

# The spread of vaccine misinformation on Twitter during 2021

*COVID-19, misinformation, social networks, Twitter, vaccines*

## Extended Abstract

The widespread adoption of vaccines is extremely important to reduce the impact of the COVID-19 pandemic, but a proportion of individuals worldwide are still hesitant about some or all vaccines. Previous literature suggests that vaccine hesitancy is linked to a number of factors that include individuals' political, cultural, and social background, as well as the information environment [7]. Lately, concerns have been raised around false vaccine-related claims spreading online that may promote vaccine hesitancy or refusal [3].

A massive amount of misinformation — the so-called “infodemic” [10] — followed the outbreak of the COVID-19 pandemic [4], including a deluge of health-related false claims and conspiracy theories that hindered the effectiveness of vaccination campaigns [5]. Motivated by recent studies, which reveal a positive association between exposure to misinformation and vaccine hesitancy at the individual and population level [6, 9], our work investigates the spread of COVID-19 vaccine misinformation on Twitter during 2021, when vaccination programs were launched in most countries around the world.

Our contribution is manifold. We identified millions of tweets sharing links to low-credibility and mainstream news websites, finding an increasing trend in the reshares of unreliable news during the year, and an opposite, decreasing trend for reliable information. However, low-credibility information about vaccines was generally less prevalent than mainstream news throughout the period of analysis. We also observed a non-negligible proportion of tweets linking to YouTube videos that had been inaccessible, most likely because they were violating platforms' policy against misinformation [1] (see Figure 1). Alarmingly, we noticed three low-credibility websites with volumes of reshares comparable to reliable sources. In particular, the most relevant source of vaccine misinformation, Children's Health Defense, received more than twice the number of tweets linking to the Centers for Disease Control and Prevention (see Figure 2). We observed the presence of many verified accounts among users who earned the most retweets when sharing low-credibility news about vaccines. In particular, the most popular spreader of misinformation was Robert Kennedy Jr., the founder of Children's Health Defense. While we found significantly fewer verified accounts sharing low-credibility vaccine content compared to those sharing vaccine content in general, we also observed that verified accounts tended to receive more retweets when posting low-credibility content than general vaccine content (see Figure 3). Leveraging a bot-detection algorithm, we found that tweets sharing links to low-credibility sources had significantly higher bot-activity levels than vaccine tweets overall. However, we did not find a trend of increased levels of automated sharing over time.

Our findings help map the landscape of online vaccine misinformation and design intervention strategies to curb its spread. In a context of widespread uncertainty such as the COVID-19 pandemic, trust is critical for overcoming vaccine hesitancy, and recent research shows how online misinformation fueled vaccine hesitancy and refusal sentiment [9]. Our results reveal the presence of a small number of main producers and repeat spreaders of low-credibility content. Given that these superspreaders played key roles in disseminating vaccine misinformation, a

straightforward strategy could be to deplatform them, as it has been done by major platforms in notable cases such as Alex Jones [2] and Donald Trump. While social media platforms have legal rights to regulate online conversations, the decisions to deplatform public figures should be made with caution. In fact, past intervention has sparked a vivid debate around free speech and caused many users to migrate to alternative platforms. An alternative path of action could be to reduce the financial incentives of those who profit from the spread of misinformation. Other effective approaches to reduce its spread include lowering the visibility of certain content or not showing that content to users, as well as adding warning labels to content that is potentially harmful or inaccurate. Our work adds to extant literature suggesting that platforms should partner with policymakers and researchers in evaluating the impacts of such different interventions [8].

