

1 **Transmission dynamics and control measures of COVID-19 outbreak in**
2 **China: a modelling study**

3

4 Xu-Sheng Zhang^{1,2,*}, Emilia Vynnycky^{1,3}, Andre Charlett¹, Daniela de Angelis^{1,4}, Zhengji
5 Chen⁵, PHE COVID-19 modelling group, and Wei Liu^{5,*}

6

7 ¹Centre for Infectious Disease Surveillance and Control, National Infection Service, Public
8 Health England, 61 Colindale Avenue, London NW9 5EQ, UK

9 ²Medical Research Council Centre for Outbreak Analysis and Modelling, Department of
10 Infectious Disease Epidemiology, Imperial College Faculty of Medicine, Norfolk Place,
11 London W2 1PG, UK

12 ³TB Modelling Group, TB Centre, Centre for Mathematical Modelling of Infectious Diseases
13 and Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical
14 Medicine, London, UK.

15 ⁴Medical Research Council Biostatistics Unit, University Forvie Site, Robinson Way,
16 Cambridge CB2 0SR, UK.

17 ⁵School of Public Health, Kunming Medical University, Kunming, Yunnan, P. R. China

18

19 * Corresponding authors:

20 Xu-Sheng Zhang

21 Email: xu-sheng.zhang@phe.gov.uk

22

23 Wei Liu

24 Email: liuweikm@qq.com

25 **Abstract**

26 COVID-19 is reported to have been effectively brought under control in China at its initial
27 start place. To understand the COVID-19 outbreak in China and provide potential lessons for
28 other parts of the world, in this study we combine a mathematical modelling with multiple
29 datasets to estimate its transmissibility and severity and how it was affected by the
30 unprecedented control measures. Our analyses show that before 29th January 2020, the
31 ascertainment rate is 6.9%(95%CI: 3.5 – 14.6%); then it increased to 41.5%(95%CI: 30.6 –
32 65.1%). The basic reproduction number (R_0) was 2.23(95%CI: 1.86 – 3.22) before 8th
33 February 2020; then it dropped to 0.04(95%CI: 0.01 – 0.10). This estimation also indicates
34 that the effect on transmissibility of control measures taken since 23rd January 2020 emerged
35 about two weeks late. The confirmed case fatality rate is estimated at 4.41%(95%CI: 3.65 –
36 5.30%). This shows that SARS-CoV-2 virus is highly transmissible but less severe than
37 SARS-CoV-1 and MERS-CoV. We found that at the early stage, the majority of R_0 comes
38 from the undetected infected people. This implies that the successful control in China was
39 achieved through decreasing the contact rates among people in general populations and
40 increasing the rate of detection and quarantine of the infected cases.

41

42 Keywords: COVID-19, mainland China, modelling, mortality, transmissibility

43

44

45 **Introduction**

46 An outbreak of severe pneumonia, an infectious disease now named as COVID-19, was
47 reported by Chinese authorities in December 2019. The aetiological agent, SARS-CoV-2, a
48 novel coronavirus, was isolated by Chinese authorities on 7 January 2020 and reported by
49 WHO on 9th January 2020. The first case of COVID-19 was reported to have symptom onset
50 on 1st December 2019 in Wuhan city of Hubei province, China¹, and the virus was quickly
51 spread to other parts of China^{2,3} and the first case was reported outside of China on 13th
52 January 2020 (see Table 1). Because of the rapid spread of the disease, WHO announced the
53 outbreak of COVID-19 as a “public health emergency of international concern” on 30th
54 January 2020 and assessed COVID-19 as a pandemic on 11th March 2020⁴. Up to 25th
55 March 2020, 405,742 cases and 18,791 fatalities associated with COVID-19 have been
56 reported globally. Of these, 81,218 cases and 3,281 fatalities have been reported in mainland
57 China. Up to 10th April 2020 (25 days late), the global numbers increase hugely to
58 1,521,252 cases and 92,798 deaths while the corresponding numbers in mainland China are
59 81,953 cases and 3,339 deaths. These numbers of COVID-19 cases and fatalities indicate that
60 the outbreaks outside of China develop quickly, whereas the daily reported number of new
61 cases in mainland China remained very low. This situation of mainland China can be
62 attributed to the draconian and rapid control measures taken in China since late January 2020.
63 The control measures started from the lockdown of the epicentre Wuhan city on 23rd January
64 2020 and then extended to the whole of mainland China³. Extreme response measures
65 included the complete shutdown and isolation of whole cities, cancellation of Chinese New
66 Year celebrations, and prohibition of attendance at school and work, broadcast of critical
67 information (e.g., promoting hand washing, mask wearing, and care seeking) with high
68 frequency through multiple channels, massive mobilization of health and public health
69 personnel as well as military medical units, rapid construction of entire hospitals for patients

70 with severe symptoms, and reconstruction of shelters for patients with none or mild
71 symptoms. Control measures integrate travel bans and restrictions, contact reductions and
72 social distancing, screening and contact tracing, early case identification and isolation^{2,5} to
73 curb the epidemic. The massive vigorous actions taken by the Chinese government and
74 Chinese people have slowed down the epidemic of COVID-19 and now the spread of the
75 disease has been brought under control in China.

76

77 As a novel coronavirus, it is important to understand its transmissibility and severity from the
78 data of the outbreak and to judge how effectively the control measures taken in China have
79 helped stop the spread of SARS-CoV-2 virus in mainland China. This understanding should
80 provide some critical information for controlling COVID-19 outbreaks in other parts of the
81 world and for its potential re-emergence in future. In particular, exploring the main driver of
82 the spread will provide practically useful information for policy makers to make effective
83 non-pharmaceutical interventions at its early occurrence. To estimate the early dynamics of
84 transmission in Wuhan, Kucharski et al⁶ fitted a stochastic transmission dynamic model to
85 four datasets up to 11th February 2020: daily number of new internationally exported cases;
86 daily number of new cases in Wuhan with no market exposure; daily number of new cases in
87 China; and proportion of infected passengers on evacuation flights. They estimated that the
88 median daily reproduction number (R_t) in Wuhan declined from 2.35 (95% CI: 1.15–4.77) 1
89 week before travel restrictions were introduced on 23rd January 2020, to 1.05 (95% CI: 0.41–
90 2.39) 1 week after. Many other modelling studies^{2,3,7,8} also used the confirmed case data at
91 the early stage to model and estimate the spread dynamics and control of COVID-19 outbreak
92 in China. As the number of new confirmed case in mainland China reached a peak only after
93 17th February 2020 and other relevant data such as fatalities and hospital discharges are
94 becoming available, a further study including more data streams is needed to show the overall

95 effectiveness of the integrated control measures executed in the whole nation. Based on the
96 observed data associated with the confirmed cases, death and recovery that were published by
97 Chinese National Health Commission, in this study we propose a Synthesis model (Figure 1)
98 which covers both transmission dynamics of SARS-CoV-2 virus and disease reporting
99 process of COVID-19 to estimate the transmissibility and severity of COVID-19 in China
100 and the impact on the spread of COVID-19 of control measures the Chinese government has
101 taken from late January 2020.

102

103 **Results**

104 The estimates of the model parameters are shown in Table 2. The detection and reporting rate
105 (or ascertainment rate) of COVID-19 cases in mainland China was about 6.9% (95%CI: 3.5 –
106 14.6%) before 29th January 2020, and then increased to 41.5% (95%CI: 30.6 – 65.1%)
107 afterwards. The early reporting rate is very low and about 93.1% of all infections were
108 undetected prior to 29th January 2020 (c.f.^{7,9,10}), which might be mainly due to the limited
109 knowledge and unclear definition of the novel disease¹¹. The reporting rate of 41.5% at the
110 late stage reflects the increased awareness of the virus and consequently escalating rate of
111 medical help seeking behaviour for respiratory symptoms; however, this rate appears not high
112 and might indicate that many infections are of mild symptoms or asymptomatic. China
113 CDC¹² used data up to 11th Feb 2020 and estimated that COVID-19 has been mild for 81%.
114 The high proportion of mild symptomatic or asymptomatic infections was further confirmed
115 by a well-investigated outbreak on Diamond Princess cruise ship during February 2020:
116 among 696 confirmed cases, 410 (58.9%; c.f.¹³) are asymptomatic¹⁴.

