TRENDING HEALTH NARRATIVES:

THE ZIKA VIRUS OUTBREAK 2015-16 IN BRAZIL

by

Luiz Henrique Moura Lopes

B.A., Universidade Federal de Alagoas, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF JOURNALISM

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2017

© Luiz Henrique Moura Lopes, 2017
Abstract

This study explores the intersections of social media and health communication, focusing on the trending narratives on Twitter during the 2015-16 Zika virus outbreak, in Brazil. It examines how social media data provides insights into current media practices, as a way of improving health communication strategies and literacy during epidemics. This project uses a qualitative study framework to illustrate the phenomenon under examination, drawing a theoretical thematic analysis as a method (Braun & Clarke, 2006). Findings from this thesis reveal how increases in Zika-related news commentaries seemed to have shifted the topics and the tone of Zika-related conversations on Twitter. This project further explores the interconnectedness between the volume of Zika-related mentions on Twitter and external media events, which echoes relevant scholarship (Fu et al., 2016; Dredze, Broniatowski & Hilyard, 2016). Finally, it also discusses the opportunities and limitations of social media for health communication, addressing traditional inequalities and power dynamics.
Preface

This thesis is original, unpublished, independent work by the author, Luiz Henrique Moura Lopes. For all figures, copyright permission was not required, as the Twitter User Agreement makes content submitted to or through its services available to the rest of the world with a worldwide, non-exclusive, royalty-free license.

Dr. Taylor Owen, Professor at the University of British Columbia School of Journalism, and Dr. Jerry Spiegel, Professor at the University of British Columbia School of Population and Public Health, Faculty of Medicine, approved this thesis on 13 April 2015.
Table of Contents

Abstract .................................................................................................................................................. ii
Preface .................................................................................................................................................. iii
Table of Contents ................................................................................................................................ iv
List of Figures ......................................................................................................................................... vi
List of Abbreviations ........................................................................................................................ vii
Acknowledgements ............................................................................................................................ viii
Dedication ............................................................................................................................................... ix
Chapter 1: Introduction ......................................................................................................................... 1
Chapter 2: A Review of Literature ......................................................................................................... 3
  2.1 What is the Zika virus: overview of the disease and previous outbreaks ................................. 3
  2.2 The 2015-2016 outbreak ................................................................................................................ 7
  2.3 The timeline of events and the Brazilian media reaction ............................................................ 12
  2.4 Social networking ........................................................................................................................... 15
    2.4.1 Intersections with public health .............................................................................................. 18
    2.4.1.1 The limitations of social media ......................................................................................... 22
  2.5 Chapter summary ............................................................................................................................ 23
Chapter 3: Methodology ....................................................................................................................... 26
  3.1 Rationale for qualitative research design, and methodology .................................................... 26
  3.2 Rationale for method ...................................................................................................................... 27
    3.2.1 Overview of method phases .................................................................................................. 29
  3.3 Data collection tool ....................................................................................................................... 31
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>The research sample</td>
<td>33</td>
</tr>
<tr>
<td>3.5</td>
<td>Overview of the research design</td>
<td>34</td>
</tr>
<tr>
<td>3.6</td>
<td>Opportunities and limitations</td>
<td>35</td>
</tr>
<tr>
<td>3.7</td>
<td>Chapter summary</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 4: Findings</strong></td>
<td>38</td>
</tr>
<tr>
<td>4.1</td>
<td>General findings</td>
<td>38</td>
</tr>
<tr>
<td>4.2</td>
<td>Thematic findings</td>
<td>44</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Proportion and distribution of topics in the analysis</td>
<td>48</td>
</tr>
<tr>
<td>4.3</td>
<td>Chapter summary</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 5: Discussion</strong></td>
<td>54</td>
</tr>
<tr>
<td>5.1</td>
<td>How can social media data provide insights into the role of the media during global health crisis?</td>
<td>55</td>
</tr>
<tr>
<td>5.2</td>
<td>How does social power play on social media?</td>
<td>56</td>
</tr>
<tr>
<td>5.3</td>
<td>What were the trending conversations on Twitter during the Zika virus epidemic in Brazil?</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 6: Final considerations</strong></td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>62</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1 – Google Trends data showing interest over time of the words Zika (blue) and Zica (red) ................................................................................................................................. 14
Figure 2 – Word cloud representing the words used in the research sample .............................................. 40
Figure 3 – Clusters representing the words used in the research sample ................................................. 41
Figure 4 – Chart representing the volume of Zika-related posts over the chosen date range, by day ......................................................................................................................... 42
Figure 5 – Tweet representing an example of the news commentary category ........................................... 45
Figure 6 – Tweet representing an example of the case reporting category ................................................. 46
Figure 7 – Tweet representing an example of the general positive category ............................................. 46
Figure 8 – Tweet representing an example of the general neutral category .............................................. 47
Figure 9 – Tweet representing an example of the general negative/fear category ................................. 48
Figure 10 – Tweet representing an example of the off-topic/bad luck category ........................................ 48
Figure 11 – Chart representing the proportion of Zika-related posts over the chosen date range, by color-coded category........................................................................................................ 50
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>ECDC</td>
<td>European Centre for Disease Prevention and Control</td>
</tr>
<tr>
<td>FIFA</td>
<td>Fédération Internationale de Football Association</td>
</tr>
<tr>
<td>FIOCRUZ</td>
<td>Oswaldo Cruz Foundation</td>
</tr>
<tr>
<td>IHR</td>
<td>International Health Regulations</td>
</tr>
<tr>
<td>PAHO</td>
<td>Pan American Health Organization</td>
</tr>
<tr>
<td>PHEIC</td>
<td>Public Health Emergency of International Concern</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>SUS</td>
<td><em>Sistema Único de Saúde</em></td>
</tr>
<tr>
<td>Tdap</td>
<td>Tetanus, Diphtheria, and Pertussis Vaccine</td>
</tr>
<tr>
<td>U.S.</td>
<td>United States of America</td>
</tr>
<tr>
<td>UN</td>
<td>United Nations</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Acknowledgements

I thank the faculty, staff and fellow students at UBC Graduate School of Journalism, who have inspired me in every way.

I thank my supervisors Dr. Taylor Owen and Dr. Jerry Spiegel, who guided me with their vast expertise and continuous encouragement.

I thank professors Peter Klein and Kathryn Gretsinger, who are truly inspiring individuals.

I thank the Faculty of Arts, the Gwyn and Aileen Gunn Bursary in Journalism, and the Alan Gayfer Memorial Bursary, whose financial support was essential.

I thank the Social Media Advanced Research Teaching and Training Lab (SMARTTLab), whose resources and staff were critical in this thesis.

Special thanks to my family for providing me with unfailing support.

Special thanks to my partner for providing me with continuous encouragement.

Thank you.
Dedication

To Journalism…
Chapter 1: Introduction

This study seeks to explore the online trending narratives during the 2015-16 Zika outbreak, in Brazil. The purpose of this study is to understand the role of social media during a global health crisis and how it relates to broader health communication strategies. My assumption is that a better understanding of Internet-based tools and social networking would allow journalists and global health actors to have a more informed perspective in terms of reaching and educating mass audiences.

This research employed a qualitative study framework to illustrate the phenomenon under examination, drawing on a theoretical thematic analysis (Braun & Clarke, 2006). The data collection was achieved through a tool developed by Crimson Hexagon to scrape social media data over a chosen time frame, languages and location. I used the tool to hand-code a set of textual contents (tweets) into categories, which were predetermined by me. After this sorting process, Crimson Hexagon’s algorithm generated the analysis results of 2,581,249 tweets, and I critically discussed them.

The results identified a set of general and thematic findings, which will be further explored in the Discussion chapter. The findings were discussed with a focus on this project’s main research questions, which were (1) How can social media data provide insights into the role of the media during global health crisis? (2) How does social power play on social media? (3) What were the trending conversations on Twitter during the Zika virus epidemic in Brazil?

The rationale for this study stems from my interests in exploring the intersections of social media and health communication. Increased understanding of the role of the media during a public health crisis may not only shed a light on current practices and processes of shaping the
public opinion, but also increase the potential for improving health communication strategies and literacy.

This project also acknowledges the traditional inequalities in overall Internet access, as well as the limitations of using social networking for health communication (Moorhead et al., 2013; Dredze, Broniatowski & Hilyard, 2016; Orizio et al., 2010; Adams, 2010). However, the worldwide increase in the use of social media (boyd & Ellison, 2007) requires a look into these Internet-based tools as a way of promoting health equality, as well as improving health communication practices and surveillance.
Chapter 2: A Review of Literature

The purpose of this case study is to explore the online trending narratives on social media in Brazil during the 2015-16 Zika outbreak. Specifically, I sought to understand the role of social media during a global health crisis and how it relates to broader health communication strategies. Therefore, to carry out this study, it was necessary to complete an ongoing critical review of (1) current literature on the Zika virus disease and outbreaks; (2) social media; and (3) health communication. This review takes as a premise the interconnectedness of social networking, media, social power, and global health.

To conduct this selected literature review, I used multiple information sources, including books, dissertations, newspapers and other internet resources, professional journals, and periodicals. There was no specific temporal period guiding the scope of this review, because of the historical nature of the topics, but priority was given to recently published sources or at least the ones published around the time frame of the 2015-16 Zika virus epidemic. I also attempted to reduce bias and allow for a cross-cultural perspective by including research published in other languages, particularly Portuguese.

This review identifies the scholarly context needed to ground the project’s main research questions, which focus on the trending online Zika-related conversations to help understand how social media data can provide insights into the role of the media during global health crisis.

2.1 What is the Zika virus: overview of the disease and previous outbreaks

The Zika virus was first discovered and named in 1947 near Entebbe, Uganda, when it was identified in samples taken from a sentinel rhesus monkey (Macacamulatta) in the Zika Forest, as part of a research project on jungle yellow fever (Dick, Kitchen & Haddow, 1952; see
also Sikka et al., 2016). For six decades, there were no outbreaks and only 14 cases of Zika virus disease had been previously documented in humans across Africa and Asia (Filipe, Martins & Rocha, 1973). The first outbreak of Zika virus disease occurred in 2007 when an estimated three-quarters of 11,250 Yap Island residents, in the Federated States of Micronesia, were infected with Zika virus. A series of house-to-house surveys among the island’s population confirmed 49 cases of suspected Zika virus disease. There were no hospitalizations or deaths reported in association with the disease. Men were more likely to have Zika infection than women across all age groups (Duffy et al., 2009).

