YouTube as a source of misinformation on COVID-19 vaccination: a systematic analysis

Heidi Oi-Yee Li, Elena Pastukhova, Olivier Brandts-Longtin, Marcus G Tan, Mark G Kirchhof

ABSTRACT

Introduction Vaccines for SARS-CoV-2 have been accessible to the public since December 2020. However, only 58.3% of Americans are fully vaccinated as of 5 November 2021. Numerous studies have supported YouTube as a source of both reliable and misleading information during the COVID-19 pandemic. Misinformation regarding the safety and efficacy of COVID-19 vaccines has negatively impacted vaccination intent. To date, the literature lacks a systematic evaluation of YouTube’s content on COVID-19 vaccination using validated scoring tools. The objective of this study was to evaluate the accuracy, usability and quality of the most widely viewed YouTube videos on COVID-19 vaccination.

Methods A search on YouTube was performed on 21 July 2021, using keywords ‘COVID-19 vaccine’ on a cleared-cache web browser. Search results were sorted by ‘views’, and the top 150 most-viewed videos were collected and analysed. Duplicate, non-English, non-audiovisual, exceeding 1-hour duration, or videos unrelated to COVID-19 vaccine were excluded. The primary outcome was usability and reliability of videos, analysed using the modified DISCERN (mDISCERN) score, the modified Journal of the American Medical Association (mJAMA) score and the COVID-19 Vaccine Score (CVS).

Results Approximately 11% of YouTube’s most-viewed videos on COVID-19 vaccines, accounting for over 18 million views, contradicted information from the WHO or the Centers for Disease Control and Prevention. Videos containing non-factual information had significantly lower mDISCERN (p<0.001), mJAMA (p<0.01) and CVS (p<0.001) scores compared with videos with factual information. Videos from government sources had higher mJAMA and CVS scores, but averaged three times the ratio of dislikes to likes, whereas videos containing non-factual information averaged 14 times more likes than dislikes.

Conclusion As the COVID-19 pandemic evolves, widespread adoption of vaccination is essential in reducing morbidity, mortality, and returning to some semblance of normalcy. Providing high-quality and engaging health information from reputable sources is essential in addressing vaccine hesitancy.

WHAT IS ALREADY KNOWN?

⇒ Throughout the COVID-19 pandemic, videos on social media questioning the efficacy and safety of COVID-19 vaccines have become more numerous and widespread. It has previously been reported that 27.5% of the most viewed YouTube videos on COVID-19 contained misinformation.

⇒ Information from reputable sources, such as the WHO or the Centers for Disease Control and Prevention, is under-represented on social media.

WHAT ARE THE NEW FINDINGS?

⇒ Approximately 11% of YouTube’s most-viewed videos on COVID-19 vaccines, accounting for over 18 million views, contradicted the reference standard from the WHO, or the Centers for Disease Control and Prevention.

⇒ Information from reputable sources may be perceived as less favourable by the public, when proxied by the ratio of likes to dislikes. On average, YouTube videos from government sources had triple the ratio of dislikes to likes, while entertainment videos had 14 times the ratio of likes to dislikes.

WHAT DO THE NEW FINDINGS IMPLY?

⇒ Despite efforts by YouTube in combatting health misinformation, a substantial portion of highly viewed content contains misinformation regarding COVID-19 vaccines continues to be hosted by the second-largest social media platform.

⇒ Public health officials and clinicians should recognise the impact of YouTube and other social media platforms as sources of health information and misinformation for patients and leverage these platforms to disseminate high-quality content to combat misinformation given poor vaccination uptake.

INTRODUCTION

COVID-19, caused by the novel SARS-CoV-2, was first noted in December 2019 and has evolved into a global pandemic with over 246 million cases worldwide and nearly 5 million deaths to date. Vaccines against SARS-CoV-2 became accessible to the public since December 2020, after being granted emergency use authorisation status by the Federal Drug Administration. Despite this, as of 5 November 2021, only 58.3% of
Americans are fully vaccinated. This is concerning as public trust in vaccination programmes is paramount in reducing morbidity and mortality, limiting the spread of highly transmissible SARS-CoV-2 variants, and putting an end to this pandemic.

In the digital age, more individuals are relying on the internet and social media as sources of healthcare-related information. In particular, YouTube is one of the most predominant social media platforms with billions of views daily and is the second most used search engine after Google. This readily accessible platform offers multiple genres of content, including entertainment, news, consumers, and education, and allows its users to interact with each other in real time.

