1	Adult mortality before and during the COVID-19 pandemic in
2	nine communities of Yemen: a key informant study
3	Mervat Alhaffar1*, Huda BaSaleem2, Fouad Othman3, Khaled Alsakkaf2, Sena Mohammed Mohsen
4	Alkhteeb ⁴ , Hussein Kolaise ⁵ , Abdullah K. Babattah ⁶ , Yaseen Abdulmalik Mahyoub Salem ⁴ , Hannah
5	Brindle ¹ , Najwa Yahya ⁷ , Pasquale Pepe ¹ , Francesco Checchi ¹
6	
7	1 Department of Infectious Disease Epidemiology
8	Faculty of Epidemiology and Population Health
9	London School of Hygiene and Tropical Medicine, London, UK
10	
11	2 Department of Community Medicine and Public Health
12	Faculty of Medicine and Health Science
13	University of Aden, Aden, Yemen
14	
15	3 Faculty of Medicine and Health Science
16	Ta'iz University, Ta'iz, Yemen
17	
18	4 Epidemiology Surveillance Office,
19	Ta'iz Governorate Health Office, Ta'iz, Yemen
20	
21	5 Department of Internal Medicine
22	Faculty of Medicine and Health Science
23	University the proprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

- 24 6 Primary Health Care Program, Health Sector
- 25 HUMAN ACCESS for Partnership and Development
- 26 Aden, Yemen
- 27 7 Yemen Public Health Network
- 28 Aden, Yemen
- 29
- 30 * Corresponding author: Mervat.Alhaffar1@lshtm.ac.uk

31 Abstract

32 Introduction

Widespread armed conflict has affected Yemen since 2014. To date, the mortality toll of seven years of crisis, and any excess due to the COVID-19 pandemic, are not well quantified. We attempted to estimate population mortality during the pre-pandemic and pandemic periods in nine purposively selected urban and rural communities of southern and central Yemen (Aden and Ta'iz governorates), totalling > 100,000 people.

38

39 Methods

Within each study site, we collected lists of decedents between January 2014-March 2021 by interviewing different categories of key community informants, including community leaders, imams, healthcare workers, senior citizens and others. After linking records across lists based on key variables, we applied two-, three- or four-list capture-recapture analysis to estimate total death tolls. We also computed death rates by combining these estimates with population denominators, themselves subject to estimation.

46 Results

After interviewing 138 disproportionately (74.6%) male informants, we identified 2445 unique decedents. While informants recalled deaths throughout the study period, reported deaths among children were sparse: we thus restricted analysis to persons aged \geq 15 years old. We noted a peak in reported deaths during May-July 2020, plausibly coinciding with the first COVID-19 wave. Death rate estimates featured uninformatively large confidence intervals, but appeared elevated compared to the non-crisis baseline, particularly in two sites where a large proportion of deaths were attributed to war injuries. There was no clear-cut evidence of excess mortality during the pandemic period.

54

55 Conclusions

We found some evidence of a peak in mortality during the early phase of the pandemic, but death rate estimates were otherwise too imprecise to enable strong inference on trends. Estimates suggested substantial mortality elevations from baseline during the crisis period, but are subject to serious potential biases. The study highlighted challenges of data collection in this insecure, politically contested environment.

61

62 Key words

Yemen, humanitarian, armed conflict, crisis, mortality, SARS-CoV-2, COVID-19, key informant, capture
 recapture, multiple systems estimation

65 Background

Yemen has been affected by political unrest throughout recent history [1–3]. The current armed conflict started in 2014. The country is split into several areas of control with contested borders and multiple local and international regional actors holding different political and ideological stances [3–6]. The protracted armed conflict resulted in targeting civilians, healthcare facilities and personnel, water and sanitation infrastructure, schools, roads and bridges thereby disrupting services and exacerbating poor living conditions experienced before the war [5, 7–10]. In addition to frequent civilian casualties, the protracted

72 conflict has resulted in widespread forced displacement, food insecurity, severe malnutrition and food 73 insecurity bordering on potential famine [11], and a heavily disrupted and fragmented health system 74 consisting of two ministries of health (controlled by two different authorities in the north and the south), one 75 based in Sana'a and the other in Aden [8, 12].

The first laboratory-confirmed case of COVID-19 in Yemen was announced in April 2020 [13]. Efforts to track the pandemic's dynamics and share information relating to its impact on morbidity and mortality have differed between areas of political control. Houthi authorities in particular have tended not to openly report cases and deaths in their areas of control [14–16].

In conflict settings, information on health status can help humanitarian actors better allocate scarce resources and record the progress of response activities [17, 18]. Population mortality is a key indicator of health status as it captures the contributions of various health risk factors [18, 19]. Additionally, documenting mortality during conflicts may support advocacy for peacebuilding and post-conflict reconciliation [20].

To contribute evidence on COVID-19 attributable mortality in Yemen, we aimed to estimate the excess death toll and death rate during the pandemic period, as well as the pre-pandemic levels, in nine selected sites in the south and central of the country by collecting lists of decedents from key community informants, and applying capture-recapture statistical analysis to these. A secondary aim was to explore the feasibility and limitations of conducting such research in this context.

