Predictive Factors of Clinical Assays during COVID-19

Matthieu MILLION[1,2,*], Pierre DUDOUET[1,2], Eric CHABRIERE[1,2], Sébastien CORTAREDONA[1,3], Philippe BROUQUI[1,2], Didier RAOULT[1,2]

Affiliations :

[1]IHU-Méditerranée Infection, Marseille, France

[2]Aix Marseille Univ., IRD, AP-HM, MEPHI, Marseille, France

[3]Aix Marseille Univ., IRD, AP-HM, SSA, VITROME, Marseille, France

*** Corresponding author**: Prof. Matthieu MILLION. MEPHI, Institut Hospitalo-Universitaire Méditerranée Infection, 19-21 Boulevard Jean Moulin 13385 Marseille Cedex 05, France. Phone: + 33 (0) 4 13 73 24 01. Fax: + 33 (0) 4 13 73 24 02.

Email: matthieumillion@gmail.com

Abstract word count: 238

Main text word count: 2635

References: 29

Figures: 2

Tables: 3

ABSTRACT

Background

The Covid-19 pandemic led to a violent debate about the efficacy of hydroxychloroquine (HCQ) and remdesivir and about randomized controlled trials (RCTs) and observational studies. Here, we wanted to determine the most influential biases on the results of the clinical therapeutic studies in this context.

Methods

Predictive criteria were identified through critical review of studies assessing HCQ and remdesivir for Covid-19 mortality from March to November 2020. Multiple correspondence analysis, comparative meta-analysis, and predictive value were used to explore and identify criteria associated with study outcomes.

Results

Among the 61 included studies, potential conflict of interest, detailed therapeutic protocol, toxic treatment (overdose or use in contraindicated patients), known centers and doctors, and private data computing company were the criteria most predictive of study results. All 18 observational studies evaluating HCQ and reporting a detailed therapeutic protocol without conflict of interest were Pro. All 4 studies with toxic treatment and the 3 studies with a private data computing company were Con. Potential conflict of interest was a perfect predictor for remdesivir efficacy. RCTs were associated with HCQ inefficacy and potential conflict of interest.

Conclusions

In therapeutic trials on COVID-19, the major biases predicting the conclusions are not methodology nor data analysis, but conflict of interest and absence of medical expertise. The herein proposed criteria should help reviewers to avoid a new scandal of retracted articles and to improve the honesty and medical quality of future clinical therapeutic studies.

INTRODUCTION

In the COVID-19 episode, one of the greatest scientific scandals of all time occurred (1) with the rapid retractions of major publications in the New England Journal of Medicine and the Lancet (2, 3). In the meanwhile, a considerable debate has emerged on Remdesivir, which the WHO finally considered useless (4), a few days after the European Commission purchased 2 billion euros worth of the drug. The putative efficacy of remdesivir was mainly published in the New England Journal of Medicine, some of whose articles looked more like advertising than science (5,6). On the other hand, more than 180 publications have been made on hydroxychloroquine (HCQ), with censorship effects such as refusal to examine the publications, including ours (7), even though it was the largest mono-centric series in the world. All publications showing a positive effect of HCQ have been published in journals that until then were not the scientific leaders in the field. All this was done in an unprecedented financial context, since remdesivir, whose futility was finally shown (4), was the subject of unprecedented speculation on a pharmaceutical product and therefore the financial stakes were colossal (1,6).

Furthermore, conflicts of interest at all levels have been neglected: that of the government, politicians, scientific advisors, appointees (5) and that of the journals and the publishers themselves, whose funding is often common with that of the pharmaceutical industry, and who receive advertising from the pharmaceutical industry (1,5). Conflicts of interest of authors are often neglected, without being penalized in scientific journals, despite the evidence of bias (8,9). Finally, conflicts of interest of reviewers are neglected, given that the milieu of people who conduct therapeutic trials is very commonly affected by conflicts of interest, as shown for infectious disease academics (9).

In this context, an objective analysis of published data requires the establishment of new criteria, which are independent of these pressures, in order to have a certain reliability.

The absence of such criteria leads to variability in meta-analyses (10,11) which have moved out of the scientific domain to enter into a passionate, ideological, and commercial domain. Finally, meta-analysis, and apparently therapeutic trial specialists, take less account of current medical practice and care, and the risk of bias related to pharmaceutical company influence, but rather focus on methodologies commonly recommended by pharmaceutical companies. RCTs are not superior to observational studies (12, 13) so that there is no transcendental methodology in therapeutic trials. Multicentric RCTs only reflect one perspective, which is not universal (12,13), and which is more in line with the needs of the pharmaceutical industry than with the reality of practice, including in episodes of acute infection epidemics.

