

1 **No excess mortality detected in rural Bangladesh in 2020**

2 **from repeated surveys of a population of 81,000**

3 Prabhat Barnwal^{1,2}, Yuling Yao^{1,3}, Yiqian Wang², Nishat Akter Juy⁴, Shabib Raihan⁴,

4 Mohammad Ashraful Haque⁴, Alexander van Geen⁵

5 **ABSTRACT**

6 **Background** Excess mortality has demonstrated under-counting of COVID-19 deaths in many countries
7 but cannot be measured in low-income countries where civil registration is incomplete.

8 **Methods** Enumerators conducted an in-person census of all 16,054 households in a sample of 135
9 villages within a 350 km² region of Bangladesh followed by a census conducted again in May and
10 November 2020 over the phone. The date and cause of any changes in household composition, as well as
11 changes in income and food availability, were recorded. For analysis, we stratify the mortality data by
12 month, age, gender, and household education. Mortality rates were modeled by Bayesian multilevel
13 regression and the strata aggregated to the population by poststratification.

14 **Results** A total of 276 deaths were reported between February and the end of October 2020 for the subset
15 of the population that could be contacted twice over the phone, slightly below the 289 deaths reported for
16 the same population over the same period in 2019. After adjustment for survey non-response and
17 poststratification, 2020 mortality changed by -8% (95% CI, -21% to 7%) relative to an annualized
18 mortality of 6.1 per thousand in 2019. However, salaried breadwinners reported a 40% decline in income
19 and businesses a 60% decline in profits in May 2020.

20 **Discussion** All-cause mortality in the surveyed portion of rural Bangladesh was if anything lower in
21 2020 compared to 2019. Our findings suggest various restrictions imposed by the government limited the
22 scale of the pandemic, although they need to be accompanied by expanded welfare programs.

23 ¹ P.B. and Y.Y. contributed equally to the study.

24 ² Department of Economics, Michigan State University, USA.

25 ³ Department of Statistics, Columbia University, USA.

26 ⁴ Innovations for Poverty Action, Bangladesh.

27 ⁵ Lamont-Doherty Earth Observatory, Columbia University, USA. Correspondence should be addressed
28 to A.v.G., avangeen@ldeo.columbia.edu.

29

30

31 **Key questions**

32 **What is already known?**

33 Civil registry data from dozens of countries, where available, indicate gaps between official death counts
34 attributed to COVID-19 and, usually, a larger increase in total mortality in 2020 compared to previous
35 years. This approach is not available to gauge the impact of COVID-19 in countries such as Bangladesh
36 where the civil registry system is slow and coverage incomplete. One year after the first COVID-19 case
37 was reported in Bangladesh in 2020, the number of deaths attributed to COVID-19 was equivalent to 1%
38 of annual mortality in previous years. Whether this low figure compared to many other countries is an
39 accurate reflection of the situation or is distorted by massive under-counting has been much debated,
40 albeit on the basis of little direct evidence. The lack of accurate mortality data has made it only more
41 difficult for policy makers to balance the public health benefit of lockdowns and similar measures relative
42 to the well-documented economic costs and hardship imposed by such measures on poor households in
43 particular. A PubMed search conducted on May 4, 2021 under (Bangladesh[Title/Abstract]) AND (excess
44 mortality[Title/Abstract]) limited to 2020-21 did not yield a single relevant study.

45 **What are the new findings?**

46 By conducting of repeated census of a large rural population over the course of 2020, once in person and
47 twice over the phone, we document if anything a slight decline in mortality across a rural area of
48 Bangladesh compared to 2019. We also place an upper limit on the level of under-reporting at the national
49 level that is consistent with our observations. At the same time, interviewed households reported a large
50 and sustained drop in income as well as reduced access to food.

51 **What do the new findings imply?**

52 The impact of the pandemic on mortality was thankfully limited in rural study area of Bangladesh in
53 2020. This suggests that officially recorded COVID-19 deaths may have been contributed largely by the
54 urban population, about a third of the country overall. At the same time, the economic and nutritional
55 impact of restrictions on trade and movement was substantial and probably underestimated in the rural

56 population. As cases surge again, as they did in March–April 2021, policy makers may want to consider

57 limiting strict restrictions to urban areas while expanding a financial support throughout the country.

58

59 INTRODUCTION

60 Excess mortality has received increasing attention over the course of the pandemic as a robust measure of
61 the net impact of COVID-19, especially when comparing countries with different reporting systems and
62 different definitions of COVID-19 deaths (1). The approach has revealed under-reporting of COVID-19
63 deaths by an order of magnitude in several countries (2). The key assumption underlying this approach is
64 that mortality data are reliably compiled within a reasonable time frame through a country's civil
65 registration system, however. This is not the case in most low-income countries where, therefore, excess
66 mortality cannot be reliably measured and under-reporting of COVID-19 deaths is particularly difficult to
67 assess (3–5).

68
69 The lack of reliable excess mortality data in low-income countries is a serious issue for at least two
70 reasons. First, citizens, policy makers, and donor agencies cannot know the actual impact of COVID-19
71 without reliable mortality data and act accordingly (6). Second, recent surveys conducted across several
72 low-income countries have shown that the poor have been particularly hurt economically by the pandemic
73 (7). Lack of reliable mortality data makes it even more difficult for policy makers to gauge the trade-off
74 between the economic costs and public health benefits of a blanket lockdown or other more targeted
75 restrictions.

76
77 During the year that ended March 1, 2021, a total of 8,400 deaths in Bangladesh had been officially
78 attributed to COVID-19 (8). This corresponds to 1% of the 820,000 annual deaths calculated for a total
79 population of 167 million and an annual mortality rate of 4.9 per 1000 estimated from the government's
80 sample vital statistics data (9). However, only a third of deaths are currently officially recorded within 45
81 days in Bangladesh (10). Considering the high excess mortality estimated from burial records in some
82 countries without a functioning civil registry, such as Indonesia (11) and Yemen (12) for instance, the
83 actual number of COVID-19 deaths could be considerably higher than the national data suggest. The
84 possibility of under-reporting of a first wave of COVID-19 cases and fatalities in South Asia has been

85 widely discussed but on the basis of limited direct evidence (5,13,14). The goal of this study was
86 determine whether the low number of official COVID-19 deaths in Bangladesh in 2020 could have been
87 the result of massive under-counting. The issue is salient for policy makers in Bangladesh as the COVID-
88 19 cases and mortality increased rapidly starting in early March 2021.

89
90 According to government data, a majority of COVID-19 cases confirmed by rRT-PCR testing have been
91 concentrated in the capital Dhaka and the next largest city Chattogram (figure 1). Thousands of cases
92 have also been reported from more remote districts, however. The pattern leaves no doubt that the
93 pandemic has spread across the entire the country. The low number of official COVID-19 deaths
94 therefore cannot be attributed to effective isolation of the most affected areas from the rest of the country.
95 The first confirmed cases of COVID-19 were reported on March 8, 2020. In spite of a de facto lockdown
96 declared on March 23, the number of confirmed cases rose rapidly to reach a maximum over 20,000 per
97 week in June 2020 (figure 2a). Over much of the summer, the number of deaths attributed to COVID-19
98 hovered around 250 per week. Both cases and deaths then gradually declined through February 2021,
99 with the exception of a temporary resurgence in November 2020. Since early March 2021, both COVID-
100 19 cases and mortality have increased rapidly to levels above their peaks reached in June 2020.

101
102 Using detailed survey data containing individual level information for a sizeable population, we estimate
103 here month-, age-, gender-, and education-specific mortality. We calculate excess mortality relative to
104 2019 after adjusting for key characteristics and averaging over the January 2019 census population. We
105 complement this analysis with data on the economic impact of the pandemic obtained during the same
106 phone surveys.

