

Original Paper

Characterization of False or Misleading Fluoride Content on Instagram: Infodemiology Study

Matheus Lotto^{1,2}, DDS, MSc; Tamires Sá Menezes¹, DDS; Irfhana Zakir Hussain^{2,3}, BTech; Shu-Feng Tsao², BSc, MSc; Zahid Ahmad Butt², MBBS, MSc, PhD; Plinio P Morita^{2,4,5,6,7}, BEng, MSc, PhD; Thiago Cruvinel¹, DDS, MSc, PhD

¹Department of Pediatric Dentistry, Orthodontics and Public Health, Bauru School of Dentistry, University of São Paulo, Bauru, Brazil

²School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada

³Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, India

⁴Research Institute for Aging, University of Waterloo, Waterloo, ON, Canada

⁵Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

⁶eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

⁷Institute of Health Policy, Management, and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Thiago Cruvinel, DDS, MSc, PhD

Department of Pediatric Dentistry, Orthodontics and Public Health

Bauru School of Dentistry

University of São Paulo

Alameda Dr Octávio Pinheiro Brisolla, 9-75, Vila Universitária

Bauru, 17012-901

Brazil

Phone: 55 14 3235 8318

Fax: 55 14 3223 4679

Email: thiagocruvinel@fob.usp.br

Abstract

Background: Online false or misleading oral health-related content has been propagated on social media to deceive people against fluoride's economic and health benefits to prevent dental caries.

Objective: The aim of this study was to characterize the false or misleading fluoride-related content on Instagram.

Methods: A total of 3863 posts ranked by users' total interaction and published between August 2016 and August 2021 were retrieved by CrowdTangle, of which 641 were screened to obtain 500 final posts. Subsequently, two independent investigators analyzed posts qualitatively to define their authors' interests, profile characteristics, content type, and sentiment. Latent Dirichlet allocation analysis topic modeling was then applied to find salient terms and topics related to false or misleading content, and their similarity was calculated through an intertopic distance map. Data were evaluated by descriptive analysis, the Mann-Whitney *U* test, the Cramer *V* test, and multiple logistic regression models.

Results: Most of the posts were categorized as misinformation and political misinformation. The overperforming score was positively associated with older messages (odds ratio [OR]=3.293, *P*<.001) and professional/political misinformation (OR=1.944, *P*=.05). In this context, time from publication, negative/neutral sentiment, author's profile linked to business/dental office/news agency, and social and political interests were related to the increment of performance of messages. Although political misinformation with negative/neutral sentiments was typically published by regular users, misinformation was linked to positive commercial posts. Overall messages focused on improving oral health habits, side effects, dentifrice containing natural ingredients, and fluoride-free products propaganda.

Conclusions: False or misleading fluoride-related content found on Instagram was predominantly produced by regular users motivated by social, psychological, and/or financial interests. However, higher engagement and spreading metrics were associated with political misinformation. Most of the posts were related to the toxicity of fluoridated water and products frequently motivated by financial interests.

KEYWORDS

eHealth; fluorides; infodemiology; information seeking behavior; internet; misinformation; social media; infoveillance; health outcome; dental caries; health information; dental health

Introduction

The analysis of big data originating from people's production and consumption of online dental information can contribute to recognizing the needs of distinct populations, aiding the planning and implementation of public health actions [1,2]. Within this context emerged the concept of infodemiology, defined as "the science of distribution and determinants of information in an electronic medium with the ultimate aim to inform public health and public policy" [3]. Specifically, internet users have adopted social media to perform queries and express their concerns, doubts, and advice about oral health conditions [4,5]. However, while these behaviors are desirable to provide empowerment and autonomy for individuals toward health education and decision-making [6,7], the content overabundance of social network ecosystems poses a challenge to the public to filter relevant posts, which leads to the consumption of false information and, consequently, the development of damaging health beliefs [8-10]. In this way, previous studies demonstrated that Instagram could be a significant source of health information, including several issues such as COVID-19 and vaccination [11], especially considering the increased popularity of this platform in recent years [12,13].

In this scenario, online false or misleading content propagates the discouragement of the consumption of fluoride-containing water and oral care products concerning their relevance, safety, and harmful consequences [14]. Notably, anti-fluoridation information is broadly shared on social media, deceiving people against fluoride's economic and health benefits [15]. Moreover, some characteristics of these false or misleading posts, such as the sense of innovation and the negative sentiment charges, favor the diffusion of falsehoods in contrast to trustworthiness [16,17]. In parallel, fluoride refusal is a growing phenomenon observed in dental offices, possibly generated or reinforced by online misinformation [18]. Divergently, there is robust scientific evidence on the beneficial effects of fluoridated water, dentifrices, and mouthwashes to prevent the demineralization and promote the remineralization of dental tissues. Fluoride is considered the most effective measure to reduce the incidence and prevalence of dental caries [19-21], which is the most prevalent oral disease worldwide, affecting the permanent and deciduous teeth of approximately 2.3 billion people and 532 million children, respectively [22].

Thus, the adoption of digital strategies to manage the oral health information disorder on social media is mandatory. Toward this end, the aim of this study was to characterize the false or misleading fluoride-related content on Instagram, regarding authors' and posts' features, interaction and spreading metrics, and the sentiment of posts.

Methods

Study Design

This longitudinal and retrospective infodemiology study analyzed and characterized the false or misleading fluoride-related content of 500 English posts on Instagram. A total of 3863 posts ranked by users' total interaction were retrieved by CrowdTangle, of which 641 were screened for the inclusion criteria. All posts were made available on Instagram between August 2016 and August 2021. Two independent investigators (ML and TSM) analyzed these posts qualitatively to define the authors' interests, profile characteristics, content type, and sentiment of posts. Topic modeling methods were applied to find salient terms and topics related to false or misleading fluoride content. Finally, statistical analysis was performed as described in detail below.

Ethics Considerations

This study did not require institutional review board approval from the Council of Ethics in Human Research of Bauru School of Dentistry because federal regulations do not apply to research using publicly available data that does not involve human subjects. It should be emphasized that the raw data presented in this manuscript have been anonymously disclosed in an open data repository [23].

Search Strategy, Data Collection, and Preprocessing Data Set

CrowdTangle is an online analytics and insights tool owned by Meta Inc that enables the study of several social media metrics such as the number of posts, data, profile information, type of posts, total interaction (sum of the number of likes, comments, and views in a post), and overperforming score through specific keywords. It is also possible to access posts from distinct periods, languages, and social media, besides ranking them into various measures.

