

RESEARCH ARTICLE

Open Access

# Using text mining and sentiment analysis to analyse YouTube Italian videos concerning vaccination



Annamaria Porreca<sup>1\*</sup> , Francesca Scozzari<sup>1</sup> and Marta Di Nicola<sup>2</sup>

## Abstract

**Background:** Recently, social networks have become a popular source of information on health topics. Particularly, in Italy, there is a lively discussion on the web regarding vaccines also because there is low vaccination coverage, vaccines hesitancy, and anti-vaccine movements. For these reasons, in 2017, Institutions have introduced a law to force children to make ten compulsory vaccines for school attendance and proposed a vaccination campaign. On social networks, this law has fostered a fierce discussion between pro-vaccinations and anti-vaccinations people. This paper aims to understand if and how the population's opinion has changed before the law and after the vaccination campaign using the titles of the videos uploaded on Youtube in these periods.

**Method:** Using co-occurrence network (CON) and sentiment analysis, we analysed the topics of YouTube Italian videos on vaccines in 2017 and 2018.

**Results:** The CON confirms that vaccinations were very disapproved before the law. Instead, after the communication campaign, people start to be less critical. The sentiment analysis shows that the intense vaccination campaign also promoted by medical doctors pushed the sentiment to change polarity from a prevailing negative opinion in 2017 (52% negative) to a positive one in 2018 (54% positive).

**Conclusion:** At the population level, the potential misinformation of social networks could be significant and is a real risk for health. Our study highlights that vaccination campaigns on social networks could be an essential instrument of health policies and a sharp weapon to fight ignorance and misrepresentations of non-qualified people influencing individuals' decision-making.

**Keywords:** Vaccinations, YouTube, Sentiment analysis, Pro-vax, No-vax

## Background

The vaccine coverage provides information on the efficiency of the vaccination system and the trust of the vaccination practice. Currently, Europe faces many challenges, including the spreading antivaccination sentiment [1]. In many countries, there are people refusing vaccinations for themselves and their children, and promoting the idea that immunization is a way to do business. Moreover, there are also other reasons why vaccination practises are not shared by everyone [2–4]. For example, from 2016 onwards, in the USA and

Europe, many antivaccination movements were pushed by the “*Vaxxed: From Cover-Up to Catastrophe*”. The latter is an American pseudoscience documentary film directed by Andrew Jeremy Wakefield, who was a British medical supporting the causal relationship between the trivalent MPR vaccine (measles, mumps, rubella) and the appearance of autism and intestinal disease. This theory stimulated people to lose their belief in vaccines, and thus it contributed to lead to the so-called “*vaccine hesitancy phenomenon*” causing a vaccination coverage reduction. The vaccines hesitancy has created health alarmism by pushing countries and international organizations, such as the World Health Organization [5, 6], to look for a way to contain this phenomenon and promote the safety of vaccines. Among the countries of the

\* Correspondence: [annamaria.porreca@unich.it](mailto:annamaria.porreca@unich.it)

<sup>1</sup>Department of Economic Studies, G. d'Annunzio University of Chieti-Pescara, Viale della Pineta 4, Pescara, Italy

Full list of author information is available at the end of the article



European Union, in Italy, from 2013 onwards, vaccination coverage has been showing a decreasing trend. Therefore, in 2017, a measles epidemic caused 5000 cases and four deaths [7]. For this reason, on the 7th of June 2017, the Italian Institutions approved the Lorenzin' decree-law, which increased the number of compulsory vaccinations from four to ten [8]. However, this legislative manoeuvre obtained dissent from a slice of the population divided into two sides: the anti-vaccine people (encompassing who supports freedom of choice in vaccination and thus promotes the non-compulsoriness of vaccinations) and pro-vaccination people who are active in promoting positive campaigns. Both anti-vaccine and pro-vaccine people use the web to share their opinions. Today, many people find information on the World Wide Web (WWW); indeed, recent studies showed that more than half of the population has access to the Internet, and most web users seek health information there (e.g [9, 10]). Moreover, health information is more and more available on the Internet [11, 12]. There are many professional health web sites, app or social network groups to support and inform population but sometimes the quality of health information on the WWW could be a real danger for people [13, 14]. Social networks permit individuals to quickly create, share, and retrieve information, and allow anyone to give opinions and spread their message (also to anti-vaccination activists). Social networks thus contribute to false myths and misinformation about vaccines with a consequent negative impact on people's willingness to be vaccinated [4, 15].

The aim of the paper is to analyze the online sentiment to understand if the Italian legislative intervention on vaccination is shared by the population and which are the arguments connected with this topic when people share information on YouTube. Indeed, because monitoring social networks could be a good proxy to evaluate vaccine hesitancy, understanding people thinking can guide policy-makers to plan effective information campaigns able to change people mindset regarding vaccinations.

Specifically, we focused on Italy because "*Among the European countries, Italy has one of the highest levels of scepticism related to effectiveness and safety of vaccines*" [16–18]. Moreover, to solve this problem, the Italian Institutions approved the Lorenzin' decree-law, and thus we are also interested in understanding the effects of this manoeuvre on the sentiment of the population.

