**Combatting COVID Controversies: An Argument for Open and Explicit Research Based on Quantitative case studies from the SARS-CoV-2 Pandemic**

**Abstract:**

Since the origins of the SARS-CoV-2 pandemic in December 2019, there have been over 263 million confirmed cases and over 5 million deaths from COVID-19 (WHO, 2021). In this difficult and unprecedented time, many researchers have worked tireless to guide efforts to reduce the burden of this disease. This paper retrospectively analyzes scientific literature that impacted responses to the COVID-19 pandemic to illustrate how scientific evidence can be both immeasurably valuable and/or potentially misleading depending on how it is presented. This is especially true when findings are reported in an information-limited setting and the topics are controversial, as in the case of , the effectiveness of facemasks, and possible treatments for COVID-19. As time has passed, however, and new confusions about the omicron variant of SARS-CoV-2 emerge, scientists are more explicit in their communication of results to others, which will hopefully set a new standard for open data and interpretations in all non-pandemic publications to come.

**Background**: **Statistical Significance Testing**

One of the hallmarks of epidemiological studies, including many COVID-19 studies, is the use of statistical significance testing (SST). At the most basic, the p-value is the probability of obtaining an estimate at least as far from a specified value as the estimate obtained, while the 95% confidence interval is the set of all parameter values for which p0.05 (Poole, 2001). Though useful for identifying unusual data, the p-value as a measure of statistical significance relies on a variety of presumptions, including appropriate sample size, correct models, and absence of bias and confounding, which are not always met. Many studies have espoused the dangers of relying solely on p-values and SSTs in biomedical research (Stang, Poole & Kuss, 2010; Greenland *et al*, 2016). Instead of depending on p-values in epidemiological studies, these researchers emphasize the importance of examining other factors, including effect size and confidence interval range (Poole, 2001). Though issues of misuse and misinterpretation have been discussed and published for years in the biomedical research community, scientists continue to misrepresent SSTs to this day, and this misinterpretation is one of the contributing factors to the added confusion and worsening of the COVID-19 pandemic.

**Assessments:**

Many of the first quantitative studies on SARS-CoV-2 centered on the virus’ basic reproductive number, or R-naught. One of the first peer-reviewed papers published on this topic was completed by Li *et al* (2020) and released on January 29 (Singh *et al*, 2021). After collecting case data between December 10 and January 4, the authors fitted a transmission model and estimated the to be 2.2 (95% CI, 1.4 to 3.9). Though based on very early case count data, and with a confidence interval indicating lower precision in results, this estimate was similar to other peer-reviewed estimates released at that time (Majumder and Mandl, 2020). While many peer-reviewed papers were being released with estimates, other scientists were also releasing pre-prints with their own estimates to expedite collaboration and information-sharing.

Chart, bubble chart

Description automatically generated In a subsequent article, Majumder and Mandl (2020) compared the estimates and corresponding confidence intervals in several articles approximating the of SARS-CoV-2 and analyzed their impact on the public and the larger scientific community. A key figure from this study is replicated in **Figure 1**. Based on the studies included in **Figure 1**, the pre-print groups and peer-reviewed groups have many differences. For one, there is considerable variability in the estimates of the pre-print group, whose values range from 2.2-6.5, while that of the peer review group ranges from 2.2-3.4. Additionally, the mean of the pre-print group (3.61 CI: 2.77-4.45) is greater than that of the peer-review group (2.54 CI: 2.17-2.91), even if the two upper limit outliers are removed from the calculation (3.02 CI: 2.65-3.39) (Majumder and Mandl, 2020). Finally, several of the studies included in the analysis by Majumder and Mandl were adjusted and reincluded, and these updates estimates each reflect greater precision after based on the confidence interval ranges in **Figure 1**. Ultimately, the authors conclude from their analysis that pre-prints, rather than peer-reviewed article were more influential on COVID-19 discourse because of how quickly their results could be disseminated, despite possible issues of credibility and misinformation.

