Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Estimating the infection-fatality risk of SARS-CoV-2 in New York City during the spring 2020 pandemic wave: a model-based analysis

Wan Yang, Sasikiran Kandula, Mary Huynh, Sharon K Greene, Gretchen Van Wye, Wenhui Li, Hiu Tai Chan, Emily McGibbon, Alice Yeung, Don Olson, Anne Fine, Jeffrey Shaman

Summary

Background As the COVID-19 pandemic continues to unfold, the infection-fatality risk (ie, risk of death among all infected individuals including those with asymptomatic and mild infections) is crucial for gauging the burden of death due to COVID-19 in the coming months or years. Here, we estimate the infection-fatality risk of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in New York City, NY, USA, the first epidemic centre in the USA, where the infection-fatality risk remains unclear.

Methods In this model-based analysis, we developed a meta-population network model-inference system to estimate the underlying SARS-CoV-2 infection rate in New York City during the 2020 spring pandemic wave using available case, mortality, and mobility data. Based on these estimates, we further estimated the infection-fatality risk for all ages overall and for five age groups (<25, 25–44, 45–64, 65–74, and ≥75 years) separately, during the period March 1 to June 6, 2020 (ie, before the city began a phased reopening).

Findings During the period March 1 to June 6, 2020, 205 639 people had a laboratory-confirmed infection with SARS-CoV-2 and 21 447 confirmed and probable COVID-19-related deaths occurred among residents of New York City. We estimated an overall infection-fatality risk of 1·39% (95% credible interval 1·04–1·77) in New York City. Our estimated infection-fatality risk for the two oldest age groups (65–74 and ≥75 years) was much higher than the younger age groups, with a cumulative estimated infection-fatality risk of 0·116% (0·072–0·148) for those aged 25–44 years and 0·939% (0·729–1·148) for those aged 45–64 years versus 4·87% (3·37–6·89) for those aged 65–74 years and 14·2% (10·2–18·1) for those aged 75 years and older. In particular, weekly infection-fatality risk was estimated to be as high as 6·72% (5·52–8·01) for those aged 65–74 years and 19·1% (14·7–21·9) for those aged 75 years and older.

Interpretation Our results are based on more complete ascertainment of COVID-19-related deaths in New York City than other places and thus probably reflect the true higher burden of death due to COVID-19 than that previously reported elsewhere. Given the high infection-fatality risk of SARS-CoV-2, governments must account for and closely monitor the infection rate and population health outcomes and enact prompt public health responses accordingly as the COVID-19 pandemic unfolds.

Funding National Institute of Allergy and Infectious Diseases, National Science Foundation Rapid Response Research Program, and New York City Department of Health and Mental Hygiene.

Copyright © 2020 Elsevier Ltd. All rights reserved.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in late 2019 in China and has subsequently spread to more than 200 other countries. As of Oct 5, 2020, over 35 million cases of COVID-19 and over 1 million COVID-19-related deaths have been reported worldwide.1 As the pandemic continues to unfold and populations in many places worldwide largely remain susceptible, understanding the severity, and, in particular, the infection-fatality risk of the virus is crucial for gauging the full impact of COVID-19 in the coming months or years. However, estimating the infection-fatality risk of SARS-CoV-2 is challenging due to the large number of undocumented infections, fluctuating infection detection rates, and inconsistent reporting of fatalities. Furthermore, the infection-fatality risk of SARS-CoV-2 could vary by location, given differences in demographics, health-care systems, and social structures (eg, intergenerational households are the norm in some societies whereas older adults commonly reside and congregate in long-term care and adult care facilities in other societies). Most estimates of infection-fatality risk for SARS-CoV-2 to date have come from data recorded in China, the Diamond Princess cruise ship, and France.2–5 As yet, the infection-fatality risk in the USA—the country currently reporting the largest number of cases—remains unclear.

