Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China

Joseph T Wu (joewu@hku.hk)
WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong

Kathy Leung
WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong

Mary Bushman
Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard TH Chan School of Public Health

Nishant Kishore
Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard TH Chan School of Public Health

Rene Niehus
Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard TH Chan School of Public Health

Pablo M de Salazar
Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard TH Chan School of Public Health

Benjamin J Cowling
WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong

Marc Lipsitch
Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard TH Chan School of Public Health

Gabriel M Leung
WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong

---

Research Article

Keywords: COVID-19, estimating clinical severity, symptomatic case-fatality risk (sCFR)

Posted Date: March 13th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-17453/v1

Version of Record: A version of this preprint was published on March 19th, 2020. See the published version at https://doi.org/10.1038/s41591-020-0822-7.

Abstract

As of February 13, 2020, there have been 59,863 laboratory-confirmed cases of COVID-19 infections in mainland China, including 1,367 deaths. A key public health priority during the emergence of a novel pathogen is estimating clinical severity. Here we estimated the symptomatic case-fatality risk (sCFR; the probability of dying from the infection after developing symptoms) of COVID-19 in Wuhan using public and published information. We estimated that sCFR was 0.5% (0.1%-1.3%), 0.5% (0.2%-1.1%) and 2.7% (1.5%-4.7%) for those aged 15-44, 45-64 and >64 years. The overall sCFR among those aged ≥15 years was 1.4% (0.8%-2.0%).

Authors Joseph T Wu and Kathy Leung contributed equally to this work

Case Definition In Clinical Severity Measures

Infection fatality risk (IFR): IFR defines a case as a person who has shown evidence of infection, either by clinical detection of the pathogen or by seroconversion or other immune response. Such cases may or may not be symptomatic, though asymptomatic ones may go undetected.

Symptomatic case fatality risk (sCFR): sCFR defines a case as someone who is infected and shows certain symptoms.

Hospitalization fatality risk (HFR): HFR defines a case as someone who is infected and needs to be hospitalized.

Introduction

On 9 January 2020, a novel coronavirus, COVID–19, was officially identified as the cause of an outbreak of viral pneumonia in Wuhan, China. As of Feb 13, 2020 (1000 h Hong Kong time), there have been 59,863 laboratory-confirmed cases of COVID–19 infections in mainland China, including 1,367 deaths.

One of the most critical clinical and public health questions during the emergence of a completely novel pathogen, especially one that could cause a global pandemic, pertains to the spectrum of illness presentation or severity profile. For the patient and clinician, it affects triage and diagnostic decision-making, especially in settings without ready access to laboratory testing or when surge capacity has been exceeded. It also influences therapeutic choice and prognostic expectations. For managers of health services, it is important for rapid forward planning in terms of procurement of supplies, readiness of human resources to staff beds at different intensities of care and generally ensuring sustainability of the health system through the peak and duration of the epidemic.

At the population level, determining the shape and size of the “clinical iceberg”, both above and below the observable threshold, is key to understanding transmission dynamics and interpreting epidemic

trajectories. Both inform the development and evaluation of public health strategies that need to be traded off against economic, social and personal freedom costs.

For the general public, the overriding concern about uncertainty can breed fear, even panic. There is arguably no greater cause of such anxiety than the relative probability of death and disability caused by infection. For a completely novel pathogen, especially one with a high (say >2) basic reproductive number, assuming homogeneous mixing and mass action dynamics, the majority of the population will be infected eventually unless drastic public health interventions are applied over prolonged periods and/or vaccines become available sufficiently quickly. Even under more realistic assumptions about mixing, at least a quarter to a half of the population will very likely become infected absent drastic control measures or a vaccine. Therefore, the number of severe outcomes or deaths in the population is most strongly dependent on how ill an infected person is likely to become, and this question should be the focus of attention.

Therefore, we have extended our previously published transmission dynamics model [1], updated with real time input data and enriched with additional new data sources, to infer a preliminary set of clinical severity estimates that could guide clinical and public health decision making as the epidemic continues to spread and evolve globally.

Methods

Estimation of true case numbers – necessary to determine the severity per case – is challenging in the setting of an overwhelmed health care system that cannot ascertain cases effectively. Therefore as in our prior work [1],our approach has been to use a range of data sources to build a picture of the full number of cases and deaths by age group. Because the health care structure has been overwhelmed in Wuhan, and milder cases were unlikely to have been tested, we used the prevalence of infection in travelers (both on the commercial flights prior to 19 January and on the charter flights from 29 January to 4 February) to estimate the true prevalence of infection in Wuhan, and used the Wuhan case numbers from the first 425 cases solely to estimate the growth rate of the epidemic (assuming these were a constant proportion of true cases). See Figure 1 for illustration.

