THE RELATION BETWEEN QUALITY OF CLINICAL TRIALS AND ACUPUNCTURE EFFICACY

RELAÇÃO ENTRE A QUALIDADE DOS ENSAIOS CLÍNICOS E A EFICÁCIA DA ACUPUNTURA

David Gonçalves Nordon*

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 a 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.

Key-words: acupuncture, placebo, efficacy.

RESUMO

Introdução: ensaios clínicos de acupuntura nem sempre têm resultados concordantes, principalmente devido à sua heterogeneidade. Dois índices foram desenvolvidos para analisar a qualidade de ensaios de acupuntura. Este estudo tem a hipótese de que, quanto mais adequadas as técnicas de acupuntura e controle, mais eficaz a acupuntura. Métodos: ambos os índices foram aplicados em 27 ensaios clínicos randomizados comparando acupuntura a placebo. Os resultados foram comparados com o uso do teste de Mann-Whitney. Resultados: estudos favoráveis à acupuntura tinham um escore de intervenção mediano de 11.5; os desfavoráveis, de 7, p: 0,0017. Artigos com e sem diferenças estatisticamente significantes, contudo, tinham a mesma mediana para os escores no índice de controle: 6. Discussão: há uma relação positiva entre um escore melhor para a técnica de acupuntura e uma diferença estatisticamente significante entre acupuntura e controle interventivo. Contudo, devido à pequena heterogeneidade no grau de efeito fisiológico de cada artigo, o índice de controle não teve diferença estatística. Conclusão: este estudo estabeleceu que, entre ECR controlados por placebo ou sham de efeito fisiológico moderado, a adequabilidade da técnica é mais importante que a adequabilidade do controle em estabelecer uma diferença estatisticamente significativa entre acupuntura e controle interventivo.

Descritores: acupuntura, placebo, eficácia.

BACKGROUND

Since it was introduced in the Occident in the middle of the 20th century, researchers all around the world have been trying to prove the efficacy of acupuncture in treating several diseases. Even though several studies have been already elaborated to prove how acupuncture works, there is still no consensus; only part of its surface is comprehended. And this is specially influenced by little concordant studies on its efficacy.

Meta-analyses (MA) and systematic reviews (SR) of acupuncture not always agree on a similar theme. However, when these studies are thoroughly evaluated, one may observe that the great heterogeneity of their randomized clinical trials (RCT) is responsible for the major part of such discordance.[1-3]

The quality of a MA is unavoidably attached to the quality of its included articles. Along this study, indexes were created in order to evaluate the quality of RCT. The first one is an index of acupuncture technique adequacy (Figure 1); its score ranges from 0 to 13 in the first part and 0 to 8 in the second. According to what is described in the text, the more the characteristics of the traditional acupuncture were followed, the higher it will be scored. The second one is an index of the adequacy of the control technique (placebo or sham) (Figure 2). Its score is also based on what is presented in the text and ranges from 0 to 20 and, the higher it is, the higher the possibility of the control exerting some physiological effect.

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.

Rev. Fac. Ciênc. Méd. Sorocaba, v. 15, n. 2, p. 6 - 10, 2013

* Acupunturista e shiatsuterapeuta pelo Colégio Brasileiro de Acupuntura; médico - FCMS/PUC-SP.

Recebido em 25/1/2012. Aceito para publicação em 13/9/2012.

Contato: d-nordon@uol.com.br

Figure 1. Index of acupuncture technique adequacy

	ITEM	VALUE	OBSERVATIONS
	1. Characteristics concerning the technique		
a.	**Diagnosis and point-selection according to patterns from TCM**	4	a) Consider acute pain as "Local blood and Qi stagnation", what may be treated either locally or systemically. b) Ifthe points selected may resolve only partof the pattern and its adequate treatment, score 2.
b.	***De Qi* achievement**	3	
c.	**Length/Frequency of sessions**	1	
d.	**Adequate manipulation of needles during session**	1	a) Consider as adequate frequency of manipulation of needles during the session. b) For electroacupuncture, evaluate if the wave type is adequate.
e.	**Adequate frequency of spins**	1	a) Whether reinforcing, reducing or balancing. b) For electroacupuncture, evaluate if the frequency is adequate to the stimulus.
f.	**Duration of treatment according to recommendations of TCM for the pattern**	3	
	Total:	**13**	
	2. Characteristics concerning the follow-up		
a.	**Evaluation of success through examination and confirmation of pattern resolution**	4	a) Confirmation through tongue and/or pulse diagnosis.
b.	**On the follow-up, re-evaluation was performed according to pattern diagnosis**	4	a) Re-evaluation through anamnesis, tongue and/or pulse diagnosis.
	Total:	**8**	

