



Research

Coronavirus Disease 2019—Article

Core Outcome Set for Clinical Trials on Coronavirus Disease 2019 (COS-COVID)



Xinyao Jin^{a,b,#}, Bo Pang^{a,b,#}, Junhua Zhang^{a,b,*}, Qingquan Liu^c, Zhongqi Yang^d, Jihong Feng^e, Xuezheng Liu^f, Lei Zhang^f, Baohe Wang^e, Yuhong Huang^e, Alice Josephine Fauci^g, Yuling Ma^h, Myeong Soo Leeⁱ, Wei'an Yuan^j, Yanming Xie^k, Jianyuan Tang^l, Rui Gao^m, Liang Du^{b,n}, Shuo Zhang^e, Hanmei Qi^o, Yu Sun^p, Wenke Zheng^{a,b}, Fengwen Yang^{a,b}, Huizi Chua^{a,b}, Keyi Wang^{a,b}, Yi Ou^{a,b}, Ming Huang^a, Yan Zhu^q, Jiajie Yu^{b,n}, Jinhui Tian^r, Min Zhao^s, Jingqing Hu^t, Chen Yao^u, Youping Li^{b,n}, Boli Zhang^{a,q,*}

^a Evidence-Based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China

^b Chinese Clinical Trials Core Outcome Set Research Center, Tianjin 301617, China

^c Beijing Hospital of Traditional Chinese Medicine, Capital Medical University, Beijing 100010, China

^d The First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou 510405, China

^e The Second Affiliated Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin 300150, China

^f The First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin 300201, China

^g Istituto Superiore di Sanità, Rome 00161, Italy

^h University of Oxford, Oxford OX1 3PT, UK

ⁱ Korea Institute of Oriental Medicine, Daejeon 34054, Republic of Korea

^j Shuguang Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai 200120, China

^k Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China

^l Affiliated Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu 610075, China

^m Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing 100091, China

ⁿ The Chinese Cochrane Center, West China Hospital of Sichuan University, Chengdu 610041, China

^o The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China

^p Suqian People's Hospital of Nanjing Drum-Tower Hospital Group, Suqian 223800, China

^q Tianjin State Key Laboratory of Modern Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China

^r Lanzhou University, Lanzhou 730000, China

^s The First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou 450099, China

^t Institute of Basic Theory for Chinese Medicine, China Academy of Chinese Medicine Science, Beijing 100700, China

^u Peking University First Hospital, Beijing 100034, China

ARTICLE INFO

Article history:

Received 3 March 2020

Revised 5 March 2020

Accepted 6 March 2020

Available online 18 March 2020

Keywords:

Core outcome set

COVID-19

2019-nCoV

Coronavirus disease

Clinical trials

ABSTRACT

Since its outbreak in December 2019, a series of clinical trials on coronavirus disease 2019 (COVID-19) have been registered or carried out. However, the significant heterogeneity and less critical outcomes of such trials may be leading to a waste of research resources. This study aimed to develop a core outcome set (COS) for clinical trials on COVID-19 in order to tackle the outcome issues. The study was conducted according to the Core Outcome Measures in Effectiveness Trials (COMET) Handbook: Version 1.0, a guideline for COS development. A research group was set up that included experts in respiratory and critical medicine, traditional Chinese medicine (TCM), evidence-based medicine, clinical pharmacology, and statistics, in addition to medical journal editors. Clinical trial registry websites (www.chictr.org.cn and clinicaltrials.gov) were searched to retrieve clinical trial protocols and outcomes in order to form an outcome pool. A total of 78 clinical trial protocols on COVID-19 were included and 259 outcomes were collected. After standardization, 132 outcomes were identified within seven different categories, of which 58 were selected to develop a preliminary outcome list for further consensus. After two rounds of Delphi survey and one consensus meeting, the most important outcomes for the different clinical classifications of COVID-19 were identified and determined to constitute the COS for clinical trials on COVID-19 (COS-COVID). The COS-COVID includes one outcome for the mild type (time to 2019 novel coronavirus (2019-nCoV) reverse transcription-polymerase chain reaction (RT-PCR) negativity), four outcomes for

* Corresponding authors.

E-mail addresses: zjhtcm@foxmail.com (J. Zhang), zhangbolipr@163.com (B. Zhang).

These authors contributed equally to this work.