107

108 **METHODS**

109 **Census survey**

110 Using an approach developed for conflict zones (15), we inferred mortality in our study area in 2020 by
111 repeatedly conducting a census of a large number of households, once in person in early 2020 and two
112 more times over the phone later in the year. The in-person census involved a team of 60 enumerators
113 working in the field for two weeks; each of the phone surveys employed full-time another 60 enumerators
114 for an entire month. At the beginning of each phone call, each consenting respondent was asked to list
115 current household members without prompting. The electronic questionnaire was set up for the
116 enumerator to check off each household member recorded in January 2020 and identify discrepancies to
117 investigate subsequently. As a reference for identifying excess mortality in 2020, each household was
118 asked at the end of the first phone census about all deaths that occurred in 2019. Only for 2019 was a
119 respondent asked directly about a death that occurred in the household.

120
121 Repeated census surveys were conducted in 135 villages, or pre-defined portions (“paras”) of larger
122 villages, located 30–100 km to the northeast of Dhaka (figure 1). The villages had been selected for a
123 randomized controlled trial, paused in February 2020, to reduce arsenic exposure from drinking well-
124 water. The 135 study villages are distributed across 16 unions, an administrative unit of which there are a
125 total of 4,563 in Bangladesh. Government data show that the average age in these 16 unions is 3 years
126 higher than in the 2,188 other rural unions of the country and that the proportion of households engaged
127 in agricultural activity is somewhat lower (Table 1). Proxies for socioeconomic status such as education
128 and the number of rooms in the house are no different in the study unions compared to other rural unions.

129
130 From January 15 to February 3, 2020, enumerators sought to contact door-to-door all 17,538 households
131 identified in the study villages or paras. Among these households, 1,478 were absent and could not be
132 reached. Of the remaining 16,060 households, only 6 declined to participate in what was presented at the
133 time as a study of arsenic mitigation. Sharing a kitchen was the criterion used to define a household.

134 Following consent, the name, age, gender, and relationship of each individual member of the household,
135 as well as GPS coordinates of each house, were recorded electronically. Up to two mobile phone numbers
136 from all consenting households were also recorded. Most households were subsequently recontacted
137 using one of these numbers, or in some cases alternative numbers provided by their neighbors.
138 During the first phone survey conducted from May 8 to June 7, 2020, a total of 14,551 (91%) of the
139 households surveyed in person in January 2020 could be contacted and consented to respond. During the
140 second phone survey conducted between October 27 and December 14, 2020, 11,933 (74%) households
141 consented. The total population of consenting households surveyed in person amounts to 81,164. This
142 number is used as the reference population and includes 7,921 household members recorded during the
143 phone surveys who had reportedly been overlooked during the January 2020 in person census, as well as
144 484 deaths in the households that occurred reportedly in 2019. The reference population does not include
145 1,068 household members who joined the household from elsewhere in 2020.

146
147 Additional information related to 2019 and 2020 deaths was collected during the phone surveys.

148 Respondents were asked if a reported death was the result of injury, e.g. a road accident, and if treatment
149 from a doctor or at a hospital was sought. Respondents were asked if death was preceded by symptoms
150 related to COVID-19 such as fever, headache, cough, sore throat, breathing difficulty, loss of sense of
151 smell, muscle aches, and chills. Respondents were also asked to attribute reported deaths to a few broad
152 categories including stroke or heart disease (combined here because they are often confused in rural
153 Bangladesh), cancer, liver, or lung disease. Some deaths, often in the case of an elderly parent, were
154 reported more than once by different households. A total of 76 such cases were identified over 2019–2020
155 based on name and proximity and then confirmed with an additional phone call to avoid duplication.

156 In order to determine the economic impact of COVID-19 in Bangladesh, we requested additional
157 information from a randomly selected 20% of 16,054 households in the first phone survey and the same
158 20% with an additional 8% of households in the second phone survey. A total of 2,608 households
159 participated in the first survey and 3,151 households in the second.

160

161 **Multilevel regression model of excess mortality and poststratification**

162 We estimate mortality using a multilevel regression model and poststratification for two main reasons:
163 first, to adjust for attrition i.e., non-response in census round 1 or 2, or both, and second, to understand the
164 variation across demographic groups. The underlying assumption is that attrition is a function of age,
165 gender and education. To estimate aggregate mortality, we therefore first estimate mortality for various
166 age, gender, and education group and then calculate a weighted sum.

167 We stratify observed death counts into cells based on a four-way structure: (a) month (22 levels: Jan
168 2019, Feb 2019, ..., Oct 2020), (b) age during the evaluation month (binned into 9 levels: 0–9 years old,
169 10–19, ..., 70–79, 80 and above), (c) gender (2 levels, 1 = female), and (d) household education level (2
170 levels indicating if less (=1) or at least half (=2) of the adult members in the household have ever received
171 some formal education.

172

173 In order to compare mortality before and during the pandemic, and to estimate the resulting excess
174 mortality, we need to specify the local onset of impact of COVID-19 in our study villages. We primarily
175 consider February 2020 as the start month but also consider later months. We specify a comparison-
176 starting month t_{start} , which splits time into the two periods (a) the baseline: Jan 2019, Feb 2019, ...,
177 $t_{\text{start}} - 1$, and (b) the potential excess period: t_{start} , $t_{\text{start}} + 1$, ..., Oct 2020. This approach allows us to
178 be agnostic about when COVID-19 started to impact mortality in the study villages.

179 *Modeling baseline and excess mortality in each strata.*

180 Both baseline and excess mortality may differ across strata. We set up a multilevel logistic regression
181 (17,18) to model baseline mortality in each stratum by a four-way interaction of month-of-year, age, sex
182 and education (Appendix A). To quantify the stratum-specific excess mortality (the additional risk in
183 2020 after month t_{start}), we set another layer of multilevel model, where the monthly excess mortality is
184 decomposed by a three-way interaction of age, sex and education.

185

186 *Partial-pooling using Bayesian inference.*

187 We complete the model with a lag-1 autoregressive (AR) prior on all age effects and month effects, and
188 weakly-informative priors on other coefficients. We perform fully Bayesian inference for our complete
189 model in Stan using 4 chains and 3000 posterior simulation draws for each chain. (19) Computational
190 diagnostics indicate all chains mix well.

191

192 *Aggregating mortality and excess mortality.*

193 To aggregate the stratum-specific mortality to the population of interest, we generate posterior simulation
194 draws of mortality in each stratum. Under the assumption that stratum-specific mortality rates are not
195 affected by attrition in later rounds of census survey, the aggregate mortality is a weighted sum of
196 mortality in each stratum, where weights are constructed using the January 2020 census data, from which
197 we compute the excess mortality in the population by the Monte Carlo method. The poststratified excess
198 mortality in equation [eq_excess_m] implicitly compares the averaged mortality in $t_{\text{start}}, t_{\text{start}} + 1, \dots, \text{Oct}$
199 2020, with the same period in 2019. We repeat the same process to calculate the population-aggregated
200 age-specific excess mortality change (Appendix A).

201

202 **RESULTS**

203 *Raw data.*

204 Considering first only the 11,256 households that could each be reached during both phone surveys, a
205 total of 639 deaths were reported between January 2019 and the end of October 2020 for a total
206 population of 58,806, excluding individuals who joined the household after the January 2020 census. This
207 corresponds to an average annualized mortality rate of 5.9 per 1000 over 22 months. For the same
208 households that could be reached twice over the phone, a total of 276 deaths were reported between
209 February and the end of October 2020, slightly below the 289 deaths reported for over the same months in
210 2019. Using all the available data, the monthly-sample-weighted annualized mortality rate between

211 January 2019 and Oct 2020 is 6.1 per thousand (figure 2). The challenge is to determine the extent to
212 which mortality differed before and after the pandemic by taking into account all the available data,
213 including households that could not be reached after the January 2020 in-person census.
214 Broken down by month, the annualized sample mortality rate in the study population varied between 3.5
215 and 11 deaths per 1000 since January 2019, but without a clear seasonal pattern (figure 2b). A longer time
216 series from official statistics suggests that mortality in rural Bangladesh is on average 20% higher during
217 winter compared to summer (20). The highest mortality rate during our study period was recorded in May
218 2019, but this could also be an artifact of the first phone survey being conducted in May 2020 and
219 households reporting that a member died about a year ago. The data do not show an increase in mortality
220 that coincides with the peak in COVID-19 cases and deaths reported centered on late May 2020 (figure
221 2a). The second round of phone surveys only partly covers the second peak in COVID-19 cases and
222 deaths centered on November 2020. The age distribution of the surveyed population (figure 2) is
223 comparable to published national trends (9). The number of deaths as a function age over the entire study
224 period shows an initial drop beyond ages 0–5 followed by a steady increase in the number of deaths to
225 age 80 and above, in spite a diminishing contribution to the overall population. Beyond these simple
226 compilations, modeling is required to infer representative mortality trends.