**Figure 1: Reproduced figure from Majumder and Mandl (2020) plotting early estimates for COVID-19 in pre-prints and peer-reviewed publications.**

While pre-prints can be invaluable, especially in situations like the COVID-19 pandemic when timely-information sharing is essential, care must be taken to ensure that the results being shared are accurate. As previously detailed, the pre-print articles in this study are much more varied than those that were peer-reviewed, and each re-examined study demonstrated greater precision upon re-evaluation. Thus, there is an additional responsibility on authors publishing pre-prints to express, as clearly as possible, the limitations of their own research, and to provide as much open-information about methods, data, estimates, and confidence intervals so readers can come to their own conclusions about the applicability of the results when peer-review is not available. Additionally, publishing sites have a responsibility to clearly indicate clearly when a paper has not yet been peer-reviewed to their readers, and when a paper has been updated, features which were not always obvious initially in the pandemic (Dinis-Oliveira, 2020). The next paragraphs examine the fallbacks that can occur when information is misrepresented and, in turn, misinterpreted by the public, at a time when clarity in communication is essential.

**Effectiveness of Mask-Wearing**:

Another controversial topic in the early months of the COVID-19 pandemic was the effectiveness of face masks as a non-pharmaceutical intervention (NPI). Because SARS-CoV-2 is thought to primarily spread from person to person via respiratory droplets, the use of facemasks is an easy and cheap method with which to slow that spread, however, there aren’t many studies which examined the effectiveness of mask-wearing prior to the COVID-19 pandemic (A PubMed search of ‘mask wearing effectiveness’ generates 1,295 results, more than 900 of which are dated between 2020 and 2022). Of the reviews published prior to the COVID-19 pandemic, however, most emphasize the need for continued research, or the recommendation that face masks be worn only “as a last resort” against respiratory viruses (Cowling *et al*, 2010; Davies *et al*, 2013). It is understandable then, why the recommendations of mask-wearing were initially controversial, especially considering the added confusion from mixed stances of authorities like the World Health Organization (WHO, 2020).

Chart

Description automatically generated Despite the complicated and contradictory evidence on the effectiveness of mask-wearing against viral transmission, however, mask mandates were quickly introduced worldwide to mitigate SARS-CoV-2 transmission. In May 2020, a systematic review and meta-analysis on the efficacy of facemasks was completed by Liang *et al* (2020). Following a Preferred Reporting Items for Systemic Reviews and Meta-Analysis (PRISMA) statement, 21 studies were included in the overall analysis, which consisted of 13 case-control studies, 6 cluster randomized trials, and 2 cohort studies (*ibid*). A forest plot from this study is replicated in **Figure 2**. Though 9 of the 21 studies were unable to find a significant difference between the risk of contracting respiratory viruses in those who wore masks compared to those who didn’t, the overall analysis of over 8,600 participants presents strong evidence of reduction of risk in mask wearers. From the pooled results, the odds of contracting respiratory viruses were 65% percent lower in the mask-wearers compared to non-mask-wearers (OR=0.35, CI: 0.24-0.51) (*ibid*). These results, with a strong effect size and narrow confidence interval generated from a large, compiled sample, demonstrate not only the effectiveness of mask wearing to prevent respiratory viral spread, but also how overall analyses are necessary to bolster claims based on frequentist methods (Gurevitch *et al*, 2018).

**Figure 2: Replicated Forest plot from Liang et al (2020) demonstrating individual and pooled odds ratios of mask wearing as a prevention for respiratory viral infection.**

**Proposed Treatments:**

While several influential studies focused on the transmissibility of SARS-CoV-2, there were also major initial research efforts into possible treatments for the COVID-19 disease, some of which were very highly publicized. One of the first treatment proposals was the immunosuppressive hydroxychloroquine, which has been approved by the Food and Drug Administration (FDA) to treat conditions like malaria and lupus. In February 2020, an article examined the effectiveness of Remdesivir and chloroquine against SARS-CoV-2 *in vitro* (Wang *et al*, 2020). The first clinical trial to test these results *in vivo* was completed by Gautret *et al* (2020) and published in July the same year. Gautret *et al* (2020) compared 20 patients with confirmed COVID-19 who were treated with 600mg of hydroxychloroquine daily to 16 control patients and, using Student’s t-test, determined there was a statistically significant difference between the clearance outcomes in both groups (Day 6 post inclusion, p=0.001).