Figure 2.Index of control adequacy (placebo or sham)

Item	Value
1. **Points selected:**	
Are the same from intervention group	**4**
Are local points, near those from intervention, in case of local pain	**3**
Are different from intervention, however traditional points	**2**
Are different from intervention, non-traditional points, however near those from intervention	**1**
Are different from intervention, non-traditional and not near those from intervention	**0**
2. Needle insertion:	
Yes, deep in general and/or shallow in local pain	**2**
Yes, shallow	**1**
No	**0**
3. De Qi achievement:	
Yes	**3**
No	**0**
4. Needle manipulation:	
Yes	**1**
No	**0**
Total:	**0-10**
5. Length of control therapy was similar or the same as the intervention +	
Score ≥ 7	**X 2**
Score 3-6	**X 1,5**
Score 0-2	**X 1**
Corrected total:	**0-20**

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.[4-30]

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.[31,32] 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 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.

AUTHOR AND YEAR	N	ACUPUNCTURE INDEX	CONTROL INDEX	RESULTS	
				Acupuncture X interventional control	Interventional control X other control group
Hedström 1998[21]	104	7	7,5	No statistical difference.	No statistical difference (no treatment).
Diener 2006[10]	794	12	6 (1)	No statistical difference.	No statistical difference (conventional treatment).
Endres 2007[8]	409	12	1	Acupuncture is superior to control in secondary outcomes. It is not possible to exclude psychological effects. Health insurance companies support.	Does not apply.
Melchart 2005[17]	270	6	1	No statistical difference. Health insurance companies support.	Control is better than waiting list.
Facco 2008[11]	127	13	7,5	Acupuncture is superior to control.	No statistical difference (conventional treatment).
Karst 2001[15]	69	7	6	No statistical difference	Does not apply.
Alecrim-Andrade 2008[20]	37	13	6	Acupuncture is superior to control.	Does not apply.
Alecrim-Andrade 2006[6]	28	11	6	No statistical difference.	Does not apply.
White 2000[7]	50	4	6	No statistical difference. The study originally intended to use a sample of 80 to present statistical difference.	Does not apply.
Xue 2004[9]	37	13	9 (4,5)	Acupuncture is superior to control at short term.	Does not apply.
Vas 2004[18]	97	11	6	Acupuncture + diclofenac is superior to control – diclofenac.	Does not apply.
Scharf 2006[15]	1007	9	4,5	No statistical difference. Health insurance companies support.	Control is superior to conventional treatment.
Witt 2005[16]	286	12	0	Acupuncture is superior to control for as long as 1 year.	Interventional control is superior to other control group.
Fink 2001[22]	67	7	7,5 (2)	No statistical difference	Does not apply.
Leibing 2002[14]	131	9	6	No statistical difference	Interventional control is superior to other control group.
Molsberger 2002[13]	174	11	6 (1)	Acupuncture is superior to control.	No statistical difference.
Wong 2006[12]	27	11	6	There was a tendency to lower pain in acupuncture group when compared to control, but the use of analgesics was significantly smaller in the first group. The author consider it may be due to inadequate frequency of electroacupuncture.	Does not apply.

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's 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, with or without a third group of comparison. It was also observed that current studies do not perform follow-ups or evaluations of patterns' resolutions according to TCM philosophy. This was already expected, as rare articles actually diagnose patterns in the beginning of the study, even though a considerable number of them are based on the TCM philosophy when choosing the points.

Data show a positive relation between a better score for acupuncture technique and a statistically signifit difference between acupuncture and interventional control. One may suppose that an effective technique scores 9 or above in the acupuncture index.

However, the control index did not show a statistical significance. This is due to the similarity of scores among the studies (median and mode for both were 6), there being mainly studies of moderate physiological effect of the control (4.5 to 9 points), little of low effect (0 to 2) and none of high effect (14 to 20). Performing an analysis that purposely includes studies of higher and/or lower effect will make it possible to establish such a relation with statistical significance.