Overall, it seemed essential to list all the evaluation criteria for scientific studies, whether comparative, randomized or not, to assess their quality not from a medical-political point of view (5), and to consider the classifications obtained, depending on whether certain criteria are retained or excluded, which seem to us to be indicative of an ideological or financial bias. The basic elements of the clinical description have led to profound errors in the interpretation of the data, such as the lack of stratification of patients according to severity, which is also a mistake related to people who no longer practice or have never practiced medicine, and who make a single entity of a disease that has different stages, different degrees of severity, and different potential risks of mortality.

METHODS

Inclusions of studies: *Search strategy*

The global strategy to identify new evaluation criteria is detailed in the Supplementary data. Briefly, the keywords “hydroxychloroquine”, “HCQ”, “chloroquine”, “coronavirus”, “COVID-19”, “SARS-Cov-2”, and “remdesivir” were entered in PubMed, Google Scholar and Google search engines on studies published in English from March to November 11,

2020. An online search was also performed using the website https://c19study.com/. Only the death outcome was considered, so studies without any death were not eligible. We reviewed studies evaluating the effects of chloroquine derivatives and remdesivir against SARS-CoV-2 in groups of COVID-19 patients as compared to control groups of patients who did not receive any experimental treatment.

Identification of characteristics and criteria

The criteria are summarized in Table 1 and detailed in the Supplementary Data. Some of these criteria have already been identified in a previous work (14,15) and have been completed as we observed critical pitfalls in studies assessed for the present work.

Multiple correspondence analysis

Multiple correspondence analysis (MCA) is a statistically-based visualization method that allows the user to graphically represent and analyze the associations among categorical variables (16). The basic idea behind our approach was to use MCA 1) to construct synthetic quantitative variables that represent the studies, their characteristics, and their criteria (see Table 1) on a two-dimensional plane 2) to identify clusters of studies that shared the same criteria and characteristics. MCA was performed with the R software and the FactomineR package (17).

Predictive value

In a qualitative approach, we evaluated the predictive value of presence or absence of the identified criterion on the positive (Odds ratio for mortality < 1; identified as Pro regardless of significance) or negative (OR ≥ 1; identified as Con) outcome of included studies. The

association of the presence or absence of each criterion with Pro or Con was tested using a two-sided Fisher exact test. A p-value < 0.05 was considered significant.

Meta-analysis and heterogeneity

In a quantitative approach, when applicable, a comparative meta-analysis was performed with a random effects model using Comprehensive Meta-Analysis v3 (Biostat, Englewood, NJ, USA) as recommended by Borenstein *et al*. (18). The most adjusted effect size, reflecting the greatest control for potential confounding factors, was extracted. When propensity score matching was used, the number of matched patients was included in quantitative analysis. Heterogeneity was considered substantial when $I^2 > 50\%$. A p-value < 0.05 was considered significant. To identify which criteria were associated with a significant difference in summary effect, the Q-value and its p-value were reported, and criteria were ranked according to Q-value.

RESULTS

Multiple correspondence analysis

Unsupervised analysis (Figure 1) of HCQ studies evidenced three clusters. First, megatrials and RCTs were associated with New England Journal of Medicine, JAMA, unclear affiliations of authors, absence of laboratory confirmation of diagnosis, toxic treatment (overdose or use in contraindicated patients), unexpected results not reported and conclusions neglecting a 25% decrease in the risk of mortality. This cluster was associated with multinational studies, USA, UK and Brazil.

A second cluster regrouped big data studies, that were associated with private data computing company of unknown financing (and therefore a likely existence of a conflict of interest), the Lancet, a potential conflict of interest, unknown centers and doctors, undeclared

funding and conflict of interests, and absence of detailed therapeutic protocol and detailed treatment monitoring. These studies were also associated with the absence of an expert in the field among the authors and a role of previous health status and severity not ruled out (confounding by indication). These studies were associated with USA, Europe and Peru.

Conversely, monocentric studies were associated with absence of potential conflict of interest, an author expert in the field, a detailed therapeutic protocol, a detailed treatment monitoring, and standard care reported. This cluster was associated with Andorra, China, Egypt, France, Iran, Italy, Mexico, and Spain. These studies were mainly observational (but not "big data" studies), with a laboratory confirmation of the diagnosis, the different stages of disease kept separate, role of severity ruled out, centers and doctors clearly reported with at least one author expert in the field. These studies were associated with 2 journals: American Journal of Tropical Medicine and Hygiene, and International Journal of Antimicrobial Agents.

