POLITECNICO DI TORINO
Repository ISTITUZIONALE

Characterizing HIV discussions and engagement on Twitter

(Article begins on next page)

18 October 2022

Characterizing HIV discussions and engagement on Twitter

Aqdas Malik[1,2] · Angi Antonino[3] · M Laeeq Khan[4] · Marko Nieminen[2]

Abstract

The novel settings provided by social media facilitate users to seek and share information on a wide array of subjects, including healthcare and wellness. Analyzing health-related opinions and discussions on these platforms complement traditional public health surveillance systems to support timely and effective interventions. This study aims to characterize the HIV-related conversations on Twitter by identifying the prevalent topics and the key events and actors involved in these discussions. Through Twitter API, we collected tweets containing the hashtag #HIV for a one-year period. After pre-processing the collected data, we conducted engagement analysis, temporal analysis, and topic modeling algorithm on the analytical sample ($n = 122,807$). Tweets by HIV/AIDS/LGBTQ activists and physicians received the highest level of engagement. An upsurge in tweet volume and engagement was observed during global and local events such as World Aids Day and HIV/AIDS awareness and testing days for trans-genders, blacks, women, and the aged population. Eight topics were identified that include "stigma", "prevention", "epidemic in the developing countries", "World Aids Day", "treatment", "events", "PrEP", and "testing". Social media discussions offer a nuanced understanding of public opinions, beliefs, and sentiments about numerous health-related issues. The current study reports various dimensions of HIV-related posts on Twitter. Based on the findings, public health agencies and pertinent entities need to proactively use Twitter and other social media by engaging the public through involving influencers. The undertaken methodological choices may be applied to further assess HIV discourse on other popular social media platforms.

Keywords HIV · AIDS · Social media · Twitter · Digital health · Public health surveillance · Engagement

1 Introduction

Human immunodeficiency virus (HIV) is a serious and challenging public health concern affecting an estimated 38 million people globally [1]. Although the infection cannot

✉ Aqdas Malik
  aqdas.a.malik@aalto.fi

  Angi Antonino
  antonino.angi@studenti.polito.it

  M Laeeq Khan
  Khanm1@ohio.edu

  Marko Nieminen
  marko.nieminen@aalto.fi

1 Department of Information Systems, Sultan Qaboos University, Muscat, Oman

2 Department of Computer Science, Aalto University, Konemiehintie 2, 02150 Espoo, Finland

3 Department of Control and Computer Engineering, Politecnico Di Torino, Torino, Italy

4 Scripps College of Communication, Ohio University, Athens, OH, USA

be completely cured, recent research has shown that vigilance by regularly monitoring personal health information (e.g., medication adherence, CDU counts, viral loads) may improve the health and wellness of people living with HIV. Likewise, preventive and cautionary measures in the form of informational resources can potentially decrease the vulnerability of acquiring HIV. Affected people, their partners, and family members utilize various sources (e.g., primary care providers, community support groups, printed material, and online channels) to seek guidance and support about HIV. Furthermore, information disseminated through some of these resources is also targeted to the general public at large.

HIV surveillance systems play a significant role in public health monitoring. They inform and support HIV prevention and intervention initiatives such as United Nations 90–90-90 and Ending the HIV epidemic: A plan for America by the U.S. Department of Health and Human Services (HHS). HIV surveillance systems have progressed exponentially over the years, yet several key challenges still prevail. The traditional data sources for monitoring HIV are often inaccurate, incomplete, unrepresentative, and partially

disconnected from each other [2, 3]. These discrepancies limit the intelligence to track, report, and respond to emerging HIV epidemiological trends periodically and effectively.

Recent advances and the prevalence of digital technologies have exponentially expanded the opportunities for sharing and seeking public health information. By aggregating and analyzing data from various sources (e.g., social media, search engines, mobile phones, location services, crowdsourcing, and wearable devices), digital surveillance complements traditional public health surveillance systems [3–5]. Novel potentials of digital data (such as real-time, organic, and volume) from location logs, web searches, and online reviews have supported in tracking, predicting, and preventing diseases [6, 7] as well as other critical public health concerns such as substance abuse [8], public attitudes towards vaccination [9], and self-injurious thoughts and behaviors [10]. Similarly, social media has recently emerged as a significant vein of digital surveillance systems. Platforms such as Twitter, Facebook, Yelp, and Instagram offer "novel settings" and have been increasingly used for communicating and understanding numerous public health and wellness issues. A wide range of information and resources facilitate social interactions between care providers and patients and among the public to support discussions about various health issues. Due to promising contributions and implications to the public health surveillance system, there is an increased interest in developing tools and techniques utilizing digital data originating outside the public health system that usually relies on the hospital, outpatient, or laboratory-based systems.

