ESTIMATING ASYMPTOMATIC AND SYMPTOMATIC TRANSMISSION OF NOVEL CORONAVIRUS DISEASE 2019 IN SELENGE PROVINCE, MONGOLIA

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Influenza and Other Respiratory Viruses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>IRV-2023-277</td>
</tr>
<tr>
<td>Wiley - Manuscript type:</td>
<td>Original Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>18-Jul-2023</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Dambadarjaa, Davaalkham; Mongolian National University of Medical Sciences</td>
</tr>
<tr>
<td></td>
<td>Mend, Tsogt; National Center for Communicable Diseases, Department of</td>
</tr>
<tr>
<td></td>
<td>Surveillance and Research</td>
</tr>
<tr>
<td></td>
<td>Shapiro, Andrew; University of California Los Angeles, Department of</td>
</tr>
<tr>
<td></td>
<td>Statistics</td>
</tr>
<tr>
<td></td>
<td>Mandakh, Undram; Mongolian National University of Medical Sciences,</td>
</tr>
<tr>
<td></td>
<td>Family Medicine</td>
</tr>
<tr>
<td></td>
<td>Enebish, Temuulen; World Health Organization Regional Office for the</td>
</tr>
<tr>
<td></td>
<td>Western Pacific, COVID-19 Incident Management Support Team</td>
</tr>
<tr>
<td></td>
<td>Le, Linh-Vi; World Health Organization Regional Office for the Western</td>
</tr>
<tr>
<td></td>
<td>Pacific, COVID-19 Incident Management Support Team;</td>
</tr>
<tr>
<td></td>
<td>Bandoy, Darwin; World Health Organization Regional Office for the Western</td>
</tr>
<tr>
<td></td>
<td>Pacific, COVID-19 Incident Management Support Team;</td>
</tr>
<tr>
<td></td>
<td>University of the Philippines Los Baños College of Agriculture and Food</td>
</tr>
<tr>
<td></td>
<td>Science, College of Veterinary Medicine</td>
</tr>
<tr>
<td></td>
<td>Amarjargal, Ambaselmaa; National Center for Communicable Diseases,</td>
</tr>
<tr>
<td></td>
<td>Department of Surveillance and Research</td>
</tr>
<tr>
<td></td>
<td>Altangerel, Bilegt; National Center for Communicable Diseases,</td>
</tr>
<tr>
<td></td>
<td>Department of Surveillance and Research</td>
</tr>
<tr>
<td></td>
<td>Chuluunbaatar, Tuvshintur; National Center for Communicable Diseases,</td>
</tr>
<tr>
<td></td>
<td>Department of Surveillance and Research</td>
</tr>
<tr>
<td></td>
<td>Guruuchin, Uugantsetseg; National Center for Communicable Diseases,</td>
</tr>
<tr>
<td></td>
<td>Department of Surveillance and Research</td>
</tr>
<tr>
<td></td>
<td>Lkhagvajav, Oyuntulkhuur; National Center for Communicable Diseases,</td>
</tr>
<tr>
<td></td>
<td>Department of Surveillance and Research</td>
</tr>
<tr>
<td></td>
<td>Enebish, Oyunsuren; Mongolia Ministry of Health</td>
</tr>
<tr>
<td></td>
<td>Handcock, Mark; University of California Los Angeles, Department of</td>
</tr>
<tr>
<td></td>
<td>Statistics</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Coronavirus disease 2019, transmission, surveillance</td>
</tr>
</tbody>
</table>

This version of the article has been accepted for publication, after peer review and is subject to the publisher's terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1111/irv.13277
This work is dedicated to Andrew Shapiro.

Andrew's intelligence, insight, skill and perseverance are all expressed in this article.
TITLE: ESTIMATING ASYMPTOMATIC AND SYMPTOMATIC TRANSMISSION OF NOVEL CORONAVIRUS DISEASE 2019 IN SELENGE PROVINCE, MONGOLIA

Article Type: Original Article

Authors:
*Davaalkham Dambadarjaa, *Tsogt Mend, Andrew Shapiro, Undram Mandakh, Temuulen Enebish, Linh-Vi Le, DJ Darwin R. Bandoy, Ambaselmaa Amarjargal, Biligt Altangerel, Tuvshintur Chuluunbaatar, Uugantssetseg Guruuchin, Oyuntulkuur Lkhagvajav, Oyunsuren Enebish, Mark S. Handcock

Author Affiliations:
1 School of Public Health, Mongolian National University of Medical Sciences, Ulaanbaatar, Mongolia
2 Department of Surveillance and Research, National Center for Communicable Diseases, Ulaanbaatar, Mongolia
3 Department of Statistics, University of California, Los Angeles, USA
4 Department of Family Medicine, School of Medicine, Mongolian National University of Medical Sciences, Ulaanbaatar, Mongolia
5 COVID-19 Incident Management Support Team, World Health Organization Regional Office for the Western Pacific, Manila, Philippines
6 College of Veterinary Medicine, University of the Philippines Los Baños, Laguna, Philippines
7 Mongolian Red Cross Society
8 Ministry of Health, Mongolia
* Equal contribution

Correspondence: Undram Mandakh
Email: undram@mnums.edu.mn

Short title: Transmission of COVID-19 in Mongolia

Keywords: Coronavirus disease 2019, transmission, surveillance
ABSTRACT

Background: Following a locally transmitted case in Sukhbaatar city, Selenge province, we conducted a study with two objectives. First, we aimed to estimate the basic reproduction number of COVID-19, leveraging the epidemiological and clinical characteristics observed in the first 67 confirmed cases. Second, we aimed to model the outbreak considering different patient profiles - asymptomatic, symptomatic, and pre-symptomatic - with the goal of predicting the ultimate scale of the epidemic in the scenario of uninterrupted transmission.

