

UC San Diego

UC San Diego Previously Published Works

Title

#HIV: Alignment of HIV-Related Visual Content on Instagram with Public Health Priorities in the US

Permalink

<https://escholarship.org/uc/item/47w070cq>

Journal

AIDS and Behavior, 24(7)

ISSN

1090-7165

Authors

Nobles, Alicia L

Leas, Eric C

Latkin, Carl A

et al.

Publication Date

2020-07-01

DOI

10.1007/s10461-019-02765-5

Peer reviewed



Published in final edited form as:

AIDS Behav. 2020 July ; 24(7): 2045–2053. doi:10.1007/s10461-019-02765-5.

#HIV: Alignment of HIV-Related Visual Content on Instagram with Public Health Priorities in the US

Alicia L. Nobles¹, Eric C. Leas², Carl A. Latkin³, Mark Dredze⁴, Steffanie A. Strathdee¹, John W. Ayers¹

¹Division of Infectious Diseases and Global Public Health, Department of Medicine, University of California San Diego, La Jolla, CA, USA

²Division of Health Policy, Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA, USA

³Department of Health, Behavior and Society, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA

⁴Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA

Abstract

Instagram, with more than 1 billion monthly users, is the go-to social media platform to chronicle one's life via images, but how are people using the platform to present visual content about HIV? We analyzed public Instagram posts containing the hashtag “#HIV” (because they are self-tagged as related to HIV) between January 2017 and July 2018. We described the prevalence of co-occurring hashtags and explored thematic concepts in the images using automated image recognition and topic modeling. Twenty-eight percent of all #HIV posts included hashtags focused on awareness, followed by LGBTQ (24.5%) and living with HIV (17.9%). However, specific strategies were rarely cited, including testing (10.8%), treatment (10.3%), PrEP (6.2%) and condoms (4.1%). Image analyses revealed 44.5% of posts included infographics followed by people (21.3%) thereby humanizing HIV and stigmatized populations and promoting community mobilization. Novel content such as the handwriting image-theme (3.8%) where posters shared their HIV test results appeared. We discuss how this visual content aligns with public health priorities to reduce HIV in the US and the novel, organic messages that public health could help amplify.

Keywords

HIV; Digital health; Health surveillance; Social media; Big data

Introduction

Although stigma about HIV has declined in the US [1], HIV infections remain a stigmatized condition [2]. People living with HIV face not only internalized stigma, but also enacted

stigma and anticipated stigma, resulting in stereotyping, prejudice, and discrimination among their social networks, health care providers, and society at large [3]. As a result, many people feel uncomfortable discussing HIV-related questions with or disclosing their HIV status to others [4] and turn to online communities to engage in conversations related to HIV [5] and social support [6]. For example, researchers have used online communities to share information about HIV prevention and control [7] and developed peer-to-peer interventions to support people living with HIV who have imperfect antiretroviral (ART) medication adherence [8], and people living with HIV have leveraged online communities for informational and emotional support [6, 9, 10]. Moreover, key population groups that are particularly vulnerable to HIV, including men who have sex with men and adolescents, are increasingly using social media to form relationships [11], access information about reproductive and sexual health [10, 11], and build virtual communities of peers for emotional support [12].

Recent advances in machine learning and natural language processing now allow us to examine HIV-related conversations on publicly available social media at-scale. For example, Young et al. built a machine learning classifier that can identify tweets that mention HIV risk behaviors (sexual activity and substance use associated with HIV infection) with similar accuracy as manual annotation [13]. These geolocated tweets were found to be associated with HIV prevalence [14] and predictive of syphilis incidence [15] indicating that surveillance of social media may complement traditional surveillance of infectious diseases. Similarly, Ireland et al. found geolocated tweets that contained future oriented language were associated with lower rates of HIV infection in counties identified as vulnerable because of the higher prevalence of behaviors associated with increased risk of HIV infection [16]. Beyond surveillance, Nobles et al. examined a Reddit community focused on sexually transmitted infections and found that posters seek informational support about how HIV is transmitted, the risk of HIV infection associated with select behaviors, and HIV testing window periods [17]. To date the majority of at-scale analyses of HIV-related content on social media are text-based and primarily oriented to complement traditional surveillance of health outcomes.

