HCQ is effective for COVID-19 when used early: analysis of 139 studies

Covid Analysis, October 20, 2020 (Version 14, November 12, 2020)
https://hcqmeta.com/

- HCQ is effective for COVID-19. The probability that an ineffective treatment generated results as positive as the 139 studies to date is estimated to be 1 in 17 billion ($p = 0.00000000006$).

- Early treatment is most successful, with 100% of studies reporting a positive effect and an estimated reduction of 64% in the effect measured (death, hospitalization, etc.) using a random effects meta-analysis, RR 0.36 [0.29-0.46].

- 100% of Randomized Controlled Trials (RCTs) for early, PrEP, or PEP treatment report positive effects, the probability of this happening for an ineffective treatment is 0.002.

- There is evidence of bias towards publishing negative results. 89% of prospective studies report positive effects, and only 73% of retrospective studies do.

- Significantly more studies in North America report negative results compared to the rest of the world, $p = 0.003$. 

---

A

Early treatment

All studies

Lower Risk

Increased Risk

B

% Increased Risk % Lower Risk

---

Probability results from ineffective treatment 1 in 1 thousand
1 in 10 thousand
1 in 1 million

Jul 3 $p=0.001$

Aug 21 $p=0.0001$

Oct 31 $p=0.0000001$

Probability results from ineffective treatment 1 in 1 million
1 in 100 thousand

Sep 9 $p=0.00001$

Nov 12 $p=hcqmeta.com$
Figure 1. A. Scatter plot showing the distribution of effects reported in early treatment studies and in all studies (the vertical lines and shaded boxes show the median and interquartile range). Early treatment is more effective. B and C. Study results ordered by date, with the line showing the probability that the observed frequency of positive results occurred due to random chance from an ineffective treatment.

Introduction

We analyze all significant studies concerning the use of HCQ (or CQ) for COVID-19 (Appendix 1), showing the effect size and associated $p$ value for results comparing to a control group. Typical meta analyses involve subjective selection criteria and bias evaluation, requiring an understanding of the criteria and the accuracy of the evaluations. However, the volume of studies presents an opportunity for a simple and transparent analysis aimed at detecting efficacy.

If treatment was not effective, the observed effects would be randomly distributed (or more likely to be negative if treatment is harmful). We can compute the probability that the observed percentage of positive results (or higher) could occur due to chance with an ineffective treatment (the probability of $\geq k$ heads in $n$ coin tosses, or the one-sided sign test / binomial test). Analysis of publication bias is important and adjustments may be needed if there is a bias toward publishing positive results. For HCQ, we find evidence of a bias toward publishing negative results.

Figure 2. Treatment stages.

Figure 2 shows stages of possible treatment for COVID-19. Pre-Exposure Prophylaxis (PrEP) refers to regularly taking medication before being infected, in order to prevent or minimize infection. In Post-Exposure Prophylaxis (PEP), medication is taken after exposure but before symptoms appear.
Early Treatment refers to treatment immediately or soon after symptoms appear, while Late Treatment refers to more delayed treatment.

Results

Figure 3, Figure 4 and Table 1 show results by treatment stage, and Figure 5 shows a forest plot for a random effects meta-analysis of all studies. Analysis excluding studies with major issues is in Appendix 2.

Early treatment. 100% of early treatment studies report a positive effect, with an estimated reduction of 64% in the effect measured (death, hospitalization, etc.) from the random effects meta-analysis, RR 0.36 [0.29-0.46].

Late treatment. Late treatment studies are mixed, with 70% showing positive effects, and an estimated reduction of 23% in the random effects meta-analysis. Negative studies mostly fall into the following categories: they show evidence of significant unadjusted confounding, including confounding by indication; usage is extremely late; or they use an excessively high dosage.

Pre-Exposure Prophylaxis. 78% of PrEP studies show positive effects, with an estimated reduction of 46% in the random effects meta-analysis. Negative studies are all studies of systemic autoimmune disease patients which either do not adjust for the different baseline risk of these patients at all, or do not adjust for the highly variable risk within these patients.

Post-Exposure Prophylaxis. 100% of PEP studies report positive effects, with an estimated reduction of 39% in the random effects meta-analysis.

<table>
<thead>
<tr>
<th>Treatment stage</th>
<th>Number of studies reporting positive results</th>
<th>Total number of studies</th>
<th>Percentage of studies reporting positive results</th>
<th>Probability of an equal or greater percentage of positive results from an ineffective treatment</th>
<th>Random effects meta-analysis results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early treatment</td>
<td>22</td>
<td>22</td>
<td>100%</td>
<td>0.000000024 1 in 4 million</td>
<td>64% improvement RR 0.36 [0.29-0.46]</td>
</tr>
<tr>
<td>Late treatment</td>
<td>61</td>
<td>87</td>
<td>70.1%</td>
<td>0.00011 1 in 9 thousand</td>
<td>23% improvement RR 0.77 [0.70-0.85]</td>
</tr>
<tr>
<td>Pre-Exposure Prophylaxis</td>
<td>21</td>
<td>27</td>
<td>77.8%</td>
<td>0.003 1 in 338</td>
<td>46% improvement RR 0.54 [0.40-0.74]</td>
</tr>
<tr>
<td>Post-Exposure Prophylaxis</td>
<td>5</td>
<td>5</td>
<td>100%</td>
<td>0.031 1 in 32</td>
<td>39% improvement RR 0.61 [0.41-0.92]</td>
</tr>
<tr>
<td>All studies</td>
<td>107</td>
<td>139</td>
<td>77.0%</td>
<td>0.00000000006 1 in 17 billion</td>
<td>32% improvement RR 0.68 [0.62-0.74]</td>
</tr>
</tbody>
</table>

