

Appealing for efficient, well organized clinical trials on COVID-19

Yang Zhao, Yongyue Wei, Sipeng Shen, Mingzhi Zhang, Feng Chen

Department of Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 210029, China

Contributions: (I) Conception and design: Y Zhao, Y Wei, F Chen; (II) Administrative support: Y Zhao, Y Wei, F Chen; (III) Provision of study materials or patients: Y Zhao, F Chen; (IV) Collection and assembly of data: Y Zhao, SP Shen, MZ Zhang; (V) Data analysis and interpretation: Y Zhao, SP Shen, MZ Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Feng Chen; Yang Zhao. Department of Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 210029, China. Email: fengchen@njmu.edu.cn; yzhao@njmu.edu.cn.

Background: The rapid emergence of clinical trials on COVID-19 stimulated a wave of discussion in scientific community. It is important to understand the characteristics of the ongoing or pending interventional clinical trials on COVID-19.

Methods: We reviewed the characteristics of interventional trials from Chinese Clinical Trial Registration (ChiCTR) and ClinicalTrials.gov. A total of 171 COVID-19-related interventional trials were identified on Feb 22, 2020. These trials are classified into 4 categories based on treatment modalities, including chemical drugs (CDs), biological therapies (BTs), traditional Chinese medicine (TCM) treatments and other therapies.

Results: Our analysis focused on the issues of stage, design, randomization, blinding, primary endpoints (PEs) definition and sample size of these trials. Although most trials use parallel-arm design (88.3%) and randomization (77.2%), blinding is applied in only 25 trials (14.6%). More than half of the trials planned to recruit ≤ 100 patients, indicating a possibility of insufficient statistical power. About one third of trials will recruit severe and critically ill patients. More trials on traditional Chinese medical treatment use 2 or more PEs than those on CDs or biological treatments (57.6%, 39.4% and 40.5%, respectively).

Conclusions: We found some studies with potential defects including unreasonable design, inappropriate PE and small sample size. Clinical trials on COVID-19 should be designed based on scientific rules, ethics and benefits for patients.

Keywords: Clinical trials; COVID-19

Submitted Mar 12, 2020. Accepted for publication Apr 26, 2020.

doi: 10.21037/atm-20-2429

View this article at: <http://dx.doi.org/10.21037/atm-20-2429>

Introduction

As of February 29, 2020, approximately 85,000 cases of coronavirus disease 2019 (COVID-19) have been confirmed worldwide, with nearly 3,000 deaths occurred. Recently, many COVID-19-related interventional clinical trials have emerged in China, a high-incidence area of COVID-19 (1). The study aims of these interventional trials include eradicating virus, relieving pneumonia symptoms and promoting recovery, etc. The rapid emergence of these trials stimulated a wave of discussion on whether it is appropriate to start so many clinical trials at the critical time point of the battle with the coronavirus.

Recently, a group of Chinese scientists made an urgent call for raising the scientific rigorousness of clinical trials on COVID-19 (2). In order to understand the characteristics of the ongoing or pending interventional clinical trials on COVID-19, we reviewed the information of these trials registered at Chinese Clinical Trial Registration (ChiCTR) or ClinicalTrials.gov (3,4).

Methods

The data of interventional trials from ChiCTR and ClinicalTrials.gov were retrieved updated on February

