Metabiota Risk Report No. 3: February 25, 2020

Monitoring and risk analytics for the 2019 novel coronavirus (COVID-19) epidemic

Executive Summary

Since January 2, 2020, Metabiota has been closely monitoring and analyzing the COVID-19 epidemic. New cases continue to be reported, indicating the epidemic is not yet contained. Metabiota is tracking and analyzing this outbreak in several ways, including through a true-to-source digital surveillance system and a global disease spread model.

Metabiota created a near-term forecasting model of disease spread, which incorporates the current known characteristics of the virus. Our forecasting model estimates, for March 3rd, a median 127,000 cumulative cases within a 95% confidence window of 81,500 and 295,000 cases. We predict that the countries most at-risk of reporting new cases for the upcoming week include China, Japan, Italy, Iran, South Korea, Thailand, United States, Taiwan, Australia, and the Philippines.

Digital Surveillance

Metabiota collects and structures data from a wide variety of reporting sources to produce a dataset with the finest spatiotemporal resolution data. Currently, Metabiota is monitoring incidence across 37 countries using 39 public data sources for this outbreak, ranging from the Hong Kong Centre for Health Protection to the World Health Organization. An aggregated view of the data is available publicly through Metabiota's Epidemic Tracker.

Box 1: Situation Report as of Tuesday, February 25

- There are currently 80,308 confirmed cases with 2,707 deaths
- Cases have been confirmed in 33 territories[1] in China
- Cases have been confirmed in 36 additional countries
- There have been 53 confirmed cases in the US: including 8 in California, 2 in Illinois, and 1 each in Arizona, Massachusetts, Texas, Washington state, and Wisconsin. 36 cases were identified among passengers repatriated from the *Diamond Princess* cruise ship quarantined in Yokohama, Japan.

High-resolution data, available for non-commercial use, has been publicly released to Metabiota's Epidemic Data repository.

[1] Including provinces, special administrative regions, and municipalities

Box 2: COVID-19: Epidemiologic Assessment

- The number of new cases reported in China has been declining.
- There has been a spike in cases in South Korea, Italy and Iran which suggest uncontrolled spread in these countries as well as the potential for undetected cases.
    - South Korea has declared this outbreak as a “severe” national crisis.
    - While Iran has only reported 64 confirmed cases they have reported 12 confirmed deaths. This high number of deaths compared to cases suggests that there is likely a large number of undetected or unconfirmed cases in this country.
- The role of subclinical, asymptomatic, and presymptomatic cases in transmission is still unclear and of great concern.

Forecasting

Metabiota developed a forecasting model to predict the near-term trajectory of the novel coronavirus (COVID-19) epidemic. Our modeling framework consists of a stochastic metapopulation compartment model coupled with human mobility networks. We simulated 200 COVID-19 outbreaks encompassing a range of plausible parameter values (Table 1) with starting conditions similar to those reported on January 26th.

Table 1. Coronavirus Model Parameters

Parameter	Mean value	Definition	Source(s)
Transmissibility (R_0) *	2.6	Disease transmission rate	[2] [3] [4]
Incubation Period (days)	5.2	Median time from exposure to symptom onset	[2]
Infectious Period (days)	9.2	The number of days an individual is infectious	[5]

*Superspreading is incorporated as dispersion of R_0 around the mean input value

[2] Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med. 2020; doi: 10.1056/NEJMoa2001316

[3] Read, JM, JRE Bridgen, D Cummings, A Ho, and CP Jewell. Novel Coronavirus 2019-NCoV: Early Estimation of Epidemiological Parameters and Epidemic Predictions. *MedRxiv*, January 1, 2020, 2020.01.23.20018549.

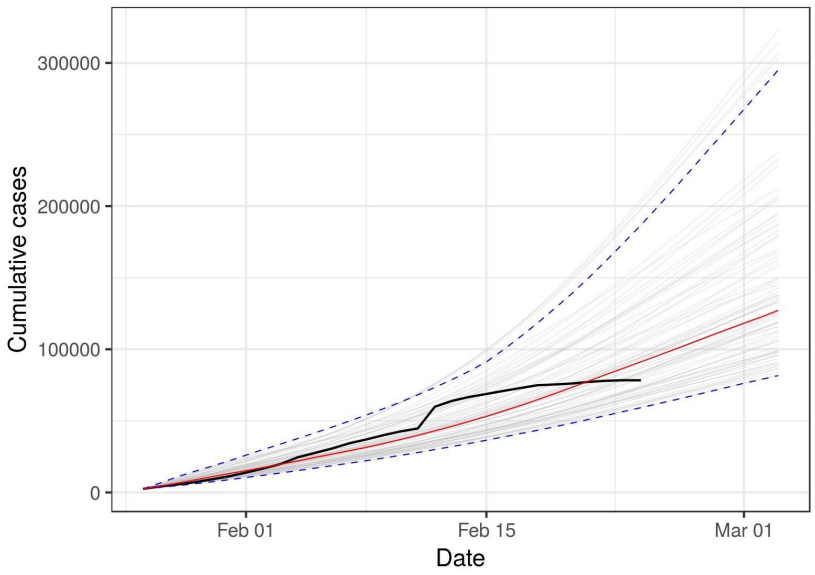
[4] Imai N, ACori, I Dorigatti, M Baguelin, CA Donnelly, S Riley, and N Ferguson. “Report 3: Transmissibility of 2019-NCoV,” https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-2019-nCoV-transmissibility.pdf

