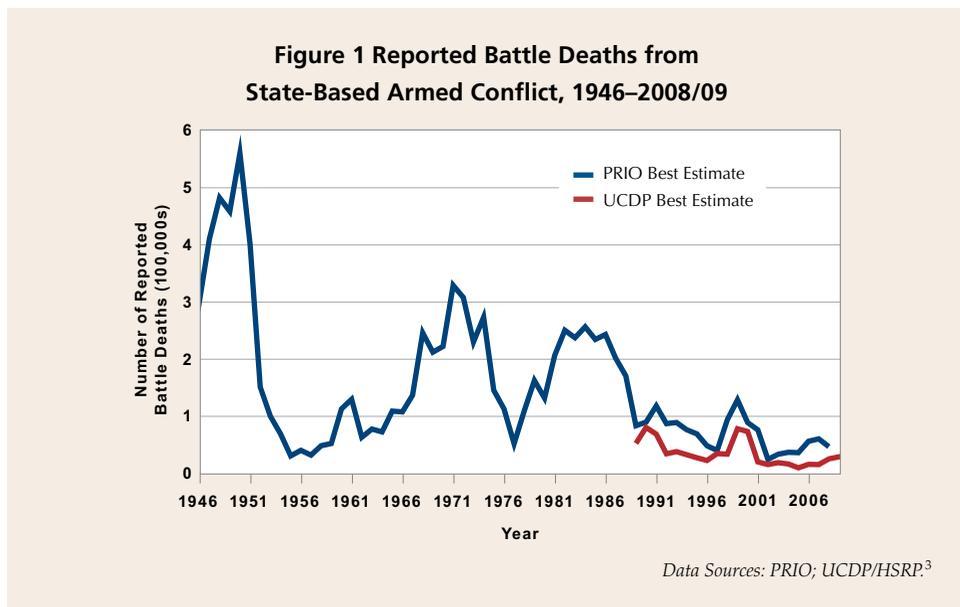


# Estimating Battle Deaths: A Challenging Exercise<sup>1</sup>

There are currently two datasets that provide estimates of the number of worldwide battle deaths—i.e., combat-related deaths—that occur in state-based armed conflicts.<sup>2</sup> (*State-based armed conflicts* are those in which at least one of the warring parties is the government of a state). Battle deaths include not only combatants but also civilians caught in the crossfire—deaths that are often referred to as *collateral damage*.

Uppsala Conflict Data Program (UCDP) has battle-death data for the period from 1989 to 2010 and the dataset is updated annually. Peace Research Institute Oslo (PRIO), on the other hand, currently has data from 1946 to 2008. This dataset is updated periodically.

The trends in battle deaths described by the two datasets are quite similar as Figure 1 shows, although the PRIO “best estimates” of worldwide battle deaths are higher than those of UCDP—several times higher in some years. (This can be seen more clearly in Figure 2, below.) As Figure 1 demonstrates, there has been a clear long-term, but highly uneven, decline in battle deaths since 1946.



The methods used by UCDP and PRIO are similar, but to understand why PRIO’s best estimates are mostly greater than those of UCDP, we need to know how each dataset is compiled.

## The Datasets

The PRIO dataset was originally created to determine if there were any long-term trends in the deadliness of armed conflict in the post-World War II era. The starting point for the coding of battle deaths was the list of armed conflicts from 1946 contained in the UCDP/PRIO Armed Conflict Dataset.<sup>4</sup> The coding was undertaken by Bethany Lacina.

The first iteration of the dataset was released in 2005 and covered the period 1946–2002.<sup>5</sup> The dataset has been updated twice, in 2006 and 2009, and the latest version covers the period 1946–2008. As of 12 March 2012, no final decision had been made on the next update.

UCDP, unlike PRIO, collects data on three types of organized violence. In 2005 data from all three new UCDP datasets were published for the first time in the *Human Security Report*. In addition to the battle-death data for state-based conflicts, UCDP introduced two new datasets—one on non-state battle deaths, and one on deaths from one-sided violence. Subsequently, each dataset has been updated annually—currently to 2010. Moreover, all three UCDP datasets have now been backdated to 1989.

Both the PRIO and UCDP provide “low,” “high,” and “best” estimates of fatality numbers for each year.<sup>6</sup>

## Summary Estimates versus Incident Counting

PRIO and UCDP both use a wide range of sources in compiling their estimates. But there is one major difference between their approaches—one that helps explain why PRIO’s estimates are generally higher than those of UCDP.

