Raising Awareness About Cervical Cancer Using Twitter: Content Analysis of the 2015 #SmearForSmear Campaign

Philippe Lenoir, Bilel Moulahi, Jérôme Azé, Sandra Bringay, François
Carbonnel

HAL Id: lirmm-01633320
https://hal-lirmm.ccsd.cnrs.fr/lirmm-01633320

Submitted on 12 Nov 2017

ORIGINAL PAPER

Authors:
LENOIR Philippe, MOULAHI Bilel, AZE Jérôme, BRINGAY Sandra, CARBONNEL François

LENOIR Philippe, MD
Department of General Practice, University of Montpellier, UFR Médecine site Nord, 641 Avenue du Doyen Giraud 34093 Montpellier cedex 52, France

MOULAHI Bilel,
LIRMM UM-CNRS
LIRMM, University of Montpellier, CNRS, 860 rue de St Priest, 34095 Montpellier cedex 5, France

AZE Jérôme,
LIRMM UM-CNRS
LIRMM, University of Montpellier, CNRS, 860 rue de St Priest, 34095 Montpellier cedex 5, France

BRINGAY Sandra,
LIRMM UM-CNRS
AMIS Group, University Paul Valéry, route de mende, 34090 Montpellier, France

CARBONNEL François, MD-MSc
Research Unit EA4556 Epsylon, University of Montpellier, Rue Henri Serre, 34000 Montpellier, France
CEPS Platform, Universities of Montpellier, Rue Henri Serre, 34000 Montpellier, France
Department of General Practice, University of Montpellier, UFR Médecine site Nord, 641 Avenue du Doyen Giraud 34093 Montpellier cedex 52, France
Avicenne Multiprofessional Health Center, 2 rue Ibn Sinaï dit Avicenne, 66330 Cabestany, France

Corresponding Author:
CARBONNEL François
Centre de Médecine Générale Avicenne, 2 rue Ibn Sinaï dit Avicenne, 66330 Cabestany, France.
+33684014834
drfrancoiscarbonnel@gmail.com

Title:
Raising awareness about cervical cancer using Twitter: a content analysis of the 2015 #SmearForSmear campaign

Running title:
Lessons learned from the #SmearforSmear campaign

Abstract:

Background: Cervical cancer is the 2nd cause of cancer among women under age 45. To deal with the decrease of smear test (ST) coverage in United Kingdom, a Twitter campaign called #SmearForSmear has been launched in 2015 for the European Cervical Cancer Prevention Week. Its aim was to encourage women to take a selfie showing their lipstick going over the edge and post it on Twitter with a raising awareness message promoting cervical cancer screening. Estimated audience was 500 million of people. Other public health campaigns have been launched on social medias such as Movember to encourage participation and self-engagement. Their result was unsatisfactory as their aim had been diluted to become mainly a social buzz.

Objective: Our objectives were to identify the tweets delivering a raising awareness message promoting cervical cancer screening (sensitizing tweets) and the characteristics of Twitter users posting about this campaign.

Methods: 3-step content analysis of the English tweets tagged #SmearForSmear posted on Twitter for the 2015 European Cervical Cancer Prevention Week. Data were collected using Twitter API. Their extraction was based on an analysis grid generated by two independent researchers using a thematic analysis, validated by a strong Cohen's kappa coefficient. 7 themes coded for sensitizing tweets and 13 for Twitter users' status. Verbatims were thematically then statiscally analyzed.

Results: 3018 tweets were collected, 1881 analysed. 70% of tweets had been posted by people living in United Kingdom. 57.4% of users were women, sex was unknown in 36% of cases. 54.4% of the users had posted at least one selfie with smeared lipstick. 32.2% of tweets were sensitizing. Independent factors associated with posting sensitizing tweets were: woman who experienced an abnormal ST (OR 13.456 [CI95% 3.101-58.378], p = 0.0005), female sex (OR 3.752 [CI95% 2.133-6.598], p < 0.0001) and live in the United-Kingdom (OR 2.097 [CI95% 1.447-3.038], p < 0.0001). Non-sensitizing tweets were statistically more posted by non-health or non-media company (OR 0.558 [CI95% 0.383-0.814], p = 0.0024).