Table 1. Results by treatment stage. 2 studies report results for a subset with early treatment, these are not included in the overall results.
Figure 3. Results by treatment stage.
Figure 4. Results by treatment stage. Study results are ordered by date, with the line showing the probability that the observed frequency of positive results occurred due to random chance from an ineffective treatment.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Score</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gautret</td>
<td>0.34</td>
<td>[0.17-0.68]</td>
</tr>
<tr>
<td>Huang</td>
<td>0.09</td>
<td>[0.01-1.44]</td>
</tr>
<tr>
<td>Esper</td>
<td>0.36</td>
<td>[0.15-0.87]</td>
</tr>
<tr>
<td>Ashraf</td>
<td>0.32</td>
<td>[0.16-1.10]</td>
</tr>
<tr>
<td>Huang (ES)</td>
<td>0.41</td>
<td>[0.26-0.66]</td>
</tr>
<tr>
<td>Guérin</td>
<td>0.57</td>
<td>[0.02-13.36]</td>
</tr>
<tr>
<td>Chen</td>
<td>0.28</td>
<td>[0.10-0.82]</td>
</tr>
<tr>
<td>Lagier</td>
<td>0.41</td>
<td>[0.27-0.62]</td>
</tr>
<tr>
<td>Denwand</td>
<td>0.21</td>
<td>[0.03-1.47]</td>
</tr>
<tr>
<td>Minjà</td>
<td>0.75</td>
<td>[0.32-1.78]</td>
</tr>
<tr>
<td>Skipper</td>
<td>0.48</td>
<td>[0.17-1.39]</td>
</tr>
<tr>
<td>Hong</td>
<td>0.35</td>
<td>[0.13-0.72]</td>
</tr>
<tr>
<td>Ly</td>
<td>0.46</td>
<td>[0.27-0.79]</td>
</tr>
<tr>
<td>Ip</td>
<td>0.54</td>
<td>[0.30-0.95]</td>
</tr>
<tr>
<td>Heras</td>
<td>0.08</td>
<td>[0.00-0.35]</td>
</tr>
<tr>
<td>Kirenga</td>
<td>0.74</td>
<td>[0.37-1.48]</td>
</tr>
<tr>
<td>Sulaiman</td>
<td>0.36</td>
<td>[0.17-0.80]</td>
</tr>
<tr>
<td>Guisado-Vasco (ES)</td>
<td>0.12</td>
<td>[0.03-0.45]</td>
</tr>
<tr>
<td>Fonseca</td>
<td>0.36</td>
<td>[0.20-0.67]</td>
</tr>
<tr>
<td>Cadegiani</td>
<td>0.17</td>
<td>[0.01-3.56]</td>
</tr>
<tr>
<td>Simova</td>
<td>0.03</td>
<td>[0.00-0.65]</td>
</tr>
</tbody>
</table>

**Early**: 0.36 [0.29-0.46]

64% improvement
**Randomized Controlled Trials (RCTs)**

RCTs are very valuable and minimize potential bias, however they are neither necessary or sufficient. [Concato] find that well-designed observational studies do not systematically overestimate the magnitude of the effects of treatment compared to RCTs. [Anglemyer] summarized reviews comparing RCTs to observational studies and found little evidence for significant differences in effect estimates. [Lee] shows that only 14% of the guidelines of the
Infectious Diseases Society of America were based on RCTs. Limitations in an RCT can easily outweigh the benefits, for example excessive dosages, excessive treatment delays, or Internet survey bias could easily have a greater effect on results. Ethical issues may prevent running RCTs for known effective treatments. For more on the problems with RCTs see [Deaton, Nichol]. Results restricted to RCTs are shown in Figure 6 and Table 2. Even with the small number of RCTs to date, there is a strong indication of efficacy. When excluding late treatment, 100% of RCTs to date report positive results.

**Figure 6.** Randomized Controlled Trials. The distribution of results for RCTs is similar to the distribution for all other studies.

**Figure 7.** RCTs excluding late treatment.
<table>
<thead>
<tr>
<th>Treatment time</th>
<th>Number of studies reporting positive results</th>
<th>Total number of studies</th>
<th>Percentage of studies reporting positive results</th>
<th>Probability of an equal or greater percentage of positive results from an ineffective treatment</th>
<th>Random effects meta-analysis results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized Controlled Trials</td>
<td>18</td>
<td>22</td>
<td>81.8%</td>
<td>0.0022 1 in 460</td>
<td>11% improvement RR 0.89 [0.76-1.04]</td>
</tr>
<tr>
<td>Randomized Controlled Trials (excluding late treatment)</td>
<td>9</td>
<td>9</td>
<td>100%</td>
<td>0.002 1 in 512</td>
<td>30% improvement RR 0.70 [0.53-0.93]</td>
</tr>
</tbody>
</table>

**Table 2. Summary of RCT results.**

**Discussion**

**Publication bias.** Publishing is often biased towards positive results, which we would need to adjust for when analyzing the percentage of positive results. Studies that require less effort are considered to be more susceptible to publication bias. Prospective trials that involve significant effort are likely to be published regardless of the result, while retrospective studies are more likely to exhibit bias. For example, researchers may perform preliminary analysis with minimal effort and the results may influence their decision to continue. Retrospective studies also provide more opportunities for the specifics of data extraction and adjustments to influence results.