<https://doi.org/10.1016/j.eng.2020.03.002>

2095-8099/© 2020 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the ordinary type (length of hospital stay, composite events, score of clinical symptoms, and time to 2019-nCoV RT-PCR negativity), five outcomes for the severe type (composite events, length of hospital stay, arterial oxygen partial pressure (PaO₂)/fraction of inspired oxygen (FiO₂), duration of mechanical ventilation, and time to 2019-nCoV RT-PCR negativity), one outcome for critical type (all-cause mortality), and one outcome for rehabilitation period (pulmonary function). The COS-COVID is currently the most valuable and practical clinical outcome set for the evaluation of intervention effect, and is useful for evidence assessment and decision-making. With a deepening understanding of COVID-19 and application feedback, the COS-COVID should be continuously updated.

© 2020 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The clinical efficacy and safety evaluation of medical interventions are usually based on the measurement and analysis of certain clinical outcomes. However, studies have found that the outcomes used in clinical research are frequently inconsistent, nonstandard, irrational, or inessential, weakening the scientific and practical nature of research results and leading to research waste [1–3]. In order to overcome these problems, experts in evidence-based medicine and clinical research methodology have put forward strategies to develop core outcome sets (COSs). A COS refers to an agreed-upon standard set of outcomes that should be measured and reported, as a minimum, in all clinical trials in specific areas of health or healthcare [4]. A COS is helpful to regularize outcomes adopted in clinical trials, so as to improve the practicability, comparability, and transparency of the results [4]. In 2010, the Core Outcome Measures in Effectiveness Trials (COMET) Initiative[†] was launched to promote research on COS by developing methodological guidelines. Thus far, a series of guidelines and handbooks have been published, including the COMET Handbook: Version 1.0 [5], the COS-Standards for Development (COS-STAD) [6], the COS-Standards for Reporting (COS-STAR) [7], and the COS-Standardised Protocol Items (COS-STAP) [8].

Since the outbreak of coronavirus disease 2019 (COVID-19), hundreds of clinical trial protocols have been registered and have begun subject recruitment. By 20 February 2020, 228 protocols were already listed in two clinical trial registries^{‡,††}. However, there were some deficiencies in the registered clinical trial protocols, especially in outcomes, such as nonstandardized descriptions, significant heterogeneity, subpar clinical value, and ambiguous measure points. Hence, it is necessary to develop a COS for clinical trials on COVID-19 (COS-COVID), which is the aim of this study.

Researchers are encouraged to apply the COS-COVID for the evaluation of different interventions (either pharmaceutical or non-pharmaceutical therapies) in clinical trials on COVID-19. A full spectrum of COVID-19 classifications is covered, ranging from mild and ordinary to severe and critical types, in addition to rehabilitation period. The COS-COVID can be used not only in clinical trials, but also in systematic reviews/meta-analyses, guidelines, and other research on evidence evaluation and decision-making for COVID-19.

2. Methods

This study was conducted and reported following the COMET Handbook, COS-STAD, and COS-STAR. A research plan was published on the websites of COMET and the Chinese Clinical Trials Core Outcome Sets Research Center^{‡‡}.

[†] <http://www.comet-initiative.org>.

[‡] <http://www.chictr.org.cn>.

^{††} <https://clinicaltrials.gov>.

^{‡‡} chicos.org.cn:1080.

2.1. Participants

In order to guarantee quality and efficiency in the development of the COS-COVID, a steering group with participants from different stakeholder groups was set up. This group comprised 20 members, including scholars in Western medicine, traditional Chinese medicine (TCM), evidence-based medicine, and clinical pharmacology; statisticians; and medical journal editors. The participants were selected based on their specialty, recognition, and region. The clinical doctors within this group are experts in the field of respiratory and critical medicine, and have experience in the clinical treatment of patients with COVID-19. Experts on behalf of different interest groups participated in the whole research process. A coordination group, which was responsible for research process coordination and data analysis, was also established.

2.2. Information sources

Two clinical trial registries were comprehensively searched to retrieve the outcomes used in clinical trials from 1 December 2019 to 12 February 2020. Randomized controlled trials, non-randomized controlled trials, case series, and cohort studies aimed at evaluating different interventions for COVID-19 were included. Studies that included suspected cases, diagnostic tests, and syndrome surveys were excluded. Clinical trials from the registries were screened by two reviewers according to the inclusion and exclusion criteria. A predesigned Excel spreadsheet was used to extract data, including design type, intervention, patient, outcome, and so forth. Information on outcomes, which was extracted by two authors independently, included the outcome name, measurement method, measurement time point, and data type. Disagreements were resolved by discussion.

