clinical practice guidelines by the Infectious Diseases Society of America. *Clin Infect Dis* 2006;43:1089–134.

26 Wills AB, Spaulding AB, Adjemian J, et al. Long-term follow-up of patients with Lyme disease: longitudinal analysis of clinical and quality-of-life measures. *Clin Infect Dis* 2016;62:1546–51.

27 Berende A, Hofstede HJM, Vos FJ, et al. Randomized trial of longer-term therapy for symptoms attributed to Lyme disease. *N Engl J Med* 2016;374:1209–20.

28 Jaulhac B, Saunier A, Caumes E, et al. Lyme borreliosis and other tick-borne diseases. Guidelines from the French scientific societies (II). Biological diagnosis, treatment, persistent symptoms after documented or suspected Lyme borreliosis. *Med Mal Infect* 2019;49:335–46.

29 Clavier VD, Oliveira J-P, Badau E. *Nommer la viande sans antibiotique: la communication de l'industrie agro-alimentaire américaine entre engagement et garantie de la qualité sanitaire. In: Alimentation et Santé: Logiques D'acteurs en Information et Communication.* ISTE: 55–77 2018 Série L'information dans la santé coordonnée par Viviane Clavier et Céline Paganelli, Collection Ingénierie de la santé et société.

30 Collins LC, Nerlich B. Uncertainty discourses in the context of climate change: a corpus-assisted analysis of UK national newspaper articles. *Communications* 2016; 41:291–313.

31 Collins L, Nerlich B. Examining user comments for deliberative democracy: a corpus-driven analysis of the climate change debate online. *Environ Commun* 2015; 9:189–207.

32 Sophie M. *Les Discours de la Presse Quotidienne. Observer, Analyser, Comprendre.* PUF (Lingu. PUF (Linguistique nouvelle), editor. Paris, 2007: 180 p.

33 Europresse [en ligne] Disponible sur. Available at: https://nouveau.europresse.com/Login/ (3 January 2018, date last accessed).

34 Perronne C. *La Vérité Sur la maladie de Lyme.* Chap 8 Mon expérience de médecin-chercheur face à la maladie de Lyme. Paragraphe: Ou est passé la recherche médicale? O Jacob, editor. Paris, 2017: 165 p.

35 Perronne C. Critical review of studies trying to evaluate the treatment of chronic Lyme disease. *Press Med* 2015;44:828–31.

36 Horowitz R. *Lyme Disease Presenting as Multiple Chronic Infectious Disease Syndrome (MCIDS/SIMC) & Co-infections: Diagnosis and Treatment JNI.* Toulouse, 2011. [Internet]. Available at: http://www.infectiologie.com/UserFiles/File/medias/JNI/JNI11/CT/JNI2011-Lyme-horowitz.pdf (21 July 2018, date last accessed).

37 Perronne C. *La Vérité Sur la maladie de Lyme.* Chap 6 La résistance s'organise Paragraphe Des experts américains rebelles au 'dogme' sont sanctionnés ou discrédités. O Jacob, editor. Paris 2017: 100 p.

38 FFMVT, Fédération Française des Maladies Vectorielles à Tiques. Fonctionnement/Bureau [Internet]. Available at: http://ffmvt.org/bureau/ (21 July 2018, date last accessed).

References 39–43 are available in the Supplementary material.

How do Twitter users react to TV broadcasts dedicated to vaccines in Italy?

Francesco Gesualdo [iD] [1], Angelo D'Ambrosio[2], Eleonora Agricola[1], Luisa Russo[1], Ilaria Campagna[1], Beatrice Ferretti[1], Elisabetta Pandolfi[1], Marco Cristoforetti[3], Alberto E. Tozzi[1], Caterina Rizzo[1]

1 Predictive and Preventive Medicine Research Unit, Multifactorial and Complex Disease Research Area, Bambino Gesù Children's Hospital, IRCCS, Rome, Italy
2 Department of Public Health and Pediatric Sciences, University of Turin, Turin, Italy
3 Information and Communication Technology Department, Fondazione Bruno Kessler, Trento, Italy

Correspondence: Francesco Gesualdo, Predictive and Preventive Medicine Research Unit, Multifactorial and Complex Disease Research Area, Bambino Gesù Children's Hospital, IRCCS, Piazza S. Onofrio, 4, 00165 Rome, Italy, Tel: +39 06 68 592852, Fax: +39 06 6859 2853, e-mail: f.gesualdo@gmail.com