For HCQ, 88.9% of prospective studies report positive effects, compared to 72.8% of retrospective studies, indicating a bias toward publishing negative results. Figure 8 shows a scatter plot of results for prospective and retrospective studies.

Figure 9 shows the results by region of the world, for all regions that have > 5 studies. Studies from North America are significantly more likely to report negative results than studies from the rest of the world combined, two-tailed z test -2.99, p = 0.003. [Berry](#) performed an independent analysis which also showed bias toward negative results for US-based research.

![Figure 8. Prospective vs. retrospective studies.](#)

We also note a bias towards publishing negative results by certain journals and press organizations, with scientists reporting difficulty publishing positive results [Boulware, Menegussa]. Although 107 studies show positive results, The New York Times, for example, has only written articles for studies that claim HCQ is not effective [The New York Times, The New York Times (B), The New York Times (C)]. As of September 10, 2020, The New York Times still claims that there is clear evidence that HCQ is not effective for COVID-19 [The New York Times (D)]. As of October 9, 2020, the United States National Institutes of Health recommends against HCQ for both hospitalized and non-hospitalized patients [United States National Institutes of Health].

**Treatment details.** We focus here on the question of whether HCQ is effective or not for COVID-19. Significant differences exist based on treatment stage, with early treatment showing the greatest effectiveness. 100% of early treatment studies report a positive effect, with an estimated reduction of 64% in the effect measured (death, hospitalization, etc.) in the random effects meta-analysis, RR 0.36 [0.29-0.46]. Many factors are likely to influence the degree of effectiveness, including the dosing regimen, concomitant medications such as zinc or azithromycin, precise treatment delay, the initial viral load of patients, and current patient conditions.
Conclusion

HCQ is an effective treatment for COVID-19. The probability that an ineffective treatment generated results as positive as the 139 studies to date is estimated to be 1 in 17 billion ($p = 0.0000000006$).

Revisions

This paper is data driven, all graphs and numbers are dynamically generated. We will update the paper as new studies are released or with any corrections.

10/21: We added studies [Dubee, Martinez-Lopez, Solh]. We received a report that the United States National Institutes of Health is recommending against HCQ for hospitalized and non-hospitalized patients as of October 9, and we added a reference.

10/22: We added [Anglemyer, Ñamendys-Silva]. We updated the discussion of [Axfors] for the second version of this study. We added a table summarizing RCT results.

10/23: We added [Komissarov, Lano]. The second version of the preprint for [Komissarov] includes a comparison with the control group (not reported in the first version). We updated [Lyngbakken] to use the mortality result in the recent journal version of the paper (not reported in the preprint).

10/26: We added [Coll, Goenka, Synolaki].

10/28: We added [Arleo, Choi].

10/30: We added [Berenguer, Faico-Filho].

10/31: We added [Fonseca, Frontera, Tehrani].

11/1: We added [Trullàs].

11/4: We added [Behera, Cadegiani].

11/8: We added [Dhibar].

11/9: We added [Self].

11/10: We added [Mathai].

11/12: We added [Simova, Simova (B)].

References


13. Altman (B) et al., BMJ, doi:10.1136/bmj.d2090, How to obtain the confidence interval from a P value, https://www.bmj.com/content/343/bmj.d2090.


48. Chen (B) et al., medRxiv, doi:10.1101/2020.07.08.20148841v1, A Multicenter, randomized, open-label, controlled trial to evaluate the efficacy and tolerability of hydroxychloroquine and a retrospective study in adult patients with mild to moderate Coronavirus disease 2019 (COVID-19), https://www.medrxiv.org/content/10.1101/2020.07.08.20148841v1.

49. Chen (C) et al., medRxiv, doi:10.1101/2020.07.08.20148841v1, A Multicenter, randomized, open-label, controlled trial to evaluate the efficacy and tolerability of hydroxychloroquine and a retrospective study in adult patients with mild to moderate Coronavirus disease 2019 (COVID-19).


64. DISCOVERY Trial, DISCOVERY Trial Preliminary Results, https://twitter.com/raoult_didier/status/131350924167529472.
65. **Dr. Goldin**, *Summary of HCQ usage in India from an MD in India*, https://www.facebook.com/groups/hydroxychloroquine/permalink/2367454293560817/.


81. **Frontera** et al., Research Square, doi:10.21203/rs.3.rs-94509/v1, *Treatment with Zinc is Associated with Reduced In-Hospital Mortality Among COVID-19 Patients: A Multi-Center Cohort Study*, https://www.researchsquare.com/article/rs-94509/v1.


97. Guisado-Vasco (B), Clinical characteristics and outcomes among hospitalized adults with severe COVID-19 admitted to a tertiary medical center and receiving antiviral, antimalarials, glucocorticoids, or immunomodulation with tocilizumab or cyclosporine: A retrospective observational study (COQUIMA cohort), https://www.sciencedirect.com/science/article/pii/S2589537020303357.


157. NPR News, Senegal pledges a bed for every coronavirus patient, https://wfuv.org/content/senegal-p...E2%80%94and-their-contacts-too.


182. Shabrawishi et al., medRxiv, doi:10.1101/2020.05.08.20095679, Negative nasopharyngeal SARS-CoV-2 PCR conversion in response to different therapeutic interventions, https://www.medrxiv.org/content/10.1101/2020.05.08.20095679v1.


200. The BL, Russia supports the use of hydroxychloroquine, the drug to treat the CCP Virus suggested by Trump, https://thebl.com/world-news/russi..oroquine-drug-ccp-virus-trump.html.


