Estimating the influence of Twitter on pre-exposure prophylaxis use and HIV testing as a function of rates of men who have sex with men in the United States

Man-pui Sally Chan[a], Alex Morales[a], Maria Zlotorzynska[b], Patrick Sullivan[b], Travis Sanchez[b], Chengxiang Zhai[a] and Dolores Albarracín[a]

Objective(s): Acceptance of pre-exposure prophylaxis (PrEP) and testing for HIV is likely to vary as a function of the norms and communications within a geographic area. This study examined associations involving county tweets, in person communications, and HIV prevention and testing in regions with higher (vs. lower) estimated rates of men who have sex with men (MSM).

Design and Methods: Ecological analyses examined (a) tweets about HIV (i.e. tweet rates per 100 000 county population and topic probabilities in 1959 US counties); (b) individual-level survey data about HIV prevention and testing and communications about PrEP and HIV ($N = 30\,675$ participants); and (c) estimated county-level MSM rates (per 1 000 adult men).

Results: In counties with higher rates of MSM, tweet rates were directly associated with PrEP use and HIV testing ($rs = .06$, $BF_{10} > 10$). Topics correlated with PrEP use ($rs = -0.06$ to 0.07, $BF_{10} > 10$) and HIV testing ($rs = -0.05$ to 0.05, $BF_{10} > 10$). Mediation analyses showed that hearing about and discussing PrEP mediated the relations between tweet rates and PrEP use ($b_{i^*} = 0.01$–0.05, $BF_{10} > 100$) and between topics and PrEP use ($b_{i^*} = -0.04$– 0.05, $BF_{10} > 10$). Moreover, hearing about PrEP was associated with PrEP use, which was in turn associated with tweet rates ($b_{i^*} = 0.01$, $BF_{10} > 100$) and topics ($b_{i^*} = -0.03$ – 0.01, $BF_{10} > 10$).

Conclusions: Rates of MSM appear to lead to HIV tweets in a region, in person communications about PrEP, and, ultimately, actual PrEP use. Also, as more men hear about PrEP, they may use PrEP more and may tweet about HIV.

AIDS 2021, **35** (Suppl 1):S101–S109

Keywords: communication, HIV testing, HIV, men who have sex with men, pre-exposure prophylaxis, PrEP use, social media

Tweeting about HIV: estimating possible influences on PrEP use and HIV testing as a function of MSM rates in the United States

Studying the impact of social media messages about HIV can provide valuable insights for policies and health communications at a time when eradicating HIV is possible. However, there are currently no estimates of the size of the influence of social media messages on behavior like pre-exposure prophylaxis (PrEP) use or the processes leading to this influence. In this paper, we ask whether social media messages within a region correlate with HIV prevention and testing, whether social media messages also correlate with other communication about PrEP and

[a]University of Illinois at Urbana-Champaign, Champaign, IL, and [b]Emory University, Atlanta, GA, USA.

Correspondence to Man-pui Sally Chan, PhD, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, USA.

E-mail: sallycmp@illinois.edu

Received: 29 May 2020; revised: 16 December 2020; accepted: 7 January 2021.

DOI:10.1097/QAD.0000000000002838

ISSN 0269-9370

HIV, and whether social media messages and communications correlate with PrEP use and HIV testing. We also ask whether these associations to be present in geographic areas with a larger men who have sex with men (MSM) community.

What are the health behavior consequences of living in a region with higher concentrations of people 'like you,' such as more sexual minorities when you are one? Prior studies have provided insights into how the presence of other sexual minorities affects psychological well-being and health behaviors. For example, sexual minorities in more (vs. less) inclusive contexts report less fear, are more willing to 'come out' (i.e. disclose their sexual identity and preferences) [1], are more likely to utilize health services [2], and have lower rates of risky sexual behaviors [3]. Likewise, communications about PrEP or HIV, both in person and on social media, may differ between communities with more MSM and communities with fewer MSM [4,5].

Imagine African-American men who date other men in Manhattan, San Francisco, Pittsburgh, and Oklahoma City. There are stark contrasts in the number of MSM, service use, and social media activity across these cities. In terms of numbers of inhabitants, 87 556 and 66 586 MSM live in Manhattan and San Francisco, whereas only 26 666 and 14 028 MSM live in Pittsburgh and Oklahoma City. In terms of service use, HIV-relevant services are also higher in Manhattan and San Francisco than in Pittsburgh and Oklahoma City [6]. For example, New York State has approximately seven times the rate of PrEP use (i.e. 103 PrEP users per 100 000 population) as Oklahoma (i.e. about 15 PrEP users per 100 000 population) [6]. Furthermore, these cities also differ in social media use [7]. Manhattan ranks 3 in terms of Twitter use, whereas Oklahoma City is much further down the line [8].

The rates of MSM in a community may be important for social media messages to exert an influence on behavior. More MSM in a region may need more tweets and in-person communications about topics of interest, including PrEP. Also, more MSM in a region may lead to greater openness in messaging and/or talking about HIV issues, which may influence HIV prevention and testing. In contrast, regions with less MSM may not reach the necessary amount of social media messages and/or other communications about HIV to ultimately promote HIV prevention and testing. This paper considered whether regions with higher MSM rates promote more social media messages about HIV or specific topics of social media conversation, and whether these conversations are associated with other relevant communications and with PrEP use and HIV testing.