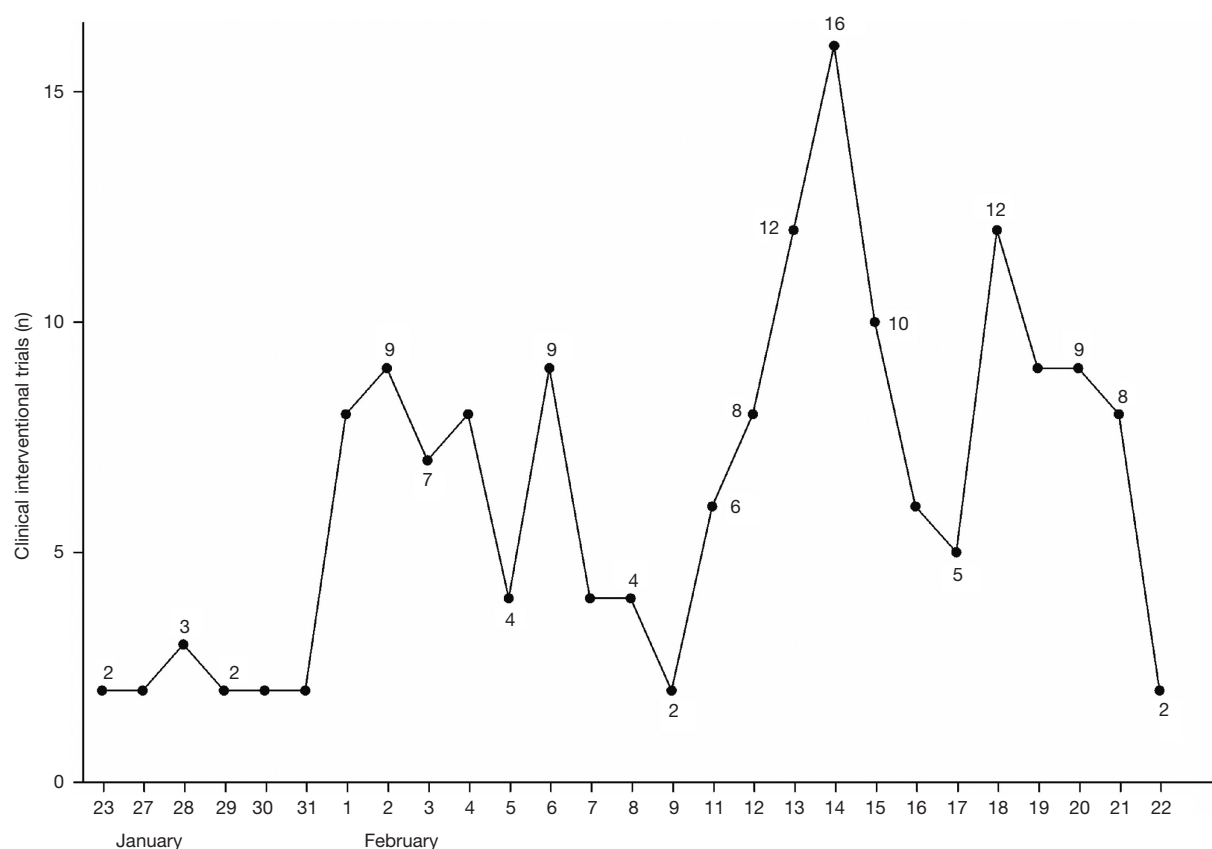


Figure 1 The registration date distribution of the interventional clinical trials.

22, 2020. Two authors (Zhao and Shen) independently extracted the relevant information from the downloaded records, including study phase, study design, presence or absence of randomization, blinding, sample size, severity of disease and source of samples.

To assess the quality of these trials, we used a self-developed scale which involved the applicable items from criteria selected from the Jadad score, Delphi and Cochrane Collaboration (5). The scale includes 7 questions:

- (I) Was the study described as randomized?
- (II) Was the method used to generate the sequence of randomization described and appropriate?
- (III) Were the eligibility criteria specified?
- (IV) Were the patients blinded to the treatment?
- (V) Were the care provider and investigator blinded to the treatment?
- (VI) Did the study provide primary efficacy endpoint?
- (VII) Did the study provide safety endpoint?

We counted the number of positive responses to each question as the score of the quality of a trial.

Descriptive statistics (frequencies and proportions) are used to present the characteristics of the trials. All analyses were performed using R (version 3.6.3).

Results

A total of 171 COVID-19-related interventional trials were identified (138 from ChiCTR and 33 from ClinicalTrials.gov). The registration date distribution of these trials was shown in *Figure 1*. Of the trials registered at ChiCTR, 120 were approved by the institutional review boards. The first 2 trials were registered on Jan 23, 2020, the first day of the “lockdown” of Wuhan city.

The 171 trials are classified into 4 categories based on treatment modalities, including chemical drugs (CDs), biological therapies (BTs), traditional Chinese medicine (TCM) treatments and other therapies. The biological treatments include PD1, Bevacizumab or other monoclonal agents, cell treatments (e.g., NK cells) and convalescent plasma therapy. TCM treatments include both Chinese

medicines (TCMs) and Chinese physical treatments, such as “Qigong” and “Taiji”. The characteristics of these trials were summarized in *Table 1*. The mean quality scores for CDs, BTs, TCMs and other therapies are 4.5, 4.0, 4.2 and 4.1, respectively.

