Journal of Conflict Resolution

http://jcr.sagepub.com

Estimating War Deaths: An Arena of Contestation

Michael Spagat, Andrew Mack, Tara Cooper and Joakim Kreutz

Journal of Conflict Resolution 2009; 53; 934
DOI: 10.1177/0022002709346253

The online version of this article can be found at:
http://jcr.sagepub.com/cgi/content/abstract/53/6/934

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Peace Science Society (International)

Journal of Conflict Resolution
Volume 53 Number 6
December 2009 934-950
© 2009 The Author(s)
10.1177/0022002709346253
http://jcr.sagepub.com

Estimating War Deaths

An Arena of Contestation

Michael Spagat
Department of Economics, Royal Holloway College
University of London, UK
Andrew Mack
School for International Studies
Simon Fraser University, Canada
Tara Cooper
School for International Studies
Simon Fraser University, Canada
Joakim Kreutz
Department of Peace and Conflict Research
Uppsala University, Sweden

In a much-cited recent article, Obermeyer, Murray, and Gakidou (2008a) examine estimates of wartime fatalities from injuries for thirteen countries. Their analysis poses a major challenge to the battle-death estimating methodology widely used by conflict researchers, engages with the controversy over whether war deaths have been increasing or decreasing in recent decades, and takes the debate over different approaches to battle-death estimation to a new level. In making their assessments, the authors compare war death reports extracted from World Health Organization (WHO) sibling survey data with the battle-death estimates for the same countries from the International Peace Research Institute, Oslo (PRIO). The analysis that leads to these conclusions is not compelling, however. Thus, while the authors argue that the PRIO estimates are too low by a factor of three, their comparison fails to compare like with like. Their assertion that there is "no evidence" to support the PRIO finding that war deaths have recently declined also fails. They ignore war-trend data for the periods after 1994 and before 1955, base their time trends on extrapolations from a biased convenience sample of only thirteen countries, and rely on an estimated constant that is statistically insignificant.

Keywords: *civil war; armed conflict; war deaths; health surveys*

Introduction

Estimating war deaths can be a highly contentious exercise as the sharply polarized debates over fatality totals in postinvasion Iraq remind us.[1] Controversy arises in part because conflict death tolls have been calculated using very different methodologies and these have produced fatality estimates that are sometimes startlingly different.

There are two main methodological approaches to estimating war deaths; one relies on collating and recording reports of war fatalities from a wide variety of sources, while the other relies on estimates derived from retrospective mortality surveys.[2] Debate over the rival merits of the two approaches was taken to a new stage in 2008 with the publication in the *British Medical Journal* of "Fifty Years of Violent War Deaths from Vietnam to Bosnia" by Ziad Obermeyer, Christopher J. L. Murray, and Emmanuela Gakidou (henceforth, OMG).

Prior to this article being published, the debate had focused on sharply differing fatality estimates in *single* conflicts, notably Iraq (see below), but also Darfur (U.S. General Accounting Office, 2006). OMG's contribution has been to compare survey-based wartime fatality estimates from *thirteen* countries with estimates of battle-deaths that are based on reported deaths from a variety of sources for the same countries in the same years (2008b, 1). OMG are sharply critical of the latter estimates, which are taken from a data set created by the International Peace Research Institute, Oslo (PRIO).[3] OMG, like other proponents of survey-based estimation methodologies,[4] claim that report-based estimation methods *systematically* undercount death tolls because large numbers of deaths go unreported.

OMG's criticisms echo those of two articles published in the U.K. medical journal, the *Lancet,* in 2004 and 2006 that produced violent-death estimates for the war in Iraq that were far higher—particularly the 2006 estimate—than counts produced by the incident-reporting methodology of the much-cited U.K.-based nongovernment organization, Iraq Body Count, which the *Lancet* articles refer to as a form of "passive surveillance" (Burnham et al. 2006, 1426; Roberts et al. 2004, 1863).[5] Here too the central criticism was that incident-reporting methodologies led to gross undercounts of violent deaths. In the 2006 *Lancet* article Gil Burnham and co authors claim that "Aside from Bosnia, we can find no conflict situation where passive surveillance recorded more than 20 percent of the deaths measured by population-based methods"[6] (Burnham et al. 2006, 1426).

However, the *Lancet* articles' findings were in turn widely criticized for greatly *over*estimating violent deaths resulting from wartime violence in Iraq. Critics noted that, while the second survey reported in the *Lancet* produced an estimate of 601,000 violent deaths for the thirty-nine months that followed the U.S.-led invasion in 2003 (Burnham et al. 2006), a major survey funded by the United Nations (UN) with technical assistance provided by the World Health Organization (WHO), produced an estimate of 151,000 violent deaths—just one-quarter that of death toll reported in the *Lancet* study (Iraq Family Health Survey Study Group 2008).[7]

Criticisms of survey-based war-death estimations are not unique to Iraq. Recently the International Rescue Committee's (IRC) much-publicized, survey-based claim that there were 5.4 million excess war deaths in the Democratic Republic of the Congo between 1998 and 2007 was critically reviewed in the *Human Security Report 2008/9* (Coghlan et al. 2008).[8] The *Report*'s authors argue

that the methodology used in the first two of the IRC's five retrospective mortality surveys was so flawed that they should be rejected. The IRC estimated that the excess death toll for the period covered by the three subsequent surveys was 2.83 million. But the *Report*'s analysis argued that this estimate was derived from an inappropriately low baseline mortality rate. When a more appropriate baseline rate was employed, the excess death toll shrank by more than two million to less than one-third of the IRC's estimate.

A New Approach to Comparing Fatality Estimates

The arguments over the pros and cons of survey-based versus incident-recording methodologies for estimating war deaths remain far from resolved. However, the *British Medical Journal* study referred to above breaks new ground by attempting to compare the two estimation methods for thirteen countries. In principle, multicountry studies are better suited than single-country analyses for determining whether or not there are significant and systematic differences in fatality estimates generated by surveys versus those produced by report-based methodologies.