We relied on publicly available and recently published information. Specifically, we inferred the epidemiologic parameters listed in Table 1 by fitting an age-structured transmission model (an extension of the model from our previous study) to the following data:

1. The epidemic curve of confirmed cases of COVID-19 in Wuhan with no epidemiologic links to Huanan Seafood Wholesale Market (which was presumed to be the index zoonotic source of the COVID-19 epidemic) between 1 December 2019 and 3 January 2020 (Figure 1; Table S1) [2].
2. The number of cases exported from Wuhan to cities outside mainland China via air travel on each day between 25 December 2019 and 19 January 2020 (Figure 1; Table S2) [1].

3. The number of expatriates who returned to their countries from Wuhan on charter flights between 29 January and 4 February 2020 and the proportion of passengers on each flight who had laboratory confirmed infection with COVID-19 upon arrival (Figure 1 and Table S3).
4. The age distribution of 425 confirmed cases of COVID-19 from the early stages of the epidemic in Wuhan [2] (Table S4).
5. The age distribution of 41 fatal cases of COVID-19 infection during the early stages of the epidemic in Wuhan [3-5] (Table S5).
6. The cumulative number of deaths among laboratory-confirmed cases of COVID-19 infection in Wuhan up to 11 February 2020 [6] (Table S6).
7. The time between onset and death or the time between admission and death for 41 death cases of COVID-19 in Wuhan [3-5] (Table S7).
8. The time between the onset dates (i.e. serial intervals) of 40 infector-infectee pairs (Table S8).

We made the following assumptions in the model:

1. The incubation period was gamma distributed with mean 6.5 days and standard deviation 2.6 days [7].
2. The infection-symptomatic probability (P_{sym}; the proportion of infections that progressed to become symptomatic cases) was the same for all age groups. We assumed P_{sym} = 0.5 in the baseline scenario and 0.25 or 0.75 in alternative scenarios.
3. The sensitivity of detecting symptomatic cases exported from mainland China (P_{det}; the proportion of exported symptomatic cases that were detected upon border entry) was the same for all countries between 25 December 2019 and 19 January 2020. We assumed $P_{det} = 30\%$ in the baseline scenario and 60% in alternative scenarios.

4. Inbound and outbound mobility in Wuhan had been reduced by around 90% for mainland Chinese cities (https://qianxi.baidu.com/) and 99% for international cities since Wuhan was quarantined on 23 January 2020.
5. Children under the age of 15 years were not susceptible to infection, based on the absence of this age group in the age distribution of cases in Wuhan [2]. We therefore omitted this age group from the transmission model and stratified the remaining population into three age groups: 15-44, 45-64 and >64 years. The dependence of susceptibility to infection (α_i) and symptomatic case-fatality risk ($sCFR_i$) on age were inferred (Table 1).
6. We assumed that the diagnostic test for the charter flight passengers was 100% sensitive and 100% specific for detecting COVID-19 infections.

7. Recent phylogenetic analyses suggested that the most recent common ancestor of the sequenced COVID-19 genomes emerged between 23 October and 16 December 2019. (http://virological.org/t/clock-and-tmrca-based-on-27-genomes/347 accessed 12 Feb 2020). As such, we assumed that the epidemic in Wuhan was seeded by a single zoonotic event which generated z_0 infections on 15 November 2019. We inferred the value of z_0 (Table 1).
8. We assumed that public health interventions in Wuhan reduced local transmissibility by φ_0. We inferred the value of φ_0 (Table 1).
9. Given that the epidemic curve in Wuhan was weeks ahead of that in other mainland Chinese cities, we ignored the effect of case importation at Wuhan.

These assumptions were reflected in following SEIR model where $S_i(t)$, $E_i(t)$, $I_i(t)$ and $R_i(t)$ were the number of susceptible, latent, infectious and recovered individuals in age group i at time t:

$$\frac{dS_i(t)}{dt} = -\alpha_i S_i(t)\pi(t) + \frac{N_i(t)}{N(t)} L_{W,inbound}(t) - \left(\frac{L_{W,I}(t)}{N(t)} + \frac{L_{W,C}(t)}{N(t)}\right) S_i(t)$$

$$\frac{dE_i(t)}{dt} = \alpha_i S_i(t)\pi(t) - \frac{E_i(t)}{D_E} - \left(\frac{L_{W,I}(t)}{N(t)} + \frac{L_{W,C}(t)}{N(t)}\right) E_i(t)$$