According to the "*We are Social - Digital in 2018 Report*", the internet users in the world exceed four billion, and in Italy, they are more than 43 million; this means that today, more than half of the world's population is online. In the world, Facebook has much greater penetration than YouTube [19] whereas, in Italy, YouTube dominates the video-based social media platforms. It

was on YouTube that the North American anti-vaccine movement was able to upload and share conference recordings free of charge to a wider audience. Compared to the past, when tapes of proceeding had to be ordered by mail, YouTube allowed autism awareness groups to host instant discussions, comments, and information-sharing easily [20–24].

The main contribution of this research to the existing literature on the problem of vaccination hesitancy is to propose the use of text mining and sentiment analysis. Indeed, to educate people increasingly inclined to use WWW as a source of information and to understand their mindset, policy-maker need appropriate tools capable of dealing with the new digital age. Effectively, these instruments could be adopted by public authorities to understand what people's concerns are about health, what opinion the population has regarding the health policies implemented by the institutions, and finally to know if the policies and campaigns implemented have had the expected results.

This research uses a networked mindset to tackle vaccine hesitancy. Specifically, we aim to understand what the sentiment towards vaccines on YouTube is, and finally to detect if and how the Italians' opinion has changed from the period before the introduction of the decree-law to the following period, i.e. after the vaccination campaign.

## Methods

We used Netvizz to collect YouTube videos [25]. Netvizz is a tool for extracting data from the YouTube platform via the YouTube API v3 [26]. The search query is "*vaccin*". The search for the query took place concerning the period from the first of May to the first of October for both the year 2017 and 2018. The iterations used to search the query were chosen equal to ten (fifty items for every iteration) and has been specified to search for each day of the timeframe. The tool provided 3777 videos for 2017 (1898 were adopted for the analysis) and 3788 for 2018 (822 were used for the analysis). The exclusion criteria were a language different from the Italian one; titles not containing the subject under investigation; and finally videos with ambiguous titles were not taken into account.

We highlight that Netvizz also provided the following variables about videos: number of dislikes and likes, duration, comment count, and view count. We used Google Trends to determine the period in which to focus the search of the query and which was the sought-after query in terms of vaccines.

We considered the maximum and minimum of the query's functions of each year to capture the time-variability of the queries before and after the vaccination campaign. This criterion led to identifying 6 months in

2017 (from May to October) while, for 2018, a two-months range (from June to July). To guarantee the comparability of the two periods and considering that the June–July period lies within the May–October interval, we refer to the May–October periods both in 2017 and 2018.

All relevant material for Italians is available only with the local language; in truth, even on the website of the Ministry of Health in Italy, it is quite difficult to find material in a language other than Italian! Indeed, to date, regarding vaccinations, the English version is still under construction [27]. Therefore, the choice of the query was quite trivial because we compared the query's functions on Google Trends for the following words: “*vaccino*” (in English “*vaccine*”), “*vaccini*” (in English “*vaccines*”), “*vaccinazioni*” (in English “*vaccinations*”), “*morbilli*” (in English “*measles*”), and “*autismo vaccini*” (in English “*autism vaccines*”). Because both “*vaccino*” and “*vaccini*” were much more searched on the web than the others, we used the query “*vaccin*” that allowed us to search for both the singular and the plural in the Italian language [28].”

To understand the main topic about vaccines on YouTube, we analyzed the data by using the co-occurrence network (CON) based on Freeman's betweenness centrality Index by KH Coder 3.0 [29, 30]. Betweenness centrality is the measure of the degree to which one node is “between” other nodes in a network; this means that this node could act as mediators on the network, regardless of the frequency of connectivity [29]. After the user fixes some coding rules, CON takes into account co-occurrence relations among terms. The degree of co-occurrence was determined using the Jaccard similarity coefficient [31]. The value of the latter index is a scalar and ranges from 0, indicating no spatial overlap between two sets of binary segmentation results, to 1, indicating complete overlap. To conduct the CON, we used lexicon coding rules to exclude logic operators such as “*e*”, “*o*”, “*non*”, demonstrative adjective, people names, other language words and meaningless words. To analyze the sentiment during the considered periods about the content of the co-occurrence networks, we used the function “*get\_nrc\_sentiment*” based on the NRC sentiment dictionary to calculate the presence of eight different emotions and their corresponding valence in a text file from the “*syuzhet*” R environment package [32]. The words in each sentence are compared to a dictionary and tags positive or negative words by +1 or -1, respectively. Before testing differences into emotions between the two periods, we rescaled the sentiment vectors to the normalized axis, and we got re-sampling vectors for comparing them. T-tests were performed to test differences in the mean between the year's sentiment and video characteristics: duration (min), view

count, number of likes, number of dislike and comment count. Statistical significance was set at  $p < 0.05$ . For each variable in the study, we have included descriptive statistics.

## Results

Table 1 shows the descriptive statistics of the YouTube videos in Italian language containing a vaccination theme in the year 2017 and 2018. On average, the duration of the videos and the number of views decrease from 2017 to 2018. It is interesting to note that the number of likes, on average, is higher in 2018 than in 2017. The same is for the number of comments, and thus 2018 videos have reached greater popularity.