## References

- [1] D. Axelrod, B. Harper, and J. Paolillo. Youtube covid-19 vaccine misinformation on twitter: Platform interactions and moderation blind spots. In *Proceedings of the Conference on Truth and Trust Online*, 2022.
- [2] N. Bilton. The downfall of alex jones shows how the internet can be saved, 2021.
- [3] T. Burki. Vaccine misinformation and social media. *The Lancet Digital Health*, 1(6):e258–e259, 2019.
- [4] R. Gallotti, F. Valle, N. Castaldo, P. Sacco, and M. De Domenico. Assessing the risks of ‘infodemics’ in response to covid-19 epidemics. *Nature human behaviour*, 4(12):1285–1293, 2020.
- [5] P. Hotez, C. Batista, O. Ergonul, J. P. Figueroa, S. Gilbert, M. Gursel, M. Hassanain, G. Kang, J. H. Kim, B. Lall, et al. Correcting covid-19 vaccine misinformation: lancet commission on covid-19 vaccines and therapeutics task force members. *EClinicalMedicine*, 33, 2021.
- [6] S. Loomba, A. de Figueiredo, S. J. Piatek, K. de Graaf, and H. J. Larson. Measuring the impact of covid-19 vaccine misinformation on vaccination intent in the uk and usa. *Nature human behaviour*, 5(3):337–348, 2021.
- [7] N. E. MacDonald et al. Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, 33(34):4161–4164, 2015.
- [8] I. V. Pasquetto, B. Swire-Thompson, M. A. Amazeen, F. Benevenuto, N. M. Brashier, R. M. Bond, L. C. Bozarth, C. Budak, U. K. Ecker, L. K. Fazio, et al. Tackling misinformation: What researchers could do with social media data. *The Harvard Kennedy School Misinformation Review*, 2020.
- [9] F. Pierri, B. L. Perry, M. R. DeVerna, K.-C. Yang, A. Flammini, F. Menczer, and J. Bryden. Online misinformation is linked to early covid-19 vaccination hesitancy and refusal. *Scientific Reports*, 12(1):1–7, 2022.
- [10] J. Zarocostas. How to fight an infodemic. *The Lancet*, 395(10225):676, 2020.

