Short Communication

Robustness of meta-analysis results in Cochrane systematic reviews: A case for acupuncture trials

Jiyoon Won

KM Science Research Division, Korea Institute of Oriental Medicine, Daejeon, 34054, Republic of Korea

ARTICLE INFO

Article history:
Received 5 July 2022
Revised 29 September 2022
Accepted 6 October 2022
Available online 7 October 2022

Keywords:
Fragility index
Cochrane systematic review
Meta-analysis
Complementary and integrative medicine
Acupuncture

ABSTRACT

Introduction: Fragility index (FI) refers to the smallest value that change statistical significance of study results. Meta-analyses of Cochrane systematic reviews are considered as the best evidence for stake holders because they enable effect size estimation that cannot be derived by individual studies, particularly in the field of complementary and integrative medicine (CIM). Thus, this study aimed to evaluate robustness of meta-analysis in Cochrane systematic reviews of acupuncture, one of the most used CIM treatment, using FI of meta-analysis.

Methods: Meta-analyses of acupuncture Cochrane systematic reviews with binary benefit outcome measures were searched in PubMed, Embase, and CENTRAL and subject to analysis. Randomized clinical trials (RCTs) or quasi RCTs adopted penetrating needles as treatment and compared with controls such as sham acupuncture, usual care, and active control were included. FI of meta-analyses was calculated in web (https://clinicalepidemiol.fr/fragility_ma/); and fragility quotient (FQ) was calculated by dividing FI by total sample size of meta-analysis.

Results: Of 248 retrieved studies, 12 Cochrane systematic reviews with 48 meta-analyses were analyzed. The median FI for statistically significant and non-significant meta-analyses was 7 [interquartile range: 3 - 12] and 6 [interquartile range: 3 - 10]. FQ was similar irrespective of statistical significance. Among controls, meta-analyses adopting sham acupuncture displayed the widest range of FI and FQ.

Conclusion: Robustness of meta-analyses in Cochrane systematic reviews for acupuncture was similar irrespective of statistical significance. Impact of control intervention cannot be determined in this study. As FI enables intuitive interpretation, applications for CIM field can be useful.

© 2022 Korea Institute of Oriental Medicine. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

Trial results are mostly presented with a p value in publications. While a threshold of 0.05 has been accepted as an empirical standard for statistical significance, there have been critics on usage of the p value alone in presentation and interpretation of results. Since Walsh et al suggested a concept of fragility index, robustness of randomized clinical trial results with binary outcome measures can be assessed in a more straightforward way. 1 Fragility index refers to the smallest value that makes statistically significant results non-significant or makes non-significant results significant. Application of fragility index has expanded to interpretation of robustness of meta-analysis. 2 There should be a difference between interpretation of trial level fragility index and meta-analysis level fragility index considering that studies included in meta-analysis of systematic reviews are not only orthodox randomized clinical trials but also other study designs such as quasi-randomized clinical trials and observational studies.

Meta-analysis of Cochrane systematic reviews are considered as the best evidence for stake holders including researchers and practitioners in that they can shed light on estimating effect size of interventions that can not be derived by individual studies. 3 Most of complementary and integrative medicine intervention (CIM) trials are relatively small compared to conventional medicine intervention trials. Accordingly, meta-analysis of Cochrane systematic reviews for CIM can have a strong implication to stake holders because it demonstrates the effect of CIM. To the knowledge, evidence on measuring how much robustness the finding of meta-analysis achieved is lacking. By evaluating robustness of meta-analyses in Cochrane systematic reviews of acupuncture, one of the most popular CIM interventions, using fragility index, 2 this study can promote intuitive understanding of the meta-analyses.