The UCDP dataset is compiled primarily by counting the annual total of combat-related fatalities (national and global) from reports of fatalities in individual violent incidents (battles, clashes, etc.) in each state-based conflict being waged around the world. To do this, UCDP uses a variety of sources, including news reports and on-the-ground reports from human rights organizations, local NGOs (nongovernmental organizations), etc.

This approach not only provides very detailed information about particular violent events, but it also gives UCDP researchers considerable confidence that they have a reliable estimate of the *minimum* number of battle deaths in a conflict. But since it is highly unlikely that all reports of battle deaths will be recorded—particularly in conflicts where outside observers are banned from war zones—this methodology will almost certainly underestimate the actual number of battle deaths.<sup>7</sup>

By contrast, the PRIO dataset relies heavily on *summary estimates*—i.e., expert assessments of overall fatalities. There is no reason to assume that summary estimates will systematically undercount battle deaths as does UCDP’s incident-based estimation method.<sup>8</sup>

In producing estimates of battle deaths for the early decades of the post-war era, PRIO researchers had no choice but to rely on summary reports of war deaths—which were typically estimates of how many people had been killed over the course of an entire conflict.<sup>9</sup> During these early decades, there was nothing remotely like the Factiva database, with its 35,000-plus

news sources, that among other sources has enabled UCDP to undertake electronic searches for reports of individuals killed in battle.

Second, as noted earlier, UCDP data coders disaggregate deaths from organized violence into several categories. This means, for example, that UCDP will not count fatalities from intentional one-sided violence against civilians in its battle-death category, since these fatalities are collected in a separate dataset. However, the summary estimates of war deaths on which PRIO relies may well include the intentional killing of civilians along with battle deaths, because their sources do not always distinguish between the two types of violence.<sup>10</sup> Where PRIO battle-death tolls include civilian deaths that UCDP counts separately, the effect will again be to increase the PRIO toll relative to that of UCDP.<sup>11</sup>

Third, UCDP's stringent coding rules mean that violent deaths will not be recorded as battle deaths unless the identity of the perpetrators is known. But sometimes it is not possible to identify the perpetrators, and where this is the case, the fatalities are not recorded. This helps explain why UCDP's battle-death counts in Iraq are lower than those of other sources, and is yet another reason why PRIO's best-estimate death tolls are higher on average than UCDP's.<sup>12</sup>

Fourth, another difference between the two datasets arises from the fact that UCDP updates its data annually. UCDP researchers may well find, and include, important new sources of data for early conflict periods that were not available at the time to PRIO researchers. The PRIO dataset cannot include these data until its next periodic update.

These four differences account for most of the variation between PRIO and UCDP battle-death estimates.<sup>13</sup>

## Conclusion

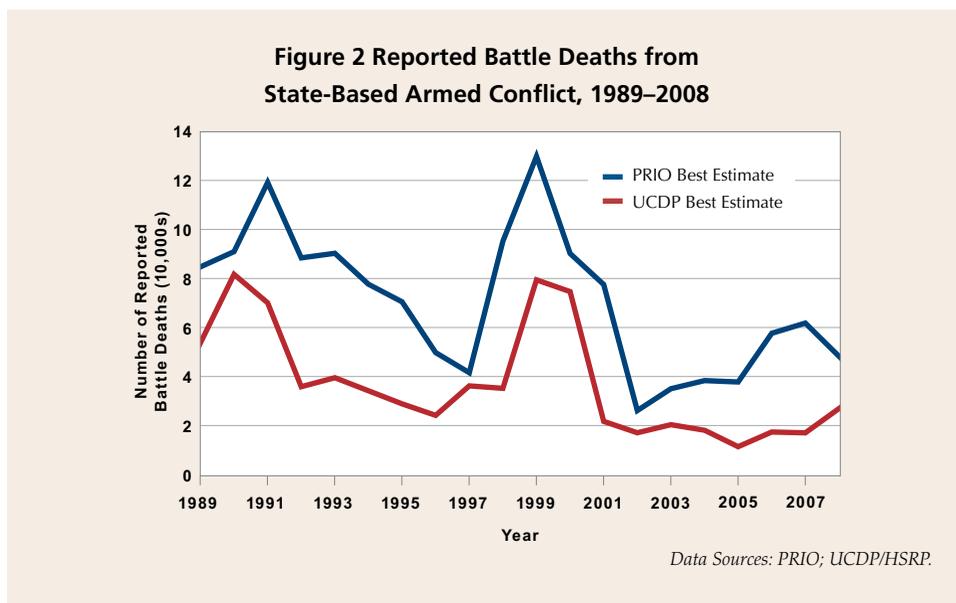
Although the differences in battle-death counts between the two datasets are often considerable, the overall trends track each other reasonably closely as Figure 2 (next page) indicates. Both datasets show a clear, though very uneven, decline in battle deaths over the past two decades, the period in which the two datasets overlap. If different data collection methodologies generate data that follow similar trends, then we can be reasonably confident that the trends are real.

