Randomized Controlled Trials for COVID-19: Evaluation of Optimal Randomization Methodologies - Need for the Data Validation of the Completed Trials, and to Improve the Ongoing and Future Randomized Trial Designs

Venkata R. Emani MD. FACC, Sanjeev Goswami MD. FCCP, Dheeraj Nandanoor MD, Shaila R. Emani BS, Nidhi K. Reddy BA, Raghunath Reddy MD

PII: S0924-8579(20)30433-7
DOI: https://doi.org/10.1016/j.ijantimicag.2020.106222
Reference: ANTAGE 106222

To appear in: International Journal of Antimicrobial Agents


This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 The Author(s). Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Highlights

- The top health experts and clinicians are relying on the results of the recently released randomized controlled trials for COVID-19 treatment.

- The Recovery trial showed mortality benefit with Dexamethasone, and the ACTT-1 trial showed that the treatment with Remdesivir reduced the Time to recovery without mortality benefit.

- The randomization methodologies of these randomized control trials were suboptimal for matching the studied groups based on disease severity, and the published literature is very limited about the disease severity metrics.

- The authors have failed to show that the data of these trials is without fatal sampling errors and sampling biases.

- There is a definite need for the validation of data in the completed trials along with the improvement of randomization methodologies for the design of future randomized controlled trials for COVID-19.

- Only double-blind placebo controlled randomized trials when done with robust randomization methodologies can yield high quality data that can give clinical guidance for treating COVID-19 patients.
Randomized Controlled Trials for COVID-19: Evaluation of Optimal Randomization Methodologies - Need for the Data Validation of the Completed Trials, and to Improve the Ongoing and Future Randomized Trial Designs.

Running title: COVID-19 Randomized Controlled Trial Methodologies - Data Validation of Current Trials and Improvement of Future Trial Designs.

Correspondence to: Venkata R. Emani, MD, FACC; 1148 Norman Drive, Suite #2, Manteca, CA 95336 vemani@yahoo.com

Authorship:
This manuscript was initially drafted by the first author, and approved by all authors.
Venkata R. Emani, MD. FACC
Sanjeev Goswami, MD. FCCP
Dheeraj Nandanoor, MD
Shaila R. Emani, BS
Nidhi K. Reddy, BA
Raghunath Reddy, MD

AUTHORS INSTITUTIONS:
Venkata R. Emani, MD. FACC
Institution: Central Valley Cardiovascular Associates, Inc
Sanjeev Goswami, MD, FCCP
Institution: San Joaquin Critical Care Medical Group

Dheeraj Nandanoor, MD
Institution: Synergy Med

Shaila R. Emani, BS
Institution: Central Valley Cardiovascular Associates, Inc

Nidhi K Reddy, BA
Institution: Stockton Primary Care

Raghunath Reddy, MD
Institution: Stockton Primary Care

**Corresponding author:**
Venkata R. Emani, M.D
Central Valley Cardiovascular Associates, Inc
1148 Norman Drive, Suite#2
Manteca, CA 95336
E-mail: vemani@yahoo.com

**Institutions address:**
Central Valley Cardiovascular Associates, Inc
1148 Norman Drive, Suite#2
Manteca, CA 95336
Phone: 209-239-7852

San Joaquin Critical Care Medical Group
1801 E March Ln c300,
Stockton, CA 95210

Synergy Med
1801 E March Ln STE D460
Stockton, CA 95210

Stockton Primary Care
1805 N California St #102
Stockton, CA 95204

Author’s e-mails address:

Venkata R. Emani, MD. FACC
e-mail: vemani@yahoo.com

Sanjeev Goswami, MD, FCCP
e-mail: sanjeevkgoswami@gmail.com

Dheeraj Nandanoor, MD
If you need any additional information, please feel free to contact me.

Thank you,

Sincerely,

Venkata R. Emani, MD.

E-mail: drdj25@gmail.com
Shaila R. Emani, BS
E-mail: shaila88@ucla.edu

Nidhi K Reddy, BA
E-mail: nidhikreddy@gmail.com

Raghunath Reddy, MD
E-mail: nininireddy@yahoo.com

ABSTRACT:
During this emerging COVID-19 pandemic, initially there were no proven treatment options. With the release of randomized controlled trial results, we are beginning to see possible treatment options for
COVID-19. The Recovery trials showed an Absolute Risk Reduction (ARR) in mortality by 2.8% with Dexamethasone, and the ACTT-1 trial showed that the treatment with Remdesivir reduced the Time to recovery by 4 days. The Hydroxychloroquine and Lopinavir/Ritonavir treatments did not show any mortality benefit in both the Recovery and WHO Solidarity trials. The NIH Hydroxychloroquine and Brazilian Hydroxychloroquine trials did not show any benefit for Hydroxychloroquine based on the 7-point ordinal scale outcomes.

