Spread of Misinformation About Face Masks and COVID-19 by Automated Software on Facebook

The dangers of misinformation spreading on social media during the COVID-19 pandemic are known. However, software that allows individuals to generate automated content and share it via counterfeit accounts (or “bots”) to amplify misinformation has been overlooked, including how automated software can be used to disseminate original research while undermining scientific communication.

We analyzed conversations on public Facebook groups, a platform known to be susceptible to automated misinformation, concerning the publication of the Danish Study to Assess Face Masks for the Protection Against COVID-19 Infection (DANMASK-19) to explore automated misinformation.

We selected DANMASK-19 because it was widely discussed (it was the fifth most shared research article of all time as of March 2021 according to Altmetric) and demonstrated that masks are an important public health measure to control the pandemic.

Methods | Sample. We obtained the names of 563 Facebook groups in which a link to the publication of DANMASK-19 on the Annals of Internal Medicine website was posted and downloaded all available posts (N = 299 925) from these groups using CrowdTangle (crowdtangle.com). We limited our study period to the 5 days following the publication of DANMASK-19.

Table 1. Examples of Misinformation From Posts Made to Public Facebook Groups That Contained a Link to DANMASK-19

<table>
<thead>
<tr>
<th>Labeled outcomes*</th>
<th>Excerpt example*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claimed mask wearing harms the wearer</td>
<td>It appears that not only does wearing a mask not provide meaningful protection against SARS-CoV-2, but also leads to an increase in infections with other respiratory viruses!</td>
</tr>
<tr>
<td>Labeled outcomesa</td>
<td>Claimed mask wearing is a political/corporate control of DANMASK-19 (labeled as yes/no). Additionally, did not claim mask-wearing harms or cite conspiracies</td>
</tr>
<tr>
<td>Conspiratorial claims</td>
<td>Corporate fact-checkers are lying to you! All this to serve their Dystopian #Agenda2030 propaganda!</td>
</tr>
<tr>
<td>Did not claim mask-wearing harms or cite conspiracies</td>
<td>The recommendation to wear surgical masks in addition to other public health measures did not reduce the rate of SARS-CoV-2 infection among individuals preferring to impose their behavior in all public spaces...[These are scientists paid by world elites who shamelessly lie to billions of people!]</td>
</tr>
</tbody>
</table>

Abbreviation: DANMASK-19, Danish Study to Assess Face Masks for the Protection Against COVID-19 Infection.

* Misinformation was of 2 types: (1) incorrectly stating the conclusions of DANMASK-19 by claiming wearing a mask harmed the wearer (labeled as yes/no) and (2) promoting conspiracies, such as claims of covert political/corporate control of DANMASK-19 (labeled as yes/no). Additionally, we present examples of posts that made neither claim (labeled as yes/no).

b Excerpts are taken from the actual public Facebook posts as examples of the specific types of content that was labeled. Posts were edited for length and to remove any potentially identifying information.

The dangers of misinformation spreading on social media during the COVID-19 pandemic are known. However, software that allows individuals to generate automated content and share it via counterfeit accounts (or “bots”) to amplify misinformation has been overlooked, including how automated software can be used to disseminate original research while undermining scientific communication.

We analyzed conversations on public Facebook groups, a platform known to be susceptible to automated misinformation, concerning the publication of the Danish Study to Assess Face Masks for the Protection Against COVID-19 Infection (DANMASK-19) to explore automated misinformation. We selected DANMASK-19 because it was widely discussed (it was the fifth most shared research article of all time as of March 2021 according to Altmetric) and demonstrated that masks are an important public health measure to control the pandemic.

Methods | Sample. We obtained the names of 563 Facebook groups in which a link to the publication of DANMASK-19 on the Annals of Internal Medicine website was posted and downloaded all available posts (N = 299 925) from these groups using CrowdTangle (crowdtangle.com). We limited our study period to the 5 days following the publication of DANMASK-19.