The comparison of interventional control groups and other control groups showed no statistical significance either; it would be expected that the higher the physiological effect from the interventional control, the bigger the difference between groups. The number of studies included, however, was too small and the scores remained too similar. Performing an analysis of three-armed studies whose interventional controls have both low and high physiological effects will establish such a relation.

Wayne's study,[30] the last included in this review, uses a Japanese acupuncture technique. Nevertheless, as it had a high score in the adequacy index, having achieved *De Qi* despite the shallow insertion of needles, we opted for including it in the analysis. This shows that, in spite of having been at first conceived for traditional Chinese acupuncture, it may be applied for other acupuncture techniques, once the differences are respected.

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. 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

1. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomized trials affect estimates of intervention efficacy reported in meta-analyses? Lancet. 1998;352:609-13.
2. Moher D, Fortin P, Jadad AR, Jünl P, Klassen T, Lorier JL, et al. Completeness of reporting of trials published in language other than English: implications for conduct and reporting of systematic reviews. Lancet. 1996;347:363-6.
3. Birch S. Systematic reviews of acupuncture – are there problems with these? Clin Acupunct Oriental Med. 2001;2:17-22.
4. Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AMK, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial. Ann Intern Med. 2004;141:901-10.
5. Haake M, Müller H-H, Schade-Brittinger C, Baster HD, Schäfer H, Maier C, et al. German acupuncture trials (GERAC) for chronic low back pain. Randomized, multicentre, blinded, parallel-group Trial with 3 groups. Arch Intern Med. 2007;167(17):1892-8.

