Understanding reactions to swine flu, Ebola, and the Zika virus using Twitter data: an outlook for future infectious disease outbreaks

Wasim Ahmed[1], Peter A. Bath[2], Laura Sbaffi[2], Gianluca Demartini[3]

[1]Newcastle University, Newcastle, UK

[2]University of Sheffield Information School, Sheffield, UK

[3]School of Information Technology and Electrical Engineering, University of Queensland, Brisbane, Australia

Infectious disease outbreaks are a serious public health threat which can disrupt world economies. This paper presents an in-depth qualitative analysis of n=15,415 tweets that relate to the peak of three major infectious diseases: the swine flu outbreak of 2009, the Ebola outbreak of 2014, and the Zika outbreak of 2016. Tweets were analysed using thematic analysis and a number of themes and sub-themes were identified. The results were brought together in an abstraction phase and the commonalities between the cases were studied. A notable similarity which emerged was the rate at which Twitter users expressed intense fear and panic akin to that of the phenomena of "moral panic" and the "outbreak narrative". Our study also discusses the utility of using Twitter data for in-depth qualitative research as compared to traditional interview-methods. Our study is the largest in-depth analysis of tweets on infectious diseases and could inform public health strategies for future outbreaks such as the coronavirus outbreak.

Keywords
Ebola, social media, swine flu, Zika

1. Introduction

Infectious disease outbreaks are severe public health threat that have the potential to cause a high number of fatalities within a very short time. They are also known to account for 29 out of 96 causes of major human mortality [1,2]. Historical cases of infectious disease outbreaks highlight the potential danger and threat to society. The Black Death, or bubonic plague, which took place from 1346 to 1353, is estimated to have taken the lives of somewhere between one-fourth and three-fourths of the world's population across Europe and Asia. In Europe alone, some records have indicated that 25 million people died from it, which included half of the London population at that time (approximately 100,000 people) [3].

The Spanish influenza pandemic (A/H1N1) is another example of a historical infectious disease outbreak. It occurred from 1918 to 1920 and infected 500 million people and claimed 50-100 million lives which, at the time, was equivalent to 3% to 5% of the world's population [4], more lives than the First World War, although the war over-shadowed the pandemic [5]. Historical cases of infectious disease outbreaks highlight how future disease outbreaks could have comparable disastrous consequences for global population health. More recently, for instance, the swine flu pandemic (H1N1/09) in 2009 infected between 43 and 89 million people and is thought to have claimed between 8,870 and 18,300 lives (CDC, 2010).

Due to growing globalisation in the 21st century through international trade and travel, the likelihood of a deadly infectious disease outbreak that could spread from one country to another and develop into a pandemic has increased considerably [6]. In 2003, for example, the severe acute respiratory syndrome (SARS) spread from China to at least 37 countries worldwide in only a matter of weeks [6]. One negative outcome of infectious disease outbreaks is the very high mortality rate; however,

ISBN: 978-91-89081-09-3
DOI: https://doi.org/10.15626/ishimr.2020.04

another is the large impact that they can have on the global economy [6]. For the SARS outbreak alone, the global macro-economic impact was estimated to be $30 to $100 billion, and which ranged from $3 to $10 million per infected case [6]. The costs were dispersed across a number of different sectors, and those that were hit hardest were the travel and tourism industries [6]. The recent epidemic of the novel coronavirus (2019-nCoV) has led to heightened public concern and has also been forecasted to be a major threat to global health and the global economy with stock-markets being affected during the outbreak [7].

Although, the coronavirus (2019-nCoV) has most recently been in the news, the 2009 swine flu and 2014 Ebola outbreaks are two health-related events which, at the time in which they occurred, had highest proportion of media coverage in the 21st century [8]. At the peak times, Google ranked 'swine flu' as the fastest rising Web search query in Google News [9], and 'Ebola' was among the most searched terms in 2014 [10]. Henceforth, because of their viral nature, events related to deadly infectious disease outbreaks are likely to lead to public views and opinions towards emerging diseases and the events that may take place during an outbreak. These views may be expressed in the 'online world', which can act as a space for people to share information, their thoughts and feelings. People began to share health information online towards the latter part of the 20th century for example, through personal websites, discussion forums and online communities. However, in the last few years there has been a shift towards sharing information via social media platforms, e.g., Twitter and Facebook, and this has changed the ways people communicate health issues [11]. Previous deadly outbreaks, such as the Spanish influenza virus and the Black Death, occurred without modern communication devices such as personal computers and mobile phones. The 2009 swine flu pandemic, the 2014 Ebola epidemic, and the 2016 Zika outbreaks all occurred in the age of social media platforms, such as Twitter. This has made it possible to examine unfiltered public views and opinions shared by people during these outbreaks and, importantly, to study the aspects of health that online communities converse about. This is a very important topic to study because slight misunderstandings among the public could potentially have dangerous consequences because social media can lead to the rapid cascading of potentially misleading, and sometimes malicious, information.

Twitter allows its users to send 280-character text updates (originally limited to 140 characters from 2006 to September 2017 [12] in the English language. These character text updates are known as 'tweets' and they can contain thoughts, feelings, activities, and opinions [13]. There are many health communities which are activity using Twitter [14]. For public health researchers, Twitter provides a unique opportunity to examine unmoderated discussions and information that are shared during deadly infectious disease outbreaks. The focus of this paper is to analyse Twitter data related to the peak of the swine flu, Ebola, and Zika outbreaks and then to compare them for potential similarities and differences. In the following three sections these outbreak cases will be briefly outlined before the use Twitter data for investigating phenomena is described.

