Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

International Journal of Antimicrobial Agents 57 (2021) 106222

Contents lists available at ScienceDirect

International Journal of Antimicrobial Agents

journal homepage: www.elsevier.com/locate/ijantimicag

Review

Randomised controlled trials for COVID-19: evaluation of optimal randomisation methodologies—need for data validation of the completed trials and to improve ongoing and future randomised trial designs

Venkata R. Emani[a,*], Sanjeev Goswami[b], Dheeraj Nandanoor[c], Shaila R. Emani[a], Nidhi K. Reddy[d], Raghunath Reddy[d]

[a] Central Valley Cardiovascular Associates, Inc., 1148 Norman Drive, Suite #2, Manteca, CA 95336, USA
[b] San Joaquin Critical Care Medical Group, 1801 E March Ln c300, Stockton, CA 95210, USA
[c] Synergy Med, 1801 E March Ln STE D460, Stockton, CA 95210, USA
[d] Stockton Primary Care, 805 N California St #102, Stockton, CA 95204, USA

ARTICLE INFO

Article history:
Received 3 September 2020
Accepted 1 November 2020

Keywords:
COVID-19
Randomised controlled trial
Methodology

ABSTRACT

During the emerging COVID-19 (coronavirus disease 2019) pandemic, initially there were no proven treatment options. With the release of randomised controlled trial (RCT) results, we are beginning to see possible treatment options for COVID-19. The RECOVERY trial showed an absolute risk reduction in mortality by 2.8% with dexamethasone, and the ACTT-1 trial showed that treatment with remdesivir reduced the time to recovery by 4 days. Treatment with hydroxychloroquine (HCQ) and lopinavir/ritonavir did not show any mortality benefit in either the RECOVERY or World Health Organization (WHO) Solidarity trials. The National Institutes of Health (NIH) and Brazilian HCQ trials did not show any benefit for HCQ based on the seven-point ordinal scale outcomes. The randomisation methodologies utilised in these controlled trials and the quality of published data were reviewed to examine their adaptability to treat patients. We found that the randomisation methodologies of these trials were suboptimal for matching the studied groups based on disease severity among critically-ill hospitalised COVID-19 patients with high mortality rates. The published literature is very limited regarding the disease severity metrics among the compared groups and failed to show that the data are 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 are accurate. We also propose proper randomisation methodologies for the design of RCTs for COVID-19 as well as guidance for the publication of COVID-19 trial results.

1. Introduction

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

A properly performed randomised trial is always superior and provides the highest quality data. Any large, well-designed RCTs should evenly distribute known and unknown factors among the intervention and control groups in order to minimise the potential for bias [1]. However, the large proportion of negative trials is a problem in critical care settings [2], which is largely due to heterogeneous patient populations and variable disease pre-

* Corresponding author. Tel.: +1 209 239 7852.
E-mail address: vemani@yahoo.com (V.R. Emani).

https://doi.org/10.1016/j.ijantimicag.2020.106222

sentations and levels of response to treatment among these patients. Preserving the integrity of clinical trials during the coronavirus pandemic is crucial [3], and these trials have to be critically examined to ensure that they are yielding valid data. The Adaptive COVID-19 Treatment Trial (ACTT-1) and the Outcomes Related to COVID-19 treated with Hydroxychloroquine among Inpatients with symptomatic Disease (ORCHID) trial by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (NIH/NIAID) are the only double-blind, placebo-controlled randomised trials, whereas all of the other trials for COVID-19 are open-label RCTs. In this manuscript, the following RCTs were reviewed to analyse the clinical impact of the treatment options.

- The ACTT-1 trial showed no statistically significant mortality benefit for remdesivir, but it reduced the time to recovery by 4 days [4].
- The Randomized Evaluation of COVID-19 Therapy (RECOVERY) trial showed an absolute risk reduction in mortality by 2.8% with dexamethasone and this beneficial effect was mainly seen in patients who required invasive mechanical ventilation [5–7]. This trial did not show any benefit for treatment with either hydroxychloroquine (HCQ) or lopinavir/ritonavir (LPV/r) [5–7].
- The halted ORCHID trial did not show any benefit for HCQ based on its seven-point ordinal scale outcomes [8,9].
- The Brazilian HCQ trial also did not show any benefit for HCQ based on its seven-point ordinal scale outcomes [10].
- The World Health Organization (WHO) Solidarity trial showed no mortality benefit using HCQ and LPV/r treatments [11].

These trials were launched rapidly in the middle of the COVID-19 pandemic and the published findings of these trials lack important details about disease severity among these critically-ill hospitalised patients. In this review, we examined whether these RCT designs live up to the expectations of yielding high-quality data that can fully guide patient treatment.

