

WikiLeaks and Iraq Body Count: the Sum of Parts May Not Add Up to the Whole—A Comparison of Two Tallies of Iraqi Civilian Deaths

Dustin Carpenter, MD, MPH;¹ Tova Fuller, MD, MS;¹ Les Roberts, PhD, MSPH²

1. Columbia University Medical Center, New York, New York USA
2. Mailman School of Public Health, Columbia University, New York, New York USA

Correspondence:

Les Roberts, PhD, MSPH
Mailman School of Public Health
Columbia University
60 Haven Ave., B-4
New York, NY 10023 USA
E-mail: lfr2102@columbia.edu

Conflicts of interest: There was no funding for this research. No author has any financial or other competing interests that are related to the findings or contents of this study.

Keywords: armed conflict; deaths; epidemiology; Iraq; mortality; passive surveillance; sensitivity analysis; under-reporting; violence

Abbreviations:

IBC: Iraq Body Count
WL: War Logs

Received: November 26, 2012

Accepted: December 2, 2012

Revised: January 12, 2013

Online publication: February 7, 2013

doi:10.1017/S1049023X13000113

Abstract

Introduction: The number of civilians killed in Iraq following the 2003 invasion has proven difficult to measure and contentious in recent years. The release of the Wikileaks War Logs (WL) has created the potential to conduct a sensitivity analysis of the commonly-cited Iraq Body Count's (IBC's) tally, which is based on press, government, and other public sources.

Hypothesis: The 66,000 deaths reported in the Wikileaks War Logs are mostly the same events as those previously reported in the press and elsewhere as tallied by iraqbodycount.org.

Methods: A systematic random sample of 2500 violent fatal War Log incidents was selected and evaluated to determine whether these incidents were also found in IBC's press-based listing. Each selected event was ranked on a scale of 0 (no match present) to 3 (almost certainly matched) with regard to the likelihood it was listed in the IBC database.

Results: Of the two thousand four hundred and nine War Log records, 488 (23.8%) were found to have likely matches in IBC records. Events that killed more people were far more likely to appear in both datasets, with 94.1% of events in which ≥ 20 people were killed being likely matches, as compared with 17.4% of singleton killings. Because of this skew towards the recording of large events in both datasets, it is estimated that 2035 (46.3%) of the 4394 deaths reported in the Wikileaks War Logs had been previously reported in IBC.

Conclusions: Passive surveillance systems, widely seen as incomplete, may also be selective in the types of events detected in times of armed conflict. Bombings and other events during which many people are killed, and events in less violent areas, appear to be detected far more often, creating a skewed image of the mortality profile in Iraq. Members of the press and researchers should be hesitant to draw conclusions about the nature or extent of violence from passive surveillance systems of low or unknown sensitivity.

Carpenter D, Fuller T, Roberts L. WikiLeaks and Iraq Body Count: the sum of parts may not add up to the whole—a comparison of two tallies of Iraqi civilian deaths. *Prehosp Disaster Med.* 2013;28(3):223-229.

Introduction

On October 22, 2010, WikiLeaks released US classified military records, which included descriptions of approximately 66,000 Iraqi civilian deaths. There have been competing estimates of civilian deaths in Iraq, with different sources at the center of the debate at distinct moments in time. Since the beginning of the war, these estimates have tended to fall into two categories. The first category comprises passive monitoring systems, which use sources such as Iraqi government records, press reports, or these sources plus a corrective factor.¹⁻³ Reports based on passive monitoring have included a range of estimates that generally cluster around 50,000 deaths by mid-2006, and >100,000 deaths as of October 2010. These estimates contrast with the second category of estimates, population-based surveys and polls. Population-based surveys and polls have generated violent death estimates up until some specific point in time that were roughly three, six, 10, and 20 times the passive estimates.⁴⁻⁷

The *New York Times* was given exclusive US pre-release review rights to the so-called WikiLeaks "Iraq War Logs" released on October 22, 2010. The *Times* was careful not to attribute these findings to supporting a particular civilian death total. But, many