90

91 Methods

92 Scoping phase

Prior to data collection, we conducted a scoping phase consisting of informal in-depth interviews and group discussions in Arabic with more than seventy people who identified as Yemenis in the United Kingdom (UK), Yemen, and other countries. The participants were identified through a snowball approach of sequential referrals while ensuring a difference in their backgrounds including profession and place of origin in Yemen. The interviews took place in person in London and Sheffield (after obtaining verbal

98 consent), and online with note-taking used to record the information collected. Interviews included questions on (i) the categories of people most likely to know about deaths and births in the community, (ii) 99 100 the ways in which people report/know about deaths in the community, (iii) whether health facilities or burial 101 sites keep any written lists on deceased people and (iv) security concerns that might impact participation 102 or bias people's answers in different areas in Yemen. This scoping phase helped us to form an independent 103 advisory group of country-based Yemeni researchers and public health actors to provide guidance on the 104 project, and establish collaborations between the London School of Hygiene and Tropical Medicine, the 105 University of Aden and Ta'iz University.

106

107 Study population and period

108 While we originally wished to select a nationally representative sample, study permission was not received 109 from authorities in Sana'a (the Supreme Council for the Management and Coordination of Humanitarian 110 Affairs and International Cooperation SCMCHA)[21], and even in southern Yemen, the geographical scope 111 of data collection was constrained by travel and security challenges. We thus opted to purposively select 112 nine accessible sites including four in Aden governorate and five in Ta'iz governorate which were identified 113 by Yemen-based authors and represented different levels and typologies of exposures to the crisis (Table 114 1). We further selected sites that were clearly delineated (e.g. city neighbourhoods, groups of villages) 115 and of an approximate population size (5000-20,000) consistent with feasible data collection within the 116 study's resources and security constraints. All the sites were under the control of the Internationally 117 Recognised Government (IRG), site names are omitted here to maintain confidentiality of the key informants. Data collection started on 18th January and ended on 31st March 2021. We inquired about 118 119 deaths among people residing within the site from January 2014 to the start of data collection.

Several key informant categories were identified: community leaders and their spouses; healthcare providers; imams; ritual burial preparers; senior citizens; and teachers. None were known to keep written death records, except for health facility managers. Moreover, in places where the conflict was intense, records from local hospitals were destroyed or were missing data from specific years or were not even collected. Depending on the site's characteristics e.g. urban versus rural or size, some or all of the above informant categories were identifiable and thus included in the study.

126	Table 1. Characteristics of the study sites and informant lists retained for capture-recapture analysis. A and T
127	refer to sites in Aden and Ta'iz governorates, respectively.

		<u> </u>		
Site	Rural/ Urban	Characteristics (estimated population, March 2021)	Informant lists (main analysis)	Informant lists (alternative analysis)
A1	Urban	Directly and heavily affected by the armed conflict; under Houthi control for around four months. Residents fled the area due to the lack of water, electricity and fuel, and shortage of food and health services. (population size uncertain)	 Community leaders Health workers Imams Burial preparers 	 Community leaders Health workers Imams + Burial preparers
A2	Urban	Directly and heavily affected by the armed conflict; under Houthi control for around four months. Residents fled the area due to the lack of water, electricity and fuel, and shortage of food and health services. (population 11,000, ≈9% of subdistrict)	 Community leaders Senior citizens Imams Burial preparers 	 Community leaders senior citizens Imams + Burial preparers
A3	Urban	Moderately affected by the armed conflict; destination of many internally displaced populations (IDPs). (population 2600, ≈2% of subdistrict)	 Community leaders Imams Burial preparers 	 Community leaders Imams + Burial preparers
A4	Urban	Moderately affected by the armed conflict; destination of many IDPs. (population 9400, ≈6% of subdistrict)	 Community leaders spouses Senior citizens Burial preparers 	 Community leaders spouses + senior citizens Burial preparers
T1	Urban	On the contested borders between the two areas of political control. (population 2700, ≈3% of subdistrict)	 Community leaders spouses + senior citizens + imams Health workers 	 Community leaders spouses + senior citizens Health workers Imams + Burial preparers
T2	Rural area	Ongoing direct conflict since the beginning of the crisis; divided between two areas of control. (population 14,700 100% of subdistrict)	 Community leaders senior citizens + imams Health workers Teachers Burial preparers 	 Community leaders senior citizens + teachers Health workers Imams + burial preparers
Т3	Semi- urban	No active conflict, but a destination for many IDPs. (population size uncertain)	 Community leaders spouses + senior citizens + imams Health workers 	 Community leaders spouses + senior citizens Health workers Imams
T4	Urban	Past direct conflict, was under bombardment and siege. (population 8600, ≈15% of subdistrict)	 Community leaders senior citizens Health workers 	 Community leaders + senior citizens Health workers
T5	Rural area	Ongoing direct conflict since the beginning of the crisis, divided between two areas of control. (population 21,700, 100% of subdistrict)	 Community leaders Senior citizens 	 Community leaders senior citizens (no analysis possible: only one list)

128

129 Participant recruitment

The in-country research team consisted of three male medical doctors and one female midwife who were residents and familiar with the study communities. They initially drew on their personal and professional connections to identify key informants and build trust before conducting the interviews. Participants provided verbal, researcher-witnessed informed consent. Those who declined to participate were mostly able to recommend others to be interviewed. All key informants provided consent in-person , mostly after 135 receiving recommendations from respected peers, e.g. another community leader or a hospital manager, 136 who either wrote letters of introduction or accompanied researchers on their first visits to potential 137 participants. In Aden, preliminary phone conversations helped to establish rapport with key informants and 138 facilitate organizing appointments for introducing the study while in Ta'iz in-person visits were necessary. 139 In Aden 13/69 people contacted (two women and 11 men) declined to participate in the study, while in 140 Ta'iz 7/89 (all men) declined. Stated reasons for non-participation included concerns that information would 141 be shared with opposing forces; negative previous experiences with researchers; insufficient time; and 142 anticipated difficulty to recall deaths.

