Vaccine 40 (2022) 953–960

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Vaccine misinformation types and properties in Russian troll tweets

Echo L. Warner [a,b,*], Juliana L. Barbati [c], Kaylin L. Duncan [c], Kun Yan [c], Stephen A. Rains [c]

[a] University of Arizona Cancer Center, 1515 N Campbell Ave, Tucson, AZ 85724, USA
[b] College of Nursing, University of Arizona, 1350 S Martin Ave. Tucson, AZ 85721, USA
[c] Department of Communication, College of Social & Behavioral Sciences, University of Arizona, 1103 E University Blvd, Tucson, AZ 85721, USA

ARTICLE INFO

Article history:
Received 19 July 2021
Received in revised form 6 December 2021
Accepted 14 December 2021
Available online 13 January 2022

Keywords:
Misinformation
Disinformation
Twitter
Tweet
Social media
Communication

ABSTRACT

Objective: To identify the content of and engagement with vaccine misinformation from Russian trolls on Twitter.
Methods: Troll tweets (N = 1959) obtained from Twitter in 2020 were coded for vaccine misinformation (α = 0.77–0.97). Descriptive, bivariate, and multivariable negative binomial regressions were applied to estimate robust incidence rate ratios (IRRs) and 95% confidence intervals (95 %CI) of vaccine misinformation associations with tweet characteristics and engagement (i.e., replies, likes, retweets).
Results: Misinformation about personal dangers (43.0%), civil liberty violations (20.2%), and vaccine conspiracies (18.6%) were common. More misinformation tweets used anti-vaccination language (97.3% vs. 13.2%) and referenced symptoms (37.4% vs. 0.5%) than non-misinformation tweets. Fewer misinformation tweets referenced credible sources (14.0% vs. 19.5%), were formatted as headlines (39.2% vs. 77.0%), and mentioned specific vaccines (11.3% vs. 36.1%, all $p < 0.01$) than non-misinformation tweets. Personal dangers misinformation had 83% lower rate of retweets (95 %CI 0.04–0.66). Civil liberties misinformation had significantly higher rate of replies (IRR: 7.65, 95 %CI 1.06–55.46), but lower overall engagement (IRR: 0.38, 95 %CI 0.16–0.88) than non-misinformation tweets.
Conclusions: Strategies used to promote vaccine misinformation provide insight into the nature of vaccine misinformation online and public responses. Our findings suggest a need to explore influences on whether users reject or entertain online vaccine misinformation.

1. Background

Defined as information that is factually inaccurate or misleading regardless of intent to deceive, misinformation can undermine pro-health behaviors [1–3]. Indeed, the spread of vaccine misinformation online threatens their collective efficacy [4]. Existing studies demonstrate that foreign and domestic groups lead and sow discord about vaccines [5]. One notable example is the Russian Internet Research Agency's (IRA) deployment of operatives on social media—referred to as "trolls"—around the 2016 presidential election [6]. Vaccination was a prominent topic raised by trolls who advocated pro- and anti-vaccination positions [4,7]. Although arguments against vaccination were common, little is known about how the public responded to misinformation shared by trolls.

A recent surge in anti-vaccine sentiments and misinformation began with the publication of Andrew Wakefield's study that inaccurately linked vaccination for measles, mumps, and rubella (MMR) to autism [8]. Even after redaction, many influential celebrities and figures continue to spread these false findings, leading to a stark decrease in public vaccination [8], and vaccine hesitancy has been amplified in the COVID-19 era. Today, anti-vaccine ideas typically revolve around four main themes: trust, alternatives, safety, and conspiracies [8]. Although these ideas are broad and reflect a general opposition to vaccination, they also frequently manifest in misinformation [9]. While anti-vaccine messages do not necessarily contain misinformation, anti-vaccine rhetoric is often based on misinformed and misinterpreted data and opinions. A single study on user engagement with troll vaccine misinformation documented low engagement from general users, but this study did not consider different types of vaccine misinformation [10].

In this project, we analyze messages posted to Twitter (i.e., tweets) about vaccination made by IRA trolls as part of their efforts to sow discord among the American public. We advance existing research by focusing on the message properties of tweets containing misinformation and allowing for misinformation to manifest as pro- and anti-vaccine messages. We examine specific properties associated with vaccine misinformation and the degree to which

* Corresponding Author at: University of Arizona Cancer Center, University of Arizona College of Nursing, 1305 Martin Ave, Tucson, AZ 85709, USA.
E-mail address: echowarner@email.arizona.edu (E.L. Warner).

https://doi.org/10.1016/j.vaccine.2021.12.040
0264-410X/

the presence of misinformation fosters audience engagement in the form of likes, retweets, and comments. Additionally, we analyze different classes of and engagement with vaccine misinformation.

Because the American public was not generally aware of the trolls' status as Russian IRA actors endeavoring to foster antipathy, these data offer a naturalistic setting in which to investigate the content and impact of vaccination misinformation. Prior studies have shed light on these data, showing that trolls generally share polarized pro- and anti-vaccine perspectives by embodying political personas [7] and erode public trust in vaccines [4]. Despite its importance, few studies have documented troll efforts to spread vaccine misinformation and to our knowledge this is the first report isolating the message features associated with misinformation and engagement with misinformation. In doing so, this project offers valuable insights about the implications of specific message features that are associated with vaccine misinformation on social media, their influence on public attitudes toward vaccination, and advances the classification of misinformation by documenting pro-vaccine tweets with misinformation.