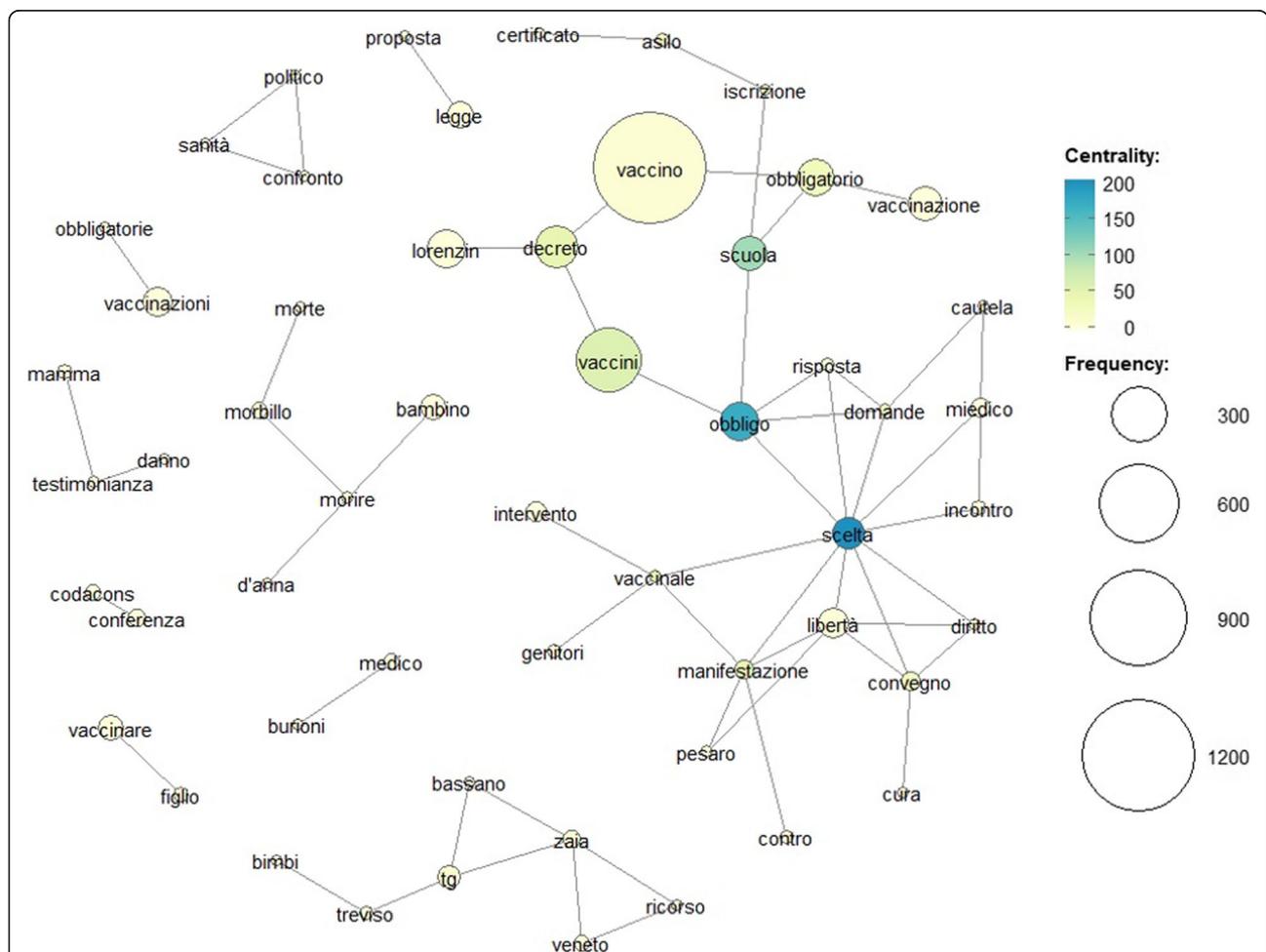
Figures 1 and 2 show the textual analysis through the CON in 2017 and 2018, respectively. Figure 1 shows that, in 2017, the terms “*compulsory*” and “*choice*” have high degrees of centrality within the network. But also the terms “*school*” and “*decree*” have a high centrality within the network although their frequency is low. Particularly interesting is the fact that the term “*choice*” is connected and many other terms pertaining to the subject such as “*vaccination*”, “*freedom*”, “*law*”, “*meeting*”, “*doctor*”, “*questions*”, “*compulsory*”, “*manifestation*”. Moreover, we observe many connected cliques and small clusters stressing strong concern and perplexities. In summary, the graph underlines that there is an interesting connection between vaccinations and freedom of choice.

Figure 2 shows that, in 2018, the most central words in the network in order of importance are: “*medical doctor*”, “*child*”, “*deadline*”, “*vaccination*”, “*ministry*”, “*health*”, and “*parent*”. The betweenness centrality index of these words does not coincide with the frequency of appearance in the text, this means that although they do not appear very frequently in the speech, they connect most of the themes acting as bridges to reach the most isolated nodes. The word “*ministry*” is also the node a high betweenness centrality being the one most connected in the network having a high degree centrality. It can be noticed how taking as a “*ministry*” centre this constitutes a centre for the formation of a connected cluster that includes topics concerning the Lorenzin decree, conferences, health, regulation at the regional and scholastic level. Interesting is the change in the importance of terms such as “*freedom*”, “*choice*”, “*risk*”, and “*compulsory*”; hence, it suggests that the previous negative discussions about vaccinations are weakened, bringing the speech back to less polemical tones.

Figure 3 shows the barplot of the emotions' polarity in 2017 and 2018. In 2017, the prevailing sentiment was the negative one (52%), whereas, in 2018, the positive sentiment was 54%.