Conclusions: This study demonstrates that the success of a public health campaign using a social media depends on its ability to get its targets involved. It also suggests the need to use social marketing to help its dissemination. The clinical impact of this Twitter campaign to increase cervical cancer screening is yet to evaluate.

Keywords:

uterine cervical neoplasms; papanicolaou test; social media; early detection of cancer; health promotion; smear test; twitter

MAIN ARTICLE BODY:

Introduction:
Cervical cancer is the second cause of cancer among women under age 45 and leads to significant mortality [1]. Cervical cancer is caused by human papilloma virus (HPV) [2]. Smear test (Papanicolaou test) detects precancerous changes and earlier stage cervical cancers. Its introduction has allowed a dramatic decline of cervical cancer incidence and death rates in many countries, such as the developed countries [3]. In England, an organized national screening program has been established in 1988. Incidence of cervical cancer in women aged 20–79 in England has almost halved from 1982 to 2006, thanks to this program. However, its incidence is now rising in 20-29-year-olds from 1996 onwards in most regions in England [4]. From 1999 to 2013, the number of women that didn't attend their smear test for a 5-year period has progressively increased from 16% to 22% [5]. It suggests that organized screening isn't intrinsically strong enough to keep a high coverage rate.
Social medias would have a great potential to improve behavior change as interactive tools, encouraging participation and self-engagement instead of a « descending » information [6-8]. Facebook, Twitter or Instagram had respectively more than 1.86 billion, 317 million and 500 million monthly active users in December 2016. For Twitter, more than 500 million tweets are traded every day [9]. These social media have become a valuable source of information for health professional and clinicians to effectively discovering health related topics and behaviors online [32, 36, 37, 38]. Public health campaigns have already tried to take advantage of the ability of social medias to make a campaign viral. The ALS Ice Bucket challenge's goal was to mediatize and fundraise for the amyotrophic lateral sclerosis. It had involved many celebrities worldwide. On 1 September 2014, more than 17 million videos had been shared on Facebook and had been watched more than 10 billion times by more than 440 million people [10]. Thanks to this campaign, more than 100 million dollars had been raised by the ALS association [11]. Hundreds of thousands of people had tweeted daily about ALS, which is a much higher number of tweets than those emitted about multiple sclerosis, a disease better known to the public [12]. Movember is an annual event organized each November since 2003 whose goal is to raise awareness and fundraise about masculine disease such as prostate or testicle cancer. Participants let their moustache grow and post selfies on social medias to raise the awareness of their contacts and show their involvement in this campaign. In Denmark, after the initiation of the 2011 Movember campaign, a significant decline in the PSA level at referral and an increase in the number of patients referred under suspicion of prostate cancer was observed. However, only minor differences in referral patterns and prostate cancer diagnosis were detected [13]. These campaigns may be parasitized by the buzz they seeked to create and may vehicle non-health-related messages. A content analysis of the 2013 Movember Canada campaign on Twitter showed that it did not meet the stated campaign objective of creating conversations about men's health and, specifically, about prostate and testicular cancers [14].
To deal with the decrease of smear coverage in United Kingdom, a Twitter campaign called #SmearForSmear has been launched in 2015 by the Jo's Cervical Cancer Trust for the European Cervical Cancer Prevention Week. Its goal was to encourage women to take a picture of themselves (selfie) showing their lipstick going over the edge and post them on Twitter with an awareness message promoting cervical cancer screening. Estimated audience was 500 million people [15].
Our objectives were to identify the tweets delivering raising awareness messages about cervical cancer screening (RAM) and the characteristics of Twitter users posting about this campaign.

Methods:
We conducted a 3-step content analysis of the English tweets posted on Twitter during the 2015 European Cervical Cancer Prevention Week.

Data collection and extraction:
To collect the tweets, we used Twitter Application Programming Interface (API). It allows the user to conduct manual searches for keywords in tweets with specific parameters such as hashtags, language, and date range. The ones used for this research were: #SmearforSmear, English language, and tweets posted between 25 January 2015 to 31 January 2015 inclusive (European Cervical Cancer Prevention Week). All tweets have been manually collected and assessed. Only original tweets, rather than

retweets, were analyzed. In the tweets, only the verbatim were transcribed. Hashtags and content preceded by "@" were removed if that action didn't make the verbatim unintelligible. We also considered all hypertexts linked to another verbatim on another web-platform (e.g., Instagram). The corresponding verbatim were transcribed only if they were informative.