1.1 H1N1

The swine influenza (flu) Pandemic of 2009 outbreak started in April 2009 and originated from Mexico [15]. It spread across the world because it was a new strain of flu and members of the public had no immunity to it [16]. The United States Centres for Disease Control and Prevention (CDC) announced on the 21st of April 2009 that two patients from California had been infected by the swine flu virus and this then lead to preparations for a swine flu pandemic. Four weeks after the initial two reports in California, 41 countries reported diagnosed cases of the virus [17]. There were an estimated 123,000 to 203,000 deaths due to swine flu from 1st April to 31st December 2009 [18]. The swine flu pandemic led to members of the public in the United States to worry about events that were taking place, and media also began to increase their reporting at this time.

1.2 Ebola

The 2014 outbreak of the Ebola virus was traced to Guinea in December 2013, and spread across West Africa. As of January 2016, there were 28,637 cases of Ebola across and 11,315 deaths [19]. The 2014 epidemic was the largest epidemic of Ebola which has ever been recorded, and the number of cases outnumbered all of the previous cases combined [20]. In June 2014, Médecins Sans Frontières (MSF) noted that the outbreak was out of control and, in August 2014, the United Nations (UN) declared that Ebola was an international public health emergency [21]. During infectious disease outbreaks there is the potential for people from non-medical backgrounds to share health information which may not be correct and could potentially be harmful. Henceforth, during the 2014 Ebola epidemic, it was reported that medical misinformation on the cures of Ebola had taken a number of lives due to the rumour that salt water had the potential to cure Ebola [22].

1.3 Zika

Zika became a public concern in early 2016, when it first spread outside of Africa and Asia, to where the virus had formerly been restricted. This geographical expansion led to the World Health Organisation (WHO) declaring the outbreak a public health emergency of international concern (PHEIC) [23]. The Zika virus is linked to adverse consequences during pregnancy, as well as birth defects, such as microcephaly and brain defects [24]. However, unlike previous infectious disease outbreaks, there appeared to be very limited public knowledge of the Zika outbreak in the United States, leading to the dissemination of health information such as mosquito-bite prevention notices and recommendations by public health authorities to avoid travelling to Zika-affected areas [24].

1.4 Twitter

Twitter has been used to find and examine whether poll data can be correlated to tweets by using sentiment analysis [25], and it has been used for earthquake detection by using semantic analysis and search queries such as ‘earthquake’ and ‘shaking’ [26]. Moreover, Twitter has been credited as being influential during the Arab Spring, a series of protests that took place across the Middle East and North African countries from 2010 [27]. Twitter has therefore been researched across a number of disciplines including public health, political science, sociology, business and management studies. Furthermore, social media platforms provide those from the field of marketing with the ability to gather insights into how consumers might communicate about products, and this is particularly useful when new products are being launched. A number of platform-based, real-time analytic tools have been built which allow organisations to monitor discussions that take place on Twitter. This highlights the social listening power of Twitter as real-time insights can be rapidly gained and understood.

Previous research has examined Twitter in relation to health topics such as dementia [28], antibiotics [29], marijuana [30], sexual risk behaviours [31], and vaccination sentiments (positive and negative views) [32]. Although previous empirical research has examined Twitter content surrounding swine flu [13, 33, 34], and Ebola [35,36], no research has conducted an in-depth qualitative analysis of how Twitter users respond during infectious disease outbreaks. This study aimed to address this gap in knowledge by conducting an in-depth thematic analysis of Twitter data related to three infectious disease outbreaks, allowing for an abstraction phase comparing the outbreaks, and evaluating the utility of Twitter data.

1.5 Study Purpose and Research Questions

A case study approach was utilized and the study operated under the pragmatic research paradigm and involved elements of interpretivism. The overall aim of this study was to provide a comparative insight into Twitter data from the peaks of Swine flu, Ebola, and the Zika virus outbreaks. More specifically, the study proposed to address the following research questions:

- What were the key discussions on Twitter during the peak of the swine flu, Zika and Ebola outbreaks?
- What similarities and differences emerged by contrasting the thematic findings of each of the outbreaks to one another?
- Does the response of users on Twitter resemble that of a moral panic and/or display evidence for the outbreak narrative?

2. Methods

2.1 Data Gathering and Filtering

Case studies were selected purposely when Google Trends was showing that there was a heightened interest in search queries for Swine flu, Ebola, and Zika respectively. Google Trends was used to identify peaks because, at the current time, there are no mechanisms to search all of Twitter for a particular keyword to examine peaks of interest. Google Trends was used because an increase in Web search queries is often associated with emerging news events [37]. As noted earlier, Twitter is known as a platform for the detection of breaking news [38]; therefore, if there is an increase in Google Trends it would not be unreasonable to assume that Twitter activity would also increase around this time. The Google Trends Score is a normalised metric based on the number of searches for a single query; however, because it is normalised, two search queries could acquire a score of 100 but have different amount of searchers [10].

2.2 Data Retrieval and Filtering for Swine flu

The initial dataset retrieved related to swine flu consisted of 214,784 tweets posted during the two-day period of April 28th and April 29th 2009 and identified using the keyword terms “swine flu”, “#SwineFlu”, and “H1N1”. As stated above, this time interval was selected because it falls when the Google Trends data showed increased social media activity in the outbreak. The approach for filtering data on Swine flu is summarised in Table 1 below. A final sample of 7,678 Tweets (10%) were included in the final sample.

Table 1 Research approach for filtering swine flu data

Stage	Total
Pre-data Cleaning	214,784
Removing Exact Duplicates	102,852
Removing Duplicates at a 60% threshold	76,783
10% sample removed for analysis	7,678

2.3 Data Retrieval and Filtering for Ebola

The initial dataset that was retrieved relating to Ebola consisted of 181,110 tweets produced during the period of 29th and 30th September 2014 identified using the keyword “Ebola”. Again, Google Trends data showed an increase in interest around Ebola web-search queries during that time. The approach that was taken for filtering data on Ebola is summarised in Table 2 below. A final sample of 5,695 Tweets was obtained.

