

An Analysis of the COVID-19 Infodemic: The Impact of American Public Sources on Sentiment, Conversation, and Physician Behaviour Towards Hydroxychloroquine

10.17975/sfj-2021-001

Emily W.L. Chan[†], Ginah Choi^{*,††}, Kendrew S.K. Wong^{§,††}, Shirley Zeng^{||,††}, and Anish Verma[#]

[†]Department of Integrated Science, University of British Columbia

^{††}Marketing and Behavioural Science Division, University of British Columbia

[‡]Department of Medicine, University of British Columbia

[§]Department of Computer Science, University of British Columbia

[#]1QB Information Technologies (1QBit), Vancouver, BC, Canada

^{††} These authors contributed equally to this work

ABSTRACT: The COVID-19 infodemic, described as an overabundance of both accurate and inaccurate information, poses a significant public health risk in spreading fear and provoking inappropriate prescription. The overwhelming and often contradictory information on as potential treatments for COVID-19 have contributed to this infodemic. Public sources including the US federal government, health organizations, and research publications have released conflicting statements on the efficacy of hydroxychloroquine. Previous research has not analyzed the influence of these sources on public attitudes and conversation towards the drug. To evaluate this impact, changes in the number and sentiment of tweets tagged with the hashtag or keyword "hydroxychloroquine" from March 12th to June 22nd, 2020 in relation to public sources were analyzed. We found that the US government had a statistically significant influence on public attitudes and behaviour ($p < 0.001$), unlike health organizations and research publications. Public sentiment on hydroxychloroquine has also been observed to become more negative over time, suggesting that public attitudes towards controversial topics can change. This study also found a positive correlation between public sentiment of hydroxychloroquine and other drugs (i.e. azithromycin and remdesivir) which indicates that public sources disseminating hydroxychloroquine-related information could also affect public attitudes towards related treatments. In a public health crisis, all statements and actions from public sources regarding contentious topics like hydroxychloroquine should be made with caution. To mitigate the disproportionate influence of public sources in an infodemic, we recommend three solutions: (a) education to empower individuals of all ages to develop critical thinking and digital literacy skills; (b) stronger action from social media platforms in labeling misinformation; (c) and cooperation between entities with strong influence (e.g. federal government) and other sources for public health measures. Together, these recommendations could resolve shortcomings existent with a single approach. Future research should be conducted with a custom trained model for sentiment analysis. It would also be valuable to conduct a similar version of the study on other social media platforms as well as for public health issues beyond COVID-19.

KEYWORDS: infodemic, sentiment, Twitter, hydroxychloroquine, COVID-19

INTRODUCTION

Caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), coronavirus disease of 2019 (COVID-19), has rapidly escalated to a global pandemic with over 10.9 million confirmed cases as of June 2020 [1]. The World Health Organization (WHO) has recognized an accompanying global

epidemic of misinformation called an infodemic [2], defined as "an overabundance of information, some accurate and some not, that makes it hard for people to find trustworthy sources and reliable guidance" [3]. The COVID-19 infodemic can lead to adverse public reactions, such as fear; mistrust in medical advice; reduced response to following public

health measures; and inappropriate prescription practices from physicians [4]. Other consequences can include a wider uptake and diminished availability of certain drugs, which can be especially harmful to vulnerable populations needing these drugs to manage existing immune-mediated inflammatory diseases [5].

The COVID-19 infodemic is further exacerbated by social media - Sylvie Briand, director at WHO's Health Emergencies Programme, has noted that social media amplifies misinformation and rumours to spread "faster and further" [2]. Twitter, as "a space for the informal sharing of health information" with the ability to mass circulate information [6], can be an especially dangerous breeding ground for misinformation and rumours. While previous research publications have measured people's changing sentiment on Twitter towards topics of misinformation [7] [8] [9], this is the first known study to apply a sentiment analysis approach to investigate the impact of public sources (i.e. the US federal government, health organizations, and research publications) on attitudes towards potential COVID-19 treatment drugs. Sentiment analysis is commonly used for determining public opinion on commercial products by analyzing reviews and comments made by consumers [10]. The same approach can be used to effectively categorize tweets by sentiment, such that negative sentiment tweets represent negative attitudes towards hydroxychloroquine, and positive sentiment tweets represent positive attitudes towards the drug.

Throughout this infodemic, communication regarding 4-aminoquinolinedrugs(chloroquineandhydroxychloroquine) as potential treatments for COVID-19 has often been overwhelming and contradictory [11]. Hydroxychloroquine, used to treat malaria and rheumatic diseases [12,13], has been shown to be effective against human coronaviruses, such as the human coronavirus (HCoV) strain OC43 infection [14] and SARS-CoV-1 [14]. Other drugs such as azithromycin [15,16], remdesivir [17,18], and dexamethasone [17,18] have also been considered. While early retrospective studies and small-scale clinical trials examining hydroxychloroquine have shown promising COVID-19 patient outcomes [19,20,21,22], other studies have shown uncertain or negative conclusions on its safety and efficacy [23,21,22]. Researchers have advised against the use of the drug for treatment of COVID-19 until clinical trials could determine its efficacy in treatment [5,25]. It should be noted that although hydroxychloroquine had

Table 1. Search strategy employed for the literature review. The search strings were modified based on the source of information.

Event type	Search string
Public figure	"last name of the public figure" AND "hydroxychloroquine"
Organization*	"full name of the organization" OR "abbreviated name of the organization if applicable*" AND "hydroxychloroquine"
Research paper	"title of the study" OR "name of the re-search journal/name of the online publisher" AND "hydroxychloroquine"

*An exception was made for the World Health Organization (WHO) as results would have included the word "who", which did not refer to the World Health Organization.

been explored as a potential treatment as early as January 2020 [24,25], these studies were published in research journals likely little-known by the public. As such, we would expect public attitudes towards the drug to be formed during its recent popularization.

In the dissemination of information regarding hydroxychloroquine, American public sources have played a central role. President of the United States (POTUS), Donald J. Trump, has repeatedly promoted hydroxychloroquine as a safe drug for treating COVID-19 [24-26], while other American public sources have communicated otherwise [22,28-30]. During the same time period, the US was experiencing one of the highest numbers and rates of new COVID-19 cases in the world [26] making potential treatments of high relevance to the public. Although Brennen et al. have linked public sources as key contributors of misinformation [27], it is unclear the level of influence that different public sources have on public attitudes and behaviour towards hydroxychloroquine. The diversity of voices within the same time period provides a unique opportunity to measure the extent of influence American public sources exhibit in an infodemic.

Thus, the objective of this study is to investigate the extent to which public attitudes, conversation, and physician behaviour are influenced by American public sources communicating conflicting information on hydroxychloroquine as a COVID-19 treatment. As a secondary aim, we extend the examination of public attitudes, conversation, and physician behaviour towards similar potential COVID-19 treatment drugs (e.g. azithromycin, remdesivir, and dexamethasone) to identify the

extent of public sources' impact beyond hydroxychloroquine when disseminating hydroxychloroquine-related information.

METHODS

Determination of significant events regarding hydroxychloroquine

A literature review was conducted to compile a preliminary list of all events where American public sources promoted or questioned hydroxychloroquine, from March 1st to June 22nd, 2020. The preliminary list (a total of 33 events) was narrowed down by searching each event on Google News. Events that returned more than 10 Google News results were deemed as significant events to include in the final timeline (Appendix A). Table 1 shows the search strategy and Boolean operators used to determine the number of Google News results for each event.

Search results were filtered to be within the same time period as the day of the event. For example, in reference to event I which occurred on March 19th, 2020 (Appendix A), the keywords were "Trump" AND "Hydroxychloroquine", and search results were restricted to the time period of March 19th from 12:00 AM to 11:59 PM. This method yielded a total of 20 significant events, and events were categorized together based on their relation to similar entities.

Tweets posted within two days from each of the 20 significant events were grouped. For example, event I (Appendix A) from the timeline was represented through grouping together all the tweets between March 19th at 12:00 AM to March 20th, 2020 at 11:59 PM. This process was repeated for every event and each event was labelled by the respective source. In the case of event I, tweets representing this event would be labelled under the category of the "US federal government". There are three groups for analysis: the US federal government (i.e. POTUS and the White House); publicly recognized health organizations (e.g. the Food and Drug Administration (FDA), the National Institute of Health (NIH), the National Institute of Allergy and Infectious Diseases (NIAID), and the WHO); and research publications freely accessible to the public. Despite Dr. Fauci's affiliation with both the NIAID and the White House Coronavirus Task Force, Dr. Fauci's statement on March 20th, 2020 (event II) was categorized under organizations given that most news articles within March through a Google search characterized Fauci as being the director of NIAID, rather than a member

of the White House Coronavirus Task Force (data not shown). An exception was made for event VII because it was an odd instance where it did not fit into one of the three main categories, and the sample size would be too small if it had a category of its own. The top 15 events that generated the most results were highlighted as the major events to include on the graphs.