2. Methods

2.1. Sample selection

We used data provided by Twitter including nearly 9 million tweets from more than 3,500 accounts. The data are from the accounts Twitter identified as IRA accounts and supplied to the U.S. Congress as part of their election interference investigation [11]. We limited our sample to tweets that contained the keyword "vacc." [7], including 1,959 tweets appearing between 2015 and 2017.

2.2. Tweet misinformation

Importantly, because, we opted not to apply the definition of disinformation to this content, which requires an intentional deliberate effort to knowingly deceive to gain "money, power, or reputation" given that we are unable to determine if the trolls posting or sharing false or misleading content were aware the content included falsehoods [12]. To begin, we coded for the presence or absence of misinformation in tweets. First, we adapted an existing codebook for anti-vaccine sentiment to the context of misinformation by reviewing published definitions for each category and reviewing vaccine misinformation topics. Because prior studies have documented that IRA trolls amplify both sides of the vaccine debate [4], our coding approach allowed for misinformation that was both pro-vaccine (e.g., "#VaccinateUS You should take a little poison to get stronger") and tweets that were anti-vaccine without misinformation (e.g., "#tech Survey: Parents increasingly ask doctors to delay vaccines" or "#News #US Anti-vaccine moms speak out amid fierce backlash"). Three authors initially identified twenty focused misinformation categories in tweets by reviewing 250 tweets. Conceptually similar content categories were combined to form six vaccine misinformation macro categories consistent with prior research [4,13]: personal dangers, civil liberties, conspiracy, educational, violations of morality, and alternative medicine (examples in Table 1 and the complete coding scheme in Supplemental Table 1).

2.3. Tweet message property variables

Each tweet was coded for multiple message properties (examples in Table 2), including citing a source with medical expertise (e.g., highly credible source including academic journals, government, scientists, clinicians), formatting the tweet as a news headline, articulating a pro-, anti-, or neutral vaccine position, referencing a vaccine symptom or side-effect, referring to legal requirements or legislation for vaccines, and identifying a specific vaccine type or brand. Tweets that mentioned a specific vaccine were coded for the presence or absence of all vaccines used in the United States [14]. In total, 24 vaccines were mentioned in the tweets, including veterinary vaccines, vaccines that did not exist at the time of data collection (e.g., Ebola, Zika, HIV/AIDS), and "other" vaccines (e.g., diabetes, cancer). A tweet's status as an original or retweet was supplied by Twitter.

2.4. Coding procedures

Multiple iterative rounds of coding were conducted to refine the classification of vaccine misinformation categories and to develop the tweet message property variables. During each round, three coders applied the coding scheme to 50 tweets, agreement was evaluated, and discrepancies discussed amongst the research team and codes were added or modified through consensus. After five rounds of coding, intercoder reliability was evaluated on a set of 250 tweets and found to be in the moderate to almost perfect range for each category (Krippendorff's α: 0.77–0.97) [15]. The coders independently applied the coding scheme to the remaining tweets.

2.5. Tweet engagement

Tweet engagement was measured using metadata provided by Twitter. For all original tweets (n = 1598), Twitter reported the number of likes, replies, and retweets (including "quote" tweets) each tweet received from other users. These measures were evaluated separately; we also created an overall tweet engagement score, which was a dichotomous variable reflecting the presence of at least one like, reply, or retweet.

2.6. Statistical analysis

We summarized the types of vaccine misinformation in the dataset and compared the prevalence of specific message properties among tweets with and without misinformation using chi-square tests and Fisher Exact tests for variables with cell sizes < n = 5. We calculated summaries of vaccine engagement among original tweets for the full sample and for misinformation tweets separately. Then we estimated negative binomial multivariate regressions with robust incident rate ratios (IRRs), which are often applied to generate estimates in datasets with count data containing a large number of zeros [16], to evaluate engagement by tweet characteristics for the full sample and misinformation tweets only. Given the small number of misinformation tweets that were promoted on Twitter, only selected effect estimates are shown that contained adequate sample size to apply regression modeling. All analyses were performed in Stata 16.0 and statistical significance was set at $p < 0.05$.

3. Results

3.1. Misinformation categories, prevalence, and examples

There were 372 (19.0%) vaccine tweets that contained misinformation and 1587 (81.0%) did not. The most common vaccine misinformation was related to personal dangers (43.0%), followed by civil liberty violations (20.2%), vaccine conspiracies (18.6%), educational misinformation (5.9%), morality violations (2.1%), and homeopathic alternatives (2.1%). Table 1 provides definitions of

Table 1
Vaccine misinformation categories, prevalence, and examples.

