Original Investigation | **Public Health**

The Fragility of Statistically Significant Results in Randomized Clinical Trials for COVID-19

Takahiro Itaya, RN, MPH; Yotsuha Isobe, RN, BHS; Sayoko Suzuki, RN, BHS; Kanako Koike, RN, MPH; Masakazu Nishigaki, RN, PhD; Yosuke Yamamoto, MD, PhD

Abstract

IMPORTANCE Interpreting results from randomized clinical trials (RCTs) for COVID-19, which have been published rapidly and in vast numbers, is challenging during a pandemic.

OBJECTIVE To evaluate the robustness of statistically significant findings from RCTs for COVID-19 using the fragility index.

DESIGN, SETTING, AND PARTICIPANTS This cross-sectional study included COVID-19 trial articles that randomly assigned patients 1:1 into 2 parallel groups and reported at least 1 binary outcome as significant in the abstract. A systematic search was conducted using PubMed to identify RCTs on COVID-19 published until August 7, 2021.

EXPOSURES Trial characteristics, such as type of intervention (treatment drug, vaccine, or others), number of outcome events, and sample size.

MAIN OUTCOMES AND MEASURES Fragility index.

RESULTS Of the 47 RCTs for COVID-19 included, 36 (77%) were studies of the effects of treatment drugs, 5 (11%) were studies of vaccines, and 6 (13%) were of other interventions. A total of 138 235 participants were included in these trials. The median (IQR) fragility index of the included trials was 4 (1-11). The medians (IQRs) of the fragility indexes of RCTs of treatment drugs, vaccines, and other interventions were 2.5 (1-6), 119 (61-139), and 4.5 (1-18), respectively. The fragility index among more than half of the studies was less than 1% of each sample size, although the fragility index as a proportion of events needing to change would be much higher.

CONCLUSIONS AND RELEVANCE This cross-sectional study found a relatively small number of events (a median of 4) would be required to change the results of COVID-19 RCTs from statistically significant to not significant. These findings suggest that health care professionals and policy makers should not rely heavily on individual results of RCTs for COVID-19.

Key Points

Question In randomized clinical trials (RCTs) of COVID-19 that report statistically significant results, what is the fragility index, ie, the minimum number of participants who would need to have had a different outcome for the RCT to lose statistical significance?

Findings In this cross-sectional study of 47 RCTs with a total of 138 235 participants that had statistically significant results, the median fragility index was 4. That is, a median of 4 events was required to change the analysis findings from statistically significant to not significant.

Meaning In this study, many RCTs for COVID-19 had a low fragility index, challenging confidence in the robustness of the results.

Author affiliations and article information are listed at the end of this article.

Introduction

Since December 2019, the number of people with COVID-19 has surged worldwide.[1] Information about this newly discovered infectious disease has been widely reported in both traditional and social media, resulting in global awareness of a previously unknown respiratory infection and increased public perception of risk. This emergency situation has pressured researchers to conduct randomized clinical trials (RCTs) immediately, at various study scales and of varied quality.[2] Regardless of the scale and quality of RCTs, the results of each received attention from the general public and health care

researchers, via different media, and people alternated between optimism and despair based on the individual findings of these trials.[3]

In particular, there is risk that the results depend on the number of outcome events, as designing a trial for an expected number of outcome events is unrealistic in an emergent situation. *P* values are likely to change if the number of events is small.[4] Furthermore, *P* values can be affected by methodological limitations, such as loss to follow-up or inadequate blinding. However, there is still a strong reliance on *P* values for quick clinical decisions, despite several statements critiquing the superficial interpretation of *P* values.[5,6]

The fragility index is helpful in interpreting the robustness of results obtained from clinical trials.[7] It outlines the minimum number of participants in a positive trial who would need to have had a different outcome for the results of the trial to lose statistical significance. A lower number on the fragility index indicates that the statistical significance of the trial depends on fewer events. For example, a score of 2 on this measure means that if 2 participants in the intervention group had different event outcomes, the RCT would not have a statistically significant result when using the conventional *P* value cutoff of less than .05 (**Figure 1**). Specifically, *P* values from studies with low fragility indexes should be carefully interpreted because they can change easily depending on the number of events. Thus, the fragility index can be an intuitive indicator for the careful interpretation of clinical trial findings conducted under emergency status. The aim of this study was to evaluate the robustness of statistically significant findings from RCTs for COVID-19 using the fragility index.