There are 11 confirmatory Phase III or Phase II/III studies (6.4%), 8 of which are for CDs including Remdesivir, Ritonavir, Triavevirin, Leflunimide, Glucocorticoid, Hydroxychloroquine, Darunavir and Cobicistat. Two Phase III trials for biological treatments investigate the effect of Bevacizumab or Intravenous Immunoglobulin. There is also a Phase III trial on Yinhu Qingwen Decoction, a combination of 12 TCMs.

Most trials (88.3%) use parallel-arm design. Randomization is applied in 132 trials (77.2%), while blinding in only 25 trials (14.6%). The proportion of using blinding (single, double, triple or quadruple) is higher for trials on CDs (24.6%) than on the others.

The sample sizes of 45.6% trials are more than 100. Two trials on TCM treatments are planned to recruit 800 or more patients. Severe and critically ill patients are respectively recruited in 70 (40.9%) and 28 (16.4%) trials. More than half of the trials (72.4%) are planned to collect blood samples, while only 19 trials (13.8%) to collect nasopharyngeal swabs.

Most primary endpoints (PEs) are efficacy-related, while only 1 trial also included safety in the PEs. There are 113 trials (66.1%) including safety as secondary endpoint. More trials on TCM treatment use 2 or more PEs than those on CDs or BTs (57.6%, 39.4% and 40.5%, respectively). Some trials even have 6 or more PEs. The most 3 commonly used PEs are associated with clinical symptoms (73.1%), pathogen (29.2%) and lung function (27.5%). Trials on CDs have greater proportions of pathogen (41.0%) or fatality (18.0%) related PEs than the others. It is worth noting that these trials have registered over 300 distinct PEs totally.

Discussion

During the COVID-19 outbreak, we are in urgent need of effective treatment strategies. Through analyzing currently registered interventional trials, we found some studies with potential defects including unreasonable design, inappropriate PE definition, small sample size and ethical issue.

There are 17 trials (9.9%) using single arm designs. Without an appropriate control arm, we cannot exclude the possibility that the recovery of patients is due to the natural

progress of the disease or other concomitant therapies. There are 132 trials registered as randomized trials. Randomization is a widely used technique in clinical trials to remove both measured and un-measured confounding effects. Without randomization, the “causal” treatment effect cannot be estimated as we are not sure whether there are un-measured confounders possibly biased the analysis. Only 17 trials used double, triple or quadruple blinding, while 8 trials used single blinding. As human behavior is influenced by what they know or believe, it is important to use blinding (sometimes called masking) to remove the bias resulted from the subjectivity in assessment. Recently, Cai *et al.* reported the result from an open-label control study on the efficacy of Favipiravir (FPV) for COVID-19 (6). Although their findings demonstrated a positive improvement on the viral clearance of FPV than the control therapy, the authors acknowledged the limitations that the study “is not a randomized, double-blind, parallel trial, further well-designed and large-scale confirmatory trials are warranted”. Methods of sensitivity analysis, such as confounding functions or bounding factors, can be applied when analyzing data from trials with potential bias (7).

Determining the appropriate PEs is crucial to address the primary scientific question of a clinical trial. About one third trials used more than 2 PEs. Trials on TCMs were more likely to use more than 2 PEs than trials on CDs or BTs. Using too many PEs in a trial may inflate the type I error rate. As an example, we noticed that some trials registered 10 PEs. Suppose that the significance level for each PE is 0.05, the overall type I error rate of that trial would be $1 - (0.95)^{10} = 0.4$, which is far more beyond the level that we can tolerate, especially in confirmatory trials. Some may argue that using more PEs may increase the probability to identify a useful treatment. However, the possibility of false discovery is also increased. Forwarding too many useless treatments to late phase trials, would exhaust the already limited medical resources. Also, the PE should be defined clearly and specifically to improve the operation quality. Objective endpoints, such as pathogen or imaging results, are preferred to avoid the bias during evaluation. We also noticed that it is possible that some investigators confused the definitions of PE and measurement of efficacy.