**Table 1** Descriptive statistics of YouTube videos about vaccines in the Italian language in 2017 and 2018, respectively and *p*-value of t.test comparisons

Variable	2017 (N = 1898)		2018 (N = 822)		<i>p</i> -value
	Number of videos	Mean ± SD	Number of videos	Mean ± SD	
Duration (min)	1885	503.40 ± 693.70	811	404.66 ± 586.18	0.278
View count	1891	2621.68 ± 19,243.40	820	1599.40 ± 9233.13	0.633
Number of Like	1799	69.24 ± 622.43	787	85.27 ± 598.93	0.853
Number of Dislike	1799	7.21 ± 31.38	786	6.21 ± 37.04	0.837
Comment count	1720	17.32 ± 80.89	822	18.41 ± 106.30	0.935



**Fig. 1** Co-occurrence Network (CON) based on Betweenness centrality index for Italian Youtube videos concerning vaccines in 2017 (from May to October). Nodes corresponding text in english is: "zaia" = Zaia (italian politician), "veneto" = Veneto (italian region), "vaccino" = vaccine, "vaccinazione, vaccinale" = vaccinations, "treviso" = Treviso (Italian city), "tg" = newscast, "testimonianza" = evidence, "scuola" = school, "scelta" = choice, "sanità" = "risposta" = answer, "ricorso" = contestation, "proposta" = motion, "politico" = politician, "pesaro" = Pesaro (italian city where took place no-vax parade), "obbligo, obbligatorio" = compulsory, "morte" = death, "morire" = die, "morbillo" = measles, "medico" = Miedico, "manifestazione" = manifestation, "libertà" = freedom, "iscrizione" = enrolment, "intervento" = intervention, "incontro" = meeting, "genitori" = parents, "figlio" = son, "domande" = questions, "diritto" = right, "dannno" = damage, "d'anna" = D'anna (senator against vaccines), "cura" = care, "contro" = against, "confronto" = confrontation, "conferenza" = conference, "codacons" = Codacons (italian non-profit association for the protection of consumers and the environment), "certificato" = certified, "cautela" = prudence, "burioni" = Burioni (Italian medical doctor pro-vax), "bimbi" = kids, "bassano" = Bassano (italian city), "asilo" = kindergarten

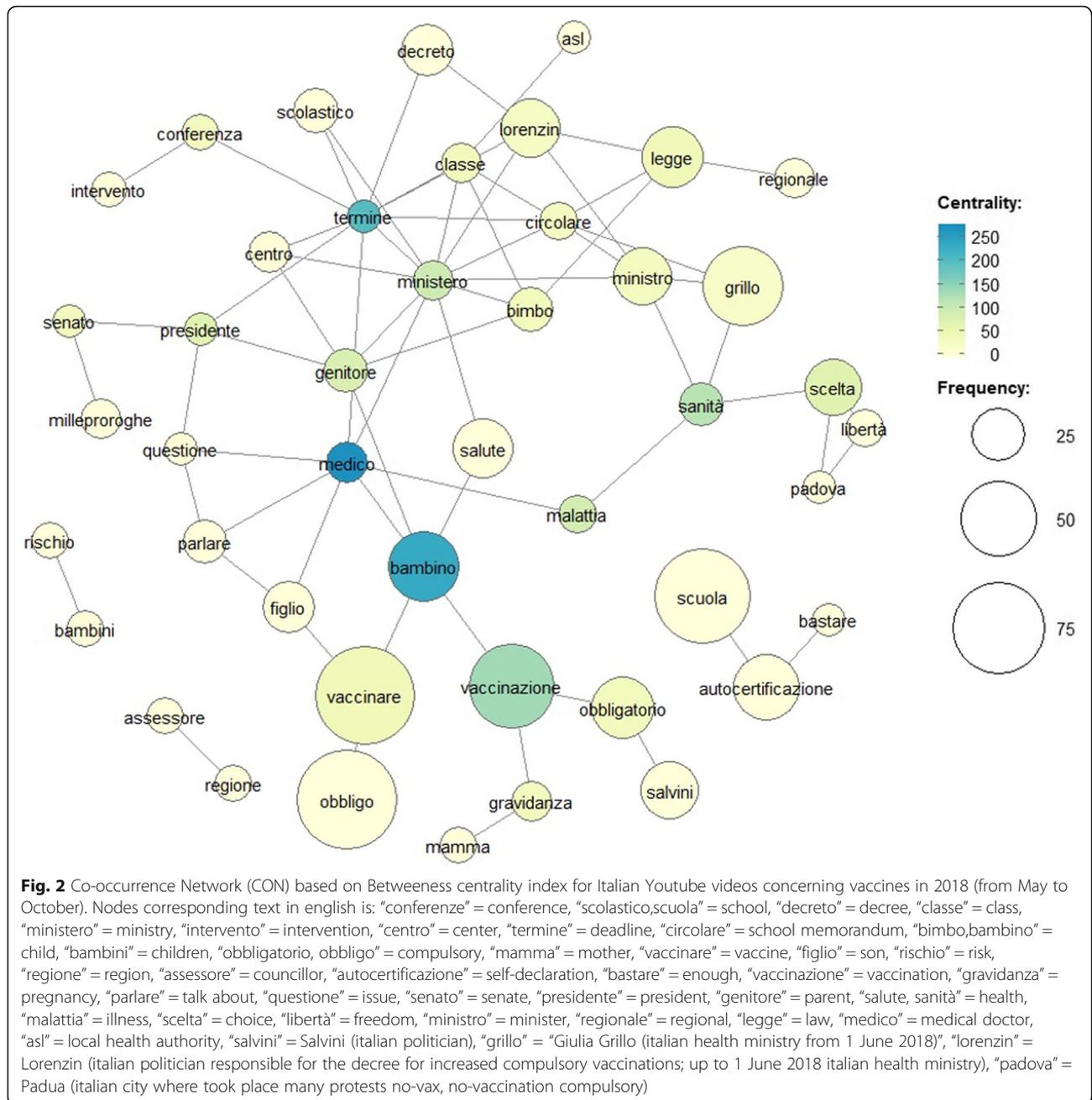
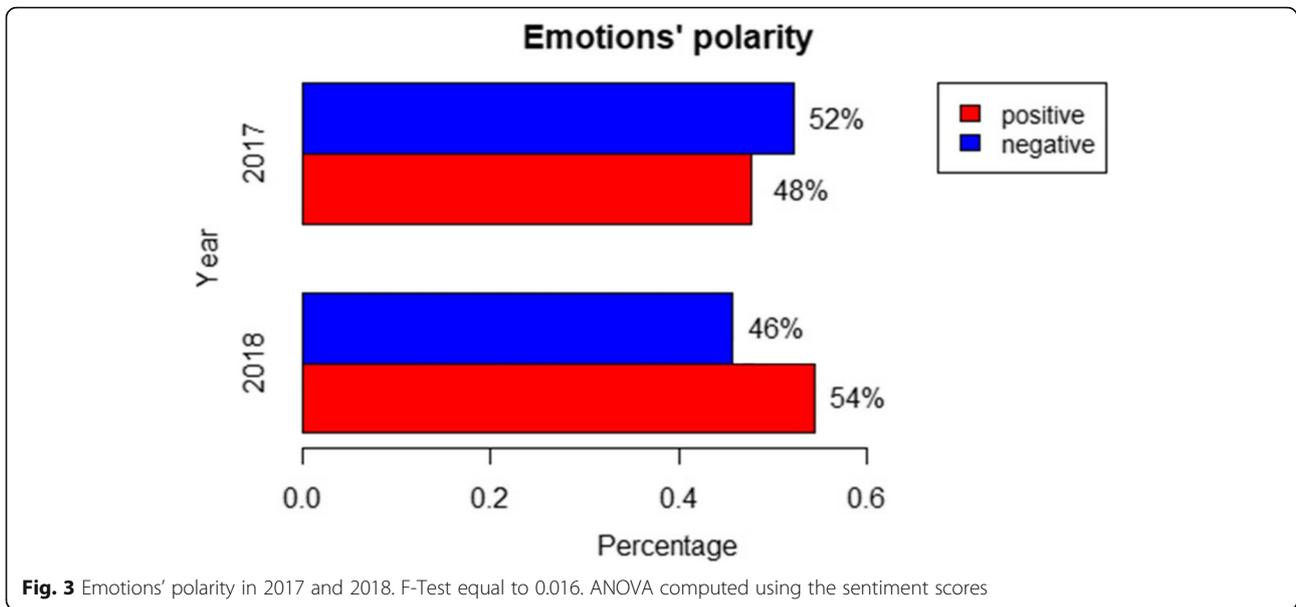