2. Randomisation strategies for critically-ill hospitalised patients and the limitations of current trials' randomisation methodologies for COVID-19

COVID-19 is a multisystemic disease with a high mortality rate in critically-ill hospitalised patients [4–7]. In critically-ill hospitalised patients with suspected infection, the Sequential Organ Failure Assessment (SOFA) score and National Early Warning Score (NEWS) were demonstrated to be superior for the prediction of mortality [12–20]. The protocol of the ACTT-1 trial adopted a daily NEWS assessment [4,21], and a recent study showed the important prognostic value of the SOFA score in predicting poor outcomes in COVID-19 patients [22]. A quick COVID-19 Severity Index (qCSI) scoring system [23] accurately predicted patients' progress to respiratory failure within 24 h of admission using bedside respiratory examination findings that employed similar respiratory parameters as the SOFA and NEWS scores. In addition to ground-glass opacities present on chest computed tomography (CT), other laboratory findings including lymphopenia, particularly in young healthy adults, thrombocytopenia, hypoalbuminaemia, elevated levels of D-dimer, C-reactive protein (CRP), erythrocytic sedimentation rate (ESR), interleukin-6 (IL-6), procalcitonin, lactate dehydrogenase and neutrophil count are considered to be other important indicators for COVID-19 disease severity and worse prognosis [22,24–32].

All of the completed RCTs only randomised patients based on the respiratory support received at randomisation [4–6,10,21,33–35]. The examples in Table 1 illustrate the flaws in the randomisation methodologies of these trials when we specifically reviewed the impact of COVID-19 disease severity. For example, a group

Table 1
Disease severity in critically ill patients and mortality: mortality rates of all subjects in the RECOVERY and ACTT-1 trials for COVID-19.

Disease severity in critically ill patients and mortality		*NEWS or NEWS 2 score [12]*		*mSOFA score [69]*	
SOFA score [15]					
Mean SOFA score	Mortality	Initial score	30-day mortality	Initial score	30-day mortality
0–1	1%	1–4	5.5%	0–7	4%
1.1–2.0	5%	5–6	11.3%	8–11	31%
2.1–3.0	16%	7–8	13.3%	>11	58%
3.1–4.0	13%	9–20	27.6%		
4.1–5.0	19%				
>5.1	27%				
The following four hypothetical patients who are on the same 4 L of oxygen on nasal canula and have the same pre-existing baseline co-morbidities, and their risk of death is based on disease severity.					
NEWS score [12]					
		Score range	**30-day mortality**		
Patient 1		1–4	5.5%		
Patient 2		5–6	11.3%		
Patient 3		7–8	13.3%		
Patient 4		9–20	27.6%		
Mortality rate (all trial subjects) in the COVID-19 randomised trials					
		Mortality rate	**Chances for randomisation sampling errors**		
Recovery–Dexamethasone (n = 6425) 28-day mortality [5]		24.8%	Very high		
Recovery–Hydroxychloroquine (n = 4716) 28-day mortality [6]		25.6%	Very high		
ACTT-1 remdesivir (n = 1059) No. of deaths at 14 days [4]		8.2%	Very high		

COVID, coronavirus disease 2019; SOFA, Sequential Organ Failure Assessment; NEWS, National Early Warning Score; mSOFA, modified Sequential Organ Failure Assessment.

Table 2
Respiratory support at randomisation and proportion of deaths in each group

	No oxygen [*n* (%)]	On oxygen [a] [*n* (%)]	Invasive mechanical ventilation/ECMO [*n* (%)]
Proportion of patients randomised based on respiratory support [4–7]			
RECOVERY–Dexamethasone ($n = 6425$)	1535 (23.9)	3883 (60.4)	1007 (15.7)
RECOVERY–Hydroxychloroquine ($n = 4716$)	1112 (23.5)	2811 (59.6)	793 (16.8)
RECOVERY–lopinavir/ritonavir ($n = 4972$)	26	70	4
ACTT-1 remdesivir ($n = 1059$) (missing data $n = 42$ from Table 2)	127 (12.0)	421+197 (58.3)	272 (25.6)
Proportion of deaths in each group–randomisation based on respiratory support			
RECOVERY–Dexamethasone (Fig. 3) Deaths ($n = 1592$)	234 (14.7)	980 (61.6)	378 (23.7)
RECOVERY–Hydroxychloroquine (Fig. 3) deaths ($n = 1206$)	156 (12.9)	724 (60.0)	326 (27.0)
ACTT-1 remdesivir (Table 2) Day 15 score data, deaths ($n = 88$)	2 (2.3)	50 (56.8)	33 (37.5)

ECMO, extracorporeal membrane oxygenation.
[a] RECOVERY trials, oxygen-only respiratory support received at randomisation; ACTT-1 trial, patients both in Group 5 (supplemental oxygen) and Group 6 (non-invasive ventilation or use of high-flow oxygen).