Category	Definition	n	% all tweets (*N* = 1959)	% misinformation tweets (*N* = 372)	Examples
Personal dangers	Addresses specific negative outcomes of vaccination	160	8.2	43.0	• People!!! just think that you could die from #vaccination for disease which you might not get at all #VaccinateUS • #politics Paul: Vaccines can cause mental disorders
Civil liberties	Focuses on personal freedoms and vaccination requirements	75	3.8	20.2	• Christie: Parents deserve measure of choice on vaccinations #news • #VaccinateUS Obligatory #vaccination? What's next? Obligatory taxes for #vaccination?
Conspiracy	Discusses sinister actions and actors operating against the public's interest	69	3.5	18.6	• have you ever thought that #vaccines are a weapon our government is using? #VaccinateUS • How could I know that it's not a biological weapon in your ampule?! #VaccinateUS
Educational	Portrays doctors as lacking knowledge or identifies anti-vaccination claims from medical professionals	22	1.1	5.9	• RT @***: Study: Unvaccinated Children Healthier Than The Rest | World https:// *** https://*** • RT @***: Truth about vaccine science fraud - http://*** http://*** via @***
Violations of morality	Argues that vaccines transgress an ethical or moral code	8	0.4	2.1	• #vaccination is interference in God's will! #VaccinateUS • RT @***: Study Links Autism to Vaccines Made With Cells From Aborted Babies https://***https://***
Alternative medicine	Focuses on homeopathic alternatives to vaccines	8	0.4	2.1	• RT @***: Curcumin Found To Outperform Pneumococcal Vaccines In Protecting Infants https://*** https://*** • RT @***: Better than vaccines? Vitamin D found to be powerful prevention vs. colds and flu https://***#antivax #nutri

Note. Categories are not mutually exclusive. Anti-vaccination tweets were coded into multiple categories when they addressed more than one misinformation topic. Mentions and URLs were replaced with "***".

categories, prevalence, and examples of vaccine misinformation tweets.

3.2. Vaccine tweet message properties by misinformation content

In Table 2, 65% of tweets with misinformation were original tweets compared to only 35% of misinformation tweets that were retweets ($p < 0.01$). Only 14.0% of misinformation tweets cited a source compared to 19.5% of tweets without misinformation ($p < 0.01$). Similarly, 39.2% of misinformation tweets were formatted as a news headline compared to 77.0% of tweets without misinformation ($p < 0.01$). Notably 97.3% of misinformation tweets were anti-vaccine (1.3% were pro-vaccine and 1.3% neutral) compared to only 13.2% of tweets without misinformation ($p < 0.01$), and 37.4% of misinformation tweets referenced a vaccine side effect or symptom compared to only 0.5% of tweets without misinformation ($p < 0.001$). Only eleven misinformation tweets referenced legislation or laws about vaccines, but those that did were primarily against vaccine legislation (90.9% vs. 9.1% that were neutral, $p < 0.001$). The most referenced vaccines in misinformation tweets were influenza (19.0%), mumps (16.7%), measles (14.3%), human papilloma virus (HPV) (14.3%), rubella (9.5%), polio (9.5%), Zika (7.1%), and Ebola (2.4%). A higher proportion of misinformation tweets referenced mumps (16.7% vs. 3.1%, $p < 0.01$) and rubella (9.5% vs. 1.7%, $p = 0.01$) compared to tweets without misinformation. Supplemental Fig. 1 shows the top 25 terms used to describe the tweet content with misinformation (top 5: parents, health, children, choice, Christie) and without misinformation (top 5: health, news, bill, politics, California).

3.3. Overall association of misinformation categories with engagement

In the full sample, 293 tweets (15.0%) received some form of engagement. In Table 3, tweets with misinformation about personal dangers of vaccination received a significantly lower number of retweets than tweets without personal danger misinformation (IRR: 0.17, 95% CI 0.04–0.66, $p = 0.01$). Similarly, tweets with civil liberties misinformation had 97% and 95% lower incidence of being liked or retweeted, respectively, and overall received a lower incidence of any type of engagement (IRR: 0.38, 95% CI 0.16–0.88, $p = 0.02$). There was, however one exception, tweets with misinformation about civil liberties had 7.65 higher incidence of replies (95% CI 1.06–55.46, $p = 0.04$) than tweets that were not about civil liberties.

3.4. Association of tweet misinformation with engagement among misinformation tweets

In Table 4, misinformation tweets formatted as a headline were 94% less likely to be liked (95% CI 0.02–0.27, $p < 0.01$), and 95% less likely to be retweeted (95% CI 0.01–0.25, $p < 0.01$) than tweets not formatted as headlines. Tweets with a reference to vaccine symptoms were significantly more likely to be liked than those that did not (IRR: 7.61, 95% CI 1.66–34.76, $p < 0.01$).

3.5. Overall association of tweet message properties with engagement

We examined how Twitter users engaged with the full sample of troll vaccine tweets through a composite variable indicating any form of engagement, replies, likes, and retweets. When it comes to any kind of engagement, tweets that cited a source were 28% more likely to receive any engagement (95 %CI 1.00–1.64, $p = 0.04$), and tweets that mentioned a specific vaccine had 46% higher incidence of receiving any form of engagement (95% CI 1.19–1.78, $p < 0.01$) than their counterparts. Tweets with an anti-vaccination stance (IRR: 0.62, 95% CI 0.47–0.83, $p < 0.01$) compared to pro-vaccine tweets and those that contained misinformation (IRR: 0.60, 95 %CI 0.41–0.88, $p < 0.01$) compared to those that did not had lower incidence of engagement. Only tweets formatted as news headlines had 89% lower incidence of receiving a reply (95% CI 0.05–0.23, $p < 0.01$, Table 4) compared to tweets that were not formatted as headlines.

Tweets formatted as headlines had 84% lower incidence of being liked (95% CI 0.06–0.41, $p < 0.01$) than tweets that were

Table 2
Vaccine tweet message properties by misinformation content (N = 1959).