Background: Social media monitoring during TV broadcasts dedicated to vaccines can provide information on vaccine confidence. We analyzed the sentiment of tweets published in reaction to two TV broadcasts in Italy dedicated to vaccines, one based on scientific evidence [Presadiretta (PD)] and one including anti-vaccine personalities [Virus (VS)]. **Methods:** Tweets about vaccines published in an 8-day period centred on each of the two TV broadcasts were classified by sentiment. Differences in tweets' and users' characteristics between the two broadcasts were tested through Poisson, quasi-Poisson or logistic univariate regression. We investigated the association between users' characteristics and sentiment through univariate quasi-binomial logistic regression. **Results:** We downloaded 12 180 tweets pertinent to vaccines, published by 5447 users; 276 users tweeted during both broadcasts. Sentiment was positive in 50.4% of tweets, negative in 37.7% and neutral in 10.1% (remaining tweets were unclear or questions). The positive/negative ratio was higher for VS compared to PD (6.96 vs. 4.24, $P<0.001$). Positive sentiment was associated to the user's number of followers (OR 1.68, $P<0.001$), friends (OR 1.83, $P<0.001$) and published tweets (OR 1.46, $P<0.001$) and to being a recurrent user (OR 3.26, $P<0.001$). **Conclusions:** Twitter users were highly reactive to TV broadcasts dedicated to vaccines. Sentiment was mainly positive, especially among very active users. Displaying anti-vaccine positions on TV elicited a positive sentiment on Twitter. Listening to social media during TV shows dedicated to vaccines can provide a diverse set of data that can be exploited by public health institutions to inform tailored vaccine communication initiatives.

Introduction

According to the World Health Organization, at a global level, immunization coverage is stable.[1] In 2017, the Italian government introduced a law making 10 vaccines mandatory for infants, which contributed to a recent, positive trend in vaccination coverage.[2]

Before the introduction of the law for mandatory vaccinations, a constant decrease in immunization coverage had been recorded in Italy for all antigens (apart from pneumococcus and meningococcus)[3] with a consequent rise in the risk of vaccine preventable disease epidemics. A measles epidemic occurred in 2017 and 2018, and 4991 cases and 4 deaths were recorded in 2017.[4]

This reduction in immunization coverage can be attributed, at least partially, to a decrease of vaccine confidence. Confidence in vaccination is dependent upon a complex system of determinants, that includes social, economic and cultural characteristics, at the individual and community levels.[5] Understanding vaccine confidence determinants, characterizing risk perception and exploring how these factors impact on vaccine hesitancy is a complex task.[6–8] Although traditional surveys remain a mainstay for studying vaccine confidence in the population,[5] methods commonly used in social marketing,[9] such as social media monitoring, allow to obtain data with a higher speed, compared to official channels.[6]

Mass media may have an impact on vaccine hesitancy. In 1982, public concern raised after the airing of a TV broadcast reporting non-fact-based data on DTP safety.[10] More recently, the delivery of negative contents on the human papillomavirus (HPV) vaccine through mass media increased Google searches on this issue, and correlated with a reduction of HPV vaccine coverage.[11] Messages on vaccines transmitted by traditional media commonly elicit a reaction on social media, therefore, monitoring social media in reaction to news transmitted by traditional media can provide a consistent and diverse set of information on vaccine confidence among social media users. This kind of information can be useful for public health agencies to tailor communication initiatives to internet users' information needs.

In 2016, two TV programmes on vaccines were broadcasted on Italian TV. Both elicited an intense discourse on social media, in particular, on Twitter. In this study, we analyze the use of Twitter during these broadcasts dedicated to vaccines, with the aim of exploring the potential of this kind of media monitoring for informing public health practice.

Methods

This is a mixed-methods study that utilizes both quantitative and qualitative methods to analyze the sentiment of vaccine-related tweets published before, during and after the airing of two broadcasts dedicated to vaccines on the Italian TV. The two programmes explored vaccines from different perspectives.

Presadiretta (PD) aired on prime time on 10 January 2016. It was followed by an audience of 1 690 000 and had a programme share of 6.93%. The show included video-interviews that addressed fears and misconceptions about vaccinations through communicating scientific evidence.

Virus (VS) aired on 12 May 2016, was followed by 1 252 000 and had a programme share of 5.61%. The broadcast was a talk show and most of it was dedicated to anti-vaccine themes that were predominantly explored by vaccine-hesitant guests.

Twitter analysis

Twitter data were extracted from the database of the VCMP platform, a system monitoring vaccine confidence in Italy based on web data. The VCMP platform was developed by the Bambino Gesù Children's Hospital (Rome, Italy) and the Bruno Kessler Foundation (Trento, Italy) for the real-time monitoring of Google searches, online news and social media discourse on vaccines in Italy.

The platform exploited the following data sources: Google Trends, Facebook (monitoring 400 open Facebook groups), Twitter, Google News and other news platforms. The platform also included a system facilitating the classification of tweet sentiment.

