How information about COVID-19 travels across
the public, media science community and policy makers and affects
public behaviour

1 Project overview

From a mysterious pneumonia-like Wuhan virus in China first reported in December 2019 to coronavirus 19-nCoV , a member of coronavirus family [] in January to COVID-19 , officially renamed in February 11, 2020, COVID-19 took less than three months to reach over 29 countries from Asia, Australia, North America, Europe and Africa, as Egypt reported the first COVID-19 case on Feb 14, 2020 []. On January 30, the World Health Organization declared COVID-19 a public health emergency of international concern. By Feb xx, over 67,000 cased confirmed and 1,500 fatality had been documented with xx% in China []. It has been a short yet long ongoing journey of the disease outbreak, caused the the world unguarded with hazy challenges, fear and anxiety epidemically [], socially [] , culturally [], economically [], and politically [] in daily life (e.g., to mask or not to mask , and etc). Dance inside the hospital

1.1 Aims

- Potential to contribute to the global response to COVID-19

- Social and policy countermeasures and Global Coordination Mechanism

1.2 Research Question

Our research questions were framed based on the object of developing potential social and policy countermeasures for COVID-19 outbreak (under the CIHR special call object).

1. Who does information (and misinformation) travel between scientists, public-health workers, mass media and social media?

- How does media cover and disseminate information?

- How media content are consumed among the general public, science community and the national and international govern units?

The aim of this question is to congregate contents to address the social and cultural dimensions of different population during the epidemic such as, examining how individuals and communities understand and react to the disease (Object 1)

2. How does communication affect public behaviour and the course of the outbreak? This research question studies and compares the public health institutes response and guidelines and how policy affect public health behavior in reacting to the infection? (Object 2)

(CS: by fitting data from google trends, news and twitter and public health authority)

3. How can scientists and policy-makers evaluate and improve the effectiveness of their communication? This research question aim to develop strategies to combat misinformation, stigma, and fear; and to improve public awareness and knowledge (Object 3)

(CS: by comparing policy, public reaction and effectiveness (reported by public health authority and by media))

The proposed research will improved existing infomation flow by (1) curating the infomation, (2) understand how interest groups are interpreting and communication information. The proposed research will work to bridge the gap between theoretical advances in epidemic modelling inference, public health practice and .. with interest groups.

2 Background

((mli: Everything above need to be sharper. Everything below are placeholder jargons I am copying over.)) COVID-19 is a coronavirus (CoV), that is pathogenic to humans, and the ability to directly transmit between humans has posed a significant threat and numbers compared to the previous coronavirus outbreaks: severe acute respiratory syndrome (SARS) coronavirus (2002-SARS-CoV) and Middle East respiratory syndrome (MERS) coronavirus (2012-MERS-CoV). Public health officials have reacted quickly in response to the outbreak; many researchers in the scientific community have published their preliminary analyses of the outbreak as pre-prints to estimate potential risk. **((mli: how are media/public responding to the news/preliminary analysis from the science folks? Below is how the scientist are responding.))** However, there are a lot of uncertainties in both modeling and effective control responses that can cause miscommunication between modellers and public and greatly affect control. **((mli: What do we know that is true?))** Without an effective vaccine available in time, the current control strategy is strictly quarantine and isolation and much are still unknown how effective these countermeasures are in controling the spread of the disease. We propose to ... and why it is useful.

((mli: The next set of subsections is layout the problem and what we are proposing to attack the problem))

2.1 Timeline and info flow

((mli: Info vs misInfo))

2.2 What do people want?

((mli: Maybe look at RQ.md?))

2.3 Communication between interest groups

((mli: Communication vs miscommunication))

2.4   Interest group's response

((mli: again, some jargon placeholder text)) As the COVID-19 continues to spread globally and recently annouced as the new pandemic, many researchers have published their preliminary analyses of the outbreak as pre-prints, focusing in particular on estimates of the basic reproductive number R0 and predictions. However, estimates of R0 vary widely across from different research groups depending on the estimation methods, data and their underlying assumptions. For example, are researchers using the same data and adaptive to changes in case-defintions (i.e. fitting window and incidence vs cumulative cases), what assumptions are researcher using for generation interval distribution (are they using serial interval, and are they considering variation), are researchers accounting for different forms of uncertainties and heterogenetities. These modelling choices often lead to disparate set of estimates which can lead to possible different communication messages to public health and assessment of the situation and intervention planning.