Methods

Study Design and Data Source

For this cross-sectional study, we systematically searched PubMed to identify articles reporting RCTs on COVID-19 until August 7, 2021, using the following search strategy: (*COVID-19* OR *COVID-19* [Medical Subject Heading (MeSH) Terms] OR *COVID-19 Vaccines* OR *COVID-19 Vaccines* [MeSH Terms] OR *COVID-19 serotherapy* OR *COVID-19 serotherapy* [Supplementary Concept] OR *COVID-19 Nucleic Acid Testing* OR *covid-19 nucleic acid testing* [MeSH Terms] OR *COVID-19 Serological Testing* OR *covid-19 serological testing* [MeSH Terms] OR *COVID-19 Testing* OR *covid-19 testing* [MeSH Terms] OR *SARS-CoV-2* OR *sars-cov-2* [MeSH Terms] OR *Severe Acute Respiratory Syndrome Coronavirus 2* OR *NCOV* OR *2019 NCOV* OR *coronavirus* [MeSH Terms] OR *coronavirus* OR *COV*) AND (*randomized controlled trial* [Publication Type] OR (*randomized* [Title/Abstract] AND *controlled* [Title/Abstract] AND *trial* [Title/Abstract])) AND (*2019/11/01* [PDAT]: *3000/12/31* [PDAT]).

Figure 1. Example of the Fragility Index Calculation for a Randomized Clinical Trial
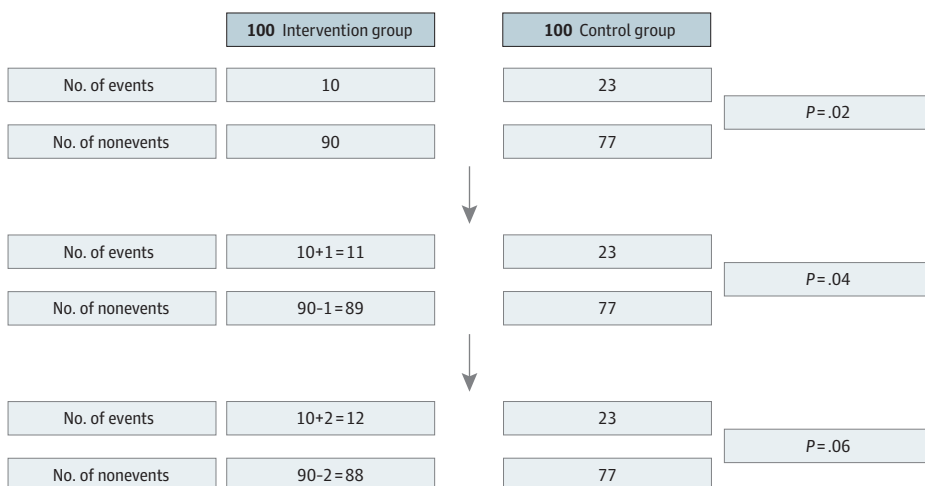
In this example, the original *P* value from the Fisher exact test was .02, and the fragility index was 2. This means that the statistically significant result would not have been significant if 2 cases had changed from nonevents to events in the intervention group.

Per the Common Rule, this study did not require ethical approval because we analyzed only published results and did not include patients. We followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines for cross-sectional studies.

Study Selection

After removing duplicate records from the initial search results, 2 pairs of reviewers (T.I. and K.K.; Y.I. and S.S.) screened the titles and abstracts of all identified articles in accordance with the following prespecified eligibility criteria. The inclusion criteria were RCTs that (1) were superiority trials, (2) randomly assigned patients 1:1 into 2 parallel groups, (3) reported at least 1 dichotomous or time-to-event outcome as statistically significant in the abstract, and (4) tested an intervention for COVID-19. Exclusion criteria were RCTs that were (1) not original articles, (2) preprint articles, (3) phase 1 or 2 trials, (4) noninferiority trials, (5) cluster or crossover RCTs, and (6) non-English articles.

Data Extraction

The 4 reviewers independently extracted data from each trial in duplicate using a prespecified data collection form. Discrepancies were discussed in pairs; if not resolved, they were addressed by a third reviewer from the review team. We extracted the following data: type of intervention (treatment drug, vaccine, or others); outcome definitions (primary or secondary, time-to-event or not, composite or not); analytical strategy (adjusted confounders or not, intention to treat or not); allocation concealment (adequate or no/unclear); the number of participants lost to follow-up; the reported P value; the number of outcome events; the sample size; funding (nonprofit, profit, both, no funding, or not reported).

Outcome

The primary outcome of this study was the fragility index. We calculated the fragility indexes in each RCT based on a previous report.[7] Using 2 × 2 contingency tables, the fragility index was calculated by the iterative addition of an event to the experimental or control group with a smaller number of events and concomitant subtraction of a nonevent from that same group. We continued this calculation until statistical significance (defined as $P < .05$) was lost, while maintaining the total number of events and nonevents. P values were recalculated using a 2-sided Fisher exact test. In terms of time-to-event outcome, based on previous studies,[7] we calculated the fragility index by the number of events and nonevents during the observation period, without considering censoring.

Statistical Analysis

To summarize study characteristics, continuous variables are presented as medians with IQRs, and categorical variables are presented as counts with percentages. We plotted the fragility index as a histogram and described the fragility index by subgroups based on trial characteristics. All statistical analyses were performed using Stata version 16.1 (StataCorp).

