groups of patients, with Crohn's disease and Coeliac disease. We ask whether there is any difference in skinfold between patients with these diagnoses and what it might be.

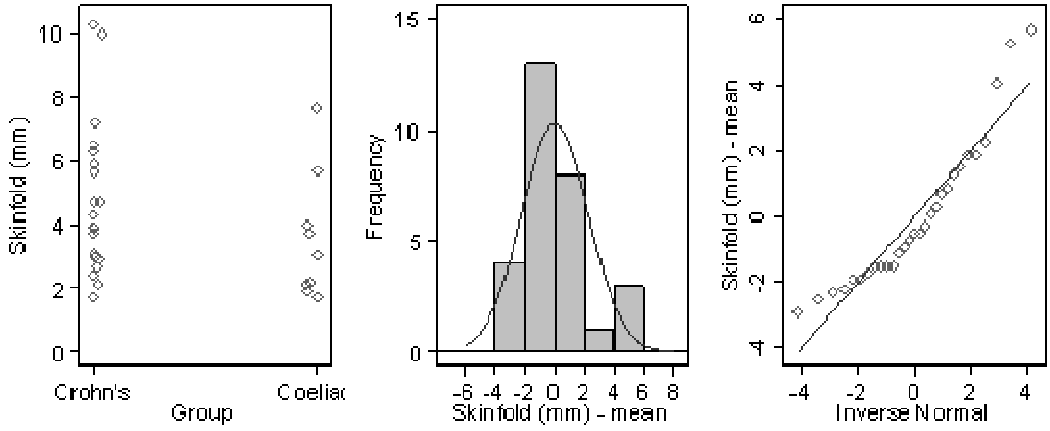

Figure 9. Untransformed biceps skinfold thickness for Crohn's disease and Coeliac disease patients, with histogram and Normal plot of residuals