Data extraction from Twitter

For this study, the frequency and extent of public conversation regarding hydroxychloroquine are measured by the number of hydroxychloroquine-related tweets over time. Data from Twitter were obtained following the extraction method described in an open-source Github repository created by Irekponor [28]. Tweets were extracted for the period from March 12th at 12:00 AM to June 22nd, 2020 at 11:59 PM. Only tweets published in English with the hashtag or keyword "hydroxychloroquine", the most common way of referring to the drug in the US, were included. Modifications were made to Irekponor's script to remove replies to tweets. Timestamp, hashtags used, text, links, username, and the number of replies or retweets of a given tweet were outputted.

A literature review was conducted to compile a preliminary list of other potential drugs used for treatment towards COVID-19. The data extraction procedure above was applied to the preliminary list of other potential drugs. Drugs with a sufficient number of tweets for analysis (i.e. 10,000 tweets or more) during this time period were chosen to be analyzed further (i.e. azithromycin, remdesivir, dexamethasone).

Data extraction of prescription rates

For this study, physician behaviour is defined as the prevalence of American physicians prescribing hydroxychloroquine for COVID-19. Data showing the prevalence of American physicians who have prescribed hydroxychloroquine, remdesivir, and azithromycin to treat COVID-19 in patients from March 25th to May 24th, 2020 were extracted from Sermo's website [29]. Original data from Sermo was derived through surveying thousands of physicians across the world with a questionnaire asking questions regarding certain treatment regimes or feelings towards COVID-19 on a weekly basis.

Sentiment analysis

For this study, public attitudes are measured by the sentiment of hydroxychloroquine-related tweets with relation to public sources. All links, numbers, and user tags were removed from the text of each tweet before the sentiment calculation. The sentiment of the tweet was analyzed using three different sentiment analysis programs, Textblob v.15.3 [30], Flair [31], and VADER (Valence Aware Dictionary and sentiment Reasoner) [32]. VADER was chosen as our final sentiment analysis program following comparison of the results between the three tools (Appendix B). VADER produces a sentiment score ranging between -1.0 and 1.0. Tweets with a sentiment score of less than 0, greater than 0, and 0, were defined to have negative, positive, and neutral sentiment, respectively.

Data analysis and visualization

Twitter data with polarity scores were imported into R software (R Core Team, Vienna, Austria) for data wrangling [33]. 108,915 duplicate tweets were filtered out. Three visualizations were made using the package ggplot2 [43]: a line graph showing the total number of tweets for hydroxychloroquine, azithromycin, remdesivir, and dexamethasone with respect to time; a line graph showing the proportion of positive and negative sentiment tweets with respect to time; and a scatter plot showing the sentiment values of hydroxychloroquine with respect to sentiment of azithromycin, remdesivir, and dexamethasone.

Tweets falling on the days of the 15 defined significant events (Appendix A) were sorted based on their association to the three public sources. Average sentiment for 24 hours was calculated for each of these groups and was treated as one sample under a given public source. Samples considered statistical outliers were removed using the Interquartile Range (IQR) method. The relative change in sentiment for a sample was calculated through comparison to the sentiment obtained from the previous day. Again, for event I, the absolute relative change in sentiment was calculated by following Equation 1. The means of these values were plotted on ggplot2 to its respective public source.

$$\begin{aligned} & \text{Relative \% change} \\ &= \frac{(\text{Average sentiment within 48 hours from event}) - (\text{Average sentiment of previous day})}{\text{Average sentiment of previous day}} \\ & \times 100\% \end{aligned}$$

(1)

Similarly, the average total number of daily tweets was calculated based on the event groups under each public source. Samples deemed as statistical outliers were removed using the IQR method. These values were visualized on ggplot2 for each of the three public sources. Tweets falling on days between March 12th and June 22nd, 2020 not part of the samples analyzed nor statistical outliers were used as the control data, which is compared with the mean absolute values of relative change and the mean total number of daily tweets.

The ggplot2 package was again used to plot the change in prescription rates of hydroxychloroquine, azithromycin, and remdesivir over time. Texts of positive and negative tweets were inputted into Tableau (Salesforce, Seattle, US) [34] for word cloud generation, displaying the top 40 most frequently occurring words. Stop words (e.g. "a", "the"), and frequently occurring words (e.g. "hydroxychloroquine", "chloroquine", "coronavirus", "COVID") were excluded (Appendix D).

Statistical analysis

One-way analysis of variance (ANOVA) was performed on the average total number of daily tweets with respect to the three public sources and the control. The data was first evaluated to determine whether ANOVA assumptions including homogeneity of variance across groups and normality of data distribution were met. Multiple pairwise comparisons between the means were then conducted through determining the Tukey Honest Significant Difference (HSD). This process was repeated for the absolute values of the mean relative change in sentiment.

The data was evaluated for the assumptions underlying linear regression (linearity, homoscedasticity, independence, and normality) through the generation of diagnostic plots (data not shown). Regression analyses were then performed between the number of tweets associated with hydroxychloroquine and the number of tweets associated with azithromycin, remdesivir, or dexamethasone using a linear model. These analyses were repeated for the prescription rates between hydroxychloroquine and azithromycin or remdesivir, as well as sentiment values between hydroxychloroquine and azithromycin, remdesivir, or dexamethasone.

RESULTS

Impact of major events on the number of tweets related to azithromycin, dexamethasone, hydroxychloroquine, and remdesivir

In total, 500,309 tweets were extracted. The first significant event (event I), where the POTUS supported hydroxychloroquine for the treatment of COVID-19, occurred during the first notable increase in the number of hydroxychloroquine-related tweets (Figure 1). Similarly, the POTUS's following announcement of his hydroxychloroquine regimen as a preventative measure on May 18th, 2020 (event VIII) occurred one day before the single largest increase in hydroxychloroquine tweets in the timeframe studied. Other peaks throughout this time period generally occurred on the same day or following day of an event, with the key exception of April 6th, 2020.

When comparing hydroxychloroquine-related tweets with azithromycin, dexamethasone, and remdesivir-related tweets, hydroxychloroquine-related tweets almost always have a higher number of tweets on any given day than other drug-related tweets. Azithromycin-related tweets' first and only peak on March 21st, 2020 occurred one day after Dr. Anthony Fauci denied the effectiveness of hydroxychloroquine (event II). Dexamethasone-related tweets' first and only peak on June 16th, 2020 occurred one day after the FDA revokes

emergency use authorization of hydroxychloroquine for COVID-19 (event XVIII). Remdesivir-related tweets had two peaks in April; however, no public sources made significant announcements on hydroxychloroquine prior to or during the peak.

Impact of public sources on changes in the number of daily hydroxychloroquine-related tweets

The US federal government showed a significantly higher average change in the number of tweets than the control ($p < 0.001$), organizations ($p < 0.001$), and publications ($p < 0.05$) (Figure 2). Although organizations and research publications each have a larger average number of tweets than the control, both differences are not statistically significant ($p > 0.05$).

Impact of major events on the percentage of physicians using hydroxychloroquine, azithromycin, and remdesivir

An increase in the percentage of physicians prescribing hydroxychloroquine and azithromycin for COVID-19 treatment occurred during the beginning of April 2020, with the biggest spike occurring over the week of April 6th (Figure 3). The approval of the emergency use of hydroxychloroquine by the FDA (event III) and the re-endorsement of hydroxychloroquine by the POTUS (event IV) transpired over the same time period

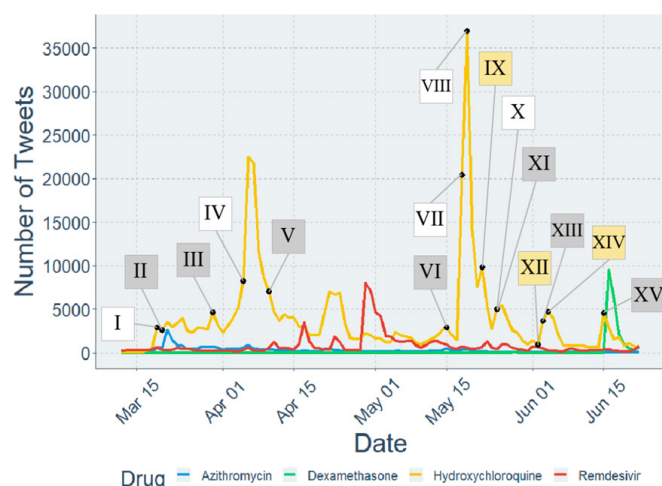


Figure 1. Number of tweets with respect to time. Colors of the different lines represent the different drug-related tweets. Roman numerals on plot signify the top fifteen events relevant to the communication of hydroxychloroquine information from public sources. White, grey, and yellow-filled boxes represent events associated with the US federal government, publicly recognized organizations, and research publications respectively.