Numerous studies have supported YouTube as a source of both reliable and misleading information during the COVID-19 pandemic. A previous study demonstrated that 27.5% of the most viewed YouTube videos on COVID-19 contained misinformation. As vaccine misinformation has been shown to negatively impact vaccination intent, propagation of misinformation regarding the safety and efficacy of COVID-19 vaccines on large platforms, such as YouTube, has potential devastating consequences on public health.

To date, the literature lacks a systematic evaluation of YouTube as a source of information on COVID-19 vaccination using validated scoring tools. The objective of this study is to evaluate the accuracy, usability and quality of the most widely viewed YouTube videos on COVID-19 vaccination.

METHODS

Search protocol
A YouTube search was performed on 21 July 2021, using keywords ‘COVID-19 vaccine’ on a cleared-cache web browser. Search results were sorted by ‘views’, and 150 videos with the greatest number of views were collected and analysed.

Study selection and data extraction
Three reviewers (HO-YL, EP and OB-L) reviewed and analysed the 150 videos. Videos that were duplicate, non-English, non-audio, non-visual or unrelated to COVID-19 vaccine were excluded. Discrepancies were resolved by discussion for consensus. Video descriptive characteristics including title, hyperlink, number of views, number of likes and dislikes, upload date, duration and publishing category, were collected.

The reliability and quality of video content were assessed using the modified DISCERN (mDISCERN) score (online supplemental table S1) and the modified Journal of the American Medical Association (mJAMA) criteria (online supplemental table S2). A novel five-point COVID-19 Vaccine Score (CVS), developed based on previously published work, was used to evaluate the usefulness of video content for the average viewer (online supplemental table S3).

Patient and public involvement
No patients or public were involved in any way in this study.

Data analysis
Following data extraction, two independent reviewers (EP and OB-L) categorised all videos into videos containing factual or non-factual information, and into their video publishing category, in duplicate, with a third reviewer (HO-YL) consultation for consensus. Video sources were categorised into healthcare professionals, network news, education, internet news, consumers, entertainment, government sources and newspapers according to previously published work (online supplemental table S4).

Videos were considered non-factual if they contained one or more non-factual statements, defined as contradictory to guidelines published by public health agencies (ie, WHO or Centers for Disease Control and Prevention) at the time of video publication. The CVS score was calculated as follows: if a CVS criterion was discussed, it was given a score of 1 if factually correct, a score of 0.5 if ambiguous and 0 if non-factual (online supplemental table S3). The total score was then divided by the number of items discussed (up to a maximum of 5) and expressed as a proportion. mDISCERN and mJAMA scores were calculated with one point given for each criterion correctly addressed (online supplemental tables S1 and S2).

Descriptive statistics including two tailed t-tests and Mann-Whitney U tests compared factual videos to non-factual videos, with a significance level of p<0.01 to offset multiple hypothesis testing. Percentages, means and ratios compared videos from different sources. All analyses were performed using Microsoft Excel and Mathworks Matlab software (V.R2019b).

RESULTS

Video characteristics
Of 150 videos screened, 122 videos (81%) were included. A total of 28 (19%) videos were excluded: 14 for not discussing any CVS criteria, 13 for non-English and one for non-audio (online supplemental figure S1).

Source of videos
Videos were classified into eight categories (figure 1). The total viewership at time of data collection was 169446382 views. 24.1% of views were from videos produced by healthcare professionals, 20.9% from network news, 16.9% from education, 10.2% from internet news, 8.4% from consumers, 8.1% from entertainment, 7.3% from government sources and 4.2% from newspapers.

Government videos had the highest average CVS score (0.98) and number of CVS items discussed per minute of video (1.89 items/min). Internet news videos had the lowest CVS score (0.83), and consumer videos had the lowest CVS items discussed per minute (0.20 items/min). Internet news also had the lowest mDISCERN and mJAMA scores, whereas government videos had the
highest mJAMA scores and educational videos had the highest mDISCERN scores (table 1).

Government videos had the shortest average time since publication of 65 days, compared with 310 days for newspaper videos. Entertainment videos on average had 14.4 times more likes than dislikes, whereas government videos averaged more dislikes than likes, with 3.1 dislikes to every like.