The United States setting is ideal to understand the real-life influences of social media for two reasons. First, the United States has the highest social media penetration rate in the world, with Twitter specifically having over 65 million active users and about 17 million posts a day. Second, the political will to halt HIV has facilitated expansion of HIV testing and promotion of PrEP use [9]. The present study examined whether tweets about HIV that posted between 2014 and 2016 and geotagged to counties correlate with reported communication about PrEP and HIV (i.e. hearing about PrEP, discussing PrEP use, and discussing HIV testing), and HIV prevention and testing (i.e. PrEP use and HIV testing) in United States counties with higher (vs. lower) rates of MSM. We studied communications and HIV prevention and testing from the American Men Internet Survey (AMIS 2014–2016). We proposed that Twitter messages about HIV may correlate with PrEP use and HIV testing, either directly or via mediating associations with reported communication about PrEP and HIV (see Fig. 1) and that these relations may differ as a function of county MSM rates. We first examined a proposed model in which communications mediate the relations between tweets (see top panel of Fig. 1) and HIV prevention and testing. We also examined a model in which hearing about and discussing HIV and PrEP facilitate HIV prevention and testing behaviors, and these behaviors then yield different tweets (see bottom panel of Fig. 1). We considered Bayes factors (BF_{10}) and posterior probabilities to gauge support for different directional models. Furthermore, we estimated if a mediator explained all of the influence of the external variable on the dependent variable (i.e. full mediation) or just part (i.e. partial mediation).

The analysis of tweets was two-fold. We first obtained a rate of HIV tweets per 100 000 county population. This rate represents the level of tweets in each county when taking the county population into account. We then conducted topic modeling that assesses the co-occurrences of words in the tweet corpus to identify a mixture of themes for each county. Thus, the models in Fig. 1 were tested with both HIV tweet rates and topics.

Methods

The present study combined (a) individual-level data obtained from AMIS, which is an annual, cross-sectional, online HIV behavioral survey of MSM in the United States, (b) county-level tweets about HIV, and (c) estimated county-level MSM rates (per 1000 adult men) [10]. Details of the estimated county-level MSM rates, Twitter data, and survey data are described in the following sections and supplemental information (SI, http://links.lww.com/QAD/C22).

Men who have sex with men rates

We used the estimates reported by Grey *et al.* [10] and the total number of adult men to calculate the rates of MSM in each county:

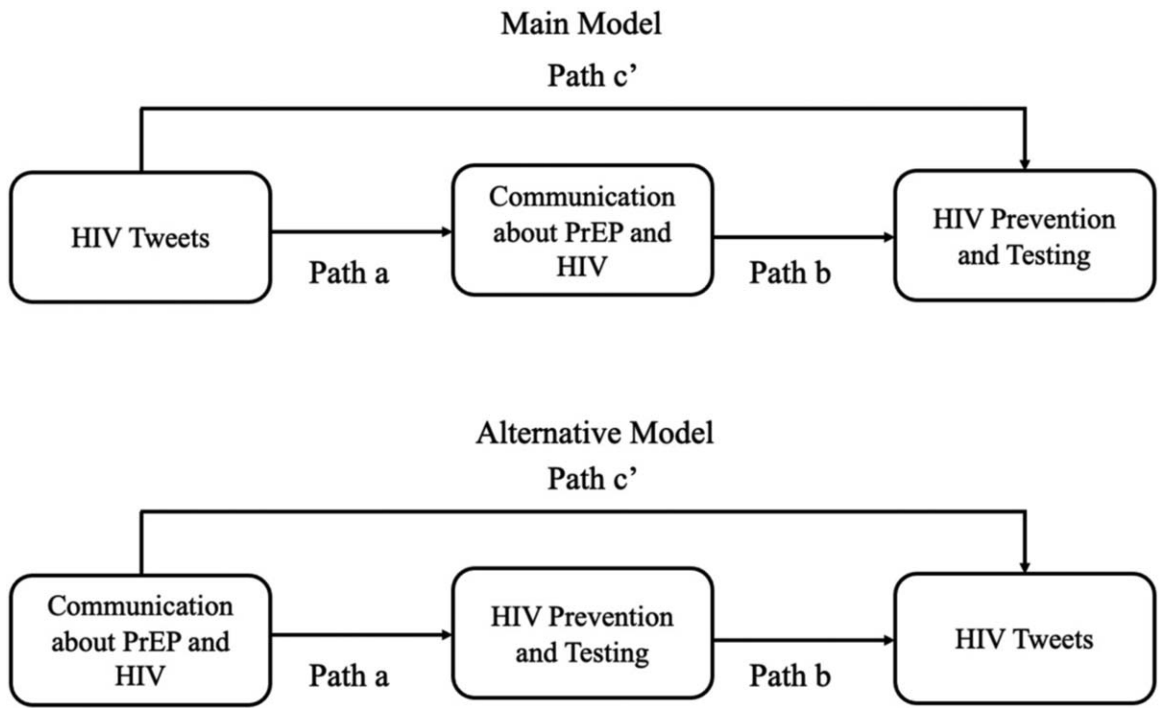

Fig. 1. Two conceptual models showing the associations from tweets about HIV (i.e. tweet rates and Twitter topics) to communications to behaviors (top) and the associations from communications to behaviors to tweets about HIV (bottom).

$$\text{Rates of MSM} = \frac{\text{MSM Estimates}}{\left(\frac{\text{Total Adul Men}}{1000}\right)}$$

We then used the median of the MSM rates (i.e. 40.87) to split the dataset into two subgroups: (a) counties with higher MSM rates and (b) counties with lower MSM rates (see Supplemental Information for details).

Twitter data

HIV-related tweets

We used the Twitter Streaming Application Programming Interface to collect 1% of random tweets from 2014 to 2016 and geotagged 43 930 487 tweets that were posted in the United States [11]. We applied an HIV support vector model classifier to the 2014–2016 tweet corpus led to the inclusion of 651 061 tweets about HIV (see Supplemental Information for details). We then calculated tweet rates per 100 000 county population and probabilities of HIV topics for those 651 061 tweets.

Tweet rates

To measure the volume of tweets about HIV per county each year, we counted the number of HIV tweets from 2014 to 2016 and obtained the population estimates from the American Community Survey to calculate a tweet rate per 100 000 population for each county.