The randomization methodologies utilized in these controlled trials and the quality of published data were reviewed to examine their adaptability to treat patients. We found that the randomization methodologies of these trials were suboptimal for matching the studied groups based on disease severity among critically ill hospitalized COVID-19 patients with high mortality rates. The published literature is very limited about the disease severity metrics among the compared groups, and failed to show that the data is without fatal sampling errors and sampling biases. We also found that there is a definite need for the validation of data in these trials along with additional important disease severity metrics to ensure that the trials’ conclusions were accurate.

We also propose proper randomization methodologies for the design of randomized controlled trials for COVID-19, and guidance for the publication of COVID-19 trial results.

**Keywords**: COVID-19 Randomized Controlled Trial Methodologies

1. **Introduction**

Coronavirus disease 2019 (COVID-19) is caused by SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2). Because of the sudden pandemic outbreak, there are no initial recommended treatments for
COVID-19 patients. Clinicians are looking for new or existing options that can help guide their patients’ treatments. In the absence of detailed literature, top health experts and clinicians who are treating the patients are relying on the results from the recently released randomized controlled trials for COVID-19 treatment.

A properly done randomized trial is always superior and provides the highest quality data. Any large, well-designed randomized control trials should evenly distribute known and unknown factors among the intervention and control groups in order to minimize the potential for bias. However, a large proportion of negative trials is a problem in critical care settings, which is largely due to heterogeneous patient populations, variable presentations of disease and levels of response to treatments among these patients. Preserving the integrity of clinical trials during the coronavirus pandemic is crucial, and these trials have to be critically examined to ensure that they are yielding valid data. The NIH/NIAID’s ACTT-1 and ORCHID trials are the only double-blind placebo controlled randomized trials, and all of the other trials for COVID-19 are open-label randomized controlled trials. In this manuscript, the following randomized controlled studies were reviewed to analyze the clinical impact of the treatment options.

1) The ACTT-1 trial showed no statistically significant mortality benefit for Remdesivir, but reduced the Time to recovery by 4 days.

2) The Recovery trial showed an Absolute Risk Reduction (ARR) in mortality by 2.8% with Dexamethasone, and this beneficial effect was mainly seen in patients who required invasive mechanical ventilation. This trial did not show any benefit for treatment with either Hydroxychloroquine or Lopinavir/Ritonavir.

3) The halted ORCHID trial did not show any benefit for Hydroxychloroquine based on its 7-point ordinal scale outcomes.
4) The Brazilian Hydroxychloroquine trial also did not show any benefit for Hydroxychloroquine based on its 7-point ordinal scale outcomes.10

5) The WHO Solidarity trial showed no mortality benefit using Hydroxychloroquine and Lopinavir/Ritonavir treatments.11

These trials were launched rapidly in the middle of the pandemic, and the published findings of these trials lack important details about the disease severity among these critically ill hospitalized patients. In this review, we examined if these randomized controlled trial designs are living up to the expectations of yielding high quality data that can fully guide patient treatments.

2. Randomization strategies for critically ill hospitalized patients and the limitations of current trials’ randomization methodologies for COVID-19:

COVID-19 is a multi-systemic disease with high mortality rates in critically ill hospitalized patients4–7. In critically ill hospitalized patients with suspected infection, the sequential organ failure assessment (SOFA) and national early warning (NEW) score were demonstrated to be superior for the prognostication of mortality12–20. The protocol of the ACTT-1 trial adopted a daily NEW score assessment4,21 and a recent study showed the important prognostic value of the SOFA score in predicting bad outcomes in COVID-19 patients22. A quick COVID-19 Severity Index (qCSI) scoring system23 accurately predicted patients’ progress to respiratory failure within 24 hours of admission using bedside respiratory examination findings which employed similar respiratory parameters as the SOFA and NEW scores. In addition to ground-glass opacities when present in the CT chest, other laboratory findings including lymphopenia particularly in young healthy adults, thrombocytopenia, hypoalbuminemia, elevated levels of D-Dimer, C-reactive protein, erythrocytic sedimentation rate (ESR), interleukin-6, procalcitonin, lactate dehydrogenase, neutrophils
count are considered to be some other important indicators for COVID-19 disease severity and worse prognosis.\textsuperscript{22,24–32}