US newspapers ran an AP wire service article that stated: "The U.S. military has recorded just over 66,000 civilian deaths, according to the documents posted by WikiLeaks. Iraq Body Count (IBC), a private, British-based group that has tracked the number of Iraqi civilians killed since the war began, said in a press release that it had analyzed the information and found 15,000 previously unreported deaths, which would raise its total from as many as 107,369 civilians to more than 122,000 civilians. The Iraqi government has issued a tally claiming at least 85,694 deaths of civilians and security officials were killed [*sic*] between January 2004 and Oct. 31, 2008."⁸ This statement suggested that there was independent consistency among the reports, and perhaps even validation of 2010 estimates in the zone of 100,000 civilian deaths in Iraq. Many papers such as the *Washington Post* included statements in their coverage such as, "There appear to be no major revelations in the latest logs."⁹ On October 26, 2010, the *Post* ran an editorial entitled "WikiLeaks' leaks mostly confirm earlier Iraq reporting."¹⁰ The editorial concluded "...claims such as those published by the British journal *The Lancet* that American forces slaughtered hundreds of thousands are the real 'attack on truth.'" Thus, the *Washington Post* explicitly, and other papers implicitly, sided with the narrative of Iraqi death tolls created by passive surveillance sources such as IBC.

Over several months, a group from the Columbia University School of Public Health (New York, New York USA) evaluated IBC's finding that the 66,000 civilian deaths reported by WikiLeaks only included approximately 15,000 new death reports, and that by implication three-quarters of the deaths in WikiLeaks were already reported and known through the IBC online database. Such analysis is crucial to evaluating the ever-growing tendency to treat news reports, morgue tallies, and other passive sources as "data" useful in understanding population-wide trends or patterns.¹¹

Methods

All analysis was conducted using data publicly available on the Web. No names or medical information were recorded or used in this analysis. These conditions exclude the report from ethical review requirements by the Columbia University IRB.

In order to compare the overlap between the Wikileaks War Logs released in October 2010 and the IBC tally that existed at that time:

1. Iraq War Log files were downloaded on October 23, 2010 from the website of The *Guardian*, a UK newspaper. The data from the WL was in MS Excel (Microsoft Corp., Redmond, Washington USA) format. Those files are available at <http://www.guardian.co.uk/news/datablog/2010/oct/23/wikileaks-iraq-data-journalism>.
2. On October 25, 2010, the IBC data file of incidents was downloaded in an attempt to preserve a file uninfluenced by the WikiLeaks data.
3. The WikiLeaks data file was filtered to contain only civilian deaths. The 33,809 entries contained reports of 65,679 civilian deaths. These deaths were recorded as "killing events." A killing event could include one or several deaths. Bombings, for example, generally included multiple deaths, while shooting events frequently entailed one death.
4. Initially, each of 25 students in a graduate class on epidemiological methods for documenting human rights abuses was systematically (based on the date field) assigned 100 killing event data entries to review and determine if the killings described in those reports were: (a) included in IBC's dataset; (b) not included; or (c) possibly included. Only 23 students fulfilled the initial review.
5. After the initial evaluation of 2,300 data records, the participants developed a set of decision rules to standardize the matching process. Participants were asked to code each WikiLeaks killing event record match probability to a corresponding incident in IBC as: 0 = no match, 1 = unlikely, 2 = likely, or 3 = high probability. To be coded above 0, the following was required:
 - a. two events must not have had contradictory locations or have been contradictory types of events (in broad terms, shootings, torture, airstrikes, or explosive devices were different types of events).
 - b. IBC killing event reports could not be dated later than two days after the WikiLeaks event time. Moreover, IBC reports could not be more than 24 hours before the WikiLeaks event time. This delay was allowed because the press reporting process might have delayed the arrival of reports in IBC. However, since IBC data are mostly based on press reports, it would be unlikely for them to have dates prior to those recorded in the WikiLeaks data, which had a military source and a specific hour of each event recorded. If there were multiple events in WikiLeaks that corresponded to multiple events in IBC, then participants were to match only the number that appeared in the lowest report. If, for example, there were four different shooting events in one neighborhood in Baghdad recorded in WikiLeaks, and only three potential matches in IBC, then only three killing event matches could be recorded. If there was a disparate number of reported deaths in a specific killing event (eg, 11 deaths in one report but 12 in another) where the event appeared to be a likely match, then only the lower number of deaths could be considered matched or likely matched based on the reviewer's judgment (eg, if one report said 17 killed and another said nine killed, the discrepancy was considered by the reviewer when assessing the likelihood of a match).
6. To aid with the matching process, Google Earth (Mountain View, California USA) was utilized and a Google Earth template was developed to identify specific neighborhoods in Baghdad.
7. With the Google Earth tool and the standardized matching rules, all students started over and matched their assigned records in the War Logs to IBC's data. Once the dataset had been examined by an individual, a second individual who was blinded to the first evaluator's identity or matching score (all 1, 2, or 3 scores were converted to "M") was given the dataset and asked to give his/her own match ranking (0, 1, 2, or 3) to any entries that were coded above 0 by the original evaluator. This was done to evaluate the consistency of the matching process.
8. A third and final arbiter was asked to review those cases where match rankings given by the first and second reviewers differed. The arbiter was not permitted to move the average rankings by two units (eg, rankings of 2 and 3 on the first two rounds could not be changed to a 0). If the arbiter felt that the score was off by two units or more, a fourth reviewer had to agree. At the third arbitration review, to receive a 3 ranking, the date of the event could