Visual social media platforms, such as Instagram, are accelerating in popularity [18], especially among young adults. Despite interest in using visual content to communicate about HIV [19], little is known about how the public already uses images on social media to communicate about HIV. Given the stigmatized nature of HIV, how do people use Instagram, a platform in which users cite their primary motivation of usage is to surveil others and document their life [20], to communicate about HIV?

In this manuscript, we take an important first step in describing how the public is sharing images that are self-tagged as related to HIV on Instagram. Instagram, described as akin to a “virtual photo album [20]”, is an image sharing platform where hashtag(s) serve as anchoring devices to label the content of Instagram posts and make the content searchable. We selected Instagram for our study because it is the fastest growing social media platform with 1 billion monthly users and most young adults in the US are likely to use it (67% of 18-29 year olds in the US) [18]. Our objectives in this study were to: (1) identify and describe the hashtags that are commonly associated with images self-tagged as HIV-relevant

and (2) explore the themes present in the images. We contrast the hashtags and themes identified in the visual content to the priorities identified by the Ending the HIV Epidemic: A Plan for America agenda (e.g., Phase II target is to reduce new HIV infections in the US by 90% by 2030 [21]) to examine alignment with public health priorities in the US and opportunities to amplify organic advocacy initiated by the public.

Methods

We assembled a dataset of public Instagram posts from January 2017 and July 2018 using InstaLooter [22], a software that scrapes Instagram for images and associated captions given a hashtag. Posts with the hashtag “#HIV” were collected to capture posts that posters self-labeled as related to HIV. The posts were then restricted to those with captions authored in English, as identified by the automated language identification software Language Identification tool [23], so that comparisons could be drawn to US public health targets and the hashtags would be interpretable by the English speaking investigative team.

We then assessed the potential focus of the posts by analyzing hashtags that co-occurred in the posts with #HIV. First, we identified the most used common hashtags by their frequency of co-occurrence with #HIV. Second, author AN pruned the 1000 most used hashtags to retain only HIV-related or HIV at-risk hashtags, reducing the set to 178 hashtags. For example, non-specific hashtags that were removed included #love, #memes, and #ifunny. Third, authors AN and JA independently grouped the remaining 178 hashtags into clusters according to their primary thematic topic based on iterative discussions between the authors, including: (a) antiretroviral therapy (ART), (b) awareness, (c) condoms, (d) non-specific prevention, (e) persons living with HIV, (f) pre-exposure prophylaxis (PrEP), (g) testing, or groups at-risk of HIV, including (h) lesbian, gay, bisexual, transgender, queer or questioning (LGBTQ). The two authors agreed on 160 out of the 178 (90%) hashtags, and discussed discordant labels to come to a consensus agreement. We report the frequency of the posts within each topic.

We then analyzed the visual content of the images in the posts using procedures first developed by Chen and Dredze [24]. First, we used automated image recognition software to assign textual tags (e.g., “man”, “woman”, “indoors”, “doctor”, “hospital”) to the content in the images. The labeling of images relies on a deep convolutional neural network model with a classification accuracy of 89.3% for the top 5 textual tags of an image [25]. After each image was assigned a textual tag, we further grouped tags into topics using Latent Dirichlet Allocation (LDA) to assign topics to each image [26, 27]. The LDA model we employed is contained in a natural language processing software called MALLET [28]. MALLET is Java-based, so we used the gensim library, which makes the model accessible using Python [26]. We identified 30 topics present in the images and, although an image can have multiple topics, we assigned the topic with the largest probability as the predominant topic of each image, consistent with previous studies [29]. The top 10 textual tags associated with each topic and five random images associated with each topic were examined by authors AN, EL, and JA to manually assign a label to each topic. Topics that had similar labels were further grouped into overarching themes for presentation, resulting in a total of

10 themes present in the images. Identifying features of images and usernames were omitted to protect user privacy [30].

The resulting hashtags, hashtag clusters, and image-themes were described descriptively using frequencies and percentages. This study was deemed exempt review by the University of California San Diego Human Research Protections Program because it solely used anonymous publicly available data.