of patients with the same baseline oxygen requirement and co-morbidities can have an expected 5–27.6% mortality risk based on their disease severity if the SOFA or NEWS score is applied [12,15]. Table 2 illustrates that all three of the groups randomised based on respiratory support, particularly the majority of patients randomised to the oxygen group in these trials, will be at an even greater risk for disease severity heterogenicity 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 randomisation [4–7,10]. Most importantly, the majority of deaths occurred among patients who were randomised 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 [4,5]. Since the mortality rate was as high as 26.6% among all subjects in these trials, the SOFA or NEWS score mismatch (5.0–27.6% variability in mortality) among the groups in these randomised 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 [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 provided in the Supplementary material). Similarly, the same type I or type II fatal sampling errors can occur in all other randomised trials since the randomisation 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 RCTs, as they did not randomise patients based on COVID-19 disease severity for varying levels of hypoxaemia (PaO_2/FiO_2 ratio) including the standardised metrics for disease severity in these critically-ill hospitalised patients and cannot yield high-quality data.

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

This was a double-blind, placebo-controlled randomised trial that evaluated the safety and efficacy of remdesivir in hospitalised 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 non-invasive ventilation or high-flow oxygen; and Group 7, receiving invasive mechanical ventilation or extracorporeal membrane oxygenation (ECMO)] based on respiratory support received at randomisation [4].

2.2. Randomized Evaluation of COVID-19 Therapy (RECOVERY) trial

This was an open-label 2:1 randomised trial to evaluate the efficacy of low-dose corticosteroids (dexamethasone), HCQ and LPV/r in hospitalised COVID-19 patients with the primary endpoint of mortality benefit. Patients were assigned to three groups based on the respiratory support received at randomisation [5–7].

The trial randomised 2104 patients in the dexamethasone trial and followed the initial goal of assigning 2000 patients to the active drug treatment arms, but they only enrolled 1561 patients in the HCQ arm and 1596 in the LPV/r arm before the trials were terminated [5–7]. The trial enrolled a varying proportion of patients to the invasive ventilation group (4% in LPV/r, 15.7% in dexamethasone and 16.8% in HCQ trials) [5–7].

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

This was an investigator-initiated, blinded, placebo-controlled, randomised trial evaluating HCQ for the treatment of hospitalised patients with COVID-19. The primary endpoint was seven-point ordinal scale outcomes at Day 15, and the protocol noted an initial plan to enrol 510 patients [8,9,35]. However, the trial was prematurely terminated with the enrolment of only 479 patients owing to an observed lack of benefit [8,9,35]. This trial was not powered for mortality benefit and the time to recovery endpoints [36].

2.4. Coalition COVID-19 Brazil I Investigators hydroxychloroquine trial

This was an open-label trial (1:1:1 randomisation) and patients with mild-to-moderate COVID-19 were allocated based on the respiratory support received at randomisation. The primary endpoint was seven-point ordinal scale outcomes and the trial enrolled 665 patients [10]. The trial did not show any benefit with HCQ for the studied seven-point ordinal scale outcomes.

2.5. World Health Organization (WHO) Solidarity trial platform

This was an open-label RCT that employed four arms in a 1:1:1:1 ratio randomisation to either the control, LPV/r, HCQ or remdesivir arms, with all-cause mortality as the primary endpoint in the Canadian Solidarity trial. The protocol allowed the individual countries participating in the trial to customise the protocol with an option to combine interferon beta-1a (IFN-β-1a) with the LPV/r arm. Following recruitment of 5500 patients as of 1 July 2020, they discontinued the trial's HCQ and LPV/r arms that showed no mortality benefit [11,34,37].

Table 3
Randomised controlled trials for hospitalised COVID-19 patients, randomisation plan and COVID-19 disease severity status at randomisation.

	RECOVERY–Dexamethasone	RECOVERY–Hydroxychloroquine	ACTT-1 remdesivir	NIH/NIAID ORCHID trial	Brazil hydroxychloroquine	WHO Solidarity trial
Trial design	Open-label, investigator-initiated	Open-label, investigator-initiated	Double-blind, placebo-controlled	Investigator-initiated, blinded, placebo-controlled	Open-label	Open-label (individual countries can customise protocol)
Allocation plan	1:2 allocation (active treatment:usual care)	1:2 allocation (active treatment:usual care)	1:1	1:1	1:1:1 randomisation	1:1:1:1 (control, lopinavir/ritonavir ± interferon-beta-1a, hydroxychloroquine and remdesivir arms)
Randomisation plan	Respiratory support (no oxygen, oxygen, invasive mechanical ventilation)	Respiratory support (no oxygen, oxygen, invasive mechanical ventilation)	Respiratory support (Groups 4–7)	Respiratory support	Patients with mild disease	Hospitalised patients
Primary endpoint	28-day mortality	28-day mortality	7-point ordinal scale changed to time to recovery	7-point ordinal scale	7-point ordinal scale	All-cause mortality
Sample size goal	2000/4000	2000/4000	572, changed to continued enrolment to assure 400 recoveries	510, stopped after 479	630	~50 000
Power of trial for primary endpoint	≥90% power at two-sided $P = 0.01$	≥90% power at two-sided $P = 0.01$	85% power for detecting a recovery rate ratio of 1.35 with a two-sided type I error rate of 5%	90% power to detect an odds ratio of 1.82 with a two-sided significance level of $P < 0.05$	80%	Not published
Final enrolment (n)	2104/4321	1561/3155	1063	479	665	5500 patients as of 1 July 2020
Early termination	No	Yes, before 2000/4000 reached	Continued after 572 to ensure ≥400 recoveries and to address subgroup analysis	Yes, lack of benefit	No	Lopinavir/ritonavir and hydroxychloroquine terminated early due to lack of benefit
Randomisation based on COVID-19 disease severity						
Varying levels of baseline hypoxaemia (PaO_2/FiO_2 ratio)	No	No	No	No	No	No
NEWS, SOFA or mSOFA score	No	No	No	No	No	No
Biochemical markers of disease severity	No	No	No	No	No	No
Total deaths–all subjects in the trial	1519	1206	87	Not published	18	Not published
Mortality rate (%)–all subjects in the trial	23.6% (28-day mortality)	25.6% (28-day mortality)	8.2% (no. of deaths at 14th day)	Not published	2.7% (hospital deaths)	Not published