not differ by more than one day between the two reports. Thus, the final data file involved data matches reviewed by at least two people, and if the first two did not exactly agree, the ranking was based on the judgment of at least three people. Reports that might only potentially match to the morgue tallies that were in IBC (IBC contained morgue tallies with no information about the individuals included therein) were given a score of 0 or 1 based on the reviewers' judgment. Previous evidence suggests that most violent deaths in Baghdad did not go to a morgue.¹²

9. An unintended natural experiment occurred when the War Logs file was imported into a US version of Excel. The software re-sorted the Iraq War Logs records by day such that all killing events on January 1 were listed first regardless of year, and the date was converted from the UK day/month/year format to the US month/day/year ordering. Because reviewers initially were told incorrectly the date field was in the UK format, and one-third of the reviewers had a day between the 1st and the 12th of the month (so the reversal of the day and month order was undetectable), some of these initial reviewers attempted to match to the wrong date. These mismatches were used to gain insight into the rate of false positives, but the data were not further used or explored.
10. Some data fields in the War Logs were incomplete and did not allow proper matching. Others involved some question as to whether the victims were civilians or combatants, or whether the event was an accident or intentional violence. When there was doubt, the data record was excluded.
11. Using a random start number, a systematic sample of 10 locations within the reviewed data was obtained. At each interval, 20 negative (match = 0) records were examined to determine the fraction of false negatives in the dataset. Upon review by outside readers, this initial analysis proved inadequate because two of the graduate students had very high rates of false negatives, and the systematic skip through one-tenth of the data set had skipped over most of their entries. After reassessing those two sets of 100 Wikileaks records (eight percent), 150 unmatched records (match = 0) were again selected by a systematic random sample. Those events where the reviewer deemed the record merited a score > 0 were independently assessed by a second reviewer. The final match number was determined by a third reviewer in those cases where the first two reviewers could not agree.
12. Match rates were defined as the number with a specific score divided by all fully evaluated reports. Match rates by the number of civilian casualties associated with an event was achieved by filtering the War Logs data set by casualty number and dividing the number with each score by the total number in that group. Total number of deaths matched by score was obtained by filtering the dataset by match number and dividing the total deaths with a specific match number by the total number of deaths.
13. An attempt was made to compare the two datasets without any human judgment involved. Because of the incompatibilities between the data sets, a complete automated comparison would require major reformatting that in itself would incorporate human judgment, and may have added little to the assessment. For example, the WL always provide the military sector (which typically included three provinces) where an incident occurred, while IBC sometimes listed

Match Score	Incidents n (%)	Of All Deaths n (%)
0 No match	1306 (63.7%)	1910 (43.5%)
1 Unlikely match	255 (12.4%)	449 (10.2%)
2 Likely match	202 (9.9%)	455 (10.4%)
3 High probability match	286 (14.0%)	1580 (36.0%)