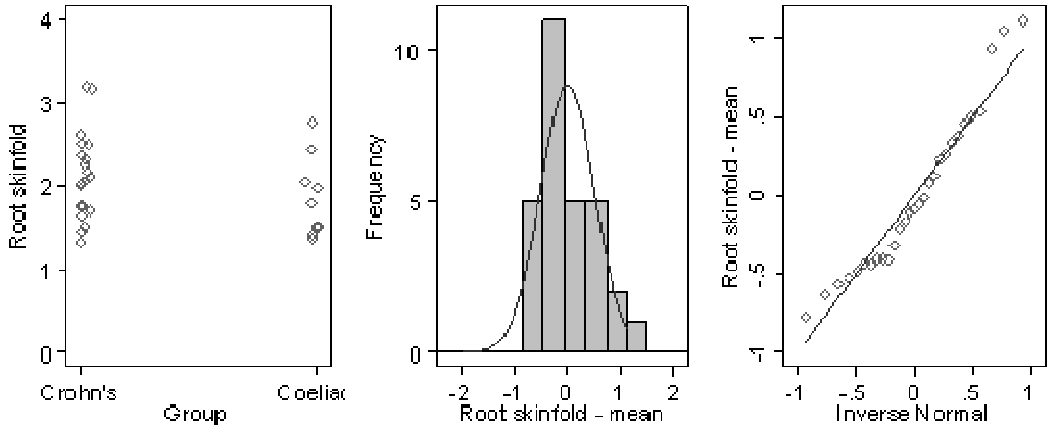

Figure 10. Square root transformed biceps skinfold thickness

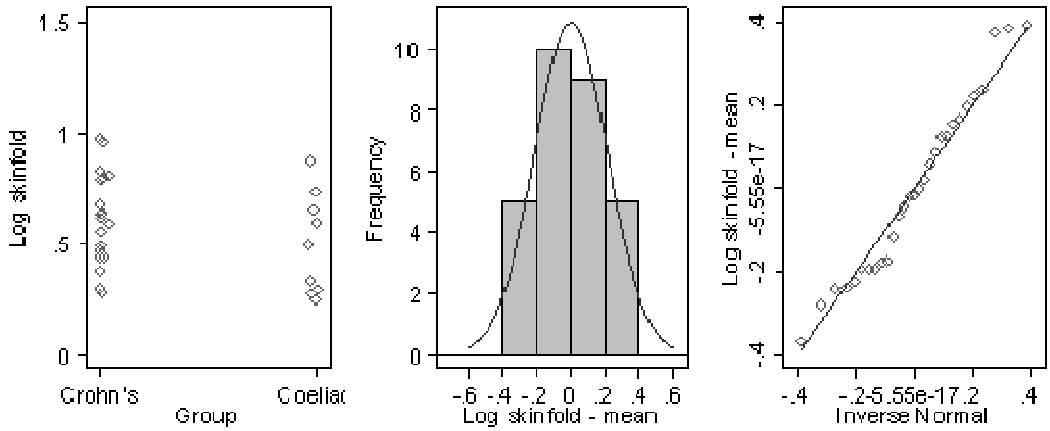

Figure 11. Log transformed biceps skinfold thickness

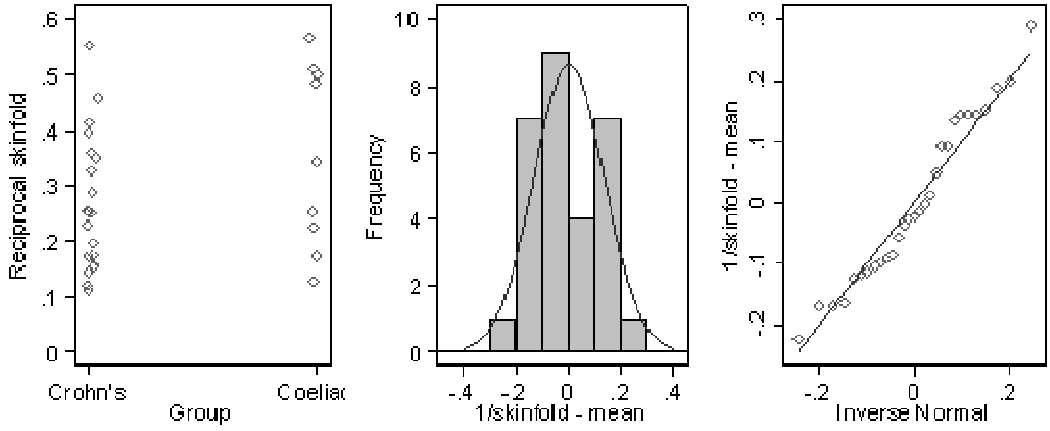

Figure 12. Reciprocal transformed biceps skinfold thickness

Table 2. Comparison of mean biceps skinfold between Crohn's disease and Coeliac disease patients using different transformations

Transform-ation	Two sample t test, 27 d.f. t	P	95% confidence interval for difference on transformed scale	Variance ratio larger/ smaller
None	1.28	0.21	−0.71mm to 3.07mm	1.52
Square root	1.38	0.18	−0.140 to 0.714	1.16
Logarithm	1.48	0.15	−0.114 to 0.706	1.10
Reciprocal	−1.65	0.11	−0.203 to 0.022	1.63

Table 3. C-reactive protein (CRP) (mg/L) before and after debridement of wounds using larval therapy (Steenmoorle and Julena, 2004)

CRP before larvae	CRP after larvae	CRP difference, after minus before	CRP average of before and after
3	2	–1	2.5
5	9	4	7.0
16	36	20	26.0
17	5	–12	11.0
26	218	192	122.0
29	193	164	111.0
29	24	–5	26.5
30	6	–24	18.0
32	77	45	54.5
47	0	–47	23.5
61	19	–42	40.0
87	42	–45	64.5
123	26	–97	74.5
124	68	–56	96.0
163	59	–104	111.0
227	26	–201	126.5

Figure 9 shows the distribution of biceps skinfold. This is clearly positively skew. Figure 10 shows the same data after square root transformation. This is still skew, though less so than the untransformed data. Figure 11 shows the effect of a log transformation. The distribution is now more symmetrical and the Normal plot is closer to a straight line. Figure 12 shows the effect of a reciprocal transformation, which looks fairly similar to the log. Any of the transformations would be an improvement on the raw data.