Predictive value of identified criteria for HCQ efficacy or inefficacy

Among the 6 studies on remdesivir, both positive and negative predictive value of potential conflict of interest with remdesivir were 100%. All 5 studies with a conflict of interest declared or not declared were in favor of remdesivir, the only study without conflict of interest reported no benefit with remdesivir. Among the 56 studies on HCQ, the following criteria were associated with a predictive value > 50% for HCQ efficacy (Table 2) : *Detailed treatment protocol* (84%), *At least one of the main authors expert in the field* (affiliated in infectious diseases, internal medicine or pneumology) (76%), *Control for severity (at least oxygen)* (75%), *Centers and doctors who take care of patients are identified* (73%), *Diagnosis formally confirmed* (PCR or serology-based diagnosis) (69%) and *Control for health status* (at least age) (63%). Conversely, the following criteria were associated with a predictive value

significantly > 50% for HCQ inefficacy: *Private data computing company* (100%), *Toxic treatment* (100%), *Potential conflict of interest* with remdesivir (73%) and *Undeclared funding or conflict of interest* (66%). The difference of predictive value according to each criterion was significant for *potential conflict of interest* (p = .001), lack of *detailed therapeutic protocol* (p = 0.011), *toxic treatment* (p = 0.013), *Unknown centers and doctors not known* (p = 0.03), and *private data computing company* (p = 0.041). The 18 observational studies with a detailed therapeutic protocol and without a potential conflict of interest had a 100% predictive value for HCQ efficacy (Table 3).

Comparative meta-analysis

Among these 18 studies, 16 provided quantitative results available for meta-analysis with a significant effect (n = 17, Odds ratio = 0.60, 95% confidence interval 0.52 – 0.70, $p = 6.7x10^{-12}$). This was not related to an isolated aberrant study as shown by one-study-removed meta-analysis (Supplementary Figure 1). Combination of HCQ with azithromycin (AZ) was associated with a significant beneficial effect compared to HCQ monotherapy (n = 5 comparisons with AZ in all patients, 0.36, 0.21 – 0.63 / n = 9 without AZ in any patient, 0.68, 0.56 – 0.82 / Q-value = 4.41, p = 0.036). Comparative meta-analysis with ranking by Q-value confirmed that potential conflict of interest, including private data computing company, was the criterion associated with the greatest and most significant difference in summary effect (Supplementary Table 1). Effect of HCQ on mortality was beneficial (n = 43, 0.75, 0.66 – 0.84, $p = 6.3 \times 10^{-7}$) or deleterious (n = 19, 1.15, 1.07 – 1.23, $p = 1.1 \times 10^{-4}$) when an absence or a presence of a potential conflict of interest was found, respectively (Figure 2).

Neglecting a non-significant but relevant decrease in mortality

We found 6 studies observing a decrease in the risk of mortality greater than 25% but this finding was not analyzed nor mentioned because it was not significant (Supplementary data), or thought to be not relevant to the outcome of the study. We previously commented this (17). Strikingly, the day-28 mortality was halved in a French RCT (20) suspended and closed after the publication of Mehra *et al.* (2). If the planned enrollment had been included (1300 patients), if the observed tendance were confirmed, the difference would have been significant (31/650 (4.8%) versus 58/650 (8.9%), Odds ratio 0.55, two-sided Mid-p exact test p = 0.003).

DISCUSSION

There is a conflict in the evaluation of therapeutics for infectious diseases between methodologists who recommend multicentric randomized controlled trials (RCTs), which are mainly used by the pharmaceutical industry, and observational studies performed by medical doctors. More recently, a third source of comparative analysis has been the analysis of large data (Big Data) collected automatically in health care centers. Interestingly in infectious diseases currently 83% of IDSA recommendations are not based on RCTs (21), although considered the gold standard. Moreover, RCTs require significant funding, and the pharmaceutical industry's willingness to demonstrate efficacy or non-inferiority is under pressure of conflict of interest because those who pay and analyze have a well-known and long-evaluated chance of having biased results in favor of the products they finance (8). Moreover, the obtention of the compound by the company for testing is commonly subject to a possible censorship as an approvement of the work presented is required (22). That may lead to dissimulate negative results (22).

The methodology of analysis used here is to our knowledge unique. Pharmaceutical industry is a major actor directly or indirectly influencing authors with conflicts of interest, declared or not (which is quite common among French authors that we were able to identify

thanks to the obligation of declaration in France). Potential conflict of interest with Gilead has a predictive value of 74% against HCQ (whereas 78% of the work with no link to this company is in favor of HCQ). This work also made it possible to identify the target journals of the work in which the remdesivir producer or its partners played an important role, which is the case of the New England Journal of Medicine.