New media technologies, including online forums, computerized counseling tools, gamified interventions, specialized portals (e.g., CDC HIV Resource Library, WHO AIDS, and HIV.Gov), and social media facilitate the delivery of a wide spectrum of information on HIV, STD, and, AIDS. These digital solutions offer evidence-based HIV education, preventive measures, testing promotions, and treatment and adherence program [4, 11–13]. The latest research focusing on HIV prevention and care point out a significant rise and interest in people seeking, sharing, and engaging with information on social media [14, 15]. Within the context of social media surveillance for HIV, most of the work focuses on developing and testing interventions aimed at risk reduction and prevention [16, 17]. In understanding the vulnerable or high-risk groups, statistics reveal that the "risk of acquiring HIV is 26 times higher among gay men or other men who have sex with men (MSM), 29 times higher among people who inject drugs, 30 times higher for sex workers, and 13 times higher for transgender people" [1]. Studies spanning social media interventions tailored for vulnerable, high-risk, and underserved communities show promising potential for disseminating information and effective delivery of interventions for HIV care and prevention [18].

Given the global prevalence of social media and its emerging role in public health surveillance, the current study sought to investigate the HIV-related discussions prevailing on social media. Intending to advance the current understanding of social media conversations about HIV, our study seeks to address the following research aims: (1) determine the characteristics of the most engaging tweets; (2) comprehend the salient events that drove engagement, and (3) identify the prevalent HIV-related discussions on Twitter. Answering these research questions will enable us to recommend relevant implications and facilitate subsequent discussion regarding HIV surveillance systems.

2 Methods

The current study consisted of three phases. We employed the Twitter API to extract data during the first stage by querying hashtag #HIV during a one-year period (November 12, 2018, to November 12, 2019). In total, we managed to collect 160,658 tweets together with associated metadata (tweet full text, numbers of retweets and favorites for each tweet, number of followers and friends, and geolocation of the user). The retrieved data was consequently cleaned by removing the duplicate tweets ($n = 2948$), tweets containing URL(s), mentions, or hashtags only ($n = 10,023$), and non-English tweets ($n = 24,880$). The cleaning process resulted in an analytical sample of 122,807 tweets.

We pre-processed the analytical sample during the second stage to transform data that can be reliably used for further analysis. All the standard techniques, including stop words removal, basic normalization, and lemmatization, were applied to the tweets. Furthermore, URLs, hashtags, and non-printable characters were removed from the corpus.

In the final stage, we conducted multiple analyses to answer our research questions. An engagement analysis provided insights about the prominent users and descriptives of the top 20 tweets concerning the associated number of retweets and favorites. The temporal analysis was conducted to chart out the volume of tweets per week. The list of hashtags separated during the pre-processing stage was used to get the 20 frequently used hashtags in the dataset. Finally, to infer the salient topics within the analytical sample about HIV-related conversations on Twitter, Latent Dirichlet Allocation (LDA), an unsupervised generative modeling technique was applied [19]. LDA is a prominent topic-modeling approach for discovering clusters of documents (tweets in the current case) through a representative set of words. Within each cluster, words bearing the highest weights provided us an overview of the topic. The algorithm further requires the number of topics as an input. The extraction of the number of topics using the LDA model was based on the coherence

score of models. The topic coherence score measures a topic by calculating the degree of semantic similarity between high scoring (high probability of the occurrence) words. A higher coherence score is used to determine the optimal number of topics and indicative of a better model as the similarity in the words within each topic is higher [20]. For the model selection process, multiple LDA models were trained to find the one that has the highest coherence score and the lowest number of topics. During the model selection process, an elbow plot is used to assess and pick the model that gives the highest score before the curve flattens (see Fig. 1). Our final model specified 8 topics across the tweets dataset used as it indicates an optimal balance between high coherence score and the number of topics.

One of the authors labeled the topics by going through 10 frequently used words ordered by their weights within 20 tweets with the highest probability about the topic. The topic labels were refined after a mutual consensus was reached among all the authors. Following Twitter's privacy policies and terms of use, usernames and tweets were not reported verbatim in the current study.

3 Results

3.1 Engagement analysis

Engagement on the #HIV tweets was assessed on two dimensions i.e., favorites and retweets count.

Engagement through favorites (or likes) and retweets (or shares) are robust measures for understanding active user participation on social media [21, 22]. To determine the most engaging tweets within the analytical sample, we identified 20 tweets with the highest numbers of favorites and retweets (see Table 1). Additionally, the user type and location of the user account were also recorded manually by visiting each of the Twitter profiles in the list. Tweets of users hailing from Europe ($n = 9$) followed by users from North America ($n = 4$) received the highest number of retweets. Likewise, the highest number of favorites were recorded for the users originating from Europe ($n = 12$) and North America ($n = 4$). Concerning user type, activists received the highest number of retweets ($n = 7$) and favorites ($n = 8$), followed by physicians and personal accounts.