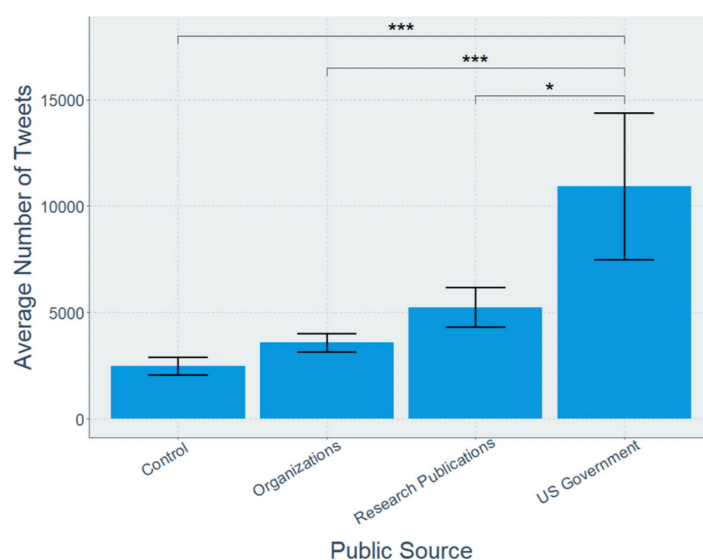


Figure 2. Average number of hydroxychloroquine-related tweets associated with public sources. Bars represent the average number of daily tweets based on 48-hour time periods after significant events associated with each type of public source. Control represents the average number of tweets based on days with no significant events. Error bars represent standard error. *** $p < 0.001$, * $p < 0.05$.

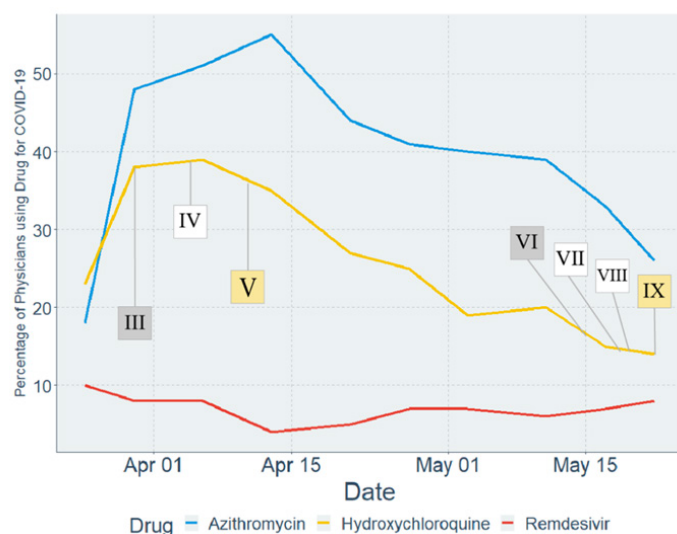


Figure 3. Percentage of physicians prescribing hydroxychloroquine, azithromycin, and remdesivir for COVID-19 treatment over time. Roman numerals on plot represent the events from the list of the top 15 most significant times when public sources communicated information relevant to hydroxychloroquine that occurred within this time period. White, grey, and yellow-filled boxes represent events associated with the US federal government, publicly recognized organizations, and research publications respectively.

leading up to the increase in the percentage of prescribing physicians. Meanwhile, declines in hydroxychloroquine and azithromycin use coincided with a study showing a lack of efficacy in hydroxychloroquine use (event VI), statements from the Department of Veteran Affairs (event VII), and Trump's statement of his consumption of hydroxychloroquine (event VIII). Prescription rates for azithromycin shared similar patterns of increase and decrease as hydroxychloroquine ($p < 0.05$) but were consistently greater than hydroxychloroquine after March 30th, 2020. For remdesivir, the percentage of physicians using the drug remained relatively static throughout the same time period.

Impact of major events on changes in the sentiment of hydroxychloroquine-related tweets

The percentage of tweets with positive and negative sentiment reaches a maximum and minimum, respectively, on March 14th, 2020 despite a lack of public sources communicating information prior to or during this extreme polarity (Figure 4). From March 19th to June 3rd, 2020 (events I, III, IV, V, VII, VIII, and XIII), there is a fluctuating downward trend in the percentage of positive sentiment tweets. Inversely, there is a fluctuating upward trend in the percentage of negative sentiment tweets despite different

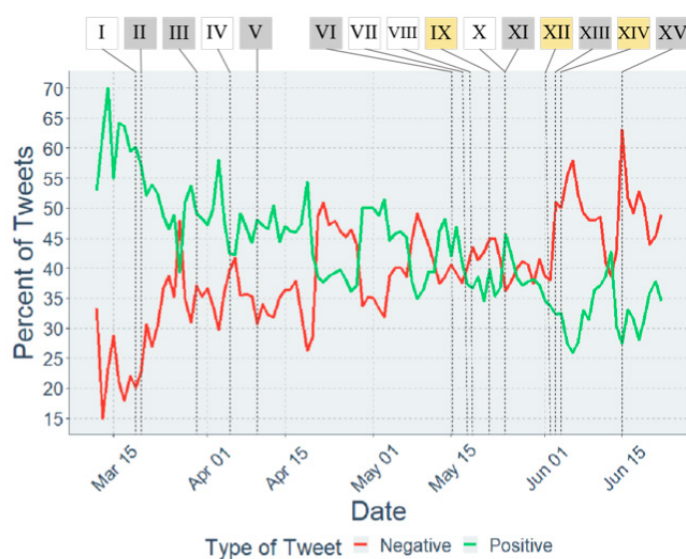


Figure 4. Proportion of positive and negative sentiment hydroxychloroquine-related tweets with respect to time. Roman numerals on the plot signify the top 15 events relevant to the communication of hydroxychloroquine information from public sources. White, grey, and yellow-filled boxes represent events associated with the US federal government, publicly recognized organizations, and research publications, respectively. Note that events X and XI occurred on the same day.

public sources supporting the drug within the same time period. The proportion of negative tweets overtakes the proportion of positive tweets from May 25th onward. The percentage of tweets with negative sentiment reaches its highest level on June 15th, 2020, which is when the FDA revokes emergency use authorization of hydroxychloroquine for COVID-19 (event XV).

Impact of public figures on changes in the sentiment of hydroxychloroquine-related tweets

Similar to the number of tweets, the US federal government had a significantly higher average percentage change in sentiment of tweets than the control ($p < 0.001$), organizations ($p < 0.05$), and publications ($p < 0.05$) (Figure 5). Although organizations have a larger percentage change in the sentiment of tweets than the control group, both the percentage change in sentiment of tweets associated with organization and research publications are statistically insignificant compared to the control ($p > 0.05$).

Analysis of common words used in the positive and negative tweets, after excluding a set of insignificant words

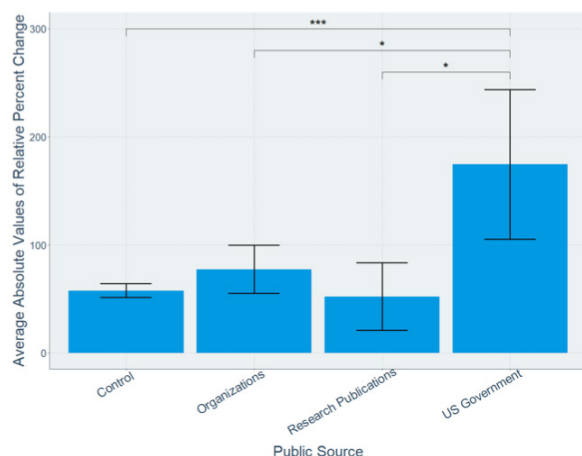


Figure 5. Average absolute values of relative percent change in sentiment of hydroxychloroquine-related tweets with respect to public sources. Control represents the absolute values of relative percent change in sentiment based on days with no significant events. Bars represent the different public sources of hydroxychloroquine information. *** $p < 0.001$, * $p < 0.05$.