	Total N = 1959		Misinformation N = 372, 19.0%		Not Misinformation N = 1587, 81.0%			Tweet examples
	N	%	N	%	N	%	p-value[1]	
Tweet origin								RT @WorldTruthTV: Study: Unvaccinated Children Healthier Than The Rest (retweet example)
Original	1598	81.6	242	65.0	1356	85.4	**<0.01**	
Retweet	361	18.4	130	35.0	231	13.6		
Cites source								#VaccinateUS American Medical Association state that #vaccines are safe (cites source example)
Yes	361	18.4	52	14.0	309	19.5	**<0.01**	
No	1598	81.6	320	86.0	1,278	80.5		
Headline								RT @NYC_Everyday: Type 1 diabetes vaccine clears first hurdle (headline example)
Yes	1368	69.8	146	39.2	1222	77.0	**<0.01**	
No	591	30.2	226	60.8	365	23.0		
Vaccination stance								People should protect their children and #vaccination helps to do it #VaccinateUS (pro-vaccine example)
Pro	911	46.5	5	1.3	906	57.1	**<0.01**	
Anti	572	29.2	362	97.3	210	13.2		
Neutral	476	24.3	5	1.3	471	29.7		
References symptom								Paul: Vaccines can cause mental disorders #politics (symptom example)
Yes	147	7.5	139	37.4	8	0.50	**<0.01**	
No	1812	92.5	233	62.6	1579	99.5		
Legislation reference[2]								Vaccine mandate bill revamped to ease home-schooling waiver #politics #news (anti-vaccine legislation example)
Pro	149	7.6	0	0.0	149	41.4	**<0.01**	
Anti	39	2.0	10	90.9	29	8.1		
Neutral	183	9.4	1	9.1	182	50.6		
No reference	1588	97.0	361	97.0	1227	77.3		
Mentions specific vaccine								Maryland, U.S. cancer centers endorse new HPV vaccine recommendations (specific vaccine mention example)
Yes	615	31.4	42	11.3	573	36.1	**<0.01**	
No	1344	68.6	331	88.7	1013	63.9		
Vaccine reference								#TopNews California school orders non-vaccinated kids home after possible measles case (measles vaccine example)
Measles	125	20.3	6	14.3	119	20.8	0.31	
Influenza (flu)	100	16.3	8	19.0	92	16.1	0.61	
Ebola[3]	60	9.8	1	2.4	59	10.3	0.07	
Zika[3]	56	9.1	3	7.1	53	9.2	0.45	
HPV	51	8.3	6	14.3	45	7.8	0.14	
Polio	29	4.7	4	9.5	25	4.4	0.13	
Meningococcal	26	4.2	2	4.8	24	4.2	0.86	
Mumps	25	4.1	**7**	**16.7**	**18**	**3.1**	**<0.01**	
Yellow Fever	17	2.8	0	0.0	17	3.0	0.30	
HIV/AIDS[3]	15	2.4	1	2.4	14	2.4	0.73	
Pertussis	14	2.3	1	2.4	13	2.3	0.63	
Veterinary	14	2.3	2	4.8	12	2.1	0.25	
Rubella	14	2.3	**4**	**9.5**	**10**	**1.7**	**0.01**	
Pneumococcal	9	1.5	1	2.4	8	1.4	0.47	
Varicella (chickenpox)	6	1.0	1	2.4	5	0.9	0.35	
Hepatitis B	6	1.0	2	4.8	4	0.7	0.06	
Rabies	5	0.8	0	0.0	5	0.9	0.70	
Hepatitis A	4	0.7	0	0.0	4	0.7	0.75	
Shingles	4	0.7	0	0.0	4	0.7	0.75	
Smallpox	3	0.5	0	0.0	3	0.5	0.81	
Anthrax	3	0.5	0	0.0	3	0.5	0.81	
Cholera	3	0.5	0	0.0	3	0.5	0.81	
Diphtheria	3	0.5	1	2.4	2	0.3	0.19	
Other[4]	60	9.8	**0**	**0.0**	**60**	**10.5**	**0.03**	

[1] p-value for chi-square tests, or Fisher's exact tests for variables with cell sizes n < 5.
[2] Only among n = 371 tweets referencing legislation.
[3] Ebola, Zika, and HIV/AIDS vaccinations did not exist at the time of the tweets.
[4] Other included references to cancer, Alzheimer, Hawaiian monk seals, MERS, mad cow disease, RSV, prairie dogs, dengue, depression, and diabetes.

not formatted as headlines. Compared to pro-vaccine tweets, those that were anti (IRR: 0.35, 95% CI 0.16–0.80, p = 0.01) and neutral (IRR: 0.36, 95% CI 0.16–0.78, p < 0.01) had lower incidence of being liked. In addition, tweets that contained vaccine misinformation were 57% less likely to be liked (95% CI 0.21–0.87, p = 0.02) than tweets without vaccine misinformation. In contrast, tweets that did not reference vaccine legislation had significantly higher incidence of being liked compared to tweets with pro-vaccine legislation (IRR:3.60, 95% CI 1.40–9.31, p < 0.01).

Tweets that were formatted as headlines (IRR: 0.26, 95% CI 0.08–0.85, p = 0.03), took an anti (IRR: 0.34, 95% CI 0.15–0.75, p < 0.01) or neutral (IRR: 0.40, 95% CI 0.17–0.96, p = 0.04) vaccination stance, and tweets that contained any vaccine misinformation (IRR: 0.31, 95% CI 0.15–0.63, p < 0.01) had significantly lower incidence of being retweeted than their counterparts.