This article is based on Twitter data only. The data set included tweets and retweets in Italian acquired through the Twitter Search application programming interface (API),[12] containing at least one keyword obtained from a set of general, health-related keywords. To increase specificity, also negative keywords were included in the query. Tweets were downloaded hourly and duplicate entries were eliminated. Text, tweet's information (whether it was an original tweet, a retweet or a reply; publication time) and user's information (number of posted tweets, followers and friends) were automatically extracted for each tweet. Tweets were filtered based on presence of the keyword 'vaccin*' (i.e. all keywords starting with 'vaccin'), in order to select those actually relevant to vaccines. For each of the two TV programmes, we selected tweets published during a period of 8 days, centred on the broadcast airing. PD aired on 10 January 2016 from 20.45 to 22.15 (UTC); VS aired on 15 December 2016 from 19.00 to 21.30 (UTC). We categorized the tweets as posted in a window of 4 h around the time of the broadcasts' airing (20.00–24.00 UTC of 10 January 2016 for PD; 19.00–23.00 of 12 May 2016 for VS), or before, or after this window. Recurrent users were defined as users who posted contents during both periods.

Two trained researchers (E.A. and E.P.) manually and independently reviewed selected tweets. Tweets not pertaining to human vaccinations were excluded. Included tweets were classified by sentiment. *Positive* sentiment was assigned to tweets promoting the benefits of vaccinations, encouraging vaccination, reporting sound scientific references or criticizing anti-vaccine movements. *Negative* sentiment was assigned to tweets against immunizations, promoting comments or references from anti-vaccine movements. Tweets were classified as *neutral* when promoting the broadcast or reporting news or events without the expression of a specific vaccine sentiment. Tweets not matching any of the previous definitions were classified as *unclear*. Finally, tweets expressing information needs were classified as *Questions*. When no agreement was obtained in the classification among the two researchers, a decision was taken by a third researcher (C.R.).

Statistical analysis

The number of tweets and users were described as number, percentage and 95% confidence intervals (95% CIs); users' characteristics (i.e.: tweets, followers and friends by user) were described as median and interquartile range (IQR); and retweets/tweets and positive/negative tweets ratios were reported. Differences were tested through Poisson, quasi-Poisson or logistic univariate regression as appropriate.

Sentiment trends were smoothed using a local regression model[13] with a span parameter of 0.75.

The association between users' characteristics and sentiment was studied through univariate quasi-binomial logistic regression models. The regression estimates were regularized with a Bayesian approach, using a weakly informative Cauchy prior distribution.[14] The analyses were performed on the global dataset and separately for both PD and VS. Odds ratios (OR) of positive over negative tweets were reported.

Since the tweets retrieved through the Twitter Search API represent an unknown sample of the total tweet production, uncertainty about all the cited statistics is displayed through 95% CI. Consequently, we are assuming that our sample is consistent with the totality of produced tweets conditional on search keywords and timeframe.

To identify themes associated with positive or negative tweets, we used the following approach: we tokenized tweets, eliminated stop words, URLs and usernames and treated hashtags as common terms; different lexical forms of the terms were joined by the relative lemma (e.g. is, are, am—be);[15] we retained only words of at least three