143

144 Data collection

145 A standardised questionnaire was developed and built using the Open Data Kit (ODK) [22] platform, with 146 in-built validation rules to collect information on (i) key informants (age, gender, typology, site and 147 residency) and (ii) names, age, gender, year of death, cause of death, place of death and origin of 148 decedents within the sites since 2014. Online two-day training sessions for researchers took place in 149 Arabic consisting of study concepts, procedures for data collection and management, and role-play. Data 150 were entered into the ODK Application on Android devices and automatically uploaded to a secure server 151 hosted at the LSHTM. Participants were also offered the option to recall and write down decedent 152 information in their own time on a printed form, if needed enlisting help from their spouses; researchers 153 collected the form at a later point and entered the data into ODK.

154 Investigators in Aden, Ta'iz and London monitored data quality, addressed issues and replied to questions 155 in real-time. After conducting 56 interviews, we paused data collection for a week to share experiences. 156 Given that the preliminary analysis showed a disproportionate reporting of adult male deaths, we agreed 157 on procedures to enhance the detection of child and women decedents during subsequent interviews. This 158 included a heavier reliance on spouses for recall and listing female names as 'mother/wife/daughter of', 159 which was felt to be more protective of their confidentiality.

161 Analysis

162 Record linkage

Data were manually cleaned, and duplicates removed from each key informant's list. Linkage criteria (Table S1, Supplementary File) were developed to establish, for each site, which lists a given decedent was included in. Individuals' names in Yemen usually consist of first name, father's name, grandfather's name and lastly tribe name or area/ district of origin. Records with missing first name, gender, age, and/or year of death were excluded from analysis. The year and age at death were averaged if informants reported discordant values.

169

170 Capture-recapture analysis

171 Capture-recapture analysis examines the overlap among lists *L* to estimate the number of individuals (in 172 this case decedents) who have not been captured by any list. This estimate, summed to the number of 173 individuals appearing on at least one list, provides the total. In this study, lists consisted for each site of 174 the records collected from specific categories of informants. After record linkage, site data consisted of 175 two, three or four lists (see below and Table 1).

176 In a two-list scenario, each decedent $x \in \{1,2,3...N\}$ has status x_{10} if named within L_1 only, x_{01} if named within L_2 only, x_{11} if named by both lists and x_{00} if not captured by either list. The resulting contingency 177 178 table consists of four cells, n_{10} , n_{01} , n_{11} and n_{00} , the last of which is unknown. We used the simple 179 Chapman estimator to estimate the total number of deaths as $\hat{N} = n_{10} + n_{01} + n_{11} + \hat{n}_{00} =$ $\left(\frac{(n_{10}+1)(n_{01}+1)}{(n_{11}+1)}\right)$ - 1; we computed a confidence interval (CI) as $[e^{-Z_{\alpha/2}\hat{\sigma}_{0.5}}\varphi, e^{Z_{\alpha/2}\hat{\sigma}_{0.5}}\varphi]$, where $\varphi = n_{10} + 1$ 180 $n_{01} - n_{11} - 0.5 + \frac{(n_{10} - n_{11} + 0.5)(n_{01} - n_{11} + 0.5)}{(n_{11} + 0.5)}$, $z_{\alpha/2}$ is the normal distribution quantile for a given significance 181 $\hat{\sigma}_{05} =$ 182 level (1.96 of interest for $\alpha = 0.05$ or 95%CI) and $\sqrt{\frac{1}{n_{11}+0.5} + \frac{1}{(n_{10}-n_{11}+0.5)} + \frac{1}{(n_{10}-n_{11}+0.5)} + \frac{n_{11}+0.5}{(n_{10}-n_{11}+0.5)(n_{10}-n_{11}+0.5)}}, \text{ as per Sadinle [23]}.$ 183

184 In a three-list scenario, the overlap among lists *L* may be represented by eight alternative candidate log-185 linear Poisson models, each of which features terms for the probability of appearing on any given list, as 186 well as two-way interaction terms representing potential dependencies among lists: these models range

187 from one with no interaction terms to a model featuring all the two-way interactions $L_1 \times L_2$, $L_2 \times L_3$ and $L_1 \times L_3$. We wished to also include in the models an exposure (period before and during the COVID-19 188 pandemic in Yemen) and potential confounding variables (age, gender). To allow for continuous 189 190 covariates, we used Rossi et al.'s [24] parametrisation of log-linear models, whereby the dataset is 191 expanded to feature, for each individual, all potential list statuses ($x_{000}, x_{100}, x_{101}, x_{001}, x_{110}, x_{101}, x_{011}, x_{111}$); 192 an outcome of 1 for the actual status, missing for status x_{000} and 0 otherwise; and any covariate values. 193 The model, once fit, is used to predict \hat{x}_{000} , interpretable as each individual's contribution to \hat{n}_{000} , the estimate of uncaptured deaths (i.e. $\hat{n}_{000} = \sum_{i=1}^{N} \hat{x}_{000}$); this quantity may of course be stratified by exposure 194 195 stratum. This estimation framework can easily be extended to the four-list scenario, which however entails 196 a larger set of models, featuring both three-list and/or (hierarchically non-redundant) two-list interactions.