The extracted outcome information was sorted by similarity. Duplicated outcomes were excluded, nonstandard outcomes were standardized, and synonym outcomes were merged. The process was carried out by two researchers independently, and the differences were resolved by discussion.

The outcomes obtained after data cleaning were assigned to seven different categories: clinical symptoms, physical and chemical detection, viral nucleic acid detection, quality of life, significant events, disease process, and safety indexes. In order to generate a preliminary list of outcomes for consensus, all the outcomes in each region were voted on for inclusion or not. An outcome would be removed from the preliminary list when 75% of the voting members voted it to be unnecessary. The remaining outcomes formed the preliminary list of outcomes. According to the different classifications of COVID-19, outcomes in the preliminary list were divided into five types: mild, ordinary, severe, critical, and rehabilitation period.

2.3. Consensus process

In this study, two rounds of Delphi survey were conducted for consensus. After each round of survey feedback, a meeting of

expert was held to discuss and determine whether to add or remove outcomes.

2.3.1. Identifying stakeholder groups

In order to ensure the efficiency and quality of the consensus process, this study invited representatives of respiratory, critical, TCM, and evidence-based medicine, in addition to medical management and journal editors, to join the Delphi survey. In consideration of geographical balance, experts were invited from different regions in China, including Wuhan (Hubei Province), Tianjin, Beijing, Jiangsu, Guangdong, Shanghai, Henan, and Sichuan, and internationally, from Italy, the Republic of Korea, the United Kingdom, and the United States. All of the participants provided informed consent and were willing to participate in the survey.

2.3.2. Questionnaire process

An electronic questionnaire by cell phone was used for the Delphi survey. The questionnaire had two main sections: ① scoring each indexes; and ② recommending the indexes to be added. After each round of information feedback, an expert meeting was held by phone to discuss and determine whether to add, delete, or merge outcomes. Due to the urgency of requirement, participants were required to provide feedback within 24 h.

2.3.3. Outcome scoring

A Likert scale with nine scores was used to evaluate the importance of the outcomes. Each outcome was scored from a scale of 1 to 9 (unimportant: 1–3; important but not essential: 4–6; essential: 7–9). At the end of each round, data analyses were carried out immediately. Based on the importance ranking, an outcome that was scored higher than 7 by more than 75% of the experts was retained for the next consensus process. The outcomes recommended by the experts might enter the second round after discussion by the steering group.

2.3.4. Consensus meeting

All outstanding representatives of different stakeholder groups, clinical experts who completed the Delphi survey, and members of the steering group were invited to the consensus meeting. If an outcome was ranked as essential (7–9) by at least 75% of the participants, it was considered to be agreed upon by consensus and was recommended into the final COS [5].

Due to the special circumstance of the disease epidemic, the consensus meeting was held by telephone conference instead of through a face-to-face meeting. The contents of the consensus meeting covered five aspects: ① reporting the research methods; ② reporting the results of two rounds of Delphi survey; ③ putting forward the key points to be discussed; ④ fully discussing the candidate core outcomes; and ⑤ voting on the outcomes and reaching a consensus to form the COS-COVID through discussion.

3. Results

We obtained 107 registered protocols of clinical studies on COVID-19: 84 from www.chictr.org.cn and 23 from clinicaltrials.gov. After screening, 78 protocols met the eligibility criteria: 52 with interventions of chemical or biological drugs and 26 with TCM plus standard treatments.

3.1. Outcomes pool

There were 259 outcomes (used 596 times) reported in the included clinical protocols. After the standardized process, 132 outcomes were obtained and assigned to seven domains. Details of the outcome pool are provided in [Table 1](#).

This list of outcomes was too long to be used for a Delphi survey. In order to improve the efficiency and quality of the Delphi survey, five experts from the steering group voted on and discussed the outcomes to be retained or eliminated. A preliminary list of outcomes for the first Delphi survey was formed that comprised 58 outcomes assigned to five types of COVID-19 (mild, ordinary, severe, critical, and rehabilitation period) [9]: 17 outcomes for mild, 33 outcomes for ordinary, 35 outcomes for severe, 22 outcomes for critical, and 6 outcomes for rehabilitation period. Details are provided in [Table 2](#).