117

118 At the early stage before 8th February 2020, the transmissibility of COVID-19 is high with R_0
119 = 2.23 (95% confidence interval (CI): 1.86 – 3.22); However, it reduced dramatically to 0.04

120 (95%CI: 0.01 – 0.10) from 8th February 2020 (95% CI: 5th – 9th February 2020). This
121 estimate of R_0 at the early stage are consistent with most of previous estimates^{3,6,7,8,9,10,15–20}.
122 The integrated intervention measures is estimated to reduce transmissibility of 1.8% of its
123 initial value (c.f.³). In our model, the basic reproduction number is contributed by two parts:
124 the confirmed cases and undetected infections. By fixing the duration from onset of
125 symptoms to death and recovery at 17.8 days and 22.6 days respectively²¹, the infectious
126 period (D_1) for these that are symptomatic and confirmed and then quarantined is 2.43 days
127 (95%CI: 2.02 – 3.64 days) and the infectious period for undetected infections (D_1+D_u) is 5.44
128 days (95%CI: 3.37 –12.62 days). Our estimate of infectious period before isolation is
129 comparable with the estimates of²²⁻²⁵; though it appears to be shorter than that obtained by³:
130 5.19 days (95%CI: 4.51 – 5.86 days). The estimate of infectious period for the infections that
131 were not detected and not quarantined is in agreement with the reviewed range^{26,27} from 1 to
132 14 days.

133

134 Both confirmed cases and undetected infections are assumed to be of the same infectivity in
135 this study; combining with the estimate that only about 6.9% of infections were confirmed at
136 the early stage, this suggests that at the early stage, about 96% of R_0 (Table 2) is due to the
137 undetected infections. If the infectivity of undetected infections is only half or one third of
138 that of confirmed cases⁷, the ascertainment rate increases to 12.2% or 17.6% (Supplementary
139 Table S2.1). And under the conditions where undetected infection are less infectious, the 88%
140 or 71% of R_0 is due to the undetected infections. This indicates that only isolating confirmed
141 cases and their contacts is not enough to stop the spread and that the main factor that stopped
142 the COVID-19 outbreak in mainland China was the dramatic drop in contact rates among the
143 general population (c.f.⁷). This may further explain why the impact of the draconian control

144 measures executed in China since the 23rd January 2020 took about 16 days (i.e., delay from
145 23rd January to 8th February 2020) to effect in stopping the transmission within community.
146
147 Our model analysis shows the confirmed case fatality rate (cCFR) is 4.41% (95%CI: 3.65 –
148 5.30%). This is slightly below the estimate of ²⁸: 5.65% (95%CI: 5.50–5.81%). Deng et al²⁸
149 used individual information of cases to obtain their estimate which is close to both the
150 corresponding crude or naïve confirmed case fatality risk: $4632/82758 = 5.60\%$ and the
151 approximator of $\text{deaths}/(\text{deaths} + \text{recoveries}) = 4632/(4632+78112) = 5.60\%$ as of 21st April
152 2020. Based on the data up to 11th February 2020, Verity et al²¹ suggested that the overall
153 CFR for China outbreak is 1.38% (95%CI: 1.23-1.53%). Wu et al²⁹ found the overall
154 symptomatic case fatality rate (the probability of dying after developing symptom) is 1.4%
155 (95%CI: 0.9–2.1%). Russell et al³⁰ based on naïve estimates to obtain CFR in China to be
156 1.2% (95% CI: 0.3–2.7). Our estimate of CFR is higher than these. This should be obvious
157 because these three studies did not include the 1290 deaths added on 17th April 2020 by
158 Chinese National Health Commission³¹. Within the Synthesis model, the hidden transmission
159 dynamics was inferred through observables associated with the disease and so was the
160 number of total infections during the outbreak in mainland China: we found that there were
161 238894 (95%CI: 127486 – 387917) infected people in mainland China up to 21st April 2020
162 (Figure 2). Using this information, our estimate of infection fatality rate (IFR) is 0.30%
163 (95%CI: 0.15 – 0.66%) prior to 29th January 2020, and it increased to 1.81% (95%CI: 1.26 –
164 3.03%) afterwards. This is comparable with estimates of ^{21,30} since both studies were
165 performed before the addition of 1290 deaths on 17th April in Wuhan.
166
167 Figure 2 shows the model fitting to the daily number of hospitalizations (confirmed cases),
168 the daily number of deaths, and daily number of recoveries. Our theoretical model well

169 reconstructs the dynamic changes in daily numbers of confirmed cases and recoveries.
170 However, the model predictions of the daily number of deaths shifted about one week late.
171 Our model inference suggests that on 1st December 2019, there were 185 (95%CI: 51 – 354)
172 people who carried SARS-CoV-2 virus. If the public were made aware of COVID-19 earlier
173 and the control measures started earlier, the size of outbreak should be kept much smaller.
174 This can be quantitatively analysed by assuming the same epidemiological characteristics but
175 moving the start time of both reporting and transmission turning time points. The results are
176 listed in Table 3. It shows that if the control measures started one week, two weeks, or three
177 weeks earlier, about 51% (60%), 76% (86%) and 89% (96%) of the confirmed cases (and
178 deaths) would be averted. However, if the control measures started one week, two weeks, or
179 three weeks late, the numbers of the confirmed cases (and deaths) would increase 2.0-fold
180 (2.2), 4.1-fold (4.7) and 8.3-fold (9.76) across mainland China, respectively. This estimate of
181 the effect of start time of control measures is smaller than but comparable with what Lai et al²
182 and Yang et al⁵ found, both studies estimated only the number of infections.

183

184 To compare how the epidemic within the epicentre differs from that over the whole nation,
185 we also obtain the estimates of model parameters for Hubei province (Supplementary Table
186 S3.1). The results show that within the epicentre, transmissibility is only slightly higher ($R_0 =$
187 2.34 (95%CI: 1.92 – 3.55) versus 2.23 (95%CI: 1.86 – 3.22)). Ascertainment rate at the early
188 stage within Hubei province is more than double the national rate (15.4% (95%CI: 6.8 –
189 28.7%) versus 6.9% (95%CI: 3.5–14.6%)); whereas the amount of transmission due to
190 undetected cases is slightly reduced (93% versus 96%). The confirmed case fatality rate is
191 higher (cCFR= 5.16% (95%CI: 4.21–6.28%) in Hubei province versus 4.41% (95%CI: 3.65–
192 5.30%) over the whole nation with the infection fatality rate at early stage of the outbreak

193 being more than double the national rate (0.79% (95%CI: 0.35–1.53%) versus 0.30% (95%CI:
194 0.15–0.66%).

195

196

197 **Discussion**

198 Understanding the transmissibility and severity of the novel coronavirus (SARS-CoV-2) is
199 paramount; further disclosing how its rapid spread was brought under control in mainland
200 China is of practical implication for other countries now facing the ongoing outbreaks of
201 COVID-19. In this study we have explored these by using a mathematical model to
202 reconstruct the COVID-19 outbreak in mainland China from 1st December 2019 to 21st April
203 2020 and assess the impact of its unprecedented control measures. Our analyses indicate that
204 the SARS-CoV-2 has a basic reproduction number of 2.23 under the situation of no
205 intervention measures and therefore is highly transmissible. The fatality rate among those that
206 are symptomatic and confirmed is about 4.41%. The draconian control measures taken by
207 Chinese government from 23rd January 2020 has brought the spread of SARS-CoV-2 under
208 control in mainland China. However, it took more than two weeks for the effect of control
209 measures to emerge.