"Trump" was the most frequently occurring word for positive sentiment tweets (with 61,895 mentions) and negative sentiment tweets (with 66,102 mentions). Both word clouds included generally similar vocabulary, such as "Twitter", "drug", "patients", and "study". However, some key differences between the word clouds included "effective", "like", and "life" for positive sentiment tweets, and "no", "death", and "media" for negative sentiment tweets.

Association between number of tweets, prescription rates, and sentiment between hydroxychloroquine to azithromycin, dexamethasone, and remdesivir

The number of azithromycin-related tweets were shown to be significantly associated with the number of hydroxychloroquine-related tweets ($p < 0.05$) (Figure 1). Similarly, azithromycin prescriptions were significantly associated with hydroxychloroquine prescriptions ($p < 0.05$) (Figure 3). Further, azithromycin ($p < 0.001$) and remdesivir sentiment ($p < 0.001$) each exhibited a significantly positive association with hydroxychloroquine sentiment (Figure 6). Dexamethasone was not associated with hydroxychloroquine throughout all parameters.

DISCUSSION

Overall, we found that the US federal government had a disproportionate amount of influence on public attitudes and

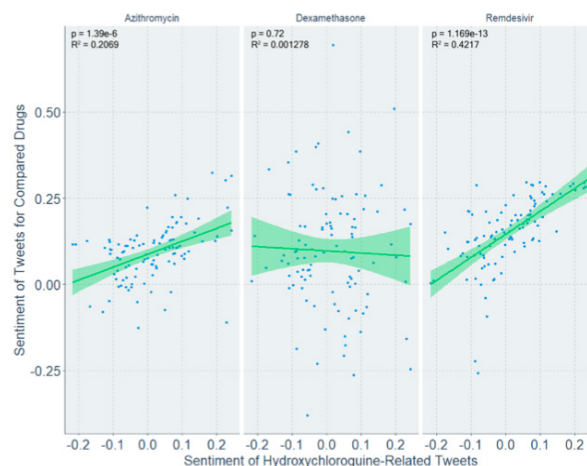


Figure 6. Linear regression analysis of sentiment associated with hydroxychloroquine to sentiment associated with azithromycin, dexamethasone, and remdesivir. Points on the plot represent average sentiment values on a given day for hydroxychloroquine with respect to one of the three drugs.

conversation towards hydroxychloroquine as compared to other evidence-based sources. Usage of hydroxychloroquine for COVID-19 treatment was found to be more discordant with recommendations from public sources, suggesting that physicians may be hesitant towards information from public sources. Additionally, we found that public sources, when disseminating hydroxychloroquine-related information have some influence on public conversation surrounding other drugs, especially azithromycin. Public attitudes towards remdesivir were significantly associated with hydroxychloroquine.

Social media conversation was shown to be driven by public sources. The correspondence between major peaks in hydroxychloroquine-related tweets and communication of hydroxychloroquine-related information from the POTUS (Figure 1) suggests that the US federal government has greater exposure and influence on public conversation on Twitter, as opposed to publicly recognized organizations and research publications. This is supported by the federal government being associated with the greatest number of daily tweets, with a statistically significant difference compared to all the other groups (Figure 2). The word cloud results also suggest the US federal government increases online discussion of hydroxychloroquine, given that "Trump" emerged as the most frequently occurring word amongst all tweets regardless of sentiment (Appendix C). This aligns with the overall popularity of the different entities where the POTUS has over 80 million followers on Twitter, in

comparison to health organizations like the WHO or research journals like the Lancet which only garner over 8 million and 500 thousand followers, respectively [35]. Aside from a strong online following, his impact on online discussions is further perpetuated by coverage from cable news segments, Google search results, and online news stories [36].

After an initial increase in the percentage of physicians using hydroxychloroquine during continued endorsement of hydroxychloroquine from POTUS (event V), the percentage steadily decreased to return to pre-infodemic rates independent of further communication from public sources (Figure 3). This suggests that physicians may have been receptive to recommendations for hydroxychloroquine usage from public sources, but quickly shifted prescription practices. Previous research on primary health information sources among American physicians has been mixed [37] have found that most physicians obtain human papillomavirus (HPV) vaccine information from professional organizations while Research and Markets [38] has found that most physicians use health information websites and physician websites. Regardless of the exact source, it seems that American physicians are not strongly influenced by information disseminated from public sources.

Regarding public sentiment, some public sources have influenced attitudes towards hydroxychloroquine. The gradual progression from a greater proportion of positive sentiment tweets to a greater proportion of negative sentiment tweets indicates that public sentiment is dynamic and can change over time (Figure 4). The reverse in the proportion of positive to negative tweets following the POTUS' last significant promotion of hydroxychloroquine (event XI) also points to a greater impact demonstrated by the US federal government. This is further supported by Figure 5, where the US government was found to have a statistically significant impact on the sentiment of tweets compared to the control and the other public sources, while organizations and research publications showed statistically insignificant changes in sentiment compared to the control. Such findings could be partly explained by the partisanship of the different entities. According to a poll conducted by Hamel et al. [39], 86% of Republicans trust the POTUS for coronavirus information compared to only 8% of Democrats. Meanwhile, organizations such as the CDC yield trust from 60% and 74% of the Republicans and Democrats, respectively. Given the

polarization of trust towards the POTUS, this likely translates to greater extremes in sentiment for statements made by Trump. As such, the POTUS' absolute influence on attitudes would be greater. The US government's disproportionate influence on hydroxychloroquine attitudes raises the concern that world leaders may misuse their platforms to make misleading or false claims [40].

In addition to hydroxychloroquine, public sources have influenced public discussion of other drugs. Most drugs showed little similarity in peak patterns for the number of tweets compared to hydroxychloroquine, suggesting that attention and public discussion towards certain drugs on Twitter are independent of hydroxychloroquine-related information communicated by public figures. Dexamethasone, for instance, only experienced a peak in tweets later in June 2020, which was likely attributed to dexamethasone-related studies pointing towards its efficacy in COVID-19 treatment [41]. This was supported by regression analyses in the number of tweets, which showed only a statistically significant association between hydroxychloroquine and azithromycin-related tweets. These trends were also reflected in the percentage of physicians who prescribe each drug, where azithromycin was the only drug positively correlated to hydroxychloroquine. This association likely occurs because they are frequently taken together as a combination therapy in COVID-19 studies [42, 43]. As such, it can be inferred effects on conversations caused by public sources will be parallel among drugs that have a co-relationship with hydroxychloroquine.

Public sources may exert a greater influence on public sentiment towards similar drugs than public conversations. While a positive correlation between azithromycin and hydroxychloroquine could be expected due to their recognition as potential combination drugs [43, 44], the unexpected correlation between remdesivir and hydroxychloroquine suggests that information disseminated by public sources have a far-reaching influence on public attitudes towards tangentially-related drugs. Multivariable regression analyses could be used to investigate whether public attitudes are solely dependent on hydroxychloroquine-related information being disclosed by public sources, or rather confounded by other events.

In analyzing public attitudes, we obtained tweets using Irekponor's method [29] paired with our sentiment analysis

pipeline to quickly obtain an unlimited number of historical tweets. We also used the Google search engine to determine significant events because of its user-friendly interface and ability to filter specific dates. Replies to tweets were not extracted for analysis because we found that they were often incomprehensible without the original tweet to provide context. We also found that for replies where the original tweet has negative sentiment, those tweets were prone to receiving an incorrect sentiment score (e.g. a positive sentiment reply to a negative sentiment tweet is usually voicing support for the negative sentiment in that tweet and should instead be categorized as negative sentiment towards hydroxychloroquine).

We evaluated three different sentiment analysis tools (Textblob v.15.3 [31], Flair [32], and VADER [33]) and chose the tool that yielded the most optimal sentiment results. These pre-trained models were selected for their convenience and accuracy [44]. After determining the average daily sentiment of each program (Appendix B), we ultimately chose VADER. Textblob did not follow the general trend of the other two programs, suggesting inaccuracies in its classification. Conversely, Flair had extremely polar sentiment which we believe was unrealistic. VADER was chosen as it is specifically designed to analyze social media data, whereas both Textblob and Flair are models trained on movie reviews. Its use was supported by a similar study that used VADER to analyze public sentiment on Twitter for topics of misinformation [45]. One analysis has suggested that custom trained models using Flair are more accurate as compared to Textblob and VADER [46]; however, the test was not conducted using the pre-trained model of Flair that we used. The accuracy of VADER was further confirmed by the words depicted by the word cloud. Words generated for the positive sentiment included words such as "use", "effective", "saved", and "life" that is consistent with strongly supportive language; meanwhile, negative sentiment had more strongly opposing language such as "side", "effects", "emergency", "anti" and "die".