3.2 Temporal analysis

Conducting a temporal analysis provided us a deeper insight into the volume of #HIV tweets over a period of one year. This analysis guided us to better understand the triggers through increased activity. Within the data corpus, we were able to identify and demarcate some of the major events behind the peaks. Some of the key events where Twitter users responded actively include World Aids Day, National Black HIV/AIDS Awareness Day, National Women and Girls HIV/AIDS Awareness Day, and National HIV Testing Day (see Fig. 2).

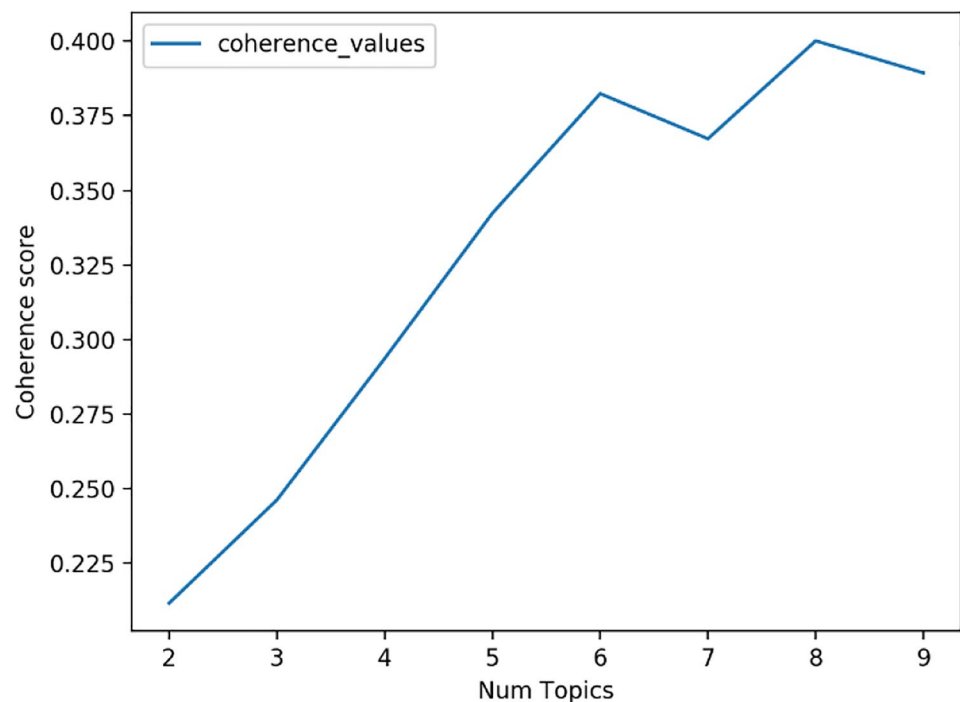**Fig. 1** Topic model graph with coherence scores and number of topics

Table 1 Top 20 tweets with highest retweets and favorites

Retweets			Favorites		
Count	**User type**	**Location**	**Count**	**User type**	**Location**
Between 2000–8000	Public Health employee	North America	Between 15,000–25,000	Public Health employee	North America
	Public Health employee	North America		HIV/AIDS activist	Europe
	HIV/AIDS activist	Europe		Public Health employee	North America
	HIV/AIDS activist	-	Between 2500–7000	Physician	Europe
Between 1000–1999	LGBTQ rights activist	Europe		LGBTQ activist	Europe
	Physician	Europe		HIV/AIDS activist	-
	Public health agency	Europe		HIV/AIDS activist	-
	Feminist activist	North America	Between 1000–2499	HIV/AIDS activist	Europe
	HIV/AIDS activist	-		Physician	North America
	Physician	Africa		Personal account	Europe
Between 500–999	LGBTQ rights activist	Europe		LGBTQ rights activist	Europe
	Public health agency	Europe		Physician	Europe
	Physician	North America		HIV/AIDS activist	Europe
	Personal account	Asia		Physician	Europe
	Personal account	-		Personal account	North America
	HIV/AIDS activist	Europe		Public health agency	Europe
	Personal account	Europe		Personal account	-
	Physician	Europe		LGBTQ activist	Europe
	Journalist	Asia		Physician	Africa
	Personal account	Africa		Personal account	Europe