197 While conventional capture-recapture analysis selects the best-fitting among candidate models, we 198 adapted Rossi et al.'s suggested approach for averaging multiple models [25]. First, we screened out 199 models that did not fit (e.g. due to sparse overlap among lists), yielded an implausible $\hat{n}_{000(0)}$ (defined as \geq 10 times the number of listed deaths) or featured a likelihood-ratio test p-value \geq 0.60 when compared 200 201 to the saturated model (indicating potential overfitting). At this stage we also assessed whether to retain 202 any potential confounder covariates, based on likelihood-ratio tests compared to the no-confounder model 203 and inspection of estimates with and without the confounder. For each shortlisted model $i \in \{1, 2, 3 \dots K\}$, we then computed a weight (equivalent to a Bayesian posterior probability) between 0 and 1 $w_i = \frac{e^{-\Delta_i/2}}{\sum_{k=1}^{K} e^{-\Delta_i/2}}$ 204 205 , where $\Delta_i = AIC_i - AIC_{\min}$, i.e. the difference between the model's Akaike Information Criterion (AIC) and 206 the lowest AIC among all shortlisted models (the AIC is a goodness-of-fit indicator that rewards predictive 207 accuracy and parsimony, i.e. the fewest possible model terms). Finally, we computed a weighted average 208 estimate of $\hat{n}_{000(0)} = \sum_{i=1}^{K} w_k \hat{n}_{000(0),i}$. We present results overall and by period.

209

210 Alternative list groupings

Capture-recapture analysis is unfeasible when the overlap among the lists is very poor, causing models not to fit (see above). To avoid these problems, for our main analysis we grouped lists together that plausibly reflected similar sources of community information: community leaders' and their spouses' lists

were combined in sites A4, T1 and T3; community leaders' and senior citizens' lists were combined in sites
T1, T2, T3 and T4; community leaders' and imams' lists were combined in sites T1, T2 and T3. This yielded
between two and four grouped lists per site (Table 1).

As an alternative analysis, we (i) combined imams' and burial preparers' lists across all sites to form one list and (ii) added records obtained from teachers and senior citizens to the community leaders' lists.

219

220 Population estimates and death rates

221 We previously estimated the all-age population of Yemen between 2014 and 2021, by month and sub-222 district (administrative level 3), using a combination of pre-crisis census, geospatial projections by 223 WorldPop (available at 100 m² resolution, and resulting from extensively validated predictive statistical 224 models [26]) and displacement flow data. Details are provided in Checchi et al. [27]. However, only two 225 sites (T2, T5) consisted of entire subdistricts. In remaining sites, researchers attempted to collect GPS 226 coordinates of all corners of the site (e.g. road boundaries) using their phones. We overlaid these polygons, 227 and the boundaries of the surrounding subdistricts, onto WorldPop projections for 2017 (the mid-point of 228 the data collection period) to compute the approximate proportion of the subdistrict's population that fell 229 within the site. We then multiplied our subdistrict estimates by this proportion to estimate the site's 230 population. Site boundaries, and thus populations, were unresolved for two sites (A1, T3). We used population estimates to compute average death rates by period (pre-pandemic and pandemic). We 231 232 specifically present crude death rates among people aged ≥15yo (CDR15+), per all-age population: this 233 may be thought of as an age-specific fraction of the all-age CDR.

234

235 **Results**

236 Decedent records

Between 14 January and 31 March 2021, 138 interviews were conducted (56 in Aden, 82 in Ta'iz: Figure
1), yielding 3022 decedent records. Across all sites, only 35/138 (25.4%) interviews were conducted with
female informants (midwives, health workers, wives of community leaders and burial preparers).



Figure 1. Number of key informant interviews done, by site and gender of the informant.

242

247

240

Apart from 2014, when the number of recalled deaths was 138, the annual number of death events recalled by informants did not appear to decrease as time became more remote (range = 269 to 372) (Figure 2). The largest number of deaths were reported for 2020 (N = 574). The highest number of deaths was reported by participants from site T1 (N = 718) and the lowest from T5 (N = 81).



Figure 2. Number of decedent records reported by key informants, by year and site.

250 Characteristics of reported deaths

After record linkage (including definite, probable and possible matches) and excluding 21 non- analysable records due to one or more missing variables, 2445 unique decedents were identified across all sites. All sites except for T3 were notable for the sparsity of data on children (only 189/2445 or 7.7% reported decedents were aged <15yo), and for a disproportion of males (1851/2445, 75.7%; Table 2). Site T4 featured an exceptionally large share of deaths in the 15-44yo stratum while in the remaining sites the highest numbers of deaths were reported among people aged \geq 45yo. Site-specific detail by list is provided in the Supplementary File.

258

site:	A1	A2	A3	A4	T1	T2	T3	T4	T5
Age	Age								
Mean	48.5y	57.3y	54.4y	58.1y	43.6y	49.7y	38.1y	33.4y	53.9y
0.4	1	0	2	2	26	9	27	4	0
Uy	(0.8%)	(0.0%)	(0.9%)	(0.6%)	(4.3%)	(2.3%)	(20.0%)	(1.2%)	(0.0%)
1 v to 4 v	3	2	1	0	6	11	7	3	0
Ty to 4y	(2.5%)	(0.9%)	(0.4%)	(0.0%)	(1%)	(2.8%)	(5.2%)	(0.9%)	(0.0%)
5y to 1 dy	6	3	1	6	29	22	6	10	2
5y to 14y	(5.0%)	(1.3%)	(0.4%)	(1.9%)	(4.8%)	(5.6%)	(4.4%)	(3.0%)	(2.6%)
15v to 14v	37	44	64	65	252	111	35	244	19
15y to 44y	(30.8%)	(18.7%)	(27.4%)	(20.9%)	(41.3%)	(28.3%)	(25.9%)	(73.5%)	(25.0%)
SAEV	73	186	166	238	297	239	60	71	55
245y	(60.8%)	(79.1%)	(70.9%)	(76.5%)	(48.7%)	(61%)	(44.4%)	(21.4%)	(72.4%)
Gender									
Fomalo	36	70	53	86	115	132	45	31	26
remale	(30.0%)	(29.8%)	(22.6%)	(27.7%)	(18.9%)	(33.8%)	(33.3%)	(9.3%)	(34.2%)
Mala	84	165	181	225	495	260	90	301	50
IVIAIE	(70.0%)	(70.5%)	(77.7%)	(72.8%)	(81.1%)	(66.3%)	(66.7%)	(90.7%)	(65.8%)
Totals	120	235	234	311	610	392	135	332	76

Table 2. Age and gender distribution of reported deaths, by site.