Concerning Big Data, this is a new problem. In some Big data studies, data acquisition is directly financed by Gilead, the pharmaceutical industry with a conflict of interest against HCQ (23,24). In another Big data study reporting a beneficial effect of remdesivir and a deleterious effect of HCQ, a direct conflict of interest is declared by several authors (25). Companies such as Surgisphere, two papers of which had to be retracted (2,3), have unknown funding, something that should have been required from the publisher. One may question if companies such as Surgisphere (2), TARGET PharmaSolutions (24), and TriNetX (26) have received funding since these Big Data studies also clearly have a predictive value in favor of remdesivir (24) and to the disadvantage of HCQ (2,24,26). This suggests that potential conflict of interest must be sought well beyond the mere declaration of conflict of interest by authors or direct funding of studies. Conversely, individual monocentric studies focusing on HCQ have multiplied and are associated with the success of HCQ.

These three elements (potential conflict of interest, private data computing company, and multi- or monocentric studies) can predict the outcome of the meta-analysis based on the choices that will be made to retain certain studies. For the first time to our knowledge, number of studies were conducted ignoring the very basis of inclusions at the medical level. i.e. clinical signs found in this disease (not yet reported in acute respiratory infection in general) such as anosmia and ageusia, and pulmonary embolisms are not in the clinical diagnostic criteria. On the other hand, some studies have been published without even having confirmatory biological tests (27,28), which for infectious diseases is a regression that has no

equivalent. Finally, in most cases the evaluation of treatments in the different stages of the disease should correspond to different therapeutic options, and this is often not evaluated.

All in all, this crisis highlighted very different therapeutic evaluation strategies. The considerable weight of the pharmaceutical industry on the results of therapeutic trials is clear and causal (1,5,8). It seemed clear to us that the credibility of medical research on therapeutic trials must take these elements into account given the considerable importance of the financial stakes (1,5). It would be naïve to rely on goodwill to fight against the power of financial interests and against the biases linked to these interests. In practice, RCTs have been set up to avoid biases, but given their massive use by the pharmaceutical industry, from our point of view, RCTs, particularly multi-centric RCTs, where no investigator can have access to all the raw data before analysis (1,5), favor biases by favoring their manipulation by the pharmaceutical industry as illustrated by Husserl in this quotation “Methods are the clothes of ideas”. Meta-analyses allow small studies to be analyzed and multicenter studies should report results by center so that investigators can ensure validity and to avoid the Simpson effect (29). The DisCoVeRy megatrial (30), recruiting patients in 32 French sites, and included in the WHO Solidarity megatrial recruiting patients in 405 hospitals in 30 countries (31) did not stratify by region of inclusion. Since the number of patients included per center (low number of inclusions in some regions (30)) and the effect for each center was not reported and may have been highly variable, both these megatrials are likely to be biased by the Simpson's effect.

Most of the criteria identified in this work (Table 1) are new, not part of usual quality checklists (STROBE, CONSORT or PRISMA – see Supplementary Table 2) and may be useful for future critical review. This comeback to independent clinical and microbiological expertise is the best lesson to be learned from the global scandal we have witnessed, for the greatest benefit of patients.

Funding

This work was funded by ANR-15-CE36-0004-01 and by ANR "Investissements d'avenir", Méditerranée infection 10-IAHU-03, and was also supported by Région Provence-Alpes-Côte d'Azur. This work had received financial support from the Mediterranean Infection Foundation.

Declaration of competing interest

The authors declare no competing interests. Funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. Our group used widely available generic drugs distributed by many pharmaceutical companies.

Acknowledgments

This manuscript has been edited by a native English speaker. The authors thank Yanis ROUSSEL for helpful discussions.

REFERENCES

1. Godlee F. Covid-19 : The lost lessons of Tamiflu. BMJ 2020;371:m4701 doi: https://doi.org/10.1136/bmj.m4701
2. 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 [published online ahead of print, 2020 May 22] [retracted in: Lancet. 2020 Jun 5;:null]. Lancet. 2020a;S0140-6736(20)31180-6. doi:10.1016/S0140-6736(20)31180-6
3. 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-e102
4. WHO Solidarity Trial Consortium, Pan H, Peto R, et al. Repurposed Antiviral Drugs for Covid-19 - Interim WHO Solidarity Trial Results [published online ahead of print, 2020 Dec 2]. N Engl J Med. 2020;10.1056/NEJMoa2023184. doi:10.1056/NEJMoa2023184
5. Abbasi K. Covid-19: politicisation, "corruption," and suppression of science. BMJ 2020;371:m4425 doi: https://doi.org/10.1136/bmj.m4425
6. Grein J, Ohmagari N, Shin D, et al. Compassionate Use of Remdesivir for Patients with Severe Covid-19. N Engl J Med. 2020;382(24):2327-2336. doi:10.1056/NEJMoa2007016
7. Lagier JC, Million M, Gautret P, et al. Outcomes of 3,737 COVID-19 patients treated with hydroxychloroquine/azithromycin and other regimens in Marseille, France: A retrospective analysis. Travel Med Infect Dis. 2020;36:101791. doi:10.1016/j.tmaid.2020.101791
8. Bhandari M, Busse JW, Jackowski D, Montori VM, Schünemann H, Sprague S, et al. Association between industry funding and statistically significant pro-industry

findings in medical and surgical randomized trials. CMAJ. 2004 Feb 17;170(4):477-80.