210

211 The massive vigorous actions taken by the Chinese government have stopped the spread of
212 COVID-19 in mainland China. This was achieved under very strict control measures, which
213 dropped the transmissibility from $R_0 = 2.23$ to 0.04. We found that the infectious period
214 before isolation for the confirmed cases is about 2.4 days and the infectious period for these
215 undetected is about 5.4 days ranging from 3.4 to 12.6 days. Although increasing the rate of
216 detection and quarantine of symptomatic cases can help reduce the sources of infection (c.f.²),
217 the main force of transmission is from the undetected cases which contributed the vast

218 majority of the transmissibility of SARS-CoV-2 during the early stage of the outbreak in
219 mainland China (c.f.⁷). Hence it is the restriction measures, which limited the mobility of the
220 general population that stopped the spread of SARS-CoV-2 virus in mainland China.
221 Through model fitting to observed data in mainland China, we found that the model predicted
222 fourteen times higher cases than were reported at the early stage of the outbreak; even at the
223 late stage since 29th January 2020, the estimated number of infected people were more than
224 twice that were reported. Based on the epidemiological characteristics obtained from
225 COVID-19 outbreak data, our analyses suggest that even if the control measures started one
226 week earlier, this will avoid 51% confirmed cases and 60% deaths. If starting three weeks
227 earlier, then 89% confirmed cases and 96% of deaths will be averted (c.f.²). This reinforces
228 our common sense: the earlier discovering and earlier response, the easier and better control
229 strategy.

230

231 As in ⁶, we assumed the latent period is equal to the incubation period. In view of evidence
232 that there is pre-symptomatic transmission^{27,32,33}, it is interesting to know whether this will
233 alter the results of our model. For this, we modified the model system (Figure 1) by dividing
234 exposure stage E into two equal sub-classes E_1 and E_2 and assuming people in E_2 can transmit
235 the virus with the same infectivity of ill cases (SI Section 4). We obtain the quite similar
236 results (Supplementary Table S4.1): For example, the basic reproduction number before the
237 control measures is 2.24 (95%CI: 1.86 – 2.88) among which 95% were due to undetected
238 cases, and the cCFR is 4.41% (95%CI: 4.41 (95%CI: 3.67 – 5.37%). This indicates our model
239 is robust to possible pre-symptomatic transmission. Nevertheless, this might imply that
240 population models of SARS-CoV-2 virus transmission cannot distinguish where or not there
241 is pre-symptomatic transmission.

242

243 The non-pharmaceutical interventions executed in mainland China have stopped the spread of
244 the virus, but the risk of another potential outbreak lies in the nearby future. This argument
245 come straightforward³. As our model estimated, at most 238894 individuals got infected up to
246 21st April 2020 during the outbreak, which is only 0.02% of the Chinese population. Even if
247 all those recovered from infection have built complete immunity³⁴, this immunity level in
248 Chinese population is too lower than the herd immunity of $1-1/R_0 > 50\%$ which is required to
249 protect the population^{35,36}. This simply implies that once the strict quarantine measures
250 currently executed in mainland China is released, the rebound of COVID-19 in China is very
251 likely especially under the current rising outbreaks across the rest of the world (c.f.³⁷).
252 Nevertheless, the non-pharmaceutical interventions have halted the spread of SARS-CoV-2
253 virus in mainland China and bought time for vaccines and drugs to be developed and used
254 late on.

255

256 It is of interest to compare the novel coronavirus SARS-CoV-2 with other two coronaviruses:
257 SARS-CoV-1 and MERS-CoV that caused large outbreaks in human populations. SARS-
258 CoV-1 has R_0 from 2–5³⁸ and case-fatality rate of 9.6% among probable cases in mainland
259 China³⁹. MERS-CoV in the 2015 outbreak in South Korea has been estimated to have a R_0
260 from 2–7⁴⁰ and case fatality rate of 34.5% among laboratory-confirmed cases⁴¹. It indicates
261 that SARS-CoV-2 is nearly as transmissible as SARS-CoV-1 and MERS-CoV, but less
262 severe. Furthermore, there was no evidence of a super-spreader event occurring in any of the
263 Chinese health facilities serving COVID-19 patients, which is distinct from the 2003
264 outbreak of SARS-CoV and 2015 MERS-CoV outbreak in South Korea. Unfortunately, due
265 to the advanced modern transportation, people can move easily and quickly across the world,
266 and the SARS-CoV-2 quickly spread to other countries. This is in a sharp contrast with both
267 SARS-CoV-1 and MERS-CoV which have been controlled and confined within relatively

268 limited areas of the world. This sharp difference may be attributed to another aspect of the
269 coronaviruses: a large proportion of SARS-CoV-2 infections are of mild symptoms or
270 asymptomatic¹²⁻¹⁴, while both SARS-CoV-1 and MERS_CoV are highly symptomatic^{20,32}.
271 This characteristic of SARS-CoV-2 virus will make it a real challenge for humans to control
272 and manage it.

273

274 The strength of our analysis comparing with previous studies^{2,3,5,6,20} lies in two aspects: our
275 investigations are based on the three datasets (confirmed cases, death and recovery) and we
276 model the outbreak over a long period (143 days) which should avoid any bias and
277 confounding arising due to observations over a short period. This study also has several
278 limitations. To model complicated processes of transmission dynamics and disease reporting,
279 Synthesis model has been simplified in several aspects. To reflect the temporal change in
280 both ascertainment rate and transmissibility of COVID-19, two different values are assumed
281 for each of them. The change in ascertainment rate and transmissibility may gradually take
282 place during the outbreak as do the public awareness and interventions³. For example, Tsang
283 et al¹¹ found the ascertainment rate changed as the case definition for COVID-19 changed
284 from initially narrow to gradually wider during the period from 15th January to 3rd March. In
285 the current study, we assume the confirmed case fatality rate (cCFR) remained unchanged
286 during the outbreak in mainland China. It may reduce with time as medical conditions and
287 clinical treatments improved. The time-to-event intervals such as the delay from symptom
288 onset to death may also changes as epidemic grows^{21,23,25}. Further, in this study we ignore the
289 heterogeneity in both geography and age^{3,7,12}. To provide more specific and practically
290 useful information for control measures, it needs to look at variation in regions^{3,7} and age
291 groups¹². A further limit is we model the overall effectiveness of integrated intervention

292 measures rather than the different types of control measures and therefore cannot provide
293 specific information for their relative impacts on stopping the spread of infection (c.f.^{2,3,15}).

294

295 In conclusion, our finding that the main driver of transmission of SARS-CoV-2 at the early
296 stage of the outbreak in mainland China came from the undetected infections provides vital
297 information for policy makers when designing the optimal intervention strategies. Under the
298 situation where vaccine and effective drugs are not available, early detection and isolation are
299 essential for containing and controlling the spread of SARS-CoV-2 but the most crucial is to
300 reduce contact rates among people in the general population.

301

302 **Methods**

303 **Data**

304 We extracted the following data relating to COVID-19 for mainland China from 3 datasets on
305 the website of Chinese National Health Commission: the daily number of confirmed cases
306 who were confirmed / admitted to hospital, the daily number of deaths and the number of
307 patients who had recovered each day. The originally reported data from 1st December 2019 to
308 21st April 2020 are collected for this study. Here the data are given by the symptom onset
309 date during the period 1st December 2019 – 1st January 2020 from¹ and by reporting date
310 thereafter from the website of Chinese National Health Commission. The reporting date is
311 assumed to be the same as the date that cases were diagnosed. Before 12th February 2020,
312 confirmed cases were defined as those who were positive for SARS-CoV-2 on PCR; from
313 12th February 2020 confirmed cases were defined (for the epicentre Hubei province) as those
314 who were either clinically diagnosed or positive for SARS-CoV-2 on PCR¹¹.