Table 2 Research approach for filtering Ebola data

Stage	Total
Pre-data Cleaning	181,110
Removing Duplicates	102,852
Removing near Duplicates at a 60% threshold	76,782
10% sample removed for analysis	5,695

2.4 Data Retrieval and Filtering for Zika

The data that were retrieved on the Zika outbreak contained 749,131 tweets, considerably more than the number of tweets on Swine flu and Ebola. However, when near duplicate clusters were removed, the dataset had the largest reduction of duplicate content. It appeared that there were a large number of news articles shared and fewer personal views and opinions were shared on the Zika outbreak. This aligns with research that has found that there was low knowledge of Zika in the United States [40]. This is summarised in Table 3 below. The final sample contained 2,042 Tweets.

Table 3 Research approach for filtering Zika data

Stage	Total
Pre-data Cleaning	749,131
Removing Exact Duplicates	76,943
Removing Duplicates at a 60% threshold	20,421
10% sample removed for analysis	2,042

2.5 Analysis Techniques and Reliability Measures

After the filtering stage, data were entered into Nvivo and the six stages of thematic analysis (Braun and Clarke, 2006) were utilised in order to analyse the data. For the data on swine flu, inter-coder reliability was checked and was 99.96%, and for Ebola it was 99.93%. This was calculated by sourcing a coder who coded 300 tweets for each dataset. In regards to test-retest reliability for tweets on swine flu the agreement rate was 99.94%. for Ebola it was 99.94%, and for Zika it was 99.24%.

Ethics approval for this study was obtained from the University of Sheffield in accordance with its research ethics policy (https://www.sheffield.ac.uk/rs/ethicsandintegrity/ethicspolicy/general-principles/homepage).

3. Results

Overall, it was found that the information which was shared on Twitter during this time period revolved around eight key themes and a number of sub-themes as described in Table 4 below.

Table 4 Thematic Findings for swine flu

Theme (N/%)	Sub-themes (N/%)
A. Emotion and feeling (253/4.4%)	A.1 General Fear (174/3.0%) A.2 Fear of Travel (54/0.9%) A.3 Anger (17/0.3%) A.4 Worry (8/0.1%)
B. Health Information (609/10.6%)	B.1 Transmission (22/0.4%) B.2 Prevalence Monitoring (158/2.8%) B.3 Prevention Techniques (134/2.3%) B.4 Prevention Products (126/2.2%) B.5 Symptoms (80/1.4%) B.6 Speculative Diagnosis (18/0.3%) B.7 Medication (14/0.2%) B.8 References to Other Infection or Disease (57/1.0%)
C. General commentary & Resources (2467/43.0%)	C.1 General Discussions (1826/31.8%) C.2 Information Seeking (145/2.5%) C.3 Economic Impact of Swine Flu (62/1.1%) C.4 Voice of Reason (109/1.9%) C.5 Frightening Scenarios (13/0.2%) C.6 Name Discussion (26/0.5%) C.7 Resources (42/0.7%) C.8 Images used in Tweets (36/0.6%) C.9 Unfollowing Users (2/0.03) C.10 Other Discussions (206/3.6%)
D. Media and Health Organisations (675/11.80%)	D.1 Health Organisations (general) (136/2.4%) D.2 Health Organisations (critical) (7/0.1%) D.3 Media Organisations (general) (444/7.7%) D.4 Media Organisations (critical) (88/1.5%)
E. Politics (124/2.2%)	E.1 Political Reference (81/1.4%) E.2 Obama (43/0.75%)
F. Country of Origin (Mexico/Travel) (211/3.7%)	F.1 Reference to Mexico and/or Mexico City (162/2.8%) F.2 Reference to Mexicans (43/0.8%) F.3 Reference to Borders (6/0.10%)
G. Food (428/7.5%)	G.1 Pork Consumption (336/5.9%) G.2 Food Humour (92/1.6%)
H. Humour or Sarcasm (975/ 17.0%)	H.1 Humour Related to Pigs (100/1.8%) H.2 Nervous Humour (18/0.3%) H.3 Popular Culture/Understanding (221/3.9%) H.4 Miscellaneous Humour (378/6.6%) H.5 Sarcasm (258/4.5%)

Table 4 shows that Twitter users discussed a number of topics around eight key themes. The most frequent discussion revolved around very general commentary and resources (43.0%). An interesting observation in this case study was the confusion that the term ‘swine flu’ created among users. The use of humour and/or sarcasm was high (17.1%) and the proportion that were dominated by media and health organisations in the discussion on the platform was also high (11.3%).

Overall, the main finding was that discussions on Twitter involving Ebola revolved around eight key themes, and a number of sub-themes as highlighted in Table 5.