Table 1 General characteristics of users and tweets by broadcast

	Broadcast		*P*-value	Total
	Presadiretta (PD)	**Virus (VS)**		
Users—*N* (%; 95% CI)	3183 (58.4%; 57.1–59.7%)	2637 (48.4%; 47.1–49.7%)	<0.001[a]	5447 (100%)
Friends by user—median (IQR)	322 (122–864)	435 (155–1090)	<0.001[c]	360 (131–949)
Followers by user—median (IQR)	224 (60–787)	411 (121–1330)	0.281[c]	284 (75–994)
Tweets by user—median (IQR)	2990 (569–12 300)	6960 (1540–23 300)	<0.001[c]	4070 (806–16 300)
Total tweets—*N* (%; 95% CI)	7960 (65.4%; 64.5–66.2%)	4220 (34.6%; 33.8–35.5%)	<0.001[a]	12 180 (100%)
Original tweets—*N* (%; 95% CI)	2494 (31.3%; 30.3–32.4%)	1546 (36.6%; 35.2–38.1%)	<0.001[b]	4040 (33.2%; 32.3–34%)
Replies—*N* (%; 95% CI)	664 (8.34%; 7.74–8.97%)	376 (8.91%; 8.07–9.81%)	0.286[b]	1040 (8.54%; 8.05–9.05%)
Retweets—*N* (%; 95% CI)	4802 (60.3%; 59.2–61.4%)	2,298 (54.5%; 52.9–56%)	<0.001[b]	7100 (58.3%; 57.4–59.2%)
Tweets before broadcast—*N* (%; 95% CI)	1119 (14.1%; 13.3–14.8%)	807 (19.1%; 17.9–20.3%)	<0.001[b]	1926 (15.8%; 15.2–16.5%)
Tweets during broadcast—*N* (%; 95% CI)	5293 (66.5%; 65.4–67.5%)	591 (14%; 13–15.1%)	<0.001[b]	5884 (48.3%; 47.4–49.2%)
Tweets after broadcast—*N* (%; 95% CI)	1548 (19.4%; 18.6–20.3%)	2822 (66.9%; 65.4–68.3%)	<0.001[b]	4370 (35.9%; 35–36.7%)
Positive tweets—*N* (%; 95% CI)	3712 (46.6%; 45.5–47.7%)	2367 (56.1%; 54.6–57.6%)	<0.001[b]	6079 (49.9%; 49–50.8%)
Negative tweets—*N* (%; 95% CI)	875 (11%; 10.3–11.7%)	340 (8.06%; 7.25–8.92%)	<0.001[b]	1215 (9.98%; 9.45–10.5%)
Neutral tweets—*N* (%; 95% CI)	3292 (41.4%; 40.3–42.4%)	1,254 (29.7%; 28.3–31.1%)	<0.001[b]	4546 (37.3%; 36.5–38.2%)
Question tweets—*N* (%; 95% CI)	58 (0.729%; 0.554–0.941%)	6 (0.142%; 0.0522–0.309%)	<0.001[b]	64 (0.525%; 0.405–0.671%)
Unclear tweets—*N* (%; 95% CI)	23 (0.289%; 0.183–0.433%)	253 (6%; 5.3–6.75%)	<0.001[b]	276 (2.27%; 2.01–2.55%)
Positive/negative ratio (95% CI)	4.24 (3.94, 4.57)[a]	6.96 (6.21, 7.82)[a]	<0.001[b]	5 (4.7, 5.33)[a]

a: Poisson regression.
b: Logistic regression.
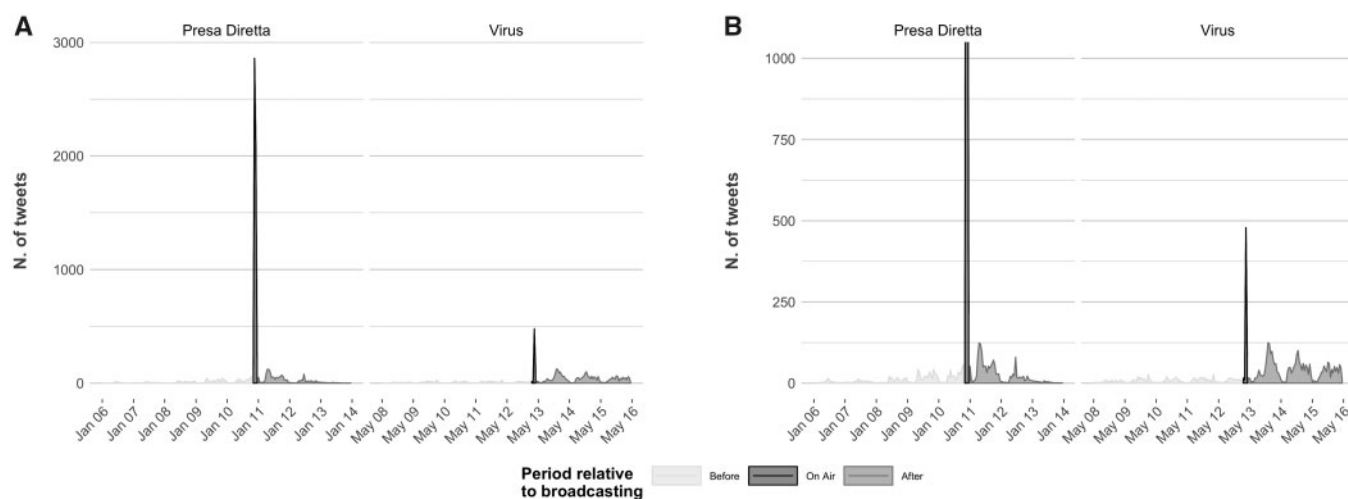c: Quasi-Poisson regression (cfr. methods).

Figure 1 Hourly trend of tweets in the time period relative to PD and VS. Panel A shows the trend in the original scale, Panel B uses a smaller scale, limited at 1000 tweets per hour

letters and included in at least 1% and in maximum 60% of tweets. For each term/document pair, a normalized term frequency—inverse document frequency score was computed. Such scores were used with a MARS logistic regression (three-degree interaction, linear predictors only),[16] to identify the set of terms which better predict a positive or negative sentiment. The resulting list was manually reviewed to identify the most relevant keywords associated with a strong vaccine sentiment.

Data management, analysis and plotting were performed using the R statistical language v.3.4.0.

