

Artificial intelligence-enabled analysis of UK and US public attitudes on Facebook and Twitter towards COVID-19 vaccinations

Amir Hussain^{1*}, Ahsen Tahir¹, Zain Hussain², Zakariya Sheikh², Mandar Gogate¹, Kia Dashtipour¹, Azhar Ali³, Aziz Sheikh^{4*}

1. School of Computing, Edinburgh Napier University, UK
2. Edinburgh Medical School, College of Medicine and Veterinary Medicine, University of Edinburgh, UK
3. NHS Forth Medical Group, UK & Harvard T.H. Chan School of Public Health, USA
4. Usher Institute, Edinburgh Medical School, University of Edinburgh, UK

*Corresponding Authors:

Professor Amir Hussain
School of Computing
Edinburgh Napier University
Edinburgh, EH10 5DT, Scotland, UK
E-mail: a.hussain@napier.ac.uk
<https://www.napier.ac.uk/people/amir-hussain>

Professor Aziz Sheikh
Usher Institute
College of Medicine and Veterinary Medicine
University of Edinburgh
Edinburgh, EH16 4UX, Scotland, UK
E-mail: aziz.sheikh@ed.ac.uk
<https://www.ed.ac.uk/profile/aziz-sheikh>

Abstract

Background: Global efforts towards the development and deployment of a vaccine for SARS-CoV-2 are rapidly advancing. We developed and applied an artificial-intelligence (AI)-based approach to analyse social-media public sentiment in the UK and the US towards COVID-19 vaccinations, to understand public attitude and identify topics of concern.

Methods: Over 300,000 social-media posts related to COVID-19 vaccinations were extracted, including 23,571 Facebook-posts from the UK and 144,864 from the US, along with 40,268 tweets from the UK and 98,385 from the US respectively, from 1st March - 22nd November 2020. We used natural language processing and deep learning based techniques to predict average sentiments, sentiment trends and topics of discussion. These were analysed longitudinally and geo-spatially, and a manual reading of randomly selected posts around points of interest helped identify underlying themes and validated insights from the analysis.

Results: We found overall averaged positive, negative and neutral sentiment in the UK to be 58%, 22% and 17%, compared to 56%, 24% and 18% in the US, respectively. Public optimism over vaccine development, effectiveness and trials as well as concerns over safety, economic viability and corporation control were identified. We compared our findings to national surveys in both countries and found them to correlate broadly.

Conclusions: AI-enabled social-media analysis should be considered for adoption by institutions and governments, alongside surveys and other conventional methods of assessing public attitude. This could enable real-time assessment, at scale, of public confidence and trust in COVID-19 vaccinations, help address concerns of vaccine-sceptics and develop more effective policies and communication strategies to maximise uptake.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Introduction

The imminent availability of COVID-19 vaccines poses a pressing need to continually monitor and better understand public sentiments in order to develop baseline levels of confidence in vaccines and enable identification of early warning signals of losses in confidence [1]. This will help address the concerns of vaccine sceptics [2,3,4] and develop required public trust in immunisation [5,6] to realise the goal of herd immunity [7].

Traditionally, governments use surveys to understand public attitude; however, these typically suffer from small sample sizes, closed questions and limited spatio-temporal granularity. In order to overcome these limitations, we argue social-media data can be used to increase numbers and enable real-time analyses of public sentiments and attitudes with considerable spatiotemporal granularity. Over half the world population, including around 70% of both the UK and US populations are active social-media users, with significantly increased usage reported during the pandemic, e.g. by 37% for Facebook [8]. Since social-media data is largely unstructured, it is amenable to application of established AI techniques, such as machine learning (ML), deep learning (DL) [9] and natural language processing (NLP) [10], to extract topics and sentiments from the social-media posts.

Sentiment analysis involves categorising subjective opinions from text, audio and/or video [10] to obtain polarities (e.g. positive, negative, neutral), emotions (e.g. angry, sad, happy) or states of mind (e.g. interested versus uninterested) towards target topics, themes or ‘aspects’ of interest [11]. A complementary approach, termed stance detection [12], assigns a stance-label (favourable, against, none) to a post towards a specific predetermined target, which in itself may not be referred to, or be the target of opinion in the post. Such approaches are currently under-utilised in healthcare. In particular, there is significant untapped potential in drawing on AI-enabled social-media analysis to inform public policy research.

Methods

Ethics

We conducted a thorough assessment of the privacy risk to individuals posed by our research, in light of [13,14], to ensure compliance with relevant sections of the General Data Protection Regulation (GDPR). Further we have striven to comply with best practices for user protection [15,16], ensuring no non-public material is included in our dataset.