[5] Riou J and CL Althaus, 2020. Pattern of early human-to-human transmission of Wuhan 2019-nCoV bioRxiv 2020.2001.2023.917351

We released forecasting results in a previous report generated by the same model and using similar methods. Our previous forecast predicted a median of 85,000 cases would be reported by February 17th, which was a difference of ~16% from the 73,000 cases reported on that day. The true case count was well within our confidence window of 58,000 and 180,000 cases. Since our previous forecast was released, the transmission rate within China appears to have slowed significantly[6], prompting us to incorporate transmission-slowing interventions in our model. See Appendix 1 for further methodological details.

With the newly implemented interventions, our forecast model predicts reported cumulative cases will continue to grow by approximately 50% between Feb 24th and March 3rd if the transmission rate is not further decreased (Figure 1). Our model currently predicts a median 127,000 cumulative cases within a 95% confidence window of 81,500 and 295,000 cases on March 3rd.

Figure 1: Short Term Forecasted Trajectory of the COVID-19 outbreak

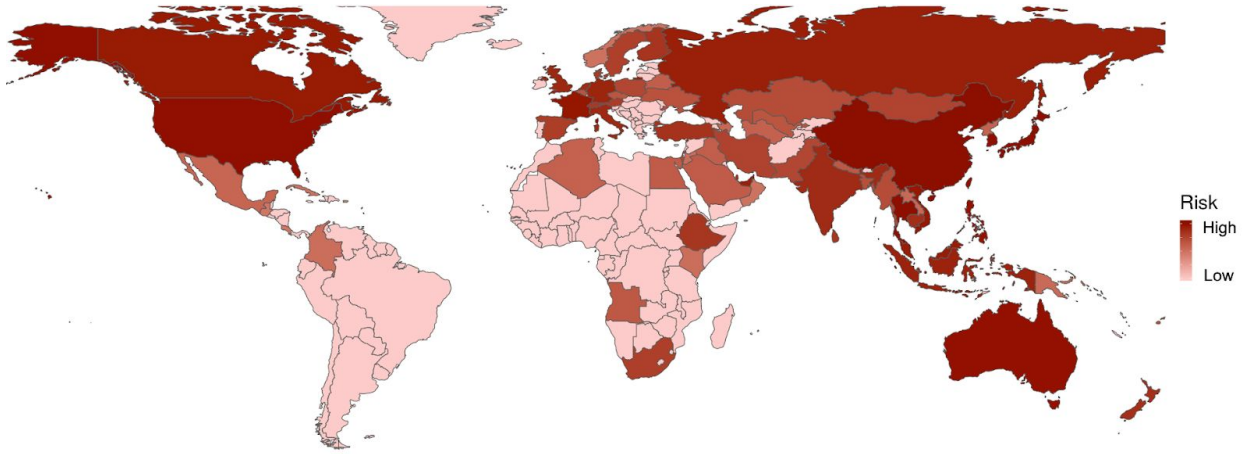


The black line shows reported cumulative infections up to March 3rd. Faded gray lines show individual model runs after filtering. The solid red and dashed blue lines show median cumulative case counts predicted by our forecast model and the 95% confidence window, respectively. The median and confidence windows were determined after filtering.

[6] Sun, Haoxuan, Yumou Qiu, Han Yan, Yaxuan Huang, Yuru Zhu, and Song Xi Chen. "Tracking and Predicting COVID-19 Epidemic in China Mainland." Preprint. Epidemiology, February 20, 2020. https://doi.org/10.1101/2020.02.17.20024257.

Our forecasting model predicts geographic spread of COVID-19 using air travel and commuting networks to simulate population mobility. Although the COVID-19 event is continuing to evolve, our latest forecasting model estimates the following countries to be at high risk for the number of new cases in the next week: China, Japan, Italy, Iran, Thailand, South Korea, United States, Taiwan, Australia, and the Philippines (Figure 2). This risk is inferred by ranking countries according to the number of new cases predicted within the forecasting window.

Figure 2: Rank of Number of New Cases by Country, February 25-March 3, 2020



The relative risk of new COVID-19 cases within the forecast window of February 25th to March 3rd at the country-level. Countries shaded in darker red shades show a higher risk of reporting new COVID-19 cases than those shown in lighter shades.