Data analysis:
A total of 3019 tweets that met the search criteria were imported into Excel for data extraction. An analysis grid had been created based on 200 tweets initially collected and thematically analyzed by two independent researchers to extract the themes (topics) of tweets' verbatim and Twitter users' statuses. Then, this grid had been tested on 50 new tweets. No new themes had been identified, confirming that category saturation was achieved. The thematic analysis methodology used consists of transforming qualitative content into a quantitative form by establishing coding categories. The number of data units that fall into each coding category are counted (such as phrases, messages, responses). Finally, they are categorized based on similar meanings and overt or inferred communication [16,17]. Themes were not restricted to preexisting themes. They emerged through an inductive process whereby open coding of data revealed themes that moved from the specific to the general [18]. The two researchers, both general practitioners and trained in qualitative study, elaborated a 7-themes codebook, based on tweets, to identify if the tweets delivered raising awareness messages about cervical cancer screening: *incentive to carry out the smear test, evocation of smear test importance without any precision, reminder of the smear test preventive nature, reminder of the low incidence of smear test, allusion to the mortality or morbidity of cervical cancer, reminder of the incidence of cervical cancer* and *testimony of an experience related to smear test or cervical cancer*. If a tweet had at least one of these awareness-raising messages, it was considered a sensitizing tweet. Reproducibility of the classification of 300 tweets by the two independent researchers was tested and calculated with Cohen's kappa coefficient. The agreement was strong and varied between 0.8842 and 1.
The following information was collected about each tweet: *verbatim, posting date, retrieval date, presence of a selfie with lipstick going over the edge, picture or video referring to the campaign, user sex, user location, number of followers at the date of retrieval, presence of a selfie with lipstick going over the edge,* and *user's status.* To classify the users, we used their Twitter status. If it didn't exist or was incomplete, we extracted these information from links on their Twitter profile, whenever possible. The analysis grid enlisted 14 themes regarding Twitter user's status: *health company, media company, non-health and non-media company, marketing company, fashion company, blogger or youtuber, health professional, National Health Service (NHS), politician, woman who experienced cervical cancer or relatives, woman with unspecified cancer or relatives, woman who experienced an abnormal smear test, general public* and *unknown*. The "*unknown*" status was attributed when no information to categorize the user was available. Only the "*unknown*", "*general public*" or "*NHS*" status were exclusive.
An initial global description of the sample has been performed, using the frequencies of the different categories for the qualitative variables. As the distribution of quantitative variables wasn't always Gaussian (Shapiro-Wilk test), they were expressed by their mean, standard deviation, median, minimum and maximum values and interquartile. Comparison of means was executed through the Student test when distribution was Gaussian, otherwise based on Mann-Whitney test. Comparison of qualitative variables was executed through the Chi2 test for parametric tests, or Fisher exact test when the conditions for applying Chi2 were not observed. A multivariate logistic modeling process was then carried out to identify the independent factors associated with the presence of a sensitizing message in the tweets and associated with each type of sensitizing message. A "step by step" selection procedure of the variable was used with an input and output variable set at 0.10 and 0.05 respectively. The significance threshold was set to 5%. Statistical analysis had been performed by the Department of Medical Information at Montpellier teaching hospital with SAS version 9.
This research did not require ethics approval since we only used publically available Twitter content.

Results:
3018 tweets met the search criteria. 1138 were removed (retweets or copies of tweets) and 1881 were analysed.

608 tweets (32,2%) were sensitizing and included from 1 to 5 raising awareness message (RAM), mean 0.54 +/- 0.93 (Table 1). Incentive to carry out the smear test was the most frequent RAM.