Videos with factual versus non-factual information

Thirteen videos (10.7%) contained non-factual information and accounted for 11.0% of total viewership, while 109 videos (89.3%) contained only factual information. Of the 13 non-factual videos, 4 were from network news, three from healthcare professionals, two each from education and internet news, and one each from newspaper and consumer sources. Two of the 13 videos (15.4%) containing non-factual information did not contain a YouTube information icon prompting the viewer to seek additional information on COVID-19 vaccination.

When comparing the two groups, there were no significant differences in views per video (p=0.94), likes per video (p=0.40), dislikes per video (p=0.36), duration (p=0.98) and mean days since publication (p=0.56) (table 2). Videos with non-factual information on average had a greater median ratio of likes to dislikes, compared with videos with factual information (13.89 vs. 2.33, p<0.01).

Videos with non-factual information were found to have lower CVS (p<0.001), mDISCERN (p<0.001) and mJAMA (p<0.01) scores, compared with videos with factual information. Classification and examples of non-factual statements are shown in table 3.

**DISCUSSION**

This study highlights that approximately 11% of YouTube’s most viewed videos on COVID-19 vaccines contradicted the reference standard, reaching over 18 million viewers. Our findings were significantly lower than other studies reporting 27.5% and 45% of videos containing non-factual information. This discrepancy is likely due to increased awareness of the impact of social media on disseminating misinformation that has resulted in YouTube’s implementation of a COVID-19 medical misinformation policy in May 2020. Furthermore, government sources are increasingly leveraging YouTube as a valuable opportunity to disseminate vaccine information, with an over twofold increase in the percentage

<table>
<thead>
<tr>
<th>Channel type</th>
<th>% of total views</th>
<th>CVS (SD)</th>
<th>mDISCERN (SD)</th>
<th>mJAMA (SD)</th>
<th>Age (Mean, SD)</th>
<th>Median likes/dislikes (range)</th>
<th>Mean subscribers (M)</th>
<th>Mean views (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health professional n=25 (20.5%)</td>
<td>24.11</td>
<td>0.93 (0.15)</td>
<td>3.7 (0.95)</td>
<td>1.2 (0.66)</td>
<td>116 (75)</td>
<td>2.8 (0.16–22.2)</td>
<td>0.62</td>
<td>1.63</td>
</tr>
<tr>
<td>Entertainment n=7 (5.7%)</td>
<td>8.10</td>
<td>0.95 (0.96)</td>
<td>3.6 (1.40)</td>
<td>1.3 (0.95)</td>
<td>127 (34)</td>
<td>14.4 (0.91–52.2)</td>
<td>6.58</td>
<td>1.96</td>
</tr>
<tr>
<td>Network news n=42 (34.4%)</td>
<td>20.92</td>
<td>0.90 (0.18)</td>
<td>3.0 (1.23)</td>
<td>1.0 (0.68)</td>
<td>99 (85)</td>
<td>0.72 (0.15–20.2)</td>
<td>4.26</td>
<td>0.83</td>
</tr>
<tr>
<td>Internet news n=6 (4.9%)</td>
<td>10.16</td>
<td>0.83 (0.20)</td>
<td>2.0 (2.0)</td>
<td>0.5 (0.84)</td>
<td>103 (64)</td>
<td>15.4 (0.30–23.2)</td>
<td>3.56</td>
<td>2.87</td>
</tr>
<tr>
<td>Government n=9 (7.4%)</td>
<td>7.26</td>
<td>0.98 (0.07)</td>
<td>3.8 (0.97)</td>
<td>1.3 (0.87)</td>
<td>65 (57)</td>
<td>0.32 (0.10–2.96)</td>
<td>0.57</td>
<td>1.37</td>
</tr>
<tr>
<td>Newspaper n=4 (3.3%)</td>
<td>4.22</td>
<td>0.84 (0.20)</td>
<td>3.8 (0.96)</td>
<td>1.3 (0.96)</td>
<td>310 (180)</td>
<td>1.58 (0.71–13.9)</td>
<td>1.71</td>
<td>1.79</td>
</tr>
<tr>
<td>Consumer n=10 (8.2%)</td>
<td>8.38</td>
<td>0.88 (0.20)</td>
<td>3.5 (1.65)</td>
<td>1.1 (0.74)</td>
<td>223 (157)</td>
<td>3.0 (1.07–22.8)</td>
<td>0.91</td>
<td>1.42</td>
</tr>
<tr>
<td>Education n=19 (15.6%)</td>
<td>16.85</td>
<td>0.98 (0.05)</td>
<td>4.3 (1.11)</td>
<td>1.2 (1.17)</td>
<td>167 (78)</td>
<td>9.7 (1.23–35.8)</td>
<td>3.11</td>
<td>1.50</td>
</tr>
</tbody>
</table>

Mean (SD), unless otherwise indicated. M represents millions.