$$\frac{dI_i(t)}{dt} = \frac{E_i(t)}{D_E} - \frac{I_i(t)}{D_I} - \left(\frac{L_{W,I}(t)}{N(t)} + \frac{L_{W,C}(t)}{N(t)}\right) I_i(t)$$

$$\frac{dR_i(t)}{dt} = \frac{I_i(t)}{D_I} - \left(\frac{L_{W,I}(t)}{N(t)} + \frac{L_{W,C}(t)}{N(t)}\right) R_i(t)$$

$$N_i(t) = S_i(t) + E_i(t) + I_i(t) + R_i(t), \qquad N(t) = \sum_{i=1}^{3} N_i(t)$$

$$\pi(t) = \beta(1 - \varphi(t)) \sum_{j=1}^{3} \frac{I_j(t)}{N(t)}$$

$$\varphi(t) = \begin{cases} 0 & \text{before 23 January 2020} \\ \varphi_0 & \text{otherwise} \end{cases}$$

The incidence rate of infection, onset and death for age group *i* at time *t* were calculated as follows:

$$A_{i,infection}(t) = \alpha_i S_i(t) \pi(t)$$

$$A_{i,onset}(t) = P_{sym} \int_0^t A_{i,infection}(u)\, f_{incubation}(t-u) du$$

$$A_{i,death}(t) = sCFR_i \int_0^t A_{i,onset}(u)\, f_{onset-to-death}(t-u) du$$

where

$f_{incubation}$ and $f_{onset-to-death}$

were the probability density function of the incubation period and the time between onset and death for those who died from the infection. We inferred the parameters listed in Table 1 assuming that the remaining parameters were fixed at the values shown in Table 2.

Results

In the baseline scenario

$(P_{scr} = 30\%, P_{sym} = 0.5)$,

we estimated that sCFR was 0.5% (0.1%-1.3%), 0.5% (0.2%-1.1%) and 2.7% (1.5%-4.7%) for those aged 15-44, 45-64 and >64 years. The overall sCFR among those aged ≥15 years was 1.4% (0.8%-2.0%). Compared to those aged 15-44 years, those 45-65 and >64 years were 3.1 (2.4-4.0) and 6.4 (4.8-8.4) times more susceptible to infection.

Figure 2 summarizes our estimates for the key epidemiologic parameters of COVID-19 in Wuhan. The basic reproductive number was 2.42 (2.11-2.88) and the mean serial interval was 7.7 (6.1-9.2) days. The mean time from onset to death was 19 (16-24) days with standard deviation 10 (7-14) days. The mean latent and infectious period was 4.5 (3.0-6.4) and 3.2 (1.9-5.2) days, respectively.

We estimated that the epidemic doubling time (the time it takes for daily incidence to double) was 5.5 (4.4-6.9) days before Wuhan was quarantined and public health interventions implemented within Wuhan reduced transmissibility by 53% (21%-82%). We estimated that only 0.3% (0.2%-0.6%) of symptomatic cases that occurred between 1 December 2019 and 3 January 2020 were laboratory confirmed.

Our estimates of sCFRs were unaffected by the infection-symptomatic probability (P_{sym}) but increased linearly with detection sensitivity (P_{det}), i.e. sCRFs would be doubled if P_{det} was 60% instead of 30%. The same trends applied to our estimates of the confirmation rate between 1 December 2019 and 3 January 2020. If P_{det} and P_{sym} were higher, our estimates of intervention effectiveness would be smaller (e.g. only around 25% (2%-43%) if P_{scr} was 60% and P_{sym} was 50%). The estimates of all the remaining parameters were mostly unaffected by P_{det} and P_{sym}.

Discussion

There is clear and considerable age dependency in infection and outcome risks, by several fold in each case. For simplicity, we only considered the adult population given empirical reports of only very rare observations of cases below age 15 years. Given that we have parameterized the model using death rates inferred from projected case numbers (from traveler data) and observed death numbers in Wuhan, the precise fatality risk estimates may not generalize to those outside of that original epicenter especially during subsequent phases of the epidemic. The experience gained from managing those initial patients and the increasing availability of newer, and potentially better, treatment modalities to more patients would presumably lead to fewer deaths, all else being equal. Public health control measures widely imposed since the Wuhan alert have also kept case numbers down elsewhere such that their health systems are not nearly as overwhelmed beyond surge capacity, thus again perhaps leading to better outcomes. Indeed, to date, the death-to-case ratio in Wuhan has been consistently much higher than that

among all the other mainland Chinese cities (Figure S1). Therefore, as and when data from outside Wuhan become available over the next few weeks, our model should be reparameterised accordingly to give a fuller account of the clinical iceberg.