Carpenter © 2013 Prehospital and Disaster Medicine

Table 1. Fraction of Incidents and Deaths Matched Between WL and IBC by Certainty of Match

Abbreviations: IBC, Iraq Body Count; WL, Wikileaks War Logs

a town or city and sometimes a province. To explore if a non-subjective matching would lead to different conclusions, Arbil- (also spelled Irbil) and Karbala-based incidents were explored; these two areas were uniquely easy to match because the Province and the urban provincial capital had the same name. Fields were matched by three different rules: by date (+/- one calendar day), date and number dead (+/-30%), and date and number dead and type of weapon used. This algorithm-based matching involved little to no human judgment. In both cases, the entire IBC data file was filtered for the Provinces of "Arbil" and "Karbala" and then matched by the three criteria to see if the incident was in WL. Then the process was reversed by filtering the entire WL data set in the event description for the appearance of the words "Irbil" and "Karbala." Thus, this analysis is on the complete datasets and not related to the sampling described above.

14. Data were analyzed in Microsoft Excel 2010 and with R Version 2.15.2 (The R Foundation for Statistical Computing, Vienna, Austria. Matching frequencies were presented by match number, and for most purposes, a score of 0 or 1 was considered not likely to be a match and scores of 2 and 3 were seen as likely to be in both datasets.

Results

In total, 2409 event records were examined by the two or more reviewers process. An additional 88 reports (3.5%) were unmatchable because they lacked information such as the date, or because they were deemed to have been an accident and not conflict related. By any measure of match (scores of 1, 2, or 3), the majority of events described in the WL were not included in IBC database. The fraction of all WL records that were matched to records in IBC is shown in the second column of Table 1. It is likely that 488 (23.8%) of records matched (ie, had a score of 2 or 3). Two hundred fifty-five (12.4%) of records had a small chance of being listed in IBC (ie, lower evidence with a score of 1). These were primarily single killings in Baghdad. Nearly two-thirds of records, 1306 (63.7%), were judged as not being in IBC (score of 0).

The fraction of total deaths (not reports of killing events) reported in WL that were also in IBC was also assessed, as shown in the third column of Table 1. Two thousand three hundred fifty-nine (53.7%) of deaths were not included in IBC, more than the overall fraction of WL, because the larger killing events were likely to be recorded by both listings. The majority of reports in both WL and IBC involved one or two deaths.