Table 1. Examples of Misinformation From Posts Made to Public Facebook Groups That Contained a Link to DANMASK-19

<table>
<thead>
<tr>
<th>Labeled outcomes*</th>
<th>Excerpt example*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claimed mask wearing harms the wearer</td>
<td>It appears that not only does wearing a mask not provide meaningful protection against SARS-CoV-2, but also leads to an increase in infections with other respiratory viruses!</td>
</tr>
<tr>
<td>Labeled outcomesa</td>
<td>Claimed mask wearing is a political/corporate control of DANMASK-19 (labeled as yes/no). Additionally, did not claim mask-wearing harms or cite conspiracies</td>
</tr>
<tr>
<td>Conspiratorial claims</td>
<td>Corporate fact-checkers are lying to you! All this to serve their Dystopian #Agenda2030 propaganda!</td>
</tr>
<tr>
<td>Did not claim mask-wearing harms or cite conspiracies</td>
<td>The recommendation to wear surgical masks in addition to other public health measures did not reduce the rate of SARS-CoV-2 infection among individuals preferring to impose their behavior in all public spaces...[These are scientists paid by world elites who shamelessly lie to billions of people!]</td>
</tr>
</tbody>
</table>

Abbreviation: DANMASK-19, Danish Study to Assess Face Masks for the Protection Against COVID-19 Infection.

* Misinformation was of 2 types: (1) incorrectly stating the conclusions of DANMASK-19 by claiming wearing a mask harmed the wearer (labeled as yes/no) and (2) promoting conspiracies, such as claims of covert political/corporate control of DANMASK-19 (labeled as yes/no). Additionally, we present examples of posts that made neither claim (labeled as yes/no).

b Excerpts are taken from the actual public Facebook posts as examples of the specific types of content that was labeled. Posts were edited for length and to remove any potentially identifying information.

The dangers of misinformation spreading on social media during the COVID-19 pandemic are known. However, software that allows individuals to generate automated content and share it via counterfeit accounts (or “bots”) to amplify misinformation has been overlooked, including how automated software can be used to disseminate original research while undermining scientific communication.

We analyzed conversations on public Facebook groups, a platform known to be susceptible to automated misinformation, concerning the publication of the Danish Study to Assess Face Masks for the Protection Against COVID-19 Infection (DANMASK-19) to explore automated misinformation. We selected DANMASK-19 because it was widely discussed (it was the fifth most shared research article of all time as of March 2021 according to Altmetric) and demonstrated that masks are an important public health measure to control the pandemic.

Methods | Sample. We obtained the names of 563 Facebook groups in which a link to the publication of DANMASK-19 on the Annals of Internal Medicine website was posted and downloaded all available posts (N = 299 925) from these groups using CrowdTangle (crowdtangle.com). We limited our study period to the 5 days following the publication of DANMASK-19.

Table 1. Examples of Misinformation From Posts Made to Public Facebook Groups That Contained a Link to DANMASK-19

<table>
<thead>
<tr>
<th>Labeled outcomes*</th>
<th>Excerpt example*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claimed mask wearing harms the wearer</td>
<td>It appears that not only does wearing a mask not provide meaningful protection against SARS-CoV-2, but also leads to an increase in infections with other respiratory viruses!</td>
</tr>
<tr>
<td>Labeled outcomesa</td>
<td>Claimed mask wearing is a political/corporate control of DANMASK-19 (labeled as yes/no). Additionally, did not claim mask-wearing harms or cite conspiracies</td>
</tr>
<tr>
<td>Conspiratorial claims</td>
<td>Corporate fact-checkers are lying to you! All this to serve their Dystopian #Agenda2030 propaganda!</td>
</tr>
<tr>
<td>Did not claim mask-wearing harms or cite conspiracies</td>
<td>The recommendation to wear surgical masks in addition to other public health measures did not reduce the rate of SARS-CoV-2 infection among individuals preferring to impose their behavior in all public spaces...[These are scientists paid by world elites who shamelessly lie to billions of people!]</td>
</tr>
</tbody>
</table>

Abbreviation: DANMASK-19, Danish Study to Assess Face Masks for the Protection Against COVID-19 Infection.