A limitation of this study is that certain peaks in the number of tweets or sentiment can-not be directly linked to communication of hydroxychloroquine-related information from a public source. Notably, the first major spike in hydroxychloroquine-related tweets on April 7th, 2020 could not be explained with existing data. Upon further investigation of other factors, we found that POTUS'

potential financial links to Sanofi, a hydroxychloroquine manufacturer, were publicly revealed on the same day as the spike in tweets [47]. This suggests that interest in the drug is not completely accounted for by the direct communication of hydroxychloroquine-related information from public sources but may also be impacted by third-party news events involving public sources and the drug. Additionally, we assumed that public sources' impact on public behaviour and sentiment was short-term, thus we only associated each respective public source with tweets up to 48 hours after an event. However, it is quite possible that their impact is longer-lasting, and that people will tweet in response days later. There is also an inherent flaw to sentiment analysis as it can be hard for the algorithm to detect sarcasm and irony, negations, jokes, and exaggerations [48], which may change the meaning of the content. As such, the conclusions derived from this study are based on the assumption that Twitter discussion reflects public beliefs expressed in real life and other media. Another potential limitation are Twitter bots and trolls on Twitter, which have been found to skew the online public discourse on vaccinations [49] - similar skew may have occurred in discourse on hydroxychloroquine. This limitation could also be a factor influencing sentiment and conversation [50] have found that humans are vulnerable to retweeting misinformation from bots, which can further exacerbate the infodemic.

CONCLUSION

Unlike health organizations and research publications, the US federal government has a statistically significant influence on public attitudes and conversations. POTUS in particular contributed several times to a notable influx in the number of tweets, fostering a larger abundance of conflicting information being disseminated on social media. This can be especially concerning given POTUS's tendency to make false and unsubstantiated claims [41]. Public sentiment gradually shifted from positive to negative, demonstrating that public attitudes towards topics are open to change. We also found that information related to hydroxychloroquine can affect public attitudes towards related treatments (i.e. hydroxychloroquine and remdesivir), which suggests the extensive effects of misleading public health information and its potential impact on similar topics.

Based on our findings, we recommend three solutions: education, stricter regulation of the dissemination of online

information, and cooperation between sources to amplify specific public health messaging (Appendix E). A combination of these solutions would help to address the effects of the infodemic and disproportionate influence of the US government. In the future higher accuracy in results could be obtained using a custom trained model such as Flair word embeddings [32] or a NLTK (Natural Language Toolkit) [51] naive bayes classifier. It would also be valuable to extend the analysis to other social media platforms. The sentiment analysis pipeline employed in this study may be applied to public health issues beyond COVID-19 where an infodemic is prevalent like vaccinations or the cure for cancer.

ACKNOWLEDGEMENTS

We would like to thank our mentor, Anish Verma, for his mentorship and support. We also thank the entire STEM Fellowship team for organizing this valuable learning opportunity. Finally, we would like to extend our gratitude to the partners, sponsors, and judges for their contribution in bringing the 2020 National Undergraduate Big Data Challenge to reality.

REFERENCES

1. World Health Organization. Coronavirus disease (COVID-19) pandemic. WHO [Internet]. 2020 May [cited 2020 May 30]. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>
2. Zarocostas J. How to fight an infodemic. The Lancet [Internet]. 2020 Feb 29 [cited 2020 Jul 6]; 395(10225):676. Available from: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30461-X/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30461-X/fulltext)
3. Coronavirus disease 2019 (COVID-19) Situation Report - 85. WHO [Internet]. 2020 Apr 15 [cited 2020 Jul 6] Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200415-sitrep-86-covid-19.pdf?sfvrsn=c615ea20_4
4. Rosenberg H, Syed S, Rezaie S. The Twitter pandemic: The critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. CJEM [Internet]. 2020 Apr 6 [cited 2020 Jul 6]. Available from: <https://www.cambridge.org/core/journals/canadian-journal-of-emergency-medicine/article/twitter-pandemic-the-critical-role-of-twitter-in-the-dissemination-of-medical-information-and-misinformation-during-the-covid19-pandemic/9F42C2D9-9CA00FBAE50A66D107322211>
5. 2011 Jan 8 [cited 2020 Jul 6];42(2):145-53. Available from: <https://link.springer.com/article/10.1007/s12016-010-8243-x>
6. MA S Daniel,MS, MPP S Vanessa, CIC L Elaine L.RN, PhD,FAAN,. Dissemination of health information through social networks: Twitter and antibiotics. AJIC [Internet]. 2010 Apr 11 [cited 2020 Jul 6] 38(3):182-8. Available from:[https://www.ajicjournal.org/article/S0196-6553\(10\)00034-9/fulltext](https://www.ajicjournal.org/article/S0196-6553(10)00034-9/fulltext)
7. Mostafa MM. Clustering halal food consumers: A Twitter sentiment analysis. Sage [Internet]. 2018 Apr 24 [cited 2020 Jul 6];61(3):320-37. Available from: <https://journals.sagepub.com/doi/abs/10.1177/1470785318771451>
8. Kušen E, Strembeck M. Politics, sentiments, and misinformation: An analysis of the Twitter discussion on the 2016 Austrian Presidential Elections. Online Social Networks and Media [Internet]. 2018 Mar 4 [cited 2020 Jul 6];5:37-50. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S2468696417301088>
9. Du J, Xu J, Song H-Y, Tao C. Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. BMC [Internet]. 2017 Jul 5 [cited 2020 Jul 6];17(Suppl 2):69-70. Available from: <https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-017-0469-6>
10. Gupta S. Sentiment Analysis: Concept, Analysis and Applications. Towards Data Science [Internet]. 2018 Jan 7 [cited 2020 Jul 6]. Available from: <https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications-6c94d6f58c17>
11. Devaux CA, Rolain J-M, Colson P, Raoult D. New insights on the antiviral effects of chloroquine against coronavirus: what to expect for COVID-19? Int J Antimicrob Agents [Internet]. 2020 Mar 12 [cited 2020 Jul 6]. Available from: <https://www.sciencedirect.com/science/article/pii/S0924857920300881#bib0006>
12. Keyaerts E, Li S, Vijgen L, Rysman E, Verbeeck J, Ranst MV, et al. Antiviral Activity of Chloroquine against Human Coronavirus OC43 Infection in Newborn Mice. Antimicrob Agents Chemother [Internet]. 2009 Jul 17 [cited 2020 Jul 6];53(8):3416-21. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2715625/?tool=pmcentrez&report=abstract>
13. Ben-Zvi I, Ben-Zvi I, Kivity S, Kivity S, Langevitz P, Langevitz P, et al. Hydroxychloroquine: From Malaria to Autoimmunity. Clin Rev Allergy Immunol [Internet]. 2011 Jan 8 [cited 2020 Jul 6];42(2):145-53. Available from: <https://link.springer.com/article/10.1007/s12016-010-8243-x>
14. Vincent MJ, Bergeron E, Benjannet S, Erickson BR, Rollin PE, Ksiazek TG, et al. Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. Virol J [Internet]. 2005 Fall [cited 2020 Jul 6];2(1):69. Available from: <https://virologyj.biomedcentral.com/articles/10.1186/1743-422X-2-69>
15. Mittal A. Battling COVID-19: using old weapons for a