Topic modeling

We used the Python package *scikit-learn* to convert a collection of documents (i.e. tweets in each county) into a frequency matrix of token counts. The matrix of token counts was then analyzed using a well-established algorithm in computer science, namely, latent Dirichlet allocation (LDA) [12]. LDA is a Bayesian mixture model, P (word | topic) and P (topic | document), that groups words that often appear together to create topics (see Fig. 2). We experimented different numbers of topics (i.e. *k*) and a model with 100 topics had the highest log perplexity score. Hence, we used this model for to calculate topic probabilities for each county.

Twitter topics indicate general social attitudes within a community rather than personal plans to engage in HIV testing and prevention (see word clouds of the selected topics in Supplemental Figure 1, http://links.lww.com/QAD/C22 and all topics in Supplemental Table 1, http://links.lww.com/QAD/C22). For example, some of the words identified in the topics refer to formal and informal communications, like 'deals' and 'coverage' (topic 69) or 'advice' and 'reply' (topic 95). Other topics refer to information and information sources, like 'media', 'nytimes', 'talk' (topic 76), and real-world events that raise awareness of HIV infections, such as 'Worldaidsday' (topic 95). Yet, other topics refer to love and celebrations, like 'romance' and 'raw' (topic 66), or 'birthday' and 'celebrating' (topic 45). More generally, topics refer to issues but have no literal interpretation (for a discussion of this issue) [11].

Survey data

Design and methods

We used the 2014–2016 AMIS data. Respondents were recruited online from various websites through banner

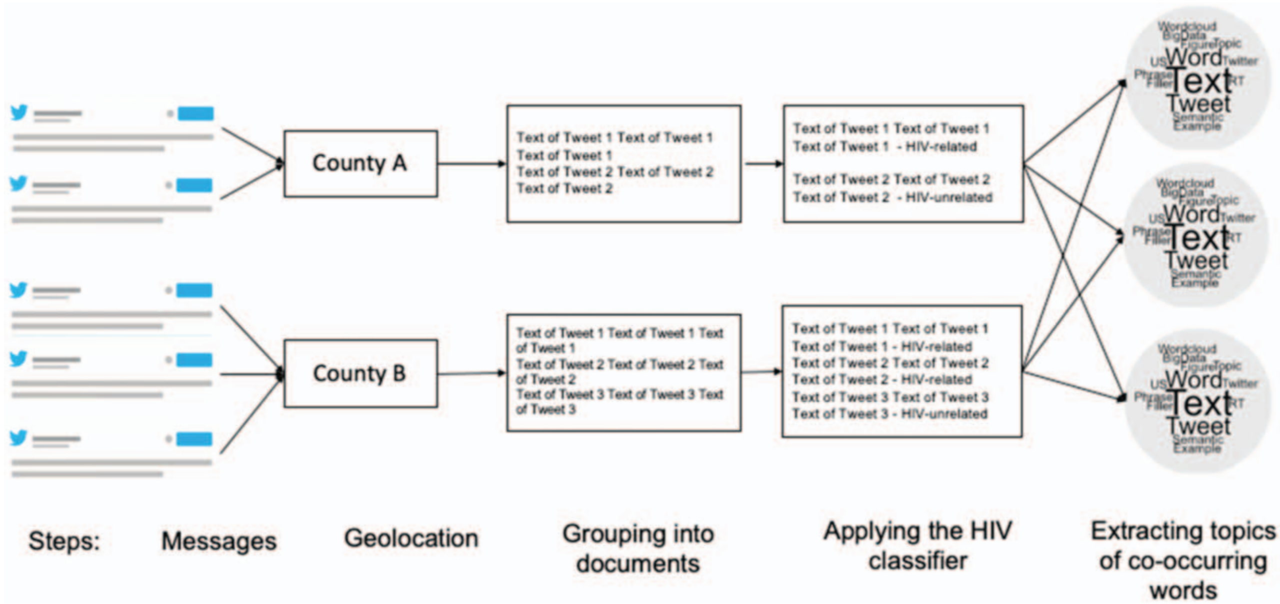

Fig. 2. Procedures of processing tweets.

advertisements and e-mail blasts to complete a self-administered online survey, which is suitable for reporting private or sensitive health behaviors [13,14]. Men were eligible to participate in the survey if they were at least 15 years old, resided in the United States, provided a valid United States zip code and reported ever having sex with a man (see Supplemental Information for details).

Measures of pre-exposure prophylaxis and HIV testing

We used answers (i.e., 'Yes' or 'No') to two questions to indicate respondents' PrEP use status and HIV testing behavior in the past 12 months (see Supplemental Information for details).

Measures of hearing about pre-exposure prophylaxis, discussing pre-exposure prophylaxis use, and discussing HIV testing

We used answers (i.e. 'Yes' or 'No') to three questions to indicate whether respondents heard about PrEP, discussed PrEP, and discussed HIV testing (see Supplemental Information for details).

Statistical analyses

We first calculated intra-class correlations (ICCs) of survey measures (see Supplemental Information for details) and then conducted Bayesian statistical analyses to estimate the Bayesian factors and posterior probabilities of our hypotheses in relation to the data [15]. The key advantage of Bayesian analyses is a direct test of both the null and alternative hypotheses to measure the strength of evidence [16]. Details of each Bayesian analysis appear next.

Bayesian correlation

We used Jeffreys's Amazing Statistics Program [17] with the default uniform prior (i.e. 1) in which the possible correlation between -1 and $+1$ is equally plausible [18]. We examined the relations of the tweet rates and topic probabilities with (a) HIV prevention and testing (i.e. PrEP use and HIV testing) and (b) reported communication about PrEP and HIV (i.e. hear about PrEP, discuss PrEP, and discuss HIV). When the correlations suggested both an association between tweets and reported communications and an association between reported communication and HIV prevention and testing (i.e. $BF_{10} > 10$), they met the condition for mediation analyses, and we proceeded to analyze mediation.

