SAIS 2023 Proceedings   Southern (SAIS)

7-1-2023

The Celebrity Factor: Modeling the Impact of Influencer Statements on the Public Perception toward Covid-19 Vaccines

Abhishek Shah

Shweta Shah

Follow this and additional works at: https://aisel.aisnet.org/sais2023

The Celebrity Factor: Modeling the Impact of Influencer Statements on the Public Perception of COVID-19 Vaccines

Abhishek Shah
Green Level High School
shahabhishek1729@gmail.com

Shweta Shah
Green Level High School
shwetashah857@gmail.com

ABSTRACT

Online social networks rapidly propagate information and opinions expressed on such platforms can influence others' decisions. These platforms were widely used during the COVID-19 pandemic, and many celebrities shared their opinions about the vaccines developed against the virus. Most encouraged vaccination, but many were skeptical about the vaccines' safety and efficacy. This study modeled the impact of eight influencers' statements on the public sentiment toward COVID-19 vaccines. Sentiment is measured for 2 million vaccine-related tweets, time-series models are used to model the influencers' impacts, and Bayesian analysis of each model's predictions was used to estimate impacts. The results found that influencers who encourage vaccination tend to increase the number of "pro-vaccination" tweets, while influencers sharing "anti-vaccination" messages are prone to unstable effects (sometimes leading to *fewer* anti-vaccination tweets). This study provides insights into the complex issue of vaccine hesitancy and may inform public health strategy.

Keywords

Vaccine hesitancy, time-series forecasting, vaccines, COVID-19, machine learning, celebrity influence, BSTS, sentiment analysis, VADER lexicon, Twitter

INTRODUCTION

The SARS-CoV-2 virus was first detected during December 2019 in Wuhan, China (Zhu et al. 2020). The virus has since spread to most of the world and was declared a global pandemic by the World Health Organization on March 11, 2020 (Ryan 2021). To end the pandemic, governments around the world funded the development of vaccines – the U.S. government alone spent roughly $13 billion on vaccines (Bloom et al. 2021). mRNA-based COVID-19 vaccines, such as the COMIRNATY vaccine (developed by Pfizer Inc., BioNTech SE, and Shanghai Fosun Pharmaceutical Co., Ltd.), and the Moderna vaccine, function by injecting genetic material from the virus (or a lab-made modification of the virus) to stimulate the immune system to produce antibodies that combat that strain — the immune system is then able to regenerate these antibodies when infected with the strain again (Hendaus and Jomha 2021). However, if governments are to achieve the goal of herd immunity (when a large part of a population has developed immunity to a disease), they need large proportions, an estimated 67%, of their populations to take the developed vaccines (Randolph and Barreiro 2020). Unfortunately, many individuals are hesitant to take the vaccine, citing beliefs (often based on misinformation) about the safety and efficacy of the vaccine to support their views (Afifi et al. 2021). Vaccine hesitancy leads to a delay in countries achieving herd immunity despite enough availability of vaccines — potentially leading to further spread and mutation of the virus, and eventually even greater loss of life (MacDonald et al. 2015). Indeed, vaccine hesitancy was recently included in the list of the ten major threats to global health by the World Health Organization (Godlee 2019). One approach employed to promote vaccine adoption is to use prominent figures in popular culture to spread accurate information and encouragement (Sullivan 2021).

Celebrities have long held influence among the public, especially among younger age groups, i.e., adolescents and young adults, who, incidentally, happen to be one of the most vaccine-hesitant groups (Adams et al. 2021). Celebrity advertising, for instance, has been a popular method of marketing used by companies for decades to reach these groups (Yannopoulos 2012). With the rise of social media platforms such as Twitter, Instagram, Facebook, Reddit, and others, celebrities have been able to spread their views further than ever before, due to the rapid diffusion rate of messages posted on their relatively massive networks (Selkie 2022). Recently, many popular celebrities and social media influencers have used their platform to push out statements supporting the vaccine, both independently and by partnering with public health authorities; however, a considerable number of celebrities have also pushed statements against the vaccine (Sloss 2021), potentially leading to increased vaccine hesitancy amongst their followers. This study explores how celebrities and social media influencers have affected the public sentiment, as expressed on Twitter, regarding COVID-19 vaccines.