227
228 The model confirms the expected increase in mortality as a function of age, expressed as a log odds ratio
229 (figure 3a). There is no clear seasonal pattern in mortality in our data according to the model (figure 3b).
230 Excess mortality is the key output from the model. At first, we assume that the onset of the pandemic in
231 the study villages was in February 2020. The model shows no excess mortality during February–October
232 2020 and, if anything, possibly a slightly lower mortality log odds ratio at the lowest and highest ages
233 (figure 3c). The model also suggest that excess mortality was likely higher for men compared to women
234 across the entire range of ages (figure 3d).

235

236 After poststratification across gender and education, the model shows no indication of an increase in
237 mortality rate during Feb–Oct for 2020 compared to 2019 in all age group (figure 4a). Comparing Feb–
238 Oct 2020 with the baseline using estimate [eq_excess_m2] in Appendix A, the monthly mortality changes
239 by -0.04 (95% CI, -0.11 to 0.04), deaths per thousand people per month averaged over all ages, and -0.12
240 (95% C.I., -0.52 to 0.29) at ages 50 and older. The baseline monthly baseline mortality is estimated to be
241 0.51 (.41–.63) for all age average, or 2.4 (1.9–3.0) deaths per thousand per month. Hence, using estimate
242 [eq_excess_m3] in Appendix A, our inferred mortality declines amount to -8% (-21% to +7%) or -5% (-
243 21% to +12%) in percentage changes. This overall decline was largely due to the mortality decline in the
244 80+ group. Excess mortality does not vary much with household education level (a proxy for
245 socioeconomic level) and gender.

246
247 The excess mortality changes as the boundary between the two comparison window is shifted to later in
248 2020. Comparing Aug–Oct 2020 to the baseline, the model indicates that monthly mortality rate increased
249 by 0.10 (95% CI, -0.06 to 0.30), death per thousand people per month). For age 50+, the monthly
250 mortality increased by 0.58 (95% CI, -0.23 to 1.60). The inferred posterior mean corresponds to a 20%
251 (25%) increase in percentage in all (50+) population, with considerable uncertainty.

252
253 The circumstances and causes of a total of 795 deaths reported for 2019–20 during the two rounds of
254 phone calls did not vary much over time. In 2019 and 2020, respectively, 73 and 67% of deaths were
255 preceded by consultation with a doctor or nurse, 17 and 18% of deaths occurred at a hospital, and 6.9 and
256 7.6% were the result of injury. Heart disease and stroke combined reportedly caused 41 and 47% of
257 deaths in 2019 and 2020, respectively, and cancer reportedly caused 9.0 and 9.1% of deaths. The
258 proportion of deaths attributed to lung disease actually went down from 6.4 to 3.7%. Among the 30
259 deaths attributed to lung disease in 2019, COVID-19 related symptoms such as fever, headache, cough,
260 sore throat, breathing difficulty, loss of sense of smell, muscle aches, and chills were reported 64 times.
261 Among the 12 deaths attributed to lung disease in 2020, the same symptoms were reported 30 times.

262

263 *Economic impacts.*

264 The lockdown had a large economic impact on our study population. Survey responses show that
265 household income where the main bread earner had a salaried job declined by 40% on average in May
266 2020, although the decline was reduced to 30% by November 2020 (figure 6b). For self-employed bread
267 earners in business, the decline in profit was 60% and barely recovered later in the year. Over the same
268 period, 25% of households reported that they couldn't obtain an essential food item in May 2020 because
269 of reduced income, although this proportion had declined to 9% by November 2020. Even when the
270 mobility of people reverted to the pre-pandemic levels, the negative economic impact of the pandemic
271 was not proportionally reversed.

272

273 **DISCUSSION**

274 The annualized mortality rate of 6.1 per 1000 calculated over the entire 22 months covered by the study is
275 somewhat higher than the official 2019 estimate for rural Bangladesh of 5.4 per 1000 (9). The difference
276 is plausible given significant geographic differences in rural mortality across the country. The posterior
277 mean of the annualized baseline mortality rates inferred from the model range from a minimum of 0.5 per
278 1000 for the 10–19 age range to a maximum of 150 per 1000 for 80 years and older (figure 3a). These
279 values are also consistent with existing national statistics collected as recently as 2019 (9).

280

281 On the basis of the census data collected on three occasions, once in person and twice over the phone, we
282 conclude that mortality did not increase in 2020 across our 135 study villages or paras. In fact, our best
283 estimate is that mortality declined by 8%. This decrease in mortality can be potentially explained by a
284 decline in mortality from other causes such as road accidents or the seasonal flu caused by reduced travel
285 and social interactions, respectively. On average, the net impact of COVID-19 therefore does not appear
286 to have come close to the levels of excess mortality of 20% and over in 2020 reported for over two dozen
287 countries including the US (2). The 95% confidence interval obtained from the model indicates a one in

288 forty chance that mortality actually increased by more than 7% in our study villages. Applied to the
289 months of February through October 2020 along with an annual mortality of 820,000 for the country
290 overall in 2019 (9), our upper bound corresponds to a one in forty chance of 43,000 additional deaths or
291 more. The official death count attributed to COVID-19 had reached 5,900 by October 31, 2020.
292 Combining these two figures indicates that there is only a one in forty chance that COVID-19 deaths were
293 under-reported by more than a factor of 7. This comparison assume that both COVID-19 mortality and
294 the reporting of COVID-19 mortality were similar in urban and rural areas, which probably was not the
295 case. Our data could suggest that all the reported COVID-19 mortality was limited to urban areas, which
296 account for about 1/3 of the Bangladesh population. Even under this scenario, reported COVID-19
297 mortality would correspond to only about of 3% or urban mortality in previous years.

298
299 We do not claim that our sample of 135 villages is representative of all of rural Bangladesh, although
300 Table 1 provide some reassurance. Another limitation of the study is that a quarter of the households
301 surveyed in January 2020 could not be reached over the phone by November 2020. We cannot exclude
302 that households with a member who recently died is less likely to pick up the phone or less willing to
303 participate in the survey. Our approach to calculate aggregate mortality using demographic group-level
304 mortality corrects for it, to the extent that mortality estimated for a particular age, sex and education
305 group is not biased by attrition. The repeated census approach may also not have entirely eliminated a
306 tendency not to report a recent death, especially if it was associated with COVID-19 symptoms because
307 of widely reported stigma, especially at the beginning of the pandemic (21). We have separate evidence
308 from our survey that COVID-19-like symptoms were under-reported by affected households on the basis
309 of responses from neighboring households integrated at the village and para level.

310
311 Various hypotheses have been proposed to explain the apparently lower impact of COVID-19 in some
312 low-income countries (22). The effect of a relatively young population cannot be a factor in our study
313 given that we are comparing the same population over two years. Spending more time outside or in well-

314 ventilated houses has been invoked as an explanation, but another possibility is that previous infections in
315 regions like rural Bangladesh could have dampened the symptoms of COVID-19 (23).