3.3 Content analysis

Based on the LDA model, 8 topics about HIV discussions on Twitter were identified. In the first topic, the discussion relates to HIV-associated stigma that contains keywords people, stigma, uequalsu, treatment, and status. Topic 2 focuses on the HIV prevention and care facilities containing keywords such as "care", "prevent", "treatment", "support", "women", and "communities". The keywords such in Topic 3 that include "epidemic", "people", "countries", and "fund" capture the commentary underlining the situation and need for support in the developing countries. Tweets from three
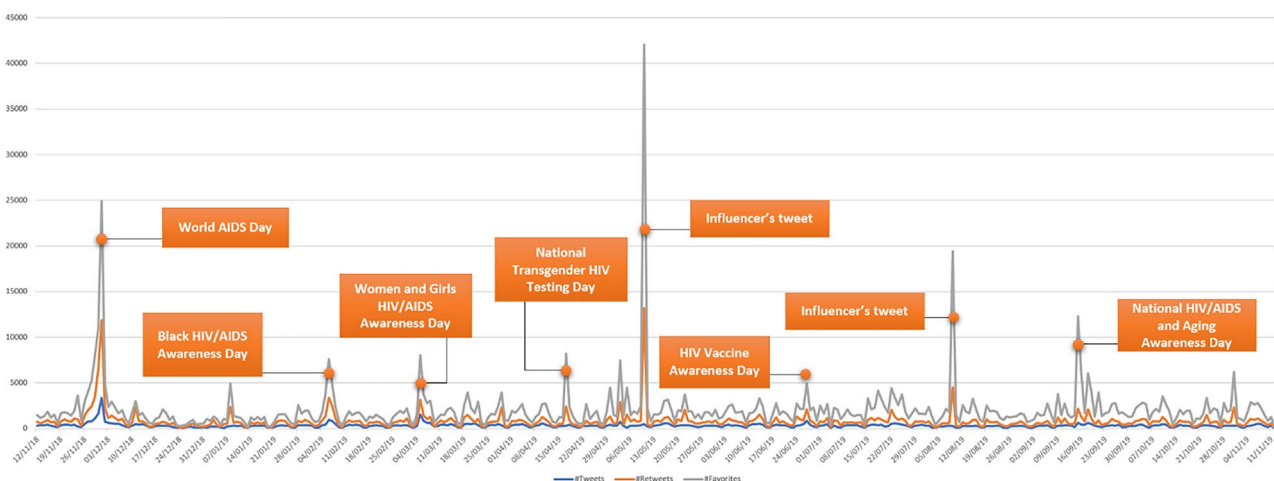

Fig. 2 Temporal Analysis of #HIV: Days view

topics draw on activities and events that include World Aids Day, HIV Testing Day, and academic conferences. Finally, two topics (5 and 7) associate collectively with tweets about HIV treatment and cure and discussions about PrEP. Table 2 displays each topic together with words identified by the topic model and a representative tweet.

Computing the frequency of hashtags within the dataset provided us further insights into the key topical themes being discussed concerning HIV. One of those topics is about diseases relevant to HIV that include #AIDS, #STI, #STD, and #TB. Another theme of tweets relates to campaigns and events including #WorldAidsDay, #UequalsU, and #KnowYourStatus. Prevention and care discussion can be observed from the #PReP and #HIVPrevention. Finally, tweets relevant to high-risk communities that are often marginalized and stigmatized are represented through #LGBTQ and #Gay.

4 Discussion

Harnessing data from new media platforms for epidemiological research is an emerging approach that complements traditional public health surveillance systems. Real-time discussions and trends from social media provide timely and accurate reflections of critical public health issues that can facilitate prevention, control, and treatment efforts. Several studies have utilized data from social media to detect and assess the spread of pandemics and infectious disease outbreaks [23, 24], delivering interventions [25], and assessing public attitudes, behaviors, and concerns towards health and wellness topics [26–29]. For instance, Fung et al., 2017 studied global health-related tweets for malaria, HIV, tuberculosis, and other tropical diseases showing the conversations by advocates, policymakers, and health professionals [30].

Table 2 Topic clusters identified through topic modeling

Topic label	Words in the topic model (in decreasing order of weight)	Representative tweet
Stigma	people, stigma, unequalsu, posit, life, treatment, status, live, medic, virus	Life of any person should never end after #HIV diagnosis, it's the stigma that really kills and forces people to hide themselves in close doors. Let's fight the ignorance and celebrate by bringing them to life #uequalsu
Prevention	care, prevent, treatment, services, support, health, access, work, women, communities	Though the #HIV situation has slightly improved across #Africa, there is a dire need for accessible care practices to carry on the momentum. We must opt for radical models of collective care especially for #women and poor communities to prevent and treat the spread of #HIV and #AIDS
Epidemic in the developing countries	end, epidemic, people, infect, case, countries, fund, report, plan, diagnosis	HIV and AIDS cases in #city is increasing rapidly as roughly 13 new cases of #HIV infections being recorded. If we plan to end the epidemic in 20 years state must increase the funding for awareness and diagnosis of #std #sti
World AIDS Day	WorldAidsDay, support, communities, stigma, work, end, fight, people, educate, aware	Its #Worldaidsday today and it's time to remember those we have lost and those people fighting their battle against #hiv & #aids. My never ending support and fight along all who volunteer for #aidsawareness to end this disease
Treatment	cure, research, patient, infect, virus, vaccine, treatment, studies, cancer, drug	Proud moments as my son joins the #HIV research team developing a therapeutic vaccine for treating HIV. I strongly believe that it will be effective and safe and likely to prevent the virus from advancing to aids. #hvad
Events	research, dr, conference, 2019, work, prevent, meet, present, join, healthcare	Youth fronts working for @username and @username calls African activities to join their cohorts for #HIV advocacy, research, education, prevention, and healthcare resource mobilization by applying at the following link
PrEP	prep, prevent, risk, infect, women, men, drug, transmission, reduce, preexposure	FDA has recently approved new #Prep drug that provides stronger protecting against #hiv while show minimal risks to kidneys and bones than the others. A daily pill can be highly effective in preventing the infection from injection drug use or sex. #hivprep #aids
Testing	test, std, today, sti, status, KnowYourStatus, prevent, treatment, PREP, condom	Love your life by getting tested and knowing your hiv/aids status. We are doing rapid #HIV tests today from 1230 to 7 pm at our lab across central station. You can also request our free home sample collection service. #msm #testing #lgbtq #gaymen #uequalsu #pride #knowyourstatus