3.2. Delphi survey

Sixty participants were invited to vote in the first round of Delphi survey and 52 responses were eventually received, for an attrition rate of 13.3%. According to the consensus standards, 10, 25, 34, 22, and 5 outcomes, respectively, were voted as essential for the types of mild, ordinary, severe, critical, and rehabilitation period. Outcomes, including body mass index (BMI), complete blood count (CBC), arterial blood gas, diarrhea, B-type natriuretic peptide (BNP), myocardial infarction index, duration of intensive care unit (ICU) admission, and immunological index for different types were recommended by participants. Based on the voting results and feedback, the steering group held a meeting to discuss which outcomes were of importance and should be included in the second round of Delphi survey. Different expressions of the same index were combined to improve the concentration of the outcomes. After discussion, 5 outcomes for mild, 15 outcomes for ordinary, 20 outcomes for severe, 15 outcomes for critical, and 5 outcomes for rehabilitation period obtained consensus for the second round of Delphi survey; none of the recommended outcomes from the first round of Delphi survey were included ([Table 3](#)).

Twenty-two experts, with an emphasis on clinicians in the front line of clinical treatment, were invited to join the second round of Delphi survey. With 20 of these experts responding to the questionnaire within 24 h, the attrition rate was 9.1%. Certain additional outcomes—chest computerized tomography (CT) test, respiratory rate, blood gas analysis, acute physiology and chronic health evaluation (APACHE) II score, lactic acid, and psychological test—were suggested to supplement the agreed-upon outcomes. Based on the results of the second round, a teleconference was held by the steering group to discuss and confirm the candidate outcomes for the final consensus meeting. After discussion, the steering group agreed that the APACHE II score would be added to the severe stage; the CURB-65 (stand for confusion, uremia, respiratory rate, blood pressure, age ≥ 65 years) score and duration of extracorporeal membrane oxygenation (ECMO) would be removed from the critical type; and the incidence of sequelae and the rate of interstitial pneumonia would be combined into the incidence of sequelae. Finally, the outcomes voted as essential for consensus included 4 for mild, 8 for ordinary, 16 for severe, 12 for critical, and 4 for rehabilitation period ([Table 4](#)).

3.3. Consensus meeting

The consensus meeting was held on 24 February 2020, and involved 20 participants. These included representatives from various stakeholder groups who were experts in respiratory, critical, TCM, and evidence-based medicine; clinical pharmacology; and statistics, in addition to medical journal editors and decision-makers. There was no conflict of interest among the different stakeholders. Before discussion on each outcome, the results of the two rounds of Delphi survey were shown to the participants of the meeting. Based on the Delphi survey results and discussion, the participants voted anonymously for each outcome, following the criteria of clinical value, clinical feasibility, and stability of

Table 1
Outcomes adopted in the protocols of clinical trials on COVID-19.