Table 1: Description of tweets and Twitter users

Variable	Total n (%)
Number of raising awareness message in a tweet	
- 0	1273 (67.68)
- 1	347 (18.45)
- 2	149 (7.92)
- 3	83 (4.41)
- 4	25 (1.33)
- 5	4 (0.21)
Sensitizing tweet	608 (32.32)
- Incentive to carry out the smear test	440 (23.39)
- Reminder of smear test preventive nature	217 (11.54)
- Allusion to the mortality or morbidity of cervical cancer	134 (7.12)
- Testimony of an experience related to smear test or cervical cancer	92 (4.89)
- Smear test importance	63 (3.35)
- Evidence of the number of cervical cancers	41 (2.18)
- Low Incidence of smear test	27 (1.44)
Categories of Twitter users:	
- Unknown	442 (23.5)
- Non-health and non-media company	396 (21.05)
- Health company	292 (15.52)
- Blogger or youtuber	262 (13.93)
- Media company	240 (12.76)
- Fashion activity	240 (12.76)
- Marketing activity	220 (11.70)
- NHS	79 (4.2)
- General public	77 (4.09)
- Woman who experienced cervical cancer or relatives	60 (3.19)
- Health professional	53 (2.82)
- Woman who experienced an abnormal ST	33 (1.75)
- Politician	12 (0.64)
- Woman who experienced unspecified cancer or relatives	6 (0.32)

Main users were people from English-speaking countries. United Kingdom accounted for 70% of the posted tweets, followed by USA (9%) and Australia (1%). Nationality was unknown in 15.2% of cases. 57.4% of users were women, sex was unknown in 36% of cases. Twitter users had a mean number of followers of 44 420.8 +/- 420 819. 54.4% of the users had posted at least one selfie with smeared lipstick. 15.6% of the tweets were associated with a picture or a video referring to the #SmearforSmear campaign.

Univariate and multivariate analysis
Significant statistic associations between emitting sensitizing tweets and Twitter user status are detailed on Table 2.

Table 2: Twitter users known characteristics

Characteristics	Total n (%)	P value
Characteristics linked to a higher probability of emitting **sensitizing tweets**:		
- United-Kingdom	1316 (82.51)	<0.000001
- Female sex	1079 (89.62)	0.000002
- NHS	79 (4.2)	0.020013

- Woman who experienced an abnormal ST	33 (1.75)	<0.000001
Characteristics linked to a higher probability of emitting non-sensitizing tweets:		
- Non-health or non-media company	396 (21.05)	<0.000001
- Media	240 (12.76)	0.044861
- Marketing activity	220 (11.70)	<0.000001
- Man	125 (10.38)	0.000002

The "step by step" selection procedure has allowed to identify independent factors influencing the sensitizing characteristic of tweet (Table 3).

Table 3: Independent factors influencing the emission of sensitizing tweets (* statiscally significant influence on the emission of non-sensitizing tweets)

Message of tweet, Variables	Adjusted OR (95% CI)	P value
Sensitizing tweet		
- Woman who experienced an abnormal ST	13.456 (3.101, 58.378)	0.0005
- Female Sex	3.752 (2.133, 6.598)	<0.0001
- United-Kingdom	2.097 (1.447, 3.038)	<0.0001
- Non-health or non-media company*	0.558 (0.383, 0.814)	0.0024
Direct encouraging people to go for ST		
- Female Sex	5.967 (2.606, 13.659)	<0.0001
- Health company	2.203 (1.042, 4.656)	0.0386
- United-Kingdom	1.997 (1.320, 3.021)	0.0011
- Selfie	1.673 (1.228, 2.280)	0.0011
- Non-health or non-mass media company*	0.481 (0.310, 0.746)	0.0011
Evocation of the importance of smear test without any precision		
- Woman who experienced an abnormal ST	7.365 (2.314,23.436)	0.0007
- NHS	4.266 (1.778,10.238)	0.0012
- United-Kingdom	2.888 (1.015,8.212)	0.0467
- Fashion	2.724 (1.430,5.188)	0.0023
- Selfie	2.158 (1.163,4.002)	0.00147
Reminder of the preventive aspect of smear test		
- Woman who experienced an abnormal ST	4.216 (1.734,10.254)	0.0015
- Politician	3.545 (1.028,12.221)	0.00451
- Female sex	2.555 (1.156,5.646)	0.00204
- Marketing activity*	0.414 (0.211,0.812)	0.00103
Evocation of the mortal or morbid aspect of cervical cancer		
- Woman who experienced unspecified cancer or relatives	6.359 (1.043,38.776)	0.0449
- Woman who experienced an abnormal ST	5.591 (2.227,14.035)	0.0002
- Female Sex	3.396 (1.050,10.982)	0.0412
- Woman who experienced cervical cancer or relatives	2.598 (1.228,5.495)	0.00125
- United-Kingdom	2.268 (1.069,4.808)	0.0328
Reminder of the low incidence of smear test		
- Politician	14.754 (3.074,70.816)	0.0008
Reminder of the incidence of cervical cancer		
- General public	2.913 (1.002,8.474)	0.0497
- Picture or a video linked to the #SmearforSmear campaign	2.701 (1.372,5.318)	0.0040
Statement from people who experienced abnormal smear test or cervical cancer		
- Woman who experienced an abnormal ST	65.364 (22.709,188.140	<0.0001
- Woman who experienced unspecified cancer or relatives	14.371 (2.335,88.417)	0.0040
- Woman who experienced cervical cancer or relatives	7.641 (3.690,15.822)	<0.0001