Figure 4 shows sentiments' change over time from 2017 (mean = 0.46 and SD = 0.29) to 2018 (mean = 0.53 and SD = 0.30) assessed in the different sixth-months periods. T-test confirms a statistically significant change ( $p = 0.02$ ).

**Discussion**

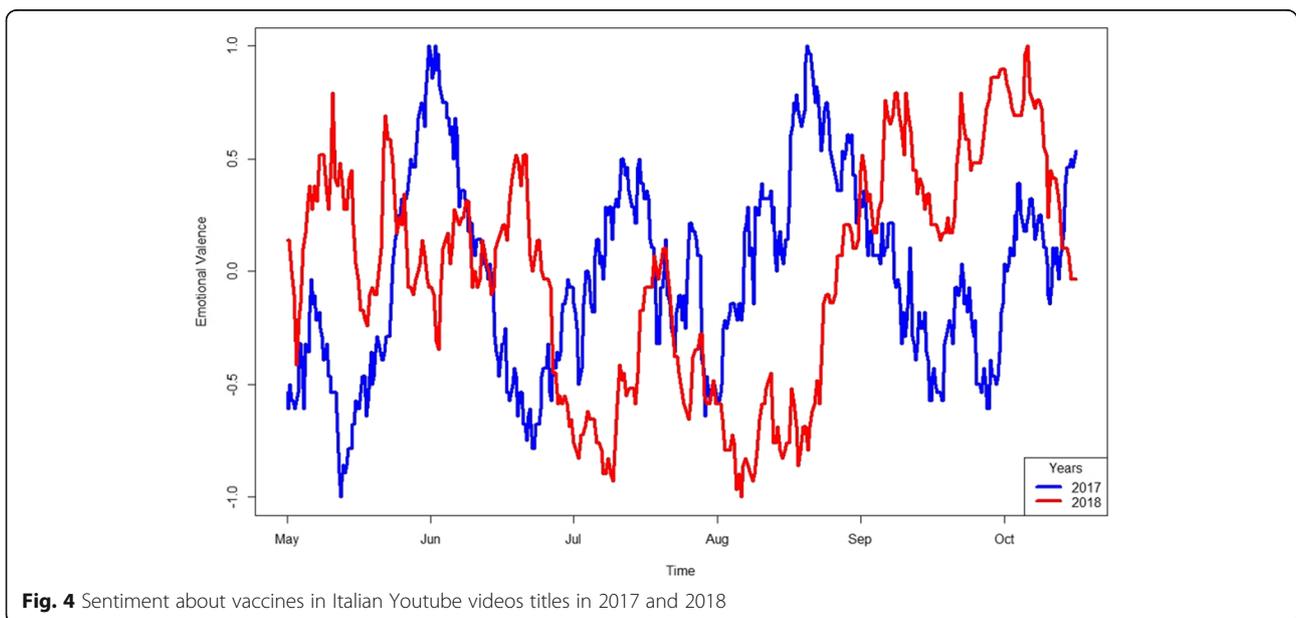
In Italy, the vaccination coverage strongly increased from the 31st of December 2017 to the first 6 months of 2018 [7]. Indeed, the minimum threshold (95% vaccination coverage) recommended by the WHO was reached