Data Sources

We used data from both Facebook and Twitter, two of the most popular and representative social-media platforms [8][17]. We used English-language Facebook posts and tweets that were posted in the UK and the US from 1st March to 22nd November 2020. Facebook posts were obtained through the CrowdTangle platform [18], and Twitter posts from a publicly available Twitter API. We utilised hydrated tweets from the global COVID-19 dataset available at [19], which collects up to 4.4 million tweets per day including retweets, and up to 1.1 million cleaned tweets without the re-tweets. The total number of tweets hydrated and utilised for this study were over 158 million. The Facebook posts and tweets were *thematically* filtered for both COVID-19 and vaccine related keywords and then geographically filtered for the UK and the US. The first step filtering with COVID-19 related keywords utilised widely used terms from [19] (see Appendix Section-A1). The vaccination terms used for second step filtering were selected by our team: vaccine, vaccination, immunise, immunize, immunisation and immunization. The two-step thematic filtering process was applied using these keywords before processing and analysis.

Analysis

The filtered dataset was initially pre-processed (e.g. removing links, hashtags, stop words) and a new hierarchical hybrid-ensemble based AI model was developed for thematic sentiment analysis. This utilised an average-weighting ensemble [20] of two lexicon-based methods: Valence-Aware Dictionary and sEntiment-Reasoner (VADER) [21] and TextBlob [22]. These were combined with a pre-trained DL-based model: Bidirectional Encoder-Representations from Transformers (BERT) [23], using a rule-based ensemble method, as illustrated in Figure 1.

A random 10% sample of posts and tweets were then manually annotated by the team, and checked against our hybrid-ensemble AI model’s sentiment classifications for refinement and validation. The hybrid-ensemble model was optimised based on the validation results, with sensitivity and specificity analysis showing that the lexicon-based methods provided generally better accuracy for positive sentiments, and the BERT model generally provided better performance for neutral and negative sentiments. The final rule based hybrid-ensemble method utilised the classification output of the BERT model for neutral and negative sentiments, while the weighted-average output of lexicon-based methods was selected for classifying positive sentiments.

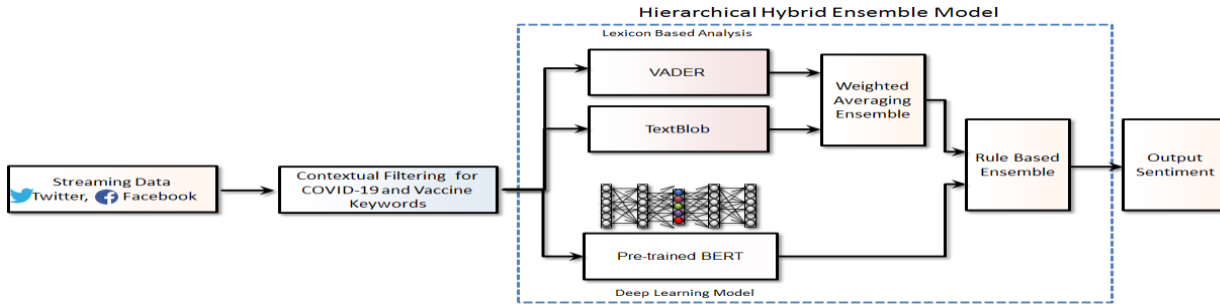


Figure 1: Hierarchical hybrid-ensemble based AI model and data pipeline for thematic sentiment analysis

A number of established NLP techniques were used to analyse the processed data (see details in Appendix Section-A2). Specifically, in addition to analysing averaged sentiment trends and their geo-spatial mappings in the UK and US, we statistically analysed the trends (using Pearson’s r) and compared findings with independent surveys. Sentiment word cloud and N-gram analysis was applied to specific time-periods of interest, around points of inflexion on sentiment trend graphs, to identify topics of discussion and glean insight into the positive and negative content of online discourses. The analysis was also carried out over the full period of study, to identify underlying themes and topics. Findings were validated, and further insights obtained, through a manual reading by our team, of randomly selected posts around target points of interest. Relevant social-media datasets and outputs were anonymised, and statistical aggregates made openly available for transparency and reproducibility (including through a publicly available dashboard [24]).

Results

Temporal Sentiment Trends

Monthly volume trends of the filtered UK and US Facebook and Twitter posts used for the target period of study are shown in Appendix (Figure-A1). Figure 2 illustrates the averaged (weekly) positive, negative and neutral Facebook sentiments from March–November 2020, for the UK and US. We identified topics of discussion around points of interest on the graphs. These are referred to in our descriptive analysis of the graphs below, and some are highlighted in Figures 2 and 3. It was interesting to note that on Facebook, the difference between the averaged positive and negative sentiment trends was more pronounced, compared to Twitter. For the UK, the positive Facebook sentiment trend graph was most prevalent (Figure 2(a)), and rose steadily since May, corresponding to the initiation of vaccine trials and the recruitment of the first volunteer. There was a peak in mid-August, possibly associated with news on vaccine developments in the UK and Russia. The negative sentiment trend graph inversely followed the positive trend, and discourse was centered around vaccine conspiracies and halting of trials. For the US, the positive Facebook sentiment trend was again most prevalent, and had a small peak in August, associated with posts relating to research on COVID-19 vaccine. Moreover, the negative sentiment trend graph had slightly increased in mid-September, associated with posts relating to the COVID-19 vaccine being rushed (Figure 2(b)). More recently, from mid-October onwards, there has been a growing positive sentiment trend for both the UK and the US in part attributed to the announcements from Pfizer and Moderna of successful vaccine trials [25].