315

316 On 17th April 2020, Chinese National Health Commission raised the total death number in
317 Wuhan city from 2579 to 3869³¹. The announcement of the Chinese National Health
318 Commission stated that these added 1290 deaths most likely occurred before 20th February
319 2020. Because, after this date, the number of hospitals that can treat COVID-19 patients
320 increased from 2 to 48, and the HOUSHEN and LEISHEN hospitals, and FANGCANG
321 shelters that have been constructed since also provided much more beds meeting the needs of
322 COVID-19 patients with different symptoms. Further, the system of data collection improved
323 rapidly, and the number of missed cases/deaths decreased greatly. As the detailed information
324 for these 1290 deaths is not available, for the sake of model fitting these 1290 deaths were
325 relocated in the proportion to the number of deaths on each day reported before from the date
326 of first death (10th January 2020) to 20th February 2020. For example, the number of deaths
327 reported on 18th January 2020 was 26, the number is now corrected to $26 + 26 \times 1290 / 2236 =$
328 $26 + 15 = 41$. Here 2236 is the cumulative number of death up to 20th February 2020. The so-
329 modified death data will be used in this study. For sensitivity analysis, we also assume that
330 these 1290 deaths added on 17th April 2020 were distributed in the proportion to the daily
331 number of deaths reported before over the entire period from 10th January to 17th April 2020.
332 The results shown in Table 2 are very similar under two ways of relocating 1290 deaths
333 added on 17th April 2020.

334

335 **Model**

336 In this study we use a synthesis model⁴² (Figure 1), which combines the hidden transmission
337 dynamics of SARS-CoV-2 virus and the reporting system of COVID-19, to investigate the
338 transmissibility and severity of COVID-19 and the efforts to contain and control the spread of
339 SARS-CoV-2. We assume the transmission dynamics of SARS-CoV-2 virus is described by
340 an SEIR compartmental model. That is, a susceptible person (*S*) can contract SARS-CoV-2

341 virus from infectious persons and then enter the latent class (E); after a latent period (L), the
342 exposed person progresses to become infectious (I) for a period and can transmit the SAR-
343 CoV-2 virus before recovering or dying (R). We assume that some of infected people (with
344 severe symptoms) will be detected and admit to hospital and will then be treated and isolated
345 from the community; while other mild or asymptomatic infections (I_u) will not be detected
346 and hence can continue to exist in the community as sources of infection. For simplicity, we
347 assume that the latent period is equal to the incubation period, which is fixed at the value of L
348 = 5.2 days as estimated by^{17,25,43}. SARS-CoV-2 virus can be transmitted by three possible
349 modes: respiratory transmission (through respiratory droplets when symptomatic people
350 sneeze or cough), aerosol transmission (through fine virus particles that were aerosolized)
351 and contact transmission (through contacting the contaminated surface). In view of this easy
352 and quick mode of transmission, random mix among people is a reasonable assumption.
353 Although China is a large country with a population of 1400050000 people residing on a
354 huge area of 9,596,960 km², the recent urbanization and the development of rapid transport
355 systems in China make it easy and quick for its people to move around the country. Hence it
356 should be rational to model the transmission of COVID-19 within the whole country as a well
357 mixing population. For comparison, we also model the spread of COVID-19 within Hubei
358 province where the outbreak started (SI Section 3).

359

360 The synthesis model is approximated by equation (1).

361

$$\frac{d}{dt}S(t) = -\beta(t)S(t)(I_1(t)(\theta_1 + \xi(1 - \theta_1)) + \xi I_u(t))/N$$

$$\frac{d}{dt}E(t) = \beta(t)S(t)(I_1(t)(\theta_1 + \xi(1 - \theta_1)) + \xi I_u(t))/N - E(t)/L$$

362 $\frac{d}{dt}I_1(t) = E(t)/L - I_1(t)/D_1 + Imported(t)$

$$363 \quad \frac{d}{dt} I_u(t) = (1 - \theta_1) I_1(t) / D_1 - I_u(t) / D_u \quad (1)$$

$$\frac{d}{dt} H(t) = \theta_1 I_1(t) / D_1 - 2\theta_2 H(t) / D_{\text{Death}} - 2(1 - \theta_2) H(t) / D_{\text{Recovery}}$$

$$\frac{d}{dt} \text{Recovery}_0(t) = 2(1 - \theta_2) H(t) / D_{\text{Recovery}} - 2\text{Recovery}_0(t) / D_{\text{Recovery}}$$

$$\frac{d}{dt} \text{Recovery}(t) = 2\text{Recovery}_0(t) / D_{\text{Recovery}}$$

$$\frac{d}{dt} \text{Death}_0(t) = 2\theta_2 H(t) / D_{\text{Death}} - 2\text{Death}_0(t) / D_{\text{Death}}$$

$$\frac{d}{dt} \text{Death}(t) = 2\text{Death}_0(t) / D_{\text{Death}}$$

364

365 Here $N=1400050000$ is the size of total population in mainland China and is assumed to be
366 constant during the outbreak. The 12 compartments are defined in Table 4 and the definitions
367 of model parameters are given in Table 2. Note that an item *Imported(t)* in equation for I_1 is
368 included to account for the imported cases which come back from other countries to China
369 since 24th March 2020.

370

371 As a novel virus, no effective pharmaceutical interventions were available. Non-
372 pharmaceutical interventions are therefore essential public health response to the outbreaks.
373 These include isolating ill persons, contact tracing, quarantine of exposed persons, travel
374 restrictions, school and workplace closures, and cancellation of mass gathering events. These
375 measures were expected to reduce transmission^{2,5}. During the period of COVID-19 outbreak
376 in China, the day, 20th January 2020, when the information that “SARS-CoV-2 virus can
377 transmit among people” was announced in public marked some huge and quick changes in
378 people contact behaviours and rates. These changes were further strengthened when the
379 lockdown of epicentre Wuhan city from 23rd January 2020 and then quick follow-up of
380 nationwide restrictions of mobility were enforced (Table 1 Timeline of the outbreak in

381 China). To model the potential changes in transmission rate due to these combined and
382 dramatic control measures taken by Chinese government, we simply assume there existed a
383 turning point in time τ_β such that the transmission coefficient varies as shown in the equation

384 (2)

385

$$386 \quad \beta(t) = \begin{cases} \beta_a & t \leq \tau_\beta \\ \beta_b & t > \tau_\beta \end{cases} \quad (2)$$

387 (c.f.^{3,40}).

388

389 At the early stage of COVID-19 outbreak in China, due to the limited knowledge and unclear
390 definition of COVID-19, the rate of detection and reporting was very low (i.e., under-
391 detection and under-reporting). However, the rate of detection and reporting increased
392 quickly with the knowledge about the COVID-19 and availability of advanced techniques to
393 test SARS-CoV-2¹¹. Especially since 20th January 2020, searches for cases, diagnosis and
394 reporting have sped up. Local governments across China encouraged and supported routine
395 screening and quarantine of travellers from Hubei Province to discover COVID-19 infections
396 as early as possible². To reflect the changes in detection and reporting of COVID-19 cases
397 around 20th January 2020 when the SAR-CoV-2 was publicly declared to be able to transmit
398 among people, we assume the proportion of detection (or rate of ascertainment) varied with
399 time as

$$400 \quad \theta_1 = \begin{cases} \theta_{1,a} & t \leq \tau_\theta \\ \theta_{1,b} & t > \tau_\theta \end{cases} \quad (3)$$

401 Here τ_θ is the turning point in detection and reporting.

402

403 In the disease reporting system of the synthesis model (see the shaded part of Figure 2), we

404 assume all confirmed cases were hospitalised. The hospitalised patients either died at the

405 probability of θ_2 or discharged because of recovery. Despite sporadic cases of COVID-19
 406 that might not have chances to be hospitalized and died at home especially at the early stage
 407 of the outbreak, our assumed flow of confirmed COVID-19 patients should well approximate
 408 actual procedures during the outbreak in mainland China⁴⁴. In this study we fix the durations
 409 from onset of symptoms to death and to hospital discharge at $(D_1+D_{\text{eath}}=)$ 17.8 days and
 410 $(D_1+D_{\text{Recovery}}=)$ 22.6 days, respectively, the estimates of²¹ from the China outbreak data. To
 411 allow the durations from hospitalization to death and from hospitalization to recovery to have
 412 gamma distribution rather than the usual exponential distribution, we introduce the in
 413 between compartments *Death_0*, and *Recovery_0*. As we use (mostly) data of reported dates,
 414 the time-event-length (duration) should be thought to include such reporting delay, which is
 415 not explicitly treated in this study. Fortunately, modelling technique can code with such
 416 external noise⁶.