E-mail address: erigh15@kiom.re.kr

https://doi.org/10.1016/j.imr.2022.100890
2213-4220 © 2022 Korea Institute of Oriental Medicine. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
2. Methods

2.1. Search strategy and inclusion criteria

A comprehensive literature search on Cochrane systematic reviews was conducted in PubMed from inception to June 2022 using terms of “Cochrane Database Syst Rev”[Journal], “acupuncture”, and “meta”. Embase and Cochrane Central Register of Controlled Trials (CENTRAL) was also searched using adjusted term for respective databases. Inclusion criteria were as follows: 1) Meta-analysis of binary benefit outcome measure including general assessment of therapy effectiveness was subject to analysis. Meta-analysis of other outcome measures such as harm or other measures was not considered in this study. 2) Randomized clinical trials or quasi randomized clinical trials adopted penetrating needles (manual acupuncture and electroacupuncture) as intervention and compared to control (sham acupuncture, usual care, no treatment, and active control) were included. If there were studies that added identical adjuvant treatment otherwise acupuncture to intervention group and control group, those studies were analyzed. Meta-analysis that compared different kind of acupuncture was not considered. 3) Any type of meta-analysis in Cochrane systematic review irrespective of outcome measure, i.e., primary outcomes and secondary outcomes was all included in this study. As fragility index for meta-analysis can be calculated meta-analysis containing more than two trials, meta-analysis of one trial was excluded.

2.2. Data extraction and statistical analysis

Extracted data were title of Cochrane systematic reviews, inclusion criteria of relevant Cochrane systematic reviews such as study design, participants and their conditions, intervention and control, details of meta-analysis including number of studies in each meta-analysis, type of analysis method including statistical method (Pero, Mantel-Haenszel, or inverse variance) effect measures (odds ratio, risk ratio, or risk difference), and analysis model (fixed effect or random effects), number of events and total number of samples. Control interventions were coded as sham (sham acupuncture), active control, and usual care. For convenience of analysis, usual care included usual care, no treatment, and waitlist. While add-on cases (e.g., acupuncture and other treatment was given to intervention group and only other treatment was given to control group) and multiple control group cases (e.g., sham acupuncture and no treatment control were simultaneously included in one control group) were included in calculating fragility index and fragility quotient of all studies, they were excluded in analyzing correlation between fragility index and fragility quotient and control groups. Sub-group analysis

Fragility index of meta-analysis was calculated in web (https://clinicalepidemio.fr/fragility_ma/) which was suggested by Atal et al. The minimum value that can change status of statistical significance of meta-analysis became the fragility index: a significant meta-analysis is neutralized by subtracting number or events of intervention group or adding number of events of control group and vice versa (Supplementary Fig.1). High fragility index indicated that the significant or non-significant meta-analysis is relatively robust compared to low fragility index if respective meta-analysis has an equal sample size. As fragility of meta-analysis with large sample size can be overestimated and those with small sample size can be underestimated, fragility quotient was calculated. A fragility quotient is adjusted value that dividing fragility index by sample size.

R Software 4.1.3 was used in calculating descriptive statistics.

3. Results

Of retrieved 248 studies, a total of 12 Cochrane systematic reviews with 48 meta-analyses were analyzed (Supplementary Fig.2). All meta-analyses adopted Mantel-Haenszel method. Details of included meta-analyses including number of trials, total sample size, and total number of events in respective meta-analyses, effect measure, and analysis model, control group interventions are described in Table 1 and supplementary file 1. The median sample size of all meta-analyses was 289 [interquartile range: 177.5 – 708]; the median sample size of statistically significant ones was 723 [interquartile range: 364 – 849]; and the median sample size of statistically non-significant ones was 288 [interquartile range: 150 – 395.5]. Almost half of all meta-analyses and statistically significant and non-significant meta-analyses used fixed effect model.

