DECLASSIFIED

RESTRICTED

From: Creon Butler

Date: 8 November 2004

Copies to: Dominic Asquith
John Williams

Special Advisers

PS

COUNTING IRAQI CASUALTIES

Following the Secretary of State's meeting this morning, I have looked through the Lancet article. My initial thoughts are as follows:

I agree that the statistical methodology appears sound. The authors accept that there is considerable uncertainty over their central estimate of excess deaths, and provide a careful assessment of various possible statistical biases. Their conclusion is that, despite these, the results are sufficiently strong to raise concern and at the very least justify further study.

The method involves taking 33 random samples from the Iraqi population, each comprising 30 households living in the same neighbourhood. Interviews were used to establish how many deaths occurred in each cluster of households in the 17.8 month period after the invasion compared with the 14.6 month period preceding it. Provided the samples are genuinely random, statistically valid inferences can be drawn for the Iraqi population as a whole. Survey techniques of this kind are widely used (e.g. in measuring unemployment in the UK). There are five main types of question worth asking in judging the validity of this particular study:

- Was the sampling genuinely random? The authors have tried hard to achieve this despite sometimes severe constraints on where their survey staff could go. One reason for dropping the Falluja sampling point from the calculations behind the headline estimate was that the authors could not be sure the cluster of households in the Falluja district was selected in a genuinely random way. Equally, while they sought to reduce the travel and risk faced by interviewers by dropping certain Governorates from the sample, they did so in such a way as to ensure the overall sample was not biased.

- Was the information provided by interviewees accurate? The paper discusses the possibility of "recall" bias – i.e. that deaths before the invasion were not remembered as accurately as those after. However, they argue – pretty convincingly – that deaths are unlikely to be forgotten. Another possibility is that families might exaggerate the number of deaths since the invasion because of hostility to the coalition/interim government. Death certificates were only sought in 78 out of 988 households. In these cases there was a high rate of confirmation (63 out of 78, with plausible explanations in all cases where

DECLASSIFIED

DECLASSIFIED

certificates could not be produced). However, it is possible that this gives a biased picture of accuracy if interviewers tended to ask for certificates mainly when the information they had received was most plausible and hence the risk of causing offence minimised.

- How accurate were the inferences made about the broader population? One possibility is that the size of households was under-estimated (because respondents wished to protect members who were insurgents) in which case the scaling up would have over-estimated the total number of deaths. However, the authors argue that the bias may just as likely go the other way as families might seek to justify higher ration distributions by overestimating household size. Another possibility is that the sampling strategy might have missed various categories of people – homeless, soldiers etc. This seems likely to have been more of an issue. But the bias introduced – particularly vis-à-vis soldiers - could be to underestimate the figure for excess deaths.

- Even if the estimates of excess deaths are sound, can the same be said of the breakdown between different causes of death, and in particular the extent to which additional deaths were caused by air attacks from coalition forces? The scaled up estimates of deaths from particular causes will be more uncertain than the estimates of excess deaths in total, simply because the samples are smaller still. It is also possible that interviewees might exaggerate the proportion of deaths caused by coalition action vis-à-vis other causes not directly linked to the activities of coalition forces.

- If the methodology is sound, how can one reconcile the 98,000 death estimate with other data. In particular (a) the much lower casualty estimates based on press reports; and (b) the lack of anecdotal evidence of much larger numbers of injured attending Iraqi hospitals. The divergence with estimates based on press reports might be explained through the passive nature of press reporting and the partial territorial coverage of journalists. However, it would seem much harder to explain the mismatch between the estimated number of deaths and the anecdotal evidence on injuries. Assuming a ratio of four injured for every death (as reflected in the Iraqi MOH figures from hospital admissions) would suggest 400,00 injuries (although given the variability in death/injury ratios with different conflict circumstances some caution needs to be exercised here too).

Conclusions

Overall, it is perhaps not surprising that the methodology appears sound, since the Lancet's pre-publication reviewing process should have revealed significant methodological weaknesses.

The authors describe a series of possible biases in both directions, but none of these (or others noted above) seems at first sight so striking as to invalidate the overall findings. The hardest discrepancy to explain is the lack of anecdotal evidence of injured people in proportion to a 98,000 central estimate for deaths.

In commenting on the study we should certainly continue to emphasise the considerable uncertainty around the central estimate (reflecting the small sample

DECLASSIFIED

DECLASSIFIED

size),[1] as well as the lack of corroborating evidence - particularly evidence of injured in the numbers one might expect. We could also highlight some of the factors which might bias the study towards an over-estimate of deaths. However, there are as many reasons why the study might be biased in the other direction (so probably safer not to go down this road).

There are various ways to try and check the validity of the estimates using data from other sources (hospital reports, casualty figures reported by soldiers and police, reports of funerals etc) and trying to refine it to remove biases. It might also be possible, as Gerard Russell has suggested, to try and validate the study's pre-invasion estimate of mortality by checking it against unpublished MOH health figures. But there is (a) no certainty at this stage that this kind of work would invalidate the Lancet findings, or (b) any guarantee that if it does produce a different answer, that the rejection of the Lancet findings would be conclusive. In the absence of a detailed census (impossible in the current security environment), the best way of narrowing down the uncertainty in the Lancet article is likely to be to conduct a similar survey with a significantly larger sample.

Creon Butler
Chief Economist

[1] While at the same time avoiding belittling the efforts of the interviewers who faced considerable obstacles and risk.

DECLASSIFIED