Bayesian mediation analyses

We next used the *BayesMed* package in R to conduct Bayesian mediational models [19]. The analyses examined the mediation effect (i.e. path ab) and the direct effect (i.e. path c'). The presence of evidence of path ab ($BF_{10} > 10$) but not path c' ($BF_{10} < 10$) indicates a *full* mediation, whereas the presence of evidence of both path ab and path c' suggests a *partial* mediation. After examining the first proposed model (see top panel of Fig. 1), we assessed the alternative model containing paths from reported communication to HIV prevention and testing to either tweet rates or Twitter topics (see bottom panel of Fig. 1).

Results

ICCs determine if the data vary primarily across individuals or across counties. All ICCs showed low levels of county variability (ICCs ranged between 0.03 and 0.14), suggesting that the associations with tweets could be assessed at the individual level. Therefore, we used single-level Bayesian analyses in the following analyses. The present study first examined the correlations between tweets about HIV and the AMIS survey data (N in 2014–2016 = 30 675) and assessed the data for counties with higher and lower MSM rates.

Associations of HIV tweets and reported communication about pre-exposure prophylaxis and HIV, and HIV prevention and testing across counties with higher and lower rates of men who have sex with men

results for higher rates of men who have sex with men

Table 1 presents the correlations between tweet rates, topic probabilities, PrEP use, HIV testing, hearing about PrEP, discussing PrEP use, and discussing HIV testing in counties with higher rates of MSM. First, hearing about PrEP and discussing PrEP use were positively correlated with actual PrEP use ($rs = 0.17-0.61$, $BF_{10} > 100$) and discussing HIV testing was positively correlated with actual testing ($rs = .16$, $BF_{10} > 100$).

Second, the rate of HIV tweets per 100 000 county population was positively correlated with all survey measures ($rs = 0.06-0.08$, $BF_{10} > 100$), except discussing HIV testing. These results thus support associations of the tweets with reported communications and HIV prevention and testing but cannot determine the direction of the relations.

Table 1. Correlational results among HIV tweets (rates and topic probabilities) and AMIS variables in counties with higher rates of MSM.

	HIV prevention and testing		Reported communication about PrEP and HIV			
Variable	PrEP use	HIV testing	Hear about PrEP	Discuss PrEP use	Discuss HIV testing	Any indication of mediating associations (yes/no) (If yes, number of mediation)
Survey measures						
HIV testing	0.22***	–				
Hear about PrEP	0.17***	0.16***	–			
Discuss PrEP use	0.61[a]	0.36***	NA	–		
Discuss HIV testing	0.09***	0.16***	0.02	0.21***	–	
Measures of tweets						
Tweet rates	0.06***	0.06***	0.08***	0.08***	0.02	Yes (2)
Twitter topics						
Topic 0	0.04***	0.02	0.02	0.05*	−0.01	Yes (1)
Topic 1	−0.04***	−0.02	−0.02	−0.05**	0	Yes (1)
Topic 8	−0.04***	−0.02	−0.02	−0.06***	0.01	Yes (1)
Topic 12	−0.04**	−0.03	−0.04***	−0.04	−0.01	Yes (1)
Topic 17	0.04***	0.02	0.02	0.05**	0.01	Yes (1)
Topic 19	0.04***	0.03	0.03	0.04	0	No
Topic 25	0.02	0.01	0.03*	0.03	−0.01	No
Topic 32	−0.03	−0.03	−0.04***	−0.03	0	No
Topic 35	−0.04***	−0.04***	−0.03	−0.06***	−0.03	Yes (1)
Topic 42	−0.04***	−0.03*	−0.02	−0.05*	0	Yes (1)
Topic 43	0.03	0.03*	0.02	0.03	−0.01	No
Topic 45	−0.04***	−0.03*	−0.03*	−0.04	0	Yes (1)
Topic 48	0.04***	0.03	0.03	0.04	0	No
Topic 49	0.05***	0.02	0.04***	0.05*	−0.01	Yes (2)
Topic 50	0.04***	0.02	0.02	0.03	−0.01	No
Topic 54	−0.04***	−0.02	−0.02	−0.05*	−0.02	Yes (1)
Topic 57	−0.04***	−0.01	−0.03	−0.03	0.01	No
Topic 62	−0.04***	−0.03	−0.04**	−0.05**	0	Yes (2)
Topic 66	−0.05***	−0.03	−0.03	−0.06***	−0.02	Yes (1)
Topic 69	0.07***	0.05***	0.06***	0.08***	−0.01	Yes (2)
Topic 74	0.04***	0.01	0.03	0.04	0.01	No
Topic 76	0.05***	0.03	0.04**	0.04	−0.01	Yes (1)
Topic 78	0.04**	0.01	0.03	0.04	0	No
Topic 83	−0.06***	−0.03**	−0.05***	−0.06***	0.01	Yes (2)
Topic 84	−0.03	−0.01	−0.03	−0.04	0.01	No
Topic 87	0.01	−0.02	0.01	−0.03	−0.04**	No
Topic 88	0.04***	0.02	0.02	0.04	−0.01	No
Topic 91	0.03	0.03*	0.04***	0.04	0.01	No
Topic 92	−0.04***	−0.02	−0.04***	−0.04	0	Yes (1)
Topic 95	0.04***	0.03	0.03	0.04	−0.01	No
Topic 96	0.04**	0.04***	0.04**	0.04	0	Yes (1)
Topic 98	−0.04***	−0.05***	−0.02	−0.03	−0.01	No
Topic 99	0.04***	0.02	0.02	0.04	−0.01	No

*$BF_{10} > 10$.
**$BF_{10} > 30$.
***$BF_{10} > 100$, [8]BF_{10} infinite. AMIS, American Men Internet Survey; MSM, men who have sex with men; PrEP, pre-exposure prophylaxis.