New York City, NY, USA, reported its first case of COVID-19 on March 1, 2020, in a traveller, and quickly became the epicentre of the pandemic in the country.6
Intense community transmission occurred during the following 3 months before a series of public health interventions brought the pandemic under control. In particular, public schools in the city were closed on March 16, 2020, and a citywide stay-at-home order was imposed on all non-essential workers starting the week of March 22, 2020. The city was able to reopen industries according to a phased schedule starting the week of June 7, 2020. By June 6, 2020, before the city’s phased reopening, over 200,000 people had been diagnosed with laboratory-confirmed SARS-CoV-2 infection and more than 20,000 COVID-19-related deaths had been reported in the city. Since the beginning of the pandemic, the New York City Department of Health and Mental Hygiene (DOHMH) and the Mailman School of Public Health at Columbia University (New York City, NY) have been collaborating to generate real-time model projections in support of the city’s pandemic response. Our model-inference system uses a meta-population network model to simulate SARS-CoV-2 transmission in the city’s 42 United Hospital Fund neighbourhoods. The model is run in conjunction with the ensemble adjustment Kalman filter and fit simultaneously to case and mortality data for each of the 42 neighbourhoods while accounting for under-detection, delay between infection, case reporting, and death, and changing interventions (eg, physical distancing). In this analysis, we applied this network model-inference system to estimate the infection-fatality risk for five age groups (ie, <25, 25–44, 45–64, 65–74, and ≥75 years) and all ages overall, from March 1 to June 6, 2020. In the process, we also estimated infection detection rates (ie, the fraction of infections documented as confirmed cases) and the cumulative infection rate by June 6, 2020.

Methods
Study design and data
In this model-based analysis, we aggregated laboratory confirmed SARS-CoV-2 infections reported to the New York City DOHMH by week of diagnosis and age group (<1, 1–4, 5–14, 15–24, 25–44, 45–64, 65–74, and ≥75 years) for each of the 42 United Hospital Fund neighbourhoods in New York City, according to the patient’s residential address at time of reporting. We aggregated mortality data for confirmed and probable COVID-19-associated deaths from deaths registered and analysed by the New York City DOHMH. Confirmed COVID-19-associated deaths were defined as those occurring in people with laboratory-confirmed SARS-CoV-2 infection; and probable COVID-19 deaths were defined as those with COVID-19, SARS-CoV-2, or a similar term listed on the death certificate as an immediate, underlying, or contributing cause of death but did not have laboratory confirmation of SARS-CoV-2 infection. Due to privacy concerns, the New York City DOHMH aggregated mortality data to five coarser age groups (<18, 18–44, 45–64, 65–74, and ≥75 years) for each neighbourhood by week of death. To match with the age grouping for case data, we used the citywide fraction of deaths occurring in each of the five finer age groups (ie, <1, 1–4, 5–14, 15–24, 25–44) to apportion deaths in the younger than 18 and 18–44 year age categories. For this study, case and mortality data were both retrieved on Aug 7, 2020.
We used mobility data to model changes in the rate of SARS-CoV-2 transmission due to public health interventions implemented during the pandemic. We sourced these data from SafeGraph\(^1,1\) and they contained counts of visitors to locations in each zip code based on mobile device locations. The released data were anonymised and aggregated in weekly intervals (with weeks defined as Sunday to Saturday). We spatially aggregated these data to the neighbourhood level.

This study was classified as public health surveillance and was exempt from ethical review and informed consent by the Institutional Review Boards of both Columbia University and New York City DOHMH.

**Meta-population network transmission model**

Our meta-population network model simulated intra-neighbourhood and inter-neighbourhood transmission of SARS-CoV-2 and assumed susceptible-exposed-infectious-removed dynamics, per the following equation system:

\[
\frac{dS}{dt} = -\sum_{j=1}^{42} \frac{b(t)\beta_{yo}(t)c(t)I_j}{N_j} E_i \\
\frac{dE}{dt} = S \sum_{j=1}^{42} \frac{b(t)\beta_{yo}(t)c(t)I_j}{N_j} - E_i / Z(t) \\
\frac{dI}{dt} = E_i / Z(t) - I_i / D(t) \\
\frac{dR}{dt} = I_i / D(t)
\]

where \(S\) is the number of susceptible individuals, \(E\) is the number of exposed (but not yet infectious) individuals, \(I\) is the number of infectious individuals, \(R\) is the number of removed individuals (either recovered or deceased), and \(N\) is the total population from a given age group in neighbourhood \(i\). Due to model complexity and a scarcity of information for parameterising interactions among age groups, we modelled each age group separately (ie, combining all sources of infection to each age group); as such, system (1) describes the spatial transmission across neighbourhoods with no interactions among age groups. \(t\) is time, and we make this time dependence explicit for the parameters to indicate that they were estimated for each week and could vary over time due to disease seasonality or public health interventions, or both; state variables (\(S\), \(E\), \(I\), and \(R\)) are inherently vary with other parameters (appendix pp 4–6). To calculate the connectivity among the neighbourhoods, we first divided the inter-neighbourhood mobility by the local mobility, which gave a relative measure of connectivity (eg, if two neighbourhoods are highly connected with lots of individuals travelling between them, inter-neighbourhood mobility would be closer to 1, but if they were not highly connected then inter-neighbourhood mobility would be much lower than 1); we then scaled these relative rates by a multiplicative factor \(m_i\), which was also estimated along with other parameters (appendix pp 4–6).