316
317 Mobility data and our economic impact data both indicate that the reach of the government's
318 interventions extended to rural areas of Bangladesh. The absence of excess mortality in our study
319 population suggests that limiting gatherings, encouraging masks, and maintaining social distance were
320 broadly successful in these areas in 2020 (24). At the same time, our data show that measures restricting
321 work, trade, and travel imposed an economic burden on rural households that extended over at least six
322 months. Similar impacts have previously been reported for 2020, including from a neighboring area of
323 Bangladesh where both a large drop in household income and increased food insecurity were documented
324 (25). There is therefore a critical need for the government of Bangladesh, and low-income countries more
325 generally, to expand their social safety net to the population affected by strong non-pharmaceutical
326 interventions to keep the COVID-19 pandemic under control (7,26).

327
328 A riskier interpretation of our results is that the de facto national lockdown imposed in March 2020 was
329 excessive given the high economic cost and that no excess mortality was observed in our study villages.
330 We do not endorse this view but the government of Bangladesh and other low-income countries might
331 want to consider in the future more regionally targeted lockdowns that distinguish urban and rural areas
332 among other factors. Such a targeted approach crucially depends, however, on monitoring across the
333 country and, therefore, on the widespread availability of COVID-19 testing. The rapid growth in the pace
334 testing at the beginning of the pandemic was arrested after the imposition of a charge for testing of
335 BDT200 (US\$2.40) at government facilities, a charge halved since, and BDT500 (US\$6.00) for samples
336 collected from home (27). This suggest free COVID-19 testing should be made available again
337 throughout the country and that increased testing capacity will be needed to handle the likely surge in
338 demand. By analogy to the impact free well-testing for arsenic had in rural Bangladesh over a decade ago

339 (28,29), widely available COVID-19 testing could have an additional impact by encouraging health-
340 protective behavior such as avoiding gatherings, wearing a mask, and maintaining social distance.
341 The lessons from the pandemic in 2020 remain relevant in 2021 and may remain so for a while to come.
342 When the number of COVID-19 cases and deaths started to increase rapidly in March 2021, the
343 government of Bangladesh faced a difficult decision between the economic cost of a new national
344 lockdown and the public health cost of not imposing restrictions (figure 2). A lockdown was re-imposed
345 in early April 2021 and, unlike in neighboring India (30), the worrisome trends fortunately reversed by
346 the beginning of May 2021.

347

348 **CONCLUSIONS**

349 Where national statistics make this possible, excess mortality data have provided a powerful measure of
350 the impact of COVID-19. In spite of their limitations, our repeated census surveys of a sizeable rural
351 population fill this gap for a country where such statistics are not available. For reasons that deserve
352 further study, it does not appear that COVID-19 had an impact on mortality in rural Bangladesh in 2020
353 anywhere close to that of the on-going pandemic in many other countries. While this suggests restrictions
354 imposed by the government had the desired effect, our study also highlights the need for corrective public
355 programs to address the very high economic burden resulting from such measures.

356 **Acknowledgments** We thank the large teams of enumerators involved in the field and phone surveys.
357 We particularly thank the residents of our study villages whose patience in a stressful situation was tested
358 by a number of long phone calls. A number of colleagues including Joel Cohen (Rockefeller University
359 and Columbia University), Mushfiq Mobarak (Yale University), Stephen Luby (Stanford University),
360 Jeffrey Shaman (Columbia University), and Moyeen Uddin (Civil Registration and Vital Statistics
361 Initiative, Government of Bangladesh) provided helpful advice over the course of the study.

362
363 **Contributors** PB and AvG conceived the original arsenic mitigation study and adapted it with YW,
364 NAJ, and SR to reconstruct mortality. NAJ, SR, MAH recruited, trained, and supervised the enumerators.
365 The data were checked for errors and compiled by NAJ, SR, and YW. YY conducted the statistical
366 analysis with input from PB, YW, NAJ, and AvG. AvG, YY, and PB drafted the manuscript, which was
367 edited based on input from all co-authors.

368
369 **Funding** The arsenic mitigation trial that set the stage for this study was supported by NSF SBE awards
370 1851928 and 1853289.

371
372 **Competing interests** None declared.

373
374 **Ethics approval.** Institutional Review Boards at both Michigan State University and Columbia
375 University approved the study.

376
377 **Data sharing**
378 Data and code are available at <https://github.com/yao-yl/mortalityPaper>. All confidential information was
379 removed from the posted survey data.

380

381 **REFERENCES**

- 382 1. Beaney T, Clarke JM, Jain V, Golestaneh AK, Lyons G, Salman D, et al. Excess mortality: The gold
383 standard in measuring the impact of COVID-19 worldwide? *Journal of the Royal Society of*
384 *Medicine*. 2020;113(9):329–34.
- 385 2. Karlinsky A, Kobak D. The world mortality dataset: Tracking excess mortality across countries
386 during the COVID-19 pandemic. *medRxiv*. 2021; doi: 10.1101/2021.01.27.21250604
- 387 3. Setel PW, Macfarlane SB, Szreter S, Mikkelsen L, Jha P, Stout S, et al. A scandal of invisibility:
388 Making everyone count by counting everyone. *Lancet*. 2007;370(9598):1569–77.
- 389 4. Mikkelsen L, Phillips DE, AbouZahr C, Setel PW, De Savigny D, Lozano R, et al. A global
390 assessment of civil registration and vital statistics systems: Monitoring data quality and progress.
391 *The Lancet*. 2015;386(10001):1395–406.
- 392 5. Chatterjee P. Is India missing COVID-19 deaths? *Lancet*. 2020;396(10252):657.
- 393 6. Checchi F, Roberts L. Documenting mortality in crises: What keeps us from doing better? *PLoS*
394 *Med*. 2008;5(7):e146.
- 395 7. Egger D, Miguel E, Warren SS, Shenoy A, Collins E, Karlan D, et al. Falling living standards during
396 the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances*.
397 2021;7(6):eabe0997.
- 398 8. World Health Organization. WHO COVID-19 dashboard [Internet]. 2021. Available from:
399 {<https://covid19.who.int>}
- 400 9. Bangladesh Bureau of Statistics. Report on sample vital registration system (SVRS) 2019 [Internet].
401 2020. Available from:
402 {<https://drive.google.com/file/d/1TtdcJaDyc7vf5u7Aza8GmqgGv8eqP6JN/view>}
- 403 10. Uddin M, Ashrafi SAA, Azad AK, Chowdhury A, Chowdhury HR, Riley ID, et al. Improving
404 coverage of civil registration and vital statistics, Bangladesh. *Bulletin of the World Health*
405 *Organization*. 2019;97(9):637.

- 406 11. Elyazar IR, Surendra H, Ekawati L, Djaafara BA, Nurhasim A, Hidayana I, et al. Excess mortality
407 during the first ten months of COVID-19 epidemic at Jakarta, Indonesia. medRxiv. 2020; :
408 <https://doi.org/10.1101/2020.12.14.20248159>.
- 409 12. Besson ESK, Norris A, Ghouth ASB, Freemantle T, Alhaffar M, Vazquez Y, et al. Excess mortality
410 during the COVID-19 pandemic: A geospatial and statistical analysis in Aden governorate, Yemen.
411 BMJ Global Health. 2021;6(3):e004564.
- 412 13. Biswas RK, Afiaz A, Huq S. Underreporting COVID-19: The curious case of the Indian
413 subcontinent. *Epidemiology & Infection*. 2020;148.
- 414 14. Mukherjee S. Why does the pandemic seem to be hitting some countries harder than others? The
415 New Yorker.; March 1, 2021.
- 416 15. Hygiene London School of, Medicine Tropical. The use of epidemiological tools in conflict-affected
417 populations: Open-access educational resources for policy-makers. [Internet]. 2009. Available from:
418 {http://conflict.lshtm.ac.uk/page_02.htm}
- 419 16. Bangladesh Bureau of Statistics. Household income and expenditure survey (HIES). 2016.
- 420 17. Gelman A, Little TC. Poststratification into many categories using hierarchical logistic regression.
421 *Survey Methodology*. 1997;23:127–35.
- 422 18. Park DK, Gelman A, Bafumi J. Bayesian multilevel estimation with poststratification: State-level
423 estimates from national polls. *Political Analysis*. 2004;375–85.
- 424 19. Stan Development Team. Stan user’s guide [Internet]. 2021. Available from: {<https://mc-stan.org>}
- 425 20. Becker S, Weng S. Seasonal patterns of deaths in Matlab, bangladesh. *International journal of*
426 *epidemiology*. 1998;27(5):814–23.
- 427 21. Mahmud A, Islam MR. Social stigma as a barrier to COVID-19 responses to community well-being
428 in Bangladesh. *International Journal of Community Well-Being*. 2020;1–7.
- 429 22. De Weerd J, Van Damme W. Health, wealth and the double paradox of COVID-19 mortality in low-
430 income countries. 2021. Available from: {<https://ssrn.com/abstract=3793427>}