260

For 2020 and 2021 only, we collected months of death: when aggregating all sites, a notable peak in monthly deaths is evident in May-July 2020, possibly corresponding to the first COVID-19 wave (Figure 3), with a majority of recalled deaths among people aged ≥ 45yo.



264

Figure 3. Number of unique decedents per month reported by informants across all sites, January 2020 to March2021.

267

Across all sites and particularly in A1, T1 and T4, a substantial proportion of deaths was attributed to intentional or war-related injuries (Table 3).

270

Table 3. Causes of death as reported by informants over the entire analysis period, by site.

Site	Intentional/war injury	Unintentional injury	Other causes	Total
A1	44 (36.7%)	15 (12.5%)	61 (50.8%)	120
A2	38 (16.2%)	3 (1.3%)	194 (82.6%)	235
A3	46 (19.7%)	15 (6.4%)	173 (73.9%)	234
A4	50 (16.1%)	11 (3.5%)	250 (80.4%)	311
T1	244 (40.0%)	31 (5.1%)	335 (54.9%)	610
T2	61 (15.6%)	23 (5.9%)	308 (78.6%)	392
T3	11 (8.1%)	11 (8.1%)	113 (83.7%)	135
T4	168 (50.6%)	106 (31.9%)	58 (17.5%)	332
T5	12 (15.8%)	0 (0.0%)	64 (84,2%)	76

272

273 Estimates of total adult mortality

Across all sites, estimates of total deaths overall and by period before and after the first COVID-19 case

was announced in Yemen featured wide 95% confidence intervals (95%CI; Table 4). When considering

- point estimates, the mean monthly mortality rose appreciably during March 2020 to March 2021 compared
- to January 2014 to February 2020 in sites A1, A2, A4, T3 and T5, remained nearly constant in A3, T1 and
- T2, and declined in T4. The alternative analysis yielded an approximately similar pattern (Table 5).
- 279
- Table 4. Estimates of total deaths and deaths per month, overall and by period (main analysis).

	esti	mated total deaths (95%	estimated mean deaths per month			
Site	Overall	Jan 2014 to Feb 2020	Mar 2020 to Mar 2021	Overall	Jan 2014 to Feb 2020	Mar 2020 to Mar 2021
A1	461 (166 to 3570)	342 (129 to 2049)	119 (38 to 1520)	5.2	4.6	9.9
A2	633 (358 to 1636)	492 (260 to 1340)	141 (98 to 296)	7.1	6.6	11.8
A3	604 (373 to 1293)	604 514 (373 to 1293) (319 to 1064)		6.8	6.9	7.5
A4	1840 (664 to 7722)	1474 (453 to 6720)	367 (211 to 1002)	20.9	19.7	30.6
T1	3127 (2203 to 4776)	2361 (1631 to 3732)	455 (227 to 1451)	35.5	31.5	37.9
T2	839 (613 to 1283)	732 (534 to 1100)	108 (79 to 183)	9.5	9.8	9.0
Т3	1739 (291 to 56,295)	545 (113 to 17,876)	322 (69 to 10,602)	19.7	7.3	26.8
T4	4920 (2102 to 15,933)	3716 (1601 to 12,023)	299 (69 to 9902)	33.1	49.6	24.9
T5	569 (200 to 3769)	299 (114 to 1985)	107 (32 to 3649)	6.5	4.0	8.9

281

Table 5. Estimates of total deaths and deaths per month, overall and by period (alternative analysis).

	est	imated total deaths (95%	estimated mean deaths per month			
Site	Overall	Jan 2014 to Feb 2020	Mar 2020 to Mar 2021	Overall	Jan 2014 to Feb 2020	Mar 2020 to Mar 2021
A1	207 (35 to 1703)	52 (7 to 544)	2.3	2.1	4.3	
A2	441 308 132 (368 to 567) (243 to 437) (106 to 196)				4.1	11.0
A3	534 (424 to 733)	6.1	6.0	6.6		
A4	845 (569 to 1587)	818 (322 to 5430)	296 (211 to 568)	9.6	10.9	24.7
T1	No results (all models yielded implausible estimates)					
T2	915 (587 to 2060)	802 (509 to 1815)	113 (78 to 244)	10.4	10.7	9.4
Т3	No results (all models yielded implausible estimates)					
T4	4126 (2074 to 10,274)	119 (64 to 485)	46.8	62.9	9.9	
T5	No results (only one list)					

Available CDR15+ estimates varied considerably across sites (Table 6). There was no consistent pattern of increase during the pandemic period, though 95%Cls are uninformatively wide. Sites T1 and T4, where

- the highest CDR15+ was observed, were also those with the highest percentage of intentional/war injury
- deaths (Table 3) and deaths among young adults (Table 2).