**Observational model**

To account for delays in diagnosis and detection, we included a lag of time from infectious to detection (ie, an infection being diagnosed as a case), drawn from a gamma distribution with a mean of \(T_m\) and an SD of \(T_m\) days. To account for under-detection, we included an infection detection rate (\(r\)—ie, the fraction of infections (including subclinical or asymptomatic infections) reported as cases. To calculate the model-simulated number of infections per day (including those from the previous weeks) by the infection detection rate, and further distributed these simulated cases in time per the distribution of time from infectious to detection. We then aggregated the daily lagged, simulated cases to weekly totals for model inference. Similarly, to calculate the model-simulated deaths per week and account for delays in time to death, we multiplied the simulated number of infections by the infection-fatality risk and then distributed these simulated deaths in time per the distribution of time from infectious-to-death lag, and aggregated these daily numbers to weekly totals. For each week, we estimated the infection detection rate (\(r\), the mean \((T_m)\) and SD \((T_m)\) of time from infectious to detection, and the infection-fatality risk on the basis of weekly case and mortality data. The distribution of time from diagnosis to death was based on observations of 15,686 confirmed COVID-19-related deaths in New York City as of May 17, 2020, from New York City DOHMH (gamma distribution with a mean of 9·36 days [SD 9·76]; appendix pp 4–6).

**Parameter estimation**

To estimate model parameters (\(\beta_{yo}, Z, D, m_i, m, T_m, T_{m0}, r\), infection-fatality risk and \(b_i\) for \(i=1,...,42\)) and state
variables ($S$, $E$, and $I$, for $i=1,...,42$) for each week, we ran the meta-population network model stochastically with a daily timestep in conjunction with the ensemble adjustment Kalman filter and fit to weekly case and mortality data from the week starting March 1, 2020, to the week ending June 6, 2020. The ensemble adjustment Kalman filter uses an ensemble of model realisations (n=500 here), each with initial set of parameters and variables randomly drawn from a prior range (appendix pp 4–6). After model initialisation, the model ensemble was integrated forwards in time for a week to calculate the model-simulated number of cases and deaths for that week; these prior estimates were then combined with the observed cases and deaths for the same week to calculate the posterior distribution of each model parameter or variable for that week per Bayes’ theorem.9 Notably, the ensemble adjustment Kalman filter also models the observational errors (eg, due to imperfect sensitivity and specificity of SARS-CoV-2 RT-PCR tests for case diagnosis) over time by specifying an error structure and using this information when calculating the posterior distribution.9

We did this parameter estimation process separately for each of the eight age groups (ie, <1, up to ≥75). To include transmission from other age groups, we used measured intra-group and inter-group contacts from the POLYMOD study22 to calculate the total number of contacts made to each age group and adjusted the prior range of the transmission rate ($\beta_{i,m}$) during the first week of the pandemic for each age group accordingly. We calculated the posterior estimate on the basis of case and mortality data for each age group, which included all sources of infection. Thus, the estimated transmission rate for each age group included all sources of transmission.

To account for stochasticity in model initiation, we ran the parameter estimation process independently ten times. We combined results for each age group from these ten runs (each with 500 realisations). To combine estimates of the infection-detection rate and infection-fatality risk for those younger than 25 years or all ages overall, we weighted the age-group specific estimates (median and credible interval [CrI]) by the fraction of estimated infections from each related age group.

### Model validation

As a model validation, we compared our estimates of cumulative infection rates to three independent serology datasets measuring the seroprevalence of antibodies to SARS-CoV-2 during the pandemic wave in New York City. Details of the serology data and matching by timing of measurement, age group, and location are in the appendix (pp 1, 8, 10).