Seminal scholarly work has also shown the utility of Twitter in a broader understanding of public opinions and sentiments to support the prevention, care, and treatment of various diseases, including diabetes and flu [31–33]. Within the context of social media surveillance for HIV, most of the work focuses on developing and testing interventions aimed at risk reduction and prevention [16, 17, 34]. In other instances, major yearly events have been studied to understand the highlighting of such global issues. A case study of HIV/AIDS-related tweets around World AIDS Day examined and found differences between user responses based on variations in income across countries [35]. In the current study, we extend the body of knowledge within the public health domain by examining the engagement factors, reasons for surges in the activity, and key topics being discussed about HIV on Twitter; a highly popular social media platform.

Based on the temporal analysis, we identified the triggers (major events and influencers) behind the amplified tweet activity (see Fig. 2). From the engagement perspective, tweets from the influencers received a higher number of retweets and favorites. In addition to World Aids Day and HIV Vaccine Awareness Days, which are celebrated to raise awareness about the HIV/Aids pandemic, many other annual campaigns dominated the tweets and engagement activity. The events detected through the activity peaks are predominantly US-based awareness campaigns targeting specific demographics such as women, blacks, transgender people, and seniors. Given the efficacy of reaching high-risk populations and stigmatized groups through social media [15, 36–38], promoting HIV, AIDS, and STD/STI-related campaigns through Twitter can raise awareness among these cohorts and deliver ensuing interventions. Concerning the engagement analysis (retweets and favorite count), the most engaging tweets were sent from user accounts based in Europe and North America. Tweets from HIV/AIDS/ LGBTQ rights campaigners and feminist activists were also highly engaging. Tweets of physicians and employees of public health agencies also received high engagement. Given the large follower base of global and local organizations, a handful of them managed to receive high engagement. This finding may suggest that Twitter users endorse and trust individuals working to safeguard the interests and rights of people affected by HIV and AIDS. Organizations involved in HIV/AIDS awareness and activism might need to team up with prominent entities such as leading activists and celebrities to project their messaging to a wider audience. Given the strong potential to reach an extensive and diverse audience through these platforms [39], further social media participation is required from entities working for HIV and AIDS awareness and treatment in developing countries, as well as media organizations.

Results from the topic modeling reveal a wide cluster of conversations about HIV on Twitter. The identified topics can be broadly categorized under five key themes i.e., events and activities, preventive measures, treatment options, stigma, and call for additional resources. Campaigns and events such as World Aids Day, HIV Testing Days, and U = U received greater projection. Among these tweets, individuals and other entities encouraged and stressed their followers to know one's status by getting tested for HIV. This finding also links to the temporal analysis where testing and awareness days for at-risk and marginalized communities such as transgender people, women, and seniors are observed. Treatment options are another significant theme of the discussion that focuses on antiretroviral therapy (ART), care locations, and the latest advances in scientific research. Even though no effective cure is available for HIV, diagnosis and getting proper medical care can slow down the progression of HIV in the human body.

Another important theme relates to tweets reflecting various modes of preventive measures. This stream of tweets reflects preventions that may slow down the spread of HIV. Many tweets point out different sites for testing and clinical care within the prevention and care facilities topic. Several tweets also discuss individual counseling, screenings, and access to biomedical interventions such as PrEP. Furthermore, several references point out self-management tools, games/gamified solutions, and smartphone apps for the prevention and treatment of HIV. These findings echo with prior research highlighting the potential of new media technologies for disseminating innovative interventions and building virtual communities, particularly among key communities such as sex workers, prisoners, transgender people, and PWID [5, 40, 41]. Similarly, social media platforms have been deemed effective for promoting general awareness, behavioral modification, and HIV testing [34, 42, 43].

PrEP is another relevant theme that deliberates various medications and subsequent concerns related to their costs and side effects in conjunction with preventive measures. Although PrEP is a relatively new mode of HIV prevention, it is regarded as one of the most significant breakthroughs for HIV prevention. Recent studies suggest that in addition to associated stigma, accessibility, and social-cultural barriers, high costs and concerns related to side effects are some of the key barriers responsible for the slow uptake of PrEP [44, 45]. Even though some of the concerns are being discussed on Twitter, there is a strong need to use social media more effectively to counter these concerns and conspiracy-related beliefs regarding PrEP. Several studies indicate that social media can facilitate PrEP uptake and care that caters to treatment initiation, monitoring, adherence, suppressions, and linkage to care facilities [5, 46, 47]. Likewise, social media interventions targeted towards the population at high risk for HIV acquisition (e.g., Transgender people, MSN, and youth) can also be an effective way to reach them.

