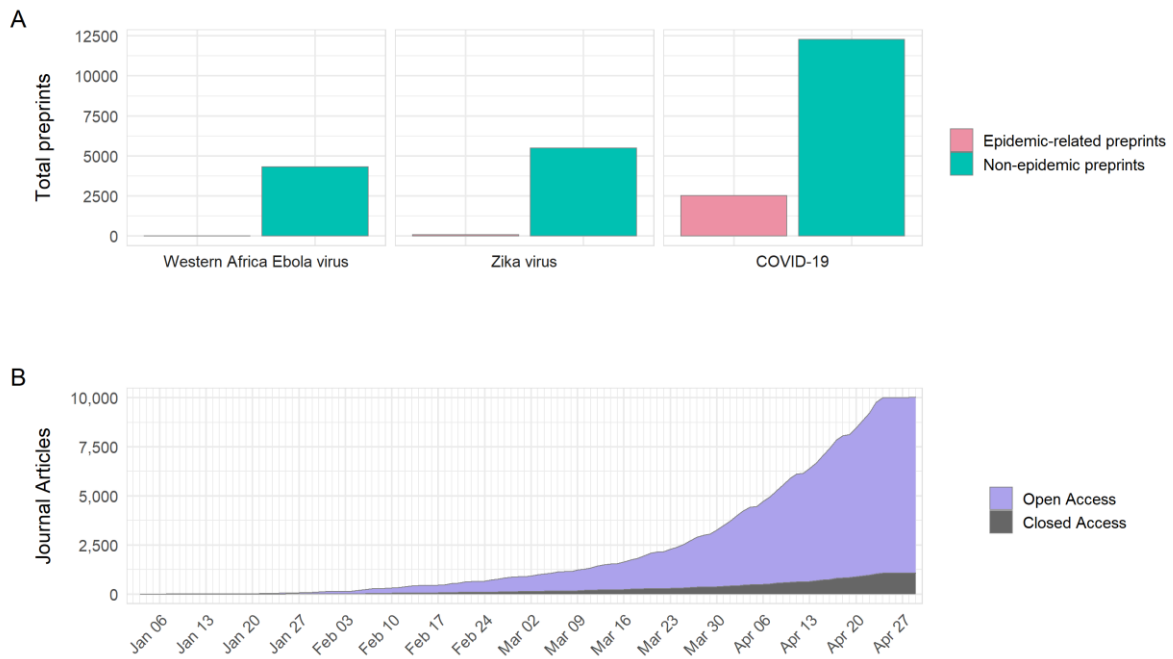
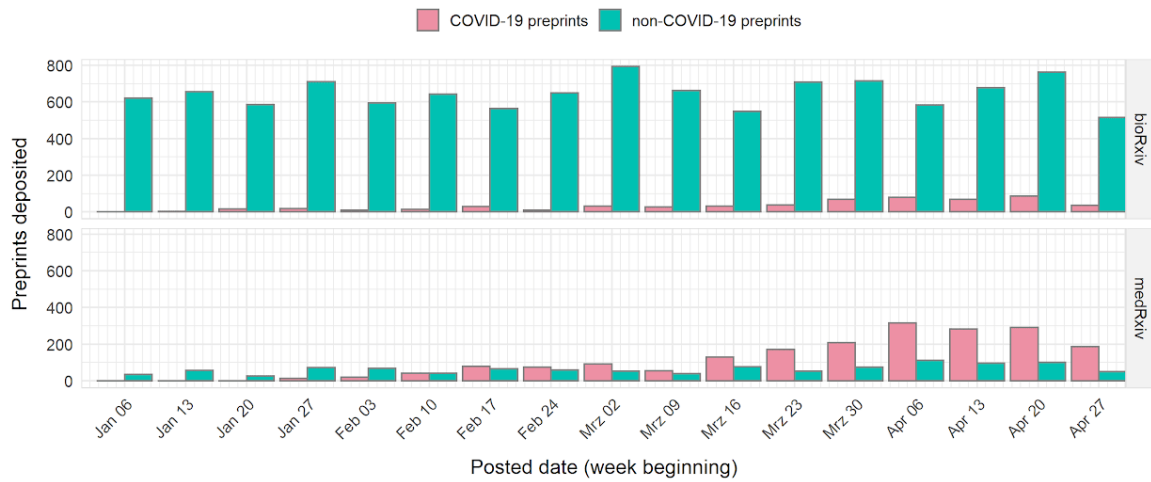