Results

During the study period, we discovered 39,233 Instagram posts containing #HIV with 26,766 of the posts authored in English which were retained for further analysis. Sixty-six percent of the posts included co-occurring hashtags in addition to #HIV, with the average post containing 2.39 additional hashtags (median = 1, min = 0, max = 19). The 25 most frequent co-occurring hashtags included #AIDS, cited in 42.7% of the posts, followed by #gay (13.9%), #love (10.3%), #hivawareness (8.8%), and #cancer (8.5%) (Table 1). About half (13) of the top 25 hashtags were primarily related to HIV (e.g., #hivpositive) or HIV at-risk groups (e.g., #instagay); the remainder were related to infections beyond HIV (e.g., #herpesdating), or the use of common social media jargon (e.g., #memes).

Twenty-eight percent of all #HIV posts included hashtags focused on fostering general awareness, including #hivawareness, #endthestigma, or #aidsday (Table 2). The second most frequent cluster of hashtags was for sexual and gender minorities, such as #gay, #LGBT, or #instagay, and were included in 24.5% of the posts. Additional hashtag clusters included persons living with HIV (17.9%), testing (10.8%), ART (10.3%), and non-specific prevention (8.1%). Surprisingly, specific prevention strategies were rarely mentioned. For example, PrEP, was represented in just 6.2% of the posts, most often with #PrEP, #Truvada, #PrEPforlove, or #PrEPworks. Condoms were only cited in 4.1% of the posts, including #condoms or #wrapitup.

Among the themes present in the images, numerous HIV-relevant narratives emerged (Table 3). Infographics were by far the most frequently occurring theme in the images, being the focus of 44.5% of all #HIV images. These images included embedded text which typically cited statistics on HIV and encouraged prevention (e.g., encouraging disclosure of status or usage of PrEP), treatment (e.g., information about the benefits of reducing viral load), or surveillance (e.g., testing). People were the second-most common image-theme, accounting for 21.3% of all #HIV images, and included images that focused on persons as the subject of the photo, such as a selfie with a clinical setting, two men holding hands, or groups of people at HIV-related events and fundraisers. Many of the images of people presented persons living with HIV, people who identify as gender minorities, and relationships of sexual minorities. The theme of people was followed by images containing outdoor (8.9%) or indoor settings (4.4%), advertising for or sharing of events (e.g., sporting events or black-tie affairs) (4.4%), pictures of handwriting (3.8%), taking place in educational contexts (3.6%), being associated with healthcare (3.4%), or containing food (3.1%). Images in the handwriting theme frequently illustrated HIV testing results, with posters sharing their status. Images in the healthcare theme frequently presented doctors or a medical setting, or

pictures of ART or PrEP pills. Images in the events theme frequently presented fundraising campaigns or awareness events.

Discussion

Our analysis indicates that there are thousands of Instagram posts that are publicly sharing HIV-related visual images that can be potentially viewed by Instagram's audience. Understanding the visual content that Instagramers are publicly sharing can inform (1) about how the public is using images to engage in HIV-related communication, (2) if the existing visual content aligns with the agenda to reduce HIV infections in the US [31], and (3) potential novel messages that are already being adopted organically that may be amplified with assistance by public health professionals.

We find that the public is presenting visual content on Instagram that addresses both clinical interventions, which have a direct effect on reducing the incidence, morbidity and mortality of HIV, as well as the social enablers, which are critical to mobilize people to engage with clinical interventions [27]. About half of all #HIV images were infographics with information about and promoting clinical interventions (testing, PrEP, and ART) provided in the embedded text as well as messages related to social enablers (reducing the stigma of sexual minorities and regular HIV testing). Almost one quarter of all #HIV images primarily included images of people. These images often focused on de-stigmatization of sexual minorities as well as persons living with HIV by presenting people in everyday and intimate contexts, thereby humanizing the person(s) in the image as a unique individual(s) and not an abstract demographic or disease. Stigma, especially among key populations where new infections are highly concentrated (men who have sex with men, racial/ethnic minorities, and those that live in the southern US) is identified as a social challenge inhibiting clinical interventions [32]. Images of groups of people highlighted people engaged in community mobilization, which connected people often in the context of educational, outreach, or interventional activities. Moreover, posters highlighted the needs of vulnerable minority groups by including hashtags relevant to the LGBTQ community in about 25% of all #HIV image captions.