417

418 The basic reproduction number R_0 , which was defined as an average number of secondary
 419 infections generated by an infectious person introduced into a completely susceptible
 420 population, is an important quantity to characterise the transmissibility of infectious agents³⁵.

421 At the early stage of our model system, it is given by

$$422 \quad R_{0,1} = (\theta_{1,a} D_1 + \xi(1-\theta_{1,a})(D_1+D_u))\beta_a \quad (4a)$$

423 while at the late stage after the time point defined as $\max(\tau_\beta, \tau_\theta)$,

$$424 \quad R_{0,2} = (\theta_{1,b} D_1 + \xi(1-\theta_{1,b})(D_1+D_u))\beta_b \quad (4b)$$

425 From equation (4a), the contribution from undetected infections to the transmissibility at the
 426 early stage is

$$427 \quad R_{0,1}(\text{undetected cases}) = \xi(1-\theta_{1,a})(D_1+D_u)\beta_a \quad (4c)$$

428 Here D_1+D_u is the total infectious period of the infected people who were not detected. The

429 parameter ξ is introduced to measure the relative infectivity of undetected infections to

430 confirmed cases, and it will take the value of 1.0 (i.e., both undetected and confirmed
431 infection are of the same infectivity). To see how our results may change with different
432 relative infectivity⁷ the results for situations of $\xi = 1/2, 1/3$ will be given in Supplementary
433 Table S2.1.

434

435 **Initial seeding:** Wu et al²⁰ Lancet assumed the epidemic during 1st – 31st December 2019
436 was seeded by constant zoonotic force of infection that caused 86 cases (twice the 43
437 confirmed cases with zoonotic exposure) per day before market closure on 1st January 2020.
438 Kucharski et al⁶ assumed the outbreak started with a single infectious case or 10 cases on 22nd
439 November 2019. In this study we assume the epidemic started from 1st December 2019 with
440 initial infections of $I_1(0) = I_0$ which is to be estimated from model fitting to data. (The detail
441 of how infections were seeded is given in SI Section1)

442

443 **Inference method**

444 We denote the model parameter to be inferred as $\Theta = \{I_0, \beta_a, \beta_b, \tau_\beta, \theta_{1,a}, \theta_{1,b}, \tau_\theta, \theta_2, D_1, D_u\}$
445 which is listed in Table 2. For each set of parameter values, the Runge-Kutta fourth order
446 method was used to solve the model equations and to obtain model predicted time series of
447 infections, confirmed cases, deaths, and recoveries. In the inference of model parameters,
448 directly observed dataset of confirmed/hospitalized/reported cases (denoting it as HOS for
449 short in the following), Death and Recovery are used as illustrated in the following. To
450 capture the large dispersion in the daily numbers of these cases, the negative binomial
451 likelihood function was assumed. The likelihood for observed number $x^C(t)$ of cases on day t
452 is given as

$$453 \quad l(x^C(t) | \Theta, \eta^C) = \frac{\Gamma(x^C(t) + r^C(t))}{\Gamma(r^C(t))\Gamma(x^C(t) + 1)} \left(\frac{1}{\eta^C}\right)^{r^C(t)} \left(1 - \frac{1}{\eta^C}\right)^{x^C(t)} \quad (5)$$

454

where

$$455 \quad r^C(t) = \frac{\mu^C(t)}{\eta^C - 1} \quad (6)$$

456 Here η^C is the dispersion parameter and $\mu^C(t)$ are the predictions of the cases on day t from

457 Synthesis model (1). Superscript C represents three different datasets: HOS, Death and

458 Recovery.

459

460 Special attention was paid to the extra high daily number of cases (15152) on 12th February

461 2020 (day 74 from 1st December 2019) due to the change in the case definition in Hubei

462 province^{11,45}. In principle, these cases might have been accumulated over the days before 12th

463 February 2020. To avoid the complexity, only the cumulative numbers of cases on the 12th

464 February and daily numbers of cases after that will be used in the model inference. Let the

465 reported daily number of HOS be represented by $x(1), x(2), \dots, x(T)$, with T being the

466 number of days from 1st December 2019 to 21st April 2020. The likelihood of the cumulative

467 number of cases on 12th February $X = \sum_{t=1}^{74} x(t)$ is assumed to be

$$468 \quad L_{74} = \frac{\exp(-M)}{X!} M^X \quad (7)$$

469 Here $M = \sum_{t=1}^{74} \mu^{\text{HOS}}(t)$ represents the cumulative number of confirmed cases predicted by

470 model. Assuming that the observational daily number of death: $y(1), y(2), \dots, y(T)$ and daily

471 number of recovery: $z(1), z(2), \dots, z(T)$ are conditionally independent, the total likelihood

472 given model parameters Θ is

$$473 \quad L(\Theta, \eta^{\text{HOS}}, \eta^{\text{Death}}, \eta^{\text{Recovery}}) =$$

$$474 \quad L_{74} \times \prod_{t=75}^T l(x(t) | \Theta, \eta^{\text{HOS}}) \times \prod_{t=1}^T l(y(t) | \Theta, \eta^{\text{Death}}) l(z(t) | \Theta, \eta^{\text{Recovery}})$$

$$475 \quad (8)$$

476

477 *Parameter inference:* We assume the uninformative prior distributions $f(\Theta)$ which are
478 uniform for parameters (Table 2). Employing Bayesian framework through the combination
479 of the prior distribution $f(\Theta)$ and the likelihood $L(\Theta, \eta^{\text{HOS}}, \eta^{\text{Death}}, \eta^{\text{Recovery}}; x, y, z)$, the
480 posterior distribution can be obtained by Markov Chain Monte Carlo simulations (MCMC)⁴⁶
481 (Birrell et al 2011). From these samples, we can obtain medians and their 95% confidence
482 intervals for the model parameters. The posteriors of the model parameters will provide the
483 estimates of the transmissibility and the severity of SARS-CoV-2 in mainland China and the
484 effects on transmissibility of control measures executed by Chinese government.

485

486 **Acknowledgements**

487 This work is jointly supported by Public Health England, UK and Innovative Research Team
488 of Yunnan Province (2019(6)), China.

489 **Author Contributions**

490 Conceived and designed the mathematical model and analysis: X.-S.Z., E.V., A.C., D.D.A.
491 and W. L.; collated data: Z.C.; performed the study: X.-S.Z. and W.L.; wrote the paper: X.-
492 S.Z. and E.V.

493 **Competing financial interests:** The authors declare no competing financial interests.