9. Roussel Y, Raoult D. Influence of conflicts of interest on public positions in the COVID-19 era, the case of Gilead Sciences. New Microbes New Infect. 2020;38:100710. Published 2020 Jun 6. doi:10.1016/j.nmni.2020.100710
10. Fiolet T, Guihur A, Rebeaud ME, Mulot M, Peiffer-Smadja N, Mahamat-Saleh Y. Effect of hydroxychloroquine with or without azithromycin on the mortality of coronavirus disease 2019 (COVID-19) patients: a systematic review and meta-analysis [published online ahead of print, 2020 Aug 26]. Clin Microbiol Infect. 2020;S1198-743X(20)30505-X. doi:10.1016/j.cmi.2020.08.022
11. 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. Published 2020a Jun 6. doi:10.1016/j.nmni.2020.100709
12. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med. 2000;342(25):1887-1892. doi:10.1056/NEJM200006223422507
13. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. Cochrane Database Syst Rev. 2014;(4):MR000034. Published 2014 Apr 29. doi:10.1002/14651858.MR000034.pub2
14. Raoult D. Lancet gate: a matter of fact or a matter of concern. New Microbes New Infect. 2020;38:100758. doi:10.1016/j.nmni.2020.100758

15. Raoult D. Rational for meta-analysis and randomized treatment: the COVID-19 example [published online ahead of print, 2020 Oct 21]. Clin Microbiol Infect. 2020a;S1198-743X(20)30643-1. doi:10.1016/j.cmi.2020.10.012
16. Greenacre, M.J. and Blasius, J. (2006). Multiple correspondence analysis and related methods. Chapman & Hall/CRC.
17. Lê, S., Josse, J. & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software. 25(1). pp. 1-18.
18. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. 2009. John Wiley & Sons, Ltd. 2009. https://doi.org/10.1002/9780470743386.
19. Million M, Chaudet H, Raoult D. Hydroxychloroquine Failure: The End does not justify the Means [published online ahead of print, 2020 Aug 6]. Clin Infect Dis. 2020c;ciaa1117. doi:10.1093/cid/ciaa1117
20. Dubée V, Roy PM, Vielle B, Parot-Schinkel E, Blanchet O, Darsonval A, et al. for the HYCOVID study group A placebo-controlled double blind trial of hydroxychloroquine in mild-to-moderate COVID-19 medRxiv 2020.10.19.20214940; doi: https://doi.org/10.1101/2020.10.19.20214940
21. Khan AR, Khan S, Zimmerman V, Baddour LM, Tleyjeh IM. Quality and strength of evidence of the Infectious Diseases Society of America clinical practice guidelines. Clin Infect Dis. 2010;51(10):1147-1156. doi:10.1086/656735
22. Kasenda B, von Elm E, You JJ, Blümle A, Tomonaga Y, Saccilotto R, et al. Agreements between Industry and Academia on Publication Rights: A Retrospective Study of Protocols and Publications of Randomized Clinical Trials. PLoS Med. 2016 Jun;13(6):e1002046.