[6]. This result can be extrapolated by a medium-term survey carried out to assess the impact of the law on compulsory vaccinations in the cohorts of children born in 2015, 2014 and 2010 [7]. The Lorenzin's decree has undoubtedly pushed up the number of compulsory vaccines; however, it would be interesting to understand if the communication campaigns and commitment by the experts on the web had some effects on the opinion regarding the vaccination practice and their possible risks. In recent years, as Italian National Institute of Statistics (ISTAT) official data confirm, the number of Internet



users in Italy has increased [33–35]. In 2017, the Italians represented 7.8% of European internet users [36]. Due to the growth in the number of people who use the WWW to look for health-related information, public health policy-makers should consider the consequences of such use at the population level. For this reason, the analysis of web search trends and social network data could represent an exciting proxy for vaccine hesitancy [37, 38]. In 2017, YouTube was identified as the most adopted social network in Italy [6]. Nevertheless, few studies have analysed how people exploit this social network for health purposes and to make decisions on their future behaviour. Indeed, these platforms and their contents

have the potential to influence people health behaviours [20, 39], and thus the influence of vaccine-critical content and increase in using these platforms to share health information is a cause for concern for public health authorities. As also stressed by MacDonald, poor or inappropriate communications can lower the vaccination coverage and contribute to the hesitation of vaccination [40]. However, the quality of vaccine-related material available online is varied, and internet search engines often bring people to anti-vaccine or low-quality web information; therefore, the potential for misinformation is relevant [41, 42]. Generally, who use a search engine is likely to focus on the potential risks of the



vaccines and thus find anti-vaccination websites. Therefore, widely spread bad contents may contribute to reduce vaccine uptakes and increase the risk of preventable diseases. Hence, at the national level, better communication in the field of vaccination is desirable to counteract the phenomenon of vaccine hesitancy, including the development of digital tools to facilitate progress towards empowerment and changes in citizens' behaviour. For these reasons, it is vital to stimulate communication campaigns to encourage people healthy lifestyles and develop materials for health education. In fact, communication plays a crucial role in vaccination hesitancy, and thus it is necessary to identify the main communication channels to use resources and strategies that can influence public opinion [43]. To the best of our knowledge, this is the first study that tries to assess the contents of YouTube videos and their associated sentiment on compulsory vaccination using text mining and sentiment analysis tools. This paper shows that, after the vaccination campaigns and promotion, and also after health ministry change (from 1 June 2018 Giulia Grillo is the Italian health ministry), the topics and sentiment about vaccines on YouTube have changed from 2017 to 2018 (see also Figs. 1 and 2). The vaccines arguments in 2017 on YouTube in Italy were very criticisms about their obligatoriness. Most titles expressed dissents and mostly negative sentiment in the linked comments. In 2018, the arguments became less harsh towards vaccination and Lorenzin's decree, and there was even an inversion of polarity in the sentiment.

The final aim of this paper is to stimulate public authorities to invest in the digital communication for spreading health information, raising awareness among increasingly younger users, and reaching a type of communication that is increasingly popular. Indeed, this kind of information is essential to educate people increasingly inclined to use WWW as a source of information. The competent authorities should invest and apply policies for correct information on the web by financing projects and qualified operators capable of coping with the new digital age.

### Limitations

Although a considerable number of videos were selected for analysis, availability and access to API, and time limitations made it unfeasible to examine all videos. For these reasons, we have conducted an analysis completely excluding human judgment. Despite these limitations, to the best of our knowledge, this is the first study to examine YouTube vaccination's contents in Italy using a text mining and sentiment analysis procedure not affected by a human classification which implies an intense subjectivity in the attribution of sentiment scores such as to cause a strong bias in the study.

### Conclusions

All European countries have vaccination policies and, in particular, Italy is the first European country to approach comprehensively the growing vaccine hesitancy. Following the disinformation about vaccines, vaccination coverage in Italy has fallen below the critical threshold of 95% for various diseases. This situation can have serious consequences, both for the health of unvaccinated children and for those around them. The issue of compulsory vaccinations introduced by the Lorenzin's decree has aroused much discussion within the WWW. Social networks express opinions from both experienced and non-experts and reflect the feeling related to the subject of vaccination in Italy. Although in 2017 vaccinations have been highly criticized regarding their safety and ethics, and thus recorded a negative opinion by the Italian YouTube users, in 2018, there is an inversion of polarity in the general sentiment. Indeed, in 2018, users discussed this topic less critically, no longer emphasising the risk and danger of the vaccine but trying to understand and address the consequences of the law-decree regarding the information of the documentation required for access to school.