It is possible that some trials lack statistical power. Assuming a two-side significance level of 0.05 and a power of 80%, a trial on mild pneumonia patients should recruit nearly 1,000 patients if the effective rate increases from 90% to 95% for the new treatment. More patients are needed for the trials planning to decrease the fatality of

Table 1 Characteristics of the 171 interventional clinical trials [n (%)]

Characteristics of the trials	Chemical drugs (n=61)	Biological therapies (n=42)	TCM treatment (n=59)	Other therapies (n=9)	Total (n=171)
Phase					
I and II	20 (32.8)	27 (64.3)	28 (47.5)	6 (66.7)	81 (47.4)
II/III and III	8 (13.1)	2 (4.7)	1 (1.7)	0 (0)	11 (6.4)
IV	22 (36.1)	7 (16.7)	16 (27.1)	0 (0)	45 (26.3)
Others	11 (18.0)	6 (14.3)	14 (23.7)	3 (33.3)	34 (19.9)
Design					
Single-arm	5 (8.2)	8 (19.0)	4 (6.8)	0 (0)	17 (9.9)
Parallel-arm	56 (91.8)	34 (81.0)	54 (91.5)	7 (77.8)	151 (88.3)
Factorial design*	0 (0.0)	0 (0.0)	1 (1.7)	2 (22.2)	3 (1.8)
Randomization					
No	4 (6.6)	4 (9.5)	9 (15.2)	0 (0)	17 (9.9)
Yes	50 (82.0)	28 (66.7)	46 (78.0)	8 (88.9)	132 (77.2)
Not applicable	7 (11.4)	10 (23.8)	4 (6.8)	1 (11.1)	22 (12.9)
Blinding					
Open	23 (37.7)	16 (38.0)	21 (35.6)	3 (33.3)	63 (36.9)
Single	5 (8.2)	2 (4.8)	1 (1.7)	0 (0)	8 (4.7)
Double/triple/quadruple	10 (16.4)	2 (4.8)	5 (8.5)	0 (0)	17 (9.9)
Unavailable	14 (23.0)	11 (26.2)	23 (39.0)	4 (44.4)	52 (30.4)
Not applicable	9 (14.7)	11 (26.2)	9 (15.2)	2 (22.2)	31 (18.1)
Sample size					
≤50	15 (24.6)	17 (40.5)	3 (5.1)	1 (11.2)	36 (21.1)
50–100	21 (34.4)	13 (31.0)	20 (33.9)	3 (33.3)	57 (33.3)
100–400	23 (37.7)	10 (23.8)	31 (52.5)	3 (33.3)	67 (39.2)
>400	2 (3.3)	2 (4.7)	5 (8.5)	2 (22.2)	11 (6.4)
Number of primary endpoints					
1	37 (60.6)	25 (59.5)	25 (42.4)	1 (11.1)	88 (51.5)
2	12 (19.7)	5 (11.9)	11 (18.6)	4 (44.4)	32 (18.7)
≥3	12 (19.7)	12 (28.6)	23 (39.0)	4 (44.4)	51 (29.8)
Type of primary endpoint [#]					
Nucleic acid amplification testing	25 (41.0)	10 (23.8)	13 (22.0)	2 (22.2)	50 (29.2)
Fatality	11 (18.0)	4 (9.5)	5 (8.5)	0 (0)	20 (11.7)
Clinical symptoms	39 (63.9)	28 (66.7)	51 (86.4)	7 (77.8)	125 (73.1)
Imaging	13 (21.3)	8 (19.0)	11 (18.6)	0 (0)	32 (18.7)
Lung function	14 (23.0)	12 (28.6)	17 (28.8)	4 (44.4)	47 (27.5)
TCM symptoms	0 (0)	0 (0)	7 (11.9)	0 (0)	7 (4.1)
Immunology	4 (6.6)	5 (11.9)	4 (6.8)	1 (11.1)	14 (8.2)

Table 1 (continued)

Table 1 (continued)

Characteristics of the trials	Chemical drugs (n=61)	Biological therapies (n=42)	TCM treatment (n=59)	Other therapies (n=9)	Total (n=171)
Safety endpoints					
No	15 (24.6)	13 (31.0)	25 (42.4)	6 (66.7)	59 (34.5)
Yes	46 (75.4)	29 (69.0)	34 (57.6)	3 (33.3)	112 (65.5)
Severity of pneumonia [#]					
Mild	46 (75.4)	27 (64.3)	36 (61.0)	4 (44.4)	113 (66.1)
Severe	19 (31.1)	32 (76.2)	17 (28.8)	2 (22.2)	70 (40.9)
Critically ill	7 (11.5)	11 (26.2)	9 (15.3)	1 (11.1)	28 (16.4)
Samples ^{**,#}					
No	1 (2.3)	2 (6.5)	6 (10.9)	4 (44.4)	13 (9.4)
Blood	28 (65.1)	25 (80.6)	44 (80.0)	3 (33.3)	100 (72.4)
Nasopharyngeal swabs	11 (25.6)	3 (9.7)	3 (5.5)	2 (22.2)	19 (13.8)
Respiratory secretions	5 (11.6)	1 (3.2)	2 (3.6)	0 (0)	8 (5.8)
Bronchoalveolar lavage fluid	2 (4.7)	0 (0)	0 (0)	0 (0)	2 (1.4)
Stool	1 (2.3)	0 (0)	0 (0)	0 (0)	1 (0.7)
Quality score					
Mean ± SD	4.5±1.1	4.0±1.2	4.2±1.2	4.1±0.9	4.3±1.2