All of the completed randomized controlled trials only randomized patients based on the respiratory support received at randomization.\textsuperscript{4–6,10,21,33–35} The examples in Table 1 illustrate the flaws in the randomization methodologies of these trials when we specifically reviewed the impact of COVID-19 disease severity. For example, a group of patients with the same baseline oxygen requirement and comorbidities can have an expected 5\% to 27.6\% mortality risk based on their disease severity if the SOFA score or NEW score is applied.\textsuperscript{12,15} Table 2 illustrates that all three of the groups that were randomized based on respiratory support, particularly the majority of patients randomized to the Oxygen group in these trials, will be at an even greater risk for the disease severity heterogeneity and unpredictability of their outcomes, creating the conditions for a very high chance of fatal sampling errors unless the disease severity metrics are assessed as part of the randomization.\textsuperscript{4–7,10} Most importantly, the majority of deaths occurred among the patients who were randomized to the Oxygen group with 61.6\% of total deaths in the Recovery-Dexamethasone trial and 56.8\% of total deaths in the ACTT-1 trial.\textsuperscript{4,5} Since the mortality rate was as high as 26.6\% among all subjects in these trials, the SOFA score or NEW score mismatch (5.0\% to 27.6\% variability in mortality) among the groups in these randomized trials can itself create fatal sampling errors (Type I or Type II). For example, the Recovery trials were only powered for an absolute difference of 4 percentage points between the two arms,\textsuperscript{5,6,33} and any sampling errors that would have caused 11 less deaths than what was observed in the Dexamethasone arm or 24 more deaths than what was observed in the usual care arm may result in the loss of statistical significance for Dexamethasone in the Recovery trial. These variances of 11 or 24 deaths among the two compared arms only account for 0.69\% - 1.51\% of the total deaths (n = 1592) in the Recovery-Dexamethasone trial, which signifies the uncertainties in the conclusions reached that would have caused a high probability of Type I and/or Type II errors (a copy of the statistical analysis is enclosed).
Similarly, the same Type I or Type II fatal sampling errors can occur in all other randomized trials since the randomization methodologies were also similar.

Based on the data from Tables 1-3, we strongly believe that there is a chance for the occurrence of more than a 4% sampling error(s) in these randomized controlled trials, as these randomized controlled trials did not randomize patients based on COVID-19 disease severity for varying levels of hypoxemia (PaO2/FiO2 ratio) including the standardized metrics for disease severity in these critically ill hospitalized patients and cannot yield high quality data.

2.1 Adaptive COVID-19 Treatment Trial (ACTT-1)

This was a double-blind placebo controlled randomized trial which evaluated the safety and efficacy of Remdesivir in the hospitalized patients with a primary endpoint of Time to recovery. This trial allocated patients to four groups (Group 4: not requiring supplemental oxygen, Group 5: requiring supplemental oxygen, Group 6: receiving noninvasive ventilation or high-flow oxygen devices, and Group 7: receiving invasive mechanical ventilation or extracorporeal membrane oxygenation [ECMO]) based on respiratory support received at randomization.

2.2 Randomized Evaluation of COVID-19 Therapy (Recovery) trial

This was an open-labeled 2:1 randomized trial to evaluate the efficacy of low-dose corticosteroids (Dexamethasone), Hydroxychloroquine and Lopinavir-Ritonavir in the hospitalized COVID-19 patients with the primary endpoint of mortality benefit. The patients were assigned to three groups based on respiratory support received at randomization.
They randomized 2,104 patients in the Dexamethasone trial and followed the initial goal of assigning 2,000 patients to the active drug treatment arms, but they only enrolled patients in the Hydroxychloroquine (n=1,561) and Lopinavir-Ritonavir (n=1,596) arms before they terminated the trials. The trial enrolled a varying proportion of patients to the invasive ventilation group (4% in Lopinavir-Ritonavir, 15.7% in Dexamethasone and 16.8% in Hydroxychloroquine trial).5–7

2.3 The Outcomes Related to COVID-19 treated with Hydroxychloroquine among In-patients with symptomatic Disease (ORCHID) trial