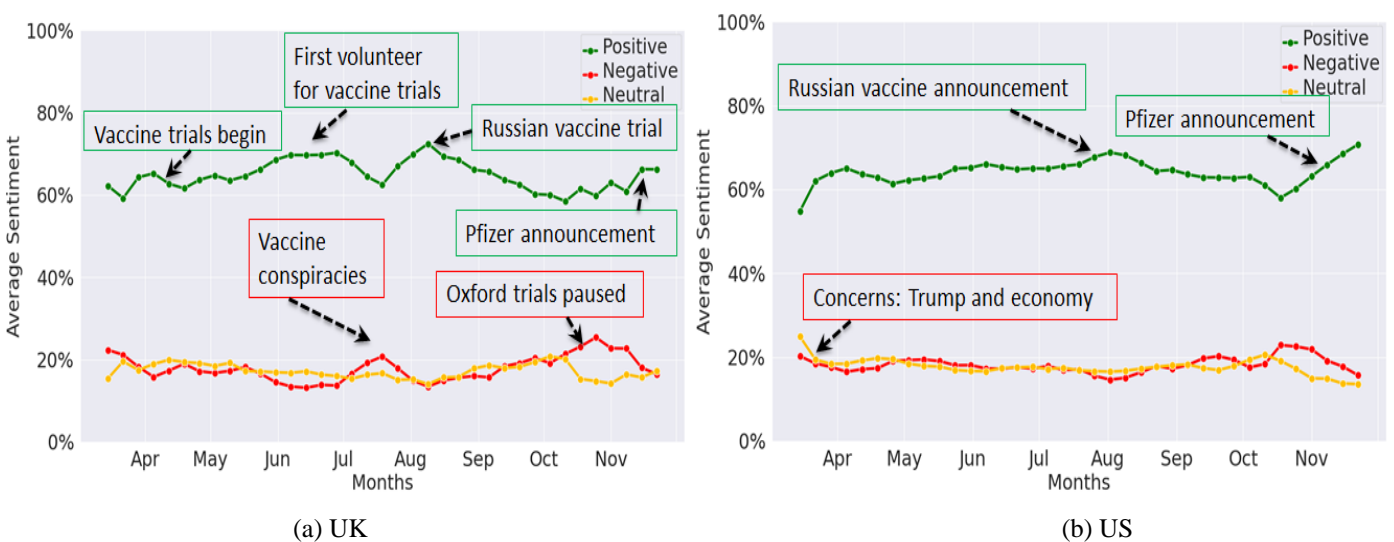


Figure 2: Averaged weekly Facebook sentiment trends for (a) UK and (b) US

Figure 3 illustrates the positive, negative and neutral Twitter sentiments from March to November 2020, for the UK and the US. We can see that in the UK (Figure 3(a)), positive Twitter sentiment was the most prevalent and its trend graph had small peaks at the end of April and July 2020. The former was found to be related to the first human vaccination trial. The negative sentiment trend had peaks in July and October 2020, simultaneous to the UK opting out of the European Union vaccination scheme, and the

Oxford vaccination trials being paused [26] due to safety concerns respectively. In the US (Figure 3(b)), positive sentiment was also more prevalent, and had a major peak from end September to end November. This was related to claims by President Trump on a vaccine being ready in a few weeks, and an increase in twitter discourse due to his reference to the “herd mentality”. A small peak in the negative trend graph in mid-September was related to pausing of the Oxford vaccination trials.

For both the UK and the US, there has been a noticeable marked increase in the positive sentiment trend, since end-October, which we found related to recent breakthrough announcements by Pfizer and Moderna. Analysis of social media conversation indicated public optimism, with trial results being hailed as “good” and “amazing”, there being “hope” for the “new year” (see Appendix Figures A2 and A3). A notable peak in the negative sentiment trend graphs for both countries, around mid-October, was associated with the growing anti-vaccination movement, and with concerns around “fake news” and “misinformation”.