- 431 23. Cepon-Robins TJ, Gildner TE. Old friends meet a new foe: A potential role for immune-priming
432 parasites in mitigating COVID-19 morbidity and mortality. *Evolution, Medicine, and Public Health*.
433 2020;2020(1):234–48.
- 434 24. Ge Y, Zhang W, Liu H, Ruktanonchai CW, Hu M, Wu X, et al. Effects of worldwide interventions
435 and vaccination on COVID-19 between waves and countries. *medRxiv* 2021033121254702. 2021;
- 436 25. Hamadani JD, Hasan MI, Baldi AJ, Hossain SJ, Shiraji S, Bhuiyan MSA, et al. Immediate impact of
437 stay-at-home orders to control COVID-19 transmission on socioeconomic conditions, food
438 insecurity, mental health, and intimate partner violence in Bangladeshi women and their families: An
439 interrupted time series. *The Lancet Global Health*. 2020;8(11):e1380–9.
- 440 26. Gerard F, Imbert C, Orkin K. Social protection response to the COVID-19 crisis: Options for
441 developing countries. *Oxford Review of Economic Policy*. 2020;36(Supplement_1):S281–96.
- 442 27. Cousins S. Bangladesh’s COVID-19 testing criticised. *The Lancet*. 2020;396(10251):591.
- 443 28. Madajewicz M, Pfaff A, Geen A van, Graziano J, Hussein I, Momotaj H, et al. Can information
444 alone both improve awareness and change behavior? Arsenic contamination of groundwater in
445 Bangladesh. *Journal of Development Economics* 2007, vol. 84, issue 2, 731-754.
- 446 29. Huhmann BL, Harvey CF, Navas-Acien A, Graziano J, Parvez F, Chen Y, et al. Changes in arsenic
447 exposure in Araihaazar, Bangladesh from 2001 through 2015 following a blanket well testing and
448 education campaign. *Environment International*. 2019;125:82–9.
- 449 30. Mallapaty S. India’s massive COVID surge puzzles scientists. *Nature*. 2021;
450 <https://www.nature.com/articles/d41586-021-01059-y>

451

452

453 **Table 1** Socioeconomic statistics in rural Bangladesh nationwide from the 2016 household income and
454 expenditure survey(16) (distributed across 2,204 unions), and in our sample survey.

455

| Variable | Unions in rural Bangladesh | | | Unions in this study | | |
|-----------------------|----------------------------|-------|-------|----------------------|-------|-------|
| | N | Mean | SD | n | Mean | SD |
| Age | 129,347 | 28 | 19 | 1,394 | 25 | 19 |
| Sex | 129,348 | 0.501 | 0.500 | 1,394 | 0.507 | 0.500 |
| Education | 31,795 | 0.85 | 0.36 | 300 | 0.85 | 0.35 |
| Rooms in house | 31,758 | 2.3 | 1.3 | 300 | 2.4 | 2.5 |
| Agricultural activity | 31,795 | 0.43 | 0.47 | 300 | 0.26 | 0.41 |

456

457 The number of observations (n) corresponds to individuals for age (years) and sex (1=Female, 0=Male),

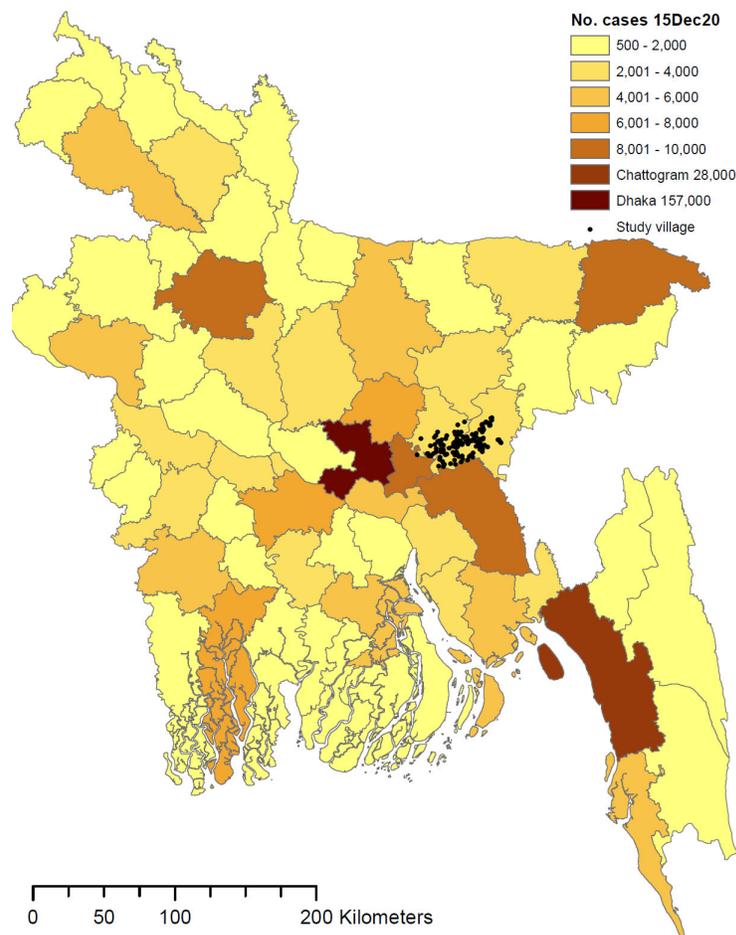
458 and to households for education (1=over half of household members received any formal education),

459 main economic activity (proportion of household members engaged in agricultural sector), and count of

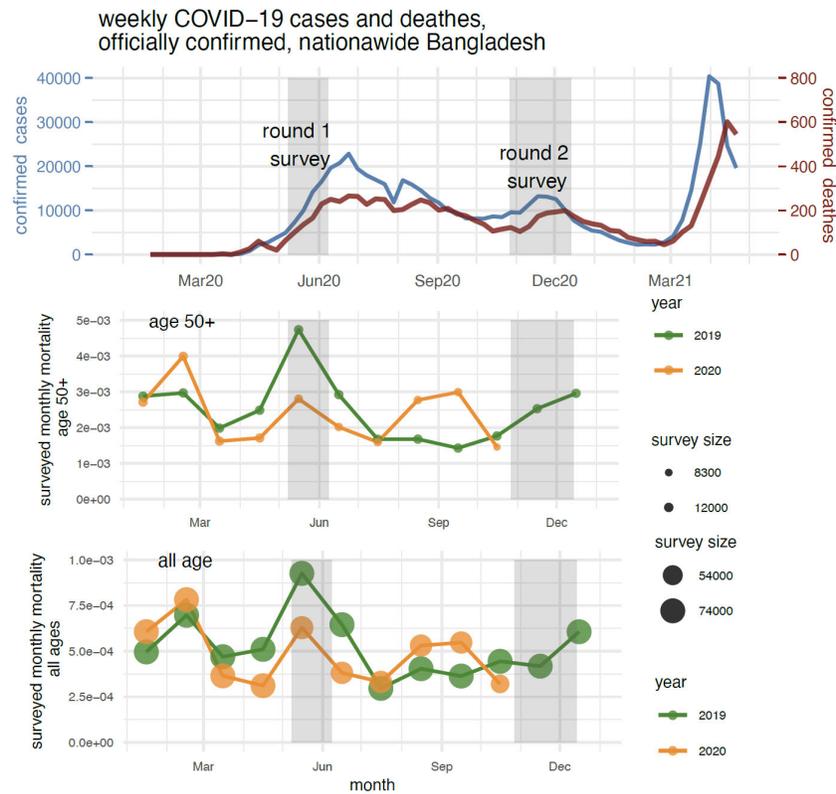
460 rooms in the house.

