1   **Title**

2   Estimating the risk of 2019 Novel Coronavirus death during the course of the outbreak in China,

3   2020

4

5   **Authors:**

6   Kenji Mizumoto, Gerardo Chowell

7   **Affiliations:**

8   Georgia State University, Atlanta, GA, USA (K. Mizumoto, G. Chowell)

9   Kyoto University, Kyoto, Japan (K. Mizumoto)

10

11

12

Abstract

14   Since the first case of Novel Coronavirus (2019-nCov) was identified in December 2019

15   in Wuhan City, China, the number of cases continues to grow across China and multiple cases

16   have been exported to other countries. The cumulative number of reported deaths is at 637 as of

17   February 7, 2020. Here we statistically estimated the time-delay adjusted death risk for Wuhan as

18   well as for China excluding Wuhan to interpret the current severity of the epidemic in China. We

19   found that the latest estimates of the death risk in Wuhan could be as high as 20% in the

20   epicenter of the epidemic whereas we estimate it ~1% in the relatively mildly-affected areas.

21   Because the elevated death risk estimates are likely associated with a breakdown of the

22   medical/health system, enhanced public health interventions including social distancing and

23   movement restrictions should be effectively implemented to bring the epidemic under control.

24

Background

26   Since the first case of Novel Coronavirus (2019-nCov) was identified in December 2019

27   in Wuhan City in the Hubei Province of China, the number of cases continues to expand across

28   China, with multiple exported cases reported in over 25 countries. In particular, the cumulative

29   number of deaths is at 637 as of February 7, 2020, a figure that will soon surpass the number of

30   people that succumbed to Severe acute respiratory syndrome (SARS) in 2002-2003 *(1)*.

31   In the context of an emerging infectious disease with pandemic potential, it is critical to

32   assess not only its ability to spread between humans, but also the associated risk of death from

33   the disease. In particular, the type and intensity of public health interventions are often a function

34   of these epidemiological metrics. Importantly, in the absence of vaccines or antiviral for SARS-

35   CoV and 2019-nCov, the effective implementation of non-pharmaceutical interventions (social

36   distancing) and movement restrictions, which are the only alternative strategies available to

37   mitigate their spread in the population, also pose significant pressure on the global economy *(2)*.

38   As interventions are gradually implemented and calibrated during the course of an

39   outbreak, early estimates of the case fatality risk (CFR) provide crucial information for policy

40   makers. Yet, the estimation of epidemiological characteristics including the CFR during the

41   course of an outbreak tend to be affected by right censoring and ascertainment bias *(3-5)*. The

42   phenomenon of right censoring is caused by the gap of illness onset to death between vulnerable

43   population and healthy population, resulting in underestimation, while ascertainment bias is due

44   to the unreported bulk of cases with mild symptoms or asymptomatic infection, potentially

45   leading to overestimation.

46   The purpose of this study is to derive estimates of the time-delay adjusted CFR for 2019-

47   nCov for Wuhan, China as well as for China excluding the city of Wuhan with quantified

48   uncertainty in order to interpret the current severity of the epidemic in China.

49

50   **Method**

51   **Data sources**

52   We use two different types of data. First, daily confirmed cases and deaths in China were

53   extracted from daily reports published by China government, Hubei Province in China and

54   Wuhan City *(7-9)*. These data were categorized by area of Wuhan City, Hubei Province

55   excluding Wuhan City or China excluding Hubei Province, considering that over 50% of the

56   deaths are occurring in Hubei Province. In particular, the great majority of the deaths have

57   occurred in Wuhan City. The diagnosis of 20190-nCov cases relies solely on polymerase chain

58   reaction (PCR), as rapid diagnostic tests for this novel coronavirus are yet to be developed.

