Improving Medical Statistics www.improvingmedicalstatistics.com
and Interpretation of Clinical Trials home author contact

Subgroup Analysis - Potential Limitations

Dr. Peter Sleight has written insightful comments in regards to the limitations of subgroup analysis.[1,2] Dr. Sleight, among numerous other accomplishments, was one of the investigators involved in the ISIS-2 trial.

The ISIS-2 trial[2] (Second International Study of Infarct Survival) was a very large randomized clinical trial which showed that both aspirin and a clot dissolving medication had important and statistically significant benefits for patients having a heart attack. Aspirin reduced the death rate by 23% in patients having a heart attack.

Inappropriate subgroup analysis can lead to ludicrous results

Dr. Sleight and the ISIS-2 trial investigators did a subgroup analysis of patient outcome by astrological sign to show the potential limitations in reliability of subgroup analysis.

The ISIS-2 investigators[2] reported that “subdivision of the patients in ISIS-2 with respect to their astrological birth sign appears to indicate that for persons born under Gemini or Libra, there was a slightly adverse effect of aspirin on mortality (9% increase, SD 13; NS), while for patients born under all other astrological signs, there was a striking beneficial effect (28% reduction, SD 5; 2p <0.00001.)”

The subgroup of analysis suggesting that Gemini and Libra had an adverse effect, rather than a beneficial effect with aspirin was not a true relation. These patients would benefit from aspirin to an equal degree as the rest of the group. Aspirin is a life saving medication in the setting of an acute myocardial infarction (heart attack).

Subgroup analysis, at times, can lead to findings that are incorrect. If a subgroup analysis results in an unexpected finding in outcome that is different from a highly significant and beneficial effect for the group as a whole, the subgroup analysis is often incorrect. *In fact, it is more likely that the unexpected subgroup finding that runs counter to the group finding is simply not valid.*

This subgroup analysis erroneously suggested that treatment with aspirin was quite effective and statistically significant for all patients except those born under the sign of Gemini or Libra. The difference in outcome with respect to astrologic sign was naturally an artifact and would not be reproducible in subsequent studies which was the point of their analysis.

Validity of subgroup analysis:

The view of this website is that subgroup analysis can be quite useful, but the validity tends to be inversely proportional to the number of subgroups which are analyzed.

A study is not immune to an incorrect subgroup analysis outcome simply because the subgroup was prespecified, particularly if there were a large number of prespecified subgroup analyses.

(If 20 subgroup analyses are prespecified, then it is expected that one of these subgroup analyses may show a false result for a P=.05 probability relationship.)

Part of the benefit of having prespecified subgroup analysis is that there are necessarily fewer such analyses than the almost unlimited number of ways to subdivide the data in a post hoc analysis after the trial results have already been obtained.

What is the reliability of a finding of a small subgroup in a trial who unexpectedly have a different outcome from the rest of the group?

As pointed out by Dr Sleight, if one of a multitude of subgroup analyses has a different outcome (in an unexpected fashion) from the outcome of the overall group in a very large trial, it is more reliable to assume that the subgroup actually had the same outcome as the overall group.

A subgroup analysis which results in variance from the overall group outcome is more likely to be true if it involves a large subgroup and there are a very limited number of prespecified analyses. Even then, the subgroup analysis findings often tend to be most valid as a starting place for subsequent clinical trials to confirm or refute the finding, rather than viewed as a definitive result.

Particularly vulnerable to error, is the post hoc analysis of trial data when a number is derived retrospectively from trial data and is then said to separate the responders from the nonresponders.

Years ago, this type of analysis of the CARE study data by prominent investigators was said to indicate that treatment of initial LDL cholesterol levels below 124 mg/dL with medication was not helpful in patients with coronary disease (blocked heart arteries). Clearly, this was later shown to be erroneous by multiple studies, including the Heart Protection Study[3].

More recently, a different post hoc subgroup data analysis concerning which patients with a cardiomyopathy (weak heart muscle) benefit from an implantable defibrillator led to erroneous conclusions by the Medicare administration.

For an enlightening look at potential limitations of subgroup analysis, the following two articles are recommended:

1. Debate: Subgroup analyses in clinical trials: fun to look at- but don't believe them! Peter Sleight. Current Control Trial Cardiovasc Med. 2000 1(1): 25-27.

2. ISIS-2 (Second International Study of Infarct Survival). Lancet 1988: ii: 349-360 (pages of interest 356-357)

---

3. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20536 high-risk individuals: a randomized placebo-controlled trial. Lancet 2002; 360: 7-22.