INFORMATION POINT:

Mann–Whitney Test

The Mann–Whitney test, also known as the Wilcoxon rank sum test, is a non-parametric test used to test for difference between the medians of two **independent groups**. This test is the non-parametric equivalent of the two-sample *t*-test. There are two derivations of the same test, one devised by Wilcoxon and the other by Mann and Whitney. To avoid confusion with the paired test also devised by Wilcoxon (see p. 584), this test is most frequently called the Mann–Whitney *U*-test, but alternative names are sometimes used.

The test requires two **independent groups of** observations. For example, in the paper above one group is the withdrawers and the other group is the completers. To carry out the test we put both groups together and rank the observations, giving 1 for the smallest, 2 for the next smallest and so on. The test statistic is based on summing the ranks for each group. Altman (1991) and Conover (1980) provide more detail about the calculation of the test statistic.

The test is based on putting the observations in order, so the data must be at least ordinal. However, to carry out the equivalent parametric test, a two-sample *t*-test, we require that the actual data values must be measured on an interval scale. In addition, for a *t*-test to be valid we also need to make the distributional assumption that values for each group follow a normal distribution. This distributional assumption is not required for the Mann–Whitney test.

Since the Mann–Whitney test makes fewer assumptions, why is it not used all the time in preference to the *t*-test? Whilst it is considered more robust, in the sense that it does not make many assumptions about the data, the cost is that the Mann–Whitney test is less powerful than the *t*-test, particularly in situations where the distributional assumptions of the *t*-test would be fulfilled. The other disadvantage, as discussed by Crichton (1998), is that estimation from non-parametric techniques, for example of the size of a treatment effect, can be difficult or impossible.

NICOLA CRICHTON

Further reading

Altman D.G. (1991) *Practical statistics for medical research*. Chapman & Hall, London. pp. 194–197.

Conover W.J. (1980) *Practical nonparametric statistics*. 2nd edn. John Wiley and Sons, New York. pp. 216–223.

Crichton N.J. (1998) Statistical considerations in design and analysis. In: *Research and development in clinical nursing practice* (Roe B. & Webb C., eds). Whurr, London. p. 209.