Despite these optimistic results, other researchers have pointed out the limitations of this trial, including a lack of randomization, blinding, and appropriate placebo control in addition to a poor sample size, which did not allow for the completion of the primary outcome (dos Santos, 2021). Further, the use of Student’s t-test was inappropriate to test significance, as Gautret *et al* (2020) collected repeated outcome measures, meaning an analysis of variance (ANOVA) should have been included (dos Santos, 2021). Gautret *et al* (2020) also neglect to include standard deviations and confidence intervals in their analysis, which, as mentioned earlier, are often more informative than the basic p-value. Despite these criticisms being released and available to the public, many people, including former United States President Donald Trump, went months believing hydroxychloroquine to be an effective treatment against COVID-19 when in fact it is inconclusive and poorly supported, as demonstrated in recent systematic reviews (Hernandez *et al*, 2020; Lewis *et al*, 2021; Das *et al*, 2020).

Other drugs were also featured for newsworthy claims of success in treatment against COVID -19. Ivermectin, an FDA-approved drug used to treat parasitic worms, is one such drug. Though the FDA now strongly advises individuals to avoid using ivermectin as a treatment or prevention against COVID -19, initial experimental results *in vitro* published in June 2020 were promising (FDA, 2021; Caly *et al*, 2020). One case-series completed after this initial *in vitro* report observed 100 patients with COVID-19 who were given 0.2mg/kg of ivermectin each day. Though Alam *et al* (2020) concluded in this case series that ivermectin was “very effective” at viral clearance in patients with mild and moderate COVID-19, it is essential to note that there was no control group for this intervention and, thus, there was also no randomization of treatment provided. Acknowledging these drawbacks, Alam *et al* (2020) provided a recommendation for further studies to clarify ivermectin’s effectiveness.

One of the first clinical trials of ivermectin against COVID-19 was tested by Ahmed *et al* (2020). This trial has many benefits over the case series of Alam *et al* (2020) in that it includes a placebo group and randomizes 72 hospitalized patients to one of three interventions: treatment with ivermectin only, treatment with ivermectin + doxycycline, and placebo treatment. Though virological clearance was significantly earlier in the ivermectin group compared to the control group (9.7 days, CI: 7.8-11.8 compared to 12.7 days, CI: 11.3-14.2, respectively), there was no difference between the mean duration of hospitalization after treatment (9.7 days, CI: 8.1-11.0, compared to 9.6 days, 7.7-11.7, respectively). It is unclear in the methods of this study what statistical methods were used to make these assessments, and whether they were analyzed using an ANOVA, as is necessary for interventions with more than one assignment group.

In addition to the duration comparisons, Alam *et al* (2021) also establish a hazard ratio (HR) for virological clearance. On day 7, the authors calculate that those in the ivermectin group reach viral clearance 4.1 times earlier than those in the placebo group (p=0.03), however, the confidence interval for this hazard ratio is huge, ranging from 1.1-14.7. The results are similar at the two-week mark, with an HR of 2.7 with a p=0.02 and a 95% confidence interval of 1.2-6.0 (Alam *et al*, 2021). Though statistically significant, appropriate caution should be taken, and emphasized to the readers of this study, as precision in these HRs is considerably lacking.

Another clinical trial examining the effectiveness of ivermectin against COVID-19 was completed by López-Medina *et al* (2021). This study included 476 participants and randomly assigned patients to either a 5-day course of ivermectin or a placebo treatment. The analyses were assessed through Kaplan-Meier plots and compared with a log-rank test, and the authors reported no significant difference between the time to resolution of symptoms between the two groups (10 days vs. 12 days in intervention vs. control groups, respectively). The HR for the resolution of symptoms was reported as 1.07 (CI: 0.87-1.37, p=0.53) (López-Medina *et al*, 2021).