## Supplemental Figures

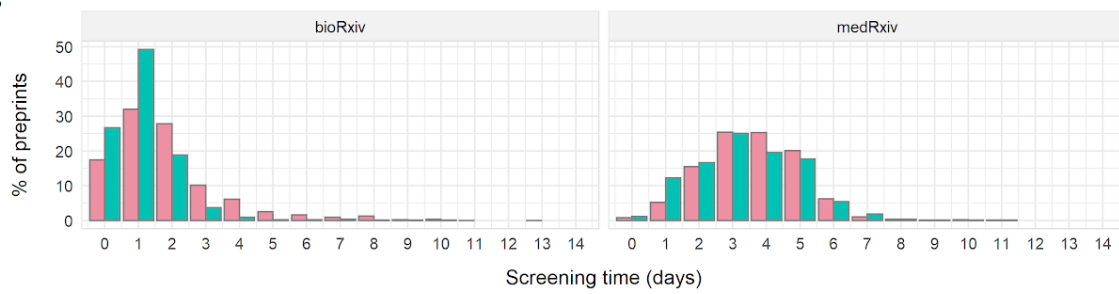


**Supplementary Figure 1. Preprint posting volume for COVID-19 in additional contexts (A)** Total preprints for Western Africa Ebola virus, Zika virus and COVID-19. The number of preprints posted that were related to the epidemic, and the number that were posted but not related to the epidemic in the same time period are shown. Periods of data collection for Western Africa Ebola virus (24th January 2014 - 9th June 2016) and Zika virus (2nd March 2015 - 18th November 2016) correspond to the periods between the first official medical report, and World Health Organization end of Public Health Emergency of International Concern declaration. The period of data collection for COVID-19 refers to the analysis period used in this study, 1st January 2020 to 30th April 2020. **(B)** Comparison of COVID-19 journal articles based on accessibility (open access vs closed access).