352 23. Roussel Y, Million M, Chabrière E, Lagier JC, Raoult D. Be careful with Big Data:
353 Re-analysis of Patient Characteristics and Outcomes of 11,721 Patients with
354 COVID19 Hospitalized Across the United States [published online ahead of print,
355 2020 Oct 22]. Clin Infect Dis. 2020;ciaa1618. doi:10.1093/cid/ciaa1618
356 24. Fried MW, Crawford JM, Mospan AR, et al. Patient Characteristics and Outcomes of
357 11,721 Patients with COVID19 Hospitalized Across the United States [published
358 online ahead of print, 2020 Aug 28]. Clin Infect Dis. 2020;ciaa1268.
359 doi:10.1093/cid/ciaa1268
360 25. Flisiak R, Zarębska-Michaluk D, Berkan-Kawińska A, Tudrujek-Zdunek M, Rogalska
361 M, Piekarska A, et al. Remdesivir-based therapy improved recovery of patients with
362 COVID-19 in the SARSTer multicentre, real-world study. medRxiv
363 2020.10.30.20215301; doi: https://doi.org/10.1101/2020.10.30.20215301
364 26. Singh S, Khan A, Chowdhry M, Chatterjee A. Outcomes of Hydroxychloroquine
365 Treatment Among Hospitalized COVID-19 Patients in the United States- Real-World
366 Evidence From a Federated Electronic Medical Record Network. medRxiv.
367 2020.05.12;20099028; doi: 10.1101/2020.05.12.20099028
368 27. Boulware DR, Pullen MF, Bangdiwala AS, et al. A Randomized Trial of
369 Hydroxychloroquine as Postexposure Prophylaxis for Covid-19. N Engl J Med.
370 2020;383(6):517-525. doi:10.1056/NEJMoa2016638
371 28. Skipper CP, Pastick KA, Engen NW, et al. Hydroxychloroquine in Non hospitalized
372 Adults With Early COVID-19: A Randomized Trial [published online ahead of print,
373 2020 Jul 16]. Ann Intern Med. 2020;M20-4207. doi:10.7326/M20-4207
374 29. Julious SA, Mullee MA. Confounding and Simpson's paradox. BMJ.
375 1994;309(6967):1480-1481. doi:10.1136/bmj.309.6967.1480

30. Ader F, Peiffer-Smadja N, Poissy J, Bouscambert-Duchamp M, Belhadi D, Diallo A, et al. Antiviral drugs in hospitalized patients with COVID-19 - the DisCoVeRy trial. doi: https://doi.org/10.1101/2021.01.08.20248149

31. WHO Solidarity Trial Consortium, Pan H, Peto R, et al. Repurposed Antiviral Drugs for Covid-19 - Interim WHO Solidarity Trial Results [published online ahead of print, 2020 Dec 2]. N Engl J Med. 2020;NEJMoa2023184. doi:10.1056/NEJMoa2023184

382 **Table 1. Twenty quality criteria proposed to assess future clinical therapeutic studies in infectious diseases**

Conflict of interest	
1. Potential conflict of interest	• Work funded by a company with a conflict of interest • At least one author compensated by a company with a conflict of interest (received fee) declared or not by this author • A private data computing company (see definition below)
2. Private data computing company	• A for-profit company collecting, aggregating, and computing data in "Big data" studies (with frequent unclear funding)
3. Undeclared funding and conflict of interest	• Funding of the work not mentioned. • A conflict of interest not declared by an author but found through transparency websites (dollarfordocs, eurosfordocs) or other means (through internet investigations). • An indirect funding through a shell company by a company with a conflict of interest.
Centers and doctors' identification	

4. Known centers and doctors	• Recruiting centers and investigating doctors who directly take care of patients in the clinical unit (at least one by center) are identified.
Clinical expertise	
5. Patients without confirmation of diagnosis by a microbiological test are excluded	• A patient is considered infected if the infection is confirmed in the laboratory (PCR, blood culture, serology). Clinical definition not sufficient.
6. Detailed standard of care (SoC)	• The standard care of patients with or without experimental treatment is reported (including criteria for admission, vital monitoring, initial check-up, anticoagulants, oxygenotherapy…). This standard care is likely to influence outcome in a greater extent than the experimental treatment itself. • Change over time of SoC should be reported.
7. Detailed therapeutic protocol	• With at least most frequent contraindications assessed, dosage, and duration
8. Treatment not toxic	• Dosage is usual (not overdosed) and known to be well tolerated, treatment is effectively not used in patients with contra-indications
9. Treatment monitoring	• Side effects are reported. • Critical (serious) side effects are reported (death, organ failure). If any death were related to experimental treatment, it should be mentioned.

	• Interruption of experimental treatment because of side effect. • Side effects are not artificially mixed. For instance, mild (diarrhea) and severe (renal failure) side effects should be analyzed separately.
10. Untreated group is not treated	• Group without experimental treatment does not receive another specific treatment, or this one may be analyzed.
11. At least one main author is a clinical expert-in-the-field	• At least one author directly takes care of patients and is specialized in this care (for a respiratory viral disease, this includes an infectious disease specialist, an internal medicine specialist or a pneumologist).
12. Confounding role of previous health status (at least age) is ruled out	• Previous health status should be assessed (at least age) and controlled for. This could be achieved using comorbidity score (Combined Charlson score). Previous health status should not be different at baseline and/or approaches should be used to control it (matching, multivariate analyses). Authors should provide evidence that this confounding has been controlled (for instance, age and comorbidities after matching are shown and not different).
13. Confounding role of severity (at least vital parameters) is ruled out	• Initial severity should be assessed (at least vital parameters) and controlled for. This could be achieved using severity score (NEWS score). Initial severity

