Caution: Please note that there are large uncertainties in our model projections due to unknown disease transmission dynamics (model misspecification), changing behavior and policies, delay in reporting, and under-reporting. In particular, the data our projections are based on reflect situations ~2 weeks ago due to time lags from interventions implemented to transmission events (a couple days to weeks), from infection to symptom onset (~2-6 days), from symptom onset to seeking treatment (~2-7 days), from seeking treatment to getting tested and then reported in the surveillance system (~2-7 days). In addition, how the epidemic would unfold also depend largely on behavior changes over time.

Main Model Settings, Assumptions, Inputs, and Outputs

Authors: Wan Yang,[1] Sasikiran Kandula,[2] Jeffrey Shaman[2]

[1]Department of Epidemiology, [2]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University

1. Model Form

1.1. City-level SEIR model: Susceptible-Exposed-Infectious-Removed (SEIR) model accounting for reporting delay for case diagnosis, and delay from infection to hospitalization, ICU admission, and death for estimating the numbers of hospitalization, ICU, and death by week, respectively. This model was used for projections generated from March 16 – April 3, 2020.

1.2. SEIR model by age group: For this model, the same SEIR construct is used; however, the model parameters account for differences by age (e.g. reporting rate and disease severity) and the model is trained using age-specific incidence data for eight age groups: <1, 1-4, 5-14, 15-24, 25-44, 45-64, 65-74, and 75+ years, separately. This system is used for projections generated from 4/3/2020 onwards.

1.3. Network SEIR model by age group: This system accounts for spatial heterogeneity using a network model capturing connections among 42 United Hospital Fund (UHF) neighborhoods in NYC. This system also incorporates mobility data provided by Safegraph.com to gauge changes in transmission within each UHF neighborhood and between pairs of neighborhoods. The model is trained using age-specific, UHF-specific incidence data.

Note that UHF locations are based on residential address and may not match with hospital locations and catchment area.

2. Data

Confirmed cases of COVID-19 in New York City (from Week 10, i.e. March 1-7 of 2020 to the most recent week, date labeled in the Results folder), provided by NYC DOHMH.

3. Model Training and Assumptions

Bayesian inference approach in which the DOHMH data are used to partially constrain the model parameters and state variables prior to making a projection. The form is similar to that used for influenza forecasting; however, here the data are very limited (3 weeks) so the model is less well constrained. Initial prior ranges are set as: transmission rate (β): [0.5, 1]; latency period (T_{ei}): [2, 5] days; infectious period (T_{ir}): [2, 5] days; mean reporting delay (i.e., from viral shedding to being diagnosed; Td.mean): [3, 9] days; standard deviation of reporting delay (Td.sd): [1, 3] days; and reporting rate (i.e., the proportion of infections that are diagnosed; α): [5, 80]%. These parameters are estimated based on the weekly confirmed case data.

In addition, for the delay from infection to hospitalization, ICU, and death, we used reported time from symptom onset of SARS-CoV-2 to the corresponding event (Yang et al. 2020; Zhou et al. 2020; Wang et al. 2020). To compute the numbers of different health outcomes from the model estimated total infections, we used the following probability ranges: 3-9% for hospitalization (severe and critical cases); 1-3.6% for ICU and 0.5-1.5% for mortality. These probabilities are based on reported numbers among diagnosed cases in NYC (information from NYC DOHMH), China (China CDC, 2020) and other countries and assuming a 20-30% ascertainment rate (Li et al., 2020). To compute the healthcare demands for each week, we used reported retention times in hospitals and ICU (Zhou et al. 2020) for corresponding estimates. See further details below.

4. Model Scenarios

Seasonality: There are 4 endemic coronaviruses infecting humans (OC43, 229E, NL63, HKU1). These viruses typically cause mild cold-like symptoms and exhibit a pronounced seasonality with peak incidence in January-February and very little incidence in summer. The cause of this seasonality is unknown, but its presence has led to speculation that SARS-CoV-2, the virus causing COVID-19, may wane during summer months in New York City. Consequently, we used the seasonality of OC43, which is well observed and a betacoronavirus, like SARS-CoV2, to estimate a seasonal reduction of transmissibility for SARS-CoV2 during summertime. We then generated projections from the 2 forms for all scenarios: 1) With seasonal changes to virus transmissibility; and 2) Without seasonality.

We generated the following projections, each with 10 model runs to provide a distribution of possible outcomes:

No Control (i.e. Worst Case) Scenario: For these projections, the model posterior (i.e. an ensemble of model simulations with parameters and state variables as estimated following training with weekly confirmed case data) *estimated with data from Week 10 (March 1-7) of 2020 (an earlier week with minimal interventions)* was integrated 8 weeks into the future to create a reference, no control, "worst case" scenario.