Table 2 shows the result of a two sample t test and confidence interval for the raw data and the transformations. The transformed data clearly gives a better test of significance than the raw data, in that the P values are smaller.

The confidence intervals for the transformed data are more difficult to interpret. The confidence limits for the difference between means cannot be transformed back to the original scale.

For the square root transformation, the lower limit is negative. We can square this, which would give a positive number, and this will happen whatever the limits because all squares are positive. Hence squaring the limits will not give a 95% confidence interval for the difference in biceps skinfold. The confidence interval must include the null hypothesis value, which would be zero. The same problem arises with the logarithmic transformation, all antilogs are positive. For the reciprocal, we could transform back, but what would this mean? The closer the limits are on the reciprocal scale, the further apart they will be on the natural scale. The upper limit for the reciprocal is very small (0.022) with reciprocal 45.5. The difference clearly could not be 45.5 mm, as all the observations are much smaller than this. The null hypothesis value, zero on the reciprocal scale, transforms back to infinity! A point to watch out for is that the square root and logarithm keep differences in the same direction as the raw data, the reciprocal reverses the direction.

Confidence limits for the difference cannot be transformed back to the original scale. However, the logarithm does give interpretable results (0.89 to 2.03) but these are not limits for the difference in millimetres. They do not contain zero yet the difference is not significant. The back-transformed 95% confidence interval using the log transformation, 0.89 to 2.03, are the 95% confidence limits for the ratio of the Crohn's disease mean to the Coeliac disease mean. When we take the difference between the logarithms of the two geometric means, we get the logarithm of their ratio, not of their difference.

Transformed data give us only a P value when comparing groups, unless we use the log, in which case we can get confidence intervals for ratios.

6. Transformations for paired data

Table 3 shows C-reactive protein (CRP) (mg/L) before and after debridement of hard to heal wounds using larval therapy (Steenmoorle and Julena, 2004). Wounds with large CRP showed much more variable CRP than did wounds with small CRP. The data are shown graphically in Figure 13. This is shown by the plot of difference against average; the differences clearly get larger as the magnitude of CRP increases. The Normal plot is not a particularly good fit, although there is no clear curve indicating skewness. Rather, the curve is first convex and rises from below to above the straight line then drops below, indicating a long tail on the left, then concave and falls from above to below the line and then rises above, indicating a long tail on the right. This distribution is symmetrical but with long tails in either direction.

Table 4. Square root transformed C-reactive protein (CRP) (mg/L) before and after debridement of wounds

√CRP before larvae	√CRP after larvae	√CRP difference, after minus before
√3 = 1.73	√2 = 1.41	–0.32
√5 = 2.24	√9 = 3.00	0.76
√16 = 4.00	√36 = 6.00	2.00
√17 = 4.12	√5 = 2.24	–1.88
√26 = 5.10	√218 = 14.76	9.66
√29 = 5.39	√193 = 13.89	8.50
√29 = 5.39	√24 = 4. 90	–0.49
√30 = 5.48	√6 = 2.45	–3.03
√32 = 5.66	√77 = 8.77	3.11
√47 = 6.86	√0 = 0.00	–6.86
√61 = 7.81	√19 = 4.36	–3.45
√87 = 9.33	√42 = 6.48	–2.85
√123 = 11.09	√26 = 5.10	–5.99
√124 = 11.14	√68 = 8.25	–2.89
√167 = 12.77	√59 = 7.68	–5.09
√227 = 15.07	√26 = 5.10	–9.97