461

462



463
464 **Figure 1.** Map of Bangladesh showing the distribution of confirmed COVID-19 cases as of December 15,
465 2020. Study villages surveyed in person and over the phone in 2020 are shown as black dots east of the
466 capital Dhaka.
467
468



469

470 **Figure 2.** (a) Weekly time series of confirmed COVID-19 cases (blue curve) and deaths (red curve) in

471 Bangladesh in relation to timing of our phone surveys, from February 2020-April 28, 2021. (b, c)

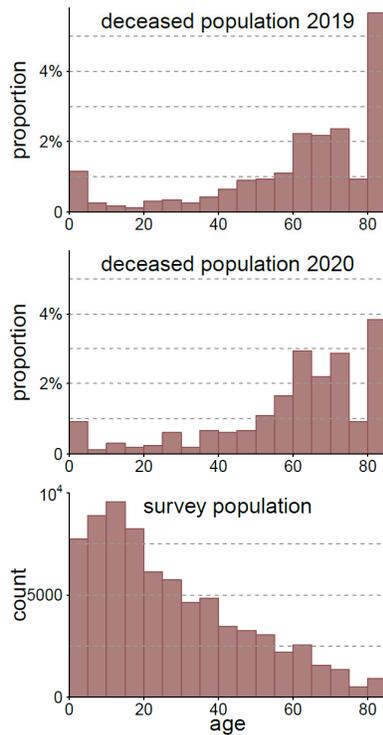
472 Monthly sample mortality rate for elderly (age 50+) and all ages in 2019 and 2020 in study villages. The

473 mortality rate is calculated relative to the size of the survey sample for that month, which is indicated by

474 the dot size.

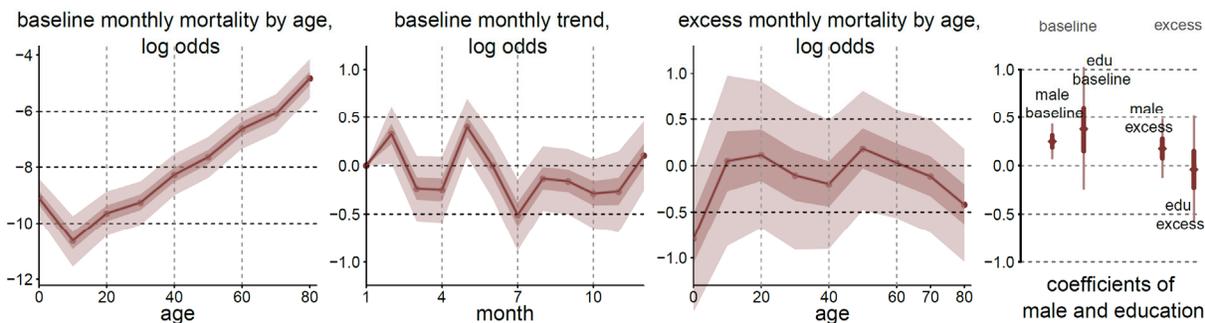
475

476



477
478 **Figure 3.** (a, b) Distribution of age at death in study households in 2019 and 2020. (c) Histogram of ages
479 of all survey individuals in Jan 2020. Reported ages above 80 are combined.

480
481



482

483 **Figure 4.** Decomposition of all modeled factors to baseline and excess mortality in log odds, visualized

484 by the posterior mean, 50% and 95% confidence intervals: (a) the age-specific baseline monthly

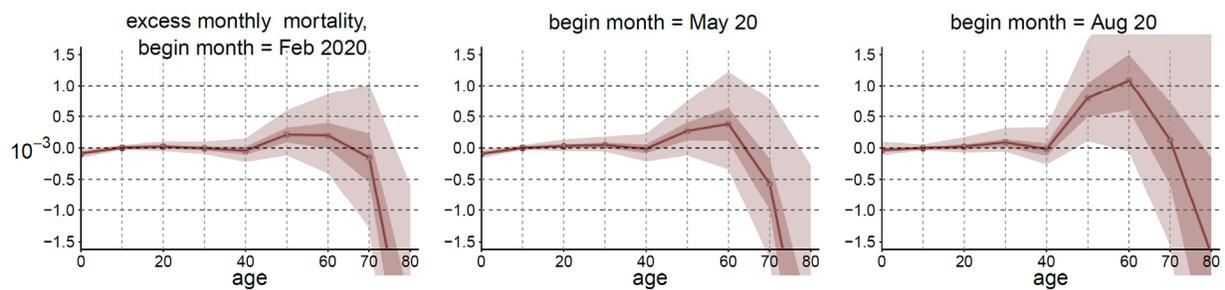
485 mortality log odds across ages; (b) the baseline month-of-year effect of the monthly mortality log odds

486 across months; (c) the age-specific excess monthly mortality during February-October 2020; (d) the

487 baseline and excess log odds associated with male or family education level.

488

489

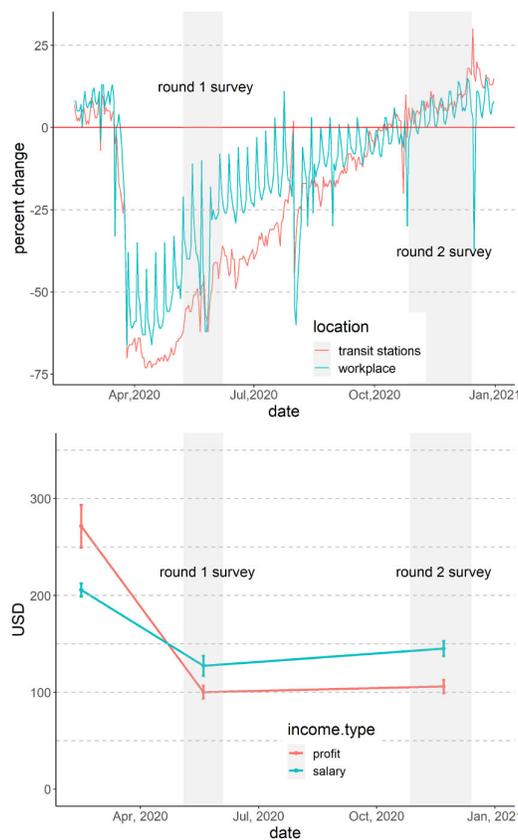


490

491 **Figure 5.** The age-specific average excess monthly mortality rate and its confidence interval. The
492 comparison window always ends in Oct 2020 while the starting month varies from February 2020 to
493 August 2020.

494

495



496

497 **Figure 6.** (a) Mobility change: the percentage change in mobility provided by Google COVID-19

498 Community Mobility Reports. The baseline mobility is the median of mobility index from January 3-

499 February 6, 2020. (b) Income change: self-reported job salary or business profit from pre-COVID to

500 November 2020.

Appendix A: Model details

Baseline mortality rate. We use the following notation for the four discrete variables: month $j_1 = 1, \dots, 22$; age during the evaluation month, $j_2 = 1, \dots, 9$; gender $j_3 = 1, 2$; and household education $j_4 = 1, 2$. Any individual stratum can be written as $j := (j_1, j_2, j_3, j_4)$. In each stratum j , we count n_j the number of surveyed individuals that were alive in the beginning of the month and accessible throughout the month, and y_j the number of deceased individuals. Assuming independent sampling, the data model is

$$y_j \sim \text{Binomial}(n_j, \theta_j),$$

where the parameter θ_j is what we want to estimate: the monthly mortality rate in month j_1 for age group j_2 , gender j_3 , and household education level j_4 .