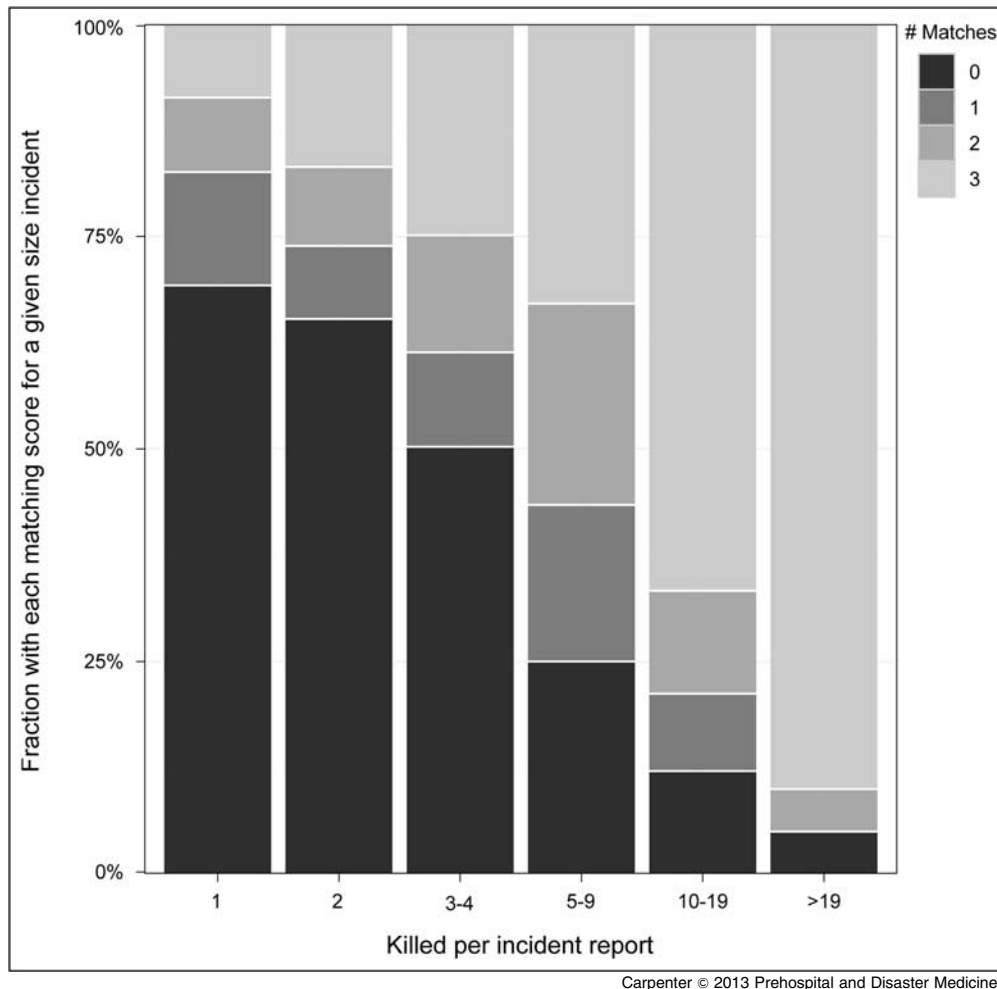


Figure 1. Match Score Distribution by Number of People Killed per Incident Report

The fraction of events that match with each score by the different number of civilians reportedly killed is shown in Figure 1. Events that killed many people were far more likely to be reported in both datasets than were killing events that involved few people; this finding is evidenced by the high fraction of matches with scores of ≥ 2 among the larger killing events. While single killing events represented 59.8% of events, only 32.8% of all deaths were singletons. Two hundred fifty-one (17.4%) of these 1440 single killing events were judged to be in IBC as evidenced by a score of 2 or 3, as opposed to 38.9% of multiple killing events. Number of deaths per WL event was highly correlated with likelihood of being in IBC; for example, 86.2% of events with ≥ 10 deaths were matched to a record in IBC, and 94.1% of events with ≥ 20 deaths were judged to be in IBC.

In the preliminary review, before decision rules and mapping tools were developed, among the 822 records that were matched to the wrong date, 72 (8.8%) were matched (eg, a score of 3) and 132 (16.1%) were matched with a score of 2. In this same initial round, 135 of 515 records (26.2%) matched to the correct date were initially given a score of 3, and 43 (8.3%) a score of 2. The process of using the Google Earth map, having more clarified matching rules, and the process of having three reviewers instead of one lowered the match rate (match of 3 or 2) from 34% to 19%. In a best-case scenario, assuming the more rigorous procedures would have caught almost all of the matches of score 2 which were

attributed to the wrong date, it is likely that several percent, perhaps six to eight percent, of the WikiLeads reports which were judged to be in IBC are in fact not, but are false positive matches. Given the different language and geographic delineations used between the two lists, especially within the City of Baghdad, it is exceedingly difficult to ensure there are no false positive matches.

Because the triple review process was only applied to the initial positive matches, and because the work of two initial reviewers (eight percent of dataset) needed to be redone, an assessment of the false negative (match = 0) rate was undertaken. Of the 150 non-matches checked by two independent reviewers, four (2.7%) were given a score of 3, three (2.0%) were given a score of 2, and 13 (8.7%) were given a score of 1. This suggests a likely false negative match rate of 4.7%.

Colleagues suggested conducting a computer-driven match between the two datasets as a comparator to the multiple reviewer process. One location in the Kurdish North (Arbil) and one in the South (Karbala) were chosen because of the ease of filtering the datasets where the Province and the major city had the same name. The results are presented in Table 2.