The overperforming score is a post's performance regarding its actual interaction divided by its expected interaction according to the number of followers of the author's profile (ie, how many ordinary followers the post reached). In this way, positive scores are associated with good performance posts, reaching a larger user's number than simply the number of the author's followers, and negative scores convey the opposite. Briefly, the algorithm of CrowdTangle generates benchmarks to identify these expected values using the last 100 posts from a given account. For this calculation, the top and bottom 25% posts are dropped and then the mean number of interactions are calculated with the middle 50% of posts in different time intervals (15 minutes old, 60 minutes old, 5 hours old, etc). Subsequently, when the account in question publishes a new post, the platform compares the post metrics to the calculated average and multiplies the difference by the weights in each dashboard [24].

The search strategy (“fluoride free”+“fluoride-free”) was defined from exploratory analyses of hashtags and terms related to a higher volume of posts that discouraged fluoride use on Instagram. A data set related to 3863 posts was downloaded as a CSV file on September 15, 2021, regarding specific language (English) and time frame (August 2016 to August 2021), and ranked by total interaction. The period for the collection was determined from the availability of data observed in a preliminary analysis using the search strategy on CrowdTangle and the number of worldwide Instagram users [25]. Furthermore, posts were ranked by total interaction to guarantee the inclusion of those accessed by a considerable volume of Instagram users (ie, those influencing a number of individuals not relativized by the potential of authors to achieve an audience).

Before the qualitative and natural language processing analyses, the raw data set was preprocessed in two ways depending on the type of investigation. First, the data set was screened to obtain a feasible number of posts (n=500), enabling a robust qualitative manual evaluation to feed artificial intelligence-based models, and preventing expected mischaracterization associated with automated tools. Thus, an investigator (ML) read a sample of collected posts (n=641) in full to obtain a list of the first 500 posts ranked by total interaction that satisfied the following inclusion criterion: nonrepeated false or misleading content published in English. The investigator excluded 139 posts due to repetition and 2 posts that were not published in English. It is noteworthy that this process aimed to characterize posts containing false or misleading content with the highest engagement rates on Instagram.

To ensure the quality of topic modeling analysis, another investigator (IZH) performed an additional preprocessing of the words of 500 selected posts, removing symbols, special characters, punctuations, URLs, numbers, personal pronouns, and keywords of the search strategy.

Data Analysis

Qualitative Analysis

The false or misleading fluoride posts were characterized through passive qualitative analysis [26], examining information patterns and interaction metrics. This approach was directed by the most accepted definitions of the categories of information disorder, as follows: (1) misinformation, defined as false information determined based on a grounding of truth and applies only to informationally oriented content [27-29]; (2) fake news, defined as intentionally misleading and biased representational information for the benefit of the messenger sender, which contains false information, with or without a blend of one or more components of omitted important information, a decontextualized content, misleading headlines, or clickbait [30]; (3) disinformation, defined as information that is false and deliberately created to harm a person, social group, organization, or country [27,28]; and (4) conspiracy theories, which are attempts to explain the ultimate causes of significant social and political events and circumstances with claims of secret plots by two or more powerful actors [31].

Additionally, false or misleading online content can be motivated by distinct types of interest such as financial (profiting from information disorder through advertising), political (attempts to influence public opinion due to political positions), social (connecting with a particular group online or offline), and psychological (seeking prestige or reinforcement) [27]. The identification of specific motivations could be a reliable and objective indicator of authors' intentionality, regarding that its determination is only based on the subjective judgment of online content founded on researchers' perspectives [27]. However, according to Poe's law, the clues left by content makers are often inadequate to differentiate between honest and dishonest mistakes (ie, the authors' intentions to deliberately produce or share misleading content to deceive people cannot be categorically identified) [8,32]. Regarding the aforementioned difficulties to establish the specific type of information disorder, misinformation was characterized by two trained and calibrated investigators (ML and TSM) (intraclass correlation coefficient for absolute concordance varying from 0.85 to 0.92), according to the following criteria: author's profile (regular users, business, dental office, or news agency), type of content (commercial or noncommercial), author's interest (social, psychological, financial, and/or political), and sentiment (negative, neutral, or positive). Commercial content was detected when associated with a business, dental office, and news agency, or with regular users identified as influencers for promoting the sales of dental products. Both investigators were trained by the discussion of representative characteristics of posts. The calibration of individual judgment criteria was confirmed by the independent classification of 10% of posts (n=50). The posts that investigators divergently qualified were reassessed until consensus. Additionally, the combination of the author's profile (dental office or others) and the detection of political interests (yes or no) defined the categories of information disorder, grouped as misinformation (posts from regular users without political interests), professional misinformation (posts from a dental office without political interests), and political misinformation (posts from authors with political interests).

Natural Language Processing

Topic modeling is an unsupervised machine learning method that is effectively used to identify patterns within a large corpus of unstructured documents, as previously observed in the health information area [33,34]. Interestingly, researchers who apply unsupervised algorithms do not need to previously define issues in topic modeling, corroborating with the automatized evaluation of social media data sets [35]. Besides a faster analysis, this process allows for identification that would not have been achieved by manual inspection because it is less prone to human biases [36].

We applied latent Dirichlet allocation (LDA) topic modeling using Python 3 in a Google Colab interface to determine the main salient terms and topics from the studied data set, examining the relationship between similar and different content. Synthetically, LDA is a probabilistic and word count-based model that analyzes the frequency of words to determine distinct topics [33]. Given the number of topics K , LDA algorithms may generate a keyword list that is most relevant to each topic individually. Although this analysis does not provide a complete

meaning of social media posts, it can contribute to a good overview of issues, facilitating data interpretation [37]. A detailed description of the LDA model is provided elsewhere [38].

We defined the ideal number of topics based on the metric proposed by Nikolenko et al [39] for qualitative studies. Thus, a higher coherence score represents topic modeling with better quality, simplifying the interpretation of outputs. In this way, the coherence values were computed for K topics, where K ranges from 2 to 50, before eventually narrowing down the consideration range to 3-15 topics. We then carefully examined the models with the highest coherence values and selected that with the most significant score [35,40]. Finally, the topics' distances were calculated to establish their similarity through an intertopic distance map.