Finally, stigma is another topic of imperative significance observed in the dataset. Despite numerous efforts ranging from global to community scale, there is a wide prevalence of HIV/AIDS stigma bundled with misinformation. Stigma not only imposes hardships on individuals living with HIV but also on their loved ones and communities. Social media platforms such as Twitter provides social (e.g., networking, communal events), emotional (e.g., seeking empathy and voicing emotional states), and informational support (e.g., requesting and sharing information and resources) that can facilitate coping with the stigma associated with HIV. Furthermore, these platforms can effectively reach and serve people at higher risk of HIV, such as bisexuals, gays, and other vulnerable groups in various world regions [1, 15, 36–38]. There is also a need to understand platform-related policies, legal frameworks, and societal norms that may lead to varying levels of discussions on certain issues such as HIV. For example, Youku, Weibo, and WeChat in China, and VK in Russia are working under different content sharing models. HIV-positive people use social media to create social ties and build a sense of community, such as accessing and sharing health information and obtaining emotional support with others alike [36, 48]. Likewise, these platforms can also be helpful in the update of treatment adherence and accessing prevention and testing services. At the same time, many tweets within the stigma topic emphasize the de-stigmatization of PLWHIV and misinformation associated with the causes and spread of the virus. One such example revealed during the analysis is the U = U campaign aiming to dismantle HIV stigma by encouraging testing and treatment [1]. Given that individuals with stably suppressed viral load cannot transmit the virus to others, social media may further reinforce the existing consensus and refute associated misinformation.

4.1 Limitations and future directions

The current study solely focuses on Twitter, a platform prevalent among adults. Future research on HIV and AIDS should further investigate the sharing behaviors, engagement patterns, and conversations on notable social media popular among youth and senior citizens such as Instagram, YouTube, Pinterest, and Facebook. Analysis of tweets in only one language could also serve as a limitation. As the current study analyzed English-only tweets, there could likely be other engaging tweets from users based in non-English-speaking countries. The addition of non-English tweets such as Chinese, Russian, Spanish, Arabic, and Urdu may result in a nuanced understanding of salient topics and issues. Likewise, comparative studies delving into social media platforms popular among communities and regions may also provide nuanced insights. Big data analytics may not capture nuanced insights into the communication style of users

originating from different geographical regions or personal backgrounds. Considering this limitation of the current study, qualitative studies employing social media data may support extending our current understanding of why specific groups communicate about HIV in a particular way. Given that social media, notably, Twitter is being used actively and deliberately to spread fake news and misinformation (e.g., vaccine safety and efficacy), researchers should investigate the extent of HIV. There is a further need to understand how effectively hospitals, health agencies, and HIV/AIDS associations are leveraging the power of social media to dispel fear and myths associated with HIV in general, transmission routes, antiretroviral treatments such as highly active antiretroviral therapy or pre-exposure prophylaxis (PrEP).

4.2 Conclusions

The current study utilized publicly available Twitter data on HIV by applying a data-driven approach to understand the conversations and key entities and events driving the engagement. Findings from the current study provide actionable insights that may further support enhancing awareness and prevention messaging, developing testing services, bolstering PrEP uptake, and optimizing clinical and community interventions. The wide adoption of social media and its use for health purposes as a participatory and interactive source hold significant implications for HIV education, care, and prevention efforts. Social media holds great potential in digital surveillance and interventions related to HIV as it is rather easier to reach diverse, underserved, ethnic, and high-risk groups. Effective use of social media by offering relevant, timely, and factual information and interventions can be highly cost-effective and can change individual behaviors and social norms about HIV.