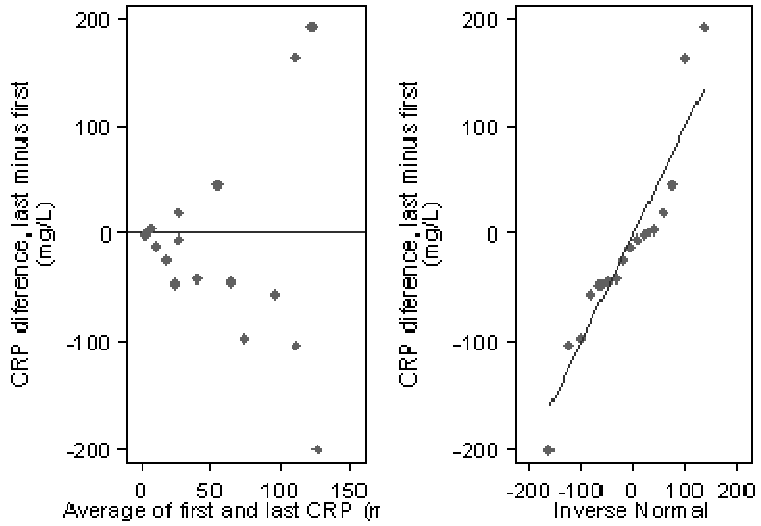

Figure 13. Difference against average and Normal plot for differences for CRP in 16 patients with hard to heal wounds treated with larval therapy

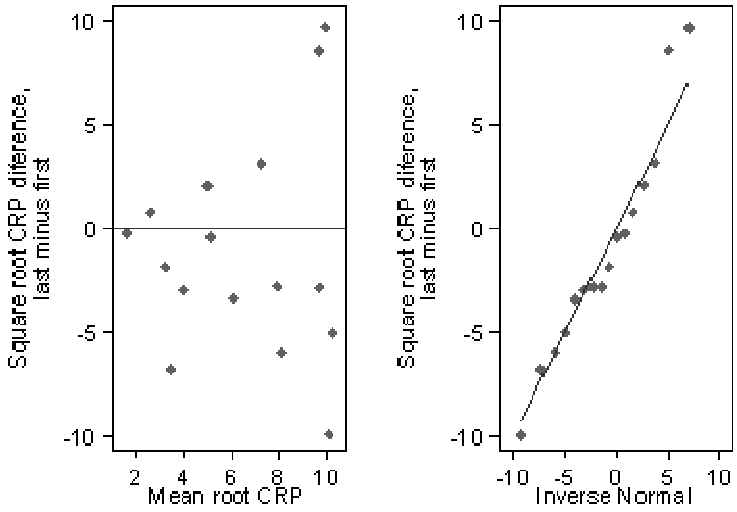

Figure 14. Difference against average and Normal plot for differences for square root transformed CRP

Table 5. Log plus one transformed C-reactive protein (CRP) (mg/L) before and after debridement of wounds

Log (CRP +1) before larvae	Log (CRP +1) after larvae	Log (CRP +1) difference, after minus before
1.386	1.099	–0.287
1.792	2.303	0.511
2.833	3.611	0.778
2.890	1.792	–1.098
3.296	5.389	2.093
3.401	3.219	–0.182
3.401	5.268	1.867
3.434	1.946	–1.488
3.497	4.357	0.860
3.871	0.000	–3.871
4.127	2.996	–1.131
4.477	3.761	–0.716
4.820	3.296	–1.524
4.828	4.234	–0.594
5.100	4.094	–1.006
5.429	3.296	–2.133

We would like to transform the data to make them fit the assumptions required for the paired t method. Differences are often negative, as several are here. We cannot log or square root negative numbers. Mathematically, there is nothing to stop us taking reciprocals, but the reciprocal has what we call a discontinuity at zero. As we go towards zero from the positive end, we get larger and larger positive numbers and zero has no reciprocal, it is an infinitely large number. As we go towards zero from the negative end, we get larger and larger negative numbers. So at zero the reciprocal switches from an infinitely large negative number to an infinitely large positive number. We cannot use any of these transformations for the differences. Instead, we transform the original observations then calculate the differences from the transformed observations.