The most likely sources of this first outbreak were attributed to an infected human or an infected mosquito, because of the abundance of the mosquito Aedes hensilli on Yap (Duffy et al., 2009; see also Savage et al., 1998). Three factors were considered by World Health Organization (WHO) researchers Kindhauser, Allen, Frank, Santhana and Dye (2016) as possible causes for the rapid spread of the virus and its epidemic behavior in Yap Island. First, the lack of population immunity in the affected area. Second, the clinic overlapping features to other pathogens such as dengue and chikungunya. Third, underreporting due to lack of surveillance systems and the often-mild nature of the Zika virus disease in adults.

The following major outbreak happened in 2013, in French Polynesia, where nearly 8,746 suspected cases of Zika virus infection and 383 confirmed cases were identified by the surveillance system of that country (Passi, Sharma, Dutta & Ahmed, 2017). Zika virus infections were also reported in other islands of the Pacific Region, at the same time as the French Polynesian outbreak. In the Cook Islands, there were 932 suspected cases and 54 confirmed cases. In New Caledonia, there were 1,400 confirmed cases. And in Easter Island, there were 89 suspected cases and 51 confirmed cases. Exposure to Zika-infected mosquitoes was likely to be
the principal risk for Zika virus infection during the French Polynesian epidemic (Roth et al., 2014).

Scientists from the Louis Malardé Institute (French Polynesia), and the WHO, Musso, Nilles and Cao-Lormeau (2014), found correlating cases in which neurological and autoimmune complications were acquired by individuals after Zika virus disease during the French Polynesian outbreak, such as Guillain-Barré syndrome, which can cause temporary paralysis in patients. However, the possibility of Zika virus as a stimulus predisposing this autoimmune disease was not confirmed at the time (Ioos, Mallet, Leparc Goffart, Gauthier & Herida, 2014).

The evidence for the link between Zika virus infection and Guillain-Barré syndrome was only presented in 2016 with retrospective data from French Polynesia, which revealed that 42 cases of Guillain-Barré syndrome also tested positive for Zika virus infection (Cao-Lormeau et al., 2016), highlighting the burden of this disease in public health services.

These two initial waves of Zika virus outbreaks in the Pacific area (2007, Yap Island; 2013, French Polynesia) were considered by leading researchers (Roth et al., 2014) as the early stages of a disease trend that would continue for several years, as there was a high risk for further spread of mosquito-borne diseases in the region due to i) little immunity because of limited or inexistent circulation of the viruses in the area; ii) the presence of local mosquitoes known to be vectors for these viruses; iii) increase in air traffic, which accelerated the spread of the mosquitoes (see also Tatem et al., 2012).

As recommendations to mitigate the disease burden on health systems and limit further spread to other parts of the world, leading scholars (Roth et al., 2014) suggested improvements in surveillance and response measures. The same study also highlighted the potential for the Zika virus infection to require international concern to support surveillance and control, as it had
potential for rapid spread. Other studies further addressed the underlying determinants of these two initial outbreaks as potential causes for the outbreak, such as the combination of socio-economic, environmental and ecological factors relating to the disease (Jones et al., 2008).

Most infections caused by the Zika virus are asymptomatic, but the disease can cause fever, red eyes, joint pain, headache, and edemas. This clinical presentation is similar to other infections transmitted by mosquito vectors, such as dengue, the West Nile virus, yellow fever, and others. The infection itself generally lasts less than seven days, and severe cases involving hospitalization are rare (Passi et al., 2017).

The diagnosis can be obtained by testing blood, urine or saliva for the presence of Zika virus essential molecules, the virus’ RNA. Transmission of the Zika virus disease happens via Aedes-type female mosquitoes’ bites (peak hours are during early afternoon and late afternoon/early evening hours), blood transfusion, unprotected sex, and from mother-to-child (prenatal).

There is no effective vaccine against the Zika disease and prevention involves controlling mosquitoes (eliminating standing water, screens on doors and windows, insecticides etc.), the use of condoms and supportive medicine therapy for mild symptoms, such as headache (Passi et al., 2017). There are currently nearly 30 companies working to develop potential diagnostic tests for Zika virus and 14 vaccine developers across five countries (U.S., France, Brazil, India, and Austria) who are actively involved in 23 projects for developing Zika virus vaccines, including the use of self-limiting genetically-modified mosquitoes to control its reproduction (Passi et al., 2017).
2.2 The 2015-2016 outbreak

In Brazil, over 1.5 million people were affected by the disease and an unusually high number of infants were born with microcephaly, a condition in which a baby is born with a small and underdeveloped head (ECDC, 2015). In 2015, there were 3,530 suspected cases of microcephaly in the country, from an annual average of 163 cases of microcephaly over the previous five years (Paixao et al., 2016).

Researchers from French Polynesia suggested the Zika virus was introduced to Brazil during the 2014 World Sprint Championship canoe race, in Rio de Janeiro, because of the international aspect of the event (Musso, 2015). Another theory suggested the Zika virus was introduced during the FIFA World Cup 2014 competition (Zanluca et al., 2015).

But data from airline traffic data combined with scientific analysis of the virus suggested the Zika virus was circulating undetected in Brazil since 2013, over a year before the first case detection of the virus in the country, and before the popular international events (Faria et al., 2016). This was explained by a spike in the number of air passengers to Brazil from Zika virus endemic areas, in 2013. Furthermore, the samples of the virus collected in the Americas were found to be identical to the Zika virus strain attributed to the French Polynesia outbreak, in 2013 (Imperato, 2016; see also Musso 2015; Zanluca et al., 2015).

Following the Brazilian Zika outbreak, the infection spread across over 22 countries and territories (Passi et al., 2017). The rapid spread of the Zika virus in Latin America and the Caribbean led health officials to issue polemic travel alerts so that pregnant women do not travel to affected areas. In El Salvador, women were advised not to have babies until 2018 but this measure has been criticized for being infeasible for many women (Schuck-Paim et al., 2012).
The evidence for causal connections between Zika virus infection and severe fetal brain defects was reviewed and finally confirmed by the U.S. Centers for Disease Control and Prevention in April 2016 (Rasmussen et al., 2016), which gave support to long-standing media speculations (Aguiar & Araujo, 2016).

In February 2016, the WHO issued a statement declaring the 2015-16 Zika virus outbreak as a Public Health Emergency of International Concern (PHEIC), which means it represented a serious and extraordinary event that required a coordinated international response. It was the fourth time the WHO declared this emergency. Past promulgations of PHEIC happened during the swine flu (2009), polio (2014), and Ebola (2014) outbreaks.

The PHEIC declaration for the Zika virus outbreak generated a report outlining a $56m strategic plan for the WHO and its partner organizations to respond to the ongoing outbreaks of Zika virus disease in the Americas (WHO, 2016). The measure also helped the WHO and its member states recommend travel limits and fast-track resources to fight the outbreak and potentially attract international donors, and improve worldwide preparedness coordination (Kindhauser, Allen, Frank, Santhana & Dye, 2016).

The United States government approved a bill seven months after the WHO’s PHEIC declaration, which included public funding for vaccine development against the Zika virus. The bill also funded research to develop rapid diagnostic tests and to explore the effects of the disease on unborn babies (Fox, 2016).

The Zika virus PHEIC declaration came some time after the WHO was widely criticized for its slow response to the 2014-15 Ebola outbreak in West Africa (Wenhan, 2016; see also Passi et al., 2017). Global health and law experts Lawrence Gostin and Eric Friedman (2014) argued the WHO had failed to fulfill its mission of global health leadership and coordination
during the Ebola outbreak, which was worsened by the agency’s limited budget, human resources, and lack of transparency. The criticism also acknowledged the saturation of international actors and the challenges faced in global health governance. Gostin and Friedman (2014) also urged the UN Security Council to use its authority to set priorities, and specify its member states’ responsibilities. A couple of years after the Ebola outbreak, the WHO responded to such criticisms by suggesting an organization-wide change to fully integrate its international partners, and to implement a platform for outbreaks and emergencies (Chan, 2015).

In November 2016, the WHO decided that Zika virus and its neurological consequences such as microcephaly and Guillain-Barré syndrome were no longer considered a Public Health Emergency of International Concern (PHEIC). The WHO published a statement saying the disease should now instead be perceived as an ongoing threat, and escalated into a sustained program of work “with dedicated resources to address the long-term nature of the disease” (WHO, November 2016). The end of the PHEIC still allows for individual member states of the WHO to declare local emergencies and work with the WHO to develop programs to address the long-term consequences of the disease. The WHO’s decision to end the PHEIC could also be explained by the seasonal aspect of the Zika virus disease, as the mosquito vector tends to be more active in higher temperatures (Caminade et al., 2016).

A recently published study used mathematical models to suggest the Zika epidemic in Latin America will likely end within three years, as individuals are unlikely to be infected with the Zika virus twice during their lifetime because of the immune system’s generation of antibodies to kill the virus (Ferguson et al., 2016; see also Imperial College London, 2016; Lessler et al., 2016). This herd immunity effect would cause the epidemic to reach a stage where there are few susceptible people to the infection for transmission to be sustained for at least
another ten years when there will be a new generation in the population who have not been exposed to the Zika virus.

The study also argued the current epidemic is not containable, but interventions could mitigate its burden (Ferguson et al., 2016). That happens because the Zika transmissibility is seasonal and modulated by climatic variations, which results in sporadic epidemics that happen when favorable conditions of transmission are met. Understanding this phenomenon is significant because it offers a “window to develop new interventions before further large-scale outbreaks occur” (p. 354).

Ultimately, the long-term solution to protect against the virus is the Zika vaccine (Lessler et al., 2016). Enhanced vector control is potentially beneficial, but it is important to consider the evidence suggesting that traditional insecticide-based control is not enough to stop similar seasonal epidemics, such as dengue, because these programs are not usually sustainable (Ferguson et al., 2016; see also Achee et al., 2015). In Latin America, campaigns led by the Pan-American Health Organization (PAHO) in the 1950s were effective in fighting mosquitoes (Llinás & Gardenal, 2012), but the interests in sustaining these campaigns were not continued (Vittor, 2016).