Table 5 Thematic Findings for Ebola

Theme (N/%)	Sub-themes (N/%)
E. Emotion and Feeling (113/2.60%)	E.1 Anger (12/0.3%) E.2 Fear (55/1.3%) E.3 Fear of travel (5/0.11%) E.4 Praying, Prayer or call to God (26/0.60%) E.5 Dead rising generates fear (15/0.35%)
F. Health Information (192/4.5%)	F.1 Transmission Reporting (41/1.00%) F.2 Transmission of Ebola (26/0.60%) F.3 Symptoms (37/0.90%) F.4 Vaccines (36/0.80%) F.5 Prevention (22/0.51%) F.6 Speculative Diagnosis (7/0.16%) F.7 Quarantine (25/0.60%)
G. Significant News Stories (282/6.60%)	G.1 Ebola Patients Rise from Dead (107/2.5%) G.2 Australia will not Send Volunteers (64/1.5%) G.3 US to Send Troops to Fight Ebola (18/0.42%) G.4 News Story uses Terrorism Analogy (6/0.14%) G.5 Doctor Exposed to Ebola (87/2.03%) G.6 FDA Warning Over Fake Drugs (30/0.70%)
H. General Commentary (2311/54.0 %)	H.1 General Discussions (2025/47.26%) H.2 Information Seeking (28/0.65%) H.3 Economic Impact of Ebola (11/0.25%) H.4 Death Count (32/0.74%) H.5 Western Privilege (11/0.25%) H.6 Link to Instagram (24/0.56%) H.7 Twitter Users Linking to YouTube (56/1.30%) H.8 Refers to iPhone (9/0.21%) H.9 Twitter Users Linking to Other Tweets (55/1.30%) H.10 Downplaying Ebola risk (9/0.21%) H.11 Conspiracy Theories (51/1.20%)
I. Refers to official organizations (75/1.80%)	I.8 WHO (17/0.40 %) I.9 CDC (37/0.90%) I.11 MSF (12/0.30 %) I.12 UNICEF (9/0.21%)
J. Refers to West African City and or region (181/4.2)	I. 1 Sierra Leone (104/2.42%) I.2 Liberia (36/0.84%) I.3 Nigeria (33/0.80%) I.4 Guinea (8/0.2%)
K. Political References (88/2.05%)	I. 1 Obama (68/1.58%) I.2 Julie Bishop (7/0.16%) I.3 Critical of Governments (13/0.30%)
L. Humour or Sarcasm (1046/24.41%)	J.1 Sarcasm (425/9.9%) J.2 Humour (418/9.8%) J.3 Zombies (77/1.8%) J.4 Zombie Apocalypse (18/0.42%) J.5 Ebola Used as an Insult (108/2.52%)

Similar to Table 4, Table 5 shows that Twitter users conversed on a diverse range of topics, which were based around eight key themes. The most frequent of these discussions was based on general commentary (54.0%). In this case study it was interesting to observe a number of news stories that were significant and which made up for 6.6% of the discussion at this time. The proportion of Tweets in which humour and/or sarcasm were referred to in the Tweets was relatively high (24.4%).

Table 6 Thematic Findings for Zika

Theme (N/%)	Sub-themes (N)
1. Pregnancy (164/8.36%)	1.1 Avoid Pregnancy Narrative (12/0.61%) 1.2 Zika Threat to Pregnant Women (19/0.96%) 1.3 Zika Virus Spreads Fear Among Pregnant Brazilians (22/1.12%) 1.4 Zika Threat to Pregnant Columbians (29/1.47%) 1.5 Abortion Debate (47/2.39%) 1.6 Pregnancy (35/1.78%)
2. Olympics (75/3.82%)	2.1 Fear (19/0.96%) 2.2 Olympics Rio 2016 (56/2.85%)
3.Mosquitoes and Conspiracy (259/13.21%)	3.1 Zika Conspiracy (61/3.11%) 3.2 GM Mosquitoes (65/3,31%) 3.3 Mosquitoes (133/6.78%)
4. Health Organisations (365/18.62%)	4.1 Critical of WHO (7/0.35%) 4.2 WHO Related News (358/18.26%)
5. Health Information (160/8.16%)	5.1 Zika Origin (8/0.40%) 5.2 Zika Symptoms (10/0.51%) 5.3 Transmission (19/0.96%) 5.4 Prevention (34/1.73%) 5.6 Zika Vaccine (37/1.88%) 5.7 Microcephaly (52/2.65%)
6. Travel and Tracking (321/16.37%)	6.1 Zika Travel Advice (29/1.47%) 6.2 Zika Spreading Explosively in South and Central America (14/0.71%) 6.3 Zika will Spread Across the Americas (17/0.86%) 6.4 Mentions Brazil and Zika Virus (19/0.96%) 6.5 Geographical Tracking (21/1.07%) 6.6 Geographical Transmission (116/5.91%) 6.7 Travel (45/2.29%)
7. General Discussions (646/32.95%%)	7.1 Zika found in Ugandan Forest (15/0.76%) 7.2 Zika and Climate Change or Global Warming (18/0.91%) 7.3 Broadcast Advert (21/1.07%) 7.4 Information Seeking (26/1.32%) 7.5 Politics (26/1.32%) 7.6 Humour and Sarcasm (77/3.92%) 7.7 Name Discussion (124/6.32%) 7.8 Zika Information (General) (363/18.52%)

Table 6 shows the results of the themes from the analysis of the Zika case and the results were noticeably different to those from swine flu and Ebola. The discussion revolved around 7 key themes and topics such as travel (16.4%), health organisations (18.6%), mosquitoes and conspiracy (13.2%), pregnancy (8.4%), and the Olympics (3.8%). Similar to the results of swine flu and Ebola, this case study also contained a substantial proportion of very general comments and discussions (33.0%).

Based on the results presented in Tables 4 to 6, it is possible to compare the thematic findings. The results of Zika appeared to be noticeably different from that of swine flu and Ebola; henceforth, the figure below compares swine flu and Ebola only and comparisons with Zika were conducted separately. Figure 1 below provides a clear overview of the similarities and differences that were found in each of the cases..

