Estimates of crisis-attributable mortality in Somalia, 2014-2018

FAQ Document

January 2020

Table of Contents

1. Who undertook this study?

A team of experts from the London School of Hygiene and Tropical Medicine (LSHTM) conducted the study. The LSHTM, a public university, is among the top 100 research universities worldwide, and was named UK University of the Year in 2016. The main study team comprised of Abdihamid Warsame, Severine Frisson, Amy Gimma and Francesco Checchi. Team members combine extensive experience working in humanitarian response with expertise in health information and statistical analysis. Prof Checchi is a recognised expert in the epidemiology of humanitarian crises and has conducted studies of mortality in crisis settings in various locations, including Angola, the Thai-Burma border, Darfur, Northern Uganda, the Central African Republic, Afghanistan, Malawi and Somalia.

2. Who funded the study?

The study was funded by the United Kingdom Government's Department for International Development. The LSHTM independently carried out data collection, analysis, publication and dissemination.

3. What are the key findings from the study?

The study found that about 45,000 Somali have died as a result of the drought crisis in Somalia from January 2017 through December 2018. We think however that the true number may be considerably higher (see below). The death rate was highest in the northeast and central regions of the country. Our estimate includes "directly" crisis-attributable deaths (i.e. people who died due to war injuries) and "indirectly" crisis-attributable deaths (i.e. resulting from disruptions in health services, worsened food security and livelihoods, displacement to overcrowded camps, etc.). Otherwise put, the estimate refers to the "excess" number of deaths above and beyond the number that would have occurred in Somalia in the absence of a drought crisis.

We were also able to estimate death tolls by district and month: as expected given the dynamics of the crisis we found that most excess mortality was concentrated in the northeast and central regions of Somalia.

4. What were the other causes of death?

We did not have any information on other (non-violent) causes of deaths. These were not recorded in the field surveys we analysed, and furthermore information on disease-related causes of death as provided by next-of-kin is known to be unreliable, particularly if the decedent did not attend a health facility and receive a medical diagnosis.

5. How did you actually come up with these estimates?

Our study report goes into detail on the methods used. Briefly, we used a statistical modelling approach based on previously collected data.

Firstly, we reviewed and re-analysed 201 so-called "SMART" surveys done by humanitarian actors within individual districts of Somalia from 2014 to 2018, mainly for monitoring the nutritional situation: these household surveys all contained a standardised mortality questionnaire, in which household respondents provided information on births, deaths and in-/out-migration within their household over a specified recent period (approximately the last 3-4 months). This information enables estimation of the death rate, i.e. how many people died per unit population and time (deaths per 10,000 persons per day is the metric used in humanitarian responses). We scored the quality of each survey based on information provided in its respective report, and also re-analysed raw datasets of the surveys whenever available (91 surveys).

On their own, the ground surveys would not have enabled us to estimate mortality across all of Somalia and for the entire period of interest, as they only covered only a small fraction of the country and period. We did however use them, in combination with a range of other data, to fit a statistical model that predicted the death rate for periods and areas of Somalia where no survey was done. A statistical model is in effect an equation that, in this case, helps us to estimate a quantity we do not know by virtue of variables, or 'predictors', we do have data on. For example, we could probably predict someone's income level if we knew his/her age, gender, ethnic group, car ownership and make, residential address, etc. In our case, we looked for predictor data that would be available throughout each district of Somalia and for the entire period of interest: these data included armed conflict incidence, displacement, main type of economic activity (livelihood), purchasing power, humanitarian presence, vaccination coverage, occurrence of epidemic diseases, etc. (see report for the full list).

Luckily, the above predictors were good at estimating ground mortality estimates as provided by the 91 SMART surveys, thereby giving us a model that we could use to predict the death rate where no ground data were available. We also made a set of assumptions about what value these predictors would have taken in the absence of a civil war: for example, we assumed that there would have been some ongoing internal displacement and insecurity in Somalia, but not at the level seen in actuality. These assumptions helped us to predict what we call the "counterfactual baseline", i.e. what the death rate *would have been* in the absence of the war. The difference between the actual predicted death rate and the counterfactual baseline death rate, for any district and month, is the *excess*, or crisis-attributable death rate.