Methods: We conducted a prospective case study following the WHO FFX cases generic protocol. The rapid response teams collected the surveillance data from November 14–29, 2020. We created a stochastic process to draw many transmission chains from this greater distribution to better understand and make inferences regarding the outbreak under investigation.

Results: The majority of the cases involved household transmissions (35, 52.2%), work transmissions (20, 29.9%), index (5, 7.5%), same apartment transmissions (2, 3.0%), school transmissions (2, 3.0%), and meetup transmissions (1, 1.5%). The posterior means of the basic reproduction number of both the asymptomatic cases, $R_0^{asy}$ and pre-symptomatic cases $R_0^{pre}$ (1.35 [95% CrI 0.88–1.86] and 1.29 [95% CrI 0.67–2.10], respectively), were lower than that of the symptomatic cases.

Conclusion: Our study highlights the heterogeneity of COVID-19 transmission across different symptom statuses and underscores the importance of early identification and isolation of symptomatic cases in disease control. Detailed contact tracing data with advanced statistical methods, can be applied to other infectious diseases, facilitating a more nuanced understanding of disease transmission dynamics.

INTRODUCTION

Mongolia is located between China and Russia and is a sparsely populated country with a population of 3.5 million. Almost 50% of the total population lives in Ulaanbaatar, making the capital city the most densely populated area in the country.14 Although Mongolia is a neighbouring country to China, it was able to delay an outbreak and domestic transmission for an extended period. The country was one of the first to close its borders in response to the Wuhan outbreak in early 2020 and eventually closed schools and kindergartens, in addition to imposing travel restrictions and allowing only charter flights.1 Initial measures, as directed by the State Emergency Committee, included screening, quarantining incoming travellers for 14 days in designated camps, cancelling national holiday celebrations, banning mass gatherings, closing rail and land crossings, and reducing domestic travel. The Government later extended the 14-day quarantine period to 21 days, as some evidence
suggested that the incubation period of the virus persists beyond two weeks, and some cases tested positive after initially testing negative.\textsuperscript{1} In response to vigorous public awareness campaigns and the promotion of protective measures, the public practised social distancing and wearing masks extensively.\textsuperscript{2}

A surveillance system was implemented for contact tracing to limit the spread of infection. Suspected cases, those who were in close contact with recently confirmed cases, or those with a high-risk travel history were quarantined at local quarantine facilities. In Mongolia, primary contact was regarded as someone who was in close contact (1 m) with a COVID-19 confirmed case 4 days before symptom onset.\textsuperscript{3}

Despite aggressive prevention and risk-reduction measures taken by the Government of Mongolia since the start of the pandemic, a truck driver under in-home observation after a 21-day mandatory stay at a quarantine facility tested positive for COVID-19 on November 10, 2020. Along with his three family members who also tested positive, they became the first local cases of COVID-19 in Mongolia.\textsuperscript{4} Concerning the unique housing conditions and the current capacity of public health response, the quarantine policy has been different in Mongolia since the beginning of the pandemic, with a 21-day mandatory quarantine of all suspected, probable, and confirmed cases and close contacts, as well as all repatriated people.\textsuperscript{5,6} To control the spread of infection, it was important to intervene in the transmission chain. Following a locally transmitted case in Ulaanbaatar, Selenge Province reported its first local transmission on November 14, 2020, without any known link to the capital city.\textsuperscript{4}

In this study, we aimed to estimate the basic reproduction number based on the epidemiological and clinical characteristics of the first 67 confirmed cases of COVID-19 in Sukhbaatar City, Selenge Province, and to simulate the outbreak based on asymptomatic/symptomatic and pre-symptomatic profiles to determine the final size of the epidemic in the case of uninterrupted transmission.

**METHODS**

**Study Design**

This was a prospective case study following the WHO FFX cases generic protocol\textsuperscript{7,8} with a few modifications based on the local context. The first 67 local COVID-19 cases in Sukhbaatar, Selenge Province, Mongolia, were identified using a real-time reverse transcription polymerase chain reaction (RT-PCR) test. Once SARS-CoV-2 infection was confirmed, the local health department monitored the cases for up to 21 days until three consecutive negative RT-PCR results were obtained. Their contacts were followed-up for 14 days with three RT-PCR tests. They were managed as confirmed cases.
according to local COVID-19 regulations if the contact test results were positive. The rapid response
teams used the national interim guidelines for COVID-19 surveillance and the WHO FFX protocol
questionnaire for data collection from confirmed cases, as well as their close contacts. They followed
the same procedure for confirmed cases if the RT-PCR test results were positive for close contacts.
The cycle threshold (Ct) value of the RT-PCR tests was reportedly dependent on the period from
infection and was useful for determining infectivity. More specifically, as the Ct value reflects the viral
load, the subjects with higher Ct-values (Ct ≥ 30) are thought to be potentially less infectious. In
particular, samples with Ct values ≥ 30 were no longer cultured and did not show infectivity. Hence,
during contact tracing, the infection potential lasted for approximately two weeks after contact, even
in asymptomatic subjects.

