

# 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 [1] 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 [2]. On January 30, the World Health Organization declared COVID-19 a public health emergency of international concern. By Feb xx, over 67,000 cases confirmed and 1,500 fatality had been documented with xx% in China [3]. 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 [4], socially [5], culturally [6], economically [7], and politically [8] in daily life (e.g., to mask or not to mask, and etc)

### 1.1 Research questions/objectives (Aims)

((mli: write a line or two after each question?))

- Who does information (and misinformation) travel between scientists, public-health workers, mass media and social media?
- How does communication affect public behaviour and the course of the outbreak? What are the side effects of changes in public behaviour?
- How policy correlate with health behavior in contributing/reducing/avoiding the infection?
- How can scientists and policy-makers evaluate and improve the effectiveness of their communication?

The proposed research will improved existing information flow by (1) curating the information, (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.

1. Improve communication between modellers, public health, and interest groups
2. Curate info

## 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 controlling 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 announced 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  $R_0$  and predictions. However, estimates of  $R_0$  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-definitions (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 heterogeneities. 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 pipelines

We have access to ongoing data streams of differing levels of detail.

**HuBei Province** Daily updates from the Provincial Health Commission of HuBei are available 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 available 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.

**Google Trend data** 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

**Science community report R0 of the virus transmission** The epidemic findings 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)

**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)

We will need to curate the data-stream.

#### Analysis

- Correlation: between data or interest groups?

#### Communicate

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

**deliverables** software, and else?

Not sure we need this, but let's put it here for now.

## 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

## 6 Challenges and Mitigation Strategies

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

- 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

## References