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.

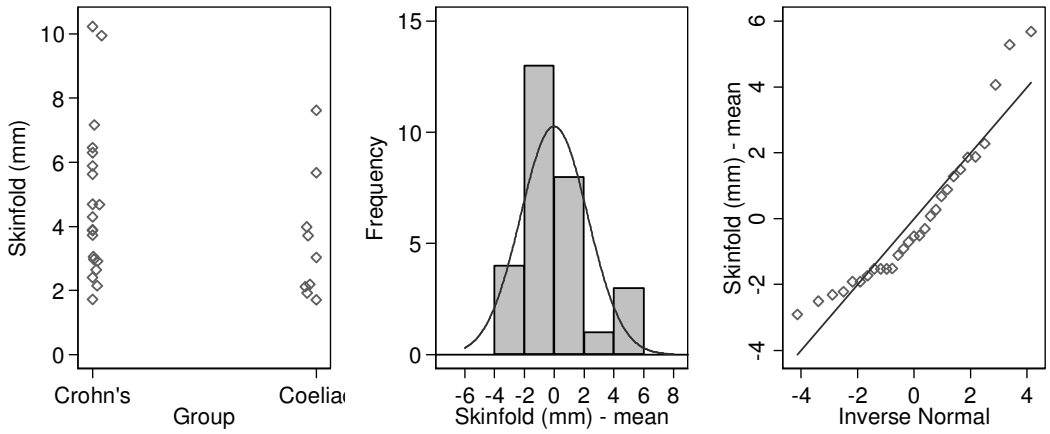

Figure 9. Untransformed biceps skinfold thickness for Crohn's disease and Coeliac disease patients, with histogram and Normal plot of residuals