COVID, coronavirus disease 2019; NEWS, National Early Warning Score; SOFA, Sequential Organ Failure Assessment; mSOFA, modified Sequential Organ Failure Assessment.

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

Many RCTs for COVID-19 were designed at the beginning of the 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 randomised trials with limitations in randomisation methodologies failed to show whether the compared groups were matched for important disease severity indicators to avoid sampling errors.

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

In the randomisation of the ACTT-1 trial, an excess of 23 very sick patients were randomised to the placebo group on mechanical ventilation ($n = 22$) or 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 randomised to remdesivir [5]. These 53 patient mismatches could easily have created a positive outcome for the studied endpoint (time to recovery) in favour of the trial drug remdesivir. Data on 42 patients from Table 2 of the publication are missing, and the allocation of these patients among Groups 4–7 is not clear. In addition, there is also potential for a mismatch in the other disease severity indicators (NEWS) 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 (NEWS 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 RCTs with mortality outcomes, sampling errors with limitations in randomisation methodologies would have caused type I and type II errors both in the RECOVERY and WHO Solidarity trials [5,6,10,11,38]. The RECOVERY–dexamethasone trial followed their goal of assigning 2000 patients to the active drug treatment arm in order to adequately power the study [5]. The HCQ and LPV/r studies were terminated early both in 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 [6,7,11] (if 28-day mortality was 20%, the allocation of at least 2000 patients to the active treatment arm would yield ≥90% power at two-sided $P = 0.01$ to detect a proportional reduction of one-fifth [5,6,33]).

Of the RCTs with time to recovery and seven-point ordinal scale outcomes, the obvious sampling biases and NEWS mismatches with suboptimal randomisation methodologies may have caused type I errors for remdesivir in the ACTT-1 trial [4]. The Brazilian HCQ 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 enrolment goal in order to sufficiently power the study [36]. Underpowering of any trial with insufficient sample size may result in the trial showing lack of benefit for an intervention even when one exists [39,40].

Sampling bias can also occur from the timing of therapy relative to the onset of illness. For example, in the RECOVERY–HCQ 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 [5,6]. A relatively high toxic dose of HCQ was used in the RECOVERY–HCQ trial [6], and the importance of the potential therapeutic synergistic mechanism of zinc sulfate with HCQ was not explored in any of the RCTs [41–43].

The Solidarity trial platform also allowed individual countries participating in the trial to customise the protocol, which can result in high heterogenicity in the enrolment of patients in the trial that can cause difficulties in analysing the data from heterogeneous populations and in interpreting the results [44].

Limited information exists in the current randomised trials about monitoring for cardiac abnormalities, and specifically how to adjudicate the cause and effect of cardiac rhythm abnormalities where a significant proportion of patients may require admission to the intensive care unit (ICU) and/or may need mechanical ventilation, vasopressor use and other drugs that can prolong the QT interval with associated hypoxaemia, electrolyte imbalance and/or acidosis that can potentially cause cardiac arrythmias [45–51]. 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 RCTs 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 predict 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 hypoxaemia (PaO_2/FiO_2 ratio), disease severity using standardised metrics (NEWS, 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 who were randomised against the placebo in favour of the study drug remdesivir. Based on analysis of the baseline disease severity data in both compared groups, a standardised statistical analysis can be performed to correct for any biases that are observed in any of the randomised groups. The data should also adjust for any bias in outcomes owing to delays in the starting of treatment relative to the onset of symptoms to ensure both the treatment group and the control group are adequately matched in each respiratory subgroup that the patient is randomised to.