While Alam *et al* (2021) and López-Media *et al* (2021) completed their reports under separate conditions and neither can make absolute conclusions about the effectiveness of ivermectin against COVID-19, there are clear differences in the reporting of their results which affect their readers’ perceptions of their conclusions. As noted, there were inequalities between the two papers in the transparency of their methods, a vital note for scientists seeking to reproduce findings from scientific publications. Additionally, while the individual HRs reported differed between the two studies, so did the size of the confidence intervals, indicating additional differences in the precision of each of the two reports, likely due in part to the sixfold difference in sample size. While these nuances will likely be reflected in any systemic review or meta-analysis including either two papers, these compilations take time, and, when scientific results such as these can and should be disseminated as quickly as possible, it is essential for both the authors of a publication and the publishing site to take additional steps to limit the misrepresentation of results. Without a background in basic epidemiology, or prior knowledge of the importance of confidence intervals and p-values, lay readers and even experienced epidemiologists may fall into common misunderstandings of the true meaning of these frequentist results, and instead may revert to the harmful reductionist stance of solely seeing reports as “significant” or not.

**Omicron**:

Now, with the newly discovered omicron variant, COVID-19 research has entered a new liminal space. From the early pre-prints currently available, however, there are already marked improvements in how research is being completed and publicized. The first of these studies comes from Pulliam *et al* (2021) who have examined the risk of reinfection among variants of SARS-CoV-2, including omicron. The study includes a large population of over 2.7 million individuals with confirmed SARS-CoV-2 and considers reinfected cases those which have two positive tests at least 90 days apart. The risk of reinfection from Beta and Delta variants had hazard rations of 0.75 (0.59-0.97) and 0.71 (0.56-0.92), respectively, while the hazard ratio for reinfection from omicron was 2.39 (1.88-3.11) (Pulliam *et al*, 2021). In comparisons to other early studies explored in this paper, this recent preprint shows many positive signs. Not only is the sample size abundantly large, but the inclusion of hazard ratios with confidence intervals provides the reader with a greater understanding of the uncertainty surrounding omicron analysis, though it should be noted that the confidence interval for the omicron hazard ratio is more than three times larger than that of the other ratios, specifying the relative lack of uncertainty regarding analyses of the new variant. Further, medRxiv, the site hosting this preprint, has included the helpful disclaimer below the title of this article labelling it as a preprint not yet certified by peer review, a feature neglected in many previous COVID-19 preprints (Dinis-Oliveira, 2020).

**Conclusion:**

This paper examines several controversial topics researched over the course of the SARS-CoV-2 pandemic, including debates over , mask mandates, potential treatments, and features of omicron. Many of these important and influential papers use appropriate methods and present results accurately and clearly, however, that cannot be said of all scientific articles examined, especially those central to these controversial topics. While it is difficult to make inferences in such an information-limited setting as the start of a novel outbreak, the publications highlighted in this paper highlight key issues in biomedical publication which, intentionally or unintentionally, obfuscate interpretations in favor of “significant” results. While p-values and confidence intervals are important in noting unusual data, these must be considered with assumptions of appropriate sample sizes and correct model usage. In a time when scientific results can and should be shared as quickly as possible, researchers and publishers have an additional responsibility to their readers to ensure their methods are clearly stated and all results clearly indicate confidence interval ranges and limitations in addition to the basic “significance”. Luckily, as the pandemic has continued, researchers have been able to point out and change poor practices for the better, hopefully in such a way that open research and clear interpretation becomes the norm in all circumstances.

**Citations:**

Ahmed, S., Karim, M. M., Ross, A. G., Hossain, M. S., Clemens, J. D., Sumiya, M. K., ... & Khan, W. A. (2021). A five-day course of ivermectin for the treatment of COVID-19 may reduce the duration of illness. *International Journal of Infectious Diseases*, *103*, 214-216.

Alam, M. T., Murshed, R., Bhiuyan, E., Saber, S., Alam, R. F., & Robin, R. C. (2020). A case series of 100 COVID-19 positive patients treated with combination of ivermectin and doxycycline. *Journal of Bangladesh College of Physicians and Surgeons*, 10-15.

Caly, L., Druce, J. D., Catton, M. G., Jans, D. A., & Wagstaff, K. M. (2020). The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 in vitro. *Antiviral research*, *178*, 104787.

Cowling, B. J., Zhou, Y. D. K. M., Ip, D. K. M., Leung, G. M., & Aiello, A. E. (2010). Face masks to prevent transmission of influenza virus: a systematic review. *Epidemiology & infection*, *138*(4), 449-456.