Table 5 shows the square root transformed data. A plot of difference against average is shown in Figure 14. This is an improvement on Figure 13, but there is still some relationship apparent.

CRP is a concentration in blood and so we might expect a log transformation to be the best bet. Also, the conical shape of the difference against average plot suggests (to the experienced eye) that the standard deviation will be proportional to the mean. There is a problem, however. We have a zero observation and zero has no logarithm. What we usually do is to add a small constant to everything. This should be somewhere between zero and the smallest non-zero observation. If there is no reason not to do so, we usually choose 1.0 for this constant. Table 5 shows the transformed data and Figure 15 shows the scatter plot of difference against average and the Normal plot for the differences. There is little to suggest a relationship between difference and magnitude, as the largest difference is for a subject with one of the smallest averages and the next largest is for a subject with one of the largest averages. In the Normal plot, the points lie much closer to the straight line than in Figure B. The transformed data appear to fit the assumptions needed for the paired t method much better than the raw data.

If we apply the paired t method to Table 5, we get mean difference = –0.495, 95% CI = –1.299 to 0.308, P = 0.2. How can we interpret this? If we had used a simple log transformation, we could antilog to get exp(–0.495) = 0.61, and say that we estimate the mean CRP to fall to 61% of the pre-treatment value, with 95% CI 27% to 136%. Does our addition of 1.0 to everything before log transformation affect this? The answer is: not much. The antilog transformation which gives the ratio becomes approximate rather than exact, but the effect of adding one before logging is not great. We will illustrate this using the data of Table 3. If we drop the zero observation, we have a sample with no zeros and we can log the data. (This is purely to illustrate the properties of the transformation, we would not do this as a method of analysis.) This gives a difference estimate = –0.291. If we use the add one then log transformation for the same data, omitting the case with a zero, we get difference = –0.270. If we take the antilogs of each of these we get exp(–0.291) = 0.75, exp(–0.270) = 0.76, so these are very similar. After the add constant then log transformation, we can antilog the estimated difference and its confidence interval and still interpret these as estimates of the ratios of the original observations.

7. Can all data be transformed?

Not all data can be transformed successfully. Sometimes we have very long tails at both ends of the distribution, which makes transformation by log, square root or reciprocal ineffective. For example, Figure 16 shows the distribution of blood sodium in ITU patients. This is fairly symmetrical, but has longer tails than a Normal distribution. The shape of the Normal plot is first convex then concave, reflecting this, just like the differences for the raw CRP data in Figure B. We can often ignore this departure from the Normal distribution, for example when using the two-sample t method, but not always. If we were trying to estimate a 95% range, for example, we might not get a very reliable answer.

Sometimes we have a bimodal distribution, which makes transformation by log, square root or reciprocal ineffective. Figure 17 shows systolic blood pressure in a sample of ITU patients. This is clearly bimodal. The serpentine Normal plot reflects this. None of the usual transformations will affect this and we would still have a bimodal distribution. We should not ignore this departure from the Normal distribution.

Sometimes we have a large number of identical observations, which will all transform to the same value whatever transformation we use. These are often at one extreme of the distribution, usually at zero. For example, Figure 18 shows the distribution of coronary artery calcium in a large group of patients. More than half of these observations were equal at zero. Any transformation would leave half the observations with the same value, at the extreme of the distribution. It is impossible to transform these data to a Normal distribution.