Discussion:

Main results:

32.2% of the tweets of the #SmearforSmear campaign were sensitizing. This result was promising as it goes well beyond the results of the 2013 Movember campaign where only 0.85% of the posted tweets may raise awareness of men's health risks [13]. Many factors may explain this gap. On one hand, this campaign had been created using social marketing in a holistic approach. Its objective was clear, its title referred to its objective. Jo's cervical cancer trust posted key messages reflecting the need to adhere to the screening of cervical cancer, used to fill the content of tweets. A slogan had also been created *"Attend your smear, reduce your risk"*, widely retweeted in this campaign. On the other hand, this campaign created to detect an exclusive feminine cancer was based on elements of two women's social construct, lipstick and selfies [19,20]. This approach was possible because this campaign had been designed for the UK, where the cervical cancer screening is organized. Targeted women automatically receive a letter explaining them what to do to get screened and where. Receiving an invitation letter is an independent sensitizing factor associated with greater likelihood of cervical cancer screening [21].

As for the Twitter users, our expectations were broadly confirmed. From a general point of view, Twitter users posting sensitizing tweets were people personally involved in cervical cancer screening: women, women concerned by a feminine cancer or their relatives, people living in the UK (where this English-speaking campaign took place), the NHS as a partner of this campaign and women who experienced an abnormal ST. As peer, women raised awareness by insisting on the preventive aspect of ST and directly encouraged other women to attend their ST. Peer influence is known as an important social lever for health-related behaviour change [22]. Likewise, women or their relatives who experienced a pathological state (abnormal ST, cervical cancer or an unspecified cancer) had the greatest potential among categories of Twitter users to post a sensitizing tweet. Hashtags, such as #SmearforSmear, tend to create communities behaving as support group [23]. Unveiling elements of private life is conducive to trust and emotional bond [24]. Fashion company was a user status that has a significant potential to post tweets about the importance of smear test without any precision. Actively participating to the campaign by posting selfies and pictures or videos linked to the #SmearforSmear campaign helped encouraging people to attend their ST and disseminate the importance of ST. Women's magazines also act as a guidebook and reinforce women's individual responsibility to create and maintain good health for themselves and their families [25]. As for the other user categories, the RAM in their tweets was in line with expectations. Politicians broadcasted information about the low incidence of ST and how it helped preventing cervical cancer, in relation with their use of social media to communicate with the press and the public [26]. Health companies' RAM was more direct, encouraging people to attend their ST. The general public cautioned about incidence of cervical cancer. The NHS insisted on the importance of ST without giving more information. It was probably in relation with the fact that NHS was only a partner of this campaign and that it only helped disseminating it. Finally, there was a scotoma of health professionals. This status didn't emerge as a relevant category. Their participation to a health campaign on social medias is interesting as it has been shown that the information contained in their posts are more likely to be true compared to other groups [27]. This under-representation was probably due to the shortness of the studied campaign period.