**Study Setting**

Selenge Province is on the country’s northern side, with a total population of 107,341 and
29,004 households. (Figure 1) It has a railway station that connects to the Russian railway system
with border crossings. Despite strict quarantine measures, a local case was reported without
connection to the first locally transmitted case in Ulaanbaatar. Later, cases arose in the province, and
local health administrative units and emergency agencies initiated rapid investigations and response
teams to detect cases and their contacts. This action was followed by a strict lockdown, mandatory
hospitalization for cases, and 21 days of contact quarantine. The Ministry of Health deployed a rapid
response team from the National Center for Communicable Diseases (NCCD) to alleviate in-province
health sector stress and workload. During an epidemiological investigation, the team found that these
cases were linked to the railway stations.
Epidemiological and laboratory investigations

We collected data on the cases and their close contacts in Sukhbaatar City, Selenge Province, from November 14–29, 2020. Demographic, clinical, and virological information was gathered based on local ethical considerations. The team used a case definition from a ministerial order at the national level. They interviewed the confirmed cases with regard to travel history, epidemiological exposure, clinical symptoms, and risk factors. A trained interviewer conducted either in-person interviews using personal protective equipment or, if possible, remote interviews using mobile phones. The hospital staff monitored the confirmed cases’ health status daily, and contacts who developed symptoms were required to provide samples and be hospitalized if SARS-CoV-2 was detected by RT-PCR.

Definition of transmissions

Below, we provide definitions that guided the data-cleaning procedures and were necessary for our modelling method. A confirmed case was a patient with laboratory confirmation of COVID-19 infection. A case was considered active when transmissive, from inception until quarantine, recovery, non-transmission, or death. A case is asymptomatic if the individual reported not developing symptoms from inception through recovery and was no longer transmissive. Asymptomatic transmission referred to the transmission of the virus from an active asymptomatic case to a secondary case. Pre-symptomatic transmission was the transmission of the virus from an active symptomatic patient to a secondary patient before the onset of symptoms. Symptomatic transmission
referred to the transmission of the virus from an active symptomatic patient to a secondary patient
after the onset of symptoms.

Data Cleaning

Each case in our dataset had a list of contact cases (all specified as not applicable [NA] if at the
start of an outbreak), a date of symptom onset (NA if asymptomatic), a range of possible exposure
dates, the end of their transmissible period (the earliest date between the date of quarantine, date of
survey, or death), and a list of symptoms, if symptomatic. If a case had a date of onset but no listed
symptoms (or minimal evidence of symptoms; this is at the discretion of the researcher), the patient
was reclassified as asymptomatic, and the date of symptom onset was NA. If the date of symptom
onset is incompatible with the range of exposure dates, occurring too soon or too late to be reasonable
for the given pathogen, either the range of exposure dates was adjusted or the case was reclassified
as asymptomatic, and the date of symptom onset was NA. If a secondary case’s range of possible
exposure dates was incompatible with the infecting case’s range of possible exposure dates, the range
of possible exposure dates for the two cases was adjusted, or the infecting case was removed from
the secondary case’s list of contact cases.

Stochastic Reconstruction of Transmission Chains

A transmission chain for the outbreak of a pandemic is a static network of confirmed cases,
where network ties go from an infecting case to a secondary case, with each transmission occurring
at a specific time and each case having a specific end in its transmissible period. In our dataset, some
patients had multiple infectors and a wide range of possible exposure dates. This means that there
was a greater distribution of all possible outbreaks that might have occurred, and a particular outbreak
was one realization of that distribution. We created a stochastic process to draw many transmission
chains from this greater distribution to better understand, and make inferences regarding, the
outbreak under investigation. Details regarding this process are provided in the Supporting
Information.

Estimation of the basic reproductive number \(R_0\)

For each realization of our stochastically reconstructed transmission chains, we estimated the
posterior distributions of \(R_0^{asy}, R_0^{pre}, R_0^{Sym}, R_0^{Tot}\), the basic reproductive numbers of cases spread via
asymptomatic, pre-symptomatic, symptomatic patients, and all transmissions. Each quantity was
modelled using a Bayesian framework, a gamma prior, and COVID-19 transmission was modeled with
a Poisson process in time. Point estimates and the corresponding 95% credible intervals (CrI) were
obtained from the posterior distributions using Monte Carlo integration. This method is an extension
of that used by Cori et al. (2013), leveraging our access to outbreak data as a transmission chain of
cases instead of aggregate daily counts of cases to determine the reproductive number of
subcategories of an outbreak in addition to the total reproductive number of an outbreak. All analyses were performed using R software (version 4.2.0). Additional details of this methodology are included in the Supporting Information, including the software code used.

