Using the R statistical software to improve the reproducibility, transparency and translation of research

***Barnett, AG*[1]**

[1] *Queensland University of Technology, Brisbane, Australia*

Background
Data management and statistical analyses are critical tasks for most health and medical studies, but are often done without making a clear reproducible record. Mistakes in data management and analyses are common and can invalidate the results, creating research waste.

Objectives
To show how the statistical software *R* can increase good research practice by creating more reliable, transparent and interactive analyses.

Method
R is free and has thousands of additional packages covering a huge variety of statistical methods. *Rmarkdown* is a user-friendly interface for *R* that combines code and text to create reports in Word or PDF that can be quickly updated as new data arrive. Clinical trial reports can be created using: i) scrambled treatment groups to help find programming errors, and ii) blinded groups to facilitate unbiased interpretations of the results. The *shiny* application in *R* can be used to create interactive web pages where results can be examined in greater detail and alternative assumptions tested.

Results
Using scrambled treatment groups allowed programming errors to be fixed without the embarrassment of recalling published results. One regression model did not converge due to the redundancy of clustering on ward within hospital. This second level of clustering was removed before examining the real treatment group, which removes any suspicion of altering the model to get a more desirable result. An interactive web page for a cost-effectiveness analysis allows users to try their own assumptions and provides results that are far more dynamic than a standard paper.