161

Herbison P[1]
1. University of Otago

ANALYSING COUNT DATA

Hypothesis / aims of study
Many outcomes used in research into incontinence are count outcomes, such as the number of micturitions and the number of leakage episodes. This data is often analysed in a less than optimum way. This abstract will mention the different methods of analysis of this data and use data from a randomised controlled trial as an example.

Study design, materials and methods
The different potential methods of analysing count data are discussed. They are then illustrated by using data from two arms of a pilot study. Ethical approval had been given for the trial by the local ethics committee.

Results
Count data can be divided into two groups, either with a large mean (such as pulse rate) or a low mean (such as episodes of incontinence in 24 hours). Count data with a large mean can be treated as ordinary continuous data, and if it is close to following a normal distribution, can be analysed with the usual parametric statistics, so will not be considered further.

The distribution of count data with a low mean almost certainly does not approximate a normal distribution. This is often because it is truncated at zero, that is, negative values are impossible, and is skewed to the right. Because of this many people use non-parametric statistics to analyse such data. This is not incorrect, but does have some disadvantages.

The three main ways of analysing count data with a low mean are:

1. Ignore the distribution and use usual methods such as the t-test
2. Use nonparametric statistics
3. Use a method that uses the likely distribution of the data such as poisson regression.

None of these methods are difficult to use as they are available in most statistical software. The advantages and disadvantages are summarised in table 1.

Table 1. Advantages and disadvantages of different methods of analysis

Method	**Advantages**	**Disadvantages**
Normal parametric	Everyone knows about it. Can be shown to produce quite good answers in this sort of data, especially if the numbers in the groups are large. Can easily adjust for confounders, such as baseline values.	It is wrong. For some deviations from normality can produce incorrect results.
Nonparametric	Everyone knows about it Does not depend on the distribution	Compares the whole distribution, not just the medians. Has less power. Difficult to adjust for confounders.
Poisson parametric	Uses the real distribution. Can easily adjust for confounders.	It is not often used. Results are relative rather than absolute.

The distributions of the numbers of leakage episodes for the two groups are shown in the figure. These are clearly not normally distributed. The results of the three types of analysis are given in table 2.
Figure. Distributions of leakage episodes at the end of the study.

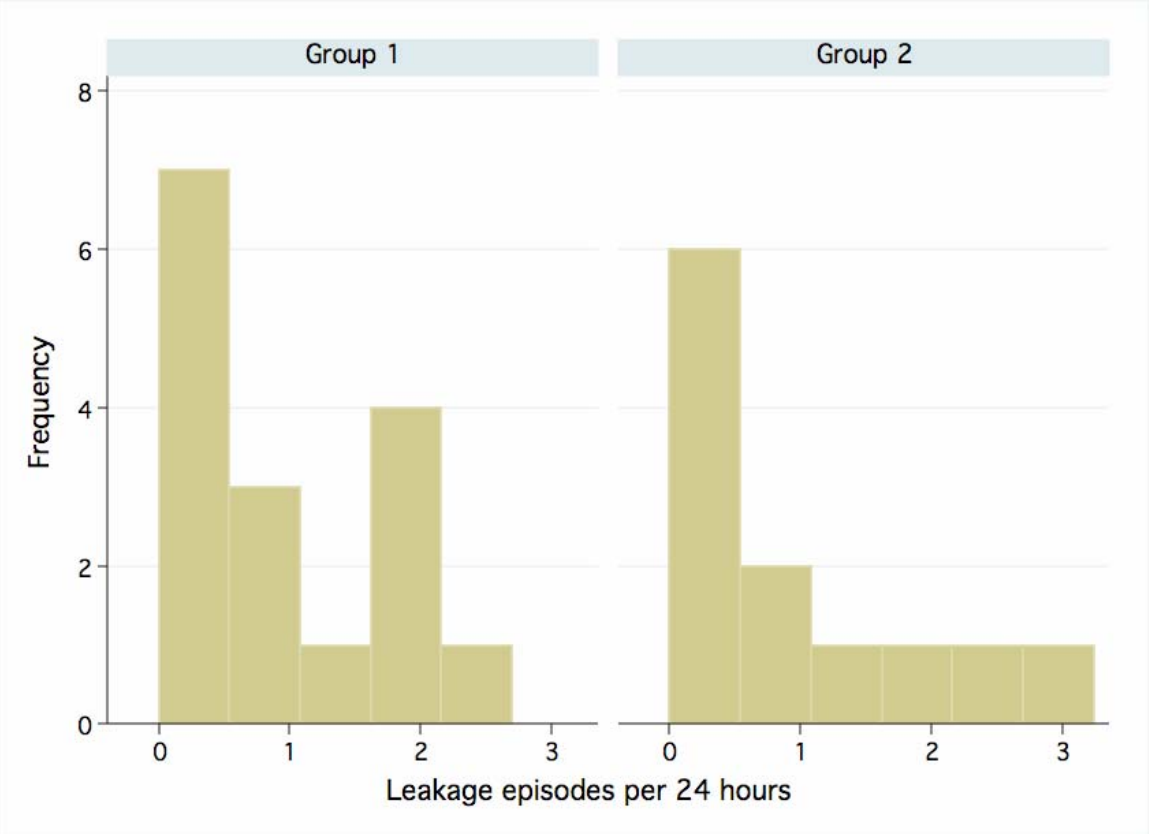

Table 2. Results of the three different analyses.

Method	Details	p-value
Normal parametric	Group 1: mean 0.91 (SD 0.22) Group 2: mean 0.86 (SD 0.28)	0.878
Nonparametric	Group 1: median 0.85 (IQR 0 to 1.9) Group 2: median 0.65 (IQR 0 to 1.6)	0.775
Poisson parametric	Incidence rate ratio = 0.94, 95% CI 0.42 to 2.09	0.880

SD = standard deviation
IQR = inter quartile range
CI = confidence interval

The incidence rate ratio of 0.94 means a reduction in the rate of leakages by 6% in group 2.

Interpretation of results

These three methods of analysis give much the same results for this set of data, even though there are relatively few observations in each group. This is what would be expected to occur in most cases but simulations would be needed to confirm when this would happen and the situations in which one method would be preferable. Both the normal parametric and the Poisson parametric approach are more versatile in that they can easily be adjusted for confounders, such as the baseline levels of leakage. In addition the poisson method can be extended if the data do not fit the poisson distribution well. For example, negative binomial regression can be used if there are more people with higher numbers of leakage episodes than expected, and the zero inflated poisson used if there are more people with no leakage than expected.

Concluding message

Poisson regression and its close relatives should be used more often in research into incontinence.

FUNDING: NONE
DISCLOSURES: NONE
HUMAN SUBJECTS: This study was approved by the Otago Ethics Committee and followed the Declaration of Helsinki Informed consent was obtained from the patients.