THE RELATION BETWEEN QUALITY OF CLINICAL TRIALS AND ACUPUNCTURE EFFICACY

ARTIGO ORIGINAL / ORIGINAL ARTICLE

INTRODUCTION
Clinical trials of acupuncture not always have concordant results, mostly due to their great heterogeneity. Two indexes have been developed to analyze the quality of acupuncture trials. This study hypothesizes that, the more adequate the intervention and the control techniques, the more efficacious the acupuncture. Methods: both indexes were applied to 27 randomized clinical trials comparing acupuncture to placebo. Results were compared by using the Mann-Whitney test. Results: studies favorable to acupuncture had an intervention score’s median of 11.5; for the unfavorable ones, it was 7, p: 0.0017. Articles with and without statistically significant differences, though, had the same median for their scores in the control index: 6. Discussion: there is a positive relation between a better score for acupuncture technique and a statistically significant difference between acupuncture and interventional control. However, due to the little heterogeneity in the degree of physiological effect from each article, the control index had no statistical significance. Conclusion: this study established that, among acupuncture RCT controlled by placebo or sham of moderate physiological effect, the adequacy of the technique is more important than the adequacy of control in establishing a statistically significant difference between acupuncture and interventional control.

METHODS
Both indexes were applied to 27 randomized clinical trials comparing acupuncture to placebo. Results were compared by using the Mann-Whitney test. Results: studies favorable to acupuncture had an intervention score’s median of 11.5; for the unfavorable ones, it was 7, p: 0.0017. Articles with and without statistically significant differences, though, had the same median for their scores in the control index: 6. Discussion: there is a positive relation between a better score for acupuncture technique and a statistically significant difference between acupuncture and interventional control. However, due to the little heterogeneity in the degree of physiological effect from each article, the control index had no statistical significance. Conclusion: this study established that, among acupuncture RCT controlled by placebo or sham of moderate physiological effect, the adequacy of the technique is more important than the adequacy of control in establishing a statistically significant difference between acupuncture and interventional control.

RESULTS
Studies favorable to acupuncture had an intervention score’s median of 11.5; for the unfavorable ones, it was 7, p: 0.0017. Articles with and without statistically significant differences, though, had the same median for their scores in the control index: 6. Discussion: there is a positive relation between a better score for acupuncture technique and a statistically significant difference between acupuncture and interventional control. However, due to the little heterogeneity in the degree of physiological effect from each article, the control index had no statistical significance. Conclusion: this study established that, among acupuncture RCT controlled by placebo or sham of moderate physiological effect, the adequacy of the technique is more important than the adequacy of control in establishing a statistically significant difference between acupuncture and interventional control.

CONCLUSION
This study hypothesizes that, the most adequate the acupuncture technique is and the less physiological effect the control exerts, the more effective the acupuncture intervention will be, when compared to control. This study intends to test this hypothesis and evaluate if the study quality is reflected in the efficacy of treatment by acupuncture.

ABSTRACT

INTRODUCTION: clinical trials of acupuncture not always have concordant results, mostly due to their great heterogeneity. Two indexes have been developed to analyze the quality of acupuncture trials. This study hypothesizes that, the more adequate the intervention and the control techniques, the more efficacious the acupuncture. Methods: both indexes were applied to 27 randomized clinical trials comparing acupuncture to placebo. Results were compared by using the Mann-Whitney test. Results: studies favorable to acupuncture had an intervention score’s median of 11.5; for the unfavorable ones, it was 7, p: 0.0017. Articles with and without statistically significant differences, though, had the same median for their scores in the control index: 6. Discussion: there is a positive relation between a better score for acupuncture technique and a statistically significant difference between acupuncture and interventional control. However, due to the little heterogeneity in the degree of physiological effect from each article, the control index had no statistical significance. Conclusion: this study established that, among acupuncture RCT controlled by placebo or sham of moderate physiological effect, the adequacy of the technique is more important than the adequacy of control in establishing a statistically significant difference between acupuncture and interventional control.

METHODS: both indexes were applied to 27 randomized clinical trials comparing acupuncture to placebo. Results were compared by using the Mann-Whitney test. Results: studies favorable to acupuncture had an intervention score’s median of 11.5; for the unfavorable ones, it was 7, p: 0.0017. Articles with and without statistically significant differences, though, had the same median for their scores in the control index: 6. Discussion: there is a positive relation between a better score for acupuncture technique and a statistically significant difference between acupuncture and interventional control. However, due to the little heterogeneity in the degree of physiological effect from each article, the control index had no statistical significance. Conclusion: this study established that, among acupuncture RCT controlled by placebo or sham of moderate physiological effect, the adequacy of the technique is more important than the adequacy of control in establishing a statistically significant difference between acupuncture and interventional control.