Conversely, "non-sensitizing" tweets had a much greater probability to be sent either by users not directly concerned by cervical cancer such as men (exclusively feminine cancer) or by users that participated but only broadcasted information, without getting involved: medias, marketing companies and non-health/non-media companies. It questions their participation to this campaign. Was it about an opportunistic appropriation of a viral campaign? It is probably one of the main limitations of the virality of health campaigns on social media. Most tweets posted for the 2013 Movember campaign and the breast cancer prevention month did not spark conversations about prostate and testicular cancer nor promote any specific preventive behavior about breast cancer [14,28]. They may also be an interesting lever for social stimulation.

Strengths & Limitations:

To our knowledge, no study analyzing the content of the #SmearforSmear campaign on Twitter has been published yet. Our findings are corroborated by the content analysis of others health campaigns on Twitter. We used a content analysis method based on a double analysis of the sensitizing capacity of each tweet, in an exploratory process. We also mined Twitter to gather information about users' characteristics and complete the tweets' content.

This highly demanding method made us decide early to restrict our study to one week. This choice was also relevant, in our opinion, as this campaign had been created for the European Cervical Cancer Prevention Week. Compared to other Twitter campaigns, our relatively high results must question its ability to keep a high proportion of sensitizing tweets in other countries (particularly where the cervical cancer screening is not organized) and if it remains high over time.

The choice to collect the tweets based on the hashtag #SmearforSmear may have limited their number, by omitting those not using it. As for the content analysis, two safeguards have been used: analyzing the content of tweets to create the categories before the study and evaluate the reproducibility of the classification by two independent researchers with Cohen's kappa coefficient, which was strong in this study. The shortness of Twitter posts, limited to 140 characters may have created a loss of information as users often used hyperlinks to be exempt from this limit. We then chose to manually mine Twitter to complete the tweets' content and gather information about users' characteristics.

Perspectives:

The #SmearforSmear campaign has allowed to disseminate sensitizing messages about cervical cancer screening and became viral. It was based on a well-designed campaign, on a facilitating audience and a facilitating health system using an organized screening.

Choosing a social media adapted to the target is a major concern for a successful campaign. Twitter is interesting as it is well-suited for appointment campaigns such as #SmearforSmear or the ALS ice bucket challenge. It also is a social media used by young adults to keep up in real-time with news [29]. But its audience is mainly male, living in urban areas. And although diverse, its percentage of users with college educations and incomes over $50,000 is much higher than those of Facebook or Instagram. Users of Instagram are mainly female but 72% of online American adults use Facebook and its audience is the most engaged with 70% logging on daily [29]. Health campaigns on social medias must be seen as a way to reduce social inequities in health. In the UK, the main decline in screening was about 25-49-year-old women and black and Asiatic ethnic minorities [11]. Targeted audience must be on the social media chosen and then adapt to the shift of the evolution of their audiences.

The impact of facilitators is to be studied. As previously shown, many Twitter users of this campaign didn't engage in this campaign as they didn't post sensitizing tweets. But they participated and helped broadcasting to their audience. Models such as Cara Delevingne also posted a selfie to support campaign to her millions of followers (8.5 million in May 2017) [30]. They may boost a campaign as influencers and role model.

The present findings show a clear need for studies that are capable of automatically analyze the data and extracting useful insights from the #Smear4Smear Twitter campaign. We propose the use of machine learning to tackle these challenges, and we suggest three perspectives for future directions. First, we plan to undertake a large-scale analysis using a collection of tweets that we are currently crawling since February 2017. Thisanalysis will include (i) the application of the Latent Dirichlet Allocation (LDA) in order to extract the topics [31-33] emerging from the discussions about the campaign, as well as (ii) the exploration of the linguistic style [34] of the Twitter' users. Second, we could benefit from statistical learning techniques [35] to predict automatically the categories of all tweets about the campaign. This study may allow us to assess our findings and generalize our results. We will learn a model with the One-Vs-the-Rest multi-class classifier based on an annotated data set and we will apply it on all tweets about the campaign. We will compare the results to the manual processing and annotation done so far. Moreover, within a sufficiently large data set, we can take advantage of machine learning models to use features that are more complex to characterize the users tweeting about the campaign. We suggest focusing on user groups including health professionals, celebrities, general public politicians. This will lead us to understand which group of users is prominent, so that it could influence others making them to action by retweeting the messages relevant to the campaign, liking and replying to tweets or more importantly donating money. Third, we plan to

