Estimating war deaths with population-based survey data

Extended Abstract
IUSSP Session 1403: Demography of armed conflict

Ziad Obermeyer, MD, MPhil, research scientist[1] and resident physician[2]
Julie Knoll Rajaratnam, PhD, research scientist[1]
Chang Park, BA, research fellow[1]
Christopher JL Murray, MD, DPhil, institute director[1]
Emmanuela Gakidou, PhD, associate professor[1]

[1] Institute for Health Metrics and Evaluation, Seattle, WA, USA
[2] Department of Emergency Medicine, Brigham & Women's and Massachusetts General Hospitals, Boston, MA, USA

Background

Accurate measurement of war deaths is crucial for political, military, and public health planning, as well as for purposes of national history and reconciliation. However, data from nationally-representative vital event registration (VR) systems, the gold standard for counting deaths, are seldom available during wartime, either because they never existed in most war-torn countries or because they ceased to function due to war.

A number of methods for measurement of war deaths have emerged in the absence of VR data, each with their own strengths and weaknesses. First, in a limited number of countries including Bosnia and Guatemala, investigations have compiled and reconciled lists of deaths based on information from eyewitnesses and well as burial sites, missing persons lists, and other sources. These thorough enquiries are useful for providing a minimum number of confirmed or likely deaths, though without comprehensive population data it is impossible to know with certainty what fraction of all deaths were actually captured.

Second, a number of efforts have attempted to compile data from media and other published eyewitness reports of war deaths. Media reports, which are increasingly readily available in digital form, circumvent most of the logistical difficulties and expense of active investigations, and are consequently available for a far larger number of countries; however, these data have been shown to suffer from a number of biases that make interpretation problematic. High levels of war related mortality occur in those violent and dangerous areas where eyewitnesses are least likely to go, resulting in systematic under-reporting of deaths by media sources. These data are also are also subject to political pressures, which can result in either exaggeration or under-reporting of deaths.

Finally, population-based surveys have also been used to estimate war deaths, both those resulting directly from violence and those caused indirectly by wartime conditions.

Surveys have a number of advantages over other methods: random population sampling avoids many major sources of bias, and standard statistical methods can accurately represent the uncertainty underlying estimates. As with systematic enquiries, a disadvantage of surveys is the time and expense required, and for some surveys the additional logistical challenges of working during active combat.

Data and methods

Most survey-based estimates of war mortality are drawn from dedicated wartime surveys studying mortality. Our approach instead uses retrospective data from two large, independent population-based survey programs to estimate war deaths. These retrospective data are readily available from a wide range of countries. First, the World Health Organization's 2002-3 World Health Survey (WHS) Programme gathered data from respondents regarding the births and deaths of their siblings, and asked specific questions about the cause of sibling deaths including whether they had died from war injuries. A proportional mortality analysis was used to determine what fraction of siblings died from war injuries over time by age and sex, correcting for under-representation of high-mortality families, the resulting set of fractions was applied to UN Population Division (UNPD) estimates of deaths to create estimates of violent war deaths.

Second, a large and growing number of Demographic and Health Surveys (DHS) collect sibling history data in a wide range of countries. We build on ongoing efforts to estimate general adult mortality on the basis of DHS sibling history data, using logistic regression to directly model probabilities of death by age group, sex, and chronological period. We correct for two known problems with estimates of mortality based on survey data: the systematic under-representation of high mortality families, and recall bias due to omission. Once we have estimated mortality by sex, age group, and year, we can identify countries in which wars have occurred during the period covered by the surveys, and construct a counterfactual of mortality in the absence of war based on trends from time periods contiguous to wartime periods. We thus obtain two sets of probabilities of death for the wartime periods, one observed (M) and one counterfactual (M'). These are then applied to UNPD estimates of population and compared, yielding a total number of deaths attributable to war that captures both direct and indirect war mortality.

Preliminary results and significance

We have presented results from our analysis of WHS data elsewhere (Obermeyer, Murray, & Gakidou, "Fifty years of violent war deaths," *British Medical Journal* 2008) but it is worth briefly reviewing and discussing the two main findings here. First, WHS data provide estimates of war mortality over time that fit with most qualitative historical accounts, and that are plausible in terms of sex distribution. Second, WHS estimates are on average higher than compiled totals based on media and other eyewitness report data. In Bosnia, the WHS estimate was also higher than a rigorous forensic investigation that collected statements of surviving victims and eyewitnesses. While further research is

needed to explain the differences between WHS and other estimates, there are at least three explanations for the discrepancies. First, it is possible that the database of media reports to which the WHS estimates were compared captures only "battle deaths," and not the one-sided violence that also characterises some conflicts; since these latter deaths were presumably captured by the WHS and not the media database, this methodological difference could account for part of the difference. Second, some aspect of the WHS sampling or methodology may have led to an exaggeration of war deaths, though to our knowledge no specific factors have been proposed. Third, it is possible that despite our efforts to conservatively represent uncertainty, true uncertainty is larger than presented. This could imply that our estimates are not statistically distinguishable from media report estimates, though again, we are not aware of any specific aspects of the surveys or methods that would account for this. Finally, a considerable part of the difference may be explained by the known downward biases inherent to media and eyewitness data. In the specific case of Bosnia, for example, it is possible that some deaths were simply not discovered by the investigation, or that evidence of war crimes was actively obscured by parties with much to lose.

Data from the DHS surveys are in a more preliminary stage, but the Figure below shows an example of the data on general adult mortality that will be used to generate our specific estimates for war mortality. The Figure shows a summary metric of adult mortality, the probability of death between ages 15 and 60 ($_{45}q_{15}$), in Rwanda from 1988 to 2003, on the basis of sibling history data from three consecutive DHS (with bars showing the 95%CI). The dramatic effects of the 1994 genocide can be seen clearly in the $_{45}q_{15}$ estimate covering 1991-1995. Values of $_{45}q_{15}$ so close to one imply that, had prevailing genocide-era death rates been applied to a cohort of adults over multiple years, nearly the entire population would have died.