Estimating Mortality in the Aftermath of the Indian Ocean Tsunami

Elizabeth Frankenberg
Thomas Gillespie
Samuel H. Preston
Bondan Sikoki
Duncan Thomas

On December 26, 2004 the Sumatra-Andaman earthquake occurred in the Indian Ocean. Registering a magnitude of 9.3 on the Richter scale, the quake's vibrations were strong enough to shift the location of the North Pole by several centimeters (Sheble, 2005; NASA 2005). The quake generated a 1200 mile rupture, displaced a trillion tons of water, and generated a tsunami surge that slammed into the island of Sumatra shortly after the earthquake (Kerr 2005; Lay et al., 2005; Marris 2005).

The tsunami ultimately wreaked havoc on 10 countries and some 4500 kilometers of coastline throughout the region. Estimates suggest that worldwide casualties number around a quarter of a million people. Indonesia was the country hardest hit. Deaths there probably account for over two-thirds of total mortality from the tsunami and the damage to public infrastructure, productive assets, and private property is estimated at a value of $US 4.5 billion.

In Indonesia the brunt of the tsunami's impact was concentrated in the province of Aceh. It is thought that some 170,000 people died, or roughly 5% of Aceh's total population. Among survivors, some 700,000 were displaced, entire communities were devastated, and damage to the built and natural environment was massive.

In this paper we focus on the impact of the tsunami on mortality levels and patterns. One result of our work will be an estimate of the total number of Indonesians killed by the tsunami. Many of the media accounts of the tsunami have focused on the vast toll it exacted with respect to mortality, with estimates based on body counts or reports of the missing. Our approach to the task of quantifying these outcomes will be much more systematic both in terms of data and methodology. It will also support a calculation of degree of excess mortality resulting from the tsunami

We will analyze data from a longitudinal survey of some 40,000 individuals in the tsunami-affected areas of Aceh and in nearby comparison areas of inland Aceh and the neighboring province of North Sumatra. Baseline data were collected in February 2004, prior to the earthquake, by Statistics Indonesia as part of their annual cross-sectional Socioeconomic Survey (SUSENAS).

Our first re-survey, the Study of the Tsunami Aftermath and Recovery (STAR) took place between May 2005 and May 2006 (STAR1). We ascertained survival status for 96% of the respondents to the 2004 baseline and interviewed 94% of known survivors. The greatest challenge arose when we were unable to find a 2004 respondent or any of the members of his or her 2004 household. In those cases, we used a special "mini" version of the preprinted household roster to obtain information about the person from multiple sources, including the village leaders, the person's immediate relatives and former neighbors, and other people in the community. This information included the informant's assessment of the individuals' current survival status and whereabouts, a note for additional information (such as whether the person was seen alive recently, or the body was seen after the tsunami), and a place for the interviewer to assess his or confidence in the informant. We also drew on village-level records of who survived and who did not, which were compiled in order to facilitate the disbursement of survivor benefits, and on centralized records maintained by the Indonesian Red Cross.

The STAR data provide an unprecedented opportunity to examine the impact of this natural disaster on patterns of mortality in far more detail than is typically possible. Additionally, by combining the STAR data with other sources of information, we can construct an estimate of total tsunami-induced mortality with a much higher degree of accuracy than can be done with the limited information on body counts available in the tsunami's immediate aftermath. Many bodies were never recovered and many of the recovered bodies were never recorded. The process of constructing these estimates will likely yield innovative methods for mortality estimation after disasters.

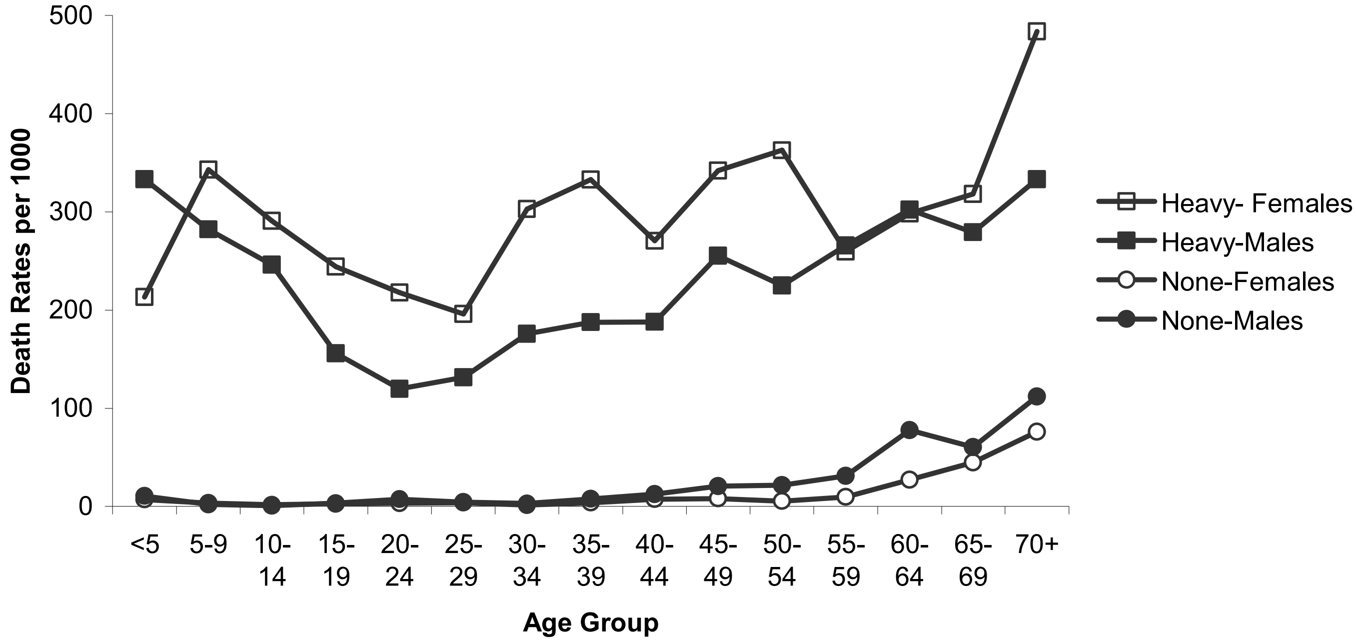
The basic methods that have been used to estimate the excess mortality associated with natural or man-made disasters are body counts, intercensal methods, and retrospective surveys. The utility of body counts is limited to circumstances where the counting has become administratively routine, as in the death of U.S. soldiers in Iraq (Preston and Buzzell, 2006). Under most circumstances, the counting function is severely disrupted by the disaster itself. Intercensal methods (e.g., Ashton et al, 1984; Heuveline, 1998) infer excess deaths by comparing the numbers of people recorded in censuses before and after the disaster, taking account of the attrition expected from "normal" mortality.