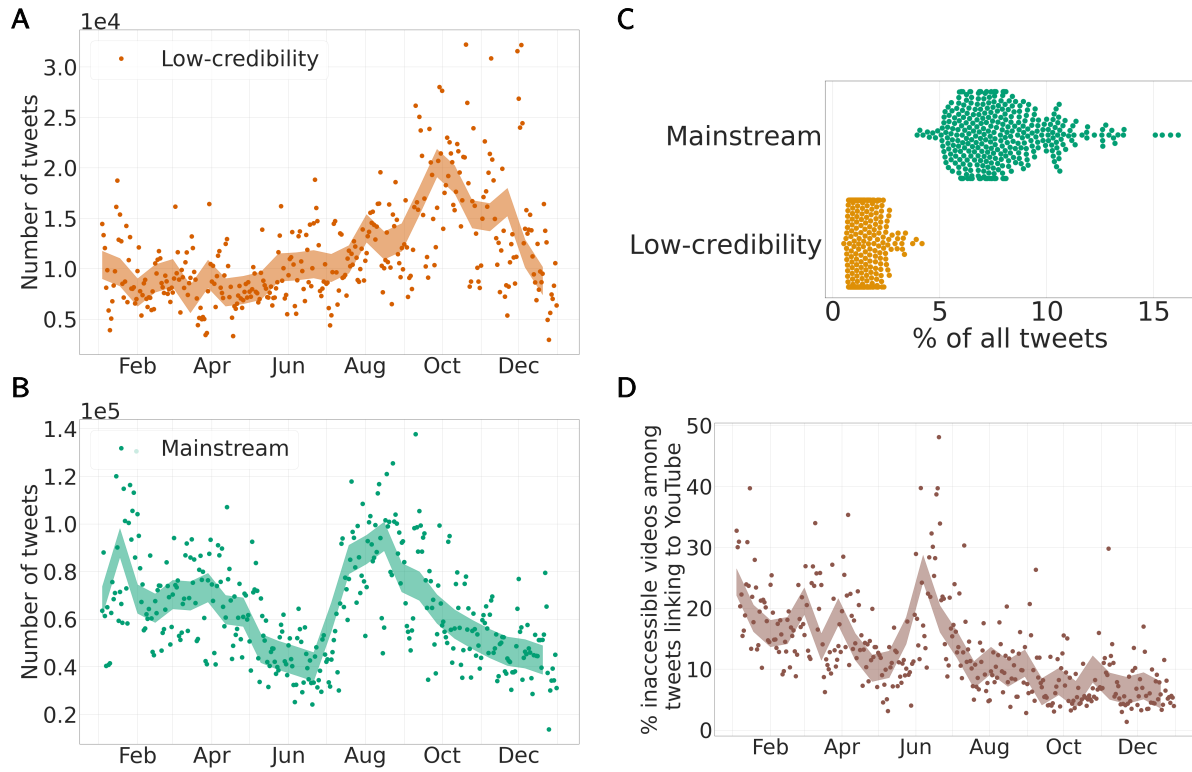


Figure 1: **(A)** Daily number of vaccine tweets sharing links to news articles from low-credibility sources. There is a significant increasing trend ( $P < .001$ ). **(B)** Daily number of vaccine tweets sharing links to news articles from mainstream sources. There is a significant decreasing trend ( $P < .001$ ). **(C)** Daily percentage of vaccine tweets and retweets that share links to low-credibility news sources (median: 1.31%) and mainstream news sources (median: 7.53%). The distributions are statistically different according to a two-sided Mann-Whitney test ( $P < .001$ ). **(D)** Daily percentage of tweets sharing links to inaccessible YouTube videos, out of all tweets sharing links to YouTube. There is a significant decreasing trend ( $P < .001$ ). We employ non-parametric Mann-Kendall tests for trends. Colored bands correspond to a 14-day rolling average with 95% C.I.

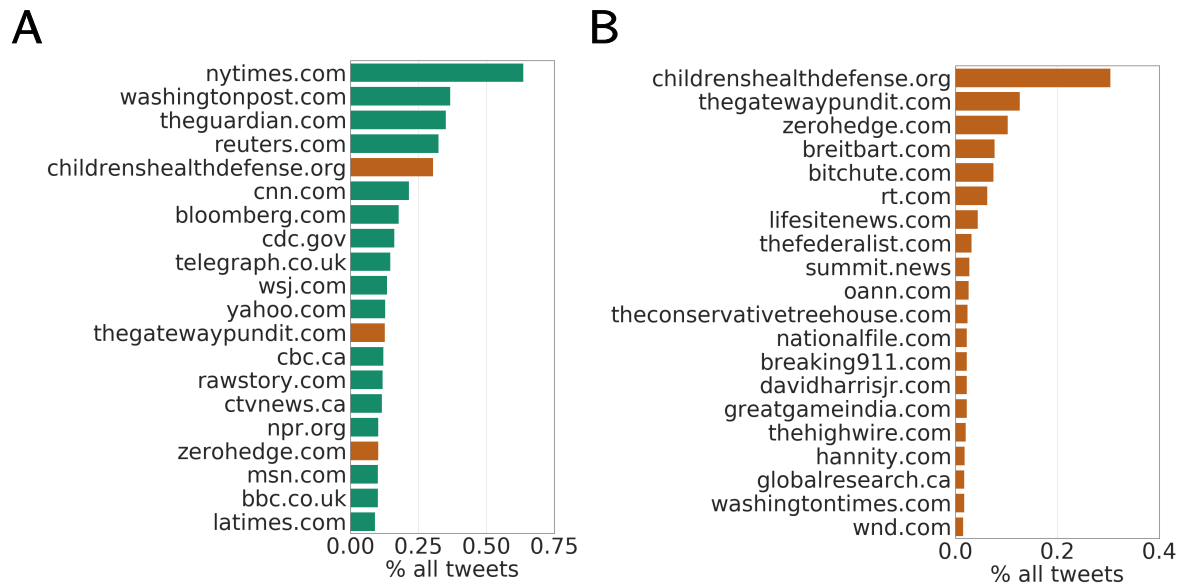


Figure 2: **(A)** The top 20 news sources ranked by percentage of vaccine tweets. **(B)** The top 20 low-credibility news sources ranked by percentage of vaccine tweets.