- new enemy. *Trop Dis Travel Med Vaccines* [Internet]. 2020 May 20 [cited 2020 Jul 6];6(1):6. Available from: <https://tdtmvjournal.biomedcentral.com/articles/10.1186/s40794-020-00107-1>
16. Gul MH, Htun ZM, Shaikat N, Imran M. Potential specific therapies in COVID-19. *Ther Adv Respir Dis*. SAGE [Internet]. 2020 May 21 [cited 2020 Jul 6];14AD;14:1753466620926853. Available from: <https://journals.sagepub.com/doi/full/10.1177/1753466620926853>
 17. Veronese N, Demurtas J, Yang L. Use of Corticosteroids in Coronavirus Disease 2019 Pneumonia: A Systematic Review of the Literature. *Frontiers* [Internet]. 2020 Apr 24 [cited 2020 Jul 6]. Available from: <https://www.frontiersin.org/articles/10.3389/fmed.2020.00170/full>
 18. Marinella MA. Routine antiemetic prophylaxis with dexamethasone during COVID-19: Should oncologists reconsider? *J Oncol Pharm Pract* [Internet]. 2020 Jun 6 [cited 2020 Jul 6]; 26(6). Available from: <https://journals.sagepub.com/doi/10.1177/1078155220931921>
 19. Yu B, Li C, Chen P, Zhou N, Wang L, Li J, et al. Low dose of hydroxychloroquine reduces fatality of critically ill patients with COVID-19. *Sci China Life Sci* [Internet]. 2020 Summer [cited 2020 Jul 6];63, 1515-1521. Available from: <https://pubmed.ncbi.nlm.nih.gov/32418114/>
 20. Lagier J-C, Million M, Gautret P, Colson P. Outcomes of 3,737 COVID-19 patients treated with hydroxychloroquine/azithromycin and other regimens in Marseille, France: A retrospective analysis. *Travel Med Infect Dis* [Internet]. 2020 Jul [cited 2020 Jul 6];6AD;101791. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7315163/>
 21. Million M, Lagier J-C, Gautret P, Colson P, Fournier P-E, Amrane S, et al. Early treatment of COVID-19 patients with hydroxychloroquine and azithromycin: A retrospective analysis of 1061 cases in Marseille, France. *Travel Med Infect Dis* [Internet]. 2020 May 5 [cited 2020 Jul 6];35:101738. Available from: <https://www.sciencedirect.com/science/article/pii/S1477893920302179?via%3Dihub>
 22. Gautret P, Lagier J-C, Parola P, Hoang VT. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *Int J Antimicrob Agents* [Internet]. 2020 Jul [cited 2020 Jul 6]; 3AD;105949. <https://www.sciencedirect.com/science/article/abs/pii/S0924857920300996?via%3Dihub>
 23. Pacheco RL, Riera R. Hydroxychloroquine and chloroquine for COVID-19 infection. *Rapid systematic review*. *Journals Bahiana* [Internet]. 2020 [cited 2020 Jul 6];3AD;2(1). <https://www5.bahiana.edu.br/index.php/evidence/article/view/2843>
 24. Prajapat M, Sarma P, Shekhar N, Avti P, Sinha S, Kaur H, et al. Drug targets for corona virus: A systematic review. 2020 Mar 11;52:56. Available from: <https://link-gale-com.ezproxy.library.ubc.ca/apps/doc/A617690979/HRCA?u=ubcolumbia&sid=HRCA&xid=a950693e>
 25. Jun C, Danping L, Li L, Ping L, Qingnian X, Lu X, et al. A pilot study of hydroxychloroquine in treatment of patients with moderate COVID-19. 2020 Mar 6;49(2):215-9. Available from: <http://www.zjujournals.com/med/CN/10.3785/j.issn.1008-9292.2020.03.03>
 26. Worldometer. COVID-19 Coronavirus Pandemic [Internet]. 2020 [cited 2020 Sep 15]. Available from: https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1?
 27. Nielsen PRK, Howard DPN, Simon F, Brennen DJS. Types, sources, and claims of COVID-19 misinformation [Internet]. *Reuter's Institute*; 2020 Apr 7 [cited 2020 May 30] Available from: https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation?fbclid=IwAR14s8tkzNVV9L1EmnCrzIFB7xoQLgFEuCzsKFhEwQF_nWgsmRVa0kFLS6U
 28. Victor Irekponor. Optimized-modified-getoldtweets3-omgot, April 16 2020.
 29. COVID-19 Trend Report: HCP Infections [Internet]. 2020 [cited 2020 May 30]. Available from: <https://app.sermo.com/covid19-barometer/global>
 30. Loria S. TextBlob. 2019 [cited 2020 May 27]. Available from: <https://textblob.readthedocs.io/en/dev/#>
 31. Akbik A, Blythe D, Vollgraf R. Contextual String Embeddings for Sequence Labeling. 2018;1638-49. <https://www.aclweb.org/anthology/C18-1139.pdf>
 32. Gilbert E, Hutto CJ. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. 2014 Jun; (Eighth International Conference on Weblogs and Social Media (ICWSM-14)). Available from: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/download/8109/8122/>
 33. Team RC. R: A Language and Environment for Statistical Computing. 2020.
 34. Tableau Software. LCC. Tableau. 2020.
 35. Twitter. Twitter [Internet]. [cited 2020 Sep 30]. (Twitter). Available from: www.twitter.com
 36. Staff T. The 25 Most Influential People on the Internet [Internet]. *TIME*; 2019 Jul 16 [cited 2020 Sep 15]. Available from: <https://time.com/5626827/the-25-most-influential-people-on-the-internet/>
 37. Bynum SA, Malo TL, Lee J-H, Guiliano AR, Vadaparampil ST. HPV vaccine information-seeking behaviors among US physicians: Government, media, or colleagues? *Vaccine* [Internet]. 2011 Jul 8 [cited 2020 Jul 6];29(32):5090-3. Available from: <https://www.clinicalkey-com.ezproxy.library.ubc.ca/#!/content/playContent/1-s2.0-S0264410X11007419?returnurl=null&referrer=null>
 38. ProQuest. Research and Markets: Physicians are using general health information websites as their primary sources of health information online [Internet]. 2001 Mar 1 [cited 2020 Jul 6]. Available from: serialssolutions.com
 39. Hamel L, Kearney A, Kirzinger A, Lopes L, Muñana C, Brodie M. KFF Health Tracking Poll - September 2020: Top Issues in 2020 Election, The Role of Misinformation, and Views on A Potential Coronavirus Vaccine. KFF [Internet].