6. Alecrim-Andrade J, Maciel-Júnior JA, Cladellas XC, Correa-Filho HR, Machado HC. Acupuncture in migraine prophylaxis: a randomized sham-controlled trial. Cephalalgia. 2006;26:520-9.
7. White AR, Resch K-L, Chan JCK, Norris CD, Modi SK, Pate JN, et al. Acupuncture for episodic tension-type headache: a multicentre randomized controlled trial. Cephalalgia. 2000;20:632-7.
8. Endres HG, Böwing G, Diener HC, Lange S, Maier C, Molsberger A, et al. Acupuncture for tension-type headache: a multicentre, sham-controlled, patient and observer-blinded, randomised Trial. J Headache Pain. 2007;8:306-14.
9. Xue CCL, Dong L, Polus B, English RA, Zheng Z, Da Costa C, et al. Electroacupuncture for tension-type headache on distal acupoints only: a randomized, controlled, crossover Trial. Headache. 2004;44:333-41.
10. Diener HC, Kronfeld K, Boewing G, Lungenhausen M, Maier C, Molserberg A, et al. Efficacy of acupuncture for the prophylaxis of migraine: a multicentre, randomized, controlled clinical trial. Lancet Neurol. 2006;5:310-6.
11. Facco E, Ligueri A, Petti F, Zanette G, Coluzzi F, Nardin M, et al. Traditional acupuncture in migraine: a controlled, randomised study. Headache. 2008;48:398-407.
12. Wong RHL, Lee TW, Sihoe ADL, Wan IY, Ng CSH, Chan SKC, et al. Analgesic effect of electroacupuncture in postthoracotomy pain: a prospective randomized Trial. Ann Thorac Surg. 2006;81:2031-6.
13. Molsberger AF, Mow J, Pawelec DB, Winkler J. Does acupuncture improve the orthopedic management of chronic low back pain – a randomized, blinded, controlled trial with 3 months fllow-up. Pain. 2002;99:579-87.
14. Leibing E, Leonhardt U, Köster G, Goerlitz A, Rosenfeldt JA, Hilgers R, et al. Acupuncture treatment of chronic low-back pain – a randomized, blinded, placebo-controlled Trial with 9-month follow-up. Pain. 2002;96:189-96.
15. Scharf HP, Monsmann U, Streitberger K, Wilte S, Krämer J, Maier C, et al. Acupuncture and knee osteoarthritis – a three-armed randomized Trial. Ann Intern Med. 2006;145:12-20.
16. Witt C, Brinkhaus B, Jena S, Linde K, Streng A, Wagenpfeil S, et al. Acupuncture in patients with osteoarthritis of the knee: a randomized trial. Lancet. 2005;366:136-43.
17. Melchart D, Streng A, Hoppe A, Brinkhaus B, Witt C, Wagenpfeil S, et al. Acupuncture in patients with tension-type headache: a randomised controlled Trial. BMJ. 2005;331:376-9.
18. Vas J, Méndez C, Perea ME, Vega E, Panadero MD, León JM et al. Acupuncture as a complementary therapy to the pharmacological treatment of osteoarthritis of the knee: a randomized controlled Trial. BMJ. 2004;329(7476):1216.
19. Karst M, Reinhard M, Thum P, Wiese B, Rolling J, Fink M. Needle acupuncture in tension-type headache: a randomized, placebo-controlled study. Cephalalgia. 2001;21:637-42.
20. Alecrim-Andrade J, Maciel Júnior J, Carnè X, Vasconcelos GMS, Correa Filho HP. Acupuncture in migraine prevention a randomized sham controlled study with 6-months posttreatment follow-up. Clin J Pain. 2008;24(2):98-105.
21. Gosman-Hedström G, Clarsson L, Lingenstierna U, Carlsson J, Olausson B, Fritzell M et al. Effects of acupuncture treatment on daily life activities and quality of life. A controlled, prospective and randomized study of acute stroke patients. Stroke. 1998;29:2100-8.
22. Fink MG, Kunbsebeck H, Wipperman B, Gehrke A. Non-specific effects of traditional Chinese acupuncture in osteoarthritis of the hip. Complement Ther Med. 2001;9:82-8.
23. Assefi NP, Sherman KJ, Jacobsen C, Goldberg J, Smith WR, Buchwald D. A randomized clinical Trial of acupuncture compared with sham acupuncture in fibromyalgia. Ann Intern Med. 2005;143:10-9.
24. Lao L, Bergman S, Langenberg P, Wong RH, Berman B. Efficacy of Chinese acupuncture on postoperative oral surgery pain. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 1995;79:423-8.
25. Jubb RW, Tukmachi ES, Jones PW, Dempsey E, Waterhouse L, Brailsford S. A blinded randomized trial of acupuncture (manual and electroacupuncture) compared with a non-penetrating sham for the symptoms of osteoarthritis of the knee. Acupunct Med. 2008;26(2):69-78.
26. Lee SW, Lion ML, Yuen KH, Leong WS, Chee C, Cheah PY, et al. Acupuncture versus sham acupuncture for chronic prostatitis/chronic pelvic pain. Am J Med. 2008;121:79e1-79e8.
27. Manheimer E, Lim B, Lao L, Berman B. Acupuncture for knee osteoarthritis – a randomised Trial using a novel sham. Acupunct Med. 2006;24(Suppl):S7-14.
28. Kennedy S, Baxter GD, Kerr DP, Bradbury I, Park J, McDonough SM. Acupuncture for acute non-specific low back pain: a pilot randomized non-penetrating sham controlled trial. Complement Ther Med. 2008;16:139-46.
29. Inoue M, Kitakoji H, Ishizali N, Tawa M, Yano T, Katsumi Y, et al. Relief of low back pain immediately after acupuncture treatment – a randomized, placebo controlled trial. Acupunct Med. 2006;24(3):103-8.
30. Wayne PM, Kerr CE, Schnyer RN, Legedza ATR, Savestky-German J, Shields MH, et al. Japanese-style acupuncture for endometriosis-related pelvic pain in adolescents and young women: results of a randomized sham-controlled trial. J Pediatr Adolesc Gynecol. 2008;21:247-57.
31. Nordon DG, Gianini RJ, Azevedo GR. Development and validation of an index for evaluating the quality of acupuncture articles. Rev Fac Ciênc Méd Sorocaba. 2012;14(2):59-63.
32. Nordon DG, Gianini RJ, Azevedo GR. Development and Validation of an Index of the Adequacy of the Control Technique for Acupuncture Studies. Rev Fac Ciênc Méd Sorocaba. 2013;15(1):186-91.

REVISTA DA FACULDADE DE CIÊNCIAS MÉDICAS DE SOROCABA

Agradecemos a colaboração da Associação dos Docentes da PUC-SP