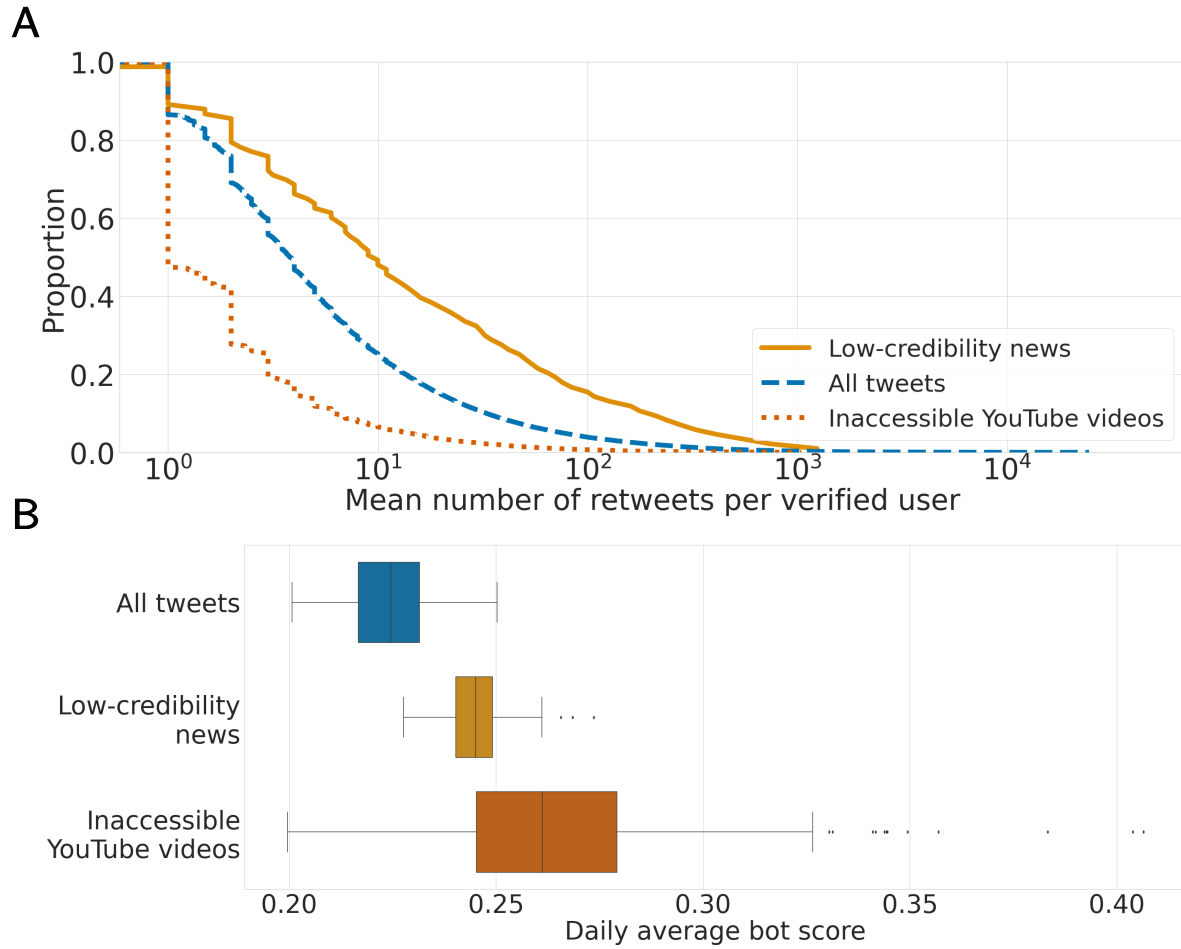


Figure 3: **(A)** Distributions of the mean numbers of retweets earned by verified accounts when sharing vaccine content (median 3.82), low-credibility news (median 9.43), and links to inaccessible YouTube videos (median 1). Since the distributions are broad, the box plots (inset) have many outliers. Therefore we also display the complementary cumulative distributions (main plot). All distributions are significantly different from each others according to two-sided Mann-Whitney tests ( $P < 0.001$ ). **(B)** Comparison between the daily average bot score of tweets sharing different categories of vaccine content. The median daily average bot scores of accounts sharing vaccine content, low-credibility news, and inaccessible YouTube videos are 0.22, 0.25 and 0.26, respectively. All distributions are significantly different from each other according to two-sided Mann-Whitney tests ( $P < 0.001$ ).