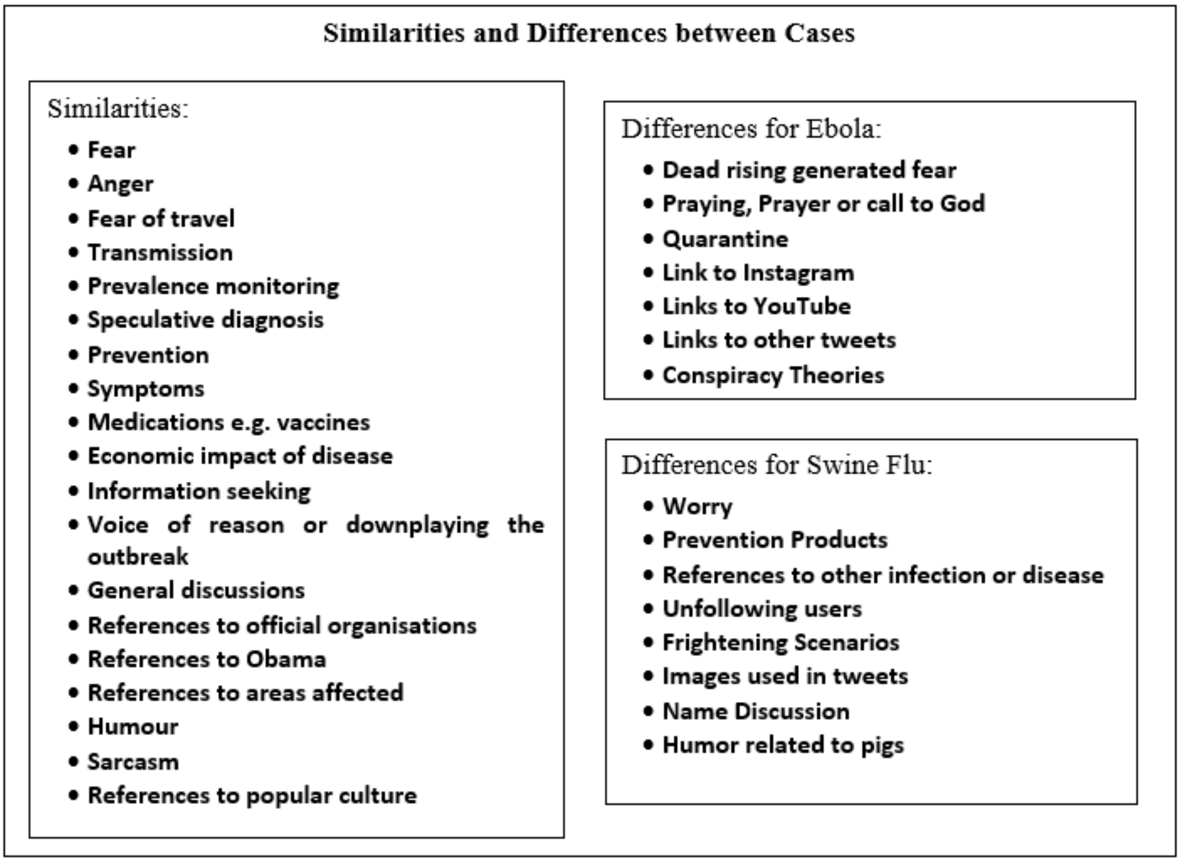

Figure 1 Similarities in themes between swine flu and Ebola

Overall, this comparison shows that there are a number of similarities as well as differences in the way in which Twitter users responded during the 2009 swine flu pandemic the 2014 Ebola epidemic and the Zika virus outbreak. Figure 1 above shows that there were similarities across at least 19 themes and the outbreaks evoked a similar response from Twitter users. Some of the differences in Figure 1 may have been due to the characteristics of the diseases and to specific events. For example, with swine flu, there were discussions surrounding its name and the potential to cause confusion. In contrast, with Ebola, the story that people who had died had come back to life was specific to events that occurred during that particular outbreak. Other differences related to the availability and popularity of video sharing and image sharing platforms such as Instagram and YouTube.

Zika was compared separately to that of swine flu and Ebola, as it was found to trigger virus-specific conversations. More specifically, the similarities for Zika primarily related to the theme of fear, mentions of the WHO, political references, transmission, prevention, and travel. Twitter users also discussed the name of Zika, as was the case with swine flu. However, discussions around Zika were significantly distinct from those for Ebola and Swine flu, and there were a number of interesting themes specific to the Zika outbreak, as follows:

- Avoid Pregnancy Narrative
- Zika Threat to Pregnant Women
- Zika virus Spreads Fear Among Pregnant Brazilians
- Zika threat to pregnant Columbians
- Abortion Debate
- Pregnancy
- Olympics Rio 2016
- Mosquitoes
- Mosquitoes
- Microcephaly
- Zika Spreading explosively in South and Central America

- Zika will Spread Across the Americas
- Mentions Brazil and Zika Virus

As mentioned previously, many of these differences in themes arose because the content within the tweets were related to the specific events and debates surrounding the Zika outbreak.

4. Discussion

This study utilised Twitter data to provide in-depth qualitative insights into the swine flu, Ebola and Zika outbreaks. In the current body of literature, it appears that the field of evidence-based health research has validated the use of Twitter for gaining real-time insights [41]. However, there appears to have been very little work examining the effectiveness of Twitter for providing in-depth qualitative insights into health topics. The difference between using quantitative methods, over qualitative methods, on Twitter data is that it may involve finding correlations between tweets and incidence rates, or that a large number of tweets are automatically processed and frequently occurring keywords are displayed. Qualitative Twitter studies provide insight into the tweets within the context in which they have been written, by using research methods associated with qualitative research: only a limited number of studies have utilised in-depth qualitative methods to analyse Twitter data [42, 43, 44].