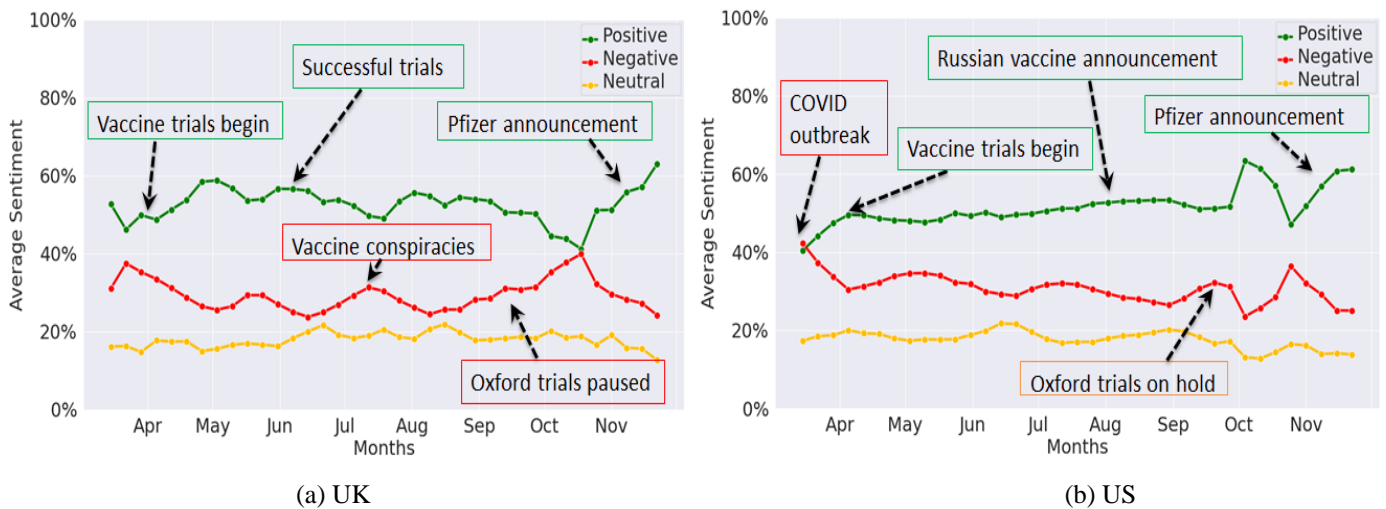


Figure 3: Average weekly Twitter sentiment trends for (a) UK and (b) US

Statistical Analysis of Sentiment Trends

Statistical findings are detailed in the Appendix Section-A3, which assessed the strength of the association between the predicted sentiment in the trend graph and the accuracy of the labelled data. In conclusion, there was a stronger sentiment on Twitter in relation to COVID-19 vaccinations for the US, with both positive and negative sentiments demonstrating stronger increasing and decreasing trends, as compared to the UK. Public sentiment on Facebook was found to represent decreasing positive sentiment and increasing neutral sentiment for both the UK and the US, with positive sentiment demonstrating a slightly stronger decreasing trend in the UK than the US.

Sentiment Word Clouds and Text N-Gram Analysis

We applied these techniques to the entire period of study, to identify and analyse notable events that were of interest to social-media users, and summarise them in the Appendix: Section-A4, Figures A2, A3 and Tables A1, A2 (some of these were also identified in the analysis above, on the sentiment trend graphs).

Geo-spatial Sentiment Analysis

The geo-spatial mapping of overall (averaged) sentiments to states in the US are shown in Figure 4 (left) and indicate that most states have a negative sentiment. The states with an overall negative sentiment towards COVID-19 vaccination were found to be concentrated in the West and Midwest of the country, namely: Idaho, Kansas, New Hampshire, West Virginia, Alabama. The states with an overall positive sentiment were in the East, namely, Maine, Colorado, Georgia, Hawaii

The geo-spatial mapping of averaged sentiments to counties in the UK is illustrated in Figure 4 (right). In contrast with the US, most counties in the UK had an overall positive sentiment towards COVID-19 vaccination. The counties with the most positive sentiment included Cornwall, Kent, East Sussex, Surrey and Dorset all in England, and Aberdeenshire, Angus and Stirlingshire all in Scotland. Furthermore, the counties with the most negative sentiment were West Sussex, Somerset, North Yorkshire and Durham, all in England.

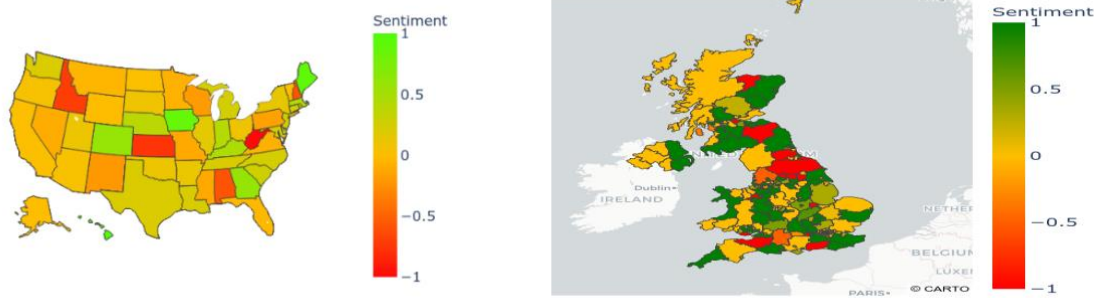


Figure 4: Geo-spatial mapping of averaged social-media public sentiment in US (left) vs UK (right) related to COVID-19 vaccination (+1 or green indicates positive sentiment, 0 neutral, and -1 or red indicates negative sentiment)

Overall Averaged Sentiment

Overall averaged sentiments in the UK and US on Facebook and Twitter are shown in Figure 5 below (and described in Appendix Section-A5).