In order to convert death rates into death tolls (i.e. numbers of who died), we need to know the population denominator, i.e. how many people were living in each district in each month. We therefore reconstructed to the extent information was available the evolution of population in each district. For each district we used the weighted average of several population estimations. We then projected population forward in time based on an assumed population growth rate (births – deaths), subtracted people who left the district (internally displaced people who went to another district, as well as refugees) and added people who came into the district (internally displaced people from other counties and any returnees). The net result was our best estimate of population at any given time. We multiplied this estimate by the death rates we estimated as per the above explanation, in order to compute the death toll.

We extended the above approach to also estimate the number of people killed, by modelling the death rate due to injury and that due to violence specifically, relying on any survey that did break down these causes of death.

6. What does the model actually mean? Can we tell which predictor caused the most deaths?

The variables retained in the final statistical models used for estimation all contribute to the overall accuracy or predictive power of the model. When reading the estimated coefficients (see Report, Table 4), a reasonable interpretation is that a positive coefficient means that the variable in question (or the category of the variable, e.g. a certain level of armed conflict intensity) is associated with higher death rate; a negative coefficient means that the variable or category is associated with lower death rate. We cannot directly tell from this how many deaths were attributable to different variables.

Furthermore, it's important to note that each of the variables in the model is mainly a "proxy" of a something more fundamental. For example, the presence of armed conflict is positively associated with death rate: however, this doesn't necessarily mean that the observed increase in death rate is due only to people dying of conflict; rather, what is more plausible is that the presence of armed conflict indicates that people had overall decreased access to humanitarian services. Similarly, measles vaccination uptake is likely to be a good proxy of the overall functionality of health services, just as terms of trade (amount of flour that can be obtained in exchange for a goat) is indicative of overall food security.

7. Why does the report conclude that the true death toll may be higher than that estimated?

We have identified a few noteworthy potential sources of bias in the data and methods we have used. These biases could lead to our analysis over- or under-estimating the true number of deaths. The main potential sources of bias are (i) likely under-reporting of child deaths by the ground surveys we analysed; (ii) uncertainty regarding population figures, especially the number of internally displaced people in different parts of Somalia; and (iii) the less-than-perfect predictive accuracy of the statistical model we came up with. On balance, most potential sources of bias would plausibly lead to under- rather than over-estimation. Therefore, we think that it is likely that the true numbers of deaths are somewhat higher than our estimates.

8. How confident can we be in these figures, given you are acknowledging they may be biased?

We are reasonably confident that our estimates are accurate enough to serve as a good basis for decision-making and action. As for nearly all research studies, our findings are subject to limitations, which we have acknowledged and explored the effect of in the report's Discussion. We have provided a range of alternative estimates derived from different assumptions, statistical techniques and levels of bias. Even in a conservative scenario, the excess death toll remains above 19,400. There are further indications that the findings are likely to be broadly valid: in particular, the statistical model we developed displays the kind of cause-and-effect associations with the death rate that we would have expected based on existing evidence and plausibility (for example, we see the death rate increase progressively with the incidence of armed conflict.

9. What would have been needed for the estimates to be more accurate?

More and better data on variables that may be associated with the death rate – in particular, any data from actors within Somalia on humanitarian access, as well as more specific and consistent data on the presence and activities of humanitarian actors. We believe such data may be collected by the UN Peacekeeping Mission in Somalia, but we were unable to access them. Moving forward, improved estimation of population denominators and work to address possible under-reporting of child deaths during field mortality surveys would also enhance the accuracy of any repeat analysis.

10. Why did you omit refugees, and doesn't that mean you are under-estimating mortality?

We excluded refugees living in Somalia but from other countries as they were numerically negligible. We excluded refugees from Somalia to other countries because we did not have enough data on these to also extend our methods to them. We don't think however that these omissions are very impactful: it is known that refugees often experience death rates similar to or even lower than if they had remained in their country of origin, and as such it is by no means clear that including refugees would have increased the excess death toll.

11. How does this number compare to excess mortality in other conflicts?

While the estimated death toll in Somalia is lower than during the 2010-2012 famine, it remains staggering in its own right. By comparison, previous studies have put the number of excess deaths at 376,000 to 389,800 in South Sudan during 2013-2018, 151,000 to 654,000 in Iraq during 2003-2006, depending on the study; 62,000 in Darfur between 2003 and 2008; and 360,000 in Syria since 2011 according to a systematic tracking project.