The strength and limitations of RCTs were detailed in a review article and, despite their strengths, RCTs have significant limitations, including lack of external validity in the application of the findings to populations outside the study [1]. These trials can take years to execute and there are difficulties in performing RCTs for any infectious disease outbreaks rapidly on the basis of limited data that are available [1]. In light of the limitations of the current randomised trials for COVID-19, it will be counterproductive to give more importance to these RCTs at the expense of other potentially useful sources of data [1], especially during this public-health emergency where observational studies can yield highly valuable information that will be helpful in designing better randomised trials for any current and/or future infectious disease pandemics. A recent meta-analysis of chloroquine derivatives in the treatment for COVID-19 showed benefit by improving clinical and virological outcomes in addition to reducing 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 [52]. Multiple other observational studies have also showed a benefit of HCQ therapy in COVID-19 patients [53–58]. It is also important to understand inclusion/exclusion criteria

and the modes of statistical analysis that were utilised in the observational studies of any intervention, particularly in the case of HCQ for COVID-19 treatment. For example, some of these observational studies that showed a lack of benefit with HCQ for COVID-19 treatment included patients with moderate-to-severe disease with clinical deterioration that were given HCQ and compared them with stable patients with mild disease who did not receive HCQ [59–63]. In addition, some of the studies did not exclude deaths that occurred during the first 24–48 h of admission that would have biased findings against HCQ [61,62]. Among the three observational studies from the US East Coast that did not exclude the first 24–48 h deaths, with similar mortality rates (21.7%, 21.8% and 20.3%) in all subjects with the majority of the patients receiving HCQ (75.9%, 76.2% and 70%), the first study performed a statistical analysis among all subjects and found that the use of HCQ was associated with decreased in-hospital mortality [57]. The other two studies compared moderate-to-severe disease patients who received HCQ to patients with mild disease who did not receive HCQ and reported a lack of benefit with HCQ for COVID-19 [60,61]. This shows that statistical analysis of raw data by normalising all the variables that could impact the outcomes of treatment intervention is a must.

Although a small number of RCTs were subjected to re-analysis in the published data, re-analysis of RCT data from a sample comprising 36 articles showed that 35% of published re-analyses resulted in different conclusions compared with those of the original articles [64]. The recent retraction of 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 clinicians and have adverse effects on society [65–68].

4. Suggestions for improvement of randomisation methodologies for randomised controlled trials

4.1. Need for a systematic approach to randomisation

COVID-19 is a multisystemic disease with the potential for rapid deterioration of the condition at any level of baseline respiratory status. In addition to baseline hypoxaemia (PaO_2/FiO_2 ratio), randomisation should also include disease severity based on the SOFA or NEWS 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, CRP, ESR, IL-6 and D-dimer on admission and at the time of randomisation. 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 randomisation methodologies to include other disease severity indicators for randomisation; and
- 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 (PaO_2/FiO_2 ratio), and all of the disease severity indicators as listed above should be included in the published data for observational studies and even more so for randomised 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 co-morbidities and all known COVID-19 disease severity indicators, and also follow standardised statistical methodologies to reach their conclusions.

5. Conclusions

The COVID-19 RCTs 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 of COVID-19 disease severity and the efficacy of therapeutic agents from all available data from various study designs including RCTs and observational studies. It is prudent to rely on the knowledge we gain both from RCTs and non-randomised studies to make treatment decisions and to design future randomised trials for COVID-19 or any new infectious disease emergencies. Only double-blind, placebo-controlled randomised trials when performed with robust randomisation methodologies and due process to match all treatment groups can yield valid data that can give clinical guidance for treating COVID-19 patients.

Funding: None.

Competing interests: None declared.

Ethical approval: Not required.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijantimicag.2020.106222.

References

[1] Frieden TR. Evidence for health decision making—beyond randomized, controlled trials. N Engl J Med 2017;377:465–75. doi:10.1056/NEJMra1614394.
[2] Santacruz CA, Pereira AJ, Celis E, Vincent J-L. Which multicenter randomized controlled trials in critical care medicine have shown reduced mortality? A systematic review. Crit Care Med 2019;47:1680–91. doi:10.1097/CCM.0000000000004000.
[3] McDermott MM, Newman AB. Preserving clinical trial integrity during the coronavirus pandemic. JAMA 2020;323:2135–6. doi:10.1001/jama.2020.4689.
[4] Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the treatment of Covid-19—preliminary report. N Engl J Med 2020 Oct 8 [Epub ahead of print]. doi:10.1056/NEJMoa2007764.
[5] Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, Linsell L, et al., RECOVERY Collaborative Group Dexamethasone in hospitalized patients with Covid-19—preliminary report. N Engl J Med 2020 July 17 [Epub ahead of print]. doi:10.1056/NEJMoa2021436.
[6] Horby P, Mafham M, Linsell L, et al. Effect of hydroxychloroquine in hospitalized patients with COVID-19: preliminary results from a multi-centre, randomized, controlled trial. medRxiv 2020 July 15. doi:10.1101/2020.07.15.20151852.
[7] RECOVERY Collaborative Group Statement from the Chief Investigators: no clinical benefit from use of lopinavir–ritonavir in hospitalised COVID-19 patients studied in RECOVERY; 2020. https://www.recoverytrial.net/files/lopinavir-ritonavir-recovery-statement-29062020_final.pdf, [accessed 19 June 2020].
[8] National Institutes of Health (NIH) NIH halts clinical trial of hydroxychloroquine. NIH; 20 June 2020. https://www.nih.gov/news-events/news-releases/nih-halts-clinical-trial-hydroxychloroquine, [accessed 19 June 2020]..
[9] Petal Network. Major US trial closes showing no benefit for hydroxychloroquine in COVID-19 22 June 2020 https://petalnet.org/node/5886, [accessed 19 June 2020]..
[10] Cavalcanti AB, Zampieri FG, Rosa RG, Azevedo LCP, Veiga VC, Avezum A, et al. Hydroxychloroquine with or without azithromycin in mild-to-moderate