288

Table 6. Estimated crude death rates among people aged \geq 15yo, overall and by period (main analysis).

Cite	deaths ≥15yo pe	er 1000 all-age pop (95%Cl)	oulation per year	deaths ≥15yo per 10,000 all-age population per day (95%CI)			
Sile	Overall	Jan 2014 to Feb 2020	Mar 2020 to Mar 2021	Overall	Jan 2014 to Feb 2020	Mar 2020 to Mar 2021	
A1	[no population estimates]						
A2	9.1	8.5	11.9	0.2	0.2	0.3	
	(5.2 to 23.5)	(4.5 to 23.2)	(8.3 to 25.0)	(0.1 to 0.6)	(0.1 to 0.6)	(0.2 to 0.7)	
A3	38.3	39.5	32.8	1.0	1.1	0.9	
	(23.7 to 82)	(24.5 to 81.7)	(20.0 to 83.4)	(0.6 to 2.2)	(0.7 to 2.2)	(0.5 to 2.3)	
A4	28.2	27.3	32.7	0.8	0.7	0.9	
	(10.2 to 118.3)	(8.4 to 124.3)	(18.8 to 89.4)	(0.3 to 3.2)	(0.2 to 3.4)	(0.5 to 2.4)	
T1	132.9	138	70.8	3.6	3.8	1.9	
	(93.6 to 203)	(95.4 to 218.2)	(35.3 to 225.9)	(2.6 to 5.6)	(2.6 to 6.0)	(1.0 to 6.2)	
T2	7.5	8.6	4.0	0.2	0.2	0.1	
	(5.5 to 11.5)	(6.3 to 12.9)	(2.9 to 6.8)	(0.2 to 0.3)	(0.2 to 0.4)	(0.1 to 0.2)	
Т3	[no population estimates]						
T4	56.0	59.8	11.7	1.5	1.6	0.3	
	(23.9 to 181.4)	(25.7 to 193.3)	(2.7 to 385.9)	(0.7 to 5.0)	(0.7 to 5.3)	(0.1 to 10.6)	
T5	3.8	2.4	4.0	0.1	0.1	0.1	
	(1.3 to 24.9)	(0.9 to 15.9)	(1.2 to 137.6)	(0.0 to 0.7)	(0.0 to 0.4)	(0.0 to 3.8)	

290

291 Discussion

292 Main findings

293 Covering a population of over 100,000 and a period of seven years, our analysis suggests that nine 294 purposively selected communities in Yemen, featuring varying exposure to the crisis, experienced highly 295 variable death rates among people aged ≥15yo over periods before and during the COVID-19 pandemic.

Due to sparse reporting of data on children, we restricted analysis to older age groups; accordingly, our CDR15+ indicators are not readily comparable with other crisis-affected settings. Based on United Nations demographic projections not accounting for crisis conditions, Yemen would have expected about 115,000 annual deaths ≥15yo during 2016-2020 within a mid-period population of 27,835,000, yielding a 'baseline' CDR15+ of 4.1 per 1000 person-years (0.1 per 10,000 person-days) [28]. In six of the seven study sites with available population denominators, the 95%CIs of our CDR15+ estimates exceed this baseline during all study periods, suggesting a broad pattern of considerably elevated mortality during the crisis period, 303 with alarmingly high levels in two sites (T1, T4) where the age distribution (15-44 yo) and cause of reported 304 deaths suggests elevated mortality directly due to armed conflict. We did not observe an obvious increase 305 in mortality (CDR15+) during the COVID-19 transmission period, but the inaccuracy of period estimates 306 hampers this comparison; it is possible that in some sites, extreme elevations in mortality during the early 307 phase of the war, followed by a less acute situation, would have counterbalanced any increase due to 308 COVID-19 .The distribution of monthly reported deaths during 2020 does suggest the study sites 309 experienced an acute epidemic peak in May-July temporally consistent with the first wave of the pandemic. 310 A second wave was reported in Yemen shortly after this study's data collection period.

Based on a PubMed search with keywords "mortality" and "Yemen", we identified only two other reports of population mortality in the country since the crisis began in 2014. In Aden, satellite imagery of cemeteries and civil registration data suggested a very similar peak in mortality as in this study during May-July 2020 [13]. Ogbu et al. analysed 56 surveys done by humanitarian actors during 2015-2019 and classified Yemen's governorates into high- and low-child mortality groups [29]; these surveys have been criticised for under-estimation [5].

The sites we investigated are unlikely to be representative of Aden and Ta'iz governorates, or indeed all of Yemen, and furthermore our estimates are subject to considerable imprecision and potential bias, highlighting challenges of primary data collection in insecure, politically contested settings (see below). The study does, however, indicate that community informants are able to recall details of local, adult decedents, even going back several years.

322

323 Data collection challenges

Various challenges of this study arose during implementation. Limited electricity and internet coverage, ongoing SARS-CoV-2 transmission and unpredictable security conditions constrained data collection, especially in rural, remote sites within Ta'iz governorate, where the female researcher had to be accompanied by her spouse to secure travel. Moreover, some informants reported concerns that nominal details of decedents would be shared with warring parties, resulting either in non-participation or in short interviews. While informants generally recalled chronologically remote deaths, recall of ages and years of death was difficult, and informants appeared to mainly remember deaths during salient events (e.g. 331 bombardments or battles) or decedents who are considered 'martyrs' in their community. This may explain the high proportion of war injury deaths among those reported. Capturing deaths among women and, in 332 333 particular, children was difficult across sites, probably reflecting the choice of key informants (mostly men 334 in positions of authority). Female researchers appeared better able to obtain this information from male 335 and female participants, suggesting future similar studies in Yemen and comparable cultural settings 336 should systematically deploy mixed-gender teams to collect data in communities and conduct formative 337 work to ensure a balanced gender distribution among key informants. Lastly, entry and management of 338 data on the ODK platform proved straightforward, though entering long lists of names onto smart phones 339 was time-consuming.