What can we do if we cannot transform data to a suitable form? If the sample is large enough, we can ignore the distribution and use large sample z methods. For small samples we can do the same, and hope that the effect of the departure from assumptions is to make confidence intervals too wide and P values too big, rather than the other way round. This is usually the effect of skewness, but we should always be very cautious in drawing conclusions. It is usually safer to use methods that do not require such assumptions. These include the non-parametric methods, such as the Mann-Whitney U test and Wilcoxon matched-pairs test, which are beyond the scope of this course. These methods will give us a valid significance test, but usually no confidence interval.

Are there data which should not be transformed? Sometimes we are interested in the data in the actual units only. Cost data is a good example. Costs of treatment usually have distributions which are highly skew to the right. However, we need to estimate the difference in mean costs in pounds. No other scale is of interest. We should not transform such data. We rely on large sample comparisons or on methods which do not involve any distributions. Economists often use a group of methods which do not rely on any assumptions about distributions called bootstrap or resampling methods.

8. Pitfalls of transformations

The most obvious pitfall is not using a transformation when one is indicated, but instead using methods requiring Normal assumptions on data which clearly do not meet them.

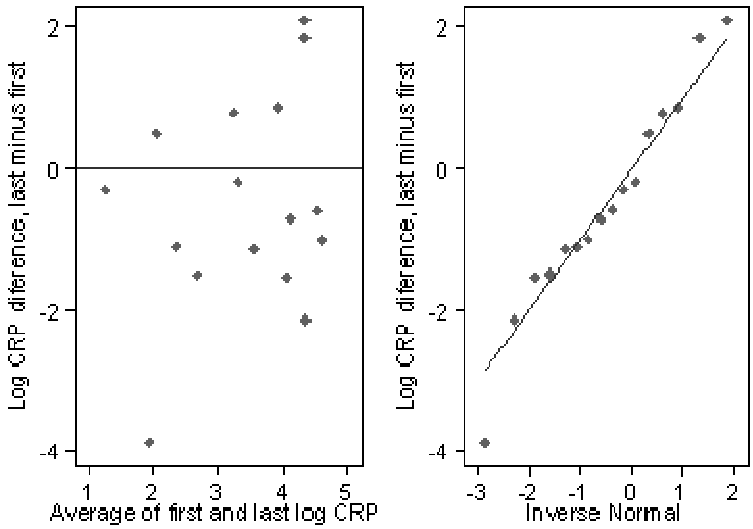

Figure 15. Difference against average and Normal plot for differences for log plus one transformed CRP

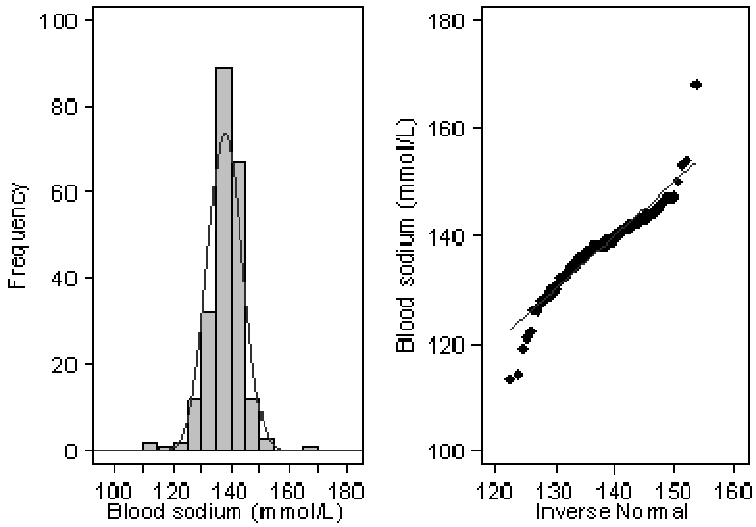

Figure 16. Blood sodium in 221 ITU patients (data of Friedland *et al.*, 1996)