This is an investigator-initiated, blinded, placebo-controlled, randomized trial evaluating Hydroxychloroquine for the treatment of hospitalized patients with COVID-19. The primary endpoint was 7-point ordinal scale outcomes at day 15, and the protocol noted an initial plan to enroll 510 patients.8,9,35 However, the trial was prematurely terminated with the enrollment of only 479 patients due to a lack of benefit8,9,35. This trial was not powered for mortality benefit and the Time to recovery endpoints.38

2.4 Coalition COVID-19 Brazil I Investigators - Hydroxychloroquine trial

This was an open-label trial (1:1:1 randomization), and patients with mild-moderate disease were allocated based on the respiratory support received at randomization. Their primary endpoint was 7-point ordinal scale outcomes, and the trial enrolled 665 patients.10 The trial did not show any benefit with Hydroxychloroquine for the studied 7-point ordinal scale outcomes.

2.5 World Health Organization (WHO) Solidarity trial platform

This is an open-label, randomized controlled trial that employed 4-arms in a 1:1:1:1 ratio randomization to either the control, the Lopinavir/Ritonavir, Hydroxychloroquine, and Remdesivir arms with all-cause
mortality as the primary endpoint in the Canadian Solidarity trial. The protocol allows the individual countries participating in the trial to customize the protocol with an option to combine Interferon beta-1a with the Lopinavir/Ritonavir arm. After the recruitment of 5,500 patients as of July 1, 2020, they discontinue the trial’s Hydroxychloroquine and Lopinavir/Ritonavir arms that showed no mortality benefit\(^{11,34,39}\).

3. Are the conclusions of these randomized trials with sampling errors valid in providing guidance for treatment options during this pandemic?

Many randomized controlled trials for COVID-19 were designed in the beginning of this pandemic when we did not know much about the COVID-19 disease process. Since then, we know more about the epidemiology of this disease and the disease severity indicators that can prognosticate patients who are at increased risk for worse outcomes\(^{12-20,22-32}\). Based on disease prognostic markers that we know now, it is apparent that the completed randomized trials with limitations in randomization methodologies failed to show if the compared groups were matched for the important disease severity indicators to avoid sampling errors.

It is understandable that there is an urgent need to find therapeutic options that do not exist in the middle of this pandemic, but it is also equally important to plan the randomized clinical trials and then critically analyze the data with currently known prognostication markers that will be helpful to improve the analysis of data and/or protocol revisions of ongoing randomized trials and the design of any future randomized trials.

In the randomization of the ACTT-1 trial, an excess of 23 very sick patients were randomized to the placebo group on mechanical ventilation (n=22) and on high flow oxygen (n=1), and 30 additional patients with less severe disease who were on oxygen nasal canula (n=23) and not on supplemental oxygen (n=7) were randomized to Remdesivir\(^5\). These 53 patients’ mismatches could easily have created a positive outcome
for the studied endpoint, in favor of the trial drug Remdesivir (Time to recovery). Data on 42 patients from Table 2 of the publication is missing, and the allocation of these patients among Groups 4-7 is not clear. Additionally, there is also potential for a mismatch in the other disease severity indicators (NEW score) along with a possible mismatch in Group 5 patients in the placebo arm and Group 6 patients in the Remdesivir arm with a high death/ventilation ratio, signifying multiorgan failure that could be the cause of more deaths. This variability of Remdesivir benefit among the treatment groups is suggestive of varying levels of disease severity (NEW score mismatch) rather than the potential efficacy of the studied drug in one group and the lack of efficacy for the same in the other group.

Of the randomized controlled trials with mortality outcomes, the sampling errors with limitations in randomization methodologies would have caused Type I and Type II errors in both the Recovery and WHO Solidarity trials. The Recovery-Dexamethasone trial followed their goal of assigning 2,000 patients to the active drug treatment arm in order to adequately power the study. The Hydroxychloroquine and Lopinavir-Ritonavir studies were terminated early in both the Recovery and WHO Solidarity trials with a smaller sample size in each arm, which may have underpowered these trials unlike the Recovery-Dexamethasone trial (if 28-day mortality was 20%, the allocation of at least 2000 patients to the active treatment arm would yield at least 90% power at two-sided P=0.01 to detect a proportional reduction of one-fifth).