Statistical Analysis

Statistical analysis was performed using the Statistical Package for Social Sciences (v. 21.0). First, the variables were dichotomized as follows: time from publication (≤ 859 or > 859 days), categories of information disorder (misinformation or professional/political misinformation), authors' profile (regular users or business/dental office/news agency), sentiment (negative/neutral or positive), type of content (commercial or noncommercial), type of publication (video or photo), total interaction (≤ 1179 or > 1179), overperforming score (≤ 1.38 or > 1.38). The continuous variables were dichotomized from their median values. Additionally, dental offices, news agency, and business profiles were dichotomized on the same side because of their common financial background.

The data normality and homogeneity were determined through the Kolmogorov-Smirnov test and Levene test, respectively.

Subsequently, as data were nonnormally distributed, the comparison of total interaction and overperforming score of dichotomized variable groups was performed by the Mann-Whitney U test. The differences in the distribution of dichotomized variables according to the categories of information disorder were assessed by the Cramer V test.

Additionally, multiple logistic regression models were developed to evaluate the association of overperforming scores and total interaction with distinct variables. Only factors with significant Wald statistics in the simple analyses were included in the multiple regression models. For all analyses, $P < .05$ was considered significant.

Results

As shown in Table 1, in general, the posts were predominantly commercial, produced by regular users, expressing positive sentiment, and published as an album/photo. The types of interests identified among the 500 selected posts were social ($n=500$, 100.0%), psychological ($n=492$, 98.4%), financial ($n=421$, 84.2%), and political ($n=79$, 15.8%). Considering the specific interests and authors' profiles, the investigators categorized the posts as misinformation ($n=413$, 82.6%), political misinformation ($n=79$, 15.8%), and professional misinformation ($n=8$, 1.6%).

Table 1 presents the comparison of total interaction and overperforming scores with the distinct dichotomized variable groups. A significantly higher number of total interaction was found for noncommercial content items, whereas a significantly higher overperforming score was detected for > 859 days, professional/political misinformation, business/dental office/news agency profiles, and negative/neutral sentiment.

Table 1. Comparison of total interaction and overperforming scores between dichotomized variable groups.

Variable	Posts (N=500), n (%)	Total interaction		Overperforming score		<i>P</i> values ^a	
		Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Total interaction	Overperforming score
Time from publication						.64	<.001
≤859 days	250 (50.0)	2720 (5031)	1149 (1265)	1.95 (7.48)	1.10 (3.22)		
>859 days	250 (50.0)	2292 (3404)	1222 (1633)	6.15 (18.27)	1.89 (3.57)		
Types of interest						.36	.01
Social, financial, and psychological	421 (84.2)	2468 (4168)	1160 (1473)	4.23 (15.20)	1.33 (4.63)		
Social and political	79 (15.8)	2710 (4449)	1271 (1138)	3.11 (5.15)	1.98 (4.46)		
Author's profile						.46	<.001
Regular users	296 (59.2)	2701 (4759)	1189 (1486)	1.65 (13.79)	1.06 (3.05)		
Business/dental office/news agency	204 (40.8)	2224 (3511)	1155 (1278)	7.54 (13.84)	3.59 (6.40)		
Sentiment						.71	.004
Negative/neutral	77 (15.4)	2699 (5003)	1263 (1118)	3.54 (5.66)	2.30 (4.62)		
Positive	423 (84.6)	2471 (4160)	1164 (1490)	4.15 (15.14)	1.33 (4.52)		
Type of content						.009	.54
Noncommercial	95 (19.0)	3408 (5687)	1554 (2240)	1.98 (4.36)	1.63 (1.63)		
Commercial	405 (81.0)	2295 (3878)	1144 (1279)	4.54 (15.49)	1.35 (5.02)		
Type of publication						.54	.61
Video	45 (9.0)	2835 (4090)	1218 (2017)	2.26 (3.95)	1.70 (4.96)		
Photo	455 (91.0)	2474 (4319)	1174 (1414)	4.23 (14.72)	1.38 (4.77)		

^aMann-Whitney *U* test (*P*<.05 considered statistically significant).

Table 2 summarizes the distribution of distinct dichotomized variable groups according to the categories of information disorder. Accordingly, the overperforming score and noncommercial content were significantly higher among professional misinformation and political misinformation

groups. Furthermore, political misinformation was frequently posted by regular users with negative/neutral sentiment. By contrast, misinformation commonly presented commercial content with positive feelings.

Table 2. Distribution of dichotomized variable groups according to the categories of information disorder.

Variable	Misinformation (n=413), n (%)	Professional misinformation (n=8), n (%)	Political misinformation (n=79), n (%)	ϕ	P value ^a
Time from publication				0.79	.21
≤859 days	209 (50.6)	6 (75.0)	35 (44.3)		
>859 days	204 (49.4)	2 (25.0)	44 (55.7)		
Overperforming				0.149	.004
≤1.38	221 (53.5)	1 (12.5)	30 (37.9)		
>1.38	192 (46.5)	7 (87.5)	49 (62.1)		
Author's profile				0.154	.003
Regular users	247 (59.8)	0 (0)	49 (62.1)		
Business/dental office/news agency	166 (40.2)	8 (100)	30 (37.9)		
Sentiment				0.712	<.001
Negative/neutral	18 (4.3)	0 (0)	59 (74.7)		
Positive	395 (95.7)	8 (100)	20 (25.3)		
Total interaction				0.061	.39
≤1179	212 (51.3)	4 (50.0)	34 (43.1)		
>1179	201 (48.7)	4 (50.0)	45 (56.9)		
Type of content				0.493	<.001
Noncommercial	42 (10.2)	6 (75.0)	47 (59.5)		
Commercial	371 (89.8)	2 (25.0)	32 (40.5)		
Type of publication				0.136	.01
Video	31 (7.5)	0 (0)	14 (17.7)		
Photo	382 (92.5)	8 (100)	65 (82.3)		

^aCramer V test ($P < .05$ considered significant).

Table 3 displays the results of the multiple logistic regression model for overperforming score. Overperforming was positively associated with older posts and professional/political misinformation. Notably, total interaction did not show significant Wald statistics for any factor in the simple analysis.