In their article, OMG extract wartime fatality data from WHO World Health Surveys for the countries in question and contrast them with estimates derived from reported battle-deaths that were collated in the PRIO study noted above. They argue that the methodology used to create the PRIO data set, which was compiled by Bethany Lacina, is flawed by its "uncritical acceptance of media reports" (2008a under A4) and that it underestimates war deaths by a factor of three.

OMG estimate a linear equation for transforming the PRIO battle-deaths estimates for the thirteen countries in their study into figures in line with those they estimate from the WHO data. They then apply this algorithm to PRIO's figures for battle-deaths in all countries and produce a revised estimate of the trend in violent war-related deaths for the period 1955 through 1994. They use this estimate to assert that "there is no evidence to support a recent decline in war deaths" (2008b, 1). They thus reject the central finding of Lacina and Gleditsch (2005) that there has been a major, decades-long decline in battle-deaths.

OMG's estimates are drawn from the sections of the WHO surveys in which respondents were asked if any of their siblings had died during periods of conflict, and if so how. These data were used to calculate the fatality rate from wartime violence sustained by each of the countries sampled.

Such estimates may potentially make a useful contribution to our knowledge of war fatalities. But, as we show below, the article's criticisms of PRIO's battle-death estimation methodology are unwarranted. The comparisons they draw are inappropriate, and the article contains serious errors of fact.

Background

The wartime mortality data that OMG use in their critique of the conflict research data are derived from WHO population health surveys taken in 2002 and 2003 in forty-five countries. Respondents reported deaths of siblings that they themselves viewed as associated with "injury . . . due to war" (OMG 2008b, 3). OMG use these data to provide estimates of "war deaths"—by which they mean fatalities from injuries that respondents attribute to warfare—for the thirteen countries out of the original forty-five for which at least five sibling war deaths were recorded during a ten-year period. (Note that this definition of "war deaths" does *not* include "indirect" deaths from war-exacerbated disease and malnutrition.) The thirteen countries are Bangladesh, Bosnia, Burma, Ethiopia, Georgia, Guatemala, Laos, Namibia, Philippines, Republic of Congo, Sri Lanka, Vietnam, and Zimbabwe. The estimates are disaggregated into ten-year periods, with 95 percent confidence intervals.

This approach allowed OMG to take advantage of the large, nationally representative surveys that the WHO had conducted, while avoiding the significant logistical and ethical challenges that confront the conduct of surveys in the midst of armed conflict. However, since the coverage periods of these surveys exceed forty-five years, recall bias may constitute a major problem. Recall bias is of sufficient concern among epidemiologists for the Standardized Monitoring and Assessment of Relief and Transitions (SMART) guidelines on survey methodology to state categorically that "Recall periods longer than one year should not be used."[9]

OMG go on to suggest that the PRIO data for these same thirteen countries are based primarily on media reports.[10] In their Technical Appendix, OMG even claim that the PRIO data set relies on, ". . . estimates based on an *uncritical acceptance* of media reports" (2008a under A4, emphasis added). They further argue that these data undercount the true war-death toll by a factor of three and that this undercounting arises because many deaths, particularly in the deadliest conflicts, go unreported in the media, and that "there is no evidence" to support PRIO's claim that there has been a long-run decline in battle-deaths.

The sections that follow review a number of methodological shortcomings in the OMG article and demonstrate why their criticisms of the PRIO data set are unwarranted.

PRIO versus OMG

In this section, we argue that OMG mischaracterize the methodology used to create the PRIO data set, that their claim that it underestimates battle-deaths by a factor of three is simply wrong, and that their choice of countries to sample may well be biased toward finding a high ratio between their death-toll estimate and PRIO's.

1. The PRIO Data Set is Not Based Primarily on Media Reports

OMG claim repeatedly that the PRIO data set is based on media reports. In fact, as the documentation for the data set makes clear, fatality estimates are *not* based primarily on media reports.[11] PRIO used all available sources, making informed judgments about their reliability before including anything in the fatality estimates. The following rough de facto hierarchy of source reliability was employed: peer-reviewed studies aimed at integrating the findings from many sources; the work of epidemiologists and demographers; military historians; other published casualty estimates, mostly in books on particular conflicts; and Keesings Record of World Events.

Clearly then, the PRIO data set's fatality estimates do *not,* as OMG claim, rely primarily on media reports, nor does PRIO accept media reports uncritically where these are drawn on. Indeed some of the sources that PRIO reviews are as likely to *overestimate* as underestimate battle-death tolls. For example, for many years, the most commonly used estimate of war deaths in Bosnia was some two hundred thousand. Subsequent in-depth investigations using all available sources put the death toll at half that number (Ball, Tabeau, and Verwimp 2007).[12]

This is not to say that PRIO's sources are wholly reliable—*no* method of estimating war deaths, including media surveillance, is wholly reliable. But we should note that even where fatalities are systematically undercounted, surveillance data can still be of great value to analysts and policy makers because they reveal trends—whether things are getting better or worse—and can do so in near real-time.

2. PRIO Estimates are Substantially Larger than OMG Claim

OMG's analysis does not support their claim that PRIO's global fatality toll indicates, "a figure of only a third" (2008b, 1) of their survey-based estimate. OMG estimate that there were 5,393,000 war deaths in the thirteen countries while ascribing to PRIO a figure of 2,784,000 for these countries. But this is a ratio of only 1.9 to 1. Moreover, OMG report a 95 percent confidence interval on their total estimate of 2,997,000 to 8,697,000 and a range of PRIO estimates of 2,695,000 to 3,026,000. Thus, OMG are not in a position to convincingly rule out ratios pushing down toward 1.0.