Of the randomized controlled trials with their Time to recovery and 7-point ordinal scale outcomes, the obvious sampling biases and NEW score mismatches with suboptimal randomization methodologies may have caused Type I errors for Remdesivir in the ACTT-1 trial. The Brazilian Hydroxychloroquine and the NIH ORCHID trials may have similar disease severity mismatches and they may not be powered enough like the ACTT-1 Remdesivir trial which continued beyond their initial enrollment goal in order to sufficiently
power the study\textsuperscript{38}. Underpowering of any trial with insufficient sample size may result in the trial showing lack of benefit for an intervention even when one exists\textsuperscript{41,42}.

Sampling bias can also occur from the timing of therapy relative to the onset of illness. For example, in the Recovery–Hydroxychloroquine trial, the median time since the onset of symptoms was 9 days in the treatment arm and usual care arm. However, in the Recovery-Dexamethasone trial, the median time since the onset of symptoms was 8 days in the treatment arm and 9 days in the usual care arm\textsuperscript{5,6}. A relatively high toxic dose of Hydroxychloroquine was used in the Recovery–Hydroxychloroquine trial\textsuperscript{6}, and the importance of the potential therapeutic synergistic mechanism of zinc sulfate with Hydroxychloroquine was not explored in any of the randomized controlled trials\textsuperscript{43–45}.

The Solidarity trial platform also allows the individual countries participating in the trial to customize the protocol, which can result in high heterogeneity in the enrollment of patients in the trial that can cause difficulties in analyzing the data from heterogeneous populations and interpreting the results\textsuperscript{46}.

Limited information exists in the current randomized trials about the monitoring for cardiac abnormalities, and specifically, how to adjudicate the cause and effect of cardiac rhythm abnormalities where a significant proportion of patients may need admission to the intensive care unit and/or may need mechanical ventilation, vasopressor use and other drugs that can prolong the QT interval with associated hypoxemia, electrolyte imbalance and/or acidosis that can potentially cause cardiac arrhythmias\textsuperscript{47–53}. The existence of these confounding factors in COVID-19 patients with severe disease and high mortality rates can independently cause cardiac arrhythmias in addition to the potential causation from the studied trial intervention.
Despite these unforeseen limitations of the randomized control trials that were not known initially, the following analysis can correct these limitations. A comprehensive review of the raw data of the whole study cohort should be performed to identify variables that can independently prognosticate worse outcomes in the study cohort. For example, in the Recovery and ACTT-1 trials, an analysis of the data can be conducted on varying levels of hypoxemia (PaO2/FiO2 ratio), disease severity using standardized metrics (NEW score, SOFA score or an equivalent metric), and biochemical markers of disease severity among patients in three respiratory support groups. For the ACTT-1 trial, the data should be adjusted for the 53 patients that were randomized against the placebo in favor of the study drug Remdesivir. Based on the analysis of the baseline disease severity data in both compared groups, a standardized statistical analysis can be done to correct for any biases that are observed in any of the randomized groups. The data should also adjust for any bias in outcomes due to delays in the starting of treatment relative to onset of symptoms to make sure both the treatment group and control group are adequately matched in each respiratory subgroup that the patient is randomized to.

The strength and limitations of randomized controlled trials were detailed in a review article, and despite their strengths, the randomized controlled trials have significant limitations, including lack of external validity in the application of the findings to populations outside the study. These trials can take years to execute, and there are difficulties in performing randomized controlled trials for any infectious disease outbreaks rapidly on the basis of limited data that is available. In light of limitations of the current randomized trials for COVID-19, it will be counterproductive to give more importance to these randomized control trials at the expense of other potentially useful sources of data, especially during this public health emergency where observational studies can yield highly valuable information that will be helpful in designing better randomized trials for any current and/or future infectious disease pandemics. A recent meta-analysis of Chloroquine derivatives in treatment for COVID-19 showed benefit by improving clinical and virological...
outcomes in addition to reducing the mortality by a factor of 3, and electronic registry data analyses that did not show benefit were associated with a lack of basic treatment definitions and conflicts of interest\textsuperscript{54}. Multiple other observational studies have also showed a benefit of Hydroxychloroquine therapy in COVID-19 patients\textsuperscript{55–59}. It is also important to understand inclusion/exclusion criteria and the modes of statistical analysis that were utilized in the observational studies of any intervention, particularly in the case of Hydroxychloroquine for COVID-19 treatment. For example, some of these observational studies that showed a lack of benefit with Hydroxychloroquine for COVID-19 treatment included patients with moderate to severe disease with clinical deterioration that were given Hydroxychloroquine and compared them with stable patients with mild disease who did not receive Hydroxychloroquine\textsuperscript{60–63}. Additionally, some of the studies did not exclude the deaths that occurred during the first 24-48 hours of admission that would have biased findings against Hydroxychloroquine\textsuperscript{62,63}. Among the three observational studies from the United States East Coast that did not exclude the first 24-48 hour deaths, with similar mortality rates (21.7%, 21.8% and 20.3%) in all subjects with majority of the patients receiving Hydroxychloroquine (75.9%, 76.2% and 70%), the first study performed statistical analysis among all subjects and found that the use of Hydroxychloroquine was associated with decreased in-hospital mortality\textsuperscript{59}. The other two studies compared moderate-severe disease patients who received Hydroxychloroquine with patients with mild disease who did not receive Hydroxychloroquine and reported lack of benefit with Hydroxychloroquine for COVID-19\textsuperscript{61,62}. This shows that statistical analysis of raw data by normalizing all the variables that could impact the outcomes of treatment intervention is a must.