Moreover, a broader look into the Zika virus outbreaks will indeed reveal a framework in which the environment and climate change have significant implications for human health, particularly for the burden of the Zika virus disease. Passi et al. (2017) note that “wiping out an entire species could also have deleterious effects on the ecosystem and existing species could be replaced by more threatening species” (p. 16).

An article published Dr. Amy Vittor (2016), a professor of medicine at the University of Florida, discussed a series of factors or layers at play during the 2015-2016 outbreak, which
coincided and led to a “system failure.” The first layer found in the 2015-16 Zika virus outbreak was the region’s fertile environment for mosquitoes, including temperature and availability of breeding sites with increased urbanization and poverty (see also Ali et al., 2017; Caminade et al., 2016). The second layer was the mosquito vector and its ability to expand their geographic range in the past few decades (see also Mordecai, 2016). The third layer was the number of susceptible hosts since the population had not been exposed yet to the virus, particularly in a continental-sized country such as Brazil where the virus can continue circulating for an extended time. This layer relates well with the idea of the Zika virus herd immunity argued by Ferguson et al. (2016), and the papers published by Roth et al. (2014), Tatem et al. (2012), and Jones (2008) about the underlying determinants of the outbreak, such as the environmental factors. The fourth and last layer is the introduction of the virus, possibly associated with the increasing air travel (Vittor, 2016; Faria et al., 2016; see also Nunes et al., 2014).

Overall, I think additional research is needed to develop a vaccine and further understand the epidemiology of the Zika virus to mitigate the disease burden. Moreover, the collaborative framework of the International Health Regulations (IHR) should foster international surveillance and approaches for early detection and response to a public health crisis of this magnitude, involving many different factors (Rodier, Greenspan, Hughes & Heymann, 2007). As noted in relevant scholarship (Jones et al., 2008; Shuman, 2010; Epstein, 2005) it is also important to address the underlying social determinants of health and the effects of globalization, urbanization and how climate change adds a new layer of complexity to health systems, resilience and capacity building.
2.3 The timeline of events and the Brazilian media reaction

The Zika virus disease received significant and unprecedented attention from the media and the public when it hit Latin America in 2015. The first mentions of an unknown infection were reported by the Brazilian media early in that year, when there was a large number of people reporting the disease’s symptoms in hospitals of impoverished northeastern states. News reports at the time suggested the “disease without diagnosis” was a new type of dengue fever but with mild symptoms, even though clinical tests conducted by health authorities suggested the “new disease” was not dengue (Aguiar & Araujo, 2016).

In May 2015, Brazil’s health ministry confirmed the presence of the Zika virus in the country, following tests made at Brazil’s national reference laboratory. The Zika virus disease was initially considered “benign” and “not a cause of concern” by the Brazilian health ministry (“Ministério da Saúde confirma 8 casos de zika virus,” 2015), which was reproduced in the media. But news reports started to gradually change the tone around the “mysterious” disease when the Brazilian health ministry announced in July 2015 that the Zika virus infection could be linked to an increase in cases of Guillain-Barré syndrome, which can cause temporary paralysis in patients (Aguiar & Araujo, 2016).

In Brazil, health is recognized in the constitution as a citizen’s right and duty of the state. Therefore, the country established in the 1990s the creation of the SUS (Sistema Único de Saúde; or Unified Health System), based on principles such as universality and social participation; although it is becoming increasingly privatized (Paim, Travassos, Almeida, Bahia & Macinko, 2011). Hence, a large proportion of the population relies on the public health system and its outreach programs. Neurological implications stemming from Zika virus infections would have immense effects on the country’s public health services (see Cao-Lormeau et al., 2016).
In October 2015, the Brazilian media started to report on a significant increase in cases of microcephaly in the country, a condition in which a baby is born with a small and underdeveloped head. The speculations surrounding this neurological disorder and the Zika virus infection became even more popular in the media, as the São Paulo State Obstetrics and Gynecology Association published a statement expressing concerns about the situation, and proposing additional investigation (Aguiar & Araujo, 2016). A paper published during a regional infectious disease conference in Brazil also circulated in the media, as it recommended increased surveillance and control of the Zika virus disease (Luz, Santos & Vieira, 2015).

In the same month, October 2015, the Brazilian government declared the Zika virus disease as an Emergência em Saúde Pública de Importância Nacional (Public Health Emergency of National Concern), followed by a press conference where the health ministry announced they were further investigating and surveilling the spike in cases of microcephaly.

In November 2015, Brazil’s health ministry issued an epidemiological alert confirming the connection between Zika virus infection and microcephaly based on two studies. The first, published by the Oswaldo Cruz Foundation (FIOCRUZ), found the virus in the amniotic fluid of two pregnant women whose fetuses were normal before the Zika virus infection (Calvet et al., 2016). The second, published by the Evandro Chagas Institute, found the virus in a newborn baby diagnosed with microcephaly, who died soon after birth (Oliveira & Vasconcelos, 2016).

The confirmation from the Brazilian authorities led the PAHO to publish an epidemiological alert about the Zika virus warning about the possible relation between the infection and other neurological, and autoimmune syndromes (PAHO, December 2015). This was the first statement coming from an international health organization about the Zika virus
outbreak in Brazil, and it represented a critical turn in public opinion, as the conversation now assumed a public health emergency tone (Aguiar & Araujo, 2016).

The Brazilian federal government waged a “war on the mosquito” in the months following the health ministry’s public confirmation of the link between Zika virus infection and microcephaly. All of these highly-publicized media events added an extra layer of fear to the outbreak and it didn’t take long for conspiracy theories to arise, such as speculations involving Tdap vaccination (diphtheria, tetanus, and pertussis) given to pregnant women or even rumors about expired vaccines (Worth, 2016). These theories were publicly denied by health authorities (Ministerio da Saude, 2015 December).

The outbreak of the Zika virus in Brazil and Latin America also sparked a prompt worldwide interest in Zika. A look into Google Trends data (Figure 1) suggests that query volumes related to Zika spiked in January 2016 when there was the confirmation of the first American infection, which was worsened by concerns with the proximity of the 2016 Summer Olympic Games (Leetaru, 2016).

![Figure 1 – Google Trends data showing interest over time of the words Zika (blue) and Zica (red)](image)
The avalanche of information and the volume of complex epidemiological data and reports published by governments and health authorities involved in the Zika epidemic have shown to oftentimes impair editorial judgments, leading to a coverage marked by sensationalism and analysis based on anecdotal evidence with wide repercussions on social media, where much of what we know is determined by what algorithms infer as relevant. (Gyawali, Bradbury & Taylor-Robinson, 2016; see also Aguiar & Araujo, 2016; Leetaru, 2016).

Leading health communication scholars in Brazil (Aguiar & Araujo, 2016) highlighted the challenges faced by the media to report and apply the journalistic rigor among a scenario of scientific uncertainty, which could have had consequences such as feeding popular misconceptions and pseudo-scientific claims. The scholars also analyzed the front pages of nine Brazilian newspapers over the course of the outbreak and noted how people who are disenfranchised were often portrayed by the media as “the poor, ignorant,” carrying the blame for what happened because of their ignorance, which could reflect and perpetuate health inequities. The study suggested the media used the words fear and risk to perpetuate “narratives of suffering,” exposing the crisis faced by impoverished mothers and their babies with microcephaly in the country’s urban centers.

2.4 Social networking

It is important to understand the ongoing increase in social media use globally (boyd & Ellison, 2007) and the opportunities it creates for social participation, communication and shared meanings. “A growing body of research points to how social media, and specifically Twitter, is emerging as a hybrid space where citizens are involved in the flow, framing, and interpretation of news” (Callison & Hermida, 2015). As people share their stories and voice their fears, hopes,
and concerns through social media, it is important to show these Internet-based tools and its uses’ implications as a relevant part of the public discussion.

The definition of social media I subscribe defines it as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content” (Kaplan & Haenlein, 2010).

Communication scholars danah boyd and Nicole Ellison (2007) argue these web-based services allow individuals to construct a profile within the system where the user can articulate, view and traverse their list of connections. In this system, there are two main frameworks at play: the *media-related component* and the *social dimension* (Kaplan & Haenlein, 2010). The first relates to the idea that a medium is effective when it can eliminate ambiguities and uncertainty to allow a continuous interchange of information, resulting in a high degree of social presence (see also Daft & Lengel, 1996; Berger & Luckmann, 1967). The second relates to Goffman’s (1959) idea of people controlling impressions in any social interaction to either gain influence or to create a consistent image of one’s personal identity, which is here applied to social media.

Andreas Kaplan and Michael Haenlein (2010) also argue that social media and social networking are not the same, even though they might be used interchangeably; social media is a channel to deliver a message, while social networking is the direct communication and sharing of information. This ultimately results in various types of social media platforms, with different ways of networking. Other scholars (Kietzmann, Hermkens, McCarthy & Silvestre, 2011) proposed a “social media ecology.” The understanding of this ecology and its building blocks would allow organizations and communities to recognize the conversations that are already happening online and develop strategies based on that.
The change in the relationship between producers and receivers of information in social media is also present in other research areas. For example, the passive consumption of media by the masses (Livingstone, 2005) is being constantly challenged by the ability people possess to make sense of the media with individual interpretative lenses, and ultimately communicate their thoughts through internet technologies (see Mangold & Faulds, 2009; Marwick and boyd, 2011).

As much as 75% of online news consumers in the United States read news through social networking sites or emails (Purcell et al., 2010). This is also reflected in the way news organizations work now by prioritizing digital publications first, adding a whole range of tools that facilitate the sharing, recommendations by readers and encouraging journalists to promote their content while linking back to the news organization's website in order to drive traffic, extend audience reach, and foster brand loyalty (Singer et al., 2011, as cited in Hermida, 2015). Candis Callison and Alfred Hermida (2015, p. 816) argue that “social media spaces allow for new relations that disrupt authorial structures and established flows of information.” Therefore, this “networked audience” changes the constructs and dynamics of the media system in a new “hybrid media system” which involves processes of transition, interdependence, and blending between new and old forms of media (Chadwick, 2013).