59   Second, fifty epidemiological descriptions of the patients who died of 2019-nCov

60   infection were obtained from several sources (7-9). After checking for duplication and missing

61   data, the sample size of delays from onset to death and from hospitalization to death was 39 and

62   33, respectively. We fitted a gamma distribution, an exponential distribution and a lognormal

63   distribution to the distributions and selected the best model based on the AIC. (Figure S1-S2,

64   Table S1-S2).

65   **Case fatality ratio**

66   Crude CFR is defined as the number of cumulative cases divided by the number of cumulative

67   deaths at a specific point in time. For the estimation of CFR in real time, we employ the delay

68   from hospitalization to death, h_s, which is assumed to be given by $h_s = H(s) – H(s\text{-}1)$ for $s>0$

69   where $H(s)$ follows a gamma distribution with mean 10.1 days and SD 5.4 days, obtained from

70   the available observed data. Let π be the time-delay adjusted case fatality ratio, the likelihood

71   function of the estimate π is

$$L(\pi_{a,t_i}; c_{a,t}, \theta) = \prod_{t_i} \left(\sum_{\substack{t=1 \\ D_{a,t_i}}}^{t_i} c_{a,t} \right) \left(\pi_{a,t_i} \frac{\sum_{t=2}^{t_i} \sum_{s=1}^{t-1} c_{a,t-s} h_s}{\sum_{t=1}^{t_i} c_t} \right)^{D_{t_i}} \left(1 \right.$$

$$\left. - \pi_{a,t_i} \frac{\sum_{t=2}^{t_i} \sum_{s=1}^{t-1} c_{a,t-s} h_s}{\sum_{t=1}^{t_i} c_{a,t}} \right)^{\sum_{t=1}^{ti} c_{a,t} - D_{a,t_i}}$$

72   where $c_{a,t}$ represents the number of new cases with reported day t in area a, and $D_{a,t}$ is the

73   number of new deaths with reported day t_i in area a (10-11). We assume that the observed

74   cumulative death, D_t is the result of the binomial sampling process with probability π.

75   We estimated model parameters using a Monte Carlo Markov Chain (MCMC) method in

76   a Bayesian framework. Posterior distributions of the model parameters were estimated by

77   sampling from the three Markov chains. For each chain, we drew 100,000 samples from the

78   posterior distribution after a burn-in of 20,000 iterations. Convergence of MCMC chains were

79   evaluated using the potential scale reduction statistic (12-13). Estimates and 95% credibility

80   intervals for these estimates are based on the posterior probability distribution of each parameter

81   and based on the samples drawn from the posterior distributions.

82   All statistical analyses were conducted in R version 3.6.1 (R Foundation for Statistical

83   Computing, Vienna, Austria) using the 'rstan' package.

84

85   **Results**

86   As of February 5th, 2020, a total of 28001 cases and 563 deaths of 2019-nCov have been

87   reported in China. Of the 28001 cases reported in China, 11618 cases (37.3%) are from Wuhan

88   City, 10494 cases (33.7%) are from Hubei Province excluding Wuhan City and 8976 cases

89   (28.9%) are from China excluding Hubei Province. Of the 642 deaths in China, 478 deaths

90   (74.5%) are from Wuhan city, 140 deaths (21.9%) are from Hubei Province and only 24 deaths

91   (4%) are from China excluding Hubei Province.

92   Figure 1 illustrates the cumulative cases in (A) Wuhan City, (B) Hubei Province

93   excluding Wuhan City and (C) China excluding Hubei Province, and the cumulative deaths in

94   (D) Wuhan City, (E) Hubei Province excluding Wuhan City and (F) Mainland China excluding

95   Hubei Province, respectively. The curve of the cumulative number of deaths grows after that of

96   the cumulative number of cases, and the increase in the number of deaths in Wuhan City is

97   occurring more rapidly and the associated mortality burden appears to be much higher relative to

98   rest of China, while the cumulative case counts for the three areas in China are similar.

99   Figure 2 displays the observed and model-based posterior estimates of crude CFR in (A)

100   Wuhan City, (B) Hubei Province excluding Wuhan City and (C) Main land China excluding

101   Hubei Province, and model-based posterior estimates of time-delay adjusted CFR in (D) Wuhan

102   City, (E) Hubei Province excluding Wuhan City and (F) Mainland China excluding Hubei

103   Province, January – February in 2020, respectively. Our model-based crude CFR fitted the

104   observed data well throughout the course of the epidemic except for the very early stage. During

105   the course of the outbreak, our model-based posterior estimates of time-delay adjusted CFR take

106   much higher values than the observed crude CFR, excluding at the early stage in Wuhan City

107   and the later stage in China excluding Hubei Province. Our estimates of the time-delay adjusted

108   CFR appear to be decreasing almost consistently in Hubei Province excluding Wuhan City and

109   in China excluding Hubei Province, whereas in Wuhan city, estimates were low at the early stage

110   and then increased and peaked in the midst of the study period, and are now following a

111   decreasing trend similar to the other two areas, reaching estimates around 20%.