	should not be different at baseline and/or approaches should be used to control it (matching, multivariate analyses). Authors should provide evidence that this confounding has been controlled (for instance, initial severity after matching is shown and not different).
14. Different stages of the disease are not mixed	• Different treatment could be associated with different effect at different stages of the disease. Results should be stratified by stage of the disease (for instance outpatient, non-severe or severe inpatient or early versus late) according to previous knowledge of the disease.
Methodology	
15. Identification of observational and interventional studies	• Observational studies may be a case / control (dead / alive) or exposed / unexposed (treated / untreated). In this case, covariables are adjusted by matching, propensity score approaches or multivariate analysis. • Interventional studies may be randomized studies, and theoretically the patient's situation is comparable.

15.1. Among observational studies, identification of electronic ("Big data") versus clinical studies	• Studies should be classified as 'electronic' or 'big data' studies when conducted on electronic medical records extracted by public-health specialists and epidemiologists who did not care for COVID-19 patients themselves. • Conversely, studies should be classified as 'clinical studies' when the authors are physicians who cared for COVID-19 patients themselves.
15.2. Among interventional studies, identification of megatrials	• Large-scale interventional trials including several centers (usually > 10).
16. Identification of monocentric and multicentric studies, and center effect is evaluated in multicentric studies.	• Multicentric observational (including Big data studies) and interventional (including megatrials) studies are sensitive to Simpson's paradox effect. In multicentric studies, adjusted results should be reported for each center, using forest plot. • Summary effect calculation should use random effects models since experimental conditions are inevitably different among different centers recruiting human patients. Indeed, in contrast to mouse lines in environmentally-controlled cages (where fixed effect model could be used), standard of care and human

	populations are always genetically, environmentally, and behaviorally different between centers.
17. Objective is objective and invariant	• The main outcome is objective, independent of human subjectivity and context (death, viral load) and should not change during study.
18. Number of events and total sample size mentioned for each group in each center	• This may improve verifiability.
Interpretation and conclusions	
19. Conclusions do not neglect a 25% difference in risk of death (in whole population or any subgroup)	• An observation of a relevant change in mortality risk in the whole population or in any secondary analysis (subgroup, *etc*…) should be reported and discussed, regardless of significance. In this case, the authors should calculate the number of participants that would be needed to significantly confirm the effect observed in the relevant group and, if data from similar studies are available, conduct a meta-analysis to eliminate a lack of statistical power.
20. Unexpected findings may be reported	• When a non-prespecified effect is observed and clinically relevant, it should be analyzed. For instance, a specific effect in a specific subgroup.
Data sharing	

Additional criterion. Data should be shared within 12 months	• Data sharing may improve verifiability.

383

384 **Table 2. Predictive value of each criterion for the issue of clinical assays for HCQ**

	Con HCQ n (row %)	**Pro HCQ** n (row %)	p-value*
Potential Conflict of interest (n=15)	11 (73.3)	4 (26.7)	0.001
No potential conflict of interest (n=41)	9 (21.9)	32 (78.1)	
Detailed therapeutic protocol (n=25)	4 (16.0)	21 (84.0)	0.011
Absence of detailed therapeutic protocol (n=31)	16 (51.6)	15 (48.4)	
Toxic treatment (n=4)	4 (100.0)	0 (0.0)	0.013
Non-toxic treatment (n=52)	16 (30.8)	36 (69.2)	
Known centers and doctors (n=41)	11 (26.8)	30 (73.2)	0.030
Unknown centers and doctors (n=15)	9 (60.0)	6 (40.0)	
Private data computing company (n=3)	3 (100.0)	0 (0.0)	0.041
No private data computing company (n=53)	17 (32.1)	36 (67.9)	
Declared Funding COI (n=47)	14 (29.8)	33 (70.2)	0.056
Undeclared funding COI (n=9)	6 (66.7)	3 (33.3)	
Observational (n=47)	14 (29.8)	33 (70.2)	0.056
Not observational (n=9)	6 (66.7)	3 (33.3)	
Role of severity ruled out (n=32)	8 (25.0)	24 (75.0)	0.090
Role of severity not ruled out (n=24)	12 (50.0)	12 (50.0)	
Big data (n=22)	11 (50.0)	11 (50.0)	0.092
No big data (n=34)	9 (26.5)	25 (73.5)	
Number of events and total mentioned for each group (n=40)	17 (42.5)	23 (57.5)	0.13
Number of events and total not mentioned for each group (n=16)	3 (18.8)	13 (81.3)	
Standard care reported (n=9)	1 (11.1)	8 (88.9)	0.136
Standard care not reported (n=47)	19 (40.4)	28 (59.6)	
Treatment monitoring (n=19)	4 (21.1)	15 (78.9)	0.143
Absence of treatment monitoring (n=37)	16 (43.2)	21 (56.8)	
Lab confirmed diagnosis (n=42)	13 (30.9)	29 (69.1)	0.198
No lab confirmed diagnosis (n=14)	7 (50.0)	7 (50.0)	
Monocentric (n=18)	4 (22.2)	14 (77.8)	0.2326
Multicentric (n=38)	16 (42.1)	22 (57.9)	
One author expert in the field (n=29)	7 (24.1)	22 (75.9)	0.345