Among the COVID-19 candidate vaccines, the information available on possible candidate vaccines and the COVID-19 epidemiology is very preliminary, no licensed vaccine currently exists and not for a long time. Prevention of COVID-19 transmission primarily relies on non-pharmaceutical interventions such as quarantine and isolation, which aim to prevent contact with the infectious individuals. The extremely high infectiousness of CoV related cases in general are in hospitals and health facilities settings makes health workers protection an important part of COVID-19 control. Thus, it is important to figure out the specific locations and settings cases and fatalities occur to guide effectiveness control.

3   Methods and feasibility

3.1   Data and Sources

We have access to ongoing data streams of differing levels of detail.
((mli: MLi's stuff))

HuBei Province   Daily updates from the Provincal Health Commission of HuBei are avalible online. These updates contains information on up-to-date cumulative cases, new cases per day, number of confirm case hospitalized, severity categorizations and deaths for all cities in HuBei province.

Surveillance   Daily updates from outside of HuBei are avalible online. These updates contains information on up-to-date cumulative cases, new cases per day, number of confirm case, severity categorizations and deaths for different locations.
((mli: CFS's stuff))
(ref theme analysis in the H1N1 paper, and PMID: 31956275) Data published from November, 2019 to present (or one or two month after the outbreak waned)

Media   Newspaper will be collected from country of Singapore, Taiwan, Hong Kong, Japan (?), Canada and USA (and England?) based on the top circulation (either in English or Chinese)

- USA: Associated Press (AP) newswire; U.S. English-language newspapers of top high-circulation; most viewed youtube (and broadcast news transcripts from top networks?). Lexis-Nexis will be used to collect texts from the first three sources, and the MIT MediaCloud database to collect texts from websites. Texts were collected from both sources using broad search terms citepLiuSieg19 ; and [World Journal](https://www.worldjournal.com/)

- Canada: (similar to USA), and [Sing Tao Daily](https://www.singtao.ca/toronto/?variant=zh-hk)

- Taiwan: the top x newspapers. https://www.ncl.edu.tw/

- Hong Kong: [Epoch Times](https://www.epochtimes.com/gb/news415.htm)

- China: [The Peoples Daily English](http://en.people.cn/), [The Peoples Daily English](http://www.people.com.cn/), [China Daily](http://global.chinadaily.com.cn/)

- Singapore ??

- Japan: [Japanese Times](https://www.japantimes.co.jp/)

- Twitter: (CS: How does it work internationally outside Canada and USA?). It will be great if we can have social media (instgram or FB in Taiwan, Singapore etc)

- Main newspapers in China and their official health bureau websites are included for analysis but not social media. Social media such as twitter and Facebook are barred in China. Yet, there are information and coverage about the outbreak inside China in the media we analyze either from direct interview or re-twittes.

Google Trend data

- Google Trends provides search around the world by region and county. Research showed that Google Trend can be influenced by the media attraction than by true epidemiological burden [], yet more evidences suggested that it reflects awareness-related burst of searches citepBousAgac17, MahrBrag19

- By analyzing google trends data, the daily infected cases and fatality and frequency of media coverage, we use Google Search Trends and media coverage as a proxy indicator for information dissemination and public reaction

- e.g.

- Google Trend data can be very confusing, for [example] Why would Wuhan topped coronavirus and in Taiwan?

Science community report R0 of the virus transmission

- Science publication: preprint (by searching Google Scholar, arXiv, bioRxiv, medRxiv, and SSRN) and peer-reviewed (by search Google Schoolar, PubMed, Embase, Medline,

and OVID)

- The epidemic findings from science community is part of the information the public, media and international community rely on. They play a powerful role during public health crises due to the time urgency with which they can disseminate new information, accurate or not. (ref: [Early in the Epidemic: Impact of preprints on global discourse of 2019-nCoV transmissibility] We can also check findings of preprint (speedy information delivery, lack of peer review) in terms of accuracy and how misinformation get circulated based on those findings. (see [examples]

- (CSs confusion: Is context same as noise in modelling (e.g., Park and Dushoff 2020): dynamical or process; noise(randomness directly or indirectly affecting disease transmission); observation noise(randomness underlying how many true cases reported). Based on their definition, how do context or media, public reaction and pubic policy-guideline fit into models? Is it a right question?)

- (CS: It will be good to incorporate modelling ideas in this project but WHAT? Also, how workable is it to compare this coronavirus to 2012 MERS and 2002 SARS in terms of the spread and reaction context ?)

