Visualization analysis of the characteristics of COVID-19 clinical trial

**CURRENT STATUS:** UNDER REVIEW

Xiaoqiong Cai  
Shanghai University of Engineering Science

Zengliang Zheng  
Shanghai University of Engineering Science

Qianmin Su  
Shanghai University of Engineering Science

Corresponding Author
suqm@sues.edu.cn

ORCID: https://orcid.org/0000-0002-9885-6485

Jihan Huang  
Shanghai University of Traditional Chinese Medicine

**DOI:** 10.21203/rs.3.rs-21472/v1

**SUBJECT AREAS**  
Infectious Diseases

**KEYWORDS**  
COVID-19, Clinical trial, Antivirals, Primary endpoint, Data visualization
Abstract
Objectives: This article points out the characteristics and trends of COVID-19 clinical trials through data collection, translation, mining and visualization to help in clinical trial design.

Method: The registered data of COVID-19 clinical trials are gathered from the Chinese Clinical Trial Registry and ClinicalTrials.gov website transformed by Python, further demonstrated by visual tools.

Results: As of 24:00 on March 28, 2020, totally 732 trial registration records have been retrieved. Overall, there are 406 (55.46%) interventional studies and 271 (37.02%) observational studies. Among interventional studies, 38.93% are randomized parallel trials, 55 (13.55%) trials considered time condition for clinical recovery as the primary endpoint, and 46 (11.33%) trials through clinical parameters and laboratory index as the primary endpoint. In the selection of intervention measures, chemical or biological agents was under the responsibility of 43.60%, of which antivirals accounted for 14.53%, antimalarials accounted for 8.87%, and 98 cases (24.14%) of studies involving Traditional Chinese Medicine or Integrated Medicine. In addition, joint network analysis of antivirals to explore the combination of drugs is further conducted.

Conclusions: By Mining characteristic information of topical COVID-19 clinical trial registration, this article deserves further trial design ideas for researchers to enhance the effects.

Introduction
Since the discovery of pneumonia cases caused by COVID-19 in December 2019, the pandemic spreads rapidly, and medical researchers and clinicians have concentrated on this without hesitation. On January 31, 2020, World Health Organization (WHO) declared in Geneva that COVID-19 was a Public Health Emergency of International Concern (PHEIC).\(^1\)

In order to test the effects of intervention, international researchers have been done a large number of related medical scientific research projects. Among them, clinical trials played an important role in the prevention and control of the COVID-19 in order to verify the efficacy of clinical treatment of drugs and form the necessary evidence.\(^2\) As more and more medical institutions conduct clinical trials, there are also some researchers who research and evaluate them. Numerous experts and researchers analyze the data, but short of profound consideration. In view of this, this article further
explores the research characteristics and problems of COVID-19 clinical trial registration by processing and assessing the relevant data from the Chinese Clinical Trial Registration Center database and ClinicalTrials.gov website.

Materials And Methods

**Data acquisition strategy**

533 trial registration data were retrieved from the Chinese Clinical Trial Registry (ChiCTR) platform using "COVID-19" as the keyword. The source code of the pages was acquired, and rules for the collection of unstructured data were developed using programming methods. Clinical trial registration fields obtained include: registration number, registration time, sample size, study type, study design, blinding, primary indicator, intervention and randomized method. In the ClinicalTrials.gov website, 199 relevant trial data were retrieved with the keyword "COVID-19" and all available field data were downloaded. All of the above data were included in 732 registrations as of March 28.

**Data processing**

This article uses Pandas, Openpyxl libraries and regular expressions in Python to normalize the data and replace the missing values, transform them into profitable discrete structures and construct data tables for subsequent analysis.

In order to prevent bias in the data screening and conversion process, the two authors of this article divided the data and compared them separately. The accuracy of the results is then ensured by consulting with the relevant medical staff on the disputed data.