| Category (amount) | Outcomes |
|--------------------------------------|---|
| Clinical symptoms (25) | Time to defervescence; remission rate of respiratory symptoms; time to cough reported as mild; time to cough reported as absent; Leicester cough questionnaire; percentage of cases without coughing; different approaches of oxygen inhalation; time to dyspnea reported as mild/absent; frequency of dyspnea; respiratory rate; frequency of hypoxia; duration (d) of supplemental oxygenation; recovery time of pulmonary function; respiratory symptoms and signs; frequency of respiratory progression; frequency of respiratory remission; time to gastrointestinal symptoms reported as absent; frequency of individual symptom absent; time to clinical remission; frequency of clinical improvement; clinical symptoms; score of clinical symptoms; TTCI; curative effects of TCM syndromes; score of TCM symptoms |
| Physical and chemical detection (28) | X-ray examination; changes in pulmonary imaging; CT and MRI scan of hip joint; chest imaging; pulmonary function; cardiac function; CBC; peripheral blood cell count; C-reactive protein; hypersensitive C-reactive protein; erythrocyte sedimentation rate; calcitonin; procalcitonin; proinflammatory cytokines; blood gas analysis; blood oxygen saturation; PaO ₂ /FiO ₂ ; improvement rate of finger oxygen; recovery rate of laboratory indexes; lymphocyte count; change curve of lymphocyte subsets count; immunological indexes; CD4 ⁺ and CD8 ⁺ T cells count; D-dimer; coagulation function; myocardial enzymes; myoglobin; creatinine kinase |
| Viral nucleic acid detection (3) | 2019-nCoV RT-PCR test; virologic indicators; virus antibody level in blood samples |
| Quality of life (9) | Modified Barthel index; health survey short form; self-rating anxiety scale; European QoL-5 dimensional questionnaire; assessment of daily living ability; social support rating scale; SF-36; psychological indexes; self-rating depression scale |
| Significant events (11) | Fatality rate; all-cause fatality rate; all-cause mortality; survival condition; mortality rate; time to treatment failure; frequency of MODS; frequency of ARDS; sequential organ failure assessment; organ support intensity; frequency of shock |
| Disease process (44) | Percentage of cases with significant improvement or meeting discharge standard; time to recovery; de-isolation rate; frequency of 2019-nCoV RT-PCR negativity; time to cure; cure rate; frequency of disease remission; frequency of disease progression; assessment of disease evolution; frequency of requirement for first aid; combination with other infections; percentage of cases administered with steroids; percentage of antibiotics use; DIC score; vasopressor days; percentage of cases turning to critical; frequency of hemodialysis filtration; frequency of requirement for pulmonary surfactant; percentage of cases turning to severe; time to cases turning to severe; percentage of cases recovering to mild type; time to cases recovering to mild type; frequency/length of severe cases turning to critical type or death; frequency of ICU admission; length of ICU stay (d); time to ICU admission; length of hospital stay (d); frequency of tube insertion; duration (d) of tube insertion; parameters of respirator; frequency of requirement for mechanical ventilation; duration (d) of mechanical ventilation; parameters of mechanical ventilation; duration (d) of ECMO; mode and parameters of ECMO; seven-point ordinal scale; APACHE II score; CURB-65 pneumonia severity score; Murray lung injury score; NEWS2 score; SOFA score; pulmonary high-resolution CT score; PSI score; St. George's respiratory questionnaire |
| Safety indexes (12) | Adverse reactions; frequency of adverse events; frequency of severe adverse events; frequency of drug withdrawal due to adverse reactions; frequency of complications; blood biochemistry; liver function; liver and renal function; frequency of renal injury; routine urinalysis; electrocardiography; blood concentration of chloroquine |

TTCI: time to clinical improvement; CT: computerized tomography; MRI: magnetic resonance imaging; PaO₂/FiO₂: the ratio of arterial oxygen partial pressure to fraction of inspired oxygen; CBC: complete blood count; 2019-nCoV: 2019 novel coronavirus; RT-PCR: reverse transcription-polymerase chain reaction; QoL: quality of life; SF-36: the medical outcome study 36-item short-form health survey; MODS: multiple organ dysfunction syndrome; ARDS: acute respiratory distress syndrome; DIC: disseminated intravascular coagulation; ICU: intensive care unit; ECMO: extracorporeal membrane oxygenation; APACHE: acute physiology and chronic health evaluation; CURB-65: confusion, uremia, respiratory rate, blood pressure, age ≥ 65 years; NEWS: national early warning score; SOFA: sequential organ failure assessment; PSI: pneumonia severity index.

Table 2
Preliminary list of outcomes for the first round of Delphi survey.