A non-subjective matching process resulted in the same conclusions: the majority of events in WL are not in IBC and vice-versa. This exploration also suggests that both datasets are more complete in the less violent pro-Coalition North than in Karbala. These two provinces had similar populations

Filter of Files	Match by Date, Deaths, Weapon Type n (%)	Match by Date and Deaths (+/- 30%) n (%)	Match by Date Only (+/- 1 Day) n (%)	Did Not Match by Any Criteria n (%)
IBC for Arbil, n = 8	6 (75%)	0	1 (12.5%)	1 (12.5%)
War Logs for Irbil, n = 10	6 (60%)	1 (10%)	1 (10%)	2 (20%)
IBC for Karbala, n = 136	13 (10.3% ^a)	1 (0.8% ^a)	27 (21.4% ^a)	95 (69.9%)
War Logs for Karbala, n = 39	1 (2.6%)	4 (10.3%)	2 (5.1%)	32 (82.1%)

Carpenter © 2013 Prehospital and Disaster Medicine

Table 2. 2004–2009 Criteria-based Computer Matching for Arbil (Irbil) and Karbala^aDenominator for this calculation excludes 10 guerre tallies.

	In WikiLeaks	Not in WikiLeaks	Total
In IBC	3,752 (6.3% ^a)	6,623 (11.1% ^a)	10,375
Not in IBC	17,772 (29.9% ^a)	31,374 (52.7% ^a)	49,146 ^c
Total	21,524	37,997 ^c	59,521 ^c

Carpenter © 2013 Prehospital and Disaster Medicine

Table 3. Overlap of Single Killing Events

Abbreviations: IBC, Iraq Body Count

^aDenotes percent of speculative total if these were independent samples.^bTotal if these were independent samples = 3752 + 6,623 + 17,772 + 31,374 = 59,521 singleton deaths.^cDenotes a total that would be true if the samples were independent, which is likely not the case.

(1.10 million vs 1.05 million) at the start of the conflict.⁶ This is consistent with the notion that passive surveillance is less sensitive in more violent places. Of note, in both places, more IBC events are in the WL than are WL events in IBC.

Discussion

There are several serious limitations to the ability of any reviewer to determine which events described in the WL are included in IBC. Some locations are described multiple ways or with different spellings. The two sources do not record the same data fields; for example, WL has a field for the kind of weapon used, which may or may not be in an IBC report. Moreover, it is possible that there are recording errors and false reports in both data sets and outside reviewers cannot assess this easily. It is estimated that approximately 4.7% of the “not match” 0 scores should have been a 2 or 3 score, ie, are false negatives. Estimating the false positive match rate in our data is more difficult. This is because the positive matches may have had two, three, four, or even five people’s judgment contributing to the rating. Thus, having one more arbiter look at positive matches would contribute different levels of additional quality control and varying certainty that any one match is a false positive.

In spite of the imprecision in matching some events, this analysis of the WikiLeaks Iraq WL suggests that most of the WL reports were not in IBC’s database. In Karbala Province, most events do not even have a potential corresponding event in the other dataset within a day either side of the incident. The widespread dissemination of a contrary message in October 2010 suggests many in the press were unable to digest detailed reports or perform detailed analysis. Thus, while this analysis represents hundreds of hours of work, a brief check of a handful

of reports (outside of Baghdad where matching is more time consuming) would have shown any reporter writing about the WikiLeaks releases that most of these violent reports were previously unreported; the imprecision of the matching process utilized in the current study is unlikely to call that conclusion into question.

It may not be possible to estimate a total death toll from these two sources because it is not possible to assess the degree to which the two datasets are correlated. If one assumes these two sources are independent, one might apply the logic of capture-recapture analysis to assess the fraction missed by both. Figure 1 suggests that both datasets disproportionately capture large events and thus are not independent. The fact that both datasets are more complete in Irbil and less complete in more violent Karbala suggests they are not independent. Borrowing from the logic of capture-recapture approaches for discussion’s sake, the data in Table 3 is based on the assumption that single killing events are independent between the two datasets.

