

1 Prof. Jin Wang (<http://orcid.org/0000-0002-0062-2489>)

2

3 **Article type:** Original Articles

4

5 **Mathematical models for devising the optimal SARS-CoV-2 strategy**
6 **for eradication in China, South Korea, and Italy**

7

8 Shuo Jiang^{1,#}, Qiuyue Li^{1,#}, Chaoqun Li^{1,#}, Shanshan Liu¹, Xiaomeng He¹, Tao
9 Wang², Hua Li³, Christopher Corpe⁴, Xiaoyan Zhang¹, Jianqing Xu¹, Jin Wang^{1,*}

10

11 ¹Shanghai Public Health Clinical Center, Fudan University, 2901 Caolang Road,
12 Jinshan District, Shanghai 201508, P.R. China; ²Wuhan Academy of Social Science,
13 Wuhan, Hubei, China; ³State Key Laboratory for Oncogenes and Bio-ID Center, School
14 of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China; ⁴King's
15 College London, London, Nutritional Science Department, 150 Stamford Street,
16 Waterloo, London, SE19NH, United Kingdom.

17

18 **Running Title:** Mathematical models for devising SARS-CoV-2 eradication

19

20

21

22 ***Correspondence should be addressed to:**

23 Jin Wang, MD. & Ph.D.

24 Shanghai Public Health Clinical Center

25 Fudan University

26 2901 Caolang Road, Jinshan District

27 Shanghai 201508, China

28 Ph: 86-21-57036495

29 Fax: 86-21-57247094

30 Email: wjincityu@yahoo.com

31

32 # These authors contributed equally to this work

33 **The email addresses of all authors:**

- 34 Shuo Jiang: js951227@163.com
- 35 Qiuyue Li: liqiuyuezh@126.com
- 36 Chaoqun Li: cqli@sibs.ac.cn
- 37 Shanshan Liu: sfphbzv@qq.com
- 38 Xiaomeng He: 2215746263@qq.com
- 39 Tao Wang: wangtao69@sina.com
- 40 Hua Li: kaiikaixin@situ.edu.cn
- 41 Christopher Corpe: christopher.corpe@kcl.ac.uk
- 42 Xiaoyan Zhang: 18801119270@163.com
- 43 Jianqing Xu: 13791795950@163.com
- 44 Jin Wang: wjincityu@yahoo.com
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55 **ABSTRACT**

56 **Background:** Coronavirus disease 2019 (COVID-19), which is caused by severe acute
57 respiratory syndrome coronavirus 2 (SARS-CoV-2), spreads rapidly and has attracted
58 worldwide attention.

59 **Methods:** To improve the forecast accuracy and investigate the spread of SARS-CoV-
60 2, we constructed four mathematical models to numerically estimate the spread of
61 SARS-CoV-2 and the efficacy of eradication strategies.

62 **Results:** Using the Susceptible-Exposed-Infected-Removed (SEIR) model, and
63 including measures such as city closures and extended leave policies implemented by
64 the Chinese government that effectively reduced the β value, we estimated that the β
65 value and basic transmission number, R_0 , of SARS-CoV-2 was 0.476/6.66 in Wuhan,
66 0.359/5.03 in Korea, and 0.400/5.60 in Italy. Considering medicine and vaccines, an
67 advanced model demonstrated that the emergence of vaccines would greatly slow the
68 spread of the virus. Our model predicted that 100,000 people would become infected
69 assuming that the isolation rate α in Wuhan was 0.30. If quarantine measures were taken
70 from March 10, 2020, and the quarantine rate of α was also 0.3, then the final number
71 of infected people was predicted to be 11,426 in South Korea and 147,142 in Italy.

72 **Conclusions:** Our mathematical models indicate that SARS-CoV-2 eradication
73 depends on systematic planning, effective hospital isolation, and SARS-CoV-2
74 vaccination, and some measures including city closures and leave policies should be
75 implemented to ensure SARS-CoV-2 eradication.

76 **Keywords:** COVID-19; SARS-CoV-2; Mathematical models; Hospital isolation.

77 **Introduction**

78 The outbreak of COVID-19 pneumonia in Wuhan, caused by the novel coronavirus
79 SARS-CoV-2, has drawn tremendous attention around the world (1). The ongoing
80 COVID-19 outbreak resulted in more than 16,171,000 SARS-CoV-2 infections and
81 over 647,350 deaths worldwide by July 25, 2020 (2). SARS-CoV-2 has never been
82 found in humans before and may not be as virulent as severe acute respiratory syndrome
83 (SARS), but in humans, it is highly infectious.

84 Coronaviruses (CoVs) are pathogens that can infect the respiratory, gastrointestinal,
85 hepatic and central nervous systems of humans, livestock, birds, bats, mice and other
86 wild animals (3, 4). Regarding the outbreaks of SARS in 2002 and Middle East
87 respiratory syndrome (MERS) in 2012, the possibility of SARS-CoV and MERS-CoV
88 transmission from animals to humans has been suggested (5, 6). However, there have
89 been no effective strategies, including therapeutics and vaccines, identified and the best
90 approach to deal with severe CoV infections is to control the source of infection, using
91 early diagnosis, isolation, and treatments and the timely dissemination of epidemic-
92 related information to avoid unnecessary panic. Thus, surveillance and outbreak
93 response management systems as a framework for COVID-19 outbreak modeling are
94 urgently needed to control SARS-CoV-2 outbreaks worldwide.

95 Mathematical modeling plays an important role in understanding the complexities of
96 infectious diseases and their containment (7) because it can rapidly meet the need for
97 assessing the potential long-term impact of such diseases and offer strategies for the
98 evaluation and prediction of the effect of possible interventions, even when available
99 data are limited. Typical examples are the abundance of early models for HIV (8),

100 pandemic influenza (9, 10), bovine spongiform encephalopathy (BSE) (11), and
101 Creutzfeldt–Jakob disease (CJD) (12). Since the outbreak of SARS, mathematical
102 models have been published (13, 14), and policy makers have learned how models can
103 help support their decision making. Based on the SARS-CoV-2 full genomic sequence
104 data released on January 22, 2020 (15), that showed that SARS-CoV-2 has a more than
105 82% identical genome to those of SARS-CoV and bat SARS-like coronavirus (SL-CoV)
106 (16), COVID-19 can now be tracked in the population, and rapid and accurate
107 mathematical models can aid epidemiologic monitoring.

108 To improve forecast accuracy and investigate the spread and eradication pathways of
109 SARS-CoV-2, a mathematical model needs to take into consideration patient
110 improvements, personal protection strategies, regulation implementation, and other
111 contributing factors. SIR (Susceptible-Infected-Removed) models the behavior of the
112 early spread of infectious diseases and is generally used for measles, mumps, rubella
113 and other infectious diseases (17). Susceptible-Exposed-Infected-Removed (SEIR)
114 model is similar to SIR, with the variables (S, I, and R) representing the number of
115 people in each compartment at a particular time, but the incubation period has been
116 added in SEIR so that it is more applicable to infectious diseases with a certain
117 incubation period (18). By considering the characteristics of SARS-CoV-2-its spread
118 trends and local condition constraints-and economic optimization, we aimed to develop
119 mathematical models that would provide the optimal COVID-19 eradication plan that
120 was sensible and feasible.

121 **Methods**

122 **Characteristics of SARS-CoV-2.** To ensure the feasibility and usefulness of our
123 approach and before constructing our models, the important attributes of SARS-CoV-2
124 were summarized. (1) Origin: according to the current etiological research, the natural
125 host may be bats; the intermediate host is currently unknown. Most of the original
126 patients were geographically linked to the Huanan Seafood Wholesale Market (19). (2)
127 Transmission: the route of SARS-CoV-2 infection is through direct, aerosol and contact
128 transmission (20). The CoV causing the current outbreak is different from the human
129 coronaviruses previously identified, and the common ancestor of SARS/SARS-like
130 coronaviruses is a virus similar to HKU9-1. The incubation period of SARS-CoV-2 is
131 1-14 days (4, 21-23).

132

133 **Assumptions of the model for SARS-CoV-2 analysis.** Basic assumptions are that the
134 outbreak began in Wuhan. After the outbreak, the city was closed on January 23, 2020,
135 the remaining total population of Wuhan was 9,000,000, and we assumed it remained
136 unchanged. The basic assumptions for South Korea and Italy are as follows: (1) The
137 date that the first case occurred in China was Dec 08, 2019, and the date that the first
138 case was diagnosed was Jan 11, 2020, with a difference of 34 days. We assumed that
139 the actual occurrence of the first case in other countries was 30 days earlier than the
140 first announced diagnosis. (2) In 2018, South Korea had a population of 51,640,000,
141 and Italy had a population of 60,430,000.

142

143 **Four models for the numerical prediction of the spread of SARS-CoV-2.** The

144 datasets generated during and/or analyzed during the study of the model for SARS-
145 CoV-2 are available on the following websites: ([https://www.msn.com/en-](https://www.msn.com/en-gb/weather/other/coronavirus-outbreak-who-names)
146 [gb/weather/other/coronavirus-outbreak-who-names](https://www.msn.com/en-gb/weather/other/coronavirus-outbreak-who-names);
147 <https://ncov.dxy.cn/ncovh5/view/pneumonia?from=timeline&isappinstalled=0>; and [http://www.zq-](http://www.zq-ai.com/#/fe/xgfybigdata)
148 [ai.com/#/fe/xgfybigdata](http://www.zq-ai.com/#/fe/xgfybigdata)). All information on the basic considerations and assumptions of
149 the SIR and SEIR models are included in the Methods section, and supplementary files
150 provide more information about our models. All related parameters are shown in Table
151 1. For the SIR model, the sensitivity of the β parameter (the probability of transfer from
152 the susceptible state to the resistant state) was estimated. There was $N \approx S$ in the early
153 stage of the epidemic, so:

$$154 \quad \frac{dI}{dt} = \beta \frac{IS}{N} - \gamma I \approx (\beta - \gamma)I.$$

155 Hence:

$$156 \quad I(t) = e^{(\beta - \gamma)t}.$$

157

158 For the SEIR model, we divided the total population into four groups, including the
159 Susceptible (S), Incubation (E), Infected (I) and Removed (R) groups, and modeled
160 SARS-CoV-2 infection as related to the number of instances of contact between
161 susceptible and infected people multiplied by the infection rate β . We can express the
162 rate of change of the susceptible group as:

$$163 \quad \frac{dS}{dt} = -\beta S \frac{I}{N}.$$

164 Note that σ represents the outflow rate of the incubation group to the infected group.

165 The rate of change of the incubation group can be expressed as:

166
$$\frac{dE}{dt} = \beta S \frac{I}{N} - \sigma E.$$

167 As γ represents the outflow rate of the infected group to the removed group, we obtain
 168 the following equation:

169
$$\frac{dI}{dt} = \sigma E - \gamma I.$$

170 We can obtain the rate of change of the removed group. Its inflow is from the infected
 171 group at the rate of γ , and it has no outflow since it is the end of the system. Hence, we
 172 obtain the following equation:

173
$$\frac{dR}{dt} = \gamma I.$$

174 The whole process set out above can be displayed in the flow chart below.



175
 176 Furthermore, we made some modifications to the basic model by including hospital
 177 isolations. The outflow rate of the infected group will change since part of this group
 178 will be moved to the hospital isolation group at the rate of α , and the other part will
 179 move to the removed group at the rate of γ . Hence, we have:

180
$$\frac{dI}{dt} = \sigma E - \gamma(1 - \alpha)I - \alpha I.$$

181 For the new hospital isolation group, its inflow is from the infected group at the rate of
 182 α , and its outflow is to the removed group at the rate of ω . We can obtain the rate of
 183 change of this group:

184

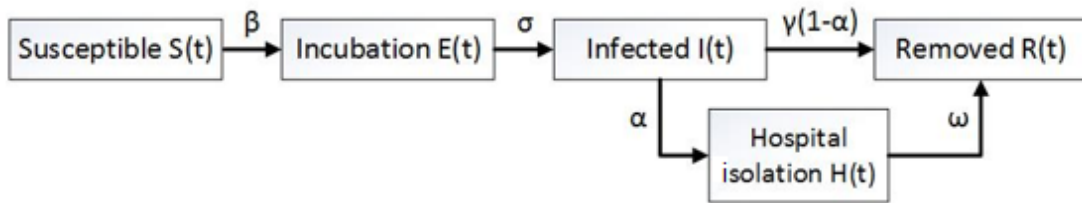
$$\frac{dH}{dt} = \alpha I - \omega H$$

185 Similarly, we made some modifications to the rate of change of the removed group:

186

$$\frac{dR}{dt} = \gamma(1-\alpha)I + \omega H$$

187 We next define α as the isolation rate, which is the rate of people moving from the
 188 infected group to the hospital isolated group, and ω as the outflow rate of the infected
 189 group to the hospital isolated group. Let H be the number of people in this group. The
 190 advanced model introduced above can be displayed in the flow chart below.



191

192 Finally, we incorporated effective medications and vaccines into this advanced model.

193 We define M as the number of people in the immunity group, I_E as the number of people
 194 in the early infection group, and I_L as the number of people in the advanced infection
 195 group. We present the system of differential equations for our advanced model
 196 considering medication and vaccination as follows:

$$\begin{aligned} \frac{dS}{dt} &= -\beta S \frac{I_E + I_L}{N} - \theta S \\ \frac{dE}{dt} &= \beta S \frac{I_E + I_L}{N} - \sigma E \\ \frac{dI_E}{dt} &= \sigma E - \gamma_E(1-\alpha)I_E - \alpha I_E \\ \frac{dI_L}{dt} &= \gamma_E(1-\alpha)I_E - \gamma_L I_L \end{aligned}$$

197

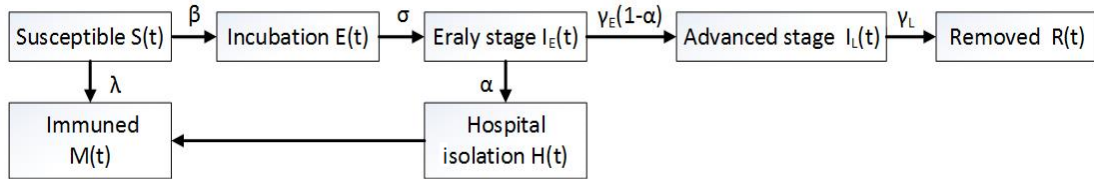
$$\frac{dH}{dt} = \alpha I_E - \omega H$$

$$\frac{dM}{dt} = \theta S + \omega H$$

$$\frac{dR}{dt} = \gamma_L I_L$$

198

199 The above model can be displayed in the flow chart shown below.



200

201

202 Results

203 Disease review of COVID-19

204 The first confirmed case of COVID-19 was on Dec 8, 2019, and the Chinese
 205 government began to continuously report cases on Jan 11, 2020. After the early outbreak
 206 of SARS-CoV-2, as of March 4, 2020, 80,424 cases had been confirmed in China,
 207 including 2,984 deaths (24); outside China, 10,566 cases were reported in 72 countries,
 208 and a large percentage of those cases occurred in South Korea and Italy (2). Figure 1
 209 shows the detailed epidemic trend until July 25, 2020, and includes the number of
 210 confirmed cases in Wuhan (Fig. 1a); Fig. 1b shows the number of cumulative treated
 211 and recovered patients and the number of deaths in Wuhan. Detailed epidemic trends
 212 that include the number of confirmed cases in South Korea (Fig. 1c) and Italy (Fig. 1d)
 213 are also shown. The severity of the situation means an accurate mathematical model
 214 needs to be established to predict epidemic trends in order to take positive and effective
 215 countermeasures. The actions these affected countries take today will be the difference

216 between a handful and a larger cluster of cases.

217 **Effect of the SARS-CoV-2 infection rate on eradicating SARS-CoV-2**

218 Under the optimistic estimation, the number of cases reported in only Wuhan is counted
219 as the confirmed cases recorded in Wuhan. In the modeling, according to the pessimistic
220 estimate, it is assumed that the infection cases found in other areas are all from Wuhan.
221 By iterating on different β values, the β value with the minimum variance is selected as
222 the fitted β value:

$$223 \min_{\beta} \sum_{i=1}^n (f_i - d_i)^2$$

224 where f_i denotes the data i on the fitted curve, d_i denotes the real data and n denotes
225 the data size. It can be concluded that the β value in the optimistic case is 0.213 (Fig.
226 2a). Therefore, the fitting of infection cases under optimistic conditions is shown in Fig.
227 2b. The pessimistic estimation indicates that the β value in the pessimistic case is 0.236
228 (Fig. 2c). Therefore, the fitting of infection cases under pessimistic conditions is shown
229 in Fig. 2d. It can be concluded that the value of the resistance parameter (the probability
230 of transfer from one susceptible individual to another) is between 0.213 and 0.236. We
231 can also estimate the basic reproduction number (R_0) of the novel coronavirus as:

$$232 R_0 = \frac{\beta}{\gamma} = \frac{0.213 \sim 0.236}{1/14} = 2.98 \sim 3.30.$$

233 The R_0 of infections is the average number of people infected with an infectious disease
234 that can spread to other people without intervention and without immunity. The larger
235 the R_0 number is, the more difficult it is to control the epidemic. In this model, we
236 estimated that the basic transmission R_0 of SARS-CoV-2 was between approximately
237 2.98 and 3.30.

238 **Effect of SARS-CoV-2 government control measures on eradicating SARS-CoV-2**

239 **(1) In Wuhan:** As of Feb 12, 2020 (day 67), a total of 19,558 cases were confirmed in
240 Wuhan. That is, $I(67) = 19,558$ can be solved as $\beta = 0.476$. In the SEIR model, the R_0
241 of the novel coronavirus is:

$$242 R_0 = \frac{\beta}{\gamma} = \frac{0.476}{1/14} = 6.66$$

243 It can be found that the disease outbreak starts to intensify around the 80th day (late
244 Feb) and reaches its peak around the 120th day (mid-April). The changing trend of each
245 group is shown in Fig. 2e. According to the actual number of cases in Wuhan, the
246 epidemic curve was fitted (Fig. 3a), and β values were evaluated. For government
247 control of the outbreak, under the SEIR model, we simulated that β was 0.476 (Fig. 3b).
248 According to the above β value, we estimated that approximately 3,200,000 people will
249 be infected in the whole city of Wuhan. When β decreases to 0.4, the peak infection rate
250 drops to 3,000,000, and when β further decreases to 0.3, the peak infection rate drops
251 to 2,500,000 (Fig. 3b).

252 **(2) In South Korea and Italy:** In the SEIR model, the basic reproduction number of
253 the novel coronavirus in South Korea and Italy is:

$$254 R_0(\text{South Korea}) = \frac{\beta}{\gamma} = \frac{0.359}{1/14} = 5.03;$$

$$255 R_0(\text{Italy}) = \frac{\beta}{\gamma} = \frac{0.400}{1/14} = 5.60.$$

256 According to the actual number of cases in Korea and Italy, the epidemic curve was
257 fitted (Fig. 3d, g), and β values were evaluated. For government control of the outbreak,
258 under the SEIR model, we simulated that β was 0.359 for South Korea (Fig. 3e) and
259 0.400 for Italy (Fig. 3h). According to the above β values, we estimated that

260 approximately 51,640,000 and 60,430,000 people would be infected in South Korea
261 and Italy, respectively. If the government implements some measures, such as city
262 closure policies and extended leave policies, it can effectively reduce the β value. In
263 South Korea, when β drops to 0.30 or 0.20, the peak infection rate drops to 13,974,030
264 or 9,313,317 (Fig. 3e). In Italy, when β decreases to 0.30 or 0.20, the peak infection rate
265 drops to 16,358,690 or 10,897,254 (Fig. 3h).

266 **Effect of hospital isolation on eradicating SARS-CoV-2**

267 Applying different α ratios, we can see that with the increase in the ratio of hospital
268 isolation, the number of infections at the highest point decreased, and the time of peak
269 occurrence was later. When not isolated, approximately 3,200,000 people eventually
270 become infected in Wuhan. When only 10% of the population is quarantined, only
271 1,200,000 become infected in Wuhan. When the quarantine rate is raised further, to 20%
272 or 30%, 400,000 or only 100,000 people, respectively, would become infected (Fig. 3c),
273 which is consistent with the current number of infected people in the report, verifying
274 that 30% of the population has been quarantined in Wuhan according to this modeling
275 analysis. Additionally, if quarantine measures were implemented from March 10, 2020,
276 and the quarantine rate of α was 0.3, the final number of infected people would be
277 11,426 in South Korea with a peak of 450 days (Fig. 3f) and 147,142 in Italy with a
278 peak of 405 days (Fig. 3i).

279 **Effect of the vaccination rate and time on eradicating SARS-CoV-2 in Wuhan**

280 Vaccination rate: Assuming the isolation rate alpha is 0.2, the changes in the number of
281 infected patients under different vaccination rates (theta) are shown in Fig. 4a. Without

282 a vaccine, at a 20% isolation rate, 400,000 people could be infected. If the vaccination
283 rate is 0.005, fewer than 20,000 people will become infected. Therefore, the emergence
284 of vaccines can greatly alleviate the spread of a virus. If a vaccine is developed within
285 two months of the outbreak (day 60) with a vaccination rate of 0.005, the changes after
286 that are shown in Fig. 4b; this implies that the SARS-CoV-2 will peak at day 150 and
287 only 15,000 people in Wuhan will be infected. However, it will take a few months to
288 develop a vaccine for SARS-CoV-2.

289

290 **Discussion**

291 To date, the COVID-19 epidemic is still in a phase of rapid dispersion worldwide, and
292 this epidemic represents a clear and ongoing global health threat. It is currently
293 uncertain whether it is possible to contain the continuing epidemic within China (25).

294 In the early stage of a SARS-CoV-2 outbreak, the medical conditions are not ideal, and
295 no effective measures have been taken. The number of infected people increases
296 exponentially, so it is particularly critical to control the value of β . In our model, the β
297 of South Korea is 0.359, which is the lowest among the three countries, followed by
298 Italy, whose β is 0.4. The β of both South Korea and Italy is lower than that of China.

299 In these three countries, the trend of R_0 was consistent with β . Thus, reducing the
300 exposure rate of β could significantly reduce the reproduction number. Next, we
301 estimated an R_0 of 6.66 in Wuhan, which represents a relatively higher value than those
302 computed so far (as of Mar 01, 2020). In several other mathematical models that have
303 been devised and released to date, R_0 varies from 1.30 to 6.47 (Table S1) (26-30). By

304 comparing the methodologies of the various investigations, these different reproduction
305 numbers reflect the dynamics of transmission, and the cases of COVID-19 fluctuate and
306 vary over time. In our models, we used these data (Jan 15, 2020 - Mar 01, 2020) to
307 estimate a relatively accurate R_0 , which was relatively reliable in our models.

308 We further took hospital isolation and SARS-CoV-2 drugs and vaccines into account in
309 our models. With other parameters unchanged, the greater the α for the isolation rate in
310 the hospital, the later the peak is reached, the smaller the peak size, and the more
311 effective it is to prevent the spread of the disease. The hospital pathway should
312 sufficiently isolate SARS-CoV-2-infected patients from other patients to decrease
313 infection, and more concern should be given about protecting doctors and nurses (31).

314 For the penetration of vaccines and drugs, the greater the parameter theta, the earlier
315 the peak would appear, the smaller the peak would be, and the lower the final total
316 number of infections would be. However, scientists have still not found effective
317 medications and vaccines for SARS-CoV-2 (32). Finally, our simulations have shown
318 that the containment outcome depends highly on the effectiveness of the intense control
319 effort now underway in China. Reducing the exposure rate of β and increasing the
320 isolation rate of α can significantly reduce the number of infected people. The
321 government should continue to tightly monitor the epidemic situation and must take
322 immediate measures against it, and this includes immediate isolation of newly infected
323 people and closures of cities with severe outbreaks, in case there are unexpected
324 outbreaks in the eradication process. Regional transmission is the root of the spread of
325 COVID-19. Local governments must have the responsibility to set a deadline for the

326 final eradication, and the SARS-CoV-2 epidemic in China revealed that China still
327 needs to strengthen the establishment of a rapid outbreak response strategy and health
328 policies. Finally, there were some limitations in our SIR or SEIR models of SARS-CoV-
329 2, such as we only used the average latency value during the latent period. We
330 recognized that some patients with COVID-19 have a longer latent period, while others
331 have a shorter latent period of only 4 days, which affects the basic reproduction number
332 (R_0) (33). In addition, we need to consider “superspreading events and superspreaders”,
333 which tend to occur at large gatherings with close contact (34), transmit infection to a
334 larger number of individuals than is typical by one individual (35), and affect R_0 . For
335 example, a man who later tested positive for COVID-19 visited several clubs in Seoul
336 and infected 170 new individuals after South Korea relaxed social distancing rules in
337 May. Moreover, the differences in case definitions and reporting measures, city closures
338 and leave policies, and viral testing would likely affect the basic reproduction number
339 and the secondary attack rate (SAR) (34, 36). It is important that the Italian government
340 implemented extraordinary measures to limit viral transmission in March 2020 and
341 minimized the likelihood that people were infected (37). Therefore, all those countries
342 that had beaten back the virus to low levels need to be especially vigilant for
343 superspreaders and superspreading events.

344

345 **Conclusions**

346 In summary, our mathematical model of SARS-CoV-2 infection can accurately predict
347 the incidence and number of cases as well as the peak and end times of the epidemic

348 and provide feasible solutions for future epidemic prevention and control, including
349 predicting future epidemic trends and providing a reference for effective control options.
350 Our results emphasized that effective SARS-CoV-2 eradication must involve active
351 cooperation between the government, pharmaceutical companies and hospital
352 organizations.

353

354 **Ethics approval and consent to participate**

355 Not Applicable.

356

357 **Consent for Publication**

358 Not Applicable.

359

360 **Availability of data and material**

361 The datasets generated during and/or analyzed during the current study are available on
362 the following websites: ([https://www.msn.com/en-gb/weather/other/coronavirus-outbreak-who-](https://www.msn.com/en-gb/weather/other/coronavirus-outbreak-who-names)
363 [names](https://www.msn.com/en-gb/weather/other/coronavirus-outbreak-who-names); <https://ncov.dxy.cn/ncovh5/view/pneumonia?from=timeline&isappinstalled=0>; and
364 <http://www.zq-ai.com/#/fe/xgfybigdata>).

365

366 **Competing interests**

367 The authors declare no competing financial interests.

368

369 **Funding**

370 The work is supported partially by a grant (2018ZX10302103-003) from the National
371 Special Research Program of China for Important Infectious Diseases and a grant from
372 the National Natural Science Foundation of China (81672383).

373

374 **Author Contributions**

375 Conceived and designed the experiments: J.W., and S. J. Constructed the mathematical
376 model and designed and carried out data analysis: S.J. Analyzed the data: S. J., C.L., X.
377 H., H. L., and T. W. Contributed analysis tools: S.L., C.C., X.Z. and J.X. Wrote the
378 manuscript: Q.L. and J.W. Revised the manuscript: J.W.

379

380 **Acknowledgments**

381 Not Applicable.

382

383 **References**

- 384 1. Coronavirus latest: WHO officially names disease COVID-19 [Available from:
385 [https://www.msn.com/en-gb/weather/other/coronavirus-outbreak-who-](https://www.msn.com/en-gb/weather/other/coronavirus-outbreak-who-names%20disease-covid-19/vi-BBZTC7q)
386 [names%20disease-covid-19/vi-BBZTC7q](https://www.msn.com/en-gb/weather/other/coronavirus-outbreak-who-names%20disease-covid-19/vi-BBZTC7q).
- 387 2. The Global COVID-19 Big Data Analysis Platform [Available from:
388 <http://www.zq-ai.com/#/fe/xgfybigdata>.
- 389 3. Wang L-F, Shi Z, Zhang S, Field H, Daszak P, Eaton BT. Review of bats and SARS.
390 Emerging infectious diseases. 2006;12(12):1834-40.
- 391 4. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, et al. Isolation and
392 characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature.
393 2013;503(7477):535-8.
- 394 5. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat Rev
395 Microbiol. 2019;17(3):181-92.
- 396 6. Cauchemez S, Van Kerkhove MD, Riley S, Donnelly CA, Fraser C, Ferguson NM.
397 Transmission scenarios for Middle East Respiratory Syndrome Coronavirus (MERS-
398 CoV) and how to tell them apart. Euro surveillance : bulletin European sur les maladies
399 transmissibles = European communicable disease bulletin. 2013;18(24).

- 400 7. Han XN, de Vlas SJ, Fang LQ, Feng D, Cao WC, Habbema JD. Mathematical
401 modelling of SARS and other infectious diseases in China: a review. *Trop Med Int*
402 *Health*. 2009;14 Suppl 1:92-100.
- 403 8. Anderson RM. The role of mathematical models in the study of HIV transmission
404 and the epidemiology of AIDS. *Journal of acquired immune deficiency syndromes*.
405 1988;1(3):241-56.
- 406 9. Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S, Meeyai A, et al.
407 Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature*.
408 2005;437(7056):209-14.
- 409 10. Longini IM, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings DAT,
410 et al. Containing pandemic influenza at the source. *Science (New York, NY)*.
411 2005;309(5737):1083-7.
- 412 11. Ferguson NM, Donnelly CA, Woolhouse ME, Anderson RM. The epidemiology of
413 BSE in cattle herds in Great Britain. II. Model construction and analysis of transmission
414 dynamics. *Philosophical transactions of the Royal Society of London Series B,*
415 *Biological sciences*. 1997;352(1355):803-38.
- 416 12. Bishop MT, Hart P, Aitchison L, Baybutt HN, Plinston C, Thomson V, et al.
417 Predicting susceptibility and incubation time of human-to-human transmission of vCJD.
418 *The Lancet Neurology*. 2006;5(5):393-8.
- 419 13. Wang W, Ruan S. Simulating the SARS outbreak in Beijing with limited data.
420 *Journal of theoretical biology*. 2004;227(3):369-79.
- 421 14. Bai Y, Jin Z. Prediction of SARS epidemic by BP neural networks with online
422 prediction strategy. *Chaos, Solitons & Fractals*. 2005;26(2):559-69.
- 423 15. National Genomics Data Center MP. Database Resources of the National Genomics
424 Data Center in 2020. *Nucleic acids research*. 2020;48(D1):D24-D33.
- 425 16. Hu B, Ge X, Wang L-F, Shi Z. Bat origin of human coronaviruses. *Virology journal*.
426 2015;12:221.
- 427 17. Chowell G, Sattenspiel L, Bansal S, Viboud C. Mathematical models to
428 characterize early epidemic growth: A review. *Phys Life Rev*. 2016;18:66-97.
- 429 18. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic
430 and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a
431 modelling study. *Lancet*. 2020;395(10225):689-97.
- 432 19. Chan JF-W, Yuan S, Kok K-H, To KK-W, Chu H, Yang J, et al. A familial cluster
433 of pneumonia associated with the 2019 novel coronavirus indicating person-to-person
434 transmission: a study of a family cluster. *Lancet (London, England)*.
435 2020;395(10223):514-23.
- 436 20. Liu J, Liao X, Qian S, Yuan J, Wang F, Liu Y, et al. Community Transmission of
437 Severe Acute Respiratory Syndrome Coronavirus 2, Shenzhen, China, 2020. *Emerging*
438 *infectious diseases*. 2020;26(6).
- 439 21. Letko M, Munster V. Functional assessment of cell entry and receptor usage for
440 lineage B β -coronaviruses, including 2019-nCoV. *bioRxiv : the preprint server for*
441 *biology*. 2020.
- 442 22. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. Discovery of a
443 novel coronavirus associated with the recent pneumonia outbreak in humans and its

444 potential bat origin. 2020.

445 23. Gralinski LE, Menachery VD. Return of the Coronavirus: 2019-nCoV. *Viruses*.
446 2020;12(2).

447 24. Real-time dynamics of COVID-19 in China [Available from:
448 <https://ncov.dxy.cn/ncovh5/view/pneumonia?from=timeline&isappinstalled=0>.
449 25. Xu S, Li Y. Beware of the second wave of COVID-19. (1474-547X (Electronic)).
450 26. Zhang C, Wang M. MRCA time and epidemic dynamics of the 2019 novel
451 coronavirus. 2020.

452 27. Riou J, Althaus CL. Pattern of early human-to-human transmission of Wuhan 2019
453 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro surveillance* :
454 bulletin European sur les maladies transmissibles = European communicable disease
455 bulletin. 2020;25(4).

456 28. Liu T, Hu J, Xiao J, He G, Kang M, Rong Z, et al. Time-varying transmission
457 dynamics of Novel Coronavirus Pneumonia in China. *bioRxiv*.
458 2020:2020.01.25.919787.

459 29. Majumder M, & Mandl, K. D. Early transmissibility assessment of a novel
460 coronavirus in Wuhan, China. *SSRN*. 2020.

461 30. Tang SY, Xiao YN, Peng ZH, Shen HB. [Prediction modeling with data fusion and
462 prevention strategy analysis for the COVID-19 outbreak]. *Zhonghua liu xing bing xue*
463 *za zhi* = *Zhonghua liuxingbingxue zazhi*. 2020;41(4):480-4.

464 31. Hui DS, I Azhar E, Madani TA, Ntoumi F, Kock R, Dar O, et al. The continuing
465 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019
466 novel coronavirus outbreak in Wuhan, China. *International journal of infectious*
467 *diseases* : *IJID* : official publication of the International Society for Infectious Diseases.
468 2020;91:264-6.

469 32. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global
470 health concern. *Lancet (London, England)*. 2020;395(10223):470-3.

471 33. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the
472 transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*.
473 2020;368(6493):860-8.

474 34. Liu Y, Eggo RM, Kucharski AJ. Secondary attack rate and superspreading events
475 for SARS-CoV-2. *Lancet*. 2020;395(10227):e47.

476 35. Al-Tawfiq JA, Rodriguez-Morales AJ. Super-spreading events and contribution to
477 transmission of MERS, SARS, and SARS-CoV-2 (COVID-19). *J Hosp Infect*.
478 2020;105(2):111-2.

479 36. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented
480 infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2).
481 *Science*. 2020;368(6490):489-93.

482 37. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? *Lancet*.
483 2020;395(10231):1225-8.

484

485

486 **Table 1.** Parameters of our mathematical models.

Parameters	Symbol	Description
Susceptible group	S	People who have no immunity against the disease. They are very likely to be infected by coming in direct contact with infected people.
Exposed group	E	People who have been infected but have not displayed any explicit symptoms. They do not transmit the virus to susceptible people.
Infected group	I	People in the infected group show explicit symptoms of SARS-CoV-2, and they can transmit the virus to susceptible people.
Removed group	R	The removed group includes people who have died of the disease or who have survived the disease. People who have survived the disease will obtain complete immunity against it.
Number	N	The total population.
Infection rate	β	The probability that a susceptible person will become ill after coming into contact with an infected person.
Outflow rate	σ	The outflow rate of the incubation group to the infected group.
Period	γ	The days from infected to removed.
Isolation rate	α	The rate of people moving from the infected group to the hospital isolated group.
Outflow rate	ω	The outflow rate of the infected group to the hospital isolated group.
Vaccination rate	θ	The percentage of susceptible people who receive a vaccination each day.
Number	H	The number of people in the hospital isolation group.
Number	I_E	The number of people in the early infection group.
Number	I_L	The number of people in the advanced infection group.
Period	γ_E	Period when patients in the early infected group display explicit symptoms of COVID-19 and can transmit the virus to susceptible people.
Period	γ_L	Period when patients in the advanced infected group display explicit symptoms of COVID-19 and can transmit the virus to susceptible people.
Number	M	The number of people in the immunity group.

487

488 **Figure Legends**

489 **Figure 1.** Review of COVID-19 in Wuhan, Republic of Korea, and Italy. (a) The
490 number of confirmed cases in Wuhan; (b) the number of cumulative cured patients and
491 deaths in Wuhan; and the number of confirmed cases in South Korea (c) and Italy as of
492 July 25, 2020 (d).

493

494 **Figure 2.** The SIR and SEIR models analyze the basic epidemic data from Wuhan: (a,
495 b) optimistic estimation and (c, d) pessimistic estimation. (a) The β value in the
496 optimistic estimation. (b) The infection cases in the optimistic estimation are reasonable.
497 (c) The β value in the pessimistic estimation. (d) The fitting of infection cases in the
498 pessimistic estimation. (e) Trends in each group of SEIR models for Wuhan.

499

500 **Figure 3.** The SIR and SEIR models were used to analyze the epidemic situation in
501 Wuhan. (a) The fitted curve (blue) for Wuhan, where red points represent the actual
502 number of cases; (b) SEIR model analysis for Wuhan; and (c) Hospital isolation
503 included in the model analysis for Wuhan. Hospital isolation model with different
504 isolation ratios showing that the peak of the outbreak occurred on day 450. (d) The
505 fitted curve (blue) for South Korea; red points represent the actual number of cases. (e)
506 SEIR model analysis for South Korea. (f) Hospital isolation model with different
507 isolation ratios included in the model analysis for South Korea. The peak of the outbreak
508 occurred on day 450. (g) The fitted curve (blue) for Italy; red points represent the actual
509 number of cases. (h) SEIR model analysis for Italy. (i) Hospital isolation model with

510 different isolation ratios included in the model analysis for Italy, and the peak of the
511 outbreak occurred on day 405.

512

513 **Figure 4.** Vaccination rates were included in the model analysis for Wuhan. (a) SEIR
514 model with different vaccination rates. (b) The changes of the SEIR model with
515 vaccination starting on day 60.