Mortality among populations of southern and central Somalia affected by severe food insecurity and famine during 2010-2012

A Study commissioned by FAO/FSNAU and FEWS NET

from the

London School of Hygiene and Tropical Medicine

and the

Johns Hopkins University Bloomberg School of Public Health

Rome, Washington, 2 May 2013

USAID FROM THE AMERICAN PEOPLE

Mortality among populations of southern and central Somalia affected by severe food insecurity and famine during 2010-2012

Lead authors:

Francesco Checchi

London School of Hygiene and Tropical Medicine, London, UK

and

W. Courtland Robinson

Johns Hopkins University Bloomberg School of Public Health, Baltimore, USA

Acknowledgements

The United Nations Food and Agricultural Organization's Food Security and Nutrition Analysis Unit for Somalia (FAO/FSNAU) and the USAID-funded Famine Early Warning System Network (FEWS NET) commissioned this study from the London School of Hygiene and Tropical Medicine (LSHTM) in the United Kingdom, and from the Johns Hopkins University Bloomberg School of Public Health in Baltimore, USA. The LSHTM (lead author: Francesco Checchi) carried out estimation of excess mortality, while the Johns Hopkins Bloomberg School of Public Health (lead author: W. Courtland Robinson) provided demographic analysis of population denominators, as part of a parallel study on the dynamics of migration during the Somalia 2011 crisis.

The authors sincerely thank Abdihamid Warsame and Séverine Frison from the LSHTM; Andrew J. Tatem of the University of Southampton, Southampton, UK; Linnea Zimmerman from Johns Hopkins University Bloomberg School of Public Health, Baltimore, USA; and Harry Shannon from McMaster University, Hamilton, Canada, for their contributions and reviews of various sections of the report.

The study primarily relied on survey data generated by the FAO/FSNAU in Somalia. We are most grateful to the FSNAU field researchers, too many to name individually, who led the more than two hundred surveys included in this analysis, often conducted in difficult circumstances; and to the many partner agencies who supported the implementation of these surveys. Several FSNAU staff made specific contributions to the analysis, including Ahono Busili, Louise Masese, Tamara Nanitashvili, Elijah Odundo, Kamau Wanjohi and Alex Koton. We are also thankful to Grainne Moloney (formerly with FSNAU) for having initiated this study and coordinated survey data collection during the 2010-2012 emergency period. Several FAO staff, including Mark Smulders, Daniel Molla, Panagiotis Karfakis and Mohamed Ag-Bendech made useful technical and editorial comments to improve the report.

At FEWS NET, we are grateful to Christopher Hillbruner for having coordinated study finalisation and for insightful comments; to Erin Martin for useful suggestions on the manuscript; and to Sonja Perakis and Blake Stabler for valuable advice on composition and interpretation of food security indicators.

We are also grateful to agencies and individuals who supplied data used in the analysis, including the UN Office for Coordination of Humanitarian Assistance, OCHA (Stephanie Bulgjacic); Epicentre (Fabienne Nackers, Jonathan Polonsky, Andrea Bernasconi); AfriPop; and staff members of the UNHCR Population Movement Tracking (PMT) unit, particularly Conor Flavin, for sharing data on Somali Internally Displaced People (IDPs) and refugees in Ethiopia and Kenya.

The authors adapted open source R statistical code on cross-validation and random effects models written by Jonathan Starkweather and Ben Bolker.

Executive summary

Background

Between late 2010 and early 2012, southern and central Somalia experienced severe food insecurity and malnutrition precipitated by a prolonged period of drought resulting in the poorest harvests since the 1992-1993 famine. The effects of the drought were compounded by various factors including decreased humanitarian assistance and increasing food prices. Furthermore, this emergency occurred against a backdrop of heightened insecurity and persistent high levels of acute malnutrition, and affected populations whose resilience mechanisms had already been weakened over the past few years by a protracted crisis featuring a combination of armed conflict, natural disasters and adverse economic conditions. The evolving humanitarian emergency situation was detected in a timely way by existing early warning systems run by the United Nations Food and Agriculture Organisation's Food Security and Nutrition Analysis Unit for Somalia (FAO/FSNAU) and the USAID-funded Famine Early Warning Systems Network (FEWS NET). By July 2011, based on criteria established by the multi-partner Integrated Food Security Phase Classification (IPC, an analysis template used globally for determining relative severity of food insecurity), the United Nations declared famine in several regions of Somalia. Based on further data and information collected on food security and nutritional status, disease and mortality, additional regions were designated as famine-affected over the subsequent two months. As a result of this emergency, during 2011 large numbers of people were internally displaced within Somalia or migrated to already overcrowded refugee camp complexes in Dollo Ado (Ethiopia) and Dadaab (Kenya). Measles, cholera and other epidemics, which typically accompany situations of greatly deteriorated nutritional status of the population, were also reported from nearly all affected regions.

There is consensus that the humanitarian response to the famine was mostly late and insufficient, and that limited access to most of the affected population, resulting from widespread insecurity and operating restrictions imposed on several relief agencies, was a major constraint. Based on numerous individual surveys conducted throughout southern and central Somalia by FAO/FSNAU and partners, and in the refugee camps by various other agencies, it was assumed that the impact of these combined events on human health would be severe. Indeed, the surveys indicated that both death rates and the prevalence of acute malnutrition among children were well in excess of emergency thresholds, and far surpassing any value observed in Somalia during the previous five years, at least. However, the estimates of mortality from available surveys did not cover the entire affected population, nor the full period during which food security emergency and famine conditions occurred. Indeed, during the emergency, the United Nations did not issue real-time death toll estimates. In 2012, improved conditions presented an opportunity to take stock of lessons learned and document the effects on health and mortality of exposure to severe food insecurity and malnutrition during 2010 and 2011. Therefore, a study was commissioned by FAO/FSNAU, with substantial technical and financial support from FEWS NET, in order to produce an estimate of the number of deaths during the 2011 Somalia famine, and among refugees displaced to camps in Ethiopia (Dollo Ado) and Kenya (Dadaab).

In this report, we refer to the 2010-2012 events as "severe food insecurity and famine." Although imperfect, this term encompasses famine, while also capturing regions and periods of time that were not classified as being in a state of famine, but nonetheless experienced extraordinary food insecurity stress, as well as other adverse conditions, such as reduced access to humanitarian assistance, that would plausibly result in abnormally high levels of mortality. The study provides estimates of overall and excess mortality over a period of 19 months between October 2010 and April 2012 during which time severe food insecurity and famine conditions prevailed.

Methods

Our analysis considered the 28-month period from April 2010 to July 2012 inclusive. The starting point for the analysis was determined by when the prevailing food security situation first began to deteriorate, while the end point reflected more pragmatically the timing of when this study was conducted. In practice, harvest and market data indicated that by July 2012, food security in nearly all regions of southern and central Somalia had returned to pre-emergency levels.

We sought to quantify deaths that occurred above and beyond the number expected in the absence of the emergency (also known as excess deaths or excess mortality). Excess mortality can be estimated by combining three pieces of information: (i) the total death rate (i.e. number of people dying per population and per unit time) during the emergency period; (ii) what this death rate would have been if the emergency had not happened (this is also known as "baseline" mortality); and (iii) the population living in the affected areas. Excess mortality is then given by the difference between the total and baseline death rates, multiplied by the population living in the region of analysis.

These pieces of information were not immediately available for Somalia due to lack of systematic birth and death registration and incomplete tracking of population movements, and hence, had to be estimated. Our study used a variety of previously collected data and statistical techniques in order to do so. The 2010-2011 drought and crop failure affected mainly southern Somalia (Bakool, Banadir, Bay, Gedo, Hiran, Lower Juba, Middle Juba, Lower Shabelle and Middle Shabelle regions), and to a lesser extent, the central regions of Galgadud and Mudug, as evident from multiple data sources. We included all of these regions within the analysis, as well as the 11 refugee camps around Dollo Ado, Ethiopia and Dadaab, Kenya. Our estimation was restricted to refugees who arrived to these camps during the analysis period. Other regions were excluded from estimation as they were far less affected by the drought, but data from surveys conducted in these regions were used to inform statistical models (see below).

For the purpose of analysis, we divided the population of each of the included regions into the major livelihood types present in the region (pastoralist, agro-pastoralist, "riverine" or agriculturalist, internally displaced [IDPs] and urban), thereby creating 42 separate "strata" within Somalia, plus a further 11 strata consisting of each refugee camp in Ethiopia or Kenya. We further split the period of analysis within each stratum into months (i.e. "stratum-months"). While this division was dictated by the structure of the datasets that we analysed, we also sought to provide estimates with relatively detailed geographic and period resolution.

For stratum-months within Somalia, we estimated total death rates "directly" whenever a field household survey had been done covering these stratum-months; alternatively, we estimated total death rates "indirectly" by means of a statistical model. Nutrition and mortality surveys were carried out by the FAO/FSNAU in partnership with other agencies; these applied standard methods for estimating mortality in emergencies. Briefly, these consisted of selecting a sample of households and interviewing each about births, deaths and other changes to the composition of the household in the prior three months. Wherever possible, we re-analysed survey datasets to confirm the original results.

Since 2007, nutrition and mortality surveys were done with some seasonal regularity throughout Somalia, though logistics and security conditions meant that not all areas were covered. The statistical model we used supplemented these data gaps by predicting mortality based on variables that were collected continuously, even when and where surveys were not done, and that were plausibly expected to have a strong association with the risk of dying. After various model construction steps, these variables included the survey location (southern and central Somalia versus Somaliland and Puntland); the livelihood type; the occurrence of incidents of armed conflict; the presence of confirmed epidemics; the availability of humanitarian assistance; and the so-called "terms of trade", an indicator of food security built from market data that, in this case, quantifies for any given time point the purchasing power (in terms of Kcal of staple cereal) of a typical daily wage for low-income labourers in Somalia. As above, terms of trade as a measure of food access were taken to be a proxy of broader food security conditions, in the absence of more complete data, such as dietary diversity and the prevalence of malnutrition. (These variables are highly relevant, but were not collected often enough to be used in the model). Models for all-ages and under 5-year mortality were built using data from all available mortality surveys from southern and central Somalia, Somaliland and Puntland since 2007 (205 surveys altogether). When selecting an appropriate model, we verified how accurately the combination of these variables would be expected to predict mortality in stratum-months without survey data.

For refugee camps, we identified all available sources of mortality data covering newly arrived refugees during the period by reviewing the published literature, conducting an internet search and contacting key agencies. These data were very sparse (two surveys and two prospective surveillance systems in Dollo Ado camps, Ethiopia; two surveys in Dadaab, Kenya) and did not allow for either direct estimation or statistical modelling. Instead, for each camp we made conservative death rate assumptions bounded by a pessimistic (worst-case) and optimistic (best-case) scenario, informed by available data and circumstantial evidence on living conditions (e.g. overcrowding, service availability) in that camp.

Available surveys showed that mortality in Somalia was chronically elevated compared to regional averages, even before the emergency period analysed here. Problems with correctly defining this "baseline" (which can only be estimated on the basis of assumptions, since deaths that would have occurred in the absence of the emergency are immeasurable) have affected previous studies of mortality in settings of crisis. Moreover, our analysis needed to isolate excess mortality occurring as a result of conditions (food insecurity, reduced humanitarian assistance, civil strife) that, compared to an epidemic disease or deaths due to war trauma, are less easily definable in quantitative terms.

In order to more accurately represent the extent of uncertainty in our analysis, we explored three alternative methods for estimating the baseline. These are presented more fully in the main body of the report. Of these, method 3 was considered the most plausible, as it applied the same statistical model as above to predict what mortality would have been in each stratum if food security (approximated through terms of trade as a measure of food access) had remained constant instead of deteriorating, and if epidemics likely to have been precipitated by severe malnutrition had not occurred. Results below are based on this third method, unless otherwise specified. We also applied baseline estimates from within Somalia to the refugee populations, but used baseline values that reflected more specifically the refugees' regions of origin.

Lastly, for both southern and central Somalia and the camps, we reviewed available population and displacement figures, and applied demographic models informed by a series of assumptions to come up with a dataset of monthly population figures. This dataset was based principally on (i) 2005 population estimates from a UN study in Somalia; (ii) 2010 fine resolution estimates of population per area produced by the AfriPop project; and (iii) data on displacement shared by the United Nations High Commissioner for Refugees. The population figure used in this study is a demographic modelling estimate conducted for the purposes of this analysis alone and it should not be used in lieu of official planning figures for Somalia, which are currently subject to a revalidation exercise. It should be noted that this exercise has resulted in population estimates that are 20-25 percent higher than the 2005 UN estimates, which hitherto has been the main reference point for population statistics in Somalia, including by FAO/FSNAU and FEWS NET.

We inputted all of the above data into a computer simulation that generated best estimates of the excess death toll, as well as margins of error around these best estimates. Data and statistical programming codes are made available alongside this report.

Results

Based on the most plausible set of population denominator data, we estimated that 258,000 (244,000 to 273,000) excess deaths attributable to the emergency occurred in southern and central Somalia between October 2010 and April 2012 inclusive, of which some 52% (133,000) among children under 5 years old (Table 1). These estimates are based on the most robust approach for calculating baseline mortality (i.e. method 3) among those explored. The two alternative baseline estimation methods, in which we have less confidence, resulted in higher baseline mortality estimates, so that altogether estimated values of the excess death toll ranged from as low as 143,000 to as high as 273,000.

Table 1. Summary of estimated excess death toll by region or livelihood type, as an absolute figure and as a percentage of the population, for all ages and children under 5 years old, in southern and central Somalia (October 2010 to April 2012). Results from baseline method 3 are shown.

Region, livelihood	Estimated excess death toll (95% percentile interval)		Estimated percentage of the population that died in excess of the baseline (95% percentile interval)	
	all ages	under 5y	all ages	under 5y
Regions				
Bakool	10,700 (8300 to 13,200)	5600 (4300 to 6800)	2.9	7.3
Banadir	58,000 (48,000 to 68,900)	27,200 (22,100 to 32,400)	6.2	16.6
Bay	20,700 (18,600 to 22,900)	20,900 (18,200 to 23,900)	6.4	12.7
Galgadud	4200 (2000 to 7700)	4300 (2700 to 8800)	1.0	4.7
Gedo	17,800 (16,000 to 19,700)	6400 (5300 to 7900)	5.0	7.1
Hiran	7400 (5600 to 9200)	2800 (1900 to 3800)	1.8	3.2
L. Juba	7900 (5900 to 10,100)	3300 (2100 to 4800)	1.8	3.2
L. Shabelle	96,200 (89,200 to 104,200)	42,900 (38,900 to 47,000)	9.0	17.6
M. Juba	11,900 (10,200 to 14,100)	5900 (4800 to 7300)	4.5	9.8
M. Shabelle	22,200 (19,800 to 24,900)	12,900 (11,600 to 14,300)	3.7	9.7
Mudug	300 (-2200 to 3500)	100 (-1500 to 3500)	0.1	0.1
Total	**257,900 (243,600 to 272,700)**	**132,900 (124,700 to 142,300)**	**4.6**	**10.1**
Livelihood types				
agro-pastoralist	69,800 (64,100 to 76,100)	41,600 (37,700 to 45,600)	5.0	11.2
IDP	68,700 (60,900 to 76,100)	34,400 (29,900 to 39,100)	6.5	14.6
pastoral	34,200 (29,800 to 39,600)	18,700 (15,500 to 24,400)	2.3	5.1
riverine	46,800 (42,200 to 51,700)	22,400 (19,500 to 25,600)	5.1	11.1
urban	38,200 (30,000 to 47,200)	15,500 (12,300 to 19,400)	4.8	10.6
Total	**257,900 (243,600 to 272,700)**	**132,900 (124,700 to 142,300)**	**4.6**	**10.1**

The highest estimated death tolls were in Banadir, Bay and Lower Shabelle regions. (Note that rare negative numbers in the lower limit of the margin of error indicate that fewer deaths were predicted to occur during the emergency than if baseline conditions had been maintained).

The full toll of the emergency is perhaps easier to visualise when considering the percentage of the population estimated to have died as a result: these are about 4.6 percent overall (Table 1), peaking in Lower Shabelle at 9 percent for all ages and at 17.6 percent among children under 5 years old.

Prior to 2011, available surveys done in Somalia yielded a crude death rate for all ages (CDR) and an age-specific death rate for children under 5 years old (U5DR) that remained consistently below 2 and 4 deaths per 10,000 people per day respectively, though many of the values recorded were already indicative of emergency conditions. In southern and central Somalia (but not in the rest of the country), a striking peak in recorded mortality is apparent in July-October 2011, with individual survey CDRs and U5DRs reaching 5-6 and 10-15 per 10,000 per day respectively, and a CDR value of around 2.5 per 10,000 per day for southern and central Somalia, as estimated for all strata combined through direct and indirect methods. By contrast, the counter-factual baseline CDR was estimated to oscillate between 0.5 and 0.8 throughout the period, while the Sub-Saharan Africa 2010 average was 0.37. A higher baseline in southern and central Somalia compared to regional averages likely reflects underlying factors related to the chronic crisis, including inappropriate feeding practices, limited access to health infrastructure, inadequate water and sanitation services, armed conflict, etc.

As shown in Figure 1, excess mortality visibly began to accrue in October 2010. (Note that in order to emphasise the difference between the emergency period and the estimated baseline, this figure shows all deaths based on the estimated CDR, not just excess deaths. See the full report below for an alternative version showing monthly excess mortality only). Between May and October 2011 inclusive, greater than 20,000 excess deaths per month (i.e. the difference between total and baseline deaths in Figure 1) were estimated to occur in southern and central Somalia. While this is considered the main famine period, it should be noted that excess mortality in the population began to rise well before, as conditions deteriorated over time, including in areas where famine was not declared.

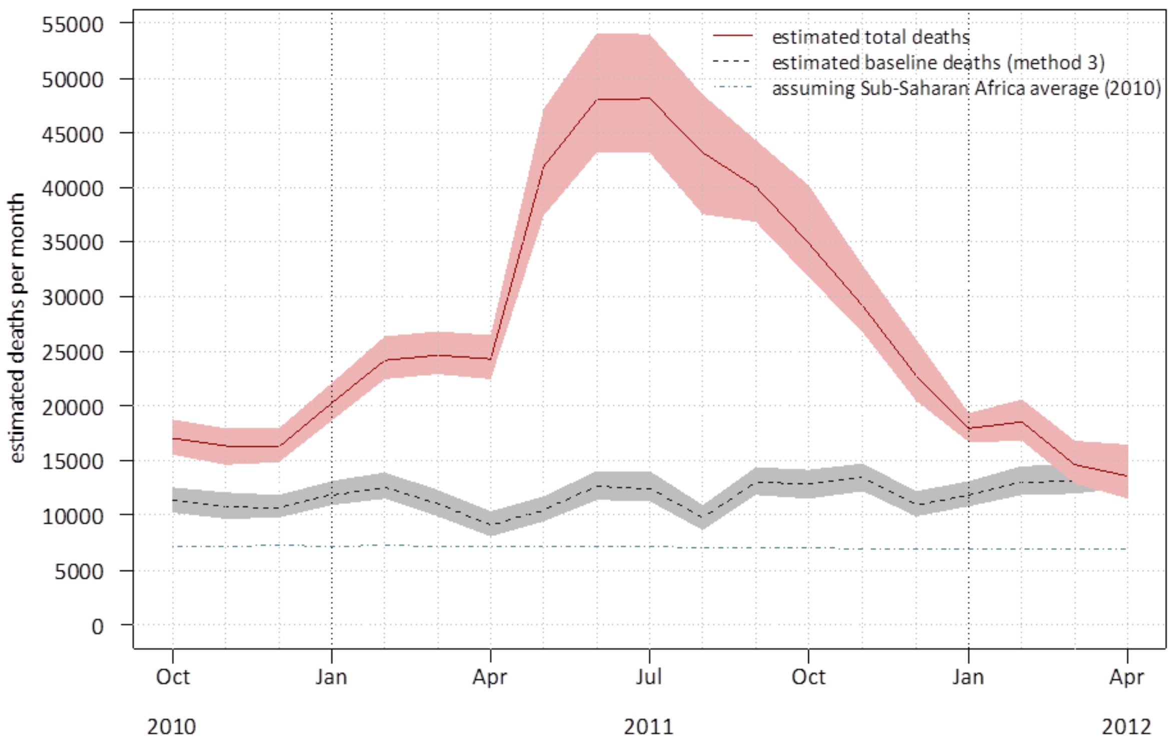

Figure 1. Estimated number of deaths per month during the emergency, compared to deaths that would have occurred if the emergency had not taken place, according to baseline method 3 criteria. Shaded areas indicate 95% percentile intervals around the point estimates. Mortality that would have resulted from a death rate equal to the Sub-Saharan average is also shown to facilitate interpretation. Note that baseline death tolls are slightly imprecise as they do not account for varying death rates during the period.

In Dollo Ado and Dadaab refugee camps, excess mortality estimates ranged between -1000 to 5700 and -1300 to 8800 respectively, suggesting that either fewer or many more deaths occurred in these camps than would have if no emergency had occurred: due to limitations in the available mortality and population data, no single best estimate can be provided for the camps, though most of the uncertainty range clearly falls within the positive region, indicating that many excess deaths (probably in the thousands) may have occurred in these camps as well. Additional results are presented in the body of this report.

Conclusions

This study suggests that severe food insecurity and famine in southern and central Somalia over a 19-month period in 2010-2012 resulted in a very large death toll, with a majority of excess deaths among children under 5 years old and a peak in excess mortality during mid-late 2011, which coincides with the declared famine period. Our study was not designed to clearly attribute portions of this death toll to any single cause, such as epidemics or high food prices. Rather, our estimates should be viewed as the combined impact of drought, reduced humanitarian assistance, high food prices and civil strife in the affected regions, and the downstream consequences of the above factors (such as disease epidemics), all in a context of persisting and/or worsening insecurity. Taken together, however, our analysis supports the notion that famine and pre-famine severe food insecurity conditions occurred in large parts of southern and central Somalia during the above period, and in 2011 particularly.

Our estimate of about 244,000 to 273,000 excess deaths (using baseline model 3) is similar to that for the 1992-1993 famine in Somalia. However, percent mortality during 2010-12 was about half, the population affected was larger and the definitions of famines were not consistent in the two events. Peak death rates were similar to those observed in other recent famines in Ethiopia and South Sudan. Excess child deaths, as estimated by this study, are about two to three times the annual amount in all industrialised countries combined. More than 90% of estimated deaths occurred inside south and central Somalia, where internally displaced people and riverine populations, particularly in Lower Shabelle, Bay and Banadir regions, were disproportionately affected. Notably, excess mortality began to increase in late 2010, well before humanitarian relief began to be mobilised, and possibly earlier than previously recognised.

Our study relied on a unique dataset of more than 200 nutrition and mortality surveys conducted by FAO/FSNAU and partners in difficult conditions within Somalia using a mostly standardised approach. These were complemented by rich data on food security and other variables generated by the FAO/FSNAU and FEWS NET. These data enabled us to adopt a robust statistical estimation approach that, for Somalia, generates estimates with a mostly quantifiable amount of accuracy.

Findings are highly dependent on the accuracy of mortality survey data, variables used for statistical modelling, and population denominators, which we could only partly assess. Estimates for the refugee camps should be considered far less robust, as they are based on investigator assumptions rather than a formal statistical approach. In general, our estimates would have been more accurate if we had had access to additional indicators of food insecurity and in more frequent intervals (monthly), other than terms of trade, and more nuanced data on the actual availability of humanitarian assistance in different regions and periods. We discuss a number of these limitations in the main body of the report. Aside from these limitations, the true number of excess deaths may be greater than that estimated here, as we did not consider refugees and other migrants other than to Kenya and Ethiopia camps, as well as more delayed mortality effects that could plausibly manifest themselves for years following the end of the emergency.

Our findings place the 2010-2012 Somalia emergency among the most severe affecting humanity over the last decades, at least using a mortality metric. They broadly illustrate the potential health effects of drought and large-scale food insecurity in the absence of a timely and adequate humanitarian response. This evidence should be used to ensure such deficiencies never occur again in the future. Our estimates provide renewed justification for ensuring that adequate humanitarian assistance reaches all populations in Somalia, and for vigorously pursuing a resolution to the ongoing armed conflict.

Table of contents

List of abbreviations

ACLED	Armed Conflict Location and Event Dataset
ARRA	Administration of Refugee and Returnee Affairs (Government of Ethiopia)
$AToT_W$	Absolute terms of trade (Kcal cereal equivalent of one average daily wage)
$AToT_G$	Absolute terms of trade (Kcal cereal equivalent of one local quality goat)
CDR	Crude death rate
CI	Confidence interval
CMB	Cost of Minimum Expenditure Basket
CPI	Consumer Price Index
ENA	Emergency Nutrition Assessment software
FAO	United Nations Food and Agriculture Organization
FEWS NET	Famine Early Warning Systems Network
FAO/FSNAU	Food Security and Nutrition Analysis Unit
HIS	Health Information System, United Nations High Commissioner for Refugees
IDP	Internally displaced person/people
IPC	Integrated Phase Classification
LOOCV	Leave one out cross-validation
MSE	Mean squared error
MSF	Médecins Sans Frontières
MUAC	Middle upper arm circumference
NICS	Nutrition Information in Crisis Situations
OLS	Ordinary least squares
PPS	Probability proportional to size
$RToT_W$	Terms of trade relative to start of time series (cereal equivalent of daily wage)
$RToT_G$	Terms of trade relative to start of time series (cereal equivalent of local goat)
SD	Standard deviation
SE	Standard error
UNDP	United Nations Development Programme
UNHCR	United Nations High Commissioner for Refugees
U5DR	Under 5 years death rate
WFP	United Nations World Food Programme

List of tables and figures

1 Introduction

1.1 Historical background

Since the fall of the Siad Barre government in 1991, much of Somalia has been affected by recurrent waves of armed conflict and insecurity. At the time of writing, southern and central Somalia had for about five years been disputed among different parties to the conflict, including the Federal Government of Somalia (broadly recognised internationally) and a number of armed groups and local administrations. Ethiopian and Kenyan armed forces and an African Union force are present inside the country with various stated objectives. The United States and other western countries also carry out occasional armed interventions. The situation is markedly different in the northwest and northeast of the country, where separate administrations have been set up (Somaliland and Puntland respectively). Though these regions have disputed borders and limited international recognition, they have over the past decade enjoyed a considerable degree of stability, economic growth and infrastructure development.

Food insecurity and high rates of malnutrition have been a major threat to the livelihoods and health of Somalis since 1991. While important factors such as climate abnormalities, limited farming technology and maternal, infant and young child feeding practices modulate its severity, armed conflict is generally understood to be a key underlying cause [1]. Moreover, its effects on health are exacerbated by other factors brought about by war, including limited preventive and curative health services, poor water and sanitation, increased disease transmission due to displacement, and exposure to both physical and mental trauma.

These adverse conditions have led to waves of emigration and forced displacement. Within the country, about 1.4 million internally displaced persons (IDPs) were estimated by UNHCR to be within all of Somalia (including Somaliland and Puntland) as of mid-2010, while a further 600,000 Somalis were counted as refugees abroad (see below). Refugee camps in Kenya (the Dadaab complex) and Ethiopia (Dollo Ado complex) host large numbers of Somalis, many of whom are recent arrivals since 2010. During the more than 20-year old protracted crisis period, Somalis living throughout the world (known as the 'Somali Diaspora') have played a pivotal role in supporting livelihoods and development in Somalia, largely through cash remittances.

A severe famine occurred in the aftermath of the disintegration of the Somali state in 1992-1993. This event was precipitated by armed conflict and featured high levels of hyper-inflation, nearly complete crop failure, 50-70% livestock losses, and closure of ports to food importation [2]. At the time, crude death rates of 7 to 23 per 10,000 person-days, among the highest ever recorded globally, were measured by the Centers for Disease Control and other epidemiologist teams, with a prevalence of acute malnutrition of 40-70% [3,4].

Other periods of severe food insecurity have affected Somalia's population since that time. Both death rates and global acute malnutrition prevalence have remained elevated compared to countries in the region [5].

In 2008, southern and central Somalia experienced severe hyperinflation partly linked to the global increase in food prices and to poor harvests. Analyses of this food insecurity period and empirical data from FAO/FSNAU (see below), suggest that the population was able to rely on coping mechanisms, thereby minimising malnutrition and mortality. Critically, at the time relief agencies were fully operational and providing assistance in nearly all of Somalia. It is suggested, however, that this crisis left large sections of the Somali population far more impoverished than previously and less able to cope with further stress. Moreover, prices stabilised at a higher level following this crisis [1].

1.2 The 2010-2012 severe food insecurity and famine period

A La Niña meteorological event created drought conditions throughout the Horn of Africa in 2010-2011. Southern and, to a lesser extent, central regions of Somalia experienced particularly reduced rainfall, resulting in a high proportion of crop failure (Figure 2) compared to previous years.

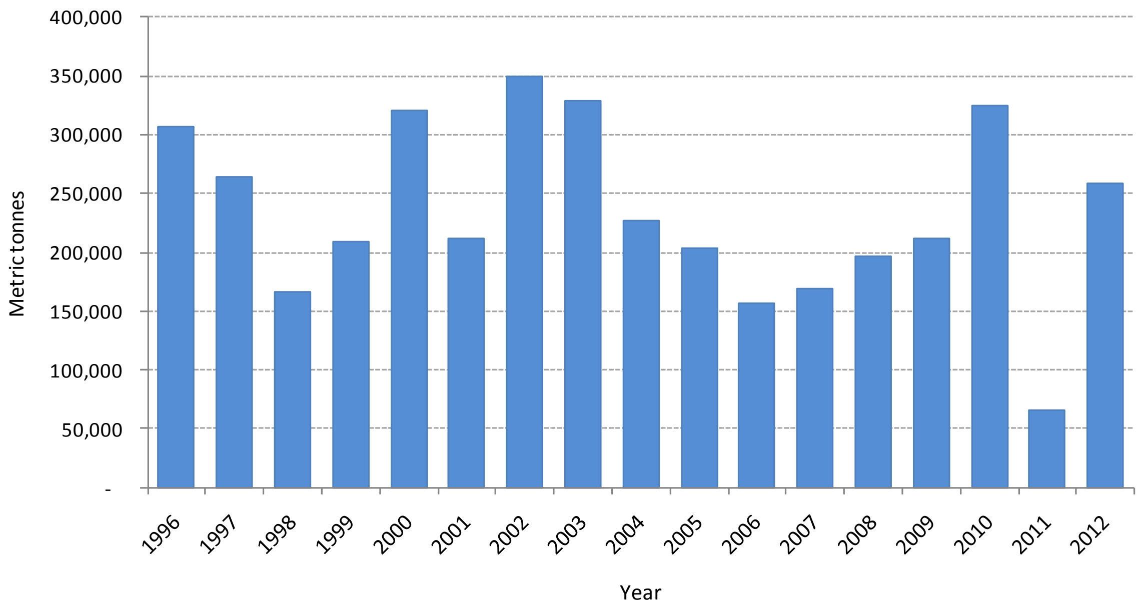

Figure 2. Estimated cereal production quantities by year (Somalia southern regions only). Source: FAO/FSNAU Technical Series Report No. VI (42) [2].

This development occurred just as the people of southern Somalia were becoming increasingly isolated from food and livelihoods assistance, following the progressive withdrawal of humanitarian agencies from most of southern Somalia, and the interruption of humanitarian funding to those regions by some donors [6]. Access to southern Somalia remained extremely difficult for a large number of agencies throughout 2010 and early 2011, including for major actors involved in food security and nutrition related interventions.

During this period, repeated warnings of deteriorating conditions were issued by two technical bodies tasked with regular monitoring of food security and nutrition in Somalia: the UN Food and Agriculture Organization's Somalia Food Security and Nutrition Analysis Unit (FAO/FSNAU) and the Famine Early Warning Systems Network (FEWS NET). Data informing these bulletins included meteorological and rainfall data, price trends at networks of sentinel markets, data on harvest size for all the major staple cereals, qualitative information collected by experienced Somali food security researchers, and multi-indicator household surveys typically performed twice per year after harvests, that estimated household food availability indicators, anthropometric status and both all-age and child mortality. These estimates and bulletins were publicly available on the FEWS NET (www.fews.net/Pages/country.aspx?gb=so&l=en) and FAO/FSNAU (www.fsnau.org/products) websites and shared widely in humanitarian fora [7,8].

FAO/FSNAU and FEWS NET jointly classify overall severity of food insecurity within any given area of Somalia based on a globally applicable Integrated Food Security Phase Classification (IPC) system that was first adopted in Somalia in 2006 after extensive expert consultation. The IPC combines all available data sources (ranging from food security and livelihoods data and information to anthropometry and mortality) into an overall grading, ranging from 1 (generally food secure) to 5 (famine/humanitarian catastrophe). As of July 2011, all of southern Somalia was classified in phase 4 or 5, an unprecedented situation since the IPC's adoption (Figure 3).

Concurrently, an increasing flow of refugees crossed into camps around Dadaab, Kenya and Dollo Ado, Ethiopia, while large scale internal displacement was reported within Somalia, including to the already overcrowded Afgooye corridor and nearby Mogadishu. In July 2011, following a round of surveys throughout Somalia revealing alarmingly high prevalence of acute malnutrition and population death rates, the United Nations declared a famine in parts of Somalia in accordance with the IPC classification - the first instance of this official designation since the famine in the Gode region of southern Ethiopia in 2002 (the Gode famine preceded the IPC framework). As our analysis scope also includes areas of Somalia that experienced unusually severe food security stress without meeting the IPC famine criteria (see below), we will hereafter use the broader term **"severe food insecurity and famine"** to characterise the events of 2010-2012. However, we recognise that no single term fully captures the spectrum of conditions and risk factors that affected the Somali population during this period. Specifically, our chosen terminology is not meant to imply that food insecurity alone was responsible for the emergency (as indeed other factors such as reduced assistance and greater insecurity played a critical role). Rather, it indicates our interest in documenting all excess mortality during this period, and not just that occurring in regions or periods officially designated as famine-affected.

The evolution of the severe food insecurity and famine period over 2010 and 2011, and the effectiveness of humanitarian responses mounted to address it, have been analysed extensively (see in particular a special series of the Global Food Security journal [9]). We will not reproduce such analyses here. Briefly, further rounds of surveys in August and October broadly showed a continued acute picture, with famine declarations in new areas. The FAO/FSNAU's operations, along with those of most humanitarian actors, were restricted after this time and surveys were no longer possible in much of southern Somalia, though market data continued to be collected. It is widely agreed that, despite extensive early warning, most humanitarian agencies responded late to the emergency. In southern Somalia, food and livelihoods relief as well as management of acute malnutrition interventions were carried out by only a handful of agencies, notably the International Committee of the Red Cross, until it, too, was asked to cease operations in late 2011. Numerous assessments also suggested acute emergency conditions among new refugees in Kenya and Ethiopia, and among IDPs around Mogadishu. Refugees, IDPs and the general population all experienced country-wide epidemics of measles, cholera and other diseases during this period, while insecurity increased due to a shifting balance of power between parties to the conflict as well as a Kenyan invasion of southern Somalia. The emergency appeared to subside in early 2012 following bumper crops, although improvements in health and nutritional status were difficult to ascertain in the absence of surveys.