RELATED WORK

Research in the areas of celebrity influence and endorsements has developed significantly in recent times. A review in the late 20th century suggested that celebrity endorsements were less effective than they had been prior to the 1980s (Kaikati 1987). In a similar context, a study conducted in 2021 found that celebrity messages significantly increased users' urges to make impulsive purchases (Zafar et al. 2021). Both analyses, however, look at celebrity messaging in a marketing context, in which the subjects being studied behave in a different manner than if they were making health-related decisions (as those explored here). A 2013 analysis concluded that celebrity messaging should be further examined as a tool to help public health authorities discredit misinformation and promote positive health practices (Hoffman and Tan 2013). In a more specific context, a study examining AIDS in the African American community found that while in many cases, celebrities were indeed effective at promoting safe practices to avoid contracting HIV/AIDS, many non-celebrities were capable of being just as effective (Robinson 2003). Such studies have begun to show interesting results, but further evaluation is needed in this context to develop a better understanding of influencers' impacts on the public's medical decisions. This study will look at a more recent setting in which to study the medical context, namely, the COVID-19 pandemic. The prevalence of this pandemic during the window of the study will help provide clearer conclusions, and both kinds of influencers will be studied: those who promote vaccinations *and* those who share messages against vaccines.

DATA

To quantify the public opinion towards COVID-19 vaccines, messages regarding these vaccines posted on the Twitter platform were collected. 1,948,191 English tweets containing keywords pertaining to vaccines (such as "vaccine" or "vax") from November 1, 2020, through March 11, 2022, were collected. Data was limited to 4,133 tweets per day (some days, fewer numbers were collected, often because of site outages) to ensure that the data collected was distributed as evenly as possible throughout the window. Once this data was collected, each tweet was cleaned by removing all hashtags, mentions, URLs, media (such as images or videos), special formatting characters included with tweets (e.g., "\n"), non-ASCII characters, emojis and punctuation. To classify a tweet as pro-vaccination or anti-vaccination, a sentiment score was assigned to each tweet based on its processed text; tweets sharing messages promoting vaccines tend to express positive sentiments, whereas tweets sharing messages against vaccinations tend to express negative sentiments. To determine the sentiment scores, the Python-based library VADER (Valence Aware Dictionary and sEntiment Reasoner) was used (Hutto and Gilbert 2014). VADER is optimized to deal with abbreviated phrases, emoticons, slang, and other elements commonly found in tweets. Once a score, denoted C, between -1 and 1 was obtained, tweets were then labeled as follows:

1. "Positive" tweets, where $C > 0.2$, which are likely to express pro-vaccination sentiments
2. "Negative" tweets, where $C < -0.2$, are likely to share messages against vaccinations or mandates

(Any tweets not falling into either category were discarded.) Two new datasets were built based on the collected tweets and the attached sentiment labels: one contained only tweets labeled "positive" and one contained only the "negative". Both datasets were formatted as time series of the number of tweets expressing each sentiment shared per day and a number of useful features were added to each dataset. By the end of this process, both the positive and the negative datasets were 495×38 matrices, storing the date and tweet count with the 36 other features for each of the 495 days. The data collection phase also involved identifying influencers that have made statements in the past regarding COVID-19 vaccines that may have affected the public opinion. An influencer is, for the purposes of this study, defined as a social media account on either the Twitter or Instagram platforms who had at least 100,000 followers on their respective platform, and had been granted "verified" status on their platform as of February 28, 2022. Furthermore, an influencer is said to be positive if the post of interest expresses a pro-vaccination sentiment (regardless of whether other posts prior or after the post of interest express conflicting sentiments) and vice versa. This study tracked a total of eight influential figures — four "positive" accounts: @SteveMartinToGo (tweet) @Bill Gates (tweet), @BigBird (tweet), and @johnlegend (Instagram post), and four "negative" accounts: @evangelinelillyofficial (Instagram post), @RobSchneider (tweet), @chethanx (video shared on Instagram on July 09, 2021) and @NICKIMINAJ (tweet). Although influencers were chosen from both platforms, their posts were discussed extensively on Twitter and their impact was studied based solely on the collected tweets.