Results

**Number of COVID-19 clinical trials registered in ChiCTR and ClinicalTrials.gov**

According to statistics on COVID-19 clinical trial registrations (Figure 1), it can be seen that very few clinical trials were registered in January. Owing to the outbreak of the pandemic, daily registrations have gone ballistic. In February, a maximum of single-day registrations is 22, with an average number of 11.69. Daily registrations peaked at 32 on March 15, and the average in this month reached 13.79. In a word, overall registrations in March are higher than the previous two months.
By analyzing the trend of overall registrations (Figure 2), it can be observed that the number was a generally uptrend over time, which has declined since March 17 and increased significantly after the 24. In addition, it was predicted that the coming period, registrations would go up continuously because of the activity of vaccine and unlisted drug clinical trials.

**Basic characteristics of clinical trials**

Across all programs, 355 clinical trials are recruiting volunteers and 316 clinical trials are not yet recruiting. A total of 436 clinical trials have been approved by ethics committees and 97 clinical trials have not been submitted for approval (Table 1). Ethics committee approval is a prerequisite for clinical trial registration, but the timeliness of COVID-19 has resulted in some trials not being registered in time.

Table 1 Statistics of basic characteristics for clinical trial registrations (n=732)

| Category                        | Information                      | Number of trials |
|---------------------------------|----------------------------------|------------------|
| Recruiting status               | Recruiting                       | 355              |
|                                 | Not yet recruiting               | 316              |
|                                 | Completed                        | 28               |
|                                 | Withdrawn                        | 22               |
|                                 | Active, not recruiting           | 5                |
|                                 | Enrolling by invitation          | 3                |
|                                 | Available                        | 3                |
| Approved by the Ethics Committee| Yes                              | 436              |
|                                 | No                               | 97               |
|                                 | N/A                              | 199              |

**Study Types and Methods**

Summary statistics for the study types are in Table 2. There are 406 interventional trials. The main research methods are Randomized Parallel Assignment (38.93%), Non-randomized Assignment
(5.33%), and Single Group Assignment (6.83%). Parallel control designs are observed experimentally at the same time, which is beneficial to eliminate errors caused by factors such as time and conditions. Since there is not any specific drug for COVID-19 and it has a longer survival period than other epidemics, most of interventional trials designed by parallel assignment. Among the 271 observational trials, Sequential Assignment (18.72%), Cohort (6.96%), and Factorial Assignment (4.51%) were preferable owing to a largeish sample size. In addition, there were 28 Diagnosis, 10 Epidemiological and 17 other trials.

Table 2 Study types for all clinical trials based on ChiCTR and ClinicalTrials.gov (n=732)

| Study Type   | Model                   | ChiCTR | ClinicalTrials.gov | Total (%) |
|--------------|-------------------------|--------|--------------------|-----------|
| Intervention | Randomized Parallel Assignment | 201    | 84                 | 285 (38.93%) |
|              | Non-randomized Assignment | 31     | 8                  | 39 (5.33%)   |
|              | Single Group Assignment  | 22     | 28                 | 50 (6.83%)   |
|              | Sequential Assignment    | 11     | 7                  | 18 (2.46%)   |
|              | Crossover Assignment     | 1      | 2                  | 3 (0.41%)    |
|              | Other                    | 11     | 0                  | 11 (1.50%)   |
| Observational| Sequential Assignment    | 137    | 0                  | 137 (18.72%) |
|              | Factorial Assignment     | 33     | 0                  | 33 (4.51%)   |
|              | Case-Only                | 12     | 9                  | 21 (2.87%)   |
|              | Non-randomized Assignment| 9      | 0                  | 9 (1.23%)    |
|              | Cohort                   | 6      | 43                 | 49 (6.69%)   |
|              | Case-Control             | 3      | 9                  | 12 (1.64%)   |
|              | Other                    | 4      | 6                  | 10 (1.37%)   |
| Diagnosis    | Sequential Assignment    | 22     | 0                  | 22 (3.01%)   |
|              | Factorial Assignment     | 5      | 0                  | 5 (0.68%)    |
|              | Cross-Sectional          | 1      | 0                  | 1 (0.14%)    |
| Epidemiological| Sequential Assignment    | 8      | 0                  | 8 (1.09%)    |
|               | Case-Control             | 1      | 0                  | 1 (0.14%)    |
|               | Single Group Assignment  | 1      | 0                  | 1 (0.14%)    |
| Other        |                         | 14     | 3                  | 17 (2.32%)   |
|              | Total                    | 533    | 199                | 732 (100%)   |
Blinding is one of the important tools for reducing subjects’ perceptions of treatment allocation schemes between groups, reducing bias and improving the scientific and validity of trials. In this statistic, as can be seen from Table 3, 268 studies were blinded and 464 (26.33%) were unspecified. In addition, the subgroups of the study are counted in this article. 356 trials involved two groups; 285 trials involved a single group. Fifty-five and 22 trials involved groups three and four, with the remaining groups used less frequently. During COVID-19, most trials were divided into two groups for controlled studies. This grouping allows differences between groups to be efficiently assessed for "potent drugs" to deal with the virus.