4. Discussion

Previous analyses of Russian IRA tweets have documented a concerted effort of trolls to polarize vaccination by generating large volumes of partisan tweets for and against vaccination [4,7]. Our research builds on these reports by examining engagement with

Table 3
Overall association of misinformation types with tweet engagement (N = 1598)[1].

	Any engagement (N = 293, 18.3%)			Reply (N = 51, 3.2%)			Like (N = 181, 11.3%)			Retweets[3] (N = 181, 11.3%)		
	IRR	95 %CI	p-value	IRR	95 %CI	p-value	IRR	95 %CI	p-value	IRR	95 %CI	p-value
Personal dangers												
No	Ref.			Ref.			Ref.			Ref.		
Yes	0.69	0.36–1.33	0.27	0.43	0.04–4.37	0.48	0.34	0.11–1.03	0.06	**0.17**	**0.04–0.66**	**0.01**
Civil liberties												
No	Ref.			Ref.			Ref.			Ref.		
Yes	**0.38**	**0.16–0.88**	**0.02**	**7.65**	**1.06–55.46**	**0.04**	**0.03**	**0.01–0.13**	**<0.01**	**0.05**	**0.01–0.25**	**<0.01**
Conspiracy												
No	Ref.			Ref.			Ref.			Ref.		
Yes	0.61	0.16–2.31	0.47	[2]	[2]	[2]	[2]	[2]	[2]	0.56	0.11–2.71	0.47
Educational												
No	Ref.			Ref.			Ref.			Ref.		
Yes	0.87	0.25–3.08	0.83	[2]	[2]	[2]	0.38	0.03–5.29	0.47	0.25	0.05–1.33	0.10
Violations of morality												
No	Ref.			Ref.			Ref.			Ref.		
Yes	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]
Alternative medicine												
No	Ref.			Ref.			Ref.			Ref.		
Yes	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]	[2]

[1] Adjusted for total number of followers and following per account.
[2] Unable to estimate due to small sample size.
[3] Retweets included quote tweets, a type of retweet that allows the user to make a statement in addition to retweeting content.

vaccine misinformation from trolls on Twitter. This work is responsive to ongoing needs that have been identified in the field of vaccine criticism on the Internet [17]. Specifically, we distinguish six different types of vaccine misinformation, specific message properties associated with misinformation, presence of vaccine misinformation in pro-vaccine tweets, and the public's engagement with all vaccine tweets and tweets with vaccine misinformation from trolls. These findings can help accelerate the identification of vaccine misinformation on social media and may inform social media-based public health vaccine campaigns.

4.1. Characteristics of misinformation tweets

The most common vaccine misinformation topics included personal dangers, civil liberties, and conspiracy theories. The prevalence of personal danger and conspiracy-related misinformation suggests that trolls used misinformation to promote fear about vaccines by highlighting the perceived negative personal or social consequences of vaccination. The prevalence of civil liberty misinformation is likely an outgrowth of the trolls' broader strategy of undermining public faith in democracy [4]. The educational, violations of morality, and alternative medicine categories appeared in less than 11% of misinformation tweets. These findings imply a need for targeted public health messaging to debunk misinformation about personal dangers of vaccines, to reframe civil liberty arguments about vaccination, and to counter conspiracy theories.

We observed several trends in tweet properties associated with misinformation. One notable characteristic was the high prevalence of tweets referencing symptoms. Specifically, 37.4% of misinformation tweets referenced a vaccine side effect or symptom compared to only 0.5% of tweets without misinformation ($p < 0.01$). We examined the overlap between tweets that referenced symptoms and those with personal harm misinformation – 90.6% of the tweets that referenced symptoms contained personal harm misinformation ($p < 0.001$). There is a high likelihood that tweets about vaccine side effects contain misinformation; 9 out of every 10 tweets about side effects included misinformation. Suggesting that, at least on Twitter, side effects were not discussed accurately by trolls. One interpretation of these findings is that trolls embellish or fabricate the dangers posed by vaccines to elevate and leverage fear. Although it remains to be seen if this trend applies to Twitter users more broadly, the potential connection between side effects and misinformation could offer a valuable means for researchers to identify and potentially counter vaccine misinformation.

Significantly more misinformation tweets mentioned a specific vaccine than tweets without misinformation. Influenza, mumps, HPV, and measles were most commonly discussed in misinformation tweets. Social media companies that are motivated to halt the spread of vaccine misinformation may consider systematically reviewing tweets about these specific vaccines. Significantly higher proportions of tweets about mumps (16.7%) and rubella (9.5%) included misinformation. We suspect this is due to heated debates about the efficacy and safety of the MMR vaccine [16]. Notably, HIV/AIDS, Ebola, and Zika vaccines, which have not received FDA approval in the United States, were common, and some of the tweets about these vaccines involved misinformation. This suggests that misinformation circulates about vaccines on social media well before public availability, underscoring the need to preemptively counter vaccine misinformation online.