Although a small number of randomized controlled trials were subjected to reanalysis in the published data, reanalysis of randomized controlled trial data from a sample comprising 36 articles showed that thirty-five percent of published reanalyses resulted in different conclusions compared to those of the original articles\textsuperscript{64}. The recent retraction of the two publications pertaining to COVID-19 from highly reputed journals
reinforces the benefits of careful analysis of the raw data from any clinical research study, instead of rushing it into the public domain that could potentially misguide the clinicians and have adverse effects on the society.65–68.

4. **Suggestions for improvement of randomization methodologies for randomized controlled trials**

4.1 **Need for a systematic approach to randomization**

The COVID-19 infection is a multi-systemic disease with the potential for a rapid deterioration of condition at any level of baseline respiratory status. In addition to baseline hypoxemia (PaO2/FiO2 ratio), the randomization should also include disease severity based on the SOFA score or NEW score (or equivalent disease severity metrics) and baseline lactic acid levels, lactate dehydrogenase, ferritin, absolute lymphocyte count, neutrophil count, renal function, liver function test including serum albumin, C-reactive protein, ESR, interleukin-6 and D-Dimer on admission and at the time of randomization. We also recommend the following:

- Any ongoing and future trials should have mortality benefit as the primary endpoint.
- Where possible, all ongoing trials should change their protocol randomization methodologies to include other disease severity indicators for randomization.
- We suggest that the sample size should be enough to adequately power the trials along the same lines as the Recovery trial [“if 28-day mortality was 20% then a comparison of at least 2000 patients allocated to active drug and 4000 to usual care alone would yield at least 90% power at two-sided P=0.01 to detect a proportional reduction of one-fifth (a clinically relevant absolute difference of 4 percentage points between the two arms)” 5,6,33].

*Since the current projected mortality is less than 20%, the sample size may have to be adjusted to higher numbers to ensure a power of 90% for the trial.*
4.2 Need for a systematic approach to publish all of the important baseline disease severity data

All of the baseline co-morbidities, including varying levels of hypoxemia (PaO2/FiO2 ratio), and all the disease severity indicators as listed above, should be included in the published data for observational studies and even more so for randomized trials. In the case of observational cohort studies, it is essential that the study authors match both the treatment and control groups equally without any bias for baseline comorbidities and all known COVID-19 disease severity indicators, and follow standardized statistical methodologies to reach their conclusions.

5. Conclusions

The COVID-19 randomized controlled trials were rapidly completed on the basis of limited and imperfect available data about COVID-19 disease severity, and have limitations in yielding high-quality data based on the disease severity information that is currently known. Like any new disease state, we are making incremental progress in our understanding about COVID-19 disease severity, and the efficacy of therapeutic agents from all available data from various study designs including randomized controlled trials and observational studies. It is prudent to rely on the knowledge we gain from both the randomized controlled trials and non-randomized studies to make treatment decisions, and to design future randomized trials for COVID-19 or any new infectious disease emergencies. Only double-blind placebo controlled randomized trials when done with robust randomization methodologies and due process to match all treatment groups can yield valid data, that can give clinical guidance for treating COVID-19 patients.

Declarations

Funding: No funding
Competing Interests: The authors have no conflict of interest or financial relationships relevant to the submitted review to disclose.