Results

Selection Flow

We identified 1187 articles. After excluding duplicate articles and applying the exclusion criteria, 401 articles were deemed eligible for the full-text review. These articles were checked according to the eligibility criteria, and 47 articles, with 138 235 participants, were included in the study.[8-54] At the full-text review stage, 73 articles were studies with binary outcomes but were excluded because they did not have statistically significant results. The detailed study selection flow is presented in **Figure 2**.

Study Characteristics

Table 1 summarizes the characteristics of the included studies. Of the 47 RCTs, 36 (77%) were studies of the effects of treatment drugs, 5 (11%) were vaccines, and 6 (13%) were other topics. The median

(IQR) sample size was 111 (72-392) participants, with a median (IQR) of 44 (18-112) outcome events. Approximately half the trials were conducted based on nonprofit funding.
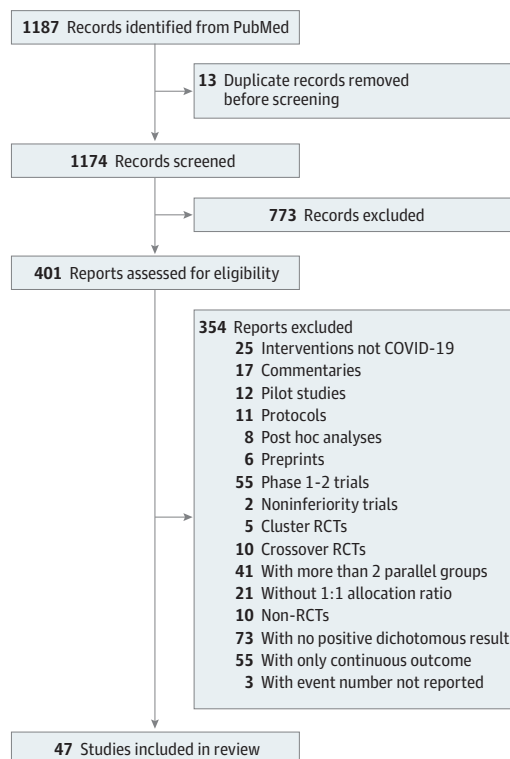
The Fragility Index in COVID-19 Trials

The median (IQR) fragility index for the 47 trials was 4 (1-11): a median of 4 events was required to change the analysis findings from statistically significant to not significant. **Figure 3** shows the distribution of the fragility index for the included studies. We describe the fragility index by subgroups of trial characteristics in **Table 2**. The median (IQR) fragility indexes of RCTs in treatment drugs was 2.5 (1-6); in others it was 4.5 (1-18). In contrast, the median (IQR) fragility index of vaccine trials was 119 (61-139). In addition, among 26 trials (55%), the fragility index was 1% or less of the total sample size.

Discussion

Our study found that the fragility index was 4 or less in 50% of binary outcomes from RCTs on COVID-19 reported in medical journals published until the beginning of August 2021. This result means that for half the COVID-19 trials, reversing the outcome status of 4 patients in the intervention group would change the result from statistically significant to not significant. In terms of types of interventions, most COVID-19 vaccine trials had a large fragility index, whereas most RCTs studying treatment drugs and other interventions had a very small fragility index. In addition, the fragility index among most of the studies was less than 1% of each sample size.

Our findings were consistent with those reported in various clinical fields surveyed before the pandemic, such as spine surgery,[55,56] anesthesia and critical care,[57-59] sports medicine and arthroscopic surgery,[60] and nephrology.[61] These previous studies reported a median fragility index of 2 to 5, which is similar to our results. In addition, consistent with that reported in previous studies,

Figure 2. Study Selection Flow

1187 Records identified from PubMed

13 Duplicate records removed before screening

1174 Records screened

773 Records excluded

401 Reports assessed for eligibility

354 Reports excluded
25 Interventions not COVID-19
17 Commentaries
12 Pilot studies
11 Protocols
8 Post hoc analyses
6 Preprints
55 Phase 1-2 trials
2 Noninferiority trials
5 Cluster RCTs
10 Crossover RCTs
41 With more than 2 parallel groups
21 Without 1:1 allocation ratio
10 Non-RCTs
73 With no positive dichotomous result
55 With only continuous outcome
3 With event number not reported

47 Studies included in review

RCT indicates randomized clinical trial.

the fragility index appeared to be associated with the sample size and *P* values. In this study, the sample size of clinical trials examining vaccines was very large, and the fragility index was large in many of these studies. These RCTs of vaccines not only had large sample sizes, but also a high number of events. This result was consistent with those of previous studies that focused on clinical trials in 5 high-impact medical journals, such as *JAMA* and the *New England Journal of Medicine*,[7] and in heart failure.[62] These RCTs also had both large sample sizes and large numbers of outcome events.