A



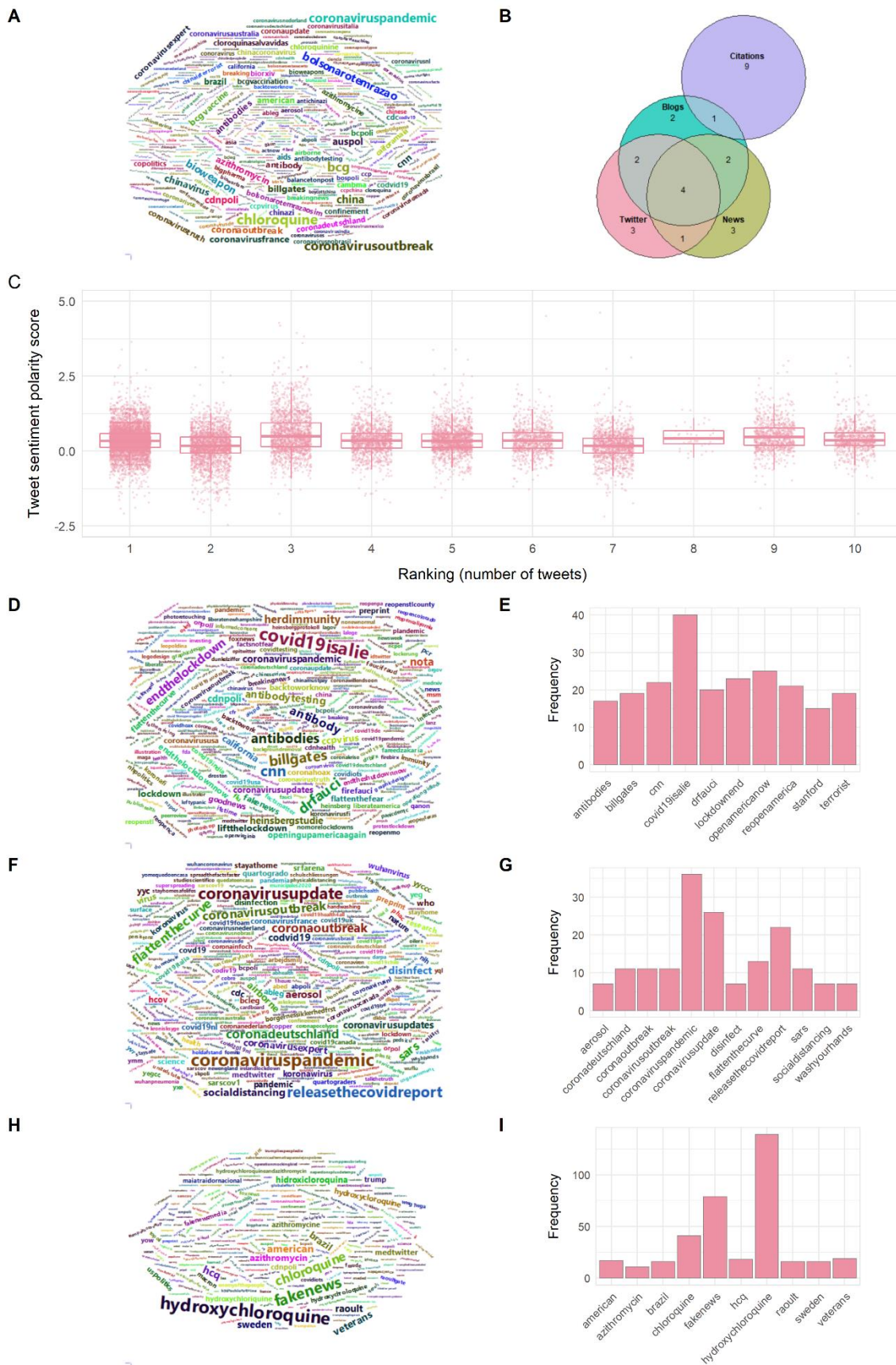
B



**Supplementary Figure 2. Preprint attributes broken down by server for bioRxiv and medRxiv.** (A) Number of preprints posted to each server in a given week. (B) Screening time for preprints for bioRxiv and medRxiv.