The use of Twitter for gaining in-depth insights for qualitative research more generally was noted by Marwick [46] in an article entitled Ethnographic Qualitative Research on Twitter, which summarised a number of textual analysis methods. In 2017, Sinnenberg published a systematic review examining Twitter as a tool for health research. A total of 137 studies were considered, and it was found that 56% of them utilised content analysis to analyse data, 26% utilised surveillance methods, 14% engagement methods, 7% used Twitter for recruitment, 7% were based on intervention, and 4% were related to network analysis [41]. The author noted that there should be standardised reporting guidelines and polices which reflect more on the ethical issues associated with researching Twitter. However, the study did not consider whether Twitter could be utilised to gain in-depth insights into health. Moreover, the study did not distinguish content analysis and thematic analysis when assessing the types of methods that were used to analyse Twitter data. This is important because thematic analysis is more associated with qualitative research. A systematic review by Hu [47] examined health communication research on digital platforms. This study found that research that utilises quantitative methods is more common, with 67% of published studies in comparison to 20% of articles which utilise qualitative research methods [47]. Therefore, much of the previous research on Twitter may have been based on quantitative methods, which could explain why previous health research has not examined whether Twitter data has the potential to provide in-depth insights into health.

We predict that for future infectious disease outbreaks public information needs will revolve around seeking information related to the safety of travel, transmission, prevention, and symptoms, and indeed, news stories on the current Coronavirus outbreak reflect these areas. Much of this information would be provided by public health authorities and governments. In the case of swine flu, we observed insensitive comments towards Mexico and those of Mexican descent, as the virus was believed to have originated from Mexico at the time. During the coronavirus (2019-nCoV) there have been several reports of Chinese citizens in the United Kingdom being mistreated (Campbell, 2020). Governments and local authorities could work to educate the public in such cases via various outreach campaigns.

4.1 Moral Panic

A well-known concept in sociology is known as the 'moral panic' and this occurs when:

> A condition, episode, person or group of persons emerges to become defined as a threat to societal values and interests (Cohen, 2002: p.1) [48]

It could be argued that during the peak of the swine flu, Ebola, and Zika outbreaks a moral panic was underway, and Twitter users were caught up in this. The tweets from the fear theme can be used in support of this argument. In fact, there appeared to be an exaggerated fear from Twitter users across all of the themes, particularly in the discussions surrounding the possibility of patients rising from the dead, and of the potential of a 'zombie apocalypse'. This reaction on Twitter, coupled with increased media attention surrounding the outbreak, could resemble a moral panic. Another defining factor of moral panic is the exaggeration of an episode by mass media, and in the outbreaks of swine flu and Ebola, there were articles shared on Twitter that appeared to sensationalise the outbreaks. Looking back now at the outbreaks, sometime after they have occurred and with the benefit of hindsight, it is easy to say that Twitter users, the general public, and the media might have over-reacted but, at the time, reactions were seen as appropriate in order to make people aware, and as the authorities became aware, of a potential global health threat. Moreover, it must be noted that social media as a platform in general may work to inflame fears on any range of events.

4.2 Outbreak Narrative

One interesting aspect to note across the results of the study was the frequency of references to popular culture, zombies, and the potential of a zombie apocalypse. These surprising finding were unreported in previous evidence-based research on Twitter in the context of the outbreaks. It is worth drawing attention to media and culture literature in order to understand these phenomena. In certain cases, it might be in the best interest of the media, especially unregulated online websites, to offer shocking narratives intentionally in order to generate page views and link clicks. This was found to be the case with the Ebola outbreak, as it could be argued that the news article on Ebola patients rising from the dead was intended to shock readers. Ostherr [49] and Wald [50] noted theoretically that outbreak narratives could have an influence on the opinions of the public, and the present study provides empirical evidence that Hollywood narratives were indeed able to influence how some Twitter users were experiencing and understanding the outbreak. Ostherr and Wald also noted that the narratives displayed in television and film might lead to negative public reactions to pandemics, and the media reports which are excessive and/or pessimistic might cause members of the public to become fearful. Importantly, they noted that narratives that provide users with information that is inaccurate might negatively affect the behaviour of the public and especially of individuals who may have limited or no formal education. Therefore, health authorities and governments should be aware of narratives surrounding infectious disease outbreaks, and respond swiftly, and appropriately, when a potential harmful narrative is shared, especially when it contains misleading or inaccurate information. Twitter data could act as a platform that could be studied by those from disciplines such as English literature and/or history, could bring more varied methodologies to analyzing texts.

4.3 Study Limitations

The study examined two-day time intervals from when there was heightened interest surrounding Swine flu, Ebola, and Zika therefore, there could be limitations in the conclusions that can be drawn from the data. This study concentrated on examining tweets in English, and therefore it is not a complete record of all users that were tweeting about the outbreaks, as other languages were not considered. It may appear that the number of tweets in the fear category was low; however, tweets across a number of themes appeared to contain an element of fear to support our assertion that a moral panic was underway. Another related limitation is that the study may not have retrieved all data from Twitter related to the outbreaks because certain users may have been talking about the outbreaks without mentioning them. The study analysed tweet content and did not categorize and/or analyse images or videos in tweets. Further research could seek to include investigating the role of multi-media elements in social media disease outbreaks.

5. Conclusions

It can be argued that social media platforms have become just as influential to certain groups as conventional media and is particularly relevant to study from a public health perspective [51,52]. For infectious disease outbreaks, incorrect information can have dangerous consequences which makes it very important to study information that is shared on the platforms. As, for example, research around the novel coronavirus (2019-nCoV) on social has found the existence of various conspiracy theories [53, 54]. Our study has developed new insight into how users respond during infectious disease outbreaks, identified potentially commonalities, and reflected on users' response in association with the sociological concept of the moral panic.