The 3.0 figure that appears in OMG's Table 3 turns out to be an unweighted average of the thirteen individual ratios of OMG's WHO-derived estimates to the PRIO figures. Thus, Georgia with a ratio of 12.0 gets the same weight as Vietnam with a ratio of 1.8 despite the fact that Georgia had only 0.6 percent of the estimated war deaths and a wide confidence interval of –71 percent to +137 percent whereas Vietnam had 71 percent of the estimated war deaths and a confidence interval of –42 percent to 56 percent.

OMG offer no argument for weighting all countries equally, and other methods for calculating the ratio yield different and lower results. As already noted, weighting all deaths, rather than the ratios, equally gives a ratio of 1.9, while taking the median, rather than the mean, of the thirteen ratios yields a ratio of 2.1. Moreover, the Georgia ratio is an extreme outlier, 2.4 times the next highest ratio and 2.6 times the third highest. Without this outlier the mean of the remaining twelve ratios falls to 2.2. Thus, OMG's factor-of-three is not robust to alternative calculations, nor to the treatment of the Georgia outlier. Three plausible alternative calculations come out in the neighborhood of two, rather than three.[13]

3. OMG Fail to Compare Like with Like

As noted above, the WHO surveys that OMG rely on measure "war deaths"—that is, fatalities from injuries that survey respondents attribute to warfare. But the PRIO data set that OMG critique does *not* measure war deaths thus defined. Rather, it counts a considerably narrower category of fatalities, namely "battle-deaths" in "state-based" conflicts—that is, those in which a government is one of the warring parties. Battle-death counts include deaths of soldiers from combat-related injuries and of civilians caught in the crossfire—so-called "collateral damage." But they exclude two important categories of "war deaths" that *are* counted by OMG. First, the PRIO data set, in common with most other conflict data sets, does not count fatalities from "nonstate" armed conflicts—those in which a government is *not* one of the warring parties. The latter category includes conflicts between rebel groups, intercommunal conflicts and violent confrontations between warlords. Deaths from these conflicts *would*, on the other hand, be recorded in the WHO survey responses.

"Nonstate" conflicts are an important source of war deaths and the focus of a new conflict data set created by Uppsala University's Conflict Data Program (UCDP). In 2002, the first year for which UCDP has recorded data on this type of conflict, UCDP recorded more "nonstate" armed conflicts than the "state-based" conflicts whose fatalities the PRIO data set counts, although most of these conflicts have considerably fewer fatalities than state-based conflicts.

Second, the PRIO battle-death data set does not include deaths from "one-sided violence"—that is, the intentional, politically motivated killing of noncombatants by either governments or nonstate armed groups—on the grounds that slaughtering defenseless citizens does not constitute armed conflict. (These data *are* now collected by the UCDP for a separate one-sided violence data set.)

Deaths from "one-sided violence" that the WHO surveys also include as "war deaths" may well account for a significant share of all war deaths. The conflicts in Bangladesh and Vietnam—both cases examined by Obermeyer and colleagues—were notorious for the slaughter of civilians, that is, for "one-sided" deaths. OMG estimated 2.6 million more deaths than PRIO did in the full thirteen-country sample.

Of these, nearly two million come from Bangladesh and Vietnam, where, a priori, we would expect war deaths to be much higher than battle-deaths. Indeed, according to Barbara Harff, eight of OMG's thirteen countries experienced genocide or "politicide" during the period covered by OMG's study (Harff 2003).[14]

Only if "nonstate" and "one-sided violence" fatalities were added to the PRIO "state-based" battle-death count would the comparison between the PRIO and the WHO-derived fatality counts be a valid one.

How much greater would the overall PRIO fatality toll be if this were done? The short answer is that no one knows. There are no reasonably comprehensive data for fatalities from "one-sided violence" prior to 1989, nor for "nonstate conflicts" prior to 2002. But it could well be the case that the total fatalities that would have been reported if these two categories had been included would have been *much* higher than the battle-death fatalities alone.

Uppsala's researchers do have data from 2002 to 2006 for both of the categories of fatality that are missing from the PRIO count of battle-deaths in "state-based" conflicts that OMG use to compare with their WHO-derived estimates. The combined fatality toll from "one-sided" violence and "nonstate" conflicts is equal to some 55 percent of the "state-based" fatality toll for the 2002 to 2006 period.

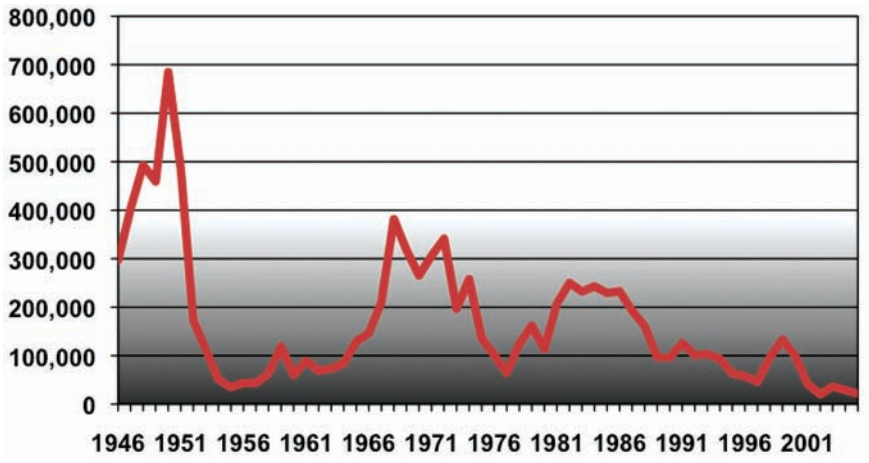
If the ratio of "nonstate" and "one-sided violence" to "state-based" fatalities that existed from 2002 to 2006 were representative of the fatality breakdown of the three types of political violence for the period 1955 to 2002, the PRIO fatality figures would need to be increased by 55 percent—that is, from 2.784 million to 4.315 million for a like-with-like comparison to be made. Making this adjustment would drop the OMG/PRIO ratio from 1.9 to 1.2, a long way from the factor of three claimed by OMG.

