Interpreting Effect Size to Estimate Responsiveness of Outcome Measures

To the Editor:

I congratulate Hsieh et al on their well-conducted study comparing the responsiveness of some common outcomes used in stroke rehabilitation research. They report large responsiveness indices for Fugl Meyer (FM) and Action Research Arm test with FM being a more responsive measure than Action Research Arm test.

Hsieh et al have used the standardized response mean (SRM), which is one of the best valid measures to estimate responsiveness. The authors have, however, used Cohen’s thresholds (≥0.8 large; 0.5 to 0.8 moderate, and <0.5 small) for grading the SRM values, which is debatable. Cohen’s thresholds are described for effect size (ESp) calculated by dividing change in scores by pooled SD (population standard deviation). SRM effect size is slightly different from ESp and is calculated by dividing the change in scores by SD of change in scores. Consequently, applying Cohen’s thresholds to SRM values rather than ESp could lead to over- or underestimation of effects.

Middel et al in their 2 studies have shown up to half of the estimates based on Cohen’s threshold applied to SRM values being either over- or underestimation of an intervention-related effect. A simple method to avoid such estimation errors has been suggested whereby the correlation between repeated measurements (baseline and follow-up) is used to determine the effect size and then apply the Cohen’s thresholds. For example, an SRM of 0.85, which could be interpreted as a large effect, changes to a moderate effect according to Cohen’s thresholds due to a correlation of 0.82 between repeated measurements.

It is also to be noted that the results of large responsiveness observed in this study by Hsieh et al are not comparable to previous FM and Action Research Arm test responsiveness studies (in stroke subjects), which have showed only a moderate effect size. The moderate responsiveness is in fact believed to be one of the limitations of FM scale.

It is quite possible Hsieh et al may find that even after calculation of ESp values, the category of effect sizes as per Cohen’s thresholds remains the same, that is, large responsiveness indices for FM and Action Research Arm test. If this is the case, the responsiveness of these scales will have to be further researched, particularly the FM scale because it is extensively used in stroke rehabilitation studies.

Disclosures

None.

Manoj Sivan, MRCS
Leeds Teaching Hospitals NHS Trust
Leeds, UK

3. Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integr Care. 2002;2:e15.
Interpreting Effect Size to Estimate Responsiveness of Outcome Measures
Manoj Sivan

Stroke. 2009;40:e709; originally published online October 29, 2009;
doi: 10.1161/STROKEAHA.109.566836

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://stroke.ahajournals.org/content/40/12/e709

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/