References

[1] Taylor, L. H., Latham, S. M., & Mark, E. J. (2001). *Risk factors for human disease emergence*. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 356(1411), 983-989.
[2] Anon. 1996. The Global burden of disease : a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020 : summary / edited by Christopher J. L. Murray, Alan D. Lopez. (January 1996). Retrieved May 31, 2018 from http://apps.who.int/iris/handle/10665/41864.
[3] Haensch, S., Bianucci, R., Signoli, M., Rajerison, M., Schultz, M., Kacki, S., ... & Carniel, E. (2010). *Distinct clones of Yersinia pestis caused the black death*. PLoS pathogens, 6(10), e1001134.
[4] Taubenberger, J. K., & Morens, D. M. (2006). *1918 Influenza: the mother of all pandemics.* Emerging infectious diseases, 12(1), 15.
[5] Bostrom, N. (2008). Global Catastrophic Risks. Oxford: Oxford University Press.
[6] Smith, R. D. (2006). *Responding to global infectious disease outbreaks: lessons from SARS on the role of risk perception, communication and management.* Social science & medicine, 63(12), 3113-3123.
[7] Landsman, S. (2020). Coronavirus is the biggest market threat right now, Wall Street bull Ed Yardeni says. [online] CNBC. Available at: https://www.cnbc.com/2020/02/09/coronavirus-is-the-biggest-threat-to-rally-right-now-ed-yardeni-warns.html [Accessed 11 Feb. 2020].
[8] McCandless, D. (2009). Information is beautiful (pp. 1–255). London: Collins.
[9] Google Zeitgeist. (2009). Google. Retrieved 9 July, 2017, from https://archive.google.com/intl/en/press/zeitgeist2009/.
[10] Google. (2017). Google Trends. Retrieved 14 May, 2017, from https://trends.google.co.uk/trends/.
[11] Ahmed, W. (2018). Using Twitter data to provide qualitative insights into pandemics and epidemics (Doctoral dissertation, University of Sheffield).
[12] Telegraph (2017, November 7). Twitter is letting everyone post 280 character tweets. The Telegraph. Retrieved from http://www.telegraph.co.uk/technology/2017/11/07/twitter-letting-everyone-post-280-character-tweets/
[13] Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PloS one, 5(11), e14118.
[14] Zhang, Z., & Ahmed, W. (2019). *A comparison of information sharing behaviours across 379 health conditions on Twitter.* International journal of public health, 64(3), 431-440.
[15] Davis, C. P. (2015). *Swine Flu: Facts, Symptoms, Treatment and H1N1 Vaccine*. MedicineNet. Retrieved 11 October, 2015, from http://www.bibme.org/citation-guide/apa/website.
[16] NHS Choices [NHS]. (2015). Swine flu (H1N1). Retrieved 20 October, 2015, from http://www.nhs.uk/conditions/pandemic-flu/Pages/Introduction.aspx.
[17] Wang, T. T., & Palese, P. (2009). *Unraveling the mystery of swine influenza virus*. Cell, 137(6), 983-985.
[18] Simonsen, L., Spreeuwenberg, P., Lustig, R., Taylor, R. J., Fleming, D. M., Kroneman, M., ... & Paget, W. J. (2013). *Global mortality estimates for the 2009 Influenza Pandemic from the GLaMOR project: a modeling study.* PLoS medicine, 10(11), e1001558.
[19] Gulland, A. (2016). Ebola outbreak in west Africa is officially over. i243. DOI:http://dx.doi.org/10.1136/bmj.i243.
[20] Frieden, T. R., Damon, I., Bell, B. P., Kenyon, T., & Nichol, S. (2014). *Ebola 2014—new challenges, new global response and responsibility.* New England Journal of Medicine, 371(13), 1177-1180.
[21] BBC News (2016, January 14). Ebola: Mapping the outbreak. Retreived from http://www.bbc.co.uk/news/world-africa-28755033.
[22] Ngade, I., Singer, M., Marcus, O., & Lara, J. E. H. (2016). Washing Away Ebola. A Companion to the Anthropology of Environmental Health, 157–172.
[23] Fauci, A. S., & Morens, D. M. (2016). *Zika virus in the Americas—yet another arbovirus threat.* New England Journal of Medicine, 374(7), 601-604.
[24] Brasil, P., Pereira Jr, J. P., Moreira, M. E., Ribeiro Nogueira, R. M., Damasceno, L., Wakimoto, M., ... & Zin, A. A. (2016). *Zika virus infection in pregnant women in Rio de Janeiro.* N Engl J Med, 2016(375), 2321-2334.
[25] O'Connor, B., & Balasubramanyan, R. (2010). From tweets to polls: Linking text sentiment to public opinion time series. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (pp. 122–129). Washington, DC: AAAI.