Finally, there were significantly more misinformation tweets that did not cite a source, did not take the form of a news headline, adopted an anti-vaccine stance, and were against vaccine legislation than non-misinformation tweets. We found misinformation in pro-vaccine tweets as well, suggesting a critical need for social media companies, researchers, and public health practitioners recognize the potential for misinformation to be masked in pro-vaccine tweets. Our data was from before the COVID-19 pandemic, so it is unclear the extent to which our results apply to misinformation about new vaccines, but examining both pro- and anti-COVID-19 vaccine content for misinformation is a priority. These findings collectively point to potential actionable steps in managing online misinformation content, such as automated flagging of tweets that contain terms related to symptoms, pro- and anti-vaccine ideas, and vaccine legislation.

4.2. Engagement

Our study contributes to existing literature by examining engagement patterns based on tweet message properties and mis-

Table 4
Association of tweet message properties with engagement[1].

Misinformation Sample (N = 242)	Any engagement (N = 23, 9.5%)			Reply (N = 8, 3.3%)			Like (N = 11, 4.5%)			Retweets[2] (N = 13, 5.4%)		
	IRR	95 %CI	p-value	IRR	95 %CI	p-value	IRR	95 %CI	p-value	IRR	95 %CI	p-value
Cites source												
No	Ref.			Ref.			Ref.			Ref.		
Yes	0.84	0.23–3.060	0.79	[3]	[3]	[3]	0.57	0.06–5.18	0.62	0.63	0.11–3.75	0.61
Headline												
No	Ref.			Ref.			Ref.			Ref.		
Yes	0.43	0.17–1.09	0.08	[3]	[3]	[3]	**0.06**	**0.02–0.27**	**<0.01**	**0.05**	**0.01–0.25**	**<0.01**
Vaccination stance												
Pro	Ref.			Ref.			Ref.			Ref.		
Anti	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]
Neutral	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]
References a symptom												
No	Ref.			Ref.			Ref.			Ref.		
Yes	1.51	0.72–3.15	0.28	0.34	0.06–1.99	0.23	**7.61**	**1.66–34.76**	**<0.01**	1.63	0.45–5.94	0.46
Legislation reference												
Pro	Ref.			Ref.			Ref.			Ref.		
Anti	2.46	0.33–18.38	0.38	5.82	0.69–48.91	0.10	[3]	[3]	[3]	[3]	[3]	[3]
Neutral	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]
No reference	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]	[3]
Mentions specific vaccine												
No	Ref.			Ref.			Ref.			Ref.		
Yes	2.63	0.93–7.44	0.07	[3]	[3]	[3]	0.50	0.07–3.27	0.47	3.18	0.55–19.60	0.19
	Any engagement (N = 293, 15.0%)			Reply (N = 51, 2.6%)			Like (N = 181, 9.2%)			Retweets[2] (N = 181, 9.2%)		
Full Sample (N = 1598)	IRR	95 %CI	p-value	IRR	95 %CI	p-value	IRR	95 %CI	p-value	IRR	95 %CI	p-value
Cites source												
No	Ref.			Ref.			Ref.			Ref.		
Yes	**1.28**	**1.00–1.64**	**0.04**	0.68	0.27–1.75	0.42	0.72	0.28–1.43	0.35	0.65	0.34–1.23	0.19
Headline												
No	Ref.			Ref.			Ref.			Ref.		
Yes	1.06	0.84–1.26	0.61	**0.11**	**0.05–0.23**	**<0.01**	**0.16**	**0.06–0.41**	**<0.01**	**0.26**	**0.08–0.85**	**0.03**
Vaccination stance												
Pro	Ref.			Ref.			Ref.			Ref.		
Anti	**0.62**	**0.47–0.83**	**<0.01**	1.17	0.33–4.05	0.81	**0.35**	**0.16–0.80**	**0.01**	**0.34**	**0.15–0.75**	**<0.01**
Neutral	0.93	0.74–1.18	0.56	0.92	0.44–1.92	0.83	**0.36**	**0.16–0.78**	**<0.01**	**0.40**	**0.17–0.96**	**0.04**
References a symptom												
No	Ref.			Ref.			Ref.			Ref.		
Yes	0.79	0.44–1.41	0.42	0.70	0.16–3.04	0.63	0.83	0.34–2.00	0.68	0.42	0.15–1.18	0.10
Legislation reference												
Pro	Ref.			Ref.			Ref.			Ref.		
Anti	0.98	0.48–2.02	0.96	2.91	0.34–25.14	0.33	1.27	0.31–5.09	0.74	0.45	0.16–1.27	0.13
Neutral	0.86	0.56–1.30	0.46	1.15	0.25–5.32	0.86	1.08	0.47–2.50	0.85	0.87	0.33–2.26	0.77
No reference	0.78	0.57–1.08	0.14	1.78	0.46–6.90	0.40	**3.60**	**1.40–9.31**	**<0.01**	1.39	0.62–3.06	0.42
Mentions specific vaccine												
No	Ref.			Ref.			Ref.			Ref.		
Yes	**1.46**	**1.19–1.78**	**<0.01**	0.72	0.43–1.52	0.39	1.21	0.48–3.06	0.68	1.37	0.63–2.99	0.43
Contains misinformation												
No	Ref.			Ref.			Ref.			Ref.		
Yes	**0.60**	**0.41–0.88**	**<0.01**	2.64	0.71–9.82	0.15	**0.43**	**0.21–0.87**	**0.02**	**0.31**	**0.15–0.63**	**<0.01**

[1] Adjusted for total number of followers and following per account.
[2] Retweets included quote tweets, a type of retweet that allows the user to make a statement in addition to retweeting content.
[3] Unable to estimate due to small sample size.

