Understanding the Temporal Effects on Tweetcussion of COVID-19 Vaccine

Chei Sian Lee
Nanyang Technological University
leecs@ntu.edu.sg

Dion Hoe-Lian Goh
Nanyang Technological University
ashlgoh@ntu.edu.sg

Han Wei Tan
Nanyang Technological University
htan092@ntu.edu.sg

Han Zheng
Nanyang Technological University
han019@ntu.edu.sg

Yin Leng Theng
Nanyang Technological University
tyltheng@ntu.edu.sg

ABSTRACT

In the fight against COVID-19, the Pfizer and BioNTech vaccine announcement marked a significant turning point. Analysing the topics discussed surrounding the announcement is critical to shed light on how people respond to the vaccination against COVID-19. Specifically, since the COVID-19 vaccine was developed at unprecedented speed, different segments of the public with a different understanding of the issues may react and respond differently. We analysed Twitter tweets to uncover the issues surrounding people's discussion of the vaccination against COVID-19. Through the use of Latent Dirichlet Allocation (LDA), nine topics were identified pertaining to vaccine-related tweets. We analysed the temporal differences in the nine topics, prior and after the official vaccine announcement.

KEYWORDS

COVID-19, Topic Modelling, Twitter, Vaccine.

INTRODUCTION

In the fight against COVID-19, 9 November 2020 perhaps marked a significant turning point in the battle as Pfizer and BioNTech (Pfizer-BioNTech) announced that their vaccine candidate was more than 90% effective in preventing the disease (Pfizer, 2020). The development of the Pfizer-BioNTech and other COVID-19 vaccines occurred at an unprecedented speed (World Economic Forum, 2020) and has led to scepticism and other concerns among some segments of the public. Indeed, conversations about the vaccine proliferated on social media platforms after the announcement, presenting an opportunity to understand online reactions to this announcement.

Given that vaccination is a critical tool in the fight against COVID-19, the present study aims to investigate the public reactions towards COVID-19 vaccines and the temporal differences in topics, prior and after the vaccination announcement. To identify the underlying topics, we employ topic modelling in our analysis. Twitter is used in our research because it is currently being actively used for sharing COVID-19 and vaccine-related content (Thelwel et al., 2021). Research has shown that content shared online not only reflects the public's understanding of health issues but also has the potential to shape perceptions and attitudes (e.g., Massey et al., 2016; Deiner et al., 2019). So, understanding topics discussed is critical in rolling out vaccination programs since prior work on vaccines has also found that consumption of online health information sources can be risky due to the amount of misinformation and unverified content available (e.g., Tomeny et al., 2017; Yuan et al., 2019). Here, we propose that tweets are useful and timely for policymakers to frame COVID-19 vaccination messages in their communication to the public. We propose to answer this research question: *What are the topics discussed surrounding the Twitter conversation on COVID-19 vaccine following the announcement?*

RELATED WORK

Prior works related to the sharing and seeking health-related content have shown that Twitter can be used for surveillance of public health topics, categorisation of health-related tweets, Twitter-based public health interventions, (Gunaratne, 2019; Osop et al., 2020; Sinnenberg et al., 2017).Recent studies have identified major topics on vaccine refusal (Bonnevie et al., 2021) and vaccine hesitation (Thelwell et al., 2021; Griffith et al., 2021) to understand the source of concern and, importantly, combat against vaccine misinformation. Our study complements prior works by incorporating the temporal component since time matters when examining tweets in a public health context.

METHODOLOGY

This study was part of a larger project that analyses tweets pertaining to COVID-19 vaccine from November 01, 2020, to November 16, 2020.English tweets were extracted from a public dataset on GitHub (Chen et al., 2020) that contains an ongoing of tweets IDs related to COVID-19, since January 28, 2020, through Twitter's Search API using a list of keywords (e.g., "Covid-19", "pandemic", etc.). The tweets were filtered for vaccine-related keywords

(e.g., "vaccine", "vaccination", etc.), and we removed emojis, "RT" text, mentions, URL links, special characters, and spaces. A list of standard stop words (e.g., "the", "does"), frequently appeared words (e.g., "covid19", "virus"), and the vaccine-related keywords were also removed. Porter stemmer was used to abbreviate the words into root words and eliminated infrequent sparse terms. The final dataset comprised 44,699 tweets with 365,288 words and 1,566 unique words. Latent Dirichlet Allocation (LDA) topic modelling was used to identify COVID-19 vaccine-related topics discussed on Twitter due to its potential and effectiveness to detect new trends on social media as an unsupervised classification technique (Ostrowski, 2015). In this process, we used two data-driven metrics to evaluate the model fit (Cao et al., 2009; Deveaud et al., 2014), and after evaluation we ran the LDA model with nine topics for our dataset.