### Role of the funding source

The funders of the study had a role in the data collection and no role in study design, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

### Results

Between March 1 and June 6, 2020, 205639 people had been diagnosed with laboratory-confirmed SARS-CoV-2 infection and 21447 COVID-19-related deaths had been reported in New York City (table). The epidemic timing (eg, peak of confirmed cases and mortality) varied substantially by age group and neighbourhood (appendix p 10). We were able to use our model-inference system to recreate the case and mortality time series for each age group and all ages overall (figure 1). For most age groups, confirmed cases peaked during the week of March 29, 2020, and the mortality rate peaked about 1 week later than the case rate, due to the time-lag from severe infection to death.

However, there was substantial under-detection of infections, variations by age group, and fluctuations of infection-detection rates over time, in part due to changing testing criteria.11 The estimated infection-detection rate for all ages overall started at a low level of 2·2% (95% CrI 0·3–4·5) during the week of March 1, 2020, and increased to 17·4% (11·3–26·1) during the week of March 15 (figure 2). However, due to shortages in testing and personal protective equipment, testing was restricted to severely ill patients in early April before it became more widely available in May.12 Consistently, the estimated infection detection rate dropped to approximately 13% in early-mid April, then gradually increased to approximately 19% in early May and stayed at similar levels through the week of May 31, 2020 (figure 2). The estimated infection detection rate was highest for the two oldest age groups and was substantially lower for younger age groups (figure 2). During the week of May 31, 2020, before the city began its phased reopening, we estimated that 29·8% (21·7–42·3) of infections among those aged 65–74 years and 36·0% (28·4–47·9) of infections among those aged 75 years and older were detected; by comparison, only

### Table: Summary estimates of cases of COVID-19 and COVID-19-related deaths in New York City, NY, USA for the period March 1 to June 6, 2020, by age group

<table>
<thead>
<tr>
<th>Age group</th>
<th>Confirmed cases</th>
<th>Confirmed and probable deaths</th>
<th>Estimated cumulative infection rate</th>
<th>Estimated infection-fatality risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;25 years</td>
<td>18332</td>
<td>45</td>
<td>8·56% (5·66–17·5)</td>
<td>0·00972% (0·00405–0·0154)</td>
</tr>
<tr>
<td>25–44 years</td>
<td>64743</td>
<td>734</td>
<td>22·6% (16·6–31·2)</td>
<td>0·116% (0·0729–0·148)</td>
</tr>
<tr>
<td>45–64 years</td>
<td>74798</td>
<td>4722</td>
<td>22·7% (18·0–29·2)</td>
<td>0·239% (0·179–0·322)</td>
</tr>
<tr>
<td>65–74 years</td>
<td>25460</td>
<td>5181</td>
<td>15·0% (11·4–21·6)</td>
<td>4·87% (3·37–6·89)</td>
</tr>
<tr>
<td>≥75 years</td>
<td>24296</td>
<td>10755</td>
<td>12·8% (9·92–18·6)</td>
<td>14·2% (10·2–18·1)</td>
</tr>
<tr>
<td>Overall</td>
<td>205639</td>
<td>21447</td>
<td>17·2% (12·9–25·1)</td>
<td>3·59% (1·04–7·77)</td>
</tr>
</tbody>
</table>

Data are n, median cumulative infection rate with 95% CrI in parentheses, and median estimated infection-fatality risk with 95% CrI in parentheses. Data are given to three significant figures. Cases and deaths were reported by the New York City Department of Health and Mental Hygiene between March 1 and June 6, 2020. Cumulative infection rate is for all those infected by June 6, 2020. And infection-fatality risk is averaged over March 22 to June 6, 2020, with estimates for March 1–21 excluded because estimates were less accurate for these earliest weeks when zero or few deaths were reported. CrI=credible interval.
Figure 1: Model fit for confirmed number of cases of COVID-19 (A, C, E, G, I, K) and model estimate of number of COVID-19 related deaths (B, D, F, H, J, L), by age group and overall. Boxes and whiskers show the median, 50% CrI, and 95% CrI. Red dots show the observed confirmed case rates (A, C, E, G, I, K) and observed mortality rates (B, D, F, H, J, L). CrI=credible interval.
11.0% (7.2–17.6) of infections among those younger than 25 years and 16.8% (11.8–23.1) of infections among those aged 25–44 years were detected.