As for topics, over 50% of the correlations ($BF_{10} > 10$) were positive and primarily involved PrEP use, hearing about PrEP, and discussing PrEP use. A total of 17 topics were associated with both PrEP use and either hearing about or discussing PrEP use ($rs = -0.06-0.08$, $BF_{10} > 10$), suggesting that hearing and discussing PrEP could be relevant to the influence of Twitter topics. As with rates, no topics were correlated with discussing HIV.

Additionally, we assessed the direct and mediational pathways in Figure 1. Direct pathways alone would imply that tweets alone can explain HIV prevention and testing, whereas indirect pathways would imply that tweets influence communications, which in turn influence HIV prevention and testing. A summary of the correlational results presented in Table 1 shows 23 possible mediating associations involving either tweet rates or topics, all of which concerned PrEP.

Results for lower rates of men who have sex with men

The results for lower rates of MSM appear in Supplemental Table 2, http://links.lww.com/QAD/C22. Hearing about PrEP and discussing PrEP use were also positively correlated with actual PrEP use ($rs = 0.14-0.58$, $BF_{10} > 100$). Likewise, discussing HIV testing was correlated with actual testing ($rs = 0.18$, $BF_{10} > 100$). However, neither tweet rates nor Twitter topics were associated reported communications or HIV prevention and testing ($BF_{10} < 10$). There was thus no evidence to conduct mediation analyses, which were then conducted for counties with higher rates of MSM only.

Summary

The correlation results were in line with our expectation that, in regions with more MSM, HIV tweets would correlate with reported communications and with HIV prevention and testing. In contrast, the associations between HIV tweets and communications and between HIV tweets and behaviors were not present in regions with fewer MSM.

Assessing mediating associations with HIV tweets in counties with higher rates of men who have sex with men

Table 2 presents the mediational analyses, first for a model from either tweet rates or topics to communications to behaviors (see panel 1) and then for a model from communications to behaviors to either tweet rates or topics (see panel 2). For each mediation model, we used Bayes factors of the indirect path (i.e. path ab) to assess the strength of evidence and Bayes factors of the direct path (i.e. path c') to indicate the type of mediation (i.e. full or partial). As shown in panel 1 (see Table 2), the model with hearing about PrEP and discussing PrEP use as mediators of the influence of the tweets on PrEP use received support (i.e. BF_{10} of path ab > 10). Specifically, tweet rates and six topics (i.e. 0, 17, 49, 69, 76 and 96) were positively correlated with reported communication about PrEP, and these relations were associated with PrEP use (tweet rates: $b_i* = 0.01-0.05$, $BF_{10} > 100$, Twitter topics: $b_i* = 0.01 - 0.05$, $BF_{10} > 10$; see top panel of Fig. 3). According to these results, in counties with higher MSM rates, respondents from areas with higher tweet rates and topic probabilities were more likely to hear about PrEP and discuss PrEP use, and these relations then facilitated higher PrEP use.

Additionally, there were 13 negative correlations between topics and reported communication about PrEP, which revealed that certain topics of discussion had negative, indirect effects on PrEP use ($b_i* = -0.04 - -0.01$, $BF_{10} > 10$). In other words, lower probabilities of various topics (i.e. 1, 8, 12, 35, 42, 45, 54, 62, 66, 83, and 92) promoted hearing about and discussing PrEP, which in turn increased PrEP use.

As shown in panel 2 of Table 2, the results provided some support of the alternative model. There were indirect paths from hearing about PrEP to PrEP use to tweet rates ($b_i* = 0.01$, $BF_{10} > 100$) and to the probabilities of eight topics (i.e. 12, 45, 49, 62, 69, 76, 83 and 96; $b_i* = -0.03-0.01$, $BF_{10} > 10$; see bottom panel of Fig. 3). Hearing about PrEP presumably promotes PrEP use, and PrEP used then increased tweets about HIV and about certain topics (i.e. 49, 69, 76 and 96) and decreased tweets about other topics (i.e. 12, 45, 62 and 83). Lastly, the mediation pathways from discussing PrEP to tweets via PrEP use yielded unidentified Bayes factors, leading to inconclusive results.

Altogether, a certain threshold of communication about the issues may be necessary for tweets to influence behavior. A larger MSM community may actually lead to more tweets or more in person conversations about HIV relevant issues. We conducted additional correlations and found extremely strong evidence ($BF_{10} > 100$) between rates of MSM and communications about PrEP (see Supplemental Information for details).

Discussion

Our analyses supported the notion that, in the United States, social media messages about HIV within a county correlate directly with HIV prevention and testing behaviors (i.e. PrEP use and HIV testing in the past 12 months) or indirectly, via reported communication about PrEP (i.e. hear about PrEP and discuss PrEP use). More importantly, the results were only present in counties with higher rates MSM, suggesting that a sizable community and favorable regional norms are important

Table 2. Mediation involving HIV tweets (i.e. tweet rates and Twitter topics), hearing about PrEP, discussing PrEP, and PrEP use in counties with higher rates of MSM.