Particularly on Twitter, the networked platform works by weaving users’ contents into constantly updated “social awareness streams” where people can reframe and reinterpret messages through social interaction (Callison & Hermida, 2015). Additionally, the use of hashtags (#) to identify a theme or topic within your tweet (message) allows users to participate, follow and filter through specific online conversations among Twitter’s massive content.
These processes of using hashtags and filtering through content mainly indicate “a desire to participate in a specific communicative exchange and make such content more visible on the network, beyond a user’s immediate circle of followers.” (Callison & Hermida, 2015, p. 698; see also Bruns & Moe, 2013). These hashtags can be ephemeral and gradually lose importance or even disappear among the new “trending topics.” But they can also be long-lasting and reflect authentic online communities with the generation of meaning and dialogue (Callison & Hermida, 2015). This diversity of conversations also points to the nature of the networking platform, which seeks to give real-time status updates to create a sense of “ambient awareness” of issues through specific formats (short messages) and protocol implications (Kietzmann et al., 2011).

Alternatively, research also demonstrated the structure of influence and passivity on social media, particularly how the correlation between popularity and influence is weaker than expected, based on how individuals actively engage and act on the messages they receive (Romero, Galuba, Asur & Huberman, 2011). This relates to the idea that everyone seems to be talking, but not everyone is being heard in social media spaces (Hermida, 2015). Hence, it is important to acknowledge the institutional power relations at play in society and how that is reflected or contested on social media.

### 2.4.1 Intersections with public health

A growing body of research also points at how social media changes the way people communicate about public health issues. The personal and professional use of social media by physicians is also increasing (Quisenberry et al., 2016), but trust through online interactions appears to be strongly based on a series of other external factors. These factors include “previous personal interactions, professional standing, authenticity and relevancy of voice, consistency of
communication, peer recommendation, and non-anonymous and moderated sites” (Panahi, Watson & Partridge, 2016, p. 70). Physicians have also used Twitter to reach a broad audience and “potentially even affect health policy decisions” (Fogelson, Rubin & Ault, 2013).

Health and information science scholars (Panahi, Watson & Partridge, 2016; 2014) identified a set of reasons for physicians to join social media, including “staying connected with colleagues, reaching out and networking with the wider community, sharing knowledge, engaging in continuing medical education, benchmarking, and branding.” They also identified challenges in adopting social media such as “maintaining confidentiality, lack of active participation, finding time, lack of trust, workplace acceptance and support, and information anarchy.” This study sought to understand the scope and impact of social media and healthcare workers while suggesting measures to maximize the benefits of using social media for specific needs of a clinical community.

Patients and healthcare workers also use social networking to communicate about health issues through social marketing campaigns, by producing and distributing information through collaborative writing and sharing, enhancing viral marketing to larger audiences (Thackeray, Neiger, Hanson & McKenzie, 2008).

A national survey conducted by the Pew Research Center (Fox, 2014) in the U.S. found that 72% of adult internet users went online for information on health issues; 26% went online to look into other people’s health issues; 16% went online to find similar people, sharing the same health issues. The latter is significant for people with rare or chronic diseases who not only information online but also similar people experiencing the same health issues through social networking, which “adds to the pile of evidence showing the psychosocial benefits of connecting online.” (Fox, 2014).
Drawing on relevant scholarship (Chou et al, 2009; Kontos, Emmons, Puleo & Viswanath, 2010), Anne Moorhead et al. (2013) notes that the use of social media makes health information more accessible, particularly for audiences with special needs in low-literacy settings; hence, facilitating the dialogue between healthcare workers and patients. Alternatively, the scholars argue the low cost of social media and its ability to provide real-time communication can track and monitor public response to health issues and outbreaks, identify misinformation and target areas for interventions.

The study also notes the traditional inequalities in terms of overall Internet access and how it remains an issue in need of “ongoing scrutiny regarding equality of access and effectiveness for different users” (p. 9). Therefore, it is important to acknowledge the benefits of social media for health communication, but there needs to be further research exploring the long-term effects of using social media in health care.

Further examples of research involving social media analysis in health communication reveal how social networking is a powerful tool to reach massive audiences and advance certain messages of control and prevention. King-Wa Fu et al. (2016) describe the findings of a computational content analysis on how people react to Zika virus outbreaks on Twitter, based on Zika virus-related English tweets in late January 2016. The study found a large number of tweets in Spanish and Portuguese in 2015, reflecting the awareness of the outbreak in those language-speaking areas during that first year of the 2015-16 Zika virus outbreak. However, the researchers noticed a spike of Zika virus-related English tweets in early 2016, when worldwide media attention was given to the WHO’s PHEIC announcement and consequent travel advisories to Zika virus affected areas, and an increase in cases in the continental United States.
The topics of these conversations revolved around the public’s concerns about the risks of microcephaly and Guillain-Barré syndrome, as well as the disease’s transmission routes and the risks associated with the infection. The study also found that user-generated content was the preferred channel of Zika-related information, which means people were not sharing links of public health agencies’ websites directly.

Another significant study in this research area found that daily news story volume predicted trends in the United States search behavior related to Zika virus, particularly health authority public announcements (Southwell, Dolina, Jimenez-Magdaleno, Squiers & Kelly, 2016). The scholars assessed the findings based on the combination of Google Trends (to assess the number of searches), Crimson-Hexagon's monitoring tool (to capture the total number of daily tweets), and The Associated Press news wire (to measure the daily volume of Zika-related news stories). The study suggested the “news coverage of public health authority announcements opens brief windows of information sharing, engagement, and searching that offer opportunities to address perceptions and provide preparation and vector control recommendations through education” (p. 3).

A third use of Internet-based data sources to develop health research found that Internet tools can be used to forecast and assess dynamics of an outbreak, and potentially mitigate disease transmission (McGough, Brownstein, Hawkins & Santillana, 2017). The study used different sources of data to assess the findings, such as Google searches, Twitter, and HealthMap (a surveillance system that provides case counts). The intent was to track and forecast weekly suspects of Zika virus infection cases during the 2015-16 Latin American outbreak. This work is particularly relevant to public health because of the delayed official Zika case reports from health authorities, and the need for “early interventions and real-time surveillance [...] necessary to curb
mosquito transmission” (p. 2). Moreover, this study’s methodology can be used in any country and help identify changes in Zika transmission for public health policymaking (see also Majumder, Santillana, Mekaru, McGinnis, Khan & Brownstein, 2016).

Similar studies using Twitter and/or internet search query data for surveillance of influenza (Polgreen, Chen, Pennock & Nelson, 2008; Ginsberg et al., 2009; Yuan et al., 2013; Nagar et al., 2014; Signorini, Segre & Polgreen, 2011; Santillana et al., 2015), dengue (Althouse, Ng & Cummings, 2011) and Ebola outbreaks (Majumder, Kluberg, Santillana, Mekaru & Brownstein, 2015) are also prominent in the literature.

2.4.1.1 The limitations of social media

Other studies have demonstrated how social media can hinder public health efforts and lead to ultimate distrust in health authorities (Dredze, Broniatowski & Hilyard, 2016). In the case of the Zika virus epidemic, the uncertainty around the context and epidemiology of the disease created a “fertile environment for conspiracy theories and pseudo-scientific claims,” which have made their way to the mainstream media and has the potential to become “entrenched, increasing the likelihood people will refuse a Zika vaccine” (p. 3441).

Leading scholars (Dredze, Broniatowski & Hilyard, 2016) found that specific mentions within the tweets that were associated with pseudo-scientific claims also coincided with overall interest increases. The study argues these pseudo-scientific claims are built on existing narratives and beliefs, which makes them more persuasive. The research sample consisted of 138,513 Zika-related Twitter messages from January 1 to April 29, 2016, which were analyzed by the scholars. The results showed an increase in the number of Zika vaccine tweets coinciding with increased media attention, which echoes the findings of Fu et al. (2016).
These findings reflect a significant limitation of social media use for health communication, which is the lack of quality found in the information shared on Internet-based platforms, and the overload caused by the large volume of information that can lead to inaccuracy (Moorhead et al., 2013). Relevant scholars (Orizio et al., 2010; Adams, 2010; as cited in Moorhead et al., 2013) also argue the lack of reliability of the health information is worsened by the anonymity behind some producers of information, and the lack of regulations for health professionals to communicate with patients online, including confidentiality (regarding data security). Alternatively, the link between external media events and increases in the volume of online conversations around health issues (Fu et al., 2016; Dredze, Broniatowski & Hilyard, 2016) offers a window to understand the role of the media and other global health actors in shaping the public opinion.

2.5 Chapter summary

This chapter outlined some of the relevant scholarship on the Zika virus disease, its history, and epidemiology. Additionally, I discussed the specificities of the outbreak in Brazil and its implications in terms of media and health communication strategies. The intent of this review was to explore the interconnectedness of social networking, media, social power, and global health, as I sought to understand how can social media data provide insights to understanding the online trending narratives during the Zika virus outbreak 2015-16.

For the purpose of this project, this review contributed to a general understanding of how can social networking offer a window for journalists and global health actors to have more informed perspectives, reaching mass audiences and enhancing health communication strategies. Therefore, I explored the literature demonstrating the potential of social media for reaching mass
audiences as a tool for health communication, intervention, education, and up-to-date surveillance system to alleviate disease burden and transmission (McGough et al., 2017).

Alternatively, social media can be a tool for patient empowerment, and health behavior change (Syed-Abdul et al., 2016; Collinson, Khan & Heffernan, 2015).

This review also considered the limitations of social media use in health communication and the need for further research examining the short and long-term effects of Internet technologies in public health. Thus, I reviewed the existing literature demonstrating how social media can hinder public health efforts and lead to ultimate distrust in health authorities (Dredze, Broniatowski & Hilyard, 2016).

The literature review also discussed the need for public health officials to tackle the concerns to debunk nonscientific claims on social media (Moorhead et al., 2013), which are worsened by the online anonymity granted by social networking platforms and the lack of regulations regarding online health communication (Orizio et al., 2010; Adams, 2010). This is particularly relevant in a context where the use of social media is exponentially growing (boyd & Ellison, 2007; Fox, 2014; Quisenberry et al., 2016). In Brazil, there were an estimated 93 million people using social media in 2016 (Forbes, 2016). Therefore, this information needs to be monitored for quality and reliability (Moorhead et al., 2013).