The baseline ($j_1 < t_{\text{start}}$) mortality rate for stratum j is modeled as a function of month and individual's age-sex-education attributes,

$$\theta_j = \text{logit}^{-1} (\alpha_{[j_1 \bmod 12]} + \gamma_{[j_2]} + \mathbb{1}(\gamma_{j_3} = 2) \mu_{\text{male}} + \mathbb{1}(\gamma_{j_4} = 2) \mu_{\text{edu}}), \quad \forall j_1 < t_{\text{start}}. \quad (1)$$

where the free parameters are

- The month-of-year factor, or the seasonal trend, denoted by $\alpha_1, \dots, \alpha_{12}$ for January to December each year. Due to periodicity, cell j processes this seasonal factor $\alpha_{[j_1 \bmod 12]} \in \mathbb{R}$. For identification, we set January as baseline such that $\alpha_1 = 0$.
- The age factor, denoted by $\gamma_1, \dots, \gamma_9$ for 9 age categories. The age factor for cell j is $\gamma_{[j_2]} \in \mathbb{R}$.
- The male factor $\mu_{\text{male}} \in \mathbb{R}$. We set female as reference, then the sex factor for cell j is $\mathbb{1}(\gamma_{j_3} = 2) \mu_{\text{male}}$.
- The education factor $\mu_{\text{edu}} \in \mathbb{R}$. We set non-education as reference, and the education factor for cell j is $\mathbb{1}(\gamma_{j_4} = 2) \mu_{\text{edu}}$.

Mortality in 2020. We take a flexible approach by specifying a comparison-starting month t_{start} , which splits time into the two periods (a) the baseline: Jan 2019, Feb 2019, \dots , $t_{\text{start}} - 1$, and (b) the potential excess period: t_{start} , $t_{\text{start}} + 1, \dots$, Oct 2020. During the second period, on top of the baseline model (1), we model the stratum-specific excess risk:

$$\theta_j = \text{logit}^{-1} (\alpha_{[j_1 \bmod 12]} + \gamma_{[j_2]} + \mathbb{1}(\gamma_{j_3} = 2) \mu_{\text{male}} + \mathbb{1}(\gamma_{j_4} = 2) \mu_{\text{edu}} + \gamma_{[j_2]}^{\text{excess}} + \mathbb{1}(\gamma_{j_3} = 2) \mu_{\text{male}}^{\text{excess}} + \mathbb{1}(\gamma_{j_4} = 2) \mu_{\text{edu}}^{\text{excess}}), \quad \forall j_1 \geq t_{\text{start}}, \quad (2)$$

where additional parameters γ^{excess} , $\mu_{\text{male}}^{\text{excess}}$ and $\mu_{\text{edu}}^{\text{excess}}$ represent the excess risk in 2020 associated in the age, sex, and education, on top of the baseline risk.

Partial-pooling and priors. We complete the model by weakly-informative priors:

$$\begin{aligned} \alpha_i - \alpha_{i-1} &\sim \text{normal}(0, \tau_1), \quad i = 2, \dots, 12; \quad \alpha_1 - \alpha_{12} \sim \text{normal}(0, \tau_1), \quad \tau_1 \sim \text{normal}^+(0, 1), \\ \gamma_i - \gamma_{i-1} &\sim \text{normal}(0, \tau_2), \quad i = 2, \dots, 9, \quad \tau_2 \sim \text{normal}^+(0, 1), \quad \gamma_1 \sim \text{normal}(-8, 4), \\ \gamma_i^{\text{excess}} - \gamma_{i-1}^{\text{excess}} &\sim \text{normal}(0, \tau_3), \quad i = 2, \dots, 9, \quad \tau_3 \sim \text{normal}^+(0, 1), \quad \gamma_1^{\text{excess}} \sim \text{normal}(0, 1), \\ \mu_{\text{male}}^{\text{excess}}, \mu_{\text{edu}}^{\text{excess}}, \mu_{\text{male}}, \mu_{\text{edu}} &\sim \text{normal}(0, 1). \end{aligned}$$

Aggregate mortality in the population. After model is fitted using Markov chain Monte Carlo simulations, we generate posterior simulations of the demographic-specific baseline mortality :

$$\tilde{\theta}_{j_2, j_3, j_4} = \text{logit}^{-1} \left(\gamma_{[j_2]} + \mathbb{1}(\gamma_{j_3} = 2) \mu_{\text{male}} + \mathbb{1}(\gamma_{j_4} = 2) \mu_{\text{edu}} + \frac{1}{12} \sum_{j_1=1}^{12} \alpha_{j_1} \right),$$

and the demographic-specific mortality in 2020 after t_{start} :

$$\begin{aligned} \tilde{\theta}_{j_2, j_3, j_4}^* &= \text{logit}^{-1} \left(\gamma_{[j_2]} + \mathbb{1}(\gamma_{j_3} = 2) \mu_{\text{male}} + \mathbb{1}(\gamma_{j_4} = 2) \mu_{\text{edu}} + \right. \\ &\quad \left. \gamma_{[j_2]}^{\text{excess}} + \mathbb{1}(\gamma_{j_3} = 2) \mu_{\text{male}}^{\text{excess}} + \mathbb{1}(\gamma_{j_4} = 2) \mu_{\text{edu}}^{\text{excess}} + \frac{1}{12} \sum_{j_1=1}^{12} \alpha_{j_1} \right). \end{aligned}$$

We use census data to construct the demographic post-stratification weight w_{j_2, j_2, j_3} : the proportion of individuals in our round-1 survey with age j_2 , gender j_2 and educational level j_4 . Then the monthly excess mortality of any given age group j_2 is computed by

$$\text{excess mortality}[j_2] = \frac{\sum_{j_3, j_4} w_{j_2, j_3, j_4} (\tilde{\theta}_{j_2, j_3, j_4}^* - \tilde{\theta}_{j_2, j_3, j_4})}{\sum_{j_3, j_4} w_{j_2, j_4, j_4}}. \quad (3)$$

Likewise, we aggregate this excess mortality across all ages in the population:

$$\text{excess mortality}_{\text{all age}} = \frac{\sum_{j_2, j_3, j_4} w_{j_2, j_3, j_4} (\tilde{\theta}_{j_2, j_3, j_4}^* - \tilde{\theta}_{j_2, j_3, j_4})}{\sum_{j_2, j_3, j_4} w_{j_2, j_4, j_4}}. \quad (4)$$

To compute the relative mortality change, we divide this excess rate by the average baseline mortality,

$$\text{excess mortality}_{\text{relative, all age}} = \text{excess mortality}_{\text{all age}} \times \frac{\sum_{j_2, j_3, j_4} w_{j_2, j_4, j_4}}{\sum_{j_2, j_3, j_4} w_{j_2, j_3, j_4} \tilde{\theta}_{j_2, j_3, j_4}} \times 100\%. \quad (5)$$