CVS, COVID-19 Vaccine Score; mDISCERN, modified DISCERN; mJAMA, modified Journal of the American Medical Association.
Table 2  Comparison of factual and non-factual videos

<table>
<thead>
<tr>
<th></th>
<th>Non-factual</th>
<th>Factual</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of videos</td>
<td>13 (10.66%)</td>
<td>109 (89.34%)</td>
<td></td>
</tr>
<tr>
<td>No of views</td>
<td>18 482 373 (11%)</td>
<td>150 964 009 (89%)</td>
<td></td>
</tr>
<tr>
<td>Views per video (in millions)</td>
<td>1.421 (1.120)</td>
<td>1.385 (1.677)</td>
<td>0.939</td>
</tr>
<tr>
<td>CVS</td>
<td>0.7 (0.202)</td>
<td>0.948 (0.122)</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>mDISCERN</td>
<td>1.769 (1.235)</td>
<td>3.651 (1.181)</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>mJAMA</td>
<td>0.538 (0.66)</td>
<td>1.183 (0.807)</td>
<td>0.00649</td>
</tr>
<tr>
<td>Median likes/dislikes (range)</td>
<td>13.89 (0.49–22.8)</td>
<td>2.33 (0.10–52.2)*</td>
<td>0.00078†</td>
</tr>
<tr>
<td>No of comments</td>
<td>9214 (9693)</td>
<td>6308 (9920)</td>
<td>0.319</td>
</tr>
<tr>
<td>No of days since publication</td>
<td>144.9 (114.1)</td>
<td>127.6 (100.3)</td>
<td>0.564</td>
</tr>
<tr>
<td>No of likes</td>
<td>25 539 (19945)</td>
<td>18 815 (33716)</td>
<td>0.398</td>
</tr>
<tr>
<td>No of dislikes</td>
<td>3400.7 (5402.1)</td>
<td>5668.2 (8596.5)</td>
<td>0.365</td>
</tr>
<tr>
<td>Average duration (min)</td>
<td>8.38 (6.10)</td>
<td>8.45 (10.59)</td>
<td>0.981</td>
</tr>
</tbody>
</table>

Category

<table>
<thead>
<tr>
<th></th>
<th>Non-factual</th>
<th>Factual</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health professional</td>
<td>3 (23.1%)</td>
<td>22 (20.2%)</td>
<td></td>
</tr>
<tr>
<td>Entertainment</td>
<td>0 (0%)</td>
<td>7 (6.4%)</td>
<td></td>
</tr>
<tr>
<td>Network news</td>
<td>4 (30.7%)</td>
<td>38 (34.9%)</td>
<td></td>
</tr>
<tr>
<td>Internet news</td>
<td>2 (15.4%)</td>
<td>4 (3.7%)</td>
<td></td>
</tr>
<tr>
<td>Government</td>
<td>0 (0%)</td>
<td>9 (8.3%)</td>
<td></td>
</tr>
<tr>
<td>Newspaper</td>
<td>1 (7.7%)</td>
<td>3 (2.8%)</td>
<td></td>
</tr>
<tr>
<td>Consumer</td>
<td>1 (7.7%)</td>
<td>9 (8.3%)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>2 (15.4%)</td>
<td>17 (15.6%)</td>
<td></td>
</tr>
</tbody>
</table>

Data are represented as n (%) or mean (SD). P value for two-tailed t-test is shown where applicable. A significance threshold of p<0.01 was used.

Significant p-values highlighted in bold.

* n=101 as 8 videos hid their likes and dislikes count.
† Mann-Whitney U test used in lieu of t-test.

CVS, COVID-19 Vaccine Score; mDISCERN, modified DISCERN; mJAMA, modified Journal of the American Medical Association.