A major problem with these methods is that there is no way of determining whether those missing from the later census had died or moved away, a frequent occurrence following a disaster. Estimates of normal mortality are also subject to error. Retrospective surveys ask respondents about the survival status of relatives (Hirschman, Preston, and Loi, 1995; Roberts et al., 2006). Any intra-family correlation in the risk of dying leads to an underestimate of the mortality impact of the disaster. Reference period biases have also plagued retrospective estimates of mortality (Preston, Heuveline, and Guillot, 2001).

All of these problems can be avoided or minimized by a data system that identifies individuals before the disaster and establishes their survival status after the disaster. From the STAR surveys, we know which of the respondents to the 2004 survey were still alive at the time of the STAR1 survey, some 5-17 months after the tsunami. To illustrate the quality of data emerging from these surveys, we have used the information on survival status at STAR1 and the date of the STAR0 and STAR1 interviews to construct annualized death rates (per 1000) by age, sex, and degree of tsunami damage in the community.

The death rates for the most heavily damaged communities, and those with no tsunami damage, are graphed below. The death rates in the communities without tsunami damage provide information on what death rates in the damaged communities might have been in the absence of the tsunami. They are low at young ages and similar for males and females until age 40 or so, when a gap between males and females appears, with males at a disadvantage that is widely observed across populations. The pattern in the heavily damaged areas is starkly different. Mortality rates are very high, reaching nearly 500 per 1000 for women in their 70s, and females are at a disadvantage at almost all ages. The female disadvantage is particularly high between the ages of 30 and 50. These patterns by age and sex are very different from what emerges in other high mortality contexts such as famine, war, and genocide, where death rates tend to be the highest for prime-age males.

Death Rates by Age, Sex, and Level of Damage




These graphs clearly depict the important influence of age and sex with respect to the degree of excess risk of death relative to baseline risks. We have conducted some preliminary multivariate analyses that confirm that age and sex matter far more than measures of socioeconomic status such as education levels or pre-tsunami levels of economic resources. As part of this paper these analyses will be refined. For example we will consider whether other measures of economic resources, such as asset ownership or, perhaps more importantly, characteristics of housing construction, offered any protection and how such protection might interact with age and sex.

The death rates presented above were calculated from the STAR data, in combination with damage measures constructed from remotely-sensed information on loss of vegetation. STAR is a survey and so by definition provides information for only a subset of the tsunami-damaged areas. To construct an estimate of total mortality, we will draw on Statistics Indonesia data from the 2000 Census for Aceh and Sumut as a source of information on the pre-tsunami population for the entire region. The satellite information is also available for the entire region and can be used to stratify all the administrative areas of Aceh into damage zones. Once the communities are stratified, we will apply the estimates of "normal" mortality (from the STAR data from undamaged communities) to the census data to estimate total mortality, by region, in the absence of the tsunami. We will then apply the estimates of mortality as a result of the tsunami (from the STAR data organized by damage stratum) to the census data by stratum. The difference in total deaths within the strata under the two different mortality regimes will yield our estimate of total tsunami deaths as well as the standard errors associated with those estimates.

In addition to estimating total mortality using the methodology outlined above, we will also take advantage of data from a special census conducted in Aceh in 2005 explicitly to enumerate the population in light of the tsunami. These data will provide an additional estimate of total mortality from the two censuses by intercensal methods (Preston, Heuveline, and Guillot, 2001). As noted above, however, data from the 2005 census will reflect migration, including that induced by the tsunami, as well as mortality. We expect that the definitive results will be derived from the STAR data.

In developing these estimates, we will carefully consider the question of how the basic methods could be used in other contexts. The question is important given that remotely sensed data is becoming more and more widely available and that the frequency of natural disasters has increased over the past several decades. Although a longitudinal survey with a large sample size might not be routinely available, under certain circumstances it may be possible to replicate the approach with a rapid assessment survey of mortality (or another phenomenon such as injury), particularly if some sort of pre-event survey was available from which to select communities thought to be representative and to generate rosters of household composition before the event that can be used to determine the survival status of household members after the event.