**Predicting the Epidemic Outcomes had Mitigation not been Applied**

To better understand the potential impact of this outbreak in the absence of social distancing, quarantining, and other COVID-19 transmission-mitigating measures, we implemented a susceptible–infected–recovered (SIR) model that outputs estimates of the prevalence, total number of cases, incidence, number of new daily cases, and cumulative deaths. The SIR model was based on $R_{tot}^{0}$. In an SIR model, there is a fixed-size population, and all members of the population are either susceptible (not infected but can be infected), infected (have the infection and can spread it to susceptible people), or recovered (had the infection but no longer have it and cannot spread it, nor can they get the infection again). We opted for an SIR model because the data collected only allowed us to understand the transmissions that occurred and not the network of contacts that did not result in new cases. As a result, we lacked information on exposure that could be credibly used in an SIR model for this outbreak. The details of this model are provided in the Supporting Information.

**Ethics approval**

Our study was undertaken as part of the national pandemic response and received MoH Ethical clearance on 2nd February 2021, with Institutional Review Board Approval No. 202.

**Role of funding source**

This exploratory study was based on surveillance data, and supported by WHO Unity Studies, a global sero-epidemiological standardization initiative.

**RESULTS**

Between November 14 and 29, 2020, 67 patients were confirmed to have the SARS-CoV-2 infection. The index case in Selenge Province was detected on November 14th, and 21 cases were identified through primary healthcare without any known connection to a confirmed or suspected case. The majority of the cases involved household transmissions (35, 52.2%), work transmissions (20, 29.9%), index (5, 7.5%), same apartment transmissions (2, 3.0%), school transmissions (2, 3.0%), and meetup transmissions (1, 1.5%). This reveals that the home environment plays a critical role in the spread of SARS-CoV-2 infection, and understanding its dynamics is vital for implementing effective public health strategies.

Of the 67 patients, 38 (56.7%) were female, and the mean age was 36.1 (SD, 18.4). Regarding symptoms, of the 67 patients, 24 (35.8%) reported nasal congestion, 14 (23.3%) had a dry cough, 14 (22.6%) had a loss of smell and taste, and seven (11.5%) had a fever and other symptoms.
In terms of comorbidities, 10 (14.9%) patients were obese, six (9.0%) had renal disease, four (6.0%) had diabetes, two (3.0%) had cancer, and there was one case each (1.5%) of asthma and liver disease. (Table 1)

Table 1: Characteristics of confirmed COVID-19 cases in Sukhbaatar City, Selenge Province, Mongolia, from 14 to 29 November, 2020.

<table>
<thead>
<tr>
<th>General characteristics of study participants</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>38</td>
<td>56.7%</td>
</tr>
<tr>
<td>Male</td>
<td>29</td>
<td>43.3%</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-9</td>
<td>4</td>
<td>6.0%</td>
</tr>
<tr>
<td>10-19</td>
<td>10</td>
<td>14.9%</td>
</tr>
<tr>
<td>20-29</td>
<td>13</td>
<td>19.4%</td>
</tr>
<tr>
<td>30-39</td>
<td>11</td>
<td>16.4%</td>
</tr>
<tr>
<td>40-49</td>
<td>12</td>
<td>17.9%</td>
</tr>
<tr>
<td>50-59</td>
<td>10</td>
<td>14.9%</td>
</tr>
<tr>
<td>≥60</td>
<td>7</td>
<td>10.4%</td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity</td>
<td>10</td>
<td>40.0%</td>
</tr>
<tr>
<td>Renal disease</td>
<td>6</td>
<td>24.0%</td>
</tr>
<tr>
<td>Liver disease</td>
<td>1</td>
<td>4.0%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>4</td>
<td>16.0%</td>
</tr>
<tr>
<td>Cancer</td>
<td>2</td>
<td>8.0%</td>
</tr>
<tr>
<td>Heart disease</td>
<td>1</td>
<td>4.5%</td>
</tr>
<tr>
<td>Asthma</td>
<td>1</td>
<td>4.5%</td>
</tr>
<tr>
<td>Symptoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal congestion</td>
<td>24</td>
<td>38.7%</td>
</tr>
<tr>
<td>Headache</td>
<td>15</td>
<td>23.8%</td>
</tr>
<tr>
<td>Dry cough</td>
<td>14</td>
<td>23.3%</td>
</tr>
<tr>
<td>Loss of smell and taste</td>
<td>14</td>
<td>22.6%</td>
</tr>
<tr>
<td>Fever</td>
<td>7</td>
<td>11.5%</td>
</tr>
<tr>
<td>Sore throat</td>
<td>7</td>
<td>11.3%</td>
</tr>
<tr>
<td>Runny nose</td>
<td>4</td>
<td>6.5%</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>4</td>
<td>6.3%</td>
</tr>
</tbody>
</table>
Joint pain 4 6.3 %
Nausea 4 6.3 %
Shortness of breath 3 4.8 %

Social status
Employee 40 59.7 %
Student 14 20.9 %
Retired 10 14.9 %
Unemployed 2 3.0 %
Unknown 1 1.5 %

Exposure
Family 32 47.8 %
Work 20 29.9 %
Index 5 7.5 %
Relative 3 4.5 %
Lives in the same building 2 3.0 %
School 2 3.0 %
Unknown 2 3.0 %
Meetup 1 1.5 %

**Estimation of $R_0$**

Figure 2 shows the estimates of the symptomatic, asymptomatic, pre-symptomatic and total reproductive numbers of the Selenge outbreak.