494 **References**

- 495 1. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in
496 Wuhan, China. *Lancet*. **395**, 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30154-](https://doi.org/10.1016/S0140-6736(20)30154-9)
497 9 (2020)
- 498 2. Lai, S. *et al.* Effect of non-pharmaceutical interventions for containing the COVID-19
499 outbreak in China. *Nature* <https://doi.org/10.1038/s41586-020-2293-x> (2020).
- 500 3. Tian, H. *et al.* 2020. An investigation of transmission control measures during the first
501 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020)
- 502 4. <https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19>
- 503 5. Yang, Z. *et al.* Modified SEIR and AI prediction of the epidemics trend of COVID-19
504 in China under public health interventions. *J Thorac Dis* **12**(3), 165–174
505 <http://dx.doi.org/10.21037/jtd.2020.02.64> (2020)
- 506 6. Kucharski, A. J. *et al.* Early dynamics of transmission and control of COVID-19: a
507 mathematical modelling study. *Lancet Infect Dis* **20**(5), 553–558 (2020).
- 508 7. Li. R. *et al.* Substantial undocumented infection facilitates the rapid dissemination of
509 novel coronavirus (SARS-CoV2). *Science*. **368**, 489–493
510 [doi:10.1126/science.abb3221](https://doi.org/10.1126/science.abb3221) (2020)
- 511 8. Yang, Y. *et al.* Epidemiological and clinical features of the 2019 novel coronavirus
512 outbreak in China medRxiv preprint doi:
513 <https://doi.org/10.1101/2020.02.10.20021675>. (2020)
- 514 9. Read, J. M., Bridgen, J. R. E., Cummings, D. A. T., Ho, A. & Jewell, C. P. Novel
515 coronavirus 2019-nCoV: early estimation of epidemiological parameters and
516 epidemic predictions. *medRxiv* preprint doi:
517 <https://doi.org/10.1101/2020.01.23.20018549>.this version posted January 28, 2020.
518 (2020)
- 519 10. Zhao, S. *et al.* Preliminary estimation of the basic reproduction number of novel
520 coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the
521 early phase of the outbreak. *Int J Infect Dis* **92**, 214–217 (2020)
- 522 11. Tsang, T. K., Wu, P., Lin, Y., Lau, E. H. Y., Leung, G. M. & Cowling, B. J. Effect of
523 changing case definitions for COVID-19 on the epidemic curve and transmission
524 parameters in mainland China: a modelling study. *Lancet Public Health*
525 [https://doi.org/10.1016/S2468-2667\(20\)30089-X](https://doi.org/10.1016/S2468-2667(20)30089-X) (2020)

- 526 12. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team.2020
527 The Epidemiological Characteristic of an outbreak of 2019 novel coronavirus disease
528 (COVID-19)—China, 2020. *China CDC Weekly*.
- 529 13. Mizumoto, K., Kagaya, K, Zarebski, A. & Chowell, G. Estimating the asymptomatic
530 proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond
531 Princess cruise ship, Yokohama, Japan, 2020. *Euro*
532 *Surveill.* 25(10):pii=2000180. (2020)
- 533 14. https://en.wikipedia.org/wiki/2020_coronavirus_pandemic_on_cruise_ships, visit on
534 01/04/2020
- 535 15. Chinazzi, M. *et al.* The effect of travel restrictions on the spread of the 2019 novel
536 coronavirus (2019-nCoV) outbreak *Science* **368**, 395–400 doi:
537 10.1126/science.aba9757_(2020).
- 538 16. Imai, N. *et al.* Report 3(imperial): Transmissibility of 2019-nCoV.
539 [https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-transmissibility-25-01-2020.pdf)
540 [fellowships/Imperial-College-COVID19-transmissibility-25-01-2020.pdf](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-transmissibility-25-01-2020.pdf) (2020)
- 541 17. Li, Q. *et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus-
542 infected pneumonia. *N Engl J Med.* **382**, 1199–1207
543 <https://doi.org/10.1056/NEJMoa2001316> (2020).
- 544 18. Liu, T. *et al.* Transmission dynamics of 2019 novel coronavirus (2019-nCoV).
545 *bioRxiv.* (published online Jan 26.) (preprint). doi: 10.2139/ssrn.3526307 (2020)
- 546 19. Riou, J. & Althaus, C. L. Pattern of early human-to-human transmission of Wuhan
547 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill.*
548 **25**(4), pii= 2000058. (2020)
- 549 20. Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential
550 domestic and international spread of the 2019-nCoV outbreak originating in Wuhan,
551 China: a modelling study. *Lancet* **395**, 689-697 [https://doi.org/10.1016/S0140-](https://doi.org/10.1016/S0140-6736(20)30260-9)
552 [6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9) (2020)
- 553 21. Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based
554 analysis. *Lancet Infect Dis* [https://doi.org/10.1016/S1473-3099\(20\)30243-7](https://doi.org/10.1016/S1473-3099(20)30243-7). (2020)
- 555 22. Linton, N. M. *et al.* Incubation period and other epidemiological characteristics of
556 2019 novel coronavirus infections with right truncation: A statistical analysis of
557 publicly available case data. *J Clin Med* **17**, 9(2). pii:E538. Doi:10.3390/jcm9020538
558 (2020).

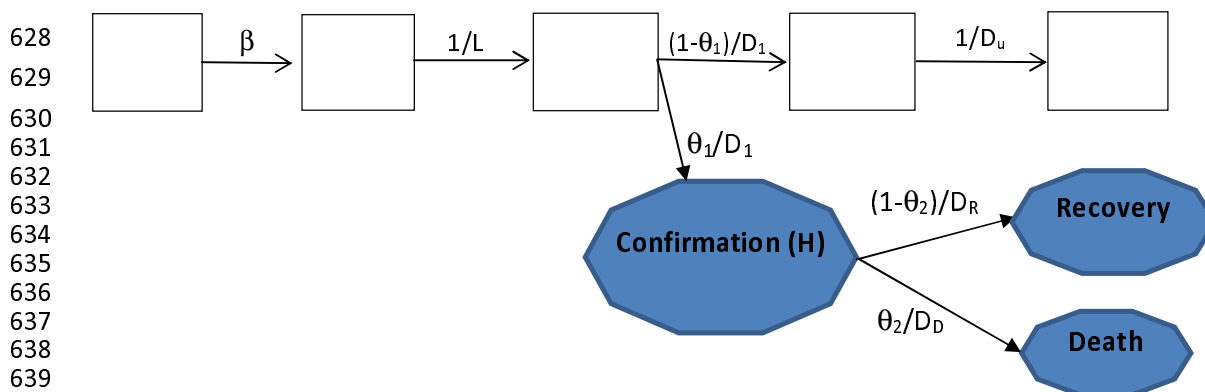
- 559 23. Sanche, S., Lin, Y. T., Xu, C., Romero-Severson, E., Hengartner, N. & Ke, R. The
560 Novel coronavirus, 2019-nCoV, is highly contagious and more infectious than
561 initially estimated medRxiv preprint doi:
562 <https://doi.org/10.1101/2020.02.07.20021154> (2020)
- 563 24. Xu, X. W. *et al.* Clinical findings in a group of patients infected with the 2019 novel
564 coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series *BMJ*
565 368:m606 <http://dx.doi.org/10.1136/bmj.m606> (2020)
- 566 25. Zhang, J. *et al.* Evolving epidemiology and transmission dynamics of coronavirus
567 disease 2019 outside Hubei province, China: a descriptive and modelling study.
568 *Lancet Infect Dis* **20**(7),793–802 [https://doi.org/10.1016/S1473-3099\(20\)30230-9](https://doi.org/10.1016/S1473-3099(20)30230-9)
569 (2020)
- 570 26. Jing, Q.-L. *et al.* Household secondary attack rate of COVID-19 and associated
571 determinants in Guangzhou, China: a retrospective cohort study. *Lancet Infect Dis*
572 [https://doi.org/10.1016/S1473-3099\(20\)30471-0](https://doi.org/10.1016/S1473-3099(20)30471-0) (2020).
- 573 27. Siordia Jr, J. A. Epidemiology and clinical features of COVID-19: A review of
574 current literature. *J Clin Virol.* 127(2020)104357. (2020).
- 575 28. Deng, X. *et al.* Case fatality risk of the first pandemic wave of novel coronavirus
576 disease 2019 (COVID-19) in China. *Clin Infect Dis.* Doi:10.1093/cid/ciaa578 (2020)
- 577 29. Wu, J. T. *et al.* Estimating clinical severity of COVID-19 from the transmission
578 dynamics in Wuhan, China. *Nat Med* <https://doi.org/10.1038/s41591-020-0822-7>
579 (2020).
- 580 30. Russell, T. W. *et al.* Estimating the infection and case fatality ratio for coronavirus
581 disease (COVID-19) using age-adjusted data from the outbreak on the Diamond
582 Princess cruise ship, February 2020. Volume 25, Issue 12, 26/Mar/2020 (2020)
- 583 31. <http://www.nhc.gov.cn/xcs/yqtb/202004/6f8eb06d959f4ab7b56fe03236920be1.shtml>
- 584 32. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19
585 *Nat Med* <https://doi.org/10.1038/s41591-020-0869-5> (2020)
- 586 33. Savvides, C. & Siegel, R. Asymptomatic and presymptomatic transmission of SARS-
587 CoV-2: A systematic review. medRxiv preprint doi:
588 <https://doi.org/10.1101/2020.06.11.20129072> (2020)
- 589 34. Xu, X. *et al.* Seroprevalence of immunoglobulin M and G antibodies against SARS-
590 CoV-2 in China. *Nat Med* 2020. doi: 10.1038/s41591-020-0949-6. (2020).
- 591 35. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and*
592 *Control.* (Oxford University Press, Oxford, 1991)

- 593 36. Vynnycky, E. & White, R. G. *An Introduction to Infectious Disease Modelling*.
594 (Oxford University Press, 2010)
- 595 37. Ferguson, N. M. *et al.* Impact of non-pharmaceutical interventions (NPIs) to reduce
596 COVID19 mortality and healthcare demand 16 March 2020 Imperial College
597 modelling report [https://www.imperial.ac.uk/media/imperial-](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf)
598 [college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf)
599 [modelling-16-03-2020.pdf](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf) (2020)
- 600 38. Wallinga, J. & Teunis, P. Different epidemic curves for severe acute respiratory
601 syndrome reveal similar impacts of control measures. *Am. J Epidemiol* **160**, 509–516
602 (2004).
- 603 39. Jia, N. *et al.* Case fatality of SARS in mainland China and associated risk factors.
604 *Trop. Med Int Health* **14**(suppl 1), 21–27. doi: 10.1111/j.1365-3156.2008.02147.x.
605 (2009)
- 606 40. Zhang, X.-S. *et al.* Estimating and modelling the transmissibility of Middle East
607 Respiratory Syndrome CoronaVirus during the 2015 outbreak in the Republic of
608 Korea. *Influenza Other Respir Viruses*, **11**(5), 434–444. Doi: 10.1111/irv.12467.
609 (2017)
- 610 41. Oh, M. D. *et al.* Middle East respiratory syndrome: what we learned from the 2015
611 outbreak in the Republic of Korea. *Korean J Inter Med*. **232**(2), 233–246 (2018).
- 612 42. Birrell, P. *et al.* Forecasting the 2017/2018 seasonal influenza epidemic in England
613 using multiple dynamic transmission models: a Case Study. *BMC Public Health* Doi:
614 10.21203/rs.2.16030/v2 (2020)
- 615 43. Lauer, S. A. *et al.* The incubation period of coronavirus disease 2019 (COVID-19)
616 from publicly reported confirmed cases: estimation and application. *Ann Intern Med*
617 **172**, 577–582 (2020).
- 618 44. Yan, L. & Li, Y. Identification and treatment strategies for patients with COVID-19.
619 *New Medicine* **51**(3), 161–167. doi:10.3969/j.issn.0253-9802.2020.03.001 (2020)
- 620 45. Chinese National Health Commission release of 7th edition of case definitions
621 [http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989.s](http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989.shtml)
622 [html](http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989.shtml) Date: 2020 Date accessed: March 12, 2020(in Chinese).
- 623 46. Birrell, P. *et al.* Bayesian modelling to unmask and predict influenza A/H1N1pdm
624 dynamics in London. *Proc Natl Acad Sci USA*. **108**, 18238–18243 (2011).

625

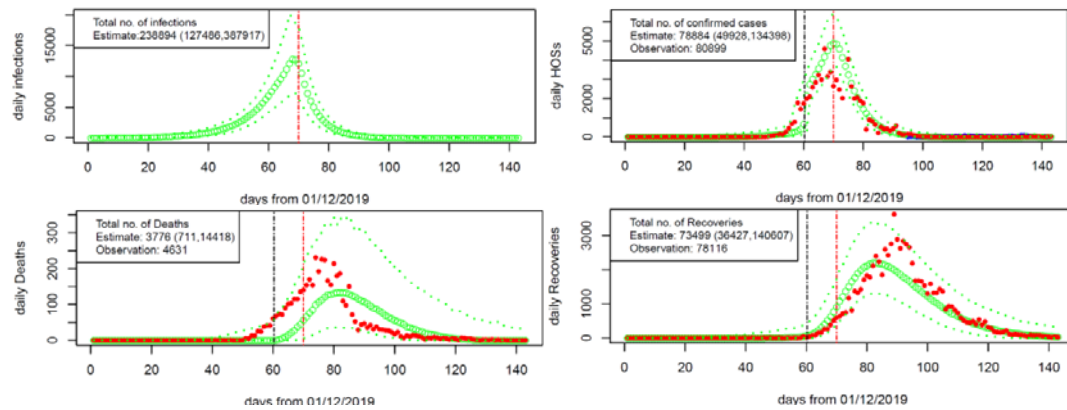
626 **Figures**

627



645 Figure 1 Flow chart of synthesis model: transmission dynamics and disease reporting
646 processes. The five rectangle boxes represent the hidden transmission dynamics process and
647 the three shaded polygons represent the quantities upon which observations were made.

648



649

650

651

652 Figure 2. Epidemic curves from 1st December 2019 to 21st April 2020. Model prediction of
653 infections (top left panel) and model fitting to the observations of confirmed cases (top right
654 panel), deaths (bottom left panel) and recoveries (bottom right panel) are shown. The model
655 predictions are obtained under the redistribution of 1290 death added on 17th April 2020 to
656 the period before 20th February 2020 in accordance to the daily number of deaths reported
657 before 17th April 2020. The green dotted lines represent the model predictions (large green
658 circles for median and thin green dotted lines for lower and upper levels of 95% confidence
659 interval). The red points are the observed data. The red vertical line denotes the estimate of
660 break point in transmission rate and the black vertical line the break point in detection and
661 reporting rate. In top right panel, the blue points represent daily number of imported cases.
662 Note that the daily number of confirmed cases (15152) on 12th February 2020 (day 74) is
663 beyond the range shown in the top right panel.

