Response to Letter by Sivan
Yu-wei Hsieh, Ching-yi Wu and Keh-chung Lin
Stroke 2009;40:e710-e711; originally published online Oct 29, 2009;
DOI: 10.1161/STROKEAHA.109.567115
Stroke is published by the American Heart Association. 7272 Greenville Avenue, Dallas, TX 72514
Copyright © 2009 American Heart Association. All rights reserved. Print ISSN: 0039-2499. Online
ISSN: 1524-4628

The online version of this article, along with updated information and services, is
located on the World Wide Web at:
http://stroke.ahajournals.org/cgi/content/full/40/12/e710
Response to Letter by Sivan

Response:

We thank Dr Sivan for his comments on different viewpoints of calculation and interpretation of effect size. Because there is a lack of consensus regarding which responsiveness index is most appropriate for paired observations (within-group change),1,2 we chose the Liang’s standardized response mean3 to evaluate the responsiveness of the 3 outcome measures in our study.

We read with interest the articles of Middle et al cited by Dr Sivan. Middel et al4–6 pointed out that using the Cohen criteria7 for the standardized response mean may over- or underestimate the effects, because Cohen’s thumb of rule was originally proposed for comparing independent samples, such as between treatment and control groups. The effect size used for paired observations (eg, repeated measures within 1 group) has to be adjusted to more accurately apply the Cohen criteria or estimate the power.7 The adjusted effect size, $d$, is obtained by dividing the $d’$ by $\sqrt{1-r}$, where the $d’$ is obtained by dividing the mean change by the pooled standard deviation ($SD_{pooled}$), and the $r$ indicates the correlation between the values for paired observations.

We considered the correlation between the scores of paired observations in our study and calculated the adjusted effect size $d$. The Table shows the results of the adjusted effect size $d$ and standardized response mean of the 3 motor outcome measures. We found that estimates of the adjusted effect size $d$ for the Fugl-Meyer Assessment (2.11), the Action Research Arm Test (1.43), and the functional ability scale of the Wolf Motor Function Test (1.86) remained in the “large” category based on the Cohen criteria: small (0.2 to 0.5), moderate (0.5 to 0.8), and large (>0.8). The results support our prior results and interpretation of the responsiveness of the 3 outcome measures.

Another concern raised by Dr Sivan is that 2 previous studies (Lin and Hsieh)8,9 demonstrated the moderate responsiveness of the Fugl-Meyer Assessment and the Action Research Arm Test. Those studies used another effect size index calculated by dividing change in scores by the baseline standard deviation for paired observations. In addition, the differences among those 2 studies and ours may be partly attributable to differences in intervention protocols and time frame of assessments. Responsiveness of one outcome measure may be affected by the subject population, treatment interventions, time frame of assessments, and the indices used. Clinical trials usually perform repeated assessments over time, and the use of valid and responsive outcome measures are important for detecting clinically important changes after treatment. Further studies should seek better interpretation of the responsiveness of outcome measures in clinical studies by considering relevant factors, including treatment intervention and the correlation between scores of the repeated measures. These efforts may also improve our knowledge about the dose-response relationship for rehabilitation regimens.

Table. The Adjusted Effect Size $d$ and the Standardized Response Mean of the 3 Outcome Measures

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean Change</th>
<th>SD change</th>
<th>SDpooled</th>
<th>SRM</th>
<th>Adjusted Effect Size $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMA</td>
<td>0.901</td>
<td>5.75</td>
<td>4.06</td>
<td>8.67</td>
<td>1.42</td>
</tr>
<tr>
<td>ARAT</td>
<td>0.915</td>
<td>4.68</td>
<td>4.95</td>
<td>11.24</td>
<td>0.95</td>
</tr>
<tr>
<td>WMFT-TIME</td>
<td>0.594</td>
<td>2.13</td>
<td>5.56</td>
<td>5.91</td>
<td>0.38</td>
</tr>
<tr>
<td>WMFT-FAS</td>
<td>0.951</td>
<td>0.30</td>
<td>0.23</td>
<td>0.73</td>
<td>1.30</td>
</tr>
</tbody>
</table>

Disclosures

None.

Yu-wei Hsieh, MS
School of Occupational Therapy
College of Medicine
National Taiwan University
Taipei, Taiwan

Ching-yi Wu, ScD, OTR
Department of Occupational Therapy
Chang Gung University
Taoyuan, Taiwan

Keh-chung Lin, ScD, OTR
School of Occupational Therapy
College of Medicine
National Taiwan University
Taipei, Taiwan


6. 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.