MODELS

To model the impact of a post by an influencer on the recorded sentiment scores, the following process was adopted:

1. Using only the data prior to the selected influencer's post, a collection of 5 models (described below) were trained on the dataset matching the post's sentiment (i.e., models for a pro-vaccine post's impact are trained on the positive dataset, and vice versa)
2. Based on these models, 5 forecasts were generated for the next 20 days after the post (producing a separate forecast for each model). These forecasts are labelled counterfactuals i.e., estimates of what "would have happened" in the

absence of the post. Counterfactuals, in this case, are time-series model with 20 time points (days) of data — they estimate the likely trajectory of public sentiment in the days following the post had it not been shared (for positive influencers, they represent a forecast of the number of positive tweets over the next 20 days, and vice versa).

3. The counterfactual data is then passed along with the observed trajectory of the time series ("what actually happened" over the next 20 days) into the CausalImpact R library. The library runs a Bayesian analysis on the two time-series and returns distributions of the average and sum of the differences between the counterfactual prediction (for the number of tweets sharing the influencer's sentiment) and the observed values (for the number of tweets sharing the influencer's sentiment) over the next 20 days. It also computes the probability that the difference could have been observed due to chance
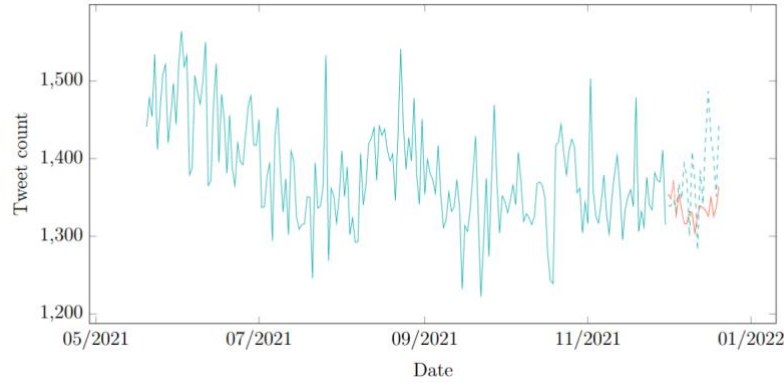

Figure 1. Plot showing the process to compute the impact for an example post on 12/01/2021 — all data points prior to this date (of which only the most recent 200 are shown in solid cyan for visibility purposes) would be used to train the models described below, which generate the counterfactual shown (in orange). It is then compared to the observed data (dashed cyan) to generate the impact estimates.

To generate the counterfactual estimates, five different models were built, each of which was evaluated based on its root mean squared error (RMSE). 80% of the data before the influencer's post was used to train the models, and the last 20% was used to evaluate their performance (by computing the RMSE); after RMSE values were obtained for a model, it was retrained on the full data before the post and the forecasts for the period after the post were obtained. The RMSE was calculated in the following manner:

$$J = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where $\hat{y}$ denotes a 20-element vector of the model's counterfactual estimates, and y represents a 20-element vector of the true values of the tweet counts over the next 20 days. The following models were used to generate the counterfactuals (represented by y in the equations below):