Das, R. R., Behera, B., Mishra, B., & Naik, S. S. (2020). Effect of chloroquine and Hydroxychloroquine on COVID-19 virological outcomes: An updated meta-analysis. *Indian journal of medical microbiology*, *38*(3-4), 265-272.

Davies, A., Thompson, K. A., Giri, K., Kafatos, G., Walker, J., & Bennett, A. (2013). Testing the efficacy of homemade masks: would they protect in an influenza pandemic?. *Disaster medicine and public health preparedness*, *7*(4), 413-418.

Dos Santos, M. R. (2021). Referring to the study: Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open label non-randomized clinical trial. *International Journal of Antimicrobial Agents*, *57*(1), 106176.

Food and Drug administration. (2021). Why You Should Not Use Ivermectin to Treat or Prevent COVID-19. Retrieved from <https://www.fda.gov/consumers/consumer-updates/why-you-should-not-use-ivermectin-treat-or-prevent-covid-19>

Gautret, P., Lagier, J. C., Parola, P., Meddeb, L., Mailhe, M., Doudier, B., ... & Raoult, D. (2020). Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *International journal of antimicrobial agents*, *56*(1), 105949.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, *31*(4), 337-350.

Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, *555*(7695), 175-182.

Hernandez, A. V., Roman, Y. M., Pasupuleti, V., Barboza, J. J., & White, C. M. (2020). Hydroxychloroquine or chloroquine for treatment or prophylaxis of COVID-19: a living systematic review. *Annals of internal medicine*, *173*(4), 287-296.

Lewis, K., Chaudhuri, D., Alshamsi, F., Carayannopoulos, L., Dearness, K., Chagla, Z., ... & GUIDE Group. (2021). The efficacy and safety of hydroxychloroquine for COVID-19 prophylaxis: A systematic review and meta-analysis of randomized trials. *PloS one*, *16*(1), e0244778.

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., ... & Feng, Z. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England journal of medicine*.

Liang, M., Gao, L., Cheng, C., Zhou, Q., Uy, J. P., Heiner, K., & Sun, C. (2020). Efficacy of face mask in preventing respiratory virus transmission: A systematic review and meta-analysis. *Travel medicine and infectious disease*, *36*, 101751.

López-Medina, E., López, P., Hurtado, I. C., Dávalos, D. M., Ramirez, O., Martínez, E., ... & Caicedo, I. (2021). Effect of ivermectin on time to resolution of symptoms among adults with mild COVID-19: a randomized clinical trial. *Jama*, *325*(14), 1426-1435.

Majumder, M. S., & Mandl, K. D. (2020). Early in the epidemic: impact of preprints on global discourse about COVID-19 transmissibility. *The Lancet Global Health*, *8*(5), e627-e630.

Poole, C. (2001). Low P-values or narrow confidence intervals: which are more durable?. *Epidemiology*, *12*(3), 291-294.

Pulliam, J. R. C., van Schalkwyk, C., Govender, N., von Gottberg, A., Cohen, C., Groome, M.J., Dushoff, J., Mlisana, K., Moultrie, H., (2021). Increased risk of SARS-CoV-2 reinfection associated with emergence of the Omicron variant in South Africa. medRxiv 2021.11.11.21266068; doi:https://doi.org/10.1101/2021.11.11.21266068

Singh, S., McNab, C., Olson, R. M., Bristol, N., Nolan, C., Bergstrøm, E., ... & Phelan, A. (2021). How an outbreak became a pandemic: a chronological analysis of crucial junctures and international obligations in the early months of the COVID-19 pandemic. *The Lancet*.

Stang, A., Poole, C., & Kuss, O. (2010). The ongoing tyranny of statistical significance testing in biomedical research. *European journal of epidemiology*, *25*(4), 225-230.

Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., ... & Xiao, G. (2020). Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell research*, *30*(3), 269-271.

 World Health Organization. Advice on the use of masks in the community, during home care and in health care settings in the context of the novel coronavirus (2019-nCoV) outbreak. Geneva, Switzerland; 2020.

World Health Organization. (2021). WHO Coronavirus (COVID-19) Dashboard. Retrieved from <https://covid19.who.int/>