* Misinformation was of 2 types: (1) incorrectly stating the conclusions of DANMASK-19 by claiming wearing a mask harmed the wearer (labeled as yes/no) and (2) promoting conspiracies, such as claims of covert political/corporate control of DANMASK-19 (labeled as yes/no). Additionally, we present examples of posts that made neither claim (labeled as yes/no).

b Excerpts are taken from the actual public Facebook posts as examples of the specific types of content that was labeled. Posts were edited for length and to remove any potentially identifying information.
(November 18, 2020, through November 22, 2020) because media interest is typically greatest initially. This study was exempted as not human participants research by the University of California, San Diego Human Research Protections Program. Additional details are provided in the eAppendix in the Supplement.

**Measures.** When identical links are posted in close succession, it suggests that automated software was used.2,3 We identified the subsets of Facebook groups that were the most or least likely to be affected by automation by calculating the frequency that identical links were posted to pairs of Facebook groups and the time that elapsed between these posts for all links (n = 251656) shared during the study period. Adapting past operationalizations,4 a pair of Facebook groups that (1) hosted identical links 5 or more times and (2) at least half of these links being posted within less than 10 seconds would be considered the most affected by automation. Comparatively, Facebook groups in which the total time elapsed between identical links was in the top 90th percentage of time between postings were considered the least affected by automation. Facebook groups that were most affected by automation had a mean (SD) of 4.28 (3.02) seconds between shares of identical links compared with 4.35 (11.71) hours for those least affected by automation.

To quantify the extent to which Facebook groups were subject to misinformation, all posts that linked to DANMASK-19 in the groups most or least affected by automation were qualitatively coded by 2 authors (B.C. and Z.Z.) for 2 types of misinformation: (1) whether the primary conclusion of DANMASK-19 was misrepresented (eg, mask wearing harms the wearer) and (2) whether conspiratorial claims were made about DANMASK-19 (eg, claims of covert political/corporate control). A separate outcome for not including either form of misinformation was computed. Table 1 presents example posts. Coders disagreed on 3.9% of labels (Cohen κ = 0.76) and resolved disagreements unanimously with the first author (J.W.A.).

**Analysis.** The percentage of posts that linked to DANMASK-19 that included each type of misinformation or neither type was calculated separately for the sets of Facebook groups most and least affected by automation along with prevalence ratios comparing these percentages. Statistical significance was set to \( P < .05 \) and 95% confidence intervals were bootstrapped. Analyses were computed with R, version 3.6.1 (R Foundation).

**Results.** A total of 712 posts that provided direct links to DANMASK-19 were shared in 563 public Facebook groups. Of these, 279 posts (39%) that linked to DANMASK-19 were in Facebook groups most affected by automation, of which 17 were deleted and unavailable for further analysis. Sixty-two posts (9%) were made in Facebook groups that were least affected by automation, and 3 were deleted.

Among posts made to groups most affected by automation, 19.8% (95% CI, 14.9%-24.5%) claimed masks harmed the wearer, 50.8% (95% CI, 44.6%-56.5%) made conspiratorial claims about the trial, and 43.9% (95% CI, 37.4%-49.6%) made neither claim (Table 2). In contrast, among posts made to groups least affected by automation, 8.5% (95% CI, 1.7%-15.2%) claimed masks harmed the wearer, 20.3% (95% CI, 10.2%-30.5%) made conspiratorial claims about the trial, and 72.9% (95% CI, 59.3%-81.4%) made neither claim.

The percentage of posts linking to DANMASK-19 that claimed that masks harmed the wearer was 2.3 (95% CI, 1.0-6.5) times higher in Facebook groups that were most affected by automation vs groups that were least affected by automation; conspiratorial claims (prevalence ratio, 2.5; 95% CI, 1.5-4.5) were also higher in Facebook groups that were most affected by automation. Making neither claim was more common in Facebook groups that were least affected by automation (prevalence ratio, 0.6; 95% CI, 0.5-0.7).

**Discussion.** A campaign that presumably used automated software promoted DANMASK-19 on Facebook groups to disseminate misinformation. The limitations of the study in-
clude that the entities responsible for organizing this automated campaign cannot be determined, only public Facebook groups were studied, and only a single high-profile study over a few days was evaluated.