340

341 Limitations

342 Capture-recapture analysis involves certain statistical assumptions. Firstly, record linkage should be 343 accurate, resulting in error-free attribution of decedents to one or more lists. Misclassification could have 344 occurred due to the complexity of names in Yemen and the high frequency of certain names, combined 345 with other errors in recall of key linkage variables. The long recall period affected recall of the exact full 346 names of decedents by the key informants. If decedents are incorrectly classified as unique (i.e. less 347 overlap among lists), overall mortality estimates would be upward-biased (over-estimated), while the 348 inverse would happen if decedents on different lists are incorrectly matched. If misclassification is a 349 function of date of death (e.g. more frequent errors for deaths in the remote past), our estimates of period 350 mortality would tend to be artefactually higher in the pre-pandemic period than during COVID-19: this may 351 be an explanation for the high pre-pandemic death rates in several sites, though it should be noted that 352 data completeness was not correlated with date of death (data not shown).

A second assumption of capture-recapture methods is that lists do not draw on each other for information: this assumption is adjusted for in three- and four-list analyses, but may have biased our estimates for T1, T3, T4 and T5, where two-list analysis was done (in the latter, senior citizen and community leader lists may plausibly be non-independent).

Across all sites, we observed surprisingly little overlap among lists (Supplementary File), in contrast to previous applications of this method [30]. This may reflect individual informants being aware of deaths only 359 within a sub-section of the site, difficulty in identifying all informants of the same typology within the site 360 and/or local kinship, access to information, trust networks and ongoing insecurity.

Lastly, population denominators used to compute death rates are also subject to inaccuracy, not captured in the 95%Cls of our estimates. While all estimates were based on a 2004 census, projections rely on several statistical models and ground displacement information, all of which will typically have higher relative error at fine geographic resolution.

365 Conclusions

This study provides some evidence of elevated mortality in parts of Yemen during the crisis period (particularly in communities with high death tolls due to war injuries), and of a mortality peak consistent with the first COVID-19 wave in 2020. However, study limitations considerably weaken this inference: as such, evidence from this study should be evaluated carefully alongside other efforts to document mortality and public health impacts of the war in Yemen. Future studies could more specifically compare sites affected directly by the conflict with those that weren't.

This application of a key informant method to document mortality illustrates the feasibility and challenges of conducting such research in Yemen, and possibly comparable crisis-affected, insecure settings. We believe that key informant interviews combined with capture-recapture analysis could be applied in conflictaffected settings by taking into consideration the contextual differences and addressing the limitations we outline above. This method also has the potential to be applied prospectively as a substitute for nonfunctional vital events registration.

378

379 Declarations

380 Ethics approval and consent to participate

The study was approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine (Ref: 22080) and Faculty of Medicine and Health Science, University of Aden, Yemen (REC-76-2020). Local approvals from the Ministry of Public Health and Health and Population Office in Ta'iz Governorate were obtained. All research participants provided verbal informed consents before the interviews.

385	
386	Consent for publication
387	Not applicable.
388	
389	Availability of data and materials
390	Due to ongoing armed conflict and insecurity in Yemen, we believe that the benefit of publishing data, even
391	if anonymised, is outweighed by the risk to participants. We may publish data in the future if the situation
392	allows. All statistical code, and a dummy dataset that can be used to implement capture-recapture
393	methods, is published at https://github.com/francescochecchi/mortality_capture_recapture_analysis
394	
395	Competing interests
396	The authors declare that they have no competing interests.
397	
398	Funding
399	The study was funded by the United Kingdom Foreign Commonwealth and Development Office, grant
400	reference 300708-139. Funders were not involved in study design, data collection, analysis, or manuscript
401	preparation.
402	
403	Authors' contributions
404	Study design: FC, MA. In-country supervision: HB, FO, KZ. External supervision: HB. Data collection: MA,
405	SA, HK, AB, YS. Data management: MA. Analysis: FC. Drafting manuscript: MA, FC. All authors
406	commented on the draft and provided substantial inputs. All authors read and approved the final
407	manuscript.

409 Acknowledgments

- 410 We are grateful to all participants who gave their time to support this study, especially those in Yemen,
- 411 who work and live under conditions of extreme hardship and adversity.
- 412 ODK servers & support were provided by the LSHTM Global Health Analytics Group (odk.lshtm.ac.uk).