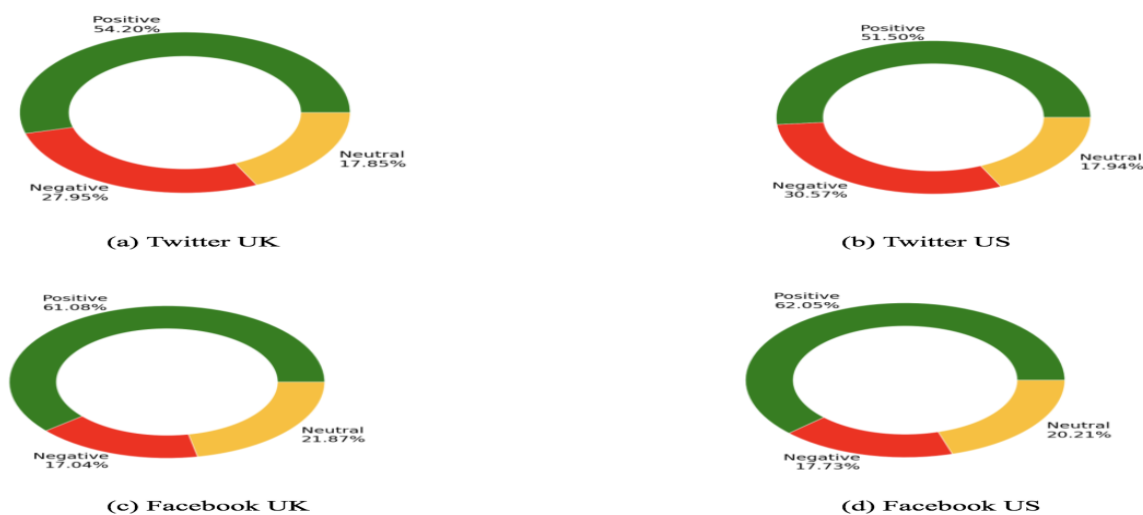


Figure 5: Overall averaged sentiment: (a) Twitter UK (b) Twitter US (c) Facebook UK (d) Facebook US

Discussion

We analysed temporal variations in public sentiments on COVID-19 vaccination in the UK and the US. Key events impacting positive, negative and neutral sentiments were identified, evaluated and mapped to the temporal trends. We also mapped spatial variations in public sentiment to regions and states of the UK and the US respectively. Our geo-spatial mappings can help identify areas with more negative sentiment towards a COVID-19 vaccine which can be further studied for potential interventions, to allay underlying public fears and concerns.

Our analysis has shown that online public discourse on Facebook and Twitter platforms across both the UK and the US is evolving, with both complementary and contrasting insights gleaned from the two popular platforms. Comparative analysis results indicate that over the nine-month period of study, averaged public sentiment towards COVID-19 vaccines has been mostly positive and similar in both the UK and the US across both platforms (57.70% average across both platforms for UK vs 56.80% for the US). Positive sentiment was found to be related to public opinion on vaccine development, related trials and news related to the availability of a vaccine.

The overall averaged negative sentiments, on both platforms, were also found to be similar for the UK (22.50%) and US (24.10%). It is interesting to note that Twitter sentiment appears to be more negatively biased, with the proportion of negative sentiment being almost double that of Facebook, for both the UK (27.95% vs 17.04%) and also the US (30.57% vs 17.73%), potentially reflecting their respective user demographics. This also appears to be consistent with the study in [17], which found public opinion on Twitter to be often more negatively biased than on Facebook, where it is more positively biased. Negative sentiments in our study were

found to relate to public apprehensions and concerns around delays or pauses in vaccine trials, vaccine safety, corporations and governments influencing vaccine availability and rights exclusivity for economic benefits.