RESULTS

Nine topics were identified from the LDA topic modelling approach comprising T1: **Administration of local vaccine program**, T2: **Complementary measures to vaccine**, T3: **Social aspects of vaccine**, T4: **Efficacy of vaccine**, T5: **Global distribution and access of vaccine**, T6: **Myths about vaccine**, T7: **Legal and economic aspect of vaccine**, T8: **Pace of vaccine development** and T9: **Political aspects of vaccine**. The evaluated trend of the topic weightage score (see Figure 1) was obtained by dividing the number of tweets per day in each topic by the total number of daily tweets in all topics. Within expectation, an increase in discussions related to efficacy and social aspects of the vaccine was detected around the vaccine on the day of the announcement, on November 9, 2020. We also observed an increase in interest, likely with concerns, over the global distribution of the vaccine (T5: Global distribution and access of vaccine),possibly leading to the debate on the legal and economic aspects of the vaccination (T7: Legal and economic aspects of vaccine), with the USA presidential election (T9: Political aspects of vaccine) being a potentially linked political topic over the situation, while reduced conversations over vaccine development and administration (T1: Administration of local vaccine Program and T8: Pace of vaccine development) was found.
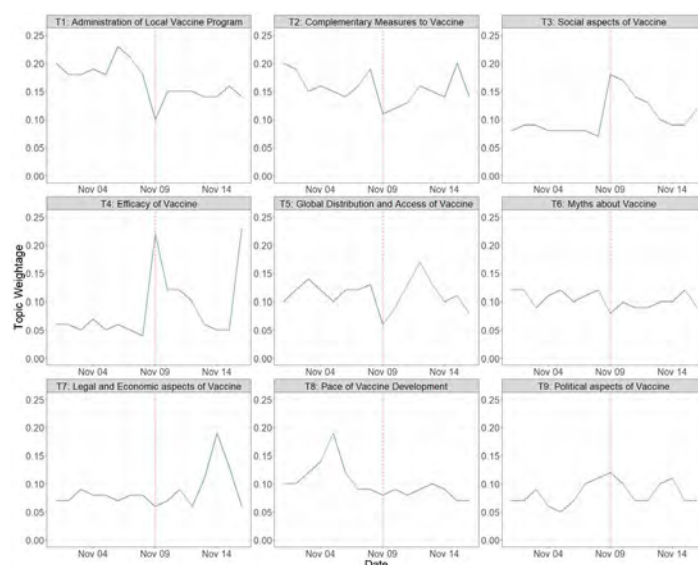

Figure 1. Change of topic over Time

DISCUSSION AND CONCLUSION

Our findings reveal the shared thinking and understanding of the public and their opinions surrounding the vaccine announcement. Our preliminary findings indicate some topics continued to proliferate on Twitter while other topics saw a decline in discussion. The findings help to better understand people's attitudes and opinions about the COVID-19 vaccine. Importantly, we found that despite efforts put in to debunk misinformation surrounding vaccination against COVID-19 (Russell, 2021), it appeared that as COVID-19 vaccine programs continued to roll out globally, the myths around the vaccines continued to proliferate on Twitter (Figure 1 Topic 6).With new developments and controversies surrounding specific vaccines and the vaccination process in various countries, we plan to collect more data and incorporate newer tweets to ascertain if and how topics have changed. We also plan to conduct sentiment analysis to shed light on the sentiments behind the topics. Comparing vaccine discussions across different social media platforms and languages would also be worthwhile. By focusing on the conversations around the announcement, this study aims to contribute to better understand reactions to vaccination against COVID-19 which may ultimately help policymakers in ensuring equity and inclusivity in implementing vaccination programs that work for everyone. Also, understanding how social media conversations shift will help in the creation of better official responses to the public.