Our point here is that even if the PRIO data set correctly identified every single battle-death, its battle-death total would, of necessity, still be much smaller than a fatality estimate from retrospective mortality surveys that accurately accounted for *all* war deaths from injuries caused by war-related violence. If fewer categories of war deaths are being counted, then the result must be a smaller total.

4. OMG May Have a Biased Sample

OMG use the World Health Survey data posted at http://www.who.int/healthinfo/survey/en/index.html, which provides access to data from sixty-seven country surveys. According to the PRIO data set, thirty-six of these sixty-seven countries suffered battle-deaths yet only thirteen of these thirty-six cases were included in OMG's analysis. We can eliminate two cases because the documentation on the WHO site states that sibling mortality data were not collected in these surveys. OMG state that (2008, 3):

Figure 1
Global State-Based Battle-Deaths, 1946-2005



Data source: Lacina and Gleditsch (2005).

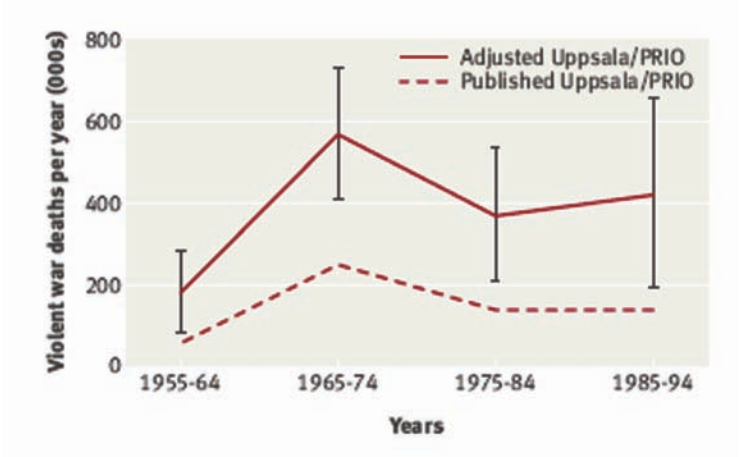
> Of the 45 surveys [including some that do not appear in the PRIO data set] with data on sibling history, 13 were found to have more than five reported sibling deaths from war injuries in a given 10 year period, and we focused on these countries in the analysis.

This implies that OMG excluded twenty-one cases that are included in the PRIO data set *because* the survey estimates were too low to allow for meaningful analysis.[15] We would expect the OMG/PRIO ratios for these omitted countries to be lower than they are for many of the thirteen countries for which OMG do report results. Thus, this procedure is biased toward finding a high OMG/PRIO ratio—when the OMG estimate turns out to be high, it is retained, but if it is low, it is discarded and does not figure in the comparison. This bias is potentially large.

Have Battle-Deaths Declined?

OMG join Sarkees, Wayman, and Singer (2003) in challenging the claim that war deaths have been declining.[16] OMG criticize what they describe as "strong claims made on the basis of current data from media reports" to the effect that "the number of deaths related to war has declined consistently since the mid-20th century and that recent wars have killed relatively few people" (2008b, 6). This claim they argue should be "reevaluated."

Figure 2
War Deaths by Decade: Uppsala Data as Published and as "Adjusted" (OMG 2008b, Figure 5)



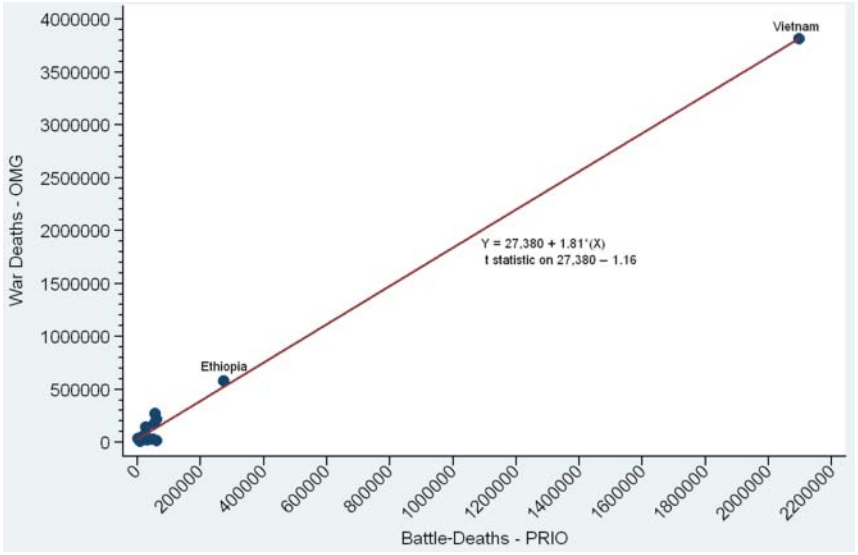
Data source: OMG (2008b, 6).

Contrary to OMG, we are not aware of any conflict researcher who has claimed that the decline has been "consistent"—it has been highly uneven as Figure 1 shows.[17] Second, the most prominent and well-documented "strong claim" in the literature is that *battle*-deaths, not *war* deaths, have declined.

OMG's claim that there is no evidence for a decline in war deaths is based exclusively on taking the PRIO data displayed in Figure 1 and then using their thirteen WHO survey estimates to transform the PRIO data so that they no longer show a clear downward trend, and even increase slightly for 1985 through 1994 compared to 1975 through 1984, shown below in Figure 2. However, the method that they use to reach this conclusion is neither sound nor convincing.

