As vaccines against Covid-19 are rolled out, there is a pressing need to better understand and monitor public sentiments and address the concerns of vaccine sceptics. This urgency has been exacerbated by the current wave of the global pandemic and growing pressures on health services. The Royal Society reviewed the behavioral aspects of vaccine uptake and misinformation and identified the need for further research into five factors: confidence in the vaccines’ efficacy and safety; complacency; convenience; sociodemographic variation; and sources of information. The survey advocated the need for public dialogue around vaccine concerns and hesitancy to replace the current passive communication strategies that tend to focus narrowly on “infodemic” challenges, because genuine knowledge voids are being filled by misinformation. Artificial intelligence (AI) can enable real-time analysis of public attitudes, and their demographic determinants, from social media and linked Web platforms. This analysis offers the opportunity to track changing public sentiments and develop proactive two-way communication strategies. In particular, an iterative learning cycle can help identify unforeseen areas of public concern as well as potential barriers and facilitators for required interventions, prior to their implementation, to maximize vaccine uptake across demographic communities.

Nearly 3 billion people worldwide currently engage with at least one social media platform, including around 70% of the United States and UK populations. Recent studies have demonstrated the utility of social media analysis for public health professionals and policy makers, for example, to predict medical conditions and track outbreaks of disease.
Social media usage has significantly increased during the pandemic (for example, by up to 37% for Facebook6), and plays an important role in providing social support, including for those suffering from long-term Covid-19 consequences.7 Social media data is largely unstructured, but natural language processing (NLP) and machine learning (ML)8 can analyze it on a large scale and identify trends in public opinion. There is currently growing interest in utilizing established NLP techniques such as sentiment analysis and topic modeling,9 in conjunction with ML, to mine social media data for health care applications.10

Sentiment analysis uses computational methods to identify opinions in text, audio, and/or video9,11 to determine the author’s attitudes toward topic(s) of discussion. These can be expressed as polarities (positive, negative, neutral), or feelings (interested versus uninterested). A complementary approach is stance detection12 where a stance label (favorable, against, or none) is assigned to a post regarding a specific target, whether or not that target is explicitly addressed in the post. For example, if someone asserts that a disease can be devastating, it can be inferred that the author has a favorable stance toward a vaccine for that disease. Topic modeling9 enables discovery of related text, such as topics and themes.

These approaches are currently underutilized in health care. Social media analysis holds significant untapped potential to inform public policy research.

Social media platforms can enable real-time assessment of public confidence in Covid-19 vaccination. Changing levels of public interest and sentiment can be tracked for a range of concerns relating to vaccine safety and effectiveness, as well as trust in science, pharmaceutical companies, and governments.13 The impact of policies and messaging can be monitored to assess public perceptions around decisions on prioritization and equitability.

Suitable search strategies can be developed to capture social media conversations and attitudes toward Covid-19 vaccinations. Anonymized data can be sourced from a wide range of social media and Web forums, such as Twitter, Facebook, news, moderated blogs, forums, and reviews. Some platforms such as Twitter give geographical locations for posts at a regional (e.g., city) level, while others such as Facebook Crowdtable14 do so at national levels.