413

414 **References**

- 415 1. Yemen profile Timeline. BBC News. 2019. https://www.bbc.co.uk/news/world-middle-east-14704951.
 416 Accessed 8 Feb 2022.
- 417 2. Montgomery M. A Timeline of the Yemen Crisis, from the 1990s to the Present. Arab Center Washington DC.
- 418 2021. https://arabcenterdc.org/resource/a-timeline-of-the-yemen-crisis-from-the-1990s-to-the-present/.
 419 Accessed 8 Feb 2022.
- 3. Dureab F, Al-Sakkaf M, Ismail O, Kuunibe N, Krisam J, Müller O, et al. Diphtheria outbreak in Yemen: The
 impact of conflict on a fragile health system. Conflict and Health. 2019;13:1–7.
- 422 4. Elnakib S, Elaraby S, Othman F, BaSaleem H, Abdulghani AlShawafi NA, Saleh Al-Gawfi IA, et al. Providing
- 423 care under extreme adversity: The impact of the Yemen conflict on the personal and professional lives of health
 424 workers. Social Science & Medicine. 2021;272:113751.
- 425 5. Maxwell D, Hailey P, Spainhour Baker L, Janet Kim J. Constraints and Complexities of Information and
 426 Analysis in Humanitarian Emergencies: Evidence from Yemen. 2019.
- 427 6. AlKarim T, Abbara A, Attal B. Armed conflict alone does not explain the devastation of Yemen's health system.
 428 BMJ Glob Health. 2021;6:e004740.
- 429 7. The Lancet Global Health. Yemen needs a concrete plan—now. The Lancet Global Health. 2019;7:e1.
- 430 8. Qirbi N, Ismail SA. Health system functionality in a low-income country in the midst of conflict: the case of
- 431 Yemen. Health Policy and Planning. 2017;32:911–22.

- 9. Dureab F, Shibib K, Al-Yousufi R, Jahn A. Yemen: Cholera outbreak and the ongoing armed conflict. The
 Journal of Infection in Developing Countries. 2018;12:397–403.
- 434 10. Group of Eminent International and Regional Experts on Yemen. Situation of human rights in Yemen,
- 435 including violations and abuses since September 2014. 2019.
- 436 11. United Nations Office for Coordination of Humanitarian Affairs. Yemen Plan to reach 5.1 million people at437 risk of famine Yemen.
- 437 risk of famine Yemen.
- 438 12. Dureab F, Hussain T, Sheikh R, Al-Dheeb N, Al-Awlaqi S, Jahn A. Forms of Health System Fragmentation
 439 During Conflict: The Case of Yemen. Frontiers in Public Health. 2021;9:659980.
- 440 13. Besson EK, Norris A, Bin Ghouth AS, Freemantle T, Alhaffar M, Vazquez Y, et al. Excess mortality during
- the COVID-19 pandemic in Aden governorate, Yemen: A geospatial and statistical analysis. BMJ Global Health.
- 442 2020;:2020.10.27.20216366.
- 443 14. Al-Hamdani R, Wilson R. Yemen's Response to Covid-19: Part I. 2020.
- 15. Yemen: Houthis Risk Civilians' Health in Covid-19. Human Rights Watch. 2021.
- 445 16. Dureab F, Al-Awlaqi S, Jahn A. COVID-19 in Yemen: preparedness measures in a fragile state. The Lancet
 446 Public Health. 2020;5:e311.
- 17. Checchi F, Warsame A, Treacy-Wong V, Polonsky J, van Ommeren M, Prudhon C. Public health information
 in crisis-affected populations: a review of methods and their use for advocacy and action. The Lancet.
 2017;390:2297–313.
- 450 18. Checchi F, Roberts L. Documenting Mortality in Crises: What Keeps Us from Doing Better? PLOS Medicine.
 451 2008;5:e146.
- 452 19. Jawad M, Hone T, Vamos EP, Roderick P, Sullivan R, Millett C. Estimating indirect mortality impacts of
 453 armed conflict in civilian populations: Panel regression analyses of 193 countries, 1990-2017. BMC Medicine.
 454 2020;18:1–11.
- 20. Ekzayez A, Alhaj Ahmad Y, Alhaleb H, Checchi F. The impact of armed conflict on utilisation of health
 services in north-west Syria: an observational study. Conflict and Health 2021 15:1. 2021;15:1–10.

- 457 21. Supreme Council for the Management and Coordination of Humanitarian Affairs and International
 458 Cooperation SCMCHA. https://www.scmcha.org/en/. Accessed 6 Feb 2022.
- 459 22. Open Data Kit. https://opendatakit.org/. Accessed 6 Feb 2022.
- 460 23. Sadinle M. Transformed Logit Confidence Intervals for Small Populations in Single Capture–Recapture
- 461 Estimation. http://dx.doi.org/101080/03610910903168595. 2009;38:1909–24.
- 462 24. Rossi G, Pepe P, Curzio O, Marchi M. Parameterization of Continuous Covariates in the Poisson Capture-
- 463 Recapture Log Linear Model for Closed Populations. Statistica. 2019;79:427–43.
- 464 25. Rossi G, Pepe P, Curzio O, Marchi M. Generalized linear models and Capture-Recapture Method in a closed
- 465 population: strengths and weaknesses. Statistica. 2010;70:371–90.
- 466 26. Stevens FR, Gaughan AE, Linard C, Tatem AJ. Disaggregating Census Data for Population Mapping Using
- 467 Random Forests with Remotely-Sensed and Ancillary Data. PLoS ONE. 2015;10:e0107042.
- 27. Checchi F, Koum Besson ES. Reconstructing subdistrict-level population denominators in Yemen after six
 years of armed conflict and forced displacement. Journal of Migration and Health. 2022;5:100105.
- 470 28. United Nations D of E, Social Affairs PD. World Population Prospects Population Division United Nations.
 471 2019.
- 29. Ogbu TJ, Rodriguez-Llanes JM, Moitinho de Almeida M, Speybroeck N, Guha-Sapir D. Human insecurity
 and child deaths in conflict: evidence for improved response in Yemen. International Journal of Epidemiology.
- 474 2022;:dyac038.

30. Roberts B, Morgan OW, Sultani MG, Nyasulu P, Rwebangila S, Myatt M, et al. A new method to estimate
mortality in crisis-affected and resource-poor settings: validation study. Int J Epidemiol. 2010;39:1584–96.