Analysis: Predictors of COVID-19 case counts

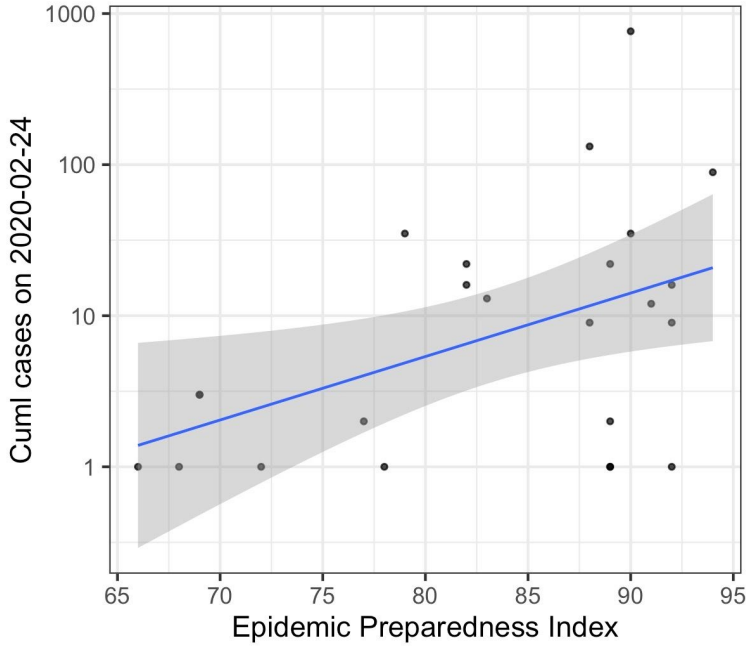
We explore the relationship between country preparedness and the number of reported cases in countries outside of China using Metabiota's Epidemic Preparedness Index (EPI), which summarizes the capacity at the country-level to detect and respond to infectious disease events based on holistic set of institutional, financial and infrastructural capabilities.[7] In a previous analysis, countries ranked higher by the EPI were found to have significantly higher vaccination rates during the 2009 H1N1 influenza pandemic and were found to report infectious disease outbreaks more quickly.[7]

We performed a linear regression on the total number of cases (as of February 24, 2020) reported at the country-level against the country-specific EPI for that country (Figure 3). Since the amount of air travel between China and the affected country will influence this relationship, it

[7] Oppenheim B, Gallivan M, Madhav NK, et al. Assessing global preparedness for the next pandemic: development and application of an Epidemic Preparedness Index. *BMJ Glob Health*. 2019;4(1). doi:10.1136/bmjgh-2018-001157

was controlled for in the model. We found a positive relationship between the total number of cases occurring as of February 24th and the country-specific EPI score (*DF* = 22, *P* = 0.02).

Figure 3: Cumulative COVID-19 Cases vs. Epidemic Preparedness Index



The relationship between cumulative COVID-19 cases (y-axis; shown here in log10 scale) and Metabiota's Epidemic Preparedness Index (EPI; x-axis). Points represent case counts and EPI scores at the country-level. The blue line shows the linear fit between these variables; the shaded gray region shows the 95% confidence interval.

These results reveal that better-prepared countries are reporting more COVID-19 cases, even when controlling for the effect of air travel. These findings may be an indication that less prepared countries are underreporting cases due to a lack of capability to detect and report these cases.

For Further Information:

Ben Oppenheim, PhD
Senior Director of Product
Email: boppenheim@metabiota.com

Jaclyn Guerrero, MPH
Senior Epidemiologist/Analyst
Email: jguerrero@metabiota.com

Appendix 1: Forecasting methodology

Each simulation began with outbreak conditions of what was known on January 26, 2020. Given the parameter combinations we are assuming based on current epidemiological knowledge (Table 1), we found that an estimate of 2% of cases identified was appropriate to mimic the observed epidemic growth. Therefore, we initiated our model with the number of cumulative cases reported on January 26th in the infectious compartment of our model, and an additional 50 times this amount in the exposed-and-unreported compartment. This assumption could inflate our model predictions within the forecast window if underreporting is not this extensive, but other estimates of underreporting have been in line with these assumptions.[8] Given the current overwhelmed state of public health services in the epicenters of the outbreak, we believe the assumption is within reason.

As the outbreak progresses, our model assumes new infections are also subject to underreporting, and that only 20.84% of cases are being reported in China, where most of the cases occur. Metabiota previously built a model that predicts pathogen and country specific reporting rates based on the case fatality ratio of the pathogen and the corresponding country's Metabiota's Epidemic Preparedness Index value. In the newest version of our forecast model, we reduced R_0 in each simulation by 60% beginning on February 14th to account for the observed slowing in transmission.

To refine our forecast based on model simulations, we excluded all model runs that did not result in at least 75% of cases reported as of February 24, 2020. Once these runs were excluded, we calculated the interquartile range (IQR) of remaining case counts occurring on February 24th, and excluded runs predicting more than 3Q + 1.5xIQR (209,876) cases.

[8] Sun, Haoxuan, Yumou Qiu, Han Yan, Yaxuan Huang, Yuru Zhu, and Song Xi Chen. "Tracking and Predicting COVID-19 Epidemic in China Mainland." Preprint. Epidemiology, February 20, 2020. https://doi.org/10.1101/2020.02.17.20024257.