Covid-19. N Engl J Med 2020 July 23 [Epub ahead of print]. doi:10.1056/NEJMoa2019014.
[11] World Health Organization (WHO) WHO discontinues hydroxychloroquine and lopinavir/ritonavir treatment arms for COVID-19, Geneva, Switzerland: WHO; 4 July 2020. https://www.who.int/news-room/detail/04-07-2020-who-discontinues-hydroxychloroquine-and-lopinavir-ritonavir-treatment-arms-for-covid-19, [accessed 3 July 2020]..
[12] Royal College of Physicians (RCP) National Early Warning Score (NEWS) 2: standardising the assessment of acute-illness severity in the NHS, London, UK: RCP; 2017. https://www.rcplondon.ac.uk/file/8636/download, [accessed 12 November 2020].
[13] Vincent J-L, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med 1996;22:707–10. doi:10.1007/BF01709751.
[14] Vincent J-L, de Mendonca A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on 'sepsis-related problems' of the European Society of Intensive Care Medicine. Crit Care Med 1998;26:1793–800. doi:10.1097/00003246-199811000-00016.
[15] Ferreira FL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. JAMA 2001;286:1754–8. doi:10.1001/jama.286.14.1754.
[16] Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score—development, utility and challenges of accurate assessment in clinical trials. Crit Care 2019;23:374. doi:10.1186/s13054-019-2663-7.
[17] Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 2016;315:801–10. doi:10.1001/jama.2016.0287.
[18] Kovach CP, Fletcher GS, Rudd KE, Grant RM, Carlbom DJ. Correction: Comparative prognostic accuracy of sepsis scores for hospital mortality in adults with suspected infection in non-ICU and ICU at an academic public hospital. PLoS One 2019;14:e0224780. doi:10.1371/journal.pone.0224780.
[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 the National Early Warning Score in non-ICU patients with/without infection. Crit Care Med 2018;46:1923–33. doi:10.1097/CCM.0000000000003359.
[20] Goulden R, Hoyle M-C, Monis J, et al. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emerg Med J 2018;35:345–9. doi:10.1136/emermed-2017-207120.
[21] ACTT-1 Study Group ACTT-1 trial protocol. A multicenter, adaptive, randomized blinded controlled trial of the safety and efficacy of investigational therapeutics for the treatment of COVID-19 in hospitalized adults. DMID Protocol Number: 20-0006; 2020. https://www.nejm.org/doi/suppl/10.1056/NEJMoa2007764/suppl_file/nejmoa2007764_protocol.pdf, [accessed 12 November 2020].
[22] Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020;395:1054–62 Erratum in: Lancet 2020;395:1038. doi: 10.1016/S0140-6736(20)30638-3.. doi:10.1016/S0140-6736(20)30566-3.
[23] Haimovich A, Ravindra NG, Stoytchev S, et al. Development and validation of the quick COVID-19 Severity Index (qCSI): a prognostic tool for early clinical decompensation. Ann Emerg Med 2020;76:442–53. doi:10.1016/j.annemergmed.2020.07.022.
[24] Xie J, Tong Z, Guan X, Du B, Qiu H, Slutsky AS. Critical care crisis and some recommendations during the COVID-19 epidemic in China. Intensive Care Med 2020;46:837–40. doi:10.1007/s00134-020-05979-7.
[25] Liu J, Liu Y, Xiang P, et al. Neutrophil-to-lymphocyte ratio predicts critical illness patients with 2019 coronavirus disease in the early stage. J Transl Med 2020;18:206. doi:10.1186/s12967-020-02374-0.
[26] Lu Y, Sun K, Guo S, et al. Early warning indicators of severe COVID-19: a single-center study of cases from Shanghai, China. Front Med (Lausanne) 2020;7:432. doi:10.3389/fmed.2020.00432.
[27] Tan L, Kang X, Ji X, et al. Validation of predictors of disease severity and outcomes in COVID-19 patients: a descriptive and retrospective study. Med (N Y) 2020 May 19 [Epub ahead of print]. doi:10.1016/j.medj.2020.05.002.
[28] Zhang H, Wang X, Fu Z, et al. Potential factors for prediction of disease severity of COVID-19 patients. medRxiv 2020 Mar 23. doi:10.1101/2020.03.20.20039818.
[29] Zhou C, Huang Z, Tan W, et al. Predictive factors of severe coronavirus disease 2019 in previously healthy young adults: a single-center, retrospective study. Respir Res 2020;21:157. doi:10.1186/s12931-020-01412-1.
[30] Zachariah P, Johnson CL, Halabi KC, et al. Epidemiology, clinical features, and disease severity in patients with coronavirus disease 2019 (COVID-19) in a children's hospital in New York City, New York. JAMA Pediatr 2020;174:e202430. doi:10.1001/jamapediatrics.2020.2430.
[31] Bhargava A, Fukushima EA, Levine M, et al. Predictors for severe COVID-19 infection. Clin Infect Dis 2020;71:1962–8. doi:10.1093/cid/ciaa674.
[32] Cao J, Tu W-J, Cheng W, et al. Clinical features and short-term outcomes of 102 patients with coronavirus disease 2019 in Wuhan, China. Clin Infect Dis 2020;71:748–55. doi:10.1093/cid/ciaa243.
[33] The RECOVERY Collaborative Group Randomised Evaluation of COVID-19 Therapy (RECOVERY). Trial protocol; 2020. https://www.recoverytrial.net/files/recovery-protocol-v7-0-2020-06-18.pdf, [accessed 12 November 2020].