The ongoing debate about the adoption of mandatory vaccination policies, especially in the light of the fragmentation of the vaccination regulatory framework still challenges the scientific community, the political stakeholders and the general population in Italy. Because the pro-vaccine campaigns are designed to raise awareness among parents about the importance of vaccinating their children, policymakers should adopt the appropriate campaigns according to the society they live in. Lawmakers and public health authorities should play a fundamental role in informing the public and should use the media to do this so people can make informed choices. But politics are primarily influenced by what appears in 'old' media, such as television, radio and newspapers, and 'new' media such as social networks and the Internet are not exploited as necessary. Besides, the link between the media and politics can have a notable influence on vaccination coverage, and thus policymakers should use the 'old' and 'new' media for enhancing health-related information that can strengthen vaccination coverage.

This study shows how the public authority must take into account the WWW and its potential dangers, adopting communication strategies to manage as much as possible the information from social networks by contrasting fake news.

### Abbreviations

CON: Co-occurrence network; ISTAT: Italian National Institute of Statistics; WHO: World Health Organization; WWW: World Wide Web

### Acknowledgements

Not applicable.

**Authors' contributions**

All authors read and approved the final manuscript. AP was the lead author and undertook the statistical analysis. FS supervised the paper and gave advice on paper structure. MD advised on the research approach and statistical analysis and provides methodological advice and interpretation of findings.

**Authors' information**

Annamaria Porreca (AP) is a PhD Student in Computer Science at the Department of Economic Studies, G. d'Annunzio University of Chieti-Pescara, Viale della Pineta 4, Pescara, Italy.  
 Francesca Scozzari (FS) is Associate Professor at the Department of Economic Studies, G. d'Annunzio University of Chieti-Pescara, Viale della Pineta 4, Pescara, Italy.  
 Marta Di Nicola (MD) is Associate Professor at the Department of Medical Oral Science and Biotechnology, G. d'Annunzio University of Chieti-Pescara Chieti, Via dei Vestini 31, Chieti, Italy.

**Funding**

Not applicable.

**Availability of data and materials**

Data are available on Netwizz tool (<https://tools.digitalmethods.net/netwizz/youtube/>) setting it on five iterations of data and Italian language, using as research query "vaccin".

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Department of Economic Studies, G. d'Annunzio University of Chieti-Pescara, Viale della Pineta 4, Pescara, Italy. <sup>2</sup>Department of Medical Oral Science and Biotechnology, G. d'Annunzio University of Chieti-Pescara Chieti, Via dei Vestini 31, Chieti, Italy.