Social media “dashboards” for policy makers15 could analyze and visualize “mentions” of specific topics or keywords with the highest levels of engagement (that is, the number of times a topic or keyword is mentioned). These analyses can assess trends in the volume and sentiment of mentions for each social media platform. NLP algorithms, such as topic modelling and sentiment word clouds (visual representations of how frequently words are mentioned, using font size to indicate relative frequency), can be utilized for time periods of interest. Insights into the content of discussions can be gleaned, particularly around inflection points on a trend (for example, correlating changes with specific events or the circulation of particular pieces of misinformation). Policy experts, as part of “Expert-in-the-loop” NLP frameworks, can identify common positive and negative themes surrounding vaccination issues, through their manual reading of random subsets of social media and related news posts. Views of underrepresented online groups can also be identified and analyzed. Relevant dashboard data sets and outputs can be anonymized and presented as statistical aggregates to ensure compliance with national privacy laws.
Open challenges include the urgent need to develop NLP techniques to detect the sources of viral spread of fake news and misinformation, including deceptive and sarcastic language sometimes used by vaccine skeptics. This effort will require new multilingual vaccination lexicons (or dictionaries mapping commonly used terms to positive, negative, and neutral sentiments), to capture cultural differences in language use across communities. These lexicons will also need to keep up with the evolution of pandemic-related terminology (for example, the term “Rona” sometimes used as a nickname for Covid-19, or terms like “vaccine nationalism” or “viral anxiety”). State-of-the-art deep learning–based NLP models, such as Deep Bidirectional Transformers,16 can be trained through transfer learning and fine-tuning for enhanced language understanding. The use of social-network analysis techniques9,17 can help identify sources of misinformation and their patterns of spreading, which can shed light on the anatomy of key strategies used by those who intend to mislead and misinform, such as use of emotion and anecdotes to impact rational decision-making.2

Users with varied demographic characteristics are likely to engage differently with social media platforms on vaccination-related topics. Established AI approaches11,17,18 can enable policy makers to monitor how fake and genuine messages resonate with specific audiences. These insights can be used to help craft messages more likely to resonate with diverse audiences. In contrast with health care, such targeted messaging strategies aimed at different social groups are already extensively and successfully used in commercial marketing and advertising.18

Recent research18 has demonstrated the potential of AI-enabled demographic analysis to estimate age, gender, ethnicity, and geographic origin of social media profiles, using a combination of self-reported location data, names, social media photographs, and inferential statistics of national census data. This can help identify and categorize demographically distinct groups, such as the elderly, ethnic minorities, and incarcerated people,19 who are likely to experience disproportionate effects of Covid-19. Such analysis can inform development of real-time monitoring metrics for demographic-level engagement and more effective communication strategies to promote diversity and inclusion in vaccination campaigns.

Other challenges include linking social media with external trustworthy sources of data such as polling, surveys, news, and clinical sources. This linkage is important because social media users may not be representative of the population at large, particularly in the context of Covid-19 vaccinations that are likely to be preferentially targeted at high-risk groups (including communities who have historically lower rates of vaccination uptake).13 In such cases, linked analysis with multiple social media and Web platforms, and use of network metadata such as likes and retweets, can extend the study population.

The linked analysis is also important to help reduce bias and increase trust20 in AI model predictions, because different demographics (ages, political affiliations, socioeconomic status) tend to use different platforms. Corroborated findings from independent sources can inform linkage decisions. For example, YouGov polls have shown that almost 1 in 5 Americans and 1 in 6 UK respondents state that they would be unlikely to get vaccinated against Covid-19.21 This reluctance can be attributed in part to concerns about politicization and widespread misinformation. Linked social media analysis can corroborate these findings and shed light on common determinants of
Global efforts to develop Covid-19 vaccines have advanced rapidly. Harnessing the opportunities offered by AI-enabled high-value decision-making and linked analysis can address the range of behavioral factors underpinning vaccine confidence and trust, and further our understanding of
their evolving impact on uptake. By analyzing social media and Web forum responses to ongoing vaccine trials and vaccine availability in real time, we can identify and prioritize demographic groups and regions where more targeted confidence and trust building is needed. Adopting continuously informed and proactive policy making and communication strategies will promote participatory dialogues to build support for ethical principles and maximize the deployment and uptake of life-saving vaccines.

Amir Hussain, PhD
Professor, School of Computing, Edinburgh Napier University, Edinburgh, United Kingdom

Aziz Sheikh, MD
Professor, Usher Institute, University of Edinburgh, Edinburgh, United Kingdom

Acknowledgements

The authors are supported by the Scottish Government Chief Scientist Office under its Covid-19 priority research program (grant ref. COV/NAP/20/07). AS is also 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).