As Is (i.e. Status Quo) Scenario: For these projections, the model posterior *estimated using data from Week 10 through the most recent week,* was integrated 8 weeks into the future to create a reference, As Is, "status quo" scenario. While very preliminary, these projections provide a rough assessment of effectiveness of current interventions, compared to the "no control" scenario.

Control Scenarios: Five control scenarios, using the model posterior as initial conditions and adjustment of model parameters (relative to the As Is scenario estimates) to represent different levels of interventions:

1. Moderate (25%) reduction in contact rate (via social distancing)
2. Moderate (25%) reduction in contact rate (via social distancing) and moderate (25%) reduction in infectious period (via case isolation/self-quarantine/treatment, etc.)
3. Large (50%) reduction in contact rate (via social distancing) and no reduction in infectious period
4. Large (50%) reduction in contact rate (via social distancing) and moderate (25%) reduction in infectious period (via case isolation/self-quarantine/treatment, etc.)
5. Large (50%) reduction in contact rate (via social distancing) and large (50%) reduction in infectious period (via case isolation/self-quarantine/treatment, etc.)

Rebound Scenarios: In light of the recent uptick in mobility and potential rebound in transmission rate, from 5/1/2020 onwards, we also simulate two rebound scenarios, using the model posterior as initial conditions and adjustment of model parameters (relative to the As Is scenario estimates) to represent different levels of transmission rebound:

1. Moderate (25%) increase in contact rate (relaxed social distancing)
2. Large (50%) increase in contact rate (relaxed social distancing)

Note there is no particular specification of how reductions in contact rates or spread are achieved. In a model of this form different reduction options (e.g. isolation vs. quarantine) are not represented explicitly; rather, they are effected by adjusting the estimated (posterior) contact rate and infectious period within the model, relative to estimates for the most recent week (the As Is scenario).

Model Output

We use the model to estimate new weekly numbers of total infections, reported/observed infections, hospitalizations, patients in ICU, and deaths. For the latter three health outcomes we accounted for delay from infection to corresponding event as described above.

- **New total infections** are directly estimated by the model without a delay and are an unobserved quantity that includes subclinical/undiagnosed infections.
- **New reported/observed infections** include a reporting delay; the reporting rate estimated at the last day of model training was used for the entire forecast period; thus, these numbers may largely differ from observations depending on changes in testing.
- **New total hospitalizations**: we assume 15-30% of *reported* infections are hospitalized and time from symptom onset to hospitalization of 5 days (mean; SD = 3 days; estimates from China: China CDC, 2020, adjusted for NYC).
- **New ICU admissions:** we assume 4-8% of *reported* infections are critical and enter ICU and time from symptom onset to ICU admission of 11 days (mean; SD=5 days; estimates from China: China CDC, 2020).
- **New non-ICU hospitalizations**: computed as the difference between new total hospitalizations and ICU admissions.
- **New patients needing ventilators**: we assume 60-100% of patients admitted to the ICU need ventilators (per data from NYC DOHMH).
- **New deaths:** we assume an infection mortality risk of 0.05-1.5% based on data in NYC and elsewhere in the world and time from symptom onset to death of 10 days (mean; SD = 4 days; per data from NYC)

To support logistics and planning, we also use the model to estimate the numbers of hospital beds and ICU beds needed each week under each scenario:

- **Demand for hospital beds:** projections are based on new hospitalizations each day and length of stay in hospital (Mean=24 days; SD=5.2 days, per data from the US and elsewhere).
- **Demand for ICU beds:** projections are based on new ICU admissions each day and length of stay in the ICU (Mean=21 days; SD=5.9 days, per data from the US and elsewhere).
- **Demand for ventilators:** projections are based on new ICU admissions each day and length of use (Mean=12 days; SD=3 days, per estimates from the CDC).

We also report the estimated attack rate as the number of New Yorkers (total population size: 8,398,744 as of 2018) infected in the next 8 weeks.

Results – See tables and figures.

References

China CDC. Vital Surveillances: The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020, 2020, 2(8): 113-122 (http://weekly.chinacdc.cn:80/en/zcustom/volume/1/2020)

Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, Shaman J. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (COVID-19) *Science* 16 Mar 2020: eabb3221. doi: 10.1126/science.abb3221

World Health Organization, Coronavirus disease (COVID-2019) situation reports, 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/.

Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. Lancet Respir Med. 2020;S2213-2600(20)30079-5; Epub ahead of print.

Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X, Guan L. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet. 2020 Mar 11.

Wang et al., Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China - JAMA, February 7, 2020