1.3 Study aim

During the humanitarian emergency and famine periods of 2010-12, the United Nations did not issue real-time death toll estimates. In 2012, improved conditions presented an opportunity to take stock of lessons learned and document the effects on health and mortality of exposure to severe food insecurity and malnutrition. This study complements existing evaluations by presenting statistical estimates of the number of people who are likely to have died as a result of an extended period of severe food insecurity and famine during the 19-month period lasting from October 2010 to April 2012.

Specifically, we aimed to estimate excess mortality among people of all ages and children under 5 years (under 5y) old in Somalia and among refugees that left Somalia for camps in Ethiopia and Kenya during the emergency period.

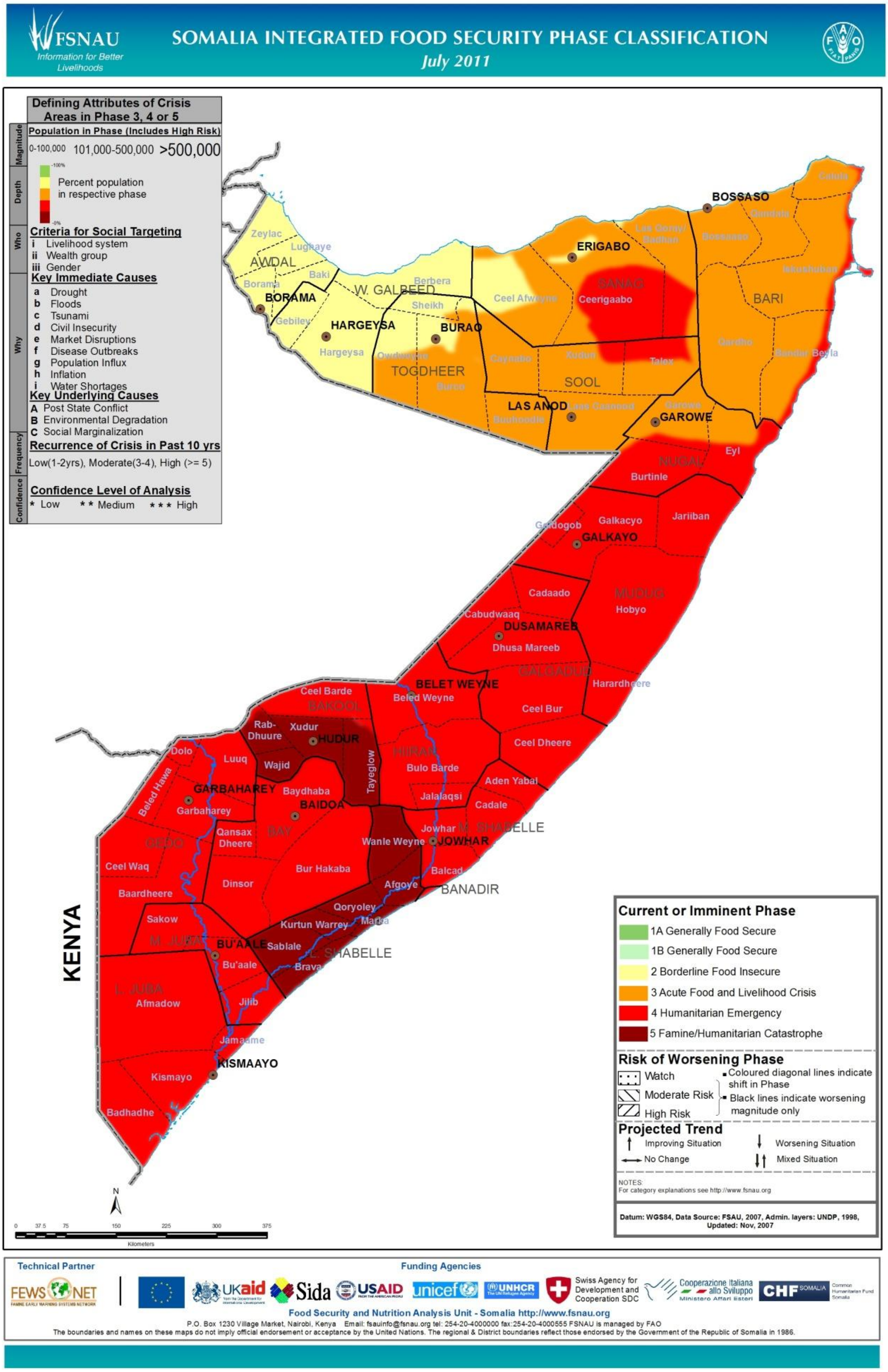

Figure 3. FSNAU/FEWS NET Integrated Food Security Phase Classification map for Somalia, July 2011. Source: FAO/FSNAU.

2 Methods

2.1 Study design

We adopted different approaches to estimation for southern and central Somalia and for the refugee camp populations in Ethiopia and Kenya. For southern and central Somalia, we divided the population into strata defined on the basis of region and livelihood. We estimated total crude and under 5y death rates (CDR, U5DR) for each month in each stratum over the analysis period (late 2010 to early 2012: see below), either directly if the stratum-month fell within the sampling frame of one or more mortality surveys; or indirectly using a predictive statistical model informed by variables plausibly associated with mortality and collected continuously throughout the emergency. For the refugee camps, no formal statistical approach was deemed feasible due to the sparseness of data. Instead, we came up with optimistic and pessimistic projections of total mortality based on expert judgment informed by available mortality information and circumstantial evidence.

We then subtracted from these total death rates values of baseline mortality computed through different approaches, so as to estimate excess death rates attributable to emergency conditions. Separately, we used demographic models and available data on displacement to construct a comprehensive dataset of population denominators for Somalia and refugee camps, which we multiplied by excess death rates so as to estimate death tolls attributable to the emergency.

For brevity's sake, parts of the methods are summarised in this section; they are presented in greater detail in a Statistical Annex. Demographic methods used to construct population estimates are likewise summarised here, and will be published separately. Data (with the exception of primary mortality datasets) and statistical programming code (written in R language) are also made available alongside this report to facilitate external scrutiny of our estimates. These data will be available for download at www.fsnau.org and wwww.fews.net. All analyses were carried out using R, version 2.12.1 [10], except for population estimation, carried out using Stata 11 (StataCorp, College Station, TX, USA). The study was entirely desk-based and did not collect primary data. Secondary analysis of existing mortality survey data was approved by the ethics committee of the London School of Hygiene and Tropical Medicine. Other datasets used were already in the public domain and were either aggregate and anonymous, or did not contain data on human subjects.

2.2 Populations and periods included in estimation

2.2.1 Study population

Somalia is divided into regions and districts. The 2010-2011 drought and crop failure, and their effect on local staple food prices, affected mainly southern Somalia (Bakool, Banadir, Bay, Gedo, Hiran, Lower Juba, Middle Juba, Lower Shabelle and Middle Shabelle regions), and, to a lesser extent, the central regions of Galgadud and Mudug, as evident from multiple data sources including rainfall and market data [8]. We thus included all of those regions in the analysis, as well as all refugee camps in Dollo Ado, Ethiopia and Dadaab, Kenya. Our estimation was restricted to refugees that arrived to these camps during the analysis period (see below). Other parts of Somalia were not included in the analysis, but data from surveys conducted in these regions were used to inform statistical models. Note that in different livelihood zones of Somalia, the effects of crop failure may vary. For example, crop producers in riverine areas might either increase their income due to rising prices or lose out on income due to reduced supply, while urban populations would have a harder time purchasing food. We considered that, through any such mechanism, deleterious effects on mortality could have occurred.

Within Somalia, the sole source of direct mortality estimates was household surveys conducted by the FAO/FSNAU in collaboration with various partners. Since 2007, these surveys have aligned with wider food security assessments by using livelihood, in addition to administrative boundaries, to define their geographic unit of analysis. Accordingly, all of Somalia is informally divided into 34 livelihood zones reflecting the main occupation of people within them (see Annex). In southern and central Somalia, three distinct rural livelihood types can be distinguished: pastoralist (goat or cattle herding), agro-pastoralist (both livestock herding and cultivation, with a different balance of the two depending on the region) and riverine, i.e. agriculturalist (cultivation of various crops, including staple cereals such as sorghum and maize, in the proximity of major rivers).

The sampling universe of each survey is typically a livelihood zone or combination thereof, sometimes intersected by a regional boundary (e.g. regular surveys are conducted within the Hawd and Addun pastoralist zones straddling several regions, while in Gedo specific and simultaneous surveys are done for the region's pastoralist, agro-pastoralist and riverine communities). Given this data structure, for Somalia we defined "least common denominator" analysis strata consisting of region-specific livelihood types (Table 2). Accordingly, for each region we defined a pastoralist, agro-pastoralist and riverine stratum (some regions only featured one or two of these, while in others different livelihood zones of the same type were combined into one. We also created varyingly sized internally displaced persons (IDP) strata within each region and included an urban stratum if the source data for the baseline population estimates at the start of the emergency, namely the AfriPop project [11], included one or more urban polygons within the region (see below; with the exception of Mogadishu, these urban zones are not explicitly included in the FSNAU/FEWS NET livelihood map). For refugees, we considered each camp as an individual stratum (Table 2).

Table 2. Geographic analysis strata used for estimation purposes.

Region	Number of strata, by livelihood type						Total strata
	Pastoralist	Agro-pastoralist	Riverine	Urban	IDP	Refugee	
Bakool	1	1	0	0	1	0	3
Banadir (Mogadishu)	0	0	0	1	1	0	2
Bay	1	1	0	1	1	0	4
Galgadud	1	1	0	0	1	0	3
Gedo	1	1	1	0	1	0	4
Hiran	1	1	1	0	1	0	4
Lower Juba	1	1	1	1	1	0	5
Middle Juba	1	1	1	0	1	0	4
Mudug	1	1	0	0	1	0	3
Lower Shabelle	1	1	1	1	1	0	5
Middle Shabelle	1	1	1	1	1	0	5
						Subtotal:	**42**
Ethiopia camps	0	0	0	0	0	6†	6
Kenya camps	0	0	0	0	0	5‡	5
						Total:	**53**

†Bokolmanyo, Bur Amino, Dollo Ado Transit/Reception Centres, Hilaweyn, Kobe and Melkadida.
‡Dahagaley, Hagadera, Ifo, Ifo 2 and Kambioos.

2.2.2 Analysis period

Given that the emergency had a progressive onset, we allowed the period of analysis (i.e. over which mortality was estimated) to vary by stratum, according to when the local deterioration in food security began. We quantified food security primarily using terms of trade indicators (as a measure of food access), built from sentinel market data: these express the Kcal staple cereal equivalent of an average labourer's daily wage, i.e. food purchasing power. As an alternative, we also explored terms of trade indicators more relevant for pastoralists, namely using a local quality goat in lieu of daily wage (see Annex). Much of the mortality attributable to food insecurity probably occurs with a lag time of months (acute malnutrition develops over weeks of consistently insufficient nutritional intake, while malnutrition episodes themselves last about 3-4 months for moderate malnutrition and 1-2 months for severe malnutrition [12]). However, statistical models (see below and Annex) suggested a strong association between low terms of trade and mortality even with a lag of zero, possibly reflecting immediate deterioration of children who are nutritionally "on the brink". For this reason, we decided to start our estimation of excess mortality as soon as terms of trade started to decline within each stratum. After considering various statistical approaches, we decided to define this time point through simple graphical inspection of each stratum's time series, which followed a consistent pattern (Figure 4). Practically, analysis started for all strata between April and December 2010.

Defining an appropriate end for the analysis period was less straightforward. In practice, our study considered data up to July 2012, when an annual round of surveys was completed. Therefore, we arbitrarily censored our estimates at July 2012 inclusive. We could not find data that conclusively demonstrated improvements in health or anthropometric status by this time point (we analysed trends in admissions to supplementary feeding programmes up to November 2012, but these mainly reflected Mogadishu and appeared to be confounded by improved humanitarian access towards the end of the time series). Statistical models showed a weak but statistically significant association of mortality with low terms of trade up to a lag of six months (Annex); 22/42 strata had not yet returned to the pre-emergency terms of trade (defined as the average level over the three months prior to the start of the analysis period in each stratum) and sustained this for a further six months by July 2012 (of these, seven had not returned to the baseline at all, though all had shown considerable improvement). Overall, the median analysis period was 24 months (a range of 20-28 months).

After estimation, it became apparent that excess mortality was almost completely restricted to the period between October 2010 and April 2012, inclusive (Figure 14).T: Therefore, cumulative excess death toll estimates are presented for this somewhat shorter 19-month period.

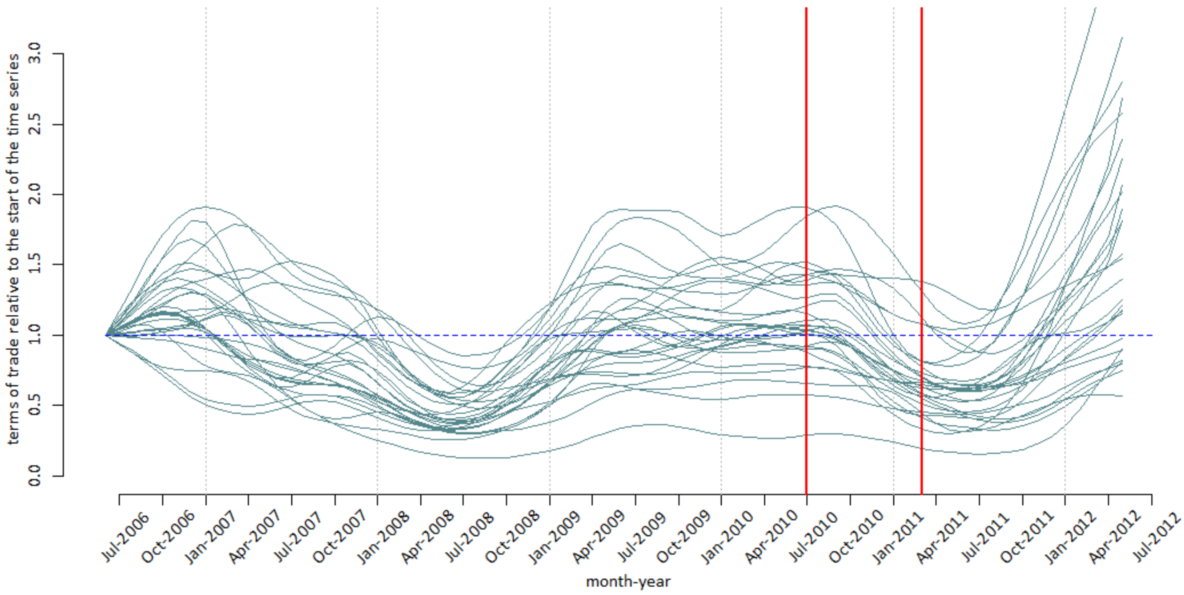

Figure 4. Trends in terms of trade relative to the start of the time series across all Somalia analysis strata. Each line represents a stratum. Thick vertical lines show the range in analysis start months.

For refugees in Kenya and Ethiopia camps, we also defined the analysis period as April 2010 to July 2012, i.e. the same time period as for the Somalia strata.

2.3 Total death rate estimation

2.3.1 Direct estimation (southern and central Somalia)

We defined the crude and under 5y death rates (CDR and U5DR) as the number of deaths due to all causes (among all ages for CDR; among children under 5 years for U5DR) occurring in a given population (total population for CDR; children under 5 years for U5DR) over a period of interest, divided by total person-time spent by the population during the same period. We expressed both as deaths per 10,000 person-days, the metric familiar to humanitarian agencies working in emergencies [13], though both the CDR and U5DR can readily be converted to deaths per thousand per year, which is a more common mode of expression for demographic rates outside of humanitarian settings. Crude death rates are by definition not age-standardised. Along with the prevalence of acute malnutrition, they are considered a key indicator of overall physical health status in emergencies [14].

The FAO/FSNAU and partners conduct livelihood-specific multi-indicator household surveys throughout Somalia. These are typically scheduled in May-July and November-January around the end the *Gu* (main) and *Deyr* (minor) rainy seasons, which are also periods of comparatively greater food security in pastoral areas and the peak lean season in crop-producing areas. While surveys are supposed to assess the same sampling universe during each round, in practice, logistical and security constraints as well as specific information needs generated many exceptions to this pattern. Surveys are carried out by experienced Somali field researchers supervised mostly remotely by FAO/FSNAU staff in Nairobi, where data are analysed. Survey methods have been described elsewhere [7]. Briefly, the mortality component of the questionnaire elicits information from household respondents on the composition of the household at the time of the survey and demographic events (births, deaths, arrivals and departures) within the household over a systematic retrospective recall period of 90 days. Simpler versions of this general questionnaire, consisting of aggregate rather than individual-based data collection and questions on household size and deaths only, were also administered during the analysis period with the stated purpose of simplifying data collection during the emergency (see Annex and discussion). Generally, no information on sex and age of household members (other than whether they were aged under five) was collected.

We systematically reviewed FAO/FSNAU publications for reports of surveys that (i) included a mortality component, (ii) had a specific livelihood zone or type as their sampling universe and (iii) were carried out between January 2007 and July 2012 inclusive. For each such survey, we extracted summary meta-data including survey characteristics, CDR and U5DR estimates and results for other indicators. Of 207 surveys meeting the above criteria, nearly all relied on a simple two-stage cluster sampling design with self-weighting allocation of clusters by probability proportional to size (PPS) [15,16]; 11 surveys (all among IDPs) were exhaustive (i.e. all households in the target population were visited). We verified the original FAO/FSNAU analysis for a sample of 82 surveys for which primary data were shared with the investigators (see Annex). Few serious discrepancies were noted, and none for surveys within the analysis period. We nonetheless adopted estimates from our verification for further analysis.

A total of 61 surveys covered the analysis period, but most stratum-months were not covered by data (Figure 5). Similarly, the proportion of the total person-time in the analysis that was covered by a survey, and therefore, for which mortality was estimated directly, was limited, though higher in some of the most populous strata, including pastoralists in Mudug, IDPs in Lower Shabelle and inhabitants of Banadir region, both urban and IDPs (Figure 6).

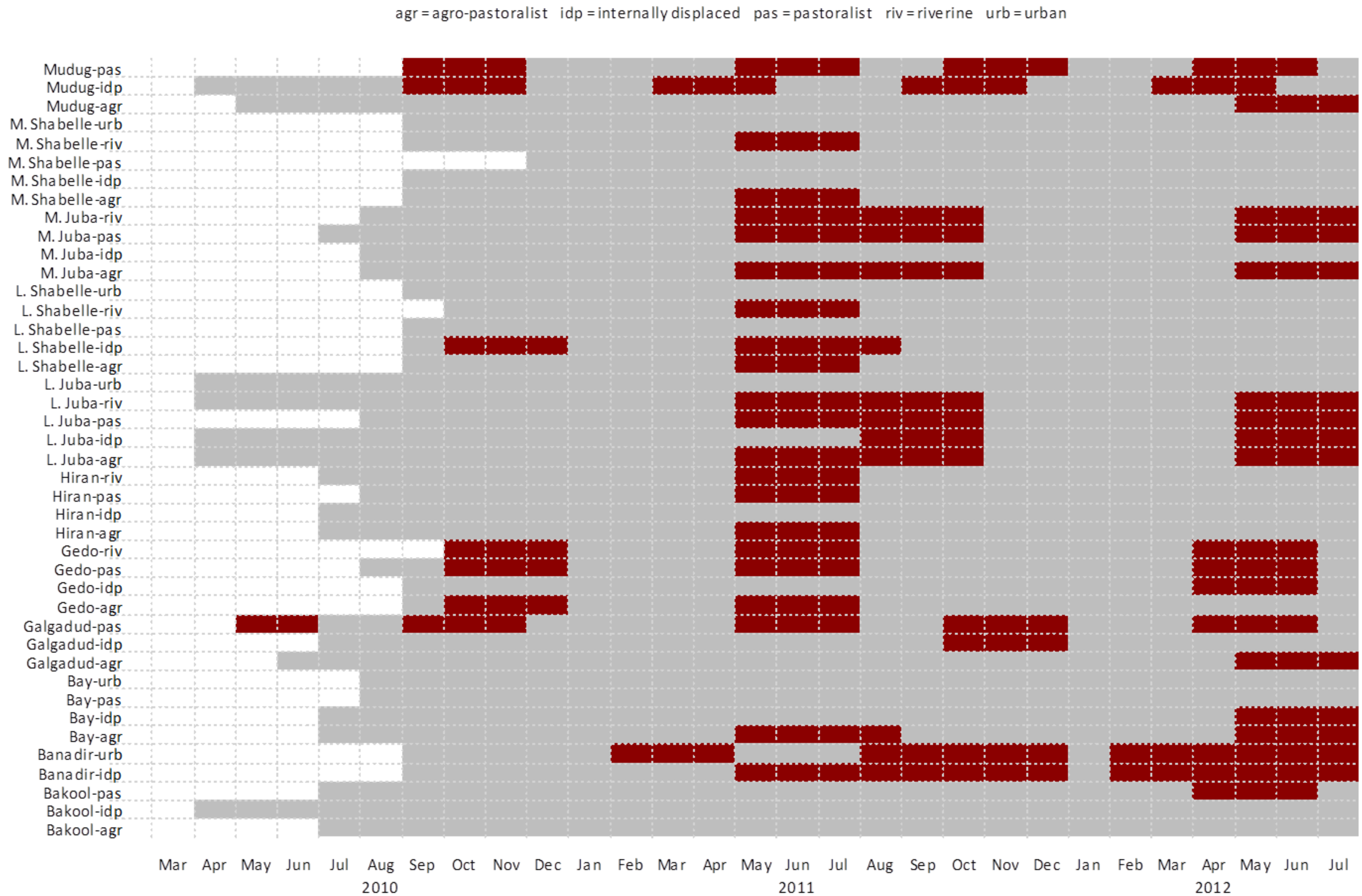

Figure 5. Months within each Somalia stratum's analysis period covered by the recall period of at least one mortality survey. Light grey cells indicate months within period but not covered by mortality data; red cells indicate months in period that are covered by data.

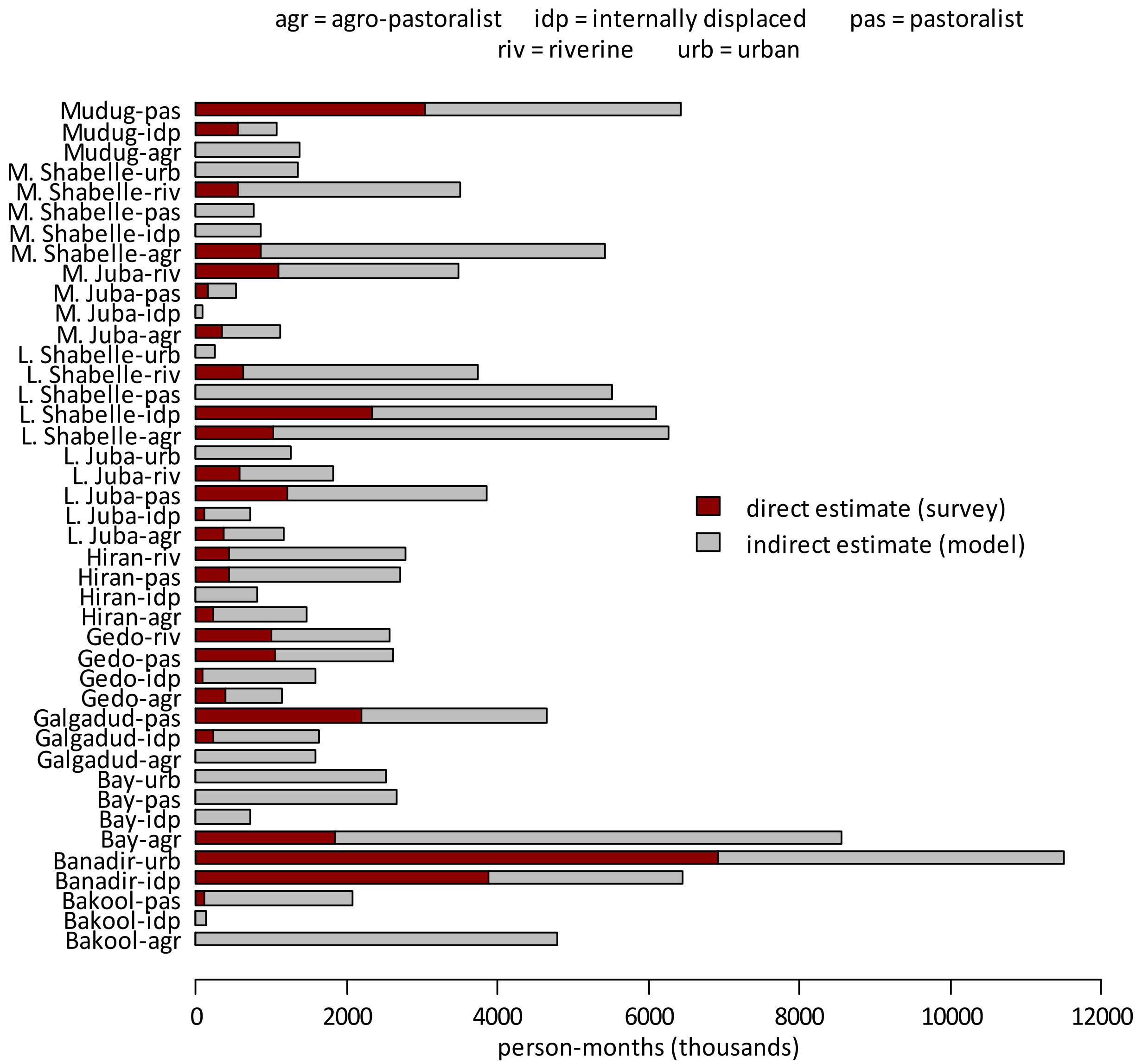

Figure 6. Amount of person-time in Somalia (October 2010 to April 2012 inclusive) for which mortality was computed based on direct (survey) or indirect (predictive model) estimation, by stratum.

2.3.2 Indirect estimation (southern and central Somalia)

To estimate death rates during months not covered by a survey, we fit predictive models with CDR or U5DR as the dependent variable respectively, and predictors selected from a list of variables that were plausibly associated with mortality and for which data were available continuously throughout the analysis period across all strata. These variables included (i) main survey location (southern and central Somalia versus Puntland/Somaliland, as a proxy of overall political stability and availability of government infrastructure and services); (ii) survey year, representing a putative secular trend; (iii) survey season, based on the typical Somali calendar cycle of minor lean and major lean periods (and periods in between, defined here as post minor and post major lean periods); (iv) incidence of armed conflict events occurring within the sampling universe during the recall period or with a specified lag (delay) between this incidence and the recall period, as extracted from the Armed Conflict Location and Event Dataset (ACLED), produced by an academic initiative to track conflict-related violence based on media and other reports (www.acleddata.com/about-acled/) [17]; (v) livelihood type (agro-pastoralist, pastoralist, riverine, IDP, urban); (vi) occurrence of epidemics during the recall period; (vii) availability of humanitarian assistance during the recall period or with a given lag (due to lack of comprehensive data, this variable was simplistically represented by whether the WFP was present, using the accurate data collected by WFP on its relief activities throughout Somalia); and (viii) average terms of trade during the recall period or with a lag (after univariate analysis, we settled on the Kcal cereal equivalent of an average daily wage as the single terms of trade indicator to use across all surveys). Note that the WFP's presence was used as a proxy for humanitarian assistance, in the absence of comprehensive data on the presence, output or funding level of relief agencies by stratum and period, the kind of data resolution that would have been necessary to inform the model. The association of mortality and WFP presence/absence should thus be interpreted as more broadly representing the effect of humanitarian relief agency interventions more broadly, not those of the WFP alone. Similarly, terms of trade were selected as a variable representing the broader concept of food security, due to the lack of suitable alternatives collected with sufficient consistency. We found in exploratory analyses that terms of trade correlated well with the prevalence of acute malnutrition (data not shown), indicating a plausible causal pathway to mortality, and thus their relative appropriateness as a proxy of food security and food access.

In order to maximise data available for model fitting and thus the model's precision, we included all livelihood-based surveys meeting eligibility criteria (see Section 2.3.1) done since January 2007 in all of Somalia (including Puntland and Somaliland). After excluding two surveys with suspected poor quality of data collection, the dataset consisted of 205 surveys.

For details of the model and source data, see the Annex. Briefly, we fit an ordinary least-squares linear regression, as this was the best distributional assumption available given that for most surveys we did not have access to source household-level datasets and could thus only analyse meta-data. Observations were weighted as a function of the surveys' quality and precision. The dataset had a longitudinal structure given that many surveys repeatedly sampled the same population over time. Thus, we tested various random effects with the surveys' sampling universe as the grouping factor, but ultimately excluded these as the model's fit worsened (surveys also drew independent samples and thus were only repeated at the group level). Since the model was used to predict mortality on external data, we used different predictive power statistics, including R^2, as the main criteria for including variables and interactions in the model. Furthermore, we estimated the model's predictive power on external data (i.e. external validity) using two cross-validation techniques. We used models to estimate CDR or U5DR in months without data within the Somalia strata.

2.3.3 Estimation in refugee camps

Data capture

We conducted a systematic search for primary reports of mortality in any of the camps listed in Table 2 published between January 2010 and September 2012. We consulted the Complex Emergency Database (CE-DAT) of surveys maintained by the Centre for Research on the Epidemiology of Disasters (www.cedat.be) and the Nutrition Information in Crisis Situations (NICS) database of the UN Standing Committee for Nutrition (www.unscn.org/en/publications/nics/database.php); carried out Google advanced searches using mortality or death rate and camp names as the search terms, screening .pdf, .doc and .docx files for relevant content; consulted the Medline medical database using the same search terms; extracted all available monthly camp reports from the UNHCR Health Information System (HIS) database (an integrated system to record, collect and analyse data on population, health and nutrition among refugees: see http://his.unhcr.org/his; monthly camp data include vital event surveillance); and contacted Epicentre to obtain unpublished assessment reports.

We retained for analysis all primary reports of mortality data that were based on either exhaustive data collection or representative sampling among refugees that had arrived to the camps during the analysis period (hereafter referred to as "new arrivals"). We specifically extracted death rate estimates for the period post arrival in the camp. Available mortality reports are listed in the Annex (Table 12, Table 13, Figure 17). Out of these, surveys done in Bokolmanyo and Melkadida camps, Ethiopia were excluded as the report [18] did not present disaggregated estimates for new arrivals; two Kenya reports were likewise not used for estimation: a survey in Dahagaley camp [19] featured probable methodological flaws (very low number of clusters and biased last-stage household selection); and a rapid survey of newly registered refugees [20], while methodologically sound, appeared to have a limited sampling universe. Data shared by Epicentre included unpublished results of a prospective demographic surveillance system implemented in Kobe and Hilaweyn camps between August 2011 and January 2012, and relying on weekly home visitor visits to newly arrived refugee households [21,22]. Other prospective surveillance data available in monthly HIS reports were excluded from analysis as they covered a minority of months in the period, did not disaggregate long-term camp residents from new arrivals, and were implausibly low for both CDR and U5DR (between one half and one fourth of concurrent survey estimates).

Ethiopia estimation

Available data were judged insufficient for statistical modelling or other formal indirect estimation methods. Thus, estimation for the refugee camps was conducted based on qualitative assumptions guided by available data and contextual information on health conditions in the camps.

For simplicity, for both Ethiopia and Kenya, we ignored mortality between departure and arrival to camps. Data on mortality during migration were not available for Ethiopia. Death rates during migration to Kenya appeared very elevated [20,23], possibly due to the arduous journey on foot, but the exposure (i.e. travel time to Dadaab camps) was short (mean 8 days, inter-quartile range 5-16 [23]; median 8 days [20]), resulting in a very small contribution to overall death tolls. We assumed that a similar pattern applied to Ethiopia.

Post-arrival death rate estimates were available only for Hilaweyn and Kobe camps, both established on an emergency basis to reduce overcrowding in the existing camps (Bokolmanyo and Melkadida) and host new arrivals. No other reliable data were found, and as such, we made very broad projections for remaining camp person-time on the basis of these data and a qualitative review of available UNHCR and other agency situation reports. We attempted to define optimistic and pessimistic scenarios that would on the whole probably encompass the plausible range of death rates (Table 3). The most influential assumptions are as follows:

- For Bokolmanyo and Melkadida, we assumed that the emergency phase during which camp services were under severe strain from the sudden influx of new arrivals lasted until October 2011, when a measles epidemic subsided and many of the new arrivals were moved to Kobe (as of July 2011), Hilaweyn (August 2011) and Bur Amino (January 2012). We assumed that mortality during this phase was in a range twice as broad as the 95%CI of a UNHCR survey done in Kobe and covering the first few months after its opening, when services were known to be insufficient. Remaining months in these camps saw a very limited influx of new arrivals, as most were diverted to one of the new camps. The assumptions of elevated mortality are also partly supported by March-April 2011 surveys done in these camps among older and new arrivals combined, showing CDR and U5DR of 0.78 and 2.26 per 10,000 person-days in Bokolmanyo, and 1.47 and 4.04 in Melkadida [18].

Table 3. Projected optimistic and pessimistic scenario death rates (per 10,000 person-days) among newly arrived Somali refugees in Dollo Ado, Ethiopia camps, April 2010-July 2012.