We need to carefully interpret the results of COVID-19 trials with a small fragility index. A small fragility index means that the results may be less robust in terms of statistical significance; in other

Table 1. Characteristics of Included Studies

Characteristic	Studies, No. (%) (N = 47)
Intervention	
Treatment drugs	36 (77)
Vaccines	5 (11)
Others	6 (13)
Outcome	
Primary	23 (49)
Secondary	24 (51)
Time-to-event	6 (13)
Composite	7 (15)
Total sample size, median (IQR)	111 (72-392)
Loss to follow-up, median (IQR)	3 (0-37)
Outcome events, median (IQR), No.	44 (18-112)
Reported *P* value	
<.05-.01	22 (47)
<.01-.001	9 (19)
<.001	11 (23)
Unclear (eg, reported only 95% CI)	5 (11)
Intention-to-treat analysis	25 (53)
Adjusted analysis	8 (17)
Allocation concealment	40 (85)
Funding	
Nonprofit	24 (51)
Profit	5 (11)
Both	6 (13)
No funding	8 (17)
Not reported	4 (9)

Figure 3. Distribution of the Fragility Index for All Studies
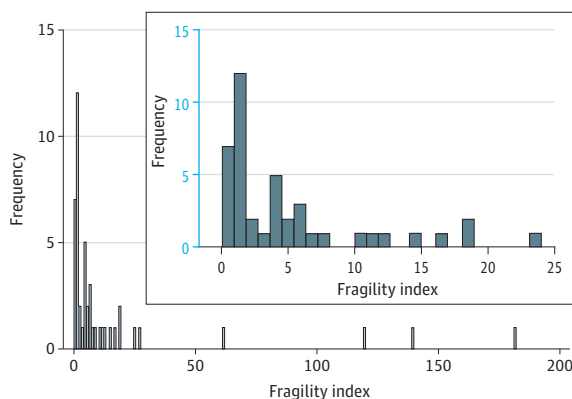
words, a change in the outcome occurrence for a small number of participants in an intervention group can easily change the study result. However, a small fragility index does not imply that the study is not trustworthy. Small RCTs with low fragility indexes may still prove useful if the aggregated

Table 2. Fragility Index by Subgroups Based on Trial Characteristics

Characteristic	No.	Fragility index, median (IQR)
All trials	47	4 (1-11)
Type of intervention		
Treatment drugs	36	2.5 (1-6)
Vaccines	5	119 (61-139)
Others	6	4.5 (1-18)
Outcome		
Primary	23	5 (1-12)
Not primary	24	1.5 (1-6)
Time-to-event	6	4.5 (4-14)
Not time-to-event	41	3 (1-10)
Composite	7	4 (1-11)
Not composite	40	4 (1-11)
Analysis		
Adjusted	8	9 (4.5-129)
Not adjusted	39	2 (1-8)
Intention to treat	25	4 (1-8)
Not intention to treat	22	1 (1-14)
Allocation concealment		
Adequate	40	3.5 (1-7.5)
Unclear	7	14 (1-61)
Loss to follow-up		
≤1%	18	4 (1-7)
>1%-5%	8	1 (0.5-3)
>5%-10%	9	6 (3-11)
>10%	12	3.5 (1-19)
P value		
<.05-.01	22	1 (0-1)
<.01-.001	9	4 (4-6)
<.001	11	12 (6-24)
Unclear	5	61 (4-119)
Outcome events, No.[a]		
6-18	12	1.5 (1-4)
19-44	12	1 (0-7)
45-112	12	5 (1-10)
113-839	11	12 (5-119)
Sample size, No.[a]		
34-72	12	2.5 (0.5-4.5)
73-111	12	1 (1-8)
112-392	12	4 (1-9.5)
393-39 058	11	12 (4-119)
Funding		
Nonprofit	24	3 (1-6)
Profit	5	18 (1-61)
Both	6	5.5 (1-12)
No funding	8	3 (0.5-12.5)
Not reported	4	5.5 (2-16)

[a] The number of events and sample size were divided by IQR into 4 groups.

or the individual patient data they provide can be combined on evidence synthesis platforms, such as the COVID-NMA project.[63]

Strengths and Limitations

Our study had several strengths. We used a systematic and rigid approach to identify all RCTs related to COVID-19. We systematically identified the articles using a predefined search strategy for all articles in PubMed, which is the most commonly used medical literature database. In addition, we included all eligible COVID-19 trials, regardless of publication period; this makes our findings relatively comprehensive for COVID-19 research and reflects the overall state of the evidence currently available.