| Type (amount) | Outcomes |
|---------------------------|--|
| Mild (17) | Time to defervescence; time to cough disappearance; time to dyspnea reported as absent; time to gastrointestinal symptoms reported as absent; TTCI; chest X-ray examination; chest CT scan; CBC; C-reactive protein; calcitonin; proinflammatory cytokines; lymphocyte count; D-dimer; time to 2019-nCoV RT-PCR negativity; frequency of hospital discharge; length of hospital stay (d); score of TCM symptoms |
| Ordinary (33) | Time to defervescence; time to cough reported as absent; time to cough reported as mild; time to dyspnea reported as absent; remission rate of respiratory symptoms; time to gastrointestinal symptoms reported as absent; TTCI; chest X-ray examination; chest CT scan; myocardial enzymes; CBC; C-reactive protein; erythrocyte sedimentation rate; calcitonin; proinflammatory cytokines; blood gas analysis; blood oxygen saturation; PaO ₂ /FiO ₂ ; lymphocyte count; CD4 ⁺ and CD8 ⁺ T cells count; D-dimer; coagulation function; all-cause mortality; frequency of MODS; frequency of hospital discharge; time to 2019-nCoV RT-PCR negativity; time to hospital discharge; percentage of cases turning to severe or critical; frequency of ICU admission; duration (d)/ frequency of mechanical ventilation; duration (d) of oxygen inhalation; curative effects of TCM syndromes; score of TCM symptoms |
| Severe (35) | Time to defervescence; respiratory rate; duration (d) of supplemental oxygenation; frequency of clinical improvement (turning to ordinary or recovery); TTCI; chest CT scan; CBC; C-reactive protein; erythrocyte sedimentation rate; calcitonin; proinflammatory cytokines; blood gas analysis; blood oxygen saturation; PaO ₂ /FiO ₂ ; lymphocyte count; CD4 ⁺ and CD8 ⁺ T cells count; D-dimer; coagulation function; myocardial enzymes; time to 2019-nCoV RT-PCR negativity; all-cause mortality; frequency of MODS; frequency of shock; time to hospital discharge; DIC score; percentage of cases turning to critical; frequency of ICU admission; length of hospital stay (d); frequency of mechanical ventilation; duration (d) of mechanical ventilation; duration (d) of ECMO; NEWS2 score; CURB-65 pneumonia severity score; PSI score; curative effects of TCM syndromes |
| Critical (22) | Time to defervescence; frequency of clinical improvement; TTCI; proinflammatory cytokines; blood gas analysis; blood oxygen saturation; PaO ₂ /FiO ₂ ; lymphocyte count; CD4 ⁺ and CD8 ⁺ T cells count; coagulation function; 2019-nCoV RT-PCR test; frequency of MODS; frequency of shock; all-cause mortality; length of ICU stay (d); length of hospital stay (d); frequency of clinical improvement (turning to ordinary or recovery); duration (d) of mechanical ventilation; duration (d) of ECMO; APACHE II score; CURB-65 pneumonia severity score; PSI score |
| Rehabilitation period (6) | Chest X-ray examination; chest CT scan; pulmonary function; SF-36; frequency of interstitial pneumonia; frequency of other sequelae |

indicators in different classifications of COVID-19. Each outcome that met the consensus standards was included in the final COS. At the end, the COS-COVID consisted of one outcome for mild (time to 2019 novel coronavirus (2019-nCoV) reverse transcription-

polymerase chain reaction (RT-PCR) negativity), four outcomes for ordinary (length of hospital stay, composite events, score of clinical symptoms, and time to 2019-nCoV RT-PCR negativity), five outcomes for severe (composite events, length of hospital stay,

Table 3

List of outcomes for the second round of Delphi survey.

| Type (amount) | Outcomes |
|---------------------------|--|
| Mild (5) | Defervescence (rate/time); clinical symptoms remission (rate/time/score); lymphocyte (count/ratio); 2019-nCoV RT-PCR negativity (rate/time); hospital discharge (rate/time) |
| Ordinary (15) | Frequency of composite events (ICU admission, MODS, shock, death, deterioration to severe or critical type); all-cause mortality; time to hospital discharge (meet discharge standards); 2019-nCoV RT-PCR negativity (rate/time); chest CT scan; length of hospital stay (d); mechanical ventilation (rate/rate/time); blood oxygen saturation; PaO ₂ /FiO ₂ ; C-reactive protein; proinflammatory cytokines (TNF- α , IL-1 β , IL-6, IL-8); lymphocyte (count/ratio); immunological indexes; clinical symptoms remission (rate/time/score); defervescence (rate/time) |
| Severe (20) | All-cause mortality; hospital discharge (rate/time, meet discharge standards); frequency of composite events (ICU admission, MODS, shock, death, critical deterioration) composite; 2019-nCoV RT-PCR negativity (rate/time); lymphocyte (count/ratio); immunological indexes; proinflammatory cytokines (TNF- α , IL-1 β , IL-6, IL-8); mechanical ventilation (rate/rate/time); blood oxygen saturation; PaO ₂ /FiO ₂ ; NEWS2 score; CURB-65 pneumonia severity score; PSI score; SOFA score; DIC score; chest CT scan; clinical symptoms remission (rate/time/score); defervescence (rate/time); respiratory rate; C-reactive protein |
| Critical (15) | All-cause mortality; frequency of shock; length of ICU stay (d); time to hospital discharge (meet discharge standards); blood oxygen saturation; PaO ₂ /FiO ₂ ; duration (d) of mechanical ventilation; APACHE II score; CURB-65 pneumonia severity score; PSI score; SOFA score; lymphocyte (count/ratio); 2019-nCoV RT-PCR negativity (rate/time); defervescence (rate/time); duration (d) of ECMO |
| Rehabilitation period (5) | Chest CT scan; pulmonary function; SF-36; frequency of interstitial pneumonia; frequency of other sequelae |

TNF: tumor necrosis factor; IL: interleukin.