REFERENCES

Bonnevie, E., Gallegos-Jeffrey, A., Goldbarg, J., Byrd, B., & Smyser, J. (2021). Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. Journal of Communication in Healthcare, 14(1), 12–19. https://doi.org/10.1080/17538068.2020.1858222

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. Advances in Machine Learning and Computational Intelligence, 72(7), 1775–1781. https://doi.org/10.1016/j.neucom.2008.06.011

Chen, E., Lerman, K., & Ferrara, E. (2020). Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. JMIR Public Health and Surveillance, 6(2), e19273. https://doi.org/10.2196/19273

Deiner, M. S., Fathy, C., Kim, J., Niemeyer, K., Ramirez, D., Ackley, S. F., Liu, F., Lietman, T. M., & Porco, T. C. (2019). Facebook and Twitter vaccine sentiment in response to measles outbreaks. Health Informatics Journal, 25(3), 1116–1132. https://doi.org/10.1177/1460458217740723

Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. Document numérique, 17(1), 61–84. Cairn.info. https://doi.org/10.3166/DN.17.1.61-84

Griffith, J., Marani, H., & Monkman, H. (2021). COVID-19 Vaccine Hesitancy in Canada: Content Analysis of Tweets Using the Theoretical Domains Framework. Journal of Medical Internet Research, 23(4), e26874. https://doi.org/10.2196/26874

Gunaratne, K., Coomes, E. A., & Haghbayan, H. (2019). Temporal trends in anti-vaccine discourse on Twitter. Vaccine, 37(35), 4867–4871. https://doi.org/10.1016/j.vaccine.2019.06.086

Massey, P. M., Leader, A., Yom-Tov, E., Budenz, A., Fisher, K., & Klassen, A. C. (2016). Applying Multiple Data Collection Tools to Quantify Human Papillomavirus Vaccine Communication on Twitter. J Med Internet Res, 18(12), e318. https://doi.org/10.2196/jmir.6670

Osop, H., Hasan, R., Lee, C. S., Neo, C. Y., Foo, C. K., & Saurabh, A. (2020). Diabetweets: Analysis of Tweets for Health-Related Information. In C. Stephanidis, M. Antona, & S. Ntoa (Eds.), HCI International 2020 – Late Breaking Posters (pp. 500–508). Springer International Publishing. https://doi.org/10.1007/978-3-030-60703-6_65

Ostrowski, D. A. (2015). Using latent dirichlet allocation for topic modelling in twitter. Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015), 493–497. https://doi.org/10.1109/ICOSC.2015.7050858

Pfizer. (2020, November 9). Pfizer and BioNTech Announce Vaccine Candidate Against COVID-19 Achieved Success in First Interim Analysis from Phase 3 Study | pfpfizeruscom. https://www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-announce-vaccine-candidate-against

Russell, A. (2021). The Fight Against Vaccine Misinformation. The New Yorker. https://www.newyorker.com/news/letter-from-the-uk/the-fight-against-vaccine-misinformation

Sinnenberg, L., Buttenheim, A. M., Padrez, K., Mancheno, C., Ungar, L., & Merchant, R. M. (2017). Twitter as a Tool for Health Research: A Systematic Review. American Journal of Public Health, 107(1), e1–e8. https://doi.org/10.2105/AJPH.2016.303512

Thelwall, M., Kousha, K., & Thelwall, S. (2021). Covid-19 vaccine hesitancy on English-language Twitter. El Profesional de La Información, e300212. https://doi.org/10.3145/epi.2021.mar.12

Tomeny, T. S., Vargo, C. J., & El-Toukhy, S. (2017). Geographic and demographic correlates of autism-related anti-vaccine beliefs on Twitter, 2009-15. Social Science & Medicine, 191, 168–175. https://doi.org/10.1016/j.socscimed.2017.08.041

World Economic Forum. (2020). 5 charts that tell the story of vaccines today.https://www.weforum.org/agenda/2020/06/vaccine-development-barriers-coronavirus/.

Yuan, X., Schuchard, R. J., & Crooks, A. T. (2019). Examining Emergent Communities and Social Bots Within the Polarized Online Vaccination Debate in Twitter. Social Media + Society, 5(3), 2056305119865465. https://doi.org/10.1177/2056305119865465