This study also has limitations. First, the concept of the fragility index can only be applied to trials performing 1:1 randomization and reporting statistically significant findings for binary outcomes.[7] Although many clinically relevant end points have binary outcomes, many articles in this study were excluded because they had more than 2 parallel arms (n = 41), no positive dichotomous outcome (n = 73), and only continuous variables (n = 55). Second, we included only articles written in English. This restriction may have led to selection bias, but as the leading studies on COVID-19 are often published in international journals that are PubMed-listed in English, it is unlikely to have caused major problems. Third, the current study did not assess the study quality and the study protocol of individual RCTs in detail and only focused on the fragility index. We only considered a few major aspects of study quality, such as intention-to-treat analysis and allocation concealment. A study with a large fragility index does not necessarily indicate a good study. A larger sample size is likely to result in a larger fragility index, but ethical considerations require that RCTs recruit the minimum number of participants necessary based on the findings of previous studies. The fragility index is only a metric to ascertain the robustness of clinical trials and should not be used alone to judge the merits of a study. Furthermore, there is no clear cutoff point for the fragility index.[64] Although we have to pay attention to these limitations, the fragility index is an intuitive aid for interpreting RCT results because the simple metric is easy to interpret and may help allay complex concerns regarding smaller trials with fewer events that are difficult to understand intuitively.

Conclusions

In this study, we found that the statistically significant findings of many COVID-19 trials depended on few events. Therefore, health care professionals and policy makers should not rely heavily on individual results of RCTs on COVID-19. The fragility of RCT results should be considered before applying them to clinical settings. Nevertheless, small RCTs with low fragility indexes may still provide robust and useful findings using evidence synthesis platforms.

Corresponding Author: Takahiro Itaya, RN, MPH, Department of Healthcare Epidemiology, Graduate School of Medicine and Public Health, Kyoto University, Yoshida Konoe-cho, Sakyo-ku, Kyoto, 606-8501, Japan (itaya-kyt@umin.ac.jp).

Author Affiliations: Department of Healthcare Epidemiology, Graduate School of Medicine and Public Health, Kyoto University, Kyoto, Japan (Itaya, Yamamoto); Department of Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan (Isobe, Suzuki); Department of Medical Genetics, International University of Health and Welfare Graduate School, Tokyo, Japan (Koike, Nishigaki).

REFERENCES

1. World Health Organization. Coronavirus disease (COVID-2019) weekly epidemiological update and weekly operational update. Accessed February 9, 2022. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/

2. Honarmand K, Penn J, Agarwal A, et al. Clinical trials in COVID-19 management and prevention: a meta-epidemiological study examining methodological quality. *J Clin Epidemiol*. 2021;139:68-79. doi:10.1016/j.jclinepi.2021.07.002

3. The Lancet Infectious Diseases. The COVID-19 infodemic. *Lancet Infect Dis*. 2020;20(8):875. doi:10.1016/S1473-3099(20)30565-X

4. Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med*. 1979;301 (26):1410-1412. doi:10.1056/NEJM197912273012602

5. Wasserstein RL, Lazar NA. The ASA statement on *P* values: context, process, and purpose. *Am Stat*. 2016;70 (2):129-133. doi:10.1080/00031305.2016.1154108

6. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019;567(7748): 305-307. doi:10.1038/d41586-019-00857-9

7. 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 Epidemiol*. 2014;67(6):622-628. doi:10.1016/j.jclinepi.2013.10.019

8. Alessi J, de Oliveira GB, Franco DW, et al. Telehealth strategy to mitigate the negative psychological impact of the COVID-19 pandemic on type 2 diabetes: a randomized controlled trial. *Acta Diabetol*. 2021;58(7):899-909. doi:10.1007/s00592-021-01690-1

9. Aref ZF, Bazeed SEES, Hassan MH, et al. Clinical, biochemical and molecular evaluations of ivermectin mucoadhesive nanosuspension nasal spray in reducing upper respiratory symptoms of mild COVID-19. *Int J Nanomedicine*. 2021;16:4063-4072. doi:10.2147/IJN.S313093

10. Baden LR, El Sahly HM, Essink B, et al; COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med*. 2021;384(5):403-416. doi:10.1056/NEJMoa2035389

11. Barnabas RV, Brown ER, Bershteyn A, et al; Hydroxychloroquine COVID-19 PEP Study Team. Hydroxychloroquine as postexposure prophylaxis to prevent severe acute respiratory syndrome coronavirus 2 infection: a randomized trial. *Ann Intern Med*. 2021;174(3):344-352. doi:10.7326/M20-6519

12. Beigel JH, Tomashek KM, Dodd LE, et al; ACTT-1 Study Group Members. Remdesivir for the treatment of COVID-19—final report. *N Engl J Med*. 2020;383(19):1813-1826. doi:10.1056/NEJMoa2007764

13. Cadegiani FA, McCoy J, Gustavo Wambier C, Goren A. Early antiandrogen therapy with dutasteride reduces viral shedding, inflammatory responses, and time-to-remission in males with COVID-19: a randomized, double-blind, placebo-controlled interventional trial (EAT-DUTA AndroCoV Trial—Biochemical). *Cureus*. 2021;13(2): e13047. doi:10.7759/cureus.13047

14. Cheng LL, Guan WJ, Duan CY, et al. Effect of recombinant human granulocyte colony-stimulating factor for patients with coronavirus disease 2019 (COVID-19) and lymphopenia: a randomized clinical trial. *JAMA Intern Med*. 2021;181(1):71-78. doi:10.1001/jamainternmed.2020.5503