- 2020 [cited 2020 Sep 30]. Available from: <https://www.kff.org/coronavirus-covid-19/report/kff-health-tracking-poll-september-2020/>
40. Factcheck.org. Person: Donald Trump [Internet]. Fact Check; 2020 [cited 2020 Jul 6]. Available from: <https://www.factcheck.org/person/donald-trump/>
 41. Kupferschmidt K. One U.K. trial is transforming COVID-19 treatment. Why haven't others delivered more results? [Internet]. Science Mag; 2020 Jul 2 [cited 2020 Jul 6]. Available from: <https://www.sciencemag.org/news/2020/07/one-uk-trial-transforming-covid-19-treatment-why-haven-t-others-delivered-more-results>
 42. Juurink DN. Safety considerations with chloroquine, hydroxychloroquine and azithromycin in the management of SARS-CoV-2 infection. Can Med Assoc J [Internet]. 2020 Apr 27 [cited 2020 May 28];27;192(17):E450-3. Available from: <https://www.cmaj.ca/content/192/17/E450>
 43. Mégarbane B, Scherrmann J. Hydroxychloroquine and Azithromycin to Treat Patients With COVID-19: Both Friends and Foes? J Clin Pharmacol [Internet]. 2020 Summer [cited 2020 May 28];60(7):808-14. Available from: <https://accp1.onlinelibrary.wiley.com/doi/full/10.1002/jcph.1646>
 44. Terry-Jack M. NLP: Pre-trained Sentiment Analysis. Medium, [Internet]. 2019 May 1 [cited 2020 May 28]. Available from: <https://medium.com/@b.terryjack/nlp-pre-trained-sentiment-analysis-1eb52a9d742c>
 45. Garay J, Yap R, Sabellano MJ. An analysis on the insights of the anti-vaccine movement from social media posts using k-means clustering algorithm and VADER sentiment analyzer. IOP Conf Ser: Mater Sci Eng. IOP Science, [Internet] 2019 [cited 2020 May 28];482. Available from: <https://iopscience.iop.org/article/10.1088/1757-899X/482/1/012043/pdf>
 46. Rao P. Fine-grained Sentiment Analysis in Python (Part 1). Toward Data Science, [Internet]. 2019 Sept 4 [cited 2020 May 28]. Available from: <https://towardsdatascience.com/fine-grained-sentiment-analysis-in-python-part-1-2697bb111ed4>
 47. Trump Has 'Small,' 'Distant Link' To Sanofi, French Drugmaker Of Hydroxychloroquine [Internet]. Forbes; 2020 Apr 7 [cited 2020 Jul 6]. Available from: <https://www.forbes.com/sites/lisettevoytko/2020/04/07/trump-has-small-distant-link-to-sanofi-french-drugmaker-of-hydroxychloroquine/#35709f687260>
 48. Martin R. Can sentiment analysis detect sarcasm on the internet? Oh yeah, totally... [Internet]. Infegy; 2016 Jun 8 [cited 2020 May 28]. Available from: <https://blog.infegy.com/sentiment-analysis-and-sarcasm>
 49. PhD B David A, MPH J Amelia M, MAA, SM Q SiHua, SM A Lulwah, PhD C Tao, MS B Adrian, et al. Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. Am J Public Health [Internet]. 2018 Oct [cited 2020 Jul 6];108(10):1378-84. Available from: <https://ajph.aphapublications.org/doi/10.2105/AJPH.2018.304567>
 50. Shao C, Ciampaglia GL, Varol O, Yang K-C, Flammini A, Menczer F. The spread of low-credibility content by social bots. Nat Commun [Internet]. 2018 Nov 20 [cited 2020 Jul 6];9(1):4787-9. Available from: <https://www.nature.com/articles/s41467-018-06930-7>
 51. Perkins J. NLTK (Natural Language Toolkit) Trainer. 2020 Apr.
 52. House W. President Pence, and Members of the Coronavirus Task Force in Press Briefing [Internet]. White House; 2020 Mar 19 [cited 2020 Jul 6]. Available from: <https://www.whitehouse.gov/briefings-statements/remarks-president-trump-vice-president-pence-members-coronavirus-task-force-press-briefing-6/>
 53. House W. Remarks by President Trump, Vice President Pence, and Members of the Coronavirus Task Force in Press Briefing [Internet]. White House; 2020 Mar 20 [cited 2020 Jul 6]. Available from: <https://www.whitehouse.gov/briefings-statements/remarks-president-trump-vice-president-pence-members-coronavirus-task-force-press-briefing/>
 54. FDA Approves Emergency Use Of Malaria Drug Trump Touted Despite Scant Evidence That It Works [Internet]. KHN Morning Briefing; 2020 Mar 30 [cited 2020 May 30]. Available from: <https://khn.org/morning-breakout/fda-approves-emergency-use-of-malaria-drug-trump-touted-despite-scant-evidence-that-it-works/>
 55. House W. Remarks by President Trump, Vice President Pence, and Members of the Coronavirus Task Force in Press Briefing [Internet]. White House; 2020 Apr 6 [cited 2020 Jul 6]. Available from: <https://www.whitehouse.gov/briefings-statements/remarks-president-trump-vice-president-pence-members-coronavirus-task-force-press-briefing-20/>
 56. Gulf News. Coronavirus: Clinical trial of hydroxychloroquine for COVID-19 treatment begins in US [Internet]. Gulf News; 2020 Apr 10 [cited 2020 May 28]. Available from: <https://gulfnews.com/world/americas/coronavirus-clinical-trial-of-hydroxychloroquine-for-covid-19-treatment-begins-in-us-1.1586491544130>
 57. Magagnoli J, Narendran S, Pereira F, Cummings T, Hardin JW, Sutton SS, et al. Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19. Med [Internet]. 2020 [cited 2020 May 30]. Available from: <https://www.medrxiv.org/content/10.1101/2020.04.16.20065920v2>
 58. Yen H. U.S. Veterans Affairs says it won't stop use of unproven malaria drug on vets [Internet]. CTV News; 2020 May 15 [cited 2020 May 30]. Available from: <https://www.ctvnews.ca/health/coronavirus/u-s-veterans-affairs-says-it-won-t-stop-use-of-unproven-malaria-drug-on-vets-1.4941648>
 59. O'Reilly A. Trump reveals he's taking hydroxychloroquine in effort to prevent coronavirus symptoms [Internet]. Fox News; 2020 May 18 [cited 2020 May 30]. Available from: <https://www.foxnews.com/politics/trump-reveals-taking-hydroxychloroquine-coronavirus>

60. House W. Remarks by President Trump in Cabinet Meeting [Internet]. White House; 2020 May 19 [cited 2020 Jul 6]. Available from: <https://www.whitehouse.gov/briefings-statements/remarks-president-trump-cabinet-meeting-17/>
61. Mehra MR, Desai SS, Ruschitzka F, Patel AN. Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *The Lancet* [Internet]. 2020 May 22 [cited 2020 May 28]. Available from: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)31180-6/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31180-6/fulltext)
62. Sorace S. WHO temporarily halts hydroxychloroquine, chloroquine trials for coronavirus treatment pending safety review [Internet]. Fox News; 2020 May 25 [cited 2020 May 30]. Available from: <https://www.foxnews.com/health/who-hydroxychloroquine-chloroquine-coronavirus-safety-review>
63. Timm JC. Trump says he's no longer taking hydroxychloroquine [Internet]. NBC News; 2020 May 25 [cited 2020 Jul 6]. Available from: <https://www.cnbc.com/2020/05/25/trump-says-hes-no-longer-taking-hydroxychloroquine.html>
64. Spetalnick M. U.S. sends Brazil 2 million doses of hydroxychloroquine, drug touted by Trump [Internet]. National Post; 2020 May 31 [cited 2020 Jul 6]. Available from: <https://nationalpost.com/pmnh/health-pmn/u-s-sends-brazil-2-million-doses-of-hydroxychloroquine-drug-touted-by-trump>
65. Herper M. Top medical journals raise concerns about data in two studies related to Covid-19 [Internet]. STAT; 2020 Jun 2 [cited 2020 Jul 6]. Available from: <https://www.statnews.com/2020/06/02/top-medical-journals-raise-concerns-about-data-in-two-studies-related-to-covid-19/>
66. Joseph M. Vinetz. Lack of efficacy of hydroxychloroquine in covid-19. *BMJ* [Internet]. 2018 Summer [cited 2020 Jul 6]; 369. Available from: <https://www.bmj-com.ezproxy.library.ubc.ca/content/369/bmj.m2018>
67. Boulware DR, Pullen MF, Bangdiwala AS, Pastick KA, Lofgren SM, Okafor EC, et al. A Randomized Trial of Hydroxychloroquine as Postexposure Prophylaxis for Covid-19. *The New England Journal of Medicine* [Internet]. 2020 Fall [cited 2020 Jul 6]; 383:517-525. Available from: <https://www.nejm-org.ezproxy.library.ubc.ca/doi/10.1056/NEJMoa2016638>
68. Joseph A. Lancet, New England Journal retract Covid-19 studies, including one that raised safety concerns about malaria drugs [Internet]. STAT; 2020 Jun 4 [cited 2020 Jul 6]. Available from: <https://www.statnews.com/2020/06/04/lancet-retracts-major-covid-19-paper-that-raised-safety-concerns-about-malaria-drugs/comment-page-3/>
69. Reuters T. FDA revokes emergency use authorization of hydroxychloroquine for COVID-19 [Internet]. CBC; 2020 Jun 15 [cited 2020 Jul 6]. Available from: <https://www.cbc.ca/news/world/fda-revokes-hydroxychloroquine-covid-19-1.5612499>
70. Mole B. WHO gives up on hydroxychloroquine for COVID-19, stops trials [Internet]. Ars Technica; 2020 Jun 17 [cited 2020 Jul 6]. Available from: <https://arstechnica.com/science/2020/06/who-gives-up-on-hydroxychloroquine-for-covid-19-stops-trials/>
71. Jackson H. Coronavirus: U.S. National Institutes of Health halts trial of hydroxychloroquine [Internet]. Global News; 2020 Jun 20 [cited 2020 Jul 6]. Available from: <https://globalnews.ca/news/7088606/coronavirus-nih-hydroxychloroquine-trial/>
72. Greater Langley Chamber of Commerce. Events Calendar - Chamber Event [Internet]. Greater Langley; 2020 [cited 2020 Jul 6]. Available from: <https://www.langleychamber.com/events/catgid/3>
73. City of Vancouver. Educational activities in Vancouver [Internet]. City of Vancouver; 2020 [cited 2020 Jul 6]. Available from: <https://vancouver.ca/news-calendar/calendar-of-events.aspx>
74. Bomey N. Helping seniors fight "infodemic": Poynter's MediaWise combats misinformation [Internet]. USA Today; 2020 Jun 24 [cited 2020 Jul 6]. Available from: <https://www.usatoday.com/story/money/2020/06/24/poynter-mediawise-aarp-facebook-misinformation-fake-news-seniors/3241164001/>
75. YouTube Help. Information panel giving topical context [Internet]. YouTube Help; 2020 [cited 2020 Jul 6]. Available from: <https://support.google.com/youtube/answer/9004474?hl=en-GB>
76. YouTube Help. Coronavirus disease 2019 (COVID-19) updates [Internet]. YouTube Help; 2020 Jul 6 [cited 2020 Jul 6]. Available from: <https://support.google.com/youtube/answer/9777243?hl=en>
77. Roth Y, Pickles N. Updating our Approach to Misleading Information [Internet]. Twitter; 2020 May 11 [cited 2020 Jul 6]. Available from: https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html

APPENDIX A

Table A.1. Timeline of Significant Events. Top 15 events shown on the left.