Camp, period	Optimistic scenario		Pessimistic scenario	
	Projected death rates	Assumptions	Projected death rates	Assumptions
Bokolmanyo, Melkadida				
Apr 2010 - Oct 2011	CDR 0.71 U5DR 2.20	Relatively high mortality due to overcrowding and pressure on camp services, but rates lower than in Kobe due to better established services and lower measles incidence; arbitrarily halved lower 95%CI of Kobe UNHCR survey [24].	CDR 5.10 U5DR 16.00	High mortality due to overcrowding and pressure on camp services, and rates higher than in Kobe due to overwhelmed services and more recently arrived refugees; arbitrarily doubled upper 95%CI of Kobe UNHCR survey.
Nov 2011 - Jul 2012	CDR 0.16 U5DR 0.51	Mortality decreased after camp decongestion and thanks to greater humanitarian relief, stabilising at lower 95% percentile of Kobe and Hilaweyn Epicentre/MSF weekly surveillance estimates [21,22] after measles epidemic (weeks 44-2011 to 7-2012 for Kobe; weeks 44-49 2011 for Hilaweyn); assumed that these data also applied to the rest of 2012.	CDR 0.78 U5DR 2.12	Mortality decreased after camp decongestion and thanks to greater humanitarian relief, stabilising at upper 95% percentile of Kobe and Hilaweyn Epicentre/MSF weekly surveillance estimates after measles epidemic (weeks 44-2011 to 7-2012 for Kobe; weeks 44-49 2011 for Hilaweyn); assumed that these data also applied to the rest of 2012.
Bur Amino				
Jan-Mar 2012	CDR 0.16 U5DR 0.51	Mortality immediately under control due to planned nature of camp; adopted lower 95% percentile of Kobe and Hilaweyn Epicentre/MSF weekly surveillance estimates (weeks 44-2011 to 7-2012 for Kobe; weeks 44-49 2011 for Hilaweyn).	CDR 1.56 U5DR 4.24	Mortality somewhat higher during first 3mo of camp opening as services were established; arbitrarily doubled upper 95% percentile of Kobe and Hilaweyn Epicentre/MSF weekly surveillance estimates (weeks 44-2011 to 7-2012 for Kobe; weeks 44-49 2011 for Hilaweyn).
Apr-Jul 2012	CDR 0.16 U5DR 0.51	As above	CDR 0.78 U5DR 2.12	Upper 95% percentile of Kobe and Hilaweyn Epicentre/MSF weekly surveillance estimates (weeks 44-2011 to 7-2012 for Kobe; weeks 44-49 2011 for Hilaweyn).
Dollo Ado Transit/Reception Centres				
Apr 2010 - Jul 2012	CDR 2.84 U5DR 8.80	Higher than in other camps due to dire lack of food, shelter, health care and other services, and recent arrival of refugees; arbitrarily doubled lower 95%CI of Kobe UNHCR survey. Assumed all refugees stayed one week in the reception centres.	CDR 5.10 U5DR 16.00	Higher than in other camps due to dire lack of food, shelter, health care and other services, and recent arrival of refugees; arbitrarily doubled upper 95%CI of Kobe UNHCR survey. Assumed all refugees stayed one month in the reception centres.
Hilaweyn				
Aug-Oct 2011	CDR 0.96 U5DR 3.15	Lower 95%CI of UNHCR Hilaweyn survey [24]; assumed that post-arrival mortality was equal to pre-arrival.	CDR 1.89 U5DR 6.57	Upper 95%CI of UNHCR Hilaweyn survey; assumed that post-arrival mortality was equal to pre-arrival.
Nov 2011 - Jul 2012	CDR 0.21 U5DR 0.64	Lower 95% percentile of Epicentre/MSF weekly surveillance estimates (weeks 44-49 2011).	CDR 0.78 U5DR 2.05	Upper 95% percentile of Epicentre/MSF weekly surveillance estimates (weeks 44-49 2011).
Kobe				
Jul 2011	CDR 2.84 U5DR 8.80	Lower 95%CI of UNHCR Kobe survey arbitrarily doubled to account for measles epidemic peaking during this month and recent establishment of camp.	CDR 5.10 U5DR 16.00	Upper 95%CI of UNHCR Kobe survey arbitrarily doubled to account for measles epidemic peaking during this month and recent establishment of camp.
Aug-Oct 2011	CDR 1.42 U5DR 4.40	Lower 95%CI of UNHCR Kobe survey.	CDR 2.55 U5DR 8.00	Upper 95%CI of UNHCR Kobe survey.
Nov 2011 - Jul 2012	CDR 0.16 U5DR 0.49	Lower 95% percentile of Epicentre/MSF weekly surveillance estimates (weeks 44-2011 to 7-2012); assumed that these data also applied to the rest of 2012.	CDR 0.67 U5DR 1.89	Upper 95% percentile of Epicentre/MSF weekly surveillance estimates (weeks 44-2011 to 7-2012); assumed that these data also applied to the rest of 2012.

- For Dollo Ado transit and reception centres, located closest to the border and where refugees anecdotally spent up to several weeks without basic services and conditions were worst [21,25], we assumed a consistently high level of mortality throughout the period, as no reports suggested an improvement in conditions over time in these sites.
- For all other camps after October 2011, we assumed that greater humanitarian relief and camp decongestion would have mitigated most excess mortality, and adopted CDR and U5DR levels in the range measured by the Epicentre/MSF surveillance systems in Kobe and Hilaweyn [21,22]. A key assumption here was that these systems maintained consistently high sensitivity, i.e. that the observed decline was real (with time, prospective surveillance is known to degrade and capture an increasingly smaller proportion of deaths [26]).

Kenya estimation

Available reports only covered the period from February to September 2011, and only new arrivals in Dahagaley were included in the sampling universe of surveys. The most informative of these was an exhaustive survey of new arrival settlements within Dahagaley camp [23], which also stratified mortality by time since arrival. We used this survey to compose pessimistic and optimistic scenarios, assuming that its estimated death rates were applicable to the entire analysis period, and, unlike for Ethiopia, projected mortality by time since arrival to the camp, irrespective of when arrival actually took place (Table 4). While overall conditions in the early part of the emergency (i.e. before September 2011) were plausibly worse due to overstretched relief services having to cope with a sudden acute influx, later months featured measles and cholera epidemics, and a sudden scaling back of agency presence in the camps due to insecurity and kidnappings. On balance, we assumed that health risks were similar in these two periods.

The main differences between the optimistic and pessimistic scenario were:

- Assumptions about mortality in Ifo 2 and Kambioos, two camps opened during the emergency to house new arrivals only: In the optimistic scenario, we assumed that people arriving or relocated to these camps would have enjoyed relatively good living conditions. (The camps were planned well in advance to satisfy minimal sectoral relief standards). In the pessimistic scenario, we assumed that the rapid relocation had exposed refugees to a comparatively high risk of death. This was corroborated by HIS mortality surveillance data from Kambioos during its first three months of existence, along with other reports of problems accessing food, water and health services in this camp [27].
- Assumptions about mortality among new arrivals in Hagadera and Ifo: while according to HIS reports these camps appeared comparable to Dahagaley in terms of living conditions as of 2010 (data not shown), we roughly assumed that mortality in these camps was either half (optimistic) or double (pessimistic) that in Dahagaley. We assumed that this possible range covered the plausible extent of any differences among the camps.
- Assumptions about mortality 12mo after arrival and beyond: The range chosen broadly reflects variability in CDR and U5DR observed in surveys done in Dadaab before 2010 as listed on the NICS database (CDR estimates: 0.20, 0.28, 0.33, 0.39 and 0.66 per 10,000 person-days; U5DR: 0.85, 0.90, 1.20, 1.30 and 1.46 per 10,000 child-days).

Table 4. Projected optimistic and pessimistic scenario death rates (per 10,000 person-days) among newly arrived Somali refugees in Dadaab, Kenya camps, April 2010-July 2012.

Period, camp	Optimistic scenario		Pessimistic scenario	
	Projected death rates	Assumptions	Projected death rates	Assumptions
<3mo after arrival				
Dahagaley	CDR 1.5 U5DR 2.5	From Polonsky et al. [23]; assumed to apply to all months in analysis period.	CDR 1.5 U5DR 2.5	From Polonsky et al.; assumed to apply to all months in analysis period.
Hagadera, Ifo	CDR 0.75 U5DR 1.25	Older Dadaab camps had comparable services as of 2010 according to HIS, but assumed that new arrivals in Hagadera and Ifo (not represented in available surveys) had better living conditions than in Dahagaley; arbitrarily halved Dahagaley death rates.	CDR 3.0 U5DR 5.0	Older Dadaab camps had comparable services as of 2010 according to HIS, but arbitrarily assumed that new arrivals in Hagadera and Ifo (not represented in available surveys) had worse living conditions than in Dahagaley; arbitrarily doubled Dahagaley death rates
Ifo 2, Kambioos	CDR 0.3 U5DR 0.7	Assumed that refugees arriving to these camps or relocated from other camps immediately benefited from services comparable to those settled in the older Dadaab camps for at least 6mo, i.e. applied rates for 6-11mo period (see below).	CDR 3.0 U5DR 5.0	Assumed that these camps did not have well-developed services, as suggested by UNHCR reports and high mortality in Kambioos during Sep-Dec 2011 according to camp surveillance; arbitrarily doubled Dahagaley death rates.
3-5mo after arrival				
Dahagaley	CDR 0.9 U5DR 2.2	As above	CDR 0.9 CDR 2.2	As above
Hagadera, Ifo	CDR 0.45 U5DR 1.1	As above: Dahagaley rates / 2	CDR 1.8 CDR 4.4	As above: Dahagaley rates x 2
Ifo 2, Kambioos	CDR 0.3 U5DR 0.7	As above	CDR 1.8 CDR 4.4	As above: Dahagaley rates x 2
6-11mo after arrival				
Dahagaley	CDR 0.6 U5DR 1.4	As above	CDR 0.6 U5DR 1.4	As above: Dahagaley rates x 2
Hagadera, Ifo	CDR 0.3 U5DR 0.7	As above: Dahagaley rates / 2	CDR 1.2 U5DR 2.8	As above: Dahagaley rates x 2
Ifo 2, Kambioos	CDR 0.3 U5DR 0.7	As above: Dahagaley rates / 2	CDR 1.2 U5DR 2.8	As above
≥12mo after arrival				
Dahagaley	CDR 0.14 U5DR 0.28	UNHCR, Dahagaley [28] (lowest of three "old" camp estimates); assumed to apply to all months in analysis period.	CDR 0.41 U5DR 0.94	UNHCR, Ifo (highest of three "old" camp estimates); assumed to apply to all months in analysis period.
Hagadera, Ifo	CDR 0.14 U5DR 0.28	As above	CDR 0.41 U5DR 0.94	As above
Ifo 2, Kambioos	CDR 0.14 U5DR 0.28	As above	CDR 0.41 U5DR 0.94	As above

2.4 Baseline death rate estimation

2.4.1 General considerations

Excess mortality may be defined as deaths above and beyond those that would have occurred in the absence of a given event. Past studies of mortality attributable to crises have used disparate approaches to estimate this counterfactual and thus immeasurable baseline, ranging from applying an arbitrary humanitarian standard threshold [29] to adopting pre-emergency regional death rates from a previous census [30] to estimating the within-sample relative risk of mortality compared to the period before the emergency, based on a retrospective study covering a sufficiently long pre-emergency period [31]. These different approaches imply varying understandings of what should be considered excess: for example, the baseline for refugees in a camp might be taken as their pre-displacement death rate, or alternatively as the typical death rate that should be aspired to in a camp if minimum humanitarian relief standards are met. Furthermore, all of these approaches tend to under-estimate excess deaths as they ignore likely secular improvements in mortality from the pre-emergency baseline that might have occurred in the absence of an emergency, i.e. they treat the baseline as static (given the relatively short period of our analysis, these improvements would probably be small).

In our case, defining the baseline was further complicated by several issues: (i) the emergency “event” could not be circumscribed to an isolated cause such as an earthquake, an epidemic or a heat wave, but rather was a combination of factors including abnormally severe food insecurity, armed conflict and the coverage and effectiveness of humanitarian assistance; (ii) Somalia had been in crisis for decades, with no census since the 1980s and unclear demographic trends since then; (iii) factors above were themselves varying in intensity during the analysis period, and were almost impossible to disentangle from each other; and (iv) epidemics occurring during the analysis period, while very plausibly attributable to the event, could conceivably have taken place even in its absence (i.e. deaths due to these epidemics, even if they had been quantified, would have been difficult to include in or subtract from excess mortality).

Given the above difficulties, we used several methods to estimate the baseline, reflecting different decisions about what should be considered excess mortality. These methods are outlined below.

2.4.2 Method 1

We defined the baseline strictly in terms of food security, irrespective of other factors. Accordingly, we populated an empirical distribution of baseline CDR and U5DR values using surveys done within Somalia since 2007 and for which the livelihood-specific terms of trade (see Annex) based on market data from within the survey’s sampling frame were equal to or above their baseline value at the start of the time series (July 2006) during the entire recall period of the survey, as well as the previous six months.

Method 1 assumes naively that all excess mortality is attributable to food insecurity, and considers relative rather than absolute levels of food security, with a different reference level for each stratum. It may either over- or under-estimate mortality depending on whether other factors responsible for excess mortality (e.g. armed conflict) are more or less influential during periods defined by Method 1 as baseline. Generally, it is likely to yield conservative estimates, as the baseline will for some strata consist of fairly low terms of trade. More specifically, this method ignores the much reduced presence of humanitarian assistance during 2010-2012.

2.4.3 Method 2

We defined the baseline as phases 1a, 1b or 2 out of the five defined by the Integrated Phase Classification (IPC) system for areas used in Somalia and other countries to grade the severity of food security and nutritional conditions – phase 5 being the most severe (famine). These phases are defined as:

- Phase 1 (Minimal): More than four in five households are able to meet essential food and non-food needs without engaging in atypical, unsustainable strategies to access food and income, including any reliance on humanitarian assistance.
- Phase 2 (Stressed): Even with any humanitarian assistance at least one in five households in the area have minimally adequate food consumption but are unable to afford some essential non food expenditures without engaging in irreversible coping strategies.

While different versions of the IPC were in use in recent years in Somalia, version 1.1 was used during the analysis period. In IPC version 2.0, 1a and 1b were condensed into a single phase 1. For more detail, see the IPC website (http://www.ipcinfo.org/). We composed empirical distributions as above, but using surveys conducted within Somalia populations and periods that were entirely classified into any of the above phases according to FAO/FSNAU technical series documents presenting projected populations by region, livelihood and IPC phase.

Method 2 isolates instances of acceptable food security, nutrition and health (defined with an absolute metric), and assumes that these would have been prevalent in the absence of the event. As such, it may yield a high-end estimate of excess mortality which, however, arguably reflects more closely the true gap between the situation in Somalia and what should be achievable with improved security and access to food. The IPC classifications are likewise more holistic, as they consider a variety of factors including not just food security but also nutritional status, feeding practices, epidemic occurrence, etc. However, our adoption of phases 1 and 2 as the baseline is arbitrary.

2.4.4 Method 3

We used predictive statistical models developed above to estimate a counter-factual baseline for each stratum-month. To do so, we generated model predictions of CDR and U5DR based on a hypothetical dataset that differed from the true data in that the following variables were artificially modified as follows: the occurrence of epidemics was set to zero, while the absolute terms of trade (Kcal cereal per living wage) was held constant throughout the analysis period at the average level over the three months before the start of the period in each stratum (as an alternative, we considered setting some sort of seasonal average as the baseline, but no seasonal trend was apparent and the association between terms of trade and season was non-significant).

Method 3 qualitatively defines the emergency as severe food insecurity only, but additionally assumes that epidemics would not have occurred without it. This assumption is based on empirical evidence from past crises that famines and nutritional emergencies greatly increase the risk and severity of epidemics, particularly for airborne droplet (measles, pertussis) and faecal-oral (cholera, shigellosis) transmission diseases [14,32]. Clearly it is strictly incorrect to assume that no epidemics would have occurred; however, for the purposes of statistical modelling, we had no way to decide which epidemics might have been part of the normal baseline. Method 3 treats armed conflict as a confounder, and likewise assumes that humanitarian assistance would have been constrained even without an emergency (indeed, WFP's and other agencies' withdrawal occurred well before the emergency). Because the same model is used to estimate total and baseline mortality, the effects of these factors are captured by this method.

As this definition seems plausible and conceptually consistent with the remainder of our analysis, we based our best estimates upon it.

2.5 Population dataset construction

2.5.1 Data sources

There has been no census of the Somali population over the past three decades. The population experiences a high birth rate (see below), but this has been counter-balanced by periodic increases in mortality due to various crises (including the 1991-1992 famine) and successive waves of internal and external migration. In 2005, the United Nations Development Programme (UNDP) carried out a household-based validation exercise that generated the first updated population dataset since armed conflict began in 1991. This population estimate (7.5 million) is used to date by the United Nations, though an exercise to update figures to the current time is ongoing.

The AfriPop project, a research initiative to develop geographically informative and up to date population estimates for various developing countries, has used a remote sensing approach to generate more current data for Somalia. Methods behind the AfriPop Somalia estimates are presented elsewhere [11]. Briefly, the project sought to distribute geographically the 2005 UNDP total figures for all of Somalia (including Puntland and Somaliland) according to satellite imagery-generated data on land use pattern and information from neighbouring Kenya on the relationship between different land use patterns and relative population density. The AfriPop estimates are centred in June 2010 and update the 2005 figures to this date by applying UN birth and death rate projections. The project also defines major urban populations as those falling within a set of urban land use polygons identified on satellite images.

The UNHCR tracks both internal displacement in Somalia and refugee movements in Kenya and Ethiopia (as well as some limited return or onward resettlement in third countries). Specifically, we used data provided by the UNHCR-Population Movement Tracking (PMT) initiative, some of which are made publicly available through its Horn of Africa Somali Displacement Crisis Information Sharing Portal (http://data.unhcr.org/horn-of-africa/country.php?id=197). Some more detailed data were shared in the form of Excel files, particularly monthly data on internal displacement by region and district, as well as arrival data in the Dollo Ado camps.

2.5.2 Baseline estimates for Somalia

The following steps were followed to compute baseline populations by analysis stratum for June 2010:

1. **Population figures for each stratum.** We took as our starting baseline (June 2010) population estimates produced by the AfriPop project (9,331,000 people), which represent an updated figure from a previous AfriPop estimate of 8.7 million for the 2010 Somalia population. We used ArcGIS software (ESRI, Redlands, CA, USA) to compute AfriPop estimates for each analysis stratum by overlaying shape files of the livelihood zone borders provided by the FAO/FSNAU into the AfriPop pixel map of Somalia's population.
2. **Adjustment for IDPs arriving into strata.** As of June 2010, 1,464,000 IDPs were reported by UNHCR to be within Somalia. However, this figure was considered likely to be an over-estimate as it resulted from cumulating new displacement waves without accounting for IDPs that may have returned to their region of origin. Reviewing monthly IDP movement data

from January 2008 to June 2010 (as well as estimates of displacement in 2007), we estimated from UNHCR-PMT data that 519,000 IDPs had been displaced during the period June 2009 through May 2010 (one year prior to the start of the analysis period) and that 946,000 IDPs had been displaced during the period June 2008 to May 2009 (the sum of these two years of cumulative displacement is 1,465,000, though this may be coincidental). Using FAO/FSNAU data on duration of displacement, we made a simplifying modelling assumption that half of IDPs were displaced for more than one year and the other half for less than one year. To calculate a baseline (June 2010) estimate of IDPs we took the total of the cumulative monthly displacements in June 2008-May 2009 (519,000) and added to that number half of the cumulative monthly displacements between June 2007 and May 2008 (half of approximately 1.5 million=750,000) for a baseline total of 1,269,000. We then allocated this revised total number to each region of arrival based on the reported geographic distribution for the 1,464,000 UNHCR figure.

3. **Adjustment for IDPs leaving strata.** In order to subtract the baseline IDP population of people from their regions and districts of origin, we used data from UNHCR-PMT on the place of origin of the 519,000 displaced from June 2009-May 2010 and made the assumption that this breakdown of origins also applied to IDPs displaced before this period. Figures for given regions and districts were then distributed proportionally into livelihood zone strata.
4. **Adjustment for refugees.** UNHCR estimated a total of 594,694 refugees in 7 countries (including 323,370 in Kenya and 72,054 in Ethiopia) as of June 2010. However, we did not subtract Somali refugees already in other countries from the 2010 Somalia population baseline. The assumption behind this decision was that the refugee movements had taken place over a number of years and that the refugees were thus already discounted from the population as of 2005.

2.5.3 Population growth during the analysis period

In order to estimate monthly population changes due to births and deaths, we applied the CDR and U5DR values estimated for each stratum-month based on methods outlined above, and the UNICEF (2009) estimate of the annual crude birth rate (44 per 1000 per year, calculated monthly), which was also close to the median crude birth rate observed in FAO/FSNAU surveys (43 per 1000 per year) for which we were able to do re-analysis. While death rates specific to each stratum-month were used, the crude birth rate was assumed to be the same for all stratum-months analysed. For the all-age population, a growth rate equal to the difference between CDR and crude birth rate was applied each month to adjust the stratum population.

The under 5y population at the beginning of the analysis period was assumed to be 25% of the population (namely the median value in FAO/FSNAU surveys re-analysed: see Annex). Each month, the under 5y population was adjusted by adding the number of births estimated from the crude birth rate and subtracting the number of deaths estimated from the U5DR. Additionally, a constant proportion (1/59) of the under 5y old population was assumed to age out into age 5y and above each month.

2.5.4 Internal displacement during the analysis period

In order to adjust monthly totals for new IDP displacement during the analysis period, we used data from UNHCR-PMT, which estimated displacement into and from regions and districts by month. New IDPs in a given month were thus added to the IDP stratum in the region where UNHCR-PMT recorded them arriving, and were correspondingly subtracted from the regional IDP stratum where UNHCR-PMT recorded them having departed from (these district and regional numbers were distributed proportionally by livelihood stratum, as above).

Based on data collected by FAO/FSNAU surveys, which showed that approximately half of IDPs were displaced for more than one year and about half were displaced for less than one year, we assumed that the IDPs displaced during the interval, June 2010 - August 2012, would remain in displacement for one year. Each monthly cohort of new IDPs thus remained in the IDP stratum for 11 subsequent months, after which they were "returned" to their stratum of origin. This formula was also applied to the 519,000 IDPs displaced from June 2009 to May 2010.

To balance the population of IDPs who were displaced in the model for one year and then "returned" to their stratum of origin, we left 750,000 of the baseline IDP population in displacement for the full duration of the interval.

2.5.5 Refugee movement during the analysis period

For refugees entering Ethiopia (Dollo Ado) or Kenya (Dadaab) during the interval, we used monthly camp populations obtained from UNHCR-PMT. We calculated monthly new arrivals by subtracting each month's camp population from that of the previous month, making the assumption that refugee returns to Somalia or resettlement to other countries or locations were negligible during the analysis period (we could not obtain data on these movements out of camps). In Kenya, this assumption is probably not robust, as some refugees were known to have migrated to urban centres in Kenya or returned to Somalia during the analysis period [33].

To identify regions of origin of refugees in Ethiopia, we used data from UNHCR-PMT showing districts of origin for 128,570 Somali refugees newly arrived to Dollo Ado camps during the analysis period. These district estimates were then redistributed to regional strata, as above. We did not, however, subtract refugee movements from IDP strata; although one study by Jureidini et al. [34] indicated that up to 49% of refugees were IDPs before they became refugees, these data were for Somali refugees fleeing to Turkey and we had no empirical basis for assessing the proportion of refugees in Ethiopia and Kenya who may have come from IDP camps as opposed to their region and district of origin. No data on region of origin of Dadaab refugees were shared by UNHCR. We instead relied on survey data by Polonsky et al. [23], which provided region of origin for new arrivals in Dahagaley camp.

Several of the camps were established during the analysis period in order to reduce crowding in existing camps, but data on which camp populations were moved first and when were likewise not accessible to us. For Kenya, where the death rate projections were stratified by time since arrival, we transitioned all new arrivals through successive strata (<3mo, 3-5mo, 6-11mo, ≥12mo since arrival); whenever an older camp (e.g. Ifo) experienced a sudden population drop with concurrent sudden large increase in a new camp (e.g. Ifo 2), we assumed that all departures from the former were relocated to the latter; that only new arrivals since 2010 were being moved; and that people were relocated in order of arrival, i.e. those who had arrived the earliest were relocated first.

2.5.6 Children under 5 years old

While we were able to estimate the population under 5y old for Somalia strata, this was not possible for the refugee camps due to very limited data available on age and gender distribution. Instead, during estimation for Ethiopia and Kenya camps (see below), we drew a random proportion of children under 5y old from the empirical distributions of under 5y proportions in available mortality reports for Ethiopia and Kenya camps respectively.

2.6 Excess death toll estimation

We used a computer simulation to generate best estimates and confidence intervals for the death toll (overall and for different strata and periods). We used this approach so as to account for (known) uncertainty in the distributions of parameters used to inform the estimate. These parameters arose from various statistical processes with different variance distributions (direct estimations from survey data, indirect predictions from linear regression, empirical baseline mortality distributions).

Detailed simulation steps are provided in the Annex. Briefly, for each stratum-month, and for either CDR or U5DR, we programmed the following sequence of simulation steps:

1. Draw a random value of the total death rate from either direct estimates (for stratum-months covered by survey data) or model predictions. For refugee camps, we used estimates for either the optimistic or pessimistic scenario.
2. Draw a random value of the baseline death rate using any of the above three methods.
3. Subtract the output of step 2 from the output of step 1 to obtain the excess death rate.
4. Multiply by the period being analysed and population denominator to obtain the excess death toll.

We repeated the above steps 10,000 times to generate an uncertainty distribution, the median and 95% percentiles of which provided point estimates and confidence intervals for the excess death toll over any stratum and period or combination thereof.

3 Results

3.1 Direct survey estimates

We firstly show key results from FAO/FSNAU-led mortality surveys done in all of Somalia since 2007 and used in our analysis, so as to illustrate broad trends over time and region. Mortality findings of the surveys are shown in Figure 7 (every point on the figure represents the point estimate of a survey). Prior to 2011, many surveys yielded a CDR and U5DR greater than 1 and 2 deaths per 10,000 person-days respectively (in Sub-Saharan Africa, these thresholds are typically considered as indicative of emergency conditions). However, nearly all surveys yielded estimates below 2 and 4 deaths per 10,000 person-days respectively, threshold levels considered indicative of an out of control emergency. Moreover, no increase is apparent during the 2008-2009 period of food security stress (see Figure 4). By contrast, a very atypical peak is observed in 2011 in southern and central Somalia. No such peak is evident for Somaliland and Puntland.

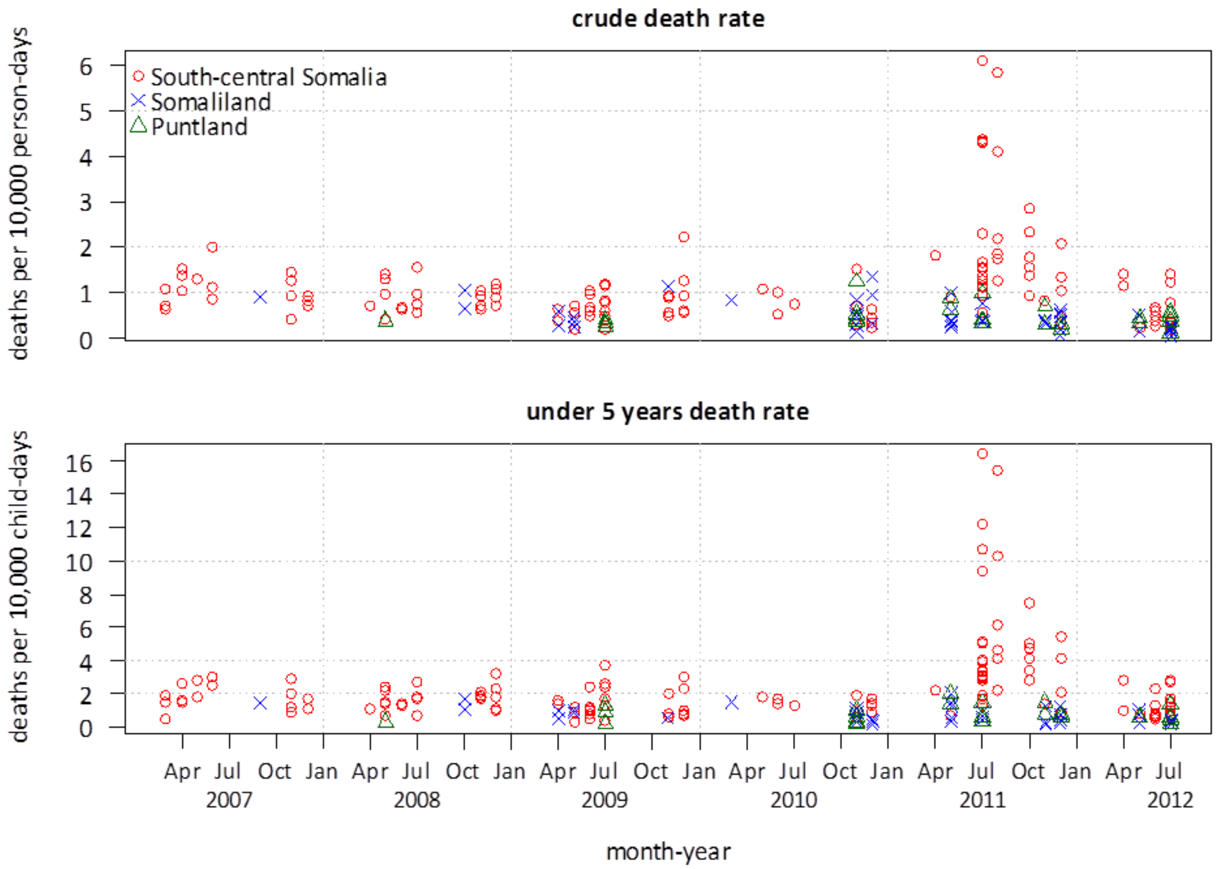

Figure 7. Crude and under 5 years death rate estimates from 205 surveys, by main survey location and time (mid-point of the recall period).

A similar pattern is observed for acute malnutrition prevalence data collected through surveys over the same period, irrespective of whether prevalence is computed using weight for height or middle-upper arm circumference (MUAC) as the anthropometric index of choice (Figure 8). Most Somalia estimates from 2011 are well in excess of any value observed in previous years. Trends possibly also indicate an unusual rise in the proportion of bilateral oedema (indicative of *kwashiorkor* malnutrition, a form rarely seen in the Horn of Africa during non-famine times) during 2011 (bottom panel), though this pattern is less clear.

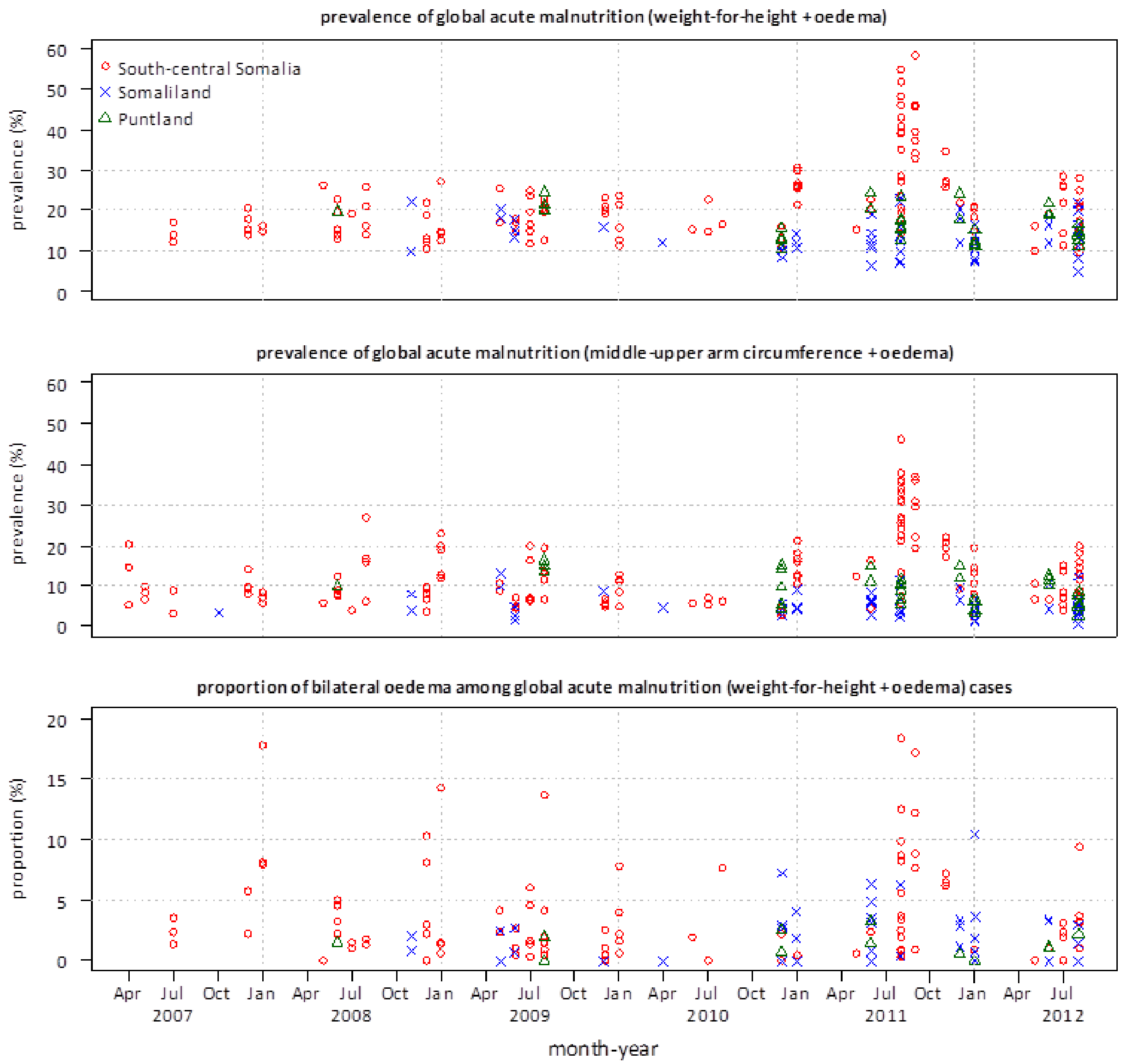

Figure 8. Prevalence of acute malnutrition from 237 surveys, using different anthropometric indicators based on WHO 2006 standards, by main survey location and time (survey month).

Table 5 summarises results of mortality surveys done in southern Somalia during the second half of 2011, which partly informed the decision to progressively declare famine conditions in different regions. Repeated surveys performed within the same stratum suggest a protracted period of very high mortality up to December at least. Particularly high death rates were estimated in Banadir and Lower Shabelle (by comparison, a typical CDR for countries in Sub-Saharan Africa not affected by crises is on the order of 0.2 to 0.6 per 10,000 person-days, i.e. some 10-20 times lower than the highest values from Somalia in 2011).

Table 5. Mortality results of surveys conducted in southern Somalia between July 2011 and December 2011.

Region / Livelihood	July 2011		August 2011		October 2011		December 2011	
	CDR (95%CI)	U5DR (95%CI)	CDR (95%CI)	U5DR (95%CI)	CDR (95%CI)	U5DR (95%CI)	CDR (95%CI)	U5DR (95%CI)
Banadir								
IDPs (Mogadishu)	4.37 (3.42-5.59)	12.21 (9.49-15.69)	5.83 (4.73-7.20)	15.43 (11.92-19.96)	1.79 (1.30-2.47)	5.00 (3.39-7.39)	2.08 (1.62-2.67)	5.47 (4.00-7.48)
urban (Mogadishu)					2.84 (2.12-3.82)	7.49 (5.30-10.61)	1.34 (0.89-2.02)	4.12 (2.66-6.39)
Bay								
agro-pastoralist	1.10 (0.81-1.50)	2.93 (1.93-4.44)	2.17 (1.64-2.88)	6.16 (4.42-8.57)				
Gedo								
agro-pastoralist	1.68 (1.25-2.27)	3.88 (2.85-5.29)						
pastoralist	1.25 (0.90-1.73)	3.37 (2.28-4.98)						
riverine	1.68 (1.04-2.70)	5.16 (3.05-8.72)						
Hiran								
agro-pastoralist	1.54 (1.21-1.95)	3.07 (2.22-4.26)						
pastoralist	1.54 (1.15-2.05)	2.85 (2.03-4.02)						
riverine	1.38 (1.09-1.75)	3.16 (2.39-4.18)						
Lower Shabelle								
agro-pastoralist	4.29 (3.16-5.82)	9.39 (7.17-12.31)						
IDPs (Afgooye)	4.33 (3.38-5.54)	10.72 (8.44-13.60)	4.10 (3.22-5.21)	10.30 (7.95-13.34)				
riverine	6.08 (4.64-7.99)	16.38 (11.99-22.38)						
Middle Shabelle								
agro-pastoralist	2.31 (1.81-2.94)	5.04 (3.82-6.65)						
riverine	1.52 (1.11-2.09)	3.99 (2.62-6.07)						
Middle and Lower Juba								
agro-pastoralist	1.14 (0.86-1.51)	3.01 (2.20-4.13)	1.84 (1.36-2.49)	4.62 (3.28-6.50)	1.38 (1.04-1.82)	3.43 (2.45-4.81)		
pastoralist	1.24 (0.90-1.73)	3.31 (2.35-4.68)	1.25 (0.84-1.86)	2.26 (1.33-3.85)	0.94 (0.65-1.35)	2.77 (1.82-4.23)		
riverine	1.15 (0.87-1.51)	3.43 (2.55-4.62)	1.75 (1.32-2.32)	4.16 (3.09-5.61)	1.55 (1.22-1.97)	4.14 (3.02-5.69)		
IDPs (Kismayo)					2.33 (1.82-2.97)	4.76 (3.63-6.25)		

3.2 Indirect estimation

Table 6 shows ordinary least-squares fixed effects predictive models used to indirectly estimate mortality for stratum-months not covered by surveys (see the Annex for more detail on model development and fitting). Here, negative coefficients imply that the factor in question, when present, is associated with reduced mortality (and vice versa); small p-values indicate strong (significant) associations. As this is a model primarily used for prediction (see below), however, the strength of individual associations is of secondary importance, and should mostly be considered so as to verify the model's plausibility against what would reasonably be expected to happen in reality.