A comparative analysis with independent surveys was carried out. Our findings relating to averaged positive and negative sentiment trends across both the UK and US were found to correlate broadly. In the US, during the early stages of the pandemic, polling indicated that a significant minority had low trust in a vaccine – for example, a Yahoo News/YouGov survey in May [27] found only 55% of Americans planned to get vaccinated against COVID-19, whilst almost one in five (19%) would not get vaccinated. A similar survey in July [28] indicated 42% would get vaccinated (27% would not get vaccinated), whilst a survey in September [29] found only 36% were certain they would get vaccinated (32% would not get vaccinated). More recently, an Axios-Ipsos survey in November reported, consistent with our social-media analysis findings, a marked increase in those likely to get vaccinated (51% were ‘very’ or ‘somewhat’ likely to get the first generation vaccine, increasing to 70% if the vaccine was proven to be safe and effective by public health officials) [30].

In the UK, a YouGov survey in June [31] found 41% of respondents would ‘probably’ or ‘definitely’ get vaccinated, whilst one in six (16%) of respondents ‘definitely’ or ‘probably’ would not get vaccinated. The survey also found individuals who used social-media more than traditional media as their source for news, were nine percentage points less likely to be in favour of being vaccinated. A more recent UK YouGov survey in November [32], in relation to the Pfizer COVID-19 vaccine, found 67% were ‘very’ or ‘fairly’ likely to take the vaccine when available, with one in five (21%) unlikely to take it. Whilst there has been a slight increase in the proportion unlikely to take the vaccine, the proportion of those likely to get vaccinated has increased, indicating a decrease in the number of those who were previously unsure. This could be attributed in part to the recent announcements by vaccine manufacturers, and our results corroborate this with a marked increase reported in positive sentiment since mid-October, in both the UK and US, across the two social-media platforms. Further study into the change in sentiment could help us better understand factors that have contributed to this, in particular the impact of government education programmes.

It is important to consider the limitations of our data sources and techniques, and related challenges and opportunities they present for required future research. Whilst we attempted to gauge country-wide public sentiments in the UK and US, by analysing English-language posts on both Facebook and Twitter, our data may not be representative of the broader population. Users are known to differ in their social-media platform preference and usage, based on their socio-demographics (e.g. age, socio-economic status, political affiliation). Vaccinations are likely to be preferentially targeted at older populations and possibly ethnic minorities, communities which have historically lower rates of vaccination uptake (e.g. [33,34]). Further exploration is therefore imperative to increase our understanding of public perception towards vaccines, and their underlying behavioral determinants [35]. Social network analysis [36] can be used in conjunction with DL methods to effectively identify sources of fake news/misinformation and their social networks, to help deal with ‘infodemic’ challenges [37]. Demographic information, such as age, gender, race and geographic origin can also be inferred from user social-media profiles using AI techniques [38]. This can help categorise distinct groups and inform the development of demographic-level engagement and tailored communications strategies to promote diversity and inclusion in vaccination campaigns. These can also effectively account for the fact that there are genuine knowledge voids being filled by misinformation [35].

Technical limitations of our approach include challenges in determining the geographical location of users, and issues relating to the accuracy of the AI techniques (e.g. dealing with sarcasm, implicit context). The two-step keyword-based thematic filtering process and use of geo-tagged posts in the study resulted in relatively small sample sizes. This could be improved by using more sensitive filtering and data-driven search mechanisms, network meta-data (such as likes, retweets) and additional social-media and web platforms. On account of the current limitations, our approach should only be used in conjunction with other techniques for understanding public sentiments, such as focus groups, input from civil society organisations, surveys and public consultations.

Future studies could consider conducting periodic public surveys over the period of interest being explored by social-media analysis. This would ensure both methodologies were informed by each other over the course of the study to enable more fine-grained spatio-temporal analysis, allowing more robust comparisons from reciprocal findings and deeper insights for policy makers. These could also complement other qualitative methods, such as in-depth interviews and ethnographic studies as part of mixed-study approaches. Manual annotation/labelling of datasets is imperative when training AI models for NLP tasks, to ensure accuracy and generalisability. These can be affected by the skill of annotators and the proportion of the dataset that is labelled. Confounding factors, such as political affiliations, should also be included in future studies, by applying further filters to search strategies and through targeted demographic analysis, to better understand underlying determinants of public sentiment. Attitudes towards different vaccine manufacturers could also be explored, to identify and assess effective public engagement strategies to build support for ethical principles, and maximise uptake of the imminently available vaccines.

One of the main threats to the resilience of vaccination programmes globally is the rapid and global spread of misinformation. The public's confidence in COVID-19 vaccine is known to be exacerbated by unproven vaccine safety scares seeding doubt and distrust. There are also cases where vaccine debates have been purposefully polarised, exploiting the doubting public and system weaknesses for political purposes, while waning vaccine confidence in other places may be influenced by a general distrust in the government and scientific elites. Recent surveys and polls in the UK and the US have indicated the fragility of support for

vaccination, which furthers the case for a better understanding of underlying public concerns and attitudes, both at scale and in real time. Retrospective analysis of two popular and most representative social-media platforms in this study demonstrates the potential of AI-enabled real-time social-media monitoring of public sentiments and attitudes to help detect and prevent such fears, and also to enable policymakers better understand the reasons behind why some social groups may be reluctant to be vaccinated against COVID-19. This can inform more effective policy making and promote participatory dialogues around complex vaccine deployment issues, under conditions of uncertainty, including decisions on prioritisation and equitability, to help maximise the uptake of imminently available vaccines.