information status. We revealed that tweets with misinformation had 57–69% lower incidence of being liked or retweeted, respectively. There were also differences in the type of audience engagement (see Tables 3–4). In general, misinformation tweets had a higher incidence of receiving comments but lower incidence of likes and retweets. Tweets with misinformation related to civil liberties were more likely to receive replies, but significantly less likely to be retweeted or liked. Tweets containing personal danger misinformation were similarly less likely to be retweeted. These different types of engagement may imply differing underlying motivations. Whereas likes and retweets may indicate endorsement of the tweet's content, the comment feature could be used to draw attention to misinformation in a tweet or alternatively to further misinformation discussions. Although beyond the scope of our study, the pattern of results may reflect users' efforts to respond to and perhaps correct misinformation from trolls.

Overall, engagement with vaccine tweets containing misinformation was relatively low, similar to at least one prior study [11]. Within the misinformation sub-sample (Table 4), there were only significant differences in engagement between tweets including a news headline and tweets referencing a symptom. We speculate these differences may be due to logical persuasion in the case of news headlines and emotion-driven perceptions (e.g., fear) in the case of vaccine symptoms. Conversely, there were more nuances in engagement within the full sample of tweets (Table 4). Advancing from the assumption that likes and retweets indicate endorsement, these patterns may reveal user's ability to recognize misinformation on Twitter. Previous research indicates that indi-

viduals with lower education, lower health literacy, distrust in healthcare, and an affinity for alternative medicine are less likely to detect online misinformation [18–21]. Partisan identity also influences whether misinformation will be refuted, in that people are more likely to share misinformation if it aligns with their political ideology [22]. While we did not take these variables into account in the current study, we encourage future research to explore what characteristics influence whether people reject or entertain online vaccine misinformation. Results of this magnitude would have implications for public health officials and journalists alike.

4.3. Public health implications

The advent of social media has enabled the rapid spread and deep pervasion of vaccine misinformation worldwide [2]. Public health efforts can be thwarted when vaccine misinformation prevails in online social networks. Digital misinformation spread is particularly dangerous when misinformation is high-risk and causes damage before accurate information can be shared [1]. Thus, identifying the strategies used to permeate online vaccine information spread is critical. Public health practitioners should be strategic about developing accurate vaccine messages that appeal to consumers across the vaccine perception continuum so that vaccine advocacy messages are not siloed away from vaccine critic online communities. There is potential to harness the features that were associated with misinformation tweets in our results. For example, identifying tweets that refer to vaccine related side effects may be an actionable cue to misinformation, given that 9 out of 10 tweets about vaccine side effects contained misinformation. Building on this idea, the fact that misinformation was engaged via comments suggests interventions that engage the public in identifying misinformation may be worthwhile. Indeed, this approach is being undertaken by Twitter in a project called Birdwatch [23], which aims to address misinformation through a community-driven approach that allows users to flag and add notes about tweets with misinformation. However, there is mixed documentation of the efficacy of fact-checking on social media for deterring the spread of misinformation. Fact-checking is most successful if it includes a source and is affiliated with trusted universities, health institutions, and government organizations [24,25]. Correction may be less effective if provided by individual users [26] and when correcting conspiracy theories [27], although studies vary in documenting these effects.

Finally, vaccine misinformation included substantial ties to civil liberties and personal freedoms suggesting political undertones, a notably different rhetoric than vaccine symptom misinformation. Public health practitioners and clinicians alike should be prepared to refute this line of reasoning in online spaces and clinical encounters. At minimum, individual users can be cognizant of tweet characteristics that are associated with misinformation. Future studies should further evaluate associations between misinformation the presence of images or videos, and comparing vaccine misinformation from troll tweets to other types of bots and real users would facilitate more sophisticated comparisons that could have direct implications for public health intervention.

5. Limitations

Perceptions surrounding vaccines are constantly changing and, because this dataset is from 2015 to 2017, the findings may not be applicable to new and emerging vaccinations (e.g., COVID-19). A second limitation is that we lacked individual characteristics of Twitter users who engaged with the trolls' tweets. Geographic location, age, and gender often are associated with vaccine hesitancy and acceptance [28], and we did not account for these factors in our analysis. Our sample size was a relatively small portion of all troll tweets from this time, which may limit generalizability. However, given the important influence of vaccine misinformation on public health this level of detailed analysis is needed. Evaluating individual factors may further inform the amplification of targeted public health messages on social media. Although we applied a rigorous coding structure with specific definitions achieved through multiple rounds of coder consensus, the coding of tweets for misinformation could have been influenced by subjective perceptions of the researchers. Finally, it is possible that the engagement we describe is inflated due to the self-inflation practices of online propaganda campaigns, which often leverage cross-linking and cross-retweeting to artificially inflate engagement.

6. Conclusions

Vaccination of populations against infectious diseases is one of the greatest accomplishments of modern-day medicine and public health. By describing the characteristics of online vaccine misinformation tweets, we point to actionable steps in managing online misinformation content, such as automated flagging of potentially misleading vaccine information in tweets that contain terms about side effects, pro- and anti-vaccine ideas, and vaccine legislation. These findings also build upon the argument that to have the broadest and most penetrable reach, vaccine promotion on Twitter should be strategically designed to mirror the target population's rhetoric so that vaccine supportive messages are not siloed away from vaccine hesitant online communities. The ongoing introduction of new vaccines presents an ideal environment for rapid misinformation spread. Preemptively mitigating vaccine misinformation through active monitoring of misinformation trends is critical to promoting public trust in new vaccines well before vaccine availability.