Scientific journals are easy targets of automated software. Possible approaches to prevent misinformation due to dissemination of articles by automated software include legislation that penalizes those behind automation; greater enforcement of rules by social media companies to prohibit automation; and counter-campaigns by health experts.

John W. Ayers, PhD, MA
Brian Chu, BS
Zechariah Zhu
Eric C. Leas, PhD, MPH
Davey M. Smith, MD, MAS
Mark Dredze, PhD
David A. Broniatowski, PhD, SM

Author Affiliations: Division of Infectious Diseases and Global Public Health, University of California, San Diego, La Jolla (Ayers, Smith); Perelman School of Medicine, University of Pennsylvania, Philadelphia (Chu); Department of Mathematics, University of California, Los Angeles (Zhu); Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, La Jolla (Leas); Department of Computer Science, Johns Hopkins University, Baltimore, Maryland (Dredze); Department of Engineering Management and Systems Engineering, The George Washington University, Washington, DC (Broniatowski).

Accepted for Publication: April 15, 2021.

Published Online: June 7, 2021. doi:10.1001/jamainternmed.2021.2498

Corresponding Author: John W. Ayers, PhD MA; #333 CRSF 9500 Gilman Drive, La Jolla, CA 92039 (ayers.john.w@gmail.com).

Author Contributions: Drs Ayers and Broniatowski had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of Interest Disclosures: Dr Ayers reported owning equity in Directing Medicine, Health Watcher, and Good Analytics outside the submitted work.

Role of the Funder/Sponsor: This work was supported by the Burroughs Wellcome Fund, National Institutes of Health grant AI036214, The John S. and James L. Knight Foundation to the George Washington University Institute for Data, Democracy, and Politics, and The John and Mary Tu Foundation.

Funding/Support: This work was supported by the Burroughs Wellcome Fund, National Institutes of Health grant AI036214, The John S. and James L. Knight Foundation to the George Washington University Institute for Data, Democracy, and Politics, and The John and Mary Tu Foundation.

Role of the Funder/Sponsor: The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Sharing Statement: The data used in the study are public in nature and are available from CrowdTangle. CrowdTangle prohibits providing raw data to anyone outside of a CrowdTangle user’s account. Anyone with a CrowdTangle account may access these corresponding data. Researchers may request CrowdTangle access at https://help.crowdtangle.com/en/articles/4302208-crowdtangle-for-academics-and-researchers.

Additional Contributions: We appreciate discussions and editing with Benjamin M. Althouse, PhD, ScM (Gates Foundation), Alicia Nobles, PhD (University of California, San Diego) and Christopher Longhurst MD (University of California San Diego). They were not compensated for their contributions.

LESS IS MORE

Frequency of Administration of Standardized Screening Questions in Federally Qualified Health Centers

Federal funders and managed care insurers often provide financial incentives for health care clinicians to perform standardized screening questionnaires. The goal of rapidly identifying patients who may have depression, anxiety, or substance use disorders or who may need extra help in understanding medical terms is noble. However, the financial incentives may lead to overuse of these screening tools, creating an unnecessary burden for patients and primary care clinicians and decreasing the efficiency of the clinical practice. To assess the potential overuse of screenings, we evaluated the frequency of standardized screenings tied to performance metrics within a national network of federally qualified health centers (FQHCs).

Editor’s Note page 1255

Methods | We retrospectively analyzed electronic health record data to determine the frequency by which patients within 24 FQHCs were asked and completed 6 standardized screening questionnaires (Patient Health Questionnaire-2 for depression, Generalized Anxiety Disorder 2-item for anxiety, tobacco use, passive smoke exposure, preferred learning style, and health literacy) from January 1 through December 31, 2019. The study population included patients 18 years and older with at least 1 visit in 2019. The main outcome was the presence or absence of an excess screening. Definitions of “excessive” were derived from performance metrics for recommended screening frequencies (depression and tobacco use screening: annual, except 6 months for...