Predictor	Mediator	Outcome	Path a b_i^*	Path a Posterior probability	Path b b_i^*	Path b Posterior probability	Path c′ b_i^*	Path c′ Posterior probability	Path ab b_i^*	Path ab Posterior probability	Strength of evidence of the indirect path (mediation type)
Main model (see top panel of Fig. 1)											
Tweet rates											
Tweet rates	Hear about PrEP	PrEP use	0.08***	1.00	0.17***	1.00	0.06***	1.00	0.01***	1.00	Extremely strong (partial)
Tweet rates	Discuss PrEP	PrEP use	0.08***	1.00	0.61***	1.00	0	0.01	0.05***	1.00	Extremely strong (full)
Twitter topics											
Topic 0	Discuss PrEP	PrEP use	0.05*	0.96	0.61***	1.00	0.02	0.10	0.03*	0.96	Strong (full)
Topic 1	Discuss PrEP	PrEP use	−0.05**	0.99	0.61***	1.00	−0.02	0.23	−0.03**	0.99	Very strong (full)
Topic 8	Discuss PrEP	PrEP use	−0.06***	1.00	0.61***	1.00	−0.01	0.06	−0.04***	1.00	Extremely strong (full)
Topic 12	Hear about PrEP	PrEP use	−0.04***	1.00	0.17***	1.00	−0.03	0.83	−0.01***	1.00	Extremely strong (full)
Topic 17	Discuss PrEP	PrEP use	0.05**	0.97	0.61***	1.00	0.03	0.51	0.03**	0.97	Strong (full)
Topic 35	Discuss PrEP	PrEP use	−0.06***	1.00	0.61***	1.00	−0.01	0.03	−0.04***	1.00	Extremely strong (full)
Topic 42	Discuss PrEP	PrEP use	−0.05*	0.97	0.61***	1.00	−0.02	0.25	−0.03*	0.97	Strong (full)
Topic 45	Hear about PrEP	PrEP use	−0.03*	0.97	0.17***	1.00	−0.03***	1.00	−0.01*	0.97	Strong (partial)
Topic 49	Hear about PrEP	PrEP use	0.04***	1.00	0.17***	1.00	0.04***	1.00	0.01***	1.00	Extremely strong (partial)
Topic 49	Discuss PrEP	PrEP use	0.05**	0.97	0.61***	1.00	0.03	0.67	0.03**	0.97	Very strong (full)
Topic 54	Discuss PrEP	PrEP use	−0.05*	0.95	0.61***	1.00	−0.02	0.24	−0.03*	0.95	Strong (partial)
Topic 62	Hear about PrEP	PrEP use	−0.04**	0.98	0.17***	1.00	−0.04***	1.00	−0.01**	0.98	Very strong (partial)
Topic 62	Discuss PrEP	PrEP use	−0.05**	0.98	0.61***	1.00	−0.01	0.08	−0.03**	0.98	Very strong (full)
Topic 66	Discuss PrEP	PrEP use	−0.06***	1.00	0.61***	1.00	−0.02	0.32	−0.04***	1.00	Extremely strong (full)
Topic 69	Hear about PrEP	PrEP use	0.06***	1.00	0.17***	1.00	0.07***	1.00	0.01***	1.00	Extremely strong (partial)
Topic 69	Discuss PrEP	PrEP use	0.07***	1.00	0.61***	1.00	0.03	0.55	0.05***	1.00	Extremely strong (full)
Topic 76	Hear about PrEP	PrEP use	0.04**	0.99	0.17***	1.00	0.04***	1.00	0.01**	0.99	Very strong (partial)
Topic 83	Hear about PrEP	PrEP use	−0.05***	1.00	0.17***	1.00	−0.05***	1.00	−0.01***	1.00	Extremely strong (full)
Topic 83	Discuss PrEP	PrEP use	−0.05***	1.00	0.61***	1.00	−0.03	0.80	−0.03***	1.00	Extremely strong (full)
Topic 92	Hear about PrEP	PrEP use	−0.04***	1.00	0.16***	1.00	−0.02***	1.00	−0.01***	1.00	Extremely strong (partial)
Topic 96	Hear about PrEP	PrEP use	0.04**	0.99	0.17***	1.00	0.03***	1.00	0.01**	0.99	Extremely strong (partial)
Alternative model (see bottom panel of Fig. 1)											
Tweet rates											
Hear about PrEP	PrEP use	Tweet rates	0.17***	1.00	0.06***	1.00	0.07***	1.00	0.01***	1.00	Extremely strong (partial)
Discuss PrEP	PrEP use	Tweet rates	–	–	–	–	–	–	–	–	Unidentified BF
Twitter topics											
Discuss PrEP	PrEP use	Topic 0	–	–	–	–	–	–	–	–	Unidentified BF
Discuss PrEP	PrEP use	Topic 1	–	–	–	–	–	–	–	–	Unidentified BF
Discuss PrEP	PrEP use	Topic 8	–	–	–	–	–	–	–	–	Unidentified BF
Hear about PrEP	PrEP use	Topic 12	0.17***	1.00	−0.03	0.83	−0.03**	0.97	−0.03**	0.97	Extremely strong (partial)
Discuss PrEP	PrEP use	Topic 17	–	–	–	–	–	–	–	–	Unidentified BF
Discuss PrEP	PrEP use	Topic 35	–	–	–	–	–	–	–	–	Unidentified BF
Discuss PrEP	PrEP use	Topic 42	–	–	–	–	–	–	–	–	Unidentified BF
Hear about PrEP	PrEP use	Topic 45	0.17***	1.00	−0.03*	0.92	−0.03	0.81	−0.01*	0.92	Strong (full)
Hear about PrEP	PrEP use	Topic 49	0.17***	1.00	0.04***	1.00	0.03**	0.98	0.01***	1.00	Extremely strong (partial)
Discuss PrEP	PrEP use	Topic 49	–	–	–	–	–	–	–	–	Unidentified BF
Discuss PrEP	PrEP use	Topic 54	–	–	–	–	–	–	–	–	Unidentified BF
Hear about PrEP	PrEP use	Topic 62	0.17***	1.00	−0.04***	1.00	−0.03	0.78	−0.01***	1.00	Extremely strong (full)
Discuss PrEP	PrEP use	Topic 62	–	–	–	–	–	–	–	–	Unidentified BF
Discuss PrEP	PrEP use	Topic 66	–	–	–	–	–	–	–	–	Unidentified BF
Hear about PrEP	PrEP use	Topic 69	0.17***	1.00	0.07***	1.00	0.05***	1.00	0.01***	1.00	Extremely strong (partial)
Discuss PrEP	PrEP use	Topic 69	–	–	–	–	–	–	–	–	Unidentified BF
Hear about PrEP	PrEP use	Topic 76	0.17***	1.00	0.04***	1.00	0.03	0.85	0.01***	1.00	Extremely strong (full)
Hear about PrEP	PrEP use	Topic 83	0.17***	1.00	−0.05***	1.00	−0.04***	1.00	−0.01***	1.00	Extremely strong (partial)
Discuss PrEP	PrEP use	Topic 83	–	–	–	–	–	–	–	–	Unidentified BF
Hear about PrEP	PrEP use	Topic 92	0.17***	1.00	−0.03	0.86	−0.04**	0.98	−0.01	0.86	No
Hear about PrEP	PrEP use	Topic 96	0.17***	1.00	0.07***	1.00	0.05***	1.00	0.01***	1.00	Extremely strong (partial)