Table 3 is constructed on the assumption that there are 21,524 singleton deaths in WL and 10,375 in IBC (the results of filtering the data as available in October 2010 by setting minimum deaths = 1 and maximum deaths = 1). This assumption of independence of these historical records would suggest most (52.7%) singleton killings were missed by both datasets, with IBC missing far more. It is probable that both datasets become more sensitive with larger and larger events, and captured more events in the most violent areas, and thus are not independent.

If the report collection processes were positively correlated, such as both IBC and WL disproportionately recording deaths in Baghdad, in major towns, or associated with bombings, then the overlap between the two would be artificially high, and the true

Report Source	Match to IBC			No Match IBC	Total
	Match = 3	Match = 2	Match = 1	Match = 0	
US source	2	3	1	14	20
Iraqi source	2	2	3	13	20

Carpenter © 2013 Prehospital and Disaster Medicine

Table 4. Number of Matches Out of a Random Sample of 20, US Military vs Iraqi Government Source
Abbreviations: IBC, Iraq Body Count

total number of deaths would be higher than implied by a simple overlap comparison as seen in Table 3. If the two data collection processes were negatively correlated (eg, one tended to record Sunni deaths and the other Shia), then the true total number of deaths would be less than implied by the overlap analysis. Ignoring the specific singleton death count, this table emphasizes that both systems are insensitive to singleton killing events.

A minority of WL records explicitly stated that an event was reported by a US military source or by the Iraqi Government. To explore the potential that one of the datasets was more closely aligned with one of these sources of reports, a systematic sample of 20 records from each of these sources was taken, using a random start. As shown in Table 4, there was no evidence (Fisher exact = .895) that one of the datasets is more strongly associated with either source.

These two listings of fatal events in Iraq, Wikileaks WL and IBC, were likely created for contrasting purposes. In the best of "Western liberal democratic" traditions, IBC was created by volunteer efforts to increase accountability and compassion for events unfolding in Iraq with the available sources of information. While not fully known, it is likely that the War Logs compilation resulted from efforts within the Coalition forces to be both accountable and responsive to the security needs of those within Iraq. In spite of differing processes and motivations, the data depicted in Figure 1 suggest that it is unlikely these two passive surveillance systems were independent or representative samples of all killing events. Other than suggesting that the records created by both Wikileaks and IBC are missing many, and probably the majority of deaths in Iraq, these sources alone cannot tell us the true death toll following the 2003 invasion. Others have already noted that in the first four years of the conflict, IBC missed many and perhaps most newspaper-reported deaths.¹² The analysis provided here should instead act as a cautionary note for scholars attempting to draw population-wide conclusions from convenience samples such as IBC.^{11,13,14} The possibility that large events, or certain kinds of events (eg, car bombs) are overrepresented might allow attribution that one side in a conflict was more recklessly killing civilians, when in fact, that is just an artifact of the data collection process. For example, of the 28,000 bombs dropped in 2003, almost none produced event reports within IBC, suggesting that Coalition bombs may have been systematically underreported.¹⁵ Likewise, none of the US Air Force post-attack assessments of bombing raids are included in the Iraq War Logs. Obermeyer et al have similarly shown that the widely cited Uppsala war and conflict dataset, in

part based on convenience, non-sample data sources, consistently underestimates death tolls from conflict.¹⁶

Conclusion

The two largest datasets of deaths from the Iraq conflict, the Wikileaks Iraq War Logs and the press-based Iraq Body Count, both are insensitive and probably have missed the majority of violent deaths. This comparison of two collections of fatal event reports in Iraq beckons two responses. First, press members and scientists alike should be cautious about assuming the completeness and representativeness of tallies for which no formal evaluation of sensitivity has been conducted. Citing partial tallies as if they were scientific samples confuses the public, and opens the press and scholars to being manipulated in the interests of warring parties. Secondly, academics in general and epidemiologists in particular need to develop widely accepted techniques for assessing sensitivity and bias in passively-collected data. Given the growing popularity of crowdsourced data and indirect evidence such as press reports, providing systematic reviews and critiques of convenience samples will likely become a significant challenge for public health communication for health officials and scholars alike.