Table 3 Blinding and subgroups for all clinical trial registrations (n=732)

| Category       | Information   | Number of trial |
|----------------|---------------|-----------------|
| Blinding       | Open Label    | 196             |
|                | Single        | 23              |
|                | Double        | 21              |
|                | Triple        | 7               |
|                | Quadruple     | 21              |
|                | N/A           | 464             |
| Numbers of group | 1           | 285             |
|                | 2             | 356             |
|                | 3             | 55              |
|                | 4             | 22              |
|                | 5             | 9               |
|                | 6             | 2               |
|                | 7             | 1               |
|                | 8             | 2               |

**Sample size**

Sample size is an important aspect of clinical trial design, directly related to the reliability, reproducibility, and efficiency of the study. It was noted that the sample size of clinical trials was mainly concentrated in the range of 0-299, with a larger proportion of interventional and
observational trials (Figure 3). Whereas, the number of observational trials increased with the expansion of the sample size interval in the larger sample size interval.

**Primary endpoint of Interventional study**

When designing a clinical trial, one of the most challenging and critical issues is how to select the primary endpoint used to assess efficacy. In view of the reliable evidence for benefits and risks should be provided, the primary endpoint is preferably a outcome measure that clearly reflects the benefit to the patients. Among 406 interventional studies, time to clinical recovery accounted for 13.5%, clinical parameters and laboratory index for 11.33%, the change of pneumonia severity for 11.08%, questionnaire or scales for 9.85%, virus negative conversion rate of time for 8.13 %, cure rate for 7.39%, and the mortality for 6.16% (Figure 4).

**Interventions**

This article counted the interventions of 406 interventional studies. Details are shown in Table 4. Chemical or Biological drugs accounted for 43.60%, of which 14.53% were antivirals, 8.87% were antimalarial drugs, and 7.88% were antineoplastics and immunomodulators. The trials involving Traditional Chinese Medicine (TCM) or Integrated Medicine accounted for 24.14%, and Cell Therapy, Behavioral Intervention and Medical Instruments accounted for 8.62%, 7.14% and 4.93%. In addition, five vaccine clinical trials for COVID-19 are underway and their registration numbers are NCT04299724, NCT04276896, NCT04283461, NCT04313127 and NCT04324606.

Table 4 Statistics of interventions (n=406)
| Category                                      | Number of trials | Percentage of total(%) |
|----------------------------------------------|------------------|------------------------|
| Chemical or Biological                       | 177              | 43.60%                 |
| Antivirals                                   | 59               | 14.53%                 |
| Antimalarials                                | 36               | 8.87%                  |
| Antineoplastics and Immunomodulators         | 32               | 7.88%                  |
| Antipyretic Analgesics                       | 19               | 4.68%                  |
| Glucocorticoids                              | 9                | 2.22%                  |
| Mucolytics                                   | 6                | 1.48%                  |
| Other                                        | 16               | 3.94%                  |
| Traditional Chinese Medicine or Integrated Medicine | 98           | 24.14%                 |
| Cell Therapy                                 | 35               | 8.62%                  |
| Plasma                                       | 15               | 3.69%                  |
| Vaccine                                      | 5                | 1.23%                  |
| Medical Instruments                          | 20               | 4.93%                  |
| Behavioral Intervention                      | 29               | 7.14%                  |
| Psychological intervention                   | 4                | 0.99%                  |
| Other                                        | 23               | 5.67%                  |
| Total                                        | 406              | 100%                   |

**Combination of antiviral drugs and other drugs**

This part of the research further studied the clinical trials of the 59 antiviral drugs mentioned, and analyzed the combined use of antivirals with other drugs (Figure 5). Among them, the most commonly used antivirals are Lopinavir/ritonavir (LPV/r), Arbidol and Ribavirin. The drug with which they are most frequently combined is Interferon, followed by Favipiravir with Tocilzumab and Darunavir with Cobicistat.