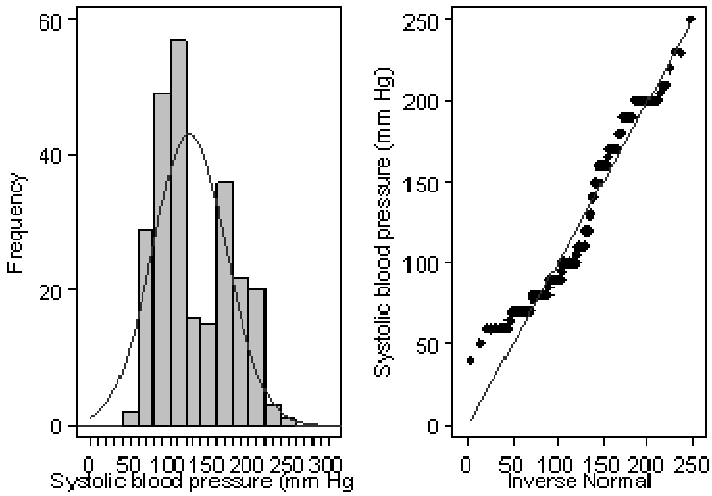

Figure 17. Systolic blood pressure in 250 ITU patients (data of Friedland *et al.*, 1996)

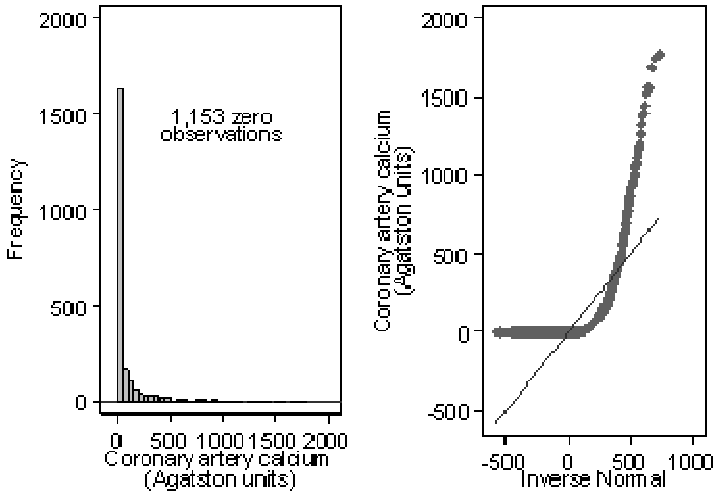

Figure 18. Coronary artery calcium in 2217 subjects (Data of Sevrukov *et al.*, 2005)

References

Bland M. (2000) *An Introduction to Medical Statistics.* Oxford University Press.

Cutting CW, Hunt C, Nisbet JA, Bland JM, Dalgleish AG, Kirby RS. (1999) Serum insulin-like growth factor-1 is not a useful marker of prostate cancer. *BJU International* **83**, 996-999.

Friedland JS, Porter JC, Daryanani S, Bland JM, Screaton NJ, Vesely MJJ, Griffin GE, Bennett ED, Remick DG. (1996) Plasma proinflammatory cytokine concentrations, Acute Physiology and Chronic Health Evaluation (APACHE) III scores and survival in patients in an intensive care unit. *Critical Care Medicine* **24**, 1775-81.

Kiely PDW, Bland JM, Joseph AEA, Mortimer PS, Bourke BE. (1995) Upper limb lymphatic function in inflamatory arthritis. *Journal of Rheumatology*, **22**, 214-217.

Markus HS, Barley J, Lunt R., Bland JM, Jeffery S, Carter ND, Brown MM. (1995) Angiotensin-converting enzyme gene deletion polymorphism: a new risk factor for lacunar stroke but not carotid atheroma. *Stroke* **26**, 1329-33

Sevrukov AB, Bland JM, Kondos GT. (2005) Serial electron beam CT measurements of coronary artery calcium: Has your patient's calcium score actually changed? *American Journal of Roentgenology* **185**, 1546-1553.