[34] ClinicalTrials.gov Treatments for COVID-19: Canadian arm of the Solidarity trial (CATCO); 2020. https://clinicaltrials.gov/ct2/show/NCT04330690?term=solidarity&cond=Covid19&draw=2&rank=1, [accessed 12 November 2020].
[35] Casey JD, Johnson NJ, Semler MW, et al. Rationale and design of ORCHID: a randomized placebo-controlled trial of hydroxychloroquine for adults hospitalized with COVID-19. Ann Am Thorac Soc 2020;17:1144–53. doi:10.1513/AnnalsATS.202005-478SD.
[36] ACTT-1 Study Group Supplementary appendix to manuscript entitled 'Remdesivir for the treatment of COVID-19–preliminary report. New England Journal of Medicine; 2020. https://www.nejm.org/doi/suppl/10.1056/NEJMoa2007764/suppl_file/nejmoa2007764_appendix.pdf, [accessed 12 November 2020].
[37] World Health Organization (WHO) 'Solidarity' clinical trial for COVID-19 treatments, Geneva, Switzerland: WHO; 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov/solidarity-clinical-trial-for-covid-19-treatments, [accessed 12 November 2020]..
[38] ClinicalTrials.gov Outcomes Related to COVID-19 treated with Hydroxychloroquine among In-patients with symptomatic Disease (ORCHID); 2020. https://clinicaltrials.gov/ct2/show/NCT04332991?term=ORCHID&cond=Covid19&draw=2&rank=1, [accessed 12 November 2020].
[39] Freiman JA, Chalmers TC, Smith H, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. N Engl J Med 1978;299:690–4. doi:10.1056/NEJM197809282991304.
[40] Moher D. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA 1994;272:122–4. doi:10.1001/jama.1994.03520020048013.
[41] Carlucci PM, Ahuja T, Petrilli C, Rajagopalan H, Jones S, Rahimian J. Zinc sulfate in combination with a zinc ionophore may improve outcomes in hospitalized COVID-19 patients. J Med Microbiol 2020;69:122–34. doi:10.1099/jmm.0.001250.
[42] Xue J, Moyer A, Peng B, Wu J, Hannafon BN, Ding W-Q. Chloroquine is a zinc ionophore. PLoS One 2014;9:e109180. doi:10.1371/journal.pone.0109180.
[43] Maret W. Zinc in cellular regulation: the nature and significance of 'zinc signals'. Int J Mol Sci 2017;18:2285. doi:10.3390/ijms18112285.
[44] 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 2020 June 2. doi:10.1101/2020.06.01.20118927.
[45] Goyal P, Choi JJ, Pinheiro LC, et al. Clinical characteristics of Covid-19 in New York City. N Engl J Med 2020;382:2372–4. doi:10.1056/NEJMc2010419.
[46] Richardson S, Hirsch JS, Narasimhan M, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City Area. JAMA 2020;323:2052–9. doi:10.1001/jama.2020.6775.
[47] Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ 2020;369:m1966. doi:10.1136/bmj.m1966.
[48] Lewnard JA, Liu VX, Jackson ML, et al. Incidence, clinical outcomes, and transmission dynamics of severe coronavirus disease 2019 in California and Washington: prospective cohort study. BMJ 2020;369:m1923. doi:10.1136/bmj.m1923.
[49] ISARIC (International Severe Acute Respiratory and Emerging Infections Consortium) COVID-19 report: 03 September 2020. ISARIC; 2020. https://media.tghn.org/medialibrary/2020/09/ISARIC_Data_Platform_COVID-19_Report_20AUG20.pdf, [accessed 2 September 2020]..
[50] Cummings MJ, Baldwin MR, Abrams D, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. Lancet 2020;395:1763–70. doi:10.1016/S0140-6736(20)31189-2.
[51] Richards-Belle A, Orzechowska I, Gould DW, et al. COVID-19 in critical care: epidemiology of the first epidemic wave across England, Wales and Northern Ireland. Intensive Care Med 2020;46:2035–47. doi:10.1007/s00134-020-06267-0.
[52] Million M, Gautret P, Colson P, et al. Clinical efficacy of chloroquine derivatives in COVID-19 infection: comparative meta-analysis between the big data and the real world. New Microbes New Infect 2020;38:100709. doi:10.1016/j.nmni.2020.100709.
[53] Gautret P, Lagier J-C, Parola P, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents 2020;56:105949. doi:10.1016/j.ijantimicag.2020.105949.
[54] Arshad S, Kilgore P, Chaudhry ZS, et al. Treatment with hydroxychloroquine, azithromycin, and combination in patients hospitalized with COVID-19. Int J Infect Dis 2020;97:396–403. doi:10.1016/j.ijid.2020.06.099.
[55] COVID-19 RISK and Treatments (CORIST) CollaborationUse of hydroxychloroquine in hospitalised COVID-19 patients is associated with reduced mortality: findings from the observational multicentre Italian CORIST study. Eur J Intern Med 2020 Aug 25 [Epub ahead of print]. doi:10.1016/j.ejim.2020.08.019.
[56] Catteau L, Dauby N, Montourcy M, et al. Low-dose hydroxychloroquine therapy and mortality in hospitalized patients with COVID-19: a nationwide observational study of 8075 participants. Int J Antimicrob Agents 2020;56:106144. doi:10.1016/j.ijantimicag.2020.106144.
[57] Mikami T, Miyashita H, Yamada T, et al. Risk factors for mortality in patients with COVID-19 in New York City. J Gen Intern Med 2020 June 30 [Epub ahead of print]. doi:10.1007/s11606-020-05983-z.
[58] Bernaola N, Mena R, Bernaola A, et al. Observational study of the efficiency of treatments in patients hospitalized with Covid-19 in Madrid. medRxiv 2020 July 21. doi:10.1101/2020.07.17.20155960.

