

EDITOR'S PAGE

Reporting and interpreting results of systematic reviews



L Harvey and JJ Wyndaele
Antwerp University Hospital, Antwerp, Belgium
E-mail: spinalcord@uza.be

Dear *Spinal Cord* reader,

In recent years *Spinal Cord* has published increasingly more systematic reviews that evaluate the effectiveness of interventions for people with spinal cord injuries. Many methodological issues need to be considered when conducting systematic reviews: to identify all relevant trials, not just those with favourable outcomes; to carefully define the scope; and to appropriately rate trials for bias. The most important findings of systematic reviews are the estimates of treatment effectiveness. For continuous data, this is reflected in between-group differences and corresponding 95% confidence intervals (CI).¹ If the 95% CI spans the minimally worthwhile treatment effect, the results are inconclusive. If the entire 95% CI is larger than the minimally worthwhile treatment effect, the treatment is clearly effective (again on average when applied to the target population); if the entire 95% CI is smaller than the minimally worthwhile treatment, the treatment is clearly ineffective. This all assumes that a favourable outcome is reflected by a positive number and a detrimental outcome is reflected by a negative number. This is not always the case, so data need to be carefully examined.

Statistically combining individual trials with a meta-analysis provides a pooled summary of treatment effectiveness, taking into account the relative contribution of each trial. It gives a more precise estimate of treatment effectiveness. Sometimes systematic reviews just report *P* values for outcomes of individual trials. The Cochrane Society strongly discourages this approach because authors (and hence readers) erroneously interpret findings with $P > 0.05$ as evidence of treatment ineffectiveness. As example,² we reported a non-significant between-group difference on bone density after three months of regular standing. An erroneous interpretation of these results is that passive standing is ineffective and does not prevent bone loss. This may or may not be the case; the results of our study do not permit a conclusion either way, due to an insufficient sample size.³ That results can be statistically significant but might not be clinically important is also problematic. In our study above, a statistically significant effect of regular standing on ankle joint mobility was found, but the size of the treatment effect was too small to be intrinsically worthwhile. Contrary, the data provided quite conclusive evidence that regular standing for ankle joint mobility was not worthwhile, despite a statistically significant between-group difference. Of course, readers may dispute authors' definitions of minimally worthwhile treatment effects in which case they are free to interpret results with respect to what they believe to be clinically worthwhile. However, readers of systematic reviews can do neither, if only *P* values are provided, and it is not a satisfactory solution to report pre- to post-change data for experimental and control groups without also articulating the critical between-group differences.

The growth of systematic reviews and clinical practice guidelines must be encouraged, but at the same time these publications should provide sufficient data to support informed and sensible clinical decisions. Most important are the inclusion of between-group differences and corresponding 95% CIs. If not, there is a real risk of misinterpreting the evidence and advocating for interventions that may have trivial effects or worse still, abandoning interventions that may be worthwhile.

If you want to learn more about how to interpret data on effects of interventions in systematic reviews, freely available on-line resources are provided by The Cochrane Collaboration.

References

- 1 Matthews JN, Altman DG. Statistics notes. Interaction 2: Compare effect sizes not *P* values. *Br Med J* 1996; 313: 808.
- 2 Ben M, Harvey L, Denis S, Glinsky J, Goehl G, Chee S *et al*. Does 12 weeks of regular standing prevent loss of ankle mobility and bone mineral density in people with recent spinal cord injuries? *Aust J Physiother* 2005; 51: 251–256.
- 3 Altman DG, Bland JM. Absence of evidence is not evidence of absence. *Br Med J* 1995; 311: 485.