RESULTS: studies favorable to acupuncture had an intervention score’s median of 11.5; for the unfavorable ones, it was 7, p: 0.0017. Articles with and without statistically significant differences, though, had the same median for their scores in the control index: 6. Discussion: there is a positive relation between a better score for acupuncture technique and a statistically significant difference between acupuncture and interventional control. However, due to the little heterogeneity in the degree of physiological effect from each article, the control index had no statistical significance. Conclusion: this study established that, among acupuncture RCT controlled by placebo or sham of moderate physiological effect, the adequacy of the technique is more important than the adequacy of control in establishing a statistically significant difference between acupuncture and interventional control.

CONCLUSION: this study hypothesizes that, the most adequate the acupuncture technique is and the less physiological effect the control exerts, the more effective the acupuncture intervention will be, when compared to control. This study intends to test this hypothesis and evaluate if the study quality is reflected in the efficacy of treatment by acupuncture.
METHODOLOGY

All the RCT used for the other articles from this study were selected for analysis, adding up to 34 articles. From these, those that compared acupuncture to a control by placebo or sham (including here only control techniques that involved using needles, such as retractable needle or sham acupuncture) were included, resulting in 21 articles. As this was too small a number for the statistical analysis, a new search in the MEDLINE database was performed during the month of December, 2010, using the keywords “Acupuncture” and “Placebo” as words in the title, with restrictions for the years of 2005 to 2011, without any restriction of language of publication or theme. The articles were once more selected by title and abstract and those that sufficed the inclusion criteria were obtained for analysis, what resulted in another six articles.

For each article only the part concerning the study methodology was extracted. All RCT were scored with both indexes on a chart, being identified by number, without any author identification, following the instructions of application presented in the second and third parts of this study. In case there was no information concerning a certain item, it was scored the worst, as recommended. Next, once all articles were scored, the Results session was analyzed in order to register whether: a) there was any statistical difference between interventional control and acupuncture; and b) there was any difference between interventional control and other control group, when available.

Last, the whole article was analyzed in order to add important data on limitations, such as authors’ commentaries concerning the sample or support of health insurance companies.

Results were statistically analyzed with the Mann-Whitney test; results were considered significant when \( p < 0.05 \).
RESULTS

The scores for each article are found on table 1. It is shown the name of the main author, year of publication, number of sample statistically analyzed, the score for the acupuncture technique, the score for control and finally the results, first the comparison between acupuncture and interventional control, and then interventional control and other control group. When there is another number between parentheses beside the score value, it is the corrected score, for information on De Qi achievement.

<table>
<thead>
<tr>
<th>AUTHOR AND YEAR</th>
<th>N</th>
<th>ACUPUNCTURE INDEX</th>
<th>CONTROL INDEX</th>
<th>RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hedstrom 1998</td>
<td>104</td>
<td>7</td>
<td>7.5</td>
<td>No statistical difference.</td>
</tr>
<tr>
<td>Diner 2006</td>
<td>794</td>
<td>12</td>
<td>6(1)</td>
<td>No statistical difference.</td>
</tr>
<tr>
<td>Entre 2007</td>
<td>409</td>
<td>12</td>
<td>1</td>
<td>Acupuncture is superior to control in secondary outcome. It is not possible to evaluate psychological effects.</td>
</tr>
<tr>
<td>Medhurst 2007</td>
<td>270</td>
<td>6</td>
<td>1</td>
<td>No statistical difference.</td>
</tr>
<tr>
<td>Pacco 2008</td>
<td>127</td>
<td>13</td>
<td>7.5</td>
<td>Acupuncture is superior to control.</td>
</tr>
<tr>
<td>Kaur 2006</td>
<td>69</td>
<td>7</td>
<td>6</td>
<td>No statistical difference.</td>
</tr>
<tr>
<td>Aerteme-Andrade 2008</td>
<td>3.7</td>
<td>13</td>
<td>6</td>
<td>Acupuncture is superior to control.</td>
</tr>
<tr>
<td>Aerteme-Andrade 2006</td>
<td>25</td>
<td>11</td>
<td>6</td>
<td>No statistical difference.</td>
</tr>
<tr>
<td>White 2008</td>
<td>50</td>
<td>4</td>
<td>6</td>
<td>No statistical difference. The study originally intended to use a sample of 80 to present statistical difference.</td>
</tr>
<tr>
<td>Xue 2004</td>
<td>37</td>
<td>13</td>
<td>9 (4.5)</td>
<td>Acupuncture is superior to control at short term.</td>
</tr>
<tr>
<td>Vos 2004</td>
<td>97</td>
<td>11</td>
<td>6</td>
<td>Acupuncture + de qi is superior to control - de qi alone.</td>
</tr>
<tr>
<td>Scherf 2006</td>
<td>1007</td>
<td>9</td>
<td>0.5</td>
<td>No statistical difference. Health insurance companies support.</td>
</tr>
<tr>
<td>Witt 2005</td>
<td>286</td>
<td>12</td>
<td>0</td>
<td>Acupuncture is superior to control for as long as 1 year.</td>
</tr>
<tr>
<td>Link 2001</td>
<td>62</td>
<td>7</td>
<td>5.1(2)</td>
<td>No statistical difference.</td>
</tr>
<tr>
<td>Leitang 2002</td>
<td>131</td>
<td>9</td>
<td>6</td>
<td>No statistical difference.</td>
</tr>
<tr>
<td>Midberger 2002</td>
<td>134</td>
<td>11</td>
<td>6(1)</td>
<td>Acupuncture is superior to control. There was a tendency to lower pain in acupuncture group when compared to control, but the use of analgesia was significantly smaller in the first group. The author consider it may be due to inadequate frequency of electroacupuncture.</td>
</tr>
<tr>
<td>Wang 2006</td>
<td>27</td>
<td>11</td>
<td>6</td>
<td>No statistical difference.</td>
</tr>
</tbody>
</table>

