**BJS stats filler version 1 23-07-15**

**Spurious findings**

It is well recognised that data will eventually "confess" if it is interrogated enough. Provided enough variables and tests are deployed, a "statistically significant" result can usually be obtained. Consequentially, many of the "discoveries" in clinical research later transpire to be spurious findings reflecting only chance occurrences and the idiosyncrasies of the dataset and analysis strategies used, and do not reflect a real, let alone clinical useful, relationship.(ref Ionnadis) Three main strategies can be used to limit the scope for misleading findings. First, only relationships which have plausibility should be assessed with existing knowledge applied to weed out inappropriate analyses. Just because a statistical analysis can be done doesn't mean it should be. He/she who seeks may finds but they do not necessarily have anything useful to show. Second, statistical analyses should be pre-specified as far as possible and those which are exploratory or stimulated by incidental findings should be clearly labelled as such. Third, the simplest statistical analysis which is reliable and sufficiently informative should be adopted. Greater complexity opens up the potential for a "precisely wrong" result due to assumptions that do not hold. More complex analyses may well be appropriate as a secondary analysis approach but the preference should be for simplicity as far as possible to address the research question.

Ref

Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2(8): e124.