112   The latest estimates of the time-delay adjusted CFR are at 18.9% (95%CrI: 17,1-20.8%)

113   in Wuhan City, 6.1% (95%CrI: 5.2-7.1%) in Hubei Province excluding Wuhan City and 0.9%

114   (95%CrI: 0.6-1.3%) in China excluding Hubei Province, while the observed crude CFR takes

115   4.1% (95% Confidence Interval: 3.8%-4.5%) in Wuhan City, 1.3% (95%CI: 1.1-1.6%) in Hubei

116   Province excluding Wuhan City and 0.27 % (95%CI: 0.2-0.4%) in China excluding Hubei

117   Province, respectively. (Figure 3, Table)

118

119   **Discussion**

120   In this paper we have derived estimates of the CFR for the ongoing 2019-nCov outbreak

121   in China. We have estimated time-delay adjusted CFR in three different areas in China and found

122   that the most severely affected areas are Wuhan City as well as the Hubei Province excluding

123   Wuhan City while the rest of China exhibits a milder impact (China excluding Hubei Province).

124   We found that the latest estimates of the delay adjusted CFR in Wuhan City is as high as

125   18.9% (95%CrI: 17.1-20.8%), an estimate that his 3-fold higher than our estimate for Hubei

126   Province excluding Wuhan City and around 20-fold higher than our estimate for China excluding

127   Hubei Province. These findings suggest that the ongoing situation in Wuhan is particularly dire

128   compared to the other affected areas. Upward trend of CFR at the early phase generally indicates

129   a deterioration of an ascertainment bias.

130   However, this upward trend should be interpreted with caution. In particular, the

131   difficulties in diagnosis cases of 2019-Cov because the associated symptoms are not specific.

132   Further, the fraction of 2019-Cov patients who develop mild symptoms including asymptomatic

133   patients is not minor, which complicates detection and diagnosis early after symptoms onset,

134   leading to ascertainment bias (*14-15*). Indeed, out of a total of 566 Japanese returnees who

135   evacuated Wuhan city by government-chartered plane during January 29-31 in 2020, a total of 5

136   asymptomatic and 4 symptomatic 2019-nCov cases were detected after detailed medical

137   examination (*16*). However, considering that this underestimation occurred during the course of

138   outbreak and the number of deaths is relatively accurately reported, the upward trend indicates

139   that the temporal disease burden exceeded the capacity of medical facilities and the surveillance

140   system likely missed many cases during the early phase. Additionally, hospital-based

141   transmission is likely occurring, affecting healthcare workers, inpatients and visitors to medical

142   facilities, which could explain an increase trend and elevated CFR estimates. Indeed, nosocomial

143   outbreaks of MERS and SARS nosocomial outbreaks have been reported (*17-18*). Inpatients with

144   underlying disease or seniors infected in the hospital setting have raised the CFR to values as

145   high as 20% for MERS (*19-20*). More generally, 2019-nCov transmission is facilitated in

146   confined settings. For instance, a large cluster of 2019-Cov secondary infection has been

147   reported aboard a cruise ship, with 61 confirmed cases representing about a fifth of the people

148   tested for the virus, supporting high transmissibility of 2019-nCov in confined conditions (*21*).

149   A downward trend in CFR is suggestive of the extent of improvements in

150   epidemiological surveillance. In addition, this pattern indirectly supports a significant number of

151   mild or asymptomatic cases in Wuhan City, and the underlying transmission might prolong the

152   end of the outbreak or further transmission to other areas unless effective social distancing

153   measures are implemented while a vaccine becomes available. Furthermore, given that both the

154   delay adjusted CFR and crude CFR estimates in Wuhan City are around 20-fold higher than our

155   estimates for China excluding Hubei Province suggests a breakdown in medical care delivery

156   and points to the critical need for urgent medical support in the epicenter of the epidemic.

157   We also found that the estimates of delay adjusted CFR both in in Hubei Province

158   excluding Wuhan City and in China excluding Hubei Province show a time-dependent declining

159   trend as the epidemic progresses. Indeed, a similar trend has been previously reported for the

160   Middle East Respiratory Syndrome (MERS) outbreak in the Republic of Korea in 2015, where a

161   significant fraction of the cases had comorbidities or were senior patients (*16-17*). If the

162   proportion of vulnerable cases is high at the early phase and gets proportionally smaller in the

163   later stage of the outbreak, this partly explains the time dependent decline. However, since the

164   epidemic has yet to peak, this time dependent decrease is likely caused by ascertainment bias.