Table 6. Least-squares models used to estimate mortality in person-time not covered by data.

Variable	Association with crude death rate (Box-Cox transformed, $\lambda = 0.4$)			Association with under 5 years death rate (Box-Cox transformed, $\lambda = 0.4$)		
Category	**Coefficient**	**SE**	**p-value**	**Coefficient**	**SE**	**p-value**
Intercept	-0.53	0.13	<0.001	-0.56	0.28	<0.001
Main location of surveyed population						
Somaliland or Puntland	ref.	-	-	ref.	-	-
Southern and central Somalia	0.30	0.08	<0.001	0.67	0.18	<0.001
Rate of armed conflict incidents (per 100,000 person-days; trimester mean; lag = 3 months)						
0	ref.	-	-	ref.	-	-
0.01-0.99	0.27	0.09	<0.001	0.55	0.21	0.010
1.00-1.99	0.18	0.11	0.109	0.30	0.26	0.252
≥ 2.00	0.47	0.11	<0.001	1.17	0.26	<0.001
Main livelihood						
Agro-pastoralists	ref.	-	-	ref.	-	-
Pastoralists	-0.08	0.07	0.226	-0.18	0.15	0.226
Riverine (agriculturalists)	0.13	0.07	0.067	0.34	0.16	0.034
Urban	0.13	0.17	0.448	-0.28	0.40	0.484
Internally displaced	0.17	0.08	0.044	0.11	0.17	0.501
Confirmed epidemics during recall period						
None	ref.	-	-	ref.	-	-
One	0.10	0.06	0.100	0.36	0.14	0.013
two or more	0.42	0.08	<0.001	0.86	0.18	<0.001
Presence of humanitarian assistance (World Food Programme as proxy) during most of a trimester period (lag = 2 months)						
Present	ref.	-	-	ref.	-	-
Absent	-0.35	0.18	0.051	0.08	0.36	0.831
Absolute terms of trade (Kcal cereal equivalent of daily wage; trimester mean; lag = 2 months)						
≥25,000	ref.	-	-	ref.	-	-
20,000-24,999	-0.26	0.09	0.004	-0.19	0.22	0.375
15,000-19,999	-0.33	0.10	<0.001	-0.23	0.23	0.307
10,000-14,999	-0.18	0.09	0.042	-0.20	0.21	0.345
0-9999	0.20	0.11	0.062	1.23	0.25	<0.001
Effect modification: presence of humanitarian assistance (World Food programme as proxy) **x absolute terms of trade**						
HA‡ absent and 20,000-24,999 Kcal/wage	-0.30	0.30	0.320	-0.31	0.53	0.560
HA absent and 15,000-19,999 Kcal/wage	0.30	0.23	0.195	-0.57	0.46	0.217
HA absent and 10,000-14,999 Kcal/wage	0.60	0.21	0.006	0.82	0.46	0.073
HA absent and 0-9999 Kcal/wage	0.93	0.22	<0.001	0.97	0.45	0.033

Model validity	Internal (fit to training data)	External (prediction of new data)	Internal (fit to training data)	External (prediction of new data)
Mean squared residuals (untransformed)	0.128	0.156 (LOOCV†) 0.144 (bootstrap)	0.580	0.715 (LOOCV) 0.650 (bootstrap)
Adjusted R^2	71%	65% (LOOCV)	78%	73% (LOOCV)
Proportion of predictions that are accurate within ± 0.5 deaths per 10,000 person-days	81%	77% (LOOCV)	47%	43% (LOOCV)
Proportion of predictions that are accurate within ± 1.0 deaths per 10,000 person-days	97%	95% (LOOCV)	77%	72% (LOOCV)

†Based on leave-one-out cross-validation technique. ‡Humanitarian Assistance.

Models for CDR and U5DR were specified identically and suggest a common pattern. In southern and central Somalia (as opposed to Somaliland and Puntland), an increasing rate of armed conflict incidents over the three months prior to the survey's recall period, and the occurrence of epidemics during the recall period all show highly significant positive associations with mortality; the association with riverine livelihood is weakly significant. We also found a strong interaction between the absence of humanitarian assistance and increasingly low terms of trade, resulting in exponential associations with mortality in an untransformed metric. Accordingly, the combined effect of low terms of trade and lack of humanitarian assistance is greater than the sum of the two effects taken in isolation.

Both models have reasonably good predictive power (adjusted R^2, indicating the proportion of variance in mortality explained by the model), and are expected to retain this accuracy when applied to predict new data. Specifically, model CDR estimates for stratum-months not covered by a survey are expected to be accurate within ± 0.5 deaths per 10,000 person-days about 77% of the time.

3.3 Baseline mortality

The criteria specified under method 1 to define periods of baseline mortality (i.e. based on when terms of trade were equal to or higher than at the start of 2007) identified 35 surveys providing baseline CDR and U5DR estimates, of which were most done among pastoralists and agro-pastoralists (Figure 9). As shown, these survey estimates featured considerable variability, particularly for U5DR (e.g. from 0.5 to 3.9 per 10,000 child-days for agro-pastoralists), and included values normally considered indicative of emergency conditions (> 1 per 10,000 person-days for CDR and > 2 for U5DR).

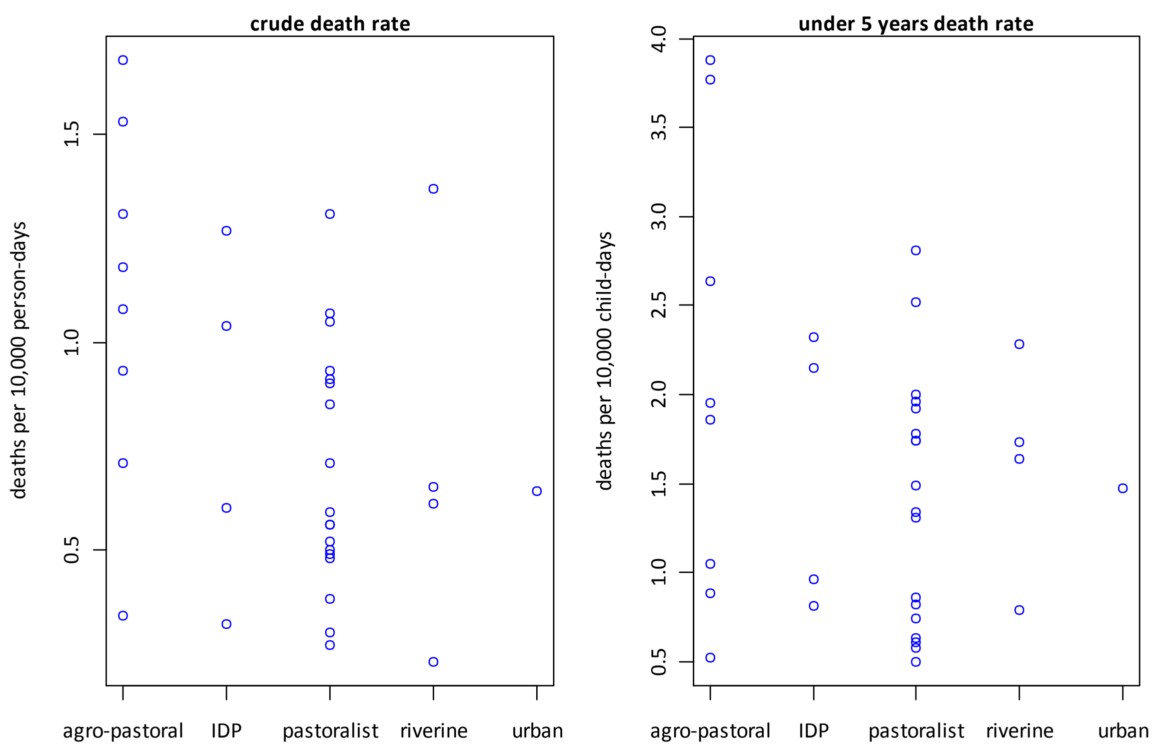

Figure 9. Crude and under 5 years death rate estimates from surveys used for baseline method 1 (terms of trade), by livelihood type. Each point represents the point estimate of a survey.

Using the IPC phases 1 and 2 to define baseline periods (method 2), only 12 surveys were available (Figure 10).

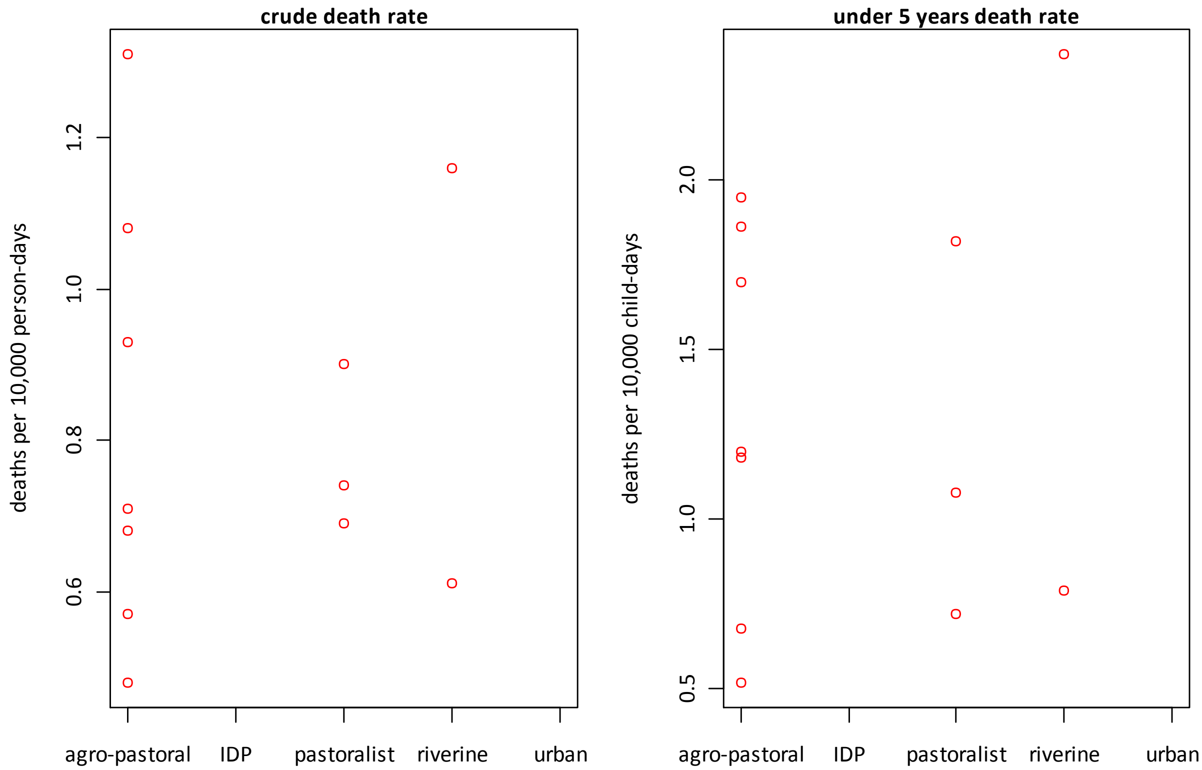

Figure 10. Crude and under 5 years death rate estimates from surveys used for baseline method 2 (IPC phases), by livelihood type.

Applying baseline method 3, the counterfactual baseline CDR for all of southern and central Somalia combined was estimated to oscillate between 0.5 and 0.8 deaths per 10,000 person-days throughout the analysis period, about 0.2 to 0.5 deaths per 10,000 person-days higher than the average CDR for Sub-Saharan Africa in 2010 (Figure 11). By contrast, CDR was predicted to be elevated compared to the baseline from October 2010 to April 2012, peaking around 2.5 per 10,000 person-days around June-July 2011.

As shown in Figure 12, estimated CDR and U5DR across southern and central Somalia followed similar patterns, although changes in U5DR were more pronounced. Most of the period before May 2011 was not covered by surveys, and as such was subject mainly to indirect estimation: the models for CDR and U5DR had the same predictors, though different effect sizes, and as such would not be expected to be able to identify an earlier rise in U5DR, compared to CDR.

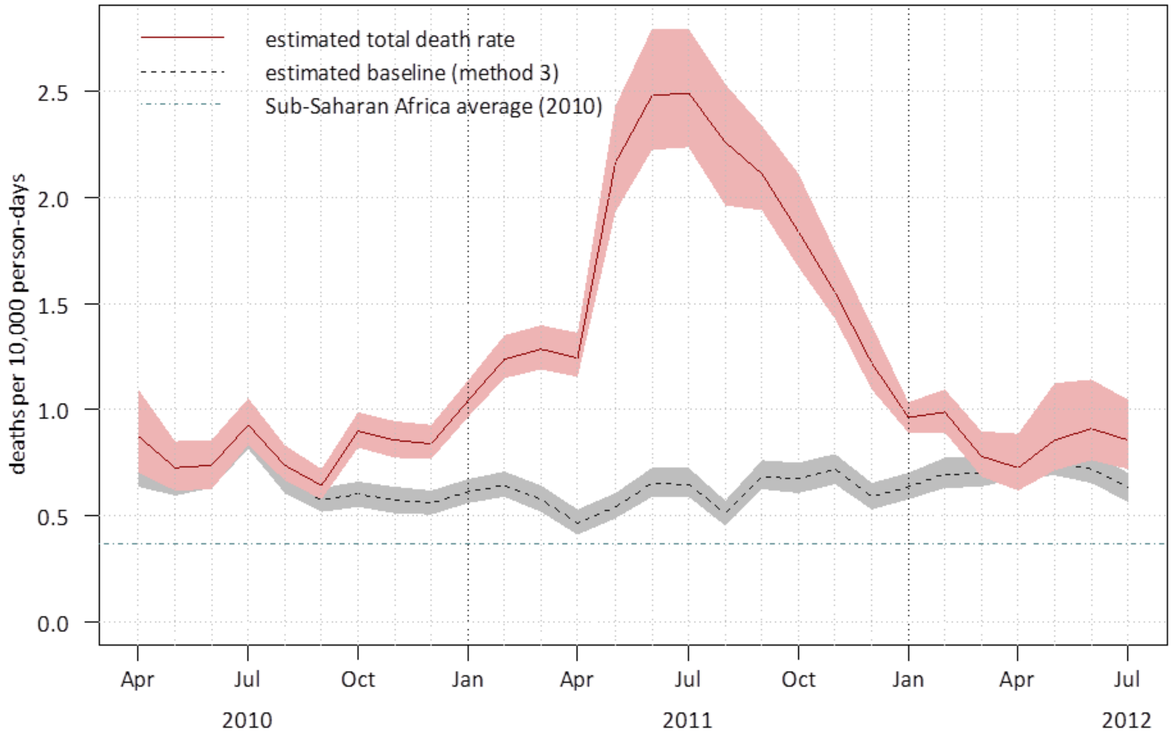

Figure 11. Estimated monthly total crude death rate in southern and central Somalia, compared to counter-factual baseline (method 3) and Sub-Saharan Africa average. Shaded areas indicate 95% percentile intervals around the point estimate.

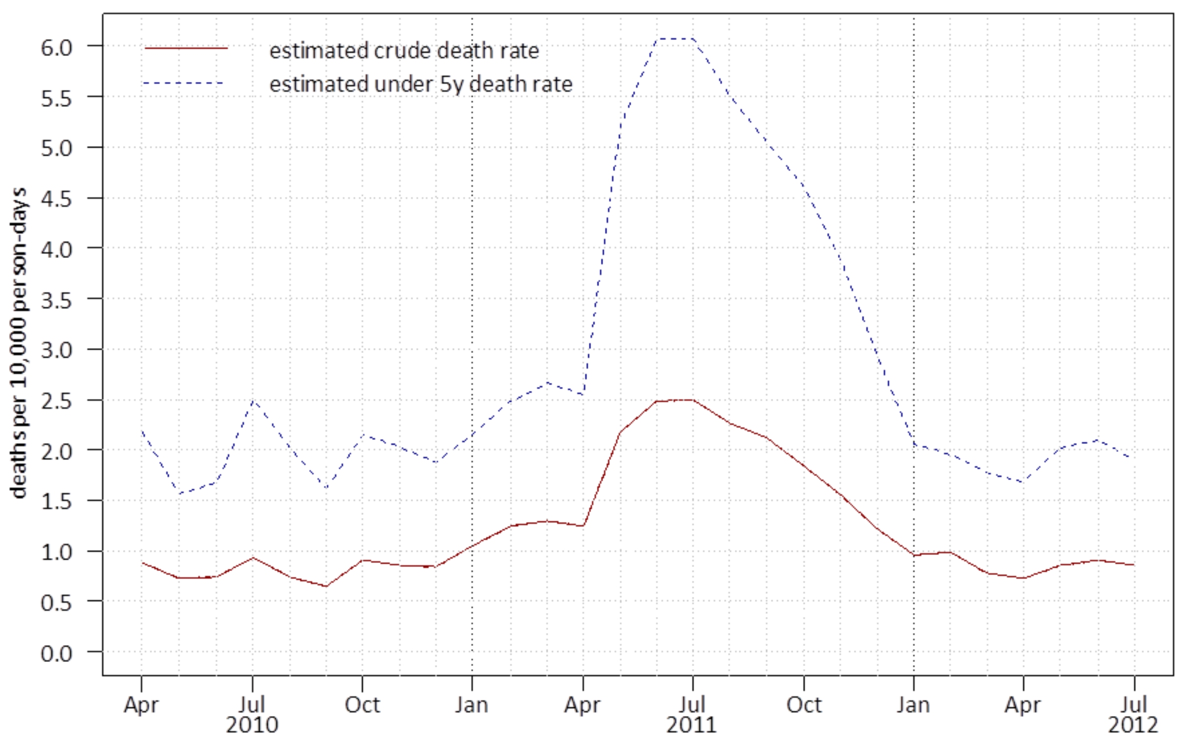

Figure 12. Comparison of trends in estimated crude and under 5 years death rates across southern and central Somalia.

3.4 Excess mortality estimates

3.4.1 Somalia

Table 7 shows estimated excess death tolls for each stratum, according to the baseline computation method used. The three baseline methods broadly agree, although method 3, considered the most plausible, mostly yields higher death toll estimates, particularly for agro-pastoralists in Bay and Middle Shabelle, where the baseline death rate according to method 3 was far lower than according to the other methods, which did not compute a stratum-specific baseline. The proportion of children under 5y old among estimated deaths was 54.8% under method 1, 65.6% under method 2 and 51.5% under method 3.

For some strata, the estimated death toll is negative, indicating fewer people are estimated to have died than would have, given the baseline CDR or U5DR estimate being applied. As expected, central Somalia features lower mortality than the southern strata. Overall, the largest excess death tolls are projected for Bakool, Banadir, Bay, Lower Shabelle and Middle Shabelle regions, where all baseline methods suggest excess death tolls greater than 10,000 during the analysis period. These geographic patterns are shown in Figure 13, where each stratum is represented.

The total death toll considering all three baseline methods and their 95% percentile intervals ranges between a minimum of 143,100 and a maximum of 272,700 considering the period October 2010 to April 2012 inclusive. Baseline method 3, considered the most plausible, yielded a point estimate of 257,900. While this figure represents the excess death toll, overall 595,300 (95% percentile interval 580,400 to 610,800) total deaths (i.e. including the baseline and excess) were estimated to have occurred in southern and central Somalia during the period. Of these, according to baseline method 3 some 290,700 (285,100 to 297,300) would have occurred irrespective of the emergency.

Of the total excess deaths estimated using method 3, 39.8% (102,700/257,900) were computed "directly", i.e. for stratum-months that were covered by one or more mortality survey, while the remainder were computed indirectly by applying the predictive statistical model. The corresponding percentages for baseline methods 1 and 2 were 51.6% (84,500/163,600) and 46.6% (91,000/195,500) respectively.

Table 7. Estimated excess death toll in southern and central Somalia (all ages and under 5y old) between October 2010 and April 2012 inclusive, by analysis stratum and method used to define the baseline. Numbers are, respectively, medians and 95% percentile intervals of simulation output distributions.

Region, livelihood	baseline 1 (surveys done in periods of high food security)		baseline 2 (surveys done in IPC phases 1 and 2)		baseline 3 (model predictions assuming no food security deterioration and no epidemics)	
	all ages	under 5y	all ages	under 5y	all ages	under 5y
Bakool						
agro-pastoralist	22,000 (18,300 to 25,500)	11,700 (9700 to 13,700)	25,400 (22,500 to 28,100)	14,400 (13,000 to 15,700)	9000 (6700 to 11,400)	4600 (3400 to 5800)
IDP	700 (600 to 800)	400 (300 to 400)	700 (600 to 800)	400 (300 to 400)	600 (500 to 700)	300 (300 to 400)
pastoralist	800 (-400 to 2000)	700 (0 to 1400)	600 (-400 to 1600)	1000 (400 to 1600)	1100 (500 to 1700)	700 (400 to 1100)
Banadir						
IDP	28,500 (20,200 to 35,700)	16,400 (12,000 to 20,900)	29,800 (23,300 to 35,500)	18,100 (14,100 to 22,700)	27,200 (20,800 to 33,300)	16,100 (12,200 to 20,100)
urban	31,100 (21,200 to 41,700)	8700 (4500 to 13,100)	34,000 (24,800 to 43,300)	11,000 (7600 to 15,100)	30,800 (22,800 to 39,800)	11,000 (7900 to 14,800)
Bay						
agro-pastoralist	5300 (-1900 to 12,300)	4700 (500 to 9200)	11,500 (6700 to 16,400)	10,200 (6900 to 13,700)	13,400 (11,500 to 15,300)	13,900 (11,400 to 16,900)
IDP	1800 (1300 to 2400)	900 (600 to 1200)	1900 (1500 to 2300)	1100 (900 to 1400)	3000 (2800 to 3300)	1400 (1300 to 1600)
pastoralist	600 (-1000 to 2200)	700 (-200 to 1600)	600 (-700 to 1800)	1000 (300 to 1800)	1400 (1000 to 1800)	2600 (2100 to 3000)
urban	3200 (1200 to 5000)	800 (-400 to 1700)	3600 (2100 to 5200)	1500 (700 to 2400)	3000 (2200 to 3800)	2900 (2300 to 3600)
Galgadud						
agro-pastoralist	-200 (-1300 to 900)	200 (-400 to 900)	1000 (200 to 1700)	1200 (800 to 1700)	2100 (1700 to 2500)	1800 (1500 to 2000)
IDP	900 (-300 to 2200)	100 (-600 to 1100)	1200 (200 to 2300)	600 (0 to 1500)	700 (100 to 1800)	700 (300 to 1400)
pastoralist	-600 (-3700 to 3300)	-100 (-2300 to 4200)	-1200 (-4100 to 2700)	600 (-1300 to 4900)	1300 (-700 to 4700)	1700 (300 to 6200)
Gedo						
agro-pastoralist	-400 (-1200 to 500)	-100 (-500 to 400)	500 (-200 to 1100)	700 (300 to 1000)	2400 (2000 to 2900)	900 (600 to 1100)
IDP	1500 (300 to 2800)	800 (100 to 1400)	1800 (1000 to 2700)	1300 (800 to 1800)	4300 (3600 to 4900)	1300 (1000 to 1700)
pastoralist	300 (-1400 to 1900)	600 (-500 to 1500)	-400 (-1700 to 1100)	900 (100 to 1700)	4500 (3600 to 5400)	1700 (1100 to 2400)
riverine	2100 (-100 to 4300)	2000 (800 to 3400)	2400 (800 to 4300)	2500 (1600 to 3800)	6600 (5200 to 8200)	2500 (1600 to 3700)
Hiran						
agro-pastoralist	-700 (-1800 to 300)	-200 (-900 to 500)	300 (-400 to 1200)	700 (300 to 1100)	1400 (800 to 2000)	500 (200 to 900)
IDP	700 (0 to 1200)	400 (0 to 700)	800 (400 to 1300)	600 (300 to 800)	1000 (600 to 1400)	400 (200 to 600)
pastoralist	400 (-1300 to 2000)	500 (-500 to 1400)	300 (-1000 to 1800)	900 (0 to 1700)	2600 (1600 to 3700)	900 (400 to 1500)
riverine	1200 (-900 to 3100)	1500 (500 to 2500)	1500 (0 to 3100)	2100 (1300 to 2900)	2400 (1100 to 3600)	1000 (300 to 1600)
Lower Juba						
agro-pastoralist	-700 (-1600 to 200)	-200 (-800 to 400)	300 (-500 to 1100)	600 (100 to 1100)	1200 (600 to 1900)	500 (200 to 1000)

Region, livelihood	baseline 1 (surveys done in periods of high food security)		baseline 2 (surveys done in IPC phases 1 and 2)		baseline 3 (model predictions assuming no food security deterioration and no epidemics)	
	all ages	under 5y	all ages	under 5y	all ages	under 5y
IDP	600 (0 to 1100)	200 (-100 to 500)	700 (300 to 1100)	400 (200 to 600)	900 (600 to 1200)	300 (100 to 500)
pastoralist	-800 (-3300 to 1900)	400 (-1200 to 1900)	-800 (-3000 to 1500)	900 (-300 to 2300)	3100 (1600 to 4800)	1400 (500 to 2500)
riverine	300 (-1200 to 1900)	700 (-100 to 1800)	500 (-600 to 1900)	1200 (500 to 2100)	1800 (800 to 3000)	900 (300 to 1700)
urban	0 (-1000 to 800)	-400 (-1000 to 0)	200 (-500 to 800)	-100 (-400 to 300)	800 (300 to 1400)	200 (-100 to 500)
Lower Shabelle						
agro-pastoralist	11,600 (6300 to 18,200)	6300 (3100 to 9500)	16,300 (11,200 to 22,000)	10,100 (7800 to 12,600)	25,100 (21,000 to 30,300)	11,200 (9100 to 13,500)
IDP	18,700 (12,800 to 23,700)	10,600 (7500 to 13,500)	19,600 (15,000 to 24,400)	12,200 (9800 to 15,300)	29,200 (25,400 to 33,700)	12,800 (10,700 to 15,300)
pastoralist	9100 (5500 to 12,400)	5800 (3900 to 7700)	8400 (5600 to 10,900)	6400 (5000 to 7900)	18,200 (16,400 to 20,000)	8100 (7000 to 9100)
riverine	16,300 (12,400 to 20,900)	10,400 (8000 to 13,100)	17,000 (13,600 to 20,900)	11,200 (9100 to 13,900)	22,500 (19,200 to 26,500)	10,300 (8300 to 12,800)
urban	600 (400 to 800)	200 (100 to 300)	600 (500 to 800)	300 (200 to 400)	1000 (800 to 1100)	400 (300 to 400)
Middle Juba						
agro-pastoralist	200 (-700 to 1200)	200 (-300 to 800)	1100 (500 to 1900)	1000 (500 to 1400)	2500 (1900 to 3100)	1200 (900 to 1700)
IDP	100 (0 to 200)	100 (0 to 100)	100 (100 to 200)	100 (100 to 100)	200 (200 to 300)	100 (100 to 100)
pastoralist	500 (-100 to 1500)	400 (0 to 1200)	500 (-100 to 1500)	500 (100 to 1300)	1200 (800 to 2400)	600 (300 to 1500)
riverine	3900 (1100 to 6600)	3000 (1700 to 4500)	4400 (2400 to 6500)	3800 (2600 to 5100)	8000 (6400 to 9800)	3800 (2900 to 5000)
Middle Shabelle						
agro-pastoralist	900 (-2900 to 5000)	100 (-2300 to 2400)	5600 (2800 to 8200)	3400 (1800 to 5100)	11,600 (9700 to 13,700)	6600 (5700 to 7700)
IDP	900 (300 to 1500)	300 (-100 to 600)	1100 (600 to 1500)	500 (300 to 800)	1700 (1400 to 2100)	1000 (800 to 1100)
pastoralist	2800 (2400 to 3300)	1500 (1300 to 1800)	2800 (2400 to 3200)	1600 (1400 to 1800)	800 (500 to 1200)	400 (200 to 600)
riverine	2100 (-500 to 4900)	1600 (100 to 2900)	2500 (700 to 4200)	2300 (1300 to 3400)	5400 (4100 to 6800)	3800 (3000 to 4700)
urban	1200 (200 to 2200)	-100 (-600 to 400)	1500 (600 to 2300)	400 (0 to 800)	2600 (2000 to 3300)	1000 (800 to 1300)
Mudug						
agro-pastoralist	-1400 (-2500 to -500)	-1100 (-1700 to -600)	-400 (-1100 to 100)	-200 (-600 to 100)	1000 (700 to 1300)	200 (100 to 400)
IDP	-200 (-1000 to 600)	-300 (-800 to 100)	100 (-500 to 600)	0 (-400 to 300)	-400 (-700 to 100)	-200 (-500 to 100)
pastoralist	-2400 (-6800 to 2200)	-900 (-3600 to 2800)	-3200 (-6800 to 700)	-200 (-2200 to 3200)	-400 (-2700 to 2800)	100 (-1400 to 3600)
Totals	**163,600 (143,100 to 185,600)**	**89,700 (78,600 to 101,400)**	**195,500 (178,100 to 213,600)**	**128,200 (118,600 to 139,000)**	**257,900 (243,600 to 272,700)**	**132,900 (124,700 to 142,300)**

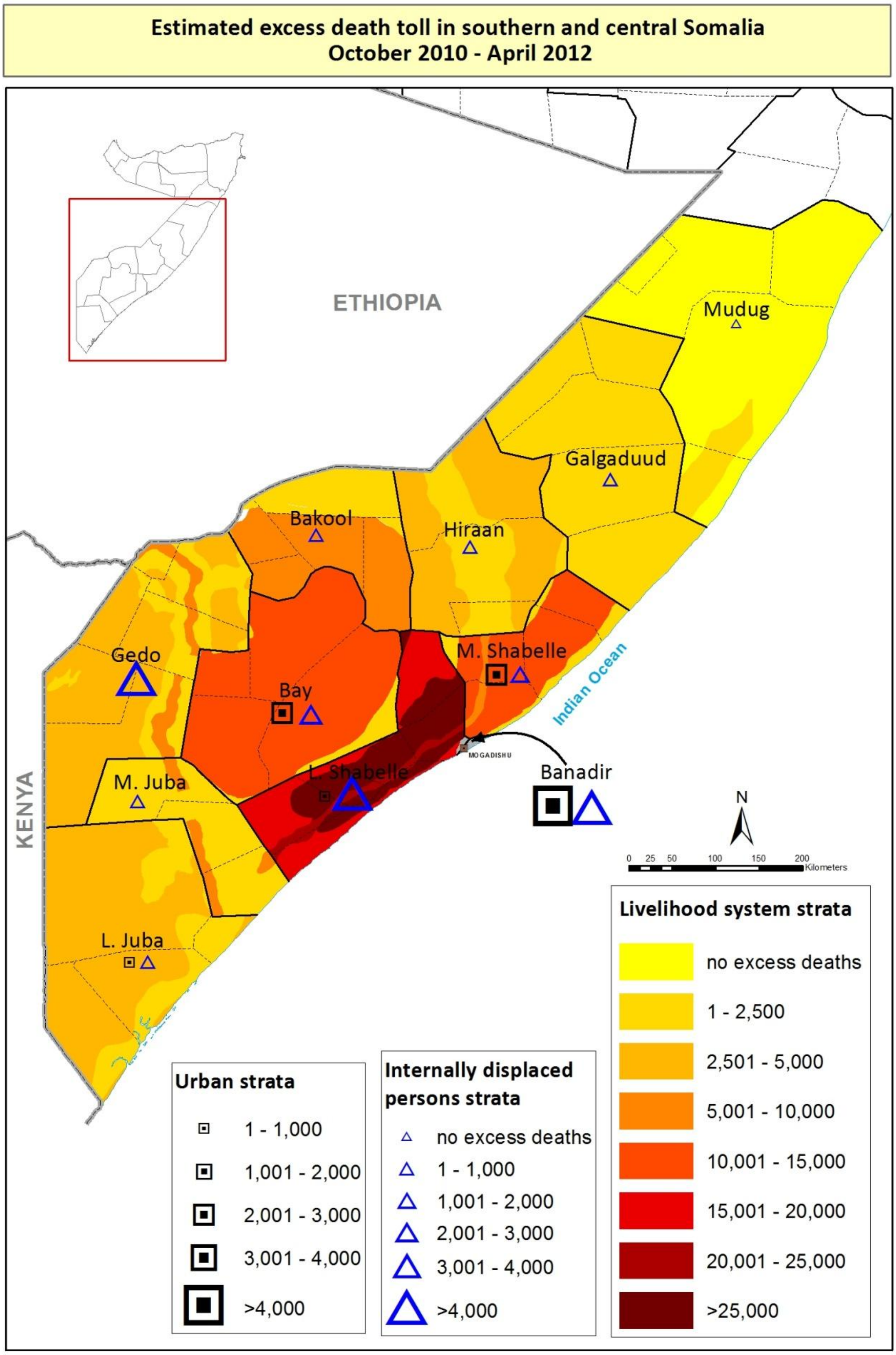

Figure 13. Geographic distribution of estimated excess death toll (according to baseline method 3) in southern and central Somalia, October 2010 to April 2012. Regional livelihood strata are represented in colour, corresponding to the number of estimated deaths occurring within them. Regional IDP and urban strata are shown as icons (blue triangle and black square, respectively; the placing of these icons on the map does not indicate the actual location of IDP settlements or towns).

As shown in Figure 14, the rise in monthly excess mortality was estimated to begin on a country-wide level in October 2010 under baseline method 3, and December 2010 under methods 1 and 2; by April 2012 mortality had returned to baseline levels, though baseline method 3 predicted some continued elevation up to July 2012. During May-September 2011, greater than 20,000 people were estimated to die per month throughout southern and central Somalia as a result of the emergency. A similar pattern was observed for IDP and urban residents in Lower Shabelle (Afgooye corridor) and Banadir (Mogadishu; Figure 15).