Table 1

| Characteristics of included meta-analysis. | All (n = 48) | Statistically significant (n = 16) | Statistically non-significant (n = 32) |
|------------------------------------------|-------------|----------------------------------|--------------------------------------|
| Number of trials                         | 2 [2 – 4]   | 3 [2 – 4]                         | 2 [2 – 4]                            |
| Total sample size                        | 289 [177.5 – 708] | 723 [364 – 849]                 | 288 [150 – 395.5]                     |
| Total number of events                   | 122.5 [72.5 – 219.5] | 230 [178 – 364]                 | 90 [53.5 – 149.5]                     |
| Effect measure                           |             |                                  |                                      |
| Risk ratio                               | 39 (81.3)   | 13 (81.3)                        | 26 (81.3)                            |
| Odds ratio                               | 9 (18.7)    | 3 (18.7)                         | 6 (18.7)                             |
| Analysis model                           |             |                                  |                                      |
| Fixed effect                             | 23 (47.9)   | 9 (56.3)                         | 14 (43.8)                            |
| Random effects                           | 25 (52.1)   | 7 (43.7)                         | 18 (56.2)                            |
| Control                                  |             |                                  |                                      |
| Sham                                     | 12 (22.2)   | 5 (31.3)                         | 7 (21.9)                             |
| Usual care                               | 4 (11.2)    | 2 (12.5)                         | 2 (6.3)                              |
| Active control                           | 16 (33.3)   | 6 (37.5)                         | 10 (31.3)                            |
| Others                                   | 16 (33.3)   | 3 (18.7)                         | 13 (40.6)                            |
| Condition                                |             |                                  |                                      |
| Musculoskeletal disorders or chronic pain| 11 (22.9)   | 7 (43.8)                         | 4 (12.5)                             |
| Neuropsychiatric disorders               | 19 (39.6)   | 4 (25.0)                         | 15 (46.9)                            |
| Reproductive system                      | 11 (22.9)   | 2 (12.4)                         | 9 (28.1)                             |
| Others                                   | 7 (14.6)    | 3 (18.8)                         | 4 (12.5)                             |
| Fragility index                          | 6 [2 – 10.5] | 6.5 [2.5 – 12]                  | 6 [2 – 10]                           |
| Fragility quotient                       | 0.017 [0.011 – 0.028] | 0.017 [0.008 – 0.022] | 0.021 [0.011 – 0.033] |

Data are expressed as median [interquartile range] or n (%). Usual care control includes no treatment control. While fragility indices and fragility quotients were calculated in all of included meta-analysis, fragility indices and fragility quotients for respective control groups are calculated only when control group intervention met inclusion criteria of this study. Other control groups include add-on (i.e., intervention group: acupuncture plus other treatment and control group: other treatment only), and combination of control groups (e.g., sham and no treatment were aggregated to one control group).
Of 16 statistically significant meta-analyses from 9 Cochrane systematic reviews, results of 2 statistically significant meta-analyses favored active control over acupuncture.

Fragility indices and fragility quotients are described in Table 1 and Fig. 1. Fragility index and fragility quotient of included meta-analyses were similar irrespective of statistical significance. The median of fragility index of all meta-analyses was 6 [interquartile range: 2 – 10.5] (range: 1-24); the median of fragility index that neutralized 16 statistically significant meta-analysis was 7 [interquartile range: 3 - 12] (range: 1 – 16); and fragility index that can change 32 statistically non-significant meta-analysis was calculated as 6 [interquartile range: 2 - 10] (range: 1 – 24). In two instances, the web tool ² provided the fragility index as the number of events that made non-significant meta-analysis significant while the direction of statistical significance favored control group intervention rather than acupuncture.

Associations of fragility index and fragility quotient with each condition and control group are presented in Fig. 1. Of conditions, meta-analyses of reproductive system related indications (e.g., male and female infertility) demonstrated the highest median fragility index and fragility quotient. While the median of fragility index was 4 in neuropsychiatric disorder and 5 in musculoskeletal disorder or chronic pain, the median of fragility quotient of neuropsychiatric disorder was relatively high. Among control interven-
tions, meta-analyses comparing acupuncture with sham exhibited the widest distribution in fragility index (range of sham: 1 – 24; range of active control: 1 - 21 range of usual care: 1 - 12) and fragility quotient (range of sham: 0.001 - 0.080; range of active control: 0.001 - 0.047; range of usual care: 0.003 - 0.032).

4. Discussion

Out of 248 records retrieved, 12 reviews with 48 meta-analyses were analyzed. Fragility index and fragility quotient for meta-analysis was similar regardless of statistical significance. Among control interventions, meta-analyses comparing acupuncture with sham exhibited minimum and maximum values of fragility index and fragility quotient.