First, if the authors wished to challenge the PRIO finding that battle-deaths have declined, they should have used the same period for comparison in both cases. They fail to do this. The series of ten-year periods OMG use to support their claim that fatalities are *not* declining starts in 1955 and ends in 1994. But, as is obvious from an inspection of the plots below, the PRIO data set shows much higher deaths before 1955, and much lower deaths after 1994 than it does for the period between these years covered in the OMG picture. The WHO fatality tolls miss the Korean War—the second deadliest conflict since World War II. They also miss the decline in battle-deaths between 1994 and 2002. The battle-death count in 2002 was only 22 percent of the 1994 count according to the UCDP/PRIO version of the PRIO data set.[18] Once again OMG fail, quite inappropriately, to compare like with like.

Figure 3
OMG's War Deaths Versus PRIO's Battle-Deaths



Second, no claims about *global* fatality trends can be made with any confidence based on a nonrandom convenience sample of just thirteen countries. The OMG sample contains only 18 percent of the seventy-one countries that experienced conflict during the period covered by the WHO surveys (1955-2002) according to the *Correlates of War* version of the PRIO data set used by OMG.[19] Such a small convenience sample cannot provide a reliable basis for extrapolation. The PRIO data set, however, records battle-death fatalities in *all* countries that experience armed conflict. These latter data can legitimately be used to make claims about global trends; the former cannot.

Third, OMG perform statistical analysis on their small and dubiously selected convenience sample, fitting a line to a mere thirteen data points that give OMG estimates and PRIO figures for the thirteen countries in the OMG sample. Figure 3 gives a scatter plot of these data points that shows Ethiopia to be an outlier and Vietnam to be a huge outlier. The remaining eleven estimates become an indistinct "splotch" of clustered points near the origin. Essentially, OMG fit a line to only three data points, Vietnam, Ethiopia, and the "splotch" that represents the other eleven countries.

The key point here is that OMG's transformation of the PRIO trends shown in Figure 3 is highly sensitive to precisely how this line is drawn. Of course, a line is pinned down by a slope and an intercept. The appropriate slope here is unambiguous. This has to be about 1.8 because the line has to more or less go through the massive Vietnam point (2,784,000, 5,393,000) as well as some point near the origin.

Visual inspection of Figure 3 suggests that the origin (0, 0) is as reasonable a point as any to pick for the intercept of the line. OMG's own statistical analysis, shown in their appendix, confirms that the intersection of the line is statistically indistinguishable from 0. Nevertheless, OMG accept and lean heavily on their statistically insignificant intercept estimate of (0, 27,380).

OMG then take the PRIO battle-death figures and transform them into estimated war deaths using their special and particular choice of lines to fit the data shown in Figure 2:

$$\text{Estimated war deaths for war } W = 27{,}380 + 1.81 \times (\text{PRIO figure for war } W)$$

OMG's Figure 5 simply plots the sums of these transformed figures alongside the original PRIO figures for a truncated period. The trends differ slightly because each PRIO figure is multiplied by 1.81 and has 27,380 added to it.

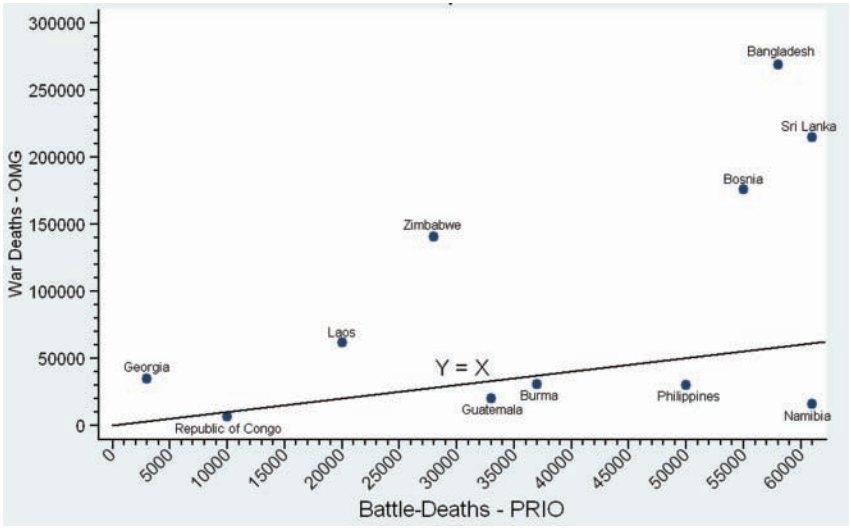
Crucially, if each PRIO figure were only multiplied by 1.81 (or any other number) *without* having something substantial added in as well, then the PRIO curve would simply make a parallel shift upward and all trends would be preserved exactly. Adding the constant is vital to OMG's argument. Without it, the trend in their war death estimates would be completely identical to the PRIO battle-death trends. Everything rests on the constant, 27,380, which according to both the formal statistics and the naked eye can easily be equal to 0.[20]

Even if this constant had been statistically significant, it would still make little sense to apply it mechanically to every conflict, since doing so implies that a conflict in which just a few hundred battle-deaths have been recorded will be estimated to have nearly thirty thousand war deaths. That is, the OMG formula implies that the claimed PRIO underestimation is proportionally much larger in small conflicts than in large ones. However, if anything like the OMG formula was true then we would rarely see conflicts with less than twenty thousand war deaths! In fact, the average "state-based" conflict in 2007 killed less than five hundred people according to Uppsala's Conflict Data Program. Yet, even if we nevertheless accept the formula and apply it to a full run of PRIO data from the Korean War to 2002, there is still a distinct decline over time.