^{*}, three trials are registered as “factorial design”, one for TCM, one for psychological intervention to doctors and nurses, and one for probiotics. However, unlike the general purpose of factorial design, none of the 3 trials aim to find the best “combination” of the treatments. ^{**}, trials registered at ClinicalTrials.com do not provide information about the collected samples. [#], percentages for PE type are calculated by the number of trials using this PE divided by the number of corresponding type of trials. Percentages of severity and samples are calculated in the same manner.

critically ill patients. Many trials have high probabilities of failing as most registered trials only try to recruit less than 400 patients. Given that there are less and less confirmed COVID-19 patients in China, these powerless trials would exhaust limited patients resources.

Some drugs or therapies under study are short of previous theoretical support for the treatment of COVID-19. As an example, a trial which would recruit more than 200 patients tried to investigate the effect of “Qigong” on the recovery of lung functions for critically ill patients. There is no solid evidence that these therapies may be beneficial to the patients, leading to non-compliance to ethical standards. These unnecessary trials may also waste medical resources, as well as diverting patients’ resources. Recently, two trials on the efficacy of Remdesivir on critically ill Chinese patients had been terminated due to the lack of patients. WHO has suggested an establishment of a centralized research program to ensure the most promising

researches (8).

During the preparation of this article, we have noticed that while the number of registered trials keep rising (ChiCTR: 282 on April 18, 2020), only 1,734 confirmed COVID-19 patients are still under treatment as on April 18, 2020. On April 6, 2020, the state council of the People’s Republic of China has issued a file urging the medical institutions to regulate the clinical trials for COVID-19. Clinical trials on COVID-19 should be designed based on scientific rules (appropriate controls, randomization, blinding, and sufficient sample size), ethics and patients’ benefits. The standards for clinical trials should not be compromised even in such a special time.

Acknowledgments

The authors are grateful to Miss Yuanping Yue and Mrs. Yue Wang for their help on the information extraction

during the preparation of the manuscript.

Funding: This study was funded by the Special Program of National Natural Science Foundation of China on tracing, pathogenesis, prevention and treatment for COVID-19 (No. 82041024 to FC, the National Natural Science Foundation of China (No. 81872709 to YZ) and Key Project of the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No. 18KJA110004 to YZ).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm-20-2429>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the

original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Maxmen A. More than 80 clinical trials launch to test coronavirus treatments. *Nature* 2020;578:347-8.
2. Chen F, Hao YT, Zhang ZJ, et al. An urgent call for raising the scientific rigorousness of clinical trials on COVID-19. *Zhonghua Liu Xing Bing Xue Za Zhi* 2020;41:301-2.
3. Chinese Clinical Trial Registration. Available online: <http://www.chictr.org.cn>. Accessed February 22 2020.
4. ClinicalTrials.gov. Available online: <https://clinicaltrials.gov/>. Accessed February 22 2020.
5. Berger VW, Alperson SY. A general framework for the evaluation of clinical trial quality. *Rev Recent Clin Trials* 2009;4:79-88.
6. Cai Q, Yang M, Liu D, et al. Experimental Treatment with Favipiravir for COVID-19: An Open-Label Control Study. *Engineering* 2020. doi: 10.1016/j.eng.2020.03.007.
7. Kasza J, Wolfe R, Schuster T. Assessing the impact of unmeasured confounding for binary outcomes using confounding functions. *Int J Epidemiol* 2017;46:1303-11.
8. WHO. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19) 2020 February 24.

Cite this article as: Zhao Y, Wei Y, Shen S, Zhang M, Chen F. Appealing for efficient, well organized clinical trials on COVID-19. *Ann Transl Med* 2020;8(10):632. doi: 10.21037/atm-20-2429