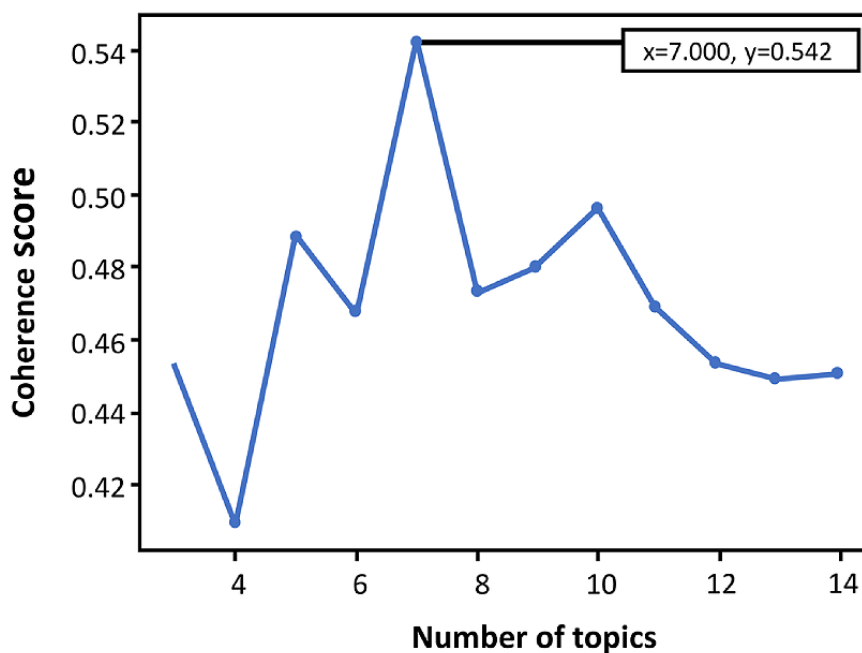
Table 3. Multiple logistic regression model for overperforming score (>1.38).

Variable	B ^a (SE)	Wald statistic	OR ^b (95% CI)	P value
Time from publication (>859 days)	1.192 (0.189)	39.84	3.293 (2.274-4.768)	<.001
Information disorder (professional/political misinformation)	0.664 (0.336)	3.900	1.944 (1.005-3.758)	.05
Sentiment (positive)	-0.143 (0.335)	0.163	0.867 (0.434-1.731)	.69
Constant (y-intercept)	-0.605 (0.367)	2.717	0.546	.10

^aUnstandardized coefficient.

^bOR: odds ratio.

We adopted an exploratory process to select the topic modeling algorithm with the best performance concerning the coherence score. Figure 1 depicts these values from different number of topics, demonstrating the most significant value for 7 topics (0.54).

Figure 1. Coherence scores for distinct numbers of topics.

Thus, the LDA algorithm was executed with all posts ($N=500$) through the configuration $K=7$, which generated 7 different fluoride-related topics. Based on the salient keywords of each topic, we attributed a brief description to determine their meaning and subsequently stratified them regarding the main issues, as presented in Table 4. Figure 2 shows the topics' distances to establish their similarity through an intertopic distance map. There was higher proximity of topics 3, 4, and 5; a similarity between topics 1 and 7; and a considerable distance of topics 2 and 6 from the others. Overall, the topics that emerged from the analysis were related to discouraging the consumption of fluoridated products and water by adults and children, justified by their toxicity, using arguments on the

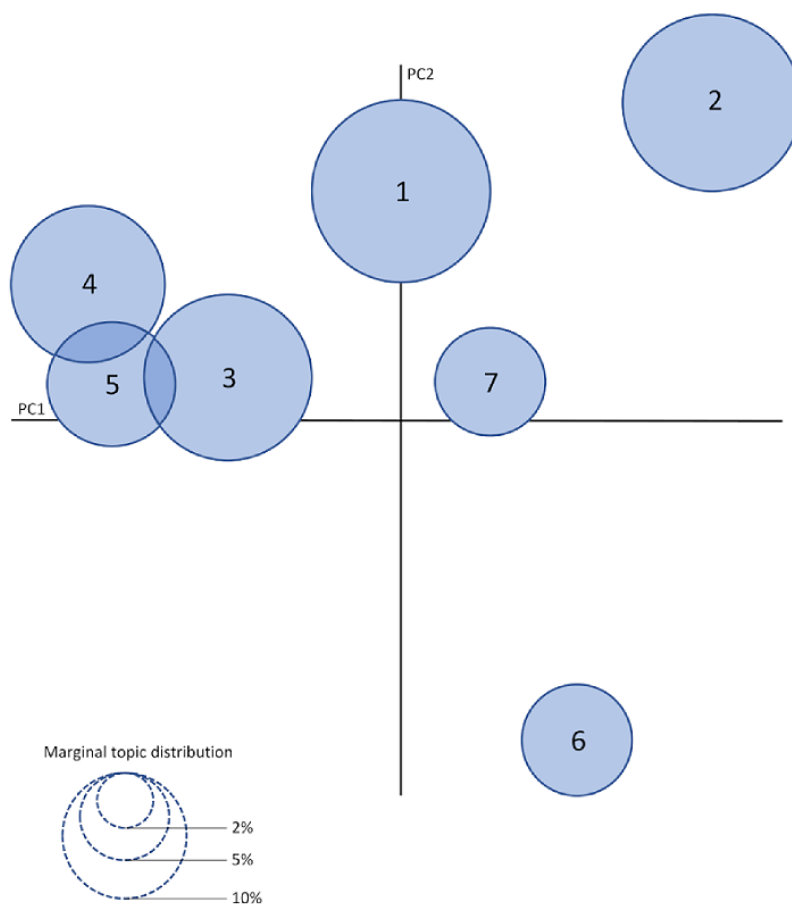
improvements of oral health habits (topics 1 and 7), side effects of fluoride (topic 2), the use of dentifrice containing natural and/or vegan ingredients (topics 3, 4, and 5), and propaganda of fluoride-free oral care products (topic 6).

The most representative words of each topic were employed to determine its issues, depending on the specific context of posts, as verified in the manual analysis. For example, the word "giveaway" was linked to posts about dentifrices containing natural and/or vegan ingredients because several authors promote draws of this kind of products, as follows: "It is GIVEAWAY TIME! Baby care is simplified with Dr. Brown's wide range of health and hygiene products."

Table 4. Fluoride-related salient topics stratified according to number of posts, most frequent words, issues, and examples.