Ultimately, I explored the challenging scenario of reporting and applying journalistic rigor among scientific speculations and uncertainty, which could lead to mainstream media perpetuating and feeding popular misconceptions, or fear (Aguiar & Araújo, 2016). Additionally, I explored research demonstrating the link between external media events and increases in the volume of online conversations around health issues (Fu et al., 2016; Dredze, Broniatowski &
Hilyard, 2016), which offers a window to understand the role of the media and other global health actors in shaping the public opinion.

Overall, the intent was to identify the scholarly content that was necessary to ground this project’s main research questions, which focus on the trending online Zika-related conversations to help understand how social media data can provide insights into the role of the media during a global health crisis.
Chapter 3: Methodology

The purpose of this study is to explore trending narratives on social media in Brazil during the 2015-16 Zika outbreak. Specifically, I sought to understand the role of social media during a global health crisis and how it relates to health communication strategies. My assumption is that a better understanding of Internet-based tools and social networking would allow journalists and global health actors to have a more informed perspective in terms of reaching and educating mass audiences.

In seeking to understand this phenomenon, the study addressed three research questions (1) How can social media data provide insights into the role of the media during global health crisis? (2) How does social power play on social media? (3) What were the trending conversations on Twitter during the Zika virus epidemic in Brazil?

This chapter describes the study’s research methodology and includes discussions around the following areas: the rationale for a qualitative research design and methodology, the rationale for the method, the overview of the method phases, data collection method, the research sample, opportunities and limitations, and a chapter summary.

3.1 Rationale for qualitative research design, and methodology

The rationale for this qualitative study was based on a social constructivist epistemology, which is underpinned by relativist ontological leanings (see Berger & Luckmann, 1967; Kuhn, 1962). My approach focused on context to understand the complexities of social interactions of everyday life, at the same time as remaining sensitive about the researcher’s subjectivity and engaging in systematic reflection to conduct the study (Marshall & Rossman, 2006).

Therefore, there was no search for a singular objective meaning, truth, or central core
structure, “explained by referring to some supposed inner truth or essence” (Taylor & Ussher, 2001). Instead, I looked at ways in which people and their interactions constructed meaning for themselves. This echoes the core idea of intentionality in social constructivism, conveying that people make meaning through the sense of being and participating in an interactive space among other people and objects, in the available social stock of knowledge (Berger & Luckmann, 1967). Hence there was an emphasis on discovery, description and the “multiplicity of interrelated understandings, each with their own inherent validity” (Taylor & Ussher, 2001).

The social constructivist framework also considers the way humans index subjective meanings through signification, or the production of signs. Thus, language is seen as the most important sign system in the human society as its complexity and variety allow access to the other’s subjectivity. Language is a “repository of vast accumulations of meaning and experience, which it can then preserve in time and transmit to following generations” (Berger & Luckmann, 1967, p. 52). This is particularly relevant for this project as it also sought to make sense of the language through categories used to reflect the online conversations around the 2015-16 Zika virus epidemic, specifically in Portuguese, and the exchange of meanings across different people.

3.2 Rationale for method

Within the theoretical framework of a qualitative research, the method most suited for this project was a thematic analysis, as I sought to identify the topics and overarching themes in the online conversation to ground the project’s main research questions. The research questions focused on the trending online Zika-related conversations to help understand how social media data can provide insights into the role of the media during global health crisis.
Therefore, thematic analysis as a method seeks to identify, analyze, and report patterns, themes, and meanings within data (Braun & Clarke, 2006). This method seeks to identify themes and patterns across an entire data set, as opposed to individual texts or interviews which are the case for other forms of analysis. The framework for thematic analysis developed by Virginia Braun and Victoria Clarke (2006) “is not wed to any pre-existing theoretical framework, and so it can be used within different theoretical frameworks (although not all), and can be used to do different things within them.” However, I adopted a constructionist approach for this thesis, which involved considering experiences, structural conditions, and meanings shared in an intersubjective reality, and its effects on the online narratives surrounding the Zika virus outbreak in Brazil.

The definition of theme within a thematic analysis is one that “captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set” (Braun & Clarke, 2006, p. 10). In terms of size, the scholars put forward that it should be taken into account theme’s prevalence within each data item, and across the entire dataset. It is important to notice that more prevalence here doesn’t necessarily make a theme more crucial, as highlighted by Braun and Clarke (2006). Researcher judgment is also relevant to determine what a theme is, as it is a qualitative analysis. For the purposes of this project, the categories and consequential prevalence were not defined in a linear process. Instead, they were a result of going back and forth between the data to construct them over some time.

The approach used for the theoretical thematic analysis was semantic, where the themes are identified within the “explicit or surface meanings of the data” (p. 13) and involves “a progression from description, where the data have simply been organized to show patterns in
semantic content, and summarized, to interpretation, where there is an attempt to theorize the significance of the patterns and their broader meanings and implications” (p. 13).

3.2.1 Overview of method phases

Braun and Clarke (2006) provide an outline of six phases consistent with a thematic analysis. The researchers also recognize the flexibility of these “basic precepts” to fit different research questions and data sets, consistent with a qualitative analysis framework (Patton, 1990, as cited in Braun & Clarke, 2006).

Phase one consists of familiarizing with the data. The scholars argue that it is important to “immerse yourself” in the data sufficiently enough to be familiar with it, even if the researcher collected the data through interactive means and therefore having a prior knowledge or analytic thoughts about the data. “Identification of possible patterns will be shaped as you read through” (p. 16). In this project, however, the massive amount of big data collected makes familiarization with all its content unfeasible. This phase was conducted by reading through the most shared and most popular conversations, measured by the data collection tool used in this study.

Phase two consists of generating initial codes from the data, which identify the semantic features of the data. The authors said these codes will develop into themes in the next phase, which are often broader, and are also “where the interpretative analysis of the data occurs, and in relation to which arguments about the phenomenon being examined are made” (Boyatzis, 1998, as cited in Braun & Clarke, 2006, p. 18). Since the approach of this project was theoretical and driven by my interest in the area, the themes will be approached with specific research questions to code around. Coding will be done with a tool from Crimson-Hexagon, which will be explored later in this chapter.
Phase three consists of analyzing the codes to form an “overarching theme,” which can be a combination of different codes. This phase “re-focuses the analysis at the broader level of themes, rather than codes, involves sorting the different codes into potential themes, and collating all the relevant coded data extracts within the identified themes” (p. 19).

Phase four consists of refining those themes and reviewing evidence to support them, including reviewing at the level of the coded data extracts. “Data within themes should cohere together meaningfully, while there should be clear and identifiable distinctions between themes” (p. 20). This phase involved two levels. The first one consists of an analysis at the coded data extracts level and considers if they appear to be in a coherent pattern; if not, they should be reorganized or discarded. The second level consists of using the resulting “thematic map” and analyze its coherence pattern with the entire dataset. It is also important to check it the map reflect meaning that’s evidence in the data or check to see if there are missing themes. Braun and Clarke (2006) note “the need for recoding from the data set is to be expected as coding is an ongoing organic process” (p. 21).

Phase five consists of defining and naming themes, which is “the essence of what each theme is about”, and “what determining aspects of the data each theme captures” (p. 22)

Phase six consists of producing the actual report, which involves the final analysis with sufficient evidence within the data. “The task of the write-up of a thematic analysis, whether it is for publication or for a research assignment or dissertation, is to tell the complicated story of your data in a way which convinces the reader of the merit and validity of your analysis” (Braun & Clarke, 2006, p. 22). The researchers also highlight the importance of using vivid examples in the report to help construct the argument.
3.3 Data collection tool

The tool used in this thesis for big data collection was developed by Crimson-Hexagon, a social media analytics company founded in 2007 with technology developed at Harvard University’s Institute for Quantitative Social Science (Hitlin, 2013). This platform uses a supervised text analysis model, which relies on an algorithm and method developed by Hopkins and King (2010):

These new methods take as data a potentially large set of text documents, of which a small subset is hand coded into an investigator-chosen set of mutually exclusive and exhaustive categories. As output, the methods give approximately unbiased and statistically consistent estimates of the proportion of all documents in each category. (p. 229)

Crimson Hexagon’s software is called ForSight and it is a “Twitter-certified” product, which means it has access to the “Twitter firehose,” or every public tweet on Twitter, in any language and location. Retweets are included in its analysis by default, but the researcher can choose to leave it out (Jamal, Keohane, Romney & Tingley, 2015). I chose Twitter because it an accessible social media platform, with fewer privacy settings and more public posts when compared to other social networking sites (Dredze, Broniatowski, Smith & Hilyard, 2016).

The software from Crimson Hexagon analyzes textual content from Twitter and classifies it with statistical patterns in words, as mentioned before. In the software’s interface, each study or query is named as a monitor. In order to create a new monitor, the researcher has to decide the time frame and the content that is going to be examined, including the language of the tweets and/or their location. The researcher then uses the monitors to familiarize with the data and hand-code a set of textual contents (tweets) into categories, which are predetermined by the researcher.
This process of hand-coding tweets into categories is called “training” and its purpose is to show the proportion in the universe of tweets that fit in the researcher’s predetermined categories.

The training is made using a random selection of tweets collected by Crimson Hexagon and presented to the researcher, who then sorts the posts individually into categories. The sorting continues until the researcher feels the monitor is trained or until the fulfillment of the software’s minimum number of tweets required for this process. After the training, the researcher runs the algorithm and Crimson Hexagon generates the analysis results, which contains the option of bulk exporting a limited number of tweets from the monitor. As Jamal et al. (2015) argues, the option of exporting the tweets is useful for getting a sense of how well the algorithm performed. The scholars also point other advantages of using Crimson Hexagon, such as the high level of filtering out spam and bots that could contaminate the analysis because of Crimson Hexagon’s commercial customers’ demands.

It is also important to notice that Crimson Hexagon uses sampling to generate some snapshots of the conversation online, rather than measuring all the conversation (except for volume, which is always based on total). This happens because of the large numbers of data points included in such social media analysis. According to the company’s website, each metric available in the software’s monitor (except for volume) has the potential to show either the total conversation or a random sample of the total conversation. The sample is randomly selected based on the proportion to the prevalence of that data source in the user’s total volume.