165   Moreover, both of the latest estimates of the delay adjusted CFR and crude CFR in Hubei

166   Province are around 6-7 fold higher than that in China excluding Hubei Province, where the

167   health care system has not reached peak capacity. This also indicates the need to anticipate for

168   additional medical support to deliver medical care to the most vulnerable and at highest risk of

169   succumbing to the disease from underlying medical conditions.

170   Our study is not free from limitations. First, our CFR estimate is influenced by

171   ascertainment bias and this might influence estimates upwards. For those infectious diseases

172   characterized by a large fraction cases with mild or asymptomatic illness, the infection fatality

173   risk, e.g., the number of death divided by the total infected, is a more appropriate index of

174   disease burden (*22-23*). For this purpose, mass serological surveillance and surveys to assess the

175   presence/absence of symptoms is strongly recommended to disentangle the threat of emerging

176   infectious diseases including 2019-nCov.

177   Second, in our estimation we employed the time delay distribution from illness onset to

178   deaths (N=39), which was obtained from secondary sources, but the available data does not

179   include either the date of illness onset or the confirmed date. For this reason, we utilized the time

180   delay from hospitalization to death (N=33).

181   In conclusion, our latest estimates of the risk of 2019-Cov deaths in China could be as

182   high as 20% in the epicenter of the epidemic whereas this estimate is around 1% in the relatively

183   mildly-affected areas in China as of February 5th, 2020. Because it is likely that the death risk

184   from 2019-nCov is associated with a breakdown of the medical/health system in the absence of

185   pharmaceutical interventions (vaccination and antiviral drugs), enhanced public health

186   interventions including social distancing, quarantine, enhanced infection control in healthcare

187   settings and movement restrictions should be effectively implemented to rapidly contain this

188   epidemic.

189

190   Acknowledgments

199
200

201   References

202   1. World Health Organization. Summary of probable SARS cases with onset of illness from

203   1 November 2002 to 31 July 2003 [cited 2020 Feb 7].

204   https://www.who.int/csr/sars/country/table2004_04_21/en/

205   2. Castillo-Garsow MA, Castillo-Chavez C. J Theor Biol. 2003;224(1):1-8. SARS outbreaks

206   in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control

207   mechanism. Chowell G, Fenimore PW,

208   3. Ghani AC, Donnelly CA, Cox DR, Griffin JT, Fraser C, Lam TH, et al. Methods for

209   estimating the case fatality ratio for a novel, emerging infectious disease. Am J

210   Epidemiol. 2005; 162: 479-486

211   4. Kucharski AJ, Edmunds WJ, Case fatality rate for Ebola virus disease in west

212   Africa.Lancet. 2014; 384: 1260

213   5. Jewell NP, Lei X, Ghani AC, Donnelly CA, Leung GM, Ho LM et al. Non-parametric

214   estimation of the case fatality ratio with competing risks data: an application to Severe

215   Acute Respiratory Syndrome (SARS). Stat Med.;26(9):1982-1998.

216   6. Garske T, Legrand J, Donnelly CA, Ward H, Cauchemez S, Fraser C, et al.Assessing the

217   severity of the novel influenza A/H1N1 pandemic. BMJ. 2009; 339: b2840

218   7. The State Council, The People's Republic of China. [cited 2020 Feb 7].

219   http://www.gov.cn/

220   8. Health Commission of Hubei Province, China. [cited 2020 Feb 7].

221   http://wjw.hubei.gov.cn/

222   9. Health Commission of Wuhan City, Hubei Province, China [cited 2020 Feb 7]

223   http://wjw.hubei.gov.cn/

224   10. Nishiura H, Klinkenberg D, Roberts M, Heesterbeek JA. Early epidemiological

225   assessment of the virulence of emerging infectious diseases: a case study of an influenza

226   pandemic.PLoS One. 2009;4(8):e6852. doi: 10.1371/journal.pone.0006852.

227   11. Tsuzuki S, Lee H, Miura F, Chan YH, Jung SM, Akhmetzhanov AR, Nishiura H.