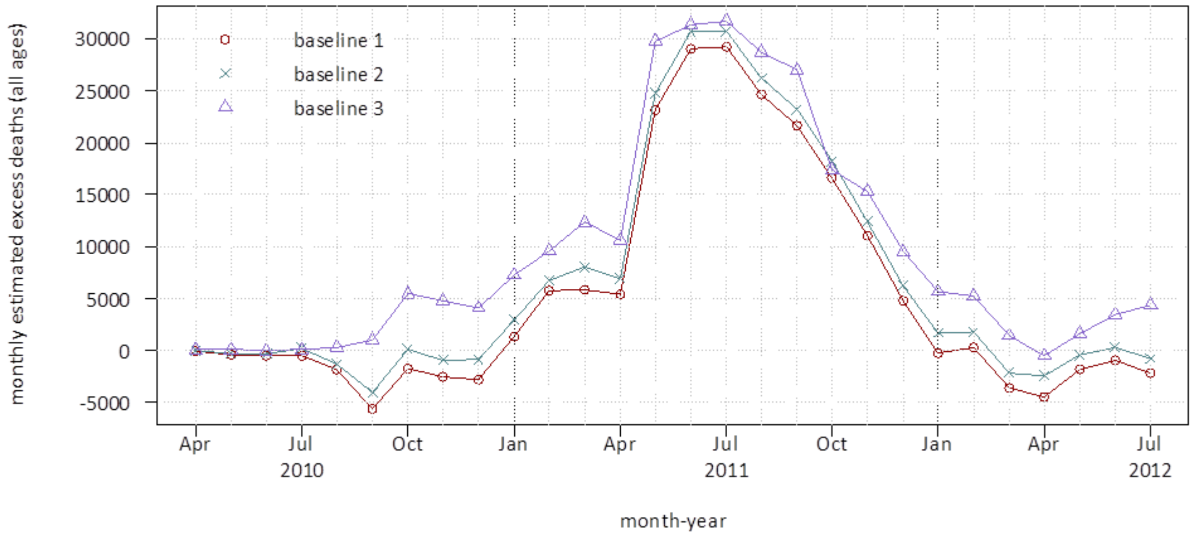

Figure 14. Estimated monthly excess death toll (all ages) in southern and central Somalia, by baseline method used.

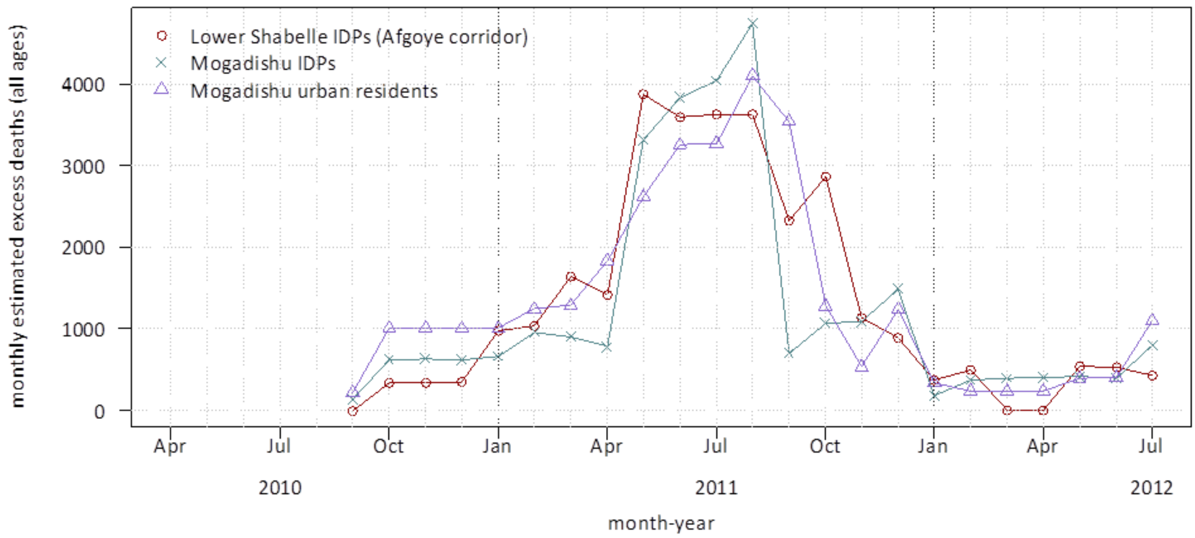

Figure 15. Estimated monthly excess death toll (all ages) in selected strata (baseline method 3 only).

When expressed as percentage of the population at the start of each month, excess mortality over the analysis period was exceptionally high among children and in IDP strata (Table 8). For example, applying results for baseline method 3, more than one in ten IDP children living in Bakool, Banadir, Bay, Lower Shabelle and Middle Juba were estimated to have died during the analysis period as a result of the emergency; more than one in five were estimated to have died in IDP strata of Bakool and Banadir and among riverine populations in Lower Shabelle, the region with consistently highest percentages. These geographic patterns are also shown in Figure 16.

Table 8. Estimated percentage of the population in southern and central Somalia that died in excess of the baseline between Oct 2010 and April 2012 inclusive, by analysis stratum and baseline method used.

Region, livelihood	baseline 1 (surveys done in periods of high food security)		baseline 2 (surveys done in IPC phases 1 and 2)		baseline 3 (model predictions assuming no food security deterioration and no epidemics)	
	all ages	under 5y	all ages	under 5y	all ages	under 5y
Bakool						
agro-pastoralist	8.6	22.1	10.0	27.5	3.5	8.3
IDP	10.9	24.7	11.3	27.1	9.6	22.7
pastoralist	0.8	3.1	0.5	4.2	1.0	2.9
Banadir						
IDP	8.7	22.6	9.1	25.0	8.3	22.4
urban	5.2	9.6	5.6	12.1	5.1	12.1
Bay						
agro-pastoralist	1.3	5.3	2.6	10.9	3.0	14.7
IDP	4.7	10.5	5.0	12.9	7.9	16.2
pastoralist	0.5	2.2	0.4	3.5	1.0	8.5
urban	2.4	2.7	2.8	5.2	2.3	10.2
Galgadud						
agro-pastoralist	0.0	1.2	1.2	6.7	2.5	9.7
IDP	1.2	0.9	1.5	3.5	0.9	3.9
pastoralist	0.0	0.0	0.0	1.0	0.5	3.1
Gedo						
agro-pastoralist	0.0	0.4	1.0	6.1	4.3	7.4
IDP	1.7	4.0	2.0	6.4	5.0	7.0
pastoralist	0.4	2.6	0.0	3.7	3.4	6.0
riverine	1.8	8.3	2.0	10.0	5.1	9.5
Hiran						
agro-pastoralist	0.0	0.0	0.4	4.1	1.8	3.1
IDP	1.6	3.7	2.0	6.1	2.3	4.3
pastoralist	0.3	1.7	0.2	2.8	1.8	3.0
riverine	0.8	4.7	1.0	6.7	1.6	3.0
Lower Juba						
agro-pastoralist	0.0	0.0	0.5	4.3	2.0	4.1
IDP	1.1	1.8	1.4	4.2	2.1	3.1
pastoralist	0.0	1.1	0.0	2.2	1.6	3.1
riverine	0.3	3.6	0.6	5.7	2.0	4.3
urban	0.0	0.0	0.3	0.0	1.3	1.3
Lower Shabelle						
agro-pastoralist	3.4	8.8	4.8	14.3	7.5	15.7
IDP	5.9	15.4	6.2	17.8	9.2	18.5
pastoralist	3.1	9.4	2.8	10.5	6.2	13
riverine	8.2	25.2	8.5	27.4	11.2	24.9
urban	4.8	7.8	4.8	10.3	7.5	13.2
Middle Juba						
agro-pastoralist	0.4	2.4	2.0	8.0	4.3	9.9
IDP	3.6	8.0	3.9	10.5	5.6	11.8
pastoralist	1.7	6.9	1.7	7.9	4.4	10.8
riverine	2.3	8.4	2.5	10.5	4.5	10.4
Middle Shabelle						
agro-pastoralist	0.3	0.0	2.0	5.4	4.1	10.6
IDP	2.0	2.6	2.3	5.0	3.7	9.6
pastoralist	7.1	18.7	6.9	19.8	2.1	5.0
riverine	1.1	3.8	1.3	5.8	2.9	9.4
urban	1.7	0.0	2.1	2.4	3.7	6.5
Mudug						
agro-pastoralist	0.0	0.0	0.0	0.0	1.3	1.2
IDP	0.0	0.0	0.1	0.0	0.0	0.0
pastoralist	0.0	0.0	0.0	0.0	0.0	0.1
Totals	**2.6**	**6.8**	**3.1**	**9.7**	**4.5**	**10.1**

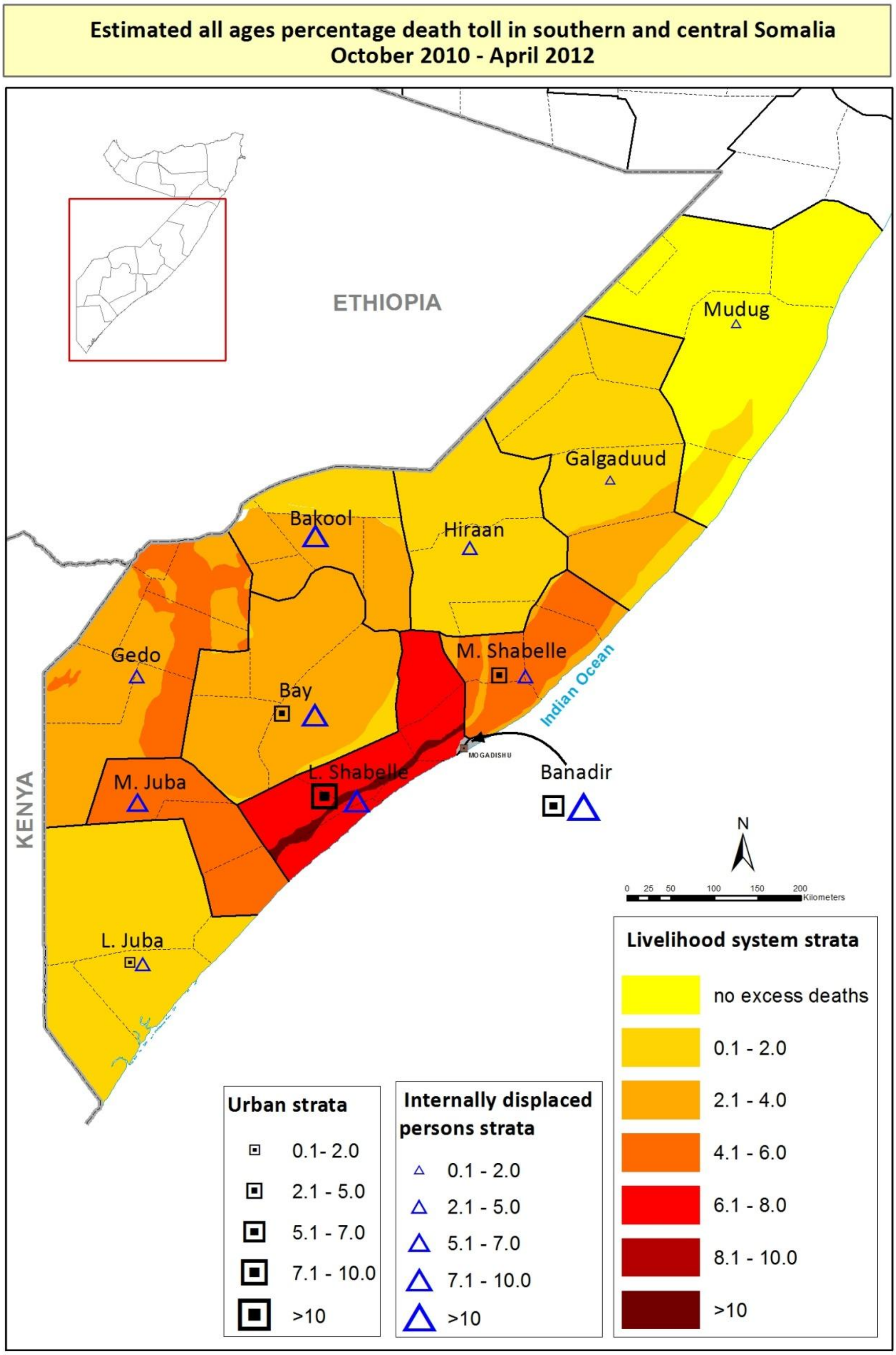

Figure 16. Geographic distribution of the estimated percentage of people who died in excess of the baseline (according to baseline method 3) in southern and central Somalia, October 2010 to April 2012.

3.4.2 Refugee camps

The overall projected ranges of excess mortality (optimistic to pessimistic scenario) for both Ethiopia and Kenya camps were far lower than within Somalia, though most of the range contained values in excess of 1000 (Table 9). The wide spectrum of assumptions made resulted in uninformative estimates that include negative excess mortality.

Table 9. Estimated excess death toll in Ethiopia and Kenya refugee camps (all ages and under 5y old), by analysis stratum and method used to define the baseline. Numbers are, respectively, the median simulation outputs from the optimistic and pessimistic scenarios.

Country, camp	baseline 1 (surveys done in periods of high food security)		baseline 2 (surveys done in IPC phases 1 and 2)		baseline 3 (model predictions assuming no food security deterioration and no epidemics)	
	all ages	under 5y	all ages	under 5y	all ages	under 5y
Ethiopia						
Bokolmanyo	-600 to 1900	-300 to 1700	-600 to 2000	-100 to 1900	-100 to 2500	-100 to 1900
Bur Amino	-300 to 0	-200 to 100	-300 to 100	-100 to 100	-100 to 200	-100 to 100
Dollo Ado transit/ reception centres	-100 to 1400	0 to 1200	-100 to 1400	0 to 1200	100 to 1500	0 to 1200
Hilaweyn	-500 to 0	-200 to 200	-500 to 100	-100 to 300	-100 to 500	-100 to 300
Kobe	-300 to 500	0 to 600	-200 to 600	100 to 700	200 to 1000	100 to 700
Melkadida	-300 to 800	-100 to 700	-300 to 800	-100 to 800	-100 to 1000	-100 to 800
Totals	**-2200 to 4600**	**-800 to 4500**	**-1900 to 4900**	**-300 to 5000**	**-200 to 6700**	**-300 to 4900**
Kenya						
Dahagaley	0 to 0	-100 to -100	100 to 100	100 to 100	600 to 600	200 to 200
Hagadera	-1600 to 1500	-1000 to 800	-1400 to 1600	-600 to 1200	-500 to 2600	-600 to 1200
Ifo 2	-1400 to 2200	-800 to 1200	-1300 to 2300	-600 to 1400	-600 to 3000	-600 to 1400
Ifo	-1100 to 1900	-700 to 1000	-900 to 2100	-400 to 1300	-200 to 2800	-300 to 1400
Kambioos	-300 to 400	-200 to 200	-200 to 400	-100 to 300	-100 to 500	-100 to 300
Totals	**-4400 to 5900**	**-2800 to 3100**	**-3700 to 6500**	**-1600 to 4300**	**-800 to 9500**	**-1400 to 4500**

4 Discussion

4.1 Main findings

To our knowledge, this is the first attempt to estimate mortality comprehensively during the 2010-2012 period of severe food insecurity and famine in Somalia, and it is one of the first statistical analyses of mortality due to famine and severe food insecurity crises affecting humanity in recent decades. Our estimate of about 244,000 to 273,000 excess deaths according to the most plausible baseline (out of a total estimated 580,000 to 611,000 deaths including baseline and excess) suggests that the emergency was very severe, and that efforts to mitigate it, as is widely agreed, fell woefully short. Our study suggests that estimates lower than the above are also plausible, depending on the baseline mortality estimate adopted. All estimates, however, are broadly within a similar order of magnitude. Empirically, trends in survey-estimated death rates and prevalence of acute malnutrition since 2007 strongly suggest that the events of 2010-2012 were highly exceptional, even for Somalia, resulting in death rates far exceeding any mortality estimate seen over previous years, and consistent with the exceptionality usually attributed to famine events.

Findings further suggest that the vast majority of excess mortality occurred inside Somalia and that mortality began rising in late 2010, well before humanitarian agencies began scaling up their response. On this basis, it might be appropriate to revisit the typical schedule of nutrition and mortality surveys done in Somalia, by building in more flexibility to carry ad hoc surveys if deteriorating conditions are suspected, and anticipating the first yearly round of surveys to February-April rather than June-July as done presently. Furthermore, a review of how IPC classifications are used in practice may be required. All phases in the IPC scale require action to prevent further deterioration, but in this Somalia emergency, assistance appears to have been mobilised to an appreciable extent in a mostly reactive way, and only once mortality and other indicators included in the IPC framework reached very critical (i.e. phase 5) levels.

The high percentage of the population that we estimate to have died as a result of the emergency is particularly disturbing: the highest percentages noted were among IDPs and in Bakool, Banadir and Lower Shabelle regions. Among children under 5y old, peak death rates were more than a hundred times greater than those observed in developed countries. We estimate that about 133,000 children under 5y old may have died during the analysis period in southern and central Somalia (population approximately 6-7 million) according to the most plausible baseline method. By comparison, about 65,000 children under 5y old were estimated to have died in 2010 throughout all industrialised countries combined (population 990 million), and 1.3 million in Eastern and Southern Africa (population 399 million) [35].

Particularly high death tolls were noted for Lower Shabelle region and IDPs. While this is not captured explicitly by our estimation, a large number of displaced people left the worst-affected regions on foot, seeking labour and assistance in Mogadishu. Some of these stopped while on the way in Lower Shabelle, so as to seek farm labour. These IDPs might have been particularly vulnerable.

The rapid decline in mortality observed in late 2011 and early 2012 is difficult to explain, though a 1985 study of historical and more recent famines suggested that famines generally are local events and are characterized by high mortality of relatively short duration [36]. Our study was not designed to attribute portions of the estimated mortality, or of its decline, to individual factors. (This would be a very arduous task given the inextricable nature of most of these. Data are likewise insufficient with which to draw any conclusions about the extent to which increased humanitarian assistance in late 2011 and onwards (including potentially greater income from diaspora remittances) may have contributed to the observed mortality decline. Our statistical model nevertheless provides some clues. Firstly, epidemics mainly occurred in the first nine months of 2011 and subsided by late 2011. Particularly among poor agricultural populations, "famines and epidemics often interact synergistically" [37]; that is, famines can lead to epidemics through migration and overcrowding (in urban areas and IDP settlements, for example), while epidemics can precipitate famines by disrupting agricultural activities and further compromising nutritional status. Secondly, the bumper crops of early 2012 greatly improved terms of trade, and by implication, food access: this improvement was very rapid, and would plausibly have dramatically cut the risk of new episodes of malnutrition and consequent health effects. Thirdly, improved access to parts of Somalia (e.g. Mogadishu) and the scale-up of relief wherever access was possible, including by a variety of non-traditional donors and actors, is likely to have made a considerable difference. Indeed, our model suggests strongly that humanitarian assistance was a critical modulator of the effects of food insecurity, helping to contain its effects in 2008 (when assistance was mostly present) while potentiating its negative impact in 2010-2011 (when mostly absent). These three developments occurred around the same period and may together explain the rapid improvement in mortality observed in this study. Above and beyond this, our model findings should be interpreted as supporting the pivotal role of adequate humanitarian assistance in the response to the Somalia crisis.

4.2 Findings in context

Our estimates for Somalia are similar in magnitude to those constructed for the 1992-1993 famine (212,000 to 248,000) [38], although the 1992-1993 study relied mainly on extrapolation of limited existing data, similarly to our estimates for the refugee camps, and famine definitions as well as affected populations were different. Surveys conducted in southern Somalia during this earlier famine offer a second basis for comparison, namely percentages of the pre-famine population estimated to have died in excess of the baseline: in 1992-1993, these percentages were 12-25% for Bay region, 6-16% for Lower Juba, Lower Shabelle and Bakool, and 2-8% elsewhere [38], i.e. generally higher than in our study. In Gode district, Ethiopia, 1999-2000, where the most recent UN-designated famine had occurred prior to 2011, CDR was 3.2 per 10,000 person-days (U5DR 6.8), exacerbated as in this case by measles and other epidemics [39]. In Bahr el Ghazal state of South Sudan (1998), CDRs of 5 to 23 per 10,000 person-days were noted among displaced populations [40].

In nearly all large crises, death toll figures are issued by different bodies, with varying aims and supported by more or less robust estimation methods. Crisis-wide mortality estimates broadly fall into five typologies:

- “Body counts” resulting from systematic listing of mortality reports in the media, other publicly available sources or local initiatives [41,42]. These analyses are only applicable for intentional violence deaths but establish a reasonably robust minimum number.
- Multiple systems estimation (or capture-recapture analysis), an improvement over simple body counts whereby different individual lists of violent deaths are analysed for overlap so as to estimate the number of violent deaths not captured by any list, and thereby the total [43].
- Estimation based on a single survey covering the entire (or most of the) affected region and spanning a sufficiently long period of time; recently this has been a common approach, resulting for example in the oft-quoted estimate of 5-6 million war-related deaths for the Democratic Republic of Congo [29]. Individual surveys are vulnerable to a variety of potential biases and inherent inaccuracy due to limited sample sizes, and, as shown in Iraq, can results in relatively disparate estimates [31,44-46].
- Demographic modelling, usually possible only years or decades after the crisis, and provided that either vital events registration, pre- and post-crisis census, or large-scale demographic survey estimates are available with which to construct a modelled, age-structured population fitted to both the pre- and post-crisis data. As an example, publication of a new 2008 census in North Korea enabled demographic estimation of between 240,000 and 420,000 excess deaths during the 1996-1997 famine [47].

In addition, governments, the UN or other groups typically issue estimates that may not necessarily have a statistical basis, but that nonetheless attract the same or more attention than scientific estimates derived through any of the above methods. As an example, in Haiti the government issued a figure of 316,000 dead as a result of the earthquake, but did not supply a methodology supporting this estimate. At the same time, alternative surveys came up with lower and greatly different totals

(46,000 to 85,000 [48]; 158,000 [49]), although only the government figure continued to be cited in the mainstream media [50].

Discrepancies among mortality estimates for the same crisis can be perplexing to stakeholders in the public, relief agencies, donors and the media. In most crises statistical objectivity struggles to impose itself over conflicting tendencies to either inflate the death toll or under-report it, depending on the perspective of the actor promulgating figures. Generally, estimates that are not accompanied by some uncertainty range and supported by a publicly available description of methods containing a fair assessment of the study's potential limitations should be viewed as potentially less accurate. Estimates for which uncertainty and full study methods are presented may not converge to a single consensus figure, but should all be viewed as contributing to a varying extent towards a more robust body of evidence, and range of possible values, than any single study considered in isolation.

Our study is somewhat distinct from the methodological approaches outlined above in that it analyses a large number of relatively small-area surveys conducted using similar procedures under a common meta-analytic framework. We know of only two other similar analyses, both conducted in Darfur [30,51] (of the two, only Degomme & Guha-Sapir used models to derive an excess death toll, estimated at 298,000). However, neither of those studies reported the predictive power of models used to fit survey data, a key parameter to assess the accuracy of estimates (see below). Our approach was possible because, relative to most other crises [52], and in spite of formidable operating challenges for the FAO/FSNAU and partners, the amount of mortality and other data available for Somalia was considerably greater, finer in geographic resolution, and more systematic. (This was not the case for Dadaab and Dollo Ado camps, despite the latter being far more amenable to the conduct of quality surveys or prospective surveillance). In this respect, we believe that the combined FAO/FSNAU and FEWS NET data collection systems demonstrate that essential and integrated data collection covering long-range early warning, food security, anthropometry, essential service (e.g. vaccination, water and sanitation) coverage and mortality is valuable and feasible even in a very difficult environment, provided that necessary resources and access are available.

4.3 Study validity

We believe that the statistical framework underpinning the estimates within Somalia is acceptably robust. Assuming that (i) survey estimates are unbiased, that (ii) the statistical model used to predict mortality is informed by valid data and that (iii) population denominator data are reasonably accurate, our analysis should yield estimates that have a mostly quantifiable level of statistical confidence. However, the estimation is subject to potential sources of unseen error, as follows:

- The ordinary least-squares model we fitted does not, in fact, account for all of the variability in the household survey mortality data, since it uses meta-data (survey point estimates) as the dependent variable. Had we had access to all of the primary source survey datasets, we would have been able to fit a model (e.g. Poisson or negative binomial) that is better suited to numerically rare events and that would have accounted for household-level variance and additional variance due to cluster sampling design effects. As such, the standard errors produced by our model, and by extension the confidence intervals associated with death toll estimates, are very likely to be overly narrow, meaning our estimates are in fact less precise than we are able to quantify, though probably not biased on this account.
- We do not have complete information on the validity of the survey estimates we relied on, although survey datasets obtained during the critical July-October 2011 period were available to us and also underwent external validation by the United States Centers for Disease Control and Prevention. Even where full survey datasets were available, we did not have sufficient details on survey sampling designs, attrition and potential sources of non-sampling error (such as insufficient interviewer training or problems with questionnaire administration) that would have allowed us to adequately assess the quality of FAO/FSNAU

surveys. Previous reviews have identified frequent quality problems with mortality surveys performed in humanitarian settings [3,53,54]. Moreover, aspects of the questionnaires and sampling designs used by retrospective mortality surveys are known or suspected to feature potential biases: these are reviewed at length elsewhere [15,16]; we note merely that most of these methodological issues have not yet been formally investigated through operational research, despite the increasingly pivotal role of mortality data and surveys in informing humanitarian responses worldwide. In particular, the very abridged questionnaires used in southern Somalia during the 2011 emergency are likely to have featured greater than usual inaccuracy due to their limited probing questions around deaths and births occurring during the recall period (which would lead to under-estimation) and to insufficient provisions for respondents to accurately place events within the start of the recall period (which could lead to either direction of bias). Similarly, detailed age-sex distributions of household members and deaths were missing, precluding more refined life table and age-stratified analyses. On the other hand, we found numerous indications than, on the whole, the quality of FAO/FSNAU surveys is likely to be acceptable and comparatively higher than that of surveys done in other crises worldwide, despite an arguably more difficult operating environment. Assuming the set of surveys we were able to re-analyse was representative of those we did not have access to, we believe that the coherency of different demographic statistics estimated by the surveys, such as birth rate, household size or proportion of children under 5y old suggests a pattern of acceptable quality (see Annex for more comments). Furthermore, the ongoing nature of FAO/FSNAU's work, with emphasis on training local researchers and systematic quality control of anthropometric datasets on Emergency Nutrition Assessment (ENA) software (see http://www.nutrisurvey.de/ena/ena.html), would be expected to increase accuracy compared to one-off surveys done by agencies in settings unfamiliar to researchers.

- When analysing surveys done among IDPs, we assumed due to lack of additional information that all deaths had occurred in the stratum where the survey was done: however, this was clearly an imperfect assumption, as many IDPs recently arrived would have died en route (e.g. many IDP deaths in surveys done in Banadir camps could have occurred in Lower Shabelle or other regions). This assumption would not have caused bias in the overall estimate for southern and central Somalia, but may have resulted in a biased distribution of estimated IDP deaths among the strata, with strata receiving IDPs over-represented in the death toll.

- Data on model predictor variables most likely feature considerable bias. In particular, the ACLED dataset of armed conflict events, while comprehensive, relies on media and other remote accessible reports that may result in differential reporting rates for different districts of Somalia; reports of epidemics were non-specific and did not allow clear identification of the populations most affected, and of the relative severity of the outbreak; while terms of trade indicators used were the result of numerous assumptions and were built upon data collected in sentinel markets that may or may not be representative of their surrounding populations. Moreover, both terms of trade and the presence of humanitarian assistance were imperfect proxies for broader dynamics of food security and humanitarian assistance. We were restricted in our choice of model variables by those that were collected continuously for all stratum-months under analysis: therefore, we were unable to meaningfully use the full extent of information collected by FAO/FSNAU surveys, including prevalence of acute malnutrition, feeding practices, vaccination coverage, and access to water and sanitation, all of which were plausibly predictive of mortality but none of which were collected in strata or during months not covered by a survey. The model's fit to the data and its strong biological plausibility, however, suggest, that these biases are unlikely to have a large influence on the findings. When applied to a large number of surveys, the model fits the data reasonably well, does not appreciably over- or under-estimate mortality and maintains a good fit when applied to new datasets (i.e. upon cross-validation): this may be taken as some evidence that it is informed by reasonably accurate data.

- Population denominators are subject to various assumptions and may be inaccurate. Somalia has not had a national census since 1975 and more recent surveys and estimation exercises have yielded a range of figures. Perhaps the most widely used is a figure of 7.5 million based on a 2005 survey carried out by UNDP. Other population estimates for that same year, however, yielded numbers of 8.5 and 7.9 million [11]. We used a figure of 9.3 million for the total population of Somalia as of June 2010, which was generated by AfriPop based on most recent UN population estimates and spatially allocated using land cover information derived from satellite imagery. Miscalculations of total or livelihood zone populations for the analysis period could have led to over- or under-estimations of death tolls.
- While there is considerable evidence that famines contribute to a decline in fertility [36,37], the FSNAU surveys done in 2011 and 2012 suggested that birth rates had not decreased significantly (data not shown). If birth rates, in fact, had declined, then our model would overestimate births and thus overestimate the number of children under-five who were exposed to higher death rates during the emergency, i.e. result in artificially high death toll estimates.
- As regards IDPs specifically, a key limitation was the general lack of data on patterns and duration of displacement during the analysis period, other than monthly estimates of arrivals into and departures from particular regions in Somalia (note that these were figures issued with a caution by UNHCR-PMT not to merely cumulate them in order to estimate IDP totals). Better data on the duration of displacement and the various trajectories of displacement (return to place of origin, movement elsewhere in Somalia, or movement out of Somalia to refugee camps or other destinations) might have enhanced the robustness of the model. Given the assumptions made about displacement, for example, subtracting IDPs from strata proportionally based on data on regions of origin in 2009-2010 left a total of 417,000 people in Mogadishu (Banadir urban) as of June 2010, which may be an under-estimate. Under 5y old populations have been estimated crudely (25% of the total population at the start of the analysis period), although we don't believe that the estimated range of the percent of the population under 5y old excludes likely values for any of the strata. Similarly, more "granular" data on the age-sex composition of the population were too sparse to attempt age adjustment and further age or sex stratification of the estimates.
- By contrast, the estimates for Kenya and Ethiopia refugee camps do not arise from a quantifiable statistical process, but rather from best judgement informed by limited mortality data and mostly non-quantitative circumstantial evidence. Moreover, population estimates of new arrivals to the camps were subject to numerous assumptions required to fill gaps in available data (in particular, we are likely to under-estimate new arrivals in Kenya as we ignore resettlement and return flows; and we may have misallocated resettlement to new camps). Therefore, the refugee camp estimates should be viewed as considerably less robust than those for Somalia. We believe however that our approach of delineating optimistic and pessimistic scenarios is likely to have bounded the true death toll within a realistic, albeit very broad range.

We attempted to define alternative baseline mortality estimates that actually reflected the aim of the study, namely to isolate excess mortality due to food insecurity. As a result, all our baseline estimates are specific to Somalia (and, for method 3, to each stratum) and already incorporate chronically raised mortality, compared to other countries in the region. We believe that these baseline estimation methods encompass the likely plausible range, as long as the basic premise holds that the empirically observed spike in mortality during 2011 was indeed mainly a result of food insecurity greatly in excess of typical levels: surveys of refugees arriving to Kenya suggest that 91% left Somalia because of the drought [23], corroborating this premise.

More broadly, our death toll estimates reflect the geographic and period scope of our analysis. Specifically, we did not quantify excess mortality among:

- Informal migrants and refugees that left Somalia as a result of the emergency for destinations other than Dadaab and Dollo Ado camps: we are not aware of reliable sources of data on these migratory flows.
- Refugees already living in the camps at the outset of the emergency, and who were exposed to a higher risk of mortality due to the new refugee influx (e.g. as a result of increased pressure on public health services, overcrowding or seeding of cholera and measles outbreaks): we did not identify any reliable means of projecting excess mortality among these older arrivals.
- Populations within Somalia and camp-based refugees beyond July 2012. It is plausible to assume that factors such as acute malnutrition becoming chronic in a larger than normal percentage of children; maternal and neonatal mortality due to poor nutrition during pregnancy; and forced migration (resulting for example in greater exposure to infectious diseases such as diarrhoea, pneumonia [55] and tuberculosis [56], or higher risk of suicide due to increased frequency of mental disorders [57]) will continue to have pernicious effects on Somalis' health for years and decades to come.

4.4 Conclusions

Our statistical estimates should be viewed with due appreciation of the potential sources of error that may have influenced them, in addition to the known imprecision that we were able to quantify and present. Our methods are open to criticism, as are those of other studies of mortality in recent wars and disasters. To this end, data, statistical code and detailed methods we used are presented alongside this report. We hope that this report contributes to global evidence and methodology to better estimate the effects of armed conflict, food insecurity and crises on human health. We note that much of the inaccuracy in our methods would be reduced if agencies carry out well-planned, coordinated surveys or establish prospective mortality surveillance on a regular basis wherever humanitarian access makes this feasible, while contemporaneously tracking population movements, including return flows and length of displacement; and if data collected are made available publicly, to the extent that this would not compromise beneficiaries' security or relief operations.

Despite limitations, we believe that our study is sufficiently robust to conclude that a very large number of people (very probably between 143,000 and 273,000, and more likely to be closest to our best estimate of 258,000) died in southern and central Somalia as a result of the food insecurity emergency of 2010-2012. Our study was not designed to explore causes of excess mortality, and others are better placed to interpret these findings in context and translate them into action. Nevertheless, an obvious rational use of our estimates would be to ensure that in Somalia and food security-related emergencies elsewhere, future early warning and surveillance alerts do translate into tangible, immediate relief interventions to support livelihoods, health and nutrition in the affected communities.

While this analysis has attempted to circumscribe its estimate of excess mortality to that attributable to (severe) food insecurity, our findings should be viewed within a broader framework in which armed conflict and insecurity probably hold the greatest share of ultimate attribution. The return of peace and stable government to all parts of Somalia is likely to be the most effective and durable solution to reduce future threats to the livelihood and well-being of Somalis everywhere. This study offers a compelling case for all parties to the armed conflict to pursue its resolution as a matter of the greatest urgency.

5 References

1. Maxwell D, Fitzpatrick M (2012) The 2011 Somalia famine: Context, causes, and complications. Global Food Security 1: 5-12.

2. Food Security and Nutrition Analysis Unit - Somalia (2011) Technical series report no. VI. 42. Food security and nutrition analysis Post *Gu* 2011. Nairobi: FSNAU. http://www.fsnau.org/downloads/FSNAU-Nutrition-Technical-Series-Report-Post-Gu-2011-Analysis.pdf (accessed 12 Feb 2013).

3. Boss LP, Toole MJ, Yip R (1994) Assessments of mortality, morbidity, and nutritional status in Somalia during the 1991-1992 famine. Recommendations for standardization of methods. Jama 272: 371-376.