However, the weakness of OMG's discussion of trends runs even deeper than their reliance on a statistically insignificant slope term calculated by extrapolating from a small convenience sample. OMG note in their technical appendix that one of the two "basic assumptions" that underpins their work is that there is a "*consistent relationship* of media report data to less-biased population based data" (2008a, under A4, emphasis added). Yet, the authors' own data comparisons demonstrate that *in five of the thirteen cases PRIO's fatality total is actually greater than the WHO-based estimates.*

If the relationship between the two sets of estimates really was consistent, such reversals should be rare. Moreover, consider Figure 4, which removes the Vietnam and Ethiopia outliers from the figure 3 scatter plot so we can see the variation within

Figure 4
OMG War Deaths versus PRIO Battle-Deaths:
Without Ethiopia and Vietnam



the "splotch." Absent the outliers it is clear that there is no consistent relationship between OMG's data and PRIO's. In other words, one of the two "basic assumptions" of the methodology that OMG employ is simply wrong.

Conclusion

Survey-based and incident-based methodologies for estimating and documenting deaths resulting from warfare are complementary in principle. Each has strengths and weaknesses; both serve differing analytic, policy, and advocacy purposes. But in practice, dialogue between proponents of the two approaches has been marked by misunderstanding and considerable controversy. This article has sought to contribute to the ongoing discussion about the pros and cons of these very different methodologies via a detailed examination of the first multicountry study to compare the two approaches.

The critique of PRIO's methodology mounted by Obermeyer and his co-authors' much-cited study has likely contributed to the misunderstandings noted above. The *British Medical Journal* is a prestigious, refereed journal with a high level of credibility—and a very high circulation. As such, its readers have little reason to

question the accuracy its content. And yet, as we have pointed out, OMG fail to substantiate their claim that the PRIO data set underestimates war deaths by a factor of three; their "basic assumption" that there is a consistent relationship between their survey-based fatality data and the PRIO data is flatly contradicted by the very data that they invoke in making this claim, and their startling assertion that "there is no evidence to support a recent decline in war deaths" is simply incorrect. In fact, nothing in their article challenges the PRIO finding that there has been a remarkable, though far from even, decline in battle-deaths since the 1950s.

OMG argue that their study demonstrates the bias they claim is inherent in report-based conflict data and that it does this by comparing PRIO fatality data to "less-biased population-based data." We have argued that this comparison was inappropriate because the PRIO data, which are drawn from a wide variety of sources, measure a narrower category of war deaths than do the surveys that OMG rely on. This, in combination with the flawed and arbitrary statistical choices they make when transforming the PRIO data, leaves us without any real insights into the strengths and limitations of report- versus survey-based methodologies for monitoring conflict deaths.

Estimating the number of fatalities caused by wartime violence remains an extraordinarily challenging task whatever methods are used. With incident-based methodologies—as is well understood by all those involved with conflict data collection—there is a real risk that many deaths will go unreported, even though access to fatality data has improved dramatically since the advent of the Internet and powerful search engines.

The reasons are not difficult to discern. Governments may forbid reporting of war deaths—particularly of their own forces. Journalists and other observers are sometimes banned from war zones, as in Sri Lanka in 2009, or may stay away because conditions are simply too dangerous. And in wars with very high daily death tolls—like Iraq in 2005 and 2006—violent incidents with very small numbers of deaths may go unreported.[21] There is also a risk, although we believe it is a lesser one, that death tolls will be overreported.

Furthermore, report-based methodologies, no matter how accurate, can never determine the number of "indirect" war deaths—that is, those arising from war-exacerbated disease and malnutrition. This is because the diseases that kill people in wartime also kill them in peacetime. An individual killed in combat on the battlefield is *unambiguously* a victim of war; an individual who dies from—say—malaria during the same war is not. She or he might well have died from the same disease had there been no war. But "excess deaths" from war-driven disease can, in principle at least, be estimated with survey-based methodologies that can compare the difference between prewar and wartime mortality rates from disease and malnutrition.[22]

Notwithstanding the known limitations of data generated by report-based methodologies like those created at PRIO and UCDP and the Correlates of War Project, conflict researchers who use quantitative methods rely on them overwhelmingly. As

a consequence of collaboration between PRIO and UCDP, consistently coded annually updated data (to 2008) are now available for all "state-based" conflicts from 1946 to 2008. Data on "one-sided violence" are now available from 1989, and for "nonstate" conflicts since 2002. There are data on battle death numbers, war onset and termination years, types of conflict termination, the duration of the fighting, and the identity of the warring parties. These data can be used to compare regional trends in the number and deadliness of wars, whether—and if so, how—conflicts escalate over time, and changing patterns in war terminations.

Retrospective mortality surveys, which are undertaken by different organizations using different methodologies and coding protocols, are too few in number and too disparate in methodology to have much utility for the large-*N* statistical exercises that have become standard tools of conflict research. Moreover, surveys usually do not code the identity of warring parties nor describe the different types of conflict; nor do they collect data on conflict termination types. And because they record mortality retrospectively, often with long recall periods, they are not particularly useful for tracking global and regional conflict, or war death trends. Incident-reporting methodologies, like Iraq Body Count, on the other hand, can report fatality trends in near-real time.

These very different methodologies reflect the different data concerns of scholars in the conflict research community and researchers whose major focus is global public health issues. The former are interested in exploring questions that have to do with the onset, duration, and deadliness of armed conflicts and how they end. Conflict data sets like those produced at PRIO and UCDP, notwithstanding their acknowledged limitations, are well suited for these purposes. Global health researchers who use population health surveys, like the World Health Surveys, on the other hand, are interested in a far broader range of population health outcomes than war deaths. Their research interests are well served by population health surveys, much less so by the sorts of data that PRIO and UCDP collate.