investigate the temporal distribution of messages to focus on the campaign dynamics over time. We may study the temporal correlations between the reactions of twitter users and real world events such as media coverage of the campaign. This analysis is exploratory and it could help in identifying the factors contributing to the raising the awareness. For example, a televised promotion of the campaign or a promotion published by a celebrity may stimulate a huge volume of tweets and reactions online. Beyond this, we can also analyze the geographical distribution of tweets during the campaign. Health campaigns on social networks may raise awareness of public health issues. Their clinical impact is yet to evaluate. It will be a hard task in a hyper-connected world to be able to individualize the effect. This scientific step is important to convince stakeholders, health professionals and general public to get involved and use Web 3.0 as a collective intelligence to drive back chronic diseases [37-38], particularly for the most fragile ones.

Multimedia appendix of supplementary files:
Supplementary file 1

References:
1. Ferlay J SI, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources,methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015 Mar 1;136(5):E359-86. PMID: 25220842
2. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. 1999 Sep;189(1):12-9. PMID:10451482
3. Thun MJ, DeLancey JO, Center MM, Jemal A, Ward EM. The global burden of cancer: priorities for prevention. Carcinogenesis 2010 Jan;31(1):100-10. PMID: 19934210
4. Foley G, Alston R, Geraci M, Brabin L, Kitchener H, Birch J. Increasing rates of cervical cancer in young women in England: an analysis of national data 1982-2006. Br J Cancer 2011 Jun 28;105(1):177-84. PMID:21654675
5. Marlow L a. V, Wardle J, Waller J. Understanding cervical screening non-attendance among ethnic minority women in England. British Journal of Cancer 2015 Sep 1;113(5):833-9. PMID:26171938
6. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008 Nov 4;10(5):e36. PMID:18984557
7. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. PMID:21189422
8. Neiger BL, Thackeray R, Burton SH, Giraud-Carrier CG, Fagen MC. Evaluating social media's capacity to develop engaged audiences in health promotion settings: use of Twitter metrics as a case study. Health Promot Pract 2013 Mar;14(2):157-62. PMID:23271716
9. Bagadiya J. 151 amazing social media statistics you should know in 2017. Social pilot. https://socialpilot.co/blog/151-amazing-social-media-statistics-know-2017. Archived at: http://www.webcitation.org/6qPIbslqL

10. Facebook newsroom. The ice bucket challenge on Facebook. https://newsroom.fb.com/news/2014/08/the-ice-bucket-challenge-on-facebook. Archived at: http://www.webcitation.org/6qPLqVSlx

11. Worland J. ALS Ice Bucket Challenge Donations Just Topped $100 Million. Time. http://time.com/3222224/als-ice-bucket-challenge-donations-million. Archived at: http://www.webcitation.org/6qPM5xnIE

12. Koohy H, Koohy B. A lesson from the ice bucket challenge: using social networks to publicize science. Frontiers in Genetics. 2014 Dec 15;5:430. PMID: 25566317

13. Thomsen FB, Mikkelsen MK, Hansen RB, Brasso K. The Movember campaign: Impact on referral patterns and diagnosis of prostate cancer. Scand J Public Health 2016 May;44(3):228-32. PMID : 26511591

14. Bravo CA, Hoffman-Goetz L. Tweeting About Prostate and Testicular Cancers: Do Twitter Conversations and the 2013 Movember Canada Campaign Objectives Align? J Cancer Educ 2016 Jun;31(2):236-43. PMID:25649663

15. Smith L. Cervical Cancer Prevention Week : Facts, statistics and the #SmearForSmear awareness campaign. International Business Times. http://www.ibtimes.co.uk/cervical-cancer-prevention-week-facts-statistics-smearforsmear-awareness-campaign-1539347. Archived at: http://www.webcitation.org/6qPEnSY2T

16. Kondracki NL, Wellman NS, Amundson DR. Content analysis: review of methods and their applications in nutrition education. J Nutr Educ Behav 2002;34(4):224-230. PMID:12217266

17. Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-88. PMID:16204405

18. Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-15. PMID:18352969

19. Etcoff NL, Stock S, Haley LE, Vickery SA, House DM. Cosmetics as a feature of the extended human phenotype: modulation of the perception of biologically important facial signals. PLoS ONE 2011;6(10):e25656. PMID:21991328