Declarations

References

1. Global HIV & AIDS statistics — 2020 fact sheet. https://www.unaids.org/en/resources/fact-sheet. 2020. Accessed 23 Nov 2020.
2. Ayers JW, Althouse BM, Dredze M. Could behavioral medicine lead the web data revolution? JAMA. 2014;311:1399–400.
3. Paul MJ, Dredze M. Social monitoring for public health. Synth Lect Inf Concepts Retr Serv. 2017;9:1–183.
4. Nwagwu WE. Usability of mobile phones for personal health care by people living with HIV/AIDS. Health Technol. 2021;11:491–504. https://doi.org/10.1007/s12553-020-00490-x.
5. Strathdee SA, Nobles AL, Ayers JW. Harnessing digital data and data science to achieve 90–90–90 goals to end the HIV epidemic. Curr Opin HIV AIDS. 2019;14:481–5.
6. Bansal S, Chowell G, Simonsen L, Vespignani A, Viboud C. Big data for infectious disease surveillance and modeling. J Infect Dis. 2016;214:S375–9.
7. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. N Engl J Med. 2009;360:2153.
8. Kalyanam J, Katsuki T, Lanckriet GR, Mackey TK. Exploring trends of nonmedical use of prescription drugs and polydrug abuse in the Twittersphere using unsupervised machine learning. Addict Behav. 2017;65:289–95.
9. Porreca A, Scozzari F, Di Nicola M. Using text mining and sentiment analysis to analyse YouTube Italian videos concerning vaccination. BMC Public Health. 2020;20:1–9.
10. Kleiman EM, Nock MK. Real-time assessment of suicidal thoughts and behaviors. Curr Opin Psychol. 2018;22:33–7.
11. Chiasson MA, Hirshfield S, Rietmeijer C. HIV prevention and care in the digital age. JAIDS J Acquir Immune Defic Syndr. 2010;55:S94–7.
12. Daher J, Vijh R, Linthwaite B, Dave S, Kim J, Dheda K, Peter T, Pai NP. Do digital innovations for HIV and sexually transmitted infections work? Results from a systematic review (1996-2017). BMJ open. 2017;7(11):e017604.
13. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, Gilliam M. Interventions using new digital media to improve adolescent sexual health: a systematic review. J Adolesc Health. 2012;51:535–43.
14. Maloney KM, Bratcher A, Wilkerson R, Sullivan PS. Electronic and other new media technology interventions for HIV care and prevention: a systematic review. J Int AIDS Soc 2020;23:e25439.
15. Taggart T, Grewe ME, Conserve DF, Gliwa C, Isler MR. Social media and HIV: a systematic review of uses of social media in HIV communication. J Med Internet Res 2015;17:e248.
16. Lelutiu-Weinberger C, Pachankis JE, Gamarel KE, Surace A, Golub SA, Parsons JT. Feasibility, acceptability, and preliminary efficacy of a live-chat social media intervention to reduce HIV risk among young men who have sex with men. AIDS Behav. 2015;19:1214–27.
17. Tso LS, Tang W, Li H, Yan HY, Tucker JD. Social media interventions to prevent HIV: a review of interventions and methodological considerations. Curr Opin Psychol. 2016;9:6–10.
18. Yun K, Xu J, Leuba S, Zhu Y, Zhang J, Chu Z, Geng W, Jiang Y, Shang H. Development and Validation of a Personalized Social Media Platform–Based HIV Incidence Risk Assessment Tool for Men Who Have Sex With Men in China. J Med Internet Res 2019;21:e13475.
19. Blei D, Ng A, Jordan M. Latent Dirichlet allocation. J Mach Learn Res. 2003;3:993–1022.
20. Mimno D, Wallach H, Talley E, Leenders M, McCallum A. Optimizing semantic coherence in topic models. In: Proceedings of the 2011 conference on empirical methods in natural language processing. UK. Association of Computational Linguistics, Edinburgh, Scotland. 2011;262–72.
21. Khan ML. Social media engagement: What motivates user participation and consumption on YouTube? Comput Hum Behav. 2017;66:236–47.
22. Khan ML, Ittefaq M, Pantoja YIM, Raziq MM, Malik A. Public Engagement Model to Analyze Digital Diplomacy on Twitter: A Social Media Analytics Framework. Int J Commun. 2021;15:29.
23. Aiello AE, Renson A, Zivich PN. Social Media–and Internet-Based Disease Surveillance for Public Health. Annu Rev Public Health. 2020;41:101–18.
24. Daughton AR, Paul MJ. Identifying protective health behaviors on Twitter: observational study of travel advisories and Zika virus. J Med Internet Res. 2019;21(5):e13090.
25. Pretorius K, Johnson KE, Rew L. An integrative review: understanding parental use of social media to influence infant and child health. Matern Child Health J. 2019;23:1360–70.
26. Ketonen V, Malik A. Characterizing vaping posts on Instagram by using unsupervised machine learning. Int J Med Inf 2020;141:104223.
27. Malik A, Khan ML, Quan-Haase A. Public Health Agencies Outreach through Instagram during COVID-19 Pandemic: Crisis and Emergency Risk Communication Perspective. Int J Disaster Risk Reduct 2021;102346.
28. Malik A, Li Y, Karbasian H, Hamari J, Johri A. Live, love, Juul: User and content analysis of Twitter posts about Juul. Am J Health Behav. 2019;43:326–36.
29. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. Emerg Themes Epidemiol. 2014;11:1–7.
30. Fung IC-H, Jackson AM, Ahweyevu JO, Grizzle JH, Yin J, Tse ZTH, Liang H, Sekandi JN, Fu K-W. # Globalhealth Twitter conversations on# Malaria,# HIV,# TB,# NCDS, and# NTDS: a cross-sectional analysis. Ann Glob Health. 2017;83:682–90.
31. Karami A, Dahl AA, Turner-McGrievy G, Kharrazi H, Shaw G Jr. Characterizing diabetes, diet, exercise, and obesity comments on Twitter. Int J Inf Manag. 2018;38:1–6.
32. Molaei S, Khansari M, Veisi H, Salehi M. Predicting the spread of influenza epidemics by analyzing twitter messages. Health Technol. 2019;9:517–32. https://doi.org/10.1007/s12553-019-00309-4.
33. Stefanidis A, Vraga E, Lamprianidis G, Radzikowski J, Delamater PL, Jacobsen KH, Pfoser D, Croitoru A, Crooks A. Zika in Twitter: temporal variations of locations, actors, and concepts. JMIR Public Health Surveill 2017;3:e22.
34. Jones J, Carter B, Wilkerson R, Kramer C. Attitudes toward HIV testing, awareness of HIV campaigns, and using social networking sites to deliver HIV testing messages in the age of social media: a qualitative study of young black men. Health Educ Res. 2019;34:15–26.
35. Fung IC-H, Yin J, Pressley KD, Duke CH, Mo C, Liang H, Fu K-W, Tse ZTH, Hou S-I. Pedagogical Demonstration of Twitter Data Analysis: A Case Study of World AIDS Day, 2014. Data. 2019;4:84.
36. Chen L, Guo Y, Shi J. Social Support Seeking on Social Media Among Chinese Gay Men Living with HIV/AIDS: The Role of Perceived Threat. Telemed E-Health. 2018;25:655–9. https://doi.org/10.1089/tmj.2018.0136.