The posterior means of the basic reproductive number for symptomatic cases, $R_0^{sym}$, was 2.00 (95% CrI 1.38–2.76), indicating that individuals displaying symptoms of the disease have a higher transmission potential compared to asymptomatic and pre-symptomatic cases. Notably, the posterior means of the basic reproduction number of both the asymptomatic cases, $R_0^{asy}$ and pre – symptomatic cases $R_0^{pre}$ (1.35 [95% CrI 0.88–1.86] and 1.29 [95% CrI 0.67–2.10], respectively), were lower than that of the symptomatic cases. This suggests that individuals who have not yet developed symptoms but are infected can also transmit the disease to others. Therefore, this resulted in an aggregated $R_0^{tot}$ value of 1.53 (95% CrI 1.22–1.89). (Figure 2)
Figure 2: Estimates of the reproductive numbers for three types of transmission (Asymptomatic, Pre-symptomatic, and Symptomatic). Each distribution summarizes the knowledge we have about the corresponding reproductive number and the relative probability of that value. More concentrated distributions represent more accurate knowledge, based on the study.

**Estimation of Potential Epidemic Size and Impact**

The SIR model was run with the starting conditions of $S(0) = 20,000$, $I(0) = 1$, and $R(0) = 0$ to mimic the conditions of the outbreak in Sukhbaatar City, where a single infection that started the outbreak in a city with a population of approximately 20,000. COVID-19 was modelled with a mean infectious time of $1/\gamma = 5.22$, and a case fatality rate of 0.05. Figure 3 includes estimates of the mean Prevalence, Incidence, and Cumulative Deaths according to the number of days since the outbreak (the solid black line). We also included estimates of Prevalence, Incidence, and Cumulative Deaths for a lower boundary value (the 2.5% percentile of our sample), an upper boundary value (the 97.5 quantile of our sample), and a highly probable value (the median of our sample) of $R^*_T$ to give reasonable expectations of what an outbreak might have looked like in best case, worst case, and realistic case scenarios. An important caveat is that the SIR model is run with a fixed case fatality rate. The health care facilities of a small city of 20,000 would likely be overwhelmed by a prevalence of three thousand cases, resulting in infected individuals 1) not receiving adequate care and 2) a higher case fatality rate. Therefore, for the worst-case scenarios, it is reasonable to assume that the
Cumulative Deaths are underestimated, and the effect of the outbreak in a worst-case scenario could be far more severe.

**Figure 3:** Estimates of the Prevalence, Incidence, and Cumulative Deaths over time had mitigation not occurred. The lower/upper boundaries are the 2.5%/97.5% probability events. The difference between the mean and median outcomes reflects the skewness of the distributions.
DISCUSSION

Our study found that symptomatic individuals had a higher basic reproduction number compared to asymptomatic and pre-symptomatic individuals. This discovery underscores the potential for symptomatic individuals to be significant drivers of COVID-19 transmission. However, the reproductive number for pre-symptomatic individuals was lower than that found in a similar study conducted by Seyed. M. Moghadas, et al. The lower reproduction number in our study could be attributed to the stringent public health and social measures implemented in Sukhbaatar city during our study period. We posit that the early case and contact tracing initiatives carried out in our study significantly helped in minimizing negative health outcomes.

At the beginning of the pandemic and before the availability of vaccines, most countries with emerging cases limited the spread of COVID-19 through the strict closure of schools, public services, and organizations. The efforts to contain the disease and maintain the pathogen reproductive number \( R_t \) at <1 placed a major strain on economies and societies. Nevertheless, some countries managed to successfully control transmission without imposing a mandatory lockdown. A notable example is Taiwan, which, despite its close proximity to China, had among the lowest COVID-19 incidence and mortality rates globally. Population-based measures such as face mask use, social distancing, maintaining hygiene, and case-based strategies, including case detection, contact tracing, quarantine, and surveillance, were used to adequately decrease COVID-19 transmission. Contact tracing aims to detect possible cases that were in contact with a newly identified COVID-19 patient. However, unless contact tracing capabilities are sufficient, the efficacy is reduced, mostly because of pre-symptomatic cases or delays between symptom onset and detection. In another example, Rwanda was one of the first countries in Africa to take action against COVID-19 transmission by screening all passengers from countries with confirmed cases and implementing a countrywide lockdown. The country rigorously utilizes contact tracing for early detection, and uses data to estimate secondary attack rates and spatial analysis to determine high-risk areas.