Results

A total of 32 597 tweets were downloaded; 12 180 (37%) tweets were pertinent to human vaccinations and published by a total of 5447 users. Reported results refer to pertinent tweets only, which were manually reviewed by researchers. Table 1 reports a general description of the characteristics of users and tweets included in the analysis.

Tweets and users

Of the 12 180 pertinent tweets, 4040 (33.2%) were original tweets, 7100 (58.3%) were retweets and 1040 (8.6%) were replies. Retweet/tweet ratio was 1.76 (95% CI 1.69–1.83).

A total of 7960 tweets (65.4%) were published in the time period relative to PD (5293 during the show airing), by a total of 3183 users; 4220 tweets were published in the time period relative to VS (591 during the show airing), by a total of 2637 users.

A total of 276 users (5% of the total) tweeted during both broadcasts ('recurrent users'). Based on the number of tweets by user, users tweeting during VS were in general more active compared to those tweeting during PD (median: 6960 tweets by user for VS vs. 2990 for PD, *P* < 0.001).

Tweet trends

Figure 1A and B reports the tweet trend for each broadcast, on two different scales to allow comparison between the broadcasts.
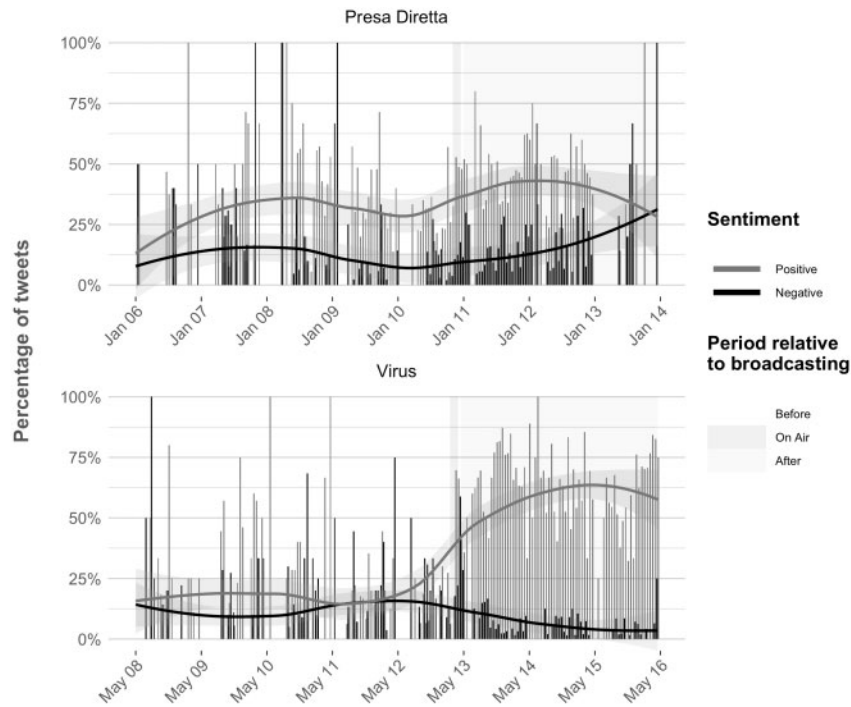
Figure 2 Proportion of tweets with positive and negative sentiment, by hour. The green and the red lines represent the overall trends of positive and negative tweets, respectively, computed by Loess regression

Table 2 Odds ratio and 95% CI for tweets' positive sentiment (vs. negative sentiment) by users' characteristics

	Presadiretta	Virus	Total
Recurrent (ref. 'No')	3.26 (2.22–4.79)	1.47 (0.951–2.27)	2.59 (1.93–3.46)
Number of statuses (tens of tweets)	1.12 (1.02–1.22)	1.25 (0.854–1.83)	1.06 (0.988–1.15)
% Tweets[a]	0.885 (0.862–0.909)	0.928 (0.895–0.963)	0.895 (0.876–0.914)
% Retweets[a]	1.21 (1.18–1.25)	1.18 (1.13–1.23)	1.2 (1.17–1.23)
% Replies[a]	0.84 (0.808–0.874)	0.844 (0.807–0.883)	0.845 (0.82–0.87)
% Tweets during broadcast[a]	1.03 (0.999–1.06)	0.865 (0.835–0.897)	0.935 (0.915–0.954)
% Tweets after broadcast[a]	0.974 (0.947–1)	1.16 (1.12–1.2)	1.07 (1.05–1.09)
User followers[b]	1.95 (1.72–2.21)	0.971 (0.802–1.18)	1.68 (1.52–1.86)
User tweets[b]	1.61 (1.44–1.79)	0.879 (0.732–1.05)	1.46 (1.33–1.59)
User friends[b]	2.08 (1.8–2.41)	1.14 (0.909–1.43)	1.83 (1.63–2.06)

a: The univariate analysis took into account change in odds for a 10% increase.
b: Change in odds for an unit increase on the log 10 scale.