15. Cohen MS, Nirula A, Mulligan MJ, et al; BLAZE-2 Investigators. Effect of bamlanivimab vs placebo on incidence of COVID-19 among residents and staff of skilled nursing and assisted living facilities: a randomized clinical trial. *JAMA*. 2021;326(1):46-55. doi:10.1001/jama.2021.8828

16. Davoudi-Monfared E, Rahmani H, Khalili H, et al. A randomized clinical trial of the efficacy and safety of interferon β-1a in treatment of severe COVID-19. *Antimicrob Agents Chemother*. 2020;64(9):e01061-e20. doi:10.1128/AAC.01061-20

17. Deftereos SG, Giannopoulos G, Vrachatis DA, et al; GRECCO-19 investigators. Effect of colchicine vs standard care on cardiac and inflammatory biomarkers and clinical outcomes in patients hospitalized with coronavirus disease 2019: the GRECCO-19 randomized clinical trial. *JAMA Netw Open*. 2020;3(6):e2013136. doi:10.1001/jamanetworkopen.2020.13136

18. Dilogo IH, Aditianingsih D, Sugiarto A, et al. Umbilical cord mesenchymal stromal cells as critical COVID-19 adjuvant therapy: a randomized controlled trial. *Stem Cells Transl Med*. 2021;10(9):1279-1287. doi:10.1002/sctm.21-0046

19. Edalatifard M, Akhtari M, Salehi M, et al. Intravenous methylprednisolone pulse as a treatment for hospitalised severe COVID-19 patients: results from a randomised controlled clinical trial. *Eur Respir J*. 2020;56(6):2002808. doi:10.1183/13993003.02808-2020

20. Gonzalez-Ochoa AJ, Raffetto JD, Hernández AG, et al. Sulodexide in the treatment of patients with early stages of COVID-19: a randomized controlled trial. *Thromb Haemost*. 2021;121(7):944-954. doi:10.1055/a-1414-5216

21. Grieco DL, Menga LS, Cesarano M, et al; COVID-ICU Gemelli Study Group. Effect of helmet noninvasive ventilation vs high-flow nasal oxygen on days free of respiratory support in patients with COVID-19 and moderate to severe hypoxemic respiratory failure: the HENIVOT randomized clinical trial. *JAMA*. 2021;325(17):1731-1743. doi:10.1001/jama.2021.4682

22. Heath PT, Galiza EP, Baxter DN, et al; 2019nCoV-302 Study Group. Safety and efficacy of NVX-CoV2373 COVID-19 vaccine. *N Engl J Med*. 2021;385(13):1172-1183. doi:10.1056/NEJMoa2107659

23. Hu K, Guan WJ, Bi Y, et al. Efficacy and safety of Lianhuaqingwen capsules, a repurposed Chinese herb, in patients with coronavirus disease 2019: a multicenter, prospective, randomized controlled trial. *Phytomedicine*. 2021;85:153242. doi:10.1016/j.phymed.2020.153242

24. Kalil AC, Patterson TF, Mehta AK, et al; ACTT-2 Study Group Members. Baricitinib plus remdesivir for hospitalized adults with COVID-19. *N Engl J Med*. 2021;384(9):795-807. doi:10.1056/NEJMoa2031994

25. Kratzke IM, Rosenbaum ME, Cox C, Ollila DW, Kapadia MR. Effect of clear vs standard covered masks on communication with patients during surgical clinic encounters: a randomized clinical trial. *JAMA Surg*. 2021;156(4):372-378. doi:10.1001/jamasurg.2021.0836

26. Lenze EJ, Mattar C, Zorumski CF, et al. Fluvoxamine vs placebo and clinical deterioration in outpatients with symptomatic COVID-19: a randomized clinical trial. *JAMA*. 2020;324(22):2292-2300. doi:10.1001/jama.2020.22760

27. Li C, Luo F, Liu C, et al. Effect of a genetically engineered interferon-alpha versus traditional interferon-alpha in the treatment of moderate-to-severe COVID-19: a randomised clinical trial. *Ann Med*. 2021;53(1):391-401. doi:10.1080/07853890.2021.1890329

28. Li L, Zhang W, Hu Y, et al. Effect of convalescent plasma therapy on time to clinical improvement in patients with severe and life-threatening COVID-19: a randomized clinical trial. *JAMA*. 2020;324(5):460-470. doi:10.1001/jama.2020.10044

29. Libster R, Pérez Marc G, Wappner D, et al; Fundación INFANT–COVID-19 Group. Early high-titer plasma therapy to prevent severe COVID-19 in older adults. *N Engl J Med*. 2021;384(7):610-618. doi:10.1056/NEJMoa2033700

30. Lopes MI, Bonjorno LP, Giannini MC, et al. Beneficial effects of colchicine for moderate to severe COVID-19: a randomised, double-blinded, placebo-controlled clinical trial. *RMD Open*. 2021;7(1):e001455. doi:10.1136/rmdopen-2020-001455