Visual content of *specific* clinical interventions (early detection, usage of ART for viral suppression, and usage of PrEP for prevention - identified as three of the four strategies essential to achieving the President's agenda) are not well represented on Instagram relative to public health priorities. Just 6% of all posts include hashtags that indicated the poster was intending to promote PrEP, even though PrEP is a prevention priority among leaders because of its efficacy in reducing the risk of HIV acquisition [33]. Similarly, less than 1 in 20 of all posts included hashtags focused on condoms, and only 1 in 10 of posts included hashtags encouraging engagement in ART. However, we observed that some visual content is promoting specific clinical interventions that can reduce the exposed audience's risks, sometimes in surprising ways. For instance, some posters shared their HIV test results to normalize the experience of getting tested to protect themselves and the community.

Social media offers a valuable opportunity to communicate about HIV and reach key populations [29, 34]. Public health professionals might adopt the promotion of novel organic

messages on social media, such as test sharing, using a photovoice method [30] combined with a hashtag to diffuse the campaign. Photovoice methods are a community-based participatory research approach, where community stakeholders (in this case, lay persons that are tested for HIV) combine their own photographs documenting their actions with grassroots social action, thereby promoting community-based action. The right message on social media, such as #MeToo that went viral on social media to increase awareness of sexual violence [31], can quickly spread and have a profound impact on social discourse. Leaders could create a hashtag such as #MyHIVTest and encourage the public to share their HIV tests (with personal information removed) or a sticker indicating they were just tested, much like campaigns to encourage voting. Although sharing test results have been advocated in some settings, such as among African American church leaders [35], it has not been promoted online by any coordinated campaign.

Selecting a strategic hashtag for hashtag activism movements is critical [32]. Despite more than half of the posts (66%) containing an additional two hashtags on average, a predominant, unique hashtag does not stand out for each hashtag theme. Instead, we found inconsistency among the usage of various hashtags for each theme. The lack of predominant hashtag(s) in these clusters may indicate difficulty with the dissemination of a public health campaign. Future research should focus on the hashtags that are commonly used in target audiences as well as potential peer health influencers on Instagram that have large target audiences that can be leveraged to increase the visibility of information that could be diffused through their network.

Finally, the visual content and hashtags humanize HIV and stigmatized groups, such as the LGBTQ community, but there is room for improvement. Although transgender women have the highest HIV prevalence among key populations [36], there was only one hashtag among the top 1000 that was exclusively dedicated to ‘trans’, and there were few posts involving gender minorities in general.

Limitations

First, we limited our analysis to posts authored in English, and therefore limit our comparisons to public health targets in the US. We acknowledge that content relevant to other countries and cultural contexts may differ, such as a need to focus on social enablers in countries with restrictive gender norms. Since geolocation is not readily available, future studies could focus on methods to extract geolocation from the text of a user’s profile. Additionally, we do not extend inferences to other social media platforms because the Instagram community cannot be considered representative of all social media platforms. Future work could focus on how visual content on Instagram compares to other social media platforms and other cultural contexts. Second, not everyone uses social media; however, Instagram’s primary demographics overlaps with the age group of those most at risk for a new infection in the US [37]. Third, we only used public images, therefore our analysis is subject to self-selection bias (i.e., those that elect to publicly share their images) and did not capture visual content that is being broadcast in private networks. Fourth, we did not consider how contextual factors (e.g., demographics of the poster) may impact the theme of the visual content, but consideration of intersectionality (e.g., the macroaggressions and

microaggressions towards racial minorities and persons living with HIV) is critical to fully understand the willingness of key populations to discuss topics in online communities and an important area for future work. Fifth, although we observed an encouraging amount of organic engagement with HIV by the public, we may be potentially undercounting engagement if a post does not contain #HIV and instead uses other hashtags. Finally, although some argue that social media data are a misrepresentation of the public, numerous studies show that social media monitoring can validly infer public health trends [38-44].

Conclusions

In this study, we leverage automated methods (automated image recognition and topic modeling) to examine visual content on Instagram, which is the fastest growing image sharing platform, at-scale. We identified that visual content primarily focuses on presenting infographics about clinical interventions (testing, ART, and PrEP) and presenting images of people aimed at reducing stigma of sexual minorities and people living with HIV. We also identified potential messages, such as people sharing their anonymized test results, that could be amplified with the assistance of a unified campaign to reduce the stigma associated with being tested for HIV. We identified several areas important to future research (such as identification of appropriate hashtags for hashtag activism and peer health influencers) to better understand how visual content can be best leveraged to achieve targets for reduction in HIV incidence in the US.