Table 3  Non-factual claims made in YouTube videos about COVID-19 vaccination*

<table>
<thead>
<tr>
<th>Reference standard</th>
<th>No of videos containing non-factual statements that contradict reference standard</th>
<th>Examples of non-factual statements</th>
</tr>
</thead>
</table>
| COVID-19 is a serious disease | 1                                                                               | ‘Immunity (natural and through healthy habits) and vitamin D are enough to survive COVID-19’
|                     |                                                                                 | ‘This virus is like pollen in the air…God is the real healer.’                                    |
| COVID-19 vaccine is effective | 3                                                                               | ‘Right now we are in the middle of a storm, and a vaccine is not going to help.’
|                     |                                                                                 | ‘I had COVID-19 so I don’t need the vaccine’                                                        |
| COVID-19 vaccine is generally safe | 6                                                                               | ‘I have autoimmune disease so I can’t have it’                                                     |
|                     |                                                                                 | ‘Vaccine affects fertility’                                                                          |
| There are very few contraindications to the COVID-19 vaccine | 3                                                                               | ‘Those who have any history of …. fever (may be due to any other cause), bleeding disorder, on blood thinners or immunocompromised…. or anyone suffering from serious health related issues…must not take vaccines’
|                     |                                                                                 | ‘Old people should avoid getting vaccinated’                                                          |

of government videos in this study compared with a study published earlier in the pandemic on COVID-19 misinformation. Despite producing the most reliable content, videos from government sources had over three times the ratio of dislikes to likes. In comparison, entertainment videos had a ratio of over 14 times the number of likes to dislikes. Similarly, the median ratio of likes to dislikes for videos containing non-factual information was almost six times higher than that of videos containing factual information. This suggests the presence of an active community of users who are sceptical of the COVID-19 vaccine and government sources. It is important to note that this may also represent a selection bias, as a previous study demonstrated that participants who relied on private sources for COVID-19 information, such as network news and social media, were less knowledgeable about COVID-19 and less likely to adhere to social distancing guidelines. Additionally, another study demonstrated that YouTube users were less likely to receive the COVID-19 vaccine compared with non-users.

Developing timely strategies to monitor and optimise the quality of information on YouTube during the COVID-19 pandemic is crucial for two reasons. First, it may be difficult for the YouTube algorithm to discern the accuracy and reliability of information conveyed. For example, the infamous COVID-19 film, Plandemic, which contained significant misinformation and conspiracy theories, was only removed from YouTube after public outcry, but not before it had already garnered millions of views at the time of removal, with the potential for serious consequences. Second, the YouTube algorithm allegedly makes recommendations for videos with a higher number of views and a greater proportion of likes. Moreover, the algorithm also allegedly adjusts video recommendations based on the user’s history of videos watched. Thus, it is plausible that individuals who encounter an initial video containing misinformation, may continue to receive recommendations for other videos containing similar misinformation. To address these concerns, in 2020, YouTube implemented an automated flagging system to demonetise and remove videos containing COVID-19 misinformation. However, this system’s reliance on automation has resulted in enforcement errors, such as mistakenly removing videos which sought to refute COVID-19 misinformation. Potential improvements could include increasing viewership of reputable government sources, as they have been shown to be the most reliable, by suggesting these videos to users. Furthermore, given that over 85% of those aged 13–17 years old use YouTube, algorithms suggesting youth-friendly government videos on COVID-19 vaccination to this audience may help increase vaccine compliance in the younger population, which has been reported to be much lower than among adults.

CONCLUSION

As the COVID-19 pandemic evolves, widespread vaccination is essential in reducing morbidity, mortality and returning to some semblance of normalcy. Providing high-quality health information from reputable sources that is engaging to a variety of viewers will be essential in addressing vaccine hesitancy and improving the public’s understanding and adherence to public health measures. Policy-makers should consider implementing mandates surrounding medical misinformation on social media platforms. Public health officials and clinicians should recognise the impact of YouTube and other similar social media sites as sources of health information for patients from all demographic backgrounds and leverage these platforms to disseminate high-quality content to combat misinformation.

Contributors Conception and design: HO-YL. Data collection: HO-YL, EP and OB-L. Analysis and interpretation: all authors. Manuscript writing: HO-YL, EP, OB-L and MGT. Critical revision of the article: all authors. Final approval of the article: all authors. Obtained funding: not applicable. Overall responsibility and guarantor: MGK.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was not required for this study as it does not involve human or animal subjects.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs Heidi Oi-Yee Li http://orcid.org/0000-0001-6495-088X
Mark G Kirchhof http://orcid.org/0000-0002-4015-8377

REFERENCES

23 Ohlheiser A. Facebook and YouTube are rushing to delete “Plandemic”, a conspiracy-laden video. MIT Technol Rev 2020 https://www.technologyreview.com/2020/05/07/1001469/facebook-youtube-plandemic-covid-misinformation/