4. Moore PS, Marfin AA, Quenemoen LE, Gessner BD, Ayub YS, et al. (1993) Mortality rates in displaced and resident populations of central Somalia during 1992 famine. Lancet 341: 935-938.

5. Cambrezy C (1997) Nutrition surveys in Somalia, 1980-1996. Nairobi: FSNAU. http://www.fsnau.org/downloads/Historical-Nutrition-Trends-Somalia-1980-1996.pdf (accessed 5 Feb 2013).

6. Menkhaus K (2012) No access: critical bottlenecks in the 2011 Somali famine. Global Food Security 1: 29-35.

7. Salama P, Moloney G, Bilukha OO, Talley L, Maxwell D, et al. (2012) Famine in Somalia: evidence for a declaration. Global Food Security 1: 3-19.

8. Hillbruner C, Moloney G (2012) When early warning is not enough - Lessons learned from the 2011 Somalia Famine. Global Food Security 1: 20-28.

9. Maxwell D, Haan N, Gelsdorf K, Dawe D (2012) The 2011-12 famine in Somalia: introduction to the Special Edition. Global Food Security 1: 1-4.

10. R Core Team (2012) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing

11. Linard C, Alegana VA, Noor AM, Snow RW, Tatem AJ (2010) A high resolution spatial population database of Somalia for disease risk mapping. Int J Health Geogr 9: 45.

12. Isanaka S, Grais RF, Briend A, Checchi F (2011) Estimates of the duration of untreated acute malnutrition in children from Niger. Am J Epidemiol 173: 932-940.

13. Checchi F, Roberts L (2008) Documenting mortality in crises: what keeps us from doing better. PLoS Med 5: e146.

14. Salama P, Spiegel P, Talley L, Waldman R (2004) Lessons learned from complex emergencies over past decade. Lancet 364: 1801-1813.

15. Cairns KL, Woodruff BA, Myatt M, Bartlett L, Goldberg H, et al. (2009) Cross-sectional survey methods to assess retrospectively mortality in humanitarian emergencies. Disasters 33: 503-521.

16. Working Group for Mortality Estimation in Emergencies (2007) Wanted: studies on mortality estimation methods for humanitarian emergencies, suggestions for future research. Emerging Themes in Epidemiology 4: 9.

17. Raleigh C, Linke A, Hegre H, Karlsen J (2010) Introducing ACLED-Armed Conflict Location and Event Data. Journal of Peace Research 47: 1-10.

18. United Nations High Commissioner for Refugees (2011) Joint UNHCR, WFP, ARRA and MSF-Spain annual health and nutrition survey at Boqolmayo and Melkadida refugee camps. Addis Ababa, Ethiopia: UNHCR.

19. Sackl A (2011) Rapid assessment: New settlements of Dagahaley camp, Kenya, MSF-OCG. Geneva: Médecins Sans Frontieres Operational Centre Geneva.

20. Centers for Disease C, Prevention (2011) Notes from the field: mortality among refugees fleeing Somalia--Dadaab refugee camps, Kenya, July-August 2011. MMWR Morb Mortal Wkly Rep 60: 1133.

21. Bernasconi A (2011) Somalian refugees in Ethiopia: Impact of emergency health relief activities in Dolo Ado. Paris: Epicentre.

22. Sorkin N, Bernasconi A (2012) Mortality surveillance system: case study from Liben refugee camps, nutrition emergency, Somali Region 2011. MSF-UK Scientific Day. London. http://issuu.com/msfuk/docs/naomi_sorkin?mode=window&pageNumber=1 (accessed 30 December 2012).

23. Polonsky JA, Ronsse A, Ciglenecki I, Rull M, Porten K (2013) High levels of mortality, malnutrition, and measles, among recently-displaced Somali refugees in Dagahaley camp, Dadaab refugee camp complex, Kenya, 2011. Confl Health 7: 1.

24. United Nations High Commissioner for Refugees, Administration for Refugee and Returnee Affairs, World Food Programme, United Nations Children's Fund, GOAL (2011) Health and Nutrition Survey Report - Kobe and Hilaweyn Refugee Camps, Dollo Ado Woreda, Somali Regional State of Ethiopia, Mid October to Early November 2011. Addis Ababa: UNHCR.

25. International Organisation for Migration (2011) IOM Response to the Horn of Africa Crisis external situation report - 25 October 2011. Nairobi: IOM. http://nairobi.iom.int/resources/sitreps/item/download/50_5edb164d7c12a93cf60d2b13c5747a13 (accessed 28 November 2012).

26. Spiegel PB, Sheik M, Woodruff BA, Burnham G (2001) The accuracy of mortality reporting in displaced persons camps during the post-emergency phase. Disasters 25: 172-180.

27. United Nations High Commissioner for Refugees (2011) Kambioos Camp Profile - April 2012. Nairobi: UNHCR. http://data.unhcr.org/horn-of-africa/download.php?id=862 (accessed 3 January 2013)

28. United Nations High Commissioner for Refugees (2011) Nutrition Surveys - Dadaab Refugee Camps - Hagadera, Ifo, And Dagahaley Camps & Dagahaley Outskirts. Nairobi: UNHCR.

29. Coghlan B, Brennan RJ, Ngoy P, Dofara D, Otto B, et al. (2006) Mortality in the Democratic Republic of Congo: a nationwide survey. Lancet 367: 44-51.

30. Degomme O, Guha-Sapir D (2010) Patterns of mortality rates in Darfur conflict. Lancet 375: 294-300.

31. Burnham G, Lafta R, Doocy S, Roberts L (2006) Mortality after the 2003 invasion of Iraq: a cross-sectional cluster sample survey. Lancet 368: 1421-1428.

32. Connolly MA, Gayer M, Ryan MJ, Salama P, Spiegel P, et al. (2004) Communicable diseases in complex emergencies: impact and challenges. Lancet 364: 1974-1983.

33. Lindley A, Hasley A (2011) Working paper 79 - Unlocking protracted displacement: Somali case study. Oxford: University of Oxford. http://www.rsc.ox.ac.uk/publications/working-papers-folder_contents/RSCworkingpaper79.pdf/view (accessed 2 April 2013).

34. Jureidini R (2010) Mixed Migration Flows: Somali And Ethiopian Migration To Yemen And Turkey - Final Report. Cairo: Center for Migration and Refugee Studies, American University in Cairo. http://www.drc.dk/fileadmin/uploads/pdf/IA_PDF/Horn_of_Africa_and_Yemen/Mixed%20Migration%20through%20Somalia%20and%20across%20Gulf%20of%20Aden%20-%20Background%20study.pdf (accessed 20 February 2013).

35. United Nations Children's Fund (2012) The state of the world's children 2012: children in an urban world. New York: UNICEF. http://www.unicef.org/sowc2012/pdfs/SOWC%202012-Main%20Report_EN_13Mar2012.pdf (accessed 12 March 2013).

36. Watkins SC, Menken J (1985) Famines in Historical Perspective. Population and Development Review 11: 647-675.

37. Dyson T, O'Grada C (2002) Famine Demography: Perspectives from the Past and Present. Oxford: Oxford University Press.

38. Hansch S, Lillibrdge S, Egeland G, Teller C, Toole M (1994) Lives lost, lives saved: Excess mortality and the impact of health interventions in the Somalia emergency. Washington, DC: Refugee Policy Group. http://www.fews.net/docs/.../Lives%20Lost,%20Lives%20Saved.pdf (accessed 10 October 2012).

39. Centers for Disease C, Prevention (2001) Mortality during a famine--Gode district, Ethiopia, July 2000. MMWR Morb Mortal Wkly Rep 50: 285-288.

40. Biong Deng L (1999) Famine in the Sudan: causes, preparedness and response : a political , social and economic analysis of the 1998 Bahr el Ghazal famine. Brighton: Institute of Development Studies. http://www.ids.ac.uk/files/dmfile/Dp369.pdf (accessed 15 February 2013).

41. Hicks MH, Dardagan H, Bagnall PM, Spagat M, Sloboda JA (2011) Casualties in civilians and coalition soldiers from suicide bombings in Iraq, 2003-10: a descriptive study. Lancet 378: 906-914.

42. Price M, Klingner J, Ball P (2013) Preliminary Statistical Analysis of Documentation of Killings in the Syrian Arab Republic. Benetech Human Rights Program. https://www.benetech.org/download/Benetech-final-SY-report.pdf (accessed 22 February 2013).

43. Comisión para el Esclarecimiento Histórico (1999) Guatemala: Memoria del Silencio. Guatemala City: United Nations Office of Project Services.

44. Zeger SL, Johnson E (2007) Estimating excess deaths in Iraq since the US-British-led invasion. Significance 4: 54-59.

45. Alkhuzai AH, Ahmad IJ, Hweel MJ, Ismail TW, Hasan HH, et al. (2008) Violence-related mortality in Iraq from 2002 to 2006. N Engl J Med 358: 484-493.

46. Tapp C, Burkle FM, Jr., Wilson K, Takaro T, Guyatt GH, et al. (2008) Iraq War mortality estimates: A systematic review. Confl Health 2: 1.

47. Spoorenberg T, Schwekendiek D (2012) Demographic changes in north Korea: 1993–2008. Population and Development Review 38: 133-158.

48. Schwartz TT, Pierre Y-F, Calpas E (2011) Building Assessments and Rubble Removal in Quake-Affected Neighborhoods in Haiti. Washington, DC: USAID. http://pdf.usaid.gov/pdf_docs/PNADY468.pdf (accessed 19 February 2013).

49. Kolbe AR, Hutson RA, Shannon H, Trzcinski E, Miles B, et al. (2010) Mortality, crime and access to basic needs before and after the Haiti earthquake: a random survey of Port-au-Prince households. Med Confl Surviv 26: 281-297.

50. O'Connor MR (2012) Two Years Later, Haitian Earthquake Death Toll in Dispute. Columbia Journalism Review: http://www.cjr.org/behind_the_news/one_year_later_haitian_earthqu.php?page=all (accessed 21 February 2013).

51. Nielsen J, Prudhon C, de Radigues X (2011) Trends in malnutrition and mortality in Darfur, Sudan, between 2004 and 2008: a meta-analysis of publicly available surveys. Int J Epidemiol 40: 971-984.

52. Checchi F (2010) Estimating the number of civilian deaths from armed conflicts. Lancet 375: 255-257.

53. Prudhon C, de Radigues X, Dale N, Checchi F (2011) An algorithm to assess methodological quality of nutrition and mortality cross-sectional surveys: development and application to surveys conducted in Darfur, Sudan. Popul Health Metr 9: 57.

54. Prudhon C, Spiegel PB (2007) A review of methodology and analysis of nutrition and mortality surveys conducted in humanitarian emergencies from October 1993 to April 2004. Emerg Themes Epidemiol 4:10.

55. Bellos A, Mulholland K, O'Brien KL, Qazi SA, Gayer M, et al. (2010) The burden of acute respiratory infections in crisis-affected populations: a systematic review. Confl Health 4: 3.

56. Kimbrough W, Saliba V, Dahab M, Haskew C, Checchi F (2012) The burden of tuberculosis in crisis-affected populations: a systematic review. Lancet Infect Dis 12: 950-965.

57. Tol WA, Barbui C, Galappatti A, Silove D, Betancourt TS, et al. (2011) Mental health and psychosocial support in humanitarian settings: linking practice and research. Lancet 378: 1581-1591.

58. Centers for Disease C, Prevention (2011) Notes from the field: malnutrition and mortality--southern Somalia, July 2011. MMWR Morb Mortal Wkly Rep 60: 1026-1027.

6 Statistical annex

6.1 Data sources and management

6.1.1 Geo-coding and administrative divisions

Somalia is divided into regions and districts. Datasets used for this analysis utilised various anglicised spellings of Somali administrative names. We adopted a simple Westernised spelling of region, district and city names (see analysis datasets).

Datasets also featured inconsistent listings of districts, reflecting the creation of new administrative divisions by different local administrations since the ousting of the Barre government. Despite extensive internet searches we were unable to determine the exact boundaries of many of these new districts, and likewise had no reliable population figures for them. Therefore, our analysis adopts the administrative district and regional boundaries of Somalia before the ousting of the Barre government in 1991 (see analysis datasets). Within Banadir region (i.e. Mogadishu), we aggregated all city districts into a single super-district, as we did not have any data broken down below this division. Where appropriate we also divided Somalia into a central sector (Mudug and Galgadud regions) and a southern sector (all other southern regions).

Administrative divisions and nomenclature adopted for this analysis are purely functional to our estimation purposes and do not imply recognition of different administrative authorities, or lack thereof. Our choices also ultimately do not influence the results of this analysis.

6.1.2 Livelihood zones

We adopted the classification, nomenclature and borders of livelihood zones throughout Somalia as defined by FAO/FSNAU and FEWS NET: see map on www.fews.net/docs/Publications/-SO_Livelihoods.pdf and corresponding geographical datasets www.fews.net/livelihood/so/-Shapefiles.zip. However, we ignored the fishing, Awdal border and potato zone livelihoods (codes 2, 10, 24 on the above map) as they were small and no surveys to our knowledge had covered these livelihood zones.

6.1.3 Survey metadata

We systematically reviewed all of FAO/FSNAU's publicly available publications (see www.fsnau.org/products) for reports of household surveys done between January 2007 and July 2012 in Somalia (including Puntland and Somaliland). We restricted data collection to this period as surveys done previously mostly included a mixture of different livelihood zones with their sampling frames without any stratification, whereas starting in 2007 livelihood zones were used as the survey sampling frames or as explicit strata within survey samples, in line with the wider FAO/FSNAU analysis framework. We included surveys with an administrative rather than livelihood-based sampling universe (e.g. district) only if the population within the universe had a dominant livelihood (> 75% of the population based on FAO/FSNAU regional livelihood profiles). We further excluded 111 surveys that relied on non-representative sampling designs (e.g. rapid anthropometric screening assessments; in practice only two of these estimated death rates).

Before 2009, FAO/FSNAU published individual, detailed reports on each survey (see www.fsnau.org/products/survey-reports/nutrition/archives). Thereafter, so as to streamline the burden of reporting as the frequency of surveys increased, summary results for each surveys were presented within Technical Series (www.fsnau.org/products/technical-series) or Nutritional Updates (www.fsnau.org/products/nutrition-update) covering bi-yearly cycles of integrated assessment. As a result, the methods and actual sampling output (e.g. attrition, replacement of clusters) of surveys since 2009 are mostly not described and meta-data are more limited.

For each survey included in the review, we extracted meta-variables that we considered plausibly related to mortality. Typically, surveys featured at a minimum a mortality, anthropometry, childhood illness and measles vaccination module, though the amount of information collected varied very widely over time and space, with surveys in stable regions of Somaliland and Puntland collecting detailed data on food security and feeding practices, and those done during emergencies in insecure areas of Somalia restricted to mortality and anthropometry.

Table 10. Meta-variables collected for each survey included in the review (N = 207 surveys total). Data availability as per publicly available FAO/FSNAU reports.

Variable	Description	Number of surveys with sufficient data (% of total)
Survey characteristics		
Survey year	year in which the survey was done	207 (100.0)
Survey month	month in which the survey was done	207 (100.0)
State(s)	state(s) included in the sampling universe of the survey (Somalia, including Puntland, Somaliland)	207 (100.0)
Region(s)	region(s) included in the sampling universe	207 (100.0)
Other location	other information on which specific district(s), towns or other divisions (e.g. parts of a livelihood zone) were included in the sampling universe	207 (100.0)
Livelihood	main livelihood of the population in the sampling universe (agro-pastoralist, pastoralist, riverine [agriculturalist], internally displaced, urban)	207 (100.0)
Sampling design	two-stage cluster sampling with probability proportional to size allocation of clusters assumed unless specified as an exhaustive sample	207 (100.0)
Number of clusters	number of clusters if cluster sampling was done	75 (36.2)
Bias and quality issues	free text notes on various problems with quality (see section 6.3.4)	207 (100.0)
Essential service coverage		
Bed net ownership	proportion of households owning at least one bed net	81 (39.1)
Clean water access	proportion of households that have access to a protected source of drinking water (self-reported)	150 (72.5)
Sanitation access	proportion of households that have access to a latrine or other sanitary excreta disposal device (self-reported)	145 (70.0)
Measles vaccination coverage	proportion of children 9-59m who have received at least one dose of measles vaccine according to card or caregiver	170 (82.1)
Polio vaccination coverage	proportion of children 9-59m who have received at least one dose of polio vaccine according to card or caregiver	95 (45.9)
Child health		
Illness attack rate	proportion of children 6-59m that have been ill in the past two weeks according to caregiver	190 (91.8)
Diarrhoea attack rate	proportion of children 6-59m that have had diarrhoea in the past two weeks according to caregiver	160 (77.3)
Health-care seeking	proportion of children 6=-59m reported ill in the past two weeks who received care in a public health facility	45 (21.7)
Occurrence of epidemics	whether any confirmed infectious disease outbreak was affecting the surveyed population during the recall period, and if so which aetiologic agent was responsible for the outbreak	207 (100.0)
Feeding practices		
Nutritional diversity	proportion of households consuming four or more food groups	149 (72.0)
Breastfeeding prevalence	proportion of children 6-24mo who are still breastfed	108 (52.2)
Anthropometric status		
Prevalence of global acute malnutrition (WHO 2006)	point estimate of global acute malnutrition prevalence among children 6-59mo old according to WHO 2006 reference standards (definition: weight-for-height Z score < 2 standard deviations and/or presence of bilateral oedema)	192 (92.8)
Prevalence of severe acute malnutrition (WHO 2006)	point estimate of severe acute malnutrition prevalence among children 6-59mo old according to WHO 2006 reference standards (definition: weight-	193 (93.2)

Variable	Description	Number of surveys with sufficient data (% of total)
	for-height Z score < 3 standard deviations and/or presence of bilateral oedema)	
Prevalence of global acute malnutrition (NCHS)	point estimate of global acute malnutrition prevalence among children 6-59mo old according to National Centre for Health Statistics (pre-2006) reference standards (definition as above)	202 (97.6)
Prevalence of severe acute malnutrition (NCHS)	point estimate of severe acute malnutrition prevalence among children 6-59mo old according to National Centre for Health Statistics (pre-2006) reference standards (definition as above)	202 (97.6)
Prevalence of global acute malnutrition according to middle-upper arm circumference (MUAC)	point estimate of global malnutrition prevalence expressed using MUAC (definition: MUAC < 125 mm and/or bilateral oedema)	201 (97.1)
Prevalence of severe acute malnutrition according to MUAC 110mm threshold	point estimate of severe malnutrition prevalence expressed using old MUAC threshold (definition: MUAC < 110 mm and/or bilateral oedema)	46 (22.2)
Prevalence of severe acute malnutrition according to MUAC 115mm threshold	point estimate of severe malnutrition prevalence expressed using new MUAC threshold (definition: MUAC < 115 mm and/or bilateral oedema)	137 (66.2)
Prevalence of oedema	proportion of children assessed who have bilateral oedema	167 (80.7)
Proportion of acute malnutrition among pregnant and lactating women	proportion of pregnant and lactating women with MUAC <23cm	173 (83.6)
Mortality		
Recall period duration	three months for all surveys	207 (100.0)
Number of households	number of households to whom the mortality questionnaire was administered	68 (32.9)
Number of children under 5y old	number of children under 5y old included in the mortality analysis	38 (18.4)
Deaths under 5y	number of deaths of children under 5y old reported by households interviewed	82 (39.6)
Number of people	number of people of all ages included in the mortality analysis	43 (20.8)
Deaths	number of deaths of people of all ages reported by households interviewed	45 (21.7)
CDR point estimate	point estimate of crude death rate (deaths per 10000 person-days)	207 (100.0)
CDR lower confidence interval	lower 95% CI of CDR	207 (100.0)
CDR upper confidence interval	upper 95% CI of CDR	207 (100.0)
U5DR point estimate	point estimate of under 5y death rate (deaths of children under 5y old per 10000 child-days)	207 (100.0)
U5DR lower confidence interval	lower 95% CI of CDR	207 (100.0)
U5DR upper confidence interval	upper 95% CI of CDR	207 (100.0)

6.1.4 Sentinel market data

FAO/FSNAU field researchers collect data on a monthly basis in 54 sentinel markets located in urban centres throughout Somalia. Data collected include sale prices of standard quantities of various food and non-food items, as well as the typical daily wage of a labourer. We extracted data from these markets for the period Jun 2006 to May 2012 inclusive (72mo), and for the following price variables: 1 Kg white sorghum; 1 Kg red sorghum; 1 Kg yellow maize; 1 Kg white maize; 1 Kg imported red rice; 1 Kg wheat flour; 1 goat of local quality (i.e. not for export); daily wage in the local currency. Sorghum and maize are locally produced while wheat and rice are mostly imported. After removing obvious outlier values based on various range checks, we defined market time series as eligible for further analysis if they contained price data on at least one of the above cereals and the price of a local quality goat and the daily wage for ≥ 75% of the months in the series (54mo), with no continuous break in data coverage of > 3mo. Forty markets met these eligibility criteria. We noted that sufficient time series for red sorghum were mostly unavailable, despite this cereal having been severely affected by the drought and also being a preferred staple in many parts of southern Somalia. However, data on white sorghum, probably a reasonable replacement, were almost complete.

We determined which of the eligible markets fell within the sampling frame of each mortality survey and within each analysis stratum. These markets were used to quantify food security indicators specific to each survey and analysis stratum. One analysis stratum (Bay pastoralist) had no market within it, so nearby Baidoa and Dinsor markets were associated with this stratum instead. Similarly, the sampling frames of 16 surveys did not contain any market: we assigned to each of these one or two nearby markets, reflecting plausible commercial routes (see variables <backup_district1> and <backup_district2> in the analysis datasets).

If data were missing at the start or end of the time series, they were assumed equal to the first or last available monthly observation. The price series were then mildly smoothed in order to enhance trend signals over random fluctuation and interpolated for missing months, using the spline R function "*smooth.spline*" with "*spar*" (resolution of smoothing) parameter set at 0.5. Smoothing was specific to each market series and did not account for trends in neighbouring markets.

We combined the above variables to compute several so-called terms of trade food security indicators for each month and for each analysis stratum or survey. Terms of trade within the context of this analysis, and consistent with food security analyses done by FAO/FSNAU and FEWS NET, refers to the amount of staple cereal (expressed in Kcal) that can be purchased at any given time by a typical daily wage (or, in the case of pastoralists, the selling price of a local quality goat, a key income source for this livelihood). Terms of trade obviate the need to adjust for inflation and currency exchange rates, and provide a directly relevant quantification of the ability to procure food other than that available to households from subsistence agriculture or stored supplies. The main indicator we considered is the monthly absolute terms of trade defined as the Kcal cereal equivalent of a typical daily wage (AToTW).

Table 11 summarises the computation of this indicator and input parameters. The corresponding equation is as follows:

$$ATоT_{W,S,t} = \frac{\sum_{m}^{M_S} \frac{\sum_{c}^{C_m} \left(\frac{W_{m,t}}{I_{c,m,t}} r_c K_c\right)^2}{\sum_{c}^{C_m} \frac{W_{m,t}}{I_{c,m,t}} r_c K_c} U_m}{\sum_{m}^{M_S} U_m}$$

In the absence of data on relative purchase quantities of different cereals, we assumed based on simple elasticity principles that, in the context of a poor harvest and high poverty levels, people would have had a strong preference for purchasing the cheapest cereal, and would have shifted their preference according to cereal prices, with a greater tendency to shift to cheaper staples, the lower the daily wage. Generally, market data showed that local cereal prices increased steeply in 2010-2011 (presumably as supplies dwindled) and decreased very dramatically after the late 2011 bumper harvest, while imported cereal prices were higher on average considering the entire time series, and much more stable, possibly as they were much less subject to limited supply (data not shown). We represented this assumed elasticity by using the AToT for each cereal itself as a weight when averaging all cereals (mathematically this is equivalent to averaging the squares of each cereal's AToT).

Table 11. Composition of terms of trade indicators used in the analysis.

Step	Explanation	Parameters	Notes
1	For each market in the analysis stratum or survey sampling frame, for each month and for each cereal, determine the Kg cereal equivalent of one daily wage.	$AToT_W$ = absolute terms of trade (cereal vs. wage) S = stratum or survey sampling frame t = month m = market (out of M_S in stratum or survey sampling frame) c = cereal (out of C_m cereals with sufficient data coverage in market m) I_c = price of 1 Kg of cereal c W = average daily labour wage	Used smoothed data.
2	For each cereal, transform Kg into the Kcal equivalent after adjusting for loss due to home milling.	r_c = home milling recovery rate of cereal c K_c = Kcal per Kg milled of cereal c	Milling recovery proportions and Kcal/Kg equivalents provided by the Famine Early Warning Systems Network (see analysis datasets)
3	Average all cereals-specific AToT into a single terms of trade indicator, weighting for assumed relative consumption of each cereal.		Only included in the average cereals for which data coverage over the series was ≥ 75%. See text for choice of weights.
4	Average terms of trade over the entire stratum or survey sampling frame, by weighting.	U_m = urban population of the district housing market m	District urban population based on UNDP 2005 data, and taken to be a proxy of the relative importance of the market.

In addition to $AToT_W$, we computed in a similar way other terms of trade indicators, as follows:

- Absolute terms of trade as Kcal cereal equivalent of a local quality goat ($AToT_G$), calculated as above but replacing daily wage with the price of a goat G;
- Relative terms of trade for both cereal vs. wage ($RToT_W$) and cereal vs. goat ($RToT_G$): these were computed as the ratio of AToT in month t to AToT in the first month of the series, i.e. as a relative change from the baseline; no weighting was applied in this case when averaging different cereals;
- Livelihood-specific relative terms of trade ($RToT_{liv}$), equal to the $RToT_G$ for strata or survey frames mainly populated by pastoralists, and $RToT_W$ otherwise.

AToT has the advantage of providing a direct measure of food security that can be compared across time and location, while RToT indicates relative food security stress over time within the same population (stratum or survey sampling frame).

So as to consider exposures of comparable duration to the recall period of the surveys (three months), we computed trimester averages of each terms of trade indicator, centred at varying lags from the mid-point of the survey's recall period.

Separately, we also extracted data collected and analysed by FAO/FSNAU on the price of a minimum basket of essential food and non-food items, adjusted for inflation (Cost of Minimum Expenditure Basket or CMB: see www.fsnau.org/downloads/FSNAU_CMB_CPI_for_Somalia.pdf for more detail on how this indicator is calculated). Because CMB was quantified for only some regions and with yearly or quarterly frequency only during 2007 and 2008, we did not use this indicator for statistical modelling. However, we included it in descriptive analysis of trends in different strata. Specifically, we computed a Consumer Price Index (CPI) by smoothing CMB time series as described above, and, using smoothed data, computing the CPI for any given month as the ratio CMB during that month to CMB during the first month of CMB data collection in a given region.

6.1.5 Armed conflict incidents

We collected data on individual armed conflict incidents in Somalia from January 2007 to July 2012 inclusive from the open access ACLED (Armed Conflict Location and Event Dataset) project (see http://www.acleddata.com/about-acled/). We settled on this data source as publicly available United Nations security bulletins were not issued with sufficient frequency over the time period of analysis, and did not present systematic data and variables on insecurity events.

After correcting geo-coding errors, we aggregated data by district and month in the time series (we removed <20 observations for which the district could not be determined). We ignored events labelled in the dataset as "Headquarters or base established", "Non-violent activity by a conflict actor" and "Riots/Protests", as they were unlikely to have major effects on mortality. We then computed event incidence rates (per 100,000 person-months) for each stratum or survey sampling frame, by summing events in the districts falling fully or partly within the stratum/survey frame, and dividing these by the UNDP June 2005 population estimates of the combined districts, adjusted upwards assuming a yearly growth rate of 3%. Note that these incidence rates are over-estimated because many districts are only partly contained within a stratum or survey sampling frame; however, we expect that such bias would project across all analysis units and thus have a limited impact on the analysis.

As above we also computed trimester event incidence rates with different lags.

6.1.6 Presence of humanitarian assistance

Based on humanitarian bulletins, we assumed that WFP, by far the largest provider of food and cash transfer relief in Somalia, was present everywhere between July 2006 and December 2009. Thereafter, the WFP did not have access to the southern sector of Somalia, with the exception of Banadir, for the remainder of the time series. As above, we determined whether WFP had been present (defined as presence during two or three months out of three) over different trimester periods.

Note that the WFP's presence, due to the agency's relative importance, was plausibly expected to be a major factor influencing the impact on food insecurity (see below). However, it is stressed that this variable is to a large extent a proxy of wider humanitarian access to the affected population, especially given that WFP's lack of access was followed by the interruption of activities of a large number of other agencies and a reduction in funding by major donors such as the United States [6].

6.1.7 Mortality reports from refugee camps

Table 12 and Table 13 show mortality reports available for Ethiopia and Kenya camps, respectively. Figure 17 shows trends in mortality and measles incidence in selected Ethiopia camps.

Table 12. Primary mortality reports covering Dollo Ado, Ethiopia refugee camps during the analysis period.

Author	Camp (opening date)	Population type	Data collection dates	Recall period	Methods	Results	Notes
Retrospective surveys							
UNHCR and partners [18]	Melkadida (Feb 2010)	Entire camp population (old and new arrivals).	28 Mar-2 Apr 2011	90 days	Systematic random sampling	CDR 1.47 (95%CI 0.96-2.23) U5DR 4.04 (95%CI 2.50-6.47)	Did not disaggregate data for new arrivals.
UNHCR and partners [18]	Bokolmanyo (Jan 2009)	Entire camp population (old and new arrivals).	22-27 Mar 2011	90 days	Systematic random sampling	CDR 0.78 (95%CI 0.43-1.44) U5DR 2.26 (95%CI 1.15-4.41)	Did not disaggregate data for new arrivals.
UNHCR and partners [24]	Kobe (24 June 2011)	New arrivals from Bay, Bakool, Gedo regions (unknown proportions).	Oct-Nov 2011	≈3mo (since 1 August 2011)	Systematic random sampling (empty shelters excluded a priori)	CDR 1.90 (95%CI 1.42-2.55) U5DR 5.95 (95%CI 4.40-8.00)	Individual household roster. All deaths and all person-time were within camp. Non-response <10%.
UNHCR and partners [24]	Hilaweyn (5 August 2011)	New arrivals from Bay, Bakool, Gedo regions (unknown proportions).	Oct-Nov 2011	≈3mo (since 1 August 2011)	Systematic random sampling (empty shelters excluded a priori)	CDR 1.35 (95%CI 0.96-1.89) U5DR 4.57 (95%CI 3.15-6.57)	Individual household roster. 9/33 (27.3%) deaths occurred before arrival to the camp (person-time before arrival not presented). Non-response <10%.
Prospective surveillance							
MSF/Epicentre [21]	Kobe	New arrivals from Bay, Bakool, Gedo regions (unknown proportions).	5 Aug 2011-19 Feb 2012	n/a	Weekly active shelter to shelter surveillance by home visitors, triangulated with grave counting	See Figure 17.	
Sorkin & Bernasconi [22]	Hilaweyn	New arrivals from Bay, Bakool, Gedo regions (unknown proportions).	22 Aug 2011-11 Dec 2011 (U5DR since 10 Oct 2011)	n/a	Weekly active shelter to shelter surveillance by home visitors, triangulated with grave counting	See Figure 17.	

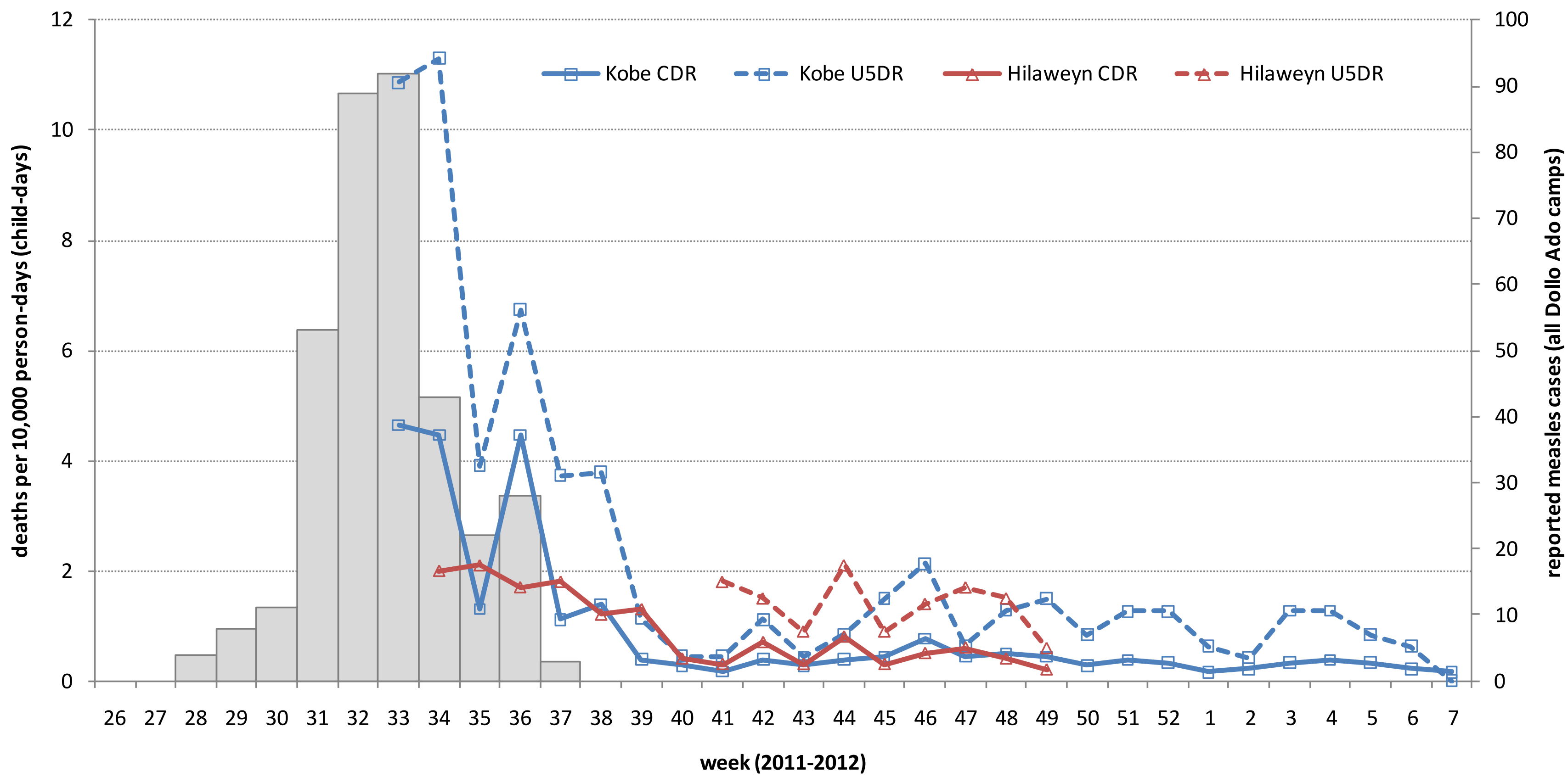

Figure 17. Trends in mortality and measles incidence in Dollo Ado, Ethiopia camps, 2011-2012. Source: Epicentre/Médecins Sans Frontières [21,22].