Considering the risk estimates in context, Table 3 compares infection/case/hospitalization fatality risks for pandemic influenza in 1918 and 2009, SARS and MERS. SARS causes moderate to severe disease requiring hospitalization, thus the infection fatality risk and case fatality risk are essentially the same as the hospitalisation fatality risk. The hospitalization fatality risk for MERS is well documented although the shape and depth of the clinical iceberg remains less well defined. In contrast, the majority of COVID–19 infections do not cause severe disease requiring hospitalization, and the symptomatic case fatality risk would be substantially lower than the hospitalisation fatality risk. However, despite a lower symptomatic case fatality risk, COVID–19 is likely to infect many more people and ultimately cause many more deaths than SARS and MERS. Compared with the 1918 and 2009 influenza pandemics, our estimates are intermediate but substantially higher than 2009, which was generally regarded as a low severity pandemic. Like 2009 we find symptomatic case-fatality risk is highest in the oldest age group; unlike any previously reported pandemic or seasonal influenza, we find that infection risk is also increasing with age group, though this may be in part due to preferential ascertainment of older and thus more severe cases. One largely unknown factor at present is the number of asymptomatic, undiagnosed infections. These do not enter our estimates of the *symptomatic* CFR, but if such asymptomatic or clinically very mild cases existed and were not detected, the infection fatality risk would be lower than the sCFR. Further clarifying this requires new data sources not yet available, including serologic studies.

Our inferences were based on a variety of sources, and have a number of caveats highlighted below, but considering the totality of the findings nevertheless indicate that COVID–19 transmission is difficult to control. With a reproductive number in the range of 2–2.5, we might expect at least half of the population to be infected even with aggressive use of community mitigation measures. Perhaps the most important target of mitigation measures would be to “flatten out” the epidemic curve, reducing the peak demand on healthcare services and buying time for better treatment pathways to be developed. In due course, vaccines may also be available to protect against infection or severe disease. While our estimates of symptomatic case fatality risk are concerning, these could be reduced if effective antivirals were identified and widely adopted for treatment of severe cases. Timely data from clinical trials of remdesivir, lopinavir/ritonavir, and other potential chemotherapies, as well as supportive care modalities, would be extremely informative.

Several important caveats bear mention, as follows. First and most importantly, our modelled estimates have necessarily relied on numerous strong assumptions, given the paucity of definitive data elements such as serosurveys, serial viral shedding studies, robust ascertainment of sufficient transmission chains, incomplete testing of travelers and returnees from Wuhan, all of which needed to be underpinned by systematic unbiased sampling of the underlying population and by important age- and other sub-groups.

Our estimates of symptomatic case-fatality risk are inevitably affected by underascertainment of cases and deaths of COVID–19. On one hand, overstretched and overwhelmed health care surge capacity in Wuhan could result in symptomatic case-fatality risks that are higher than they would be in a less stressed health care setting. We have accounted for limited sensitivity of detection of cases among travelers, and our sensitivity analyses show that lower detection among travelers implies larger populations infected in Wuhan and correspondingly lower severity. On the other hand, the numerator of the number of deaths could also have been undercounted although much less likely so by comparison to enumerating the denominator, for the same surge capacity reason or due to imperfect test sensitivity especially during the first month of the outbreak [8]. If deaths in Wuhan were under-ascertained, this would bias our severity estimates downward.

Another caveat concerns one of our key inputs of the infection prevalence among expatriates airlifted out of Wuhan. Their point prevalence might well be lower than that among local residents, because of a generally more advantaged socioeconomic background, thus it would be a lower bound of the cross-sectional disease prevalence. If this were the case, then we would have overestimated the reduction in transmissibility conferred by public health interventions in Wuhan and overestimated severity.

Based on only publicly available data, there is necessarily substantial uncertainty in our estimates of the effectiveness of intra-Wuhan public health interventions in reducing transmissibility. Calculating the instantaneous reproductive number from a set of line lists that are updated daily would be the most reliable method for detecting changes in transmissibility associated with interventions.

There has been refinement of case definitions at both the national and provincial levels, such as excluding RT-PCR test-positive asymptomatics (or perhaps in fact very mildly symptomatics) from being labelled an officially “confirmed” case [9] or including test-naïve clinically diagnosed cases with clear epidemiologic links as “confirmed” [10]. While these should not affect our estimation given our data sources from the earlier phase of the epidemic, such changes in the reporting criteria may influence the interpretation of future data. Finally, given that Wuhan is no longer the only, albeit the first, location with sustained local spread, it would be important to assess and take into account the experience from elsewhere both domestically in mainland China and overseas. These secondary epicenters, having learnt from the early phase of the Wuhan epidemic, might have had a systematically different epidemiology and response that could impact the parameters estimated here.

