

Statisticians Discuss Mortality in Iraq Survey

Mary Gray, Professor of Mathematics and Statistics, American University

At a well-attended February 8, 2007, meeting sponsored by the Washington Statistical Society, the Science and Human Rights Program of the American Association for the Advancement of Science (AAAS), DC chapter of the American Association for Public Opinion Research, and CASPA, Gilbert Burnham, Shannon Doocy, and Scott Zager of The Johns Hopkins University Bloomberg School of Public Health reported on their survey work on mortality in Iraq. Discussants Jana Asher of AAAS and David Marker of Westat critiqued the design of the survey and its ethical and cultural sensitivity aspects.

In September of 2004, a cluster sample survey written about in *The Lancet* undertook to compare mortality in Iraq before and after the March 2003 invasion. The estimate of excess deaths in the post-invasion period was 98,000, with a confidence interval of 8,000–194,000. The subject of the February presentation was a follow-up study conducted in May and July of 2006 and also written about in *The Lancet*, which estimated excess deaths at 654,965 (392,979–942,636). These estimates are as much as 10 times higher than those from other sources and have been the subject of considerable public attention. The popular press generally has focused on the magnitude of the estimates and the size of the samples. That only 33 clusters with 7,868 residents in the first instance and 47 clusters with 12,801 individuals in the follow-up study constituted the samples generated substantial criticism.

While the survey design may be theoretically adequate, given the conditions in Iraq, particularly at the time of the second survey, whether cluster sampling is appropriate is a more difficult issue. However, given time and travel concerns, as well as the security situation, it seems the most reasonable choice. Nonetheless, the incidence of violent deaths, which constituted 601,027 of the total estimate, is highly variable by location so that the sample may fail to be representative. In fact, in the first survey, a Fallujah cluster was tagged as an outlier and the analysis was recalculated without it. The most common cause of violent death was gunfire; although the actual numbers of deaths due to coalition forces went up from



The audience of statisticians listen to discussants Gilbert Burnham, Shannon Doocy, and Scott Zager of The Johns Hopkins University report on their survey work on mortality in Iraq.



David Marker of Westat discusses the design of the mortality in Iraq survey and its ethical and cultural sensitivity aspects.

“Any study that needs to report ‘no interviewers died or were injured during the survey’ might well be characterized as heroic.”

the 2004 to the 2006 survey, they declined as a percentage of all violent deaths.

Unfortunately, if the confidence intervals were noted at all in general press coverage of the survey, it was usually in order to characterize the results as lacking precision, rather than with any attempt to explain what confidence intervals mean. What was not well-explained in most discussions was how the mere 629 deaths reported in the survey formed the basis of the excess mortality estimate.

Because of the public attention that had been focused on the issue of mortality estimates in Iraq, the audience at the meeting obviously hoped to gain a more in-depth understanding of the survey design. To some extent, their expectations

may have been met, but there are clearly issues that are unresolved. Nonetheless, the survey team must be commended for undertaking an important task under extremely difficult conditions.

Any study that needs to report “no interviewers died or were injured during the survey” might well be characterized as heroic. Sadly, the recommendation in the first study to establish an independent monitoring process by an agency such as the International Committee of the Red Cross or World Health Organization has not yet been taken.

A significant factor in the estimates of the second study is the population instability caused by massive internal and external migration, estimated by now to be as high

as 4 million of a total estimated population of 27 million. Not only is this a large number, but, to a certain extent, the migration is localized. Such population instability undercuts the reliability of mortality estimates. In addition, as one who has been responsible for surveys in Iraq as far back as July 2003, I had difficulty with several aspects of the project. In the second survey, GPS was not used due to security concerns. Thus, the cluster selection relied on what could be a biased selection of main streets and tributaries. Moreover, if the team members felt uncomfortable, they were free to choose another neighborhood. Once the cluster was selected, interviewers went from house to house until 40 households were interviewed. That is a lot of in-depth interviews to be conducted in one day, but, more importantly, no follow-up was done on nonresponses due to empty houses or houses where no one answered. While this is understandable from a security point of view, it can introduce bias.