[59] Geleris J, Sun Y, Platt J, et al. Observational study of hydroxychloroquine in hospitalized patients with Covid-19. N Engl J Med 2020;382:2411–18. doi:10.1056/NEJMoa2012410.
[60] Ip A, Berry DA, Hansen E, et al. Hydroxychloroquine and tocilizumab therapy in COVID-19 patients—an observational study. PLoS One 2020;15:e0237693. doi:10.1371/journal.pone.0237693.
[61] Rosenberg ES, Dufort EM, Udo T, et al. Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with COVID-19 in New York State. JAMA 2020;323:2493–502. doi:10.1001/jama.2020.8630.
[62] Magagnoli J, Narendran S, Pereira F, et al. Outcomes of hydroxychloroquine usage in United States veterans hospitalized with COVID-19. Med (N Y) 2020 June 5 [Epub ahead of print]. doi:10.1016/j.medj.2020.06.001.
[63] Geleris J, Sun Y, Platt J, et al. Observational study of hydroxychloroquine in hospitalized patients with Covid-19. N Engl J Med 2020;382:2411–18. doi:10.1056/NEJMoa2012410.
[64] Ebrahim S, Sohani ZN, Montoya L, et al. Reanalyses of randomized clinical trial data. JAMA 2014;312:1024–32. doi:10.1001/jama.2014.9646.
[65] Mehra MR, Ruschitzka F, Patel AN. Retraction—Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. Lancet 2020;395:1820. doi:10.1016/S0140-6736(20)31324-6.
[66] Mehra MR, Desai SS, Ruschitzka F, Patel AN. Retracted: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. Lancet 2020 May 22. doi:10.1016/S0140-6736(20)31180-6.
[67] Mehra MR, Desai SS, Kuy S, Henry TD, Patel AN. Retraction: Cardiovascular disease, drug therapy, and mortality in Covid-19. N Engl J Med DOI: 10.1056/NEJMoa2007621. N Engl J Med 2020;382:2582. doi: 10.1056/NEJMc2021225.
[68] Mehra MR, Desai SS, Kuy S, Henry TD, Patel AN. Cardiovascular disease, drug therapy, and mortality in Covid-19. N Engl J Med 2020;382:e102. doi:10.1056/NEJMoa2007621.
[69] Grissom CK, Brown SM, Kuttler KG, et al. A modified Sequential Organ Failure Assessment score for critical care triage. Disaster Med Public Health Prep 2010;4:277–84. doi:10.1001/dmp.2010.40.