Relevant scholarship has used Crimson Hexagon for the purposes of social sciences research and analyzing big data, particularly Twitter (see Jamal et al., 2015; Ceron, Curini, Iacus & Porro, 2014; Runge et al., 2013; Etling, 2014; Lipizzi, Iandoli & Marquez, 2015; Raynauld & Greenberg, 2014; Hodges & Stocking, 2016). These studies lend support to the claim that
Crimson Hexagon is a powerful and effective tool for making a sense of the online conversations on social media and sort through the overwhelming amount of data involved in analyzing social media content.

### 3.4 The research sample

I limited the time frame for the data collection of tweets based on the duration of the 2015-16 Zika virus outbreak in Brazil. Therefore, the date range entered into Crimson Hexagon was between January 2015 (around the time when there were the first reported cases of Zika virus disease in Brazil) until November 2016, when the WHO announced the Zika virus and associated consequences no longer represented a Public Health Emergency of International Concern (PHEIC).

The geographical location was set to Brazil, and the language was set to English and Portuguese. The research sample included the following terms, or keywords, used in the query, which represents the name of the virus in both languages, and with the use of hashtags (a) Zika; (b) #Zika; (c) Zica; and (d) #Zica.

Within this delimiting time frame, there were 2,581,249 tweets in the universe of the sample identified by Crimson Hexagon, therefore relating to the monitor’s’ parameters, including retweets. In this universe of tweets, 46% (or 1 million posts) had identifiable gender; 56% were male users, 44% were female users. Almost 53% of these messages could be geographically located on a city level. The cities with the most volume of tweets were in the southeast area of Brazil.
3.5 Overview of the research design

The following list is based on Braun and Clarke’s (2006) thematic analysis phases combined with Crimson Hexagon’s workflow. It was meant to summarize the steps I used to carry out this project:

1. Delimit a time frame, location, and keywords to define the content sources that will be analyzed.
2. Familiarization with the data by reading through random samples, topic wheel, word clouds and authors according to their Klout Scores, which is provided by Crimson Hexagon’s monitor and consists of a third-party application that ranks users in a numerical value between 1 and 100 according to influence, measuring the size, content created and interaction. I sought to read the most and the least influential content generators to have a better understanding of the context.
3. Generating initial codes and searching for themes by taking notes with thick description.
4. Reviewing themes and checking their relationship with the entire dataset.
5. Defining and naming themes on Crimson Hexagon’s monitor.
6. Training Crimson Hexagon’s algorithm into my predefined categories to measure the proportion of conversations of each category.
7. Familiarization with the analysis generated by Crimson Hexagon and comparing that with the external timeline of events relating to the 2015-16 Zika virus outbreak.
8. Producing the reports
3.6 Opportunities and limitations

This project contains certain limitations, which stem from the study’s research methods and tools. The data collection tool developed by Crimson Hexagon can only access posts that were made public, so I did not have access to tweets posted in private accounts. Additionally, only half of the universe of users indicated their city, which makes a regional breakdown and analysis harder. But this could be considered a minor issue, since Twitter is an accessible social media platform with fewer privacy settings when compared to other social network platforms (Dredze, Broniatowski, Smith & Hilyard, 2016). Braun and Clarke (2006) note the method used in this thesis (thematic analysis) does not require a detailed theoretical knowledge to identify where patterns in the study are socially produced; although this provides a framework for expanding this study into a discursive analysis, in the future. Finally, the researcher’s subjectivity in a qualitative project should also be considered along with concerns with bias and presumptions.

Given these limitations, I sought to minimize their impact on the analysis and address issues of validity by applying certain procedures for rigor in qualitative research, based on the work of John Creswell and Dana Miller (2000). In order to evaluate the trustworthiness of this study, I applied the following strategies: (a) peer debriefing; (b) developing an audit trail; (c) reflexivity; (d) thick description; (e) triangulation; and (f) comparison. The procedures for ascertaining the criteria were: (a) discussing the methods, research design and findings with supervisors and colleagues who have done research using similar data collection tools; (b) documenting the evolution of the research by keeping a research journal and personal notes about data collection, interpretation, and overall rationale of the research process; (c) engaging in an ongoing reflexive process about the research questions, intentions, and researcher’s
subjectivity and potential bias; “turning back on one's experience upon oneself wherein the self to which this bending back refers is predicated and must also be understood as socially constructed” (Steier, 1991, p. 3); (d) using thick description in the findings with rich details, which can be relevant in terms of applying the findings to broader contexts (Schram, 2003; see also Denzin & Lincoln, 2017); (e) gathering information from different sources (Mathison, 1988); and (f) comparing my findings with existing literature.

3.7 Chapter summary

This chapter discussed the methodological approach of this project and demonstrated how a qualitative framework was used to understand how can social media data provide insights to understanding the online trending narratives during the Zika virus outbreak 2015-16. Additionally, the intent was that this project would contribute for understanding how social networking would allow journalists and global health actors a window to have more informed perspectives, reaching mass audiences and improving health communication strategies.

The method employed was thematic analysis, based on a social constructivist epistemology. The method was combined with a tool developed by Crimson Hexagon to collect big data and the sort online conversation around the Zika virus into predetermined themes. The sample collected resulted in 2,581,249 tweets published within the delimiting time frame of January 2015 and November 2015, when significant events related to the Zika virus outbreak occurred in Brazil and internationally, including the WHO’s PHEIC declaration.

Finally, I applied certain procedures to evaluate trustworthiness and rigor in qualitative research, and to minimize the limitations of this study, including peer debriefing, developing an audit trail, reflexivity, thick description, triangulation, and comparison. The procedures for
ascertaining these criteria were also laid out by me. I hoped this project would shed a light into understanding the initial research questions, and developing a framework for further research into deploying Internet-based applications to help put forward health communication strategies.
Chapter 4: Findings

The purpose of this study is to explore the narratives on social media in Brazil during the 2015-16 Zika outbreak. I sought to understand the role of social media during a global health crisis and how it relates to broader health communication strategies.

This chapter presents the key findings obtained from analyzing 2,581,249 Twitter posts published from January 2015 until November 2016; thus, overlapping with critical events in the timeline of the Zika virus outbreak in Brazil. The Twitter posts were obtained using a software developed by Crimson Hexagon, which has tools to store and analyze textual content from social media platforms.

This chapter is divided into the following sections: (a) general findings (b) thematic findings. The general findings section describes the interpretation of quantitative data generated by the software Crimson Hexagon, including charts, word cloud, clusters, and list of most influential authors. The thematic findings section describes the rationale for the categories and the respective major findings.

4.1 General findings

From the data collected, three major initial findings emerged from this study and the visualizations:

1. The overwhelming number of Zika-related posts published within the delimiting time frame, which indicates a high level of engagement among Twitter users in Brazil.
2. The correlation between the volume of Zika-related tweets and the timeline of external events of the outbreak, particularly in Brazil.
3. The most influential authors of Zika-related tweets were Brazilian news organizations. However, Brazilian politicians and popular entertainers also made to the list.

The next paragraphs will follow with a discussion explaining the findings in detail. I will pull out illustrative examples from the data to document the users’ experiences and provide the reader with the possibility of better understand the findings, and ultimately give basis to the discussion. As a native speaker of Portuguese, I translated keywords identified in this section to English.

**Finding 1: The overwhelming number of posts published within the delimiting time frame, which indicates a high level of engagement among Twitter users in Brazil.**

Within the delimiting time frame (January 2015 until November 2016), there were 2,581,249 Twitter posts. The jump observed in Figure 4 from less than 10,000 Zika-related mentions per day in 2015 to over 50,000 a day in some critical weeks of 2016 could reflect a high level of awareness and concern of the Zika virus among Twitter users in Brazil. The geographical distribution of the conversation was identifiable at a city level on 1,362,786 posts (52%). Within this percentage, the top five cities were in the southeast region, which leads among all Brazilian regions in terms of population and GDP per capita: Sao Paulo (21.79%), Sete Lagoas (13.97%), Belo Horizonte (6.97%), Rio de Janeiro (6.93%), and Brasilia (3.65%). The result was similar when I focused on the first four, and last four months of the original time frame.
Most topics discussed in the conversation were identified by the process of sampling the posts and visualizing them in a word cloud, as shown in Figure 2. The intention was to get an initial overview of the discussions and topics used with high frequency in the sampled posts. I found the conversation involved talking about the risks associated with the infection, such as microcephaly, represented in the word cloud by the words *Abortion, Death, Epidemic, Uterus, Pregnant, Detection, Guess, Cause, Fear, and Counselling*. The users were also concerned with the political impacts of the Zika virus outbreak, represented in the word cloud by the words *Government, Olympics, Rio, Research, Scientist, Control, Mosquito, WHO, U.S., Minister, and Dilma* (reference to the Brazilian president at the time, Dilma Rousseff).

![Figure 2 – Word cloud representing the words used in the research sample](image)

The words were then separated into clusters, to help visualize the interconnectedness of the terms, as shown in Figure 3. This indicated a strong relation between the words *Zika, WHO, Microcephaly, Chikungunya, and Dengue*; as well as the words *Counselling, Detection, Speak*...
Out, Gynecologist, and Testing, on a separate cluster. The visualization also indicated a third cluster, suggesting the link between the terms Zika, Death, and God.

![Figure 3 – Clusters representing the words used in the research sample](image)

To obtain this visualization, I set a maximum limit of 20 words within the selected sample. I found that a larger number of words in clusters represented similar results but with less effectiveness because of the number of off-topic individual words found in the data.

**Finding 2:** The correlation between the volume of Zika-related tweets and the timeline of external events of the outbreak, particularly in Brazil.

An overview of the volume of total posts over the chosen time frame of this project revealed a close relationship between increased media attention to the Zika virus disease and increases in the volume of Zika-related Twitter posts, as shown in Figure 4.
The largest number of Zika-related tweets were between January and April 2016, when major public announcements were made involving the Zika virus outbreak. This relationship is also reflected in the findings of Dredze et al. (2016). Alternatively, Fu et al. (2016) argued the increase in the number of suspected cases of Zika virus infection reported to WHO/PAHO by several countries and territories with active cases coincided with increases in the number of Zika-related tweets.