This is important because the most prominent—and most contested<sup>14</sup>—finding that has emerged from more than six decades of PRIO battle-death data is that there has been a long-term, but uneven, secular decline in deadliness of warfare around the world.<sup>15</sup>

Notwithstanding their differences—and in part because of them—the two datasets are clearly complementary. Only the PRIO dataset has trend data in battle deaths for the first four decades of the post-World War II era, but only UCDP provides updates and revisions every year. This means that UCDP's data are usually more up to date than PRIO's and thus of particular interest to policy-makers. Plus, UCDP provides separate, but compatible, datasets on violent deaths from one-sided violence and non-state conflicts. Because of their incident-based method of data collection, UCDP is not only able to differentiate between distinct types

of organized violence but the data can also provide more detailed information on where and when the deaths occurred.

As pointed out below, a primary objective of the PRIO dataset is to provide data on long-term historical trends in the number of battle deaths, while the UCDP dataset focuses on recent time periods and allows for more detailed and disaggregated analysis of violent events. Users must take this into account when choosing the appropriate dataset for a specific research question.



Policy-makers and other researchers can be confident that UCDP’s data accurately reflect the minimum number of battle deaths that occur in each conflict each year, but its overall estimates are likely to undercount the true extent of the death toll for reasons spelled out above—and by UCDP itself.

PRIO’s data, on the other hand, not only cover six-plus decades of conflict but will tend to provide higher, and often more realistic, estimates of overall death tolls than UCDP’s battle-death counts, precisely because PRIO’s methodology is not affected by the factors that tend to reduce battle-death estimates in the UCDP dataset. But it is also possible that PRIO’s estimates, which are based in large part on judgments about the accuracy of sources, rather than on counting reports of deaths in individual incidents, will overestimate the extent of battle deaths in some cases. It is highly improbable that this would happen with the UCDP estimates.

We should also be clear that these datasets are neither intended—nor suitable—for the sort of painstaking in-depth investigations of human rights violations and deaths from organized violence that Benetech’s Human Rights Data Analysis Group (HRDAG) undertakes—often for truth and reconciliation commissions in post-conflict settings.<sup>16</sup>

However, HRDAG’s intensive investigations often take multiple years to complete. They are essentially one-off exercises and are only carried out in a minority of war-affected countries

in each year. As such, they cannot be used for mapping global *trends* in organized violence on an ongoing basis.

For researchers and policy-makers interested in understanding trends in death tolls from organized violence and what drives them, and in determining the impact of violence-reduction strategies, there are no substitutes for the trend data revealed by the PRIO and UCDP battle-death datasets. The datasets serve different but complementary purposes.