1. A linear regression model
2. A Prophet model with regressors (Taylor and Latham 2018)

$$y = g(t) + s(t) + h(t) + \varepsilon(t)$$

where $g(t)$ denotes non-periodic trends visible throughout the data, $s(t)$ represents seasonal patterns, $h(t)$ is used to account for the effects of holidays, and ε denotes the error term, storing white noise unaccounted for by the model. If the data seems to exhibit "non-linear, saturating growth" (which levels out at some carrying capacity C), $g(t)$ can be most easily modeled by

$$g(t) = \frac{C}{1 + e^{-k(t-m)}}$$

where C denotes the carrying capacity, the growth rate is represented by k, and m denotes an offset. For non-logistic trends, below is a piece-wise linear trend model with S change-points (s_j, where $j = 1, \ldots, S$) at which the growth factor changes

$$g(t) = (k + \boldsymbol{a}(t)^{\top}\boldsymbol{\delta})t + (m + \boldsymbol{a}(t)^{\top}\boldsymbol{\gamma}$$

where $\boldsymbol{a}(t) \in {0,1}^S$, such that $aj(t) = 1$ for all $t \geq 1$ and $a_j(t) = 0$ otherwise. The change in the growth rate at sj is denoted by δ_j and included in the vector δ; to ensure the function is continuous, γ is defined where $\gamma_j = -s_j\delta_j$. As before, the growth rate is represented by k and the offset parameter is denoted m.

3. An Extreme Gradient Boosting (XGBoost) model (Chen and Guestrin 2016)

Unlike the prior models described, XGBoost builds an additive ensemble of k trees to compute the following prediction:

$$y = \phi(\boldsymbol{x}) = \sum_{k=1}^{K} f_k(\boldsymbol{x}), \, f_k \in \mathcal{F}$$

Here, CART (the space of trees), is represented by $\mathcal{F} = f(\boldsymbol{x}) = w_{q(x)}$, where $q: \mathbb{R}^m \to T$ represents the tree structure, mapping a group of features to their leaf index, T represents the tree's leaf count and f_k represents a tree with structure q and weights $w \in \mathbb{R}^T$. To update the model's functions based on previous examples, XGBoost works to minimize the following objective:

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \left[\gamma T + \frac{1}{2}\lambda||w||^2\right],$$

and $\gamma T + \frac{1}{2}\lambda||w||^2$ is commonly abbreviated $\Omega(f)$. Here, $l(\hat{y}_i, y_i) = \frac{1}{2}(\hat{y}_i - y_i)^2$, but any differentiable l is acceptable. λ is the regularization parameter and γ controls the pruning of the model, with both parameters working together to penalize extra complexity (and avoid overfitting).

4. A Prophet model with gradient-boosted errors

This model uses a two-pronged approach towards modeling the time series. First, a Prophet model as described above is used to approximately represent the time series (ignoring all extra regressors). Then, an XGBoost model is applied to the residual terms of the Prophet model. In this way, approaches (2) and (3) above are combined to generate a fourth model.

5. A weighted ensemble combining all four previous models based on their ranked RMSE:

$$y = \theta_1 h_1 + \theta_2 h_2 + \theta_3 h_3 + \theta_4 h_4$$

where h_i represents the predictions of the ith model, and $\theta_i \in {0.1; 0.2; 0.3; 0.4}$ represents the weight the ith model's predictions receive. The model with the lowest RMSE received a weight of 0.4, the next lowest received a weight of 0.3 and so on. These values were chosen since they sum to 1.0 and provided an easy way to weight the accuracy of the models — in future work, alternative weights may be explored.

To perform the intervention analysis, two distinct time periods were defined: if d represents the date of an influence post regarding the vaccine, the 1) pre-intervention period is defined to run from November 1, 2020, to $d - 1$, and the 2) post-intervention period is defined to run from d to $d + 20$. Then, using the CausalImpact R package (Brodersen et al. 2015), the information from these counterfactuals were combined with a Bayesian structural time series model. A Bayesian structural time series model takes the following form:

$$y_t = Z_t^T \alpha_t + \varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, H_t)$$
$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \qquad \eta_t \sim \mathcal{N}(0, Q_t)$$

where y_t represents the observed data, and α_t represents the unobserved latent state. The priors for this model were the distributions of the model parameters which were chosen to be Gaussian random walks with standard deviation 0.01 (as per the CausalImpact recommendation for stable time series). This model was used to estimate 3 values over the next 20 days: 1) the average impact, which was computed as the average daily difference between the observed number of tweets sharing the influencer's sentiment and models' estimates throughout post-intervention period; 2) the total impact, which was computed as the sum of the differences between the observed data and counterfactual estimate over the post-intervention period; and 3) the *p-value*, which represents the Bayesian one-sided tail area probability of obtaining the computed differences by chance.