Ethical Approval: Not required

REFERENCES


19. Redfern OC, Smith GB, Prytherch DR, Meredith P, Inada-Kim M, Schmidt PE. A Comparison of the Quick Sequential (Sepsis-Related) Organ Failure Assessment Score and


46. TOUMI M, LIANG S, DABBOUS M, WANG Y, QIU T, HAN R. COVID-19: Why SOLIDARITY and DisCoVeRy trials may fail to bring informative and timely results. *medRxiv*. Published online January 1, 2020:2020.06.01.20118927. doi:10.1101/2020.06.01.20118927


Table 1: The disease severity in critically ill patients and mortality. Mortality rates of all subjects in the Recovery and ACTT-1 trials for COVID-19.

<table>
<thead>
<tr>
<th>SOFA Score</th>
<th>NEW score or NEW2 Score</th>
<th>mSOFA score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean SOFA score</td>
<td>Mortality</td>
<td>Initial Score</td>
</tr>
<tr>
<td>0-1</td>
<td>1%</td>
<td>1-4</td>
</tr>
<tr>
<td>1.1-2.0</td>
<td>5%</td>
<td>5-6</td>
</tr>
<tr>
<td>2.1-3.0</td>
<td>16%</td>
<td>7-8</td>
</tr>
<tr>
<td>3.1-4.0</td>
<td>13%</td>
<td>9-20</td>
</tr>
<tr>
<td>4.1-5.0</td>
<td>19%</td>
<td></td>
</tr>
<tr>
<td>&gt;5.1</td>
<td>27%</td>
<td></td>
</tr>
</tbody>
</table>

The following four hypothetical patients who are on the same 4L of oxygen on nasal canula and have the same preexisting baseline comorbidities, and their risk of death is based on disease severity.

<table>
<thead>
<tr>
<th>NEW Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score range</td>
</tr>
<tr>
<td>Patient 1</td>
</tr>
<tr>
<td>Patient 2</td>
</tr>
<tr>
<td>Patient 3</td>
</tr>
<tr>
<td>Patient 4</td>
</tr>
</tbody>
</table>

Mortality rate (all trial subjects) in the COVID-19 Randomized trials

<table>
<thead>
<tr>
<th>Mortality rate</th>
<th>Chances for randomization sampling errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery-Dexamethasone (n=6425) 28-day mortality</td>
<td>24.8%</td>
</tr>
<tr>
<td>Recovery-Hydroxychloroquine (n=4716) 28-day mortality</td>
<td>25.6%</td>
</tr>
<tr>
<td>ACTT-1 Remdesivir (n=1059) No. of deaths at 14th days</td>
<td>8.2%</td>
</tr>
</tbody>
</table>
Table 2: Respiratory support at randomization and proportion of deaths in each group.

<table>
<thead>
<tr>
<th>Proportion of patients randomized based on respiratory support¹,²</th>
<th>No oxygen (n) %</th>
<th>On oxygen * (n)%</th>
<th>Invasive mechanical ventilation/ECMO (n)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery-Dexamethasone (n=6425)</td>
<td>1535 (23.9%)</td>
<td>3883 (60.4%)</td>
<td>1007 (15.7%)</td>
</tr>
<tr>
<td>Recovery-Hydroxychloroquine (n=4716)</td>
<td>1112 (23.5%)</td>
<td>2811 (59.6%)</td>
<td>793 (16.8%)</td>
</tr>
<tr>
<td>Recovery-lopinavir-ritonavir (n=4972)</td>
<td>26%</td>
<td>70%</td>
<td>4%</td>
</tr>
<tr>
<td>ACTT-1 Remdesivir (n=1059) (Missing data n=42 from Table 2)</td>
<td>127 (12.0%)</td>
<td>421+197 (58.3%)</td>
<td>232 (25.8%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proportion of deaths in each group - Randomization based on respiratory support</th>
<th>No oxygen (n) %</th>
<th>On oxygen * (n)%</th>
<th>Invasive mechanical ventilation/ECMO (n)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery-Dexamethasone (Fig 3) Deaths (n=1592)</td>
<td>234 (14.7%)</td>
<td>980 (61.6%)</td>
<td>378 (23.7%)</td>
</tr>
<tr>
<td>Recovery-Hydroxychloroquine (Fig 3) Deaths (n=1206)</td>
<td>156 (12.9%)</td>
<td>724 (60.0%)</td>
<td>326 (27.0%)</td>
</tr>
<tr>
<td>ACTT-1 Remdesivir (Table 2) day 15 score data, Deaths (n=88)</td>
<td>2 (2.3%)</td>
<td>50 (56.8%)</td>
<td>33 (37.5%)</td>
</tr>
</tbody>
</table>