Table 4

List of outcomes for the consensus meeting.

| Type (amount) | Outcomes |
|---------------------------|---|
| Mild (4) | Defervescence (rate/time); clinical symptoms remission (rate/time/score); lymphocyte (count/ratio); 2019-nCoV RT-PCR negativity (rate/time) |
| Ordinary (8) | Composite events (ICU admission, MODS, shock, death, deterioration to severe or critical type); length of hospital stay (d); 2019-nCoV RT-PCR negativity (rate/time); chest CT scan; blood oxygen saturation; PaO ₂ /FiO ₂ ; clinical symptoms remission (rate/time/score); defervescence (rate/time) |
| Severe (16) | All-cause mortality; time to hospital discharge; composite events (ICU admission, MODS, shock, death, critical deterioration); 2019-nCoV RT-PCR negativity (rate/time); lymphocyte (count/ratio); immunological indexes; mechanical ventilation (rate/rate/time); blood oxygen saturation; PaO ₂ /FiO ₂ ; PSI score; SOFA score; chest CT scan; clinical symptoms remission (rate/time/score); defervescence (rate/time); respiratory rate; APACHE II score |
| Critical (12) | All-cause mortality; frequency of shock; length of ICU stay (d); length of hospital stay (d); blood oxygen saturation; PaO ₂ /FiO ₂ ; duration (d) of mechanical ventilation; APACHE II score; PSI score; SOFA score; 2019-nCoV RT-PCR negativity (rate/time); defervescence (rate/time) |
| Rehabilitation period (4) | Chest CT scan; pulmonary function; SF-36; incidence of sequelae |

arterial oxygen partial pressure (PaO₂)/fraction of inspired oxygen (FiO₂), duration of mechanical ventilation, and time to 2019-nCoV RT-PCR negativity), one outcome for critical (all-cause mortality), and one outcome for rehabilitation period (pulmonary function), as shown in Table 5.

4. Perspectives

This was a fast COS study conducted under special requirements and in a special environment. Nevertheless, the study was rigorously conducted and reported according to the COS-STAD and COS-STAR. The COS-COVID was accomplished on time and with clinical significance. We hope that all clinical trials and research on evidence transformation for COVID-19 can refer to the COS-COVID during protocol design and decision-making.

Three points must be illustrated for the rational application of the COS-COVID. First, although the COS is the minimum, it is not the only index that should be reported in every clinical study. Studies with different purposes can add other outcomes if necessary.

Table 5

The COS for clinical trials on COVID-19.

| Type (amount) | Outcomes |
|---------------------------|---|
| Mild ^c (1) | Time to 2019-nCoV RT-PCR negativity ^a (d) |
| Ordinary ^c (4) | Length of hospital stay ^b (d); composite events (total number of patients diagnosed as the types of severe, critical, and all-cause death); score of clinical symptoms ^d ; time to 2019-nCoV RT-PCR negativity |
| Severe ^c (5) | Composite events (total number of patients diagnosed as type critical and all-cause death); length of hospital stay (d); PaO ₂ /FiO ₂ ; duration (d) of mechanical ventilation; time to 2019-nCoV RT-PCR negativity |
| Critical ^c (1) | All-cause mortality |
| Rehabilitation (1) | Pulmonary function |

^a Negativity: two consecutive negative results (sampling interval of at least 24 h) of the 2019-nCoV nucleic acids tests of respiratory pathogens.

^b Discharge standards: ① normal body temperature for more than three days; ② significant recovered respiratory symptoms; ③ lung imaging showing obvious absorption and recovery of acute exudative lesion; ④ negativity of nucleic acids tests performed twice.

^c Mild type: the clinical symptoms are mild and no pneumonia manifestation can be found in imaging. Ordinary type: patients have symptoms like fever, respiratory tract symptoms, and pneumonia manifestation can be seen in imaging. Severe type (meeting any of the following): ① respiratory rate ≥ 30 times·min⁻¹; ② oxygen saturation < 93% at a rest state; ③ PaO₂/FiO₂ ≤ 300 mmHg (1 mmHg = 0.133 kPa); ④ patients with > 50% lesions progression within 24–48 h in pulmonary imaging. Critical type (meeting any of the following): ① respiratory failure occurred and mechanical ventilation required; ② shock occurred; ③ complicated with other organ failure, ICU treatment required.