203. The Indian Express, Vadodara administration drive: HCQ helping in containing Covid-19 cases, say docs as analysis begins, https://indianexpress.com/article/..y-docs-as-analysis-begins-6486049/.


Appendix 1. Methods and Study Results

We performed ongoing searches of PubMed, medRxiv, ClinicalTrials.gov, The Cochrane Library, Google Scholar, Collabovid, the reference lists of other studies and meta-analyses, and submissions to the site c19study.com, which regularly receives submissions of both positive and negative studies upon publication. Search terms were hydroxychloroquine or chloroquine and COVID-19 or SARS-CoV-2, or simply hydroxychloroquine or chloroquine. All studies regarding the use of HCQ or CQ for COVID-19 that report an effect compared to a control group are included in the main analysis. This is a living analysis and will be updated regularly.

We extracted effect sizes and associated data from all studies. If studies report multiple kinds of effects then the most serious outcome is used in calculations for that study. For example, if effects for mortality and cases are both reported, the effect for mortality is used, this may be different to the effect that a study focused on. If mortality results are given at multiple times, we used the latest time. Mortality alone is preferred over combined outcomes. Outcomes with zero events in both arms were not used. Clinical outcome is considered more important than PCR testing status. For PCR results reported at multiple times, preference is given to results mid-recovery (after most or all patients have recovered there is no room for an effective treatment to do better). When results provide an odds ratio, we computed the relative risk when possible, or converted to a relative risk according to [Zhang]. Reported confidence intervals and p-values were used when available, using adjusted values when provided. When needed, conversion between reported p-values and confidence intervals followed [Altman, Altman (B)], and Fisher's exact test was used to calculate p-values for event data. If a study separated HCQ and HCQ+AZ we used the combined results were possible, or the results for the larger group. Results are all expressed with RR < 1.0 suggesting effectiveness. Most results are the relative risk of something negative. A few studies report relative times, where the results are expressed as the ratio of the time for the HCQ group versus the time for the control group. One study reports the rate of reduction of viral load, where the result is based on the percentage change in the rate. Calculations were done in Python (3.8.5) with scipy (1.3.3), pythonmeta (1.11), numpy (1.19.1), statsmodels (0.12.0), and plotly (4.10.0). The forest plot is computed using PythonMeta [Deng] with the DerSimonian and Laird random effects model (the fixed effect assumption is not plausible in this case). We received no funding, this research is done in our spare time. We have no affiliations with any pharmaceutical companies or political parties.

We have classified studies as early treatment if most patients are not already at a severe stage at the time of treatment, and treatment started within 5 days after symptoms, although a shorter time may be preferable. Antivirals are typically only considered effective when used within a shorter timeframe, for example 0-36 or 0-48 hours for oseltamivir, with longer delays not being effective [McLean, Treanor].

A summary of study results is below. It is easy to propose excluding certain papers for various reasons, for example [Fried, Kelly, Kuderer, McGrail] report negative results but do not themselves consider the results comparable - they note that treated patients were significantly more ill and do
not make adjustments. To avoid potential bias in evaluation we currently include all studies. HCQ research exhibits a negative bias as shown above and addressing this bias will increase the observed efficacy. Given the state of scientific discussion about HCQ, we feel that a conservative approach is appropriate, especially since efficacy is clear even with this approach. For reference, a draft analysis excluding studies with major issues can be found in Appendix 2.

Please submit updates and corrections with the form at https://hcqmeta.com/.

**Pre-Exposure Prophylaxis**

Only one result per study is included in calculations, as per the details above.

[Abella], risk of COVID-19 case, RR 0.95, \(p = 1.00\).

[Arleo], all patients, RR 0.50, \(p = 0.67\).

[Arleo], inpatients, RR 0.48, \(p = 0.64\).

[Behera], risk of COVID-19 case, RR 0.72, \(p = 0.29\).

[Bhattacharya], risk of COVID-19 case, RR 0.19, \(p = 0.001\).

[Cassione], risk of COVID-19 case, RR 1.50, \(p = 0.59\).

[Chatterjee], full course vs. unused risk of COVID-19 case, RR 0.33, \(p < 0.001\).

[de la Iglesia], risk of hospitalization, RR 1.50, \(p = 1.00\).

[de la Iglesia], suspected COVID-19, RR 1.43, \(p = 0.15\).

[de la Iglesia], confirmed COVID-19, RR 0.92, \(p = 0.84\).

[Ferreira], risk of COVID-19 case, RR 0.53, \(p < 0.001\).

[Ferri], risk of COVID-19 case, RR 0.37, \(p = 0.01\).

[Gendebien], risk of COVID-19 case, RR 0.96, \(p = 0.93\).

[Gendelman], risk of COVID-19 case, RR 0.92, \(p = 0.88\).

[Gentry], risk of death, RR 0.13, \(p = 0.10\).

[Gentry], risk of COVID-19 case, RR 0.79, \(p = 0.27\).

[Gianfrancesco], risk of hospitalization, RR 0.97, \(p = 0.82\).

[Goenka], risk of IgG positive, RR 0.13, \(p = 0.03\).

[Grau-Pujol], risk of COVID-19 case, RR 0.30, \(p = 0.47\).

[Huang], risk of hospitalization, RR 0.20, \(p < 0.001\).

[Huh], risk of COVID-19 case, RR 1.48, \(p = 0.09\).
[Khurana], risk of COVID-19 case, RR 0.49, \( p = 0.02 \).