After accounting for the infection-detection rate, the epidemic peak for new infections occurred 1–2 weeks before the peak in confirmed cases, during the week of March 22, 2020, for those younger than 65 years and in all ages combined (figure 2). This peak was coincident with the timing of public health interventions in New York City (ie, public schools closing and the citywide stay-at-home order was imposed). Tallied over the entire study period, the estimated overall cumulative infection rate was 17.2% (95% CrI 12.9–25.1) by June 6, 2020 (table). However, estimated cumulative infection rates varied substantially across age groups and neighbourhoods in the city (figure 3). Specifically, the highest cumulative infection rates were in people aged 25–44 years and 45–64 years, and those aged 65–74 years and 75 years and older had the second highest cumulative infection rates, and those younger than 25 years had the lowest cumulative infection rate (table). Spatially, among the five boroughs in New York City, estimated cumulative infection rates were highest in neighbourhoods in the Bronx and lowest in neighbourhoods in Manhattan (figure 3).

Our model estimates of cumulative infection rates have large uncertainties. To assess the accuracy of our model, we compared our model estimates with three datasets of serological data. After accounting for changing infection detection rates and excluding the first 3 weeks of the study period (ie, March 1–21, 2020, in which none or few deaths were reported, hence making the model estimates less accurate), we estimated that the overall infection-fatality risk, including both confirmed and probable deaths, was 1.39% (95% CrI 1.04–1.77) during March 22 to June 6, 2020 (table). If only confirmed COVID-19-related deaths were included, given that 16,924 (78.9%) deaths were confirmed COVID-19-related, the model-inference system estimated higher infection rates (95% CrI 12.9–25.1) by June 6, 2020 (table).
to be due to SARS-CoV-2 infection, the overall infection-fatality risk would be around 1·10% (ie, 1·39% × 0·789).

Examining estimates by age group, the estimated infection-fatality risk was lowest in young age groups, increasing substantially with age (table; figure 4). These estimates were similar to infection-fatality risks reported for China for corresponding age groups. However, the estimated infection-fatality risk for the two oldest age groups was much higher than the younger age groups and about twice as high as the rates reported for these age groups in China.3,4 Additionally, the estimated infection-fatality risk fluctuated substantially over time for the two oldest age groups. For those aged 65–74 years, the estimated infection-fatality risk was 6·72% (95% CrI 5·52–8·01) during the week of April 5, 2020, but decreased to 4·20% (2·22–7·01) during the week of May 31, 2020 (figure 4). For those aged 75 years and older, estimated infection-fatality risk was 19·11% (14·70–21·92) during the week of April 5, 2020, but decreased to 10·38% (6·17–14·96) during the week of May 31, 2020 (figure 4).

Discussion
In light of the large uncertainties in infection-fatality risks for SARS-CoV-2 due to under-detection of infections, we used a model-inference system, developed to support the pandemic response in New York City, to estimate local infection-fatality risks. During the 2020 spring pandemic (March 1–June 6, 2020), New York City recorded the largest number of COVID-19 cases and related deaths in the USA. Despite public health efforts to slow the pandemic (eg, via physical distancing), and to increase health-care capacity, 21447 people died due to COVID-19 in the city in the short span of 3 months. Based on this large number of deaths, the estimated overall infection-fatality risk in New York City was 1·39% if both confirmed and probable COVID-19-related deaths were included or 1·10% if only confirmed COVID-19-related deaths were included. Both estimates were higher than previously reported elsewhere (eg, about 0·7% in both China1 and France5). Importantly, New York City has nosologists who rapidly review all death certificates and record deaths into a unified electronic reporting system (the average death certification time was 22·2 h and 95% of deaths were certified within 3·1 days during the pandemic wave; unpublished data, New York City DOHMH, Huynh M). This mortality surveillance infrastructure and enhanced nosology thus allow more rapid and complete death reporting in New York City than other places in the world.
As such, our estimates here probably reflect the underlying fatality risk of SARS-CoV-2 infection more accurately than do those in previous studies. Furthermore, because the public health infrastructure and health-care systems in New York City are probably stronger than in many other places, the higher infection-fatality risk estimated here suggests that the fatality risk from SARS-CoV-2 might be higher in the USA and some other countries than has been previously reported. Notably, despite the large surge in cases and admissions to hospital, through quick expansion of health-care systems, most hospitals in New York City were able to meet demand for patient care during the pandemic. Because COVID-19 continues to pose pandemic risk in many places worldwide, governments must account for and closely monitor the infection rate and health outcomes, including admissions to hospital and mortality, and take prompt public health responses accordingly.