Tables

Table 1. Epidemiologic parameters fitted in the model

Parameters	Description
R_0	Basic reproductive number, which was linearly related to the transmission rate β
z_0	The number of infections generated by the seeding zoonotic event
φ_0	Relative reduction in transmissibility after vs. before lockdown of Wuhan on 23 January 2020
α_2	The relative susceptibility for those aged 44-64 years compared to those aged 15-44 years
α_3	The relative susceptibility for those aged >64 years compared to those aged 15-44 years
$sCFR_1$	The symptomatic case-fatality risk for those aged 15-44 years
$sCFR_2$	The symptomatic case -fatality risk for those aged 45-64 years
$sCFR_3$	The symptomatic case -fatality risk for those aged >64 years
(μ_D, σ_D)	Mean and standard deviation for the time between onset and death
D_E	Mean latent period
D_I	Mean infectious period
ε	The proportion of human-to-human infections that were confirmed between 1 December 2019 and 4 January 2020

Table 2. Assumed constants in the model

Data	Description	Value
$L_{W,inbound}(t)$	Daily number of persons entering Wuhan (assumed all susceptible)	International: 3,546; decreased by 99% after Wuhan was quarantined on 23 January 2020 Mainland China: 487,310 and 810,500 before and during *chunyun*; decreased by 90% after Wuhan was quarantined on 23 January 2020
$L_{W,I}(t)$	Daily number of persons leaving Wuhan to destinations outside mainland China	3,633; decreased by 99% after Wuhan was quarantined on 23 January 2020
$L_{W,C}(t)$	Daily number of persons leaving Wuhan to destinations in mainland China	502,013 and 717,226 before and during *chunyun*; decreased by 90% after Wuhan was quarantined on 23 January 2020
$N_1(0)$	Number of Wuhan residents aged 15-44	See Table S9
$N_2(0)$	Number of Wuhan residents aged 45-64	See Table S9
$N_3(0)$	Number of Wuhan residents aged 65+	See Table S9
P_{sym}	Infection-symptomatic probability	0.5 in the baseline scenario; 0.25 or 0.75 in alternative scenarios
P_{det}	Sensitivity of detecting exported symptomatic cases in countries outside China	30% in the baseline scenario; 60% in alternative scenarios

Table 3. Severity estimates of SARS (2002-3), MERS (2014-), 1918 influenza pandemic (1918-20) and 2009 influenza pandemic (2009-10)

Parameter	SARS*	MERS†	1918 influenza pandemic	2009 influenza pandemic
Infection fatality risk (IFR) (risk of death among all infections)	Worldwide (WHO)** [1] 9.6% (774/8096) Mainland China** [11] 6.4% (343/5327) Hong Kong*** [12,13] Overall: 17.2% (302/1755) <60 yrs: 13.2% (9.8-16.8) ≥60 yrs: 43.3% (35.2-52.4) Taiwan*** [14] Overall: 27.6% (180/664) <60 yrs: 15.3% (72/470) ≥60 yrs: 48.6% (88/181)	--	Worldwide [15,16] 2.5% Copenhagen [17] 1.7%	Hong Kong [18] <60 yrs: <0.1% ≥60 yrs: 1.1% (0.2-4.7)
Symptomatic case fatality risk (sCFR) (risk of death among symptomatic infections)		--	--	United Kingdom [19] Overall: 26 (11-66) per 100,000 5-14 yrs: 11 (3-36) per 100,000 ≥65 yrs: 980 (300-3200) per 100,000
Hospitalization fatality risk (HFR) (risk of death among infections that require hospitalization for medical reasons, not only for case isolation)		Worldwide (WHO)*** 34.4% (858/2494) Saudi Arabia*** [20] 40.7% (726/1783) South Korea*** [20] 20.4% (38/186)		North America [21] Overall: 2.6% (1.6, 3.9) ≤19 yrs: 0.8% (0.5, 1.1) 20-64 yrs: 5.4% (3.5, 7.5) ≥65 yrs: 10.7% (5.3, 17.6)

* IFR=CFR=HFR given virtually every infected person required hospitalization and seroprevalence amongst close contacts and in general community approximated zero

** Among probable cases

*** Among laboratory-confirmed cases

† There are more infections “undetected” in MERS compared with SARS. The seroprevalence amongst individuals exposed to camels was estimated as 6.2% in Arabian Peninsula [22].

Declarations

Acknowledgement

We thank Di Liu, Miky Wong and Chi-Kin Lam from School of Public Health, The University of Hong Kong for technical support. This research was supported by a commissioned grant from the Health and Medical Research Fund from the Government of the Hong Kong Special Administrative Region.