Funding statement

This work was supported by the National Cancer Institute under award number T32CA078447 (Warner). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Dr. Warner was also supported, in part, by a University of Arizona Postdoctoral Research Development Grant.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.vaccine.2021.12.040.

References

[1] Global Risks 2013: Eighth Edition: Digital Wildfires in a Hyperconnected World. World Economic Forum; 2013.
[2] Sylvia Chou W-Y, Gaysynsky A. A prologue to the special issue: health misinformation on social media. Am J Public Health 2020;110(S3):S270–2.
[3] Southwell BG, Niederdeppe J, Cappella JN, Gaysynsky A, Kelley DE, Oh A, et al. Misinformation as a misunderstood challenge to public health. Am J Prev Med 2019;57(2):282–5.
[4] Broniatowski DA, Jamison AM, Qi SiHua, AlKulaib L, Chen T, Benton A, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. Am J Public Health 2018;108(10):1378–84.

[5] Linvill DL, Warren PL. Troll factories: manufacturing specialized disinformation on Twitter. Polit Commun 2020;37(4):447–67.
[6] Jamieson KH. Cyberwar: how Russian hackers and trolls helped elect a president: what we don't, can't, and do know. Oxford University Press; 2018.
[7] Walter D, Ophir Y, Jamieson KH. Russian Twitter accounts and the partisan polarization of vaccine discourse, 2015–2017. Am J Public Health 2020;110 (5):718–24.
[8] Hussain A, Ali S, Ahmed M, Hussain S. The anti-vaccination movement: a regression in modern medicine. Cureus 2018;10(7):e2919.
[9] Hoffman BL, Felter EM, Chu K-H, Shensa A, Hermann C, Wolynn T, et al. It's not all about autism: The emerging landscape of anti-vaccination sentiment on Facebook. Vaccine 2019;37(16):2216–23.
[10] Dunn AG, Surian D, Dalmazzo J, Rezazadegan D, Steffens M, Dyda A, et al. Limited role of bots in spreading vaccine-critical information among active Twitter users in the United States: 2017–2019. Am J Public Health 2020;110 (S3):S319–25.
[11] Exposing Russia's Effort to Sow Discord Online: The Internet Research Agency and Advertisements. U.S. House of Representatives. https://intelligence.house.gov/social-media-content/. Published 2018. Accessed 04/28/2021.
[12] Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020;41 (1):433–51. https://doi.org/10.1146/annurev-publhealth-040119-094127.
[13] Kata A. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. Vaccine 2010;28(7):1709–16.
[14] List of the Vaccines Used in the United States. Centers for Disease Control and Prevention. https://www.cdc.gov/vaccines/vpd/vaccines-list.html. Published 2018. Accessed 04/05/2021.
[15] McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012;22(3):276–82.
[16] Dong C, Clarke DB, Yan X, Khattak A, Huang B. Multivariate random-parameters zero-inflated negative binomial regression model: An application to estimate crash frequencies at intersections. Accid Anal Prev 2014;70:320–9.
[17] Ward JK, Peretti-Watel P, Verger P. Vaccine criticism on the Internet: Propositions for future research. Hum Vaccin Immunother 2016;12 (7):1924–9.
[18] Scherer Laura D, McPhetres Jon, Pennycook Gordon, Kempe Allison, Allen Larry A, Knoepke Christopher E, et al. Who is susceptible to online health misinformation? A test of four psychosocial hypotheses. Health Psychol 2021;40(4):274–84. https://doi.org/10.1037/hea0000978.
[19] Jones-Jang SM, Mortensen T, Liu J. Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. Am Behav Sci 2021;65(2):371–88.
[20] Khan ML, Idris IK. Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. Behav Inform Technol 2019;38(12):1194–212.
[21] Scherer LD, McPhetres J, Pennycook G, et al. Who is susceptible to online health misinformation? A test of four psychosocial hypotheses. Health Psychol 2021.
[22] Osmundsen M, Bor A, Bjerregaard Vahlstrup P, Bechmann A, Bang PM. Partisian polarization is the prrimary psychological motivation behind political fake news sharing on Twitter. Am Political Sci Rev 2020.
[23] Coleman K. Introducing Birdwatch, a community-based approach to misinformation. https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation.html. Published 2021. Accessed.
[24] Zhang J, Featherstone JD, Calabrese C, Wojcieszak M. Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. Prev Med 2021;145:106408. https://doi.org/10.1016/j.ypmed.2020.106408.
[25] Vraga EK, Bode L. Using expert sources to correct health misinformation in social media. Sci Commun 2017;39(5):621–45.
[26] Vraga EK, Bode L. I do not believe you: how providing a source corrects health misperceptions across social media platforms. Inform, Commun Soc 2018;21 (10):1337–53.
[27] Miller JM, Saunders KL, Farhart CE. Conspiracy endorsement as motivated reasoning: the moderating roles of political knowledge and trust. Am J Polit Sci 2016;60(4):824–44.
[28] MacDonald NE. Vaccine hesitancy: Definition, scope and determinants. Vaccine 2015;33(34):4161–4.