37. Dong Y, Zhou X, Lin Y, Pan Q, Wang Y. HIV-related posts from a Chinese internet discussion forum: An exploratory study. PLoS One. 2019 Feb 28;14(2):e0213066. https://doi.org/10.1371/journal.pone.0213066.

38. Nobles AL, Leas EC, Noar S, Dredze M, Latkin CA, Strathdee SA, Ayers JW. Automated image analysis of instagram posts: Implications for risk perception and communication in public health using a case study of# HIV. PLoS One 2020;15:e0231155.

39. Johri A, Karbasian H, Malik A, Handa R, Purohit H. How diverse users and activities trigger connective action via social media: Lessons from the twitter hashtag campaign# ilooklikeanengineer. In: Proceedings of the 51st Hawaii International Conference on System Sciences 2018.

40. Cao B, Gupta S, Wang J, Hightow-Weidman LB, Muessig KE, Tang W, Pan S, Pendse R, Tucker JD. Social Media Interventions to Promote HIV Testing, Linkage, Adherence, and Retention: Systematic Review and Meta-Analysis. J Med Internet Res. 2017;19:e394. https://doi.org/10.2196/jmir.7997.

41. Merchant RC, Clark MA, Liu T, Rosenberger JG, Romanoff J, Bauermeister J, Mayer KH. Preferences for oral fluid rapid HIV self-testing among social media-using young black, Hispanic, and white men-who-have-sex-with-men (YMSM): implications for future interventions. Public Health. 2017;145:7–19.

42. Huang E, Marlin RW, Young SD, Medline A, Klausner JD. Using Grindr, a smartphone social-networking application, to increase HIV self-testing among Black and Latino men who have sex with men in Los Angeles, 2014. AIDS Educ Prev. 2016;28:341–50.

43. Iribarren SJ, Ghazzawi A, Sheinfil AZ, Frasca T, Brown W, Lopez-Rios J, Rael CT, Balán IC, Crespo R, Dolezal C. Mixed-method evaluation of social media-based tools and traditional strategies to recruit high-risk and hard-to-reach populations into an HIV prevention intervention study. AIDS Behav. 2018;22:347–57.

44. Mayer KH, Agwu A, Malebranche D. Barriers to the Wider Use of Pre-exposure Prophylaxis in the United States: A Narrative Review. Adv Ther. 2020;37:1778–811.

45. Pacífico de Carvalho N, Mendicino CCP, Cândido RCF, Alecrim DJD, Menezes de Pádua CA. HIV pre-exposure prophylaxis (PrEP) awareness and acceptability among trans women: a review. AIDS Care. 2019;31:1234–40.

46. Macapagal K, Kraus A, Korpak AK, Jozsa K, Moskowitz DA. PrEP awareness, uptake, barriers, and correlates among adolescents assigned male at birth who have sex with males in the US. Arch Sex Behav. 2020;49:113–24.

47. Patel VV, Ginsburg Z, Golub SA, Horvath KJ, Rios N, Mayer KH, Kim RS, Arnsten JH. Empowering with PrEP (E-PrEP), a peer-led social media–based intervention to facilitate HIV preexposure prophylaxis adoption among young Black and Latinx gay and bisexual men: protocol for a cluster randomized controlled trial. JMIR Res Protoc 2018;7:e11375.

48. Cifor M, McKinney C. Reclaiming HIV/AIDS in digital media studies. First Monday. 2020.