31. Lopes RD, de Barros E Silva PGM, Furtado RHM, et al; ACTION Coalition COVID-19 Brazil IV Investigators. Therapeutic versus prophylactic anticoagulation for patients admitted to hospital with COVID-19 and elevated D-dimer concentration (ACTION): an open-label, multicentre, randomised, controlled trial. *Lancet*. 2021;397(10291):2253-2263. doi:10.1016/S0140-6736(21)01203-4

32. Luo Z, Chen W, Xiang M, et al. The preventive effect of Xuebijing injection against cytokine storm for severe patients with COVID-19: a prospective randomized controlled trial. *Eur J Integr Med*. 2021;42:101305. doi:10.1016/j.eujim.2021.101305

33. Mahmud R, Rahman MM, Alam I, et al. Ivermectin in combination with doxycycline for treating COVID-19 symptoms: a randomized trial. *J Int Med Res*. 2021;49(5):3000605211013550. doi:10.1177/03000605211013550

34. McCoy J, Goren A, Cadegiani FA, et al. Proxalutamide reduces the rate of hospitalization for COVID-19 male outpatients: a randomized double-blinded placebo-controlled trial. *Front Med (Lausanne)*. 2021;8:668698. doi:10.3389/fmed.2021.668698

35. Mesri M, Esmaeili Saber SS, Godazi M, et al. The effects of combination of *Zingiber officinale* and echinacea on alleviation of clinical symptoms and hospitalization rate of suspected COVID-19 outpatients: a randomized controlled trial. *J Complement Integr Med*. 2021;18(4):775-781. doi:10.1515/jcim-2020-0283

36. Polack FP, Thomas SJ, Kitchin N, et al; C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. *N Engl J Med*. 2020;383(27):2603-2615. doi:10.1056/NEJMoa2034577

37. Policarpo S, Machado MV, Cortez-Pinto H. Telemedicine as a tool for dietary intervention in NAFLD-HIV patients during the COVID-19 lockdown: a randomized controlled trial. *Clin Nutr ESPEN*. 2021;43:329-334. doi:10.1016/j.clnesp.2021.03.031

38. Rahmani H, Davoudi-Monfared E, Nourian A, et al. Interferon β-1b in treatment of severe COVID-19: a randomized clinical trial. *Int Immunopharmacol*. 2020;88:106903. doi:10.1016/j.intimp.2020.106903

39. Ranjbar K, Moghadami M, Mirahmadizadeh A, et al. Methylprednisolone or dexamethasone, which one is superior corticosteroid in the treatment of hospitalized COVID-19 patients: a triple-blinded randomized controlled trial. *BMC Infect Dis*. 2021;21(1):337. doi:10.1186/s12879-021-06045-3

40. Ravikirti RR, Roy R, Pattadar C, et al. Evaluation of ivermectin as a potential treatment for mild to moderate COVID-19: a double-blind randomized placebo controlled trial in eastern India. *J Pharm Pharm Sci*. 2021;24:343-350. doi:10.18433/jpps32105

41. Réa-Neto Á, Bernardelli RS, Câmara BMD, Reese FB, Queiroga MVO, Oliveira MC. An open-label randomized controlled trial evaluating the efficacy of chloroquine/hydroxychloroquine in severe COVID-19 patients. *Sci Rep*. 2021;11(1):9023. doi:10.1038/s41598-021-88509-9

42. Rocco PRM, Silva PL, Cruz FF, et al; SARITA-2 investigators. Early use of nitazoxanide in mild COVID-19 disease: randomised, placebo-controlled trial. *Eur Respir J*. 2021;58(1):2003725. doi:10.1183/13993003.03725-2020

43. Roostaei Firozabad A, Meybodi ZA, Mousavinasab SR, et al. Efficacy and safety of levamisole treatment in clinical presentations of non-hospitalized patients with COVID-19: a double-blind, randomized, controlled trial. *BMC Infect Dis*. 2021;21(1):297. doi:10.1186/s12879-021-05983-2

44. Roozbeh F, Saeedi M, Alizadeh-Navaei R, et al. Sofosbuvir and daclatasvir for the treatment of COVID-19 outpatients: a double-blind, randomized controlled trial. *J Antimicrob Chemother*. 2021;76(3):753-757. doi:10.1093/jac/dkaa501

45. Rosén J, von Oelreich E, Fors D, et al; PROFLO Study Group. Awake prone positioning in patients with hypoxemic respiratory failure due to COVID-19: the PROFLO multicenter randomized clinical trial. *Crit Care*. 2021; 25(1):209. doi:10.1186/s13054-021-03602-9

46. Sadeghipour P, Talasaz AH, Rashidi F, et al; INSPIRATION Investigators. Effect of intermediate-dose vs standard-dose prophylactic anticoagulation on thrombotic events, extracorporeal membrane oxygenation treatment, or mortality among patients with COVID-19 admitted to the intensive care unit: the INSPIRATION randomized clinical trial. *JAMA*. 2021;325(16):1620-1630. doi:10.1001/jama.2021.4152