Compared to previous study by Atal et al that analyzed robustness of 906 Cochrane systematic reviews, 2 the overall fragility index of acupuncture meta-analyses in Cochrane systematic reviews was somewhat low (this study: 6 [2 – 10.5] Atal et al: 12 [4 – 33]) while proportion of meta-analyses having fragility index of 5 or less was similar (21 out of 48 meta-analyses: 43.8%, Atal et al: 29%). The proportion of meta-analyses where fragility quotient was less than 0.01 was 20.8% of all meta-analyses, 31.3% of statistically significant meta-analyses (5 out of 16 meta-analyses) and 20.0% of statistically non-significant meta-analyses (5 out of 25 meta-analyses). The findings of this study regarding statistically significant meta-analyses was consistent with that of Atal et al while the proportion of statistically non-significant meta-analyses was relatively fragile in this study. Findings from another previous study that analyzed fragility of meta-analyses of pediatric surgery 5 was also similar to this study in that they reported median and interquartile range of fragility index as 5 [2 – 11] but median and interquartile range of fragility quotient was somewhat low (0.0077 [0.0029 – 0.019]). Similar fragility index and higher fragility quotient of acupuncture meta-analyses can be interpreted that robustness of statistical significance of acupuncture meta-analyses is not inferior to that of other field or other interventions addressed in meta-analyses of Cochrane systematic reviews.

Comparison of fragility index and fragility quotient according to conditions implied that the direction of impact can be inconsistent. Further studies are needed to determine whether condition itself is attributable to these results or other variables related with conditions such as patient characteristics, control interventions, outcome measures are attributable to these.

Among control interventions, meta-analyses comparing acupuncture with sham exhibited the widest range in fragility index and fragility quotient. As sham acupuncture has been adopted in randomized clinical trials as ‘inactive control like placebo’, types are varied and responses following sham acupuncture are different accordingly. 6, 7 Wide range of fragility index and fragility quotient of sham acupuncture reconfirm the varied responses following sham acupuncture.

There are some limitations in this study. First one is about the concept of fragility index itself. There has been some criticism against the usage of fragility index because it is strongly correlated with p value 5 and only addresses robustness of statistical significance of binary outcomes, 1 its intuitive nature can be more easily accepted into stakeholders including clinicians and policy makers than p values. Second, this study only analyzed meta-analysis of Cochrane systematic reviews. Fragility index of meta-analysis of non-Cochrane reviews needs further studies.

Robustness of meta-analyses in Cochrane systematic reviews for acupuncture was similar regardless of statistical significance. Impact of control intervention on fragility index cannot be determined in this study. As fragility index enables intuitive interpretation of meta-analyses, applications on other fields can be useful.

Authorship contribution

This is the sole author's work.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the Korea Institute of Oriental Medicine (grant no. KSN20222210). The funding source had no input regarding the interpretation or publication of the study’s results.

Ethical statement

No ethical approval was required for this study.

Data availability

The data associated with this article will be made available from the author upon reasonable request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.imr.2022.100890.

References

1. Walsh M, Srinathan SK, McAuley DF, et al. The statistical significance of randomized controlled trial results is frequently fragile: a case for a Fragility Index. J Clin Epidemio. 2014;67:622–628 2014/02/11. doi:10.1016/j.jclinepi.2013.10.019.
2. Atal I, Porcher R, Boutron I, et al. The statistical significance of meta-analyses is frequently fragile: definition of a fragility index for meta-analyses. J Clin Epidemio. 2019;111:32–40 2019/04/04. doi:10.1016/j.jclinepi.2019.03.012.
3. Tanjong-Ghogomu E, Tugwell P, Welch V. Evidence-based medicine and the Cochrane Collaboration. Bull NYU Hosp Joint Dis. 2009;67:198.
4. Ahmed W, Fowler RA, McCredie VA. Does Sample Size Matter When Interpreting the Fragility Index? Crit Care Med. 2016;44:e1142–e1143 2016/10/19. doi:10.1097/ CCM.0000000000004197.
5. Schröder A, Muensterer OJ, Oetzmann von Sochaczewski C. Meta-analyses in paediatric surgery are often fragile: implications and consequences. Pediatr Surg Int. 2021;37:363–370 2021/01/18. doi:10.1007/s00383-020-06427-5.
6. Fei YT, Cao HJ, Xia RY, et al. Methodological challenges in design and conduct of randomised controlled trials in acupuncture. Imj. 2022;376:e064345 2022/02/27. doi:10.1136/ijm-2021-064345.
7. Zeng D, Yan X, Deng H, et al. Placebo response varies between different types of sham acupuncture: A randomized double-blind trial in neck pain patients. Eur J Pain. 2022;26:1006–1020 2022/02/08. doi:10.1002/ejp.1924.