Acknowledgements

This research was supported by funds from the California HIV/AIDS Research Program Office of the University of California (OS17-SD-001). Dr. Nobles is also supported by the National Institute of Drug Abuse (T32 DA023356). The content is solely the responsibility of the authors and does not necessarily represent the official views of the California HIV/AIDS Research Program Office or National Institute of Drug Abuse.

References

1. Herek GM, Capitanio JP, Widaman KF. HIV-related stigma and knowledge in the United States: prevalence and trends, 1991–1999. *Am J Public Health*. 2002;92:371–7. [PubMed: 11867313]
2. Berger BE, Ferrans CE, Lashley FR. Measuring stigma in people with HIV: psychometric assessment of the HIV stigma scale. *Res Nurs Health*. 2001;24:518–29. [PubMed: 11746080]
3. Eaton LA, Allen A, Maksut JL, Earnshaw V, Watson RJ, Kalichman SC. HIV microaggressions: a novel measure of stigma-related experiences among people living with HIV. *J Behav Med*. 2019. 10.1007/s10865-019-00064-x.
4. Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one's HIV-positive status, stigma and social support. *AIDS Care*. 2008;20:1266–75. [PubMed: 18608080]
5. Mo PKH, Coulson NS. Living with HIV/AIDS and use of online support groups. *J Health Psychol*. 2010;15:339–50. [PubMed: 20348355]
6. Mo PKH, Coulson NS. Exploring the communication of social support within virtual communities: a content analysis of messages posted to an online HIV/AIDS support group. *Cyberpsychol Behav*. 2008;11:371–4. [PubMed: 18537512]
7. Muessig KE, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman LB. A systematic review of recent smartphone, Internet and Web 2.0 interventions to address the HIV continuum of care. *Curr HIV/AIDS Rep*. 2015;12:173–90. [PubMed: 25626718]
8. Horvath KJ, Oakes JM, Rosser BRS, Danilenko G, Vezina H, Amico KR, Williams ML, Simoni J. Feasibility, acceptability and preliminary efficacy of an online peer-to-peer social support ART adherence intervention. *AIDS Behav*. 2013;17:2031–44. [PubMed: 23553347]

9. Coursaris CK, Liu M. An analysis of social support exchanges in online HIV/AIDS self-help groups. *Comput Hum Behav.* 2009;25:911–8.
10. Magee JC, Bigelow L, Dehaan S, Mustanski BS. Sexual health information seeking online: a mixed-methods study among lesbian, gay, bisexual, and transgender young people. *Health Educ Behav.* 2012;39:276–89. [PubMed: 21490310]
11. Lim MSC, Vella A, Sacks-Davis R, Hellard ME. Young people’s comfort receiving sexual health information via social media and other sources. *Int J STD AIDS.* 2014;25:1003–8. [PubMed: 24616114]
12. Saha K, Kim SC, Reddy Carter AJ, Sharma E, Haimson OL, De Choudhury M. The language of LGBTQ + minority stress experiences on social media. *Proc ACM Hum-Comput Interact.* 2019;3:89. [PubMed: 32935081]
13. Young SD, Yu W, Wang W. Toward automating HIV identification. *J Acquir Immune Defic Syndr.* 2017;74:S128–31. [PubMed: 28079723]
14. Young SD, Rivers C, Lewis B. Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes. *Prev Med.* 2014;63:112–5. [PubMed: 24513169]
15. Young SD, Mercer N, Weiss RE, Torrone EA, Aral SO. Using social media as a tool to predict syphilis. *Prev Med.* 2018;109:58–61. [PubMed: 29278678]
16. Ireland ME, Schwartz HA, Chen Q, Ungar LH, Albarraçin D. Future-oriented tweets predict lower county-level HIV prevalence in the United States. *Health Psychol.* 2015;34S:1252–60. [PubMed: 26651466]
17. Nobles AL, Dreisbach CN, Keim-Malpass J, Barnes LE. “Is this a STD? please help!”: online information seeking for sexually transmitted diseases on Reddit. In: International AAAI Conference on Weblogs and Social Media. 2018. pp 660–663.
18. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. In: Pew Research Center. <https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/>. Accessed 25 Sep 2019.
19. Koegler E, Thomson TJ, Speno AG, Teti M. Image-sharing via social media: reflections from an ethnically- and age-diverse sample of people living with HIV in the Midwest. *J HIV/AIDS Soc Serv.* 2018;17:249–62.
20. Sheldon P, Bryant K. Instagram: motives for its use and relationship to narcissism and contextual age. *Comput Hum Behav.* 2016;58:89–97.
21. What is “ending the HIV epidemic: a plan for America”? In: HIV. gov. (2019). <https://www.hiv.gov/ending-hiv-epidemic> Accessed 30 Sep 2019.
22. Larralde M. InstaLooter. <https://github.com/althonos/InstaLooter>. Accessed 27 Aug 2018.
23. Lui M. Language Identification (LangID). <https://github.com/saffsd/langid.py>. Accessed 27 Aug 2018.
24. Chen T, Dredze M. Vaccine images on twitter: analysis of what images are shared. *J Med Internet Res.* 2018;20:e130. [PubMed: 29615386]
25. Clarifai. <https://www.clarifai.com/technology>. Accessed 27 Aug 2018.
26. Rehurek R, Sojka P. Software framework for topic modelling with large corpora. Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks. 2010.
27. Schwartländer B, Stover J, Hallett T, et al. Towards an improved investment approach for an effective response to HIV/AIDS. *Lancet.* 2011;377:2031–41. [PubMed: 21641026]
28. McCallum AK. MALLET: A Machine Learning for Language Toolkit. <http://mallet.cs.umass.edu>. 2002.
29. 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. [PubMed: 29175811]
30. Catalani C, Minkler M. Photovoice: a review of the literature in health and public health. *Health Educ Behav.* 2010;37:424–51. [PubMed: 19797541]
31. Caputi TL, Nobles AL, Ayers JW. Internet searches for sexual harassment and assault, reporting, and training since the #MeToo Movement. *JAMA Intern Med.* 2019;179:258–9. [PubMed: 30575847]