[Konig], risk of hospitalization, RR 0.97, \( p = 0.88 \).

[Laplana], risk of COVID-19 case, RR 1.56, \( p = 0.24 \).

[Macias], risk of hospitalization, RR 0.74, \( p = 1.00 \).

[Macias], risk of COVID-19 case, RR 1.49, \( p = 0.53 \).

[Mathai], risk of COVID-19 case, RR 0.10, \( p < 0.001 \).

[Mathai], risk of COVID-19 case, RR 0.12, \( p < 0.001 \), symptomatic.

[Mitchell], risk of death, RR 0.01, \( p < 0.001 \).

[Rajasingham], risk of hospitalization, RR 0.50, \( p = 1.00 \).

[Rajasingham], risk of COVID-19 case, RR 0.73, \( p = 0.12 \).

[Rentsch], risk of death, RR 1.03, \( p = 0.83 \).

[Singer], risk of COVID-19 case, RR 1.09, \( p = 0.62 \).

[Zhong], risk of COVID-19 case, RR 0.09, \( p = 0.04 \).

**Post-Exposure Prophylaxis**

Only one result per study is included in calculations, as per the details above.

[Boulware (B)], risk of COVID-19 case, RR 0.83, \( p = 0.35 \).

[Boulware (B)], probable COVID-19 case, RR 0.75, \( p = 0.22 \).

[Dhibar], risk of COVID-19 case, RR 0.59, \( p = 0.03 \).

[Dhibar], risk of COVID-19 case, RR 0.50, \( p = 0.04 \), PCR+.

[Dhibar], risk of symptomatic case, RR 0.56, \( p = 0.21 \).

[Mitjà], risk of death, RR 0.68, \( p = 0.58 \).

[Mitjà], baseline pcr- risk of cases, RR 0.70, \( p = 0.15 \).

[Polat], risk of COVID-19 case, RR 0.43, \( p = 0.03 \).

[Simova], risk of COVID-19 case, RR 0.04, \( p = 0.01 \).

**Early treatment**

Only one result per study is included in calculations, as per the details above.

[Ashraf], risk of death, RR 0.32, \( p = 0.15 \).
[Cadegiani], risk of death, RR 0.17, p = 0.21, control group 1.

[Cadegiani], risk of ventilation, RR 0.05, p < 0.001, control group 1.

[Cadegiani], risk of hospitalization, RR 0.02, p < 0.001, control group 1.

[Chen], median time to PCR-, RR 0.28, p = 0.01.

[Derwand], risk of death, RR 0.21, p = 0.12.

[Derwand], risk of hospitalization, RR 0.18, p < 0.001.

[Esper], risk of hospitalization, RR 0.36, p = 0.02.

[Fonseca], HCQ vs. nothing, RR 0.36, p < 0.001.

[Fonseca], HCQ vs. anything else, RR 0.49, p = 0.006.

[Gautret], risk of no virological cure at day 6, RR 0.34, p = 0.001.

[Guisado-Vasco], risk of death, RR 0.12, p = 0.001.

[Guérin], risk of death, RR 0.57, p = 0.73.

[Guérin], risk of no recovery, RR 0.35, p < 0.001.

[Heras], risk of death, RR 0.08, p < 0.001.

[Hong], risk of prolonged viral shedding, RR 0.35, p = 0.001.

[Huang (B)], risk of no virological cure, RR 0.41, p < 0.001.

[Huang (C)], risk of no recovery at day 14, RR 0.09, p = 0.09.

[Huang (C)], risk of no improvement in pneumonia at day 14, RR 0.17, p = 0.22.

[Ip], risk of hospitalization, RR 0.54, p = 0.03.

[Izoulet], risk of death, RR 0.15, p < 0.001.

[Kirenga], median time to recovery, RR 0.74, p = 0.20.

[Lagier], risk of death, RR 0.41, p = 0.05.

[Ly], risk of death, RR 0.46, p = 0.03.

[Mitjà (B)], risk of hospitalization, RR 0.75, p = 0.64.

[Mitjà (B)], risk of no recovery, RR 0.83, p = 0.38.

[Simova (B)], risk of hospitalization, RR 0.03, p = 0.01.

[Simova (B)], risk of viral+ at day 14, RR 0.02, p = 0.001.
[Skipper], risk of hospitalization, RR 0.48, \( p = 0.19 \).

[Skipper], risk of no recovery at day 14, RR 0.80, \( p = 0.21 \).

[Sulaiman], risk of death, RR 0.36, \( p = 0.01 \).

[Sulaiman], risk of hospitalization, RR 0.61, \( p = 0.001 \).

**Late treatment**

Only one result per study is included in calculations, as per the details above.

[Abd-Elsalam], risk of death, RR 1.20, \( p = 1.00 \).

[Abd-Elsalam], risk of no recovery at day 28, RR 0.70, \( p = 0.009 \).

[Alamdari], risk of death, RR 0.45, \( p = 0.03 \).

[Alberici], risk of death, RR 0.57, \( p = 0.12 \).

[Almazrou], risk of ventilation, RR 0.35, \( p = 0.16 \).

[Almazrou], risk of ICU admission, RR 0.79, \( p = 0.78 \).

[An], time to viral clearance, RR 0.97, \( p = 0.92 \).

[Annie], risk of death, RR 0.96, \( p = 0.83 \).

[Annie], risk of death, RR 1.21, \( p = 0.46 \).

[Aparisi], risk of death, RR 0.37, \( p = 0.008 \).