b_i^* denotes standardized coefficient estimate.
*$BF_{10} > 10$.
**$BF_{10} > 30$.
***$BF_{10} > 100$.
MSM, men who have sex with men; PrEP, pre-exposure prophylaxis.

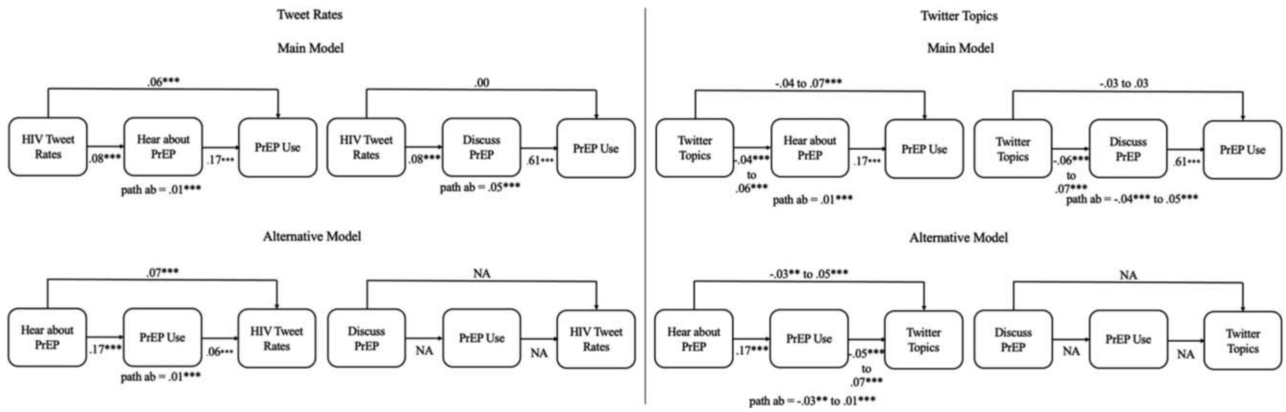

Fig. 3. A summary of mediation results of main (top) and alternative (bottom) models. NA indicates unidentified Bayes factors and standardized estimates.

for in person and social media communications about PrEP or HIV. Tweet rates had consistent positive associations with reported communication about PrEP and actual PrEP use. However, Twitter topics had mixed correlations with reported communication about PrEP and actual PrEP use. Some topics were positively correlated with communication about PrEP and PrEP use, whereas others were negatively correlated with communication and PrEP and PrEP use.

To the best of our knowledge, this study is the first to assess how social media messages influence communications about PrEP that promote PrEP use. There was also some support for the alternative pathway in which hearing about PrEP can promote PrEP use, which influences HIV tweets (see Supplemental Figure 1, http://links.lww.com/QAD/C22). Additionally, these models received support only in counties with higher rates of MSM, suggesting that a real-life community provides the context for in-person communications that are keys for social media to influence behavior. Altogether, our results provided empirical evidence that social media messages are associated with reported communication about PrEP and PrEP use only when MSM rates are higher.

Several caveats are in order.

(1) We examined *ecological* associations between tweets and service use across United States counties. They did not evaluate the longitudinal sequence of tweets and service use, which will only be possible as more years of social media and survey data accumulate and given sufficient change over time.
(2) The study analyzed over 600 000 HIV tweets and survey data from more than 30 000 respondents in 1959 United States counties. Despite the careful study design and statistical analysis, this study may be limited by the analyses of a subset of HIV tweets and the sample representativeness of AMIS, thus reducing the generalizability of the findings. Future research may examine these associations through experiments or large-scale randomized controlled trials.
(3) Our data established associations between county-level tweets and individual-level variables, such as discussing PrEP use and actual PrEP use. Future research should investigate the associations at the individual level.
(4) We examined the presence of messages about HIV based on generative probabilistic models, which calculate the likely topics in a county. Therefore, no identified topics (i.e. clusters of words) can be used to monitor individuals' messaging or biomedical service use.
(5) This work provides preliminary evidence about the effects of tweets about HIV and HIV prevention and testing. The results, therefore, may not generalize to associations with other themes (e.g. drug use) or other outcome variables (e.g. uptake of an eventual HIV vaccine). Future studies should explore generalizability in other contexts and to other media. For example, HIV topics could also be identified from Tumblr posts, Flickr captions, Instagram captions, or other text-based social media when data are available. The strength of the associations with communications may differ across sites, but Twitter remains one of the most popular platforms [20], and its users are representative of the social media population [7] and the younger, diverse populations at risk for HIV [21].

Implications

Our findings offer new insights and have practical implications for HIV prevention and testing, which are likely applicable to other private and sensitive health behaviors. Specifically, regional diversity may create an open atmosphere to discuss topics that empower people to use PrEP and reduce the stigma that often delays help seeking. In addition to new and conventional health strategies, social media conversations hold promise to reduce HIV infections by encouraging communications about prevention and testing strategies and ultimately facilitating actual PrEP use. The rates of MSM appear to explain the relations between tweeting about HIV and having in-person conversations, probably by the sheer volume of MSM in the county or perhaps by mitigating prejudice and stigma [22]. For example, Brooks *et al.* [23] concluded that social biases in a community hinder conversations about healthcare, and these differences, in turn, maintain health disparities. Our analyses provide strong evidence that social media can act as a societal force that sparks dialogues about health and that these dialogues then promote health behaviors. Moreover, they show that individual communications and behaviors may also influence regional social media messages about HIV. Understanding these mechanisms is critical to designing effective large-scale, digital health campaigns that facilitate communication, promote HIV prevention behaviors, and ultimately end the HIV epidemic in the United States.