228   Dynamics of the pneumonic plague epidemic in Madagascar, August to October

229   2017.Euro Surveill. 2017;22(46). doi: 10.2807/1560-7917.ES.2017.22.46.17-00710.

230   12. Gamerman, D. & Lopes, H. F. Markov Chain Monte Carlo: Stochastic Simulation for

231   Bayesian Interference. 2nd edn (Chapman & Hall/CRC, 2006).

232   13. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences.

233   Stat Sci 7, 457?472, doi:10.1214/ss/1177011136 (1992).

234   14. Rothe C, Schunk M, Sothmann P, et al. Transmission of 2019-nCoV Infection from an

235   Asymptomatic Contact in Germany. N Engl J Med. 2020. doi: 10.1056/NEJMc2001468.

236   15. Kai Kupferschmidt. Study claiming new coronavirus can be transmitted by people

237   without symptoms was flawed. Feb.3,2020, 5:30 PM   Available from:

238   https://www.sciencemag.org/news/2020/02/paper-non-symptomatic-patient-transmitting-

239   coronavirus-wrong [Accessed February 4, 2020]

240   16. Ministry of Health, Labour and Welfare, Japan. Available from

241   https://www.mhlw.go.jp/index.html [in Japanese]

242   17. Chowell G, Abdirizak F, Lee S, Lee J, Jung E, Nishiura H, Viboud C. Transmission

243   characteristics of MERS and SARS in the healthcare setting: a comparative study. BMC

244   Med. 2015;13:210. doi: 10.1186/s12916-015-0450-0.

245   18. Abdirizak F, Lewis R, Chowell G. Evaluating the potential impact of targeted vaccination

246   strategies against severe acute respiratory syndrome coronavirus (SARS-CoV) and

247   Middle East respiratory syndrome coronavirus (MERS-CoV) outbreaks in the healthcare

248   setting. Theor Biol Med Model. 2019 ;16(1):16. doi: 10.1186/s12976-019-0112-6.

249   19. Mizumoto K, Endo A, Chowell G, Miyamatsu Y, Saitoh M, Nishiura H. Real-time

250   characterization of risks of death associated with the Middle East respiratory syndrome

251   (MERS) in the Republic of Korea, 2015. BMC Med. 2015;13:228. doi: 10.1186/s12916-

252   015-0468-3.

253   20. Mizumoto K, Saitoh M, Chowell G, Miyamatsu Y, Nishiura H. Estimating the risk of

254   Middle East respiratory syndrome (MERS) death during the course of the outbreak in the

255   Republic of Korea, 2015. Int J Infect Dis. 2015;39:7-9. doi: 10.1016/j.ijid.2015.08.005.

256   21. Gale A, Bhattacharya S. Japan Reports 41 New Coronavirus Infections on Quarantined

257   Cruise Liner. The wall street Journal. Feb. 7, 2020. https://www.wsj.com/articles/japan-

258   reports-41-new-coronavirus-infections-on-quarantined-cruise-liner-11581045439

259   22. J.Y. Wong, P. Wu, H. Nishiura, E. Goldstein, E.H. Lau, L. Yang, et al. Infection fatality

260   risk of the pandemic A(H1N1)2009 virus in Hong Kong Am J Epidemiol., 177 (8)

261   (2013), pp. 834-840

262   23. Presanis AM, De Angelis D; New York City Swine Flu Investigation Team, Hagy A,

263   Reed C, Riley S, Cooper BS, et al. The severity of pandemic H1N1 influenza in the

264   United States, from April to July 2009: a Bayesian analysis. PLoS Med.

265   2009;6(12):e1000207.

266

267

268   Address for correspondence: Kenji Mizumoto, Graduate School of Advanced Integrated

269   Studies in Human Survivability, Kyoto University, Yoshida-Nakaadachi-cho, Sakyo-ku,

270   Kyoto 606-8306, Japan; email: mizumoto.kenji.5a@kyoto-u.ac.jp

271

272   Table. Summary results of time-delay adjusted case fatality ratio of 2019-nCov in the 3 areas in China,
273   2020 (As of February 6, 2020)