[Arshad], risk of death, RR 0.49, \( p = 0.009 \).

[Ashinyo], risk of hospitalization, RR 0.67, \( p = 0.03 \).

[Ayerbe], risk of death, RR 0.48, \( p < 0.001 \).

[Barbosa], risk of death, RR 2.47, \( p = 0.58 \).

[Berenguer], risk of death, RR 0.38, \( p < 0.001 \).

[Bernaola], risk of death, RR 0.83, \( p < 0.001 \).

[Bousquet], risk of death, RR 0.57, \( p = 0.15 \).

[Catteau], risk of death, RR 0.68, \( p < 0.001 \).

[Cavalcanti], HCQ+HCQ/AZ risk of death, RR 0.84, \( p = 0.77 \).

[Cavalcanti], HCQ+HCQ/AZ risk of hospitalization, RR 1.28, \( p = 0.30 \).

[Chen (B)], risk of no virological cure, RR 0.76, \( p = 0.71 \).
[Chen (B)], median time to PCR-, RR 0.50, p = 0.40.

[Chen (C)], risk of no virological cure, RR 1.29, p = 0.70.

[Chen (D)], risk of no improvement in pneumonia at day 6, RR 0.43, p = 0.04.

[Chen (E)], risk of radiological progression, RR 0.71, p = 0.57.

[Chen (E)], risk of viral+ at day 7, RR 2.00, p = 1.00.

[Choi], median time to PCR-, RR 1.22, p < 0.001.

[Coll], risk of death, RR 0.54, p < 0.001.

[Cravedi], risk of death, RR 1.53, p = 0.17.

[D'Arminio Monforte], risk of death, RR 0.66, p = 0.12.

[Davido], risk of combined intubation/hospitalization, RR 0.45, p = 0.04.

[Di Castelnuovo], risk of death, RR 0.70, p < 0.001.

[DISCOVERY], 29 day mortality estimated from graph, RR 0.69, p = 0.35.

[DISCOVERY], risk of 7-point scale status, RR 0.83, p = 0.40.

[Dubee], mortality at day 28, RR 0.54, p = 0.21.

[Dubee], combined mortality/intubation at day 28, RR 0.74, p = 0.82.

[Dubee], HCQ+AZ from day 0 subgroup combined mortality/intubation, RR 0.16, p = 0.21.

[Dubernet], risk of ICU admission, RR 0.12, p = 0.008.

[Faíco-Filho], Δt7-12 ΔCt improvement, RR 0.19, p = 0.40.

[Faíco-Filho], Δt<7 ΔCt improvement, RR 0.76, p = 0.36.

[Faíco-Filho], Δt>12 ΔCt improvement, RR 1.15, p = 0.52.

[Fontana], risk of death, RR 0.50, p = 0.53.

[Fried], risk of death, RR 1.27, p < 0.001.

[Frontera], PSM, RR 0.63, p = 0.01.

[Frontera], regression, RR 0.76, p = 0.02.

[Geleris], risk of combined intubation/death, RR 1.04, p = 0.76.

[Goldman], risk of death, RR 0.78, p = 0.46.

[Gonzalez], risk of death, RR 0.73, p = 0.06.
[Guisado-Vasco (B)], risk of death, RR 0.80, p = 0.36.

[Gupta], risk of death, RR 1.06, p = 0.41.

[Heberto], risk of death, RR 0.46, p = 0.04.

[Heberto], risk of ventilation, RR 0.34, p = 0.008.

[Huang (D)], risk of no virological cure, RR 0.33, p < 0.001.

[Ip (B)], risk of death, RR 0.99, p = 0.93.

[Kamran], risk of disease progression, RR 0.95, p = 1.00.

[Kamran], with comorbidities, RR 0.45, p = 0.30.

[Kamran], risk of viral+ at day 14, RR 1.10, p = 0.52.

[Kelly], risk of death, RR 2.43, p = 0.03.

[Kim], risk of hospitalization, RR 0.49, p = 0.01.

[Kim], risk of no virological cure, RR 0.44, p = 0.005.

[Komissarov], risk of viral load, RR 1.25, p = 0.45.

[Kuderer], risk of death, RR 2.34, p < 0.001, HCQ+AZ.

[Lammers], risk of combined death/ICU, RR 0.68, p = 0.02.

[Lano], risk of death, RR 0.67, p = 0.28.

[Lano], risk of combined death/ICU, RR 0.61, p = 0.23.

[Lano], not requiring O2 on diagnosis, RR 0.31, p = 0.11.

[Lauriola], risk of death, RR 0.27, p < 0.001.

[Lecronier], risk of death, RR 0.58, p = 0.24, HCQ vs. control.

[Lecronier], risk of treatment escalation, RR 0.94, p = 0.73, HCQ vs. control.

[Lecronier], risk of viral+ at day 7, RR 0.85, p = 0.61, HCQ vs. control.

[Luo], risk of death, RR 1.02, p = 0.99.

[Lyngbakken], risk of death, RR 0.96, p = 1.00.

[Lyngbakken], improvement in viral load reduction rate, RR 0.29, p = 0.51.

[Magagnoli], risk of death, RR 1.31, p = 0.28.

[Mahévas], risk of death, RR 1.20, p = 0.75.
[Martinez-Lopez], risk of death, RR 0.67, \( p = 0.20 \).

[McGrail], risk of death, RR 1.70, \( p = 0.69 \).

[Membrillo de Novales], risk of death, RR 0.45, \( p = 0.002 \).