Topic	Posts, n	Most frequent words	Issues	Examples
1	115	Love, Day, Use, Product, Body, Skin, Time, Natural, Get, Feel, Help, Work, Life, Know, Try	Improvements of oral health habits	“Do your kids enjoy brushing their teeth?! My boys used to fight it until we made it a fun routine!! @grinnatural has now become part of our routine and not only has it helped our kids oral care but has also become a fun activity they look forward to!”
2	62	Water, Drink, Health, Body, Level, Use, Filter, Pineal Gland, Study, High, Know, Brain, Cause, Bone, Source	Side effects of fluoride	“Intentional poisoning of the municipal water sources with toxic fluoride and other toxins/heavy metals of primary source for pineal gland calcification”
3	32	New, Ingredient, Love, Ad, Clean, Formula, Kid, Fresh, Target, Adult, Know, Try, Smile, Toothpaste, Flavor	Use of dentifrice containing natural/vegan ingredients	“Brush-brush the germs away from your baby’s teeth with the help of Mee Mee’s Fluoride Free Strawberry Flavour Toothpaste”
4	93	Natural, Whiten, Charcoal, Product, Smile, Activate Charcoal, Use, Vegan, Giveaway, Ingredient, White, Toothbrush, Winner, Follow, Coconut	Use of dentifrice containing natural/vegan ingredients	“The Grounded Activated Charcoal Teeth Powder is a 100% natural and fluoride free teeth whitening formula to brighten your teeth shade, remove plaque, cleanse the mouth, remove toxins & make your mouth feel sparkling clean”
5	77	Kid, Brush, Brush Tooth, Love, Baby, Fun, Toothbrush, Fruit, Natural, Ad, Flavor, Routine, Start, Child, Safe Swallow	Use of dentifrice containing natural/vegan ingredients	“#ad Chloe’s favorite part of her morning routine is brushing her teeth. Thankfully @toms_of_maine makes brushing her teeth fun with their Silly Strawberry toothpaste. Chloe loves the delicious taste and I love that it’s natural free from artificial flavors, colors and preservatives”
6	93	Oral Care, Gum, Mouth, Product, Mouthwash, Bacteria, Disease, Cavity, Oral, Plaque, Bad, Natural, @Garnensgarden, Breath, Garnens Garden	Propaganda of fluoride-free oral care products	“Make the Switch, to an all-natural oral care products from Garners Garden (@garnersgarden)! Protect your gums and teeth from cavities and bad bacteria!”
7	28	Organic, Use, Tongue, Add, Oil, Healthy, Daily, Toxin, Routine, Clean, Coconut Oil, Brush, Day, Tap, and Antibacterial	Improvements of oral health habits	“How many of you guys Oil Pull? It’s one of my favorite ways to detox and keep my teeth healthy/white”

Figure 2. Intertopic distance map of the topic modeling analysis. Note that the bubbles are denominated according to the number of the specific topic. PC: principal component.



Discussion

Principal Findings and Comparison With Prior Work

These findings indicate that the predominant false or misleading fluoride Instagram posts were categorized as misinformation ($n=413$) and political misinformation ($n=79$). In this context, several characteristics were related to the increment of overperforming scores of messages, such as time from publication, negative or neutral sentiment, business/dental office/news agency author's profile, and social and political interests. In particular, older messages (odds ratio [OR]=3.29) and professional/political misinformation (OR=1.94) were associated with better performance of spreading among Instagram users. Remarkably, commercial content was significantly more prevalent in the misinformation category than in the professional and political misinformation categories. Furthermore, regular users preponderantly published political misinformation presenting negative or neutral sentiment, whereas misinformation was linked to positive commercial posts. The messages generally addressed the toxicity of fluoridated products and water, focusing on improving oral health habits, side effects of fluoride, dentifrice containing natural and/or vegan compounds, and propaganda of fluoride-free oral care products. Although previous studies have analyzed fluoride-related information on social media, including Instagram [13-15,17,41], this study differs regarding only

focusing on analyses of false or misleading fluoride information, identified based on contemporary concepts and methods on information disorder.

From these outcomes, we confirmed that oral health information seekers engage more with political fluoride misinformation, even after excluding the influence of time as a confounding factor. Indeed, social media consumers tend to connect with others similar to themselves regarding political ideology [42]. People motivated by specific political overviews, influenced by personal characteristics such as beliefs and values, are predisposed to be more interactive with congruent arguments and assimilate them uncritically (confirmation bias) [43,44]. Thus, greater political homophily is associated with increased user interaction since it reinforces similar ideologies [45]. It is important to note that individuals are susceptible to believing and sharing misinformation regardless of their underlying political creed [44].

Moreover, LDA topic modeling categorized most of the political misinformation in topic 2, covering the possible side effects from fluoride toxicity, as exemplified by the following posts:

over three hundred studies have found that fluoride is literally a neurotoxin
fluoridated water provides no benefits, only risks.
Babies given fluoridated water in their formula may have reduced IQ scores.

This demonstrates that Instagram users were strongly influenced by concerns and fears surrounding fluoridated products and the water supply, interacting with negative sentiment posts that emphasized the adverse health aspects of fluoride. These outcomes are in agreement with posts of Twitter users [17].

The positive impact of the time of availability of posts on overperforming scores is an expected result because users have more opportunities to access these posts in comparison to more recent posts. Likewise, authors' profiles linked to economic activities, such as companies, dental offices, and news media, usually structure their messages to attract customers, besides probably paying money to promote their content on Instagram, which increases people's engagement and thus raises content diffusion. Surprisingly, we detected financial interest in most posts, including a substantial portion of regular users (digital influencers) that publicized fluoride-free products. Moreover, several salient topics that emerged from modeling were closely connected to brands. Indeed, the distribution of information disorder often has a close relationship with economic gains [27]. Specifically, our findings suggest that the anti-fluoridation proposals strongly connect with financial concerns beyond the above-discussed ideological aspects. In this sense, distinct oral care companies have been focused on developing products that meet the individual wishes of consumers, even with the absence of scientific evidence [46].

Practical Implications

These findings can support the development of methods and models to automatically identify false or misleading content items and assess their propagation on social media. In addition, outcomes such as topic modeling can subsidize the elaboration of eHealth and mobile health fluoride-related educational approaches to guide social media users toward the consumption of adequate online oral health information [47]. In this context, dental professional teams need to be conscious of fluoride-related misinformation toward improving the quality of their relationship with patients. Additionally, universal access to oral health, improving eHealth and electronic literacy, and offering high-quality dental information are desirable to prevent the consumption of deceptive messages. Certainly, policymakers should recognize the negative influence of these false posts on communities, creating guidelines and laws to control the spread

of information disorder. Specifically, social media managers should be encouraged to develop mechanisms for screening posts to detect false or misleading content before considering messages eligible for sponsorship, avoiding the dissemination of misinformation. Despite the difficulties in determining the authors' intentions, society needs to start discussing education measures and possible penalties for misinformation propagators, within the confines of democratic values, mainly when disseminated by health professionals.