Five waves in the volume of posts are represented the visualization. The first wave started in November 2015, following a declaration by the Brazilian health ministry confirming the link between Zika virus infection and microcephaly, and news reports showing compelling photos of babies born with abnormally small heads. The second and largest wave of posts started in February 2016, following the PHEIC declaration by the WHO early in that month. The third wave happened in April, when the CDC reviewed evidence and found causal connections between Zika virus infection and severe fetal brain defects. The fourth wave happened between
May and June, following the WHO declaration about the Zika virus outbreak and the 2016 Summer Olympics, in Rio de Janeiro. The fifth and last wave of posts was between August and September, following the 2016 Summer Olympics, in Rio de Janeiro.

The highest peak of Zika-related tweets over the chosen time frame was on February 5, 2016, when there were 57,423 posts relating to the conversation. On this same day, the Brazilian Oswaldo Cruz Foundation (FIOCRUZ) announced during a press conference that the Zika virus was found actively in human urine and saliva. Consequently, the scientists leading the testing said there was the possibility of passing the infection from a person who has the Zika virus systematically through those fluids. A significant number of sampled posts during that day discussed the implications of the FIOCRUZ findings, particularly the possibility of transmission through saliva and the proximity to Carnival season.

**Finding 3:** The most influential authors were Brazilian news organizations. However, Brazilian politicians and popular entertainers also made to the list.

Within the Zika virus conversation, data showed that 7 out of the 10 most influential Twitter authors were news organizations. A Brazilian popular singer also made to the list, along with a satire/humor YouTube channel profile, and the official account of the Brazilian president at the time, Dilma Rousseff. The Klout Scores was used to rank the most influential users by measuring the size, content created and interaction. The top five influential authors are listed below, along with their Klout Scores and illustrating sample tweet with a large number of impressions.
The most influential author was Rede Globo (@redeglobo, Klout score 90), which is the Twitter handle of the largest TV network in Brazil. Their sample tweet talks about a news report looking back one year after the Zika virus confirmation in the country and explores the implications of disease in the cases of microcephaly.

The second most influential author was Dilma Rousseff (@dilmabr, Klout score 87), which is the Twitter handle of the former president of Brazil and the first woman to have held the Brazilian presidency. Her sample tweet talks about the WHO Director-General who praised the “transparency and commitment” with which Brazil is treating the Zika virus epidemic.

The third most influential author was Folha de S. Paulo (@folha, Klout score 86), which is an influential Brazilian daily newspaper. Their sample tweet talks about a news story reporting the first case of Zika virus infection in the state of Texas, U.S.

The fourth most influential author was Revista Epoca (@revistaepoca, Klout score 85), which is a Brazilian weekly news and analysis magazine part of the Rede Globo media conglomerate. Their sample tweet talks about how new research points to the end of the Zika virus epidemic within the next three years.

The fifth most influential author was Ivete Sangalo (@ivetesangalo, Klout score 84), which is a renowned Brazilian singer. Her sample tweet links to a publicity piece she did with SBP (a company that produces pesticides and insect repellents) promoting their repellent against the Zika-virus mosquito.

4.2 Thematic findings

A closer look at the data collected revealed the presence of five main topics, which were (i) news commentary; (ii) case reporting; (iii) general positive; (iv) general neutral; (v) general
negative; and (vi) off-topic. Stemming from these topics, two main themes became clear, which
will be explored further in the discussion (a) Zika according to them (b) Zika according to me.

I trained Crimson Hexagon’s software by hand-coding individual tweets into the five
predefined topics/categories, to measure the proportion of conversations of each category. The
topics/categories will be briefly described below with the use of illustrative examples from the
data in order to provide the reader with the possibility of better understand the findings, and
ultimately give basis to the discussion. As a native speaker of Portuguese, I translated keywords
identified in this section to English.

(i) News commentary

In this category, I included posts that either reproduced the content of news reports or
represented a commentary about news reports, as shown in Figure 5. The tweets in this category
shared characteristics such as citing a source and using a hyperlink to direct the reader of the post
to the original news report. Messages that mentioned or cited public releases from governments,
health authorities or research institutions were also included.

![Figure 5 – Tweet representing an example of the news commentary category](image-url)
(ii) Case reporting

In this category, I included posts that reported cases of Zika virus infection generated by users. I also included posts that reported symptoms or suspects of the disease by the individuals. The post depicted in Figure 6 reads “Fuck, I have Zika, I cannot believe,” in Portuguese.

![Tweet representing an example of the case reporting category](image)

(iii) General positive

In this category, I included posts that mentioned something positive in relation to the Zika virus outbreak, such as vector-control and government programs, vaccines and research development, and even just positive wishes, praises, and prayers so God will help them overcome the disease. The post depicted in Figure 7 reads “Recipe for homemade insect repellent to drive out the mosquito of the dengue, Zika, and chikungunya,” in Portuguese.

![Tweet representing an example of the general positive category](image)
(iv) **General neutral**

In this category, I included posts that included general comments about the disease, without being too specific about their sentiment. The tweets in this category mentioned symptoms and general information about the disease. In the following conversation, a Twitter user explains to another that Zika cannot be transmitted through a hug. The post depicted in Figure 8 reads “*That’s not how it is transmitted, man* (characters denoting laugh). *It is through a mosquito, cousin of the dengue*” in Portuguese.

![Figure 8 – Tweet representing an example of the general neutral category](image)

(v) **General negative**

In this category, I included posts that were negative about the disease or denoted fear, alarm, and distress. The post depicted in Figure 9 reads “*Seriously, I am using a lot of insect repellent because I am so afraid of this so called Zika virus*” in Portuguese.
(vi) Off-topic

In this category, I included posts that were off-topic. In Brazil, the word Zika was oftentimes associated with bad luck, particularly in soccer-related commentaries. In other cases, the word Zika was positively associated with the qualities of someone incredible. I subscribe to the idea that this phenomenon could have a relation with the rapid growth of the disease in the country, and its pervasiveness. The post depicted in Figure 10 tags another user on Twitter and reads “@Thiago10s my idol is Zika, man” in Portuguese.

4.2.1 Proportion and distribution of topics in the analysis

The application of the six predetermined categories to the data using Crimson Hexagon software’s algorithm revealed another three major findings, which reflect the way in which topics were distributed and shifted over the delimiting date range of this thesis:
4. The large proportion of off-topic Zika-related mentions was overshadowed and decreased significantly with the spike in the proportion of Zika-related news commentaries, which could suggest a change in the tone of the conversations.

5. The spike in the proportion of Zika-related news commentaries caused significant changes to all the other Zika-related categories, which could suggest a change in the topic of the conversations motivated by the 24-hour news cycle.

6. The waves and spikes in the proportion of Zika-related news commentaries coincided with external events and coincided with increases in positive and negative general comments. This could suggest that the relation between the media and external events have great influence on inducing positive or negative online public behavior.

The next paragraphs will follow with a brief discussion and rationale for the findings in details. I will use visualizations generated by Crimson Hexagon’s software to illustrate the findings and provide the reader with the possibility of better understand the relationships, and ultimately give basis to the discussion.

**Finding 4:** The large proportion of off-topic Zika-related mentions was overshadowed and decreased significantly with the spike in the proportion of Zika-related news commentaries, which could suggest a change in the tone of the conversations.

The proportion of off-topic mentions was consistently high from January until early November 2015, as shown in Figure 11. However, the situation started to change in late
November, when there was a significant increase in the proportion of news commentaries (from 29% between the weeks of November 15 until 21 to 53% between the weeks of November 29 until December 5). The proportion of off-topic mentions starts to increase again and the news commentaries around the Zika virus start to decline only between August and September 2016, which coincides with events such as the 2016 Summer Olympic Games, in Rio de Janeiro. This relation between off-topic mentions and news commentaries could suggest a change in the tone of the Zika-related conversations, mostly driven by the spike in news commentaries.

Figure 11 – Chart representing the proportion of Zika-related posts over the chosen date range, by color-coded category

**Finding 5:** The spike in the proportion of Zika-related news commentaries caused significant changes to all the other Zika-related categories, which could suggest a change in the topic of the conversations motivated by the 24-hour news cycle.
The chart shown in Figure 11 reveals how the high proportion of Zika-related news commentaries seemed to have influenced various shifts in all the other Zika-related categories. Most of these shifts could be related to external events that might have had positive, neutral, or negative influence in the online conversation (this will be explored further in the discussion). The shrink in the proportion of case reporting is also noticeable, which could suggest the importance social media monitoring would have had during early stages of the outbreak.

**Finding 6:** The waves and spikes in the proportion of Zika-related news commentaries coincided with external events and coincided with increases in positive and negative general comments. This could suggest that the relation between the media and external events have great influence on inducing positive or negative online public behavior.

The spikes of news commentaries coincide with news events, such as the confirmation of the Zika virus in Brazil by the country’s health ministry in mid-April 2015. The proportion then retracts and stays dormant until early November 2015, when news reports began to emerge of babies born with microcephaly, followed by the confirmation of the link between Zika virus infection and microcephaly by Brazil health ministry in late November. The general negative and positive also followed trends such as the events leading up to the Olympics and the Games (August 2016) itself and programs for vector control launched by Brazil’s government in December 2015.
4.3 Chapter summary

This chapter presented six major findings that emerged from this thesis. The first three findings stemmed from a general overview of quantitative and qualitative data generated by Crimson Hexagon’s data collection tool. These initial three findings provided an insight on issues such as the overwhelming number of Zika-related posts published within the delimiting time frame, which indicates a high level of engagement among Twitter users in Brazil. The correlation between the volume of Zika-related tweets and the timeline of external events of the outbreak, particularly in Brazil. And that the most influential authors of Zika-related tweets were Brazilian news organizations; however, Brazilian politicians and popular entertainers also made to the list.