[Mikami], risk of death, RR 0.53, \( p < 0.001 \).

[Nachega], risk of death, RR 0.72, \( p = 0.17 \).

[Nachega], risk of no improvement, RR 0.74, \( p = 0.13 \).

[Paccoud], risk of death, RR 0.89, \( p = 0.88 \).

[Peters], risk of death, RR 1.09, \( p = 0.57 \).

[Pinato], risk of death, RR 0.41, \( p < 0.001 \).

[RECOVERY], risk of death, RR 1.09, \( p = 0.15 \).

[Rivera], risk of death, RR 1.02, \( p = 0.90 \).

[Roomi], risk of death, RR 1.38, \( p = 0.54 \).

[Rosenberg], risk of death, RR 1.35, \( p = 0.31 \).

[Saleemi], median time to PCR-, RR 1.21, \( p < 0.05 \).

[Sbidian], risk of death, RR 1.05, \( p = 0.74 \), whole population HCQ AIPTW adjusted.

[Sbidian], risk of no hospital discharge, RR 0.80, \( p = 0.002 \), whole population HCQ AIPTW adjusted.

[Self], risk of death, RR 0.93, \( p = 0.84 \).

[Serrano], risk of death, RR 0.57, \( p = 0.14 \).

[Shabrawishi], risk of no virological cure at day 5, RR 0.85, \( p = 0.66 \).

[Shoaibi], risk of death, RR 0.85, \( p < 0.001 \).

[Singh], risk of death, RR 0.95, \( p = 0.72 \).

[Singh], risk of ventilation, RR 0.81, \( p = 0.26 \).

[Solh], risk of death, RR 1.18, \( p = 0.17 \).

[SOLIDARITY], risk of death, RR 1.19, \( p = 0.23 \).

[Soto-Becerra], risk of death, RR 1.84, \( p = 0.02 \).

[Synolaki], risk of death, RR 0.76, \( p = 0.27 \).

[Sánchez-Álvarez], risk of death, RR 0.54, \( p = 0.005 \).


Appendix 2. Draft Analysis with Exclusions

Many meta-analyses for HCQ have been written, most of which have become somewhat obsolete due to the continuing stream of more recent studies. Recent analyses with positive conclusions include [IHU Marseille] which considers significant bias from an understanding of each trial, and [Garcia-Albeniz, Ladapo, Prodromos] which focus on early or prophylactic use studies. Meta analyses reporting negative conclusions focus on late treatment studies, tend to disregard treatment delay, tend to follow formulaic evaluations which overlook major issues with various studies, and end up with weighting disproportionate to a reasoned analysis of each study’s contribution. For example, [Axfors] assigns 87% weight to a single trial, the RECOVERY trial [RECOVERY], thereby producing the same result. However, the RECOVERY trial may be the most biased of the studies they included, due to the excessive dosage used, close to the level shown to be very dangerous in [Borba] (OR 2.8), and with extremely sick late stage patients (60% requiring oxygen, 17% ventilation/ECMO, and a very high mortality rate in both arms). There is little reason to suggest that the results from this trial are applicable to more typical dosages or to earlier treatment (10/22; the second version of this study released 10/22 assigns 74% to RECOVERY and 15% to SOLIDARITY [SOLIDARITY], which is the only other trial using a similar excessive dosage).

We include all studies in the main analysis, however there are major issues with several studies that could significantly alter the results. Here, we present a draft analysis excluding studies with significant issues, including indication of significant unadjusted group differences or confounding by indication, extremely late stage usage >14 days post symptoms or >50% on oxygen at baseline, very minimal detail provided, excessive dosages which have been shown to be dangerous, significant issues with adjustments that could reasonably make substantial differences, and reliance on PCR which may be inaccurate and less indicative of severity than symptoms. We welcome feedback on improvements or corrections to this. The studies excluded are as follows, and the resulting forest plot is shown in Figure 10.

[Tang], risk of no virological cure at day 21, RR 0.79, p = 0.51.

[Tehrani], risk of death, RR 0.87, p = 0.63.

[Trullàs], risk of death, RR 0.64, p = 0.12.

[Ulrich], risk of death, RR 1.06, p = 1.00.

[Wang], risk of death, RR 0.94, p = 0.63.

[Xia], risk of no virological cure, RR 0.62, p = 0.17.

[Yu], risk of death, RR 0.40, p = 0.002.

[Zhong (B)], risk of no virological cure at day 10, RR 0.20, p < 0.001.

[Ñamendys-Silva], HCQ+AZ vs. neither HCQ or CQ, RR 0.68, p = 0.18.

[Ñamendys-Silva], CQ vs. neither HCQ or CQ, RR 0.63, p = 0.09.

[Ñamendys-Silva], HCQ+AZ or CQ, RR 0.66, p = 0.006.
[Alamdari], substantial unadjusted confounding by indication.

[An], results only for PCR status which may be significantly different to symptoms.

[Annie], confounding by indication is likely and adjustments do not consider COVID-19 severity.

[Barbosa], excessive unadjusted differences between groups.

[Cassione], not fully adjusting for the different baseline risk of systemic autoimmune patients.

[Chen], results only for PCR status which may be significantly different to symptoms.

[Chen (B)], results only for PCR status which may be significantly different to symptoms.

[Chen (C)], results only for PCR status which may be significantly different to symptoms.

[Cravedi], substantial unadjusted confounding by indication.

[de la Iglesia], not fully adjusting for the different baseline risk of systemic autoimmune patients.

[Fried], excessive unadjusted differences between groups, substantial unadjusted confounding by indication.