The United Kingdom was one of the first European countries to have new emerging cases. While most cases were imported, most secondary cases were close contacts. In a single-center retrospective analysis of the first 500 confirmed cases in Manila, the Philippines, 133 (26.6%) were healthcare workers (HCW), and 367 (73.4%) were non-HCW. Similarly, in Vietnam, among the first few hundred confirmed cases, 60% were imported and 43% of cases remained asymptomatic for the duration of infection. Results of previously published work on nationwide sero-prevalence of SARS-COV-2 in Mongolia showed that sero-positivity was associated with symptomatic cases and higher hospitalization rates.
A recent meta-analysis estimated household secondary infection attack rates of studies using the WHO Unity Studies Household Transmission Investigations (HHTI) and FFX methods. Study results highlighted HHTI and FFX are critical tools to characterize novel pathogens at the early pandemic stage and tailor public health and social measures around the evidence generated.\textsuperscript{31}

To conduct an FFX study, countries must prepare and plan carefully from the beginning. In future FFX studies, the local data collection capacity should be strengthened and the quality should be monitored regularly. Despite robust individual case interviews and contact-tracing records, missing data was the main challenge in this study. After a series of government responses to the pandemic, people were exhausted and refused to cooperate in filling in any missing information to collect data retrospectively.

Our model used a common relative infectivity profile for all individuals. The relative infectivity profile may vary according to an individual’s age and other characteristics. However, there is insufficient knowledge of the infectivity profiles to allow for this level of detail.

An important contribution of this work is the creation of open-source software that implements novel statistical methods. This software was written in the open-source R\textsuperscript{14} statistical language and is available in Supporting Information. While this study was based on a population where mitigation was possible, it was applied to outbreaks where mitigation was more difficult (e.g., Vietnam).\textsuperscript{32}

To avoid challenges and barriers encountered in our study, countries should integrate and digitalize health information systems (HIS) to ensure long-term possible pandemic preparedness and readiness. Resiliency should be carefully adapted to the local health systems to avoid socio-economic impacts caused by public health and social measures. These can be achieved rapid response mechanism with quality ensured epidemiological and other relevant sectors data. Despite good multi-sectoral collaborative efforts lack of digitalization and modernization of data collection was the root of the problem for further investigation and analysis.

Our approach can be applied not only to other variants of SARS-CoV-2 but also to other infectious diseases. Understanding the varying reproductive numbers among different groups can enhance contact tracing strategies and public health messaging for emerging infectious diseases. Future research should investigate how these findings can be applied across different pathogens and epidemiological contexts.

In conclusion, our findings provide new insights into the reproductive numbers among symptomatic, asymptomatic, and pre-symptomatic individuals, and underscore the importance of
robust public health measures and advanced data management systems in controlling infectious
disease transmission.

**CONTRIBUTORS**

UM, TM served as the principal investigators for the project. UM, DD, TM conducted planning
and coordination. L-VL and DD has been supervisors. AA, BA, TC, UG, OL provided coordination for the
data collection. AS, TM, did data analysis. UM, DD, TM, AS, MSH, TE came up with first manuscript.
Writing and conceptualizing has been conducted by UM, DD, TM, AS, MSH, OE, and L-VL. L-VL, MSH,
DJ DRB, UM, TM and all authors reviewed and edited the manuscript.

**DECLARATION OF INTERESTS:**

We declare no competing interests.

**ACKNOWLEDGMENTS**

This publication has been supported by WHO Unity Studies, a global sero-epidemiological
standardization initiative, with funding to WHO by the COVID-19 Solidarity Response Fund and the
German Federal Ministry of Health (BMG) COVID-19 Research and development.

We are grateful to the following collaborating partners for their contribution to the study:
Ariuntuya Ochirpurev, Nyamkhuu Dulmaa, Naranzul Dambaa, national and provincial Rapid Response
Teams, Selenge province authorities, and the World Health Organization.

**REFERENCES**


2. Dagvadorj A, Jantsansengee B, Balogun OO, Baasankhuu T, Lkhagvaa B. Health emergency

3. Ministry of Health. Health Ministerial order no.269: Interim guideline for COVID-19 surveillance,
prevention and rapid response. Published online April 29, 2020.

5. Campi A. Mongolia and the Coronavirus Outbreak -- Rethinking Strategies for Regional Integration. Mong J Int Aff. 2020;21(0):13-49. doi:10.5564/mjia.v21i0.1432


Stochastic Reconstruction of Transmission Chains

For each case, we knew the date of symptom onset (for symptomatic cases), the date of quarantining (the date when a confirmed positive case enters quarantining and can no longer spread COVID-19 to others), each case’s infector or possible infectors (in our data, some cases had as many as three contacts with confirmed positive cases that could have been the infector for the given cases; for cases with a single contact, that contact is assumed to be the infector), and the range of dates of possible incidence of each case. Additionally, we have infectivity profiles $P_{inf}^{asy}(s)$, $P_{inf}^{pre}(s)$, and $P_{inf}^{sym}(s)$, which are the probability distributions dependent on time, $s$, and independent of calendar time, for asymptomatic, presymptomatic, and symptomatic cases, respectively. The infectivity profiles for asymptomatic, presymptomatic, and symptomatic cases were those modeled from 77 transmission pairs obtained from publicly available sources within and outside mainland China and given in the supplementary materials.