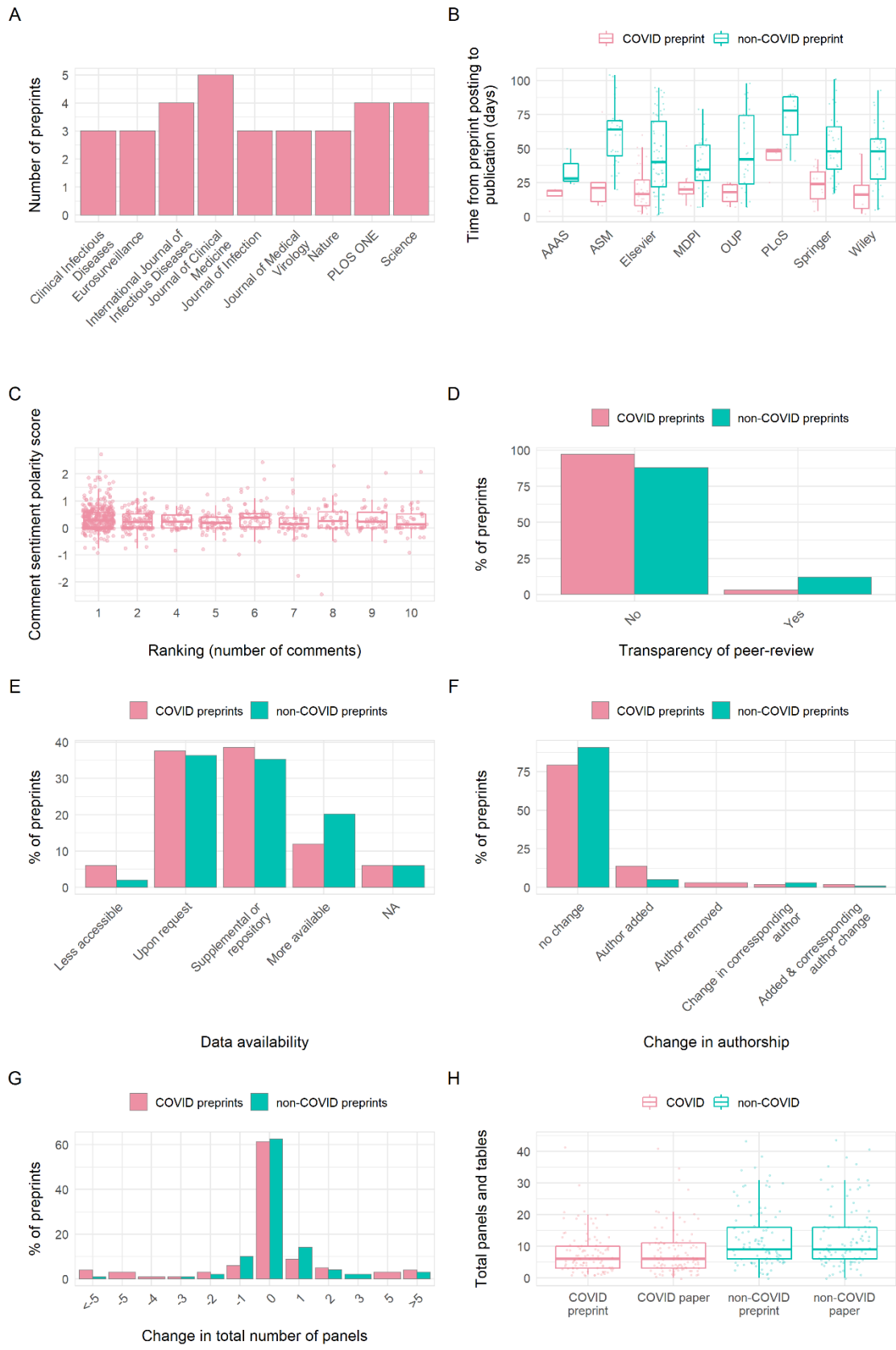


**Supplementary Figure 3. Additional abstract view and PDF download data.** (A) Abstract views for non-COVID-19 preprints posted on bioRxiv and medRxiv between September 2019 and April 2020. (B) PDF downloads for non-COVID-19 preprints posted on bioRxiv and medRxiv between September 2019 and April 2020. (C) Comparison of PDF downloads for COVID-19 and non-COVID-19 preprints across multiple preprint servers. Red shaded area in (A) and (B) represents the pandemic period. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to 1.5\*IQR. All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility.



**Supplemental Figure 4. Analysis of Twitter hashtags associated with COVID-19 preprints. (A)** Wordcloud of hashtags for the 10 most tweeted COVID-19 preprints. **(B)** Venn diagram showing

overlap between the 10 most tweeted COVID-19 preprints, 10 most covered COVID-19 preprints in the news, 10 most blogged about preprints and the 10 most cited COVID-19 preprints. (C) Sentiment score for the 10 most tweeted COVID-19 preprints. (D) Wordcloud of hashtags associated with the preprint "COVID-19 Antibody Seroprevalence in Santa Clara County, California" (highest tweeted, 3rd highest news coverage). (E) Frequency of the 10 most used hashtags in (D). (F) Wordcloud of hashtags associated with the preprint "Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1" (5th highest tweeted, highest news coverage). (G) Frequency of the 10 most used hashtags in (F). (H) Wordcloud of hashtags associated with the preprint "Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19" (7th highest tweeted, 2nd highest news coverage). (I) Frequency of the 10 most used hashtags in (H). Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to  $1.5 \times \text{IQR}$ . All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility.



**Supplemental Figure 5. Preprint-publication pairs are not significantly different.** (A) Destination journals for COVID-19 preprints that were published within our timeframe. (B) Time from posting on bioRxiv or medRxiv to publication broken down by publisher. (C) Sentiment scores of the top 10

preprints with the most comments, excluding rank three due to paper retraction (Table 2). (D) Presence of open peer-reviews for preprint-paper pairs. (E) Data availability for preprint-paper pairs. (F) Changes in author lists for preprint-paper pairs. (G) Difference in the total number of panels and tables for between preprint and published article. (H) Total number of panels and figures per preprint or paper. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to  $1.5 \times \text{IQR}$ . All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility.