

1 **Case fatality risk of novel coronavirus diseases 2019 in China**

2 Xiaowei Deng, MSc^{1#}, Juan Yang, PhD^{1#}, Wei Wang, MSc^{1#}, Xiling Wang, PhD^{1#},
3 Jiaxin Zhou, BSc¹, Zhiyuan Chen, BSc¹, Jing Li, BSc¹, Yinzi Chen, BSc¹, Han Yan,
4 BSc¹, Juanjuan Zhang, PhD¹, Yongli Zhang, MSc², Yan Wang, MSc¹, Qi Qiu, MSc¹,
5 Hui Gong, BSc¹, Xianglin Wei, BSc¹, Lili Wang¹, Kaiyuan Sun, PhD³, Peng Wu, PhD⁴,
6 Marco Ajelli, PhD⁵, Benjamin J. Cowling, PhD⁴, Cecile Viboud, PhD³, Hongjie Yu,
7 PhD¹

8 #These authors contributed equally to this work.

9 Corresponding author to Prof. Hongjie Yu, yhj@fudan.edu.cn

10 **Affiliations:**

11 1. School of Public Health, Fudan University, Key Laboratory of Public Health
12 Safety, Ministry of Education, Shanghai, China

13 2. Savaid Medical School, University of Chinese Academy of Sciences, Beijing,
14 China

15 3. Division of International Epidemiology and Population Studies, Fogarty
16 International Center, National Institutes of Health, Bethesda, MD, USA

17 4. WHO Collaborating Centre for Infectious Disease Epidemiology and Control,
18 School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong,
19 Hong Kong Special Administrative Region, China

20 5. Bruno Kessler Foundation, Trento, Italy

21 **ABSTRACT**

22 **Objective** The outbreak of novel coronavirus disease 2019 (COVID-19) imposed
23 a substantial health burden in mainland China and remains a global epidemic
24 threat. Our objectives are to assess the case fatality risk (CFR) among COVID-19
25 patients detected in mainland China, stratified by clinical category and age group.

26 **Methods** We collected individual information on laboratory-confirmed COVID-19
27 cases from publicly available official sources from December 29, 2019 to
28 February 23, 2020. We explored the risk factors associated with mortality. We
29 used methods accounting for right-censoring and survival analyses to estimate
30 the CFR among detected cases.

31 **Results** Of 12,863 cases reported outside Hubei, we obtained individual records
32 for 9,651 cases, including 62 deaths and 1,449 discharged cases. The deceased
33 were significantly older than discharged cases (median age: 77 vs 39 years,
34 $p < 0.001$). 58% (36/62) were male. Older age (OR 1.18 per year; 95%CI: 1.14 to
35 1.22), being male (OR 2.02; 95%CI: 1.02 to 4.03), and being treated in less
36 developed economic regions (e.g., West and Northeast vs. East, OR 3.93; 95%CI:
37 1.74 to 8.85) were mortality risk factors. The estimated CFR was 0.89-1.24%
38 among all cases. The fatality risk among critical patients was 2-fold higher than
39 that among severe and critical patients, and 24-fold higher than that among
40 moderate, severe and critical patients.

41 **Conclusions** Our estimates of CFR based on laboratory-confirmed cases

42 ascertained outside of Hubei suggest that COVID-19 is not as severe as severe
43 acute respiratory syndrome and Middle East respiratory syndrome, but more
44 similar to the mortality risk of 2009 H1N1 influenza pandemic in hospitalized
45 patients. The fatality risk of COVID-19 is higher in males and increases with age.
46 Our study improves the severity assessment of the ongoing epidemic and can
47 inform the COVID-19 outbreak response in China and beyond.

48

49

50

51

52 **Introduction**

53 As of March 3, 2020, a total of 80,270 cases of novel coronavirus disease 2019
54 (COVID-19) have been reported in mainland China, including 2,981 deaths. The
55 outbreak is caused by a novel coronavirus of presumed zoonotic origin, the
56 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)¹. COVID-19 cases
57 have now been identified in 72 countries, some of which have reported onward
58 local transmission and deaths². The unprecedented scale of the epidemic has
59 prompted an urgent need for clinical severity assessment, of which the case
60 fatality risk (CFR) is a key metric.

61 A few studies have assessed the fatality risk of COVID-19 but estimates have been

62 highly variable. Wu et al. estimated that the fatality risk among hospitalized cases
63 was 14% during the early phase of outbreak in Wuhan³. Dorigatti et al. estimated
64 that the CFR among laboratory-confirmed cases was 18% in Hubei province and
65 ranged from 1.2-5.6% outside mainland China⁴. A recent report of the World
66 Health Organization (WHO)-China Joint Mission on Coronavirus Disease 2019
67 estimated the case fatality risk as 3.8% by dividing the number of deaths at the
68 time of analysis by the number of laboratory-confirmed cases at the time of
69 analysis⁵. They also reported a higher case fatality risk in Hubei than that in other
70 provinces (5.8% vs. 0.7%)⁵. However, those estimates would be a lower bound on
71 the CFR for the laboratory-confirmed cases because many cases were still in the
72 hospital and had not reached a final outcome of either death or discharge after
73 recovery⁶.

74 In the present study, we aimed to assess the CFR among laboratory-confirmed
75 COVID-19 cases detected in mainland China, stratified by different clinical
76 categories (e.g. mild-, moderate-, severe- and critical-patients) and by age group.
77 We also explored the risk factors associated with fatal outcomes.

78

79 **Methods**

80 **Case definitions and surveillance**

81 The National Health Commission of China (NHC) and the Chinese Center for
82 Disease Control and Prevention (China CDC) have launched a new surveillance

83 system to record information on COVID-19 cases since the start of the outbreak
84 of atypical pneumonia cases in Wuhan in late December 2019. A description of
85 the surveillance system is provided elsewhere⁷. As the epidemic evolves, a total
86 of six versions of case definitions for suspected- and laboratory-confirmed-cases
87 have been issued by NHC⁷⁻⁹. Details are provided in the Appendix table S1.

88 Four clinical categories of laboratory-confirmed COVID-19 patients have been
89 identified by NHC, including mild-, moderate-, severe-, and critical-patients⁷⁻⁹.
90 Mild patients, introduced in the fifth and sixth versions of COVID-19 case
91 definition, refer to patients with mild symptoms and no radiographic evidence of
92 pneumonia. Moderate patients, introduced in the fourth version of the case
93 definition, refers to patients with fever, respiratory symptoms, and radiographic
94 evidence of pneumonia. Severe patients, introduced in the second version, refers
95 to patients with any breathing problems, finger oxygen saturation, and low
96 PaO₂/FiO₂ (PaO₂ denotes partial pressure of oxygen in arterial blood; FiO₂
97 denotes fraction of inspired oxygen), etc. Critical patients, a definition used from
98 the very beginning of the outbreak, refer to patients having any respiratory
99 failure, shock, and any other organ failure that requires ICU admission.

100 Patients were discharged when they met all the following criteria: 1) normal
101 body temperature for more than 3 days, 2) significantly improved respiratory
102 symptoms, 3) significant inflammation absorption in lung radiographic findings,
103 and 4) negative nucleic acid detection by real-time RT-PCR using respiratory

104 specimens on two consecutive days, with a sampling interval ≥ 1 day⁹.

105

106 **Data collection**

107 Daily aggregated data (hereafter called aggregated dataset) on the cumulative
108 number of cases were extracted from the websites of national, provincial, and
109 municipal Health Commissions. Individual records on laboratory-confirmed
110 COVID-19 cases (hereafter called individual dataset) were collected from two
111 official publicly available sources from December 29, 2019 through to February
112 23, 2020, including: 1) the websites of national, provincial, and municipal Health
113 Commission; 2) the websites of national and local government affiliated medias.
114 Individual information was extracted and entered into a structured database
115 comprising demographic characteristics, dates of symptom onset, first healthcare
116 consultation, hospital admission, official announcement (reporting date), as well
117 as outcome information (e.g. death/discharge and corresponding dates). Each
118 individual record was extracted and entered by three coauthors and was
119 cross-checked to ensure data accuracy. Conflicting information was resolved
120 based on the Health Commission data. Details on the collection of individual data
121 and assessment of completeness of variables used in the study are provided in
122 Appendix Tables S2-3.

123

124 **Statistical analysis**

125 We restricted analyses of demographic characteristics, risk factors associated
126 with fatal outcome, and key time to event intervals to the provinces outside
127 Hubei, where the majority of individual records were obtained (97.6%,
128 9,651/9,886) as of February 23, 2020. We implemented a multivariate logistic
129 regression model to explore the risk factors associated with death. We included
130 age, sex, economic region¹⁰, time interval from symptom onset to first medical
131 consultation, first hospital admission, and laboratory diagnosis. We categorized
132 China into three economic regions (East, Central, West and Northeast) according
133 to gross domestic product per capita in 2018 (see Appendix Figure S1 for map)¹⁰.
134 We estimated key time-to-event distributions including symptom onset to first
135 healthcare consultation, hospital admission, laboratory diagnosis, and death or
136 discharge, and from hospital admission to death or discharge. We fitted three
137 parametric distributions (Weibull, gamma, and lognormal) to time-to-event data
138 and selected the best fit based on the minimum Akaike information criterion.
139 We used three methods to estimate CFR among COVID-19 cases. Firstly, we
140 calculated a crude CFR based on the cumulative number of deaths divided by the
141 cumulative number of laboratory-confirmed cases, ignoring the time-lag between
142 symptoms onset and death and resulting right-censoring of outcomes⁵.
143 In a second approach, we adjusted for delays between hospitalization and death
144 to obtain more accurate estimates of CFR, using the method described by Garske
145 et al. for pandemic influenza A/H1N1 in 2009¹¹. For above two methods, we used

146 the aggregated dataset as of February 23, and binomial distributions were used
147 to estimate the 95% CIs.

148 Thirdly, to allow for incomplete information about outcomes, we used survival
149 analyses to allow inclusion of all cases admitted to hospital in the individual
150 dataset, incorporating data for patients who were still in hospital at the time of
151 analysis. In our individual dataset, the outcome was unavailable for some
152 patients because the information was not communicated through public
153 channels, although these patients may have been discharged or died at the time
154 of this writing (hereafter denoted as missing outcome). This is different from the
155 issue of right-censoring for patients still hospitalized whose illnesses have yet to
156 be resolved. The cases who were still hospitalized and those with missing
157 outcome were treated as unresolved in our analysis. A multiple imputation was
158 used to generate outcomes for these patients.

159 For each date t , we calculated the number of discharged/deceased patients that
160 required imputation by subtracting the number of discharged/deceased patients
161 in our individual dataset from that in the aggregated dataset (Table S4 in
162 Appendix). All these patients with missing data for outcomes before date t were
163 considered for imputation on date t . They were randomly selected as discharge
164 or death according to probability calculated using the density of interval from
165 hospital admission to discharge/death. This imputation procedure was repeated
166 100 times to generate 100 imputed datasets for further estimation of CFR.

167 We employed a dual-outcome (discharge or death) time to event framework to
168 estimate CFR based on the fraction $F1/(F1+F2)$ ¹². F1 and F2 stand for the
169 admission to death distribution and the admission to discharge distribution,
170 respectively.

$$\hat{F}_j(t) = \sum_{t_i \leq t} \frac{d_{ij}}{n_i} \hat{S}(t_i)$$

171 Where, $t_1 < \dots < t_k$ denotes the distinct observed event times for outcomes
172 (discharge or death), with d_{ij} representing the number of outcome j that occur
173 at time t_i . \hat{S} is the Kaplan-Meier estimator of overall survival function
174 (combined event of discharge or death)¹². Then we implemented a 1,000 times
175 bootstraps for estimation of 95% CIs, and used Rubin's formula to pool all
176 estimates across 100 imputed datasets¹³.

177 For the survival analysis, we restricted analyses to the provinces outside Hubei,
178 considering the completeness of individual records obtained. When estimating
179 CFR, we excluded cases hospitalized beyond 17 days on each date in the baseline
180 analysis. The choice of 17 days was based on the 90th percentile of the
181 distribution of the time from hospitalization to outcome (discharge/death)
182 among COVID-19 cases in the provinces other than Hubei as of February 23. As a
183 sensitivity analysis, we also considered the 80th and 50th percentiles of this
184 distribution, corresponding to 14 days and 10 days, respectively. As mentioned
185 above, patients were discharged only after testing negative by nucleic acid
186 detection tests on two consecutive days⁹. Hence, we assumed that these patients

187 had biologically recovered three days prior to the reported date of discharge,
188 accounting for one additional day for delay of laboratory confirmation and
189 official reporting. Accordingly, for the discharged patients, their time from
190 hospitalization to discharge was cut down by three days when estimating the
191 90th, 80th and 50th percentiles of the distribution of the time from hospitalization
192 to outcome.

193 All deaths occurred among critical cases, as reported by China CDC⁵. Separately
194 for mainland China, Hubei Province, and the provinces outside Hubei, we further
195 estimated CFRs among severe and critical patients by dividing the above derived
196 CFR by proportions of severe and critical patients among all reported COVID-19
197 cases. We used the average of daily proportions among COVID-19 cases who were
198 still in hospitals on each day other than the clinical severity on admission, which
199 were obtained from the aggregated dataset and showed very stable (Table S2,
200 and Figure S2-3 in Appendix). Only Guangdong Province officially reported
201 aggregated data on mild-, moderate-, severe- and critical patients. And thus, for
202 the provinces outside Hubei, we additionally estimated the CFR by these clinical
203 categories using the corresponding proportions in Guangdong Province.
204 Statistical analyses were performed with R (version 3.6.0).

205

206 **Ethics**

207 The study was approved by the Institutional review board from School of Public

208 Health, Fudan University (IRB#2020-02-0802). All data were collected from
209 publicly available sources and did not contain any personal information.

210

211 **Results**

212 As of February 23, 2020, a total of 77,150 laboratory-confirmed cases with 2,592
213 deaths, 24,711 discharged and 49,847 patients who were still hospitalized were
214 reported in mainland China (see Table S2 for details of each province). Of these,
215 provinces outside Hubei accounted for 12,863 (16.7%, 12,863/77,150)
216 laboratory-confirmed cases including 97 deaths (3.7%, 97/2,592), 7,973 (32.3%,
217 7,973/24,711) discharged cases and 4,793 (9.6%, 4,793/49,847) patients who
218 were still hospitalized. We collected individual information from publicly
219 available official sources on 9,651 laboratory-confirmed cases detected outside
220 Hubei by February 23, accounting for 75.0% (9,651/12,863) of total cases
221 reported, 63.9% (62/97) of deceased patients, 18.2% (1,449/7,973) of recovered
222 patients. Of 9,651 cases, unresolved patients accounted for 84.3% (8,140/9,651)
223 (Table 1). See Figure S4 for the epidemic curve of cases with available individual
224 information.

225 The median age of the cases outside Hubei was 45 years (range, four days-97
226 years), and 51% (4,956/9,651) were male. Those who died were significantly
227 older than discharged cases (median age: 77 vs 39 years, $p < 0.001$). 77% (48/62)
228 of deaths occurred in the older adults aged 65 years or above, and 58% (36/62)

229 were male. (Table 1) Multivariate logistic analysis revealed that increasing age,
230 being male, and living in less developed economic regions (e.g. Central region or
231 West and Northeast region) were risk factors for mortality (Table 2). The
232 univariate logistic analysis is shown in Appendix Table S5.

233 The intervals from symptom onset to first healthcare consultation, from
234 symptom onset to hospitalization, and from symptom onset to laboratory
235 confirmation were consistently longer for deceased patients than for those who
236 recovered. Overall, the time interval from symptom onset to death was estimated
237 to be 12.9 days (95%CI: 2.2 to 40.2), and from symptom onset to discharge was
238 16.7 days (95%CI: 8.6 to 28.9). (Table 3)

239 Based on the total patients reported to the surveillance system, the CFR
240 estimated by Garske's method¹¹ and survival analyses were all higher than the
241 crude CFR (Table 4 and Appendix Table S6). The CFR estimated by Garske's
242 method¹¹ was 4.52% (95%CI: 4.47% to 4.67%) in mainland China, with highest
243 estimate in Wuhan (6.19%, 95%CI: 6.12% to 6.41%), and lowest in the provinces
244 outside Hubei (0.89%, 95%CI: 0.83% to 1.06%). The CFR estimated by survival
245 analyses was 1.24% (95%CI: 1.24% to 1.24%) among all cases, and 11.21%
246 (95%CI: 11.21% to 11.21%) among severe and critical patients outside Hubei.
247 There was no difference in both overall CFR and that among severe and critical
248 patients outside Hubei estimated by survival analyses or Garske's method
249 ($p>0.05$) (Table 4). In sensitivity analyses, we excluded all cases hospitalized for

250 longer than 14 and 10 days, and the estimates were all consistent with those of
251 the baseline analysis, in which we excluded cases with hospitalizations longer
252 than 17 days ($p>0.05$) (Appendix, Figure S5-6).

253 In the provinces other than Hubei, the CFR increased with age, with highest
254 estimates among patients aged ≥ 70 years (Figure 1 panel B). The fatality risk
255 among critical patients was 23.8-33.3%, which was 2-fold higher than that
256 among severe and critical patients, and 24-fold higher than that among moderate,
257 severe and critical patients (Figure 1 panel B). The CFR among all cases
258 estimated by survival analyses declined rapidly from 8% on January 25 to around
259 1% on January 28, and remained at 1.2-1.5% afterwards. Patterns were similar
260 for estimates using Garske's method (Figure 2).

261

262 **Discussion**

263 We have shown that the fatality risk among detected cases was 0.89-1.24% in the
264 provinces outside Hubei in mainland China and increased with clinical severity.
265 Further, the CFR was estimated at 8.02-11.21% among severe and critical
266 patients. Estimates accounting for right-censoring of unresolved cases were
267 higher than crude estimates. Male patients, older age, and less developed regions
268 were factors associated with a higher CFR. These estimates could represent the
269 most accurate estimates of CFR in China so far.

270 Our study is strengthened by accounting for unknown outcomes among patients

271 who were still in hospital at time of data cutoff. We used Garske's method
272 developed for pandemic influenza A/H1N1 in 2009¹¹, as well as survival analyses,
273 both of which consider censoring. Compared to the crude CFR, Garske's method
274 improved estimates by adjusting for the cumulative density of intervals from
275 hospital admission to death/discharge. Survival analysis was a useful tool for
276 comparison as it relied on a very large individual dataset comprising a total of
277 9,651 reported cases. Availability of individual data enabled us to explore
278 mortality risk factors and estimate CFR by age group. Estimates have not been
279 reported previously based on such a large sample size and a competing risk
280 model of survival analysis with 90th quantile truncation. There was no difference
281 in CFR estimates by the two methods, which lends support to our estimates.

282 Our study has some limitations. First, in the individual dataset, the clinical profile
283 of patients was not available. Hence, we could not provide direct estimates of
284 fatality risk stratified by clinical categories using survival analysis. Instead, we
285 divided the estimated CFR among all cases by the proportions of different clinical
286 categories obtained from the aggregated dataset. This is a reasonable approach
287 method because all deaths occurred among critical cases ⁵.

288 Second, the analyzed individual records were retrieved from publicly available
289 official sources, ensuring accuracy and reliability of information. However, we
290 were only able to collect few individual records in Hubei because they did not
291 release complete individual information. And thus, we were unable to estimate

292 CFR in Hubei using survival analyses. Moreover, assessment of clinical severity in
293 Hubei, especially in the epicenter of the outbreak in Wuhan, is challenging
294 because disease severity may be increased by bottlenecks in local healthcare
295 capacity, as COVID-19 cases surged. In addition, case surveillance and clinical
296 management were biased towards severe cases in Hubei, especially in the early
297 phase of the epidemic. Our estimates of the CFR in Hubei and Wuhan using
298 Garske's method¹¹ should be viewed cautiously as the sensitivity of surveillance
299 of both deaths and cases remains unclear.

300 Our study only addresses CFR among detected cases. The level of ascertainment
301 of mild cases remains unclear. More estimates that include fatality risk among
302 syndromic patients and asymptotically infected individuals can only be
303 available through enhanced routine surveillance, such as increased testing of
304 patients with influenza-like-illnesses, and by analysis of future
305 sero-epidemiological studies.

306 Our CFR estimates of 0.89-1.24% among detected COVID-19 patients outside
307 Hubei province are higher than the crude CFRs reported by WHO and China CDC,
308 which is 0.4-0.7%^{5,14}. It is expected that the crude CFR obtained by dividing the
309 cumulative number of reported deaths by the cumulative number of reported
310 cases is an underestimate due to the inevitable delay between symptom onset
311 and death. Our findings reveal that older individuals and male patients
312 experience higher fatality risk, which is consistent with the WHO report¹⁴.

313 Additionally, WHO reported that patients with underlying conditions had much
314 higher fatality rates¹⁴. Our study was unable to address the relative risk of fatal
315 outcome among patients with underlying diseases compared to healthy people,
316 because limited information on underlying conditions was available from
317 publicly available data sources.

318 A clinical study conducted in Wuhan showed that 4.3% of hospitalized patients
319 died¹⁵. Another study relying on patients from 552 hospitals across 30 provinces,
320 found that 1.4% of patients died¹⁶, in which less study participants (28%) were
321 from Wuhan. The estimates of these clinical studies would be the lower bounds
322 for the CFR since separately 62% and 94% of patients were still in hospitals. The
323 fatality risk from these clinical studies is higher than our CFR estimates, probably
324 due to shortage of health services in Wuhan, e.g., advanced health care facilities
325 for critically ill patients as extracorporeal membrane oxygenation.

326 Our CFR estimates outside Hubei province indicate that the severity of
327 SARS-CoV-2 is lower than that of other diseases caused by zoonotic
328 coronaviruses, including Middle East respiratory syndrome (MERS), which had
329 an estimated CFR of 34.4% globally¹⁷, and severe acute respiratory syndrome
330 (SARS) with an estimated CFR of 10.9% across the world and 7% in mainland
331 China¹⁸. In contrast, the CFR of SARS-CoV is in the same order of magnitude as
332 that of pandemic 2009 influenza A(H1N1) virus hospitalizations, which has an
333 estimated CFR of 1.4% among hospitalized patients in Asia¹⁹. In the long run,

334 depending on how much of the severity pyramid of SARS-CoV2 is captured in our
335 data, the absolute severity of SARS-CoV2 may prove to be similar, to somewhat
336 more severe, than the 2009 influenza pandemic, albeit with a different age profile.
337 Comparison of clinical data from China and other countries will prove useful to
338 settle this question.

339 Outside Hubei province, close contacts of laboratory-confirmed cases were kept
340 in quarantine for 14 days, and local hospitals tested patients with respiratory
341 symptoms (e.g., fever and cough) and epidemiological links to Hubei province or
342 other cases. This strategy would have enabled detection of many mild cases.

343 However, a small number of mild cases were captured. In our aggregated dataset
344 for Guangdong province for instance, only 8% of reported cases were mild, while
345 the majority (83%) of reported cases had moderate disease severity with
346 presence of pneumonia. And thus, our CFR estimates could approximately
347 represent the fatality risk among laboratory-confirmed COVID-19 cases with
348 chest x-ray confirmed pneumonia. Even though clinical information for these
349 patients was not available from publicly available sources, we believe that our
350 CFR estimates could be viewed as the fatality risk among hospitalized COVID-19
351 cases. Chest x-ray confirmed pneumonia is a threshold for hospital admissions in
352 in China. This may vary among countries due to different clinical practices and
353 health service capacity.

354 Clinical studies have reported a higher proportion of severe patients among older

355 age group (29% vs. 14%)¹⁶. No specialized treatment for COVID-19 patients has
356 been identified, and the mainstay clinical management has been supportive care.
357 For non-critically ill patients, close follow-up is likely to be sufficient to manage
358 the disease. But critically ill patients were more likely to develop ARDS and
359 require ICU admission²⁰. That could explain our findings that severe patients had
360 a higher fatality risk. The high observed CFRs of COVID-19 in older adults is
361 consistent with the age profile of MERS, SARS, pandemic H1N1 2009, and
362 seasonal influenza.¹⁹

363 Compared to the Eastern region, cases detected in the less developed Central
364 region had a 2.45-fold higher risk of death, and those in West and Northeast
365 region had a 2.93-fold higher risk. It is important to note that those variations in
366 CFR do not reflect underlying differences in clinical disease severity. CFR will
367 vary regionally depending on the sensitivity of surveillance systems to detect
368 cases at different levels of the severity pyramid and clinical care offered to severe
369 and critical patients. More attention should be paid to less developed settings
370 with limited health services like Iran, which reports a larger ratio of deaths to
371 cases than other countries².

372 Notably, the definition of suspected cases eligible for laboratory testing used in
373 China shifted from a narrow clinical criteria based on three symptoms early in
374 the outbreak (fever; radiographic findings of pneumonia; normal or reduced
375 white blood cell count, or reduced lymphocyte count at early onset of symptoms),

376 to a broader criteria including any two of three symptoms by January 27. This
377 would bias our sample towards more clinically severe cases before January 27. In
378 addition to improvement in therapeutic capacity, the shift in surveillance
379 definition could partially explain the declining trend of CFR from 8% to around 1%
380 at the end of January, which remained stable afterwards. Accordingly, our CFR
381 estimate for February could provide a true picture of the severity of
382 laboratory-confirmed cases of COVID-19.

383 In conclusion, our estimates of CFR among laboratory-confirmed cases suggest
384 that COVID-19 is not as severe as SARS and MERS, but similar to that of pandemic
385 2009 H1N1 among hospitalized patients. The fatality risk of COVID-19 cases is
386 higher in male, and increases with age, particularly in adults aged 70 years and
387 above. Our findings can inform the severity assessment and response to the
388 on-going COVID-19 outbreak, and assist preparations for a global epidemic of
389 COVID-19.

390

391 **Contributors**

392 H.Y. conceived, designed and supervised the study. W.W., J.L., Y.C., H.Y., Y.Z., Q.Q.,
393 H.G., Xiang.W., L.W. and K.S. participated in data collection. X.D., J.Y., X.W., JX.Z., Z.C.,
394 J.Z., and Y.W. analyzed the data, and prepared the figures. J.Y. prepared the first
395 draft of the manuscript. X.D., P.W., M.A., B.C., C.V., and H.Y. commented on the data
396 and its interpretation, revised the content critically. All authors contributed to

397 review and revision and approved the final manuscript as submitted and agree to

398 be accountable for all aspects of the work.

399

400 **Disclaimer:** The findings and conclusions in this study are those of the authors

401 and do not necessarily represent the official position of the National Institutes of

402 Health or U.S. Department of Health and Human Services.

403

404 **Acknowledgments**

405 We thank Xin Chen, Jiaxian Chen, and Sihong Zhao, from School of Public Health,

406 Fudan University, and Yuheng Feng from School of Basic Medical, Sciences, Fudan

407 University for providing assistance with data collection.

Reference

1. National Health Commission of the People's Republic of China. Update on COVID-19 as of 24:00 on March 3, 2020.
<http://www.nhc.gov.cn/xcs/yqtb/202003/7a5f57b3f1b94954b1fc25f81dacc874.shtml>
(accessed March 4 2020).
2. World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report – 343 2020.
https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200303-sitrep-43-covid-19.pdf?sfvrsn=2c21c09c_2 (accessed March 4 2020).
3. Wu P, Hao X, Lau EHY, et al. Real-time tentative assessment of the epidemiological characteristics of novel coronavirus infections in Wuhan, China, as at 22 January 2020. *Euro Surveill* 2020.
4. WHO Collaborating Centre for Infectious Disease Modelling and Imperial College London. Dorigatti I, Okell L, Cori A, Imai N, Baguelin M, Bhatia S, et al. Report 4: Severity of 2019-novel coronavirus (nCoV).
<https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/>. (accessed Feb 25 2020).
5. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020. *China CDC Weekly* 2020; 2(8):113-22.
6. Lipsitch M, Donnelly CA, Fraser C, et al. Potential Biases in Estimating Absolute and Relative Case-Fatality Risks during Outbreaks. *PLoS Negl Trop Dis* 2015; 9(7): e0003846.
7. Prevention CCFDCa. Epidemic update and risk assessment of 2019 Novel Coronavirus. 2020. <http://www.chinacdc.cn/yyrdgz/202001/P020200128523354919292.pdf> (accessed Jan 31 2020).
8. National Health Commission of China. The diagnosis and treatment scheme of novel coronavirus diseases 2019 (Trial version 5th).
http://www.gov.cn/zhengce/zhengceku/2020-02/05/content_5474791.htm. (accessed Feb 25 2020).
9. National Health Commission of China. The diagnosis and treatment scheme of novel coronavirus diseases 2019 (Trial version 6th).
http://www.gov.cn/zhengce/zhengceku/2020-02/19/content_5480948.htm. (accessed Feb 25 2020).
10. National Bureau of Statistics of China. China Statistical Yearbook 2019.
<http://www.stats.gov.cn/tjsj/ndsj/2019/indexch.htm>. (accessed Feb 25 2020)
11. Garske T, Legrand J, Donnelly CA, et al. Assessing the severity of the novel influenza A/H1N1 pandemic. *BMJ* 2009; 339: b2840.
12. Jewell NP, Lei X, Ghani AC, et al. Non-parametric estimation of the case fatality ratio with competing risks data: an application to Severe Acute Respiratory Syndrome (SARS). *Stat Med* 2007; 26: 1982–98.
13. Rubin DB. Multiple imputation for nonresponse in surveys. New York, NY: J Wiley and Sons, 1987.
14. World Health Organization. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19).

<https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>. (accessed on Feb 29 2020).

15. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* 2020: 10.1001/jama.2020.1585.

16. Guan W-J, Ni Z-Y, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *The New England journal of medicine* 2020: 10.1056/NEJMoa2002032.

17. World Health Organization. Middle East respiratory syndrome coronavirus (MERS-CoV) – The Kingdom of Saudi Arabia.

<https://www.who.int/csr/don/18-december-2019-mers-saudi-arabia/en/>. (accessed on Feb 29 2020).

18. World Health Organization. Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003.

https://www.who.int/csr/sars/country/table2003_09_23/en/. (accessed on Feb 29 2020).

19. Wong JY, Kelly H, Cheung C-MM, et al. Hospitalization Fatality Risk of Influenza A(H1N1)pdm09: A Systematic Review and Meta-Analysis. *Am J Epidemiol* 2015; **182**(4): 294-301.

20. Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020: S2213-600(20)30079-5.

Table 1. Demographical characteristics of COVID-19 cases outside Hubei province in mainland China, as of February 23, 2020

| Characteristic | Died n=62 | Discharged n=1,449 | Unresolved^a n=8,140 | All cases n=9,651 |
|--------------------------------------|----------------------|-------------------------------|---|------------------------------|
| Median age (year, range) | 77 (25–94) | 39 (0.15–95) | 46 (0.01–97) | 45 (0.01–97) |
| Age group (year) (n, %) ^b | | | | |
| 0-6 | 0 (0) | 30 (2) | 107 (1) | 137 (1) |
| 7-17 | 0 (0) | 54 (4) | 211 (3) | 265 (3) |
| 18-24 | 0 (0) | 110 (8) | 391 (5) | 501 (5) |
| 25-49 | 2 (3) | 815 (56) | 3,746 (46) | 4,563 (47) |
| 50-64 | 10 (16) | 283 (20) | 2,141 (26) | 2,434 (25) |
| ≥65 | 48 (77) | 103 (7) | 1,028 (13) | 1,179 (12) |
| Missing | 2 (3) | 54 (4) | 516 (6) | 572 (6) |
| Sex (n, %) | | | | |
| Male | 36 (58) | 742 (51) | 4,178 (51) | 4,956 (51) |
| Female | 26 (42) | 669 (46) | 3,795 (47) | 4,490 (47) |
| Missing | 0 (0) | 38 (3) | 167 (2) | 205 (2) |
| Region (n, %) ^c | | | | |
| East | 24 (39) | 751 (52) | 3,259 (40) | 4,034 (42) |
| Central | 15 (24) | 364 (25) | 3,160 (39) | 3,539 (37) |
| West and Northeast | 23 (37) | 334 (23) | 1,721 (21) | 2,078 (22) |

^a Including these cases who may had outcomes (i.e., death/discharge), but their information unavailable from public data sources. ^b Significant difference was observed among patients who died and the discharged ($p < 0.001$). ^c Significant difference was observed among patients who died and the discharged ($p < 0.05$). East: Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong and Hainan provinces; Central: Shanxi, Anhui, Jiangxi, Henan, and Hunan provinces; West: Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang; Northeast: Heilongjiang, Jilin and Liaoning.

Table 2. Risk factors associated with fatal outcome among COVID-19 patients

| Variables | OR (95%CI) | Z-value | P-value |
|--|-------------------|----------------|----------------|
| Age, per year increase | 1.18 (1.14-1.22) | 10.2 | <0.001 |
| Sex | | | |
| Female | ref | / | / |
| Male | 2.02 (1.02-4.03) | 2.00 | 0.045 |
| Unknown | 0 (0-Inf) | -0.01 | 0.990 |
| Economic regions ^a | | | |
| East | ref | / | / |
| Central | 3.45 (1.32-9.03) | 2.53 | 0.012 |
| West and Northeast | 3.93 (1.74-8.85) | 3.30 | <0.001 |
| Time from symptom onset to first healthcare consultation | | | |
| ≤2 days | ref | / | / |
| >2 days | 1.11 (0.33-3.71) | 0.17 | 0.863 |
| Unknown | 0.40 (0.14-1.14) | -1.72 | 0.086 |
| Time from symptom onset to hospital admission | | | |
| ≤3 days | ref | / | / |
| >3 days | 0.65 (0.20-2.11) | -0.72 | 0.471 |
| Unknown | 0.55 (0.18-1.69) | -1.04 | 0.298 |
| Time from symptom onset to laboratory confirmation | | | |
| ≤6 days | ref | / | / |
| >6 days | 1.41 (0.42-4.74) | 0.56 | 0.575 |
| Unknown | 1.41 (0.44-4.55) | 0.58 | 0.561 |

^a East: Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong and Hainan provinces; Central: Shanxi, Anhui, Jiangxi, Henan, and Hunan provinces; West: Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang; Northeast: Heilongjiang, Jilin and Liaoning. /not applicable.

Table 3. Key time to event intervals of COVID-19 patients outside Hubei province in mainland China, as of February 23, 2020 (mean, 95%CI)

| Key time to event interval | Died n=62 | Discharged n=1,449 | Unresolved ^a n=8,140 |
|---|------------------|-----------------------|------------------------------------|
| Time from symptom onset to first healthcare consultation (days) | n=27 | n=493 | n=2,838 |
| Estimates from empirical data | 3.0 (0.5, 11.1) | 1.0 (0.5, 11.0) | 2.0 (0.5, 11.0) |
| Estimates by fitting | 2.0 (0.2, 18.2) | 1.6 (0.2, 13.2) | 1.6 (0.2, 12.5) |
| Time from symptom onset to hospital admission (days) | n=36 | n=572 | n=2,148 |
| Estimates from empirical data | 4.0 (0.5, 13.2) | 3.0 (0.5, 12.0) | 3.0 (0.5, 12.0) |
| Estimates by fitting | 3.4 (0.1, 17.0) | 3.1 (0.2, 13.1) | 2.2 (0.3, 18.0) |
| Time from symptom onset to laboratory confirmation (days) | n=35 | n=654 | n=4,955 |
| Estimates from empirical data | 6.0 (0.5, 16.0) | 5.0 (0.7, 15.0) | 5.0 (0.5, 16.0) |
| Estimates by fitting | 5.6 (0.6, 17.2) | 5.1 (0.6, 15.3) | 5.2 (0.6, 15.8) |
| Time from symptom onset to death/discharge (days) | n=41 | n=769 | / |
| Estimates from empirical data | 13.0 (3.0, 39.0) | 17.0 (8.0, 29.0) | / |
| Estimates by fitting | 12.9 (2.2, 40.2) | 16.7 (8.6, 28.9) | / |
| Time from hospital admission to death/discharge (days) | n=49 | n=980 | / |
| Estimates from empirical data | 8.0 (0.6, 35.4) | 13.0 (6.0, 24.5) | / |
| Estimates by fitting | 8.0 (0.8, 30.2) | 13.1 (6.0, 24.2) | / |

^a Including these cases who may had outcomes (i.e., death/discharge), but their information unavailable from publicly data sources; / not applicable.

Table 4. Fatality risk of COVID-19 among all reported cases, and among severe and critical cases ^a

| | Number of cases | | Fatality risk among all reported cases (%, 95%CI) | | | Fatality risk among severe and critical patients (%, 95%CI) ^b | | |
|------------------------|-----------------|----------------------|--|---|--------------------------------|---|---|--------------------------------|
| | Death | Total cases reported | Crude | Estimated using Garske's method ¹¹ | Estimated by survival analyses | Crude | Estimated using Garske's method ¹¹ | Estimated by survival analyses |
| Mainland China | 2,592 | 77,150 | 3.36 (3.23, 3.49) | 4.52 (4.47, 4.67) | / | 17.97 (17.30, 18.66) | 24.18 (23.91, 24.96) | / |
| Hubei province | 2,495 | 64,287 | 3.88 (3.73, 4.03) | 5.37 (5.31, 5.55) | / | 18.57 (17.86, 19.30) | 25.71 (25.42, 26.54) | / |
| Wuhan | 1,987 | 46,607 | 4.26 (4.08, 4.45) | 6.19 (6.12, 6.41) | / | 18.54 (17.75, 19.35) | 26.93 (26.61, 27.88) | / |
| Outside Wuhan | 508 | 17,680 | 2.87 (2.63, 3.13) | 3.54 (3.44, 3.81) | / | 23.36 (21.4, 25.45) | 28.75 (28.00, 30.99) | / |
| Outside Hubei province | 97 | 12,863 | 0.75 (0.62, 0.92) | 0.89 (0.83, 1.06) | 1.24 (1.24, 1.24) | 6.79 (5.57, 8.28) | 8.02 (7.48, 9.55) | 11.21 (11.21, 11.21) |

^a crude fatality risk was calculated as the cumulative number of deaths divided by the cumulative number of laboratory-confirmed cases. /not estimated using survival analyses due to limited individual data.

^b estimated using the proportion of severe and critical patients among reported cases in corresponding areas.

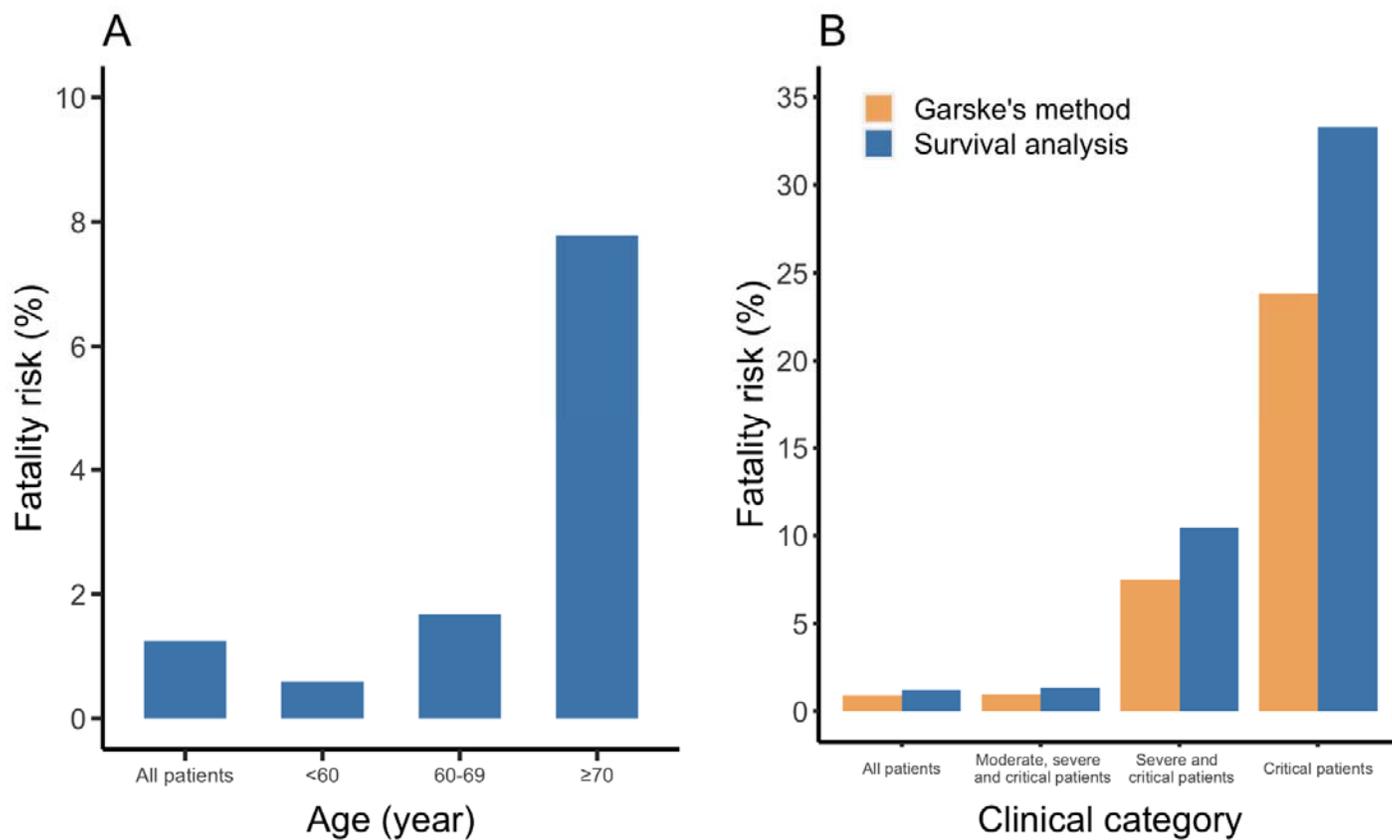


Figure 1. Case-fatality risk (mean) outside Hubei province in mainland China. A: by age group; B: by clinical categories (All patients includes mild, moderate, severe and critical patients). 95%CI was narrow and thus not presented here.

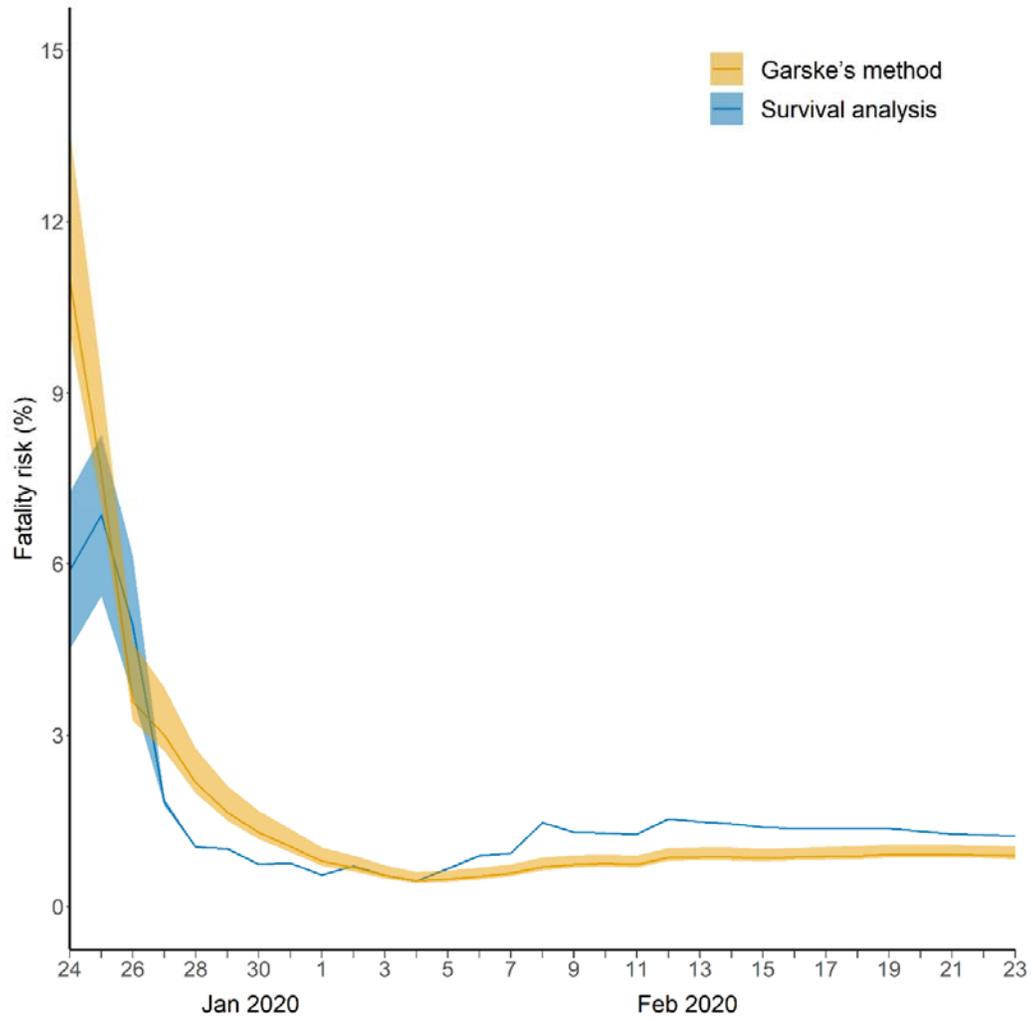


Figure 2. Case-fatality risk over time outside Hubei province in mainland China (%) (mean, 95% CI).