Table 13. Primary mortality reports covering Dadaab, Kenya refugee camps during the analysis period.

Author	Camp (opening date)	Population type	Data collection dates	Recall period	Methods	Results	Notes
Sackl [19]	Dahagaley (self-settled areas; August 2008)	New arrivals settling in camp outskirts (12% arrived since 1mo, 27% 1-3mo, 33% 4-6mo, 25% 7-12mo, 3% >12mo). Origin: 69% Lower Juba, 2% Middle Juba, 8% Banadir, 2% Lower and Middle Shabelle, 7% Bay, 6% Gedo, 6% other.	16-17 Jun 2011	365 days	Two-stage cluster sample survey; 8 clusters of 8-54 households (1 cluster per sector, number of households proportional to sector size); random walk to select cluster households.	CDR 0.41 U5DR 0.85	Low sample size (200 households, 8 clusters). Random walk not consistent with recommended method (heavily biased toward central households). Person-time not stratified between pre- and post-camp period. Standard errors not reported. Aggregate household roster.
Polonsky et al. [23]	Dahagaley (self-settled areas)	New arrivals (87% within previous 6mo) settling in camp outskirts. Origin: 26% L. Juba, 20% Bay, 19% M. Juba, 11% Gedo, 9% L. Shabelle, 5% Bakool, 5% Galgadud/Mudug, 3% Banadir, 2% other.	31 Jul-10 Aug 2011	167-177 days (since 15 Feb 2011)	Exhaustive survey.	CDR 1.0 U5DR 2.2 <3mo since arrival: CDR 1.5, U5DR 2.5 3-6mo: CDR 0.9, U5DR 2.2 >6mo: CDR 0.6, U5DR 1.4	Non-response <10%. Individual household roster.
Anonymous [20]	All Dadaab camps (1991)	New arrivals being registered (median 24 days since arrival)	24 Jul-3 Aug 2011	114-124 days (since 1 Apr 2011); median 24 days in post-arrival period	Systematic sampling during refugee registration (sampled every other household).	CDR 0.44 (95%CI 0.00-0.93) U5DR 1.53 (95%CI 0.00-3.25)	Non-response not reported. Survey representative of a small fraction of new arrivals (about 1500 households).
UNHCR [28]	Dahagaley (self-settled areas)	New arrivals (78% within previous 6mo). Origin of children 6-59mo old: 28% L. Juba, 26% M. Juba, 16% Gedo, 6% Bay, 2% L. Shabelle, 23% other.	Sep 2011	97 days (since end of May 2011)	Two-stage cluster sample survey (30 clusters of 15 households). Cluster starting points determined by random GPS coordinates.	CDR 1.23 (95%CI 0.73-2.06) U5DR 3.02 (95%CI 1.72-5.24)	Non-response <10%. Individual household roster. Person-time not stratified between pre- and post-camp period (about 10-30% pre-camp). North-east of camp excluded.
UNHCR [28]	Hagadera (1991)	Entire camp population (8% arrived within previous 6mo).	Aug 2011	86 days (since end of May 2011)	Two-stage cluster sample survey (35 clusters of 15 households); clusters allocated to blocks by PPS.	CDR 0.14 (95%CI 0.04-0.46) U5DR 0.33 (95%CI 0.08-1.36)	Non-response <10%. <5% person-time pre-arrival.
UNHCR [28]	Ifo (1991)	Entire camp population (4% arrived within previous 6mo).	Aug 2011	86 days (since end of May 2011)	As above	CDR 0.41 (95%CI 0.21-0.80) U5DR 0.94 (95%CI 0.45-1.98)	Non-response <10%. <2% person-time pre-arrival.
UNHCR [28]	Dahagaley formal camp blocks (1991)	Entire camp population (3% arrived within previous 6mo).	Sep 2011	97 days (since end of May 2011)	As above	CDR 0.14 (95%CI 0.05-0.36) U5DR 0.28 (95%CI 0.07-1.17)	Non-response <10%. <2% person-time pre-arrival.

6.2 Re-analysis of FAO/FSNAU mortality datasets

6.2.1 Data management

We obtained source mortality datasets for 82/205 (40.0%) of the surveys incorporating a mortality questionnaire carried out by FAO/FSNAU and partners between 1 Jan 2007 and 31 July 2012 and published on the FAO/FSNAU website. Remaining datasets were not stored centrally and could not be retrieved. The datasets obtained included 51/69 (73.9%) of the surveys done in between June 2010 and July 2012 inclusive in Somalia, i.e. within the person-time of interest for our analysis. We reanalysed all 82 datasets so as to verify the reported estimates and explore data quality.

Surveys employed different questionnaires (Table 14), reportedly as a function of the skills level of interviewers and security at the survey location. All datasets featured one row per household questionnaire. Variables entered included cluster number, household number and different aggregate totals of household numbers and different demographic events depending on the questionnaire (see Table 14). No information on which interviewers visited which household, on replacement of clusters and on survey attrition was available.

We calculated the total person-time in days for each household as shown in Table 14. As dates of entry or exit events were not reported, we assumed that each had occurred at the mid-point of the period and thus multiplied them by 0.5.

After inspecting frequency distributions of each variable, households were removed from the analysis if any of the following implausible values was encountered: $n_{all} = 0$; $n_{all} > 20$; $n_{u5} > 10$; $n_{u5} > n_{all}$; $j_{all} > 10$; $j_{u5} > 5$; $l_{all} > 10$; $l_{u5} > 5$; $b > 5$; $d_{all} > 10$; $d_{u5} > 5$; $d_{0\text{-}11mo} > 2$; $d_{12\text{-}59mo} > 3$; $d_{12\text{-}59mo} > 3$; $pt_{all} \leq 0$; $pt_{u5} \leq 0$. We also removed any household records with missing cluster, n_{all} , n_{u5} (except questionnaire type 1b), d_{all} and/or d_{u5} . Altogether, these criteria resulted in the exclusion of 1094/51,830 (2.1%) household records from the analysis.

6.2.2 Estimation

Two of the surveys collected an exhaustive sample, while the remainder had a self-weighting two-stage cluster design with allocation of clusters by probability proportional to size. We estimated all rates by fitting null Poisson models with the natural log of person-time as an offset, and robust standard errors for cluster designs, using R package `survey`. Point estimates $\hat{y}$ of each rate and their 95%CIs were computed as follows from model coefficients:

$$\hat{y} = e^{(\ln \hat{y})} \; y_{0.025} = e^{(\ln \hat{y} - 1.96SE\,(\ln \hat{y}))} \; y_{0.975} = e^{(\ln \hat{y} + 1.96SE\,(\ln \hat{y}))}$$

Rates were then expressed using the customary metric (per 10,000 person-days for CDR and U5DR in emergency contexts; per 1000 person-years for the crude birth rate).

During the 2011 emergency, some surveys in southern Somalia only collected the size of the household, with no breakdown of under 5y olds, so as to minimise interview time (type 1b): for this reason, U5DR estimates initially published for these surveys were subsequently retracted [58]. We kept these surveys in the analysis and imputed n_{u5} by simulation (see Table 14); we present the median point estimate and 95%CI of U5DR arising from 1000 randomly sampled values of n_{u5} .

Table 14. Types of questionnaire used by mortality surveys re-analysed.

Questionnaire type	Number of surveys	Information collected/entered	Household census	Person-time (days) calculation for each household† (recall period = 90 days for each survey)
1a	14	Number of people living within household on survey date; number under 5y Deaths during recall period, by age group (0-11mo, 12-59mo, older)	Aggregate	$pt_{\mathrm{all}} = n_{\mathrm{all}} \cdot 90$ $pt_{\mathrm{u5}} = (n_{\mathrm{u5}} + 0.5d_{\mathrm{u5}}) \cdot 90$‡
1b	18	Number of people living within household on survey date (no age breakdown) Deaths during recall period, by age group (0-11mo, 12-59mo, older)	Aggregate	$pt_{\mathrm{all}} = n_{\mathrm{all}} \cdot 90$ $pt_{\mathrm{u5}} = (n_{\mathrm{all}} p_{\mathrm{u5}} + 0.5d_{\mathrm{u5}}) \cdot 90$‡ where p_{u5}, the proportion under 5y, was sampled randomly from the empirical distribution of p_{u5} values in surveys done among populations with the same livelihood (IDP, urban, agro-pastoralist, etc.).
2a	3	Number of people living within household on survey date; number under 5y Deaths during recall period, by age group (under 5y, older)	Aggregate	$pt_{\mathrm{all}} = n_{\mathrm{all}} \cdot 90$ $pt_{\mathrm{u5}} = (n_{\mathrm{u5}} + 0.5d_{\mathrm{u5}}) \cdot 90$‡
2b	3	Number of people living within household on survey date; number under 5y Deaths during recall period, by age group (under 5y, older) Births during recall period	Aggregate	$pt_{\mathrm{all}} = (n_{\mathrm{all}} + 0.5d_{\mathrm{all}} - 0.5b_{\mathrm{all}}) \cdot 90$ $pt_{\mathrm{u5}} = (n_{\mathrm{u5}} + 0.5d_{\mathrm{u5}} - 0.5b_{\mathrm{u5}}) \cdot 90$
3a	37	Number of people living within household on survey date; number under 5y Deaths during recall period, by age group (under 5y, older) Births during recall period People that joined or left household during the recall period (under 5y, older)	Individual	$pt_{\mathrm{all}} = (n_{\mathrm{all}} + 0.5d_{\mathrm{all}} + 0.5l_{\mathrm{all}} - 0.5b_{\mathrm{all}} - 0.5j_{\mathrm{all}}) \cdot 90$ $pt_{\mathrm{u5}} = (n_{\mathrm{u5}} + 0.5d_{\mathrm{u5}} + 0.5l_{\mathrm{u5}} - 0.5b_{\mathrm{u5}} - 0.5j_{\mathrm{u5}}) \cdot 90$
3b	7	Number of people living within household on survey date; number under 5y Deaths during recall period, by age group (under 5y, older) Births during recall period People that joined or left household during the recall period (no age breakdown)	Individual	$pt_{\mathrm{all}} = (n_{\mathrm{all}} + 0.5d_{\mathrm{all}} + 0.5l_{\mathrm{all}} - 0.5b_{\mathrm{all}} - 0.5j_{\mathrm{all}}) \cdot 90$ $pt_{\mathrm{u5}} = (n_{\mathrm{u5}} + 0.5d_{\mathrm{u5}} - 0.5b_{\mathrm{u5}}) \cdot 90$

† pt = person-time; b = births; d = deaths; j = joined; l = left; n = number in household at survey date; all = all ages; u5 = under 5 y.

‡ deaths were added to ensure that households without any surviving children on the survey date but at least one child death during the recall period had a non-zero person-time and thus were included in the analysis.

6.2.3 Comparison with FAO/FSNAU results and quality assessment

Table 15 presents a comparison of the original FAO/FSNAU estimates and those arising from our reanalysis. For CDR, differences between the two were small (<20%) with the exception of 4/82 (4.9%) surveys. For U5DR, 8/82 (9.8%) sets of results were discrepant, while for a further 18 surveys the difference was explained by the fact that these surveys adopted the type 1b questionnaire above: our imputed proportion of children under 5y (empirical based on other surveys) was systematically higher than that used by FAO/FSNAU (about 20%).

Other findings were mostly unremarkable and suggested a consistent demographic pattern indicative of comparable data collection practices across FAO/FSNAU field researchers. The ratio of U5DR to CDR (median 2.3) and the high proportion of children among all decedents (median 0.61) and of infants among all child decedents (0.46) were consistent with a pre-epidemiologic transition setting with high child mortality. The crude birth rate estimates (median 42.2 per 1000 person-years) were likewise comparable with those of countries in the region over 2005-2010 (45.9 in Chad, 38.6 in Yemen, 38.0 in Kenya, 37.5 in Eritrea, 33.8 in Sudan, 33.3 in Ethiopia, 29.4 in Djibouti), particularly given that Somalia probably has a lower prevalence of family planning use http://data.un.org/-Data.aspx?d=PopDiv&f=variableID%3A53 . An implausibly high birth rate in mortality surveys may suggest over-reporting of household demographic events and lack of understanding of the recall period, which would also lead to upward bias of the death rate. An outstanding finding was the relatively high proportion of children under 5y old (0.25 compared to a more typical range of 0.17 to 0.23 in Sub-Saharan Africa), but this is possibly explainable through high rates of adult emigration to find work abroad.

Generally, datasets and information on survey methods and conduct were insufficient to carry out an in-depth quality assessment of mortality data. Information available does not indicate obvious problems with quality and bias, although the relative lack of documentation may itself be considered a source of concern [53].

6.3 Statistical models to indirectly estimate mortality

6.3.1 Conceptual frameworks

We firstly arranged all variables at our disposal into a conceptual framework (Figure 18) depicting the main plausible causal pathways and possible effect modifications (interactions) leading to mortality, recognising that some lie along the same pathway but at different levels (distal, intermediate or proximal). Since predictive models would only be useful if built upon variables that were available even where/when mortality data were not collected, the framework for predictive modelling reduced into a simpler version (Figure 19) that obviated the need for multi-level or structural modelling.

6.3.2 Eligibility criteria

Surveys were included in predictive and baseline models if they had a non-zero quality weight (see below), i.e. no evidence of a critical methodological flaw, and if at least the point estimate and 95%CIs (exhaustive surveys exempted) of both CDR and U5DR were reported.

Table 15. Results of re-analysis of FAO/FSNAU mortality datasets. All rates are per 10,000 person-days except for the crude birth rate (per 1000 person-years). Markedly discrepant results are highlighted in red. All other findings reported are from reanalysis.

Survey characteristics				Crude death rate (95%CI)		Under 5y death rate (95%CI)		Proportion of children under 5y old among all deaths (infants among deaths under 5y old)	Crude birth rate	In-migration rate	Out-migration rate	Mean household size (proportion under 5y old)
Year	Month	Location	Livelihood	FAO/FSNAU	Reanalysis	FAO/FSNAU	Reanalysis					
2007	3	Somalia	urban	0.65 (0.42-0.80)	0.64 (0.35-1.17)	2.19 (1.35-3.01)	1.47 (0.81-2.68)	0.58 (n/a)	25.79	0.10	1.64	5.96 (0.26)
2007	4	Somalia	agropastoral	1.53 (0.75-2.30)	1.53 (0.91-2.58)	2.65 (1.11-4.18)	2.64 (1.57-4.47)	0.48 (n/a)	54.29	1.75	2.86	5.21 (0.28)
2007	4	Somalia	pastoral	1.05 (0.51-1.60)	1.05 (0.66-1.68)	1.48 (0.22-3.18)	1.49 (0.60-3.66)	0.41 (n/a)	50.20	0.34	1.15	5.37 (0.30)
2007	4	Somalia	riverine	1.09 (0.50-1.67)	1.37 (0.82-2.26)	1.63 (0.68-2.57)	1.64 (0.92-2.92)	0.36 (n/a)	45.32	55.41	2.92	5.13 (0.25)
2007	5	Somalia	agropastoral	1.88 (0.90-2.90)	1.31 (0.94-1.84)	1.84 (0.56-3.10)	1.86 (1.11-3.10)	0.39 (n/a)	37.88	2.58	5.1	5.30 (0.28)
2008	11	Somalia	IDP	0.70 (0.37-1.34)	0.70 (0.38-1.30)	1.69 (0.90-3.17)	1.70 (0.93-3.09)	0.50 (n/a)	39.37	2.33	3.34	6.21 (0.21)
2008	11	Somalia	riverine	1.01 (0.66-1.55)	1.02 (0.68-1.53)	2.15 (1.17-3.94)	2.16 (1.21-3.84)	0.50 (n/a)	33.04	1.24	2.89	5.64 (0.24)
2009	5	Somalia	pastoral	0.55 (0.31-0.98)	0.57 (0.33-0.98)	1.36 (0.67-2.78)	1.25 (0.59-2.66)	0.61 (n/a)	54.18	3.35	10.52	5.33 (0.26)
2009	5	Somaliland	pastoral	0.35 (0.10-1.18)	0.35 (0.11-1.08)	1.06 (0.36-3.08)	1.03 (0.39-2.75)	0.75 (n/a)	54.85	0.61	3.34	4.33 (0.25)
2009	5	Somalia	pastoral	0.71 (0.40-1.25)	0.71 (0.41-1.22)	0.92 (0.48-1.77)	0.93 (0.50-1.71)	0.35 (n/a)	63.58	2.20	6.97	5.30 (0.28)
2009	5	Somaliland	pastoral	0.46 (0.27-0.79)	0.46 (0.28-0.77)	0.82 (0.34-1.99)	0.82 (0.35-1.92)	0.33 (n/a)	28.11	0.85	4.49	5.67 (0.19)
2009	7	Somalia	agropastoral	1.17 (0.75-1.81)	1.18 (0.77-1.79)	3.82 (2.20-6.58)	3.77 (2.23-6.38)	0.73 (n/a)	39.09	1.10	2.82	5.39 (0.22)
2009	7	Somalia	riverine	0.81 (0.44-1.51)	0.82 (0.45-1.48)	2.90 (1.67-5.00)	2.64 (1.55-4.47)	0.68 (n/a)	30.93	3.63	4.98	5.34 (0.19)
2009	7	Somalia	riverine	1.19 (0.87-1.63)	1.16 (0.84-1.59)	2.62 (1.71-4.00)	2.37 (1.5-3.74)	0.62 (n/a)	65.90	2.42	2.85	4.78 (0.29)
2009	7	Somalia	agropastoral	0.17 (0.06-0.48)	0.18 (0.07-0.48)	0.37 (0.08-1.59)	0.38 (0.09-1.58)	0.86 (n/a)	42.21	2.69	5.71	4.97 (0.25)
2009	7	Somalia	pastoral	0.80 (0.45-1.43)	0.48 (0.27-0.84)	2.19 (0.99-4.81)	1.70 (0.64-4.50)	0.60 (n/a)	49.20	2.09	2.45	4.67 (0.28)
2010	3	Somaliland	agropastoral	0.97 (exhaustive)	0.83 (exhaustive)	2.22 (exhaustive)	1.50 (exhaustive)	0.36 (n/a)	43.33	0.53	4.99	5.30 (0.20)
2010	6	Somalia	pastoral	0.52 (0.30-0.83)	0.52 (0.31-0.87)	1.74 (0.88-3.41)	1.74 (0.91-3.34)	0.67 (n/a)	45.32	0.55	6.58	5.12 (0.20)
2010	11	Somalia	pastoral	0.71 (0.41-1.21)	0.71 (0.43-1.18)	1.91 (0.93-1.87)	1.92 (0.97-3.80)	0.46 (n/a)	61.64	1.75	7.23	5.12 (0.18)
2010	11	Somaliland	pastoral	0.61 (0.20-1.87)	0.54 (0.33-0.89)	0.53 (0.32-0.90)	0.41 (0.11-1.56)	0.23 (n/a)	42.41	1.45	7.68	5.21 (0.21)
2010	11	Somaliland	pastoral	0.45 (0.25-0.80)	0.45 (0.25-0.79)	1.10 (0.49-2.47)	1.10 (0.51-2.41)	0.54 (n/a)	33.95	0.59	5.34	5.36 (0.22)
2010	12	Somalia	riverine	0.22 (0.11-0.46)	0.23 (0.11-0.46)	1.69 (0.78-3.61)	1.73 (0.84-3.60)	1.62 (n/a)	27.27	0.80	1.44	5.31 (0.22)
2011	5	Somalia	IDP	0.89 (0.59-1.36)	0.89 (0.60-1.33)	1.01 (0.47-2.17)	0.68 (0.27-1.66)	0.18 (n/a)	41.45	1.18	4.06	4.48 (0.25)
2011	5	Puntland	IDP	0.89 (0.58-1.37)	0.88 (0.59-1.32)	2.23 (1.20-4.11)	2.03 (1.08-3.84)	0.61 (n/a)	39.09	1.45	5.09	4.92 (0.25)
2011	5	Puntland	IDP	0.61 (exhaustive)	0.61 (exhaustive)	1.39 (exhaustive)	1.39 (exhaustive)	0.38 (n/a)	24.94	0.91	22.78	5.43 (0.17)
2011	5	Somaliland	IDP	0.56 (0.30-0.91)	0.56 (0.36-0.90)	2.02 (1.21-3.34)	2.02 (1.25-3.28)	0.79 (n/a)	36.90	0.68	2.62	5.16 (0.22)
2011	5	Somaliland	IDP	0.28 (0.15-0.54)	0.28 (0.15-0.54)	0.31 (0.09-1.13)	0.31 (0.08-1.24)	0.22 (n/a)	22.88	0.13	2.70	5.66 (0.20)
2011	5	Somaliland	IDP	0.37 (0.19-0.73)	0.38 (0.20-0.72)	0.59 (0.22-1.55)	0.58 (0.23-1.48)	0.31 (n/a)	26.55	0.49	2.73	5.73 (0.20)
2011	7	Somalia	agropastoral	1.09 (0.74-1.45)	1.10 (0.81-1.50)	4.04 (2.31-5.78)	2.93 (1.93-4.44)	0.75 (0.48)	n/a	n/a	n/a	5.91 (n/a)
2011	7	Somalia	agropastoral	1.67 (1.14-2.19)	1.68 (1.25-2.27)	5.29 (3.61-6.96)	3.88 (2.85-5.29)	0.65 (0.58)	n/a	n/a	n/a	5.99 (n/a)
2011	7	Somalia	agropastoral	1.50 (0.90-2.10)	1.54 (1.21-1.95)	4.24 (2.83-5.65)	3.07 (2.22-4.26)	0.56 (0.19)	n/a	n/a	n/a	5.30 (n/a)
2011	7	Somalia	agropastoral	4.21 (2.89-5.53)	4.29 (3.16-5.82)	12.48 (9.14-15.81)	9.39 (7.17-12.31)	0.62 (0.44)	n/a	n/a	n/a	6.37 (n/a)
2011	7	Somalia	agropastoral	1.13 (0.51-1.75)	1.14 (0.86-1.51)	4.20 (3.06-5.33)	3.01 (2.20-4.13)	0.74 (0.34)	n/a	n/a	n/a	5.43 (n/a)
2011	7	Somalia	agropastoral	2.28 (1.71-2.86)	2.31 (1.81-2.94)	6.84 (4.91-8.76)	5.04 (3.82-6.65)	0.61 (0.44)	n/a	n/a	n/a	7.08 (n/a)
2011	7	Somalia	IDP	4.29 (3.22-5.36)	4.37 (3.42-5.59)	14.09 (10.65-17.53)	12.21 (9.49-15.69)	0.69 (0.39)	n/a	n/a	n/a	5.98 (n/a)
2011	7	Somalia	IDP	4.24 (3.17-5.31)	4.33 (3.38-5.54)	12.47 (9.56-15.38)	10.72 (8.44-13.60)	0.61 (0.42)	n/a	n/a	n/a	6.37 (n/a)
2011	7	Somalia	pastoral	0.56 (0.34-0.91)	0.56 (0.35-0.90)	1.92 (1.02-3.59)	1.74 (0.97-3.13)	0.69 (n/a)	102.61	2.35	4.60	5.21 (0.21)
2011	7	Somalia	pastoral	1.89 (1.60-2.19)	1.91 (1.65-2.22)	5.06 (3.80-6.32)	3.96 (3.10-5.06)	0.54 (0.51)	n/a	n/a	n/a	5.09 (n/a)
2011	7	Somalia	pastoral	1.18 (0.79-1.57)	1.25 (0.90-1.73)	4.06 (2.47-5.64)	3.37 (2.28-4.98)	0.70 (0.65)	n/a	n/a	n/a	6.43 (n/a)
2011	7	Somalia	pastoral	0.91 (0.60-1.37)	0.91 (0.61-1.34)	2.12 (1.19-3.77)	1.96 (1.13-3.39)	0.52 (n/a)	39.75	0.91	3.85	4.75 (0.23)
2011	7	Somalia	pastoral	1.53 (1.07-1.98)	1.54 (1.15-2.05)	3.67 (2.38-4.97)	2.85 (2.03-4.02)	0.49 (0.18)	n/a	n/a	n/a	4.51 (n/a)
2011	7	Somalia	pastoral	1.25 (0.68-1.81)	1.24 (0.90-1.73)	4.33 (3.23-5.43)	3.31 (2.35-4.68)	0.70 (0.59)	n/a	n/a	n/a	5.26 (n/a)
2011	7	Somalia	riverine	1.62 (1.00-2.25)	1.68 (1.04-2.70)	6.20 (4.21-8.19)	5.16 (3.05-8.72)	0.76 (0.65)	n/a	n/a	n/a	5.85 (n/a)

Survey characteristics				Crude death rate (95%CI)		Under 5y death rate (95%CI)		Proportion of children under 5y old among all deaths (infants among deaths under 5y old)	Crude birth rate	In-migration rate	Out-migration rate	Mean household size (proportion under 5y old)
Year	Month	Location	Livelihood	FAO/FSNAU	Reanalysis	FAO/FSNAU	Reanalysis					
2011	7	Somalia	riverine	1.37 (1.04-1.71)	1.38 (1.09-1.75)	4.13 (2.95-5.31)	3.16 (2.39-4.18)	0.61 (0.04)	n/a	n/a	n/a	5.64 (n/a)
2011	7	Somalia	riverine	5.93 (4.28-7.57)	6.08 (4.64-7.99)	18.64 (13.05-24.22)	16.38 (11.99-22.38)	0.67 (0.46)	n/a	n/a	n/a	6.46 (n/a)
2011	7	Somalia	riverine	1.18 (0.50-1.82)	1.15 (0.87-1.51)	4.76 (3.38-6.14)	3.43 (2.55-4.62)	0.79 (0.52)	n/a	n/a	n/a	5.28 (n/a)
2011	7	Somalia	riverine	1.71 (1.10-3.20)	1.52 (1.11-2.09)	5.19 (2.96-7.41)	3.99 (2.62-6.07)	0.70 (0.41)	n/a	n/a	n/a	6.74 (n/a)
2011	7	Somalia	agropastoral	2.20 (1.70-2.70)	1.98 (1.70-2.30)	7.00 (5.20-8.80)	4.98 (4.13-6.00)	0.71 (0.57)	n/a	n/a	n/a	5.01 (n/a)
2011	8	Somalia	agropastoral	2.15 (1.50-2.80)	2.17 (1.64-2.88)	6.16 (3.91-8.40)	6.16 (4.42-8.57)	0.69 (0.76)	n/a	n/a	n/a	5.19 (0.24)
2011	8	Somalia	agropastoral	1.82 (1.23-2.41)	1.84 (1.36-2.49)	4.6 (2.87-6.37)	4.62 (3.28-6.50)	0.73 (0.44)	n/a	n/a	n/a	4.60 (0.28)
2011	8	Somalia	IDP	5.68 (4.48-6.88)	5.83 (4.73-7.20)	15.43 (11.40-19.50)	15.43 (11.92-19.96)	0.78 (0.35)	n/a	n/a	n/a	5.53 (0.28)
2011	8	Somalia	IDP	4.02 (3.05-4.99)	4.10 (3.22-5.21)	10.30 (7.57-13.00)	10.30 (7.95-13.34)	0.65 (0.49)	n/a	n/a	n/a	6.19 (0.25)
2011	8	Somalia	pastoral	1.25 (0.77-1.73)	1.25 (0.84-1.86)	2.61 (1.32-3.88)	2.26 (1.33-3.85)	0.62 (0.46)	n/a	n/a	n/a	5.12 (0.34)
2011	8	Somalia	riverine	1.76 (1.20-2.32)	1.75 (1.32-2.32)	4.22 (2.56-5.87)	4.16 (3.09-5.61)	0.66 (0.36)	n/a	n/a	n/a	4.76 (0.27)
2011	10	Somalia	agropastoral	1.37 (0.88-1.86)	1.38 (1.04-1.82)	3.43 (2.00-4.87)	3.43 (2.45-4.81)	0.73 (0.45)	n/a	n/a	n/a	4.98 (0.29)
2011	10	Somalia	IDP	1.78 (1.19-2.36)	1.79 (1.30-2.47)	5.00 (3.00-7.00)	5.00 (3.39-7.39)	0.70 (0.45)	n/a	n/a	n/a	6.07 (0.25)
2011	10	Somalia	IDP	2.30 (1.60-3.00)	2.33 (1.82-2.97)	4.76 (3.08-6.44)	4.76 (3.63-6.25)	0.76 (0.52)	n/a	n/a	n/a	4.32 (0.36)
2011	10	Somalia	pastoral	0.93 (0.52-1.34)	0.94 (0.65-1.35)	2.76 (1.41-4.12)	2.77 (1.82-4.23)	0.80 (0.56)	n/a	n/a	n/a	4.84 (0.27)
2011	10	Somalia	riverine	1.54 (1.02-2.06)	1.55 (1.22-1.97)	4.12 (2.53-5.71)	4.14 (3.02-5.69)	0.76 (0.54)	n/a	n/a	n/a	5.03 (0.28)
2011	10	Somalia	urban	2.81 (1.97-3.64)	2.84 (2.12-3.82)	7.49 (4.83-10.20)	7.49 (5.30-10.61)	0.64 (0.55)	n/a	n/a	n/a	6.63 (0.24)
2011	11	Somalia	IDP	0.80 (0.45-1.42)	0.80 (0.46-1.39)	1.39 (0.62-2.08)	1.38 (0.64-2.99)	0.45 (n/a)	52.19	1.35	4.43	5.23 (0.27)
2011	11	Puntland	IDP	0.68 (0.44-1.07)	0.69 (0.44-1.06)	1.51 (0.84-2.72)	1.51 (0.86-2.66)	0.56 (n/a)	48.03	0.33	3.48	4.98 (0.26)
2011	11	Puntland	IDP	0.30 (0.15-0.59)	0.30 (0.16-0.57)	0.77 (0.31-1.88)	0.76 (0.32-1.77)	0.60 (n/a)	63.64	0.06	1.08	4.79 (0.24)
2011	11	Somaliland	IDP	0.31 (0.18-0.54)	0.32 (0.19-0.54)	0.14 (0.02-1.11)	0.14 (0.02-1.02)	0.09 (n/a)	11.50	0.37	2.75	5.78 (0.20)
2011	11	Somaliland	IDP	0.38 (0.20-0.71)	0.36 (0.19-0.69)	0.44 (0.14-1.40)	0.3 (0.08-1.19)	0.17 (n/a)	32.13	0.06	1.82	6.08 (0.20)
2011	12	Somalia	IDP	2.06 (1.60-2.66)	2.08 (1.62-2.67)	5.46 (3.95-7.51)	5.47 (4.00-7.48)	0.66 (0.58)	n/a	n/a	n/a	6.36 (0.25)
2011	12	Somalia	pastoral	0.32 (0.16-0.66)	0.33 (0.17-0.65)	0.87 (0.31-2.42)	0.89 (0.34-2.37)	0.56 (n/a)	34.96	2.36	5.75	5.63 (0.21)
2011	12	Puntland	pastoral	0.26 (0.10-0.65)	0.30 (0.12-0.77)	0.36 (0.04-2.82)	0.56 (0.08-3.99)	0.38 (n/a)	30.37	0.23	6.58	4.62 (0.21)
2011	12	Puntland	pastoral	0.19 (0.08-0.43)	0.19 (0.09-0.43)	0.77 (0.28-2.08)	0.78 (0.30-2.03)	1.00 (n/a)	25.32	0.12	4.16	4.58 (0.25)
2011	12	Somalia	pastoral	0.49 (0.28-0.84)	0.49 (0.29-0.82)	0.86 (0.30-2.41)	0.86 (0.32-2.33)	0.38 (n/a)	52.05	1.16	5.33	5.41 (0.22)
2011	12	Somalia	urban	1.33 (0.88-2.02)	1.34 (0.89-2.02)	4.12 (2.62-6.43)	4.12 (2.66-6.39)	0.83 (0.72)	n/a	n/a	n/a	6.23 (0.27)
2012	4	Somalia	IDP	1.42 (1.05-1.92)	1.42 (1.06-1.91)	2.80 (1.87-4.17)	2.80 (1.89-4.14)	0.57 (n/a)	30.47	4.08	6.31	5.97 (0.29)
2012	4	Somalia	urban	1.22 (0.88-1.69)	1.16 (0.84-1.60)	1.06 (0.56-2.00)	0.97 (0.50-1.88)	0.22 (n/a)	55.20	5.18	5.73	5.95 (0.28)
2012	5	Somaliland	IDP	0.50 (0.28-0.88)	0.51 (0.29-0.87)	1.01 (0.36-2.80)	1.01 (0.38-2.66)	0.42 (n/a)	21.53	0.17	2.91	5.55 (0.21)
2012	6	Somalia	pastoral	0.31 (0.15-0.61)	0.27 (0.13-0.56)	0.86 (0.43-1.73)	0.74 (0.35-1.56)	0.75 (n/a)	42.19	n/a	n/a	5.84 (0.27)
2012	6	Somalia	pastoral	0.59 (0.35-1.01)	0.59 (0.36-0.98)	1.36 (0.77-2.40)	1.31 (0.76-2.24)	0.55 (n/a)	89.41	n/a	n/a	6.26 (0.25)
2012	6	Somalia	riverine	0.78 (0.43-1.43)	0.65 (0.33-1.25)	2.34 (1.22-4.43)	2.28 (1.22-4.24)	0.95 (n/a)	60.85	n/a	n/a	5.96 (0.27)
2012	7	Somalia	agropastoral	1.40 (0.93-2.10)	1.41 (0.95-2.09)	2.70 (1.90-3.89)	2.70 (1.90-3.84)	0.56 (n/a)	n/a	n/a	n/a	5.48 (0.29)
2012	7	Somalia	IDP	1.41 (0.99-2.02)	1.41 (1.00-2.00)	2.81 (1.82-4.33)	2.81 (1.84-4.30)	0.51 (n/a)	44.60	4.24	5.38	5.56 (0.26)
2012	7	Somalia	IDP	0.42 (0.27-0.66)	0.42 (0.28-0.65)	1.52 (0.91-2.53)	1.52 (0.93-2.49)	0.88 (n/a)	n/a	n/a	n/a	5.52 (0.24)
2012	7	Somalia	pastoral	0.44 (0.20-0.99)	0.40 (0.18-0.89)	0.81 (0.29-2.27)	0.66 (0.25-1.70)	0.44 (n/a)	n/a	n/a	n/a	5.65 (0.27)
2012	7	Somalia	urban	1.23 (0.81-1.83)	1.23 (0.83-1.84)	1.54 (0.82-2.85)	1.53 (0.84-2.81)	0.34 (n/a)	49.69	3.23	3.28	5.85 (0.28)

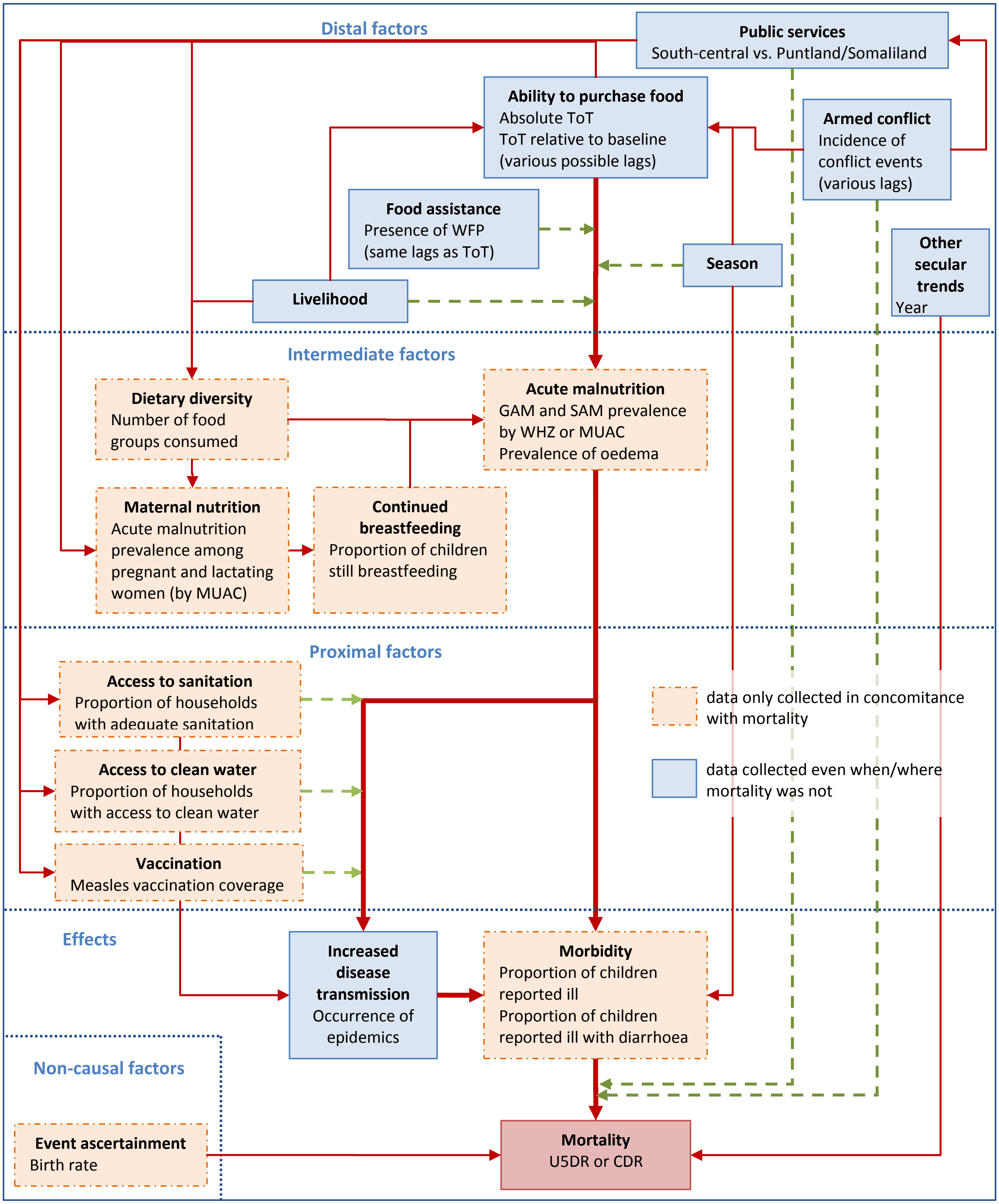

Figure 18. Conceptual framework of factors predictive of mortality used to guide model specification. Only variables available for analysis (listed under each box heading) and causal links deemed relevant for estimation are included.