Funding

This research was supported by a commissioned grant from the Health and Medical Research Fund from the Government of the Hong Kong Special Administrative Region and Award Number U54GM088558 from the US National Institute of General Medical Sciences. P.M.D was supported by the Fellowship Foundation Ramon Areces. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. The funding bodies had no role in study design, data collection and analysis, preparation of the manuscript, or the decision to publish. All authors have seen and approved the manuscript. All authors have contributed significantly to the work. All authors report no conflicts of interest. The manuscript and the data contained within have not been published and are not being considered for publication elsewhere.

References

1 Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, doi:10.1016/S0140-6736(20)30260-9.

2 Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine*, doi:10.1056/NEJMoa2001316 (2020).

3 Data Platform of Shanghai Observer. *Line list of 2019-nCoV confirmed fatal cases (from publicly available information)*, < http://data.shobserver.com/www/datadetail.html?contId=1000894> (2020).

4 Wuhan Municipal Health Commission. *Wuhan Municipal Health Commission's briefing on the current pneumonia epidemic in the city*, < http://wjw.wuhan.gov.cn/> (2020).

5 Hubei Municipal Health Commission. *Hubei Municipal Health Commission's briefing on the current pneumonia epidemic in the province*, < http://wjw.hubei.gov.cn/fbjd/dtyw/> (2020).

6 Chinese Center for Disease Control and Prevention. *Dashboard of reported 2019-nCoV cases*, < http://2019ncov.chinacdc.cn/2019-nCoV/> (2020).

7 Backer, J. A., Klinkenberg, D. & Wallinga, J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance* **25**, 2000062, doi:doi: https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062 (2020).

8 The State Council of The People's Republic of China. *Press Conference of the Joint Prevention and Control Mechanism of the State Coucil*, < http://www.gov.cn/xinwen/2020-02/09/content_5476513.htm> (2020).

9 National Health Commission of People's Republic of China. *Notice of the General Office of the National Health Commission on the Distribution of the Plan of Prevention and Control of the Pneumonia Caused by the Novel Coronavirus (Version 4)*, < http://www.gov.cn/zhengce/zhengceku/2020-02/07/content_5475813.htm> (2020).

10 Hubei Municipal Health Commission. *Situation of the epidemic of pnumonia caused by the novel coronavirus in Hubei, as of Feb 12*, < http://wjw.hubei.gov.cn/fbjd/dtyw/202002/t20200213_2025581.shtml> (2020).

11 Jia, N. *et al.* Case fatality of SARS in mainland China and associated risk factors. *Tropical Medicine & International Health* **14**, 21-27 (2009).

12 Donnelly, C. A. *et al.* Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *The Lancet* **361**, 1761-1766, doi:10.1016/S0140-6736(03)13410-1 (2003).

13 Leung, G. M. *et al.* The epidemiology of severe acute respiratory syndrome in the 2003 Hong Kong epidemic: an analysis of all 1755 patients. *Annals of internal medicine* **141**, 662-673 (2004).

14 Lau, E. H. *et al.* A comparative epidemiologic analysis of SARS in Hong Kong, Beijing and Taiwan. *BMC infectious diseases* **10**, 50 (2010).

15 Taubenberger, J. K. & Morens, D. M. 1918 Influenza: the mother of all pandemics. *Revista Biomedica* **17**, 69-79 (2006).

16 Collins, S. D. Age and Sex Incidence of Influenza in the Epidemic of 1943-44, with Comparative Data for Preceding Outbreaks: Based on Surveys in Baltimore and Other Communities in the Eastern States. *Public Health Reports (1896-1970)* **59**, 1483-1503, doi:10.2307/4585095 (1944).

17 Andreasen, V., Viboud, C. & Simonsen, L. Epidemiologic Characterization of the 1918 Influenza Pandemic Summer Wave in Copenhagen: Implications for Pandemic Control Strategies. *The Journal of Infectious Diseases* **197**, 270-278, doi:10.1086/524065 (2008).

18 Wong, J. Y. *et al.* Infection Fatality Risk of the Pandemic A(H1N1)2009 Virus in Hong Kong. *American Journal of Epidemiology* **177**, 834-840, doi:10.1093/aje/kws314 (2013).

19 Donaldson, L. J. *et al.* Mortality from pandemic A/H1N1 2009 influenza in England: public health surveillance study. *BMJ* **339**, b5213, doi:10.1136/bmj.b5213 (2009).

20 Oh, M.-d. *et al.* Middle East respiratory syndrome: what we learned from the 2015 outbreak in the Republic of Korea. **33**, 233 (2018).