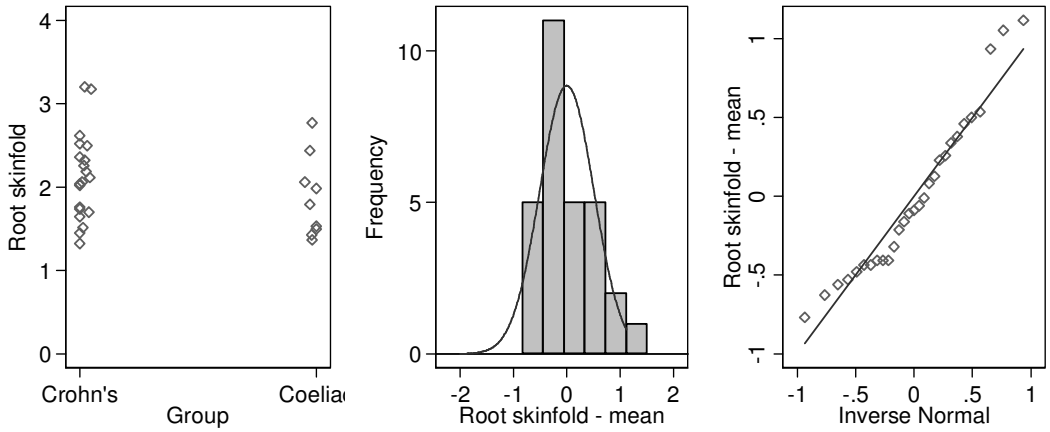

Figure 10. Square root transformed biceps skinfold thickness

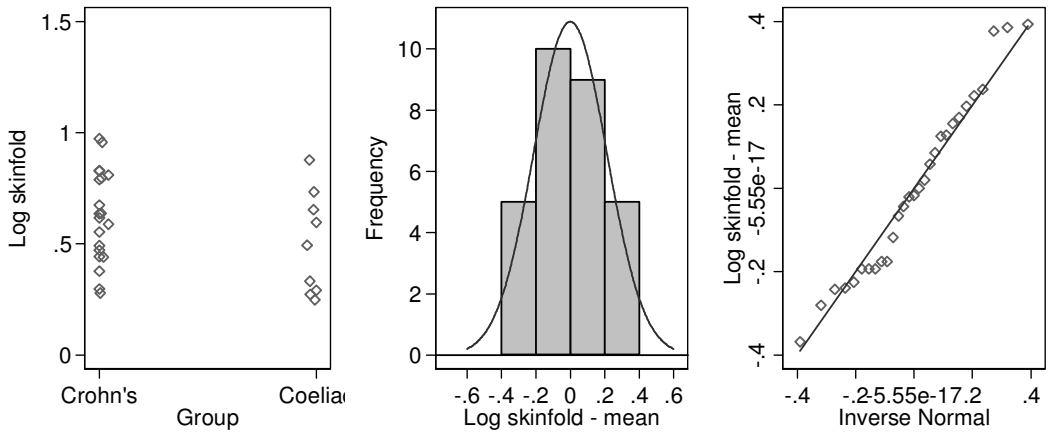

Figure 11. Log transformed biceps skinfold thickness

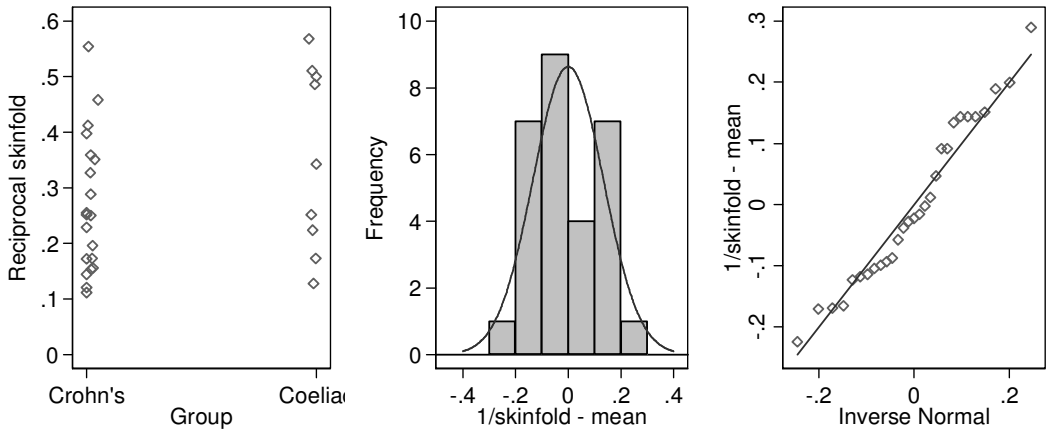

Figure 12. Reciprocal transformed biceps skinfold thickness

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

Transformation	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

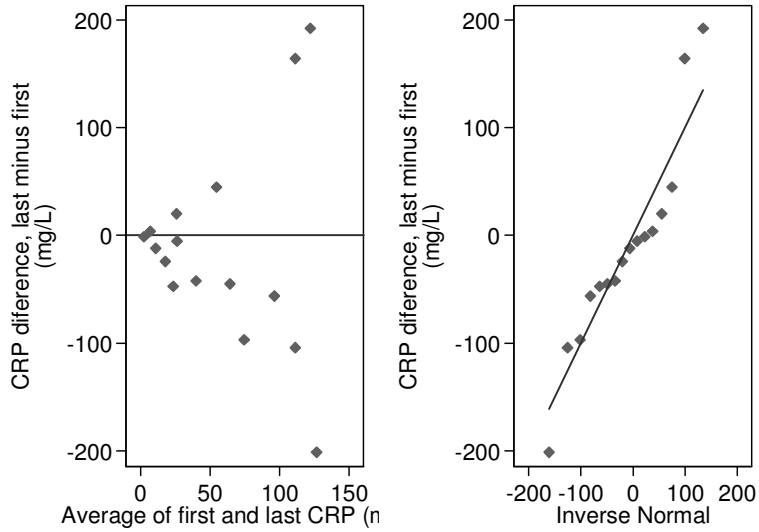

Figure 13. Difference against average and Normal plot for differences for CRP in 16 patients with hard to heal wounds treated with larval therapy