No author expert in the field (n=20)	8 (40.0)	12 (60.0)	
Different stages mixed (n=21)	9 (42.9)	12 (57.1)	0.405
Different stages not mixed (n=35)	11 (31.4)	24 (68.6)	
Unexpected results reported (n=48)	16 (33.3)	32 (66.7)	0.437
Unexpected results not reported (n=8)	4 (50.0)	4 (50.0)	
Conclusions neglect a 25% decrease in mortality (n=12)	3 (25.0)	9 (75.0)	0.506
Conclusions do not neglect a 25% decrease in mortality (n=44)	17 (38.6)	27 (61.4)	
Megatrial (n=6)	3 (50.0)	3 (50.0)	0.6553
Not a megatrial (n=50)	17 (34.0)	33 (66.0)	
Role of previous health status ruled out (n=45)	17 (37.8)	28 (62.2)	0.728
Role of previous health status not ruled out (n=11)	3 (27.3)	8 (72.7)	
Untreated group with specific treatment (n=2)	1 (50.0)	1 (50.0)	1.000
Untreated group without specific treatment (n=54)	19 (35.2)	35 (64.8)	
Death as a clear outcome (n=47)	17 (36.2)	30 (63.8)	1.000
Death not a clear outcome (n=9)	3 (33.3)	6 (66.7)	

*: Two-sided p-value (Fisher's exact test). n = 56 studies

385

386 **Table 3. Observational studies with a detailed therapeutic protocol without potential conflict of interest**

Study name	INPATIENTS/OUTPATIENTS/BOTH	Country	Pro Con HCQ
Alberici, Kidney International, 2020	BOTH	Italy	Pro
Arshad, Int J Infect Dis, 2020	INPATIENTS	USA	Pro
Ashraf, MedRxiv, 2020	INPATIENTS	Iran	Pro
Ayerbe, Intern Med Emerg, 2020	INPATIENTS	Spain	Pro
Catteau, Int J Antimicrob Agents, 2020	INPATIENTS	Belgique	Pro
Davido, Int J Antimicrob Agents, 2020	INPATIENTS	France	Pro
Derwand, Int J Antimicrob Agents, 2020	OUTPATIENTS	USA	Pro
Di Castelnuovo, Eur J Intern Med, 2020	INPATIENTS	Italy	Pro
Guerin, Asian J Med Health, 2020	OUTPATIENTS	France	Pro
Lagier, Trav Med Infect Dis, 2020	BOTH	France	Pro
Lauriola, Clinical Transl Sci, 2020	INPATIENTS	Italy	Pro
Lecronier, Critical Care, 2020	ICU	France	Pro
Membrillo de Novales, Preprints, 2020	INPATIENTS	Spain	Pro
Mikami, J Gen Intern Med, 2020	INPATIENTS	USA	Pro
Nachega, Am J Trop Med Hyg, 2020	INPATIENTS	Congo	Pro
Paccoud, Clin Infect Dis, 2020	INPATIENTS	France	Pro

Sulaiman, MedRxiv, 2020	OUTPATIENTS	Saudi Arabia	Pro
Yu, Sci Chi Life Sci, 2020	ICU	China	Pro

387 All these 18 studies were in favor of a HCQ efficacy (100% predictive value). ICU: Intensive care unit.

Figure legends

Figure 1. Multiple Correspondence Analysis (MCA) including all the characteristics of 56 studies (n=56)

Unsupervised approaches (such as MCA for qualitative variables) allow graphical representation without a priori that takes together the variables and observations (biplot). Studies and their characteristics can be identified and analyzed according to an additional variable (such as direction of effect of studies pro/Con). Direction of effect of each study is indicated in green (Pro) and red (Con). Ellipses cluster 90% of the points belonging to the two groups chosen. *For these studies, it could not be easily determined whether at least one main author is a clinical expert-in-the-field who directly take care of Covid-19 patients (see Table 1).

Figure 2. HCQ meta-analysis according to potential conflict of interest

95% CI: 95% confidence interval. Random effects model.