^d Score of clinical symptoms: a total score of six common and important clinical symptoms, including fever, cough, fatigue, shortness of breath, diarrhea, and body pain, each of which can be scored as 0 (no), 1 (mild), 2 (moderate), or 3 (significant).

Second, the COS is not equivalent to the primary outcomes. According to the main purposes of different studies, one or more outcomes in a COS can be selected as the primary outcomes. Third, there is no restriction on the treatment course and measure point in a given trial. However, they should be well-defined and based on scientific and feasible principles. For COVID-19, more than two weeks of treatment course are suggested. Safety outcomes were not included in the COS-COVID, because different drugs might have different adverse reactions. In addition, we suggest that researchers report all adverse events encountered during clinical trials.

There are several limitations in this study. First, the development of the outcome pool was only based on clinical trial protocols listed in two registry platforms. Doctors and patients were not consulted to collect indicators. Therefore, there is a potential risk of missing important outcomes. Second, due to the prevalence of new infectious diseases, patients were not invited to join in the

survey and consensus process. As a result, patients' opinions may not have been fully reflected. Third, the number of representatives for different stakeholders may not be fully adequate. The fact that the majority of the experts were from China weakens the regional representation. Fourth, the process of consensus was conducted via conference calls instead of through face-to-face meetings, which may have led to insufficient discussion and affected the consensus results. Finally, the current understanding of COVID-19 is still incomplete and in the process of being established, so the relevant evaluation outcomes and COS must be updated with the process of practice. Furthermore, we wish to strengthen communication with relevant international academic organizations to promote the application and update of the COS-COVID.

Acknowledgements

We are deeply grateful to the front-line clinicians who participated in the questionnaire while directly fighting the epidemic. This work was supported by the National Science and Technology Emergency Project (2020yfc0841600).

Compliance with ethics guidelines

Xinyao Jin, Bo Pang, Junhua Zhang, Qingquan Liu, Zhongqi Yang, Jihong Feng, Xuezheng Liu, Lei Zhang, Baohe Wang, Yuhong Huang, Alice Josephine Fauci, Yuling Ma, Myeong Soo Lee, Wei'an Yuan, Yanming Xie, Jianyuan Tang, Rui Gao, Liang Du, Shuo Zhang, Hanmei Qi, Yu Sun, Wenke Zheng, Fengwen Yang, Huizi Chua, Keyi Wang, Yi Ou, Ming Huang, Yan Zhu, Jiajie Yu, Jinhui Tian, Min Zhao,

Jingqing Hu, Chen Yao, Youping Li, and Boli Zhang declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Sinha IP, Gallagher R, Williamson PR, Smyth RL. Development of a core outcome set for clinical trials in childhood asthma: a survey of clinicians, parents, and young people. *Trials* 2012;13(1):103.
- [2] Williamson PR, Altman DG, Blazeby JM, Clarke M, Gargon E. Driving up the quality and relevance of research through the use of agreed core outcomes. *J Health Serv Res Policy* 2012;17(1):1–2.
- [3] Xing DM, Zhang JH, Zhang BL. Approach to developing core outcome sets for clinical trials of traditional Chinese medicine. *Chin J Tradit Chin Med Pharm* 2014;29(5):1352–5.
- [4] Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13(1):132.
- [5] Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET handbook: version 1.0. *Trials* 2017;18(Suppl 3):280.
- [6] Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core outcome set-standards for development: the COS-STAD recommendations. *PLoS Med* 2017;14(11):e1002447.
- [7] Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Devane D, et al. Core outcome set-standards for reporting: the COS-STAR statement. *PLoS Med* 2016;13(10):e1002148.
- [8] Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core outcome set-standardised protocol items: the COS-STAP statement. *Trials* 2019;20(1):116.
- [9] National Health Commission of the People's Republic of China, National Administration of Traditional Chinese Medicine. Diagnosis and treatment protocol for novel coronavirus pneumonia (trial version 5). Beijing: The State Council of the People's Republic of China; 2020 Feb 18 [cited 2020 Mar 1]. Available from: <http://www.nhc.gov.cn/yzygj/s7653p/202002/8334a8326dd94d329df351d7da8aefc2/files/b218cfeb1bc54639af227f922bf6b817.pdf>. Chinese.