RESULTS

Before interpreting the model results, it is important to make two distinctions: 1) An influencer is said to have had a *positive impact* when the estimated average and total impacts for an influencer are greater than 0. If the impact estimates are below 0, the influencer is said to have had a negative impact. This means that a "negative" influencer (based on the sentiment their vaccine-related post expresses) may have a positive impact (implying that the number of negative tweets increased overall in the post-intervention period), and vice versa. 2) An influencer is said to have had a *significant impact* when the p-value (Bayesian one-sided tail area probability) of their impact is below 0.050 (i.e., $\alpha = 0.05$).

Influencer	RMSE	Average Impact	Total Impact	p-value
@SteveMartinToGo	0.785	620.02	12400.32	0.022
@BillGates	0.690	238.59	5671.80	0.001
@BigBird	0.665	12.10	241.97	0.013
@johnlegend	0.612	14.204	284.97	0.003

Table 1. Weighted ensemble estimates for influencers sharing pro-vaccination message

In both the instances of @SteveMartinToGo and @BillGates, all models agree on the influencers having significant, positive impacts, albeit to varying degrees. These similar results make intuitive sense, given that both posts were shared within a short time of each other. For @BigBird, four out of five models seemed to agree on the tweet having a positive impact, with another four out of five models also electing that the tweet had a significant impact. For @johnlegend, only three out of five models believed the post had a significant impact, although every model showed that the impact was positive (the XGBoost counterfactual computed almost an imperceptible impact).

Influencer	RMSE	Average Impact	Total Impact	p-value
@evangelinelillyofficial	1.256	352.94	7058.87	0.001
@chethanx	0.728	48.12	962.35	0.001
@RobSchneider	0.742	48.99	979.85	0.001
@NICKIMINAJ	0.487	-123.84	-2476.88	0.001

Table 2. Weighted ensemble estimates for influencers sharing anti-vaccination messages

For each influencer in this category, all models predicted a significant impact — furthermore, each impact was fairly large in magnitude (note that not all of these impacts were positive). The weighted ensemble (the most robust of the models built) did indeed perform near the best in every case.

DISCUSSION AND CONCLUSIONS

Based on the results from the "positive" influencers, it seems as though statements made in early 2021 were more impactful than those made later. This may be because the global focus on vaccines decreased as time went on. Indeed, the estimated impacts mostly decrease as one moves through the "positive" influencers chronologically. Overall, the results for the "negative" influencers were more volatile than the "positive", with many influencers' results being disputed amongst the models. However, the patterns found amongst these influencers imply that "negative" posts made by influencers did indeed affect the conversations Twitter users had regarding the vaccines. In fact, in almost every model in the "negative" section, the estimated cumulative impacts were on the order of 10^3 or greater. However, not all these impacts were positive; in fact, in multiple instances, a "negative" post by an influencer was able to significantly increase the number of "positive" tweets in the following days. This phenomenon can be directly observed by reading selected "negative" posts by influencers on a selected platform — many users, in response to the statements in the original post, begin to provide their own counterarguments in favor of the vaccine. This generates a "counter-effect", where a "negative" post leads to a higher influx of "positive" tweets. Some influencers (e.g., @NICKIMINAJ) fell victim to this counter-effect according to all models, whereas others (e.g., @evangelinelillyofficial) had conflicting results between models.