Significant Events	Date (2020)	Event
I	March 19	President Trump publicly supports the use of hydroxychloroquine as a COVID-19 treatment for the first time and says that the drug is "not going to kill anybody" [52]
II	March 20	Dr. Anthony Fauci warns against following President Trump's medical advice on hydroxychloroquine [53]
III	March 30	The Food and Drug Administration (FDA) approves emergency use of hydroxychloroquine [54]
IV	April 5	President Trump voices positive outlook on hydroxychloroquine as a COVID-19 treatment and reiterates that the drug "doesn't kill people" [55]
V	April 10	The National Institute of Health (NIH) begins a clinical trial to test the effectiveness of hydroxychloroquine as a COVID-19 treatment [56]
	April 23	A study on outcomes of hydroxychloroquine usage finds no evidence that mechanical ventilation risk is reduced in American veterans hospitalized with COVID-19. [57]
VI	May 15	The US Department of Veterans Affairs (VA) states that fewer of its patients are now taking hydroxychloroquine and acknowledges that VA Secretary Robert Wilkie previously wrongly asserted that hydroxychloroquine had shown benefit on younger veterans. [58]
VII	May 18	President Trump announces that he has been taking hydroxychloroquine as a preventative measure for COVID-19 and is optimistic for its success [59]
VIII	May 19	President Trump defends his promotion of the drug, saying he "worked with doctors" and that he did not know that the FDA issued a warning that the drug "should not be used outside of a hospital setting" [60]

Significant Events	Date (2020)	Event
IX	May 22	The Lancet publishes a study that finds no benefit of hydroxychloroquine for COVID-19 [61]
X	May 25	WHO temporarily halts hydroxychloroquine trials for COVID-19 to allow experts to review evidence concerning safety of the drug [62]
XI	May 25	President Trump completes his cycle of taking hydroxychloroquine [63]
	May 31	US sends Brazil 2 million doses of hydroxychloroquine [64]
XII	June 2	The Lancet editor and other experts raise concerns about the reliability of the Lancet study published on May 22 (Event IX) [65]
XIII	June 3	The WHO resumes hydroxychloroquine trials for COVID-19 after reviewing safety concerns [66]
	June 3	The New England Journal of Medicine publishes a study that finds that hydroxychloroquine did not prevent infection after exposure when used as post-exposure prophylaxis within four days [67]
XIV	June 4	The Lancet retracts its hydroxychloroquine study published on May 22 (Event X) due to concerns on the veracity of primary data [68]
XV	June 15	The FDA revokes emergency use authorization of hydroxychloroquine for COVID-19 due to growing evidence on the ineffectiveness of the drug [69]
	June 17	The WHO halts trials on hydroxychloroquine as evidence emerges that the drug does not lower risk of death and provides no benefit to hospitalized COVID-19 patients [70]
	June 20	The NIH halts trial of hydroxychloroquine due to the drug's unlikely benefits to hospitalized COVID-19 patients [71]

APPENDIX B



Figure B.1. Comparison of average sentiment of tweets over time using different sentiment analysis tools. Graph was generated by plotting the calculated average daily sentiment of Textblob v.15.3 [31], a pre-trained classification model from Flair [32], and VADER [33].

APPENDIX C

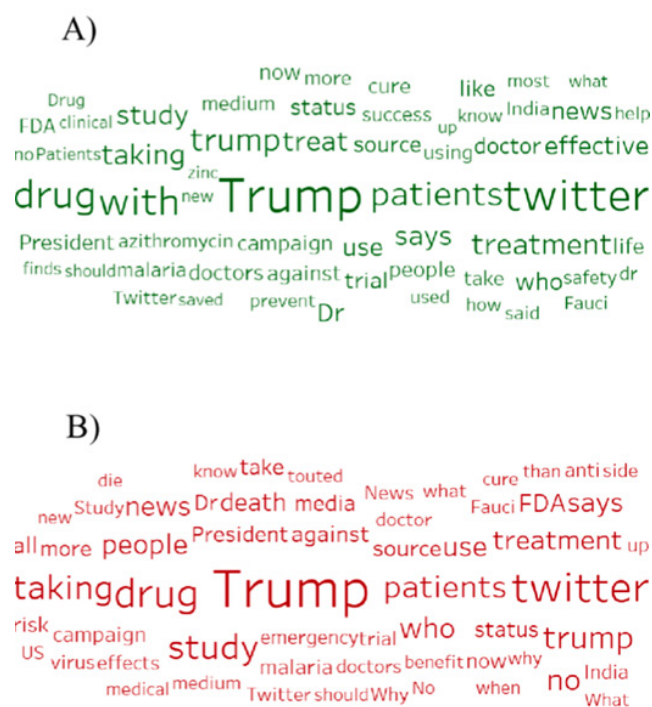


Figure C.1. Visualizations of top 60 most frequently occurring words from tweets categorized as A) positive

and B) negative sentiment. Word clouds were generated on Tableau v.20.2.25110 [35] using the tweets generated from March 12th to June 22nd, 2020 with the hashtag or keyword "hydroxychloroquine". Tweets were analyzed for sentiment using VADER [33]. Size of the word corresponds to its frequency.

APPENDIX D

Table D.1. List of excluded words from word clouds generated in Appendix C. Word cloud was generated on Tableau v.20.2.25110 using the tweets generated from March 12th to June 22nd, 2020, with the hashtag or keyword "hydroxychloroquine". Tweets were analyzed for sentiment using VADER [33].

a	get	or
A	has	out
about	have	over
after	He	pic
all	he	s
an	her	so
and	him	t
are	his	that
as	html	The
at	https	the
be	hydroxychloroquine	their
because	Hydroxychloroquine	there
been	I	they
but	if	This
by	If	this
can	in	to
chloroquine	Is	us
com	is	utm
Coronavirus	it	via
coronavirus	It	was
covid	its	we
Covid	just	We
COVID	my	will
d	not	with
do	of	would
for	on	you
from	one	your

APPENDIX E

Supplementary research regarding possible solutions.

Based on our findings, we propose solutions under three themes to be implemented in conjunction.

Education

The first solution revolves around education and focuses on helping individuals develop life-long critical thinking and digital literacy skills. This ideally would take the form of creating innovative educational content to engage the audience. Free courses and seminars should be offered by local organizations and community centers to provide access for people of all ages and demographics. The goal would be to provide relevant skills in making accurate judgements about the legitimacy of any news source or statements made by public sources. For example, in Langley, BC, Canada, free seminars on communication, awareness, and other versatile topics are regularly offered by the local Chamber of Commerce [72]. Similarly, the City of Vancouver, BC, Canada, also offers numerous free seminars on a variety of topics annually [73]. However, neither of these two organizations cover topics regarding misinformation. Thus, we believe that the incorporation of a digital literacy program to these accessible platforms would be a great way to begin engaging and educating the general public on recognizing and identifying sources of misinformation. This solution is also being considered by big organizations like Facebook, which has funded and vowed to promote online educational videos created by MediaWise on misinformation to help seniors "fight infodemic" and "combat misinformation" [74]. In addition to educational seminars and online material, such content could be offered as part of the school curricula to develop crucial habits and digital literacy skills beginning from early childhood. This long-lasting solution would be effective in combating the current infodemic as well as any others in the future. The major shortcoming of this approach is that it will likely take a long period of time for positive effects to show, which ignores the immediate issue of ongoing misinformation.

Stricter regulation of online misinformation

To mitigate this, we propose a secondary solution focused on Twitter and other social media companies in enforcing stricter regulation of online misinformation. For example, in response to COVID-19, YouTube has curbed misinformation on the disease by positioning an information panel at the top of the page that gives topical context of COVID-19 whenever COVID-19-related keywords are searched on YouTube [75]. YouTube has also "updated its Community Guidelines to

include a page on COVID-19 mis-information" [76]. Similarly, Twitter has acted by placing labels and warning messages on some disputed and misleading tweets [77]. Social media companies should continue their current practices created in response to COVID-19 for future contentious topics with the potential to cause public harm. A potential issue with this approach is that it heavily depends on the willingness of social media companies to invest and take initiatives. However, it is in the companies' best interest to prevent, address, and mitigate misinformation to avoid potential backlash from misinformed users.

Cooperation between sources

Our final solution is based on the theme of cooperation. We propose that organizations partner with entities (i.e. President Trump or prominent celebrities) that have large followings to boost evidence-based public sources' reach and influence as well as show a broader spectrum of views. These three proposed themes, when implemented together, can provide long-term solutions and address any current misinformation to mitigate shortcomings that a single approach alone would present.