20. Fox J, Vendemia MA. Selective Self-Presentation and Social Comparison Through Photographs on Social Networking Sites. Cyberpsychol Behav Soc Netw 2016 Oct;19(10):593-600. PMID:27732079

21. Tavasoli SM, Pefoyo AJK, Hader J, Lee A, Kupets R. Impact of invitation and reminder letters on cervical cancer screening participation rates in an organized screening program. Prev Med 2016 Jul;88:230-6. PMID:27143497

22. Webel AR, Okonsky J, Trompeta J, Holzemer WL. A systematic review of the effectiveness of peer-based interventions on health-related behaviors in adults. Am J Public Health 2010 Feb;100(2):247-53. PMID:20019321

23. Park H, Reber BH, Chon M-G. Tweeting as Health Communication: Health Organizations' Use of Twitter for Health Promotion and Public Engagement. J Health Commun 2016 Feb;21(2):188-98.

24. Ussher J, Kirsten L, Butow P, Sandoval M. What do cancer support groups provide which other supportive relationships do not? The experience of peer support groups for people with cancer. Soc Sci Med 2006 May;62(10):2565-76. PMID:16303220

25. Roy SC. « Taking charge of your health »: discourses of responsibility in English-Canadian women's magazines. Sociol Health Illn 2008 Apr;30(3):463-77. PMID:18194356

26. Lyons BA, Veenstra AS. How (Not) to Talk on Twitter: Effects of Politicians' Tweets on Perceptions of the Twitter Environment. Cyberpsychol Behav Soc Netw 2016 Jan;19(1):8-15. PMID: 26779660

27. Alnemer KA, Alhuzaim WM, Alnemer AA, Alharbi BB, Bawazir AS, Barayyan OR, et al. Are Health-Related Tweets Evidence Based? Review and Analysis of Health-Related Tweets on Twitter. J Med Internet Res 2015 Oct 29;17(10):e246. PMID:26515535

28. Thackeray R, Burton SH, Giraud-Carrier C, Rollins S, Draper CR. Using Twitter for breast cancer prevention: an analysis of breast cancer awareness month. BMC Cancer 2013 Oct 29;13:508. PMID:24168075

29. Walton J. Twitter vs Facebook vs Instagram : who is the target audience ? Investopedia. http://www.investopedia.com/articles/markets/100215/twitter-vs-facebook-vs-instagram-who-target-audience.asp. Archived at: http://www.webcitation.org/6qQTVyI8b

30. Delevingne C. Twitter. https://twitter.com/caradelevingne. Archived at: http://www.webcitation.org/6qQZAjG1m

31. Wang S, Paul MJ, Dredze M. Exploring health topics in Chinese social media: An analysis of Sina Weibo. Proceedings of the AAAI Workshop on the World Wide Web and Public Health Intelligence 2014. p. 20–23.32. Zhan Y, Liu R, Li Q, Leischow SJ, Zeng DD. Identifying Topics for E-Cigarette User-Generated Contents: A Case Study From Multiple Social Media Platforms. Journal of medical Internet research 2017;19(1).

33. Lu Y, Zhang P, Liu J, Li J, Deng S. Health-related hot topic detection in online communities using text clustering. Plos one 2013;8(2):e56221.

34. Zeng QT, Tse T. Exploring and developing consumer health vocabularies. Journal of the American Medical Informatics Association 2006;13(1):24–29.

35. Himmel W, Reincke U, Michelmann HW. Text mining and natural language processing approaches for automatic categorization of lay requests to web-based expert forums. Journal of medical Internet research 2009;11(3):e25.

36. Attard A, Coulson NS. A thematic analysis of patient communication in Parkinson's disease online support group discussion forums. Computers in Human Behavior 2012;28(2):500–506.

37. Lemire M, Paré G, Sicotte C, Harvey C. Determinants of Internet use as a preferred source of information on personal health. International journal of medical informatics 2008;77(11):723–734.

38. Ybarra ML, Suman M. Help seeking behavior and the Internet: a national survey. International journal of medical informatics 2006;75(1):29–41.

Abbreviations:
ST: smear test
GPs: General Practitioners
NHS: National Health Service
UK: United-Kingdom