Appendix B: Additional graphs

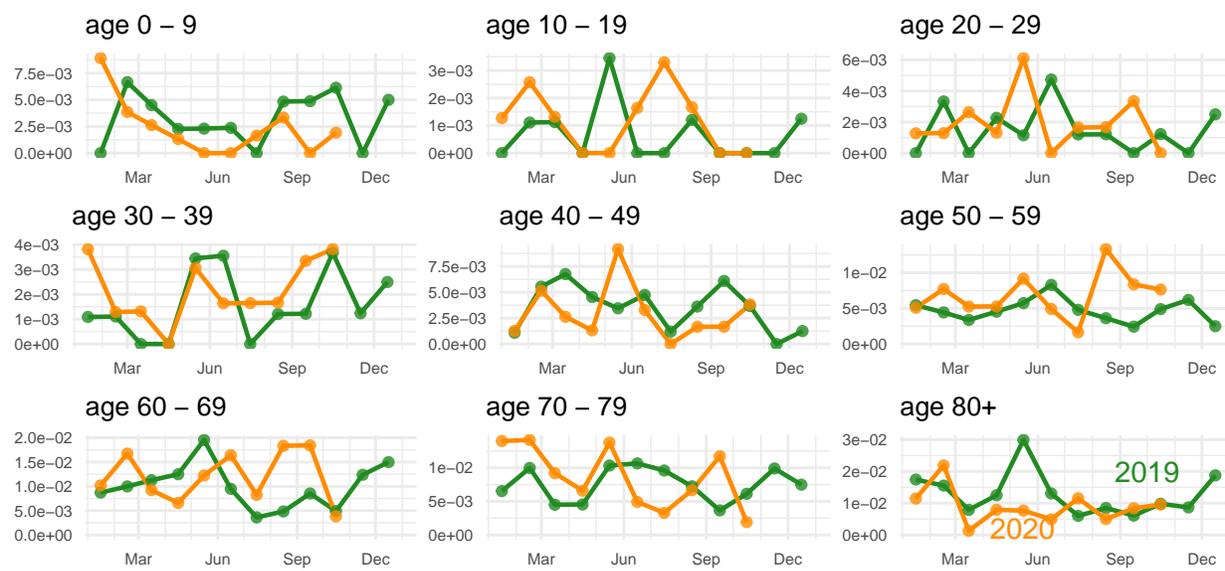


Figure 7: The empirical monthly mortality rate by age in 2019 (green line) and 2020 (orange). The mortality decline during Feb–Oct 2020 was evident for age group 0–9 and 80+.

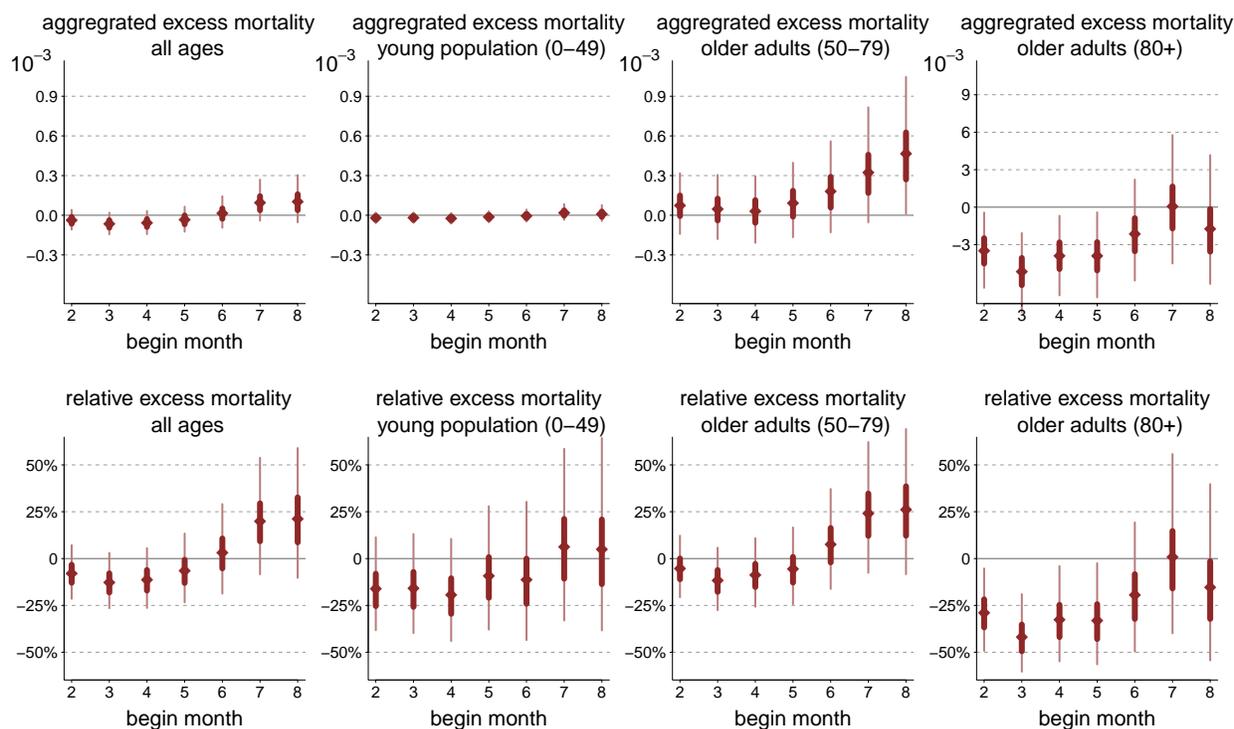


Figure 8: The posterior inference of the age aggregated (all ages, 0–49, 50–79, and 80+) monthly excess mortality rates and confidence intervals. The upper row is absolute change compared with baseline (extra death per thousand per month), and the lower row is the relative changes in percentage. The starting month of the comparison window t_{start} varies from Feb 2020 to Aug 2020.

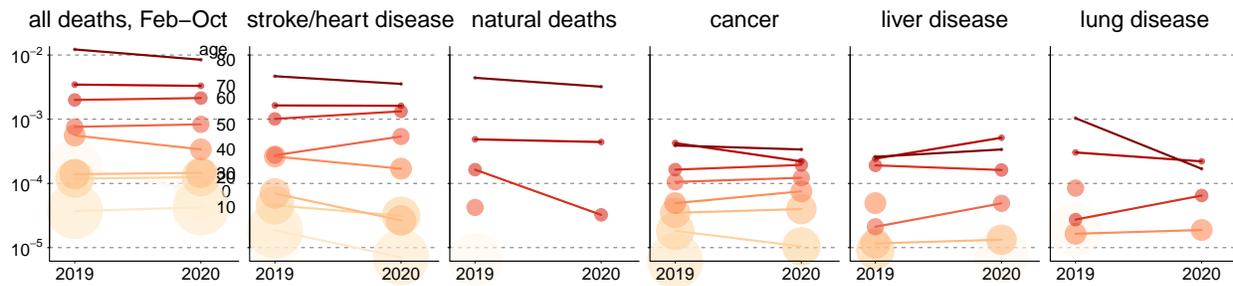


Figure 9: The empirical comparison of the age-specific proportion of survey samples (among all survey objects) who (1) died from all causes, (2) died from “stroke/heart disease”, (3) died from “natural death”, (4) died from “cancer”, (5) died from “liver disease”, (6) died from “lung disease”. The comparison is between Feb–Oct 2019 and Feb–Oct 2020. The dot sizes are the accessible population of that age group in the survey.

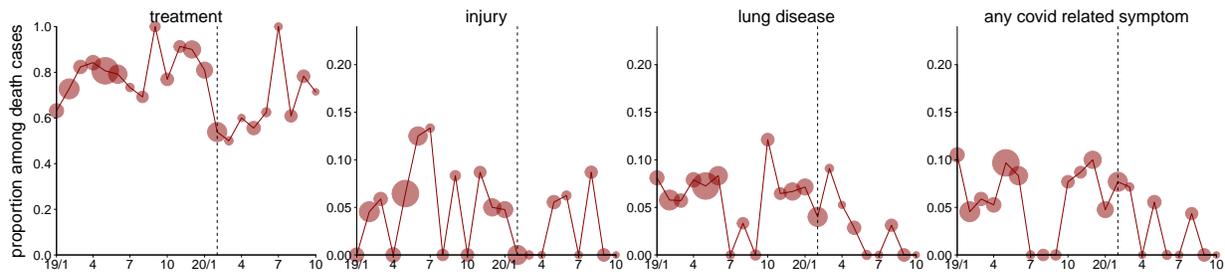


Figure 10: The empirical percentage of surveyed death case (among all deaths) of who (a) received any medical treatment, (b) were injured, (c) went through lung disease, and (d) went through any one of the COVID-related symptoms (headache, muscle aches, chill, cough, sore throat, lose-of-sense-of-smell, breathing difficulty, and fever). The dot size indicates the total death cases during each month. The category (c) and (d) largely overlap, while none of them exhibited an increasing trend in 2020.