32. Wang R, Liu W, Gao S. Hashtags and information virality in networked social movement. *Online Inf Rev.* 2016;40:850–66.
33. Cohen MS, Baden LR. Preexposure prophylaxis for HIV—where do we go from here? *N Engl J Med.* 2012;367:459–61. [PubMed: 22784041]
34. Taggart T, Grewe ME, Conserve DF, Gliwa C, Roman Isler M. Social media and HIV: a systematic review of uses of social media in HIV communication. *J Med Internet Res.* 2015;17:e248. [PubMed: 26525289]
35. Berkley-Patton J, Thompson CB, Moore E, Hawes S, Simon S, Goggin K, Martinez D, Berman M, Booker A. An HIV testing intervention in African American Churches: pilot study findings. *Ann Behav Med.* 2016;50:480–5. [PubMed: 26821712]
36. Herbst JH, Jacobs ED, Finlayson TJ, McKleroy VS, Neumann MS, Crepaz N, HIV/AIDS Prevention Research Synthesis Team. Estimating HIV prevalence and risk behaviors of transgender persons in the United States: a systematic review. *AIDS Behav.* 2008;12:1–17. [PubMed: 17694429]
37. U.S. statistics. In: [HIV.gov.](https://www.hiv.gov/hiv-basics/overview/data-and-trends/statistics) (2019). <https://www.hiv.gov/hiv-basics/overview/data-and-trends/statistics>. Accessed 30 Sep 2019.
38. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. *J Mach Learn Res.* 2003;3:993–1022.
39. Ayers JW, Althouse BM, Dredze M. Could behavioral medicine lead the web data revolution? *JAMA.* 2014;311:1399–400. [PubMed: 24577162]
40. Strand C Blogging: a new tool for coping and accessing psycho-social support for people living with HIV? In: Åsa, editor. *E-health communities and online self-help groups: applications and usage.* PA: IGI Global; 2011. p. 106–20.
41. Paz-Bailey G, Hoots BE, Xia M, Finlayson T, Prejean J, Purcell DW, NHBS Study Group. Trends in internet use among men who have sex with men in the united states. *J Acquir Immune Defic Syndr.* 2017;75(Suppl 3):S288–95. [PubMed: 28604430]
42. Paul MJ, Dredze M. *Social monitoring for public health.* CA: Morgan & Claypool Publishers; 2017.
43. Mehrotra R, Sanner S, Buntine W, Xie L. Improving LDA topic models for microblogs via tweet pooling and automatic labeling. In: *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval.* New York: ACM; 2013. p. 889–92.
44. Ayers JW, Caputi TL, Nebeker C, Dredze M. Don't quote me: reverse identification of research participants in social media studies. *Digit Med.* 2018;1:30. 10.1038/s41746-018-0036-2.