Disclosures: Amir Hussain has nothing to disclose. Aziz Sheikh 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.

References

1. Hunter DJ. Trying to “Protect the NHS” in the United Kingdom. N Engl J Med.

2. Covid-19 vaccine deployment: behaviour, ethics, misinformation and policy strategies. The Royal Society. October 21, 2020. Accessed January 13, 2021. https://royalsociety.org/-/media/policy/projects/set-c/set-c-vaccine-deployment.pdf.

3. Bloom BR, Nowak GJ, Orenstein W. “When Will We Have a Vaccine?” — Understanding questions and answers about Covid-19 vaccination. N Engl J Med. 2020;383(6):2202-4

4. Merchant RM, Asch DA, Crutchley P. Evaluating the predictability of medical conditions from social media posts. PLoS One.

5. Samaras L, García-Barriocanal E, Sicilia MA. Comparing Social media and Google to detect and predict severe epidemics. Sci Rep.
6. Perez S. Report: WhatsApp has seen a 40% increase in usage due to Covid-19 pandemic. TechCrunch. March 26, 2020. Accessed January 13, 2021. https://techcrunch.com/2020/03/26/report-whatsapp-has-seen-a-40-increase-in-usage-due-to-covid-19-pandemic/.

7. Yelin D, Wirtheim E, Vetter P. Long-term consequences of COVID-19: research needs. Lancet Infect Dis. 2020;20(6):1115-7

8. Ganguli I, Gordon WJ, Lupo C, et al. Machine learning and the pursuit of high-value health care. NEJM Catal Innov Care Deliv 2020;1(6) https://catalyst.nejm.org/doi/10.1056/CAT.20.0094.

9. Chen X, Xie H. A structural topic modeling-based bibliometric study of sentiment analysis literature. Cognit Comput. 2020;12(6):1097-129

10. Correia RB, Wood IB, Bollen J, Rocha LM. Mining social media data for biomedical signals and health-related behavior. Annu Rev Biomed Data Sci 2020;3:433-458 https://www.annualreviews.org/doi/abs/10.1146/annurev-biodatasci-030320-040844.

11. Hussain A, Cambria C. Semi-supervised learning for big social data analysis. Neurocomputing. 2018;275(6):1662-73

12. Al-Ghadir AI, Azmi AM, Hussain A. A novel approach to stance detection in social media tweets by fusing ranked lists and sentiments. Inf Fusion. 2021;67(6):29-40

13. de Figueiredo A, Simas C, Karafilakis E, Paterson P, Larson HJ. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. Lancet. 2020;396(6):898-908

14. Garmur M, King G, Mukerjee Z, Persily N, Silverman B. CrowdTangle platform and API. Harvard Dataverse 2019;3 https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/SCCQYD.

15. Edinburgh Napier University, University of Edinburgh. COVID-19 AI-powered dashboard for sentiment and opinion mining in social media platforms. CovidTracker. Last updated November 23, 2020. Accessed January 13, 2021. http://covidtracker.cloud.

16. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv 2018:1810.04805 [preprint] https://arxiv.org/abs/1810.04805.

17. Block P, Hoffman M, Raabe IJ. Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. Nat Hum Behav. 2020;4(6):588-96

18. Brandt J, Buckingham K, Buntain C. 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. 2020;3(6):167-88

19. Strassle C, Jardas E, Ochoa J. Covid-19 vaccine trials and incarcerated People — the ethics of inclusion. N Engl J Med. 2020;383(6):1897-9
20. Silcox C, Dentzer S, Bates DW. AI-enabled clinical decision support software: a trust and value checklist for clinicians. NEJM Catal Innov Care Deliv 2020;1(6) https://catalyst.nejm.org/doi/10.1056/CAT.20.0212.

21. Gellin B. Why vaccine rumours stick—and getting them unstuck. Lancet. 2020;396(6):303-4

22. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. Digit Health.