Although the infection-fatality risk we estimated here was similar to that previously reported elsewhere for younger age groups, we found that the infection-fatality risk for individuals aged 65 years and older in New York City were about twice as high as in reports from other locations. These higher infection-fatality risk estimates might be in part due to differences in population characteristics, in particular, the prevalence of underlying medical conditions such as diabetes, chronic lung disease, and cardiovascular disease. Regardless, our estimated weekly infection-fatality risk was as high as 6.7% for those aged 65–74 years and 19.1% for those aged 75 years and older. These dire estimates highlight the increased risk of COVID-19-related mortality in older populations and the importance of infection prevention in congregate settings. Thus, early detection and adherence to infection control guidance in long-term and adult care facilities should be a priority for COVID-19 response as the pandemic continues to unfold.

Over 5000 COVID-19-related deaths occurred among adults aged 25–64 years during the study period. Despite this large number of deaths, estimated cumulative infection rates in these age groups were only around 20% by the week of May 31, 2020, much lower than the 50–70% herd immunity needed to prevent large epidemics of COVID-19 (assuming the basic reproductive number for SARS-CoV-2 is around 2–3.5 and infection confers long-lasting immunity). By July, 2020, many places where lockdown-like measures were lifted saw increases in the number of cases of COVID-19 among young adults. These continuous infections could ignite new epidemics of COVID-19 and lead to further devastating effects in older populations and in younger adults (in particular, those aged 45–64 years) given the remaining high
population susceptibility in many places and transmission across age groups. As such, young adults must strictly adhere to physical distancing and preventive measures (eg, mask wearing) in places with continuous transmission, despite their relatively low infection-fatality risk.

In this study, we incorporated multiple data sources, including age-grouped, spatially resolved case and mortality data and mobility data, to calibrate our model-inference system. Notably, the timing of the peak of the COVID-19 pandemic varied substantially among New York City neighbourhoods. For instance, peak mortality rates occurred up to 8 weeks apart among the 42 neighbourhoods. Fitting the model-inference system simultaneously to these diverse case and mortality time series thus enabled improved constraint of key model parameters (eg, infection detection rate and infection-fatality risk).

We note there remain large uncertainties in our model estimates. A full assessment of COVID-19 severity will require comprehensive serological surveys of the population by age group and neighbourhood due to the large heterogeneity of infection rates across populations and space. Additionally, we only included deaths that were laboratory confirmed as related to SARS-CoV-2 infection or explicitly coded as related to COVID-19. Previous studies have reported that excess deaths in New York City during about the same period were over 24000,26–27 which are more than the 21447 COVID-19-related deaths included in this study. Furthermore, studies have reported severe sequelae of COVID-19 in children—ie, multisystem inflammatory syndrome in children.28 Thus, monitoring health outcomes in younger age groups after infection is important as the pandemic unfolds, despite the low infection-fatality risk in these age groups noted to date.

Contributors
WY conceived the study with input from JS, SK, MH, SKG, GVW, and AF. MH, GVW, WL, and HTC oversaw the collection of and provided the human endemic coronaviruses data used for estimating the seasonality of the pandemic. EM contributed to data management of confirmed cases and confirmed and probable COVID-19-associated deaths. AF oversaw all data collection processes at the DOHMH. SK compiled the human mobility data from SafeGraph and wrote the initial version of data processing scripts. SKG and AF provided critical input on parameter estimation. DO provided input on the mortality pattern. WY did all modelling analyses and wrote the first draft of the manuscript. All authors critically reviewed and revised the manuscript.

Declaration of interests
JS and Columbia University disclose partial ownership of SK Analytics. JS discloses consulting for BNI and Merck. All other authors declare no competing interests.

Acknowledgments
This study was supported by the National Institute of Allergy and Infectious Diseases (AI145883), the National Science Foundation Rapid Response Research Program (RAPID: 2027369), and the New York City DOHMH. We thank the New York City DOHMH Bureau of Vital Statistics team and staff members in the New York City DOHMH Incident Command System Surveillance and Epidemiology Section for data management. We thank Miranda S Moore at New York City DOHMH, who wrote and maintained the code to generate weekly extracts of COVID-19 case data for provision to Columbia University under a data use agreement. We thank Jasmine Shaff at New York City DOHMH and Helen Alesbury and Kathryn Klein at the Office of Chief Medical Examiner of the City of New York for initiating and coordinating discussions on modelling COVID-19 mortality in New York City. We also thank Columbia University Mailman School of Public Health for high performance computing and SafeGraph for providing the mobility data used in this study.

References