Table 1

25 most common co-occurring hashtags

<i>Hashtag</i>	<i>Percentage of posts</i>
aids	42.7
gay	13.9
love *	10.3
hivawareness	8.8
cancer *	8.5
std *	8.3
hivpositive	7.9
herpes *	7.6
health *	7.3
aidsawareness	7.2
knowyourstatus	6.8
lgbt	6.7
instagay	6.2
prep	5.8
memes *	5.7
gettested	5.5
positivesingles	5.2
hpv *	5.2
herpesdating *	4.8
lgbtq	4.8
herpesdatingsites *	4.4
dank *	4.1
hivprevention	4.0
hsv *	3.9
ifunny *	3.8

Note: Each hashtag was labeled for being primarily HIV or HIV at-risk group related, or not, with the latter designated by an *

Table 2


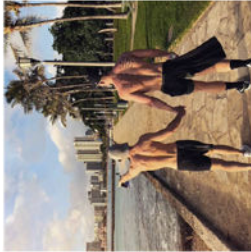

Rank order of co-occurring hashtag themes on #HIV posts


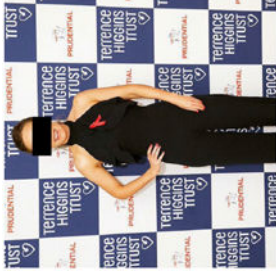
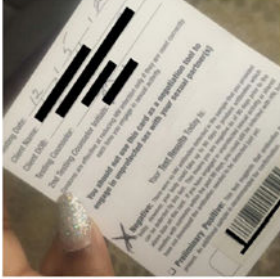
Category	Percentage of posts	Hash-tags
Awareness	28.5	hivawareness, endthestigma, aidsday, hiveducation, hivaid
LGBTQ	24.5	gay, lgbt, intsagay, lgbtq, gaymen, gaylife, bisexual
Persons living with HIV	17.9	hivpositive, livingwithhiv, poz, positive, gaypoz
Testing	10.8	gettested, hiv testing, testing, nationalhivtestingday, nhdt
ART	10.3	uequalsu, art, treatment, undetectable, untransmittable, hivtreatment
Non-specific prevention	8.1	hivprevention, prevention, preventhiv, playsafe, safeissex
PrEP	6.2	prep, truvada, prep4love, prepworks
Condoms	4.1	condoms, condom, wrapitup, freecondoms


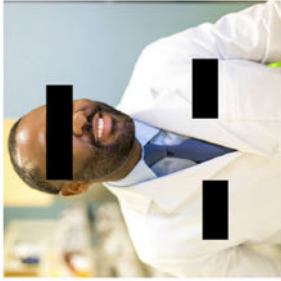
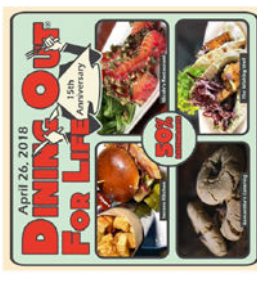
Note: Posts were obtained from all original posts (N = 26,766) with #HIV written in English. Co-occurring hashtags represent the top 1000 after pruning to remove those unrelated to HIV (e.g., #love)


Table 3

Rank order of image-topics in #HIV posts

Topic	Percentage of posts	Example textual tags	Example images
Graphics	44.5	symbol, design, template, presentation, banner	
People	21.3	man, people, woman, adult, child	
Outdoors	8.9	outdoors, nature, sky, travel, landscape	

Topic	Percentage of posts	Example textual tags	Example images
Indoors	4.4	room, sitting, people, furniture, indoors	
Events	4.42	competition, race, people, soccer, crowd	
Handwriting	3.79	paper, text, document, writing, business	

Topic	Percentage of posts	Example textual tags	Example images
Education	3.59	education, group, school, administration, leader	
Healthcare	3.38	medicine, healthcare, science, health, hand	
Food	3.14	person, food, health, drink, glass	

Topic	Percentage of posts	Example textual tags	Example images
Technology	2.54	business, technology, internet, communication, computer	

Note: Posts were obtained from all original posts (N = 26,766) with #HIV written in English