[26] Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). *Earthquake shakes Twitter users.* Proceedings of the 19th International Conference on World Wide Web (pp. 851). New York: ACM Press.
[27] Howard, P. N., Duffy, A., Freelon, D., Hussain, M. M., Mari, W., & Maziad, M. (2011). Opening closed regimes: What was the role of social media during the Arab Spring? Project on Information Technology and Political Islam Data Memo 2011.1. Seattle: University of Washington.
[28] Robillard, J. M., Johnson, T. W., Hennessey, C., Beattie, B. L., & Illes, J. (2013). *Aging 2.0: Health information about dementia on Twitter.* PloS ONE, 8(7), p.e69861. doi: 10.1371/journal.pone.0069861.
[29] Scanfeld, D., Scanfeld, V., & Larson, E. L. (2010). *Dissemination of health information through social networks: Twitter and antibiotics.* American Journal of Infection Control, 38(3), 182–8. doi: 10.1016/j.ajic.2009.11.004.
[30] Cavazos-Rehg, P. A., Krauss, M., Fisher, S. L., Salyer, P., Grucza, R. A., & Bierut, L. J. (2015). *Twitter chatter about marijuana.* The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 56(2), 139–45. doi: 10.1016/j.jadohealth.2014.10.270.
[31] Young, S. D., Rivers, C., & Lewis, B. (2014). *Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes.* Preventive Medicine, 63, 112–5. doi: 10.1016/j.ypmed.2014.01.024.
[32] Salathé, M., & Khandelwal, S. (2011). *Assessing vaccination sentiments with online social media: Implications for infectious disease dynamics and control.* PLoS Computational Biology, 7(10), p.e1002199. doi: 10.1371/journal.pcbi.1002199.
[33] Signorini A., Segre A. M., Polgreen P. M. (2011) *The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic.* PLoS ONE 6(5), p.e19467. doi: 10.1371/journal.pone.0019467.
[34] Szomszor, M., Kostkova, P., & St Louis, C. (2011). *Twitter informatics: Tracking and understanding public reaction during the 2009 Swine Flu pandemic.* Proceedings of the 2011 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2011 (Vol. 1, pp. 320–323). New York: ACM Press. doi:10.1109/WI-IAT.2011.311.
[35] Jin, F., Wang, W., Zhao, L., Dougherty, E., Cao, Y., Lu, C.T., & Ramakrishnan, N. (2014), *Misinformation Propagation in the Age of Twitter.* Computer, 47(12), 90–94.
[36] Odlum, M., & Yoon, S. (2015). *What can we learn about the Ebola outbreak from tweets?* American Journal of Infection Control, 43(6), 563–71. doi: 10.1016/j.ajic.2015.02.023.
[37] Carneiro, H. A., & Mylonakis, E. (2009). *Google trends: A web-based tool for real-time surveillance of disease outbreaks.* Clinical infectious diseases, 49(10), 1557–1564.
[38] Phuvipadawat, S., & Murata, T. (2010). *Breaking news detection and tracking in Twitter.* In Proceedings Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010. IEEE/WIC/ACM International Conference 3, 120–123: IEEE.
[39] Rasmussen, S. A., Jamieson, D. J., Honein, M. A., & Petersen, L. R. (2016). *Zika virus and birth defects — reviewing the evidence for causality.* New England Journal of Medicine, 374(20), 1981-1987.
[40} Paulus, T., Woods, M., Atkins, D. P., & Macklin, R. (2017). *The discourse of QDAS: Reporting practices of ATLAS. ti and NVivo users with implications for best practices.* International Journal of Social Research Methodology, 20(1), 35–47.
[41] Sinnenberg, L., Buttenheim, A. M., Padrez, K., Mancheno, C., Ungar, L., & Merchant, R. M. (2017). *Twitter as a tool for health research: A systematic review.* American Journal of Public Health (ajph).
[42] Burch, L. M., Frederick, E. L., & Pegoraro, A. (2015) *Kissing in the Carnage: An Examination of Framing on Twitter During the Vancouver Riots.* Journal of Broadcasting & Electronic Media, 59(3), 399–415. doi: 10.1080/08838151.2015.1054999.
[43] Hewis, J. (2015). *Do MRI Patients Tweet? Thematic Analysis of Patient Tweets About Their MRI Experience.* Journal of Medical Imaging and Radiation Sciences. 46(4), 396–402.
[44] Shepherd, A., Sanders, C., Doyle, M., & Shaw, J. (2015). *Using social media for support and feedback by mental health service users: thematic analysis of a twitter conversation.* BMC psychiatry, 15(1), 29.
[45] Branthwaite, A., & Patterson, S. (2011). *The power of qualitative research in the era of social media.* Qualitative Market Research: An International Journal, 14(4), 430–440.
[46] Marwick, A. (2013). Ethnographic and Qualitative Research on Twitter. In: K. Weller,
A. Bruns, C. Puschmann, J. Burgess, J. & M. Mahrt (Eds.), Twitter and Society, (pp. 109–122). New York: Peter Lang.
[47] Hu, Y. (2015). *Health communication research in the digital age: A systematic review.* Journal of Communication in Healthcare, 8(4), 260–288.

[48] Cohen, S. (2004). Folk Devils and Moral Panics (3rd edn). London: Routledge.
[49] Ostherr, K. (2005). Cinematic prophylaxis: Globalization and contagion in the discourse of world health. Duke University Press.
[50] Wald, R., Khoshgoftaar, T. M., Napolitano, A., & Sumner, C. (2013). *Which users reply to and interact with twitter social bots?* In Proceedings Tools with Artificial Intelligence (ICTAI), November 2013 IEEE 25th International Conference (pp. 135–144). IEEE.
[51 Ahmed, W., Bath, P. A., Sbaffi, L., & Demartini, G. (2019). *Novel insights into views towards H1N1 during the 2009 Pandemic: a thematic analysis of Twitter data.* Health Information & Libraries Journal, 36(1), 60-72.]
[52] Ahmed, W., Bath, P. A., Sbaffi, L., & Demartini, G. (2018, July). *Moral panic through the lens of Twitter: An analysis of infectious disease outbreaks.* In Proceedings of the 9th International Conference on Social Media and Society (pp. 217-221).
[53] Ahmed, W., Vidal-Alaball, J., Downing, J., & Seguí, F. L. (2020). *COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data.* Journal of Medical Internet Research, 22(5), e19458.
[54] Ahmed W, López Seguí F, Vidal-Alaball J, Katz MS. *COVID-19 and the "Film Your Hospital" Conspiracy Theory: Social Network Analysis of Twitter Data.* JMIR Preprints. 10/07/2020:22374 DOI: 10.2196/preprints.22374. URL: https://preprints.jmir.org/preprint/22374