Author Roles

Dustin Carpenter took a lead role in organizing the capture-recapture assessment by the students involved in the project, and participated in the writing of the manuscript. Tova Fuller led the analysis of the data and participated in the writing of the manuscript. Les Roberts conceived of and organized the project, and participated in the writing of the manuscript.

Acknowledgements

This work was only able to be conducted thanks to the volunteers and efforts of both the Wikileaks and Iraq Body Count organizations. Their efforts to provide transparency during the complex and obscured setting of wartime Iraq is greatly appreciated. This work was undertaken by a Columbia University graduate class to document human rights abuses. The efforts of the students involved in this analysis were invaluable. Unfortunately, a message sent by a US State Department employee was broadcast by a career placement official at Columbia University on November 30, 2010, threatening that examination of or discussion of the Wikileaks War Logs may jeopardize students' future ability to be employed by the US Government. Thus, it would be unwise and perhaps unethical to acknowledge all the participating students by name.

References

1. While the Official Iraqi estimates have tended to be released sporadically through the press, the Congressional Research Service has compiled reports through 2008 at <http://www.fas.org/sgp/crs/mideast/RS22537.pdf>. Accessed November 21, 2012.
2. Iraq Body Count. www.Iraqbodycount.org. Accessed November 21, 2012.
3. Brookings Institution Iraq Index. <http://www.brookings.edu/saban/iraq-index.aspx>. Accessed November 21, 2012.

4. Iraq Family Health Survey Study Group. Violence-related mortality in Iraq from 2002 to 2006. *N Engl J Med.* 2008;358(5):484-493.
5. Roberts L, Lafta R, Garfield R, Khudhairi J, Burnham G. Mortality before and after the 2003 invasion of Iraq: cluster sample survey. *Lancet.* 2004;364(9448):1857-1864.
6. Burnham G, Lafta R, Doocy D, Roberts L. Mortality after the 2003 invasion of Iraq, across-sectional cluster sample survey. *Lancet.* 2006;368(9545):1421-1428.
7. ORB. More than 1,000,000 Murdered Since 2003 Invasion. <http://www.zcommunications.org/more-than-1-000-000-iraqis-murdered-since-2003-invasion-by-orb>. Accessed November 26, 2012.
8. See for example, AP Wire Service. Iraq war leaks: No U.S. investigation of many abuses. *Washington Times.* October 22, 2010. <http://www.washingtontimes.com/news/2010/oct/22/wikileaks-plans-major-announcement-europe/>. Accessed November 26, 2012.
9. Miller G, Finn P. Secret Iraq war files offer grim new details. <http://www.washingtonpost.com/wp-dyn/content/article/2010/10/22/AR2010102201682.html?sid=ST2010102206224>. Accessed November 26, 2012.
10. *Washington Post.* "WikiLeaks's leaks mostly confirm earlier Iraq reporting." <http://www.washingtonpost.com/wp-dyn/content/article/2010/10/25/AR2010102504643.html>. Accessed November 26, 2012.
11. Hicks MH, Dardagan H, Bagnall PM, Spagat M, Sloboda JA. Casualties in civilians and coalition soldiers from suicide bombings in Iraq, 2003-2010: a descriptive study. *Lancet.* 2011;378(9794):906-914.
12. Siegler A, Roberts L, Balch A, et al. Media coverage of violent deaths in Iraq: an opportunistic capture-recapture assessment. *Prehosp Disaster Med.* 2008;23(4):369-371.
13. Bohorquez JC, Gourley S, Dixon AR, et al. Common ecology quantifies human insurgency. *Nature.* 2009;462:911-914.
14. Hicks MH, Dardagan H, Serdán GG, et al. The weapons that kill civilians - deaths of children and noncombatants in Iraq, 2003-2008. *N Engl J Med.* 2009;360:1585-1588.
15. Pape RA. The true worth of airpower. *Foreign Affairs.* 2004;83(2):116-130.
16. Obermeyer Z, Murray C, Gakidou E. Fifty years of violent war deaths from Vietnam to Bosnia: analysis of data from the world health survey programme. *BMJ.* 2008;336(7659):1482-1486.