Other methods of transmissivity modeling (Cori et. al. 2013) use the serial distribution, or the number of days between the days of onset between an infector and infectee case. This method works well for pathogens that are largely symptomatic. The large number of asymptomatic cases for COVID-19 makes it unreasonable to use the serial intervals, and instead we must use the generation interval, or the number of days between the incidence of a case in the infector and the incidence of a case in the infectee. This poses an
additional problem since the time of incidence is virtually impossible to witness. To overcome this, we can use the data we have to stochastically reconstruct the transmission chains of the COVID-19 breakout and effectively Monte Carlo integrate out the times of incidence to model infectivity. This also provides the additional benefit of being able to model transmissivity of different types of transmission, more specifically, but not limited to, asymptomatic and symptomatic transmission.

To create these stochastically reconstructed transmission chains, for each case $i$ and each possible infector $j$, we determine the range of viable days $j$ could have infected $i$, $t_j$.

Then, according to the infectivity profile, the probability that $j$ infected $i$ on the $k^{th}$ day of $t_j$ is $P_{inf}^{A_j k}(t_{jk})$, where $A \in \{Asy, Pre, Sym\}$ is the symptomatic class of $j$ on day $k$.

Then the probability that $j$ infected $i$ during $t_j$ is

$$P_j = \sum_{k=1}^{t_j} P_{inf}^{A_j k}(t_{jk})$$

for each $j \in J$, the set of all possible infectors of $i$. We can then reweight $\hat{P}_j = P_j / \sum_{j \in J} P_j$, then make a simple random draw from $\hat{P}$ to determine which $j$ infected $i$ in our reconstructed chain and make a simple random draw from $P_{inf}^{A_j k}(t_j)$ to determine which day $j$ infected $i$. Note, the date of incident of the infector case will affect the viable range of dates the infectee case could have been infected on, so all possible infector cases must have their incident date drawn before drawing an infectee’s incident date.

Estimation of the basic reproductive number ($R_0$)

For each realization of our stochastically reconstructed transmission chains, we estimate posterior distributions of $R_0^{Asy}, R_0^{Sym}, R_0^{Tot}$, the basic reproductive numbers of asymptomatic cases, symptomatic cases, and of all cases. We modeled transmission with a Poisson process in time, so that the instantaneous rate at which a case became infected or had symptom onset at time $t - s$ is $R_0^A P_{inf}^{A}(t - s), A \in \{Asy, Pre, Sym, Tot\}$. Hence $Y_t^A$, the number of people infected at time $t$ from cases of symptomatic class $A$, is Poisson
with mean \( R^A_0 \sum_{s=0}^{t-1} Y^A_s P^A_{\text{Inf}}(t - s) \). The instantaneous reproduction number for each symptomatic class \( A \) is assumed to be constant at \( R^A_0 \) throughout the study. The likelihood of \( Y^A_t \) given \( R^A_0 \) and \( Y^A_0, \ldots, Y^A_{t-1} \) is

\[
P(Y^A_t|R^A_0, Y^A_0, \ldots, Y^A_{t-1}, P^A_{\text{Inf}}) = \frac{\left( R^A_0 \sum_{s=0}^{t-1} Y^A_s P^A_{\text{Inf}}(t - s) \right)^{Y^A_t} \exp \left( -R^A_0 \sum_{s=0}^{t-1} Y^A_s P^A_{\text{Inf}}(t - s) \right)}{Y^A_t!}
\]

For a given outbreak, each case \( i \) has a start time \( s^A_i \), and end time \( e^A_i \) during which it was of symptomatic class \( A \) (note, if a case was never of a particular class, \( s^A_i = e^A_i = 0 \), but for cases that transition between classes, in this case from presymptomatic to symptomatic, then \( s^A_i = t^\text{onset}_i - 1 \), and \( s^A_i = t^\text{onset}_i, e^A_i = t^\text{quarantine}_i \). We define the infectivity mass at time \( t \) of symptomatic class \( A \).

\[
M^A_t = \sum_{s=0}^{t} Y^A_s P^A_{\text{Inf}}(t - s).
\]

Similarly, we define the infectivity mass of case \( i \) for symptomatic class \( A \) as

\[
M^A_i = \sum_{j=s^A_i}^{e^A_i} P^A_{\text{Inf}}(j).
\]

For an outbreak that occurs during \([t_{\text{start}}, t_{\text{end}}]\) and has cases \( i = 1, \ldots, n \) for class \( A \),

\[
\sum_{t=t_{\text{start}}}^{t_{\text{end}}} M^A_t = \sum_{i=1}^{n} M^A_i.
\]

We can simplify the likelihood of \( Y^A_t \) given \( R^A_0 \) and \( Y^A_0, \ldots, Y^A_{t-1} \) as

\[
P(Y^A_t|R^A_0, Y^A_0, \ldots, Y^A_{t-1}, P^A_{\text{Inf}}) = \frac{(R^A_0 M^A_t)^{Y^A_t} \exp(-R^A_0 M^A_t)}{Y^A_t!}.
\]