Acknowledgements

M.-p.S.C., Ph.D., Department of Psychology, University of Illinois; A.M., Department of Computer Science, University of Illinois; M.Z., Ph.D., Rollins School of Public Health, Emory University; P.S., Ph.D., Rollins School of Public Health, Emory University; T.S., Ph.D., Rollins School of Public Health, Emory University; C.Z., Ph.D., Department of Computer Science, University of Illinois; D.A., Ph.D., Department of Psychology, University of Illinois.

Research reported in this publication was supported by multiple grants: The National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Numbers R56AI114501 and R01AI147487, the National Institute of Mental Health of the National Institutes of Health under Award Number R01MH114847, and the National Institute on Drug Abuse of the National Institutes of Health under Award Number DP1DA048570. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of interest

There are no conflicts of interest.

References

1. Ragins BR, Singh R, Cornwell JM. **Making the invisible visible: Fear and disclosure of sexual orientation at work.** *J Appl Psychol* 2007; **92**:1103–1118.
2. Whitehead J, Shaver J, Stephenson R. **Outness, stigma, and primary healthcare utilization among rural LGBT populations.** *PLoS One* 2016; **11**:1–17.
3. White D, Stephenson R. **Identity formation, outness, and sexual risk among gay and bisexual men.** *Am J Mens Health* 2014; **8**:98–109.
4. Eagly AH, Wood W. Social role theory. Handbook of Theories of Social Psychology.London, UK: SAGE Publications Ltd; 2012. pp. 458-476.
5. Buffie WC. **Public health implications of same-sex marriage.** *Am J Public Health* 2011; **101**:986–990.
6. Sullivan P, Mera-GIler R, Mouhanna F, Pembleton E, Guest J, Jones J, *et al.* **Trends in use of oral emtricitabine/tenofovir disoproxil fumarate for preexposure prophylaxis against HIV infections, United States, 2012–2017.** *Ann Epidemiol* 2018; **28**:833–840.
7. Pick J, Sarkar A, Rosales J. **Social media use in American counties: geography and determinants.** *ISPRS Int J Geo-Information* 2019; **8**:424–449.
8. Tweeplers.com. Top tweeting cities in the United States right now. tweeplers.com. 2020. Available at: https://www.tweeplers.com/cities/?cc=US [Accessed 1 December 2020].
9. ten Brinke L, Liu CC, Keltner D, Srivastava SB. **Virtues, vices, and political influence in the U.S. Senate.** *Psychol Sci* 2015; **27**:85–93.
10. Grey JA, Bernstein KT, Sullivan PS, Purcell DW, Chesson HW, Gift TL, *et al.* **Estimating the population sizes of men who have sex with men in US states and counties using data from the American Community Survey.** *JMIR Public Heal Surveill* 2016; **2**:1–12.
11. Chan MS, Lohmann S, Morale A, Zhai C, Ungar LH, Holtgrave DR, *et al.* **An online risk index for the cross-sectional prediction of new HIV, chlamydia, and gonorrhea diagnoses across U.S. counties and across years.** *AIDS Behav* 2018; **22**:2322–2333.
12. Blei DM, Ng AY, Jordan MI. **Latent Dirichlet Allocation.** *J Mach Learn Res* 2003; **3**:993–1022.
13. Gnambs T, Kaspar K. **Disclosure of sensitive behaviors across self-administered survey modes: a meta-analysis.** *Behav Res Methods* 2015; **47**:1237–1259.
14. Burkill S, Copas A, Couper MP, Clifton S, Prah P, Datta J, *et al.* **Using the web to collect data on sensitive behaviours: a study looking at mode effects on the British National Survey of Sexual Attitudes and Lifestyles.** *PLoS One* 2016; **11**:1–12.
15. Dienes Z. **Bayesian versus orthodox statistics: which side are you on?** *Perspect Psychol Sci* 2011; **6**:274–290.
16. Wagenmakers E-J, Wetzels R, Borsboom D, van der Maas HLJ. **Why psychologists must change the way they analyze their data: the case of psi.** *J Pers Soc Psychol* 2011; **100**:426–432.
17. JASP Team. JASP computer software (Version 0.8.6). 2018.
18. Marsman M, Wagenmakers E-J. **Bayesian benefits with JASP.** *Eur J Dev Psychol* 2017; **14**:545–555.
19. Nuijten MB, Wetzels R, Matzke D, Dolan CV, Wagenmakers E-J. **A default Bayesian hypothesis test for mediation.** *Behav Res Methods* 2015; **47**:85–97.
20. Perrin A, Anderson M. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. 2019. Available at: https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/.
21. Wojcik S, Hughes A. Sizing up Twitter users. 2019. Available at: https://www.pewresearch.org/internet/2019/04/24/sizingup-twitter-users/.
22. Bauermeister JA, Muessig KE, LeGrand S, Flores DD, Choi SK, Dong W, *et al.* **HIV and sexuality stigma reduction through engagement in online forums: Results from the healthmpowerment intervention.** *AIDS Behav* 2019; **23**:742–752.
23. Brooks RA, Etzel MA, Hinojos E, Henry CL, Perez M. **Preventing HIV among Latino and African American gay and bisexual men in a context of HIV-related stigma, discrimination, and homophobia: Perspectives of providers.** *AIDS Patient Care STDS* 2005; **19**:737–744.