This study aimed to evaluate the effect influencers had on the public sentiment of COVID-19 vaccines from late 2020 to early 2022. Based on the results, pro-vaccination messages sent out by influencers tend to considerably, and statistically significantly, increase the number of pro-vaccination tweets in the following days. Influencers that send out anti-vaccination messages have large, statistically significant impacts, although sometimes in the opposite. Future work in this area may include adding more influencers to the analysis, to develop a more comprehensive picture of the impact

– looking at more features (e.g., seasonality) would also help make the results more robust. Also, this study did not establish causality — although the CausalImpact library was used, the counterfactuals were built solely based on prior data. Designing a randomized experiment to determine whether a post actually *causes* the observed differences should be explored in the future. The results of this study provide a clearer picture of the impact celebrities and social media influencers have on the public perception of vaccines. These findings can inform public health strategy and aid health officials in better understanding vaccine hesitancy; this understanding will be crucial as information — and misinformation — spreads with increasing speed on social media.

REFERENCES

1. Adams, S. H., Schaub, J. P., Nagata, J. M., Park, M. J., Brindis, C. D., and Irwin Jr, C. E. (2021). Young adult perspectives on COVID-19 vaccinations. Journal of Adolescent Health, 69(3):511–514.Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

2. Afifi, T. O., Salmon, S., Taillieu, T., Stewart-Tufescu, A., Fortier, J., and Driedger, S. M. (2021). Older adolescents and young adults willingness to receive the COVID-19 vaccine: Implications for informing public health strategies. Vaccine, 39(26):3473–3479.

3. Bloom, D. E., Cadarette, D., Ferranna, M., Hyer, R. N., and Tortorice, D. L. (2021). How New Models Of Vaccine Development For COVID-19 Have Helped Address An Epic Public Health Crisis. Health Affairs, 40(3):410–418.

4. Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., and Scott, S. L. (2015). Inferring causal impact using Bayesian structural time-series models. The Annals of Applied Statistics, pages 247–274.

5. Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pages 785–794.

6. Godlee, F. (2019). What should we do about vaccine hesitancy?

7. Hendaus, M. A. and Jomha, F. A. (2021). mRNA vaccines for COVID-19: A simple explanation. Qatar Medical Journal, 2021(1).

8. Hoffman, S. J. and Tan, C. (2013). Following celebrities' medical advice: meta-narrative analysis. Bmj, 347.

9. Hutto, C. and Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media, volume 8(1), pages 216–225.

10. Kaikati, J. G. (1987). Celebrity advertising: A review and synthesis. International Journal of Advertising, 6(2):93–105.

11. MacDonald, N. E. et al. (2015). Vaccine hesitancy: Definition, scope and determinants. Vaccine, 33(34):4161–4164.

12. Randolph, H. E. and Barreiro, L. B. (2020). Herd immunity: understanding COVID-19. Immunity, 52(5):737–741.

13. Robinson, W. A. (2003). AIDS in the African-American community: celebrities as an ineffective communication strategy.

14. Ryan, J. M. (2021). Timeline of COVID-19. COVID-19: Global pandemic, societal responses, ideological solutions, pages xiii–xxxii.

15. Scott, S. L. and Varian, H. R. (2013). Predicting the present with Bayesian structural time series. Available at SSRN 2304426.

16. Selkie, E. (2022). Influence at the Intersection of Social Media and Celebrity. JAMA Network Open, 5(1):e2143096–e2143096.

17. Sloss, M. (2021). 17 Famous People Who Got A COVID-19 Vaccine, And 7 Who Absolutely Refuse To.

18. Sullivan, K. (2021). Pop star Olivia Rodrigo visits White House to urge young people to get vaccinated against Covid-19.

19. Taylor, S. J. and Letham, B. (2018). Forecasting at scale. The American Statistician, 72(1):37–45.

20. Yannopoulos, P. (2012). Celebrity advertising: Literature review and propositions. World Review of Business Research, 2(4):24–36.

21. Zafar, A. U., Qiu, J., Li, Y., Wang, J., and Shahzad, M. (2021). The impact of social media celebrities' posts and contextual interactions on impulse buying in social commerce. Computers in Human Behavior, 115:106178.

22. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. New England journal of medicine.