The second group of findings stemmed from a more specific thematic analysis of the content, which was made possible by using the steps described by Braun and Clarke (2006) combined with Crimson Hexagon’s workflow in order to analyze the textual content, and classify the statistical patterns into predetermined categories, which were trained and hand-coded by me. These findings were relevant for this thesis, as I sought to make sense of the online conversations around the 2015-16 Zika virus epidemic, in Brazil. The second group of findings included insights on a possible change in the tone of the Zika-related conversations, reflected on the large proportion of off-topic mentions being overshadowed and decreasing significantly with the spike in the proportion of news commentaries. The data also suggested a change in the topic of the Zika-related conversations, plausibly fueled by an active influence of the media. And finally, the spikes in news commentaries coincided with news events, such as the confirmation of the Zika virus in Brazil by the country’s health ministry, and the WHO’s PHEIC declaration. There were
also shifts in the proportion of general positive, negative, and neutral categories, coinciding with the news events.
Chapter 5: Discussion

The purpose of this study was to explore the trending conversations on Twitter during the Zika virus epidemic in Brazil. This project was based on the following three research questions:

1. How can social media data provide insights into the role of the media during global health crisis?
2. How does social power play on social media?
3. What were the trending conversations on Twitter during the Zika virus epidemic in Brazil?

I thought the understanding of these questions would allow more informed decisions by the media and global health actors, helping put forward health communication strategies and reach mass audiences.

The previous chapter presented the findings of this research stemming from various categories and topics. In this discussion, I looked for interconnectedness between those findings and the research questions, as well as the contrasts raised by existing literature.

It is important to highlight that I subscribe to a social constructivist framework, which has an emphasis on discovery, description and the multiplicity of understandings. Therefore, there was not a search for the objective truth in the following discussion. Instead, I looked into different ways of constructing meaning from the data provided by the data collection process. This is particularly relevant for this thesis because it relies on language, which is a complex sign system that allows access to the other’s subjectivity (Berger & Luckmann, 1967). Ultimately, I sought to make sense of the language and the consequent predetermined categories to make sense of the conversation around the 2015-16 Zika virus epidemic.

As a means of illustrating the analytic process, I revisited the findings from the perspective of the three proposed research questions of this thesis. This process is listed below.
5.1 How can social media data provide insights into the role of the media during global health crisis?

The findings provided a rich framework for understanding how social media data can help rethink the role of the media in health promotion. The results seemed to reflect that news stories have significant impact on how the online community engaged in the Zika conversations. It is also important to notice in this process the increases in news commentaries, which were occasionally followed by peaks in positive or negative sentiment associated with external media events, such as the 2016 Summer Olympics, and public announcements from governments or health authorities.

Moreover, the media appeared to have the capacity of changing the tone of the Twitter conversations, as the large proportion of off-topic mentions decreased significantly with the spike in the proportion of news commentaries associated with the Zika virus. The occasional shifts in the 24-hour news cycle likely gives support to this claim, as different waves representing increases in news commentaries could be observed in the data, associated with decreases in off-topic messages.

The similarities between peaks in news commentaries and external media events appear to give support to the claim that they were relatable, which could have significant implications. Research has shown that mass media can provide important information during outbreaks and consequently induce positive healthy behavior practices in individuals, helping reduce the probability of contracting certain diseases and affecting vaccination uptake (Collinson, Khan & Heffernan, 2015). However, the same study also describes the effects of a media fatigue, which makes individuals “relax their health behaviors” (p. 1). These different phenomena can have a significant influence on epidemic outcomes.
Mainstream news outlets also played a key role in affecting the online conversations on the Zika virus, as findings demonstrated that 7 out of the top 10 most influential Twitter authors were news organizations, which reinforces the idea of inherited structural power of the most dominant actors (Callison & Hermida, 2015).

5.2 How does social power play on social media?

The findings revealed that the most influential authors were news organizations, politicians and popular entertainers. These dominant actors appeared to reflect on social media the established elites with inherited structural power (Callison & Hermida, 2015). On a separate note, the data also revealed a large number of user-generated content that was not directly linked to news commentaries, which means that people were not directly sharing Zika-related information directly linked to news websites or public health agencies (see Fu et al., 2015).

Over 34% of the total proportion of Zika-related posts were news commentaries, while 66% of the content shared was “generated” by the users. Therefore, most the conversation observed over the chosen time frame seemed to be divided between two different perspectives: Zika according to me (individual), as opposed to Zika according to them (the media, government, and international actors). This information could shed a light on emerging actors in the conversation, and what is their role in shaping it.

It is also necessary to reflect on who is this social media data is serving. The Twitter data gathered for this project with Crimson Hexagon is not available to all individuals. Instead, it is accessed through expensive data collection tools, such as the one I used. These tools are mostly targeted to commercial users at a high price tag, which helps them reinforce market ideologies and the neoliberal agenda, where individuals’ data is made available in the open market. It is
important to acknowledge this political dynamic of exploitation in the neoliberal agenda and that it is not necessarily a “win-win situation” for global public health, particularly in countries with lower capacity. Therefore, even though all these social networking platforms are freely available for users worldwide, the methods for accessing and using this vast source of information to promote health equity are not.

The growing body of research (Southwell et al., 2016; Majumder et al., 2016; Fu et al., 2016; McGough et al., 2017) using these tools to better health communication strategies represents undeniable advances, but they could be considered highly experimental, as most of these methods developed by researchers are not implemented or made available in a sustainable way to global health actors. Additionally, most of the research is published from a neocolonial and high-resource settings perspective, which means it often ignores the underlying determinants of health and local capacity building.

5.3 What were the trending conversations on Twitter during the Zika virus epidemic in Brazil?

The findings presented in page 38 revealed a wide range of topics among the conversations, and how they shifted over the chosen time frame. As mentioned before, the link between external media events seemed to be highly influential to the shifts observed in the data.

The General findings pointed at two streams of topics in the conversation. The first were the risks associated with the infection, such as microcephaly, represent by words such as abortion, death, and fear. The second stream were the concerns with the political implications of the Zika virus outbreak, represented by words such as Olympics, government, and WHO. Both streams coincide with the topics found by Aguiar and Araújo (2016) in their analysis of the
Brazilian media. Therefore, it is possible to make a connection between the Brazilian media discourse with the topics found on Twitter during the Zika virus outbreak. And even though news commentaries were a significant proportion of the Twitter conversation (34%), most of it was generated by users themselves (66%). This could suggest an active role of the media in indirectly influencing the online conversations (see Collinson, Khan & Heffernan, 2015).

The Thematic findings pointed at more specific topics and how they changed over the time frame. The main findings from this section pointed to a change in the tone of the conversations, as off-topic Zika-related mentions were overshadowed and decreased significantly with the spike in the proportion of Zika-related news commentaries. The findings also pointed to a change in the topic of the conversations, coinciding with external news events such as government announcements. Changes in the proportion of positive, negative, or neutral commentaries were also found in the data. This information further supports the relation between media influence on the Twitter conversations, discussed in the previous paragraph.

These processes of changes in the Twitter conversation happened in a challenging scenario of political and scientific uncertainty in Brazil, which was reflected in a media coverage marked by the use of “narratives of suffering,” portraying disenfranchised mothers as “the poor, ignorant” (Aguiar & Araújo, 2016). Ultimately, this leads to perpetuating health inequities and even feeding popular misconceptions based on fear.

Exploring the trending conversations on Twitter during the Zika virus epidemic in Brazil provides a framework for understanding the role of the media and other global health actors in shaping the public opinion during a health crisis. Moreover, it offers a window for rethinking and reinforcing the journalistic rigor of having in-depth reporting, methodological research, and contrast through the multiplicity of voices. This is particularly important since the findings of
this project pointed at how the media still seemed to stand as a significant and influent actor in Brazil.
Chapter 6: Final considerations

My research examined the intersections of social media and health communication, focusing on the 2015-16 Zika virus outbreak in Brazil. The conclusions from this study follow the research questions and the findings of this project, which provided insights on how changes in the topics and tone of online Zika-related conversations were plausibly fueled by an active influence of the media. Therefore, this thesis provides a framework for understanding the role of the media and other global health actors in shaping the public opinion during a health crisis. Additionally, it offers a window for rethinking and reinforcing the journalistic rigor to better health communication strategies and literacy.

I also explored the nuances of how social power plays on social media. The data revealed that mainstream news outlets were among the most influential Twitter authors, but the highest proportion of Zika-related mentions were generated by Twitter users themselves, which means they were not necessarily linking their message to a news or government report. This can either reinforce the active role of the media in influencing the online conversation, or it could suggest the emergence of new and independent voices on social media.

Furthermore, relevant scholarship explored in the literature review points at how Twitter and other Internet-based platforms have helped soften, or even reverse patterns of political engagement inequality by making health information more accessible in low-literacy settings (Moorhead et al., 2013; see also Raynauld & Greenberg, 2014; Loader, Vromen & Xenos, 2014). Also, the low-cost of social media showed its ability to provide real-time effective surveillance systems (Southwell et al., 2016; Majumder et al., 2016; Fu et al., 2016; McGough et al., 2017).

On the other side, traditional inequalities in terms of overall Internet access need to be considered, as well as developing capacities that deal with these inequalities in the first place...
(Moorhead et al., 2013). A broader look into this issue suggests the necessity to look upstream at a policy level to motivate a collaborative systemic change and promote human security, particularly in countries with high income disparity, such as Brazil. Moreover, the limitations of social media use should also be acknowledged, as it can be a source of unreliable information that can reproduce and further spread popular misconceptions (Dredze, Broniatowski & Hilyard, 2016; Orizio et al., 2010; Adams, 2010). Finally, I think further research is needed to assess the long and short-term impacts of social media on health communication practices, as I could see from my research that a lot of the studies and methods used are not equally available to all individuals, and are not escalated into sustained programs that build capacity.
Bibliography


Ceron, A., Curini, L., Iacus, S. M., & Porro, G. (2014). Every tweet counts? how sentiment analysis of social media can improve our knowledge of citizens’ political preferences
with an application to Italy and France. *New Media & Society, 16*(2), 340-358.
doi:10.1177/1461444813480466

doi:10.1093/acprof:oso/9780199759477.001.0001


doi:10.1371/journal.pone.0141423

doi:10.1371/journal.pone.0141423


doi:10.1038/nature06536


dynamics during the 2015-2016 colombian zika virus disease outbreak. *JMIR Public Health and Surveillance, 2*(1), e30. doi:10.2196/publichealth.5814


World Health Organization (2016, November), *Fifth meeting of the Emergency Committee under the International Health Regulations (2005) regarding microcephaly, other neurological*