For both broadcasts, we observe a high peak of tweets on vaccines during the show airing.

We also observe an increase in the number of tweets after the broadcast compared to tweets before the airing, with VS eliciting a higher production of tweets after the show compared to PD. The proportion of tweets published during the show airing was higher for PD (66.5 vs. 14%, $P > 0.001$), while the proportion of tweets published after the show airing was higher for VS (66.9 vs. 19.4%, $P < 0.001$). The number of tweets produced after PD airing decreases quickly in the following days, while for VS the decrease of interest is much slower.

Vaccine sentiment analysis

Vaccine sentiment was distributed as follows: 6079 (50.4%) positive, 4546 (37.7%) neutral, 1215 (10.1%) negative, 169 (1.4%) unclear and 64 (0.5%) questions. The proportion of positive tweets was higher during VS (56.1 vs. 46.6%, $P < 0.001$); negative and neutral tweets were more represented during PD (11 vs. 8.1%, $P < 0.001$; 41.4 vs. 29.7%, $P < 0.001$). The positive/negative ratio was higher for VS (6.96 vs. 4.24, $P < 0.001$) and the ratio increased after the VS airing. Figure 2 reports the trends of positive and negative tweets.

Table 2 reports the predictors of positive vs. negative tweets by user's profile characteristics.

Very active users tended to have a more positive sentiment: the number of followers (OR 1.68, $P < 0.001$), friends (OR 1.83, $P < 0.001$) and published tweets (OR 1.46, $P < 0.001$) were predictors of positive sentiment. The proportion of retweets (as opposed to original tweets and replies) was also a predictor of a positive sentiment (OR 1.2, $P < 0.001$). Finally, being a recurrent user was associated to a positive sentiment for users tweeting during PD (OR 3.26, $P < 0.001$).

Compared to non-recurrent users, recurrent users tweeted less frequently before and more frequently after the broadcast airing. Recurrent users' sentiment was significantly more positive, and the positive/negative tweet ratio was higher compared to non-recurrent users.

Associated keywords

We investigated keywords associated with a positive or negative sentiment. Among those associated with positive tweets (ordered by relevance): responsibility, disease, thanks, ignorance, science, medicine

and save. Among those associated with negative tweets: damage, pharmaceutical, doctor, mercury, baby, drug, law and oblige.

Discussion

We used Twitter to explore the interest in two broadcasts on vaccines aired on the Italian TV. We recorded an increase of vaccine-related tweets during the broadcast airing and in the following days, and we exploited this data to analyze the tweet trends, understand the characteristics of users involved in the conversation, study the vaccine sentiment and users' characteristics, identify recurrent keywords.

The information acquired through social media monitoring can be used to inform public health communication campaigns, as it can facilitate a better understanding of a specific though relevant sub-group of the population (social media users), both in terms of community structure and in terms of shared contents (interests and concerns). This kind of information can be useful to plan real-time responses to doubts and misconceptions emerging during the TV broadcasts, but also as a knowledge basis to plan—and monitor—vaccine promotion campaigns in the long term.

The frequency of vaccine-related tweets after both broadcasts was higher compared to tweets published before the broadcast airing, confirming that TV programmes can trigger an intense discourse among Twitter users. During the first half of 2017, an average of 5.4 M Italian users commented on Italian TV broadcasts through original posts/tweets or through engagement interactions on Facebook or Twitter.[17]

In our case, based on audience share data, a relatively low portion of the total broadcast public posted vaccine-related tweets in the time relative to the TV shows: 0.18% for PD and 0.21% for VS. Although this data are not representative of the whole public of the shows, it represents an interesting sample of users actively engaged in a social media discourse on vaccines and can therefore be informative for a campaign targeting social media users.

Through natural language processing (NLP) techniques, a more specific profiling of Twitter users (including age and sex) could be inferred from characteristics of tweets and users' accounts[18] and could be used to better tailor communication campaigns. In our case study, we did not specifically explore Twitter users' characteristics, but other researchers showed that 25–34-year-old males are the most active group commenting TV broadcasts (mainly sports programmes), though, in the case of talk-shows or political-themed broadcasts (like the two broadcasts we analyzed), 42% of the audience belongs to the >55 age class, mainly of female sex.[17]

We also show a generally positive sentiment in the majority of tweets and a high positive/negative ratio for both broadcasts, which is apparently not directly affected by the quality of the information presented. Both broadcasts were aired on prime time on the Italian public TV and were conducted by a presenter. PD reported evidence-based information on vaccines through previously recorded videos of experts in the field; VS was a talk show hosting a discussion between pro-vaccine persons and two personalities from the show business (an actress and a music critic) with anti-vaccine positions. The airing of VS elicited an intense discussion on social and traditional media for the time dedicated to anti-vax positions.