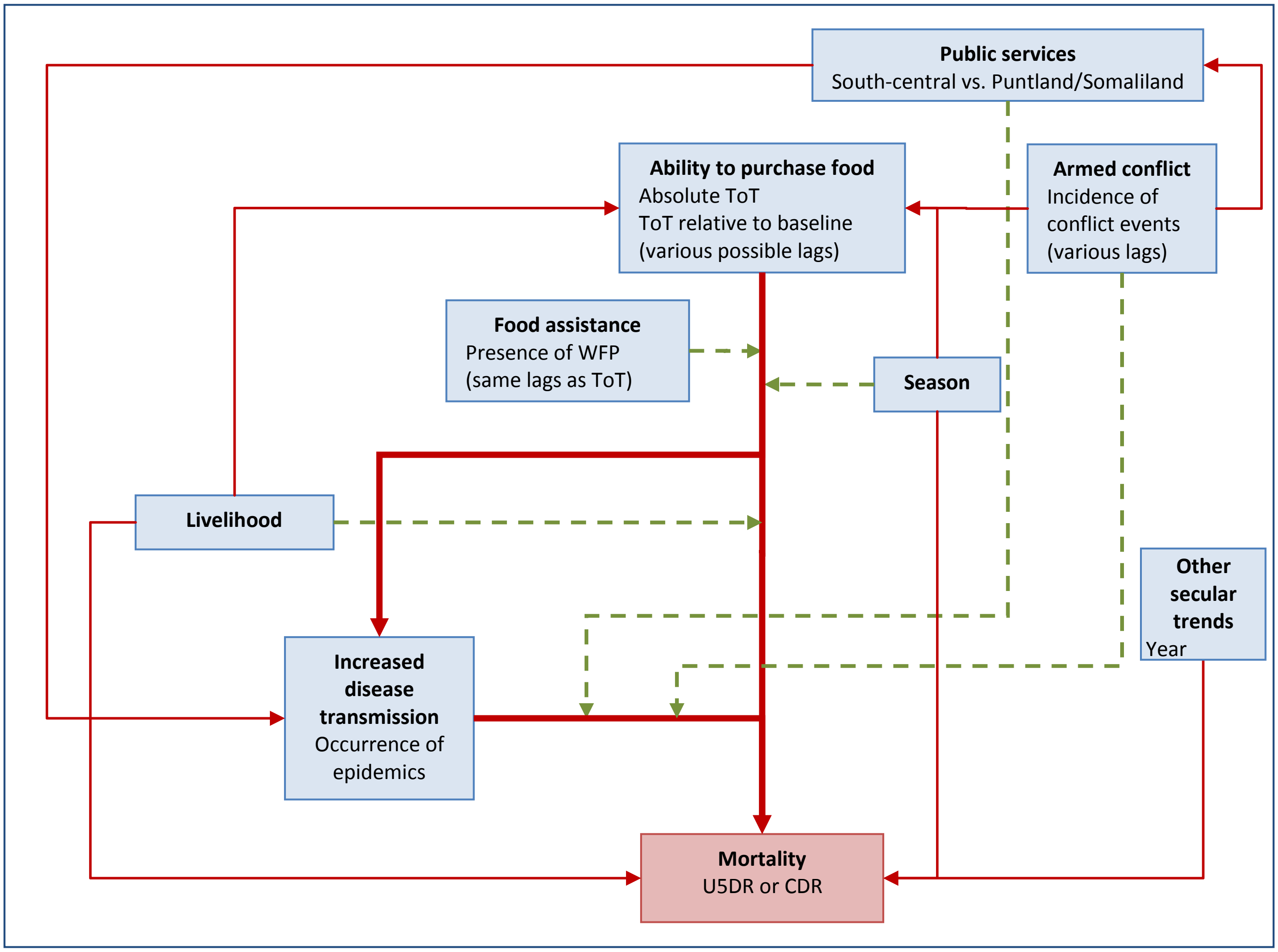

Figure 19. Reduced conceptual framework of factors predictive of mortality used to guide predictive model specification. Only variables available for analysis and only links deemed relevant for estimation purposes are included.

6.3.3 Univariate exploration

The relationship between CDR, U5DR and all the variables included in Figure 19 was explored graphically and through simple univariate linear regression, with untransformed death rate or its natural log as the dependent variable. We used this analysis to identify variables to be included in further statistical modelling and construct appropriate categories for each. Where relevant we considered various possible lags between the predictor and mortality.

Notable decisions taken on the basis of this univariate exploration included the following:

- The location variable (proxy of public services) was collapsed into southern and central Somalia versus Somaliland/Puntland.
- Armed conflict event incidence rate was made categorical so as to avoid non-linear relationships and excess of zeroes, and a lag of 3mo appeared to have better fit than 0mo, which is plausible given that armed conflict attacks can result in medical supply and staffing

shortages and other disruptions in services that may lead to delayed mortality (due to drugs running out, gradual increases in malnutrition or infectious disease transmission, etc.).

- While data on survey month treated individually were too sparse, regrouping months into various categories reflecting plausible seasonality as shown in FEWS NET calendars (http://www.fews.net/pages/timelineview.aspx?loc=1&gb=so&l=en; in particular, a hunger season in February-April; a post-hunger season in May-July; a lean season in August-October; and a post-lean season in November-January) did not show a strong fit with mortality. There was no graphical evidence of seasonality in CDR or U5DR over the five year time series of observations (data not shown).
- Epidemic occurrence was categorised simply as no ongoing outbreak during the stratum-month in question, 1 ongoing outbreak or >1 ongoing outbreak.
- AToT had greater correlation with mortality than RToT, and $AToT_W$ was more predictive than $AToT_G$. A 2mo lag showed the strongest correlation, with a clear threshold effect of steeply increasing mortality below 20,000 Kcal/daily wage (see Figure 20). To avoid non-linearity the variable was split into five categories.
- While humanitarian presence showed a high correlation with mortality using different lags from 0mo to 9mo, the lag chosen for statistical modelling was set at 2mo as for AToT (see above), since a statistical interaction between the two was considered plausible and supported by exploratory analysis (for such an interaction to make sense theoretically, it seemed reasonable to consider humanitarian presence and AToT values referring to the same time period).
- Log-transformed mortality was preferable as it showed more linear correlations with predictors.

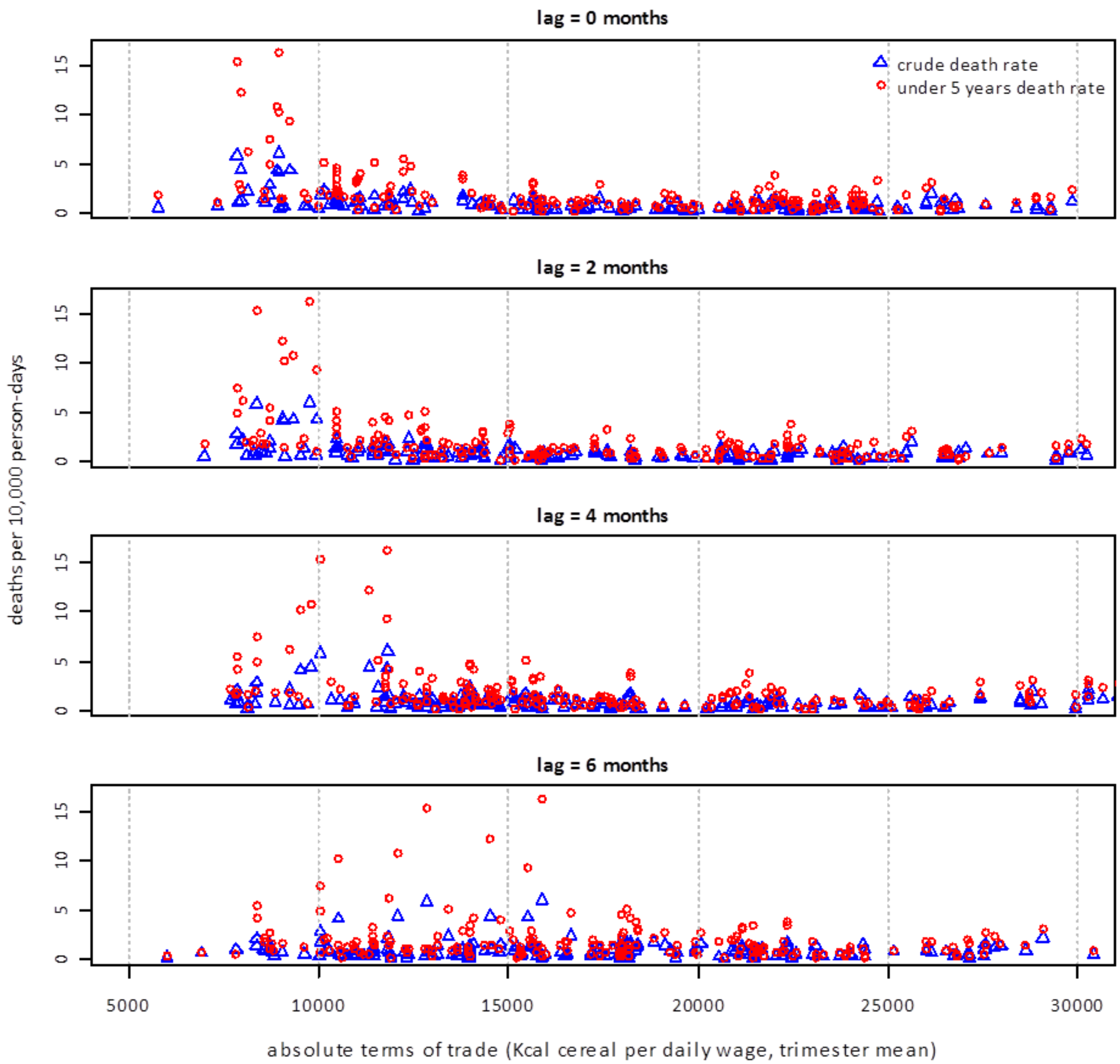

Figure 20. Graphical relationship between mortality and absolute terms of trade (Kcal cereal per daily wage), for different lags between terms of trade and the survey recall period.

6.3.4 Model fitting

Distributional assumptions

Mortality data are generally modelled through generalised linear models assuming discrete distributions such as the Poisson, quasi-Poisson or negative binomial; data generated through cluster surveys are further analysed using robust standard errors or random effects to account for intra-cluster homogeneity (design effect). In our case, however, we could not rely on the above techniques as we did not have source datasets for all the eligible mortality surveys, and instead had to use metadata, namely the reported death rate and its 95% confidence interval. We thus opted for ordinary least squares (OLS, linear) regression, in which the point estimates of each survey, rather than household-level data, were taken as the basic unit of analysis (we also attempted generalised linear models assuming a gamma distribution, but these yielded worse fits and were discarded; data not shown).

Because we did not have all source datasets, we were forced to ignore survey cluster in the analysis, which may lead to an underestimation of the true variance. Our data had an additional clustered structure as they mostly arose from repeated though independent survey samples of given livelihood zones or combinations thereof. To account for this, we grouped observations into 45 groups, namely the typical livelihood zones across Somalia within which surveys were repeatedly done (this grouping was somewhat imperfect: in some years surveys covered the entire livelihood zone whereas in others they covered only some of the zone; moreover, one survey was usually done covering Middle and Lower Juba, and Middle and Lower Shabelle respectively, but in 2011 region-specific surveys were done; in all such cases we maintained the geographically largest group, i.e. kept the number of groups to a minimum). We accounted for the data's longitudinal structure by fitting mixed models with group as the grouping random effect variable (see below).

Transformations

As the dependent variables (CDR and U5DR) were heavily right-skewed with a few extreme outliers, we applied a Box-Cox transformation (of which the natural log transformation is a special case) to both:

$$y_{i,\lambda} = \frac{y_i^{\lambda} - 1}{\lambda(GM(y))^{\lambda-1}}$$

where $y_{i,\lambda}$ are transformed observations, y are the original data, $GM(y)$is the geometric mean of y, and λ (= 0.4 in our case, yielding the most normal transformations) is a constant that we estimated by maximum likelihood on the full fixed effects model using R function `boxcox`.

Weighting

We attributed to each survey a composite weight $W_1 = W_q W_a$, where

- W_q is a semi-quantitative quality weight (range 0-1) based on information about the survey presented in FAO/FSNAU reports. W_q was 1 for 177 surveys since no obvious quality problems or sources of bias were identified for these. For remaining surveys, W_q was arbitrarily multiplied by 0.5 for each of the following quality/bias issues identified: (i) the reported number of deaths did not match with the CDR or U5DR point estimate (three surveys); (ii) anthropometric data did not meet all ENA software quality checks (five surveys); and (iii) a possibly non self-weighting sampling design was reported (one survey). Accordingly, a survey with two quality issues identified would have had its quality weight cut to one fourth (0.5 x 0.5), etc.

  So as to partially account for selection bias arising from incomplete sampling of the survey's target population, W_q was further multiplied by the approximate proportion of the nominal sampling universe that was actually included in the sampling frame (this was mostly 1, but in 21 cases the reported sampling frame was smaller than the universe due to insecurity or unspecified reasons): this proportion was approximated either using UNDP 2005 population figures or, for Mogadishu, as the proportion of urban districts included in the sampling frame; for three surveys insufficient information on the sampling frame was provided and the proportion was assumed to be 0.5. As an example, a survey with a discordant number of deaths and CDR estimate and which sampled only about 60% of the sampling universe would have had a $W_q = 0.5 \cdot\ 0.6 = 0.3$.

Lastly, W_q was set to 0 (i.e. the survey was omitted from analysis) if data fabrication was suspected by FAO/FSNAU supervisors (two surveys, plus a further three that were not included in our database as their publication was withdrawn by FAO/FSNAU).

- $W_a = \frac{1}{(SE(\ln \hat{y}))^2}$, where $\hat{y}$ is the CDR or U5DR point estimate. This analytic weight, equal to the inverse of the variance of the natural log of the death rate (where variance is the standard error squared), penalised surveys according to their relative level of imprecision.

  For most surveys, calculating $SE(\ln \hat{y})$ was straightforward. The ENA software used by FAO/FSNAU to analyse mortality data issues 95%CIs consistent with those arising from discrete distributions (negative binomial or Poisson). Accordingly, $y_{0.975} = e^{(\ln \hat{y} + 1.96SE\,(\ln \hat{y}))}$ and $y_{0.025} = e^{(\ln \hat{y} - 1.96SE\,(\ln \hat{y}))}$. After rearrangement, $(\ln \hat{y}) = \frac{\ln y_{0.975} - \ln \hat{y}}{1.96}$.

  Some surveys (21 for CDR and 28 for U5DR), all done in Puntland or Somaliland and mostly before 2009, presented implausible 95%CIs (i.e. heavily asymmetric when log-transformed), probably due to systematic mistakes in confidence interval estimation before the adoption of ENA software:

  - For some of these surveys (15 for CDR, 18 for U5DR), meta-data on either the number of deaths d or the number of people N sampled were available (the latter was used to deduce d from $\hat{y}$ and the recall period of 90 days: $d = 90N\hat{y}$). If simple random sampling (SRS) and a Poisson distribution of deaths are assumed, $SE\,(\ln \hat{y})_{SRS} = \frac{1}{\sqrt{d}}$. We multiplied this quantity by an adjustment constant K to account for design effect. K was estimated as the mean ratio $\frac{SE\,(\ln \hat{y})}{SE\,(\ln \hat{y})_{SRS}} = SE\,(\ln \hat{y})\sqrt{d}$ among cluster surveys for which we had source datasets.
  - For the remaining surveys (6 for CDR, 10 for U5DR), we first imputed d using mean characteristics of the set of source datasets we reanalysed (678 households per survey, 5.4 people per household, proportion under 5 years = 0.25), and then proceeded as above.

  Eleven surveys collected an exhaustive sample of the population. While these surveys' estimates have variance 0, to avoid infinite weights we treated these as SRS samples from a very large population. If d or N were reported (seven surveys), we proceeded as above but without a K adjustment. For the remaining four surveys, we imputed d as above but assuming 450 households (mean sample size in the other exhaustive surveys).

  Note that we would have preferred to use the variance of the untransformed death rate for weighting, but this was not calculable for most surveys as source datasets were unavailable and/or meta-data on the number of deaths or people sampled were not reported.

We specified W_1 in all models. We also fitted models without weights or with an alternative, less extreme analytic weight ($W_2 = W_q \frac{1}{SE(\ln \hat{y})}$), but model performance was mostly worse (data not shown).

Model selection

We built both CDR and U5DR predictive models using the following sequence of steps:

1. Fit the full fixed effects model containing the short-list of variables identified during univariate analysis: survey location, rate of armed conflict incidents (lag = 3mo), livelihood, year, occurrence of epidemics, humanitarian presence (lag = 2mo) and absolute terms of trade (lag = 2mo).
2. Try adding other variables one by one that might further improve predictive power: season, humanitarian presence (lag = 5mo) and absolute terms of trade (lag = 5mo).
3. Try replacing variables with alternatives: relative instead of absolute terms of trade, year as categorical rather than continuous.
4. Test each effect modification considered plausible on theoretical grounds (Figure 19), one by one.
5. Try removing variables one by one, in order of strength of association with mortality, to come up with a more parsimonious model.
6. Try adding plausible random effects (see below).

As models were intended for estimation of new data rather than explanation of causal chains, our main criterion for model selection (and thus decision on whether to include a variable or effect modification), aside from biological plausibility, was predictive power, which we observed graphically through plots of estimated versus actual data, and quantified through different statistics, including the mean squared residuals (or mean squared error, MSE); the adjusted R^2 (namely the proportion of variability in the data explained by the model, adjusted for loss of degrees of freedom); and the proportion of predictions that were within a given level of absolute precision from the data. Accordingly, we retained variables in the model if they substantially increased predictive power even when they had weak (non-significant P-value) associations with mortality.

For each candidate model, we verified key linear model assumptions including homoskedasticity and normality of residuals (Figure 21 and Figure 22 show diagnostic plots of the final models selected). As the models' purpose was predictive, we also implemented two cross-validation techniques to quantify the models' external validity, i.e. their expected predictive power if used to estimate mortality based on new data:

1. The bootstrap method, whereby B (in our case 100, a computationally feasible number) samples of the same size as the original data are drawn <u>with</u> replacement from the original data; each bootstrap sample is used as the training dataset to fit the model, which is then used to predict the original data: the mean of the B MSE values obtained from B fits to the original data is compared to that of the model being cross-validated as a measure of how optimistic the reported model accuracy based on the training dataset is, compared to its likely value when used with new data. Note that the bootstrap method has a tendency to under-estimate the model's optimism.
2. Leave-one-out cross-validation (LOOCV), whereby each observation i in the dataset is left out in turn: the model is trained on all the other data and tested by predicting mortality for the observation left out. This provides a set of predictions that may be thought of as arising from application of the model on a new dataset. We used the LOOCV output to calculate expected external predictive power statistics: in particular, the expected R^2 on external data can be calculated roughly as $R^2_{\text{external}} = 1 + \frac{MSE_{\text{LOOCV}}}{MSE_{\text{internal}}}\left(R^2_{\text{internal}} - 1\right)$.

We fit various linear mixed models using the R `lme4` package incorporating different biologically plausible random effects (a random effect of super-cluster on mortality, and additional random effects of terms of trade, main survey location, occurrence of epidemics or livelihood, all grouped by super-cluster), assuming an unstructured covariance matrix. These models all had higher predictive power than the fixed effects only option based on the training data, but this advantage disappeared when models were cross-validated, suggesting over-fitting; furthermore, some mixed models resulted in markedly non-normal and heteroskedastic residuals (data not shown). We therefore retained fixed effects only in the final analysis (see Table 6).

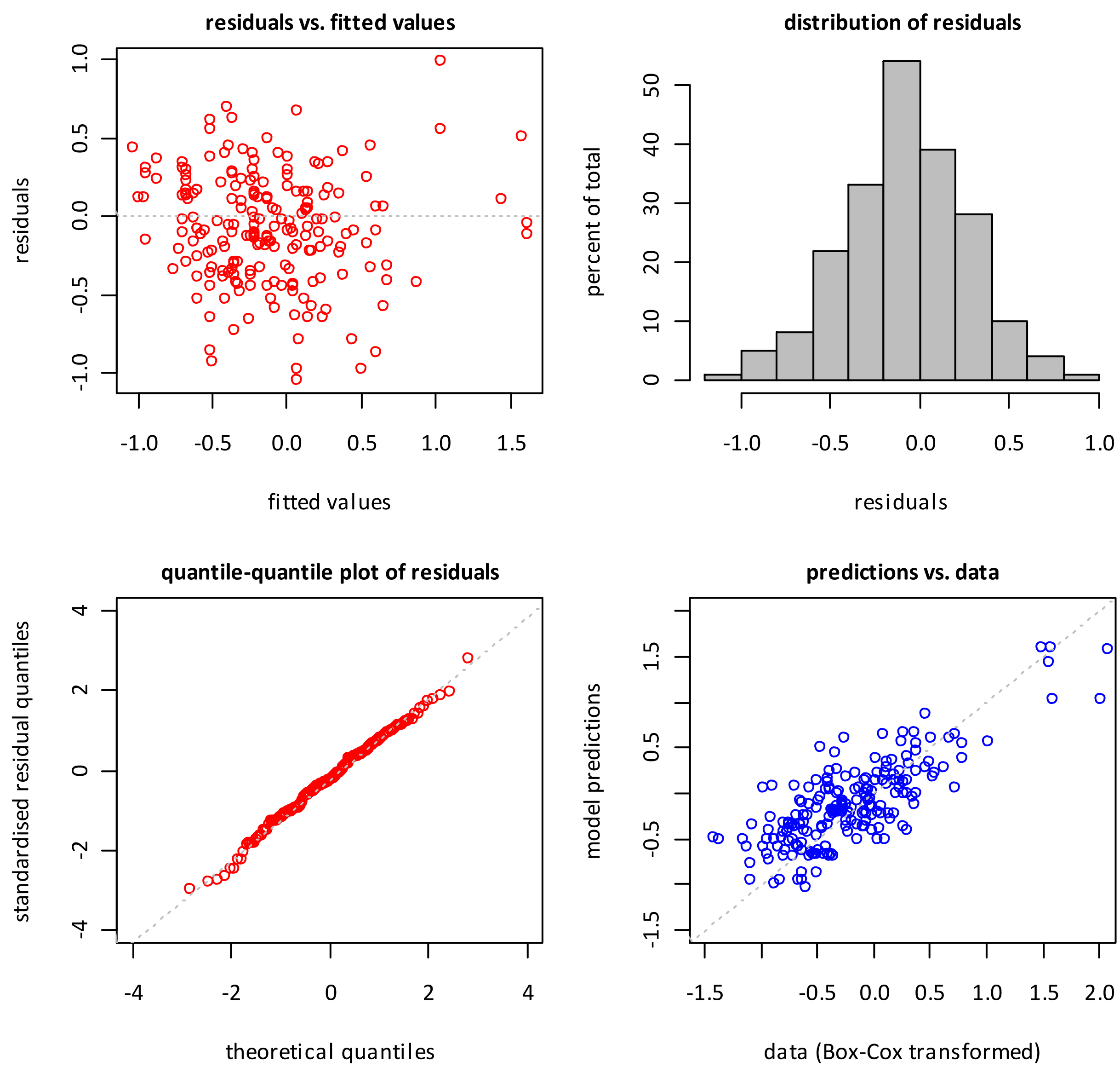

Figure 21. Diagnostic plots for final model to estimate CDR in person-time without mortality data coverage.

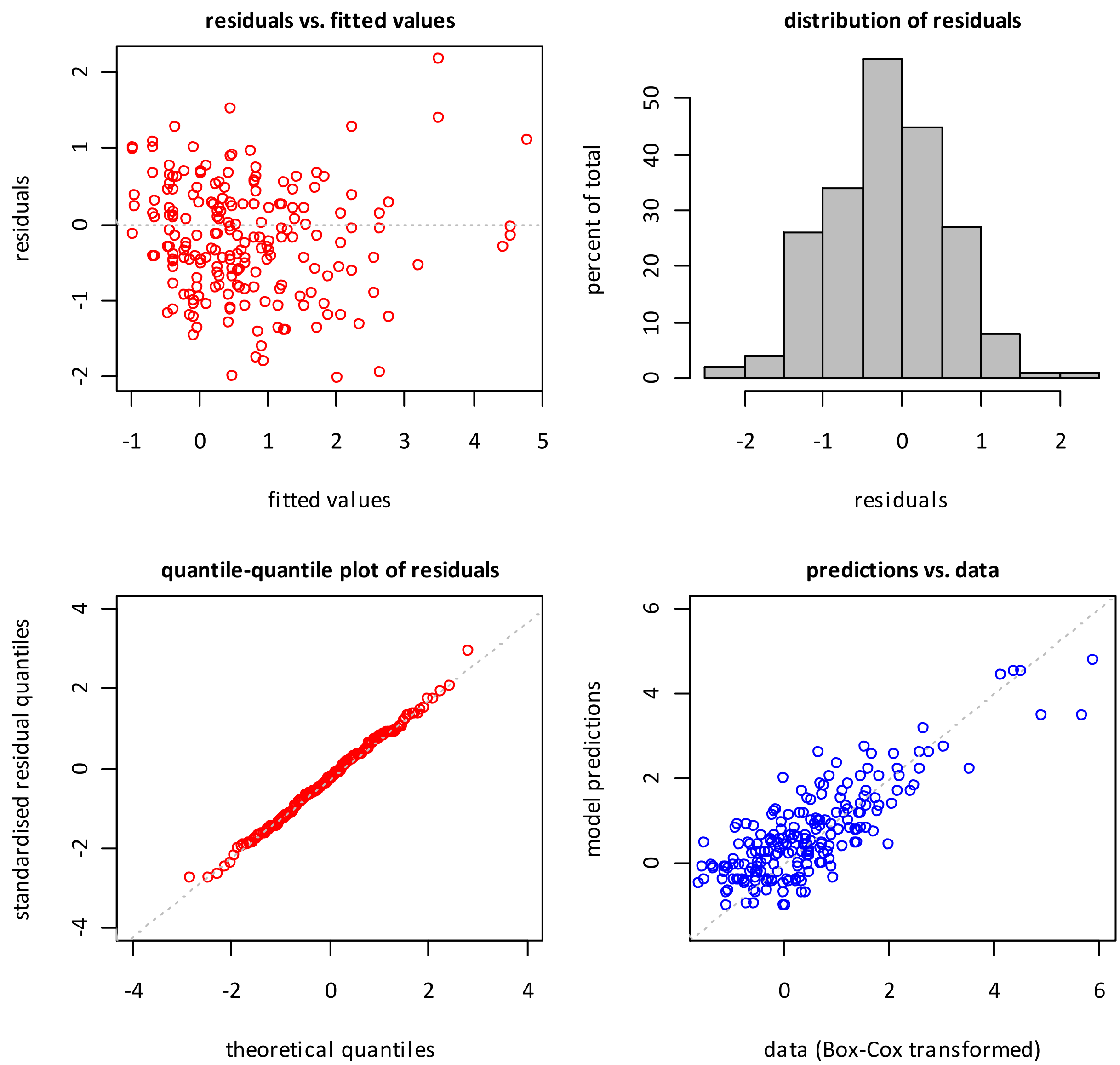

Figure 22. Diagnostic plots for final model to estimate U5DR in person-time without mortality data coverage.

6.4 Steps in the simulation to estimate excess death tolls

6.4.1 Draw a random value of the total death rate

Stratum-months covered by data

During each simulation run, we sampled random natural log CDR and U5DR values from a normal distribution given by the log of the survey point estimate and the SE of the log estimate. We then back-transformed death rates to the linear metric.

We simplistically assumed that CDR and U5DR were constant across all strata and months covered by the survey. However, for each stratum-month we adjusted the log standard error as follows:

$$SE\ (\ln \hat{y})_{i,j} = \frac{1}{\sqrt{d_{i,j}}} K = \frac{1}{\sqrt{d_{i,j}}} SE\ (\ln \hat{y})\sqrt{d} = SE\ (\ln \hat{y}) \frac{\sqrt{\hat{y}NT}}{\sqrt{\hat{y}n_i t_j}} = SE\ (\ln \hat{y}) \frac{\sqrt{NT}}{\sqrt{n_i t_j}}$$

where i is a stratum of population n_i within the survey's sampling universe of population N, and j is a month of duration t_j days within recall period T (other notation is presented above). This adjustment makes stratum-month estimates less precise than the survey as a whole, since they consist of less person-time.

In practice the above assumptions and adjustment have minimal influence on the main result of this study (i.e. the death toll across Somalia over the entire emergency period) since stratum-months are re-aggregated in order to compute this result.

A few stratum-months were covered by multiple surveys. In this case, we sampled from one of the survey distributions at random, with probability of selection for each survey s (out of S total surveys covering the stratum-month) given by

$$p_s = \frac{W_{1,s} \frac{n_{i,s} t_{j,s}}{N_s T_s}}{\sum_s^S W_{1,s} \frac{n_{i,s} t_{j,s}}{N_s T_s}}$$

Accordingly, surveys that were more precise and of higher quality, and for which a greater proportion of the sampled person-time fell within the stratum-month, were attributed higher probabilities of selection.

Stratum-months without data

We used predictive models (see above) to estimate Box-Cox transformations of CDR and U5DR for each such stratum-month, given predictor variable data for that month. We drew random transformed CDR and U5DR values from the normal distribution given by the prediction and its standard error, and back-transformed these to the linear metric.

6.4.2 Draw a random value of the baseline death rate

Method 1

For each run we sampled a CDR and U5DR value as above from one of the surveys within the set of surveys meeting criteria for the baseline, selected at random. For each survey, the probability of selection was

$$p_s = \frac{W_{1,s} L_s}{\sum_s^S W_{1,s} L_s}$$

Where L_s is an importance weight arbitrarily set at 5 if the survey was done in the same livelihood type as the stratum, and 1 otherwise. This ensures that the baseline is to some extent specific to each livelihood (data were too sparse to define a baseline distribution for each livelihood).

Method 2

For each run we sampled a CDR and U5DR value from one of the surveys within the set of baseline surveys, selected at random. For each survey the probability of selection was

$$p_s = \frac{W_{1,s}}{\sum_s^S W_{1,s}}$$

Method 3

For each run we sampled a CDR and U5DR value from the model-predicted Box-Cox transformed death rate and its standard error, as above. The prediction was done on an artificial dataset reflecting baseline conditions (see Section 2.4.4).

6.4.3 Compute the excess death rate

For each run we subtracted the random CDR and U5DR baseline value from the random CDR and U5DR total death rate.

6.4.4 Compute the excess death toll

For each run we multiplied the excess death rate by the estimated population in the stratum-month and 30.42 days on average per month to yield excess deaths.

6.4.5 Modifications for the refugee camps

For the refugee camp estimation, the total death rates were fixed at either the optimistic or pessimistic scenario. Random sampling of a baseline CDR or U5DR under method 3 was slightly modified: for each camp-month in Ethiopia, we selected a corresponding stratum-month in Somalia and used estimated baseline values for this stratum-month. Selection of the stratum-month was random, with selection weights provided by the proportion of the refugee new arrivals during that month that originated from the region containing the stratum, as per UNHCR data. For Kenya we did not have a month by month breakdown of new arrival origins, and therefore only used a single set of origin data [23] to compose the stratum selection weights.

The population under 5y old for refugee camps was sampled from under 5 y old proportions reported by eligible mortality surveys included in the analysis (25.8% and 25.6% for Dollo Ado; 31.6% and 24.4% for Dadaab).

This study was jointly commissioned (and funded) by FAO/FSNAU and FEWS NET from the London School of Hygiene and Tropical Medicine and the Johns Hopkins University Bloomberg School of Public Health.

FAO and the Somalia Food Security and Nutrition Analysis Unit express their appreciation for close collaboration with FEWS NET and the financial and technical support received from the many partners supporting FSNAU.