Since \( R^A_0 \) is considered to be constant over the course of the outbreak, the likelihood of transmission over \([t - \tau + 1, t]\) by cases of symptomatic class \( A \), \( Y^A_{t-\tau+1}, \ldots, Y^A_t \), given \( R^A_0 \) and \( Y^A_0, \ldots, Y^A_{t-\tau} \) is
\[ P(Y_t^{A}, \ldots, Y_t^{A} \mid Y_0^{A}, \ldots, Y_{t-\tau}^{A}, R_0^{A}, P_{inf}^{A}) = \prod_{s=t-\tau+1}^{t} \frac{(R_0^{A}M_t^{A})^{Y_s^{A}} \exp(-R_0^{A}M_t^{A})}{Y_s^{A}!} \]

If we give \( R_0^{A} \) a prior distribution of \( \Gamma(shape = a, rate = b) \), then under a Bayesian Framework, the joint posterior distribution of \( R_0^{A} \) is

\[ P(Y_t^{A}, \ldots, Y_t^{A} \mid Y_0^{A}, \ldots, Y_{t-\tau}^{A}, R_0^{A}, P_{inf}^{A}) = \prod_{s=t-\tau+1}^{t} \frac{(R_0^{A}M_s^{A})^{Y_s^{A}} \exp(-R_0^{A}M_s^{A})}{Y_s^{A}!} \]

\[ \times \frac{(R_0^{A})^{a-1} \exp(-R_0^{A}b + \sum_{s=t-\tau+1}^{t} M_s^{A})}{\Gamma(a)Y_t^{A}!} \]

Hence, the posterior distribution of \( R_0^{A} \) is \( \Gamma(a + \sum_{s=t-\tau+1}^{t} Y_s^{A}, b + \sum_{s=t-\tau+1}^{t} M_s^{A}) \).

Thus, over the entirety of the outbreak, with cases \( 1, \ldots, n \), and \( Y_t^{A} \) being the number of cases infected by \( i \) while of symptomatic class \( A \),

\[ R_0^{A} \sim \Gamma\left(shape = a + \sum_{i=1}^{n} Y_t^{A}, rate = b + \sum_{i=1}^{n} M_t^{A}\right). \]

We estimated \( R_0^{Asy}, R_0^{Pre}, R_0^{Sym} \), and \( R_0^{Tot} \) separately. The prior distribution for each \( R_0^{A} \) is Gamma with mean 2.5 and standard deviation 2 (corresponds with \( shape = 1.25 \) and \( rate = 0.625 \)), expressing large uncertainty about the basic reproductive number in this context (Zhang et.al. 2020). For each chain, we calculate the realized total number of people infected from case \( i \) while of symptomatic class \( A \), \( y_t^{A} \), and the infectivity mass of each case for each symptomatic class \( A \), \( m_t^{A} \). The posterior distribution results in

\[ R_0^{A} \sim \Gamma\left(shape = 1.25 + \sum_{i=1}^{n} y_t^{A}, rate = 0.625 + \sum_{i=1}^{n} m_t^{A}\right), \]

for \( n \) total cases in the outbreak. The posterior distribution can be computed directly for asymptomatic, presymptomatic, symptomatic, and all cases separately.
All analyses were done using R software (version 4.2.0). All quantities were estimated in a Bayesian framework. Point estimates and the corresponding 95% credible intervals (CrI) were obtained from the posterior distributions.

**Predicting the Epidemic Outcomes had Mitigation not been Applied**

To better understand the potential impact this outbreak could have had in the absence of social distancing, quarantining, and other Covid-19 transmission mitigating measures, we implement an Susceptible-Infected-Recovered (SIR) model that outputs estimates of the prevalence, total number of cases, the incidence, number of new daily cases, and cumulative deaths. This SIR model is based on $R_0^{Tot}$. In an SIR model, there is a fixed sized population, and all members of the population are either Susceptible (not infected but can be infected), Infected (have the infection and can spread it to Susceptible people), or Recovered (had the infection but no longer have it and cannot spread it, nor can they get the infection again). We opt for an SIR model since the data collected only allows us to understand transmissions that occurred, but not contacts that didn’t result in new cases. As a result, we lack information on exposure that can be credibly used in an SEIR model for this outbreak.

At any time $t$, suppose there are $S$ susceptible, $I$ infectious, and $R$ recovered individuals. Denote the mean infectious time as $1/\gamma$ and recovery rate $\gamma$, for $\gamma > 0$, and the constant disease transmission rate $\beta$, so that $\beta I$ is the rate of infection proportional to the number of infected, and $\beta SI$ is the number of newly infected at each time $t$, proportional to the number of susceptible individuals. The SIR model is specified by the ODEs:

$$
\frac{dS}{dt} = -\beta SI, \quad \frac{dI}{dt} = \beta SI - \gamma I, \quad \frac{dR}{dt} = \gamma I,
$$

along with initial conditions $S(0) = S_0$, $I(0) > 0$, $I(0) << S(0)$, and $R(0) = 0$.

This model results in a reproductive number of $R_0 = \beta S_0 / \gamma$. We assume a constant $R_0$ over the course of the outbreak, and we already have estimated a posterior $R_0^{Tot}$ distribution. By taking many random draws from this distribution, we use the SIR model
to estimate the prevalence, incidence, and cumulative deaths over time, along with best and worst case scenarios, over the course of the outbreak.