21 Wong, J. Y. *et al.* Hospitalization fatality risk of influenza A (H1N1) pdm09: a systematic review and meta-analysis. *American journal of epidemiology* **182**, 294-301 (2015).

22 Abbad, A. *et al.* Middle East respiratory syndrome coronavirus (MERS-CoV) neutralising antibodies in a high-risk human population, Morocco, November 2017 to January 2018. *Eurosurveillance* **24** (2019).

23 Wallinga, J. & Lipsitch, M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences* **274**, 599-604, doi:doi:10.1098/rspb.2006.3754 (2007).

24 Chan, J. F.-W. *et al.* A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The*

Lancet (2020).
25 Hubei Provincial Bureau of Statistics. *Wuhan Statistical Yearbook 2010-2018*, < http://tjj.hubei.gov.cn/tjsj/sjkscx/tjnj/gsztj/whs/> (2020).

Figures

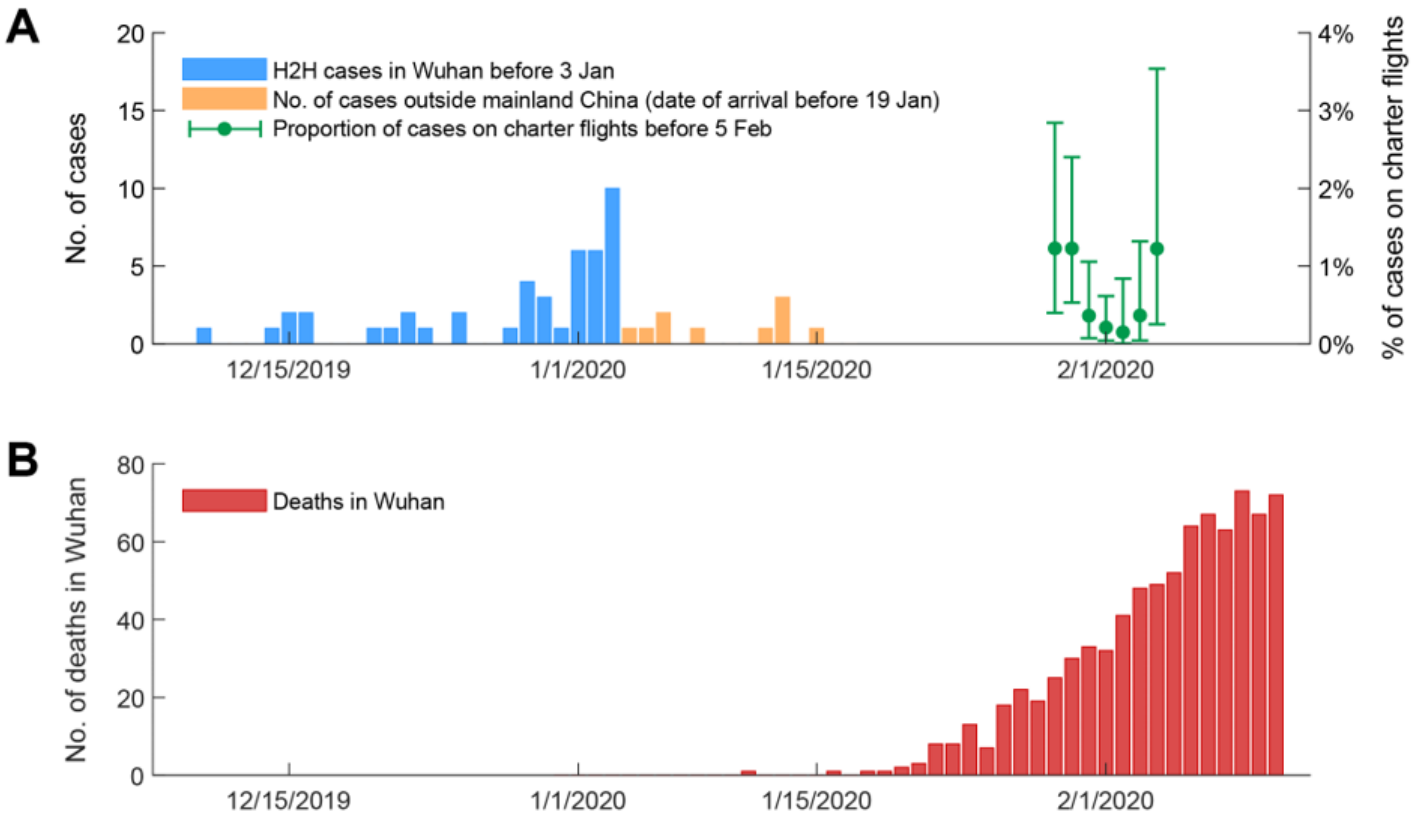

Figure 1

Data used in the inference. (A) The daily number of confirmed cases in Wuhan with no epidemiologic links to Huanan Seafood Wholesale Market between 1 Dec 2019 and 3 Jan 2020 (blue), the daily number of cases exported from Wuhan to cities outside mainland China via air travel between 25 Dec 2019 and 19 Jan 2020 (orange), and the proportion of expatriates on charter flights between 29 Jan and 4 Feb 2020 who were lab-confirmed to be infected (green). (B) The daily number of death cases in Wuhan reported between 1 Dec 2019 and 11 Feb 2020.