Limitations

This study has some limitations. First, we collected the sample from a specific search strategy composed of two keywords, limiting the findings' generalization to all false or misleading fluoride content. However, we performed an exploratory analysis to determine the most representative keywords with the greatest spread for the thematic analysis in data collection. Second, the two independent investigators analyzed only 500 posts due to work restrictions associated with human analysis, in accordance with previous dental studies [4]. In addition, the manual labeling of data sets is imperative to training artificial intelligence models for natural processing language tasks, ensuring high accuracy and data generalizability [48]. Third, as previously described, we cannot differentiate misinformation from other types of information disorder because of the incapacity of determining authors' intentionality objectively and precisely [49]. Notwithstanding, the characterization of misinformation was improved, verifying the association of specific interests and authorship with interaction metrics. Fourth, these interpretations were based on content published in English. Although English is the most spoken language worldwide, cultural aspects likely influenced the detection of falsehoods.

Conclusions

False or misleading fluoride posts available on Instagram were predominantly characterized as misinformation produced by regular users motivated by social, psychological, and/or financial interests; however, misinformation with social and political interests was associated with higher engagement and spreading metrics. In general, the content of posts was related to the toxicity of fluoridated water and products, frequently motivated by financial interests.

Acknowledgments

This work was supported by the São Paulo Research Foundation (grants 2019/27242-0 and 2021/03226-6). The authors are grateful to Meta Inc for granting use of the CrowdTangle platform.

Conflicts of Interest

None declared.

References

1. Cruvinel T, Ayala Aguirre PE, Lotto M, Marchini Oliveira T, Rios D, Pereira Cruvinel AF. Digital behavior surveillance: monitoring dental caries and toothache interests of Google users from developing countries. *Oral Dis* 2019 Jan;25(1):339-347. [doi: [10.1111/odi.12986](https://doi.org/10.1111/odi.12986)] [Medline: [30270556](https://pubmed.ncbi.nlm.nih.gov/30270556/)]
2. Rizzato VL, Lotto M, Lourenço Neto N, Oliveira TM, Cruvinel T. Digital surveillance: the interests in toothache-related information after the outbreak of COVID-19. *Oral Dis* 2021 Aug 27;1-10 [FREE Full text] [doi: [10.1111/odi.14012](https://doi.org/10.1111/odi.14012)] [Medline: [34448289](https://pubmed.ncbi.nlm.nih.gov/34448289/)]

3. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009 Mar 27;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
4. Heavilin N, Gerbert B, Page JE, Gibbs JL. Public health surveillance of dental pain via Twitter. *J Dent Res* 2011 Sep;90(9):1047-1051 [FREE Full text] [doi: [10.1177/0022034511415273](https://doi.org/10.1177/0022034511415273)] [Medline: [21768306](https://pubmed.ncbi.nlm.nih.gov/21768306/)]
5. Graf I, Gerwing H, Hofer K, Ehlebracht D, Christ H, Braumann B. Social media and orthodontics: a mixed-methods analysis of orthodontic-related posts on Twitter and Instagram. *Am J Orthod Dentofacial Orthop* 2020 Aug;158(2):221-228. [doi: [10.1016/j.ajodo.2019.08.012](https://doi.org/10.1016/j.ajodo.2019.08.012)] [Medline: [32576427](https://pubmed.ncbi.nlm.nih.gov/32576427/)]
6. Tan SS, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. *J Med Internet Res* 2017 Jan 19;19(1):e9 [FREE Full text] [doi: [10.2196/jmir.5729](https://doi.org/10.2196/jmir.5729)] [Medline: [28104579](https://pubmed.ncbi.nlm.nih.gov/28104579/)]
7. Barber SK, Lam Y, Hodge TM, Pavitt S. Is social media the way to empower patients to share their experiences of dental care? *J Am Dent Assoc* 2018 Jun;149(6):451-459. [doi: [10.1016/j.adaj.2018.01.007](https://doi.org/10.1016/j.adaj.2018.01.007)] [Medline: [29656806](https://pubmed.ncbi.nlm.nih.gov/29656806/)]
8. Giglietto F, Iannelli L, Valeriani A, Rossi L. 'Fake news' is the invention of a liar: how false information circulates within the hybrid news system. *Curr Sociol* 2019 Apr 08;67(4):625-642. [doi: [10.1177/0011392119837536](https://doi.org/10.1177/0011392119837536)]
9. Strieder AP, Aguirre PEA, Lotto M, Cruvinel AFP, Cruvinel T. Digital behavior surveillance for monitoring the interests of Google users in amber necklace in different countries. *Int J Paediatr Dent* 2019 Sep;29(5):603-614. [doi: [10.1111/ipd.12500](https://doi.org/10.1111/ipd.12500)] [Medline: [30920686](https://pubmed.ncbi.nlm.nih.gov/30920686/)]
10. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. *Annu Rev Public Health* 2020 Apr 02;41:433-451. [doi: [10.1146/annurev-publhealth-040119-094127](https://doi.org/10.1146/annurev-publhealth-040119-094127)] [Medline: [31874069](https://pubmed.ncbi.nlm.nih.gov/31874069/)]
11. Rovetta A, Bhagavathula AS. Global infodemiology of COVID-19: analysis of Google web searches and Instagram hashtags. *J Med Internet Res* 2020 Aug 25;22(8):e20673 [FREE Full text] [doi: [10.2196/20673](https://doi.org/10.2196/20673)] [Medline: [32748790](https://pubmed.ncbi.nlm.nih.gov/32748790/)]
12. Massey PM, Kearney MD, Hauer MK, Selvan P, Koku E, Leader AE. Dimensions of misinformation about the HPV vaccine on Instagram: content and network analysis of social media characteristics. *J Med Internet Res* 2020 Dec 03;22(12):e21451 [FREE Full text] [doi: [10.2196/21451](https://doi.org/10.2196/21451)] [Medline: [33270038](https://pubmed.ncbi.nlm.nih.gov/33270038/)]
13. Niknam F, Samadbeik M, Fatehi F, Shirdel M, Reza zadeh M, Bastani P. COVID-19 on Instagram: A content analysis of selected accounts. *Health Policy Technol* 2021 Mar;10(1):165-173 [FREE Full text] [doi: [10.1016/j.hlpt.2020.10.016](https://doi.org/10.1016/j.hlpt.2020.10.016)] [Medline: [33723502](https://pubmed.ncbi.nlm.nih.gov/33723502/)]
14. Mackert M, Bouchacourt L, Lazard A, Wilcox GB, Kemp D, Kahlor LA, et al. Social media conversations about community water fluoridation: formative research to guide health communication. *J Public Health Dent* 2021 Jun;81(2):162-166. [doi: [10.1111/jphd.12404](https://doi.org/10.1111/jphd.12404)] [Medline: [33058200](https://pubmed.ncbi.nlm.nih.gov/33058200/)]
15. Mertz A, Allukian M. Community water fluoridation on the Internet and social media. *J Mass Dent Soc* 2014;63(2):32-36. [Medline: [25230407](https://pubmed.ncbi.nlm.nih.gov/25230407/)]
16. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med* 2019 Nov;240:112552 [FREE Full text] [doi: [10.1016/j.socscimed.2019.112552](https://doi.org/10.1016/j.socscimed.2019.112552)] [Medline: [31561111](https://pubmed.ncbi.nlm.nih.gov/31561111/)]
17. Oh HJ, Kim CH, Jeon JG. Public sense of water fluoridation as reflected on Twitter 2009-2017. *J Dent Res* 2020 Jan;99(1):11-17. [doi: [10.1177/0022034519885610](https://doi.org/10.1177/0022034519885610)] [Medline: [31682777](https://pubmed.ncbi.nlm.nih.gov/31682777/)]
18. Eliacik BK. Topical fluoride applications related posts analysis on Twitter using natural language processing. *Oral Health Prev Dent* 2021 Jan 07;19(1):457-464. [doi: [10.3290/j.ohpd.b2048359](https://doi.org/10.3290/j.ohpd.b2048359)] [Medline: [34546013](https://pubmed.ncbi.nlm.nih.gov/34546013/)]
19. Walsh T, Worthington HV, Glenny A, Marinho VC, Jeronic A. Fluoride toothpastes of different concentrations for preventing dental caries. *Cochrane Database Syst Rev* 2019 Mar 04;3:CD007868 [FREE Full text] [doi: [10.1002/14651858.CD007868.pub3](https://doi.org/10.1002/14651858.CD007868.pub3)] [Medline: [30829399](https://pubmed.ncbi.nlm.nih.gov/30829399/)]
20. Belotti L, Frazão P. Effectiveness of water fluoridation in an upper-middle-income country: a systematic review and meta-analysis. *Int J Paediatr Dent*. Preprint posted online on September 26, 2021. [doi: [10.1111/ipd.12928](https://doi.org/10.1111/ipd.12928)] [Medline: [34564916](https://pubmed.ncbi.nlm.nih.gov/34564916/)]
21. Whelton HP, Spencer AJ, Do LG, Rugg-Gunn AJ. Fluoride revolution and dental caries: evolution of policies for global use. *J Dent Res* 2019 Jul;98(8):837-846. [doi: [10.1177/0022034519843495](https://doi.org/10.1177/0022034519843495)] [Medline: [31282846](https://pubmed.ncbi.nlm.nih.gov/31282846/)]
22. GBD 2017 Oral Disorders Collaborators, Bernabe E, Marcenes W, Hernandez CR, Bailey J, Abreu LG, et al. Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the Global Burden of Disease 2017 Study. *J Dent Res* 2020 Apr;99(4):362-373 [FREE Full text] [doi: [10.1177/0022034520908533](https://doi.org/10.1177/0022034520908533)] [Medline: [32122215](https://pubmed.ncbi.nlm.nih.gov/32122215/)]
23. Lotto M, Menezes TS, Hussain IZ, Tsao SF, Butt ZA, Morita PP, et al. Raw data of the manuscript: Characterization of false or misleading fluoride content on Instagram: Infodemiology study. Figshare. 2022 Jan 31. URL: https://figshare.com/articles/dataset/Raw_data_of_the_manuscript_Characterization_of_misleading_fluoride_information_on_Instagram/19099733 [accessed 2022-02-22]
24. How is overperforming calculated? CrowdTangle. 2022. URL: <https://help.crowdtangle.com/en/articles/1141056-how-is-overperforming-calculated> [accessed 2022-02-22]