**Discussion**

As main means of clinical evidence, clinical trials play an important part of clinical research. For safety concerns, clinical data should be strictly and efficiently managed during COVID-19. Accurate decisions can only be made if medical records are uniformly submitted in a timely manner and analyzed after reasonable cleaning. At present the number of clinical trials continued to increase, with
732 relevant clinical trials registered as of 24:00 on 28 March 2020.

Based on the trial data, it was found that clinical trial studies related to novel coronaviruses are very active. The approach of randomized parallel control and single group trials has been taken among interventional studies, most of which chose not to be blinded. Randomized controlled trial studies are the "Gold Standard" for evaluating causal effects in clinical research, researchers should choose the appropriate design type according to actual situation when conducting the study. Overall, the majority of interventional studies in registered clinical trials for COVID-19 have adopted randomized parallel controlled trials so that subjects in clinical trials have an equal chance of being assigned to either trial group or control group, independent of their subjective will, to ensure the objectivity of clinical trial results. In view of the current diversity of cases, the scientific and efficient completion of clinical trials using different design combinations can be considered as a very important research value for both treatment and prevention of the virus.

The primary clinical endpoint should be targeted based on factors such as the nature of the trial and subject status. For example, mild patients have a lower mortality rate, and researchers can set clinical recovery time or cure rate as primary endpoint. One of the reasons for the difficulty of COVID-19 prevention and control is the high specificity and low sensitivity of the nucleic acid test for the virus, which does not exclude the presence of some false-negative patients. In order to improve the accuracy of detection, some institutions conduct clinical studies on virus detection methods, so that the primary endpoint can be chosen as the virus detection.

The current TCM or integrated medicine trial are relatively active, accounting for 24.14% of the 406 intervention studies. Integrated medicine can both complement and improve clinical efficacy. At present, some researchers have found that Lianhuaqingwen has antiviral and anti-inflammatory properties against COVID-19, on the basis of which an in-depth study of may bring some positive effects to the treatment of COVID-19.

In terms of antiviral drug selection, most trials have chosen to combine LPV/r with interferon, a similar design approach that was used in clinical trials for the treatment of Middle East Respiratory Syndrome
Whereas in critically ill COVID-19 adult hospitalized patients, study team did not observe a significant therapeutic effect of LPV/r compared to standard treatment. Since novel coronaviruses are not yet available as specific drugs, the present study has visualized the use of antivirals in combination with other drugs to find multiple drug combination options, such as Favipiravir with Tocilizumab, Darunavir with Cobicistat, etc., to provide more ideas for drug combination design in relevant clinical trials.

Also, antimalarials (such as Chloroquine and Hydroxychloroquine) may be helpful for viral therapy. Cell therapy and clinical trials related to plasma in recovered individuals are also ongoing.

Considering the fact that frontline health care workers also face greater psychological pressure under long, high-intensity, demanding and high-risk working conditions, some researchers mainly used health questionnaires to understand their psychological changes and the reasons for them, in order to establish timely mind-set adjustment programs in case of emergency epidemics. In general, there are more treatment options for COVID-19 than during SARS, with a wider range of options and considerations for subjects.

Innovation of this article is that this article used data visualization to mine and analyze the registration characteristics and current progress of COVID-19 clinical trials, and finally to find potential antivirals combinations for investigators through a joint network. However, most of the trials are inconclusive, and we will be following them up continuously to examine the validity and reliability of their design in depth.

**Conclusion**

In this article, trial registration data from ChiCTR database and ClinicalTrials.gov website were collected and processed, followed by research and Visualization analysis of the 732 COVID-19 clinical trial registration features included. This article provides clinical investigators with recent trial design information and new ideas for early validation of efficacy, providing data support to guide the next phase of virus control efforts.