Area	Latest estimate	Range of median estimates during the study period	Crude CFR
Wuhan City	18.9% (95%CrI§: 17.1-20.8%)	4.1%-34.7%	4.1% (95%CI!: 3.8-4.5%) 478/11618
Hubei Province excluding Wuhan City	6.1% (95%CrI: 5.2-7.1%)	6.1%-88.5%	1.3 % (95%CI: 1.1-1.6%) 140/10494
Mainland China excluding Hubei Province	0.9% (95%CrI: 0.6-1.3)	0.67%-55.9%	0.27% (95%CI: 0.2-0.4%) 24/8976

274   §CrI: 95% credibility intervals (CrI), !95%CI: 95% confidence interval

275

276

277 **Fig 1: Temporal distribution of cases and deaths by area due to 2019-nCov, China, January –**

278 **February 2020.**

279 Cumulative cases in (A) Wuhan City, (B) Hubei Province excluding Wuhan City and (C) Main land China

280 excluding Hubei Province, and cumulative deaths in (D) Wuhan City, (E) Hubei Province excluding

281 Wuhan City and (F) Main land China excluding Hubei Province, over time. Day 1 corresponds to January

282 1st in 2020. As the dates of illness onset were not available, we used dates of reporting.

283

284 **Fig 2: Temporal variation of risk of death caused by 2019-nCov, China, January – February 2020.**

285 Observed and posterior estimated of crude case fatality ratio in (A) Wuhan City, (B) Hubei Province

286 excluding Wuhan City and (C) Main land China excluding Hubei Province, and time-delay adjusted case

287 fatality ratio in (D) Wuhan City, (E) Hubei Province excluding Wuhan City and (F) Main land China

288 excluding Hubei Province. Day 1 corresponds to January 1st in 2020. Black dots shows crude case fatality

289 ratio, and light and dark indicates 95% and 50% credible intervals for posterior estimates, respectively.

290

291 **Figure 3. Latest estimates of time-delay adjusted risk of death caused by 2019-nCov by area, 2020,**

292 **China.**

293 Distribution of time-delay adjusted case fatality risks derived from the latest estimates (February 6, 2020)

294 are presented. Top: Wuhan City, Middle: Hubei Province excluding Wuhan City, Bottom: Mainland China

295 excluding Hubei Province.

296

297 **Appendix Table S1. Summary of delay distribution from hospitalization to death, 2020, China.**

298 Gamma, exponential and exponential distribution were fitted to the available observed data, and

299 estimated parameters and AIC for each distribution are presented (N=33).

300

301   **Appendix Table S2. Summary of delay distribution from illness onset to death, 2020, China.**

302   Gamma, exponential and exponential distribution were fitted to the available observed data, and

303   estimated parameters and AIC for each distribution are presented (N=39).

304

305   **Appendix Figure S1. Delay distribution from hospitalization to death, 2020, China.**

306   Gamma (dark, dashed), exponential distribution (light black, dashed) and kernel density distribution

307   (black, full) were fitted to the available observed data (N=33).

308   **Appendix Figure S2. Delay distribution from illness onset to death, 2020, China.**

309   Gamma (dark, dashed), exponential distribution (light black, dashed) and kernel density distribution

310   (black, full) were fitted to the available observed data (N=39).
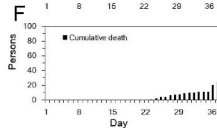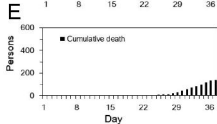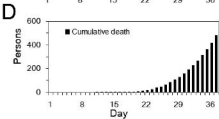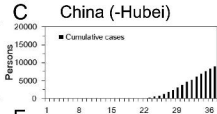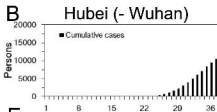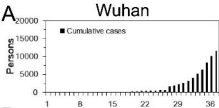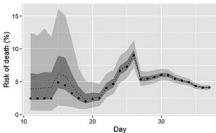
311
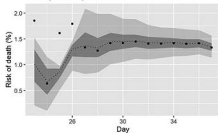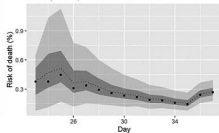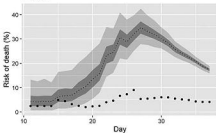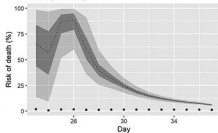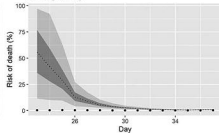
Time delay adjusted CFR by area
The latest estimate