There is nothing wrong in principle with using nationally representative population health surveys to seek to estimate war deaths, but there *is* a real issue in practice. In the extraordinarily few countries *where more than one* such survey has been carried out over the same period in war-affected countries, the resulting death toll estimates have usually been *radically* different. In Iraq, as noted earlier, the reported violent death toll in the second survey by Burnham and colleagues whose findings were featured in the *Lancet* was four times larger than that reported by a major UN-funded survey (Burnham et al. 2006; Iraq Family Health Survey Study Group 2008). Both estimates cannot be true. There were comparably large differences in estimated death tolls in the two nationwide surveys in Iraq in 2004. In the case of the Democratic Republic of the Congo, the child mortality rate estimated by the International Rescue Committee was double that of the 2007 Demographic and Health Survey over the same period.[23]

The extraordinary divergence in these estimates and the lack of any consensus as to their cause, plus the problems we have identified with the only multicountry study comparing survey-based fatality estimates with estimates from the conflict research community, suggest that the utility of survey-based approaches to measuring war deaths, while clear in principle, still confronts major challenges in practice.

Notes

1. For example, see coverage on Iraq in *Science* and *Nature* (Bohannon [2006, 2008] and Giles [2007]).

2. In principle, census data can also be used to generate war death estimates, but in practice this rarely happens. The Benetech Human Rights Data Analysis Group (http://www.hrdag.org/) has developed its own methodology for estimating violent deaths based on documented deaths taken from multiple sources.

3. This data set, that is, the "PRIO data set," is available at http://www.prio.no/CSCW/Datasets/Armed-Conflict/Battle-Deaths/. The key reference is Lacina and Gleditsch (2005). OMG confuse what we have called the PRIO data set with the Uppsala-PRIO data set, which is available at http://www.prio.no/CSCW/Data sets/Armed-Conflict/UCDP-PRIO/. Moreover, the PRIO dataset comes in three versions created for use with different lists of wars/conflicts. OMG appear also to switch confusingly between the version of the PRIO data set compatible with the Correlates of War list of wars and the version of the PRIO data set compatible with the UCDP/PRIO list of conflicts. See our online appendix.

4. Burnham et al. (2006).

5. They were also far higher than the estimates of the Uppsala Conflict Data Program that are based on incident reporting.

6. Spagat (2009) uncovers many ways in which this statement, and the evidence cited to back it up, is incorrect.

7. The second *Lancet* article by Burnham et al. (2006) attracted the most criticism. See Daponte (2007), Guha-Sapir and Degomme (2007), Iraq Family Health Study Group (2008), Johnson et al. (2008), Laaksonen (2008), Rosenblum and van der Laan (2008), and Spagat (2009).

8. Two of the authors of this article, Andrew Mack and Tara Cooper, are researchers at the Human Security Report Project.

9. See www.smartindicators.org/SMART_Methodology_08-07-2006.pdf (2006, 31). Recall bias refers to the well-documented fact that respondents systematically make mistakes in recalling past events and that the probability of making mistakes increases with time. In the case of wars, deaths occurring during war are not necessarily the same thing as deaths caused by war, but the two types of deaths may well be confused by respondents especially after many years have passed. Indeed, the boundaries between the two categories are often blurred in fact as well as in individuals' memory. If deaths occurring during periods of war have been incorrectly identified as "war deaths," as is quite possible, OMG's war deaths totals will be biased upward.

10. The section titled "Discussion" refers repeatedly to the PRIO data as being based on media reports.

11. Most of the documentation for the PRIO data set is available at http://www.prio.no/sptrans/-671630282/Documentation_PRIO-UCDP.pdf. Additional documentation relevant to the Correlates of War version of the PRIO data set, which is the main version used by OMG, is available at http://www.prio.no/sptrans/-420217012/Documentation_COW.pdf.

12. Conflict researchers working in this field are well aware that fatality counts that are drawn solely from media reports tend to undercount the true extent of the death toll. The much-cited Iraq Body Count project, for example, says explicitly that its reporting of civilian fatalities in Iraq, which relies heavily on media reporting, undercounts the extent of the death toll (Sloboda, Dougherty, and Dardagan 2007).

13. Moreover, the Georgia estimate itself is unreliable. See Spagat et al. (2009), the working paper version of the present article, for details.

14. These countries are Bangladesh (i.e., East Pakistan), Bosnia, Burma, Ethiopia, Guatemala, Philippines, Sri Lanka, and Vietnam. "Politicides" are attempts to exterminate, in whole or in part, groups that are defined in terms of their politics. Genocides target groups defined on the basis of their ethnicity or religion.

15. In all twenty-one excluded cases, the PRIO data set contains battle-deaths during the 1955 to 2002 period covered by OMG. They are Burkina Faso, Chad, China, Comoros, Cote d'Ivoire, Dominican Republic, Ecuador, Hungary, India, Kenya, Malaysia, Mali, Mexico, Morocco, Nepal, Paraguay, Russia, Senegal, South Africa, Spain, and Tunisia.

16. See Lacina, Gleditsch, and Russett (2006) for a rebuttal of the Sarkees et al. (2003) argument.

17. For figure 1, see http://www.prio.no/CSCW/Datasets/Armed-Conflict/Battle-Deaths/The-Battle-Deaths-Dataset-version-20/.

18. The *Correlates of War* version of the PRIO data set only extends to 1997 at the present moment and, therefore, is not usable for this comparison.

19. According to the UCDP/PRIO version of the PRIO data set that, unlike the *Correlates of War* version, includes low-intensity conflicts, OMG only work with 11 percent of the 114 countries that experienced conflict during the period covered by the WHO surveys.

20. The standard error on the constant, 27,380, in OMG's regression is 23,118, so this constant is not anywhere close to any standard significance level.

21. However, an analysis by Urlacher (2009) studies journalistic coverage of civil wars, 1992 to 1999, and finds that "news coverage is largely unaffected by violence, except in the most extreme circumstances (186)".