**Declarations**

**Ethics approval and consent to participate**
Not applicable

**Consent for publication**
Not applicable

**Availability of data and materials**
All data generated or analysed during this study are included in this published article.

**Competing interests**
All authors declare no conflicts of interest in this work.

**Funding**
Not applicable

**Authors’ contributions**
X.Q., Cai collated and visualized the data, and was a major contributor in writing the manuscript. Z.L., Zheng collected and collated the data, and adjusted the format of the paper. J.H., Huang provided professional guidance. All authors read and approved the final manuscript. Q.M., Su is the correspondent for this article.

**Acknowledgments**
We would like to thank all the health care workers on the front lines of the fight against the pandemic, and the researchers involved in clinical trials.

**References**
1. Disease outbreak news (DONs). Geneva: World Health Organization, 2020 (. opens in new tab).
2. Gao SONF, Meng-qun CHENG, WEI Xian-wen. Analysis of new coronavirus clinical study registration features: Based on ClinicalTrials.gov and ChiCTR[J/OL]. Chinese Journal of Zoonoses: 1-6 [2020-.20200224.1525.002.html.
3. Huang J, Su Q, Yang J, et al. Sample sizes in dosage investigational clinical trials: a systematic evaluation[J]. Drug Design Development & Therapy, 2015.
4. Fu-jun FENG. Research on the Technology of Data Cleaning in Big Data[C]. Science
and Engineering Research Center. Proceedings of 2018 2nd International Conference on Applied Mathematics, Modeling and Simulation (AMMS 2018). Science and Engineering Research Center: Science and Engineering Research Center, 2018: 186-191.

5. Alex John London, Jonathan Kimmelman. Clinical Trial Portfolios: A Critical Oversight in Human Research Ethics, Drug Regulation, and Policy[J]. Hastings Center Report, 2019, 49(4).

6. Schober P. Thomas R. Vetter. Sample Size and Power in Clinical Research. 2019, 129(2): 323-323.

7. Thomas RF. Surrogate Endpoints And FDA’s Accelerated Approval Process[J]. Health Aff. 2005; 24(1): 67-78.

8. Xuxu Wei M, Zhao C, Zhao X, Zhang R, Qiu Y, Lin Y, Sun M, Guan. Hongcai Shang. The global registry of COVID-19 clinical trials: indicating the design of traditional Chinese medicine clinical trials. TMR Modern Herbal Medicine, Online. .

9. Tian L, Guan X, Ma A. Systematic Comparison of Randomized Controlled Trial Study and Observational Study[J]. China Pharmacy. 2018; 29(04): 493-6.

10. Li Y, Yao L, Li J, Chen L, Song Y, Cai Z, Yang C. Stability Issues of RT-PCR Testing of SARS-CoV-2 for Hospitalized Patients Clinically Diagnosed with COVID-19. J Med Virol Accepted Author Manuscript. 2020. doi:.

11. Li Runfeng H, Yunlong, et al. Lianhuaqingwen exerts anti-viral and anti-inflammatory activity against novel coronavirus (SARS-CoV-2). Pharmacological Research. 2020: 104761.

12. Baden Lindsey R, Rubin Eric J. Covid-19 — The Search for Effective Therapy. N Engl J Med. 2020. DOI:.

13. 10.1056/NEJMo2001282
Cao, et al., (2020). A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19. NEJM, DOI:.

14. Colson P, Rolain JM, Lagier JC, Brouqui P, Raoult D. Chloroquine and hydroxychloroquine as available weapons to fight COVID-19. International journal of antimicrobial agents. 2020:105932.

15. Zhang Q, Wang Y, Qi C, Shen L, Li J. Clinical trial analysis of 2019-nCoV therapy registered in China. J Med Virol. 2020; 1–6. .

Figures

Figure 1

Statistics on single-day clinical trial registrations for COVID-19
Figure 2

Trend of clinical trial registrations for COVID-19
Figure 3

Relationship between different study types and sample size
Figure 4

Statistics of primary endpoint
Figure 5

Relationship between antivirals and other drugs