[Gautret], excessive unadjusted differences between groups, results only for PCR status which may be significantly different to symptoms.

[Geleris], significant issues found with adjustments.

[Gendebien], not fully adjusting for the baseline risk differences within systemic autoimmune patients.

[Gendelman], not fully adjusting for the different baseline risk of systemic autoimmune patients.

[Gianfrancesco], not fully adjusting for the baseline risk differences within systemic autoimmune patients.

[Gupta], >50% on oxygen/ventilation at baseline.

[Hong], results only for PCR status which may be significantly different to symptoms.

[Huang], significant unadjusted confounding possible.

[Huang (B)], results only for PCR status which may be significantly different to symptoms.

[Huang (D)], results only for PCR status which may be significantly different to symptoms.

[Huh], not fully adjusting for the different baseline risk of systemic autoimmune patients.

[Izoulet], excessive unadjusted differences between groups.

[Kamran], excessive unadjusted differences between groups.

[Kelly], substantial unadjusted confounding by indication.
[Konig], not fully adjusting for the baseline risk differences within systemic autoimmune patients.

[Kuderer], substantial unadjusted confounding by indication.

[Laplana], not fully adjusting for the different baseline risk of systemic autoimmune patients.

[Lecronier], >50% on oxygen/ventilation at baseline.

[Luo], substantial unadjusted confounding by indication.

[Lyngbakken], results only for PCR status which may be significantly different to symptoms.

[Macias], not fully adjusting for the baseline risk differences within systemic autoimmune patients.

[McGrail], excessive unadjusted differences between groups.

[Mitchell], excessive unadjusted differences between groups.

[Peters], excessive unadjusted differences between groups.

[RECOVERY], excessive dosage, results do not apply to typical dosages.

[Rentzsch], not fully adjusting for the baseline risk differences within systemic autoimmune patients, medication adherence unknown and may significantly change results.

[Roomi], substantial unadjusted confounding by indication.

[Saleemi], results only for PCR status which may be significantly different to symptoms, substantial unadjusted confounding by indication.

[Sbidian], significant issues found with adjustments.

[Shabrawishi], results only for PCR status which may be significantly different to symptoms.

[Singer], not fully adjusting for the baseline risk differences within systemic autoimmune patients.

[Singh], confounding by indication is likely and adjustments do not consider COVID-19 severity.

[Solkh], >50% on oxygen/ventilation at baseline, substantial unadjusted confounding by indication.

[SOLIDARITY], excessive dosage, results do not apply to typical dosages, >50% on oxygen/ventilation at baseline.

[Soto-Becerra], confounding by indication is likely and adjustments do not consider COVID-19 severity.

[Tang], results only for PCR status which may be significantly different to symptoms.

[Tehrani], substantial unadjusted confounding by indication.

[Ulrich], >50% on oxygen/ventilation at baseline.

[Wang], confounding by indication is likely and adjustments do not consider COVID-19 severity.
[Xia], detail too minimal.

[Zhong (B)], results only for PCR status which may be significantly different to symptoms.
<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huang</td>
<td>0.09</td>
<td>[0.01-1.44]</td>
</tr>
<tr>
<td>Esper</td>
<td>0.36</td>
<td>[0.15-0.87]</td>
</tr>
<tr>
<td>Ashraf</td>
<td>0.32</td>
<td>[0.10-1.10]</td>
</tr>
<tr>
<td>Guerin</td>
<td>0.57</td>
<td>[0.02-1.34]</td>
</tr>
<tr>
<td>Lagier</td>
<td>0.41</td>
<td>[0.27-0.62]</td>
</tr>
<tr>
<td>Derwand</td>
<td>0.21</td>
<td>[0.03-1.47]</td>
</tr>
<tr>
<td>Mitja</td>
<td>0.75</td>
<td>[0.32-1.78]</td>
</tr>
<tr>
<td>Skorpen</td>
<td>0.48</td>
<td>[0.17-3.39]</td>
</tr>
<tr>
<td>Ly</td>
<td>0.46</td>
<td>[0.27-0.79]</td>
</tr>
<tr>
<td>Ip</td>
<td>0.54</td>
<td>[0.36-0.95]</td>
</tr>
<tr>
<td>Hieras</td>
<td>0.08</td>
<td>[0.03-0.25]</td>
</tr>
<tr>
<td>Kirenga</td>
<td>0.74</td>
<td>[0.37-1.48]</td>
</tr>
<tr>
<td>Sulamani</td>
<td>0.36</td>
<td>[0.17-0.80]</td>
</tr>
<tr>
<td>Guisado-Vasco (ES)</td>
<td>0.12</td>
<td>[0.03-0.45]</td>
</tr>
<tr>
<td>Fonseca</td>
<td>0.36</td>
<td>[0.20-0.67]</td>
</tr>
<tr>
<td>Cadejgiani</td>
<td>0.17</td>
<td>[0.01-3.56]</td>
</tr>
<tr>
<td>Simova</td>
<td>0.03</td>
<td>[0.00-0.60]</td>
</tr>
<tr>
<td><strong>Early</strong></td>
<td>0.36</td>
<td>[0.26-0.49]</td>
</tr>
<tr>
<td><strong>Late</strong></td>
<td>0.64</td>
<td>[0.57-0.73]</td>
</tr>
</tbody>
</table>

64% improvement

36% improvement
Figure 10. Forest plot (random effects model) excluding studies with significant issues. (ES) indicates the early treatment subset of a study (these are not included in the overall results).