Furthermore, it was reported that only 15 households actually refused to participate in the survey. This seems very low, even though the interviewers were confidence-inspiring local Iraqi physicians—male and female. Unfortunately, at the presentation, there was little discussion of the religious and ethnic divides that might have affected the representativeness of the clusters or the match of the interviewers to the neighborhoods.

The project reports that 80% of the reported deaths were verified by reference to death certificates in order to counter the criticism that the number of reported deaths was too far in excess of the number of death certificates available officially to be credible. However, that official Iraqi records are inadequate in many ways is well-recognized. Nonetheless, concern remains about the accuracy of the data. Apparently, those in charge of the survey did not, themselves, visit Iraq to supervise the operation. From my experience, no

matter how carefully selected and trained those who actually conduct a survey might be, it is problematic when there is no hands-on check from the authors. Although this did not come up at the WSS presentation, J. Bohannon reported in a *Science* article that Burnham does not know exactly how the Iraqi team conducted its survey and that the details about the neighborhoods surveyed were destroyed because of security concerns.

On a "bad day," we may read of 102 pilgrims killed on their way to Kerbala, 40 people dead in a car bomb explosion in a Baghdad Sunni market, 25 bodies in a roadside ditch... However, if the estimates of the JHU group are correct, 560 Iraqi civilians have died each day, on average, since the invasion. Although we may feel that numbers from other sources underestimate the situation, it is still an open question for many statisticians of whether we can feel comfortable with the process that produced these estimates. ■



Scientific Computing Associates Corp.
525 N. Lincoln Avenue
Villa Park, IL 60181 U.S.A.

Website: www.scausa.com
E-mail: sca@scausa.com
Telephone: 630-834-4567 Fax: 630-834-1450

SCA Statistical System (8.0) ^{New}

25% Discount on all new purchases before June 30, 2007

Educational Edition (\$395 – Academic List Price)

Provides all the features you need to import and prepare time series data, identify and build ARIMA, transfer function, intervention, exponential smoothing, and regression models. Also includes new methods for seasonal ARIMA modeling and forecasting. Extensive general statistical analysis rounds out this edition as a multi-purpose educational software.

Practitioner Edition

Builds on the Educational Edition adding time series data mining, automatic time series modeling and advanced automatic outlier detection and adjustment. Also includes new date functions to facilitate daily, weekly, and monthly modeling and forecasting.

Professional Edition (A)

Builds on the Practitioner Edition adding time series analysis and forecasting using vector ARIMA models, simultaneous transfer function models, causality tests, and more.

Professional Edition (B)

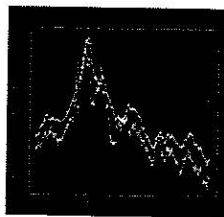
Builds on the Practitioner Edition adding nonlinear time series identification, modeling, forecasting, and analysis features using TAR, threshold ARIMA, threshold TF, and segmented time series models. Also includes new analysis methods for time-varying parameters, GARCH modeling, and model-based seasonal decomposition.

Advanced Edition

Encompasses all features of the SCA Statistical System combining Professional Edition A and B into a unified and power-packed software to handle most any research topic, or business and industrial application involving time series analysis and forecasting.

SCAB34S SPLINES

- Multivariate Adaptive Regression Splines
- Time series Adaptive Spline Threshold Autoregression (ASTAR)
- Alternating Conditional Expectations (ACE)
- General Additive Models (GAM) and PI-Splines
- Data mining using the above methods
- General nonlinear function solving, and much more



Second Edition Time Series Analysis and Forecasting

Lon-Mu Liu, PhD.
Univ. of Illinois at Chicago

This important book distills and integrates important research results into cohesive and comprehensible methodologies.
Regular: \$54.95 Limited Offer: \$44.95 - See SCA website