When considering the overall excess death toll, we know from a large number of studies that the death rate increases in different types of crisis, particularly among internally displaced people. Death rates have been known to spike to up to 20-50 times the likely baseline during the acute phase of displacement. We do not see such extreme levels of mortality in our analysis: on the one hand this may indicate that factors that would cause extremely high death rates, such as famine conditions, were less widespread than feared, and conversely that humanitarian action is having a mitigating effect; on the other hand, this may simply reflect our analysis approach: we estimated average death rates within fairly large population units (counties), and we may therefore be "smoothing out" peaks and troughs.

Indeed, what makes protracted crises such as that in Somalia so deadly is not so much acute peaks in death rates, but rather that they affect very large populations over a very long period of time: on average (i.e. across the entire crisis-affected area), the combination of different factors (disruption of public health services; displacement; loss of livelihoods; violence; etc.) results in a moderate but persistent increase in the death rate, which mathematically will inevitably result in staggering numbers of deaths: for example, in a population of 10 million over 10 years of conflict, even a small increase in the death rate, say from 0.4 to 0.5 per 10,000 person-days, would translate to 365,000 excess deaths. By comparison with our study, previous studies estimated an excess death toll of 605,000 in the Democratic Republic of Congo during 2003-2004, 298,000 in Darfur from 2003 to 2008 and 258,000 in south-central Somalia during 2010 to 2012.

12. Has this methodology been used in other contexts?

Broadly speaking, yes: our approach is similar to "indirect small area estimation", a recognised way to come up with very geographically granular predictions of parameters of interest, such as the prevalence of tobacco smoking or the level of child poverty in each district of the United States, without conducting expensive data collection within each district.

In terms of mortality estimation in humanitarian settings, we have used this approach once before to estimate the death toll attributable to food security and famine in Somalia during 2010-2012. We are working on a follow-on project to replicate the method in up to four other crisis contexts.

13. Have this methodology and the findings been peer-reviewed?

We have submitted a scientific paper on a similar study for South Sudan to a leading peer-reviewed journal. The methods for South Sudan and Somalia have been reviewed by two senior statisticians. We look forward to more feedback and suggestions for improvement from other researchers. We are choosing to release the report before a journal paper comes out because we wish to make findings available as soon as possible to policy-makers and humanitarian stakeholders.

14. Who authorised the study?

The study was approved by the Ethics Committee of the London School of Hygiene & Tropical Medicine (ref. 15334) and the Research and the Ethics Review Committee of the Ministry of Health and Human Services, Somali Federal Republic (ref. MOH&HS/DGO/1944/Dec/2018).

15. Who provided the data for this study?

Most of the data used for this study were already in the public domain. In addition, we obtained reports and raw datasets of ground mortality surveys, population displacement, food distributions, humanitarian presence, vaccination and epidemic surveillance from United Nations agencies.

16. Why are you not releasing all of the data used for the analysis?

We cannot release non-public datasets we used unless the agency that provided them feels it appropriate to do so. Data are not released because of an overriding concern with preserving the safety and ability to operate of source agencies. However, scientists interested in replicating our findings are encouraged to get in touch with the study team (francesco.checchi@lshtm.ac.uk), who will explore on a case-by-case basis whether data sharing is an option.

Of note, we are releasing all statistical code used for the analysis.

17. Have patients' or people's data been used without their consent?

The study used only anonymised data collected for routine programme purposes. No individual patient data containing identifier information have been used, and only aggregate data (e.g. by district) are used for the statistical model.

18. What is the estimate actually useful for?

The death rate is a key indicator to benchmark the gravity of a crisis. In past crises (e.g. Eastern DRC, Northern Uganda, Darfur), crisis-wide mortality estimates have had a beneficial effect in terms of attracting more resources for humanitarian action; estimates for Somalia in 2010-2012 have been used prominently to advocate for earlier response to the food insecurity crisis of 2016-2017. We cannot say whether combatants or other actors with the ability to influence the armed conflict will use our findings to pursue conflict resolution or, at a minimum, stop attacking civilians, but our findings do provide a very objective reason for urgently adopting such steps.

19. Do these findings indicate that the humanitarian response in Somalia is inadequate?

The findings themselves do not allow us to infer precisely to what extent the humanitarian response has been impactful. However, the high death toll obviously indicates considerable unmet need, i.e. a response that has been unable to avert a large number of excess deaths through public health, food security and other humanitarian services. We did not specifically evaluate the humanitarian response in terms of its performance.