Official websites to analyze social and policy countermeasures   How does the national and international community react to the disease outbreak and with what guidelines (analyzing official websites, such as WHO, CDC, Chinese Center for Disease Control and Prevention, etc)

Official health websites
- WHO, [US CDC]

- [National Health Commission of PRC

- Canada PHAC

- Taiwan CDC

- Singapore CDC

- [Hong Kong CDC]

We will include the disease control and prevention centre in Japan later if ??.

We will add sources (e.g., Central China Normal Univeristy, Wuhan University etc for our contextual analysis.

We will need to curate the data-stream.

3.2   Data mining

textual analysis of news media and Twitter are implemented by AI based on coding book which will be based on a pilot study. The findings of textual analysis Human Coders and AI

(?):

- (fill in what it is about)

- (fill in sentences on AI part)

- (fill in sentences on coders)

- The textual analysis will be conducted based on a coding book we will draft when the funding is approved. The coding book should list themes and frames for analyzing the texts, which will be categorized based on a pilot study as part of drawing the coding book. The basic structure of the themes will outline what (e.g.), in addition to information types (e.g., news, information and misinformation/fake news based on sources: verified, anonymous, speculation), who (e.g. active vs. passive subject), where (e.g., country, region and city), and theme tone (e.g., negative, positive, neutral). (see attached Basic Coding frame for what themes are like)

3.3   Analysis

Contextual analysis by researchers

- (fill in what contextual analysis is)

- How media content and information are framed are often socially and politically cultivated []. We will implement contextual analyses of the quantitative findings and elaborate how information is framed and disseminated in this uncertain circumstances into context.

- summarizing and comparing policy, guidelines and recommendation by national and international health organizations.

Modelling   (Can we do this?: incorporating media coverage, public and institutional reaction and report (which can be based on modelling results by scientific community and WHO etc), time travelling and spread network (geographical)

- key words search for Google trend analysis, and official public health websites: coronavirus, 19-nCov, COVID-19, influenza, (Wu Virus) (Wuhan Virus)(Wuhan) (Wuhan pneumonia) (pneumonia) (coronavirus), (novel coronavirus), (Wuhan novel coronavirus)

general public

- Data: travel, movies, restaurants in Taiwan, Singapore, H.K., Canada and USA; or should we focus on North America and on travel and movie box office?

deliverables

- software

- Communication: Writing papers is certainly one, but we should be tweeting and etc. What are communication platforms we should be communicating in?

4   Research Setting & Personnel

The principal applicants have a long history of influential research on disease outbreak:

The research will principally take place at McMaster University. Nominated principal applicant Dr. David J.D. **Earn** (6 hours per week) led the creation of the International Infectious Disease Data Archive [33] and has expertise in gathering and curating infectious disease data, and in dynamical modelling. Principal applicant Dr. Jonathan **Dushoff** (5 hours per week) is an internationally recognized expert in infectious disease modelling, has extensive experience with statistical frameworks for fitting models to data, and has been involved in the Ebola challenge and other forecasting projects. Co-Applicant Dr. Benjamin **Bolker** (5 hours per week) is a highly accomplished ecological statistician with extensive experience in spatial-dynamical modeling, statistical modeling and statistical software. Co-Applicant Dr. Chyun-Fung **Shi** (40 hours per week), post-doctoral researcher, has ... Co-Applicant Dr. Michael **Li** (5 hours per week), post-doctoral researcher, has focused his research on epidemic forecasting and is experienced working with large databases.

((mli: fill in the rest later))

4.1   Collaborators and Knowledge Users

((mli: fill in later))

5   Research Time line

5.1   Year 1

5.2   Year 2

5.3   Potential Outcomes

Transfer findings to peer-reviewed publications. In addition to a grand one paper, we plan to transfer findings into subgroups, such as by affected country (North America, Asia-Taiwan, H.K. and Singapore), types of media (news media and social media (twitter)).

- Potential to contribute to the global response to COVID-19

- Social and policy countermeasures and Global Coordination Mechanism

6   Challenges and Mitigation Strategies

Data curation   We will go back and document clearly all the policy changes and case defintions.

- data collected from National Health Commission

- Figure out how to use the data effectively (e.g. we are not using death, severity categorizations, number of tested, number of positive, and etc)

- Case definitions

- Media content

- Social media platforms

- Google trends

Analysis/ Pipelineing/ mainstreaming   Info delays, misinformation and miscommunications

Communication   how people are interpreting

7   Conclusion/summary

7.1   Note

We want to emphasize our bilingual (English and Chinese) background, which is essential in this study.

References