Data and Code availability: The code for data analysis and figures generation are openly available at: <https://gitlab.com/covid19aidashboard/covid-vaccination/> for reproducibility and transparent analysis. Our Twitter dataset was obtained using the publicly available Twitter API. Our Facebook dataset was obtained using the CrowdTangle platform, for which access must be requested from the organisation. All analysis was carried out in Python. Our code can be used as a reference for conducting similar studies. Due to the computationally expensive nature of the code, we recommend using a high performance computing resource.

Acknowledgements: The study reported here is funded by the Scottish Government Chief Scientist Office under its COVID-19 priority research programme (grant ref. COV/NAP/20/07) and is jointly led by co-Principal Investigators: AH and AS. AS is supported by BREATHE - The Health Data Research (HDR) Hub for Respiratory Health [MC_PC_19004], which is funded through the UK Research and Innovation Industrial Strategy Challenge Fund and delivered through HDR UK. AH is supported by the UK Government's Engineering and Physical Sciences Research Council (EPSRC) grants (Ref. EP/T021063/1, EP/T024917/1).

Conflicts of Interest: As is a member of the Scottish Government's Chief Medical Officer's COVID-19 Advisory Group and the UK Government's New and Emerging Respiratory Virus Threats (NERVTAG) Risk Stratification Subgroup. This work does not represent the views of the Scottish or UK Governments.

Abbreviations

BERT: Bidirectional Encoder Representations from Transformers

COVID-19: coronavirus disease

DL: Deep Learning

ML: Machine Learning

NLP: Natural Language Processing

P: p-value

VADER: Valence Aware Dictionary and sEntiment Reasoner

References

1. de Figueiredo A, Simas C., Karafillakis E., Paterson P., Larson, H. J., Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study, *The Lancet* 396 (2020) 898–908
2. Gellin, B., 2020. Why vaccine rumours stick—and getting them unstuck. *The Lancet*, 396(10247), pp.303-304.
3. Horder, J. Toll of vaccine hesitancy. *Nat Hum Behav* 4, 335 (2020). <https://doi.org/10.1038/s41562-020-0874-1>
4. WHO, Ten threats to global health in 2019, <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>, 14th Jan, 2019. Accessed: 26th Oct, 2020.
5. Lazarus, J.V., Ratzan, S.C., Palayew, A. *et al.* A global survey of potential acceptance of a COVID-19 vaccine. *Nat Med* (2020). <https://doi.org/10.1038/s41591-020-1124-9>
6. Bloom B R., Nowak, G. J. and Orenstein, W, “When Will We Have a Vaccine?” — Understanding Questions and Answers about Covid-19 Vaccination, *N Engl J Med*, Sep 8, 2020 DOI: 10.1056/NEJMp202533
7. Fontanet, A. and Cauchemez, S., COVID-19 herd immunity: where are we?. *Nature Reviews Immunology*, 20(10), pp.583-584 (2020)
8. Available from:
<https://techcrunch.com/2020/03/26/report-whatsapp-has-seen-a-40-increase-in-usage-due-to-covid-19-pandemic/>
9. Adeel A, Gogate M, Hussain A, Contextual deep learning-based audio-visual switching for speech enhancement in real-world environments, *Information Fusion*, Volume 59, Pages 163-170, (2020)
10. Dashtipour K, Gogate M, Li J, Jiang F, Kong B, Hussain A. A hybrid Persian sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks, *Neurocomputing*, Vol.380, (2020):1-10.
11. Majumder, N., Bhardwaj, R., Poria, Gelbukh A, Hussain A. Improving aspect-level sentiment analysis with aspect extraction. *Neural Computation & Applications* (2020) <https://doi.org/10.1007/s00521-020-05287-7>
12. Al-Ghadir A.I, Azmi A.M, Hussain A, A novel approach to stance detection in social-media tweets by fusing ranked lists and sentiments, *Information Fusion*, 67:29-40 (2021)
13. Zimmer M. "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, 12(4):313–325, (2010)