Despite the prevalently vaccine-hesitant issues emerging during VS, the positive/negative ratio increased after the airing of VS, suggesting that anti-vax messages in TV broadcasts do not necessarily elicit negative opinions on vaccines on Twitter. This confirms the observation that the effect of communication on vaccines and other health issues is not straightforward. Fair and balanced information on vaccines can generate a hostile media perception.[19] A corrective communication intervention on measles, mumps and rubella vaccination, based on institutional documents, reduced the intent to vaccinate among vaccine sceptical parents.[20] Health promotion through fear can generate a sense of constriction that can reduce the adoption of the promoted behaviours.[21] A 'one size fits all' approach is not applicable to vaccine communication for preventing vaccine hesitancy.[8] Therefore, the determinants of the effect of different vaccine communication initiatives on vaccine confidence should be investigated at the local level and in specific population sub-groups, as such effect is likely affected by specific characteristics of both the message and the target audience. This can also be obtained by analyzing specific population groups on social media, both prior to the creation of a vaccine promotion campaign and during the monitoring phase.

Users tweeting during both broadcasts had a more positive sentiment towards vaccines compared to non-recurrent users and the number of followers, friends and published tweets were predictors of a positive sentiment. This observation suggests that active users could spread information on vaccines to a large target audience. Other authors have previously studied the network of users tweeting about vaccines, identifying main influencers of the network both with pro-vax and anti-vax positions.[22] Such analysis should be performed at a country-level to identify network's nodes that could more effectively distribute positive vaccine information, or to monitor those nodes that could represent a risk in case they should decide to promote anti-vaccine contents.

In our study, we opted for a manual classification of tweets' sentiment. Another option for speeding up the classification process is to exploit automatic sentiment classification algorithms. Automatic sentiment recognition in social posts has been widely investigated, with vaccination being one of the most studied topics.[23–25] Moreover, there are a number of available platforms for NLP and machine-learning-powered social network surveillance, specifically oriented to vaccine conversations, or easily customizable for health topics.[26,27] It would be relevant for health institution to adopt vaccine intelligence platforms to automatically spot burst of anti-vaccine messages to promptly organize targeted information campaigns.

Finally, we analyzed the content of tweets through NLP techniques to automatically identify relevant keywords for tweets with positive/negative sentiment. Studying these keywords could guide the selection of contents and enhance the communication about vaccines through new media. The use of mixed methods and qualitative analysis could offer an effective framework for acquiring deeper insights from social media post,[28] and for translating data acquired through social media listening into data-based communication.[29]

This study has several limitations. First, we classified tweets only based on the tweets' text. Links to external resources were not evaluated. This could have biased our estimation of tweets' sentiment. On the other hand, the sentiment classification has been performed manually by two researchers; therefore, the attribution of sentiment is characterized by a high level of quality and precision. Another limit concerns the fact that we have no baseline data on the proportion of polarized tweets in Italy, to be used as a reference for a better evaluation of tweets published during the broadcasts. Another important limitation is that we did not try to identify bots that could have a significant impact on online communications in broad events limited in time and on the vaccination topic.[30]

In conclusion, we show that discussion on social media in reaction to public events is rich and could be used to inform communication initiatives for vaccine promotion. Public health agencies should take into account the adoption of a system for monitoring the public's sentiment towards vaccines on social media. Such a system is feasible, would require limited resources if automated by NLP techniques, and could offer the possibility of continuously assessing vaccine sentiment of social media users and provide a prompt assessment of (and possibly a contrast to) misinformation.

Acknowledgements

Funding

Key points

- Twitter users were highly involved in reaction to Italian TV broadcasts dedicated to vaccines.
- The display of anti-vax messages on traditional media can elicit Twitter reactions characterized by a positive vaccine sentiment.
- Monitoring social media activity during TV shows dedicated to vaccines can provide insights for informing vaccine communication initiatives by public health institutions.

References

1   World Health Organization. Available at: https://www.who.int/news-room/fact-sheets/detail/immunization-coverage (10 December 2019, date last accessed).

2   D'Ancona F, D'Amario C, Maraglino F, et al. Introduction of new and reinforcement of existing compulsory vaccinations in Italy: first evaluation of the impact on vaccination coverage in 2017. *Euro Surveill* 2018;23:1800238.

3   Signorelli C, Odone A, Cella P, et al. Infant immunization coverage in Italy (2000-2016). *Ann Ist Super Sanita* 2017;53:231–7.

4   Filia A, Bella A, Del Manso M, et al. Ongoing outbreak with well over 4,000 measles cases in Italy from January to end August 2017—what is making elimination so difficult? *Euro Surveill* 2017;22:30614.