Received: 12 September 2019 Accepted: 10 February 2020

Published online: 19 February 2020

**References**

- Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res*. 2006;8(2):e9.
- Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. *BMC Public Health*. 2007;7(1):53.
- Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW. Patients' use of the internet for medical information. *J Gen Intern Med*. 2002;17(3):180–5.
- Wangberg S, Andreassen H, Kummervold P, Wynn R, Sørensen T. Use of the internet for health purposes: trends in Norway 2000–2010. *Scand J Caring Sci*. 2009;23(4):691–6.
- WHO (World Health Organization). (2014). [https://www.who.int/immunization/newsroom/press/immunization\\_coverage\\_july2014/en/](https://www.who.int/immunization/newsroom/press/immunization_coverage_july2014/en/). Accessed 22 July 2019.
- WHO (World Health Organization). (2018). SAGE issues its 2018 assessment report of the Global Vaccine Action Plan. <http://www.who.int/immunization/sage/sagewg/vaccinehesitancyapr12/en>. Accessed 22 July 2019.
- Ministero della Salute. (2017). [http://www.salute.gov.it/portale/news/p3\\_2\\_1\\_1\\_1.jsp?lingua=italiano&menu=notizie&p=dalministero&id=3633](http://www.salute.gov.it/portale/news/p3_2_1_1_1.jsp?lingua=italiano&menu=notizie&p=dalministero&id=3633). Accessed 25 June 2018.
- Decreto Lorenzin. (n.d.) Legge n.119/17. Disposizioni urgenti in materia di prevenzione vaccinale <http://www.gazzettaufficiale.it/eli/id/2017/08/5/17G00132/sg>. Accessed 22 July 2019.
- Bujnowska-Fedak MM, Waligóra J, Mastalerz-Migas A. The internet as a source of health information and services; 2019.
- Korp P. Health on the internet: implications for health promotion. *Health Educ Res*. 2005;21(1):78–86.
- Benigeri M, Pluye P. Shortcomings of health information on the internet. *Health Promot Int*. 2003;18(4):381–6.
- Bundorf MK, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? *Health Serv Res*. 2006;41(3p1):819–36.
- Cline RJ, Haynes KM. Consumer health information seeking on the internet: the state of the art. *Health Educ Res*. 2001;16(6):671–92.
- Hüfken V, Deutschmann M, Baehring T, Scherbaum W. Use of the internet for health care information: results from a national telephone survey. *Sozial- und Präventivmedizin*. 2004;49(6):381–90.
- Renahy E, Parizot I, Chauvin P. Determinants of the frequency of online health information seeking: results of a web-based survey conducted in France in 2007. *Inform Health Soc Care*. 2010;35(1):25–39.
- Giambi C, Fabiani M, D' Ancona F, et al. Parental vaccine hesitancy in Italy – results from a national survey. *Vaccine*. 2018;36(6):779–87.
- Larson HJ, de Figueiredo A, Xiaohong Z, et al. The state of vaccine confidence 2016: global insights through a 67-country survey. *eBioMedicine*. 2016;12:295–301.
- Maltezou HC, Theodoridou K, Ledda C, Rapisarda V. Vaccination of healthcare personnel: time to rethink the current situation in Europe. *Future Microbiol*. 2019;14(9s):5–8. <https://doi.org/10.2217/fmb-2018-0262>.
- We are social. Digital Report 2018. (2018). <https://digitalreport.wearesocial.com>. Accessed 22 July 2019.
- Covolo L, Ceretti E, Passeri C, Boletti M, Gelatti U. What arguments on vaccinations run through YouTube videos in Italy? A content analysis. *Hum Vaccin Immunother*. 2017;13(7):1693–9.
- Ruiz JB, Bell RA. Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information. *Vaccine*. 2014;32(44):5776–80.
- Siliquini R, Ceruti M, Lovato E, Bert F, Bruno S, De Vito E, et al. Surfing the internet for health information: an Italian survey on use and population choices. *BMC Med Inform Decis Mak*. 2011;11(1):21.
- Wiley KE, Steffens M, Berry N, Leask J. An audit of the quality of online immunisation information available to Australian parents. *BMC Public Health*. 2017;17(1):76.
- Yiannakoulis N, Slavik CE, Chase M. Expressions of pro-and anti-vaccine sentiment on YouTube. *Vaccine*. 2019;37(15):2057–64.
- Rieder, B. (2015). YouTube data tools. Computer software. version. 1.0. <https://tools.digitalmethods.net/netwizz/youtube>. Accessed 22 July 2019.
- Youtube Data API. (n.d.) <https://developers.google.com/youtube/v3/>. Accessed 1 December 2018.
- Ministero della Salute 2019b. English Version. [http://www.salute.gov.it/portale/p5\\_11.jsp](http://www.salute.gov.it/portale/p5_11.jsp). Accessed on the 1 Jan 2020.
- Ministero della Salute 2019a. Elenco pagine per tag: Vaccini. [http://www.salute.gov.it/portale/temi/p2\\_2\\_0.jsp?lingua=italiano&id=881](http://www.salute.gov.it/portale/temi/p2_2_0.jsp?lingua=italiano&id=881). Accessed on 1 Jan 2020.
- Freeman LC. A set of measures of centrality based on betweenness. *Sociometry*; 1977. p. 35–41.
- Higuchi K. Analysis of free comments in a questionnaire survey: quantitative analysis by KH coder. *Shakai Chosa*. 2012;8:92–6.
- Jaccard P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull Soc Vaudoise Sci Nat*. 1901;37:547–79.
- Jockers, M. (2017). Package 'syuzhet'. URL: <https://cran.r-project.org/web/packages/syuzhet>.
- Betsch C. Innovations in communication: the internet and the psychology of vaccination decisions. *Eurosurveillance*. 2011;16(17):19849.
- ISTAT. (2007). Le tecnologie dell'informazione e della comunicazione: disponibilità nelle famiglie e utilizzo degli individui. [http://www.istat.it/salastampa/comunicati/non\\_calendario/20080116\\_00/testointegrale20080116.pdf](http://www.istat.it/salastampa/comunicati/non_calendario/20080116_00/testointegrale20080116.pdf). Accessed 22 July 2019.
- ISTAT. (2008). Statistiche demografiche. <http://demo.istat.it/bil2008/index.html>. Accessed 22 July 2019.
- Internet World Stats. (2019). <http://www.internetworldstats.com/stats.htm>. Accessed 22 July 2019.
- Censis. (2006). La nuova domanda di comunicazione sulla salute. <http://www.webcitation.org/5uX12lbQS>. Accessed 22 July 2019.
- Flash Eurobarometer. European citizens' digital health literacy. (2014). [http://ec.europa.eu/commfrontoffice/publicopinion/flash/fl\\_404\\_en.pdf](http://ec.europa.eu/commfrontoffice/publicopinion/flash/fl_404_en.pdf). Accessed 2017 July 19.

39. Covolo L, Passeri C, Gelatti U. YouTube Italian videos on vaccinations: a content analysis: Loredana Covolo. *Eur J Public Health*. 2016;26(suppl\_1):ckw174–227.
40. MacDonald NE. Vaccine hesitancy: definition, scope and determinants. *Vaccine*. 2015;33(34):4161–4.
41. Abrahamson JA, Fisher KE, Turner AG, Durrance JC, Turner TC. Lay information mediary behavior uncovered: exploring how nonprofessionals seek health information for themselves and others online. *Journal of the Medical Library Association: JMLA*. 2008;96(4):310.
42. Sak G, Diviani N, Allam A, Schulz PJ. Comparing the quality of pro-and anti-vaccination online information: a content analysis of vaccination-related webpages. *BMC Public Health*. 2015;16(1):38.
43. Odone A, Fara GM, Giammaco G, Blangiardi F, Signorelli C. The future of immunization policies in Italy and in the European Union: the declaration of Erice. *Human Vaccin Immunother*. 2015;11(5):1268–71.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