14. Williams M.L, Burnap P, and Sloan L. Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views, Online Context and Algorithmic Estimation. *Sociology*,51(6):1149–1168 (2017)
15. Franzke A.S., Bechmann A, Zimmer M, and Ess C.M.. Internet Research: Ethical Guidelines3.0. <https://aoir.org/reports/ethics3.pdf> (2020)
16. Available from: <https://esrc.ukri.org/files/about-us/policies-and-standards/esrc-research-data-policy/> (2018)
17. Waterloo S.F, Baumgartner S.E, Peter J., Valkenburg P.M, Norms of online expressions of emotion: Comparing facebook, twitter, instagram, and whatsapp, *New Media & Society*, 20:1813–1831 (2018)
18. Garmur, M; King, G; Mukerjee, Z; Persily, N; Silverman, B, 2019, "CrowdTangle Platform and API", <https://doi.org/10.7910/DVN/SCCQYD>, Harvard Dataverse, V3 (2019)
19. Banda J.M, Tekumalla R., Wang G., Yu J., Liu T., Ding Y., Chowell G., A large-scale covid-19 twitter chatter dataset for open scientific research—an international collaboration, arXiv preprint arXiv:2004.03688 (2020)
20. Yoo, B.H., Kim, J., Lee, B. *et al.* A surrogate weighted mean ensemble method to reduce the uncertainty at a regional scale for the calculation of potential evapotranspiration. *Sci Rep* **10**, 870 (2020).
21. Gilbert C., Hutto E., Vader: A parsimonious rule-based model for sentiment analysis of social media text, in: *Eighth International Conference on Weblogs and social-media (ICWSM-14)*. Available at (20/04/16) <http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf>, volume 81, 2014, p. 82.
22. Loria S, Keen P, Honnibal M, Yankovsky R, Karesh D, Dempsey E,et al., Textblob: simplified text processing, Secondary TextBlob: simplified text processing 3 (2014).
23. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, in: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume1 (Long and Short Papers), pp. 4171–4186. (2019)
24. Available from: <http://covidtracker.cloud>
25. Doshi, P., 2020. Peter Doshi: Pfizer And Moderna'S "95% Effective" Vaccines—Let'S Be Cautious And First See The Full Data - The BMJ. [online] The BMJ. Available at: <<https://blogs.bmj.com/bmj/2020/11/26/peter-doshi-pfizer-and-modernas-95-effective-vaccines-lets-be-cautious-and-first-see-the-full-data/>>
26. Mahase E. Covid-19: Oxford researchers halt vaccine trial while adverse reaction is investigated. *BMJ*. 370:m3525. doi:10.1136/bmj.m3525 pmid:32907856 (2020)
27. *Who will get a COVID-19 vaccine when it arrives.* (2020, May 11). YouGov | What the world thinks. <https://today.yougov.com/topics/health/articles-reports/2020/05/11/who-will-get-covid-vaccine>
28. *Fewer Americans say they'll get vaccinated for COVID-19.* (2020, August 5). YouGov | What the world thinks. <https://today.yougov.com/topics/politics/articles-reports/2020/08/05/vaccinated-covid-19>
29. *Concerns over fast-tracked COVID-19 vaccine has Americans unsure about vaccination.* (2020, September 15). YouGov | What the world thinks. <https://today.yougov.com/topics/health/articles-reports/2020/09/15/concerns-over-fast-tracked-covid-19-vaccine-has-am>
30. Majority of Americans expect to not pay for COVID-19 vaccine. (2020, November 24). Ipsos. <https://www.ipsos.com/en-us/news-polls/axios-ipsos-coronavirus-index>
31. Available from: https://docs.cdn.yougov.com/5mkju0kxbj/CCDH_RESULTS_062620_PR%20%28002%29.pdf
32. *How many britons are willing to take a coronavirus vaccine?* (2020, November 16). YouGov | What the world thinks. <https://yougov.co.uk/topics/health/articles-reports/2020/11/16/how-many-britons-are-willing-take-coronavirus-vacc>
33. Forster AS, Rockliffe L, Chorley AJ, et al. Ethnicity-specific factors influencing childhood immunisation decisions among Black and Asian Minority Ethnic groups in the UK: a systematic review of qualitative research. *J Epidemiol Community Health*. 71(6):544-549. (2017)
34. Malik, A.A., McFadden, S.M., Elharake, J. and Omer, S.B., Determinants of COVID-19 Vaccine Acceptance in the US, *US. Eclinical Med* 2020;100495. (2020)
35. COVID-19 vaccine deployment: Behaviour, ethics, misinformation and policy strategies. *The Royal Society*. 21 October 2020. Available from: <https://royalsociety.org/-/media/policy/projects/set-c/set-c-vaccine-deployment.pdf>
36. Block, P., Hoffman, M., Raabe, I.J. *et al et al.* Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nat Hum Behav* 4, 588–596 (2020). <https://doi.org/10.1038/s41562-020-0898-6>
37. Gallotti, R., Valle, F., Castaldo, N. *et al.* Assessing the risks of 'infodemics' in response to COVID-19 epidemics. *Nat Hum Behav* (2020). <https://doi.org/10.1038/s41562-020-00994-6>
38. Brandt J et al. Identifying social-media user demographics and topic diversity with computational social science: a case study of a major international policy forum. *J Comput Soc Sc* 3,167–188 (2020)