5   State of Vaccine Confidence in the EU 2018. Available at: https://ec.europa.eu/health/sites/health/files/vaccination/docs/2018_vaccine_confidence_en.pdf (10 December 2019, date last accessed).

6   Larson HJ, Smith DMD, Paterson P, et al. Measuring vaccine confidence: analysis of data obtained by a media surveillance system used to analyse public concerns about vaccines. *Lancet Infect Dis* 2013;13:606–13.

7   Larson HJ, Clarke RM, Jarrett C, et al. Measuring trust in vaccination: a systematic review. *Hum Vaccin Immunother* 2018;14:1599–1609.

8   Jarrett C, Wilson R, O'Leary M, et al. Strategies for addressing vaccine hesitancy - a systematic review. *Vaccine* 2015;33:4180–90.

9   Nowak GJ, Gellin BG, MacDonald NE, et al. Addressing vaccine hesitancy: the potential value of commercial and social marketing principles and practices. *Vaccine* 2015;33:4204–11.

10 González ER. TV report on DTP galvanizes US pediatricians. *JAMA* 1982;248:12–22.

11 Suppli CH, Hansen ND, Rasmussen M, et al. Decline in HPV-vaccination uptake in Denmark—the association between HPV-related media coverage and HPV-vaccination. *BMC Public Health* 2018;18:1360.

12 Developer Twitter. Available at: https://developer.twitter.com/en/docs/tweets/search/overview/standard (16 December 2019, date last accessed).

13 Cleveland WS, Grosse E, Shyu WM. In: JM Chambers, TJ Hastie, editors. *Statistical Models in S*. 1992.

14 Gelman A, Jakulin A, Pittau MG, Su Y-S. A weakly informative default prior distribution for logistic and other regression models. *Ann Appl Stat* 2008;2:1360–83.

15 Màrquez L, Rodríguez H. Part-of-speech tagging using decision trees. In: C Nédellec, C Rouveirol, editors. *Machine Learning: ECML-98*. Berlin, Heidelberg: Springer, 1998: 25–36.

16 Friedman J. Multi-variate adaptive regression splines (with discussion). *Ann Stat* 1991;19.

17 Nielsen Social. Available at: https://www.nielsen.com/it/it/press-releases/2017/young-adults-are-most-active (12 December 2019, date last accessed).

18 Wang Z, Hale S, Adelani D, et al. Demographic inference and representative population estimates from multilingual social media data. 2019; 2056–67.

19 Gunther AC, Edgerly S, Akin H, Broesch JA. Partisan evaluation of partisan information. *Commun Res* 2012;39:439–57.

20 Pediatrics. Available at: https://pediatrics.aappublications.org/content/133/4/e835?sso=1&sso_redirect_count=1&nfstatus=401&nftoken=00000000-0000-0000-0000-000000000000&nfstatusdescription=ERROR%3a+No+local+token (12 December 2019, date last accessed).

21 Erceg-Hurn DM, Steed LG. Does exposure to cigarette health warnings elicit psychological reactance in smokers? *J Appl Soc Psychol* 2011;41:219–37.

22 Sanawi JB, Samani MC, Taibi M. Vaccination: identifying influencers in the vaccination discussion on Twitter through social network visualisation. *Bus Soc* 2017;18:718–26.

23 Gohil S, Vuik S, Darzi A. Sentiment analysis of health care tweets: review of the methods used. *JMIR Public Health Surveill* 2018;4:e43.

24 Mamidi R, Miller M, Banerjee T, et al. Identifying key topics bearing negative sentiment on Twitter: insights concerning the 2015-2016 Zika epidemic. *JMIR Public Health Surveill* 2019;5:e11036.

25 Rosselli R, Martini M, Bragazzi NL. The old and the new: vaccine hesitancy in the era of the Web 2.0. Challenges and opportunities. *J Prev Med Hyg* 2016;57:E47–50.

26 Bahk CY, Cumming M, Paushter L, et al. Publicly available online tool facilitates real-time monitoring of vaccine conversations and sentiments. *Health Aff (Millwood)* 2016;35:341–7.

27 Müller MM, Salathé M. Crowdbreaks: tracking health trends using public social media data and crowdsourcing. *Front Public Health* 2019;7:81.

28 Andreotta M, Nugroho R, Hurlstone MJ, et al. Analyzing social media data: a mixed-methods framework combining computational and qualitative text analysis. *Behav Res Methods* 2019;51:1766–81.

29 Bahri P, Fogd J, Morales D, et al.; ADVANCE Consortium. Application of real-time global media monitoring and 'derived questions' for enhancing communication by regulatory bodies: the case of human papillomavirus vaccines. *BMC Med* 2017;15:91.

30 Stella M, Cristoforetti M, Domenico MD. Influence of augmented humans in online interactions during voting events. *PLoS One* 2019;14:e0214210.