664

665

666 **Tables**

667

668 **Table 1 Timeline of outbreak in mainland China**

date	Event
November 2019	Several pneumoniae of unknown aetiology were discovered in Wuhan city, Hubei province, China
01/12/2019	Symptom onset of the first case reported
31/12/2019	An alert was issued by the Wuhan Municipal Health Commission and a rapid response team was sent to Wuhan by Chinese Centre for Disease Control and Prevention (China CDC) The WHO China Country Office was informed of 27 cases of pneumonia of unknown aetiology detected in Wuhan city, Hubei Province of China.
01/01/2020	Wuhan's Huanan Seafood Wholesale Market, the most probable index source of zoonotic COVID-19 infection, was shut down and disinfected China emergency response team was constructed.
02/01/2020	The first 41 confirmed cases were identified as laboratory-confirmed COVID-19 in Wuhan city
03/01/2020	A total of 44 patients with pneumonia of unknown aetiology have been reported to World Health Organisation (WHO) by Chinese authorities. Chinese National Health Commission (CNHC) issued Diagnosis and Treatment plan for pneumonia of unknown aetiology
04/01/2020	One exported case from Wuhan city into Weng Zhou, Zhejiang Province, China
05/01/2020	A total of 59 patients with pneumonia of unknown aetiology have been reported to WHO by Chinese authorities
07/01/2020	The causative pathogen was identified as a novel coronavirus and genomic characterisation and test method development ensued.
12/01/2020	Notification of the novel coronavirus and its sequence was made to WHO
13/01/2020	First exported case was reported in Thailand
17/01/2020	A total of 62 patients with pneumonia of unknown aetiology have been reported
19/01/2020	A total of 198 patients with pneumonia of unknown aetiology have been reported
20/01/2020	Person-to-person transmission was confirmed and announced in China. China's "National Infectious Disease Law" was amended to make COVID-19 a class B notifiable disease and its "Frontier Health and Quarantine Law" was amended to support the COVID-19 outbreak response effort. By doing this, Chinese law required all cases to be immediately reported to China's Infectious Disease Information System. CNHC started daily situation report of COVID-19 on their official website.

23/01/2020	Lockdown of the epicentre Wuhan city, limiting mobility of people in and out of Wuhan, and then the control measures expanded quickly to the neighbouring areas
25/01/2020	The Chinese government raised the response level of the “Preparedness and Response Plan for Novel Infection Disease of Public Health Significance” to the Emergency level, based on the assessment that the risk of health impact caused by COVID-19 on the local population is high and imminent. Chinese government announced its highest-level commitment and mobilized all sectors to respond to the epidemic and prevent further spread of COVID-19
31/01/2020	WHO announced the outbreak of novel coronavirus as a “public health emergency of international concern”
12/02/2020	Definition of confirmed cases for Hubei province was changed to include clinically diagnosed cases or PRC tested positive. The extra high daily number of cases 15152 was reported from Hubei province. On 29/02/2020 this changed back to “clinically diagnosed plus PCR tested positive “
11/03/2020	WHO assessed COVID-19 as a pandemic
17/04/2020	CNHC raised the total death number in Wuhan city from 2579 to 3869

669

670

671

672

Table 2 Estimates of model parameters for the COVID-19 outbreak within mainland China

The 1290 deaths within Hubei province added on 17th April 2020 were distributed over either the period before 17th April (i.e., the whole period) or before 20th February 2020 in proportional to the daily number of deaths reported before 17th April 2020. The relative infectivity of undetected cases $\xi = 1$ (i.e., both undetected and confirmed infections are of the same infectivity).

Name	Definition	Prior	Posterior	
			Before 20 February	Whole period
τ_β	Break point in transmission rate	U[60,143]*	70.0[67.2,71.4]	69.5[66.3,71.7]
β_a	Transmission rate before the effective control measures from at day τ_β	U[0.020,1.00]	0.424[0.267,0.584]	0.412[0.236,0.607]
β_b	Transmission rate after the control measures	U[0.001,0.20]	0.008[0.001,0.030]	0.007[0.001,0.027]
$R_{0,1}$	Reproduction number before τ_β	–	2.23[1.86,3.22]	2.25[1.65,3.31]
R_{0_con}	R_0 due to confirmed cases	–	0.072[0.027,0.207]	0.076[0.016,0.213]
R_{0_und}	R_0 due to undetected cases	–	2.14[1.79,3.14]	2.14[1.62,3.21]
$R_{0,2}$	Reproduction number after τ_β	–	0.035[0.007,0.103]	0.033[0.007,0.098]
I_0	Initial number of infectious people on 1/12/2019	U[1,500]	39.3[13.9,72.5]	28.0[11.5,57.2]
	Initial total number of people carrying the virus on 1/12/2019	–	184.9[51.2,354.6]	106.8[47.1,247.0]
τ_θ	Break point in reporting rate	U[40,62]*	60.3[56.0,61.9]	59.7[55.1,61.9]
$\theta_{1,a}$	Reporting/hospitalization rate before τ_θ	U[1%,30%]	6.89%[3.47%,14.61%]	6.54%[2.18%,14.92%]
$\theta_{1,b}$	Reporting/hospitalization rate after τ_θ	U[30%,100%]	41.53%[30.60%,65.06%]	38.64%[30.42%,63.30%]
θ_2	confirmed case fatality rate (cCFR)	U[0.5%,50%]	4.41% [3.65%,5.30%]	4.38%[3.77%,5.10%]

D_1	Infectious period before hospitalization	U[2.0,10]	2.43[2.02,3.64] days	2.61[2.03,4.08] days
D_1+D_u	Infectious period of undetected infections	U[3.0,25.0]	5.44[3.37,12.62] days	5.52[3.42,12.28] days
η^{HOS}	Dispersion parameter for reported cases	U[1.01,1500]	78.2[51.2,127.9]	75.3[49.7,124.2]
η^{Death}	Dispersion parameter for deaths	U[1.01,5000]	39.7[28.2,57.1]	25.9[18.2,37.4]
η^{Recovery}	Dispersion parameter for recoveries	U[1.01,1500]	121.3[83.7,180.7]	130.1[88.0,196.2]
IFR_1	Infection fatality rate ($\theta_{1,a}, \theta_2$) before τ_0	–	0.30%[0.15%,0.66%]	0.29%[0.10%,0.66%]
IFR_2	Infection fatality rate ($\theta_{1,b}, \theta_2$) after τ_0	–	1.81%[1.26%,3.03%]	1.71%[1.26%,2.82%]

*: The epidemic was assumed to start from 1st Dec 2019¹

Observation: the two ways of distributing 1290 deaths added on 17th April 2020 by Chinese National Health Commission are nearly the same for estimating model parameters except the dispersion parameter η^{Death} changing from 39.7 (before 20th February) to 25.9 (the whole period of the outbreak).

Table 3 Impact of start time of control measures during the outbreak in China. Here the total numbers of different cases from 1st December 2019 to 21st April 2020 were shown under different start dates of control measures.

Start date	Number of infections	Number of confirmed cases	Number of deaths
2 nd Jan 2020	27036[8296, 61744]	8556[2433, 25656]	129[0, 3914]
9 th Jan 2020	56037[22404, 108590]	18296[7566, 42499]	527[0, 6042]
16 th Jan 2020	115065[55009, 197000]	38126[20377, 72772]	1512[90, 9279]
23 rd Jan 2020	238894[127486, 387917]	78884[49928, 134398]	3776[711, 14418]
30 th Jan 2020*	474519[271147, 730166]	161119[111010,241228]	8292[2578, 22777]
6 th Feb 2020*	957486[564094, 1489883]	326008[231093,482038]	17689[7580, 37699]
13 th Feb 2020*	1926104[1145812,3128808]	656380[458844,1009816]	36314[18731,66699]

* Noted: As from 25th January 2020, China will start its Chinese New Year holidays, the contacts between people become much higher than the other periods of the year, and hence the estimated numbers of cases here are only conservative.

Table 4 Definitions of model compartments

Name	Definition
Transmission dynamics	
$S(t)$	Number of susceptible people at time t
$E(t)$	Number of exposed (infected but not yet infectious) people at time t
$I_1(t)$	Number of all infectious people at time t who have not yet been detected
$I_u(t)$	Number of infectious people at time t who have not been detected and will remain undetected
Disease reporting	
$H(t)$	Number of people who were hospitalized/reported due to COVID-19 at time t
$Death_0(t)$	Number of people in the “Death_0” compartment (between H and $Death$) at time t .
$Death(t)$	Number of people who have died from COVID-19 at time t
$Recovery_0(t)$	Number of people in the “Recovery_0” compartment (between H and $Recovery$) at time t
$Recovery(t)$	Number of recovered people at time t