Table 1. Articles analyzed, statistically analyzed sample number, scores according to the indexes, results according to comparison between acupuncture and interventional control and interventional control and other control group.
**Characteristics of the analyzed RCT**

Sixteen among the 27 RCT were favorable to acupuncture. However, 15 among all presented a sample smaller than 100; in two articles, the authors commented that the sample was below the needed for a statistical significance of the primary outcome. Most of them were two-armed trials. Only ten were three-armed (one of them started with an anti-depressant controlled group which was excluded since the beginning due to little adhesion) and compared the interventions to another control group (no treatment, conventional treatment or waiting list). Four studies had a support from health insurance companies, an important confounder for it is believed that the participants became aware of it, what contaminated the evolution of intervention groups. Therefore, the variables were analyzed with and without these confounders.

**Acupuncture index scores**

No article had any score in the second part of this index. Studies favorable to acupuncture had a score median of 11.5; for the unfavorable ones, it was 7. When compared, the calculated Z was 3.13 and p was 0.0017, which showed a positive relation between the highest scores and a statistically significant difference between acupuncture and interventional control. When adjusting the scores according to De Qi achievement (considering that there was no De Qi achievement for the control and that there was for the acupuncture group) in studies that lacked such information, Z calculated rose to 3.28 and p lowered to 0.001. Excluding the articles with inadequate sample number (according to the authors) and with health insurance companies support, the analysis showed the lowest, though still statistically significant difference: calculated Z was 2.90, and p was 0.0037. All these results show a statistically significant positive relation.

**Control technique index scores**

1. Comparison between the index scores and the difference between acupuncture and interventional control: none of these analyses reached statistical significance. Articles with and without statistically significant differences had the same median for their scores: 6. Calculated Z was of 0.03, and p was of 0.98. The correction that showed the highest significance was excluding those supported by health insurance companies: Z: 0.70 and p: 0.48. Thus, a worse (i.e., higher) score in the study was not related to an absence of difference between the acupuncture and the interventional control groups.

2. Comparison between index scores and the difference between interventional control and other control groups: in this case, there was no statistical difference either. Only ten RCT were included in this analysis and calculated Z was of 1.92, and p was 0.055, which showed only a tendency of RCT with differences between both groups having a better (i.e., lower) score in the control technique index.

**DISCUSSION**

This study shows that both indexes may be used for the analysis of acupuncture RCT controlled by placebo or sham of moderate physiological effect, the adequacy of the technique is more important than the adequacy of control in establishing a statistically significant difference between acupuncture and interventional control. An adequate technique reflects a positive result for the intervention group. More studies are needed in order to correlate the adequacy of control and the outcome of trials.

**Acknowledgements**

We are thankful for the acupuncturists who contributed to this study: Drs. Carla A. D. Antonia, Sohaku R. Bastos and Viviane S. P. Giovaninni.

**REFERÊNCIAS**