22. However, what may be possible in principle may be impossible in practice because of major difficulties in establishing baseline mortality rates. See "The Hidden Costs of War" by Human Security Report Project (2009).

23. For a detailed analysis, see "The Hidden Costs of War" in Human Security Report Project (2009).

References

Ball, P., E. Tabeau, and P. Verwimp. 2007. The Bosnian book of the dead: Assessment of the database: Executive summary. *Human losses in Bosnia and Herzegovina 91-95.* Sarajevo: Research and Documentation Center. http://www.idc.org.ba/presentation/content.htm.

Bohannon, J. 2006. Iraqi death estimates called too high; methods faulted. *Science* 314 (October 20): 396-97.

Bohannon, J. 2008. Calculating Iraq's death toll: WHO study backs lower estimate. *Science* 319 (January 18): 273.

Burnham, G., R. Lafta, S. Doocy, and L. Roberts. 2006. Mortality after the 2003 invasion of Iraq: A cross-sectional cluster sample survey. *The Lancet* 368 (9545): 1421-428.

Coghlan, B., P. Ngoy, F. Mulumba, C. Hardy, V. N. Bemo, T. Stewart, J. Lewis, and R. Brennan. 2008. *Mortality in the Democratic Republic of Congo: An ongoing crisis.* New York: International Rescue Committee.

Daponte, B. O. 2007. Wartime estimates of Iraqi civilian casualties. *International Review of the Red Cross* 89 (868): 943-57.

Giles, J. 2007. Death toll in Iraq: Survey team takes on its critics. *Nature* 446 (7131): 6-7.

Guha-Sapir, D., and O. Degomme. 2007. Estimating mortality in civil conflicts: Lessons from Iraq: Triangulating different types of mortality data in Iraq. *CRED working paper* (June), http://www1.cedat.be/Documents/Working_Papers/CREDWPIraqMortalityJune2007.pdf.

Harff, B. 2003. No lessons learned from the Holocaust. *American Political Science Review* 97 (1): 57-73.

Human Security Report Project. 2009. *Human security report 2008/9.* New York: Oxford University Press.
Iraq Family Health Study Group. 2008. Violence-related mortality in Iraq from 2002 to 2006. *The New England Journal of Medicine* 358 (January 31): 484-93.
Johnson, N. F., M. Spagat, S. Gourley, J. Onnela, and G. Reinert. 2008. Bias in epidemiological studies of conflict mortality. *Journal of Peace Research* 45 (5): 653-63.
Laaksonen, S. 2008. Retrospective two-stage cluster sampling for mortality in Iraq. *International Journal of Market Research* 50 (3): 403-17.
Lacina, B. A., and N. P. Gleditsch. 2005. Monitoring trends in global combat: A new dataset of battle deaths. *European Journal of Population* 21 (2-3): 145-66.
Lacina, B. A., N. P. Gleditsch, and B. M. Russett. 2006. The declining risk of death in battle. *International Studies Quarterly* 50 (3): 673-80.
Obermeyer, Z., C. J. L. Murray, and E. Gakidou. 2008a. A half-century of violent war deaths from Vietnam to Bosnia: Analysis of data from the World Health Survey Programme, technical appendix. *Institute for Health Metrics and Evaluation.* http://www.healthmetricsandevaluation.org/print/articles/2008/WarDeaths-TechnicalAppendix.pdf.
Obermeyer, Z., C. J. L. Murray, and E. Gakidou. 2008b. Fifty years of violent war deaths from Vietnam to Bosnia: Analysis of data from the world health survey programme. *British Medical Journal* 336 (June 19). http://www.bmj.com/cgi/content/full/bmj.a137.
Roberts, L., R. Lafta, R. Garfield, J. Kudhairi, and G. Burnham. 2004. Mortality before and after the 2003 invasion of Iraq: Cluster sample survey. *The Lancet* 364 (9448): 1857-64.
Rosenblum, M., and M. J. van der Laan. 2008. Confidence intervals for the population mean tailored to small sample sizes, with applications to survey sampling. *U.C. Berkeley Division of Biostatistics Working Paper Series.* Working Paper 237. June. http://www.bepress.com/ucbbiostat/paper237/.
Sarkees, M. R., F. W. Wayman, and J. D. Singer. 2003. Inter-state, intra-state, and extra-state wars: A comprehensive look at their distribution over time, 1816-1997. *International Studies Quarterly* 47:123-44.
Sloboda, J., J. Dougherty, and H. Dardagan. 2007. How accurate is IBC? In *Iraq Body Count: The state of knowledge on civilian casualties in Iraq: Counts, estimates, and government responsibility.* Slides for presentation made to the United States Institute of Peace in Washington, DC, on January 10, 2007. http://www.iraqbodycount.org/analysis/beyond/state-of-knowledge/7.
Spagat, M. 2009. Ethical and data-integrity problems in the second *Lancet* survey of mortality in Iraq. *Defence and Peace Economics*. http://personal.rhul.ac.uk/uhte/014/Research.htm (accessed January 22, 2009).
Spagat, M., A. Mack, T. Cooper, and J Kreutz. 2009. Estimating War Deaths: An Arena of Contestation. Unpublished manuscript.
Standardized Monitoring and Assessment of Relief and Transitions (SMART). 2006. *Measuring mortality, nutritional status, and food security in crisis situations: SMART methodology.* Version 1. June. http://www.smartindicators.org/SMART_Methodology_08-07-2006.pdf (accessed July 25, 2009).
United States Government Accountability Office. 2006. Report to congressional requesters. *Darfur crisis: Death estimates demonstrate severity of crisis, but their accuracy and credibility could be enhanced.* http://www.gao.gov/new.items/d0724.pdf.
Urlacher, B. R. 2009. Wolfowitz conjecture: A research note on civil war and news coverage. *International Studies Perspectives* 10 (2): 186-97.