47. Sadoff J, Gray G, Vandebosch A, et al; ENSEMBLE Study Group. Safety and efficacy of single-dose Ad26.COV2.S vaccine against COVID-19. *N Engl J Med*. 2021;384(23):2187-2201. doi:10.1056/NEJMoa2101544

48. Shahbaznejad L, Davoudi A, Eslami G, et al. Effects of ivermectin in patients with COVID-19: a multicenter, double-blind, randomized, controlled clinical trial. *Clin Ther*. 2021;43(6):1007-1019. doi:10.1016/j.clinthera.2021.04.007

49. Skipper CP, Pastick KA, Engen NW, et al. Hydroxychloroquine in nonhospitalized adults with early COVID-19: a randomized trial. *Ann Intern Med*. 2020;173(8):623-631. doi:10.7326/M20-4207

50. Supady A, Weber E, Rieder M, et al. Cytokine adsorption in patients with severe COVID-19 pneumonia requiring extracorporeal membrane oxygenation (CYCOV): a single centre, open-label, randomised, controlled trial. *Lancet Respir Med*. 2021;9(7):755-762. doi:10.1016/S2213-2600(21)00177-6

51. Suppan M, Abbas M, Catho G, et al. Impact of a serious game (Escape COVID-19) on the intention to change COVID-19 control practices among employees of long-term care facilities: web-based randomized controlled trial. *J Med Internet Res*. 2021;23(3):e27443. doi:10.2196/27443

52. Wang D, Fu B, Peng Z, et al. Tocilizumab in patients with moderate or severe COVID-19: a randomized, controlled, open-label, multicenter trial. *Front Med*. 2021;15(3):486-494. doi:10.1007/s11684-020-0824-3

53. Wang JB, Wang ZX, Jing J, et al. Exploring an integrative therapy for treating COVID-19: a randomized controlled trial. *Chin J Integr Med*. 2020;26(9):648-655. doi:10.1007/s11655-020-3426-7

54. Zhou S, Feng J, Xie Q, et al. Traditional Chinese medicine shenhuang granule in patients with severe/critical COVID-19: a randomized controlled multicenter trial. *Phytomedicine*. 2021;89:153612. doi:10.1016/j.phymed.2021.153612

55. Evaniew N, Files C, Smith C, et al. The fragility of statistically significant findings from randomized trials in spine surgery: a systematic survey. *Spine J*. 2015;15(10):2188-2197. doi:10.1016/j.spinee.2015.06.004

56. Muthu S, Ramakrishnan E. Fragility analysis of statistically significant outcomes of randomized control trials in spine surgery: a systematic review. *Spine (Phila Pa 1976)*. 2021;46(3):198-208. doi:10.1097/BRS.0000000000003645

57. Ridgeon EE, Young PJ, Bellomo R, Mucchetti M, Lembo R, Landoni G. The fragility index in multicenter randomized controlled critical care trials. *Crit Care Med*. 2016;44(7):1278-1284. doi:10.1097/CCM.0000000000001670

58. Mazzinari G, Ball L, Serpa Neto A, et al. The fragility of statistically significant findings in randomised controlled anaesthesiology trials: systematic review of the medical literature. *Br J Anaesth*. 2018;120(5):935-941. doi:10.1016/j.bja.2018.01.012

59. Grolleau F, Collins GS, Smarandache A, et al. The fragility and reliability of conclusions of anesthesia and critical care randomized trials with statistically significant findings: a systematic review. *Crit Care Med*. 2019;47(3):456-462. doi:10.1097/CCM.0000000000003527

60. Khan M, Evaniew N, Gichuru M, et al. The fragility of statistically significant findings from randomized trials in sports surgery: a systematic survey. *Am J Sports Med*. 2017;45(9):2164-2170. doi:10.1177/0363546516674469

61. Shochet LR, Kerr PG, Polkinghorne KR. The fragility of significant results underscores the need of larger randomized controlled trials in nephrology. *Kidney Int*. 2017;92(6):1469-1475. doi:10.1016/j.kint.2017.05.011

62. Docherty KF, Campbell RT, Jhund PS, Petrie MC, McMurray JJV. How robust are clinical trials in heart failure? *Eur Heart J*. 2017;38(5):338-345.

63. Boutron I, Chaimani A, Meerpohl JJ, et al; COVID-NMA Consortium. The COVID-NMA project: building an evidence ecosystem for the COVID-19 pandemic. *Ann Intern Med*. 2020;173(12):1015-1017. doi:10.7326/M20-5261

64. Chaitoff A, Zheutlin A, Niforatos JD. The fragility index and trial significance. *JAMA Intern Med*. 2020;180(11):1554. doi:10.1001/jamainternmed.2020.4787