# ENDNOTES

- 1 The discussion about the differences between the Peace Research Institute Oslo (PRIO) and Uppsala University's Conflict Data Program (UCDP) datasets on battle deaths from state-based armed conflict was initiated by Gerdis Wischnath and Nils Petter Gleditsch, with subsequent input from UCDP and the Human Security Report Project (HSRP). This overview, prepared by HSRP, reflects the views of all three institutions.
- 2 State-based conflicts from 1946 to 2010 are recorded in the now widely used UCDP/PRIO Armed Conflict Dataset.
- 3 Centre for the Study of Civil War, International Peace Research Institute, Oslo, (PRIO), Battle Deaths Dataset 3.0, <http://www.prio.no/CSCW/Datasets/Armed-Conflict/Battle-Deaths/The-Battle-Deaths-Dataset-version-30/> (accessed 14 March 2012), updated from Bethany Lacina and Nils Petter Gleditsch, "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths", *European Journal of Population* 21, no. 2–3 (2005): 145–166; Uppsala Conflict Data Program (UCDP), Uppsala University, Uppsala, Sweden/ Human Security Report Project, School for International Studies, Simon Fraser University, Vancouver, Canada.
- 4 See Nils Petter Gleditsch, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand, "Armed conflict 1946–2001: A new dataset," *Journal of Peace Research* 39, no. 5 (2002): 615–637, and the two websites, [http://www.pcr.uu.se/research/ucdp/datasets/ucdp\\_prio\\_armed\\_conflict\\_dataset/](http://www.pcr.uu.se/research/ucdp/datasets/ucdp_prio_armed_conflict_dataset/) and <http://www.prio.no/CSCW/Datasets/Armed-Conflict/UCDP-PRIO/> (accessed 27 April 2012).
- 5 Bethany Lacina and Nils Petter Gleditsch, "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths," *European Journal of Population* 21, nos. 2–3 (2005): 45–166. One version of the PRIO dataset has data going as far back as 1900. See Bethany Lacina, Nils Petter Gleditsch, and Bruce Russett, "The Declining Risk of Death in Battle," *International Studies Quarterly* 50, no. 3 (2006): 673–680. The replication data are found at <http://www.prio.no/cscw/datasets> (accessed 27 April 2012).
- 6 More information on the coding of the low, high, and best estimates for each dataset can be found by consulting the respective PRIO and UCDP codebooks, available on the websites cited above.
- 7 For a description of UCDP's data collection methodology, see UCDP, "How are UCDP data collected?" [http://www.pcr.uu.se/research/ucdp/faq/#How\\_are\\_UCDP\\_data\\_collected\\_](http://www.pcr.uu.se/research/ucdp/faq/#How_are_UCDP_data_collected_) (accessed 27 April 2012).
- 8 Summary estimations of battle deaths may be based in part on body-count data from the warring parties. Typically, both sides will tend to minimize their own casualties and maximize those of their enemies. Reliance on such data is not likely to lead to any systematic upward or downward bias, however.
- 9 To produce annual battle-death estimates in these cases, PRIO had little choice but to divide estimates for the entire period of the war by the number of years it had lasted. The researchers recognized, of course, that in reality, the probability that each year of conflict would have the same number of battle deaths was low. This is another cause of the difference between PRIO's annual estimates and UCDP's—the latter are always based on incident data of reported battle deaths.

- 10 They may also include some *non-state* conflict deaths—fighting between rival militias, for example, which would be counted separately by Uppsala.
- 11 Conflicts that are likely to include a high level of one-sided violence are noted in the documentation for the PRIO dataset.
- 12 This is an even greater problem with estimating deaths from one-sided violence, since few armed groups are likely to boast about killing civilians.
- 13 There are additional technical issues that we do not elaborate on here. Most notably, PRIO researchers do not publish best estimates for conflict years where they lack reliable information based on their set of sources. Using an average of high and low estimates to replace the missing best estimates as we do in Figure A.2 exaggerates the differences between PRIO and UCDP. Note also that the PRIO and UCDP battle-death data currently available are based on different versions of the UCDP/PRIO Armed Conflict Dataset, which means that some conflict years are coded in one, but not the other, battle-death dataset.
- 14 See Z. Obermeyer, C. J. L. Murray, and E. Gakidou, “Fifty years of violent war deaths from Vietnam to Bosnia: analysis of data from the world health survey programme,” *British Medical Journal* 336, no. 7659 (2008): 1482. See rebuttal by Michael Spagat, Andrew Mack, Tara Cooper, and Joakim Kreutz, “Estimating War Deaths: An Arena of Contestation,” *Journal of Conflict Resolution* 53, no. 6 (2009): 934–950. See also M. R. Sarkees, F. W. Wayman, and J. D. Singer, “Inter-State, Intra-State, and Extra-State Wars: A Comprehensive Look at Their Distribution over Time, 1816-1997,” *International Studies Quarterly* 47, no. 1 (2003): 49–70. See rebuttal by Bethany Lacina, Nils Petter Gleditsch, and Bruce Russett, “The Declining Risk of Death in Battle,” *International Studies Quarterly* 50, no. 3 (2006): 673–680. See also Anita Gohdes and Megan Price, “First things first: Assessing data quality before model quality,” *Journal of Conflict Resolution*, forthcoming (a response by Bethany Lacina, Nils Petter Gleditsch, and Bruce Russett will be published in the same issue).
- 15 See Steven Pinker, *The Better Angels of Our Nature* (New York: Viking, 2011); Joshua Goldstein, *Winning the War on War: The Decline of Armed Conflict Worldwide* (New York: Dutton, 2011); and HSRP, *Human Security Report 2009/2010: The Causes of Peace and the Shrinking Costs of War* (New York: Oxford University Press, 2011).
- 16 See HRDAG, “Projects,” <http://www.hrdag.org/about/projects.shtml> (accessed 30 April 2012).