25. Number of monthly active Instagram users 2013-2021. Statista. 2022. URL: <https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/#:~:text=Social%20media%20usage%20worldwide,with%20114.9%20million%20active%20users> [accessed 2022-02-22]
26. Franz D, Marsh HE, Chen JI, Teo AR. Using Facebook for qualitative research: a brief primer. *J Med Internet Res* 2019 Aug 13;21(8):e13544 [FREE Full text] [doi: [10.2196/13544](https://doi.org/10.2196/13544)] [Medline: [31411143](https://pubmed.ncbi.nlm.nih.gov/31411143/)]
27. Wardle C, Derakhshan H. Information disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe. 2017. URL: <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html> [accessed 2022-02-22]
28. Journalism, 'Fake News' and Disinformation: A Handbook for Journalism Education and Training. UNESCO. 2018. URL: <https://en.unesco.org/fightfakenews>
29. Molina MD, Sundar SS, Le T, Lee D. "Fake News" is not simply false information: a concept explication and taxonomy of online content. *Am Behav Sci* 2019 Oct 14;65(2):180-212. [doi: [10.1177/0002764219878224](https://doi.org/10.1177/0002764219878224)]
30. Lim S. Academic library guides for tackling fake news: a content analysis. *J Acad Librariansh* 2020 Sep;46(5):102195. [doi: [10.1016/j.acalib.2020.102195](https://doi.org/10.1016/j.acalib.2020.102195)]
31. Douglas KM, Uscinski JE, Sutton RM, Cichocka A, Nefes T, Ang CS, et al. Understanding conspiracy theories. *Polit Psychol* 2019 Mar 20;40(S1):3-35. [doi: [10.1111/pops.12568](https://doi.org/10.1111/pops.12568)]
32. Aikin SF. Poe's Law, group polarization, and argumentative failure in religious and political discourse. *Soc Semiot* 2013 Jun;23(3):301-317. [doi: [10.1080/10350330.2012.719728](https://doi.org/10.1080/10350330.2012.719728)]
33. Asmussen CB, Møller C. Smart literature review: a practical topic modelling approach to exploratory literature review. *J Big Data* 2019 Oct 19;6(1):93. [doi: [10.1186/s40537-019-0255-7](https://doi.org/10.1186/s40537-019-0255-7)]
34. Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T. Public discourse and sentiment during the COVID 19 pandemic: using latent Dirichlet allocation for topic modeling on Twitter. *PLoS One* 2020;15(9):e0239441 [FREE Full text] [doi: [10.1371/journal.pone.0239441](https://doi.org/10.1371/journal.pone.0239441)] [Medline: [32976519](https://pubmed.ncbi.nlm.nih.gov/32976519/)]
35. Pang PC, McKay D, Chang S, Chen Q, Zhang X, Cui L. Privacy concerns of the Australian My Health Record: implications for other large-scale opt-out personal health records. *Inf Process Manag* 2020 Nov;57(6):102364. [doi: [10.1016/j.ipm.2020.102364](https://doi.org/10.1016/j.ipm.2020.102364)]
36. Hagen L. Content analysis of e-petitions with topic modeling: how to train and evaluate LDA models? *Inf Process Manag* 2018 Nov;54(6):1292-1307. [doi: [10.1016/j.ipm.2018.05.006](https://doi.org/10.1016/j.ipm.2018.05.006)]
37. Jockers ML, Mimno D. Significant themes in 19th-century literature. *Poetics* 2013 Dec;41(6):750-769. [doi: [10.1016/j.poetic.2013.08.005](https://doi.org/10.1016/j.poetic.2013.08.005)]
38. Blei DM. Probabilistic topic models. *Commun ACM* 2012 Apr 01;55(4):77-84. [doi: [10.1145/2133806.2133826](https://doi.org/10.1145/2133806.2133826)]
39. Nikolenko SI, Koltcov S, Koltsova O. Topic modelling for qualitative studies. *J Inf Sci* 2016 Jul 10;43(1):88-102. [doi: [10.1177/0165551515617393](https://doi.org/10.1177/0165551515617393)]
40. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. 2015 Presented at: 8th ACM International Conference on Web Search and Data Mining; 2015; Shanghai, China p. 339-408. [doi: [10.1145/2684822.2685324](https://doi.org/10.1145/2684822.2685324)]
41. Basch CH, Milano N, Hillyer GC. An assessment of fluoride related posts on Instagram. *Health Promot Perspect* 2019;9(1):85-88 [FREE Full text] [doi: [10.15171/hpp.2019.11](https://doi.org/10.15171/hpp.2019.11)] [Medline: [30788272](https://pubmed.ncbi.nlm.nih.gov/30788272/)]
42. Mitchell A, Gottfried J, Kiley J, Matsa KE. Pew Research Center. 2014 Oct 21. URL: <https://www.pewresearch.org/journalism/2014/10/21/political-polarization-media-habits/> [accessed 2022-02-22]
43. Kim A, Moravec PL, Dennis AR. Combating fake news on social media with source ratings: the effects of user and expert reputation ratings. *J Manag Inf Syst* 2019 Aug 04;36(3):931-968. [doi: [10.1080/07421222.2019.1628921](https://doi.org/10.1080/07421222.2019.1628921)]
44. Scherer LD, Pennycook G. Who is susceptible to online health misinformation? *Am J Public Health* 2020 Oct;110(S3):S276-S277. [doi: [10.2105/AJPH.2020.305908](https://doi.org/10.2105/AJPH.2020.305908)] [Medline: [33001736](https://pubmed.ncbi.nlm.nih.gov/33001736/)]
45. Boutyline A, Willer R. The social structure of political echo chambers: variation in ideological homophily in online networks. *Polit Psychol* 2016 May 05;38(3):551-569. [doi: [10.1111/pops.12337](https://doi.org/10.1111/pops.12337)]
46. Bauler LD, Santos CSD, Lima GS, Moraes RR. Charcoal-based dentifrices and powders: analyses of product labels, Instagram engagement, and altmetrics. *Braz Dent J* 2021;32(2):80-89 [FREE Full text] [doi: [10.1590/0103-6440202104233](https://doi.org/10.1590/0103-6440202104233)] [Medline: [34614064](https://pubmed.ncbi.nlm.nih.gov/34614064/)]
47. Lotto M, Strieder AP, Ayala Aguirre PE, Oliveira TM, Andrade Moreira Machado MA, Rios D, et al. Parental-oriented educational mobile messages to aid in the control of early childhood caries in low socioeconomic children: A randomized controlled trial. *J Dent* 2020 Oct;101:103456 [FREE Full text] [doi: [10.1016/j.jdent.2020.103456](https://doi.org/10.1016/j.jdent.2020.103456)] [Medline: [32827598](https://pubmed.ncbi.nlm.nih.gov/32827598/)]
48. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward COVID-19 vaccines in the United Kingdom and the United States: observational study. *J Med Internet Res* 2021 Apr 05;23(4):e26627 [FREE Full text] [doi: [10.2196/26627](https://doi.org/10.2196/26627)] [Medline: [33724919](https://pubmed.ncbi.nlm.nih.gov/33724919/)]
49. Son GHW, Rashid EIA. Classification of information disorder. Khazanah Research Institute. 2022. URL: <http://www.krinstitute.org/assets/contentMS/img/template/editor/DP%20-%20Classification%20of%20Information%20Disorder.pdf> [accessed 2022-02-22]

Abbreviations**LDA:** latent Dirichlet allocation**OR:** odds ratio

Edited by A Mavragani; submitted 23.02.22; peer-reviewed by R Ratto Moraes, PCI Pang, L Bouchacourt, J Chen; comments to author 21.03.22; revised version received 01.04.22; accepted 14.04.22; published 19.05.22

*Please cite as:**Lotto M, Sá Menezes T, Zakir Hussain I, Tsao SF, Ahmad Butt Z, P Morita P, Cruvinel T
Characterization of False or Misleading Fluoride Content on Instagram: Infodemiology Study**J Med Internet Res 2022;24(5):e37519**URL: <https://www.jmir.org/2022/5/e37519>**doi: [10.2196/37519](https://doi.org/10.2196/37519)**PMID:*

©Matheus Lotto, Tamires Sá Menezes, Irfhana Zakir Hussain, Shu-Feng Tsao, Zahid Ahmad Butt, Plinio P Morita, Thiago Cruvinel. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 19.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.