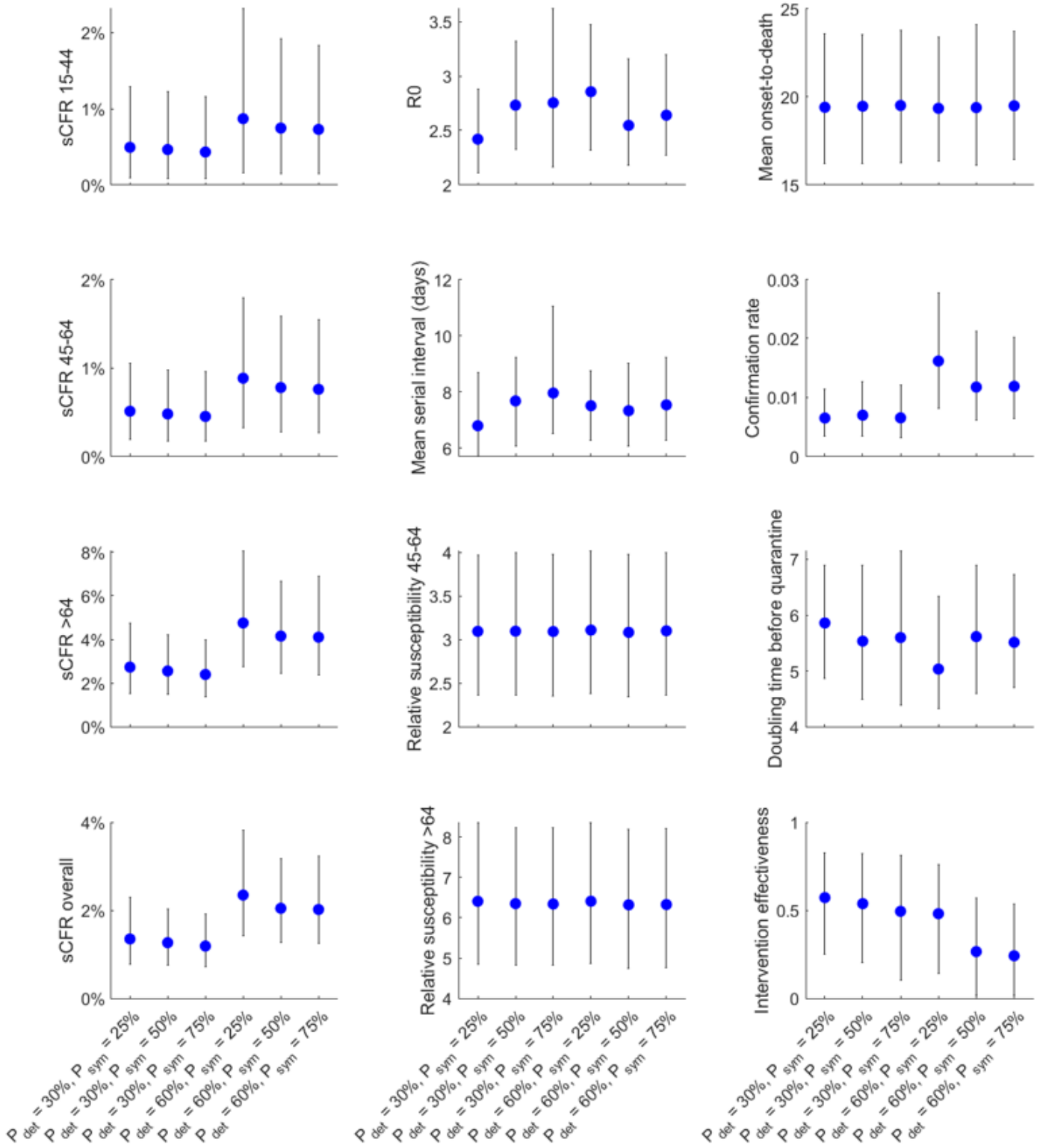

Figure 2

Inferred values of key epidemiologic parameters of COVID-19 in Wuhan.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinformation.pdf