¹Recovery trials - Oxygen only respiratory support received at randomization, ACTT-1 trial, patients in both Group 5 (supplemental oxygen) and group 6 (noninvasive ventilation or use of high-flow oxygen).
Table 3: Randomized controlled trials for the hospitalized COVID-19 patients, randomization plan and COVID-19 disease severity status at randomization.

<table>
<thead>
<tr>
<th>Trial design</th>
<th>Recovery- Dexamethasone</th>
<th>Recovery- Hydroxychloroquine</th>
<th>ACTT-1 Remdesivir</th>
<th>NIH/NIAID ORCHID Trial</th>
<th>Brazil Hydroxychloroquine</th>
<th>WHO Solidarity trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allocation plan</td>
<td>1:2 allocation (active treatment: usual care)</td>
<td>1:2 allocation (active treatment: usual care)</td>
<td>1:1</td>
<td>1:1:1 randomization</td>
<td>1:1:1:1 randomization</td>
<td>1:1:1:1 (control, Lopinavir/Ritonavir and or Interferon beta-1a, Hydroxychloroquine, and Remdesivir arms)</td>
</tr>
<tr>
<td>Randomization plan</td>
<td>Respiratory support (no oxygen, Oxygen, Invasive mechanical ventilation)</td>
<td>Respiratory support (no oxygen, Oxygen, Invasive mechanical ventilation)</td>
<td>Respiratory support (group 4-7)</td>
<td>Respiratory support</td>
<td>Patients with mild disease</td>
<td>Hospitalized patients</td>
</tr>
<tr>
<td>Primary endpoint</td>
<td>28-day Mortality</td>
<td>28-day Mortality</td>
<td>7-point ordinal scale changed to Time to recovery</td>
<td>7-point ordinal scale</td>
<td>7-point ordinal scale</td>
<td>All-cause Mortality</td>
</tr>
<tr>
<td>Sample size goal</td>
<td>2000/4000</td>
<td>2000/4000</td>
<td>572, changed to continued enrollment to assure 400 recoveries</td>
<td>510, stopped after 479</td>
<td>630</td>
<td>~50,000</td>
</tr>
<tr>
<td>Power of trial for primary endpoint</td>
<td>at least 90% power at two-sided P=0.01</td>
<td>at least 90% power at two-sided P=0.01</td>
<td>85% power for detecting a recovery rate ratio of 1.35 with a two-sided type-I error rate of 5%</td>
<td>90% power to detect an odds ratio of 1.82 with a two-sided significance level of p&lt;0.05</td>
<td>80%</td>
<td>Not published</td>
</tr>
<tr>
<td>Final enrollment n=</td>
<td>2,104/4,321</td>
<td>1,561/3,155</td>
<td>1,063</td>
<td>479</td>
<td>665</td>
<td>5,500 patients as of July 1, 2020</td>
</tr>
<tr>
<td>Early termination</td>
<td>No</td>
<td>Yes, before 2000/4000 reached</td>
<td>Continued after 572, to ensure at least 400 recoveries and to address</td>
<td>Yes, lack of benefit</td>
<td>No</td>
<td>Lopinavir/Ritonavir and Hydroxychloroquine terminated early due to lack of benefit</td>
</tr>
<tr>
<td>Varying levels of baseline hypoxemia (PaO2/FiO2 ratio)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>NEW score, SOFA, or mSOFA Score</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Biochemical markers of disease severity</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Total deaths* - all subjects in the trial</td>
<td>1519</td>
<td>1206</td>
<td>87</td>
<td>Not published</td>
<td>18</td>
<td>Not published</td>
</tr>
<tr>
<td>Mortality rate(%) - all subjects in the trial</td>
<td>23.6% (28-day mortality)</td>
<td>25.6% (28-day mortality)</td>
<td>8.2% (No. of deaths 14th day)</td>
<td>Not published</td>
<td>2.7% (hospital deaths)</td>
<td>Not published</td>
</tr>
</tbody>
</table>